Evergreen:Library deOSSification (pt. 2)

Dan Scott dscott@laurentian.ca / http://coffeecode.net

FSOSS 2008 October 23, 2008



Who I am

Dan Scott

Systems Librarian, Laurentian University

Project Manager, Project Conifer

Evergreen developer

Who I have been

IBM tech writer / information architect / software developer / product planner / open source guy

Library (information systems) student

English and Philosophy student

High school punk in the smoking section

A tour of library technology



MARC metadata format

00962njm 2200253 a 4500001001100000005001700011008 004100028035002000069050000900 0891000024000982450046001222600 038001683000036002063490020002 4250501340026251101280039665000 3300524700001500557700001700572 7000019005897000020006089490074 0062859600060070203-00181372002 0121132036.0020121s1999 cauuuu eng d a(Sirsi) a353305 a5490 aLovano, Joe, d1952- aFriendly fire /cJoe Lovano and Greg Osby aHollywood, CA:bBlue Note,c1999 a1 sound disc :bdigital, stereo aCOMPACT DISC(S) aGeo J Lo --The wild east -- Serene -- Broad Way blues -- Monk's mood -- Idris -- Truth be told -- Silenos -- Alexander the Great aJoe Lovano, saxophone

```
<collection
xmlns="http://www.loc.gov/MARC21/slim">
<record>
 <leader>00962njm a2200253 a 4500</leader>
 <controlfield
tag="001">03-0018137</controlfield>
 <controlfield
tag="005">20020121132036.0</controlfield>
 <controlfield tag="008">020121s1999
                                        cauuuu
       eng d</controlfield>
 <datafield tag="035" ind1=" " ind2=" ">
  <subfield code="a">(Sirsi) a353305</subfield>
 </datafield>
 <datafield tag="050" ind1=" " ind2=" ">
  <subfield code="a">5490</subfield>
 </datafield>
 <datafield tag="100" ind1=" " ind2=" ">
  <subfield code="a">Lovano, Joe,</subfield>
  <subfield code="d">1952-</subfield>
 </datafield>
```

unAPI.info

- A Web standard for returning alternate representations
- Example: "In search of the common good"

- Supported formats:
 - HTML
 - Atom
 - RSS2
 - RDF/DC
 - OAI/DC
 - MARCXML
 - MODS*
 - OPAC

Character encoding

- Very slow transition to Unicode
 - Majority of library metadata is still encoded in MARC8 (aka ANSEL – genealogists take note)
 - Not all Unicode characters can be represented in MARC8
 - Upgrades to Unicode often drag along other expensive requirements (hello Oracle!)
- Evergreen is UTF8 from database through to display

Rich silos

- Libraries led the way in peer-to-peer sharing technology in 1984 with Z39.50 (huh?) protocol
 - SRW and SRU are less arcane variations on Z39.50 – still inter-silo communication
 - FUD about record sharing by large, co-operative non-profit agencies chilled much of the potential
- Evergreen offers OpenSearch as well as silo technology
 - Conifer will bring all of our records together

Legacy API access

A single transaction:

E<TIMESTAMP>R
^S52CVFFCIRCBASE^FcNONE^FW<PROFILE>^FE<LIBRARY>^NQ<BARCODE>^FW<
PROFILE>^UO<BARCODE>^dC3^Og^Fv200000^^O

Command line:

echo "input" | command1 -flags | command2 -flags

- This is not a pleasant way to build or extend applications.
- -Build your own API wrapper and Web access? Ugh.

OpenSRF API (Perl)

Fetching a library hierarchy:

```
$tree = OpenILS::Utils::CStoreEditor->new-
>search actor org unit(
    {"parent ou" => undef },
       flesh
                   => -1,
       flesh fields => { aou => ['children'] },
       order by => { aou => 'name'}
```

OpenSRF API (Python)

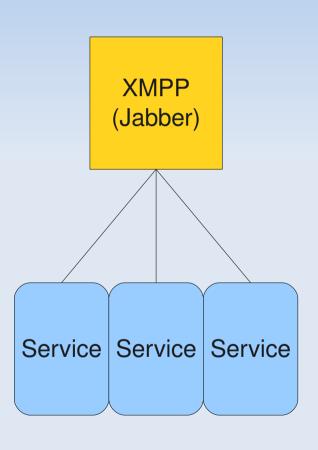
Fetching a library hierarchy (Python):

