



VISCOSITY  
NORTH AMERICA



# Grid Infrastructure Management Repository & Cluster Health Advisor

# Sean Scott

Working with Oracle technology since 1995  
Development :: DBA :: Reliability Engineering :: DevOps  
Oracle OpenWorld :: Collaborate/IOUG :: Regional UG

RAC/MAA :: Data Guard :: Sharding :: Exadata/ODA  
Diagnostic Tools (AHF, TFA, RDA, CHA, CHM)

DR, HA, Site Reliability/Continuity

Upgrade :: Migration :: Cloud

DevOps :: Infrastructure as Code :: Automation

Containers :: Virtualization





# Grid Infrastructure Management Repository (GIMR)

# GIMR - Stores Diagnostic, Performance Data

- Real time monitoring for clusters & RAC databases
- Provides early detection for system failures
  - Diagnoses, identifies likely causes
  - Recommends corrective actions
  - Generates alerts and notifications
- Little/no administration required
- Automatically monitored & managed by CRS
- Optional in 19c+

# GIMR - Stores Diagnostic, Performance Data

- Early versions used BerkleyDB
- Since 12.1, uses Oracle (multitenant) -MGMTDB
  - CDB runs on one node
  - Automatically relocated on node stop/failure
  - Default storage target is OCR/Voting disk
- Diagnostic data saved in partitions
  - Size of GIMR is related to number of targets & retention
  - Database size remains fixed

# GIMR - Clients

- Cluster Health Advisor (CHA)
  - Real-time performance data
- Cluster Health Monitor (CHM)
  - Metrics, fault, and diagnostic collections
- Oracle Clusterware (GI logging)
  - Events for all Clusterware resources
- Quality of Service Management (QoS)
  - Workload performance data

# GIMR - Clients

- Diagnostic tools
  - Autonomous Health Framework (AHF)
  - Trace File Analyzer (TFA)
  - Enterprise Manager Cloud Control (EMCC)
  - OraCheck, ExaCheck
- Oracle Fleet Patching & Provisioning (Metadata)



# GIMR - New in Oracle Database 21c

- GIMR *must be* deployed to a separate `ORACLE_HOME`
  - During new install or upgrade of Grid Infrastructure
- Centralized remote GIMR support
  - Many clusters, one GIMR
  - Separates data store, targets
- Local mode for Cluster Health Monitor
  - Run `oclumon dumpnodeview` without GIMR
  - Gathers limited OS metrics for individual nodes

# GIMR - FAQ

- Cluster & database availability unaffected if GIMR fails
- GIMR clients cache metrics locally during failures
- Uses ~376 hugepages (when available)
- Patches included in GI RUs
  - No separate patching is required
- No backups required
  - Archive data with `ocLumon` utility

# GIMR - FAQ

- Leading character of SID & PDB name are protected
  - Prevents access by DBCA, DBUA, and similar tools
  - Only MGMTCA and utilities can manage GIMR
- What resources does GIMR use?

	First 5 Targets	Additional Targets
12.1	5.2G	500M each
12.2	36G	4.7G each
19c	28G	5G each



# Cluster Health Advisor (CHA)

# CHA - Oracle Cluster Health Advisor

- Introduced in 12.2
- Monitors the OS on each cluster node
- Optionally monitors RAC database instances
- Integrated with OEM
- Stores its data in GIMR

# CHA - Oracle Cluster Health Advisor

- Monitors nodes automatically *once a RAC DB starts*
  - Reads Cluster Health Monitor data directly from memory
- RAC, RAC One Node monitoring *must be explicitly enabled*
- Reads Database ASH from SMR (no DB connection)
- Data point collection
  - 150+ signals every second per target
  - Data is synchronized, smoothed
  - Results aggregated to 5 second intervals

# CHA - Modeling

- Compares OS, Database activity against models
  - 30+ node & database problem models
  - 150+ OS & database metric predictors
  - Interconnect, Global Cache, Cluster
  - Host CPU & Memory
  - PGA memory stress
  - I/O and storage performance
  - Workload and session variations

# CHA - “Normality Model”

- Models continuously adjusted by target activity
- Normality Model considers load *similarity*, not *absolute thresholds*
  - Time/Day
  - Signal persistence
  - Observed vs predicted
  - Vector interdependency
- Differentiates momentary spikes from “deviant behavior”



# Default vs. Custom Models

- Default models are conservative
  - DEFAULT\_CLUSTER
  - DEFAULT\_DB
  - Minimize noise and false alerts
- Calibrate models to improve diagnostic sensitivity and accuracy
  - Recommended: Minimum six hour “normal” workload
  - Cluster calibration should cover representative DB activity



# GIMR Best Practices

# GIMR Best Practices - *DO NOT*:

Disable or drop GIMR!

