



JCR in Action

Content-based Applications with Apache Jackrabbit

Carsten Ziegeler
chiegeler@apache.org

Apache Con NA Presentation – November 2010 - Atlanta

About



- Member of the ASF
 - Sling, Felix, Portals, Sanselan, Excalibur, Incubator (Cocoon)
 - PMC: Felix, Portals, Sling, Incubator, Excalibur (Chair)
- RnD Team at Adobe (Day Software)
- Article/Book Author, Technical Reviewer
- JSR 286 Spec Group (Portlet API 2.0)



Motivation

- Tried and trusted **NOSQL** solution
- Standard Java API
 - First spec released in May 2005
 - Various implementations, products, and solutions
- Open Source Implementation
 - 1.0 Release in April 2006
- Think about your data storage use cases – JCR might help!



Agenda

- **JCR and Apache Jackrabbit**
- Basic Content Modeling
- References and Search
- Advanced Features
- Demo: Sample Application
- Summary and Questions



(Nearly) Everthing is Content

- Application Domain Specific Content
- Presentation Support (HTML, CSS, JavaScript, Images)
- Documentation, Translations
- ...



Content Silos without JCR

- Structured content
 - Usually schema based
- Unstructured content
- Large data (images, movies etc.)
- Different storage for each use case



Content Repository



- Single storage for all use cases!



Content Storage

- Hierarchical content
 - Nodes and properties
- Structured
 - Nodetypes and typed properties
- And/or unstructured
- Fine and coarse-grained
- **Single repository for all content!**

