Grid Task 2006

Grid Architecture Analysis for NECC and the U.S. Department of Defense

Version 1.0 final, 1 December 2006

Prepared for:

Mark Kuzma
Defense Information Systems Agency
5275 Leesburg Pike, Falls Church, VA 22041
NECC Program Management Office

Prepared by:

Michael V. Behrens
Office of the CTO
R2AD, LLC

"Possibly the single-most transforming thing in our forces will not be a weapons system, but a set of interconnections and a substantially enhanced capability because of that awareness. "-Secretary of Defense Donald Rumsfeld, August 9,2001.

1. Grid Task 2006 Final Report

1.1. Executive Summary

NECC is evolving along with web service and grid standards and is therefore also involved in the net-centric transformation which is affecting the architecture, design, and implementation of command and control systems. DISA is cautious about adopting technologies which are immature or those which are not supported by the commercial marketplace. Grid computing is evolving from a science technology into a business technology and, as such, is moving from the research labs into corporate America as noted by eBay's grid focus[1] and keynote address at Grid World 2006. Grid technology is also being redefined during this process and the grid community is starting to extend the role of grids beyond parallel execution of jobs into a paradigm where the "grid" is the backbone of a new distributed computing enterprise in which secure data sharing and dynamic servicing are core strengths. This confluence of technologies is being dubbed Grid 2.0 (the second grid generation).

1.2. Report Overview

This report investigates several areas offered by grid systems which can directly positively impact NECC. These include:

- Grid World Review
- Grid Vision
- Key Grid Technologies affecting NECC
- Web Service Evolution and Grid Standards
- Recommendations for NECC

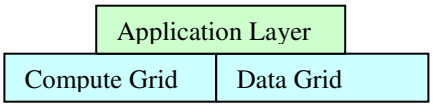
This report is a follow-on to the Grid Study, which was delivered to DISA last fiscal year. It is available on-line from: http://www.r2ad.com/whitepapers.html.

1.3. Grid World 2006 Review

On behalf of DISA, NECC and PMO, R2AD attended and participated in the Globus World and OGF-18 event. A trip report of this event has been delivered and is included in Appendix B of this document for reference. The conference was highlighted by the announced merger of Enterprise Grid Alliance (EGA) and Global Grid Forum (GGF) to form the Open Grid Forum (OGF). This is helping to add stronger *commercial* involvement, as indicated by eBay's Key-Note address. Also, the DoD is now a sponsor of OGF.

In addition to attending Grid World 2006, R2AD has been involved in weekly teleconferences with the Open Grid Services Architecture (OGSA) working group which is helping to set a global standards system for better interoperability in both the commercial and government markets.

2. The Grid Vision



As noted in the first DISA Grid Study [**Grid Study**], there are different types of grids: Compute grids, data grids, collaboration grids, and other vertical grids. One important concept which came out in the Grid World conference, and especially while participating in the Globus Toolkit Tutorials, is that grid applications do not have to be parallel in nature and can just be regular ordinary applications and services which operate within the grid environment.

We might describe such a grid as an general purpose grid[2]. It is a vision of a real distributed object environment which utilizes virtualized computation, memory/storage, and networking. It brings together the distributed computers by creating an abstraction around them which exposes a simple interface to create data, find data, and operate on data. This concept can be achieved with current grid software and it is evolving to allow end user to dynamically create data or organizations of data, and to be able to share dynamically that information with many others around the world in a secure manner within dynamic virtual organizations or Communities of Interest (COI).

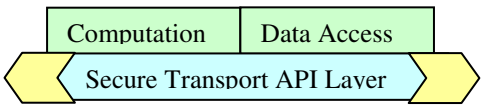
[1] http://www.eweek.com/print_article2/0,1217,a=134210,00.asp
[2] CIO magazine Oct 2005: http://www.cio.com/archive/101505/et_development.html

This application layer is built on top of compute and data grids to create a series of dynamic virtual spaces which might be analogous to NECC's **Operational Context** and the contained entities.

2.1. Key Grid Elements

In order for the NECC Grid to be fully realized, a series of steps need to be taken for which advances have already been made:

1. Global identity which can survive distance and time
    a. WS-Naming
    b. Handle.NET
    c. Globus MyProxy (Generates temporary PKI Certificate as a temporary resource ticket)
2. Global creation and management of data (state)
    a. WSRF (WS-Resource Framework)
    b. WSRT (WS-Resource Transfer)
    c. OGSA DAI (data architecture)
3. An abstract language which can describe workflow and algorithms which operation on patterns of data and events
    a. OGSA Execution Monitoring Service (EMS) (higher level)
    b. WS Choreography and BPEL
    c. WS Eventing
    d. Other grid specific languages and mechanisms (currently being discussed in OGSA)
4. A set of underlying system constructs to provide for the secure streaming of bytes to any node in the grid
    a. OGSA DAI (standardized access to relational and XML data)
    b. OGSA Byte-IO
    c. TLS (Transport Layer Security)
    d. WSS (WS Security)
    e. Secure GridFTP
    f. **Reliable File Transfer (RFT)**
    g. Stream Control Transmission Protocol (SCTP)[3]
    h. WSRM (WS Reliable Messaging)
    i. WS-RLS (WS Replica Location Services)
    j. Extensible Input/Output (XIO) from Globus



All items together form a powerful framework for build a net-centric system for NECC. The Globus Toolkit is GOTS solution backed up by COTS providers/supporters to provide an advanced framework that can be used today to build grid enabled applications. Globus consists of many modules which provide abstractions to common problems found in Net Centric architectures. These libraries and tools can be used to help transform the common patterns that are typical in command and control systems (i.e. add/update a battlefield object) into generic interfaces that decouple the application into discrete components. This would enable interoperability at a service level and across domains.

The following subsections cover some of the key items in more detail.