- OSS requires Tier One clusters 12c+ to run GIMR

Connect to MGMTDB through SQL\*Plus!

- “Contains no user serviceable parts”
- Only under direction of OSS

Manage passwords manually!

- Credentials automatically generated and managed
- Use `mgmtca` to regenerate, do not set via SQL\*Plus/clients

# GIMR Best Practices - *DO NOT*:

Add MGMTDB or MGMTLSNR as EMCC targets!

- DB and listener automatically monitored by CRS
- EMCC will treat MGMT\* as SI targets

Use `srvctl modify mgmtdb|mgmtlsnr!`

- Use `mgmtca` to set/correct password/connection issues
- Use `mbdutil.pl` script to:
  - Add or recreate MGMTDB
  - Move data files

# GIMR Best Practices - *DO*:

Verify GIMR is running and healthy

- `srvctl status mgmtldb`
- `srvctl status mgmtlsnr`
- `oclumon dumpnodeview -all`

Insure MGMTDB and MGMTLSNR run on the same node

# GIMR Best Practices - *DO*:

Use a dedicated disk group

- External redundancy is adequate
- Use `mdbutil.pl` to change storage location

Maintain at least 72 hour retention for clients

Check retention and set size:

- `oclumon manage -repos checkretentiontime 86400`
- `oclumon manage -repos changereposize <Size MB>`



# Cluster Health Advisor Calibration

# CHA Models and Calibration

- CHA evaluates activity against models
- Default models are conservative
- Models “learn” over time
- Calibration allows:
  - Accelerated learning
  - Multiple model profiles
  - Define KPI
- Only one active/monitored model per target



# Calibrate Models

Create & modify models

- KPI can be combined
  - Set *performance goals* for training
  - They are not thresholds!

Multiple models can exist for a target

Available KPI Names:

- CPUPERCENT
- IOREAD
- IOWRITE
- IOTHROUGHPUT
- DBTIMEPERCALL (DB only)

```
chactl calibrate [-cluster | -db <db_unqname>]
                [-model <model name>] [-force]
                [-timeranges 'start=<time>,end=<time>']
                [-kpiset 'name=<kpi> min=<minval> max=<maxval>, ...']
```

# Calibration Tips

Targets can have multiple models

- Daytime, nighttime, month-end
- Each model requires GIMR space
- May need to increase size of repository, number of targets

“No sufficient calibration data exists...” error

- Increase or change the time period
- Change KPI (if specified used)
- Allow CHA to collect more data

# Query Calibration Models

- Larger intervals: Faster, less detailed
- KPI sets: Identical to `chactl calibrate`
  - Do not have to match the model
  - Use to filter results
  - May be combined

```
chactl query calibration [-cluster | -db <db_unqname>]  
    [-interval <hours>] [-timeranges 'start=<time>,end=<time>']  
    [-kpiset 'name=<kpi> min=<minval> max=<maxval>, ...']
```

# Calibration Query Tips

## Specify a time range

- no time range = all target data
- YYYY-MM-DD HH24:MI:SS

## Larger intervals typically run faster

## Queries may take 30-60 minutes

- Run with nohup

## Output is lengthy

- Redirect output to a file

```
$ chactl query calibration -cluster \  
-timeranges 'start=2020-08-21 00:00:00,end=2020-08-21 12:00:00' \  
-interval 6
```

```
Cluster name : prod01db01  
Data Start time : 2020-08-21 00:00:00  
Data End time : 2020-08-21 06:00:00  
Total Samples : 4321  
Percentage of filtered data : 0.0%
```

### 1) CPU utilization (total) (%)

MEAN	MEDIAN	STDDEV	MIN	MAX	
27.70	24.60	11.41	8.80	72.10	
<14.40	<23.90	<33.40	<42.90	<52.40	>=52.40
5.00%	41.10%	29.92%	11.39%	7.57%	5.02%

```
Cluster name : npx01dbc01  
Data Start time : 2020-08-21 06:00:00  
Data End time : 2020-08-21 12:00:00  
Total Samples : 4321  
Percentage of filtered data : 0.0%
```

