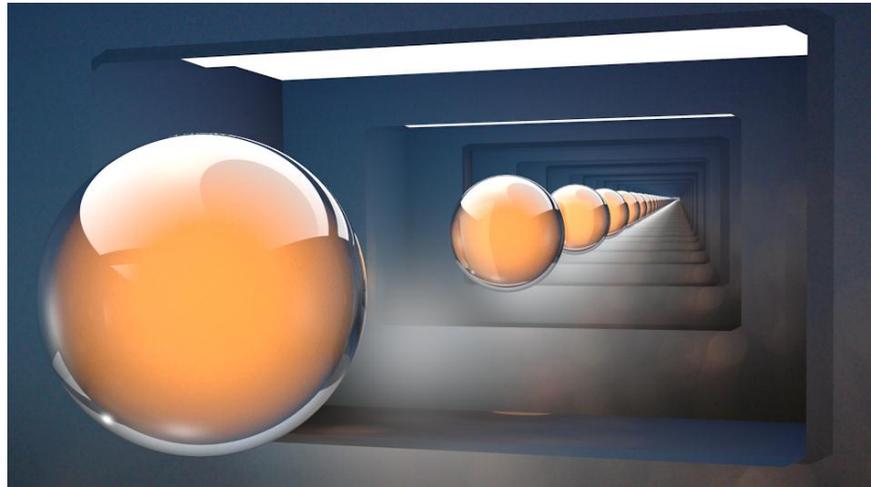


# IMPROVING DEDUPLICATION PERFORMANCE IN PHYSICAL AND VIRTUAL ENVIRONMENTS WITH SYMANTEC BACKUP EXEC 2010 R3

Companies of all sizes are looking to move away from traditional inefficient and costly backup technologies. Such methods back up data indiscriminately, failing to differentiate between new and previously saved



data. This is particularly problematic with remote offices, where the dutiful transfer of all backup data to the central office wastes time, WAN bandwidth and storage resources. By comparison, source deduplication methods back up only changed or new data to the central office, saving time, improving bandwidth utilization, using storage more efficiently and lowering storage costs. Symantec™ Backup Exec™ 2010 R3 includes powerful source deduplication technologies that let companies back up their data efficiently.

Principled Technologies, a company specializing in technology assessment, measured the advantages that Symantec Backup Exec 2010 R3 delivers when using source deduplication to back up databases, mail and other files. The solution drastically improved backup time and data storage needs in both physical and virtual environments, compared to data backup without source deduplication. For instance, in a physical database server environment, we observed an average deduplication ratio of 24.6 to 1. The solution improved the average time required to complete backup by almost 75 percent, or 3 hours. Additionally, in a Microsoft® Hyper-V™ file server virtual machine environment, we observed an average deduplication ratio of 32.3 to 1. The solution improved the average time required to complete backup by almost 50 percent, or 3 hours. Because Symantec Backup Exec 2010 R3 backs up the entire virtual machine (VM) and all associated virtual disks, we found that timing and deduplication ratios are even higher in the Microsoft Hyper-V testing scenarios than in a stand-alone application-server testing scenario.

To read both the virtualized and stand-alone server reports see:

Physical Deduplication Test:

[http://principledtechnologies.com/clients/reports/Symantec/BackupExec2010R3\\_0711.pdf](http://principledtechnologies.com/clients/reports/Symantec/BackupExec2010R3_0711.pdf)

