Grid Enabled Service Infrastructure (GESI)

Isaac Christoffersen, Christopher Dale, Doug Johnson, Dan Santillo, David Schillero,
Booz Allen Hamilton
christoffersen_isaac@bah.com, dale_christopher@bah.com, johnson_doug@bah.com,
daniel_santillo@bah.com, schillero_david@bah.com

Abstract

At the 2006 Network Centric Operations Industry Consortium Conference on Utility Computing, Grids and Virtualization, the Office of the Secretary of Defense for Networks & Information Integration (OSD NII) presented a roadmap for transformation of the Global Information Grid (GIG) to the Net-Centric Environment (NCE). The planned transformation includes a federation of distributed computing resources, available when and where they were needed, and would be built on such technologies as grid computing, server clustering and virtualization. The Grid Enabled Services Infrastructure (GESI) meets OSD NII roadmap requirements as it currently performs mission critical operations at a government client site.

1. Introduction

At the 2006 Network Centric Operations Industry Consortium Conference on Utility Computing, Grids and Virtualization, the Office of the Secretary of Defense for Networks & Information Integration (OSD NII) presented a roadmap for transformation of the Global Information Grid (GIG) to the Net-Centric Environment (NCE). The planned transformation includes a federation of distributed computing resources, available when and where they were needed, and it would be built on such technologies as grid computing, server clustering and virtualization. The Grid Enabled Services Infrastructure (GESI) meets OSD NII roadmap requirements as it currently performs mission critical operations at a government client site.

The technology enablers of the NCE will be:

- **Clustering**: Collections of computers in a fixed configuration designed to be operated and managed as a unified, high-performance machine.
- **Virtualization:** Through virtualization of server, network and storage resources, GESI more efficiently uses under utilized hardware resources.
  - **Server Virtualization:** Provides the ability to deploy a discrete number of "virtual machines" on a single hardware platform.
  - **Network Virtualization:** Allows for the establishment of Virtual LANs (VLANs) between Virtual Machines.
  - **Storage Virtualization:** Pooling of physical storage from multiple network storage devices into what appears to be a single storage device that is managed from a central console.
- **Agile Architectural Framework:** Providing a very agile grid framework upon which to build a Service Oriented Architecture (SOA).

The GESI is designed to be a highly clustered, highly virtualized, highly agile grid of computing and storage resources. The GESI can be scaled from a transit case solution to a large data center configuration. Additionally, GESI provides the essential elements of a Service Oriented Infrastructure (SOI) as defined by the SOI Working Group of The Open Group. These elements are

- Life cycle support to manage the deployment of SOI components
- Virtualization of infrastructure resources to SOI users
- Service management to assure the SOI solution provides the required service characteristics

2. Background

The initial project started as a proof of concept for a document ingestion system. The development cycle was 90 days, which included infrastructure and software development activities. Shortly thereafter, the prototype transitioned to an operational proof of concept and then into a production system.

The mission data ingestion rates grew exponentially, but, because the initial prototype was developed as a proof of concept, the agility and flexibility to grow with mission requirements were not designed and built into the system. The next generation system needed to address these issues.

The solution involved both a software and infrastructure redesign and evolved into the GESI infrastructure.

3. Grid Enable Services Infrastructure (GESI)

The GESI is based on the layered architecture shown in Figure 1. The GESI was designed to be agile in both computing resources and software architecture.