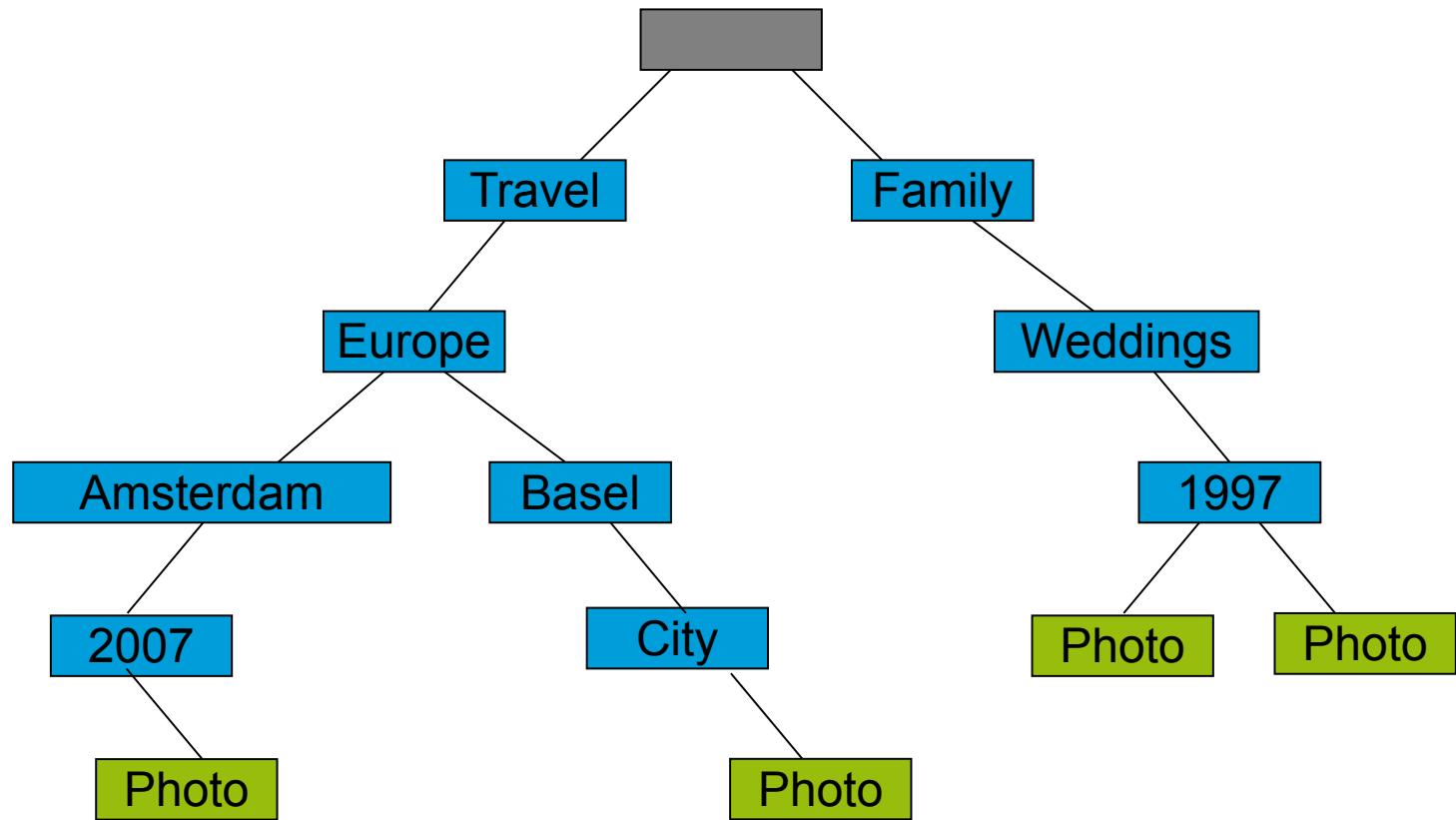


Sample Application: Slingshot

- Digital Asset Management
 - Hierarchical storage of pictures
 - Upload
 - Tagging Poor man's flickr...
 - Searching
 - Automatic thumbnail generation
- Sample application from Apache Sling



Slingshot Content Structure



Facts About Slingshot

- Java web application
- Uses Apache Sling as web framework
- Content repository managed by Apache Jackrabbit
- Interaction through the JCR API



JSR 170: Content Repository for Java™ technology API

- (Java) Standard – Version 1.0
 - Supported by many vendors
 - Used by many products and projects
 - Several open source solutions
- How do you connect to a CR?
- How do you interact with a CR?



JSR 283 : JCR 2.0 is final

- New features
- Improved specification
 - Cleaned up API (deprecation)
 - Revised specification document
- Binary compatible
 - JSR 170 apps run without modification
- More later...



CR : Combines Advantages



Content Repository Features

- Query (XPath, SQL)
- Export/Import (XML)
- Referential Integrity
- Authentication
- Access Control
- Versioning
- Observation
- Locking and Transactions (JTA)

