Grid Technology and Information Management for Command and Control

Dr. Scott E. Spetka
Dr. George O. Ramseyer*
Dr. Richard W. Linderman*

ITT Industries
Advanced Engineering and Sciences
SUNY Institute of Technology
PO Box 3050
Utica, NY 13504
scott@cs.sunyit.edu

*Air Force Research Laboratory/Information Directorate
26 Electronic Parkway
Rome, NY 13441-4514
George.Ramseyer@rl.af.mil
Phone: (315) 330-3492
Richard.Linderman@rl.af.mil
Phone: (315) 330-2208

Abstract

Over the next few years, the Command and Control (C2) community will face increased demands for processing data as increasing numbers of data sources are deployed. It is expected that skilled analysts will continue to play an important role in gleaning useful information from imagery and other sources that are by and large produced as digital multimedia. Using current information systems would require significant increases in the numbers of analysts needed to manage the increased workload that is expected. Because the number of analysts can be expected to remain fairly constant over the next few years, scalable and extensible C2 processing technology will be required to achieve productivity gains needed to keep up with increasing demands. Two complementary information systems, a Distributed Processing (DP) system and an Information Management (IM) system, form the basis for a new architecture that can meet this requirement. A combined system (DPIM) that addresses the performance and the scalability required for future C2 systems is introduced in this paper.

Introduction

In responding to changing battlefield conditions, commanders require the immediate transformation of data into information. Two complementary information systems, a Distributed Processing (DP) system and an Information Management (IM) system, form the basis for a new architecture that can meet this requirement. This combined system (DPIM) addresses the performance and the scalability required for future C2. Under DPIM, increased processing power can be transparently added to the communications backbone and to high performance computer (HPC) systems to address increased

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **JUN 2005**	2. REPORT TYPE	3. DATES COVERED **00-00-2005 to 00-00-2005**
4. TITLE AND SUBTITLE **Grid Technology and Information Management for Command and Control**		5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Air Force Research Laboratory,Infromation Directorate,26 Electronic Parkway,Rome,NY,13441-4514**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR’S ACRONYM(S) 11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**33**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

processing and data flow requirements. The DPIM system is extendable to allow new codes to be added that will improve productivity.

The DPIM architecture extends an existing DP/IM-based FrameWork architecture [1-8] by incorporating Globus grid technologies [9] to enhance the control of distributed resources. The FrameWork demonstrated that a set of core publication and subscription services can provide a base information management platform for system development to support C2. The FrameWork used HPC status publications to direct processing over available processors at distributed HPC Centers. The DPIM system uses a catalog of published processing-related information objects to generate processing plans. Additional information published by grid services includes resource reservations, security requirements, etc. The IM subsystem in the original FrameWork architecture handled communications from HPC-based agents to Common Object Request Broker Architecture (CORBA) code-specific objects [10], regarding HPC status, new data object availability, processing requests and results. The new architecture uses the IM system to support both front-end and back-end grid services. The paper presents the DPIM architecture, discusses scalability and *net-centricity*, and concludes with a discussion of Globus grid support for DPIM.

The DPIM Architecture

Figure 1 illustrates the DPIM Architecture. The architecture is implemented across client systems, the internet and HPC systems. Adaptor services for publish and subscribe can be implemented in the client and HPC systems or can be based remotely on the Internet. Clients communicate directly with grid services to implement administrative functions. For example, an authorized client can request a change in the publication rate for status information or can query HPC job launching services directly to check resource availability. Interfaces for grid services can vary according to the service provided and the implementation. Experimental implementations, enhanced functionality and custom features require variations in the supported interface.

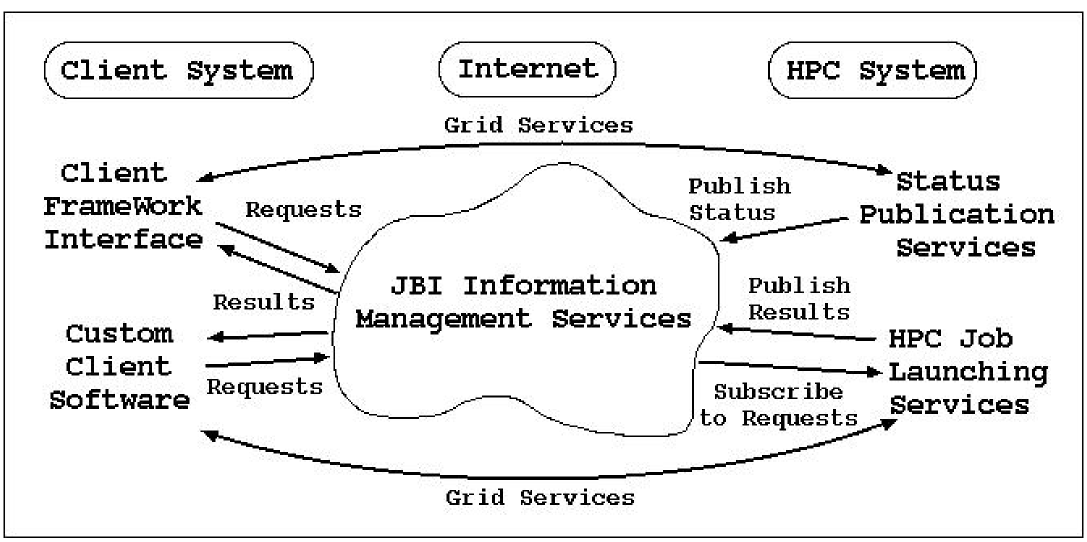


Figure 1. The DPIM Architecture

Both the client's software and HPC services use the JBI information management services [11] for publication and subscription. Clients publish requests for service and subscribe to results. Requests and results are information objects that are routed through the JBI, allowing the system to adapt to dynamically changing resources. HPC-based services can both publish and subscribe as well. Status publication services do not currently subscribe but could subscribe to information objects that control publication rate, information content, logging functions, etc. The HPC job launcher receives requests, and then dispatches them for execution and publishes the results. When an HPC center goes off-line due to network failure or scheduled maintenance, processing requests can be routed to other HPC service subscribers (service providers).

Scalability

Grid Technologies enhance scalability for our new architecture. Part of the DP subsystem was also previously based upon CORBA technology. It implemented standard interfaces to CORBA distributed objects to access spectral and other parallelized codes executed on high performance computers. By replacing the CORBA software with the Globus Toolkit software, the new DPIM architecture enhances scalability. For example, by allowing limits to be placed on services, load balancing is encouraged. Grid services monitoring abets in determining when additional resources become available to support HPC-based applications. Grid services also replace the part of the FrameWork DP subsystem that implemented HPC-based agents that processed work requests from HPC code-specific CORBA distributed objects. HPC-based grid services perform code-specific pre- and post-processing. In addition to submitting codes to HPC's for execution, the services return the results to DPIM front-end service objects.

The new DPIM architecture enhances basic functionality and supports scalability to satisfy C2 performance requirements. Grid software enhances both front-end, code-specific user interface implementation and backend services. In each case grid software enhances the ability to manage services. Using a single IM system for communication of processing requests and results as well as HPC status information assures scalability for all communications. Overall DPIM scalability depends on the scalability of the IM and DP subsystems. DP scalability is assured by the ability to initiate and manage grid services as needed. Scalability of the underlying IM subsystem depends mainly on the ability to increase data flow by adding hardware resources. In addition, "fuselets" [12] can be implemented that can be used to drive enhanced services and to build new services, based on existing information objects.

Figure 2 shows the implementations that we are comparing in this work. The new Globus grid-based implementation replaces the CORBA middleware. Middleware is used in the architecture to extend the FrameWork Web services for each of the code interfaces supported. In addition to issuing processing requests and retrieving results from the grid services code interface, FrameWork Web services can administer grid services, enabling and disabling them dynamically. New grid services implementing the code interfaces for codes as well as code launchers can be shut down for maintenance or turned on to service high volumes of traffic. The grid services-based code launcher can be dynamically configured to implement changing processing policies and priorities.