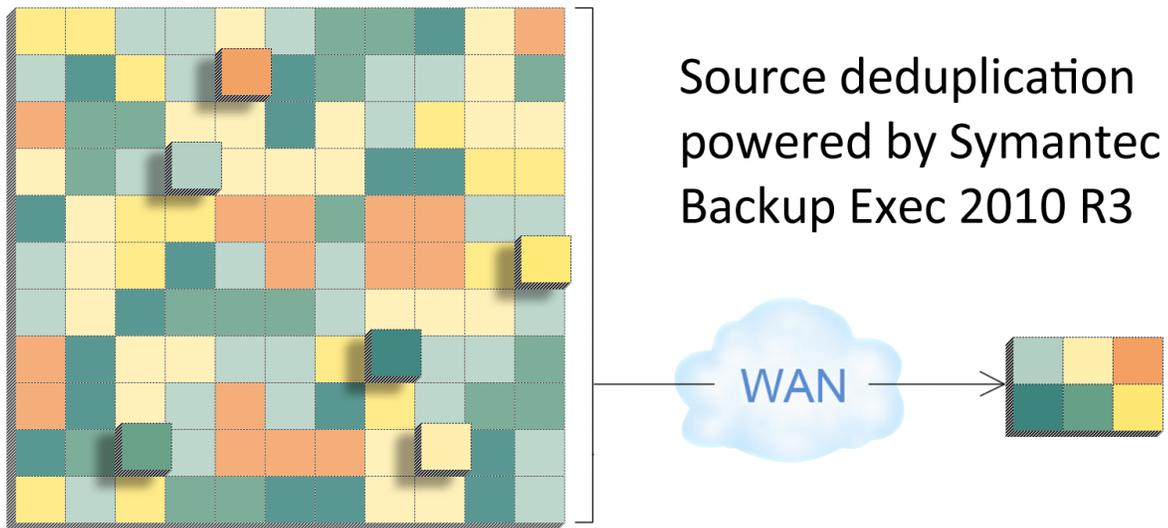
Virtual Deduplication Test:

[http://principledtechnologies.com/clients/reports/Symantec/Virt\\_BackupExec2010R3\\_0711.pdf](http://principledtechnologies.com/clients/reports/Symantec/Virt_BackupExec2010R3_0711.pdf)



## SOURCE DEDUPLICATION EXPLAINED

Source deduplication eliminates backup inefficiencies by transferring only unique blocks of data across the WAN.



Source deduplication offers an efficient way for companies to back up their data. After an initial full backup, deduplication software searches for duplicate data in subsequent backups and then backs up only blocks of data that are new or changed. “Source” means the data is deduplicated at its source, likely a server or desktop at a remote location, before it is passed to the central backup site.

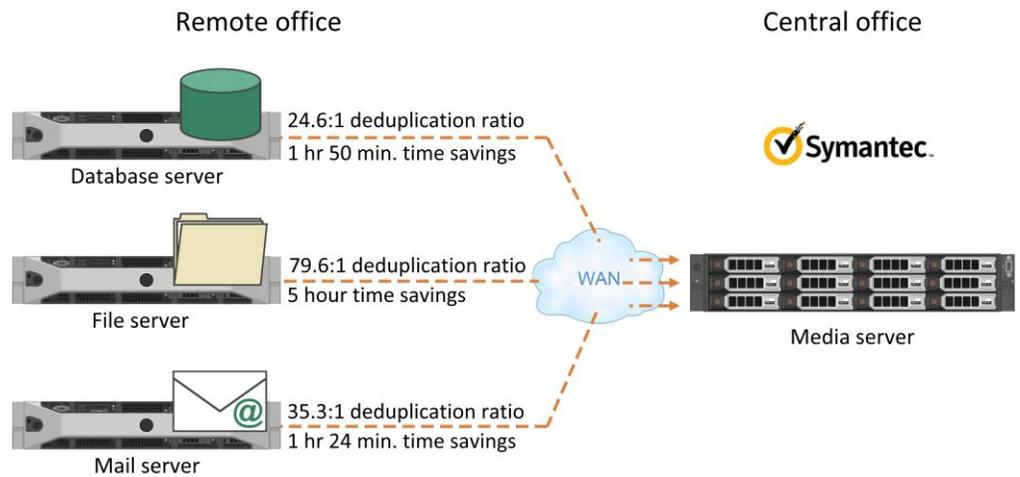
Because the large majority of data typically remains constant from day to day, transferring only new or changed blocks of data from the remote office to the central office with source deduplication is the wise choice when bandwidth, time and space are at a premium.

## DEDUPLICATION BENEFITS IN PHYSICAL ENVIRONMENTS

By deduplicating the data at the source, Symantec’s Backup Exec 2010 R3 solution was far more efficient than disk-based backup without source deduplication, which we also tested. It reduces the amount of data traveling over the WAN by a factor of more than 24 when backing up a database server, a factor of more than 79 when backing up a file server, and a factor of more than 35 when backing up a mail server.

