

Installing a new OVM

A War Story on installing a small sized
OVM on new machinery whilst
upgrading 9i2 → 12c
.... and Real Life™

Unsafe Harbour

- This is my view of reality, my experiences and most of it my colleagues agree with me on.
- Feel free to disagree and sue me.

Me ?!

- Michael Möller (aka. M-square)
- Oracle DBA consultant at Oracle and Miracle for about 20 years.
- 7.8.9.10.11.12
- Certified Master
- OakTable member

The Customer

- Customer is Selling non-food merchandise in 100 shops, most in DK
- Procurement, Logistic and Sales analysis handled in homegrown system.
 - Includes Far East site
 - Includes warehouse and distribution centre
- Has recently "streamlined" internal IT, and is in the next 5 years "rebooting" on a new (big) CRM

Current systems

- 10 year old hardware,
 - Linux V3, V4 ; Oracle 9i2 (and one 11g)
 - Clients are old Forms, shellscripts with SQLconnection, C-code, odds and ends.
 - Want the databases to run “supported”, but no upgrades until new CRM implemented.
-
- 12c, New hardware, “redundant”

Project Plan

- POC-Test migration of database(s) to 12c
 - export/import of all users schemas
- Set up new systems
 - Hardware, present SAN LUNs
 - OVM
 - Oracle 12c empty databases, new exp/imp
- Test application against new databases
- Bigbang migration weekend. (The real exp/imp)

The new System(s)

- We asked for two servers and one SAN
 - CT had some (fuzzy) high availability requirements
- Two HP Enclosures with one Server in each (and a few others not part of our systems)
 - HP BL460, 2 CPU á10 core 2,5 Ghz, 256G RAM
- 3PAR 7200 SAN with SSD, FC and NL disks
 - Approx 120T raw space, 80+ usable, we only need 10
- Shared partially with a (large?) Hyper-V cluster
- The New Computer Room....

SAN

- The 3PAR 7200 has a good performance
- Whitepaper from Oracle and HP
- Automagic everything

OVM Overview

- OVM Manager
 - Controls the server and virtual machines
 - Runs on a separate machine (Can be virtualised, too)
- OVM Server
 - The hardware server software.
 - Slave driven by OVM Manager
- OVM “Machine(s)”
 - Separate bootable virtual machines. Any OS.

Timeplan

- May – Buy & install Hardware, Install base system
 - June – Migrate test copy, start test
 - July – Holidays
 - August – Fine tune
 - Mid September – Big Bang day
-
- Hardware finally powered up late May

Install OVM Server

- CD image – how to transfer to bare iron?
 - There is no CD on server
 - HP will only boot if HP formatted USB
 - Remote Console (ILO) allows remote CD/IMG 😊
- Oracle ½Gb CD image
 - Boot, next, next, and enter IP.
- Local configuration much discouraged
- Local 300Gb disk is only <5G used.
 - (Wanted to use as storage for test)

Install OVM Manager

- Needs an additional machine.
- Eventually given virtual machine (Hyper-V)
- CD image had to be transferred to Hyper-V
- Map CD-image, boot virtual, install Linux
- Transfer Install image OVMMgr
- Install OVM – Next, next 😊 and an IP
- Needs a Database
 - Latest version includes mySQL