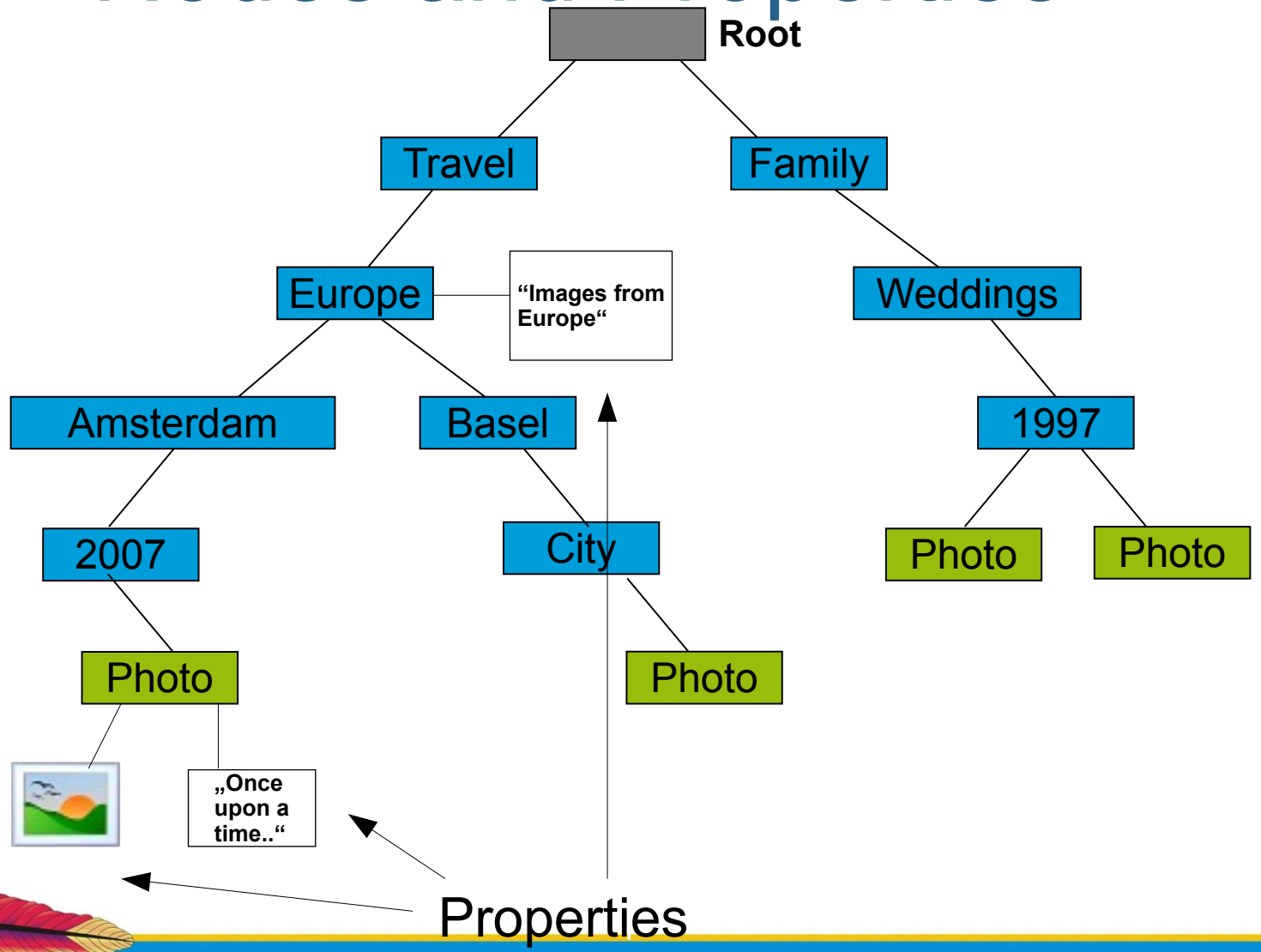


The Repository Model

- Repository: one (or more) workspaces
- Workspace contains a tree of items
- Item: Node or property
- Nodes provide the content structure
 - May have children
- Actual data is stored as values of properties
- Types and namespaces!



Nodes and Properties



Connecting to the Repository

- JCR 2.0 provides RepositoryFactory
- Uses Service Provider Mechanism
 - *META-INF/services/javax.jcr.RepositoryFactory*
- Just use
 - *RepositoryFactory.getRepository(null)*
- Or specify connection parameters
 - *RepositoryFactory.getRepository(Map)*



Working with the Repository

- Interaction is session based
 - Assemble credentials
 - Login into workspace

```
Credentials myCredentials =  
    new SimpleCredentials("USERID",  
                          "PASSWORD".toCharArray());  
Session mySession =  
    repository.login(myCredentials, "WORKSPACE");
```



Traversing the Content

- Traverse the repository
 - From the root or any node

```
Node rootNode = mySession.getRootNode();  
Node albumNode =  
    rootNode.getNode("slingshot/albums/travel");  
Node europeNode = albumNode.getNode("Europe");
```



Retrieve a Property

- Various ways to get a property
 - Different methods for each type
 - Type conversion if possible

```
Property prop =  
    albumNode.getProperty("slingshot:description");  
Value value = prop.getValue();  
String desc = value.getString();
```

```
value.getBoolean();  
value.getStream();  
value.getLong();  
value.getDate();  
value.getDouble();
```



Changing Content

- Change everything you want
 - Add/Remove nodes
 - Add/Remove/Change properties
 - Transient space
 - Then save

```
// change
albumNode.addNode("newAlbum");
europeNode.setProperty("slingshot:description",
                        "something");

// save...
mySession.save();
// ... or revert all changes
mySession.refresh(false);
```

Interaction Summary

- Get the repository
- Login to a workspace
 - Provides a session
- Use the session to
 - Access nodes and their properties
 - Change and save them



Apache Jackrabbit

- JSR 170 and 283 reference implementation
- Apache TLP since 2006
- Releases
 - 1.6 (JSR 170 based)
 - 2.1.2 (JSR 283)
- Components
 - Commons, API
 - RMI, WebDAV, WebApp, JCA
 - OCM
 - And more...



