

Distributed Resource Management Application API (DRMAA) Working Group

DRMAA Proposed Recommendation Document 1.0

Cochairs

John Tollefsrud [j.t@sun.com], Sun Microsystems
Hrabri Rajic [hrabri.rajic@intel.com], Intel

GGF8

Seattle, WA USA

June, 2003

“This is one boring spec”

- Charlie Catlett, Opening Plenary, GGF8



Agenda

Introduction to DRMAA

John Tollefsrud

The DRMAA 1.0 specification

Hrabri Rajic

DRMAA adoption

Altair Engineering/PBS

Bill Nitzberg

GridIron Software

Steve Forde

enFusion

<Hrabri Rajic>

Intel

Hrabri Rajic

Sun Microsystems

John Tollefsrud

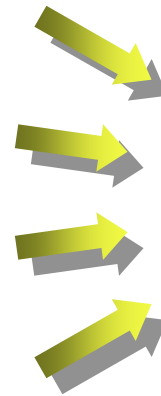
United Devices

Jikku Venkat

Q&A

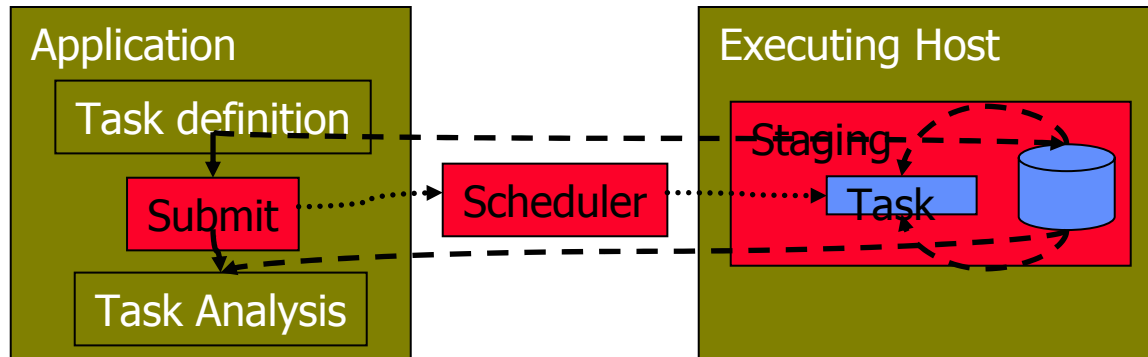
Distributed Resource Management (DRM) Systems

- Batch/job management systems
- Local Job schedulers
- Queuing systems
- Workload management systems



All are DRM Systems

Resource Management Systems Differ



- **Core services are highly similar**
 - especially from the users perspective
- **DRM programming interfaces differ**
 - ISVs are disinclined to use

Why DRMAA?

Adoption of distributed computing solutions in industry is widespread in industries dominated by software from Independent Software Vendors (ISVs)

Sector	% Revenue in 2005
Manufacturing (includes Auto, Aero, EDA, etc.)	45%
Life Sciences	15%
Financial	10%
Extraction (Oil/Gas/Mining)	8.0%
Chemistry/Chemical	6.0%
Telecommunications	5.0%
Other (CRM, Supply Chain, gaming, etc.)	11%

Global Grid Computing Report 2002
Technology Computing Partners



Why DRMAA?

Distributed computing deployments in industry can be characterized as widespread and yet ‘early adopter’

- **Direct interfacing of ISV apps to DRM systems is not the norm**
- **End users script command-line integrations**

• Fundamental improvements in the adoption pattern requires shifting the DRM integration to the ISV



The DRMAA goal

Create a well-considered DRM component interface to facilitate distributed computing deployments

DRMAA WG Charter

- Develop an API specification for the submission and control of jobs to one or more Distributed Resource Management (DRM) systems.
- The scope of this specification is all the high level functionality which is necessary for an application to consign a job to a DRM system including common operations on jobs like termination or suspension.
- The objective is to facilitate the direct interfacing of applications to today's DRM systems by application's builders, portal builders, and Independent Software Vendors (ISVs).

DRMAA Working Group

DRMAA discussed at DRMAA BOF, GGF3, Oct 2001

- Presentation and discussion on two first-draft proposals
 - Hrabri Rajic, Intel
 - Andreas Haas, Sun Microsystems
- WG status subsequently granted by GGF Steering Committee

Much committee work to integrate the proposals!

- Multiple working sessions at GGF4 – GGF7
- Weekly and Biweekly con calls
- Editors from Cadence, IBM, Intel, Robarts Res. Inst, Sun, Veridian Systems/PBS
- Other contributors from Argonne, HP, NASA, Platform Computing, United Devices, USC, others

DRMAA v.1 document public comment period completed

The current WG focus is a DRMAA C binding doc



Agenda

Introduction to DRMAA

John Tollefsrud



The DRMAA 1.0 specification

Hrabri Rajic

DRMAA adoption

Altair Engineering/PBS

GridIron Software

enFusion

Intel

Sun Microsystems

United Devices

Bill Nitzberg

Steve Forde

<Hrabri Rajic>

Hrabri Rajic

John Tollefsrud

Jikku Venkat

Q&A

Scope: Run a Job API

(Steps from: Ten Actions when SuperScheduling", GGF SchedWD 8.5, J.M. Schopf, July 2001)

- Phase 1: Resource Discovery
 - Step 1 Authorization Filtering
 - Step 2 Application requirement definition
 - Step 3 Minimal requirement filtering
- Phase 2 System Selection
 - Step 4 Gathering information (query)
 - Step 5 Select the system(s) to run on
- Phase 3 Run job
 - Step 6 (optional) Make an advance reservation
 - **Step 7 Submit job to resources**
 - Step 8 Preparation Tasks
 - **Step 9 Monitor progress (maybe go back to 4)**
 - **Step 10 Find out Job is done**
 - Step 11 Completion tasks

