Web Network Architecture

Developing a System Strategy
for
Application Performance Management

February 2004

Introduction

As an industry there have been some dramatic enhancements in application architecture, which is leading the path to the New Data Center. Increasingly powerful and flexible application environments are accelerating the realization of the 'web world'.

The focus of this paper will be to take a look at how to extend the power and flexibility within the application environments out to the users and customers it was developed for. 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 New Data Center 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.

A Little History

The foundation of the architectural goals of WNA come from IBM's System Network Architecture (SNA), which was the dominant business IT environment from the mid-70's through the early 90's. The reference is important because there were many things IBM learned about insuring SNA delivered; availability, responsiveness and operational efficiencies.

It's really quite interesting to look at the parallels in the evolution of SNA and the developments in the new WNA. Many of the piece parts of a new system architecture are in place or emerging within the industry. The challenge for IT organizations is weeding through the breadth of offerings and associated vendors' claims to identify valuable solutions. Then once they are located, the more challenging task is implementing the various 'point solutions' and creating from them, a system solution.

Web Network Architecture

Again, the WNA is more of a concept on how to design a system that incorporates the challenges of implementing a system architecture supporting the breadth of web architected applications along with client/server, legacy and other realities of an IT environment. With the impact of media rich content within a transaction based environment, Application Performance Management (APM) is a foundation of the WNA approach.

Architectural Challenges

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 the New Data Center.

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
    - *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.
    - *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.

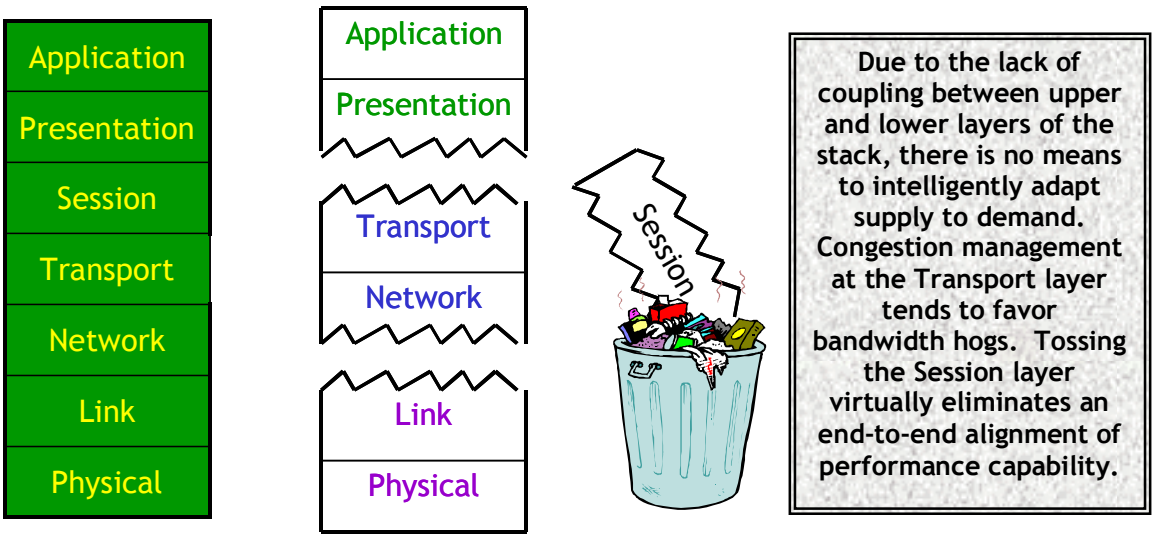
Dude Where's my Stack?

Any system architect can talk someone's ear off, if ask about the functions performed at the various layers of the protocol stack. The good ones can tell you the cause and effect of those layers on each other. Then there are the great ones who have a plan for overcoming the shortcomings of the torn and tattered TCP/IP stack.

The lack of any coupling between the various layers of the stack is the dominant technical issue to overcome in resolving application performance problems in the network. As contention for available bandwidth or application resources grows, so does the problem.

Therefore to achieve the APM Goals of developing *information, resolving* performance problems and implementing effective *controls* requires the *instrumentation of the infrastructure*.

