



***The Modern Mainframe...  
At the Heart of Your Business***

**A Mainframe Primer – Continuous Availability**



© 2006 IBM Corporation

## **Service Oriented Finance Needs a World Class Availability Solution**

**Our new loan application must be available to serve agents around the world 24 by 7**



**Service Oriented Finance  
CIO**

**You can't do better than the availability provided by System z**



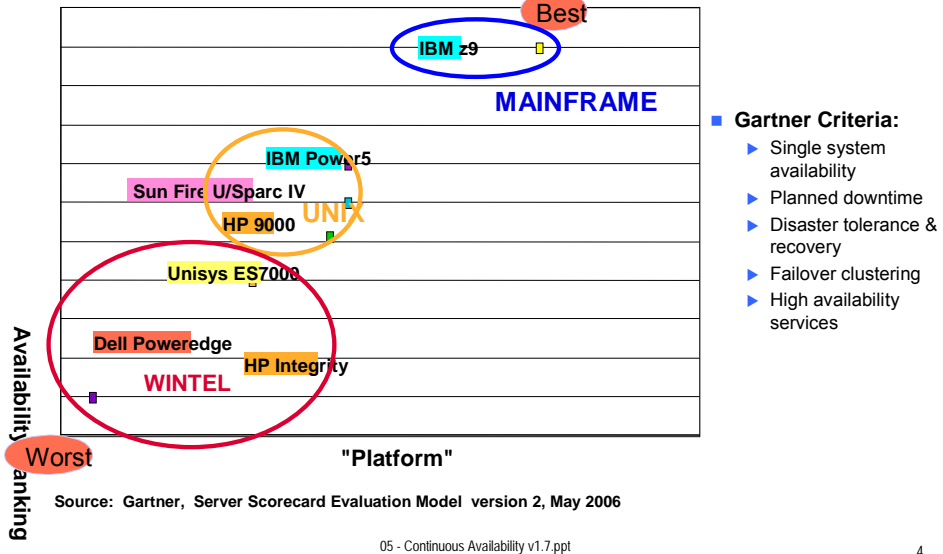
**IBM**

05 - Continuous Availability v1.7.ppt

3

# Gartner Ranks System z Tops in Availability

Availability Rankings- Selected Platforms



# Five Nines is the Gold Standard of Availability

- 99.999% availability = 5 minutes downtime per year out of 24x365
- 99.999% availability is sometimes referred to as “continuous operation”
- Survey of 28 companies with mixed environments
  - ▶ Average mainframe system availability = 99.993% or 36 minutes per year downtime
  - ▶ Average distributed server availability = 99.909% or 8 hours per year per server downtime
- Small improvements in the “nines” become more and more difficult to achieve
- Comprehensive design for availability is required for continuous operation

Downtime	mainframe	distributed	Cost impact
hours per year	.6 (99.993% availability)	7.98 (99.909% availability)	13 times downtime costs

March 12, 2007 IDC Survey of 28 customers with mixed environments

## How is Continuous Operation Achieved?

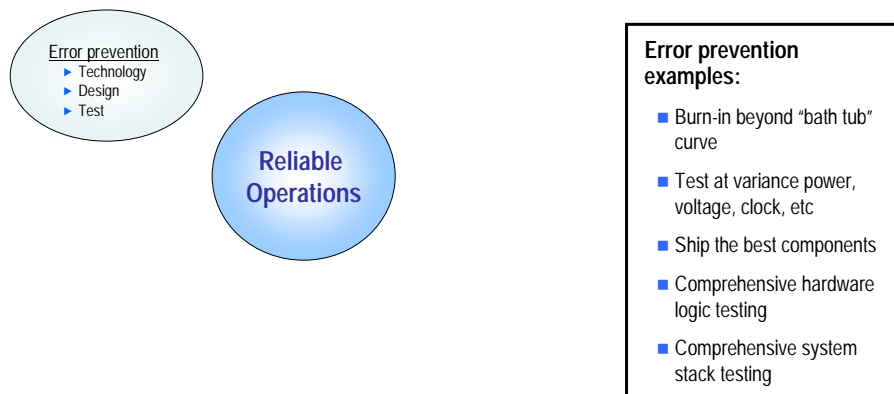
A comprehensive four part strategy is required:

