Telstra

ViewFinder
X.500 Electronic Directory System

ViewFinder™

X.500 Case Study — Telstra Corp. Directory

November 1994

Summary

Although the X.500 standards have been widely implemented, and many groups have experimented with X.500 services, few organizations have experience with the adoption of X.500 as the basis of corporate electronic directory services within their company. Telecom Australia committed to adoption of X.500 internally in 1989 and began the process in the same year. Three years on, the Telecom Corporate Directory is a mainstream information resource within the company, used daily by thousands of employees. This talk discusses Telecom's experience, and in particular points to the steps that other organizations must take to successfully move to X.500.

Contents

Corporate X.500 Electronic Directories

Technologies, Applications and Services

Hardware and Software

DSA System Architecture

DUAs and Networks

Naming Tree Design

Attributes and Object Classes

Data Strategy

User Strategy

Administration

Corporate Interfaces

Conclusion

X.500 Case Study — Telecom Corporate Directory

R. Exner

Presented at Standards Australia Conference on Naming and Addressing in the OSI Environment 1992

Corporate X.500 Electronic Directories

Corporate electronic directories are systems that allow employees of a company access to on-line directory information about other employees. At its simplest the information is merely the telephone number of the employee and the users of such systems are mainly switchboard operators.

But with the advent of X.500 and a growth in user expectations, a corporate directory can provide on-line access to all the information about an organization's units and people that others in the company may require. This includes the names, location, managers and staff of any unit of the organization, the job titles, functions, electronic mail addresses of its employees, the responsible parties for all official functions in the organization etc.

Corporate electronic directories can also serve as a means of automatically publishing traditional printed directories, or other lists for management purposes, and as a means of controlling and monitoring communications and infrastructure costs.

In recent years, Telecom (AOTC) has introduced such a corporate electronic directory (CED) into its operations. This directory, built on X.500 technology and developed within the Telecom Australia Research Laboratories, has since become a mainstream corporate information resource, used daily by thousands of employees. Much experience was gained in the course of developing and introducing this system. This talk will share some of that.

Technologies, Applications and Services

The basis of Telecom's corporate electronic directory is an X.500 directory engine developed in the Telecom Australia Research Laboratories. However, this system provides only the basic *technology* for X.500 corporate directories.

The directory engine needs to be integrated into a complete corporate directory *application*, complete with suitable user interfaces, networks, administration tools, user and administration documentation etc. before it can be used for its specific purpose.

Finally, for a corporate directory *service* to be successfully introduced into the organization requires attention to a myriad of additional factors, including initial data acquisition, data maintenance strategies, user and administrator training courses, strategies for cost recovery, consideration of printed directories, and interfacing of the corporate directory application and its data with other information processing systems in the organization.

By way of analogy, the basic X.500 technology is equivalent to the engine of a motor car, and the corporate directory application to the whole

automobile. To actually drive the car requires driver training, roads, garages, an entire infrastructure. The engine alone is not a transport service.

Hardware and Software

Telecom's X.500 directory engine was developed as a high performance system based on a custom database and search processor optimized for X.500. It features very fast search times, extensive approximate matching (phonetic matching, keyword matching etc.) and a friendly user interface. It runs under Unix System V Release 2 or later and has been successfully ported to half a dozen hardware platforms.

The current hardware platform is a Sun 670 multiprocessor with 128 MB of main memory (sufficient to hold the whole database in cache memory). It is expected to adequately support several hundred simultaneous users, with sub-second response times for most searches.

DSA System Architecture

The Telecom CED is architected as a single centralized Directory System Agent (DSA), together with many co-resident or remote Directory User Agents (DUAs), accessed over a network.

X.500 makes it possible to implement a single conceptual directory as a single DSA or as a network of cooperating DSAs optionally sharing one another's data. The advantages of having a single DSA include simplicity of administration, faster response through the avoidance of chained operations that pass a query on to a further system, and a lower investment in hardware and networking; the disadvantages include potentially more long-distance traffic, higher performance demands at the DSA, and a lack of redundancy to cope with software or hardware failures.

Note that if local control and ownership of information is required, then a distributed architecture is the only choice; a centralized DSA is only possible if all information is owned by a central authority. Likewise a distributed architecture is the only choice where the transaction rate at a central machine would simply be too great for the planned hardware platform and network.