As depicted in the following graphic, the decoupling of the protocol stack in IP systems occurs between layers 2 & 3 and 4 & 6. So what? Glad you ask!



The concept of the godfathers of TCP/IP was pretty straight forward, in that TCP would have the mechanisms to detect and accommodate performance conditions (congestion in the network or between the host). The mechanism works fine where all applications / transactions are treated equal. However, beyond the fact that everything traveling over a network isn't equal from a business perspective, many IP applications can abuse the mechanism and even further compromise the performance of critical transactions.

The lack of coupling between the upper and lower layers of the stack invites chaos since there is no concept of supply and demand. Furthermore, without the Session layer there isn't a structured mechanism to facilitate communication between the layers and systems.

Application Adaptation

Another challenge is the characteristics of applications either as protocols or as implemented. Looking back at the SNA model, the changes in the application environment are significant; hence the 'New Data Center'. With the multi-tiered architectures of many application domains and the distributed nature of information a new approach to 'front-ending' the applications is required.

Fortunately the movement to web-enabled applications is providing a common set of challenges, versus the breadth of application specific requirements introduced in a client / server model. With that said, there is now a new network on port 80 that introduces a significant number of challenges for any architect.

Essentially, once the browser and the application connect, the breadth of application services is significant. Some applications are generic HTTP, while others may include various electronic media or an application specific 'applet' may be loaded and the exchange is anything but HTTP.

Application Domains

The New Data Center application domain provides a very flexible and powerful environment, where web servers can leverage any number of application systems, which in turn can access a broad range of data resources. This architecture has introduced it's own set of challenges in how to manage those resources to achieve efficiencies, while maintaining availability and performance.

The complexity of how transactions are processed and the interaction between the multitude of platforms, applications and data resources is significant. Similar to the challenges of the 'broken' protocol stack, this decoupling of components within an application domain dictates a new approach in *instrumenting* the environment.

For the architect of the WNA the breadth of vendors providing the application domain components and the specific implementation of transaction processes further complicates the ability to obtain meaningful *information, resolve* problems and implement *control*. All of which are important in establishing a solid Web Network Architecture.

The Web Network Architecture

The most important aspect of the WNA is the inclusiveness of every element within the system. From the back-end database servers through to a dial-up user, every element of the system is part of the 'network'. Virtually every application within an organization is either networked or relies upon the network on either a transactional basis or as an extension of application services.

This infers that those elements can be part of the problem, so take the offensive and make them part of the solution. As outlined below, there are a number of options in products that enable instrumentation. The trick is proper instrumentation and leveraging it for the appropriate task.

