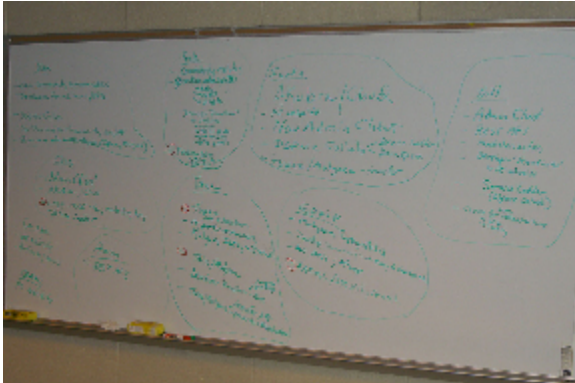
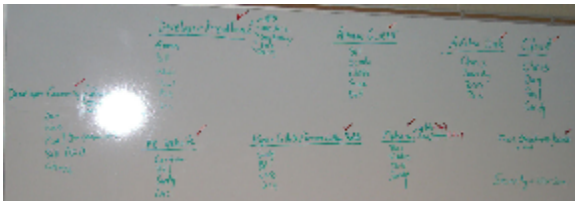


On-Site Nov 3-7 Meeting Notes



Discussion Interests by attendee

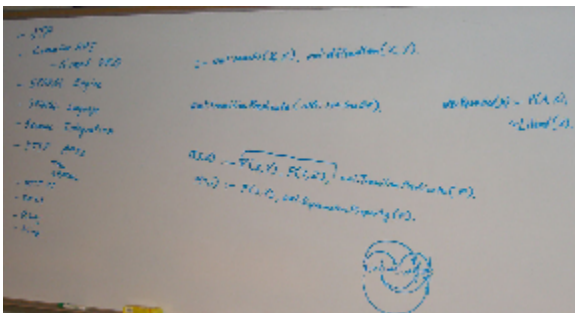


Final Discussion Topics

Bill - FedoraShare

- Missing - Call disseminations using REST API
- Web 2.0 - Add your content easily.
- Relate the stuff from multiple sites.
- What is the user experience?
- Eddie - Goal is Fedora API improvements or FedoraShare application?
- Inform, Include, Provide easy simple application (Sandy - note Timeline from Simile)
- YouTube as a dissemination. (example dissemination)
- Google has a different OpenID impl.
- With FedoraShare HTML page with YouTube embed tag.
- Eddie - Yahoo Pipes? Rick - Investigate Kepler.
- Simple back-end.
- Users as first class objects.
- Messaging does API-A.
- Dan - Rebuild slider in FedoraShare.
- Triple benefits.
- Replication into multiple sources.
- Flex is well supported.
- Low impact frontend.
- Dan - Local App/Server
- Paul - Apple Cappucino.org written in Javascript as alternative
- Developers - My shop does not support?
- Reusable code?
- Separable components.