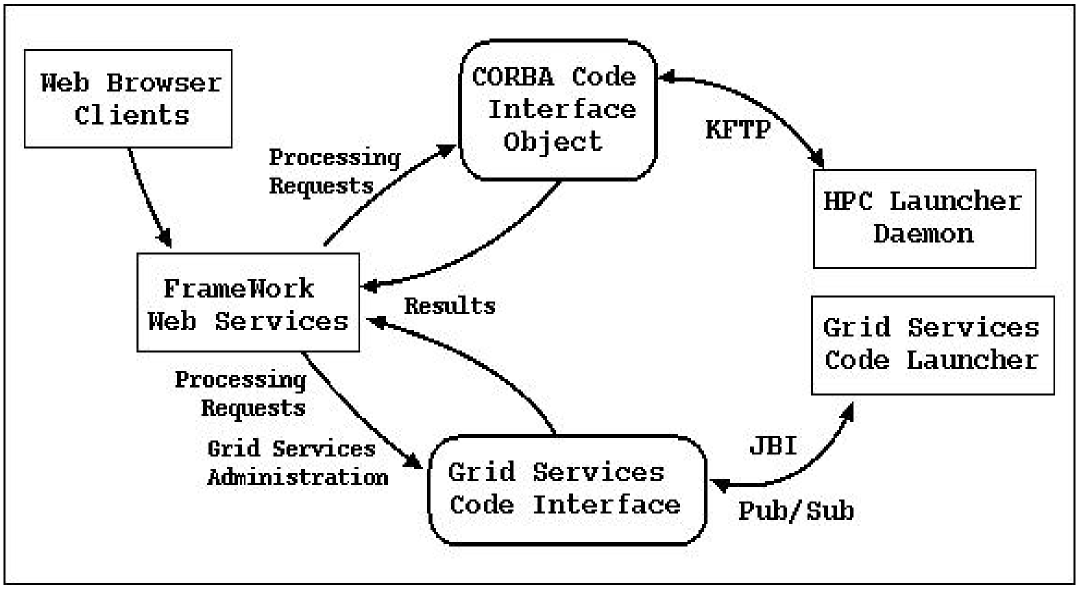


Figure 2. CORBA v. GLOBUS Middleware

Net-Centric Issues

Network-centric systems [13] support warfare in the modern information age by enhancing shared situational awareness and effects-based operations and generate increased combat power by networking sensors, decision makers, and shooters. The success of net-centric systems depends on maintaining software interoperability while adopting new technologies. The new FrameWork architecture demonstrates that it is robust and flexible enough to support new technologies while continuing to support legacy codes that have already been implemented for the CORBA-based implementation.

Distributed systems have grown to a scale that could not have been foreseen 10 years ago. CORBA standardized interfaces across heterogeneous platforms for high-performance interfaces where complex data types needed to be translated. New interfaces for Web applications require significantly less data marshalling services since most data is character type. Modern Web services implement significant enhancements for data discovery and functionality to allow dynamic binding to interfaces for accessing objects. The Globus grid technology incorporates Web services into an environment for distributed system development that supports security, service discovery, service management and other functions in a robust environment. It offers the tools to support integration of forces in dynamic environments like battlefields.

The information management services of the DPIM are the key to net-centric operations. Figure 3 illustrates the use of the JBI information management to support shared access to codes to support communities of interest (COI) for net-centric operations. In the diagram, two communities of interest share their technologies and HPC center resources to improve battlefield analysis capabilities. New information is made available to facilitate improved sensor-decision maker-shooter efficiency.

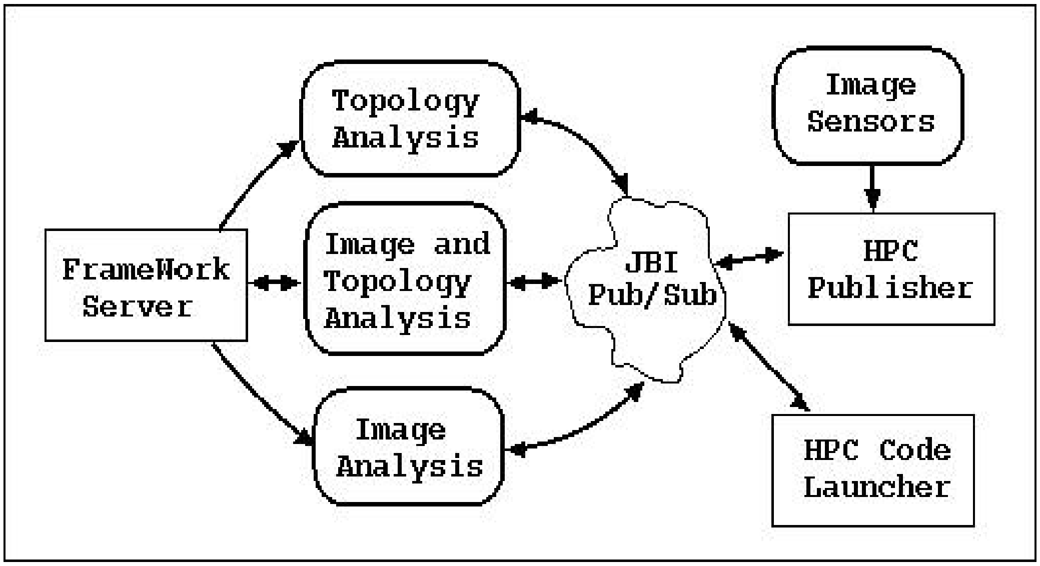


Figure 3. JBI Information Management

Grid Technologies

Globus grid software technologies offer support for distributed processing that enhances service interfaces for DPIM. Adopting Globus grid technologies complements the choice of a pub/sub-based information management system. Globus software is operating system independent to provide uniform interfaces across heterogeneous systems. It supports: reliability and other forms of quality of service (QoS), authentication and delegation of credentials as well as services like notification, dynamic service creation, lifetime management and improved general manageability [14-15]. Grid services also support a robust communication interface [16-17]. This enhances Grid communication with JBI pub/sub services and discusses advantages of pub/sub for workload distribution and access to HPC services. We are also considering adopting grid metadata management services to complement the JBI metadata repository [18].

Figure 4 shows how grid interfaces can be used to manage HPC-based launching services for distributed computations. Notification services enable DPIM operators to monitor incoming processing requests over a set of HPC's. Notifications can include HPC processing load information, available disk space statistics, projected resource requirements, recent network bandwidth usage, etc. The DPIM Grid client can respond by reassigning queued processing requests to other HPC's, changing the frequency of publication for overall HPC status and DPIM service status, requesting additional processors be assigned to DPIM, canceling or suspending running jobs.

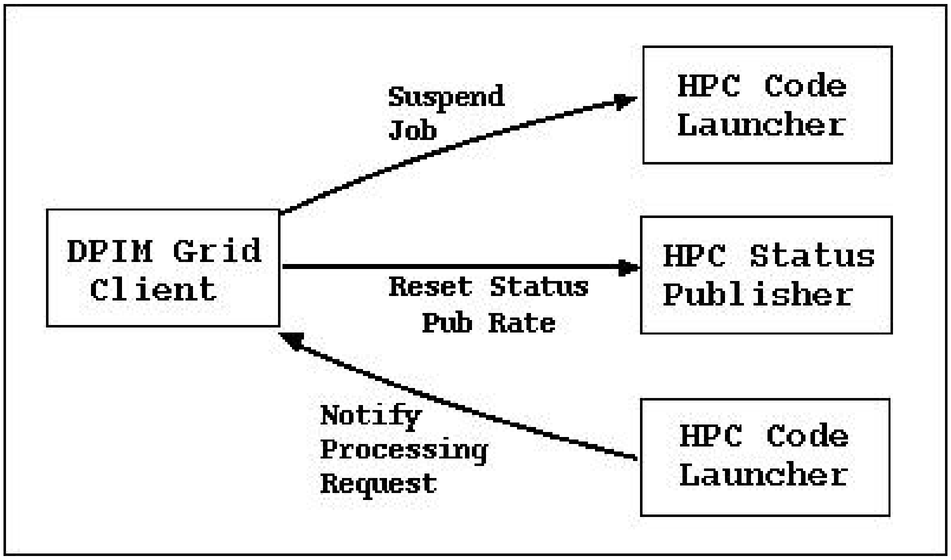


Figure 4. DPIM Grid Interfaces

Conclusions

The Distributed Processing Information Management system serves as a grid portal to access and manage codes needed by the C2 community. The extensible, scalable DPIM system provides for additional processors to be incorporated into the system without disrupting 24/7 service, and is scalable to meet increased demands. Improved access to HPC's through a system such as this is required for C2 systems when real-time processing is essential. A set of codes can be executed in parallel, across multiple HPC centers, using the DPIM framework to provide information in a fraction of the time that would normally be required. This system addresses the need of increased demand for C2 processing that will be required in the next few years.

References

[1] G.O. Ramseyer, S.E. Spetka, R. Linderman and B.C. Romano, **Rapid C4I High Performance Computing for Hyperspectral Imaging Exploitation**, Proc. 6th International Command and Control Conference (2001).

[2] S. Spetka, G.O. Ramseyer, R. Linderman and D. Gilmour, **The FrameWork Architecture for Hyperspectral Image Exploitation**, Proc. 27th Annual GOMAC Technical Conference (2002).

[3] G.O. Ramseyer, S.E. Spetka, R.E. Linderman and B.C. Romano, **The FrameWork: An Open-Architecture for Very Large Image Exploitation**, Proc. 2002 Command and Control Research and Technology Symposium (2002).

[4] S.E. Spetka, G.O. Ramseyer, D. Fitzgerald and R.E. Linderman, **A Distributed Parallel Processing System for Command and Control Imagery**, Proc. 7th International Command & Control Research& Technology Symposium (2002).

