INTERNET-DRAFT
Intended Category: Informational
Expires in six months

Kurt D. Zeilenga
OpenLDAP Foundation

3 February 2004

LDAP Multi-master Replication Considered Harmful

<draft-zeilenga-ldup-harmful-02.txt>

1. Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026.

Distribution of this memo is unlimited. Technical discussion of this document may take place on the IETF LDUP Working Group mailing list at <ietf-ldup@imc.org>. Please send editorial comments directly to the document editor at <Kurt@OpenLDAP.org>.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts. Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as ``work in progress.''

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/1id-abstracts.txt>. The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

Abstract

Over the last few years there has been significant development of Lightweight Directory Access Protocol (LDAP) replication mechanisms supporting multi-master service models. While multi-master replication may be useful in some situations, the deployment of multi-master replication alters the standard LDAP service model in a manner which can be harmful. Specifically, the atomicity, consistency, isolation, and durability (ACID) properties of the LDAP service model would be lost.

This memo discusses the LDAP service model, how multi-master replication alters the service model, and how this alteration is harmful to existing directory applications.

1. Introduction

The Lightweight Directory Access Protocol (LDAP) [RFC3377] is a protocol for accessing directory services which act in accordance with the X.500 [X.500] information and service models [X.501][X.511]. There has been significant consumer demand for "multi-master" replication of LDAP-based directory servers. However, there appears to be continued consumer confusion over data consistency issues introduced by the forms of multi-master replication being developed. Consumers tend to want "high availability", "scalability", "strong data consistency" and other qualities all at once. When engineering an information service, a balance between these qualities must be found which meets the design objectives.

The designers of X.500 and LDAP specified an information service which offers "high-availability" and "scalability" of read-access through shadowing (replication) to slave (read-only) servers and "strong data consistency" through a "single master" (authoritative) server.

The introduction of multi-master replication, as described in [RFC3384], to LDAP will significantly change the service model. In particular, as no one server is authoritative over an object, the protocol would not guarantee strong data consistency between its peers. That is, the directory service would no longer be capable of managing the concurrency of independent modifiers of directory content.

Changing the service model change will break applications which rely on current semantics and, hence, should not be made. Instead, a new directory access protocol should be developed to accommodate the desired semantics.

To understand why the introduction of multi-master replication to LDAP-enabled directories is harmful, one must first understand the X.500 information and service models as used in LDAP. These models are discussed in Section 2.

The introduction of multi-master replication would significantly alter these models. Section 3 discusses these alterations.

These alterations will break existing directory applications. A couple of examples of affected applications are provided in Section 4.

Security Considerations are discussed in Section 5.

Conclusions are discussed in Section 6.

2. X.500/LDAP Models

The X.500 information model [X.501] is hierarchical, object-oriented, and designed to distributed directory systems. The model also supports single-master replication [X.525]. LDAP is defined in terms of X.500 as an X.500 access protocol [RFC2251]. The X.500 service model [X.511] provides atomicity, consistency, isolation, and durability properties ([ACID]).

The X.500 service model requires atomicity (i.e. "all or nothing"). That is, either all the parts of the update operation are committed to the Directory or none are.

The X.500 service model requires consistency. That is, an successful update operation creates a new and valid directory state and a failed update operation leaves the directory unchanged.

The X.500 service model requires isolation. That is, no part of the update operation becomes visible to other operations until its been committed to the directory.

The X.500 service model assumes durability ("updates will not be lost"). That is, the X.500 assumes that updates committed to the Directory are held by the responsible directory system agents (DSAs or servers). However, the specification does not explicitly state a requirement that servers ensures correct state is maintained in the face of unexpected and/or unusual faults (like power outages).

It is noted that X.500 replication (shadowing) model allows for transient inconsistencies to exist between the master and shadow copies of directory information. As applications which update information operate upon the master copy, any inconsistencies in

shadow copies are not evident to these applications.

3. Multi-master Changes to LDAP Service Model

RFC 3384 defines multi-master replication as follows:

> Multi-Master Replication - A replication model where entries can be written and updated on any of several master replica copies without requiring communication with other master replicas before the write or update is performed.

For example, if two directory user agents (DUAs or clients) independently attempt to add different entries with the same name but against different masters, both operations could indicate a successful result despite the name conflict. Likewise, if two clients independently attempted to add the same attribute but with different values, both attempts could be successful despite the attribute value conflict if issued against different masters.

Depending particulars of the multi-master replication system, such conflicts are resolved either automatically or manually. Generally, automated reconciliation procedures are used which rely simply ignoring certain updates [LDUPURP]. These procedures can lead to reconciliation to a directory state not requested by the user.

Obviously, the introduction of multi-master significantly changes the X.500/LDAP service model. Atomicity is lost as the final state of the directory may not incorporate all portions of an update request. Consistency is lost because a successful update operation may not result new and valid directory state being created. Isolation is moot. No durability is provided as updates may be lost under normal operating conditions.

4. Harm to existing directory applications

All directory applications which are designed to support concurrent administration of user application information rely, to some degree, on the service model's ACID properties. The severity of the harm done to these applications will depend on a number of factors. In many cases, the harm is irreparable. This section offers a few simple examples intended to demonstrate the kind of harm that would can be inflicted. In many other cases, the harm done may be quite subtle but no less real.

