Shire Pharmaceuticals Deploys Oracle Grid Computing Model to Lower Costs and Increase Reliability

Advans IT Services, Inc. White Paper
September 2009

Shire IT Case Study / Oracle Grid Computing / Maximum Availability Architecture

Shire Pharmaceuticals Deploys Oracle Grid Computing Model to Lower Costs and Increase Reliability

Shire IT Case Study / Oracle Grid Computing / Maximum Availability Architecture: This Shire Pharmaceuticals ("Shire") case study describes the transformation of its database architecture by Advans IT Services, Inc. ("Advans"), a Westborough MA, solutions provider. In deploying Oracle's Grid Computing Model for the datacenter, Shire has adopted Oracle Real Application Clusters ("RAC") as its corporate standard. Also included is a uniform multi-site disaster recovery solution implementing DataGuard and a standardized backup and recovery strategy employing Recovery Manager ("RMAN"). Lastly, Oracle Enterprise Manager Grid Control ("Grid Control") provides a single point of management and monitoring for the entire database infrastructure.

Background

Founded in 1986, Shire plc is a global specialty biopharmaceutical company that focuses on attention deficit and hyperactivity disorder (ADHD), gastrointestinal (GI) and Genetic disorders. In 2008, Shire engaged Advans to standardize the database tier, for both hardware and software, across the entire company.

Advans utilized Oracle's Maximum Availability Architecture ("MAA") technologies and associated best practices to re-architect Shire's database infrastructure. By combining these best practices for achieving greater reliability, performance and flexibility with techniques for Oracle RAC server consolidation, Advans reduced the number of database servers by approximately 50%, while meeting or exceeding the response rate and uptime demanded by Shire.

Challenge

Shire wanted to standardize its databases tier and take advantage of technology trends in distributed and parallel processing to achieve better performance. Secondly, to reduce costs and maximize the utility of previous infrastructure investments in hardware, existing Wintel stand alone database servers were repurposed as Oracle RAC nodes running on Red Hat Enterprise Linux ("RHEL") 5 Advanced Platform.

Moreover, having outsourced the administration of its IT infrastructure to a third party, there was a need for streamlining the administration of IT assets to manage costs. Lastly, having added a datacenter recently, Shire wanted to standardize backup and recovery procedures and implement a comprehensive Disaster Recovery strategy for taking advantage of the multiple datacenters in the UK and the US.

Scalability and Reliability

As a growing global company, applications are now utilized over a larger part of the day. Coupled with a growing user base, the ability to add computing resources in real time to meet future demands became desirous. In addition to greater uptime, platform instability in some environments fueled the need for resiliency or failover in the database tier.

Standardization and TCO

Shire recognized that reducing operational complexity through standardization and decreasing the variety of technologies deployed in the datacenter is a means to lower its Total Cost of Ownership ("TCO") for IT assets. Combined with server consolidation, Shire sought to decrease its administration costs and gain other indirect savings, such as reduced power and cooling needs as well as optimizing the use of space in the datacenters, etc. This also included, as part of the drive to standardize, a strategy to automate the monitoring and management of the infrastructure through a single point. Simplifying the infrastructure serves to facilitate the deployment of new applications, since standard components are created with common methodologies thus reducing the risk and the time to deploy them.

As a biopharmaceutical company, in addition to the usual regulatory requirements of a public company, Shire must comply with a multitude of industry regulations in the US, EU and elsewhere, for the development, manufacture and distribution of its products. Compliance with these regulations extends not only to the products themselves, but also to documenting the installation, configuration and maintenance of IT assets. Auditing the operation of some components is required as well. Since IT components are standardized, compliance functions, which are often very time consuming, can be accomplished by common mechanisms. Likewise, securing these components can also be performed by a few standard best practices which easily extend across the entire infrastructure.

Solution

Implementing Oracle's Grid Computing Model to meet Shire's business needs, Advans utilized the following principals in designing the database architecture:

- Implement Maximum Availability Architecture for all technologies
- Virtualize the database tier so that additional computing resources can be added online
- Automate monitoring, alerting and reporting to reduce the need for manual intervention and to better manage the virtualized services
- Standardize the Hardware and Software platforms
- Provide for a comprehensive Disaster Recovery solution that takes advantage of Shire's three datacenters

RAC

The primary technology underlying this MAA implementation is Oracle RAC. Many of the servers utilized as RAC nodes were re-purposed WinTel servers. Considerable cost savings were achieved by utilizing existing Windows servers and re-purposing them with Red Hat Enterprise Linux 5 as the operating system. All of the servers are HP ProLiant DL580 G4 or G5 Server series. Typically they have four quad core CPU's and 32 GB of RAM

Clusters were originally built with two nodes, using OCFS2 for the clusterware and other database related files. In all, five OCFS2 volumes were created, a Voting Disk, an OCR Disk, an OCR Mirror disk, a volume for spfiles and dump files along with a volume for backup files. RMAN Backups were initially configured to disk but this was later modified to have RMAN backup to a Virtual Tape Library ("VTL").

