SAP – Network Services Best Practices

COMMUNITY

Network Services Advisory Group

27 September 2006	1.0	

esc@sap.com

http://esc.sap.com

Table of Contents

1 Introduction

SAP is the recognized leader in providing collaborative business solutions for all types of industries and for every major market. SAP's software is the operational lifeblood of over 33,000 global enterprises today. As the internet and network technologies have rapidly evolved, so has the way enterprises deploy and use SAP applications to effectively run their business.

SAP recognizes the importance that the underlying network infrastructure plays in ensuring SAP applications run smoothly and securely in this new environment. Therefore, SAP has worked closely with its network infrastructure partners in an Advisory Group of the SAP Enterprise Services Community to deliver this 'best-practices' document which it feels will benefit all its enterprise customers.

This document is intended to educate the reader on networking best practices and concepts for deploying SAP applications in today's extended enterprise environment. The goal is to help application and network infrastructure departments to better cooperate during the earliest planning stages of SAP deployments to ensure these mission-critical applications are implemented in the most secure, reliable, and highest performing ways possible.

2 General Application and IT Trends

There are a number of important trends affecting the way today's enterprises use and deploy information technology and software applications in order to remain competitive.

Globalization and Mobility: Organizations have established branch offices in far-flung locations around the world. Moreover, the typical employee is now very mobile, needing access to critical business applications from anywhere, anytime, and from any device (laptop, PDA, etc.).

The Extended Enterprise: Corporate boundaries are no longer static and easily definable as the new, extended enterprise has been opened up to include partners, suppliers, customers, and contractors.

Web-Enabled Applications: As the Internet and the World Wide Web have permeated both our personal and business lives, TCP/IP and HTTP have become the de-facto protocols of choice for applications. As a result, most software vendors, including SAP, have migrated their most critical applications to a web-based architecture. In addition, the next generation of web-based communications has also begun to gain momentum – Web Services and Service Oriented Architectures.

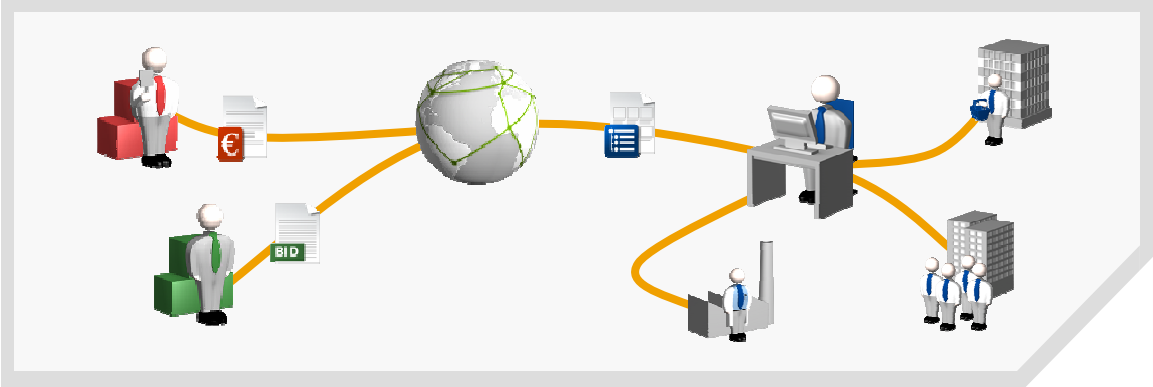
Centralization and Compliance: Although the enterprise itself is now extended and globally distributed throughout the world, the applications themselves are being moved back to the corporate datacenter. This centralization is being driven not only by the need to reduce cost, but even more by new regulatory compliance laws such as Sarbanes-Oxley, GLBA, HIPAA, and others.

All these trends have created new challenges for IT departments both small and large as they try to satisfy the demands of both its end users and the CFO. These demands include providing seamless, secure access to applications, as well as a consistent user experience, with high performance and high-availability for users everywhere at the same or reduced costs.

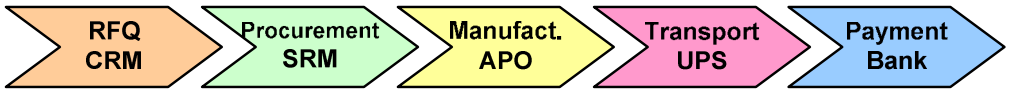
Therefore, it is important that application and network infrastructure groups work closely together from the earliest planning phases of new application implementation projects. This paper will help to give each side an introduction into the other side's world.

