

Gentran Integration Suite: Clustering Architectures

*Written by:
Raj Kumar
GIS Architect
Integration Management, Inc.*

*Contributions by:
Jason Miller
Shannon Bright
System Integration Architects*

*Mark Murnighan
Sterling Commerce*

March 2009

Contents

Server Architectures in Sterling Integrator	3
What is clustering?.....	3
Benefits of clustering.....	3
Architecture behind clustering	6
Controllers and Components	6
Queue Operation in Cluster	7
Default Queue Resource Allocation	8
Clustering in GIS/SI	9
Types of Clustering Architectures	9
Clustering Configuration.....	10
Capacity Planning	10
Active- Active (Preferred)	10
Active- Passive	11
Hybrid	11
Multicast	11
JGroups.....	11
Clusterable and Non-Clusterable Adapters	11
Typical Server Architectures for Clustering.....	13
High Availability Five tiers	13
Best Practices.....	18
Database Level	18
Application Level.....	18
Glossary.....	20

Server Architectures in Sterling Integrator

What is clustering?

Application clustering (sometimes called *software clustering*) is a method of turning multiple computer servers into a cluster (a group of servers that acts like a single system). Clustering software is installed in each of the servers in the group. Each of the servers maintains the same information and collectively they perform administrative tasks such as load balancing, determining node failures, and assigning failover duty. The other clustering method, hardware clustering, requires that specialized hardware be installed in a single server that controls the cluster.

Because servers can easily be added or removed from the cluster as needs dictate, an application cluster is more scalable than its hardware-based counterpart. Furthermore, because it doesn't require specialized hardware, an application cluster tends to be easier and cheaper to configure. For these reasons, application clustering is the more commonly practiced method.

A clustered installation of Gentran Integration Suite (GIS) connects one database to more than one installation of GIS. In clustering, servers are typically in one location, while installations in a federated system may be at several locations. Sterling Commerce's Gentran Integration Suite provides proprietary application clustering architecture. Clustering architecture in GIS version 4.3 has evolved since version 4.0 and Sterling has made significant advances in this area with moving from multi-cast to ip based architecture.

Benefits of clustering

Clustered environments can improve the three essentials for every system: Reliability, Availability and Scalability—often called RAS. In order for organizations to meet and beat their customer Service Level Agreements, they require the reliability that clustered GIS installations provide. Mission-critical business processes within GIS rely on other components, including network, disk storage and database. Adding redundancy improves GIS reliability significantly.

Improved **Reliability** means that systems handle payments, invoices and orders on time with fewer failures that can be addressed quickly if they do occur. Reliability often is referred to as the performance of Gentran Integration Suite.

Increased **Availability** reduces any scheduled and unscheduled downtime. Availability refers to being able to use Gentran Integration Suite.

Greater **Scalability** optimizes GIS technical and operational processes to handle future changes in volume. Clustering the database, file system, and GIS application servers can provide significant scalability while improving availability and reliability.

Small clusters are fairly easy to implement and are a better solution than simply using a larger server. However, clustering can add additional load processing to your backend database.

With two nodes, generally twice as many business processes are executing at the same time, so the database and all of the other components must be scaled to match. Improve the scalability of a business process by:

- Controlling Process Data within business processes.
 - Modulating business process components.
 - Avoiding locking or synchronization between processes
 - Taking reusability approach to minimize stress testing
- **High Availability** architecture helps minimize GIS downtime. Making an application highly available means minimizing the amount of scheduled/un-scheduled downtime. Highly Available architectures are expensive to implement and operate. First step towards making HA systems involve improving reliability. Following list can help you along with the decision making of implementing HA architectures
 - What is the real cost of downtime?
 - Does the length of the individual incidents make a difference?
 - Are there times when it is acceptable to be down (such as for maintenance)?
 - Does the system always have to be available at full capacity?
- **Load balancing and Performance Tuning** can be optimized through Vertical or Horizontal Clustering:
 - Vertical Clustering: Increase resources on one machine, e.g., another CPU, larger hard drive and increased memory.
 - Horizontal Clustering: Add multiple machines to existing architecture.