The screenshot shows the Apache Jackrabbit website. At the top, there is a navigation bar with the Apache Jackrabbit logo on the left and 'The Apache Software Foundation' logo on the right. Below the navigation bar, the main content area is titled 'Welcome to Apache Jackrabbit'. On the left side of the main content, there is a 'Download' button with a rabbit icon and the text 'Download Apache Jackrabbit 1.6.0'. To the right of the main content, there is a sidebar with a list of links including 'Welcome', 'Downloads', 'FAQ', 'Jackrabbit Wiki', 'Jackrabbit History', 'Documentation', 'Getting Started', 'Standalone Server', 'Jackrabbit Components', 'First Hops', 'JCR & API', 'Jackrabbit Architecture', 'Deployment Models', 'Jackrabbit Configuration', 'Node Types', and 'Object Content Mapping'. Below the sidebar, there is a section for 'Apache Jackrabbit News' with three news items: 'October 30, 2009: Apache Jackrabbit 2.0 beta1 released', 'September 23, 2009: jackrabbit-core 1.4.11 released', and 'September 23, 2009: Apache Jackrabbit 2.0 alpha11 released'. At the bottom right of the sidebar, there is a section for 'Upcoming Events' with a link to 'ApacheCon US 2009' on November 2-6 in Oakland, CA.

<http://jackrabbit.apache.org/>



Words of Advice

- Read (or browse) the JCR specification
 - `jcr-1.0.jar/jcr-2.0.jar` included
- Getting started with Jackrabbit
 - `jackrabbit-webapp`: Drop-in deployment
 - First Hops: Embedded repository
 - Take your time
 - Think about your use cases
- Have a look at Apache Sling



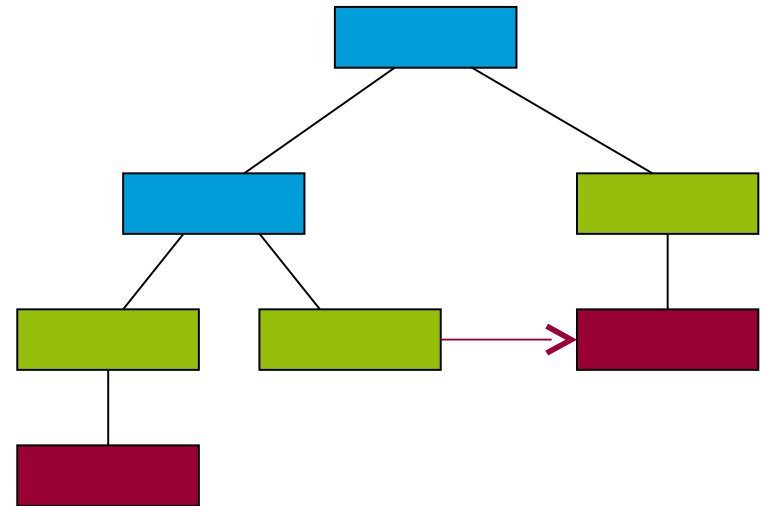
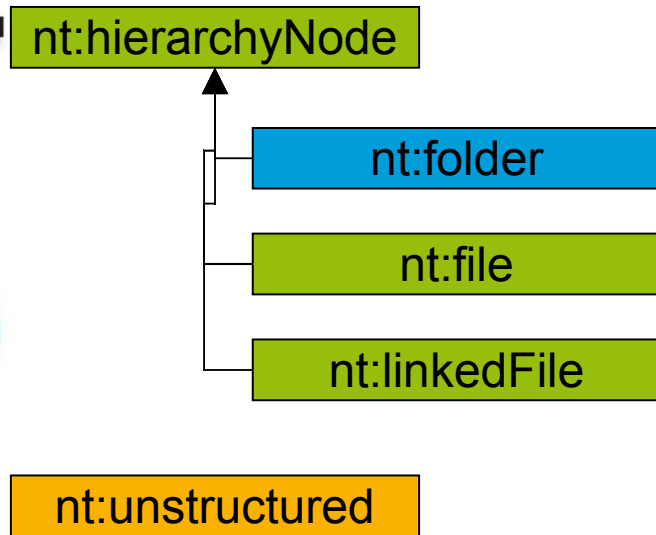
Agenda

- JCR and Apache Jackrabbit
- **Basic Content Modeling**
- References and Search
- Advanced Features
- Demo: Sample Application
- Summary and Questions



