IBM Technical Forum & Executive Briefing

17 al 21 Octubre 2011

Imagine PODER Imagine CAPACIDAD

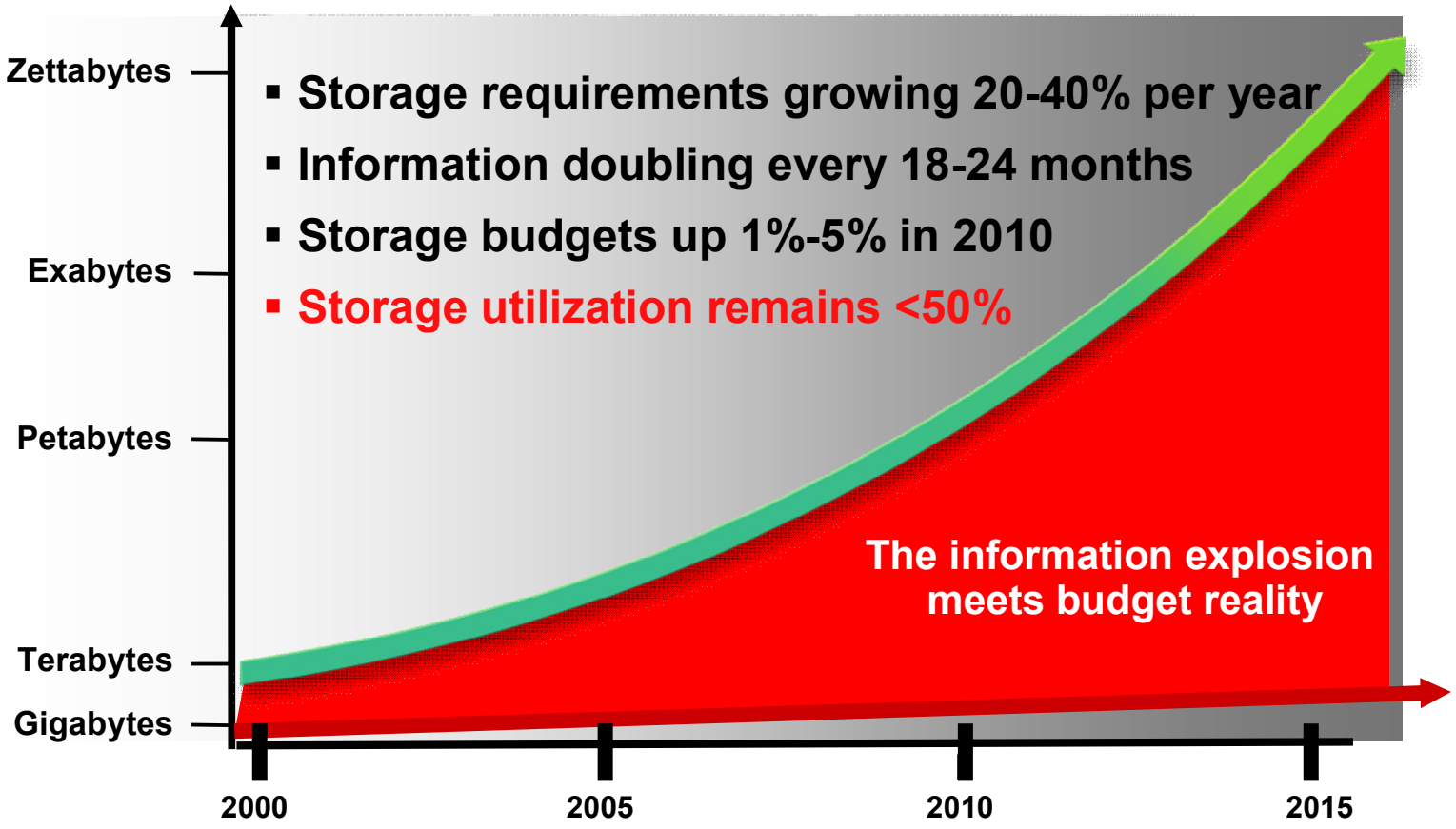
IBM System Storage DS8800

and Business Continuity

for the World's Most Demanding Customers

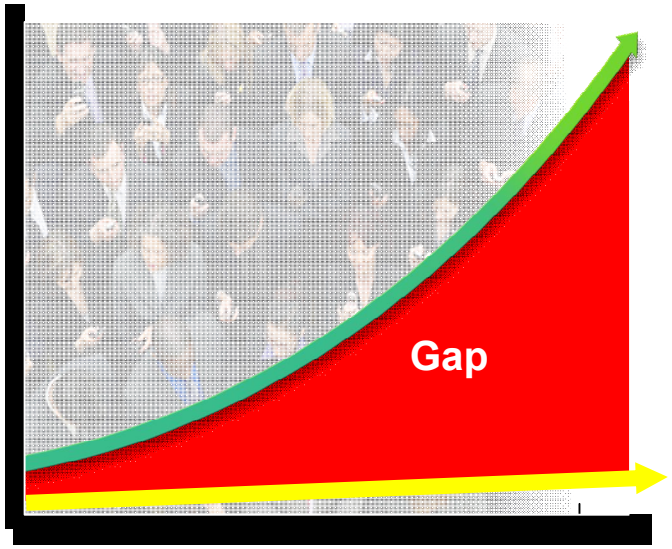
October 2011

Smarter systems are creating an information explosion



Instrumented.
Interconnected.
Intelligent.

"Just Buy More Storage"



- ***Proliferation of disk systems***, to keep up with performance/capacity growth
- Great deal of ***allocated but unused space***, to avoid provisioning delays
- Many ***unnecessary copies*** of data increase operating expenses
- Most ***data hasn't been accessed*** in months, but may be on top tier storage

IBM System Storage DS8000 Series

Enterprise Disk for the World's Most Demanding Clients

Built on 60 Years of Enterprise Class Innovation

- IBM's Flagship Enterprise Storage Device
- Strong Synergy with IBM Servers (z, i, p)

Over 15,000+ DS8K systems sold worldwide!!!

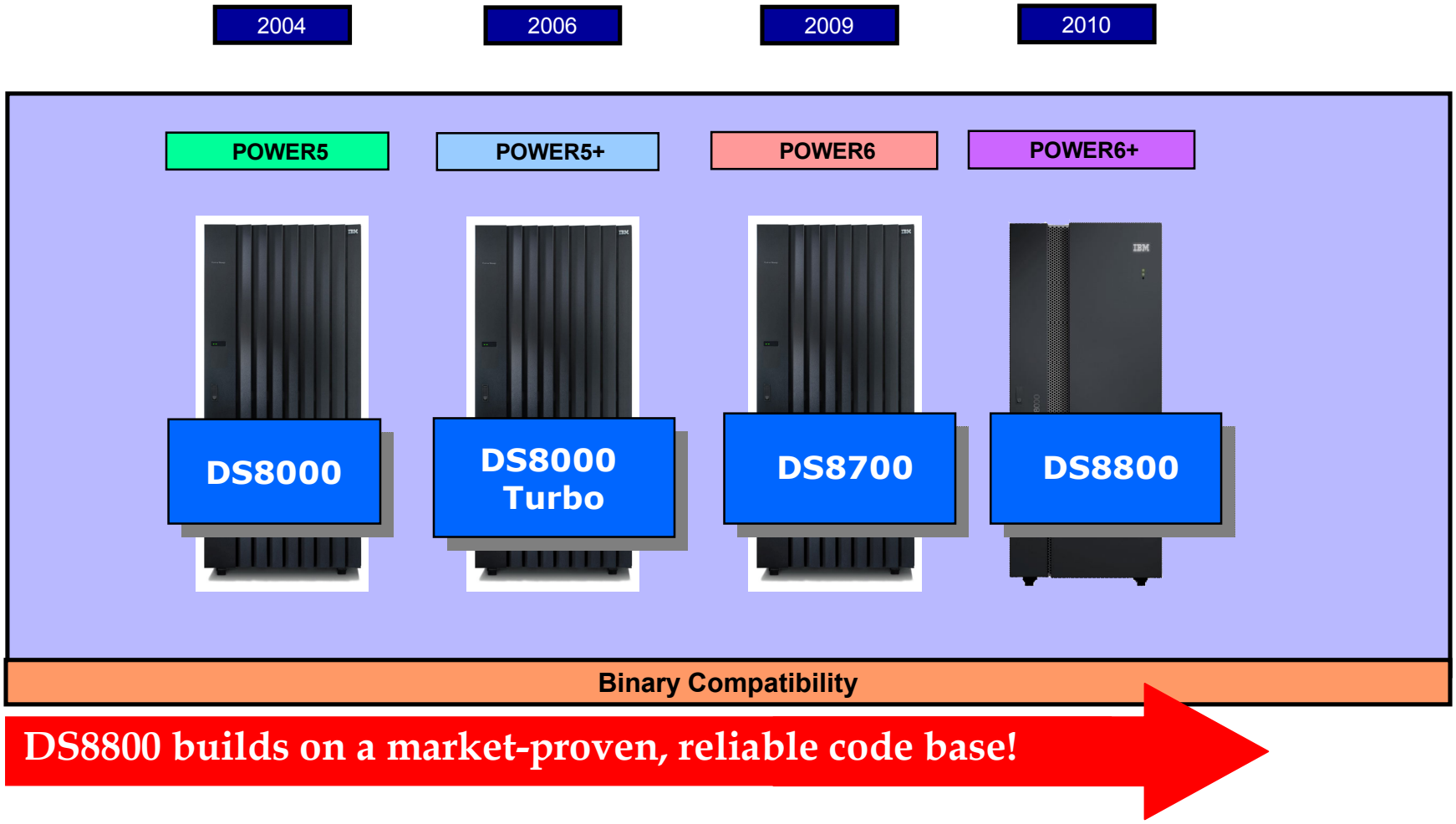
Over 1,000+ DS8K system sold in LA

Performance, Resiliency, and Flexibility to Satisfy the World's Most Demanding Clients

- ***Performance*** – Architected for highest total throughput
- ***Availability*** – Designed for 24X7 Environments
- ***Resiliency*** – Outstanding Copy and Mirroring Capability
- ***Flexibility*** – High Performance, Online & High Capacity, Nearline Disk options to satisfy tiered storage objectives
- ***Storage efficiency*** – Up to 2.3 PB storage consolidation and Easy Tier for storage optimization
- ***Heterogeneous Server Support*** - *IBM z/OS, z/VM, OS/400, i5/OS, AIX, Linux, HP-UX, Sun SOLARIS, Novell, KVM, VMware and Microsoft, among others*
- ***Security*** – *Self-encrypting Disk Drives*
- ***Long-Term Cost Advantage*** – Enterprise Choice Warranty

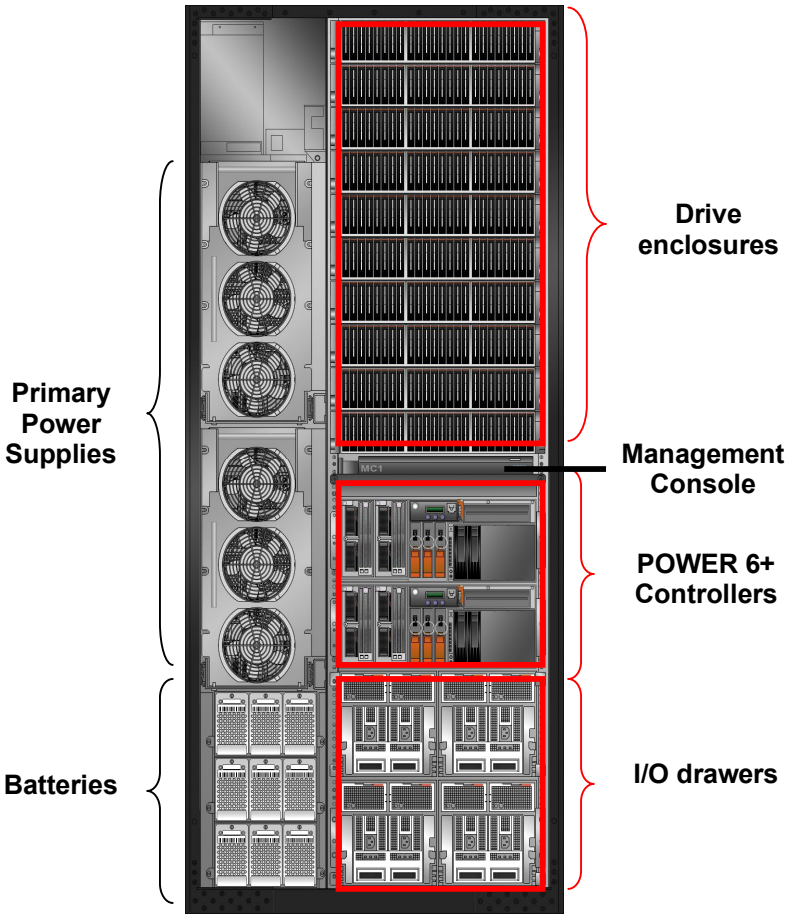
4th Generation DS8000 enterprise disk system

The IBM POWER processor has been behind the success of IBM enterprise storage beginning with the Enterprise Storage Server in 1999



DS8800 Hardware Upgrades

- Compact and highly efficiency drive enclosures
  - 2.5” small-form-factor drives
  - 6 Gb/s SAS (SAS-2)
  - Enclosures support 50% more drives
- Upgraded processor complexes
  - IBM POWER6+ for faster performance
- Upgraded I/O adapters
  - 8 Gb/s host adapters
  - 8 Gb/s device adapters
- More efficient Front-to-Back airflow cooling

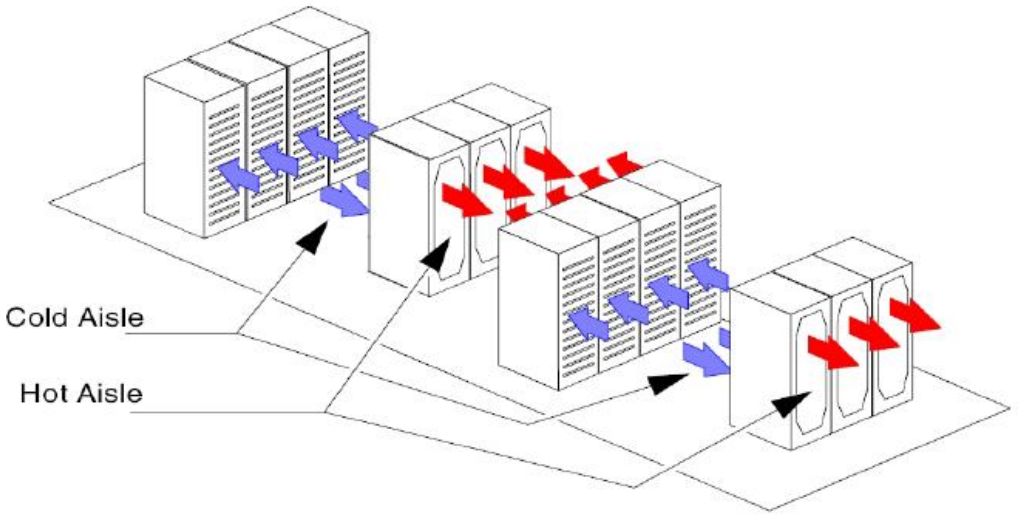


Front view

Dramatic efficiency and performance benefits

New airflow design is also more energy efficient

- More data centers are moving to hot aisle / cold aisle designs to optimize energy efficiency
- DS8800 is now designed with complete front-to-back airflow



Benefit: Greater energy efficiency and contributes to lower energy costs

DS8800 Power Improvements

- Table below takes into account controller card power, power efficiencies, power for cooling, and power for disks.

	DS8700 3.5” drives	DS8800 2.5” drives
Power per Disk	18.4 Watts	10.2 Watts
Power per Enclosure	310 Watts	245 Watts

Energy consumption comparison

DS8800 with 1056 drives

- Base frame: 6.8kW
- Exp frame: 5.4kW
- Exp frame: 6.5kW

TOTAL: 18.7kW

DS8700 with 1024 drives

- Base frame: 6.8kW
- Exp frame: 7.1kW
- Exp frame: 6.1kW
- Exp frame: 6.1kW
- Exp frame: 3.0kW

TOTAL: 29.1kW

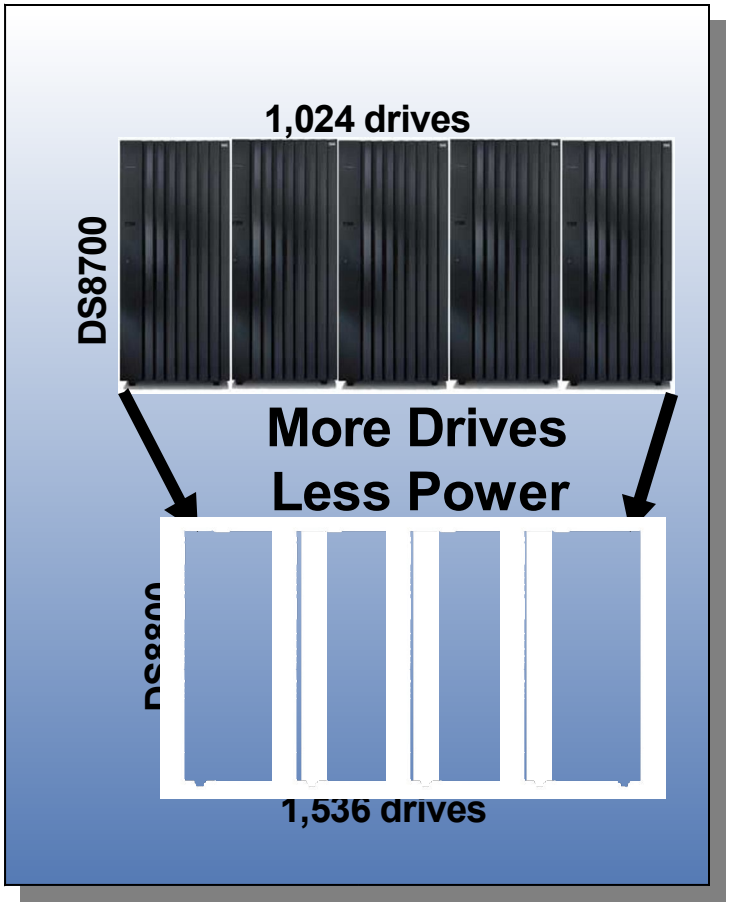
3-year cost savings

- 36% less energy usage; 40% less floor space
- $41,336 less for power/cooling (KW = $ 0.147)
- $71,624 less for floor space

Saves $112,960

Storage efficiency with space-saving design

- Space-saving design
  - Small-form-factor drives
  - High-density drive enclosures
  - Almost double the drives in same frame footprint
- Benefits
  - More effective consolidation can lower operating costs
    - Support more workloads with smaller footprint
    - Reduce number of systems to manage
    - Reduce power and cooling costs



• ~ 90% more drives in a single frame
• 3-Frame: More drives in 60% floor space fully-configured
•4-Frame: 50% more drives in 80% of the floor space

Substantial footprint reduction

I/O Priority Manager

Automated quality of service management delivers performance when and where it's needed and enables greater storage efficiency and consolidation

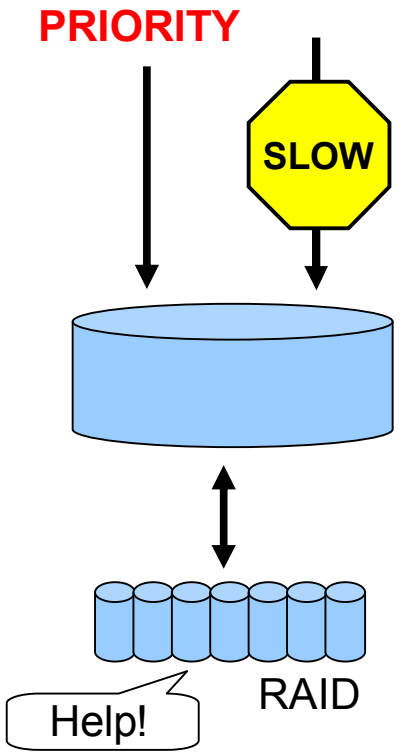
- **I/O Priority Manager aligns service levels to separate workloads in the system**
  - Administrators select from 4 Performance Groups (service levels) to assign to each volume
    - '1' for highest;
    - '2' for standard;
    - '3' for low priority;
    - '0' for no priority (default)
  - All volumes are associated with a Performance Group and all I/Os are monitored
  - System resources are dynamically allocated to higher priority volumes (applications) when there is resource contention

- **How I/O Priority Manager works**
  - **I/O Priority Manager delays the right amount of I/O from lower priority volumes**, so higher priority volumes get more throughput
  - Automatic and only when there is contention for a resource between multiple volumes

Performance when and where you need it

I/O Priority Manager

- Support CKD I/O priority with software input
  - User assigns a performance policy to each CKD volume that applies in the absence of additional software support
  - OS can optionally specify parameters that determine priority of each I/O operation in Prefix command
    - Allows multiple workloads on a single CKD volume to have different priorities **(priority at Data Set level**)
  - In general, should expect that CKD priority is managed by z/OS
- Supported on z/OS V1R11, V1R12, V1R13, and above..



