



Met Office experiences with Linux on z



Topics for discussion

- Brief History of Linux on z at the Met Office.
- The Met Office implementation on Linux on z.
- Resilience of z/VM and Linux on z.
- Why Oracle on z ?
- Challenges, Issues and experiences.



Brief history of Linux on z at MO.

- All started with the Linuxfest back in 2000.
- SUSE 6&7 installed in year 2000/01. (z800,LPAR only)
- Installed SLES 9 during 2004. (LPAR only)
- First appearance of an IFL in 2004 on a z990.
- Switched to using RHEL 4 during summer 2006.
- First instances of Oracle 10g summer 2007. (z9, LPAR)
- Transferred all LPARed Linux instances to z/VM 5.3 hypervisors, summer 2008.



Brief history of Linux on z at MO.

- Originally applications shared within a single LPAR.
 - Mainly Java applications running within tomcat.
 - Apache web serving.
 - Open source HtDig search engine, now Google
 - Open source Linux graphics manipulation, ImageMagik
 - Enterprise Service Bus (ESB) on z990 IFL in LPAR.
 - Wiki's such as MediaWiki
- Each application now lives in its own z/VM guest. The scalability of z/VM allows this !



Why the one application per guest model?

- Only one main application per guest avoiding...
 - Too many eggs in a basket.
 - Monolithic guests and disk configurations.
 - Conflict of multiple application SLA's per guest.
 - Complex resource management per guest. (WLM not so well architected as is in z/OS)
- Bonus Advantages...
 - Easier to manage for Operations.
 - Vastly speeds up time to deploy.
 - Easier for security and life cycle management.