### 1) CPU utilization (total) (%)

MEAN	MEDIAN	STDDEV	MIN	MAX	
26.20	23.60	11.67	8.20	75.00	
<13.00	<22.73	<32.45	<42.18	<51.90	>=51.90
4.77%	42.03%	30.50%	11.06%	6.60%	5.05%



# Query Diagnostic Information

```
chactl query diagnosis [-cluster | -db <db_unqname>]
                        -start <time> -end <time>
                        [-htmlfile <filename>]
```

```
chactl query diagnosis -cluster
                        -start "2020-01-01 00:00:00" -end "2020-08-21 12:00:00"
                        -htmlfile ~/cha_cluster.html
```

```
chactl query diagnosis -db ORCL
                        -start "2020-01-01 00:00:00" -end "2020-08-21 12:00:00"
                        -htmlfile ~/cha_db_ORCL.html
```

Timestamp	Target Information		Event Name	Detected/Cleared
2021-05-10 09:53:50.0	Host	01	<a href="#">Host CPU Utilization</a>	detected
2021-05-10 10:02:10.0	Host	01	<a href="#">Host CPU Utilization</a>	cleared
2021-05-10 14:59:00.0	Host	01	<a href="#">Host CPU Utilization</a>	detected
2021-05-10 15:00:55.0	Host	01	<a href="#">Host CPU Utilization</a>	cleared
2021-05-10 16:25:35.0	Host	01	<a href="#">Host CPU Utilization</a>	detected
2021-05-10 16:45:35.0	Host	01	<a href="#">Host CPU Utilization</a>	cleared
2021-05-10 16:47:05.0	Host	01	<a href="#">Host CPU Utilization</a>	detected
2021-05-10 16:50:40.0	Host	01	<a href="#">Host CPU Utilization</a>	cleared
2021-05-10 17:50:00.0	Host	01	<a href="#">Host CPU Utilization</a>	detected
2021-05-10 17:52:20.0	Host	01	<a href="#">Host CPU Utilization</a>	cleared
2021-05-10 17:54:05.0	Host	01	<a href="#">Host CPU Utilization</a>	detected
2021-05-10 17:57:25.0	Host	01	<a href="#">Host CPU Utilization</a>	cleared
2021-05-10 18:47:50.0	Host	01	<a href="#">Host CPU Utilization</a>	detected
2021-05-10 19:48:15.0	Host	01	<a href="#">Host CPU Utilization</a>	cleared
2021-05-10 19:58:35.0	Host	01	<a href="#">Host CPU Utilization</a>	detected
2021-05-10 20:09:15.0	Host	01	<a href="#">Host CPU Utilization</a>	cleared
2021-05-10 20:47:20.0	Host	01	<a href="#">Host CPU Utilization</a>	detected
2021-05-10 20:49:20.0	Host	01	<a href="#">Host CPU Utilization</a>	cleared
2021-05-10 22:01:35.0	Host	04	<a href="#">Host CPU Utilization</a>	detected
2021-05-10 22:08:25.0	Host	04	<a href="#">Host CPU Utilization</a>	cleared

Problem	Description	Cause	Action
Host CPU Utilization	CHA detected larger than expected CPU utilization on this node. The available CPU resource may not be sufficient to support application failover or relocation of databases to this node.	The Cluster Health Advisor (CHA) detected an unexpected increase in CPU utilization by databases or applications on this node.	Identify CPU intensive processes and databases by reviewing Cluster Health Monitoring (CHM) data. Relocate databases to less busy machines, or limit the number of connections to databases on this node. Add nodes if more resources are required.

Timestamp	Target Information	Event Name	Detected/Cleared
2021-05-10 08:52:25.0	Database (PDB2)	<a href="#">DB Recovery Area Space</a>	detected
2021-05-10 12:19:30.0	Database (PDB1)	<a href="#">DB Control File I/O Performance</a>	detected
2021-05-10 12:24:25.0	Database (PDB1)	<a href="#">DB Control File I/O Performance</a>	cleared
2021-05-10 13:30:15.0	Database (PDB3)	<a href="#">DB Control File I/O Performance</a>	detected
2021-05-10 13:33:15.0	Database (PDB3)	<a href="#">DB Control File I/O Performance</a>	cleared
2021-05-10 14:31:25.0	Database (PDB2)	<a href="#">Control File Enqueue Wait Hang</a>	detected
2021-05-10 14:34:30.0	Database (PDB2)	<a href="#">Control File Enqueue Wait Hang</a>	cleared
2021-05-10 15:24:35.0	Database (PDB1)	<a href="#">Private Network Traffic</a>	detected
2021-05-10 15:25:00.0	Database (PDB1)	<a href="#">Private Network Traffic</a>	cleared
2021-05-10 15:47:30.0	Database (PDB3)	<a href="#">Private Network Traffic</a>	detected
2021-05-10 15:54:10.0	Database (PDB3)	<a href="#">Private Network Traffic</a>	cleared
2021-05-11 13:47:15.0	Database (PDB2)	Ended monitoring	
2021-05-11 13:47:20.0	Database (PDB3)	Ended monitoring	
2021-05-11 13:47:25.0	Database (PDB1)	Ended monitoring	