Starting point: Leverage the standard node types

- Type hierarchy
- Content hierarchy



Bottom-up modeling: Node types

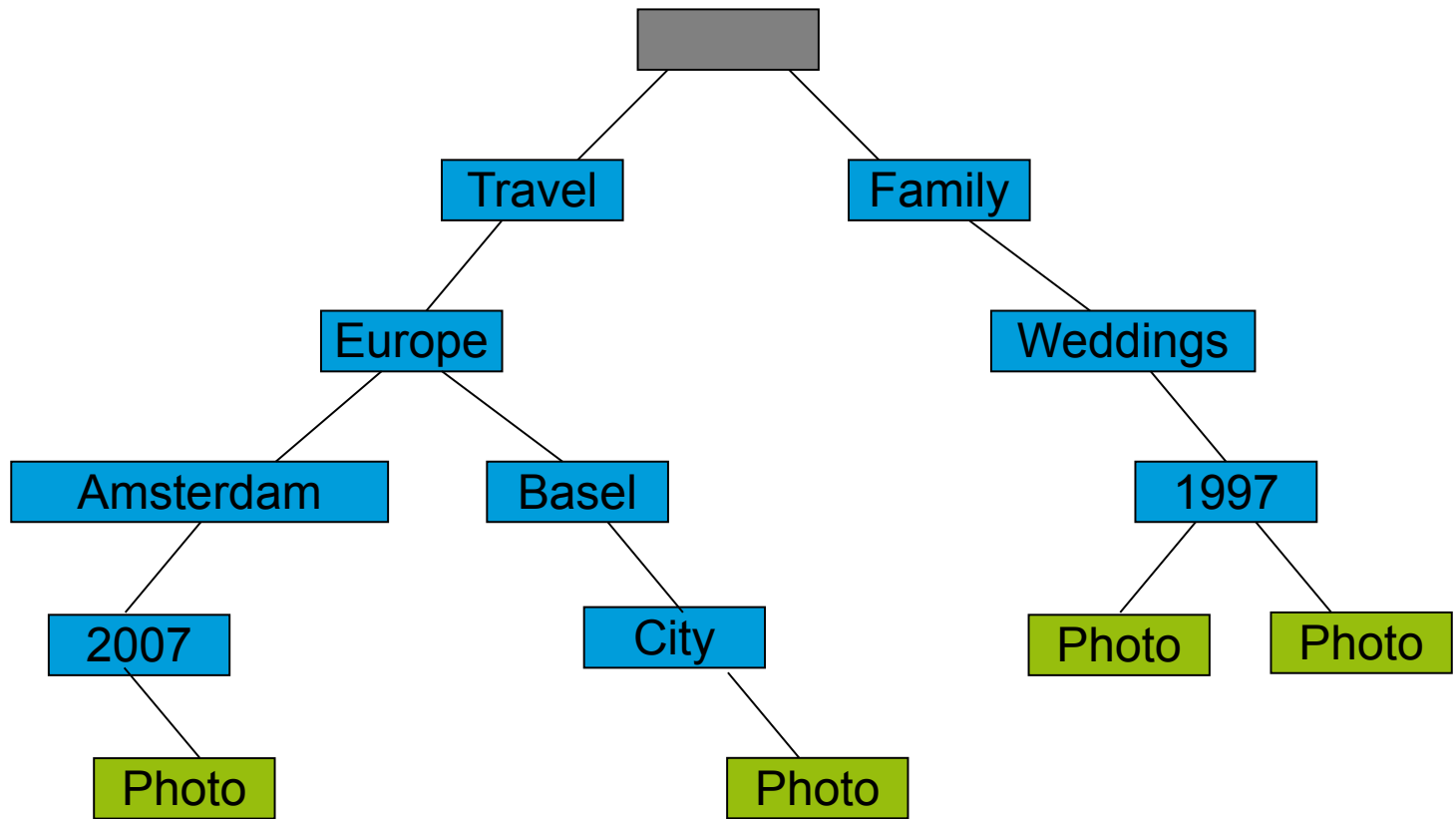
```
slingshot:album > nt:folder  
- slingshot:description (string)  
- slingshot:date (date)
```

```
slingshot:photo > nt:file  
- slingshot:description (string)  
- slingshot:location (string)  
- slingshot:tags (string[])
```

```
slingshot:tag  
- slingshot:description (string)
```



Top-down modeling: Content hierarchies



Namespaces

- Namespaces can be used for
 - Node types
 - Node and property names
- Single namespace per company or app
- Reasonably unique namespace prefix
- Prefixed names for structured content
- Default namespace for unstructured content



Content Modeling: Words of advice

- Use an application root node
 - /my:content or slingshot
 - Good for searching, backup, and migration
- Avoid flat hierarchies
 - User interface complexity
- Content-driven design
 - Design your content before your application



Content Modeling: Words of advice

- Look at existing node types (JSR 2.0)
- Make use of mixin node types
- Checkout Apache Jackrabbit wiki and mailing lists
 - "Davids Model"