By backing up only new data, Backup Exec 2010 R3 reduced the database backup window from 2 hours to an average of 21 minutes, reduced the file server backup window from almost 6 hours to an average of 47 minutes, and reduced the mail server backup window from almost 4 hours to an average of 2 hours 21 minutes.

**Figure 1: Symantec Backup Exec 2010 R3 significantly reduced the amount of data storage and time needed for data backups in physical environments.**



**The Symantec Backup Exec console made backing up data from a remote location easy.**

The Symantec remote agent compared the data on the remote server to the existing backup data on the media server. When it found a duplicate, it sent only a marker to the media server at the central office. As a result, the remote server transmitted far less data to the media server. This not only saved WAN bandwidth, but also decreased the volume of data the media server had to store—both of which add up to savings. In a straightforward process, we used the Symantec Backup Exec console to simply select the deduplication storage folder and select the option for remote access of the media server while configuring the backup job.

The difference between the volume of data and the volume of data actually stored increases over time, benefiting the business by requiring significantly less long-term storage capacity.

### Database server backup scenario

The initial complete database backup took roughly 1 hour and 35 minutes. Subsequent full database backups using source deduplication took an average of 21 minutes, achieving a deduplication ratio of at least 23.8:1. For comparison, we ran a single backup using a backup-to-disk storage device without source deduplication by Symantec, which took roughly 2 hours and 11 minutes.

Figure 2 shows the comparison of protected data versus actual deduplication data stored for our database server scenario.

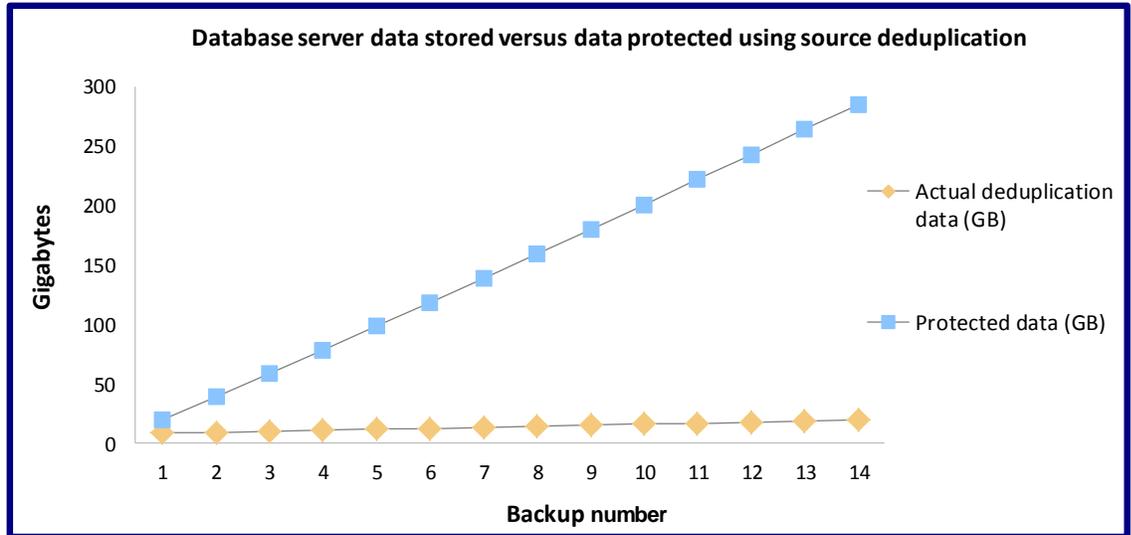


Figure 2: Comparison of protected data versus the actual data stored from our 2-week database backup simulation of a SQL Server database.

### File server backup scenario

The initial complete backup took just over 5 hours. Subsequent full file server backups using source deduplication took an average of just over 47 minutes, achieving a deduplication ratio of at least 24.4:1 up to 117.4:1 and an average of 79.6:1. For comparison, we ran a single backup using a backup-to-disk storage device without source deduplication, which took almost 6 hours.

Figure 3 shows the comparison of protected data versus actual deduplication data stored for our file server scenario.

