E89

# IM36 IMS Disaster Recovery Hardware Techniques

David Moore (mooredw@us.ibm.com)


**Las Vegas, NV** **September 15 - September 18, 2003**



**RETURN TO INDEX**

# Trademarks

IBM Trademarks or registered Trademarks

IMS/ESA, IMS

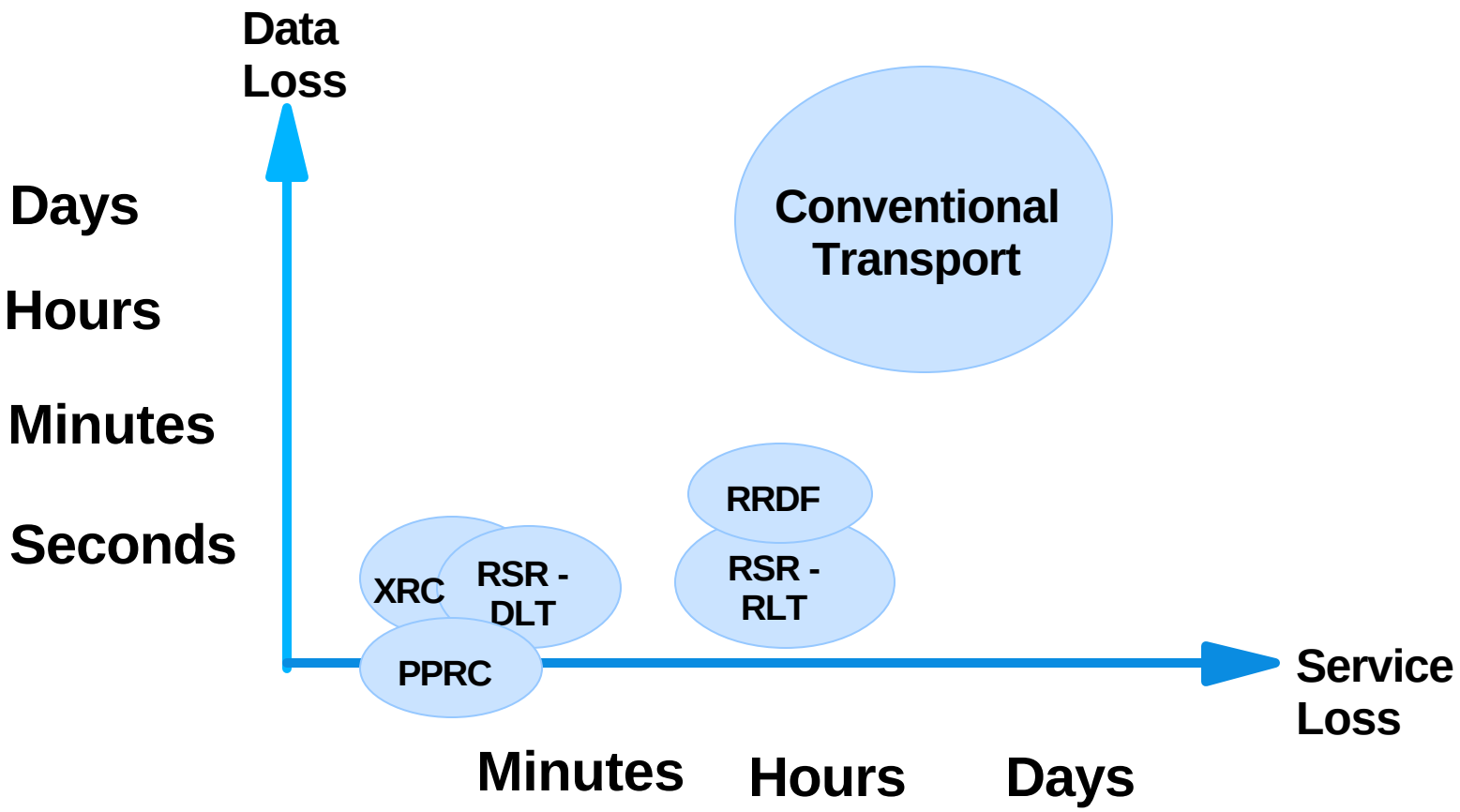
Other Trademarks

ENET - ENET Corporation

# Topics

- Software Approaches
  - RSR
  - E-Net
- Hardware Approaches
  - XRC Hardware
  - PPRC Hardware
  - GDPS Implementation
  - Other DASD
- IMS Tuning Recommendations

