

BACKUP-RECOVERY REMEDIATION: Its Value and Criteria for Success

Daymark Solutions, Inc.

Introduction

Ever increasing demands on corporate backup infrastructures have even the most proficient backup managers struggling to define the current state of their backup environment. They worry about:

- Scaling to meet data growth
- Meeting stricter Service Level Agreements
- Complying with emerging regulations
- Recovering from a disaster

With daily backup operations consuming more and more time, these are daunting issues. Keeping the backup and recovery environment in optimal operating condition is a difficult and often thankless job. In most companies, it is simply expected that all data is properly backed up, and available for recall upon demand. It is of little concern to most users how backup and recovery occurs, or just how hard it is to keep all the “moving parts” in proper place and order – they just want their data back when they accidentally delete it, or an auditor calls for it, or they realize that last month’s version of the spreadsheet was the one they really needed.

And yet, the backup window – if it exists at all – is tighter than ever. As data volumes increase, backup times become painfully long, and failures are not uncommon. Schedule conflicts, backup database performance issues, networking connectivity problems, lack of application integration, and recurring bottlenecks all drive up operating costs and, worse, increase the risk of exposed data.

At the same time, complexity of today’s computing infrastructures make it extremely difficult to keep the environment in proper operating condition. Hardware and software manufacturers continuously introduce improvements and options to address customers’ requirements, but measuring the value or risk to a particular environment can become overwhelming. Should you backup to disk? Implement another SAN node or media server? Upgrade versions? Migrate to a new tape technology? Add online Exchange backups? Backup your remote sites locally? Which new options and capabilities will provide your company the best return on investment and least disruption to your users?

The answer to these demands and questions can be derived from a backup-recovery remediation, an incredibly valuable analysis that provides a detailed review of the entire backup-recovery environment. Done correctly, a backup-recovery remediation helps IT quickly identify problems and limitations that exist in the current environment. Armed with this information, the backup team can create an informed resolution plan and implement the necessary changes to bring the backup-recovery environment back to optimal state.

This paper will explore the critical components of an effective backup-recovery remediation. It will also introduce a new approach to backup-recovery remediation that helps companies quickly document and implement the improvements or upgrades necessary to enhance backup performance and reduce operating costs.



Data Collection

The first part of the problem facing a company with backup-recovery challenges is collecting the necessary information to evaluate and understand what is occurring in the backup-recovery environment. Without the ability to collect the necessary data to review the operations and condition of the existing environment, it is nearly impossible to effect substantial improvement.

Getting the Data about Data -Backup Application Databases and Logs

Most enterprise class backup applications can be configured to capture more information about a backup it has attempted than the end user will ever need. However, it can be stored in cryptic log files and proprietary data bases that can make it very difficult to access, never mind analyze when trying to define metrics about the backup environment. Without the addition of layered software reporting modules (whether third-party or vendor-developed) most of this information lays unmined and often unused in the quest to improve backup/recovery readiness. Even with the implementation of layered reporting tools, custom manipulation is often required. Because much of this work falls well outside backup administration and operations, and is not often conducted on a daily basis, many operations find it hard to justify an investment in this crucial component of their backup/recovery infrastructure.

Understanding Backup Application Daemons and Services

A clear understanding of the inner workings of the backup application and how it manages the transfer of backup data is essential when collecting data for analysis. Without this understanding, it can quickly become frivolous to parse thousands of lines of log files in search of problem identification and root cause analysis, and can lead to an inaccurate diagnosis of the environment. One needs to understand how the different daemons/services of a backup application talk to each other and what comprises a backup job for example. The understanding of these interdependencies is not usually required for daily administration and operations, and as such, makes it difficult or cost-prohibitive to build these resources internally. However, if collection/analysis is conducted on a quarterly, biannual or annual basis it can yield significant results in achieving a healthy backup environment. More and more companies are starting to invest in quarterly “health checks” which are provided by outside firms specializing in this service. Often, these companies can quickly develop and implement valuable tools for analysis (scripts, extractions from existing backup-recovery logs, or specially licensed software products) to insure that the proper data is regularly collected, analyzed and “farmed” to provide information and support for ongoing changes.

