Starfish Storage, Inc.

- Project launched in 2011 as a prototyping exercise
  - Inspired by IRODS
  - Built on top of IRODS
- V. 1 put into production in mid 2013
  - Removal of all IRODS code
- V. 2 released as production software in late 2014 with limited client base.
- Early 2014 - Spun out of Cambridge Computer into its own company.
  - Still in “stealth mode” but making a lot of noise anyway
  - Currently in beta for v.3

Differences from IRODS

STARFISH

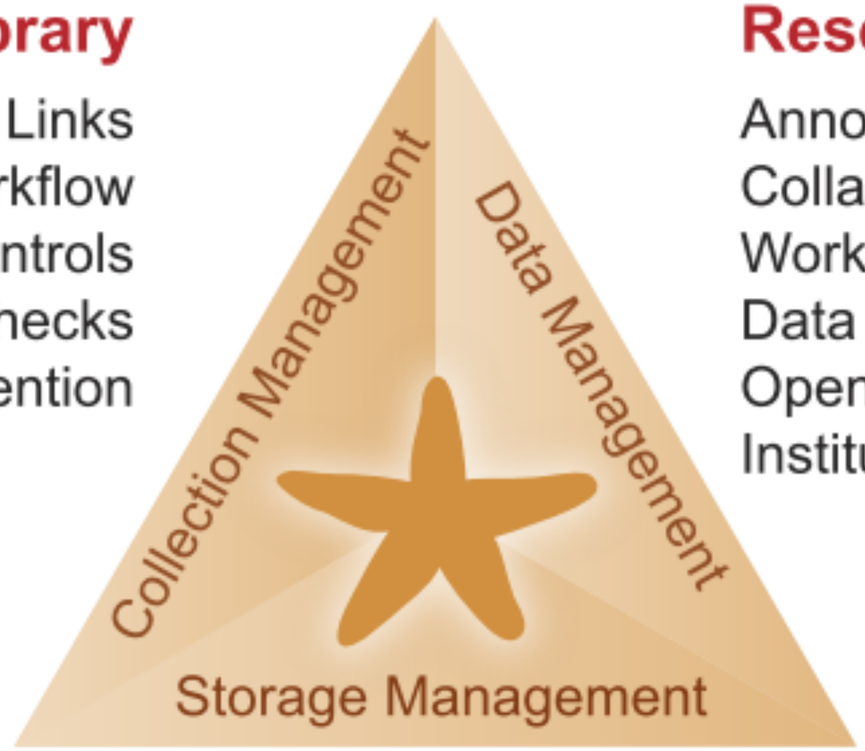
- Built for very large scale
  - Billions of files
  - Petabytes
- Flexible job scheduler rather than rules engine
  - If you need to do anything fancy, write it in Python, or Java or whatever, or we will write it for you.
- Metadata catalog is decoupled from the application
  - Change database technologies over time
- Closer to storage management than collection management
  - We get intimate with the storage infrastructure
- Currently Starfish does not federate user directories across institutions
  - We will, however, integrate with In Common . . . soon