Fetching a library hierarchy (srfsh command line):

```
request open-ils.actor open-ils.actor.org_tree.retrieve
```

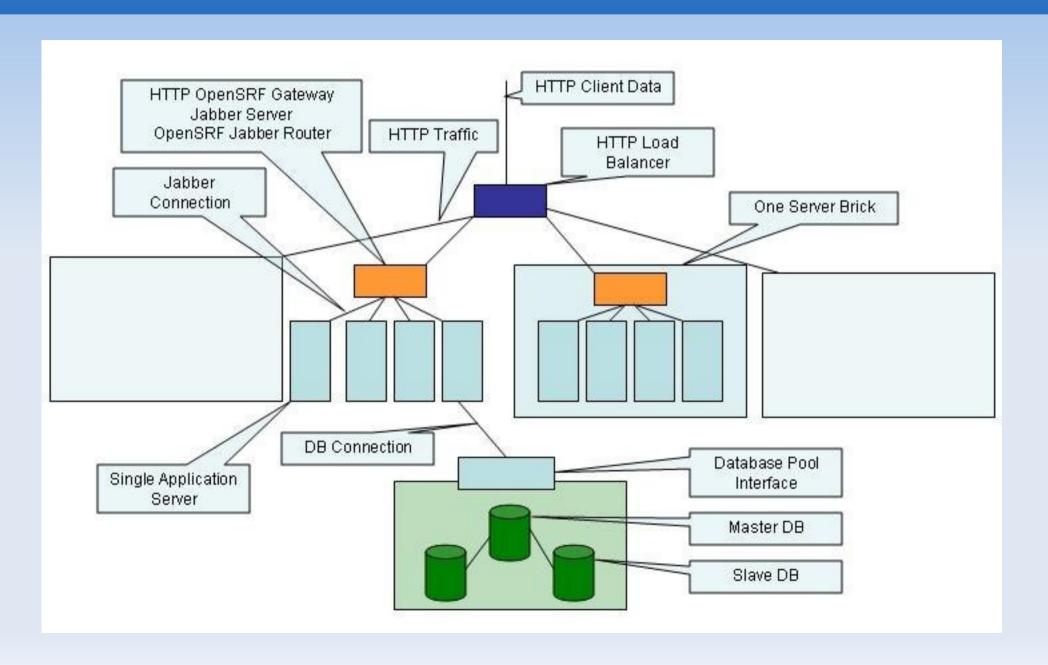
Okay, Python and srfsh just act as OpenSRF clients here, calling the Perl service. But you didn't want to implement that thing twice anyway, right?

OpenSRF architecture



- JSON-over-XMPP
- Applications register available services with a Jabber router
- Services can be clients of other services
- Also HTTP gateway and transport