Backup policies were created with Veritas NetBackup, which in turn executes RMAN scripts. While typically, NetBackup defines backups by server, to take advantage of RAC databases, a Virtual IP Address ("vip") should be specified in the policy. In this way, should a node be lost in a cluster, the vip can fail over to a surviving node and execute the backup. However, the backup volume on disk was retained for emergency

System Configuration

- **HP ProLiant DL580 G4 or G5 64-bit x86 Servers 32GB memory**
- **Oracle Database 10.2.0.4.0**
- **Red Hat Enterprise Linux 5**
- **Oracle Cluster File System 2**
- **Automatic Storage Management**

backups. Having this space readily available is also useful for creating RMAN clones, which was often employed to migrate databases. Both the Voting Disk and OCR Disk were mirrored to the OCR Mirror Disk using Normal Redundancy, thereby creating at least two copies of each.

Automatic Storage Management ("ASM") is used to manage database file storage, with Oracle Managed Files ("OMF"), simplifying administration and yielding better I/O performance. Creating pools of storage rather than using specific volumes abstracts the storage layer and allows ASM to detect and eliminate hotspots automatically. Two Diskgroups, DATA and FLASH, were created for storing data, archiving redo logs, etc., and to provide for a Flash Recovery Area.

As more servers were re-purposed, clusters were expanded from two to four nodes and then later to six nodes. In addition to providing better performance and reliability, the availability of greater computing resources in RAC clusters was also implemented as a means of server consolidation. Through this technique, Advans reduced the total number of Oracle database servers in Shire's datacenters by approximately 50%.

Server consolidation was achieved using the concept of an "Instance Group". In a six node cluster, as many as forty-five databases can be hosted. For example, servers were denoted as nodes 1 through 6. For each pair of nodes, fifteen databases are created with two instances each. Nodes 1 and 2, therefore, host databases one through fifteen in Instance Group #1. Node 1 has the instances numbered one and node 2 has the second instance. Nodes 3 and 4, in turn, host databases numbered sixteen to thirty, again with each database having two instances. Lastly, nodes 5 and 6 host databases thirty-one to forty-five. As a result, a six node cluster has the resources to host forty-five databases, each with two instances as a base configuration, (see Figure 1).