Configure OVM System



Oracle® VM Getting Started

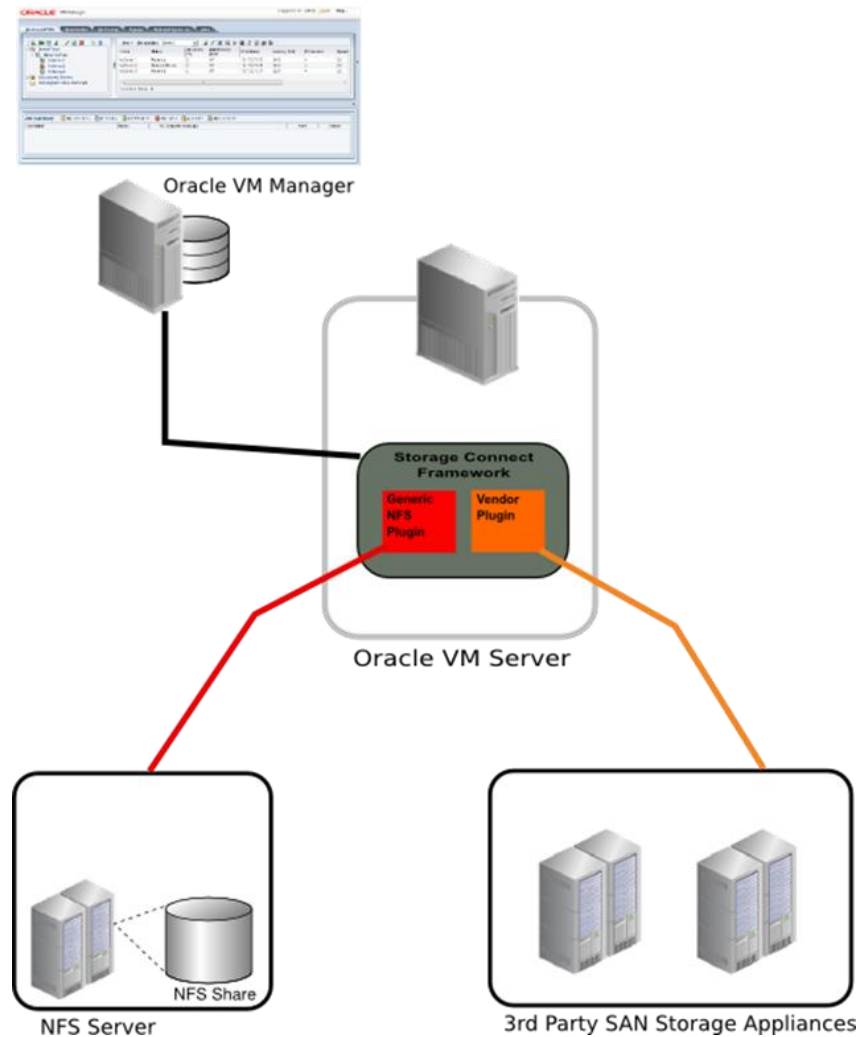
Table of Contents

- Discover Oracle VM Servers
 - Discovering Oracle VM Servers
- Discover Storage
 - Discovering a file server
 - Discovering a SAN server
- Create a Virtual Machine Network
 - Creating a virtual machine network
- Create VNICs
 - Creating VNICs
- Create a Server Pool
 - Creating a server pool
- Create a Storage Repository
 - Creating a storage repository

Configure OVM System

- Enter the IP of the OVM Server
 - (Autodiscovery possible)
- Create an OVM machine ... ☹️
- Storage?!
 - Did not have SAN yet
 - Could not use the spare OVM Server disk
 - Could use NFS
- Created NFS storage on OVM Mgr

