PACKETEER®

White Paper

Delivering High Performance Networked Applications Over the WAN

Intelligent LifeCycle Series

Contents

Perspectives in the Intelligent LifeCycle Series

In this introductory Intelligent LifeCycle white paper, we will cover a general approach that can be used for assessing application performance over any WAN and Internet link. Companion Intelligent LifeCycle Series white papers address specific application and emerging architecture issues:

Intelligent LifeCycle Perspectives	
CIFS and File Collaboration	Voice and Video Convergence
Server and Storage Consolidation	MPLS Transitions
Real Time Transactions	Branch Office Intrusion Prevention
Recreational Traffic	XML and Web Services
Malicious Traffic	Application SLAs

Delivering High Performance Networked Applications

Enterprises Rely On Applications To Run Their Businesses

Business processes have become automated to speed the delivery of information, decision making and execution of all business tasks. The result is that many applications are now needed to run the enterprise. From ERP packages, to CRM, manufacturing software, financial transactions, document collaboration and communications, and numerous IT services, industry analysts estimate that over 200 applications may be found running on a typical enterprise network.

Networks connect the distributed business. Multiple distributed offices, mobile users, suppliers, and customers are tied together via Wide Area Networks (WANs) and the Internet, and these networks have become a critical part of the business infrastructure. Enterprises have invested heavily in routers, WAN service, and Internet connections all in an effort to enable the delivery of applications to many locations.

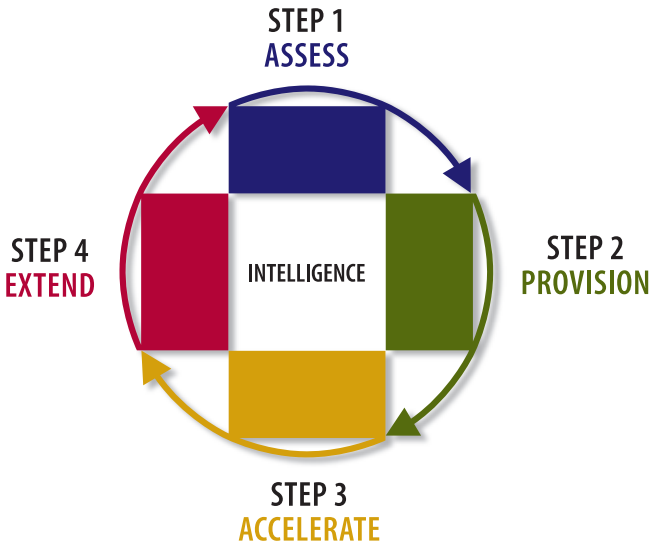
The challenge is to deliver business applications across the distributed enterprise, but too often, these applications fail to perform adequately undermining the productivity of the distributed enterprise. Indications that a problem exists include:

- Slow or unavailable applications
- Complaints from end users
- Increased bandwidth demand and costs
- Complex distributed IT systems

Grasping for a quick fix can result in long-term problems. When an enterprise's help desk starts ringing off the hook with urgent problems, it is tempting to look for easy answers. But quite often tools that fix one problem can have unintended consequences. The wrong traffic, including malicious and recreational traffic, might be accelerated, depriving critical business applications of the bandwidth they need to keep the business running effectively.

The Intelligent LifeCycle—The Smart Approach to Performance Issues

So how do you sort through hundreds of applications, focus on the key issues, and determine what tools to employ to fix the problems? The Intelligent LifeCycle provides a guideline to answer those questions and steer you through WAN and Internet performance issues. The approach is four simple steps that begin with Intelligence and lead to high performance application delivery.



1. Assess: Identify what applications are running on the network, what approaches to take to resolve issues, and continuously monitor performance.
2. Provision: Create network resource policies to align network resources with the business.
3. Accelerate: Apply technologies to enhance performance and capacity.
4. Extend: Create an intelligent overlay that extends and adapts current infrastructure to new and emerging issues.