1. Hardware reliability and serviceability
2. Cluster redundancy supported by middleware
3. Concurrent software maintenance eliminates planned outages
4. Disaster recovery

The mainframe parallel sysplex plays a key role in these capabilities

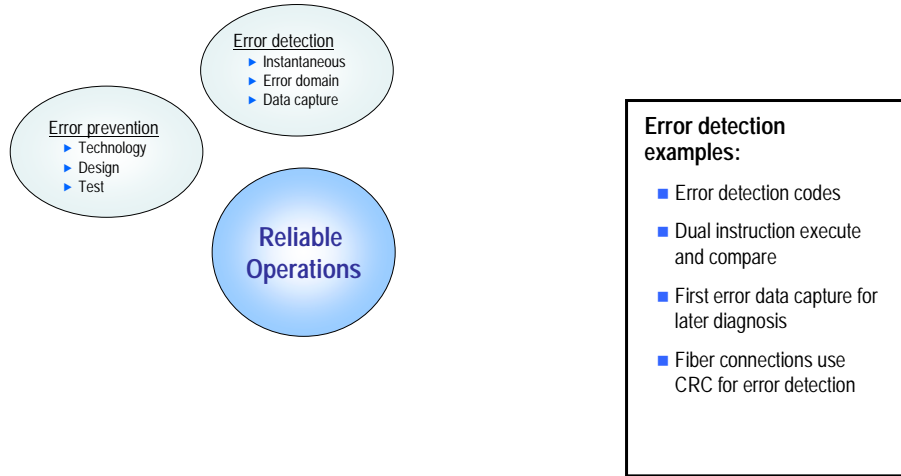
## A Comprehensive Design for RAS

### Examples of Hardware Reliability and Serviceability Features



# A Comprehensive Design for RAS

## Examples of Hardware Reliability and Serviceability Features

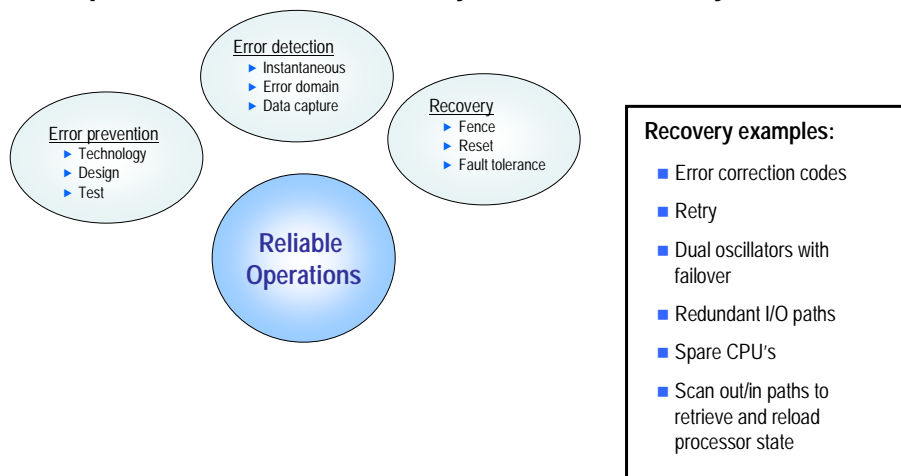


05 - Continuous Availability v1.7.ppt

8

# A Comprehensive Design for RAS

## Examples of Hardware Reliability and Serviceability Features

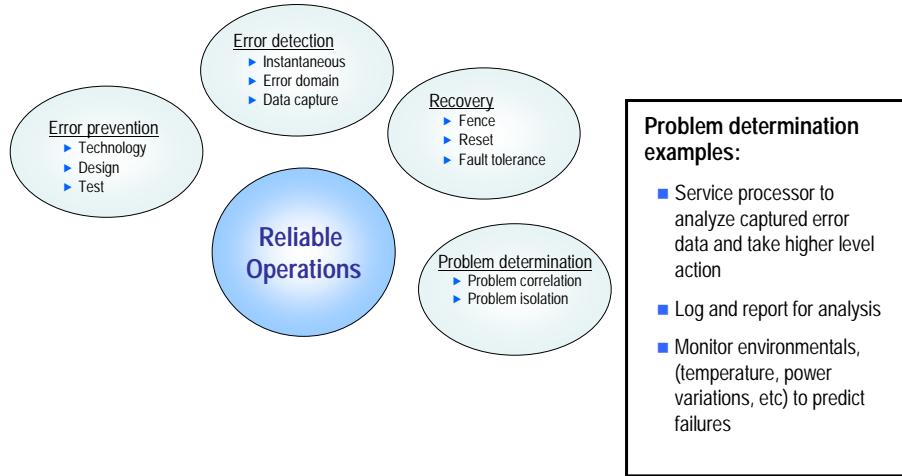


05 - Continuous Availability v1.7.ppt

9

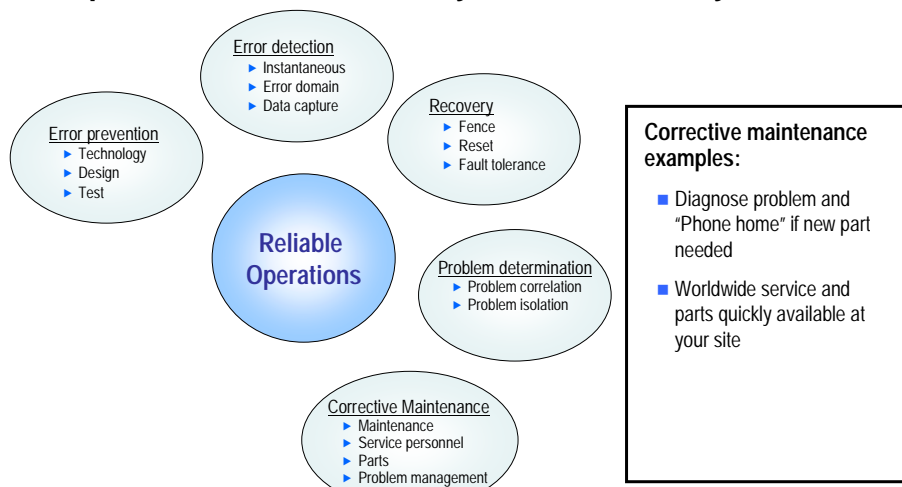
