

The datacenter of the future

Matthias R. Bangert

matthias.bangert@de.ibm.com

Principal IT-Specialist

IBM zSystems Worldwide Technical Sales

Tailored Fit Pricing, Cyber Resiliency, z Hardware

Karen Smolar

ksmolar@us.ibm.com

Principal Solution Architect

Client Engineering for Systems

Agenda topics

General

- Demand
- Sustainability
- Performance
- Simplification

Availability

- Status
- Combined HA and DR
- 2-sites vs. 3 and more sites
- Future items

DASD

- Response time
- Essential features and functions
-

TAPE

- Which data should be on TAPE ?
- TCT

Next Steps

- Get a clear status of where YOU are on the journey
- Compare your status with your goals
- Prioritize activities

If you do things like you always did them - why do you expect different results ?

General

To start with, let me ask you a couple of questions

Example: DASD Performance (Sysplex ?)

- 36 years ago, when I started as SYSPROG on Mainframes, everything below 20msec DASD response time was considered “GOOD” (3880 controller with 3380 disks).
- Today, everything better than 0,5msec is considered “GOOD”. But nowadays we are talking about “all flash”, 32GB Ficon etc.
- We have not found a way to increase the “speed of light” (yet). Which means every kilometer distance between your controllers (for PPRC) adds 10µsec. If you have 10 kilometers distance between primary and secondary datacenter this is $10\text{km} \times 10 \mu\text{sec} = 100\mu\text{sec}$ or 0,1 msec (best case).
- If you run an “active - active” Sysplex configuration the remote machine needs 0,2 msec.
- Do you think this will be “good enough in the future ?

To start with, let me ask you a couple of questions

Logical Corruption

- A newspaper in Germany did a survey among 400 medium size businesses. 2/3 of them got hacked, 42% out of the 2/3 (112) paid the ransom to gain access to their data.
- What would you do in such a case ?
- Do you have the infrastructure to test different recovery approaches ? Or would you rather pay the ransom ?

“D/R datacenter” or “Second Production Site”

- Would you run for an extended period of time in your D/R datacenter - just for test ?
- Can you provide the same quality of service in your D/R datacenter ?

Recommendation:

To get prepared for future demands

- **Ask yourself**
 - Is my infrastructure able to scale
 - DASD Response Time
 - LPAR Size
 - Number of systems per CEC
 - Overhead (Sysplex, Capture Ratio,
 - Did I miss the right “switch point” to continue on the right track ?
 - Am I able to run my mainframe in a couple of years once the “old guys” retired ?

Availability

D/R and High Availability Considerations

The last 30 years

- Introduction of datasharing parallel Sysplex (single site first, multi site later)
- Introduction of PPRC and XRC around 1995
- GDPS got introduced in Fall 1998 to automate LPAR re-ipl and manage DASD
- RPO of Zero is ok, and RTO of several hours was considered ok.
- It seems “logical corruption” was not invented back then