Assess Constantly. Set up a closed-loop system to constantly monitor performance against established baselines.

Assess

Before you can solve application delivery problems, you need intelligence about the network. Understanding your users, your applications, and your network issues is the first step to resolving current dilemmas and planning for the future. Narrowly focused technology products will often promise much more comprehensive solution sets, only to leave you stranded when the next problem emerges. Assess can be used in a variety of situations:

- Diagnosing existing issues in operational networks
- Deploying a new application or re-architecting an existing process
- Benchmarking service level agreements for applications

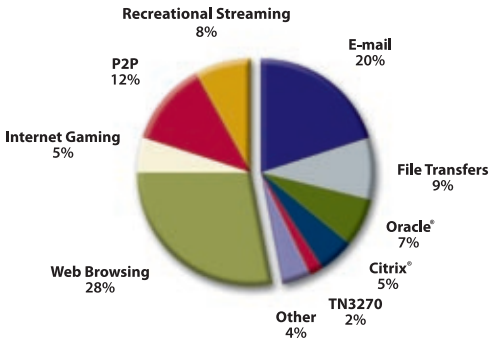
The level of visibility that assessment provides allows you to focus on the one or two most "business-critical" applications (and conversely on the small number of non-business applications that are causing problems).

Armed with this information and insight, you will be in a unique position to recommend and implement optimum acceleration technologies, or further interpret the information and evaluate the impact of the right set of tools to fix performance problems and application delivery issues.

Understand what applications are running on your network

It sounds basic, but 70 percent of organizations do not know what traffic is running on their network. It's not surprising: most network monitoring tools and infrastructure have a very limited amount of application intelligence. These tools—probes, routers, acceleration equipment—map port numbers to applications believing, that port 80 is web traffic. In todays application world, everything works over port 80. SAP is delivered on HTTP along with hundreds of recreational applications. Citrix, another example, delivers dozens of different applications and processes. So port numbers are inadequate for making intelligent decisions. The graph exemplifies shows the shocking amount of network resources that support applications not core to the business.

Bandwidth Consumption



53% of bandwidth being used by recreational applications
14% of bandwidth is "business critical" (Oracle, Citrix & TN3270)

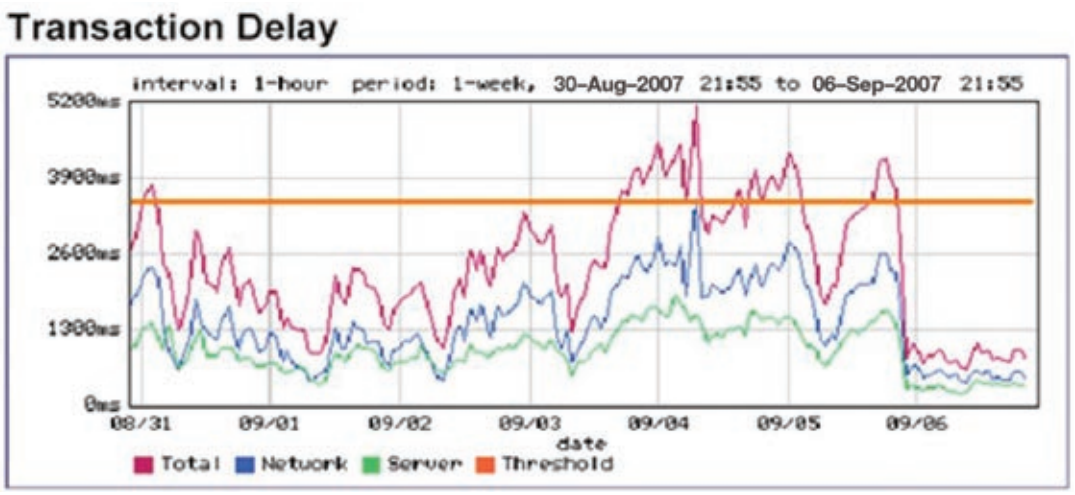
Bandwidth consumption on most enterprise WANs rarely matches business needs, with more than half being wasted on non-business traffic that can potentially impinge on transactions, VoIP, and other critical traffic.

