

OpenLDAP Replication Strategies

Gavin Henry

Managing Director, Suretec Systems Ltd.

- ghenry@suretecsystems.com

Documentation Developer, OpenLDAP Project

- ghenry@openldap.org



Agenda

- Overview
- The History of Replication
- Replication Technology
- Deployment Alternatives
- Replication Best Practices
- Questions and Answers

Overview: A word about Suretec

- Founded 2003
- Part of Suretec Group
 - Suretec Systems – Consultancy and support
 - Suretec Telecom – VoIP and Telecom products/services
 - Suretec Training - Asterisk/OpenLDAP training
- Joined OpenLDAP Engineering team in 2007
- Joined Asterisk team in 2009
- Varied client base

Overview: OpenLDAP Project

- OpenLDAP is an open source code project
- Founded 1998
- Three core team members
- A dozen or so contributors (engineering team)
- Feature releases every 12-18 months
- Maintenance releases roughly monthly

Overview: What is LDAP?

- The Lightweight Directory Access Protocol is an application protocol for querying and modifying directory services running over TCP/IP.
- A directory is a set of objects with attributes organised in a logical and hierarchical manner.
- Used to locate organisations, individuals, and other resources such as files, hosts, application configuration (think Samba) and devices in a network environment.
- Basically a centralised hierarchical (tree) data store that uses standards based access methods.

Overview: Replication

- Replicated directories are a fundamental requirement for delivering a resilient enterprise deployment.
- Slurpd is now completely removed from 2.4
- There's new terminology – Provider/Consumer
- MirrorMode and Multi-Master now available
- Replication needs to support complex environments
- Wonderful things can be done with the “Dynamic Configuration Backend”

The History of Replication

- Slurpd was the first form of replication
- Slurpd was a standalone daemon plagued with problems (in brief):
 - slurpd never rerouted requests
 - It was not reliable
 - It was extremely sensitive to the ordering of records in the relog
 - more....

The History of Replication

- It could easily go out of sync, at which point manual intervention was required
- It wasn't very tolerant of unavailable servers.
- It only worked in push mode
- It required stopping and restarting the master to add new slaves
- It only supported single master replication
- **Slurpd is no longer part of OpenLDAP 2.4**

The History of Replication

- Syncrepl has none of those weaknesses
- Syncrepl was born on April Fools Day 2003 and is documented in the Admin Guide and RFC 4533 - “LDAP Content Synchronization Operation”
- It is extremely flexible and “JGOWI”
- Push Based, Pull based, Proxies.....

The History of Replication

- OpenLDAP 2.4 adds:
 - MirrorMode (Active-Active Hot-standby)
 - N-Way Multimaster Replication
 - More sophisticated Syncrepl configurations
 - Replicating slapd Configuration (syncrepl and cn=config)



Deployment Alternatives

- Syncrepl
- Delta-syncrepl
- Syncrepl Proxy Mode
- MirrorMode
- N-Way Multi-Master

LDAP Sync Replication (Syncrepl)

- LDAP Sync Replication engine, syncrepl for short
- Consumer-side replication engine
- Resides at the consumer and executes as one of the slapd(8) threads.
- Uses the LDAP Content Synchronization protocol (or LDAP Sync for short) - RFC4533
- LDAP Sync provides a stateful replication which supports both pull-based and push-based synchronization and does not mandate the use of a history store.

LDAP Sync Replication (Syncrepl)

- Pull-based replication - periodically polls the provider for updates.
- Push-based replication - consumer listens for updates that are sent in realtime
- Syncrepl tracks status of the replication content by maintaining and exchanging synchronization cookies
- Consumer replica can be constructed from a consumer-side or a provider-side backup at any synchronization status.
- Syncrepl can automatically resynchronize the consumer replica up-to-date with the current provider content.

LDAP Sync Replication (Syncrepl)

- Session log can be used in the provider which stores the entryUUIDs of a finite number of entries deleted from a database in order to use the delete phase
- LDAP Sync provider maintains a contextCSN in the suffix entry (change sequence number = CSN)
- It is the largest entryCSN in the provider context (depending on outstanding transactions)
- contextCSN maintained primarily in memory and written at shutdown, but can be checkpointed.
- The format of a CSN string is: `yyymmddhhmmssz#s#r#c` where `s` is a counter of operations within a timeslice, `r` is the replica id (normally zero), and `c` is a counter of modifications within this operation.