# Data Loss and Takeover Speed



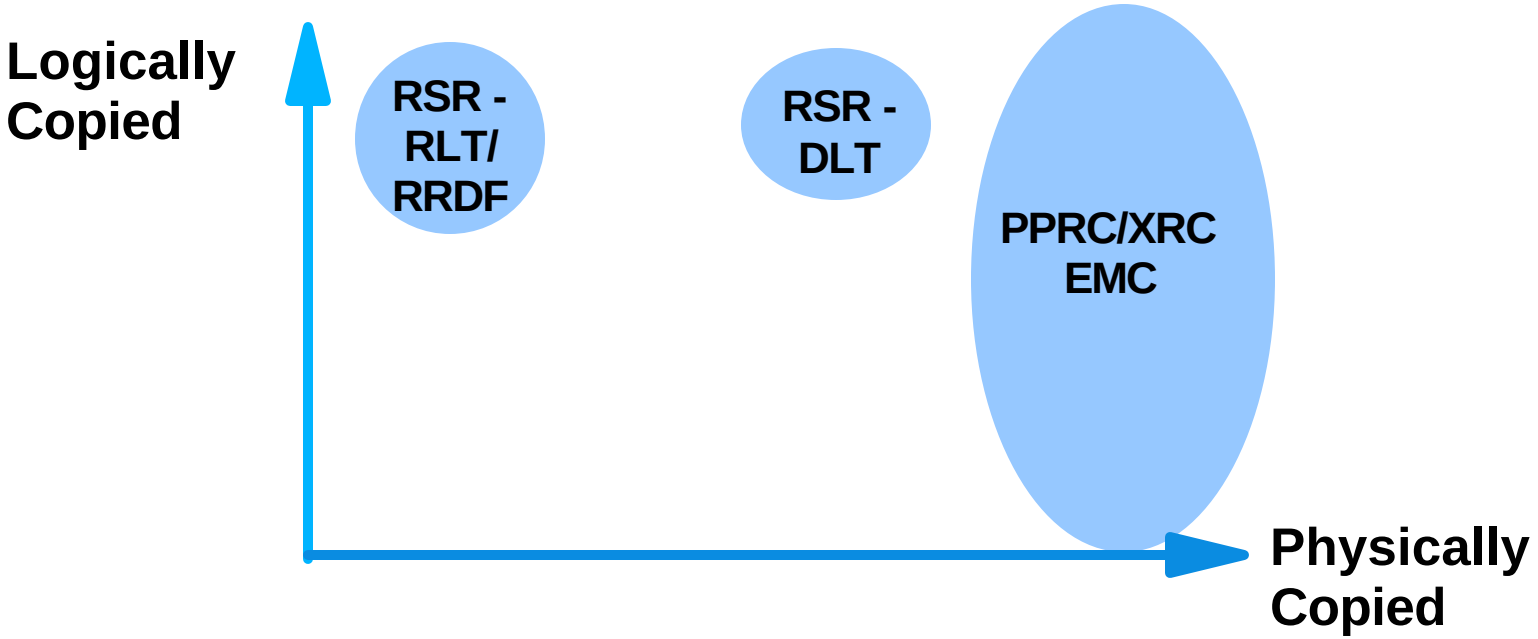
# SHARE Defined Tiers

- **Tier 0 - No Disaster Recovery Plan**
  **All data lost, no recovery possible**
- **Tier 1 - Pickup Truck Access Method**
  **Secured backup but no D/R site. New data after backup is lost**
- **Tier 2 - PTAM and Hot Site**
  **As Tier 1 but with D/R site. Recovery is 24-48 hours**
- **Tier 3 - Electronic Vaulting**
  **Remote tape library. Can reduce data loss and recovery window**
- **Tier 4 - Electronic Remove Journalling/Logging**
  **Small data loss. Applying updates minimizes recovery time**
- **Tier 5 - Two Site Two Phase Commit**
  **Application controlled. Secondary outage affects Primary**
- **Tier6 - Zero Data Loss (Synchronous Remote Copy)**
  **Zero data loss if Synchronous, seconds if Asynchronous**
  **Rolling failure means inconsistent data at secondary site**