Configure OVM System (temporary) NFS storage



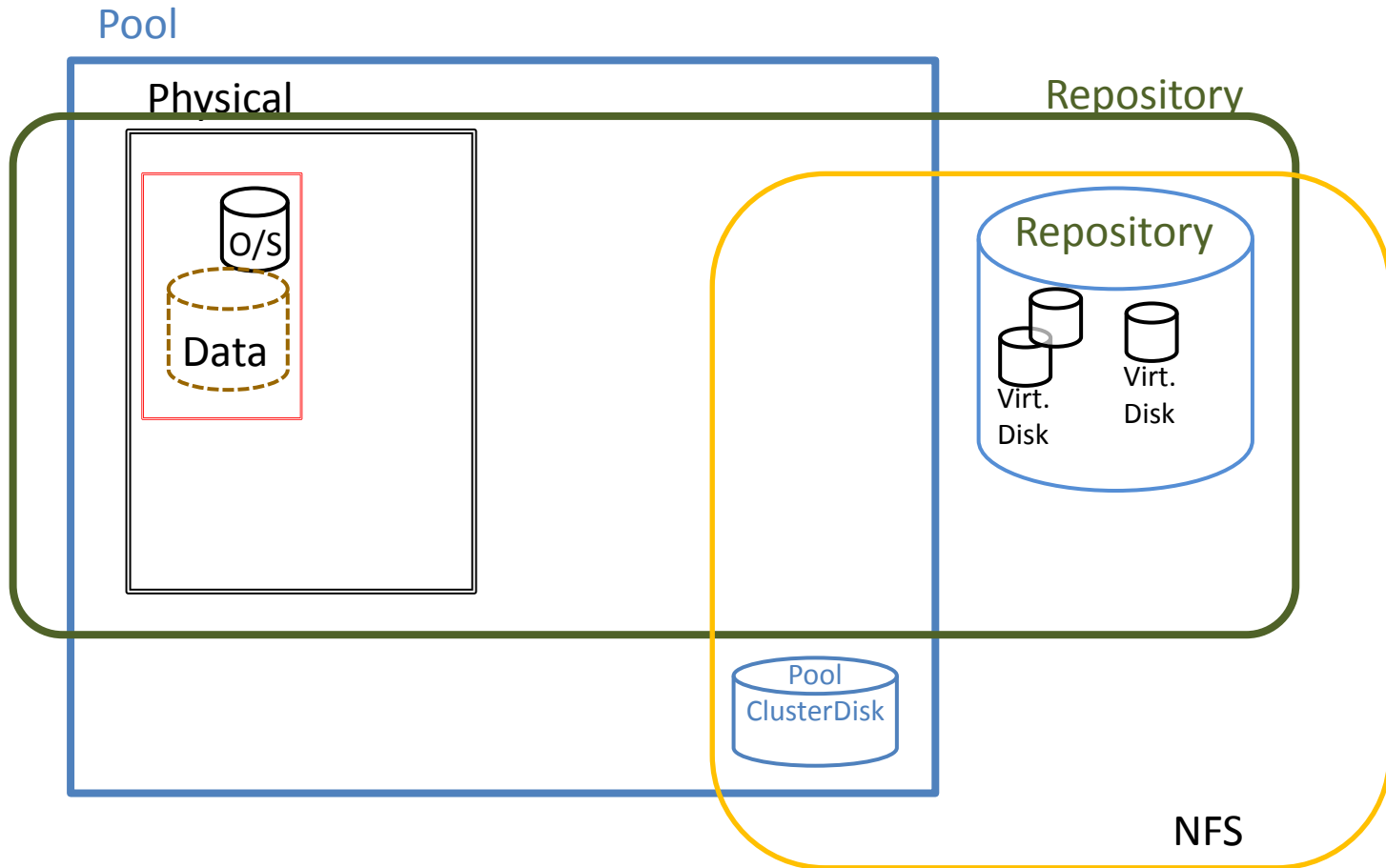
Configure OVM System Storage Repository

- To allocate storage (virtual disk) to OVM virtual machine need a repository
- NFS discovered as the repository
- To map the CD-image of Linux boot it must be in the Repository
- Copy to Repository must be from FTP or HTTP
- Create FTP Server on the OVM Mgr
- FTP CD image (via OVM Server!) to Repository
- Map CD to disk-slot.

Configure OVM System Pool

- Add discovered server to Pool
- Which has a cluster disk
- And a cluster IP
- Now able to create VM 😊
 - Which needed an IP

Temporary setup



Architecture

- OVM pool with the two servers : Prod / Test
But with Server failover
 - Forgot the SAN needs to be seen from both
 - Network needs to be same VLAN
- OVM supplies storage for OS disk
via a SAN disk
- Database disks direct SAN
 - Not via Repository
 - Not via OVM Server
 - Many LUNs