Identify key applications and establish performance benchmarks (user experience)

Enterprises often have several critical applications that, if not available, create a serious disruption to the business. There are many applications needed to run the business, but which are driving revenue—SAP order entry and financial reporting; Oracle CRM; Citrix; Financial Trades or Credit Transactions? For these key applications, it is important to set benchmarks that represent the expected user experience.

Whether during normal operations, in a pilot, or in a lab environment, measure what is considered a "great" user experience for applications. That is, what is the total delay for specific work tasks and transactions? Once established, you can create Application Service Level Agreements (Application SLAs) that you can track against the benchmark. Whether you use Application SLAs purely to alarm and track applications for internal analysis or actually publish them out of the business unit, they can be an effective benchmark for measuring performance. Below is an example of an SAP transaction, with the SLA expectation as a straight line and the actual measured performance over time.

Once established, you can also create alarms to proactively alert to the change in performance.



Diagnose the issues and identify the tools needed to fix them

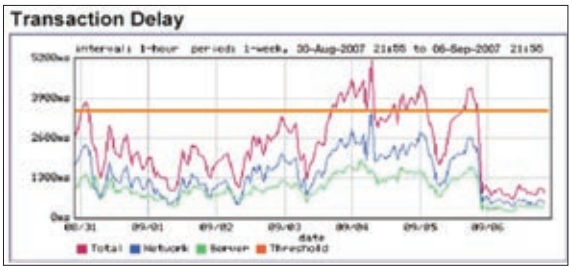
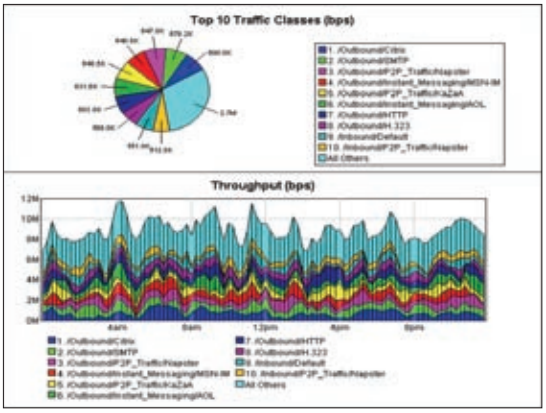
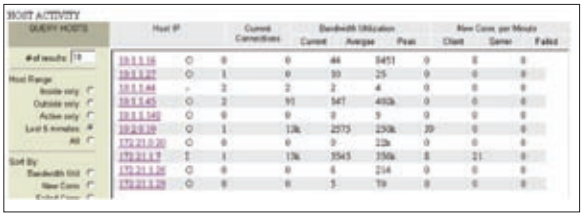
Different applications have different issues. What makes CIFS and file access go fast over a WAN can actually disrupt SAP and transactional traffic. Appendix A provides a more detailed view of applications, issues, and the best tools to fix them.

Common Problems	
Jitter & loss driven by congestion (in COS and out)	**Convergence**
Crowded out by "bursty" collaborative apps & congestion	**Real Time Transaction**
Chatty protocols (CIFS) & bulk traffic burdens WAN	**Collaboration**
Centralized delivery burdens WAN, slows performance	**Server & Storage Consolidation**
Aggressive applications create congestion	**Recreational Applications**
Storms that bring down network – availability problems	**Malicious Traffic**

Constantly monitor and apply intelligent troubleshooting to ensure future success

Today's environments are incredibly dynamic. New applications, changes in usage patterns, and shifts in application architectures create a challenging environment to effectively deliver key applications. An intelligent tool helps you adapt to the changing world, and it starts with ongoing monitoring that can range from simple, automated reporting to active data collection and analysis for major new projects.

Examples of intelligent tools.