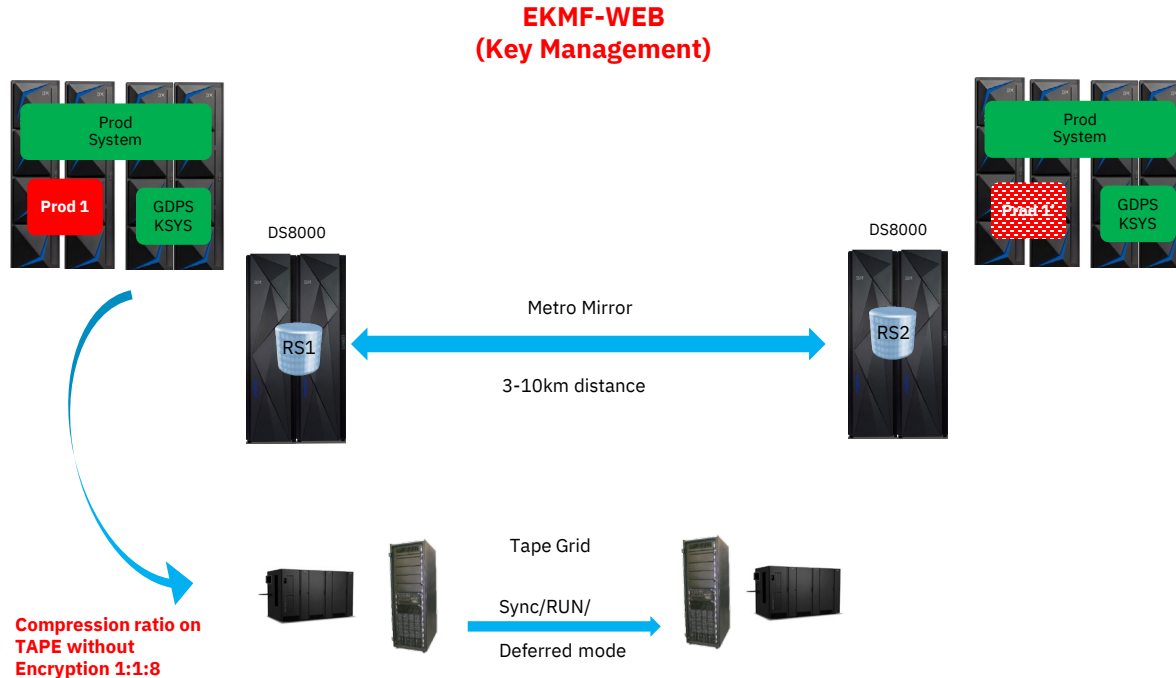


Cover Hardware outages (component or site) and software problems.

Sorry for being very blunt: This is not good enough in today's world

- If you don't believe me - ask your kids about “availability” requirements
- If I see my bank announcing a “maintenance window” and therefore “service not available” I am getting nervous - seriously.
- Logical corruption is a very serious threat in these days. And if you think “not on the mainframe” - it is the same story as for security. The mainframe is the most “securable” platform

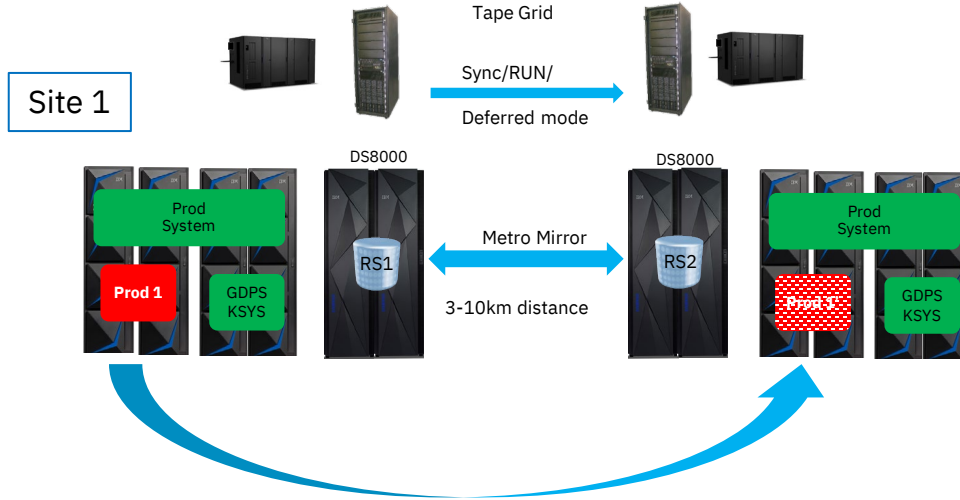
Encryption / Compression



Encryption / Compression

- Compression on the storage systems is fine - as long as the sole use case is “safe data once storage controller gets replaced”
- If you want to do a “better” encryption - it must run on the host to protect data in flight, on the network, in the CF.
- If you decide for encryption on the host - compression must be done before encryption (encrypted data cannot be compressed). Therefore, storage controller compression is useless for encrypted data.
- You also need to “manage” your keys when doing encryption.
- **Using host compression means you are prepared for any future decision towards encryption.**

