Fine-Grained Access Control for GridFTP Using SecPAL

Marty Humphrey#, Sang-Min Park#, Jun Feng#, Norm Beekwilder#, Glenn Wasson#, Jason Hogg*, Brian LaMacchia*, and Blair Dillaway*

#Department of Computer Science, University of Virginia
*Microsoft Corporation, 1 Microsoft Way, Redmond, WA
Information Access in Campus Grids

- Only campus-certified individuals
- Only faculty
- NOT students
- “emergency access”: Fred is allowed to act as Susan
  - W/O direct intervention from Susan
  - Susan actively delegates to Fred
- Anyone with a particular (virtual) key (“capability”)
The Problem: Policy

• Policy: the rules [governing access]

• Problem: Access control languages are generally either:
  • Too coarse / simple
  • Too complicated / unintuitive

• Implementation of policy languages are non-trivial as well
Grid Data Access: The Players

- **Data Owner**: scientist who creates the data
- **Resource Providers**: make the data available on the Grid
- **Virtual Organization**: Can impose additional policies
- **Data Requester**
Requirements: General (1/2)

- **[G-R1]** It must be possible to express both fine-grained access control policy (e.g., method-level, file-level, data-record-level) as well as coarse-grained access control policy (e.g., service-level, host-level, or VO-level).
- **[G-R2]** Authorization decisions must be provably correct and should be guaranteed to terminate.
- **[G-R3]** It must be agnostic to existing and future authentication policies (e.g., “no cleartext passwords on the wire”) and mechanisms (e.g., SSH) as well as information providers.
- **[G-R4]** It must be possible to broadly express policies for information sources used in the evaluation of access control policies (e.g., to specify that a particular source is authorized to provide certain information).
Requirements: General (2/2)

• [G-R5] Policies must be composable and extensible without requiring centralized policy authoring.

• [G-R6] Users should not be required to be an expert in computer security to author or understand an access control policy.

• [G-R7] It must be possible to specify a lifetime on a policy and policies should be able to be modified during their lifetime.

• [G-R8] It must be possible to delegate a subset of a principal’s rights.
Requirements: Data Owner

- [DATA-OWNER-R1] *It must be possible to specify role- and attribute-based authorization policies (for scalability)*
- [DATA-OWNER-R2] *It must be possible to specify policy specific to an access mode and purpose of the attempted access.* (e.g., access for particular operations must be over an encrypted channel).
Requirements: Resource Owner

- [RESOURCE-OWNER-R1] *It must be possible to specify policy based on time and access mode.*
Requirements: VO

• [VO-R1] *It must be possible to define an acceptable set of authorities for the virtual organization as a whole.*
SecPAL Introduction

• Declarative, logic-based, security policy language
  • Easily read as English sentences with a restricted grammar
  • Formal model designed by MSR Cambridge to guarantee policy composability and tractibility
  • Simple syntactic checks ensure evaluation safety
  • XML-based
    • Serializes as XML, XMLDSIG, XMLENC.
    • Integrates with WS-Security, WS-Trust (etc)

• Complete solution for Grid access control requirements
  • Trust, authorization, delegation policies, auditing, PKI for identity management
  • Can work with existing identity management mechanisms

• SecPAL is research, it is not a product
  • A public research release has been made available for security researchers to evaluate SecPAL, develop proof of concepts and provide feedback
SecPAL Grammar Overview

- **Assertion (Token, Policy, AuthorizationQuery)**
  - A says claim

- **Claim (conditional fact)**
  - \(\text{fact if fact1, fact2, \ldots , factn where constraint}\)

- **Fact**
  - \(P \text{ can verb resource [qual]}\) (action)
  - \(P \text{ possesses attrib=value [qual]}\) (possession)
  - \(P \text{ can say fact}\) (delegation)
  - \(P \text{ can act as } P'\) (alias)
  - \(P \text{ revokes [ClaimId]}\) (revocation)

- **Variables**
  - Support creation of generic policies
  - Prefixed by % signs

**Example:**

ResourceGuard says Bob can read [http://foo.com/] if Bob can write [http://foo.com/]

Integrating SecPAL into GridFTP.Net

1. Data Access Request
2. Engage SecPAL
3. Get attributes
4. Get additional tokens, if any
5. Get policies
6. Get AuthZ Decision
7. AuthZ Response
8. Requested Data (If Authorized)
Use-case 5 - Constrained Delegation

• How does a file repository grant access to users from the University of Virginia – plus allow such users to delegate some subset of their rights to another user

• Requires three types of policies:
  • Trust relationship
  • Resource access policy
  • Delegation policy
# Use-case 5 - Constrained Delegation

## Security Policies

| Trust policy | LA says UVa-CA **can say** %p possesses %a (from %t1 until %t2)  
where  
%t2 - %t1 <= "365.00:00:00",  
%t1 <= CurrentTime() <= %t2,  
%a matches rfc822Name:":.*@virginia\edu" |
| --- | --- |
| Resource access policy | LA says %p **can read, write** digitalContent:"ile://GridFTPRoot/" if  
%p possesses %a  
where %a matches rfc822Name:":.*@virginia\edu" |
| Enable delegation | LA says %p **can say** %x can %v  
digitalContent:"ile://GridFTPRoot/" if  
%p **can** %v digitalContent:"ile://GridFTPRoot/" |