2.1.1. Security

Grid Security is based on Web Services Security (WSS). The profile specified by OGSA provides an agreement of security which is required for interoperability. The Globus implementation for instance, defines a near compliant set of services called Grid Security Infrastructure (GSI)[4].

Current specifications for security in grid systems are aligned with the WS-I Basic Security Profile[5] and are being updated as needed to reflect the minor changes in the recent 1.1 version. OGSA is now submitting a final Secure Channel Profile for publication which depends on the OGSA Basic Security Profile. These standards help ensure interoperability in grid systems such as NECC.

---

[3] SCTP info: http://www.sigtran.org/
[4] GT4 Security: http://www.globus.org/toolkit/docs/4.0/security/
[5] http://www.ws-i.org/Profiles/BasicSecurityProfile-1.0.html

2.1.2. Identity

Each object (entity) should have its own unique identifier and these millions of objects change state and behaviour over time in a replicated battle space. For example, the U.S.S. Reagan might have a unique ID, as also would a unit that might assigned to her (a UIC perhaps) for a period of time. However, the identity for global access to an entities properties and behaviour might be implemented for better machine to machine communication and manipulation such as a WS-Address (WSA) or a WS-Name. As each entity can be itself a service endpoint, Service Addressing would be used to identify a service's network end-point. Since networks are fluid, machines that are hosting a service is accessible via a WSA and in time would more than likely be migrated to a different host or might not be available at all. This problem is addressed by the WS-Naming specification and other specifications like the Handle system: http://www.globus.org/toolkit/projects/handle_system.html. Grid computing is making good strides in solving this problem which has largely been ignored by the rest of the Web Service community.

Identity correlation is not currently being addressed however (as in Track correlation). Bringing this requirement to the standards groups would be recommended as then COTS implementations would support correlation out of the box, thus reducing the burden on the tax payers.

A good case-study to example is a project that the National Institute of Health (NIH) is sponsoring which provides unique identity to over 100 million entities, are network accessible. The caBIG[6] project uses the Model Driven Architecture (MDA) approach by modeling in Unified Modeling Language (UML) all of the semantics of their taxonomy or Enterprise Semantics. If their objects were replaced with what GCCS-J refers to as "tracks" or with NECC Entities, the result would be a vastly scalable command and control system. An example system being used today for the scientific community is the NERC DataGrid[7]. Commercial applications use keys typically in grid systems like eBay or Google and also within COTS libraries for data caching (Java Cache).

2.1.3. State

A key grid standard is WSRF. It is one of several web service specifications which support the notion of web service state. Instead of having to use cookies (per-se) in headers and elsewhere, WSRF provides access to state directly, using simple setter and getter operations. This is an advancement well beyond most web services because it standardizes the way in which properties are exposed and communicated. Instead of writing custom code for methods such as getCountryCode, an object can simply expose a managed property called CountryCode. Each instance would have that as an available attribute.

> ***Note**: Attributes could be defined outside of code in the metadata using XML Schema or other graphical models and could extend existing meta-models for common semantics. OGSA is embracing the Common Information Model (CIM) and extending it where appropriate. NECC would mine the DoD XML Registry and support the various metadata working groups in their efforts to standardize a DoD model.*

Furthermore, WSRF provides a container, in effect, for many objects of the same type. For example, a single web service for tracks could actually expose all tracks as their own service with their own endpoint which can be passed around in the form of a WS-Address. Whoever has the WS-Address can obtain the state of the track. WSRF therefore provides an elegant container model and is efficient by exposing a scalable brokering service to handle the requests for all its contained entities.

In the longer run of grid state, WS-Resource Transfer is going to play an important role. Also, it is important to note that manageability of a service to include service endpoints lifecycles can be controlled via these sort of the state management service descriptions. WSRF is an OASIS standard and is used by Globus.

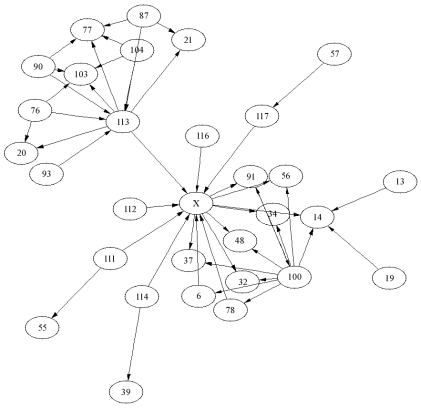
Migration of services is already being address by the grid community. This involves the moving of the service to another computer. This would be done to help meet service level agreements or QoS, or simply because the hosting container is going away. As standardized mechanisms are used to keep state, the service needs to be flexible in terms of where it is. More than one service could, in fact, be responsive to the same requests. The "scheduling" or management of which service instance is invoked is the job of the grid engine with particular focus on the scheduler and the Execution Monitoring Service (EMS) as described in the OGSA architecture[8] document.

---

[6] caBIG: http://ncicb.nci.nih.gov/NCICB/projects
[7] http://ndg.nerc.ac.uk/
[8] OGSA 1.5 Architecture: https://forge.gridforum.org/sf/go/doc13553

2.1.4. Secure Transport

Communities grow and have interactions. The node diagram illustrates this and it is just a reminder that although in NECC supports COI's, there is a reality of Cross-COI and Extended COI's which should be kept in mind when building and deploying the architecture.



Having a common transport layer to connect these nodes in a manner which is *independent* of the data schema is extremely valuable. This is analogous to using HTTP/S for web page retrieval - it's simple and allows for rapid growth of data sharing. A secure transport is the foundation of a C2 system. In current systems, COP Sync-Tools (CST) is one mechanism which is specific to tracks. It is not difficult to adapt the Globus XIO capabilities in NECC along with WS Reliable Messaging and one or two other protocols into a core reliable communication channel. This channel would be accessible via a standardized API for C2 systems and would offer a variety of delivery options based on the need of the sender and receiver. Quality of Service can be managed and optimized. This in effect, is an Enterprise Service Bus (ESB).

