Acuitive, Inc.

Virtual Resource Management*

(*AKA Server Load Balancing and Non-Server Load Balancing)

Which Vendor is Right for You?

Version 3.0

TABLE OF CONTENTS

1 INTRODUCTION.....4
1.1 About This Document 4
1.2 Acuitive Objectivity 4
1.3 Document Structure 5
1.4 What Is Virtual Resource Management? 5
1.5 Who Is Acuitive? 6
1.4 About Our Documents and Subscription Service 6
1.6 About The Authors 7
1.6.1 Mark Hoover 7
1.6.2 Dave Logan 7
1.6.3 Jared Berglund 8
1.6.4 The Rest Of Acuitive 8
1.6.5 The Vendors Themselves 8
2 VENDORS, PRODUCT TYPES AND INDUSTRY DYNAMICS.....9
2.1 Who Sells VRM? 9
2.2 VRM Product Categories 10
2.2.1 Software Products 10
2.2.2 “Software-On-A-Stick” Products 11
2.2.3 Switching VRMs 11
2.3 1998 In Retrospective – Visible Trends 13
2.4 Predictions 14
3 VENDOR SUMMARIES.....15
3.1 Evaluation Methodology 15
3.2 Summary Format-Explanation of Each Section 15
Company X (www.companyx.com) 15
3.2.1 Company Overview 15
3.2.2 Product & Support Overview 15
3.2.3 Product Functionality 16
3.2.4 Product Features at-a Glance 16
3.2.5 Key Relationships 18
3.2.6 Issues 18
3.2.7 Final Analysis 18
3.3 Alteon WebSystems (www.alteon.com) 19
3.3.1 Company Overview 19
3.3.2 Product Overview 19
3.3.3 Product Functionality 20
3.3.4 Product Features at-a Glance 24
3.3.5 Key Relationships 25
3.3.6 Issues 25
3.3.7 Final Analysis 26
3.4 ArrowPoint (www.arrowpoint.com) 27
3.4.1 Company Overview 27
3.4.2 Product & Support Overview 27
3.4.3 Product Functionality 28
3.4.4 Product Features at-a Glance 32
3.4.5 Key Relationships 33
3.4.6 Issues 33
3.4.7 Final Analysis 34
3.5 Cisco (www.cisco.com) 35
3.5.1 Company Overview 35
3.5.2 Product & Support Overview 35
3.5.3 Product Functionality 36
3.5.4 Product Features at-a Glance 38
3.5.5 Key Relationships 39
3.5.6 Issues 39
3.5.7 Final Analysis 40

3.6 Coyote Point (www.coyotepoint.com) 41
3.6.1 Company Overview 41
3.6.2 Product & Support Overview 41
3.6.3 Product Functionality 42
3.6.4 Product Features at-a Glance 43
3.6.5 Key Relationships 44
3.6.6 Issues 44
3.6.7 Final Analysis 44
3.7 F5 Networks (www.f5.com) 45
3.7.1 Company Overview 45
3.7.2 Product & Support Overview 45
3.7.3 Product Functionality 46
3.7.4 Product Features at-a Glance 48
3.7.5 Key Relationships 49
3.7.6 Issues 49
3.7.7 Final Analysis 50
3.8 Foundry Networks (www.foundrynet.com) 52
3.8.1 Company Overview 52
3.8.2 Product & Support Overview 52
3.8.3 Product Functionality 53
3.8.4 Product Features at-a Glance 55
3.8.5 Key Relationships 56
3.8.6 Issues 56
3.8.7 Final Analysis 56
3.9 HolonTech (www.holontech.com) 57
3.9.1 Company Overview 57
3.9.2 Product & Support Overview 57
3.9.3 Product Functionality 57
3.9.4 Product Features at-a Glance 59
3.9.5 Key Relationships 60
3.9.6 Issues 60
3.9.7 Final Analysis 60
3.10 HydraWeb (www.hydraweb.com) 61
3.10.1 Company Overview 61
3.10.2 Product & Support Overview 61
3.10.3 Product Functionality 62
3.10.4 Product Features at-a Glance 63
3.10.5 Key Relationships 63
3.10.6 Issues 64
3.10.7 Final Analysis 64
3.11 IBM (www.software.ibm.com) 65
3.11.1 Company Overview 65
3.11.2 Product & Support Overview 65
3.11.3 Product Functionality 66
3.11.4 Product Features at-a Glance 68
3.11.5 Key Relationships 68
3.11.6 Issues 69
3.11.7 Final Analysis 69
3.12 IPivot (www.ipivot.com) 70
3.12.1 Company Overview 70
3.12.2 Product & Support Overview 70
3.12.3 Product Functionality 71
3.12.4 Product Features at-a Glance 74
3.12.5 Key Relationships 75
3.12.6 Issues 75
3.12.7 Final Analysis 76
3.13 Radware Ltd. (www.radware.com) 77
3.13.1 Company Overview 77
3.13.2 Product & Support Overview 77
3.13.3 Product Functionality 79
3.13.4 Product Features at-a Glance 83
3.13.5 Key Relationships 84
3.13.6 Issues 84
3.13.7 Final Analysis 85
3.14 Resonate (www.resonate.com) 87
3.14.1 Company Overview 87

3.14.2 Product & Support Overview 87
3.14.3 Product Functionality 88
3.14.4 Product Features at-a Glance 91
3.14.5 Key Relationships 92
3.14.6 Issues 92
3.14.7 Final Analysis 93

3.15 Others 94
3.15.1 Introduction and Clarification 94
3.15.2 Allot (www.allot.com) 94
3.15.3 Bright Tiger 95
3.15.4 CheckPoint (www.checkpoint.com) 96
3.15.5 CinTel (www.cintel.co.kr) 96
3.15.6 Eddie (www.eddieware.org) 97
3.15.7 Microsoft (www.microsoft.com) 97
3.15.8 netSTOR (www.netstor.com) 98
3.15.9 WebManage (www.webmanage.com) 98
3.15.10 WebSpective (www.webspective.com) 99
3.15.11 Xedia (www.xedia.com) 100

4 HOW TO CHOOSE A VRM SOLUTION 101

5 VENDOR SELECTION GUIDELINES 104

5.1 Characterizing Your Application 105
5.1.1 Local VRM, Static Content 106
5.1.2 Local VRM, Downloaded Content 108
5.1.3 Local VRM, 1-Way Dynamic Application Environment 110
5.1.4 Local VRM, 2-Way Dynamic Application Environment 113
5.1.5 Local VRM, Server Optimization 116
5.1.6 Multi-Site VRM, Site Redundancy 117
5.1.7 MS-VRM, Global Content Distribution 119
5.1.8 Co-Location/Hosting 120

6 VRM FEATURES: CHART AND DEFINITIONS 123

1 Introduction

1.1 About This Document

This document is written for site and network architects everywhere who may benefit from the use of Virtual Resource Management (VRM) in their network and application/web server deployments. The document can help technical readers who need to better understand the considerations for using Virtual Resource Management, the various architectural and feature-level options associated with state-of-the-art products available today, and how to choose an appropriate vendor and product set to meet your requirements.

Anyone who has been associated with the Networking industry for more than 15 minutes has lived thorough the adage: everything changes, except change itself. With this is mind some may ask; why would you spend so much time and effort to put together a document about the VRM *Vendors*? Won't this stuff be out-of-date before I've finished reading it?

Here's the perfect consulting answer: yes and no.

Yes, prices, platforms, memory, processor speeds and various other tactical options change quite frequently. The vendors are constantly tweaking their products to keep pace with the industry trends and your needs.

No, the vendors overall architecture doesn't change that often. They have invested a huge amount of resources to bring their vision of the perfect VRM product to the table. So while we have included pricing and processor ratings and various other easily changed parameters, we have primarily focused on each vendor's product architecture and how their solution fits your application environments.

We'll provide updates on the fast-changing stuff – features, capacities, etc., via quarterly e-mailed updates.

If you haven't already guessed, this document is primarily about the vendors and their products. We make some judgments about which solutions fit which situations best. If you want to cut through all the marketing haze, learn of each vendors' strengths and weaknesses, compare the most critical feature sets, and benefit from our experience and pain, then keep reading!

This document is not about detailed technological descriptions. We have created a companion report that is purely focused on the technology of VRM. This report, entitled "**Key Technologies, Tricks of the Trade, and Application Requirements**" can be used to guide you through the technical jungle and help identify for your specific application requirements.

1.2 Acuitive Objectivity

The information in this document is based on our own research, things we have learned from helping some vendors define and deliver products, and numerous real world applications that we have designed and implemented.

Our focus is on helping end users benefit from VRM. We have accepted money from several of the vendors evaluated in this document. The list includes Alteon Websystems, F5 Networks, Resonate, and Radware. The tasks we have performed for these vendors includes everything from one-day strategy

reviews to writing white papers to defining product requirements and to extending their Professional Services capabilities. Does this give us a bias towards those vendors? Yes, as long as they listen to and execute on our suggestions, because we are trying to get them to build the best products possible for our end user customers. But, we also have good and less formal relationships with most of the other vendors surveyed in this document. Ultimately, our views are shaped by the experience our end-user customers have in deploying the various vendors' products that we recommend and design in. Those real world experiences force us to be ruthlessly objective, because the end user experience reflects on us and we like to look good.

1.3 Document Structure

Section 2 provides an overview of the market segment, who plays in this market and what basic categories their products fit into. This section also reviews the trends currently visible, discusses why they are important to you and concludes with some of our famous predictions. Read this to brace yourself for the coming onslaught of details.

Section 3 provides summaries of all the key vendors and their product lines. Section 3.1 explains our methodology for researching and presenting the vendor summaries. We also provide a primer to help you understand what the various Vendor-Sections mean. This will aid in the rapid assimilation of the Vendor information, and help you focus on the next step – deciding which Vendor is right for you

Section 4 highlights the critical areas to consider when making a Vendor choice.

Sections 5 drills down into the application of VRM technology, and how it plays into the most common application environments. It guides the reader into defining their own environment and then suggests which Vendors will best meet these requirements.

Section 6 combines all the Technology and Vendor options into one easily-accessed Feature Chart. Use this to learn quickly, grill your short list of vendors and review your Vendor decisions.

1.4 What Is Virtual Resource Management?

Acuitive uses the term Virtual Resource Management[1] (VRM) for the methods and procedures associated with making multiple networked devices "appear" to users and related components as one larger device. Want to make six NT servers and two Solaris servers appear as one large web server? VRM. Want to make three routers look like one large one? The answer's VRM. Want to make multiple firewalls, VPN devices, or proxy cache servers look like one big device? You should consider VRM.

---

1 You'll find some other terms used in the industry. The most common might be the term Server Load Balancing (VRM). But the concept has broadened to include infrastructure devices like routers, cache servers, and firewalls. So the *Server* aspect of VRM term doesn't seem encompassing enough. Also, many of the techniques today are designed not to primarily just balance load, but to use resources in the most effective manner, consistent with the nature of the application, it's security mechanisms and user-state mechanisms. As you apply these techniques to a transaction-oriented site, for instance, the goal is to ensure transaction integrity. If you can also balance the load while meeting that goal, so much the better, but it's not the primary goal.

On the other hand, Collaborative Research, who is the industry leader in market research associated with this category, uses the term Internet Traffic Management (ITM). We like that term for the collection of capabilities that they encompass with it. But it's a bit broader than what we are addressing here, because it includes the management of traffic over inherently non-virtual entities such as WAN access links. So we have a slightly narrower focus and use the term Virtual Resource Management.

1.5 Who Is Acuitive?

Started in 1997, Acuitive is a strategic consulting firm focused on the development and application of emerging networked computing technologies. Acuitive provides a wide variety of technical and marketing advisory services to equipment vendors, service providers, and enterprise network planners. Acuitive also publishes qualitative research reports on emerging technologies, to educate end-users and stimulate market development.

The Acuitive's Network Planning and Management Group assists enterprise network clients with strategic planning for IP services. The NPMG services address two key needs in enterprise network planning: Network Requirements and Application Requirements. The services include baselines for networks and applications, certification for new application rollouts, capacity planning, design of complete network management systems, advice on Virtual Resource Management and WAN bandwidth management strategies.

The Acuitive Business Strategy Group helps vendor clients with market strategy, business cases, target customer requirements, product line plans, product and service requirements, alliance partnerships, competitive evaluations, product and program management, company and product positioning, and outbound marketing.

We've been involved in the development and application of VRM technology from day one. As technologists, VRM intrigues us because it operates at the convergence point between applications and networks. As advisors to end users, it's critical because it is an important technical underpinning for making new businesses and business processes based on web technology more reliable and scaleable. As strategists consulting to vendors, it's a technology that has spawned new product categories and companies, and will be a key component of the IP services functions larger vendors need to offer to rise out of the increasingly commoditized network "plumbing" market.

As a result, we've been involved with or have observed hundreds of deployments of the technology, we have hands-on experience with almost all the products available on the market, and we have consulted with several of the vendors to help them orient their products to emerging customer needs.

1.4 About Our Documents and Subscription Service

We've "bottled" our experiences in this area to help you get educated on the technology issues of VRM, characterize your application, and choose the right vendor solution for your application. To that end, we have published two comprehensive research reports.

The first, entitled *Virtual Resource Management: Key Technologies, Tricks of the Trade, and Application Requirements* provides a detailed tutorial of the various approaches to building VRM solutions. The values of various techniques and features available in the market, as related to specific application types, are discussed. The result is a "hit list" of key attributes to look for in a VRM solution for your particular application. The attributes are organized by **policies**, **mechanisms**, **feedback**, **performance**, **redundancy**, and **management**. These areas of consideration are the key aspects of a VRM system to evaluate when architecting a solution. Some areas of technology, such as preferential services and security, are still changing rapidly. So those who acquire this research report will automatically be enlisted into a subscription program providing bi-monthly technical position papers on a technology area or application note of interest through April 2000. Some subjects under consideration are "Using and Abusing Preferential Services," "Practical Capacity Planning," "Famous Sites We Have Known" (case studies), and "Designing In Iron-Clad Security." We invite you to suggest topics. The technical position

papers will be sent to you via e-mail. We will also send out regular corrections and additions to this document on an as-needed basis, via e-mail at vrm@acuitive.com.

Next you need to choose a specific vendor to implement your solution. To help in these considerations, we offer this research report entitled "Virtual Resource Management: Which Vendor Is Right For You?" This document summarizes the capabilities available from each of the key vendors in the market today and maps their capabilities against application requirements to create a "short list" of vendors and products to consider for various types of applications. This information is always changing, so those who purchase this document will automatically be enlisted into a subscription program providing quarterly updates (via e-mail) through April 2000.

For more information on the research reports, go to www.acuitive.com. There you will find additional detailed information, including a complete Table of Contents, for each report. These reports are orderable from the web site.

Please direct any comments, questions, opinions regarding VRM to vrm@acuitive.com.

1.6 About The Authors

1.6.1 Mark Hoover

Mark Hoover is the President and co-founder of Acuitive. Mark worked at AT&T Bell Laboratories for about ten years, and was involved in the development of satellite transmission and fiber optic devices and systems, high speed packet switches, and LAN products based on emerging 10-BASET and FDDI standards. Mark also ran the team that provided technical support for the AT&T OEM agreement with Cisco. In 1990, Mark left AT&T to join SynOptics. At SynOptics, Mark was initially responsible for the definition and development of the (at the time) new generation hub platform – the System 5000 (although Jim Vogt did most of the hard work). Mark went on to form the internetworking product line at SynOptics, which ultimately resulted in the merger with Wellfleet to create Bay Networks. At Bay Networks, Mark formed the Internet/Telco Business Unit to define, develop, and market products for the service provider community.

At the end of 1995, Mark retired from Bay Networks to train for the Seniors golf tour, figuring ten years was plenty of time to prepare. After finding out that golf is a lot harder than it looks, Mark formed Acuitive at the beginning of 1997, along with Dave Danielson.

Mark now spends his time running Acuitive (which doesn't take much effort), studying technology and market trends, providing strategic consulting advice to vendors in the general area sometimes called "IP Services," and acting as a "trusted advisor" to several companies at the CIO level. Mark stays in contact with almost all of the key vendors in the VRM space and provides consulting advice, both formally and informally, and learns a lot in return.

1.6.2 Dave Logan

Dave joined Acuitive in June 1997 as a Senior Consultant. He has more than 12 years of experience in the networking industry, focusing on end-to-end network designs and the design and creation of networking technologies. Dave currently specializes in consulting with networking equipment vendors on IP Services product strategies, especially in the areas of server load balancing, bandwidth management, policy-based networking, and device/network management.

From 1993 to 1997 Dave held various positions at Bay Networks/SynOptics Communications. Most recently, Dave was a Senior Product Manager within the Network Management Group. In this role, he first managed Bay's next-generation embedded management technology and RMON2 strategy, and within a few months was running Bay's Optivity LAN team. Dave was instrumental at focusing the team on its next-generation goals and processes and designing the next major release of Optivity.

1.6.3 Jared Berglund

Jared joined Acuitive in March 1998 as a Senior Consultant. He has more than 10 years of experience in the networking industry, focusing on end-to-end network designs and solutions, large enterprise network management strategies, and the application of advanced networking technologies. Jared currently divides his time between consulting with networking equipment vendors on product strategies and designing large scale, mission-critical networks.

From 1989 to 1998 Jared held various positions at Lattice Semiconductor, Anixter, Inc., 3COM and SAIC. Most recently, he acted as the chief architect to define client requirements, direct executive staff decisions, translate business needs into technology applications and architect end-to-end enterprise solutions.

1.6.4 The Rest Of Acuitive

While Mark, Dave, and Jared were the main authors of this document, in a sense it was written by many people in Acuitive whose cumulative first-hand knowledge of the various vendor’s products in actual field application environments have been captured in this document.

1.6.5 The Vendors Themselves

All of the major VRM vendors were given a chance to review a draft of their sections, and to provide corrections, updates, rebuttals, etc. Most of the vendors participated enthusiastically in this process and their section almost always changed significantly due to this interactive process. We thank them for their time and effort.

2 Vendors, Product Types and Industry Dynamics

We'll say this – VRM is a fun industry. Everywhere you turn - whether it's customers, vendors, consultants, analysts, you run into smart, dedicated, ambitious, fun people. The situation reminds us of the early days of the routing industry, where some of the top minds of the industry were innovating at a fast pace. No one was real sure what the final form of the proper solution was, customers were trying out stuff and feeding back requirements at a furious pace, and companies were fixing things and delivering new capabilities as quickly as the laws of physics would allow.

Now, we're not saying the VRM industry will ever be as large as the routing industry has become or will spawn the next Cisco. The technology is too powerful. You simply don't need as many VRM products in a global network as you need routers. Plus, the market is more fragmented, crowded with viable vendors. But it is still a healthy, expanding market. We could see it becoming a $1B industry. Maybe half that. How's that for precise market forecasting? We're not a market forecasting company. The company that has performed the most extensive market sizing in this area is Collaborative Research (www.collaborativeresearch.com). They estimate the global market was about $132M in 1998 and will grow to about $830M by the year 2002. That sounds about right. More important, they estimate the market was about 5,000 units in 1998, growing to about 80,000 units in the year 2002, where a unit is defined roughly as a redundant scheduler. That, too, sounds about right. Any way you slice it, the use of VRM technology was much greater in 1998 than it was in 1997, and appears to be growing at a fast and furious rate so far in 1999.

2.1 Who Sells VRM?

VRM solutions are offered by a wide variety of companies and in many forms. There are huge companies in the game, and garage shops.

Let's start with the gorillas. Cisco, Microsoft and IBM all participate in this space. They each come at the problem from a slightly different perspective. Microsoft acquired Valence and is integrating that capability into versions of NT to provide services for web server, application server, and storage access redundancy. But it's only for NT, of course. IBM's VRM product is usually sold as a component of a larger package of software that is oriented toward complete E-Business site solutions. Cisco is the only one of the three gorillas that has a stand-alone product that competes head-to-head against other VRM products.

But the gorillas also have some common characteristics. For each of them, the success of their company obviously does not depend on the success of this particular product category. Therefore, although hugely bigger than the other companies delivering product in this space, they invest less in R&D and internal training than the other more focused vendors. As a result, these companies do not have the "best" products, with all the latest bells and whistles. And their support team may not be as savvy about the unique issues associated with the application of this technology. But if the product is good enough for your needs, and the support is adequate, the gorillas represent stability you can't get anywhere else and they can also provide product and support far beyond the VRM component.

Contrast the gorillas with a set of smaller, more nimble and focused start-up companies. These vendors include (in alphabetical order) Alteon Websystems, ArrowPoint, CoyotePoint, F5 Networks, HydraWeb, IPivot, Radware, and Resonate. As you'd expect, the start-ups tend to innovate a lot faster and tend to be able to provide more depth to their support because virtually everyone in the company eats and breathes the technology 24 hours a day. The real "in-the-know" customers, whose sites and jobs depend on the

capabilities of their VRM solution, choose from this vendor set. But the gorillas cast a long shadow over the market, eating up a lot of the revenue associated with applications that don't need the latest innovations. Luckily for the smaller vendors, the pace of innovation needed in many areas is still high -- the smaller companies have sold many products after customers tried the offering from a gorilla and either it didn't work or they outgrew it.

2.2 VRM Product Categories

Different vendors have taken different architectural and packaging approaches to "productizing" VRM functions. The main categories of products are:

- Software products
- "Software-on-a-stick" products
- Switching VRMs

2.2.1 Software Products

Software products are delivered on CD-ROMs or other media. You install them on computer platforms and operating systems that you choose, and the VRM system utilizes the memory, CPU, and I/O resources of that platform.

The biggest advantages of software-based products are:

- They don't require the introduction of another type of box into the mess we call web sites
- Capacity can scale by adding memory, etc. to the base computer platform
- Product performance scales immediately with the release of new server technology (Moore's Law is working for you). The latest technology can be used to increase scheduler performance (if needed), and last month's latest and greatest can be re-used as a content server to increase overall site scale
- Site functionality and topology can generally be rapidly changed, often remotely with no wiring or other physical changes

The biggest disadvantages of software-based products are:

- You need to make sure the software is supported on the type of hardware and OS version you have
- Extra work is required to install the product and customization may be necessary, which can result in extra deployment cost. Part of the problem is making sure you haven't impacted another application or resource running on the same platform
- Engineering guidelines are difficult to define. Throughput and capacity are dependent on the hardware platform, operating system, systems tuning, co-resident functions, etc. You don't know what you bought until you get it and install it. And even then it might change tomorrow
- Reliability is determined by the reliability of the computer platform and all co-resident software

Different approaches to software-based products exist. Microsoft (via Valence) offers a completely distributed scheduler system where a scheduler is deployed on each content server and all traffic is sent to all servers. IBM offers a more traditional solution, where the software is centralized and simply loaded onto a hardware platform to become the site scheduler. All traffic passes through the single scheduler to get to the content servers. Resonate offers a solution where the scheduler function is partitioned between a central scheduler and the content servers themselves. This off-loads some functions from the scheduler, making it faster. The Resonate solution also provides the option of turning on as many schedulers on as many servers as needed to meet the site performance requirements.

2.2.2 "Software-On-A-Stick" Products

Some of the negative aspects of software products: platform support, unclear engineering guidelines, and installation headaches, have been mitigated by vendors who essentially have software products, but deliver them to the marketplace pre-packaged onto (usually) PC platforms. These "software-on-a-stick" (SOAS) products allow the manufacturer to select peripherals and drivers, memory speeds and sizes, and operating system parameters for the customer and ship a working system with known performance and capacity characteristics.

The biggest advantages of the "Software-On-A-Stick" products are:

- Well-characterized performance and capacity parameters
- Easy to buy and install. No issues of OS, platform, or driver compatibility
- No impact on the performance of existing applications on servers (no co-residence)
- Broad feature set with a good hope of future enhancements

The biggest disadvantages of "Software-On-A-Stick" products are:

- Mid range performance. A Pentium CPU running BSD UNIX can only run so fast, even if the vendor pushes most of the functionality down into the kernel/driver space
- Performance degradation with feature deployment. All functions on a SOAS share a common CPU. Packet processing must be performed by that CPU, as well as all management functions, server health and load feedback monitoring, bandwidth management, packet filtering, or any other feature the vendor may provide that you want to turn on. Performance degrades with each feature enabled
- Expensive – by the time the vendor buys and stocks PCs and then turns around and resells for the required margin, you can end up buying a $20,000+ "PC"

The SOAS approach has been the simplest and fastest way to get product and new features to market. A vendor requires little hardware engineering effort and can focus on software. Free or readily available software can be leveraged when a standard Operating System such as BSD UNIX is used. Unlike the software vendors, the SOAS vendors don't have to port and test their products on a multiplicity of different operating systems. Most of the initial vendors in this space (Cisco, F5 Networks, Radware) have chosen the SOAS path. These products have helped define the market and their use has helped uncover issues and add sophistication to the feature sets. Newer entrants such as IPivot and CoyotePoint have also chosen to leverage this model.

As a result of these vendor choices and early market customer choices, roughly 70-80% of the VRM systems in deployment right now are SOAS-based.

2.2.3 Switching VRMs

Switching VRMs are devices that integrate VRM functionality with fully featured Layer 2 and/or Layer 3 (routing) features.

The advantage of a Switching VRM is that it curtails the proliferation of multiple boxes at a site. Depending on the site and product chosen, a Switching VRM can eliminate the need for separate router, switch, bandwidth management, and VRM boxes. If all these functions were implemented with redundancy, that saves a lot of boxes, wiring, and provisioning. Again, depending on the product chosen, the cost for such consolidation can just be an incremental cost over what you'd pay for a state-of-the-art Layer2/3 switch anyway.

The disadvantages of a switching VRM are:

- It takes the vendor a little longer to integrate and test new features. The SOAS vendors have always tended to be six months or so ahead of the Switching VRM vendors in terms of features, and we suspect that will always be generally true
- Generally, less capacity than SOAS and software solutions. Capacity, for services, numbers of simultaneous sessions, state management, traffic accounting, bandwidth management, etc., takes memory. You can't just load up a switching VRM with commodity memory to increase capacity. And what memory there is gets used for various Layer 2 and Layer 3 functions as well

One general myth about Switching VRMs is that they are inherently the highest performing VRM product category. That may or may not be true. ASICs built for VRMs usually provide some kind of hardware assist for low-level packet processing like MAC and IP address substitution, CRC and FCS re-calculation, basic packet classification (e.g., identification of flows), and table look-ups, because these are common functions needed by routers. Only a little VRM capability beyond that (sequence number substitution for TCP Diddling, SYN and FIN packet detection and a few other things) can be hard-coded in an ASIC. Beyond that, it comes down to software processing. And it is in that aspect that Switching VRM vendors vary widely.

Foundry uses their internal management CPU for all VRM functions. Thus, while they offer a blazing fast Layer 2 switch, the VRM performance is only about the same as a SOAS product using the latest and greatest Pentium technology.

Alteon, on the other hand, has two RISC processor cores integrated into each of their ASICs, one of which is used per physical port. Thus in a 10-port box, you have 20 processors working for you, performing all the heavy-duty VRM tasks. As a result, Alteon can achieve levels of performance (as measured by connections/second served) which are an order of magnitude more than most other products can achieve.

ArrowPoint boasts performance numbers which approach Alteon's. Arrowpoint uses an MMC chipset internally, which can be micro-coded to provide software-based processing at hardware kinds of speeds.

Holontech is somewhere between Foundry and Alteon/ArrowPoint. Holontech uses a Galileo chip set for basic Layer 2 switching, but augments it with FPGAs to accelerate some VRM functions.

There is a price for everything, however. The number of VRM options enabled and the amount of memory available will serve to limit the performance and capacity of a distributed architecture. Sharing state information between distributed CPUs with their own memory can also be a burden. If the memory architecture is distributed, then available memory in the system may not necessarily be pooled together, and thus you may be limited to what is available on the portions of the system that are active.

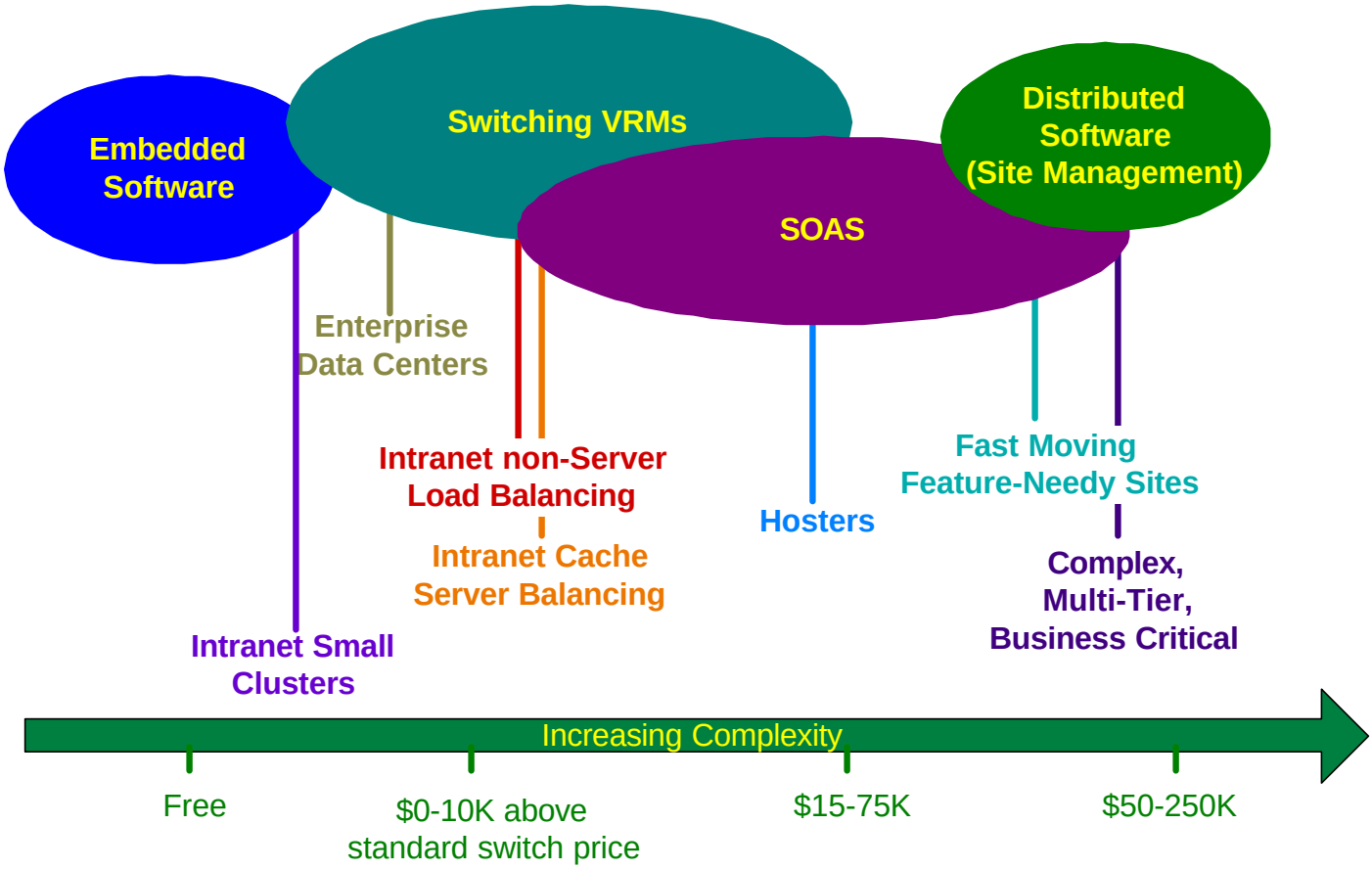
2.3 1998 In Retrospective – Visible Trends

Some interesting trends became apparent in 1998 that we believe are harbingers of things to come.

- **The emergence of non-server load balancing**. Most vendors originally designed their solutions to enable multiple web servers to act as one virtual, large server. But in 1998, many vendors tuned their solutions to support load balancing of other functions, such as routers, VPN devices, firewalls, proxy caches, switches, IP mainframe front-ends, and other devices. Using these techniques, almost any element of the network infrastructure can be made more scaleable and reliable via load balancing. Several vendors have reported that over 50% of their revenue in 1998 were for such applications. Up from roughly 0% in 1997
- **The "featurization" of load balancing**. In 1998 vendors proved that server load balancing (and non-server load balancing) could be effectively supported on high speed Layer 2/3 switches, on multi-function network appliances, in operating systems, and as part of application suites. These are examples of the shift of some load balancing applications from dedicated products to **featurized** solutions
- **The movement towards Site Management solutions**. Meta Group issued a report early last year about a set of functions that they collectively called Site Management. This included load balancing, but also included bandwidth management, firewalling and security management, performance monitoring and management, and content management. The trend among vendors is to add some of these capabilities to server load balancing to head more towards being a complete site management solution
- **The increasing influence of web hosters and outsourcers**. In 1998, web hosters, to make their hosting services more reliable and scaleable, deployed many VRM products. Of even more significance, however, was the inclusion of load balancing as a service provided by outsourcers (like Exodus and Frontier/Global Center, and Digital Island). The advantage to the end customer is that they can turn a fairly high capital expense into a more drawn out leasing/service expense
- **A glimpse of the intranet market**. All market growth projections are based on an assumption of incorporation of VRM solutions into the intranet over the next one to four years. The good news is, in 1998, we saw evidence of the emergence of intranet use. Some vendors report that 20-40% of their 1998 installs were for intranet uses. This is good for the market. It portends vigorous growth

2.4 Predictions

- **Increasing rate of featurization**. We expect a larger number of switch and router vendors, operating system vendors, distributed application vendors, service level management vendors, content management vendors, and others to support some form of server load balancing in their solutions as an integral (and optional) feature
- **Decreasing Average Selling Prices (ASPs)**. Largely due to the accelerating featurization, we expect ASPs for many load-balancing applications to drop – as low as $0 in some cases if the capability supported in the operating system or application suite meets your needs
- **Accelerated deployment of non-server load balancing.** Partially due to the lowered ASP associated with featurization, we expect the technique of load balancing infrastructure devices to take off, becoming a standard design technique in service provider and enterprise networks of all types
- **Market consolidation and maturation**. The market is not big enough to support the number of vendors presently contending, especially if half or more of the revenue goes to the gorillas. Some vendors will drop out. Expect Northern Telecom, Lucent, 3COM and maybe a few server companies to acquire some of the start-ups in 1999 or 2000
- **Application segmentation.** We expect the technology approaches and associated vendors to start to sort themselves out along application segments in 1999



3 Vendor Summaries

3.1 Evaluation Methodology

Our approach was pretty straightforward. To evaluate each vendor, we first gathered everything available to the public. This included web site material, published white papers and marketing briefs. We then contacted the vendors about this project, explained what we were trying to accomplish and asked for their participation. This generally met with enthusiastic support from the various vendors.

We then gathered information from our hands-on field experiences and discussions with others on their experiences.

We next interviewed the key executive and technical staff to probe for deeper technical points (i.e. How, *exactly*, does this work; what are your short and long term architectural directions, why did you paint the box pink, etc.). The only restriction we placed on these conversations is that we would only publish information about products/features that are currently shipping. Therefore, you don't have to sort through any vaporware in these reports.

We then interviewed one or two customers of each vendor to validate some of the information and to try to gather some dirt.

The next step was to write the vendor sections. After each section was complete, and had been completely chewed-up during Acuitive's internal reviews, we shipped each section to the corresponding vendor. These sections included our own observations about issues and gave our final analysis of their capabilities and direction. Vendors then had, more or less, two weeks to respond to their specific section.

Upon receipt, Acuitive evaluated the vendor responses and set up follow-up conference calls. The purpose of these calls was to make sure we understood the various comments (and complaints) and helped us gain even more insight into the various product offerings. This back-and-forth with each vendor sometimes went on for a month or two. In the end, while we didn't always reach consensus, the results are as unbiased and objective as we could make them.

3.2 Summary Format-Explanation of Each Section

The following section gives a brief description of each part of the various vendor summaries. This should help acclimate you to our little review-world, and expedite your search for information.

