Security Requirements Model for Grid Data Management Systems

Syed Naqvi\textsuperscript{1}, Philippe Massonet\textsuperscript{1}, Alvaro Arenas\textsuperscript{2}

\textsuperscript{1}Centre of Excellence in Information and Communication Technologies (CETIC) 
{syed.naqvi, philippe.massonet}@cetic.be

\textsuperscript{2}CCLRC Rutherford Appleton Laboratory 
a.e.arenas@rl.ac.uk
Outline

• Introduction
• Grid Data Management Systems (GDMS)
• GDMS Security Requirements Model
• Deriving Security Policies from Requirements Model
• Conclusions
Outline

• Introduction
• Grid Data Management Systems (GDMS)
• GDMS Security Requirements Model
• Deriving Security Policies from Requirements Model
• Conclusions
Critical Information Infrastructure (CII)

- **Critical infrastructures** consist of those physical and information technology facilities, networks, services and assets which, if disrupted or destroyed, would have a serious impact on the health, safety, security or economic well-being of citizens or the effective functioning of governments in the Member States.

  *Commission of the European Communities*  
  *Critical Infrastructure Protection in the Fight against Terrorism*  

- **Critical information infrastructures** are communications or information service whose availability, reliability and resilience are essential to the functioning of a modern economy, security, and other essential social values. Markets depend on them, as much as governments, to function properly.

  Ensuring (and Insuring?) Critical Information Infrastructure Protection, Report of the 2005 RUESCHLIKON Conference on Information Policy
Gird Computing – Wikipedia definition

- Grid computing is an emerging computing model that provides
  - the ability to perform higher throughput computing
    - by taking advantage of many networked computers
      - to model a virtual computer architecture that is
        - able to distribute process execution across a parallel infrastructure.

- Grids provide
  - the ability to perform computations on large data sets
  - the ability to perform many more computations at once than would be possible on a single computer

- Today resource allocation in a grid is done in accordance with SLAs (service level agreements)
The GRID
Grid as a Critical Information Infrastructure

• Now Grid computing is becoming a critical component of day-to-day business. Today’s challenging business requires continuous innovation …

IBM Grid Computing

• Grid computing is critical infrastructure for a knowledge based economy, in which the raw materials are data. Therefore, it is critical that Grid infrastructure be built to accommodate the data explosion …

Walter Stewart,
Grid Today Magazine, 21 March 2005
Outline

• Introduction
• Grid Data Management Systems (GDMS)
• GDMS Security Requirements Model
• Deriving Security Policies from Requirements Model
• Conclusions
**Functional View of Grid Data Management**

taken from www.twgrid.org

- **Application**
- **Planner:**
  - Data location,
  - Replica selection,
  - Selection of compute and storage nodes
- **Executor:**
  - Initiates data transfers and computations
- **Metadata Service**
- **Replica Location Service**
- **Information Services**
- **Security and Policy**
- **Data Movement**
- **Data Access**
- **Compute Resources**
- **Storage Resources**

**Location based on data attributes**
- Location of one or more physical replicas
- State of grid resources, performance measurements and predictions
**FileStamp Architecture**

- **Layer 0**: Key-based routing (KBR) layer: *Pastry*
- **Layer 1**: Distributed hash table (DHT): *PAST*
- **Layer 2**: Distributed file system: *Pastis*
- **Application**
FileStamp Architecture

Based on a Peer-to-Peer network
Self managing data storage location
File Redundancy

Dynamic replica regeneration
BitTorrent Technology

Moreover transfers can be interrupted and restarted from the last transferred bytes
Outline

• Introduction
• Grid Data Management Systems (GDMS)
• GDMS Security Requirements Model
• Deriving Security Policies from Requirements Model
• Conclusions
Higher-level Goals

- Files always securely available
- Security level always maintained
- Files readily accessible
- Data confidentiality always maintained
- File redundancy maintained
- Files moved to new node
- New nodes identified
- Node failures detected
- File transfer encrypted
Agents do the Attainment of Goals
Operations

- Files moved to new node
- File transfer encrypted
- Move files
- Node failures detected
- New nodes identified
- Monitor
- Notify
- Input
- Node
- Find available node
Dealing with Obstacles

- Node failures detected
  - Node failure undetected
    - Monitor failure
    - Network down
Overall Model
Outline

• Introduction
• Grid Data Management Systems (GDMS)
• GDMS Security Requirements Model
• Deriving Security Policies from Requirements Model
• Conclusions
## Policy Templates

<table>
<thead>
<tr>
<th>ID</th>
<th>Policy identifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Explanation of the policy parameters (optional)</td>
</tr>
<tr>
<td>Subject</td>
<td>Active entity that manages object(s) through a set of actions</td>
</tr>
<tr>
<td>Object</td>
<td>Passive entity that is managed by subject(s) through a set of actions</td>
</tr>
<tr>
<td>Action</td>
<td>Task to be executed by a subject on object(s)</td>
</tr>
<tr>
<td>Authorization</td>
<td>Privileges given to the subject to perform actions on the object. Authorization maybe restricted by constraints</td>
</tr>
<tr>
<td>Constraint</td>
<td>Conditions that need to be fulfilled before an action is initiated.</td>
</tr>
<tr>
<td>Event</td>
<td>Condition that triggers the policy</td>
</tr>
</tbody>
</table>
Example Policy

New replica of file is generated when an existing storage node is failed

<table>
<thead>
<tr>
<th>ID</th>
<th>NFRG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>NFRG: New File Replica Generation</td>
</tr>
<tr>
<td>Subject</td>
<td>Data Monitor</td>
</tr>
<tr>
<td>Object</td>
<td>Grid data storage nodes</td>
</tr>
<tr>
<td>Action</td>
<td>Replica generated</td>
</tr>
<tr>
<td>Authorization</td>
<td>Create files replica</td>
</tr>
<tr>
<td>Constraint</td>
<td>Availability of nodes</td>
</tr>
<tr>
<td>Event</td>
<td>Replica-host node failed</td>
</tr>
</tbody>
</table>
Further Treatment of the Derived Policies

- **Refinement of High level Policies into Operational Policies**
  - Determination of the resources that are needed to satisfy the requirements of the policy.
  - Translation of the high-level policies into operational policies that the system can enforce.
  - Verification that the lower level policies actually meet the requirements specified by the high level policies.

- **Implementation of Policies**
  - Requires specific details of a particular system.
  - Formal representation techniques are employed.
  - Implemented in a specification language.
Outline

• Introduction
• Grid Data Management Systems (GDMS)
• GDMS Security Requirements Model
• Deriving Security Policies from Requirements Model
• Conclusions
Conclusions

• Grid exhibits all the characteristics of Critical Information Infrastructure.
• Grid data management requires well-secured infrastructure like any other critical infrastructure.
• The complexity of Grid architecture makes it impossible to evaluate its security requirements by simple examination.
• We propose the use of formal requirements engineering techniques to assess the grid security requirements.
• A formal security requirements model provides a sound basis for the derivation of system’s security policy.

• Our future directions include:
  – Refinement of security policies derived from the requirements model.
  – Work on negotiation protocols to assure service level security agreements.
  – Implementation of security policies in a grid data management system.
Thank you!