
System z and Storage Synergy

A view through the decades and into the future
Spring Share 2011 Session #8958

Scott Drummond – spd@us.ibm.com



- IBM System Storage and System z are a winning combination and collaborating to support growth and protection of mission critical data. This session will highlight how DFSMS and z/OS are exploiting Storage Disk and Tape hardware features and functions. We'll cover High Performance FICON with Multi-track, Dynamic Volume Expansion, EAV, SSD, Space Efficient FlashCopy, HyperSwap, Encryption, Virtual Tape Policies and other topics during this presentation.

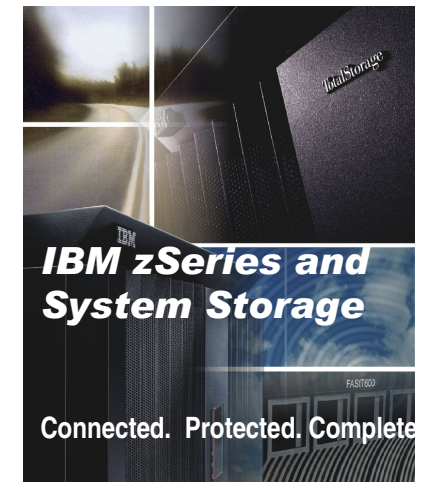
IBM System Storage: 50+ years of IBM storage innovations leading to today's broad range of storage solutions for zSeries



- **Offers one of the broadest portfolios of storage solutions in the industry**
Includes disk, tape, SAN, NAS, software, and services. (one-stop-shopping)
- **A leader in the disk and tape storage innovation** –
More than 50 years in the business, receiving more US patents, including patents in storage technology, than any other company.
- **Addresses the needs of a variety of storage capabilities required for zSeries servers**
Supports transaction processing (OLTP, batch, sequential), backup and recovery, and long term data retention for both stand-alone and networked servers as well as support for a wide variety of zSeries operating systems including Linux, z/OS, VM, VSE, TPF
- **Designed to provide high performing, efficient technology**
Includes Storage Systems zSeries performance/tuning enablers and balanced performance with published benchmark results, as well as many tape advanced functions that improve access times.
- **Designed to have low Total Cost of Ownership and excellent investment protection**
Examples include Flexible hardware and software warranty and many popular features included at no additional charge, tape features which reduce the need for drives, physical space, and personnel.
- **Provides a strong base for infrastructure simplification and consolidation.**
System Storage products are designed for high scalability, non-disruptive upgrades, flexible configuration capabilities to support multiple servers and a mix of applications.
- **Supports round the clock availability.**
The System Storage Resiliency Family ties together a broad set of hardware, software and services that are designed to be integrated with zSeries and other eServers to help provide a resilient, automated 24x7 infrastructure. Includes support for GDPS with disk and tape functions.



- **zSeries and System Storage have a unique relationship, not shared by other storage vendors.**
 - Collaborate on architecture and design
 - Conduct early, rigorous and comprehensive stress testing in zSeries lab for every enterprise storage release. Tests are designed to push the limits of functionality and robustness.
 - Share cross support by skilled resources with enterprise class experience and expertise
- This helps IBM System Storage and zSeries development to:
 - Better design products that work well together with more robust interlocking
 - Bypass problems that other vendors may fall into
 - Gain insight into potential pitfalls with fixes implemented before customers are ever impacted
 - Implement streamlined, efficient, integrated product offerings
 - Certify and qualify products that have passed stringent testing together
- This provides value to zSeries and System Storage customers by helping to:
 - Verify product reliability even in heavy stress environments
 - Speed server and storage implementation
 - Maximize infrastructure investments
 - Avoid problems and reduce risk



Where else could you find this complex server environment, tool set and skills?.

- 1964 - Count Key Data (CKD) This is a self-defining record format used on DASD devices. The count field contains the length of the key (optional) and the length of the data, the key field is optional and if present contains a string used to uniquely identify the data record (ex. account number), the data field contains the actual record. This format was introduced in the IBM 360 as a way to off-load low level I/O functions from the main processor. A key search is performed by the channel and device saving the processor from becoming involved.
- 1964 – IBM 2311 Direct Access Storage Facility
- 1964 – 2321 Data Cell Drive (Noodle Picker – also called Cherry Picker)
- 1966 - 2314 DASD
- 1968 - 2305 DASD

- 1970 - 3330 DASD with RPS, rotational position sensing
- 1975 - MVS 3 Selectable Units (SUs) – rolled up by early 1980's
- 1976 - 3350 DASD
- 1977 - Hierarchical Storage Manager (HSM) announced – supports the 3850 Mass Storage System
- 1976-7 - DSS (later DF/DSS)

- 1981 – 3380 introduced with ECKD
- 1980 Data Facility Extended Function (DFEF) for ICF support – Re-announced 1981
- 1980 Device Support – Data Facility Device Support (DFDS) – included VTOC INDEX support and Media Manager (IMS and DB2 1st usage).
- 1983 DFP/370 (DFEF and DFDS)
- Dynamic Path Reconnect (DPR) with 3880's
- 1989 MVS/DFP V3 System-Managed Storage

- 1991 ESCON
- 1993 DFSMS/MVS
 - DFSMSrmm – Major new tape management system (1.1)
 - DSN Striping (1.1), concurrent copy (1.1), Tape Library Dataserver (1.1), Tape Mount Methodology (TMM) and Compression (1.2), Manual Tape Library (1.2), XRC - z/OS Global Mirror (1.2)
- OS/390 2.10 inclusion of DFSMS/MVS
- Initial PAV and Multiple Allegiance Function (OS/390 1.3)
 - Multiple Allegiance functionality started earlier with TPF
- Dataset Level FlashCopy
- Dynamic PAVs (OS/390 2.7)
- Multiple Allegiance
- 3494/3495 VTS and VTS rules
- Control-unit initiated reconfiguration (CUIR)

- 2000 Native FICON - OS/390 2.8 with ESS
- 2001? z/OS
- 3390 mod 9, "27" and "54" Formats
- HyperPAVs
- EAV
- SSD Support and Easy Tier (DS8000)
- Dynamic Volume Expansion
- High Performance FICON with Multi-track
- Dataset separation by volume (DB2)
- DFSMSHsm Fast Replication
- Encryption and Encryption Management
- OMEGAMON XE for Storage
- Space-efficient volume copies
- HyperSwap
- z/OS Metro/Global Mirror (MzGM) incremental resync after HyperSwap
- Remote Pair FlashCopy
- TPC for Replication
- I/O Priority Queuing ?2000's
- DFSMSHsm and DFSMSrmm Fast Tape Positioning
- 2005 – MIDAWS and subchannel set support
- DFSMSHsm synergy with TS7680 Deduplication

