BNL Oracle database services status

Carlos Fernando Gamboa

Storage management group
RHIC/ATLAS computer Facility
Brookhaven National Laboratory, US

Replication Technology Evolution for ATLAS Data Workshop
CERN, Geneva. June 2014

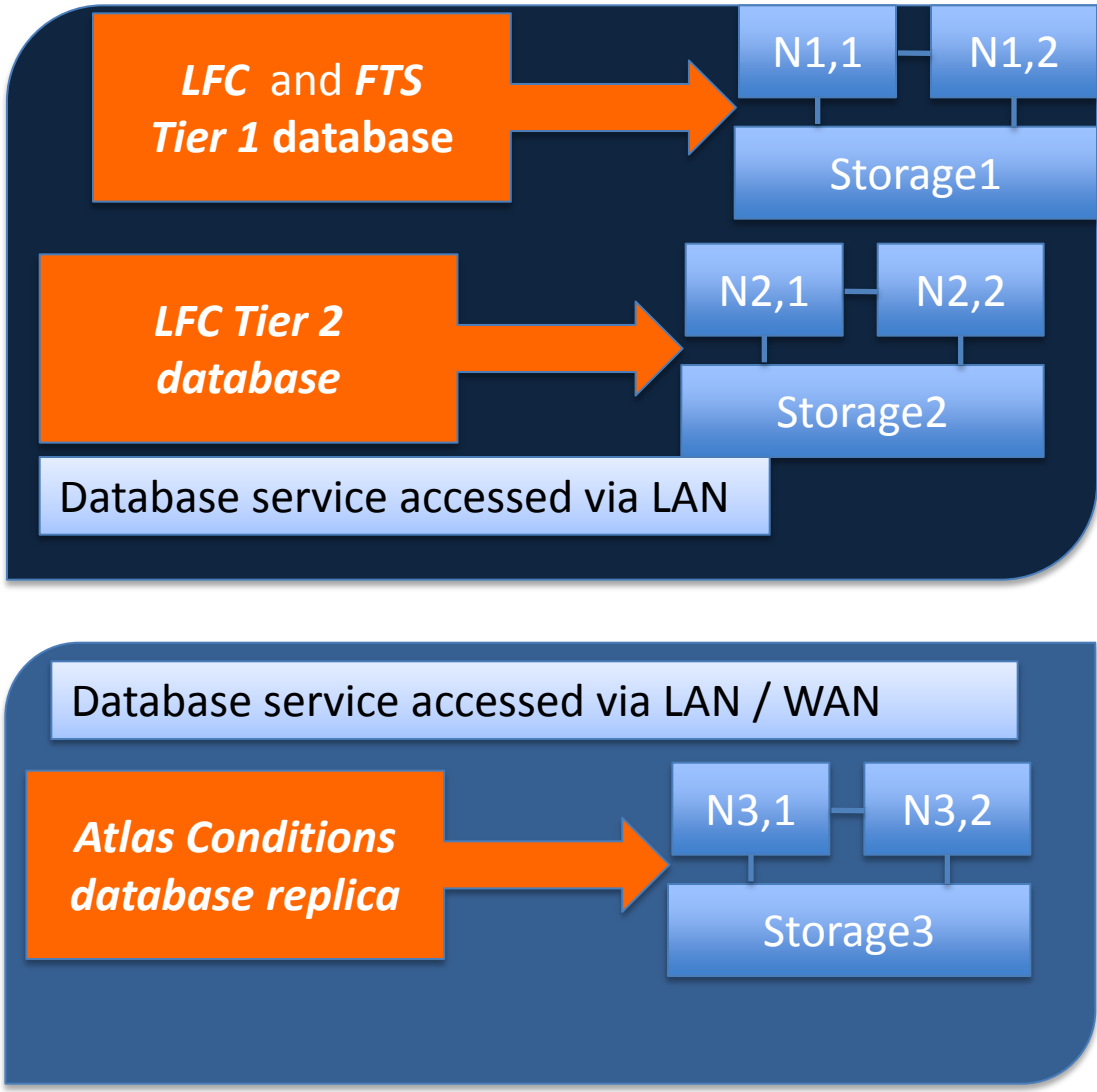
General Topology Oracle Database services hosted at BNL

Independent clusters set per application service. 3 Production primary clusters and 2 Standby Cluster (not shown)