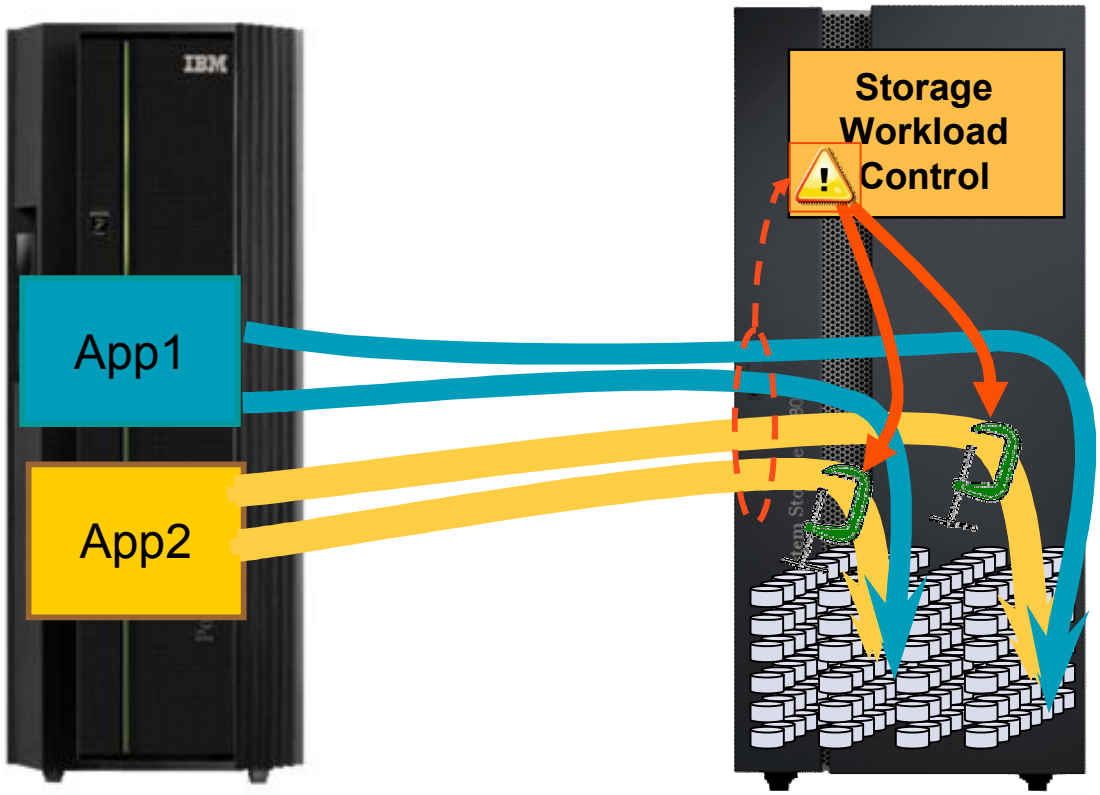
Disruption Workloads on DS8000

- By itself, critical App1 succeeds
- Non-critical App2 begins
- Critical App1 slows due to DS8000 contention

Application host *DS8000*

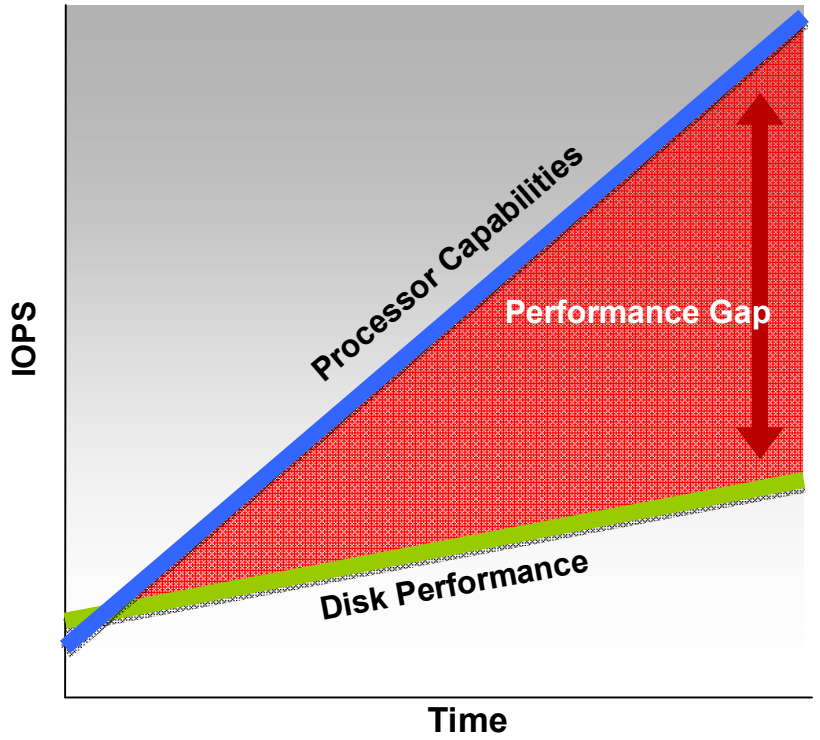
Automatic control of disruptive workload

- I/O Priority Manager detects critical App1 QoS impact
- App2 usage of critical DS8000 resources is controlled
- Controls on App2 restore App1 performance



Performance continues constrained

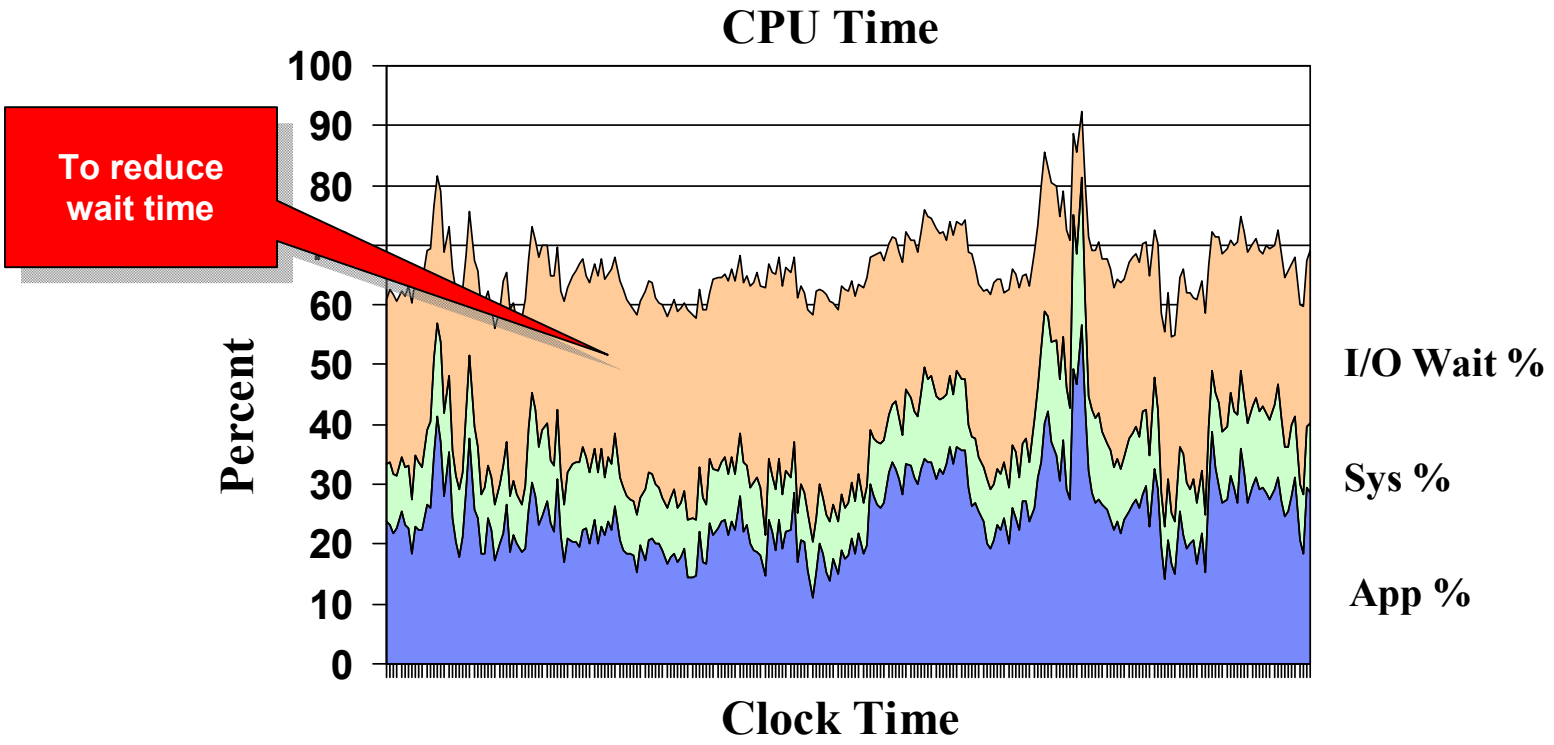
- Processor capabilities are out-stripping disk drive and RAID controller performance (rotational speed and IOPS)
- Servers and storage systems become more unbalanced between CPU/controller capability and storage performance
- Clients add more drive spindles to improve performance



Performance gains through HDDs has become ineffective and wasteful

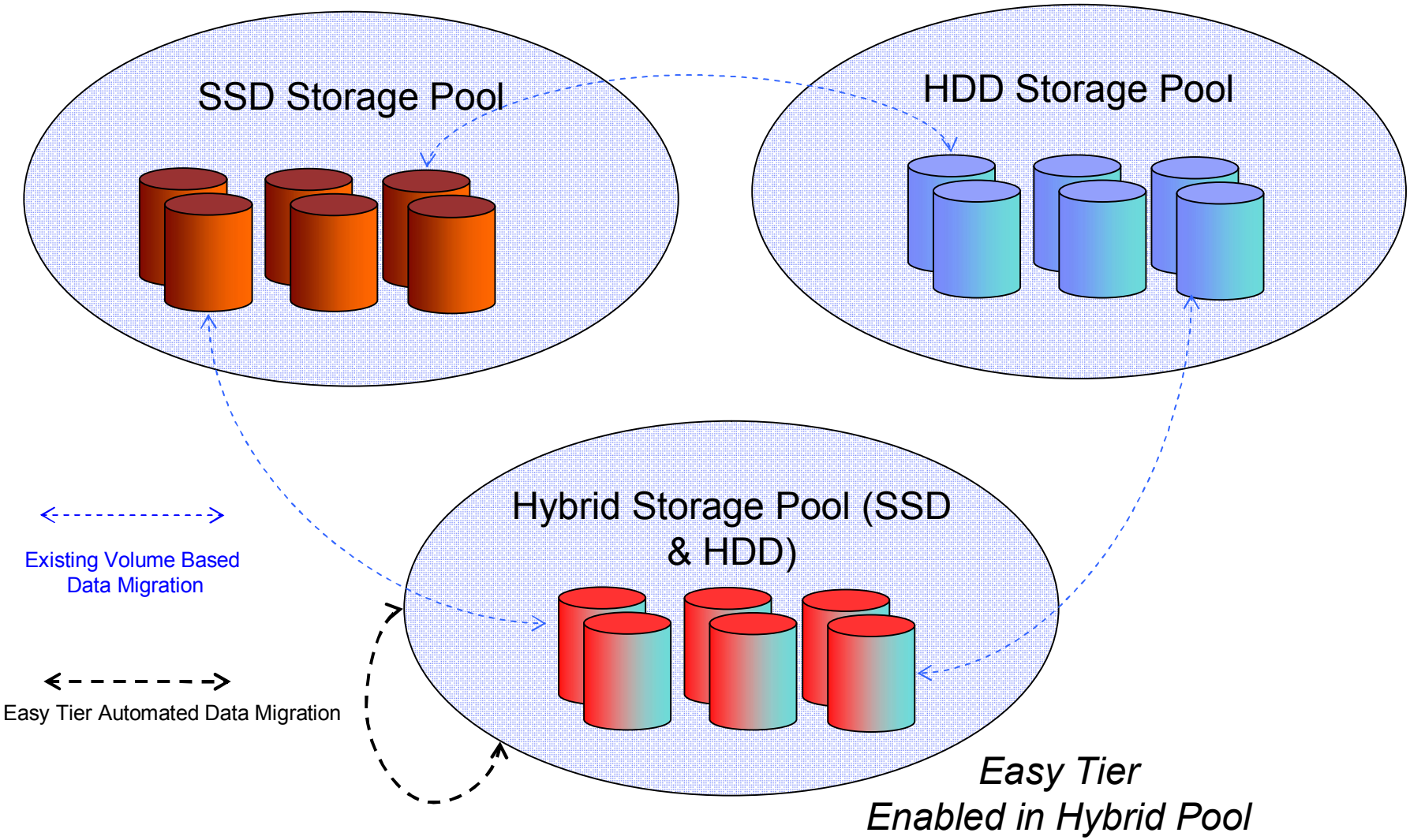
How to reduce IO wait time?

- Database Example – Use of Rotating Disk Drives
  - Reducing I/O wait time can allow for higher server utilization

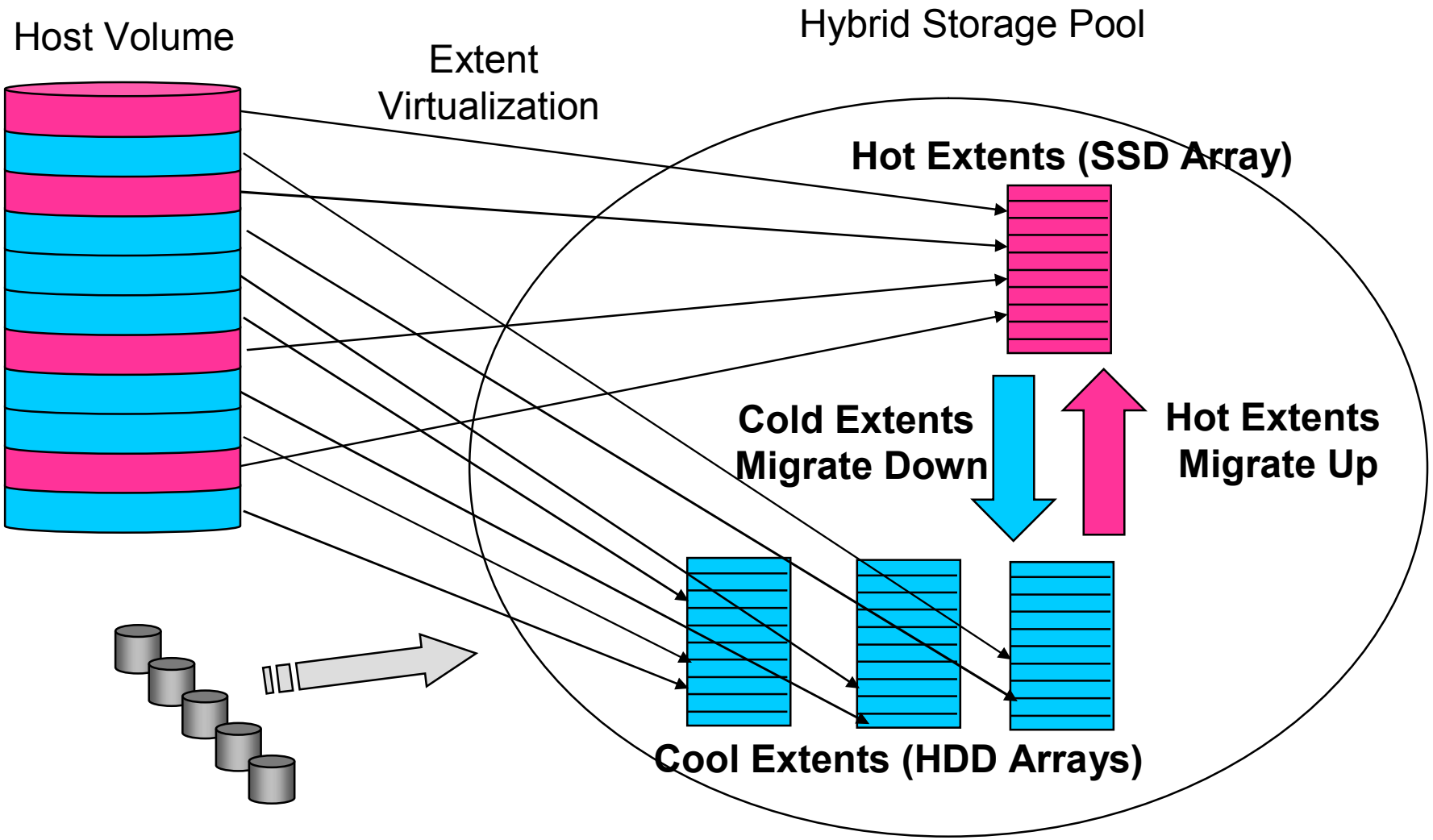


Even well-tuned databases have the opportunity to improve performance and reduce hardware resources