Yesterdays / Today's topology

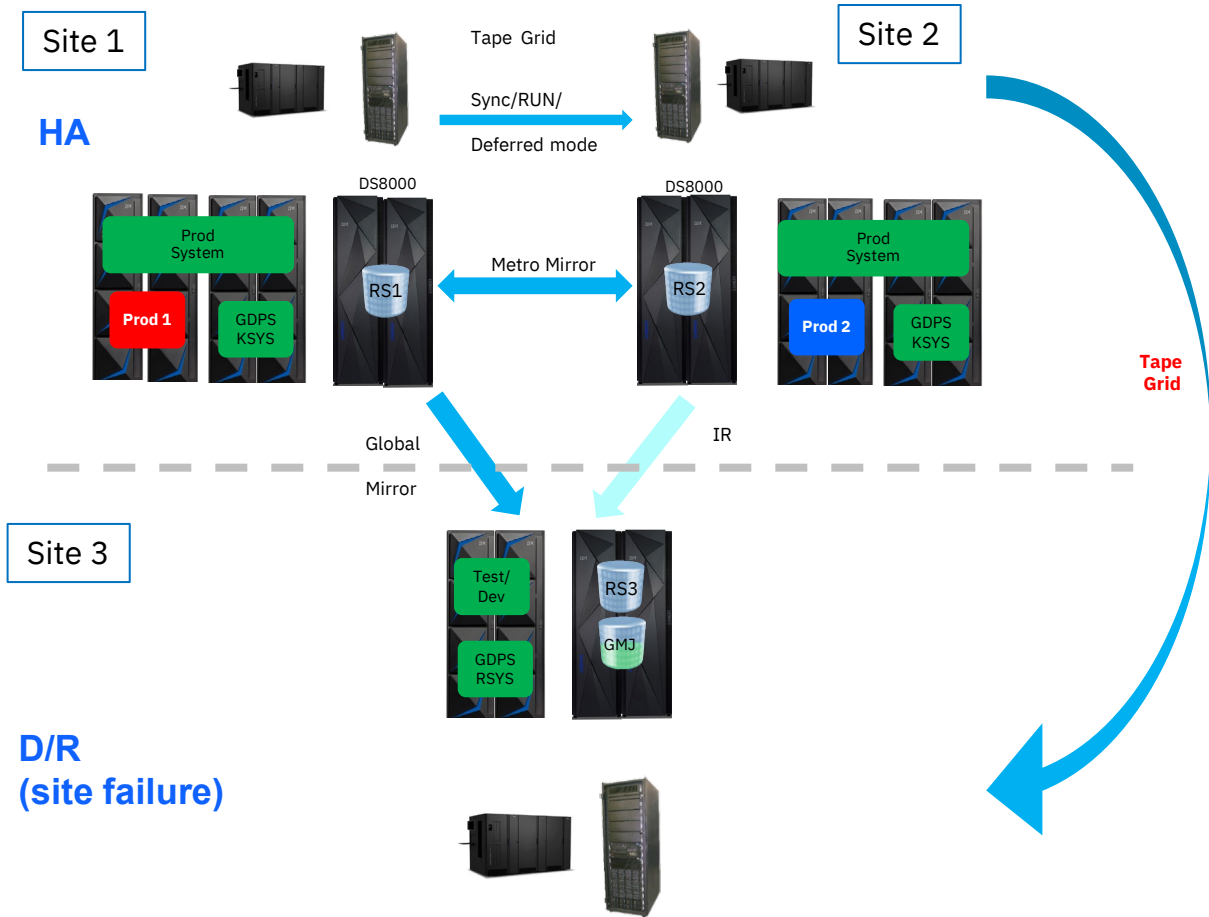


- LPARs get moved from one datacenter to another in case of a CEC failure
- This results in at least 30 minutes downtime until service is restored
- DASD response times can hardly be improved - Hyperlink, read from secondary is the best you can do.

