WAFS Market: Why Is It So Hot?

Arun Taneja

Products that provide wide area file services (WAFS) have taken off in the past few years. Here's why.

Wide Area File Services (WAFS) products have now been shipping for about two years. IT and network managers who initially wondered, "What in the world is WAFS?" became curious ("What do they do?"), and then excited ("Wow!"). Today, most of them are deploying or evaluating WAFS solutions.

The main benefit to be derived from WAFS is the ability to improve data delivery to workers at remote sites, while saving money compared to alternative configurations and solutions.

The innovations of the four original WAFS players—Tacit Networks, Actona (now part of Cisco), DiskSites (now part of Expand Networks) and Riverbed Technology—have been recognized by the WAN optimization and other performance-enhancing-appliance makers, as well as by the mainstream server and network equipment providers. Consequently, WAFS features are appearing on additional platforms and among the feature sets of additional devices.

This article explains how WAFS solve remote data access problems. We will also look at the four original WAFS players and a few others, then see where WAFS players are headed.

What Is WAFS?

The acronym WAFS refers to a category of appliance-type products that deliver LAN-like file access across a WAN. Usually WAFS products are deployed in pairs, one at the remote site's LAN edge and the other at the datacenter's LAN edge.

Depending on the WAFS product design, a proprietary protocol is spoken between the two appliances to streamline client/server file access and updates. No changes are needed to the application, the client or the server in order to use WAFS appliances, which can be remotely managed.

Not only do WAFS improve performance across the WAN, they also help simplify and reduce the cost of the IT infrastructure. WAFS eliminate the need for distributed file servers and related equipment at the remote site(s), which, in turn, saves remote staffing costs and supports centralized data and network management.

Note that there are several different ways to solve the remote file access problem, and each has pros and cons. We will contrast these approaches further below, but first let's look at the basic issues that WAFS are designed to address.

The Need For WAFS

The need for WAFS stems from the increasing importance of branch offices, subsidiaries and other remote-site enterprise operations. Even small companies today are apt to be distributed, with remote workers and outsourced functions.

The outsourcing of research and development, customer support, and testing to organizations in India and China, for example, is a reality for thousands of organizations in the U.S. Financial and retail institutions also have many remote offices. Auto and pharmaceutical manufacturers have multinational locations, each depending upon and needing to share data with the others. Virtual locations and workers are just a reality of our times.

Not only do companies worry about remote business operations, costs and worker productivity, they also have additional pressure from regulators to manage the information privacy and security aspects of their remote offices. These regulations and the penalties for infractions apply to all sensitive customer data, regardless of its location—at headquarters, the outsourcer's datacenter, at third-party locations such as an Internet datacenter or at corporate remote offices. This fact alone has changed the way C-level management thinks about data in remote locations.

As important as these remote access, privacy and security issues are, IT has had only two basic alternatives, both of which are ill suited for today's environment:

- Consolidate all data in the central datacenter, where it can be managed and protected, using the

Arun Taneja is founder and consulting analyst with Taneja Group (www.tanejagroup.com), a firm specializing in providing analyst and consulting services to the storage and server markets. He can be reached at arunt@tanejagroup.com.

Remote mini-datacenters can provide the best service, but they have to be kept in synch by the central IT staff

IT expertise available there, and let the remote users get access to this data over the WAN.

- Replicate the data needed by remote users to separate, typically smaller file servers or network attached storage (NAS) boxes in the remote offices. Then periodically synchronize the remote and central data.

Centralizing all the data works best for IT, but this option can perform poorly for the remote office users. Many enterprises still do not have high-speed VPN access at their remote sites, so application performance can be glacially slow. Even when high-speed VPNs are in place, performance can still be slow because the network enhancements do not address the file-level protocols (more on this below).

The other alternative, installing smaller file servers (or NAS) at the remote sites and replicating data to/from the datacenter NAS (or file server), has been a fairly popular option, and is shown in Figure 1. End users love the fast response, but keeping remote data in synch with central data stores and supporting mini-datacenters at remote sites can be a nightmare for centralized IT staff.

Even if there is a business unit budget to equip the remote site with all the necessary file, application, backup and storage servers, there is rarely enough to also pay for an on-site IT expert. Central IT staffers end up talking remote non-experts through complex operational procedures for backup, synchronization, troubleshooting, etc. The results are as one would expect: inconsistent and occasionally disastrous.

