OptimOSS Ltd
OptimOSS House
Victoria Rd, London
UK N22 7XB

Tel +44 (0) 208 888 8280
email info@optimoss.com
Fax +44 (0) 208 889 9723

A WHITE PAPER INTRODUCING THE PERISPHERE ARCHITECTURE

Only when network elements work seamlessly together can IT succeed in providing LAN-quality application delivery across distributed enterprises. WAN optimisation must solve issues related to bandwidth congestion and the redundant transmission of large files, high latency and loss, application contention for access to the WAN, inflexible transport options, and an overall lack of insight about what's happening over the WAN.

Using the PeriSphere Architecture, OptimOSS can deliver an integrated approach that provides the full range of elements needed for effective WAN optimisation, all packaged in an easy-to-use single platform that delivers:

- Increased WAN capacity via patented 'plug and play' hardware devices
- Sequence caching that reduces 99% of repetitive data in data transfers
- Latency reduction using Packet Flow Acceleration (PFA)
- Bandwidth management using QoS and bandwidth allocation features
- Path optimisation with Policy-Based Multipath
- Visibility via WebView and CMS (Central Management Software)

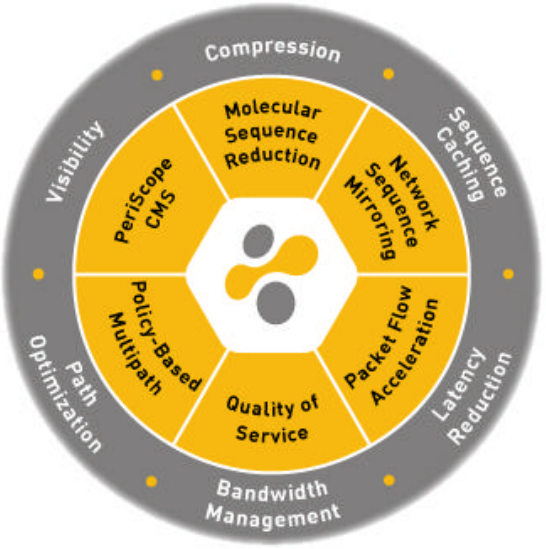
Ease-of-use and operational scalability features include routing protocol snooping for topology learning, continuous communications among the devices to dynamically update the entire network, and deployment flexibility to implement the devices without changing the existing network or current Service Provider.

For application performance across the WAN to improve, the WAN fundamentally has to behave more like the LAN. Making that improvement requires a combination of increasing WAN capacity through compression and removing the limitations of distance by speeding up the underlying transport protocols. A PeriSphere deployment immediately improves:

- Application rollout and webification
- Data centre consolidation
- Disaster recovery/backup
- Regulatory compliance

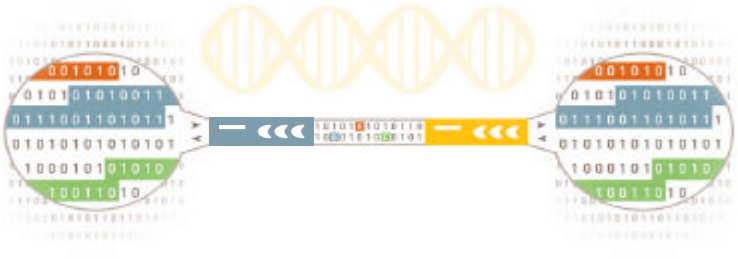
What is PeriSphere?

PeriSphere is an integrated feature set for optimising WAN's and comprises six key elements in a single platform, details of which are found in the following sections. These features are delivered via a choice of two hardware platforms, Sequence Reducers and Sequence Mirrors, over which the patent-pending software tools run.



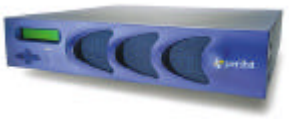
1) Increased WAN Capacity

The Sequence Reduction System (SRS™) software is the patent-pending compression algorithm at the heart of all the PeriSphere hardware devices that delivers instant increased bandwidth over existing WAN links. Having its roots in DNA pattern matching, SRS™ recognises repeated data patterns and replaces them with labels, dramatically reducing WAN transmissions. SRS™ operates in memory and its dictionary can store hundreds of megabytes of patterns. The software integrates the range of technologies needed to optimise WAN links, including compression, sequence caching, latency reduction, bandwidth management, path optimisation and visibility.



The SRS™ software powers both the Sequence Reduction (SR) and Sequence Mirror (SM) product lines. All devices operate bi-directionally, providing both reduction and assembly processes within each device. The SR delivers compressed output in speeds from 128Kbps to 155 Mbps and supports a range of connections to other SR's and SM's. The SM is focussed on enhancing throughput of large repeated data patterns and the SM-500, which provides compressed output to 20 Mbps includes on-board hard drives to increase the capacity for stored data patterns.