LDAP Sync Replication (Syncrepl)

- Whole database scanned if contextCSN not found and new one generated
- Consumer also stores its replica state, which is the provider's contextCSN received as a synchronization cookie
- New Consumer doesn't change provider config
- No provider restarts needed
- Consumer replication can stop without the need for provider-side changes and restart.

LDAP Content Synchronization Protocol (LDAP Sync)

- RFC 4533 - LDAP Content Synchronization Operation - June 2006
- refreshOnly and refreshAndPersist
- Eventually-convergent content synchronization
- Client can be notified that a complete reload is needed (used in Delta-syncrepl)
- Copy of “DIT Fragment”

LDAP Content Synchronization Protocol (LDAP Sync)

- Inconsistencies resolved on subsequent syncs
- Not good for bandwidth challenged apps/deployments
- Not for use with apps that need “transactional data consistency”
- Example of refreshOnly and refreshAndPresist to follow

LDAP Content Synchronization Operation - refreshOnly

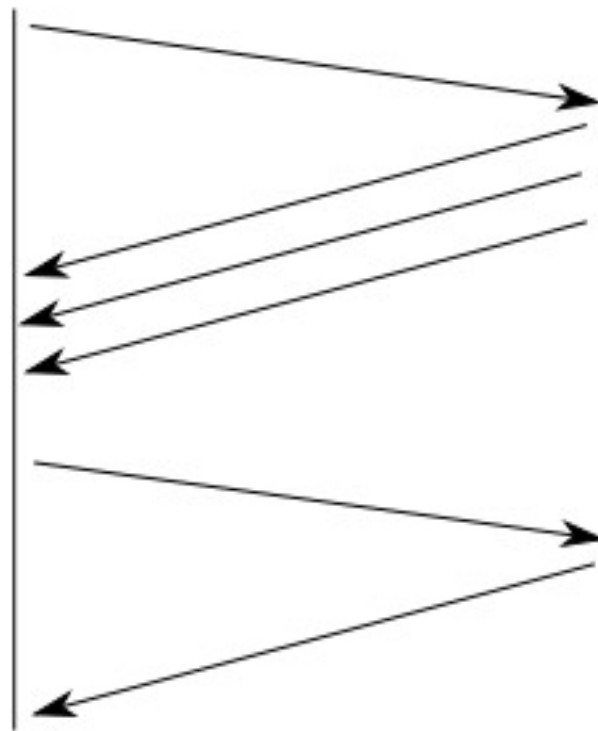
Client

Server

1. Initial client copy Sync request - search request with Sync Request Control with mode set to 'refreshOnly'

3. Polls for updates providing the previously issued syncCookie

5. Repeat using syncCookie, i.e. go back to step 3.



2a. Returns content matching search and with each entry provides a Sync State Control which contains the 'entryUUID'

2b. Follows with a SearchResultDone with a 'Sync Done Control' which provides the syncCookie - this cookie represents the session state.

4a. Use present or delete phase? Both can be used, present brings client copy up to a point where delete can begin.

4b. Server uses syncCookie as an indicator of what client got before and then sends copies of entries that have changed. **All** attributes are sent.

LDAP Content Synchronization Operation - refreshAndPersist

Client

Server

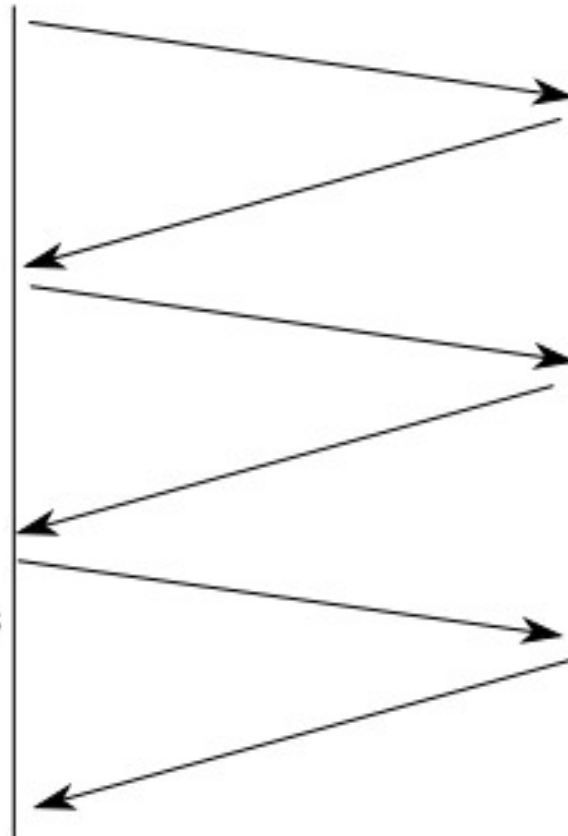
1. Same as refreshOnly request, but refreshAndPersist mode set.