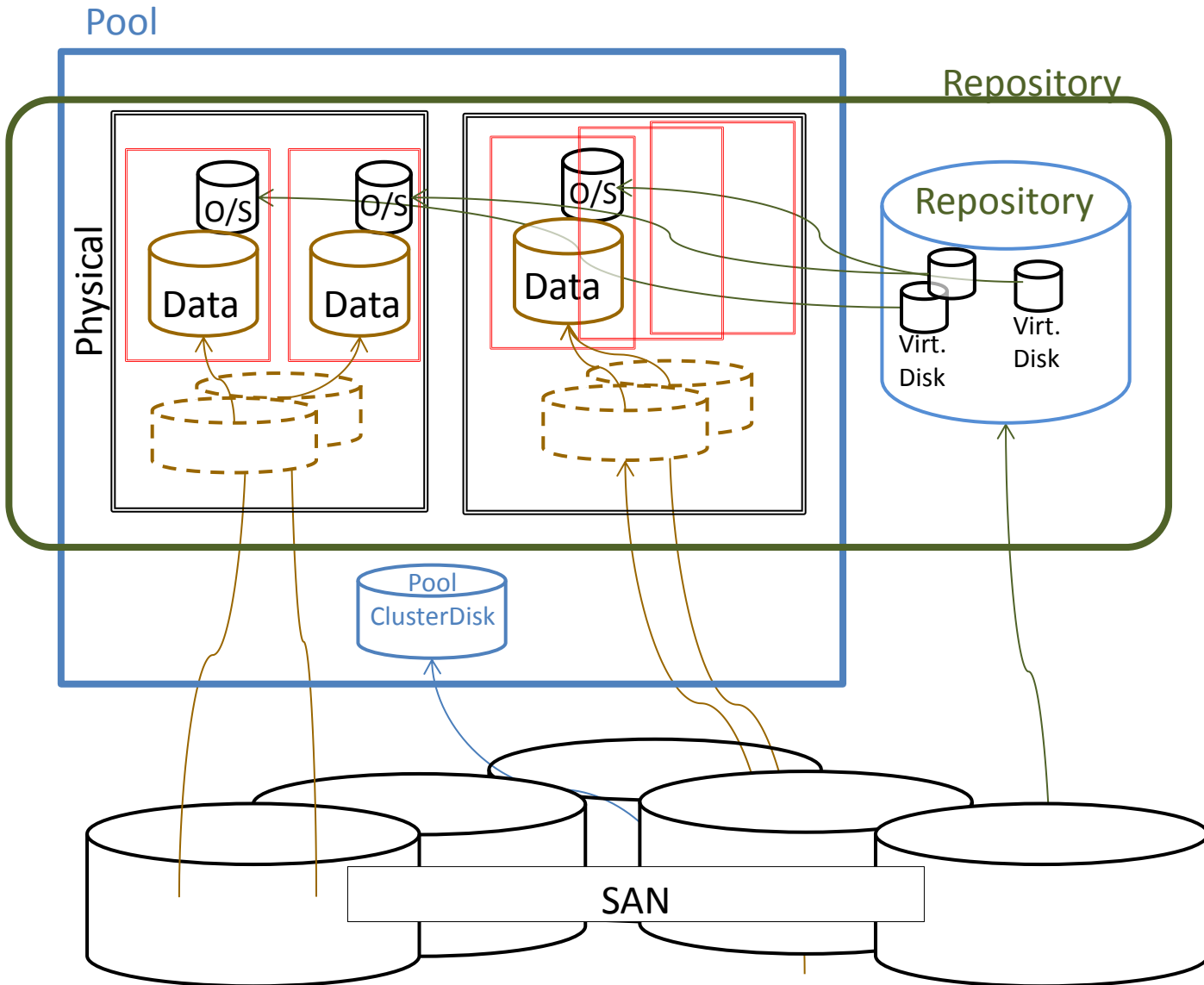
SAN available

- Hardware was iSCSI
- Could not allocate LUN before OVM requested
- Added iSCSI manually in OVM Server ☹
(Google+Internet is indeed a friend)
- Correctly: Discover SAN storage
add IP Presentation addresses
 - This present the iSCSI initiators of the OVM Servers

Project Status

- Mid July...
- To speed up I got all passwords SAN, HyperV, iLO, DNS and the OVM system I had
- "Reboot" of install; Clean.
- The *whole* install in about a day:
OVM Srv, Mgr, Config, VM machine
 - SAN disk for Pool and Repository
 - Native iSCSI setup
- Days later configuration ready for test

Configuration (near final)