# Data Loss and Takeover Speed

- RSR Offers two options"
  - RLT where log data is transported and RECON states are maintained
  - DLT where DBs are shadowed in real time
- RRDF from ENET offers RLT type functionality
- XRC is a hardware synchronous copy methodology
- PPRC is a hardware synchronous copy methodology
- EMC DASD has synchronous and semi=synchronous copy methodology
- Traditional methods send copies of logs and ICDSs to the remote site
  by overland transport
- ALL asynchronous methods risk the loss of some in-flight data

# Logical Recovery Vs Physical Recovery




Logical copying

all data needed for recovery is available and logically in synch, even if back level

Physical copying

individual physical devices are copied, but not necessarily in logical synchronization

IMS requires data be logically in synch


*Notes*

# Logical Recovery Vs Physical Recovery

It is important that the state datasets for IMS be logically correct at the remote site.

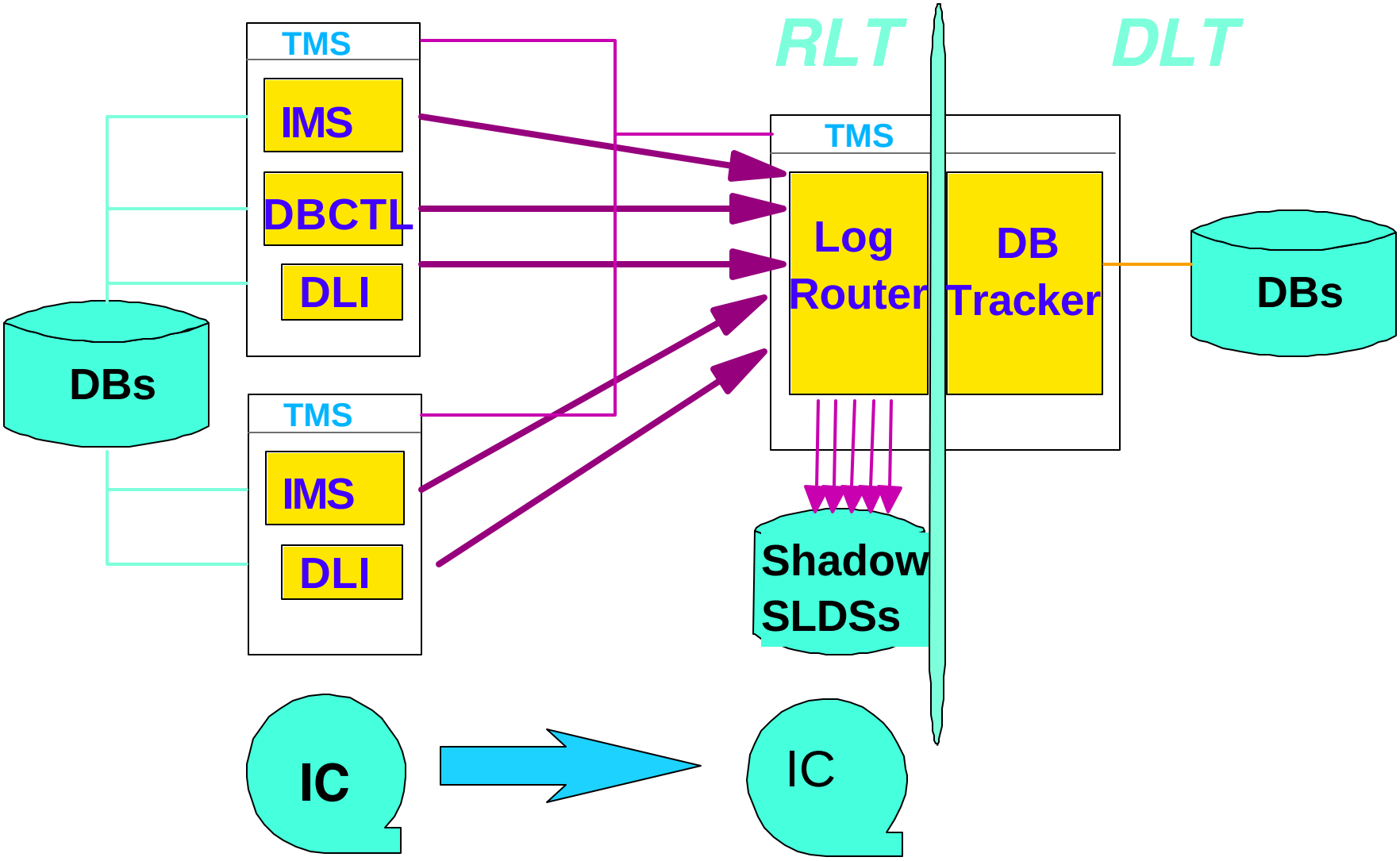
By that I mean that the RECON at the remote site must accurately reflect the DB states at the remote site, NOT at the active site. The same is true for OLDS datasets. Special data events like log archiving must be handled properly at both sites. For this reason hardware copy methods must be configured such that all the data or none of the data is transferred, including WADS.

