WAN-Optimized Replication

Contents

Abstract

Replication has become part of the fabric of IT due to its multiple uses and benefits. However, as IT and business managers know, replicating data can quickly increase bandwidth and storage costs. WAN optimization technologies can help reduce the amount of bandwidth and storage capacity used for replication, enabling cost savings while maintaining application performance. However, in most cases, WAN optimization requires the purchase and installation of expensive hardware. X-IO solutions include built-in advanced data replication technology that is far more WAN-efficient than other types of solutions, providing significant savings in bandwidth, storage capacity, and costs.

The Many Uses of Data Replication

Data replication has become increasingly popular in small and large organizations, primarily for minimizing downtime associated with backup and disaster recovery (DR). Replication also enables organizations to use data in more than one way at a time. For example, if you replicate a database, you can continue using the original data set for production operations while using the copy to load a data warehouse or perform application testing. As a result, you get more value for your investment in information creation and management.

Remote replication can improve the backup process and shrink the backup window to virtually nothing. Many organizations take snapshots of their data, replicate it to a remote location, and complete the backup-to-tape process from the remote site. The primary site can immediately return to normal business functions, reducing downtime and enabling the speedy continuation of production (and revenue-generating) operations. With data replication, backups are no longer a disruptive process. Similarly, remotely replicated data can be used for disaster recovery. With data replicated offsite, it is protected from local corruptions and failures. Many organizations count on their replication technologies to deliver the data protection required for corporate and regulatory compliance.

Replication is also extremely useful for keeping up with maintenance schedules. Some organizations run the risk of ignoring important system maintenance because it is disruptive to take down critical applications, such as e-mail. By replicating your e-mail systems, you have to bring down your primary system, start-up operations on the secondary system, and then install patches, hardware upgrades, and perform other maintenance tasks on the primary system before bringing it back online. Without replication, these systems would need to be inoperable for long periods of time, interrupting business operations. Data replication is equally helpful for datacenter moves. You can fail over your systems to a remote site, move your production datacenter to its new location, and then fail back to the new datacenter.

With data volumes growing significantly in remote and branch offices, replication provides another benefit: manageability. Since most branch offices lack adequate IT staff and the skills to maintain and protect data, remote data is often highly vulnerable to loss and corruption. With remote replication, branch office data can be copied over the WAN to a central repository or datacenter, where IT can manage it according to corporate policies for optimal protection.

The Effect of WAN Link Saturation

Replication offers such business benefits as improved protection and simultaneous uses of information to help minimize total cost of ownership (TCO). However, growing replicated data volumes need more storage to retain them, and their bulk can saturate the WAN, causing increasing latency and poor response times. Plus, in today's business environment, there are many new applications vying for WAN bandwidth — not only backup and disaster recovery, but file sharing between locations, Web 2.0 applications such as YouTube and Salesforce, video files, training applications, and so on. Furthermore, bandwidth is expensive — few organizations can afford an unlimited supply of it.

Increasing WAN traffic can slow down application performance — a situation for which users have little tolerance. The only answer for many is to perform expensive WAN upgrades to handle the increased network traffic and maintain adequate application performance.

With so many applications competing for bandwidth and storage capacity being costly, any way to reduce the amount of data passed across the WAN and stored remotely is helpful. There are several solutions on the market to accelerate WAN bandwidth, but most require additional equipment. The question for users is: Do you want to spend a lot of money implementing a complex solution to solve the problem, or do you want to solve the problem simply and easily with software that already includes WAN optimization?

WAN Resiliency

The nature of IP transport often makes replication problematic. Even the best networks can get noisy or busy, packets can drop, and data has to be retransmitted. Many replication products have difficulty dealing with transport inconsistencies.

X-IO provides proven data protection solutions with built-in quality of service (QoS) tuning to handle network disruptions. X-IO replication with iglu blaze can automatically switch from continuous to periodic replication when transport quality falls below a configured level. Network throttling is also available to ensure that bandwidth is adequately allocated to the replication traffic if the link is shared with other applications without allocating more bandwidth than is necessary.

X-IO iglu blaze is also able to gracefully resume services following an outage. For example, if you are in the middle of replicating 100 GB of accumulated data and the WAN connection is lost near the end of the transmission, many replication tools will re-start at the beginning, resending the full 100 GB back over the wire while falling further behind in light of the accumulation of new data. In extreme cases, systems can fall so far behind that they cannot catch up. With X-IO replication technology, instead of resending the full 100 GB, only the remaining data is sent, allowing the process to continue without interruption.

WAN Optimized X-IO Solutions

X-IO solutions such as X-IO® iglu blaze includes advanced replication technology that minimizes the amount of data sent over the WAN as well as the amount of storage capacity required. There is nothing additional to purchase because ReplOptimize™ technology from X-IO is built into these solutions. This patented data deduplication mechanism eliminates redundant, block-level changes that occur due to application and file system inefficiencies.

