

The APM Paper

2nd EDITION

An Architectural View
Of
Application Performance Management



*April
2004*

Table of Contents

I.	Introduction.....	3
II.	APM Overview	4
III.	An APM Solution	6
IV.	APMA Issues.....	8
V.	Industry Trends.....	10
VI.	An Architectural View.....	12
VII.	Items of Architecture	14
VIII.	Web Network Architecture.....	16
IX.	APM Infrastructure.....	23
X.	A Look Ahead.....	26
XI.	Conclusion.....	28

I. Introduction

APM Advisors covers the Application Performance Management marketplace, which led to the development of the APM Paper. This is the 2nd Edition of the APM Paper, which is structured to briefly overview the foundation conveyed in the 1st Edition, and then delves into the state of the industry and the factors shaping its direction. Therefore the paper has value to a broad audience, including enterprises of any size, private and public investors and naturally the companies who make up the industry, or have ambitions of entering it.

This version of the 2nd Edition APM Paper is licensed for distribution by various vendors of APM solutions. There is another version of this paper that contains a detailed list and review of all known vendors of APM *control* solutions, which is available directly from APM Advisors at www.apmadvisors.com.

In this paper we'll take a look at the breadth of the marketplace, rather than to dig into the details of each category / product / company. As buyers of APM solutions the goal is to provide a high level of service to users, with a focus on business applications. The other focus is to maximize the investments in infrastructure (networks, servers, applications and personnel), by minimizing the impact of non-critical or non-business usage. With that as a foundation an IT organization needs to prioritize it's spending in APM solutions that achieve their cost performance goals.

The key elements of a comprehensive APM solution are:
Information - Resolution - Control

In the 2nd Edition of the APM Paper there is much more attention on the system architecture issues encountered when looking at developing an APM solution within any system. The foundation of our approach is based on something we call the Web Network Architecture (WNA), which will be referenced in the document. The premise is that basic *information, resolution and control* are critical elements in any networked application environment. It's our position that the *infrastructure* must be *instrumented* in a method that provides optimal value at the lowest possible cost.

The future of APM solutions will be based on
Instrumenting the Infrastructure

If the future is based on *instrumenting the infrastructure*, the next question is 'What is *infrastructure*?' Obviously, there is a requirement to have an application environment and user desktops, as well as a piece of wire to connect them. After those basics, the size of the environment, locations, breadth of application environments, etc. begin to drive the critical elements of infrastructure.

Fundamentally we view the *control* solutions as having the most value in IT systems and therefore will almost exclusively focus on that aspect of the industry. While this leaves a number of solutions that meet the *information* or *resolution* qualification of APM outside the scope of this document, the need and value of those products drove the decision to stay with *control*.

You don't have an APM solution without *control*!

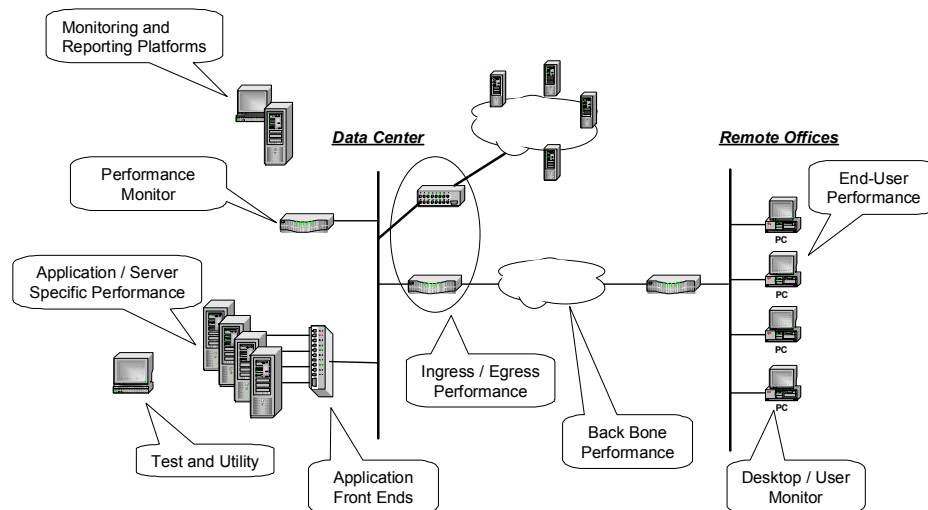
II. APM Overview

Since the founding of APM Advisors, we have found it interesting how Application Performance Management is viewed within the industry. Most of the industry seems to view things from a product perspective rather than an architectural or solution perspective. This is especially true in the large analyst firms and larger enterprise systems, where specialist abound. However, many functions required to deliver secure and high performing applications are consolidated in one product as a solution rather than just a functional requirement.

Product Scope

For definition purposes, APM solutions include any product or solution that either provides performance *information*, *resolves* a performance problem or *controls* some aspect of application performance. All the subcomponents of a system, such as the network, desktops, etc., have a direct relationship to the user's experience associated with an application and are therefore an APM solution.

The following diagram outlines the solution structure we developed for the 1st Edition of the APM Paper. It was a good starting place and still applies to the breadth of the marketplace.



*Application Performance Management Solution Categories
1st Edition APM Paper*

Monitoring & Reporting Platform

Offerings within this category are centralized *information* mining platforms that obtain the majority of their *information* for providing real-time monitoring or reports, from other products or systems.

Application / Server Specific Performance

Offerings that are tailored to monitor or manage a specific application and/or server environment. Any product within this category has a software 'agent' that executes on the server or application platform. For the most part these solutions specialize in *information*, *resolution* or *control*, but newer entries are beginning to address the complete APM solution requirement.

Ingress / Egress Performance

Offerings that monitor or manage traffic between the LAN & the WAN. This may include stand-alone products or functionality enabled in an edge device, such as a router. The products within this category may provide traffic *control* (flow and routing), caching, compression or filtering. Overall, their functional services have potential application at any Ingress / Egress point within an intranet or the Internet.

Application Front End

Offerings that improve the performance of a cluster of web server or application platforms. This category of solutions has replaced the basic Server Load Balancer by incorporating core functions such as SSL Acceleration, compression, TCP optimization, caching and IDS/IPS. Virtually all the products offer these basic functions, but there are a number of unique value-added functions as well.

Desktop/User Performance

Offerings within the Desktop Performance category provide desktop level software that gathers *information* regarding the performance characteristics of the desktop platform or user experience.

End-User Performance

These offerings are delivered as desktop software, and perform specific tasks that affect the user's experience of networked applications.

Performance Monitor

These products monitor traffic flows to develop application specific performance metrics, such as traffic statistics, response time, server and application processing times.

Test and Utility

These products provide a variety of synthetic transactions or other simulations to test networks, servers and applications. These products are classified as 'passive' since they do not perform a function that alters performance.

Backbone Performance

Backbone Performance is isolated to specific functions that are monitored or performed on traffic flows between the Ingress / Egress points in the network.