# A Comprehensive Design for RAS

## Examples of Hardware Reliability and Serviceability Features



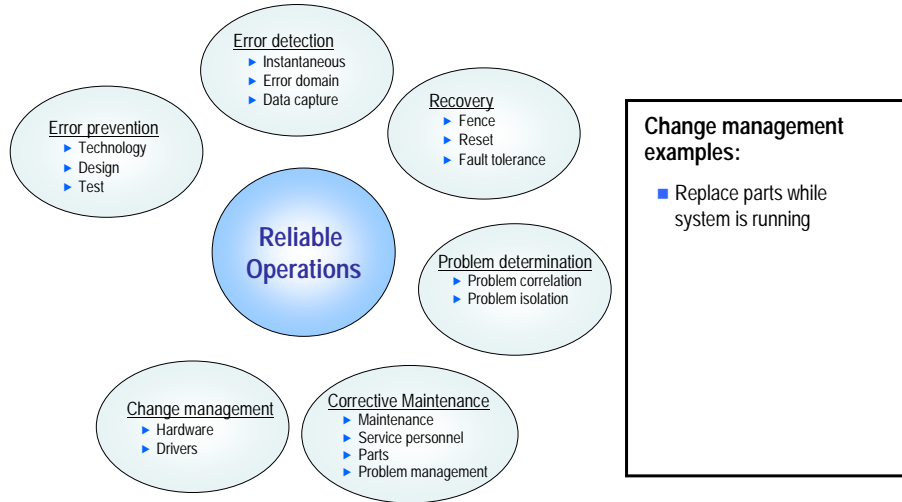
# A Comprehensive Design for RAS

## Examples of Hardware Reliability and Serviceability Features



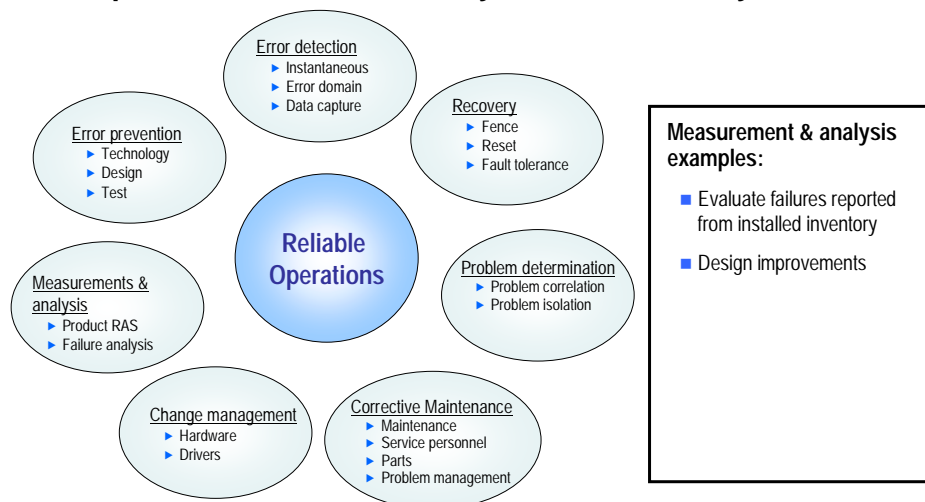
# A Comprehensive Design for RAS

## Examples of Hardware Reliability and Serviceability Features



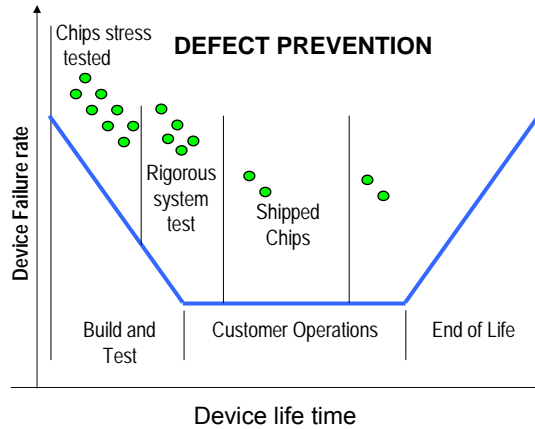
# A Comprehensive Design for RAS

## Examples of Hardware Reliability and Serviceability Features



## Rigorous Testing Beyond the “Bathtub Curve”

- Device Failures are more likely early in life
- Stress testing is designed to accelerate device aging
  - ▶ All chips are stress tested under extreme conditions
    - Burn in at 150% of the typical voltage- millions of test patterns
    - Power on hours tested at over 140 degrees Fahrenheit
- Result: Shipped chips have 10X more reliability than other industry standard components

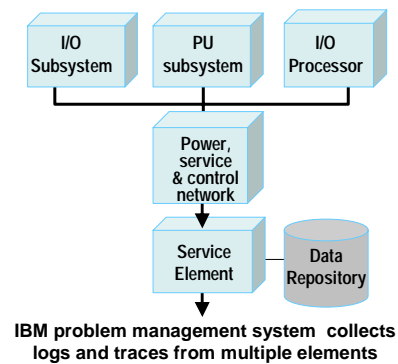


05 - Continuous Availability v1.7.ppt

14

## Error Detection: First Error Data Capture

- First error data capture (FEDC) is a diagnostic feature used to gather data for analysis when there is *any* failure in the system
- Early detection - up to 70% of all test-floor-found problems for the z990 were found through FEDC
- Customer Value - obtain the right data up front if a problem occurs, to avoid asking customers to re-create a problem
- All subsystems send logs, traces, and dumps to the service processor
- Common infrastructure to collect and store data



