Lab Validation Report

IBM Real-time Compression

Reducing Storage Capacity and Costs without Compromise

By Brian Garrett with Claude Bouffard

January 2009

Contents

ESG Lab Reports

The goal of ESG Lab reports is to educate IT professionals about emerging technologies and products in the storage, data management and information security industries. ESG Lab reports are not meant to replace the evaluation process that should be conducted before making purchasing decisions, but rather to provide insight into these emerging technologies. Our objective is to go over some of the more valuable feature/functions of products, show how they can be used to solve real customer problems and identify any areas needing improvement. ESG Lab's expert third-party perspective is based on our own hands-on testing as well as on interviews with customers who use these products in production environments. This ESG Lab report was sponsored by IBM.

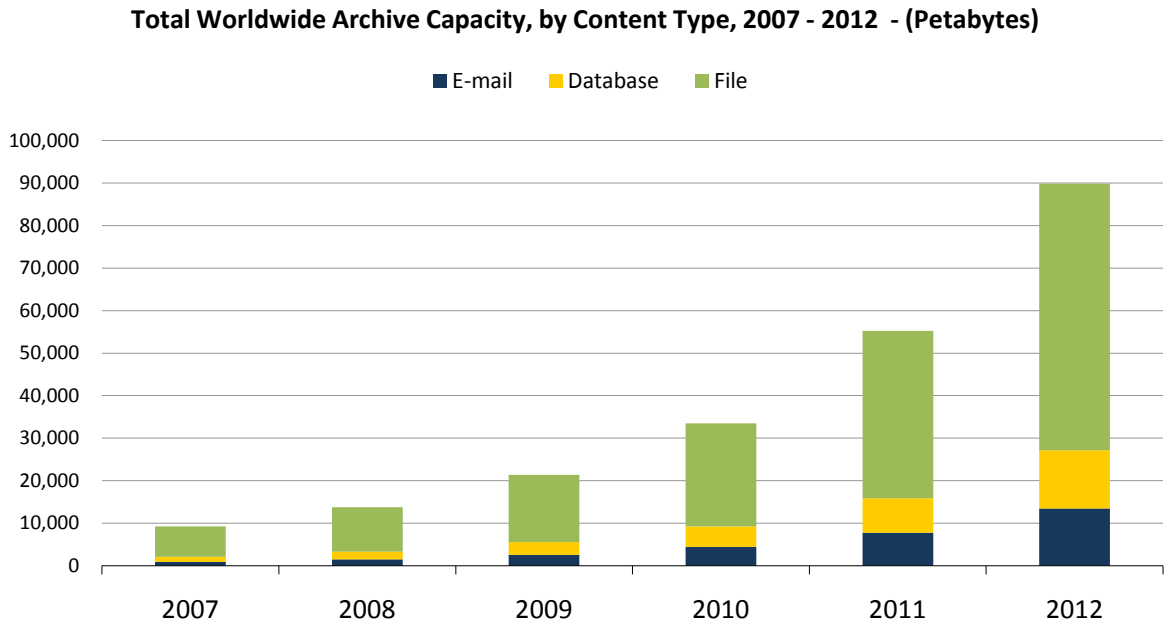
Introduction

Explosive data growth is driving IT managers to look for new ways to reduce the cost, complexity and space requirements of their storage infrastructure. This ESG Lab report examines an innovative high speed data compression appliance from IBM Real-time Compression. Hands-on testing is used to demonstrate non-disruptive deployment, loss-less capacity savings, real-time performance, and non-stop fault tolerance.

Background

Recent ESG research indicates that IT managers are drowning in a sea of file data. Unstructured file content accounted for 77% of global digital archive capacity in 2007 and is expected to constitute the bulk of digital assets for the foreseeable future.[1] Driven by regulatory compliance, corporate governance, litigation support, records management, and data management, total worldwide archive capacity is projected to increase nearly tenfold between 2007 and 2012.

Figure 1. World Wide Digital Archive Capacity



Source: Enterprise Strategy Group, 2009.

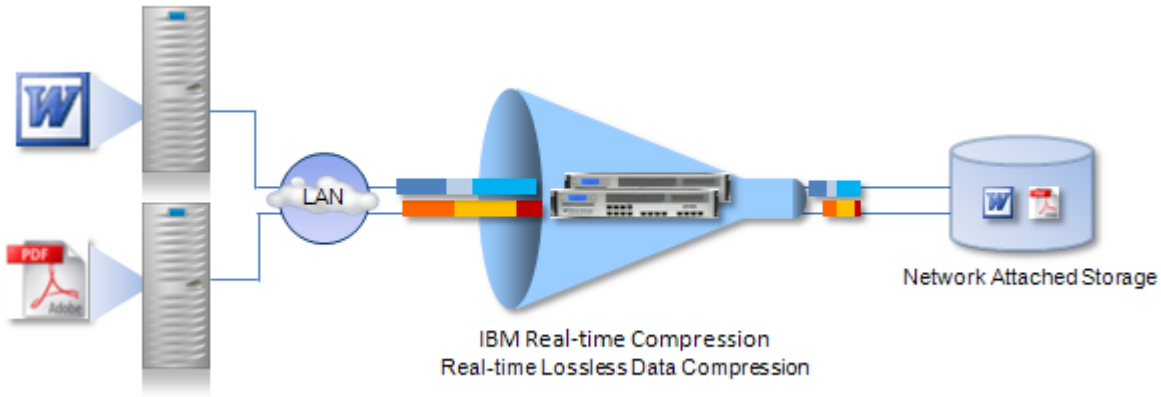
Network-attached file systems are the media of choice for each of the types of digital archive data shown in Figure 1. Network-attached file systems are also the media of choice for a number of primary storage applications, including home directories, shared corporate data, and shared project files. A rapidly growing number of organizations are also using network-attached disk arrays for virtual server images and for database applications. Put it all together and ESG sees a perfect storm of challenges for IT administrators forming on the horizon. According to a recent ESG survey, IT managers report that the challenges include the cost of storage systems (33%), the pace of overall data growth (31%), and the rapid-growth of unstructured file data (22%).[2]

[1] Source: ESG Research Report, *Digital Archiving Survey,* November 2007
[2] Source: ESG Research Report, *Medium-sized Business Server and Storage Priorities*, June 2008.