III. An APM Solution

With nine solution categories it's clear there are a number of companies and even more products to cover. Rather than provide detailed *information* on each product, we believe the value of the research was better tailored at the state of the industry, the trends and the end goals.

As outlined earlier, there are three fundamental elements to establishing an APM system solution. (*Information, Resolution and Control*) That is much easier to say than do for many reasons. A key aspect of the challenge is that the environment that any organization is attempting to stabilize is constantly changing. Therefore an APM solution needs to be as adaptive as the environment it is attempting to provide stabilization for.

Information, Resolution and Control

Information

This is the most basic and initial step for an IT organization should achieve. It's quite basic, that at the business level and within the various IT disciplines, an organization needs to understand it's networked application environment. What is running on the infrastructure, what are the baseline or current operational characteristics, who is using it and what are the associated cost.

Resolution

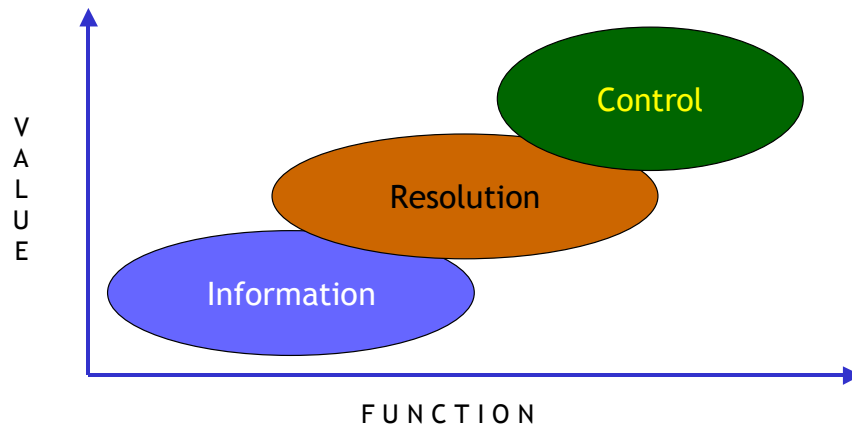
Regardless how well the system is designed and managed, there will be events that have an impact on application performance. Additionally, changes within the system, such as the introduction of new applications, user population growth, etc., will have a direct impact on performance levels. Being able to quickly identify those conditions, isolate the cause and restore stability are important capabilities to have ready to initiate.

**APM Solutions Need to Include *Control*
The Alternative is Increased Cost**

Control

In order to achieve the APM goal of aligning the business of IT, an organization needs to have the means to *control* the variables in their environment that require alignment. The breadth of *control* options is perhaps the most powerful aspect of APM and as outlined in the paper, is also the segment of the industry in the most flux. With the ability to *control* everything from the application to the behavior of a user, the power of APM *controls* are just beginning to develop.

The real challenge for any organization is making sure that the *information* they are gathering is meaningful to their business, the ability to react is instrumented and the methods of *control* are appropriate. The common feedback from IT organizations is that they have more *information* than they need, but find it lacks a system performance context. In respect to isolating anomalies, given enough time and money, any problem can be resolved, but the key is to have an infrastructure instrumented for facilitating efficiency in resolving those problems.



Establishing and leveraging *control* mechanisms is the holy grail of APM, since the alternative for not having *controls* is increased cost.

Information and Resolution in Control

We had this buried much later in the document, but brought it forward because we've been amazed that most vendors of *control* solutions don't seem to understand the value of *information* and *resolution*. Since *control* systems are instrumented into the infrastructure they have a significant amount of *information* that IT organizations pay lots of money to obtain through *information* products that can't perform the *control* functions.

What most vendors say when ask about this is "Well that's not our role. We don't want to become a reporting product." Well no one is asking them to be the reporting product, but the insight they have would make most of the reporting products better and much less expensive for IT to deploy.

Our favorite example is **Compuware's** Vantage offering, which provides one of the best top-level views across a system that we've seen. However, it will cost an IT organization more than a lot of money to obtain that picture because it requires the deployment of desktop software, network probes and server agents. So if an IT organization were deploying *control* products that instrument the same resources within the system, wouldn't it make sense for those products to provide comparable *information*? We think so.

Beyond just *information* it's all about context, since much of the *information* isn't relative to system performance rather it's more often used as 'CYA information'. Or in other words "It's not my fault." Without any type of formal survey a majority of the information systems are packaged for sale within a discipline in an IT organization. By merely obtaining information about the elements within a system, it's value across the IT organization or to the business is close to zero.

IV. APMA Issues

With well over 100 companies claiming to have APM solutions there is quite a bit of perception building going on. So this is our, “let’s call a spade a spade” section, where we withhold the names to protect the naïve, but it isn’t too much of a stretch to figure out which companies are making claims much bigger than their reality.

Furthermore, there is a lot of hidden value in various APM solutions, but many of the vendors lack the vision to take their value to the finish line. This is perhaps a bigger crime than hype, since there is so much potential in many solutions, but companies get focused on the ‘trees’ and forget they are part of the ‘forest’. We would spell this out more, but after all we make our money consulting.

Business Context

This is perhaps the largest gap in what is being said and what the potential is for many offerings in the APM marketplace. Quite often vendors and analyst will position various solutions as being business oriented or structured to provide value to the business. While there shouldn’t be a product purchased that doesn’t add value to the business writing the check, there is quite a bit of conceptual stretching going on to mold some of the solutions into a business context.

What vendors are missing is that business just wants the technology to work and is purely focused on the fiscal operations of the company. Over the past year at APMA we have probably taken about 100 briefings from vendors and out of those less than 10 can explain their core value to the business they target. For the most part, they clearly understand their value to IT they struggle to provide either the structure or method for conveying that to the business.

Synthetic Transactions

Stop the insanity! Closely aligned with the Business Context discussed above a synthetic transaction is not a customer. Services like those from **Keynote** or products like Loadrunner / Topaz from **Mercury Interactive** don’t replace the real experience of end-users. If you talk to the business side of any organization providing a B2C, B2B or internal service, they care about their customers and not a machine.

While it used to be virtually impossible to obtain end-user response time because solutions didn’t scale or the overhead of maintaining them was unsupportable, we’ve moved past that. So since the solutions have advanced, it’s time to start getting a view of the end-users experience rather than some machine.

End-to-End

As clearly stated in the 1st Edition APM Paper, there are no end-to-end solutions and nothing has changed since then. Our challenge to vendors is to explain how your solution is ‘end-to-end’ or if a shop, tell us how you’ve implemented ‘end-to-end’. Now before you pick up the phone, remember an APM ‘solution’ includes *information*, *resolution* and *control*. At best and with a lot of money you can obtain quite a bit of *information* across a system, but can it be put into the context of performance of applications ‘end-to-end’. We don’t think so.....

Resolution

For the most part vendors consider *resolution* of APM issues as being either “it’s not here” or “it’s over there somewhere”. As outlined in the 1st Edition of the APM Paper, that type of *information* is better classified as CYA reporting.

