



MAKLEE

software engineering
solutions

ORACLE PARTNER

Oracle on OpenVMS

Guy Peleg

President

Maklee Engineering

guy.peleg@maklee.com

Place Yourself in the Hands of the Experts

Disclaimer

Ich hatte eine großartige Präsentation geplant für heute

Aber, leider die meisten von Ihnen haben bereits gesehen.

Dies ist das zweite Mal, dass Herr Kriebel zwingt mich zu Schreiben Sie eine neue Präsentation Was gibt's Freunde sind für ;-)



Agenda

- Characteristics of an Oracle workload on OpenVMS
- OpenVMS Optimization Techniques
- General Optimization Techniques
- Statspack / AWR reports



Who we are?

- Pledge to provide the highest level of technical consulting
 - No Results, no payment.
- Our staff contains former members of OpenVMS engineering
- Specialize in:
 - Performance Tuning (focusing on applications not only the operating system)
 - Oracle & Oracle tuning (Oracle Partner)
 - Platform migration (Alpha to Integrity)
 - Custom engineering



MAKLEE

Who we are?

- Serving mission critical customers all over the world, including the US, France, Germany, Switzerland, Israel, Belgium, Argentina, Italy & Sweden.
- Read some of our success stories:
<http://www.maklee.com/news.html>



MAKLEE

Oracle on OpenVMS

- Oracle is “just another” user mode application
 - (Almost) O/S agnostic
 - Data files compatible between platforms
 - Does not rely on OpenVMS specific mechanism (locking, ASTs, etc.)
- Well tuned Oracle server would scale up nicely



Oracle on OpenVMS

- The following slides demonstrate the behavior of an Oracle server from an OpenVMS perspective.
- Alpha Server 32P GS1280 7/1300
- OpenVMS V8.3
 - 128GB RAM
 - 40GB SGA
- Server is running at 100% utilization
 - End of day processing

