

SHARE

Technology • Connections • Results

Real Life Disaster Recovery

Mark Wilson
RSM Partners

2nd March 2011
Session Number: **9142**



Agenda

- Introduction
- Language
- So what's it all about?
- Configuration
 - Hardware
 - Software
- What Happened?
- Recovery Steps
- Issues
- Lessons Learnt
- Questions

Introduction

- I am a mainframe technician with some knowledge of zOS
- I have been doing this for almost 30 years
- This session will describe a real event and explains what error recovery and issues we had
- Happy to take questions as we go

Language!

- And I don't mean bad language!
- Two countries separated by a common language!
- When is a ZEE not a ZEE?
- When it's a ZED
- What is PARMLIB(e)?
- When its PARMLIB

What's this?



- Zeebra?
- No it's a Zebra!
- Hopefully this will help you understand me 😊

Real Life Disaster Recovery

So what's it all about?

So what's it all about?

- We all have (hopefully) disaster recovery plans documented and tested
- And then ...
 - a disaster really happens
- What then?
- We all like to think that our disaster recovery process will kick in and all will be fine
- But what happens when it doesn't?

So what's it all about?

- I will detail a real life example of a second data center loss
- The impact it had on an organization that considered themselves well prepared for disaster
- I will detail the lessons learnt from the experience

Real Life Disaster Recovery

Configuration

Hardware Configuration

■ At the time the incident happened:

- Two Data centers some miles apart
- DC1
 - Z10
 - Several Sysplexes
 - Prod, Dev & Test
 - ICF LPAR's
- DC2
 - Z10
 - Several Sysplexes
 - Prod, Dev & Test
 - ICF LPAR's
- Plus a whole host of none mainframe kit

Software Configuration

- At the time the incident happened:
 - zOS V1.9
 - DB2 V9
 - Group Bufferpools Duplexed
 - Lock1 & SCA Structures in DC2 NOT Duplexed
 - MQ V6

Configuration



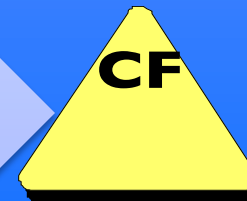
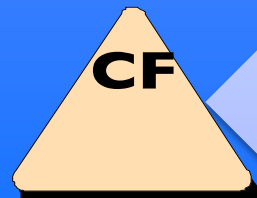
DC1

- PROD LPAR's
- ICF LPAR
- DB2
- MQ



DC2

- PROD LPAR's
- ICF LPAR
- DB2
- MQ



Real Life Disaster Recovery

What Happened

What Happened?

- One quiet Sunday afternoon
- When all seemed peaceful
- And the the traditional English pastimes of football & beer were being enjoyed!
- OK so the team were snoozing!



What Happened?

- Electrical maintenance was being performed in DC2
- A somewhat unforeseen issue meant that all power to DC2 was lost
- UPS kicked in but.....
 - Very quickly the batteries ran out!

What Happened?

- So the power to the z10 in DC2 was cut!
- The initial prognosis was that we had a network error
- Operations assumed that the HMC's in the remote command center had lost network connectivity to the SE's on the z10 in DC2
- Technical support were called and asked to assist ---- So this is where the techies step in!

What Happened?

- As previously stated the technical team were at home enjoying a peaceful Sunday afternoon 😊...
- They attempted to logon remotely but this just hung
- They then used the HMC to get to Production LPAR in DC1

What Happened?

- They noticed several IXC402D messages for all of the Production LPAR's in DC2

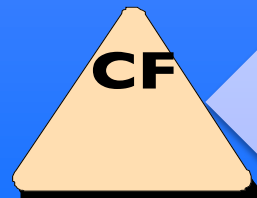
```
IXC402D sysname LAST OPERATIVE AT hh:mm:ss. REPLY DOWN  
AFTER SYSTEM RESET OR INTERVAL=SSSSS TO SET A RE  
PROMPT TIME
```

What Happened?