Combined HA and DR

- ✓ Save infrastructure money because of combined HA and DR
- ✗ Increased Sysplex and DASD response time overhead
- ✗ No preparation for logical corruption recovery
- ✗ CEC or LPAR Failure results in longer downtime as LPARs get moved / reipled
- ✗ Regional disasters (flooding, fire, power outage, earthquake,...) are not covered
- ✗ In many countries a remote datacenter is either required (US) or "strongly recommended" (Germany)

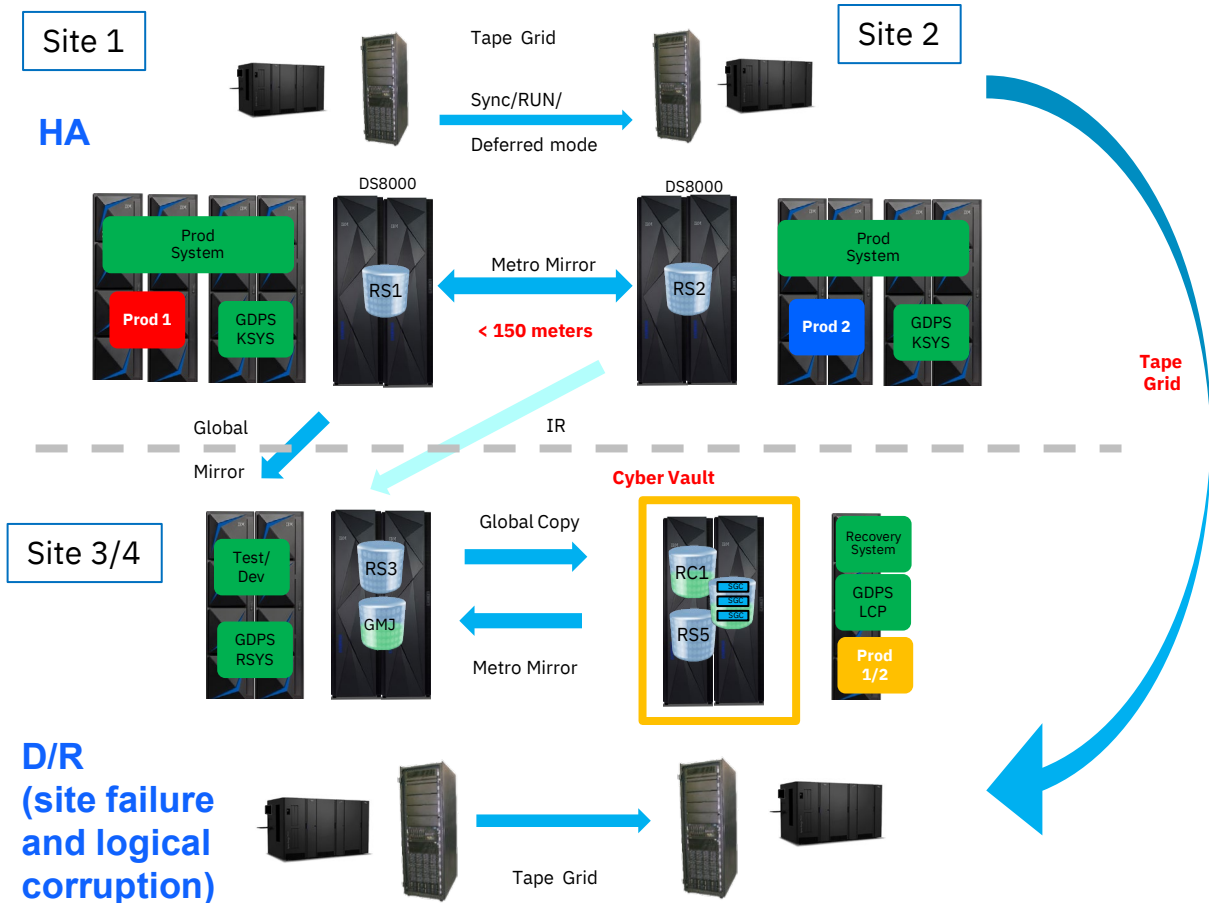
Optimized datacenter for HA and D/R



Combining local HA and remote DR

- ✗ Reduced Sysplex Overhead
- ✗ Optimized DASD performance
- ✓ Extended TAPE availability
- ✓ Coverage for regional disasters
- ✗ Cyber Vault implementation with no impact to production
- ✗ HA can be achieved by using two “datacenter cells” in the same building
- ✗ Two DASD controllers in D/R Site necessary to allow for HA in that site