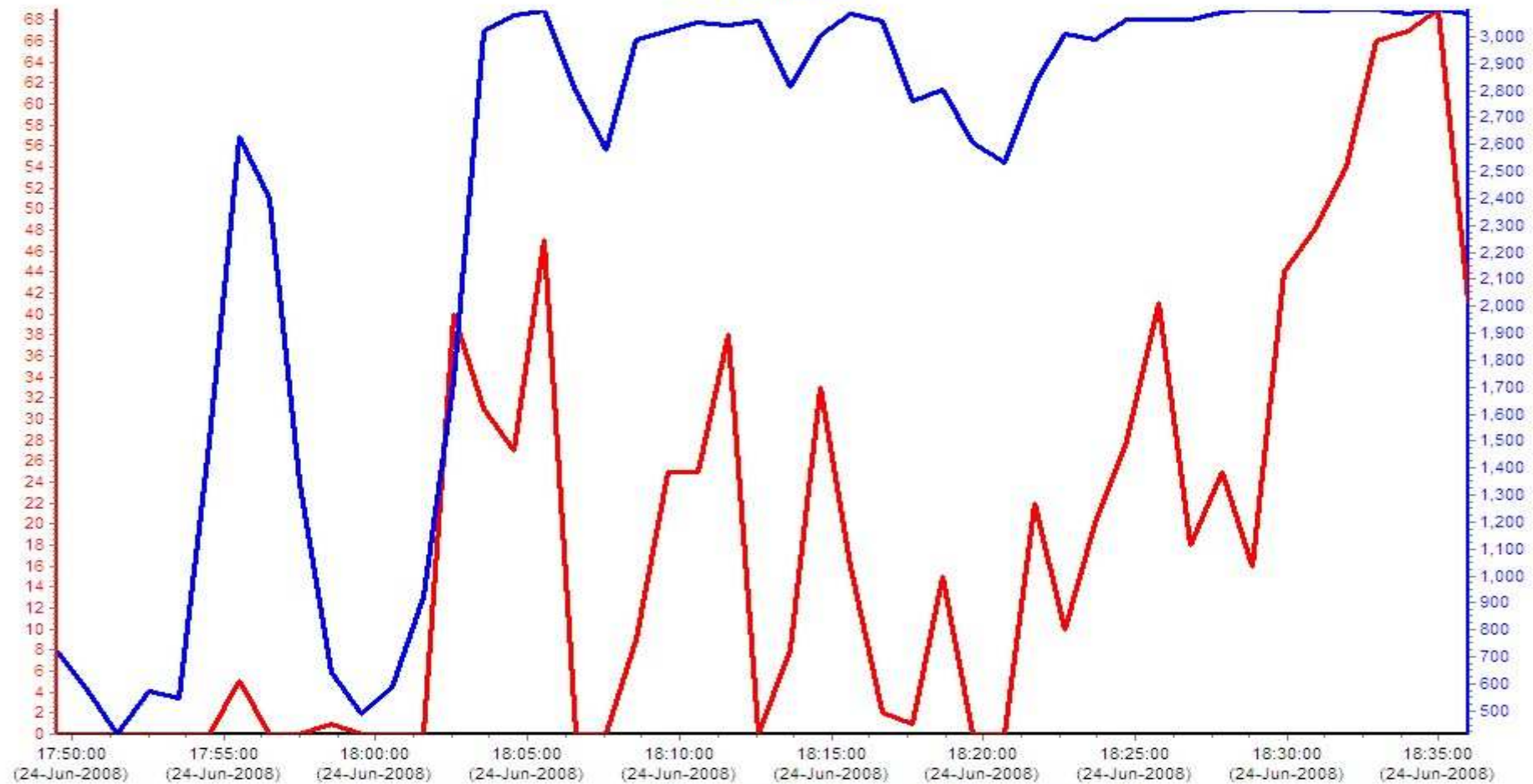


CPU Utilization



Oracle Server

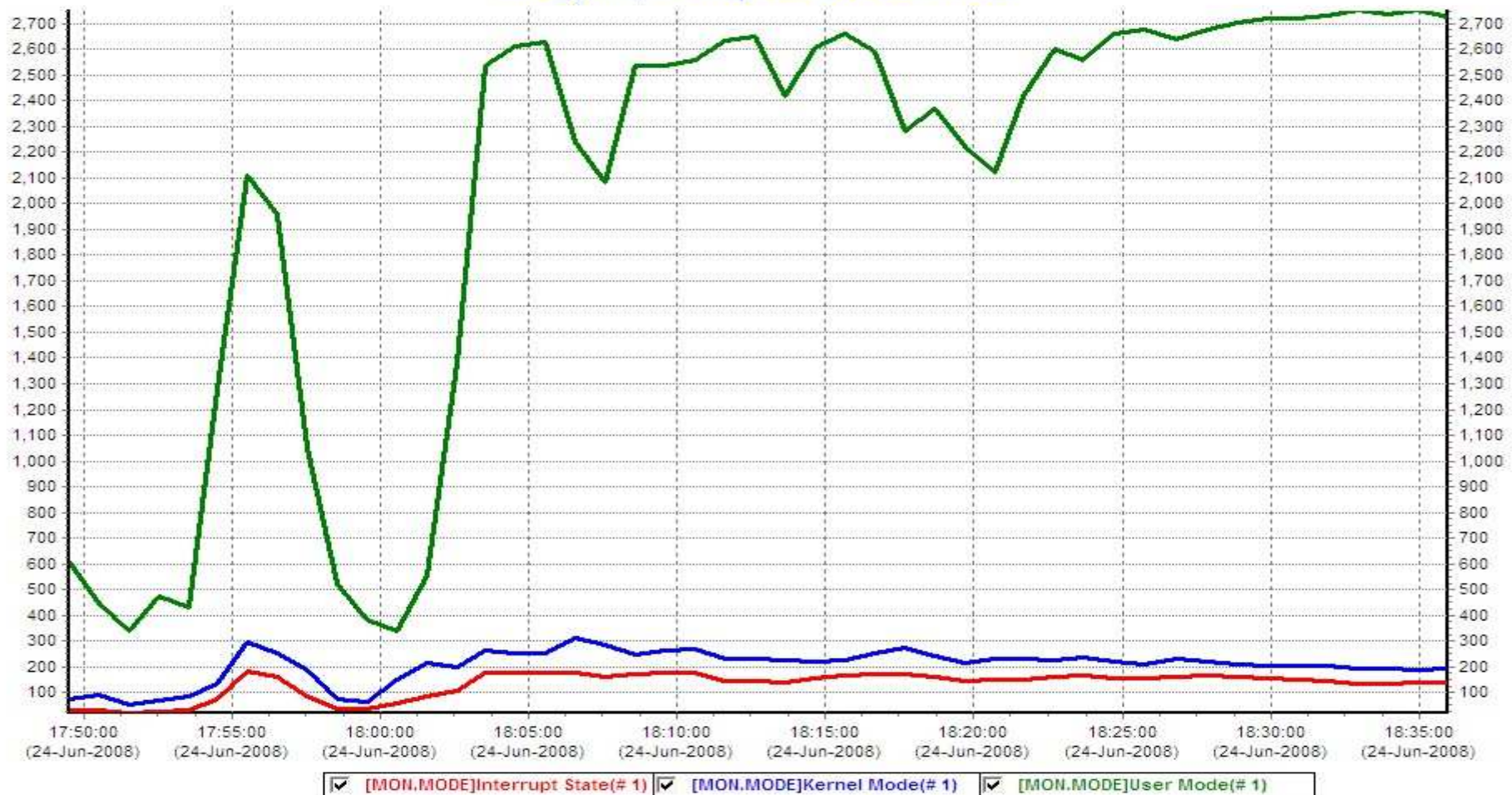
CPU Utilization



☒ [MON.STAT]Compute(# 1) ☒ [MON.SYST]Cpu Busy(# 1)

CPU Utilization

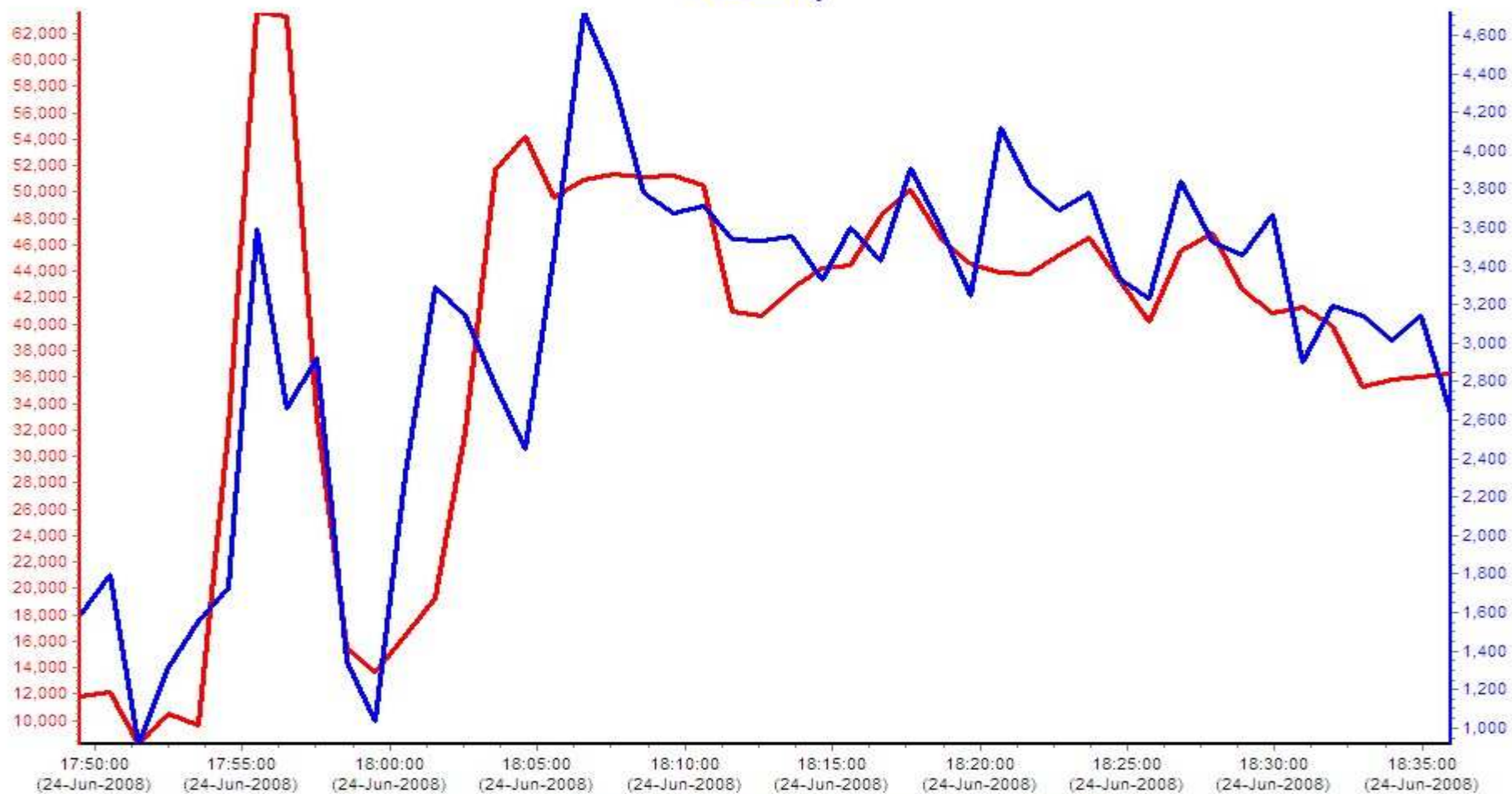
Oracle Server



I/O Activity

Oracle Server

I/O activity



☒ [MON.SYST]Buffered I/O Rate(# 1) ☒ [MON.SYST]Direct I/O Rate(# 1)

I/O Activity

- A well tuned Oracle server does not perform many physical I/Os
- Inspire to write-only physical I/Os
 - Quick quiz
 - Which I/O operation is faster, read or write?
- Redo log files size & performance is critical