05 - Continuous Availability v1.7.ppt

15

## Spare Central Processing Units (CPU's) are Shipped Pre-Installed

- CPU Sparing –Transparently substitute a good processor for one that is suspected of being defective
- This on-line switch is seamless, applications are not impacted
- Sparing is performed if a CPU fails and instruction retry is unsuccessful
- CPU Sparing copies state information from the failed CPU to the spare
- The spare can begin executing at the exact same instruction where the other CPU failed
- The bad processor is fenced off
- Using the service element the checkpoint state is scanned from the failed SPU into the spare.
- The Spare CPU is tagged with the CPU address of the failed CPU

### Service Processor automatically finds a new "spare"

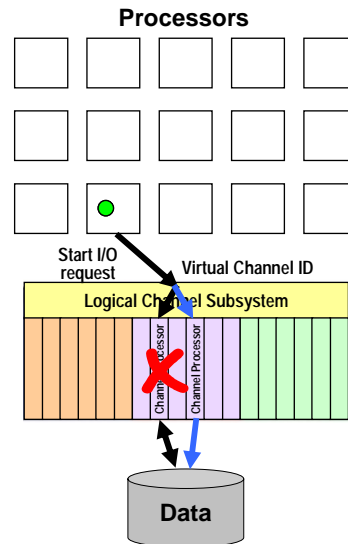
- Uses spare PU on the same book
- If none are available uses an adjacent book
- Uses any available spare in the server.
- If no spares are available, recovers the application onto *another* CP and automatically phones home for a replacement.

05 - Continuous Availability v1.7.ppt

16

## Virtualization of I/O Channels Enables Redundant I/O Paths

- Large number of physical I/O channels are available
- Logical channels subsystem provides virtual channel identifiers
- Virtualization enables an alternative physical I/O channel to be used
  - ▶ Dynamic channel paths
  - ▶ Logical channel subsystem
  - ▶ Protection against channel processor failure
- Recover I/O operations in progress
  - ▶ I/O subsystem transparently processes I/O with built in checks



05 - Continuous Availability v1.7.ppt

17



## Concurrent Hardware Repair and Upgrade Capability

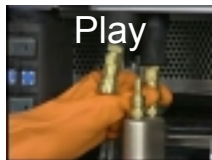
Capability	System z9
ECC on Memory Control Circuitry	Transparent While Running
Oscillator Failure	Transparent While Running
Microcode Driver Updates	Replace While Running
Book Replacement	Replace While Running
Memory Replacement	Replace While Running (Book Offline)
Memory Bus Adapter (MBA) Replacement	Replace While Running Connectivity to I/O Domain remains
Self Timed Interface Failure to I/O	Replace While Running Connectivity to I/O Domain remains
Processor Upgrades	Replace While Running
Memory Upgrades	Replace While Running
I/O Upgrades	Replace While Running
Spare CPU's	2 Pre-installed per System

05 - Continuous Availability v1.7.ppt

18

## DEMO: How Does Concurrent Hardware Repair and Upgrade Work?

- ▶ Example video of a memory upgrade while the system continues to run
- ▶ Service engineer dispatched through “phone home”
- ▶ Service engineer has part already ordered through our global parts replacement program
- ▶ Notice book is removed while the system is operational
- ▶ Even the service tray is included!
- ▶ Memory cards can be added easily similar to PC servicing



Types of Replacements:

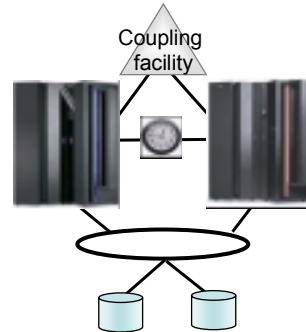
1. Add a single book for processors, memory, and I/O Connections
2. Remove and replace a book
3. Allocate physical resources on other books

05 - Continuous Availability v1.7.ppt

19

## 2. Cluster Redundancy Supported by Middleware

- Hardware reliability is not enough
- Mainframe systems can be clustered to further improve availability
- Workload is dynamically moved from a failed system to another system in the cluster
- What is required?
  - ▶ Parallel Sysplex Cluster technology
  - ▶ Middleware enabled to run in a parallel sysplex cluster with failover and recovery