3 The SAP Application Environment

The business needs of today's enterprises, in particular the push for enterprises to globalize, are driving the evolution of ERP (Enterprise Resource Planning) software as a platform for transitioning to an Enterprise Service Oriented Architecture (E-SOA). E-SOA is intended to support the business processes of the extended Enterprise, which encompasses a corporation and all its global subsidiaries, business partners and customers.



While classical ERP applications are mostly headquarter-centric software deployments with application servers and end users in a shared local area network environment, E-SOA application software components and end users are spread out world-wide and connected via a wide-area network infrastructure. An ERP transaction is usually executed by a specialized employee, for instance a procurement manager, who creates a purchase order. By web–service enabling such rather atomic "create PO" types of transactions, it becomes easier to integrate applications, chaining many such transactions together to achieve the overarching business process of a 'composite application'.



Classical ERP systems (CRM, SRM, APO) remain the backbone of composite applications, which are made up of web-service wrapped ERP transactions. In addition, business functions from outside partners (e.g. transportation, banking organizations) can be integrated with the same style of standards-based web-service calls.

Coming from this business application-centric view, it is important to recognize that the different parts of a composite application need to connect end-users with different business functions in ERP systems and also connect a variety of ERP systems which reside in one or more different company datacenters and might also reside outside a company's network, for instance linking to a business partner's ERP systems.

Not having all business connectivity in a local LAN adds fundamentally new requirements for supporting network services in a distributed application deployment:

- **Performance**: Transmitting data around the world can be slowed significantly due to unavoidable network latency times and network congestion.
- **Reliability**: The network links and services themselves need to be designed for end-to-end reliability.

- **Security**: Since company boundaries have been expanded to include connectivity to external parties and access to a company's confidential business applications, all communications should be secured.
- **Costs**: The increased overall complexity of IT needs to be considered in estimating the overall cost of application development.

Security and performance are sometimes contradictory requirements in the sense that strong security often has a negative impact on performance and vice versa. Some performance optimization techniques require unencrypted, thus un-secured, data access. The always important questions of how to build a highly reliable IT solution at reasonable costs also exist for E-SOA based solutions.

Therefore, it is essential for any web-enabled and web-service based application implementation project to consider the impact on, and requirements for a company's network infrastructure from the earliest planning stages. It cannot be overstated - application and network groups in the IT department need to cooperate very well for smooth world-wide application deployments at reasonable costs.

4 SAP Application Landscape Characteristics

One major characteristic of E-SOA is that besides end-user WAN traffic there will be a growing amount of application-to-application (A2A), web-services-call facilitated communications over long distances. The business application processing itself will be more distributed due to:

- Operation of multiple corporate data centers throughout the world.
- Communication with the external datacenters of hosting and business partners.
- Small local business applications in "branches" such as retail stores, warehouses and other remote offices, reporting to ERP applications in regional and global datacenters.
- The need to integrate applications from small businesses with those of large enterprises.

The scale of a global network infrastructure for an enterprise can be very large - interconnectivity of up to a dozen datacenters, hundreds or thousands of branches and 10,000 or more users might need to be supported. Besides business application specific network traffic, companies might want to route VoIP, media streaming and other data over their company extranet and the Internet.

Most organizations operate a mixture of current and older SAP releases. Technologies like the SAPGUI front-end still need to be supported and in some instances might even have additional uses in remote sites. In the future, further user interface innovations for Microsoft Office tool integration, flash support and other features will be moved to productive system landscapes. Network products and services should be extendable to future use cases.

A good network infrastructure can help to satisfy the key requirements of web-centric SAP applications for reliability, security, performance and cost effectiveness with features which can not be provided by application software layers alone.

More details are laid out in the following chapters, in particular for the crucial topics of performance (Section 5), security (Section 6), and reliability (Section 7).

5 Performance Considerations

In order to evaluate a solution that both addresses the problem you are experiencing and provides the best fit with your application and network infrastructure environment, it is important to understand the different underlying technologies leveraged by available solutions. For this pur-

pose, it is important to assess the available technologies considering their impact on performance, security, manageability, reliability and availability.

To address the performance obstacles of delivering SAP applications in this new global, distributed environment, there are a number of technologies that have been developed to eliminate or mitigate the different causes of poor performance. As discussed later in this white paper, these technologies are used in various degrees and combinations by point, integrated and managed service-based solutions. The different categories of performance improving technologies include: WAN Optimization, Server and Data Center Optimization, Route Optimization and Data Prioritization.

