19ᶜ ORACLE®
Database

Oracle Clusterware –
What's New with Oracle Clusterware

ORACLE®

Table of Contents

Introduction

Oracle Clusterware enables the clustering of otherwise independent servers so that they co-operate as a single system. As a cluster, these servers provide the integrated foundation to enable high availability and scalability for Oracle Real Application Cluster (RAC) databases and user applications.

The cluster of servers is coordinated via the Oracle Clusterware, with cluster resources made available as required in support of the high availability requirements of the Oracle RAC databases and applications running on the cluster, on one or more of the clustered servers, or nodes. Introduced in Oracle 10g Release 1, Oracle Clusterware has evolved and broadened its capabilities to meet the demand for a more versatile and capable infrastructure, while automating much of the co-ordination of events and actions behind the scenes.

Oracle Clusterware 12c Release 2 provided new features and enhancements aimed at reducing the effective cost of ownership, especially with regards to larger cluster estates. The introduction of Oracle Clusterware 18c continued that enrichment, focusing primarily on ease of use and the reduction of operating effort.

Now, with the introduction of Oracle Clusterware 19c, deployments are simpler still, as are upgrades and patching, and further refinements have been added for established functionality.

What's New with Oracle Clusterware

The focus for Oracle Clusterware 19c is simplifying the deployment process and improving the functionality provided in earlier releases.

The most notable change is removal of Leaf nodes from the Oracle Flex Cluster architecture. This had been announced in the previous release, and is now in effect. Customers upgrading their Grid Infrastructure to 19c will be required to follow a simple procedure to convert their Leaf nodes to Hub nodes prior to the upgrade. This procedure will also accommodate any read-mostly database instances that existed on those Leaf nodes. After the upgrade, all nodes in an Oracle Flex Cluster will be designated as Hub nodes, with complete access to shared storage.

This step forward simplifies how customers deploy and manage mixed-load clusters, in that they no longer have to worry about which cluster nodes were Leaf nodes and which were Hub nodes. And, as there will no longer be Leaf nodes, some Flex Cluster deployments may no longer need GNS support (that had been required for Leaf node deployments).

Another features added to improve the upgrade/patching experience is the Zero-Downtime Grid Infrastructure Patching utility. This new feature enables cluster administrators to patch the Grid Infrastructure stack without impacting the database instances and user sessions on any of the cluster nodes.

Previously, Oracle Clusterware 18c introduced the following features & enhancements:

» **Cluster Domain architectures enhancements**:

  » *Conversions*,

    ○ From Standalone Cluster to Database Member Cluster, and

    ○ From one storage access configuration to another for Database Member Clusters.

  » *Oracle ACFS Remote Service*: enable Oracle ACFS file systems on local Member Clusters utilizing remotely hosted ASM storage services on the Domain Services Cluster.

  » *Single network support for Application Member Clusters*.

» **Shared SCAN** – the optional election of SCAN VIP's on one cluster to act as the SCAN VIP's for other clusters, reducing the number of IP addresses required by using a single SCAN VIP setup for multiple clusters.

» **Node VIP-less Clusters** – optionally, deploy new clusters without Node VIP's, thus freeing up cluster resources and reducing the number of IP addresses required.

» **Cross-Cluster Dependency Proxies** – support for Clusterware resource dependencies that span clusters.