Assess all solutions against this measure, especially failure scenarios.

# Special Concerns Areas

- WADS
- RECON at the alternate site
- Logs
- ARCHIVED LOGS
- System datasets (ACB,MFS, PGM libs, etc.)
- Online Change activity
- anything important that is on tape

# RSR Overview



# RSR Overview

- RSR Comprises
  - The Transport Manager  in each MVS
  - The Tracker System at the Remote site
  - Appropriate log data sets and RECON data sets
- The TMS assists the IMS subsystem loggers in establishing APPC conversations on which log buffers are sent at OLDS write time
- User Exit driven by TMS filters log records that are of no interest such as DB not covered, trace records, etc.
- Log Router produces shadow SLDS copies at the Tracking site for RLT
  - Separate logg stream for each Active subsystem
- Log Router passes DB log records to Database Trackers for DLT
- Image Copies are transported to the Tracking site outside of RSR
- Tracking site RECON data sets maintain recovery status for the Tracker

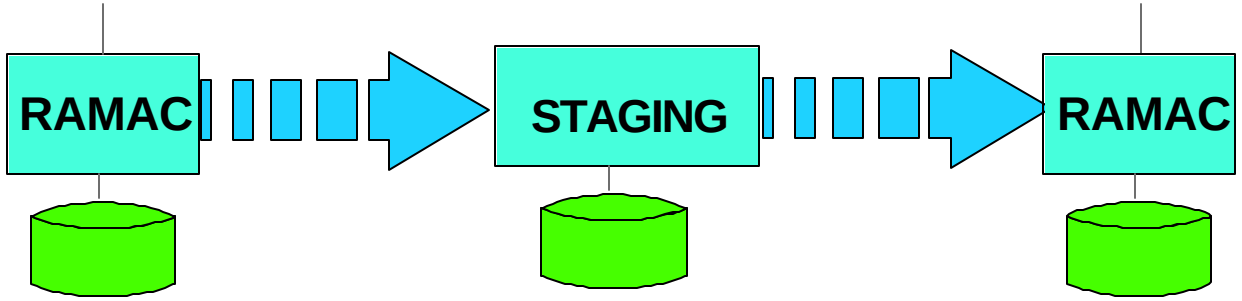
# Other Software Transport Products

ENET

- Tier 4
- Functionality similar to IMS RLT
- Support for IMS and DB2
- Maintains IMS RECON DB Info
- Handles Transmission Outages
- Understands Logical vs Physical Copy
- Internally Differs markedly from RSR

ENET was marketed as IMS RRDF

# eXtended Remote Copy - XRC




- Allows time stamp consistent copying of multiple DASD volumes to a remote site in real time.
- IMS recommendation is copy OLDS, WADS, DBs, RECON - all or nothing.
- IMS recommendation of ERRORLEVEL=SESSION to ensure both logical and physical views at the remote site.
- Use multiple redundant ESCON
- Enlarge the OLDs capacity to avoid remote site need for tape SLDS.
- Make sure you do ICs after a disaster since you may not have enough log data to go back to the oldest valid IC without archived log data.


# eXtended Remote Copy - XRC

*Notes*

- XRC provides asynchronous, multiple DASD, transport to a remote site that will keep all packs at a time consistent state. Because the copying of the IMS state datasets (like RECON) must be also LOGICALLY valid it is the IMS recommendation that you copy all or nothing for IMS.
- XRC works for DASD not tape, so Log Archiving needs to be addressed.
- Make sure you don't need archived logs for recovery after a disaster or make sure you transport a copy of archived OLDS to the disaster site.
- Enlarge your OLDS/SLDS DASD capacity.
- In the case of a disaster you need to ensure you have current image copies of all DBs.
- Remember that you only have the log data that is currently on DASD unless you've added procedures for archived logs.
- Any recoveries that need archived log data may have problems - the reason to have a large OLDS/SLDS capacity on DASD.
- There is support available for mutual data movers. IF YOU USE THIS you need to use something like GDPS to ensure that all data from the plex is kept in time sequenced order and that data is moved on an all or nothing basis.