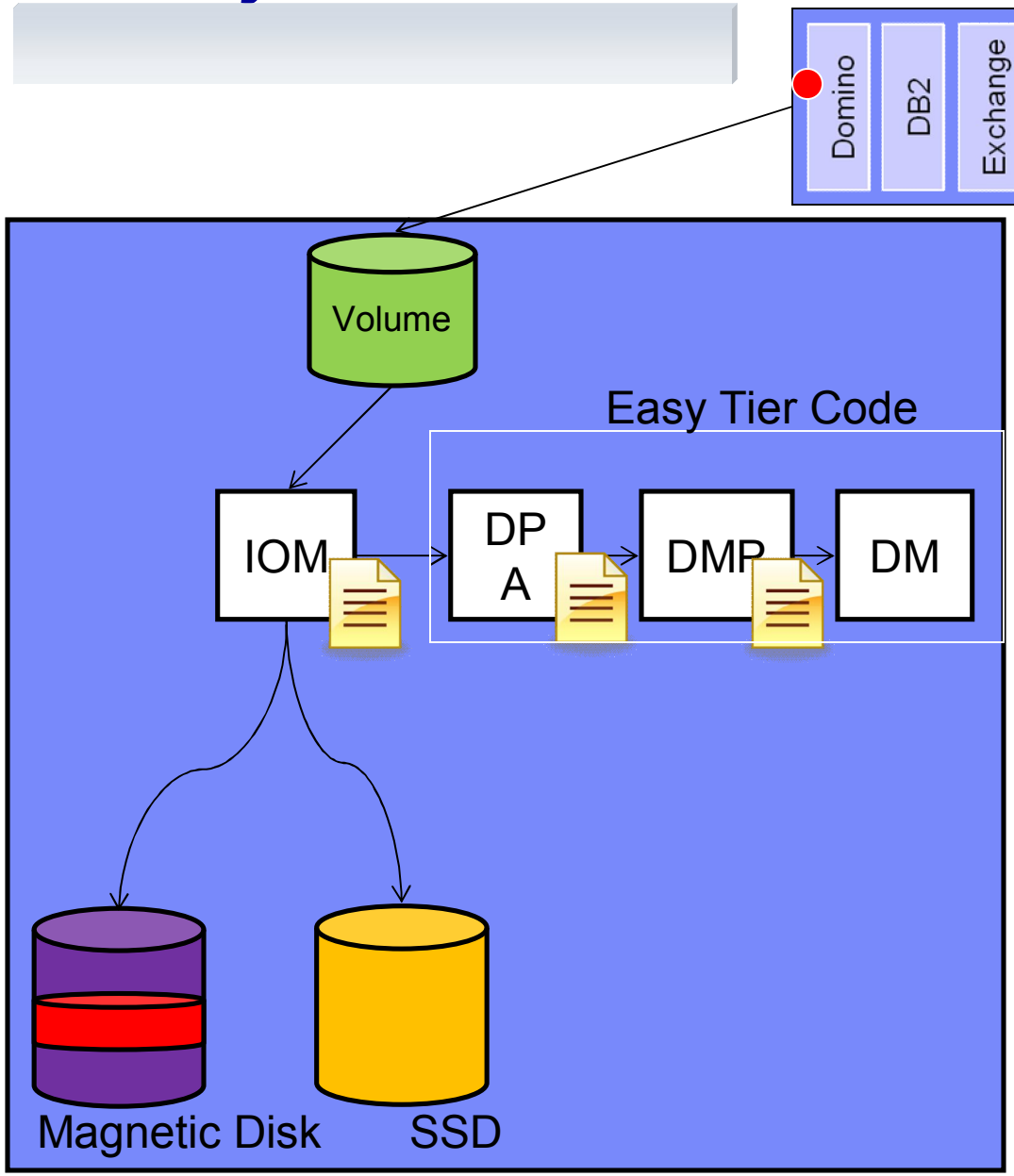
Storage Efficiency Dynamic Data Migration



Storage Efficiency IBM Easy Tier



IBM Dynamic Workload Learning



An application makes frequent access to the same area or extent of a volume in the Magnetic Disk

The **IOM (I/O Monitor)** captures access patterns and generates usage statistics Send to the **DPA (Data Placement Advisor**)

The **DPA** identifies hot extents with high IO latencies for data migrations to the **DMP** (**Data Migration Planner)**

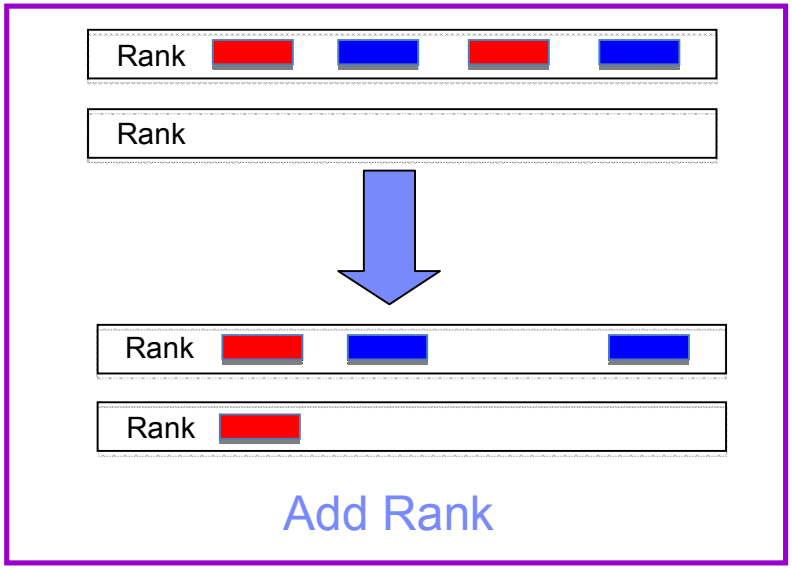
The **DMP** performs analysis and deliver migration plan to the **DM (Data Migrator**)

DM (**Data Migrator)** confirms and schedules data migration activity relocating the data to higher performing storage without any application interruption

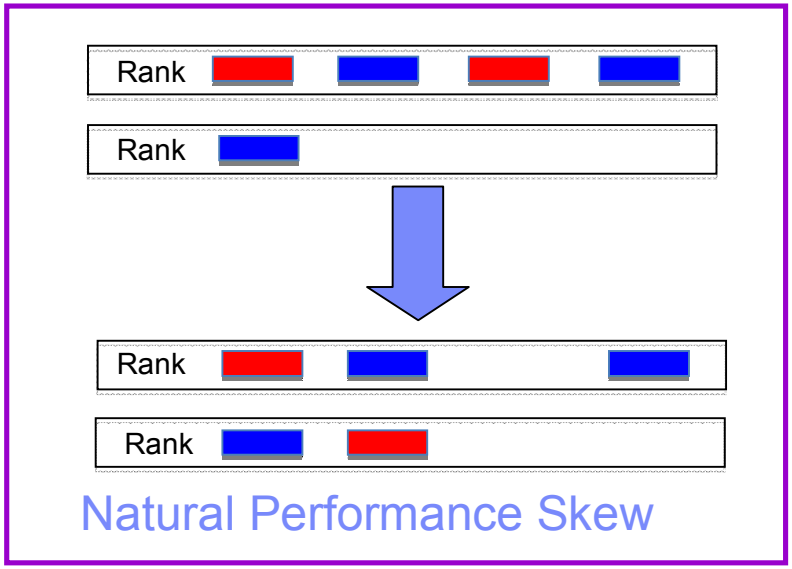
IBM Auto Rebalance Performance

1. Support both **homogeneous pool** and hybrid pool.
2. Automatically maximize storage performance primarily for IOPS. This includes all I/O types so ranks are also secondarily optimized for bandwidth.
3. Auto Performance Rebalance on any storage tier
   a) After new resource is added to or removed from the storage pool.
   b) Natural Performance Skew : Caused content of some extent has more access of the other extent.

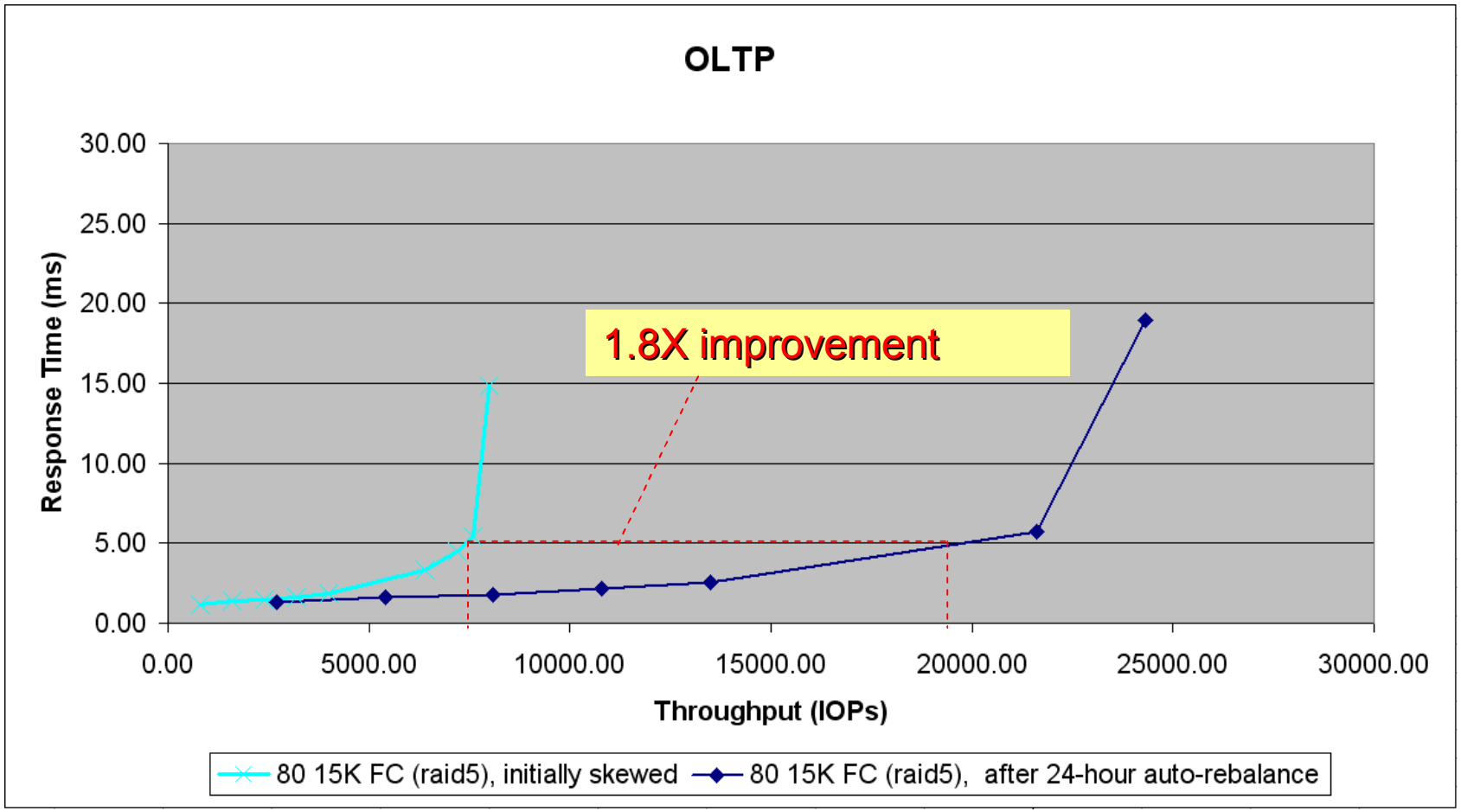
a



b



IBM Auto Rebalance Performance



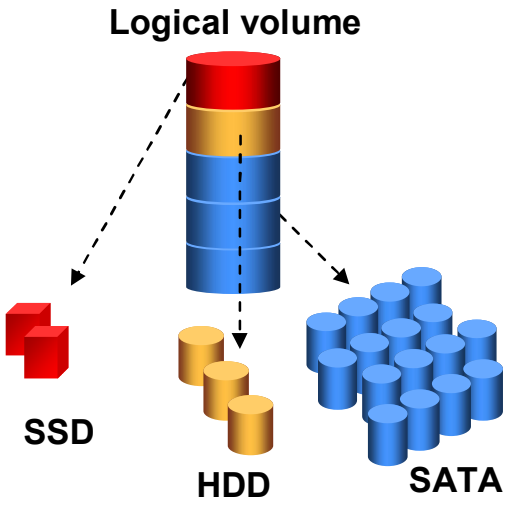
IBM Easy Tier - Smart data placement

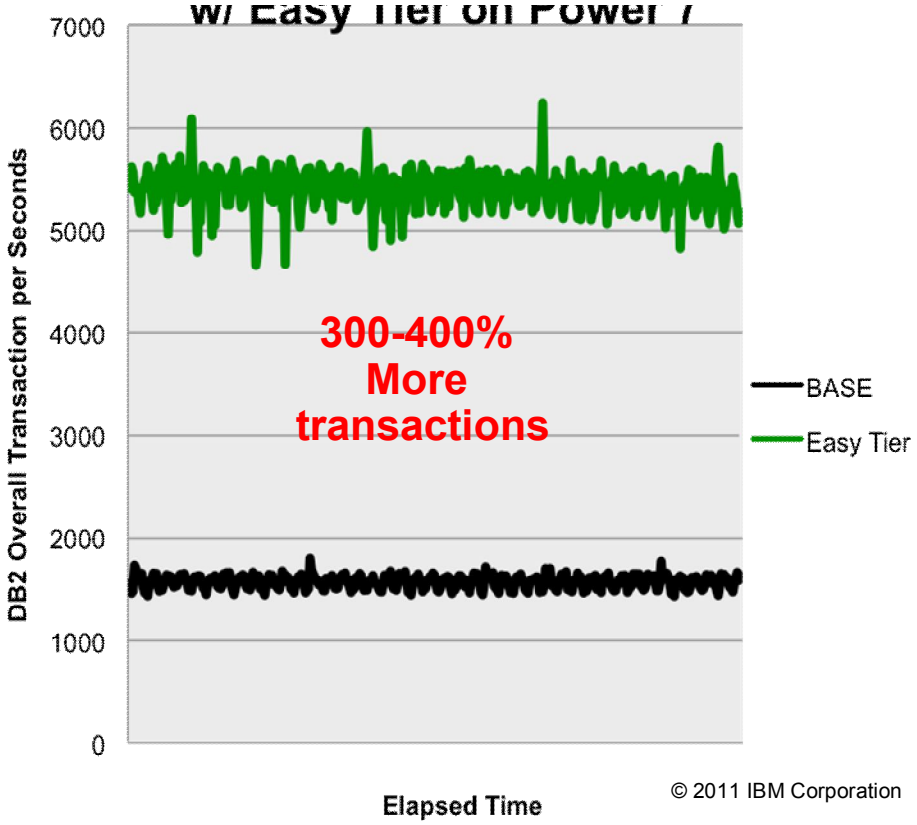
- Today's multi-core, ex. IBM POWER 7 massive thread processors demand fast access to data
- Moving only 10% of data to SSDs results:

400 % Transactions Increase

70% RT Reduction

Easy Tier





Whitepaper on database benchmark with **POWER 7, DS8700** and **DB2.**

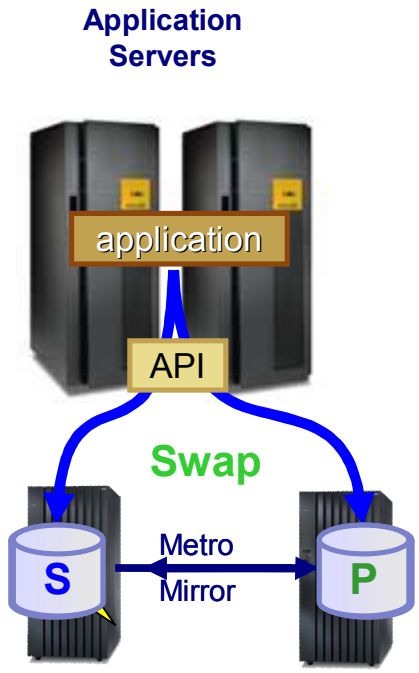
System configuration (base):
POWER7 (Model 9117-MMB)
DB2 9.7 FP1
DS8700 w/ 4.6 TB SSD + 38.4 TB 15K FC HDD)

Example of Performance w/o and with IBM Easy Tier

Without Easy Tier	With Easy Tier
Total IOs/Sec = 10,000	Total IOs/Sec = 10,000
R/W ratio = 3:1	R/W ratio = 3:1
Total Read IOs/Sec = 7500	Total Read IOs/Sec = 7500
Total Write IOs/Sec = 2500	Total Write IOs/Sec = 2500
Write Cache Hit = 100%	Write Cache Hit = 100%
Read Cache Hit = 50%	Read Cache Hit = 50%
Cache RT = 290 us	Cache RT = 290 us
SSD RT = 838 us	SSD RT = 838 us
HDD Avg RT = 5930 us	HDD Avg RT = 5930 us
Avg RT =(2500+3750) x 290 + 3750 x 5930 / 10.000 = **2,4 ms**	**Avg I RT =**(2500+3750) x 290 + 3375 x 838 + 375 x 5930 /10.000 **= 0,52 ms**

High Availability with TPC for Replication V4.2

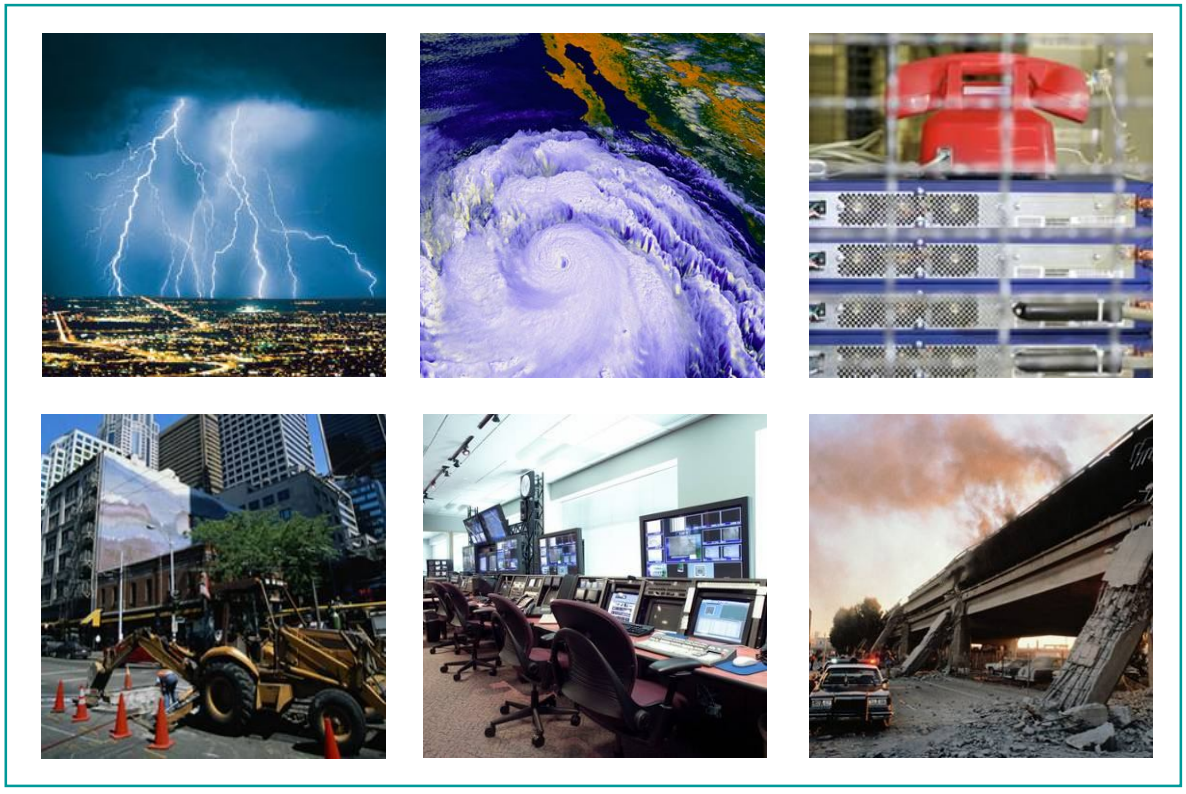
- **Ability to swap IBM DS8000 volumes in seconds**
  - Can be command driven or can be automated upon a storage system failure
  - Designed to scale to multi-thousands of volumes
- **Switches *Metro Mirror* primary storage system to the secondary storage system**
  - No operator interaction is needed for event driven operation
  - Function is configured and managed by TPC-R
- **Feature is non-disruptive**
  - Applications keep using same device addresses
- **Integration with AIX 5.3 (or later) provides higher availability for AIX environments**



Available on TPC-R Windows, AIX, VMWare, Linux and z/OS installations with the TPC-R 2 Site BC License

Importance of IT Business Continuity

Regulations, security threats, and service outages — Uptime for IT business is increasingly essential.