The overall performance of Gentran Integration Suite in a clustered environment is limited by four factors:

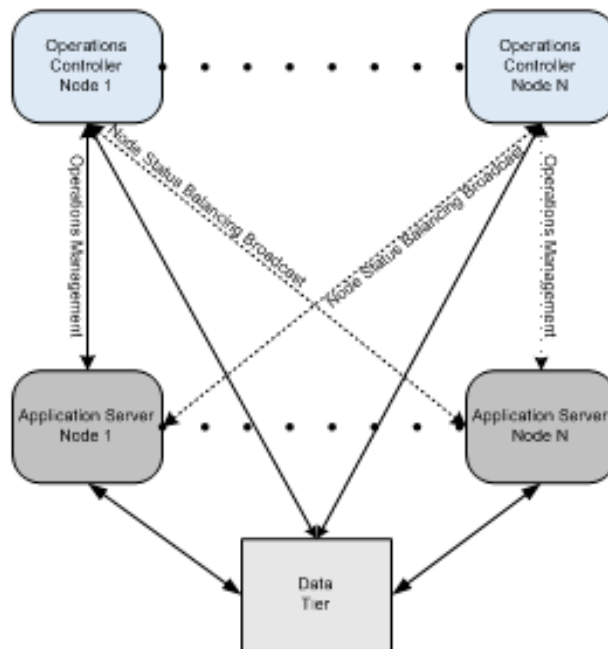
- The granularity of the work being done for example number of steps executed per business process and the amount of time it took to complete the process.
- Adapter arrangement
- Database performance
- Process interaction, e.g., locks and synchronization

Clustering Architecture

Controllers and Components

Gentran Integration Suite uses the following components in clustered installations:

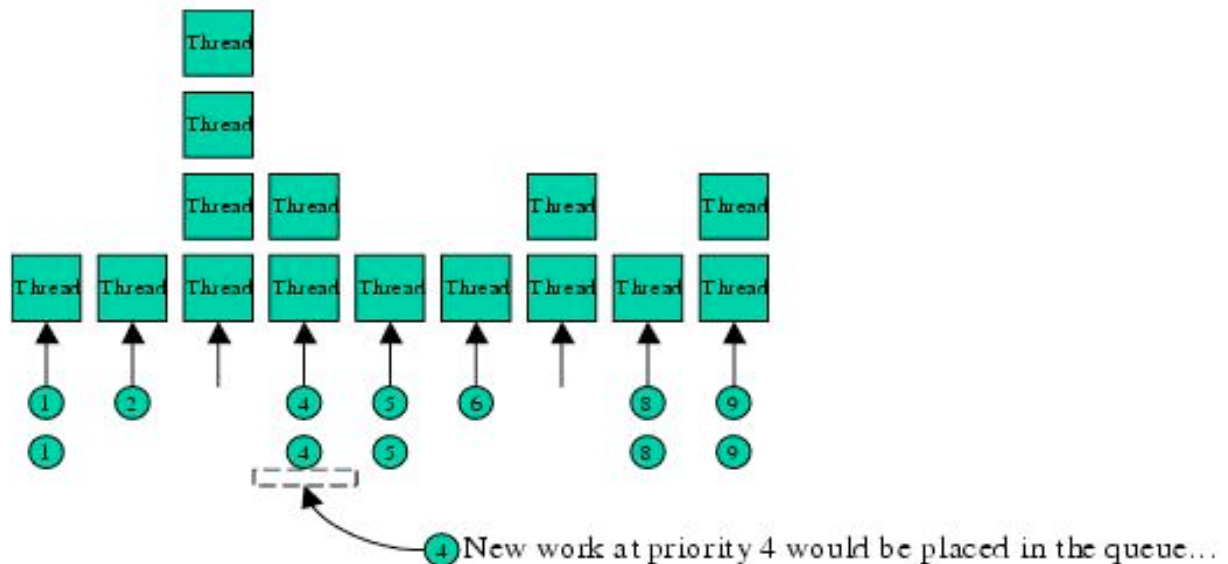
- **Core Engine Components:** Executes Business Processes within GIS.
- **Service and Adapters:** Executes Business Process steps and communicating with the external systems such as databases and ERP systems.
- **Operations Controller:** Manages JVM (Java Virtual Machine) resources. Each GIS node has an Operations Controller, which is responsible for listening across all queues. The Operations Controller uses JNDI (Java Naming Directory Interface) technology to communicate directly with other Operations Controllers in the cluster.
- **Service Controller:** Responsible for managing, configuring, querying and caching information for all services and adapters.
- **Central Operations Controller:** Communicates with the Operations Controller and contains heartbeat mechanism for the entire cluster. It runs on a separate JVM and each node has its own CentralOps Controller. During GIS startup the Central Operations Controller coordinates the local cluster node startup operations and its corresponding components, such as the Service Controller and the Scheduler. Central Ops-Controller also manages the following operations:
 - **Heartbeat mechanism:** Provides support for GIS central operations failover and node status update. It gathers queues information from cluster nodes and updates status. It runs on the same JVM as Central Operations Controller and starts as part of startup procedures.
 - **Token nodes:** Allows individual nodes to assume the role of being the Central Operations Controller.