Globus ships today with a service known as Reliable File Transfer (RFT), which in can invoke GridFTP. It manages the network connections and data cache to ensure that the data is transferred in conditions when the network is not reliable. This can be built upon for other protocols as well. The key is to shield the applications from having to worry about transport and secure transport and instead let the applications focus on the business logic which should be event driven.

Caching of data is important for efficiency. There are a number of commercial grid providers that can assist NECC with this aspect of grids (IBM, HP, Oracle, Platform, United Devices, Data Synapse, Univa, Voltair, eXludus, Mellanox, Tangosol, GigaSpaces, Cluster Resources, Appistry, Altair, ASpeed, GemStone, Digipede, Egenera, Ibrix, Azul Systems, Sun Microsystems) . The combination of secure transport and grid data access yields the concept that data and behavior is availability everywhere, even though this is an illusion, the effect is real. The Internet cloud becomes a Grid Cloud.

2.1.5. Abstract Grid Language

R2AD has hinted that a new programming language is eventually needed to codify the grid paradigm and to help abstract the basic functions into human readable constructs (Web Services, for example is already too complicated). R2AD shall continue to research this topic and might solicit NSF for R&D support. An example of this is Sun's research in to the DASL[9] language. Another is the Simple scripting language which offers mobile behavior as well as mobile state. Simple is based on the Java language.

A very interesting commercial efforts is Zero-C (used by eBay perhaps): http://www.zeroc.com/. There are other examples as well (i.e.; Sun Microsystems' DASL or from embedded computing: Javolution[10]) and exploring them would be a part of an analysis of alternatives to defining a new language to meet the needs of the DoD (ala the effort with Ada, however in a commercially/grid supported space).

In recent teleconferences within the OGSA working group, the abstraction of work-flow (BPEL, WS-Choreography, JSDL, etc) have been discussed and timing might soon be good for a submission of a new language for standardization. OGF/OGSA would not be interested in designing the language themselves, however if industry or an agency were to create it (DISA, NSF, OSD, DISA, etc) then they would be happy to accept it and work to make it a standard as that would be in their interest as well.

In the end, R2AD believes that a new language will emerge which shall contain all the net-centric principles. This language will greatly help to make the entire process of using the global grid very simple by abstracting the complexities currently found in current frameworks. Also, graphical representation of workflow might also provide the a level of simplicity[11].

---

[9] http://research.sun.com/techrep/2005/abstract-128.html
[10] DASL: http://research.sun.com/techrep/2005/abstract-128.html and http://javolution.org
[11] Grid WorkFlow with Globus: http://www.gridworkflow.org/kwfgrid/gwes/docs/

3. Evolving Standards

One of the most important changes to the overall distributed grid architecture model in recent years has been the adoption of web services as a foundation for method invocation and data exchange over a network. These services are built upon well known web service standards and provide access to grid resources using XML messages communicated typically using the HTTP/1.1 protocol. XML Schema provides a common typing system which combined with Web Services Description Language (WSDL) provide the building blocks to exchange information across disparate systems.

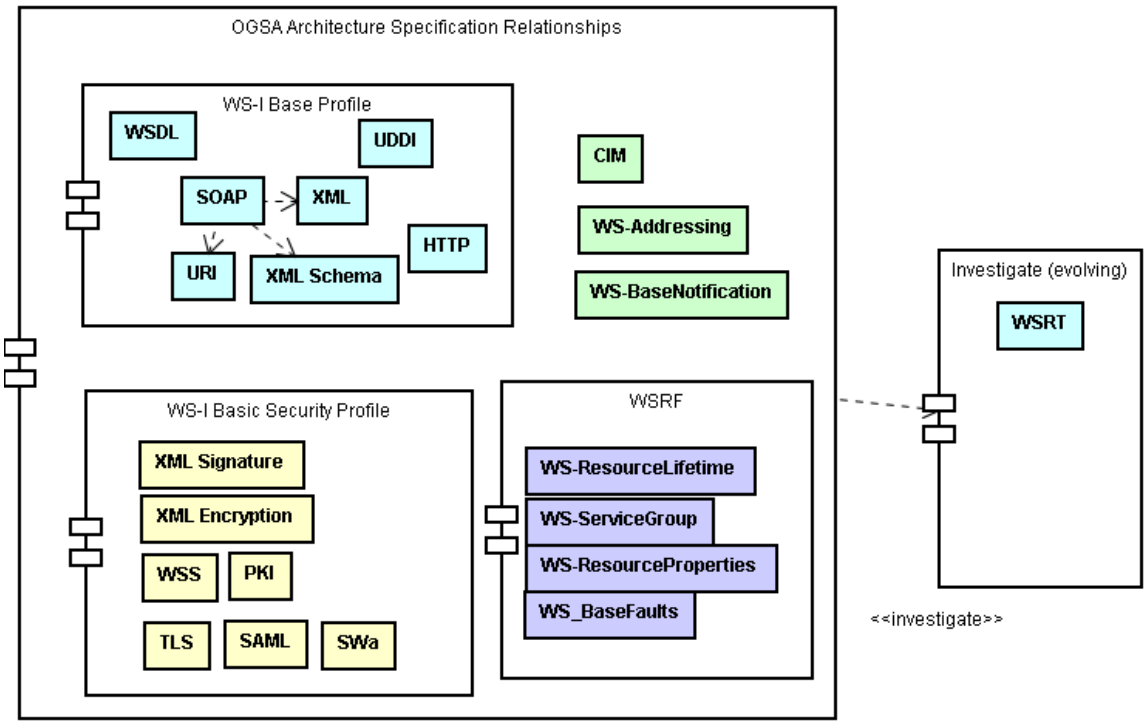


Figure 31 - Container Relationship of standards and specifications

The evolving[12] web service specifications cover many different functional areas to include messaging content, transport, security, transactions, metadata, and workflow. Since individually, these specifications may be interpreted and combined in ways which inhibit interoperability, key web service specifications have been composed together with guidelines to form profiles. The Web Services-Interoperability Organization (WS-I) has published the WS-I Basic Profile and the Basic Security Profile (BSP). OASIS has published the WS-Resource Transfer Framework (WSRF) which adds the notion of state to web services which can be remotely accessed and monitored.

