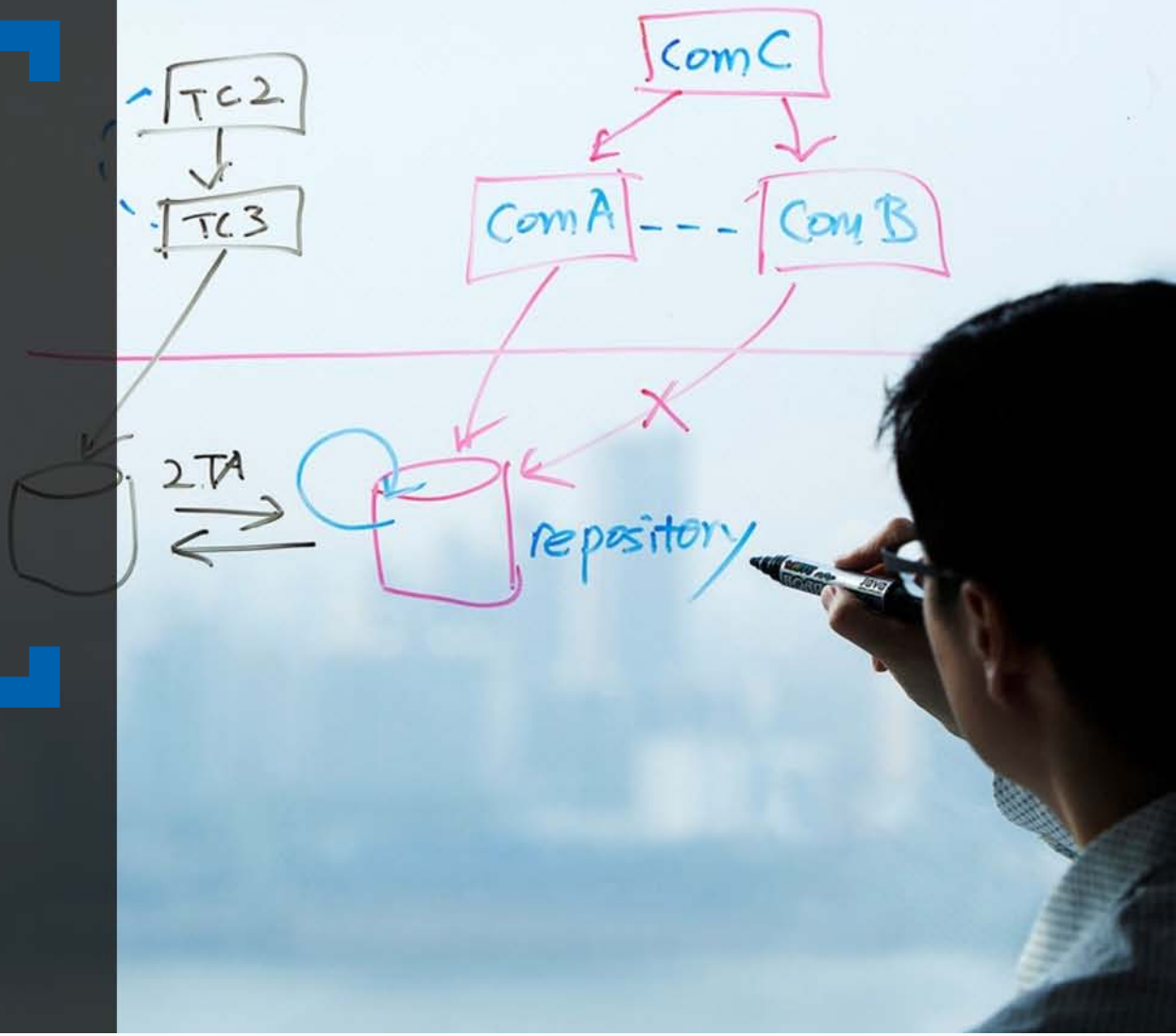
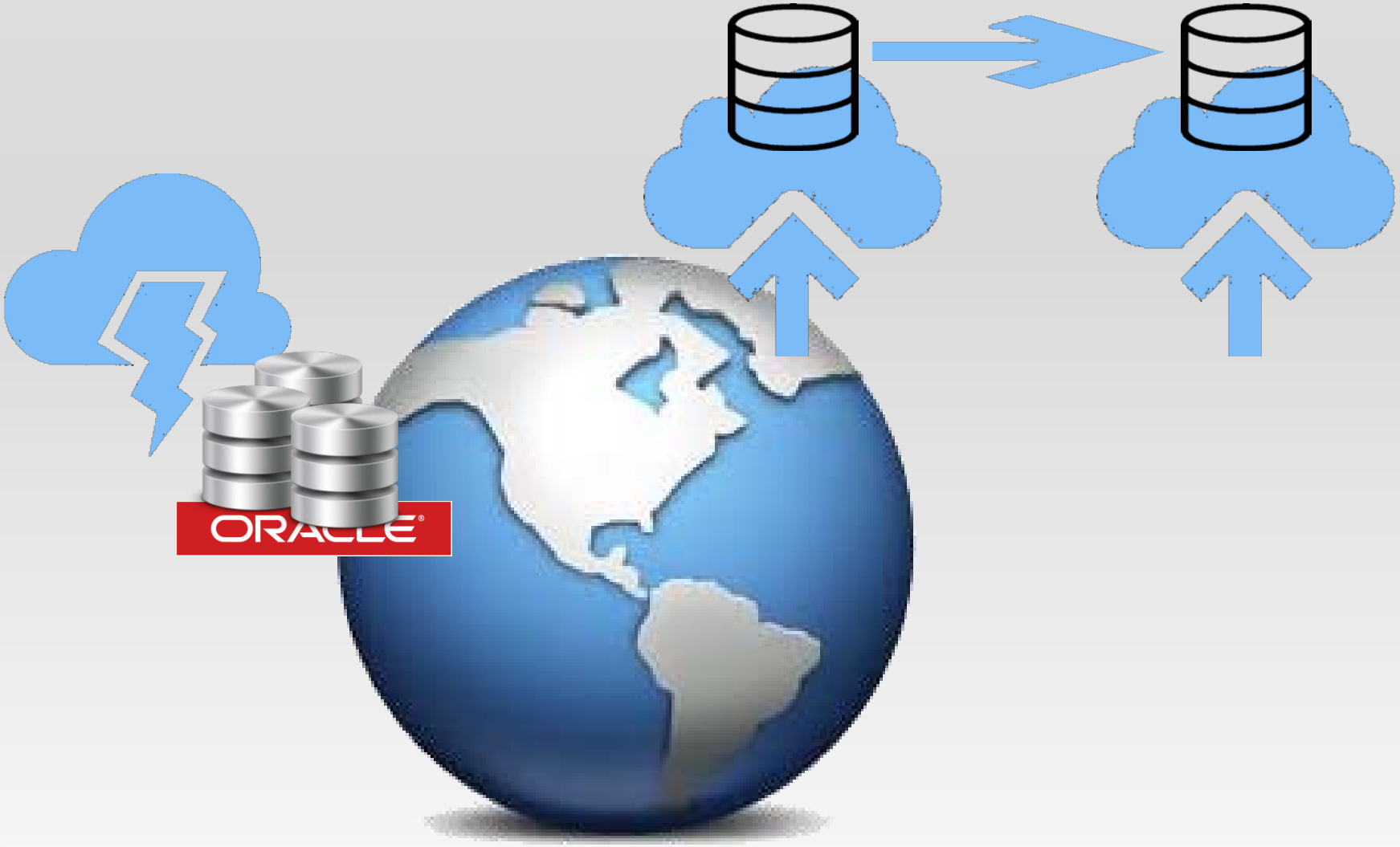
Oracle Restart and FSFO in Cloud