Sequence Reducer



Sequence Mirror

SRS™'s reduction capabilities benefit a broad-cross section of application types. Depending on the application mix, businesses typically gain a two- to four-fold increase in capacity on their existing WAN links, while some enterprises have seen as dramatic an improvement as a ten-fold increase. SRS™ effectively reduces both short, chatty applications such as Citrix and HTTP as well as larger data patterns, such as Word files or common images in PowerPoint presentations. Because SRS™ is dictionary based, it is able to identify repeat patterns even when they are separated by large amounts of other data. Eliminating this redundant non-informative data, which in many cases consumes up to 90% of network resources, generates an ***immediate*** and dramatic increase in effective WAN transmission capacity.

SRS™ ground-breaking contribution is its efficiency – even though its memory of repeated patterns is very large, the algorithm adds very little latency, typically around 2 milliseconds. Another key attribute of SRS™ is its ability to maintain its compression capabilities even on very large amounts of bandwidth, scaling to support OC-3 links. These attributes are traditionally mutually exclusive, when using compression techniques such as Lempel-Ziv, its derivatives or Predictor.

2) Sequence Caching

Complementing the award-winning SRS™ technology above, and delivered together in a single device, PeriSphere offers a new compression technology called sequence caching.

With sequence caching, the hardware platforms increase compression rates by recognising repetition across much larger data patterns than SRS™ uses. Unlike SRS ™, which operates entirely in a memory based dictionary, this feature, called Network Sequence Mirroring (NSM), relies on an embedded hard-disk, found in the SM devices, to store the longer data sequences, store them for longer periods of time, and replace them with a flag during transmission over the WAN.

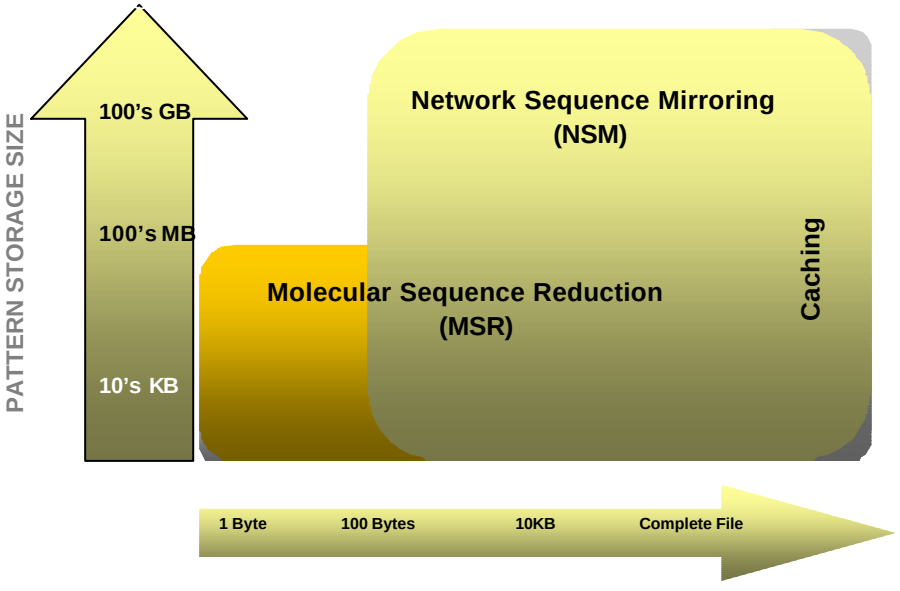
By operating on large patterns of data, NSM appears to be similar to file caching. File caching is designed to eliminate large redundant file transmissions, but it often fails to deliver because of two key limitations. First, file caching works only on a single application, and since enterprises have a heterogeneous mix of applications, the overall impact that file caching has on reducing WAN transmissions is limited. Second, file caching operates only on exactly repeated files.

NSM, in contrast, works on any IP-based application. Also, it recognises repeated data sequences and eliminates them, even when a file has been modified. Since most large files transmitted over the WAN are simply modified versions of previously sent files, NSM is far more effective than file caching. For example, a couple of bullets in a 30-page PowerPoint file may be changed and the file resent over the WAN – NSM will see the repetition and eliminate 99% of the WAN transmission, but file caching will miss it.

In addition, its use of on-board hard disks enables NSM to provide hundreds of gigabytes of persistent storage, so even sequences seen several days earlier can be eliminated.

NSM complements SRS™ in that SRS™ is able to recognise and eliminate different types of repeated data patterns than NSM, and IT will enjoy maximum traffic reduction by combining NSM and SRS™ and they provide compelling benefits in contrast to traditional compression techniques.