Problem	Description	Cause	Action
DB Recovery Area Space	The Cluster Health Advisor (CHA) detected that database sessions are waiting for a log switch to complete.	Log archiving was unable to continue because there was no more space available in the recovery area. A cluster-wide hang is likely if this problem is not resolved.	Add disk space and increase the db_recovery_file_dest_size parameter to reflect the new space. Delete unnecessary files using the RMAN delete command. Check all alert logs of this database for incidents of ORA-19815 for further recommendation.
DB Control File I/O Performance	CHA has detected that reads or writes to the control files are slower than expected.	The Cluster Health Advisor (CHA) detected that reads or writes to the control files were slow because of an increase in disk I/O. The slow control file reads and writes may have an impact on checkpoint and Log Writer (LGWR) performance.	Separate the control files from other database files and move them to faster disks or Solid State Devices.
Control File Enqueue Wait Hang	The Cluster Health Advisor (CHA) detected a hang. Some processes are waiting for the Control File enqueue for several seconds.	The Cluster Health Advisor (CHA) detected that a background process has been waiting several seconds for a Control File enqueue because the Log Writer, Archiver or Checkpoint process on another node is unresponsive, blocked by other processes, or is waiting for CPU or memory.	Check whether incidents are detected and diagnosed on other nodes and instances in the cluster using the command 'chactl query diagnosis'. Look for reports of any issues detected by the Hang Manager in the DIA0 trace file and follow the suggestions to resolve those problems.
Private Network Traffic	The Cluster Health Advisor (CHA) detected an abrupt, significant decrease in message traffic on the Cluster Interconnect.	The decrease in traffic on the private interconnect was probably caused by a problem with the underlying network.	Check for network link or connectivity failures in the system message logs.



# Management Database Utility (MDBUtil)



# MDBUtil - MGMTDB Utility (2065175.1)

- `mdbutil.pl`
  - Checks MGMTDB and listener status
  - Creates, recreates Management Databases
  - Migrates disk groups

# GIMR - MGMTDB Utility

```
# mdbutil.pl --status
```

```
MGMTDB is not configured
```

```
MGMTLSNR is not configured
```

```
# mdbutil.pl --addmdb --target=+DATA
```

```
mdbutil.pl version : 1.99
```

```
Starting To Configure MGMTDB at +DATA...
```

```
Container database creation in progress...
```

```
Plugable database creation in progress...
```

```
Executing "/tmp/mdbutil.pl --addchm" to configure CHM.
```

```
MGMTDB & CHM configuration done!
```

# GIMR - MGMTDB Utility

```
# mdbutil.pl --mvmgmtdb --target=+DATA
```

```
mdbutil.pl version : 1.99
```

```
Moving MGMTDB, it will be stopped, are you sure (Y/N)? y
```

```
Checking for the required paths under +DATA
```

```
...
```

```
Stopping mgmtdb
```

```
Copying MGMTDB DBFiles to +DATA
```

```
Creating the CTRL File
```

```
The CTRL File has been created and MGMTDB is now running from +DATA
```

```
Modifying the init parameter
```

```
Removing old MGMTDB
```

```
Restarting MGMTDB using target SPFile
```

```
MGMTDB Successfully moved to +DATA!
```



## GIMR Tips

# Identify & Remove Berkley Artifacts

- < 12.1 used BerkleyDB for its repository
- Files could grow > 100G
- Remove old/obsolete files:
  - `rm $GRID_HOME/crf/dbf/$(hostname)/*.bdb`
  - Could be on any node

# Reading Logs and Traces

- `$GRID_HOME/diag/rdbms/_mgmtdb/-MGMTDB/trace`
- Trace files prefixed with `-MGMTDB`
- \*nix tries to interpret `-` as a command flag/option
- Use `./` to manage files

```
# less -MGMTDB_mmon_1277.trc
Unknown option argument "-MGMTDB_mmon_1277.trc"

# less ./-MGMTDB_mmon_1277.trc
# rm ./-MGMTDB_mmon_1277.trc
etc.
```