Zero-Downtime Oracle Grid Infrastructure Patching

The functionality for the Zero-Downtime Oracle Grid Infrastructure patching capability was introduced in order to remove all impacts to the user sessions while applying the Grid Infrastructure patches to each of the cluster nodes. As the Grid Infrastructure stack of a node is being patched, using an out-of-place patching approach, the local database instances are not impacted (ie. they stay fully operational), meaning that the user sessions connected to those instances are also not impacted. Then, once the new old GI stack is stopped and the new version started, all without impact to the local database instances. (Note that this capability will initially apply only to RAC clusters of two or more active nodes, and for one-off patches only.)
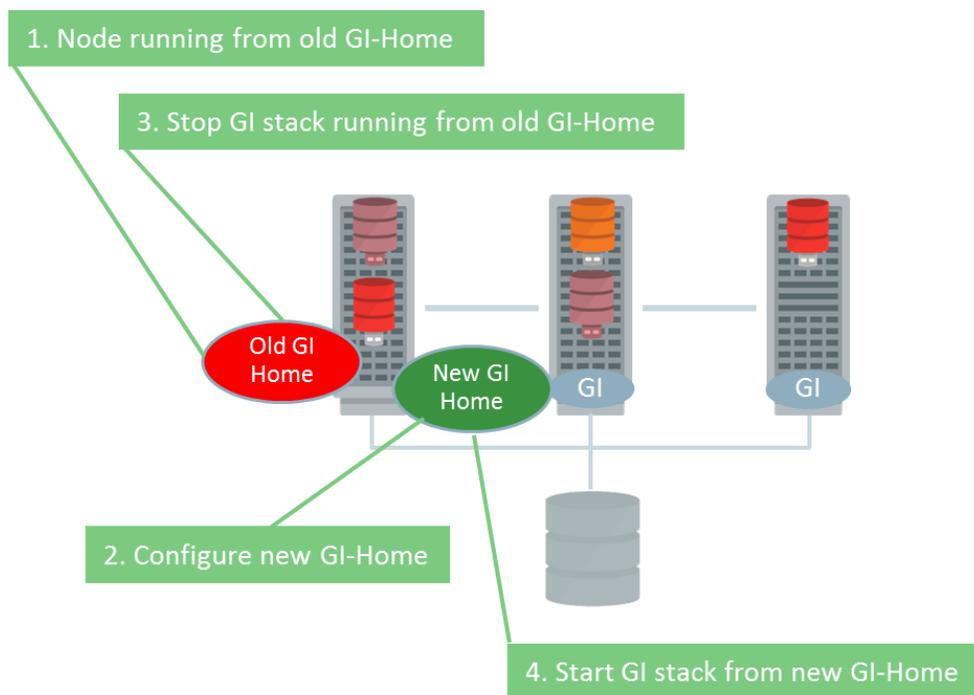

Figure 1: Zero-Downtime Oracle Grid Infrastructure Patching

Oracle Cluster Domains

The Cluster Domain architecture was introduced in Oracle Clusterware 12c Release 2 as an optional deployment model enabling simpler, easier deployments, reduced storage management effort and performance gains for I/O operations, especially useful when managing larger estates of Oracle Clusters. Oracle Clusterware 18c builds on this architecture by adding support for Oracle ACFS file systems mounted locally on Member Clusters, though they are utilizing remote ASM storage, and providing support for converting existing Standalone Clusters to Member Clusters.
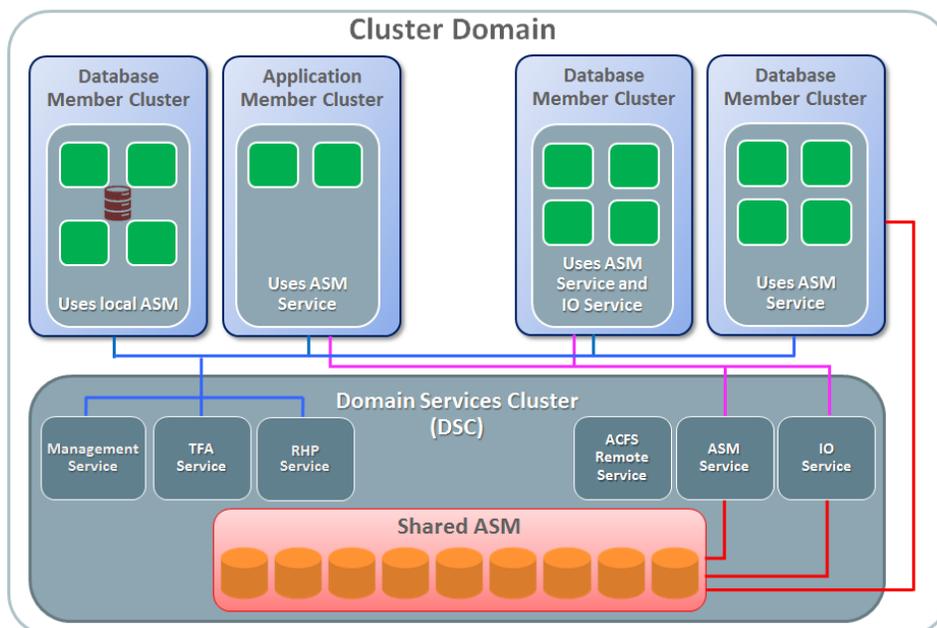
Figure 2: Cluster Domain Architecture

Conversions

When the Cluster Domain architecture was introduced in Oracle Database 12c Release 2, it was presumed that adoption of the new architecture would be for new deployments only, thus requiring support only via the Oracle Universal Installer (OUI). Now, as they are being adopted, the demand for conversions has been addressed with new functionality added in Oracle Database 18c.