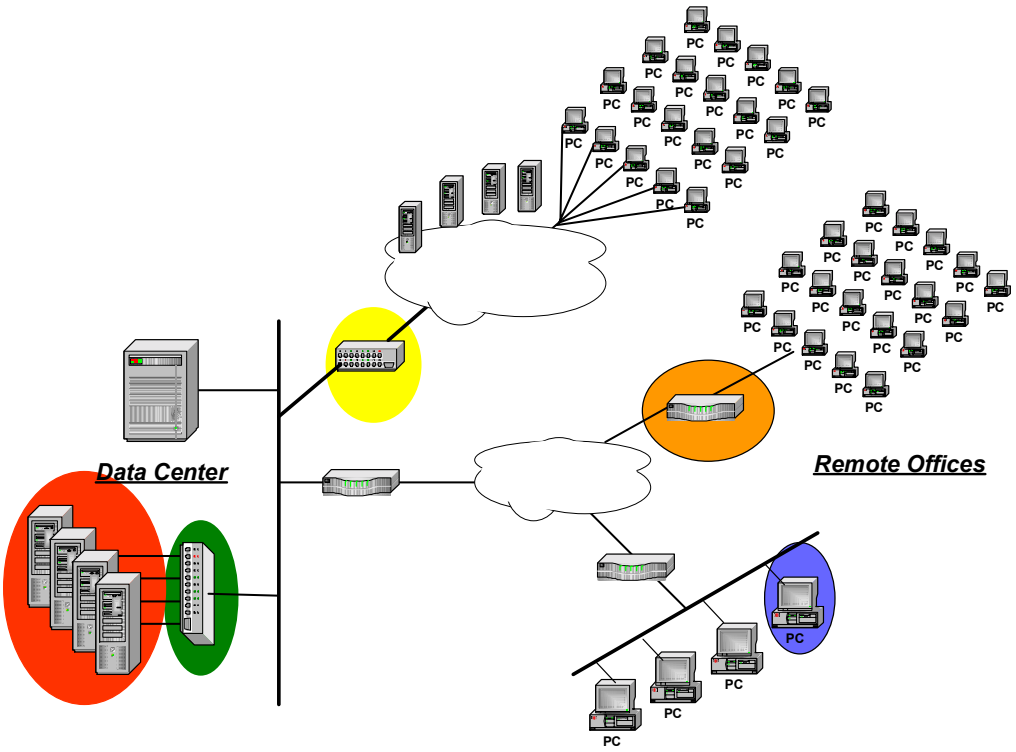
Instrumentation of the Infrastructure

The fundamentals behind the instrumentation of the infrastructure are based on improving the granularity of information, accurate resolution and optimal control. Also while instrumenting the fabric of your IT system, the capital investment and operational overhead of a WNA are significantly lower than overlaying performance specific products or addressing issues reactively.

Information – Resolution – Control

The following diagram provides a high level representation of critical elements in any system, where instrumentation is necessary. Without the instrumentation, the lack of information will lead to uninformed decisions. *Resolution* will be reactive and therefore inefficient. Without *control* the task of managing the chaos will incur cost.

As noted earlier, most of the instrumentation should take place at the edges of the system. This improves all aspects of *information, resolution and control*, by applying the functional services at the source.



Reference the table below to associate the various WNA elements

	Information	Resolution	Control
Platforms	Platform & OS Performance Info	Real Time & Historical Thresholds	Automated Tuning & Resource Allocation
Systems	Web / Appl / DB Performance	Performance Isolation	Automated Tuning & Resource Allocation
Application Front Ends	Application Availability and Transaction Performance		Server & Content Optimization – Appl Protection
Internet Gateway	Application Type & Location Internet Utilization		URL & Content Management – Traffic Optimization - Security
Remote Branch Gateway	Application & User Network Utilization	Network / Link Performance – Contention	Traffic Management – URL & Content Management
Platform	Platform & OS Performance Info	Remote Support Access	Platform Integrity
End User Application Performance	Usage & Response Time	Performance Isolation	Traffic Optimization

Thin not Dumb

There is always a significant amount of discussion around the 'thin client', but it's important to understand that thin and dumb are two different things. Any client within a system should support a number of key functional services in order to support a WNA.

- *Remote Support Access* – Most desktops supported in a system have client software that facilitates remote access by support personnel.
- *Security* – Whether maintaining the integrity of the image or protecting the system from harm, the client is a key component of security.
- *Platform Management* – Being able to account for resources and verifying the platform condition is core to any IT organization.
- *End User APM* – If instrumenting the infrastructure, the client is a very powerful component to have armed within the system.

Last Mile

The access point at a remote office is the first gateway a user depends on and it is also the most common point in a system where congestion occurs. Quite often these locations are provisioned with the least amount of bandwidth within a system and therefore it doesn't take much activity to incur performance-degrading congestion.

Typically the network service to remote locations is provided by one or more Service Providers and the operational conditions on that network are key in understanding application performance. Many organizations rely on traditional SLA metrics, but those fall quite short in having the *information* relative to application performance and provides little to no assistance in *resolving* problems.

Application Front Ends

As outlined earlier the adaptation of an application to the system is another must in the design of a WNA. In general web based applications are not very efficient and addressing those issues within the server is rarely cost effective and manageable.
The new generation of the Application Front End (AFE) delivers a very broad set of functions that significantly improve performance through efficiencies on servers and over networks. The core value propositions vary in each environment, but the following outlines the basic value features in this new generation of AFE's.

