

Network Performance Optimization

LTI DataComm Practice Area Overview

WIDE AREA DATA SERVICES: A PROVEN BUSINESS TECHNOLOGY

Wide Area Data Services (WDS) is about much more than file services across the WAN; it encompasses TCP optimization, data reduction for replication, email acceleration, and a range of other application-level optimization. This technology is distinct and important because the converged WDS capabilities embodied in the WAN optimization controller solution provide us a path to the truly serverless, lean, network-efficient branch office.

As WDS matures and gains wider market traction the long-held vision of an efficiently distributed enterprise with completely centralized IT functionality is achievable. When this happens, WDS will be viewed as the enabling technology set. This brief provides an overview of the constituent components of WDS, and a brief look at the role LTI DataComm is playing in bringing this technology to market.

CHALLENGES OF THE DISTRIBUTED ENTERPRISE

Distributed organizations come in all sizes and shapes. Even small organizations may have multiple locations connected by data networks around the world. For the largest enterprises, even the simplest usable network is a complicated arrangement, while the most elaborate and sophisticated networks have unbelievable layers of replicated and interlocking functionality.

Whether a distributed organization is small or large, whether its networks are simple or complex, each faces similar business challenges: making the distributed organization effective and efficient in its use of the Wide-Area Networks (WANs) that tie the organization together.

Often these business challenges show up as distinct IT and WAN issues at remote offices, satellite offices, or branch offices. Poor application performance and poor control of information at branch offices are common examples of ineffective use of the WAN, which leads to reduced productivity and dangerous exposure to liability. Excessive spending on network bandwidth and high administrative costs at branch offices are both common examples of inefficient use of WAN resources: throwing away money that could be put to better use.

What are the underlying causes of these challenges? The network of a distributed organization typically consists of an identifiable richly-connected core and an identifiable weakly-connected fringe. The core typically includes most servers, many clients, and most of the organization's available bandwidth. In contrast, the fringe typically includes most of the organization's remote locations. Operations between clients and servers in the core occur at the speeds typical of Local-Area Networks (LANs), whereas operations involving clients in the fringe occur at much slower speeds typical of WANs.



Technology Overview



WDS Solutions Requires Minimal Branch Office Infrastructure:

“Wide-Area Data Service (WDS)” Appliance

- Minimize Bandwidth
- Maximize Throughput

“Network” Box

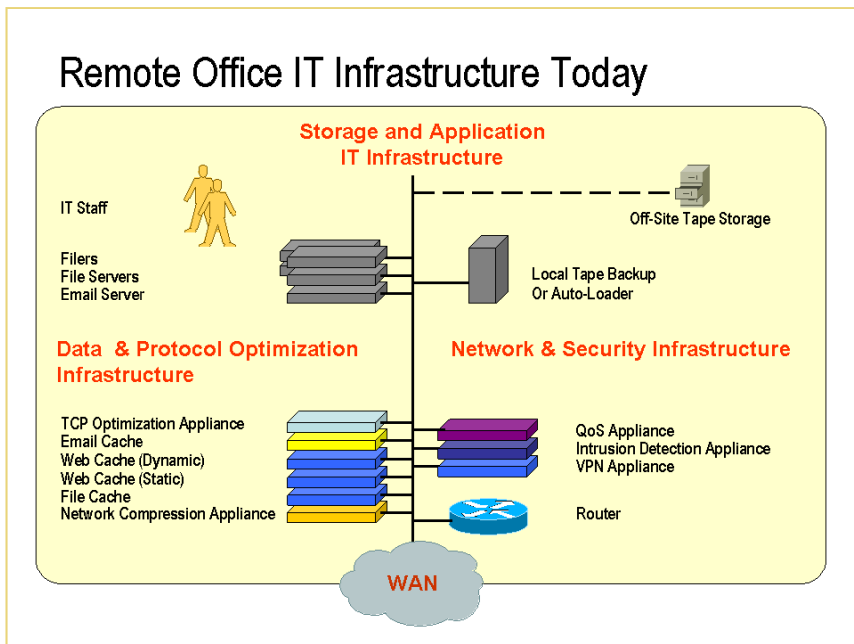
- Routing and VPN
- Security and QoS

Most other Infrastructure Centralized to DataCenter

Application of WDS:

