## Developing a crush on Evergreen

Dan Scott, Laurentian University

October 9, 2009

My relationship to Evergreen

About Conifer...

#### **Evergreen and Conifer**

ca. 2007

#### Cautionary tales (#1)

If you want to get things done, you need to commit resources to the tasks

### Cautionary tales (#2)

Giving your primary developer the additional role of project manager, on top of full-time librarianship responsibilities, might not give you the results you desire

#### **Evergreen and Conifer**

ca. 2008

How academics learn

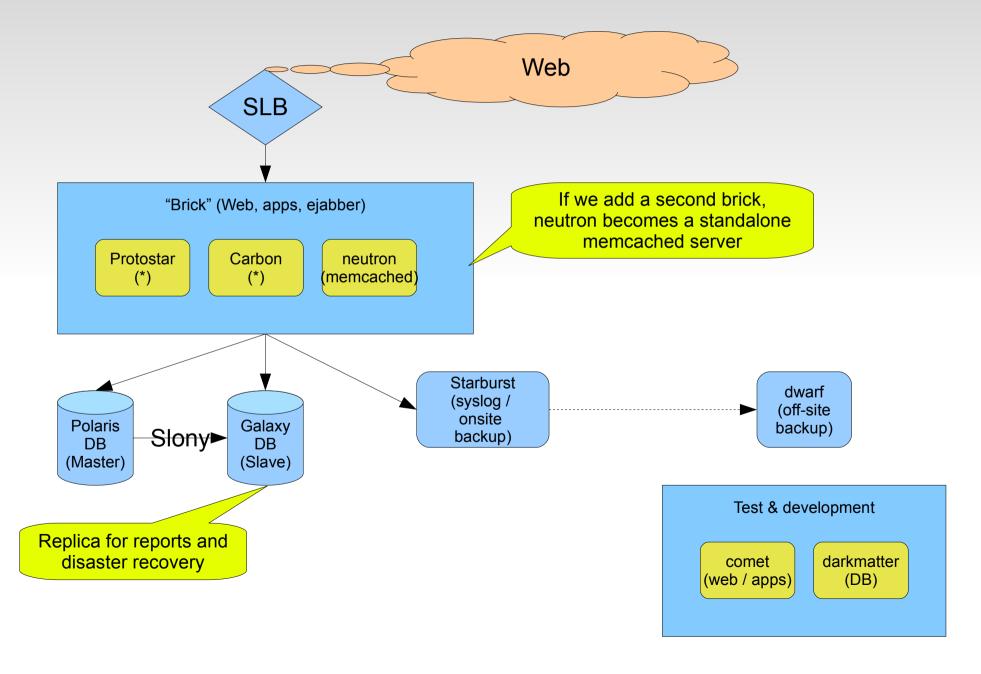
#### **Evergreen and Conifer**

ca. 2009

In production as of May 2009!

2 million bib records, 2.5 million copies 40,000 users

#### **Conifer hardware**



#### Cautionary tales (#3)

Nobody likes a prenuptial agreement... until things get messy

#### Cautionary tales (#4)

Free software is never *gratis*, and not necessarily cheap.

#### Cautionary tales (#5)

Budgets change – and they don't often grow.

Your plans may have to change accordingly.

Reservations about reserves

#### Cautionary tales (#6)

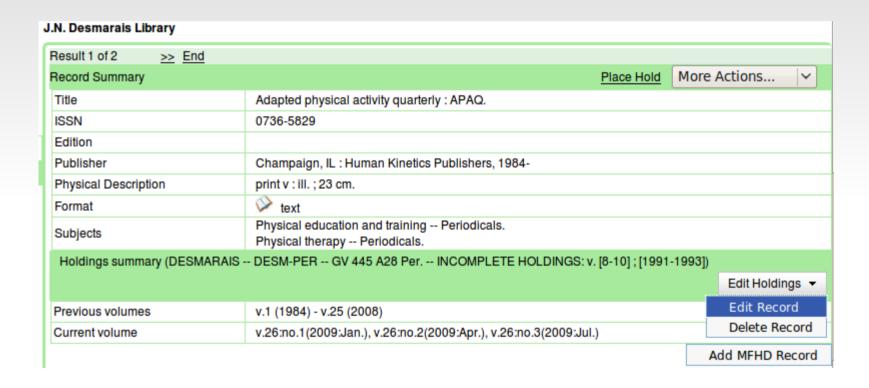
Loosely coupled means loosely coupled, and that's not always good

OLE, OCLC, take note!