Cluster Queue Operations

Cluster performance is affected by the number of threads allocated for each queue in each cluster node. Load balancing is dependent on the following factors:

- The number of threads.
- The number of steps in a Business Process that execute before rescheduling and then possibly distributed to another node.



Default Queue Resource Allocation

Out the box, GIS's default configuration for all queues is listed below. This information can be modified after installation to meet business Service Level Agreement's.

Scheduling Policy and Variable	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
ALL Policies									
AE_ExecuteCycle	1,000	100	100	100	100	10	5	2	50
AE_ExecuteCycleTime	1,000,000,000	1,000,000,000	10,000,000	2,000,000	100,000	100,000	500	300	100,000
QueueDepth	500,000	500,000	500,000	500,000	500,000	500,000	500,000	500,000	500,000
MaxPoolSize	1	3	5	8	2	8	2	2	4
CachingThreshold	0	10	10	20	30	50	50	100	10
MinPoolSize	0	0	0	1	1	8	2	2	2
JavaPriority	5	5	5	5	5	5	5	5	5
EnableDeadlines	False	True	True	True	True	True	False	False	False
Rescheduling	True	True	True	True	True	True	False	False	False
ReschedulingInterval	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000	5,000
MaxWaitTime	72,000,000	1,000,000,000	1,000,000,000	72,000,000	3,600,000	3,600,000	1,200,000	60,000	3,600,000
FairShare Scheduling Policy									
ResourceAllocation	50	20	40	40	50	60	60	120	50
InitialCycles	5	5	5	5	5	5	5	10	5
DeadlineThreads	0	0	0	0	0	0	0	0	0
StealThreads	False	False	False	False	False	False	True	False	False
AllowStealing	False	False	False	False	False	True	True	False	False

Perimeter Server

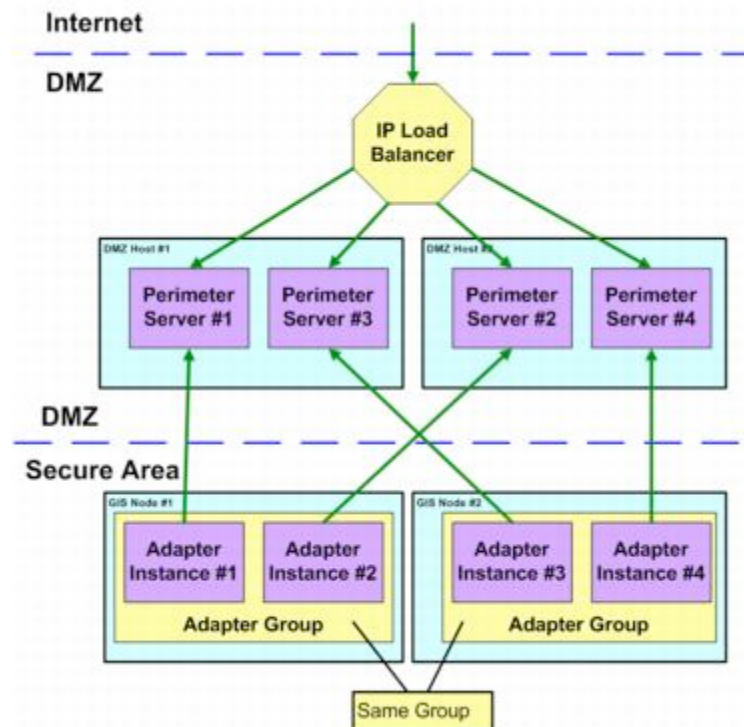
In a clustered environment, each node must be configured with one or more perimeter servers, which is communication management software installed in a Demilitarized Zone (DMZ).

Note: If a cluster node is not active, you cannot configure a perimeter server on that node.