Those who are focused on specific elements or tiers within a complex environment are generally better at providing specific *resolutions* within their area of expertise. Remember that this isn’t ‘kids stuff’ and it takes quite a bit of expertise and granular instrumentation to provide specific fault *resolution* within a system.

Policy

As outlined in the 1st Edition APM Paper, the DEN (Directory Enabled Network) initiative was the best example of good intention gone bad. However, the issue is more relative than ever and is begging for some leadership. Just take a look at any and all of the *control* solutions within the marketplace and each one of them requires policy. However as an organization policy is common across the system, therefore it would make sense that it only needs to be configured once.

Measuring Performance

As anyone will tell you they have a lot of *information*, it just doesn’t tell them anything. So the issue is ‘context’. There is a huge difference between having a resource that is operating efficiently and knowing if that resource is part of the problem or part of the solution. What 90% of the solutions we have looked at miss is the ability to obtain *information* about contention for a resource and as importantly what is contending for that resource. Understanding that gives an organization a lot of insight into what needs to be resolved, which is where *control* comes in.

Sleeping Service Providers

With very few exceptions, the breadth of the xSP marketplace seems to be asleep at the wheel, while nothing but opportunity is out in front of them. No matter how you look at it, APM is a value added extension that provides revenue and more importantly bridges a huge gap that xSP’s have with their customers. The bulk of the xSP business is based on providing network connectivity, which is the application delivery foundation for organizations. It’s this application / network point of intersection where APM can provide a significant amount of value both as technology that improves and differentiates service, as well as a means to better meet their customers needs. There are many offerings that require little to no capital investment and when rolled out would lower the cost of doing business and/or drive revenue.

Layer 7

An application is something that an APM product can be aware of, but it doesn’t operate at layer 7. Only applications operate at layer 7; period end of story! Applications receive request they process them, pull stuff from databases and return *information* to whomever made that request. All the stuff in the middle is providing a function below that. Many products improve performance by accelerating the presentation layer of applications or improve TCP flows to improve network transport, but again none of them are doing anything at layer 7. That used to be called ‘middleware’ and we all know what happened to that.

V. Industry Trends

There are several trends underway that have an impact on the how an architect addresses a performance problem.

Application Architecture

We have been shouting from our 'soap box' the fact that application architectures are shaping the requirement for the new IT infrastructure. We refer to this as the Web Network Architecture (WNA), but it's important to understand that this architecture is web enabled rather than pure web or browser based. The methods in which users access resources, how they are secured, the methods of transport and how system management is facilitated is based on the applications rather than the network. So in the real world this includes a mix of web, client/server, terminal and file services.

Security

This one can give you a headache. The value of the open architecture facilitates access, but also exposes resources to unauthorized and malicious access. While there are more than enough methods to build protection around networks, resources and applications the important trends are again associated with the application architecture. The integrated management of authentication services with the devices providing not only authentication but IDS, VPN, etc. is streamlining the management and instrumentation of infrastructure security. Starting with the network and moving up to the applications, the implementation of layered security needs to align with the application architecture.

Security is one of those areas, where there is a gap between an analyst and an architect can be significant. As we outline in the WNA, there are right and wrong places to implement various services because of the application architecture, rather than a laundry list of features or functions that should be in a 'security' device regardless of where it is within the infrastructure.

Network Services

Basically, just give me a pipe! Network intelligence needs to be limited to the definition of the network, which is layer three. Route the traffic and support the service level markings of the packets sent; period end of story. However as a provider of networks there are extensions to that core service that compliment the delivery of that pipe. As a side note, we are intrigued about the Infranet initiative being driven by **Juniper Networks** which is focused on instrumented the public infrastructure to support performance and security requirements.

Management

There are a couple of things worth noting on the management front that are just beginning to have an impact on the marketplace. The most significant is the shift toward the structures of WBEM (Web Based Enterprise Management), which is a DMTF initiative for lack of a better term. (Not sure why they chose 'enterprise', but we can get over that if the adoption and delivery is there.) At a high level, this approach provides a common schema for gathering *information* from managed systems using the Common Information Model (CIM) model. On the

back-end it's all about Web Services, which will significantly improve the challenge of integrating systems.

Without a strong 'back-end' on an APM solution, it is an IT tool and not an integral part of the business it ultimately serves.

In respect to actual management products and especially the whole concept of MoMs, etc., we are still looking for something to be encouraged about. The framework companies still have a grip on large systems, but the cost associated with implementing these systems and the weak ROI is beginning to have a toll. One of the biggest crimes here is that the large integrators push these solutions because of the huge services contracts they can wrap around them, rather than the real value to their customers. Now clearly if they can get them up and running they have value, but the dynamics of most systems keep them behind the curve rather than ahead of it. That is why we focus on the innovative companies, who are focused on delivering immediate value and being adaptive to the inevitable changes that come along.

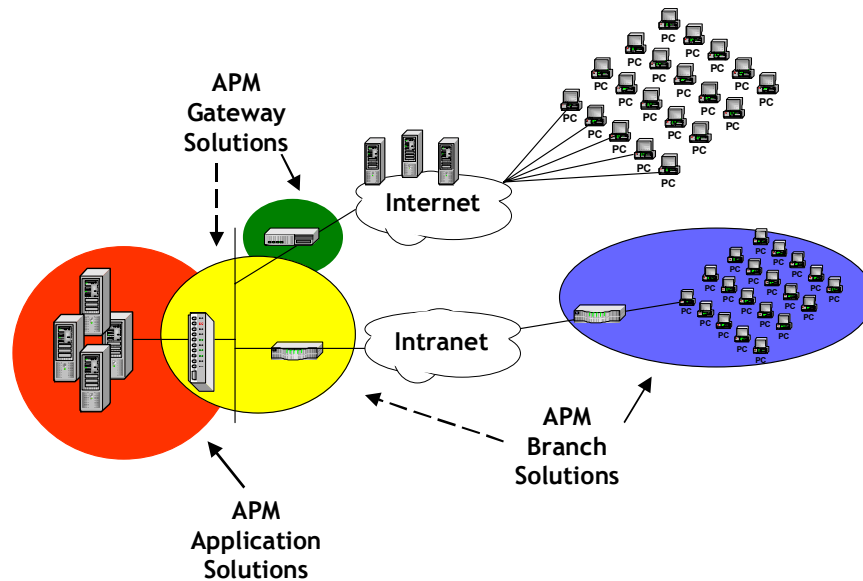
Applications

As noted earlier it's the application architecture that is defining the shape of things to come. While many look at this as a web world, the realities are that there is still a significant amount of client server, terminal and file services that can't or won't become pure web based. However, as being played out in SSL VPN services most if not all of these services can be supported through web portals for managing secure access, etc..

There are many valuable applications that could be deployed if IT was comfortable about how they could manage their performance and minimize the contention that they create. With this introduced into the ROI model, most products would be paid for before they were installed.

VI. An Architectural View

Since we are moving onto the architectural issues involving these solutions we are putting a new model in place for reference. With this focus on architecture, it's all about the solutions that provide *control* services, since the majority of the *information* and *resolution* products are very 'silo' oriented.