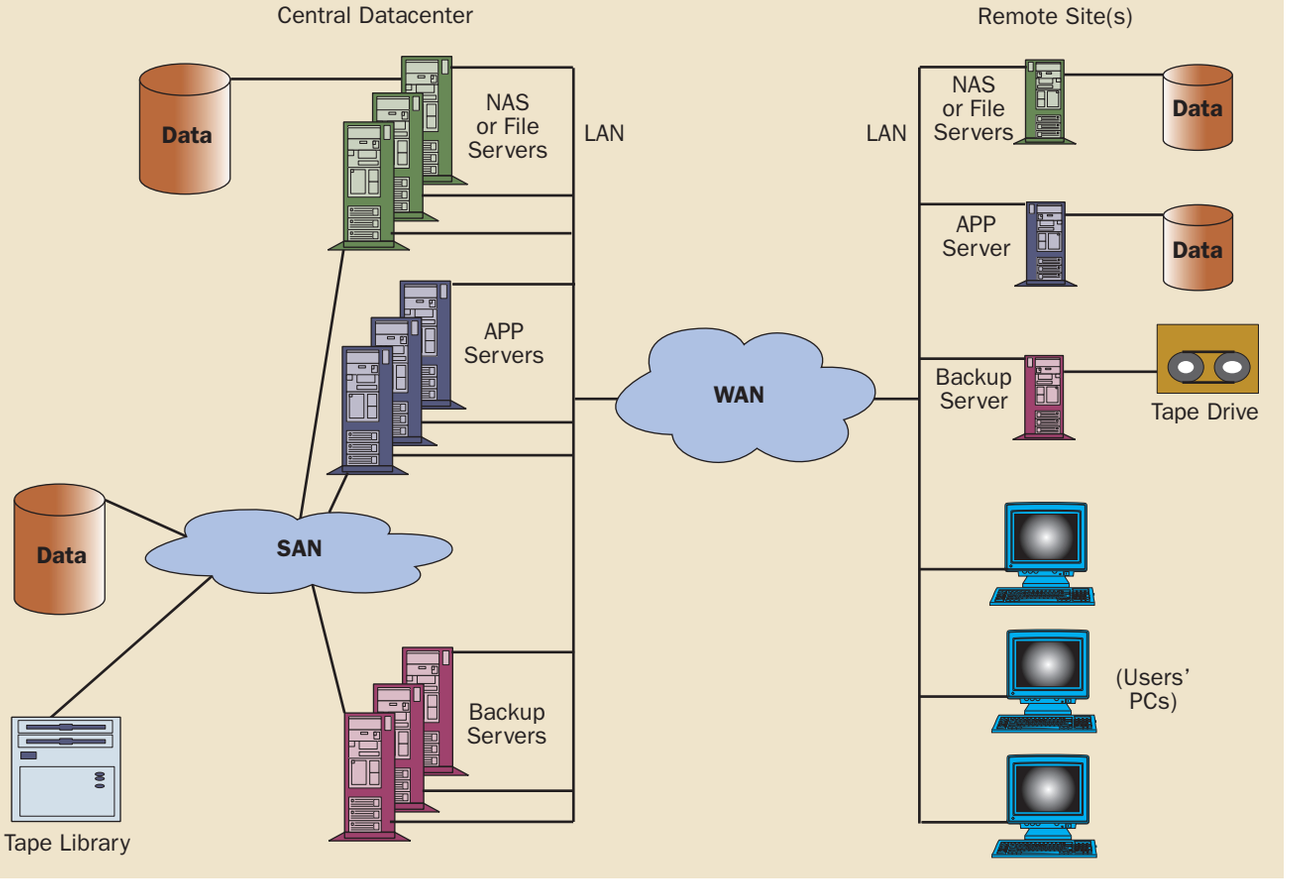
Nonetheless, the remote mini-datacenter has been the better alternative of the two, or it was until WAFS came along. This is largely due to the nature of two old protocols that are still in use in almost all commercial file-sharing applications.

The Trouble With NFS And CIFS

File servers use the Network File System (NFS) or the Common Internet File System (CIFS) protocols to securely access files located on a common file server or a NAS box. These protocols are designed so that two individuals can share files, but so that they cannot work on the same block of data at the same time. They allow read, write and/or execute access to a file, on a per-user basis, based on permissions set by the administrator.

