

**adaptTo()**

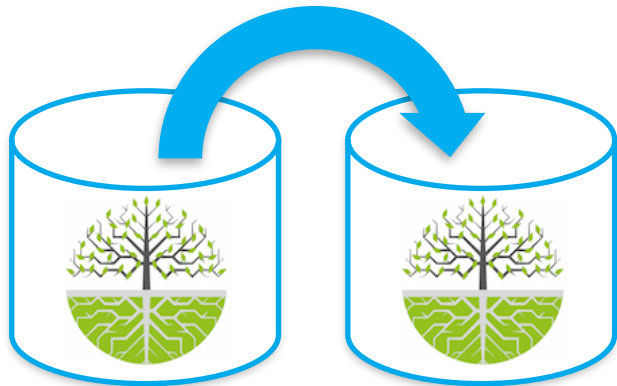
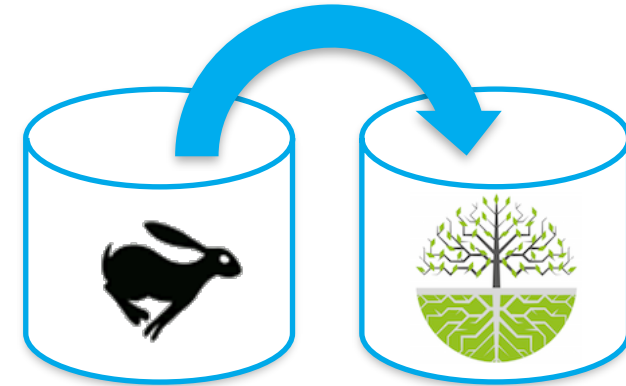
APACHE SLING & FRIENDS TECH MEETUP  
BERLIN, 28-30 SEPTEMBER 2015

# Oak Upgrade

Julian Sedding

# Scenarios

**Upgrade: Jackrabbit 2 to Oak**  
(aka crx2oak)



**Copy: Oak to Oak**



# Objectives

- **Be fast**
  - Copy only differences
  - Move processing off the critical path
- **Be flexible**
  - Copy only what you want
  - Consolidate content from multiple repositories



# Phases of an Upgrade

- Content copy phase
  - Workspace copy
  - Version copy (including orphaned versions)
- Post-processing phase
  - Index-generation
  - Version copy (except orphaned versions)
  - Validation

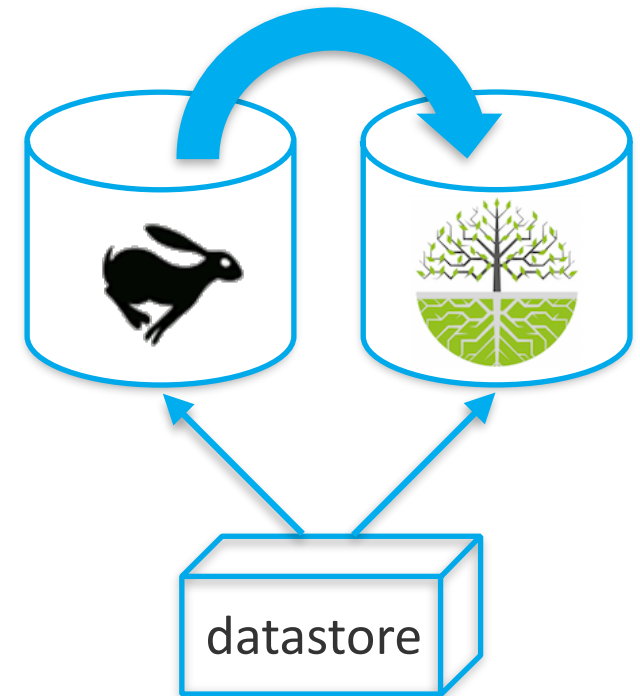


# Phases of an Upgrade - Characteristics

- Content copy phase
  - Typically fast – I/O bound
  - Single full traversal of copied (sub-)trees
- Post-processing phase
  - Typically slow – I/O + CPU bound
  - Depends on the amount of content copied

# Keeping the JR2 Datastore

- Oak supports JR2 datastores
  - via DataStoreBlobStore
- Shared datastore for upgrade
  - Copy binaries by reference
  - Compare binaries by reference