IBM Real-time Compression Data Compression Appliance

The IBM Real-time Compression STN-6000 is a family of real-time loss-less data compression appliances for primary network-attached storage (NAS) systems. STN-6000 appliances are installed as a bridge between clients and NAS systems, as shown in Figure 2. In this example, a Windows client is storing a Microsoft Word document on a network-attached disk array using the CIFS protocol. The document is compressed as it passes through an IBM Real-time Compression appliance using high-speed compression algorithms. Repeating data patterns, shown as colored blocks in Figure 2, are compressed in real-time before being written to disk. A similar operation is shown for a Linux client storing a PDF file on a network-attached disk array using the NFS protocol.

Figure 2. Introducing IBM Real-time Compression Real-Time Loss-Less Data Compression



While the hardware and algorithms at the heart of IBM Real-time Compression are well known, the technology used to perform the compression in real-time is novel and powerful. Instead of performing compression as a disruptive background task, IBM Real-time Compression performs compression at real-time speeds as files are initially created and stored. As a result, a high-speed real-time IBM Real-time Compression appliance delivers:

- Capacity savings of 75% on average
- No performance compromise
- In most cases, performance actually improves due to a decrease in disk activity
- Non-disruptive installation that works with existing networks and NAS infrastructure
- Clustered fault tolerance
- Loss-less data integrity
- Reduced power, cooling, and space requirements

The balance of this report presents the results of ESG Lab validation tests designed to demonstrate how IBM Real-time Compression appliances reduce storage capacity and costs without compromise.

ESG Lab Validation

ESG Lab performed hands-on evaluation and testing of STN-6000 data compression appliances at an IBM Real-time Compression engineering facility. Testing began with an examination of ease of deployment.

Planning with PrediSave

The ESG Lab evaluation began with a quick look at the IBM Real-time Compression PrediSave utility. PrediSave is a host-based utility which predicts the capacity savings that can be expected after deploying an IBM Real-time Compression appliance. As shown in Figure 3, a high level prediction of overall capacity savings is presented along with a graph depicting the capacity savings for the each type of file detected.

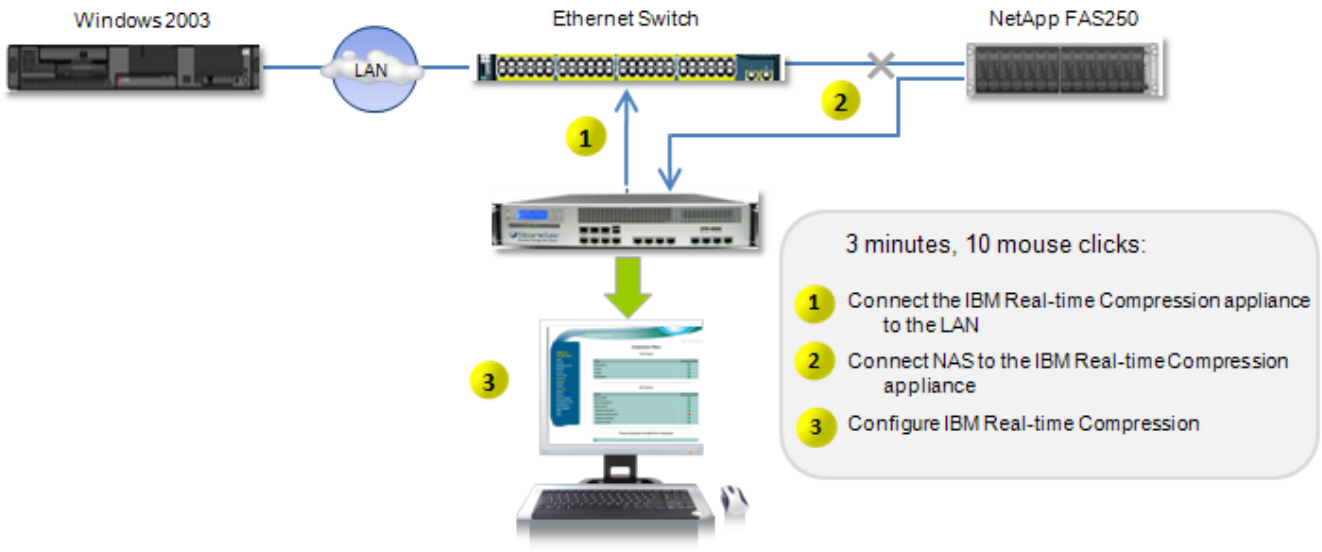
Figure 3. Planning with the PrediSave utility

The information provided by the PrediSave utility can be used to define policies, which exclude the processing of files that have already been compressed (e.g., zip files). In the example shown in Figure 3, the utility predicted capacity savings of 83.62% for all data and 89.38% for compressible data. The difference between these two values (5.76% in this example) is the capacity consumed by already compressed data.

Non-disruptive Installation

IBM Real-time Compression data compression appliances are positioned between a network switch and one or more NAS storage devices as an inline, transparent bridge shown in Figure 4.ESG Lab used a simple three-step process to install and configure an IBM Real-time Compression appliance between an IBM x305 PC running the Windows 2003 operating system and a NetApp FAS250 disk array.

Figure 4. Non-disruptive Installation



First, the STN-6000 appliance was plugged into the Ethernet switch. Next, the NetApp NAS system was disconnected from the Ethernet switch and connected directly to the STN-6000 appliance. The IBM Real-time Compression management console was launched from a web browser to finish the configuration. After basic network, password, and permission settings had been provided, the exportable file systems presented by the NetApp FAS250 were configured for compression. The ability to define exports and file type to be excluded from compression was noted.

Three minutes and ten mouse clicks after getting started, files being written to the NetApp NAS system were being automatically compressed in real-time by the IBM Real-time Compression appliance. From an end-user standpoint, the networked file share looked and behaved exactly as before IBM Real-time Compression had been installed.