Agenda

I. Background
II. Core Technology
  - Hypervisor
  - Hypervisor HA
  - Oracle Restart
  - Oracle ADG
  - Oracle FSFO
III. Restart and FSFO in Cloud
IV. Issues
V. Non Stop Cloud Active Data Center

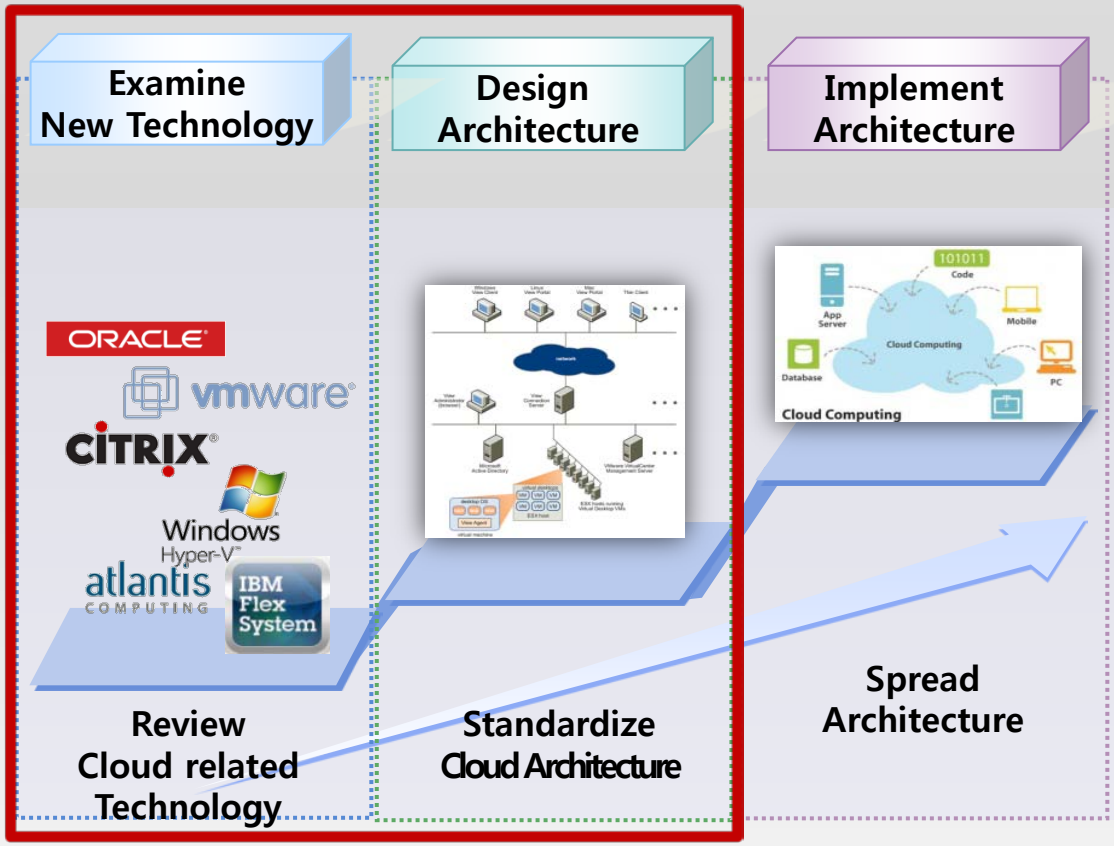
SAMSUNG



How to build Oracle in the Cloud

SAMSUNG

Background



Users Needs

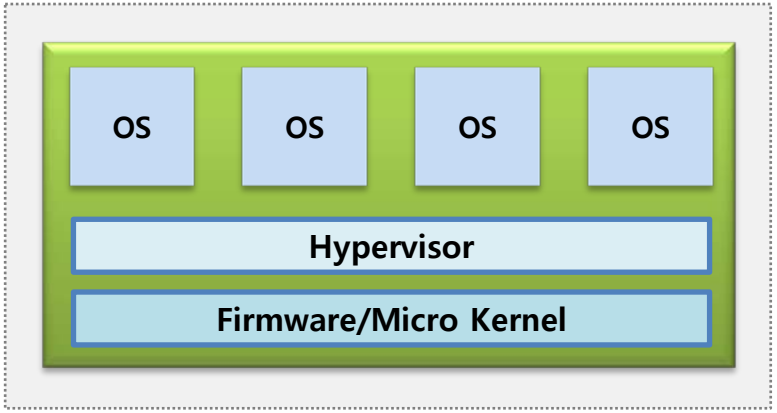
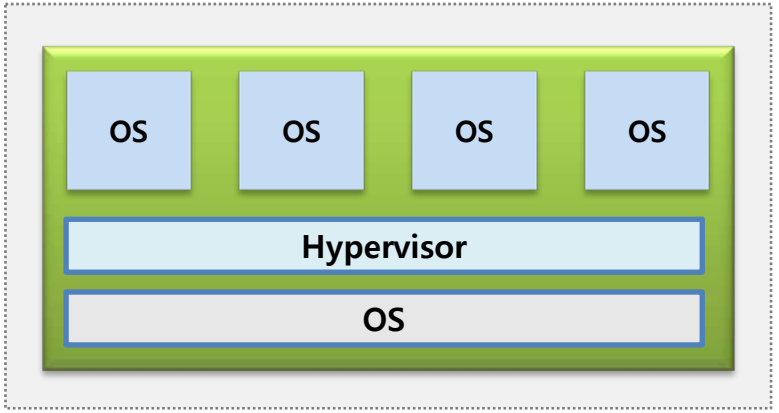
Retain Oracle DB