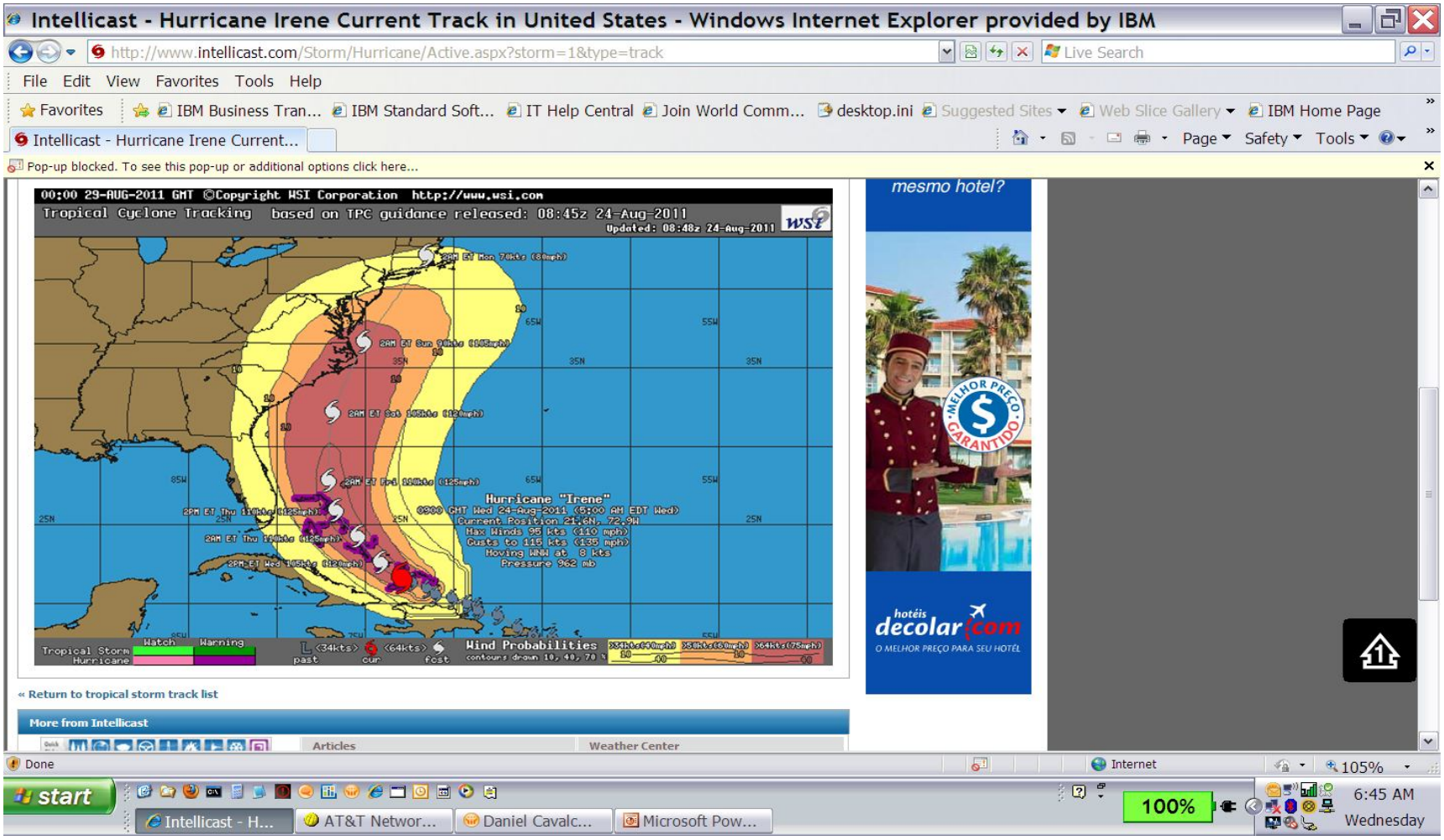
Non-resilience affects:

- ***Growth***
- ***Business risk***
- ***Competitive posture***
- ***Compliance to regulations***

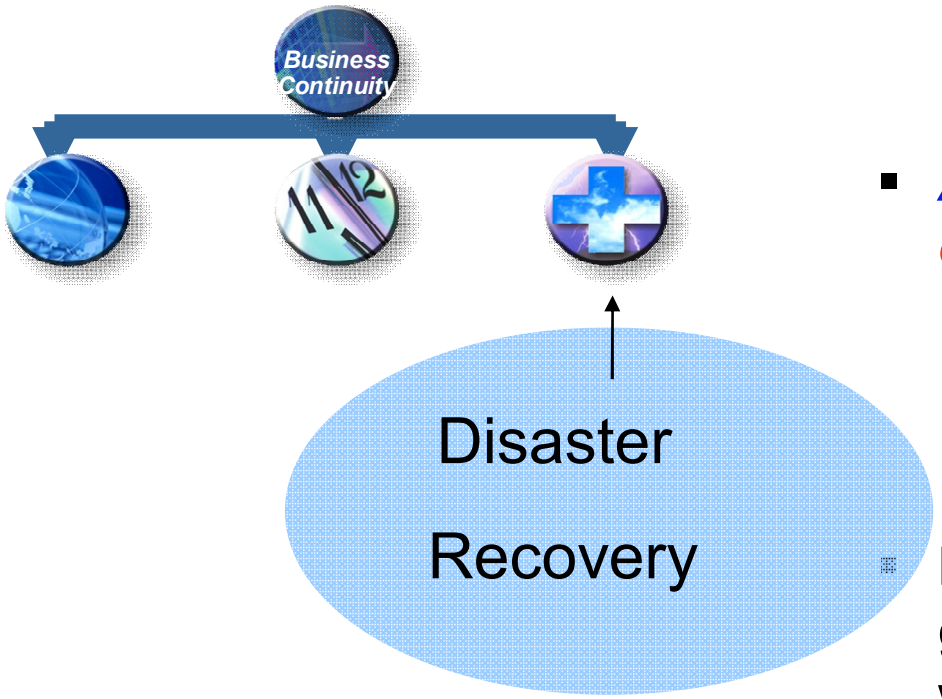
"Business Resilience is an integral thread that runs through the entire operation."

Source: FactPoint Study, May 2004

IRENE Hurricane (position Aug 24th 2011 at 5am)



Business Continuity - Disaster Recovery



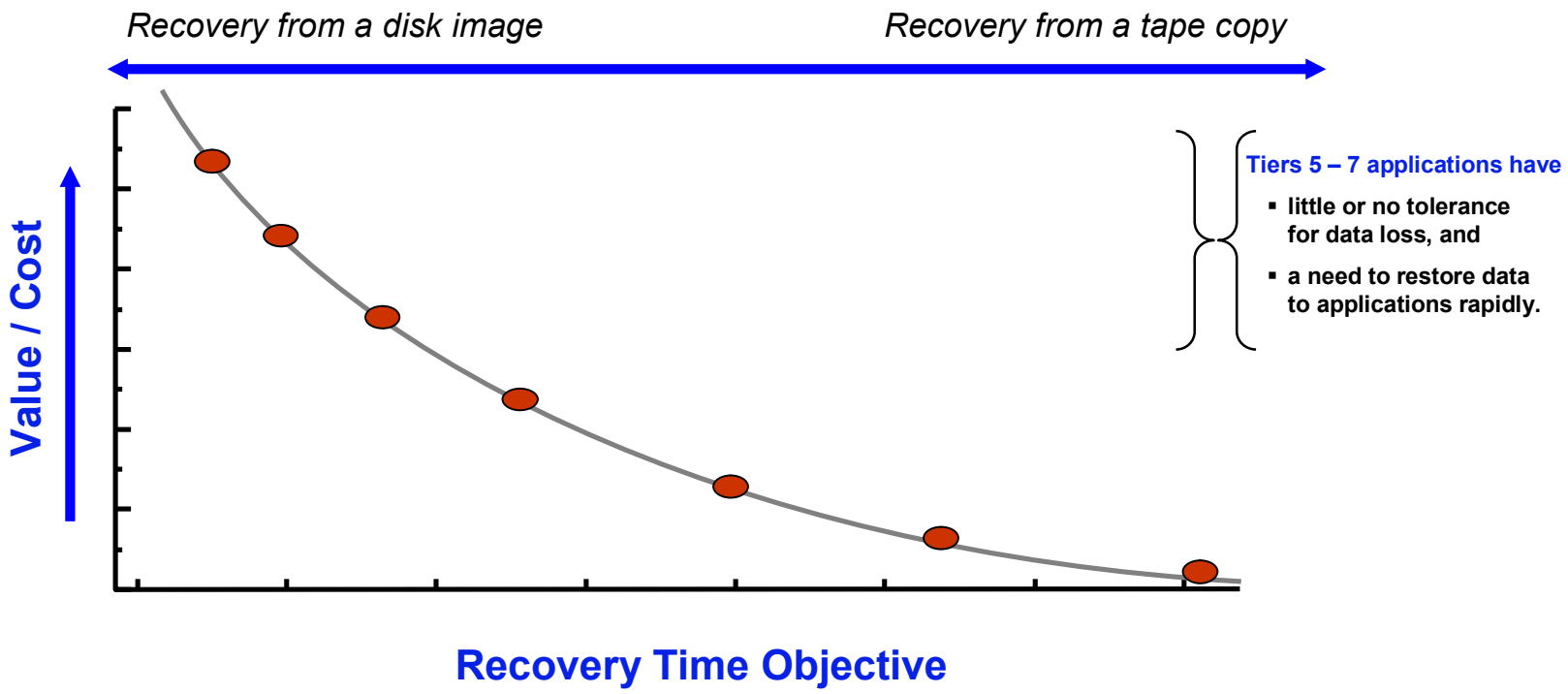
- *Ability to recover from unplanned outages at a different site*
  - Usually on *different hardware and location*

Performed after something has gone wrong on a primary site-wide basis

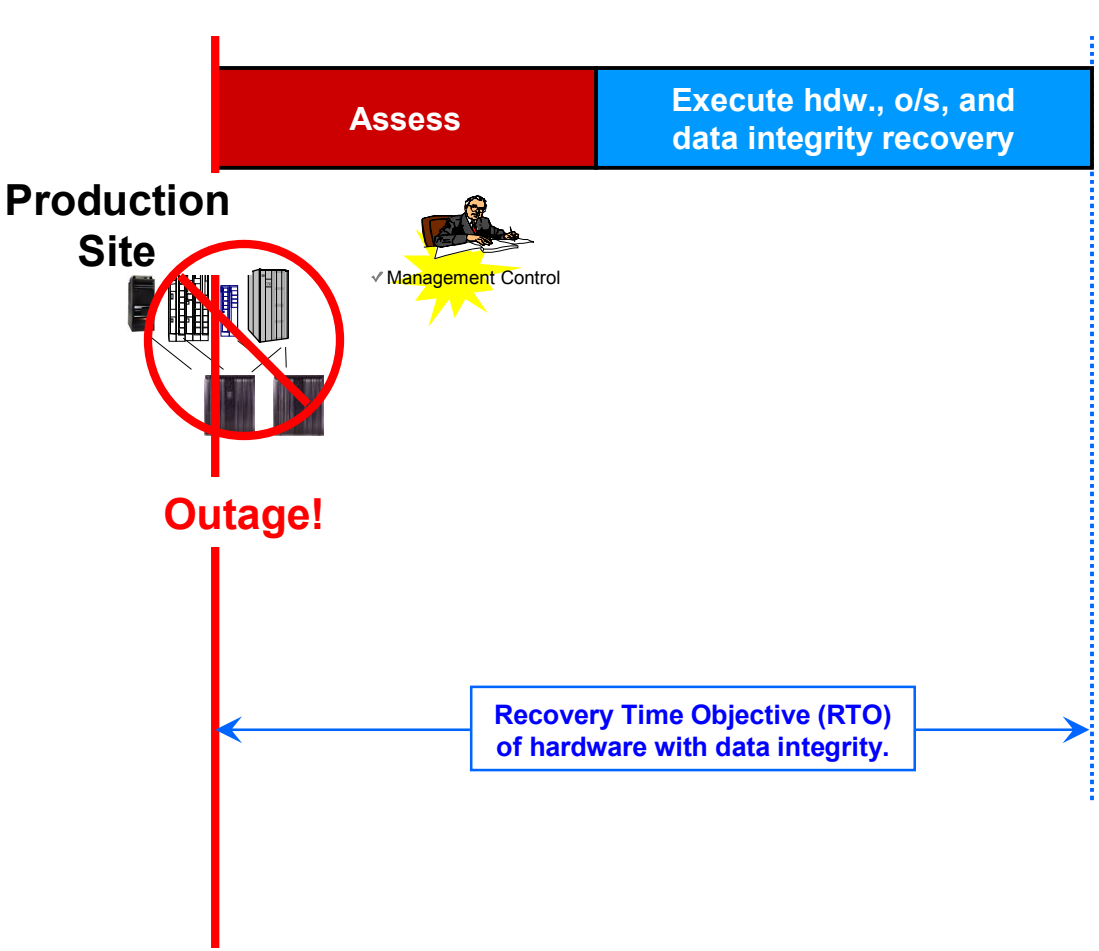
IT Business Recovery Metrics

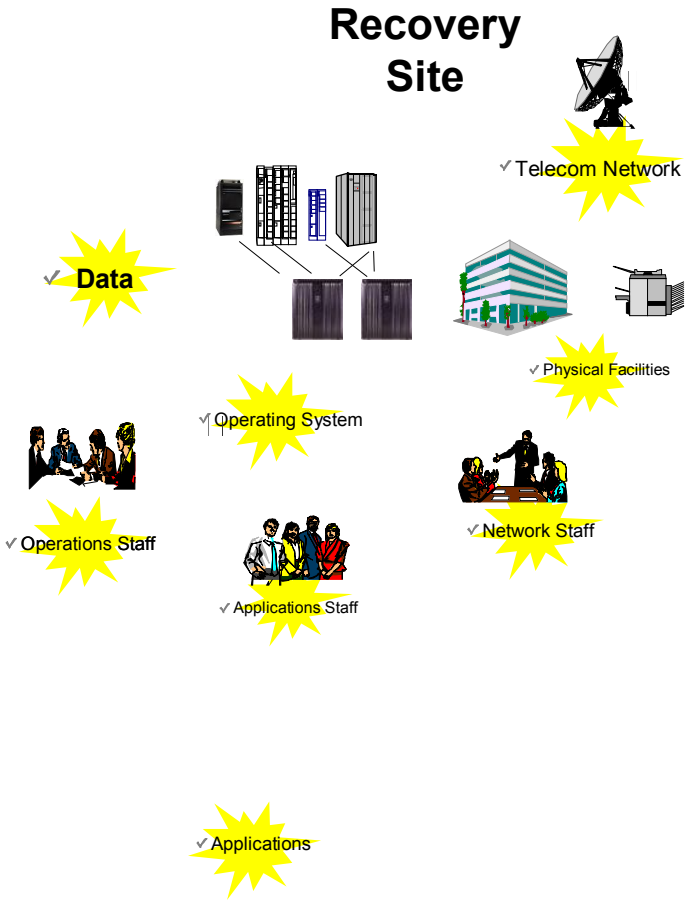
- Recovery Time Objective (RTO)
  - How long can I afford to be without my systems and *business-critical* applications?
- Recovery Point Objective (RPO)
  - How much data can I afford to *recreate* (or *lose*)?
    - Applications may be down until some or all of the data is recreated.
    - Denotes the time interval between when last Consistency Group was formed and when the storage system is again operational.
- Network Recovery Objective
  - How long it takes to get the network *operational*.

Business Continuity Tiers



Recovery Timeline (RTO)





Estimated Downtime Cost by Industry

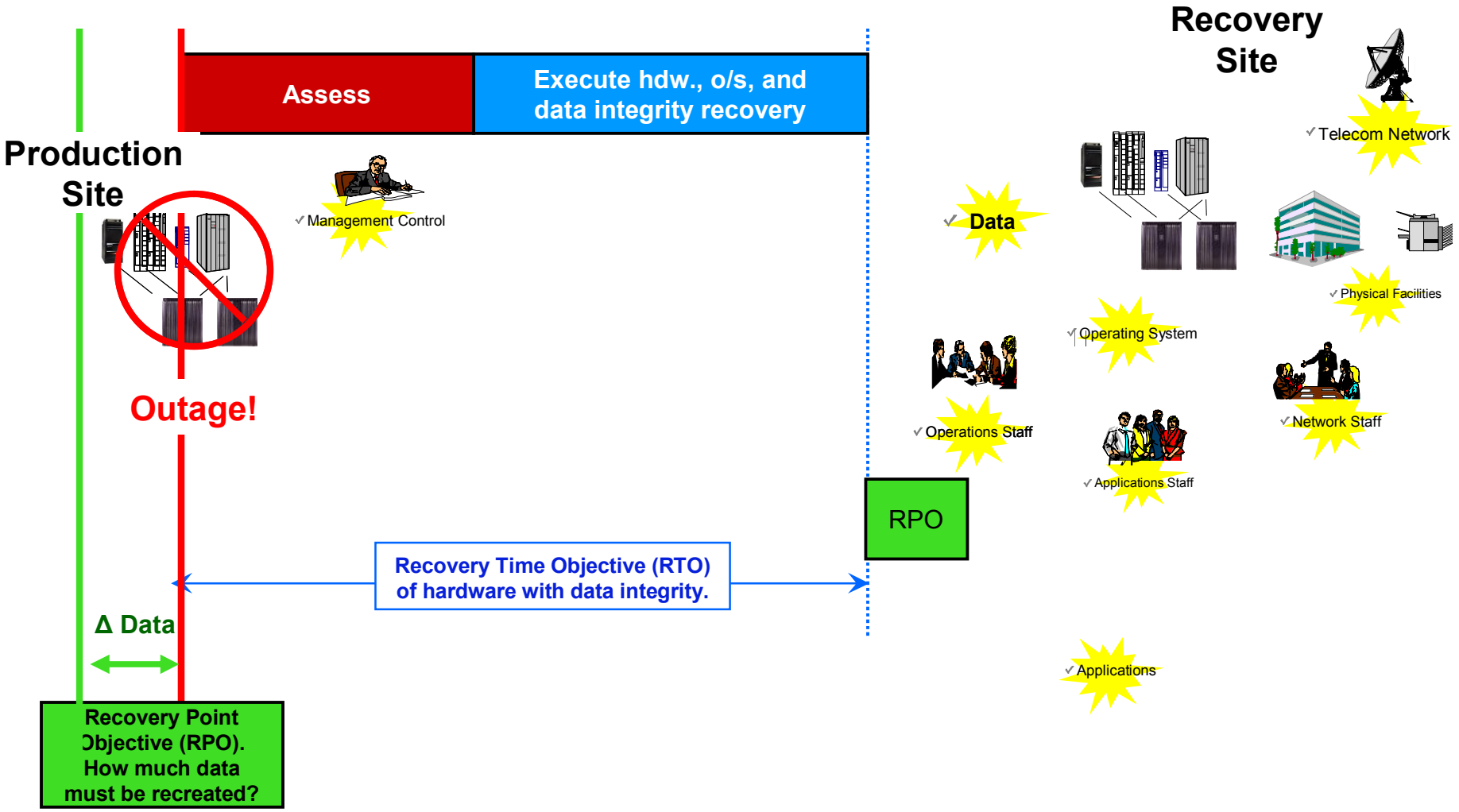
How much revenue do your business critical applications generate per hour?