The perimeter server is tied to a node; adapters and services that use perimeter services (such as FTP, Connect Direct, and the HTTP adapter) must use service groups for load balancing and failover

The perimeter server has a one-to-one relationship with the client. Because the perimeter server is tied to a particular node, adapters and services that are using

the perimeter server also must be deployed in that particular node. Multiple adapters and services can use a single perimeter server.



Clustering in GIS/SI

Clustering Architecture Types

In GIS, cluster architectures can be optimized vertically or horizontally.

Vertical Clustering: Increase resources on same machine e.g., add a CPU, hard drive, and/or more RAM.

- More CPUs allows more concurrent threads to run simultaneously, effectively increasing concurrent throughput.
- Faster CPUs adds additional horsepower, which increases execution time per Business Process.
- More RAM improves the overall performance. Memory is critical when caching objects and parameters at run time. The number of threads a node can support is related to its processing power and its memory.

Horizontal Clustering: Adding multiple machines to the GIS architecture. Scaling horizontally has the following advantages and disadvantages:

- Advantages:
 -
- Disadvantages
 - Adding nodes is always less efficient than enhancing the existing nodes due to cluster overhead.
 - Additional overhead management with more nodes.

Clustering Configuration

Capacity Planning

When planning a GIS system, take into consideration the following elements:

- Disk space for locally stored documents
- Disk space in the database (which is significantly affected by the archive policy):
 - For documents
 - For tracking information
 - For the persisting state of running business processes
- Database processing power
- Database connections (along with memory and other database resources)
- GIS node memory
- GIS node processing power
- Network bandwidth:
 - To and from GIS adapters
 - To and from database(s)
 - Between cluster nodes
- Number of GIS nodes in the cluster

Configuration Types

Active- Active

This is the preferred clustering configuration. Active-Active application level failover setup specifies that all nodes assigned to the cluster are in an “Active-Active” state. All nodes are available and working all the time, sharing the workload equally.

Active-Passive

Active-Passive setup includes a second GIS/hardware node that's configured exactly like the first and if the first instance fails, the second instance is ready to take over.

Hybrid

An Active-Active cluster is backed up by a passive node that can replace either of the active nodes in the event of hardware failure.

Multicast

Multicast is a communication method used by a cluster node to share information about its workload to rest of the cluster. Multicast is similar to broadcast communication except that the information is not sent to the entire cluster, but only to the nodes that are registered to receive information on specific addresses and port. Multicast requires routing components for communicating to rest of the nodes in the cluster. Follow the configuration properties below:

- In order to multicast to work effectively, all nodes must be registered on same network subnet.
- One port and socket is used for multicast communication across all queues in a node. Similarly, for the distribution of workflows, one port and socket per node is used across all queues. The cluster distributes the information to specific queues within a node.

JGroups

JGroups is used for broadcasting the BP load factor across all nodes in a cluster. JGroups communication method has two distinct settings:

- UDP (IP Multicast)
- TCP (Unicast and multicast)

Note: Some organizations restrict multicast usage for security reasons. IP multicast can serve as a bottleneck in cluster support across WAN for distributed clustering.

Clusterable and Non-Clusterable Adapters

Some adapters are non-clusterable and tied to specific nodes for the following reasons:

- **Resource Affinity:** File system adapters should not be deployed on a specific node.
 - FSA should not be deployed to a specific node because if it's pinned to node and if that node fails, then the process will not failover to node2 but halt the process.

- **Licensing restrictions:** There might be restrictions on number of instances that can be deployed in an installation.

