Mike Dutch

HITACHI
DATA SYSTEMS

Point-in-time copy solutions

Leveraging time and data to advantage

This paper discusses the following replication solutions:

AMDAHL	Transparent Data Migration Facility
EMC²	TimeFinder
HITACHI DATA SYSTEMS	ShadowImage
IBM	SnapShot

Introduction

A point-in-time copy provides access to data as it existed at a specific point in time. Different replication technologies impact the timing and extent to which physical data movement is required to create such a copy. The purpose of this paper is to compare various approaches to creating a point-in-time copy and to outline implementation differences to help you choose the replication solution most appropriate for your business.

Value

Simply put, effective use of point-in-time copies allows your business to focus on revenue producing activities while remaining confident your mission-critical data is protected. Although component redundancy and failure handling technologies like RAID and Remote Copy offer protection from physical failures they do not offer protection from logical data errors. Point-in-time backups allow applications to recover to a valid state when logical inconsistencies are discovered. Point-in-time copies can also be used to meet the challenges of non-stop business processing and to enhance business agility by enabling work to be completed in parallel and by leveraging your business intelligence activities. Specifically, point-in-time copies can be used to improve productivity, quality, and readiness through the following applications:

- **Data Sharing** – Transforming data into effective business decisions involves exchanging data between platforms, processing it to discover significant relationships, and placing the resulting information in context to better understand important factors in the competitive environment. Point-in-time copies allows data exchange to occur without impacting revenue producing transactions.
- **Time-Shifting** – Applications that need to process the same data may be able to use point-in-time copies to work in parallel or to shift processing to a more optimal time. Productivity is improved by overlapping processing in the same elapsed time. Availability is improved by reducing application outages to the time necessary to create a point-in-time copy after reaching a logically consistent state. For example, a backup window can be reduced by first creating a fast point-in-time copy and then resuming the applications which require access to the data. Offsite backups or data transfer can then be performed using the copy without impacting the availability or performance of the applications and without fear of missing an important backup.
- **Application testing** – Point-in-time copies allow multiple test activities to be performed in parallel and with a complete, consistent, and reusable copy of operational data. Exposure to data corruption and system failures can be isolated to the test environment by using point-in-time copies and host facilities which are physically separate from the production environment. Each application can clone their own copy of a standardized and comprehensive set of test data as often as they wish to improve the quality and speed of their test activities. This capability is particularly important for successful Year 2000 compliance efforts. On a business level, Meta Group stressed the competitive significance of such efforts in "*Y2K: A Practical Business Guide*": "*We believe each industry will see significant consolidation, as Y2K survivors acquire Y2K victims that have been devalued by their Y2K woes.*"
- **Fast replication** – Traditional replication utilities begin copying data when a point-in-time copy is requested and may disrupt applications until physical data movement is completed. More recent replication technologies allow data movement to occur either before or after the point-in-time copy is requested, thereby dramatically reducing the effective copy time window. This fast copy capability not only enhances recovery and

exchange capabilities by enabling faster and/or more frequent backups or refreshes, it allows pre-formatted data to be rapidly cloned.

- **Logical error protection** – Faster and more frequent backups allow more rapid and up-to-date recovery from logical errors. It is important to distinguish this requirement from physical failure protection provided by real-time recovery techniques such as RAID. A business cannot operate successfully using invalid data, no matter how many copies of this invalid data it has or can generate.
- **Disaster protection** – A business disaster occurs when the extent of physical failures or logical errors threatens the viability of continuing as an ongoing concern. Potential causes are many, both environmental and of human origin, and the impact may be sudden or gradually spreading. Business recovery plans will involve safety, skill, procedural, and infrastructure issues in addition to recovery of Information Technology resources like applications, networks, and storage. Point-in-time copies can play a vital role in ensuring data access during an actual recovery situations and perhaps just as important during simulated business recovery scenarios where outages to production activities must be minimized.
- **Point-in-time record** – Records retention policies and vital records requirements may make it essential to obtain copies as of a specific point-in-time such as at fiscal year end. Such archive copies are not intended to be used for recovery purposes but rather to support historical, audit, or legal concerns.