APM 'Control' Solutions Model

What we looked at were the issues of supporting the delivery of business applications to users, which simplified the solutions model considerably. Without a doubt there are dozens of variations to this model depending upon the environment, but by maintaining the focus on the application-to-user relationship, the model adapts quite well.

APM Application Solutions

For any system this has to be the foundation for application delivery. With very few exceptions the application domain is made up of a variety of applications and supporting systems. From terminal, e-mail, client/server, web and multi-media, the production environment is diverse.

Therefore an architect needs to assure that this environment is as stable as possible and instrument the domain with a complete APM Solution suite of *information*, *resolution* and *control*. The *information* focus should be at a system level, which is the responsiveness of the applications minus any external complications, such as the network. *Resolution* in this environment is initially triage at the system level, where it is then detail work within a specific discipline. *Control* is similar, in that the initial goal is to instrument the domain for system level service performance and then if necessary drill down into a particular element.

There is also a boundary function that occurs between the application domain and the network, which we labeled as the Application Front End (AFE). The AFE's core functions include load distribution, compression, IDS/IPS, SSL acceleration and TCP Offload. For the most part these solutions have very strong APM *control* services, but haven't figured out their value in the *information* and *resolution* part of the equation.

APM Gateway Solutions

The gateway between the Internet and either the Intranet or application domain has a unique set of requirements to secure and efficiently support sessions across this logical boundary. As outlined later, we advocate the instrumentation of remote desktops, but not all the desktops within a system; therefore there is some important *information* available at the gateway in respect to usage. Additionally, if the business is supporting Internet based access for business applications it's important to pick up response times for those application services. From a *resolution* perspective the most important *information* to know is the performance of the Internet services that support user access.

As far as *control* at the gateway (beyond the firewall), there is often a clear requirement for traffic management, which insures that business bandwidth is available. The availability of bandwidth within the system and on the Internet facilities is enhanced by using Web Filtering / User Content Management, which can keep a lot of non-business traffic off the net.

APM Branch Solutions

For the most part the branch location is still 'odd man out' within many systems. While there has been some movement to improve the bandwidth with less expensive broadband facilities that is about as far as it's gone. In fact we saw one analyst's conclusion was that the 'last mile' was no longer a factor in performance. To that we say 'What the _ _ _?' Organizations buy bandwidth to enable applications, not to make the existing applications work better. Therefore the supply and demand issue is always present which equates to contention. We feel there are some core APM requirements for extending performance to end-users that is uniquely coupled with the AFE or a supporting traffic manager in the data center.

The most critical *information* to have is response times for the business transactions that support branch office services. This should be supported by usage *information* to better understand how the various applications / users consume rather expensive business resources. From a *resolution* perspective it's all about instrumentation at the branch gateway as well as the desktops, since it's expensive and slow to send technicians to remote locations. The core *control* service is traffic management, which deals with the inevitable issue of contention. The more users / applications there are in branch offices the more contention there will be regardless of the network service.

APM Services Overlap

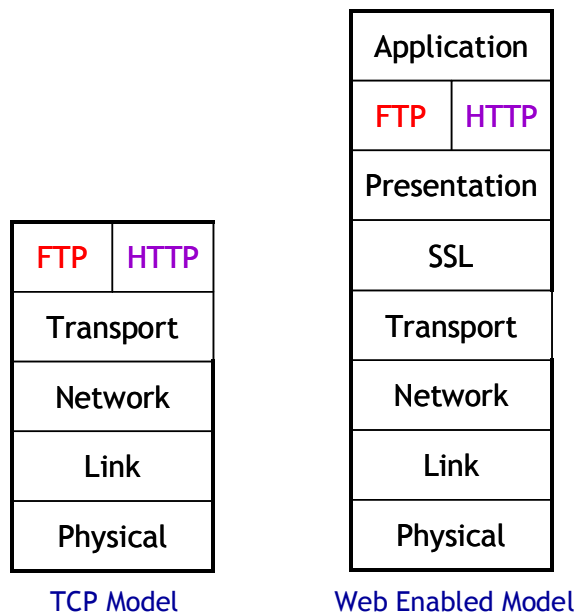
As shown on the diagram there are overlaps of responsibility or functions between all three solution groups, which challenges the vendors as well as the architect. These conflicts and trends will be covered later in the document.

VII. Items of Architecture

Web Enabled Applications

A web enabled application environment changes the traditional model and services of the network. Historically, the treatment of applications was done at the TCP port level, where various applications could be identified and managed by association with a port number.

The web technologies have changed that model and therefore the level of where value added services need to be applied. The breadth of functionality within a browser has enabled the support for virtually any application service (file transfer, transaction, streaming media) to occur within the browser. Additionally, the use of end-to-end SSL encryption further 'blinds' the resources within the network and therefore any device architected on a lower level foundation must compromise the architecture to accommodate services at the upper layers.



With web-enabled applications everything rides on Port 80, where the blend of traditional and newer proprietary application services exchange between hosts. This issue is important for any architect when looking at various APM solutions, since many vendors can't go beyond the port level or classify Port 80 services by URL / Header strings.

SSL

In the context of APM, Secure Sockets Layer (SSL) introduces a number of important considerations. The use of SSL for various applications or transactions is a component of the application architecture driving the requirements within the network.

There are two very important things to keep in mind when architecting performance management around SSL. First of all SSL in its native form (non-SSL VPN) encrypts everything in a packet beyond the TCP header, which significantly limits the visibility a product has into a packet for making performance decisions. This may or may not be important within any particular system based on how much of what traffic is SSL'd. Secondly, once SSL'd a packet can't be compressed, in any shape or form, so if you want to use compression it has to occur pre and post SSL.

SSL VPN

There is quite a bit of vendor enthusiasm behind SSL VPN, which may or may not be matched within many enterprise networks. The combination of the specific vendor implementation and the features required to support any particular environment requires some detailed research.

Basically, a user connects to a SSL VPN proxy, is authenticated against any number of services, such as LDAP, Active Directory, etc. Once authenticated the user is presented with a page of authorized resources that can be accessed.

To provide support for resources that aren't browser based, the proxy downloads an agent (typically Active X) to provide the tunneling services to support client/server, terminal and file services. These tunnels are set up at Layer 3 / 4 so most application and file services can be supported.

VIII. Web Network Architecture

Under the umbrella of what APM Advisors (APMA) is labeling the **Web Network Architecture (WNA)** the assumed goal is to maximize the efficiencies of the infrastructure to deliver a highly responsive and available IT system.

The foundations of WNA are:

Availability - Insuring that when applications are needed; they are available.

Responsive - Application transactions that drive business must perform.

Efficient - IT is an extension of the business and therefore must operate efficiently.

Flexible - Providing support for the wide range of web and 'non-web' applications.

These foundations are also the fundamental attributes of an IT system that has applied Application Performance Management (APM) solutions successfully. Starting with the Application Domain and ending up at the user, there are many challenges in architecting a system that delivers the foundations above, but it's a worthy goal.