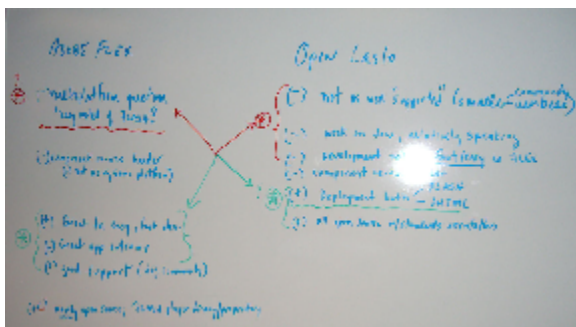
Paul - Mulgara



- JTA (Andre)
 - Handle all transactions
 - One write at a time - need to move to multiple writers
- Connection API
 - Graph URIs (permits arbitrary URIs - does not identify host)
 - Rmi://host.com/server1?graph=foo:bar as opposed to # (fragment ID)
- SPARQL Engine
- SPARQL Language
 - Focused on join optimizations rather than filtering which is needed for complex SPARQL queries
 - Meeting W3C SPARQL spec especially what goes in must come back exactly (cannot be canonicalized)
- Sesame Integration
 - James at Sesame giving good feedback
 - Sesame is 100% SPARQL but not Mulgara. SPARQL has distinct clause which causes problems with set semantics. Not easily fixed. No support for Regular Expression (different from Java). Order-by arbitrary expression support.
- HTTP APIs
 - TQL (iTQL name going away)
 - SPARQL
 - Using Servlets
 - Throw exception if "insert" into Get operation
 - Post expects an RDF file
 - May be output in JSON
- HTTP UI
 - Cleaned up no longer implemented in Barracuda
 - Tutorial mentions graph URI: (not needed for real) Select Query: Query: Submit Button
- XA 1.1
 - Scalability
 - Multiple writers
 - No problem with recycling Identifiers on 64bit systems
 - 192 to 32 files for storage
 - Better block writes
 - Get test info from Ronald/Paul for presentations
 - Change term from Model to Graph (W3C)
 - Can index on all three orders SPO, POS, OSP
 - Can index SPOG, POSG, OSPG, GSPO, GPOS, GOSP mostly these
- RLog - RDF Prolog
 - hasUncle(X, Z) :- hasFather(X, Y), hasBrother(Y, Z). head :- body
 - (X Y Z) y(X, Z) [foo:bar rdf:type foot:person]
 - Foo:Person(foo:bar).
 - Supports transitivity, symmetry, will support consistency
 - User transitivity to implement a resolver
- JLine - Library for Java but must use native code to do it. Supports a lot of platforms. Andy wrote

Bill - Admin Client

- Minimal Version
- Public Comment
- Alpha Distribution
- Keep the old version for at least one iteration



- Flex vs. Laszlo for implementation

Eddie - funAPI

- Make Fedora Zotero friendly
- Easiest to use OAI-PMH

Eddie - AtomPub

- Potential input servlet
- Simple get/post to URL
- How do we want to set all the other properties on the object

- Separate from SWORD, they may do the Fedora implementation themselves

Eddie - Zotero as Client of Fedora

- Uses WebDav, Raymond, does he think they will support Atom PP.
- Do we want to rethink supporting WebDav?

Matt - REST API

- Look at same platforms, MS Silverlight too small installed base, using Flex and ROR and simple HTML.
- Cannot get full management info on datastreams
- Timestamp checked for lazy lock
- Cannot do relationships
- No getDatastreams
- Setting the content model
- Indexing (e.g Search Engine)
- Query by Relation (RISearch)
- Find by Content Model hard to do
- Find by Relationship
- Directly support the DC datasets
- Find by DC Terms
- Find by a Specific Property or Value
- Leverage off OTM finding objects query language

Aaron - NSDL Implementation in Fedora 3.0

- In NSDL Wiki
- Move to FR30 to reduce pain points.
- OAIprovider does not have an understanding of NSDL Information model
 - Makes it difficult to find related items
 - Views are inconsistent
- Resource -> Aggregator -> Aggregator path can be traversed, queried, or be a service itself
- Eddie - Repository should know its own structure without depending on client
- Chris - Also like an inferencing problem
- Both Content Model and SDef represent kinds of "typeiness"
- For access this is easier to see but what about create/ingest/write
- Could we generate Client code?
- What REST style is most effective for Fedora?
- [Thought Experiment](#)

Rick - Plone & CMS

- Plone - CMS in Python
- Collections are captured searches
- Links (except the built-in generic isLinkedTo) are user defined strings
- Content Modeler widely used.
- Active content display server

Paul - Mulgara Roadmap

- Rules engine has been around. Compile Horn Clauses to rules.
- Has uptake.
- Distributed queries but need improvement. Can improve performance especially network.
- Cloud optimizations, wants help
- Clustering, multi disk optimizations (RAID 1, configuration)
- Hadoop
- Loading RDF is still the bottleneck but clustering is the coming problem
- SPARQL update, modify missing (delete then insert), no transactions
- Virtual graphs through resolvers, Mulgara does not have filters
- Rules, too course-grained for right now
- Consistency Checking
- RLog - needs to be more like a scripting engine, move to move parser then impl as resolver
- RETE engine for dynamic updates (most expensive)
- XA 1.1
- Security

Development Community, Processes and Infrastructure

- Chris - covered how we are provisioning projects
- Need HA solution
- Keep delivery on SourceForge
- Paul Google Groups has some bad perceptions.

- Backups of Wiki, Jira
- Rsynch SVN.
- Eddie - Mercurial
- Rick Plone - Collective
- Paul - Google Code
- Dan - Developers Pages and Experience, Wiki, CMS
- Dan - Developers Forums
- Eddie - Using Maven
- Chris - Still want a source release.