Company X (www.companyx.com)

3.2.1 Company Overview

This is a brief overview of how the company came into existence, how long it has been around, which divisions it is associated with (in the case of a small department of a larger corporation), a description of any published mission-statement, and who their key (or marquee) clients are.

3.2.2 Product & Support Overview

This section highlights what products are actually shipping, and what kind of support you can expect (order). We give a description of each orderable product; the interfaces, memory, processor ratings,

protocols-supported, etc. The detailed performance numbers and functionality are explained in the following sections.

We also give an overview of how each product fits into the entire vendor offering (where applicable). Finally, we mention the price of each product (valid as of our last conversation with the various vendors). This, obviously, is the most volatile piece of information. Please use these prices to compare cost-ranges, not to put together a budget. For up-to-the-second pricing, please contact your vendor(s) of choice.

3.2.3 Product Functionality

In this section we provide more detailed information about how each mode-of-operation works. Graphics are supplied to illustrate the operation of any proprietary vendor solution. In some cases, to avoid being terribly redundant (i.e. How DNS Redirection works, over and over again), we refer the reader to the Acuitive VRM Technology Report.

For each vendor, we supply an “Other Important Functions” section. This explains any mode of operation that doesn’t fall into the basic traffic-forwarding mechanisms mentioned in the table. These subsequent technologies may include QoS capabilities, effective redundancy schemes, content management/discovery, security, interesting WAN solutions and so forth. These are things you need to know about while evaluating the vendors’ basic capabilities. Some are second-order in terms of functionality, but they might just satisfy an important requirement for your site.

3.2.4 Product Features at-a Glance

This chart allows you to quickly gauge each vendors’ features and capabilities. The example provided below shows the typical answers for each Feature query.

We have used the Comments section to provide further details about each Capability. At times, we felt that some extra information was needed to fully capture the sense of the product offering. Some of these comments highlight deficiencies while others point out hidden benefits to any particular approach.

For the purposes of these tables, we have treated the entire vendor product set as one amorphous product line. For example, we have inter-mixed features associated with local VRM and multi-site VRM. The reason for this is so that you can get an at-a-glance overview of the entire solution set available from these vendors. A more detailed feature-by-feature breakdown of features will be mailed to you (and updated quarterly), which not only provides more detail but also identifies more specifically which products provide which functions in the overall product line.

FEATURE	CAPABILITY	COMMENTS
TYPE SOLUTION	Software-on-a-Stick, Switch, etc.	
OPERATING SYSTEM	Embedded BSD, proprietary, etc.	
DELAYED BINDING	YES/NO	
PERFORMANCE	(a rating for each product)	Performance data as provided by the vendor.
ROUTER OR BRIDGE	Router	
REDUNDANCY	Active-Standby, Active-Active	
AGENT TECHNOLOGY?	NO/YES	
MECHANISMS	• TCP Gateway (LAN) • MAT (LAN) • DNS Re-Direction (WAN)	(Mechanisms are the modes-of operation, or forwarding – A complete description is given in the Acuitive VRM Technology Report.)
POLICIES	**Local** • Weighted Round Robin • Least Connections • Response Time • "Observed" • "Predictive" • SSL session ID **Multi-Site** • Random • Round Robin • Least Users • Site Load Feedback • Round trip delay • Dropped packets	(Policies are items that are used to make a forwarding decision. A complete description is given in the Acuitive VRM Technology Report)
FEEDBACK	• External Monitoring • Network packet rate • TCP Connections (to port) • Active Content Verification • Dynamic Application Verification	(Feedback methods are metrics that are received/tracked by the device to effect forwarding Policies. A complete description is given in the Acuitive VRM Technology Report)

Network Management Features

This section explains the various means of collecting management data from the product. We explain which approaches are proprietary, which are standards-based and what platforms/OS are supported. If the approach is sufficiently advanced, we mention which policies can be set/governed by the management application and how content is dealt with.

Key Additional Features

This brief bullet-point list highlights any feature that is of import, but is not a first-order Load Balancing capability. Some examples are listed below.

- Packet filtering and firewalling
- Bandwidth management
- Port mapping
- Fast Etherchannel and Gigabit Ethernet network interface options
- Redundant schedulers

3.2.5 Key Relationships

This section discusses the important technology partnerships for each vendor. These relationships may augment the vendors' offerings by providing total-site control. It may simply be a way for various vendors to utilize 3rd party server agents to collect more pertinent information about a site's health.

3.2.6 Issues

Here's where we expose any shortcomings (from our point of view) for each vendor's approach. This section, naturally, drew the most fire from each vendor during the review process. However, this section is also where we display our objectivity and independence from marketing belief-structures. Let the reader use discernment. After reading a particular vendor section, you may feel that there are more issues than we have listed. On the contrary, you may not agree with our list of issues at all. In any event, please consider this section to be a catalyst for drawing your own conclusions.

For issues, we tried to limit ourselves to key architectural or feature issues that impact a long-term decision. We did **not** attempt to capture issues associated with known bugs. All the vendors' products in the VRM space basically work, but as you'd expect from complex products, there are always a few bugs. Since we view that each vendor in this space is diligently working on fixing their bugs, we don't think that today's bug list should influence a decision for a product you are intending to use for years.

If an issue is one that prevents the vendor from being suitable for a particular type of application, we have **bolded** that issue.

3.2.7 Final Analysis

Based on experience, here is where we offer our humble opinions.

Without further ado, we give you the Vendor Summaries.

3.3 Alteon WebSystems (www.alteon.com)

3.3.1 Company Overview

Alteon Networks was formed in May 1996 to develop what they termed "server switching" – technologies and solutions geared towards optimizing server availability, performance and server-application scalability. They recently changed their name to Alteon WebSystems to more closely tie the branding of the company to solutions they provide.

Their efforts have produced two sets of products: a Gigabit Ethernet NIC product line, and a line of Ethernet switches with load balancing and application redirection capabilities. Their first generation VRM-capable switch shipped in December 1997 with local server load balancing features – the first VRM switch in the marketplace. Software upgrades have added application redirection, multi-site load balancing, and URL parsing as additional VRM capabilities. They recently announced their 3rd generation switch technology, which has higher port density, VRM-mechanism functions burned into new ASICs and more software capabilities.

Alteon has received $37M in 3 rounds of funding from VCs such as Sutter Hill, Matrix Partners, Onset Ventures, New Enterprise Associates, GE Capital. They currently have approx. 120 employees, who come from both server/OS companies (Auspex, Sun, SGI) and networking companies (Bay, Fore, 3COM).

Key VRM customers of Alteon's includes Yahoo!, CMP Media, UUNet, Telstra, Rogers Communications and Digex.

3.3.2 Product Overview

Alteon WebSystems' VRM product line currently consists of the three switches with slightly different port densities and Ethernet flavors.

ACEdirector-2 / ACEdirector-3 / ACEswitch-180E

Their primary product is the ACEdirector-2, which is an 8-port 10/100 Ethernet switch that supports all of their VRM functions. The ACEdirector-3 adds a Gigabit Ethernet uplink. The ACEswitch-180E is a Gigabit Ethernet clustering switch which features 8-ports of 10/100/1000 (with dual connectors – TX/copper and SX/fiber) and a Gigabit Ethernet uplink. The AD2 is priced at $10,995; the AD3 is priced at $12,995 ; the A180E is priced at $21,995. The heart of this product line's architecture is a switching ASIC which has (2) RISC CPU cores embedded within the ASIC, allowing them to perform wire-speed L2/L3 frame and L4 session processing on any port.

Alteon 708 / Alteon 714

Alteon recently announced the 708 and 714 modular Ethernet switches, which are (4) module-slot and (8) module-slot switches, respectively. The switch is designed around a new L3/L3/L4 switching ASIC, which provides L2/L3 switching and VRM mechanisms such as NAT and TCP Diddling in hardware, while still maintaining their multiple CPU's per ASIC for software flexibility. The 708 and 714 supports all existing Alteon VRM features plus Bandwidth Management, enhanced Server Security, enhanced persistency features, new Multi-site Load Balancing features, Class of Service support plus significant routing protocol improvements. These switches are based on Alteon's 3rd generation architecture that features L4 session processing capabilities embedded in their new ASICs, and should set a new standard

for performance and price. Alteon will initially offer 16-port 10/100 Ethernet modules and 4-port Gigabit Ethernet modules for server and network connectivity. List pricing for this product line starts at $9,600 for a base system, with L4-capable 10/100 ports priced at $550.00 per port. These products were only recently announced by Alteon, and are not yet available.

Alteon WebOS

Alteon's WebOS base software is not a separately charged line item on the price list -- it's the functional software that comes with the hardware platforms described above. The software supports the following features:

- L2 Switching and Spanning Tree
- Local Server Load Balancing
- Application Redirection ; Web Cache Redirection
- URL parsing, supporting SLB and Cache Redirection
- IP Routing ; L3/L4 Filtering ; Default Gateway Load Balancing
- VRRP for L3/L4 Redundancy (beta)
- Full support for SNMP, a command-line interface and a Web User Interface
- IP Filtering for Server Security

Multi-Site Load Balancing (what they call Global Server Load Balancing) is available for an additional $3,000 per switch.

Other WebOS features such as Bandwidth Management and the support for the A7xx series of switches have not been priced.

Support

Alteon provides free support for the first 90 days after product purchase, including technical support via phone, fax and email, as well as software updates. Alteon also has relatively industry standard support options available to support customers and resellers beyond the first 90 days. These options include technical support during standard business hours, with a further option of 24x7 support; switch software upgrades; advanced replacement for all hardware.

3.3.3 Product Functionality

Alteon's VRM features are supported across their entire switch product line.

Local Server Load Balancing

The Alteon local server load balancing implementation is typically installed in a HNAT configuration. The switch publishes a V_IP which users send requests to, which are then HNAT rewritten by the switch and forwarded to the servers. They also support Full NAT and IP Routing for topological flexibility, as well as MAT to support IPSec servers and non-server load balancing. Their feedback methods include Active Content Verification, DAV and an API. Their load balancing policies include Least Connections, Round Robin, Weighting with secondary policy support of MaxConns and Backup/Overflow servers.

Their newest release of software supports the Delayed Binding mechanism for local load balancing, which allows them to perform URL parsing for server load balancing. This URL parsing feature allows the administrator to assign specific or wildcard matched URLs to individual server groups, allowing tuning of their server resources.

Multi-Site Load Balancing

Alteon has supported Multi-Site Load Balancing (what they call Global Server Load Balancing) since November 1998, which combines local server load balancing with multi-site load balancing using an optional GSLB software license key. They do not use or require a centralized multi-site VRM scheduler; instead, all multi-site decision-making is distributed among switches at each site performing both local VRM and multi-site VRM functions. Active Content Verification and DAV extensions test and verify multi-site/server health and response time. An inter-VRM protocol (what they call the Distributed Switch State Protocol or DSSP) communicates server and site load, health and response time information between switches, to allow the distributed decision-making process to work smoothly. WebOS GSLB provides DNS Redirection and HTTP Redirection mechanisms, and multi-site policies of Site Performance Measurement as well as what Alteon calls IANA Source Awareness. IANA Source Awareness uses the source address of the client/client LDNS request to locate the client within a rough geography based on published IANA IP subnet information.

A WebOS software release currently in beta allows the use of the IP-Proxy as a site overload/server-failure backup mechanism.

Application Redirection

Alteon also supports the transparent interception and redirection of application specific flows. This allows the transparent load balancing (which gets you high availability) of almost any IP-enabled system. The most common applications for this technology includes:

- firewall load balancing
- transparent proxy web cache load balancing
- DNS redirection
- router load balancing

The network administrator creates one or more filter rules which is described by any of the following L3/L4 frame headers:

- source and/or destination IP address (specific host or ranges)
- Protocol type (TCP, UDP, ICMP, etc)
- source and/or destination TCP/UDP port numbers (or ranges)
- TCP flags

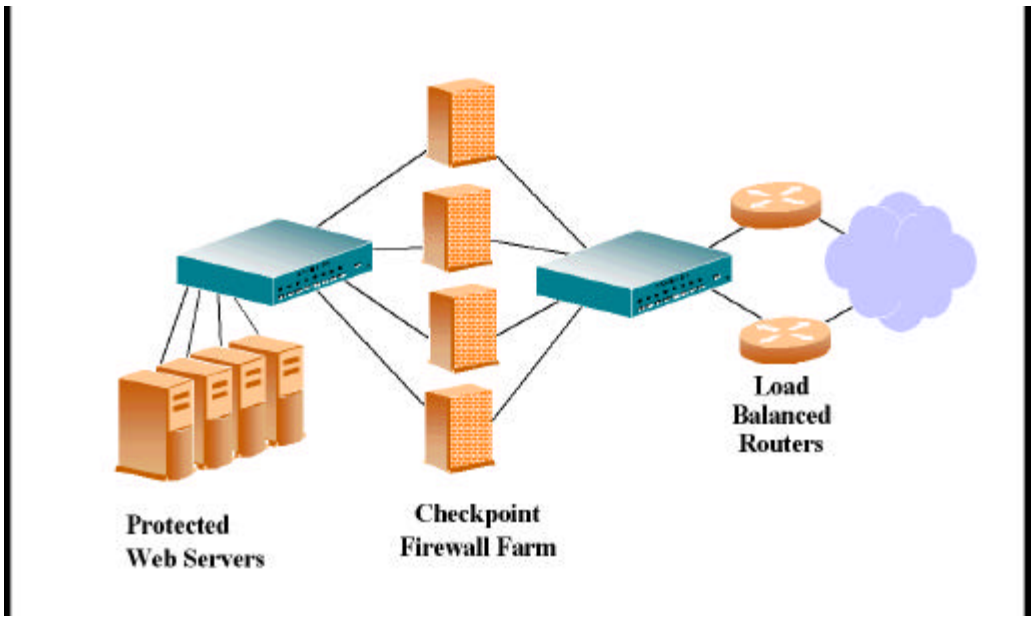
Then, the rule has an **action** associated with it, which may be set to DENY, ALLOW, NAT or REDIRECT (with MAT/ NAT, URL options). DENY and ALLOW are commonplace functions– they can be used for classic packet filtering, which Alteon provides to offload routers from the task of protecting servers. These filter rules may be applied individually per port, allowing both symmetric and asymmetric filtering.

The REDIRECT action allows specific frames or sessions to be redirected to a server group using a VRM mechanism such as MAT or NAT. MAT allows the transparent redirection to devices such as routers,

firewalls or transparent proxy cache servers, while NAT is used to redirect sessions to remote server systems. URL parsing may be enabled for web cache redirection, where the use of the cache server may be optimized by determining which requests are cacheable, and which are not.

The server group specified in the redirection rule can have between one and 256 servers associated with it. If there is more than one server (which may be a server, or may be an IP forwarding device) in the group, load balancing can take place between the servers using appropriate server-selection policies. Feedback from the load balanced servers/devices takes place either through Active Content Verification and Device ICMP Ping, or via the Alteon WebOS API.

Local load balancing and application redirection can be combined together to allow the load balancing of firewalls on the "front-end" and application servers (such as web servers) on the "back-end."



In the figure above, client requests enter the network from the Internet, passing through one of the routers into the Right Alteon switch. Using a firewall load balancing policy, frames are transparently redirected to one of the four firewalls, which perform appropriate firewall functions and forward appropriate frames to the Left Alteon switch. Alteon's firewall load balancing policy ensures that frames between the same communicating client/server always pass through the same firewall. Client requests arrive at a V_IP in the Left Alteon switch, and are load balanced among the available web servers. Response frames from the web servers pass through the Left Alteon switch, which ensures that they pass through the same firewall as the clients' request frames. The forwarded frames arrive at the Right Alteon switch, and are load balanced between the two Internet routers. The Alteon switches either use Device ICMP Ping to test forwarding path connectivity through the firewalls and/or agents on the firewalls communicate to the switches using the Alteon Server Switch API.

Other Important Functions

Redundancy

Since Alteon ships switch products, they are easily capable of offering redundancy configurations that feature fully meshed topologies.

Alteon's current WebOS software supports Hot Standby redundancy, where one switch is actively performing VRM functions and the other is either totally in standby mode, or is passing traffic around a topology break. Alteon has a release of software currently in beta that supports both Active-Standby and Active-Active modes of redundant switch operations through the use of the VRRP protocol.

Alteon's new A700 series of switches are designed with fully redundant power supplies, switch fabrics and switch management processors, and are designed for NEBS-1 compliance.

IP Routing

Alteon's switch products support IP routing between subnets/VLANs. For instance, servers may be placed in a privately addressed subnet, and virtual servers may be available on a different publicly available subnet, with the Alteon switch providing routing services between the subnets.

EtherChannel

Alteon offers Cisco-compatible Etherchannel (called Link Aggregation in IEEE-speak) to provide redundant wire connectivity and to provide higher throughput. Alteon has improved on the original Etherchannel method of operation by performing "transmit-wire" decision-making using both IP and L4 header information.

NAT

Network Address Translation is available as a filtering rule action, where frames/sessions matching the filter rule may have their source or destination addresses translated. This is useful for both server administration, where the servers are on private networks and must be administered from a public network, and for address translation services for privately addressed client workstations.

Server Switch API

Alteon provided a Server Switch API in November '98 to allow customers to create server-resident monitoring software to provide feedback to their switches. Simple application of the technology may take into account basic application availability information, or server load information, to cause the switch to perform different server selection behaviors. More complex API usage might involve the use of content distribution software that dynamically controls which real servers are available and which are not, allowing simple distribution of new content to all servers in a controlled and automated manner.

Packet Filtering

Alteon's filtering system may be used for server protection, where frames/sessions are matched against the filter rules applied to a port and appropriate ALLOW/DENY/REDIRECT actions are performed. Alteon has said they will expand this capability on the A7xx products to include dynamic filtering with session state recognition.

3.3.4 Product Features at-a Glance

FEATURE	CAPABILITY	COMMENTS
TYPE SOLUTION	ASIC-based Switch, with multiple CPUs per switch port for VRM functions.	A7xx series embeds L4 session processing in new ASICs.
OPERATING SYSTEM	Embedded	
DELAYED BINDING	Yes	For URL parsing, HTTP Redirect.
PERFORMANCE	ACEDirector-2/3: 20,000 connections per second at 800Mb/s. ACEswitch-180E: 80,000 connections per second at 4.5 Gb/s. 256,000 sessions per switch. A7xx – 350,000 connections per second. 32 million sessions per switch.	Vendor supplied data. PC Week labs tested a similar Alteon switch at 10,000 HTTP requests per second/630Mbs in Jan '98. Each request had an 8k payload.
ROUTER OR BRIDGE	Both/Either	Load balancing of traffic to up to (4) default gateways.
REDUNDANCY	Hot Standby	Software release in beta supports VRRP that allows both Active-Active and Active-Backup configurations.
AGENT TECHNOLOGY?	NO	Alteon provides the Server Switching API which allows server-side agents or third-party monitoring systems to affect switch behavior. Requires systems integration.
MECHANISMS	• NAT • Full NAT • MAT • IP Encapsulation • DNS Authoritative Server for VIPs • HTTP Redirect	DNS, HTTP Re-direction used in WAN load balancing. IP-Proxy supported in upcoming software release, currently in beta.
POLICIES	**Local** • LeastConns • Round Robin • Weighting • MaxConns • URL Parsing **Non-Server LB** • Hashing • MinMisses • URL Parsing **Multi-site** • SPM, plus IANA Source	Backup and Overflow servers
FEEDBACK	• Device ICMP Ping • TCP Connection Verification • Active Content Verification • Server Switch API	Alteon supports HTTP, DNS, POP3, SMTP, NNTP, RADIUS and IMAP for DAV server health checking.

Management Features

All of Alteon's features can be configured and monitored through a menu-driven/command driven CLI, a browser-direct Web user interface or through SNMP. Alteon also provides a HP-OpenView integration kit for monitoring and reporting from the HP-OpenView network management platform.

Other Key Features

- As part of either the IP routing capability or the application redirection capability, Alteon's WebOS is capable of load balancing outbound traffic to multiple routers. This allows the network administrator to augment their local load balancing solution by ensuring that adequate and reliable egress bandwidth is available for return traffic being sent to users.

- WebOS supports "Shopping Cart" applications, binding client IP addresses for both port 80 and port 443 connections.

- WebOS Supports deterministic session assurance based on IP address. When a fail-over occurs, the (now) active switch is capable of determining which users are associated with which servers using an algorithm based on source IP address and protocol.

- Server slow start, Graceful deactivation

3.3.5 Key Relationships

Alteon has sales and marketing relationships with all of the leading web proxy cache vendors, Inktomi, Network Appliance, Cacheflow, and Cobalt.

Alteon is a Checkpoint OPSEC partner.

Alteon recently formed an alliance with Hewlett-Packard to deliver their product offerings in the HP Covision channel program. HP uses this program to bring together "best-of-breed" technology (from applications to service providers) to address their client's web-centric needs.

3.3.6 Issues

- Alteon has not specified their performance when performing delayed binding nor have we seen any independent 3rd party testing on the performance in that mode. Based on our knowledge of the architecture of the Alteon products, we suspect the performance is pretty good, but for design purposes, we'd like to see some quantitative information
- Alteon does not support a packet/byte rate based load balancing policy
- While they have announced support for Bandwidth Management and Class of Service for their 7xx switches, Alteon does not support any preferential service capabilities today
- Their current switch technology only supports up to 256k concurrent sessions. While it is enough capacity to support many applications, Alteon needs to deliver their next generation switch (which will support millions of sessions) to be able to support every application
- Alteon's integrated local/multi-site VRM features allows for a decent level of user-site decision-making, but we feel that Alteon needs to add finer granularity of control of user-site request distribution, such as Static Client Preference

3.3.7 Final Analysis

Alteon was the first vendor to ship a switch-integrated VRM product and they were the first VRM vendor to demonstrate super-fast session binding rates.

We like the present Alteon price/performance very much. They have also solved some of the feature issues we noted in our last report. They are shipping a good multi-site solution, they have added URL parsing capabilities to their software, and they have added Active Content Verification, DAV and an API for server feedback. Alteon's solutions can meet most customer VRM needs and they are almost always among our short list of vendors to consider for any VRM application.

Alteon has taken some risk in the marketplace by announcing their new products well before availability. On paper, the next generation switches look excellent in design and capability. But they need to deliver it. It could be fall or later before these products are generally available. In the mean time, it's unclear how many critical features will be supported in the future on the Alteon platforms you can buy and deploy now. The good news is that the present products are well featured and mature. You should get them to agree to a high percentage trade-in value for switching to the new products when they become available (and time tested).

When the new products do become available, you'll want to consider how to provision servers relative to these boxes. Right now, if the site consists of less than six servers or so, we plug them directly into Alteon switch ports. If the site requires more than a modest number of servers, we generally surround the Alteon switches with lots of dirt cheap (but very reliable) Fast Ethernet hubs so that one switch can support many tens or even hundreds of servers. This works because none of the servers (especially if they are NT) individually are running at anywhere close to 100 Mbps speeds. But as the server capabilities increase, you'll want to consider plugging them directly in to switch ports, which will be more plentiful in the new products.

3.4 ArrowPoint (www.arrowpoint.com)

3.4.1 Company Overview

ArrowPoint was founded in September 1997 with the goal of providing integrated website management solutions for Web Hosting, E-commerce and ISP Points-of-Presence or intranet Web farms. To date, they have received $33.5 million in venture capital from firms such as North Bridge Venture Partners, Matrix partners, Accel Partners, Bowman Capital and Pequot Capital. They have 65 employees and are located in Westford, MA.

ArrowPoint has developed what they call Content Smart™ Web switching, which includes L2/L3/L4 and URL switching, Server Load Balancing, Bandwidth Management, firewalling and content management. These capabilities are delivered on two different switching platforms: the CS-100, a 12/16 port 10/100 Ethernet switch and the CS-800, a modular Ethernet switch that can support up to 64 10/100 Ethernet ports. ArrowPoint began shipping the CS-100 in late 1998, and has been shipping the CS-800 since February 1999. ArrowPoint's key customers and partners include:

Service Providers		Content Providers	E-commerce
- UUNET	- Cube	- MPath	- ToySmart
- Navisite	- Exodus	- ECAL	- Raging Bull
- RoadRunner	- World Online		- MotherNature.com
- Germany.net	- STIC		

3.4.2 Product & Support Overview

CS-100

The CS-100 is an Ethernet Web switch with either 12 or 16 10/100 Ethernet ports, or with 4 port 100Base-FX ports, depending on port options purchased. It is built upon the MMC 5500 chipset, the same switching chipset found in the Cisco 8500, which gives them a 5 Gbps switching fabric, per-flow QoS, wire-speed L2/L3/L4 switching, and NAT processing. A RISC (MIPS) CPU with 128MB of RAM is tasked with HTTP flow setup, URL parsing and server determination, their other content-aware switching capabilities, background tasks such as routing protocols and switch management interfaces. Full IP routing, including routing protocol support for RIP, RIP2 and OSPF is also provided. The switch can be used for both Web Cache Redirection and bi-directional Server Load Balancing, with both solutions taking advantage of the switch software's URL processing capability. The switch's US list prices range from $17,995 to $21,995 depending on purchase options and including base software, and the US list price for their Enhanced Feature set software (Distributed Web site services, Content Replication) is $7,995.

CS-800

The CS-800 is built on the same general architecture as the CS-100, but in a modular form with greater port density capabilities. This modular switch is outfitted with a 20 Gbps switching fabric (with optional redundancy), and can be outfitted with DC power and is NEBS-1 compliant. It can house up to 64 full duplex switched 10/100 ports, or up to 48 10/100 and 16 100Base-FX ports (modular – 8 per card), and up to 32 Gigabit Ethernet ports (4 per card). This platform has the same software capabilities as the CS-100, including URL parsing and IP routing, which runs on 4 RISC (MIPS) CPUs with access to 1024 MB of DRAM, and a management RISC processor with 128 MB of DRAM. The US list price of the CS-800 (includes (1) 10Gb/s fabric, (1) power supply and (1) System Control module with base software) starts at

$35,000, 8 port Fast Ethernet modules start at $5,495, and the US list price for their Enhanced Feature set software (Distributed Web site services, Content Replication) is $19,995.

Support

The basic warranty provided is a one-year hardware replacement warranty. Flexible support plans can be built and custom tailored to meet the customers needs from the following services: 7x24 phone support, secure access to a their web site, training and software updates. Each of these may be purchased for a separate fee.

3.4.3 Product Functionality

The ArrowPoint CS product set is full-featured, including local URL-aware load balancing and multi-site load balancing, ICSA compliant firewalling, QOS management and traffic shaping capabilities, wire-speed L2/L3/L4 switching and IP routing support.

Local Load Balancing

ArrowPoint supports local server load balancing with the servers directly attached to their switch or off of subordinate L2/L3 topologies behind the switch. They were the first VRM hardware vendor to popularize and deliver local server load balancing with URL parsing capabilities (what they call Content Aware Switching). URL parsing allows the administrator to optimize the server farm by sending requests of different types to different servers, and also handle "flash crowds" by dynamically replicating hot content to overflow servers.

URL parsing also allows ArrowPoint to improve server performance by sending sequential requests for the same content to the same server, as it is likely that server still has it in it's local cache. If you use this feature, you wouldn't even need to consider front ending your server farm with a Reverse Cache Proxy.

For e-commerce and authenticated services, the CS can provide "sticky" or persistent connections using Source IP, SSL application ID, and most significantly using cookies. This allows AOL and other mega proxy clients to be stuck to the best server throughout their transaction.

TCP Diddling

To forward requests and manage connections, the CS switch products perform *delayed binding*. When the client makes a request to a content server, the CS receives the incoming request, and then responds on behalf of the server, so the client thinks it has a direct connection. When the client next sends the URL (or content) request, the CS can inspect the packet in order to determine what content (server location) is requested. The Content Switch will then set up a separate connection with the most capable and available server of the moment and pass on the client's request. The fulfillment server sends the traffic directly back to the CS, where it is transferred to the client using the TCP parameters of the original connection, with wire speed bi-directional NAT. This process does not require any software on the server. We call this mechanism "TCP Diddling." It requires that the CS be in the traffic path in both directions – client-to-server and server-to-client.

Cache Redirection

Using their content-aware capabilities, ArrowPoint is capable of intercepting, pre-processing and forwarding user HTTP requests to a farm of web cache servers. Because they have URL parsing capabilities, they can determine whether or not the requests are cache-able, and can therefore determine whether they should be sent to the cache or to the origin server. This feature allows optimization of cache-use and cache processing time -- the cache only receives requests it is capable of quickly handling.

They are able to perform sophisticated URL analysis algorithms on the users HTTP request to determine the location of the requested objects. This further optimizes the cache servers by ensuring that requests for the same Web object always go to the same cache server, thereby increasing cache hit rate. They can properly and efficiently support all forms of proxy caching modes: Transparent Proxy, Forward Proxy and Reverse Proxy.

Multi-Site Load Balancing

DNS REDIRECTION

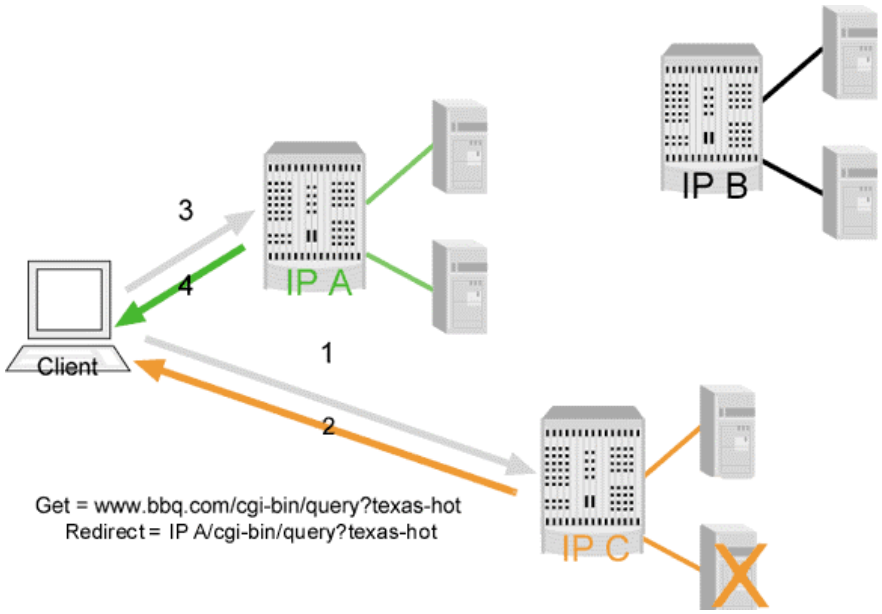
The CS product supports the standard DNS redirection mechanisms as described in the VRM Technology Report. ArrowPoint's unique value-add to this technique results from their use of their proprietary Content and Application Peering Protocol (CAPP). This protocol allows the switches to exchange a wealth of information related to site and content availability. When a DNS request comes into a CS switch, it is capable of directing the user to the best site based upon its information such as:

- Server Health
- Network Health and Content availability
- Content Quality of Service Requirements
- Server and Network Load
- Topological Proximity

Please refer to the ArrowPoint document *Content Smart Switches for Distributed Web Sites* for a detailed explanation of this capability. The document is available on their web site.

HTTP REDIRECTION

The ArrowPoint switches can detect content or server availability problems and then use HTTP Redirect to redirect users away from one site/server to another site/server. For an in-depth discussion of the HTTP redirect process, please refer to the Acuitive Virtual Resource Management Technology Report. ArrowPoint's value-add for this feature is their content awareness. They are capable of determining that a particular request will not be locally fulfilled due to resource unavailability, and can redirect the user's request to a known working server that has the content available.



1. In this example, after opening a TCP connection, a client sends a request for [www.bbq.com/cgi-bin/query?texas-hot] to the switch labeled **IP C**, which performs TCP termination and URL inspection. Upon inspecting the request, the switch determines it is not able to fulfill the request because a particular server (the CGI server) is unavailable.
2. The switch formulates an HTTP Redirect packet and sends it to the client, telling the client to go to **Site A** (and switch IP A) to retrieve the content for this URL. The TCP connection is closed.
3. The client opens a different TCP connection to the switch labeled **IP A**, and makes the same request. The request is fulfilled from IP A.

Other Important Functions

Redundancy

The CS-800 is fully NEBS-1 compliant, can support optional redundant switch-fabrics, power supplies and System Controller modules and offers no single point of hardware failure. All modules are hot swappable.

Both CS-800 and CS-100 can be configured with box-to-box redundancy using VRRP with <3 second fail over on box failure. On a single box, redundant, load sharing uplinks can be configured using ECMP (Equal Cost Multipath Routing.) Active-Active redundancy will be available 2H'99.

Intelligent Content Discovery and Optimization

By watching incoming requests and the subsequent server responses, the CS products can "learn" which servers provide the best responses for specific content. The CS switches can also probe (open TCP connections) with the servers to round-out this content availability picture. ArrowPoint also uses web-spider content discovery technology that can retrieve the content from all known servers to build a content map. This content map can be stored in the switch's memory/disk and is consulted during URL intelligent operations.

Content Based Quality of Service

The administrator can assign various response time/bandwidth levels for various types of content. For example, you may wish to assign higher priority to SSL traffic than FTP traffic, or assign HTML files a higher transfer priority than GIF images. As the CS observes response times for a critical application heading towards the pre-defined threshold, an attempt to maintain the service level will be made using one or more of the following mechanisms:

- Stop sending traffic for lower priority services to some servers which also support the higher priority service
- Throttle back requests for lower priority services

This capability also allows you to offer certain customers or content Service Level Agreements, based on response times.

Content Replication and Management

ArrowPoint provides tools for replication content between sites and between disk drives and file systems locally. These tools can be used to update sites, rollback changes, and generally reduce the cost of operations of managing frequent content changes.

For sites where content is stored on the web servers' local drives, ArrowPoint's switches can replicate content on additional servers and balance the load to those servers in response to "flash crowds," which are situations where requests for some particular content dramatically increase, such as a news report divulging that the president of some world power has been discovered to have done something naughty.

Security

The CS products provide wire-speed, ICSA-compliant firewall, URL Blocking and bi-directional NAT capabilities on any port. Their TCP termination capability allows them to provide Denial of Service attack protection.

3.4.4 Product Features at-a Glance

FEATURE	CAPABILITY	COMMENTS
TYPE SOLUTION	ASIC/Microcode-Based Switch with multiple Centralized CPUs for VRM functions	128 MB RAM per processor standard
OPERATING SYSTEM	Embedded	
DELAYED BINDING	YES	Including URL parsing and cookie support
PERFORMANCE	20,000 connections/sec (delayed binding) 40,000 connections/sec (immediate binding) 1,000,000 HTTP sessions	Vendor-supplied information.
ROUTER OR BRIDGE	ROUTER and Switch	
REDUNDANCY	Active-Standby	VRRP Active-Active announced for 2H99 OSPF Symmetrical Path Routing for Active-Active
AGENT TECHNOLOGY?	NO	Partnership with Bright Tiger and WebSpective for more granular server feedback.
MECHANISMS	• DNS Redirection • HTTP Redirection • TCP-diddling • MAC Address Translation	
POLICIES	• WAN least cost/closest • QOS maximum allowed response time metrics • URL parsing/content specific • Predictive and observed server response time (socket level)	Backup and Overflow servers, dynamic content replication
FEEDBACK	• TCP Connection Verification • Response time monitoring • Active Content Verification • Dynamic Application Verification • Passive Content Verification	

Management Features

Both the CS-100 and CS-800 are managed and configured through the ArrowPoint “FlowMinder” management system. This is a Java-enabled browser-based tool for configuration, policy setting, statistics gathering and reporting for resource and site utilization. Additional capabilities are command line interface support, SNMP and RMON (port-mirroring included) and a log file.