05 - Continuous Availability v1.7.ppt

20

## Mainframe Middleware Exploits Parallel Sysplex for High Availability

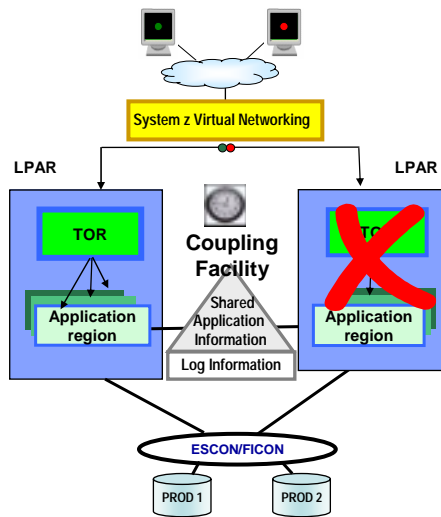
- Clusters of CICS, IMS, DB2 or MQ exploit parallel sysplex
  - ▶ Multiple images behave like a single, logical image via clustering
  - ▶ Up to 32 z/OS images coupled to one or more Coupling Facilities
  - ▶ The loss of *any one* image does not bring down the cluster
- Applications are enabled for data sharing to allow for workload balancing
- I/O subsystems support multiple I/O paths with dynamic switching
- The result is a fault-tolerant system:
  - ▶ Dynamic workload balancing; workloads run *on any image* in the sysplex
  - ▶ All images have access to all devices
  - ▶ Failover and recovery processes are fully automated

05 - Continuous Availability v1.7.ppt

21

## How CICS Exploits Parallel Sysplex for High Availability

- Dynamic routing of transactions to multiple images balances the CICS workload
- Images are designed for availability:
  - Terminal regions distribute traffic
  - Clones application regions balance the workload should one fail
- The workload of a failed image is dynamically routed to the remaining image
- State information, such as current log information, is used by the remaining image to speed recovery
- Sysplex Coupling Facility uses log information to resume processing where transactions left off

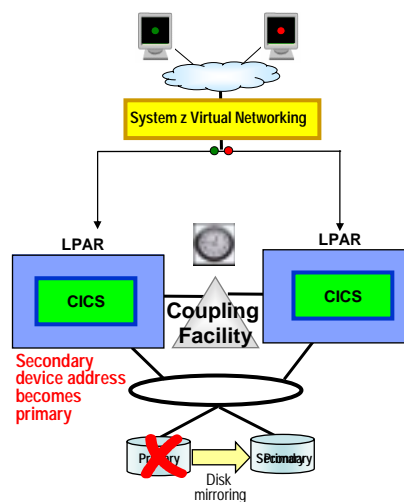


05 - Continuous Availability v1.7.ppt

22

## HyperSwap Can Protect Against Disk Failure and Data Loss

- Allows data to remain available to applications during disk failures
  - Disk subsystem failures are no longer a single point of failure for a Parallel Sysplex
- HyperSwap automates switches between primary and secondary copies of data maintaining data availability during disk failures
- HyperSwap can swap a large number of devices quickly with disruptions measured in seconds
- Enables planned maintenance without requiring applications to be quiesced
- No operator intervention needed:
  - Applications can transparently use same device addresses during failover

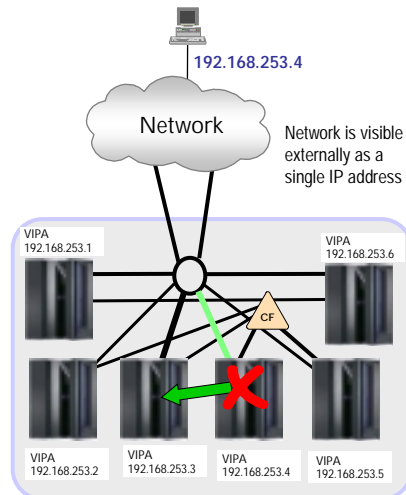


05 - Continuous Availability v1.7.ppt

23

## Virtual IP Addressing Provides Virtualization to provide Network Availability

- Virtual IP Addressing (VIPA) enables the use of IP addresses not tied to any specific physical interface
- Benefits
  - ▶ If a host suffers an unplanned outage, the entire stack is automatically moved to another host
  - ▶ For a planned outage, it allows non-disruptive relocation of stack to another host and a subsequent restore
  - ▶ Provides dynamic workload balancing through Workload Manager
  - ▶ Workload can be moved without operator intervention