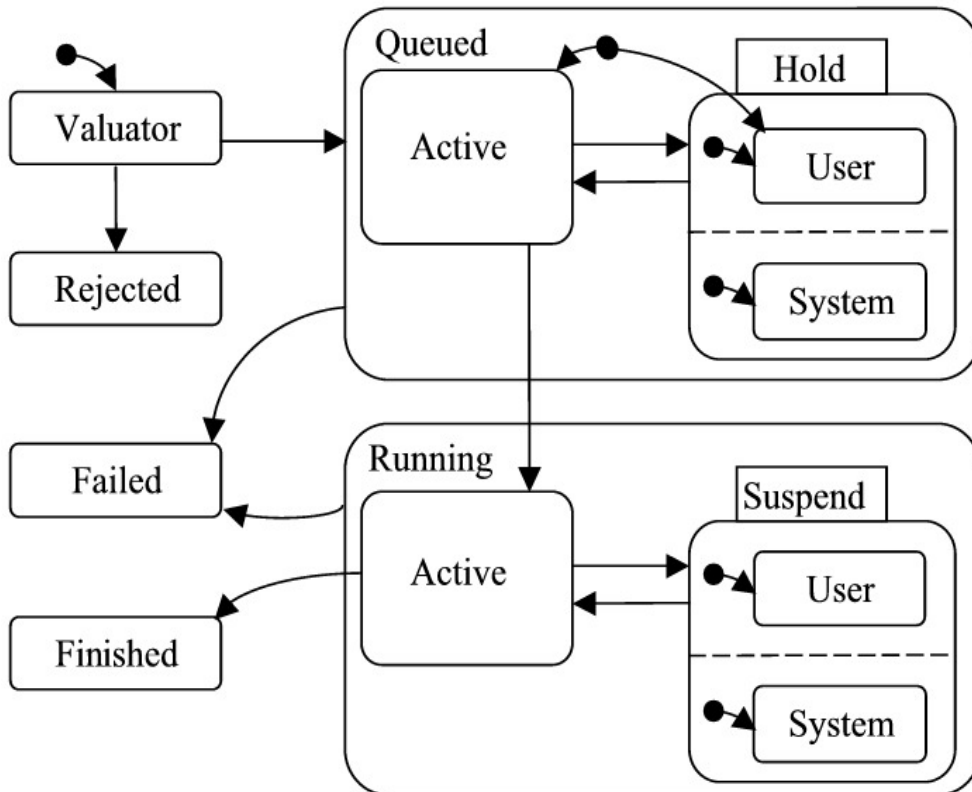


C Implementation requirements

- **C-API library interface - no protocol**
 - Simplifies utilization by ISV
 - Not transactional
 - Object oriented wrappers/classes specification in the works
- **Shared library binding**
 - Prerequisite to allow end user to select DRM technology of their choice
- **One session at the time**
- **Library supports only one DRM system per implementation**
 - Simultaneous support of different DRM systems is beyond the scope of our project



DRMAA State Diagram



The remote job could be in following states:

- system hold
- user hold
- system and user hold simultaneously
- queued active
- system suspended
- user suspended
- system and user suspended simultaneously
- running
- finished (un)successfully

DRMAA Job Attributes

Mandatory job attributes:

- Remote command to execute
- Remote command input parameters, a vector parameter
- Job state at submission
- Job environment, a vector parameter
- Job working directory
- Job category
- Native specification
- Standard input, output, and error streams
- E-mail distribution list to report the job completion and status, a vector parameter
- E-mail suppression
- Job start time
- Job name to be used for the job submission

Optional job attributes:

- transfer files
- absolute job termination time
- wall clock time limit
- soft wall clock time limit
- job run duration hlimit
- job run duration slimit



API groups

- **Init/exit**
- **Job template interfaces**
 - Allocate/delete
 - Setter/getter job template routines
- **Job submit**
 - Individual jobs
 - One time
 - Multiple times – templates (version 2)
 - Bulk jobs, implicit parameterization
- **Job monitoring and control**
- **Auxiliary or system routines**
 - Error message routine
 - Informational interfaces



Site specific requirements

- **Application Developers and DRMS vendors are not involved in the local environment specification**
 - Execution policies
 - Physical environment
- **Two hierarchical mechanisms**
 - Job categories
 - Vendor determines the name and application parameter guidelines
 - Administrators, installation people
 - Native specification
 - Opaque string that DRMAA impl. resolves
 - The burden is on the end users to define the execution environment
 - Need to know DRM
 - Need to know the remote application installation



DRMAA Adoption

Introduction to DRMAA

The DRMAA 1.0 document



DRMAA adoption

Altair Engineering/PBS

GridIron Software

enFusion

Intel

Sun Microsystems

United Devices

Q&A

John Tollefsrud

Hrabri Rajic

Bill Nitzberg

Steve Forde

<Hrabri Rajic>

Hrabri Rajic

John Tollefsrud

Jikku Venkat



Altair Engineering

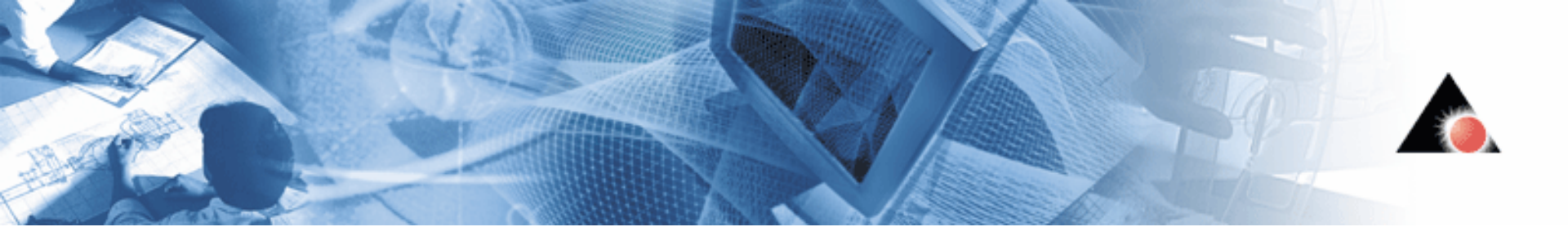
The Shortest Distance Between Concept and Reality™



PBS Pro & DRMAA

Bill Nitzberg, CTO, Altair Grid Technologies

June 2003



Who is Altair?