SRS™ and NSM reduce traffic for any IP traffic – not just TCP or UDP – so combined they deliver against a broader set of applications than many other compression techniques. In addition, with traditional compression approaches, the WAN endpoints store data-replacement flags on a per-tunnel basis, so a hub site, with connections to multiple remote locations, cannot transfer knowledge of repeated data patterns on one tunnel to other tunnels connecting to other locations. Ultimately, the efficiency of traditional compression techniques is limited, since they typically can buffer only a limited number of repeated patterns. SRS™ and NSM, in contrast, store many more and longer repeated data streams. They also remove another limitation traditional compression tools create, which is the introduction of additional latency as the CPU works to recognise data patterns and replace them with a flag. SRS™ and NSM dramatically reduce traffic flows with negligible latency implications.



3) Latency Reduction

To speed transmissions across the broadest range of business software, optimisation platforms need to benefit applications based on either short-lived or long-lived TCP connections. PeriSphere's Packet Flow Acceleration (PFA) includes a series of features to boost TCP's performance in many dimensions:

• ***Fast Connection Setup*** improves the performance of short-lived connections by eliminating one round-trip time from the TCP connection setup, speeding up applications that use short connections and have chatty protocols.

• ***Flow Pipelining*** alleviates the challenge associated with smaller window sizes on older desktop and server systems. That small window size constrains the amount of data a system will send, and Flow Pipelining transparently opens up the TCP receiver window size to 64K and buffers as needed on the receive side

• ***Active Flow Pipelining*** extends the TCP performance improvements by terminating the TCP connection local to the sender and using a more efficient transport protocol between PeriSphere devices. This feature significantly benefits application performance on high-bandwidth, high-latency connections

• ***Forward Error Correction*** limits the need for retransmissions on lossy networks. It makes use of recovery packets, sent alongside data packets that index those data packets, allowing for reconstruction of lost packets

4) Bandwidth Management

Since speed differences between the LAN and the WAN differ by orders of magnitude, no amount of compression or TCP acceleration will solve all problems. This dramatic discontinuity in bandwidth means that contention for WAN real-estate is a very real problem that needs to be addressed with an effective and implementable QoS (Quality of Service) and bandwidth allocation model that enforces business priorities.

PeriSphere's Bandwidth Management includes both QoS capabilities and bandwidth allocation to allow IT to prioritise business-critical and latency sensitive applications.

PeriSphere defies the cliché that effective QoS must be difficult to implement because its intuitive wizard/template-based approach enables IT to easily ensure that business policies are met through QoS techniques.

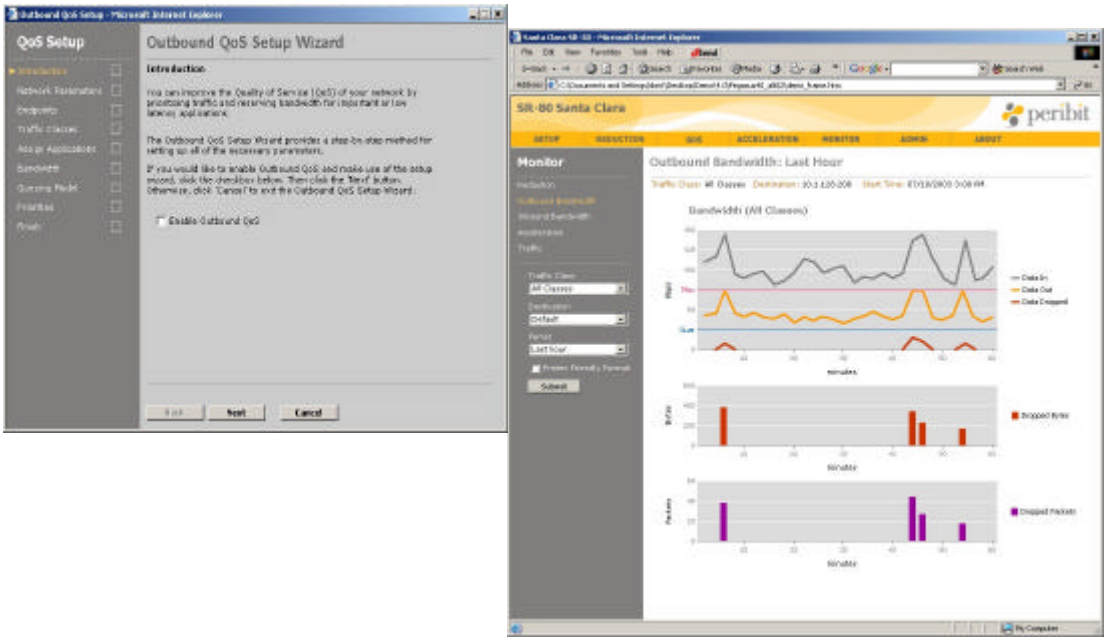
IT can assign priority status and bandwidth-allocation metrics to applications – allowing IT to classify traffic by looking at not just the IP header or ToS/DiffServ information but also inside the data payload to act on Layer 7 application information.