OGSA builds on this modular non-proprietary foundation to define specifications which enable grid systems to interoperate and to share resources between organizational boundaries, defining a standardized grid service infrastructure. OGSA ties these web service specifications together in order to describe distributed grid systems using the taxonomy and modeling principles defined in the Common Information Model (CIM) published by the Distributed Management Task Force (DMTF).

OGSA publishes this combination of evolving web service standards and OGF developed specifications as OGSA profiles, establishing the definition of a standardized interoperable grid system.

4. Recommendations and Way Ahead

The following sub-sections provide a few recommendations or options for NECC on how to being using Grid Technology to meet the global C2 challenges. Some of these could perhaps be implemented as part of the planned FY-08 efforts.

[12] http://roadmap.cbdiforum.com/reports/protocols/

4.1. Grid Day At DISA

As per past conversations with key individuals at DISA and at the Grid World/OGF event in September, it would be prudent to consider a grid day at DISA where a teleconference or a Face2Face meeting can occur to dive into the details of how grids can help the DoD. This could help kick off a design period and a Design Team which could produce a comprehensive model drive design of a global command and control system. Such a design would be impartial to older stove pipe systems and would embrace a modern data sharing principles (need to share) typical grid based global enterprise systems.
A call for participation could be posted to the OGSA mailing list or simply mentioned during the next teleconference. OGSA has traveled to various places before to help support grid computing in general and would probably officially support such as visit (OGF sponsored). It would be wise to invite both academia and commercial entities for a deep dive session. A list of recommended invitees is shown here (see Appendix A for contact info):

Standards Representation

- Ian Foster, Argonne National Labs (ANL), Father of Grid Computing
- Andrew Grimshaw, UVa, OGF steering group, OGSA, etc.
- Steve Crumb, OGF vice chair of operations

Commercial Representation

- Jay Unger, IBM, OGSA and Liason for standards development
- Paul Strong, Ebay
- Don Campbell, Tangosol
- Steve Tueke, Univa Globus Commercialization
- Christopher Smith, Platform Computing

4.2. Globus Demo

Globus is a service framework which provides a great foundation to build a C2 system and is very standards oriented and compliant. We suggest that a small task be initiated which would:

- Demonstrate Globus in action
  - Combine with a portal
  - Service Monitoring and Migration
- Provide a ground footing for other NECC developers to start using
  - virtual images for them to use in development environments
  - Guidance and help
- Provide a functioning GridFTP environment for secure file replication
  - Supports Content Staging and provisioning

Such a demo could be executed in short order if sample/demo grid applications were shown. However it might be more valuable to NECC if a more complete demonstration or real capability were developed so that other capability modules could benefit over time. Perhaps a grid based Video Mosaic example with help from I3 to rapidly process live sensor video?

A basic demo for NECC might be to show how services can be migrated and managed which is the essence of service virtualization. The demo would combine Globus and the Handle system (Identity) to show how a web service can be moved to a totally different machine and still be used by clients. For more specific information about this specific demo, refer to the following wiki: http://piggy.cnri.reston.va.us/CNRIwiki/FrontPage.

For more detailed information on Globus, please read the first Grid study paper and/or consult the Globus web sites.

4.3. Grid Node Design (Appliance) for NECC

NECC requires an effective means to collect and analyze vast amounts of information collected from geographically dispersed sensors in order to accomplish a variety of missions such as situational awareness, strategic defense, command and control, intelligence, and overall information superiority. A comprehensive design of a NECC grid system which can meet the requirements in the NECC capabilities document is recommended at this time. The output of the design effort would be a MDA (UML) based Design Document and would be fully traceable.

A key benefit of designing a grid system for NECC would be a standardized scalable environment to globally distribute data securely, distribute computational load, manage the enterprise, and provide service reliability. Current grid specifications use web services as a foundation, building a grid infrastructure. Command and Control and intelligence mission applications can take advantage of grids in several areas such as data distribution & edge caching (data grids) and massive computational problems in the areas of terrain and weather simulation, image processing, planning, intelligence preparation of the battlefield (IPB), and entity/track correlation (computational grids). The NECC grid can capture unused computer resources to increase efficiency, the sharing of data resources, and the ability to create virtual command and control centers at greatly reduced costs, and higher fault tolerance.