[5] G.O. Ramseyer, **Rapid C4I High Performance Computing for the Joint Battlespace Infosphere**, invited paper: Proc. Airborne C4I Conference, London, England (2002).
[6] S.E. Spetka, G.O. Ramseyer and R.W. Linderman, **A Framework for High-Performance Image Exploitation**, Proc. 17th Annual ACM Conference on Object-Oriented Programming, Systems, Languages and Applications (2002).
[7] G.O. Ramseyer, R.W. Linderman, S.E. Spetka, D.J. Fitzgerald and M.J. Moore, **An Open-Architecture Middleware for Hyperspectral Image Exploitation**, Proc. 28th Annual GOMAC Technical Conference (2003).
[8] S.E. Spetka, G.O. Ramseyer and R.W. Linderman, **Redeveloping a High-Performance Computing FrameWork**, Proc. 18th Annual ACM Conference on Object-Oriented Programming, Systems, Languages and Applications (2003).
[9] http://www-unix.globus.org/toolkit/
[10] http://www.corba.org/
[11] http://www.rl.af.mil/programs/jbi/
[12] http://www.fuselet.org/
[13] Office of Force Transformation, **The Implementation of Network-Centric Warfare** http://www.oft.osd.mil/library/library_files/document_387_NCW_Book_LowRes.pdf
[14] I. Foster, C. Kesselman, J. Nick, S. Tuecke. **Grid Services for Distributed System** Integration, *Computer*, 35(6) (2002).
[15] I. Foster, C. Kesselman, J. Nick, S. Tuecke, **The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration**, Open Grid Service Infrastructure WG, Global Grid Forum (June 22, 2002).
[16] B. Allcock, J. Bresnahan, R. Kettimuthu, M. Link, C. Dumitrescu, I. Raicu, and I. Foster, **The Globus Striped GridFTP Framework and Server**, *Submitted to the 2005 High Performance Distributed Computing Conference (HPDC 14)*.
[17] W.E. Allcock, I. Foster and R. Madduri, **Reliable Data Transport: A Critical Service for the Grid**, *Building Service Based Grids Workshop, Global Grid Forum* (11 June 2004).
[18] E. Deelman, G. Singh, M.P. Atkinson, A. Chervenak, N.P. Chue Hong, C. Kesselman, S. Patil, L. Pearlman, and M. Su, **Grid-Based Metadata Services**, *16th International Conference on Scientific and Statistical Database Management (SSDBM04)* (June 2004).

Grid Technology and Information Management for Command and Control

Scott Spetka - ITT Industries and
SUNY Institute of Technology
George Ramseyer - AFRL/Information Directorate
Richard Linderman - AFRL/Information Directorate

The Tenth International Command and Control Research and Technology Symposium

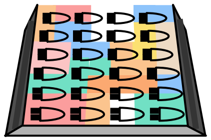
June 14, 2005 1:00pm, Room 4

Outline

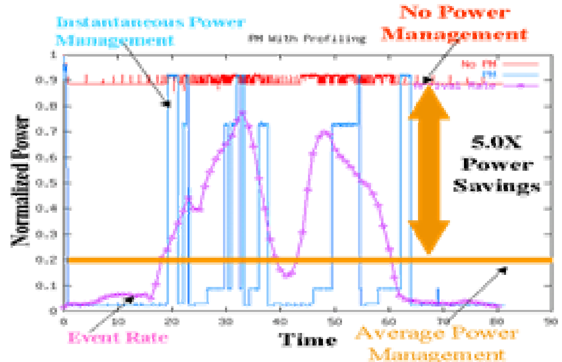
- Introduction
- Architecture
  - Sharing Resources
  - Scalability
  - Globus Grid Services
  - Interactive Model
- Net-Centricity
- Security
- Conclusion

Emerging HW Architectures

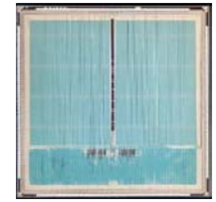
DoD C2 SW requirements lag HW power



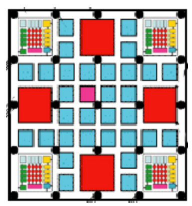
Adaptive/Reconfigurable Computing Systems (1991-2010)



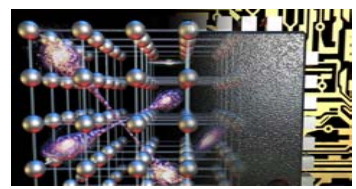
Power Aware Computing (1999-2006)



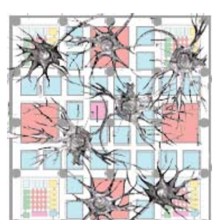
Data Intensive Computing (1998-2006)



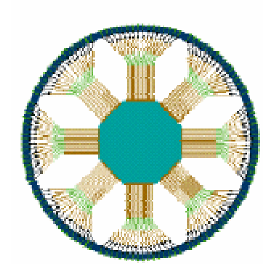
Polymorphous Computing Architectures (1999-2015)



High Productivity Computing Systems (2001-2012)



Cognitive Computing (2003-2023)



Peta-Scale Systems (2005-2018)

Hyperspectral Imaging Portfolio

ITT Industries
Engineered for life

Objective: To develop a scalable, portable, high performance computing software framework with 8 algorithms for rapidly accessing and processing hyperspectral data.

Impact

- **Near Real-Time Access to Hyperspectral Imagery Data for Battlespace Awareness**
- **Imagery Products for Intelligence Analysts and Battlefield Decision Makers**

Approach

- **Leverage serial DoD hyperspectral imagery codes to rapidly process raw hyperspectral data to produce imagery products**
- **Client for JBI to provide rapid query, publish and subscribe with secure webserver application interface**

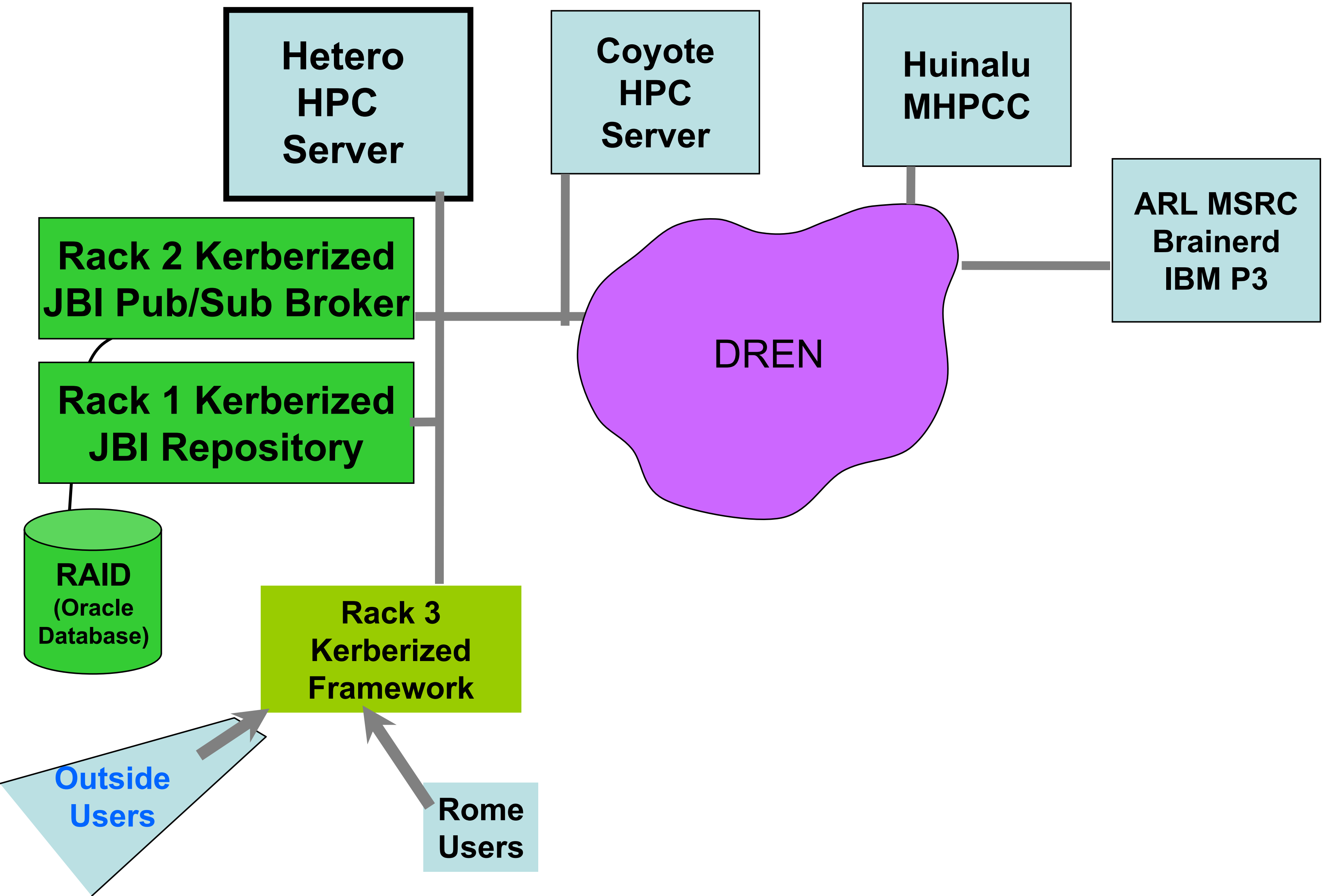
Status

- **8 Codes Tested and Integrated with Framework**
- **JBI Core Services Demonstrated**

