

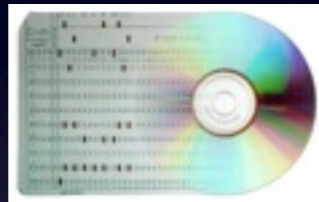
NEON

TF-Storage, March 5Th 2010, Utrecht
Maarten Koopmans for UNINETT Sigma
maarten@vrijheid.net

Goals

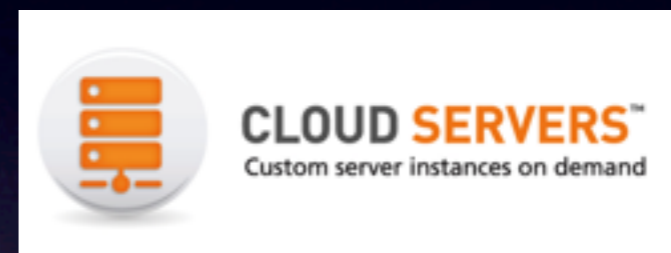
- **state-of-the art** of cloud computing;
- **cost** of moving and running **non-HPC** jobs on a cloud computing environment;
- how to do this in **practice**;
- a list of **identified risks/benefits** on a short/long perspective.

Areas



...and management

Shortlists





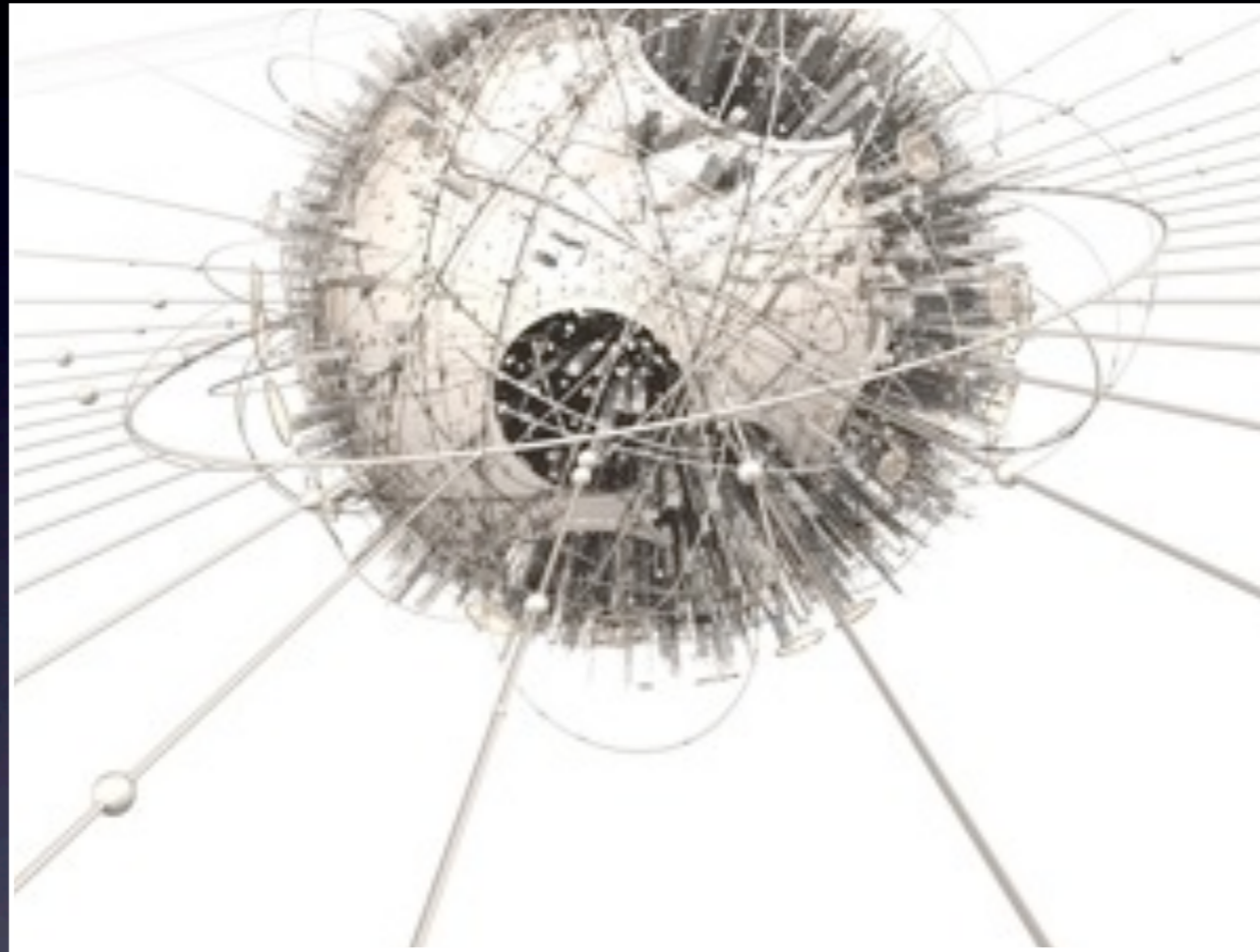
Why again?