The Data Collection Process

The success of the data collection phase will hinge on correctly identifying pain points in the backup infrastructure. With the pain points identified, a targeted data collection plan can be created that maps problem areas to an accurate data sample. The data collection activity can then be scheduled and the sample data compiled to a previously agreed upon location. The data is often copied to removable media, and either deleted from or left in the collection location. This allows for either onsite or remote analysis. On average the data collection phase of a backup-recovery remediation should take less than a day of interaction on the actual system interface, keeping any demand on internal client resources or the backup server to a minimum.



Data Analysis/Reporting

Backup Data Parsing – Mining the Data for its Value

Once the backup sample data is collected and compiled, a thorough analysis should be initiated. With a comprehensive understanding of the backup application data management mechanism, a set of tools can be designed to quickly parse the data into a standard format. A flexible parsing process can be used to target key pre-defined metrics or to determine the general health of the backup environment. To further elaborate on this concept, flexible parsing can enable the focus of the analysis to target an already defined problem element of the environment and determine root cause, or to present overall backup metrics that identify characteristics in an environment with numerous backup problems. Most companies today are not skilled in this critical area as it is not a part of their every day operations. With the ability to parse the data correctly, customers can quickly view only the relevant data about their environments and draw the necessary correlations within their environment. Without it, customers can quickly become confused and discouraged while trying to wade through a vast sea of seemingly incomprehensible data. In many cases, hiring an outside firm to provide this very valuable capability will ensure that you can properly leverage the data you have been able to extract from the environment.

Backup Metric/Trend Analysis

Formatted data then requires human analysis to identify backup environment trends and metrics. Metrics such as backup job success rate, duration, volume, and throughput are common elements to review when determining general environmental state. An overlay comparison of duration to volume by client backup, for example, can identify an at-capacity environment versus individual client backup problems. To further elaborate, a client with a large amount of backup data would be expected to have a longer duration, whereas a client with a small amount of backup data should have a shortened duration. When a client with a small amount of backup data exhibits a longer than expected duration, it usually indicates the need for individual client troubleshooting to correct the problem. A client or number of clients with long durations and large amounts of data might reflect the need for a technology refresh or configuration tuning. These trends are compared to the expected metrics of the backup infrastructure and inconsistencies highlighted. With suspect trends highlighted in the initial analysis, it is quite common to have the data re-parsed or re-modeled to take a closer look at a problem element of the backup environment.

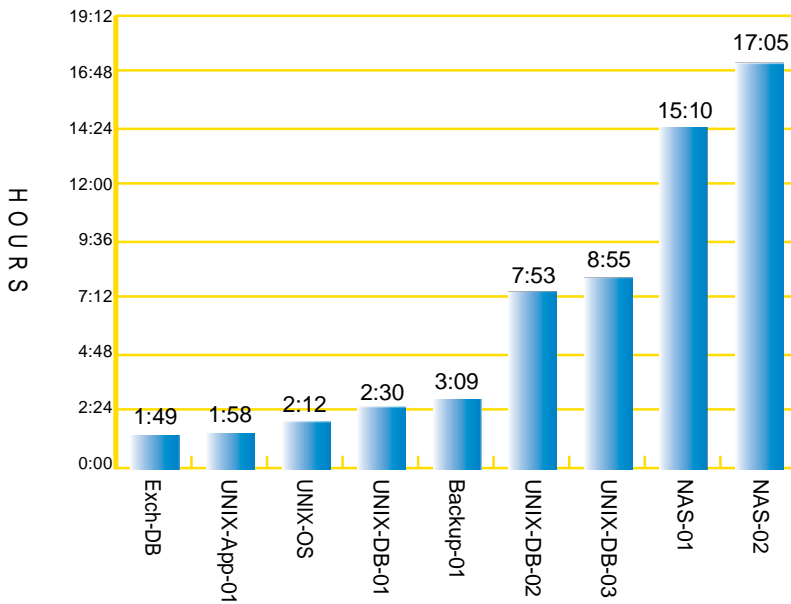
Analysis Report

The final component of the analysis phase is the creation of a findings report with recommendations for improvements to the backup environment. The recommendations are accompanied by supporting data in an easy to understand format. The supporting data can be presented in a number of different formats including; charts, graphs, log and configuration file outputs, as well as Visio diagrams. The visual summary is used to highlight key areas that, when addressed, will be instrumental in creating a more predictable backup infrastructure.