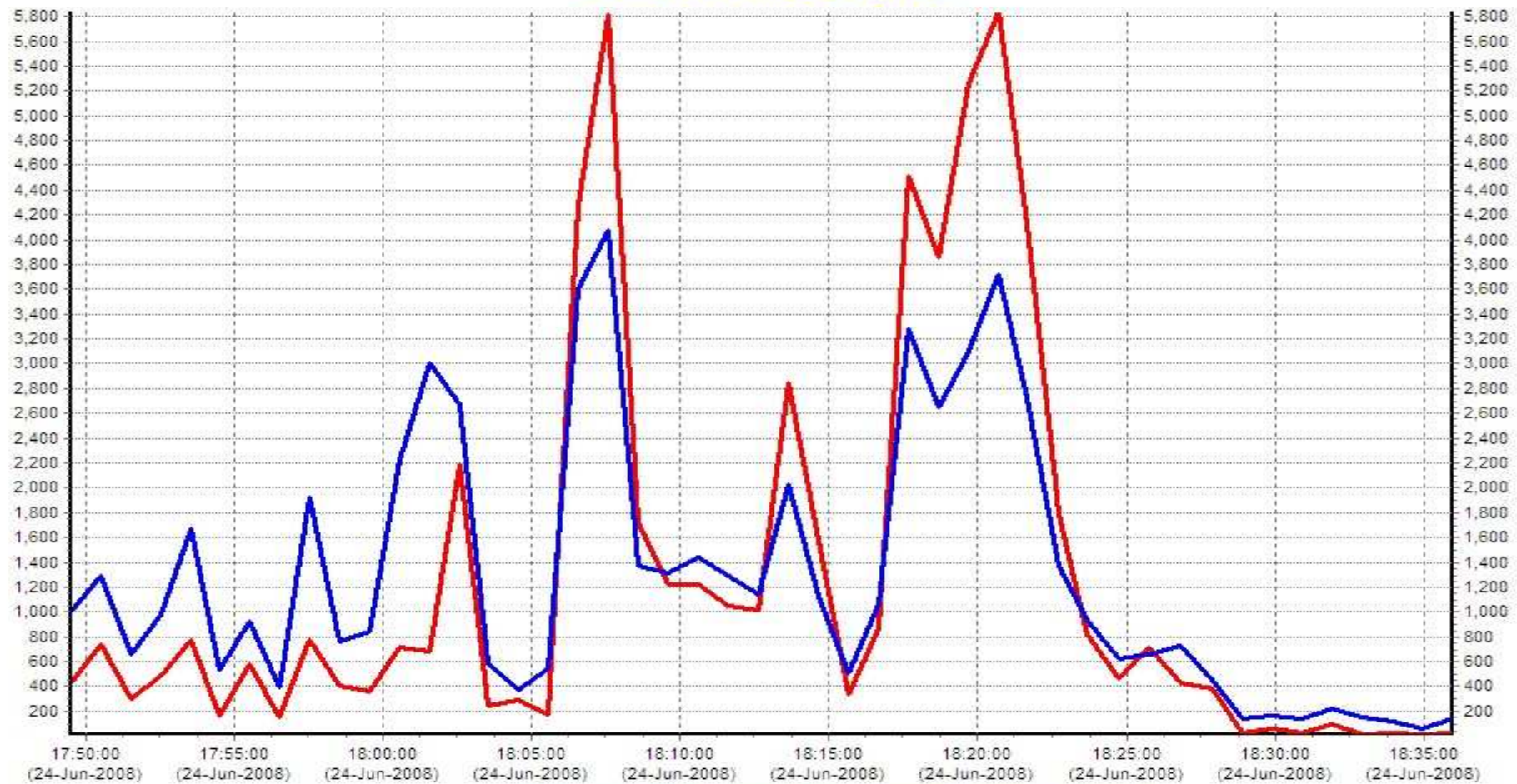


Locking



Oracle Server

Converted & New ENQ rate



☒ [MON.LOCK]Converted ENQ rate(# 1) ☒ [MON.LOCK]New ENQ rate(# 1)

Locking

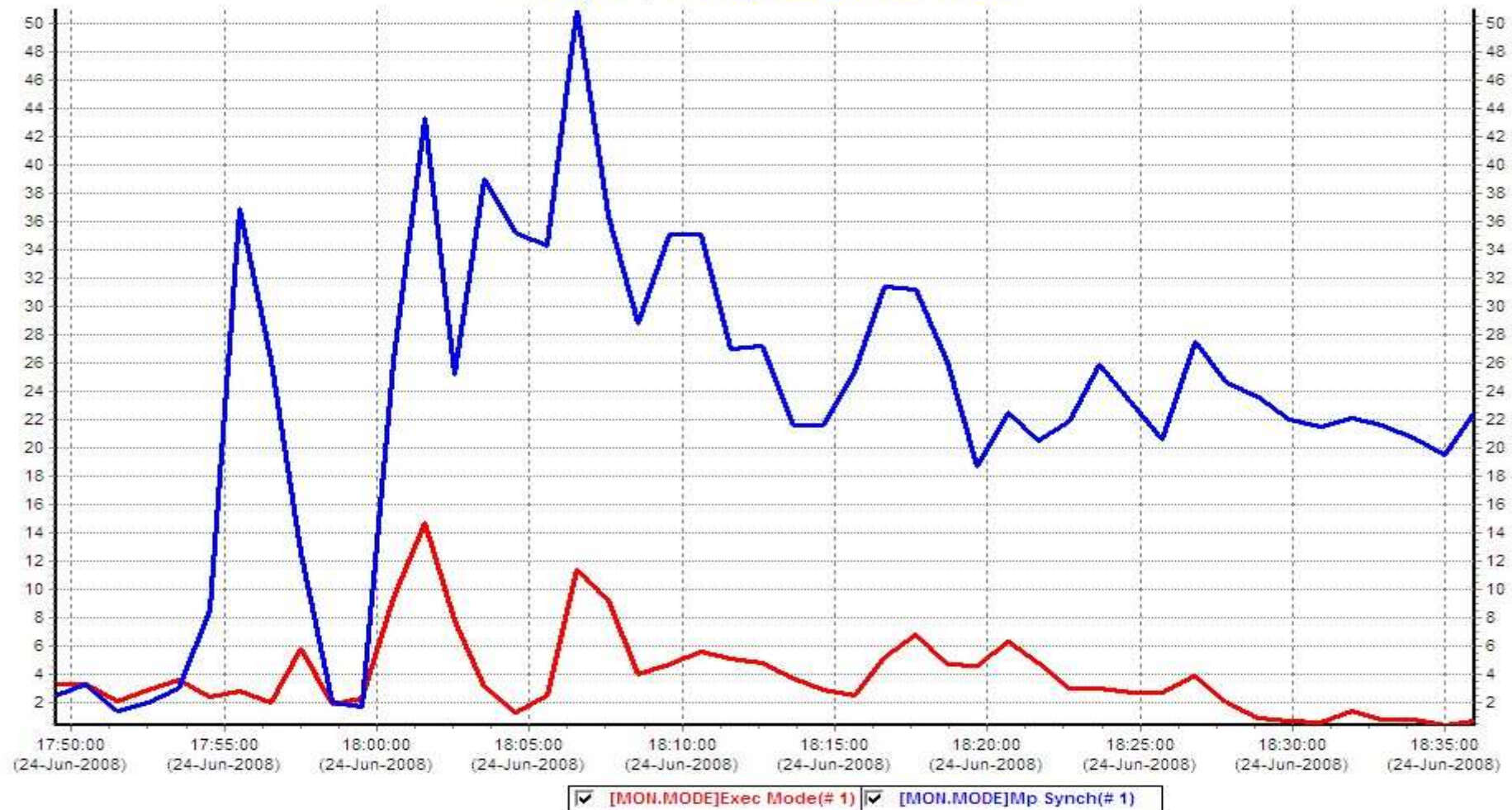
- Oracle implements it's own locking mechanism.
- Locks are done in user mode.
- Dedicated lock manager is not required for an Oracle server.



