
BGOUG Autumn conference 2009



Oracle^{Unix-like}

Svetoslav Gyurov
CNSys PLC

ORACLE[®]
Certified Professional

13-15 November 2009

Comparison



◆ HP-UX

- ❑ Commercial UNIX for 25 years
- ❑ Working on PA-RISC / EPIC platforms (HP)
- ❑ A lot of man-hours in design and implementation.
- ❑ Comes with (optional) support contracts.

◆ Linux

- ❑ Free as in freedom, not as in free beer
- ❑ Open source
- ❑ GNU's not Unix!



How Linux penetrated the market



- ◆ 1998 - Oracle8 released on the Linux platform – first commercial database on Linux
- ◆ 2003 - Oracle10g supports 64-bit Linux (Red Hat 3.0) with IPF (Itanium Processor Family)
- ◆ 2006 - Oracle announces the Unbreakable Linux program
- ◆ 2007 - Oracle released Oracle Database 11g for Linux
- ◆ 2009 - Oracle release Oracle Database 11gR2 first for Linux

Open Source into the Enterprise



- ◆ Low costs servers and OS
- ◆ High-performance - near the UNIX systems
- ◆ No HW dependant - support of ISS
- ◆ Open source - constantly finding/fixing bugs.
- ◆ Biggest community wins
- ◆ Fast and easy deploy
- ◆ Commercial support
- ◆ Opportunity to reduce IT infrastructure costs

Certified Linux for Oracle



RHEL (Red Hat Enterprise Linux)



SLES (SuSE Linux Enterprise)



OEL (Oracle Enterprise/Unbreakable linux)

Memory



- ◆ For 32-bit: Use VLM, PAE
- ◆ For 64-bit: Use Huge pages
 - Not swappable
 - Decreased page table overhead
 - Eliminated page table lookup overhead
 - Faster overall memory performance
- ◆ For 11g use AMM - /dev/shm

Network



- ◆ Bonding
 - ❑ mii-tool – obsolete
 - ❑ Use LACP (if possible)
- ◆ Jumbo frames
 - ❑ 50% reduction in CPU overhead
 - ❑ 75% reduction in Global cache buffer waits
 - ❑ IP Reassemblies dropped by 10x
 - ❑ Use it only for RAC private interconnects
 - ❑ Not an IEEE standard

Storage options



- ◆ Cluster File system – OCFS2, Veritas CFS
- ◆ Almost ZFS - Btrfs – B-Tree or “butter” FS
- ◆ ADVM and ACFS – certified for OEL 5 / Red Hat 5
- ◆ NFS, Direct NFS
- ◆ LVM
- ◆ (no) RAW

Multipath Linux



- ◆ Connectivity:
 - ❑ FC HBA driver
 - ❑ Auto compilation
- ◆ Load balancing:
 - ❑ Multipath driver and tools
 - ❑ Device Mapper Multipath Enablement Kit