# XRC Pros & Cons

- Tier 4

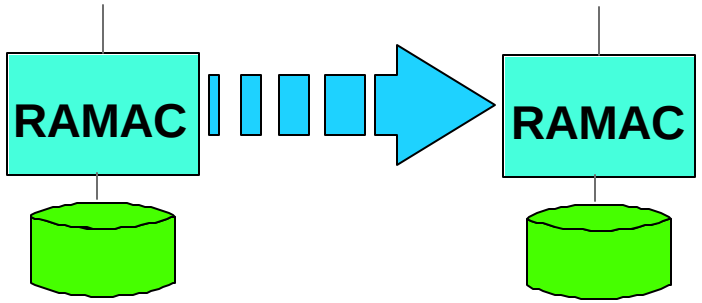
- Pro
  - Hardware only solution and works on any type of DASD content.
  - Application independent
- Con
  - Requires more transmission bandwidth (10x)
  - Does not handle IMS log archiving, but solutions can
  - ESCON (Channel Extender) distances only
  - You may have DBs that lack enough log data to be recovered until you take new image copies


# XRC Pros & Cons

*Notes*

- Because XRC will transmit entire altered blocks (not just the altered data) the bandwidth can be substantially larger than with RSR.
- Log archives are a problem for both IMS and DB2.
  - The IBM TotalStorage Vitual Tape Server duplexing function may help
- Easier to operate than a software approach
- It works for all DASD data
- CF needs consideration. You may be able to duplex Shared Queues sturctures and the MVS Logger Structure.


# Peer to Peer Remote Copy - PPRC




- Allows control unit level synchronous copying of DASD data to a remote site.
  - NO DATA LOSS in a disaster.
- Because active site writes must wait for remote site duplication there is an unavoidable active site performance impact for both throughput and turnaround.
- Remote Write failure results in Write failure reported to MVS
- IMS recommendation is CRIT=Y to ensure an all or nothing copy mode.
- Much of the response time degrade can be tuned out with additional WADS size, more dependent regions, more OLDS datasets, etc.


# PPRC Pros & Cons

- Tier 6

- Pro
  - Only synchronous method and only method to ensure no data loss in a disaster
  - Hardware only and minimal interaction with operators.
- Con
  - Impacts IMS throughput and turnaround
  - Unable to handle logs archived to tape
  - Requires more transmission bandwidth than RSR or Enet (10x)
  - ESCON distances only
  - CF needs consideration. You may be able to duplex Shared Queues sturctures and the MVS Logger Structure.


