

Enterprise Backup and Recovery Considerations

How to Maintain Your Restore-Readiness

White Paper
Prepared by

Pete Stephens
Senior Technical Consultant
Sirius Computer Solutions



Table of Contents

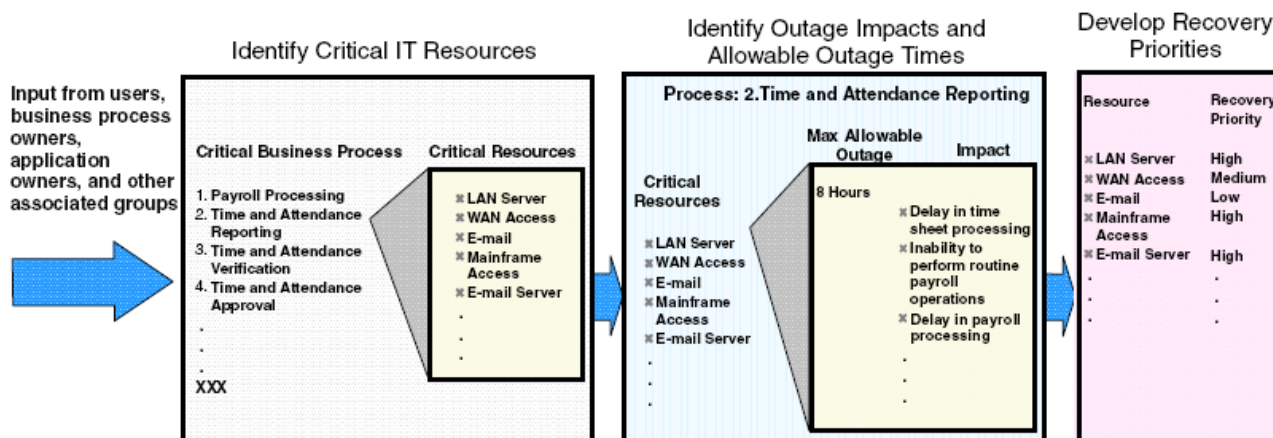
| | |
|--|-----------|
| 1.0 Overview..... | 3 |
| 1.1 Backup/Restore Objectives | 3 |
| 1.2 Maintenance of the Backup/Restore System..... | 4 |
| 1.3 Recovery Time Objectives..... | 5 |
| 1.4 Daily Restores vs. Disaster Recovery | 5 |
| 2.0 Near-Line File and Application Recovery | 7 |
| 2.1 Identify and Prioritize Critical IT Systems and Components | 7 |
| 2.2 Hierarchical Storage | 8 |
| 2.3 Reporting on the Backup Status | 9 |
| 2.3.1 Client Activity Reports | 9 |
| 2.3.2 Backup Server Status Reports | 10 |
| 2.3.3 Backup Media Status Reports | 10 |
| 2.3.4 Disaster Recovery Preparedness Reports | 11 |
| 3.0 Off-Line Disaster Recovery Media | 12 |
| 3.1 Off-site Disaster Recovery Tape Media Rotation | 13 |
| 3.2 Recovering the Backup Server in a DR Scenario | 14 |
| 3.3 Recovering Data and Systems in a DR Scenario | 15 |
| 3.3.1 Operating System and Application Restores | 15 |
| 3.3.2 Data and filesystem Restores..... | 16 |
| 4.0 Summary | 19 |

1.0 Overview

1.1 Backup/Restore Objectives

The IT systems that support most organizations vary widely in size and complexity. However, one characteristic common to most IT systems is that they are highly dynamic. Business processes change, organizations change and external requirements change. Therefore, the plans and processes that are in place to ensure the continuity and recoverability of the organization’s IT systems must also be continually reviewed, maintained and tested.