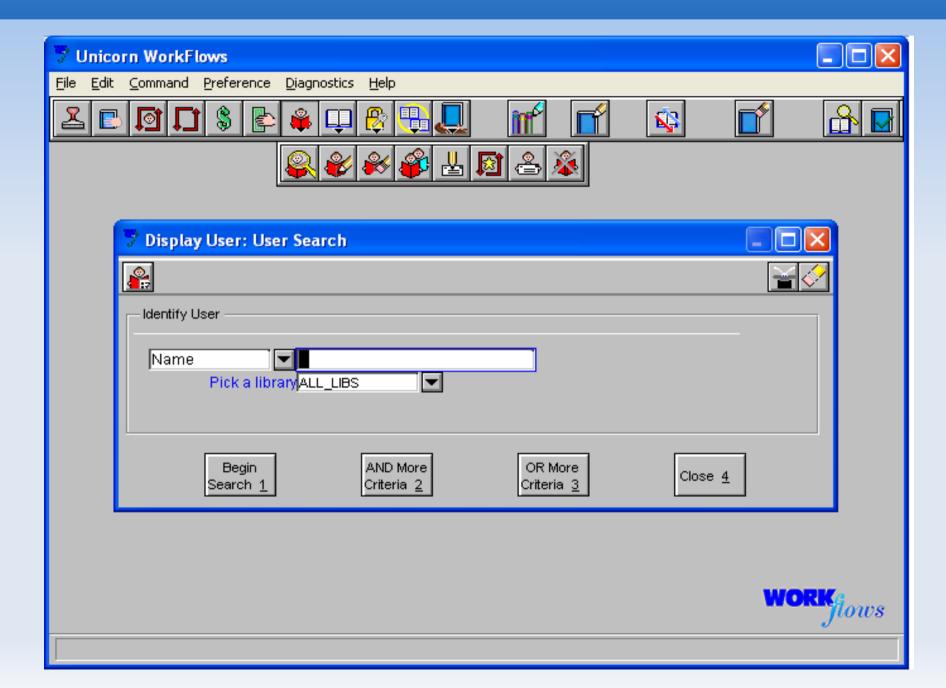
Evergreen architecture



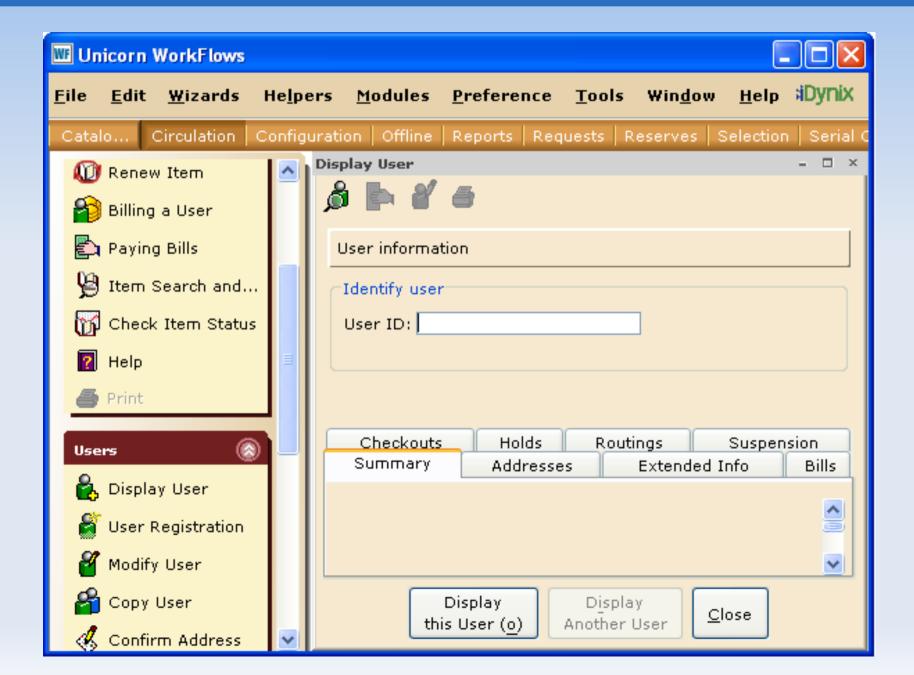
Library data stores

- The proprietary options run the gamut:
 - From non-relational databases with horrible reporting capabilities
 - To proprietary "embedded" databases with additional license required to write custom apps
- Evergreen runs on PostgreSQL
 - Uses PostgreSQL full-text search
 - Includes a powerful report builder