05 - Continuous Availability v1.7.ppt

24

## 3. Concurrent Software Maintenance Eliminates Planned Outages within a Sysplex

- IBM middleware is designed to keep running during upgrades and patches
  - ▶ Utilizes parallel sysplex to take image offline
  - ▶ Enables multiple software versions to co-exist in the same sysplex
  - ▶ Enables rolling upgrades, one node at a time
  - ▶ Patches as well as releases can be applied without causing disruption to application availability

05 - Continuous Availability v1.7.ppt

25

## Avoid Planned Outages with Operations Designed for Continuous Availability

- **CICS**
  - ▶ **CICS backup while open** allows for file backup while CICS has it open for update.
  - ▶ **CICS RDO** dynamically add or change resource definitions and have them immediately usable without the need for a scheduled outage.
  - ▶ **CICS Auto-install** provides the ability for support resources to be defined dynamically on their first use
- **IMS**
  - ▶ Update Type 2 and Type 4 SVC's without requiring an IPL
  - ▶ IMS V9 Dynamic Resource Definition
  - ▶ IMS V9 provides a dynamic resource manager which is implemented without having to IPL
  - ▶ IMS ACBLIB online change
- **DB2**
  - ▶ Backup
  - ▶ Image copy
  - ▶ Add or modify a table schema
  - ▶ Reorganize the database
  - ▶ Partition a secondary index

05 - Continuous Availability v1.7.ppt

26

## 4. Systematic Disaster Recovery – The Last Layer of Protection

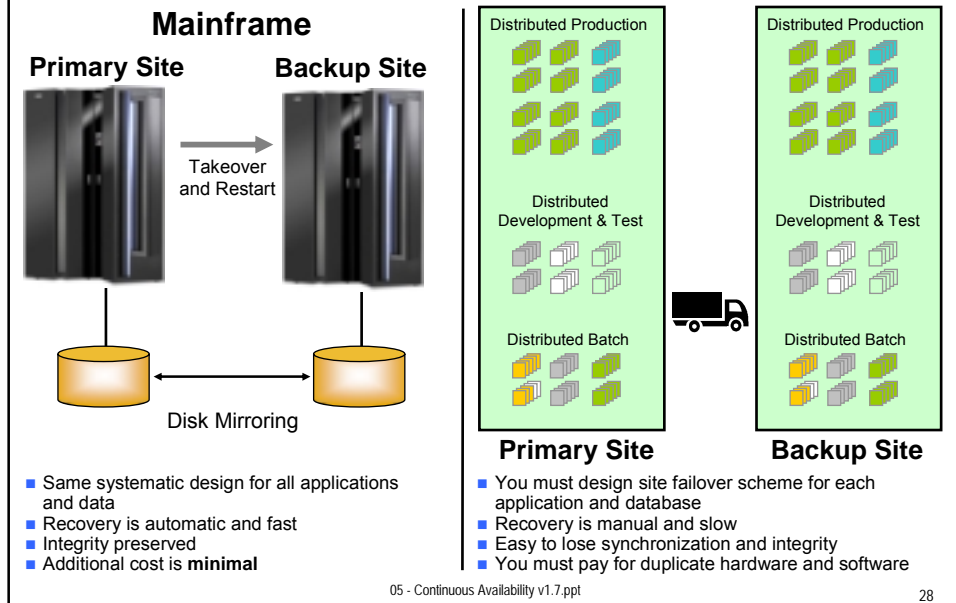


- What's needed:
  - ▶ Backup site capable of sustaining business operations
  - ▶ Automated, systematic site failover with data integrity preserved

05 - Continuous Availability v1.7.ppt

27

## Disaster Recovery on the Mainframe vs Distributed



## Geographically Dispersed Parallel Sysplex (GDPS) – Disaster Recovery

- GDPS provides systematic site failover capabilities in the event of data center disaster
  - Uses disk mirroring to failover site
- GDPS supports different recovery time requirements and different site distances with several options available
  - Metro Mirror
  - Global Mirror
- In addition to storage, sysplex, system software, GDPS provides automation of site failover tasks to expedite recovery
- GDPS is comprised of IBM products and technologies delivered along with service components