A Business Impact Analysis (BIA) allows IT management to understand and document the financial and non-financial value of a business process and its supporting application environments. This information can be used to classify systems and create policies and strategic plans to support the business requirements for availability and reliability. The BIA is a critical element for balancing the cost of supporting infrastructure versus the cost of downtime. Development of the BIA, illustrated below, is beyond the scope of this paper, but it is an important starting point for understanding the system requirements, processes, and interdependencies. It is important that the DRP team have access to this information for subsequent planning and policy-creation phases.



Above: Business Impact Analysis Process

IT systems are vulnerable to a variety of disruptions, ranging from mild (e.g., short-term power outage, disk drive failure) to severe (e.g., equipment

destruction, fire) from a variety of sources such as natural disasters, employee errors, and internal or external sabotage. While much of this vulnerability may be minimized or eliminated through technical, management, or operational solutions as part of the organization's risk management effort, it is virtually impossible to completely eliminate all risks. These risks drive the need for reliable and efficient backup / restore systems. This paper will focus on common processes and procedures that should be part of any effective enterprise backup/restore system.

1.2 Maintenance of the Backup/Restore System

Just like there are several modes of system failures as discussed above, there are also several different platforms or types of applications that require backup/restore services.

The National Institute of Standards and Technologies has published a "Contingency Planning Guide for Information Technology Systems" (NIST Special Publication 800-34 available at:

<http://csrc.nist.gov/publications/nistpubs/800-34/sp800-34.pdf>) in which seven IT platform or application types are defined. These seven platforms are:

1. Desktops and portable systems
2. Servers
3. Web sites
4. Local area networks
5. Wide area networks
6. Distributed systems
7. Mainframe systems

The full scope of Contingency Planning is beyond the scope of this paper. However, the final two points on the NIST's recommended seven-step contingency process are particularly important for maintenance of the DR Plan. NIST Special Publication 800-34 defines the following planning and maintenance steps for a good DR Contingency Plan:

6. Plan testing, training, and exercises. Testing the plan identifies planning gaps, whereas training prepares recovery personnel for plan activation; both activities improve plan effectiveness and overall agency preparedness.

7. Plan maintenance. The plan should be a living document that is updated regularly to remain current with system enhancements.

1.3 Recovery Time Objectives

The Recovery Time Objective (RTO) is the amount of time within which a system or application should be restored. The RTO will be based on the business value of the system and information, and the type of platform or application that is being backed-up. The seven types of application platforms listed above have different characteristics (amount of data, accessibility of the data, security, etc.) which will make the backup and restore process take more or less time. Of course, there are products and processes available which can speed up any backup/restore process.

1.4 Daily Restores vs. Disaster Recovery

Backups and Restores have two distinct objectives which should be considered independently:

1. During the course of normal operations, files and systems and databases will need to be restored for any number of reasons. These files and systems will be backed-up and stored such that the data can be accessed quickly. This type of storage is often called “near-line” backup storage.
2. The Disaster Recovery Plan (DRP), as suggested by its name, applies to major—usually catastrophic—events that deny access to the normal facility or site for an extended period of time. This data will be backed-up and stored such that the data can be accessed and restored during a major outage, usually involving many systems and often from another location. This type of storage is often called “off-line” backup storage.

This paper considers these two backup/recovery scenarios and suggests maintenance and processes that should be reviewed periodically and tested to support the readiness and reliability of backed-up data if and when a recovery is needed.

1. Near-line file-level and application-level backup/restores
2. Off-line disaster recovery media management

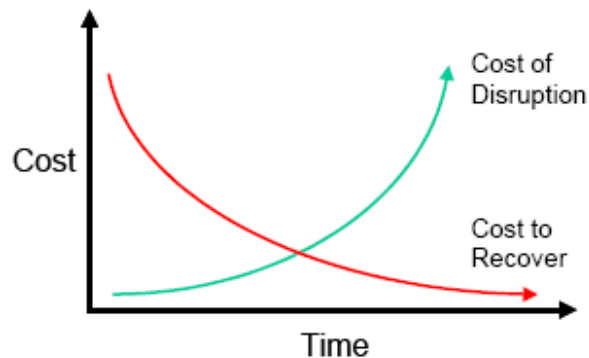
2.0 Near-Line File and Application Recovery

The majority of data restores is usually limited to a small set of related servers, and the amount of data is usually small relative to the volume of data backed-up. In this case, the speed and accuracy of recovery is usually very important.