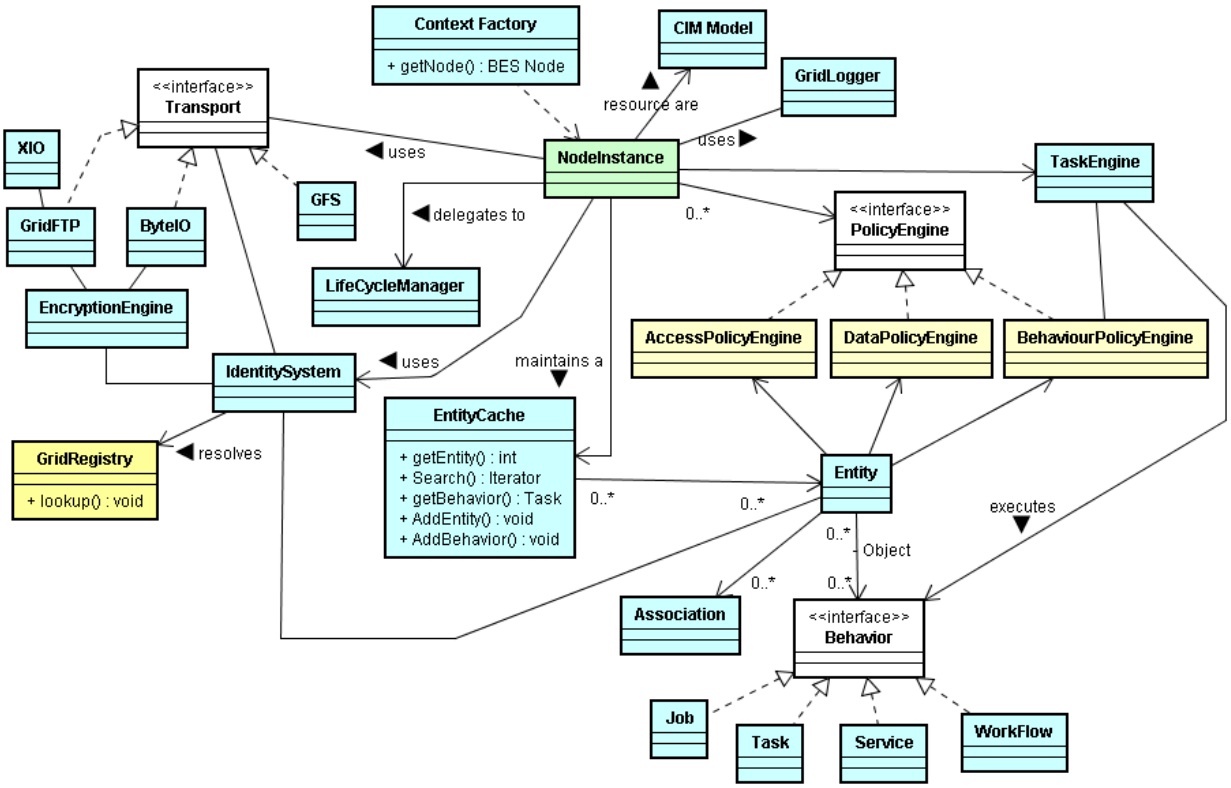


Figure 4-1 Sample NECC Grid Node Design

Figure 4-1 depicts in UML an example high level design view of what each grid node (appliance) might contain. There are modular components which have specific responsibilities to ensure that services are properly being executed and that data is available. Policy engines provide a security context and are used to validate data movement, execution, and access control[13]. Policy is the commanders intent and is expressed via a policy language.

Each node communicates via a secure standard transport interface with other nodes to provide cross-node interoperability and exchange of messages (message/service bus). This communication is in addition to normal service behavior as might be rendered in a Behavior using SOAP over HTTPS or TLS.

Since this is a rough initial design, it is not comprehensive and needs re-factoring which would be accomplished during a dedicated NECC design phase to leverage industry and standards body work products within the context of a NECC Design Team. A Design goal is to ensure plug-n-work interoperability with other grid systems. Globus would be a likely base from which to start an implementation as many of the components are already available in the Globus Toolkit. A UNIX variant would be best suited for scalability and security reasons. However the design would not need to be tied to a specific operating system and would in fact support multiple operating system variants and underlying chip designs.

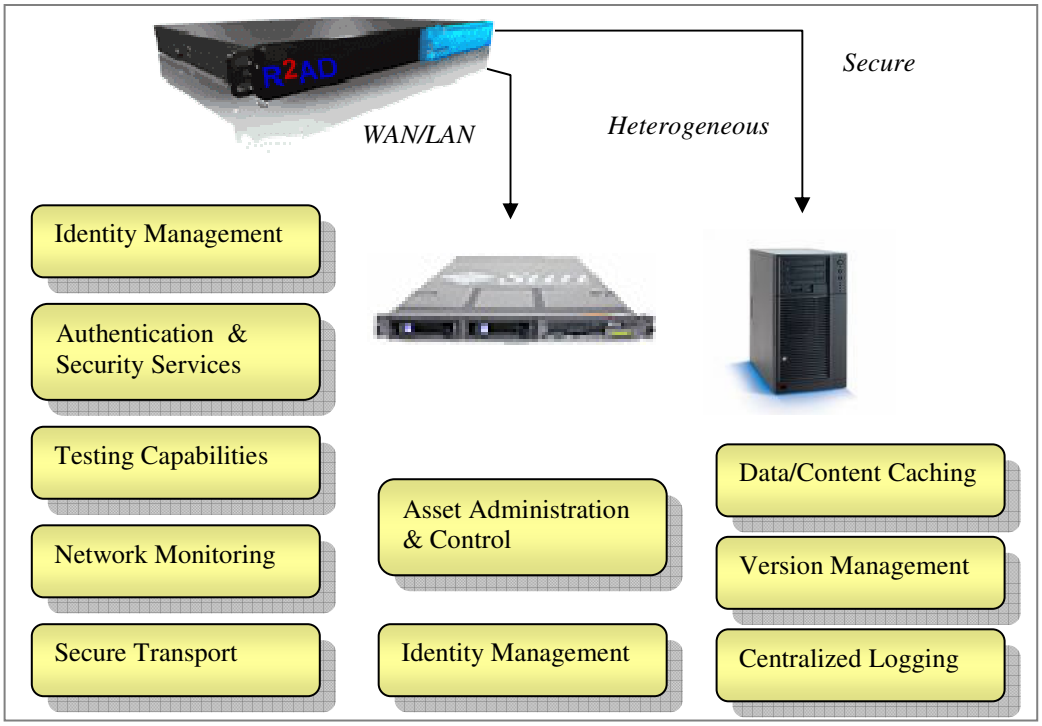
R2AD encourages NECC to lead this design effort to help ensure interoperability and to reduce the tax payer burden by eliminating duplicity in the current stove pipe implementations that might arise from the acquisition policies.

