Portals and Authentication

Issues and Solution Directions
from a CA and IGTF Perspective

David Groep
NIKHEF

Issues

- **Authentication**
  - a federated CA structure
  - Identity vetting and ‘classic’ AP requirements
  - Relying party requirements
  - Certificate ‘classes’

- **Linking Authentication and Portals**
  - automated clients
  - user credential caches
  - AAI-backed Short-Lived Credential Service CAs

Authentication model

Design and implementation choices made in production-oriented grids:

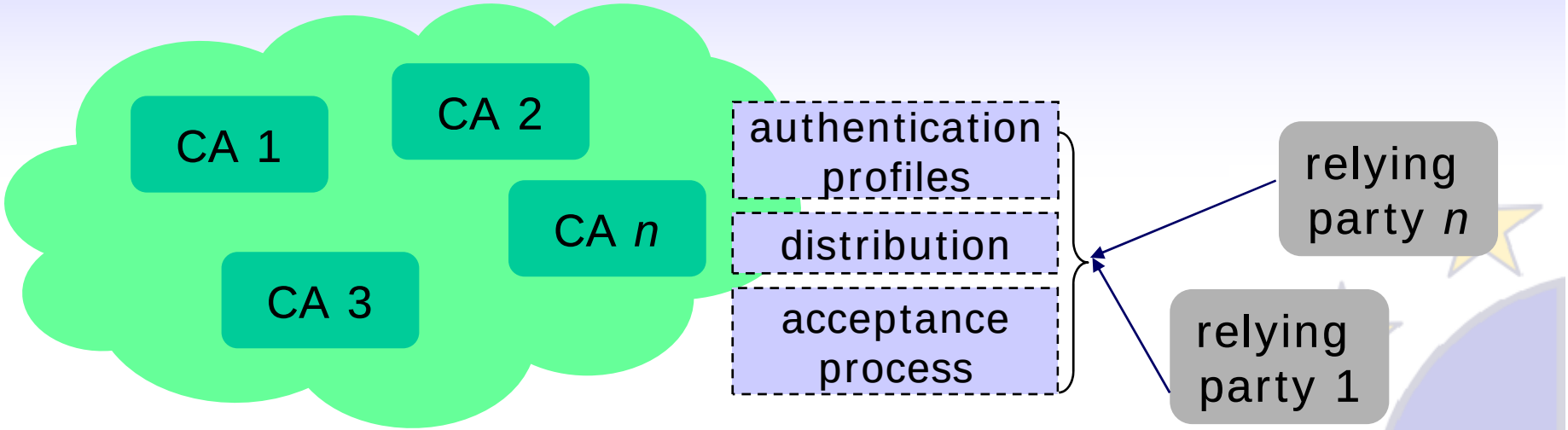
focus on providing *cross-national* trust
(initially in the context of the EU DataGrid and CrossGrid projects)

- National PKI
  - in general uptake of 1999/93/EC and e-Identification is slow
  - where available a national PKI could be leveraged, but not happened yet
- Various commercial providers
  - Main commercial drive: secure web servers based on PKI
  - Entrust, Global Sign, Thawte, Verisign, SwissSign, …
  - primary market is *server* authentication, not end-user identities
  - but use of commercial CAs solves the 'pop-up' problem
    ... so for (web) servers a pop-up free service is needed (i.e. SCS)
- on the other end of the spectrum: 'grass-roots' CAs
  - usually project specific, and without any documented policies
  - unsuitable for the 'production' infrastructure

The first grid authentication infrastructures

- Grid (academic) PKIs
  - started off with pre-existing CAs, and some new ones, late 2000
  - ‘reasonable’ assurance level based on ‘acceptable’ procedures
  - a single assurance level inspired by grid-relying party** requirements
  - using a threshold model: *minimum requirements*
- Grid CA coordination driven by actual and current needs
  - separation of AuthN and AuthZ allowed progress
  - published policies convince resource providers to ‘trust’ CAs
  - started with 6 authorities (NL, CZ, FR, UK, IT, CERN)
  - *a fundamentally federated (i.e. distributed) effort*