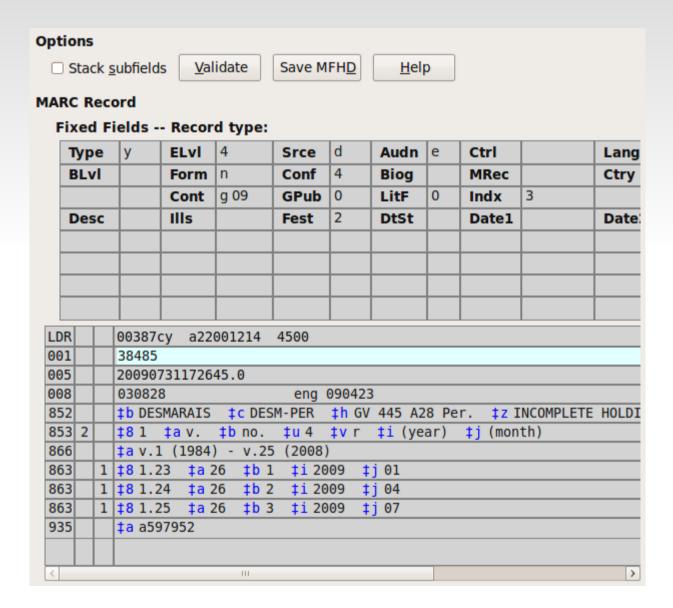
#### What hath Conifer wrought

- Internationalization support
- Migration scripts, customized skins, discussions
- Localized URIs\*
- Basic serials display and editing
- Reserves module\*\*
- Lots of input on how acquisitions works in Canadian academic institutions
- Z39.50 server maturity\*\*\*
- Early testing (and fixing) of the 1.6 release

### Serials in Evergreen 1.6



### Serials, in full MFHD glory



# Challenges at the academic level

- Circulation rules
  - Set date/time instead of just interval
  - Override time as well as date
- Interlibrary loan
  - Integration of incoming ILL items
  - Z39.50 server needs finer-grained indices
  - Z39.50 server has some i18n holes

# Challenges at the academic level

- OPAC refined metadata display in templates
- Reporting no default set of report templates
- Adjustment in general to user-friendly search
- Controlled process to roll out updates

#### **Downtime**

Cor	Conifer downtimes																
File Edit View Insert For					nat Tools Form				Help								
	K OI	\$ %	123 ▼	10pt	B	Abc	<u>A</u> •		<b>•</b>	<b>■</b> ▼		₽	Σ 🕶				
	A				В				С			D					
1	Date / time				Duration (minutes)			)	Planned?			Description					
2	1	7/27/2	2009 22:	30:00			2	40	Yes				Scott: upgraded Evergreen to OpenSRF rel_1_2 / rgreen rel_1_6_0				
3		7/27/2	2009 12:	30:00				30	No		th	em	entical queries chewing up massive CPU time, killed in manually via pg_top, then open-ils.storage needed to estarted				
4		7/28/2	2009 22:	00:00				30	Yes			•	inox services: added clustered services on carbon, set ane logging on starburst				
5		8/13	/2009 6:0	00:00				90	Yes				in Isard: Bad memory module in polaris needed acing				
6		9/10/2	2009 18:	21:00				24	No		re	port	in Isard: problem with opensrf; postgres fine; logs orted problems with osrf_json transport; restarting perl ices solved problem				
7		9/14/2	2009 10:	50:00				30	No		ea	ach (	in Isard: carbon and protostar weren't able to speak to n other (transport errors in osrferr.log) restarting perl ices on both machines solved the problem				
8		9/16	/2009 8:	35:00				35	No		R		in Isard: same issue as above. Opened ticket with esi,				
9									yes, but		ni ej no E Se 5 th	umbe abbe otice quin erver ney s	thought is that ejabberd isn't set to handle the right ber of connections. The fix was simple increase the ber of allowed connections form 1000 to 10 000 in berd's conf file. The problem occured when it was sed that something was routing traffic to carbon, and inox thought that we were running apache on two ers (which we are not) when Equinox believed this shut the instance on carbon down and things				
		9/17/2	2009 21:	00:00			1	20	not 120 r	nins	re	star	arted normally				

# Problems, features, and docs

Conifer consortial	_					Search		
				ogged in as dscott	Logout	Preferences	Help/Guide	About Trac
	Wiki	Timeline	Roadmap	View Tickets	N	lew Ticket	Search	Admin

Available Reports | Custom Query

#### {1} Active Tickets (92 matches)

- List all active tickets by priority.
- · Color each row based on priority.

Edit report

Copy report

Delete report

Ticket	Summary	Component	Version	Milestone	Туре	Owner	Status	Created
#280	Missing MFDH records	Cataloguing	None		defect	artunit	assigned	21/09/09
#38	Hanging Conifer: Adanced Search using exact phrase subject + search in another field	Public catalogue	None		defect	dscott	new	19/05/09
#215	Problems with the display of 856 fields in new version	Cataloguing	None		defect	dscott	new	29/07/09
#276	710-Added Entry- Corporate Name	Cataloguing	None		defect	risard	assigned	18/09/09

## Canada: a nation of Evergreen early adopters

- David Fiander, UWO (SIP & MFHD developer)
- BC Sitka (second live system)
- UPEI (first academic library)
- Innisfil Township, Ontario (first known self-hosted public library)
- Conifer (first academic library consortium)
- Natural Resources Canada
- Manitoba SPRUCE (public library consortium)
- Mohawk College, Hamilton, Ontario