- EAV support – 2010 z/OS R1.12
- DFSMSHsm TS7680 ProtecTier Deduplication – 2010 z/OS R1.12
- z/OSMF DASD Management z/OS R1.13

- EAV support – 2010 z/OS R1.12
- HyperPAVs
- SSD Support (DS8000)
- Dynamic Volume Expansion
- High Performance FICON with Multi-track
- Dataset separation by volume (DB2)
- DFSMSHsm Fast Replication
- DFSMSHsm TS7680 ProtecTier Deduplication
- 7720/7740 policies and Grid
- Encryption and Encryption Management
- OMEGAMON XE for Storage
- z/OSMF
- Dataset Level FlashCopy
- HyperSwap
- Space-efficient volume copies
- z/OS Global Mirror (XRC)
- z/OS Metro/Global Mirror (MzGM) incremental resync after HyperSwap
- Remote Pair FlashCopy
- TPC for Replication

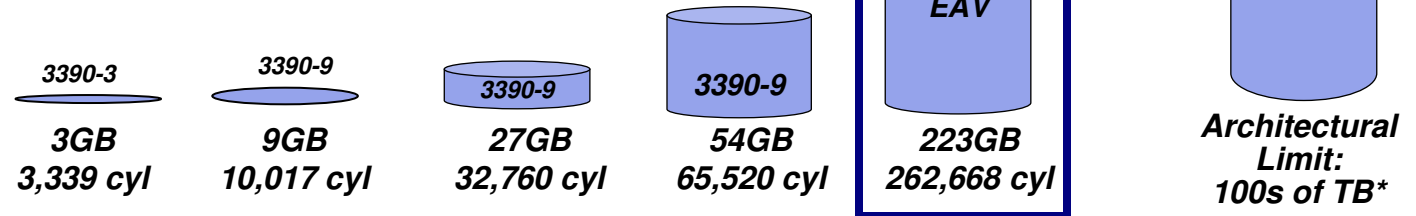
Extended Address Volumes (EAV)



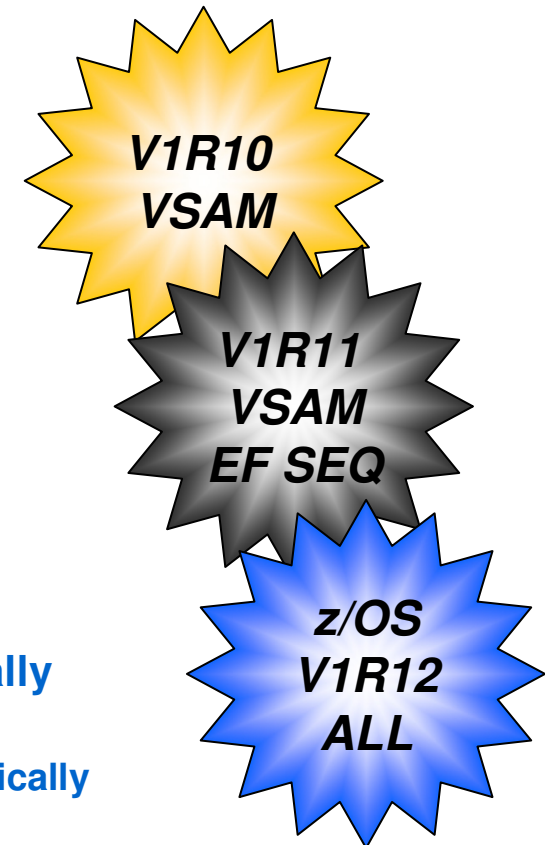
Increases support for System z volume sizes by over 400%!

Value

- Management - Reduces number of volumes to manage
 - Complementary PAV/HyperPAV preserves/enhances performance
- Growth - Removes an inhibitor to growth
 - Store more data on the same number of volumes
- Management - DS8000 automatically converts 3390-9s to the new EAV models when growing a 3390-9 by using DS8000 dynamic volume expansion



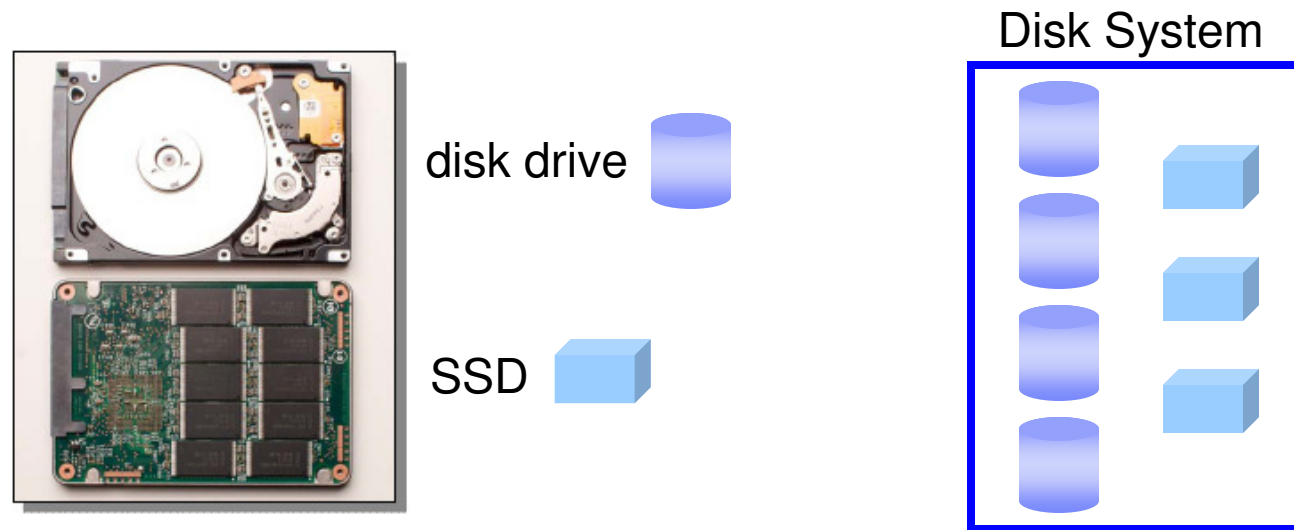
- **EAS Eligible:** A data set on an EAV that is eligible to have extents in the extended addressing space and described by extended attribute DSCBs (format 8/9)
- Can reside in track or cylinder-managed space
- SMS-managed or non-SMS managed
- Any data set type can reside in track-managed space
- Data set types supported
 - **VSAM data types (KSDS, RRDS, ESDS and linear)**
This covers DB2, IMS, CICS, zFS and NFS
CA sizes: 1, 3, 5, 7, 9 and 15 tracks
 - **Sequential (Extended Format)**
 - **Sequential (Basic and Large Format)**
 - **Direct (BDAM)**
 - **Partitioned (PDS, PDSE)**
 - **Catalog (VVDS and BCS)**
 - **All data sets used by DFSMSrmm (journal and dynamically allocated temporary files)**
One exception is the RMM CLIST data set when created automatically by SEARCH subcommand processing.