- Product development, consulting and software technology company
 - Founded in 1985
 - 2,000+ Clients
 - Offices in 10 Countries
 - Worldwide Customer Base



Tokyo, Japan

Austin, USA

London, UK

München, Germany

Bangalore, India

Shanghai, China

Los Angeles, USA

Paris, France

Torino, Italy

Lund, Sweden

Toronto, Canada

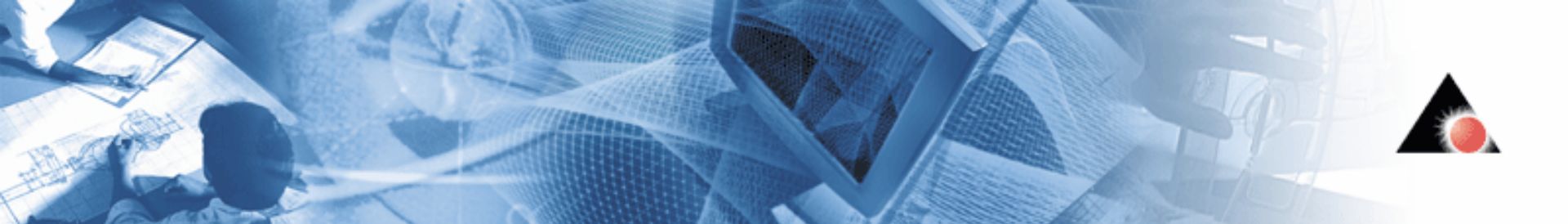
Detroit, USA

Stuttgart, Germany

Boston, USA

Lecce, Italy

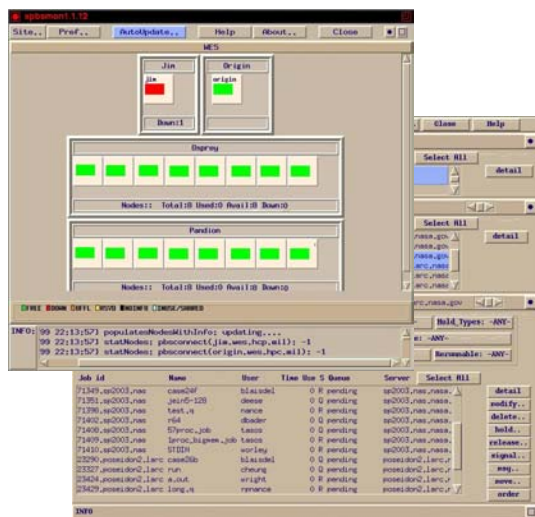
Milwaukee, USA



PBS Pro - The Portable Batch System

Flexible workload management and job scheduler

- Desktops → Clusters → Datacenters → Multi-site Grids
 - UNIX, Windows, Linux, ... && Serial, SMP, MPI, ...
- Peer Scheduling, Cycle Harvesting
- Advance reservations, File staging, X.509 certs
- Front & back-end to The Globus Toolkit™
- Fully configurable scheduler module
- Detailed accounting
- GUIs, High Availability, ACLs
- POSIX 1003.2d compliant



DRMAA

- Supports early “gridification” of commercial applications
- Altair is one of the DRMAA v1.0 authors
- We are committed to fostering & supporting GGF standards