As someone wise once said; "You have to know where you're going, so you'll know when you get there." This says it all from the perspective of developing a WNA. Many organizations are in reactive mode in respect to keeping pace with the requirements of their ever evolving application architecture.

The following outlines a few of the challenges organizations face in defining and implementing a viable architecture for their system environment. As any seasoned IT professional will tell you, their challenges are rarely isolated to being technical ones. Much of the job involves the development of consensus and obtaining clear decisions from the business.

IT Policy

In general, every company needs to have an IT Policy for its organization. This is a very important aspect for any system architect to understand prior to getting started with a plan.

Without an IT Policy, plan on spending more money!

APMA has developed a structure for the definition of an IT Policy that is a prerequisite of a WNA.

1. Prioritize Applications - This is a challenge for many organizations, since there can be multiple business units or any number of Intranet / Internet applications that drive the organization. At the very least organize them in no more than three priority classes.
2. Prioritize Transactions - Quite often an application doesn't provide enough granularity in defining priorities. Therefore many applications need to be broken down into the next level groups of transactions within an application that better represents the business.
3. Usage Policies
 - o *Internet* - Quite often the initial step is to understand how the Internet is being used within the organization. Then define guidelines for use that may include many variables, such as time of day, departments, content, etc.
 - o *E-Mail* - Like many of the elements of the IT Policy, the e-mail system has many shades of gray when it comes to appropriate use. Out of many of the elements of an IT Policy, e-mail is typically well defined within many organizations.

- *Desktop* - From a system architectural perspective, many of the other policies have a direct relationship to the use of the desktop. However, there are many uses of the desktop that are isolated to the desktop, but have a direct impact on system integrity.
- *Voice* (in converged networks) - Basically, the use of phone systems have fairly clear policies within many organizations. In converged IP systems, there is a much greater challenge in achieving and maintaining performance; therefore the usage policy needs to integrate into the prioritization of applications and transactions.

Architects Approach

If it's not clear by now at APMA we are all about solutions that apply various *control* mechanisms within an IT system that either drive efficiencies, improve application performance or both. As always, there is a little bit of up-front work required to understand the environment and establish a baseline to determine if you're gaining ground.

There are a number of methods to apply *controls* within a system that improve efficiencies of resources and end-user application service levels. Quite often organizations are in reactive mode when choosing solutions, rather than developing a plan to address performance at a system level.

What we work with our clients on is the development of a plan that addresses today's points of pain and also puts a foundation in place for the future. The basics of it follow the overall solutions model that was extensively covered in the 1st Edition of the APM Paper.

Information

Basically, any organization needs to develop *information* that has context about performance from a system perspective. The problem with most *information* today is that it is developed at an element level or how well is this 'thing' operating. This type of *information* is packaged and sold to a disciple within an organization, where they are watching over their piece of the puzzle.

Fundamentally, the most important *information* any organization needs to start with, is what traffic is going across their network. The goal here is to classify traffic from a business perspective, with the underlying performance metrics.

- What are the traffic types and what percentage of the overall traffic does it represent? It's important to understand that from a branch location, data center and Internet perspective.
- How are the core business applications performing? For a production solution we feel it's important to collect that *information* at the user desktop, but initially this *information* could be collected virtually anywhere.
- What are the variants in this *information* from a time-of-day or day-of-month perspective?

This is core *information* that provides facts about what is going on, rather than assumptions or 'squeaky wheel' factors. Quite often what you'll find is that the squeaky wheels are the tip of the ice burg, so be prepared for some eye opening *information*.

Resolution

The transformation from *information* into *resolution* is granularity of *information*, which can only be accomplished through instrumentation. The deeper the instrumentation is within a system the more detailed the *information* and therefore the more accurate. As we have pointed out, the ultimate measure of performance is response time, but if that response time is bad, what are you going to do about it? We often find organizations pointing at both the application itself and the network it rides over, but pointing is about all they're doing. They don't have enough granular *information* to isolate the problem, which has done nothing but verified the existence of a problem.

Information / Resolution Instrumentation

Fundamentally, good *information* and granular *resolution* capabilities need to be instrumented into the infrastructure. Again, good performance *information* is hard to come by, but once properly instrumented, the ability to resolve problems improves dramatically.

The most important *information* (after response times) that can be obtained within a system is contention for resources. Just follow a transaction from the 'enter' key all the way through the network, in and out of the application domain and back to the user's screen. If it's a clear and un-contended transaction the performance is as good as it's going to get. Introduce contention at any point and it drops a bit. Introduce it at several points along the way and you've got a problem.

Knowing that contention contributes to performance problems is a simple concept, but identifying it and figuring out if that contention contributed to a performance condition is more than a challenge. However, if you are able to isolate the resources where contention results in poor performance, it's pretty clear how to apply *control*.

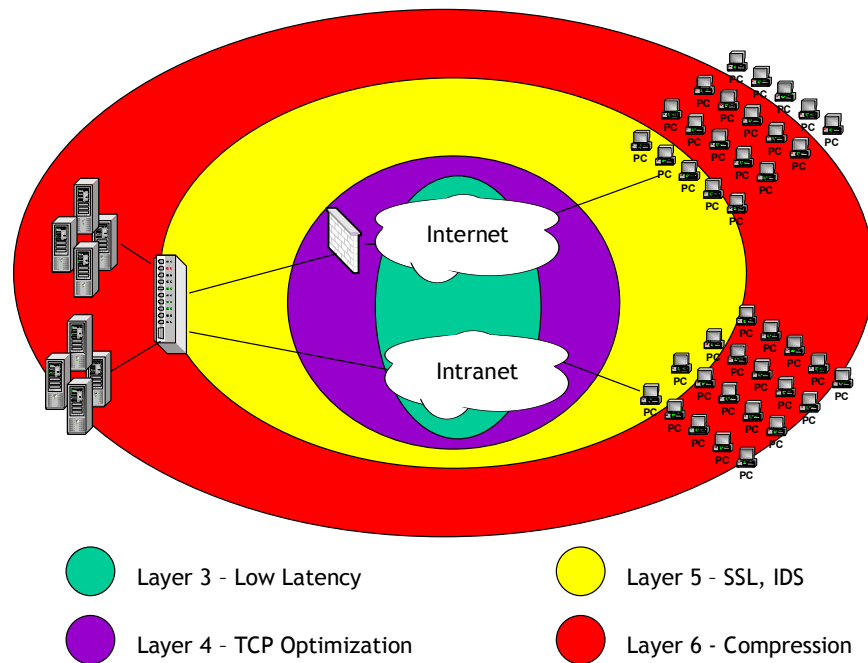
It's our position, that an organization needs the initial *information* to first of all provide the business with insight for policy decisions and secondly to provide the architect with the performance baseline to work from. Resolution is for dealing with anomalies within the system rather than a means to identify areas within the system to fix.

Operating in a resolve and repair mode is like trying to deal with gophers in those arcade games where they pop up randomly and you have to whack them with a hammer. You can whack them all day long, but they just keep popping up.

