



Enabling Grids for  
E-science in Europe

[www.eu-egee.org](http://www.eu-egee.org)

*JRA1 All Hands Meeting, 15-17 Nov 2004*

# GLUE Schema

**Sergio Andreozzi**  
JRA1 IT/CZ Cluster  
[sergio.andreozzi@cnafr.infn.it](mailto:sergio.andreozzi@cnafr.infn.it)



**EGEE is a project funded by the European Union under contract INFSO-RI-508833**  
*Copyright (c) Members of the EGEE Collaboration. 2004.*

# Contents

- Problem Statement
- GLUE Schema
  - From the beginning to specification version 1.1
  - Current Revision Process
  - What's new in the current draft for spec version 1.2



# Problem Statement

- **Resources** available in Grid systems must be **described** in a **precise** and **systematic manner** if they are to be able to be **discovered** for subsequent management or use
- A **shared** description allows multiple experts to contribute to the problem and serves as a communication mean between different knowledge domains

## INFORMATION MODEL

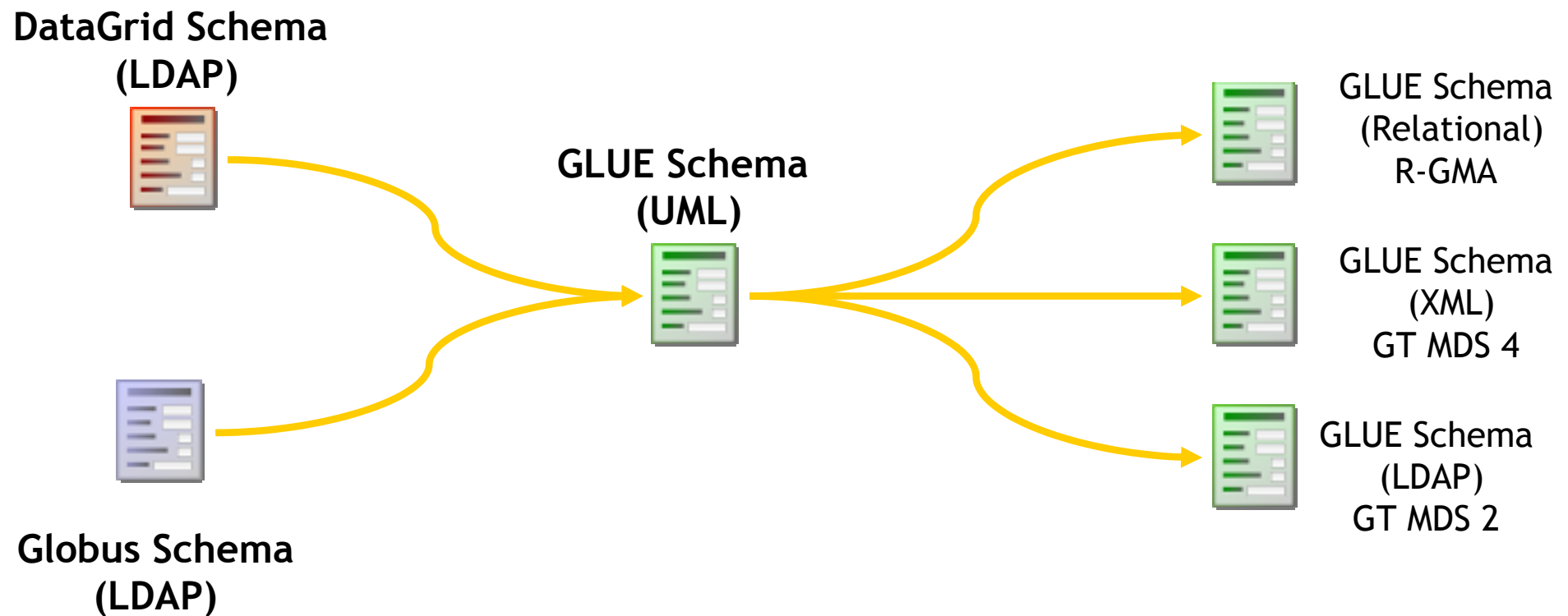
- Abstraction of real world into constructs that can be represented in computer systems (e.g., objects, properties, behavior, and relationships)
  - Not tied to any particular implementation
  - Used to exchange information among different domains

# Problem Statement

- Main Use Cases:
  - Discovery for brokering and access:
    - “what are the Computing Elements available to the VO CMS and that offer the SL3 operating system with installed the CMKIN software package?”
    - “what are the Storage Elements that offer 20 gigabytes of disk space for the VO ATLAS?”
  - Discovery for monitoring
    - “how many CPUs the site XYZ is offering to the EGEE Grid?”
    - “what is the success rate of job submitted per site?”