Why is the WAN optimisation platform the best point in the network to perform QoS and bandwidth allocation? Because it's the last point in the network that sees uncompressed traffic. If an edge WAN router performs QoS, for example, instead of the WAN optimisation device, and it enforces QoS on already compressed traffic, the QoS policy will be incorrectly applied since compressed traffic will have a disproportionate share of the bandwidth.

For example, if the QoS policy specifies that voice traffic should receive 10 percent of the available bandwidth and text traffic should receive 5 percent, but those parameters are applied after compression, text will constitute far more of the total bandwidth since it's highly compressible while voice traffic is not very compressible. The QoS policy will be accurate over the WAN but will not provide the appropriate user experience on the far end of the WAN link. Consequently, a WAN optimisation platform that lacks QoS will ultimately break the QoS policies as defined in a separate platform.

The PeriSphere devices also integrate their QoS functionality with visibility features. Getting constant feedback on WAN dynamics is essential to implementing QoS successfully. QoS

must understand, for example, how much compression is happening on the traffic – knowing the capacity of the WAN is a prerequisite for knowing when to invoke prioritisation techniques.



The alternative, required by many WAN optimisation platforms lacking this insight, is for IT to manually adjust the QoS policies after investigating and analyzing the compression results. Other platforms simply enforce QoS without regard to capacity at all, throttling back all traffic all the time to avoid contention entirely. This approach, however, makes the fundamental goal – better utilization of the WAN link – impossible.

In addition, for QoS to operate effectively throughout the enterprise, the WAN optimisation platform requires a holistic view of the WAN. That broad perspective includes seeing both those sites outfitted with an optimisation device and those without them. Many WAN optimisation platforms have no understanding of the destination location. That kind of "one-sided" QoS deployment is important for including sites without WAN optimisation platforms. However, in the enterprise QoS strategy, that mode should not be the only QoS technique allowed.

Understanding both ends of the WAN link provides extensive advantages. That kind of "dual-sided" deployment allows the PeriSphere platforms, for example, to automatically map traffic according to the QoS template assigned to that destination device. The dynamic link knowledge also enables IT to add additional PeriSphere platforms to the network very simply, with just a few mouse clicks, and have them adopt the appropriate QoS policies.

While the dual-sided approach provides for the greatest link understanding and dynamic behaviour, the platforms also support "one-sided" deployments of QoS as well, to enable a consistent QoS policy across the distributed enterprise.

Insight into the traffic type is essential for applying QoS appropriately. Many WAN optimisation platforms lack the ability to identify business applications at all levels of the network stack, from simple addresses to deep inspection within the payload. All Citrix applications, for example, look the same at Layer 3 but by looking inside the payload, a WAN optimisation platform can distinguish critical ERP traffic from simple print jobs.

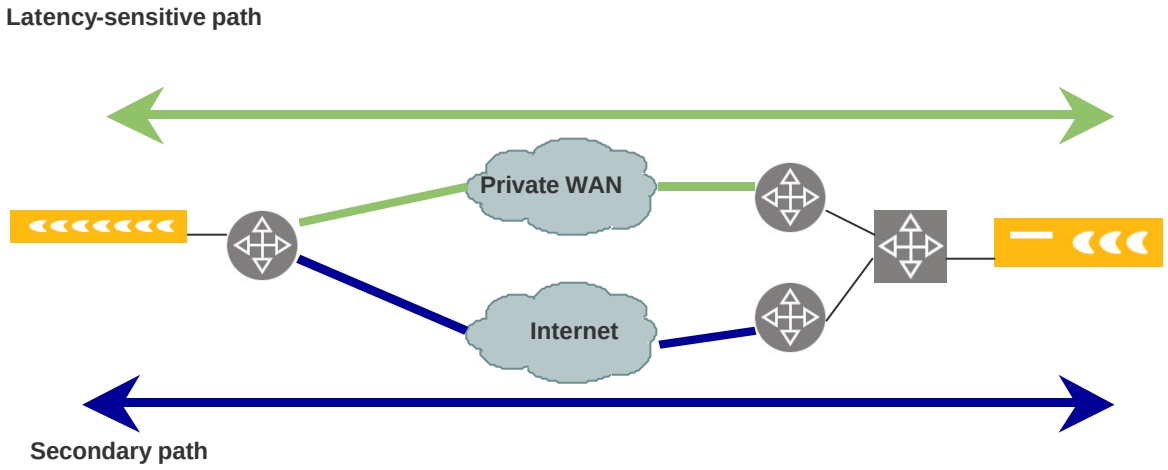
Enterprises need the flexibility to augment the QoS embedded within a WAN optimisation platform with other techniques for marking traffic, and the optimisation device should not do anything that disables the markings set by other network devices. Rather than overwrite any

needed information, PeriSphere's QoS features have been designed such that ToS/DiffServ settings can be mapped to other network devices, tunneled traffic is still identifiable by application, and MPLS CoS information can be communicated to the edge devices in service provider networks without permanently overwriting anything in the original packet.