NFS and CIFS operate on top of TCP/IP, the networking protocols of choice for both LANs (Ethernet) and WANs. All these protocols were created decades ago, when both LANs and WANs were quite slow and unreliable. To compensate,

FIGURE 1 Today's IT Infrastructure



the protocols were designed to make many short back-and-forth exchanges, or "turns," in order to accomplish each application-level activity.

We refer to such protocols as "chatty," and on a LAN, they do not cause performance problems since LAN latencies are miniscule (about 0.1 millisecond). Chatty protocols cause performance problems across the WAN, however, partly because bandwidths are typically so much less than LANs, but also because of the speed of light.

Adding WAN bandwidth can't solve the problem, because it can't make the chatty protocols any quicker or more efficient, nor can it make the speed of light any faster. Even the very highest-bandwidth WAN between New York and San Francisco will always have at least a 16-millisecond speed-of-light latency, often closer to 50 milliseconds or more each way (depending on the path the data takes, the performance of the WAN routers, etc.). This is what makes the performance of file serving essentially unbearable for all except those with LAN access to the datacenter. (For more, see *BCR*, May 2006, pp. 8–10.)

Common Approaches To WAFS

To improve file serving performance, WAFS pioneer Tacit Networks (soon to be part of Packeteer), and later Actona (now Cisco) and DiskSites (now Expand Networks), combine caching with the use of a low-latency, proprietary protocol between the pair of WAFS appliances. Each vendor's approach differs slightly, but they all make the WAFS appliance at the datacenter look like a client to the central file servers (or NAS boxes) in such a way that the latter has no knowledge of the existence of the remote users or systems (Figure 2).

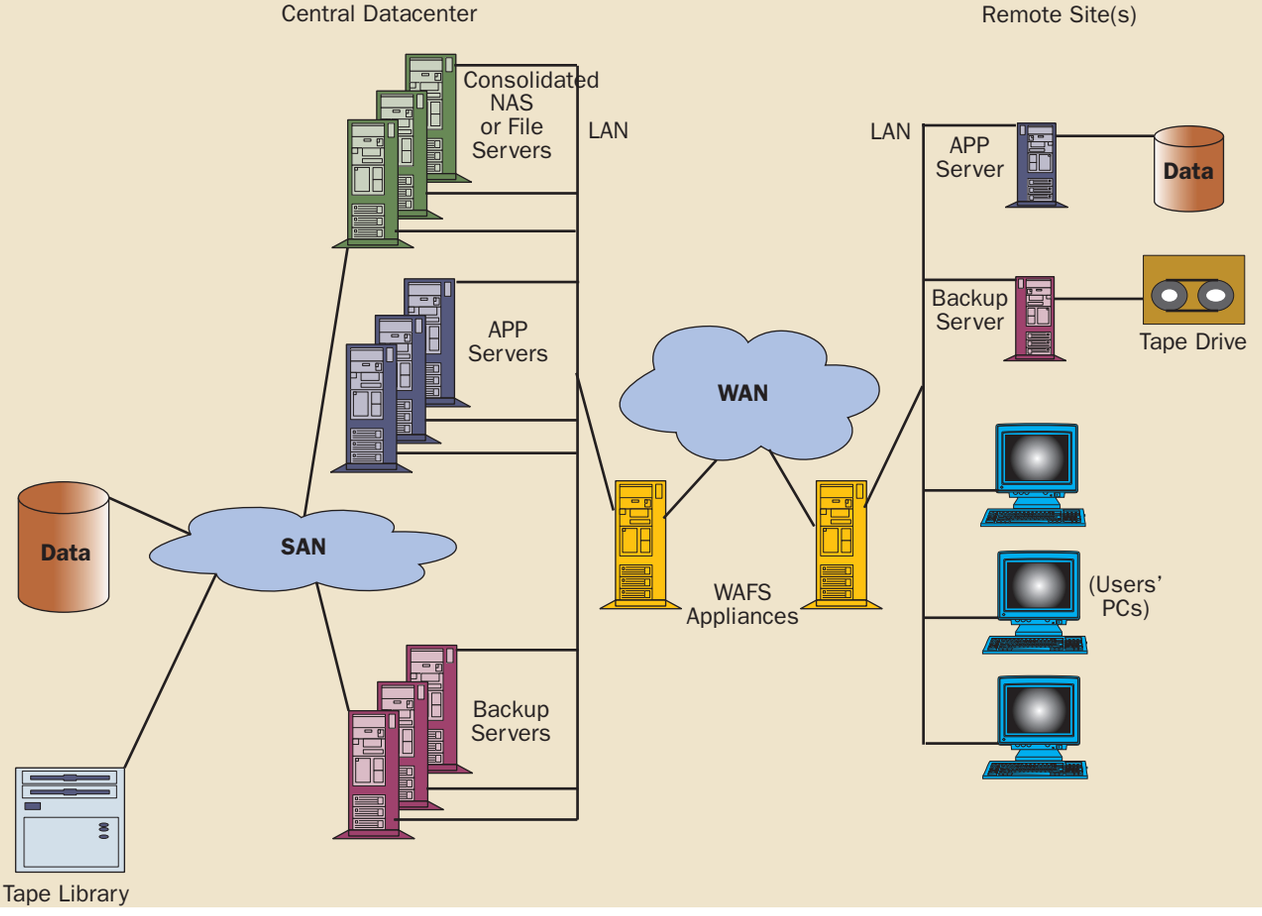
The WAFS appliance at the datacenter also manages everything downstream. It communicates very efficiently with the remote WAFS appliance, such that it eliminates much of the "chattiness" of NFS and CIFS protocols. It also proactively and intelligently sends files over to the remote appliance for caching. For example, it knows what is already in the remote cache and therefore does not send that data over again.