Convert from Standalone Cluster to Member Cluster

With Oracle Clusterware 18c, the first of the conversion capabilities had been introduced, whereby a Standalone Cluster can be converted to a Member Cluster with local ASM. The conversion effectively removed the requirement for a local Grid Infrastructure Management Repository (GIMR), thus freeing up local shared storage, and moving it to the Domain Services Cluster by subscribing to the Cluster Domain's Management Service.
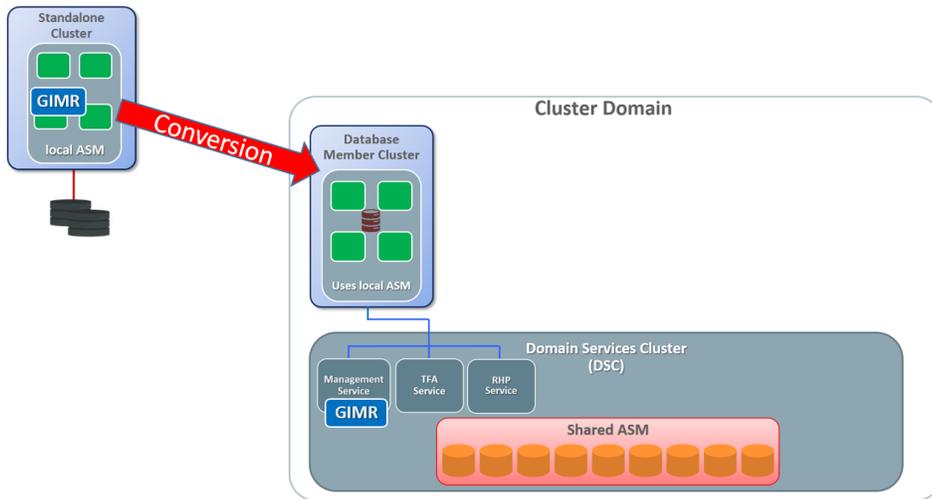
Figure 3: Convert Standalone Cluster (local GIMR) to Database Member Cluster (GIMR offloaded to DSC)

Storage Access Conversion for Database Member Clusters

Introduced with Oracle Clusterware 18c was the ability to convert the storage access methods for established Database Member clusters that used remote ASM-managed storage. This conversion allowed the Database Member Clusters to convert from using the IOServer on the Domain Services Cluster (DSC) to accessing the storage directly, or the reverse.
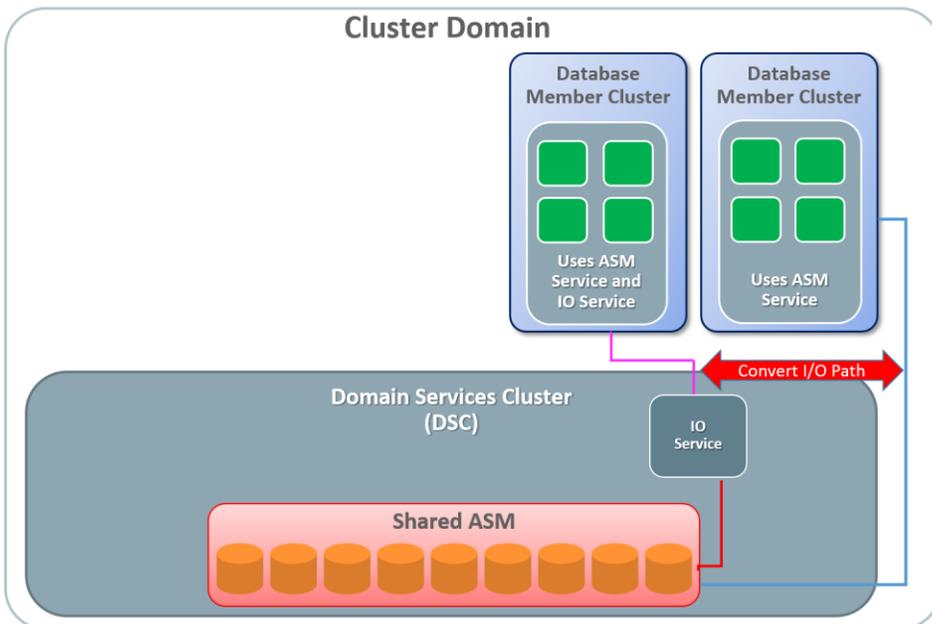