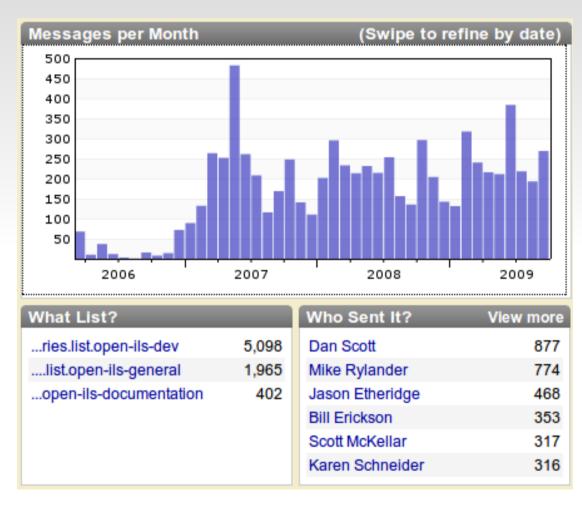
#### Measuring project health

- Number of committers & affiliation
- Number of submitted patches
- Documentation
- Testing
- Community and commercial support
- Current releases & release history

#### Good things to come

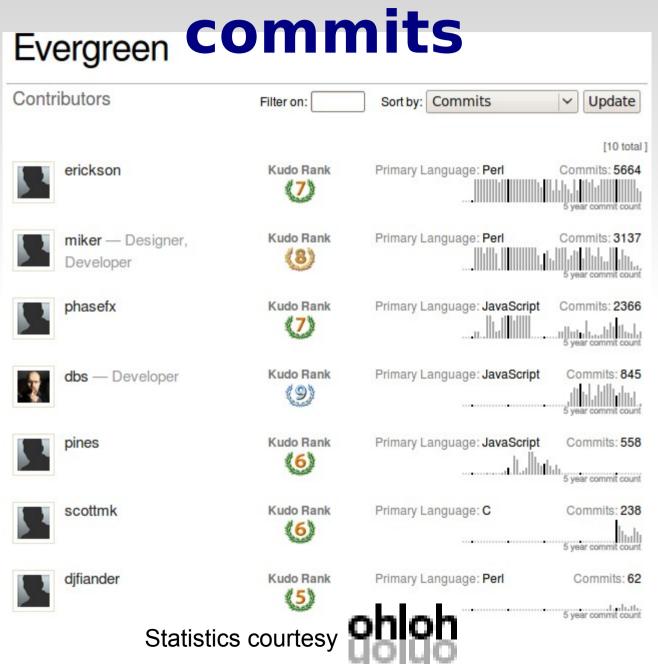
A maturing community

# Measuring community: email





# Measuring community:



WANTED: skilled developers!

Building an on-ramp for developers



Rethinking the release process

Better living through automation

I look forward to tests

Shorter release cycles

#### Documentation

A solid foundation

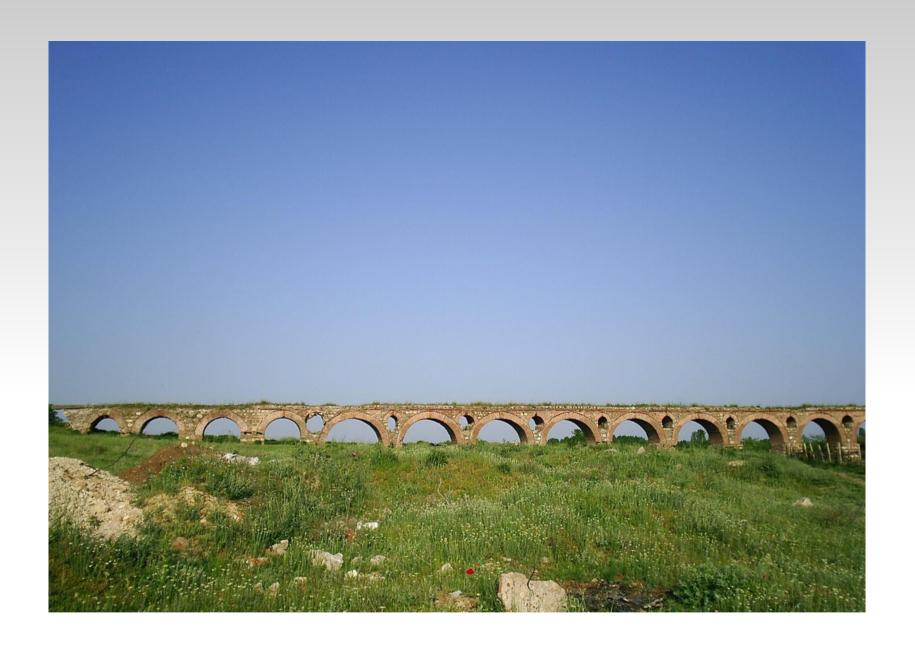
# Please bear with us...