The Big Vision

STARFISH

Library

Open Links
Curation Workflow
Access Controls
Fixity Checks
Repository Retention



Research

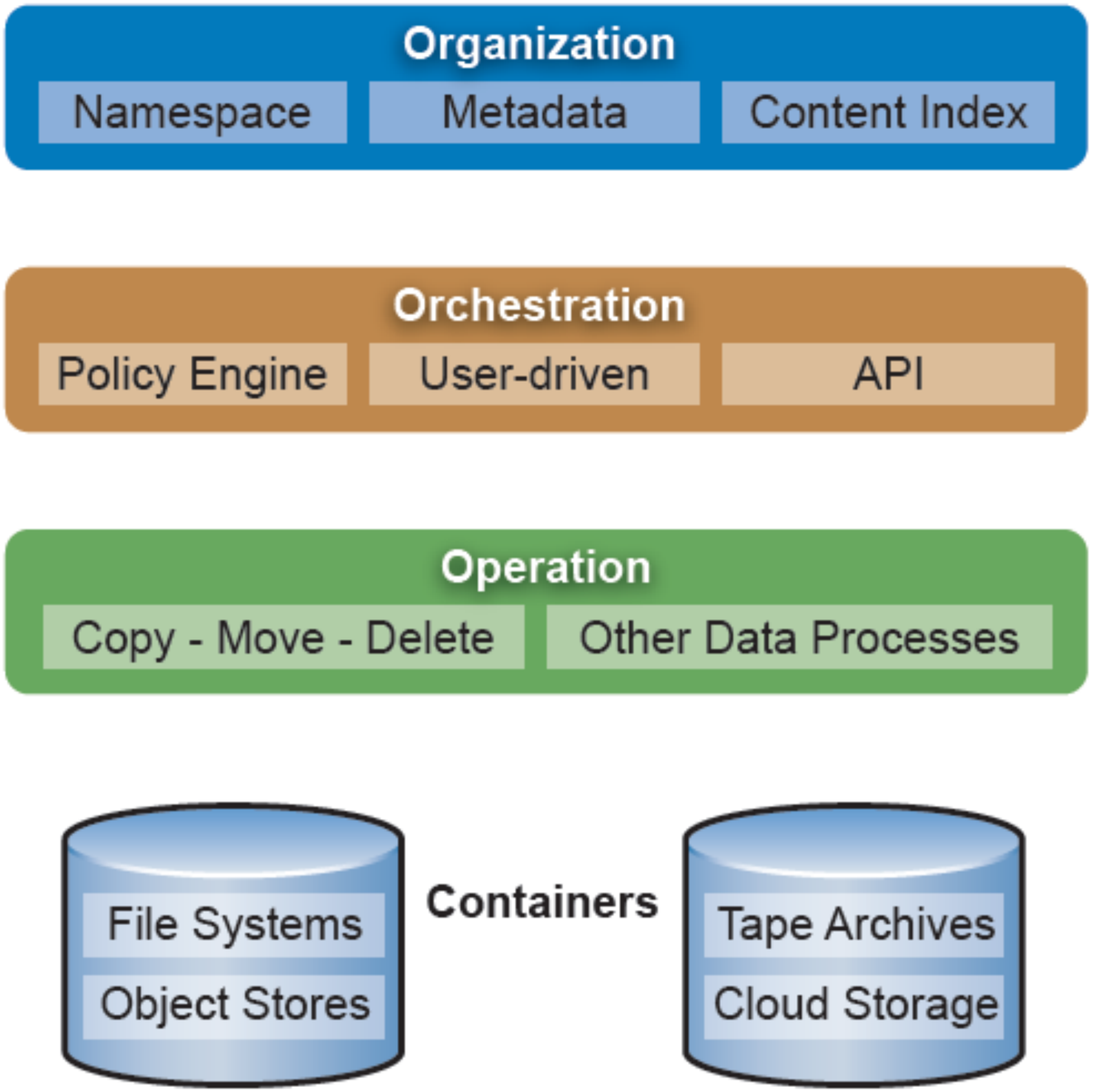
Annotation
Collaboration
Workflow
Data Management Plans
Open Access
Institutional Retention

IT Operations

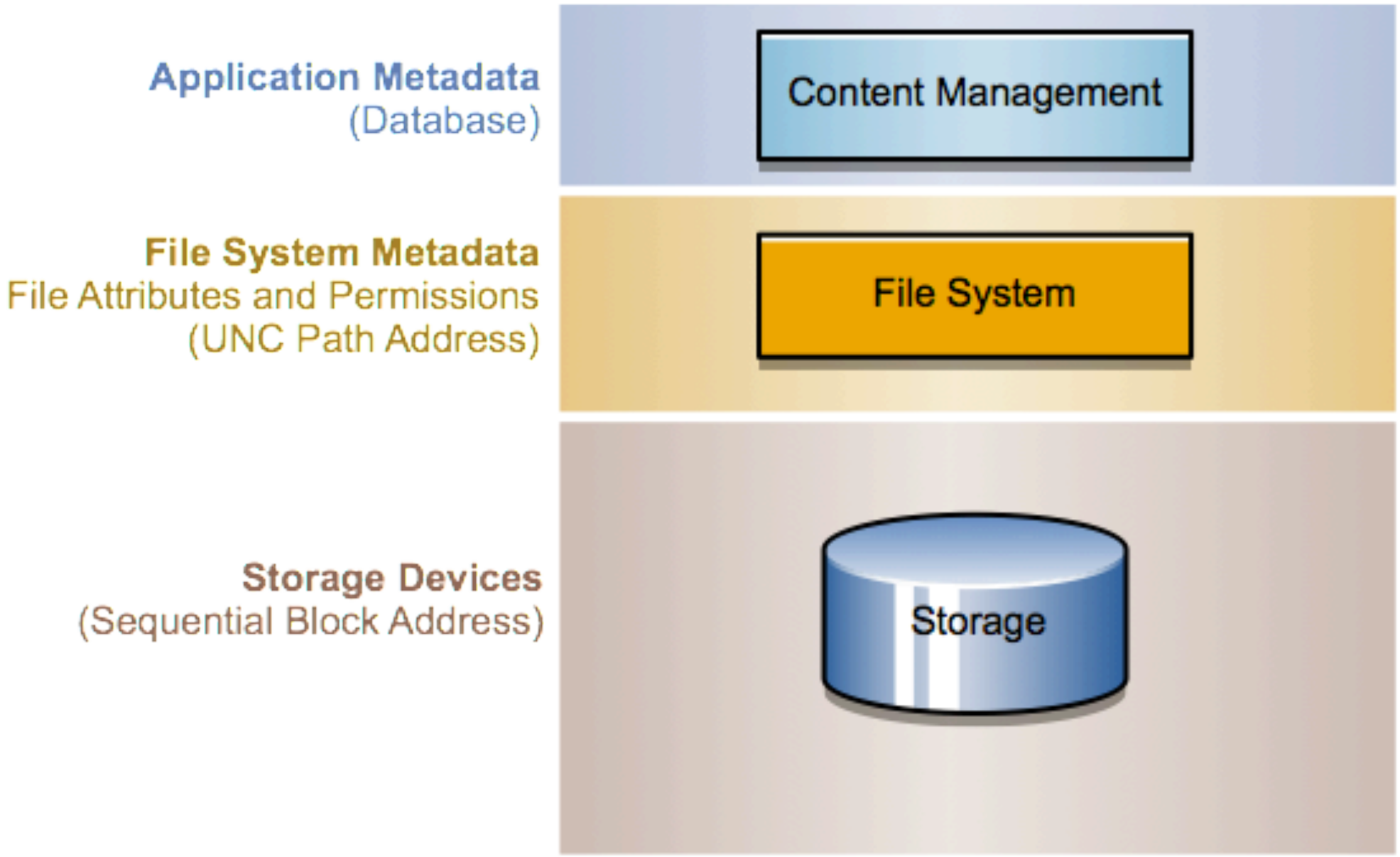
Storage Management
Backup/Restore
Tiered Storage
Data Migration
Cloud Interface