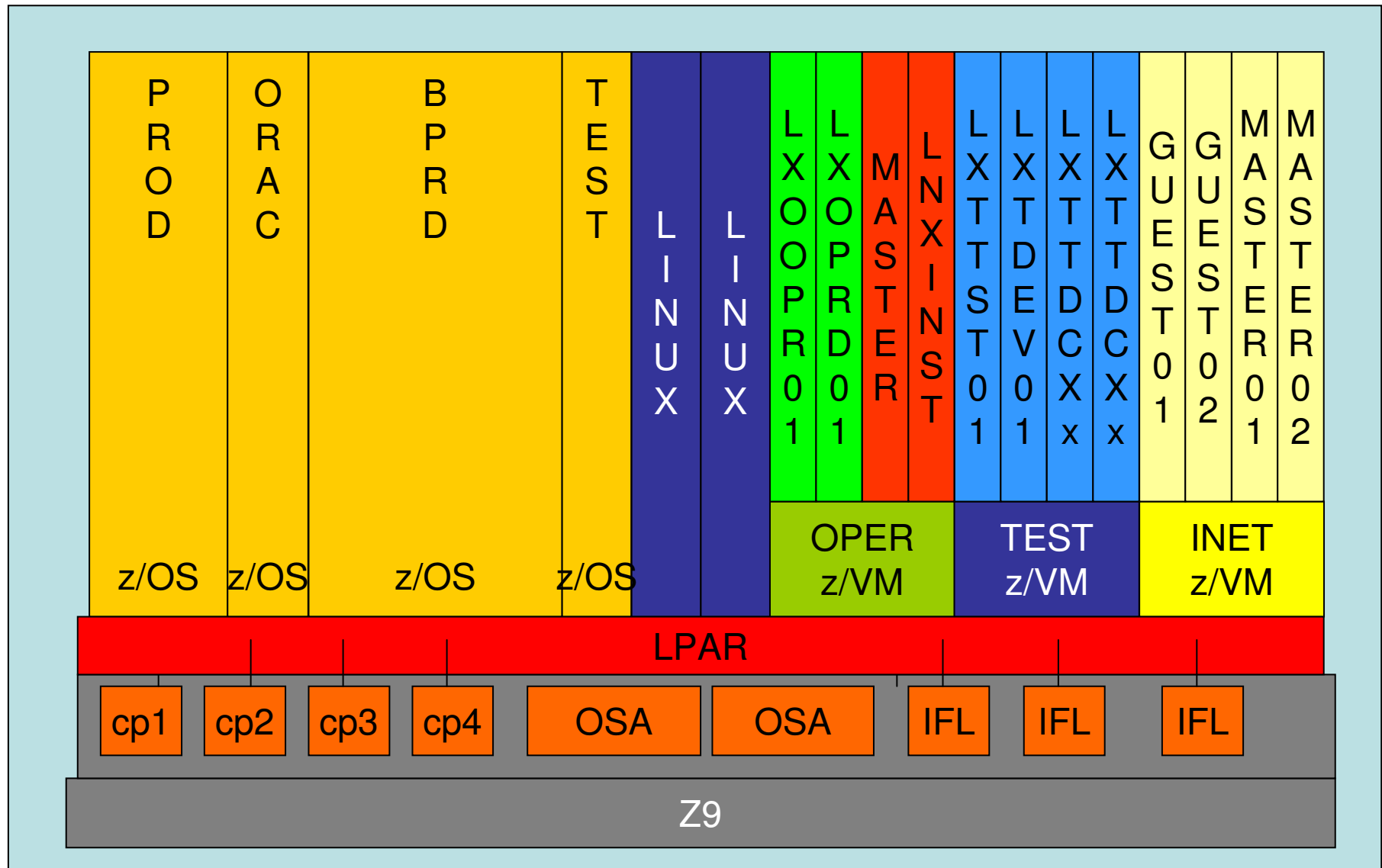
Where we are at today

- Operational deployment of z/VM late summer 2008.
- Linux guests live in two z/VM LPARs, OPER and TEST.
- Some 30 Linux guests deployed so far, 10 operational, on 2 IFLs, including key applications such as....
 - Several COGNOS finance Oracle databases.
 - Tivoli Data warehousing. (Using Oracle DB back end)
 - Remedy Oracle database.
 - Metadata Oracle databases for climate archives.
 - Encryption Key Management services.
 - Many Java applications, MediaWiki, MySQL etc, etc....



Where we are at today. z/VM configuration

- z/VM OPER LPAR configuration. (version 5.3 of z/VM)
 - 10 active guests, 8 with Oracle on them.
 - 2 IFLs.
 - 16 GB of real storage allocated with 2 GB of expanded.
 - Linux virtual storage use 26GB (Virtual to real ratio 1.5 : 1)
 - 4 x 3390-9 paging volumes and 1 x 3390-9 spool volume.
- z/VM TEST LPAR configuration. (version 5.3 of z/VM)
 - 20 active guests, 12 or so with Oracle on them.
 - 1 shared IFL with OPER.
 - 12 GB of real storage allocated with 3GB of expanded
 - Linux virtual storage use 20GB (Virtual to real ratio 1.7 : 1)
 - 9 x 3390-9 paging volumes and 1 x 3390-9 spool volume.





Implementation of Linux on z at the Met Office.

- **IBM z/VM and Linux on system z**, Virtualization Cookbook for RHEL 4 SG24-7272-00
 - <http://www.redbooks.ibm.com/redbooks/pdfs/sg247272.pdf>
- Making full use of “Master Images” within z/VM.
 - Master images initially built from RedHat kickstart scripts.
 - Oracle master images created from bare-bones kickstart configuration, with Oracle applications installed on top.
 - Has been working well for Oracle DB’s.
 - Does not work so well for lots of applications of differing nature.
- Each guest has an associated post install script.
 - Configuration scripts are a “document” of individual guests.