High Availability

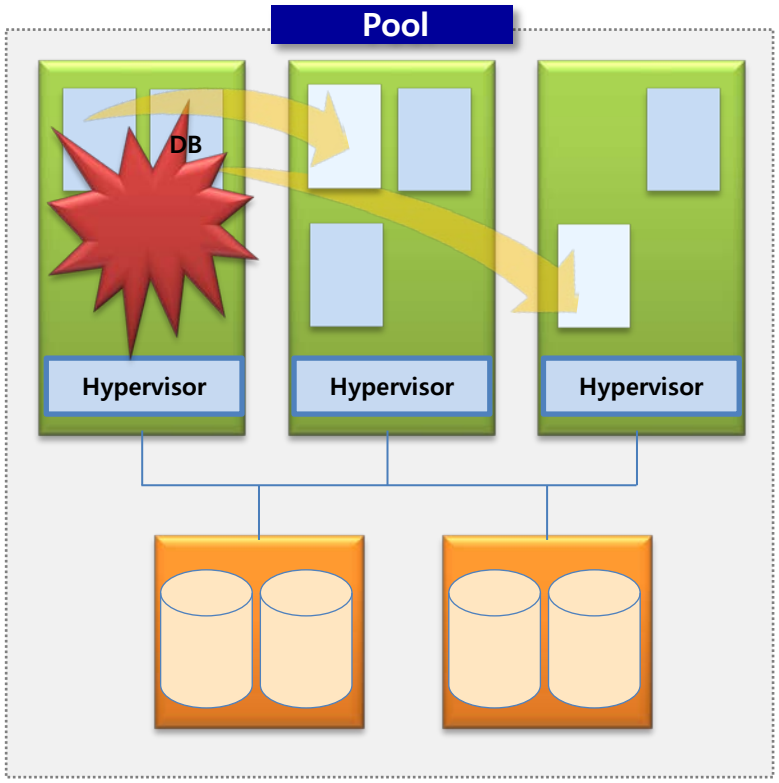
Low Cost

Need for high availability architecture in cloud

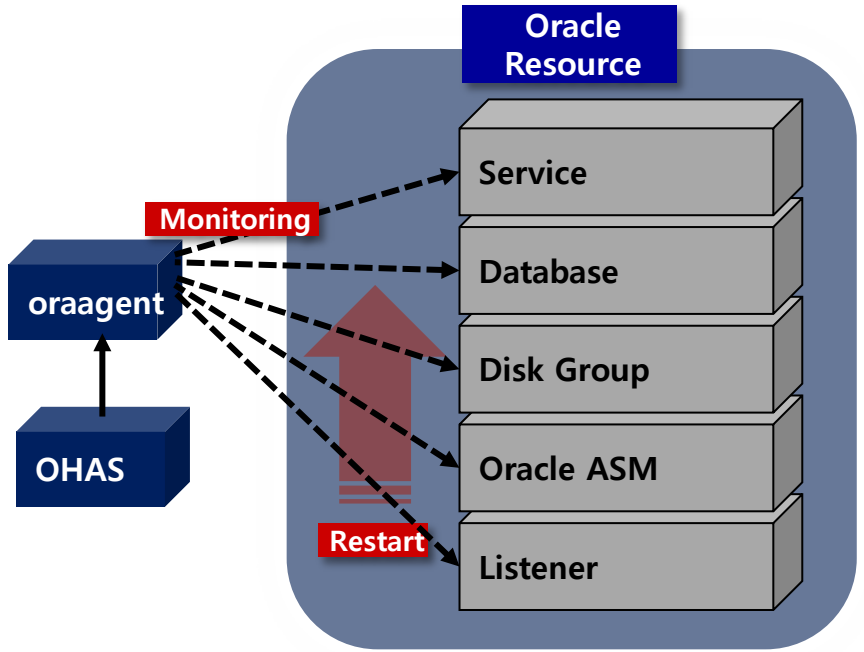
☐ Hypervisor : Virtualization solution that manages multiple operating systems

Hypervisor Type1 (Bare-metal)	Hypervisor Type2 (Hosted)
	
- Run on the Bare-metal Host Computer - Bare-metal Hypervisors . Unix IBM PowerVM(Micro Partitions/VIO) Oracle VM Server for SPARC(LDOM) . x86 VMware ESX/ESXi, MS Hyper-V Citrix XenServer, Oracle VM Server for x86	- Run on the Host OS - Hosted Hypervisors . Unix HP Integrity VM . x86 VMware Server/workstation/Fusion MS Virtual Server/PC Oracle Virtual Box KVM

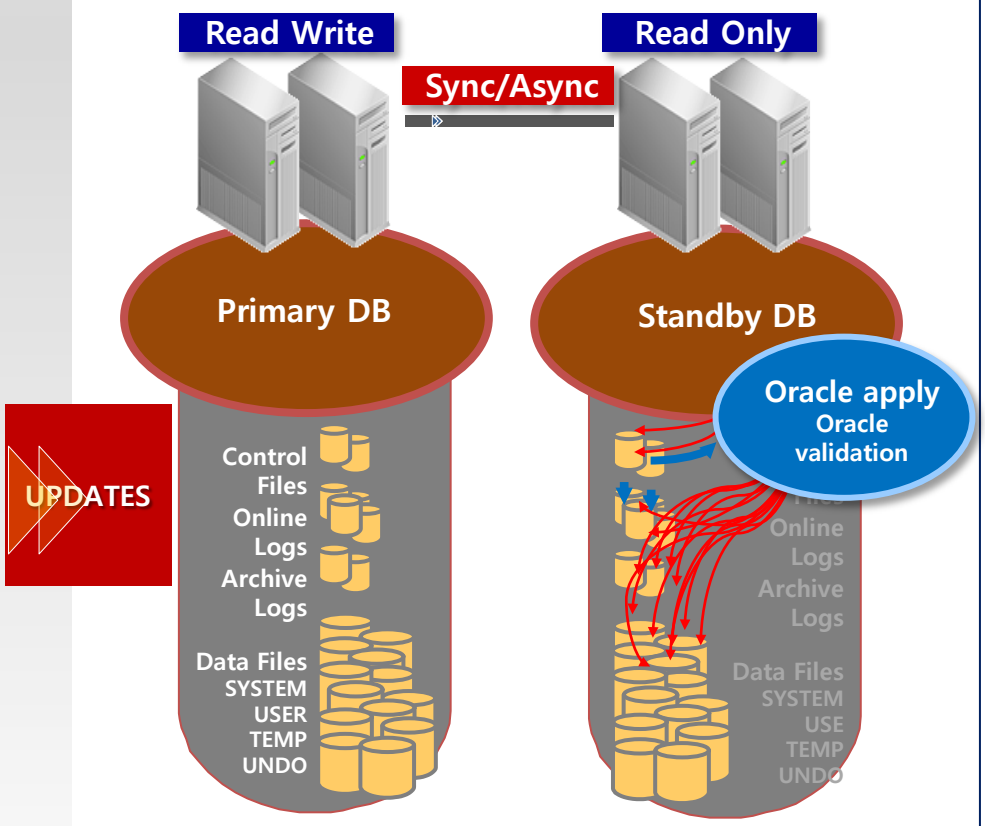
☐ Hypervisor HA

Diagram	Features
	- Fail Over . If one of the hosts fails, its VMs restart automatically on other hosts in the same pool - Live Migration . Move a VM from one host to another with no downtime - Hypervisor HA . VMWare HA . XenServer HA . Oracle VM HA