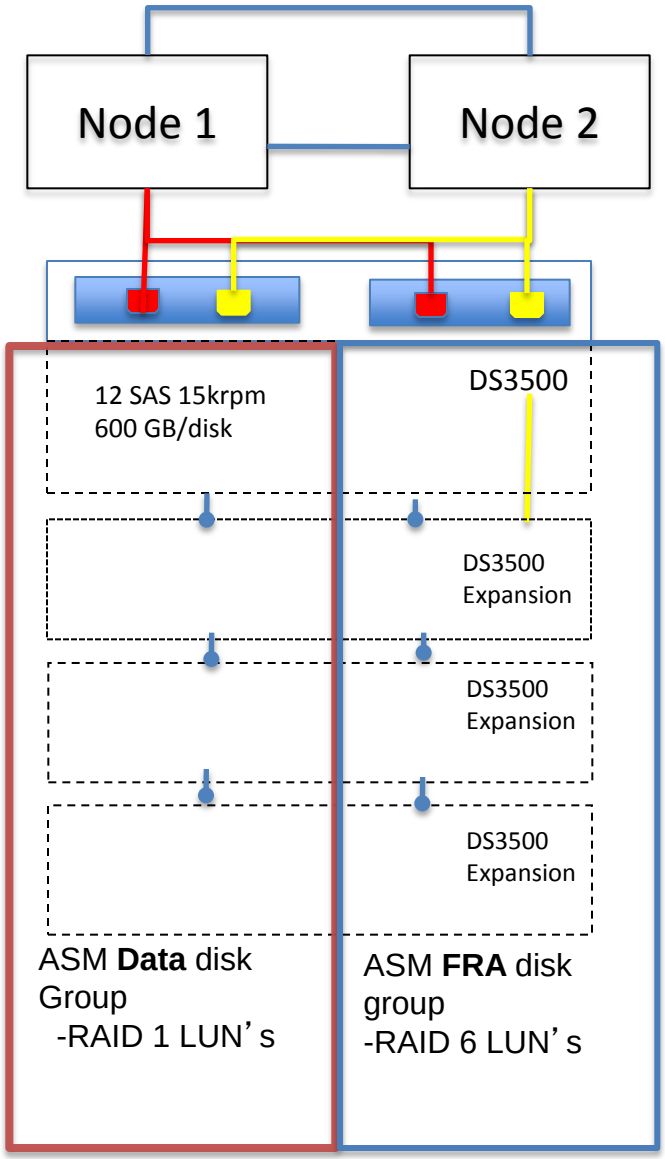
- Dual nodes, Direct Attach Storage (DAS)
- Storage distribution adjusted to application needs
  - Hardware RAID levels
  - Storage and spindles
- Flexible architecture that allows to increase nodes and storage per application needs.

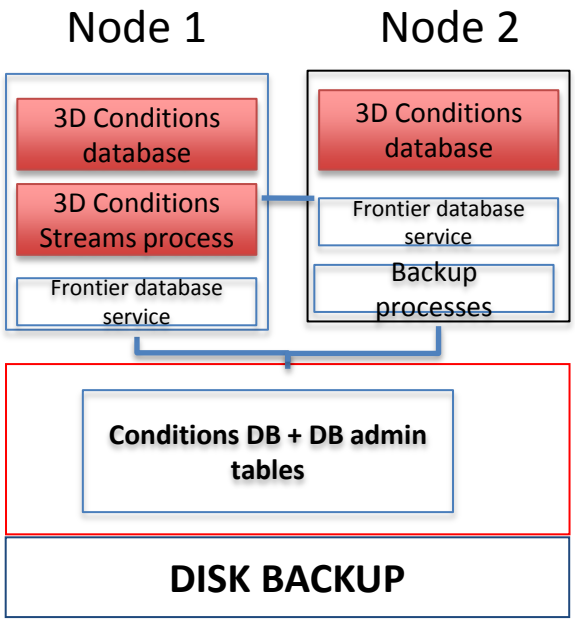
- Homogenous software stack deployed:

  Real Application Cluster (RAC) 11gR2.
  - Database server
  - Clusterware
  - ASM file system



Distribution of database services per production cluster
Example for ATLAS Conditions Database





In general, hardware distributed as:

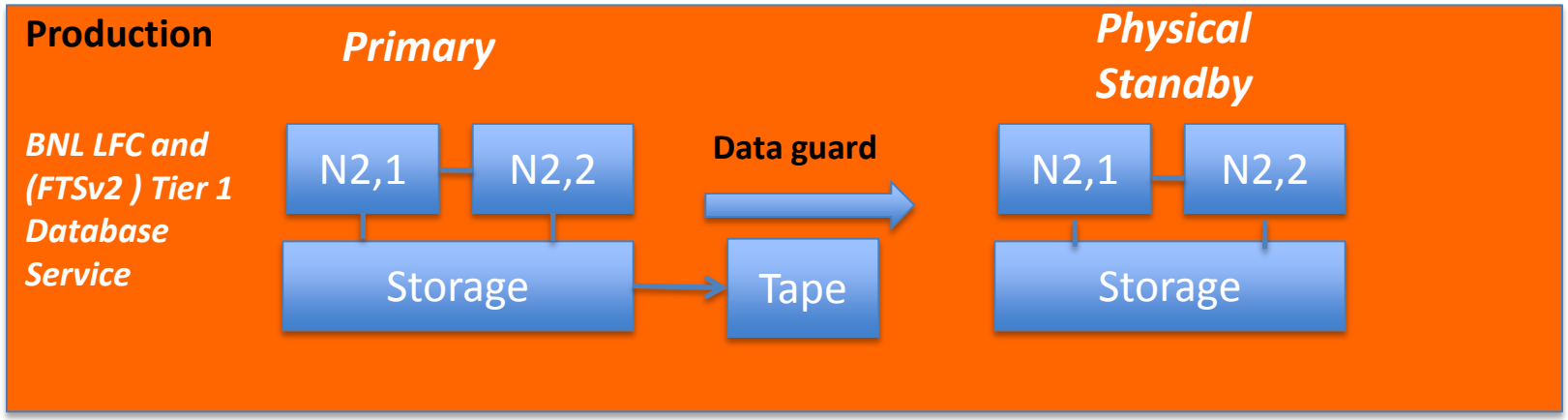
3 Primary clusters uses

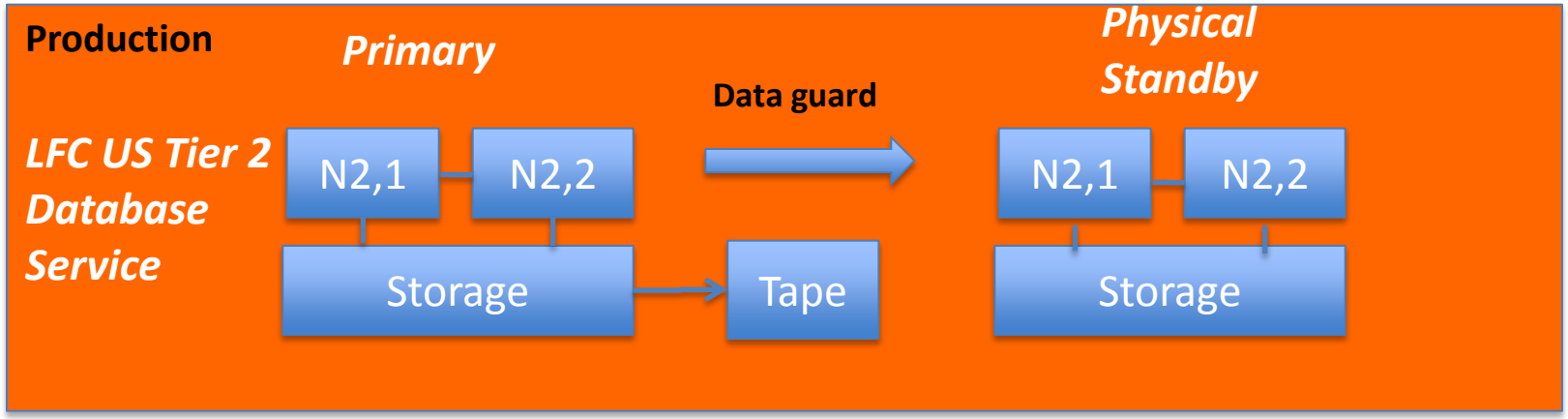
2 IBM x3650 Nodes LFC Tier 2, 2 IBM x3650 M3 nodes
LFC Tier 1 and 2 IBM x3650 M3 nodes for Conditions DB.
NIC 1Gb/s
Storage IBM 3500 DAS, 8Gb/s FC, SAS 15Krmp

2 Standby Clusters

2 IBM x3650 M3 (for LFC and FTS BNL Tier 1) and 2 IBM 3650 (LFC Tier2)
NIC 1Gb/s
Storage IBM 3400 DAS, 4Gb/s FC, SAS 15Krmp

Distribution of database services per production cluster
USATLAS LFC Database service





BNL LFC service in transition to be decommissioned.