Why This Matters

Reducing the cost of primary and archive NAS capacity is a good idea, but not if it requires massive changes to existing network and NAS infrastructure. Replacing existing NAS systems before they reach the end of lease or depreciation schedule would be a waste of time and money. Changing the way the network is configured between clients and NAS systems introduces risk.

ESG Lab has confirmed that IBM Real-time Compression appliances can be installed and configured in a matter of minutes. Leveraging existing investments in NAS technology, the non-disruptive deployment of IBM Real-time Compression appliances reduces manpower and risk as it sets the stage for dramatic capacity savings.

Capacity Savings

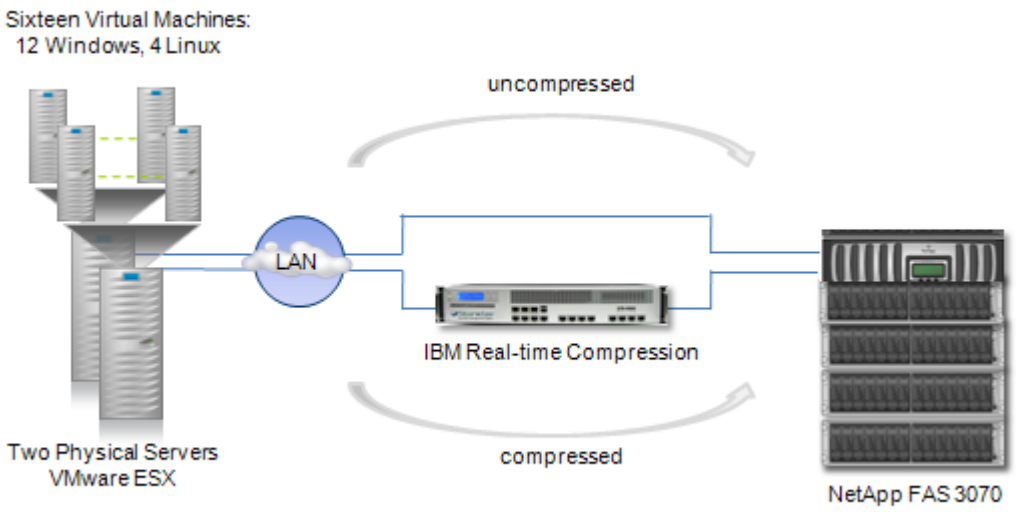
IBM Real-time Compression data compression appliances are designed to reduce the capacity of growing volumes of network file data by providing real-time data compression on-the-fly. Files being written are compressed in real-time as they pass through the appliance. Already compressed files are de-compressed as they are read. IBM Real-time Compression claims that network file capacity can be reduced by 75% on average and up to 90% or more for highly compressible data types (e.g., databases). Unlike data deduplication technology, which is designed to eliminate multiple copies of data that accumulate over time (e.g., a month's worth of accumulated backup data), IBM Real-time Compression reduces capacity for primary and archive data as it is stored for the first time.3

ESG Lab Testing

Capacity savings were measured using a variety of test beds as shown in Figure 5. Each test bed was designed to compare the capacity utilization of a file system before IBM Real-time Compression was introduced (uncompressed) to the same file system after it had been written through an IBM Real-time Compression appliance (compressed).

In this example, the capacity savings that can be achieved for virtual server images stored on a VMware .vmdk file mounted via NFS was evaluated. Two physical servers running VMware ESX software were configured. Each physical server was configured with six Windows and two Linux virtual machines. The guest operating system and local drive data for each virtual machine was stored in a .vmdk file on a NetApp FAS3070 accessed using the NFS protocol.

Figure 5. The ESG Lab VMware Test Bed



Capacity savings were measured as a 500 MB source file directory tree was copied to the NFS-mounted Windows and Linux images. The files were copied to two different directories on a NetApp NAS system: one uncompressed and the other compressed with an IBM Real-time Compression appliance. The 10.6 GB collection of source files on the Windows platform consumed only 2.6 GB after IBM Real-time Compression, as shown in Figure 6 and Table 1. A similar test was performed from a Windows 2003 client using a 562 MB collection of common office productivity files (e.g., documents, text files, xml files, presentations, etc.). In this example, a capacity savings of 90% was

[3] While IBM Real-time Compression is different than data deduplication, it should be noted that the two technologies complement each other. When deployed together, the capacity savings are magnified as data deduplication eliminates multiple copies of data that have been compressed with IBM Real-time Compression.

recorded. Last, but not least, a 287 GB ERP database instance created by the Oracle Apps R12 Vision demo consumed only 30 GB after IBM Real-time Compression.