Backup Trend Examples

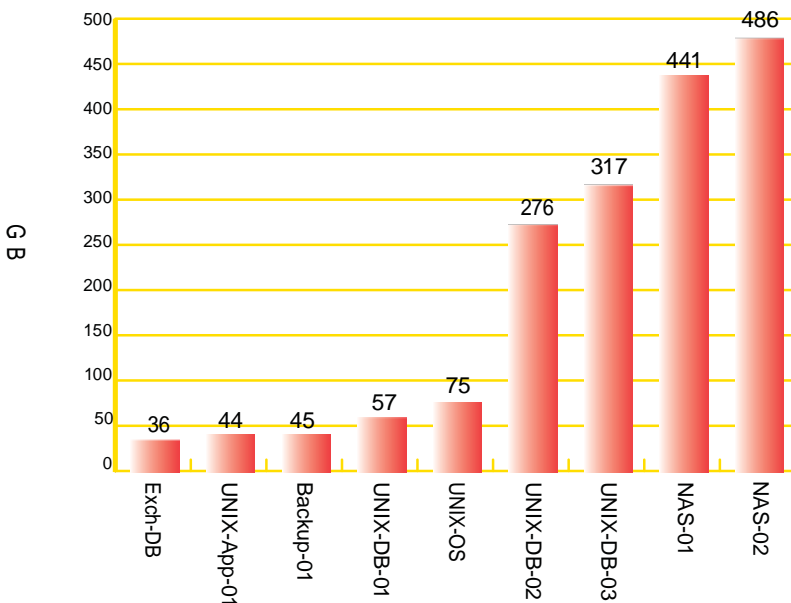
The following charts, descriptions, and summaries are samples of typical outputs contained in an analysis report. The charts contain actual data samples from a compilation of client environments with specific client information removed. The charts contain common data sets encountered when evaluating backup environments.

Backup Duration



This chart is a sample of backup durations and contains data from UNIX, NAS, database, and MS-Exchange clients. It graphs the duration of a backup job by client in hours. The backup jobs in this figure run from 1:49 to 17:05 hours in length. The two heavy hitters are NAS01 and NAS-02 with backup durations of 15:10 and 17:05 respectively. UNIX-DB03 and UNIX-DB-02 are the second heaviest with the remainder of clients well below this grouping. These clients account for the majority of backup load in this environment. The duration must now be compared to the backup volume to further understand the environment.

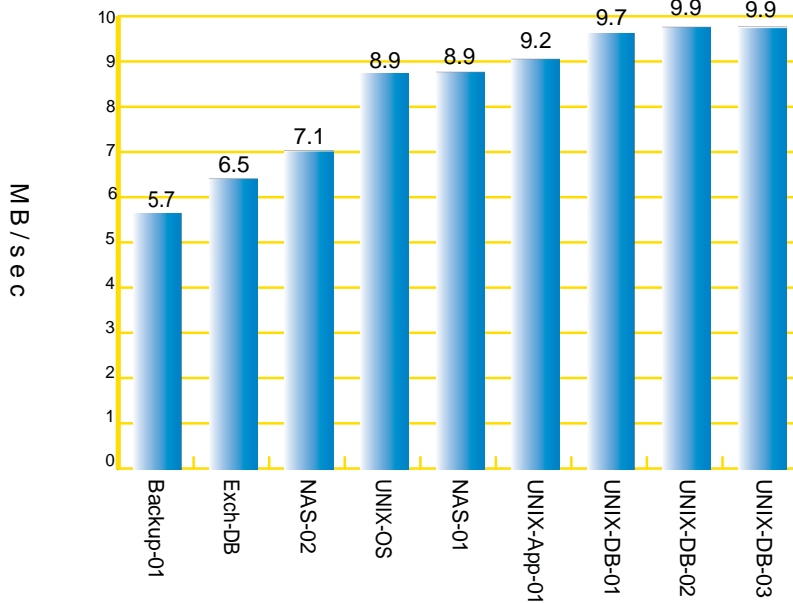
Backup Volume



This is a sample of the backup volume as compared to the backup duration chart above. The volume is represented in GB per client and is the same sample set of clients as the previous chart. The volume ranges from 36 GB on client Exch-DB to 486 GB on client NAS-02. The duration in this environment is aligned to the volume, with the larger volume clients taking more time to backup up than the smaller volume clients. If smaller volume clients had appeared in the two charts with long durations, individual troubleshooting would need to be undertaken on those clients. Because the duration is in alignment with the volume, a closer look at the throughput would be a prudent next step.