- Application Acceleration
- Remote Office Data Backup
- Data Replication
- Site Consolidation
- In-Sourcing
- Bandwidth Optimization
- Server Consolidation, COOP and Disaster Recovery

Typical WAN bandwidth at the fringe is only 1% or less of the typical LAN bandwidth, while latency (the time required to take a single end-to-end round trip) is often 100 times longer or more on a WAN than on a LAN. Because the fringe has network capacity (bandwidth) that is 100 times narrower and 100 times longer (latency) than the core, it is not surprising that performance is often a problem. With this problem in mind, we can consider how to solve the effectiveness and efficiency challenges at branches by making the WAN perform more like a LAN. Previous attempts to address these challenges have led to the deployment of a variety of local servers, or Wide Area File Services (WAFS) for file content, and an assortment of compression or TCP optimization devices. These discrete technologies can be considered a subset of a larger technology category that encompasses all of them. Namely, Wide Area Data Services (WDS).



THE WIDE AREA DATA SERVICES SOLUTION

Making the WAN behave more like a LAN has two components: 1) delivering dramatic performance improvements, especially for those applications or protocols that have the worst degradation when running over a WAN, and 2) contending with a wide spectrum of protocols that are traversing the diverse networks of a distributed organization.

Solutions that deliver this full range of optimization for WANs go by the name of Wide-area Data Services (WDS). Wide refers not only to the WAN but also to the wide applicability of WDS, and the wide spectrum of bottlenecks improved by WDS. WDS is not only a solution to attack problems of ineffectiveness or inefficiency, but also a means to realize goals that would otherwise be impossible:

- Centralizing distributed infrastructure like file servers, mail servers, Network Attached Storage (NAS), and remote office backup systems – all without affecting remote users.
- Sharing large files among colleagues on different continents – as if they were in the same building. Performing backup and replication over long distance WAN links – and completing them during backup windows that were previously unachievable.
- Delivering significantly more services on existing WANs – without upgrading the bandwidth

If desired, an organization can eliminate branch servers and devices, replacing them with only network services and WDS appliances. Additionally, a branch without local servers can dramatically improve its productivity by utilizing WDS.

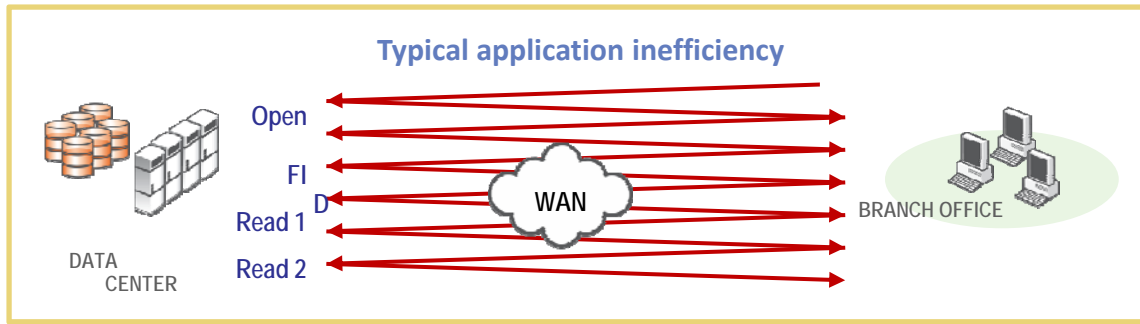
BANDWIDTH AND LATENCY BOTTLENECKS LIMIT WAN PERFORMANCE

Let's consider the bottlenecks in some more detail. WAN connections typically have lower bandwidth and higher latency than LAN links, but how do those constraints actually affect application performance? There are four distinct bottlenecks, one relating to bandwidth and three relating to latency. The bandwidth bottleneck is straightforward: no application can send data faster than the available bandwidth. However, there are several forms of latency bottlenecks that may impact applications, even when the bandwidth appears to be plentiful.