Hardware dedicated for this service in extended warranty

Monitoring

Oracle Cloud Control 12C

Oracle database service

PGBadger PostgreSQL Log Analyzer

Testing Activities
GoldenGate

Hardware

2 Node Cluster IBM 3550 DAS topology

4 Cores 3.00GHz, 16GB RAM

NIC 1Gb/s Broadcome

SAS 10k rpm, 300GB,16 disks dedicated to DATA storage

an ACFS volume

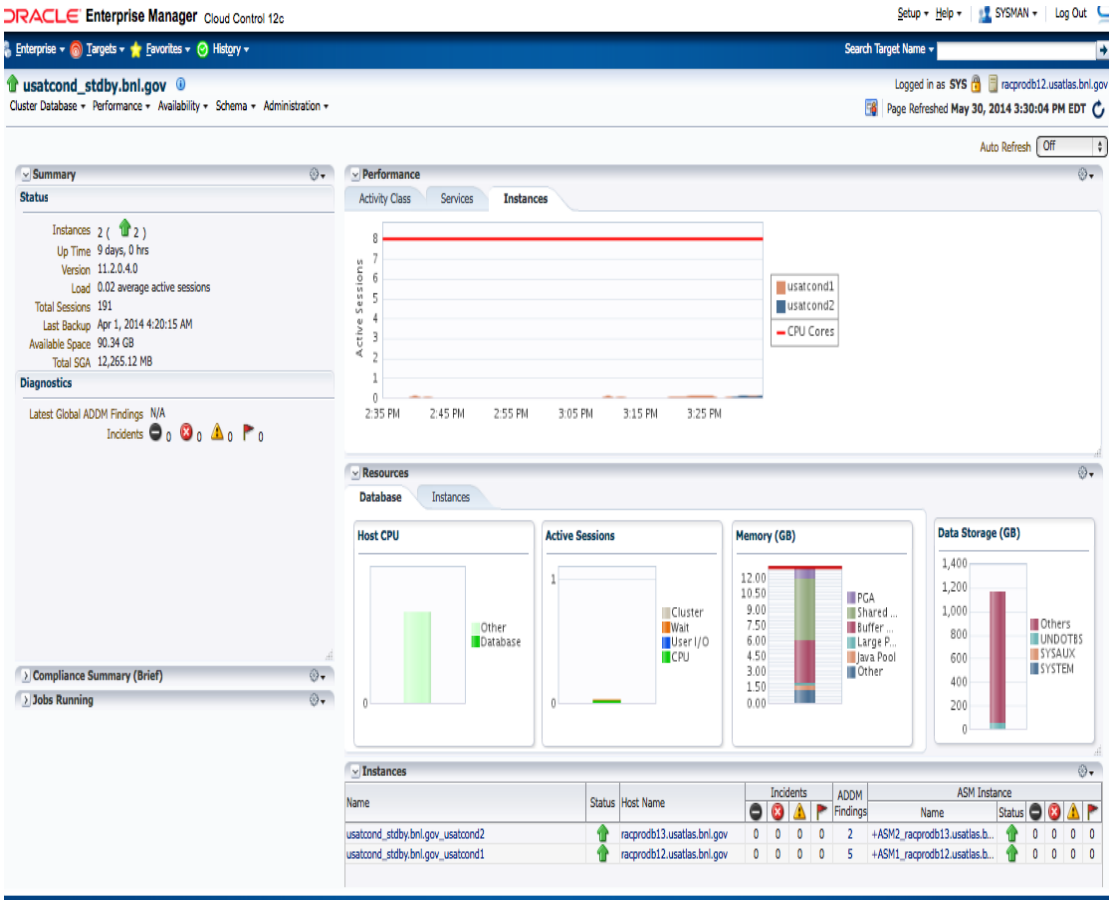
Software

OS RHEL 6 (2.6.32-431.5.1.el6.x86_64)