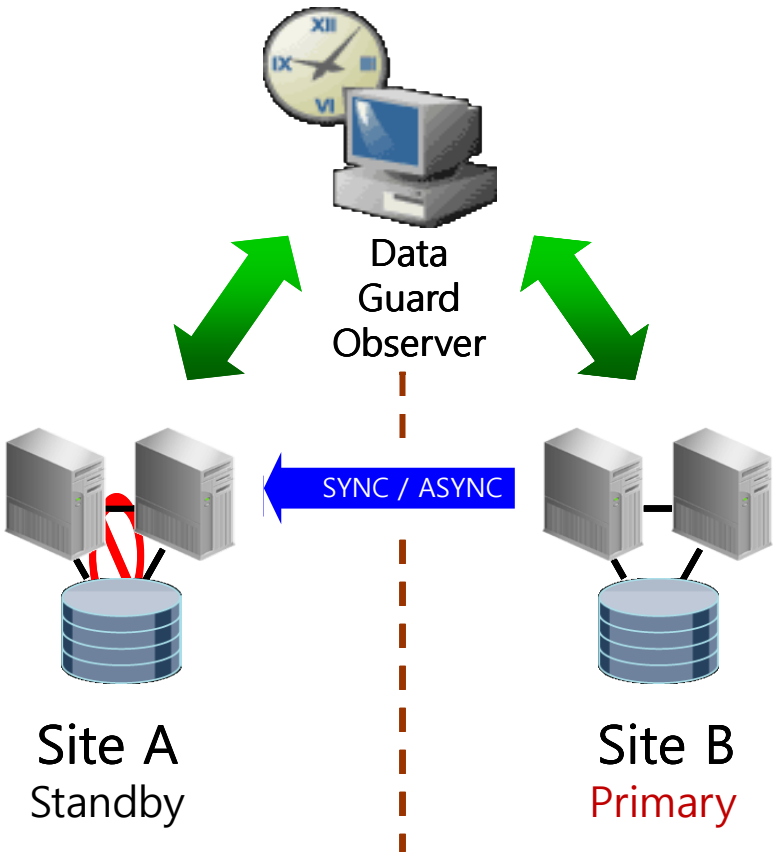
☐ Oracle Restart : HA Solution for Standalone Oracle DB

Diagram	Features
	- Oracle Restart runs periodic check operations to monitor the health of Oracle components such as Database, Listener, ASM, ASM DG and Service - If the check operation fails for a component, the component is restarted - Oracle components can be automatically restarted, whenever your database host computer restarts

☐ Active Data Guard

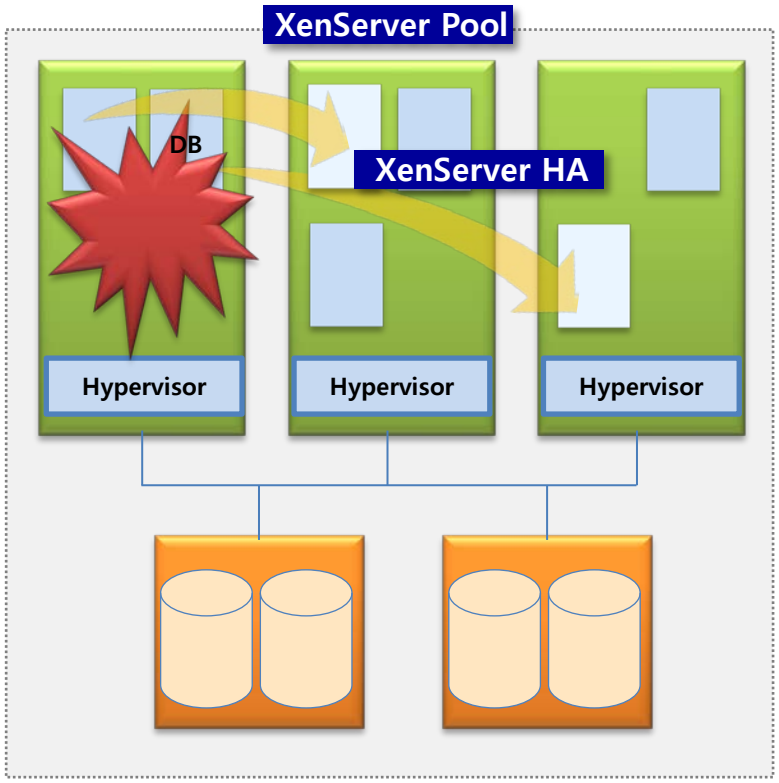
Diagram	Features
	- Protection from database file corruptions - Active Data Guard enables a physical standby database to remain open read-only, while actively applying updates received from the primary database - When Oracle detects corrupted blocks at the primary database, it will repair them online by copying the good version from an active standby database (and vice versa)

☐ Fast-Start Failover : Oracle Data Guard HA Solution (DG + HA)

Diagram	Features
	- If Observer detects primary DB failure, automatically fails over to nominated standby database - Failover is triggered under following scenarios: . Instance failures . Shutdown abort . Offline datafile due to errors . Dictionary Table corruptions - Once the primary is accessible again, Observer will re-connect and re-create a new standby DB using flashback database technology - Integrated with GI (RAC or Restart) . Failed primary automatically reinstated as standby database . Automatically start role-based services

Restart and FSFO in Cloud

☐ Cloud Architecture before Restart and FSFO

Diagram	Weak points
	- No support for OS level HA solutions - Only Hypervisor HA is applied . Does not detect DB crashes . After a DB Crash, OS Reboot or Hypervisor failover is needed; DBA must start DB manually - DR Solution is not applied in case of: . Hypervisor Pool Down . Storage or Database file Failure