It is important to carefully consider these technologies because they can dramatically improve the combined SAP application and network performance.

5.1 Transport and Application Optimizations

TCP Optimization: TCP optimization minimizes one of the greatest single causes of WAN-induced latency by reducing the number of round trips required to deliver data. The level and impact of TCP optimizations vary based on the underlying architecture of the appliance or managed service, but they are all collectively focused on eliminating the effects of TCP 'chattiness' by reducing the number of connection set-ups and tear-downs. This is achieved through establishing persistent connections, eliminating delays due to sequencing through pipelining, optimizing throughput by maximizing data packet block size ("TCP/IP window size") and by eliminating retransmit delays resulting from packet loss.

Compression: Compression provides two benefits. First, compression techniques reduce the total size of the data payload which results in an increase in total data delivery throughput and a reduction in data delivery times. Compression technologies include more than simple file compression – for instance, *de-duplication* recognizes repeated data patterns at the block or file level and replaces duplicated blocks with small symbols, greatly improving throughput. Some systems can perform de-duplication across multiple TCP or user sessions, further reducing the amount of data transmitted and increasing efficiencies. Some appliances or managed services can offload compression from the Web server, freeing server resources that would otherwise be required to perform this function. Compression can require significant processing power, especially when the goal is to operate at wire speeds, compressing data in real time without adding latency.

Caching: Caching technologies also deliver dual benefits. First, caching frequently requested objects closer to the end-user reduces the total time to download a page because the network proximity to the object is reduced. The level of improvement in this case will vary based on the underlying architecture and deployment location of the appliance (in the data center or branch office), or at an Internet point of presence (POP), which caches data at locations near the end user. Second, caching frees server resources by offloading the serving of the cached content to the appliance or managed service, thus giving the server more resources to process page requests.

Application Layer Protocol Optimization: In addition to being able to optimize the TCP protocol, WAN Optimization technologies apply acceleration techniques to improve the performance of applications that use chatty protocols and formats such as HTTP/HTML and SOAP/XML. These protocols send data in small blocks that each require an acknowledgement before the next block can be sent. A single transaction may need hundreds or even thousands of round-trip times (RTTs) to complete. As a result, performance drops dramatically across a WAN link with even modest latency — 20 ms or 30 ms — frustrating users and seriously hampering productivity. Prefetching and pipelining data blocks and web objects across the WAN sends as many in quick succession as needed to fill the available bandwidth capacity so that data blocks and web objects are available locally when requested.

5.2 Server and Data Center Optimization

Server Load Balancing: Server load balancing optimizes server performance by balancing traffic among several servers in one data center, using various algorithms to send the next request to the least loaded or fastest responding server to ensure maximum application performance and availability.

Global Load Balancing: Global Load Balancing, which can also be referred to as global traffic management, expands the load-balancing concept by optimizing performance and availability among multiple data centers by routing traffic, based on RTT and availability algorithms, sending user requests to the best performing or most available data center from which the application is being served.

Link Load Balancing: Link load balancing technologies can also be applied to manage traffic for a data center across multiple links to the Internet, choosing the optimal link for transmitting data to a given location based on performance.

SSL Offload: SSL offload technologies perform the CPU-intensive task of encrypting and decrypting SSL traffic for a server, leaving the server to process page requests. This may be an external appliance or a co-processor card installed in the server.

TCP Connection Management: Like SSL offload, TCP connection management technologies reduce the burden placed on the Web server by reducing the total number of TCP connections that must be opened, managed and closed through multiplexing techniques, thus providing the Web server with a greater amount of resources to dedicate to serving requested pages.

Terminal Services/User Interface Virtualization: Using terminal services to provide hosted applications rather than deploying applications to every user's terminal can have substantial performance benefits, aside from the savings realized in management, updating, patching and configuration of applications. Rather than the entire file being transmitted from a client to the server every time the user saves, only screen information is transmitted over the network – the application and data are all on the server to begin with. Both latency and bandwidth utilization can be reduced with this technology.

5.3 Route Optimization and Data Prioritization

Internet route optimization: Internet route optimization technologies improve performance and availability by routing traffic around broken or poorly performing routes to ensure the user's request is served using the best performing and most available path.