The slower response times of a distributed system can avoided by fully replicating all data — thus, every system has a copy of the data of every other system. This can result in increased network traffic (one of the reasons for adopting a distributed system) as updates are conveyed to every site, but provides maximum redundancy and is probably the best configuration. It has the further advantage that transaction rates at any site are reduced. However, X.500 standard protocols for data replication are available only with the 1992 version of the standards, and implementations do not yet exist. In the future, it is likely to be the most widely adopted architecture, but for the present, it requires adoption of non-standard technology. Although Telecom's system is a centralized one, it is likely to move to a fully-replicated distributed architecture in the future as the replication protocol of 1992 X.500 is implemented.

DUAs and Networks

Telecom's system supports co-resident and remote DUAs. The co-resident DUAs implement a VT100 interface, and are accessed over a network most typically by personal

computers emulating VT100 terminals. The co-resident DUAs can run on the same host as the DSA, or can run on a front-end processor co-located with the DSA's host. Remote DUAs running as applications on the users' PCs are planned as soon as suitable implementations can be provided (in the forthcoming months).

X.500 is an Open Systems Interconnection (OSI) application layer protocol, and the X.500 standard calls for conforming implementations to use an OSI stack for the protocols below the application layer. However, the acceptance of OSI in the marketplace — i.e. in the products readily and economically available for PCs and desktop computers — is still somewhat patchy. The DUAs and DSA therefore implement an architecture that permits OSI or non-OSI (i.e. TCP/IP) protocols to be used below the application layer, according to whether an OSI Association / Presentation Layer Interface (APLI) is available on the host.

Communications between PCs and the host are supported over X.25 and Telecom's TCP/IP backbone Corporate Data Network. Most users do not have TCP/IP access, but are on a local area network (LAN) which offers X.25 access, so this is the predominant access network at present.

Naming Tree Design

X.500 provides a great deal of flexibility to someone setting up a directory. Not only is there great freedom to choose the attributes and object classes that one wishes to have on-line, but there are several possible choices of naming hierarchy. X.500 requires entries in its database to be arranged in a hierarchy, and the distinguished name of any entry is the path through the hierarchy, beginning with the root at the top of the naming tree.

The first arcs in the path are dictated by CCITT and ISO, and national regulations, and were the subject of the preceding talk. But names below the name of the organization are entirely for the organization to choose. National organizations that consist of a collection of loosely coupled separate businesses actually have an additional choice to make: whether to represent their organization as one organization with sub-units, or as a collection of organizations. In Telecom's case, it chose to represent itself as one organization, with fully-owned subsidiaries appearing as sub-units.

For most organizations, the choices are as follows:

1. The naming tree can reflect the natural organizational hierarchy of the organization. Thus, intermediate entries in the naming tree represent the business units, divisions, sub-divisions, departments, branches, and sections of the organization. The advantage of this approach is that organizational unit information forms a part of the naming information of every entry, and queries about the unit to which an entry belongs, or its manager, are easy to ask and answer. The disadvantage is that names can be very long and unwieldy, and change whenever the company reorganizes itself or the individual moves to another unit.
2. The naming tree can reflect the location of the organization's premises and people. Thus, intermediate entries in the naming tree represent site names. This approach is natural if the company deemphasizes its organizational structure, and it is an individual's site rather than section that is the better known. It has the advantage that names are more stable

(provided people remain in jobs on the same site) and shorter. But it makes it very difficult to extract hierarchy information if this is needed.

3. The naming tree can be mostly flat, showing neither organizational units nor sites. Such a directory is likely to be simplest to construct from available data, and names are simplest and most stable, but it provides least information, requires artificial methods of disambiguating users with the same name (Telecom has 18 David Smith's nationwide), and fits the X.500 model most poorly causing poor performance in some implementations.

In Telecom's case, a high emphasis was placed on providing organizational hierarchy information, and method 1 was adopted. It has proven very satisfactory. The CED hides the resultant long distinguished names, and the changing of names with organizational changes has not proven to be a problem.

Attributes and Object Classes

The Telecom system implements a mixture of standard X.500 attributes and custom (Telecom-specific) attributes. The standard attributes include Common Name, Surname, Telephone, Postal Address etc. The custom attributes include Employee Number, Org Unit Code, Personal Title, Given Name, and Print Code. The last of these is used to control the automatic production of printed directories from on-line data.