Exec mode & MP Synchron



Oracle Server



MP Synch

- MP Synch time is lost processor time
 - CPU A holding a spinlock while CPU B is forced to wait for the same spinlock.
- MP Synch significantly impacts scaling
 - (in a bad way of course ;-)
- Oracle implements spinlocks in user mode as latches.
 - A process spins or sleeps when a latch is not available
 - Typically a sign for:
 - Hot block in the database
 - One of the areas in the SGA is too small





OpenVMS Optimization Techniques



MAKLEE

OpenVMS Optimization Techniques

- Resident images
- Reserved memory for the SGA
- Enable HyperThreads
 - (when appropriate, not suitable for all workloads)
- Increase the size of the VHPT
- Process quotas
 - Properly size working sets
 - Typically 8MB per connection

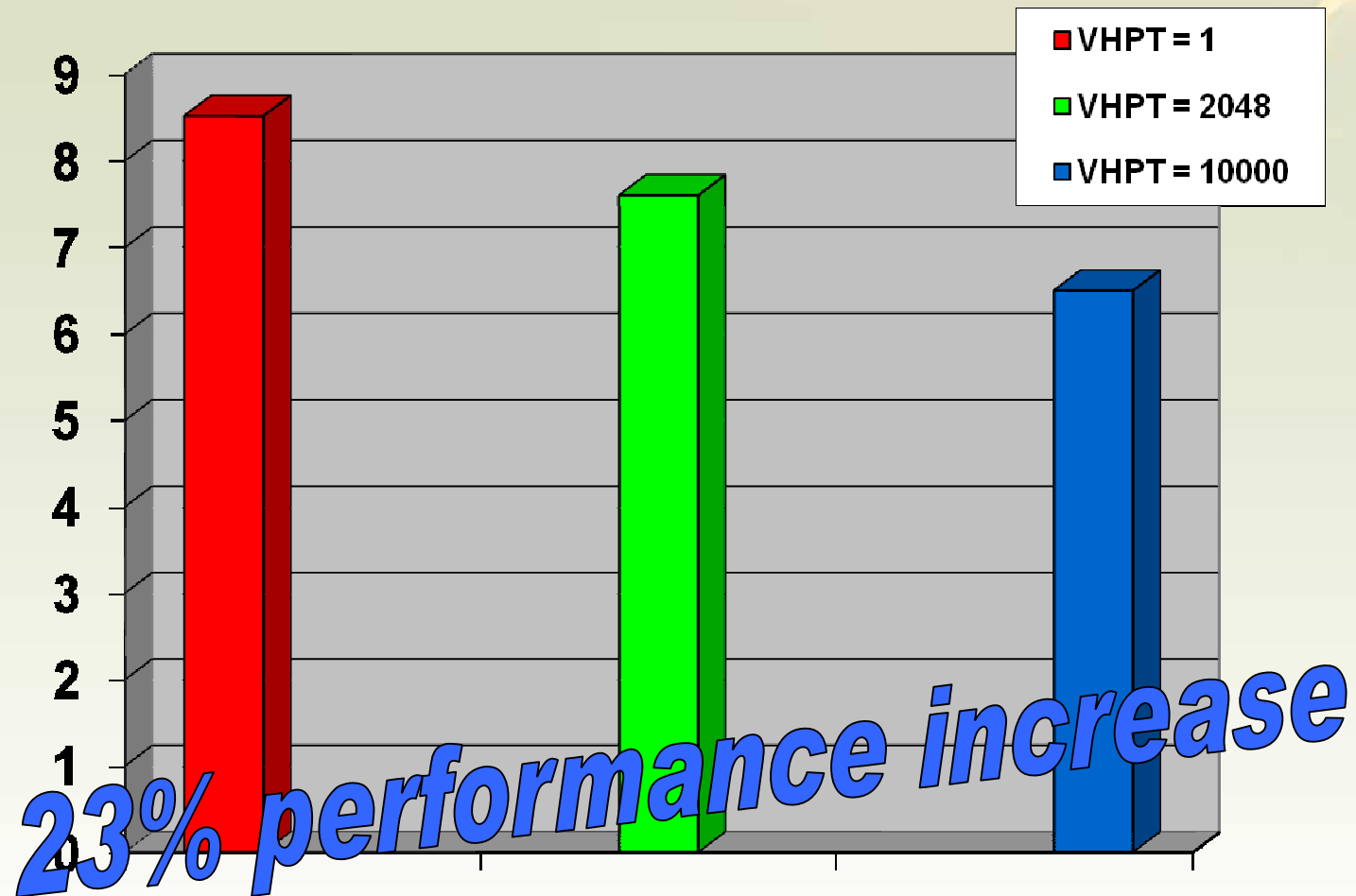


VHPT Benchmark

- The following charts illustrate the impact of increasing the VHPT made on Oracle batch jobs
- rx6600 – 8 cores
 - OpenVMS V8.3-1H1
 - EVA8000
 - Oracle 10gR2
 - HyperThreads Enabled
 - 64 GB of physical memory



Oracle Batch job A

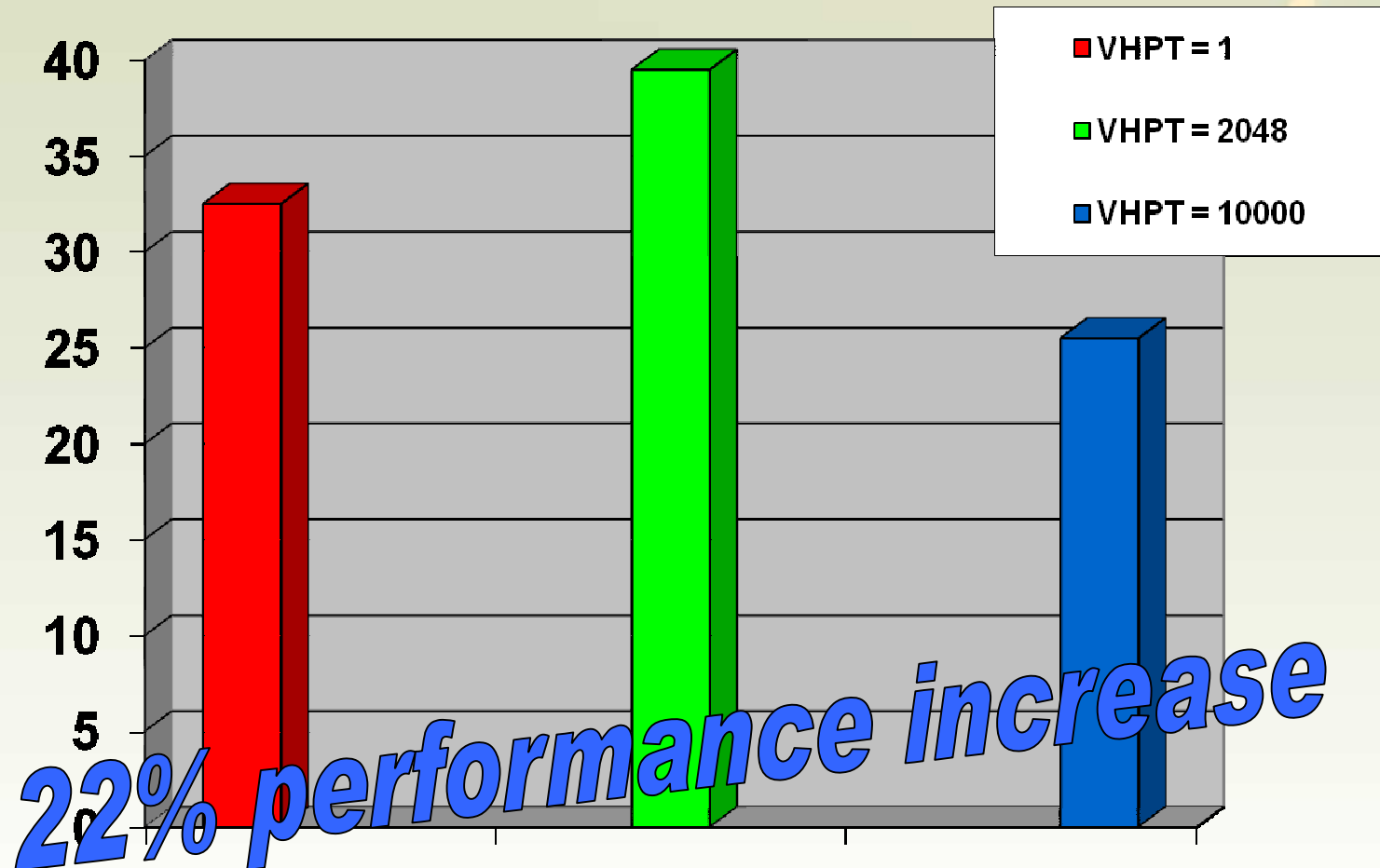


Elapsed Time in Minutes (less is better)