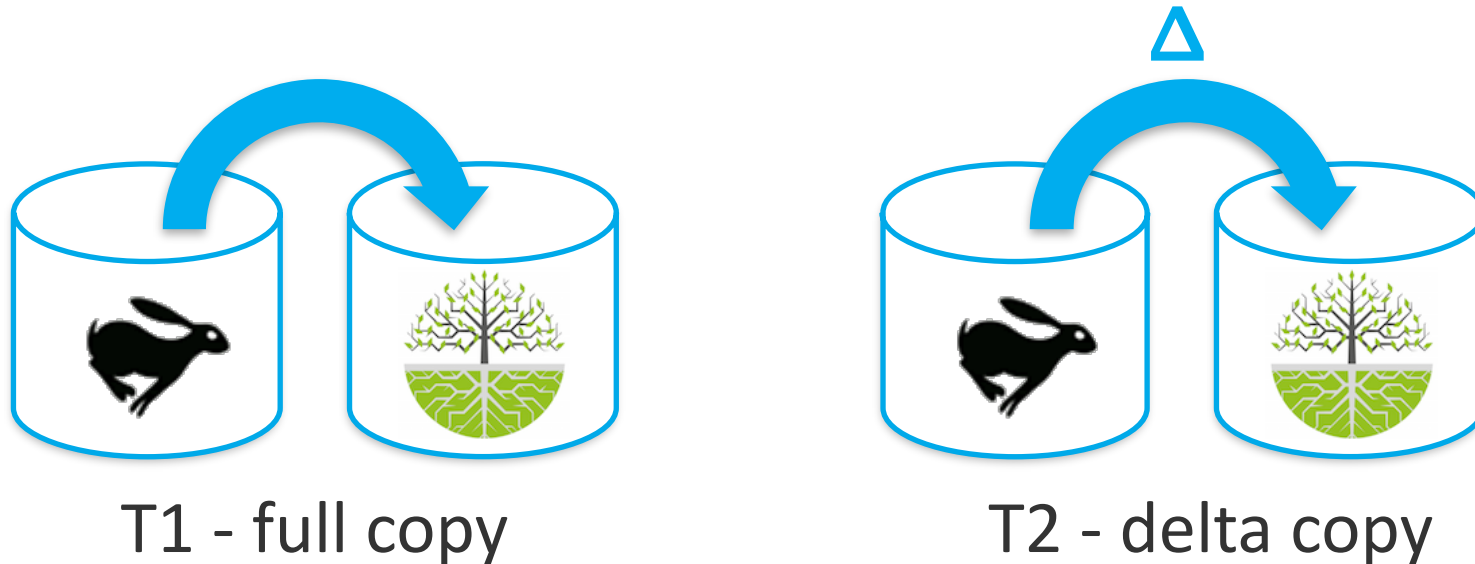




# Include/Exclude – Copy less!

- Limit scope of copy by paths
  - Include paths (e.g. /content, /etc/tags)
  - Exclude paths (e.g. /var/audit, /var/classes)
  - Merge paths (advanced use-cases)
    - Add content while preserving existing content
    - Existing nodes are never deleted

# Repeated Upgrade



- full traversal + comparison for delta copy
- post-processing only needed for differences
- can speed up critical path!



# Version Histories

- **Keep version histories?**
  - Skip all if you don't need them
  - Skip version histories older than yyyy-mm-dd
- **Keep orphaned version histories?**
  - Used by the tree restore feature in AEM
  - Skip all if not needed or skip old orphaned version histories

→ non-orphan versions copied during post-processing



# Command Line Tools

- **oak-upgrade.jar (formerly oak-run.jar)**
  - see <https://jackrabbit.apache.org/oak/docs/migration.html>
  - src <https://github.com/apache/jackrabbit-oak/tree/trunk/oak-upgrade>
- **crx2oak**
  - see <http://docs.adobe.com/docs/en/aem/6-1/deploy/upgrade.html>



## Want More?

- Pre-extract text from binaries to speed up indexing
  - see OAK-2892
- Online BlobStore migration
  - see OAK-3148
- Avoid FS access – LengthCachingDatastore
  - Configured in repository.xml, see OAK-2882



# The End

---

# Questions?