Latency Bottleneck #1: The first latency bottleneck is caused by the end-to-end acknowledgement behavior of TCP. TCP has a window of packets that can be in flight from one end to the other (i.e. between client and server). After the window is full, the sender cannot send additional packets until the destination acknowledges receipt of at least some of what has already been sent. If the maximum window is too small, the throughput of the link will be limited by the rate at which each full window can be sent to the other side and acknowledged.

In theory, this bottleneck should be rare, because well-specified mechanisms exist to allow TCP to use large windows, and most recent operating systems implement those mechanisms. However, settings on both clients and servers are usually more attuned to LANs than WANs, and it is unusual to find clients or servers with TCP stacks matched to WAN latencies.

Assuming an 8k TCP window and 75ms latency over a T3 between a file server in L.A. and a tape library in N.Y., bandwidth can be calculated by dividing the TCP window by latency: $(8\text{kbyte}/75\text{ms}) * 8 = 853\text{kbps}$.



At 853kbps, a 500Mbyte payload would require 62,500 round trips and approximately 80 minutes to complete. When optimization is leveraged, the latency between two sites is learned and the TCP window is dynamically adjusted to maximize WAN bandwidth: $(420\text{kbyte}/75\text{ms}) * 8 = 45\text{Mbps}$.

A WDS-enabled network would allow a 500Mbyte backup to complete in 1200 round trips and in less than 90 seconds.

Latency Bottleneck #2: The second latency bottleneck is caused by the slow start and congestion control behaviors of TCP. The first latency bottleneck, explained above, is a limit based on the maximum window possible. This second latency bottleneck is caused by TCP not even running at that (probably inadequate) maximum window size all the time. Instead, TCP gradually ramps up its window size when transmission appears to be successful and sharply cuts back its window size when transmission appears to be unsuccessful. In networks with both high bandwidth and high latency this behavior leads to extended periods in which available bandwidth goes unused.

Latency Bottleneck #3: The third latency bottleneck is caused by so-called application protocols that are running on top of TCP. Recall that with the first latency bottleneck, the availability of bandwidth didn't matter if TCP was limited by the size of a window of data and the need to acknowledge that data.

Analogously, the availability of bandwidth and the avoidance of the first and second latency bottlenecks (at the TCP layer) are irrelevant if the application is limited by the size of application messages that must be acknowledged. Application protocols that were originally designed for wide-area environments – such as HTTP and FTP do not encounter this third latency bottleneck. However, application protocols originally designed for use on LANs – such as Microsoft Windows file sharing via the CIFS protocol – are often severely affected by this third latency bottleneck.

WDS: WIDE APPLICABILITY AND WIDE AREA

Historically, the various single point approaches to WAN improvement have addressed specific bottlenecks, or have only offered improvement for only a narrow set of protocols.

Some examples of those approaches are summarized below. Note that WAFS technology, while excellent for alleviating bandwidth issues and application latency, does not address TCP latency. Conversely, TCP optimization software addresses the TCP latency challenge, but does not have applicability to bandwidth and application issues.

This table illustrates that prior to the conception of a unified WDS approach, no single point technology can address a broad spectrum of protocol challenges. WDS combines four distinct threads that have been present in these

Approach	Bandwidth?	TCP Latency?	Application Latency?
WAFS	Yes		Yes
WAN Optimization	Yes		
Web Caching	Yes		Yes
Data Reduction	Yes		
Email Caching	Yes		Yes
Block-Replication	Yes		
Dynamic Caching	Yes		Yes
TCP Optimization		Yes	
QoS			
SSL Acceleration			Yes

constituent approaches, but always as separate elements: High-performance disk-based data reduction; TCP optimization; Application or protocol specific optimization; and caching on local servers.