Multipath Linux again



- ◆ Ignore /dev/sd*, use /dev/mapper
- ◆ Partition the disks:
 - Use parted, not fdisk
 - Use kpartx
- ◆ EMC Powerpath - /dev/emcpowerxx
- ◆ IBM SDD - /dev/vpath
- ◆ HP SecurePath

```
linux
File Edit View Scrollback Bookmarks Settings Help
linux: / # multipath
Nov 05 11:05:27 | Invalid device number 100
Nov 05 11:05:27 | Invalid device number 100
Nov 05 11:05:27 | vdisk-fc: failed to access path 100
reload: vdisk-fc (36001438002a54fa10001400000050000) n/a HP,HSV300
[size=1.0G][features=0][hwandler=0][n/a]
\_ service-time 0 [prio=100][undef]
\_ 3:0:0:1 sda      8:0  [active][ready]
\_ 2:0:0:1 sdb      8:16 [undef][ready]
\_ service-time 0 [prio=20][undef]
\_ 2:0:1:1 sde      8:64 [undef][ready]
\_ 3:0:1:1 sdf      8:80 [undef][ready]
reload: vdisk-fata (36001438002a54fa100014000000c0000) n/a HP,HSV300
[size=1.0G][features=0][hwandler=0][n/a]
\_ service-time 0 [prio=100][undef]
\_ 2:0:1:2 sdg      8:96 [active][ready]
\_ 3:0:1:2 sdh      8:112 [undef][ready]
\_ service-time 0 [prio=20][undef]
\_ 3:0:0:2 sdc      8:32 [undef][ready]
\_ 2:0:0:2 sdd      8:48 [undef][ready]
linux: / # top -b -i
top - 11:05:30 up 1:23, 1 user, load average: 4.87, 4.09, 3.60
Tasks: 409 total, 7 running, 402 sleeping, 0 stopped, 0 zombie
Cpu(s): 0.4%us, 3.4%sy, 0.0%ni, 96.2%id, 0.0%wa, 0.0%hi, 0.0%si, 0.0%st
Mem: 131907404k total, 571808k used, 131335596k free, 13764k buffers
Swap: 2104472k total, 0k used, 2104472k free, 127848k cached

  PID USER   PR  NI  VIRT  RES  SHR  S %CPU %MEM    TIME+  COMMAND
10242 root    RT   0 34156 3500 2788 R  100  0.0   7:20.33 multipathd
10475 root    RT   0 34160 3504 2788 R  100  0.0   0:44.32 multipathd
10580 root    20   0 8868 1272  744 R   11  0.0   0:00.12 top
   22 root    15  -5    0    0    0 R    0  0.0   0:02.84 ksoftirqd/9
   26 root    15  -5    0    0    0 R    0  0.0   0:00.00 ksoftirqd/11
   76 root    15  -5    0    0    0 R    0  0.0   0:00.08 events/9
   78 root    15  -5    0    0    0 R    0  0.0   0:00.06 events/11
```

Multipath HP-UX



- ◆ FC drivers included
- ◆ HP-UX 11iv2 – additional software (license)
- ◆ HP-UX 11iv3 – native (active/active)
- ◆ Multipath is transparent
- ◆ EMC Powerpath driver

```
unix
File Edit View Scrollback Bookmarks Settings Help

root@anduril:~# autopath display

=====
HPswsp Version      : A.3.0F.02F.00F
=====
Auto Discover       : ON
=====
Array Type          : EVA-AA
Array WWN           : 5000-1FE1-5003-0620
Read Optimize       : ON
Path Verification Period : 00:10
=====
Lun WWN             : 6005-08B4-0005-50F4-0000-B000-0013-0000
Virtual Device File : /hpa/dsk/hpap1
Load Balancing Policy : Shortest Service Time
Lun Timeout         : Infinite Retry (-1)
=====
Device Path          Status
=====
/dev/dsk/c9t0d1      Active
/dev/dsk/c11t0d1     Active
/dev/dsk/c5t0d1      Active
/dev/dsk/c7t0d1      Active
=====

Lun WWN             : 6005-08B4-0005-50F4-0000-D000-0007-0000
Virtual Device File : /hpa/dsk/hpap7
Load Balancing Policy : Shortest Service Time
Lun Timeout         : Infinite Retry (-1)
=====
Device Path          Status
=====
```

ASMLib (for Linux)



- ◆ Alternative interface for ASM to identify and access block devices
- ◆ Kernel driver
- ◆ Library and tools
- ◆ Multipath disks - /dev/dm; /dev/mapper
- ◆ Disk scan order – FSFU (first scanned, first used)
- ◆ Do not delete /opt/oracle/extapi

CRS11g / ASM 11g installation



- ◆ Oracle 11.1 - vote and ocr disks can be block devices and not raw
- ◆ Oracle 11.2 – Desupport of block and raw devices
- ◆ Oracle Clusterware - ASM - Database Version Compatibility Doc ID: 337737.1

Patches and updates



- ◆ up2date
 - Public yum server: <http://public-yum.oracle.com/>
- ◆ swinstall / update-ux
 - Standard patch bundles

System backup and recovery



- ◆ Linux
 - Clonezilla, Acronis – offline
 - MondoRescue - online
- ◆ HP-UX
 - Ignite-UX – deploy, create, recover
 - make_tape_recovery
 - make_media_install

Administration & Support



- ◆ Persistent device naming
 - ❑ Flex10
 - ❑ udev
- ◆ Monitoring
 - ❑ OnlineDiag, PSP
 - ❑ MP, iLO
 - ❑ Openview, Nagios
 - ❑ OSWatcher, OSWg – Doc ID 301137.1

Q&A



Thank you!

s.gyurov@cnsys.bg