Federation Model for Grid Authentication

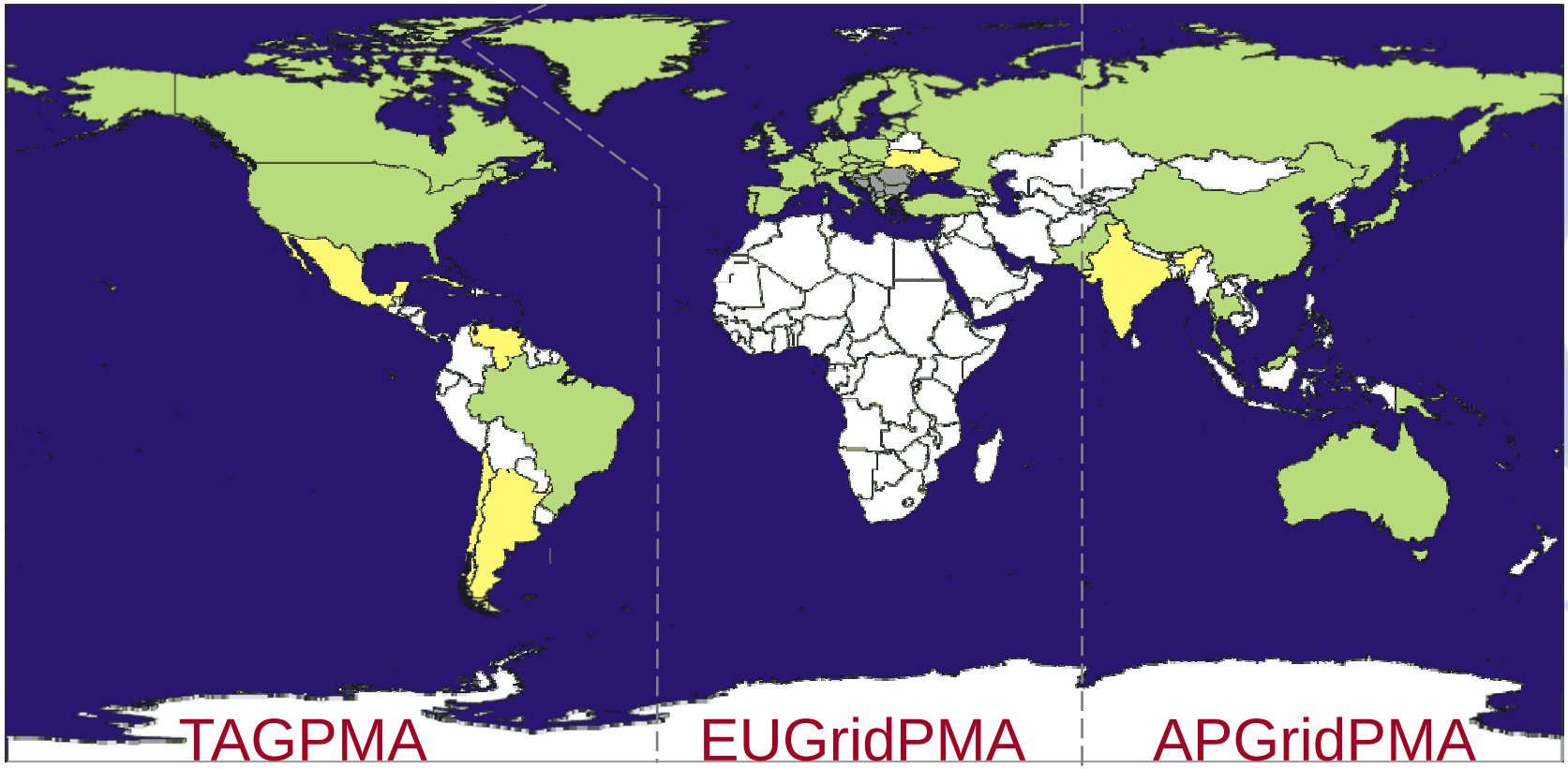


- A Federation of many independent CAs
  - common **minimum requirements** (in various flavours)
  - trust domain as required by users and relying parties *where relying party is (an assembly of) resource providers*
  - defined and peer-reviewed acceptance process
- No strict hierarchy with a single top
  - spread of reliability, and failure containment (resilience)
  - maximum leverage of national efforts and complementarities

International Grid Trust Federation

IGTF
International Grid Trust Federation
AP | EU | TAG

Federation of 3 Regional "PMAs", that define common guidelines and accredit credential-issuing authorities



Grid Relying Parties & resource providers

- In Europe
  - Enabling Grid for E-sciencE (EGEE) (~ 200 sites)
  - Distr. Eur. Infrastructure for Supercomputer Apps (DEISA) (~15 sites)
  - South Eastern Europe: SEE-GRID (10 countries)
  - *many national projects (NL BIG-GRID, VL-e, UK e-Science, Grid.IT, …)*
- In the Americas
  - EELA: E-infrastructure Europe and Latin America (24 partners)
  - WestGrid (6 sites), GridCanada, …
  - Open Science Grid (OSG) (~ 60 sites)
  - TeraGrid (~ 10 sites + many users)
- In the Asia-Pacific
  - AP Grid (~10 countries and regions participating, and growing)
  - Pacific Rim Applications and Grid Middleware Assembly (~15 sites)

data as per mid 2006

Relying Party issues to be addressed

Common Relying Party requests on the Authorities

1. standard accreditation profiles sufficient to assure **approximate parity**

   effectively, a single level of assurance sufficed then for relying parties – is changing today, as more diverse resources are being incorporated

2. monitor [] signing namespaces for **name overlaps**
3. a **forum** [to] participate and raise issues
4. [operation of] a **secure collection point** for information about CAs which you accredit
5. **common practices** where possible

list courtesy of the Open Science Grid

Pending requirements

The CP/CPS MUST describe

- How the identity (DN) assigned in the certificate is unique within the namespace of the issuing CA