Quality of Service (QoS): Quality of Service technologies classify and prioritize traffic based on traffic type to control bandwidth allocation based on the importance of the traffic. This technology is well suited for networks that transport a variety of protocols and traffic types, (e.g. HTTP, NFS, CIFS) where priority should be given to a certain class of traffic to ensure end-user response times or meet Service Level Agreements (SLAs). Tightly integrated with QoS technology is policy management, which provides the ability to easily manage policies and priorities, based on user, application or other category attributes, and apply the policies to the different classes of traffic on the network.

5.4 Security of Performance Solutions

For any performance optimization technology, the security implications of deploying a solution that incorporates the technology should be carefully considered. For example, the deployment of point, integrated or managed service solutions that accelerate application delivery should not in turn introduce new security vulnerabilities. Planners should ensure that solutions are not easily hijacked or susceptible to man in the middle attacks or distributed denial of service attacks.

5.5 Performance Monitoring and Management

In addition, a critical part of any strategy for improving performance is monitoring how well the application is performing and to respond quickly when problems occur, ensuring reliability. It is critical to ensure that both individual applications and the network infrastructure are available and responding properly, since both influence end user performance experiences.

It is important to monitor the performance, availability and utilization of servers and applications at the individual level, as well as in the context of the entire application infrastructure. In many cases, this can be done using native monitoring probes, APIs or reporting consoles provided with the hardware or software, or using management standards such as SNMP to collect and integrate monitoring data into an enterprise-wide management framework. In the case of managed services, customer portals and APIs are provided to achieve a similar level of visibility.

A management framework can also often be used to actively manage servers, applications, appliances and other network nodes, enabling technicians to receive alerts and respond within a single application. Whether this is the case or not, it is not enough to instrument applications, it is also necessary to ensure action is taken when necessary.

It is also important to monitor what end-users are experiencing, both availability and response times of applications, for critical functions, such as logging in, or submitting forms. There are two classes of technology available to provide this perspective: a) synthetic; and b) real-user.

Synthetic monitoring is a proactive management approach that uses agents (typically deployed at various points around the Internet) to mimic end-users by requesting a page or sequence of pages to measure response times and availability against defined thresholds, sending alerts if performance degrades or servers become unavailable. Real-user monitoring captures the actual response time and availability experience of real end users by logging the traffic flow between the end-user and server. Since both technologies have unique capabilities, deploying a combination of the two will ensure optimal visibility into the health of the application and network.

6 Security considerations

Security is a vertical topic which touches both the network layer and the application software layer. In the simplest case, the SAP web application server can provide basic security features. Common SAP web application servers can be configured with the following:

- SSL termination, if security from the data source onward is needed.
- User authentication and authorization, to give each end user access to business processes and data based on their job function or role.
- The composite application layer of E-SOA itself also acts effectively as an access control filter for ERP backbone data.

These security features will address the fundamental concerns for access to the application, but in reality a secured application server cannot protect and guarantee full-time availability of business data as a standalone device. It cannot address industry-wide issues seen by enterprises on a daily basis by itself.

These issues include outside hackers, employees attempting to access unauthorized material, and denial of service attacks (DoS), protocol-based attacks that attempt to overwhelm company systems. Hackers may try to guess passwords, intercept unencrypted communications, bypass application security with attacks that exploit a weakness in the software, find ‘backdoor’ weaknesses including unprotected accounts or admin accounts that have the default passwords and more.

A commonly accepted security strategy is to build up a multilayered defense where network security services and application security complement each other. This security strategy should span

across three domains – the network, the client side end point, and the data center. Security considerations should focus on threats, vulnerabilities, and methods of mitigation. The following sections will address the areas of access, endpoint security, firewalls, and vulnerability assessment and management.

6.1 Access (VPNs and SSL encryption)

Without question, remote access is a critical component of an enterprise application. Mobile users (including internal employees, consultants, business partners, and customers) require secure and easily configured access to corporate applications over the Internet. Connectivity to corporate applications through VPNs and SSL encryption has become the industry standard for secure access to public-facing applications.

For many years IPSec VPNs were the predominant method of secure remote access, but SSL VPN adoption has been steadily increasing. There are substantial architectural differences between SSL and IPSec VPNs. Since IPSec is based on network layer tunneling technology - tunneled traffic is indiscriminately passed from end to end and supports all types of application traffic. This tunneling technique is great for supporting many types of applications, but user transparency is reduced since users depend on the operation and maintenance of an installed IPSec client. SSL VPNs make use of the ubiquitous web browser as the client interface and are considered more user transparent than the IPSec client. Using the web browser allows SSL VPNs to offer virtually anywhere, anytime access without requiring a pre-installed client, enhancing business productivity and reducing client support costs.