Let's zoom in on storage

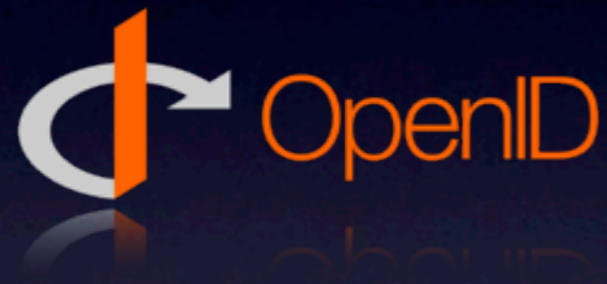
Storage: requirements

- No client to access the data.
- Transparent versioning.
- Transparent encryption - both transport and storage
- AAI integration
- Allow sharing of resources

Current Cloud storage



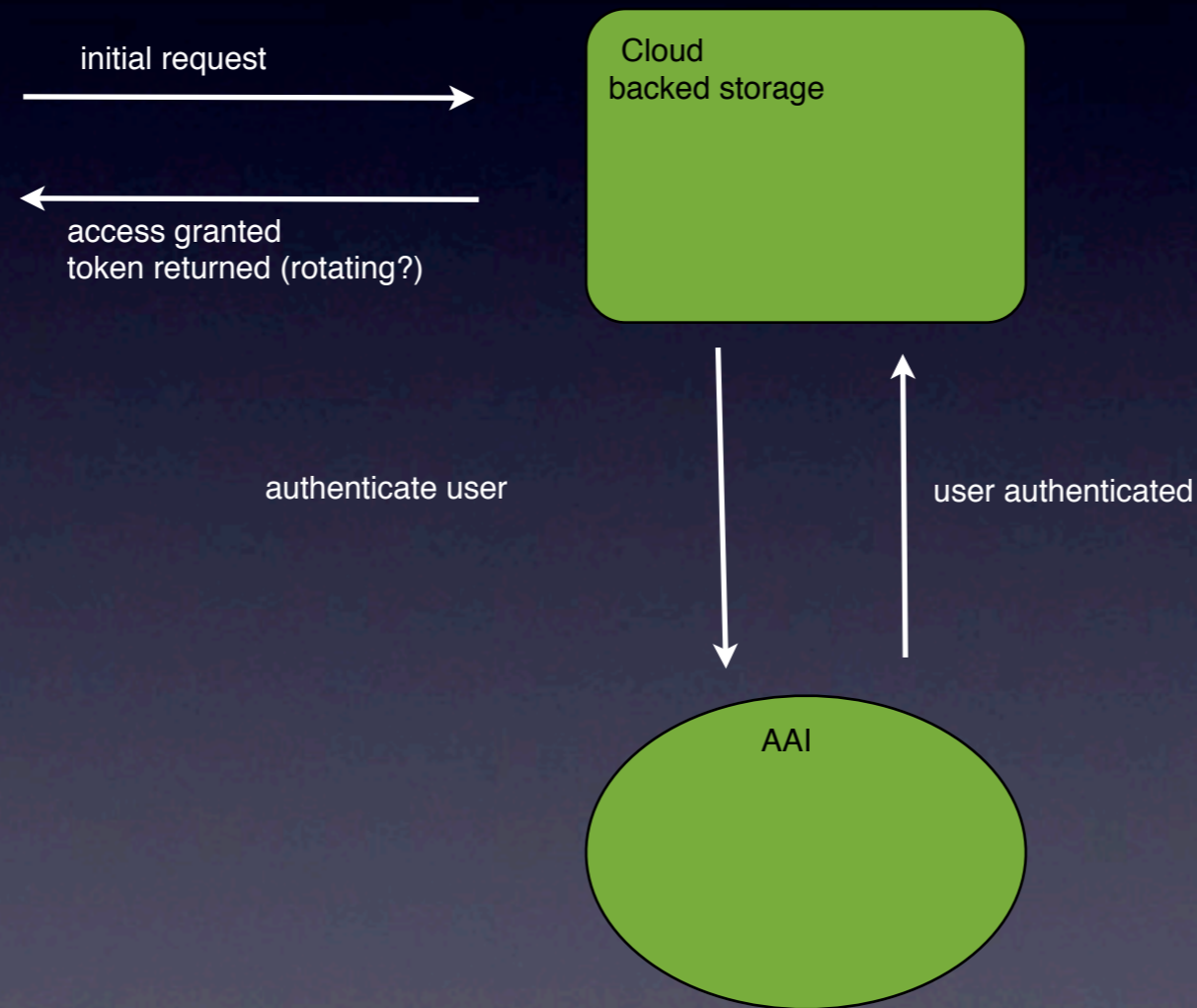
API based - complex for end users



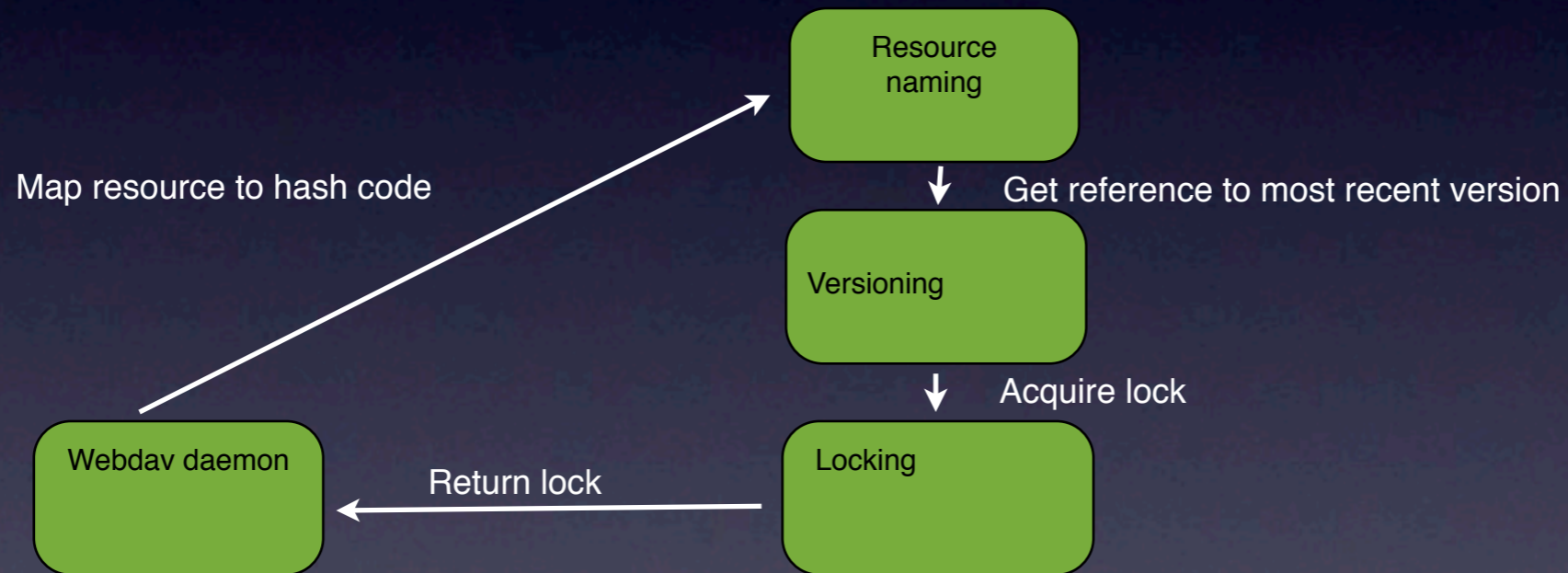
No AAI integration at all

So...