[13] Policy-Directed Data Movements in Grids: www.cs.virginia.edu/~jf4t/jfeng-policy.pdf

4.4. *Development of a Management Grid Appliance*

An appliance is a way to field hardware and software together to provide a manageable factory or presence in the field which can help maintain, monitor and provision the compute model from the core to the edge. This appliance would be a physical implementation of the node design and would meet all key performance and security criteria.

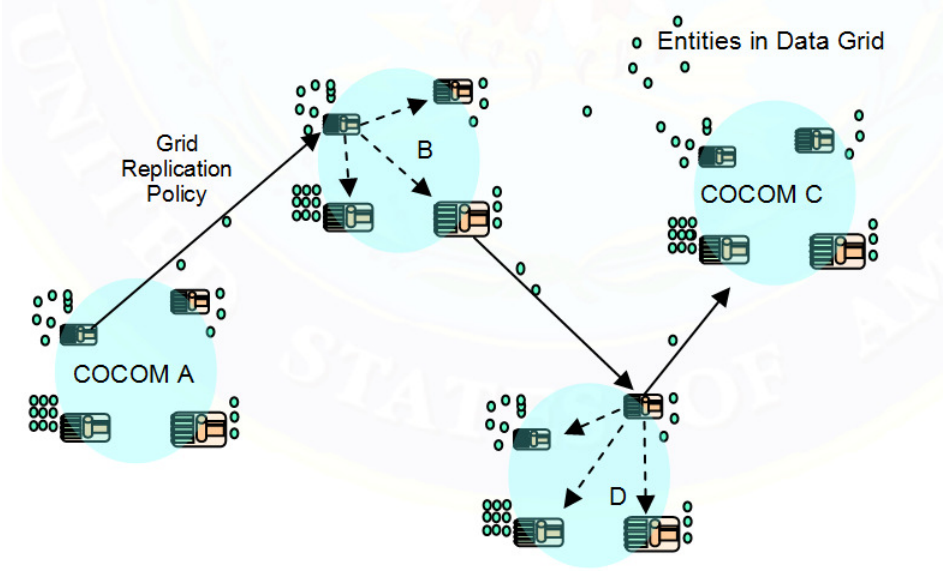
Previous briefings to DISA have proposed the concept of a provisioning appliance. The diagram below captures the possible layout of the appliance and the primary functions which it can provide. This information was also recently briefed at OGF-18 and well received by the Application Content Services (ACS) working group and the OASIS SDD working group representative. A description of deployment models follows using information from the CDDLM specification which has been submitted to the OGF recently by the CDDLM working group.



Notional Grid Appliance

4.4.1. Service Deployment via the Grid Appliance

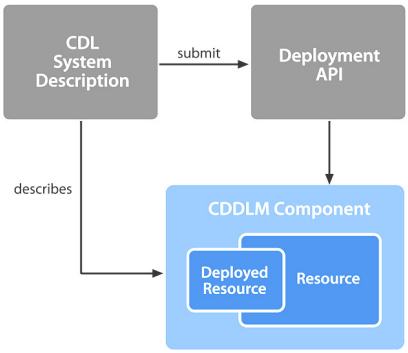
Grid standards are being developed with the support of commercial companies such as MacroVision, IBM, MacroVision, and others to help bring forth a standardized automated enterprise. The appliance would take advantage of Globus as a foundation and build on top of it an implementation of ACS. The ACS engine could then use OASIS Solution Deployment Descriptor (SDD) as needed to deploy software or Configuration Description,



Deployment, and Lifecycle Management (CDDLM). These standards provide engineering and administrators standardized data structures and programmatic methodologies for describing entire systems in XML, and packaging the XML and deployment artifacts via a web network service such as ACS. ACS also supports embedded WS-Policy, which can be used to control how the contents are used or replicated. This same model or pattern is also applicable to data or entities in the grid.

An Integrated Developer Environment (IDE) could be aware of the Solutions Deployment Descriptor XML file format, which is designed to describe the deployment of entire systems, unlike the COE, which was designed to track software dependencies on a box by box basis. This XML descriptor file, along with the network service for packing the XML and binaries into "System/Application Archive" provides the basis for two things:

- Full automation of software provisioning/installation
- Offloading of installation design and physical/manual installation burdens from Administrators.



CDDLM follows the pattern of defining a specification schema and templates from which instances are created:

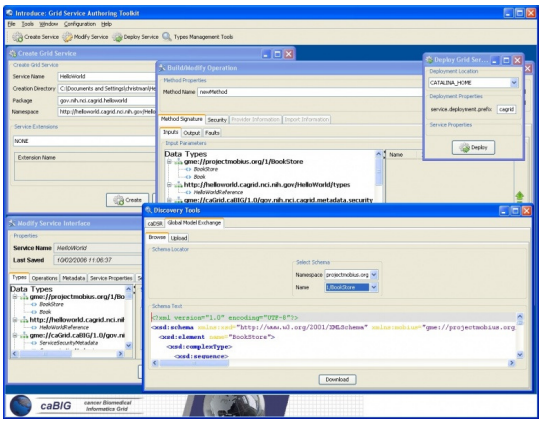
- A Component Description Language (CDL) document describes a deployment
- It is submitted to a node implementing the Deployment API
- This node is then responsible for the deployment lifecycle of the Components described in the CDL document

SOA requires definitions of dependencies at a system level, or at sets of system dependencies. Because the grid (as defined by GGF/EGA/I. Foster, et al) sees all software as jobs to be deployed automatically into the grid (not to specific boxes), it is defining the mechanisms for autonomic provisioning such as those embodied in the component model of the Configuration Description, Deployment and Lifecycle Management (CDDLM)[14] specification and current implementations. Keep in mind the grid can be sets of collections of heterogeneous hardware from multiple administrative domains. In the CDDLM specification for the description and deployment of service:

- A System is a group of 1…n Components that each contain 0…n Services
- Services are defined as WS-I or WS-RF compliant, but must have at least one valid WS-Addressing Endpoint Reference.