Figure 6. IBM Real-time Compression Capacity Savings

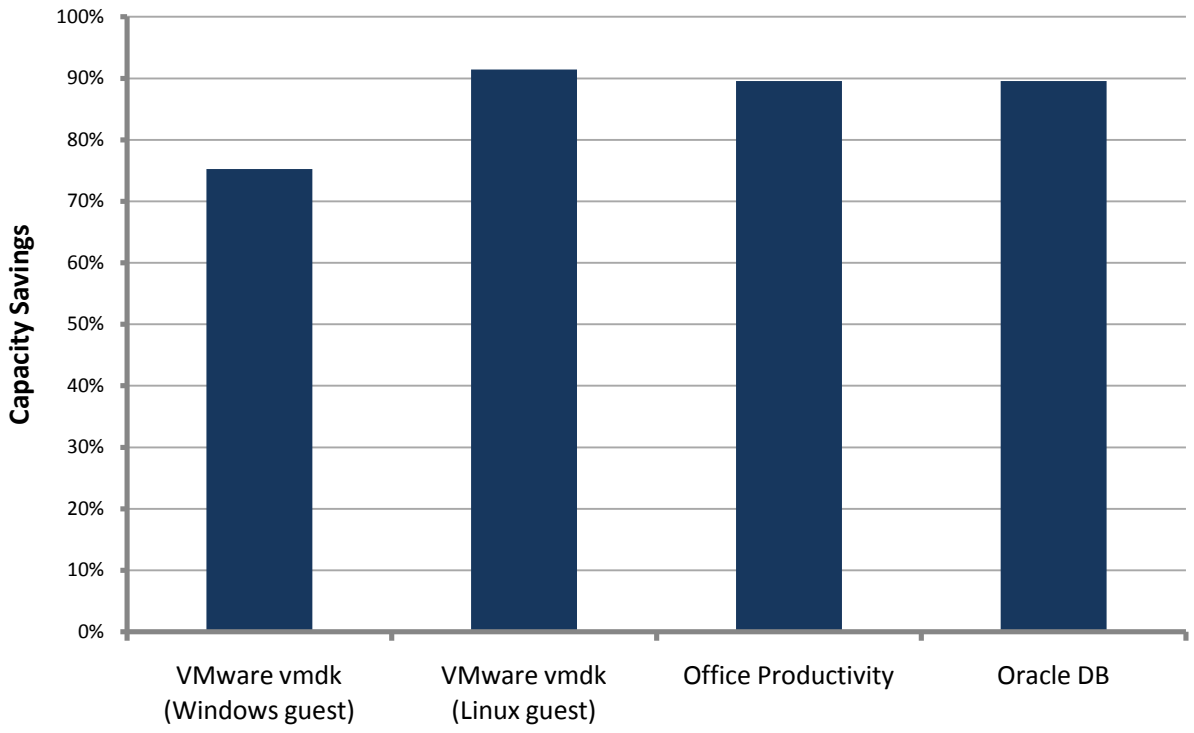


Table 1: IBM Real-time Compression Capacity Savings

	Before IBM Real-time Compression (GB)	After IBM Real-time Compression (GB)	Capacity Savings (GB)	Percent Saved
VMware VMDK (Windows)	10.5	2.6	7.9	75%
VMware VMDK (Linux)	10.5	0.9	9.6	91%
Office Productivity	564	59	505	90%
Oracle Database	287.0	30.0	257	90%

Why This Matters

Reducing the cost of storage capacity is a key concern for IT managers, especially for digital file content which accounted for 77% of global digital archive capacity in 2007 and is expected to constitute the bulk of digital assets for the foreseeable future.[4]

ESG Lab testing with real-world data has confirmed that a non-disruptive IBM Real-time Compression appliance can not only be used to used to reduce primary NAS disk capacity between 75% and 91%, it can also be used to create cost and operational savings on multiple tiers of storage throughout the data lifecycle.

[4] Source: ESG Research Report, *Digital Archiving Survey*, November 2007.

Performance

While data compression has been widely adopted in the IT industry, it has not historically been used with primary storage due to performance concerns. IBM Real-time Compression used a novel approach to tackle this problem. Based upon decades of expertise developing real-time processing and network optimization algorithms, a purpose-built appliance was developed to implement industry-standard compression algorithms in real-time. The IBM Real-time Compression approach was also designed to mitigate the performance drawbacks of mechanically slow disk. As disk capacity is reduced, so are the number of I/O requests that need to be performed to complete a file operation. As the number of disk accesses goes down, the performance for each file operation speeds up. Put it all together and a STN-6000 appliance not only avoids the performance penalty of real-line compression, it can actually be used to improve application-level performance.

ESG Lab Testing

The VMware test bed shown earlier in Figure 5 was used for the first performance test. This simple test recorded the wall clock time required to copy a 500 MB source file directory. As shown in Figure 7, the file copy operation running in a Windows guest operating system completed 28% faster after the IBM Real-time Compression appliance was installed between the VMware ESX client and the NetApp FAS3070 disk array. A copy of the same data within a Linux guest operating system completed 19% faster.