Key Additional Features

- Full TCP/UDP/IP protocol suite support
- Source Group NAT
- Bi-directional balancing of flows
- Cookie-based sticky connections for E-commerce (we think ArrowPoint is the only company claiming this capability right now)

3.4.5 Key Relationships

ArrowPoint recently formed an alliance with Hewlett-Packard to deliver their product offerings in the HP Covision channel program. HP uses this program to bring together "best-of-breed" technology (from applications to service providers) to address their client's web-centric needs.

They, like the other switching vendors in the document, have announced sales and marketing relationships with each of the major cache vendors, including Cacheflow, Inktomi, and Network Appliance.

Lucent and ArrowPoint have announced an OEM deal that will allow Lucent to brand the ArrowPoint Web switches as Lucent Traffic Director WD-100 and WD-800. They will be offered through all lucent channels and as part of Lucent's IP WorX caching product family, developed internally by Bell Labs.

3.4.6 Issues

- ArrowPoint needs to develop and support an API for integration to 3rd party server-resident instrumentation. {ArrowPoint says that as part of 3.0 software release in beta now, ArrowPoint will provide a full, secure API that will allow partners or customer applications to control all aspects of the switch}
- ArrowPoint uses the same MMC chipset featured in the Cisco 8500 for their L2/L3/L4 switching capabilities and the switch fabric. It provides per flow QoS support, but in an ATM-like way, which is queuing oriented. We prefer TCP Rate Shaping. But we'll admit it's somewhat a religious issue at this point. More importantly, ArrowPoint's bandwidth management functions, although useful for limiting the flow of traffic to certain applications, or slowing down traffic flows to stressed servers, does not generally provide the means to manage the WAN link bandwidth. Another product must be introduced for that function, which complicates the administration of consistent policies
- ArrowPoint claims wirespeed L2, L3, L4 and URL parsing support. Our experience with the MMC chip set tells us that wire speed L2/L3 should be no problem. But we'd like to see some verification of their L4 NAT and delayed binding performance claims by an independent 3rd party. ArrowPoint tells us that they agree and they are working on it. Stay tuned…
- Customers tell us that it can take minutes to boot up an ArrowPoint switch. That's an awful long time. ArrowPoint tells us that they are looking to address that issue

3.4.7 Final Analysis

ArrowPoint is a visionary company. They were the first, in our opinion to define the ultimate requirements for website management in terms of:

- The collections of functions required (local load balancing, multi-site load balancing, routing, firewalling, bandwidth management, content awareness)
- Content management, content awareness, and content access optimization
- The optimal target architecture – a scaleable software platform with hardware assist for key functions

We like the idea of URL awareness for server optimization and content-aware decision making. Many vendors support this now. But ArrowPoint has taken the concept of content awareness even further, providing means for ensuring the integrity of content returned to clients, the means to identify content being accessed regularly and to adjust the server bandwidth available to serve that content and to add value to all forms of caching that a user might want to deploy. ArrowPoint is wrapping their entire positioning and corporate strategy around adding value from being aware of the location, status, and integrity of content.

It's very possible that ArrowPoint will be the first vendor to deliver "the dream machine."

But – ArrowPoint has taken a huge bite here and it's not clear whether they are going to be able to swallow and digest. They have been late in delivering products to market, almost one year behind Alteon and nine months behind Foundry, who themselves were 6-12 months behind the software and SOAS vendors. Many customers have bought into the ArrowPoint story only to wait and wait for delivery, sometimes only to be disappointed when early releases were shipped.

We hope the development and delivery problems are behind ArrowPoint. Their vision and value-added features are alluring. But make sure they demonstrate the features you need using production code. If they don't have the features you need now, we'd wait until they do, at least until they demonstrate a consistent ability to deliver on promised products and features.

3.5 Cisco (www.cisco.com)

3.5.1 Company Overview

Cisco (maybe you've heard of them) was the first to market a Server Load Balancing product, introducing their "Local Director" technology in 1996. Cisco has since announced a multi-site product, Distributed Director.

Marquee customers include Digital Island, Excite, SportsLine USA, and Post Communications.

3.5.2 Product & Support Overview

The Cisco product line is segregated into Local and Distributed areas of functionality. A separate device services each area.

Local Director

The Local Director product line provides local load balancing capabilities. Cisco has recently updated the hardware platforms, where they now provide two products: the Local Director 416 and the Local Director 430. The LD-416 supports up to 12 Ethernet/Fast Ethernet interfaces. Redundancy is carried out via a second 416 and an RS-232 cable. The LD-416 is powered by a Pentium 166MHz, and ships with 32MB RAM, and has a 3.5" disk drive and 2 MB of Flash memory.

The Local Director 430 has the same general design as the LD-416, but supports up to 16 Ethernet/ Fast Ethernet interfaces (TR and FDDI support optional). The LD-430 is powered by a Pentium 300MHz, and ships with up to 384MB RAM. Unlike the 416, the Local Director 430 is EtherChannel capable. This interface-trunking scheme allows the 430 to offer a 400Mbps connection (more on this in the "Product Features at a Glance" section).

Distributed Director

The Distributed Director product line provides multi-site load balancing capabilities. It is built on an entirely different hardware/software platform than the Local Director products. Cisco ships three different Distributed Director products: the DD-2501, DD-2502 and the DD-4700.

The DD-2501 supports one Ethernet interface and 2 Serial interfaces, is powered by a 20MHz 68030 processor, and ships with 8MB DRAM RAM. The DD-2501 also provides a RJ-45 console port and 8 MB of Flash memory, and can use either AC or DC power supplies. The DD-2502 is identical to the 2501 with two exceptions: It comes with a Token Ring port and 2 Serial ports, and cannot accept DC power.

The Distributed Director 4700 supports one of the following interfaces (depending on which 4700 you order): Ethernet, Fast Ethernet, Token Ring, Dual Attached FDDI (multi-mode). The DD-4700 is powered by a 133MHz IDT Orion RISC processor, and ships with 32 MB DRAM. The 4700 also provides 16 MB Flash memory and 128 KB of NVRAM, and can use either AC or DC power supplies.

Support

Service plans include toll-free 24x7 technical help line support, free software updates, full access to Cisco Connection Online for do-it-yourself support and 24-hour advance ship replacement system should a unit fail.

3.5.3 Product Functionality

The Local and Distributed Directors have various modes of operation.

Local Load Balancing

Local Director is typically installed in a two-port configuration, with one port connected to the backbone network and one port connected to the server hubs/switches.

Half-NAT Address Translation

Cisco Local Directors are mostly commonly configured to perform HNAT. This eliminates the need to configure IP addresses of the server loopback interfaces. But it does require the server-to-client traffic to return through the Local Director. Care must be taken when establishing the topology to ensure traffic flows back correctly.

MAC Address Translation

This option enables direct-return of traffic from the server-to-the-client. This is often preferred for high performance systems because there can be an order of magnitude more traffic in the server-to-client path than the client-to-server path. In this mode of operation, the content servers are configured with the Virtual IP address on their loopback interface. The Local Director performs MAC Address Translation of the TCP SYN packet and all ensuing packets from the client, to send the packet to the intended server, which can then respond directly to the client.

TCP Diddling

With the most recent software release, the Local Director is capable of performing SSL Session ID Tracking via *delayed binding*, using "TCP Diddling." It requires that the Local Director be in the traffic path in both directions – client-to-server and server-to-client.

Local Director uses combinations of TCO, TCP Connection Verification and Device ICMP Ping for feedback monitoring of the servers. They recommend least connections or weighted round robin load balancing policies.

Multi-Site Load Balancing

Distributed Director provides Cisco's multi-site load balancing capability. One or more DD units are setup in a generally central location for either DNS Redirection or HTTP Redirection methods of request distribution. Distributed Director uses TCP Connection Verification for feedback monitoring of sites/servers, and can use an IVP between itself and router agents at each distributed content site. Distributed Director and Local Director do not communicate with each other using and IVP, effectively isolating the local and multi-site load balancing functions.

DNS Redirection

To provide continuity for a multiple geographic site web presence, the primary site Distributed Director can act as the Authoritative DNS server for the domain. WAN routers can be equipped with special agents that allow them to communicate directly with the Distributed Directors. These router-agents can provide information back to the Director on "WAN Distance," as measured by number of Autonomous System hops and Interior Routing Protocol path costs. WAN latency measurements are also supported. By maintaining a table containing the previous information and related V_IPs, the Authoritative Director will respond to iterative DNS queries with the V-IP addresses of a healthy target site. This provides transparent request distribution across servers located around the world. This also acts as a WAN fail-over mechanism in the case of catastrophic site failure.

HTTP Redirection

Alternatively, the DNS system can be configured to "point" to the Distributed Director, whereupon it will re-direct client HTTP GET requests using standard HTTP redirection. Using the same feedback from site routers discussed above, the Distributed Director will determine which site is most capable of handling the request, and send the client an HTTP redirect code (with appropriate URL). For an in-depth discussion of the HTTP redirect process, please refer to the Acuitive Virtual Resource Management Technology Report.

Other Important Functions

Redundancy

Cisco Local Director products in a redundant configuration consist of two identical pieces of equipment connected via a serial cable. Configuration changes made to the Active unit will automatically be transferred to the Standby. When a failure occurs, the units toggle operations as one picks up for the other and exchange Media Access Control (MAC) addresses. "Hello" packets are sent over the serial cable, along with power-supply status information. All networks interfaces send "Hello" packets every 30 seconds as well. The network activity is monitored to help the Director determine what type of failure has occurred. Using one of four predetermined metrics, the Local Directors can determine what action (if any) is required.

Local Director also supports Session Assurance (what they call "stateful failover"), where the connection (binding) tables are shared between the Active and Standby units so that if the Standby becomes Active, no sessions are dropped. This requires a higher bandwidth connection between the boxes. A dedicated LAN connection or an in-band network connection can be used for this purpose.

Distributed Director redundancy is quite different from Local Director. The backup Distributed Director can be provisioned anywhere in the network. It does not have to be located with the primary DD. However, Session Assurance is not supported between Active and Standby Distributed Directors.

Etherchannel

Up to 4 Ethernet or Fast Ethernet interfaces can be logically "grouped" together on a Local Director. This means that if a box is equipped with four 100BaseT interfaces, and they are connected to another EtherChannel capable device, traffic can be passed along as if this were one big pipe. Traffic is spread across the interfaces on an IP SA/DA pair basis. For further information on this topic, please refer to the Cisco web site. There, you will find a wealth of detailed EtherChannel information.

3.5.4 Product Features at-a Glance

FEATURE	CAPABILITY	COMMENTS
TYPE SOLUTION	Router platform (DD) PC Platform/SOAS (LD)	**LD**- 300-166MHz P2, 32-384 MB RAM **DD**- 133 MHz Orion RISC, 20 Mhz 68030, 8-32 MB RAM
OPERATING SYSTEM	Embedded OS	"lean & mean" (vendor quote)
DELAYED BINDING	YES	For SSL Session Tracking only
PERFORMANCE	• Up to 1 million simultaneous connections (LD 430) • 80 MB throughput (LD 416) • 200 MB throughput (LD 430)	EtherChannel required to improve the layer 2 performance.
ROUTER OR BRIDGE	Router/Bridge	Allows content servers to reside on different subnets.
REDUNDANCY	Hot Standby	Serial Cable, network hello packets. Power supply monitoring
AGENT TECHNOLOGY?	YES (DD)	Cisco IOS has been expanded to include communication protocols and metrics that are used by the Distributed Directors. Not exactly an "agent," but a software module nevertheless.
MECHANISMS	• HNAT (LD) • TCP-diddling (LD) • MAT (LD) • DNS Redirection (DD) • HTTP Redirection (DD)	
POLICIES	• Client-assigned (bind source IP address to specific servers) • Round Robin (LD) • Least Connections (LD) • Predicted Least Connections (LD) • WAN least cost/closest (DD) • Server response time metrics (DD)	LD supports weighting.
FEEDBACK	• TCO (LD) • TCP Connection Verification (LD) • Device ICMP Ping (LD) • Response time monitoring (DD) • Routing Tables (from routers running proper IOS versions) (DD)	

Management Features

The Local Director supports CLI management for configuration and monitoring, and SNMP management for monitoring

The Distributed Director features CLI management, support SNMP for monitoring.

Key Additional Features

- Local Director supports transparent proxy devices, including firewall and cache servers
- LD supports IP packet filtering
- LD provides UDP support
- LD supports Shopping Cart application (or what Cisco calls "sticky buddies"), binding client IP addresses for both port 80 and port 443 connections

3.5.5 Key Relationships

Cisco has made an investment with Webspective, who specialize in content generation and content management. We're not sure what the extent of this relationship is.

Cisco has also announced a relationship with Hewlett-Packard, integrating their WebQoS functions in with Local Director in some way.

3.5.6 Issues

Local Director

- **No Active Content Verification (ACV) or Dynamic Application Verification (DAV)**
- As with other single CPU solutions from other vendors, we wonder whether Intel/BSD route will remain viable during a period such as this where available WAN bandwidth is increasing at a pace greater than Moore's Law. This architecture will be stressed in high-end situations, especially if resource-intensive features such as SSL Session ID Tracking are turned on and used extensively
- The use of a short RS-232 cable for failover prevents deploying back-up boxes in different data centers, in different buildings, on different power grids, etc.
- Lack of an Active-Active redundancy scheme limits scheduler scalability. You are limited to what you can get out of one box, which is already stressed due to the single CPU/Intel architecture
- No URL-based scheduling. No cookie-based policies
- No preferential service capabilities, such as bandwidth management
- The Local Director and Distributed Director products have no common integration points. They are developed and supported out of two entirely different organizations. LD provides no feedback into DD policies. You have to install and configure each entity separately to perform their respective functions independently of one another. Good luck doing that. They have different user interfaces and different management tools. Given that DD comes out of the router organization, it leverages some of the Cisco SNMP-based management tools. Local Director does not
- No open and documented interface for receiving feedback from server-resident instrumentation

Distributed Director

- No site persistence value-add features
- Must choose between DNS Re-Direction or HTTP Re-Direction. We prefer a layered approach where we can use DNSR as the primary mechanism and a 2nd mechanism (such as HTTP Re-direction) to handle short term transition problems which DNSR cannot solve
- For useful multi-site metrics/polices, DD requires that you install recent router software releases and turn on an IVP reporting agent, which can cause a lot of churn if that's the *only* reason you are upgrading the software code. Also adds yet another load on the already potentially overloaded routers

3.5.7 Final Analysis

All of the vendors in the VRM space should bow down to Cisco and give them due respect. Cisco established this product category. They released one of the first, if not **the** first, VRM product. As an industry giant, they certainly accelerated the acceptance of Server Load Balancing as a product category, which has led to the continued and accelerated growth we are seeing today. And, most importantly for the vendors who should pay tribute to Cisco, they stopped investing in the technology.

We think of Cisco as a huge technology marketing/technology machine, and they are. But ironically, they have far less engineering assigned to their VRM products than any other vendor in this space, even the smallest. There are a handful of engineers working on maintaining Local Director and one maintaining Distributed Director, *part time*!

If it's support you are after, you might be a little better off. Some percentage of Cisco field people understand the issues of VRM and the LD/DD product capabilities. But they are diluted into a sea of people who have a huge product line to support. If you find someone who really understands the issues and has the time to help you, grab him or her fast. There used to be a dedicated set of people in the field focused on this area. Now there are not. That means less support for training, new product roll-outs, etc. Contrast that to the people who come from a company that eats, breathes, and drinks VRM 24 hours a day and you'll see a huge difference in capability.

If you have a nuts and bolts VRM application, which doesn't push the state-of-the-art feature-wise or performance-wise, *and you think it remain that way*, then using Cisco Local Director is a viable option. This is especially true of the application requires switches and routers and other stuff that is core Cisco business so that you can go to one place for shopping and support. Otherwise, buy the switches and routers from Cisco (that's almost always a good decision) but select someone else for your VRM solution.

3.6 Coyote Point (www.coyotepoint.com)

3.6.1 Company Overview

Coyote Point Systems was founded by leading industry consultants in New York and the San Francisco Bay Area in 1995. In 1996, Coyote Point launched the **Equalizer**™ intelligent load-balancing router -- among the first to transcend the limited Round-Robin DNS based methods, which randomly distributes incoming traffic to available servers.

Some marquee customers include: USWeb, Netscape and amazon.com.

3.6.2 Product & Support Overview

The Coyote Point product set is comprised of three different models of the Equalizer line. All of the products are based on PCI bus, Intel processor technology running an enhanced version of freeBSD.

The original offering, the Equalizer 150 is being phased out due to its limited capabilities. The 150 hit the market to introduce Coyote Point as a viable product provider. Since then, the company has rapidly expanded the Equalizers feature set and capabilities.

Equalizer E250

The E250 is the current entry-level product. It offers two 10/100 Mb Ethernet ports. The E250 is powered by a 300MHz Pentium and comes standard with 32 MB RAM. E250 supports a maximum of 64 Clusters and eight servers per cluster. All Coyote Point features are supported by this platform, which has a list price of $3,999.

Equalizer E350

The E350 provides two 10/100 Mb Ethernet ports, runs a 366 MHz Pentium processor and comes standard with 32 MB RAM. E350 supports an unlimited number of Clusters and 16 servers per cluster. All Coyote Point features are supported by this platform, which has a list price of $9,999.

Equalizer E450

The E450 is Coyote Point's latest product offering. It is designed to support the needs of larger, more performance-needy clients. It provides the standard feature set and two 10/100 Mb Ethernet ports. The 450 is powered by a 400 MHz Pentium processor and comes standard with 64 MB RAM. E340 supports an unlimited number of Clusters and 64 servers per cluster, and has a list price of $14,999

Envoy

These Coyote Point local load balancing products have recently been augmented by the release of the Envoy software for WAN support. This software package runs on top of any of the Equalizer platforms and acts as the authoritative DNS server for the chosen domain. List price is $2,500 per site.

Support

Service plans include toll-free 24x7 technical help line support, onsite service and priority hardware replacement. Coyote Point engineers are also standing by to add custom applications or protocols to their offering at the client's request. This service is handled on a case-by-case basis.

3.6.3 Product Functionality

The Coyote Point Equalizer uses HNAT to forward requests to the target servers locally and DNS-Re-Direction to forward between sites.

Other Important Functions

Coyote Point will provide (without charge) a lightweight server agent upon client request. This will monitor CPU and Memory utilization. The agent can be modified/enhanced by the end user to collect more sophisticated information. Basically, Coyote Point will provide script writing support on a case-by-case basis or will direct their client's to an outside resource. If the client is capable of writing their own scripts, they can collect anything they want. As of 3/99, Coyote Point has no plans to formalize this agent distribution or support program. However, this second-order agent capability proves that the company understands the value of direct server feedback.

Because they don't do delayed binding, Equalizers use sticky timers to support persistent sessions such as SSL sessions.

3.6.4 Product Features at-a Glance

FEATURE	CAPABILITY	COMMENTS
TYPE SOLUTION	PC Platform/ SOAS	Intel Pentium 300-400 MHz, 32 to 64 MB RAM standard
OPERATING SYSTEM	free/BSD	
DELAYED BINDING	NO	
PERFORMANCE	Unpublished	Two 10/100 MB Ethernet ports standard on all products
ROUTER OR BRIDGE	Router	
REDUNDANCY	Hot stand-By	Plug two Equalizers into a hub/switch. Backup waits for a lack of TCP/IP traffic in given interval, tries to diagnose problem with it's sibling LB, assumes identity of primary Equalizers IP address.
AGENT TECHNOLOGY?	YES	They offer a free UNIX script to monitor CPU utilization, memory, etc. The customer can later write any script to gather any information they want.
MECHANISMS	• Half –NAT • DNS Redirection	
POLICIES	• Number of users assigned to server • Least connections • Round Robin • Response Time • Server Agent Priority	User selectable weight scheme based on agent-collected data
FEEDBACK	• TCP open connections • Pings • Passive TCP connection monitoring • Active TCP port probing • Agent Responses	

Management Features

The administrative interface for the Equalizers is a browser-based solution. This allows not only configuration information to be set up from any browser-enabled workstation, but provides a compilation of statistics from the Equalizers. These can be used for trending and site analysis. Coyote Point's web application comes with some charts and graphs to help visualize what the raw data means.

3.6.5 Key Relationships

Coyote Point has two primary technology partnerships. These are designed to augment the current product capabilities and address some site-architecture issues.

Their key relationships lie with Bright Tiger and Webspective (formally Atreve). The Bright Tiger alliance allows CP to provide content management for the entire site. The partnership with Webspective addresses the need for complete site-and-content management. Coyote Point is used to achieve fine-grained load balancing while the other partner technologies supply the overview of site activity.

3.6.6 Issues

- Lack of MAC Address Translation support limits product application (no firewall, cache LB). However, development of this feature is currently under consideration at Coyote Point.
- No delayed binding means no SSL Session ID Tracking, no URL-based scheduling, no cookie-based scheduling, etc.
- No bandwidth management.
- **No Active Content Verification or Dynamic Application Verification.** However, the server-agents can be tweaked to provide these features.
- Failover time seems very long, and algorithmically strange**.** We'd recommend testing the failover capability during any product evaluation.

3.6.7 Final Analysis

Coyote Point is focused in offering a straightforward and reasonably priced product set which is easy to install and use. As such, it could be considered for many basic applications.

3.7 F5 Networks (www.f5.com)

3.7.1 Company Overview

F5 Labs is a growing company of about 150 employees based in Seattle WA. They have gained many key wins, with over 350 customers and a strong presence in the ISP space, among others. Some key customers are USA Today, Egghead Software, Lycos, UUnet, and Tower Records.

3.7.2 Product & Support Overview

Local Load Balancing: Big/IP

BIG/ip is a Software-On-A-Stick product that uses an Intel Pentium processor running BSD UNIX. Most of the packet processing functions are performed in the operating system kernel, which increases performance and provides some security advantages.

Big/IP comes in several forms:

- BIG/ip LB, a single, stand-alone box using a 300 MHz Pentium II processor. Shipped with 128MB RAM. Presently priced at $9,990
- BIG/ip, a single, standalone box using a 300MHz Pentium II processor. Shipped with 128MB RAM. Priced at $19,990. Pricing includes one-day installation and integration and a one year 7x12 service contract. BIG/ip HA sold as a redundant system (two boxes). The same hardware as the BIG/ip LB, priced at $35,990. Pricing includes one-day installation and integration and a one year 7x12 service contract
- BIG/ip Pro sold as a BIG/ip HA bundled with see/IT network management software. Priced at $40,990
- BIG/ip+, a single, standalone box using a 450MHz Pentium II processor, shipped with 256MB RAM. Priced at $24,990. Pricing includes one-day installation and integration and a one year 7x24 service contract. BIG/ip HA+ is also sold as a redundant system (two boxes). The hardware is a 450 MHz Pentium II processor, shipped with 256 MB RAM. Presently priced at $44,990. Pricing includes one-day installation and integration and a one year 7x24 service plan
- BIG/ip Pro+ sold as a BIG/ip HA+ bundled with see/IT network management software. Priced at $49,990

All units have a 4GB hard drive, which is used for the OS and to store long term statistics. All units are upgrade-able to 1 GB RAM, and ship with two 10/100 Ethernet NICs. Other interface cards supported include Fast Etherchannel, FDDI, and Gigabit Ethernet. BIG/ip can support up to eight 10/100 Ethernet ports or four Gigabit Ethernet or FDDI ports in any combination.

Multi-Site Load Balancing: 3DNS

3DNS is a separate device that performs multi-site traffic direction via the DNS Re-Direction mechanism. 3DNS can run stand-alone, but using BIG/ip as an agent to report on site health and load to the 3DNS enhances the system. 3DNS is presently priced at $27,490, and includes two days installation and integration and a one year 7x24 service plan.

Monitoring and Planning: see/IT

see/IT is a browser-based Network Management Console. see/IT provides real-time charting and graphing of key site and server statistics. see/IT also includes reporting, trending, correlation,

forecasting, and analysis tools to aid in long term troubleshooting and capacity planning. The software is presently priced at $9,990.

Content Management: global/SITE

global/SITE is a recently announced product expected to ship in late 2Q99. global/SITE manages the distribution of web server content and applications. Based on the same hardware as BIG/ip HA, global/SITE comes in two forms; global/SITE LAN, presently priced at $19,990, stages, publishes, and replicates web server content at a single site. global/SITE WAN, presently priced at $49,990, provides the same functions for multiple sites. Content and application files are encrypted when sent across the WAN.

Services

- On-site installation and integration support for $2,000/day
- BIG/ip LB 7x12 annual service for $2,000
- BIG/ip HA and HA+ 7x12 annual service for $6,000
- BIG/ip HA and HA+ 7x24 annual service for $8,000
- BIG/ip LB or 3/DNS 7x24 annual service for $3,000
- 1-5 days of various forms of installation and management training range from $995 to $2,595

Note that in many cases, one or more of the above services are bundled into the product pricing.

3.7.3 Product Functionality

Local Load Balancing

BIG/ip performs TCP Diddling (with half-NAT) for TCP sessions. BIG/ip supports a rich set of features, including routing, bandwidth management, packet filtering and firewalling, and delayed binding functions.

BIG/ip uses a variety of feedback methods to determine server and site health and load. All of these methods are of the external-probe type, and are executed on the BIG/ip box itself. The methods include observing traffic on the BIG/ip network interfaces (Layer 2 test), Device ICMP Ping, TCP Connection Verification, Active Content Verification and DAV.

F5s DAV implementation is particularly robust. BIG/ip can emulate the process the end user experiences at an Internet site. Some examples:

- Log onto multiple accounts; place items into a shopping cart just a customer would, complete the on-line financial transaction
- Directory & Authentication - BIG/ip allows users to load balance multiple directory and/or authentication services (LDAP, Radius, and NDS). DAV extends BIG/ip's capabilities by allowing administrators to verify that these services are providing the correct information
- Portals/Search Engines - Using DAV, BIG/ip allows administrators to perform key-word searches. If response is inadequate (a missing advertising banner for example), BIG/ip directs end users to properly working servers while administrators take down the problem server for repair without affecting end user access
- Legacy Systems - BIG/ip controllers can load balance services to multiple interactive services (i.e. Telnet, TN3270, TN5520). DAV verifies legacy host connectivity such as login screens are functioning properly

- Gateways - BIG/ip lets you load balance gateways (SAA, SNA, etc.) - BIG/ip's EAV feature verifies that the services provided by gateways are available. If not, DAV marks the problem gateway down for correction and directs end users to a functioning gateway
- E-mail (POP, IMAP, SendMail) - BIG/ip balances traffic across a large number of mail servers - Using DAV, BIG/ip verifies that those mail servers are accepting connections and responding properly

F5 also offers an API (BIG/api, with an SDK on its way) that allows one to build applications so that 3rd party instrumentation, such as BMC Knowledge Modules, can provide feedback on server or application health and load to BIG/ip.

When BIG/ip detects that a particular service on a server has failed, it will re-direct traffic only for that service, leaving the operation of other applications on the server unaffected.

Local load balancing can be performed by weighted round robin, least connections, fastest response time, "observed" response time or "predictive." "Observed" is a combined metric composed of measured response times and # open connections. "Predictive" is a metric based on trended server performance. Predictive is interesting because it provides a way to essential tune server weightings by assessing their ability to service requests over time.

BIG/ip is also capable of load balancing transparent devices such as firewalls, caches and transparent proxy servers.

BIG/ip also offers some useful bandwidth management functions. These can be used at hosting sites to hard-limit how much bandwidth is provided for access to each virtual service. Even more powerfully, as the level of traffic approaches those limits, policies can be created which control actions to "slow down" specified users, via queuing and rate shaping, to ensure that adequate bandwidth is available for other users.

BIG/ip can also be used as a firewall, providing packet filtering, SYN attack protection, source route tracing, illegal access logging, and other critical security functions.

BIG/ip has built-in redundant power supplies and a fast fail-over redundant architecture option.

Multi-Site Load Balancing

3DNS performs DNS Re-Direction. 3 DNS can operate stand-alone. When it does, it uses Active Content Verification to verify the health of each site. When used in conjunction with BIG/ip at each site, some additional feedback and policies are enabled. These include the ability for BIG/ip to measure round trip delays, site load, path packet drop rate, and other useful metrics.

3DNS supports the following load-balancing modes:

- Round Robin, Ratio, Least Connections mode, Random
- User-defined Quality-of-Service
- Topological, Round Trip Time
- Hit Rate, BIG/ip Packet Rate

3.7.4 Product Features at-a Glance

FEATURE	CAPABILITY	COMMENTS
TYPE SOLUTION	Software-on-a-Stick	
OPERATING SYSTEM	Embedded BSDI	
DELAYED BINDING	YES	SSL Session ID. Does not currently support URL-based scheduling or cookie-based scheduling.
PERFORMANCE	90 Mbps (LB and HA) 170 Mbps (HA +)	Performance data as provided by the vendor.
ROUTER OR BRIDGE	Router	RIP, OSPF, BGP
REDUNDANCY	Hot standby	Redundant power supplies. RS-232 keep-alive cable between boxes.
AGENT TECHNOLOGY?	NO	Market heavily *against* the use of agents.
MECHANISMS	• TCP Diddling w/ Half NAT (LAN) • MAT for transparent device load balancing. • DNS Re-Direction (Multi-site)	
POLICIES	**Local** • Weighted Round Robin • Least Connections • Response Time • "Observed" • "Predictive" • SSL session ID **Multi-Site** • Random • Round Robin • Least Users • Site Load Feedback • Round trip delay • Dropped packets • Combinations	
FEEDBACK	• Network packet rate • TCP Connections (to port) • Active Content Verification • Dynamic Application Verification • API for 3rd party instrumentation integration	

Network Management Features

Administration can be performed locally (TTY) or remotely. The remote options are via a command line interface using SSH for privacy or via a browser graphical interface, using SSL.

Key Additional Features

BIG/ip

- Packet filtering and firewalling
- Fast Etherchannel and Gigabit Ethernet network interface options
- Bi-Directional NAT (not just for VRM, but to support enterprises with non-routable internal addressing by managing a small pool of routable addresses for external communications)

3/DNS

- Can be used with or without BIG/ip (BIG/ip enables site load feedback, round trip delay, and dropped packets policies)
- Redundant schedulers

3.7.5 Key Relationships

F5 Networks has strong relationships with the major ISP/co-locator players such as PSINet, DIGEX and MCI Worldcom that build F5 products into "their plumbing," as well as Exodus, and Frontier Global Center. The Exodus and Frontier Global Center sales forces pro-actively sell value add high availability services to their customers, which are predominantly implemented using F5 products. In addition, F5 has recently signed distribution agreements with NTT to represent them in Japan and Singapore Technology to represent them in Singapore, Malaysia, Hong Kong, China and Thailand. F5 also has OEM agreements in place with Cabletron and Packet Engines/Alcatel.

3.7.6 Issues

The F5 products have few major feature deficits that would prevent F5 Labs from being a candidate for almost any application. Some issues and concerns include:

- The limitations of standard Intel hardware platforms and BSD Unix. F5 has done a good job of optimizing in this environment as much as possible, by pushing packet processing functions into the driver/kernel level and by churning out new hardware platforms at almost the same pace Intel provides new processor capabilities. Ironically, all the work to drive functionality into the kernel probably prevents them from multi-threading their code and leveraging SMP platforms

  F5 vociferously defends their approach. We hope that's just good defensive marketing and that they are not blind to the mega-trends in WAN bandwidth availability, peak site usage, and other factors which are resulting in rates of demand growth far outstripping the pace of processor enhancement.

  The solution for F5 is fairly simple. Just provide a capability, as IPivot has, for multiple boxes to support a common virtual service, with load sharing between the boxes. That way, a single box can support the 80-90% of the applications for which performance is adequate, and multiple boxes can be deployed for those applications that require it. This would also make

us a little more comfortable with the cost of the second box, presently now only used as a warm spare

- **Lack of support for URL-based scheduling.** This isn't to say F5 Labs doesn't support delayed binding – they do (e.g. they support SSL 3.0 Session ID Tracking). But they don't support URL-based scheduling or cookie-based scheduling. Thus you can't use F5 for the applications we put into the category of "Server Optimization"
- The products are somewhat pricey compared to alternatives on the market
- The boxes use space inefficiently. This is really only an issue of you are co-locating and are paying for rack space. It's also a big issue to the hosters and co-locators themselves
- 3DNS has a rich set of options now, but we believe two enhancements should be considered.

  1. Add the use of either Static Client Preferences or BGP-4 metrics as policies to minimize the number of Autonomous Systems traversed as a policy. This is a simple way to keep European traffic primarily in Europe, for instance.
  2. Provide a "fail-safe" mechanism, such as HTTP Re-Direct or encapsulation, at the site level (perhaps as part of BIG/ip) to re-direct traffic that the DNS Re-Direction scheme has improperly directed. This would allow the DNS-based system to return intermediate-to-large TTLs when resolving requests, which could prevent overload on the DNS system (see the VRM Technology Report for a detailed discussion of these issues)

3.7.7 Final Analysis

150 people totally focused on building and supporting state-of-the-art Virtual Resource Management products. Savvy, quality people throughout the organization. Lots of interactions with customers, which drives feature priorities, which are nimbly delivered with value-add nuances (instead of just feature checklist minimal capability). Leverage of a standard hardware platform and operating system to facilitate rapid software-based feature development.

This seems familiar to us. It reminds us of another company we worked with at a similar stage of development - Cisco. The strategy is to identify and solve real world problems before anyone else and over time create such a barrier-to-entry due to software sophistication that the competition just melts away. We're not saying that the VRM market is anywhere near the size of the router market, or that F5 Labs will dominate the market like Cisco did. It's just that the energy, focus, capability, and strategy feel familiar.

It's working. F5 Labs is either #2 in the market or awful close to #2. The only vendor they lag significantly is Cisco, due to the marketing might of that gorilla. However, F5 Labs and others have been eating significantly into Cisco market share by having more featured products and by providing more focused and value-added support.

Product-wise, other than Server Optimization scenarios, there isn't a server-oriented application that F5 Labs isn't suited for. We like the extensive health and load feedback features and the extensive policy choices. BIG/ip is especially suited for hosting and co-location sites, due to the ability to support large numbers of virtual services and content servers, and the ability to manage bandwidth to allocate just the portion of the site resources the customer contracted for. Also, for these applications, it's pretty easy to carve the site up so that you can use multiple boxes to avoid any performance issues you might see with a single box.

3DNS is also a solid product for multi-site applications with some value-add features, such as providing feedback on packet-loss that can have a real impact on system operation.

The management applications (see/IT) sound good on paper (we've only used the configuration tool so far, not the monitoring or planning tools) and F5's investment in them is yet another indication of their sensitivity to customer needs. Providing more insight into present operation and helping guide future architecture planning with simple, easy-to-use tools, are two of the most pressing needs we see in the end user space toady.

F5 Labs has been very successful over the past couple of years and are growing nicely. Given the quality of the software, and the quality of the people, we expect F5 Labs to be around for a long time.

3.8 Foundry Networks (www.foundrynet.com)

3.8.1 Company Overview

Foundry Networks was founded in 1996 by a number of ex-Centillion/Bay Networks executives and engineers. The company's charter is to build high performance, next generation Gigabit Ethernet switches and routers. The company has received $33 million in three rounds of funding from venture funds and private investors, including Accel Partners, Crosspoint Venture Partners, Dixon Doll, Institutional Venture Partners, Mitsui & Co., Ltd. and Vantage Point. They have approximately 120 employees and are headquartered in Silicon Valley.

In April 1998, Foundry released the ServerIron VRM platform that offers both server load balancing and cache server redirection capabilities.

Some of Foundry's marquee VRM customers include AOL, Yahoo!, Mindspring, AT&T Worldnet and Sprint Global One.

3.8.2 Product & Support Overview

Foundry Networks began shipping their first ServerIron product in April 1998, which was an 8-port local load balancing 10/100 Ethernet switch with Gigabit uplink options. Since that time, they have announced and shipped additional software releases, providing increased functionality, and have expanded the product line to include higher port density units. A May 1999 recent product announcement outlines Foundry's chassis-based higher performance and higher port density solution.