- How the identity (DN) assigned in the certificate will never be re-issued to another end entity during the lifetime of the CA
- How the CA attests to the validity of the identity

In order for a (RA) to validate the identity of a person, the subject

- SHOULD contact the RA face-to-face and
- present valid government or employer issued photo-id and/or official documents.

If face-to-face is not possible then the CP/CPS MUST describe:

- How the CA provides accountability, showing that they have verified enough identity information to get back to the physical person any time during the lifetime of the certificate.

Building the federation

- Trust providers ('CAs') and relying parties ('sites') together shape the common requirements
  - Several *profiles* for different identity management models
  - Authorities demonstrate compliance with profile guidelines
  - Peer-review process within the federation to (re-) evaluate members on entry & periodically
  - reduces effort on the relying parties
    - single document to review and assess for all CAs under a profile
  - reduces cost for the authorities
    - but participation does come at a cost of involved participation …
- Ultimate trust decision *always* remains with the RP
- An authority is not necessarily limited to just 'grid' use

Guidelines: secured X.509 CAs

Aimed at long-lived identity assertions, the 'traditional PKI' world

- Identity vetting procedures
  - Based on (national) photo ID's
  - Face-to-face verification of applicants via a network of distributed Registration Authorities
  - Periodic renewal (once every year)
  - revocation and CRL issuing required *and we have all RPs actually downloading the CRLs several times a day*
  - subject naming must be a reasonable representation of the entity name
- Secure operation
  - off-line signing key *or* HSM-backed on-line secured systems
- Audit requirements
  - data retention and audit trail requirements, traceability of certified entities
- Technical implementation
  - need to limit the number of issuing authorities for technical reasons *(most software and browsers cannot support $\mathcal{O}(1000)$ issuers)*
  - certificate profile and interoperability