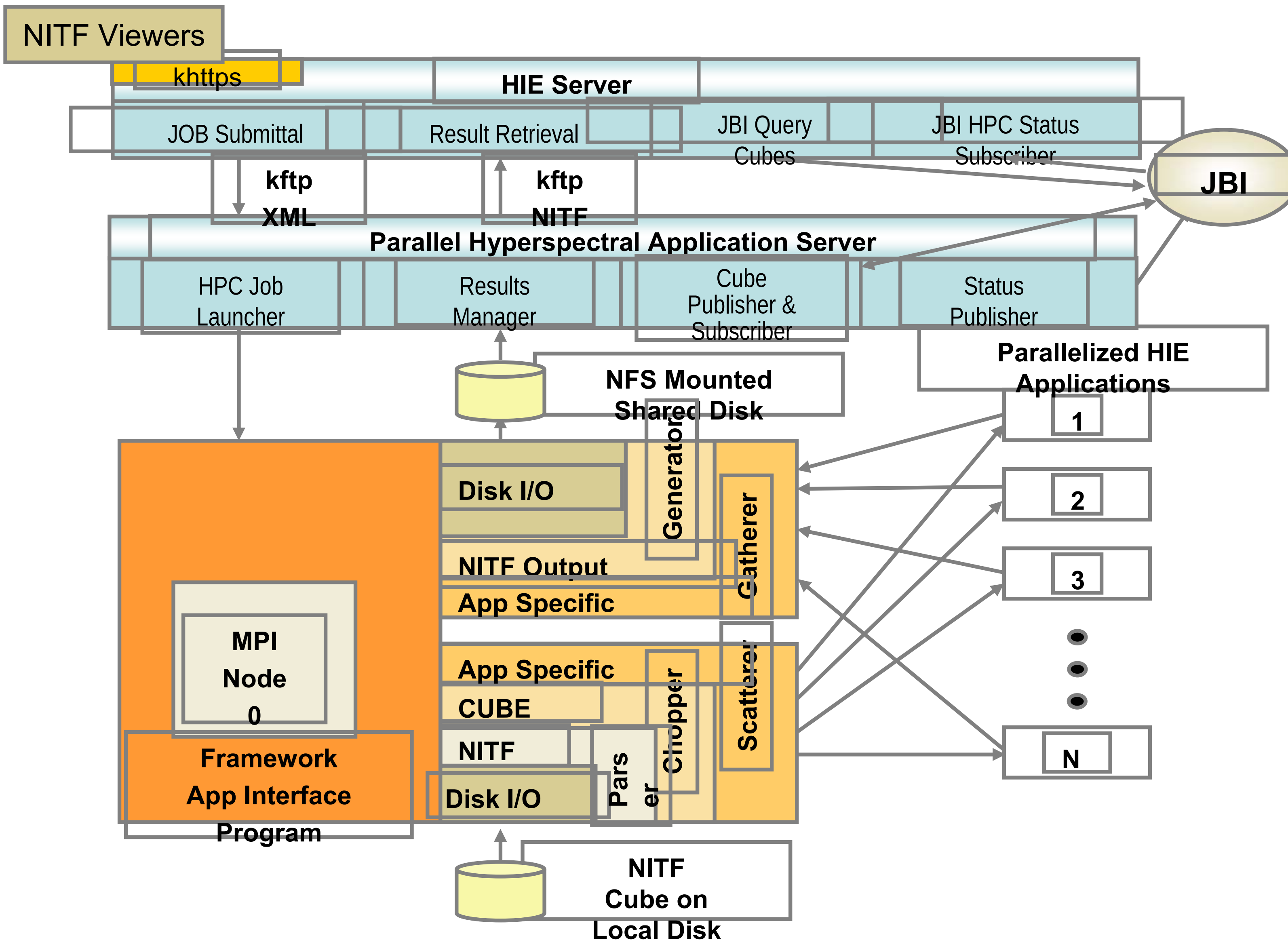
Hyperspectral Cubes

Cube Characteristics for Hyperspectral Image Exploitation						
	Sensor	Format	Size	Samples	Lines	Bands
Cuprite	Avaris	Bil	140MB	614	512	224
Jasper	Avaris	Bip	140MB	614	512	224
Virgin	Avaris	Bip	140MB	614	512	224
FortHill1	Nvis	Bip	181MB	256	6798	26
FortHill2	Nvis	Bip	172MB	256	6456	26
FortHill3	Nvis	Bip	138MB	256	5178	26
FortHill4	Nvis	Bip	165MB	256	6198	26
FortHill5	Nvis	Bip	128MB	256	4812	26
Terrain	Hydice	Bil	64MB	307	500	210
Urban	Hydice	Bil	40MB	307	307	210
Cube	Nvis	Bip	2GB	256	256	384