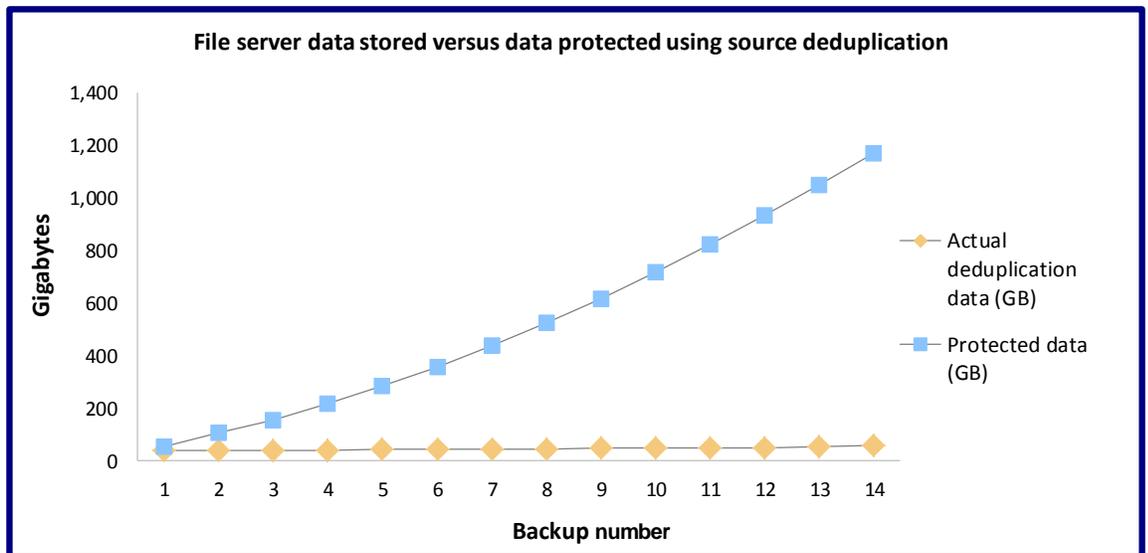


Figure 3: Comparison of protected file server data versus the actual data stored from our 2-week file server backup simulation.

## Mail server backup scenario

The initial complete backup took just over 7 hours. Subsequent full mail-storage-group backups using source deduplication took an average of just over 2 hours, achieving a deduplication ratio of at least 29.1:1. For comparison, we ran a single backup using a backup-to-disk storage device without source deduplication, which took roughly 3 hours and 45 minutes.

Figure 4 shows the comparison of protected data versus actual deduplication data stored for our mail server scenario.