? Why it Matters: Scalability for customers constrained by 4 character UCB's

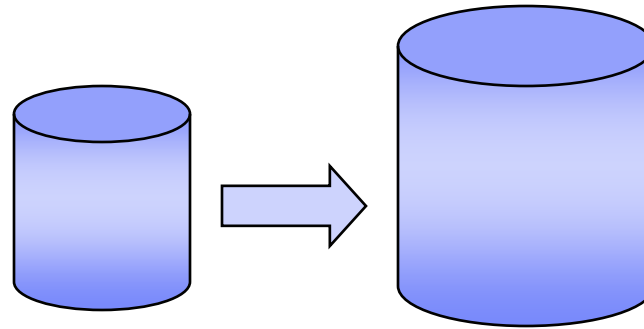
- Description
 - Heritage PAV design designed and implemented for ESCON attachments
 - Relatively static binding of PAV-alias to PAV-base
 - Dynamically Tunable by WLM
 - Complex rules of thumb for allocating the number of PAV-aliases
 - Many times PAVs are over allocated needlessly
 - HyperPAV maintains a shared pool of PAV-aliases
 - PAV-aliases bound to PAV-bases on an I/O request basis
 - Up to 10X increase in efficiency

- Reduce number of required aliases
 - Give back addressable device numbers
 - Use additional addresses to
 - support more base addresses
 - larger capacity devices.
- z/OS can react more quickly to I/O loads
 - React instantaneously to situations like ‘market open’ conditions
- Overhead of managing alias exposures reduced
 - WLM not involved in measuring and moving aliases
 - Alias moves not coordinated throughout Sysplex
- Initialization doesn’t require “static” bindings
 - Static bindings not required after swaps
- IO reduction, no longer need to BIND/UNBIND to manage HyperPAV aliases in the DS8000
- Increases I/O Parallelism



Value

- Management - Helps users allocate new data sets to the desired type of disk
- Management - Supports tools and commands that report on types of disk installed



Value

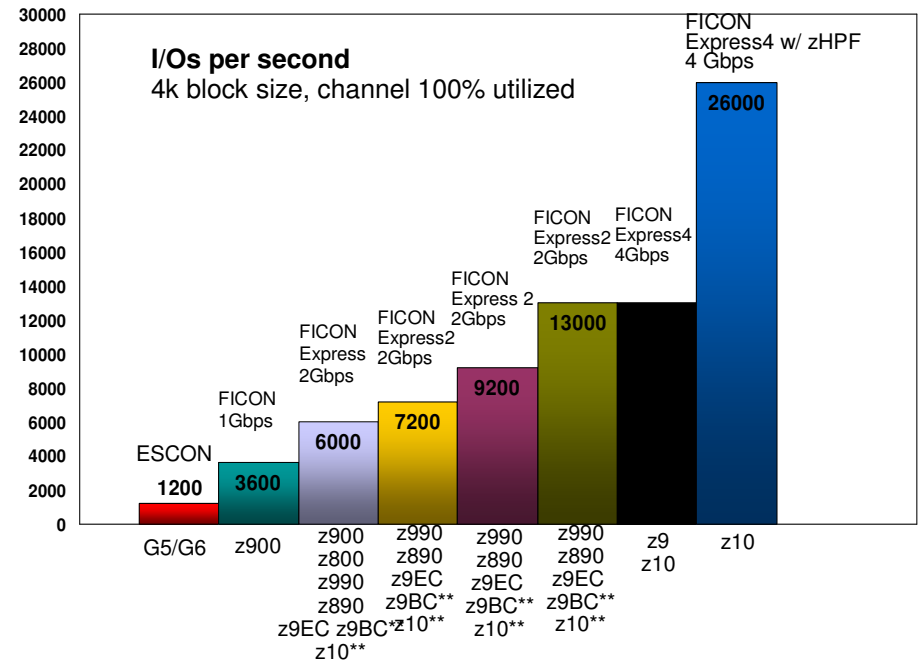
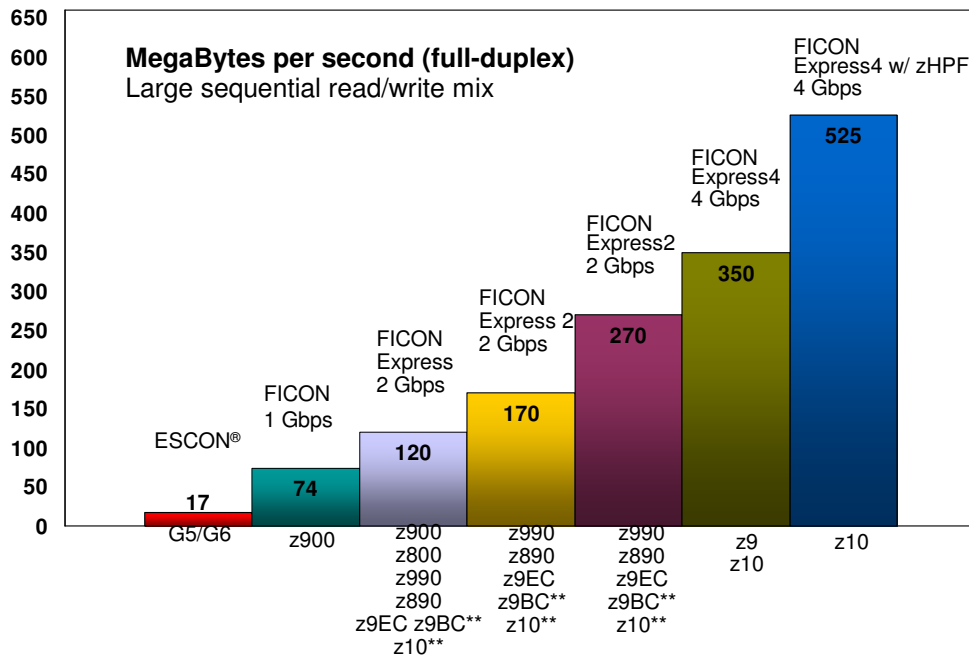
- Availability - Enables increasing volume capacity while data remains online to applications
- Availability - Enables faster response to user requests for more capacity
- Management - Simplifies operations
 - No longer need to backup data, bring volume offline, delete and recreate volume to expand it
 - Can be done on-line with a single command or GUI screen
- Growth/Management - Enables non-disruptive migration from 3390-9 to large EAV 3390s