Ongoing Auto-Discovery of Applications	• Auto-discover applications and protocols • Break down compound applications into parts that you can treat differently (for example, Citrix real time versus Citrix print)
Tracking Network Efficiency	
Monitoring Application SLAs	
Intelligent Capacity Tracking	
Finding Who is Doing What	

Provision

The next step in the Intelligent LifeCycle creates a network resource policy that aligns business goals with the actual network for both WAN service budget and bandwidth. Each organization's environment is somewhat unique with different applications and different business priorities. Real-time gaming may be recreational to a manufacturing business, or it may be critical to a defense contractor running war games on a network. By establishing a clear set of policies, you are able to demonstrate compliant usage of very expensive WAN and Internet service resources. The best approaches to setting up these policies are as follows. (More detailed provisioning guidance is provided in Appendix A).

Approach	Consideration
Contain impact of recreational and malicious traffic	• Internet traffic is generally backhauled on the private WAN, due to difficulty in securing distributed Internet access. • Usage spikes—music and video downloads (iTunes, P2P), online radio and video (YouTube or news updates), and browsing can disrupt business applications. • Most enterprises want to allow employees to use the Internet, but don't want it to impact the business applications on the network.
Protect availability and performance of key business applications	• Utilization can spike dramatically at different times. • While utilization of the network might average at 60 percent, sudden spikes can disrupt SAP, Oracle, financial transactions and other key business processes. • Guaranteeing critical application 40 percent of the bandwidth at all times, with high priority, you can make sure it's available during usage spikes without wasting bandwidth. • Likewise, this helps contain spikes from recreational traffic, as well as bulky file transfers.
Quarantine infected hosts	• Periodically, worm or virus outbreaks create large amounts of traffic that act as denial-of-service attacks. • As these infected PCs try to infect others, they send out large amounts of traffic that flood the network, rendering applications unavailable. • To quarantine requires: a) the intelligence to identify them, and b) the adaptability to quarantine them. • Packeteer's Adaptive Response provides the ability to do this, while optimizing performance.
Map and validate applications to proper MPLS class of service	• As private networks transition away from Frame Relay and ATM networks to MPLS, there is an increasing requirement to sort through applications and map them to the correct class of service (CoS) to operate on carrier networks. • Most network tools, as cited earlier, base their decisions from basic port numbers that don't correctly identify or validate traffic. • The result is applications getting the wrong level of service across the carrier backbone. • An intelligent system accurately identifies, validates, and provides enough granularity to break down large applications into smaller parts.

Accelerate

After provisioning your network, the third step is to now apply optimization technologies to solutions that enhance performance and capacity. But what are acceleration and compression, and which techniques should be applied where? By definition, acceleration technologies solve the problem of poor protocol design (chattiness) and mitigate network latency—although vendor marketing often tries to obscure the type of approaches used by optimization devices.

To help cut through the clutter, here is a brief overview of common acceleration technologies and their applicability. (For a quick guide of which technologies to apply to what traffic, see Appendix B.)