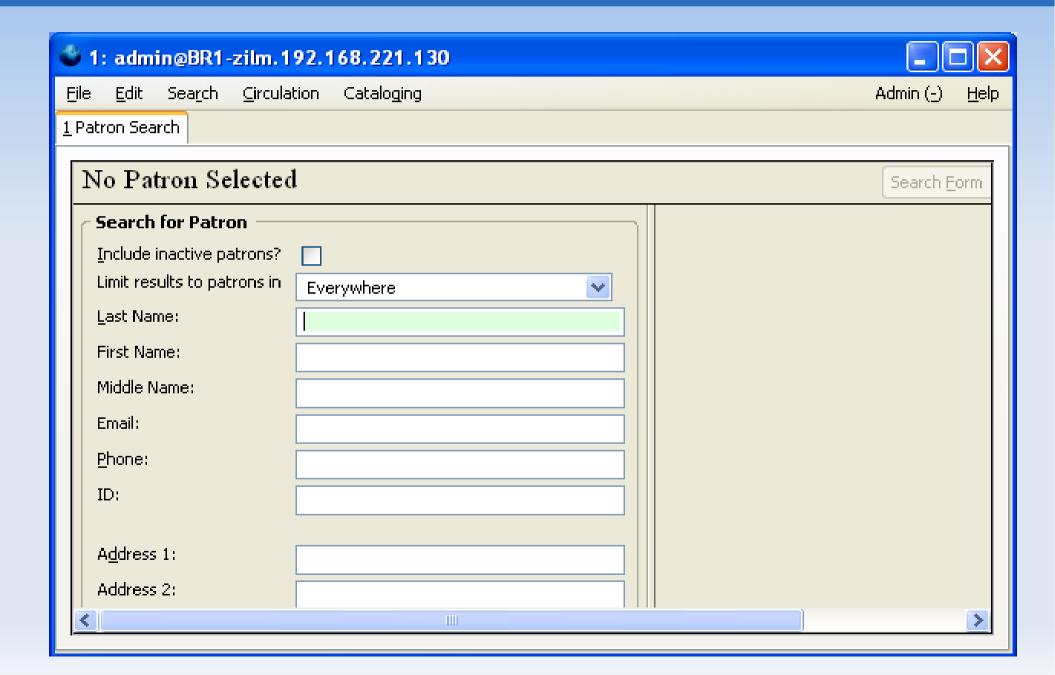
Aesthetics (staff side)