Data Strategy

Perhaps the most difficult step in setting up a corporate electronic directory is filling it with data. (The next most difficult step is keeping the data up to date). Like most organizations, Telecom had pre-existing directory information in electronically-readable form. Indeed, there were many different directories: most describing a division or department; some built on databases, others as word-processor readable files; some detailing senior managers only, others describing all employees; some quite current, others badly out of date.

Moreover, some of the information deemed desirable in a corporate directory did not exist in these directories at all; it was held in payroll or personnel databases, in PABX databases, or simply did not exist in electronic form at all. An example of the latter is the names of organizational units. The official record showed around 2000 organizational units, but for every officially recorded organizational unit, there were another three designated and staffed units which operated as units but which were unknown to the official record. The current X.500 directory lists all 8000 units and has become the definitive source of organizational unit names and relationships.

Telecom took a three-pronged approach to initially loading its X.500 directory:

1. *Automatic conversion of pre-existing directory* data. Typically this meant dumping the data as a text file, developing a small program or Unix® script to massage the data into a bulk-loadable format, and loading it into the CED. The technique was only practical for large data sets (usually over 1000 entries) and resulted in data of medium quality. Many attributes were missing.

2. *Bulk loading of data*. In some cases it was most practical to have a typist prepare a text file in a bulk-loadable format containing directory data

gathered from various printed lists. This file would then be batch loaded into the CED.

3. *Interactive updating.* All other entries were added interactively. This results in data of highest quality, but is most labour-intensive.

Data is kept up to date mostly by interactive updates (see below for the human strategy for making this successful). In addition to this, there is a monthly consolidation of entries with the payroll database, on the basis that this database is the first that gets updated when someone joins or leaves the organization!

User Strategy

Successful adoption of a corporate-wide X.500 system requires considerable attention to be given to user issues. These include: the rate of introduction of the system, training, security, how the data is to be kept up to date, requirements for printed directories, and other factors.

There are several categories of user: casual users for whom the system must above all be trivially easy to use, full-time users such as PABX operators for whom the system must be fast, update users who are responsible for keeping information up to date, and users administering the system. Each of these categories has special needs, and feedback from representative users in each category is necessary to sort out difficulties, e.g. with training or connection-setup delays.

The corporate directory data is kept up to date by a large group of update users (several hundred) who have responsibility for keeping a designated portion of the database up to date — most usually the branch or division to which the updater belongs. Compliance with their responsibilities varies from updater to updater, but twice a year notification is given of the intent to take a snapshot of the data for production of the printed directory, and this event always brings a flurry of update activity as updaters scramble to catch up.

Telecom has a small core group of staff who oversee the system, including the arrangement of training courses, the authorization of new users, operation and maintenance of the hardware, preparation of the system for the production of printed directories, etc.

Administration

The administration functions that must be performed on a large corporate directory system include system configuration, backup and recovery, system performance and usage monitoring, managing changes to the system schema, and preparation of printed directories. Good tools to manage all these aspects of system operation are essential.

Corporate Interfaces

A corporate electronic directory is a major information resource with value extending well beyond its immediate applications. To realize this value requires integration of the system into other corporate information systems.

Typically the data flow is two way: Some attributes are updated in other systems and can be automatically provided to the directory; e.g. personnel information and electronic mail addresses. Others are mastered in the directory and need to be provided to external systems; e.g. postal addresses. And for some there is a two-way flow with the directory serving as a

publishing mechanism to make information coming from external sources available to all.
Two systems to which interfaces are very useful are personnel or payroll systems, and PABX systems.

Conclusion

To build a successful X.500 corporate directory requires more than a good base technology, and careful attention must be given to the user and data strategies and other administrative considerations.
Although implementation for Telecom has been somewhat protracted, Telecom's experience with X.500 corporate electronic directories has been altogether positive, and Telecom now reaps significant cost savings and the benefits of greatly improved corporate communications from its efforts in this area.

The Author

Rolf Exner is a Principal Engineer at the Telecom Australia Research Laboratories in Melbourne, Australia, where he leads a project team concerned with directory standards development and implementation. He is currently the CCITT Special Rapporteur for Directories and chairs the CCITT committee responsible for the current extensions work in X.500; he has been an active member of that committee since work began on X.500 in 1985.