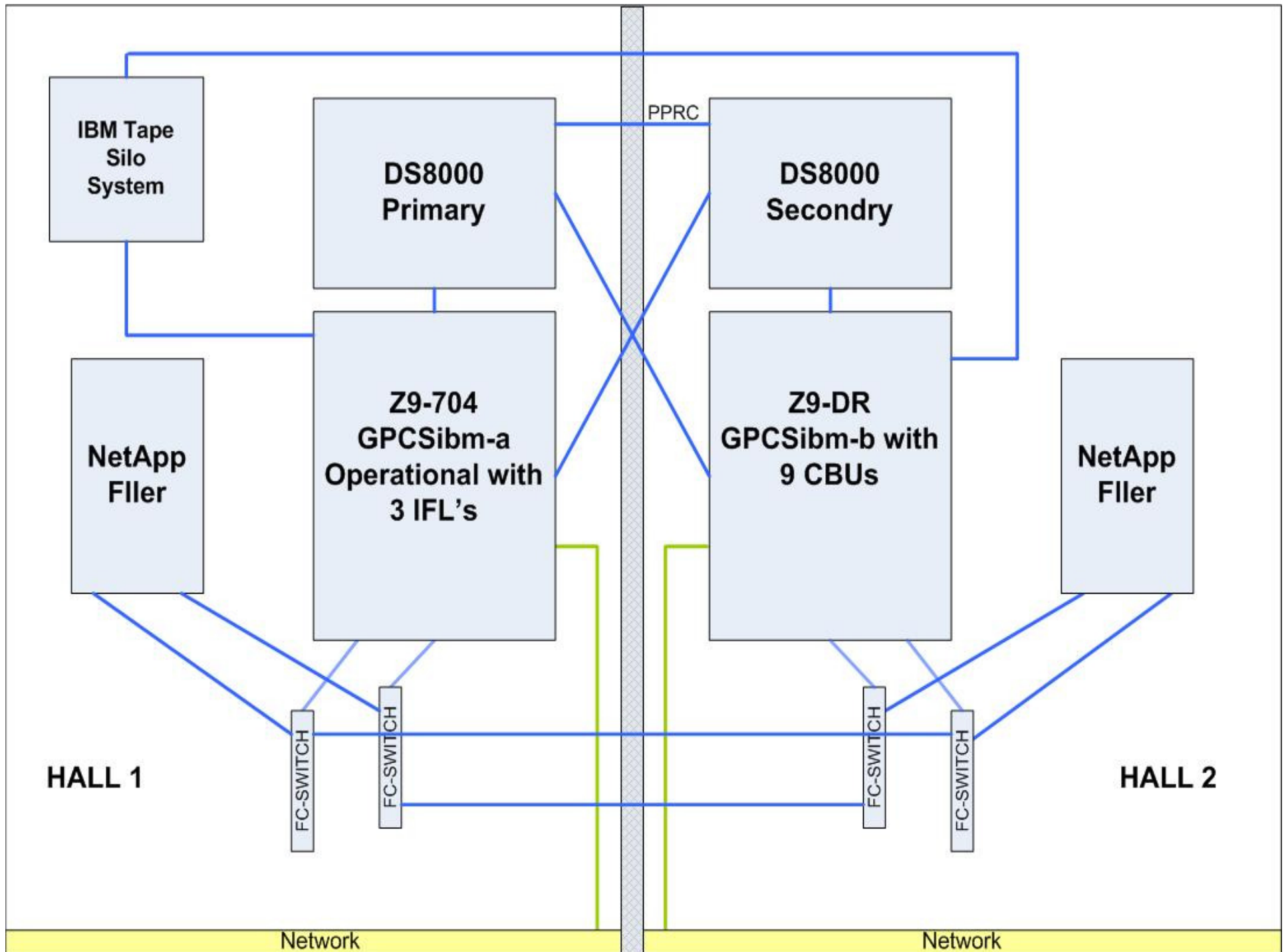
Implementation of Linux on z at the Met Office.

- DS8000/CKD storage being used for business critical services.
 - Also used for all Root and Oracle file systems. Defined by master image!
 - Accessed via the linux driver “Dasd_eckd_mod”, the standard Extended Count Key DASD device driver.
 - Consideration being given to using the “dasd_diag_mod” device driver. Simplifies I/O instruction path by letting z/VM handle I/O directly. (x'250')
- Where appropriate cheaper SATA storage (NetApp) is being used for non critical services and online backups. (FC-SW protocol)
 - Reasons for doing this are more tactical/strategic rather than financial.
 - NetApp LUNs mounted direct from Linux using the zfcps driver.
 - The Queued-I/O Assist feature provides a performance boost for virtual guests that use FCP. Default for your FCP “dedicate” statement.



Resilience/backups of z/VM and Linux on z at the Met Office.

- Not taken from the RedBook.
- z/VM resides on mirrored (PPRC) DS8000 storage.
- z/VM and all Linux master images are made online to z/OS and dumped using FDR/ABR.
- Post install config scripts of each individual guests are held within the Linux installation guest and further documented elsewhere.
- Linux guests are backed up at a file level using FDR Upstream.
- Oracles RMAN backs up databases to Netapp storage (via FC-SW). FDR upstream dumps to tape (via network).