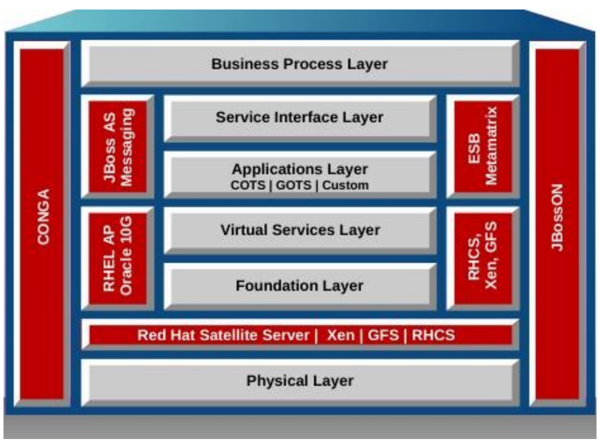

Figure 1. GESI Layered Architecture

3.1 Physical layer

The physical layer is made up of the hardware components such as computers, Storage Area Networks (SANs), Ethernet, Fibre Channel switches, etc. As much as possible, the GESI is built using commodity hardware which is defined as the most economical replaceable hardware components available to meet system requirements. Building GESI on commodity hardware means lower start up, maintenance and upgrade costs.

3.2 Foundation layer

The foundation layer forms the physical resource pool available for use at the virtual layer and above. This pooling of resources allows for better utilization while also providing high-availability and fail-over capabilities without disruption to the applications running in the infrastructure.

The foundation layer begins to move from an infrastructure that uses dedicated resources to one that allows for dynamic resource allocation. By pooling processing, network, storage, memory and database resources, provisioning of resources to applications can be done on an "as needed" basis.

3.3 Virtual services layer

Through virtualization of server, network and storage resources, GESI more efficiently uses under utilized hardware resources.

Server virtualization within GESI uses the open source XEN virtualization technology. The GESI consists of a number of XEN VMs built from prebuilt kickstart installation files. Each VM is designed to perform a specific function within the overall system architecture. As an example, there are VMs which perform only data ingestion.

Network virtualization within GESI uses bonded interface adapters with tagged VLANs.

Storage virtualization within GESI uses the Global File System (GFS) technology on top of clustered logical volumes (clvm) both of which are open source.

The virtual services layer provides the ground work necessary to create a "What you need, When you need it" environment.

Each suite of VMs is clustered. The clustering services are provided through the Red Hat Cluster Suite. The Cluster Suite provides application failover for VMs on a particular computing resource. This ensures that machines are always available to perform a specific function.

Cluster Management and Storage Management are controlled by the Red Hat Conga Cluster Management application. Cluster Management and the kickstart installations allow the cluster administrator a great deal of flexibility to redefine the grid. As an example, the cluster can be configured for initial mission requirements using a well defined set of resources. Through the cluster management suite the cluster administrator can essentially reconfigure the cluster on the fly. The cluster administrator can add or remove VMs from the cluster based on factors such as expected work load to meet surging data ingestion rates or high user activity. This also enables the administrator to consolidate the work load in real time and power down resources (physical servers) until they are needed again.

3.4 Application layer

The current implementation of GESI is developed on the JBoss Application Server. Out of the box, JBoss provides the ability to create a clustered application server environment. Not only is redundancy and fail-over now part of the default infrastructure, but the workload can be distributed in a way that allows members of the cluster to pull work when they can. In addition, JBoss Instances, running in XEN VMs on top of a GFS share, take the concept of High-Availability to a whole new level.

It should be noted that the “As Is” GESI is built on top of a J2EE Application Server, but because of the layered architecture, GESI can support any other type of application. For example, because of the ability to reconstruct the infrastructure at will, GESI is an ideal platform for an Information Assurance (IA) testing infrastructure.

3.5 Service Interface layer & Business Process layer

To take advantage of this highly clustered infrastructure, the software redesign transitioned to an event-driven Message Oriented Middleware (MOM) approach using enterprise integration best practices and a SOA.

The current implementation of GESI is a SOA using the JBoss Enterprise Service Bus (ESB) to connect services within the grid. Services are wired together using a Business Process Management (BPM) paradigm. This approach has facilitated an environment where reuse is favored over reinventing. Using this approach, the software infrastructure is a solution that is 85% COTS and GOTS integration.