VALUE

- Performance/Growth - Up to 2x higher I/O throughput per channel
 - Realistic workloads with a mix of data set transfer sizes can see 30 – 70% of FICON I/Os utilizing zHPF, resulting in up to a 10-30% channel utilization savings.

- Management - Can reduce manual tuning

- TCO – May require fewer channels on the z server and on the disk system



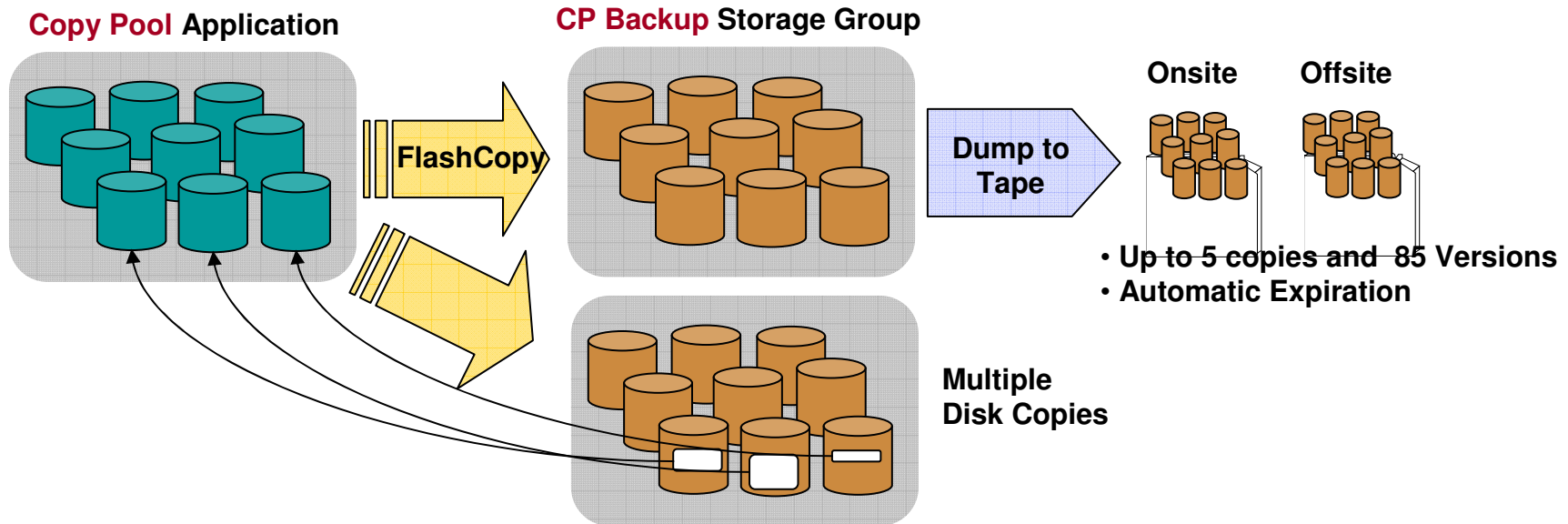
▪ **SMS Data Set Separation by Volume**

- SMS will support the allocation of critical data sets (such as DB2partitions) on different volumes
 - Reduces I/O contention, provides improved performance and helps to avoid single points of failure
 - Designed to expand the existing data set separation function, to allow you to specify that critical SMS-managed data sets be separated across extent pools and volumes that are not used by other data sets specified in the separation group.

? Why it Matters: The system will ensure DSN's that need to be on separate Disk "spindles" will actually be physically allocated that way.

HSM function that manages Point-in-Time copies

- Combined with DB2 BACKUP SYSTEM, provides non-disruptive backup and recovery to any point in time for DB2 databases and subsystems (SAP)



★ Recovery at all levels from either disk or tape!

- Entire copy pool, individual volumes and ...
- Individual data sets

- **TS7680 ProtecTier Deduplication**
 - Virtual Tape Library
 - Inline data deduplication
 - Replication – Only deduplicated data is transmitted

- **DFSMSHsm data**
 - DFSMSHsm ‘wraps’ blocks of native data with meta data, making all blocks unique
 - TS7680 has logic specific for DFSMSHsm blocks so that it can deduplicate them

? Why it Matters: Allows HSM data to be a candidate for deduplication

▪ **SMS Data Set Separation by Volume**

- SMS will support the allocation of critical data sets (such as DB2partitions) on different volumes
 - Reduces I/O contention, provides improved performance and helps to avoid single points of failure
 - Designed to expand the existing data set separation function, to allow you to specify that critical SMS-managed data sets be separated across extent pools and volumes that are not used by other data sets specified in the separation group.

? Why it Matters: The system will ensure DSN's that need to be on separate Disk "spindles" will actually be physically allocated that way.