3. After receiving the message, the client will construct a synchronized copy as described in the refreshOnly mode.

5a. For returned entries the SearchResultEntry will have the Sync State Control set to either; add, delete or modify

5b. Waits for server to send entries

7. Client refreshes if disconnects and provides last syncCookie if it has one.



2a. Same as refreshOnly mode.

2b. This time, send a Sync Info Message to client indicating refresh stage is done and then enters the persist stage

4. Server can now send change notifications based on original Sync Search Request

6. Server may terminate Sync Operation. If it doesn't provide a cookie, a full refresh is needed by client.

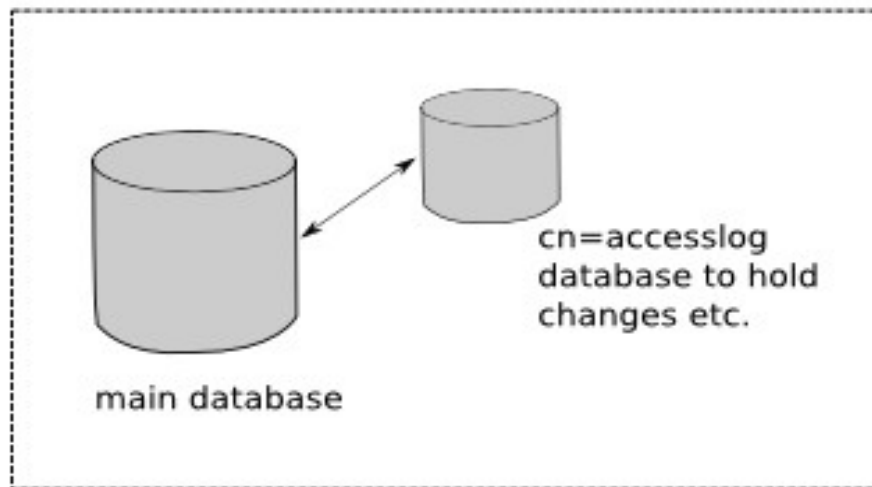
Delta-syncrepl

- Disadvantages of LDAP Sync
 - LDAP Sync replication is an object-based replication
 - Both the changed and unchanged attribute values are processed
 - Excess traffic generated for small changes
- changelog-based variant of syncrepl

Delta-syncrepl

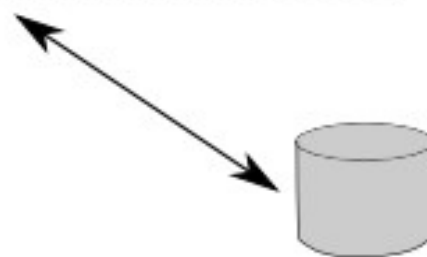
- Maintains a changelog on the provider
- Consumer checks the changelog for the changes it needs
- If a replica is too far out of sync, switches to conventional syncrepl
- Switches back to the delta-syncrepl mode when fully sync'd

Delta-syncrepl



Master/Provider

Delta-syncrepl is a changelog-based variant of syncrepl. It works by maintaining a changelog of a selectable depth on the provider. The replication consumer checks the changelog for the changes.

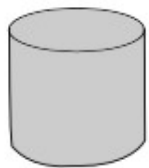


Consumer which uses syncrepl and the "syncdata=accesslog" setting. Switches back to normal syncrepl if gets too far out of sync, then once caught up goes back to delta.