MAKLEE

Oracle Batch job B



MAKLEE

Elapsed Time in Minutes (less is better)

Cluster interconnect Performance

- Oracle requires a dedicated cluster interconnect for RAC traffic.
 - Latency should be lower than 15ms
- Enabling Jumbo Frames is a must !
- OpenVMS achieved 0.5ms on
 - blades RAC (BL860)
 - V8.3-1H1
 - Gigabit Ethernet
 - Jumbo Frames enabled



Cluster interconnect Performance

- Use the following query to measure the latency of the interconnect:

```
set numwidth 20
column "AVG CR BLOCK RECEIVE TIME (ms)" format 9999999.9
select
  b1.inst_id,
  b2.value "GCS CR BLOCKS RECEIVED",
  b1.value "GCS CR BLOCK RECEIVE TIME",
  ((b1.value/b2.value) * 10) "AVG CR BLOCK RECEIVE TIME (ms)"
from gv$sysstat b1,
     gv$sysstat b2
where b1.name='gc cr block receive time'
and b2.name='gc cr blocks received'
and b1.inst_id=b2.inst_id;
```



Cluster Interconnect performance

- AWR reports the following statement as the top statement generating cluster wait time in a RAC

SQL ordered by Cluster Wait Time

Cluster Wait Time (s)	CWT % of Elapsed Time	Elapsed Time(s)	CPU Time(s)	Executions	SQL Id	SQL Module	SQL Text
28.15	85.80	32.81	6.14	1	92x4ys2kta27t	SQL*Plus	select xxxxxxxxxxxxxxxxxxxxx

- After enabling Jumbo Frames the query is running 4 times faster

SQL ordered by Cluster Wait Time

Cluster Wait Time (s)	CWT % of Elapsed Time	Elapsed Time(s)	CPU Time(s)	Executions	SQL Id	SQL Module	SQL Text
1.15	15.97	7.2	6.14	1	5xphi2ktrsw2	SQL*Plus	select xxxxxxxxxxxxxxxxxxxxx



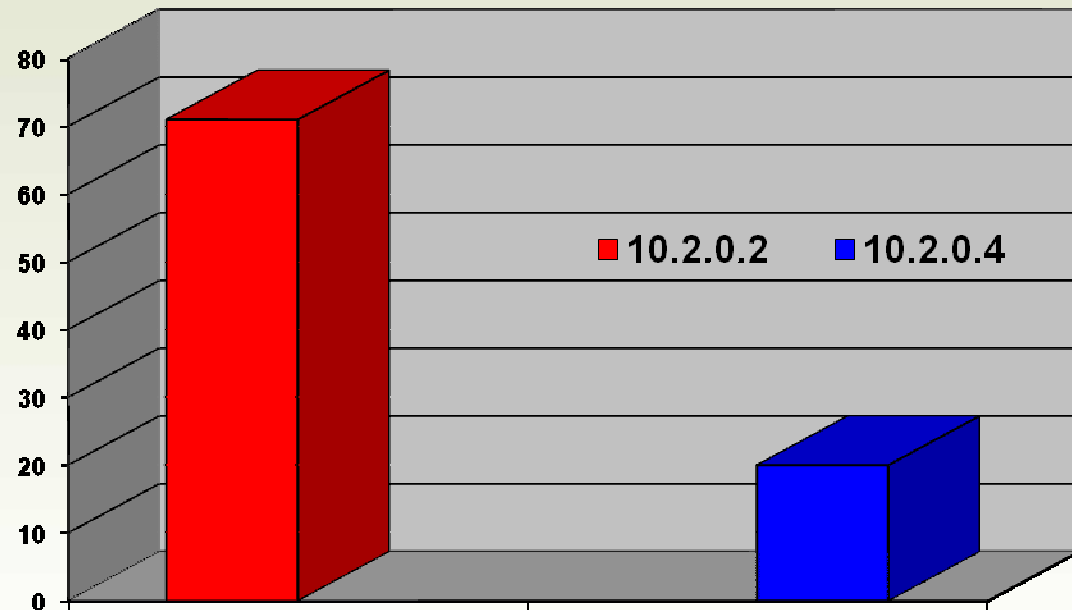
Stay current

- OpenVMS V8.3-1H1 on Itanium
- OpenVMS V8.3 on Alpha
- Oracle 10.2.0.4 on both platforms



10.2.0.4

- Many bugfixes + performance enhancements
- Multiple DB writers
- Database Vault



MAKLEE

Elapsed time (minutes) to export 25GB
Less is better

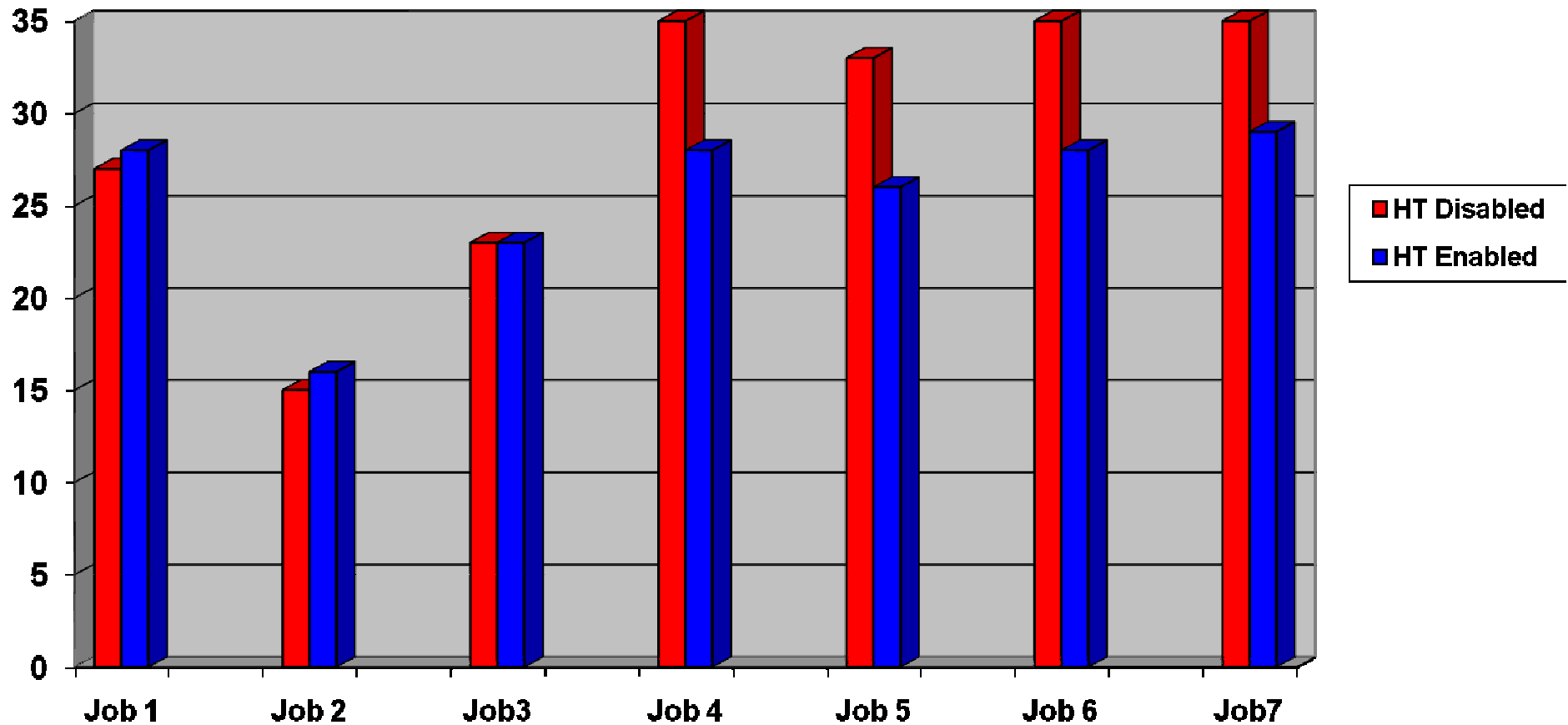
HyperThreads

- > According to Intel's marketing numbers, HTs provide up to 25% performance increase
- > Applications with poor locality are good candidates to benefit from HyperThreads
 - An Oracle OLTP DB perfectly fits the profile
- > Not suitable for all applications
 - May degrade performance for some applications

> YMMV !!!