HIE Test Framework Grid

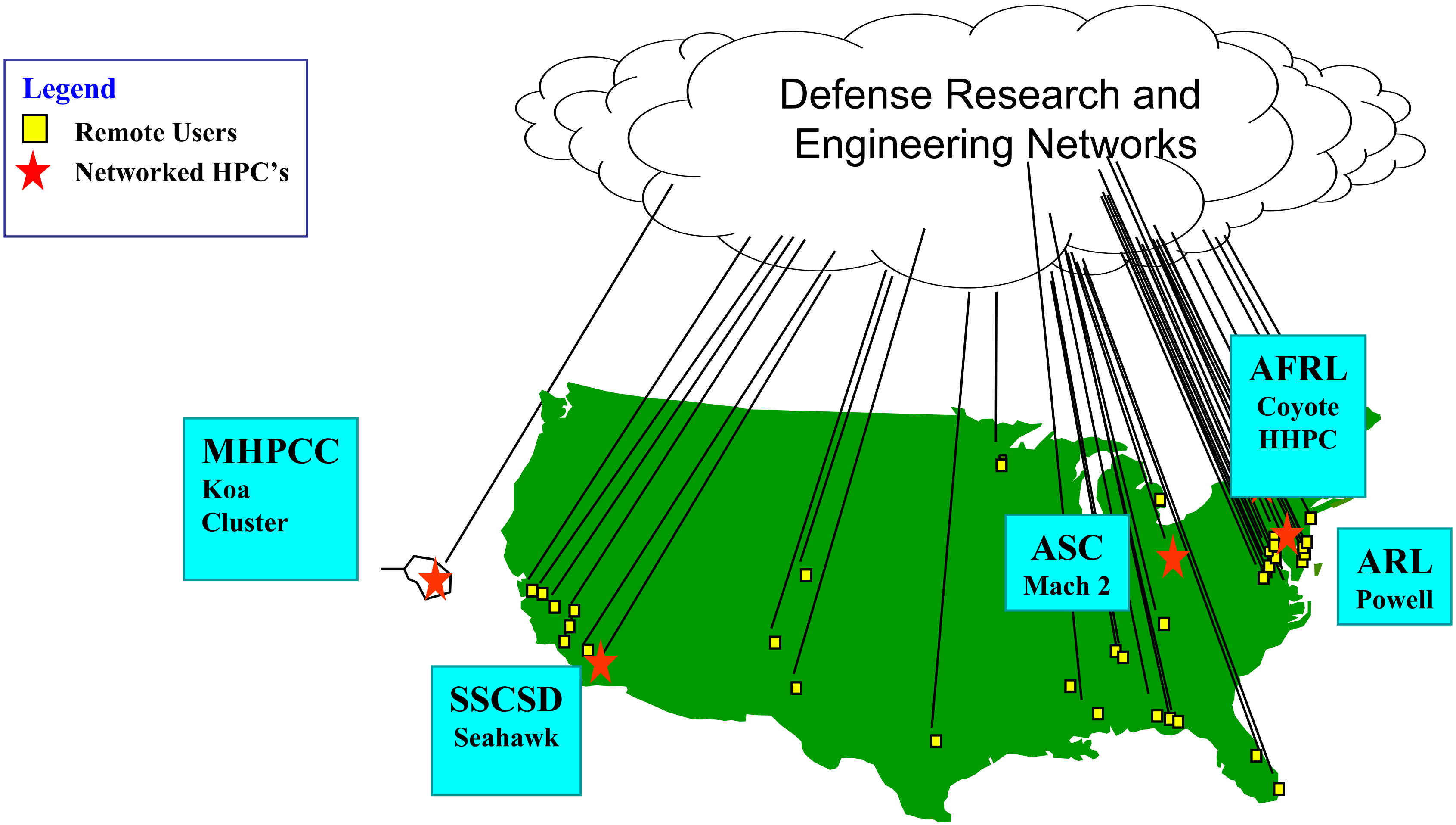


HPC/JBI/Server Interplay



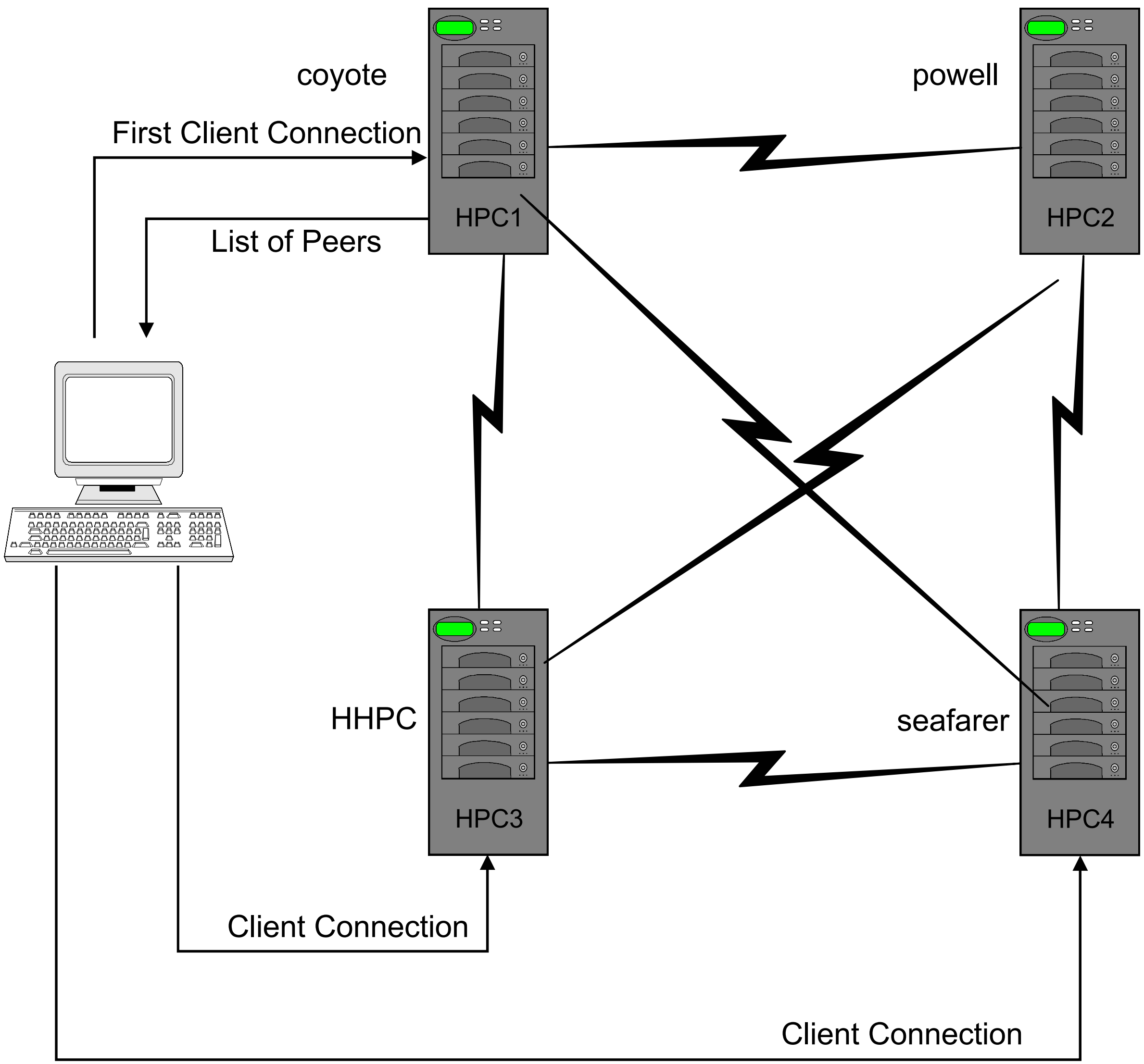
NSF TeraGrid

DPIM Testbed



- ✓ Distributed HPC's
- ✓ Accessed by authorized users anywhere on the DREN and Internet
- ✓ Interactive and time critical problems

DPIM-DREN Interactive Network

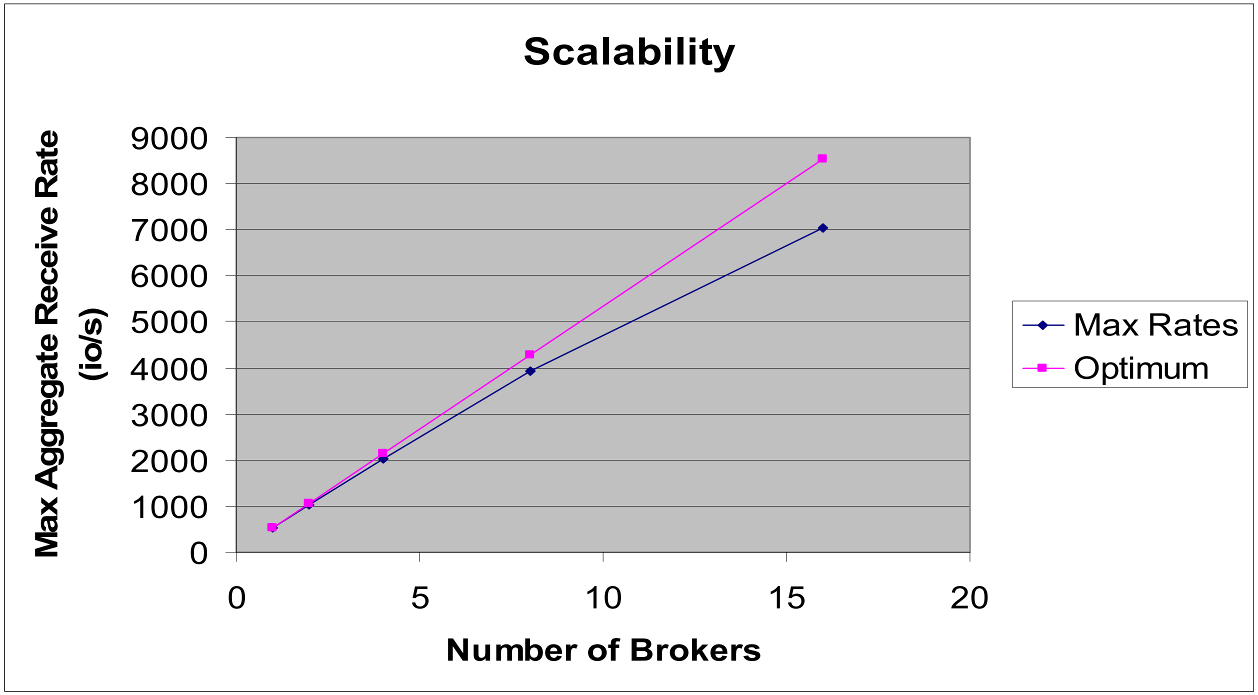


- HPCs are all interconnected
- Client gets list of HPCs from first connection
- Client maintains up to three connections
- Dynamic load balancing between HPCs based on I/O type and activity
- The client receives copies in triplicate of each publication

Joint Battlespace Infosphere

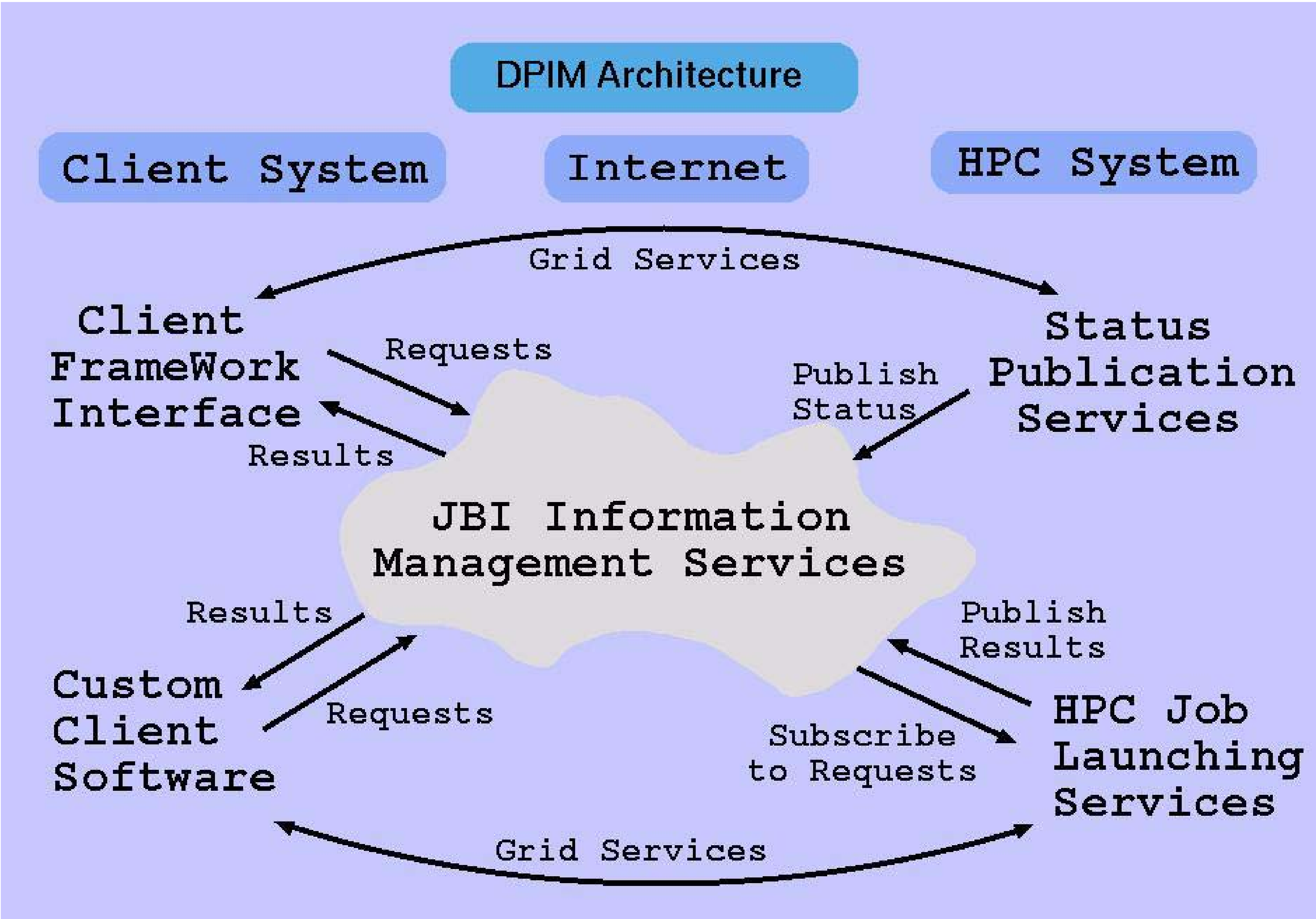
- Current JBI:
  - **Provide low latency and high capacity of the core services of a distributed, fault tolerant information management system**
- Goal:
  - **Use distributed interactive HPC resources on DREN**
  - **Fault tolerance**
  - **Scalability**