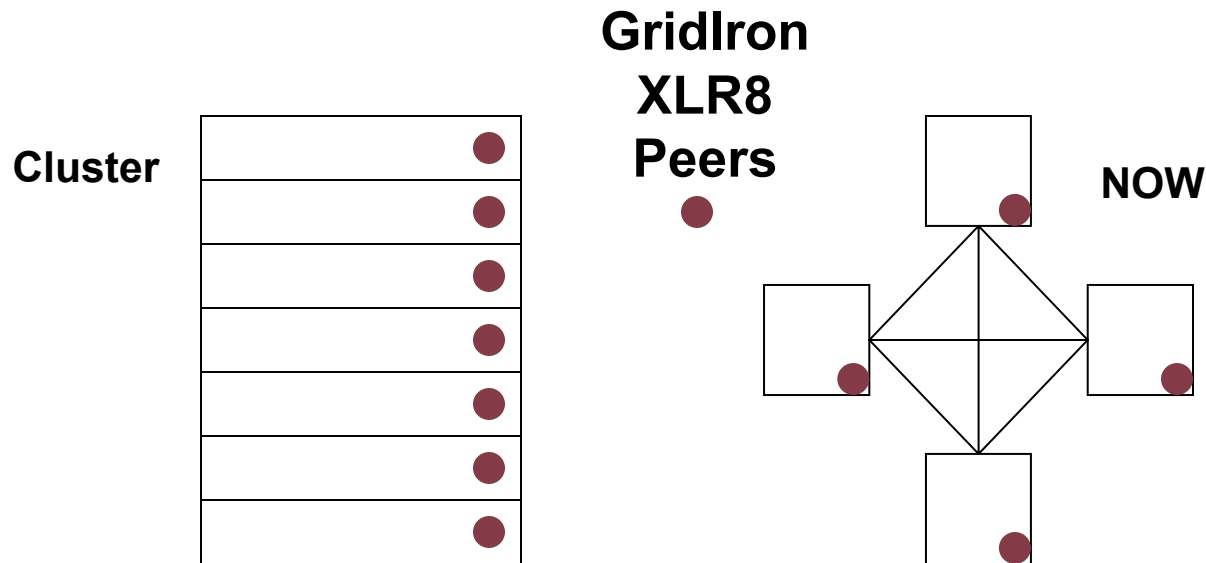


www.pbspro.com

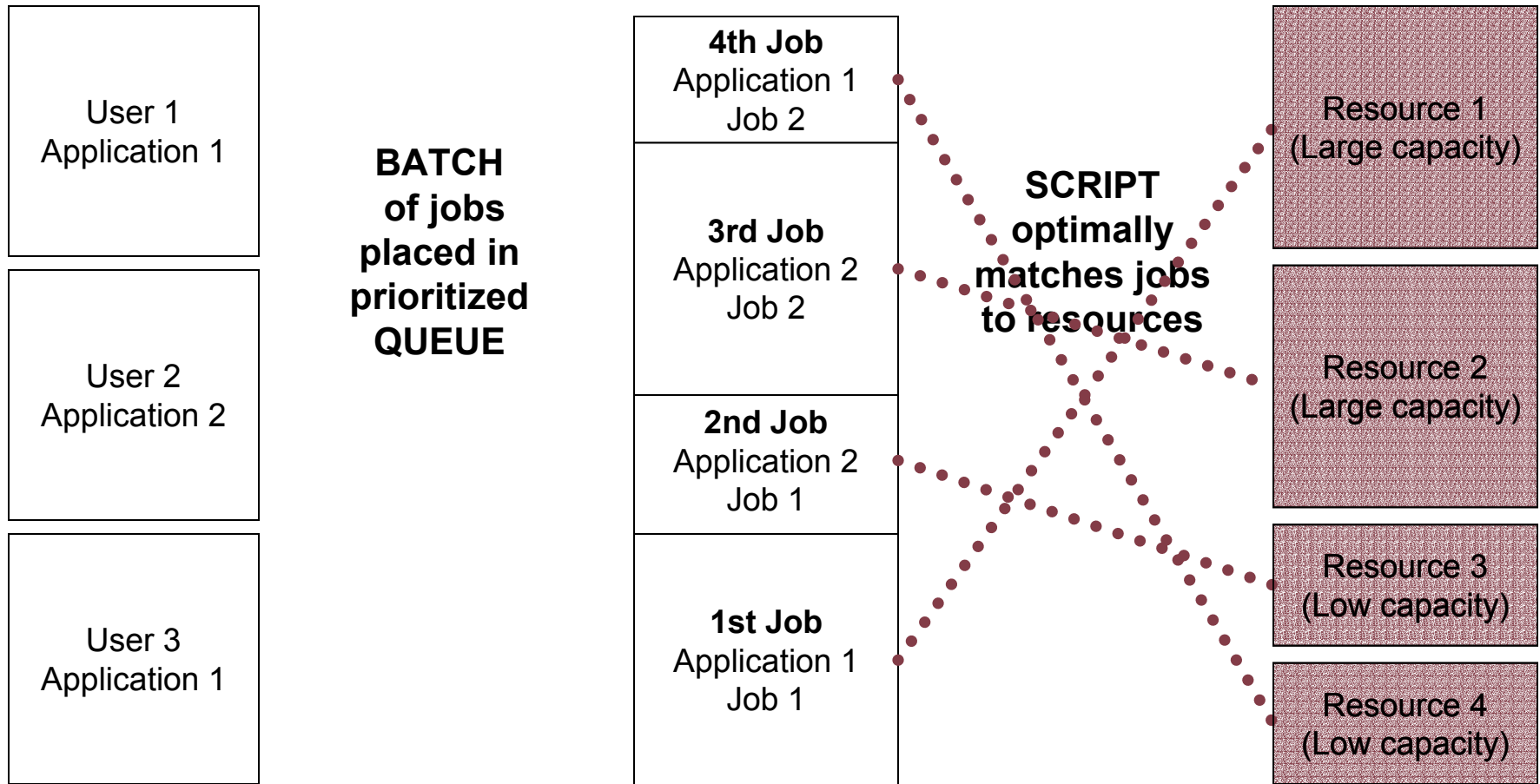
GridIron XLR8 - Context

Digital Content Creation Applications

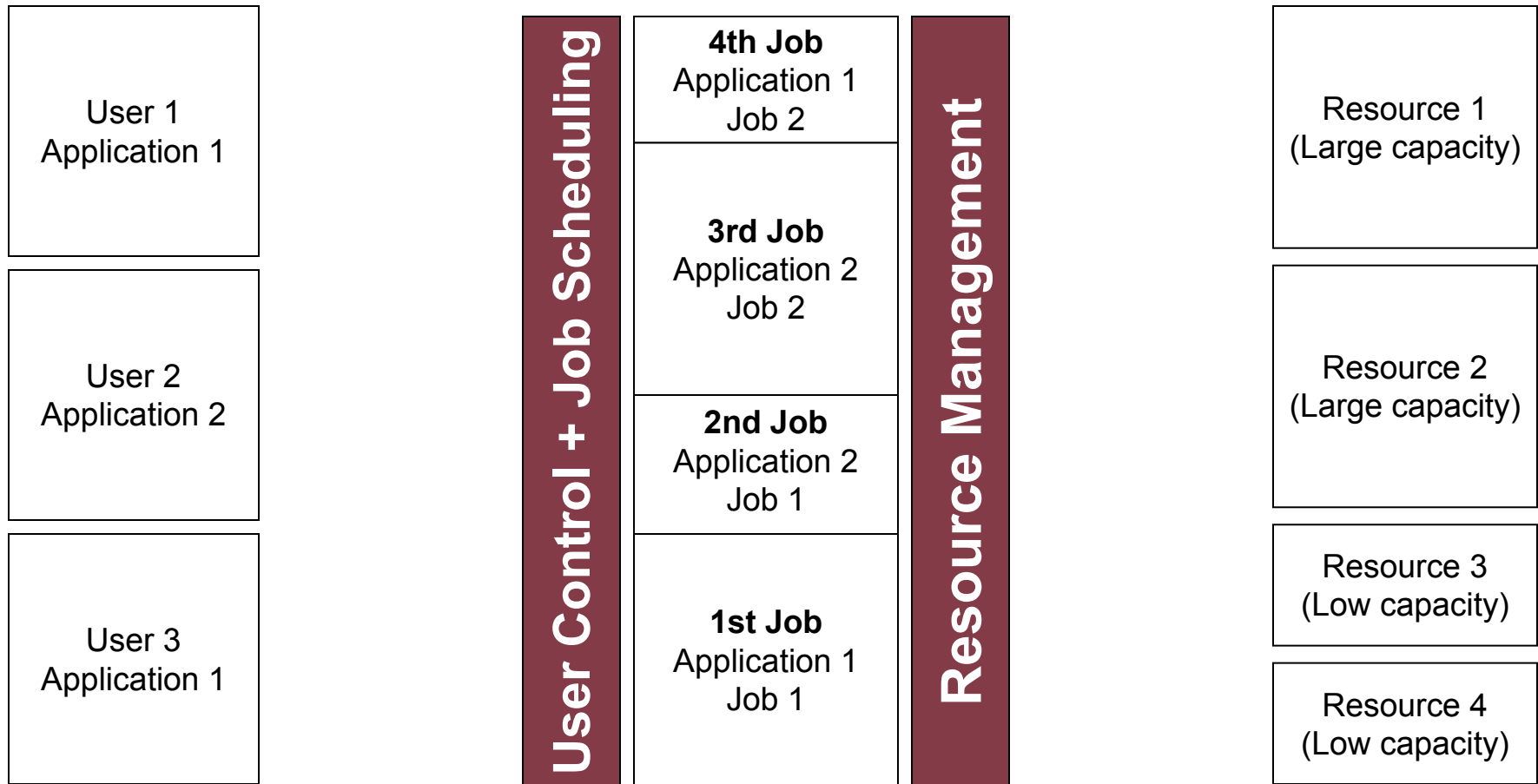
GridIron XLR8 Parallel API's



Scripted Distribution



Scripted Distribution



DRM MIDDLEWARE
(e.g. Sun Grid Engine)