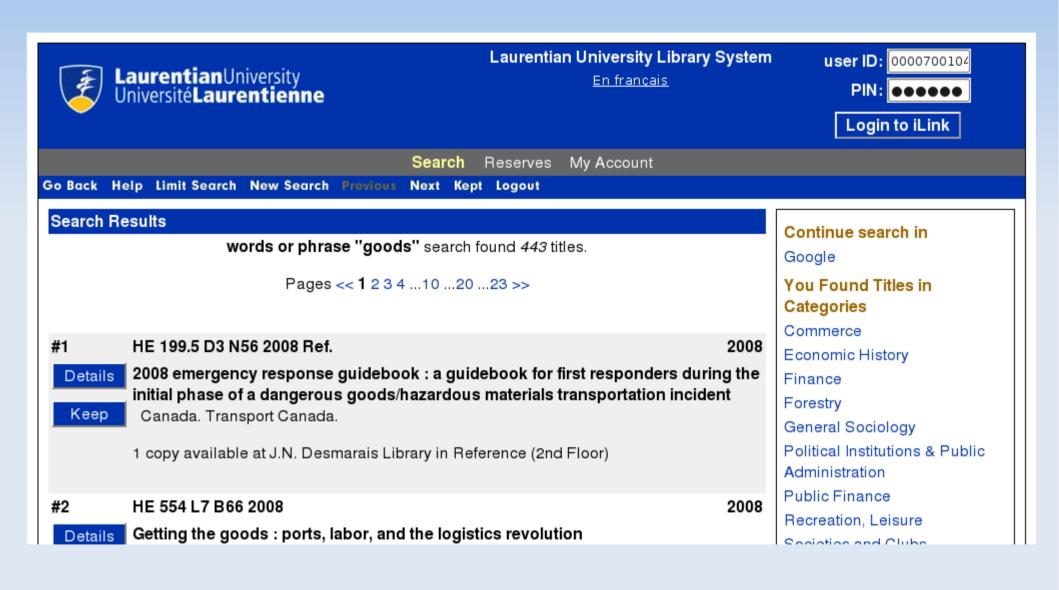


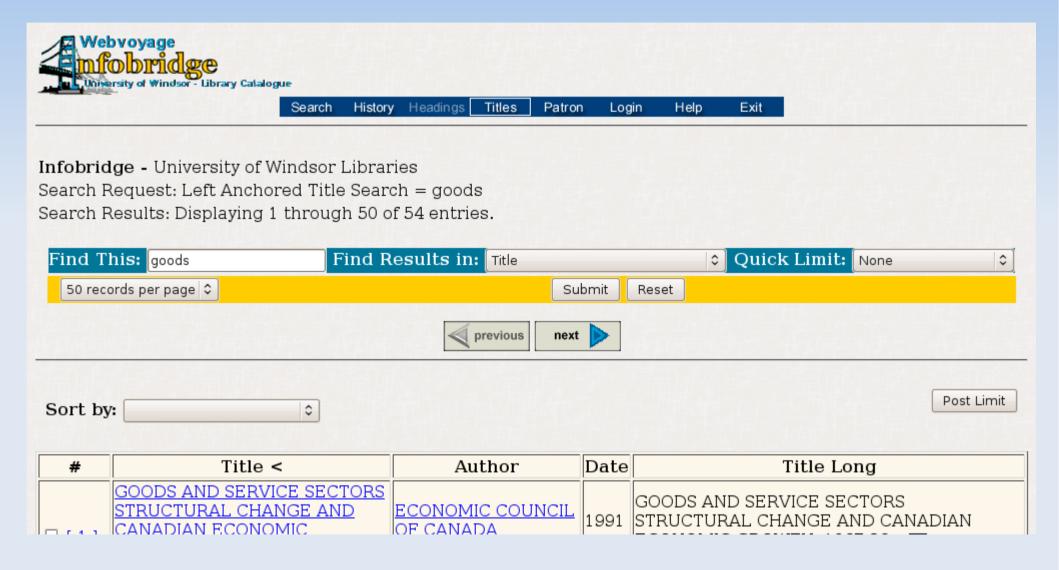
Aesthetics (staff side)