# PPRC Pros & Cons

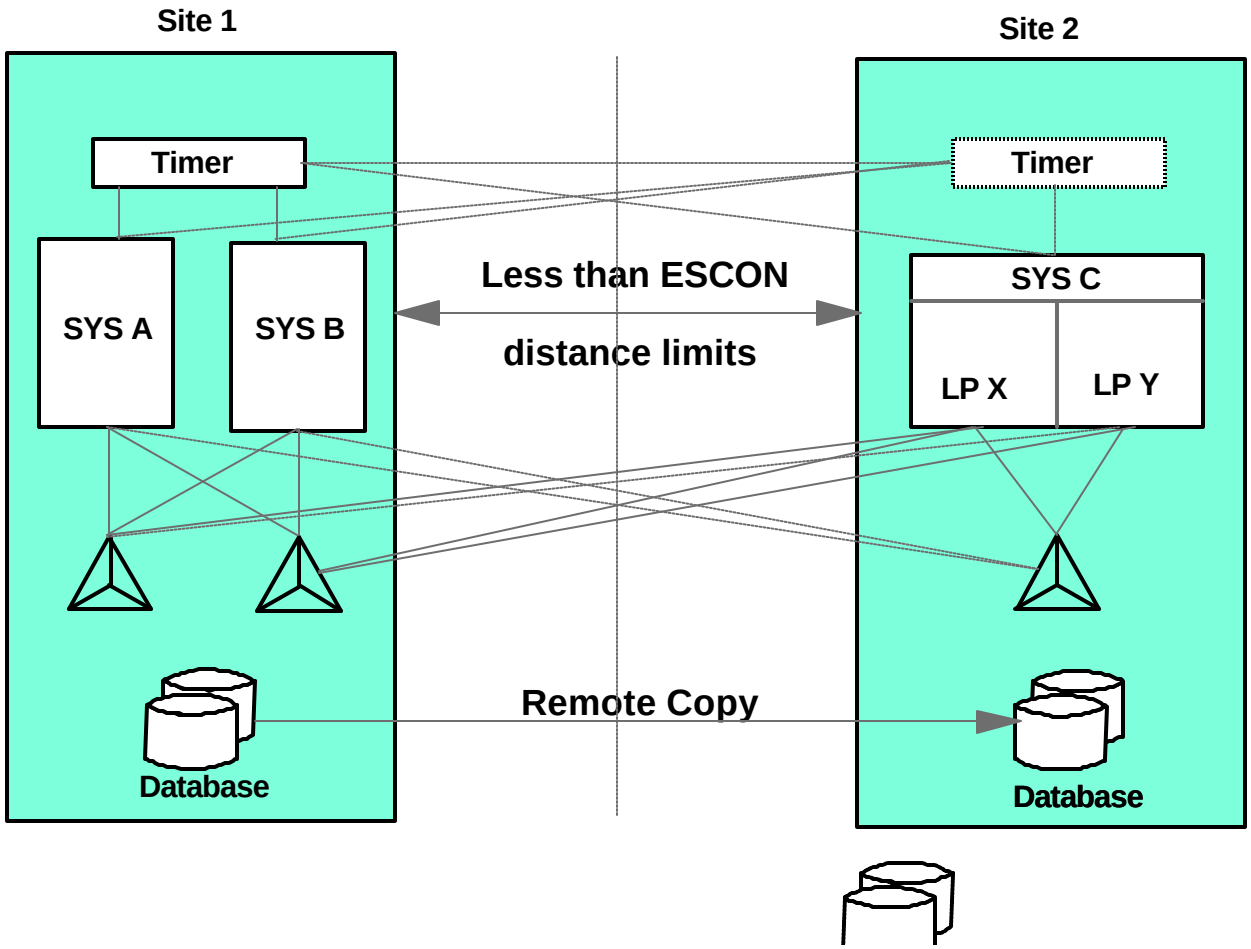
*Notes*

- Guarantees no lost data.
- Easier to operate than software solutions
- If the link between sites goes down - the active site is unavailable
  - (but see Geoplex later)
- Log archives are a problem for both IMS and DB2.
  - The IBM 3494 Magstar Vitual Tape Server duplexing function may help
- Like XRC you need more bandwidth than with RSR
- Like XRC you need to ensure you have image copies after a disaster.
- If your data is on multiple controllers you need something (like GDPS) to ensure all controller work in unison and that ALL data is in time sequenced order

# Geographically Dispersed Parallel Sysplex (GDPS)

Service offering based on enhancements with PPRC (also supports XRC environments)

- Automation expects/tolerates rolling failures




Notes

# Geographically Dispersed Parallel Sysplex (GDPS)

★ Additional PPRC operating characteristics when remote write failures occur

- FREEZE & GO - Tier 6
  - stops all secondary site writes but continues primary service
  - preserves consistency status at secondary site for later switch
  - backlevel data unless PPRC resync
- FREEZE & STOP UNCONDITIONAL - Tier 4
  - stops all writes (primary and secondary) and switches
  - interrupts service until switch complete but no missing data
- FREEZE & STOP CONDITIONAL - Tier 4 or Tier 6
  - stops all secondary writes if secondary site write error, otherwise switches
  - switches if problem is not with write into secondary control unit
  - preserves consistency status at secondary site for later switch
  - backlevel data unless PPRC resync


# GDPS Pros & Cons

- Tier 6 plus or Tier 4

- Pro
  - Based on PPRC or XRC but tolerates rolling failures
  - Hardware only and minimal interaction with operators
  - Application independent
  - Allows multiple controller PPRC or multiple datamover XRC to work with DB sysplexes like IMS and DB2
- Con
  - Impacts IMS throughput and turnaround
  - Unable to handle logs archived to tape
    - The IBM 3494 Magstar VTS duplexing function may help
  - Requires more transmission bandwidth than RSR (10x)
  - ESCON distances only with PPRC
  - Must restart on a Parallel Sysplex
  - GDPS using PPRC gets messy with partial failures


# GDPS Pros & Cons

*Notes*

- Easier to operate than software solutions
- Installation choices for failure management
- Applies to all applications
- If the link between sites goes down - the active site can continue
  but Primary data may get marooned (like asynchronous)
- Like XRC/PPRC you need more bandwidth than with RSR
- Like XRC/PPRC you need to ensure you have ICs after a disaster.