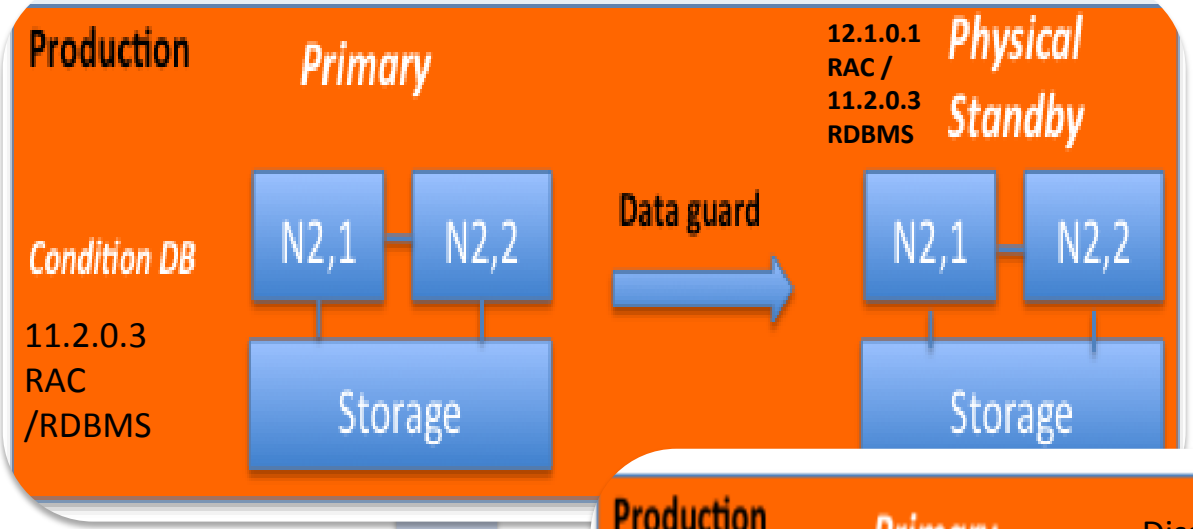
GoldenGate 12.1.0.2

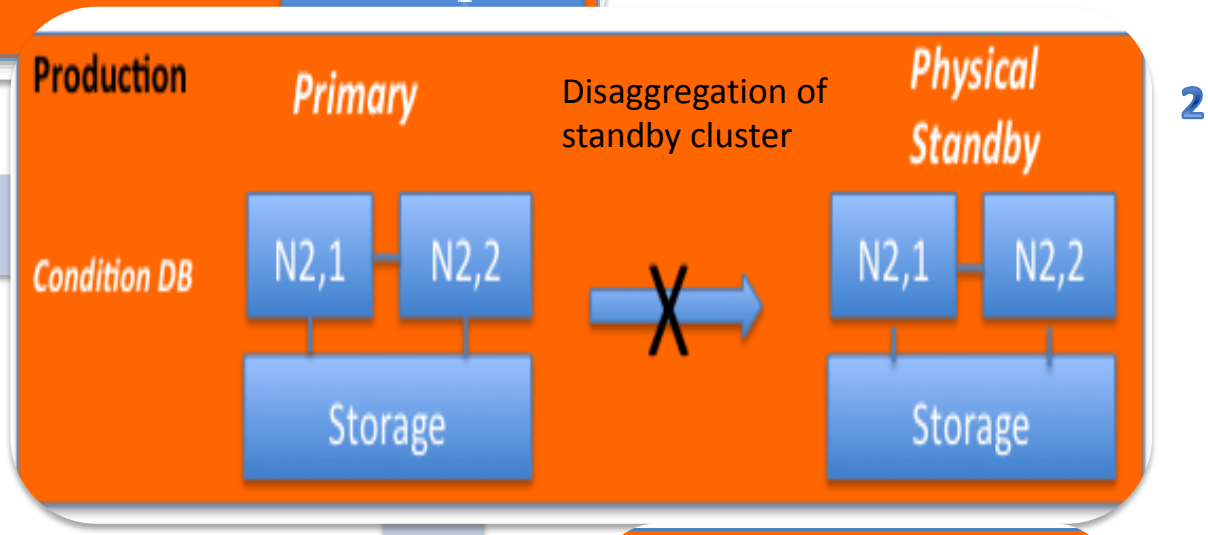
GRID INFRASTRUCTURE 12.1.0.1

RDBMS 11.2.0.4

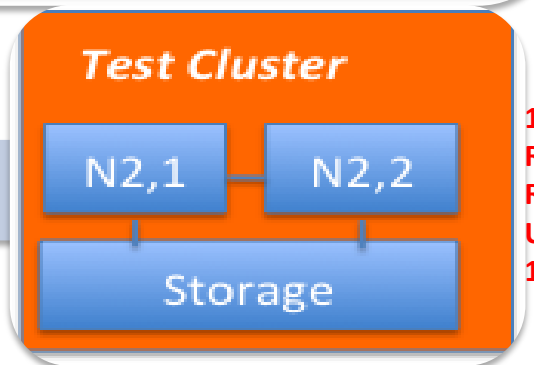


Testing Activities / GoldenGate





Using Dataguard, a snapshot from production US ATLAS Conditions database was temporarily enabled using the test cluster.