To be able to recover quickly a file or a set of files, the data should be stored on near-line storage. This means that evaluation of the cost of storing the data vs. the likelihood of recovering the data should be performed. The basis for this kind of evaluation is a set of defined application/system tiers that prioritize critical IT systems and components.

2.1 Identify and Prioritize Critical IT Systems and Components

There are several methodologies that are designed to identify and prioritize data and systems so that the business' most important assets can be identified and recovered first. This paper will not discuss the pros and cons of these methodologies. But it is vitally important that some method be used to identify related information and assign some point on the Cost vs. Time scale illustrated below.



This classification of data and systems should be reviewed at least every six months, since new applications are constantly being added and other applications and systems move through their life cycle and become obsolete.

The classification of data and applications can then be used to define the backup storage hierarchy. The storage hierarchy allows different storage resources to support different classifications of data and systems.

2.2 Hierarchical Storage



From its inception, Tivoli Storage Manager (TSM) has been built around a storage hierarchy. TSM has used the storage triangle depicted at left to represent its storage hierarchy since Version 3 came out in the late 1990s. In recent times the cost differences of various storage media types like tape, serial disk and virtual tape have narrowed. But the price-performance differences in the different media types still exist and should be considered when deciding where and how long to store various classifications of information assets.

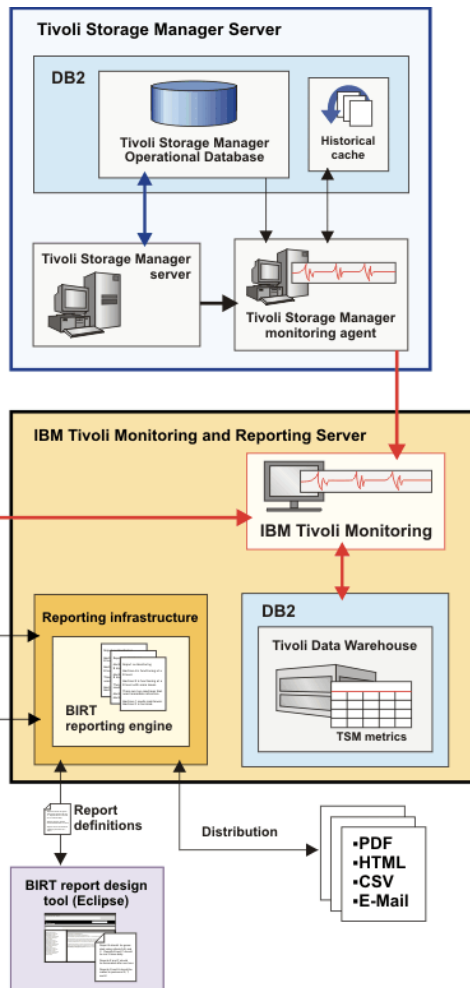
You should organize your backup server's storage resources into one or more hierarchical structures. This storage hierarchy allows flexibility in a number of ways; for example, you can set policy to have clients send their backup data to disks for faster backup operations, and then later have the server automatically migrate the data to tape.

Just like the data it contains, the Storage Hierarchy of the backup server should be dynamic enough to allow adjustments as new data or new types of media are introduced. TSM's storage hierarchy can be changed and data migrated at any time to support the changing needs of the business.