## Security Token

| Identity token | UVa-CA says K-User1 **possesses**  
rfc822Name:"humphrey@virginia.edu"  
(from "2007-01-01T00:00:00Z" until "2007-12-31T00:00:00Z") |
| --- | --- |
| Delegation token | K-User1 says K-User2 **can read**  
digitalContent:"ile://GridFTPRoot/" |

## Authorization Query

LA says K-User2 **can read** digitalContent:"ile://GridFTPRoot/data1.txt"?
Demonstration
(time permitting)
Evaluation (Qualitative)

• We believe that most of the requirements are clearly met
  • **Uniquely strong support** for provably correct decisions (GENERAL-R2), policies on info sources (GENERAL-R4), and delegation (GENERAL-R8)
  • **Needing further study**: the ability to author and compose policies (GENERAL-R5, GENERAL-R6)
Evaluation (Quantitative – 1/3): Legacy GUI “Connect”

<table>
<thead>
<tr>
<th>gridmap</th>
<th>LDAP</th>
<th>SecPAL-based system</th>
</tr>
</thead>
<tbody>
<tr>
<td>652.57 ms</td>
<td>893.06 ms</td>
<td>712.38 ms</td>
</tr>
</tbody>
</table>

- The two authorization policies we used in this experiment were:
  - *@virginia.edu can access the service
  - role == faculty can write in gridFTPRoot
- SSL Handshake was dominant cost, irrespective of the authorization system, requiring approximately 650 ms in our tests
**Evaluation (Quantitative – 2/3): Durations for SecPAL-Based System**

<table>
<thead>
<tr>
<th>SecPAL Token Gen.</th>
<th>UVa LDAP</th>
<th>Security Token Web Service</th>
<th>Policy Repo.</th>
<th>SecPAL Engine</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.89 ms</td>
<td>3.96 ms X3</td>
<td>2.72 ms X4</td>
<td>2.50 ms</td>
<td>11.58 ms</td>
<td>49.95 ms</td>
</tr>
</tbody>
</table>

- We conclude that the SecPAL engine is efficient -- over 75% of the overall cost is incurred in information-gathering prior to the SecPAL engine invocation.
Evaluation (Quantitative – 3/3): “In the Large”

<table>
<thead>
<tr>
<th></th>
<th>Gridmap</th>
<th>SecPALT</th>
<th>% diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>10M</td>
<td>2.505 sec</td>
<td>2.61 sec</td>
<td>4.2%</td>
</tr>
<tr>
<td>100M</td>
<td>10.73 sec</td>
<td>10.84 sec</td>
<td>1.0%</td>
</tr>
</tbody>
</table>

- We believe this difference is unlikely to impact most clients.
Summary

• Specific use cases in Campus Grid ➔ requirements from first principles

• Application of SecPAL to meet challenges has shown to be successful
  • Qualitative
  • Quantitative

• Next steps
  • Building the infrastructure, centered with the SecPAL engine
  • Extend to science collaborations more directly…
SecPAL for Collaborative E-Science

Welcome to the Web Page for the FLUXNET-La Thuile Data set.

This new database consists of over 900 site-years of data from over 170 eddy covariance measurement sites. Major efforts were made over the past year to harmonize, standardize, and gap-fill the growing 30-minute data records submitted by participants. The database also includes value added products like gross primary productivity, ecosystem respiration, climate and site characteristic information. This database updates and supercedes the Marcini FLUXNET database that was produced in 2000 (with 97 years of data from 38 European and North American sites).

This database was produced in association with the FLUXNET-TCO synthesis workshop held in La Thuile, Italy, from February 18 to 22, 2007. The aim of the workshop was to produce a new and extensive FLUXNET dataset, discuss its scientific potential, propose new synthesis activities and start first analyses. The workshop was attended by 60 scientists, who represented a cross section of younger and established scientists and were members of regional networks from around the world: CarbonEurope, Ameriflux, FLUXnet-Canada, LBA, Asaflux, Chinflux, USCC, Czflux, CarboEurope, Karflux, NECC, TCCS-Siberia and AfriFlux.

A policy on data distribution to the wider scientific community was discussed and established during the Workshop. It was agreed that this dataset will be accessible for one year (until August 2008).
Research Availability

• **Public Availability**
  • SecPAL Formal Model and Whitepaper
  • SecPAL Preview Binaries and Developer Documentation
  • SecPAL Schema and Schema Specification
  • [http://research.microsoft.com/projects/secpal](http://research.microsoft.com/projects/secpal)

• **Community Workspace**
  • Forums for questions and feedback
  • Additional downloads
    • Parser for simplified English grammar (F#)
    • Query editor available
  • [http://www.codeplex.com/secpal](http://www.codeplex.com/secpal)