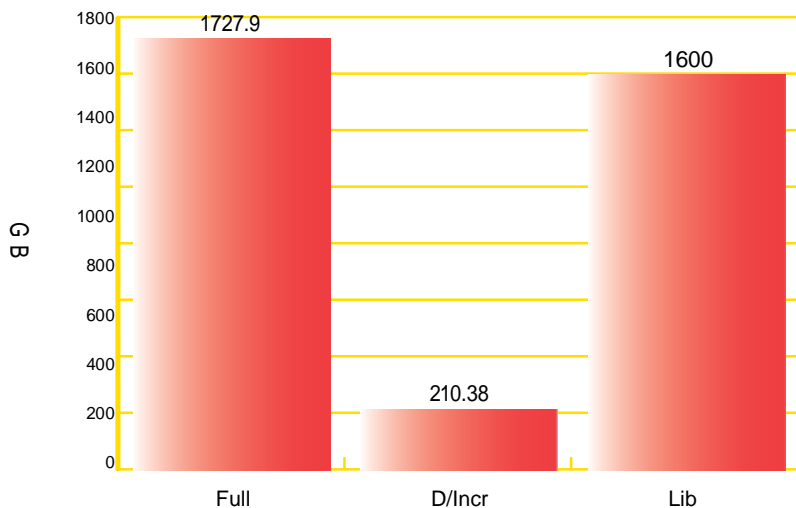


Backup Throughput



This chart illustrates the backup data throughput rates for the same client set used in the previous charts. The throughput is in megabytes per second and ranges from 5.7 to 9.9 MB/sec. The expected throughput rate for a single drive (at native capacity) for the technology deployed in this environment is 5 MB/sec. The throughput displayed is indicative of the deployed technology. Backup-01 using a single local attached tape drive is exceeding the native throughput. The server Exch-DB has a dedicated gigabit backup network to Backup-01 and with the exchange agent deployed is also exceeding the native throughput and benefiting from compression. The remaining servers running from 7.4 to 9.9 are backing up across the network to drive pools containing two tape drives. Though slightly lower than native rates at a per drive throughput analysis, the drive pools enable transfers that would not be attainable with a single tape drive per client backup.

Library Capacity



This chart represents the capacity of this environment's supporting tape library as compared to the backup data load. The library capacity is represented on the right side of the chart in gigabytes, (1600GB or 1.6TB). The average full backup is on the left side at 1728GB or 1.7TB. In the middle is the average daily incremental volume at 210GB. The library capacity in this environment is exceeded with the average weekly full backup cycle. When the Monday through Friday incremental cycle is added to the formula ($5 \times 210\text{GB}$) it comes close to doubling the library capacity. The library in this solution has become extremely undersized due to rapid data growth. This situation introduces excessive manual intervention creating a less reliable backup infrastructure. A properly sized library capable of meeting current and projected data growth requirements would address this highly inefficient and unreliable situation.



Implementation

Implementation Services

With a clearly defined improvement plan in place derived from the analysis/report phase of the project, a level-of-effort plan for implementation can be created. The implementation plan should contain project management as well as deployment effort. If properly mapped against current internal project load, staff skill set, and knowledge of any proposed new technology, the plan will highlight the need for any external services. If external resources are required, the benefits of partnering with a trusted, competent and experienced solutions partner become obvious. To facilitate an on-time, successful backup-recovery remediation project, relying on a partner whose services span from gap augmentation to program management of a turnkey solution greatly improve the chances of success.

Documentation

A key and often overlooked component of the implementation phase is comprehensive documentation. The documentation must be easily referenced and should augment product manuals. Operational tasks specific to the environment or not covered in the product guides should be included, as should design details such as new components or recent alterations to the existing infrastructure. The process to initiate vendor support should also be covered in the document. This becomes even more important in a multi-vendor solution. The ability to easily update this document as enhancements or changes are incorporated is also important as it helps to ensure ongoing ease of operation.

Knowledge Transfer

For a smooth transition from implementation to administration and operations, a comprehensive knowledge transfer is critical. Key operational personnel should be included in implementation phases that cover regularly performed administration or operational tasks. User interfaces and policy settings should be covered in detail. Architecture changes made to the environment should be covered including any new troubleshooting concepts introduced by these changes. A review of the solution document should be conducted at the conclusion of the knowledge transfer phase.