WDS IN ACTION

WDS can be applied to many areas affecting enterprise networks today. Because WANs are such an integral part of a distributed organization's infrastructure, they affect many critical business processes and in some cases stand as a real obstacle to the plans and goals of the organization. Here are a few of the key usage models that a WDS solution can provide:

Application Acceleration- Many business processes are dependent on applications deployed across WANs. The application can be as simple as sharing Windows files, or as complex as a custom-built application. In either case, if a WAN is in the middle, it's almost guaranteed that the application won't work as well as it was intended. WAN optimization controller (WOC) appliances can accelerate many applications like Exchange, HTTP, FTP, backups and

Technology Overview

several others by as much as 100 times. With LAN-like performance, users can work together no matter which remote office they are located in.

Remote Office Data Backup- Backing up remote office servers requires the transfer of what normally wouldn't be too much data. But when it needs to be sent over a low bandwidth, high latency WAN link, the throughput drops dramatically, as we've seen. What would be a quick job on the LAN can easily take longer than the available backup window when implemented across the WAN. For this reason, most IT managers rely on local tape auto-loaders or other backup schemes in the branch office. Of course, local backup is fraught with poor execution, equipment failure, and operational difficulty. With a WDS solution, backups can be completed in a fraction of the time they currently take, which allows you to take a different approach to protecting your company's data.

Data Replication- Whether data needs to be replicated to support redundancy and business continuity plans, or to mirror data so it's available for users around the world, data replication is critical. A WDS solution can accelerate data replication processes by a factor of 10 or more. A WDS appliance will remove all the redundant traffic from the WAN, and optimize TCP. The two combine to make a huge difference in reducing the time required to complete data replication.

Site Consolidation- WDS appliances allow successful consolidation of file servers, email servers, NAS, and local tape backup. Deployments can start by consolidating a single infrastructure, with incremental consolidation of other infrastructure thereafter. Alternatively, it is possible to consolidate everything and achieve the serverless office.

In-sourcing- Distributed organizations collaborating together often discover that they have spare capacity in some locations and insufficient capacity in others. By enabling more flexible sharing of information and applications across geographic boundaries, WDS makes it possible to use idle internal resources to assist in situations that might otherwise require additional local consultants or other temporary help. This in-sourcing saves money and helps lower the volatility of work life by spreading work more evenly across locations.

Bandwidth Optimization- Sometimes the goal is as simple as avoiding a WAN upgrade, and an effective WDS appliance can help with that too. LTI's experience indicates a reduction in WAN traffic by 60% to 98%, which means that an existing WAN can often support many more users, new applications like VoIP can be rolled out, and an expensive WAN upgrade can be delayed or avoided.

NETWORK PERFORMANCE ENGINEERING

When addressing network performance it is often better to take a holistic approach, rather than addressing specific applications or pain points. LTI can provide greater visibility into your network and allow it to run efficiently by using comprehensive network behavior analysis tools, such as behavioral analytics and network performance metrics. Our network performance solution will help discover unknown dependency mapping, improve security, streamline data center consolidation efforts and manage WAN resources towards accelerating productivity. In summary, it will tell you what is on your network to assist you in using your network resources more effectively.

Our network performance solution collects network flow data that traverses your network to provide detailed information pertaining to which applications run on which servers, layer 7 visibility on how bandwidth is being consumed, and determining bandwidth requirements. This approach is simpler and less invasive than installing taps and probes at different points along your network. The improved visibility can answer relevant questions that network administrators need to know; Are there unauthorized users on my network? What services does the data center deliver? What servers deliver these services and who consumes them? How is bandwidth consumed and can unnecessary bandwidth be eliminated? Investments and deployments can be optimized, usage policies enforced, service-level objectives defined and service-level events can be monitored more effectively.

LTI DataComm
23020 Eaglewood Ct. #100
Sterling, VA 20166
www.ltidata.com
800-677-5050