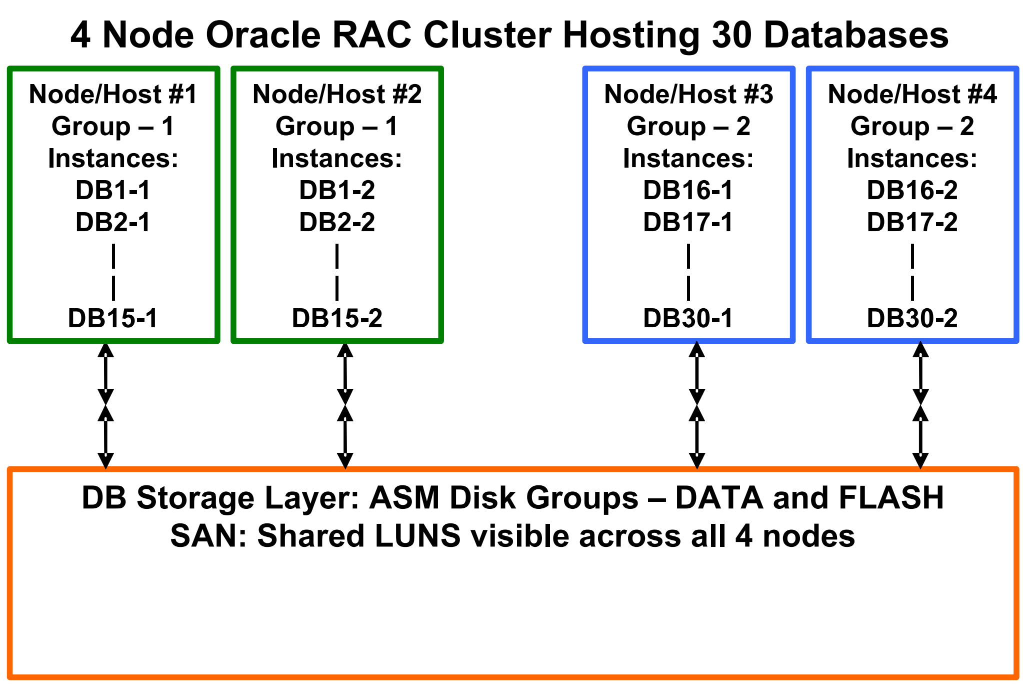

Figure 1, Instance Groups
Instance Group 1, Nodes 1 & 2 host databases 1-15, each with 2 instances
Instance Group 2, Nodes 3 & 4 host databases 16-30, each with 2 instances

The obvious constraint is physical memory. Advans observed that hosting 15 databases with a System Global Area ("SGA") up to 1.5 GB each was reasonable for an Instance Group. If databases required larger SGA's, then obviously, less databases could be hosted within an Instance Group. The same could be said for databases requiring more resources than two instances could provide. In that case, additional instances can be created on the other nodes of the cluster and this again reduces the total number of databases that can be hosted in the cluster. Advans found that for most of Shire's databases two instances were sufficient, and this technique provided a simple method for consolidating servers hosting small to medium-sized databases.

Using this technique, however, requires an adjustment to the ASM instance running on the cluster. By default, the parameters for sessions, processes and memory are set to relatively low values when an ASM instance is created. These parameters will be inadequate for the demands placed on the ASM instance when more than a few databases are hosted simultaneously. As a result, both the sessions and processes parameter were set to 1000 as a minimum, and the minimum size for the shared_pool_size parameter was 300 MB. Hosting databases in a cluster with large numbers of concurrent users may necessitate setting these values higher.

Another consideration must also be made for executing backups, when using the Instance Group technique. Since all databases do not have instances on all nodes of the cluster, a backup policy can not simply be created for the entire cluster. For instance, if the vip for node "2" in a six node cluster is referenced, then only the databases in the first instance group running on nodes "1" and "2" will actually get backed up. Therefore, when creating backup policies with third-party software, a vip must be specified for each instance group. In this example, three policies would be required for a six node cluster.

This configuration, however, provides for a great deal of flexibility since a database can "span" more than one Instance Group and Instance Groups of any number of nodes can be created. While most of the databases hosted in this environment performed very well with two instances, and actually better than they did on individual database servers, there will be occasions where more computing resources are required. If a database requires more instances, they can be quickly created on any of the other nodes in the cluster. Importantly, since additional nodes can be added, Shire did not have the need to "overprovision" its servers. Only the current computing capacity was originally created and when more resources were required, more nodes were added to the appropriate cluster.

Grid/RMAN

A dedicated two node RAC cluster was built for the purpose of hosting the RMAN catalogue and the Oracle Enterprise Manager. Both the RMAN catalogue and the Oracle Enterprise Manager Repository are RAC databases.

To achieve a Maximum Availability Architecture with the Grid Control, the Grid/RMAN cluster was placed behind a Server Load Balancer ("SLB") using the "Active/Active" configuration, (see Figure 2). Since both the Oracle Management Server ("OMS") and the Grid Control Repository database are version 10.2.0.4, the OMS and repository were installed on the same hardware. Not only is this a cost saving measure, but this eliminates network latency between the OMS tier and the repository tier, which can occur when installing the OMS on separate hardware from the database tier.

To achieve this, the OMS is installed with the option for using an existing database which is the only method for creating the repository as a RAC database. After the first OMS is installed, the install continues with the "Add Additional Management Service" option for the second OMS.

Agent data is uploaded to the virtual IP provided by the SLB. Since both OMS's point to the same database, the uploaded information is available to the Grid Console regardless of which OMS received the data from the agents.