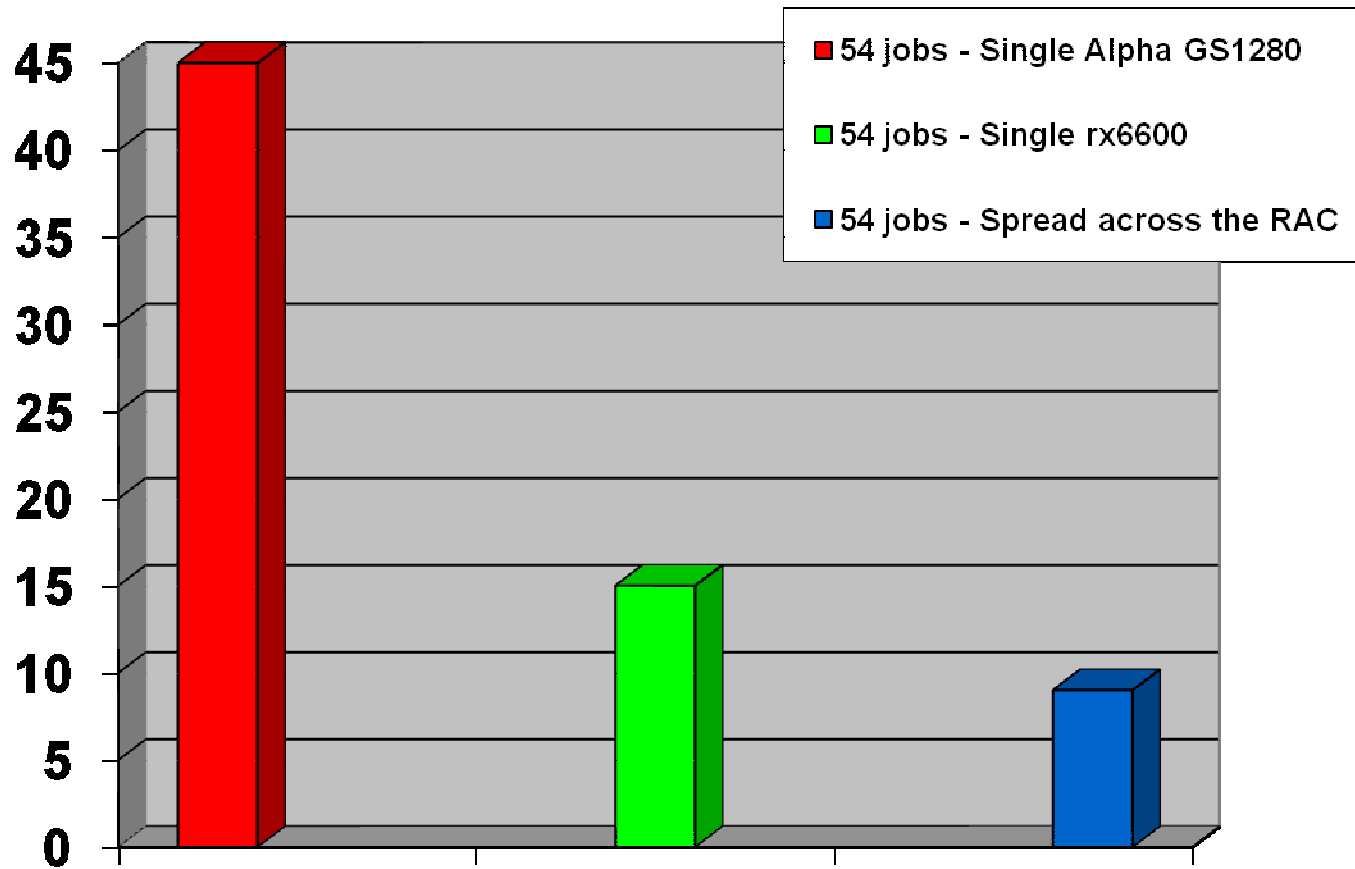


HyperThreads – Impact on Oracle Jobs



Elapsed time (minutes) to execute 7 jobs
Less is better

End of Day Benchmark



Minutes to complete batch processing cycle

Less is better

End of Day Benchmark

- > Itanium outperformed Alpha
- > RAC allows scaling outside of the box
- > Second RAC node adds 40% more throughput



Alpha Vs. Itanium - Summary



- > All customers we worked with doubled performance by moving from Alpha to Itanium.
- > Oracle on OpenVMS Itanium outperforms Oracle on Alpha
 - Memory bandwidth
 - Larger caches on the CPU
 - HyperThreads
 - Montvale
- > Oracle 10g RAC increases the availability of your database.
 - Allows scaling outside of the box
 - YMMV !!





General Optimization Techniques (not VMS specific)



System statistics

- Gathering schema statistics is a must. System statistics is equally important.
- Prior to 9i the CBO based it's calculations on the number of I/O requests that would needed to satisfy a query.
- Starting with 9i, CPU cost has been added to the algorithm
 - Turned off by default unless system statistics available
- In Oracle 10g system statistics collects more information about I/O
 - New CPU speed
 - Seek time
 - Throughput



Critical!! for certain optimization options in 10g

System statistics

- System statistics is common for all the nodes in the RAC
 - Do not collect system statistics if you are using a non symmetrical hardware configuration.
- The information will be used by the CBO to produce better execution plans.



Missing Indexes

- It is a common knowledge that indexes are key to good database performance.
- With complex SQL statements, finding a missing index is not always trivial.
- Oracle DBMS_ADVISOR (introduced with Oracle 10.1) can assist in finding a missing index.
 - May be used with a representative workload or a single statement.
 - Recommends indexes or materialized views.



Using DBMS_ADVISOR

- Create a directory for the results
 - SQLPLUS> CREATE OR REPLACE DIRECTORY advisor
AS '/ora_root/advisor_data';
 - SQLPLUS> GRANT READ,WRITE ON DIRECTORY advisor TO PUBLIC;
- Execute the advisor against the target SQL statement
 - SQLPLUS> EXECUTE dbms_advisor.quick_tune (-
DBMS_ADVISOR.SQLACCESS_ADVISOR, -
task_name => 'TASK1', -
attr1 => 'SELECT c2 FROM t1 WHERE c1 = :b1');



Using DBMS_ADVISOR

- Generate the advice script
 - SQLPLUS> EXECUTE dbms_advisor.create_file (-
buffer => dbms_advisor.get_task_script (task_name => 'TASK1'),
location => 'ADVISOR', -
filename => 'task1.sql');



Using DBMS_ADVISOR

- The advice script contains SQL statements for creating the missing index

Rem SQL Access Advisor: Version 10.2.0.2.0 - Production

Rem

Rem Username: GUY

Rem Task: TASK1

Rem Execution date: 09/23/2008 22:42

Rem

```
CREATE INDEX "US01"."T1_IDX$$_15180001"  
ON "US01"."T1"  
("C1")  
COMPUTE STATISTICS;
```



Unused Indexes

- Oracle allows monitoring index usage
 - SQLPLUS> ALTER INDEX myIdx MONITORING USAGE;
- To check which indexes have been used by the optimizer:
 - SQLPLUS> SELECT index_name,used FROM v\$object_usage;
- To disable index monitoring
 - SQLPLUS> ALTER INDEX myIdx NOMONITORING USAGE;





Statspack / AWR reports



MAKLEE

Statspack / AWR reports

- The statspack and AWR reports provide all the information required for performance analysis.
- Typical usage:
 - Create a snapshot
 - Run workload
 - Create another snapshot
 - Generate a report
- The first 2 pages provide an overview of performance, highlighting areas require attention
 - Followed by detailed information on all aspects of the database
- Look at the results at least once a month
 - The database keeps changing
 - Small tables grow large...



