National Institute of Advanced Industrial Science and Technology
AIST

S-MDS: A Semantic Monitoring and Discovery System for the Grid

Yusuke Tanimura
National Institute of AIST

Outline

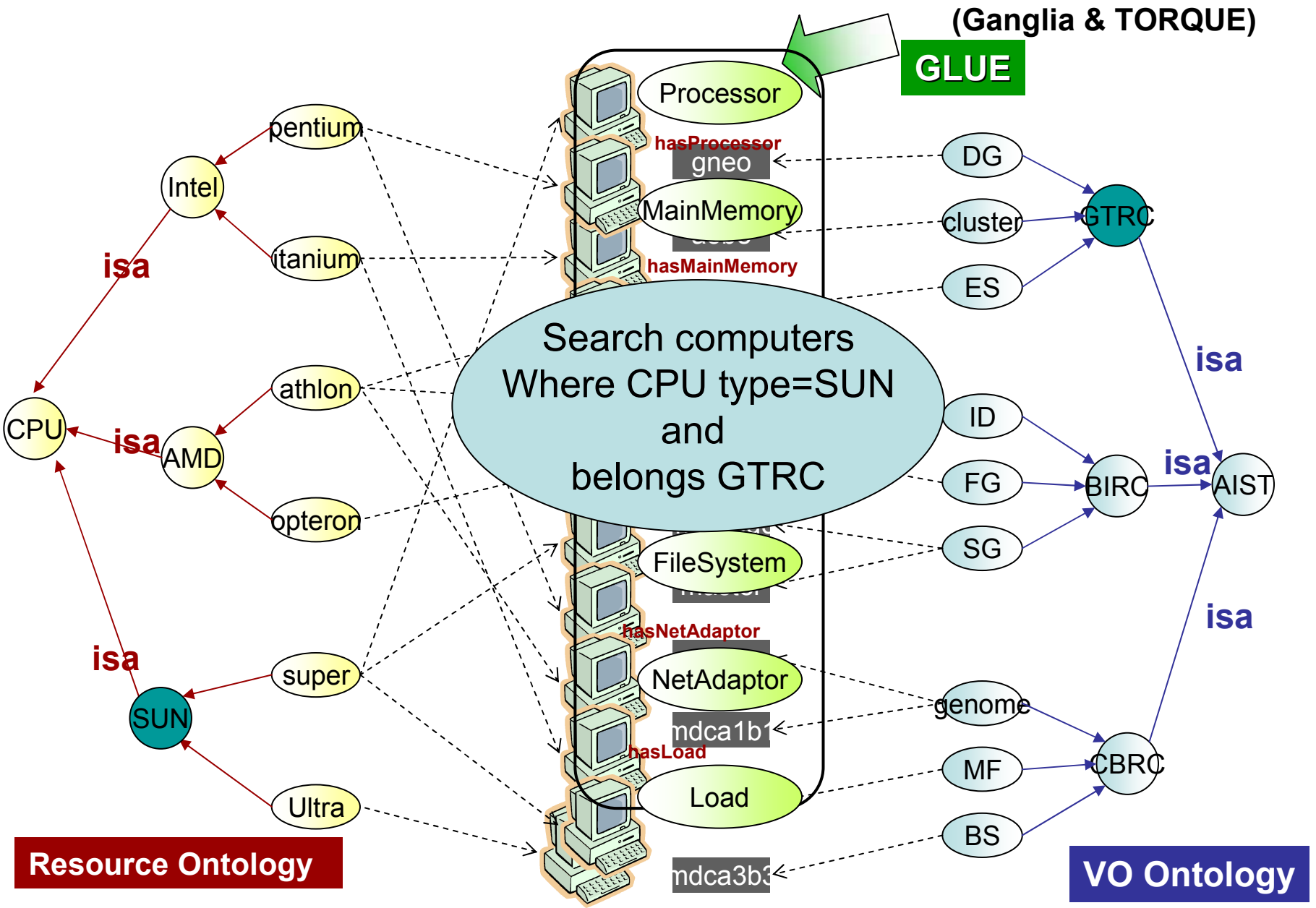
- S-MDS Introduction
- Setting up/Using S-MDS
  - Demos
- Recent Functional Enhancements

What does S-MDS do? (1)

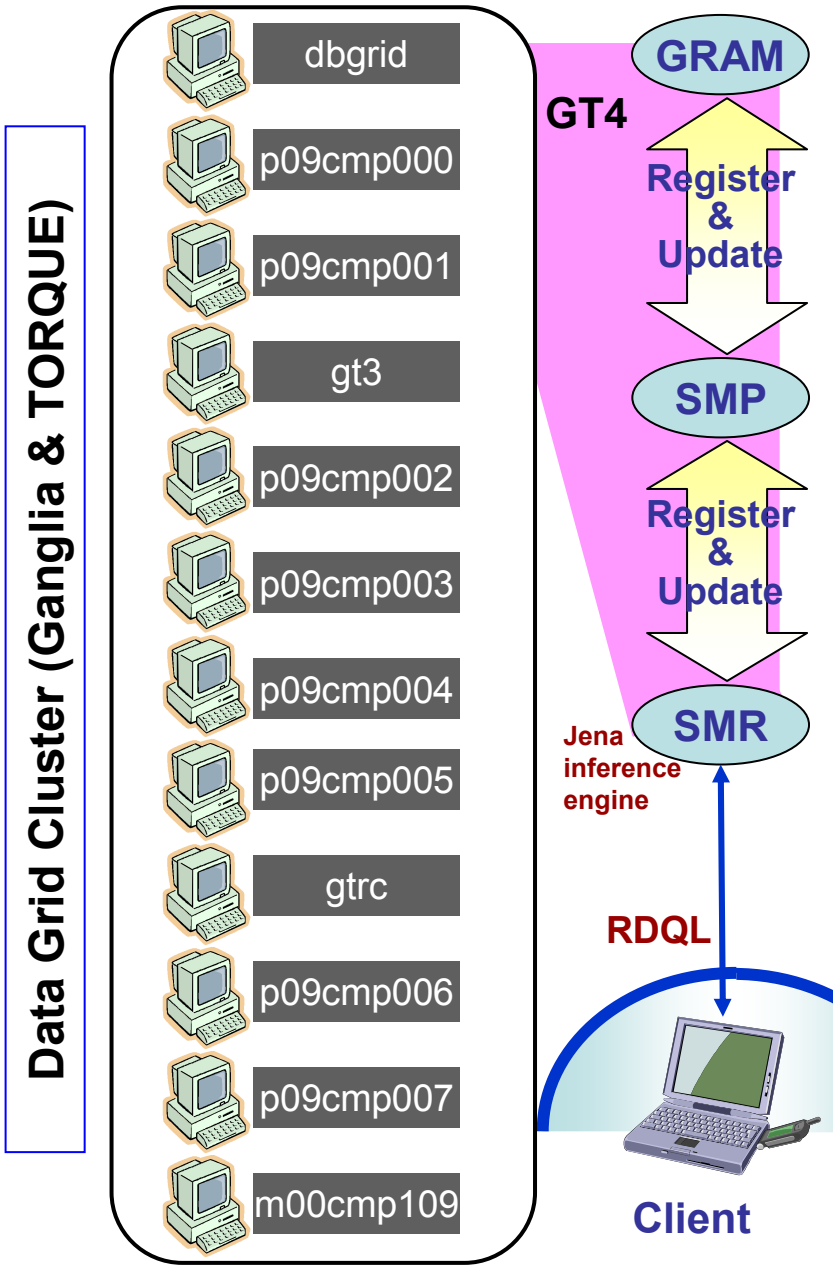
- You can define **Ontologies** over current monitoring system (Ganglia, TORQUE, etc.).
  - Ontology: a set of concepts within a domain and the relationships between those concepts
  - Ex. Athron and Pentium is both Intel-architecture.
  - Ex. GTRC and ITRI both belongs AIST organization.

