Oracle 11G SCAN: Concepts and implementation experience sharing

Yury Velikanov
Senior Oracle DBA
There are Two SCAN related news

- **Good**
  - SCAN is based on known components you worked for years now

- **Other news**
  - SCAN uses those components in different way
Few words about myself …

• Oracle **ACE** and **RAC SIG** regional leader
  • @yvelikanov

• Started as Oracle DBA
  - with 7.2 (in 1997, 14+)

• First international appearance
  - 2005 - Hotsos Symposium 2005

• First RAC experience
  - 2000 FIFA - Oracle Parallel Server

• Education *(Master Degree in Computer science)*
  - OCP 7/8/8i/9/10 + OCM 9i/10g
  - Awaiting results from 11G OCM exam

• Several 11GR2 RAC projects in production
  - Including GNS implementation
SCAN Agenda

- Introduction
- SCAN Infrastructure Main Components
- SCAN troubleshooting
- Advanced points
- Q & A

Please tweet about the event @yvelikanov @racsig
Please give your feedback velikanovs@pythian.com
I will continue to blog on SCAN topics with all your support
Single [Client Access] Name

scan.clustgrid-prod.yourdomain.com

+ service
SCAN Introduction

• Single Client Access Name
• Addresses the TNSNAMES multi address issue

• Old - 10G FAILOVER
  • Complex TNS entries
  • Complex to manage (add a node)
  • Previous Oracle Clients support

• New - 11GR2
  • One Simple TNS entry on client side
  • Easy to add nodes (address forwarding)
  • Transparent to Oracle Client versions
  • No static listener.ora file
SCAN and PREV tnsnames.ora

PROD_HR.yourdomain.com =
  (DESCRIPTION =
    (ADDRESS_LIST =
      (FAILOVER=on)
      (LOAD_BALANCE=TRUE)
      (ADDRESS = (PROTOCOL = TCP)(HOST = vip.node1)(PORT = 1523))
      (ADDRESS = (PROTOCOL = TCP)(HOST = vip.node2)(PORT = 1523))
      (ADDRESS = (PROTOCOL = TCP)(HOST = vip.node3)(PORT = 1523))
    )
    (CONNECT_DATA = (SERVICE_NAME = HR) )
  )

scan.clustgrid-prod.yourdomain.com:1523/HR
Ora*Net: Easy Connect

scan.clustgrid-prod.yourdomain.com
scan.clustgrid-prod.yourdomain.com:1521
scan.clustgrid-prod.yourdomain.com:1521/HR
scan.clustgrid-prod.yourdomain.com:1521/HR:dedicated/ERP1

PROD _HR.yourdomain.com =
  (DESCRIPTION =
   (ADDRESS = (PROTOCOL = TCP)(HOST = scan.clustgrid-prod)(PORT = 1521))
   (CONNECT_DATA =
     (SERVICE_NAME = HR)
   )
  )

Oracle® Database Net Services Administrator’s Guide
11g Release 2 (11.2)
Part Number E10836-06
http://download.oracle.com/docs/cd/E11882_01/network.112/e10836/naming.htm#BABJBFHJ
Oracle Client Connection process
1. Oracle Client sends a request to DNS asking for SCAN IP address
2. DNS delegate the request processing to GNS service
3. GNS returns one of the SCAN IPs in round-robin
4. The client sends TNS request (IP/PORT/Service) to one of the SCAN listeners
5. SCAN listener forwards request to one of the DB Listener (Virtual IPs)
6. DB Listener creates a foreground process and connect the Oracle Client with that process
SCAN Infrastructure Main Components

- Single Client Access Name + Oracle Services (Definitions)
- DNS - resolving SCAN to 3 IP addresses (Round Robin)
  - Primary / Secondary
  - NameServer configuration
- SCAN Listeners
  - Keeps records on available Local Listeners and Services those serve
  - Forwards connections to less loaded Local Listener
- Local (VIP) Listeners
  - Creates foreground processes
  - Manages sockets
- RAC (SCAN / VIP / Interconnect) ip addresses
- Grid Name Service
  - registers and resolves RAC ip addresses
- DHCP
  - Assign dynamically IP addresses
DNS resolves `scan-cluster_a.mycompany.com` to 3 SCAN IP addresses.