The value that a customer enjoys from using point-in-time copies depends upon the extent to which usage is integrated with their ongoing operations and the particular replication technology employed. This paper is intended to help clarify the benefits of each technology and suggest where particular implementations may be best used for achieving competitive advantage.

Approaches

High-end storage subsystem vendors have taken three general approaches to extending the capabilities of traditional copy utilities which sequentially copied data from one media to another. The first approach, referred to here as *Split Copy*, was an outgrowth of customers splitting volume mirrors to make point-in-time copies non-disruptively. Designs intended specifically for making point-in-time copies have added the capability to track changes to the source and target volumes so they can be synchronized by only copying changed data. Another approach, referred to here as *Asynchronous Copy*, builds upon traditional copy utilities but separates logical from physical copy completion. A more recent approach, referred to here as *Pointer Copy*, does not copy data at all; instead it creates a new set of pointers to the same blocks of data. Updates are made by writing changed blocks to another location and updating the pointers associated with the copy that was updated; pointers representing other copies continue to reference the original blocks. Each of these approaches offers different advantages and introduce tradeoffs which should be understood prior to determining the best replication solution for a particular application.

- **Split Copy** – This approach copies data *before* the copy point-in-time. A copy is created by simply stopping to maintain synchronization between the source and target disks. Planning is required to identify the data that is to be copied and ensure that sufficient resources are available to maintain the desired number of copies. The time to re-synchronize source and target data (allowing reuse of storage) can be optimized by copying only changed data after the initial synchronization. A significant advantage of this approach is that a physical copy is guaranteed to be immediately available when the copy is split. Different implementations may provide other performance enhancements, offer multiple copies of the same point-in-time, and support copies to other subsystems and sub-volume granularity.

- **Asynchronous Copy** – This approach copies data *after* the copy point-in-time. A copy is created in a manner similar to traditional copy utilities but the source data may be updated while the physical point-in-time copy is being completed. This is achieved by copying source data to a temporary holding area before an update is allowed to complete and later writing the data in proper sequence to the output device. This approach cannot guarantee a copy at a specific point-in-time because the physical copy may not complete successfully. Copies made in this manner are also not available for use until the physical copy completes so applications must distinguish between logical and physical copy completion. However, this approach does offer significant benefits. Backup windows can be virtually eliminated and copies may be made at the file or the volume level. The point-in-time copy may also include data from multiple subsystems from multiple vendors and be written to tape as well as disk.
- **Pointer Copy** – This approach *does not copy data*. A point-in-time copy is created by providing a separate path to the same data. The virtual copy is guaranteed and the source and target copies are both immediately available for use. Updates to each copy may be made independently because changed data is written to a new location while retaining the original data for reference by other copies. This approach supports file and volume level copy as long as the source and target data both reside on a single subsystem. The amount of storage consumed will vary depending on the number of updates made to each copy. Performance will be impacted according the number of applications trying to read the same data at the same time. This approach must be used in conjunction with other replication technologies when production and test facilities must be isolated or when disaster protection is paramount due to its “single subsystem” and “uncertain impact” limitations.

Finding the right solution

Selecting the most appropriate replication technology to solve a business problem involves matching the requirements of a solution with the capabilities of the technology. Once an appropriate technology is selected the alternative implementations of each technology should be explored to understand unique advantages and financial considerations. The “right” solution for one company may not be the most appropriate for another when unique factors are considered.

Split Copy technology is well-suited for applications where synchronization times, storage consumption, and operational integration issues are not problematic. This means it is most appropriate for data sharing, time-shifting, application testing, protection, and point-in-time record applications. It may also be appropriate for fast replication depending on the lead time and performance optimizations that a particular implementation provides. Storage consumption issues are usually less relevant to a business decision when total financial impacts are considered.