Scripted Distribution

PROs

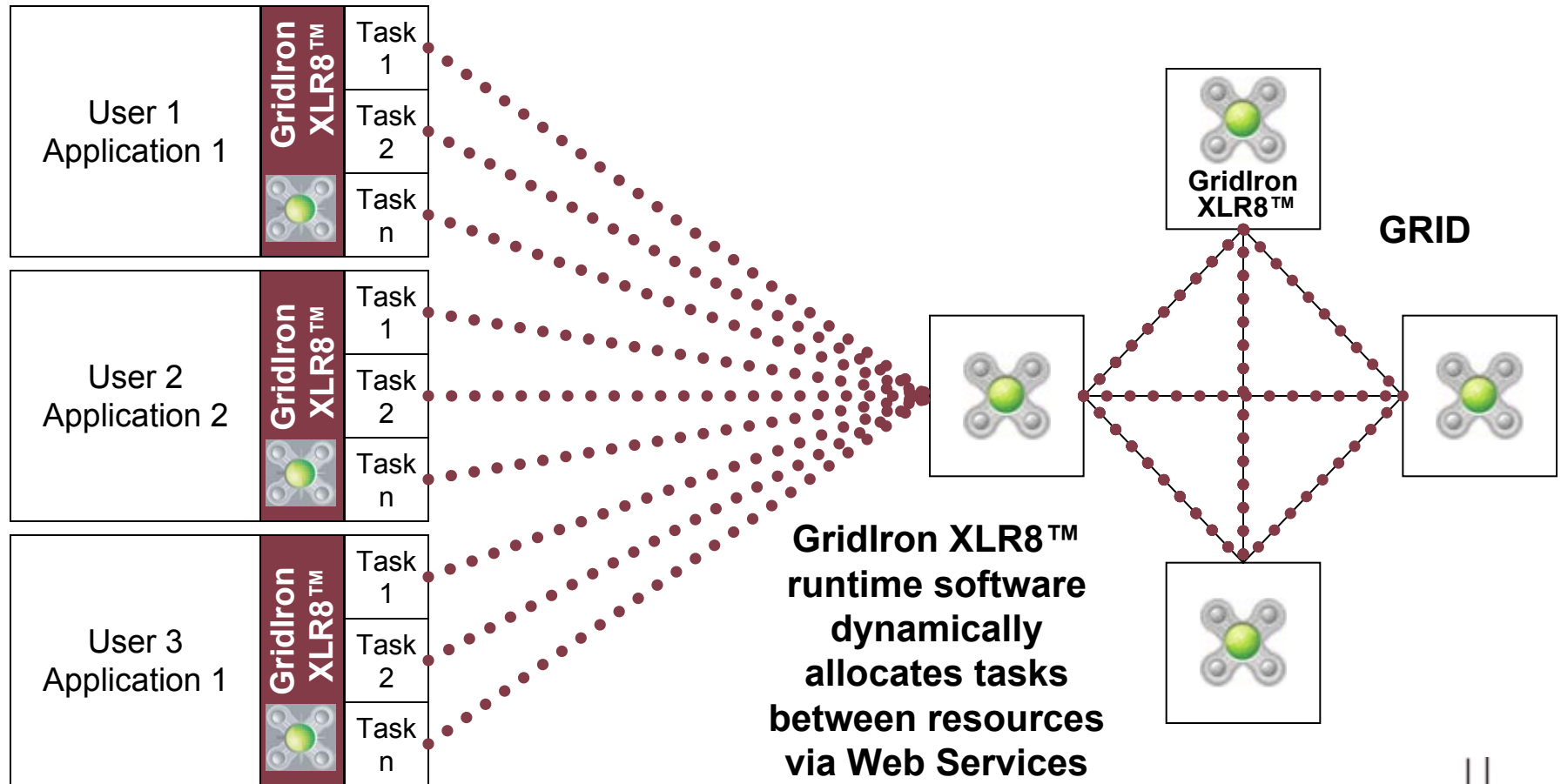
- No code modification required
- User control
- Resource management

CONs

- Only scriptable applications
- Large jobs can only run as fast as the fastest resource
- Only efficient with a large queue (many jobs)
- Requires installation of application on each resource