- Discover the resource using a **Query Language SPARQL**.
  - Ex. Search intel-architecture computers where current load average is under 1 and belong AIST.

Ex. Grid Resource Discovery Using Ontology

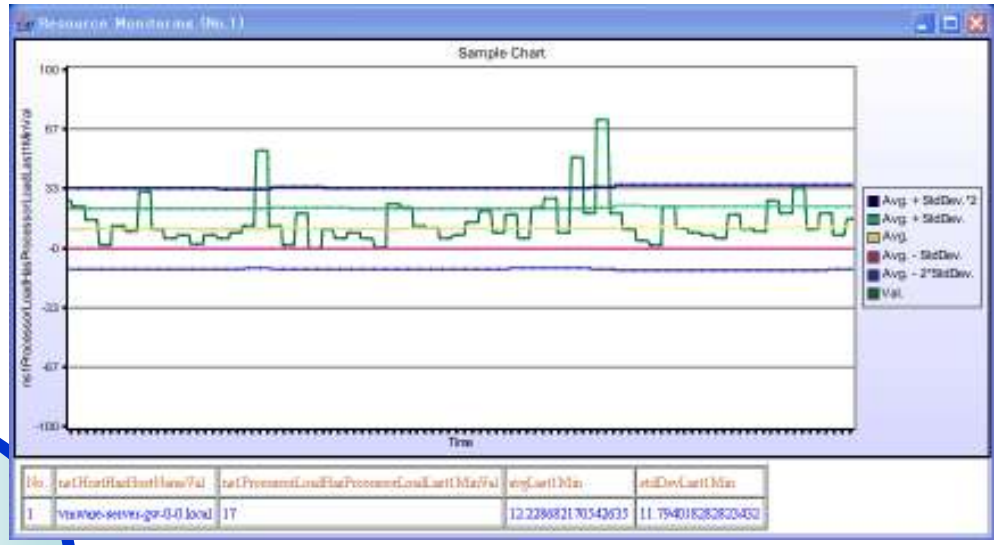


What does S-MDS do (2) ?



- Monitor the resource **continually**
  - Ex. Search loadavg of intel-architecture (monitor mode)

The following graph shows history of the CPU load average and StdDev, which were retrieved by continuous query.



Resource Properties

- Resource/service state is described by Resource Properties (RPs).
  - GetResourceProperty etc.
- Problem: deal only with the document structures but NOT with its contained semantics/concept!!

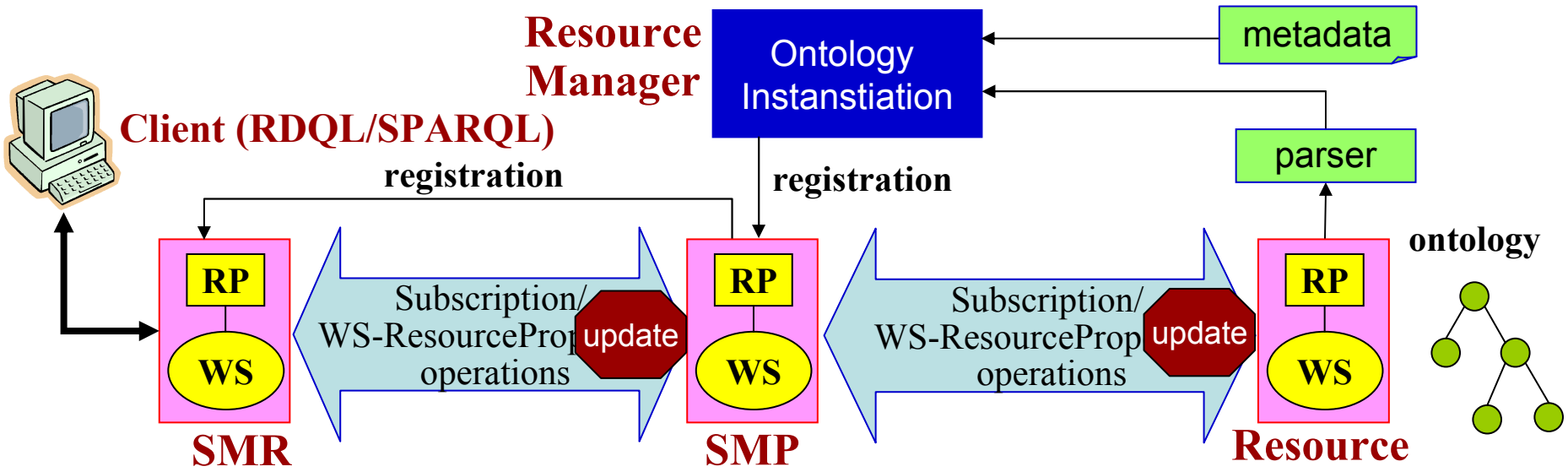
Querying Problems

```
<Cluster Name="datagrid" UniqueID="datagrid">
 <SubCluster Name=" datagrid " UniqueID=" datagrid ">
  <Host Name="dbgrid.x.x" UniqueID="dbgrid.x.x">
   <Processor CacheL1="0" CacheL1D="0" CacheL1I="0"
    CacheL2="0" ClockSpeed="3001" InstructionSet="x86"/>
   <MainMemory RAMAvailable="19" RAMSize="1484"
    VirtualAvailable="332" VirtualSize="2041"/>
   <OperatingSystem Name="Linux"
    Release="2.6.0-test5_2smp"/>
   <Architecture SMPSize="2"/>
   <FileSystem AvailableSpace="65511" Name="entire"
    ReadOnly="false" Root="/" Size="82459"/>
   <NetworkAdapter IPAddress="163.220.2.54"
    InboundIP="true" MTU="0" Name="dbgrid.hpcc.jp"
    OutboundIP="true"/>
   <ProcessorLoad Last15Min="10" Last1Min="6"
    Last5Min="8"/>
  </Host>
  <Host/>
  …
```

