Oracle  RAC 同时具备HA(High Availiablity) 和LB(LoadBalance).

》理解LB(LoadBalance)

Oracle是如何知道每个节点的负载情况的呢？

通过监听日志和pmon profile

每次数据库启动pmon都会把实例注册到LOCAL\_LISTENER and REMOTE\_LISTENER，每次发生连接pmon都会将信息更新到监听日志中

**service\_register指示初始注册信息**

$ grep service\_register /u01/app/11.2.0/grid/log/diag/tnslsnr/rac2/listener\_scan1/alert/log.xml

<txt>16-APR-2018 11:04:32 \* service\_register \* orcl1 \* 0

<txt>16-APR-2018 11:04:32 \* service\_register \* orcl2 \* 0

<txt>08-MAY-2018 14:45:48 \* service\_register \* LsnrAgt \* 0

<txt>08-MAY-2018 14:46:38 \* service\_register \* orcl1 \* 0

<txt>08-MAY-2018 14:46:39 \* service\_register \* orcl2 \* 0

**service\_update指示连接负载信息**

$ grep service\_update /u01/app/11.2.0/grid/log/diag/tnslsnr/rac2/listener\_scan1/alert/log.xml |wc -l

15379

负载重的db 每1min要去update监听，负载少的可以每10min中update监听

**establish指示连接数（总数）**

Local instance connections (same node as listener)

13-FEB-2004 11:05:02 \* (CONNECT\_DATA=(SERVER=DEDICATED)(SERVICE\_NAME=prod)

(CID=(PROGRAM=)(HOST=clienthost)(USER=oracle))) \*

(ADDRESS=(PROTOCOL=tcp)(HOST=10.10.10.10)(PORT=20108)) \* establish \* prod \* 0

Remote instance connections (redirected connections from non-local listeners)

13-FEB-2004 11:05:02 \* (CONNECT\_DATA=(SERVER=DEDICATED)(SERVICE\_NAME=prod)

(CID=(PROGRAM=)(HOST=clienthost)(USER=oracle))(INSTANCE\_NAME=Instance\_B)) \*

(ADDRESS=(PROTOCOL=tcp)(HOST=10.10.10.10)(PORT=20110)) \* establish \* prod \* 0

Notice that line 2 of the Listener\_B.log entry has the additional string "INSTANCE\_NAME=Instance\_B" identifying it as a connection that was directed from a different listener.

**INSTANCE\_NAME指出链接是由别的监听分发过来的，但无法指出是哪个监听where**

Listener\_A

grep establish listener\_A.log | wc -l ==> 800(连接总数)

grep INSTANCE\_ listener\_A.log | wc -l ==> 300(重定向分发连接数)

Listener\_B

grep establish listener\_B.log | wc -l ==> 600(E)

grep INSTANCE\_ listener\_B.log | wc -l ==> 130(F)

(total establish) - (redirected) = direct client connections

Listener\_A: 800(C) - 300(D) = 500(G)

Node\_A had 500 direct client connections

Listener\_B: 600(E) - 130(F) = 470(H)

Node B had 470 direct client connections

》纯技术分散负载

在用户连接这个层次进行，在连接时根据每个节点的负载决定把链接分配到那个实例

1 Client-side load balancing

配置方法是在客户端tnsnames.ora文件中加入LOAD\_BALANCE=YES

当客户端发起连接时，从地址列表中随机选取一个，在使用随机算法把连接请求分散到各个实例

TESTCT =

  (DESCRIPTION =

    (LOAD\_BALANCE = yes)

    (ADDRESS\_LIST =

      (ADDRESS = (PROTOCOL = TCP)(HOST = 192.168.6.113)(PORT = 1521))

      (ADDRESS = (PROTOCOL = TCP)(HOST = 192.168.6.114)(PORT = 1521))

    )

    (CONNECT\_DATA =

      (SERVICE\_NAME = orcl)

    )

  )

这种方式缺点明显，因为在分配连接时没有考虑每个节点真实负载，最后分配结果不一定平衡。且随机算法需要长时间片，短时间内同时发起多个连接，很可能被分配到同意节点。

2 Server-side load balancing

当使用DBCA创建rac时，自动配置Server-side load balancing

会自动配置REMOTE\_LISTENER=scan\_name:scan\_port

11gR2之后remote\_listener自动注册到scan监听中，所以使用scan ip即可实现server-side LB

Oracle Database 11g Release 2 and later, by default, instances register with SCAN listeners as remote listeners.