Figure 7. Less Disk Improves Performance for a VMware File Copy

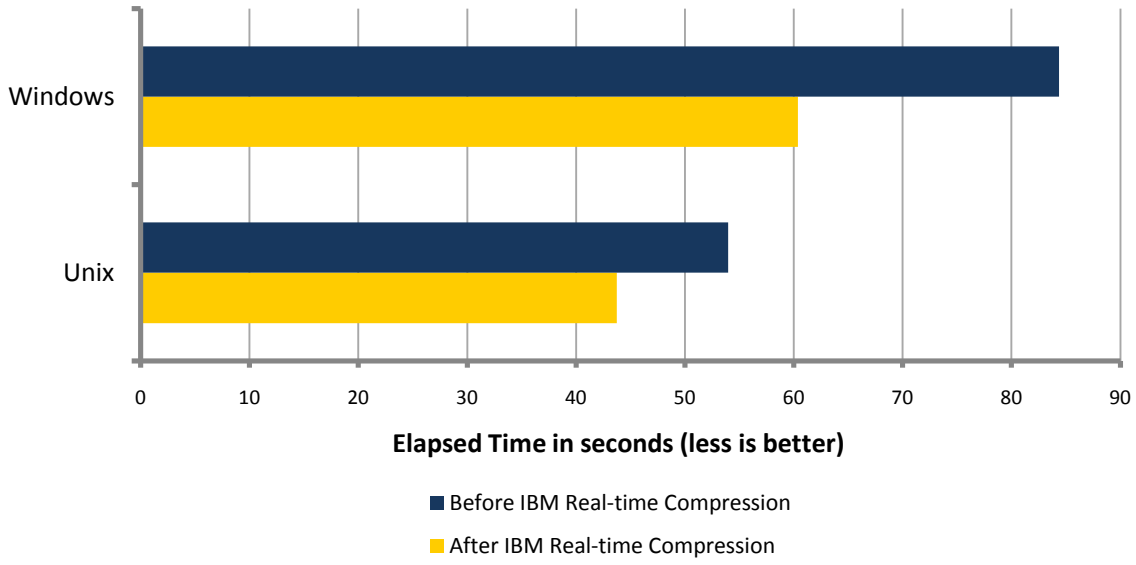


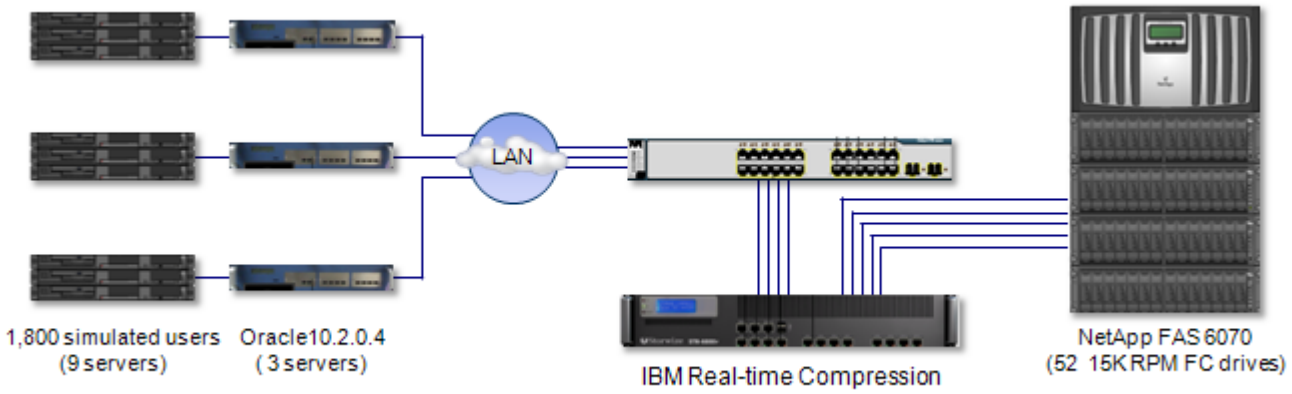
Table 2: VMware VMDK Copies

	Before IBM Real-time Compression	After IBM Real-time Compression	Percent Faster
Windows	84.36	60.39	28%
Linux	53.96	43.75	19%

An online transaction processing (OLTP) database application which simulates a warehouse order entry application for 1,800 simulated users was tested. Performance was analyzed before and after the introduction of an IBM Real-time Compression STN-6000 appliance. The configuration after the IBM Real-time Compression appliance had been installed is shown in Figure 8.[5]

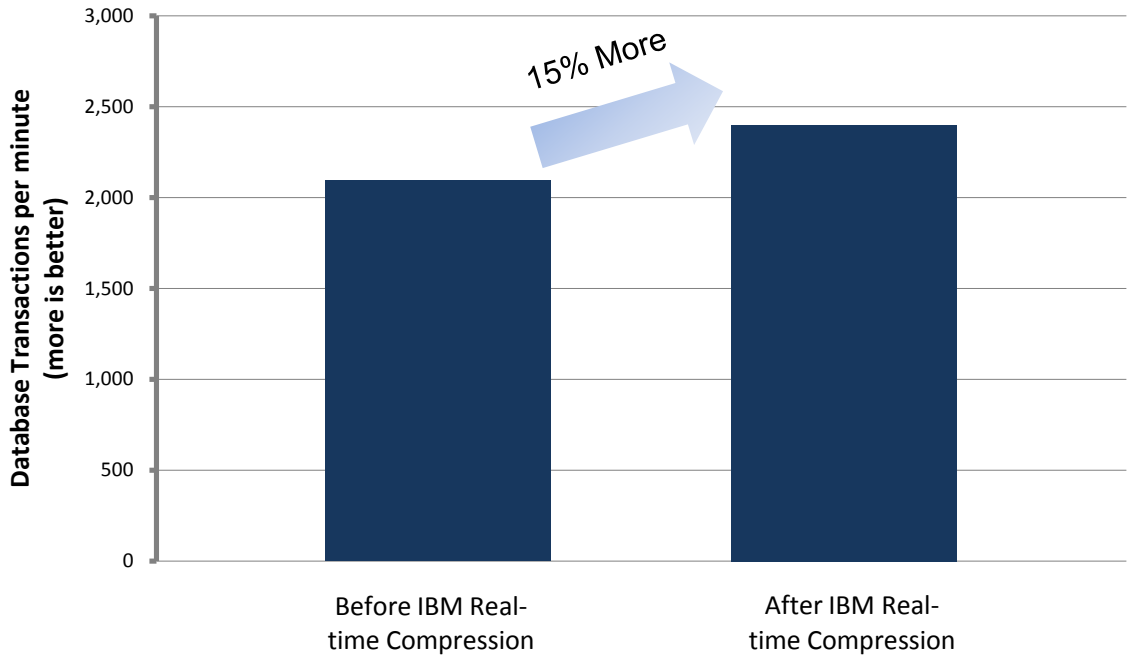
[5] Configuration details are listed in the Appendix

Figure 8. Evaluating Oracle OLTP Database Performance



The number of database transactions per minute for 1,800 simulated users was recorded. The average number of transactions per minute recorded during three test runs was used to analyze the impact of IBM Real-time Compression. As shown in Figure 9, IBM Real-time Compression increased the number of user level transactions by 15%. In other words, the database application was able to do 15% more work after the introduction of IBM Real-time Compression—as it simultaneously reduced disk capacity requirements by 80%.