Asynchronous Copy technology is well-suited for applications where guaranteed point-in-time and immediate copy reuse are not required or when data from multiple vendor subsystems are involved. The ability to support file level copy and tape output devices can be used to particular advantage in disaster protection and point-in-time record applications. It is not well-suited for data sharing, time-shifting, or application testing applications which require immediate access to the point-in-time copy. Most implementations are also restricted to data residing on an OS/390 host.

Pointer Copy technology is well-suited for inline application backups in which backups only need to be retained until a batch update process completes successfully. It is also well suited for cloning pre-formatted data quickly. Other applications can use this technology to advantage but only when used in conjunction with other point-in-time replication solutions. The relevant issue is whether the costs associated with Pointer Copy implementations can justify the incremental benefits.

Point-In-Time Copy Solutions

Implementation differences can play a crucial role in selecting the optimal solution for your business. Amdahl, EMC, Hitachi Data Systems, and IBM each offer multiple point-in-time copy solutions and understanding the technical differences between them will allow product selection to move beyond the "what should I buy" question to focus on financial and relationship considerations.

Split Copy solutions are offered by most of the high end storage vendors and significant differences exist between them though all implementations copy a source disk to a target disk of the same format. Real-time copy solutions from each vendor (APRC and XRC from Amdahl, SRDF from EMC, Remote Copy from HDS, and PPRC and XRC from IBM) could also be used to achieve a similar result though additional hardware, software, and skills would be required and performance and capabilities would be impacted. Service offerings, such as HS-DataPlex, are also available for customers wishing to implement turn-key point-in-time copy solutions.

- **Amdahl** – offers the *Transparent Data Migration Facility (TDMF)* and although originally intended to provide non-disruptive OS/390 data migration between disk subsystems from any vendor, it has recently been marketed as a generic copy and movement solution for batch window reduction, Y2K testing, data warehousing, and I/O load balancing. TDMF is an OS/390 application that copies all data on a volume once, then makes repeated passes to capture recent changes, and breaks the pair relationship when the final refresh can be performed within a specified time. A guaranteed point-in-time copy can be achieved by starting TDMF before the online systems have ended and then splitting the pairs after logical consistency is reached by waiting for operator approval or by using automation packages. TDMF differs from other Split Copy offerings in that the target volume is accessible from the host before it is split so care must be taken to ensure that no other I/O is directed to this volume. This approach allows the source and target volumes to reside on disk subsystems from different vendors but is slower than a hardware-based facility and is restricted to OS/390 data. It does not offer the ability to re-synchronize pairs or to create multiple copies of the same data at the same time and is most appropriate when copy time and memory are not an issue. Amdahl indicated they had over 140 TDMF customers and 430 licenses worldwide when they announced Release 1.3 on 23-March-1998.
- **EMC –** offers *TimeFinder* to copy data from a standard volume to a Business Continuance Volume (BCV) in the same Symmetrix disk subsystem. Customers must plan accurately for the peak capacity that will be used for making copies because separate disks must be dedicated for BCV use. EMC must then configure storage using a disruptive configuration process.

  A point-in-time copy is created by splitting an active BCV from a standard volume. A standard volume must always be protected, either using RAID 1, RAID S, or by using the Symmetrix Remote Data Facility (SRDF), a $100,000+ option. A standard volume is paired and synchronized with a single active BCV using the ESTABLISH command. Support for "3 active copies" refers to the source disk and its parity information as two additional "copies". In SRDF configurations, either or both volumes in the SRDF pair can each be paired with one active BCV.

An active BCV cannot be mirrored. Although a split BCV can be mirrored using RAID 1 or SRDF, a BCV cannot be split if the standard volume protection scheme is compromised. Recovery from an active BCV disk failure will require full volume copy. The QUERY command reports the device number for all BCVs and paired standard devices, invalid track counts, and BCV availability status (available, incomplete, establish-in-progress, in-use, or split).

Data is copied in the background from the standard volume to the BCV when a pair is established. EMC has stated that his process requires about 1.5 minutes per GB in a best case scenario. A message indicates when initial synchronization is achieved. Updates to the standard volume are mirrored to the active BCV until the pair is split.