2.3 Reporting on the Backup Status

The ability to define and distribute reports on the status of the enterprise backup status is important to maintaining an efficient backup/restore infrastructure. At a minimum, the reports should cover the following four key areas:

1. A daily status of the backup activity of the clients should be distributed to the stakeholders of each application or system group.
2. A summary of the status of the backup server and the steps and resources needed to recover the backup server should be sent to a secure site.
3. The status of the off-site DR media (to vault, return from vault, inventory in vault, etc.) should be distributed to the operations group responsible for off-site media management.
4. The current status of the backup server and the resource utilization trends (capacity, performance, etc.) should be sent to the backup systems administrator.



2.3.1 Client Activity Reports

Client activity reports should include information about your client activity as well as schedule status, filespace information, backup, and other detailed activity history for your backed-up servers.

These reports are generated by the backup server and should be available in HTML, .pdf, PostScript®, and Microsoft Excel format to be distributed to interested system stakeholders.

2.3.2 Backup Server Status Reports

Server status and trend reports include current information about the utilization of your backup server and historical information about your backup server's trends, including server throughput, resource usage, database details, and tape usage and analysis. These types of reports are essential for performance monitoring and capacity planning for current and future backup requirements. These reports should be available in a variety of formats to the backup systems administrators.

2.3.3 Backup Media Status Reports

The media on which your backed-up data is stored is constantly changing as new backups are created and expired and as data moves through the storage hierarchy. The backup storage media will move through the following states as data is backed-up, stored on near-line media, copied for off-line storage, expired, and brought back on-site. The media states are generally defined as follows.

1. Mountable – Media that is on-site and on-line and available for use.
2. Courier – Media that is being moved to another location.
3. Vault – Media that is stored off-line for DR purposes.
4. Vault Retrieve – Media that has expired data or reclaimed data and is ready to be brought back on-site.
5. Courier Retrieve – Media that has been checked out of the Vault and is being returned.
6. On-site Retrieve – Media that has been returned to the primary location and is ready to be brought back online.

Not all media will move between each of these states, so it is important that the inventory of media at each state is reported and available to the storage administrator who is responsible for storage media. The Backup Media Status reports should be created at least daily and sent to the storage administrator(s) so that he/she can track the movement and inventory of storage media in each state.

2.3.4 Disaster Recovery Preparedness Reports

The Disaster Recovery Reports should be designed to provide configuration and status information that would be needed to recover the Backup Server itself in the event of a disaster scenario (i.e. latest backup tape, steps needed to recover, locations of the backup media). This Disaster Recovery Report should be sent off-site at least daily to a secure location so that it will be available in the event of a disaster.

3.0 Off-Line Disaster Recovery Media

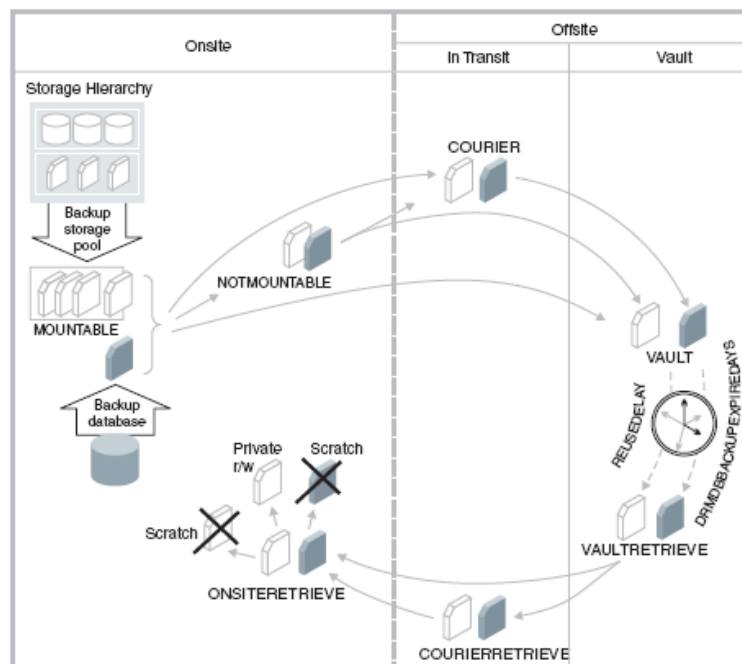
Data that is backed-up and stored off-site has distinctly different storage requirements from backup data stored on-site or “near-line.” Off-site or off-line data is stored with the requirement that the original data and systems could be restored entirely from the off-line media if a disaster or loss of access to the original site should occur. To meet these different (and frequently opposing) storage hierarchy requirements, the off-line data is frequently a copy of the near-line data and is stored in a different storage hierarchy. To support this important requirement, off-line data should be checked periodically to see that it meets the following requirements:

1. The data should be moved to the off-line location within the Recovery Point Objective that has been defined (usually at least daily).
2. The Backup/Restore server at the off-site location will have to be restored first before any data can be restored for the backed-up systems. Therefore the instructions and latest backup media information should be available and accessible from the off-site location.
3. The data at the off-site location will usually be restored in complete system or filesystem increments. Therefore, the process of restoring complete systems or filesystems to the storage systems should be available and accessible from the off-site location.

3.1 Off-site Disaster Recovery Tape Media Rotation

Data backed-up by the backup server should be moved to off-site media quickly and tracked through the data lifecycle as illustrated below. The off-site media is usually tape, but for high-value data it could be another storage subsystem using “Flash Copy” or some other type of disk transfer technology.

For Tivoli Storage Manager, the off-site storage media is managed by TSM’s DR Manager. The backup server’s Disaster Recover Manager function should move DR media through the following life cycle:

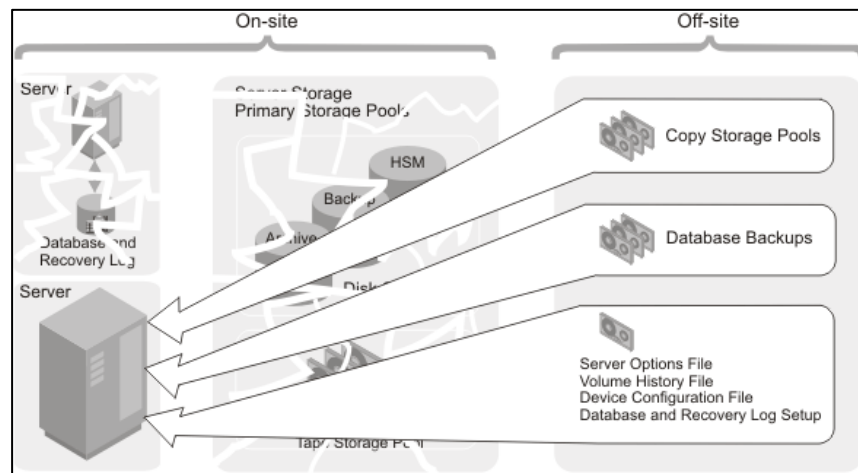


3.2 Recovering the Backup Server in a DR Scenario

Recovering the backup server in a DR scenario will require information about the backup server's latest backup media and the configuration files and procedures needed to restore the server. This information should be sent to the DR site or to a secure site on a daily schedule. A complete backup recovery plan would contain the following information:

- The backup server's recovery procedure.
- A list of required database volumes, copy storage pool disaster recovery volumes, and active-data pool volumes should be listed. In addition, the devices (tape drives, disk drives, virtual libraries or other devices) needed to read those volumes should be documented daily.
- The backup server's database and recovery log space requirements.
- Copies of the server options file, device configuration file, and volume history information file.
- Commands for performing database recovery and primary storage pool recovery.
- Commands for registering licenses.
- Instructions that you define for your specific site.
- Machine and recovery media information that you define.