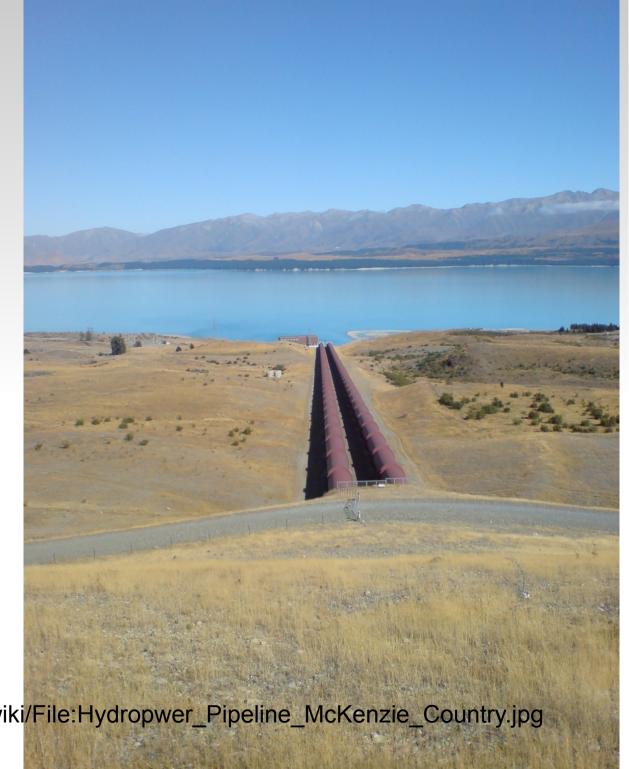




http://www.flickr.com/photos/julien\_harneis/3027936866/



http://www.flickr.com/photos/33619843@N08/3125532414/



http://commons.wikimedia.org/wiki<mark>/File:Hydropwer\_Pipeline\_McKenzie\_Country.jpg</mark>

## **More information**

- Evergreen home: http://evergreen-ils.org
- Blog aggregator: http://planet.evergreen-ils.org
- Mailing lists: http://evergreen-ils.org/listserv.php
- IRC: #evergreen on chat.freenode.net

## Developing a crush on Evergreen

Dan Scott, Laurentian University
October 9, 2009

1

My relationship to Evergreen

2

Long story short – I worked in the commercial software industry for 8 years after getting my MISt, rejoined the library world as an academic systems librarian with stars in my eyes, then ran into the ugly reality of modern ILS.

Even after taking the very expensive API training course for our ILS, I very nearly turned tail and fled – and would have, were it not for an introduction to (first) the code4lib community and (second) the emergence of the Evergreen project.

These days, I'm a core committer and getting my hands mixed up in most parts of the community... and project manager for Conifer.

About Conifer...

3

"Conifer" is a high-falutin' academic term for an evergreen.

Conifer started in July 2007 when some Ontario academic libraries got together to discuss their next library systems. We had a combination of libraries needing to migrate to new hardware (with a six-figure fee to move to the same software running on the new hardware), libraries paying exorbitant fees for a system that they had mostly had to work around, and libraries on systems with no prospect for future enhancements.

At the time, the obvious choice for a consortial open source system was Evergreen. Koha at that point was still working towards 3.0 and had not demonstrated scalability that would meet our needs, whereas Evergreen had Georgia PINES. 'nuff said.

Evergreen and Conifer

ca. 2007

4

Circulation, cataloging, and reports were present and accounted for.

All that we needed was acquisitions and some support for serials and reserves. Simple!

FANTASY-LAND: We committed a team of developers to get to work on fleshing out that support, in collaboration with the good folks from Equinox, and targeted a May 2008 migration to Evergreen.

REALITY: We didn't commit much at all in the way of skilled development resource to the effort. After a brief search for a project manager, that role was added to my workload.

## **Cautionary tales (#1)**

If you want to get things done, you need to commit resources to the tasks

4

Don't mistake light for heat. We had a number of people who were contributing resources to our goals on paper, but who were otherwise occupied with other tasks, or who simply didn't have the skills required to contribute base code.

A positive highlight: one of our summer students, Craig Ricciuto, put a lot of effort into the grunt work of making the staff client ready for translation.

I spent a lot of my time on setting up a test system for the purposes of hardware estimation and training. Some of that included migrating bib records and holdings.

There wasn't much development capacity available in the community to contract work out to...

### **Cautionary tales (#2)**

Giving your primary developer the additional role of project manager, on top of full-time librarianship responsibilities, might not give you the results you desire

(

A project manager is a marvelous addition to a project! They can bring objectivity, clarity, and the ability to focus on the moving parts.

Adding that role to your one development resource, however, usually won't work out that well. Objectivity becomes a lot more challenging. One thing that developers thrive on is long periods of uninterrupted concentration. Providing status updates, preparing meeting agendas and minutes, following up on questions, and tracking budgets and resources can interfere pretty heavily with that whole coding thing.

Early on, I raised this as an issue at status meetings, but we were never able to figure out how to shift the role to someone else.

Evergreen and Conifer

ca. 2008

•

Circulation, cataloging, and reports were present and accounted for.

All that we needed was acquisitions and some support for serials and reserves. Simple!