# GLUE Schema

- approach to the information modeling of Grid resources started in April 2002 by the DataTAG and iVDGL projects
- Contributions from DataGrid, Globus, PPDG, GryPhyn



# GLUE Schema - current spec

- 31 March 2003: GLUE Schema version 1.1
  - Computing Resources
    - Computing Element
    - Cluster
    - SubCluster
    - Host
  - Storage Resources
    - Storage Element
    - Storage Space
    - Data Access Protocol
    - Storage Library
  - Computing/Storage Relationship
    - CESEBind

[ 1 ]

# GLUE Schema Current Revision Process

# GLUE Schema - next steps

- There is a number of issues to be addressed:
  - Simplification
  - “bug fixing”
  - better documentation
  - new use cases
  - Extensions:
    - Per-VO view of computing resources (e.g., Estimated Traversal Time, Free Job Slots) (Jeff Templon)
    - Grid3 (some already in the new draft) [ 4 ]
    - GridICE (Grid monitoring)
    - INFN (Monitoring the Connectivity of a Grid) [ 5 ]
    - JRA1 Advance Reservation (being defined with T. Ferrari and E. Ronchieri)



# GLUE Schema - current revision process

## Communication:

- Mailing list: [glue-schema@hicb.org](mailto:glue-schema@hicb.org)
- Tracker: [http://inf Forge.cnaf.infn.it/tracker/?atid=118&group\\_id=9&func=browse](http://inf Forge.cnaf.infn.it/tracker/?atid=118&group_id=9&func=browse)
- Phone Calls or face-to-face meeting are not scheduled since one year

## Revision Process:

- Items for schema revision can be added in the tracker by anyone
- People can post their comments in the tracker
  - Each tracker update is forwarded to the mailing list
- When an agreement is reached, the change is inserted in the next schema revision

# GLUE Schema - current revision process

- Even though the active participation in this activity is still a problem, there is a periodical expression of interest in this work and its evolution
- In order to go ahead, each involved project should *refresh* **who are the representatives** that should actively **participate** and **make decision on behalf of the project**
- Area of interests vs. resource categories
  - we should propose names for EGEE

	Computing	Storage	Network
Brokering	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Data Management	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	?
Advance Reservation	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Monitoring	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

# GLUE Schema - next steps

- Possible approach:
  - **Minor revision** taking into account only bug fixing and extensions that maintain backwards compatibility (e.g., GLUE Schema 1.2 draft)
  - **Major revision** including refactoring for simplification (e.g., GLUE Schema 2.0 to be written)
- gLite should consider the major one

[ 3 ]

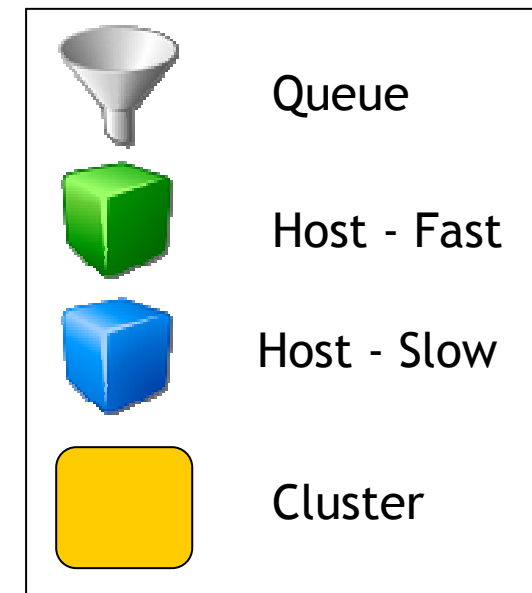


## GLUE Schema - The importance of Use Cases

- Use Cases are important
- So far, we lack of a Use Case document
- Use cases are for information models what Unit Tests are for software
- People participating in the deployment should help in collecting a number of significant use cases

