

Tech Note: Replication Performance vs. Bandwidth

The Problem with Remote Data Replication

Ideally, enterprises would like to perform frequent replications during the business day in order to reduce the data loss risk window. If the size of a replication becomes large, this traffic can consume a large amount of WAN bandwidth, interfering with other business traffic. The result can be that the number of replications is minimized during the business day, or even postponed until after business hours. This increases the data loss risk window. To solve this problem, WAN accelerators can be deployed on the WAN link between data centers and the remote replication site.

WAN Accelerators to Accelerate Replication Traffic

All WAN accelerators work on the same principle of intercepting traffic and optimizing it before it traverses the WAN. The optimization is accomplished using two techniques: TCP optimization and compression. TCP optimization ensures that on a high latency network, TCP's protocol inefficiencies do not slow down the transfer. Compression reduces the amount of data that needs to be transmitted across the WAN.

For compression, WAN accelerators fall into two categories: memory-based systems and disk-based systems. Memory-based WAN accelerators typically pass the data through a compression algorithm and compress it on the fly. Since they have a limited compression database, they achieve a lower compression ratio. More advanced systems are disk-based; they can store gigabytes of data patterns on disk, and when data passes through they see whether the data pattern has been seen before. If it has, they send a token to represent it and recreate it on the device on the other side of the WAN. Disk-based systems get a much higher compression ratio, since a single token can represent gigabytes of data and enable a significant reduction in the amount of data being replicated. This results in reduced WAN bandwidth requirements. This method of deduplication associated with disk-based compression of data across a WAN is more efficient than typical block or file-level deduplication deployed in backup and replication products.

Selecting the Best WAN Acceleration Solution for EqualLogic Replication

What are the characteristics of an ideal WAN accelerator? The best WAN accelerators are those that are able to provide the fastest replication with the most minimal use of resources (i.e., fastest replication and lowest WAN bandwidth usage).

To get the fastest replication you want to have the highest throughput on the low latency LAN between the filer and the WAN accelerator appliance so that replication over the WAN can proceed as fast as possible. Then you want the best compression and data deduplication system in order to reduce the traffic over the high latency WAN:

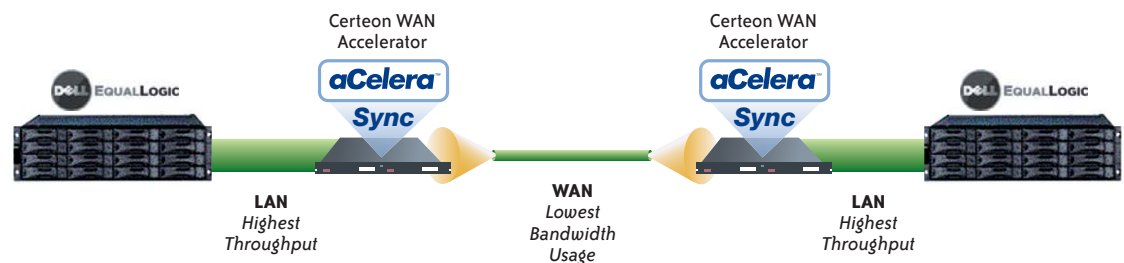


Figure 1: An ideal replication environment: Fastest replication with lowest WAN bandwidth usage

Using just any disk-based compression system is not sufficient to get the fastest replication and lowest WAN bandwidth usage. Certeon has optimized disk-based compression to provide the industry's leading system for replication environments by using the following characteristics:

- 1. Scalability** — The history store is scalable to terabytes in size. The larger the history store, the better the compression ratio.
- 2. Memory Matching** — During compression all matching is done in fast RAM, rather than by seeking on disk, which slows down a system as the disk starts to fill up.
- 3. Instant Access** — *aCelerate* stores the data in a manner that allows it to be retrieved quickly off the disk without scanning the disk.
- 4. No Defrag** — As data is added to disk, the more fragmented it becomes, and the more time the disk head spends seeking. Certeon's *aCelerate* ensures that data is written in a way that will not fragment.
- 5. Single Instance Store** — *aCelerate* keeps only a single copy of data, even with multiple peers.
- 6. Cross Protocol Capability** — *aCelerate* stores only a single copy of data, even if it is transmitted using different protocols.

Tech Note: Replication Performance vs. Bandwidth

Achieving Optimum Speed of Replication

The goal of backup and replication is to ensure secure off-site backup of company data. Designing a solution that allows frequent replication without disruption of normal business day traffic is critical. The technologies discussed in this document — optimal use of disk and memory resources, granular deduplication of data via the single history store and TCP protocol optimization — collectively ensure the best possible replication data in the shortest period of time. An important point to make here about Certeon’s WAN optimization technology is that the amount of data that has to be replicated across the WAN is reduced as well as the bandwidth required to send that data. The result is that replications can be run more frequently during the business day, narrowing the data loss risk window while still providing more bandwidth capacity to support additional users and other applications.

Metrics for Measuring the Effectiveness of Replication

To understand how much WAN optimization is benefiting replication times and required bandwidth requirements, it is important to make some baseline measurements prior to deploying the optimization solution. To do this, measure the total replication time and bandwidth consumed across the WAN for several replications prior to installing a WAN optimization solution. After installing an optimization solution, repeat the measurements for several replications and compare to the baseline. Both bandwidth consumed and replication times are important to note. If you will be replicating during the business day, the bandwidth consumed must be reduced to allow other business traffic to traverse the WAN.

Beyond Generic WAN Optimization

Certeon has created a specific WAN optimization product for data replication called *aCeler*a Sync. This software-based accelerator is preconfigured for optimal performance in replication scenarios. Certeon has also added specific software in the Sync product for iSCSI traffic acceleration, an iSCSI “blueprint” that improves performance beyond what normal TCP acceleration techniques can offer. These “Designed for Dell® EqualLogic™” features optimize the iSCSI transport over the WAN and improve the history store characteristics for more intelligent management of the history store. Below is an example of how an EqualLogic customer has been able to take advantage of *aCeler*a Sync to increase the frequency of replications and reduce the data loss risk window.

Customer experience: Change in the replication schedules due to *aCeler*a

	Previous	Current with <i>aCeler</i> a Sync
Exchange	Every 6 hours	Every hour
SQL	Every 4 hours	Every 15 minutes
File Server	Every 6 hours	Every hour
C Drives	Every week	Every day

Average compression ratios seen are between 8:1 and 9:1.

Summary — The Best of Both Solutions

When you select EqualLogic and Certeon *aCeler*a Sync, you will have the highest performance backup and replication solution available. The full featured EqualLogic appliances, combined with the “Designed for EqualLogic” *aCeler*a Sync WAN optimization virtual machine software, will give you a total solution that allows more frequent business day replications, uses less overall bandwidth and narrows the data loss risk window. As a software virtual machine, *aCeler*a Sync is a low-cost solution to improving replication performance.

For more information, contact us at info@certeon.com.



4 Van de Graaff Drive
Burlington, MA 01803
USA: 781.425.5200
Toll Free: 877.221.6688
EMEA: +33 (0)9.75.18.01.56
info@certeon.com
www.certeon.com