Sound familiar? We hadn't made a huge amount of progress in a year. We had taken a detour on acquisitions and serials by trying to graft a complete ERM (OpenTaps/Open for Business) onto Evergreen, but it turned out not to be a good fit.

There was some nascent acquisitions functionality – modeled in the database, and with a Pylons-based basic Web interface that was again subsequently abandoned. There was also a reserves RFC posted to the Evergreen community.

We moved our migration target to May 2009

How academics learn

:

Deploy. Test test test.

Arrange for a training session for each site about 4 months before you go live.

Throw it open to your library staff to test and train with and to report problems before you go live.

Throw it open to your library users to test and train with and to report problems 2 months before you go live.

Get very little feedback.

When you go live, discover that your frontline staff didn't really take advantage of that testing and training opportunity. Uh oh. Scramble.

**Evergreen and Conifer** 

ca. 2009

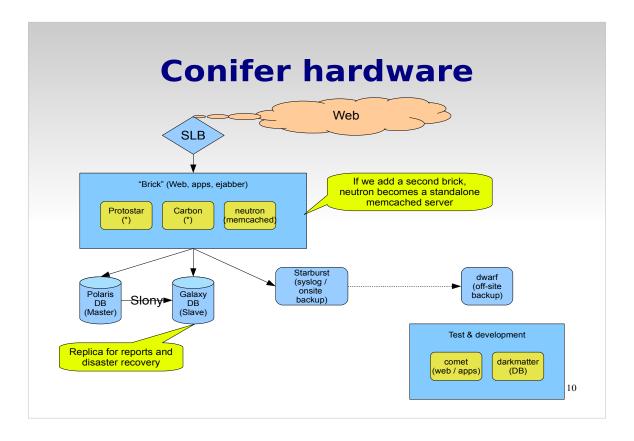
In production as of May 2009!

2 million bib records, 2.5 million copies 40,000 users

9

This time, we stuck to our guns and migrated in May 2009, despite a number of obstacles that came up:

- One of our largest partners decided to opt out of the project in late January, forcing the remaining smaller partners to share much more of the costs for hosting and hardware.
- Due to the recession, our budgets were slashed and we were forced to migrate our own data rather than use professional services.
- The 1.6 release, with serials and acquisitions, was not yet ready
- We had built a reserves module that was not really functional for our needs.



Most of the servers are 2x4-core Intel Xeon machines with 16 GB of RAM and a RAID 1 array of hard drives.

Polaris is a 4x4-core Intel Xeon machine with 48 GB of RAM and a RAID 10 array of hard drives.

We run Debian Lenny on all of our machines.

## Cautionary tales (#3)

Nobody likes a prenuptial agreement... until things get messy

11

A handshake always seems like enough at the beginning of a relationship. If you're entering into a multi-year partnership that involves the outlay of hundreds of thousands of dollars in hardware, hosting, and support services, though, it's a really good idea to enter into a written agreement that spells out what happens if someone pulls out of the partnership.

We didn't have a written agreement. We had to scramble to make things work. We're still absorbing the costs.

## **Cautionary tales (#4)**

Free software is never *gratis*, and not necessarily cheap.

12

It was clear to me that the main reason our largest partner dropped out was because they were surprised that hardware, professional hosting, and migration costs significant money. They had not contributed much towards development, so the delayed features that we had expected to be ready shouldn't have been a surprise either.

### **Cautionary tales (#5)**

Budgets change - and they don't often grow.

Your plans may have to change accordingly.

13

Making the decision to migrate our own data was a hard call, for a number of reasons. Although we had not signed a contract, we had obtained a quote for professional migration and the company in question fully expected us to follow through. That's business, I guess, but it's certainly awkward when those people are your co-developers and friends.

What was more challenging was the opportunity cost. Instead of focusing on enhancements, or primary development resource was now shouldering the burden of munging, and testing, a massive amount of data.

Reservations about reserves

14

I knocked out an RFC for an integrated reserves system in the summer of 2008 based on the needs of our library and those expressed on the Evergreen mailing lists. I had started by looking at ReservesDirect, an available open source reserves system, but had rejected it as being too loosely coupled.

Conifer hired a top-notch contractor to develop Syrup, our reserves system, and the plan that emerged was... to copy ReservesDirect, but to do it right. Z39.50 + SIP rather than OpenSRF calls.

Unfortunately, that means that we got a loosely coupled reserves system, that looks great, but doesn't have some of the core features that I had written up in the RFC.

### **Cautionary tales (#6)**

Loosely coupled means loosely coupled, and that's not always good

OLE, OCLC, take note!

15

The reason loose coupling isn't great with a reserves system built for a specific library system is that there aren't any standards for modifying library system data to do things like:

- ·Move an item from one location to another
- ·Change the circulation modifier for a given item

There are horrible workarounds that have been used in other systems, like checking the items out to magic users that then causes the attributes of the item to be changed. That's not good. Evergreen has nice Web services that enable you to do these things – the reserves system should use them.

Also, the interface is completely separate from the Evergreen catalogue. Yikes. So we're not using it. Something like VuFind could be tightly coupled, though, and would be nice if done right.