Optimized datacenter for HA and D/R



Combining local HA and remote DR

- ✓ Reduced Sysplex Overhead
- ✓ Optimized DASD performance
- ✓ Extended TAPE availability
- ✓ Coverage for regional disasters
- ✓ Cyber Vault implementation with no impact to production
- ✓ HA can be achieved by using two “datacenter cells” in the same building
- ✗ Global Mirror will be paused for some seconds each time a Safeguarded Backup is taken
- ✗ Two DASD controllers in D/R Site necessary to allow for HA in that site

Let's THINK for a second

- Speed of light will not improve in the foreseeable future. That just means distance becomes a much more important factor for response times.
- **Think about:**
 - Sysplex distance
 - Moving to asynchronous IO's costs at least elapsed time, meaning transactions and especially batch jobs are running much longer
 - If you loose one site (incl. CPU and CF) you need to reipl everything anyway - even on the surviving site. So, it does not matter too much for DR if the distance is 2 km or 2.000 km. The time to restore your service in case of a site failure is more or less the same.
 - Synchronous disk replication
 - In an "all flash" disk environment a 10 km distance means 0,1 msec - just for the signal travelling time.
 - If we want to improve the DISK response times, we need to exploit concepts which do "synchronous IO's" (like Hyperlink) instead of asynchronous (FICON)

Futures

Let's have a look into the crystal ball

- IBM has already made certain features dependent on TFP for Software. This trend will continue in the future with more flexibility etc.
 - Many customers WW, in many different countries, all GEOs run TFP-SW
 - Fastest adoption of any newly introduced SW pricing metric
 - **Recommendation:** Start planning for a migration, start negotiating with your ISVs
- The speed of light will not change in the next few years
 - But the DASD / TAPE / Server technology will !
 - More TB, MIPS etc. per footprint, faster communication
 - **Recommendation:** Rethink your setup !

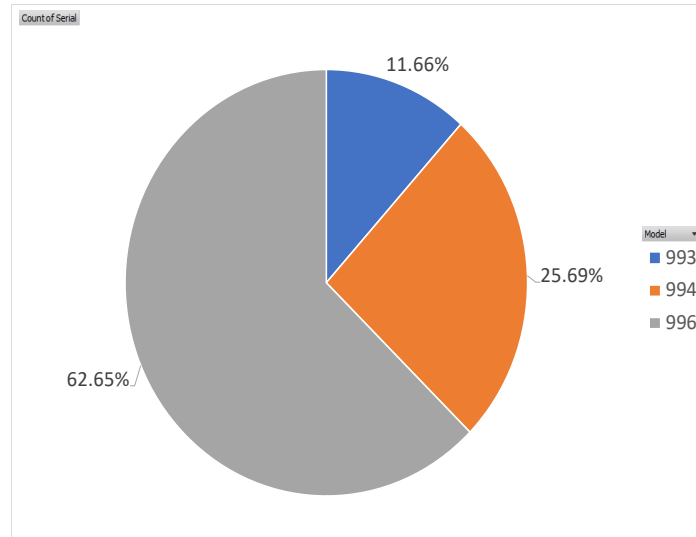
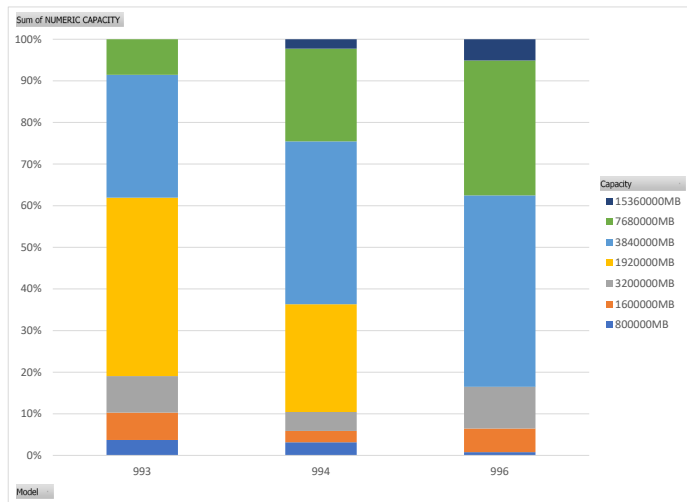
Primary Storage