Programmatic Distribution



Programmatic Distribution

PROs

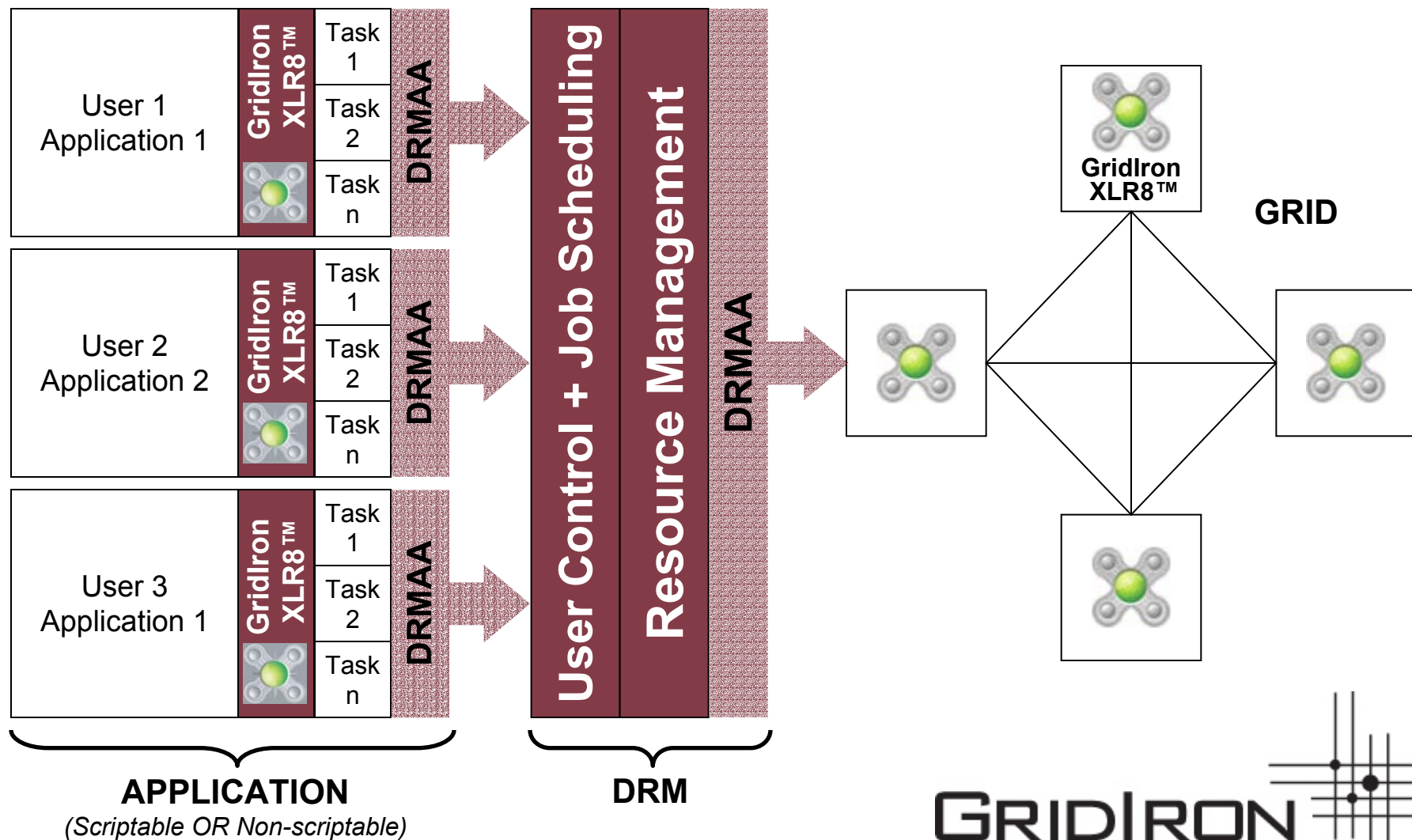
- Scriptable and non-scriptable applications
- User autonomy
- Zero configuration, zero administration runtime environment
- Effective with few jobs and/or few/small resources
- Thin runtime client

CONs

- Limited user control and resource management
- Requires code modification



DRMAA



The Best of Both Worlds

	Batch Queue	Parallel Programming	DRMAA
Scriptable applications	✓	✓	✓
Non-scriptable applications	✗	✓	✓
User administration and resource control	✓	✗	✓
User independence and autonomic resource management	✗	✓	✓
User to resource optimization (many jobs to many resources)	✓	✗	✓
Big jobs on small resources	✗	✓	✓
Few jobs on many resources	✗	✓	✓
Thin runtime client & application not required on resources	✗	✓	✓



Summary

- DRMAA allows **synergistic integration** of batch queue scriptable and programmatic solutions, e.g.
 - **Sun ONE Grid Engine** and **GridIron XLR8™**
- Delivers the **best of both worlds**:
 - Scriptable and non-scriptable applications
 - User control and resource management
 - User autonomy
 - Zero configuration autonomic runtime environment
 - Many-to-many job to resource optimization
 - Few/large jobs to few/small resources
 - Thin client runtime software



Thank-you!

Steve Forde

CTO/President

GridIron Software Inc.

sforde@gridironsoftware.com





axceleon™

EnFuzion

From **axceleon**™

The high performance
distributed computing
solution optimized for
parametric studies

GGF8 DRMAA DISCUSSION...

- Axceleon's EnFuzion is the high performance distributed computing solution optimized for parametric studies
- Axceleon views DRMAA as a very important enabling technology
- DRMAA will enable developers to utilize the power of computational grids in a wide range of environments.
- Axceleon supports the DRMAA standardization activity and is planning to implement DRMAA in its EnFuzion product.

DRMAA at Intel

Hrabri Rajic

- Intel is one of the founders of the GGF DRMAA Working Group.
- One DRMAA co-chair is from Intel.
- Intel actively contributes to DRMAA documents.
- Intel talks to ISVs about DRMAA.
- Intel is backing and participating in DRMAA activities.