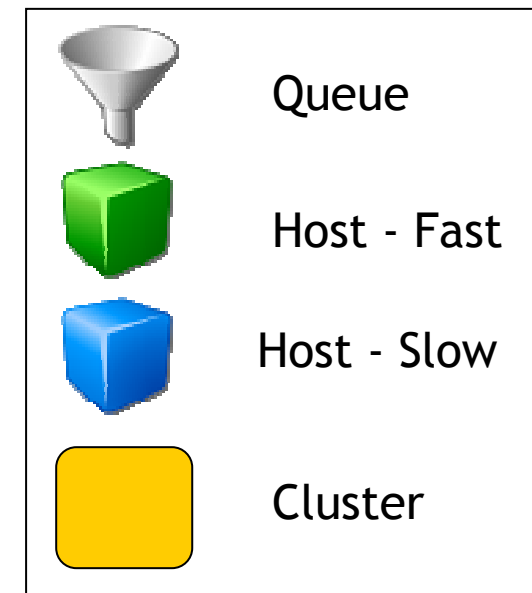
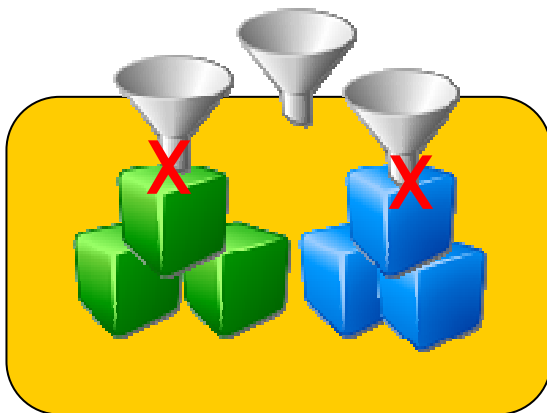
# GLUE Schema - The importance of Use Cases an example

- Site A has 6 worker nodes (3 fresh new and fast, 3 old and slow)
- The farm is configured as follows:
  - a high-end queue to the 3 fast WN's
  - a slow queue to the 3 slow WN's
  - a background queue to the 6 WN's (lower priority)



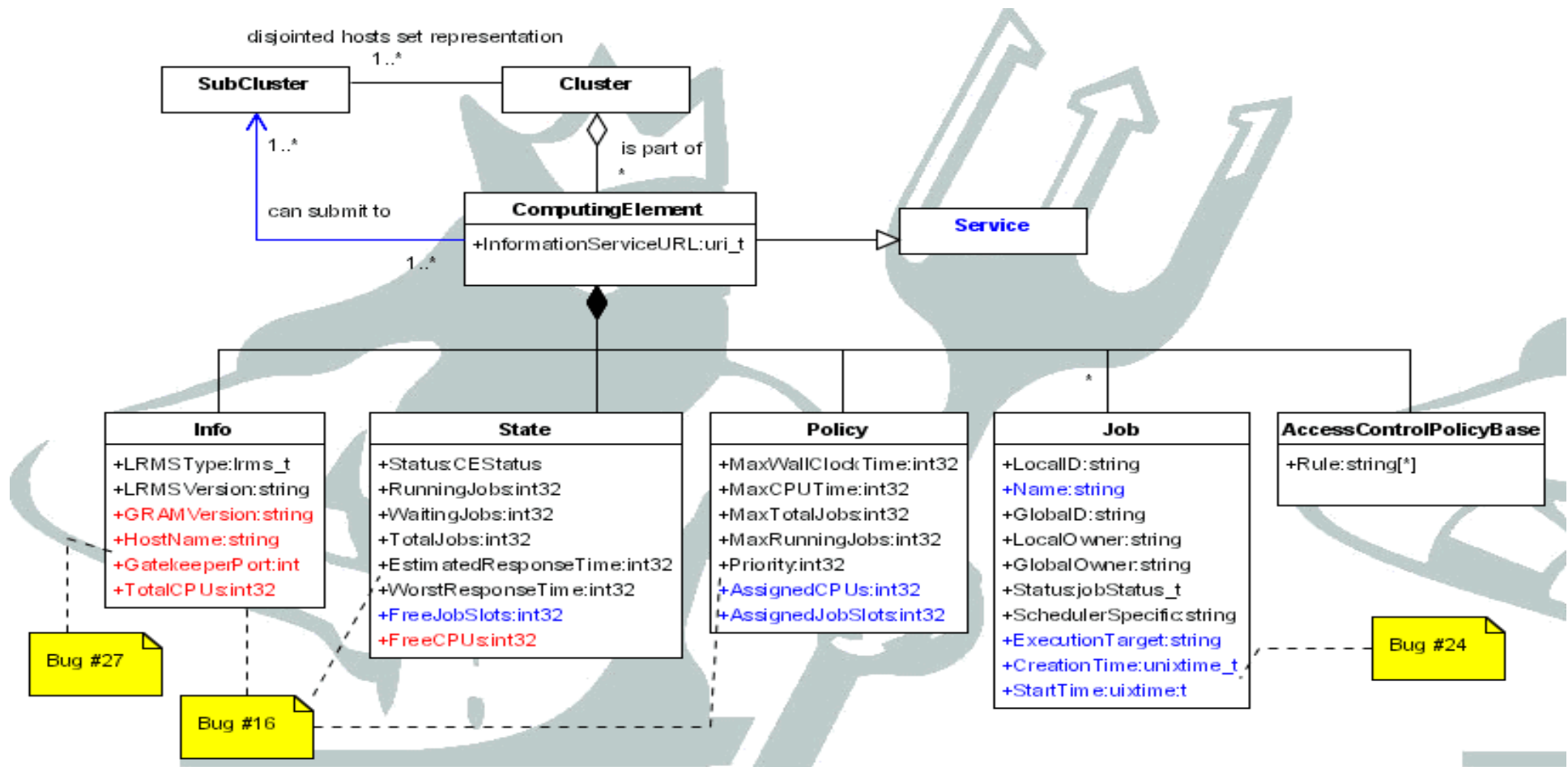
# GLUE Schema - The importance of Use Cases an example