### What hath Conifer wrought

- Internationalization support
- Migration scripts, customized skins, discussions
- Localized URIs\*
- Basic serials display and editing
- Reserves module\*\*
- Lots of input on how acquisitions works in Canadian academic institutions
- Z39.50 server maturity\*\*\*
- Early testing (and fixing) of the 1.6 release

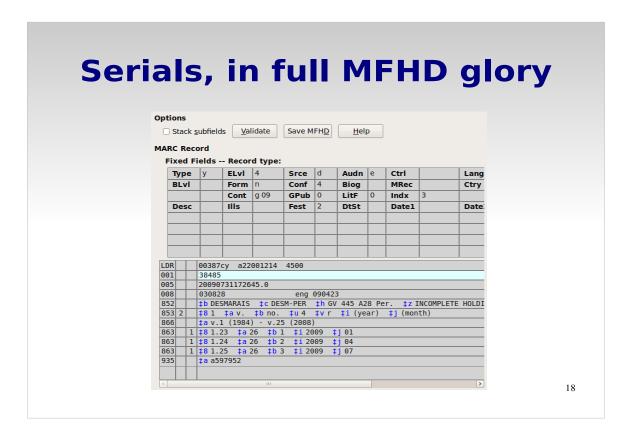
16

- \* Sponsored by Conifer, built by Equinox
- \*\* Built, but not currently being used
- \*\*\* A snap once Mike had built an SRU server

J.N. Desmarais Library	
Result 1 of 2 <u>&gt;&gt; End</u>	
Record Summary	Place Hold More Actions
Title	Adapted physical activity quarterly : APAQ.
ISSN Edition	0736-5829
Publisher	Champaign, IL: Human Kinetics Publishers, 1984-
Physical Description	print v : III. ; 23 cm.
Format	text
Subjects	Physical education and training Periodicals. Physical therapy Periodicals.
Holdings summary (DESN	MARAIS DESM-PER GV 445 A28 Per INCOMPLETE HOLDINGS; v. [8-10] ; [1991-1993])
	Edit Holding
Previous volumes	v.1 (1984) - v.25 (2008) Edit Record
Current volume	v.26:no.1(2009:Jan.), v.26:no.2(2009:Apr.), v.26:no.3(2009:Jul.)
	Add MFHD Rec

Those dialogue boxes on the bottom right only show up in the staff client, naturally. But the holdings display comes straight from the corresponding MFHD record.

Note that Windsor has taken a different route, and has opted to maintain all of their print serials holdings in their OpenURL resolver – so users search one interface for one consistent set of serial holdings (whether electronic or print).



Yes, 853/863 in full effect!

## Challenges at the academic level

- Circulation rules
  - Set date/time instead of just interval
  - Override time as well as date
- Interlibrary loan
  - Integration of incoming ILL items
  - Z39.50 server needs finer-grained indices
  - Z39.50 server has some i18n holes

19

ILL items are currently added to placeholder bib records with duplicated call numbers & barcodes, then automatically deleted once the item has circulated. A bit hokey.

## Challenges at the academic level

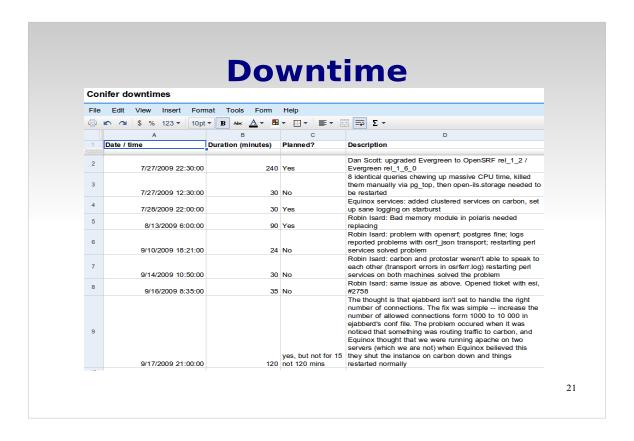
- OPAC refined metadata display in templates
- Reporting no default set of report templates
- Adjustment in general to user-friendly search
- Controlled process to roll out updates

20

OPAC "bibtemplate" allows display of one subfield or all subfields concatenated in a given field, but no demonstration yet of different displays based on, say, codes in numeric subfields (without displaying said subfield)

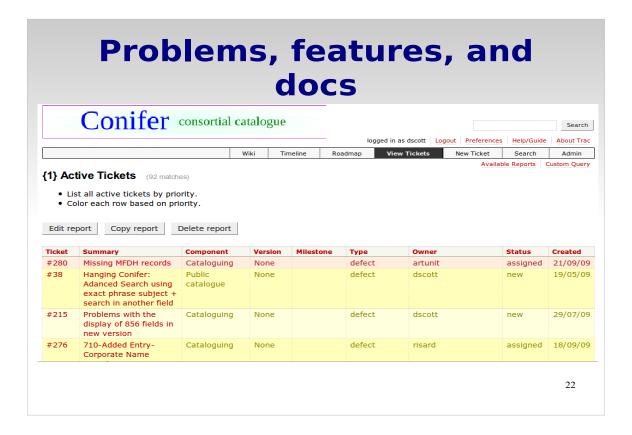
User-friendly search – for example, a publisher listed in a 710 \$4 with a relator code of 'pbl' is /not/ a corporate author, and no sane user would expect a search against an author index to pull up a publisher as a result.