Short-lived or member integrated services

Aimed at short-lived ‘translations’, that are organisation/federation bound

- Identity vetting procedures
  - based on an existing ID Management system of sufficient quality
  - Original identity vetting must be of sufficient quality to trace the individual for as long as name is in active use
  - If *documented* traceability is lost, the subject name can never be re-used
  - revocation and CRL issuing not required for assertion lifetimes << 1 Ms
  - subject naming must be a reasonable representation of the entity name
- Secure operation
  - HSM-backed on-line secured systems
- Audit requirements
  - data retention and audit trail requirements, traceability of certified entities
- Technical implementation
  - scaling of this model still needs to be demonstrated, and needs higher-level coordination
  - *most software and browsers cannot support $\mathcal{O}(1000)$ issuers*
  - *and a peer-review based trust fabric cannot do that either …*
  - certificate profile and interoperability

ID management system requirements

Technical and IT security requirements

The identity management (IdM) system containing the identity information used to issue the assertions must meet the following conditions

- Re-usable private information used to authenticate end-entities to the IdM system must only ever be sent encrypted over the network when authenticating to any system (including any non-CA systems) that are allowed to use the IdM for authentication.
- A *not-published second authentication factor* must be used to authenticate the end-entity for certificate issuance
- The end-entities must be notified of any certificate issuance, using contact information previously registered in the IdM (for example by electronic mail)
- From the information stored in the IdM it must be possible to determine if the requestor's identity has originally been validated using a face-to-face meeting as described above

ID management system requirements

Identity vetting requirements
convincing the world that you're OK

Documentation of how the IdM is populated, maintained and cleaned MUST be documented and agreed to by the PMA. Two modes

By example:

The IdM used by the CA should be a system that is *also used to protect access to critical resources*, e.g. payroll systems, for use in financial transactions, granting access to highly-valuable resources, and be regularly maintained. ⬅ *tries to 'catch' the quality of the system without having to report to formal audits*

By review:

Alternatively, equivalent security mechanisms must be provided, described in detail and presented to the PMA and *are subject to PMA agreement*.

and again the data for those entities in the IdM that qualify for 'MICS' assertions must be of a quality that allows unique tracing, name uniqueness and persistency – and a mechanism to clean 'stale' entries must be defined.

Example: the UvAmsterdam does not trust its own system even for grading!

Profile matrix: towards multiple LoAs

- All grid technical security mechanisms meet the technical protocol requirements of level 3 (but even soft tokens meet level 3 …)

- Identity vetting requirements for Classic and MICS meet ~ level '2 –'
  - only in-person allowed
    (remote option is not allowed, Authorities cannot check financial records &c)
  - except that address and DoB are not necessarily retained by the RA
    to ease data protection issues, and copies not always retained
  - but the ID number (and issuing country) is recorded,
    so 'relevant' agencies can get to the applicant
  - *VOs need to collect this information and more anyway for incident response*

- Both more stringent and looser LoAs needed for other resource classes
  - but *e-Auth* level 1 is too low, and NIST doesn't define anything in between…

Users vs Hosts

Profiles distinguish between 2-3 ‘classes’

- Users
  - high-quality identity vetting, so that the same subject name is quite surely bound to the person
  - ‘all’ CAs under the classic profile meet this bar
- Hosts (or ‘service’, e.g. ‘`CN=gatekeeper/ce.example.org`’)
  - the concept of ‘ownership’ of the (DNS) name is vague
  - can be a group of system admins, where the local RA will ensure (‘somehow’, ‘vaguely’) that the requestor is authorized
  - for some CAs, ‘service’ certificates can be requested by ‘service owners’, and no thorough checking is done with the system administrators
  - assurance level for host and service certs is really bound to the use of the DNS name only
  - when used outside securing TLS network-endpoint, the assurance level is ill-defined and varies widely across the IGTF

Hosts vs Robots

If hosts/service assurance level is so ill-defined, what then?