- **Tape Volume Cache Management**
 - Used to influence the retention of virtual volumes in cache
- **Volume Pooling**
 - Used to group logical volumes on a set of physical volumes
- **Selective Logical Volume Copy**
 - Used to create a duplicate logical volume copy on a different cartridge
- **Cross-site Replication**
 - Used in a GRID configuration to create a copy of a logical volume at a different site
- **Logical Volume Sizes**
 - Used to select larger logical volume sizes (up to 4,000 MB)
- **Secure Data Erase**
 - Limits how long expired data remains on a cartridge before being erased
- **Copy Export**
 - Used to export logical volumes from a standalone or GRID TS7700
- **Logical WORM**
 - Used to provide write-once-read-many support for logical volumes similar to physical WORM

Self-encrypting disks



Value

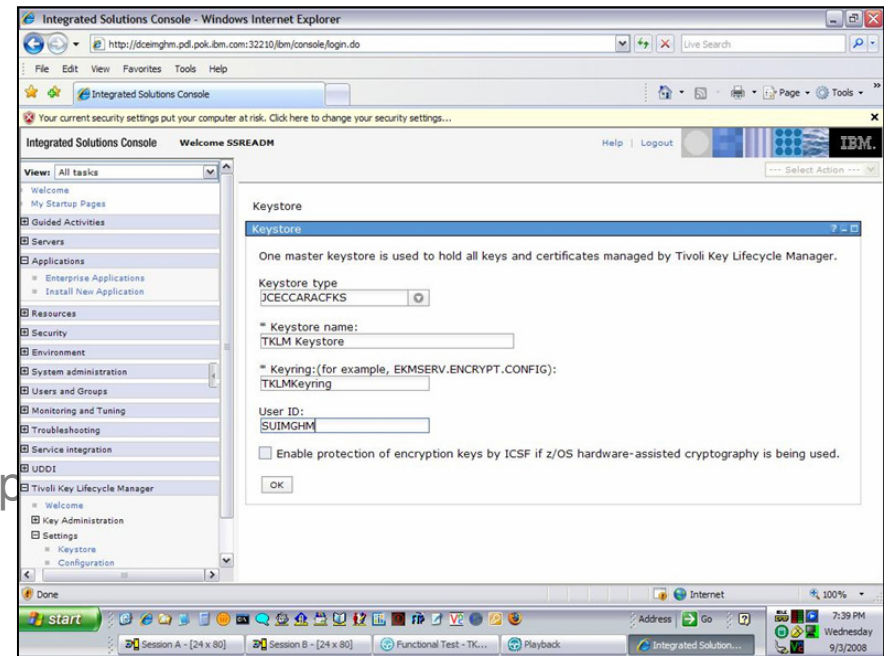
- Management/Security - Protects data on disks removed due to failure, or in a redeployed or retired system
- Performance - No performance impact
- Security - Unique key per drive
- Security - Trusted Computing Group standard
 - Fujitsu, Hitachi GST, Seagate, Samsung, Toshiba, Western Digital, Wave Systems, LSI Corp., ULink Technology, IBM
- Management - Emerging common technology regardless of where disk drives are used (e.g., DAS, SAN, NAS, PCs)
- Management - Supports the principle that encryption should be the last transformation done to maintain efficiency of compression, deduplication

IBM Tivoli Key Lifecycle Manager v.1.0

Feature Function



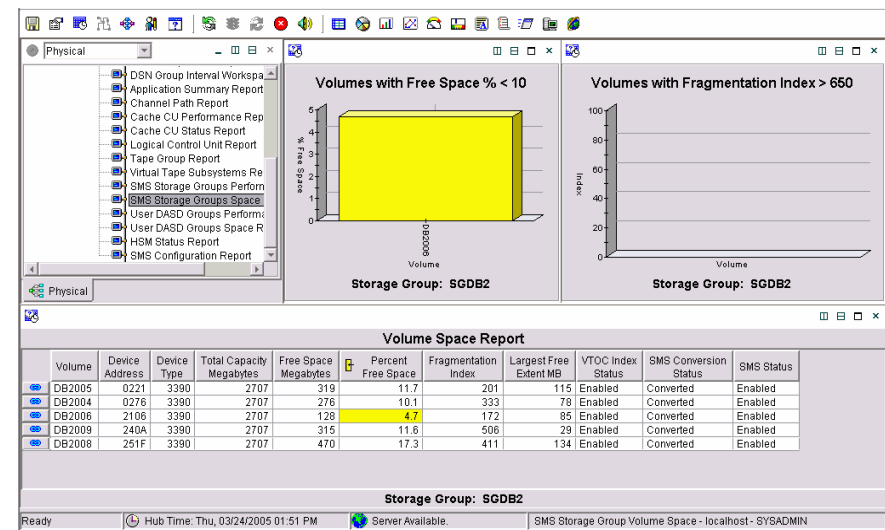
- Focused on device key serving
 - IBM encrypting tape – TS1120, TS1130, LTO gen 4, **LTO gen 5**
 - IBM encrypting disk – DS8000
- Lifecycle functions
 - Notification of certificate expiry
 - Automated rotation of certificates
 - Automated rotation of groups of keys
- Designed to be Easy to use
 - Provide a Graphical User Interface Initial configuration wizards
- Easy backup and restore of TKLM files
 - TKLM backup, DB2 backup, Key backup
 - Simple to clone instances
- Installer to simplify installation experience
 - Simple to use install, can be silent
- Platforms for V1
 - z/OS 1.9, 1.10, 1.11
 - AIX 5.3, 6.1 or later
 - Red Hat Enterprise Linux 4.0 (32 bit)
 - Red Hat Enterprise Linux 5.0 (32 bit and 64 bit)
 - SuSE Linux 9 (32 bit)
 - SuSE Linux 10 (32 bit and 64 bit)
 - Solaris 9, 10 Sparc
 - Windows Server 2003 (32 bit and 64 bit)
 - Windows Server 2008 (32 bit and 64 bit),



Reducing the impact of failures : IBM Tivoli OMEGAMON XE



- A mainframe STORAGE monitor, real-time and historical
- XE user interface, comes with the CUA UI component
- A wide breadth of mainframe storage information:
 - Space (storage groups or user groups ... define your own)
 - Performance (storage groups or user groups ... define your own)
 - Tape / VTS
 - DFSMSrmm support
 - CACHE
 - Channels (FICON)
 - Control Units
 - DFSMSHsm (View your HSM queues, control Datasets, etc.)
 - DFSMSHsm/DFSMSdss/DFSMSrmm/ICKDSF online toolkit
 - SMS constructs
 - DS8000 support
 - Ability to see all logical volumes on a physical disk
 - Powerful applications view
 - Powerful dataset view and action capability
 - Integration capabilities from TEP interface
 - Submit JCL in response to monitored situations



Tivoli Enterprise Portal (TEP) Enabled

- The z/OSMF DASD Management task is designed to simplify SMS storage management. The first stage focuses on the task of adding storage capacity to an SMS Pool storage group through a single user interface.