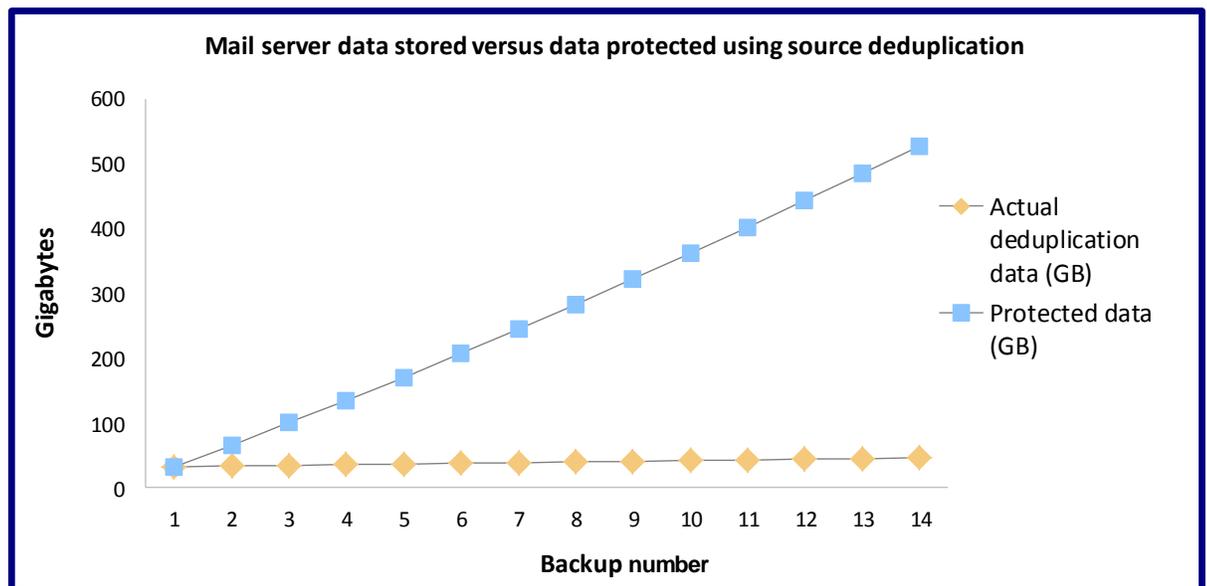


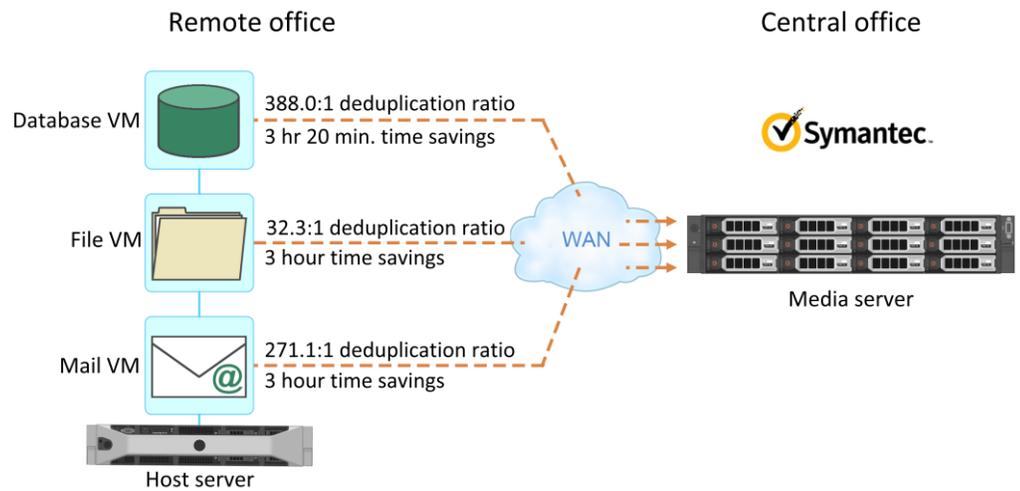
Figure 4: Comparison of protected mail server data versus the actual data stored from our 2-week Exchange 2007 backup simulation.

## DEDUPLICATION BENEFITS IN VIRTUAL ENVIRONMENTS

By deduplicating the data at the source, Symantec's Backup Exec 2010 R3 solution was far more efficient than disk-based backup without source deduplication. It reduced the amount of data traveling over the WAN by a factor of more than 32 when backing up a file server virtual machine and even more when backing up the database- and mail-server virtual machines.

By backing up only new data, Symantec BackupExec 2010 R3 reduced the database backup window from over 6.5 hours to an average of around 3 hours, reduced the file server backup window from over 6 hours to an average of 3 hours and reduced the mail server backup window from over 6 hours to an average of 3 hours.

**Figure 5: Symantec Backup Exec 2010 R3 significantly reduced the amount of data storage and time needed for data backups in physical environments.**



The Symantec Remote Agent compared the data on the remote server to the existing backup data on the media server. When it found a duplicate, it sent only a marker to the media server at the central office. As a result, the remote server transmitted far less data to the media server. This not only saved WAN bandwidth, but also decreased the volume of data the media server had to store—both of which add up to savings. In a straightforward process, we used the Symantec Backup Exec console to simply select the deduplication storage folder and select the option for remote access of the media server while configuring the backup job.

The difference between the volume of data and the volume of data actually stored increases over time, benefiting the business by requiring significantly less long-term storage capacity.

### Database server backup scenario

The initial complete database backup took almost 5 hours. Subsequent full database backups using source deduplication took an average of 3 hours and 13 minutes, achieving a deduplication ratio of at least 179.8:1. For comparison, we ran a single backup using a backup-to-disk storage device without source deduplication by Symantec, which took roughly 6 hours and 41 minutes.