Lets Talk SAN

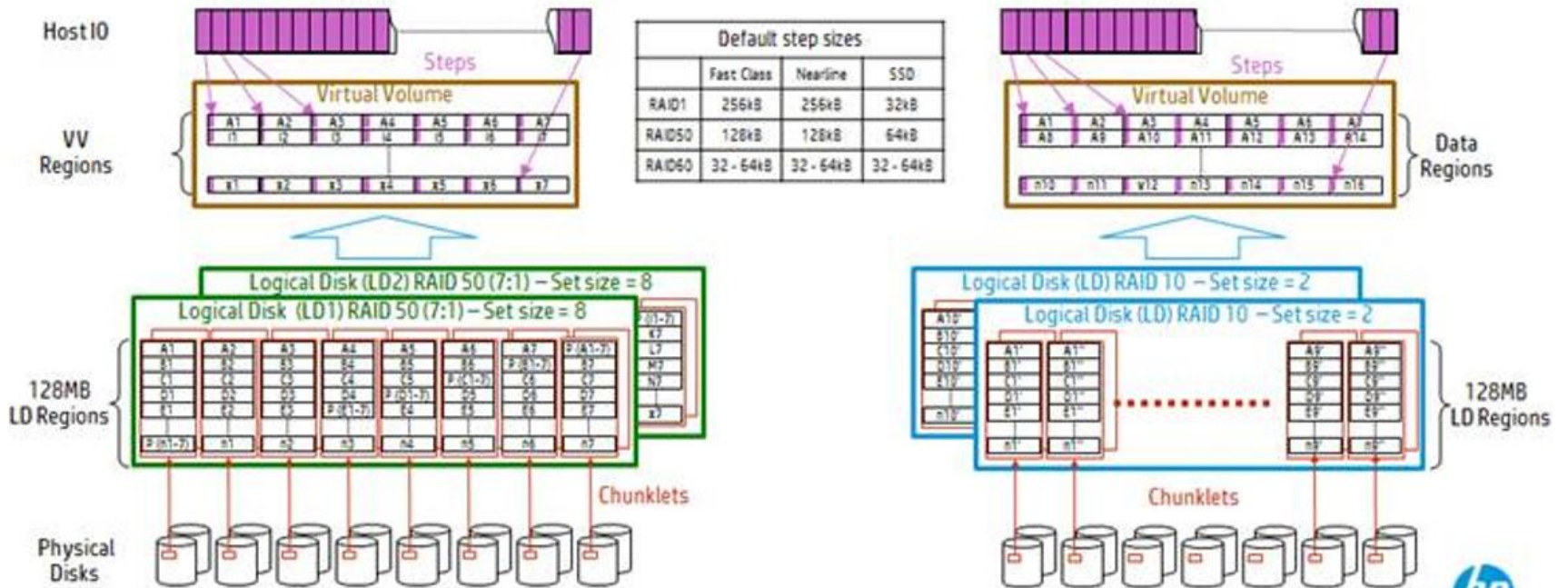
- SSD/FC/NL auto-tiering
- 12 noon rebalance (?!)
 - Monday uses Sundays load ...
 - Day/night average load is not the performance critical afternoon nor heavy night batch load
 - Includes the other system
- Fixed SAN disk to FC (15K spindles, lots, Raid5)
- Couldn't change Raid5

SAN Performance test

- Only "dummy load" with homegrown SQL
- Unable to get repeatable results
 - Sparse allocation in SAN
 - That autotune-rebalance
 - Never got round to iSCSI direct to VM.

3PAR SAN virtualization

Chunklets, Regions, Virtual Volumes and Steps



9i2 → 12c

- Chosen exp/imp, only schemas
- Platform move and upgrade
- POC test OK (done i February)
- Finetuned scripts for BigBang Day (limit downtime)
 - From POC of 3 days down to 14 hours.

exp (9i)

- Sqlplus script to get Tablespaces, Roles, Users, Sys-privs, Dblinks (no public synonyms)
- exp with list of non-system schemas
 - One with the biggest user and one with the rest in parallel (tried more, but it gave no speedup)
 - With pipe-compres (tried to NFS mounted dest but it faster to local disk)

imp (12c)

- Empty DB created before
- Created Tablespaces, users, roles
- Files transferred
- imp – same list of schemas
 - with pipe-decompres (70G instead of 350G)