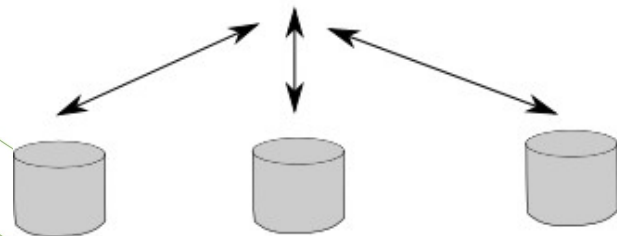
Syncrpl Proxy Mode

Push Based Replication
(replacing slurpd)

Master/Provider



Primary directory also contains back-ldap databases that replicate from the Master directory and push out changes to the replicas



Replicas

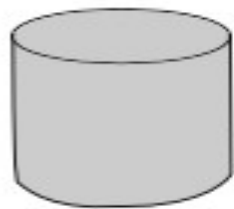
Replicas are readonly, but referrals can be handled by clients or using the chaining overlay.

- *refreshAndPersist* must still be initiated from the consumer
- Firewalls may need provider initiated push-mode replication
- slapd-ldap proxy is set up near or with the provider that points to the consumer
- syncrpl engine runs on the proxy and points to provider

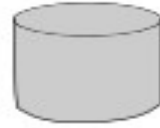
Syncrpl Proxy Mode

Push Based Replication
(replacing slurpd)

Master/Provider



Standalone
LDAP Proxy



Replicas

Primary directory is a standard OpenLDAP Master, ldap proxy using Syncrpl pulls in changes from the master and pushes out to replicas. Useful if you don't have access to original master.

Replicas are readonly, but referrals can be handled by clients or using the chaining overlay.