Figure 6 shows the comparison of protected data versus actual deduplication data stored for our virtualized database server scenario.

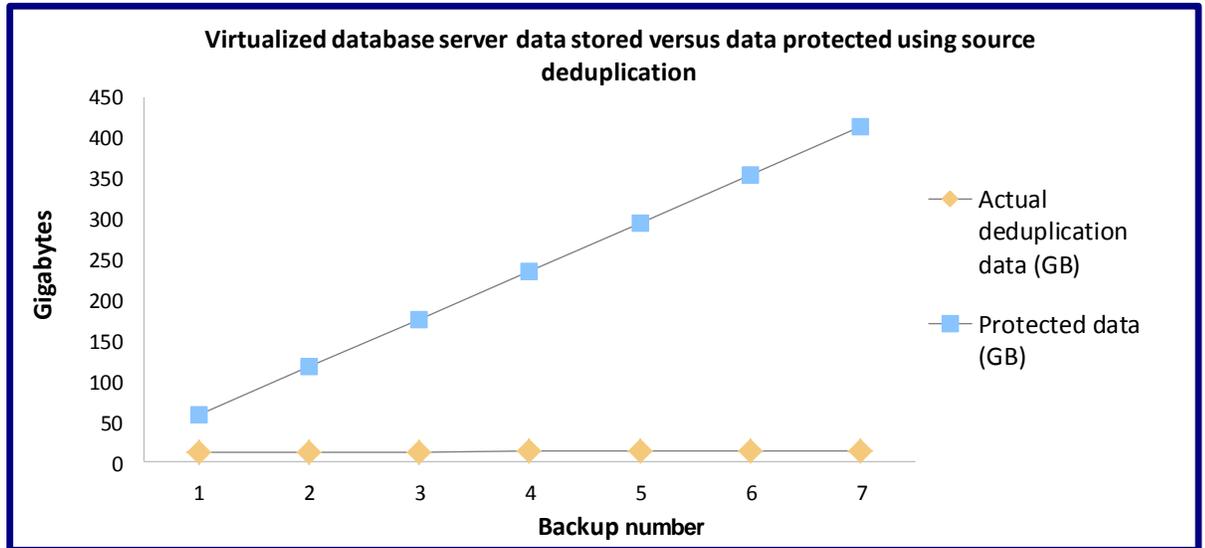


Figure 6: Comparison of protected data versus the actual data stored from our 7-day database backup simulation of a virtualized SQL Server database.

### File server backup scenario

The initial complete backup of our file server took just over 6 hours. Subsequent full file server backups using source deduplication took an average of just over 3 hours, achieving a deduplication ratio of at least 24.4:1. For comparison, we ran a single backup using a backup-to-disk storage device without source deduplication, which took just over 6 hours.

Figure 7 shows the comparison of protected data versus actual deduplication data stored for our virtualized file server scenario.