Technology	Features	Characteristics
TCP Acceleration	• TCP is constrained by needing acknowledgements before sending the next set of data. • Latency with this protocol construct can limit the amount of bandwidth to a flow. Longer latency, lower bandwidth maximum. • Limits are a function of the operating system (different TCP window sizes for different OS's) and link latency. • Acceleration fixes limitation in TCP's ability to scale bandwidth.	**Best for:** a) big jobs that require lots of bandwidth especially over b) long-latency links (satellite and intercontinental), and c) big links of moderate latency (between two data centers). **Limitations:** Low latencies (within continent) won't see a benefit unless with very big pipes or very big jobs (data center replication). Minimal benefit to transactional applications. **Watch out for:** Because 80 percent of traffic is over TCP, vendors will inaccurately claim abillity to accelerate specific applications with TCP acceleration.
CIFS Acceleration	• CIFS is used by Microsoft file access (network shares, Network Neighborhood). • *Very* chatty, resulting in very sluggish file access performance over WAN links with even moderate latencies. • Acceleration improves protocol problems, while compression/ caching minimizes bandwidth demand. • NFS is a protocol with similar problems for UNIX- and Linux-based networks.	**Best for:** File access over WAN links with latency of 60 ms or more. **Limitations:** File traffic only (or applications on CIFS). **Watch out for:** Spoofing acceleration approaches that get disabled by Microsoft security—SMB signing.
HTTP Acceleration	• HTTP has many unique (sometimes very large) objects and often several connections for a given page. • Over high latency, the processing of the page is serialized, creating slow performance.	**Best for:** Web-based applications built with multiple, serial connections and large objects over medium to high latency. Also, good for web based document management. **Limitations:** Web traffic only, requires a degree of latency or *very* large objects to realize benefits.
MS Exchange Acceleration	• Email programs—especially older versions—have very inefficient (chatty) protocols and client structures that freeze with medium latency on the link. • Email acceleration helps fix these problems, while caching can minimize bandwidth from multiple attachments.	**Best for:** Older email protocols, like Exchange 5.5 and 2000. **Limitations:** Re-architecture of Exchange has provided many configuration options, client-side improvements, and improvements to MAPI to minimize protocol issues.

Similarly, there are some *data reduction* technologies that differ in their approach and effectiveness for different applications. Data Compression Technologies (that is, data reduction to minimize bandwidth) compress data or substitute small references to previously seen data blocks in shared dictionaries. Some common data compression technologies and their characteristics are listed below, along with what type of applications most benefit.

Technology	Features	Characteristics
Basic Compression	Like GZIP or LZ derivatives, use static techniques to reduce data.	**Best for:** HTTP browser compression. **Limitations:** Process intensive, lower gains in compression.
Real-Time Dictionary Compression (RAM Based)	Stateful dictionaries dynamically build more efficient compression.	**Best for:** Real-time, latency-sensitive traffic. **Limitations:** Dictionaries built across a wide set of applications can be "diluted," lowering gains. Application-specific dictionaries can increase efficiency.
Byte Caching (disk based large dictionaries)	Large dictionary caches, often multiple gigabytes for large 'chunks' of data.	**Best for:** Big bulk applications—large files, large email attachments. **Limitations:** Need repeated access of same or similar files to get benefit differential. **Watch out for:** Minimally useful for transactional applications. High-speed disks create cost disadvantage to real-time compression.
Web Object Caching	Cache HTTP web objects, usually deployed at Internet access points.	**Best for:** HTTP-based applications. **Limitations:** Browser caching and SSL minimize gains.

Extend

The world doesn't seem to get any less complicated, especially when it comes to networked applications. New applications, more demanding users, changing architectures, emerging threats, new ways to approach old problems—all are challenges and opportunities.

Deploying an intelligent system focused on delivering high performance applications over the WAN and Internet allows you to adapt your infrastructure to a dynamic world. For this reason the fourth and final step in the Intelligent LifeCycle is to extend an intelligent overlay across your network to minimize new technology risks, extend the life of infrastructure, and ensure high performance delivery of applications to every location.

Distribute services to enhance delivery, re-engineer key processes and lower costs

Server consolidation projects aim to simplify the branch office and reduce management headaches and costs. Besides file access performance problems, however, server consolidation can introduce a number of issues that need to be addressed in order to achieve the desired impact. The table below lists key services and how an intelligent system can provide a more effective approach to application delivery.