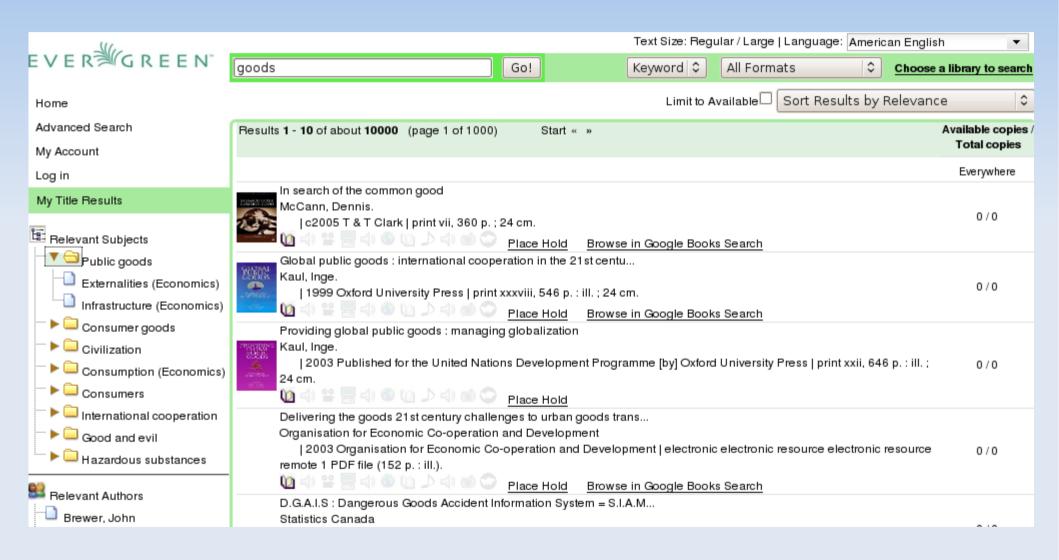


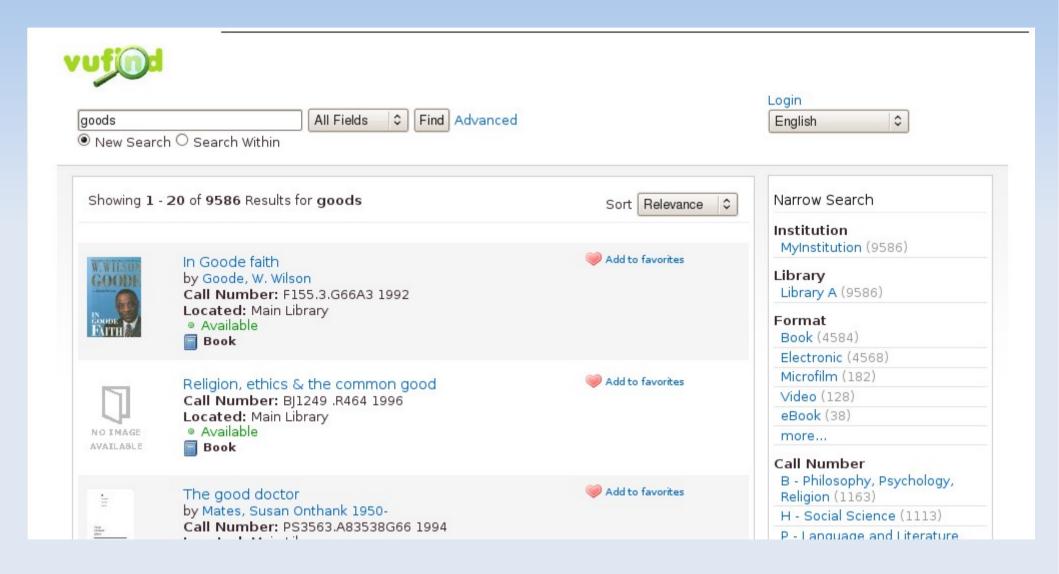
Aesthetics (staff side)



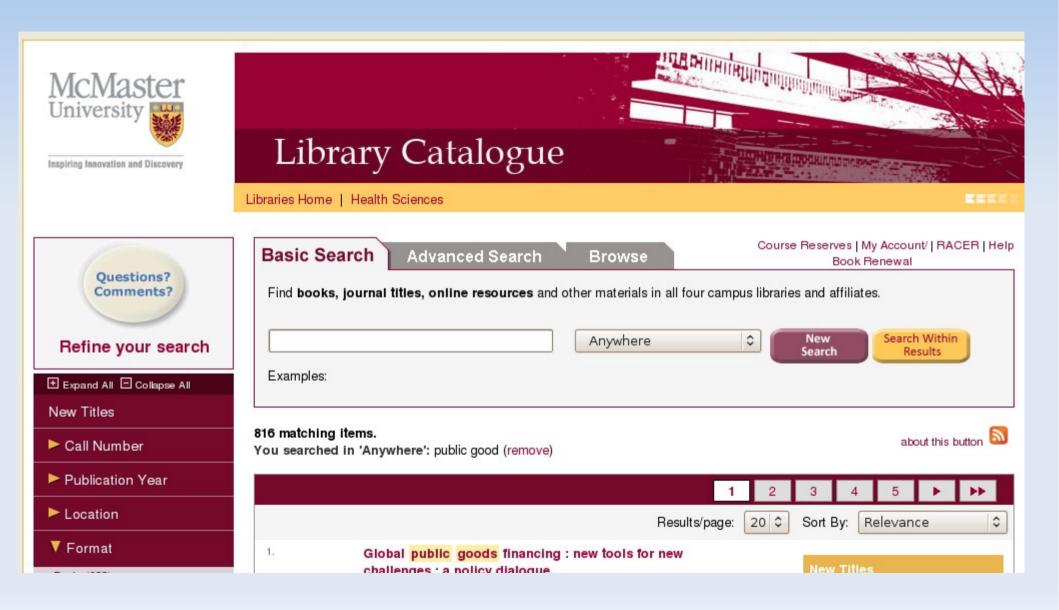












Books (809) Online (365) Research Collections (49) Videos & DVDs (40) Journals, Magazines, News... (7) Manuscripts (4) Audio (3) Subject Publication Type Language Geographical Region Subject Era Author