Because the staff client uses much code in common with the OPAC, and several sites use different approaches for serials, we've broken each other's functionality at times.



5 unplanned downtimes in three months, most under one hour. Not too bad for a new system, although some of the repeated problems with communication across the cluster are a bit disturbing. Once we nail those down, though, we should be golden.

We use Google Docs to host a shared spreadsheet for tracking our down times & the reasons for those down times.



We use Trac to document our processes and track problem reports / feature requests.

At the moment, we have about 100 open issues. I try to open corresponding issues in the common Evergreen bug tracker when they are general Evergreen issues that would be of interest or concern to the entire community.

## Canada: a nation of Evergreen early adopters

- David Fiander, UWO (SIP & MFHD developer)
- BC Sitka (second live system)
- UPEI (first academic library)
- Innisfil Township, Ontario (first known self-hosted public library)
- Conifer (first academic library consortium)
- Natural Resources Canada
- Manitoba SPRUCE (public library consortium)
- Mohawk College, Hamilton, Ontario

23

James Fournie and Brandon Uhlmann from BC Sitka have contributed patches, and BC has contributed tons of documentation.

UPEI has contributed some patches that have led to cool new functionality.

Innisfil contributed a set of training screencasts.

Warren Layton from NRC has contributed a number of patches.

Robert Soulliere from Mohawk has edited a webcast.

#### Measuring project health

- Number of committers & affiliation
- Number of submitted patches
- Documentation
- Testing
- Community and commercial support
- Current releases & release history

24

"Introducing "Health" Perspective in Open Source Web-Engineering Software Projects, Based on Project Data Analysis"(
http://www.schatten.info/pub/iiwas/iiwas2006\_1.pdf)

"Monitoring the "Health" Status of Open Source Web Engineering Projects", http://www.schatten.info/pub/ijwis/ijwis2007.pdf

Balnaves, Edmund. "Open source library management systems: a multidimensional evaluation", Australian Academic & Research Libraries, March 2008.

## **Good things to come**

A maturing community

25

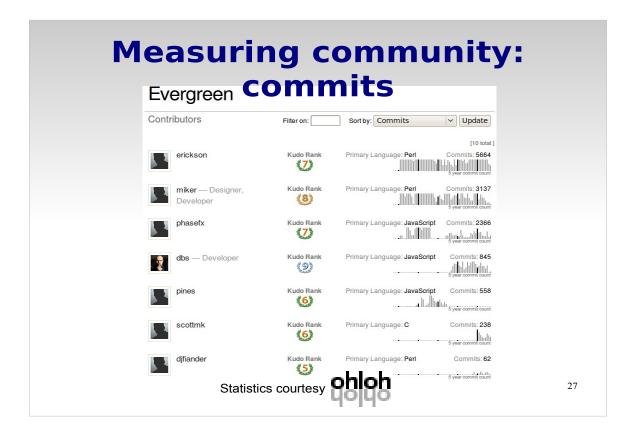
Differentiation of roles – and a broader sharing of responsibility

Revisiting community infrastructure



We've been averaging about 200 messages per month on the developer and general lists for the past few years. Some metrics (e.g. Business Readiness Rating) suggest 720 / month is the sign of a healthy community – that's based on a project with an adoption and developer base as large as Apache, which is insane.

200 messages per month, in my opinion, is a healthy base for a niche community like library system software. More would be welcome, of course!



There's a fairly concentrated set of committers for the project at the moment. We need to mitigate the "bus factor", for one thing, and broaden the diversity of contributions. The recent hiring of Galen Charlton and Joe Atzberger by Equinox should help, but it would be fantastic to see more contributions from other vendors and the community.

One should note that people like Warren Layton, Bill Ott, James Fournie, and others have contributed a number of patches for bug fixes and new features.

WANTED: skilled developers!

28

There are still a lack of developers able to contribute to Evergreen today.

There's no lack of developers interested in contributing, but it's hard to get a developer ramped up on Evergreen today. There's no good on-ramp, other than spending lots and lots of time with the code, on IRC, the lists...

Bill Erickson (from Equinox) and I gave a one-hour presentation at the Ontario Library Association SuperConference in February 2009 that covered a huge amount of ground very rapidly.

One hour is not enough! And the session wasn't recorded.

Building an on-ramp for developers

29

At the inaugural Evergreen International Conference, Mike Rylander gave an introduction to OpenSRF and Bill Erickson gave an expanded talk on developing with the new Dojo (JavaScript) Web interface widgets. I covered the current internationalization approach and challenges.

Better, but still not enough. And the sessions weren't recorded.

I'm going to be giving a four-hour workshop at the 2010 SuperConference, and I plan to develop a set of tutorial materials that (hopefully) others can work from solo / correct & extend. If at all possible, I'll also have the session recorded and make it freely available.