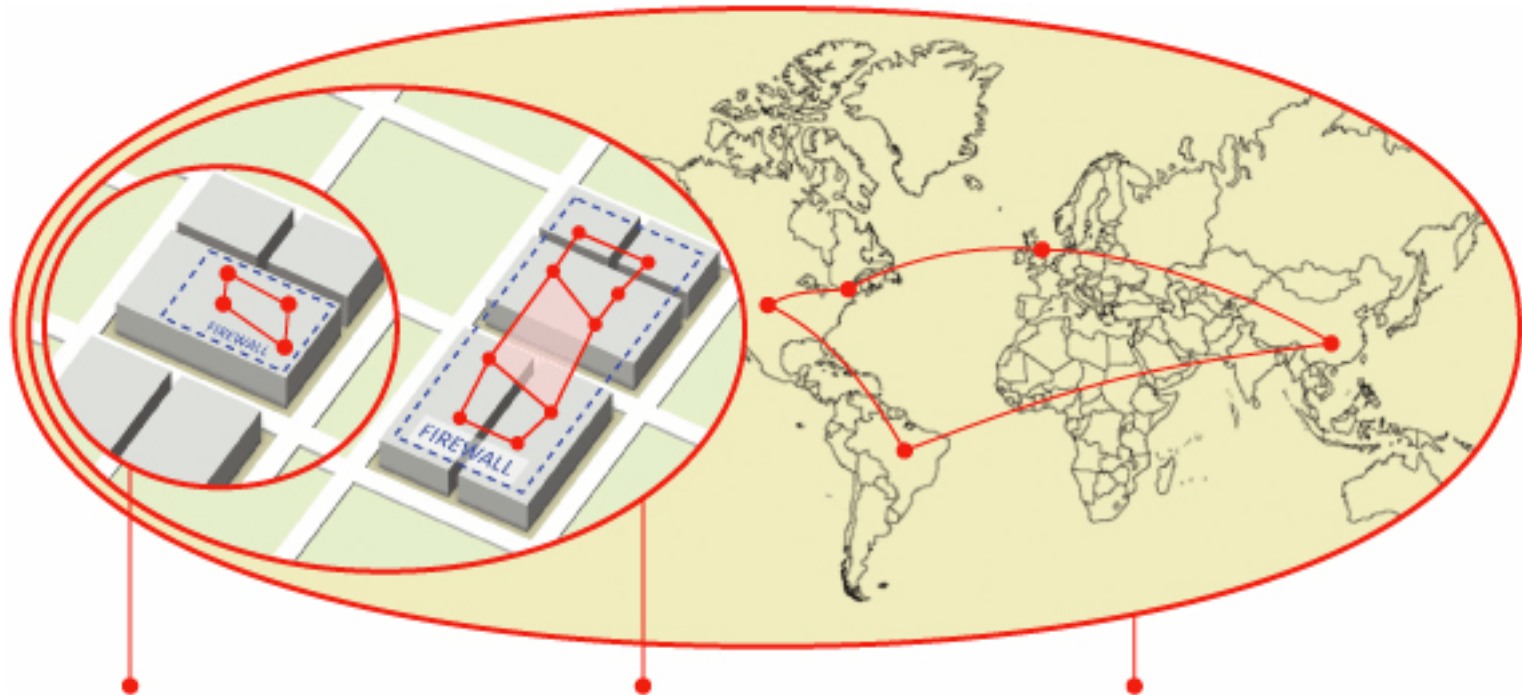


DRMAA adoption at Sun Microsystems

John Tollefsrud
j.t@sun.com



Grid Computing at Sun



Cluster Grids Departmental Computing

- Simplest Grid deployment
- Maximum utilization of departmental resources
- Resources allocated based on priorities

Enterprise Grids Enterprise Computing

- Enterprise resource sharing
- Policy driven
- Gives multiple groups seamless access to enterprise resources

Global Grids Internet Computing

- Resource sharing over the Internet
- Global view of distributed datasets
- Built on Cluster Grids and Enterprise Grids

Grid Engine software

Sun ONE Grid Engine

- Queuing, prioritizing, scheduling
- Solaris Sparc, Solaris X86, Linux
- Free download from the Sun Download Center

Sun ONE Grid Engine, Enterprise Edition

- Equitable, enforceable sharing between groups and projects
- Alignment of resources with business goals via policies
- Solaris Sparc, Solaris X86, Linux

Grid Engine Open Source Project

www.gridengine.sunsource.net

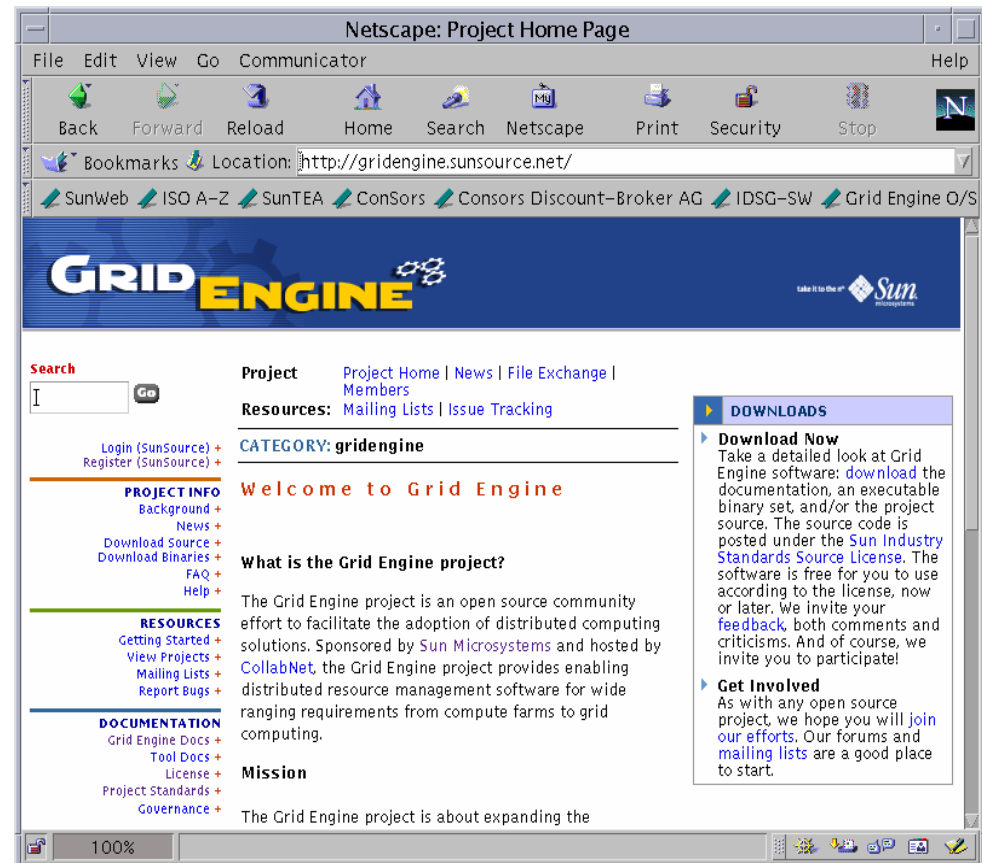
.500,000+ lines of source code

.Binaries: Solaris, Linux, AIX, HP/ux, IRIX, Tru64

.User and developer community aliases

.SISSL source license

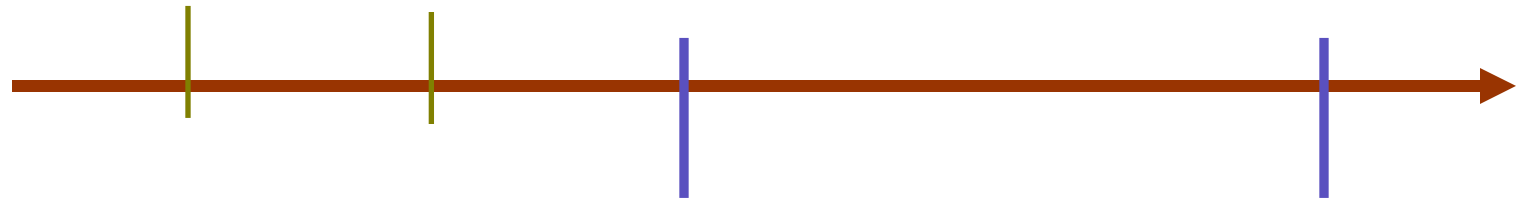
.Code basis for S1GE/EE



Planned Grid Engine DRMAA support

DRMAA 1.0 spec

C binding



Grid Engine Project

- DRMAA implementation source stable snapshot on current SGE 5.3
- DRMAA test-client source