Because the datacenter WAFS appliance interacts with the Active Directory or Network Information Server on the local side, it learns which users are authorized to receive access to which files. This enables the datacenter WAFS appliance to manage file security via the remote appliance.

When a remote user asks for a file, the local WAFS appliance either returns the file immediately (if it is in cache), or requests it from the data-

WAFS appliances streamline and speed up old chatty protocols like NFS and CIFS

FIGURE 2 IT Infrastructure With WAFS Appliances



Scores of customers have told us they are quite happy with the performance improvements WAFS provide

WAFS + WAN and App Optimization = WDS (Wide Area Data Services)

The existing WAFS players, before they merged/bought/got acquired (except for Riverbed), could help the user only with speedier file access. While the feedback coming in from early WAFS implementations is very positive, it also is apparent that customers want help with all remote office applications—including packaged email, ERP, CRM and homegrown applications as well.

Simplification is incomplete without solving problems for all applications that run in the remote location. This is why WAFS players have been adding functionality and partnering with WAN optimization and application acceleration players to bring the larger set of remote office performance improvements to IT. We refer to the combination of these three functions as wide area data services (WDS). Figure A shows how these new, fully-featured platforms will eliminate the need for additional servers at the remote sites.

Progress toward WDS is already occurring. Recently, for example, we have seen:

- The acquisition of Tacit Networks by Packeteer, a public company that specializes in delivering quality of service (QOS) on WAN. Tacit also recently added support for Exchange to their product line.
- DiskSites has been acquired by Expand Networks, another player in the WAN optimization space.
- Actona has long been a part of Cisco, a player we anticipate will play additional cards in the WAN optimization space.
- Juniper already owns Peribit, a WAN optimization player that has yet to show its WAFS cards.

All these players recognize that WAFS solutions alone are useful, but they are not the complete WDS solution. Nor can WAN optimization and application acceleration solutions compete effectively without adding support for file-level optimization. At the moment these solutions are all distinct, even those which are being delivered by the same vendor. But we expect the next 12–18 months to bring strong integration from several vendors, and for WDS appliances to become a broad market reality.

Leading this integrating trend is Riverbed, which is enjoying particularly good growth, in large measure because they thought of the remote office in totality and built one architecture that covers the entire set of applications that typically run there. They come close to delivering a WDS appliance today, having designed their solution in layers, removing chattiness at each level such that even a brand new application will benfit from the efficiencies added in the lower layers. We would not be surprised to see an IPO from Riverbed very soon□

center WAFS appliance and then stores it in cache for quick access the next time it is requested. The appliance on the datacenter side ensures that a file is not made available to two users for changes at the same time.

