

Application-Delivery.org

Glossary of Application Delivery Terminology

May 1, 2006

A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T-U-V-W-X-Y-Z

Access Control List (ACL) – a set of policies that govern access rights. ACLs are often used in application acceleration appliances to determine which traffic flows are accelerated and which flows pass through the device unmodified.

Advanced Encryption Standard (AES) - a symmetric 128-bit block data encryption technique. The U.S government adopted the algorithm as its encryption technique in October 2000, replacing the DES encryption. AES works at multiple network layers simultaneously, leveraging a fixed block size of 128-bits and a key size of 128, 192, or 256-bits.

Appliance – a specialized device that is designed for ease of installation and maintenance. Appliances have their hardware and software bundled together, with all applications and databases pre-installed. The device is plugged into an existing network and can begin working almost immediately, with little configuration.

Application Acceleration- A variety of techniques that improve perceived application response time across a WAN. These techniques include data reduction, compression, QoS, and latency /loss mitigation.

Application Front End (AFE) – AFEs reside in the data center, generally in front of Web servers. They are designed to enhance the performance of Web-based and related applications by providing a variety of features, including Layer 4 through Layer 7 redirection, load balancing, server offload, data compression, Network Address Translation (NAT), and various security functions.

Asynchronous Transfer Mode (ATM) - a WAN technology based on transferring data in cells or packets of a fixed size. The cell used with ATM is relatively small compared to units used with older technologies, allowing ATM equipment to transmit video, audio, and computer data over the same network, and assure that each traffic type gets the appropriate bandwidth and performance guarantees.

A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T-U-V-W-X-Y-Z

Bulk TCP – client/server applications that use the TCP protocol for transport. The term “bulk” refers to the fact that the size of the transactions are typically quite large – several Megabytes (MB) on average. Bulk TCP applications are not interactive, whereby individual keystrokes are transmitted across the WAN. Furthermore, they do not have the same latency and jitter requirements as real-time traffic, such as VoIP and other multimedia. Bulk TCP applications include file (CIFS), email (MAPI), and web downloads.

Bulk UDP – client/server applications that use the UDP protocol for transport. The term “bulk” refers to the fact that the size of the transactions are typically quite large – several Megabytes (MB) on average. Bulk UDP applications are not interactive, whereby individual keystrokes are transmitted across the WAN. Furthermore, they do not have the same latency and jitter requirements as real-time traffic, such as VoIP and other multimedia. Bulk UDP applications include NFS and Veritas Volume Replicator.

A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T-U-V-W-X-Y-Z

Cache – a device that locally simulates an application server, enabling local delivery of specific content. A cache sits between clients and application servers (sometimes on both ends of the WAN link, but more often just at the client location), watching all requests and locally saving copies of the responses. If another request is made for the same “object”, the cache responds directly, avoiding the need to have to go back across the WAN to the original server.

Common Internet File System (CIFS) - a protocol developed by Microsoft for remote file access. CIFS, which is based on the Simple Message Block (SMB) protocol, allows most applications to open and share files across the Internet or other IP based networks. Some specific capabilities of CIFS include file access, record locking, read/write privileges, change notification, server name resolution, request batching, and server authentication

CIFS is a fairly “chatty” protocol, requiring hundreds, or even thousands, of round trips to successfully transfer a single file. This is typically not an issue when file servers are deployed on the same Local Area Network (LAN) as clients. However, when CIFS is used across a Wide Area Network (WAN), as is the case when branch offices are accessing file servers located within a centralized data center, both latency and bandwidth constraints across the WAN can adversely impact file sharing performance.

CIFS Read-ahead - when a user is working with a file, acceleration devices can generate read ahead requests within the file in order to pipeline operations to the server, thus eliminating the round-trip delay associated with waiting for acknowledgement . This minimizes the latency associated with read operations.

CIFS Write-behind – acceleration appliances can pipeline write operations on behalf of a client, thus eliminating round-trip delays associating with waiting for acknowledgements. This minimizes the amount of round trips required to perform write operations, improving performance without risking data integrity.

