


Tru64 UNIX Monitoring und Tuning

27. DECUS Symposium 2004 in Bonn

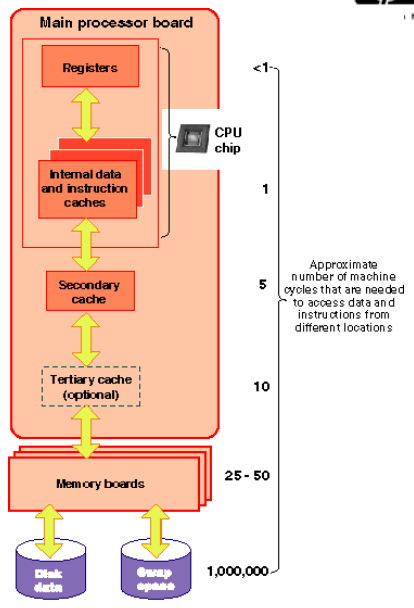
Reinhard Stadler
HP Services

© 2004 Hewlett-Packard Development Company, L.P.
The information contained herein is subject to change without notice



„System Tuning“

- Monitoring
- Erkennen von „Bottlenecks“
 - Application
 - Hardware
 - Operating System
- Bottleneck auf die schnellste Komponente verschieben



26-182U-4

26 April 2004

2



Agenda

- Allgemeines zum Tuning
- Monitoring
- Verschiedene Subsysteme
- Tuning am Beispiel eines Oracle DB Servers

26 April 2004

3



grundsätzlich ...

- typische Antwortzeiten, möglicher Datendurchsatz, ...
 - was kann man überhaupt erreichen?
 - Subjektiver Eindruck / gemessene Daten
- Schwankungen über einen kurzen Zeitraum
 - Lastverteilung über den Tag, Woche, Monat, ...
- langfristige Trends
 - besteht Handlungsbedarf für die Zukunft ?

26 April 2004

4

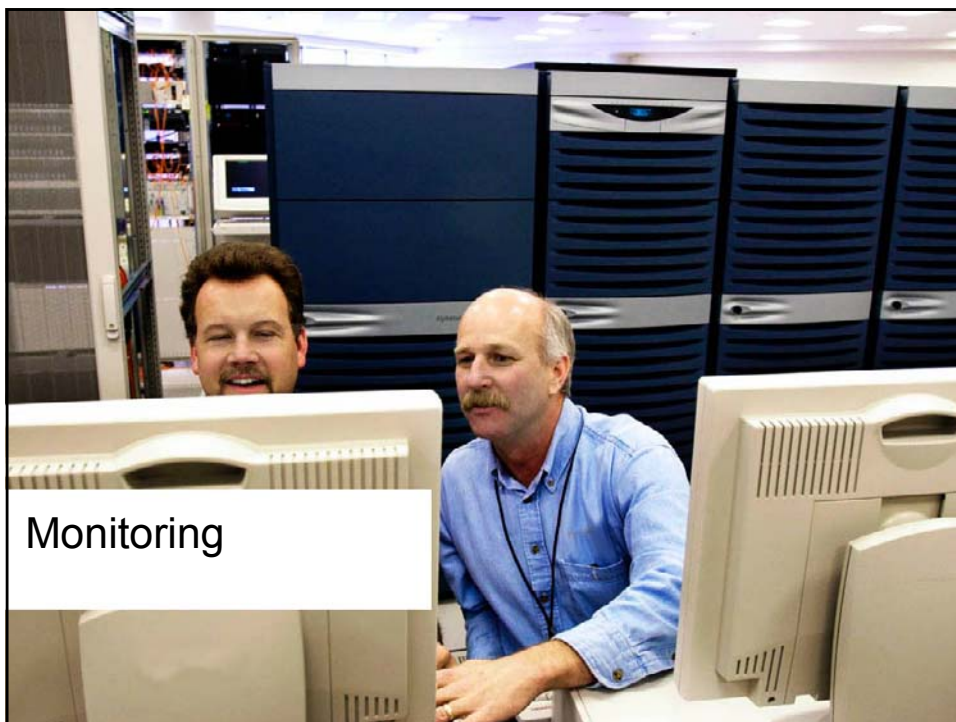


Grundregeln

- möglichst wenig I/O auf möglichst viele Komponenten verteilen
- Ausreichende Ressourcen für alle Prozesse
 - CPU Typ und Anzahl
 - I/O Subsystem
 - Netzwerk