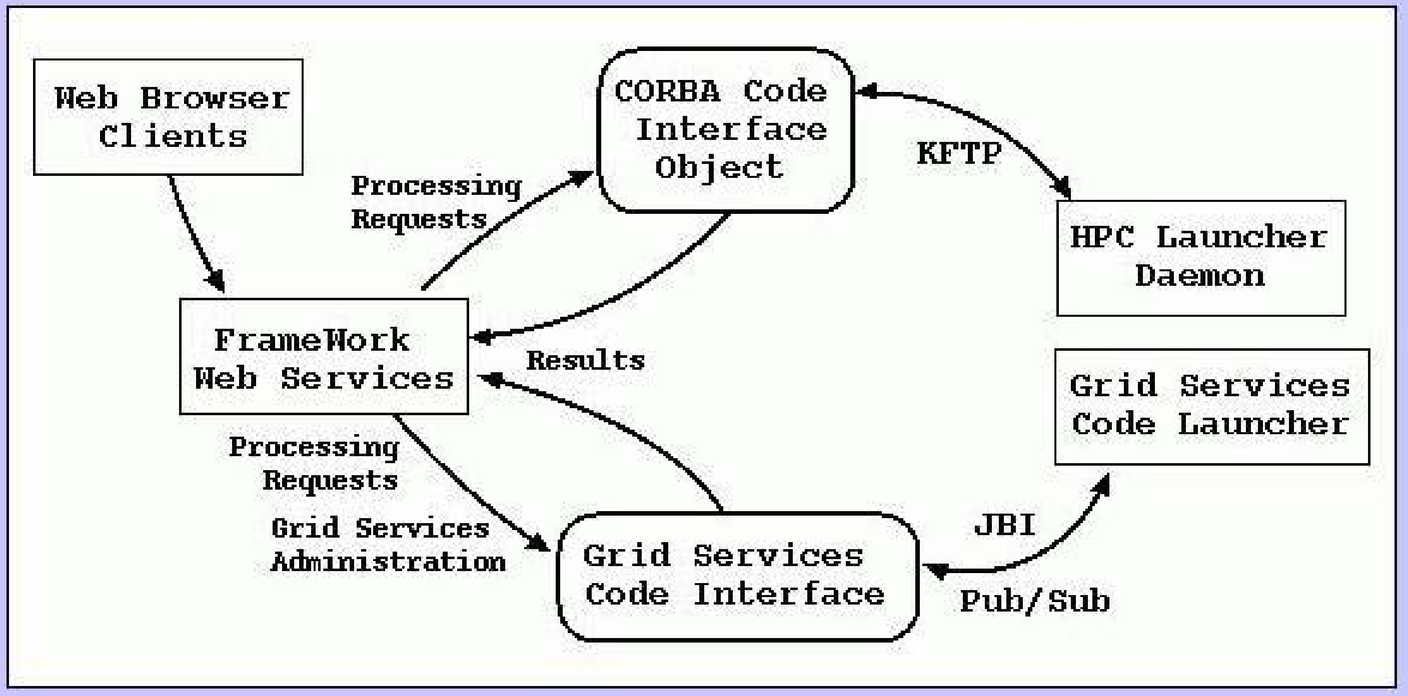
Scalability



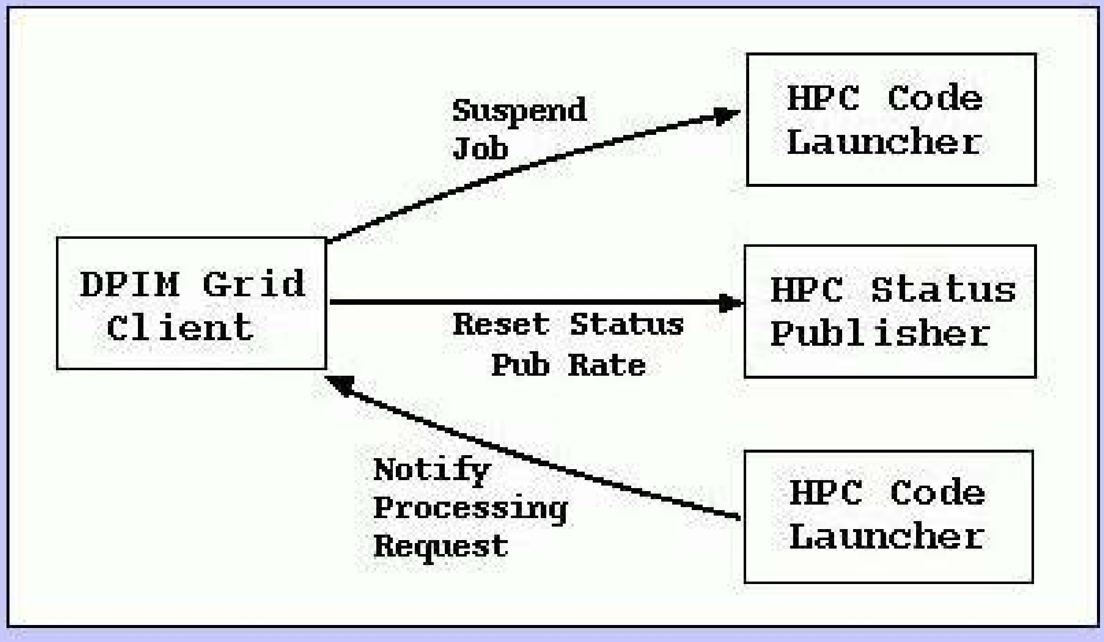
- 300 subscribers
- Predicate ensures that each publication is delivered to exactly one subscriber

- 300 Predicates
- 600 Clauses
- 0.3% hit ratio (1/300)
- 4-6 Publishers were used
- 2 KB IOs

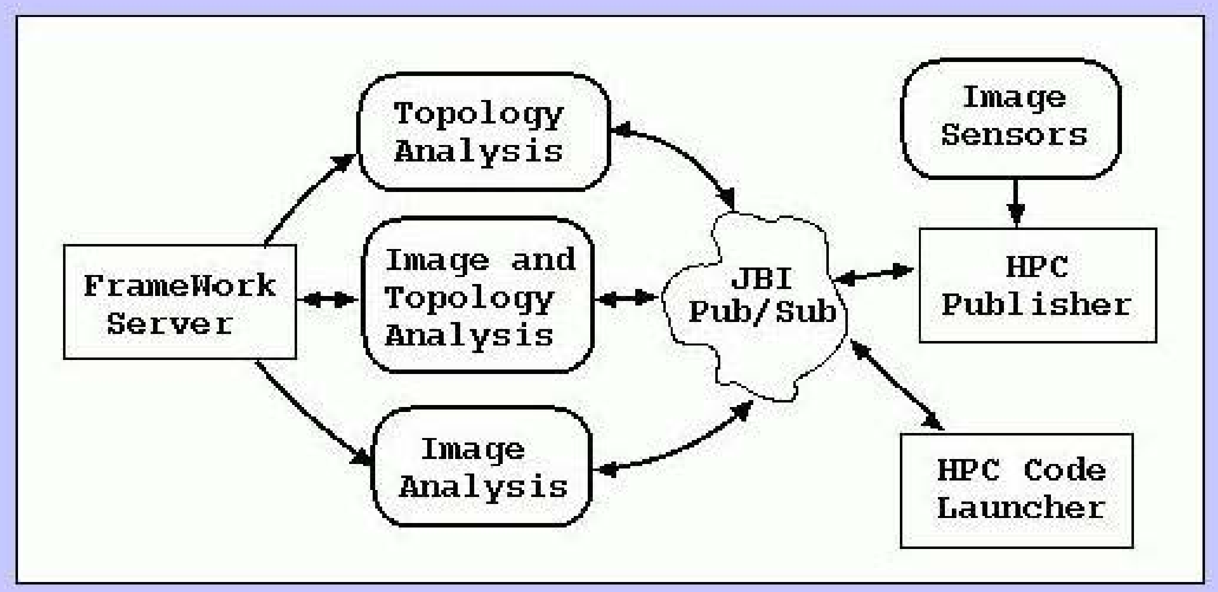

DPIM Architecture
Client System
Internet
HPC System
Grid Services
Client FrameWork Interface
Requests
Results
Publish Status
Status Publication Services
JBI Information Management Services
Results
Requests
Custom Client Software
Publish Results
Subscribe to Requests
HPC Job Launching Services
Grid Services