SSL VPNs are implemented at the transport layer and can also be deployed as a proxy, providing granular levels of control at the transport layer as well as maintaining a logical separation between the user and the application environment. SSL VPNs also provide granular application, file and URL-level access control and client security capabilities that can enable administrators to dynamically control access to applications based on the identity of the user and an assessment of the client endpoint. This allows enterprises to use SSL VPNs to securely provision access in a diverse range of use cases including employee remote access, partner and customer extranets, and emergency or disaster scenarios. SSL VPNs also utilize the common HTTPS (port 443) protocol, eliminating the need to pass non-standard ports through the firewall. This alleviates many of the common connectivity issues associated with IPSec VPNs.

In addition to securing access to applications through VPNs, SSL encryption can be implemented on web servers, using secure HTTPS rather than HTTP. Early on, servers typically handled all of the SSL connections, but often the CPU processing overhead associated with encryption/decryption and key handling of SSL transactions had a negative impact on CPU utilization, affecting performance and scalability. SSL offload appliances or co-processor cards can address scalability issues and provide encryption for the entire application transaction. For environments that require end-to-end encryption, SSL appliances can encrypt/decrypt connections from the client, but can also re-encrypt the connections to the back-end servers, which is important for other network services which need clear text data for their operation like load balancers and WAN accelerators.

6.2 Client security

Enterprises have several new challenges when it comes to protecting sensitive data. A wide diversity of client end points, such as laptops, PDAs, cell phones, browsers, coupled with highly mobile enterprise data, and a variety of delivery methods, combine to dramatically increase the risk of significant data theft.

In particular, inadequate enforcement of endpoint security has been directly responsible for the success of the recent spate of worms and other fast-moving threats. These security breaches happened because of a lack of direct enterprise control and enforcement of endpoint policies,

including up-to-date antivirus protection, host intrusion prevention, acceptable configuration of hardware and software and a limit on running processes.

Client-side security systems allow security managers to enforce compliance initiatives by determining the location of the system a client is requesting information from, or even taking inventory of a client system and controlling access based on whether the required security software is installed or not. Administrators should be able to set access policies based on a combination of variables. For example, a user may be authorized to have read/write access to information within the enterprise LAN environment, but when the same user seeks access to the same resources from a different environment, such as a public Web browser or an airport wireless access point, the access policy may deny access to sensitive information. This capability can be particularly valuable for compliance-sensitive organizations and those dealing with personal data that can be used for identify theft.

A security policy may also mandate that sensitive information remain in the data center and that users must view the data without being able to create a local copy that could move beyond the control of the enterprise. End-point analysis and access control through policies will limit non-compliant access and can mitigate the risk of exposing sensitive information.

Another approach to secure access and control of enterprise information is to host the end user environment in a centralized data center. This enables the data center administrators to directly control the deployment of endpoint security measures within that environment. For example, using a remotely presented secure browser can prevent downloading of uncontrolled Java applets and ActiveX controls to the client. Not only does this help to control unauthorized data transmission between the client and the data center, it can also be an effective protection from threats that seek to leverage browser “hijacking”, a technique used in phishing and other attacks.

In short, proper end point security begins with access-specific policies that dictate which access methods each user must use for each type of data.

6.3 Firewalls

An effective firewall strategy consists of protecting the various layers of the enterprise application infrastructure, particularly the network layer and the application layer. Protection of the network layer from protocol-layer attacks such as Denial of Service (DoS) and limiting access to the corporate network using Network Address Translation (NAT) and other technologies should be implemented at the perimeter firewall. Firewalls are also being deployed in the LAN to provide departmental security and as part of network access control schemes in conjunction with client security.

Network-layer firewalls are designed to provide access control and statefully inspect IP packets to determine what traffic is permitted to enter or exit a network. Protocols like telnet and ftp, and specific URL accesses via HTTP or HTTPS are only allowed through the firewall when legitimate requests for application access are made. Application-layer firewalls, which can be intrusion detection or prevention devices or network-based firewalls combined with intrusion prevention functionality perform “deep inspection” of IP packets in order to understand the application traffic within. These devices can detect patterns (attack signatures) or anomalies (traffic and/or protocol) created by malicious users or malicious code (e.g. worms) to exploit application vulnerabilities and then can block these attacks before they reach their target. Application-layer firewalling can protect all types of packaged and custom client-server or web-based applications. .