26 April 2004

5





Monitoring: vmstat, uptime

```
# vmstat 2
```

```
Virtual Memory Statistics: (pagesize = 8192)
```

procs			memory			pages			intr			cpu					
r	w	u	act	free	wire	fault	cow	zero	react	pin	pout	in	sy	cs	us	sy	id
2	66	25	6417	3497	1570	155K	38K	50K	0	46K	0	4	290	165	0	2	98
4	65	24	6421	3493	1570	120	9	81	0	8	0	585	865	335	37	16	48
2	66	25	6421	3493	1570	69	0	69	0	0	0	570	968	368	8	22	69
4	65	24	6421	3493	1570	69	0	69	0	0	0	554	768	370	2	14	84
4	65	24	6421	3493	1570	69	0	69	0	0	0	865	1K	404	4	20	76

```
. . .
```

```
# uptime
```

```
1:48pm up 7 days, 1:07, 35 users, load average: 7.12, 10.33, 10.31
```

26 April 2004

7



Monitoring: ps

```
# ps aux
```

USER	PID	%CPU	%MEM	VSZ	RSS	TTY	S	STARTED	TIME	COMMAND
chen	2225	5.0	0.3	1.35M	256K	p9	U	13:24:58	0:00.36	cp /vmunix /tmp
root	2236	3.0	0.5	1.59M	456K	p9	R	+ 13:33:21	0:00.08	ps aux
sorn	2226	1.0	0.6	2.75M	552K	p9	S	+ 13:25:01	0:00.05	vi met.ps
root	347	1.0	4.0	9.58M	3.72	??	S	Nov 07 01:26:44		/usr/bin/X11/X -a
root	1905	1.0	1.1	6.10M	1.01	??	R	16:55:16	0:24.79	/usr/bin/X11/dxpa
mat	2228	0.0	0.5	1.82M	504K	p5	S	+ 13:25:03	0:00.02	more
mat	2202	0.0	0.5	2.03M	456K	p5	S	13:14:14	0:00.23	-csh (csh)
root	0	0.0	12.7	356M	11.9	??	R	< Nov 07 3-17:26:13		[kernel idle]

26 April 2004

8



Monitoring: Netzwerk

netstat

```
# netstat -i      Interface Statistik
# netstat -m      Memory
# netstat -r      Routing
# netstat -a      open sockets
```

- Socket Statistik

```
# sysconfig -q socket
```

- Routing

```
# traceroute
```

26 April 2004

9



Monitoring: I/O

iostat

– Wenig hilfreich

volstat

– LSM liefert bessere I/O Statistiken!
– Anzahl I/Os, KB read, write und „Average Service Time“
– Erste Ausgabe liefert Zusammenfassung über den gesamten Beobachtungszeitraum

