

ORACLE 10G RAC
CLUSTERWARE 内部原理
分析

BY SHOUG. 章左中

SH'OUG

SHANGHAI ORACLE USERS GROUP

上海ORACLE用户组

How to Find SHOUG?



The image shows a Google search interface. The search bar contains the text '上海oracle用户组'. Below the search bar, there are navigation tabs for 'Web', 'Images', 'Maps', 'Shopping', 'More', and 'Search tools'. The 'Web' tab is selected. The search results show 'About 5,960,000 results (0.36 seconds)'. The first result is titled '上海Oracle用户组| SHOUG, 走近全系Oracle技术和数据库专家' with the URL 'www.shoug.info/'. Below the title, there is a snippet of text: 'SHOUG的全称是ShangHai Oracle Users Group, 中文为上海Oracle用户组。 SHOUG的成员仅仅局限于上海地区吗? 上海是国际化大都市, 我们将以上海为中心, ...'. The second result is titled 'Oracle 12c新特性- ORACLE数据库数据恢复、性能优化、故障诊断来 ...' with the URL 'www.askmaclean.com/archives/.../oracle/oracle-12c'. Below the title, there is a snippet of text: 'Feb 26, 2013 - 《Oracle 12c新特性》-作者: Maclean Liu, 首发于Ask Maclean中文Oracle博客. ... 手机: 13764045638, ORA-ALLSTARS Exadata用户组QQ群:23549328 ... Database 12c进入release发布的倒计时, 可能在今年7月在上海举行 ...'. Both results include a 'Translate this page' link.

Google 上海oracle用户组

Web Images Maps Shopping More Search tools

About 5,960,000 results (0.36 seconds)

[上海Oracle用户组| SHOUG, 走近全系Oracle技术和数据库专家](#)
www.shoug.info/ Translate this page

SHOUG的全称是ShangHai Oracle Users Group, 中文为上海Oracle用户组。 SHOUG的成员仅仅局限于上海地区吗? 上海是国际化大都市, 我们将以上海为中心, ...
You visited this page on 5/20/13.

[Oracle 12c新特性- ORACLE数据库数据恢复、性能优化、故障诊断来 ...](#)
www.askmaclean.com/archives/.../oracle/oracle-12c Translate this page

Feb 26, 2013 - 《Oracle 12c新特性》-作者: Maclean Liu, 首发于Ask Maclean中文Oracle博客. ... 手机: 13764045638, ORA-ALLSTARS Exadata用户组QQ群:23549328 ... Database 12c进入release发布的倒计时, 可能在今年7月在上海举行 ...
You've visited this page 4 times. Last visit: 4/25/13

Oracle 10g RAC Clusterware 内部原理分析

阅读本文之前需要了解 10gcrs 的基本结构：

可参考文档

Oracle Clusterware: Components installed. (Doc ID 556976.1)
How To Manually Start Oracle CRS Clusterware for Debugging Purposes [ID 758183.1]

首先介绍 crs 启动过程以及注意点：

1、init.crs 通过判断调用者来判断是自动启动还是手工启动

```
if [ "$CMD" = "init.crs" ]; then
    $LOGMSG "Oracle Cluster Ready Services starting by user request."
    $ID/init.cssd manualstart
else
    $ID/init.cssd autostart
```

当调用者为 rc 时，跳转到 init.cssd autostart

当调用者为 init.crs 时，跳转到 init.cssd manualstart

这里需要说明的是，我们平时启动 crs 用的命令是 crsctl start crs，为什么判断调用者是用 init.crs 来判断，其实 crsctl start crs 命令其实就是调用 init.crs start，执行 init.crs start 同样可以启动 crs。

2、scls_src 控制文件

These files are used to control some aspects of Oracle Clusterware like:

- enable/disable processes from the CSSD family (Eg. oprocd, oslsvmon)
- stop the daemons (ocssd.bin, crsd.bin, etc).
- prevent Oracle Clusterware from being started when the machine boots.
- etc.

In a Linux installation those files are located as follows:

```
[oracle@mbrac1 scls_scr]$ ls -lR
total 4
drwxr-xr-x 4 root root 4096 Oct 28 10:17 mbrac1
./mbrac1:
total 8
drwxr-xr-x 2 oracle root 4096 Oct 28 10:17 oracle
drwxr-xr-x 2 root root 4096 Oct 28 10:19 root

./mbrac1/oracle:

total 4
-rw-r--r-- 1 oracle root 7 Oct 28 10:32 cssfatal

./mbrac1/root:
total 12
-rw-r--r-- 1 root root 39 Oct 28 10:19 crsdboot
-rw-r--r-- 1 root root 7 Oct 28 10:17 crsstart
-rw-r--r-- 1 root root 39 Oct 28 10:17 cssrun
-rw-r--r-- 1 root root 0 Oct 28 10:19 noclsmon
-rw-r--r-- 1 root root 0 Oct 28 10:19 nooproc
```

关于控制文件如何控制 crs 的启动，将在后面详细的描述，需要注意的是，控制文件的目录和文件权限是不能随意修改的，如不正确可能导致启动 crs 的过程中不断的重启机器。