Coalescing – an acceleration technique whereby multiple small packets are aggregated into a single larger packet. When packets are small, packet headers consume substantial bandwidth in comparison to the amount of end-user data transferred. Packet coalescing amortizes a single header over multiple packets thus decreasing overhead, and therefore bandwidth requirements. Packet coalescing is particularly beneficial for Web applications, VoIP, and interactive applications, like Citrix.

Compression – when a string of characters is transformed into a new string that contains the same information but whose length is as small as possible. Various algorithms can be used to compress data, such as Lempel-Ziv/Huffman (LZH).

A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T-U-V-W-X-Y-Z

Data Coherency - Data that is coherent is the same everywhere across a network. In other words, data is synchronized across all clients and servers, preventing inconsistent or outdated information from being delivered. Because caches act as proxy servers, caching technology is often subject to data coherency issues.

Data reduction – an acceleration technique whereby duplicate data is delivered from local data stores in acceleration appliances instead of being transferred across the WAN. This is achieved by fingerprinting data in real-time and sending reference pointers to remote appliances when duplicate data is detected. By preventing repetitive information from traversing the WAN, data reduction provides order of magnitude performance improvements. See local instance networking (LIN).

Data Encryption Standard (DES) - a symmetric-key encryption method developed in 1975 and standardized by ANSI in 1981 as ANSI X.3.92. DES uses a 56-bit key and uses the block cipher method, which breaks text into 64-bit blocks and then encrypts them. Due to recent advances in computing, it has become possible to crack DES encryption in a relatively short period of time. As a result, AES or triple DES is typically recommended instead of DES for data security.

Differentiated Services (Diffserv) - a protocol for specifying and controlling network traffic so that certain types of traffic get precedence.

Differentiated Services Code Point (DSCP) - A six-bit field in the IP header that specifies the per hop behavior for a given flow of packets.

Disk Encryption – protecting local data stores using common encryption techniques. This prevents data at rest in acceleration appliances from being exposed to unauthorized access. Disk encryption is best performed in hardware to minimize latency.

A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T-U-V-W-X-Y-Z

Encryption - The translation of data into a secret code. Encryption is the most effective way to achieve data security. To read an encrypted file, you must have access to a secret key or password that enables you to decrypt it. Unencrypted data is called plain text; encrypted data is referred to as cipher text.

There are two main types of encryption: asymmetric encryption (also called public-key encryption) and symmetric encryption.

A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T-U-V-W-X-Y-Z

Fail open – A condition whereby the acceleration appliance behaves as an unconnected port, presenting no link level carrier to other network devices. This causes other devices to route around the failure.

Fail to wire – Fail-to-wire is a failsafe mechanism that ensures that traffic continues to flow through an acceleration appliance in the event of device failure as if the appliance was not in the network at all.

Fingerprinting – the process of examining data in real-time looking for duplicate information stored in Network Memory. If a fingerprint match exists, reference pointers are sent between appliances, enabling the duplicate information to be delivered locally.

Five-tuple - Applications can be classified using the five basic elements of a flow: source IP, destination IP, source port, destination port, and protocol.

Forward Error Correction (FEC) – a loss mitigation technique that works by adding an additional error recovery packet for every “N” packets that are sent across the WAN. This FEC packet contains information that can be used to reconstruct any single packet within the group of N. If one of these N packets happens to be lost during transfer across the WAN, the FEC packet is used on the far end of the WAN link to reconstitute the lost packet. This eliminates the need to retransmit the lost packet across the WAN, which dramatically reduces application response time and improves WAN efficiency.

Frame Relay - packet-switching protocol for connecting devices on a WAN. Frame Relay is quite popular in the US because it is relatively inexpensive. It is often replaced by ATM in other parts of the world.

A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T-U-V-W-X-Y-Z

Health Insurance Portability and Accountability Act (HIPAA) – An act passed in 1996 whereby the US Congress called for regulations promoting administrative simplification of healthcare transactions as well as regulations ensuring the privacy and security of patient information.