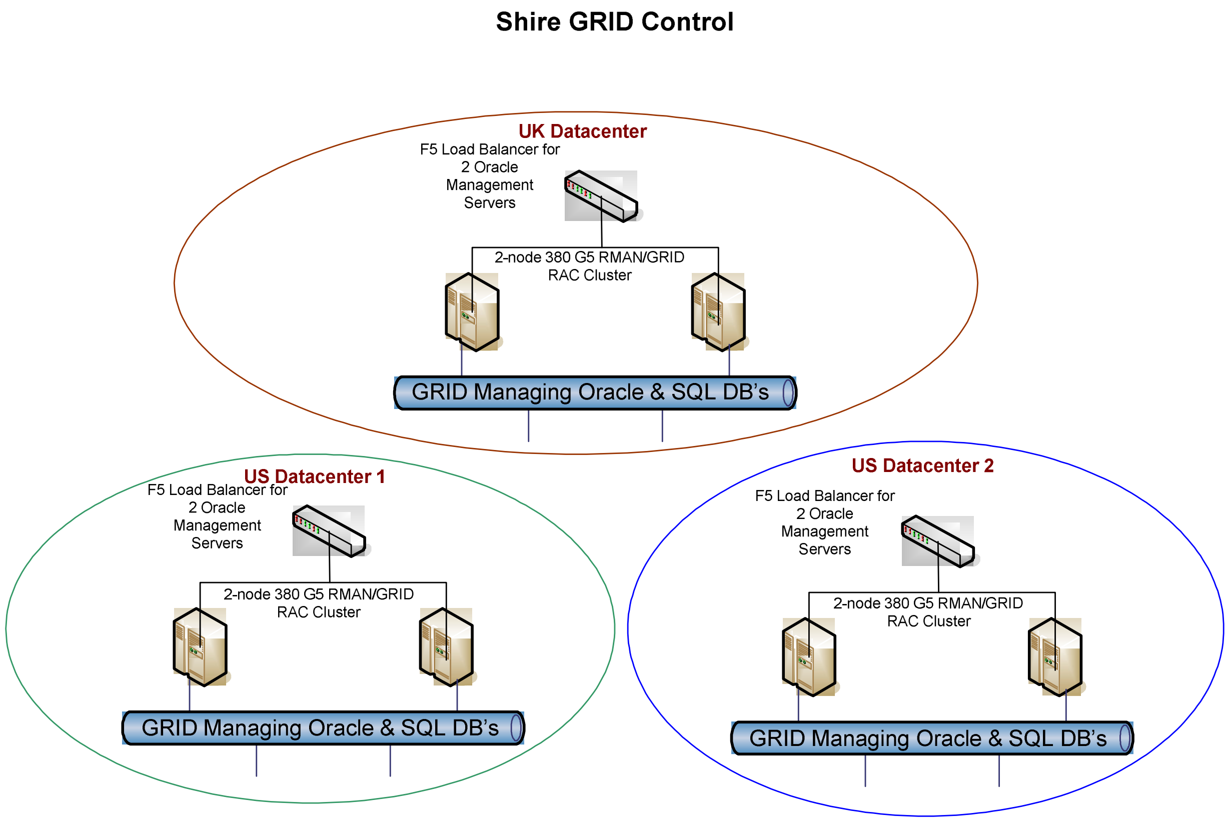

Figure 2, Grid Controls with Server Load Balancers

FDA CFR Title 21 Part II compliance for data protection and availability

Disaster Recovery

Shire maintains a datacenter at three sites, one in the UK and two in the US. Each datacenter is used to host primary databases as well as standby databases from another site's datacenter, (see, Figure 3). This topology obviates the need for a dedicated failover site. Furthermore, having failover sites aids in the compliance with US and European regulatory agencies. All standby databases are physical standby's where each site's primary databases failover to standby databases in the following rotation:

- US Datacenter 1 fails over to UK Datacenter
- UK Datacenter fails over to US Datacenter 2
- US Datacenter 2 fails over to US Datacenter 1

The redo logs are transported asynchronously in Maximum Performance Mode. While no Recovery Objective or Recovery Point has been specifically defined, the goal is to minimize the loss of data as much as possible. The decision for failover to the standby databases is a business one and that decision may be made for an individual database or for an entire datacenter, depending upon the circumstances.

Network Bandwidth

Inside the US
45 mb/sec (10 ms)

Across the Atlantic
34 mb/sec (90 ms)

Within the US, the network link for Data Guard communication is 45 mb/sec bandwidth with 10 ms RTT network latency. Across the Atlantic, the bandwidth is 34 mb/sec with 90 ms RTT network latency. Plans are underway to increase the amount of bandwidth across the WAN, although transporting redo logs in asynchronous mode has not been problematic.

The Grid Control proves to be very useful in two ways. First, redo transport and redo apply are monitored for any delays and that information is readily available. Delays in either process also trigger alerts. Secondly, the primary to standby switchover can be performed from the Grid Control Console.