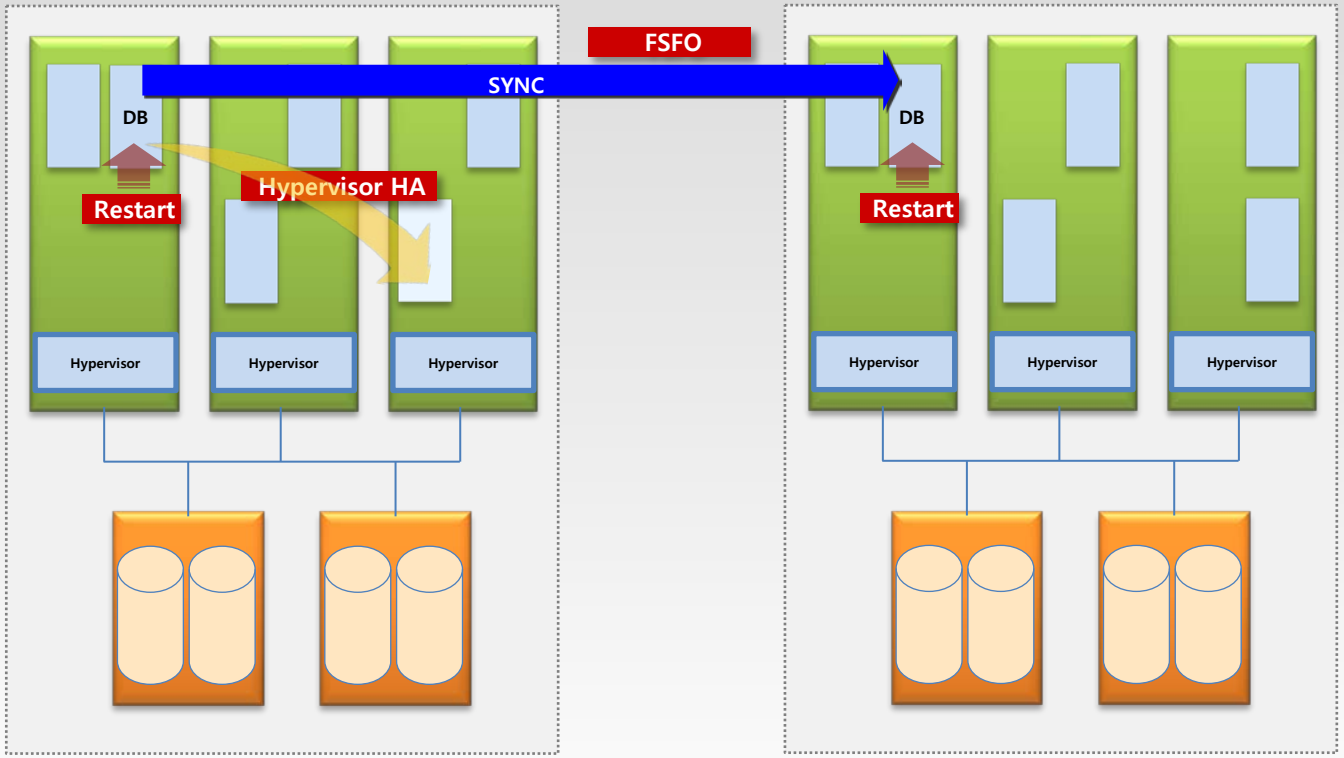
Restart and FSFO in Cloud

□ Architectural Improvement using Restart and FSFO

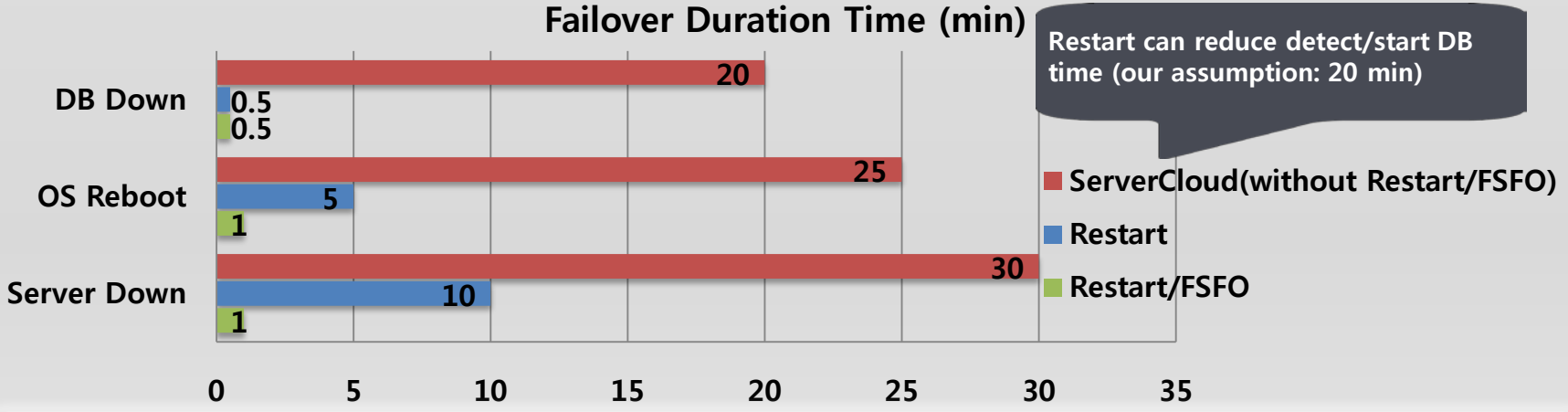
XenServer HA : Hypervisor Down

Oracle Restart : DB Crash in 30sec

Oracle FSFO : Hypervisor Down, Hypervisor Pool Down, DB File Corruption, DB Crash over 30sec, OS Reboot



Restart and FSFO in Cloud



Restart and FSFO Reduce Failover Time !!

Item	Server Cloud(As-Was)	Only Restart	Restart+FSFO
Block Corruption	✓ Manual Recovery(Over 3hr)	✓ Manual Recovery(Over 3hr)	✓ Auto Repair
DB File Corruption	✓ Manual Recovery(Over 3hr)	✓ Manual Recovery(Over 3hr)	✓ Failover to Standby (in 1min)
Storage Down	✓ Manual Recovery(Over 3hr)	✓ Manual Recovery(Over 3hr)	✓ Failover to Standby (in 1min)
Hypervisor Pool Down	✓ Manual Recovery(Over 3hr)	✓ Manual Recovery(Over 3hr)	✓ Failover to Standby (in 1min)

FSFO Reduces Recovery Time Significantly !!

Restart and FSFO in Cloud

☐ The Result of Availability Test

Category	Item	Recovery Time	HA Solution	Description
OS	DB Server Reboot	✓ 46sec	✓ FSFO	Executed Failover to Standby and Standby Reinstatement automatically
	Observer Server Reboot	✓ 0sec	✓ MonObserver.sh	Observer Restarted automatically after rebooting
	DB LAN Card Fail	✓ 44sec	✓ FSFO	Executed Failover to Standby and Standby Reinstatement automatically
DB	DB Instance Crash	✓ 26sec	✓ Restart	DB Instance was restarted automatically
	DB Listener Crash	✓ 0sec	✓ Restart	Listener was restarted automatically
	GI Stop	✓ 39sec	✓ FSFO	Executed Failover to Standby, but Standby should be reinstated manually
	Datafile Write Fail	✓ 32sec	✓ FSFO	Executed Failover to Standby, but Standby should be reinstated manually
Observer	Observer Fail	✓ 0sec	✓ MonObserver.sh	Observer Restarted automatically
DG Broker	Manual Switch Over	✓ 15sec	✓ DG Broker	Executed Switch Over by DG Broker
	Manual Fail Over	✓ 15sec	✓ DG Broker	Executed Failover and Automatic Standby Reinstatement
Hypervisor	Live Migration	✓ 0sec	✓ XenServer	Migrated to other Hypervisor online

MonObserver.sh : Observer Restart Script, registered as a cron job

Maximize Availability using Restart and FSFO

SAMSUNG

Issues

□ Observer Monitoring and Restart Script (MonObserver.sh)