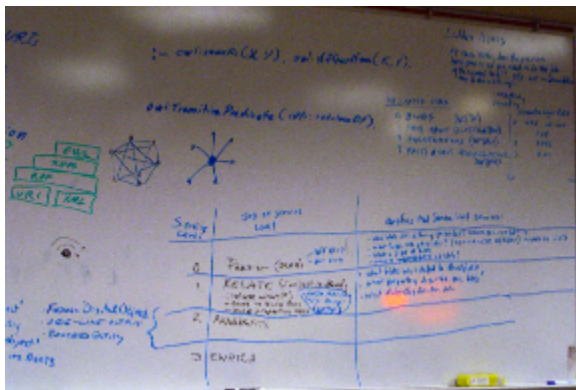
Chris - Akubra

- Eddie - Is Honeycomb dead. Chris - It represents XAM compliant storage.
- Eddie - Why is "quiesce" function required
- Sandy - Concerned with so little metadata. Chris - Many implementations with varying capabilities but other example may not be so concerned with disaster recovery.
- Sandy - Consider "last-modified-date" as required read-only metadata

Cloud Storage

- Bill - Summarized Amazon EC2/EBS/S3 experiments to date
- Reviewed - Peter Murray article on OCLC vs Amazon S3 Storage
- Reviewed Amazon SLA
- Eddie - What about Fedora Clusters (or Fedora Federation)

Chris - Ladder



Ladder Issues

1. At each level, does the previous level provide all you need to do the job of the current level? Yes, but the implementation may do this in different ways.
- 0 - Persist (blob)
 - was the blob successfully persisted?
 - what blobs are persisted? (enumerated maybe with filter)
 - what is the size of the blob?
 - what is the modified date of the blob?
 - give me the content of the blob?
 - get blob
 - put blob
 - 1 - Relate (Subject is blob)
 - (relate what?) facts about the blob
 - blob to blob
 - blog properties
 - what blogs are related to this blob
 - what properties describe this blob
 - (get/put)
 - 2 - Aggregate
 - Kahn Wilensky -> Digital Object -> Fedora Digital Objects -> ORE-Like Output
 - Compound object
 - 3 - Enrich

Symmetry with Semantic Web Stack?

Get/Put

- blobs - URI Unicode
- Facts about blobs - RDF

- Aggregations - RDFS
- Facts about Aggregations - OWL
- 3 OWL
- 2 RDFS
- 1 RDF
- 0 URI XML

Developers, Project and Product Website



- Deploy Plone
- Create a simple entry point
- Create a registry of all related items

Additional Friendliness

Installer

- Platform Specific
 - Windows
 - deb, rpm, sun
 - fink, dmg
 - Java-Only
 - Web-based

Scope

- Control Scripts (start/stop) - "Jump" to Start/Config
- JRE Decision
- Uninstaller
- "Start Menu"
- Database - Derby

Project

- core-rep -> war
- generic-installer -> jar
- windows-installer -> ^

Wrap Up Actions

- Finish Challenges end of December,
- Continue Conceptualizing FedoraShare, Prototype Deliverable
- Continue with NSDL on Fedora 3.0 Conceptualization, Design/Report Deliverable
- Admin GUI Prototype (Test REST API, Flex)
- REST API Analysis and prototype some operations, findBy
- Zotero Integration Prototype
- unAPI and APP Prototype
- ORE Prototype (with Collaborators) (First Quarter 09)
- Test Maven to decide to do wholesale conversion
- Due diligence on hot server plug-in framework technology
- Investigate Mercurial
- Backup of Topaz site but leave resources alone, Find out hosting.
- Do we integrate Mulgara more closely? What are the boundaries?
- Field search (see REST findBy enabled by Triplestore)
- Akubra next steps, Implementation Deliverable
- Formalize provisioning for projects
- Choose a test candidate to "modularize"
- Install a production Plone and move the Website. High Priority
- Redesign "Developer" Web presence High Priority

- Plone-Fedora Interface
- Integrated testing framework
- Do we want to rethink supporting WebDav? (with Zotero integration provides great integration opportunity)
- Put Fedora Clusters and Fedora Federations in the Roadmap, Mulgara Clusters, EC2 Mid 2009.
- Mulgara Roadmap Update, Express Priorities (SPARQL Update very high)
- Technology Transfer Mulgara, Topaz, and Fedora
- Rick will build a Windows installer for Fedora Repository.
- Security Fedora, Muradora and Mulgara (SPARQL Interceptor, Gert Query rewriting)
- A lot of modest user-friendliness features
- Build a wish list of potential projects and features to develop, rationalize and trim list for trackers/roadmap
- Plan the next meeting