- Raise the assurance level
  - leads to intricate problems when used for the current purpose of securing network endpoints
- Look at the ‘automated client’ class
  - identities for programs and services that act in an aotmated way towards the grid infrastructure
  - concept introduced by Mike Helm in 2002
  - criteria developed by Jens (see next talk)
  - not yet supported by all CAs, but interest is growing (actually, today only UK and NL do, with CZ coming up)

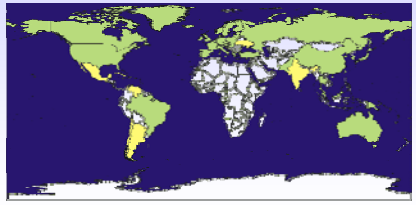
Profile matrix: where we stand

Multiple Authentication Profiles: where the IGTF stands today

Identity vetting	With govt photo-ID Only by in-person F2F meeting of RA	With govt photo-ID With proven documented traceability to individual at any time (no definite F2F requirement)	…
Subject: soft-tokens allowed Issuer: off-line or online HSM 140.2-3	**Classic AP** ***near-inline Id vetting***		
Subject: soft-tokens allowed Issuer: online HSM 140.2-3	**'MICS' (proposed)** ***time-shifted Id vetting***	**SLCS** ***time-shifted Id vetting***	
…			

note: certificate classes are orthogonal to the Profiles

Evolution and revolution



- Each CA is independent
  - constraints of manpower, local funding, national legislation &c
  - compliance is to minimum requirements

- Introduction of new features
  - through demand from within the subscriber base (per CA)
    most effective, especially if you bring along effort
  - through cross-fertilisation by peer CAs
    also effective, but can take a lot of time if effort is lacking
  - by raising the minimum requirements
    does not work well for this kind of innovation …

DutchGrid CA Policy v3

- Need for automated clients
  - from the bioinformatics domain (NBIC BioRange/BioAssist)
  - other BIG GRID application domains (e.g. astronomy)

- Supported classes of certificates (within the Classic X.509 secured profile)
  - Users: certificates for natural persons
  - Hosts: networked systems, applications or services – solely to *identify network endpoints in communications*
  - Servers: (internal)
  - **Robots: agents that perform automated functions** *protected in a secure hardware token ~ FIPS140-2 level 2*

Possible alternatives

Current authentication landscape

- **Service certs**
  - the CA *may* allow its use as an automated client
  - but the infrastructures should be wary of accepting them!
  - check of the policy may be needed
    - i.e. in NL, the ‘hosts’ class identifies network endpoints, as the verification is limited to finding the appropriate system admin; in DoEGrids they are quite weakly linked
- **User certs**
  - generate a proxy from the personal proxy of the portal owner
  - needs the owner to regularly provide the passphrase
  - but works in virtually all scenarios
- **Robots certs (see Jens’ talk)**
  - where available (UK, NL, *soon CZ*) these are the preferred choice
  - protects private key from abuse outside the portal system

and, of course, these options can be mixed
downside: requires new Grid AUP/Policies (but no new CA requirements)

Possible alternatives

Traditional Portal approach

- **Use the MyProxy solution**
  - all jobs are traceable to the requesting user
  - portal MyProxy server becomes a valuable target
  - entirely within the current policy space
  - downside: ‘real’ users cannot handle any kind of credentials

In a pervasive AAI environment (or in wonderland?)

- **Federation backed SLCS integrated with the portal**
  - SWITCHaai-like solutions
  - excellent for those countries that have a working AAI *that actually reaches all your researchers* (i.e. CH)
  - Authorize to portal based on AAI account, then generate a cert on the fly form the SLCS service
  - also entirely within current policy space
  - not too many countries have something pervasive …