4.1. Allocation of service entries

Many directory applications allocate unique service entries for users. For instance, white pages application may allow concurrent addition of users (using the naming plan for Internet directory applications [RFC2377] and inetOrgPerson schema [RFC2798]) and rely on the directory service to ensure that each DN uniquely identifies a user. One client interacting with master server might attempt to add an entry for Joe Smith called <uid=joe@example.com,dc=example,dc=com> and another client interacting with a second master server might attempt to add an entry for Joe Jones <uid=joe@example.com,dc=example,dc=com>. Both of these additions could be successful.

The introduction of multi-master replication would cause great harm to such deployments as it would allow both adds to succeed.

4.2. Allocation of serial numbers

Many directory applications require each object (in a particular class or set of classes) to have a unique serial number assigned to it. For instance, in Network Information Service [RFC2307] system, uidNumber associated with a user must be unique within an administrative domain.

One approach which allows multiple instances of the administrative client to allocate unique serial numbers, is to have an entry in the directory which holds the last assigned uidNumber. Then clients can read the uidNumber and attempt to increment it as follows:

```
dn: cn=Last UID,dc=example,dc=com
changetype: modify
modify: delete
delete: uidNumber
uidNumber: 1020
-
modify: add
add: uidNumber
uidNumber: 1021
-
```

where 1020 was the value uidNumber read and 1021 is the desired value. If the modify fails because the value to be deleted no longer exists, the client can repeat as necessary.

(Another approach is to use a modify/increment with atomic read entry features [X.511][Increment][ReadEntry].)

The introduction of multi-master replication would cause great harm to such applications, resulting in same serial number being assigned to different objects.

4.3. Allocation of single-valued authority information

Some applications rely on the value of a single valued attribute to indicate which service or process currently has authority over an object. For example, say the single valued attribute 'authority' is defined in the schema to represent the service or process which is currently responsible for administration of the object. If one client tries to add "authority: A" and another tries to add "authority: B" to an entry which presently has no authority attribute, both of these operations cannot be successful.

The introduction of multi-master replication would cause great harm to such applications, resulting in exclusive authority being granted to multiple services or processes.

4.4. Entry resurrection

Applications and administrator generally do not expect entries they delete to be resurrected. For example, if an administrator deletes a user entry, the administrator would likely be very surprised if it later found that user entry had been resurrected.

The introduction of multi-master replication can lead to such as a replication conflict, due to the addition of a child entry subordinate the user entry on another master, can result in the user entry being resurrected.

5. Security Considerations

It is unclear how one can build secure directory applications where update operations do not have the atomicity, consistency, isolation, and durability properties.

It is unclear how one can secure the directory when updates to authentication credentials and security and other policy information may be lost.

6. Conclusions

The X.500/LDAP information and service models does not support multi-master replication and cannot be altered to support multi-master replication without causing significant harm to existing directory applications. LDAP developers should heed this implementation absolute imperative [RFC 2251, Section 3.3]:

This document defines LDAP in terms of X.500 as an X.500 access mechanism. An LDAP server MUST act in accordance with the X.500(1993) series of ITU recommendations when providing the service. However, it is not required that an LDAP server make use of any X.500 protocols in providing this service, e.g. LDAP can be mapped onto any other directory system so long as the X.500 data and service model as used in LDAP is not violated in the LDAP interface.

7. Normative References

[RFC2251] Wahl, M., T. Howes and S. Kille, "Lightweight Directory Access Protocol (v3)", RFC 2251, December 1997.

[RFC3377] Hodges, J. and R. Morgan, "Lightweight Directory Access Protocol (v3): Technical Specification", RFC 3377, September 2002.

[X.500] International Telecommunication Union - Telecommunication Standardization Sector, "The Directory -- Overview of concepts, models and services," X.500(1993) (also ISO/IEC 9594-1:1994).

[X.501] International Telecommunication Union - Telecommunication Standardization Sector, "The Directory -- Models," X.501(1993) (also ISO/IEC 9594-2:1994).

[X.511] International Telecommunication Union - Telecommunication Standardization Sector, "The Directory: Abstract Service Definition", X.511(1993).

[X.525] International Telecommunication Union - Telecommunication Standardization Sector, "The Directory: Replication", X.525(1993).

[RFC3384] E. Stokes, et. al., "LDAPv3 Replication Requirements", RFC3384, October 2002.

8. Informative References

[ACID] Section 4 of ISO/IEC 10026-1:1992.

[RFC2307] Howard, L, "An Approach for Using LDAP as a Network Information Service", RFC 2307, March 1998.

[RFC2377] Grimstad, A., R. Huber, S. Sataluri, and M. Wahl,

"Naming Plan for Internet Directory-Enabled Applications", RFC 2377, September 1998.

[RFC2798] Smith, M., "The LDAP inetOrgPerson Object Class", RFC 2798, April 2000.

[Increment] Zeilenga, K., "LDAP Modify/Increment Extension", draft-zeilenga-ldap-modify-increment-xx.txt (to be submitted soon), a work in progress.

[READENTRY] Zeilenga, K., "LDAP Read Entry Controls", draft-zeilenga-ldap-readentry-xx.txt, a work in progress.

9. IANA Considerations

No IANA actions are requested.

10. Authors' Address

Kurt D. Zeilenga
OpenLDAP Foundation

Email: Kurt@OpenLDAP.org