- Representation in the GLUE Schema v. 1.1
- Queue -> Computing Element (CE\_highend, CE\_slow, CE\_background)
- Characteristics of new/fast WN's -> Subcluster A
- Characteristics of old/slow WN's -> Subcluster B
- A+B=Cluster
- **Problem:** there is no explicit relationship between CE and SubCluster; CE are associated to Cluster, hence to all underlying SubCluster



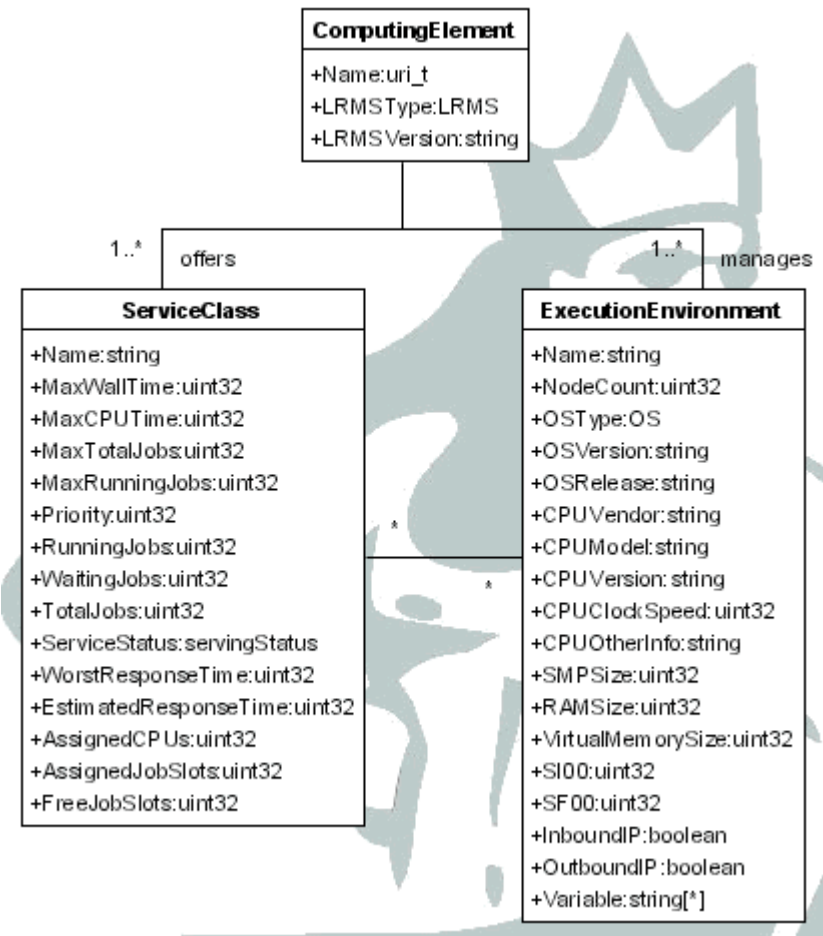
# GLUE Schema - The importance of Use Cases an example

- Possible Fix in GLUE Schema v. 1.2 (Additive approach)
- Add a new relationship: CE (\*) ---can submit to--- (\*) SubCluster



# GLUE Schema - The importance of Use Cases

- Possible Major Refactoring in GLUE Schema v.2.0
  - CE is a site cluster (EGEE Architecture document)
  - Queues are used to differentiate the service
  - The service offers access and management of available execution environments
- ServiceClass:
  - HighEnd, Slow, Background
- ExecutionEnvironment:
  - charact node A, charact node B