DS8000 population trends

Majority of Capacity is now shipped in High Capacity Flash drives with size of drive depending on size of system

Smaller DS8910F system using 1.9TB and 3.8TB drives

Larger DS8950F systems using 7.6TB and 3.8TB drives



Increasing percentage of DS8910F systems

Move to All Flash has enabled more clients to use smaller systems with large Flash drives.

DS8910F is 37% of population compared to 29% for DS8884 and rack mounted systems also increasing

Implications of Flash trends

What can we expect in the coming years.....

- Flash media sizes will increase rapidly even in mainframe environments
 - Large drives will become prevalent in the near future
 - During this decade we would expect to see very large Flash drives in optimized Flash form factors
 - Peak sequential write workloads are often the primary driver when sizing Flash configurations
- Eventually even the largest parallel sysplex environments will be contained on a single primary storage system
 - Today this applies to a large majority of environments and in many cases multiple parallel sysplexes are consolidated on a single primary storage system
- Reduction in the total cost of storage has enabled expansion of storage topologies and this trend will continue and accelerate
 - Local Metro Mirror and HyperSwap for HA with remote replication for DR
 - Symmetrical environments with HA in all locations
 - Storage for full scale test environments
 - Storage for Cyber Resilience

Simplification and Exploitation

Is my configuration ready for the 21st Century

- What can be done to simplify z/OS storage environments?
 - Do we need 1,000s and 10,000s of volumes in a sysplex?
 - How many FICON channels do I need for a storage system?
 - Is my DFSMS configuration optimal for today's requirements?
- Am I exploiting all the capabilities that exist today?
 - HyperPAV and SuperPAV enable reduction to 16-32 aliases per LSS/LCU
 - zEDC compression can significantly reduce peak write workload
 - Starting the journey to enable and then exploit thin provisioning
 - Latency reduction technology such as zHyperwrite, read from secondary, zHyperlink

Backup and Archive

Backup and restore

- What is my strategy for backup and restore?
 - Can I restore in the times I need to? – the purpose of backup is to restore
 - How will Cyber Resiliency requirements change what I do today?
- Where can I test and practice restoration of data?
 - Did you last test a large scale restore in the 1990s?

Archive

- How much am I using Virtual Tape for batch and read type processing?
 - Would this be better on primary storage?
 - Does this make DR and Cyber Resilience easier?
- When did I define my strategy for DFSMSHsm migration?
 - Is there data being migrated that would easily be kept on primary storage?
 - Am I using MIPS to save some storage? Does this make sense with TFP?

Next Steps

Recommendation:

- **Get a clear status of where YOU are on the journey**
 - Be realistic. Are you able to test at scale, can you test logical corruption incidents. Would you survive a regional disaster ?
 - Even if things are not critical yet. Sometimes it becomes difficult to change directions.
- **Compare your status with your goals**
 - Be aware of demands and start early to get on the right track
 - Don't underestimate demand changes. Sometimes things don't develop linearly
- **Prioritize activities**
 - You can't change everything at once. Therefore, a prioritization is important:
 - Which changes are most important to answer future demands ?
 - Which changes can be easily started today to be prepared for the future ?
 - Revisit old implementations like "I always used mod 3 DASD"

z/End

Q & A / discussion