David's Model

- Rule #1: Data First, Structure Later. Maybe
- Rule #2: Drive the content hierarchy, don't let it happen
- Rule #6: Files are Files are Files
- Look at [http://wiki.apache.org/jackrabbit/Davids Model](http://wiki.apache.org/jackrabbit/DavidsModel)

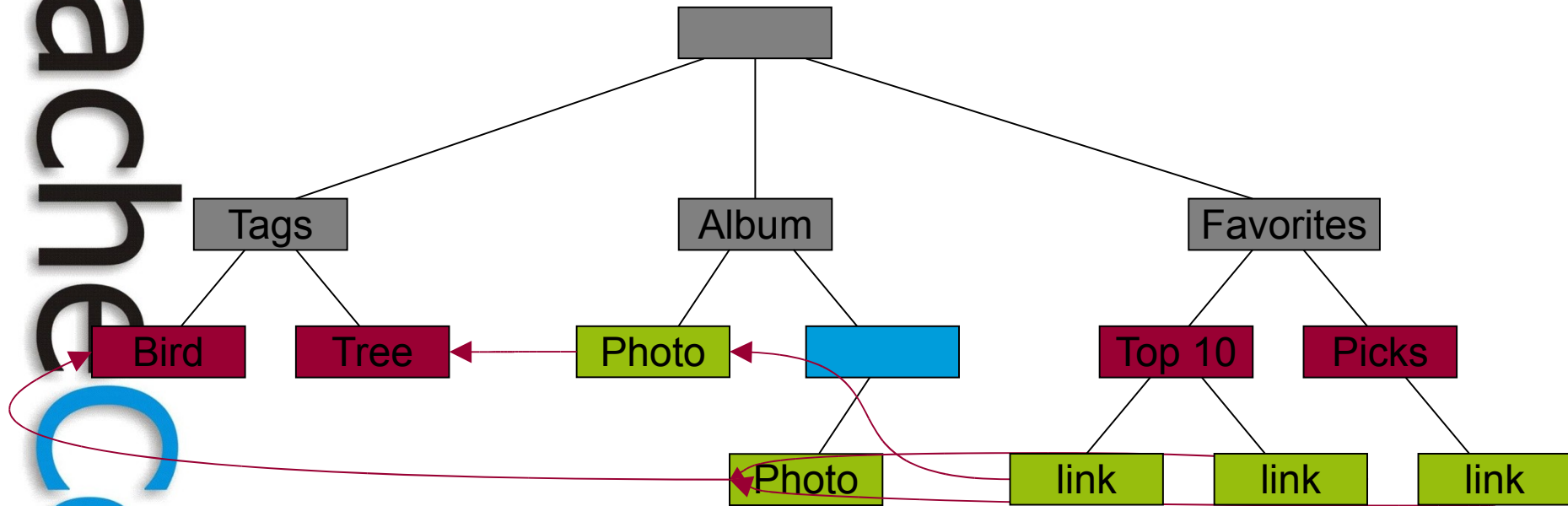


Agenda

- JCR and Apache Jackrabbit
- Basic Content Modeling
- **References and Search**
- Advanced Features
- Demo: Sample Application
- Summary and Questions