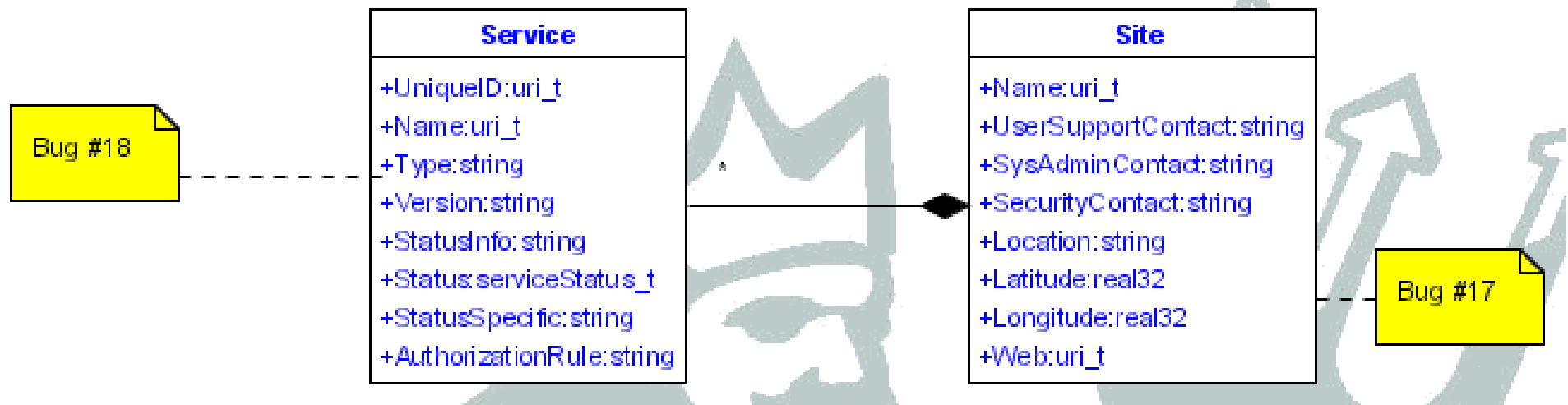


# **GLUE Schema Version 1.2**

## **Draft proposal**

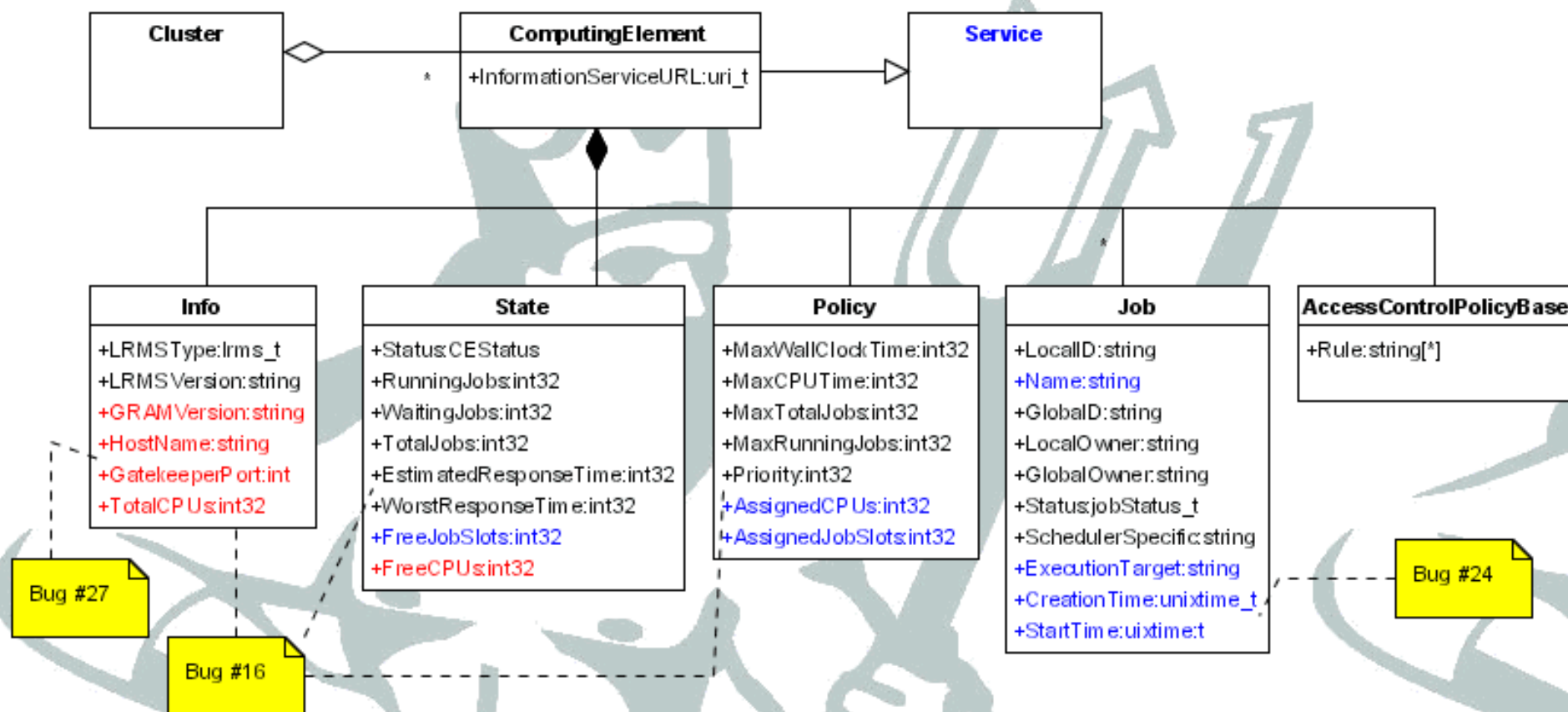
# GLUE Schema 1.2 - Core

New/modified  
Unmodified  
Deprecated



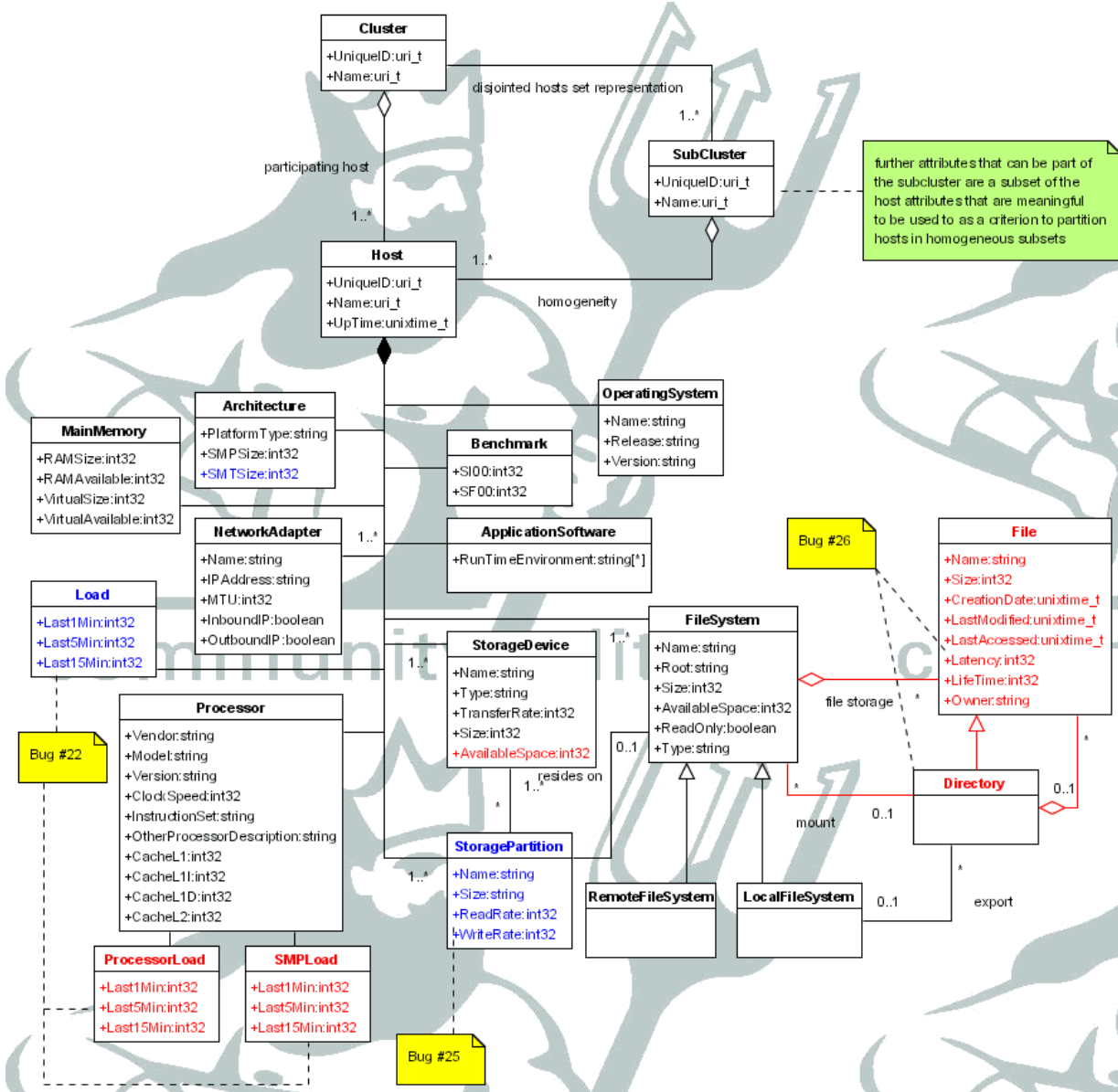
# GLUE Schema 1.2 - Computing

New/modified  
Unmodified  
Deprecated

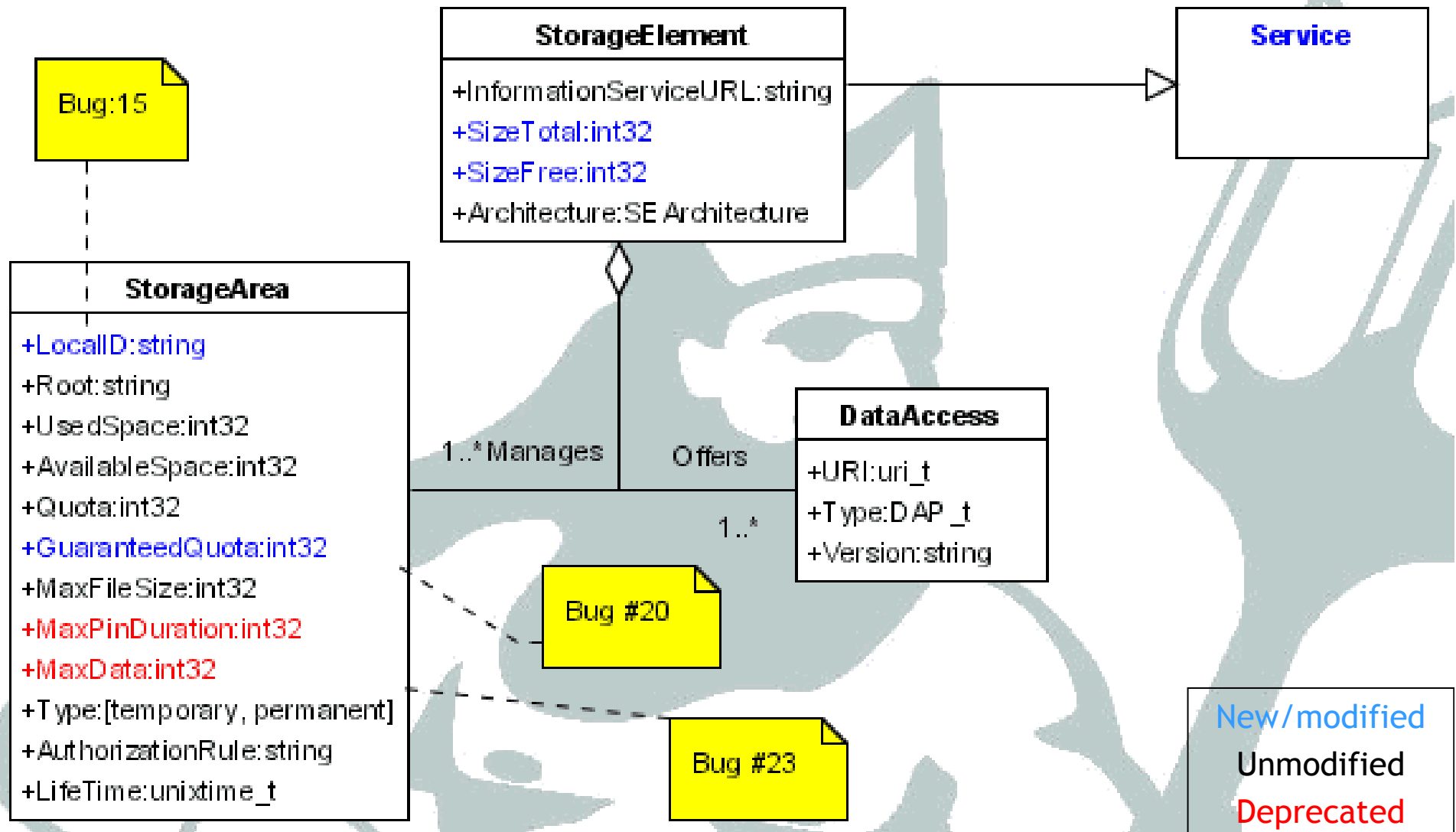


# GLUE Schema 1.2 - Cluster/SubCluster/Host

New/modified  