challenges: a policy dialogue By: Allen, Mark. Published: [New York: United Nations Development Programme, Office of Development Studies, 2002] MILLS Gov Pubs (2nd floor) UN2.. DE..90-Check Availability 2002G46 2. Advancing Public Goods [electronic resource]. By: Touffut, J.-P. Published: Northampton: Edward Elgar Publishing, 2006. Link to online resource - Preview this e-book for 5 minutes and request extended use e-Book Online 3. Public goods and public welfare By: Head, John G. Published: Durham, N.C.: Duke University Press, 1974. MILLS Bookstacks HJ 192 .H4 1975 Check Availability Competition in the Provision of Local Public Goods: Single Function Jurisdictions and Individual Ch... By: Reifsschneider, A. Petermann. Published: Northampton: Edward Elgar Publishing, 2006. Link to online resource - Preview this e-book for 5 minutes and request extended use e-Book Online Providing global public goods: managing globalization

New Titles Business Ethics of Innovation Hanekamp, Gerd Handbook of Environmental Economics : Environmental Degradation and Institutional Response... M@aler, Karl-G@oran Intellectuals and the Public Good: Creativity and Civil Courage Misztal, Barbara A. Structure and Agency in the Neoliberal University Canaan, Joyce E. Handbook of Development Economics Schultz, T. Paul

Cruel and unusual requirements

- Proprietary systems' have a weird fondness for proprietary hardware
 - One major product boldly stepped forward to support Linux (RHEL) in 2006, another in 2008
 - Incredibly expensive for a test environment
- Linux and commodity x86-64 hardware help make Evergreen scalable
 - Example: \$8K for 16 GB RAM, 8 CPU cores, 6 disk
 RAID arrays running Debian or Ubuntu

Support

- Many proprietary library system vendors offer indifferent or non-existent support
 - Best support often comes from other users via (usually closed) mailing lists
 - Monolithic service packs
- Commercial support is available for Evergreen
 - Open mailing lists
 - Ability to pick and choose fixes and features

So, obviously, FOSS!

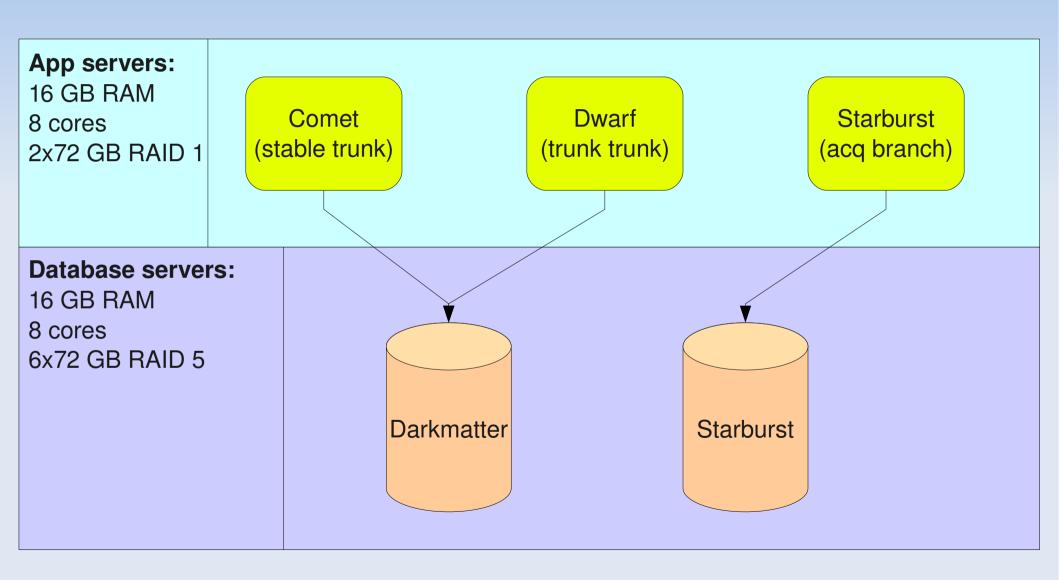
- Started investigating Evergreen as a serious option in Fall 2006
 - Project Conifer launched in Summer 2007
 - Published Evergreen Business Readiness Ratings and evaluation of multilingual support in Fall 2007
 - dbs added as a core Evergreen developer Fall 2007
 - First "internationalization" release Fall 2008



It's not just about books...