- With the DASD Management task, the storage administrator can perform the following:
 - Define new pool storage group z/OSMF DASD Management attributes (referred to as SMA) to be used as policy with this task
 - Define, display and update these attributes via the z/OSMF DASD Management task, ISMF and Navquest
 - View these attributes through DCOLLECT record type 'SG'
 - View storage group and volume information associated with the active SMS configuration
 - Add storage to an SMS pool storage group via a wizard that guides the user through steps that will simplify the task through a single UI

"All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only."

- Many batch jobs will include intermediate backup steps for the working set data – these steps can add significant wall clock time to the batch job – especially if the backups are to stand alone tape drives
- Using Dataset Level FlashCopy can significantly speed up the wall clock time by making those intermediate backup steps “painless”.
- Analyze the JCL of product jobs who take a long time in wall clock time to complete. It may be worth it to change the JCL in those jobs.

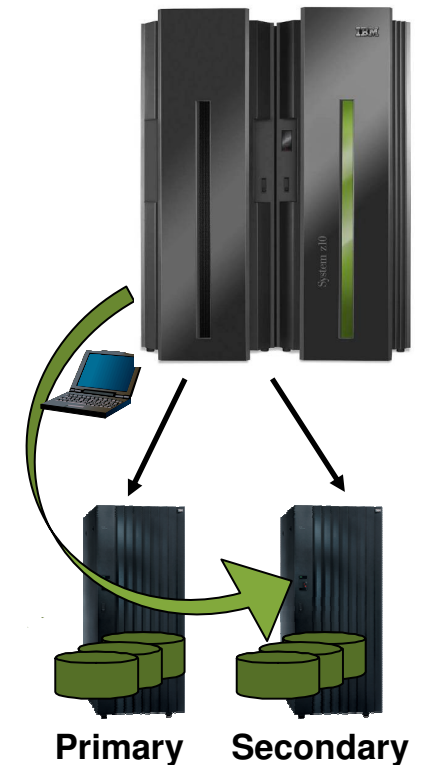
IBM Basic HyperSwap



Higher, cost-effective information availability for System z

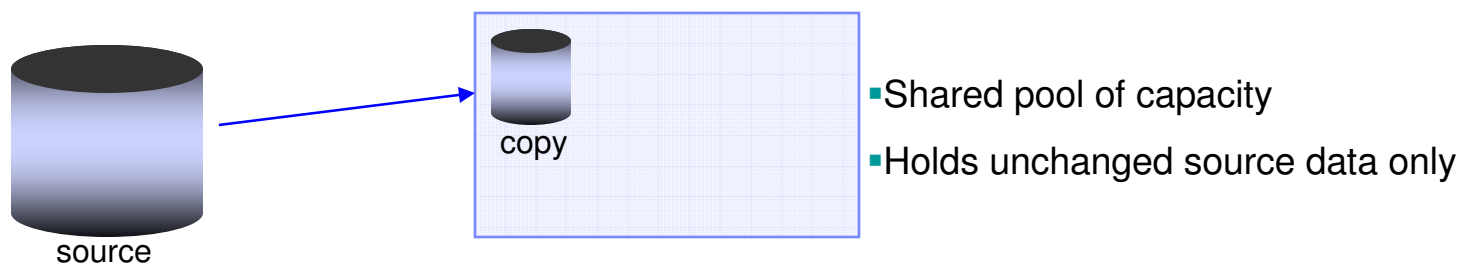
VALUE

- Availability - “Hides” disk system outages from applications
- TCO - A standard z/OS capability for single-site configurations, when full GDPS function is not required
- Management - Managed by Tivoli Storage Productivity Center for Replication (TPC-R) for System z, available at no extra charge with z/OS



IBM FlashCopy SE (Space-Efficient)

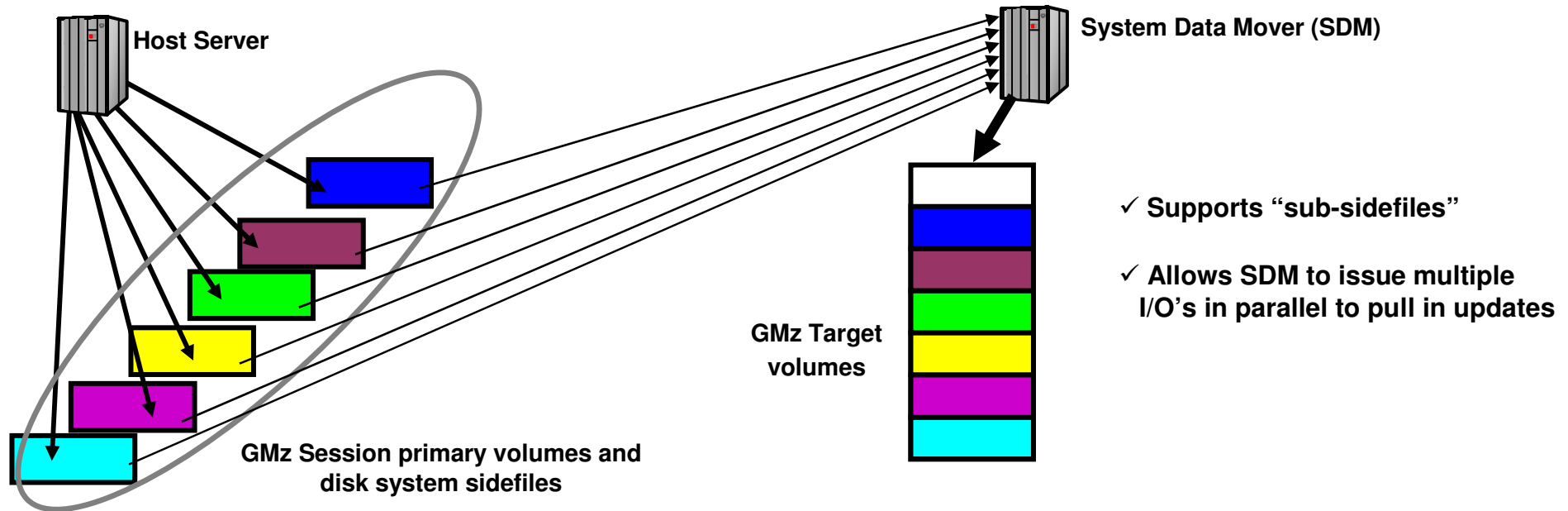
A technology optimization for short-term point-in-time copies