Summary

Keeping your backup-recovery environment in proper operating condition and meeting backup-recovery windows is a tough task, even for the most accomplished backup-recovery administrator. The time demands of staying abreast of new advances in technology, managing the day-to-day demands of expanding storage usage, and dealing with often constrained server, disk, tape or network capacity can effectively prohibit IT professionals from improving their underlying infrastructure. Unfortunately, when it is needed most and would provide the most impact, finding the time or resources to conduct a backup-recovery remediation is often quite difficult. Yet, it is wholly necessary to break the chain of having to repeatedly suffer from less than optimal operation and the accompanying exposure to data loss.

Performing a successful backup-recovery remediation provides a critical review of the backup-recovery environment, recommendations for improvement, and seamless implementation of necessary changes to bring the backup-recovery environment back to optimal state. Below is a list of questions that will help you prepare.

Checklist for a Successful Backup-Recovery Remediation

- Have you developed a plan and schedule with milestones?
- Have you defined responsibilities and gained required access/rights?
- Have you determined what data can be gathered and through what process or tool set?
- Do you have monitoring capability for services/daemons and job initiation and completion?
- What are your Recovery Time Objectives (RTO) and Recovery Point Objectives (RPO)?
- Have you established what data is relevant for your review?
- Have you established a process for mining the data?
- What ability do you have to trend your performance, usage, and growth?
- What metrics have you decided are most relevant for your business and infrastructure?
- Are you familiar with the manufacturer's stated metrics for each component of your backup architecture?
- What types of comparisons, charts, graphs or summarizations are required for presentation?
- How frequently will you remediate your environment to ensure optimal operation?
- What is your tactical plan for implementation of short term and/or critical improvements?
- What is your strategic plan for implementation of long term or significant architectural changes?
- What level of documentation will you require to capture the changes/improvements?
- How often will you schedule additional remediation efforts to keep an optimal infrastructure?



Backup-Recovery Remediation Service

Daymark Solutions Provides Expertise and Resources for Successful Remediation

Clearly, periodic backup-recovery remediation will substantially enhance backup performance, reduce operating costs and ensure the highest levels of data protection. In many companies, however, shrinking backup windows, increasing data volumes and complex technologies make it hard for IT organizations to conduct a successful remediation and implement the recommended improvements.

Daymark’s Backup/Recovery Remediation Service is a comprehensive service that enables companies to quickly identify and document the problems that cause backup failures or painfully long backup times, and implement the improvements or upgrades necessary.

Delivered by certified professionals with extensive hands-on experience in all aspects of backup/recovery, the service includes:

Assess

- Determine backup-recovery success and failure rate
- Understand existing service level agreements
- Evaluate existing hardware/software infrastructure
- Understand disaster recovery dependencies

Recommend

- Provide recommended configuration changes
- Provide recommended process improvements
- Discuss alternative backup strategies
- Discuss potential infrastructure modifications

Implement

- Implement configuration changes
- Install relevant upgrades and additional functionality
- Document process improvements

By leveraging Daymark’s technical experience and proven methodology, companies are assured of reaching key milestones, meeting completion dates and realizing their overall objectives.



About Daymark

Daymark Solutions, Inc. is an experienced technology solutions provider focused on architecting, providing and implementing effective server, storage, network and security solutions for businesses in the Northeastern United States. Our mission is to help you use technology effectively to solve the business challenges you face today, while making sure your systems are agile enough to adapt to future requirements. For more information, call 781-359-3000.

Backup-Recovery Remediation In Action

CUSTOMER PROFILE

Varian Semiconductor Equipment Associates, Inc. is a leading producer of ion implantation equipment used in the manufacture of semiconductors.

CHALLENGE

Consistent data growth expanded the storage environment to 20 terabytes, overburdening the backup infrastructure requiring constant IT attention to maintain performance. Backups that started on Friday nights often extended into production time on Monday and required round-the-clock monitoring throughout the weekend.

AFTER BACKUP-RECOVERY REMEDIATION WITH DAYMARK

- Backup window reduced from 2 days to 13 hours
- Backup success rates commonly approach 100%
- Complexity of the backup environment reduced to improve performance
- Existing licenses redeployed for significant savings
- Overall storage management reduced, freeing IT to focus on other critical tasks