SCAN (remote) listeners register and distribute information about local listeners serving a SERVICE.

Oracle Instance:
1. registers SERVICES it running in LOCAL listeners
2. registers LOCAL listeners serving a service in SCAN listeners
SCAN troubleshooting

• **Service Names**
  - DO NOT MODIFY init.ora:service_name
  - USE srvctl to configure and manage services

```bash
srvctl config service -d <DB Name>
```

... Service name: DEVERP_APEX.GGT.COM Service is enabled Failover type: NONE Preferred instances: DEVERP1 Available instances: DEVERP1, DEVERP2, DEVERP3, DEVERP4, DEVERP5, DEVERP6 ...

```sql
show parameter service_name
```

```
NAME            TYPE                 VALUE
-----------------------------------------------------------------------------------------
service_names    string               DEVERP_CDC.GGT.COM, SYS$APPLSYS.WF_CONTROL.DEVERP.WORLD, SYS$STREAMS_ADMIN.CDC$Q_ERP.DEVERP.WORLD, DEVERP_WEBM.GGT.COM, DEVERP_WWT_B2B.GGT.COM, DEVERP_RFUI.GGT.COM, DEVERP_IBI.GGT.COM, DEVERP_GENERAL.WWT.COM, DEVERP_BI.GGT.COM, DEVERP_APEX.GGT.COM, DEVERP_10g, DEVERP1, DEVERP

SQL>
```
SCAN troubleshooting

• Oracle Listeners
  • Running under **grid** OS user
    • Don’t start it under ORACLE user
    • If you do you end up with a mess
  • Manage (start/stop) by **srvctl**
    • Be careful with manual start/stop (TNS_ADMIN)
  • **listener.ora** is dynamic configuration file by default
    • All parameters managed by Cluster
    • Use srvctl to configure

• Make sure listeners listen on corresponding IPs
  >lsnrctl status LISTENER_SCAN2
  ...
  (DESCRIPTION=(ADDRESS=(PROTOCOL=ipc)(KEY=LISTENER_SCAN2)))
  (DESCRIPTION=(ADDRESS=(PROTOCOL=tcp)(HOST=10.2.9.122)(PORT=1523)))
  ...
  • LISTENER_SCAN1/2/3 on SCAN IPs
  • LISTENER on VIP and Public IPs
SCAN troubleshooting

- **init.ora:local_listener**
  - It is an OLD good parameter
  - The same rules applies
    - Specify LOCAL listener only!
    - Cant stress enough !!! NO SCAN !!!
  - You can use TNS address directly or TNS alias
  - !!! If can’t resolve an instance won't start !!!

<table>
<thead>
<tr>
<th>NAME</th>
<th>TYPE</th>
<th>VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>local_listener</td>
<td>string</td>
<td>(DESCRIPTION=(ADDRESS_LIST=(ADDRESS=(PROTOCOL=TCP) (HOST=devracdb1-vip) (PORT=1534)) (ADDRESS=(PROTOCOL=TCP) (HOST=devracdb1-vip) (PORT=1521))) )</td>
</tr>
</tbody>
</table>

```sql
SQL> show parameter local_listener
NAME            TYPE     VALUE
--------------------------
local_listener  string   (DESCRIPTION=(ADDRESS_LIST=(ADDRESS=(PROTOCOL=TCP) (HOST=devracdb1-vip) (PORT=1534)) (ADDRESS=(PROTOCOL=TCP) (HOST=devracdb1-vip) (PORT=1521))) )

SQL> show parameter local_listener
NAME            TYPE     VALUE
--------------------------
local_listener  string   devracdb1-vip

tnsping devracdb1-vip
```
SCAN troubleshooting

