



Enabling Grids for  
E-science in Europe

*JRA1 all-hands meeting, 29-June-2004*

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

## Information & Monitoring

Rob Byrom



# Contents

- Summary of I&M Use cases
- Questions for discussion

## Some I&M Use Cases

- Service discovery and status enquires
- Job scheduling
- L&B
- Network monitoring
- Accounting
- Job monitoring

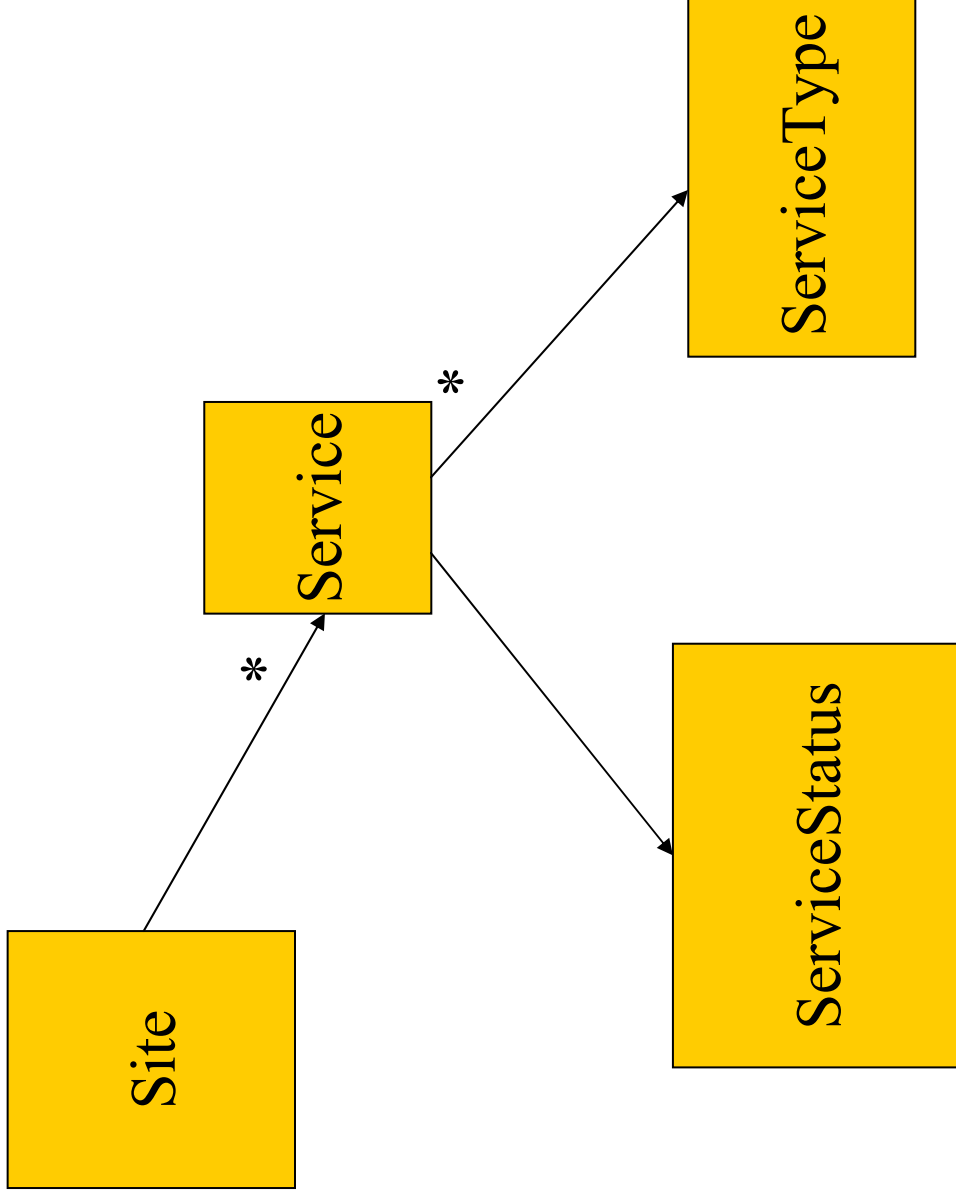
# Service Discovery

- Grid components need a mechanism to lookup and query the status of active services
- Services announce their availability by publishing to a Service table
  - Table contains the URL to locate the Service, a URL of the WSDL and URL of the semantics of the Service
- Dynamic status of the Service is then published to the ServiceStatus table
  - By locally invoked status query

# Service Schema Design Issues

- Service schema designed to separate slow and rapidly changing information
  - Service name and details change relatively slowly (order of day/months)
  - Actual Status of Service continually changes over short intervals (order of minutes)
- Dividing 'static' from dynamic information considered a good implementation principle!

# Service Schema Cont ...



# Job Scheduling

- WMS needs to support both push and pull mechanisms when scheduling jobs to a particular site
- In the push model
  - Each site must publicize its available resources
  - CE/SE info published according to a schema (such as a simplified GLUE Schema)
  - This enables the WMS to ‘match make’ the job to an appropriate site
- In the pull model
  - Each CE encodes its resources and availability when contacting the WMS?

# Logging & Bookkeeping

- Need to track the status of a job once submitted via the UI
- Status updates are logged when a job is allocated to a RB/CE etc
- R-GMA and L&B teams working closely
  - R-GMA to provide guaranteed delivery mechanism



# Networking Monitoring

- Need to provide reliable network profiling to analysis throughput and availability
- Statistics used for finding bottlenecks and for QOS provisioning
- Tools used to analyse network characteristics
  - UDPMon, Ping, IPerf etc
- Output from network analysis tools published via I&M service

# Accounting

- Want to track usage of computing/storage resources
- Accounting application being developed for **LCG** which uses R-GMA
  - Stub program on each site parses event/message logs
  - Important info such as JobId, Username, files requested etc are extracted from logs and then published via R-GMA (using a standard relational schema)
  - GOC collates all published records per site

# Job Monitoring

- Need to capture the output of the job itself, this can be carried using two approaches
  - By accessing the JM Service directly (which then forwards message updates to the I&M Service)
  - Wrapping the I&M calls within the job itself
- Both approaches are conceptually the same
- Suggested approach is to use the apache logging Service which then communicates to the I&M Service
  - User code is instrumented with log4j calls (for example)
  - Apache Logging Service publishes message using the I&M Service

# Open Issues

- Are there more I&M use cases?
- If the WMS adopts a pull model what (if any) information needs to be published by the CE/SE?
  - This isn't clear!