Reporting
Chargeback / Showback
General Reporting
IT Governance

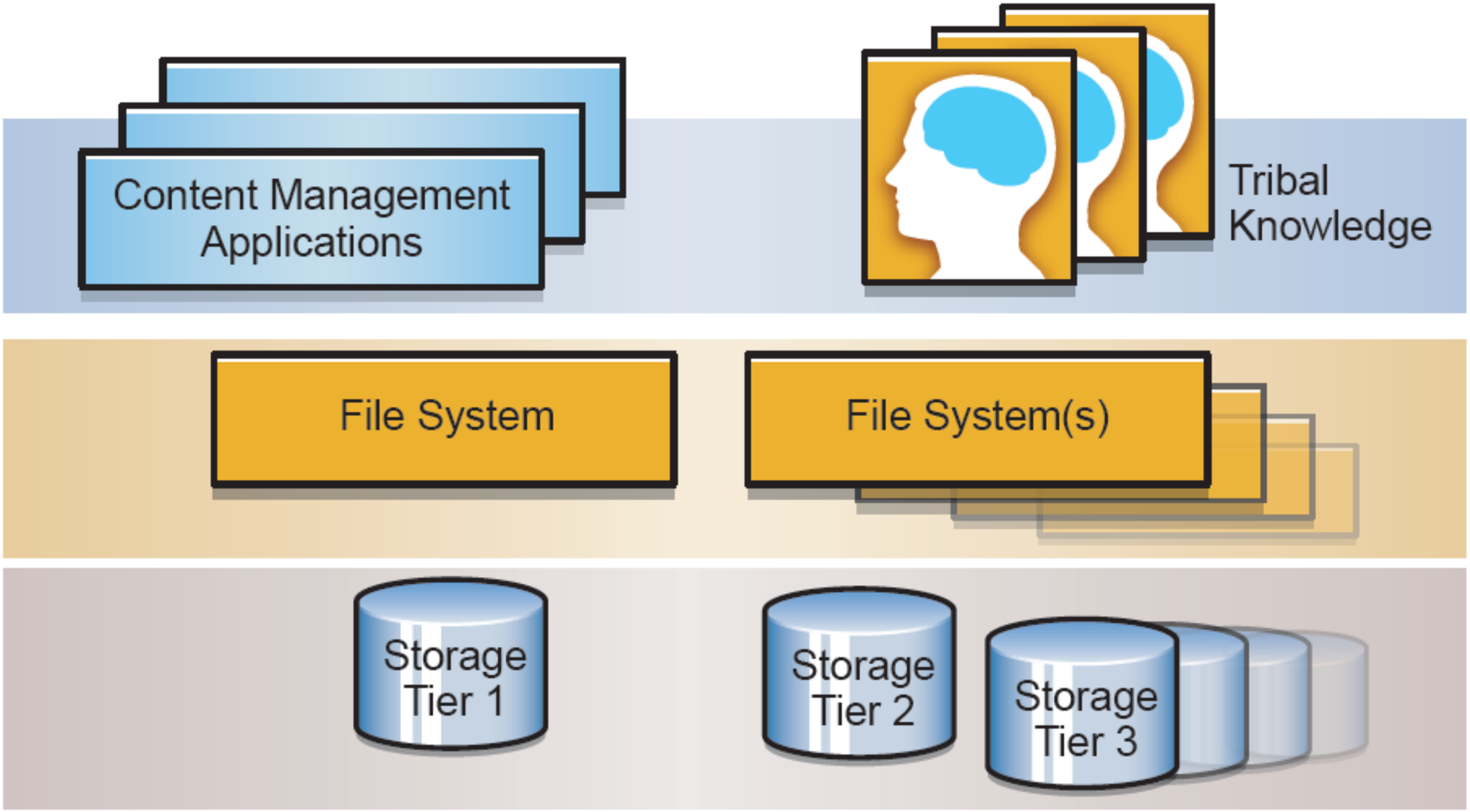
The OOO Model For Data Life Cycle Management