The ServerIron product line includes an 8-port, 16-port or 24-port 10/100 Fast Ethernet switch with optional gigabit Ethernet uplinks and server load balancing and transparent caching software in a stackable form-factor. All of the ServerIron's software features run on a single PowerPC CPU that accesses between 32MB and 128MB of RAM, depending on the ServerIron configuration purchased. The 8-port ServerIron is U.S. list priced at $6,295, the 16-port 10/100 Mbps port version is $9,995, and the 24-port version is $18,995. Uplink options for the 8, 16 or 24-port stackable version include a 1-port Fast Ethernet expansion module at $1,995 and 2-port Gigabit Ethernet expansion module at $3,695. Software features include local and remote server load balancing and transparent cache switching.

Foundry also ships the BigIron 4000 and BigIron 8000 chassis-based, modular switches that feature high port density switching and wire-speed routing. VRM functions will be supported on these platforms later in 1999.

Support

Service plans include:

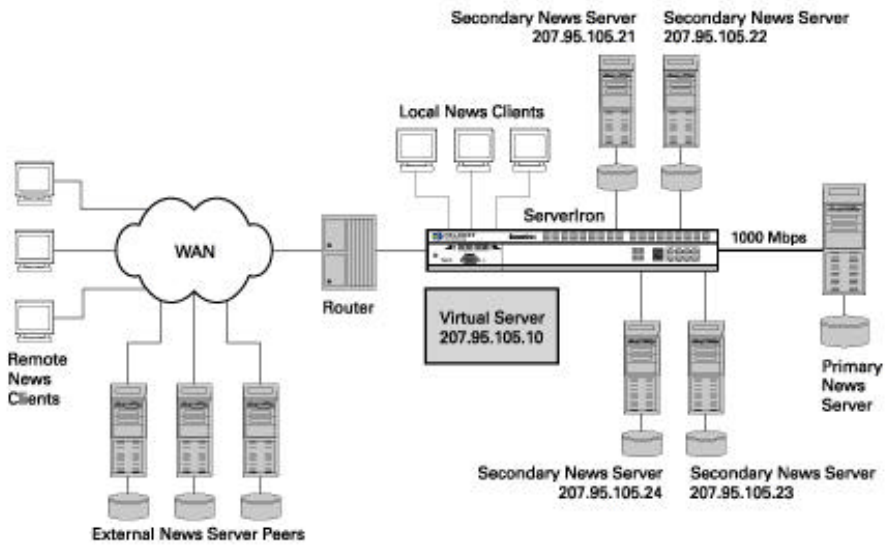
- Toll free phone support and web access to system software updates for 90 days from date of shipment
- Advanced hardware replacement within the first 30 days
- Hardware repair (after the first 30 days) for 1 year (3 day turn-around)
- There are options to extend the support plans for 12 months and obtain 24x7 support toll-free 24x7

3.8.3 Product Functionality

The ServerIron supports two primary applications, regardless of the physical port configuration of the switch: Local & remote load balancing of UDP/TCP and HTTP applications, and transparent cache switching of HTTP traffic to a web cache server farm.

Local Load Balancing Functions

Like most load balancing devices, the ServerIron attaches in-line in the topology between the backbone network and the servers. Unless port density is an issue, it is likely that the servers will be directly attached to the ServerIron.



One or more virtual servers are configured in the ServerIron and are associated with a group of servers attached (typically) to the switch. The ServerIron uses Active Content Verification (ACV), TCP Connection Observation (TCO) and ICMP pings to verify the status of the servers/applications that are part of a server farm. In addition, ServerIron performs application level health checks for HTTP, SMTP, DNS and Radius. Server load is measured by number of clients connected (by IP address) and by the number of open connections (by TCP connection).

The ServerIron primarily supports Half-NAT as its primary local SLB mechanism. The ServerIron can also be configured for direct-path return as an option. In direct-path return, traffic for the VIP is Destination NAT translated to an available server, but the return traffic is sent directly to the next hop router, bypassing the ServerIron. This allows for high-bandwidth return traffic to flow directly to the clients. Direct-path return requires configuring an IP address on the loopback interface of each physical server, which is also the same as the Virtual Server's VIP.

A recent announcement by Foundry on the support of URL-based load balancing allows for load balancing based on the URL information being requested by the clients. Direct-path return is not supported simultaneously with URL-based scheduling.

For HTTP and other types of sessions that require processing on the same server, the ServerIron supports an attribute called "sticky" whereby all TCP connections from a given client, of a

configured type of TCP Port Type (e.g. SSL), are all sent to a single server. This is typically used in Secure Socket Layer type of applications such as credit card forms or shopping cart forms where a user's transactions must be managed by a single server (Vs having multiple connections load balanced across identically configured servers).

The ServerIron switch supports both the Hot-Standby and the Active-Active modes of redundancy, which allows significant deployment flexibility. If shared media hubs or switches are used to connect the backbone side topology, user-to-VIP session state can be maintained in both switches because the switches can share the same MAC address in response to ARPs for the configured Virtual IP addresses.

Multi-Site Load Balancing

To support server farms that may reside on different networks other than that of the ServerIron switch, the ServerIron can be configured for Source NAT. Traffic sent to these remote (off-local network) servers return their traffic back to the ServerIron for proper Source NAT re-translation to prevent the user from discovering the true IP address of the server farm. This provides for the following:

- Expand SLB to geographically dispersed data centers and servers
- Backup Server Farms when the local server farm becomes unavailable

The ServerIron can also be configured to send back to the client an HTTP redirect message that enables the client to access the server directly. This may be desirable when the path to a remote server is more efficient than sending the traffic back to the ServerIron and ultimately to the requesting client.

Transparent Cache Switching

Soon after releasing their local server load balancing solution, Foundry began shipping a software upgrade for the ServerIron to allow transparent cache switching. A ServerIron is placed in the data-path of users that are generating HTTP requests to the Internet, and the switch intercepts and rewrites the requests using MAT. These requests are directed at a group of cache servers; server selection is typically done through a statistical selection method based on IP destination address. Foundry’s announcement of URL parsing implies that they will be able to perform cache switching based on URLs or hostnames.

3.8.4 Product Features at-a Glance

FEATURE	CAPABILITY	COMMENTS
TYPE SOLUTION	Switch with Internet IronWare (SLB and TCS Software)	PowerPC running at 240 MHz with 32MB RAM. Plus per port ASICs performing L2 functions for high speed switching.
OPERATING SYSTEM	Embedded	
DELAYED BINDING	YES	
PERFORMANCE	8,000 connections per second (immediate binding) 800 Mb/s	PC-Week Labs test data
ROUTER OR BRIDGE	Bridge	
REDUNDANCY	Hot Standby Active-Active standby	Keep-alives and session information is shared across the local network (no RS-232 cable). Session state can be maintained between the switches configured as Active/Active or Active/Standby pairs using hubs.
AGENT TECHNOLOGY	NO	
MECHANISMS	• HTTP Re-direction • Half-NAT • Full-NAT • MAT	MAT for both local LB (out of path return) and cache switching.
POLICIES	• # of clients (by IPSA) • Least connections • Round Robin • Weighting • Hashing (cache switching) • Statistical (cache switching)	
FEEDBACK	• ICMP pings • TCO • Active Content Verification	• DNS health check • SMTP health check • RADIUS health check

Management Features

The ServerIron can be managed through a Cisco-like CLI, a web-based interface and a standalone management software package called IronView, which runs on Windows NT and under the HP OpenView for the most popular UNIX implementations.

Key Additional Features

- SYN-Attack protection through connection thresholds
- Flexible IP packet filtering by L2/L3 addresses, protocols, and applications
- Cross-subnet directing (Real Servers can be on any subnet)
- Radius Authentication for Administrator logins

3.8.5 Key Relationships

All of Foundry's ServerIron-related relationships are within the caching market. They have partnerships with all of the major cache vendors, including CacheFlow, Cobalt Networks, Inktomi, Network Appliance and Novell.

3.8.6 Issues

- **Foundry needs to support real multi-site VRM capabilities. Their Full-NAT approach requires traffic to pass over WAN links multiple times and does not protect against site failure**
- Foundry does not presently support Dynamic Application Verification, but they have told us it is planned
- SYN-Attack protection is based on limiting the number of open connections to a VIP, which is OK for server protection, but does not discriminate between valid requests and SYN attacks from an IP host. If the SYN threshold were crossed, every request from that host would be dropped
- No supplemental features to support VRM, such as bandwidth management or routing available on the current ServerIron platform. Foundry will offer integrated switching, routing and VRM features in their upcoming high density, chassis-based platform
- Foundry's products are essentially a SOAS system embedded in a switch. They have the advantage of providing L2 forwarding with integrated switching ASICs, but all VRM functions take place on a central CPU. Because they have built their own switch platform to deliver VRM solutions, they will have a longer platform update cycle than the real SOAS vendors will. Thus, while the SOAS vendors can improve their platform performance and capacity by simply certifying another PC, Foundry will have to do a manufacturing respin.

3.8.7 Final Analysis

Foundry has fixed a few of the issues we noted in the last VRM document; they now support better forms of local health checking (ACV), they support more flexible local load balancing with Full NAT and Active-Active redundancy configurations.

The switch provides good SLB performance at an excellent price; their architecture is based on software running on a single IBM PowerPC CPU embedded within the switch hardware. Similar to the other load balancers with central CPU's, their overall performance will be affected by what functions are active and how much traffic is passing through the system.

Foundry's primary uniqueness in this product space is their ability to deliver multiple types of switch products, in addition to their Server Load Balancing and Transparent Caching functionality on the ServerIron. If you need high density Gigabit Ethernet switching or routing products **and** local Server Load Balancing, you should consider Foundry.

Their recent Internet IronWare V4.0 release now supports multi-site load balancing and URL parsing, so it will be worth keeping track of them to see what capabilities they provide in those features.

3.9 HolonTech (www.holontech.com)

3.9.1 Company Overview

HolonTech is a start-up company in San Jose, CA. that was formed out an NEC R&D organization. HolonTech still performs some contract work for NEC, but their focus is on developing, marketing, and supporting a VRM product line.

3.9.2 Product & Support Overview

HyperFlow 2 is the HolonTech product. It is a compact multi-port hardware-assisted appliance that can be purchased in two forms; an 8-port device which lists at $17,995, and a 16 port device which lists at $24,995.

3.9.3 Product Functionality

The HyperFlow 2 performs MAC Address Translation (MAT) to re-direct incoming traffic to targeted content servers. A loopback interface is configured on the Real Servers for each V_IP configured in the HF-2. Multiple Real Servers (supporting different virtual services) can be deployed off of each HyperFlow port by using Ethernet hubs or switches as port replicators. This makes HyperFlow 2 cost effective on a per-server basis.

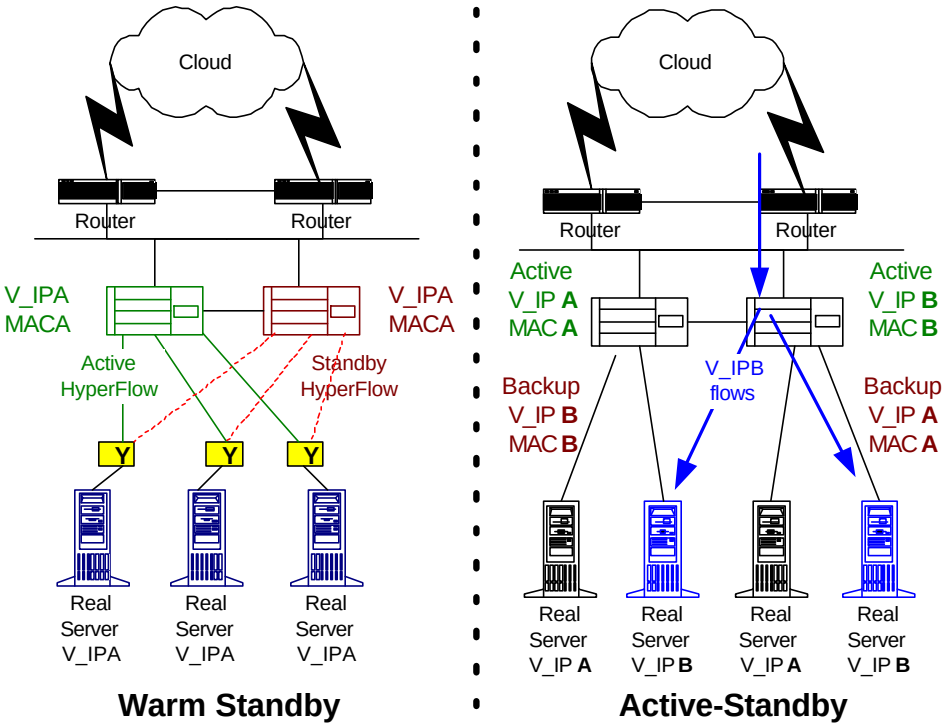
The health and load of the real servers can be evaluated by TCP Connection Verification and by Active Content Verification. The load on the network connection to the Real Servers is also monitored to make sure that network overload conditions are avoided. In addition, optional agents for NT and Solaris Real Servers can be used to provide health status information as well as CPU load and other system resource utilization information. A scripting interface allows Dynamic Application Verification-like tests and rules.

Incoming traffic can be parsed by TCP Destination address. Different target Real Server groups can be assigned by TCP Port. For instance, this allows an FTP service, HTTP service and an e-mail service all to be supported by the same V_IP.

Another application for this capability is to direct HTTP traffic to transparent cache servers. When an HTTP (TCP port 80) request is received, the request can be dispatched to a locally attached cache server.

HyperFlow 2 supports two types of failover topologies, Warm Standby and Active-Standby. In the Warm Standby case, a small, passive Y-cable connector is used to create a triangulated connection between the Real servers and each HyperFlow. The Hyperflows are connected to one another by both an RS-232 cable and a LAN connection (e.g. Fast Ethernet). If the passive HyperFlow is told of a link failure by the active, or if the passive detects that the active has completely failed, the passive HyperFlow will turn its links on.

In the Active –Standby case, each HyperFlow will be active for some subset of the V_IPs and will act as a backup for the other subset. Usually, it makes sense to spread the Real Servers associated with each V_IP between the active Hyperflows. When everything is healthy, as traffic comes in, it will be sent either to servers directly connected to the responsible scheduler (the one that "owns" the V_IP) or to servers connected to the other HyperFlow (via the LAN link between boxes) in order to spread the load. If a HyperFlow should fail, the remaining active HyperFlow will take on the responsibility for the V_IPs previously scheduled by the failed HyperFlow. In this state, only the servers connected to the remaining healthy and active HyperFlow will be used, so some overall site throughput is lost, but service remains available.



3.9.4 Product Features at-a Glance

FEATURE	CAPABILITY	COMMENTS
TYPE SOLUTION	Multi-port appliance	
OPERATING SYSTEM	Embedded real-time OS	
DELAYED BINDING	NO	
PERFORMANCE	6000 connections/sec (8-port) 12000 connections/sec (12-port)	Performance data supplied by the vendor.
ROUTER OR BRIDGE	BRIDGE	But does not support Spanning Tree. Cannot be put into an arbitrary mesh topology with other bridges.
REDUNDANCY	Warm Standby Active-Standby	Uses a special Y-cable to connect Real Servers to both the active and passive devices Both schedulers are active, for all V_IPs. They act as back-ups for one another in the event of a failure.
AGENT TECHNOLOGY?	YES, optional	Solaris and NT. Provide feedback on CPU and available memory
MECHANISMS	• MAC Address Translation	
POLICIES	• Weighted Load (server, network interface) • Response Time (HTTP) • Least Connections • Round Robin	Back-up servers
FEEDBACK	• Agent feedback (see above) • TCP and HTTP connection response times • Active Content Verification • Dynamic Application Verification	Scripts are used to define DAV tests and rules.

Management Functions

HolonTech provides both a CLI and Web browser management for both monitoring and control. Provides a graphical representation of statistical data.

Key Additional Features

- Server access can be blocked for a source IP address on a specific port or system-wide
- Redundant power supplies
- *Service* failover. If a particular service on a Real Server fails, but others remain operational, requests for that service only are re-scheduled

3.9.5 Key Relationships

Given their heritage, HolonTech has been able to establish some strong channel relationships in Japan. Also, they have some kind of marketing relationship with Oracle, which we don't believe has been fully defined or exploited yet.

3.9.6 Issues

- **HolonTech presently has no solution for multi-site applications. This is the single biggest area in which they lag the market**
- **The 2nd biggest area they lag the market is in delayed binding capability.** As a result they cannot perform SSL Session-ID persistence, URL-based scheduling, or cookie based persistence or preferential services
- **HyperFlow 2 has no general purpose API to allow integration with 3rd party server co-resident agents or other value-add functions**
- The HTTP cache-server redirection mode to is not state-of-the-art because the product cannot distinguish between static content requests and active application requests. Therefore a lot of .cgi and .asp functions and HTTP PUTs/POSTs will be sent to the cache servers, just to be re-directed back out to the ultimate proper destination

3.9.7 Final Analysis

HolonTech has come a long way in HyperFlow 2, fixing a lot of limitations we perceived in the original HyperFlow. Among the state-of-the-art capabilities we applaud in HyperFlow 2 are optional server co-resident agents, TCP port filtering to allow multiple services to be indexed to a single V_IP, active-active failover, Active Content Verification, and Extended Application Verification with script-oriented customization.

However, presuming we are pre-disposed to a multi-port device, we can get all the state-of-the-art features from Alteon or ArrowPoint, plus more, and at the same time get full fledged switching and routing – at a lower cost. So we're not sure at this point exactly where HolonTech fits in the overall scheme of things. But, given the tremendous progress they have made in the last 6-8 months, we think it's entirely possible that HolonTech will clearly establish their differentiation and positioning in the coming year.

3.10 HydraWeb (www.hydraweb.com)

3.10.1 Company Overview

HydraWeb is a company that was formed in 1996 to address the need for highly-available and reliable enterprises. From inception, HydraWeb has focused on solutions for e-commerce environments.

3.10.2 Product & Support Overview

Hydra 5000

This is their top-of-the-line box. It features from 4 to (a maximum) of 8 interfaces of 10/100 Ethernet, supports FDDI , standard WAN functionality and a built in modem for out-of-band management. Four server licenses are included with the cost of this unit. This means that out-of-the-box the Hydra 5000 is ready recognize and accept traffic from 4 servers. More server licenses can be purchased if you wish to use more servers. There is no maximum server limit when using this product. The 5000 (with spare unit) lists for $50,000.

Hydra 2000

This “workhorse load balancing appliance” falls into the mid-range category. It features from 2 to 8 (maximum) 10/100 Ethernet interfaces, FDDI support (however, only 80Mbps throughput with this interface), optional WAN functionality and a built in modem for out-of-band management. It also provides an option for dual power supplies and Hot Spare configuration. Four server licenses are included with the cost of this unit. This means that out-of-the-box the Hydra 2000 is ready recognize and accept traffic from 4 servers. More server licenses can be purchased if you wish to use more servers. There is no maximum server limit when using this product. The 2000 (with Spare unit) lists for $20,000.

Hydra 900

This low-end device is designed to meet the needs of smaller ISPs and departments that require support for online, mission critical systems. It comes with 2 10/100 Ethernet interfaces and hot spare configuration capabilities. There is no support for WAN services with this product. Four server licenses are included with the cost of this unit. This means that out-of-the-box the Hydra 900 is ready recognize and accept traffic from 2 servers. More server licenses can be purchased if you wish to use more servers. However, the server limit of this product is 4. The 900 (with Spare unit) lists for $10,000.

Hydra Commerce and Performance Agents

These agents are not required to use load balancing at the hands of a Hydra Unit (pick one from the above list) but they are highly recommended. These on-board agents provide the expected insight into server “health.” These agents can run on many UNIX OS (SunOs 4.x, Solaris 2.x, Digital Unix, Irix 5.x/6.x, BSDOS 2.x/3.x, Linux 1.3/2.x, HP-UX 9.x/10.x) and, of course, Windows NT.

Support

The HydraWeb service offerings fall into three categories. Standard Care = 24x7 email and phone support with four hour response time, overnight replacement for dead units, web based support services and software upgrades. Premium Care = same as Standard, with two hour response time.

3.10.3 Product Functionality

Local Load Balancing

TCP Diddling

HydraWEB products perform *delayed binding*. When the client makes a request to a content server, the Hydra receives the incoming request. It then responds for the server, so the client thinks it has a direct connection. When the client sends the URL (or content) request, the Hydra can inspect the packet in order to determine what content (location) is requested. The Hydra will then set up a separate connection with the most capable and available server of the moment and pass on the client's request. The fulfillment server sends the traffic directly back to the Hydra, where it is transferred to the client using the TCP parameters of the original connection. This process does not require any software on the server. It requires that the HydraWEB box be in the traffic path in both directions – client-to-server and server-to-client.

For an in-depth discussion of this technology, grab a copy of the Acuitive VRM Technology Report.

Multi-Site Load Balancing

DNS Re-direction

To provide continuity for a multiple geographic site web presence, the primary site Hydra can act as the Authoritative DNS server for the domain. By maintaining a table of active sites (with the help of the recommended Hydra Agents) and related V_IPs, the Authoritative Hydra will respond to DNS queries with the V_IP addresses of a healthy target site. This provides transparent request distribution across servers located around the world. This also acts as a WAN fail-over mechanism in the case of catastrophic site failure.

Other Important Functions

Redundancy

HydraWEB uses a serial cable and spare box for their redundancy scheme. This allows the back-up Hydra to have all state information copied to it from the primary. The spare is included in the price of every "primary" Load Balancing device.

Agents

HydraWEB offers the Performance and Commerce agents to round out the feedback capabilities of their product set. These agents can provide the following information: CPU/process usage, Load Average, Memory Utilization, Machine State (i.e. shutting down?), Service Latency (request/response turnaround time), Response State (whether the servers even responds at all). This information is passed to the HydraConsole (an in-band management application) via UDP. The algorithm fed by all this input is a patent-pending item called the PLA-Multiplexor Load Average Algorithm. This user can assign different priorities to various inputs (i.e. "Load Average") to modify the algorithm.

The Commerce Agents use RSA technology to obtain operations information from secure servers, process it locally and send the information to the HydraWeb Console via SSL encrypted channels. In this fashion, HydraWeb agents allow tight administrator security to be implemented in regards to changes in the enterprise server environment/SLB policy settings.

3.10.4 Product Features at-a Glance

FEATURE	CAPABILITY	COMMENTS
TYPE SOLUTION	PC Platform/SOAS	Intel Pentium 400 MHz, 128 MB RAM standard
OPERATING SYSTEM	BSD	
DELAYED BINDING	YES	SSL ID Load Balancing
PERFORMANCE	45-155 Mbps throughput (depending on product) Up to 3 million simultaneous connections (Hydra5000)	
ROUTER OR BRIDGE	ROUTER	
REDUNDANCY	Hot Standby	Special Serial Cable
AGENT TECHNOLOGY?	YES	Not required, but *highly recommended* (from HydraWEB site)
MECHANISMS	• DNS Redirection • TCP-diddling	
POLICIES	• WAN closest • Protocol priority w/ response time metrics • URL parsing/content specific • Predictive and observed server response time (socket level)	Backup and Overflow servers
FEEDBACK	• Socket connection requests • Response time monitoring • Standard and Custom Agent feedback • Active Content Verification	(See Agents heading under the Other Important Functions section above for details)

Management Features

All devices can be managed using the HydraConsole application. This is optional feature allows an administrator to log into any HydraWeb box, and then access any web server console to make adjustments, check management statistics. It also allows for email notifications, automatic pager messages and console/syslog messages whenever a threshold is exceeded.

Key Additional Features

- IP packet filtering
- Encrypted agent-HydraWeb unit communication feedback

3.10.5 Key Relationships

They currently feature Sybase as a strategic partner. This relationship, started in February of 1998, allows for joint marketing efforts and referrals, and Sybase recommends HydraWeb as their preferred SLB solution. We believe that HydraWeb and Sybase are early in the co-development of a Sybase-optimized agent that will provide feedback on things like database-query-queues.

Another strategic partner is Digital (Compaq). Digital's service arm directs clients with high-availability needs to HydraWeb. This relationship has been in place since October 1997.

3.10.6 Issues

- As a SOAS product, suffers from all the advantages and disadvantages of leveraging a PC platform that all other like products suffer/benefit from. Since HydraWeb does not support an Active-Active redundancy scheme (a'la IPivot), yet they support process-intensive delayed binding functions, their susceptibility to performance degradation is high
- **No support for Dynamic Application Verification**
- Serial connection required for the redundancy configuration prevents the standby unit from being physically distanced from the active unit. Also, persistent session failover is not supported
- Their multi-site features are not as full featured as many of the other vendors

3.10.7 Final Analysis

We frankly hesitate to either promote or trash HydraWeb. We just haven't had much experience with them. HydraWeb started out as a consulting company and continues to provide a lot of service to their customers, mostly Wall Street financial companies. As a result, HydraWeb's customers don't need much help from the likes of us. The products seem good and have stood up to the test of time and changing requirements among some pretty needy and bleeding edge customers. But HydraWeb's focus is pretty narrow and we're not sure how much support they offer to customers outside of their target geographical space, which seems to be bordered by Long Island Sound and the Hudson River.

3.11 IBM (www.software.ibm.com)

3.11.1 Company Overview

IBM is a big East Coast company. You may have heard of them. They do other stuff besides make VRM products.

IBM doesn't really focus purely on VRM or SLB; instead they offer complete e-commerce solutions. Their Web Server Performance Pack includes a lot of stuff beyond typical VRM features. Some examples are: Web Server capabilities, E-commerce applications, content management, QoS stuff, and proxy services. However, in the space below, we are going to *specifically* address their VRM software. (The Web Sever Performance Pack includes the SecureWay software described below). We would need another 20 pages to cover the range of options someone could put together from the IBM product quiver. They do maintain a nice web-site to help with the process.

Some marquee customers include Nagano Olympics site, IBM Corporate site, Masters Golf site, and the U.S Open/Wimbledon/French Open Tennis tournament sites.

3.11.2 Product & Support Overview

IBM's product offering is called SecureWay Network Dispatcher (the artist formally known as eNetwork Dispatcher; a.k.a. Interactive Network Dispatcher - IBM wanted to brand the product as part of their overall e-business software strategy.) This is a software solution that can be run on AIX, NT or Solaris platforms. The list for a SecureWay CD (both UNIX and NT software) with an unlimited server license is $7499.

SecureWay Dispatcher

The SWD product has two primary functions.

The first is the **Dispatcher** function. The Dispatcher is responsible for passing incoming requests to the appropriate server. The Dispatcher consists of three components:

- The executor
- The advisors
- The manager

These functions will be explained in the Product Functionality section.

The second element of the SecureWay product is the **Interactive Session Support** (**ISS**) server. ISS can be configured on content servers to provide resource-resident monitoring, or it can "front end" the Dispatcher function to provide multi-site VRM. In the multi-site mode, it acts basically as a replacement (or augmentation) of a traditional DNS server. Its job is to provide an IP address of a content server or Dispatch device (more on that later in the Product Functionality section)

Support

Support from IBM is highly customizable. It ranges from 7x24 phone support to on-site, product customization. The IBM web site contains tons of configuration and product issues information. Free code fixes, tweaking tips and newsgroups are also available on-line.

3.11.3 Product Functionality

Local Load Balancing

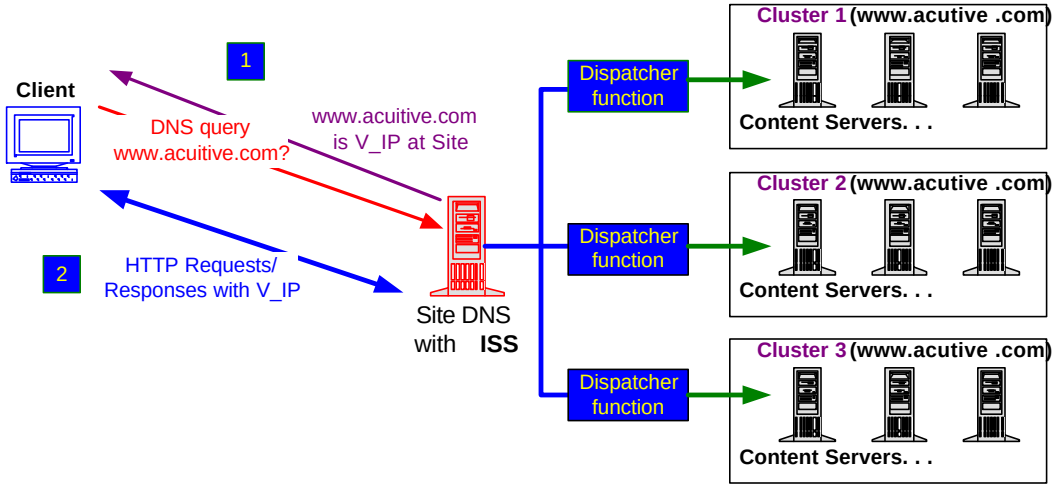
MAC Address Translation

In this mode of operation, the content servers are configured with the Virtual IP address on their loopback interface. The Executor process within each Dispatcher performs MAC Address Translation of the TCP SYN packet and all ensuing packets from the client, to send the packet to the intended server, which can then respond directly to the client. To avoid stale connections building up in the TCP table, the IBM solution times out bindings based on activity timers. Using this approach, you enable direct-return of traffic from the server-to-the-client. This is often preferred for high performance systems because there can be an order of magnitude more traffic in the server-to-client path than the client-to-server path.

Multi-Site Load Balancing

ISS (DNS) Redirection

The second scenario allows for high-availability at each site. The ISS process can front multiple Dispatchers to load balance the load-balancers. The ISS can be used (actually, *must* be used) in conjunction with a DNS server in this scenario. When the DNS query comes into the DNS server, the ISS software is on board and ready to make the call. When a DNS request comes in for an ISS domain, the ISS software will respond for the DNS process. The ISS software will then make its IP address response determination based on the internal health of each site that can service the domain. Once it makes a decision, the ISS process responds with the chosen IP address. This approach enables a "two-tiered" network configuration and provides transparent request distribution across servers located around the world. This also acts as a WAN fail-over mechanism in the case of catastrophic site failure. The approach is illustrated below.



The figure above illustrates the ability to use a single domain name across geographically disperse locations, or across multiple Dispatchers for a single site. This configuration allows for WAN re-direction and fail-over, thereby increasing the chances of a satisfactory client experience.

The client's request to resolve the IP address of acuitive.com is passed to the DNS/ISS server. The ISS server returns one of three V_IPs depending on which Dispatcher/site it decides the client should use. Again, requests can be directed in a round robin fashion, or based on a user-

definable server "weight" metric. Once at the appropriate site, the Dispatchers act as local load balancers, directing the request to the server that can best serve the client's need.

Other Important Functions

For the sake of clarity, we'll describe the overall functions/responsibility of each piece of the IBM software solution. Remember that these are all a part of the SecureWay product.

- The **executor** always runs and is responsible for recognizing new connection requests, forwarding connection requests to chosen servers, identifying packets which are associated with an existing connection, and keeps track of existing connection bindings. This is an OS kernel extension
- **Advisors** are application specific agents that query the servers to evaluate their availability and load. Advisors presently exist for HTTP, FTP, SSL, NNTP, Telnet, SMTP, and POP3. You can also write custom advisors. Advisors are optional, but their use does provide better fault detection and load balancing. These do not reside on each content server, but are resident on the same platform as the executor. They probe the servers remotely. These Advisors are optional, but **recommended**
- The **manager** sets weights for load balancing based on executor counters and/or feedback from advisors and/or feedback from system monitoring programs such as IBM's ISS (Interactive Session Support). The manager is also an optional chunk of software. However, according to IBM – "If the manager is not used, load balancing will be performed using weighted round-robin scheduling based on current server weights." Basically, you lose a lot of flexibility by not condensing and using (manager function) the detailed feedback to schedule incoming requests

Redundancy

IBM uses a heartbeat protocol on all interfaces of a SecureWay server. This proprietary method is more than a simple "keep alive" between a primary and secondary server. All state information for the site is shared by the primary with the secondary. If the primary tanks between heartbeat intervals (meaning there are a few new connections that the primary knows about, but the secondary doesn't), IBM claims that about 98% of existing connections will be saved when the secondary comes on-line.

State Preservation

IBM supports a "sticky" port option for user session persistence. (i.e. for SSL communications). Please refer to the Acuitive VRM Technology Report for a detailed explanation of "sticky ports" and their implications.

3.11.4 Product Features at-a Glance

FEATURE	CAPABILITY	COMMENTS
TYPE SOLUTION	Software	AIX, NT and various Solaris flavors are supported.
OPERATING SYSTEM	NT, Solaris, AIX	
DELAYED BINDING	NO	
PERFORMANCE	2,200 connections per second	Obviously depends on the hardware platform selected. The software introduces minimal latency.
ROUTER OR BRIDGE	Bridge	MAT
REDUNDANCY	Active-Standby	With basic state preservation through heart beat protocol.
AGENT TECHNOLOGY?	YES	Highly configurable. If you can script it, it's an agent feature. Not required, but *highly* recommended.
MECHANISMS	• ISS (DNS) Redirection • HTTP Re-Direction • MAT	
POLICIES	• Round Robin, Weighted RR • Open connections • Response time metrics • Predictive and observed server response time (socket level) • Content response time metrics (HTML "get") • WAN best health (least loaded)	
FEEDBACK	• TCP Connection Verification • Response time monitoring • Active Content Verification • Host based	Based on Agent technology (CPU util., I/O, etc.) Highly flexible.

Management Features

The SecureWay management function can be accessed via SNMP Management software. A Java-base management system support is also included with the SecureWay package.

Key Additional Features

- supports transparent proxy devices, including firewall and cache servers
- IP packet filtering

3.11.5 Key Relationships

IBM can call upon its many OEM relationships to augment its professional services.

3.11.6 Issues

- **No Active Content Verification or Dynamic Application Verification**
- **No delayed binding functions**. No SSL Session ID Tracking, no URL-based scheduling, no cookie-based persistence or preferential services
- No Active-Active or Active-Standby redundancy options
- Single CPU scheduler architecture. Lacks scalability

3.11.7 Final Analysis

It is sometimes hard to get a fix on IBM's VRM capabilities because it tends to be buried in a whole lot of other functionality and service IBM can bring to the table. That's not necessarily a bad thing. If you are looking for a state-of-the-art VRM solution to augment your bleeding edge web site, IBM is probably not the answer. But if you are looking for a turnkey solution for design, applications, management systems, and infrastructure including (possibly) outsourced operations, IBM may be a great choice.

3.12 IPivot (www.ipivot.com)

3.12.1 Company Overview

IPivot is a start-up company in San Diego that has received over $14 million in funding. The product goal is to improve web site reliability, scalability, performance and quality of service (QoS). As of April 1999, IPivot has around 80 employees worldwide.

Marquee customers include 1-800-Flowers, TheStreet.com, and Time Warner.

3.12.2 Product & Support Overview

Intelligent Broker 4000

This entry-level load balancer is for local applications only, but can be augmented by a software upgrade to include multi-site load balancing and request routing. It provides response-time /priority based Layer 4 load balancing with full out-of-path return capabilities (6,600 c/s). Up to 4 Brokers can be used to provide linear scalability to 26,000 c/s. No URL parsing is supported in this software suite. The hardware provided is a 400Mhz Intel processor that comes with a standard 128 MB RAM. The currently supported network interfaces are 10/100 Ethernet. The OS is based on BSD, with some heavy kernel tweaking. The 4000 is priced at $14,995.

Intelligent Broker 7000

This box supports everything that the 4000 does plus adds URL parsing and Intelligent Session Recovery. This is the same physical hardware and OS as the IB 4000. The 7000 is priced at $24,995.

Intelligent Broker 7000M

This is the multi-site variation of the 7000. With this specific software option, all WAN features are enabled. Again, this is the same physical hardware and OS as the IB 4000 and 7000. The 7000M is priced at $29,995.