5) Path Optimisation

Enterprises increasingly seek to take advantage of hybrid public/private WAN transports, but maintain the assurance that key performance criteria will still be met. To make effective use of both paths, IT needs to apply business policies to each link and monitor their performance. Most WAN optimisation platforms overlook this WAN deployment scenario and fail to help IT make full use of these dual links.

PeriSphere offers a unique feature called Policy-Based Multipath™ (PBM™) that enables IT to define which applications traverse which link and under what conditions. For example, IT can designate that latency-sensitive traffic such as VoIP will run over the private link while delay-tolerant applications such as e-mail and bulk file transfers will use the Internet link. But in addition to enabling this simple allocation, PBM™ also allows IT to set latency and loss thresholds for each link. IT can use PBM™ to define, by class of application, how to treat traffic when a performance threshold is exceeded. IT can designate, for example, that one application class will switch to the other link when performance suffers while another class of traffic should never leave the private link. PBM™ demonstrates the integrated nature of the PeriSphere architecture. For example, when it diverts traffic from one link to another, QoS policies ensure that applications already flowing over the second link are not negatively impacted.



6) Visibility and Monitoring

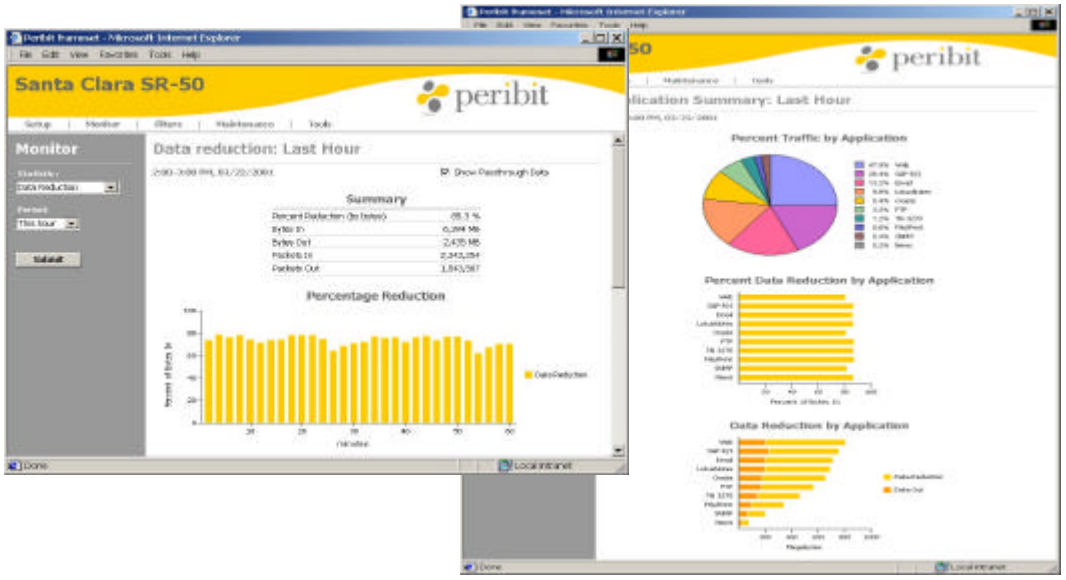
PeriSphere provides the broadest and deepest set of analysis tools for understanding WAN traffic characteristics and performance levels. IT can choose to view information per PeriSphere device or in aggregated form, and they can gain insight into such aspects as packet size distribution, error rates, throughput statistics, and TCP acceleration data.

What truly sets apart PeriSphere's monitoring capabilities, however, is its combination of data reduction and capacity improvement statistics with an understanding of the QoS statistics. It's crucial for IT to see what's happening to traffic when QoS is being invoked. If the system is applying a QoS policy, that means congestion is present. Too many applications are contending for too little bandwidth, so some traffic has to be restricted to enable priority applications to transmit. IT needs to understand that impact to better tune QoS policies.

In addition to these statistics, PeriSphere provides other unique views including a sophisticated pass-through monitor to communicate what traffic is not compressed and why, a link SLA monitor showing path latency and path packet loss, packet-size histograms that display incoming and tunneled traffic, the results of TCP acceleration, and a tunnel summary showing the status of all the tunnels on the device.