Figure 4: Storage Access Conversion for Database Member Clusters using Remote ASM

This capability further enhances the flexibility and adaptability of the Cluster Domain, without requiring significant downtime. The conversion requires that the underlying infrastructure be changed to support the new access paths, which will require a short outage on the Database Member Cluster.

Oracle ACFS Remote Service

Oracle ACFS 18c introduced support for native Oracle ACFS functionality on database member clusters and application member clusters. Oracle Member Clusters with no local shared file system can leverage all Oracle ACFS feature and functionality through the Oracle ACFS Remote Service hosted on the Domain Services Cluster.

In using this service, the Member Clusters can have shared local storage, though it is using remote storage, without having to manage NFS mounts. The Oracle ACFS file system can be used for flat files, logs, trace files, extracts and even database files.
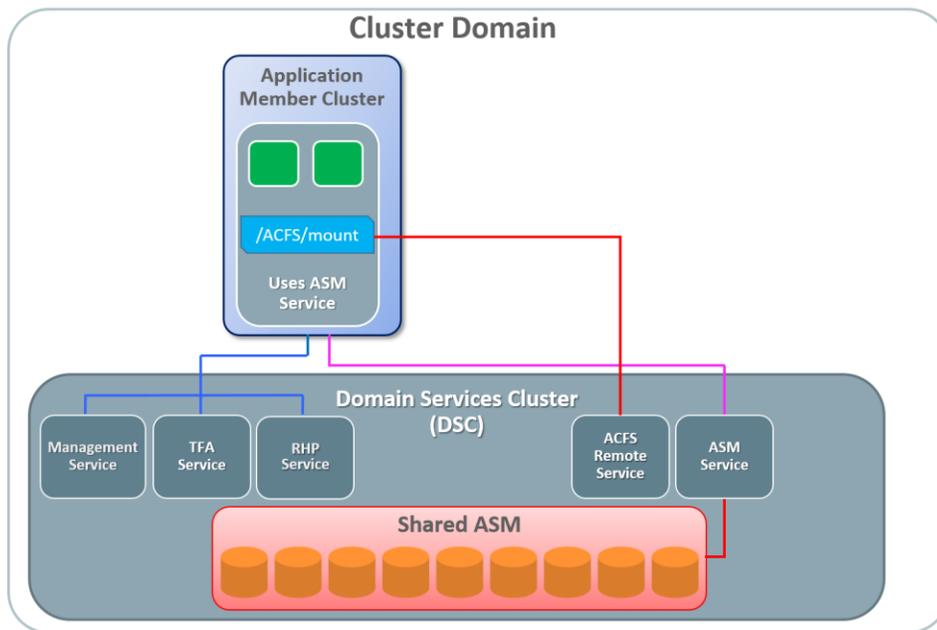

Figure 5: Oracle ACFS Remote file system as mounted on Application Member Cluster

Single Network Support for Application Member Clusters

For Application Member Clusters there is no longer a requirement for a dedicated private network for the Cluster Interconnect. During deployment, if the option to provide the private network information is left blank (the default), then all Cluster Interconnect traffic will be routed across the public network. This will significantly ease the deployment of Application Member Clusters.