- Processable queries
  - Hosts with x86 processor instruction set
    - //Host[Processor@InstructionSet='x86']
  - Host with Linux OS
    - //Host[OperatingSystem[@Name='Linux']]
- Non-processable queries
  - Hosts with intel compatible architecture
  - Hosts with Unix compatible OS

S-MDS

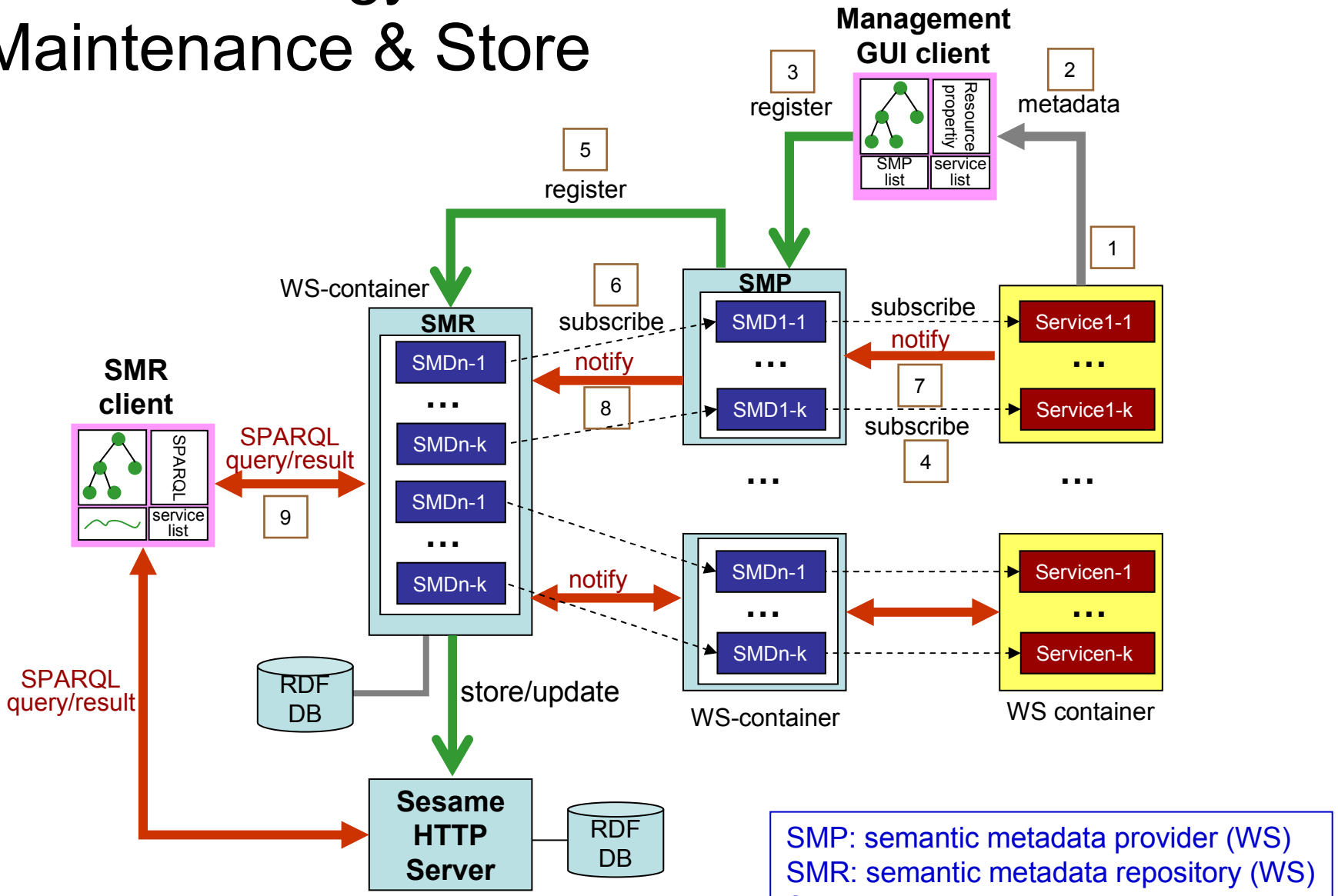
- Define an ontology for each service domain.
  - Resource ontology, VO ontology, GRAM ontology, DAI ontology, …
- Map resource property values into the ontology classes.
- Semantic Metadata Provider (SMP): maintain ontology instance (semantic metadata)
- Semantic Metadata Repository (SMR): instance aggregation & query



What does S-MDS do? (3)

- Describe the resource properties (RP) using ontology : RP←→Ontology mapping.
- Store and maintain the ontology (instances) in a WS or ontology (RDF) repository so that it can be semantically queried.
  - SPARQL or RDF query language (with inference)
- Provide a GUI for easy SPARQL creation.

Ontology Maintenance & Store



SMP: semantic metadata provider (WS)
SMR: semantic metadata repository (WS)
SMD: semantic metadata

Setting Up/Using S-MDS

- Problem & Motivation

Grid user is not an expert of the Semantic Web.

