



Best practices for data migration.

Methodologies for planning, designing, migrating and validating data migration

Contents

2 Executive summary
4 Introduction
6 Data migration methodology
12 Choosing data migration software
15 Summary

Executive summary

Minimizing the effect of mainframe, UNIX®, Linux® and Microsoft® Windows® data migrations on business operations is a critical part of the data center manager’s job. In the past, migration projects were scheduled to be performed during off-hours; however, today’s 24x7 environment has no application downtime window. Rather than delay the purchase or deployment of new storage technology because of potential migration problems, IT organizations now can employ new tools and methodologies that minimize the risk. New software allows nondisruptive migration, meaning that applications remain online during data movement without significant performance delays. And new methodologies that help with planning, migrating and validating data migrations enable organizations to take advantage of customization, speed the process and ensure data integrity.

To promote an understanding of what it takes to ensure nondisruptive migration – and how Softek® Transparent Data Migration Facility (TDMF™) software can help meet this need – this paper identifies five key factors to consider in choosing data migration software.

Highlights

There are many challenges to consider when migrating data—and Softek TDMF software can help handle them.

Key factors	Description	Softek TDMF capability
Performance	How quickly can data be copied from the source to the target and balanced against system overhead?	TDMF technology includes a throttling/pacing capability that can speed up or slow down data movement depending on system demands in order to easily balance migration with system and network requirements.
Primary volume/ source data protection	If something goes wrong, the migration can be terminated and application processing restarted or continued on the source data/device.	Protection can be easily done with the TDMF solution through backup (IBM z/OS® platform) or fallback (UNIX platform) commands.
Tiered storage	Moving data to different array or to different storage media to improve price/performance without disruption.	TDMF technology allows the migration of data between different disk storage media (e.g., DASD or S/ATA) for optimum price/performance.
Multivendor environments	Many data centers use hardware from several vendors, which can result in source and target hardware being from different vendors.	TDMF technology is completely hardware independent, supporting nondisruptive migration in multivendor environments.
Application downtime	Applications have different levels of business criticality and therefore have varying degrees of acceptable downtime.	TDMF technology allows for nondisruptive data migration so applications can stay online and continue to process data throughout the migration process.

Highlights

In a Softek survey, 60 percent of respondents said that they migrate data at least every quarter.

Introduction

Data migration is the process of making an exact copy of an organization's current data from one device to another device – preferably without disrupting or disabling active applications – and then redirecting all input/output (I/O) activity to the new device. There are a variety of circumstances that might cause an organization to undertake a data migration, including:

- *Server or storage technology replacement or upgrade*
- *Server or storage consolidation*
- *Relocation of the data center*
- *Server or storage equipment maintenance, including workload balancing or other performance-related maintenance.*

The above scenarios are fairly routine parts of IT operations in organizations of virtually any size. They are so routine, in fact, that more than 60 percent of respondents to a recent survey¹ indicated that they migrate data quarterly or more often – with 19 percent migrating weekly. However, even routine processes can cause problems for IT administrators and managers. More than 75 percent of respondents to the same survey said they had experienced problems during data migration. These problems included, but were not limited to:

- *Extended or unexpected downtime*
- *Data corruption, missing data or data loss*
- *Application performance issues*
- *Technical compatibility issues.*

Highlights

When systems must be taken down for migration, business operations can be seriously affected.

A key way to minimize the business impact of data migration is to use best practices that incorporate planning, technology implementation and validation.

In order to prevent these problems from affecting business operations, the vast majority of data migration projects typically are scheduled to occur during off-hours, primarily during weekends. However, this can increase migration costs as a result of staff overtime, and it can negatively impact IT staff morale. Furthermore, taking systems down for migration, even over the weekend, can severely affect business operations, especially if there are problems bringing the systems back online.

In fact, the potential problems with data migration cause some organizations to delay the deployment of new technology, or even to delay purchasing new technology. Such delays can, in and of themselves, be detrimental, because older hardware may require more hands-on maintenance, generally has lower performance and is more prone to failure. Most organizations seek to deploy new technology to eliminate such issues; therefore, delays in implementing new technology present a business risk. In addition, delaying deployment of a new storage device that has already been purchased or leased raises its effective cost, as the company is amortizing the cost of both the old and new devices or is paying lease fees for both old and new devices.

How can organizations minimize the business impacts of data migration – downtime, data loss and increased cost? The best way is to employ a consistent, reliable and repeatable methodology for migrations that incorporates planning, technology implementation and validation. The rest of this paper focuses on migration best practices, as well as supporting software technologies.

Highlights

A robust data migration methodology progresses from plan to migrate to validate.