Specify that the interconnect traffic should use the Public Network during deployment:
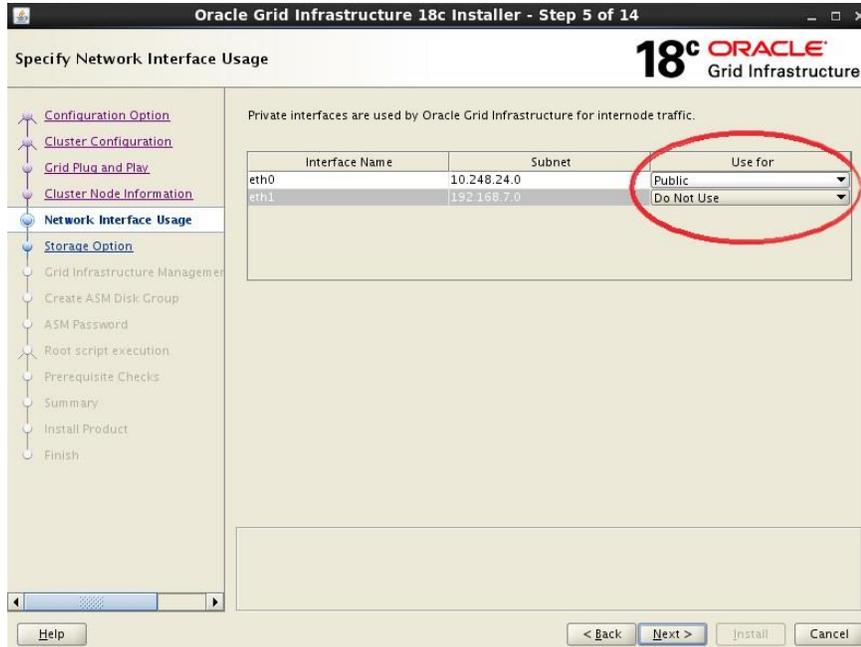

Figure 6: Select only the Public Network during Application Member Cluster deployment

Interconnect traffic for Application Member Clusters is limited to cluster health and membership messages. There is no Cache Fusion traffic. Thus, enabling the use of a single network (i.e. the Public network) for Application Member Clusters is a relatively simple enhancement. However, this is not the case for clusters that host RAC databases, where Cache Fusion makes up a considerable percentage of the traffic across the interconnect network. The performance and responsiveness of RAC databases depend upon the isolation of the Cache Fusion traffic and its reliability. In addition, by using a private interconnect network, there is not the requirement to encrypt that Cache Fusion traffic. This is not necessary for the interconnect traffic for Application Member Clusters.

Shared SCAN

Shared SCAN had been introduced in 18c to primarily reduce the number of IP-addresses required during cluster deployments. With a Shared SCAN VIP configured on one cluster (called the 'Shared SCAN Cluster'), existing and new 18c and newer clusters can be configured to utilize this single SCAN VIP setup, rather than having them configured for each of these clusters (called 'Shared SCAN clients').

The Shared SCAN is setup exactly as a local SCAN is setup; in other words, an existing SCAN setup can be utilized as a Shared SCAN. What differs is how the connections are routed through this Shared SCAN to the Shared SCAN clients. The key difference is the location of the SCAN Listeners; with a local SCAN setup, they are configured on the nodes of the local cluster, but with a Shared SCAN they are configured on the Shared SCAN Cluster.

Database clients (users) connect via a Shared SCAN VIP to the respective SCAN Listener on the Shared SCAN Cluster. The SCAN Listener on the Shared SCAN Cluster would then redirect new connections to the registered endpoints of the service/database on the Shared SCAN client.

> **What is SCAN?**
>
> Single Client Access Name (SCAN) is a feature used in Oracle Real Application Clusters (RAC) environments that provides a single name for clients to access any Oracle Database running in a cluster. It was introduced in Oracle 11g Release 2 in order to insulate users from cluster node changes (i.e. adding or removing nodes).
>
> RAC Databases register with the SCAN LISTENERS (remote_listener), allowing the SCAN LISTENERS to redirect user connections to specific databases according to their connect string.
>
> For more information about SCAN, see:
>
> Oracle Single Client Access Name whitepaper

Configuring Clusters to Utilize a Shared SCAN:

» Setup Resources on Shared SCAN Cluster (for each of the Shared SCAN clients) – issued manually:

- Create SCAN Listener and ONS configurations for the new Shared SCAN client:
  - *srvctl add scan_listener -clientcluster <cluster_name>*
  - *srvctl add ons -clientcluster <cluster_name>*
- Note: There will be a SCAN Listener created for each Shared SCAN client. (For instance, if there are 6 database clusters configured to use a Shared SCAN Cluster, then there will be 6 SCAN Listeners configured on that Shared SCAN Cluster.)
- Create the credential file(s) for the Shared SCAN client:
  - *srvctl export scan_listener -clientcluster <cluster_name> -clientdata <filename>*
  - *srvctl export ons -clientcluster <cluster_name> -clientdata <filename>*
- Note: A single credential file can be designated for both the ONS and SCAN Listener. Also, if the Shared SCAN is configured on a Domain Services Cluster (DSC), then the SCAN Listener and ONS credentials will be included in the credential file for the Database Member Cluster that is being configured.