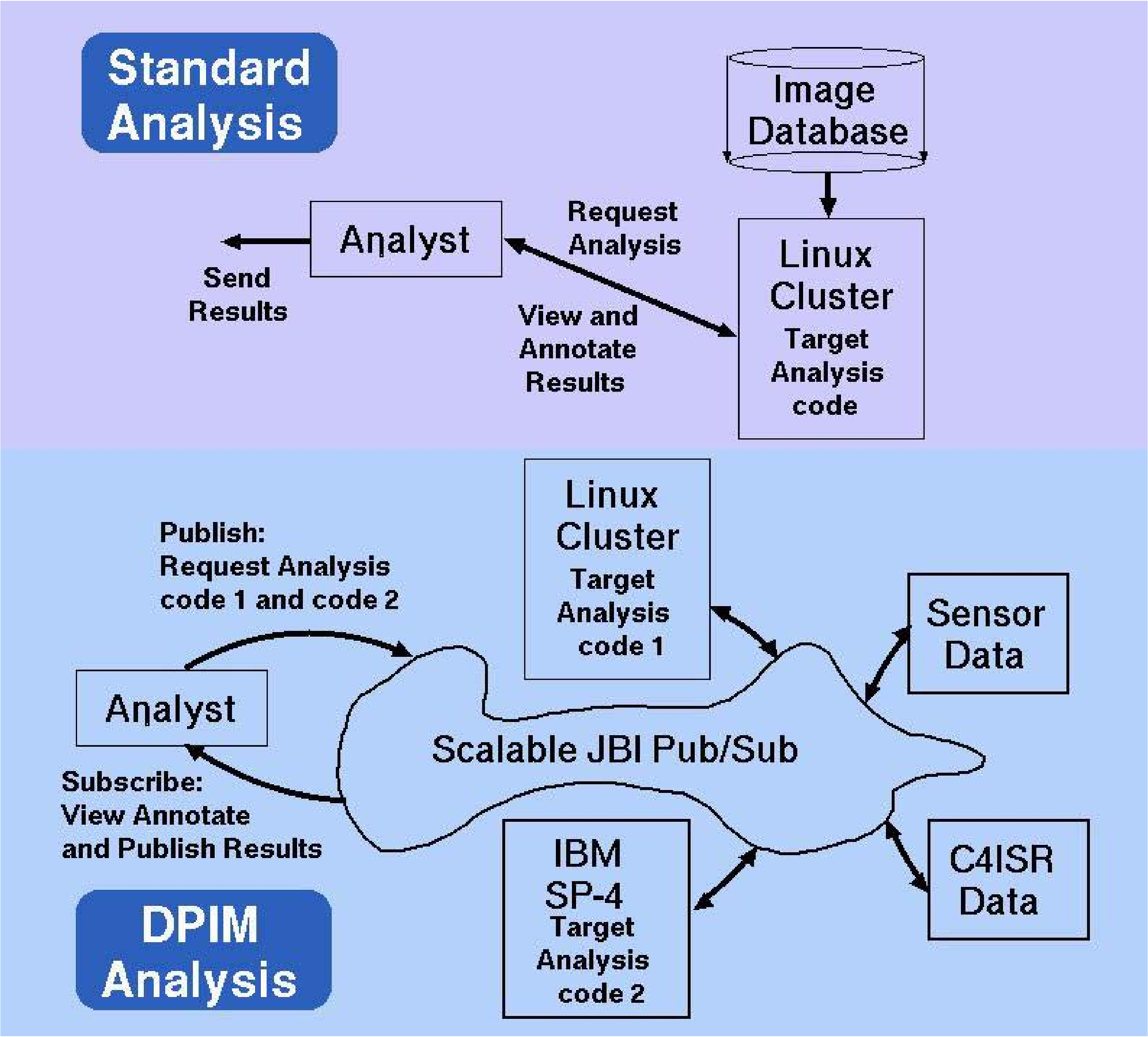
Replacing CORBA by Grid Services for Improved Control over Execution



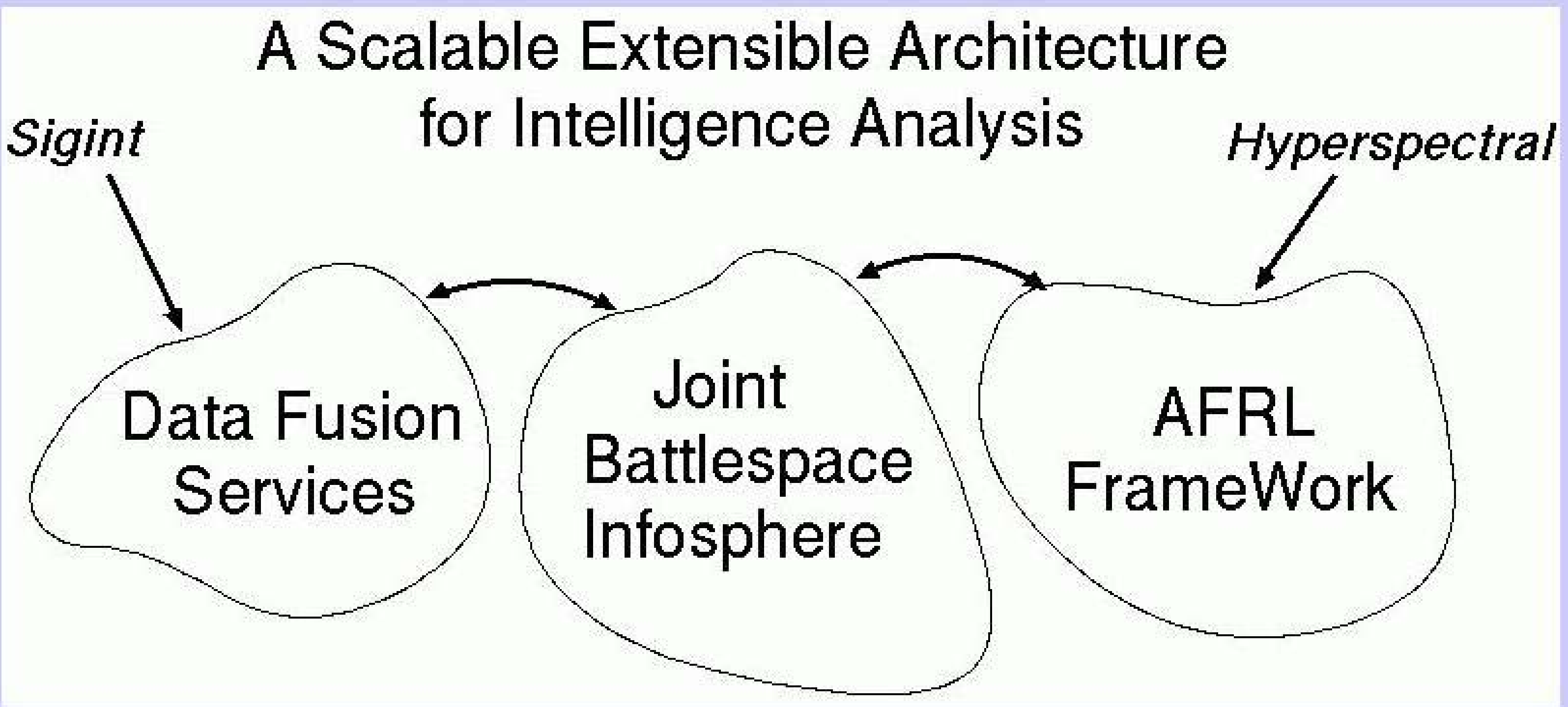
DPIM Grid Interface Used to Control Remote Operations