Data migration methodology

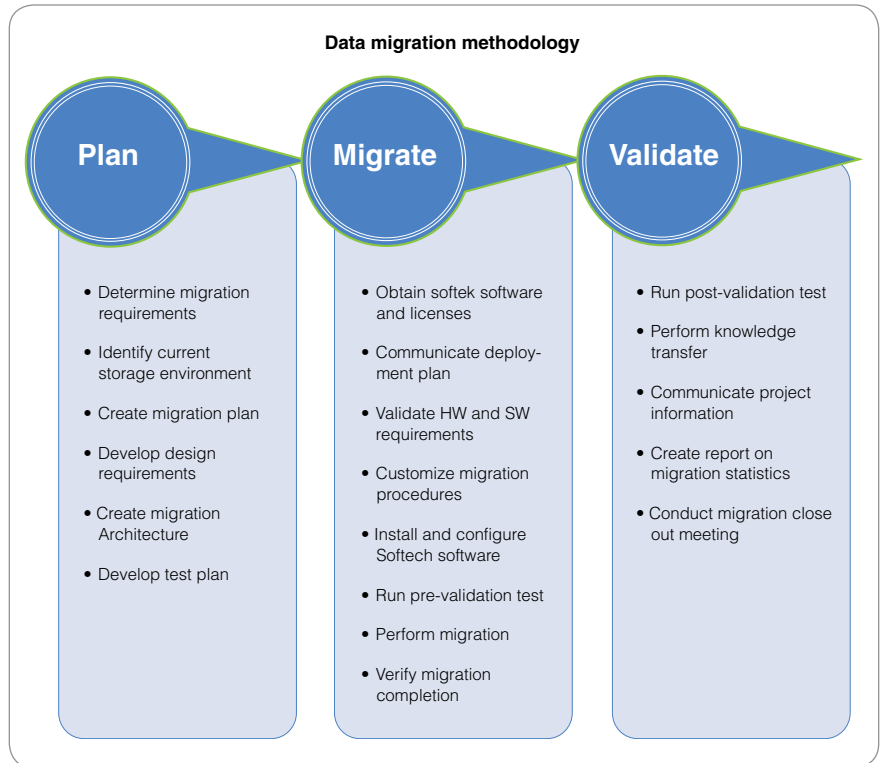


Figure 1: Migration methodology

Plan

IT organizations recognize the critical importance of planning in technology deployment. In fact, over 40 percent of respondents to the survey reported that planning and staging a migration project took two to four weeks.

Highlights

Solid migration planning can help identify potential problems and how to avoid them or, if problems are unavoidable, help IT professionals define migration strategies.

While the amount of planning depends on the size and scope of the migration, the planning process generally should involve determining the requirements of the migration, identifying the current and future environment, and creating and documenting the migration plan. During the planning stage, a determination of what hardware or software is needed to successfully perform the migration is required. The design requirements include migration architecture, specific hardware and software requirements, migration procedures, and deployment and test plans. As necessary, the IT organization should also obtain any software licenses it needs to perform the migration.

The more important the data is to business operations and the greater the complexity of the environment, the more critical migration planning is. Solid migration planning can help identify where potential problems might occur and how to avoid them, and can help IT professionals define mitigation strategies if problems are unavoidable. Migration planning also can help define which data to migrate first, whether and how long to take applications offline, and which internal and external audiences should be informed regarding the migration.

Proper migration planning involves more than just the IT staff. The business owners of the applications and data being migrated should also be included—particularly as the IT organization determines how important a given application or

Highlights

This table depicts a suggested list of action items for a migration plan.

Table 1: Example migration plan

Action item	Assigned to	Status	Date
Establish a migration management team			
Gather availability and production schedules			
Document change control procedures so they can be incorporated into the migration procedures and plans			
Document the timeline for activities for both hardware changes and the data migration			
Announce the migration at least 30 days before the intended target migration date			
Gather information about the storage server environment and applications (list and/or drawing)			
Work with the storage vendor to understand the new storage configuration			
Create a technical migration team			
Inform the security and compliance groups about the migration			
Schedule a premigration rehearsal that includes all the members of the migration team and a data sampling that will enable the application groups to appropriately conduct the pre- and post-migration verification process			
Follow the required change control process			
Establish a migration status call-in process			
Utilize a migration planning checklist to ensure that all the premigration planning steps have been executed			

Highlights

Migration design requirements include migration/replication requirements, time schedule, vendors and hardware configuration.

set of data is to the business. Otherwise, for example, the IT department might plan a migration of the financial system on the same weekend that the finance department is finalizing quarterly numbers. While this may represent a unique situation, the point remains that keeping all interested parties in the loop about data migrations can minimize or eliminate possible conflicts.