Commerce Accelerator 1000

We're not sure we have the product name right, but at the spring '99 Interop (in Las Vegas), IPivot announced a web server cluster "front end" box which offloads the key management and encryption/decryption functions associated with SSL, using a hardware assisted platform. If you were performing such functions in software on your web servers, you can potentially get much better system performance with this product. If you were performing such functions with hardware assist on your servers, the advantage is that it makes encrypted cookies, URLs, etc., available to the VRM scheduler to make more informed load balancing decisions. Several of these boxes can be deployed in a serial manner to increase overall SSL processing throughput and to provide redundancy.

Commerce Accelerator 8000

This is a combination of a fully featured VRM unit with the accelerator described above.

Support

The basic warranty is a one-year repair/replacement guarantee. The software has a 3-month warranty, which covers free bug-fix/minor change upgrades and discounts major upgrade purchases. There are two optional support plans. The Standard Support plan costs $5,000

annually. Standard highlights are: next day shipping for replacement product, 6am-6pm (PST) phone support, secure web site access, and one-day on-site support. The Premium Support plan costs $7,500 annually. Premium support covers all Standard options with the addition of 7x24 phone support, unlimited on-site support during problem periods, and next-available airplane replacement product shipping. Both plans cover free software upgrades. These support offerings benefit from IPivot's decision to use software keys to turn on more advanced features. To turn your IB 4000 into a 7000M, you simply place a support call, get your key and use it to turn on the 7000 series features.

3.12.3 Product Functionality

The IPivot Intelligent Broker offering is a full-featured load balancing solution.

Local Load Balancing

There are two options for local load balancing mechanisms: MAC Address Translation and TCP Diddling. These mechanisms correspond to two service-level options offered by IPivot:

1. Hot Mode (MAT): Traffic Classes defined by application or protocol type (i.e. UDP/TCP Port)
2. Rich Mode (TCP Diddling): Traffic Classes defined by content (URL, file type, file path, file name)

MAC Address Translation

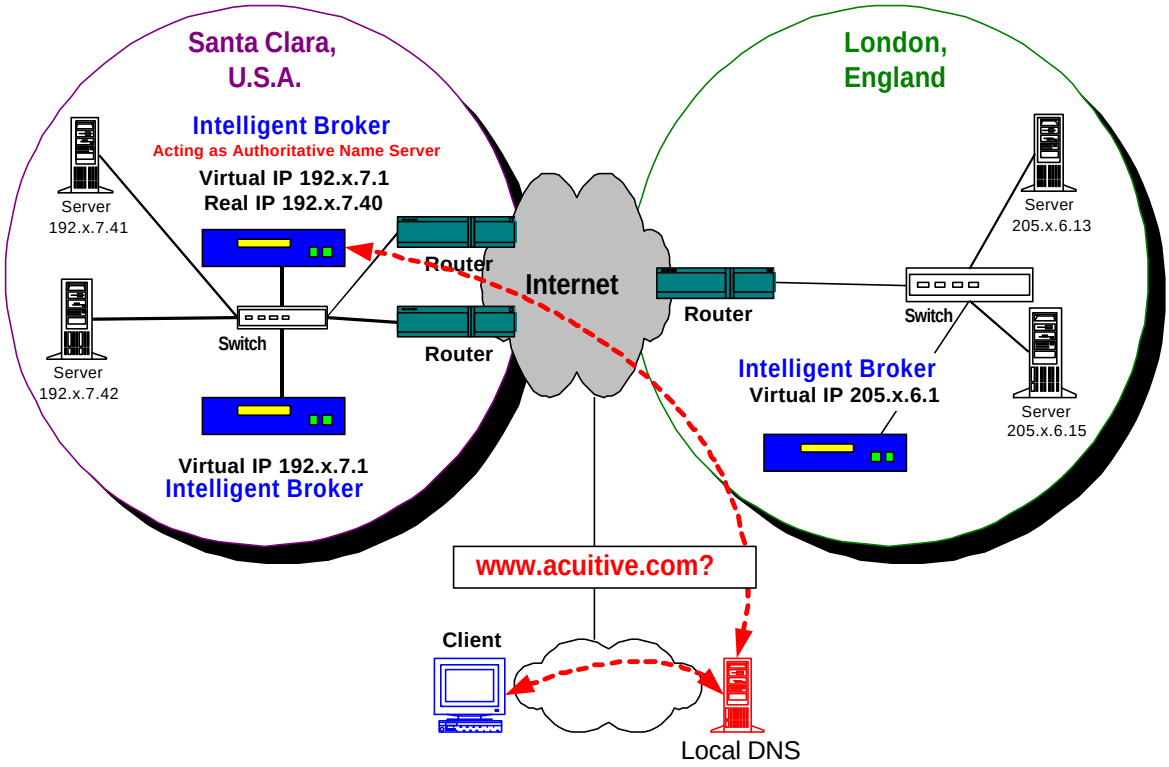
In the "hot mode" of operation, you have the option to enable direct-return of traffic from the server-to-the-client. This is often preferred for high performance systems because there can be an order of magnitude more traffic in the server-to-client path than the client-to-server path. In this mode of operation, the content servers are configured with the Virtual IP address on their loopback interface. The Broker performs MAC Address Translation of the TCP SYN packet and all ensuing packets from the client, to send the packet to the intended server, which can then respond directly to the client.

TCP Diddling

For Rich Mode operation, the Intelligent Broker performs *delayed binding*. When the client makes a request to a content server, the Broker receives the incoming request. It then responds for the server, so the client thinks it has a direct connection. When the client next sends the URL (or content) request, the Broker can inspect the packet in order to determine what content (location) is requested. The Broker will then set up a separate connection with the most capable and available server of the moment and pass on the client's request. The fulfillment server sends the traffic directly back to the Broker, where it is transferred to the client using the TCP parameters of the original connection. This process does not require any software on the server. We call this mechanism "TCP Diddling." It requires that the Intelligent Broker be in the traffic path in both directions – client-to-server and server-to-client.

Multi-Site Load Balancing

To provide continuity for a multiple geographic site web presence, the primary site Broker can act as the Authoritative DNS server for the domain. By maintaining a table of active Broker sites and related V_IPs, the Authoritative Broker will respond to iterative DNS queries with the V-IP addresses of a healthy target site. This provides transparent request distribution across servers located around the world. This also acts as a WAN fail-over mechanism in the case of catastrophic site failure.



Brokers support DNS Re-Direction in the conventional manner. The figure above illustrates the ability to use a single domain name across geographically disperse locations. This configuration allows for WAN re-direction and fail-over, thereby increasing the chances of a satisfactory client experience.

The top Broker at the Santa Clara facility is acting as the Authoritative Name server for the acuitive.com domain. The client's request to resolve the IP address of acuitive.com is passed to this Broker (using it's Real IP address), where it returns one of two V_IPs (192.x.7.1 or 205.x.6.1), depending on which site it decides the client should use. Requests can be directed in a round robin fashion, least-hop count or to the site that is the "closest" to the client's Local DNS server, as measured by WAN latency. Once at the appropriate site, the Brokers act as local Load Balancers, directing the request to the server that can best serve the client's need.

In this manner, only one domain name is required, each location can use different providers and the overall administration effort of this multi-national presence can be reduced.

Other Important Functions

Redundancy

IPivot has two methods for implementing scheduler redundancy.

- The first uses Route Advertisement as the method for fail-over between local Brokers. At the Santa Clara site in the diagram above, the active Intelligent Broker advertises a route to a V_IP. The site's WAN routers view that active box as the next router hop for the V_IP subnet and direct all traffic to it. Meanwhile, a 2nd Broker can be configured to be the standby Broker for the V_IP. A UDP-based keep alive protocol is used between the boxes. If the

standby Broker stops receiving keep-alives from the active, it will start to advertise a router to the V_IP subnet. The WAN access routers will shortly re-calculate their routing tables to use the remaining healthy Intelligent Broker

- The 2nd is standard serial cable failover, where one box is active and the 2nd is in a warm standby mode, configured with the same tables and policies as the active box so that service disruptions are minimized upon failover

Load Sharing

Up to four Brokers can be logically "ganged" to increase overall scheduler throughput. In this scenario, all four Brokers advertise equal path routes to the V_IP subnet. The access router, if it is running OSPF, can be configured to balance sessions among the four equal cost paths.

Intelligent Session Recovery

In Rich Mode, Brokers support the unique capability to do "error correction." For example, when a server responds to a client with a 400, 500 or 600 series error (i.e. 404-file not found) the Broker can intercept this error message. Instead of passing the error on to the client, the Broker simply tries to establish a connection with another server that contains the same content/function. In this way, the client will receive the requested information, but will never know the original server could not fulfill the request.

We can extrapolate the benefits of this capability by looking at a dynamic application scenario. When a client's request passes into the site it soon makes its way to an application server. Although all other aspects of the site are functioning properly, the application server itself may send an error message to the client (i.e. 606 ODBC error). The Brokers will capture this error and attempt to send the request to a different application server that can process it. In this fashion, the Intelligent Session Recovery feature can shield the client from common errors and keep the session moving along, transparent to the client.

QoS Metrics

Another core product function is the ability to make forwarding decisions while applying a preset maximum-server-response-time metric. This is currently a rather unique feature for the Load Balancing arena and the Brokers are designed to use this "QoS" metric extensively. The two most basic factors a Broker uses to service a client request are Server Response Time and the predetermined Request (type) Priority.

A further mechanism of import is related to how the Brokers actually *enforce* "Quality of Service." IPivot provides the option of defining a response time for virtual services and an order of "preference" between these services. The Broker constantly monitors the response time associated with each virtual service (by looking at live traffic). As the Intelligent Broker observes response times for a critical application heading towards the pre-defined threshold, an attempt to maintain the service level will be made using one or more of the following mechanisms:

- Bring a back-up server on-line for the high priority service
- Stop sending traffic for lower priority services to some servers which also support the higher priority service
- Throttle back requests for lower priority services

3.12.4 Product Features at-a Glance

FEATURE	CAPABILITY	COMMENTS
TYPE SOLUTION	PC Platform	Intel Pentium 400 MHz, 128 MB RAM standard
OPERATING SYSTEM	BSD	
DELAYED BINDING	YES	
PERFORMANCE	6,600 connections per second, 95 MB throughput per Broker (MAT, direct path return) 3,000 connections per second (HNAT) 600 connections per second (delayed binding and PCV)	Performance data supplied by vendor. Up to 4 brokers can be connected in parallel to provide over 26,000 connections per second using Symmetrical Multi-path Routing
ROUTER OR BRIDGE	ROUTER	
REDUNDANCY	Active-Active	OSPF Symmetrical Multi-Path Routing
AGENT TECHNOLOGY?	NO	No plans to add agents
MECHANISMS	• DNS Redirection • TCP-diddling • MAT	
POLICIES	• WAN least cost/closest • QoS maximum allowed response time metrics • URL parsing/content specific • Predictive and observed server response time (socket level)	Backup and Overflow servers
FEEDBACK	• Socket connection requests • Response time monitoring • Active Content Verification	

Management Features

The Intelligent Broker supports “traditional” (SNMP) management. Administrative and status functions are also supported by command-line interface and a Java-based browser interface. This feature also allows dynamic software loading/unloading (for upgrades/updates).

Key Additional Features

- supports transparent proxy devices, including firewall and cache servers
- full TCP/UDP/IP protocol suite support
- IP packet filtering
- secure cross platform GUI

3.12.5 Key Relationships

IPivot has signed an OEM deal with Nortel that allows Nortel to resell their load-balancing solution. As part of this announcement, the companies announced the intention to deliver the IPivot capability as a "blade" which can be plugged into various Accelar switches. They also announced, as part of the Nortel agreement, that they will be working to integrate their QoS polices with the QoS polices of a number of switch and router manufacturers. They are looking at mapping VRM polices to 802.1p, 802.1q and TOS bit information that is available to us in both the incoming and outgoing paths.

3.12.6 Issues

The main issue with IPivot is the relative newness in the marketplace and therefore reduced customer validation across a wide range of environments. But that's happening at a rapid pace right now.

- As with other single CPU solutions from other vendors, we wonder whether Intel/BSD route will remain viable during a period such as this where available WAN bandwidth is increasing at a pace greater than Moore's Law. IPivot exacerbates this issue by supporting some powerful features which require an inordinate amount of CPU resource. In Rich Mode all packets are being examined deeply into the URL fields *in both flow directions*, due to PCV. In addition, significant processing for management and QoS functions is required. As a result, the Broker performance in this mode is a full order of magnitude lower than in the simpler immediate binding and MAT mode of operation
- IPivot's ability to gang up to four schedulers together to create a larger "virtual scheduler" would pretty much mitigate the above issue, except that to make that function work in Rich Mode requires that the Source IP of the client packet be replaced with a unique IP associated with the Broker. Doing this reduces the granularity of information that can be gleaned from server-side logs, which impacts many site procedures for analyzing both technical and marketing information related to site operation. In many cases, it simply means you can't use IPivot's Active-Active option and therefore are either stuck with low performance or with not using the Rich Mode option. You also have to be careful of topologies where OSPF path reconvergence could lead to the "movement" of users between schedulers in an undesirable way
- Passive Content Verification with Re-Direction (what IPivot calls Intelligent Session Recovery) can be a little tricky. HTTP Return Codes are often useful for identifying a specific server problem, but just as often they are indications of a site-level problem (often associated with the back-end database or file server) or are ambiguous. Re-direction through a different web server to the same back-end problem doesn't help anybody. We're not yet bought into the value of this feature relative to the performance degradation, and the possibilities of the feature being exploited as a Denial-of-Service attack.
- Internally, IPivot supports an API based on their Java-based Portlet technology. They use this to work with third parties to integrate server-based information with the IPivot decision making logic. The Java portlets communicate using RMI. IPivot should consider making this toolkit more readily available, to key customers and consultants, so that 3rd party integration specific to the user's environment can be performed without loading down IPivot engineering or services

3.12.7 Final Analysis

IPivot has come up with a few clever technological ideas:

- Intelligent Session Recovery (it's kind of like HTTP Re-Direction without the delays and address advertisements)
- Using routing protocols for fail-over
- (Largely due to the above), the ability to use multiple schedulers, scheduling for the same service to the same Real Servers. They are the only hardware product which can do that today
- The QoS features

We also like the direct-to-client option and feel that the multi-site features are competitive for the most part, although not necessarily differentiating.

However, there are compromises and trade-off associated with using some of the features that make them a little less appealing. You can't get direct-path return in Rich Mode (i.e you can't do delayed binding and direct path return simultaneously). The Active-Active scheduler scheme requires that you modify the source IP address as it passes through the scheduler (if you are doing delayed binding) which limits the value of any management activity that leverages server-side logs. The Intelligent Session Recovery requires delayed binding (which means no direct path return and no Active-Active scheduler scale-ability) *plus* requires looking deeply into every packet in both directions of transmission, which can really impact performance. So while we like the features in concept, there is pain associated with every gain.

Until recently we would not have predicted good success for IPivot in the marketplace because they were a late entrant. Even today, their customer pool is small compared to people who have been in the business for two years or more. So we would have shied people away from IPivot for non-technical reasons. But the OEM deal with Nortel Networks changes all that. That deal should really help IPivot get access to customers, sell product, and invest in evolving the product line. But OEM deals for system-level products are tricky. Especially when a small company is supporting a large one, and the technology in question is complex and changing fast (as in VRM). In the beginning, everyone is enthusiastic and announces all kinds of plans for future value-added integration. Sometimes that happens, usually it does not. What you can probably count on is that quickly Nortel gets one or two software releases behind IPivot and that the so-called "value-add" that Nortel brings becomes "value-subtract" when it comes down to real VRM design and deployment issues. The good news is that the OEM deal probably established IPivot as a supplier that will be around for awhile. If you are going to buy the product from Nortel (at a higher price, by the way), you should still establish a technical-level relationship with IPivot so that you have a pipeline to the people that live and breath this technology 24 hours a day. We only hope that Nortel doesn't load down IPivot with a bunch of big company overhead and prevent them from moving as quickly in the future as they have in the recent past.

3.13 Radware Ltd. (www.radware.com)

3.13.1 Company Overview

Radware Ltd., is part of the multi-million dollar RAD group, a family of 13 companies serving various niches with-in the networking and communications industries. Radware is a subsidiary of about 60 people focused purely on load balancing of servers, cache servers, and infrastructure devices such as firewalls, and has been shipping products since early 1997. Although Radware is focused entirely on the Load Balancing arena, it has access to the various worldwide RND resources.

Some of RadWare's marquee customers include Dow Jones-Interactive, The Discovery Channel, Vanguard, Tickets.com, StockSmart, and Gateway.

3.13.2 Product & Support Overview

Radware is fundamentally a software company, and produces several different products that are really individual software loads on top of a networking appliance platform. They produce a line of server load balancing products that range in capabilities and price (WSD-Pro, WSD-DS, WSD-NP) and two other traffic management appliances (Fireproof and Cache Server Director). The only difference between these products is a EPROM, which makes upgrading from one capability to another pretty simple.

WSD Pro

The Web Server Director (WSD Pro) is a local SLB that provides support for multiple Virtual Servers, URL parsing, and has been optimized in features and price for local load balancing. WSD Pro is priced at $15,280 for the latest hardware version and $7,340 for the same software running on the previous hardware version.

WSD-DS

The Web Server Director for Distributed Sites (WSD-DS), has the WSD Pro features for local SLB, plus the ability to re-direct traffic to other WSD-DS units at other sites via DNS Re-Direction, HTTP Re-Direction or Triangulation. The load and health of each site is reported to all other sites via a Load Report Protocol, and user requests are distributed based on the load/resource availability at all sites.

WSD-NP

The Web Server Director for Network Proximity (WSD-NP) adds round trip delay as feedback for multi-site (via the Proximity Report Protocol). The "best site" is determined by a combination of router hops e and delay factors, unless that site is down or overloaded. "Static" proximity based source IP address ranges can also be used to determine the site that a given user should connect to.

Fireproof

Fireproof is a state-monitoring MAC Address Translation device that can be used to load balance firewalls, VPN devices, WAN bandwidth managers, routers, and other infrastructure devices. It allows network administrators to create highly available and somewhat scaleable firewall environments. Radware is currently listed as a Checkpoint OPSEC partner.

Cache Server Director

Cache Server Director completes the picture with the ability to direct selected IP/HTTP traffic to proxy or transparent proxy servers. It performs Delayed Binding as part of its mechanisms, and is capable of URL inspection and rewriting to optimize cache server usage and deployment options.

Hardware Platforms

Radware's current products ship on top of network appliance devices built by another division of the RAD Group. They feature Intel i960 CPU's. Users may choose between a two-port Ethernet model, a four-port Ethernet model a two-port Fast Ethernet model, or a 4 port Fast Ethernet model Each ships with 8 MB RAM, but they can be upgraded to 32 MB of RAM. More memory gives you larger table sizes. For example with 8 MB of RAM, the WSD Pro can handle up to 30,000 simultaneous clients, while with 16 MB it can handle 80,000.

Recently, they started shipping a faster appliance unit based on a 75Mhz i960Hd CPU. There is no price uplift associated with this upgrade, so the price/performance of the products is improved by about a factor of 2.5x with this enhancement. Also, RND announced a price reduction across the product line for products running the older 33 MHz i960 processor. The prices quoted above are for the newer processors.

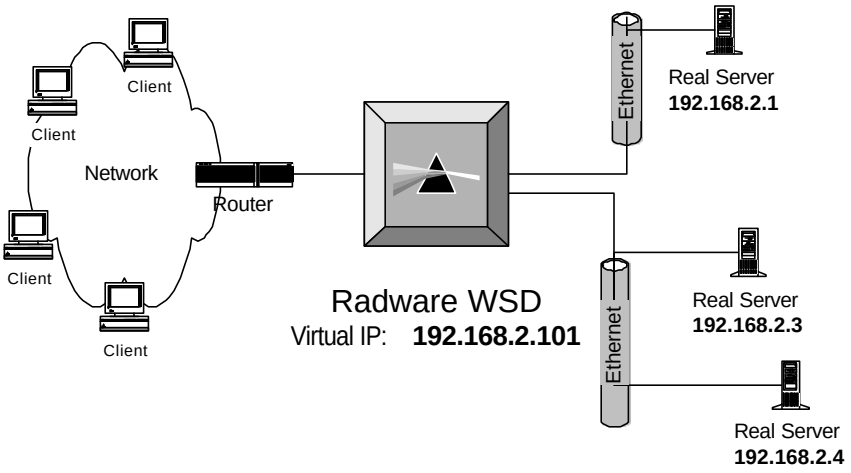
Support

Service plans include toll-free 24x7 technical help line support, free software updates and 24-hour advance ship replacement system should a unit fail.

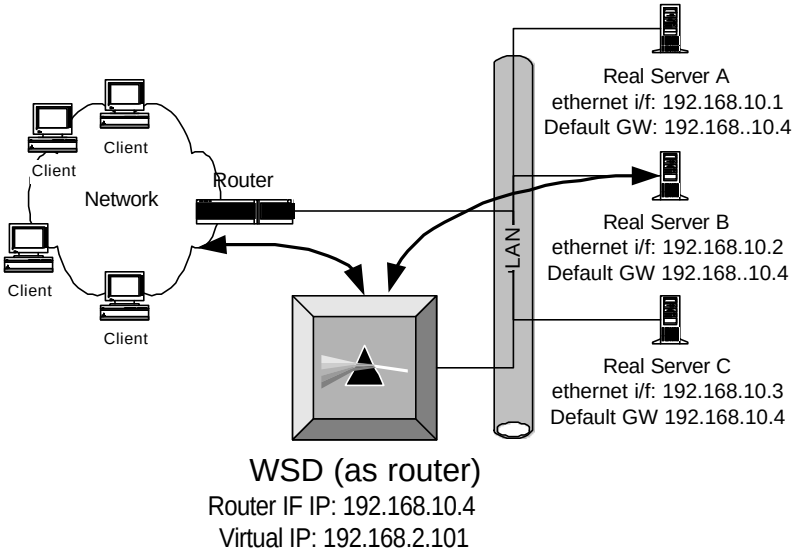
3.13.3 Product Functionality

Local Load Balancing

WSDs can be configured in a variety of different ways for local load balancing. Historically, most SOAS products were deployed in the physical data path, between the site access router and the content server LAN. This is one option for a WSD.

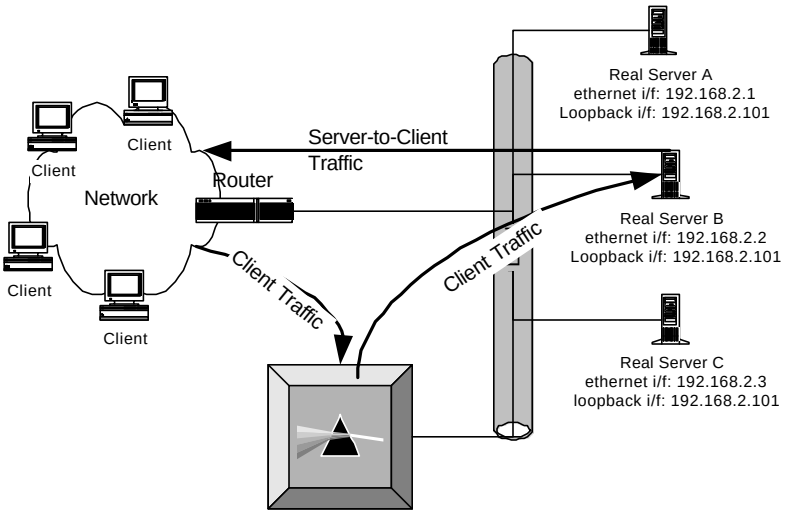


However, the WSDs can also be configured in two alternative configurations. Since the WSD is a router, the content servers can be configured to point to the WSD as their default gateway. In this scenario, the WSD performs Half NAT. While the WSD is in the data path for both client-to-server and server-to-client directions, it need not be physically installed between the access router and the content servers. It is just another device on the network. This eliminates the need for wiring changes and makes redundant and multiple scheduler configurations simple to deploy.



The configuration above is the one most often used when the WSD is configured to perform delayed binding via **TCP Diddling**. This option is required for URL parsing and SSL Session ID-based load balancing.

When delayed binding is not required, WSDs can be configured to support **MAC Address Translation**, which allows for the content servers to send client-bound traffic directly back to the access router, by-passing the WSD. Greater throughput can be achieved in this mode.



WSD (direct return)
Virtual IP: **192.168.2.101**

See the Technology Report for more information about all these modes of operations.

The WSD uses Active Content Verification (ACV), TCP Connection Verification (TCV), UDP Application Verification, and ICMP pings to verify the status of the servers/applications that are part of a server farm. Server load is measured by number of clients connected (by IP address), number of open connections (by TCP connection), packets in/out of the server, and/or feedback from SNMP agents on servers.

Multi-Site Load Balancing

Radware WSD-DS and WSD-NP products support more multi-site re-direction mechanisms than almost any other vendor does: DNS Re-Direction, HTTP Re-Direction and Triangulation. Usually, a complete solution will use more than one of these approaches, with one implementation acting as a back up for the other at different levels of the architecture. Also, it's important to note that in an overall system, any of the Radware products can be used at any site (PRO, DS, NP), with the DS and NP products acting as the controllers for traffic flow to the different sites.

To provide information to aid in "best site" selection, Radware supports two IVP protocols that run between the sites. Site health and load information is shared via **Load Report Protocol** packets. These LRPs contain site load statistics (current connections, preset source IP threshold capacity). All products (PRO, DS, NP) can communicate LRP to a DS or an NP.

The WSD-NP is capable of directing WSD units (of all types) at each site to measure the number of hops and path latency back to the client (or the client's local DNS server). The site-to-client path information learned by each WSD is then communicated back to the WSD-NP by another proprietary protocol called the **Proximity Report Protocol**. This information can be used at a later time to determine which site is the "best site" to send future requests from that same source IP to.

Using the information attained from these protocols, sophisticated "best site" determinations can be performed. These determinations can be used independent of whether the re-direction mechanism is DNS Re-Direction, HTTP Re-Direct, or Triangulation.

DNS Re-Direction

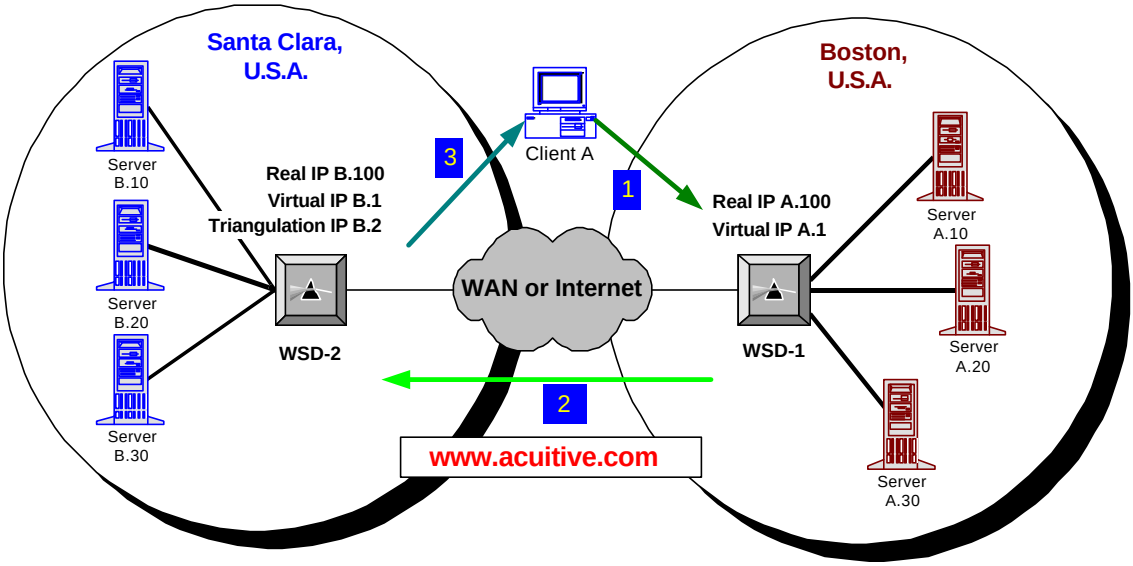
DNS Re-Direction is the most commonly used of the various multi-site options provided by RadWare. DNS Re-Direction is explained in detail in the Technology Report, but summarized here.

WSD-NP acts as an Authoritative Name Server for DNS systems. RadWare provides a lot of flexibility as to how such a system can be set up. In one scenario, the site DNS is set up to point to a single WSD-NP (with a 2nd as back-up) and that WSD-NP provides resolves all the name requests. In another scenario, the site DNS can be set up with a list of WSD-NP addresses, which it rotates through, directing each resolution request to each WSD-NP in a round robin fashion. The WSD-NPs can then be configured to simply return their V_IP (thus proving site health and accessibility) or it can itself have a list of WSD-NPs and load/health/proximity information from which it selects the "best site" and returns the V_IP used by that site.

Triangulation

Triangulation is a method where arriving client TCP connection requests (SYN packets) can be sent to another site better suited to service the request. The Triangulation method uses special "shadowed" V_IP addresses at each other site that traffic might be forwarded to. When that site sees something arriving to the unique "shadowed" VIP, it knows that the packet was re-directed to it, and what site specifically performed the re-direction. That information is necessary, because on the return path (server-to-client), the WSD must put the original V_IP, from the site the client originally addressed, into the source IP field. That way, the traffic can return directly to the client. That's what creates the triangle; client to original site, site A to site B, site B back to client.

It works something like this…



1. Client A has received the IP address of Virtual Server A.1 through a DNS query for the domain *acuitive.com*. When the client sends an FTP-Get request to V_IP A.1, the WSD-1 realizes that this site cannot service this request. Therefore, WSD-1 decides to send this request to the closest site that can process and respond.
2. WSD-1 sends the original request to the WSD –1 **Triangulation** V_IP of WSD-2 (B.2). WSD-2 recognizes packets arriving to this address to be associated with the virtual service provided by V_!) B1, but that the packet has been sent from WSD-1. WSD-2 forwards the packet to a local server that can best handle its needs. When the packet is ready to be returned to Client A, WSD-2 uses the source IP address of A.1, the original virtual address for WSD-1 that the client originally used.
3. The responses are returned to Client A, which thinks it is still communicating with A.1. When Client A sends another packet to the acuitive.com domain (ack, etc.), it uses the A.1 address. In this manner, WSD-1 maintains a table of Triangulated sessions, performs the necessary header manipulations, and forwards the packet to the unique "Triangulation IP Address" of the WSD-2 site until the Client A session is complete.

The Value of Triangulation and HTTP Redirect

Triangulation and/or HTTP Re-Direct can be useful options to augment the foundation DNS Re-Direction capability. If a site becomes temporarily overloaded, possibly because the DNS mechanisms doesn't have enough granularity to differentiate between users (a heavy load all coming from users behind one local DNS, for example), one can re-direct traffic to another site to spread the load. Or, if a site's content servers fail, Triangulation or HTTP Re-Direction can be used to re-direct traffic from clients who have cached the IP address from a previous DNS resolve. Finally, if you can create a topology where a WSD can interdict traffic before it gets to a site, entire site failures or WAN link failures can be quickly re-directed around.

Redundancy

WSD products support the Active-Active mode of redundancy, which allows significant deployment flexibility. WSDs can be set up so that one box actively supports V_IP_A and the other V_IP_B, but they act as secondary back-ups for one another. When using direct-path-return (which implies no delayed binding), V_IP_A and V_IP_B can be supporting the same service, using the same pool of content servers, which is a form of scaleable, multiple schedulers. The only flaw in that application scenario is that one WSD is generally unaware of the traffic being sent to servers by the other WSD, so load might not be evenly distributed at all times.

3.13.4 Product Features at-a Glance

FEATURE	CAPABILITY	COMMENTS
TYPE SOLUTION	Embedded Platform	i960-33Mhz or 75Mhz, 16 MB RAM standard
OPERATING SYSTEM	Embedded	
DELAYED BINDING	YES	• SSL Session ID Balancing • URL Parsing for both Caching and SLB
PERFORMANCE	• 120+ Mbps, 3500 connections/sec with immediate binding • Direct-Path-Return can achieve many 100s of Mbps throughput for applications with larger file sizes • Roughly 50% of the performance with delayed binding	Performance data provided by the vendor
ROUTER OR BRIDGE	IP ROUTER, BRIDGING	RIP, OSPF
REDUNDANCY	Active-Backup	Keep-alives and session information is shared across the local network (no RS-232 cable)
AGENT TECHNOLOGY?	NO	Can leverage SNMP agent technology on NT and other OS.
MECHANISMS	• NAT • MAT • TCP-diddling • HTTP Re-direction • DNS Redirection • Triangulation	
POLICIES	**Local** • # clients • Server load (packets) • Least connections • SNMP-based server feedback • URL-based parsing/content specific **Multi-site** • Fewest router hops • Least round trip delay • Client table fullness • Site health (% servers up, current load)	• Back-up servers
FEEDBACK	• Device ICMP Pings • TCP Connection Verification • Active Content Verification • Response time monitoring • SNMP-based feedback • Server packet load	

Management Features

Full configuration and monitoring is all available through the local CLI, which can be accessed remotely via Telnet.

Radware also offers a Java-based management tool called ConfigWare for both configuration and monitoring. ConfigWare can operate on any platform. If installed on a compatible OS/web server, it provides browser-based management. ConfigWare also has an HPOV module which can be loaded on top of HPOV to provide an integrated solution on that platform.

Key Additional Features

- Flexible IP packet filtering by L2/L3 addresses, protocols, and applications
- SSL 3.0 Session ID Tracking
- Cross-subnet directing (Real Servers can be on any subnet)
- SYN attack protection
- Server slow start, Graceful deactivation
- NAT for private IP addresses servers accessing the Internet
- Direct-Path-Return

3.13.5 Key Relationships

Radware has formed some solid alliances with firewall vendors, specifically CheckPoint and Secure Computing. Firewall load balancing is an area of marketing emphasis for Radware.

3.13.6 Issues

Radware has a generally complete product line, and there are few major feature deficits that would prevent Radware from being a candidate for any application.

Some issues and concerns include:

- The use of the i960 processor and single CPU architecture. In some ways this is better than Software-on-a-Stick. The embedded real-time executive is less overhead than UNIX. The operating system doesn't require a disk drive to boot. But the i960 enhancement path is lagging Pentium. It looks like RadWare's performance is about 60-70% of what you can get with the latest Pentium, which is a gap that will widen over time. And for the most performance intensive applications, we don't even see Pentiums cutting it

  Radware has a feature-rich set of product offerings, but (as with all single CPU implementations) we worry about performance and capacity when many functions are active simultaneously. For instance, imagine running a WSD-NP product as both a local load balancer and a multi-site load balancer, performing delayed binding for multiple servers, while providing routing and OSPF support. The system must simultaneous perform NAT, maintain session state for tens of thousands of clients, maintain site/client state tables, run a potentially heavyweight routing protocol and forward traffic. Lab tests never provide performance information under these kind of conditions

- Having said that, we must say that at least 80% of the applications in the world don't need the extreme kinds of performance under the extreme system configuration scenarios that we

worry about. So, unless you really feel you are pushing the limit (which takes a combination of fat WAN links and lot's of features turned on), then you needn't worry about performance too much

- The platform also doesn't have the memory scalability of a PC-based architecture. That hasn't been a problem for Radware in the past, because they used a fairly simple (but effective) approach of tracking client IP addresses only. As they get into functions enabled by delayed binding – SSL Session ID tracking, URL-based scheduling, cookie-based persistence or (possibly) QoS, they'll find that much more state information needs to be maintained per user and per connection. The same is true of large multi-site situations if significant router hop and delay information needs to be cached for a large number of active clients. Memory constraints could effect their ability to scale. We've seen lot's of situations with other vendors where ½ Gig of RAM or more is needed to support appropriate table sizes
- Radware has a flexible method for accepting SNMP-based feedback, but it does not have an API or equivalent capability to accept other forms of server co-resident feedback. SNMP agents do not universally support all operating system information, typically do not support application status or performance monitoring, can be unreliable across WAN links and will not typically be allowed to flow across ISP Internet connections. Additionally, using SNMP to poll too often will cause the systems that are being monitored to spend a significant portion of time responding to SNMP. Poll too infrequently, and the information learned through SNMP may be "old news." In general, such information should only be used for long term adjustments to load balancing, and for capacity planning. But for these purposes, they are very powerful. So this isn't a knock against the product, just advice on how to use this particular function
- Radware does not have the ability to perform Dynamic Application Verification. A user could write an SNMP-capable application on a server to perform health checking, but this requires significantly more effort than if they included server-side agents and/or an API. This may remove them from being considered for load balancing in a multi-tier application environment
- Radware does not have any bandwidth management capability, which is often useful (but not absolutely necessary as part of the VRM product) at hosting sites
- Radware's products only ship as discrete functionality units. If a customer wanted to perform firewall or cache server load balancing in addition to web server load balancing, they would have to purchase, install and manage more than one type of WSD device. Some VRM vendors offer the ability to provide this solution with a single VRM unit.