-> Provide GUI-based Support Tools

1. I don't want to describe OWL (Web Ontology Language) for my cluster resource.

Provide Automatic Ontology Mapping Tool

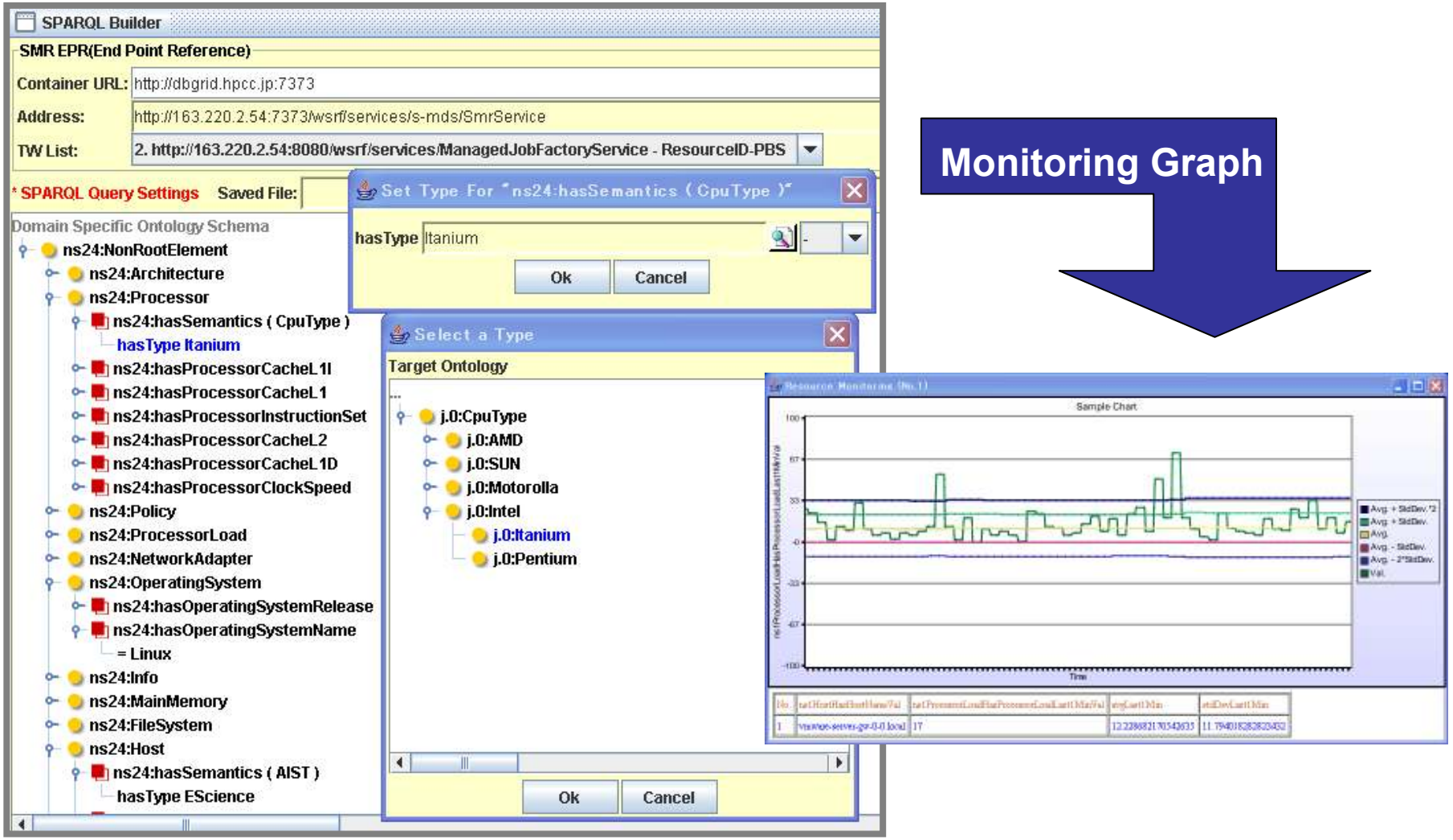
- Convert your ***ResourceProperty*** to ***OWL Ontology***
- Easy to add your ***custom ontology***