With most solutions, changing even a single bit of data requires a cluster of kilobytes to be written over the wire, but the bulk of this is redundant data (re-writes) or white space. As a result, when replicating changed data, these solutions will transfer more data than was actually changed. This wastes storage capacity, clogs the network, and creates unnecessary costs. Regardless of which files or databases your organization wants to replicate, X-IO solutions will transfer the least amount of data. Compared to other replication methods, users see bandwidth reductions of 80-90% or even more, depending on the characteristics of their data.

Array-based systems typically require the same disk array at both sites. All disk resources and capacity expansion must be done with the same types of disk subsystems, limiting corporate buying options and price negotiating. Because X-IO offers HDD, Hybrid, and All-Flash solutions, lower-cost storage can be used at the remote site in order to meet budget constraints or to extend disaster recovery protection to more applications.

The following graphic illustrates the difference between X-IO data replication and storage array-based solutions.
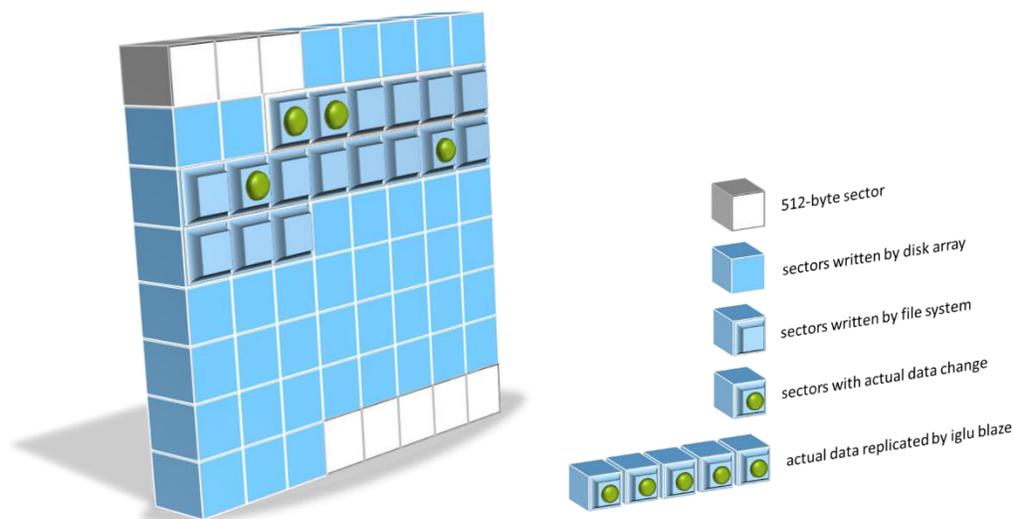
Each square represents a 512-byte sector on a disk array. This is the standard size for most disk systems and represents the smallest possible segment of data written to a disk. When an actual data change is made, the application needs to write one or more 512-byte sectors. In the illustration, four sectors are actually updated. However, file systems work in larger segments, so the file system may write an 8 KB section or 16 sectors in order to make the change. The storage array works in even larger segments, so in order to store the change it must transfer an entire 32 KB track from the source to the target disk (the

shaded blocks). In this case, an array-based replication tool would copy the entire 32 KB track, or 64 sectors, in order to replicate the four sectors that represent new data. The 60 sectors are simply the same data written again for the sake of disk array efficiency. While this may be efficient from the perspective of the disk array, it is terribly inefficient for replication, as it would send 15 times more sectors over the WAN than what is actually needed to replicate the true data changes.

Compare this to the next graphic. Instead of sending 32 KB of data over, it sends only the 512 bytes that were actually changed. That makes the X-IO solution 64 times more efficient than storage array-based solutions on a per-sector basis in terms of WAN traffic and storage capacity. Similarly, WAN-optimized replication from X-IO is significantly more efficient than other types of solutions.

Replication method	Minimum data amount moved for 512-byte change	How X-IO compares
X-IO iglu blaze	512 bytes	
Array-based	32 kilobytes	64 times more efficient
Appliance-based	16 kilobytes	32 times more efficient
Storage-appliance based	8 kilobytes	16 times more efficient

Save Bandwidth, Storage, and Money with X-IO Technology

By optimizing WAN replication, X-IO replication solutions can help you boost network performance without having to purchase expensive hardware. This improved network response time works in both directions — whether you are replicating data to a target or you need to recover data after a corruption or failure. With significantly less data to move, your recovery operations occur much faster, speeding your organization's return to productivity. Because this technology is built into our solutions, your investment is minimal and the return on investment (ROI) is immediate.

To demonstrate the potential ROI, let's review the following use case. Over an 83-day period, an organization replicated data changes that took up 1.2 TB of capacity per day. However, the X-IO solution removed 1 TB of redundant data on a daily basis, for a remarkable 84% savings in WAN bandwidth and storage requirements. If the firm's staff members were to copy all the data changes made at the file system level, it would have taken 68 hours to replicate each day's worth of data, an untenable situation. X-IO iglu blaze would reduce the firm's daily replication time from 68 hours to 10.7 hours and save $600,000 annually in WAN costs.