It is very important to also take into account backward compatibility in the design of services as it must be engineered into software through the implementation of requirements to do so. Version Management must be considered and the appliance must support the ability to deploy and support multiple versions until it is proven that all dependencies for a particular version have been dissolved. The SDD specification can be used with the description of these system and version dependencies.

4.4.2. Development Tools for Grids



The appliance would also be shipped with the full Globus engine and developers can use it to test their services in preparation for integration testing and eventual fielding as dictated by FDCE.

Tooling is the implementation of automated means to help the developer and administrators create, use, and manage a software technology. In the case of normal Web Services, many Integrated Developer Environments (IDEs) now provide built in support for services or plug-ins which can be added to tailor the product (Eclipse, etc) to the needs of the project.

The Grid community has also invested time and energy into tooling for grid applications. The "Introduce" project is one example which builds a simple tool to generate Grid Services.[15]

[14] CDDLM Architecture Slides: https://forge.gridforum.org/sf/go/doc12763
[15] Introduce Project: http://dev.globus.org/wiki/Incubator/Introduce

At the Grid World event, we saw this tool introduced. It provided not only the ability to create services, but also to maintain the meta-data about the service by generating Meta-Data-Service (MDS) hooks. NECC really needs to start using this for all generated services.

Another tool demonstrated was Grid Develop Tools (GDT)[16] which provided full integration into Eclipse and provides a Model Drive Architecture (MDA) approach to the development of Grid Services (see figure below). GDT is developed by Thomas Friese to take advantage of Java 1.5's new feature: Annotations. This technique which is quite revolutionary, is a great example of how DoD Wide services can be rendered using Java.

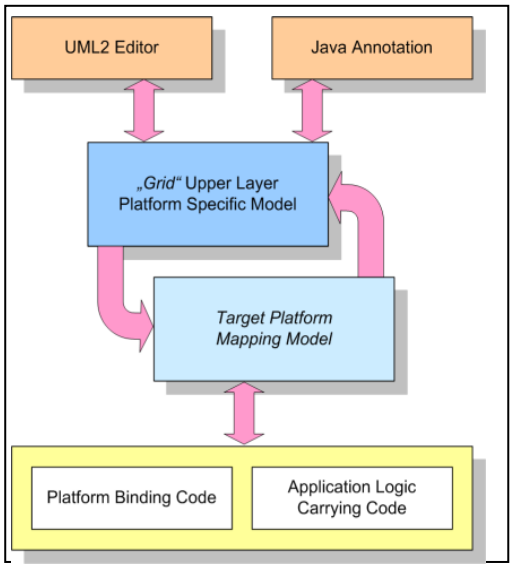


Figure 4-2 MDA approach

Since the current development cycle is complicated, as DISA is familiar with, there is on-going research into the tooling area by the WS and Grid communities[17] (and others of course). The immediate benefits of tooling can be applied to NECC to help developers efficiently generate grid compliance services which could run in the single standardized grid engine appliance.

Ian Foster, the father of grid computing, mentioned the importance of automation of grid management in one of his recent speeches. His blog site (http://ianfoster.typepad.com/blog) also makes references to this need and his efforts are starting to take root in projects such as caBIG in the National Cancer Institute (NCI) which is a key sponsor of Grids.
Some other developer tools that were mentioned during Grid World are available via these links:

- http://www.geclipse.org/
- http://dev.globus.org/wiki/Java_WS_Core_Eclipse
- http://mage.uni-marburg.de/trac/gdt/wiki

5. Conclusion

Grid systems offer many solutions to distributed systems which are composed of complex interactions which need to be rapidly composed to solve the needs of the commanders in the field. In general, emerging technologies such as enterprise computing, web services, semantic technologies, and grid computing offer the potential to make transformational advances in DoD's ability to rapidly field interoperable distributed systems such as NECC.

In order for NECC to be an effective system, it must support rapid integration of heterogeneous, geographically distributed computing facilities at data centers and at the Combatant Command centers and their subordinates. As an example, heterogeneous distributed track/entity processing and simulations require advanced geo-spatial and time management algorithms to ensure proper synchronization of data exchanged among subsystems. This challenge is directly in line with the capabilities of grid systems, which offer scalability to handle hundreds of millions of transactions with fault tolerance and high availability. Grid systems are compatible with DoD standards and can operate effectively in the Global Information Grid.

Based on the research performed under this task and the previous Grid Study task, R2AD strongly supports the adoption of grid technology in NECC.

6. References

1. [OGF] Open Grid Forum http://www.ogf.org
2. [Globus] Globus Toolkit and the Globus Alliance http://www.globus.org/
3. [OGSA-WG] Open Grid Services Architecture WG http://forge.gridforum.org/projects/ogsa-wg
4. [Grid Study] First DISA Grid Study: http://www.r2ad.com/whitepapers.html

[16] GDT support Page: http://ds.informatik.uni-marburg.de/MAGE/gdt/
[17] Borja's Mini-Symposium on tools: http://people.cs.uchicago.edu/~borja/gw06/devtools/

APPENDIX A

The following business cards have been scanned and are provided for reference.

Also, please note this UNIVA POC:
Steve Minisini, UNIVA
847-644-1050
Minisini@univa.com

APPENDIX B

GridWorld/OGF-18 Trip Report
DC September 11-15, 2006

For

Mark J Kuzma
Chief Engineer
Net-Enabled Command Capability (NECC)
DISA
PEO for C2 Capabilities
P.O. Box 4502
Arlington, VA 22204-4502

by
Michael Behrens, R2AD, LLC, CTO
Report Date: 25 September 2006

6.1. Summary

On behalf of DISA NECC PMO, I attended and participated in the Globus World and OGF-18 event. I spent most of my time in two of their main tracks: Standards and Technology. Globus, an open-source GOTS grid engine, celebrated their 10th anniversary.