- **init.ora:remote_listener**

SQL> show parameter remote_listener

<table>
<thead>
<tr>
<th>NAME</th>
<th>TYPE</th>
<th>VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>remote_listener</td>
<td>string</td>
<td>scan.clustgrid-prod.yourdomain.com</td>
</tr>
</tbody>
</table>

- The same management principals apply
  - Make **SURE** it points to SCAN IP addresses only
    - Cant stress enough **!!! NO VIP !!!**
  - Any valid TNS config is acceptable
    - tnsnames alias
    - sqlnet.ora
      - NAMES.DIRECTORY_PATH=(TNSNAMES, EZCONNECT)
  - Use **SCAN** or IPs (for static SCAN conf only)

**!!! If can’t resolve an instance won't start !!!**
SCAN troubleshooting

• DNS
  • dig (Linux os command)

```
[oracle@host01 admin]$ dig scan.clustgrid-prod.yourdomain.com
[oracle@host01 admin]$ nslookup scan.clustgrid-prod.yourdomain.com | grep Address | tail -1
Address: 172.30.193.218
[oracle@host01 admin]$ nslookup scan.clustgrid-prod.yourdomain.com | grep Address | tail -1
Address: 172.30.193.216
[oracle@host01 admin]$ nslookup scan.clustgrid-prod.yourdomain.com | grep Address | tail -1
Address: 172.30.193.218
```

```
[oracle@host01 admin]$ dig scan.clustgrid-prod.yourdomain.com
; <<>> DiG 9.3.6-P1-RedHat-9.3.6-4.P1.el5 <<>> scan.clustgrid-prod.yourdomain.com
;; global options: printcmd
;; Got answer:
;; <<HEADER>>
;; opcode: QUERY, status: NOERROR, id: 15137
;; flags: qr rd ra
;; QUERY: scan.clustgrid-prod.yourdomain.com. 1 IN A
;; ANSWER SECTION:
scan.clustgrid-prod.yourdomain.com. 58 IN A 172.30.193.218
scan.clustgrid-prod.yourdomain.com. 58 IN A 172.30.193.216
scan.clustgrid-prod.yourdomain.com. 58 IN A 172.30.193.217
```
GNS Advanced points

- SCAN + GNS implementation
- Most probably you do not need it
- Makes the configuration 100% dynamic
- Unlimited number of nodes with simple Oracle Client Configuration
- Oracle retrieves new IPs from DHCP for SCAN / VIP / [Interconnect] components at startup time
- The only static RAC IP is GNS IP
GNS Advanced points

- Additional components
  - Grid Name Service
  - DNS and GNS integration (SCAN/VIP)
  - Dedicated DHCP service
    - Separate Network Segment
    - DHCP redundancy could be an issue
  - RAC and DHCP integration
    - Make DHCP assigning the same IPs (or range) each time per RAC process (Joseph Griffiths)
  - DHCPDISCOVER from 00:00:00:00:00:00:00 via eth0

- Many things could go wrong !!!
- [GNS Troublesooting](http://blog.jgriffiths.org/?p=24) - see my blog
There are Two SCAN related news

- **Good**
  - SCAN is based on known components you worked for years now

- **Other news**
  - SCAN uses those components in different way
Pythian Facts

• Founded in 1997, over 14 years
• 100+ employees
• 5 offices in 5 countries
• Employ
  • 6 Oracle ACEs (Including 1 ACE director)
  • Several Oracle Masters
  • Plenty of technical geeks
• Platinum level partner in the Oracle Partner Network
• Actively supports technical communities via
  • Blogging
  • Conferences
  • SIGs and other events
Please tweet about the event @yvelikanov @racsig
Please give your feedback velikanovs@pythian.com
I will continue to BLOG on SCAN and other topics with all your support