2. I don't want to write an SPARQL query for monitoring.

Provide GUI-based Query & Monitoring Tool

- Just view the properties and give conditions

S-MDS: GUI-based Resource Monitoring and Access



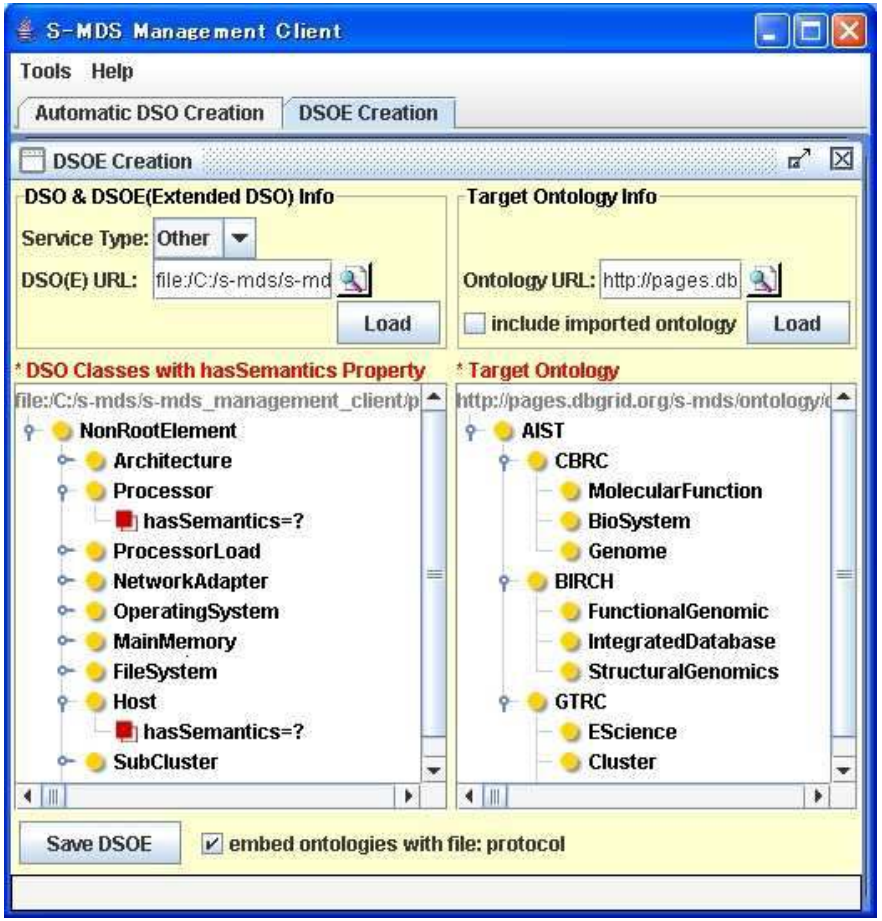
SPARQL Query
(generated by our GUI)

```
SELECT distinct ?ns24ProcessorLoadHasProcessorLoadLast1MinVal
WHERE {
      ?ns24Host ns24:hasProcessor ?ns24Processor.
      ?ns24Processor ns24:hasSemantics ?ns24ProcessorHasSemantics.
      ?ns24ProcessorHasSemantics rdf:type j.0:Pentium.
      ?ns24SubCluster ns24:hasHost ?ns24Host.
      ?ns24ProcessorLoad ns24:hasProcessorLoadLast1Min
          ?ns24ProcessorLoadHasProcessorLoadLast1Min.
      ?ns24Host rdf:type <http://www.dbgrid.org/namespaces/s-mds#Host>.
      ?ns24ProcessorLoadHasProcessorLoadLast1Min ns24:value
        ?ns24ProcessorLoadHasProcessorLoadLast1MinVal.
      ?ns24Cluster rdf:type <http://www.dbgrid.org/namespaces/s-mds#Cluster>.
      ?ns24Host ns24:hasProcessorLoad ?ns24ProcessorLoad.
      ?ns24Processor rdf:type <http://www.dbgrid.org/namespaces/s-mds#Processor>.
      ?ns24Cluster ns24:hasSubCluster ?ns24SubCluster.
      ?ns24SubCluster rdf:type <http://www.dbgrid.org/namespaces/s-mds#SubCluster>.
      ?ns24GLUECE rdf:type <http://www.dbgrid.org/namespaces/s-mds#GLUECE>.
      ?ns24Host ns24:hasSemantics ?ns24HostHasSemantics.
      ?ns24HostHasSemantics rdf:type j.1:DataGrid.
      ?ns24GLUECE ns24:hasCluster ?ns24Cluster.
      ?ns24ProcessorLoad rdf:type <http://www.dbgrid.org/namespaces/s-mds#ProcessorLoad>.
}
```

Demo of S-MDS GUI (1)

- Resource Discovery
- Resource Monitoring

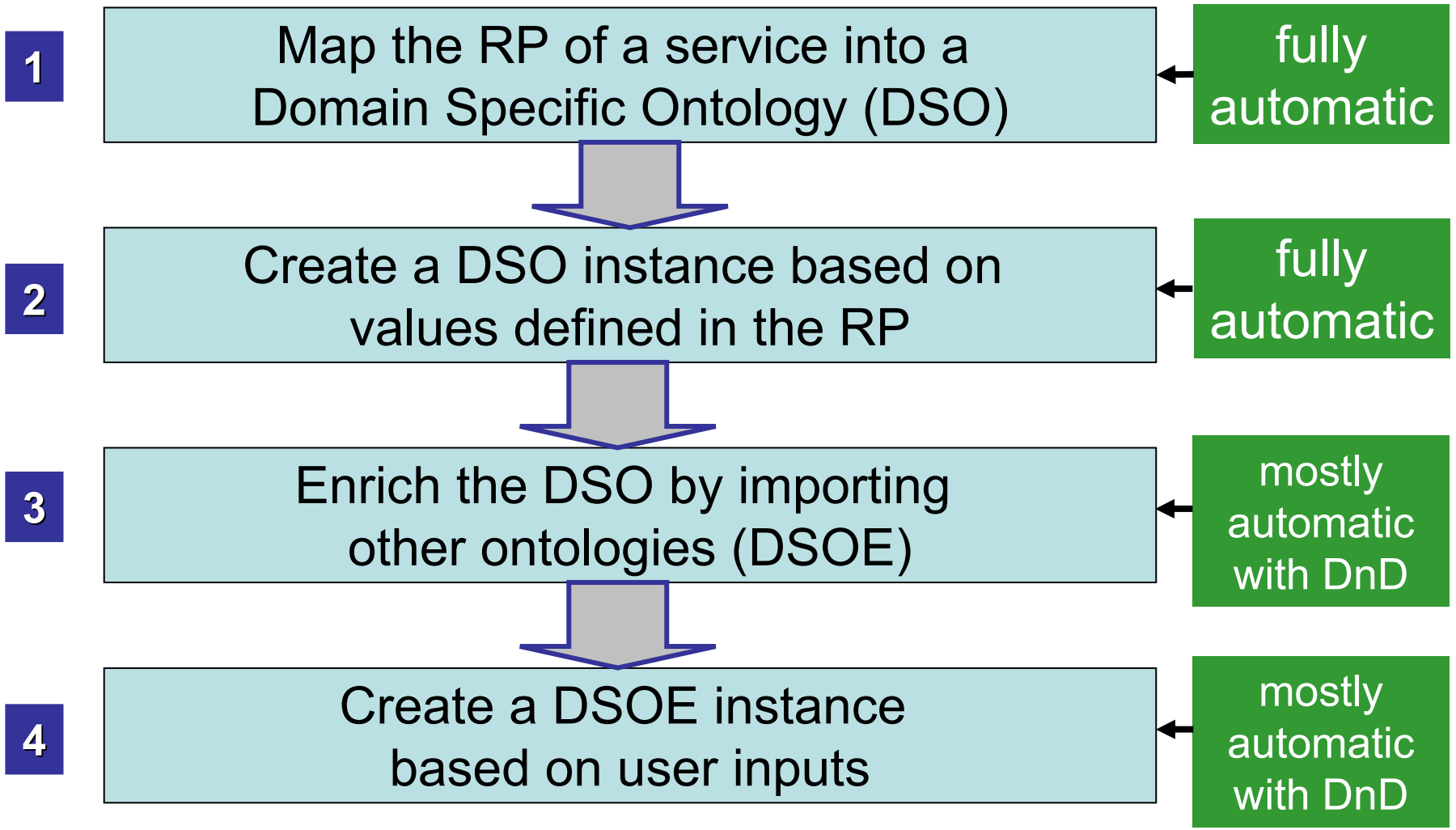
GUI-based Tool for Ontology Mapping



Setup procedure with GUI

1. Map the ResourceProperty to OWL-S ontology (Domain Specific Ontology, DSO).
2. Create Instances for Target Resources (DSO Instances).
3. Add a user-defined ontology to DSO (DSO enriched, DSOE).
4. Create Instances with DSOE (DSOE Instances).