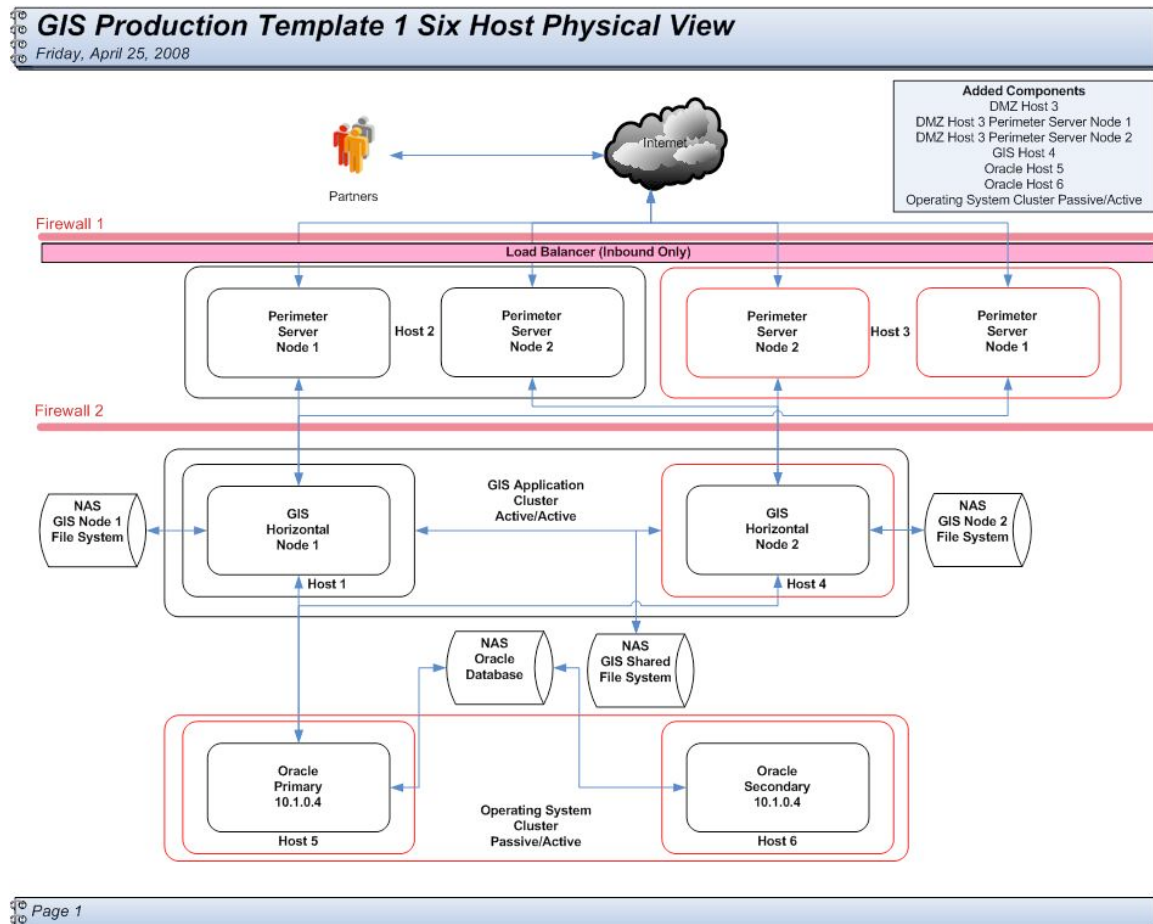
Adapters that use perimeter services are not clusterable, and are deployed to a specific node attached to their assigned perimeter server. However, these kinds of adapters can be grouped together using service groups and accessed as a single entity. These adapters include:

- FTP server adapter
- HTTP server adapter
- FTP client adapter
- HTTP client adapter
- Connect:Direct adapter
- Connect:Enterprise adapter
- B2B Communication adapter