Service	Benefit
Application and Network Services Deliver locally, with no additional server	• Print services—avoid WAN impact and speed performance • Microsoft System Management Server (SMS) • DNS and DHCP delivery—delivered locally • Domain Controller services
File Services Beyond Acceleration	• Manages file versioning and maintains data integrity • Allows for fast browsing of file shares • Enables offline access when WAN is down
Virtual Backup Services	• Re-engineer the backup process with a full fledged Wide Area File Services • Maintains constant replication of document changes to the central site, doing away with traditional backup
Virtual Storage	• Provide access to terabytes of storage at the branch with no performance penalty • Eliminates the need for storage capacity planning and provisioning at the branch office
Distributed Security	• Delivers Intrusion Prevention System (IPS) to the branch office in conjunction with WAN Optimization device for effective ROI • Maintains Microsoft security for file access (SMB Signing) • Provides extensible platform for new security services required at branch office in future

Deliver Applications Effectively to Every Location

To intelligently extend your network's effectiveness, applications must be able to be delivered everywhere. That means everywhere—to the toad Warriors in the field, to branch and remote offices across the globe, to the high-capacity data centers wherever they may be. Product flexibility is the key to meeting the needs of the enterprise today…and in the future. The following table describes typical solution sets to look for when assessing vendors.

Solution Set	Purpose
Client Software	• For mobile and home office workers where appliances aren't practical • Laptops as well as emerging use of PDA and smartphones
Server Software	• Take advantage of servers in field • Existing servers can be converted to WAFS and intelligent services platforms
Scalable Appliance Family	• Lower cost appliances for deployment at branch consolidating entire solution set • Highly scalable solutions for large capacity data centers and main sites
Central Management System	• Provide ability to manage distributed, intelligent solutions from a central location • Achieve and maintain low TCO (total cost of ownership)

Deploy a Flexible Overlay

It's hard to predict the future. Shifts in the business requirements, new applications, and changes in delivery architectures—both network and applications—challenge IT to effectively deliver applications over the WAN and Internet.

An intelligent overlay helps your infrastructure adapt to the changing world of networked applications, extending the life of your infrastructure, removing risk, and lowering costs. An intelligent system is characterized by:

- Intelligence to recognize applications, new and old
- Flexibility to "plug in" new advances and technologies for continuous improvement
- Modularity to meet the changing needs of your environment
- Scalability to manage a large distributed installation with low TCO

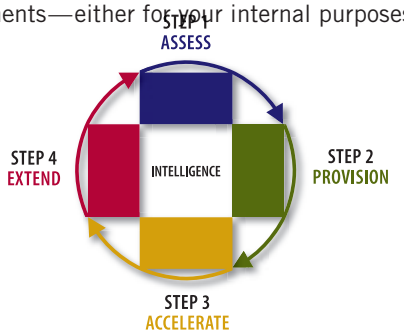
Assess (Constantly)—A Closed-Loop System for High Performance Delivery

The dynamics of users, networks, and applications change constantly. Best practices call for continuous monitoring of key applications in order to track changes in the environment and should be incorporated into a comprehensive WAN Optimization strategy. A few key concepts can help you proactively manage application delivery and fit well into ITIL and other management frameworks.

Application Performance Baselines and Monitoring—Service Level Management

By establishing baselines of expected performance, you can measure material changes in the operating environment and more consistently maintain expected service levels—Application Service Level Agreements—either for your internal purposes or to publish to business units.

- Identify key business applications to track.
- Establish key metrics and expected thresholds as internal SLAs for specific applications to notify when conditions change—for example, SAP Total Transaction Time of 800 ms.
- Set an alarm—into your SNMP trap system, e-mail, or other—to notify you when expected Application SLAs are violated.



Event and Fault Management

As you track performance to service levels, it's important to incorporate them into your event and fault management process:

- **Alert:** Feed departures from expected service level baselines into event management systems. Alarms sent to existing SNMP systems are among several alternatives to accomplish this, but are dependent on tools that provide that level of intelligence.
- **Troubleshoot:** Once the alert is made, troubleshooting processes should be available to isolate problems, as discussed in Assess.
- **Resolve:** Employing the right tool sets among many possibilities is a key benefit of a more comprehensive WAN application delivery system. Sometimes, however, resolution points to other places—router configurations, server infrastructure, or application design issues.