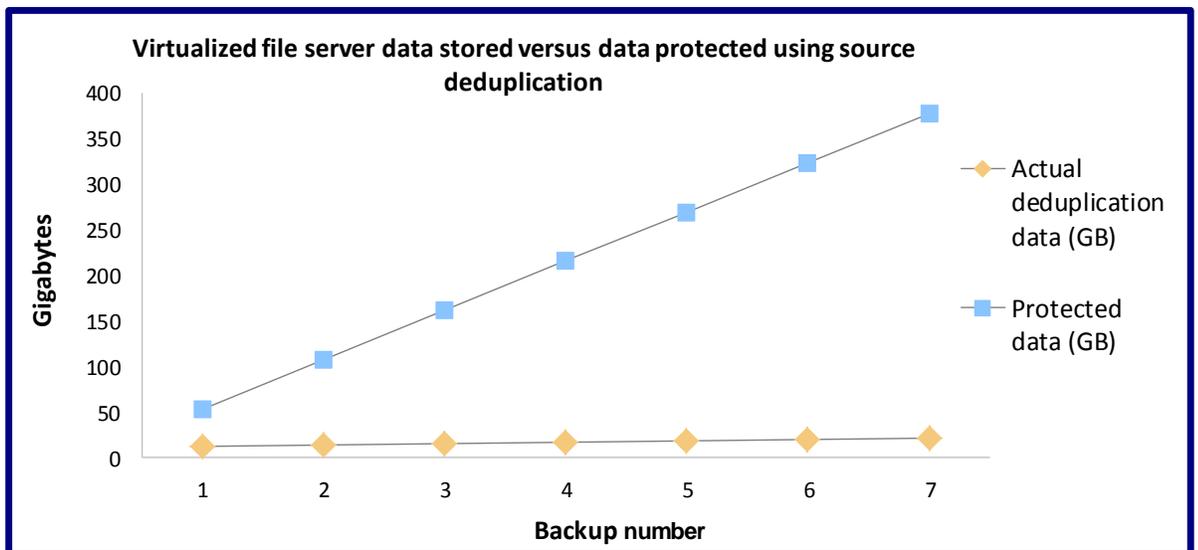


Figure 7: Comparison of protected file server data versus the actual data stored from our 7-day virtualized file server backup simulation.

## Mail server backup scenario

The initial complete backup took just 7 hours. Subsequent full mail-storage-group backups using source deduplication took an average of 3 hours, achieving a deduplication ratio of at least 177.4:1. For comparison, we ran a single backup using a backup-to-disk storage device without source deduplication, which took roughly 6 hours.

Figure 8 shows the comparison of protected data versus actual deduplication data stored for our virtualized mail server scenario.

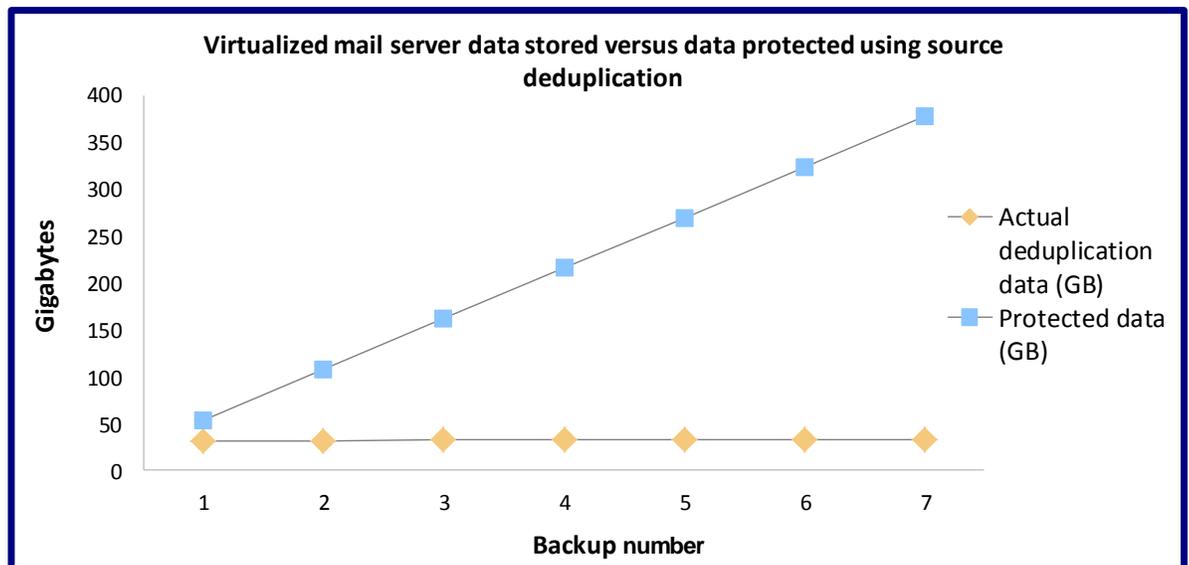


Figure 8: Comparison of protected mail server data versus the actual data stored from our 7-day virtualized Exchange 2007 backup simulation.