Summer 2003

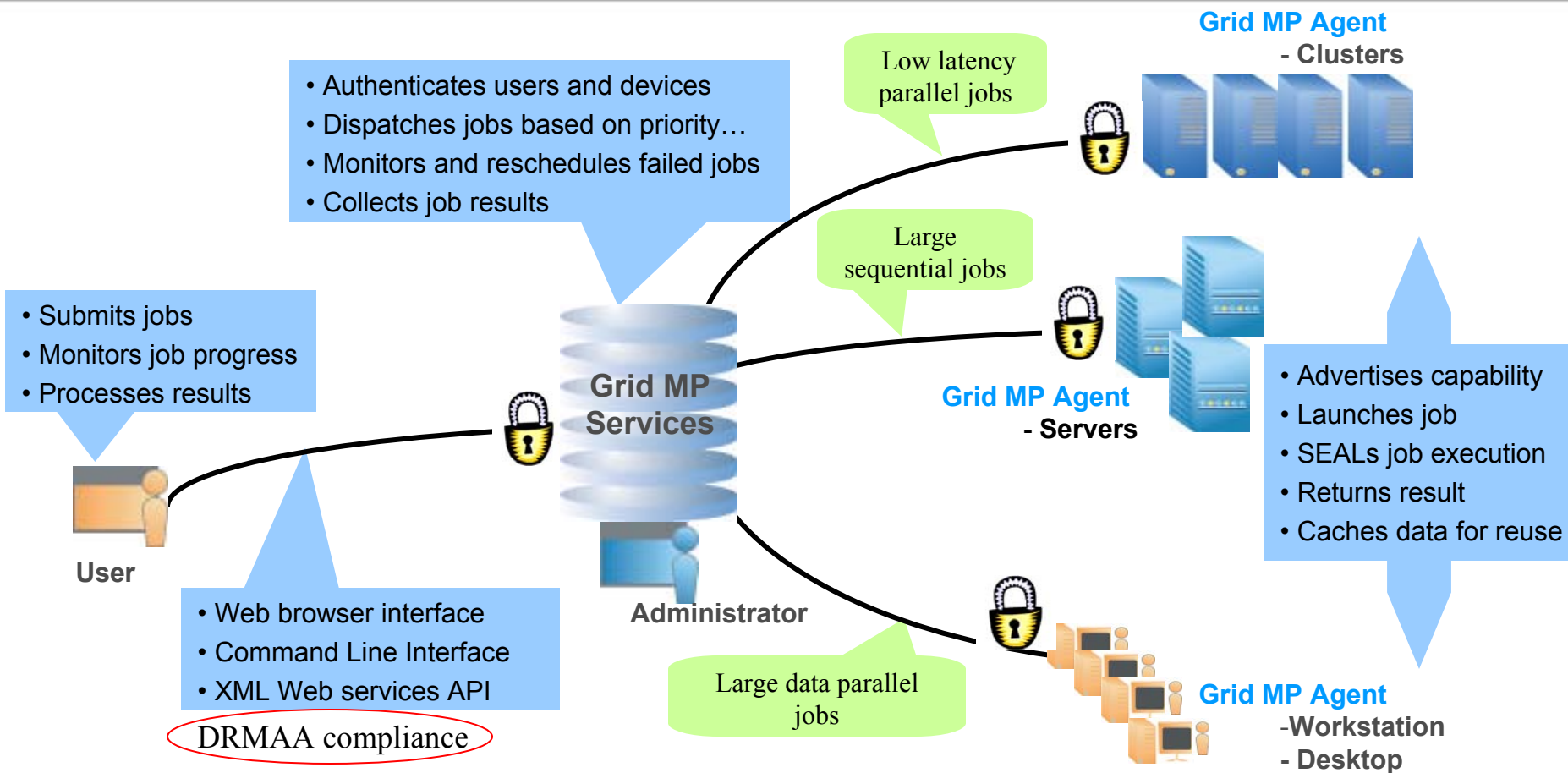
Sun ONE Grid Engine

- Support in SGE/EE 6.0

Release date tbd



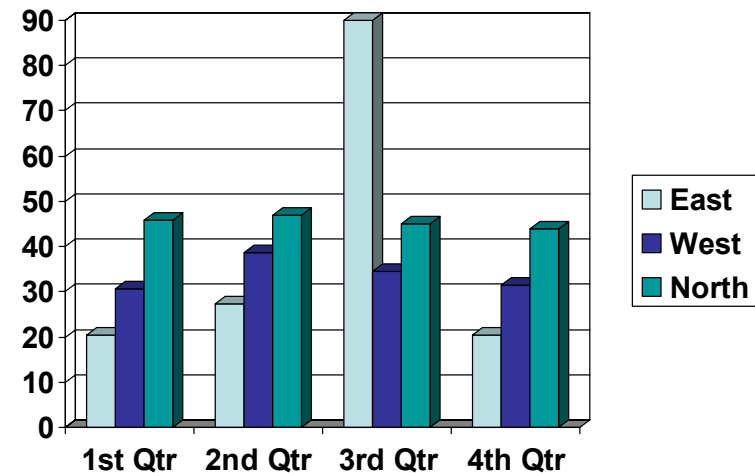
How Grid MP[®] Works





UD and DRMAA

- UD customers and partners will benefit from DRMAA
 - Protects their investment in developing application scripts and application utilities
 - Provides a common job submission layer in a heterogeneous environment – multiple DRMs
- UD supports the DRMAA specification
- Will comply with the standard in a future release of the Grid MP



DRMAA Q&A

Introduction to DRMAA

The DRMAA 1.0 document

DRMAA adoption

Altair Engineering/PBS

GridIron Software

enFusion

Intel

Sun Microsystems

United Devices

John Tollefsrud

Hrabri Rajic

Bill Nitzberg

Steve Forde

<Hrabri Rajic>

Hrabri Rajic

John Tollefsrud

Jikku Venkat



Q&A