DC1

- PROD LPAR's
- ICF LPAR
- DB2
- MQ



Duplexed Data



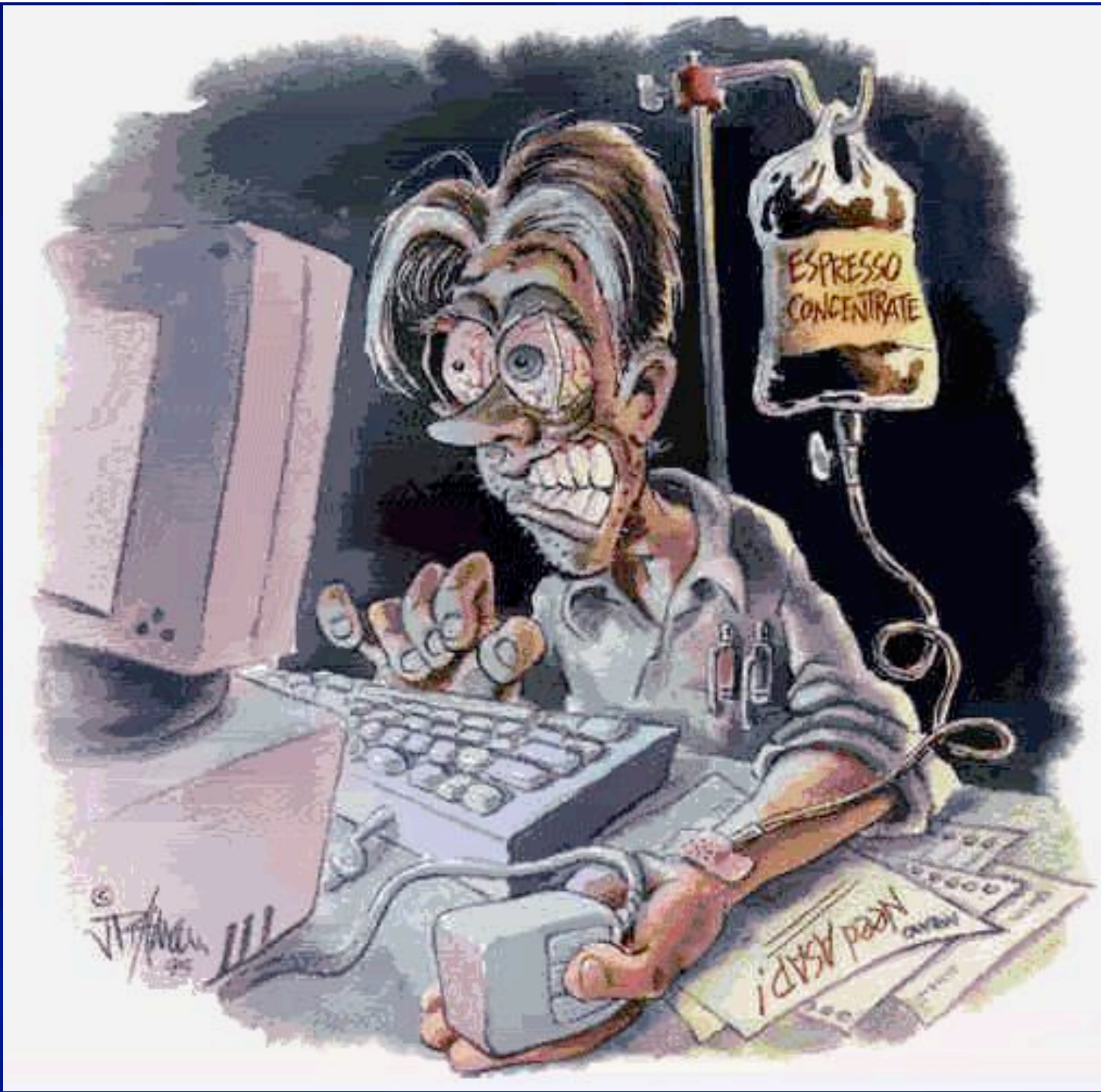
DC2



What it meant

- Lost the CF structures
 - Duplex & None Duplexed
- Lost one member of data-sharing group
- Some very unhappy DB2 subsystems
- A very fast drive to the office
- And a very long night!





Real Life Disaster Recovery

Recovery Steps

Recovery Steps

- First step was to stabilise the Production LPARS still running in DC1
 - Failed LPAR's were removed from the Production Sysplex, Replied DOWN to all IXC402D messages
 - JES2 Checkpoints reconfigured back to DASD and other such manual steps.....
 - RRS had to be completely deleted/redefined as it was in a complete mess

Recovery Steps

- At this point we thought we may be OK!
- However; some 25 minutes after the power failure in DC2 we noticed that DB2 in DC1 had failed!
- Initially the DB2 that had failed in DC2 was started in DC1 in MAINT mode to clear any outstanding locks etc
- It was then closed down to allow the normal DB2 in DC1 to function correctly

Recovery Steps

- Although DB2 had restarted it wasn't functional due to RRS and pending lock issues
- Eventually we had to take an outage to the main data sharing group in order to recover

Recovery Steps

- An action plan was put in place to start failed Production LPAR in DC1
 - On a Sunday we have enough capacity to do this, However.....
- CBU and OOCoD procedures made ready just in case DC2 not back on Monday morning to allow failed Production LPAR to run in DC1 to cope with online day