» Setup Resources and Connectivity on the Shared SCAN client – manually with these commands, or via OUI:

- Copy the credential file(s) from the Shared SCAN Cluster to the Shared SCAN client
- Create the local resources and save connectivity information for databases on the Shared SCAN client:
  - *srvctl add scan -clientdata <filename>*
  - *srvctl add scan_listener -clientdata <filename>*
  - *srvctl add ons -clientdata <filename>*

Simpler still, specify the Shared SCAN credential file during deployment:
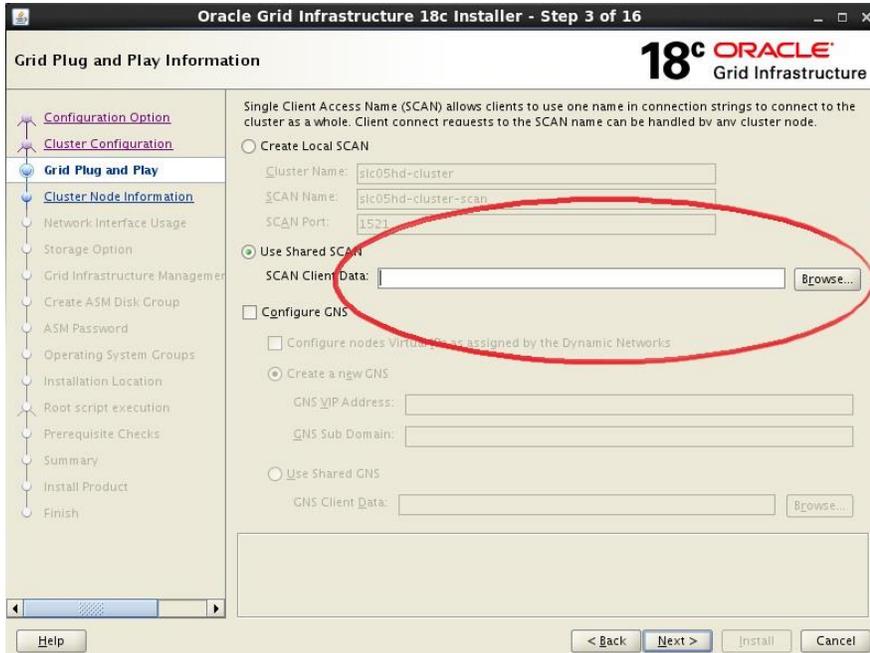
Figure 7: Enter File Name for the SCAN/ONS Credential File

Database clients (users) would connect to a specific database on a Shared SCAN client as they would to a database on a cluster with locally configured SCAN. Their connection string would specify the SCAN Name, the port and the service, exactly as they do with a local SCAN configuration. Note that it is the port and the service that differentiates the SCAN Listeners on the Shared SCAN Cluster.
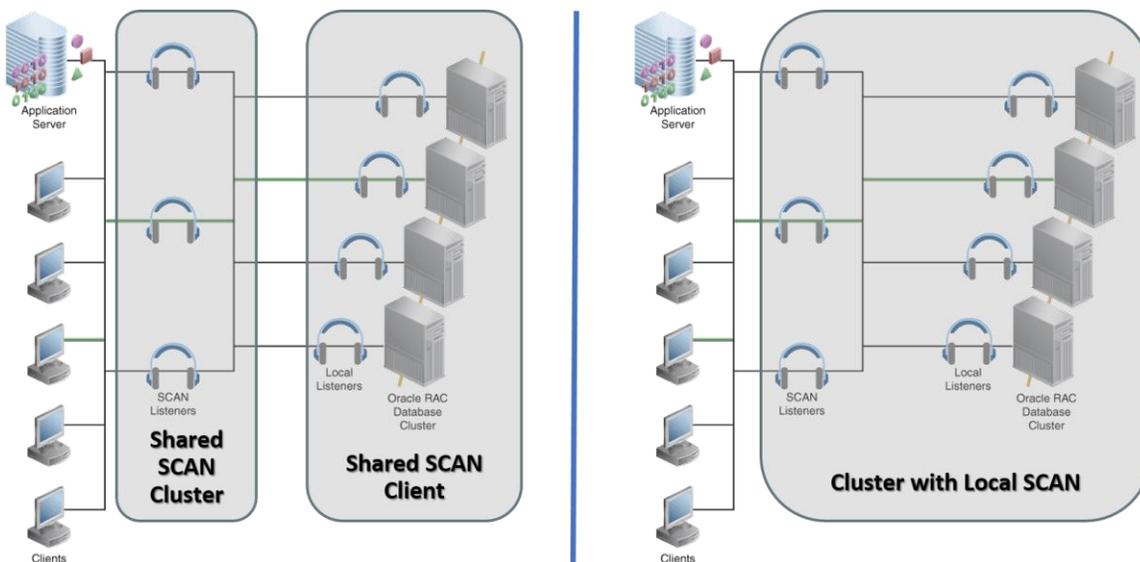
Figure 8: Database User Connections with Shared SCAN and Local SCAN