Figure 9. Less Disk Improves Performance for an Oracle Database Application

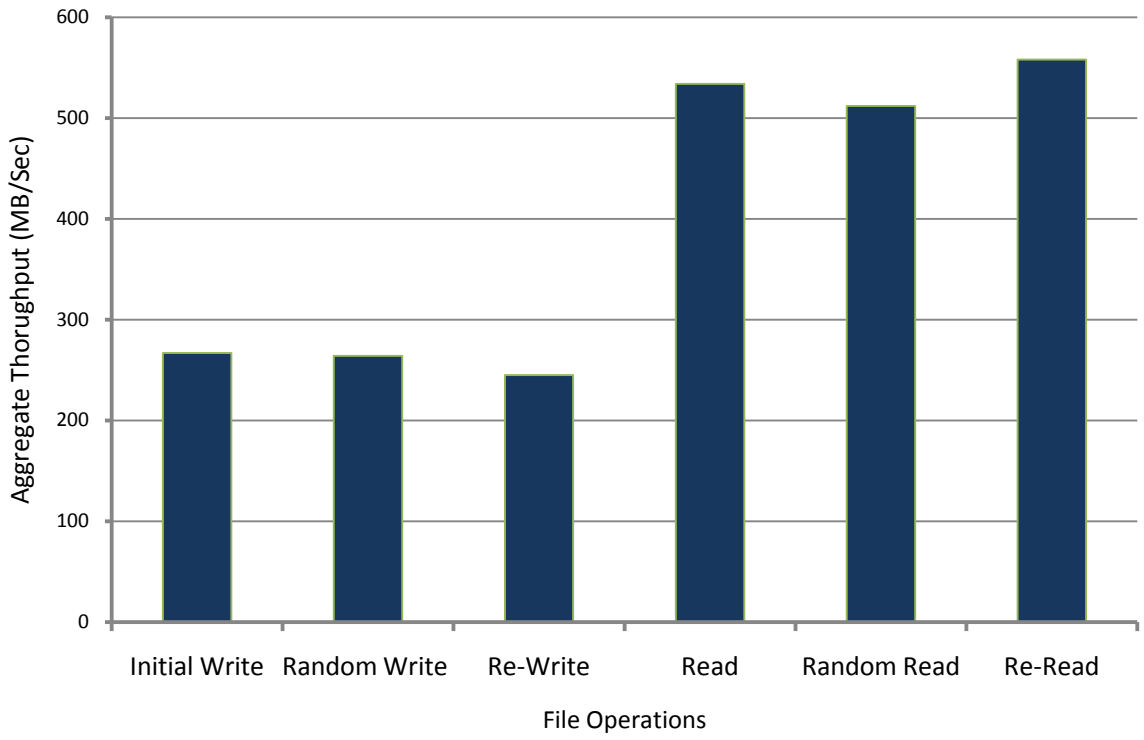


ESG Lab noted that IBM Real-time Compression not only increased the amount of transactions that the system could handle, it also improved response times. Response time improved between 20% and 90% depending on the type of transaction. For example, an order-status transaction completed 43% quicker and a new order transaction completed 90% faster.

A final test was performed to determine how well an IBM Real-time Compression appliance can maintain sustained performance levels when attached to a NAS system that is extremely busy. The industry standard Iozone utility was used to perform typical file operations running on twenty IBM x335 servers in parallel. The servers were connected to a NetApp FAS6070 with 56 15K RPM FC drives configured as a single aggregate. Ten gigabit Ethernet ports on a

high-end IBM Real-time Compression STN-6800 appliance were used to create a bridged connection between the IBM x335 clients and the NetApp FAS6070. With two IBM Real-time Compression STN-6800 ports per logical network connection between a Cisco Catalyst 3750G Ethernet switch and the disk array, five gigabit Ethernet ports were being compressed in parallel by a single IBM Real-time Compression appliance. The results are shown in Figure 10.

Figure 10. Real-Time Aggregate Throughput Over Five Gigabit Ethernet Interfaces



ESG Lab was impressed with a peak throughput of more than 558 MB/sec. A peak aggregate throughput of more than 500 MB/sec is close to the theoretical limit of NFS running over five gigabit Ethernet interfaces. A similar set of results peaking at 621 MB/sec over six gigabit Ethernet interfaces was audited. Near wire-speed performance is a great result for any NAS solution. In this case, it's outstanding given the fact that near wire-speed results were achieved as the IBM Real-time Compression appliance was compressing data in real-time.

Why This Matters

Using compression to reduce the cost of disk capacity is a great idea, but not if it degrades performance for primary and archive applications that rely on fast NAS storage capacity.

With aggregate throughput capabilities of more than 100 MB/sec per compressed gigabit Ethernet interface and the ability to increase the number of Oracle database transactions by 15%, ESG Lab has confirmed that IBM Real-time Compression can be used to improve the performance of primary and archive applications that rely on NAS storage capacity.

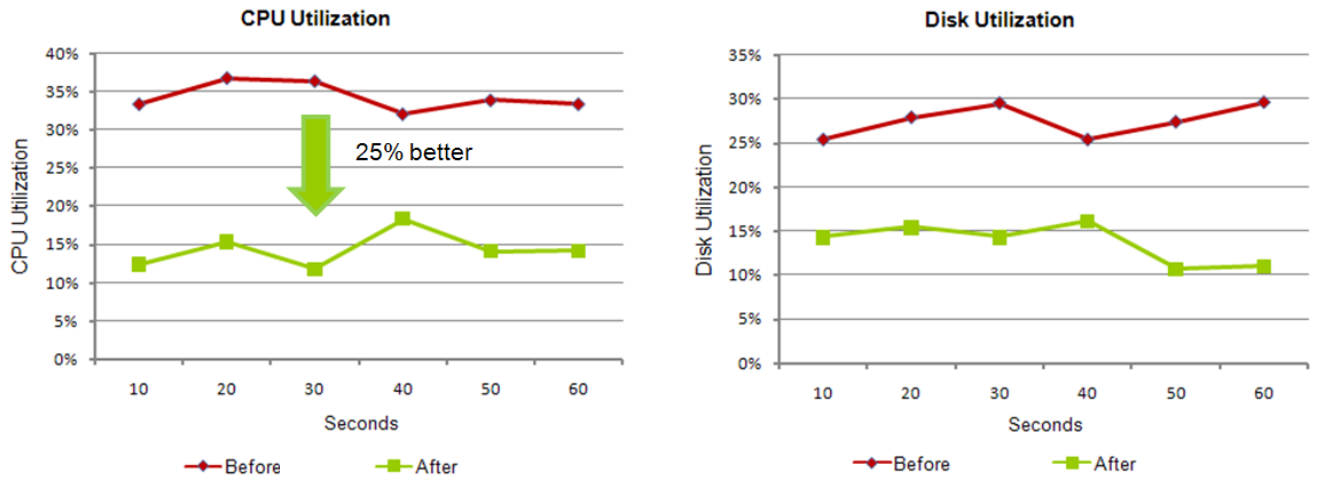
NAS Resource Efficiency

Reducing an application's disk capacity requirements with compression can improve the efficiency of a NAS storage system. With less disk operations to process per file operation, the NAS system has less work to do. With more data being served with each compressed I/O request, the effective efficiency of the cache memory in the NAS system is improved. This is particularly helpful for primary applications and workflows that tend to push the CPU or disk utilization limits of an existing NAS system.