After a pair is split the standard volume and BCV can be accessed independently. Later, the same or another BCV can be paired with the standard volume. If the BCV was previously paired with the same standard volume re-synchronization can be achieved by copying tracks which changed on either the source or target volumes from the source volume. The RE-ESTABLISH command uses the standard volume as the source; the RESTORE command uses the BCV volume as the source.

TimeFinder, a $65,000 option, provides software to manage BCVs from MVS, UNIX, and Windows NT hosts. The MVS Batch Utility, executed as a batch job, allows manual and automated control of BCV operations. A Conversion Utility is included so volumes and data sets can be renamed and used on the same MVS host as the original data. The SRDF Host Component V3+, a $22,000 option, also provides control of BCVs from an MVS started task. A library of UNIX and Windows NT commands provides a command line interface for Open Systems hosts. Shell scripts are included for supported UNIX hosts. Symmetrix Manager for Open Systems V2+, a $15,000 extra charge, offers a $22,000 Control Option which adds a graphical user interface (both Motif-based and a Windows NT GUI). Starting with 5x64 microcode, BCV volumes residing on any Symmetrix ESP-supported host system can be managed from one of the supported TimeFinder management platforms. Note that Enterprise Storage Platform (ESP) support also carries an extra charge ($18,000 for 5300, $36,000 for 5400, and $42,500 for 5700).

The ESTABLISH, SPLIT, RE-ESTABLISH, RESTORE, QUERY, and VERIFY commands are common across all supported hosts. The GLOBAL, USEREXIT, and Conversion Utility commands (DEBUG, SIMULATE, RELABEL, PROCESS, CATALOG, RENAME) are also used on MVS hosts. The SETUP command is used on Open Systems hosts.

Consultants report EMC sold about about 1000 TimeFinder licenses in its first year and currently sells about 150 licenses per quarter, 58% for Open Systems. 63% of the licenses are from North America, 30% from Europe, and 7% from Asia Pacific.

- **HDS** – offers *ShadowImage* to create point-in-time copies using an enhanced version of Split Copy. ShadowImage is supported on all 7700E subsystems with WP102B level ACP pairs. It offers these enhancements over the TimeFinder implementation**:**
  - Any volume can be used to create point-in-time copies. This has two significant advantages: planning requirements are reduced and storage is not wasted.
  - Disruptive configuration of dedicated disks is not required.
  - All volumes are RAID protected at all times supporting higher application uptime.

- Up to three active copies may be paired with the same source volume. This supports higher availability and faster access for multiple applications that use copies of data reflecting the same point-in-time.
- ShadowImage copies can be created from Remote Copy volumes. This allows up to 6 point-in-time copies to be created from the same source volume: 3 on the primary 7700E and 3 on the secondary 7700E. TimeFinder supports a maximum of only 2 active BCVs in an SRDF configuration, 1 in each Symmetrix.
- Up to 16 data mover “jobs” can be used to establish volume synchronization within the 7700E disk subsystem allowing a full 3390-3 format volume (about 2.8GB) to be copied in about 2 minutes, over 50% faster than TimeFinder.
- Performance of maintaining mirrored copies is optimized by batching updates after initial synchronization is achieved. Updated tracks will be mirrored every 5 minutes if 1000 tracks have been changed or when 15 minutes has past since the last synchronization. By reducing the load on the subsystem more active pairs can be supported while other applications can achieve higher I/O performance.
- A graphical user interface is supplied as part of the complete solution. It is not an optional extra cost add-on. This interface can be used to obtain ShadowImage status for multiple copies or when an HRC volume is part of a ShadowImage pair.
- A subset of the current Remote Copy commands are used to support ShadowImage operation allowing staff to become productive quickly by providing an opportunity to leverage skills. The CESTPAIR, CDELPAIR, CSUSPEND, and CQUERY PPRC commands can be used to add/resume, delete, and split pairs and to check status. The 7700E distinguishes ShadowImage and Remote Copy operations by noting whether the volume serial numbers are the same or not. Command details are provided in the Remote Copy Administrator’s Guide and Reference (SC35-0169) and in the ICKDSF R16 User’s Guide (GC35-0033).
- The 7700E provides multiplatform capability as a standard capability. Symmetrix 5000 series requires the extra cost ESP feature to add Open Systems capability. Additional costs are associated with unused disks dedicated to the BCV pool.