It's all about Control

Any IT system can benefit from APM *control* solutions. Right out the gate, the ROI on most *control* solutions is in months and not years, so you don't need to beg for the money. Secondly, the TCO (Total Cost of Ownership) is also strong, since most of the *control* solutions are adaptive. This is critical, since the dynamics of application services and demand will push the capabilities of any system.

The following diagram provides a structured look at applying *control*. This structure is being driven by the need to protect applications, *control* access and encrypt flows within Intranets. For the most part the structure is aligned with the various APM services layers of the protocol stack. This is why it's important to understand the difference in the visibility a product has versus the layer in which it operates.



While this diagram is quite simple as compared to the complexities of many systems, there is a simplicity to the diagram that associates functionality with the layers of the protocol stack. There are enough complexities within any system, so if the architecture can be simplified; all the better. Simplification in our book is limiting the number of boxes and insuring that the ones you have to install can provide the suite of *APM information, resolution and control*.

TCP Optimization

The application of TCP optimization has value in the data center and across the WAN, but they are two different services. At the data center, the value is offloading the servers from having to manage connection set-up and flows, which can be a significant amount of overhead. Additionally, the AFE's that typically provide the TCP offload also provide some connection management services toward the users that minimizes dropped connections, etc.

Within the network, it is the management of flow latency over various WAN services and to ensure certain applications like VoIP packets get their slot in packet flows to maintain quality. Additionally on higher latency network services such as satellite, benefit from TCP optimization to reduce the amount of packets required for connection and flow management.

Route Optimization

In our last paper we stated that this is a feature and not a product, which slowly seems to be coming true. Within the context of the WNA there is value in utilizing the most efficient routes through across the Internet, but the value proposition within the Intranet is still a push unless this is a feature extension of an infrastructure solution. It's interesting that the vendors haven't figured out how to package their technology for partnerships and are still trying to drive value propositions directly to customers.

Traffic Management

The core traffic management function we're focusing on here is prioritizing traffic to manage contention for bandwidth. Virtually any product providing an *APM control* function has incorporated some type of QoS mechanism. Granular *control* is achieved higher up in the protocol stack, so this puts IP queuing mechanisms at the bottom of the pile. This is due to the fact that if operating at the transport or session layer, the traffic can be 'back pressured' toward the host. This minimizes the amount of queuing that has to occur and therefore reduces congestion / contention, which leads to dropped packets, retransmissions, etc.. (Now we know the vendors will argue with that, but load creates queues, which equate to delays and there is only so much time to clear queued packets and a fixed amount of memory.)

Most of the traffic management solutions are box based, so the SSL issue applies here as well. The ramifications are quite different though in that traffic can be prioritized based on IP addresses, port numbers, etc. which is more efficient than doing deep layer pattern matches for URL's and headers any way. Naturally, compression also complicates the visibility as well.

Compression

The rule of compression is that it takes two to tango. One compresses and the other decompresses; it's pretty straight forward. Fortunately there has been some proactive work done within the web world and virtually any browser worth using supports Gzip compression, which was incorporated in the HTTP 1.1 specification. Since most media is pre-compressed, the value of Gzip compression is on the HTTP payload, ie. HTML or XML. In most cases the payload reductions are in the 50% plus range, so it's value is quite real. Beyond the reduction of bandwidth, there are fewer packet exchanges required to deliver the data, which adds to the performance improvements.

Now it's important to understand that Gzip is a compression scheme that is applied to HTTP, rather than an exclusive HTTP service. Therefore, Gzip can be applied to any protocol and payload, which is beginning to show up in products. The trick here is that the compression 'head end' downloads an extension to the client for handing off the decompressed traffic to the non-HTTP socket.

The other compression option is site-to-site or box-to-box compression, which is focused on maximizing the use of bandwidth between two points. These products also manage TCP connections on both ends, so packet latency is minimized as well.

As pointed out earlier, SSL traffic can't be compressed, so depending upon the amount of traffic that is SSL'd, the compression value proposition may not be maintained. To get around this, the vendors of box-to-box compression will be moving further back into the data center in a play to become the SSL Accelerator.

Caching

Similar to Route Optimization, Caching is a feature within a product and is an important extension to compression. By caching compressed objects, there are a lot of performance gains on both ends of a connection. The mechanisms for keeping cache up to date blur the lines between static and dynamic cache, since nothing is truly static.

APMA on Traffic Management / Compression / Caching

The APMA bottom line is that if you are serving up web traffic in or outside your system buy some compression. The bandwidth savings alone will pay for the AFE to do the job and you've also taken a step forward in lowering the contention potential within the network. For box-to-box you've got to take a close look at what traffic is going over the link and have a pretty clear idea of the application trends that will change that over time.

Link compression also has a significant value to many organizations, where bandwidth demand is exploding, but budgets are creeping along. The combination of compression, TCP optimization and QoS services are moving these solutions from band aides to core infrastructure status.

From the vendor perspective, there will be some push and shove into and out of the data center, where SSL acceleration belongs. It won't make any sense to see network platforms de- and re-encrypting packets, just so they can either compress or prioritize traffic. This is another example of how the application architecture will dictate what and where APM functions are applied.

User Content Management and Web Filtering

We lumped these together because from an architectural perspective they are the same requirements. As a business, there are certain things that are within or outside of IT usage policy and drawing the line between these things is impractical. We view this as a critical place to start with APM *control*, since once committed to it makes business define just what they expect IT to deal with in the system. If the business policies allow for open access on the Internet, the Traffic Management issues within the Intranet are significant due to the dynamics. If the policy is rather constrictive the dynamics may not be any less, but at least the architect will have the opportunity to manage both ends of the application.

APMA on Content Management and Web Filtering

In this area what is really evolving is the importance of the proxy function for virtually all application services. Just look at any application service, from the browser, instant messenger, e-mail and streaming video / audio. Each one of these application delivery services can be used for critical business services, but also casually. So IT organizations will be looking for methods to enforce business policies within their IT systems and like today's content management solutions, the only way you can get there from here is to proxy the connections. This can be complimented and to some extent executed on the client desktop, but between Active X and Java it isn't that easy to 'lock down' the environment.

Server Optimization

On the surface, this would seem to violate our criteria as a solution, because it's value is limited to a specific device within a system. However, one of the core requirements of developing and implementing an APM solution is having a stable application environment. This stability is challenged by *controlling* platform specific resources, like CPU and Memory in what are chaotic environments. Between run away processes and leaky memory even the most stable environments need the means to protect the platform from the software they host.

Application Domain Optimization

Giving credit where credit is due, this functional category was created by **VIEO**, who has introduced a solution that manages the entire application domain. Through a combination of hardware and software this solution defines instrumentation of the application domain infrastructure. With this instrumentation, all of the 'elements' that comprise the application domain are monitored and managed as a system. This granular instrumentation enables managers to inventory the domain; understand the relationship between those resources and tune the resources for efficiency and performance.