Only EM supports Observer HA

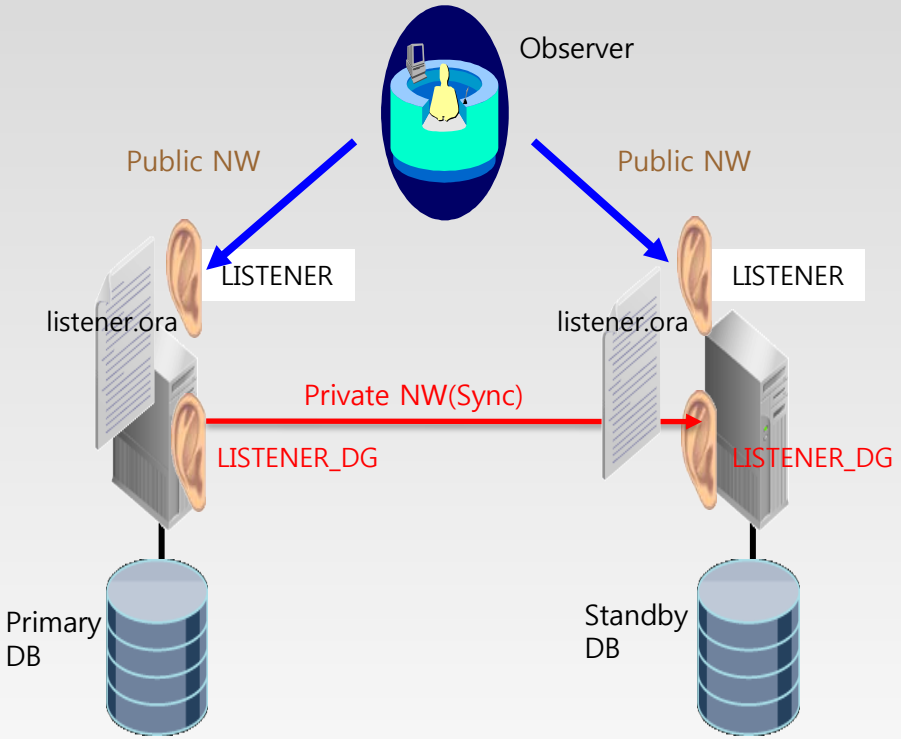
```
RESTART_NORMALDOWN="N"

$ORACLE_HOME/bin/dgmgrl -silent sys/$PW@$TNSALIAS "show database '$DBNAME'" > $ObHome/ChkObserver.log
ObserverDown=`grep -c "ORA-16819" $ObHome/ChkObserver.log|sed 's/ //g'`
ObserverCrash=`grep -c "ORA-16820" $ObHome/ChkObserver.log|sed 's/ //g'`


 if [ "$ObserverCrash" -ne "0" ]
 then

echo "Restarting Crashed Observer at `date`" >> $ObHome/MonObserver.log
$ORACLE_HOME/bin/dgmgrl -silent -logfile $ObHome/StopObserver.log sys/$PW@$TNSALIAS "stop observer;"
$ORACLE_HOME/bin/dgmgrl -silent -logfile $ObHome/Observer.log sys/$PW@$TNSALIAS "start observer FILE='$ObHome/Observer.dat'" &

 fi

 if [[ "$ObserverDown" -ne "0" && "$RESTART_NORMALDOWN" = "Y" ]]
 then

echo "Starting Shutdown Observer at `date`" >> $ObHome/MonObserver.log
$ORACLE_HOME/bin/dgmgrl -silent -logfile $ObHome/StopObserver.log sys/$PW@$TNSALIAS "stop observer;"
$ORACLE_HOME/bin/dgmgrl -silent -logfile $ObHome/Observer.log sys/$PW@$TNSALIAS "start observer FILE='$ObHome/Observer.dat'" &

 fi
```

ORA-16819 : The observer process was shut down normally
ORA-16820 : The observer process was terminated unexpectedly

Issues

☐ Add Listener (LISTENER_DG) for private NW

Oracle Restart does not support network resources



1. Modify listener.ora

```
SID_LIST_LISTENER =
  (SID_LIST =
    (SID_DESC =
      (SID_NAME = TAEKDB)
      (GLOBAL_DBNAME = TAEKDB_DGMGRL)
      (ORACLE_HOME = /oracle/app/oracle11/product/11.2.0/dbhome_1)
    )
  )

LISTENER_DG=(DESCRIPTION=(ADDRESS_LIST=(ADDRESS=(PROTOCOL=TCP)(HOST
=Private IP)(PORT=1531))))
SID_LIST_LISTENER_DG =
  (SID_LIST =
    (SID_DESC =
      (SID_NAME = TAEKDB)
      (GLOBAL_DBNAME = TAEKDB_DGB)
      (ORACLE_HOME = /oracle/app/oracle11/product/11.2.0/dbhome_1)
    )
  )
```

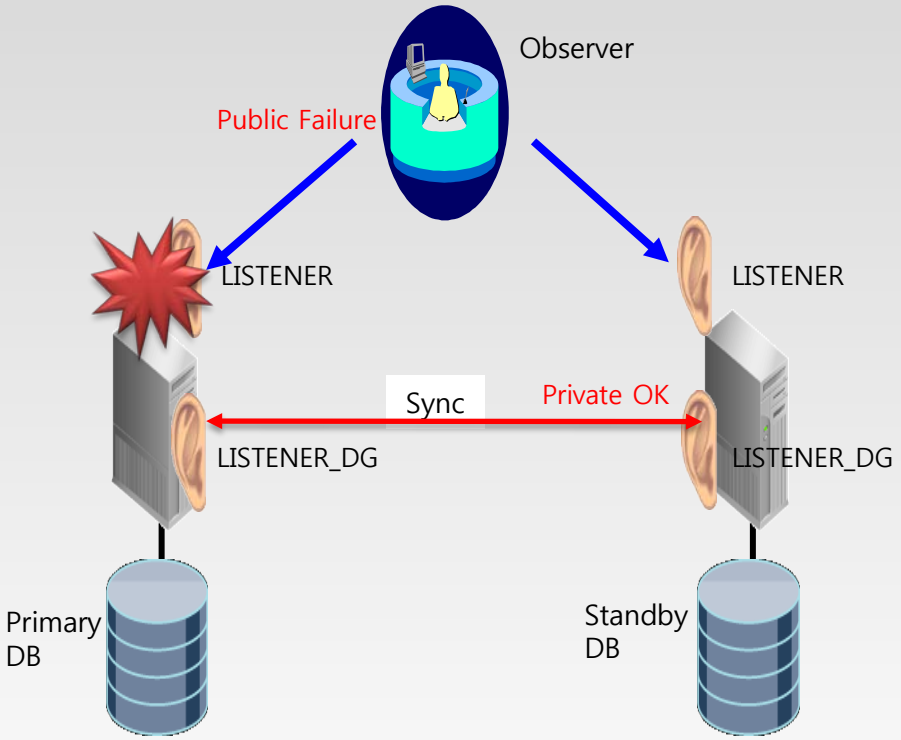
2. Add LISTENER_DG

```
$ srvctl add listener -l LISTENER_DG -p 1531
$ srvctl start listener -l LISTENER_DG
$ crsctl status res -t
$ crsctl stat res ora.LISTENER_DG.lsnr -p
$ crsctl modify resource ora.LISTENER_DG.lsnr -attr "ENDPOINTS="
  → for private endpoint only (MetaLinkID:1544433.1)
$ srvctl stop listener -l LISTENER_DG
$ srvctl start listener -l LISTENER_DG
```

Issues

☐ Public NIC Failure

Even when the Primary Public NIC Fails, the Private LAN continues to be available
→ Failover to Standby not executed, even when service down scenario encountered



ObserverOverride (New Features in 11.2.0.4)

```
Properties:
   FastStartFailoverThreshold      = '30'
   OperationTimeout                = '30'
   FastStartFailoverLagLimit       = '30'
   CommunicationTimeout            = '180'
   ObserverReconnect               = '10'
   FastStartFailoverAutoReinstate  = 'TRUE'
   FastStartFailoverPmyShutdown    = 'TRUE'
   BystandersFollowRoleChange      = 'ALL'
   ObserverOverride                = 'TRUE'
   ExternalDestination1            = ''
   ExternalDestination2            = ''
   PrimaryLostWriteAction          = 'CONTINUE'
```

ObserverOverride
The ObserverOverride configuration property, when set to TRUE, allows an automatic failover to occur when the observer has lost connectivity to the primary, even if the standby has a healthy connection to the primary.

