**CSS 功能简介**

By: [Allen Gao](https://blogs.oracle.com/author/allen-gao)

最近处理了一些和CSS(Cluster

Synchronization Service)相关的问题，进而对CSS 的功能有了一些粗浅的认识，尤其是GM(Group Management)部分的功能，在这里和大家分享一下，希望对我们今后处理CSS 问题和理解CSS 这个oracle

集群的核心组件有所帮助。

CSS的功能主要包括节点管理（Node Management，以下简称NM）和组管理（Group Management，以下简称GM）两部分，都是由守护进程ocssd.bin 来实现的，这是个多线程的进程（我们可以通过命令pstack获得更多线程的信息，本文并不会详细的介绍每个线程的功能）。

首先，NM主要负责集群中节点的管理。集群是由若干个节点构成的，NM要解决的问题就是如何让集群的各个节点处于一致的状态，以及，当集群成员发生改变时，如果维护集群的一致性。所以，NM要实现下列的功能：

1. 节点间的网络心跳：通过集群私网，每一个节点都定期向其他所有的节点发送网络心跳，以便确认节点之间的通信和健康性。

2. 磁盘心跳：只有网络心跳是不够的，因为，一旦出现了网络问题，当节点间互相无法发现对方的网络心跳时，无法判断那一个（些）节点的状态正常，哪一些不正常，所以，我们还需要磁盘心跳。每个节点会定期向表决盘（VF）中注册本地节点的状态信息。这样，当网络心跳出现问题时，我们就可以基于表决盘中的信息，判断节点的状态并作出正确的决定，并防止脑裂（split brain）的发生。

3. 集群节点重新配置：ocssd.bin会每隔一段时间都对集群中的每个节点的最后一次网络心跳和磁盘心跳进行判断，以决定是否需要对集群进行重新配置（也就是说，是否有某一个/些节点需要被evict掉）。当然，当一个节点加入集群时，集群的其他节点需要收到集群重新配置的信息，得到最新的节点成员列表。同样的道理，当集群的某一个节点离开集群时，其他节点需要知道最新的成员列表。当然，无论是节点加入或者离开集群，都会有一个RMN(reconfig)节点负责集群的重新配置，并 知集群中的其他节点。

注意：以上内容是针对cssd.bin的基本功能进行的介绍，对已11gR2的新特性，如：本地心跳，cssdagent/cssdmonitor, rebootless restart,在本文并不包括。如果您对这些新特性感兴趣，请参考之前的文章。

 接下来，我们介绍一下GM这部分的功能。在一个集群中，除了节点之外，还有很多资源是需要管理的，（当然，主要的资源管理工作是由crsd.bin来完成的）而这些资源会作为一个一个的组（在css的世界里，我们把组称为grock）注册到css上，每一个组会由若干个成员构成，而且每一个组或者组中的一些成员都要向外共享一些信息。例如，集群中的每一个数据库都会作为一个组注册到css上，而这个组的主要的成员就是lmon进程（这就是为什么，有的时候我们会发现lmon进程有的时候会报ora-29701错误的原因。）,当一个进程被启动的时候，本地实例的lmon进程需要把自己注册到css对应的组当中，它才能够知道，这个数据库同时又多少个实例在运行，每个实例运行在哪一个节点，对于这种行为，我们一般称之为组内共享。

再举一个例子，当ASM实例启动之后，ASM对应的组会把自己的信息共享出去，以便数据库实例能够发现运行的ASM实例，并进行通信，对于这种行为，我们一般称之为组间共享。如果一个组的所有成员都来自于同一个节点，我们称之为本地组；如果组中的成员来自不同的节点，我们称之为全局组。

于NM类似，当组中的某一个成员离开或者加入组的时候，也会有一个master节点（一般称这个节点为GM master）来完成组成员的重新配置。另外，组管理中另一个重要的功能就是fencing,也就是说，当组中的某一个成员离开的时候（例如:这个进程被terminate掉），css 必须确保这个进程在OS 级别已经离开，而且进程pending的所有I/O 操作都被清理掉，因为对于集群来说，我们需要在节点间保证I/O的一致性，这也是我们经常在很多文档当中会讲到，只有I/O capability的进程才会注册到css的原因之一。

最后，由于CSS的首要功能是构建集群和确保集群的一致性，所以，当集群的节点数量发生改变的时候（节点加入集群，或者离开集群），css 首先会进行NM层面的重新配置，之后才能进行GM层面的重新配置，这个顺序是不能颠倒，也不能同时进行的。这也从某种程度上解释了，为什么我们会看到，当ocssd.bin 没有完成工作的时候，crsd.bin是不能启动所管理的资源的。

以上的内容希望对大家理解CSS这个oracle集群核心进程的功能有些帮助。在今后的文章中我们会尝试通过一些具体的例子来对上面的内容进行解释。

<https://blogs.oracle.com/database4cn/css>