APMA on Server and Application Domain Optimization

Overall, these solutions will become part of infrastructure within production application domains, with the Server Optimization eventually shifting to becoming an integral function in the OS. Uniquely, the Application Domain Optimization solution incorporates the suite of services required to be an APM Solution, which covers *information, resolution* and *control*. Once organizations get the chance to see the comprehensive value a solution like **VIEO's** can bring to their application environment, it quickly becomes an indispensable part of the system.

IX. APM Infrastructure

We posed the question earlier about what defines infrastructure in a WNA, which translates to what should IT organizations focus their fiscal and human resources on? The following outlines those solutions and some basis for those selections. This follows the foundations we laid down earlier in respect to the *APM control* services required for providing secure and high performing applications to end-users.

Application Domain Infrastructure

Within the application domain, there always has been and will be a significant amount of money spent on testing, deploying and supporting applications. We look at the higher level architectural aspects of those requirements, but within the various disciplines of an organization the proper tools are a necessity.

System Domain Management (Information, Resolution, Control)

To our knowledge there is only one solution that can provide a complete APM solution of *information, resolution* and *control* within a web tiered application domain, which is **VIEO Inc.** With the combination of hardware instrumented in the application network infrastructure and software residing on the various web, application and database platforms, their VIEO 1000 offering doesn't leave many holes.

System Platform (Control)

To provide stability of the OS memory and CPU services on various server platforms, both **RTO Software** and **AppSense** are primarily focused on the Terminal Server platforms. Applying these resource management solutions on those platforms helps organizations obtain the scalability and availability organizations depend upon. Both companies have also expressed expanding their platform support, which will provide stronger vendor value within the organizations they serve.

Application Front Ends

We have been consistent in the requirement for fully functional AFE's as opposed to function-specific product offerings. The combination of IDS/IPS, compression, SSL acceleration, load distribution and TCP optimization make these solutions a core part of any infrastructure. There are three vendors leading the charge in providing these high performance and feature rich products; **Array Networks, NetScaler and Redline Networks**. Each company provides the bulk of the core services and has certain competitive differentiating features for specific environments. Each of these companies could improve their *information* and *resolution* value proposition, but there really isn't an alternative to those services. This isn't a negative, but rather the natural evolution of products and companies.

Newer entries from **Zeus Technologies Inc.** and **Crescendo Networks** look promising, so look for updates in the 3rd Edition APM Paper and the next update of our Application Front End Paper. **Zeus's** background in the web server marketplace insures that they understand the web environment and **Crescendo's** extended value proposition into the application fabric will mix things up a bit.

APM Branch Services

To optimize WAN services, the two companies with a clear value proposition and strong vision about their evolution are **Expand** and **Peribit**. We have been particularly impressed on how rapidly these companies have expanded their functional value and are defining what the next generation of Traffic Management is. This has put the incumbents (**Packeteer** and **Allot**) in a tough spot from a competitive perspective.

While we can't talk about specifics the plans both of these companies have shared with APMA make us comfortable that the application architectural issues we've discussed in this paper are well understood.

The other solution that has to be mentioned for supporting branch APM service levels is **Centrisoft Corporation**, who provides a software Traffic Management alternative. This offering provides an ideal alternative, where the sheer number of branches or population per branch challenges the viability of box based solutions. Furthermore, they promote a subscription based pricing model and work closely with Service Providers and System Integrators to fulfill delivery.

APM Gateway Services

The primary gateway service is the firewall, but we draw the line at the network. The extension into protecting applications is a function that belongs with the AFE's from both an authorization perspective as well as IDS/IPS. Moving to APM services, the other core services are Web Filtering, User Content Management and Traffic Management.

When looking at Web Filtering we support the **Cerberian** service model, because of the instrumentation flexibility associated with their agent / managed data base structure, as opposed to the **WebSense** approach. Furthermore, the User Content Management solution offered by **Blue Coat Systems** is rapidly becoming the de facto standard for proxy gateway services. About the only hole they leave at the gateway is peer-to-peer, which then dictates some type of Traffic Management solution, that not only 'grooms' the traffic but can shut down 'out of policy' Internet usage as well.

On the gateway Traffic Management front, this currently leaves **Packeteer** as the front-runner, due to the granularity and adaptive architecture. So as new content sharing mechanisms become available, that the IT organization can respond in blocking or limiting use.

APM Honorable Mention

As outlined earlier, the solutions mentioned above are considered core infrastructure, but each environment may dictate more granular or system specific solutions. This opens up the door to a number of insightful solutions that again may or may not be appropriate for every system, but overall we think they have their 'ducks in a row'.

Compuware Vantage -

We are very impressed with the system level visibility that an organization can obtain with this offering. Our core issue with the solution is the cost of instrumentation. So we have and will continue to advise them to look for methods of leveraging partnerships to deliver their core value proposition. Since the company makes about 50% of their revenue from services, it makes one wonder if the solution is packaged to 'feed the kitty'.

NetQoS -

When we initially looked at the **NetQoS** solution we saw a probe for lack of a better term, but after a look under the covers discovered quite a bit more. The core value is the ability to gather a significant amount of insight about application performance and the infrastructure that supports application delivery. They achieve this through their appliance as well as being able to extract and sift through *information* generated by Cisco's NetFlow (which is a lot to sift through). They uniquely can provide this solution as a product as well as a service offering, which has a natural fit within the xSP marketplace.

Quest Software -

As we mentioned, *resolution* is a matter of detailed instrumentation and once a problem has been isolated to an application or database a specialist within that discipline needs to have the tools to dig into the problem. **Quest Software's** suite of performance management and debugging solutions get to the root of complex platform problems in a number of web tier platforms.

ProactiveNet -

What **Compuware** doesn't get, **ProactiveNet** does, which is that their value proposition is based on their ability to resolve problems based on *information*, rather than their instrumentation. They have moved beyond the management of events to applying a structured method of analyzing trends to resolve performance problems within complex environments. What we have made clear to **ProactiveNet** is that their *resolution* value drops off dramatically as they leave the data center and the further out the weaker it is. Again, this is a product / company evolutionary issue and not a negative on their core value proposition.

Visualware -

Their core value proposition is based on supporting Internet based users of e-commerce services. Through very light-weight software agents, **Visualware** can quickly analyze performance problems on virtually any client desktop. For companies supporting a customer facing web application service, **Visualware** should be a core component of the operations group. In their evolutionary path they need to clean up the product clutter and pricing models, as well as focus on how their solution compliments others for extending value.

X. A Look Ahead

In the 1st Edition of the APM Paper it seems like we had a pretty good handle on the structure of the marketplace, the trends and themes. Prior to that paper and since, we haven't seen anything as comprehensive as that, but after some reflection on how much work that was there is probably a good reason. Since most of the material in the marketplace is published or product focused, the feedback was quite positive in respect to taking a look at solutions versus products. In respect to the trends, it was and is the consolidation of functionality that is driving the industry side of the equation. The instrumentation of the infrastructure theme is solid, but it's definition of infrastructure that's tough.