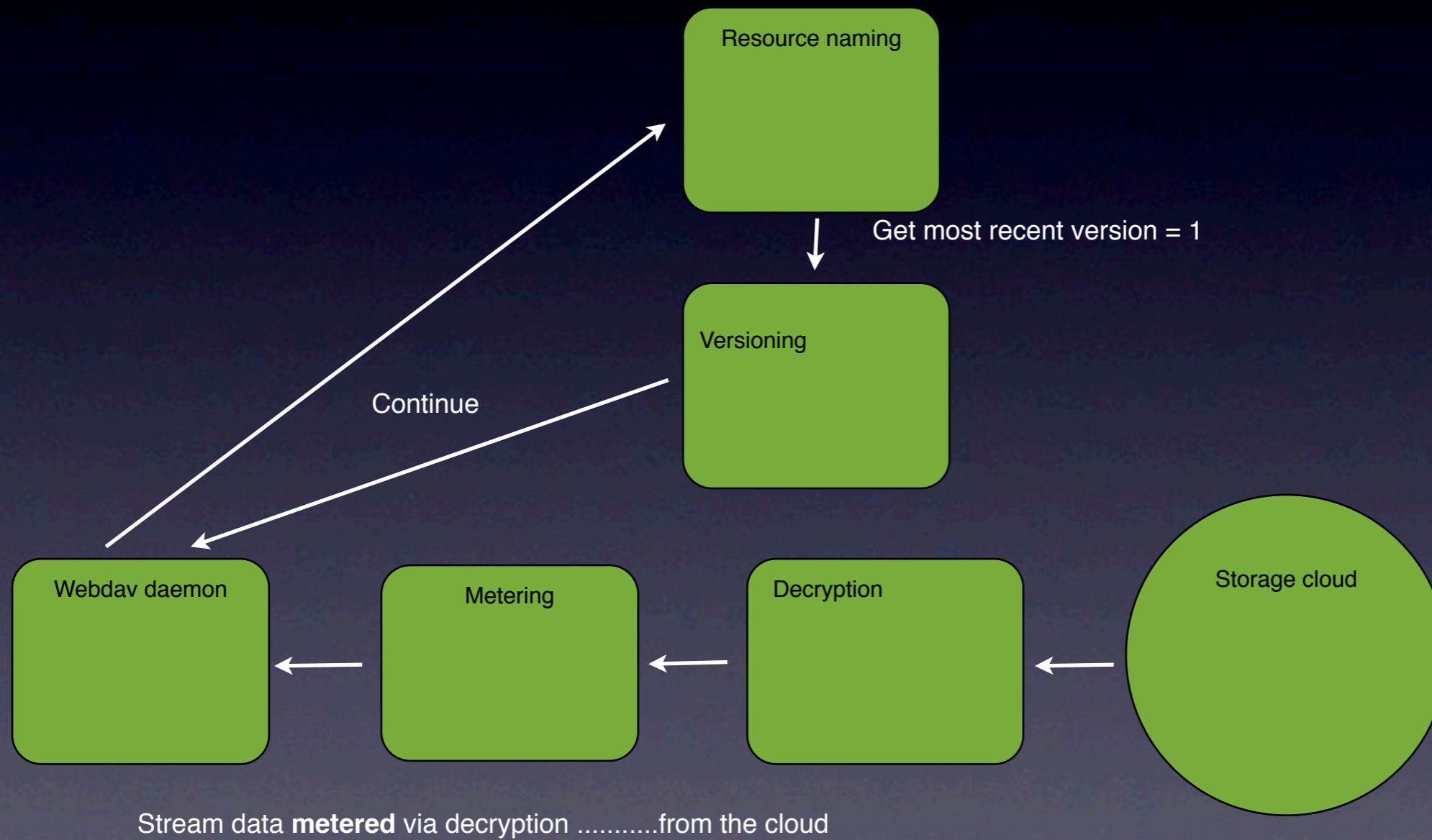
AAI: enrollment



...Locking...



...downloading...



Stand on the shoulders of giants

- WebDAV widely deployed, lots of 3rd party clients.
- Service on top of Java VM
- Scala (integration language)
- Cloud access libraries (often Java based)
- Apache ZooKeeper (configuration management, locking)
- Apache Cassandra or HBase (metering)
- AAI integration components
- ...

Lessons learnt so far

- WebDAV is a nice start for client-less access to file based resources.
- OS X requires DAV level 2 and does a lot of locking. But: from 10.5.x onwards it also does HTTP 1.1 chunked encoding; that broke a lot of servers.
- Windows works best with digest authentication.
- Linux seems to be most forgiving and least demanding.
- all clients support SSL.
- WebDAV's XML is relatively simple but the usage may differ per client type.
- CPU intensive due to the encryption per "stream". A language (model) with concurrency support is a big plus. This breaks the trend of asynchronous I/O based network services.
- Stand on the shoulders of giants: Apache Zookeeper, Bookkeeper, the JVM, Scala language, libraries for cloud access

Questions so far?



Cloud management

Computing

- Torque/PBS on AWS via RightScale
- OSGi on Eucalyptus
- Challenge: Matlab, BLAST etc - how to deal with licensing?

Where we'll be

- TF Storage (March 5Th)
- OGF session (leading, March 16Th)
- NOTUR workshop (leading, May 19Th) -
first NEON WORKSHOP
- TNC 2010 (May 31St)
- Two more workshops in H2-2010
- More presentations... who knows?

Questions II