LDAP gateway with partial caching (LDAP-PC)

Jiří Sitera

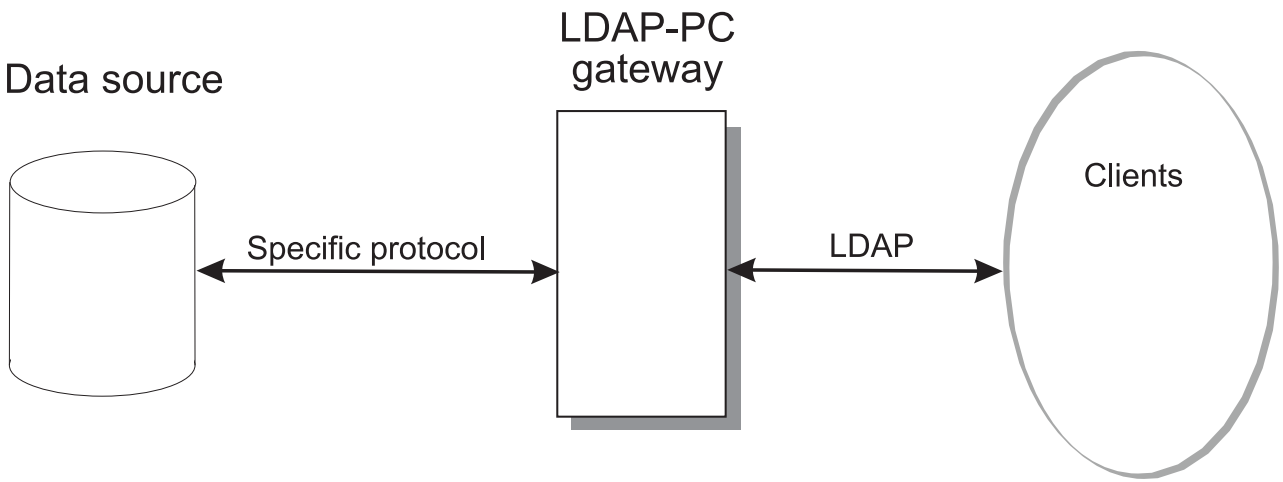
Contact: sitera@civ.zcu.cz

CESNET, Prague, Czech Republic, and
University of West Bohemia, Pilsen, Czech Republic

CGW'06, Cracow, 17.10.2006

Talk outline

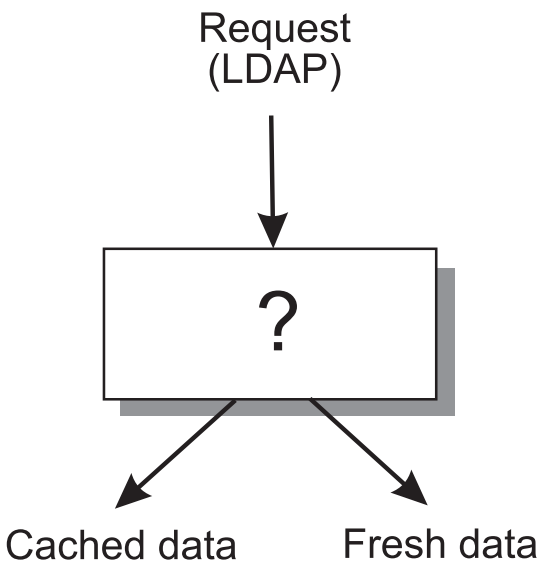
LDAP-PC quick overview



LDAP-PC concepts

- A gateway from source into LDAP protocol.
  - Source = API, protocol, data model, etc.
  - There are many different sources behind one gateway.
- Concept of partial caching.

Partial caching – basic idea



- Two scenarios of data handling.
  - Fresh data = request transformed to on-line query to the primary data source.
  - Cached data = request answered using data stored in the cache.
- Behaviour derived from query scope.
  - Simple query $\rightarrow$ fresh data.
  - Complex query $\rightarrow$ cached data.

- Refresh of all data in regullar manner.
  - All cached data are from one time slice / consistent.
  - Cached data maximal age is guaranteed (refresh period).

Motivation

- The situation
  - There are information sources with valuable data available.
  - There are clients able to profit from the data.
  - Different API, protocols, data representation, security models.
  - Clients lacks the ability to connect to appropriate data sources.
- The context
  - Demand for lightweight solution.
  - Information infrastructures are too complex, but the ideas behind are useful (unification, one way to different sources).
  - Previous work in the area of publication of data via LDAP.

Why LDAP

- LDAP is really lightweight and widely available/used.
- LDAP offers many useful features.
  - Data and naming model.
  - Authentication options, authorization model.
- Unification and junction point
  - LDAP gateway is the point where data transformation can be done
  - Transformation of syntax, semantics, cooking of raw data from different sources.

Why partial caching

- Many other caching strategies considered.
- Natural, implicit strategy.
  - Simple query
    - Some latency tolarable.
    - Fresh data expected.
  - Complex query
    - No real need for fresh data.
    - Any speedup helpful.
    - Usually there is demand for consistent view.
- No hints from client expected.
  - LDAP “query scope” parameter used.

LDAP-PC example – filesystem quota & usage reports

- Very simple and valuable data.
- Different sources, not easy to access.
  - AFS – clients not very common.
  - Local filesystem/NFS – many different tools for quota management and reporting.
- Typical usage fits the LDAP-PC strategy:
  - A particular user quota/usage query – fresh value.
  - Summary queries, for example server's admin report – consistency and speed required, data age tolerable.
- Adds authorization layer.
- Choice of autentization methods valuable.
- Speedup of summary queries using prefetched data needed anyway (AFS API too slow).

The place of LDAP-PC

- *"Keep things as simple as possible."*
  - Current GRIDs can in most cases live with much simpler information infrastructure that they try to deploy.
  - LDAP-PC as a start-up booster.
    - Deployment investment is low, future replacement doesn't hurt.
    - Can start with (and build on) small, isolated islands.
  - LDAP-PC can be used for prototyping, integration of legacy systems, etc.
- *"Keep things as simple as possible, but not simpler."*
  - LDAP-PC is not cure, only another particular solution available.
  - Covering idea (published as C-GMA)
    - Coexistence of many solutions.
    - (meta)infrastructure to provide way how to cooperate.
    - On demand selection of the most fitting solution for each purpose.

What we have

- Proof-of-concept implementation.
- Pilot deployment experience.
  - Kerberos – account states.
  - Filesystem quota/usage.
- Plans for future.
  - Modular implementation
    - Goal: Provide SW package simple to use and customize.
  - Experiments with support of write operations.
  - Deployment in MetaCenter environment.
    - Provide data for monitoring and user support tools (portal).

Summary LDAP-PC

- Lightweight but powerful tool, easy to use and to deploy.
- Ideas behind: LDAP (protocol, data model, security), partial caching.
- Purpose: Similar to information infrastructures (unification, additional integrating layers around data sources).
- Application area: everywhere the lightweight nature is an advantage.

References

- CESNET technical report (in preparation).

CESNET

ZÁPADOČESKÁ UNIVERZITA V PLZNI

Thank you for your attention.