In planning a migration, it is important to understand design requirements such as migration/replication requirements, time schedule, vendors involved, and the configuration of the hardware. When sizing data migrations, there are many key items to consider such as the number of servers, the operating system levels, the amount of storage, the volume managers, types of databases and applications, network speeds, and server clusters. In looking at the time schedule, the IT organization should create estimates for planning (planning typically accounts for 25 percent of the total schedule), installation and setup time, data copy time and production cutover.

An example of a schedule for a simple data migration is shown here:

- *Planning: two days*
- *Installation and setup: two hours*
- *Data copy: two hours and 30 minutes for one server*
- *Production cutover: minimal, performed in offline mode*
- *Vendor onsite: eight hours*
- *Professional services: three days*

There are a variety of software products that can be used to migrate data, including volume management products, host- or array-based replication products, and relocation utilities, as well as custom-developed scripts. Each of these has strengths and weaknesses—including performance, operating system support, storage vendor platform support, and whether or not application downtime is required to migrate the data. Some of these products enable online migration

Highlights

IT organizations must carefully explore software options in order to determine the best software technology to use for each migration.

This table shows a requirements list that can be used to help simplify migrations.

of data – meaning that applications don’t need to be taken offline during the migration process. A subset of these provides nondisruptive migration – meaning not only that applications remain online, but also that application processing continues without interruption or significant performance delays. Therefore, IT organizations should carefully explore software options. Specific requirements will help determine the best software technology to use for each migration.

Table 2: Design requirements—understand these requirements to simplify migration

Server environment	
	Server manufacturer
	Number of CPUs
	Number of logical partitions (LPARs) or domains
	Type of file system (UFS, VxFS, HFS, JFS, JFS2 [inline or outline], NFS, NTFS, FAT, FAT32)
	Operating system (OS) version (IBM AIX® 5.1, z/OS 1.4 platform)
	OS addressing (31-bit, 32-bit, 64-bit)
	Databases to be moved (IBM DB2®, IBM Informix®, Oracle, SQL, Sybase databases)
	Database version
	Database size
	Availability requirements of databases (any existing SLAs, downtime issues to consider)
	Cluster environment (such as MSCS, VERITAS, Sun, IBM HACMP™, and MC/Service Guard environments)
Storage environment	
	Storage vendor and model
	Channel type (ESCON, FICON, Fibre, iSCSI, SAN) SAN HBA and Model (Qlogic, Emulex, JNI)
	Number of channel paths
	Logical to physical mapping (i.e. RAID-1 vs. RAID-5)
	Number of source volumes to be migrated
	Volume sizes
	Target volumes to receive source data
Network environment (if applicable)	
	Topology
	Speed of network

Highlights

Documenting the migration process can help train staff and simplify or streamline the next migration, helping to reduce both expense and risk.

Migrate

During the migration phase, the IT organization will need to communicate its plans; obtain, install and configure any necessary software; and perform the actual data migration. A premigration data validation test is recommended, in addition to post-migration validation testing. These tests confirm that the data is in the same state after the migration as it was before.

Clearly, the most important part of this stage is the migration itself. As outlined above, software technology can simplify this process by enhancing the speed of migration, by minimizing or eliminating application downtime, and/or by enabling migration during regular business hours, helping the organization to get back to business as quickly as possible.

Validate

After the migration has been completed, the IT organization should compile migration statistics and prepare a report to highlight what worked, what didn't work and lessons learned. The report should be shared with all members of the migration team. These types of reports are critical in building a repeatable and consistent process through continuous process improvement – building on what worked and fixing or changing what didn't work. Further, documenting the migration process can help train staff, and simplify or streamline the next migration, reducing both expense and risk.

Table 3: Migration methodology and approach

Migration and validation methodology
Based on the information gathered in the planning phase, structure the migration architecture to match the production requirements
Use checklists to ensure any operating patches and software are at the correct levels
Build detailed migration procedures following the chosen architecture
Put together a schedule of events with timelines to implement the migration procedures
Establish an initial test plan to validate the initial installation of the Softek software
Develop a cooperative deployment plan
Install and configure Softek software
Run the simple initial test plan that validates the software installation
Implement the migration procedures and timeline built in the design phase
Verify the migration completion by checking the successful completion and status of the migration jobs

Highlights

Performance must be balanced against network bandwidth and system overhead.