Shared SCAN implementations still require three IP-addresses, exactly as a locally configured SCAN. Deploying more than three SCAN VIP's is not required. The Shared SCAN Client must belong to the same network sub-net as does the Shared SCAN Cluster with which it registers. Note also that for Cluster Domain implementations, the Shared SCAN will be hosted on the Domain Services Cluster.

Node VIP-less Clusters

Node VIP's have long been used to provide HA connectivity for database clients, but have no place in clusters that are not hosting Oracle databases. Such clusters, often referred to 'application clusters', rely on application VIP's for to provide fault tolerant connectivity to the applications or processes on the cluster nodes.

With the introduction of Cluster Domains in Oracle 12c Release 2, Application Member Clusters were the first instance of deployments that did not include Node VIP's.

With Oracle Clusterware 18c, Standalone Clusters without Node VIP's are an option during the deployment. The default is still to configure Node VIP's, as this preserves backwards compatibility (plus clusters upgraded from previous versions will, of course, have Node VIP's). This deployment option is recommended for clusters that do not host databases or for non-production deployments (in which failure management is not as critical as it might be for production systems).

During deployment, simply do not specify the virtual hostname:
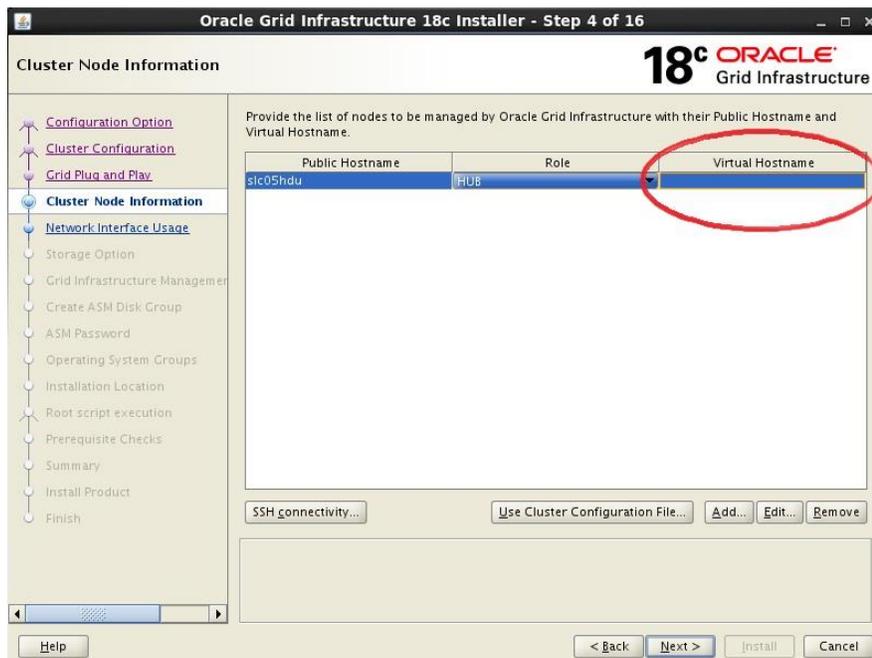
Figure 9: For Node VIP-less Deployments, omit the Virtual Hostname

Cross-Cluster Dependency Proxies

Cross-Cluster Dependency Proxies were created in Oracle Clusterware 18c to allow Clusterware resource dependency management to extend beyond the domain of a single cluster. This means that resources on one cluster may have dependencies upon the resources on other clusters.

