Oracle on HP-UX – Best Practices

Herbert Dorer
Senior Technical Consultant, HP

Agenda

- Disk Layout
- HP-UX Kernel Parameters
- Some Patches
- Performance Services

Disk Layout for SAP R/3 on Oracle

(optimized for use with large disk arrays)

Traditional approach

Disk 0:   saparch   Offline redo log file (Archive logs)

Disk 1:   origlogA   Online redo log files from the first and third group (Set A)

Disk 2:   origlogB   Online redo log files from the second and fourth group (Set B)

Disk 3:   mirrorlogA Mirrored online redo log files from the first and third group (Set A)

Disk 4:   mirrorlogB Mirrored online redo log files from the second and fourth group (Set
B)

Disk 5:   sapdata1   Database files, mirror of the control file

Disk 6:   sapdata2   Database files, mirror of the control file

Disk 7:   sapdata<n>   Other database files in sapdata3 up to sapdata<n>

....

+ special considerations for:

PSAPROLL, PSAPTEMP.....

+ strict separation of tables and associated indices

The problem

- Disk size when layout was proposed: 1 GB – 2 GB

- Typical Sizes today: 36 GB – 146 GB

  ✎ impossible to satisfy all requirements with respect to performance and security

- Difficult to achieve good I/O load distribution with traditional disk layout

- Goal: simplest possible layout, which guarantees
  - Optimal data security
  - Optimal data distribution and performance

Data Security Considerations

- Archive logs must be physically separated from data files and online redo logs
  - Very large size of array groups (XP) or a whole redundancy group (VA)
    - ✍ not practical to enforce this requirement
  - Very advanced data security features
    - ✍ archive logs in same array groups as data and online redo logs is acceptable

- Store 2 – 3 copies of control files on separate physical disks

- Avoid SW striping of online redo logs

- Separating online redo logs from datafiles is a performance issue and NOT one of data security!

Performance Considerations

- Configure enough LUNs to allow I/O load balancing over all available interfaces

- No less than 1 LUN per 4 physical disks

- On Virtual Arrays as of HP-UX 11.i, one large LUN per Redundancy Group (RG) for sapdata and increasing the SCSI Queue Depth for those LUNs to 2 – 4 times physical disks in RG is a perfect choice

- Traditional LVM striping (e.g. 64K) might interfere with pre-fetch algorithms of modern disk arrays

- When using statistical extent based distribution over all devices, separating tables and indexes gives no additional benefit

Performance Considerations (cont.)

- Online logs are written sequentially and should always be written to array cache.
  - With enough cache, no need to reserve physical disks
  - XP: Use CVS (Custom Volume Size) for placing redo logs in ‚left-over' space in some array groups for data files
  - EMC: No redo logs on hyper volumes smaller than average hyper volumes (statically bound cache)
- Use dedicated LUNs (in dedicated VG vgSIDlog) for redo logs
  ✍ log writer I/Os in different SCSI queue than db writer processes
- Never place archive logs on same LUNS as data files (same reason)
- Maximum archiver performance: Extent distribute saparch-LV over multiple array groups (XP)/both RG (VA)

Data Granularity

- Extent (or stripe element) size small compared to cache of array

- Extent (or stripe element) size small compared to size of critical Oracle segments
  - Better statistical I/O distribution
  - Critical tables/indexes typically several 100 MB

- Extent (or stripe element) size > 512 KB
  - no interference with pre-fetch algorithm

Layout Suggestion

- One device or device group for archive logs (vgSIDarch)

- Fill up with ‚non-sapdatas‘

- No dedicated array group on XP (practical reason, large array group size)

- Create dedicated VG (vgSIDlog) for online redo logs

- All other devices into VG vgSIDdata

- Create FS origlogA, origlogB in vgSIDlog

Layout Suggestion (cont.)

arch_put_reorg /sapmnt/SID /usr/sap /oracle/SID /usr/sap/trans	origlogA	origlogB	3	4	5
			8	9	10
	1	2	13	14	15
	6	7	18	19	20
	11	12	23	24	25
	16	17	26	27	28
	21	22	29	30	31

vgSIDarch

vgSIDdata

vgSIDlog (if exists, else: vgSIDdata; never: vgSIDarch)

Mount options

- HP-UX 11.0: JFS 3.1, HP-UX 11.i: JFS 3.3

- mincache=direct (bypass buffer cache on read)

- convosync=direct (force direct I/O for DB writers)

- Use for data files and redo log files

- For JFS 3.3, patch PHKL_29115 is required

- Mount options can be changed online
  - Mount –F vxfs –o \
    remount,nodatainlog,mincache=direct,convosync=direct \
    /dev/vgSIDdata/lvsapdata1 /oracle/SID/sapdata1

- Permanent in /etc/fstab or dbci.cntl (MC/SG)

HP-UX Kernel Parameters for SAP R/3 on Oracle

HP-UX Kernel Parameters

- Most: Works ✍ ✍ Does not work
- Only a few are relevant to performance
- Small/Medium systems: SAP recommendations
- Large systems: Further increase necessary for some parameters
- Consolidated systems

Memory Usage

- dbc_min_pct, dbc_max_pct (dynamic buffer cache)
  - Default 5, 50
  - Set to: 2 – 5, 3 – 10
  - Dependencies: Physical memory, mount options
- vx_ninode (VxFS inode cache)
  - Default much too big for large memory (> 4GB)
  - Set to: 20000 – 40000, depending on number of files
  - Requires patch PHKL_28185
- shmmax (maximum size of a shared memory segment)
  - Adjust to maximum required (> MAX(SGA, EM)

Miscellaneous

- nfile (number of concurrently open files)
  - Set to: (# of Oracle data files) X (# of shadow processes) + safety
- Make sure to allow enough swap space
  - Max Swap = maxswapchunks * 1024 * swchunk (typically 2048)
- Semmns (maximum number of semaphores)
  - Different formulas
  - Depends on Oracle parameter 'processes'
  - ‡(processes) + MAX(processes) + (#instances * 10)

Consolidated Systems

- Servers with MANY Oracle databases in one partition
- Main parameters to be summed up:
    - maxusers
    - maxvgs
    - msgmap
    - msgseg
    - msgtql
    - nfile
    - nflocks
    - nkthread
    - nproc
    - semmni
    - semmns
    - shmmni
    - swchunk

Memory Management Patches for HP-UX 11.11

Memory Management Patches

- New patches important for performance and functionality

- Latest patch sets should be applied from time to time

- Memory management patches for 11.11
  - PHKL_28695
  - PHKL_28410
  - PHKL_25212
  - PHKL_30259

Performance
Services

Performance Services Team

- SAP EarlyWatch and GoingLive Services
- Performance Analysis for SAP
- Transaction Volume Measurement Service
- Storage Performance Analysis
- Performance Workshop (HP-UX and SAP)
- Application Analysis (Oracle, MS SQL)
- SAP Tuning – Oracle Shared Cursor Cache Analysis
- SAP Tuning – Customer Programs
- Capacity Planning and Resizing
- System Monitoring

Contact

- Phone: +49 (2102) 90-5442

- Fax: +49 (2102) 90-5927

- E-Mail: performance.team@hp.com

- Or just call:
  - Martin Post   +49 2102 90 5717
  - Manuel Padilla   +49 2102 90 5907
  - Georg Krug   +49 2102 90 5694

- Bi-monthly HP-SAP newsletter:
  - Send mail to hpsap.newsletter@hp.com

Questions?

Thank
You!