With the increasing importance of web-based software technologies, web-based applications are receiving increased scrutiny

For example, most applications do not validate user inputs to an HTML application. The failure to inspect and validate inputs can render an application vulnerable to a myriad of application-layer attacks, such as buffer overflows or cross-site scripting. An application-layer firewall can stop these attacks.

The most common web application layer security threats are well summarized in the "OWASP Top Ten Project." Further details on these important top ten security threads can be found at:

http://www.owasp.org/index.php/Category:OWASP_Top_Ten_Project

HTML-based applications are not the only targets. Many companies use XML interfaces to provide access to older legacy applications that were never meant for external users and have never been hardened against attack. Applications originally designed to serve only trusted users are now accessed via the Internet by customers and business partners – and by hackers who can easily exploit self-describing XML interfaces to inject malicious inputs into applications.

6.4 Vulnerability assessment and management

Given the frequent discovery of new vulnerabilities, there is a need for a management system to provide assurance that all assets are monitored for policy compliance and for the presence of vulnerabilities, and then isolated and remediated as required, regardless of operating system or availability of security patches. Tasks such as detailed asset inventories, patch management, configuration management, attribute monitoring, and audit logging for network and system administration are implemented as needed to support the core mission of the security administrator.

The benefits of a vulnerability assessment and management system are to minimize risk by proactively providing updated patches and countermeasures to attack propagation. The vulnerability window can also be minimized by log consolidation and log correlation.

7 Reliability/Availability Considerations

Application reliability and availability can be improved at several levels. In addition to purchasing redundant WAN links through a telecom provider, companies can also implement load balancing at the local and global level, as well as application and server virtualization.

7.1 Load balancing

Load balancing is a ubiquitous technology and its deployment is quite straightforward in purely local scenarios. In a local load balancing deployment, the incoming traffic is distributed (balanced) among a group of servers within the same datacenter. Multiple load balancers can also be deployed for reliability and high availability so that if one fails there are others to service incoming requests. Beyond basic distribution of loads, most load balancers can check for availability of specific servers and applications, protect against DDOS attacks and provide other additional functionality to accelerate application performance.

With more and more focus on 24x7 availability, 100% redundancy, and disaster recovery even in case of entire datacenter failures, local load balancing is often not enough. If the entire datacenter network becomes unavailable or if the datacenter becomes inoperative due to natural or man-made disaster, then users will not be able to access their applications. This is where global server load balancing (GSLB) comes into play.

Enterprises maintain multiple datacenters across the globe with full data mirroring between sites. Every datacenter still supports its own local load balancing for the local servers. In addition, a second layer of load balancing service is added across all the datacenters. This GSLB service monitors the available datacenters and ensures that traffic is not directed to a an unavailable datacenter. In case one of the datacenter goes offline, traffic is diverted to the remaining datacenters. With this infrastructure, even with multiple, simultaneous failures, users are still served transparently. In addition, GSLB systems can direct user's traffic to the closest datacenter geographically or to the datacenter that responds the fastest or to a localized version of a site.

7.2 Terminal Services/User Interface Virtualization

Terminal services allows enterprises to run applications on servers instead of local desktops. This can be accomplished through a terminal server or via web-enabled applications accessed via browser. Users access applications remotely and never have to worry about backups, desktop crashes, security etc. They also do not have to worry about installing and managing their own application instances. With application virtualization, hundreds to thousands of users can be served from an application farm without installing applications on individual desktops. Further, when such applications farms are run at multiple datacenters, disaster recovery and high availability become easier. Basically, the user sees a virtual application being served from a remote site rather than a local application on their desktop.

Another advantage to terminal services is that if a terminal session is interrupted, when the user is able to re-connect, their session is available in the same state it was in before service was interrupted. In addition, administration of user services becomes easier, since changes are made to one set of applications on the server, rather than thousands of individual PCs. Bandwidth utilization can also be improved, since the effects of latency for page load and throughput can be reduced by using the specialized protocols available with terminal services.

8 Overview of Solutions

The modern enterprise is a varied and dynamic organization. No single solution will provide the optimum answer in every case. Often, planners will combine SAP built-in services, point solutions, integrated solutions and managed services to come up with the best solution for their network.

The differentiations of needs are often separated by geography, client limitations and security constraints. These dynamic environments force IT to look for solutions that can offer immediate benefit and be deployed with minimal lead time. A hybrid approach can gradually introduce point solutions or outsourced services to an existing environment without commitment to long term investments or overhauling the current system architecture.