How many business critical transactions will be lost per minute of downtime?

Business Operations	$ Revenue per Hour	30 Seconds of data is worth
Manufacturing	$1,610,000	$13,417
Finance	$1,495,000	$12,458
Retail	$1,107,000	$9,225
Pharmaceuticals	$1,082,000	$9,017
Healthcare	$636,000	$5,300
Utilities	$643,000	$5,358

Source: MetaGroup

Recovery Point Timeline (RPO)



The importance of Data Consistency

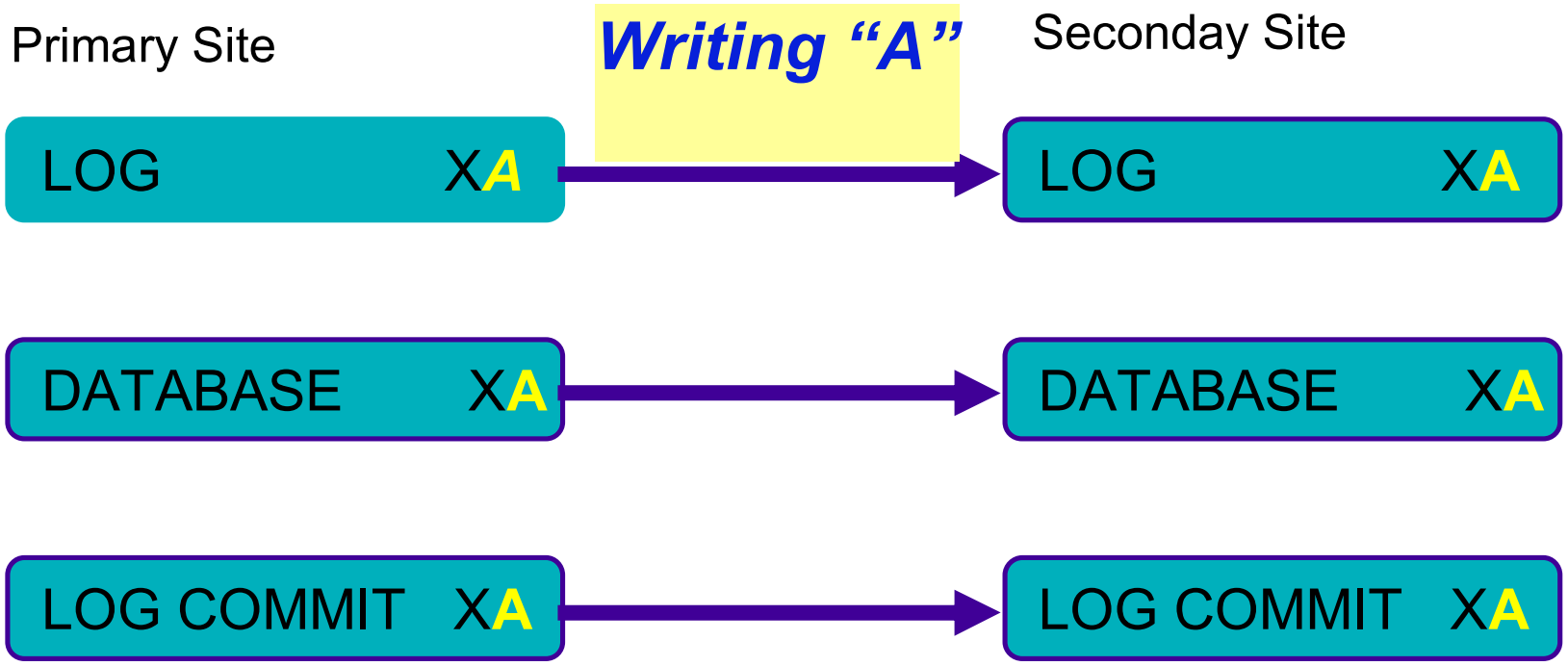
Database for performance does asynchronous multi-path and multi-thread
It DOESN'T STOP on a disk I/O error

Databases, when seeing a disk error, re-dispatch the write on alternate paths and volumes, usually on a different LUN disk frame

Result: Bypassing the synchronous nature of the storage hardware replication has been the cause of out-of-sync database conditions in synchronous mirroring environments

- ***IBM disk mirroring has specific technology to solve:***
  - It's called CONSISTENCY GROUPS (CG) and FREEZE
  - FREEZE suspends writing data to secondary, consistently, across all volumes in the Consistency Group
  - GDPS and TPC for Replication have integrated support of CG and FREEZE technology

Example of Database Mirroring and Consistency
Planned and Normal DB Recovery

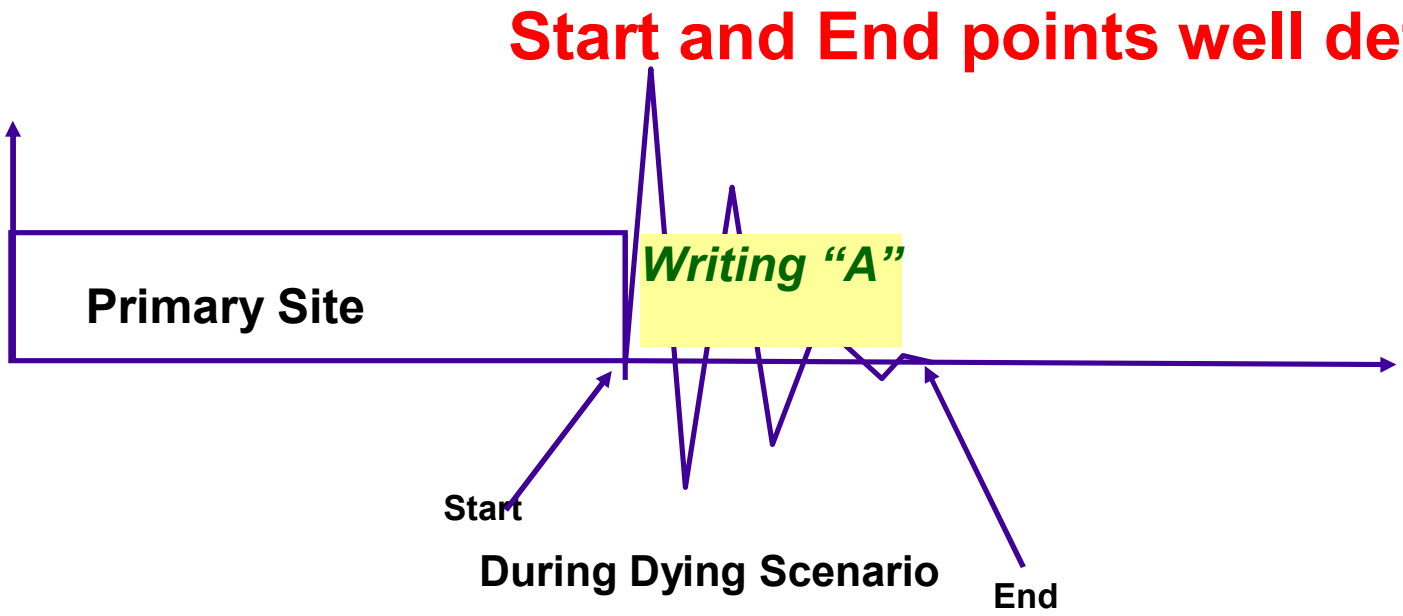


Planned DB Recovery at Secondary site:

Re-start command resumes operation in a matter of minutes

Concept of "Rolling Disaster"

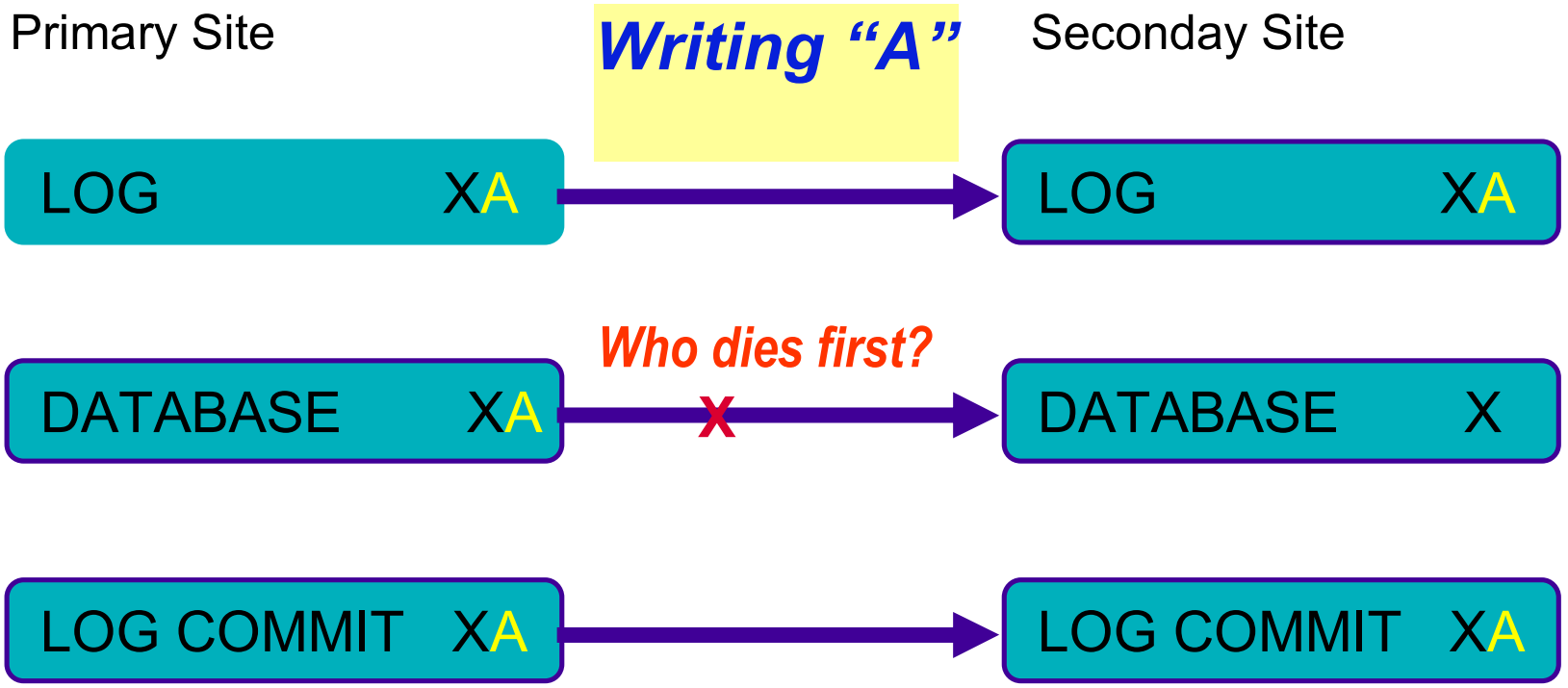
Dying scenario



During Dying Scenario any outage provokes unpredictable results (UNPLANNED OUTAGE)

Example of Database Mirroring and Non Consistency

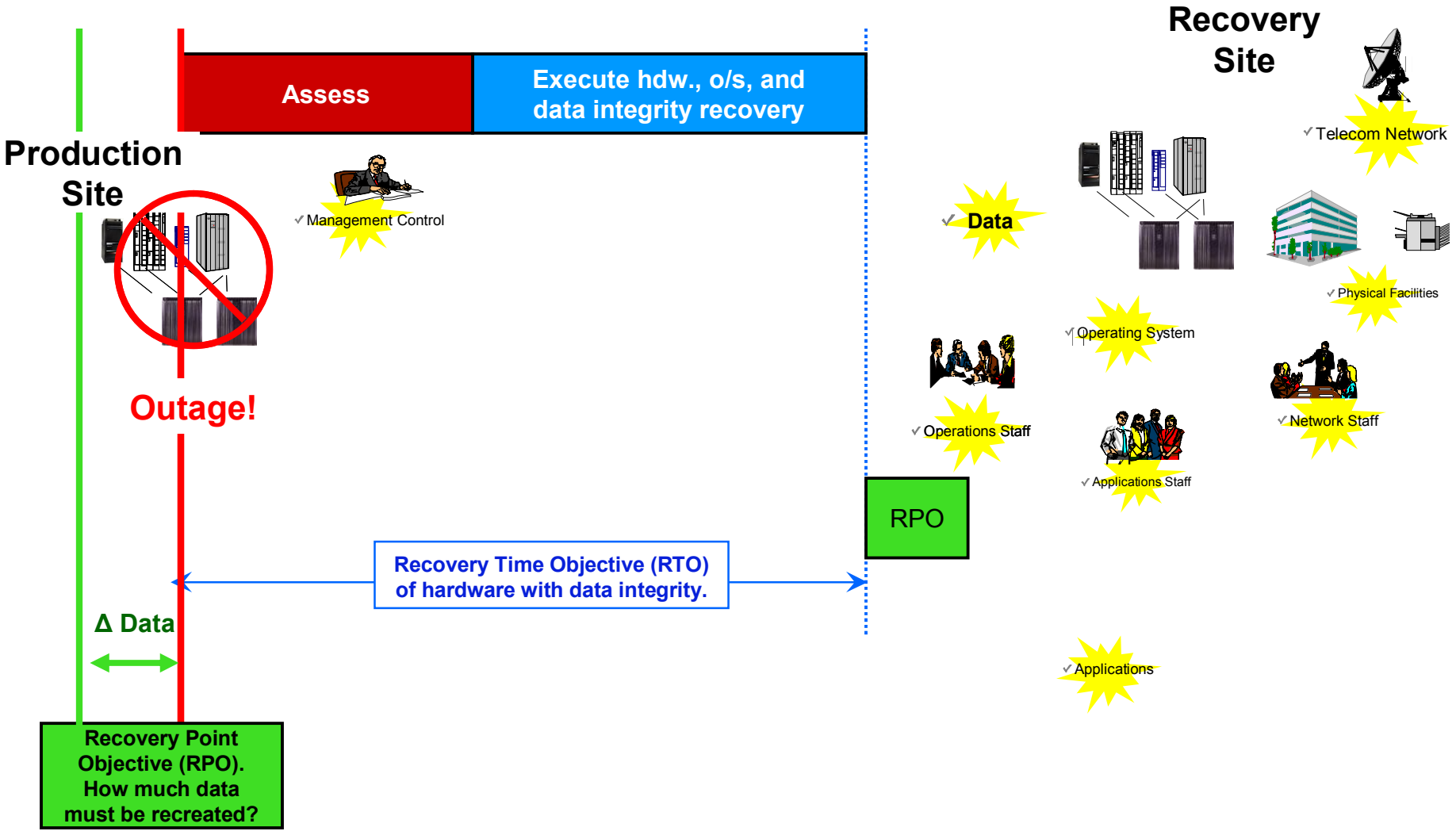
Unplanned outage



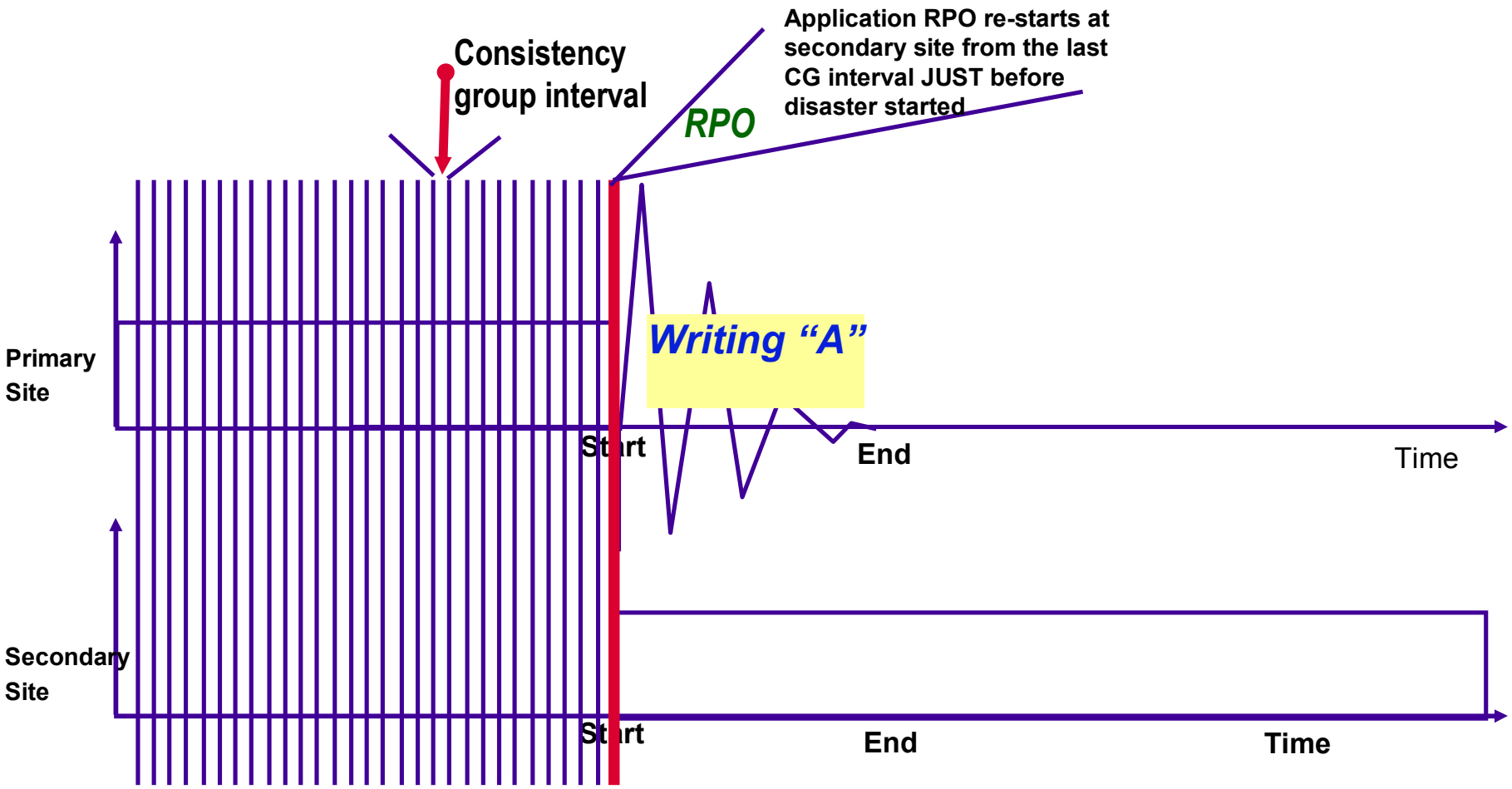
RECOVERY AT SECONDARY SITE WILL FAIL.