Steps for Ontology Mapping

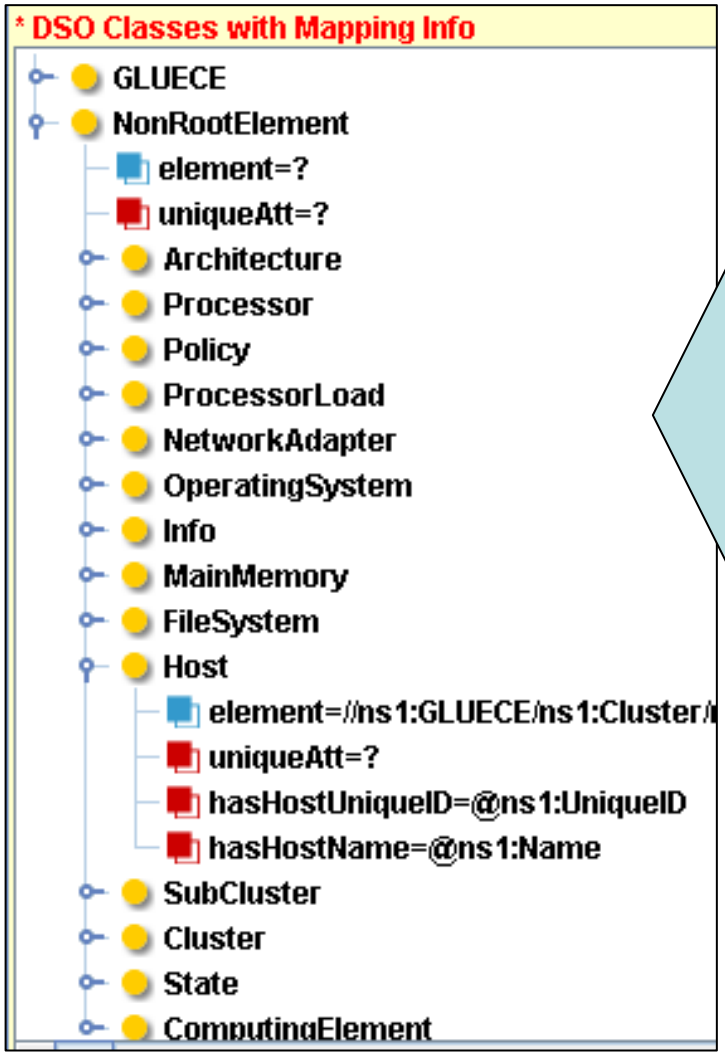


Automatic DSO Creation

1

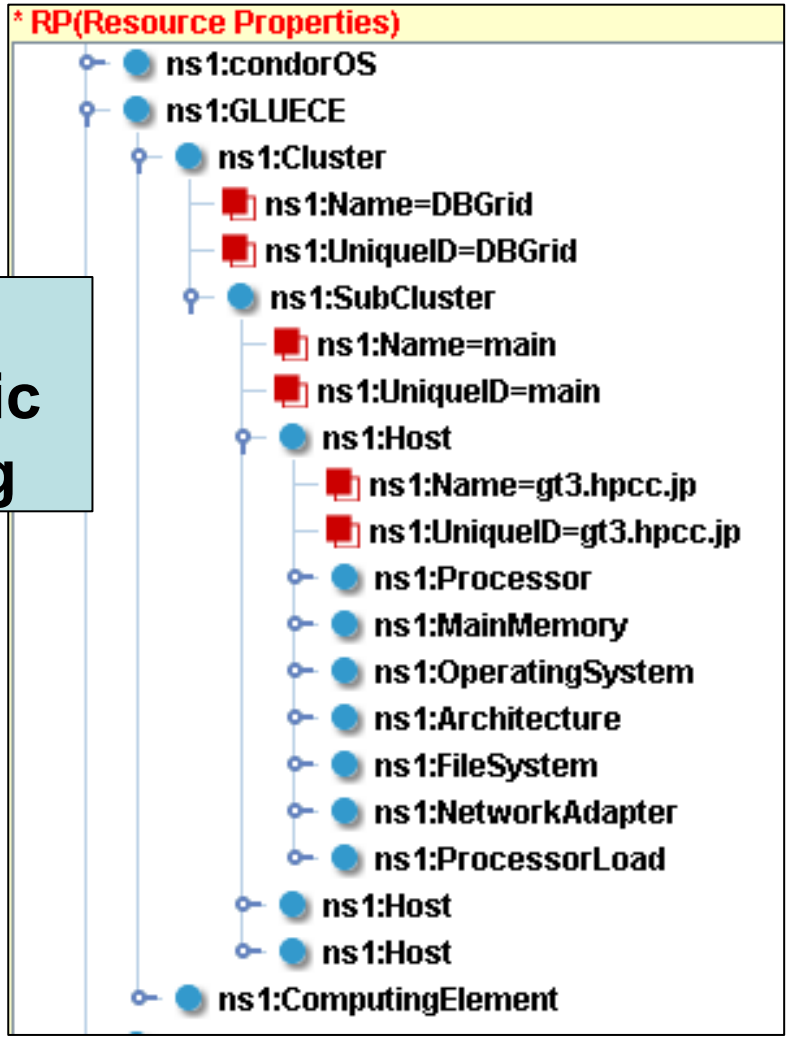
(Create a Class)

Domain Specific Ontology (OWL-S)