The TSM DB backup media, the TSM Copy Storage Pool media, and copies of the TSM Server Device Configuration File, Volume History File and Options File are stored in the TSM "Prepare File" when the "Prepare" command is run. The Prepare command should be run daily and a copy of the Prepare File should be automatically sent off-site.



Components and information required to restore the backup server – when all the on-site resources are all down

3.3 Recovering Data and Systems in a DR Scenario

When restoring systems in a DR scenario, the recovery of the operating system and application executables should be considered separately from the recovery of the business data files and data bases. The operating system and application configuration are usually not very dynamic (it is installed, then configured, and not changed often). However, recreating a specific system and application configuration could be difficult. In contrast, application data is very dynamic, but is usually fairly easy to restore, given that the correct components were backed-up (data and its logs for example).

3.3.1 Operating System and Application Restores

Depending on the platform used, there are a variety of choices which will support consistent operating system and application executable backups and restores. These consistent operating system backup and restore tools are often called “Bare Metal Restores” since the objective is to restore a complete and functioning system on a different machine. These Bare Metal Restore tools have strengths and weaknesses which allow them to accomplish this goal in different ways. A careful consideration of your environment should be made to match the features / functions of these tools to your needs. A sample list of Bare Metal restore tools that work with IBM® Tivoli® Storage Manager (TSM) is listed below:

| Software | Platform(s) | Highlights |
|-----------------------------------|-------------------------------------|---|
| ASR “Automated System Recovery” | Windows XP, 2003 | Windows and TSM support “out of the box” Not very automated. Some manual configuration |
| AIX® Sysback | AIX | IBM product for boot/install of AIX |
| AIX NIM “Network Install Manager” | AIX | Comes with AIX |
| Cristie Bare Metal Restore | Windows®, Linux, Solaris™ and HP-UX | Handles loading device drivers while restoring |
| Ultrabac | Windows, Linux, Solaris | |
| Storix SBAdmin | Linux, AIX & Solaris | |
| VMware® VCB | ESX and supported Guest OSs | Backs-up the Virtual Servers |

3.3.2 Data and filesystem Restores

When data is backed-up and restored it is very important that the full data “consistency group” is backed-up together. The “consistency group” is defined as the minimum amount of data that will return an application or data store to a specific point in time. For a database this might be a set of table spaces and the associated logs, or it might be the whole database if the backup was taken off-line. For data files this might be an individual file, or just a portion of the data blocks that make up that file, or it may be an entire filesystem.

Consistency Groups should be set up and backed-up using methods that are consistent with the type of data being backed-up. All objects within the consistency group should have the same data storage characteristics and the same data retention characteristics. With a backup solution like TSM, this is accomplished by assigning the backed-up objects to the same management class. The following sections provide a summary of the processes and procedures used to back-up various types of data.

3.3.2.1 Individual file backup/restores

Files and subfiles should be backed-up with related files in the same backup operation. To save backup time and space, only files or subfiles that have changed since the last backup need to be backed-up. This is called an incremental backup. Full incremental, partial incremental, incremental-by-date, and journal-based backups all back-up new and changed files. All related files that are backed-up from a server or group of servers should be assigned to the same backup group or management class, so that they will be migrated through the storage hierarchy and expired using the same criteria.

Open file support for backup operations can also be useful for situations where a file is held open. Care should be taken that the file does not change during the backup so that the file-level consistency group is maintained. There are two snapshot providers that can be used for open file support: LVSA and VSS (VSS is not supported on Windows XP). VSS is the recommended solution, since it utilizes Microsoft’s strategic snapshot solution.

3.3.2.2 Data Stores

Data stores are groups of files that are managed by an application as a single entity. Databases and e-mail systems are examples of data stores. Often data stores will have a backup/restore method or utility built-in, which will guarantee consistency across the data store during backup and restore operations. Using the native backup methods or utilities for each data store is the preferred way to ensure consistency and integrity of the backed-up and restored data.