Project Status

- Late August. Ready on migration but....
 - CT had not had time to test.
-
- "Plan B" – migrate to new hardware only
 - Reinstall VMs with Oracle 9i (one 11g1)
 - Copy the datafiles.
 - OR – Do a P2V (Physical-to-virtual)

”Plan B” challenges

- 2 weeks.
- 9i2 would not install on Linux 6, 5 ...
- Linux 4 as OS
- Linux 4 would not use paravirtual drivers and only 4 drives in HWM mode
- Fresh set of SAN LUNs
- (Found/Solved drive problem, too late)

Big Bang Day

- Shutdown prod
- Shutdown databases
- Copy DB files
- Startup databases
- Startup prod

Big Bang Day

- "The operation was a success, but the patient died"
- Abysmal performance – Disk IO wait 30ms
 - Linux 4 core? (NB: Crash Oracle with SETALL)
 - OVM HWM (multipath) ?
 - iSCSI problems (Flow control, loadshare)?
 - SAN config (it is not using the SSD)
 - SAN FC disks only 20-40% loaded

SAN improvement

- Migrated the FC disks to SSD
 - Took SSD out of 3 tier, made SSD/FC group, placed database disk thereon
 - Made SSD only group for Redo
 - Hyper-V stuff moved to other SAN
- Created a few extra disks for temp/undo/arch
- Fixed a FlowControl problem on iSCSI net

...

- After system settled
 - 30% of ARCH lies on SSD (auto-tune ?!)
 - SSD is used to the MAX

- Currently running stable

To Be Continued

- Project is ongoing (after Christmas)
- Will do the 9i2 to 12c (as soon as they test)
- Will continue with that SAN
 - Further tuning of tiering, better disk layout

Summary

- OVM works, is easy
- Exp/imp – tried and trusted
- SAN box throughput "spec" was an average assuming a particular mix of SSD/FC/NL use
- If you know something is a bad design, but do it anyhow, it will fail the way you hoped it would not

The Battery Fix

- Same CT. Different system. Similar age.
- Had become "slow" around the same time.
 - Was dependent on feed from migrated system
- Big taskforce to trace night run.
- Statspack showed lots of LGWR Parallel Write
- Volume of transaction seemed same
- Local disk controller blinked battery low and one disk was offline

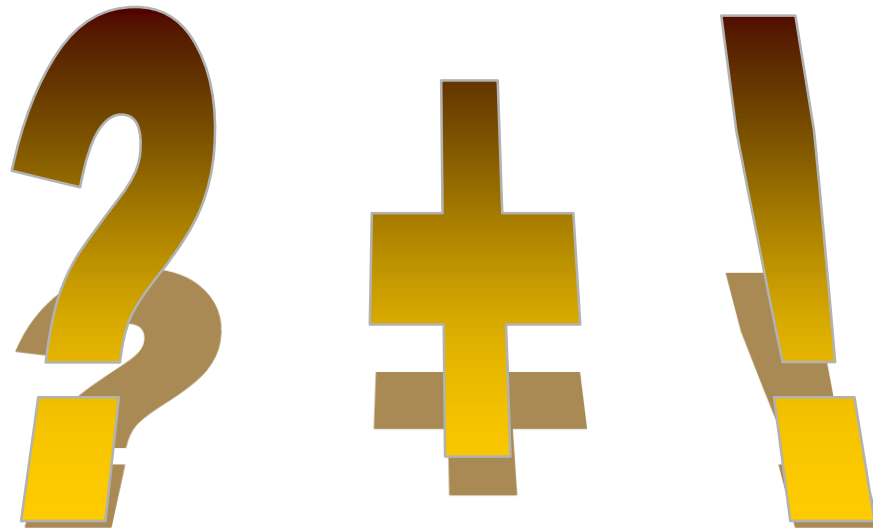
The Battery Fixed

- Battery was changed
- Disk changed
- In Statspack the LGWR Parallel Write gone
- Performance good again

Local disk Cache

- System had 6 disk in 3 Raid 1 sets
- Raid controller had writeback cache
- Good hardware disabled cache
 - writes at disk speed
- Good battery, cache enabled
 - writes at cache speed
- Thought for day:
Is this safe?

... and finally,



Thank you,
M2@Miracle.DK