- A platform for library application development
 - Example: course reading lists and reserves integration with course management software (Sakai, Moodle, Blackboard, WebCT)
- Managing and making massive amounts of digital content usable
- Consolidating our collective efforts
 - Develop a process to do one thing the right way once that serves all consortial partners

Pilot hardware



Projected costs to go live

Co-op tech writer

Travel

\$60,000.00

\$12,000.00

\$10,000.00

Servers

Migration

De-duping

1st year support

Training

Redacted

Redacted

Redacted

Redacted

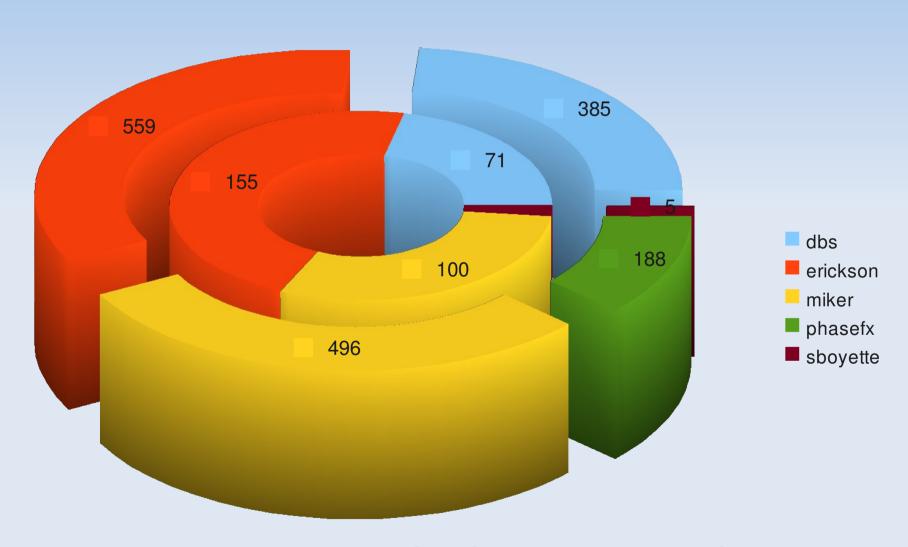
Total

\$199,000.00 (+ services)

Funding models

- Various pots of existing funds for development
 - Charge back for Art Rhyno's Knowledge Ontario service
 - Existing ILS and hardware budgets
 - "Percentage" model for existing employees
- Tried for some external funding but came away disappointed
- Cost-sharing for support & migration TBD

Are we making a difference?



of commits to Evergreen (outer ring) and OpenSRF repositories since September 2007

Challenges

- Resources and skills
 - Conifer can hire if you have skills and time
 - Developers, testers, doc writers, UCD specialists...
 - Building mutual support skills in community
 - One commercial support company
 - Evolving software engineering processes
- Time frames and legacy vendor licenses
 - Evergreen 2.0 in early 2009, go-live in May 2009
- Cost-sharing metrics: FTE vs. hospital beds

Whither Conifer?

- We've been fielding requests from institutions that want to join Conifer since shortly after the project launched
 - Colleges, hospitals, and other universities
 - Scale of economy once we reach a critical mass, we will be able to justify supporting ourselves
 - Public libraries? School libraries?
 - Knowledge Ontario?

Other FOSS library systems

Koha

- Launched in 2000, now hundreds of adopters world wide
- ca. 2006 / 2007, Koha 2.x was still trying to support MySQL 5 and was hitting MySQL 4 scalability limits
- Koha 3.0, released in 2008, rearchitected around MySQL limitations with a decoupled search layer
- Solid commercial support options

Other FOSS library systems

NewGenLib:

- Formerly proprietary system developed in India, released under GPL in early 2008
- Impressive feature support
- Written in Java
- Still finding a community

More information

- http://evergreen-ils.org (community site)
- http://conifer.mcmaster.ca (Drupal in the raw)
- http://coffeecode.net (my blog now with less online poker content!)