Real Life Disaster Recovery

Issues

Issues

- No idea how long power failure was going to take to recover
- It eventually took 10 hours to get power restored to DC2
- When power was restored all vendor engineers were onsite to verify hardware, which added additional time
- Outage to Production services due to initial failure

Issues

- IBM reviewed the diagnostic data and concluded:
 - Looks like the second failure was because of a failed ILRM rebuild
 - The rebuilds were started but there were no DXR146I messages to show they completed
- DB2 switched to simplex mode OK
- But the LOCK1 failed to rebuild in the surviving DC1 CF and that lead to it taking abend U2025 and DB2 came crashing down

Real Life Disaster Recovery

Lessons Learnt

Lessons Learnt

- We didn't have a dump to analyse or send to IBM
- We had discussed setting a SLIP, just in case!
- However, on further investigation of EREP data we found
 - THE RECOVERY ROUTINE REQUESTED THAT TERMINATION PROCESSING CONTINUE. AN SVC DUMP **WAS NOT REQUESTED**. NO LOCKS WERE REQUESTED TO BE FREED

Lessons Learnt

- What had been considered a resilient system had some design issues
- The recommendation from IBM was to either:
 - Duplex the LOCK1 & SCA structures; but at what performance cost
 - Have a stand alone CF; with separate power and totally isolated; again at what cost
 - Live with the risk of requiring a Sysplex wide DB2 restart in the event of a major outage in DC2

Lessons Learnt

- Had the incident not happened we would have been oblivious to this issue
- A whole raft of additional documentation has been created to ease the recovery process
- As you can imagine a lot was learnt/ discovered throughout the recovery process

Real Life Disaster Recovery

How do we really test our DR
process?

How do we really test our DR process?

- The biggest question we have been asked since the issue is how to do REAL DR TESTING?
- Normally we perform a clean shutdown and move of applications to prove they can move
- But just how good a test is that?

How do we really test our DR process?

- Should we just hit the EPO button and see what happens?
- Do you think management will buy this?
- How costly to do a real DR test?
- We are still having these debates with senior management today

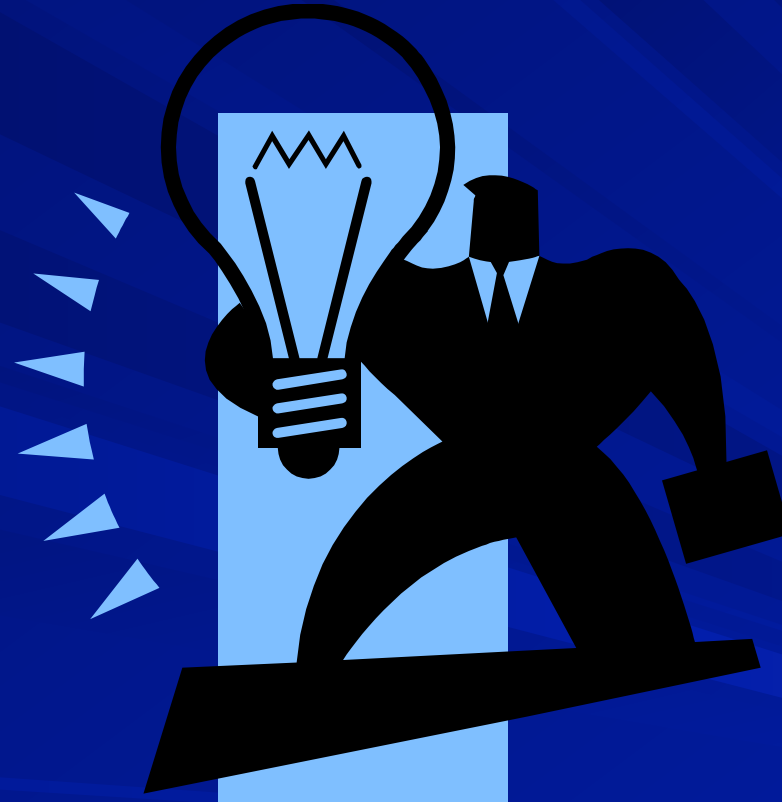
How do we really test our DR process?

- Currently we are reviewing all of the current configuration to see if we can spot any other weaknesses
- However, you don't know what you don't know and that can be painful!
- And don't forget a test is just that, it's not real until it actually happens!!!
- I just hope I am on holiday many miles away if it ever happens again.....☺

IBM Redbook

- There is some very useful information in:
 - <http://www.redbooks.ibm.com/abstracts/sg247817.html>

Questions?



And finally...

Now you can...



...get outta here!!!

Contact Details

Mark Wilson
RSM Partners

Markw@rsmpartners.com

Mobile +44 (0) 7768 617006

www.rsmpartners.com