A Flexible Component based Access Control Architecture for OPeNDAP Services

Philip Kershaw
STFC Rutherford Appleton Laboratory
Focus on two areas alluded to in the abstract:

1) “Without ready means to restrict access to data for such services, data providers and data owners are constrained from making their data more widely available.”
   - Paradox: provision of access control can open access?!
   - Illustrated by the BADC and OPeNDAP
   - But in the wider context of changing attitudes: data.gov.uk

2) “The range of different security technologies available can make interoperability between services and user client tools a challenge.”
   - Vision of seamless access can seem remote:
   - Services too complicated, get in the way or simply broken
   - A need for flexible approach to enable support for different technologies
Inception

- Seeds planted with discussion started at this conference last year
- Solution based on the use of HTTP redirects
- OPeNDAP implementations have existing support for this*:
  1. Initiate request,
  2. server side security challenge
  3. Redirect to authentication endpoint
  4. Redirect authenticated client back to original requested URI
- NERC DataGrid Security and Python WSGI middleware
  - Security filter components arranged to give the desired access control configuration

PyDAP: http://pydap.org/client.html
Alternative Approaches to Authentication

• Why use redirects at all? Host service over HTTPS:
  – but then performance limitation for large datasets,
  – changes from a well known HTTP address to HTTPS – user confusion?
• HTTP Digest
  – Not secure enough
• What about SOAP?
  – An invasive approach which would change the interfaces breaking existing client tools
  – Unsuit for large dataset transfers
**CMIP5** is a framework for co-ordinated climate change experiments

- Will input into the IPCC 5th Assessment Report (AR5) scheduled for 2013

- **Software infrastructure under development:**
  - 20 modelling centres
  - 50 numerical experiments
  - 86 simulations (total ensemble members) within experiments
  - 6500 years of simulation

  Data to be available from “core-nodes” and “modelling-nodes” in a global federation.

  Users need to find & download datasets, and discriminate between models, and between simulation characteristics.

  **Simulations Starting**
  - mid-2009

  **Model and Simulation Documentation needed**
  - end of 2010

  **Data available**
  - early to mid 2012

  **Scientific Analysis, Paper Submission and Review**
  - early 2013

  **Reports**
  - early 2013
Earth System Grid Security: Single sign on

OpenID

- Users have an identity URI:
  - Identifies them and
  - an OP (OpenID Provider), a service where they can be authenticated
- An OpenID Relying Party trusts the authentication assertion of a given OP
- SSL for mutual authenticate and enable RPs to whitelist trusted OPs

MyProxy
Credential Management Service

- OpenID less suited to non-browser based clients
- PKI based authentication - Grid based applications, OPeNDAP clients? ...
Filter Architecture

- OPeNDAP service is protected by an Authentication filter
- This redirects unauthenticated requests to an authentication service
- Authentication Service receives the client request
  - SSL client certificate based or
  - Default to OpenID based sign in
Filters in Action – ncopen client

Philip Kershaw, EGU 2010
More About Clients

• A server-side solution but what about the client side?
  – Browser
  – WGet and Curl – replacement to legacy ftp
  – NetCDF nopen - patch

• Demo:
  – Browser based access
  – Command line access with wget