It is apparent that ShadowImage offers customers significant benefits over alternative Split Copy implementations.

- **IBM** – invented the disk drive, the floppy disk, and the MR head, filed the first patent for RAID storage, and introduced RAID 1 systems in 1990, RAID 3 systems in 1991, and RAID 5 systems in 1992. Unfortunately, its current product offerings have not kept pace with its past history of innovation and no form of Split Copy is currently marketed. It had previously offered user control of Dual Copy hardware mirroring on its 3990 storage control which some customers used for creating point-in-time copies. It is expected that IBM will provide a point-in-time capability when it introduces a multiplatform model of its Versatile Storage Subsystem; however, compatibility with Concurrent Copy or SnapShot is more likely than introduction of Split Copy capability.

Asynchronous Copy solutions are offered by each of the high end storage vendors and are all compatible with IBM’s Concurrent Copy, though EMC refers to it as the Symmetrix Backup/Restore Facility (SBRF). No vendor charges extra for this capability and although marketing efforts have been minimal the majority of OS/390 customers who have implemented DFSMS are also using Concurrent Copy. Current implementations do not support concurrent use of Asynchronous Copy and Split Copy for the same data. Support to distinguish logical from physical completion has been added to IMS, DB2, and CICS Backup-While-Open (BWO) and the interface has been modified to use SnapSnap Copy instead of Concurrent Copy when possible; this capability is referred to as Virtual Concurrent Copy. Perhaps the most important implementation consideration to note is that Concurrent Copy is supported on the Amdahl Spectris, EMC Symmetrix, HDS 7700/E, and IBM RAMAC Array subsystems. It is not supported on the RAMAC subsystems manufactured by StorageTek (RVA and RSA). *The business value provided by the complete disk subsystem should be evaluated when considering Concurrent Copy because each vendor provides virtually identical capability for this stabilized point-in-time replication technology.* Similar technology is being introduced into the “Open Systems” environment by other vendors using different terminology.

Pointer Copy solutions are offered by IBM through its distribution arrangement with Storage Technology Corporation. Although not strictly limited to log-structured file system disk subsystem implementations, this approach is currently implemented solely on the StorageTek Iceberg (also marketed as the IBM RAMAC Virtual Array) disk subsystem.

- **IBM** – offers SnapShot to create virtual copies within a single RVA disk subsystem. SnapShot suffers the inherent limitations of any Pointer Copy solution as discussed earlier, namely, variable impacts on capacity and performance and the need to be used in conjunction with other point-in-time replication technologies. The SnapShot offering suffers additional implementation-specific shortcomings which make it inappropriate for many business applications.
  - Overall capabilities of the RVA are not competitive with the rest of the industry. Lower vendor costs achieved by using compression/compaction techniques have not resulted in demonstrable savings for the customer. The log-structured file architecture used to enable compression, self-tuning, and SnapShot has given rise to performance and functional characteristics that make RVA better suited for data archival than for commercial data storage. Open Systems support is severely limited by requiring S/390 host attachment and communication through a performance and capacity challenged ESCON/SCSI converter (Cross Platform Extension). Unfortunately, learning and managing a disk subsystem with unique characteristics often results in raising skill requirements and labor costs.
  - Limited device addressing (256 logical volumes) gives rise to SnapShot limitations because a volume cannot be copied if there are insufficient targets. The delays in resolving this high priority customer requirement indicate a major restructure of the microcode is required. Experience suggests code stability would be compromised and customers would be exposed to availability issues.
  - Perhaps the most significant problem associated with SnapShot is its impact on current operations. We have already discussed why Pointer Copy solutions are best suited for inline application backups (including control data set backup) and for rapid cloning of pre-formatted data. The difficulty is that applications must be changed to make use of SnapShot for even these limited uses. Of course this is true with other point-in-time copy solutions but these solutions are more generally applicable and skills and experience can be transferred to other applications. To be fair, it should be mentioned that applications that were already changed to use Concurrent Copy will be candidates for *Virtual Concurrent Copy*. This simply means that SnapShot will be used when a Concurrent Copy is requested if the source and target data both reside on the same RVA. The practical value of this may be questioned but it is laudable an attempt was made to improve usability.
  - Experience has shown that SnapShot performance can be materially affected by the number of extents and the number of SNAP DATASET commands run in parallel. The degradation stems from system overhead involved in allocation and catalog management not with the size of an extent. IBM has recommended that customers take actions to reduce the number of extents and increase the number of jobs run in parallel. Many customers will find that the time spent in micro-managing RVA storage will not prove to be a cost-effective investment.
  - A command line interface is provided to “SNAP” MVS data sets and volumes and VM minidisks. The Concurrent Copy MVS interface can also be used to invoke