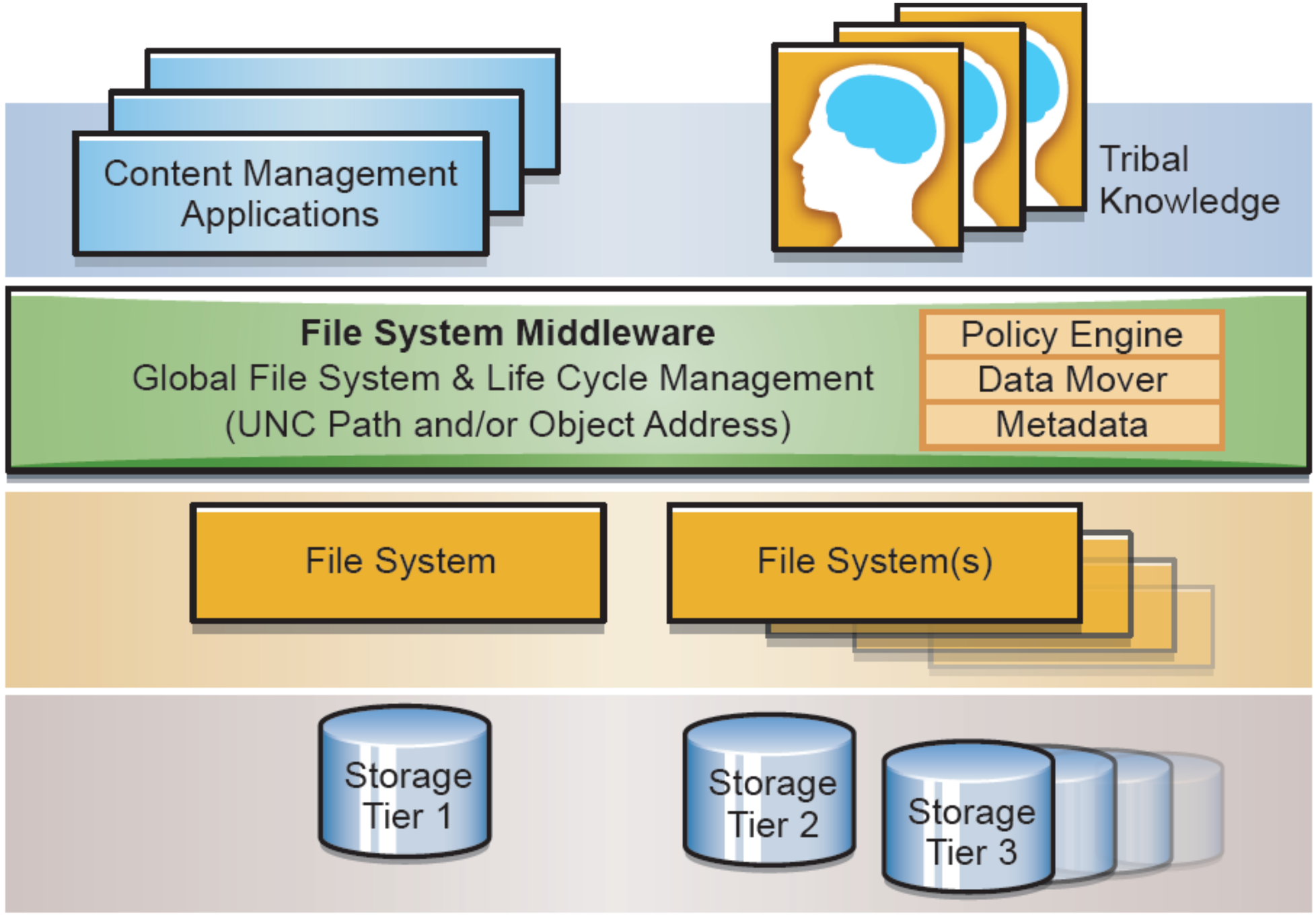
Typical Content Management "Stack"