While HIPAA does not specifically refer to application acceleration solutions, it does state that all entities must use “network controls to protect sensitive communication that is transmitted electronically over open or private networks so that it cannot be easily intercepted and interpreted by parties other than the intended recipient.” Application acceleration products naturally fall into this category.

A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T-U-V-W-X-Y-Z

In-line – often referred to as “bump-in-the-wire” - a deployment mode whereby acceleration appliances are inserted between a WAN router and a LAN Ethernet switch and look like a transparent layer 2 device.

Interactive applications – applications that generate requests in real-time (e.g. individual keystrokes, screen updates and/or mouse movements) and require real-time responses over the WAN. These applications are typically highly sensitive to latency, jitter and bandwidth availability. Examples include Citrix Presentation Server and Microsoft remote desktop.

IP Security (IPSec) - a set of protocols developed by the IETF to support secure exchange of packets at the IP layer. IPSec has been deployed widely to implement Virtual Private Networks (VPNs).

IPsec supports two encryption modes: Transport and Tunnel. Transport mode encrypts only the data portion (payload) of each packet, but leaves the header untouched. The more secure Tunnel mode encrypts both the header and the payload. On the receiving side, an IPSec-compliant device decrypts each packet. For IPSec to work, the sending and receiving devices must share a public key.

IP Virtual Private Network (IP-VPN) - a WAN that is constructed by using the Internet or other shared IP network. Encryption and other security mechanisms (e.g., IPSec) are used to ensure that only authorized users can access the network and that the data cannot be intercepted.

A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T-U-V-W-X-Y-Z

Latency - the amount of time it takes a packet to travel from source to destination. Together, latency and bandwidth define the speed and capacity of a network

Latency mitigation - techniques used to maintain application performance in high latency WAN environments.

Local Area Network (LAN) - a local network that connects computers located on the same floor or in the same building or nearby buildings. LAN speeds typically range from 10 Mbps to 1000 Mbps and are based on the Ethernet protocol.

Local data store – hard drive capacity used by acceleration appliances for the purpose of storing information and delivering it locally whenever possible.

Local Instance Networking (LIN) – a new approach to application delivery. Local Instance Networking delivers the performance and end user satisfaction of distributed servers, without the cost and complexity. This is done by inspecting all WAN traffic and storing a local instance of information at the appropriate enterprise location. The local instance is used to deliver information to remote users when appropriate, maximizing WAN efficiency and ensuring LAN-like performance.

Local Instance Networking is typically defined by the following characteristics:

- Stores a single local instance of information at each location
- Retains 100% coherency with existing servers
- Remains transparent to all applications and network infrastructure
- Provides latency and loss mitigation
- Classifies and prioritizes traffic with Quality of Service (QoS) guarantees
- Secures network traffic and locally stored information

Loss mitigation – techniques, such as Forward Error Correction (FEC), that maintain application performance in WAN environments experiencing high packet loss.

A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T-U-V-W-X-Y-Z

Mbps - short for megabits per second, a measure of data transfer speed across a WAN. A megabit is equal to one million bits.

Mobile-based Messaging API (MAPI) – an architecture that enables applications to interact with multiple messaging systems seamlessly across a variety of hardware platforms. Client applications use MAPI to access user mailbox and public folder information stored in Microsoft Exchange servers and user directory information stored in Active Directory.

In the MAPI architecture, clients communicate with a back-end system through the client interface of the MAPI subsystem. This is done via “sessions”, where a client will perform various tasks, including logging on and off of the server, reading text data, accessing attachments, and sending information. Each of these tasks requires its own function call and appropriate acknowledgements.

While MAPI is a mature and robust solution for accessing mail and folder information, it was not designed to work across a Wide Area Network (WAN). That is because it maintains an open connection for the length of the entire MAPI session and involves lots of back and forth calls that can be susceptible to the latency and bandwidth constraints inherent to most WAN links.

Multi-Protocol Label Switching (MPLS) - an IETF initiative that integrates Layer 2 information about network links (bandwidth, latency, utilization) into Layer 3 (IP) in order to simplify and improve IP-packet exchange across a wide area.