TAPE RECOVER - applies logs to get last consistency point (RPO) before outage
That causes longer time to take the application back ONLINE (RTO)

Recovery Point Timeline (RPO)



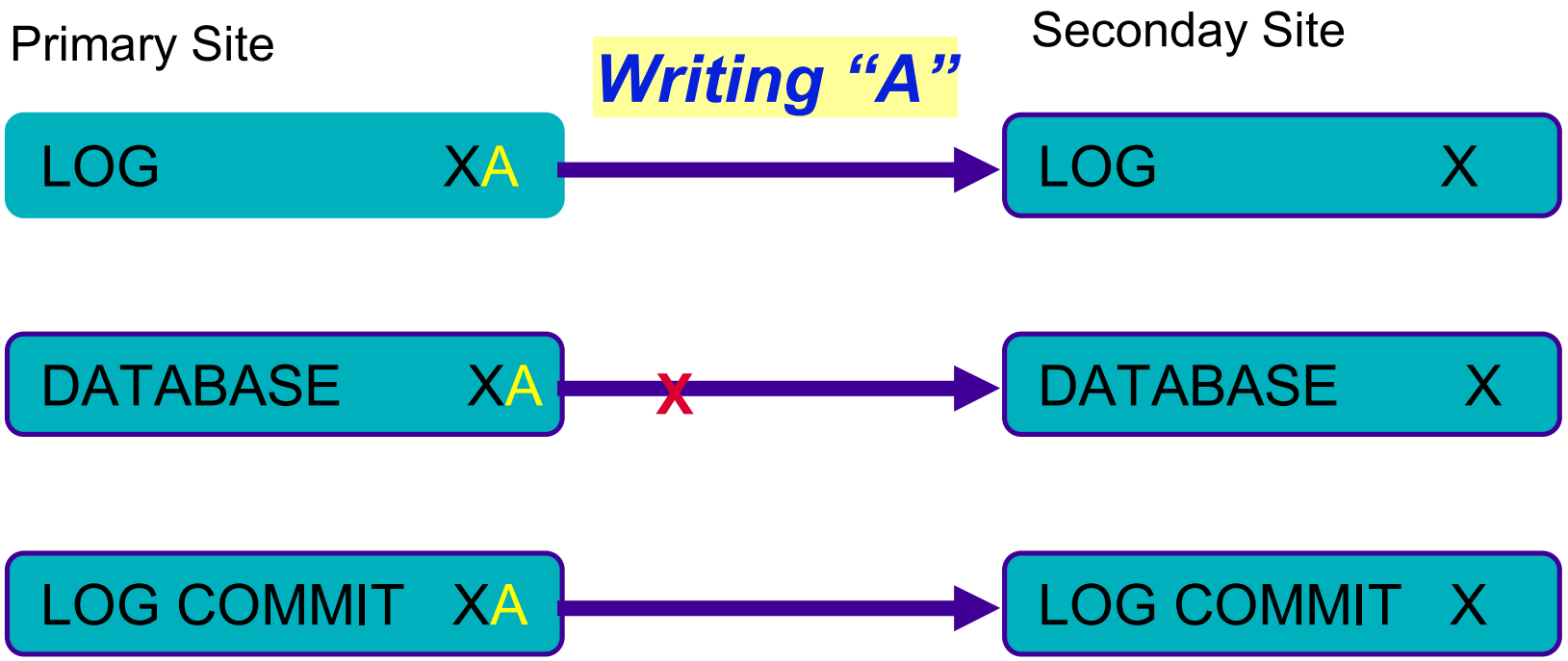
IBM DS8000 Solution Keeps consistency



Consistency allows to face unplanned outage-disaster at anytime

Example of Database and Consistency

IBM Global Mirror



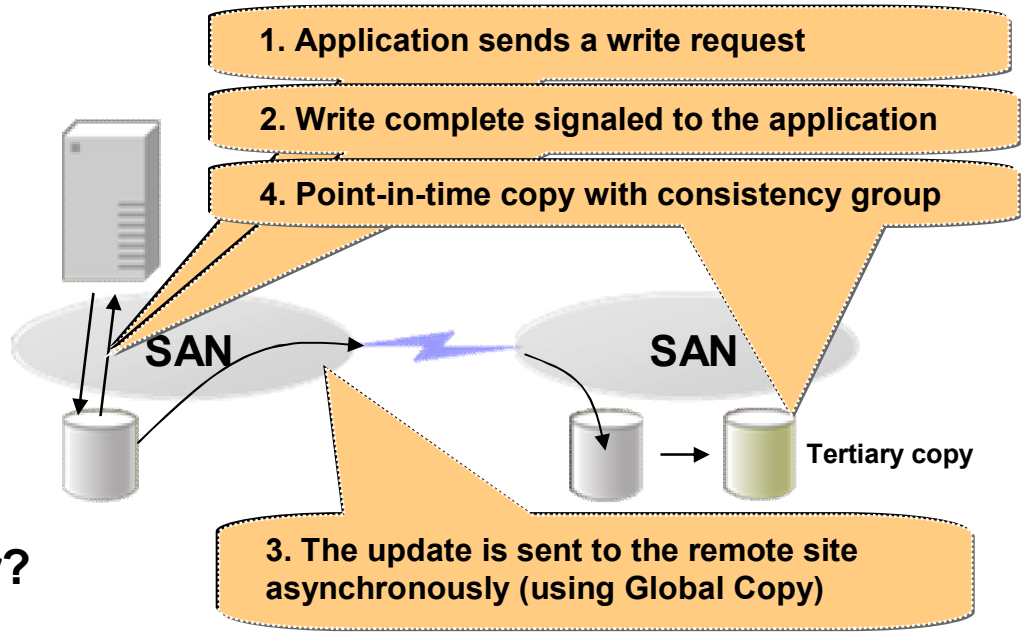
Freeze & consistency keep data consistent at the secondary site.

During Rolling Disaster dying scenario, DS8000 doesn't propagate mirror updates

RECOVERY AT SECONDARY SITE resumes in a matter of minutes

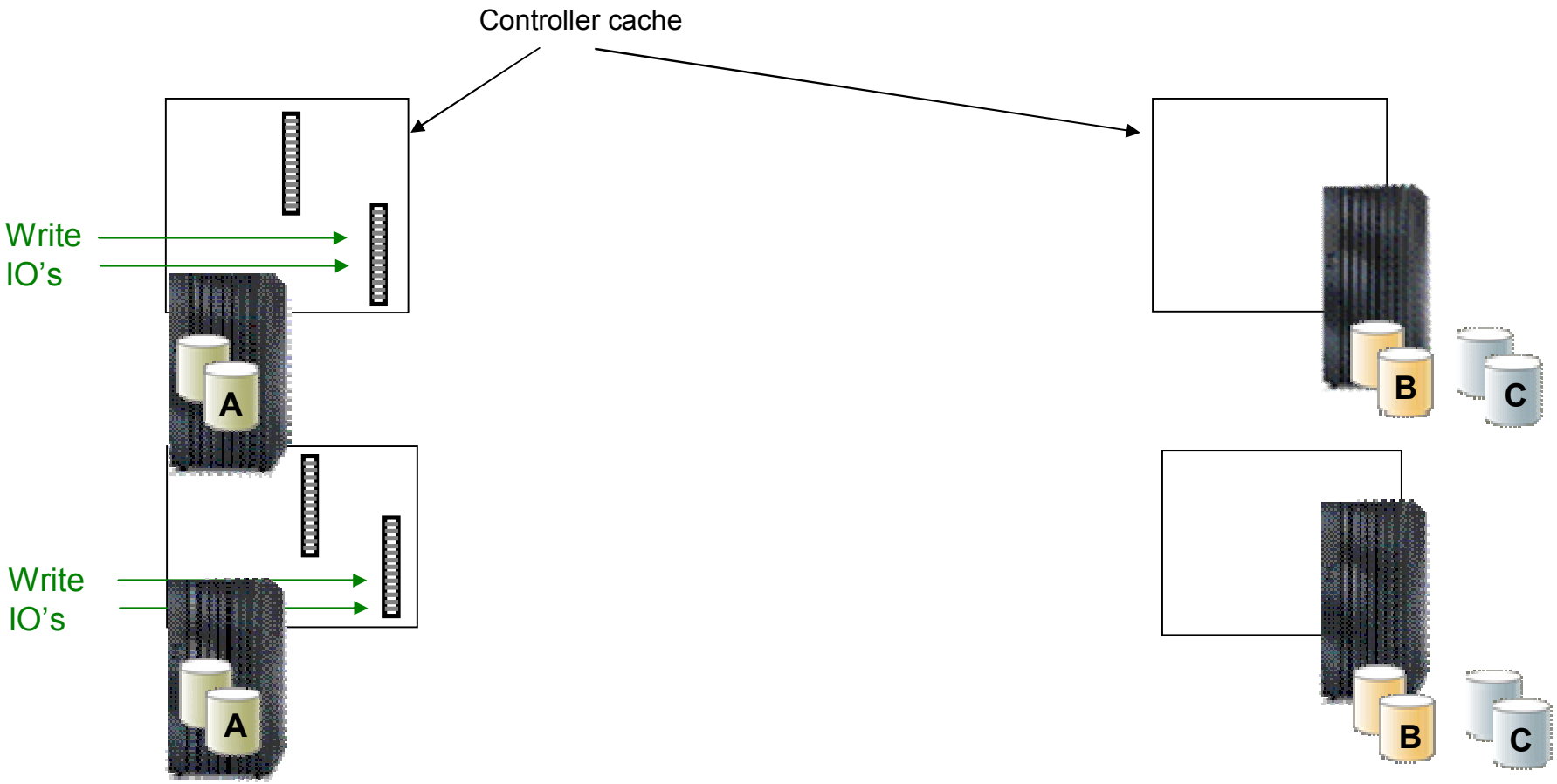
IBM Global Mirror

- **What is Global Mirror?**
  - ***Asynchronous disk mirror with continuous data integrity***
  - Performed at a volume level
  - ***Built by combining:***
    - ***Global Copy***
    - ***FlashCopy***
    - ***FREEZE***
- **What is Global Mirror used for?**
  - ***For asynchronous mirror with data consistency, at any distance***
  - Creates tertiary data consistent copy every 3-5 seconds using incremental FlashCopy



IBM Global Mirror Timeline =>

DS8000 cache NVS Bit Mapping

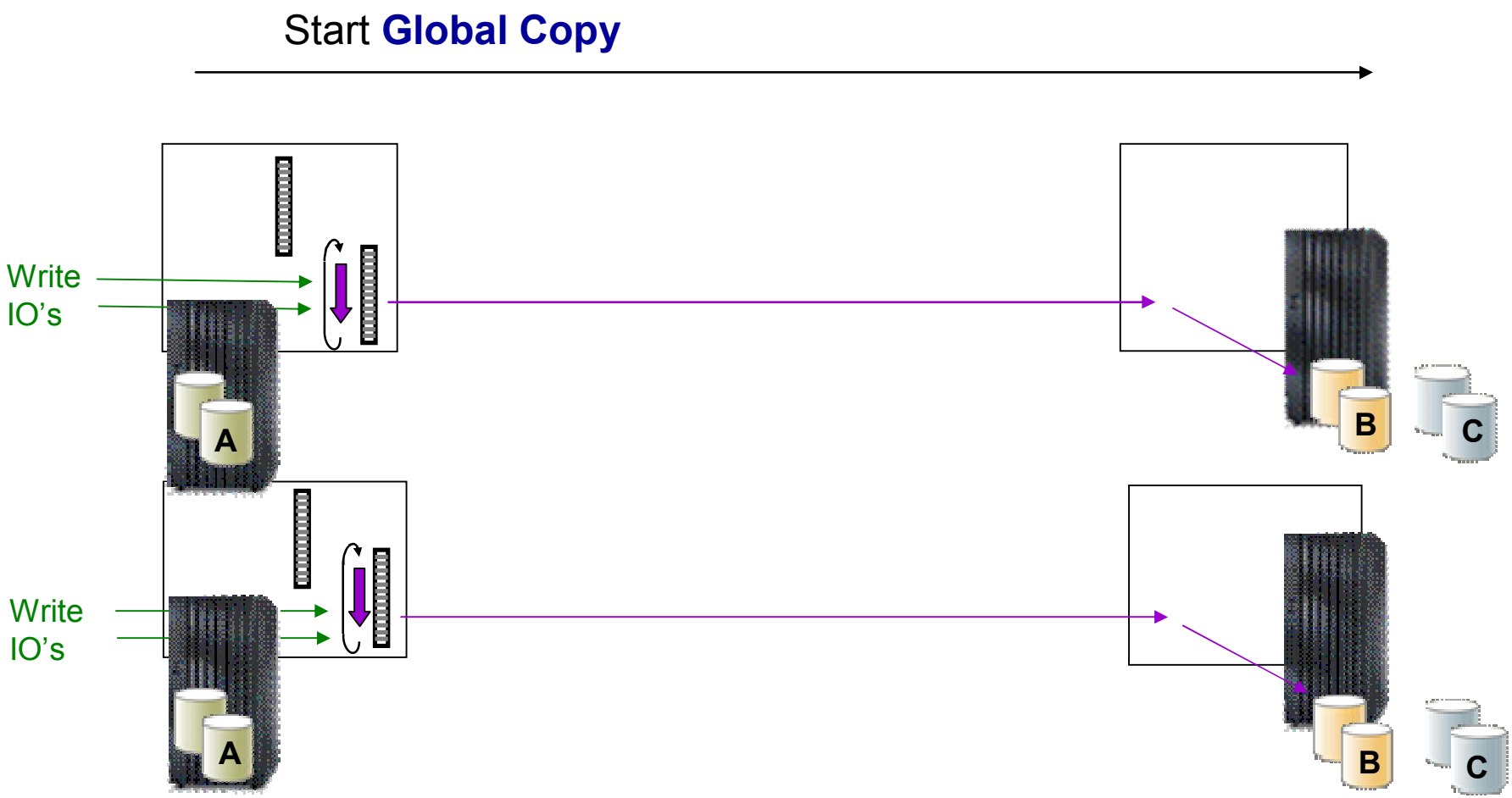


- Local site

- **Remote site**

IBM Global Mirror Timeline =>

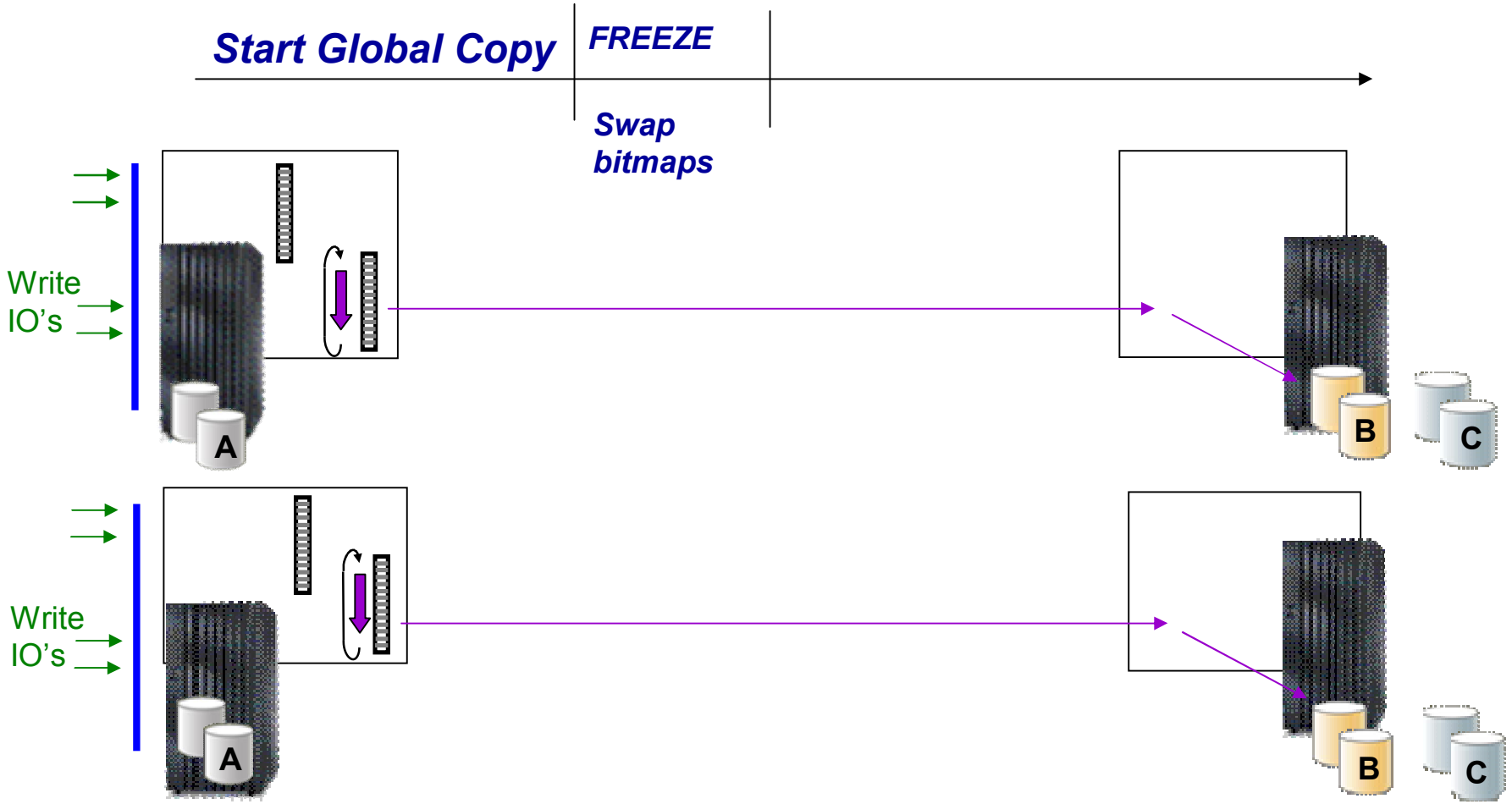
DS8000 cache NVS Bit Mapping



- **Global Copy** rotates thru bitmap, sending data in high performance (fuzzy) sequence