WAFS appliances also manage content for staleness. In other words, they manage the coherency and consistency of files and caches. Cache-based WAFS solutions also enable the remote users to continue working if there is a WAN disruption, since their active files are cached locally.

Less Common WAFS Solutions

A very different approach is taken by Riverbed Technology. Their appliances shun caching and instead gain their effectiveness by making each layer of the protocol stack much more efficient. They do this by a variety of techniques but the most obvious are:

1.) Breaking the files down to smaller pieces and never sending the same piece twice to the same remote location.

2.) Optimizing the TCP protocol with window size manipulation.

3.) Optimizing NFS and CIFS protocols to reduce their chattiness.

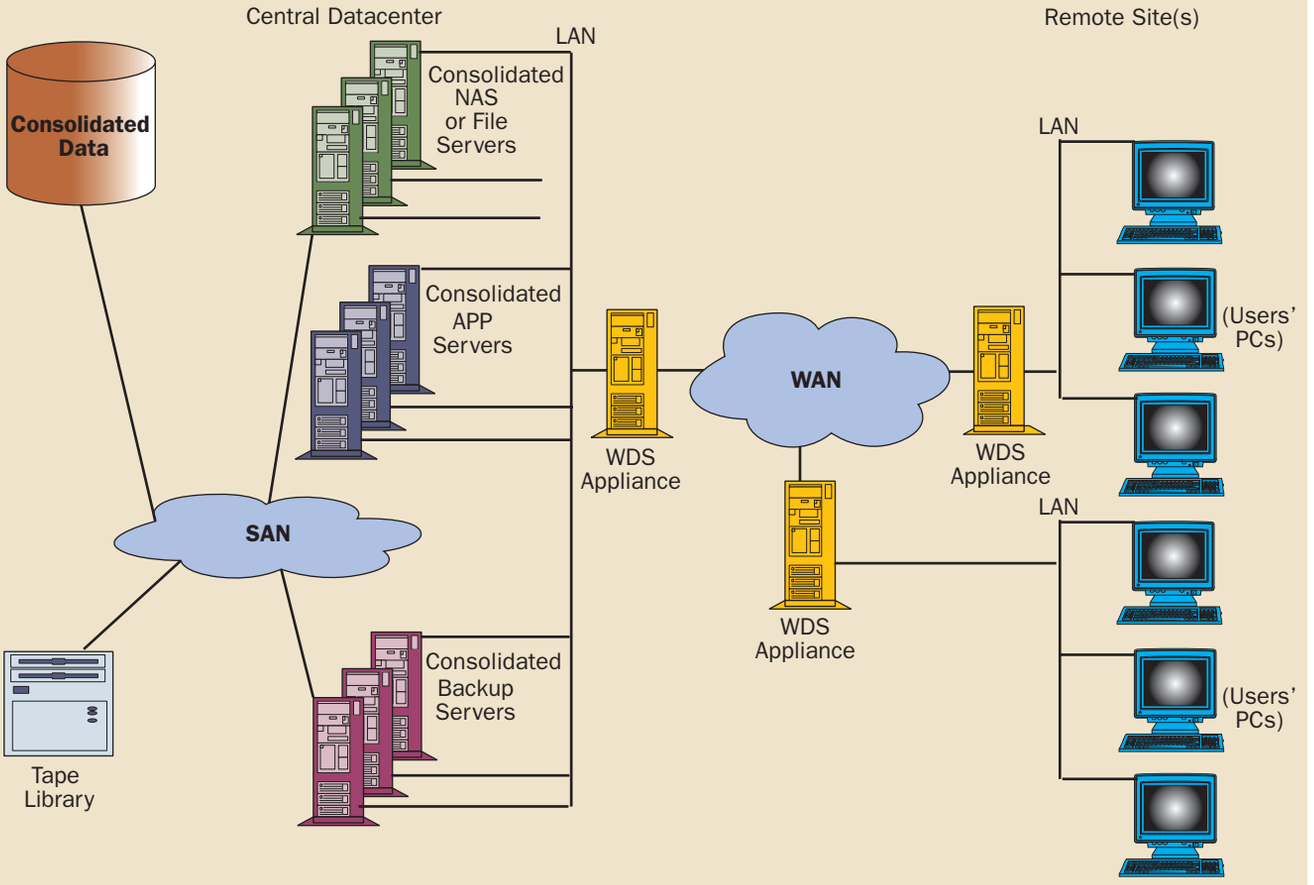
The Riverbed approach has a variety of side benefits (see "WAFS + WAN and App Optimization = WDS (Wide Area Data Services)"), chief among which is the fact that all applications that run over TCP—including Web surfing (http), file accessing (ftp), and emailing (MAPI)—automatically enjoy the benefits of TCP optimization. This is a major plus for this approach. Riverbed makes caching an optional feature.