X-IO solutions with built-in data replication are available to organizations of any size or budget and facilitate the use of existing remote locations for replication. In the past, many organizations had to implement expensive third-party disaster recovery solutions. With iglu blaze, they can replicate to a branch office by simply using their current WAN. This makes it easier to comply with increasingly stringent regulatory requirements for data protection.
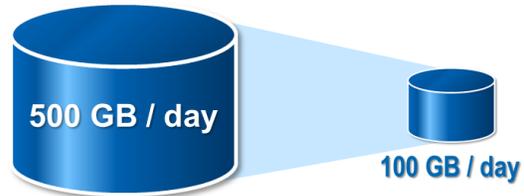
Without ReplOptimize bandwidth reduction

500 GB

DS3 = 30 hours, 9 minutes
OC3 = 8 hours, 28 minutes

An OC3 WAN link is costly and required to meet an 8-hour replication window

With ReplOptimize bandwidth reduction

**500 GB / day**

100 GB / day

DS3 = 6 hours, 1 minute
OC3 = 1 hour, 41 minutes

A lower-cost DS3 WAN link completes replication in less time than an OC3 without X-IO ReplOptimize replication

ReplOptimize technology from X-IO greatly reduces the amount of bandwidth needed for replication. If your systems generated 500 GB of data per day at the file system level, it would require an expensive OC3 WAN link to replicate all of the changes within an eight-hour window. Using ReplOptimize, over 80% of the file system changes are filtered out (the ratio is even higher with certain data types). This means you can replicate the same daily changes using a much less costly DS3 WAN replication link while actually

DS3 without ReplOptimize

Replicate only 9 application servers in 8 hours

DS3 with ReplOptimize

Replicate all 45 application servers in 8 hours
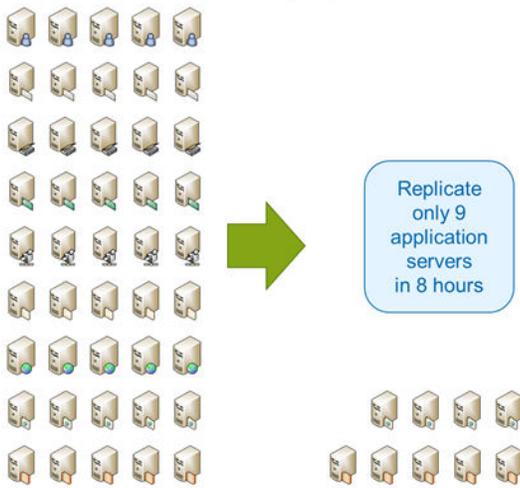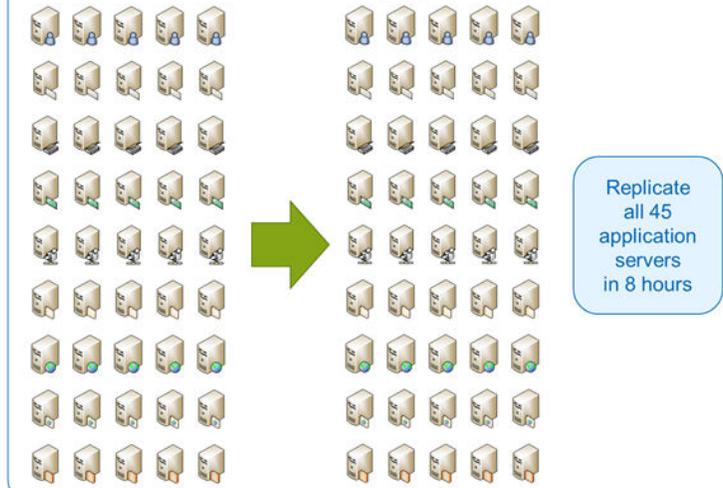
finishing two hours before the eight-hour window closes.

Due to the high cost of high-speed WAN replication links, it is common to leave many systems unprotected by replication. Often protection is limited only to the most critical systems. ReplOptimize technology lets you extend coverage without increasing bandwidth. In this example, 45 application servers created new data at an average rate of 15 GB per day each. Using a DS3 WAN link, you can protect only

nine servers within an eight-hour replication window. By adding ReplOptimize technology, you can protect all 45 application servers within the same window.

Summary

With cost savings of paramount importance in today's economy, simple, cost-effective WAN-optimized replication solutions such as those from X-IO can ease multiple IT tasks. X-IO technology enables you to improve network performance, reduce contention for network bandwidth, and meet recovery point and recovery time objectives (RPO and RTO). Now, at last, you can complete backup, disaster recovery, maintenance, remote office consolidation, data center moves, and other replication tasks at a minimal cost.