ESG Lab Testing

ESG Lab monitored the CPU and disk utilization of the NetApp FAS3070 during the VMware copy test presented earlier in this report. NetApp statistics were captured every second during the first sixty seconds of the copy job. The average CPU and disk utilization over ten second intervals is shown in Figure 11.

Figure 11. NAS Resource Efficiency



IBM Real-time Compression improved CPU and disk utilization during every time interval. Utilization improvements varied between 10% and 25%. Similar results were seen during Oracle OLTP testing where NAS utilization improvements enabled the system to do more work (15% more transactions) with quicker response times (20% to 90% faster). Much like adding a turbocharger to a car for increased fuel efficiency and more horsepower, IBM Real-time Compression not only increases the performance of a NAS engine, it also improves its efficiency.

Why This Matters

Recent ESG research indicates that the metrics used to evaluate the success of IT's contribution to an organization's green initiative are led by reductions in energy requirements (38%), IT operating expenses (38%), and IT capital costs (29%).[6] IBM Real-time Compression not only reduces energy and capital costs by using less disk capacity, it also increases the efficiency of a NAS infrastructure. A more efficient NAS infrastructure is easier to manage and can be used to forestall or possibly avoid the capital costs of an upgrade.

[6] Source: ESG Research Report, *Global Green IT and Business Initiatives*, August 2008.

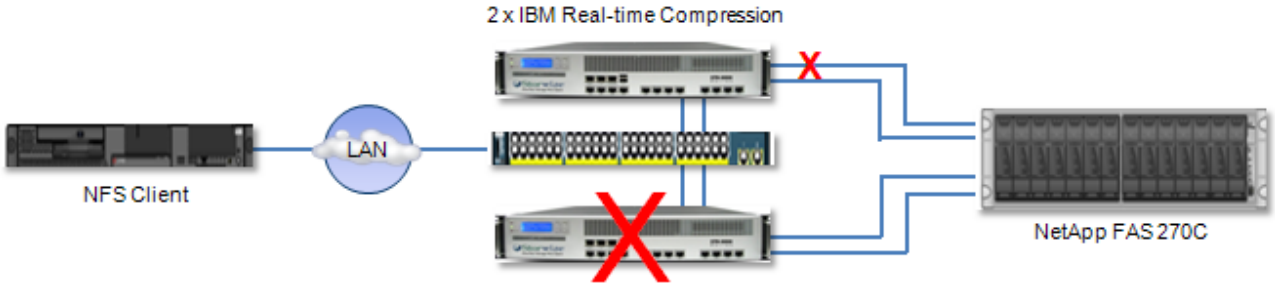
Fault Tolerance

The IBM Real-time Compression STN-6000 data compression appliance integrates into all high-availability topologies and existing configurations seamlessly, without introducing any new single point of failure or degradation of the service level offered by the storage and networking equipment. IBM Real-time Compression also provides active/active and active/passive path failover in the event of a node failure with a two-node cluster.

ESG Lab Testing

ESG Lab tested the recoverability of a two-node STN-6000 configuration in a series of fault tolerance tests. While ESG Lab performed file reads and writes through the two-node STN-6000 appliances, ESG disconnected one Ethernet cable from one of the STN-6000 nodes to trigger a failover to the other Ethernet connection as depicted in Figure 12.

Figure 12. Fault Tolerance Test Bed



The STN-6000 was able to maintain access to the NetApp FAS filer without interruption. Next, ESG Lab powered off a STN-6000 node while data was being written and read from the FAS 270C through the IBM Real-time Compression appliance. IBM Real-time Compression maintained read and write access without disruption.

Lastly, the surviving node was powered off to verify that the host-based IBM Real-time Compression Revert utility can be used to decompress files previously compressed by the STN-6000 appliance. The Revert utility can be used in the unlikely event of total failure of both appliances in cluster. The Revert utility can also be used at a remote site to decompress files that have been mirrored to a WAN-connected NAS disk array.

Why This Matters

Because IBM Real-time Compression appliances are logically configured in-band between clients and the NAS-attached disk capacity they rely on, transparent fault tolerance is a crucial requirement when deploying primary and archive applications which rely on IBM Real-time Compression capacity. ESG Lab has verified that a dual-node IBM Real-time Compression cluster survives a failed appliance transparently with no user intervention.

ESG Lab Validation Highlights

- ☑ ESG Lab confirmed that installing IBM Real-time Compression requires no changes to the network and NAS system(s).
- ☑ Installation and configuration was a quick and easy three step procedure.
- ☑ Compressed files were being stored in real-time three minutes and ten mouse clicks after starting.
- ☑ Dramatic capacity savings for real-world data sets were observed:
  - o 90% for office productivity files (docs, presentations, etc.)
  - o 80% for an order entry Oracle database application
  - o 75% for VMware VMDK file (Windows guest OS)
  - o 91% for VMware VMDK file (Linux guest OS)
- ☑ Using less disk capacity increased performance:
  - o 28% faster copy of a 500 MB file (Windows guest OS, VMware ESX)
  - o 15% more Oracle database transactions
  - o Aggregate Iozone re-read throughput in excess of 600 MB/sec was recorded as an IBM Real-time Compression appliance compressed data in real-time over six Gigabit Ethernet interfaces.
- ☑ Using less disk capacity improved NetApp resource usage:
  - o Up to 25% better CPU utilization
  - o Up to 19% better disk utilization
- ☑ A dual-node IBM Real-time Compression cluster survived a failed appliance transparently with no user intervention. The host-based Revert utility was used to confirm that compressed data can be recovered in the unlikely event of a dual-appliance failure.