SQL> show parameter listener

NAME                     TYPE     VALUE

------------------------------------ ----------- ------------------------------

listener\_networks             string

local\_listener                 string     (DESCRIPTION=(ADDRESS\_LIST=(AD

                         DRESS=(PROTOCOL=TCP)(HOST=192.

                         168.80.243)(PORT=1521))))

remote\_listener              string     rac-scan:1521

SQL>

Note:

The following features do not work with the default database service. You must create cluster managed services to take advantage of these features. You can only manage the services that you create. Any service created automatically by the database server is managed by the database server.

服务器端负载均衡的实现依赖于listener收集的负载信息。

pmon后台进程会收集负载信息，然后登记到listener中。约1min~10min更新一次,节点负载高，则更新频率高，以保证listener能够准确掌握负载情况

在listener日志中可以很清楚看到pmon的动作

\*service\_register\*      --第一次注册

\*service\_update\*       --自动更新动作

记数了动作，却没有内容

可使用1025事件追踪pmon进程

pmon不仅可以向本地注册，还可以向远程节点注册。需要参数

local\_listener,remote\_listener

这样集群每个节点listener都掌握所含有节点负载情况，收到连接时，就把连接分配到负载最小的节点。

配置方法是配置remote\_listener参数，此参数是一个tnsnames

TESTCT =

    (ADDRESS\_LIST =

      (ADDRESS = (PROTOCOL = TCP)(HOST = 192.168.6.113)(PORT = 1521))

      (ADDRESS = (PROTOCOL = TCP)(HOST = 192.168.6.114)(PORT = 1521))

    )

其中服务器端的listener.ora文件中需要去掉自动生成的 sid\_list\_listener\_dbs=() 列表

保留LISTENER\_DBS=().保证listener信息都是动态注册获得

client-side LB server-side LB 不互斥，可一起工作

》service分散负载

connection balance不足：

“分散数据”，rac环境中按照应用分离数据。

假设数据库有两个节点

两个销售 >connection balance>分散到两个节点。销售数据在CACHE FUSION作用下不断在两个节点传递数据

又有两个生成用户connection balance>分散到两个节点。销售数据在CACHE FUSION作用下不断在两个节点传递数据

可见，表面上用户被分散到了不同的instance上，似乎负载被分散了。但这种分散没有结合业务需求。是一种纯技术手段。反而加重了系统负担。

假设有“分散数据” ，销售用户都分配到1节点，销售数据都在1节点，生成用户都分配到2节点，生产数据都在2节点，那么，cache fusion的工作量就会急剧减少，从根本上解决了性能问题

此种方法可使用那个service分散负载，即，吧每个service固定在某些rac节点上。

需要开发人员和dba合作。了解业务特点。

如果使用service方法，客户端的tnsnames配置中，

connect\_data 需要使用 service\_name

TEST =

  (DESCRIPTION =

    (ADDRESS\_LIST =

      (ADDRESS = (PROTOCOL = TCP)(HOST = 172.16.1.19)(PORT = 1521))

    )

    (CONNECT\_DATA =

      (SERVICE\_NAME = orcl)

    )

  )

=============================

》 测试loadbalance

listener 无论吧用户路由到本地或远程，本地listener日志中有 establish字样，没有 instance\_name字样，如果连接请求是有远程listener装发过来，则日志条目中有establish字样，和 instance\_name字样

shell脚本

$ more testlb.sh

#!/bin/bash

i=1

while [ $i -le 20 ]; do

      i=$(($i+1))

      sqlplus system/oracle@orcl @/tmp/testlb.sql

done

$ more testlb.sql

select instance\_name from v$instance;

exit

执行测试脚本

sh testlb.sh > lb.log

$ grep orcl1 lb.log |wc -l

$ grep orcl2 lb.log |wc -l

<https://docs.oracle.com/cd/E11882_01/server.112/e10803/config_cw.htm#HABPT4860>

Understanding and Troubleshooting Instance Load Balancing (Doc ID 263599.1)