OpenLDAP Development

Setting a Course for the Enterprise

Howard Chu  hyc@symas.com
ODD/Tuebingen October 13, 2006
Status Summary

• Progress since ODD 2004
  – Many more useful overlays
  – Mostly complete back-config

• New developments
  – Syncrepl enhancements
  – Performance, further refactoring
Overlay Status

• Goals met since 2004
  – More backend entry points handled
  – SLAPI reimplemented as an overlay
  – backglue reimplemented as an overlay
Overlay Status (2)

- Enterprise-oriented features
  - In-directory password policy
  - Referential integrity
  - Translucency
  - Attribute uniqueness
  - Value sorting
  - In-directory logging
Overlay Status (3)

• Conclusions
  – Overlays have met their design goals
  – Overlays continue to improve incrementally
CN=config Status

• Goals met since 2004
  – Converted config.c to table-driven mechanism
  – Maintained backward compatibility with existing slapd.conf syntax
  – Fully dynamic capability for majority of config items
    • ACLs
    • Schema
    • Databases
    • DB indexing
    • Dynamic modules
CN=\text{config} Future

- Zero administrative downtime
  - dynamically replace/re-exec binaries
- Fine-grained syncrepl for shared configuration components
- \texttt{config\_entry} API
  - allow backends/overlays to access their own config entries and persist private state
- Your suggestions welcome...
New Developments

• Syncrepl enhancements
  – Delta-syncrepl
  – Push-mode syncrepl
  – Mirrormode

• Upcoming work
  – lessons learned from deployment, ITS’s
Syncrepl

- Delta-syncrepl
  - Addresses bandwidth concerns from plain syncrepl
  - Relies on a persistent log of changes
  - Ordering of log entries is fully serialized; no out of order updates
  - Automatic fallback to plain syncrepl if consumer loses sync with log
Syncrepl...

- Push-mode syncrepl
  - Just a syncrepl consumer sitting on back-ldap
  - Can add a customization overlay for mapping the contextCSN to a suitable remote attribute, or to store the contextCSN locally
  - Provides a simple, robust, dynamically configurable replacement for slurpd
Syncrepl...

• Mirrormode
  – Allows a single active master and many standby masters
  – Preserves single master consistency while allowing automatic promotion of alternate masters
  – Requires use of an external frontend to guarantee that writes are only sent to a single master at a time
  – Addresses the high availability/SPOF concerns with minimal fuss
  – Already in use at some Symas customer sites
Syncrepl...

- Revive support for multiple consumers/contexts in a single DB context
  - required for meaningful glue behavior
  - touches on multimaster consistency issues
    - requires synchronized clocks for all contexts
    - requires use of hostID field of CSN
    - requires per-consumer contextCSNs in addition to (*not instead of*) provider contextCSN
Performance

- Fixed Lightweight Dispatcher
  - eliminated unnecessary locking in connection manager
    - slapd-auth test against back-null yielded over 32000 binds per second on 100Mbps ethernet
    - over 128000 frames per second - ~90% of available bandwidth – essentially saturated
    - No other LDAP server we tested delivers this speed on identical hardware
Performance...

• Fixes to pcache (proxy cache) overlay
  – Fixed O(n^2) query containment behaviors
  – Optimized case where a single entry is expected
  – Added negative caching support
  – Results:
    • pcache used to be slower than a direct proxy lookup above about 500 queries
    • pcache is now always faster than passing through
Performance...

- libc malloc() still has a major impact
  - refactored Entry and Attribute management to further reduce number of calls to malloc
  - using a thread-oriented allocator like hoard provides further advantages
malloc Performance

see openldap–devel August 30 2006...
malloc Performance

- Tested on 2.6 Linux kernel with glibc 2.3.3
- Results will obviously vary by platform
- glibc malloc does not handle tight memory conditions gracefully
- libumem is good but libhoard is better
  - performance difference is minimal
  - umem on non-Solaris appears unmaintained
Performance...

- Scaling to large deployments
  - Demonstrated performance at over 150 million entries
    - November 2005: 16600 queries/second, 3400 updates/second
    - April 2006: 22000 queries/second, 4800 updates/second
  - Over 1 terabyte of real data
  - Other popular directories’ claims of scaling are provably false
    - Several other products were tested with the same data, all of them failed
    - Only OpenLDAP passed
Performance...

• benchmark details available on www.symas.com

• we may want to consider investing effort in a C-based benchmarking framework
  – existing frameworks are not credible
    • DirectoryMark in perl, fast enough to measure slow directories, not fast enough for OpenLDAP
    • SLAMD in java, same story again
A Word from Our Sponsors

• OpenLDAP is no longer only of interest to a handful of developers
  – Significant investment from Symas, HP, Sys-Net, Sendmail (pcache), others.
  – Is now running all of HP’s corporate IT, displacing previous proprietary server
  – Feature wise, performance wise, there is no credible competition
The Road Ahead

- The unmatched code quality is not matched by documentation quality
  - Working on OpenLDAP Admin book, to be published by Addison-Wesley in Spring 2007
  - The manpages need to be fleshed out, missing pages need to be written
The Road Ahead...

- More work on back-config
- Work on scale-out, vs scale-up
  - allow multi-terabyte DBs to be served without requiring a single giant server
    - page-oriented, lock-free DB to allow multiple backends to serve portions of a single shared DB
    - distributed indexing
    - cluster-friendly optimizations
Final Thoughts

• OpenLDAP is taking over the enterprise
  – reliability, flexibility, scalability beyond all users’ or competitors’ comprehension

• The OpenLDAP community continues to thrive
  – with special thanks to the corporate members of the community

• Code quality is self-evident, but needs to be balanced with documentation quality