IBM Global Mirror Timeline =>

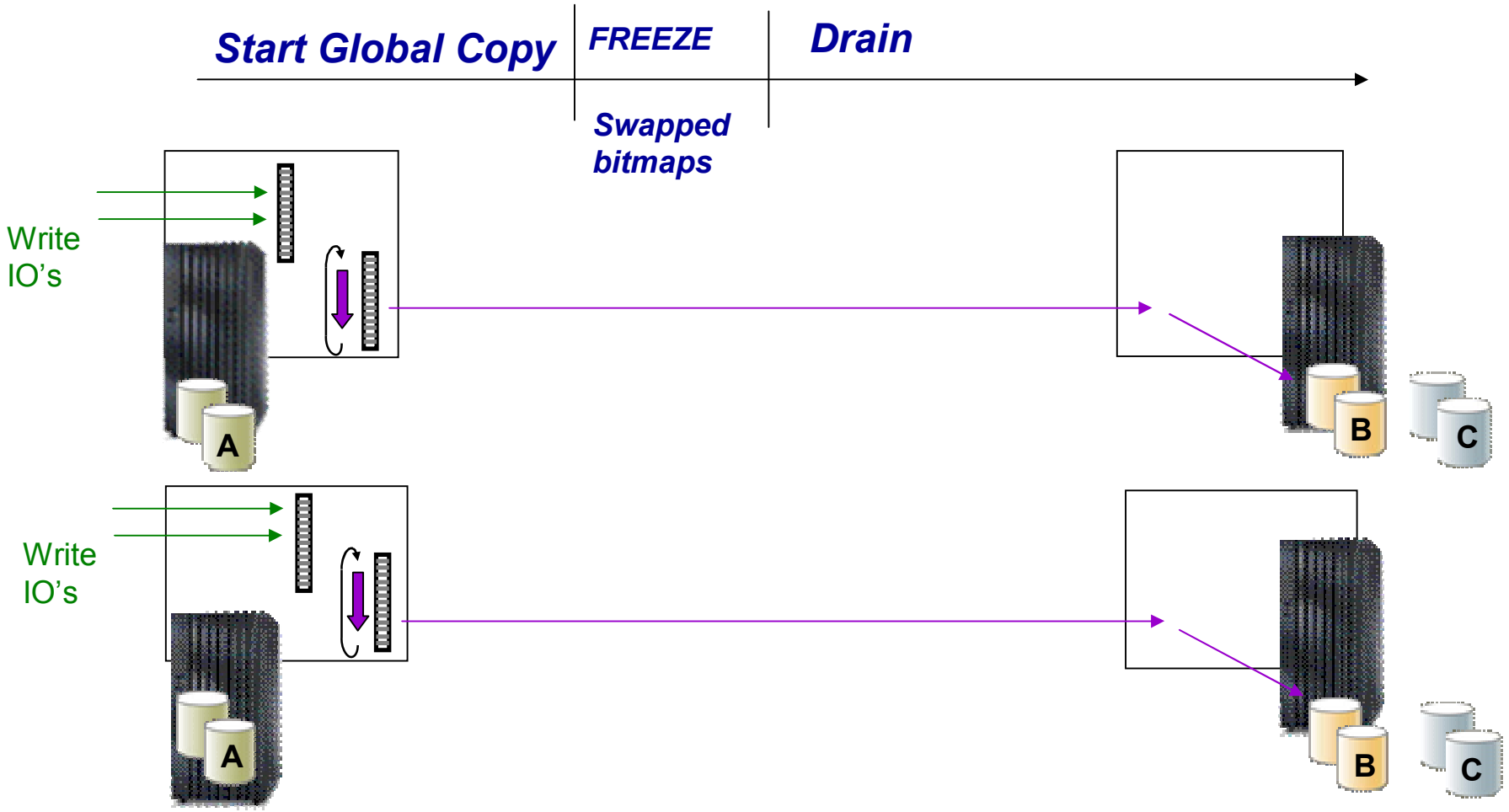
DS8000 cache NVS Bit Mapping



- *Consistency Group* initiates FREEZE **(Extended Long Busy)**
  - **Queue incoming write IOs for short period (usually < 3 ms)**
- Swap IO to other bitmap: ***First bitmap identifies the Consistency Group contents***

IBM Global Mirror Timeline =>

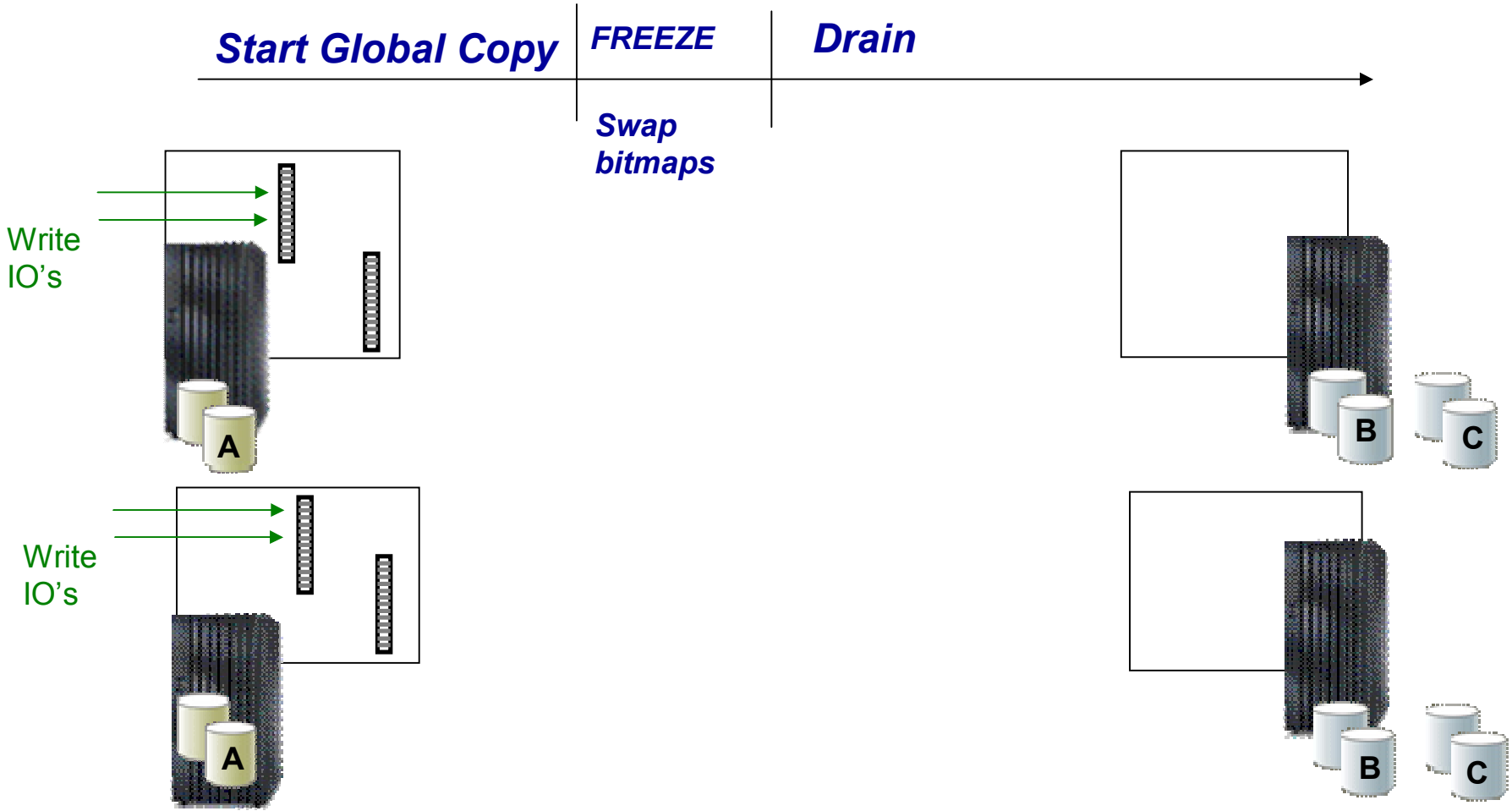
DS8000 cache NVS Bit Mapping



- ***FREEZE/RUN*** issued; releases IO to run again
  - **Incoming IO now tracked via other bitmap**
  - **Meanwhile, original bitmap used to *drain remaining CG data***

IBM Global Mirror Timeline =>

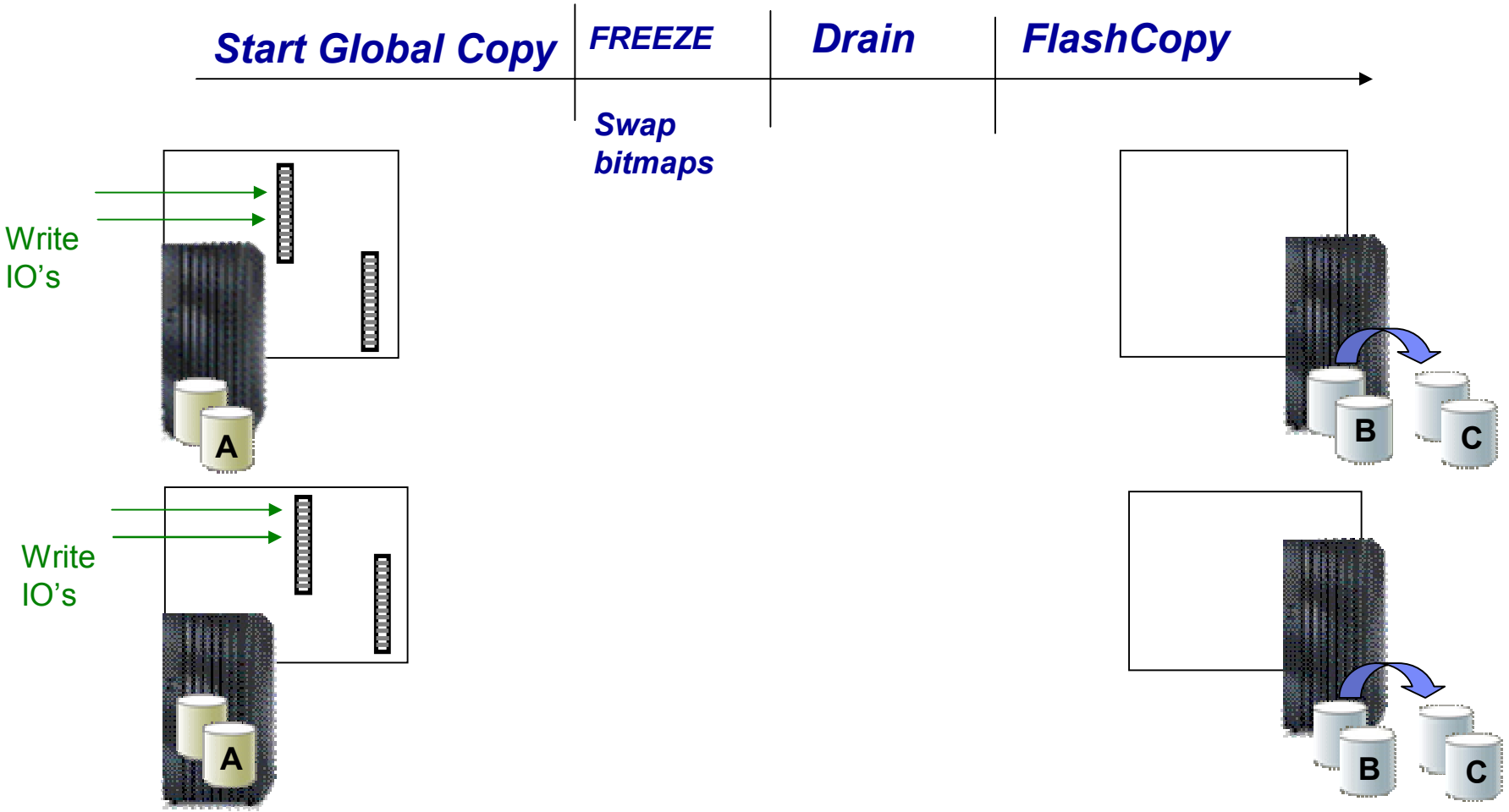
DS8000 cache NVS Bit Mapping



- All remaining data in CG drained to remote volume and at this point:
  - **B is the next CG**
  - **C is the n-1 CG**

IBM Global Mirror Timeline =>

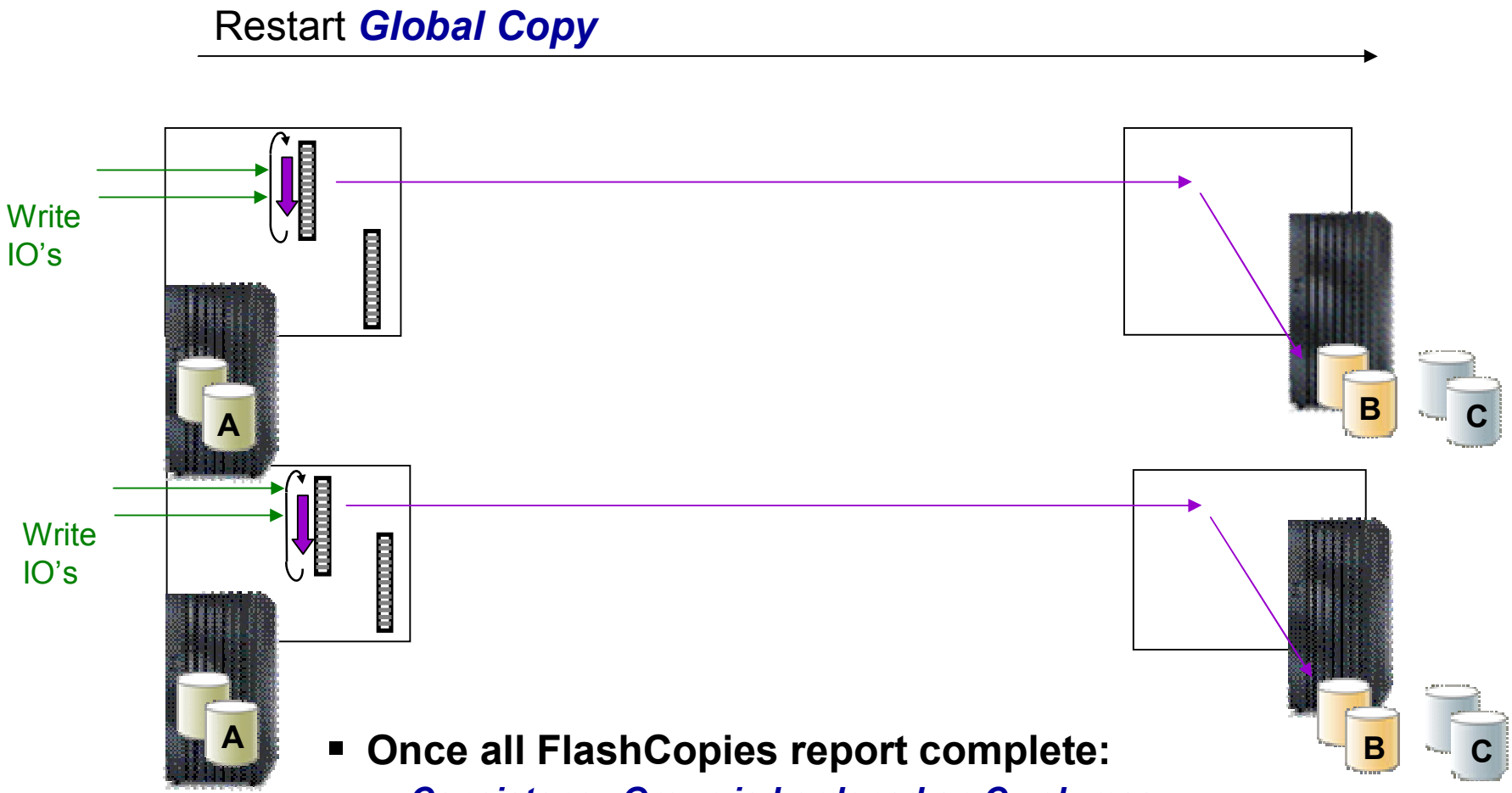
DS8000 cache NVS Bit Mapping



- Inband Flashcopy sent from primary to remote site
  - **Consistency Group is FlashCopied B->C, thus hardened on the FlashCopy targets**

IBM Global Mirror Timeline =>

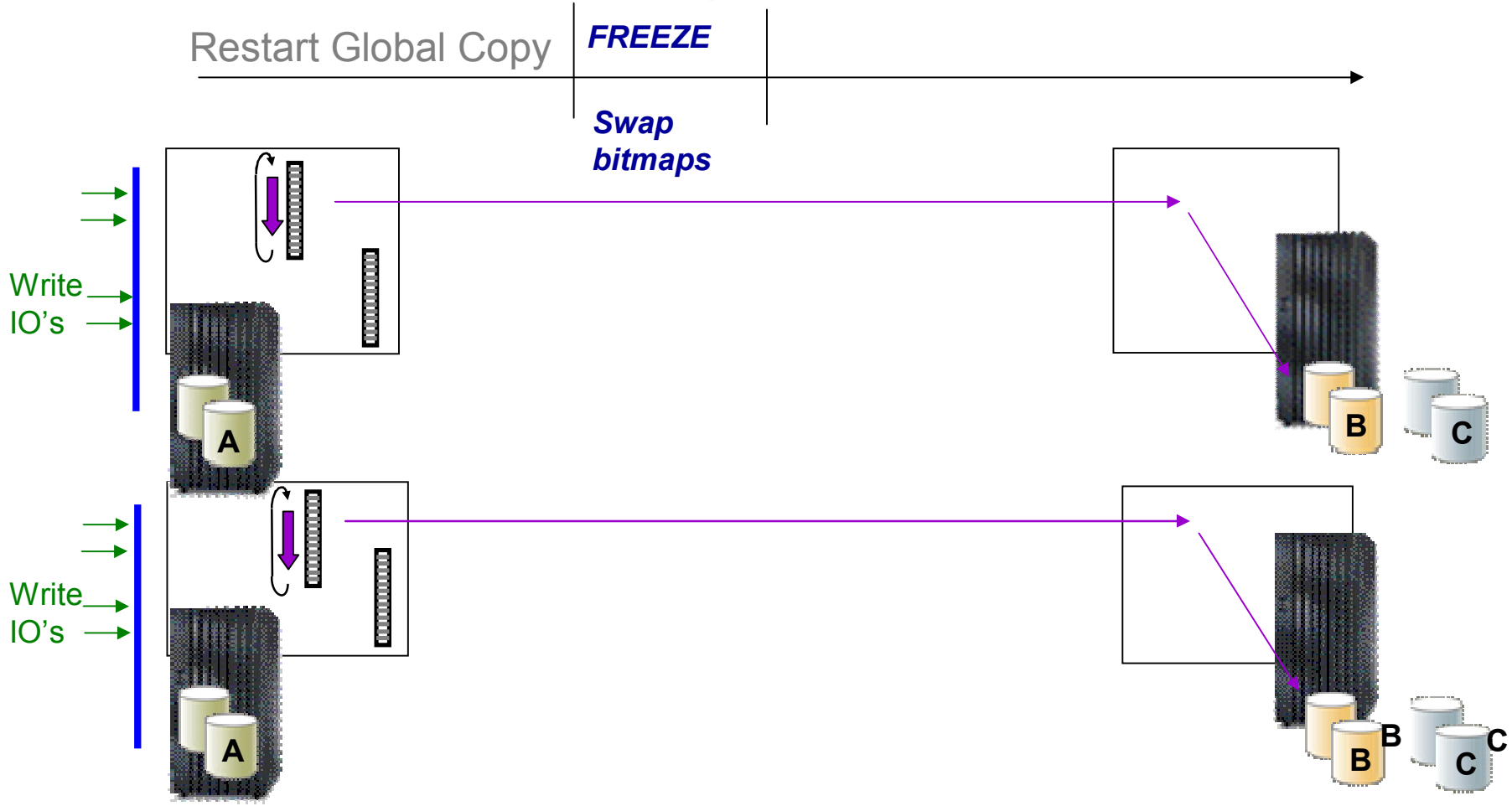
DS8000 cache NVS Bit Mapping



- **Once all FlashCopies report complete:**
  - ***Consistency Group is hardened on C volumes***
  - Global Copy restarts from the *other set of bitmaps* to B volumes
  - And the cycle repeats