3.13.7 Final Analysis

RadWare is a fast moving company who has, and continues to, put a lot of effort into improving their product. Almost every shortcoming we identified in the previous version of this document has been addressed.

We like the fact that we can support multi-site and local load balancing functions from a single product platform. We also like the fact that RadWare has a solution for all of the non-server load balancing applications of today. We also view that RadWare is one of only two vendors with a complete set of multi-site options, at least what has been invented so far. All of this adds up to flexibility and one-stop shopping.

The company is focused on this space and everyone we have met with or talked to is high energy and sharp. They remind us somewhat of F5 Networks in terms of the quality and focus of the

people. They are entirely focused on this market segment and have proven their ability to understand customers needs and quickly provide solutions. They also provide excellent customer support, both before and after sales. This is largely a reflection of their focus on this space and good hiring.

RadWare has achieved a significant position in the marketplace. It is unclear whether they or F5 Networks is #2 in the market, behind Cisco. Their biggest issue is competing with the gorilla (Cisco) but by all accounts, they and others appear to be winning significant market share. RadWare has upwards of 200 active customers. Because of this momentum, and because of their position as part of the overall RAD corporate umbrella, they are one of the vendors you can pretty much count on being around for awhile.

RadWare already has most of the industry leading single site and multi-site features you would want, and their pricing is very fair – they are almost always competitive. The only real concern we have with Radware is their hardware platform. We don't think that the i960 is up to snuff with the Pentium, we thin the gap will get even greater with time, and we're not even sure Pentiums will keep up with the pace of Internet usage and WAN bandwidth provisioning that we envision. We also think that Radware is going to run into memory constraints in larger deployments using some advanced features that require storing a lot of state information.

We'd like to see RadWare take advantage of their ability to do internal hardware development to become one of the first vendors to leverage one of the hardware-assist chips which are coming onto the market to augment their single-CPU architecture and remove performance as an issue in almost all situations.

So, we view that RadWare is rarely the wrong choice – they are usually always on our short list. Their feature set and pricing is ideal for many mid-tier applications, especially non-server load balancing applications lie firewall balancing. But, except for in some multi-site situations where their breadth of capability is differentiating, they are also rarely the clear and obvious right choice for high-end applications where a complex set of features must be deployed. They are not feature poor. A new hardware platform would make them a better candidate for the high-end applications.

3.14 Resonate (www.resonate.com)

3.14.1 Company Overview

Resonate is a company of about 80 people based in Sunnyvale, CA. They have been shipping and supporting VRM products since Q1 '97. They are one of the pioneers of this product category, positioning themselves as the experts in Internet site high performance and availability for business and mission critical deployments

Resonate reportedly has about 150 customers, supporting 100s of active sites. Customers include GeoCities, Discover Brokerage Direct, E*TRADE, Excite, N2K, eBay, Internet Shopping Network, Sun Microsystems, Merrill Lynch, and BankBoston.

3.14.2 Product & Support Overview

Resonate is a supplier of distributed software solutions for Site Management. Capabilities include server load balancing, multi-site traffic management, site performance monitoring and reporting, and service level management.

The product line is divided into three major components.

Central Dispatch is Resonates local VRM product. Central Dispatch (CD) is a distributed software solution that is loaded onto each content server in a site. CD is supported on NT, Solaris, HPUX and AIX, with future support announced for Linux on Intel platforms. CD provides scheduling of requests that arrive at a site to the “best available server” based on open connections, server CPU load, URL, session ID, and other policies.

Global Dispatch provides load balancing between sites, by acting as the authoritative DNS for the web site host name. A Global Dispatch agent at each site is aware of the health and load on servers and shares that information with the Global Dispatch scheduler. The agent also provides latency measurements to factor into its decisions the WAN latency between each site and the client’s local DNS.

Commander is a recently released product that provides three main functions:

1. Site external monitoring: Commander measures site resource performance and availability, verifying client access to a site by initiating URL tests, host access tests, and HTTP service availability tests. The results of these tests can be used as feedback into Central Dispatch and Global Dispatch traffic management functions.
2. Statistics reporting: Commander logs site performance information for historical charting or trend analysis, collecting data from agents and monitoring tests at user-specified intervals. Data is maintained in industry-standard format to facilitate analysis using popular databases, report writers, and spreadsheets.
3. Feedback Control: Commander provides a flexible interface to allow site managers to define actions in response to a wide range of monitored events. Events can be identified by Resonate agents or external site monitors. But even more powerfully, events can be identified and reported by a wide range of 3rd party agents and knowledge modules, providing application-specific and server co-resident feedback. Actions can include:

- Schedule traffic away from a Web server that is tied to a slow or failed back-end application server
- Enable a backup content server in a Central Dispatch site when one or more active content servers fail or become too busy
- Monitor applications and server processes; restart any that fail

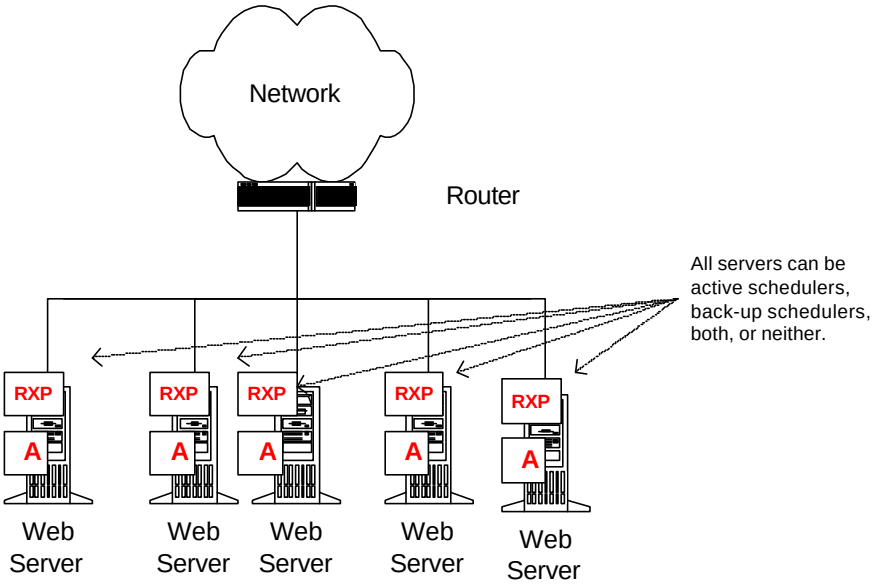
Resonate Professional Services can be contracted for to help with site design, solution customization, product training, and a variety of other services. The team is not big, but they are very good.

3.14.3 Product Functionality

Local Load Balancing

The Central Dispatch software has two components.

1. RXP, which is loaded at the driver level on each server (running in the kernel space), just under the IP stack.
2. An agent module that operates in the user space on each machine. Agents provide feedback on web server health and load (CPU utilization).



One or more of the servers (RXP functions) can be configured to be the scheduler or back-up scheduler for a specific V_IP, offering flexibility in deployment, server maintenance, and scalability One advantage of this flexibility is the ability to utilize the latest, fastest processors for the CPU-intensive scheduling functions, effectively tracking Moore's law in scheduler performance.

Central Dispatch uses a unique but proprietary method for Re-Direction called *Connection Hop*. Central Dispatch RXP completes the TCP connection process for the (as yet unidentified) Real Server as connection requests arrive from clients. When further information relevant to the active policy arrives (such as the URL), the Real Server can be chosen based on the specific content or resource requested. RXP then sends a message to that Real Server telling it that it wants to establish a TCP connection to an identified Virtual Service *as a proxy for the client*. It uses the same client IP address, source port numbers and sequence number information as was established

by the TCP connection set-up process with the client. The Real Server then "takes ownership" of the TCP connection and responds directly to the client, in a manner that is transparent to the client.

The advantages of Connection Hop are:

- As with TCP Gateways, more explicit resource, content, or user information (such as cookies) can be used to influence the decision about which server to direct the request to.
- As opposed to TCP Gateways, the SLB is only involved in traffic flow from the client to the server, which are generally just ACKs. After the initial connection set-up process, these packets require just simple encapsulation. All packets from the server-to-the-client bypass the scheduler completely.
- The "hiding" of real server IP addresses is maintained.
- Solution is very scaleable because of the sharing of functions performed between the scheduler and Real Server.

The disadvantage of Connection Hop is:

- It requires special and proprietary software on each Real Server.

Single points of failure are eliminated in CD with the installation of primary and backup schedulers, and by multiple schedulers acting as backups to each other. Automatic server and scheduler failover helps ensure that the site stays operational in the event of server or scheduler node failures

Multi-Site Load Balancing

Global Dispatch uses DNS Re-Direction for multi-site traffic management.

A GD agent at every site measures local server availability and load. Load measurements can be based on the time required to serve a typical Web page, the average CPU load across a defined range of servers, or on the time required to serve a database request. In addition, these agents make round trip delay measurements to source IP addresses (client Local DNS addresses) as directed by the GD scheduler. By comparing the results of the delay tests from each site, the GD scheduler can make more informed decisions about which site to send an arriving request to. Global Dispatch allows weighting between load and WAN latency metrics. A Quick Response mode allows latency measurements to be made in the background for cache loading, ensuring the fastest response to all inquiries.

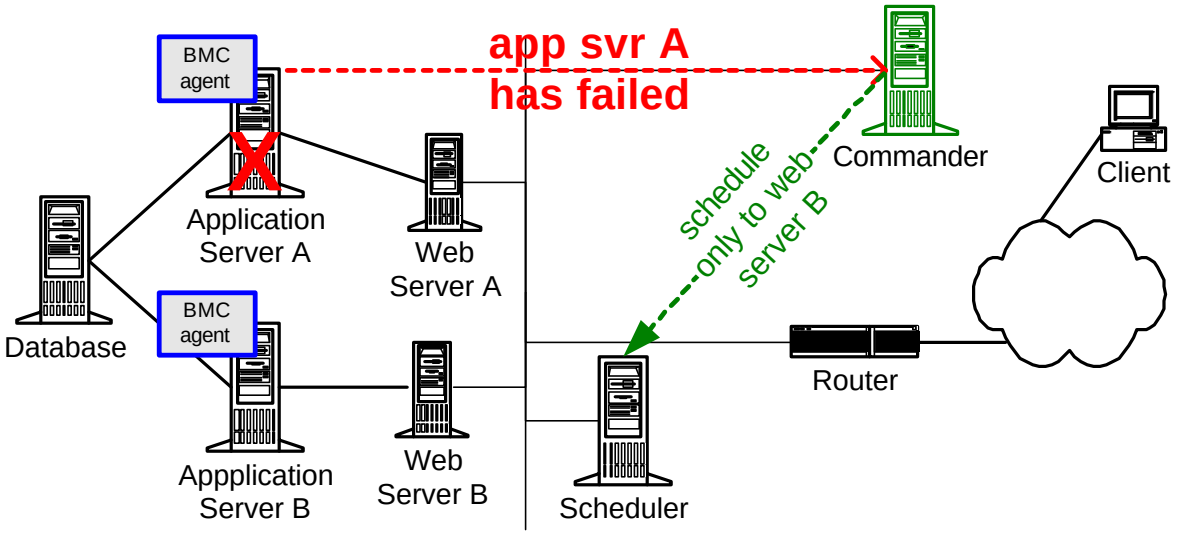
Global Dispatch allows administrators to dedicate a POP for specific users or groups of users. With support for persistent sessions, GD manages traffic between sites that maintain user state at individual POPs. This type of built-in functionality means Global Dispatch can simplify database replication and synchronization.

Support for custom scripts and APIs lets you influence Global Dispatch scheduling decisions to meet specific application requirements.

Global Dispatch can be deployed with multiple schedulers and agents (installed at each site monitoring site health) to eliminate all single points of failure. Global Dispatch does not require the deployment of Central Dispatch as a local solution, but can be integrated with Central Dispatch to share load information.

Event Handling

Commander supports an external interface to 3rd party functions via a small number of primitives, the most important being SendEvent. In many cases, a SendEvent message can be initiated by a 3rd party agent or Knowledge Module (such as from BMC or NetIQ), which, when received and interpreted by Commander can result in a specific action.



3.14.4 Product Features at-a Glance

FEATURE	CAPABILITY	COMMENTS
TYPE SOLUTION	Distributed Software	
OPERATING SYSTEM	NT, Solaris, HPUX, AIX, Linux	
DELAYED BINDING	YES	SSL Session ID, URL-based scheduling. Cookie-based scheduling is a future.
PERFORMANCE	Depends on number of active schedulers and host machine capability.	Server-to-client traffic bypasses the schedulers and therefore is supported at full wire speed.
ROUTER OR BRIDGE	N/A	
REDUNDANCY	Multiple schedulers, back-up schedulers.	No persistent session maintenance upon fail-over for CD Persistent session maintenance *is* supported between sites (via GD).
AGENT TECHNOLOGY?	YES	Provides feedback on web server health and CPU load.
MECHANISMS	• TCP Connection Hop (local) • DNS Re-Direction (WAN)	
POLICIES	• Least Connections • CPU load • WAN least cost/closest • Site persistence • URL parsing/content specific • SSL session ID	
FEEDBACK	• Server CPU utilization	Open interface via Commander enables a wide variety of feedback from a wide variety of 3rd party devices.

Key Additional Features

Central Dispatch

- Multiple primary and backup schedulers simultaneously scheduling into a single cluster, enabling increased scalability and availability without requiring the splitting of a site
- Shopping cart E-commerce capabilities, tying secure and non-secure sessions together
- Port load balancing, which allows multiple HTTP daemons to be assigned to the same virtual service, with round robin balancing between the daemons. This enables scalability within a server when the connection limitations of an HTTP process is the bottleneck, not the server itself
- Java-based GUI and command line management tools for remote management and monitoring
- Remote install/upgrade capability
- Graceful startup and shutdown
- Real-time reporting of CPU load, open connections, server status, throughput, hits per second, latency, and other metrics for monitoring and troubleshooting

Global Dispatch

- Advanced Traffic Mapping capabilities:
    - Sticky/persistent session support and sticky session failover
    - Directed Traffic Table directs users to predefined POP
- Configurable scheduling based on WAN latency and site load
- Shadow Mode for testing and tuning before deployment
- Advanced statistics including average DNS response time, POP hit rate, and other quality of service statistics
- Ability to pass multiple A-records (IP addresses) to a client for browser-based failover

3.14.5 Key Relationships

Resonate is a provider of technology to Siebel for the Siebel 99 product. This allows multiple application servers to be deployed and act as one larger virtual server. We believe this relationship portends an aspect of the future for Resonate. As a software vendor, they can easily align with other software vendors to provide more complete and powerful application management solutions.

Resonate has a marketing relationship with Freshwater to leverage Freshwater's external site monitoring capabilities. By configuring scripts on Freshwater to define Commander SendEvents, you can have situations identified by Freshwater feed back into the Central Dispatch and Global Dispatch scheduling rules.

Resonate also has marketing relationships with Sun, Hewlett Packard, BMC, and Tivoli.

3.14.6 Issues

- The requirement for software on every server, both RXP and the agent, makes one think very hard about future Operating systems, hardware platforms, and network technologies. So far, Resonate has done an excellent job of supporting all major OS releases and networks. They need to keep doing that
- The Central Dispatch agent is not optional, and in fact the failure of such an agent can take the host server out of service. Resonate does provide a Master-Child agent capability that enables the Master agent to spawn another agent should the current one fail. Agents should be optional and their failure should not affect host availability
- The 1st release of Commander, while powerful, requires quite a bit of customization to make useful. Resonate should provide more pre-packaged policies in the future
- Commander's model is event in –> action out. The flexibility for decision making in future versions of the product should be enhanced. It should be possible to create action events for combinations of events (correlation) and sequences of events. While this is possible through customized script additions to Commander today, this should be made easier for customers
- Commander provides statistics in a flat file (.CSV). It can be a slow process to generate reports. Also, you have to know beforehand what data you want to store and process. A more flexible database-oriented reporting package would be helpful to make the tremendous range of statistics available from Commander more useful

- Central Dispatch should have more pre-packaged Quality-of-Service and fail-over policy functions. With some integration effort, you can make the product do almost anything you can think of, but it's often not worth the effort except for the larger, more complex sites
- Resonate should look to support more WAN metrics than just round trip delay. Autonomous System hops and dropped packet rate are two effective measures that Resonate does not support today

3.14.7 Final Analysis

Resonate has many unique attributes.

- Commander, via integration with 3rd party agents and knowledge modules, provides greater visibility into end-to-end load and health than any other vendor. The power and flexibility of this capability should only get better with time and more releases
- The unique Connection Hop mechanism enables direct path return for server-to-client traffic, resource-based scheduling, Session ID-based persistence and (in the future) cookie-based persistence
- The distributed software approach, combined with multiple schedulers, provides a high degree of scalability. Each content server can also act as a scheduler which means that both the web server bandwidth and the scheduler bandwidth increase as you add servers
- The distributed software approach also allows you to overlay a VRM solution onto a topology without requiring wiring changes and without introducing boxes which may impact traffic flows if they fail or exhibit bugs
- Global Dispatch persistence mechanisms allow a degree of operational freedom in sites with multiple databases at different sites
- The very software nature of the solution, combined with the capability to flexibly parse packets lends itself to integration with application software suites, as evidenced by the Siebel deal. We expect to see more of such from Resonate

Resonate is a scaleable, flexible, but potentially complex solution. We expect Resonate to deepen its position at the top end of the marketplace for complex business critical web sites, application outsourcing sites, and intranet web farms, to be deployed by people who really know what they are doing.

Getting started in basic configurations with Resonate is pretty simple. But for complex sites, utilizing some of the more sophisticated features of Resonate, they are not cheap or simple to deploy. Distributed software can be difficult to maintain, and there is always the chance there won't be support for your favorite operating system or networking technology in the future.

Go into a Resonate deployment viewing them as a strategic partner, not just a supplier, and make sure both you and they have allocated enough resources to make the deployment successful.

3.15 Others

3.15.1 Introduction and Clarification

The purpose of this section is to introduce and acknowledge some “other” vendors in the VRM world. This is by no means our way of saying that these are minor or lesser players. The main reason for placing a vendor in this category is that they offer load-balancing capabilities as a *part* of their product functionality. Another bit of Acuitive logic is that some of these vendors may be (read: Linux) “do it yourself” solutions. (Please refer to Section 4 “How to Chose a VRM Solution.) Moreover, some of these vendors haven chosen to step back from the true load-balancer arena and focus on other aspects of their products technology (server-agents, OS development, etc.).

All in all, these vendors are active members of the VRM community. Some have built extensive technology relationships with the vendors highlighted in the previous sections. We would be amiss not to identify them as players, and provide a brief synopsis of their capabilities and product offerings. In the event that you require more information than is provided, we have noted the web-presence of each solution.

3.15.2 Allot (www.allot.com)

Allot Communications was founded in December 1996, and has offices in San Jose, California; Tel Aviv, Israel; and Sophia Antipolis, France. The goal of Allot Communications is to deliver Policy-Based Networking solutions.

This solution combines hardware and various software modules to offer a single device solution to various web-centric business challenges. Allot's product line consists of two hardware systems:

- The **AC200** offers two 10/100 Ethernet ports, is geared towards the small to mid size network configurations and supports RIP 1 & 2 and OSPFv2 routing protocols.
- The **AC300** offers the same network ports and routing support, but is geared towards larger network configurations (more memory, faster processor, etc.).

The Load Balancing functionality comes in the form of a software module you can purchase and load onto either of the previously mentioned hardware platforms. The **Load Balancing** module provides the following features:

- Transparently distribute client traffic to Web Server Farms
- Application-based policies allow distribution of traffic according to individual server capabilities
- Assure one hundred percent up time by automatically rerouting down-server traffic to alternate sites

Allot offers a Local Load Balancing solution. Any IP based service can be balanced, which makes this solution viable for Non-server load balancing applications as well. (i.e. Firewall load balancing, router load balancing, etc.).

There are two other software modules that can be added to the AC200 and AC300 hardware platform. The **Bandwidth Accountant** provides insight and control with the following features:

- Policy-based tracking of bandwidth and transactions
- Usage-based accounting, billing, and reporting tools
- Provide service level management guarantees and tiered services

The **Cache Enforcer**, obviously, is Allot's venture into the realm of content caching and redirection. This software module provides the following features:

- Central location for enforcing network caching policies
- Enable transparent caching and reduce network administration
- Define centralized caching policies that efficiently control WEB traffic

If you are currently using the AC product set, and need to consider load balancing, the Allot solution will take care of your local needs. Allots' goal is to provide a product that allows administrators to define, apply and enforce policies throughout the enterprise. Although they are not completely focused on Load Balancing, they do offer a compelling mix of capabilities. By using a combination of capabilities, (i.e. Load Balancing, bandwidth management, cache redirection, monitoring and accounting tools), Allot can help the "needs" of the users coexist the "needs" of the network.

3.15.3 Bright Tiger

Bright Tiger was founded in 1996, and was acquired by the Allaire Corporation early in Q2-1999. Allaire's products ColdFusion and HomeSite are used to build and deploy dynamic web applications. What this acquisition will mean for Bright Tiger's future VRM involvement is anyone's guess.

The Bright Tiger product set is a software solution targeted at managing web resources. They basically load their software onto (web, application, etc.) servers to provide a site-wide view of activity. Their marketing material states that they manage web-sites, not just web-servers (like normal load balancing devices).

They are focused on the local load balancing arena, and specialize in software agents that reside on the content servers. The product **ClusterCATS** (**C**ontent, **A**pplication and **T**ransaction **S**mart) supports the following features:

- Content Distribution and Synchronization
- Centralized Configuration and real-time Monitoring (server statistics, content distribution control, etc.)
- Error correction
- Application and Server Load Management (Load Balancing)
- Resource Monitoring and Fail-over capability
- State-sensitive session maintenance

Obviously, Bright Tiger has provided more than load balancing technology. However, they are difficult to recommend as a VRM solution. Recently, many VRM vendors have formed technology partnerships with Bright Tiger to exploit their site management and server agent technology. Often, a VRM vendor will reference Bright Tiger to their client's for overall resource visibility, while they themselves handle the fine-grained scheduling decisions. This calls into question the Bright Tiger-only VRM solution. As mentioned at the outset, the future availability of Bright Tigers ClusterCATS product and features is still a bit murky, due to the Allaire acquisition.

3.15.4 CheckPoint (www.checkpoint.com)

CheckPoint was founded in and has their International Headquarters in Ramat Gan Israel, with US headquarters based in Redwood City, California.

CheckPoint, provider of well-known security solutions, has entered into VRM arena with the inclusion of a new chunk of software. The Product is called **ConnectControl**, and comes included with Firewall-1 version 3.0 software. This software module is available for FireWall-1 or VPN-1 Gateways, and balances incoming connections among multiple servers. ConnectControl has several scheduling options based various feedback metrics. The product (module) supports a wide variety of Internet services and applications (120 to be exact). Check out their web site for full details.

They provide a local VRM solution only, but it can be used to balance anything with an IP address. This includes other Checkpoint firewalls that have not been equipped with the new VRM software. Not only will servers benefit from this feature, but also the entire sites' security can now be considered highly available.

CheckPoint may be a good choice to fill your local VRM needs if you already own their software. In fact, CheckPoint's entry into the VRM market has a very direct focus. Once the enterprise is secure, you need to address the issues of site scalability and resource optimization. CheckPoint figures that as long as they are protecting your site, why not direct traffic as well?

However, one issue you may run into is platform performance. CheckPoint states that one ConnectControl Product Benefit is the ability to "integrate traffic load balancing with enterprise-wide security policies." While there are several feedback options and scheduling rules that can be set, remember the processing/buffering loads associated with encryption, filtering, and policy enforcement. These inherent security features are enough to bury your platform as traffic increases at your site. With the additional burden of VRM scheduler software on the box, the system may not be able to deal with these new demands.

3.15.5 CinTel (www.cintel.co.kr)

CinTel was founded in 1997 and is a Korean based company. They have heavily invested in research and development to satisfy client needs in the VRM environment, and their goal is to apply various features (Bandwidth Management, etc.) as the market demands. They are working on various co-marketing schemes with US networking firms to attain a presence in the United States.

Their product is a hardware appliance called **PacketCruz** (formerly known as the NetCruz). Along with numerous other features, it offers both local and multi-site Load Balancing capabilities. The product has various modes of operation, depending on which software you load onto it. These modes of operation are:

- Redirector – load balances incoming IP traffic via NAT, provides server failure detection (pings, etc.) and basic content distribution
- Network Cache – provides transparent cache services
- IP Trapper – captures IP traffic, allows for some firewall-like functionality through filtering
- Cache – accelerates HTTP traffic (see their web site for all the details on this one)
- Web Server – just what it sounds like, but pages are stored in the device's memory

This vendor seems to care about VRM capabilities. However, we question their ability to focus on this arena.

Obviously, anyone thinking of contacting CinTel should keep in mind the Operations and Support issues that may arise. To date, we have seen no evidence that their co-marketing plan with US companies has provided exposure or an installed base. If you have seen or heard from them in a competitive situation here is the States, we'd love to hear about it.

3.15.6 Eddie (www.eddieware.org)

Here's an option for all of you die-hard, code-loving do-it-yourself individuals. Eddie is an Open Source project (set of applications) that is aimed at building highly available server clusters. It is based on the Erlang functional language, developed and recently released as Open Source by Ericsson. The basic components (applications) are as follows:

- **Enhanced DNS Server** - provides load balancing and monitoring of site accessibility for geographically distributed web sites.
- **Intelligent HTTP Gateway** - provides site based (local) Load Balancing and Quality of Service capabilities.

These components have been released as Linux beta packages (Red Hat 5.0 hosts) and Solaris 2.6 packages. As you can imagine, there are numerous bulletin boards, mailing lists, news services and development resources available at the Eddie web site. If you want to help in the further development of this technology, there is a "To-Do" list published on the site as well. (There is also an explanation of the name "Eddie.")

We don't know of any companies who have chosen Eddie as a strategic part of their IP services plan, but this is an option. As tempting as this may be to various free-spirits in the internetworking world, please keep in mind the potential Operations and Support costs when evaluating this solution.

Peace.

3.15.7 Microsoft (www.microsoft.com)

Microsoft, a company that needs no introduction, is also dabbling in the VRM space. The recently purchased a company called Valence, took their product and integrated it into Windows NT as the Windows Load Balancing Service (WLBS).

WLBS is a software solution that provides local VRM for Windows NT servers. We're not sure how its delivered, exactly. It's probably an optional software component for Windows NT 4.0 today, and will probably become part of Windows 2000. This software, which occupies about 1.2MB of disk space and 1-2 MB of memory, runs on each server that is part of the cluster. The software installs as a driver that sits between the hardware (NIC) driver and the servers' TCP/IP stack. Each server of the cluster is configured with the same IP address. User requests are multicast to the entire cluster and one of the servers in the cluster will respond and handle the user's session, a technique we describe as MAM. If a server fails, a heartbeat protocol between the servers ensures that the other servers pick up the failed server's load. The use of MAM plus the heartbeat protocol means the WLBS solution supports the Active-Active mode of redundancy.

Feedback from users that have used the product indicates a lack of scalability beyond a few servers, because of the CPU load placed on those servers by the request distribution algorithm. Its use of the multicast (MAM) protocol to distribute user requests on a LAN requires every server on that LAN to listen for requests/data that might be destined for it. This requires the server to spend CPU cycles listening to traffic that it will not need to process. If there are ten servers in a cluster, on average, 90% of the requests listened to by a server will be ignored.

Microsoft has essentially made small NT we server clusters "free." If they really worked hard at it, they could fix the technical issues with WLBS, completely integrate Back Office, their application services, IIS and their Wolfpack clustering solution. An open API between all components would possibly encourage ISVs to write software to take advantage of the entire application-smart clustering solution. This would pose a challenge to some of the VRM vendors that support more complex application environments.

Right now, though, we're not high on WLBS for anything other than the simplest of NT web server clusters (like a cluster of two).

Due to the size/strength of Microsoft, we decided to include these developments in the "Others" category. If MS decides to run after this market with their traditional fervor, they could grow into a key technology player. However, we don't predict that MS will put the VRM companies mentioned in this document out of business any time soon.

3.15.8 netSTOR (www.netstor.com)

netSTOR Technologies was founded in 1998, and is a Canadian company based in Toronto, Ontario. The company's goal is to provide solutions for IP and Internet performance and reliability issues.

The VRM product, called **IP/EDGE** is a Software-on-a-Stick appliance. This is a local load balancing solution, marketed as an "elegant, intelligent load balancer." It will direct traffic for any TCP/IP based service and offers server (iCACHE servers included) clustering for scaleable, reliable web sites.

They currently offer one other product, called the **iCACHE**, which is a fairly standard caching product.

This company is a newcomer to the VRM field. We recommend a call to headquarters to gather more information regarding the netSTOR support options before cutting a PO.

3.15.9 WebManage (www.webmanage.com)

WebManage was founded in 1997 and is based in New Hampshire, US. They are intent on providing a complete site-management and intelligent load balancing solution.

Their product is called **SiteMARC** (**M**anageable, **A**vailable, **R**eliable/**R**esponsive and **C**ontrollable), and is a software solution to whatever ails your e-commerce site. Their software works with NT and (various flavors) of UNIX.

They have attempted to go beyond earlier load balancing schemes by integrating qualitative scheduling rules into their scheduling software. That means that WebManage looks at the actual request (HTTP, SSL, FTP, etc.) and applies various metrics (customer "class," transaction nature, domain of origin, content requested) to determine how to make a load balancing decision. They also allow basic error correction (redirecting 404 "File not found" to another resource without the client's knowledge of the transfer, etc)

and include some impressive security features (single sign-on/multiple server support, scheduled site integrity checks) to round out their quiver of capabilities.

Other features include predictive server rating, reporting and management tools and the ability to apply various type-of-service/QoS-like ratings to users and domains.

WebManage seems to offer a lot of functionality from their software, and all in all they are worthy of further investigation before you select a VRM vendor. One point of caution stems from their software solution (this is the same concern that any software vendor faces). Once WebManage is installed into your site, you must buy into their solution for every server you wish to manage and control. Check into their support options. Check into their feature release plans and track record. Let us know if you select them as your VRM vendor, because we'd like more information about their solutions and ability to deliver.

3.15.10 WebSpective (www.webspective.com)

WebSpective was founded in 1997 and is headquartered in Cambridge, MA. WebSpective is focused on site management and operations, with an emphasis on content management. As a result, their solutions provide a good view of the resources on each server and statistics that can be used for capacity planning and other operational activities.

Their product is a software suite that consists of 6 integrated modules. The VRM relevant module is called the **Traffic Manager** and is a local load balancing solution. Traffic Manager provides load balancing across distributed Web Servers, based on knowledge of the available content on each server and its health/load. It bases these decisions on HTTP requests and static or near-real-time application/server/site information. Working in conjunction with the Traffic Manager is a set of Agents and Filters, which are software modules installed directly onto the content servers. The Agent collects statistics on each server (CPU load, Cache utilization, queue lengths, application performance) and sends this information the Traffic Manager. The Filter module specifically collects HTTP "hits" and associated protocol performance data.

The other five modules offer a wide range of other features:

- **Content Distributor** provides a utility for distributing and synchronizing content across live web servers. Content Distributor and Traffic Manager are integrated in that as content is updated on a site, Traffic Manager can tell which site users are connected to existing content and will allow them to finish using it. New users are routed to an updated version on another server until the content can be updated. It can also "roll-back" to an earlier content version if there are distribution problems or other software failures
- **Crisis Manager** resolves failure situations by restarting failed Web Servers and informing Traffic Manager to route traffic around them while they are down. This can include starting up additional copies of applications on other servers so that service does not suffer
- **Performance Monitor** monitors traffic, load, and performance data for all Web applications, web servers, host machines, and network interfaces
- **SiteStats Manager** collects all the statistics and creates site level reports
- **Site Administrator** is a management tool for configuring all the WebSpective features

WebSpective offers a powerful array of site management capabilities. VRM technology is just a piece of this product set. Before selecting WebSpective as your load balancing vendor of choice,

review your operations and support responsibilities and your sites requirements. You may not require everything they offer. Also, you must buy into the entire WebSpective solution (for every part of your site strategy) in order to enjoy the benefits of their VRM/site management offering.

3.15.11 Xedia (www.xedia.com)

Xedia was founded in 1993 and is based in Massachusetts, US. Their mission in life is to build high-performance IP routers for both WAN access points and LAN environments.

The LAN access products offer some local load balancing capabilities. Specifically, the **AP 100 QVPN Gateway** supports "IP Load Sharing." This simple scheme uses typical remote probing and basic metrics (round robin, least connections) to determine how to schedule a request.

There are two products in this family. These devices enable secure, reliable migration of mission-critical business applications to the Internet. They are focused on LAN, site-to-site private WAN and Remote Access requirements. Here are a few more details regarding the AP 100 and it's little brother, the AP 10:

- **AP 10 QVPN Gateway** – provides device security, policy based bandwidth management and traffic monitoring and accounting, IPSec and L2TP support for site-to-site and remote access requirements, hardware assisted encryption, OSPF, RIP 1 & 2 and Static Routing support, BGPv4 with full peering and NAT
- **AP 100 QVPN Gateway** – offers all of the AP 10's feature with the addition of two important capabilities. VRRP (Virtual Router Redundancy Protocol) support and "**IP Load Sharing**."

The WAN products (AccesssPoint T1/E1, AP 45 and the AP ATM QoS) do not provide any load balancing support. They are focused on applying policy-based control and accounting to Internet access points.

This load balancing feature is a second-order function. While Xedia does offer basic, local load balancing, the concentrate mostly on bandwidth management, VPN technologies and accounting functionality.

4 How To Choose A VRM Solution

The simple answer is – call Acuitive.

But if you are a do-it-yourself-er, here is some insight into how we do it.

First of all, we characterize the nature of the application. There are various ways to do this, but some of the key parameters to nail down are:

- Single site or multi-site? (Now and in the foreseeable future).
- What are the application requirements for state management? This is a huge issue and will determine the type of ***persistence policies and mechanisms*** you need in your solution.
- Scale expected, now and in two years. Hits per second. Number of simultaneous users. Bandwidth. This can be estimated either from usage projections or by projecting the maximum capacity of key elements of the system – WAN link size, database server capacity, application server capacity, and then mapping that back into the performance parameters.
- Application architecture. 2-Tier? N-tier? Migrating over time? This has a big impact on the way you approach monitoring resource health and load.

For a final design you'll need a lot more information than this. But the above is a good foundation to work from. It helps narrow the list of alternatives. You can then use more detailed information to make good choices within the narrower set of alternatives.

Next, you need to think about the derived requirements for the VRM solution that you will choose. To do this, we break the considerations into the following areas:

- Performance/Scalability
- Redundancy
- Operations and Management
- Functional Characteristics
  - Feedback (from servers and sites to monitor health and load)
  - Policies (to determine the "best available server," given information about health and load)
  - Mechanisms (to re-direct traffic to the "best available server")

The number of options in each of these areas is huge. That's why we have published a research report that provides a tutorial and analysis of all the various options. The right choices depend heavily on the nature of the application requirements. But there are some rules of thumb that can be applied to get in the right ballpark, that you can then tune by overlaying more features and considerations.