Unmodified  
Deprecated

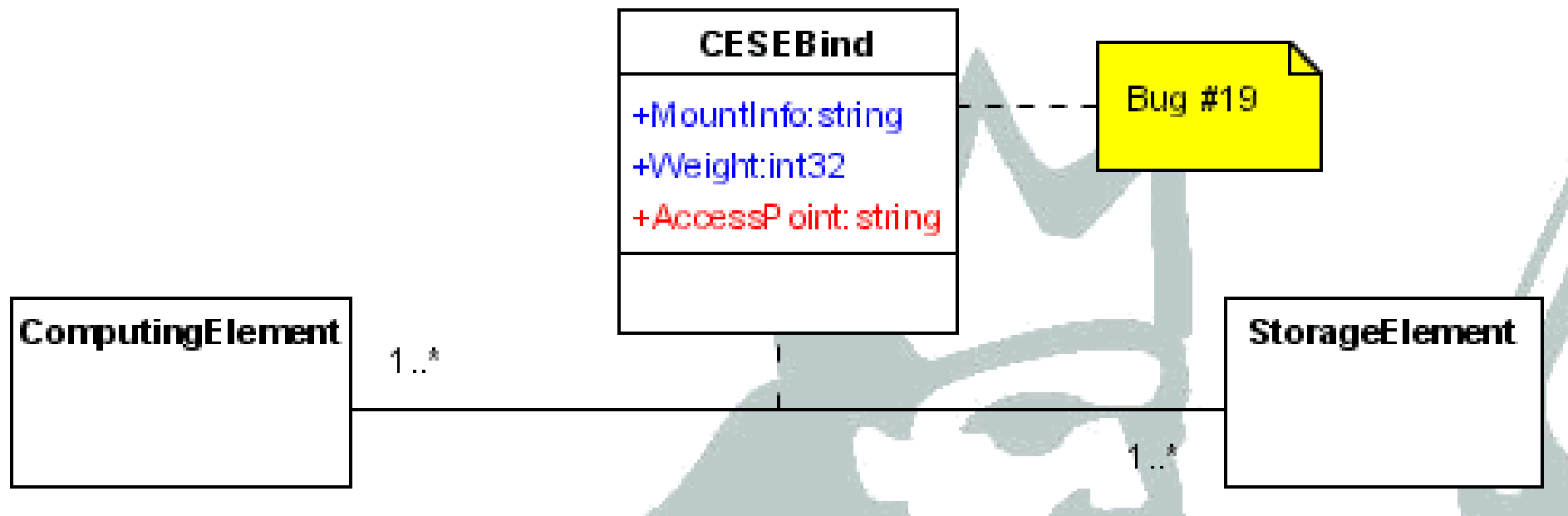


# GLUE Schema 1.2 - Storage



# GLUE Schema 1.2 - CE/SE relationship

New/modified  
Unmodified  
Deprecated



# Bibliography

- [1] **GLUE Schema Resources:** <http://www.cnaf.infn.it/~sergio/glue>
- [2] **GLUE Schema - open issues:**  
[http://infnforge.cnaf.infn.it/docman/view.php/9/65/GLUEInfoModel\\_V\\_1\\_2\\_draft\\_1.pdf](http://infnforge.cnaf.infn.it/docman/view.php/9/65/GLUEInfoModel_V_1_2_draft_1.pdf)
- [3] **GLUE Schema Version 1.2 - First Draft**  
[http://infnforge.cnaf.infn.it/tracker/?atid=118&group\\_id=9&func=browse](http://infnforge.cnaf.infn.it/tracker/?atid=118&group_id=9&func=browse)
- [4] **Grid3 metrics**  
<http://grid.uchicago.edu/metrics/metrics-table.html>  
<http://griddev.uchicago.edu/download/grid3/doc.pkg/monitoring-metrics/Grid3-metrics.doc>
- [5] **Monitoring the Connectivity of a Grid.** S. Andreatto, A. Ciuffoletti, A. Ghiselli, C. Vistoli. In Proc. of the 2nd International Workshop on Middleware for Grid Computing (MGC 2004) in conjunction with the 5th ACM/IFIP/USENIX International Middleware Conference, Toronto, Canada, October 2004.