References



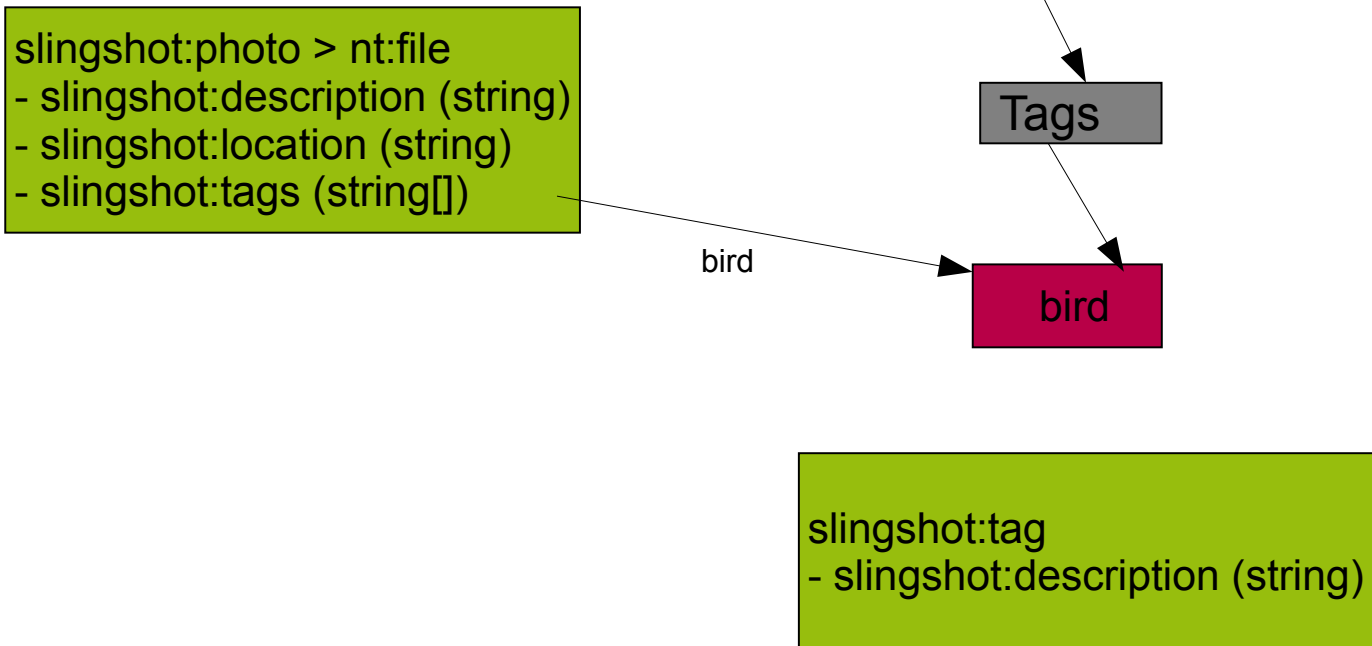
API:

`Node.getReferences():PropertyIterator`
`Property.getNode():Node`
`Node.setProperty(String name, Node)`