12.1.0.1
RAC /
RDBMS
Upgraded to
11.2.0.4

Testing Activities / GoldenGate

GoldenGate deployment (destination test point)

Data binaries installed in an ACFS volume shared across the two nodes, ASM disk group was used to create the ACFS volume:

```
[usatcond1.oracle@racprodb12 gg]$ df -lh |grep GG
/dev/asm/goldengate-438   100G  1.1G   99G   2% /GG_DATA
[usatcond1.oracle@racprodb12 gg]$ echo $GG_HOME
/GG_DATA/oracle/product/12.1.2/gg
[usatcond1.oracle@racprodb12 gg]$
```

Cluster resource was created to enable the GG application in a High Availability(HA) environment.

- Integration of a new VIP address with the cluster
- Integration of a cluster resource using a new script to handle states of the GG process (start, stop, check and clear)
- Adjusting GG internal parameters for HA

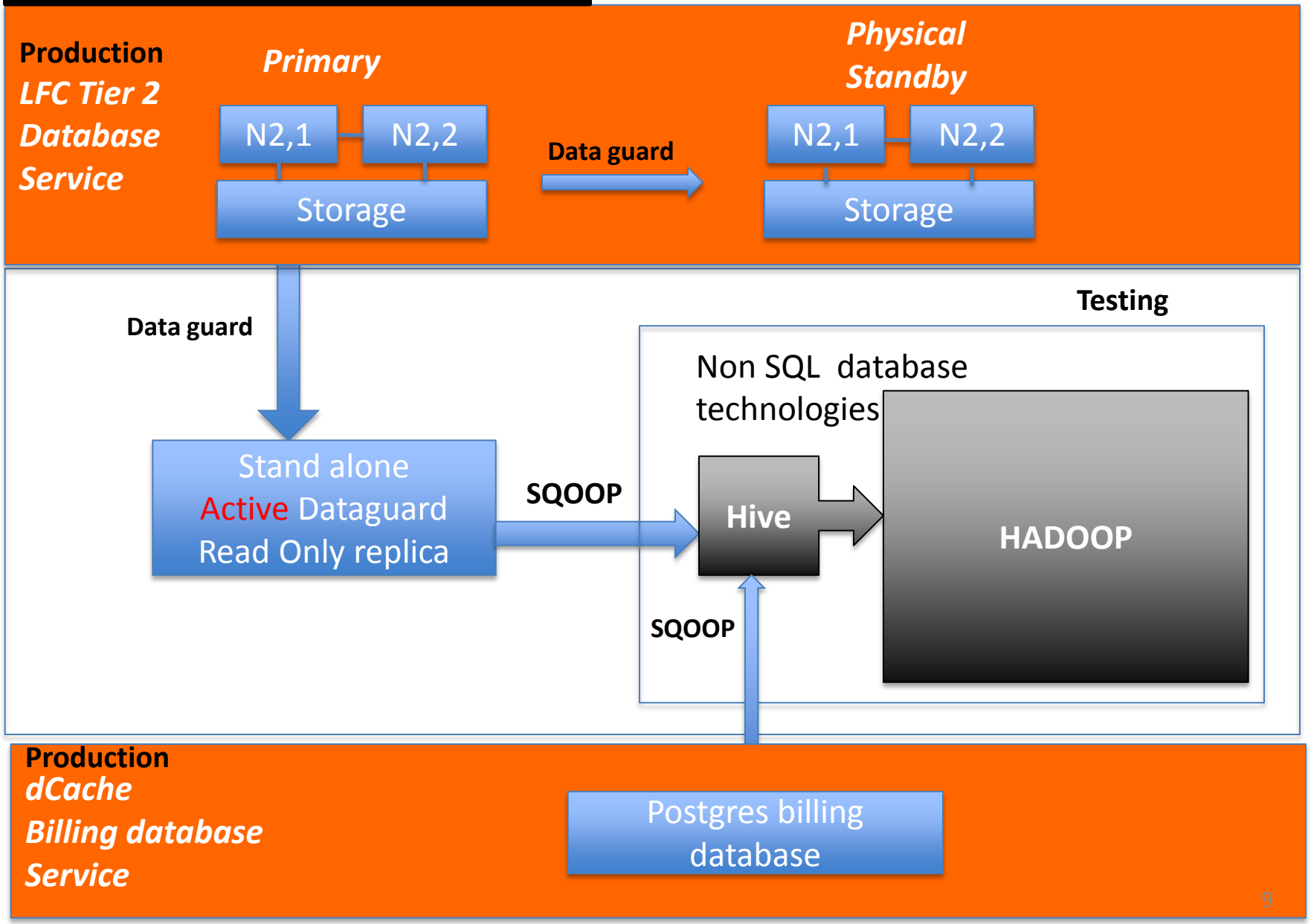
```
[usatcond1.oracle@racprodb12 gg]$ crsctl status resource vracggtestdb
NAME=vracggtestdb
TYPE=app.appvip_net1.type
TARGET=ONLINE
STATE=ONLINE on racprodb13

[usatcond1.oracle@racprodb12 gg]$ crsctl status resource ggtestapp
NAME=ggtestapp
TYPE=cluster_resource
TARGET=ONLINE
STATE=ONLINE on racprodb13
```

```
GGSCI (racprodb12.usatlas.bnl.gov) 13> stats replicat BNLTEST

Sending STATS request to REPLICAT BNLTEST ...

Start of Statistics at 2014-05-31 07:59:31.


Integrated Replicat Statistics:

        Total transactions                          19419.00
        Redirected                                      0.00
        DDL operations                                  0.00
        Stored procedures                               0.00
        Datatype functionality                          0.00
        Event actions                                   0.00
        Direct transactions ratio                       0.00%

DDL replication statistics:

*** Total statistics since replicat started     ***
        Operations                                      0.00
        Mapped operations                               0.00
        Unmapped operations                             0.00
        Other operations                                0.00
        Excluded operations                             0.00
        Errors                                          0.00
        Retried errors                                  0.00
        Discarded errors                                0.00
        Ignored errors                                  0.00

Replicating from GGUSER2.REP_TAB to GGUSER2.REP_TAB:

*** Total statistics since 2014-05-28 14:41:21 ***
        Total inserts                            6500000.00
        Total updates                            6400000.00
        Total deletes                            6500000.00
        Total discards                                 0.00
        Total operations                        19400000.00
```