Statspack / AWR reports

sp_060308_18_20 - Notepad

File Edit Format View Help

STATSPACK report for

DB Name	DB Id	Instance	Inst Num	Release	Cluster	Host
XXXXXX	3985156505	XXXXXX	1	9.2.0.7.0	NO	XXXXXX

	Snap Id	Snap Time	Sessions	Curs/Sess	Comment
Begin Snap:	18709	03-Jun-08 17:49:25	1,576	#####	
End Snap:	18713	03-Jun-08 20:00:02	1,459	#####	
Elapsed:		130.62 (mins)			

Cache Sizes (end)

Buffer Cache:	4,800M	Std Block size:	8K
Shared Pool Size:	6,560M	Log Buffer:	4,096K

Load Profile

	Per Second	Per Transaction
Redo size:	5,381,608.61	26,392.47
Logical reads:	222,084.05	1,089.14
Block changes:	22,065.85	108.22
Physical reads:	1,450.86	7.12
Physical writes:	968.61	4.75
User calls:	5,805.19	28.47
Parses:	2,117.70	10.39
Hard parses:	1.02	0.01
Sorts:	675.42	3.31
Logons:	0.24	0.00
Executes:	3,788.70	18.58
Transactions:	203.91	

% Blocks changed per Read:	9.94	Recursive Call %:	37.40
Rollback per transaction %:	22.88	Rows per Sort:	35.84

Instance Efficiency Percentages (Target 100%)

Buffer Nowait %:	99.79	Redo Nowait %:	100.00
Buffer Hit %:	99.37	In-memory Sort %:	100.00
Library Hit %:	99.96	Soft Parse %:	99.95
Execute to Parse %:	44.10	Latch Hit %:	98.87
Parse CPU to Parse Elapsd %:	14.49	% Non-Parse CPU:	98.33

Shared Pool Statistics

	Begin	End
Memory Usage %:	93.55	85.94
% SQL with executions>1:	57.91	62.14
% Memory for SQL w/exec>1:	61.60	63.67

Top 5 Timed Events

% Total

start

Wireless Network Co...

Session control windo...

Inbox - Microsoft Out...

Orade on OpenVMS

sp_060308_18_20 - ...

EN

11:41 PM

Statspack / AWR reports

sp_060308_18_20 - Notepad

File Edit Format View Help

Begin Snap: 18709 03-Jun-08 17:49:25 1,576 #####
End Snap: 18713 03-Jun-08 20:00:02 1,459 #####
Elapsed: 130.62 (mins)

Cache Sizes (end)

Buffer Cache: 4,800M Std Block size: 8K
Shared Pool Size: 6,560M Log Buffer: 4,096K

Load Profile

	Per Second	Per Transaction
Redo size:	5,381,608.61	26,392.47
Logical reads:	222,084.05	1,089.14
Block changes:	22,065.85	108.22
Physical reads:	1,450.86	7.12
Physical writes:	968.61	4.75
User calls:	5,805.19	28.47
Parses:	2,117.70	10.39
Hard parses:	1.02	0.01
Sorts:	675.42	3.31
Logons:	0.24	0.00
Executes:	3,788.70	18.58
Transactions:	203.91	

% Blocks changed per Read: 9.94 Recursive Call %: 37.40
Rollback per transaction %: 22.88 Rows per Sort: 35.84

Instance Efficiency Percentages (Target 100%)

Buffer Nowait %:	99.79	Redo Nowait %:	100.00
Buffer Hit %:	99.37	In-memory Sort %:	100.00
Library Hit %:	99.96	Soft Parse %:	99.95
Execute to Parse %:	44.10	Latch Hit %:	98.87
Parse CPU to Parse Elapsed %:	14.49	% Non-Parse CPU:	98.33

Shared Pool Statistics	Begin	End
Memory Usage %:	93.55	85.94
% SQL with executions>1:	57.91	62.14
% Memory for SQL w/exec>1:	61.60	63.67

Top 5 Timed Events

Event	waits	Time (s)	% Total Elapsed Time
CPU time		99,774	45.22
latch free	17,660,460	58,639	26.58
db file sequential read	6,115,552	26,689	12.10
buffer busy waits	3,724,937	12,191	5.52
log file sync	1,386,290	4,950	2.24

start

Wireless Network Co...

Session control windo...

Inbox - Microsoft Out...

Oracle on OpenVMS

sp_060308_18_20 - ...

EN

11:43 PM

Statspack / AWR reports

- A well tuned database should report high hit rates
- In our example Latch hit rate was lower than 99%
 - The accumulated wait time for the latch free event was 16 hours !!
 - 58,639 seconds
 - 0.5 hour per CPU in a 2.25 hours period



Latching activity

sp_060308_18_20 - Notepad

File Edit Format View Help

Latch	Get Requests	Pct Get Miss	Avg Slps /Miss	wait Time (s)	Nowait Requests	Pct Nowait Miss
Consistent RBA	1,044,757	0.4	0.0	0	0	
FAL request queue	107	0.0		0	0	
FIB s.o chain latch	167	0.0		0	0	
FOB s.o list latch	32,260	0.1	0.0	0	0	
SQL memory manager worka	268	0.0		0	0	
active checkpoint queue	60,867	0.0	0.0	0	0	
address list	2	0.0		0	0	
alert log latch	290	0.0		0	0	
archive control	203	0.0		0	0	
archive process latch	228	0.0		0	0	
begin backup scn array	53,539	0.0	0.0	0	0	
cache buffer handles	48,803,821	0.2	0.0	3	0	
cache buffers chains	3,475,261,920	0.7	0.1	8851	24,945,456	0.5
cache buffers lru chain	7,990,788	0.2	0.3	6	26,957,860	0.5
channel handle pool latc	3,041	0.0		0	0	
channel operations paren	24,898	0.0	0.0	0	0	
checkpoint queue latch	22,996,588	0.0	0.0	0	6,386,146	0.1
child cursor hash table	129,817	0.0	0.0	0	0	
commit callback allocati	6,602	0.0		0	0	
constraint object alloca	6,452	0.0		0	0	
dictionary lookup	662	0.0		0	0	
dispatcher configuration	0			0	3	0.0
dml lock allocation	13,733,186	3.3	0.4	333	0	
dummy allocation	3,857	0.0	0.0	0	0	
enqueue hash chains	22,944,859	0.5	0.4	102	0	
enqueues	5,860,684	0.3	0.1	5	0	
event group latch	1,109	0.0		0	0	
global tx hash mapping	135,546	0.0		0	0	
hash table column usage	1,472	0.0		0	636,201	0.0
hash table modification	4	0.0		0	0	
internal temp table obje	26	0.0		0	0	
job workq parent latch	0			0	580	3.3
job_queue_processes para	405	0.0		0	0	
kmcptab latch	21,979	0.0	1.0	0	0	
kmcpvec latch	0			0	21,933	0.2
ktm global data	1,218	0.0		0	0	
lgwr LWN SCN	1,050,451	1.0	0.0	0	0	
library cache	167,523,774	5.7	0.7	23609	4,116	239.3
library cache load lock	9,604	0.0		0	0	
library cache pin	142,262,584	0.6	0.4	645	0	
library cache pin alloca	42,234,047	0.5	0.3	115	0	
list of block allocation	888,169	0.1	0.3	0	0	
loader state object free	9,182	0.0		0	0	
longop free list parent	2,821	0.0		0	183	0.0
message pool operations	1,384	0.1	0.0	0	0	
messages	3,307,332	0.1	0.2	1	0	
mostly latch-free SCN	1,170,321	8.3	0.0	0	0	
multiblock read objects	2,328,566	0.2	0.1	0	21	0.0

start

Wireless Network Co...