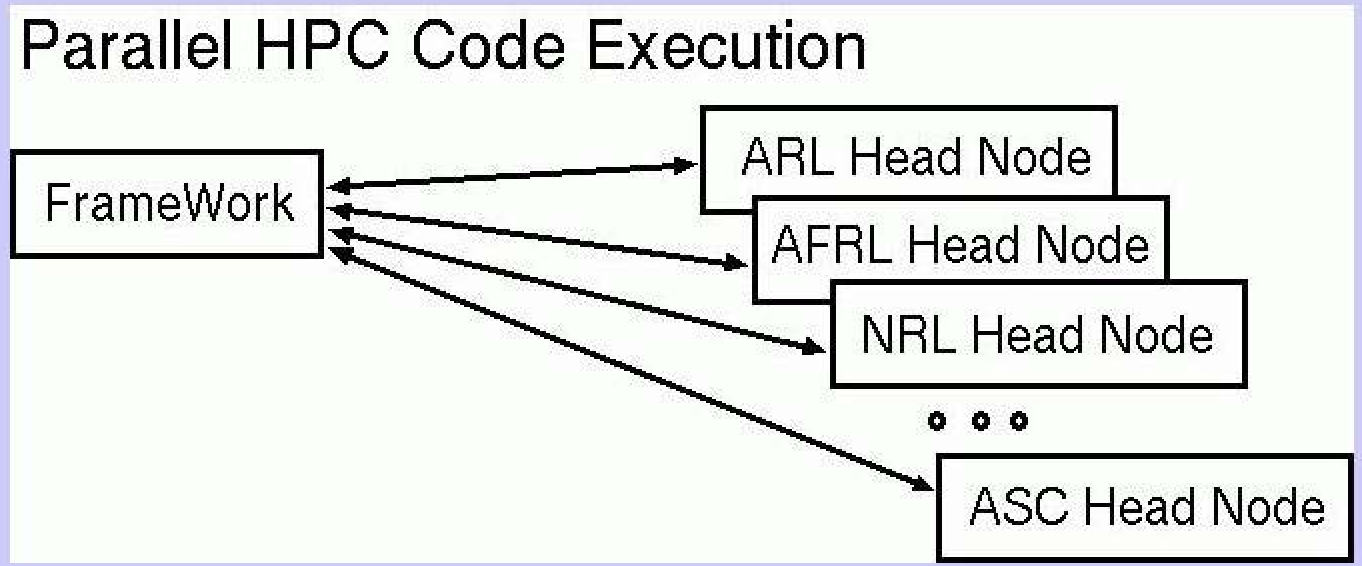
Communities of Interest Share Codes and HPC Resources to Improve Results



Standard Analysis
Image Database
Analyst
Send Results
Request Analysis
View and Annotate Results
Linux Cluster
Target Analysis code
Linux Cluster
Target Analysis code 1
Publish:
Request Analysis code 1 and code 2
Sensor Data
Analyst
Scalable JBI Pub/Sub
Subscribe:
View Annotate and Publish Results
IBM SP-4
Target Analysis code 2
C4ISR Data
DPIM Analysis



Additional Data Sources and Processing Codes can be Dynamically Configured



DPIM uses Status Publications and Resource Location Publications to Develop Execution Plan

DPIM Machines File

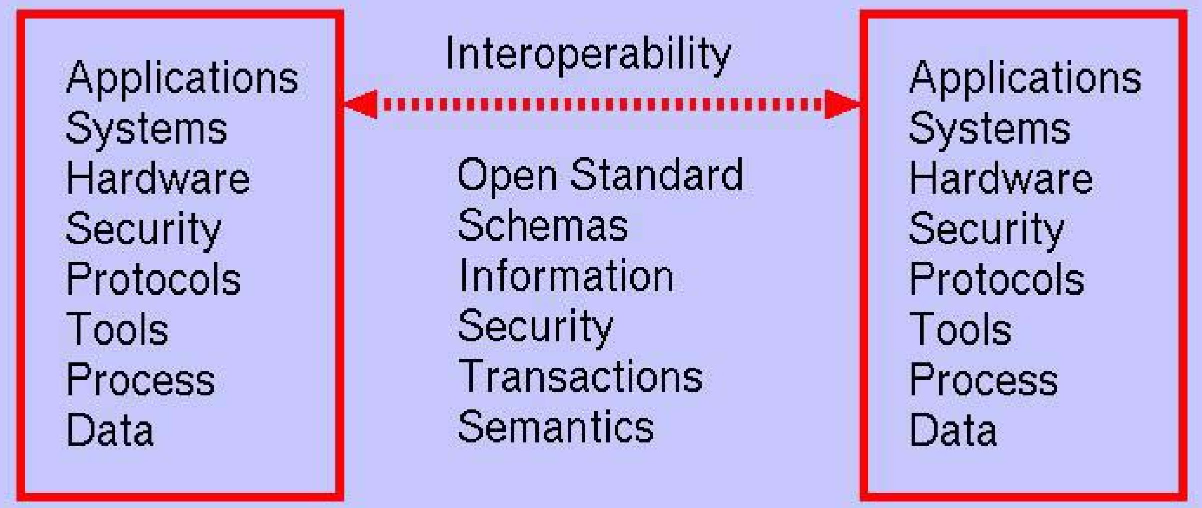
DPIM Machines File

```
http://grid.hpc.rl.af.mil/afrl/diht/coy0/launcher:20
http://grid.hpc.rl.af.mil/afrl/diht/hhpc0/launcher:24
http://grid.hpc.rl.af.mil/afrl/diht/mach2/launcher:32
http://grid.hpc.rl.af.mil/afrl/diht/seafarer/launcher:48
http://grid.hpc.rl.af.mil/afrl/diht/powell/launcher:48
```

DPIM Supports a Standard Grid Services Interface for the MPICH Globus Device

From: http://www.ncoic.org/download/NCOIC_Position_Paper_V21.pdf

Position Paper March 2005
An Introduction to the Network Centric Operations Industry Consortium (NCOIC")



"By using the open standards identified for use in the NCOIC architectural framework's lower (e.g., communications and information services) tiers, developers will be able to devote more time to the application development that provides greater value and return to the customer".

Net-Centric Issues

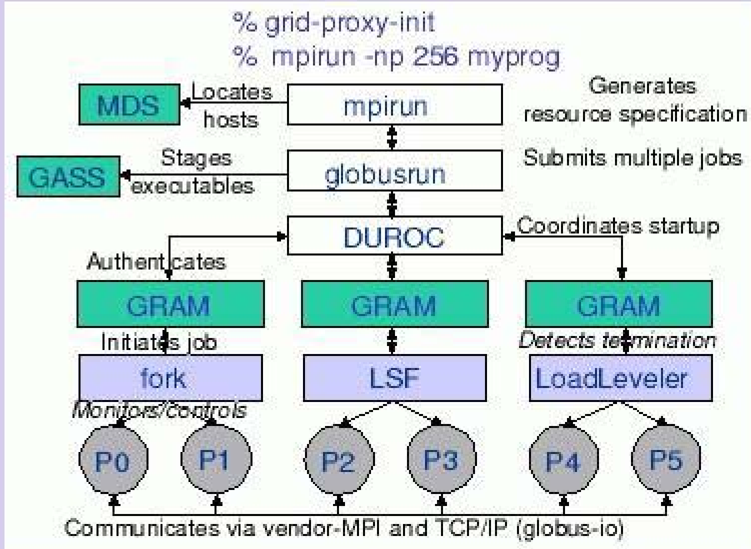
NCOIC NIF - Remote Execution Interface

MPICH-G2

NIF Remote Execution	NIF Remote Execution
Globus-IO Tcp-IP Vender-MPI	Pub/Sub

- Polling for inter-machine communication
- High-performance pub/sub

- Levels within communicator
- Transparent pub/sub

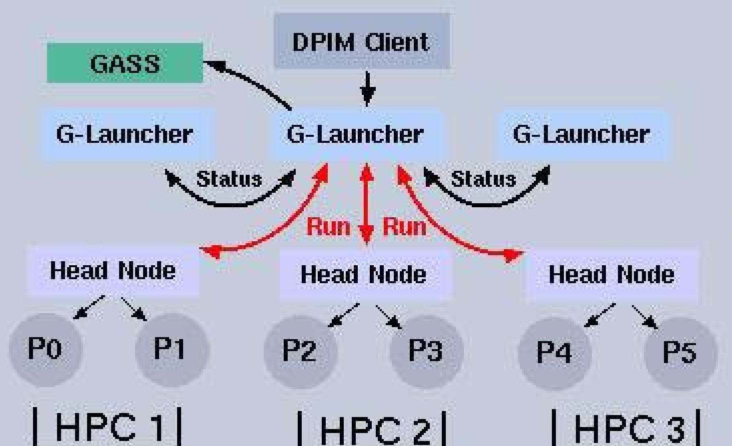


Above figure from: MPICH-G2: A Grid-Enabled Implementation of the Message Passing Interface.N. Karonis, B. Toonen, and I. Foster. Journal of Parallel and Distributed Computing, 2003.

GRID

- **General Purpose Access to HPC Systems**

  Low-level Batch Access



DPIM

- **Interactive Access to HPC Systems**

  Use of Grid Functions

Security Issues

Traditional HPC Access

Kerberos HPC Security

Single Login

Admin Logs

Other Work

Condor

Globus

Seti/Dnet

Community Framework

Nexus

GIG

JBI

Conclusion

- Increasing Demand for Interactive Access to HPC Applications
- Make use of Open Source Standards
- High-Performance Information Management is key for Command and Control