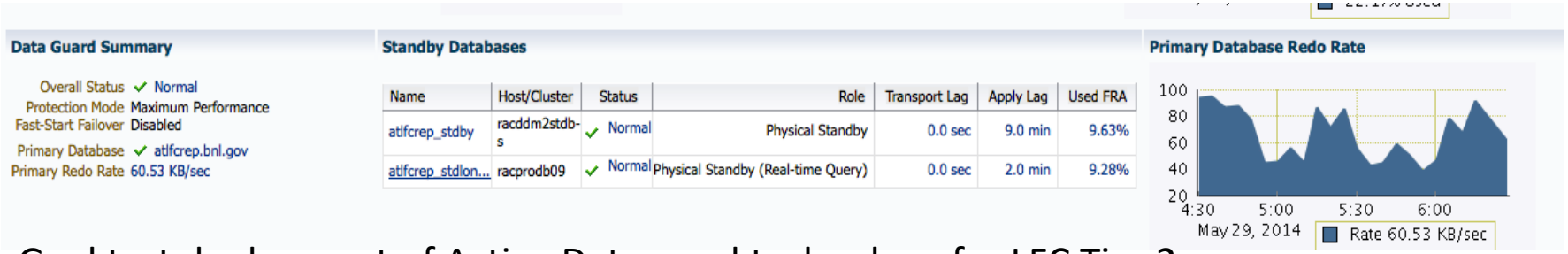
Details of this installation and functional tests will be reported tomorrow during the GG Admin Section

Testing Activities

Cross-database technology testing



Testing Activities
Active Dataguard



Goal test deployment of Active Dataguard technology for LFC Tier 2

- Single instance database replica of LFC Tier 2 database
- 24 core (2.47 Ghz), 47GB RAM Legacy hardware configured with commodity Solid State Disk data storage using ASM normal redundancy.

```
[+ASM.oracle@racprodb09 ~]$ asmcmd
ASMCMD> lsdg
State    Type    Rebal  Sector  Block       AU  Total_MB  Free_MB  Req_mir_free_MB  Usable_file_MB  Offline_disks  Voting_files  Name
MOUNTED  NORMAL  N         512   4096  1048576   1367562   582417           227927          177245              0             N  DATA/
```

Benefits

offloading primary cluster

Opportunistic use of legacy hardware for querying, reporting and testing

So far

No issues observed with the data storage configuration.