advfsstat

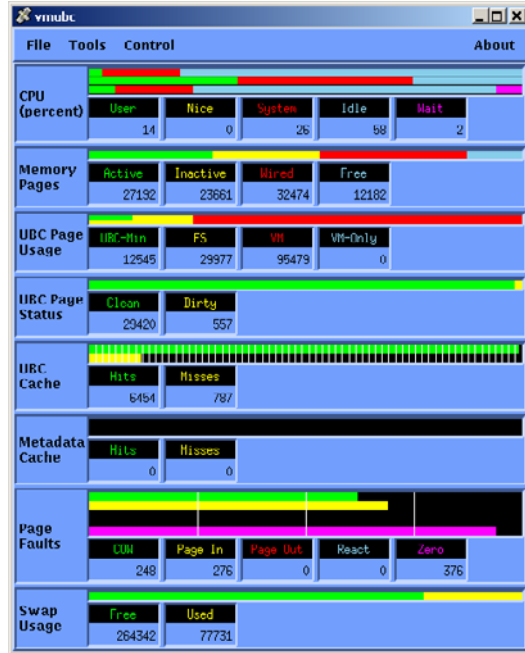
AdvFS I/O Queue Statistiken

26 April 2004

10

Tools

vmubc &



26 April 2004

11

Tools

top

```

load averages:  3.82,  3.47,  3.73                                     11:16:58
524 processes:  5 running,  53 waiting,  99 sleeping,  366 idle,  1 zombie
CPU states: 12.7% user,  0.0% nice,  3.8% system,  83.4% idle
Memory: Real: 9920M/23G act/tot  Virtual: 6544M/77588M use/tot  Free: 8279M

  PID  USERNAME  PRI  NICE  SIZE  RES  STATE  TIME  CPU  COMMAND
2438278  bscs011   44    0 1041M  947M  run    15:04  91.60%  rlh
2585143  ora011    42    0 1863M  342M  WAIT   0:35  12.90%  oracle
2097152  root      0     0   36G  1112M  run    380.5H  11.90%  kernel idle
2525877  ora011    44    0 1998M  782M  sleep  2:31  11.60%  oracle
2574379  ora011    44    0 1817M   58M  sleep  0:59  10.60%  oracle
2574902  ora011    42    0 1863M 1308M  WAIT   1:44   6.70%  oracle
2531905  ora011    42    0 1926M  358M  WAIT  12:00   5.90%  oracle
2468627  ora011    44    0 1945M  728M  sleep  1:39   3.70%  oracle
2557324  ora011    42    0 1894M  446M  WAIT  17:24   3.20%  oracle
2103572  root      42    0    0K    0K  WAIT  67:06   1.70%  icssvr_daemon_
2261541  ora011    44    0 1818M 1502M  sleep  18:04   1.10%  oracle
2567179  ora011    42    0 1824M  825M  WAIT   1:29   1.00%  oracle
2531304  ora011    44    0 1823M  470M  sleep  1:39   0.90%  oracle
2248084  dshsm99   44    0  928K  6135K  run    0:01   0.80%  top
2401896  ora011    44    0 1817M  9437K  sleep  0:00   0.80%  oracle
    
```

26 April 2004

12



Tools

```
# monitor
```

26 April 2004

13



Commands: sched_stat

```
# sched_stat -u
```

```
hostname:      diana.deu.hp.com  
processors:    3
```

Processor Usage

cpu	user	nice	system	idle	widle	scalls	intr	csw	tbsyc
0	0.6	0.0	3.9	95.1	0.4	607378654	3185720635	2391037123	89784553
1	3.8	0.0	4.3	91.4	0.4	1791345294	0	2515857324	57643768
2	2.9	0.0	4.0	92.7	0.4	1556698624	0	2570010805	61870897

26 April 2004

14



Commands: lockinfo

```
# lockinfo -sort=waitsum sleep 60
hostname:      diana.deu.hp.com
locktype:     Simple Locks
lockmode:     2 (SMP default)
processors:    3
start time:   Thu Apr 15 16:44:20 2004
end time:     Thu Apr 15 16:45:20 2004
command:      sleep 60
```

	tries	reads	trmax	misses	misses percent	sleeps	waitmax seconds	waitsum seconds
vm_umap.vm_lock (C)	17831	9900	2	61	0.3	61	0.00098	0.00293
vdT.vdIoLock (S)	955996	0	0	380	0.0	0	0.00005	0.00276
clua_port_lock (C)	403	99	1	15	3.7	15	0.00098	0.00098
vm_object.ob_lock (S)	1356064	0	0	235	0.0	0	0.00002	0.00060

26 April 2004

15



Tools: collect

- Speichert die Daten in einer komprimierten Form und/oder gibt sie sofort lesbar aus
- graphische Oberfläche (collgui) zum Auswerten
- Startet automatisch beim Boot

```
# rcmgr set COLLECT_AUTORUN 1
# rcmgr set COLLECT_ARGS ""
```

Default:

```
-i60,120 -f /var/adm/collect.dated/collect
-W 1h -M 10,15
```

26 April 2004

16



Kernel Parameter anzeigen und ändern

```
# dxkerneltuner &
```

Attribute Name	Current Value	Boot Time Value	Minimum
ubc_minpercent	10	10	
ubc_maxpercent	100	100	
ubc_borrowpercent	20	20	
vm_max_wrprio_kluster	32768	32768	8
vm_max_rdprio_kluster	16384	16384	8

Kernel Tuner on diana.deu.hp...
File Options Help
pci
platform_common
ppp
proc
ps2
psm
pts

vga
vm
ws

Select Subsystem

26 April 2004 18



System Resource Limits

- Maximale Anzahl von Prozessen / Threads
- Memory pro Prozess
- Inter Process Communication