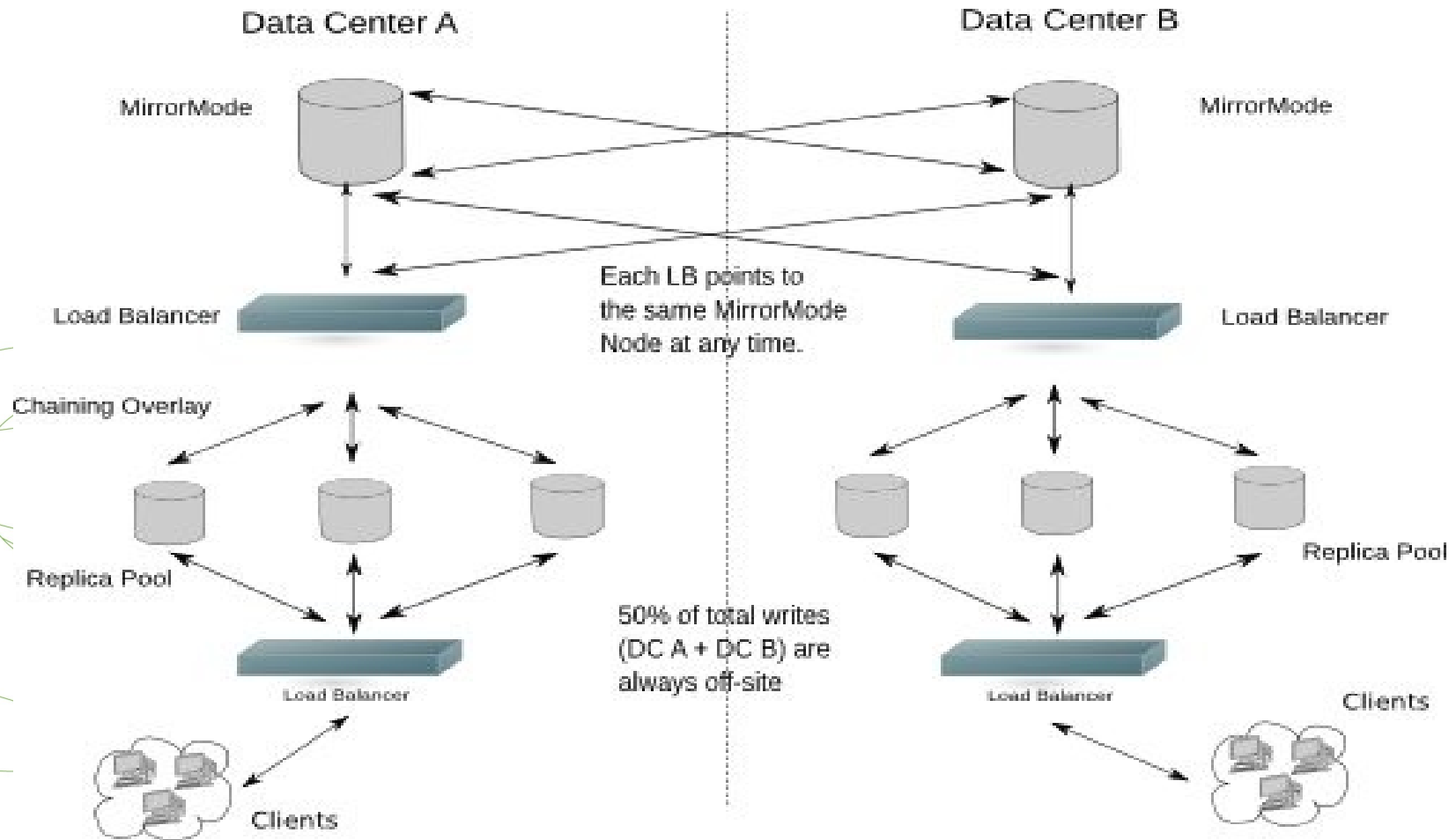
MirrorMode

- Is a Active-Active Hot-Standby solution
- External front end needed
- Not a Multi-Master solution
- Syncrepl also allows the provider nodes to re-synchronize after any downtime
- Delta-Syncrepl is not yet supported

MirrorMode

- Two providers are set up to replicate from each other
- An external frontend is employed to direct all writes to only one of the two servers.
- The second provider will only be used for writes if the first provider crashes
- automatically catch up to any changes on the running provider and resync.

MirrorMode



N-Way Multi-Master

- Uses Syncrepl to replicate data to multiple provider ("Master") Directory servers (up to 4096 to be exact!)
- Avoids a single point of failure
- Supports complex topologies
- Providers can be located in several physical sites
- Good for Automatic failover/High Availability
- Requires synchronised time source - ntp

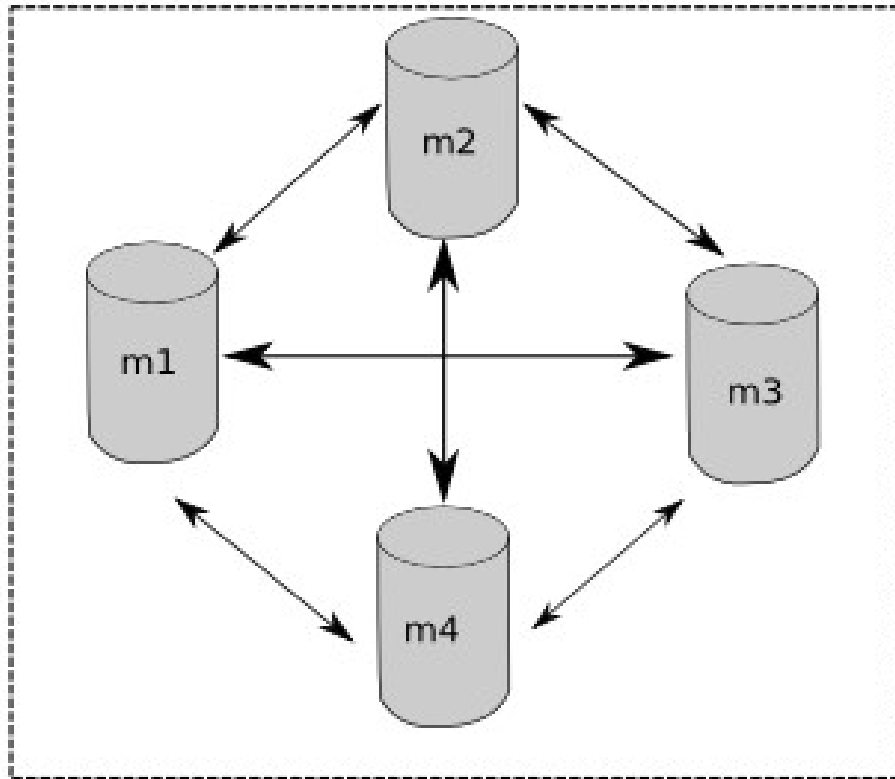
N-Way Multi-Master

- It has NOTHING to do with load balancing
- Providers must propagate writes to all the other servers
- Network traffic and write load spreads across all of the servers the same as for single-master.
- Server utilization and performance are at best identical for Multi-Master and Single-Master replication
- Single-Master is superior because indexing can be tuned differently to optimize for the different usage patterns between the provider and the consumers.