Issues to Consider

- ☑ While the ongoing compression of new files is done transparently in real-time, the process of compressing files that have already been written to disk is a one-time task that can take hours to complete and may impact performance. In other words, deploying a new application with IBM Real-time Compression is quick and easy, but migration of an existing application with a large data set already stored on disk requires a bit of planning. To help realize the great savings that can be achieved as capacity is reclaimed from an existing storage deployment, a number of tools and processes have been developed to streamline this one-time compression task, referred to as "back-file compression." The potential performance impact of back-file compression can be eliminated using a combination of scheduling, throttling and policy-based options. For example, a policy and schedule can be used to compress a targeted set of directory and file types during off-peak hours.
- ☑ STN-6000 series appliances compress network-attached file systems which use the CIFS or NFS protocol. Block-based protocols (e.g. FC, iSCSI) are currently not supported. Block-based support would bring the benefits of real-time compression to a larger universe of applications ESG Lab has been advised that block-based support is planned for a future release.
- ☑ While the number of organizations deploying databases on NAS systems is increasing, it should be noted that most database applications are currently deployed on block-based disk arrays, which are not yet supported by IBM Real-time Compression. For those organizations that have deployed database applications on NAS systems to improve manageability and reduce costs, ESG Lab strongly recommends the use of IBM Real-time Compression due to its ability to dramatically reduce capacity requirements as it increases application performance.

ESG Lab's View

Organizations are finding new and innovative ways to leverage file data. Corporate, project and home directory files are being shared on network-attached file systems. Digital archives are being used to meet legal and regulatory demands. Virtual server images are being stored on NFS mounted file systems to increase the mobility and availability of mission-critical applications. Databases, which have long been married to block-based storage technologies, are being deployed on network-attached appliances for simplified ease of management. As organizations learn to leverage each of these processes to increase the value of files shared over a network, IT managers are facing a daunting set of challenges, including increasing costs and the limitations of power, cooling, and space in the data center.

New technologies have emerged in recent years with a goal of addressing the challenges associated with the never-ending growth of networked file capacity. Some have focused on making it easier to manage more file system data from a single interface. Others have focused on reducing the amount of duplicate data that is backed up. Until now, none have focused on solving the problem at the source—as files are created and stored for the first time.

IBM Real-time Compression leverages industry-standard compression algorithms running on industry-standard servers to deliver results that are far from industry-standard. First, IBM Real-time Compression drastically reduces disk capacity requirements. ESG Lab has validated that an IBM Real-time Compression appliance can be used to reduce primary and archive NAS capacity between 75% and 91% for real-world data sets. Second, an IBM Real-time Compression appliance is drop-dead simple to install and deploy within an existing NAS infrastructure. Three minutes and ten mouse clicks after getting started, ESG Lab was storing and accessing compressed files in real-time. Third, the system is extremely fault tolerant and recoverable. And last, but not least, IBM Real-time Compression not only minimizes the potential performance impact of real-time compression, it can actually be used to increase performance for real-world applications.

ESG Lab was pleasantly surprised to learn that VMware-resident copies completed 28% quicker and an Oracle order-entry application performed 15% more transactions with IBM Real-time Compression. The reason became clear when ESG Lab saw that storage system CPU and disk utilization had improved dramatically with IBM Real-time Compression. Reducing disk capacity with real-time compression reduces the number of disk operations that need to be performed. Less disk operations reduces the amount of the work that storage systems have to do. The increased efficiency of each disk operation increases the effectiveness of the storage system's cache. Put it all together and you get the astounding results observed by ESG Lab—adding IBM Real-time Compression to your network-attached storage infrastructure reduces the cost of disk capacity as it increases application performance.

If, like most organizations, you are drowning in a sea of network-attached file data and you relish the idea of reducing primary and archive NAS capacity by 75% or more, ESG Lab recommends that you contact your IBM Real-time Compression representative to schedule a test drive with your files and applications. We're confident that you'll discover what ESG has confirmed via hands-on testing: IBM Real-time Compression minimizes capacity and costs without compromise.

Appendix

Table 3. Configuration Details

Ease of Deployment Test Bed	
Intel x306 Server 3.2 GHz – 1 GB Ram Gigabit Ethernet	Windows 2003 SP-1
IBM Real-time Compression STN-6500	12 x Gigabit ports; 4 GB Ram
NetApp FAS250 OnTap 7.2.4 512 MB Ram	13 x 136GB HDD Gigabit Ethernet
VMware Virtual Server Test Bed	
2 x Physical servers Processor 2GHz Xeon – 8GB Ram Gigabit Ethernet	VMware ESX Server Guest OS (12 Windows; 4 Linux)
IBM Real-time Compression STN-6500	
NetApp FAS3070 OnTap7.2.3.x.x 8GB Ram	Disk Capacity Gigabit Ethernet 56 x 144 GB HDD
Database Test Bed	
9 x IBM x355 Servers 2.4 GHz Xeon – 2 GB Ram Gigabit Ethernet	3 x Oracle servers Oracle 10.2.0.4 Database Size – 2 TB
IBM Real-time Compression STN-6800	12 x Gigabit ports; 32 GB Ram
NetApp FAS6070 OnTap 7.2.3 32 GB Ram	56 x 144 GB HDD Gigabit Ethernet
Cisco Catalyst 3750G Ethernet switch	
Throughput Test Bed	
20 x IBM x335 Servers 2.4 GHz Xeon – 2 GB Ram Gigabit Ethernet	Operating System
Cisco Catalyst 3750G Ethernet switch	
IBM Real-time Compression STN-6800	12 x Gigabit ports; 32 GB Ram
NetApp FAS6070 OnTap 7.2.3 32 GB Ram	56 x 144 GB HDD Gigabit Ethernet
Fault Tolerance Test Bed	
2 x IBM x306 3.2 GHz – 1 GB Ram Gigabit Ethernet	Operating System
2 x IBM Real-time Compression STN-6500	12 Gigabit Ports; 8 GB Ram
NetApp FAS270C OnTap 7.2.5.1 1 GB Ram	7 x 68 GB HDD Gigabit Ethernet

Enterprise Strategy Group | **Getting to the bigger truth.**

20 Asylum Street | Milford, MA 01757 | Tel:508.482.0188 Fax: 508.482.0218 | www.enterprisestrategygroup.com