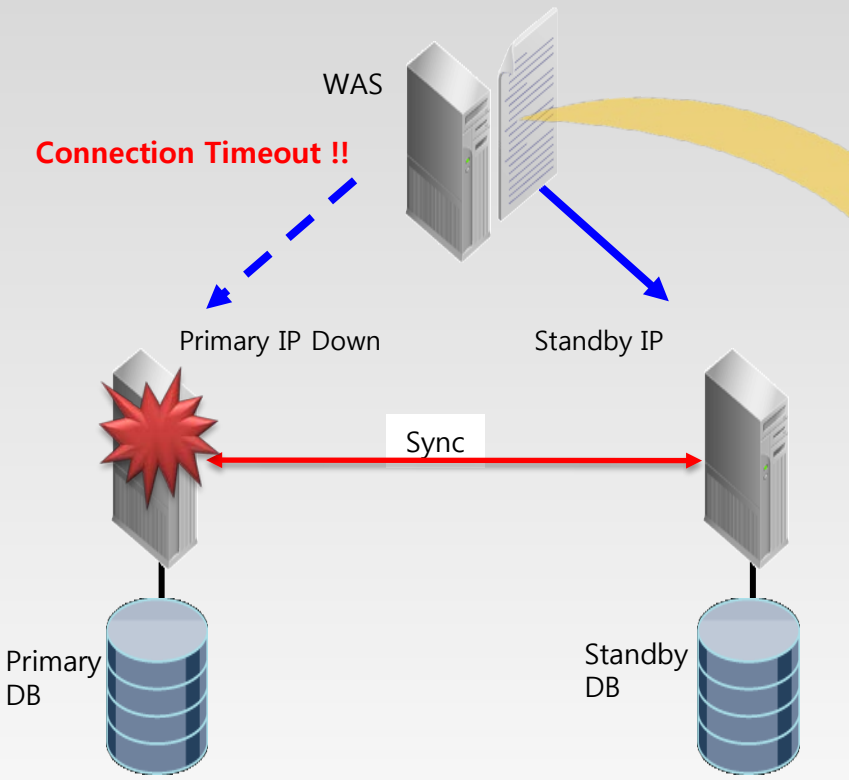
Application of Monit, an open source HA solution (<= 11.2.0.3)

```
cat /etc/monit.conf

check host myserver with address xx.xxx.xxx.xxx (← Public IP)
#  start program = ""
  stop program = "/oracle/app/oracle11/product/11.2.0/grid/bin/crsctl stop has"
  if failed icmp type echo count 10 with timeout 3 seconds
    then stop
```

☐ TCP Connection Timeout

Oracle does not support VIP between Primary DB and Standby DB

- Linux TCP Connection Timeout is 21sec : tcp_syn_retries (Default 5(3+6+12=21sec))



tnsnames.ora

```
ORCLRW =
 (DESCRIPTION_LIST =
  (LOAD_BALANCE = OFF)
  (FAILOVER = ON)
  (DESCRIPTION =
   (ENABLE = BROKEN)
   (CONNECT_TIMEOUT = 5)(TRANSPORT_CONNECT_TIMEOUT=3)(RETRY_COUNT=1)
   (ADDRESS = (PROTOCOL = TCP)(HOST = Primary Public IP)(PORT = 1521))
   (CONNECT_DATA =
    (SERVER = DEDICATED)
    (SERVICE_NAME = orclrw)
   )
  )
  (DESCRIPTION =
   (ENABLE = BROKEN)
   (CONNECT_TIMEOUT = 5)(TRANSPORT_CONNECT_TIMEOUT=3)(RETRY_COUNT=1)
   (ADDRESS = (PROTOCOL = TCP)(HOST = Standby Public  IP)(PORT = 1521))
   (CONNECT_DATA =
    (SERVER = DEDICATED)
    (SERVICE_NAME = orclrw)
   )
  )
)
```

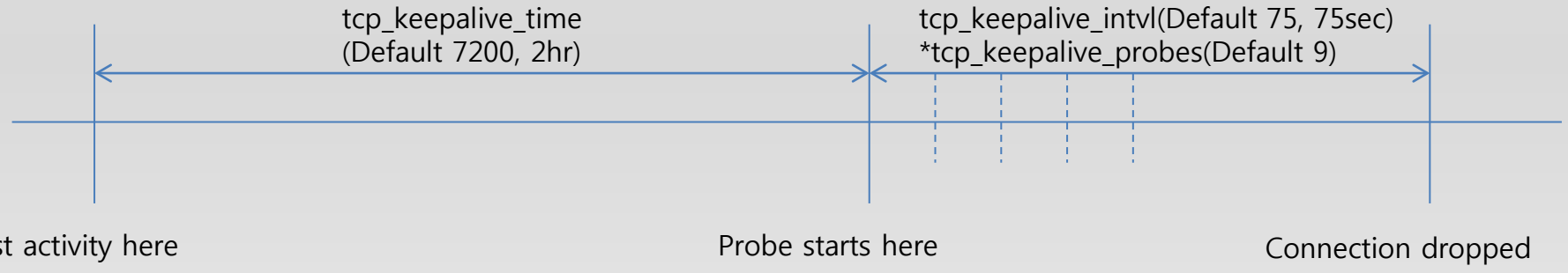
The CONNECT_TIMEOUT parameter is equivalent to the sqlnet.ora parameter SQLNET.OUTBOUND_CONNECT_TIMEOUT and overrides it.
The TRANSPORT_CONNECT_TIMEOUT parameter is equivalent to the sqlnet.ora parameter TCP.CONNECT_TIMEOUT and overrides it.