Different applications drive different VRM requirements. There are a million permutations of application environments. But when selecting a VRM solution, some things make sense no matter what the application.

Performance and Scalability

Unless you know that your application is going to have modest performance requirements *and always will*, then we recommend going with solutions that will never be a bottleneck for either processing incoming requests (measured in connections/sec) or flowing return traffic back to the clients (measured in Mbits/sec).

The first of these generally requires a method for multiple schedulers to be able to support the same service, using a common set of content servers. A single scheduler is *always* at risk of becoming the bottleneck, no matter how powerful the underlying CPU.

The second of these either requires a scheduler that can efficiently process packets in the return path from the server-to-the-client, or an approach where such traffic bypasses the scheduler entirely.

Redundancy

VRM is primarily about reliability. So the last thing you want is for your scheduler to be a reliability risk.

All the vendors have viable redundancy schemes. This is usually not the consideration that drives you towards one solution or another. Having said that, we do have preferences for redundancy schemes which allow us to:

- Actively use the multiple schedulers when all are healthy, which then take on more responsibility for more clients as schedulers fail (this is related t the performance section above)
- Physically locate back-up schedulers in different locations of a campus – in different Data Centers, on different power grids, etc.
- Maintain the state of existing persistent connections when fail-over occurs.

Operations and Management

People talk about ease-of-use all the time, and point to GUIs as examples of how to achieve that. We're not so into GUIs in and of themselves, and we don't really care if initial set-up is a little difficult, because you only do that once. What we really care about is ease-of-on-going-operations. To that end, some key features we look for are:

- Feedback on site performance from the perspective of the operations performed by an end user. Good examples of this are Site Availability and Application Response Time. Whatever the parameters, we want to know what it is now, what is was in the past, and where it is trending
- Access to data on the underlying resource loads that may be correlated to the site level performance. Ideally, we'd like the management system to point us to bottlenecks and provide guidance on how to add new resources, weight servers, modify policies, or otherwise change operations to improve performance.
- The ability to play "what-if" scenarios. Load patterns can change quickly. Traditional trending techniques do not necessarily do a good job of predicting future loads and behaviors. We'd like to be able to simulate a variety of possible loads on a system to evaluate what the critical bottlenecks may be in the future.
- Secure management interfaces are critical.

Functional Characteristics

The list of possible Feedback, Policy and Mechanism features is enormous. Our over-riding advice is "don't go nuts." Many of the techniques are unproven, can cause harm, or don't provide enough value to justify the effort to implement and maintain them.

Feedback on resource health and load is critical. We believe that all VRM deployments should use Active Content Verification (ACV) with response time monitoring as a core feedback technique.

- Within a site, other types of server-resident feedback can be overlaid on top of that core technique to create some threshold-based safeguards.
- For dynamic application environments, Dynamic Application Verification (DAV) should be performed in the background on a regular basis.
- For multi-site scenarios, ACV is generally good enough.

For site level Mechanisms, we believe that you should almost always choose a vendor that provides you the opportunity of doing delayed binding. You may not need it now, but in the future it's a good bet that some aspect of state management or site architecture will require it.

For multi-site scenarios, there is no perfect Mechanism. The ultimate solution probably has not been invented yet. Given that, we like to use DNS Re-Direction as a core technique, often supplemented by another re-direction technique (such as HTTP Re-Direct and/or or Triangulation) to deal with corner cases and short-term issues that DNS Re-Direction does not adequately address.

Policies are a complex subject. We do believe, once again, that simpler is better. Don't try to put in too much complexity to try and squeeze another 10% of capacity out of your site by performing more balanced load balancing. You'll find that you either don't achieve the goal or it wasn't worth the effort. Weighted Round Robin or Least Connections are usually the best policy choices for the "best server" choice locally, especially if supplemented with a safeguard mechanism or two related to thresholds – either thresholds on connections or thresholds on key server load parameters (e.g. CPU utilization or available memory).

Some simple QoS policies related to moving resources between applications during times of stress of a high priority application make sense to us. But only rely on them for dealing with short term issues that couldn't have been anticipated. Longer term, the best thing to do is design the site so that resources are not saturated with peak loads (sometimes that is not economically viable, but it should be the design target).

In the WAN, as mentioned above, we believe that the best policy is "use a site in your region unless it's on fire." Static user mapping, augmented by some techniques to automatically populate the static tables, is a reasonable starting point for multi-site VRM, and one that many people will never have to venture beyond. Trying to track ever-changing delays or packet drop rates or stuff like that dynamically is generally futile and can lead to more harm than good.

Overlaying all of this, make sure you have the right persistence policies. Persistence **always** takes precedence over better load balancing, because persistence determines whether the user can effectively interact with the application, while load balancing just provides incremental performance improvements. Again, persistency should be approached as simply as possible. Unless you require persistency that spans long periods of inactivity, the simplest and most general approach is to use Source IP Binding with Address Range Mapping. In some circumstances, adding SSL Session ID Tracking and/or user cookie tracking might make sense.

5 Vendor Selection Guidelines

This final portion of the report provides some guidelines for thinking about your application and what key the requirements are. We end each application section with a brief treatment of what vendors can satisfy the key requirements. The overall intent is to help characterize your circumstances and provide a "short list" of Vendor solutions to meet your requirements.

There are many criteria that are normally used to pick a particular product or service, such as performance, price, ease of use, operational characteristics, available support, size and stability of the organization. Some of these criteria are subjective, and are simply impossible to evaluate. Some of these criteria will be your first, or fourth concern, and we do not want to presume to judge what's best for you.

Take the following into account when choosing your *own* short list of vendors to evaluate:

- Determine which of our Application Descriptions best characterizes your application(s). If you fit into more than one category, combine the requirements together (i.e. you have a harder environment to support)
- Examine the list of Vendors that we recommend at the end of the application requirements description.
- Rank each of those vendors based on the following criteria, and their relative importance to you:
    - How much does vendor experience and stability affect your decision? Are you ready to trust a vendor that's been shipping product for 6 months to get the features you need, or do you want a vendor that's been in the business for awhile?
    - Is VRM technology the vendor's core business?
    - Do any of the vendors have reference accounts that you can talk to that closely match your application environment?
    - Do you need other functions/features that a vendor provides that you otherwise would purchase a separate system for?
    - What will be the purchase, install, and support cost of the product with full redundancy and the best support contract?
    - What are your application's performance characteristics, and how do they match with vendor-neutral lab tests? How does the vendor's architecture lend itself to matching you application's requirements and scaling to what you need in the future?
    - How will you buy the products? Through a reseller or VAR, or direct from the vendor?
- Evaluate your short list of vendors either by attrition (there's only one left) or by testing their products in a live environment.

One overriding principal: **We believe in simplicity.** Experience has taught us that the most simple solutions generally work, and complex solutions often have unanticipated consequences. So we'll admit right now that these guidelines to not address nuances, second order issues, and subtleties of a lot of applications and VRM systems. So be it. Following these guidelines will get you into the right ballpark. Getting to the right seat from there takes a bit of work.

You should also know that while we have some level of exposure to each of the vendor's products, we have not formally or informally tested any of them in every application environment. That job falls to you. Make sure you test any vendor you're considering in the application environment that you plan to build, with all VRM functions and background tasks active. Make sure you test all aspects of any

vendor's solution, especially ones that could affect reliability, such as redundancy and server failure detection.

In one of our first updates to this report, we will provide a detailed feature matrix on each vendor's products, which we will use to cross-reference into a more detailed short list of vendors to evaluate based on certain application requirements.

5.1 Characterizing Your Application

The first thing to do is to characterize the nature of your application at a high level. If you can define your application to be in one of the following categories, you can pretty quickly determine the "short list" of vendor/product candidates.

- **Local VRM, Static Content:** Sites that are dominated by HTTP requests for [.html] files of modest size. This is typical of informational-oriented company.com and web sites, publishing sites, and many Intranet sites.
- **Local VRM, Downloaded Content:** This is a special case where the vast majority of content downloaded is large files. This is typical of software distribution sites, document distribution sites, and many multi-media source sites.
- **Local VRM, 1-Way Dynamic Application Environment**: Sites where the user interacts with the site for a short period of time, modifying information during the period of the session, but the modifications are "forgotten" after the session is complete. Examples are search engines, comparison shopping sites, or business intelligence reporting front-ends.
- **Local VRM, 2-Way Dynamic Application Environment**: Sites where there is a two-way interaction between the user and the site, resulting in a change to the information stored at the site when the user session is complete. Examples include almost all kinds of E-commerce sites and e-mail sites.
- **Local VRM, Server Optimization**: This is an overlay requirement which could be added to any one of the applications above, where the content or application deployment behind a common domain name (or V_IP) needs to be further subdivided so that servers can be optimized for specific roles (applications) and/or content is divided up into more manageable domains.
- **Multi-Site VRM, Site Redundancy:** This is the situation where applications and content are deployed at a small number of locations, often 2, primarily to ensure application availability, but also to scale the application bandwidth somewhat.
- **Multi-Site VRM, Global Content Distribution:** This is the situation where applications and content are deployed at multiple locations, often throughout the world, to reduce latencies for content access, reduce WAN costs, and to scale content availability.

In the following sections, we identify the key features to look for in a VRM solution. We then identify the short list of vendors and products that best suit each application.

Your site might support more than one of these applications. If so, you can combine the requirements and required VRM attributes for each application to create an overall requirement. The short list of vendor choices then becomes the *intersection* of the list for each of the separate applications.

5.1.1 Local VRM, Static Content

The simplest case is one that is dominated by HTTP traffic, for accessing relatively static files. All informational-oriented web sites are of this type.

Application Characteristics

- The application is dominated by HTTP and does not record any user state information between individual user requests
- Primary requirement is to provide server redundancy. For high load sites, an additional requirement is to scale the site.
- TCP connections are fairly short. Files retrieved per TCP connection range between 4 KB and 100KB.
- Traffic flow is asymmetric, with server-to-client traffic being roughly 5-10x greater than client-to-server traffic.
- It is a non-tiered application architecture (applications run directly on the servers being load balanced)
- The system for the most part is distributing common and relatively static information to the users
- The information to be distributed is either stored locally on each server, or is shared with a file server
- Servers are local, co-located to the VRM
- Example applications are publishing sites, a lot of corporate.com informational sites, and departmental web sites
- The application flow basically consists of:
  1. client opens TCP connection
  2. user request
  3. server index response
  4. client opens TCP connections to request all objects indicated in index
  5. server responses
  6. closure of all TCP connections

VRM Scheduler Requirements

This is the simplest solution to design for, with many vendor solutions to choose from. This is the application that most low-end VRM solutions are targeted at. Pick a VRM product that matches your needs, while requiring as little environment modifications as possible. You may not even need a specific VRM product at all. Your needs may be met by features available in your server Operating System, application software suite, or multi-purpose IP appliance (e.g. firewall, bandwidth manager, intrusion detector).

Don't get mystified by all the bells and whistles that a vendor will try to sell you on. If you don't need a feature now and can't anticipate needing it any time soon, don't factor it into your decision-making. Keep it simple. Don't get caught up in the technology for technology's sake; make sure it serves a real purpose that you understand and can easily characterize.

Feedback: requires at least some form of TCP connection verification. In addition, ACV is highly desirable to ensure that the HTTP processes and storage subsystems are working properly.

Policies: Round Robin is an acceptable load balancing policy. Weighted Round Robin is recommended if there is a wide variance in server capacity. MaxConns as a 2nd order threshold policy is good to make sure that any particular server does not get overloaded

Least Connections is also very suitable, and probably better than Round Robin, because of the short duration of the TCP connections. Again, MaxConns as a 2nd order threshold policy would be useful. Response time as a secondary policy might be useful if server load fluctuates due to data rate variance.

The only persistency you have to worry about here is ensuring that all packets associated with each TCP connection are directed to the same server. But that's basic VRM. All vendors support that capability (or else their products would be broken). *If you have a small number of content servers and are worried about large groups of users coming from behind a common address-based proxy firewall*, then consider binding on Source Port number or Source Identifier instead of just the Source IP address.

If your site supports multiple different services associated with different V_IPs, and some are deemed more important than others, then some sort of simple Preferential Services policies indexed to Application Response Time would be useful. This protects you against changing load patterns that cannot be anticipated.

Mechanisms: Delayed binding is generally not required for this application. You can use any immediate binding mechanism that the vendor supports. Half NAT and MAT are the most likely candidates.

Performance: Focus on TCP connection binding rate when selecting your VRM system. The connection rate load offered to the system will be limited by the WAN access link speed, but will be a high number relative to that speed, due to the short duration of connections. Since this application does not require delayed binding or very many resource-intensive features, all vendors' products support multiple T1 access, and most have products that can support DS-3 speeds (1600 connections per second) or a little higher.

If you are a heavily accessed publishing site, or for other reasons have or plan to have OC-3 or higher access to your site, then you have to worry a bit more about scalability. Look for a solution that can support 5,000 – 10,000 connections per second (immediate binding) either in a single scheduler or via an Active-Active multiple scheduler redundancy method.

Redundancy: Hot Standby is fine for this application.

Prospective Vendor Choice

Since this application is very common and very simple, most vendors have products to fulfill the requirements for it. Check first to make sure you even need a VRM product. If you have a small homogenous environment your OS or application suite may have the features you need already built in.

(not so) Short List
Everyone

These requirements are easy to fulfill. Optimize your choice for simplicity, price, how well the product fits in your environment and manner of operations.

This *can* be an application that requires high performance (greater than 100 Mbps). Since HTTP sessions are generally short, the TCP connection binding rate is a key specification to evaluate.

Alteon, ArrowPoint, F5 Networks and **Foundry** have Gigabit Ethernet and Fast Etherchannel interfaces *and* the connection rate handling performance to at least partially fill them.

IPivot and **Resonate** do not presently support interfaces greater than 100 Mbps, but with the combination of their Active-Active redundancy schemes (to increase scheduler bandwidth) and direct return-path features, you can build topologies to support greater than 100 Mbps.

Cisco, IBM and Radware presently support multiple 100 Mbps interfaces, but do not support Active-Active redundancy schemes. However, they do support direct-path return features. In all three cases, if you split the web servers up so that they are spread across different 100 Mbps networks, you can build systems that support greater than 100 Mbps. In IBM's case, the ability to achieve such a level of performance depends on the power of the underlying host system chosen.

5.1.2 Local VRM, Downloaded Content

Another common case is one that is dominated by FTP or HTTP traffic for downloading large files, often software images of multiple megabytes.

Application Characteristics

- HTTP and FTP dominate the application.
- For FTP, persistence is required to make sure the control channel and data channels are bound to the same content server.
- Client TCP connections last from minutes to hours, depending on client access speed.
- Highly asymmetric traffic flows. Little traffic from the user to the server (small ACK packets). Many packets from the server to the user. The asymmetry can be as much as 20-40x.
- The information to be distributed is either stored locally on each server, or is generally shared with via a back-end file server or database-oriented file system.
- Servers are local, co-located to the VRM
- Example applications are shareware and retail software sites, software company customer service sites, and document management sites (such as the IETF Standards documents site).
- The application flow basically consists of:
  1. client TCP connection
  2. user index request (small)
  3. server response
  4. closure of TCP connection
  5. client TCP connection
  6. user file request (large)
  7. server response
  8. closure of TCP connection

VRM Scheduler Requirements

Feedback: The best form of feedback for this scenario is DAV, where a download request is made for a small file and the VRM system monitors the ability of the system to deliver that file. Resource resident monitoring and/or application response time monitoring (coupled with appropriate policies) may be useful if there is a wide variance in server load due to data rate fluctuations.

More for planning purposes than real-time operation, a tool for measuring and monitoring successful downloads is desirable.

Policies: The best policy here is a packet rate policy – the number of packets sourced by each server in the server-to-client path. A good secondary/threshold policy can be LAN utilization. If multiple servers are connected on the same LAN, their aggregate load can congest the attached LAN, in which case servers attached to an uncongested LAN should be preferred.

Least Connections or Round Robin Weighting don't work very well in this scenario because there is just one connection per download but each download may have a widely varying impact on server utilization, depending on the speed of connection of the client.

One type of persistency you have to worry about here is ensuring that all packets associated with each TCP connection are directed to the same server. But that's basic VRM. All vendors support that capability (or else their products would be broken).

Another form of persistency you need to worry about is download re-start. Sometimes in a long download the connection can terminate for reasons having nothing to do with the health of the server or the VRM system. When that occurs, and the client re-connects to re-establish the connection, you'll want to make sure the connection goes to the same server the client was initially attached to if the download application software you have supports some form of re-start (picks up the download where it left off). Thus you may want to bind to the Source IP address. To protect against changing Source IP addresses from clients behind proxy firewalls, you'll also need the Address Range Mapping value-added option. Response time as a secondary policy might be useful if server load fluctuates due to data rate variance.

Mechanisms: Delayed binding is generally not required for this application. You can use any immediate binding mechanism that the vendor supports. Half NAT and MAT are the most likely candidates. If the site uses FTP as a mechanisms for downloads, the VRM system must support Active FTP, ensuring that the control channels and data channels are bound to the same server.

Since most of the traffic will be from the server-to-the-client, MAT with direct-path-return might be slightly preferred here offload the VRM scheduler from any packet processing responsibility in the return path. But if the VRM devices is a high-speed switch or has lot's of excess capacity, who cares? Use the VRM mechanism that requires the least modification to the environment (network, servers). The most likely mechanism choices include Half NAT and direct server-to-client return.

Performance: Connection rate will be low. Server-to-client data rate will be high (limited by the WAN access link speed). This is the one case where vendor's throughput specifications are more useful than the connection rate specification. Almost all VRM products on the market can support the connection rates associated with this type of environment. Depending on the WAN bandwidth, they can't all necessarily handle the packet-processing rate required in the server-to-client path. If your WAN speeds are greater than DS-3, look carefully at this and either choose a high-speed switch or a VRM scheduler that allows for direct-path-return.

Redundancy: Due to the low connection rates, Hot Standby is good enough. Requires Session Assurance. After 2 hours of a download you don't want to have to start all over again.

Prospective Vendor Choices

Short List
Alteon
ArrowPoint
Cisco
F5 Networks
Holontech
HydraWeb
IBM
RadWare
Resonate

No one really has optimized their product or feature set for this particular application. However, all of the short list vendors on this list have the ability to monitor server load in a manner relevant to this application. **ArrowPoint**, **RadWare** and **HolonTech** have (perhaps) the best direct measure of server load for this application – *packet rate*. However, **Resonate** tracks CPU utilization in their agent, and the other vendors on the short list allow CPU utilization and other server-side parameters to be monitored via an API. All of these techniques work.

In this application, the scheduling throughput of the vendors' products is rarely an issue – connections/sec is very low. The only issue for high performance requirements is whether a topology that can support high-speed return-path data flow is supported. **Alteon**, **ArrowPoint**, and **F5 Networks** support high speed flows natively in their devices. **Cisco, IBM, RadWare** and **Resonate** all support a direct server-to-client return option to enhance throughput back to the client. Your topology must allow for this mode of operation.

5.1.3 Local VRM, 1-Way Dynamic Application Environment

This application environment has some amount of dynamic-ness, such as data lookup, calculation, or performing a search. The user must stay connected to the entity that is aware of the calculation for the duration of the session. But after the user is done, the site does not need to modify any information or "remember" the results of the action.

Application Characteristics

- The application uses a simple TCP or UDP protocol, but it records some state information about the user-sessions locally on the server. It may be a non-tiered application architecture (applications run directly on the servers being load balanced), or it might have access to a lightweight back-end database.
- User sessions are persistent for short duration while the recorded state information is not stale (minutes). Individual TCP connections will be short-lived (seconds to minute).
- Example applications that match this environment are a web-enabled lookup or reporting application, or a web-based search and display system
- The application flow basically consists of:
  1. new client TCP connection
  2. user dynamic request
  3. server performs lookup, records user/client state locally
  4. server response
  5. closure of TCP session

loop

6. same client connects via TCP
7. user dynamic request, based on previous state
8. server performs lookup, based on previous state
9. server response
10. closure of TCP connection

loop until application session completes

maintain binding until inactivity timer expires user state

VRM Scheduler Requirements

Feedback: Requires ACV. Enhanced by DAV. Resource resident monitoring and/or application response time monitoring (coupled with appropriate policies) may be useful if there is a wide variance in server load due to application request variance.

Policies: Least Connections is acceptable. Application Response Time or server resource as secondary/threshold policies might be required if server load fluctuates due to application request variance (which is generally true).

Persistency is required. But the persistency policy can be simple -- IP Source-based binding, with sticky timer time-out, is OK. If users are coming from behind address-based proxy firewalls, Address Range mapping as a value-added option is required.

Mechanisms: Delayed binding is generally not required for this application. You can use any immediate binding mechanism that the vendor supports. Half NAT and MAT are the most likely candidates.

Performance: This is generally a connection rate-limited application**.** Focus on TCP connection binding rate when selecting your VRM system. The connection rate load offered to the system will be limited by the WAN access link speed, but will be a high number relative to that speed, due to the short duration of connections. Since this application does not require delayed binding or very many resource-intensive features, all vendors' products support multiple T1 access, and most have products that can support DS-3 speeds (1600 connections per second) or a little higher.

If you have or plan to have OC-3 or higher access to your site, then you have to worry a bit more about scalability. Look for a solution that can support 5,000 – 10,000 connections per second (immediate binding) either in a single scheduler or via an Active-Active multiple scheduler redundancy scheme.

Redundancy: Requires Hot Standby. Ideally, the stickiness "state" should be shared with the back-up scheduler so that Persistent Session Assurance can be maintained upon scheduler fail-over.

Active-Active should be explored if the performance requirements demand it.

Prospective Vendor Choices

Short List
Alteon
ArrowPoint
Cisco
F5 Networks
Foundry Networks
Holontech
HydraWeb
IBM
IPivot
RadWare
Resonate

Each of the vendors on the short list support the general key requirement: Active Content Verification.

Some additional factors consider:

ArrowPoint, F5 Networks and IPivot provide traffic management features that provide some user groups or applications better service than others, if site overload is approached.

Cisco, F5 Networks, and Foundry support Persistent Session Assurance. This feature helps assure smooth site operation even if a scheduler should fail.

Holontech, HydraWeb, IBM, and Resonate support feedback from server co-resident agents to help balance the load. Other vendors (e.g. F5 Network, RadWare, and Alteon) support this capability via APIs, which is generally better for identifying failures than for real-time load balancing. However, we feel that ACV with response time monitoring is generally just as effective.

Alteon, F5 Networks, Holontech and IBM support various forms of Dynamic Application Verification, which can be used to test the health of key dynamic applications running at the site.

This *can* be an application that requires high performance (greater than 100 Mbps). Since HTTP sessions are generally short, the TCP connection binding rate is a key specification to evaluate.

Alteon, ArrowPoint, F5 Networks and **Foundry** have Gigabit Ethernet and Fast Etherchannel interfaces *and* the connection rate handling performance to at least partially fill them.

IPivot and **Resonate** do not presently support interfaces greater than 100 Mbps, but with the combination of their Active-Active redundancy schemes (to increase scheduler bandwidth) and direct return-path features, you can build topologies to support greater than 100 Mbps.

Cisco, IBM and Radware presently support multiple 100 Mbps interfaces, but do not support Active-Active redundancy schemes. However, they do support direct-path return features. In all three cases, if you split the web servers up so that they are spread across different 100 Mbps networks, you can build systems that support greater than 100 Mbps. In IBM's case, the ability to achieve such a level of performance depends on the power of the underlying host system chosen.

5.1.4 Local VRM, 2-Way Dynamic Application Environment

This application environment builds on the one above, and has a large amount of two-way data interaction between the user and the server. A user will regularly be pushing new or updated data towards the server, where the server will store it for a period of time, and possibly push off to a central data-store. It has the following different or additional characteristics:

Application Characteristics

- The application may use either a simple TCP protocol, or a complex set of parent-child or control-data protocols. HTTP and SSL are both often used in these application environments.
- Significant state information about the user-sessions and user-data is directed at each server, but then (probably) sent to a central datastore
- It uses a tiered application architecture, with the users interacting with an application server, which in turn interacts with a database and/or fileserver. The tiered system may be extended further, such that the users interacts with an application server, which in turn interacts with a compute server, which in turn interacts with a database.
- User sessions are persistent for a long duration (minutes to many minutes, hours). TCP sessions may be long (minutes, hours) or short (seconds, minutes) depending on the application.
- The topology is somewhat complex, with a front-end network for the VRM system and application servers, as well as a back-end network for the database and fileservers.
- Client/Server traffic won't be as asymmetric as static sites or download sites traffic patterns. Traffic direction and data rate will widely vary based on user and user action at any given time.
- Example applications that match this environment include web-based email, client/server-based email, a transaction oriented application, such as an order management system, and a browse and transact system, like a "classic" consumer E-commerce site.
- The application flow basically consists of:
  1. client TCP session connection
  2. user identification and authentication
  3. server records state about user's session

  Loop
  4. client TCP connection
  5. user request/post
  6. server lookup
  7. server response / accept
  8. data update
  9. closure of TCP session

  Loop until inactivity timer expires user state
  10. Server may need to sync with other servers
  11. Server forgets user state

VRM Scheduler Requirements

This is the environment where you generally can't allow a lot of corner cases to exist. Simplicity isn't as much of an option here as in most other cases. You'll have to pull a lot of tricks out of your bag and your VRM vendors' bag. This is the application that most high end VRM features are targeted at.

Feedback: Requires ACV and DAV, with resource-based monitoring and feedback. Sampled PCV is also a useful value-added feature.

Benefited by integration with 3rd party instrumentation at all levels of the tiered application with a rich set of correlation rules for taking action or aiding troubleshooting based upon the set of feedback received.

The management system should provide feedback relevant to the user experience such as Application Response Time, % reloads, average session duration, etc. This information isn't necessarily all for real-time site operation, but for longer term planning.

Policies: Least connections can be used as a core load balancing policy, but it must be augmented with threshold and Preferential Service secondary and tertiary policies. Some candidates:

- Server-resident resource metrics, such as CPU utilization, available memory, and disk i/o, on a threshold basis. This protection might be required if server load fluctuates due to application request variance.
- Application Response Time as a threshold metric that kicks off Preferential Services policies. Target user groups can be all users accessing a particular critical application, or subsets of users, usually identified by cookies set which identify the class of service that user has subscribed to.
- "Hot Flash" policies, where specific content is replicated to previously unused servers (for that content) when the demand for that content builds up quickly and unexpectedly.

If you are lucky, persistency policies can be simple -- IP Source-based binding, with address range mapping may be adequate. But this application generally includes SSL sessions. If SSL Sessions are allowed to be lengthy, SSL Session ID Tracking is needed. If User State is stored at the web server level, SSL Session ID tracking needs to be integrated with other persistency schemes to achieve "shopping cart" persistency. This may require cookie-based persistence.

If the site also supports a significant amount of non-persistent traffic, it is useful to be able to set different persistency policies, depending on the content and application being accessed.

Mechanisms: This application almost always requires, or significantly benefits from, delayed binding and the array of value-added features enabled by delayed binding (URL-based scheduling, SSL Session ID tracking, cookie-based persistence and/or Preferential Services).

Make sure you can use lower overhead immediate binding Mechanisms for other applications supported at the site which don't require the delayed binding value added features.

Performance: From the VRM system point of view, sites that support this type application are usually pretty heavily stressed. Performing the delayed binding functions and related value-added features creates a heavy load on the VRM scheduler. Some low-end devices (if they have the features, which they generally do not) can't handle the load even if the access speeds are only a couple of T1 lines.

You need to make sure your selected product can support the load for your present and anticipated short-to-intermediate term growth, *with the features turned on that you need to make your site robust*. That means that a VRM scheduler should be able to handle around 3,000 connections per second (with delayed binding) for DS-3 access rates, and proportionally higher for higher speed access rates. You don't necessarily have to support such performance from a single scheduler if the vendor supports an Active-Active redundancy scheme.

Redundancy: Requires Active-Backup (at a minimum), with a logical split of activity between each scheduler. Requires Session Assurance. Active-Active should be explored if the performance requirements demand it.

Prospective Vendor Choices

Short List
Alteon
ArrowPoint
F5 Networks
IPivot
Resonate

This type of application is a real test of a VRM system, and no vendor meets supports every required and desirable feature for this application. Ideally, vendors would support ACV/DAV, SSL ID persistence and/or cookie-based persistence, Persistent Session Assurance, threshold-based policies, and an API for 3rd party resource health monitoring, with an architecture to easily support both low and high performance applications.

The vendors on the short list above support most of the desirable features, but each has a few shortcomings:

Alteon does not support SSL ID or cookie-based persistence, Persistent Session Assurance, or preferential services. But Alteon does exhibit good performance, which will only get better with their new generation of products, and they have promised all of the missing features on those new platforms. We'd feel better once they are actually shipping on those promises, however.

ArrowPoint does not yet support Persistent Session Assurance. But they have everything else, including some interesting content management-related features, and report good performance as well. They definitely belong on the short list.

F5 Networks does not support cookie-based persistence (yet), but they have been a leader in the areas of ACV/DAV, content management, and Persistent Session Assurance. The big question with F5 Networks is, lacking an Active-Active redundancy scheme, how well do their products scale when you have all the rich features required by this application turned on? You'll have to measure that for yourself because no vendor or independent 3rd party tests we have seen accurately models this type of performance-intensive application.

IPivot supports a rich set of features required for this application. They do not yet have Persistent Session Assurance, Dynamic Application Verification, or an API for 3rd party agent feedback. Furthermore, their scheduler performance reduces significantly when you turn on delayed binding, as required by this application. Exacerbating that, you can't really use their Active-Active redundancy scheme in this mode to try and get scheduler bandwidth back if you use content server logs for managing your site (which we almost always do). So we view IPivot as a good choice for sites that need features more than performance.

Resonate does not yet support Persistent Session Assurance or preferential services. But they have focused on the integration of 3rd party agent feedback via their Commander product, which can be programmed to make a wide variety of load balancing policy decisions based on the cumulative feedback from instrumentation at all tiers of the application architecture hierarchy.

Resonate is the only vendor that can support delayed binding and direct-path-return simultaneously. Resonate also scales nicely due to their Active-Active scheme, which allows you to turn on (in the extreme) as many schedulers as you have content servers. Resonate should definitely be on your short list for this application.

5.1.5 Local VRM, Server Optimization

Server optimization is not usually an application in and of itself, but an architectural design choice that can be made in the context of all the local applications above.

The goal of server optimization is to reduce the role of some server clusters to a specific set of applications or content space. The way to do that in HTTP environments is to perform URL-based scheduling. Thus requests for [.html] can go to some servers and [.asp] or [cgi] requests can go to others. Directory structures can be examined to partition the overall content space.

URL-based scheduling is also required if you want to perform any cookie-inspection functions for purposes of persistence or preferential services.

If you decide to use server optimization via URL-based scheduling, look for the following attributes in a VRM system:

VRM Scheduler Requirements

- Active-Active redundancy and scalability is preferred, especially for high use sites. Delayed binding requires a lot of scheduler horsepower. Packets must be heavily processed in both directions. You must choose a solution that will scale as your needs change.
- Scheduling rules and persistency policies should be configurable on a per-service basis, where services are defined as the set of applications and content partitions that the URL-based scheduling can re-direct to.
- Ideally, the solution should allow you to perform URL-based scheduling only where needed, with immediate binding elsewhere. This should be configurable on a per-V_IP basis.

Prospective Vendor Choices

Short List
Alteon
ArrowPoint
HydraWeb
IPivot
RadWare
Resonate

All the vendors on the short list are viable candidates for a Server Optimization application. They all perform URL-based scheduling, which is the key requirement for this application.

Alteon, ArrowPoint, and Resonate are currently the vendors we'd consider as high-performance candidates for this application. In the case of Alteon and ArrowPoint, we're waiting to see more 3rd party validation of their performance, but in theory, knowing their architectures, it should be OK. Resonate already has some proof points of the ability to support a huge website that requires server optimization features, e.g.Geocities. IPivot and RadWare have the features, but the performance is low in this case, where delayed binding is required and direct-path-return is not possible.

5.1.6 Multi-Site VRM, Site Redundancy

By supporting application and content availability at two to three sites, the application can be made *even more* reliable, while providing some increase in scalability.

This section and the next one address the unique incremental requirements associated with supporting applications and/or content across multiple sites. At each site, you still need to characterize the environment of that site and apply the local site guidelines in the sections above.

Application Characteristics

- This application is fairly straightforward, building off the local static site and dynamic one-way application characteristics.
- The primary goal is high availability through the deployment of multiple VRM systems at two to three data centers, which includes the VRM schedulers, the front end servers and back end servers (if any). Secondary goals include spreading user load among multiple data centers/content sites and the operational advantages multiple data centers offers (whole data center downtime, for instance).
- The application flow basically consists of:
  1. client determines site through DNS request (DNSR assumed – the guidelines given here do not depend on that assumption)
  2. client opens TCP connection
  3. user request
  4. server index response
  5. client opens TCP connections to request all objects indicated in index
  6. server responses
  7. closure of all TCP connections

If the multi-site application is more dynamic in nature, where significant data updates go from user to application/server, the characteristics change in the following ways:

- Requires more persistence if there are databases/data stores at different sites. Users must <u>always</u> be associated with the same site/server until data can be synched between sites.
- Requires better resource monitoring to know at any given time if every request could be satisfied from each site.

MS-VRM Scheduler Requirements

Any VRM at each content site/data center has the same VRM requirements as the local applications. Follow our basic multi-site rules.

Feedback: Content site testing requires ACV from the MS-VRM to the content sites. If VRMs are at the content site, an IVP is required to communicate health and load information for use in secondary policies. Site response time monitoring (coupled with appropriate policies) may be useful as a secondary feedback method.

When dynamic applications are supported at multiple sites, content site testing requires DAV, and potentially integration with 3rd part monitoring systems to ensure that all application requests can be fulfilled from each site.

Policies: Determine site for the user based on static client preferences. Router table import is a useful option to automate this process. This approach provides automatic persistence. If secondary Mechanisms are used to re-direct traffic away from overloaded sites, a site persistence policy must override that Mechanism for selected users or applications.

Mechanisms: DNS Redirection as primary mechanism, with HTTP Redirect as secondary mechanism. Use ARF (IP Proxy, Triangulation) for non-HTTP applications. Multiple A-record return may be a useful mechanism to ensure quicker site/server failure recovery.

Redundancy: Hot Standby MS-VRM units are required. Combined MS-VRM/local VRM units with multiple NS-Record return is also viable.

Prospective Vendor Choices

Short List
Alteon
ArrowPoint
F5 Networks
HydraWeb
IBM
IPivot
RadWare
Resonate

All of the short list vendors support viable methods of distributing users to multiple sites, with the capability of ensuring user-site persistency. The vendors' solutions achieve this through either fully distributed local/MS solutions (**Alteon, Arrowpoint, Hydraweb, IBM, IPivot, Radware**) or through centralized multi-site scheduling (**F5 Networks, Resonate**). We have a hard time distinguishing between the feature sets offered by the various short list vendors for several reasons:

- The science of multi-site load balancing is relatively new
- The number of multi-site applications we have been involved with to date is much smaller than the number of local site applications
- We tend to prefer fairly simple multi-site designs – oriented towards statically directing users to sites that are in the same general geographical area, and only changing the site they are directed to if a site fails or is heavily over-loaded. Therefore many of the bells and whistles vendors provide to attempt to perform more granular load balancing or "closets site" determination are irrelevant to us.

The one vendor we can't recommend is Cisco, because of their lack of integration between local and multi-site load balancing. It's like implementing two entirely different systems.

In general, we think that if you have a multi-site requirement, you should start with the short list above and examine each vendor's suitability for your local site application requirements. This process will result in a shorter list or possibly even a specific vendor choice.