8.1 SAP In-Built Network Services

All SAP application servers of current SAP NetWeaver and older SAP releases come with built-in network services for security and WAN performance optimization. These services (gzip compression, connection keep-alive, expiration date tagging of static web content, HTTPS) are software based and make use of operating system and network APIs. As such, they need to be considered for the application server capacity planning because they use some system resources. SAP software-based solutions cannot offer hardware acceleration, TCP/IP stack based optimizations and other improvements which are truly the domain of network devices and network managed services offerings.

Besides the SAP application servers themselves, SAP delivers a few other WAN relevant software components like the classic SAPGUI front-end which features a very lean binary protocol, the SAP Router for adding security to SAP proprietary network protocols and the SAP Web Dispatcher, a software load balancer and SSL terminator.

The strength of these SAP network related features and optimizations is their availability to IT business application groups with the delivery of SAP software right away. In particular for small production, staging, development or similar systems this can be a cost and time saving advantage. For larger production deployments, a closer co-operation between application and network groups is recommended to establish company wide network services governance policies and overall network infrastructure optimizations.

8.2 Point Solutions

Point solutions are focused on resolving very specific application delivery issues. There are several benefits to deploying point solutions in an existing applications infrastructure. Point solutions are often considered best of breed, can be simpler to deploy and manage, and can address very specific issues without the complexity of redesigning the system architecture. Since they are specialized, they can provide a solution to a specific task at a high benefit/cost ratio. For instance, solutions like caching and SSL offload can help maximize application server processing capacity and can be inserted into the existing application landscape with little effort.

Point solutions can help resolve issues in a timely manner because the base management of these appliances often comes in the form of an intuitive GUI and can be deployed with virtually no formal training. As the application architecture continues to expand and offer new services, new requirements can drive demand for more point solutions. In the long run, the deployment of many disparate point solutions may lead to an increase in complexity because of issues with interoperability and management of the different systems.

In conclusion, point solutions offer a wide variety of technologies, flexibility, and can address immediate application needs, although long term planning for managing separate systems should be carefully considered.

8.3 Integrated Solutions

Integrated solutions are multiple technologies that have been converged into a single platform. This platform can become the foundation of an end-to-end architecture for addressing many different business and technical needs. These needs can either be short term and/or long term. The integrated system results in a cohesive platform for application delivery, security and management. There are advantages to combining multiple functions on a single platform.

For example, combining optimization and acceleration features like SSL offload, caching, compression and TCP offload can simultaneously improve underlying network performance issues and reduce load on back-end servers. There are platforms that combine multiple feature sets such as stateful packet inspection, intrusion detection, SSL termination and more.

An integrated architecture can secure and optimize connections from end-to-end. Client connections would no longer have to traverse separate physical devices and IP stacks for encryption/decryption in order to access cached content.

With the consolidation of features and functionality on the integrated platform, the complexity of managing a growing architecture can be reduced due to the management of fewer physical devices. In addition, having one management application that manages several functions reduces capital costs, ongoing operating costs, and training costs. The integration of technologies may affect various components of an application and thus multiple organizations within IT; careful and strategic planning will be required. Sometimes, extended planning and deployment timelines may not always meet the short term demands of a dynamic environment.

In summary, an integrated solution can improve overall performance, yield operational simplicity and efficiency, but may not always be the ideal solution for every type of deployment.

8.4 Managed Services

A third category of solution is managed services. Contrasted to a point or integrated solution, a managed service enables a company to leverage a third party network infrastructure to address the problems of performance, scalability and security for the delivery of Web applications. Like point and integrated solutions, managed services leverage a number of transport and application layer techniques to address the underlying root-cause of performance and scale bottlenecks.

These include TCP optimizations, route optimizations, caching, compression, intelligent pre-fetch, SSL offload, and accelerated SSL delivery.

Given a managed service architecture that is distributed at a global level, a managed service provides the benefits of bundling multiple technologies together and having these transport and application layer optimizations available on a world-wide scale across varying sizes of user-groups (from individual users to large enterprises) without the upfront capital investment.. In addition, some optimization techniques, such as caching, object pre-fetch and round-trip reduction, that target distance-induced latency performance issues can potentially provide a greater increase in overall application performance if delivered by a globally distributed managed service. Furthermore, the primary burden of ongoing administration and management is left to the service provider. However, a core difference between a service and point or integrated solution approach is that a service solution is limited to Internet-facing applications.