Choosing data migration software

As discussed earlier, there are a variety of commercially available software technologies, such as Softek’s TDMF technology, that can be used for data migration in mainframe, UNIX, Linux and Windows environments. What are some of the necessary features to consider in migration software? Clearly, the software chosen should support the operating system and source hardware platform on which the data resides, as well as support the target hardware. However, there are many capabilities beyond these that should be considered.

One of these key attributes is performance – how quickly the data is copied from the source to the target. However, performance must be balanced against network bandwidth and system overhead. If the data is copied at a high speed but consumes too much bandwidth or I/O, production applications or systems can be severely affected. On the other hand, if data is copied too slowly, the migration may take longer than anticipated, potentially prolonging downtime. Some migration software products, such as Softek’s TDMF solution, include a throttling or pacing capability that minimizes impact on production applications, thus enabling faster data movement when systems allow, and slowing down movement when I/O is required for other purposes. This capability helps IT organizations to more easily balance migration versus other system demands.

Highlights

It's important to consider whether or not the migration can be easily terminated or restarted.

Another migration challenge is incompatible source and target hardware.

Another requirement to consider involves the ability to “roll back” the migration. Essentially, if something should go wrong, can the migration easily be terminated and restarted or can application processing continue on the source data/device? This can be problematic with some technologies, such as volume managers, while it is easily done with others, including TDMF technology.

In many cases, one of the migration requirements is an increase in volume size from the source to the target—nearly 40 percent of the time according to the Softek survey. Therefore, if it is necessary to increase the volume size, the IT organization needs to be sure to select migration software, such as TDMF software, that supports this capability.

Another common migration situation involves unlike source and target storage hardware. While host-based products support unlike storage devices, most array-based products require that the source and the target come from the same vendor, and may require that they be the same type or generation, and/or the same firmware version on the storage device. TDMF technology, on the other hand, is designed to be hardware independent.

One of the primary reasons that data migrations occur during off-hours is to avoid application downtime during peak periods. Depending on the type of data and applications being migrated, only a narrow downtime window may be available.

Highlights

TDMF software supports nondisruptive migration—the application can stay online and continue to process data and transactions throughout the migration process.

Host-based migration technologies—such as TDMF software—generally provide more flexibility than do array-based technologies.

For example, if the HR system is offline for 24 hours over a weekend, there would be less business impact than if the online commerce system were offline for the same period. Some systems may be so critical to business operations that a few hours or even minutes of downtime—even during off-hours—may be unacceptable. A significant upgrade of such a system will require that the IT organization makes sure to minimize any downtime—and using that downtime for data migration will not be feasible. TDMF software enables nondisruptive migration—meaning that the application can stay online and continue to process data and transactions throughout the migration process.

Overall, host-based migration technologies generally provide more flexibility than do array-based technologies. In addition, they provide advantages over volume management products in that they are optimized for migration and may include capabilities such as throttling or roll back. Also, host-based migration products, such as TDMF software, do not rely upon proprietary volume managers. Finally, host-based technologies can be lightweight enough to remain installed in the storage environment to assist with day-to-day migration activities. For example, many of Softek’s customers purchase TDMF technology for a specific migration project, and then extend their license and continue to use it for additional routine migration tasks.

Highlights

Softek TDMF software enables the nondisruptive data migration that today's 24x7 data centers demand.

Summary

Data migration is a routine part of IT operations in today's business environment. Even so, it often causes major disruptions as a result of downtime or application performance problems, and it can severely impact budgets. To prevent these problems, organizations need a consistent and reliable methodology that enables them to plan, design, migrate and validate the migration. Further, they need migration software that supports their specific migration requirements, including operating systems, storage platforms and performance. In addition, migration products that maintain continuous data availability during the migration without affecting performance are desirable. Softek's TDMF data migration software provides a variety of capabilities to support such requirements in mainframe, UNIX, Linux and Windows environments – and TDMF software is designed to be storage vendor independent.

For more information

For more information about Softek's TDMF data migration software, visit:

ibm.com/services/storage/migration



© Copyright IBM Corporation 2007

IBM Global Services
Route 100
Somers, NY 10589
U.S.A.

Produced in the United States of America
06-07
All Rights Reserved

IBM, the IBM logo, AIX, DB2, HACMP, Informix and z/OS are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both.

Softek and TDMF are trademarks or registered trademarks of Softek Storage Solutions Corporation in the United States, other countries or both. Softek Storage Solutions is an IBM company.

Microsoft and Windows are trademarks of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product and service names may be trademarks or service marks of others.

References in this publication to IBM products or services do not imply that IBM intends to make them available in all countries in which IBM operates.

¹ All survey results are from Softek's 2005 Worldwide Data Migration Survey of 700 IT personnel.