Testing Activities

SQOOP

SQOOP jobs enabled to load data into HIVE

```
[hadoop@hdptest03 hive_jobs]$ sqoop-job --list
Available jobs:
  hive_prodbillinginfo
  hive_proddoorinfo
  hive_prodstorageinfo
  hivelfct2_cns_file_metadata
  hivelfct2_cns_file_replica
  lfctier2_LFC_FILE_METADATA
  lfctier2_LFC_FILE_REPLICA
```

HIVE

HIVE installed shared same node with Job tracker

```
[hadoop@hdptest03 ~]$ hive
Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4j.propertie
Hive history file=/tmp/hadoop/hive_job_log_hadoop_201405301121_1845689179.txt
hive> show databases;
OK
billing
default
lfc_tier1
lfc_tier2
lfc_us
```

dCache Billing Postgres database data

```
hive> use billing;
OK
Time taken: 0.0090 seconds
hive> show tables;
OK
billinginfo
doorinfo
storageinfo
Time taken: 0.041 seconds
hive>
```

Data being updated from dCache Billing Postgres database using SQOOP

HADOOP

4 Servers, 8 core (3GHz),16GB RAM
1 name done, 1 job tracker, 2 Data Nodes. OS RHEL6

NameNode 'hdptest01.usatlas.bnl.gov:8020' (active)

Started:	Fri May 30 09:28:18 EDT 2014
Version:	2.0.0-cdh4.1.2, f0b53c81cbf56f5955e403b49fcd27afd5f082de
Compiled:	Thu Nov 1 17:33:23 PDT 2012 by jenkins from Unknown
Upgrades:	There are no upgrades in progress.
Cluster ID:	CID-f9b7d0da-0848-44f4-9085-4addf5c78bb3
Block Pool ID:	BP-1456216560-130.199.48.165-1363117935194

Browse the filesystem
NameNode Logs

Cluster Summary

Security is *OFF*
3994 files and directories, 10680 blocks = 14674 total.
Heap Memory used 80.03 MB is 14% of Commited Heap Memory 561.44 MB. Max Heap Memory is 888.94 MB.
Non Heap Memory used 33.69 MB is 98% of Commited Non Heap Memory 34.06 MB. Max Non Heap Memory is 130 MB.

Configured Capacity	:	4.79 TB			
DFS Used	:	1.52 TB			
Non DFS Used	:	249.6 GB			
DFS Remaining	:	3.03 TB			
DFS Used%	:	31.67 %			
DFS Remaining%	:	63.24 %			
Block Pool Used	:	1.52 TB			
Block Pool Used%	:	31.67 %			
DataNodes usages	:	Min %	Median %	Max %	stdev %
		31.67 %	31.67 %	31.67 %	0 %
Live Nodes	:	3 (Decommissioned: 0)			
Dead Nodes	:	0 (Decommissioned: 0)			
Decommissioning Nodes	:	0			
Number of Under-Replicated Blocks	:	459			

HDFS:/user/hive/warehouse - Mozilla Firefox
File Edit View History Bookmarks Tools Help
HDFS:/user/hive/warehouse
hdptest02.usatlas.bnl.gov:50075/browseDirectory.jsp?dir=

Contents of directory /user/hive/warehouse

Goto : /user/hive/warehouse go

Go to parent directory

Name	Type	Size	Replication	Block Size	Modification Time	Permission	Owner	Group
billing.db	dir				2014-03-11 13:51	rwxr-xr-x	hadoop	hadoop
lfc_tier1.db	dir				2014-03-07 14:38	rwxr-xr-x	hadoop	hadoop
lfc_tier2.db	dir				2014-03-07 14:28	rwxr-xr-x	hadoop	hadoop
lfc_us.db	dir				2014-03-07 13:31	rwxr-xr-x	hadoop	hadoop

Go back to DFS home

Local logs

Log directory

Hadoop, 2013.

Future plans

Hardware migration for Conditions Database (TBD)

- Migrate OS to RHEL 6
- Upgrade to 11.2.0.4
- Migrate to GoldenGate per experiment agreement

Questions?

cgamboa@bnl.gov