9 Key Considerations for Selecting a Solution

In order to assess which solution and technology sets are best suited to meet your enterprise's needs and address the network tasks at hand, it is important to consider a number of different criteria ranging from the technology and deployment model to the total cost of ownership (TCO). In order to aid the assessment process, the following list provides a set of example points to consider when making your evaluation:

Task Analysis

1. In order to meet response time SLAs for your end users would you need to address:
    - Distance induced network latency between datacenter and end-users?
    - Server response time improvement?
    - Or a combination of these and other factors?
2. Are there particular security requirements and policies in your company which need to be followed?
3. Are there specific business events to plan for from a performance and security aspect?
4. Do you have to extend an existing application landscape in small increments, or are you building a new environment from scratch? How quickly do you need to become productive with a new solution?

Application and End-User Environment

5. What is the end-users' work environment?
    - Are they working exclusively in remote offices? Or are they connecting individually to your applications via the Internet?
    - Are they distributed evenly across a wide geographic area or clustered in a few areas?
    - Does your IT control the end-users' laptop or desktop machines? Do you also have to accommodate external, non-employee end users?
6. How are the end-users currently accessing the application?
    - Are they accessing the application over the Internet?
    - Are they accessing the application via an intranet environment?
7. What are the integration requirements to interoperate with existing applications and server infrastructures?

Total Cost of Ownership

8. What are the appliance costs or service fees for a network solution?
9. What are the initial implementation requirements and costs?
10. What are the expected ongoing administration and management costs?

While there could be a number of additional questions spawned from the questions listed above, this list, in combination with the information in this whitepaper on the available technologies and solutions will narrow down the types of solutions to evaluate. For more help on planning out your network infrastructure please contact your preferred network vendor.

10 Conclusion

Challenges for the solutions architect will continue to increase, and to spread into areas other than simply building the application. Solutions architects need to consider the broad issues of globalization, user mobility, the extended enterprise, web-enabling and service-enabling the application, centralization of IT and compliance requirements. In addition network reliability and performance, and security have to be part of any implementation planning and execution program and cannot be left entirely to others – the solutions architect must consider how they will impact each other and application usage.

The most critical lesson to be learned is that properly functioning large-scale applications will require a collaborative effort between software implementers, and developers, network management staff, and the vendors who create the applications and extensions. Contact the vendors listed in the acknowledgements section for more information on optimizing the network for SAP applications. In addition, look for further documents that will come from the SAP Enterprise Services Community in the near future that will address these issues and others.

11 Acknowledgements

The following companies have contributed to the development of this document:

Akamai Technologies, Inc., www.akamai.com

Cisco Systems, Inc., www.cisco.com

Citrix Systems, Inc., www.citrix.com

F5 Networks, Inc., www.f5.com

Juniper Networks, Inc., www.juniper.net

Netli, Inc., www.netli.com

Radware Ltd., www.radware.com

SAP AG., www.sap.com

Appendix A:

For more detailed background information on TCP/IP, please visit:
http://en.wikipedia.org/wiki/TCP/IP

TCP was not engineered for use in the modern network and WAN environment. It uses a verbose query/response architecture that can quickly increase application response times as latency increases, and the large numbers of individual sessions necessary to accomplish a single application task can produce a lot of communications overhead.

HTTP is engineered primarily as a lowest common denominator to facilitate the Web. Compression and optimization of pages are possible, but not the default behavior. Receiving one web page can require dozens of HTTP requests, and hundreds of TCP exchanges.

Optimization can happen on all layers of the overall technology stack. On application side SAP distinguishes roughly between application and the NetWeaver infrastructure layer. The NetWeaver layer in turn uses non-SAP infrastructure layer such as operating systems, databases and, increasingly importantly, the network layers. The network layer itself is commonly broken down in the so called OSI seven layer model.

The application layer of SAP can influence performance just through the design of the business process or transaction flow. Having fewer screens for the end users to go through to achieve a certain task is the most important application design advice. The application layer itself is also important for security in so far as the business logic considers an end user's authorization profile and thus allows access only to such data a user is allowed to work with.

The SAP NetWeaver infrastructure layer provides such network optimizations as can be done in software and which are mostly targeted to optimization of the HTTP protocol.

Network devices roughly fall into the categories of TCP/IP protocol level optimizations and higher network layer level optimizations.