The Maximum Availability Architecture for the Grid/RMAN cluster requires a SLB, but due to network constraints, one SLB could not span the entire network. It was necessary, therefore, to create a Grid/RMAN cluster at each site for monitoring local databases. A "global" Grid Control was considered to monitor all three sites, as having just one does simplify administration, but since it could not be configured as a MAA solution, that idea was abandoned.

Having only local Grid Controls, however, did have implications for the disaster recovery strategy. In order to use the Grid Control to monitor the application of the primary database's redo logs to its standby, the hosts for both the primary and standby databases must upload their data to the same Oracle Management Server. Since the primary and standby databases were in different locations, the application of logs to the standby databases could not be monitored.

A work-around was developed whereby two Grid Agents were installed on all the database hosts. The first agent was pointed to the local Grid Control to upload data for the primary databases. The second agent was pointed at the remote site hosting the primary databases corresponding to the local site's standby databases. As US Datacenter 1 is the failover site for US Datacenter 2, database hosts in US Datacenter 1 would host local primary databases and also standby databases corresponding to primary databases hosted in US Datacenter 2.

One agent on the RAC nodes in US Datacenter 1 pointed to the US Datacenter 1 Grid Control for monitoring primary databases. A second agent is pointed to the Grid Control in US Datacenter 2. On the US Datacenter 2 Grid Control then, only targets for standby databases hosted in US Datacenter 1 are added to the target list of local primary databases. So in this example, the US Datacenter 2 Grid Control would have targets for all the primary databases in US Datacenter 2, and their corresponding standby databases hosted in US Datacenter 1. To reduce network traffic and the load on the Grid Controls, only standby databases were monitored remotely in this manner.

In this way, every Grid Control could monitor all the local primary databases and their related remote standby databases. Obviously, DBA's would be required to log into three different Grid Control Consoles to administer the entire database tier. While this is not a recommended configuration, no anomalies were observed in having two agents installed on each host, and of course, agents were installed under different Agent Homes.