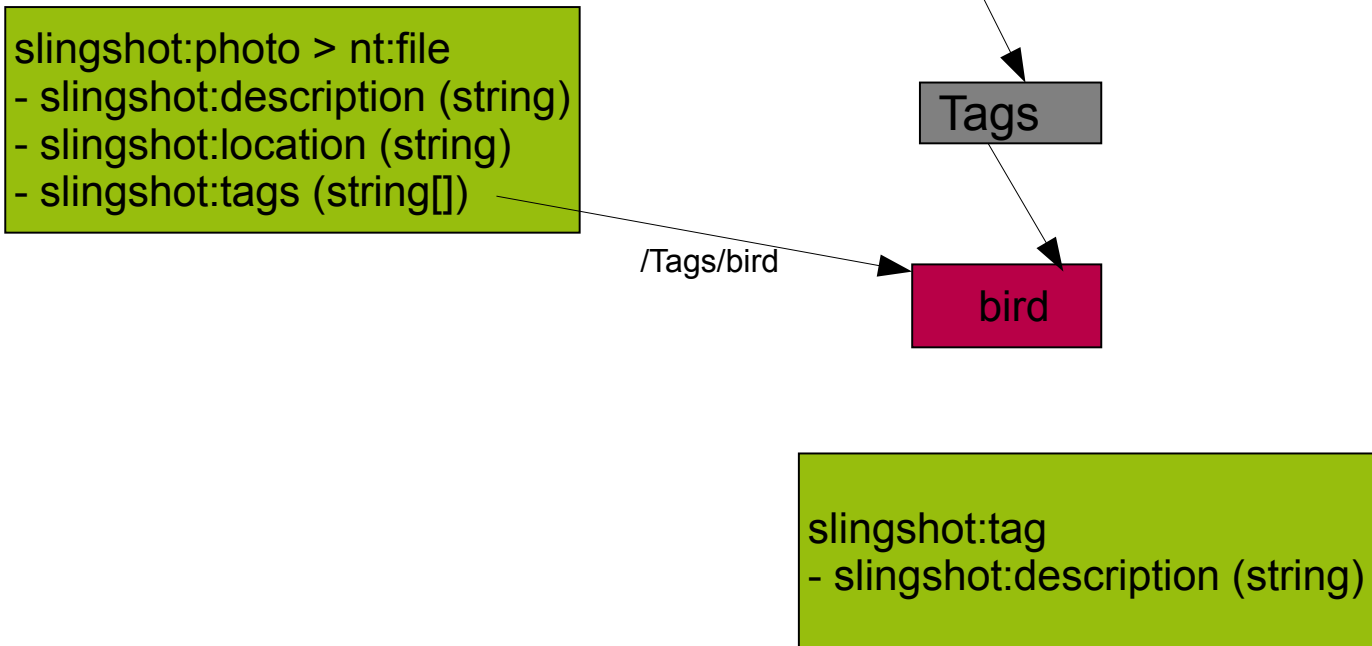
Alternative References I

- Reference by name



Alternative References II

- Reference by path



Searching

- Query API
 - Java Query Object Model
 - SQL queries
- Examples
 - `SELECT * FROM slingshot:photo WHERE jcr:path LIKE '/slingshot/%'`
 - `..AND slingshot:description CONTAINS 'atlanta'`

API:

```
Session.getWorkspace().getQueryManager():QueryManager  
QueryManager.createQuery(String stmt, String  
language):Query;  
Query.execute():QueryResult
```

Alternative Views: Words of advice

- Moderate use of references
 - Circular references only within a subtree
 - Plan for backup and content migration
- Best search performance when selecting a small subset of content
- References, path or name property



Agenda

- JCR and Apache Jackrabbit
- Basic Content Modeling
- References and Search
- **Advanced Features**
- Demo: Sample Application
- Summary and Questions



Observation

- Enables applications to register interest in events
- Monitor changes
- Act on changes

API:

ObservationManager:

```
addEventListener(EventListener listener,  
    int eventTypes,  
    java.lang.String absPath, boolean isDeep,  
    java.lang.String[] uuid,  
    java.lang.String[] nodeTypeName, boolean  
noLocal)
```

Observation Events

- Describe changes to a workspace
 - Dispatched on persistent change
- Provide the path of the item
- Provide the user ID
- Only provided to sessions with sufficient access privileges
- Events may not be complete!
 - Example: removal of a tree of nodes



Event Types

- Six different events
 - Node added
 - Node removed
 - Property added
 - Property removed
 - Property changed
 - Node moved (JCR 2.0)



Event Listeners

- Registered with a workspace
- Registration with optional filters
 - Like node types, paths
- Receive events for every change
 - Set of changes

API:

```
public void onEvent(EventIterator events);
```



Events Advice

- Events occur after a save
 - Modification based on events is a new operation
 - Such events can be filtered
- Events may not tell the complete story
 - Tree removal, copy (, move)
- Most common pattern
 - Node added, removed, changed



Versioning

- Nodes can be versioned
 - Together with their child nodes
- Version history
- Getting a specific version
- Restoring a version



JCR 2.0 Features

- Abstract Query Model
 - Java Query Object Model / SQL
- ACL and Access Control Policies
 - Plus Retention Policies and Hold
- Nodetype Registry
- New Property Types and Nodetypes
- Shareable Nodes (Graph)
- Journalling Observation



Agenda

- JCR and Apache Jackrabbit
- Basic Content Modeling
- References and Search
- Advanced Features
- **Demo: Sample Application**
- Summary and Questions



Getting Content

- Get the repository
- Login to a workspace
 - Returns a session
- Starting with the session
 - Navigate to the desired content
 - Get the query manager and search



Writing Content

- Get the repository
- Login to a workspace
 - Returns a session
- Navigate to the correct node
 - To change it
 - To add new nodes/properties
 - To remove nodes/properties
- Persist changes



Advanced Development

- Apache Sling
 - REST based web framework
 - Powered by OSGi
 - Scripting Inside
- Apache Jackrabbit OCM
 - Map content to Java objects and vice versa
 - Similar to database ORMs



Agenda

- JCR and Apache Jackrabbit
- Basic Content Modeling
- References and Search
- Advanced Features
- Demo: Sample Application
- **Summary and Questions**



And remember...

- (Nearly) everything is content
 - Application Domain Specific Content
 - Presentation Support (HTML, CSS, JavaScript, Images)
 - Documentation, Translations
 - ...
- With observation, versioning, search...



Conclusion

- Content Repositories
 - Combine advantages from FS and DB
 - Add important features
 - Structure/Access your data the way your domain requires it
 - Single repository for all your content!
- JCR – The Java API
- Apache Jackrabbit – The Implementation



Famous Last Words

- Read the specification
- JCR in your application?
- Join the Jackrabbit community!
- Seriously consider **Apache Sling** for web applications
- Check for additional stuff like OCM



ApacheCon



Q&A