Another different WAFS approach comes from Availl. Although Availl is often not included in the WAFS category, they essentially solve the same problem, albeit in a very different way.

The Availl solution makes available the entire set of files that need to be shared at the datacenter and at all remote sites simultaneously. They do this by locating the file server and the associated storage in all locations and by keeping the files synchronized at all sites, as they change.

Each time a file is accessed and changed by any user, anywhere, the changes are made at all locations, at the byte level, and in order. Since

FIGURE A Further Simplification With Wide Area Data Services (WDS) Appliances



only the bytes that are changed are transmitted, bandwidth requirements are minimal.

However, Availl's architecture requires that each site have its own complete data store. Consequently, the solution works best for customers who already have file servers and storage at each location and want to use them more efficiently. (Access rights are managed by interactions with the Active Directory or Network Information Server, as explained above.)

We have spoken with scores of customers who are quite happy with their various WAFS solutions. It is not uncommon for users to see a 50-to-1 speed-up for file access, even though the ratio varies anywhere from 10-to-1 to 150-to-1, depending on the task being performed. (Note, however, that WAFS solutions do nothing to bring WAN speeds up to LAN speeds.)

Comparing Different WAFS Solutions

How the WAFS solutions interact with the file server, the user and between the WAFS appliances determines their overall effectiveness. Several overriding characteristics are essential:

- Performance seen by the remote users must feel "LAN-like" to them. This does not always mean the speed must be identical to a LAN. A file accessed in 2 seconds feels no slower than the same file accessed in 10 milliseconds.
- Data integrity and security is a must. Data cannot be lost, or modified except by design, and then only by users who are authorized to do so.
- The appliances need to be self-managing, since the remote users are to be relieved of their IT responsibilities.
- To the remote users, access to data must be transparent. In other words, they must not be forced to interact with the system in any way other than what was customary before the appliances were installed.
- The appliances must be manageable centrally, by IT at the datacenter.
- The solution must allow two or more datacenters to participate with all the remote locations.
- The datacenter WAFS appliance must support hundreds of remote locations, whereas the remote office WAFS appliance must support from a few to 500 users.

Any enterprise that shares files with five or more locations could benefit from WAFS

Beyond that, WAFS solutions vary in three dimensions:

1.) Offering a variety of models for the appliance on the remote office side such that IT can purchase one that is the right size for the number of users in that location.

2.) Scalability of the datacenter appliance such that it can effectively deal with hundreds of remote offices.

3.) The ability to accelerate other applications besides file services.

Conclusion

Although mainstream enterprises are still evaluating or investigating WAFS, many larger enterprises have put WAFS into production, at least for their most critical sites. The financial services industry is at the forefront, but we also have seen installations in health care, manufacturing, technology and other verticals.

The results are extremely encouraging. Many of these companies have completely eliminated file servers and NAS boxes from the affected remote locations and are starting to enjoy the benefits of consolidation.

In the grand scheme of things, however, we are at very early stages of worldwide WAFS implementation. The next three years will see this technology cut across all industries and become pervasive. We believe WAFS technology applies to any enterprise, large or small, that has five or more branches and has a need to share files. The branch offices can be as small as 5 people to as large as a 500-person subsidiary or more.

New products are being developed and WAFS features added to other devices to make WAFS cost effective for even the smallest remote sites.

The good news about WAFS is that today we have technologies that are almost guaranteed to make remote users happier and IT more efficient in terms of shared file access. These technologies could not have come to the rescue at a better time□

Companies Mentioned In This Article

Availl (www.availl.com)
Cisco (www.cisco.com)
Expand Networks (www.expand.com)
Juniper (www.juniper.net)
Packeteer (www.packeteer.com)
Riverbed Technology (www.riverbed.com)
Tacit Networks (www.tacitnetworks.com)