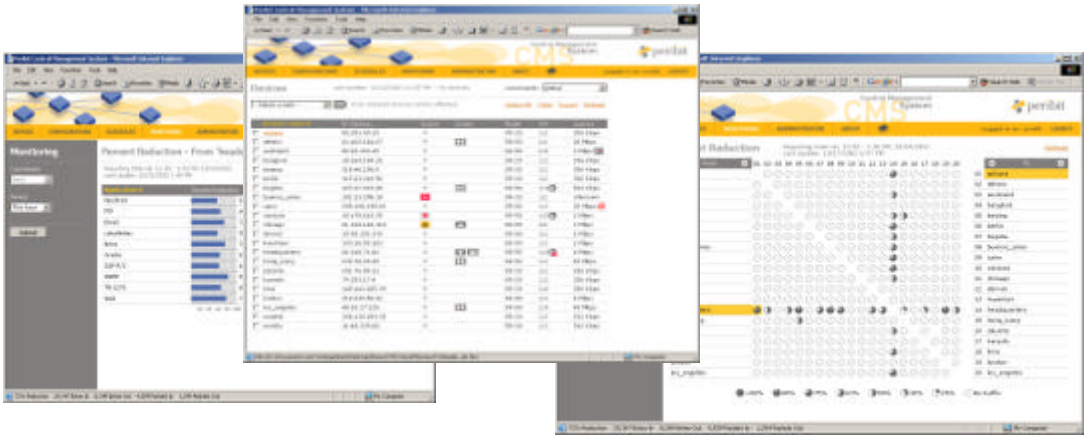
PeriSphere goes beyond simple displaying of information and allows for many methods of exporting information. Supported methods include SNMP, NetFlow, CVS, and Excel graphs. PeriSphere can facilitate in-depth troubleshooting at remote locations by taking a remote sniffer trace.

All PeriSphere devices ship with a web server to be used for configuration and monitoring reduction performance.



In addition PeriSphere offers a **Centralised Management System (CMS),** which allows IT to configure and manage multiple PeriSphere platforms from one central location. CMS provides IT with a unified view into the systems' capabilities throughout the distributed enterprise. IT can look at metrics about compression performance, application acceleration, WAN utilisation, and QoS and bandwidth allocation. CMS provides IT with information about what traffic is traversing the WAN, which applications are consuming most of the valuable WAN capacity, and which traffic is being impacted by the application of QoS. IT can also use CMS to schedule system upgrades, apply new configurations, update QoS policies, and automate license management.

To run CMS you will require a Windows 2000 Server platform, running a P3 system at 1GHz or a P4 at 1.8GHz. The system requires 768MB of RAM and 500 Mbps of disk space.



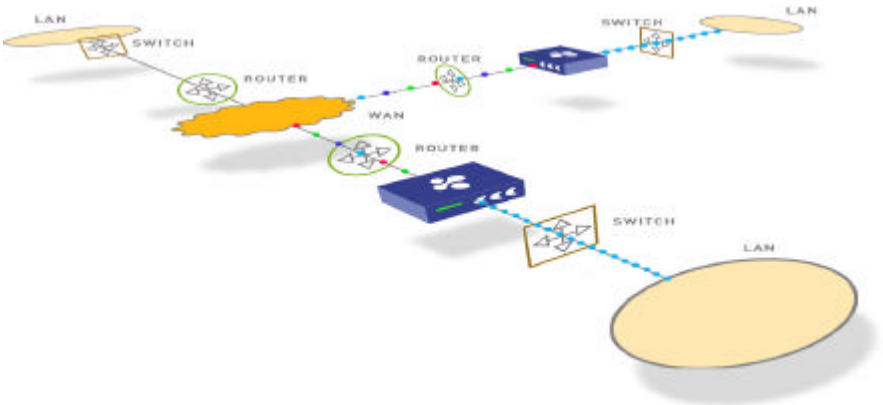
PeriSphere devices can be installed and operated in any Ethernet network, anywhere in the world. If a device encounters any hardware or software disruption, including power loss, the device immediately falls into 'pass-through' mode for all traffic that enters the box. Thus, even if a device is unplugged, it maintains network connectivity by simply behaving like a physical section of wire. This 'switch to wire' fault tolerance ensures that the worst-case result of any kind of major failure condition will be that the network assumes the same configuration that it had prior to the installation of the device. For additional levels of redundancy, the devices support dual-active redundancy with no need for extra configuration of surrounding network devices, as well as an *n*+1 backup mechanism. The platforms also work with routers configured with redundancy protocols, and the devices can load-balance tunnel traffic to redundant WAN routers or load-balance to redundant destination PeriSphere devices.