Communicate, communicate, communicate

30

It's not a failure to communicate, but it's been a challenge. Most of the Evergreen development team is inside Equinox, and it's a heck of a lot easier to talk things out face-to-face / during the course of regular meetings than it is to laboriously work things out on the mailing lists & document plans in the Trac instance and / or wiki.

We also use IRC pretty heavily, and although the logs are available and searchable via Google, it's not the prettiest interface. It also intimidates people who aren't used to IRC and who don't know who the people behind the nicknames are.

Communication is critically important, though. We need to avoid the duplication of effort in finding and resolving bugs – or in potentially working towards different implementations of the same feature.

Rethinking the release process

31

I fired a shot across the bow in early September suggesting a fairly radical reform of the way that Evergreen does releases. We've been doing feature-oriented releases for the major releases, and then bugfixes on top of those after the fact, with a rather complex a.b.c.d numbering scheme – more complex than most other software, in fact.

This is a lot of complexity to manage for a small development team, and it results in useful functionality sitting in a code repository waiting for months to see the light of day before a release comes out.

An early response suggests that the process is going to change radically - huzzah!

Better living through automation

32

The process for building a release is very laborious at the moment, and not particularly well documented.

Some light: There has been a recent move to commit database migration SQL each time the database schema changes, which will make the database migration script much less manual.

Equinox has committed to building a basic testing framework and automating the release process so that releases can be built on a nightly basis. The automated tests will be run on the release and any errors will be flagged.

I look forward to tests

33

Okay, I still have nightmares about having skipped classes for an entire semester and then showing up for the final exam, completely unprepared and expecting humiliation.

But software without tests is like a diet of junk food without healthy choices: you can live without them, but you have to be prepared to live with the unhealthy consequences.

Tests provide another onramp for would-be developers; they're a way to learn what is expected from the software, and a comparatively easy way to contribute to the project.

Adding a test for every bug that is fixed makes the project more robust.

Shorter release cycles

34

Once we have daily releases with integrated database migration scripts, it becomes much simpler for us to anoint a given day's release as the next official release.

If we do a good job of tracking bug fixes and enhancements to our issue tracker, then release notes become a simpler task and the documentation team has more to work with to write new / enhanced documentation.

Documentation

35

Once upon a time, there was a technical writer who wrote application development manuals. He also contributed to the Linux Documentation Project, and PHP documentation, and co-wrote a book about a database that almost nobody used that sold very poorly.

One day, he joined a software project that was in desperate need of documentation. "I'm going to write a manual for this software!" he told himself; he just needed to learn the software and write about it.

But as he learned the software, he learned that it had bugs, or needed features, and he started fixing those bugs and adding those features. And apart from whipping together a proof-of-concept table of contents and a very basic style manual, he did nothing more. Enter the DIG!

A solid foundation

36

At the Evergreen conference in May, 2009, there were discussions about forming an Evergreen Software Foundation for various (not necessarily a cohesive set of!) goals.

Goals (from my perspective) should include:

- · Handling trademarks and collateral
- 501(c)(3) for US tax-deductible donations directed towards Evergreen community development
- Organize Evergreen conferences

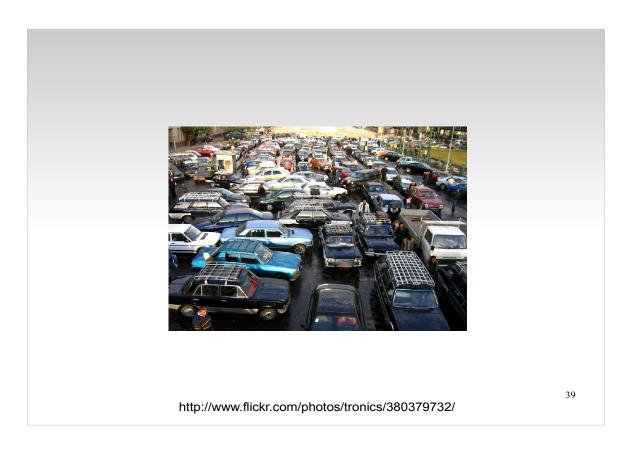
There is progress: GPLS, the current holder of the Evergreen trademarks, recently announced that it is forming a foundation in compliance with the State of Georgia regulations. They are involving the community in the birth of that foundation.

# Please bear with us...

37



I live in Northern Ontario, sure, but not that far North.



No, not this... although, sometimes development traffic jams do occur



This captures the spirit of working together to achieve a goal



Of course, we're developers and we're lazy (Roman aqueduct), so we'll spend a ton of effort building something so that we never have to do it manually again.



Once you've worked out the basic principles, it's just a matter of evolving the design. Witness storage -> cstore, permacrud -> pcrud -> PermaCrud...

## **More information**

- Evergreen home: http://evergreen-ils.org
- Blog aggregator: http://planet.evergreen-ils.org
- Mailing lists: http://evergreen-ils.org/listserv.php
- IRC: #evergreen on chat.freenode.net

43