IBM Global Mirror Timeline =>

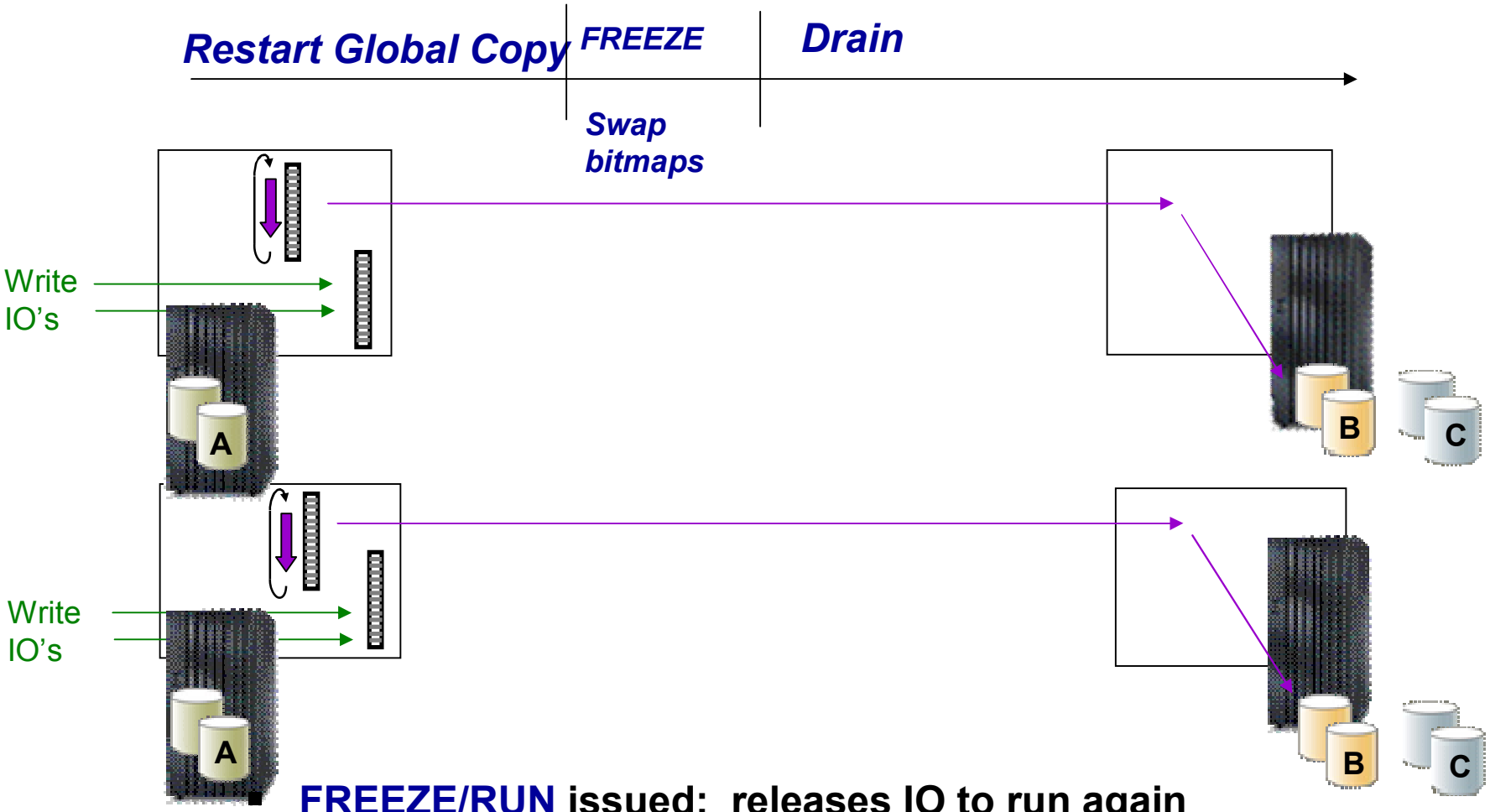
DS8000 cache NVS Bit Mapping



- *Next Consistency Group* initiates FREEZE
  - **Queue incoming write IOs for short period usually < 3 ms**
- Swap IO to other bitmap: ***bitmap identifies the Consistency Group contents***

IBM Global Mirror Timeline =>

DS8000 cache NVS Bit Mapping

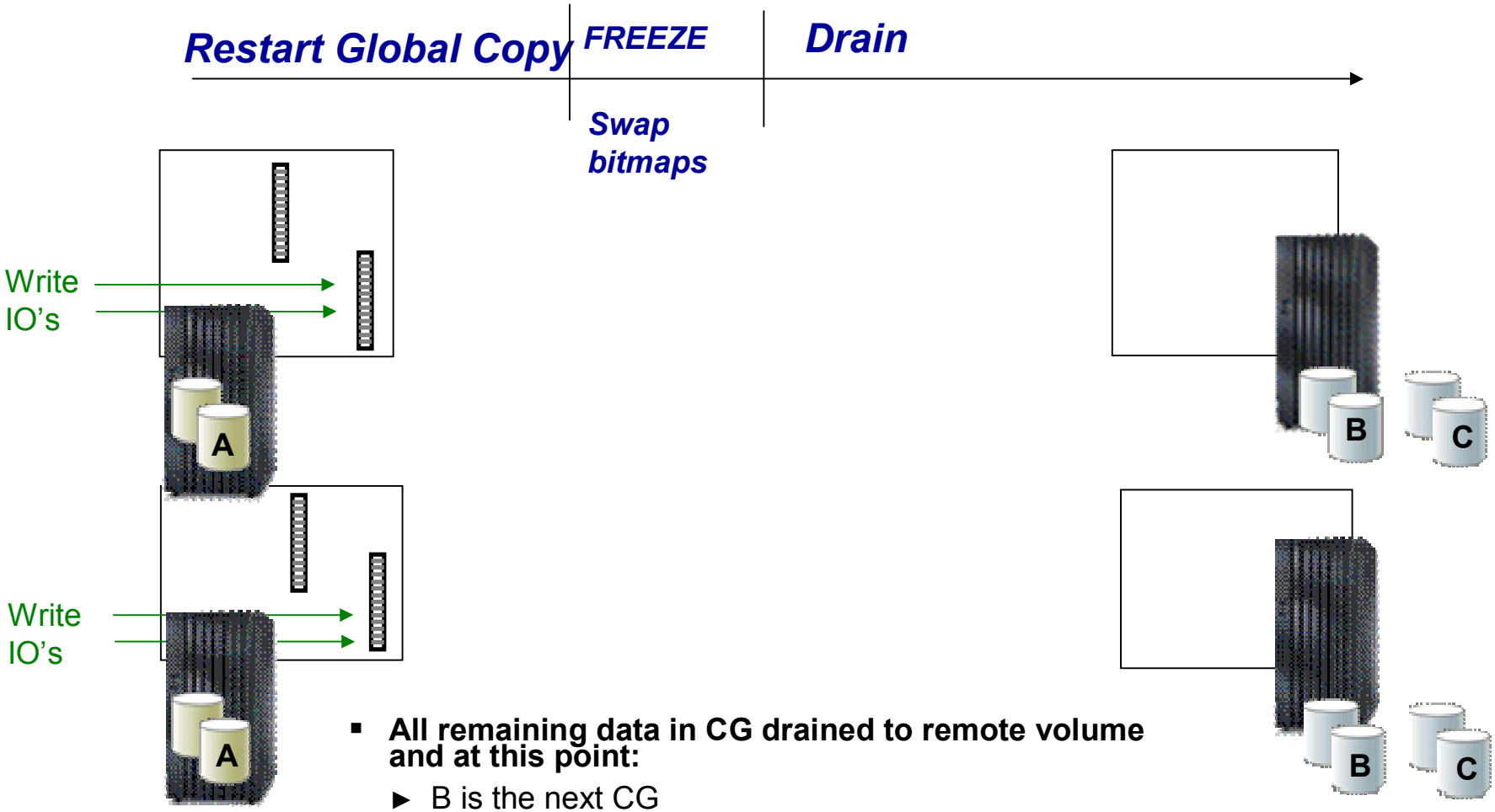


FREEZE/RUN issued; releases IO to run again

- Incoming IO now tracked via other bitmap
- Meanwhile, original bitmap used to **drain remaining CG data**

IBM Global Mirror Timeline =>

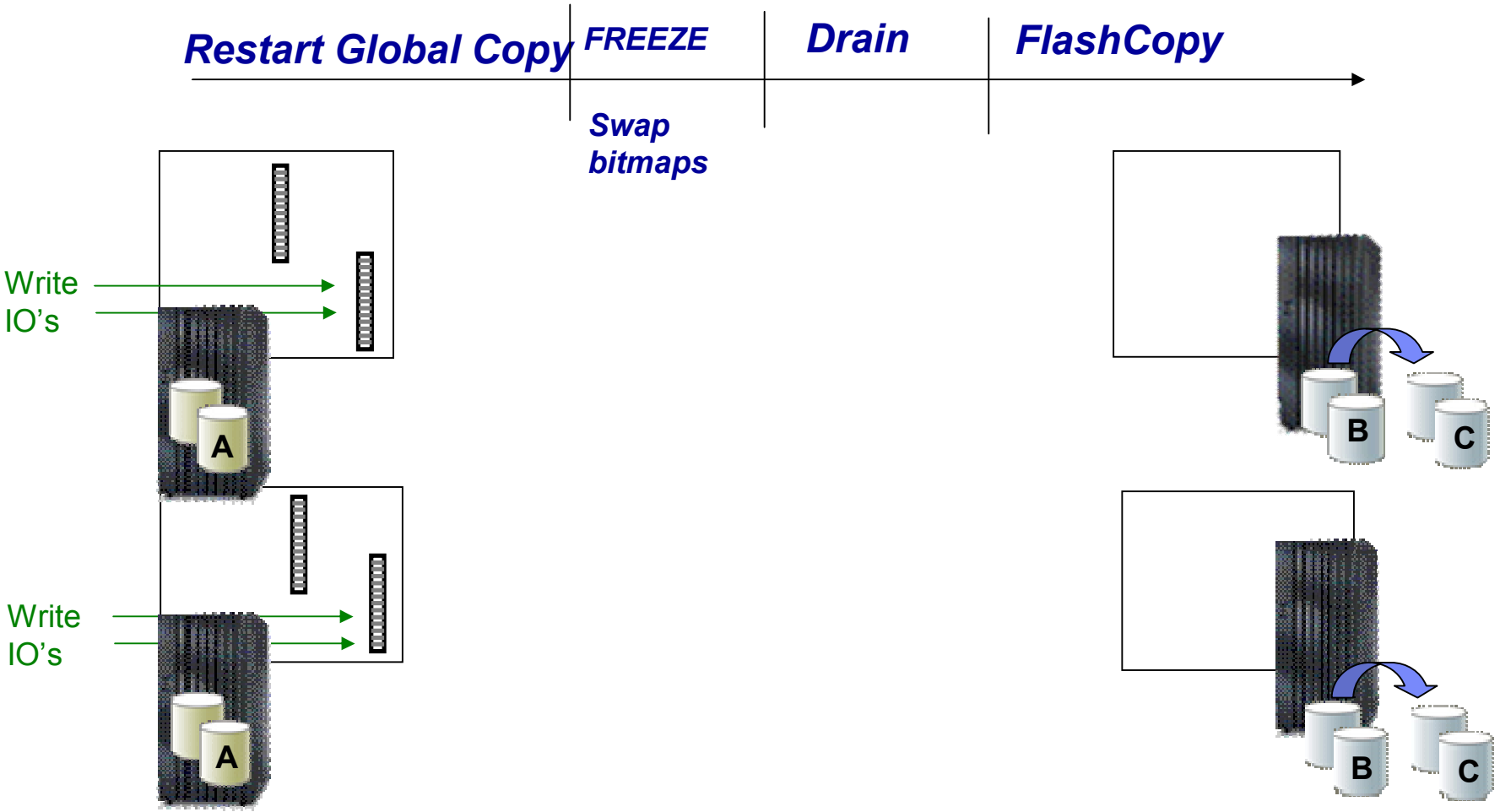
DS8000 cache NVS Bit Mapping



- **All remaining data in CG drained to remote volume and at this point:**
  - B is the next CG
  - C is the n-1 CG

IBM Global Mirror Timeline =>

DS8000 cache NVS Bit Mapping

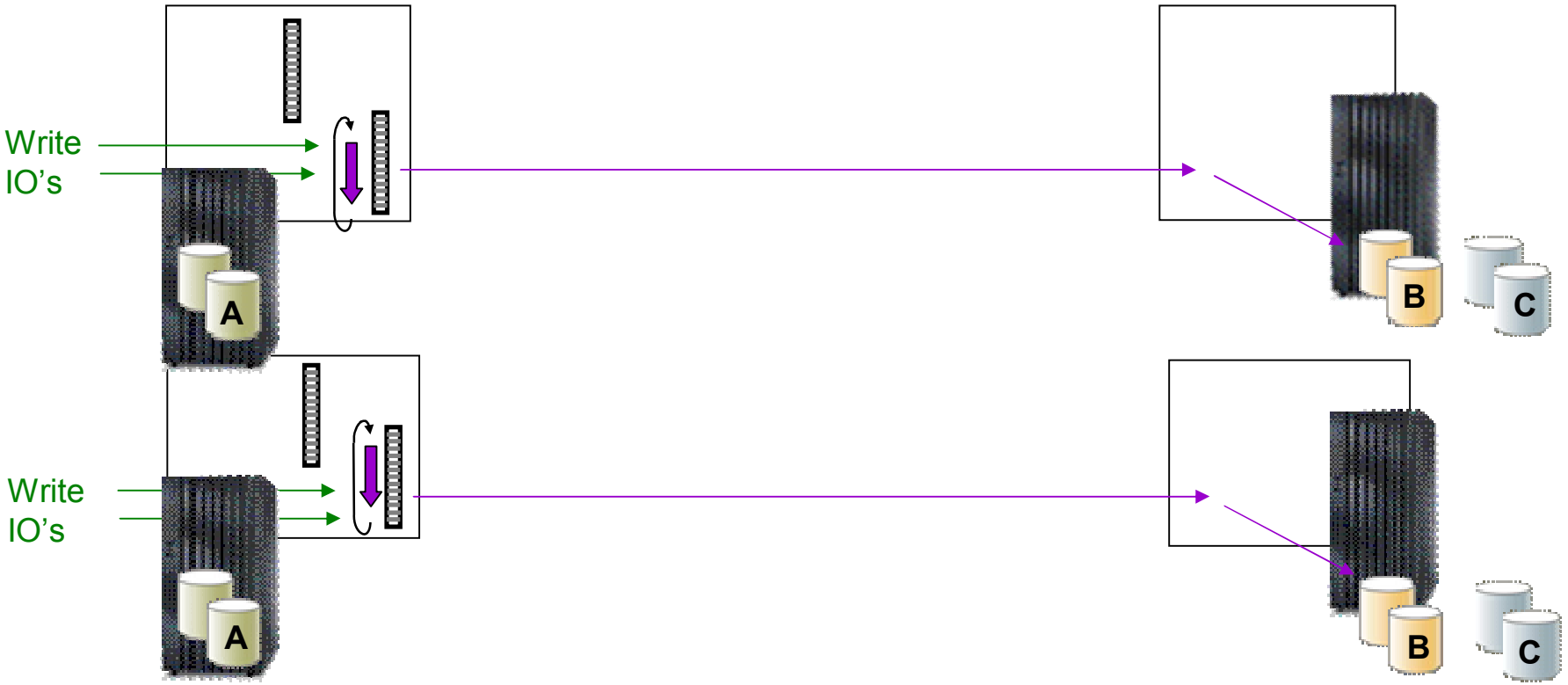


- **Inband Flashcopy sent from primary to remote site**
  - Consistency Group is FlashCopied B->C, thus hardened on the FlashCopy targets

IBM Global Mirror Timeline =>

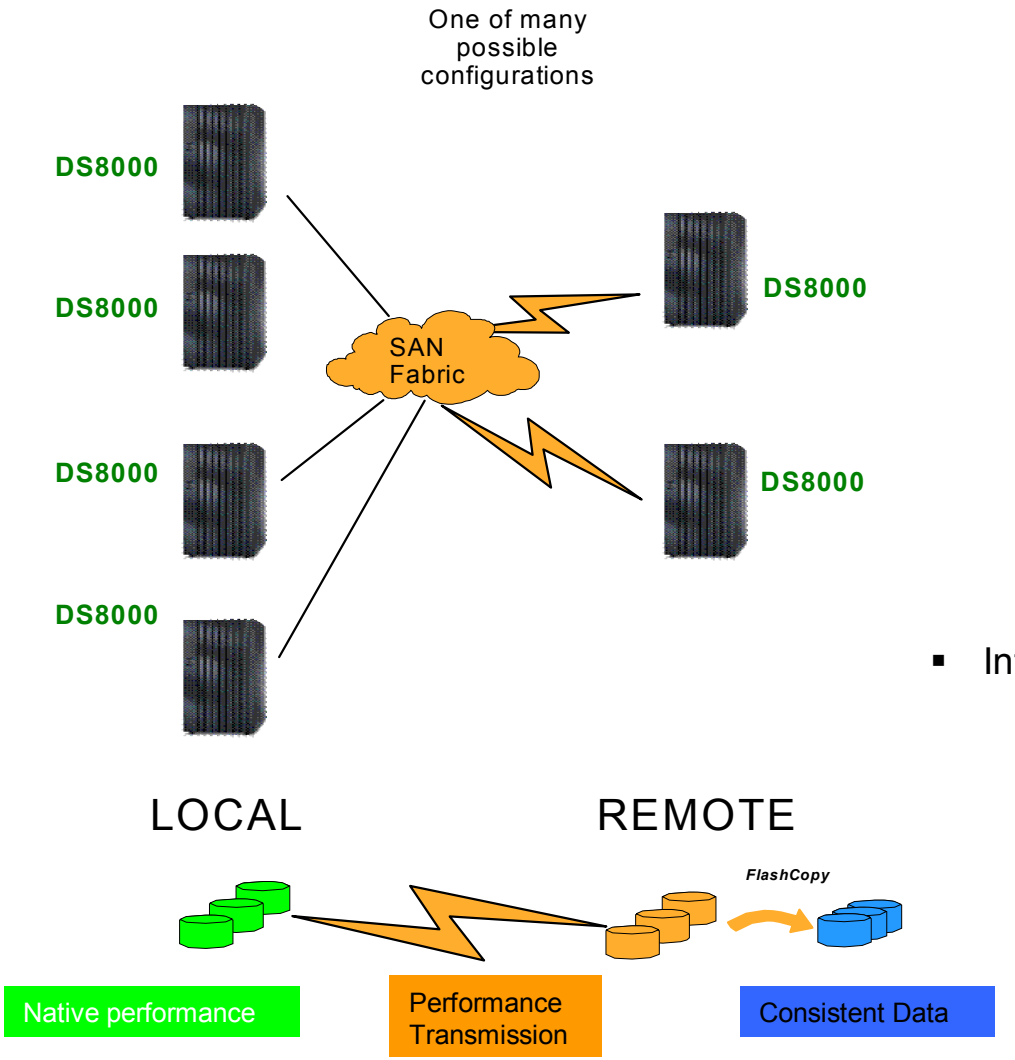
DS8000 cache NVS Bit Mapping

Restart ***Global Copy***



- **Restart Global Copy**
  - Consistency Group is hardened on C volumes
  - And the cycle repeats

IBM DS8000
Global Mirror for Business Continuity



- Designed to Provide:
  - **Copy consistency:** managed autonomic by Master Control Server in master DS8000
  - **Global Distance:** Two-site, unlimited distance, data consistent asynchronous disk mirroring
  - **Scalability:** Consistency Group supported across up to 17 total ESSs in Global Mirror session
  - **Flexibility:** Many possible configurations
  - **Heterogeneous:** Data can span zSeries® and open systems data, and can contain a mix of zSeries and open systems data
  - **Application Performance:** Native
  - **Mirroring Performance:** Two ESS Fibre Channel disk mirroring links per DS8000 sufficient for almost all workloads

- Intended Benefits
  - **Autonomic:** No active external controlling software required to form consistency groups
  - **Saves cost:** No server cycles required to manage consistency groups
  - **Lowers TCO:** designed to provide improved performance, global distances and lower costs

Business Continuity done RIGHT …

40%

60%

Technology
Hardware and software capabilities

Process
Definition/design, compliance and continuous improvement

People
Roles & responsibilities, management, skills development & discipline

… It is a Business Solution, not a technology decision. …

Technical Forum & Executive Briefing 17 al 21 Octubre 2011

GRACIAS

IBM