Value

- Growth/TCO - Reduces physical capacity needed to hold volume copies
 - Target volume capacity savings can be up to 80%
- Growth/TCO - Helps reduce energy and floor space requirements

z/OS Global Mirror (GMz) Multiple Readers for DS8000



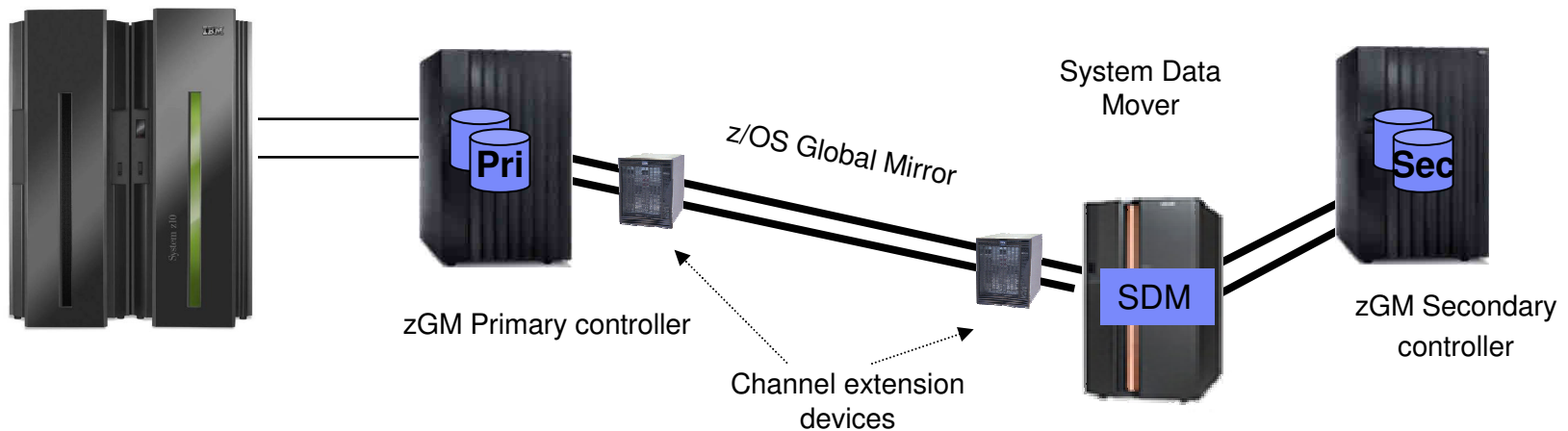
Value

- Performance/management - Increased DS8000 z/OS Global Mirror throughput
 - Increased parallelism for handling write I/Os
 - Better able to handle peak loads
 - Avoids host slowdown that can result when the disk system cannot keep up with writes
- Availability - Helps increase recovery site data currency, or maintain same currency at higher I/O loads

z/OS Global Mirror with DS8000 Extended Distance FICON



A long-distance remote mirroring solution optimized for System z

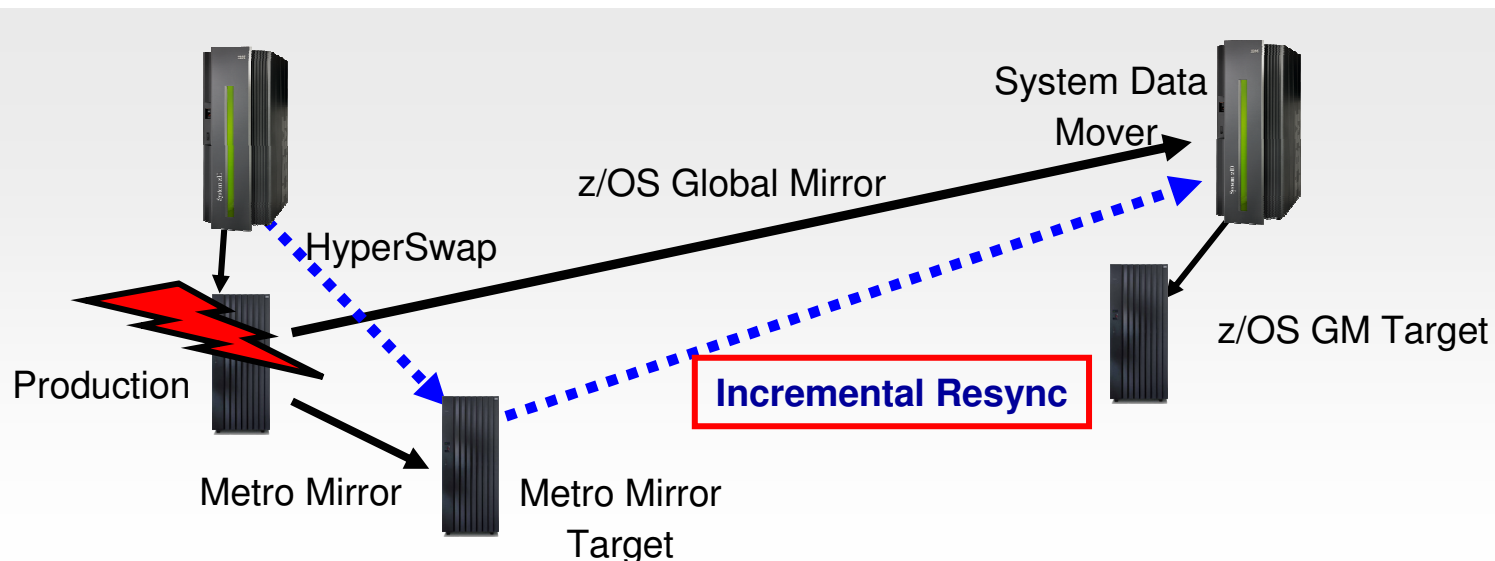


Value

- Management/Availability - Enables increased distance over channel connections
- TCO - Helps reduce costs of a z/OS Global Mirror configuration
 - Increases link efficiency (i.e., supports higher link utilization)
 - Can eliminate the need for extra channel extender devices