The GESI provides several advantages in a SOA environment. Since the virtualized infrastructure has freed up computing resources, integration of new capabilities becomes a “no cost” proposition. GESI has allowed the client to expose new services simply by deploying a new VM containing the service. Additionally, services can reside in a “sleep state” until needed. This saves the power and cooling of a computer running a service which is seldom used.

Application and service management is performed using JBoss Operation Network (JON). The JON provides the capability to administer and monitor the applications and services running on the Application and Service Interface layers.

4. GESI Grid Architecture

The “As Is” GESI infrastructure currently consists of Services, Database and Storage Grids. Each can be deployed separately and is described below.

4.1 Services Grid

The Services Grid foundational notion was to eliminate the perception of data movement and start to think of data as being available when and where necessary. The GFS, which is clustered and highly scalable, makes this possible and offers numerous advantages over a networked file server.

Using GFS resolved one of the scaling issues, moving data between applications on Ethernet. The result was an improvement in performance where an entire day's data ingest, which once required 4 four processor machines to process, could now be processed on a single server running multiple VMs in about 20 minutes.

The Services Grid implements the Server Virtualization pattern, Single logical representation of multiple resources. This pattern consists of coalescing resources to represent them as a single logical representation that provides a single interface. This is an abstraction which permits users to change the underlying resource providers without disturbing the virtualized capabilities that exist above the foundation layer.

4.2 Database Grid

The current implementation of the Database Grid uses Oracle 10G Real Application Clusters (RAC). It follows the same design approach as the Services Grid. The underlying resources are pooled (memory, cpu, networking and storage) and clients connect to services that utilize a Transparent Application Failover (TAF) mechanism to ensure that single server failures do not cause an outage from the service subscriber perspective.

The RAC environment allows GESI to scale horizontally, by adding additional servers to the cluster, as needed. Currently, the Database Grid consists of a seven-node RAC cluster that provides ample database and back-end process capacity to supply the client's needs many months in the future. All of the RAC clusters and back-end servers are managed via Grid Control. This allows the cluster administrator to monitor, manage, and optimize all back-end systems from a single management control console.

The RAC environment supports the establishment of services to the database. Users are connected to the services that allow GESI to take advantage of TAF and work load balancing between the servers in the database grid. Services allow the cluster administrator to designate primary and secondary servers to support database connectivity, failover and load balancing. Users no longer connect directly to a database instance, but to a service. Each service is supported by a number of RAC instances. If for any reason a node in the cluster fails, users who are connected to the database via a service will automatically and transparently fail over to surviving nodes. This also gives the administrators the ability to perform maintenance on the database grid with minimum down time. Services also provide the means to spread workload over a number of database servers so that jobs can be run concurrently. This "Divide-and-Conquer" approach significantly reduces the amount of time it takes to perform many database processes.

An example of this approach can be seen in the implementation of Oracle Text, which is a utility in Oracle that allows users to search on words and phrases, across millions of documents, in a matter of seconds. Oracle Text is used extensively and the system has realized major performance gains within the RAC environment. By partitioning the tables, synchronizing and optimizing the index across every server in the RAC environment, the divided partitions in the table can be handled by every server in the RAC cluster concurrently.

By applying the collective process capacity of the database grid, the processing time for maintaining the text index has improved from 84 hours to under 2 hours.

4.3 Storage Grid

The GESI "As Is" implements a Shared Storage paradigm based on Red Hat GFS on top of a virtualized storage fabric. The GFS allows all services to have direct, concurrent access to the same shared storage block. Content is accessed via globally unique identifiers. This allows the GESI to move toward a virtual Storage Resource architecture.

5. GESI as a SOI

The SOA Reference Architecture defines the 3 components (see Figure 2) of a SOA Foundation as:

• **Business Architecture:** Based on the business strategy, objectives, priorities and processes.

• **Infrastructure Architecture:** The engine that enables SOA and should address all the aspects of the infrastructure from networks, servers, data centers, and firewalls, to application infrastructure, security, monitoring, middleware, etc.