Fully Automatic Mapping

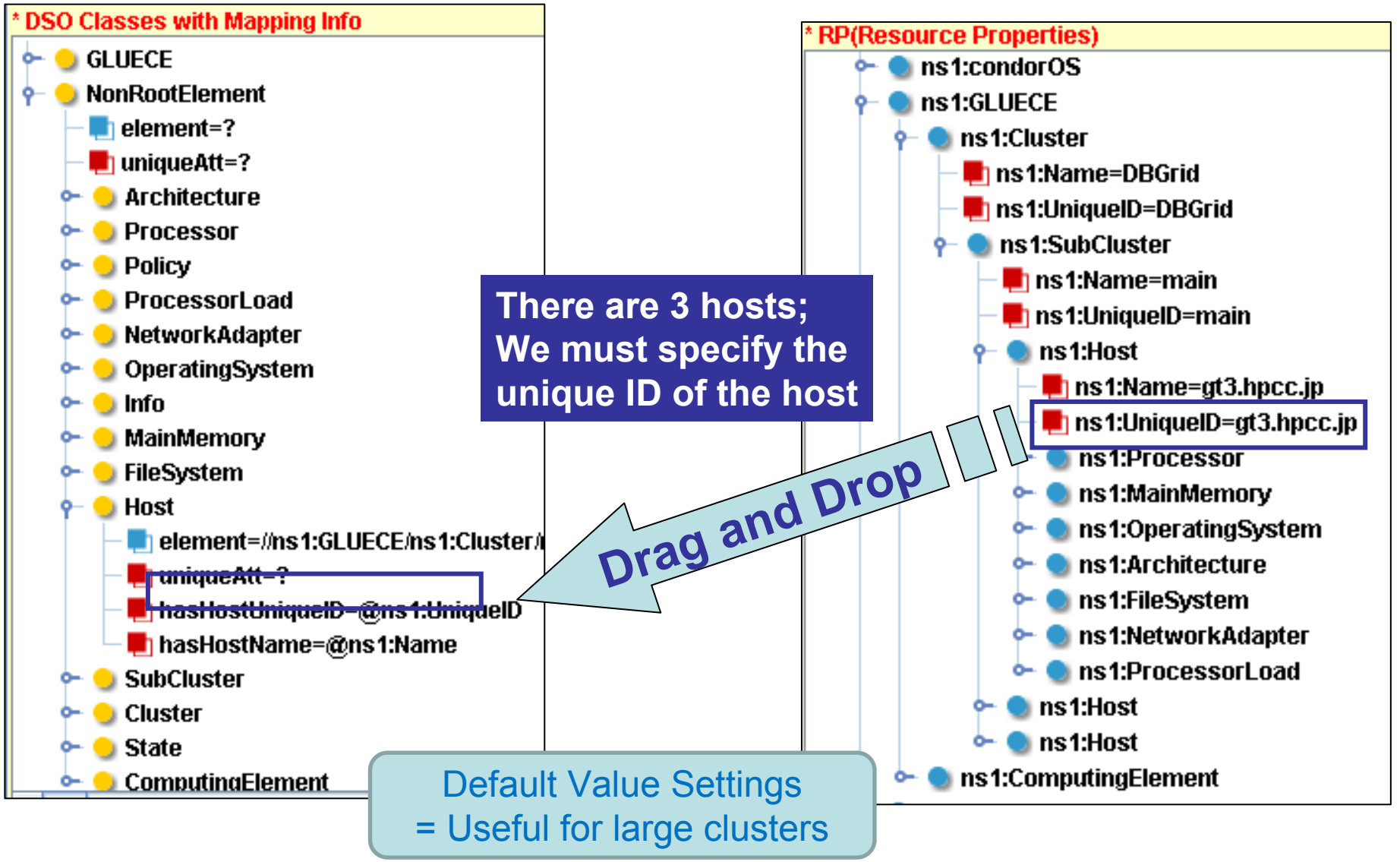
ResourceProperties (XML)