Typical Server Architectures for Clustering

High Availability Five tiers

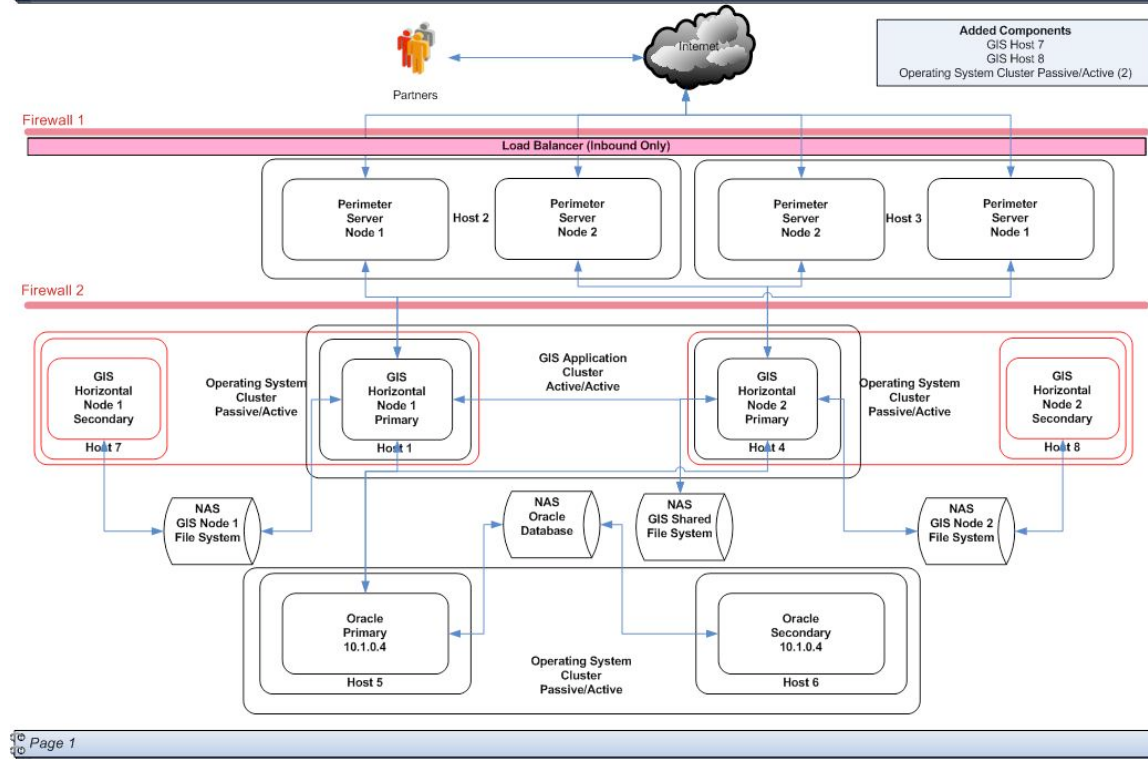
Tier 1



Tier 2

GIS Production Template 2 Eight Host Physical View

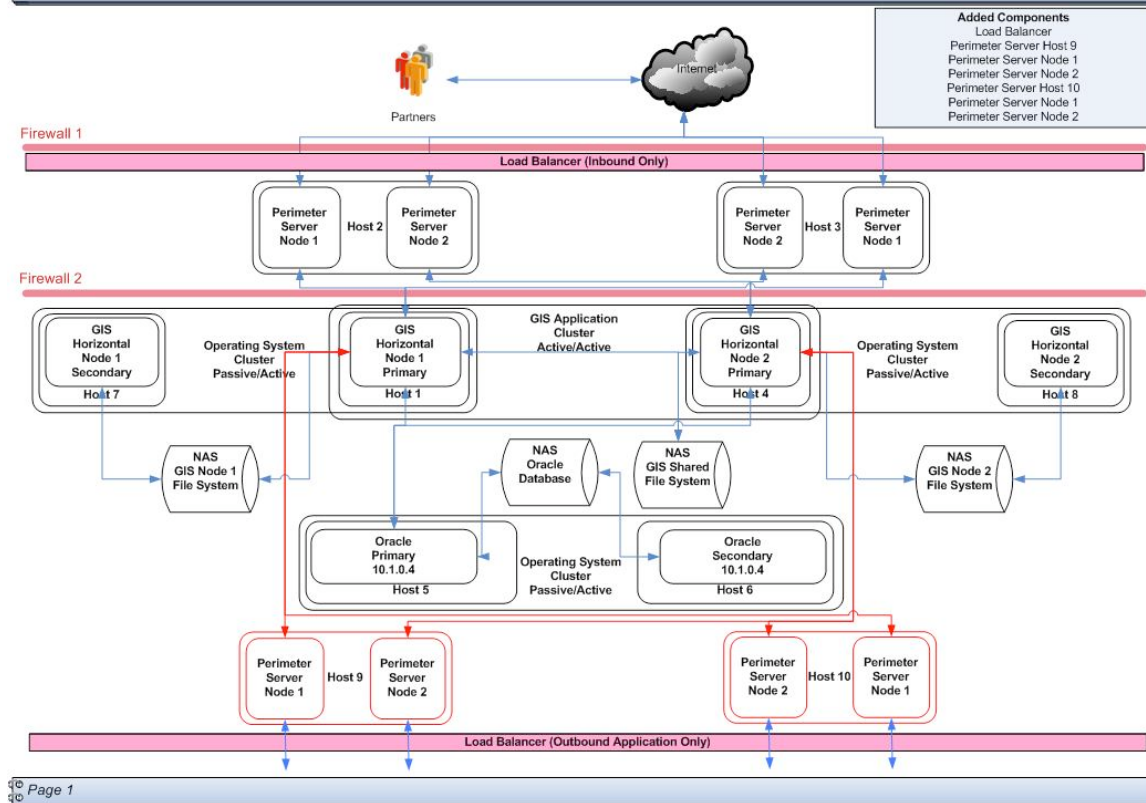
Friday, April 25, 2008



Tier 3

GIS Production Template 3 Ten Host Physical View

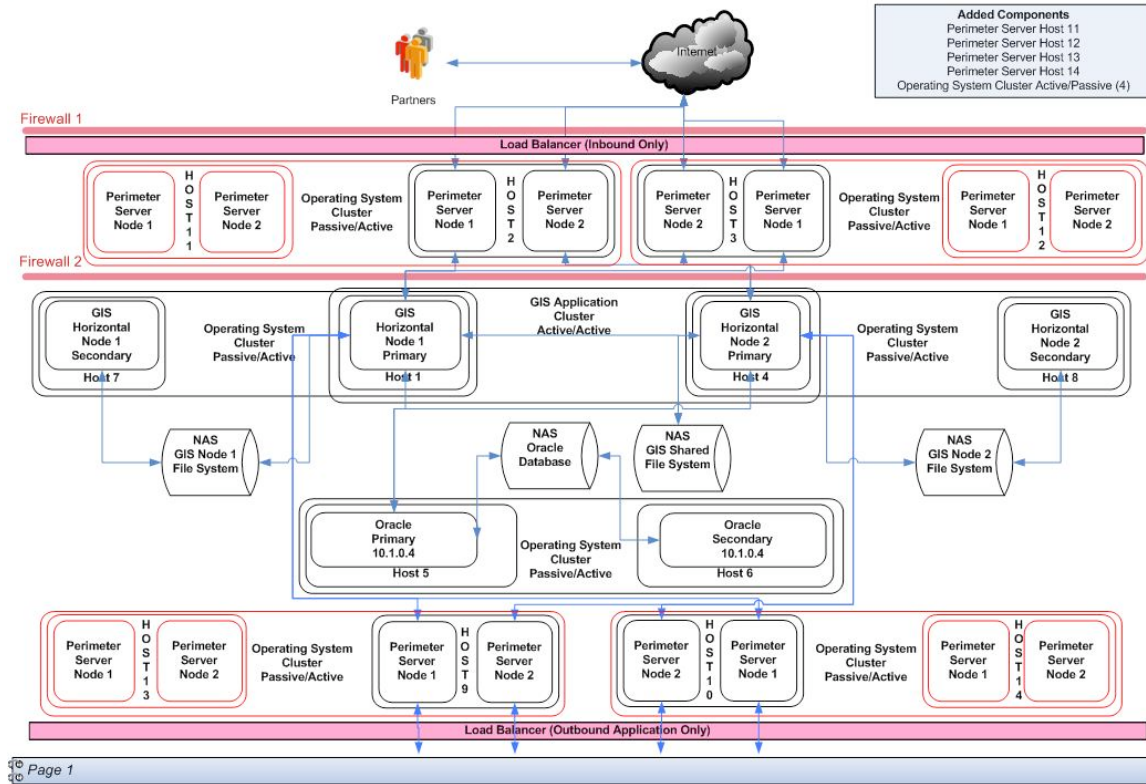
Friday, April 25, 2008



Tier 4

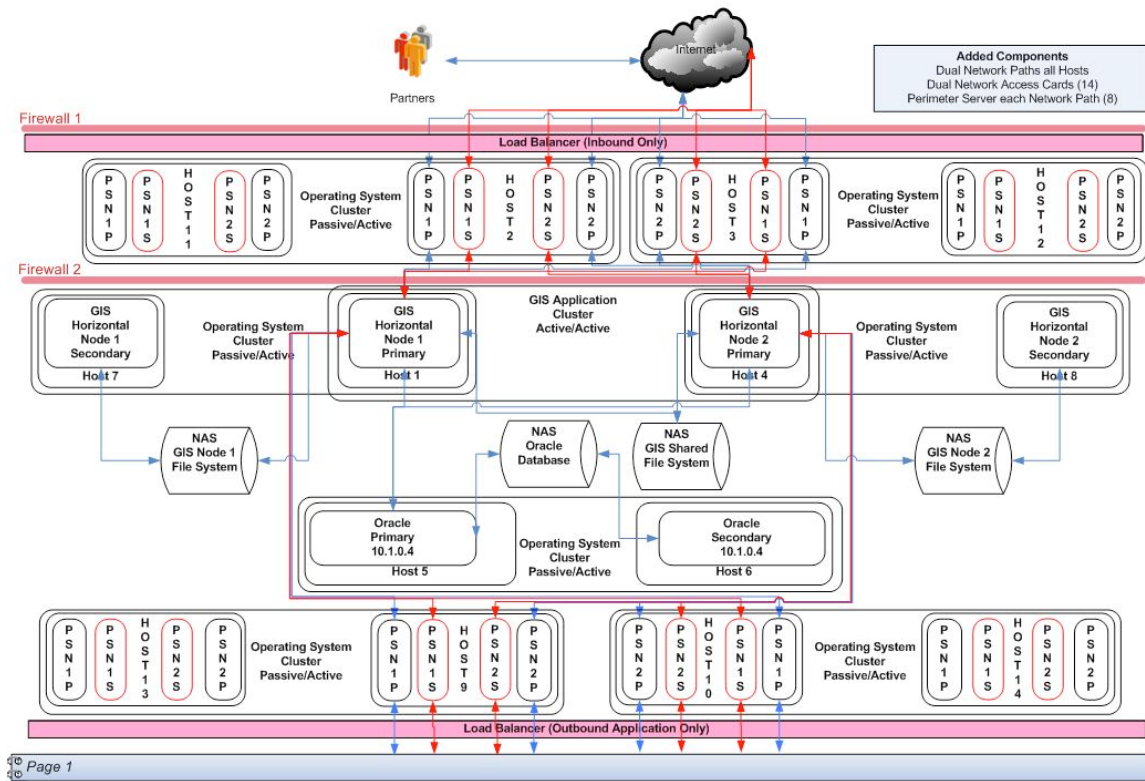
GIS Production Template 4 Fourteen Host Physical View

Friday, April 25, 2008



Tier 5

GIS Production Template 5 Fourteen Host Redundant Network Path Physical View
 Friday, April 25, 2008



Clustering Best Practices

Database Level

- Clustered environments require stable and well-scaled database servers to match your GIS architecture, because everything is pulled from databases and stresses the systems inputs and outputs.
- Maximum utilization requires a robust database engine with high processor speed and substantial memory.
- Strive for low latency between servers. Low latency maintains high availability.
- Database persistence vs. File System persistence
 - File systems can be faster but nightmare to manage.