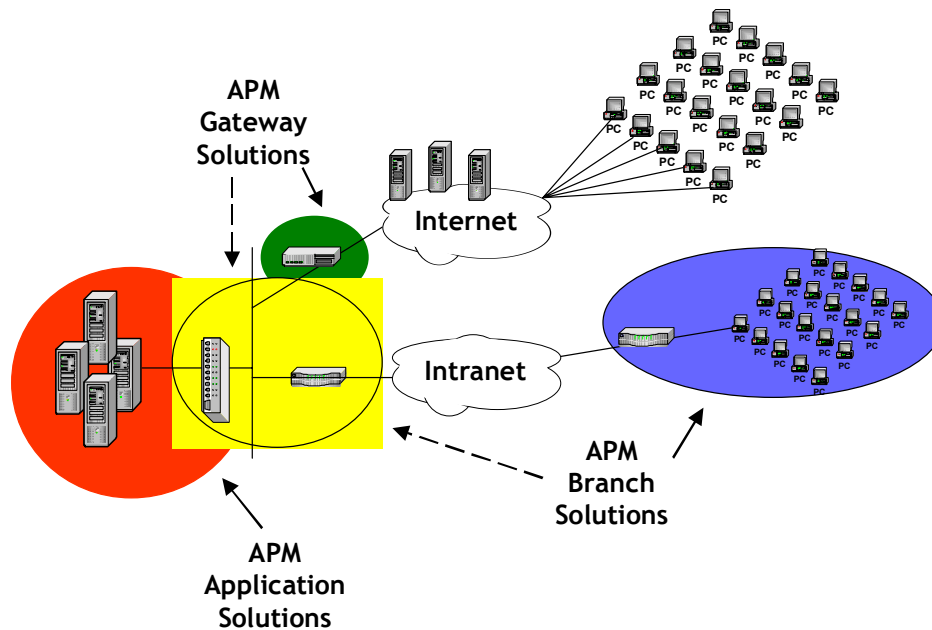
Some of the impressions we had about what will drive consolidation haven't developed yet, but the issue that supported that argument hasn't. We viewed that products requiring centralized policy administration would be natural fits, but the solutions model is having much more of an impact than administration. Thinking through that, it should have been obvious since vendors don't spend a lot of time on the broader issues their customers face and therefore it didn't get much attention.

Therefore we will be picking up the ball on the policy front and attempting to organize many of the vendors providing *control* solutions to put together the foundations for centralized policy management. The focus will be on providing a high-level business policy that would drive a number of the common policy entries that are required for *control* solutions. The simple things are what applications are important to the business and what are some of the basic policies about IT resource use. We are also very interested in developing the methods for providing business some basic *information* about usage to assist in developing cost accounting of IT resources.

APM Control Solutions - Looking Forward

With the foundation we've developed in the document it's time to return to the APM Control Solutions Model to see how the application architecture and industry factors are shaping the direction. As we've stated, the application architecture is the most significant factor shaping the trends in the network providing application performance. The impact of SSL and SSL VPN is yet to be fully understood, but it doesn't take much of it to alter the landscape. This is especially true if the application transactions that are being encrypted are the business critical ones, where response time drives revenue and productivity.

We foresee quite a bit of action on the SSL and compression fronts, with the strength of strategic partnerships taking the upper hand over the next year. Many of the companies in the mix have established their core value proposition and need to be focused on expanding market share rather than breadth of products. Smart partnerships will benefit all involved parties and especially IT organizations attempting to obtain the goals of the WNA.



APM 'Control' Solutions Model

Putting our cards on the table, we see the AFE's as being in the most strategic position in the model, due to the requirement to extend that application delivery services to the end-users. The combination of SSL and compression put AFE's at the right place within the architecture to extend those services. However, it can't be done without partnerships and alliances at this time, since it's the innovators who are delivering the impressive feature / function packages. The big guys have the ability to own both ends of the connections, but have the weakest solutions.

This doesn't count out the Traffic Management solutions, since they have a focus on extending those services out to remote offices, which is something that is just beginning to dawn on the AFE vendors. While the Traffic Management companies haven't really realized the revenue potential associated with the remote branches, they've had enough of a taste to understand the opportunity. However, as we've pointed out, it can't be exclusively a box solution due to the cost (capital and support).

Looking Beyond the APM Horizon

Any way you slice it, what the network was and what it will be are two different things. In order to manage (*information, resolution and control*) systems in the future, it's going to require much more appreciation for the application architecture. This will require that the intelligence within a system be instrumented higher in the protocol stack and as we've highlighted this isn't 'visibility' but functionality. Therefore, the role of routing and transport will be supporting functions to a proxy service that can extend a broad list of valuable services into the fabric regardless of what methods users access applications.

By providing a means to proxy user services as close to the user as possible the efficiencies of the system and performance of critical applications will be dramatically improved. If a branch office is instrumented with an intelligent proxy an architect can extend, filtering, IDS/IPS, compression, SSL and traffic management, which are the core APM *control* services. The important aspect of this from an implementation perspective is that it can be done in either hardware or software.

XI. Conclusion

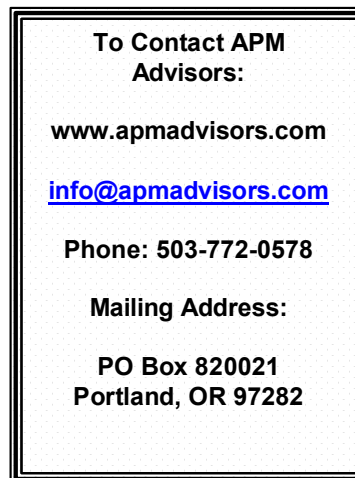
Information, Resolution and Control Instrumentation of the Infrastructure

It was our goal to move beyond the themes and put some structure to the Web Network Architecture (WNA) or an APM Solution. We really don't care what you call it, but we sure hope you can achieve it.

For the most part it isn't as complex as some of the larger vendors make it out to be. We see a lot of real value through some pretty straightforward solutions and some sound architectural planning. If you want to spend lots of money and get less, just put out an RFP to the system integrators. Don't get me wrong we understand their role within many large organizations, but achieving performance goals and maximizing IT resources doesn't have to take a few years and break the bank.

For more insight into the breadth of companies providing APM *control* solutions you'll want to pick up our 2nd Edition APM Paper (APM Control Solutions) paper, which is for the most part this document with classification and highlights on 25 vendor solutions. That version of the document is exclusively available on our web site.

www.apmadvisors.com



Legal

We don't hire lawyers if we can avoid it, so let's keep it simple.

This document contains material, which APM Advisors has sole, and exclusive control over. Any use of this material without the written permission of APM Advisors is just plain wrong. Any information contained in the document was developed with as much care as possible and sometimes even we make mistakes, so just let us know and we'll correct it.

Disclosure Statement

This paper was researched and written without vendor sponsorships. APMA does consult with some of the companies listed within, but we call 'em like we see 'em. The only disclosure worth noting is that the Founder of APM Advisors, R. Lynn Nye Jr. was a founder at Centrisoft Corporation. He left the company early in 2003 and has no active involvement with the company.