A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T-U-V-W-X-Y-Z

Network Address Translation (NAT) - an Internet standard that enables a local-area network (LAN) to use one set of IP addresses for internal traffic and a second set of addresses for external traffic. Some application acceleration devices will use NAT to communicate with one another across a WAN. The alternative to NAT is to create a tunnel between acceleration devices. Both have unique advantages. NAT, for example, provides better visibility, enabling IT staff to count individual flows and distinguish UDP vs TCP across the WAN. Tunneling is typically more secure, is better for enforcing policies, and often provides performance benefits by allowing cross-flow compression and coalescing.

Network File System (NFS) - a client/server application designed by Sun Microsystems that allows all network users to access shared files stored on computers of different types. NFS can use TCP or UDP as a transport protocol.

Network Memory —the data reduction component of Silver Peak's LIN solution. Network Memory™ uses advanced fingerprinting technology to examine data prior to it being sent across a WAN. If these fingerprints match data that is stored in a local instance at the destination location, the information will not be sent across the WAN. Instead, instructions are sent to deliver the data locally.

Network Memory™ reduces the amount of information that is traversing a WAN by orders of magnitude and enables information to be delivered with LAN-like performance. As a result, it is an extremely effective tool when delivering an application across a distributed enterprise.

A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T-U-V-W-X-Y-Z

Out-of-path – a deployment mode whereby acceleration appliance are not in the direct path of the network traffic. Redirection techniques, like PBR, WCCP, and VRRP are required to forward traffic to the appliance.

A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T-U-V-W-X-Y-Z

Packet classification - partitioning network traffic into multiple priority levels or classes of service. For example, by using the three precedence bits in the Type of service (ToS) field of the IP packet header packets can be categorized into a limited set of up to six traffic classes. After packet classification, other QoS features can be utilized to assign the appropriate traffic handling policies including congestion management, bandwidth allocation, and delay bounds for each traffic class.

Policy Based Routing (PBR) - a mechanism for expressing and implementing forwarding/routing of data packets based on the policies defined by network administrators. PBR provides a flexible mechanism for routing packets through routers, complementing the existing mechanism provided by routing protocols.

A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T-U-V-W-X-Y-Z

Quality of Service (QoS) - Quality of Service (QoS) is used to optimize performance in the presence of network impairments and to give preferential treatment to certain classes of traffic in the event that demand exceeds available bandwidth.

As each individual enterprise application has its own delivery requirements (e.g., latency and jitter) and enterprises sometimes desire that varying levels of priority be given to different types of traffic, QoS is indispensable to enterprises deploying a mix of applications across a Wide Area Network (WAN). It ensures that each application is treated appropriately as it vies for limited WAN resources.

Application acceleration solutions do not mitigate the need for QoS across the WAN. While these solutions improve bandwidth efficiency and perceived application response times, they do not eliminate the fact that a mix of traffic with varying levels of priority must be delivered over a fixed resource. As a result, QoS is still essential for guaranteeing the quality of application delivery and ensuring predictable behavior across an optimized WAN.

Queuing – the process of delaying packets until a bandwidth-related condition is met.

A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T-U-V-W-X-Y-Z

Rack Unit (RU) - unit of measurement referring to the space between shelves on an equipment rack. 1 RU is equal to 1.75 inches

Real-time applications – applications that are highly sensitive to time delay. As a result, they often have stringent jitter, latency and packet loss requirements. Examples of real-time applications include Voice over IP (VoIP) and videoconferencing.

Redundant Array of Independent Disks (RAID) - a category of disk drives that employ two or more drives in combination for fault tolerance and performance.

Router - A device that forwards data packets along networks. A WAN router connects one LAN to another via a WAN or service provider's network.

A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T-U-V-W-X-Y-Z

Sarbanes Oxley (SoX) – Sarbanes Oxley was signed into law on July 30, 2002 in response to widely publicized corporate and accounting scandals. The Act introduced stringent new rules on corporate officers and directors of publicly traded companies with the stated objective of protecting investors by improving the accuracy and reliability of corporate disclosures made pursuant to the securities laws. SoX creates various compliancy issues that increase the need for strong data security in application acceleration solutions.