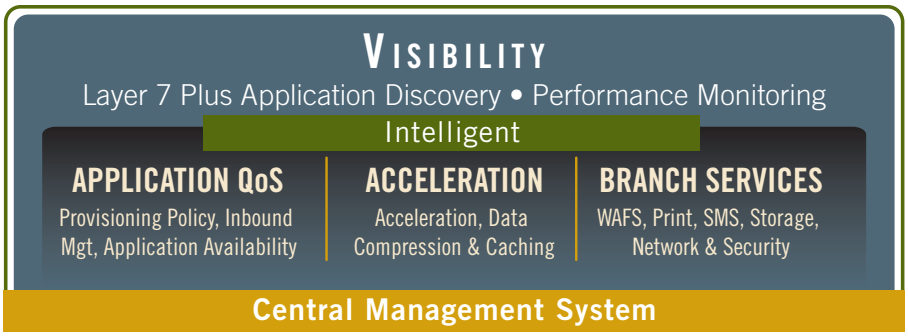
Capacity Management

Demand for resources continues to increase as more users and applications are delivered over the WAN. Leveraging an intelligent system provides a more informed decision process, as well as more alternatives:

- **Application-intelligent view.** Utilization averages are deceiving. Look at peak utilization of the network, but remove recreational applications. This gives you a better view into potential performance issues.
- **Reprovisioning:** You can regain, often, 20-50% of peak bandwidth just by containing recreational applications.
- **Data Compression** can provide quick capacity improvement without new circuit provisioning.

The Packeteer System—The Intelligent Overlay

The Intelligent LifeCycle is an approach to analyzing, planning, and executing an effective strategy for high performance application delivery over the WAN. Packeteer delivers a complete system deployed through the Intelligent LifeCycle, to ensure a great user experience for every application at every location.



With Packeteer, intelligence can be made simple. Finding hundreds of applications may seem difficult and complex but with its intrinsic application intelligence, the Packeteer system automatically discovers and measures all applications on the network. Simply plug in the appliance—inline or out of line (tap, span or mirror), enter basic IP configuration information, turn on auto-discovery and let it run. As traffic flows through, Packeteer automatically classifies and measures over a hundred statistics per application, so that, after three to five days, a wealth of detailed yet easy-to-view information is accumulated to serve as the basis to understanding WAN performance.

Application classification must play a significant part in intelligently applying WAN application acceleration technology and form the basis of your choice of technologies to fix your problems.

Especially with confusing claims and counterclaims regarding various "application acceleration" technologies, it is tempting to look for an easy answer. But, in many cases doing so will unknowingly risk eroding the performance of one class of applications when accelerating others.

Without a comprehensive application classification and visibility strategy, organizations cannot determine the broad effect that isolated application acceleration can have on network-wide performance.

Appendix A: Provision Your Networked Applications—Intelligent LifeCycle

Application Type	Application Examples	Provisioning Dependent on Ability to Identify and Validate Traffic
Convergence	IP Telephony, VOIP, Video Conferencing	• **Basic:** Guarantee % that maps to expected call capacity, prioritize top priority. • **Advanced:** Per call bandwidth guarantee for amount needed for specific voice or video codec used. • **MPLS CoS:** Real time class of service (class 1).
Real Time Transactions	ERP & CRM, SAP, Oracle, Financial Transactions, Citrix	• **Basic:** Guarantee 40% of network, burstable at second priority (or top priority if no IP Telephony or video conferencing). • **Advanced:** Per session bandwidth for key application processes. • **MPLS CoS:** Class 2—critical traffic.
Collaboration	File access, email, rich media content	• **Basic:** Prioritize to medium priority, so it doesn't impact. • **Advanced:** Per call bandwidth guarantee for amount needed for specific codec used. • **MPLS CoS:** Real time class of service (class 1).
Server & Storage Consolidation	File access (above) plus print, SMS, DNS/DHCP	• **Basic:** Deliver locally with services enabled WAFS solution. • **Advanced:** Guarantee 10-15% of bandwidth at medium-high priority. • **MPLS CoS:** Class 3 (in 4-class system) if not delivered locally.
Disaster Recovery	Storage mirror, database synch, backup	• **Basic:** Prioritize high for datacenter-to-datacenter. • **Advanced:** Enable time-based policies for specifically scheduled jobs. • **MPLS CoS:** Class 2 on dedicated links.
Recreational Traffic	iTunes, radio YouTube, IM, P2P, browsing	• **Basic:** Guarantee 0% of link, burstable to 25% at LOWEST PRIORITY. • **Advanced:** Squeeze P2P to 2 Kbps—allows it to work, renders unusable. • **MPLS CoS:** Best effort (lowest class).
Malicious Traffic	Worm, virus, DDOS	• **Basic:** Alarm on spikes in new flows per minute and failed flows. • **Advanced:** Automatically move offending hosts to "penalty box" of 1 Kbps pending further action. • **MPLS CoS:** Best effort for "Penalty Box" traffic.