Issues

☐ TCP Keep Alive

Oracle disables Keep Alive setting on Client

- Linux TCP Keep Alive Parameter



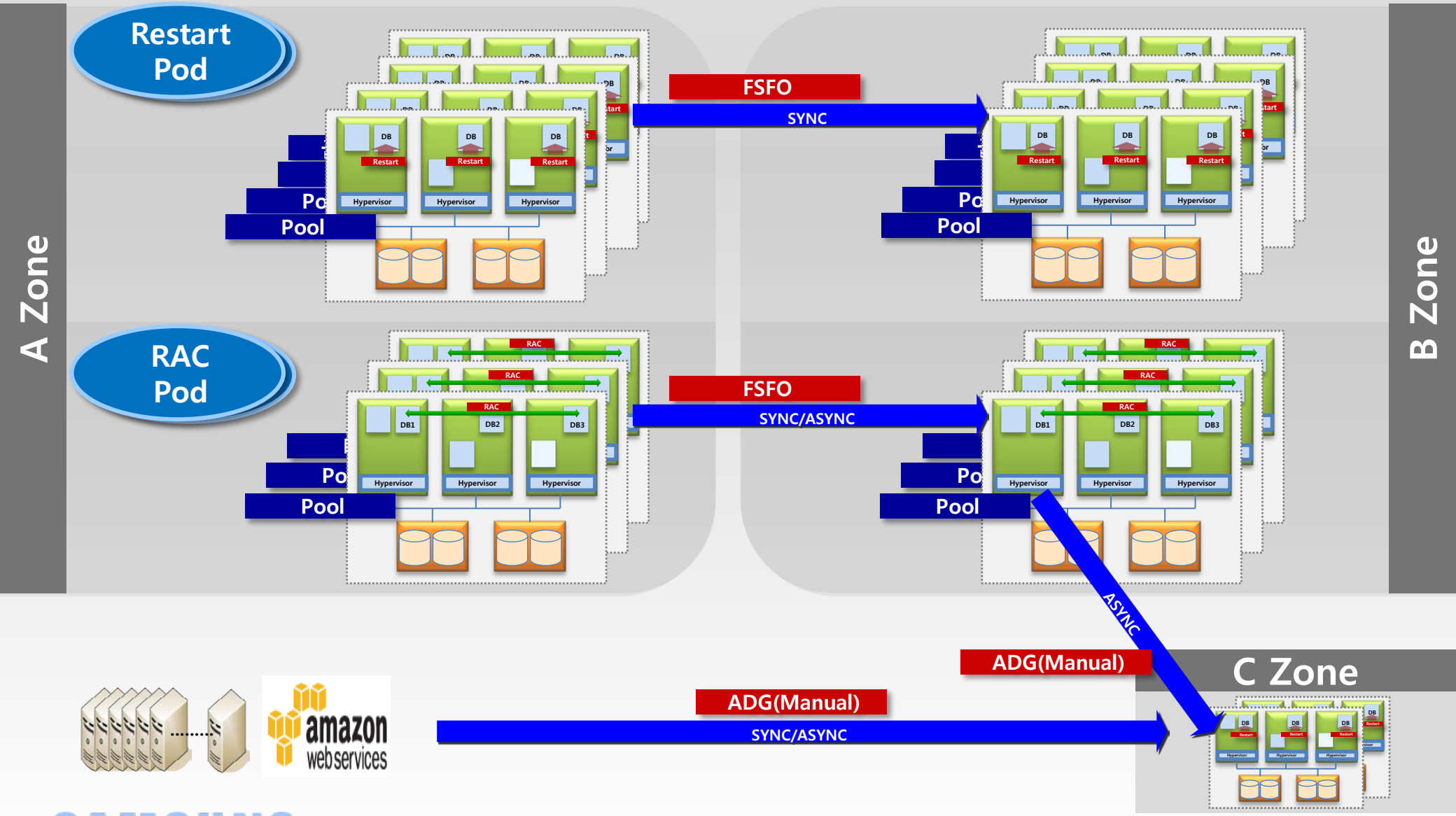
1. Modify TCP Keep Alive Parameters
/etc/sysctl.conf

```
net.ipv4.tcp_keepalive_time = 30
net.ipv4.tcp_keepalive_intvl = 5
net.ipv4.tcp_keepalive_probes = 3
```

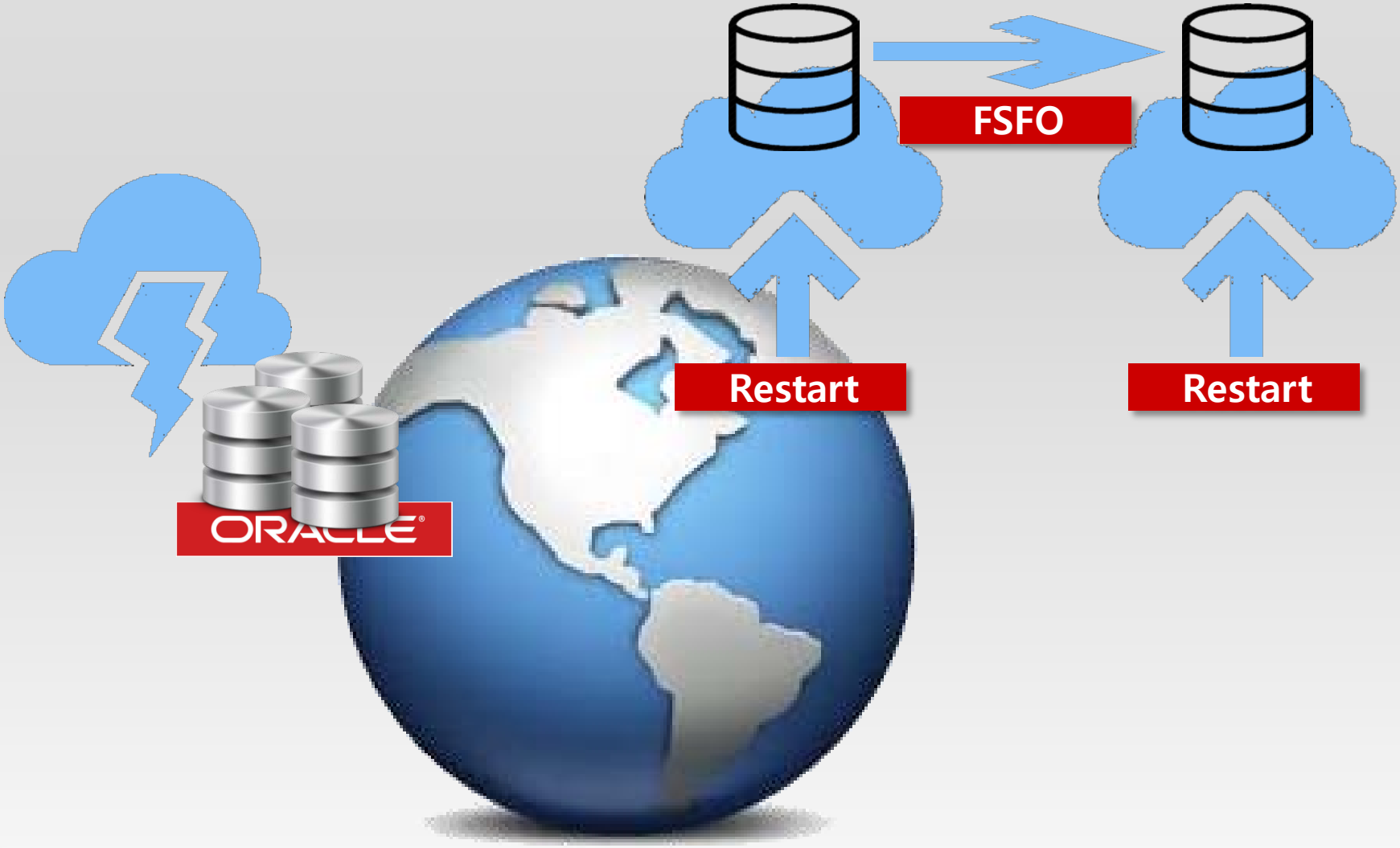
2. Enable TCP Keep Alive (ENABLE = BROKEN)
$ORACLE_HOME/network/admin/tnsnames.ora

```
ORCLRW =
 (DESCRIPTION_LIST =
   (LOAD_BALANCE = OFF)
   (FAILOVER = ON)
   (DESCRIPTION =
     (ENABLE = BROKEN)(CONNECT_TIMEOUT = 5)(TRANSPORT_CONNECT_TIMEOUT=3)(RETRY_COUNT=1)
       .
       .
       .
 )
```

Non Stop Cloud Active Data Center



SAMSUNG



We can build Oracle in the Cloud