The devices are extremely easy to install and operate, can be managed individually or collectively and offer both GUI- and CLI-based views. A device requires less than ten minutes to go from its shipping carton to being fully operational within a network. Furthermore, the device requires virtually no operator intervention or tuning once it is installed.

The PeriSphere architecture requires no network, application, server, or data modifications to be effective. Each device automatically discovers all other peer devices with which it can communicate, and begins reducing data as soon as any peer device is detected. The devices support industry standard protocols and conventions. This ensures that the PeriSphere reduction and re-assembly process operates transparently to all other processes and hardware in the network, including packet routing.

Not only is it easy to install and operate the PeriSphere devices, its important to understand that as a transparent drop-in LAN switch, they serve to seamlessly compliment existing network devices, including routers, switches, firewalls, VPN devices, and Q0S/traffic shaping equipment. Since a PeriSphere device simply accepts and transmits standard IP packets, other devices in the network are unaware that the devices exist in the network. What is noticeable by a network operator is that the network itself becomes more efficient, with greater bandwidth available over the WAN and fewer packets dropped due to congestion by other devices, including routers and QoS products. Any traffic not destined to traverse the WAN immediately gets switched through the system at wire speed.

The figure below shows where the PeriSphere devices sit within the network. Because they sit on the LAN side of the routers, application allocation and or prioritisation (shaping) take place before data hits the ISP links. Similarly statistics and reports are generated on pre-ISP link traffic. On the assumption that there is no further manipulation of the optimised traffic once it has left the device, then the statistics and reports generated will also show the core traffic going over the ISP links together with their utilisation.



Another critical feature to enabling flexible deployment options is supporting both point-to-point and point-to-multipoint configurations. Many networks are built in a hub-and-spoke design, with several branch offices feeding into regional or centralized sites. That kind of traffic aggregation must be supported by the WAN optimisation platforms so that IT maintains the efficiencies of that hierarchy. To further support network hierarchy, the application delivery

platforms should support the appropriate tunnel configurations to enable traffic aggregation from regional offices to centralized locations. The PeriSphere platforms, for example, feature Tunnel Switching, which allows IT to create multiple layers of tunnel aggregation, as is often found in frame relay networks using the hub/regional/spoke topology. To support MPLS, PeriSphere implements transparent QoS communications to place traffic in the right Label Switch Path (LSR) without ever changing the packet. To support satellite links, PeriSphere supports a series of techniques that reduce the impact that latency has on TCP

For network placement, IT must have the choice to deploy WAN optimisation equipment either on the network between a LAN switch and WAN router or attached to a switch and router in a one-armed fashion. PeriSphere supports both these modes, with its In-Line Mode and its Off-Path Mode.

Other optimisers, in contrast, cannot operate in an off-path mode. Off-path deployments are essential for interoperability with some WAN architectures. For example, if the WAN router acts as a collapsed backbone, serving both local LANs and remote networks, IT needs to attach the WAN optimisation platform directly to a port on the router. PeriSphere allows IT staff using off-path mode to selectively choose what traffic is redirected to the PeriSphere device and what traffic is left untouched.

PeriSphere is also unique in its awareness of 802.1Q and its ability to compress traffic within the full 4095 number of supported VLANs. The PeriSphere platforms can optionally preserve the VLAN tags as packets are transported to other destinations through the tunnel

Automating key functions is another critical element of PeriSphere's ease of use and its most significant contribution in this area is the platforms' auto deployment capabilities designed for branch offices. Using the CMS software, IT can pre-stage configurations centrally via templates and then have remote PeriSphere platforms download them automatically. Branch-office staff need only plug in the device and connect it to the network – straight out of the box, the device will automatically procure a network address, locate the CMS software via domain name service (DNS), request a configuration, download it, and begin operation. The lack of IT involvement in deploying devices to branch offices presents a significant savings in staff time and money and enables rapid pervasive deployment.

PeriSphere automates several other tasks that are typically cumbersome to perform on WAN optimisation equipment. After a PeriSphere device is up and running, it connects to a registration server to learn about peer other remote PeriSphere platforms and how they're configured. It learns, for example, which PeriSphere devices are hubs and spokes, whether key compression techniques are enabled, whether IPsec is enabled, and whether a device is sending out traffic over two WAN links and using path optimisation. This registration information is essential to the synchronised communications amongst the PeriSphere platforms. These communications provide IT with distributed stateful intelligence about the network state such that PeriSphere devices can act on changes such as link loss or increased congestion on a path. These automated synchronisations and communications dramatically simplify both setup and ongoing operations of the PeriSphere platforms. In contrast to the manual intervention needed to support other WAN optimisation devices, PeriSphere simplifies tasks such as applying QoS policies to remote devices, upgrading those policies, upgrading the system software, and accommodating topology changes.