Shire Global Linux Oracle RAC Environment

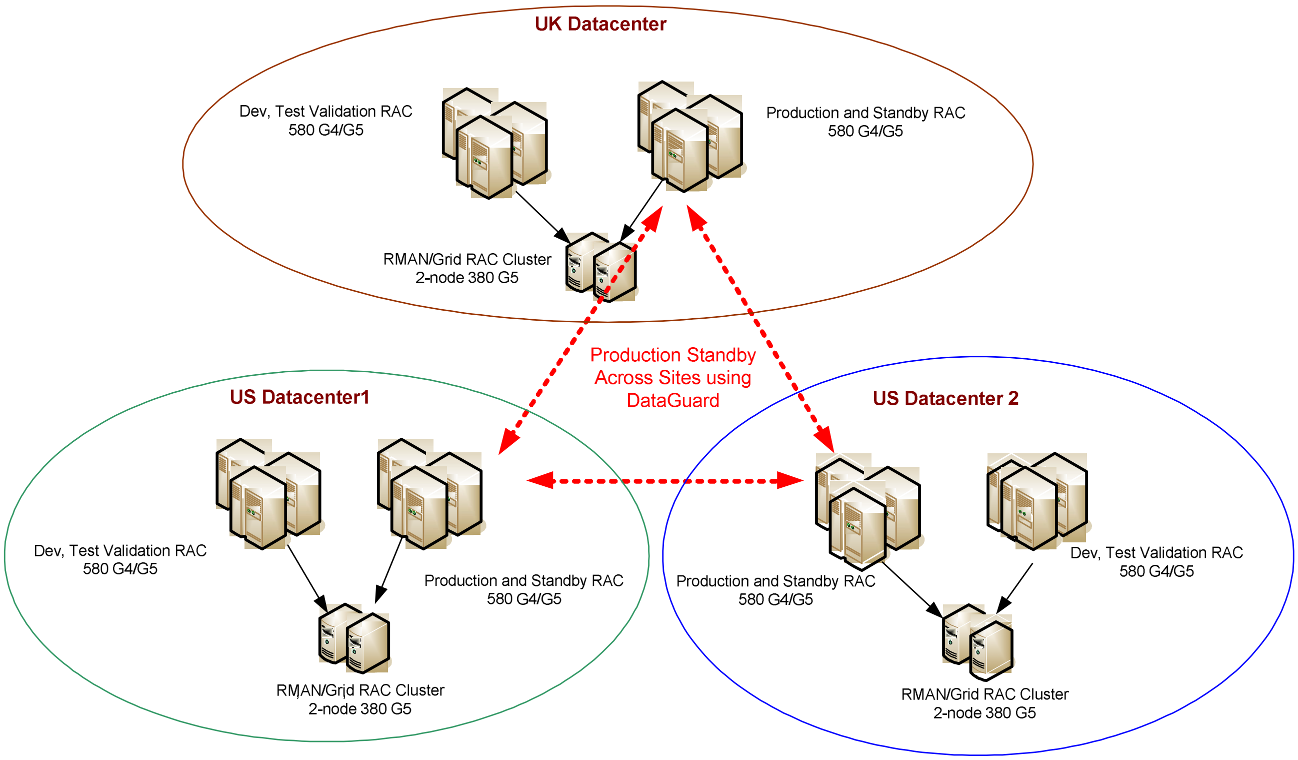

Figure 3, Architecture

Reports were run from the Grid Controls on a daily, weekly and monthly schedule, and they were a combination of customized and prepackaged reports. Reports were used by both by the DBA's and Shire's management to monitor the environment and for trend analysis. Typical daily reports consisted of backup outcomes for all databases, as well as database statistics. Weekly reports focused on performance metrics while monthly reports listed disk utilization by database so that database growth trends could be identified for managing disk space.

Conclusion

By virtualizing IT resources, Advans was able to enhance the scalability, performance and reliability of the database tier while lowering Shire's Total Cost of Ownership ("TCO") for IT assets. While an immediate benefit was obtained through server consolidation and the repurposing of existing servers, other indirect savings also resulted. Fewer servers in the datacenter meant less power and cooling needs in addition to improved space utilization.

Because of the scalability of pooled resources, less over-provisioning of servers and storage occurred, as both are added only when needed. Advans demonstrated this concept in its project plan by first building all the required RAC clusters with two nodes and migrating a subset of the total database tier to these clusters. Once built, the clusters were then later expanded to four and then six nodes as repurposed servers became available after decommissioning legacy systems. ASM Disk Groups were expanded along with the addition of RAC nodes, based on the current storage requirements of databases migrated to clusters, thus demonstrating that disk need not be provisioned a year or two in advance of its use. The effect is an increase in disk utilization while delaying both server purchases and storage provisioning which in turn increases the ROI of software applications. Furthermore, migrated applications demonstrated significant performance gains in RAC environments over dedicated servers, despite the use of Instance Groups for server consolidation.

While virtualizing the database tier is complex, requiring automated monitoring and an administration tool for its management, the result of abstracting database services to the application tier does serve to reduce overall IT operational complexity as databases services can be provisioned in real time. Given the greater reliability of pooled resources, more predictable service levels where observed, simultaneously reducing administrative costs while enhancing the service delivery of Shire's IT organization.

References

1. Oracle Maximum Availability Architecture website on OTN
   http://www.oracle.com/technology/deploy/availability/htdocs/maa.htm

2. Oracle Clusterware and Oracle Real Application Clusters Installation Guide
   http://www.mcs.csueastbay.edu/support/oracle/doc/10.2/install.102/b14207.pdf

3. Requirements For Installing Oracle10gR2 On RHEL/OEL 5 (x86_64), Metalink Note# 421308.1

4. How to configure Grid Control 10.2.0.4 Management Servers behind a Server Load Balancer (SLB), Metalink Note# 549270.1

5. Shire Pharmaceuticals: Transatlantic Deployment – Data Guard 10g Release 2
   www.oracle.com/technology/deploy/availability/pdf/ShireProfile.pdf

6. Oracle DataGuard Concepts and Administration – 10g Release 2
   http://download-west.oracle.com/docs/cd/B19306_01/server.102/b14239/toc.htm

Shire Pharmaceuticals Deploys Oracle Grid Computing Model to Lower Costs and Increase Reliability
September 2009
Published in Collaboration with Kevin Bradley and Matthew Manlove of Shire PLC
Alex Merkulov, Paul Angelo, Peo Nathan and Siva Gurunathan of Advans IT Services, Inc., Phil Grice
And
Joseph Meeks of Oracle Corporation

Authors: Paul Angelo and Peo Nathan, Advans IT Services, Inc.

Advans IT Services, Inc.
200 Friberg Parkway, Suite 2000
Westborough, MA 01581
U.S.A.
www.AdvansIT.com (or) www.GridisGreen.com

Inquiries:
Phone: +1.508.616.6900
Toll Free: +1.888.4ADVANS
Fax: +1.508.616.9901