z/OS Metro/Global Mirror Incremental Resync

For sophisticated 3-site Metro / z/OS Global Mirror (MzGM) environments



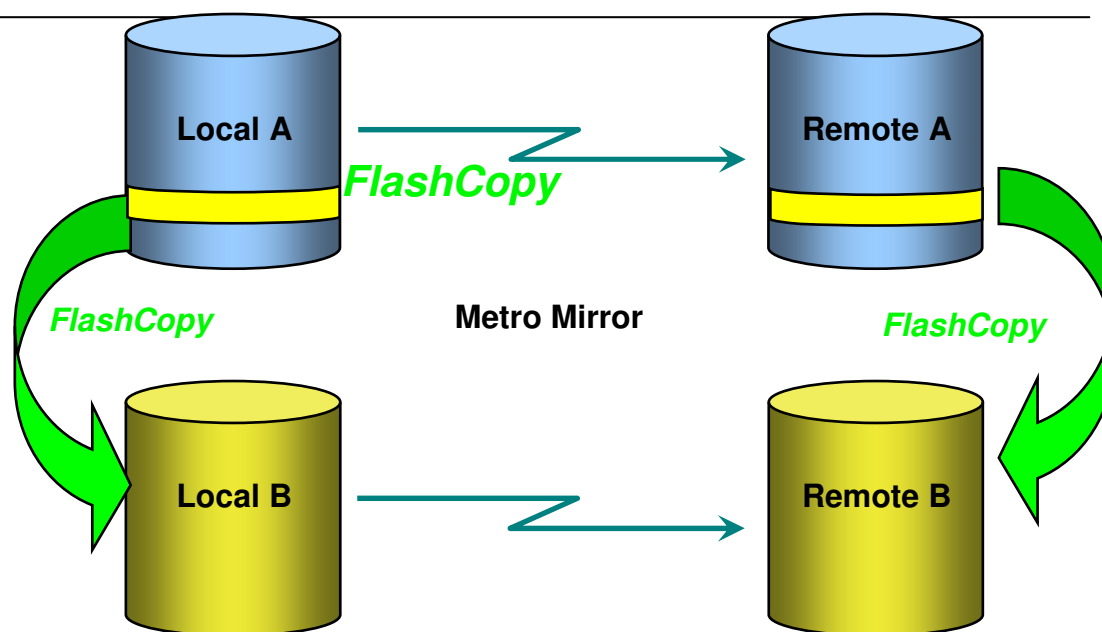
Value

- Availability - Improves remote site data currency following local site disk system outage
- Performance/TCO - Reduces use of link bandwidth
 - Only transmits changes from Metro Mirror Target to z/OS Global Mirror target after HyperSwap, instead of entire volumes
 - Can reduce resynchronization time to minutes from hours!

Remote Pair FlashCopy

Synchronous Remote Mirroring Optimization

- For z/OS data sets
- For volumes of all servers



Value

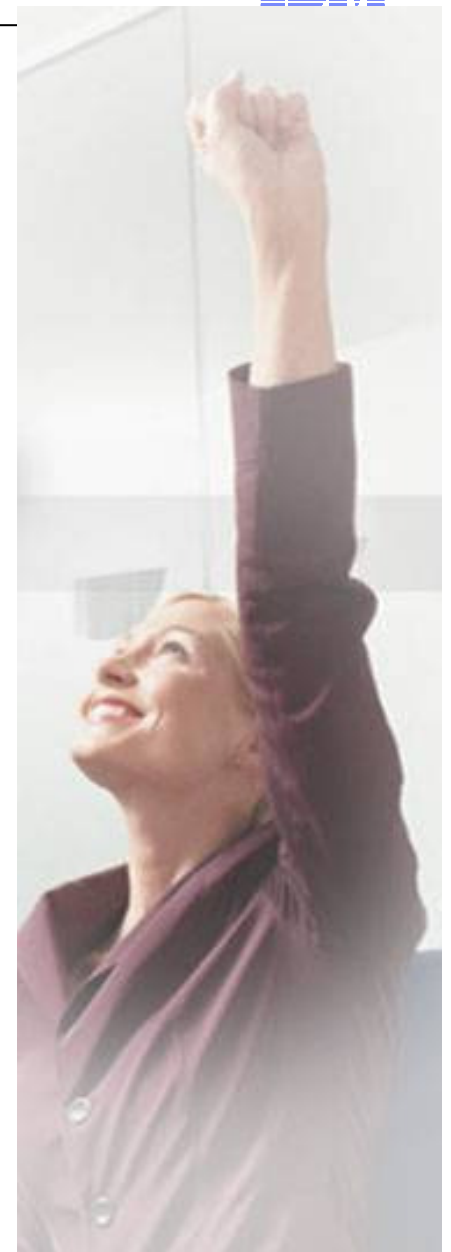
- Availability – Remote site quickly reflects same data as the local site
- Management - Promotes local-remote data consistency for ease-of-recovery
- Performance/TCO - Reduces link bandwidth use, helps improve the performance of other work

Reducing the impact of failures: TPC for Replication Family



- ▶ TPC for Replication is software running on a server that manages IBM disk based replication functions, i.e.:
 - ▶ FlashCopy
 - ▶ Metro Mirror
 - ▶ Global Mirror
 - ▶ Metro Global Mirror

- ▶ TPC for Replication
 - ▶ Provides central control of your replication environment
 - ▶ Helps simplify and automate complex replication tasks without scripts
 - ▶ Allows testing of D/R at any time using Practice Volumes
 - ▶ Provides end to end management and tracking of the copy services
 - ▶ Provides management of planned and unplanned disaster recovery procedures
 - ▶ Manage your replication from the server of your choice, z/OS or open systems servers, z/OS and open data



IBM Corporation 2011

- IBM, the IBM logo and ibm.com are registered trademarks, and other company, product or service names may be trademarks or service marks of International Business Machines Corporation in the United States, other countries, or both. A current list of IBM trademarks is available on the Web at “[Copyright and trademark information](http://www.ibm.com/legal/copytrade.shtml)” at www.ibm.com/legal/copytrade.shtml
- Adobe, the Adobe logo, PostScript, the PostScript logo, Cell Broadband Engine, Intel, the Intel logo, Intel Inside, the Intel Inside logo, Intel Centrino, the Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, IT Infrastructure Library, ITIL, Java and all Java-based trademarks, Linux, Microsoft, Windows, Windows NT, the Windows logo, and UNIX are trademarks or service marks of others as described under “Special attributions” at: <http://www.ibm.com/legal/copytrade.shtml#section-special>
- 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.