 - Databases are easier to manage and easily replicated in case of DR (Disaster Recovery) testing.

Application Level

- Node-to-Node communication:
 - If installing multiple nodes on the same machine, go to the `install_dir/properties` directory of nodes 2 and higher and change the `mcast_port` property in `jgroups_cluster.properties.in` to point to the value of node 1's `mcast_port` property in `jgroups_cluster.properties`.
- Turning the schedule off for Recovery BP will improve system performance.
- Allow File System adapters in “all” environments while configuring to prevent the FSA from failing.
- Tune and configure:
 - The location and configuration of adapters.
 - The number of threads allocated for each queue on each node.
 - The number of steps for a business process to execute before being rescheduled.
 - The location and number of adapters is important because using adapters that cannot be clustered forces all activity for a particular adapter to go through one node. In other words don't pin clusterable adapters to a node

- Use an effective network load balancing technology in front of HTTP, FTP and other network-oriented adapters to prevent activity concentrating on one IP address.
- The number of threads in combination with the execution cycle, places proportionately more workload from specific queues onto specific nodes.
- In GIS version 4.3 there are no issues with session-balancing in other words communication layer of business processes do not have to be pinned to a particular node. In the previous release, this was necessary as communication invocation would fail due to swapping of sessions in mid-stream; however in the current release clustering architecture behind the scenes will manage and keep track of the sessions.
- Pin a BP to a node and then put adapters in a service group.

Glossary

- **Cluster:** More than one instance of GIS is considered a clustered environment.
- **Node:** Each instance of GIS is referred as a node.
- **Thread:** Process execution.
- **Queue:**
- **Load Balancing:** Processes that are distributed among nodes.
- **Multicast:**
- **JGroup:**
- **Token:** Each node is assigned a token.
- **Mandatory Node:** Operation in which a Business Process can be pinned to a specific node and must be executed on that node.
- **Preferred Node:** Operation that can be assigned to be run on either node.