3、cssd 如何启动

init.cssd 首先会判断控制文件 crsstart 内容是否为 enable，如果是 enable 会调用 init.cssd manualstart,将`who -b`写入 cssrun 文件，如果为 disable，将 norun 写入控制文件 cssrun

init.cssd fatal 被 inittab 调用启动，每 RUNRECHECKTIME 时间检查一次 cssrun，通过控制文件 cssrun 文件的内容来判断是否启动 css，当 cssrun 内容为 startstack 或`who -b`时候就会打开开关。

所以：

当 cssrun 被写入 norun 之后，status 返回为非 0，导致 runcheck 直接退出，无法执行后续启动进程的操作

当 cssrun 被写入`who -b`之后，status 返回为 0，init.cssd fatal 会去调用 runcheck，
check whether basic prerequisites are met and permit the clusterware to start，如出现错误，在
/tmp 目录下产生文件名类似 crsctl.xxxx 日志，也可通过 crsctl check boot 命令检查未满足的
basic prerequisites，当所有前提条件都满足后，启动 cssd Daemons

以下是 10.2.0.3 的脚本内容：

```
'autostart')
    # Check to see whether the user has enabled automatic startup.

    # Is the flag file still there?
    if [ -r $AUTOSTARTFILE ]
    then
        case `SCAT $AUTOSTARTFILE` in
            disable*)
                $LOGMSG "Oracle Cluster Ready Services disabled by administrator."
                $ID/init.cssd norun
                ;;
            enable*)
                # Continue on to the next check
                $LOGMSG "Oracle Cluster Ready Services starting up automatically."
                $ID/init.cssd manualstart
                ;;
            *)
                $LOGERR "Oracle Cluster Ready Services disabled by damaged install."
                $LOGERR "    Unexpected settings found in $AUTOSTARTFILE."
                $ID/init.cssd norun
                ;;
        esac
    else
        # If the SCR directory does not exist, then either the user changed
        # their hostname, or they wiped sections of the disk. Give them a
        # slightly less frightening message.
        if [ ! -d "$SCRDIR" ]
        then
            $LOGERR Oracle Cluster Ready Services startup disabled.
            $LOGERR    Did the hostname change? Could not access $AUTOSTARTFILE.
        else
            $LOGERR Oracle Cluster Ready Services is disabled by damaged install.
            $LOGERR    Could not access $AUTOSTARTFILE.
        fi

        $ID/init.cssd norun
    fi
    exit 0;
    ;;
```

4、crsd 启动

init.crsd 根据控制文件 crsdboot 来判断是重启 crs 还是 crsd.bin 进程异常 (只重启 crsd.bin 进程)

当 crsdboot<>`who -b`时候，调用 crsd.bin reboot，会启动 crsd.bin 进程，并会启动资源

当 crsdboot=`who -b`时候，调用 crsd.bin restart，只会重启 crsd.bin 进程不会去启动资源

其次介绍 crs 如何维护 Daemons 进程异常：

很多人都知道，当除了 crsd.bin 进程之外的 Daemons 异常 abort 时 (比如 kill)，crs 都会自动重新启动，当 crsd.bin 进程异常 abort 时，主机重启。

是哪些进程维护的这些功能呢，其实就是 **init.* scripts**：

```
init.ocssd
init.crsd
init.evmd
```

而我们在平时起停 crs 时，特别是比较低的版本时，是否出现过 crs 无法正常起停，而又不能像起停数据库那样直接 kill 进程：

kill 非 ocspd.bin 进程会自动重启

kill ocspd.bin 进程会重启机器

而在了解 crs 如何维护 Daemons 进程异常后，我们可以通过这样的方式：

- 1、 注释掉 inittab 中 init.*scripts 的入口
- 2、 Kill 掉 init.*scripts 等进程
- 3、 然后再 kill crs 进程，crs 进程就能全部正常停掉
- 4、 然后就可以正常启动 crs

作者个人简介

章左中从2011年至今就职于甲骨文(中国)软件系统有限公司 Oracle 原厂, 为客户提供 Oracle 数据库性能调优、问题诊断、最佳实践交流以及专题分析等 Proactive&Reactive 服务。

章左中拥有近 7 年丰富的 IT 行业从业经历, 并具有大型软件项目的软件开发、系统运维、项目管理和咨询顾问等不同角色的工作经验。

特别优势: 经过不同领域(制造业、电信、咨询顾问)的项目实施以及不同工作职位角色(软件开发、系统运维、咨询顾问)的多年的经验积累, 能够全方位多角度的看待问题, 对用户的支持会更全面有效。

作为技术工程师, 有着多年的 Oracle 数据库维护和支持经验(多版本: oracle 8i、9i、10g、11g; 多产品: RAC、PL/SQL、SQL tuning、Backup & Recovery); 理论和实践积累丰富, 行业知识面宽广; 同时具有良好的方案撰写能力。

作为咨询顾问, 能够有效和客户交流, 倾听客户要求, 了解客户需求, 学习和创新能力强。

章左中具有丰富的专业经验和素质, 耐心细致负责的工作风格, 良好的计划组织和协调能力, 能够为客户提供及时、专业、有效的 Oracle 咨询实施服务。