Dynamic Routing

On the subject of dynamic routing it should be noted that as transparent 'listeners', the PeriSphere devices never have an adverse impact on a network routing protocol in a network. The information the devices get from the networking protocols is only communicated to other peer devices within their device community. This information is used to create the overlay topology for the PeriSphere tunnels that transparently operate and reduce traffic.

It is possible for a PeriSphere device to obtain its routing directly from a Cisco router. Doing this enables the device to obtain the most accurate routing table directly for its location in the network. This feature allows the device to get all types of routes (EIGRP, OSPF, static, etc)

with the appropriate metrics without requiring any redistribution on the router or having the device directly participate in route calculation.

Scalability

Many businesses have WAN capacities that range from 64KB links in small offices all the way to OC-3 connections in headquarter locations. IT should be able to not only buy a range of equipment types to deliver various bandwidth amounts but should also be able to scale a single platform to higher capacity levels by clustering. IT also needs these optimisation platforms to support large numbers of connections to other sites in these centralised hub locations. PeriSphere's SR-100 meets these demands. The platform scales to compressed output speeds of 155 Mbps and supports up to 2000 connections to other PeriSphere platforms. PeriSphere's Tunnel Switching feature also enables any-to-any communications between any two PeriSphere sites without requiring a full mesh setup of tunnels linking the sites.

Security

For WAN optimisation platforms, two aspects of security are key: securing the device itself and securing the data that traverses the device. All methods of access to PeriSphere platforms are secure, using HTTPS and SSH. IT can also define Access Control Lists (ACLs) to allow or disallow access to the platforms, and IT can deploy Authentication, Authorisation, and Accounting (AAA)-based access to the PeriSphere devices via RADIUS.

IT also has the option to disable all network access to the PeriSphere devices and support only console access. In addition, PeriSphere actively monitors the security warnings from industry security watch-dog groups to make sure that all vulnerabilities are removed.

PeriSphere's IPsec Implementation

- AES and 3DES for encryption
- HMAC-SHA-1 and HMAC-MD5 for packet authentication
- Dynamic key exchange (IKE)
- "Retail" export approval from the Department of Commerce (~15Mbps)

To secure the platforms from a physical standpoint, PeriSphere has used no exposed flash memory cards that can be stolen and compromised, and IT can choose to deactivate the front-panel configuration feature.

The PeriSphere devices also ensure the security of their data transmissions. IT can optionally deploy a standards-based IPsec encryption feature in sites without a VPN deployment, securing data sent over untrusted links such as the Internet or satellite and also securing device-to-device communications.

Summary

While there are tangible cash savings to be realised by adding PeriSphere to a network, there are other compelling benefits. By increasing effective capacity on the network infrastructure, PeriSphere serves as a true alternative to the high cost of network build-out.

With PeriSphere a dramatic increase in capacity is achieved within minutes as opposed to months waiting for Service Provider to provide extra (often costly) bandwidth. Even when extra transmission lines are finally added, that is only the beginning of the complex and logistically challenging process of upgrading the rest of the network. The new circuits could require additional network cards on routers, or in some cases, a full upgrade to a higher capacity router. Downstream of the router, other devices such as firewalls, QoS boxes, and load balancing switches may also require reconfiguration. This ripple effect on other network elements equates to network downtime which businesses can ill afford.

With PeriSphere devices all of these complexities and risks are eliminated and the network is immediately optimised to provide greater effective capacity and greater efficiency with no additional network charges.

In addition to dramatically reducing WAN traffic, PeriSphere also significantly reduces the total number of packets transmitted across the WAN. Since fewer packets traverse the WAN, router CPU loads drop dramatically, resulting in less network congestion and packet loss. Hence, adding PeriSphere devices to a network makes the network as a whole more efficient and "healthier".

PeriSphere provides a new lease on life for WAN links that are currently loaded to capacity, improving current application performance and freeing up bandwidth for a wider array of applications over the network. These benefits are both immediate as well as long term, since any new WAN bandwidth that is provisioned in the future will also be optimised by PeriSphere to provide significantly greater effective capacity.

Network before PeriSphere	Network After Adding PeriSphere
WAN-links loaded to capacity	Effective capacity increased up to 10 times
High WAN build-out cost	Reduced as much as 80%
Lengthy WAN provisioning times	Increased capacity in minutes instead of months
Frequent WAN line card and equipment upgrades	None Required
Many small packets reduce router efficiency	Packet count reduced by up to 80%
Lack of Visibility	Clear Reporting
Applications competing for limited WAN resources	Essential Quality of Service
WAN latency	Packet Flow Acceleration
Redundant Large File Transfers	Sequence Caching
Leveraging Dual Links	Path Optimisation