Restoring system z Linux services.

- Very much would depend on the nature of the incident.
- Most likely scenario is the loss of an individual Linux guest.
- In general:
 - A z/VM guest would be initially restored from its master image.
 - Post install configuration script's are re-run.
 - Upstream would restore the rest !
- In the event a hall failure or major system outage, z/VM and associated guests can be brought up on z9 in hall 2.



Why Oracle on system z Linux



Why Oracle on system z at the Met Office?

- System z well placed for high I/O with 10Gb networking, plus..
 - RAS, Reliability, availability and scalable by adding more IFLs.
 - Secure.
 - Current availability of z9 resources at the Met Office..
- Massively speeds up the time to delivery.
 - Using existing infrastructure. (No need for operational acceptance)
 - Using master/golden image to roll out another one. (< 2hrs)
 - Replication of documentation and SLA's.



Why Oracle on system z at the Met Office?

- Reduced need for RAC and other high availability solutions.
 - Oracle Real Application Clusters (RAC) is a cluster database with a shared cache architecture that provide highly scalable and available database solutions.
- Oracle not officially certified for VMware !
- Would significantly reduced licensing costs for external/internet facing DBs.
- It's green.
- Doing Oracle shared homes without the need for NFS.



The Future of Linux on system z Linux



Possible future growth of Linux on z at the Met Office

- More Oracle databases including the move of the main Met Office climate database from z/OS.
- Internet facing Linux guests via LPAR or VM ?
- Most Met Office backups done via Linux guests on z.
- Porting of some legacy HP-UX applications.
- Inherit web services from z/OS to Linux on z.
- Migration of other z/OS applications.



Future configuration work for z/VM and Linux

- Upgrading to the latest release of RHEL and Oracle.
- Making better use of the hypervisor capabilities such as Common/Global file systems and read only filesystems.
 - Oracle shared homes.
- Utilizing multipathing disks under Linux and/or SCSI drivers for z/VM.
- Increase visibility of Linux guests.
 - More use of Gui's for non z system administrators.
 - Performance monitoring, capacity planning tool sets.
- More seamless failover of applications. (HA Linux)
- Improving management, security (SELINUX) and auditing of Linux.
- Integrating with the RedHat satellite server.



Challenges and experiences.

- Highly intensive CPU workloads/applications are still being avoided on our z9 IFLs.
- Third party product availability and support for z is obscure.
 - NetApp only supporting our configuration.
 - Porting of our NAG Fortran 95 compiler went well though.
- Behind with Oracle releases and updates.
 - Makes it more difficult convincing senior architects of Oracle on z.
- Booting/building/managing direct from RedHats satellite server.
- Problems managing Linux in general (not unique to z).
 - Life cycle management and availability.



Some Met Office Business challenges with Linux on z.

- Bringing our Linux teams on board.
- Platform independent management interfaces for Linux aiding..
 - Deployment / provisioning / security patching etc...
 - General administration.
 - Capacity management.
- Continued access to installed NetApp storage. Largely down to the ongoing complexities of support/certification.
- Lack of z/VM knowledge and available skills in external market.
- MOD security accreditation for Linux on system z and z/VM.
- The need to do better disaster recovery and failover.



Our general experiences so far with Oracle and Linux on z.

- Main reason for Linux system/guest failures have been system programmers mistakes or errors of judgement.
- Very usable performance for Oracle DB but
 - Oracle on Linux more resource hungry than under z/OS.
 - The Oracle “filesystemio_options” = setall parameter, vastly improved our disk I/O performance. (Allows both asynchronous and direct I/O)
- Work closely with your DBA’s, DB developers and main users to maximize performance and resources efficiencies.



In conclusion

- It does work.
 - RHEL 4 Linux has been very reliable for us.
 - Our Oracle configuration (10.2.xx.xx) is solid with no issues.
- It's green !!
- We believe we have high availability without the costs ??
- But we've got to do more of....
 - Keeping costs as low as possible.
 - Changing minds at all levels.