# Other Hardware Transport options

- Non-IBM hardware offerings are available
- Both Synchronous and Asynchronous offerings
- Same IMS recommendation - copy all or nothing
- Same recommendations as PPRC/XRC - all controllers and DASD must act as a common unit and stay in time sequence order.

# IMS Restart at the Disaster Site

- Hardware methods (PPRC, XRC, OEM)
  - Like you pulled the power plug at your regular site (only nicer)
  - You'll need to emergency restart ALL IMS systems that were running at the originial site and let them do cleanup and backout.
  - After ere cleanups you can shutdown what you want.
  - If you don't copy all DBs then you will get and have to ignore backout failure messages

# General IMS Items

- WADS
  - Many of the hardware methods can substantially increase WADS response time
    - In several customer implementations it was shown that impact on actual IMS response can be largely tuned out and may even be self tuning due to WADS architecture.
  - Several initial plans skipped WADS and had manual methods.
    - Our reviews generally found serious holes in the plan
    - IMS recommends you do NOT skip WADS
  - Biggest impact environment is MSC networks
    - MSC environment has more serialization and is therefore more impacted
    - Use of Shared Queues function for intra plex can eliminate impact for this inter IMS messaging
    - MSC VTAM parallel session using DFSNRPT0 routing exit can improve inter plex communications. IMS V6 increases SYSID number limit in support of this option.
  - There are customer implementations where end user response time degrades have been reduced to under .1 second.
  - Larger WADS can often help = 100 cylinders is our general recommendation or..... (#_of_logbuffers * blksize)/2k = max tracks really needed.

# OLDS, DASD, Config

- Cached and Virtual DASD can substantially cut down on hardware copy impacts, but remember that this will generally increase the risk of data loss
- Higher log buffers doesn't necessarily help.
  - I/O architecture write of 10 buffers is the max (around 260K).
  - 400 buffers appears to be the top of the curve for sustainable rates. More is useful for short spikes only.
  - Blocking at 1/2 track (26K for 3390) recommended
- More dependent regions may NOT help
  - 50-60% occupancy is good
  - Larger numbers of regions increases general scheduling and dispatching overhead


# Additional IMS Tips

- If you use IMS RSR at the Version 8+ level you can support synchronization with XRC copied DB2 data
- Electronic copying of 1st image copy after any reorg or DB load ensures critical image copies are at both locations
- Low cost transport of routine image copies is OK since you don't need double copies of every IC at both locations
- Enlarge  DB2 log capacity to avoid need for archived data after a disaster
- Don't penny pinch on data transmission bandwidth. You need to handle catchups and surges.


# Testing and Planning

- Testing is difficult at best
- Testing the return is even tougher and is generally more complex
- Plan and test for disaster, slow disaster, and non-disaster


# Summary

- Viable software only and hardware only solutions are available
- Software only solutions can be effectively improved with hardware copies of related datasets
- Hardware solutions are now mature and have been implemented at actual IMS customer locations
- Archived logs are a problem for hardware only solutions


# Summary

*Notes*

- Don't forget your other required files.
- Think about your ICs and archived logs
- Understand that money is an unavoidable part of your decision.
- Everyone would like to have zero data loss and zero availability loss.
- Your actual solution will depend on a combination of your needs and your resources.


# Bibliography

GG24-4210 Disaster Recovery Library : S/390 Technology Guide

Discusses products, communications and systems management

GG24-4211 Disaster Recovery Library : Design Concepts

Discusses concept, options, strategy and implementation

GG24-3993 Disaster Recovery Library : Database Recovery

Discusses DBMS data (IMS, CICS/VSAM and DB2)

GG24-3994 Disaster Recovery Library : Data Recovery

Discusses non-DBMS data (MVS and infrastructure data)

GG24-2595 Planning for IBM Remote Copy

Discusses 3990 PPRC and XRC

GF22-5114/GF22-5063 White Papers on GEOPLEX

SG24-5338 RAMAC Virtual Array: Implementing Peer-to-Peer Remote Copy

RETURN TO INDEX