26 April 2004

19



Memory Tuning

- Physical Memory
 - Free
 - Wired
 - Active, Inactive, UBC
- Unified Buffer Cache
- Process Memory
- Page Reclamation
 - paging
 - swapping

26 April 2004

20



Shared Memory

- Segmented Shared Memory (SSM)
- Granularity Hints (large pages)

26 April 2004

21

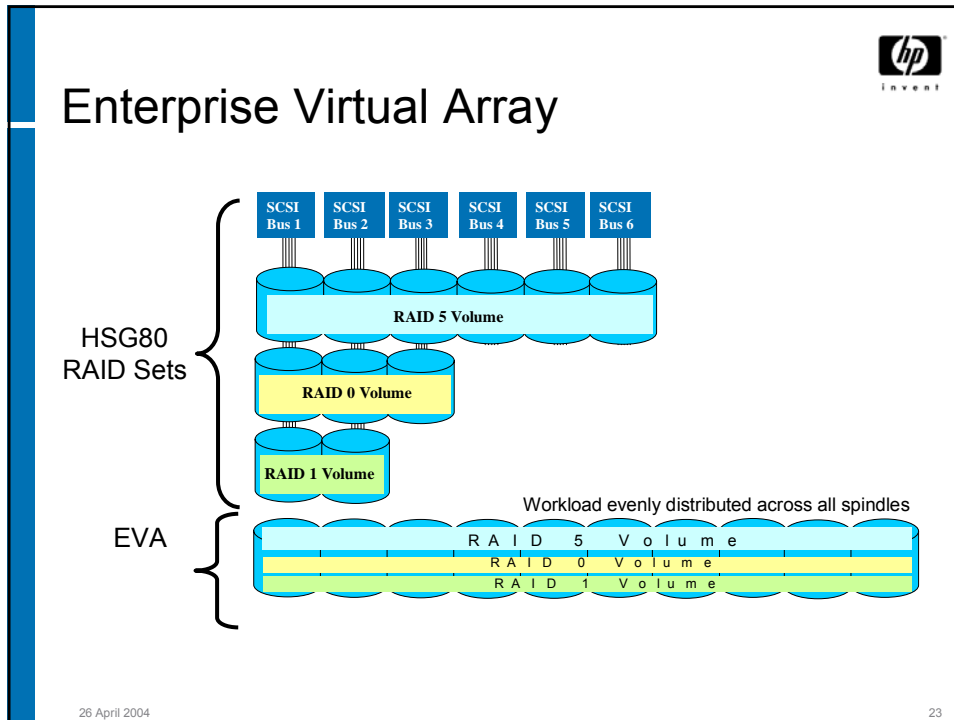


I/O Tuning

- ... Grundsätzlich:
 - Eine Disk kann ca. 150 I/O / Sekunde
- Folgerungen
 - Auf möglichst viele Drives verteilen
 - Möglichst große I/Os
- Host Based Striping als zweitbeste Lösung
- RAID Controller
 - HSZxx HSG80 EVA

26 April 2004

22



-
- The AdvFS features list includes:
- Unified Buffer Cache
 - Transaction Log
 - Direct I/O
 - Fragmentierung
 - defragment
 - vfast
- HP invent
- ## AdvFS
- Unified Buffer Cache
 - Transaction Log
 - Direct I/O
 - Fragmentierung
 - defragment
 - vfast
- 26 April 2004 24



Tuning Oracle

- Shared Memory
 - SSM oder gh_chunks
- NUMA
- UBC Tuning
- Interprocess Communication
- Internet Attributes
- Process Attributes
- gettimeofday()

26 April 2004

25




Zusammenfassung

- Einfluß der Application und der zeitlichen Aufteilung der Last
- Identifizieren einer adäquaten HW Konfiguration
- Kontinuierliches Monitoring, um Engpässe und langfristige Trends zu erkennen
- Ausreichende Leistungsreserven, um eine hohe Verfügbarkeit zu garantieren
- Anpassen der Parameter für ein Subsystem beeinflusst immer andere Subsysteme

26 April 2004

26



 invent

Weitere Informationen

- Tru64 UNIX Dokumentation
„System Configuration and Tuning“
- Best Practice Dokumentation
<http://tru64unix.compaq.com>

26 April 2004 28