The standards track revolved around the new Open Grid Forum which is now merged with the Enterprise Grid Alliance. What this means is that now more commercial support shall be present in the OGF wider community.

The other tracks offered a chance to catch up and learn more about the current state of the Globus Toolkit version 4 and other grid technologies. I attended a tutorial in which I installed a grid container and deployed web services to it which then were modified to participate with another service to form a collaboration environment (sharing notes).

Overall the conference content was excellent. I have more information which I will continue to pass onto the leadership at DISA to help bring the best capabilities to the war fighter. With support from NECC, I'll work to deliver a grid enable system capable of providing a flexible, secure, and scaleable information technology infrastructure for the warfighter.

I strongly recommend and encourage more DISA attendance in the future as there were so many briefings and because of the track schedule I was unable to cover everything.

A collection of materials (slides) that I've been able to collect is being made available on a CD.

Details of the conference are on the following pages.

For additional information about the contents of this report, please contact Michael Behrens at this e-mail address: Michael.Behrens.ctr@DISA.mil

6.2. Conference Details

6.2.1. EBAY

One of the most interesting key-notes was presented by eBay where Paul Strong (eBay Scientist) explained that once they realized that they actually were a grid they began transforming their architecture to ensure high reliability (reduce number of CNN moments) which is posturing them for adoption of more standards as they become adopted. eBay is therefore supporting OGF efforts which now shall incorporate the EGA model.
Here are some links of notes/reviews of his keynote:

1. http://www.gridtoday.com/grid/875717.html
2. eBay Sold on Grid: http://www.eweek.com/article2/0,1895,1640234,00.asp

Two week development cycle = speed: eBay has a very dynamic development model which allows them to create their site (via their home-grown grid development infrastructure "ICE Grid") in a very quick amount of time (basically the compilation, translation, HTML generation, etc. are parallelized tasks). This is then published to a QA zone where it can be tested, etc. Then, it is made available to the operational site which incorporates it every two weeks. Their architecture is very good, dividing up the visualization from the indexes, etc. I'll work to find a copy of his presentation. His big message was the need for standardized "Manageability", something that OGF, OASIS, and DMTF are working towards.

Globus Toolkit
The version of Globus which is out is GT4. It provides a C++ and a Java version, and there are others as well (Python namely). I attended several briefings about its current state and can report that GT4 is ready for use by DISA today to provide a secure service infrastructure. As part of the piloting effort, I would suggest that the grid infrastructure ECMs be developed using GT4. There is also a book out now available from Amazon, etc.

I attended an introduction tutorial (which I can present a demo of if desired!) in which I installed GT4 and deployed several Grid Services which were in fact web services. The Globus model basically uses web services as endpoints for jobs and applications. GT4 is the best "service container" in the world.

One of the presentations I attended showed some very nice Model Driven Architecture (MDA) approaches to using Globus. These are open tools available for free use. I was really amazed at some of them and will recommend that some of those tools are spot-on with where developers should be heading. It makes working with the grid very easy.

Ian Foster mentioned in his Keynote that "Grid 2.0" is now here – meaning that grid computing is much more than job scheduling, rather it is virtualize everything via SOA. (second generation grids). Ian also mentioned the need for dynamic deployment – something akin the GOE slides and ACS, etc. He used terms such as "Service Delivery". In the ACS-WG meeting, I presented slides which also called for this in the form of a new specification perhaps: WSRD for determining the location of resources even if they are yet deployed or available.

Service State
One of the announcements was that there is finally agreement in the standards about state. State is important and many of the WSRF (very powerful) and capabilities which are currently being used in grid infrastructures might eventually merge/transition[18] where needed to a new specification called WSRT[19] (WS-ResourceTransfer). This will enable future consolidation between the many currently different ways to manage service state and lifecycle. Along with this, an updated WS-MetadataExchage has also been published. I strongly recommend that NECC start considering this big impact design and architecture. See the references or Google for more information.

Future Management stack

world WSDM on

Security

[18] http://download.boulder.ibm.com/ibmdl/pub/software/dw/specs/ws-wsdmmgmt/wsdmmgmt.pdf#search=%22WSRF%20WSRT%22
[19] http://www-128.ibm.com/developerworks/webservices/library/specification/ws-wsrt/

WS-Security and XACML along with proxy certificates are proving to be very valuable in GT4. MyProxy is a powerful capability to support Single Sign-on and cross-grid interactions. GT4 also comes with a Certificate Authority (SimpleCA).

GridFTP
I met with the GridFTP developers to gain some insight into putting together a demo for DISA. They agreed that it would take a few weeks for a newbie to build and establish a good WAN secure (PKI) GridFTP demo. William Allcock is the tech lead for GridFTP at Argonne. Reliable Transfer via a SOAP protocol could also be added to ensure that FTP requests survive hard crashes, etc. This is what the Reliable communications ECM I3 submitted is meant to implement.

Univa
Univa is the commercial arm of Globus. They can provide direct support for Globus and also provide custom versions or service support as warranted. I spoke with their CEO, Steve Tuecke[20] who introduced me to a few on his team that would be able to join us at DISA for a meeting if desired. I'd like to help set up the "Grid Day" at DISA. They can provide an InstallShield version of Globus for windows or equivalent for Linux or Solaris, documentation, etc. R2AD will work to become a Univa partner to help support DISA.

DoD Supports OGF
I noticed that on the supporting sponsor poster for OGF that the Department of Defense was a supporter. I spoke with their director and VP of operations, Steve Scrumb. He indicated that Jerry Smith called one day and said they'd like to support OGF. I hope that this was the result of Mr. Kuzma's e-mail last year. This would be great. Steve also agreed that the DoD would be able to gain a lot from participation in the OGF. I would suggest that some of the ECMs could even become specifications since their use can be generic and directly applicable to many other organizations.

[20] http://www.itindepth.com/Univa_Interview.htm