Secure Sockets Layer (SSL) - a protocol developed for transmitting private documents via the Internet. SSL uses a cryptographic system that uses two keys to encrypt data – a public key known to everyone and a private or secret key known only to the recipient of the message. Most web browsers support SSL, and many Web sites use the protocol to obtain confidential user information, such as credit card numbers.

A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T-U-V-W-X-Y-Z

Transport Control Protocol (TCP) - TCP enables two hosts to establish a connection and exchange streams of data. TCP guarantees delivery of data and also guarantees that packets will be delivered in the same order in which they were sent. TCP operates at the transport layer (layer 4) of the ISO stack.

TCP acceleration – a variety of techniques used to compensate for poor performance on high latency links when using TCP based applications, including the dynamic adjustment of window and transaction sizes and selective acknowledgements (ACK).

TCP back-off - a method of throttling traffic flow specified in the Transmission Control Protocol. TCP back off enables networks to recover during periods of congestion.

TCP flow – a session between an originating device and an end device, which is defined by a unique source IP:port and destination IP:port combination. At any given time, a typical end user will have 10-20 TCP flows active on their computer.

Tunnel - virtual point-to-point links between two application acceleration devices. They work by wrapping original packets of data inside an outer IP header, which is used to specify the address of the device on the far end of the WAN link.

The alternative to tunnels is Network Address Translation (NAT). Both have unique advantages. NAT, for example, provides better visibility, enabling IT staff to count individual flows and distinguish UDP vs TCP across the WAN. Tunneling is typically more secure, is better for enforcing policies, and often provides performance benefits by allowing cross-flow compression and coalescing.

A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T-U-V-W-X-Y-Z

U – see Rack Unit (RU)

User Datagram Protocol (UDP) - a connectionless protocol that, like TCP, runs at the transport layer of the ISO stack. Unlike TCP, UDP provides very few error recovery services, offering instead a direct way to send and receive datagrams over an IP network. TCP is commonly used for time sensitive traffic, such as Voice over IP (VoIP) or video.

A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T-U-V-W-X-Y-Z

Virtual Router Redundancy Protocol (VRRP) - an election protocol that allows several acceleration appliances to utilize the same virtual IP address. In a VRRP setup, one appliance is elected as the master with the other appliance acting as a backup in case of failure.

Voice over Internet Protocol (VoIP) - hardware and software products that enable people to use IP networks, such as the Internet, to send and receive telephone calls. In a VoIP environment, information is sent in IP packets instead of via traditional methods used by the public switched telephone network. Application acceleration appliances often implement QoS, packet coalescing, and compression techniques to improve the performance of VoIP over a WAN.

A-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T-U-V-W-X-Y-Z

Wide Area File services (WAFS) - a storage technology that makes it possible to access remote file services as though they were local. WAFS leverages caching technology coupled with specific protocol acceleration techniques to improve file access and storage across the WAN. To date, WAFS has struggled to gain industry acceptance due to a lack of applications supported and potential coherency issues that come with file caching.

Wide Area Network (WAN) - A computer network that spans a relatively large geographical area. Typically, a WAN consists of two or more local-area networks (LANs). Computers connected to a WAN are often connected through public networks, such as the telephone system or the Internet. Common WAN technologies include Frame Relay, ATM, MPLS, and satellite.

WAN Optimization - WAN optimization products are most often deployed as bandwidth “band-aids”, providing short term benefits on congested WAN links where it is infeasible or too expensive to buy additional bandwidth. Although different vendors have their own proprietary implementations, WAN optimization solutions rely on two underlying technologies: compression and Quality of Service (QoS).

Web Cache Communications Protocol (WCCP) – a protocol that specifies interactions between one or more routers, acceleration appliances, or caches. The purpose of the interaction is to establish and maintain the transparent redirection of selected types of traffic flowing through a group of devices, optimizing resource usage and lowering response times.