Overload protection may be a 2nd order decision-making criteria, to be used only if all the other criteria result in a tie between one or more vendors. Overload protection is where vendors support the ability to redirect users/users requests from a site during periods of overload or server unavailability. These vendors accomplish this with combinations of HTTP Redirect, Triangulation or IP Proxy.

5.1.7 MS-VRM, Global Content Distribution

In this multi-site scenario, applications and content are deployed at multiple locations, often throughout the world, to reduce latencies for content access, reduce WAN costs, and to scale content availability.

Application Characteristics

- Building off the Site Redundancy characteristics, the desired goal in this application to make content available as close to the applications' major user populations as possible
- Primary goals include reduction of application response time for as much content as possible and the reduction of both predictable and unpredictable traffic load on the data center sites.
- This solution can easily be built with the Reverse Proxy Cache Server solution described in a previous section, were there are **Data Center** sites and **Content Sites.** The Data Center sites have the same characteristics as the Site Redundancy application. The Content Sites have one or more cache servers configured in Reverse Proxy mode. The best design would include VRM schedulers with the cache farm at each of the Content Sites, with appropriate numbers of cache servers to handle anticipated request load. These VRM schedulers would be capable of communicating with the MS-VRM system via an IVP.
- Determining optimal content site placement is a matter of statistical analysis of available user information, such as source network or source domain. For instance, you could select an arbitrary source-network traffic threshold of **5%**, such that if more than 5% of your sites traffic comes from a certain network provider or domain, you try to build a content site near that user population.
- Content site placement may require you to contract with a Service Provider that is capable of hosting your content (either on your servers or their servers) close to your designated user populations. You may want to choose the Service Provider to manage the servers to avoid the operational overhead, as long as your entire system can be built around its availability (or unavailability).

MS-VRM Scheduler Requirements

They don't change much. Here's some additional requirements to consider.

Feedback: Extended IVP capabilities between the content site VRMs and the multi-site VRMs. Consider integration with the Service Provider's (if you use one) monitoring systems as another set of feedback and input. For instance, the SP's monitoring system near a content site may help the MS-VRM system determine which Data Center site is best for the Content Site to retrieve content from.

Policies: Still use the static method, based on IP address ranges or source domains. You've already done the analysis; you know where your users are. Determine secondary policies based on the likelihood and effects of any one Content Site becoming unavailable. If a major content site goes down, how will that affect your traffic load at other sites?

Mechanisms: No suggested changes.

Redundancy: No suggested changes.

Prospective Vendor Choices

Short List
Alteon
ArrowPoint
F5 Networks
IPivot
RadWare
Resonate

The short list is comprised of vendors that meet the previous MS-VRM requirements and have an extended Inter-VRM protocol to ensure adequate load distribution among the content sites. The second list is again a nuance of the first, where capacity-availability tuning is possible using the methods mentioned in the previous section.

5.1.8 Co-Location/Hosting

Co-location services have experienced significant growth in their use over the past several years, and more recently have improved the diversity of their services. One year ago, you could rent time/space on a shared web server or rent rack space to put your own servers in. Value added options at the time included basic systems admin and operational services. Now you have the options of multi-site support where you can place your servers in multiple data centers, as well as defining service levels that you want met as part of your service. Additionally, a new type of service provider -- the Application Service Provider – has emerged to provide application hosting services, where you buy user licenses/time/bandwidth to access hosted applications like Peoplesoft and SAP.

Most online companies that were born as a result of the Internet (for instance, Yahoo) did not have the time nor the "company genetic history" to be able to build out their own data centers. No time to plan, no time to build, no time to wait. Companies like this usually have the motto of Just Do It[2]. They don't build their own data centers – they co-locate!

The primary reason to use a co-location or hosting service provider is to offload your operational staff from day-to-day activities, to lower your entry cost into 24x7 computing and to place your application near a source of wide bandwidth to the Internet. Its no longer necessary to build your own glass house data center to get redundant power and cooling, tight security and room for growth. These guys do all the heavy-lifting in those areas, allowing you to focus on what you're going to deliver to your users.

This section (and this report) is not really targeted at the architects of the co-location services and networks; it's targeted at those who need to know how to integrate a VRM system a co-location environment.

- First of all, if you are going to co-locate, it's your responsibility to select a VRM solution. The co-locator might recommend something, but that's usually because they have a deal with the vendor they are proposing and get commissioned if it is selected. If you're going to choose your own servers, your operating system, your applications because they will be the best for your purposes, you should also choose the best VRM solution for your purposes.

[2] This sounds familiar. Probably trademarked by someone.

- You're not going to be physically near the VRM system, unless you're in the same building as the co-locator. You need to choose a VRM scheduler that has good remote management and troubleshooting features.
- Choose a solution with a compact design – rack height costs you money.
- If the co-locator offers High-Availability as a service, make sure the service requirements don't make you compromise anything you would have done had you built the VRM solution into your application yourself.
- If they offer VRM solutions as part of their service offerings, they will either do so in a shared fashion (everyone shares WAN bandwidth, some number of customers share the VRM system) or they may offer a dedicated VRM system. Again, as above, make sure you're not compromising anything.

Co-location/hosters themselves have some unique things to worry about if the choose to share a VRM solution among many customers:

- They need lot's of V_IPs (100s), lot's of real servers (1000s), lot's of simultaneous connections (100,000s), lots of throughput
- They need the ability to set different rules and policies per V_IP
- They may choose to implement bandwidth management, usually in the form of hard limits to how much traffic can go to a V_IP or collection of V_IPs owned by a singe customer
- They need a management interface that allows them to view collections of V_IPs in the context of a customer (and a different collection in the context of a different customer)
- They should have the ability to give a customer-admin access to the VRM management interface for just their own partitioned services (i.e. the customer admin can look at their own stats, their own configuration, but no one else's)
- They need to ensure that as provisioning of new customers occurs, no other customer is affected
- They generally *don't* need sophisticated delayed binding features

Prospective Vendor Choices

Dedicated VRM Short List	Shared VRM Short List
Alteon ArrowPoint Cisco F5 Networks Foundry IPivot RadWare	Arrowpoint F5 Networks IPivot

The dedicated short list is comprised of vendors that should be considered if the VRM scheduler is dedicated to a single or small set of customers. The shared short list is comprised of vendors that support high capacity configurations (many sessions, many servers/services) and have some levels of preferential services available to tune resource contention between the customers, such as bandwidth management.

6 VRM Features: Chart and Definitions

To provide an easily searched review of the vendor capabilities, Acuitive has developed a Detailed Feature Chart. This Chart will be distributed via the subscription service mentioned at the outset of this document.

The chart below is a generic example of the vendor Detailed Feature Charts and serves as a Glossary of Terms for both the Technology and Vendor documents. It contains all the headings of the complete vendor Charts, and provides **common definitions** for most features. These definitions will not appear in the actual vendor Detailed Feature Charts, where the Definitions field will be replaced by a "Comments" section. This Comments section will be used to clarify any issue that requires more explanation than the standard Choices allow.

The Chart below can be an invaluable tool in speaking with the vendors, evaluating their products and comparing marketing literature to actual capabilities.

All of the technical capabilities listed here are discussed in detail in the Acuitive VRM Technology Report. If you want more information about some of the capability definitions, or are curious about a capability that has no listed definition (i.e. Single CPU – Pentium), please refer to the aforementioned Technology Report.

Local VRM	**Definitions**
LAN Environments	Includes Fast Ethernet (w/ Fast EtherChannel) and Gig Ethernet, Token Ring, FDDI.
Core Architecture	
Single CPU	
Pentium	CPU speed, available memory
Embedded	" "
Other	RISC, etc.
Scheduler OS	
BSD Unix	
Real-Time Executive	
Other	Any proprietary solution – *not* heavily customized OS (Free BSD, etc.)
Embedded Multiprocessor	
Data path processors per system (box)	
Type processors	
Hardware Assist	
Peripheral Accelerator card	
ASIC	
Functions Assisted	
Feedback	How the VRM scheduler determines server and application availability so it can make a proper and efficient user dispatch decision
External Monitoring	Feedback Monitoring that takes place externally from the servers
Device ICMP Pinging (DIP)	PING the defined IP interface(s).
TCP Connection Observation (TCO)	Examine all TCP requests as they flow through, and passively monitor all responses. The lack of a response (or many responses in a series of requests) may indicate an application or server failure.
TCP Connection Verification (TCV)	Actively execute TCP connection establishment process with configured TCP application listening port. Ex: For HTTP, perform 3-way TCP handshake to TCP port 80, wait for connection establishment, and then close the connection with a TCP FIN.
Active Content Verification (ACV)	Execute HTTP request to the server, based on predefined HTTP URL. Watch for valid response HTTP response codes VS invalid or no response.
Response Time Monitoring	As value added option to ACV, measure the app/server response time during ACV and use for feedback/policy decisions.
Passive Content Verification (PCV)	Passively monitor each client request or a sampled set of client requests for content, and examine return HTTP codes. Invalid return codes can be used for alerting, or policy decisions.
Adjustable response sampling rate	Observes return codes every *n* connection
PCV with Re-Direction	Allows VRM device to monitor return codes, intercept error codes from server to client and re-direct the original client request to another server that can fulfill the request (client is unaware the originally selected server cannot return the

	requested content)
Return Code Filtering	Allows VRM to watch for certain codes (errors), but not redirect *all* return codes, reduces load on VRM device, avoids potential malicious-use conditions.
Dynamic Application Verification (DAV)	Dynamic Application Verification (DAV) is similar to Active Content Verification, except that the entire content returned is examined by the VRM system. This enables the verification of dynamic applications, such as .asp applications, cgi-scripts, forms, etc.
Response Time Monitoring	Similar to same function under ACV.
Non-HTTP Application Verification	Applications/protocols testing other than HTTP, via DAV. Ex: verification of a DNS server/application via a DNS query test
Pre-Packaged Tests	Vendor provided content/application verification tests
Script Interface to Define Tests	End-user customization capability of Dynamic Content Verification tests
Remote External Probe Support	Allows "3rd party" probing of selected resources from a site external to (or away from) the VRM site.
Service Failure Detect VS Server Failure Detect	The VRM scheduler has the ability to detect a service failure (an individual application, such as HTTP) on an individual server and remove that service from use on that server. Contrast with removing the entire server from use when a single service fails on the server.
Resource Resident Monitoring	
Real Time Agent Feedback	Agents on the servers that are receiving load balanced requests are monitoring local resource and communicating resource and application availability information to the VRM scheduler.
CPU Utilization	
Available Memory	
Disk or storage I/o utilization	
Sanity timer time-outs	
Application queue lengths	
Data input/output (bytes and packets)	
NIC card utilization	
Internal bus utilization	
Application specific resource utilization	
Other	
Intrinsic Agent Interface	A software agent that ships with the VRM product as part of, or as an optional add on to the VRM scheduler. These agents are loaded onto the servers for application availability and performance monitoring and reporting .
Polling or Thresholding	Does the VRM need to poll the agent to determine status, values, etc. or does the agent perform "self-monitoring" and generate an alert message when a threshold is crossed?
SNMP or TCP	
Event Correlation?	Very powerful, the ability to interpret the event notification, possibly in the context of other notifications received, to determine the proper action.
3rd Party Interface	The VRM scheduler provides a 3rd party interface or API to achieve the same as above.
Polling or Thresholding	Same as above.
SNMP or TCP	
Event Correlation?	See above definition. (Event Correlation, Intrinsic Agent Interface)

Mechanisms	How the VRM scheduler dispatches users/user requests to a chosen server
MAC Address Translation (MAT)	MAC Address Translation translates the MAC Destination Address (and sometimes, but less importantly the MAC Source Address) to allow the frame to enter a NIC on the LAN with the given MAC address. IP Addresses in the frame are unchanged.
Direct to Client	MAT provides option for direct server to client response, without requiring VRM intervention on the return trip
MAC Multicast (MAM)	VRM responds to V_IP ARP requests using Multicast address, thereafter, all VRM devices in this broadcast domain will examine MAM packets and their IP contents to ""choose" if they will service the request
Half Network Address Translation (HNAT)	For Client → V_IP → Server traffic, change the destination IP address to content server real-IP address, no change to source. Change the IP source address for all return traffic from Server → Client.
Network Attached Router Configuration	VRM IP address is configured as the default gateway on R_IP content servers, in this fashion the VRM device does not have to be wired between the servers and the default router
Full Network Address Translation (FNAT)	Both source and destination IP addresses are changed by the VRM device. From Client → Server, the destination IP is changed to real_IP of content server and source IP is changed to VRM device. From Server → Client the source_IP changed to V_ IP address, destination IP is changed to client IP address.
TCP "Diddling" (TCPD) *Form of Delayed Binding.*	VRM system completes the TCP handshaking process with the client, fooling it into thinking it has established a connection with a Real Server. In the process, the VRM system records the initial Sequence # set by the client. In the handshaking process the VRM system responds with its own Sequence # and records that information for later use. VRM then makes a "best server" decision, establishes server connection (using original TCP/IP parameters). When server responds, VRM records this info. All traffic passes directly from the client's TCP process and the Real Server's TCP process, but the VRM system interdicts the traffic in the server-to-client path in order to adjust sequence numbers to match the clients expectations based on its original handshake with the VRM.
IP Fragment Handling	Properly maintain state on all TCP session transactions so that IP fragments can be appropriately handled if seen by the VRM scheduler.
TCP Gateway (TCPG)	VRM creates separate TCP sessions to the client and to the server. First, as client connection requests arrive, a TCP session is created to the client. After the handshaking process is complete and the client sends to HTTP Get request, the VRM can choose a Real Server based on the specific content or resource requested. A separate TCP session is then created to the real server. The intermediary VRM passes the data portion of the frames between the TCP Sessions.
TCP Connection Hop (TCPH)	The VRM scheduler completes the TCP connection process for the (as yet unidentified) Real Server as connection requests arrive from clients. When further information relevant to the active policy arrives (such as the URL), the Real Server can be chosen based on the specific content or resource requested. The scheduler then sends a message to that Real Server telling it that it wants to establish a TCP connection to an identified Virtual Service *as a proxy for the client.* It uses the same client IP address, source port numbers and sequence number information as was established by the TCP connection set-up process with the client. The Real Server then "takes ownership" of the TCP connection and responds directly to the client, in a manner that is transparent to the client.
Delayed Binding Features	Forwarding decisions based on variables received after TCP connection is established between client and VRM device, but *before* VRM to server connection has been made
URL-Based Scheduling	Choose a server based on configuration policy, compared to data found while parsing URLs from clients.
URL Repetition	Ensure that multiple requests within in a short period of time for the same URL go to the same server.
PCV with Re-Direction	Passively monitor response codes from each user request and redirect the user request to a different server if the request generates an error response.
SSL Session ID Tracking	Track SSL session ID's to use as a "persistency key' for persistent binding. User with same SSL session ID goes to same server, regardless of source IP address. Compare/contrast to IP Source Host/Subnet based binding.
Session Cookie Binding	Track temporary session cookies to use as a persistency key for persistent binding. User with same session cookie goes to same server, regardless of source IP address.

Permanent Cookie Binding	Examine data encoded within cookie to optimize resources, create preferential services.
HTTP1.1/HTTP1.0 "Gateway"	Translates HTTP 1.1 single TCP connection, multiple HTTP gets into multiple TCP connections (allows fine-grained forwarding decisions) for HTTP 1.0 servers
Policies	**How the VRM scheduler chooses what server to dispatch user requests to**
Best Available Server Policies	
Round Robin	Each new connection request is sent to physical servers in a round robin fashion such that, over time, each server gets the same amount of connection requests.
Weighted Round Robin	Each server is given a static weighting function that is based on some view of the capacity of each server. Servers are presented connection requests in proportion to their weighting relative to the total capacity of the system.
Maximum Connections	A Maximum Connections policy prevents a server from becoming overloaded (based on measuring the number of active connections). Once a server hits this limit, it is taken out of the Round Robin pool until the number of open connections falls beneath the limit.
Auto-Weighting	VRM system looks at the response of a server over time, as a function of offered load, and uses that information to either suggest modified weights or automatically adjust weights.
Least Connections	The number of active TCP or “user” connections supported by each server is tracked. As new connections are received, they are forwarded to the server with the fewest active connections.
Weighted Least Connections	Same principal as Weighted Round Robin, applied to Least Connections.
Maximum Connections	Sets maximum connection threshold per server. When threshold is exceeded, server is taken out of forwarding schedule until number of active connections drops
Packet Rate	By tracking the server's packet and/or byte receive/transmit rate, server load can be balanced in terms of packet processing.
Per server	Tracking per Real Server (or otherwise)
Aggregated Per Wire	Tracking based on data rate on LAN port statistics observation. Multiple Real Servers connected to a port would have packet rate observed as single set of stats, as opposed to set of stats per Real Server.
Threshold/Secondary Rule Support	Once the threshold packet rate is hit, this allows the VRM to make a decision to handle the situation (i.e activate an overflow server, etc.)
Byte Rate	See “Packet Rate” above. For byte rate instead of packet rate
Per server	
Aggregated Per Wire	
Threshold/Secondary Rule Support	
Response Time Policies	
Fastest Connection Time	By monitoring the interval between when the SYN packet is sent to the server and when the SYN ACK reply is returned, the time required for a server to respond to connection requests can be estimated.
Fastest Application Response Time	Application Response Time measurements, such as received through Active Content Verification processes, can potentially be used to drive a “best available server” policy.
Predictive	Where the trend of the connection response time is used, rather than the absolute value.
Threshold/Secondary Rule Support	Matches current and historical performance trends to pre-set limits. This in turn triggers the application of a specific rule (i.e. alert administrators that server 2 is heading for trouble, do something proactive)
Server Resource Management Policies	Rather than using such information for real-time and fine tuned adjustments, we prefer using these policies to identify overload or potential overload conditions.

CPU Utilization	
Available Memory	
Disk or storage I/o utilization	
Sanity timer time-outs	
Application queue lengths	
NIC card utilization	
Internal bus utilization	
Application specific resource utilization	
Other	
Threshold/Secondary Rule Support	
Persistency Policies	Persistency policies ensure that the same user(s) are associated with (or bound to) the same server for some period of time. A (potentially short) persistency policy is based on the requirement that all packets within a TCP session must go back to the same server. A long persistence application (one that lasts across many TCP sessions) is the example of the "shopping cart application," where users choose what they want to purchase over a period of minutes and many TCP connections. The real server they connect to records their shopping cart contents over these minutes. Thus, they must always communicate with the same server for this period of time.
Source Port Binding	VRM scheduler uses the Source TCP/UDP port as the "persistency key" for persistent binding decisions.
Source IP Address Binding	VRM scheduler uses the Source IP Address (client's IP address) as the "persistency key" for persistent binding decisions.
Source Identifier Binding	Source Identifier is the combination of the Source IP address and the Source TCP Port. Various flavors of this approach are highlighted below.
Address Range Mapping	Based upon subnet masks (e.g. matching the 1st 24 bits of the Source IP Address instead of the entire 32 bit field) or by statically configuring lists. Typically used for long session times, perhaps across many TCP connections.
TCP Connection Termination Monitoring	When SYN packets arrive, the VRM system makes a "best available server" decision and passes the SYN packet along to the server. When a FIN packet is observed, indicating the completion of the session, the connection information may be removed from the binding tables
Policies Per V_IP	Persistency policies configurable per destination (V-IP) address
Policies per Delayed Binding Scheduling Rule	Persistency policies configurable per URL-based scheduling rule, cookie ID, or SSL ID binding key.
SSL Session ID Tracking	Client and specific server handshake to begin an encrypted SSL Session, a unique SSL Session ID is assigned. When subsequent packets arrive with the same SSL Session ID, they can be re-directed to the same server that was involved in the original SSL session creation.
"Shopping Cart"	User browses though the site and indicates items he or she may buy by placing them into a "shopping cart." User must remain connected to that server for the duration of the shopping session (which may span many, many TCP sessions). Then, if the purchase requires an SSL session (for communicating credit card information, usually), the user must continue to remain connected to that server. Providing appropriate persistence for this situation requires tracking of sequences of TCP sessions from HTTP to SSL and back.
Session Cookie Tracking	When a user communicates with a website that has stored a cookie on the user's system, the browser transmits the cookie with every request. When a (cookie capable) VRM system gets involved, it terminates the TCP connection, waits for the URL requests to come from the user and then performs cookie inspection. The VRM scheduler chooses a server for the user and then records a representation of the user's cookie within its session state tables, remembering that the user with the inspected cookie is associated with the chosen server. Subsequent requests from the same user will continue to go to the same server, until some timer expires the session state entry.

Preferential Services Policies	Policies that are activated when site resources become scarce, to ensure that some users are offered better access to resources than others are.
By Application	By application, usually by TCP port number.
By Content/Directory	By content, usually by identifying the URL in the client request.
By User Group	By association of a given user with a group, and the group's preferential service policy.
Session Cookie	Detect group ID and identify preferential policy by identifying parameters in the HTTP cookie, for this session only.
Permanent Cookie	Detect group ID and identify preferential policy by identifying parameters in the HTTP cookie, for this user.
Login/Authentication	Detect group ID and identify preferential policy by snooping/participating in user authentication process.
Subscriber Database Query	Once group ID is detected, use a subscriber database for group preferential policy query.
By Subscriber	
Session Cookie	Similar to above, but identifies a particular user (and policy) instead of a group of users.
Permanent Cookie	
Login/Authentication	
Subscriber Database Query	
By Time-Of-Day	
Other	
Combinations?	
Preferential Services Targets	
Resource Oriented Targets	Any resource that can be measured and is deemed to be critical to site performance is a potential metric. (i.e. CPU utilization, available memory, open connections, etc.)
Open Connections	
CPU Utilization	
WAN BW Utilization	
LAN/Link BW Utilization	
Other	
Performance Oriented Targets	Performance-oriented targets should be related closely to the business mission of the site and the user experience associated with that mission. Application Response Time, % successful downloads, Transaction Time, % Reloads, and % user "quits" are all metrics related to the quality of the user experience.
Application Response Time	
Other	
Preferential Services Mechanisms	
Differential Server Weighting	If the performance of a high priority application degrades, differential weighting is performed, which means that quotas for resource allocation are invoked. (That could mean that the number of open connections allocated to all applications other than the highest priority one are reduced.) The sum of the allocations for all the lower priority applications should be lower than the maximum supportable on the site.
Server Pool Swapping	Different applications or user groups are initially assigned to different server groups If the performance of the higher priority application starts to degrade, servers are moved from the pool assigned to lower priority applications into the high

	priority group.
Rate Shaping	Data rate shaping is used to adjust the flow of incoming traffic to the content servers. Rate Shaping means that when servers are congested, you can send fewer packets to the servers related to lower priority user and applications, which frees up resource for the higher priority users and applications.
New Request Denial	Rather than slowing down the rate of packet transmission to the content servers (a'la Rate Shaping), one could deny requests for lower priority users during periods of congestion.
Existing Session Reset	Reset existing connections for lower priority users
Coupled to BW Management?	
VRM Redundancy and Failover	
Hot Standby	In this mode of operation, one VRM unit (the Standby unit) is waiting for the Active unit to fail. Upon Active unit failure, the Standby unit will adopt all VRM functions.
Active-Backup	In this mode both VRMs are actively performing VRM functions, but for different V_IPs. One VRM unit actively provides services for one half of the configured virtual servers, the other VRM unit handles the other half. A keep-alive protocol between the units ensures that each VRM unit knows the health of the other. If one of the units should fail, the other unit takes over all the VRM services and handles the entire load. Each unit is acting as a "standby" for the other.
Active-Active	All VRM devices are up and forwarding requests. They are each capable of forwarding requests for the same V_IPs and virtual services. When an active unit fails, some site scalability is lost, and perhaps existing user sessions are lost (unless there is the use of persistent Session Assurance).
Symmetrical Multi-Path (OSPF) Routng	VRM devices advertise a path to their respective V_IP hosts via OSPF. When a VRM device fails, an OSPF host route is no longer advertised for the V_IP.
Keep-Alives	The connectivity and mode of status communication between redundant VRM schedulers.
RS-232 Link	
Dedicated LAN Connection	
In-band Network Connection	
Combinations?	
Topologies	
In-Line	
Network Connection	
Multi-Port (Mesh)	
VRRP	Virtual Router Redundancy Protocol, which allows the sharing or pooling of IP interfaces in a router or VRM scheduler. A keepalive protocol ensures that all units in a VRRP group are aware of each other's status, so that a scheduler failure can be noticed and resolved.
Failover Features	
Persistent Session Assurance	Persistence-sensitive session tables (SSL, etc.) are shared between primary and backup VRM devices. When the primary fails, the backup has a recent copy of the session tables, and will honor the previously made client-server path agreements.
All Connection Failover	Primary scheduler shares entire connection table with backup VRM device. This is passed on a regular, pre-scheduled interval. When primary scheduler fails, all recent connections are maintained by the backup device. (in theory)
Session Table Synchronization	

Back-Up Server Activation Options	Does the VRM scheduler support activating “servers of last resort” when a server or server group fails?
Failure of all R_IPs	Activate a backup server upon Server Group failure
Threshold on available R_IPs	Activate a backup server upon a certain number of Real Servers failing
Threshold on V_IP performance	Activate a backup server upon capacity or performance problems
Server Dual-Homing	Support for servers attached to the network using more than one physical/logical connection.
New Server Slow Start	Gradual increase of traffic load to new server. Effective when a VRM scheduler brings a new server online so that the new server does not get swamped with too many user requests.
Extra Security	
Server MaxConns	Set a maximum connections policy on the servers, such that the VRM unit will not allow more than MaxConns numbers of connections to a server to be opened and active. If the MaxConns setting is set well below what the server is capable of tolerating, then this setting can be used as a method of protecting the servers.
SYN Metering	Measuring the *rate* at which user requests are coming into the system and metering their flow. User requests are only allowed to flow to the servers at a rate that is at or below the administrative meter setting.
TCP Connection Termination	VRM product is terminating all TCP connections prior to allowing requests to go through. Many DoS attacks designed to make servers crash are performed with spoofed IP source addresses. If the VRM product is terminating all TCP connections, then the TCP sessions from the spoofed sources of DoS attacks will never be completed
Integrated Firewalling	If the VRM product includes firewalling features, make sure they provide protection from the SYN and DoS attacks, and are not just simple packet filters.
Performance and Capacity	**The performance and capacity of the VRM scheduler**
HTTP hits/ sec (no delayed binding)	
HTTP hits/sec (delayed binding, URL scheduling)	
HTTP hits/sec (delayed binding, 100% SSL)	
HTTP throughput (w/ delayed binding)	
Total simultaneous flows/users	
Max V_IPs	
Max R_IPs	
Multi-Site VRM	**Definitions**
Configuration	
MS-VRM separate From Local VRM?	
MS-VRM schedules to non-VRM sites?	
MS-VRM schedules to 3rd party VRM sites?	
VRM added value feedback to MS-VRM?	

Feedback	
Content Site Testing	Extensions of Local VRM Feedback methods.
DIP	
TCV	
ACV	
DAV	
Inter-VRM Protocol (IVP)	Proprietary communication between VRM devices. Allows MS-VRM device to base user-to-site dispatch decisions on "internal" VRM site information, detailed below.
Server Health	
Client Proximity	How "close" is the client to each of the content sites? Used to make "closest site to the client" decisions. An inexact method, at best.
Site Load and Performance	Content site load and performance. Allows "best performing content site" based decisions.
Keep-Alives	
Site VRM Standby-Unit Health	
Encrypted Communications	Is the IVP protected from spoofing?
Content Availability	Does the VRM scheduler communicate content availability between the sites?
Mechanisms	
DNS Redirection (DNSR)	The MS-VRM is used as a DNS server for the content site(s) domain names, to return V_IPs as DNS A-records. This results in "dispatching" users to sites for a long/persistent period of time.
TTL Adjustable?	Can you adjust the TTL value within the VRM scheduler to mitigate potential site overload and long outage times because of site failures?
Multiple A-Record Return?	Can the MS-VRM scheduler return multiple A-records as part of its DNS services, for the same reasons as above?
Application Request Forwarding	
Triangulation	Triangulation is a method where arriving client TCP connection requests (SYN packets) can be sent to another site better suited to service the request. The Triangulation method uses special "shadowed" V_IP addresses at each other site that traffic might be forwarded to. When that site sees something arriving to the unique "shadowed" VIP, it knows that the packet was re-directed to it, and what site specifically performed the re-direction. That information is necessary, because on the return path (server-to-client), the VRM scheduler must put the original V_IP, from the site the client originally addressed, into the source IP field. That way, the traffic can return directly to the client. That's what creates the triangle; client to original site, site A to site B, site B back to client.
IP Proxy	Typically used as a mechanism of last resort between content sites, where a multi-site capable VRM unit at a site serves as a Proxy for user application requests, such that they are fulfilled from another site.
HTTP Re-Direct	An HTTP server can redirect a client HTTP request to another server for a number of different reasons. If a VRM scheduler at a content site is capable of using this feature, it will (normally) use it to redirect clients away from a site that has servers that are either overloaded or down.
URL Parsing	Local VRM unit can utilize dynamic or static URL information to determine content site to forward a user request to that it could not handle locally.
Secondary/Threshold Rules?	

MS-VRM Policies	How does the MS-VRM unit decide what site/servers to send the user to?
Static Client-Site Preference (SCP)	User will always be sent to a particular site unless the site is unavailable.
Source Address Preference	User/site decision based on statically configured site-source IP subnet associations.
Max Table Size	
Longest Prefix Match Entries	
Automatic Table Generation via Router Tables	Can the static SCP table generation process be automated?
Other Automation	
Source Domain Look-Up	User request arrives at the MS-VRM scheduler ; it performs a Reverse DNS lookup on the source address of the request. The MS-VRM unit scans the Source Domain Preference table and finds the most exact entry that matches the RDNS response. The MS-VRM scheduler then dispatches the user to the site listed within the preference entry.
Max Table Size	
L3 Topological Proximity Testing (TPT)	Each content site VRM scheduler tests the proximity of the client, usually via PING, Traceroute or a TCP connection. Once a response to the test comes back, the proximity of the client is determined by inspecting the IP TTL remaining, which (sort of) represents router hop count to the client. Proximity information is then reported from the content site VRMs to the MS-VRM via an IVP.
Initial Client Pass-Through	Does the MS-VRM wait for a "proximity answer" for a new client request, or does it send it to a default site for now?
Max Table Size	
Ping Test	
Traceroute	
TCP Connection	
Other	
Client Path Performance (CPP)	Uses measured client response time or path packet drop rate as the metric rather than TTL or hop count values. Each (capable) local VRM scheduler measures the response time (and/or packet drop rate) from each site to the requesting client by sending a PING or TCP connections request and timing the response. The response time (or packet loss rate) information is then passed off to the MS-VRM scheduler, which analyzes it to see which site experiences the most responsive communication.
Client response time	
Packet Drop Rate	
Initial Client Pass-Through	
Max Table Size	
Ping Test	
Traceroute	
TCP Connection	
Other	
Threshhold/Secondary Rule?	
Routed Path Analysis (RPA)	Relies on the routing topology tables that each Internet BGP-speaking border router and/or each interior router maintains to route traffic across the Internet and interior networks.

AS Hops	The number of BGP AS hops are calculated between each content site and the client. The shortest AS path wins.
Interior Route Path Cost	
Threshhold/Secondary Rule?	
Site Performance Measurement (SPM)	The use of ACV to perform application response time checking between content sites, which are considered to be peers. When an MS-VRM scheduler at a site performs ACV to another site, it measures the total time elapsed from the beginning of the ACV (TCP SYN) until the final frame of the session is passed (FIN ACK). Since the testing involves ACV from one site to another, most of the major components at each site will be tested for availability and throughput (routers, local networks, local load balancers, servers, etc).
Threshhold/Secondary Rule?	
Packet Retransmission (PKTR)	Use of packet retransmission data to better fine-tune the site choice for clients.
Threshhold/Secondary Rule?	
Persistency Policies	Combination of local and multi-site VRM systems must be able to persistently bind users to sites and servers.
Persistency By Group	
Persistency By Client IP	
Persistency By Content	
MS-VRM Redundancy	**Please Refer to the previous section, Local VRM Redundancy, definitions**
Hot Standby	
Active Standby	
Active-Active	
Persistent Table Failover	
MS-VRM Performance	
Redirects/Sec	
DNS Responses / Sec	
Non-Server Load Balancing	**Definitions**
Proxy Firewall Balancing	
Stateful Firewall Balancing	Stateful firewalls must process all traffic between two particular communicating systems that are on each side of the firewall. Frames that ingress the VRM from the Clean (interior) network, or the VRM from the Dirty (exterior) network are classified into session flows. These flows are balanced between both firewalls, with each VRM scheduler handling the appropriate traffic flow direction.
Router Balancing	
VPN Device Balancing	
TN 3270 Mainframe Front End (CIP) Balancing	Application gateways such as TN3270 Mainframe front-ends may be load balanced with VRM solutions.
Proxy Cache Server Balancing	These devices enable a level of scalability and redundancy for proxy cache servers, without requiring browsers to be re-configured to utilize multiple proxies. As requests come into the single IP DA the browsers are configured for, they are forwarded to a chosen proxy server for service.

Transparent Cache Server Balancing	These are devices that simply intercept requests to a specified port (usually Port 80) and send them to a Transparent Proxy Cache Server.
Proxy-to-Transparent Proxy Converter	These devices turn clusters of proxy cache servers into “apparent” transparent proxy clusters, by off-loading the routing, NAT and absolute URI re-generation functions from the cache servers, which means that standard proxy cache servers can be used, without requiring browser re-configuration to enable caching.
Non-Server LB Value-Add Functions	
TCP/UDP Destination Port based re-direction	Can the VRM scheduler redirect any protocol?
Address-based re-direction	Redirection based on source and/or destination addresses or ranges.
URL filtering - domains	Redirection / proxy server choice based on URL host header.
URL filtering - static vs dynamic	Redirection / server choice based on defined “static” vs “dynamic” content requests. Dynamic requests should usually go to the origin server rather than the proxy cache.
URL filtering - "most likely" cache	Redirection based on the knowledge of what cache server is likely to have the content request specified in the URL.
Network Address Translation asist	See Proxy-to-Transparent Converter definitions.
absolute URI re-generation	See Proxy-to-Transparent Converter definitions.
WCCP interoperability	To enable multiple cache server operation, Cisco has their own proprietary implementation to allow cooperation between their Cache Engine and their routers, with a registration protocol called WCCP. A WCCP-capable cache registers itself with a local router (and also use WCCP for keep-alives), that will dynamically start to send HTTP requests to it
Management	
Command Line Interface	
All monitoring functions supported?	
All configuration functions supported?	
Menu driven?	
Prompts?	
Serial/Console Port access?	
Telnet access?	
Password authentication?	
Encrypted?	
Browser User Interface	
Technology	
Straight HTML?	
Javascript or dynamic HTML?	
Java?	
XML?	
Security	
Password Authentication?	

Encryption?	
Functions	
All monitoring functions supported?	
All configuration functions supported?	
Require a proxy server?	
Non-CLI features (e.g. graphs)	
Installed Management Application	
SNMP?	
OS / Management Platforms supported?	
System Alerts	
SNMP Traps?	
BUI Log?	
CLI Log / Messages?	
Email?	
How many VRM specific alerts?	
Content Management	
Content Caching	
Automatic (Timed) Flushing	
"Most recent" Request Forwarding	VRM scheduler determines which server most recently handled the current request, to take advantage of disk caching that the server may be performing.
Local Reverse Proxy Caching	Can the VRM scheduler support reverse proxy caching locally in front of a website?
Auto-Configuration	Feedback from the content management system (or searching by the VRM) to determine precisely what servers can access which content and applications, and automatic maintenance of V_IP-to-content server tables with-in the VRM scheduler
Update Management	As content gets updated, narrow the target server set to the updated servers, adding servers to the group as they become updated, until all servers are again targets for the updated content
Content and Application Launching	Upon recognition of application performance degradation by the VRM system, control the content management system to deploy content and applications on new servers to increase the available resource to server the application