2

DSO Instance Creation

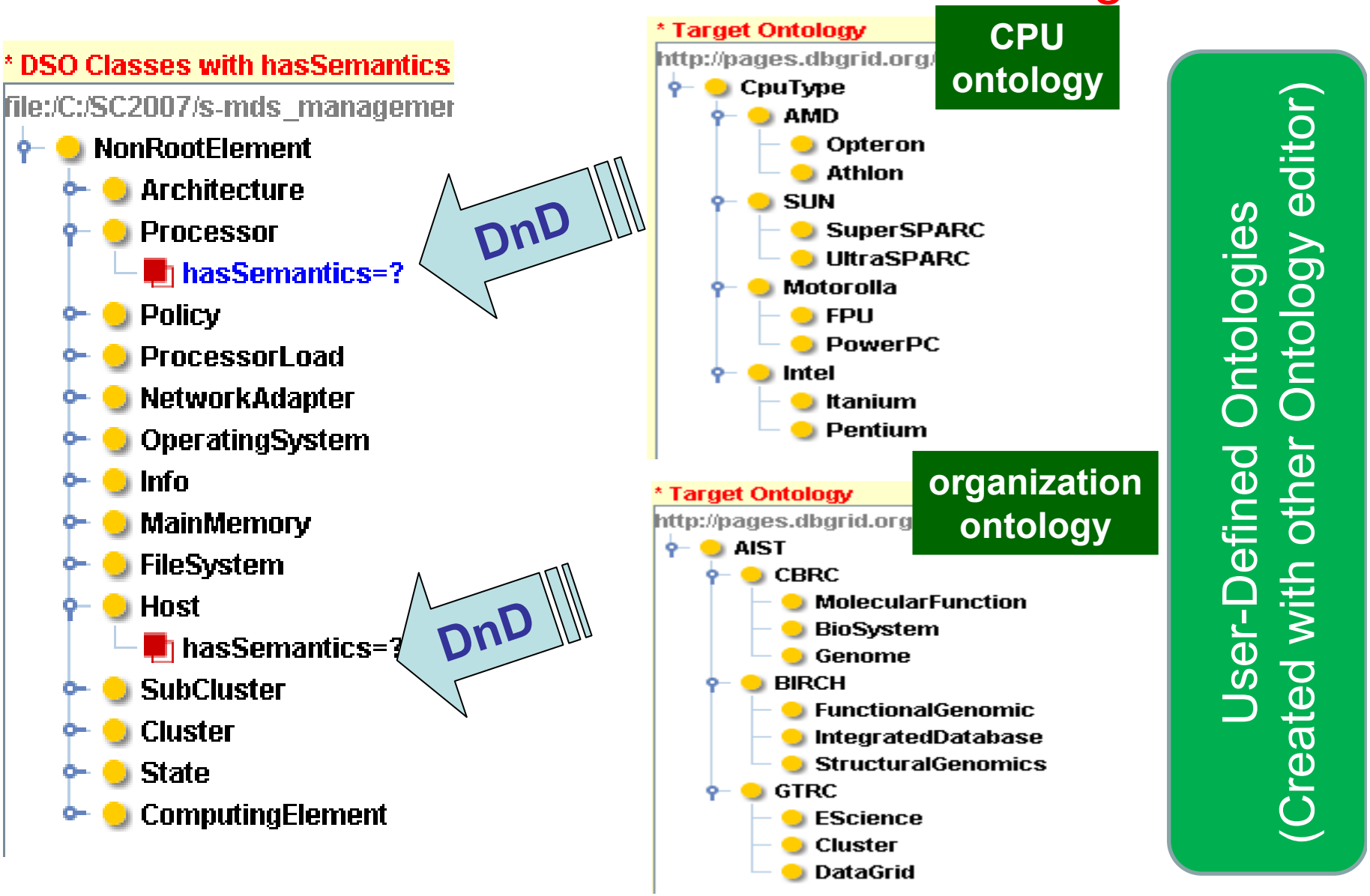
Create Instances of the resources



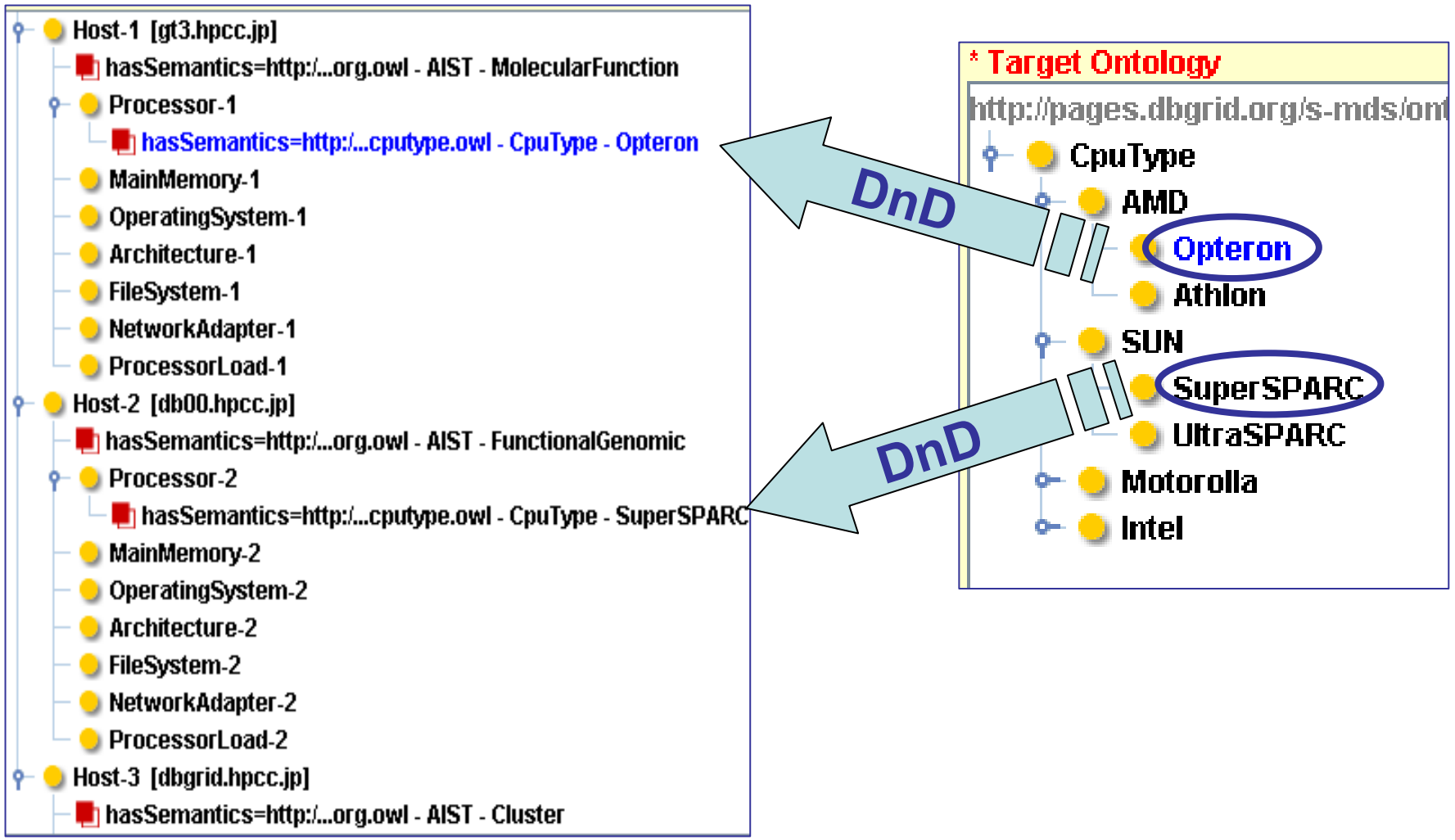
DSO Enriching/Extending

3

Associate with other user-defined ontologies



4 DSOE Instance Creation



Demo of S-DMS GUI (2)

- Ontology Mapping

Software Status

- Technical Preview 1 is available at http://dbgrid.org/smds/.
  - Please download and install if you 're interested.
  - Any feedbacks are welcome.
- Technical Preview 2 will be available soon.
  - Use of UsefulRP interface
  - Functional Enhancements
- Plan to apply Globus Incubator Project

New Functionalities for Tech. Preview 2

Rule-based Monitoring

- Record the history of the monitored value into the SMR database
  - LoadAvg, RAM/DiskFree, etc.
  - Ex: Store the history of LoadAvg for last 2 weeks.
- Statistical processing over the historical data
  - Average, Min, Max,,,
  - Ex: Caluculate the average and standard deviation of LoadAvg for every updates.
- Rule based Status Description & Detection
  - Based on Jena rule processing module
  - Ex: If LoadAvg is more than 3 x stddev+average of Loadval, then report anormaly.

Summary

- Introduction of S-MDS
  - Ontology-based information store & retrieval
  - Enhancement of retrieved-information analysis
    - Monitoring specific information by continuous query
    - Rule-based monitoring
- Setting up/Using S-MDS
  - GUI-based management & query tools (You saw the demos.)
- More information about S-MDS
  - **http://dbgrid.org/smds/**
    - Download source code
    - See papers and tutorial
  - If you have any questions, please email to **dbgrid@m.aist.go.jp**.