Appendix B: Intelligent LifeCycle Quick Diagnosis and Tool Guide

Application Group	Examples	Common Causes of Performance Problems	Key Technologies & Impact	Also Consider
Convergence	IP Telephony, VOIP, Video Conferencing	Jitter and loss driven by congestion Oversubscribed class of service (MPLS) or aggregate queues	Per call QoS removes risk of oversubscription rates in aggregate queues and CoS Traffic validation ensures correct traffic is marked into class of service (CoS)	Compression technologies make more room for IPT and Video Conferencing traffic
Real Time Transactions	ERP & CRM, SAP, Oracle, Financial Transactions, Citrix	Transactions crowded out by 'bulky' collaborative applications and congestion TCP responds to congestion by clamping down on bandwidth	Simple priority ensures bulky applications don't interfere Per session QoS is more advanced if number of users is growing and oversubscription is at risk Application sub-classification to focus on key transactions (e.g. Citrix real time vs. Citrix print)	Protocol acceleration will help on high latency links like intercontinental and satellite links Note that web object caching in browsers (for HTTP) helps equation
Collaboration	File access, email, rich media content	Latency and protocol design. CIFS and NFS are VERY poor over links with medium latency (60 ms). Big files accessed over WAN create bandwidth burden Email attachments to many users require large amount of bandwidth	CIFS protocol acceleration for Microsoft file access (NFS acceleration for NFS) Bulk compression or byte caching for large file sets that are repeatedly accessed (pulling same file multiple times or email attachment sent to multiple people in office) QoS to ensure large transfers don't block out transactional traffic	Email protocol acceleration for Exchange 2000 (Exchange 2003 eliminated many protocol issues)
Server & Storage Consolidation	File access (above) plus print, SMS, DNS/DHCP	Key services become centralized at data center, causing delay and WAN bandwidth burden	Branch delivery of key services without burden of server at branch File management services to enable offline access and maintain versions/integrity QoS to ensure performance of these infrastructure services	Microsoft SMB signing disables many acceleration devices Navigation of central file shares is also important
Disaster Recovery	Storage mirror, database sync, backup	Large jobs constrained by TCP, unable to grow to bandwidth rates to finish jobs	TCP acceleration allows for higher bandwidth rates than standard TCP allows QoS to ensure large jobs don't crowd out other applications	QoS to ensure the Compression of uncompressed traffic (block mirroring)
Recreational Traffic	iTunes, radio YouTube, IM, P2P, browsing	Use of recreational applications crowds out business traffic at wrong times	Intelligent traffic classification and validation to identify recreational traffic QoS to contain impact of recreational traffic	Allow usage of Internet, just don't block it
Malicious Traffic	Worm, virus, DDOS	Outbreak brings down network, makes applications unreachable	Ability to identify and isolate attacks in progress Response and QoS capability to contain and quarantine infected hosts	Tools that are affordable at branch but feed central IDS/SEM apps

Packeteer, Inc.
10201 N. De Anza Blvd.
Cupertino, CA 95014
United States

T + 1.408.873.4400
F + 1.408.873.4410
www.packeteer.com