Session control windo...

Inbox - Microsoft Out...

Oracle on OpenVMS

sp_060308_18_20 - ...

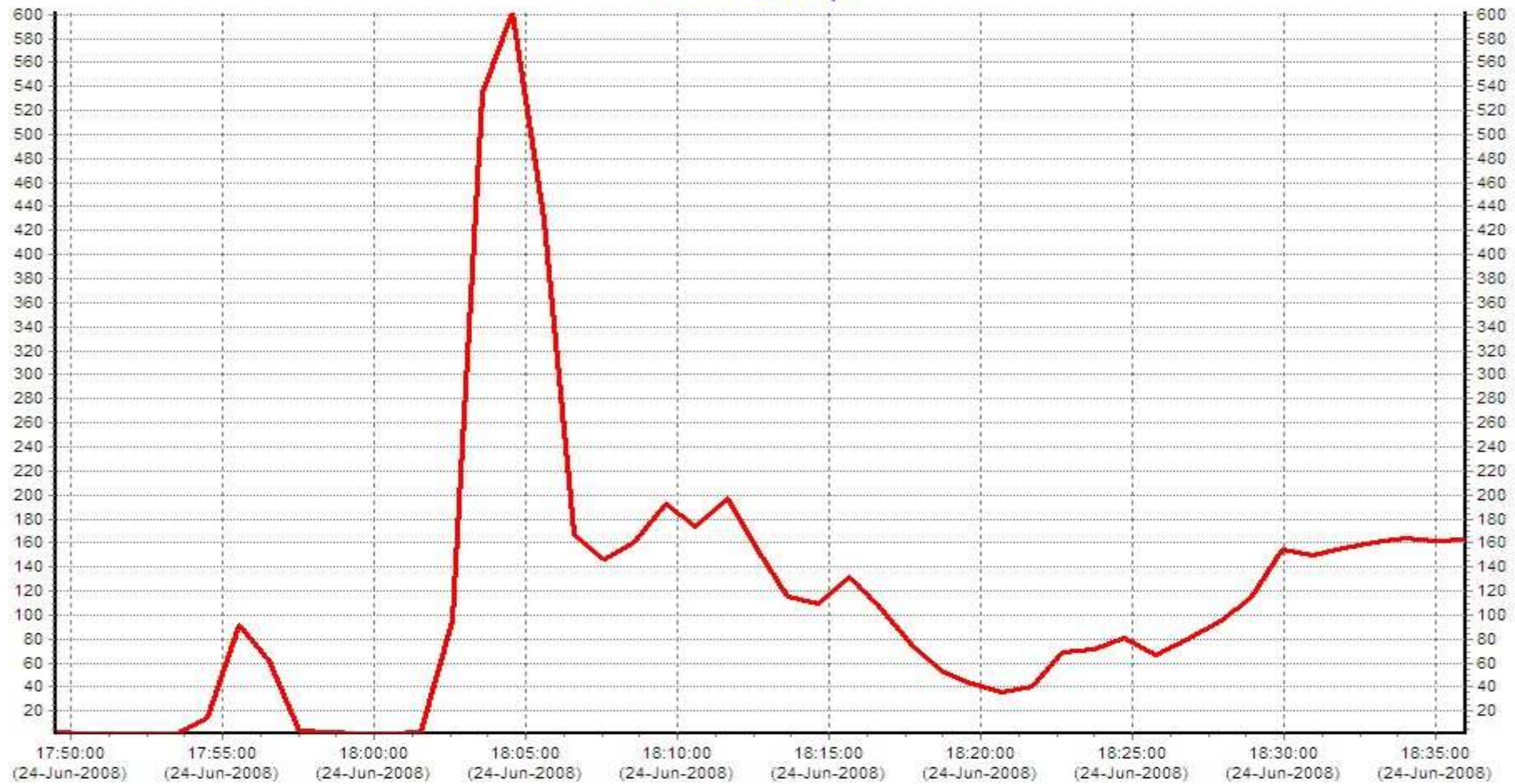
EN

11:33 PM

Latching activity

Oracle Server

Timer wakeups



☒ [MON.TIMR]TQE wakeups(# 1)

Statspack / AWR reports

- Once the latching problem has been resolved performance of the database improved significantly
 - Updates are 10 times ! faster
 - Overall performance improved 5 times



Long Connect Time

- Watch out for long connect time to the database
- Monitor CPU utilization of the BEQ listener
 - Utilization > 60% consider adding more BEQ listeners
- Possible solutions for long connection time:
 - Add BEQ listeners
 - Add listeners
 - Properly size the Flash recovery area



Automatic Memory Management

- Automatic Memory Management of the SGA is a new feature shipping with Oracle 10g
- Oracle attempts to manage available SGA memory to meet the application's requirements
- Maklee recommends disabling automatic memory management
 - Oracle may fail to respond to a sudden change in the workload
 - No good deed goes unpunished....guarantee enough memory to all areas of the SGA
 - To disable automatic memory management:
 - `SGA_MAXSIZE = X`
 - `SGA_TARGET = 0`



RDB Vs. Oracle Terminology

See Article ID: 276447.1:Quick Reference Translation from Oracle Database to Rdb Database

<ul style="list-style-type: none">•Memory Terms<ul style="list-style-type: none">–Global Buffers–Local Buffers	<ul style="list-style-type: none">•Memory Terms<ul style="list-style-type: none">–SGA–PGA
<ul style="list-style-type: none">•Storage<ul style="list-style-type: none">–Rdb root file/default storage area–RUJ / Snapshot files (SNP)–Local Storage area (Table/index)–Physical Storage Area–Storage area extent–Hot Standby	<ul style="list-style-type: none">•Storage<ul style="list-style-type: none">–System Tablespace–Rollback segments–Tablespace–Datafile (.DBF)–Extent–Dataguard
<ul style="list-style-type: none">•Non-data files<ul style="list-style-type: none">–Database root (.RDB)–RDMMONxx.log (monitor log)–Bugcheck dump	<ul style="list-style-type: none">•Non-data files<ul style="list-style-type: none">–Control file, init.ora, pwd file–ALERT log–Trace file



RDB Vs. Oracle Terminology

<ul style="list-style-type: none">• Connectivity<ul style="list-style-type: none">– Direct connect (RCI – relational call interface)– SQL/Services, Dispatcher (OCI/SQLSRV)– SQL/Services Executor	<ul style="list-style-type: none">• Connectivity<ul style="list-style-type: none">– Bequeath Connection– Listener.ora, tnsnames.ora, and sqlnet.ora– Listener dedicated process
<ul style="list-style-type: none">• Processes<ul style="list-style-type: none">– Rdb monitor (RDMMONxx)– AIJ Log Server (ALS)– AIJ Backup Server (ABS)– DBR (Database Recovery)	<ul style="list-style-type: none">• Processes<ul style="list-style-type: none">– PMON, SMON, DBW0, CKPT– LGWR– ARC0– RECO



Adabas

- Maklee acknowledge that there are other database solutions available on OpenVMS.
- Maklee continues to expand it's offering in response to customer needs.



MAKLEE

Questions?

See us at www.maklee.com for:

- Oracle Tuning
- DBA services
- Oracle RAC installations

.....Bis bald



MAKLEE

ORACLE[®] PARTNER