The default deployment in Oracle Clusterware 18c applies to Cluster Domains and the Database Member Clusters that access the ASM-managed storage on the Domain Services Cluster. The databases on the Database Member Cluster are dependent upon the availability and accessibility of the ASM diskgroups on which their database data files are stored. This dependency is mapped using the Cross-Domain Protocol (CDP).
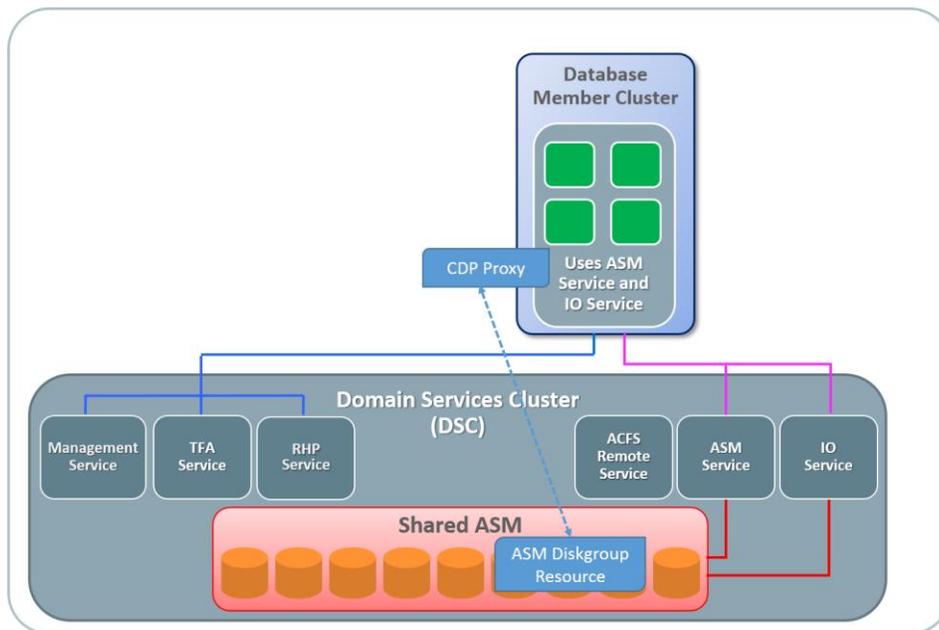
Figure 10: Cross-Cluster Dependency Proxy setup for Database Member Clusters

In addition, users may implement CDP for other uses. An example of this might be for an Application Member Cluster and its dependency upon the availability of a remotely hosted database. By creating a proxy for the remote database service, the startup of the Application Member Cluster could ensure that that remote database is also started.

CDP may also be used to monitor the status of entities in remote clusters. An example of this would be a monitoring process to keep track of which VM's in a remote cluster are active, allowing local processes to act on this information without having to establish connectivity to that remote entity.

Summary

Oracle Clusterware 12c Release 2 introduced improvements in manageability and deployment, and better mitigation and control for failure scenarios for Oracle Clusters. Oracle Clusterware 18c & 19c have continued that trend by promoting ease of use and the further reduction in operating effort.

Enhancements have been made to the Cluster Domain architecture. Introduced with Oracle Clusterware 12c Release 2, Cluster Domains are aimed at providing simplified, easier and faster management and deployment of multi-cluster infrastructures is key for larger cluster installations. New capabilities have been introduced in Oracle Clusterware to further encourage new and established customers to adopt this innovative cluster architecture

In addition, key new Oracle Clusterware 18c features allowed for the reduction of the number of IP-addresses during the deployment of new clusters, for database and non-database clusters, and the ability to manage Clusterware resource dependencies across clusters.

Now, with Oracle Clusterware 19c, the cluster topology options are further simplified with the desupport of Leaf nodes from Oracle Flex Clusters. This provides customers with greater flexibility and ease of deployment for their mixed-load clusters, while still supporting larger cluster configurations particularly for consolidated environments.

ORACLE®

Oracle Corporation, World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065, USA

Worldwide Inquiries
Phone: +1.650.506.7000
Fax: +1.650.506.7200

Integrated Cloud Applications & Platform Services

Oracle Clusterware 18c – Technical Overview
February 2019
Author: Ian Cookson
Contributing Authors: Markus Michalewicz