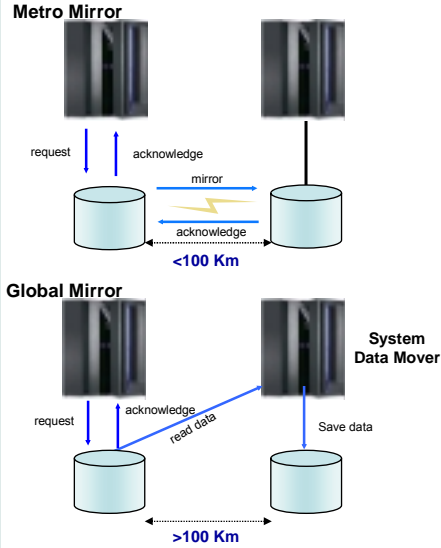
# GDPS Metro Mirroring versus GDPS Global Mirroring

## Metro Mirror

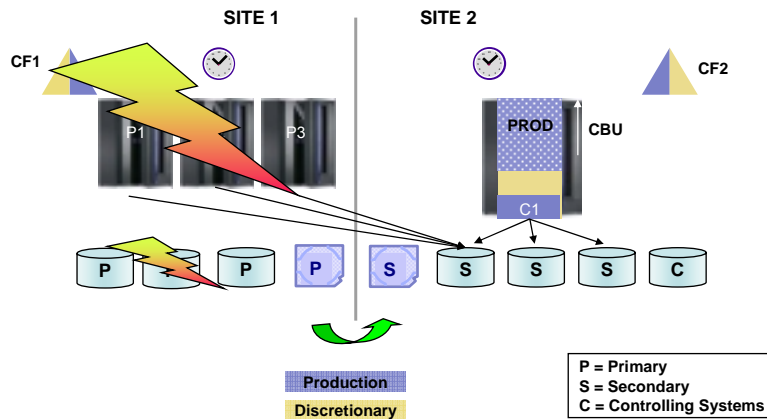
- **Synchronous** remote data mirroring
  - ▶ Application knows "I/O is complete" when both primary and secondary disks are updated
- Supports metropolitan distance (<100KM)
- GDPS freezes the image of the secondary data at the very first sign of a disaster

## Global Mirror

- **Asynchronous** remote data mirroring
  - ▶ Application knows the "I/O is complete" as soon as primary disk is updated
- Unlimited distance support
- Performance impact negligible
- System Data Mover is used to move data from Primary to Secondary disk



# Multiple Site Workload - Cross-site Sysplex Continuous Availability Configuration



Restarts failed images in Site 2  
 Designed to support continuous access to data from site 1  
 Failover with PPRC

## How is Continuous Operation Achieved?

A comprehensive four part strategy is required:

1. Hardware reliability and serviceability
2. Cluster redundancy supported by middleware
3. Concurrent software maintenance eliminates planned outages
4. Disaster recovery

## TD Bank Best Practices



### ■ Background

- ▶ TD Bank has been running Parallel Sysplex
  - no Sysplex-wide outage for **10 years**
- ▶ System z is used for Customer Account Data for applications supporting Tellers, Internet Banking and ATMs

### ■ TD Bank Recommendations

- ▶ Keep sysplex up – do not bring it down
- ▶ Practice Rolling IPLs
- ▶ Exploit concurrent hardware upgrades
- ▶ Use automation
- ▶ Configure your sysplex for availability
  - IMS/DB2 Data-sharing
  - Transaction routing
  - Sysplex Distributor for TCP/IP
  - Online database reorganizations
  - Clone each image
  - Ensure applications exploit parallel sysplex

#### ➤ Client Environment

- System z
- z/OS
- DB2
- IMS
- WMQ
- GDPS

Parallel Sysplex Deployment consists of five System z across two sites running 42 M business transactions a day



## Service Oriented Finance Needs a World Class Availability Solution



Service Oriented Finance  
CIO

Other vendors who fail to deliver this four part strategy will fall short of System z continuous availability



IBM