SnapShot (though this results in using a temporary file so a point-in-time copy can no longer be guaranteed). No graphical user interface or Open Systems interface is provided.

- Previous discussions noted that the relevant issue is whether the costs associated with Pointer Copy implementations can justify the incremental benefits. This is particularly relevant for SnapShot where acquisition costs alone include $365,000 for the RVA enabling diskette (feature code 6001), $200/month per RVA for IXFP 2.1, plus the cost of the SnapShot software for the MVS or VM host (varies by processor size; from $39-$1000/month/processor or for a one time charge of between $1,880 and $48,000).

RVA fundamentally introduces risk at a business level. The variability of storage consumption and application performance raise the technical level of uncertainty. However, customer investments in RVA are implicitly threatened by IBM's SeaScape architecture. If IBM delivers a competitive enterprise disk subsystem there can only be a future of purposeful neglect for RVA. RVA roadmaps may include features to address glaring shortcomings (PPRC, native SCSI attachment, >256 logical volumes) but what motivation does IBM have to meaningfully improve StorageTek's Iceberg? The incremental value which RVA SnapShot can offer is overshadowed by the lost opportunities for achieving business value using more competitive disk subsystems.

Operational Integration

The extent to which customers integrate the use of replication technology in their ongoing procedures directly impacts the business value they will actually achieve. Some general comments are in order. First, regardless of the technology employed all point-in-time copy solutions require that the data to be copied is in a logically consistent state. Host buffers must be written to disk and application updates must be prevented while the point-in-time copy is made. If the unit of consistency spans multiple volumes then an image of all data should be captured. Second, operational procedures and applications or job streams will often require change. Automation activities and copy procedures will have to be coordinated with the applications that use the data. No matter how trivial a change might be complexity will arise when put in the context of production workloads. Optimal business results will require some amount effort which should not be under-estimated. Third, volume level point-in-time copies will often involve file system preparation (especially in the MVS environment) and security considerations. This work may be performed automatically by replication solutions with complete file level support, it may be assisted by providing adjunct utilities, or it may be left entirely to the user. The important point is that it is not enough to just make a point-in-time copy – businesses must know how to use that copy for its intended purpose. Fourth, an obvious point but one not to be overlooked when selecting an appropriate solution is to ensure that the data you want to copy is supported by the solution. If you want to copy Open Systems data you don't want to depend on Concurrent Copy, just because it carries no extra charge. That said, you also need to examine the incremental benefits and opportunity costs that are associated with a particular solution to make a prudent business decision. You may find that other alternatives such as traditional copy utilities or real-time copy facilities are more appropriate for your replication application.

Conclusion

We have examined the tradeoffs of alternative implementations of point-in-time copy replication solutions. It has become clear that Split Copy solutions are more generally applicable than either Asynchronous Copy or Pointer Copy solutions. The case has been made that *ShadowImage* offers the most competitive solution in the industry. Finally, general remarks concerning the impact of operational integration on achieving practical results reaffirmed the adage about there being no free lunch. This paper has examined how point-in-time copy solutions can be used to leverage time and data into competitive advantage. Replication solutions may not buy you lunch but they can buy you time.

Trademarks