3.3.2.3 Filesystems

Image backups

A logical volume or a filesystem can be backed-up as a single object (image backup) on your system. The image backup has the advantage of only needing a single restore operation to put all the data back to the specific point in time when the image was made. In a Disaster Recovery scenario this is useful. The traditional static image backup prevents write access to the volume by other system applications during the operation. These volumes can be formatted FAT, FAT32, NTFS, or unformatted RAW volumes.

If a volume is NTFS-formatted, only those blocks used by the file system, or smaller than the `imagegapsize` parameter, will be backed-up. Normally you cannot restore an image backup of the system drive over itself since an exclusive lock of the system drive is not possible. However, when using WinPE, an image restore of the system drive is possible.

Network Filers

Network filers are supporting increasing amounts of centralized data. Tivoli Storage Manager (TSM) has been a major solution in Enterprise backup/restore solutions for a long time. New features and functions in TSM V6.1 are increasing the flexibility, functionality and performance of filer backup and restores.

In a Disaster Recovery scenario, when whole systems fail and need to be recovered, the ability to recover large amounts of data quickly at a volume or system level is very important.

In this type of DR scenario, the objective is to restore an entire filesystem to a known point in time. This restore can be back onto the same system, or onto an entirely new system potentially in a new location. Streaming data speeds are more important than random access to individual files in this case. Tape is still a very viable and cost-effective medium for the long-term storage of this type of backup.

TSM and NetApp® features can be combined to meet this cost/performance scenario.

4.0 Summary

The planning, execution and resources needed to restore a full system in a DR scenario is very different from the resources needed to restore a subset of files to a specific point in time because of corruption or accidental deletion.

Planning for and monitoring the performance of backups against the defined plan can help ensure the quickest possible restores in both of these scenarios.

This paper has highlighted several considerations for maintaining backup/restore readiness for both file-level restores and disaster-recovery scenarios. In both cases, planning, documentation and monitoring/reporting on the backup status are the most important steps. Tivoli Storage Manager V6.1 has several monitoring, reporting and disaster-recovery features that are provided as examples in this paper as a way to keep your backup/restore systems ready to restore your data and critical systems.

One additional step that can and should be taken for any enterprise backup/restore system is to periodically run practice exercises of the most common kind of restore scenarios. By running through the restore procedures, your people, procedures and systems become more familiar to each other, so they can perform best when you need them most... during restores.

About the author:

Pete Stephen has over 20 years of IT experience in a variety of design, development, implementation, management and consulting roles. Pete is an IBM Certified Infrastructure Systems Architect and has earned The Open Group Architecture Framework (TOGAF) Certification. He has strong IT architecture skills in: IT systems optimization and server consolidation, analyzing and defining IT standards and processes, and systems migration planning and execution. He currently performs delivery consulting for Power Systems/System p and IBM Middleware design and implementation, with AIX,TPC, DB2/UDB, TSM, HACMP, SAN and IBM DS4000, DS6000 and DS8000 storage.

About Sirius:

Sirius Computer Solutions is an IBM Premier Business Partner providing advanced infrastructure solutions to clients across the U.S. Sirius is a nationally recognized solution provider with a certified team of sales and technical professionals who have the skills, product knowledge and commitment to help clients develop and implement the right solutions to solve their business needs. With a nationwide consulting, sales and services organization, Sirius provides best-of-breed technologies from across the full spectrum of information technology, including hardware, software, storage, networking, security and voice. Sirius backs it up with post-sale support that ensures you get the maximum benefit and value from your investment. More information can be found at www.siriuscom.com or by calling 800-460-1237.

June 26, 2009

© Copyright Sirius Computer Solutions 2009. The IBM logo is a registered trademark, and the IBM Premier Business Partner emblem is a trademark, of International Business Machines Corporation, and are used together under license. IBM, AIX and Tivoli are registered trademarks of International Business Machines Corporation. All other company, service and product names are trademarks or registered trademarks of their respective companies.