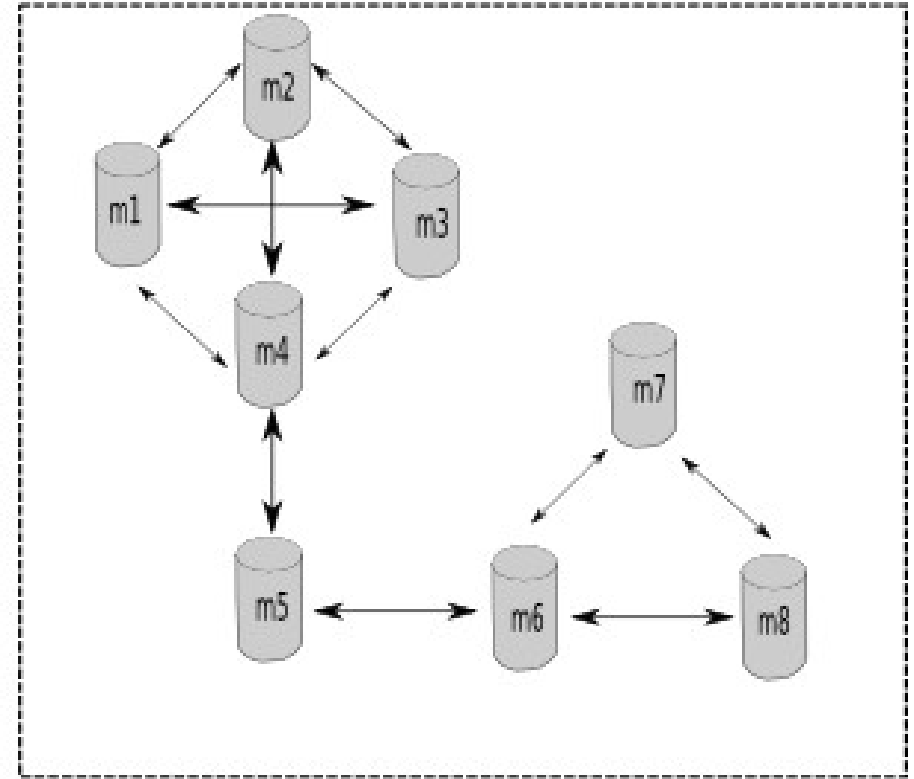
N-Way Multi-Master

- Breaks the data consistency guarantees of the directory model
- Crashed, or network link has failed???
- **More:**
 - <http://www.watersprings.org/pub/id/draft-zeilenga-ldup-harmful-02.txt>
 - <http://www.openldap.org/faq/data/cache/1240.html>

N-Way Multi-Master



Example of a Normal topology



Example of a ComplexTopology

Replication Best Practices

- Use a replication user
- You can restart replication by adding/removing `syncrepl` statement if needed when changing replication credentials/certificates
- MirrorMode always good enough for most enterprises dependent on global requirements
- Slapo-chain to chain writes back to provider

Replication Best Practices

```
[ghenry@suretec ~]$ ldapsearch -x -LLL -H ldap://master:389 -s base -b  
'dc=suretecsystems,dc=com' contextCSN dn: dc=suretecsystems,dc=com  
contextCSN: 20090217102328.285652Z#000000#000#000000
```

```
[ghenry@suretec ~]$ ldapsearch -x -LLL -H ldap://slave:389 -s base -b  
'dc=suretecsystems,dc=com' contextCSN dn: dc=suretecsystems,dc=com  
contextCSN: 20090217102328.285652Z#000000#000#000000
```

- Backups – online(ldapsearch)/offline(slapcat)/physical
- Checkpointing
- Monitoring:
 - Nagios
 - ZenOSS etc.

New cn=config replication features in the pipeline

- cn=config: sharing, conditionals
- cn=config can be fully replicated
- You may have consumers (slaves) that don't want or need all config to be replicated
- New attribute needed to tag certain config objects with the serverIDs to which they apply
- Attribute may be added to olcDatabaseConfig and olcOverlayConfig

Conclusion

- OpenLDAP is the fastest directory software in the world. Now it also provides the best replication features in the world.
- There is always a replication model to suit your needs, if not send us a patch!
- Don't hide in a corner, join our community, ask questions and help others! - <http://www.openldap.org>

Questions?