- SSL Acceleration
- TCP Offload & Optimization
- Compression
- Caching
- Application Intrusion Detection and Protection
- Load Balancing

Application Domains

Between the performance of specific platforms, to the performance of an individual transaction that involves several platforms, the instrumentation of the application domain is high on the WNA list. As with the AFE's there have been a number of companies who recognized this challenge and have developed solutions that provide insight into complex application environments.

Many of these solutions are information components of a solution, that with some work can lead to problem resolution. A newer generation of solutions entered the market and isolate performance conditions without having to make huge investments or commitments to the mega system management offerings. Controlling these environments is quite an undertaking, but a few creative solutions have recently rolled out that look promising.

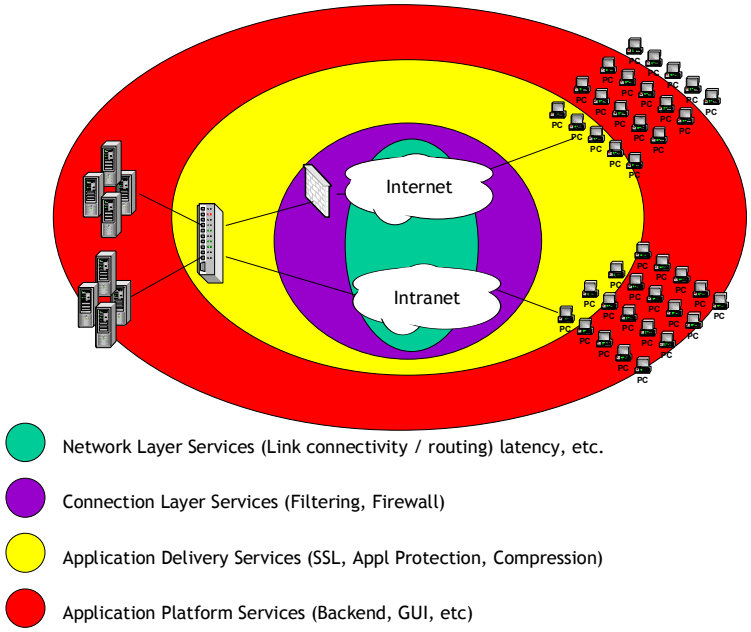
Internet and Intranet

The gateway between the inside and outside world provides incredible value to any IT system, but managing it can minimize the value or drive the cost of the value outside of acceptable business ranges. Fortunately there are a number of solutions that can step in and help an IT organization align the value with the business goals.

From an industry perspective the focus functions and value are control oriented. Whether the need is security, maximizing resources, employee productivity or prioritizing services, the breadth of capabilities is quite impressive. The issue from a WNA perspective is again based on taking a comprehensive view of the requirements and insuring proper instrumentation of services to achieve that alignment. It's not always the bigger hammer looking for nails.

Layered Intelligence

A fundamental of the WNA is the application of various layered functions at the appropriate place within the system. The following diagram provides a visual perspective on where various layered services are best implemented within a system to obtain optimal levels of *information, resolution and control*.



Basically the circles represent the highest level of appropriate instrumentation of layered services. Within each section it's important to include all the lower levels as well, since each physical link or upper layer function is the foundation for every flow across a system.

Fundamentally, various elements within the Network Layer Service circle have the capability of obtaining a significant amount of application layer information, but the context of the information lowers it's value. Knowing what is flowing across the backbone of the network has limited value to the IT group and virtually none to the business. However, knowing which user is using what applications, what their response times are and being able to associate that information to the person logged onto the machine versus it's IP address has real value.

The foundation behind this approach is pretty basic, in that instrumentation done closer to the source will always be more complete. Whether it's the gathering of *information, resolving* a problem or applying *control;* detailed insight is always more complete.

Conclusion

With capital budgets flat and the demand for resources growing, it's critical to develop a WNA for any IT organization. The important thing to do is take a step back; stop reacting and start planning. Instrument the infrastructure with the appropriate level of *information, resolution and control*. With a plan and an architected approach, the critical services IT provides can be aligned with the needs of an organization. Make sure you know where you're going, so you'll know when you get there.