## OVERVIEW OF TEST ENVIRONMENT

We tested Symantec Backup Exec 2010 R3 solution on a Dell™ PowerVault™ DL2200 server, backing up several virtual guest machines and several physical servers performing Exchange, SQL Server and File Server roles. For the virtualized test, the hypervisor was a Microsoft 2008 R2 Hyper-V installation running on a Dell PowerEdge R710. This solution offers host-level backup of Hyper-V guests, reducing the impact of backup on individual guest machines. Setup wizards guided us through setup, configuration and deployment. Additionally, we ran with Granular Recovery Technology enabled, which allows users to restore specific files, folders, and even individual emails, without running a full restore.

## SUMMARY

Our testing demonstrated the benefits of source deduplication using Symantec's Backup Exec 2010 R3 solution. We experienced significant improvements in the amount of storage space, bandwidth and time needed in our test cases. We observed these benefits in both physical and virtual environments.

## ABOUT PRINCIPLED TECHNOLOGIES



Principled Technologies, Inc.  
1007 Slater Road, Suite 300  
Durham, NC, 27703  
[www.principledtechnologies.com](http://www.principledtechnologies.com)

We provide industry-leading technology assessment and fact-based marketing services. We bring to every assignment extensive experience with and expertise in all aspects of technology testing and analysis, from researching new technologies, to developing new methodologies, to testing with existing and new tools.

When the assessment is complete, we know how to present the results to a broad range of target audiences. We provide our clients with the materials they need, from market-focused data to use in their own collateral to custom sales aids, such as test reports, performance assessments, and white papers. Every document reflects the results of our trusted independent analysis.

We provide customized services that focus on our clients' individual requirements. Whether the technology involves hardware, software, Web sites, or services, we offer the experience, expertise, and tools to help our clients assess how it will fare against its competition, its performance, its market readiness, and its quality and reliability.

Our founders, Mark L. Van Name and Bill Catchings, have worked together in technology assessment for over 20 years. As journalists, they published over a thousand articles on a wide array of technology subjects. They created and led the Ziff-Davis Benchmark Operation, which developed such industry-standard benchmarks as Ziff Davis Media's Winstone and WebBench. They founded and led eTesting Labs, and after the acquisition of that company by Lionbridge Technologies were the head and CTO of VeriTest.

---

Principled Technologies is a registered trademark of Principled Technologies, Inc.  
All other product names are the trademarks of their respective owners.

---

Disclaimer of Warranties; Limitation of Liability:

PRINCIPLED TECHNOLOGIES, INC. HAS MADE REASONABLE EFFORTS TO ENSURE THE ACCURACY AND VALIDITY OF ITS TESTING, HOWEVER, PRINCIPLED TECHNOLOGIES, INC. SPECIFICALLY DISCLAIMS ANY WARRANTY, EXPRESSED OR IMPLIED, RELATING TO THE TEST RESULTS AND ANALYSIS, THEIR ACCURACY, COMPLETENESS OR QUALITY, INCLUDING ANY IMPLIED WARRANTY OF FITNESS FOR ANY PARTICULAR PURPOSE. ALL PERSONS OR ENTITIES RELYING ON THE RESULTS OF ANY TESTING DO SO AT THEIR OWN RISK, AND AGREE THAT PRINCIPLED TECHNOLOGIES, INC., ITS EMPLOYEES AND ITS SUBCONTRACTORS SHALL HAVE NO LIABILITY WHATSOEVER FROM ANY CLAIM OF LOSS OR DAMAGE ON ACCOUNT OF ANY ALLEGED ERROR OR DEFECT IN ANY TESTING PROCEDURE OR RESULT.

IN NO EVENT SHALL PRINCIPLED TECHNOLOGIES, INC. BE LIABLE FOR INDIRECT, SPECIAL, INCIDENTAL, OR CONSEQUENTIAL DAMAGES IN CONNECTION WITH ITS TESTING, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. IN NO EVENT SHALL PRINCIPLED TECHNOLOGIES, INC.'S LIABILITY, INCLUDING FOR DIRECT DAMAGES, EXCEED THE AMOUNTS PAID IN CONNECTION WITH PRINCIPLED TECHNOLOGIES, INC.'S TESTING. CUSTOMER'S SOLE AND EXCLUSIVE REMEDIES ARE AS SET FORTH HEREIN.

---