• **Information and Data Architecture:** This deals with identifying the Key Performance Indicators and the information needs that drive the enterprise.

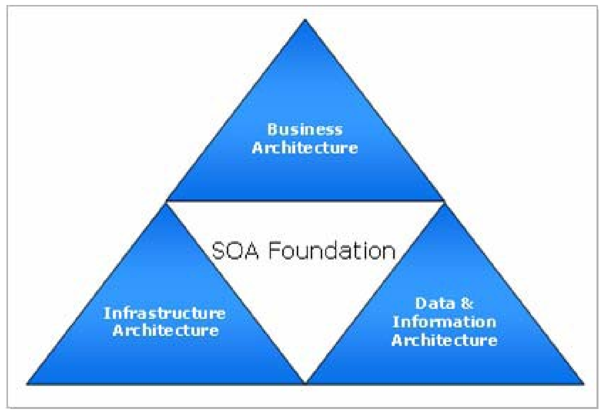

Figure 2. SOA Foundation

GESI addresses the Infrastructure Architecture aspects of the SOA Foundation by providing a very agile framework upon which to build an SOA.

5.1 Advantages of a SOI

GESI has already provided the client with many of the benefits of a SOI outlined by the SOA Reference Architecture. It has allowed the client to move from the static, "one-box-per-application" approach to a dynamic resource allocation in which virtual processing, storage and network resources are assigned to applications as needed. This approach reduced the number of servers running in the client's system from 50 to 15. This reduced capital costs through better resource utilization and increased reliability, since applications fail over to available resources without disruption.

GESI delivers more consistent service levels because the cluster administrator can allocate additional resources in real time as an application workload increases. It has been able to handle a 600% increase in workload by reallocating VMs across the grid infrastructure.

GESI has already allowed the client to overcome the “complexity barrier” that limits an organizations’ ability to design and implement new business processes. We have proven that GESI users can deploy new services in a short period of time.

As the lifecycle of GESI moves forward, it will provide many other benefits of a SOI, some of which users are beginning to see, including:

• Reduced operations workload, due to reduced manual reconfiguration of hardware and software

• Higher productivity, resulting from the ability to perform enterprise-wide platform, network, data and application management from a single, standardized management console

• Better resource utilization leading to reduced capital expenditures to accomplish IT objectives

• Greater flexibility, through dynamic resource allocation

• Simple and cost-effective upgrades, thanks to a modular, loosely-coupled SOI architecture. This allows IT to refresh its infrastructure to take advantage of new technologies without complications resulting from hard-wired management dependencies. This eliminates painful “rip-and-replace” (aka “forklift”) upgrades, allowing IT to respond to changes in a rapid and graceful manner. This reduces the inertia that causes IT to resist incorporation of new technologies because of expected disruptions.

6. Conclusions

The Grid Enabled Services Infrastructure, which is actually an implementation of a more generic framework, improves user capability and flexibility. Its scalability, efficient power consumption and efficient handling of large datasets makes it an ideal model for the next generation data center.

7. References

[1] Daly, John. Department of Defense Computing Infrastructure Brief NCOIC Stakeholder Outreach Working Group, 6 December 2007 http://colab.cim3.net/file/work/SICoP/Virtualization/2007-12-06/JDaly12062007.pdf

[2] The Open Group, Service Oriented Infrastructure: Charter, 31 July, 2007, http://www.opengroup.org/projects/soa-soi/uploads/40/14171/OG-SOI-Charter-v1-1.pdf

[3] Maldonado, Martin F. Virtualization in a nutshell: A Pattern Point of View, 20 June 2006 http://www.ibm.com/developerworks/grid/library/gr-virt/

[4] SOA Practitioner’s Guide Part 2: SOA Reference Architecture, 9 September, 2006, http://colab.cim3.net/file/work/SOACoP/2007_05_0102/SOAPGuide/SOAPGPart2.pdf

[5] SOA Alliance, SOA Blueprint SOA Reference Architecture Version 1.1, 2006