# ORA-28000 from oclumon dumpnodeview

Usually caused by:

- Failed GI install post-steps
- Incomplete drop/add MGMTDB

Run (or re-run) `mgmtca` to update wallets in OCR

```
Querying for the local host  
CRS-9118-Grid Infrastructure  
Management Repository connection error  
ORA-28000: the account is locked
```

```
# 12.2+, set/reset GIMR wallets:  
mgmtca [-allusers |  
        -user [ CALOG, CHA, CHMOS  
              GRIDHOME, QOS ]]
```

# Connect to MGMTDB (Don't do this!)

You may use OS authentication to connect to MGMTDB but *Oracle advises against this! There is no reason to access MGMTDB under normal conditions!*

```
export ORACLE_SID=\-MGMTDB  
sqlplus / as sysdba
```





# CHA and GIMR Command Glossary

# Management and Configuration Commands

```
# Start CHA
srvctl start cha [-node <node>]

# Stop CHA
srvctl stop cha [-node <node>]
                [-force]

# Show status and configuration
srvctl status cha
srvctl config cha
chactl status [-verbose]

# Show GIMR DB status
srvctl status mgmtdb [-verbose]
```

```
# Add, remove database monitoring
chactl monitor database
        -db <db_unqname>
        [-model <model name>]
chactl unmonitor database
        -db <db_unqname>

# Gather query repository
chactl query repository

# Change KEEP retention, repo size
chactl set maxretention
        -time <hours_to_keep>
chactl resize repository
        -entities <total_targets>
```

# Configure, Monitor, and Manage GIMR Resources

```
# Start, stop MGMTDB:
srvctl start mgmtdb
srvctl stop mgmtdb

# Start, stop MGMTDB Listener
srvctl start mgmtlsnr
srvctl stop mgmtlsnr

# Get DB & Listener status
srvctl status mgmtdb
srvctl status mgmtlsnr

# Get DB & Listener configuration
srvctl config mgmtdb
srvctl config mgmtlsnr
```

```
# Identify repository path
oclumon manage -get reppath
srvctl status mgmtdb
```

```
# Locate GIMR master
oclumon manage -get MASTER
srvctl status mgmtdb
```

```
# Do not modify MGMT via srvctl!
NO: srvctl modify mgmtdb
NO: srvctl modify mgmtlsnr
# Use only when directed by MOS!
```

# Get Diagnostics - oclumon dumpnodeview

## Information types

- **cpu**  
Per-CPU statistics
- **device**  
R/W rate, queue length, wait/IO
- **filesystem**  
Total, used, available space
- **nic**  
Bandwidth, send/receive & error rates

```
oclumon dumpnodeview [-v]
# Control nodes
[-allnodes |-node <node list>]
# Limit time
[-last "<duration>" |
-s "YYYY-MM-DD HH24:MI:SS"
-e "YYYY-MM-DD HH24:MI:SS"]
[-i <interval>]
# Information types:
[-system] [-process] [-cpu]
[-device] [-filesystem] [-nic]
[-protoerr] [-topconsumer]
# Formatting and output
[-format legacy|tabular|csv]
[-dir <directory> [-append]]
# Aggregate by category
[-procag]
```

# Get Diagnostics - oclumon dumpnodeview

## Information types

- **process**  
PID, name, threads, memory use
- **protoerr**  
Protocol errors
- **system**  
CPU & memory statistics
- **topconsumer**  
Top process utilization

```
oclumon dumpnodeview [-v]
# Control nodes
[-allnodes |-node <node list>]
# Limit time
[-last "<duration>" |
-s "YYYY-MM-DD HH24:MI:SS"
-e "YYYY-MM-DD HH24:MI:SS"]
[-i <interval>]
# Information types:
[-system] [-process] [-cpu]
[-device] [-filesystem] [-nic]
[-protoerr] [-topconsumer]
# Formatting and output
[-format legacy|tabular|csv]
[-dir <directory> [-append]]
# Aggregate by category
[-procag]
```



[oraclesean.com](http://oraclesean.com)



<https://www.linkedin.com/in/soscott/>



[@oraclesean](https://twitter.com/oraclesean)



<https://github.com/oraclesean>



[sean.scott@viscosityna.com](mailto:sean.scott@viscosityna.com)



Search "OracleSean" on YouTube