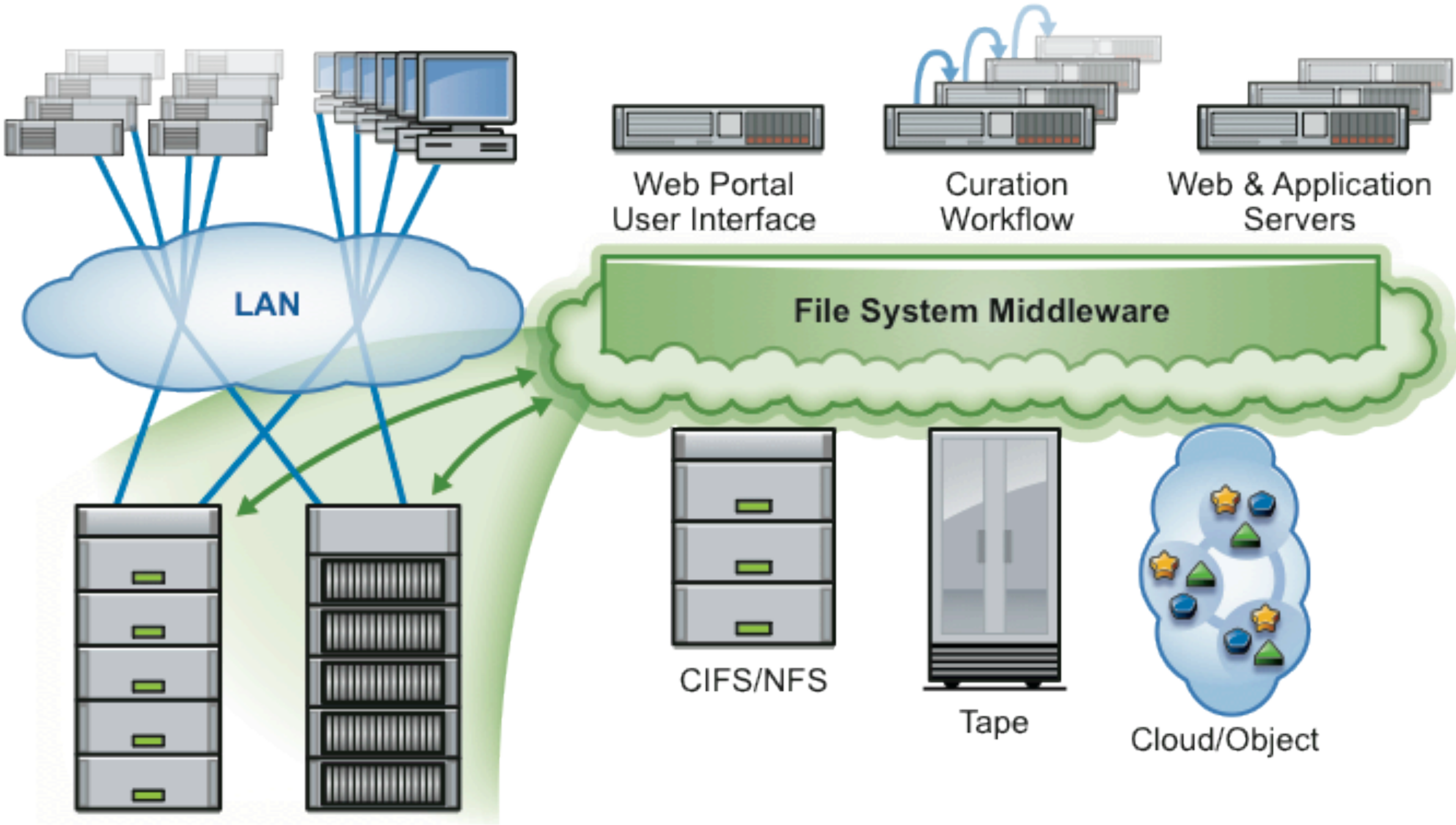
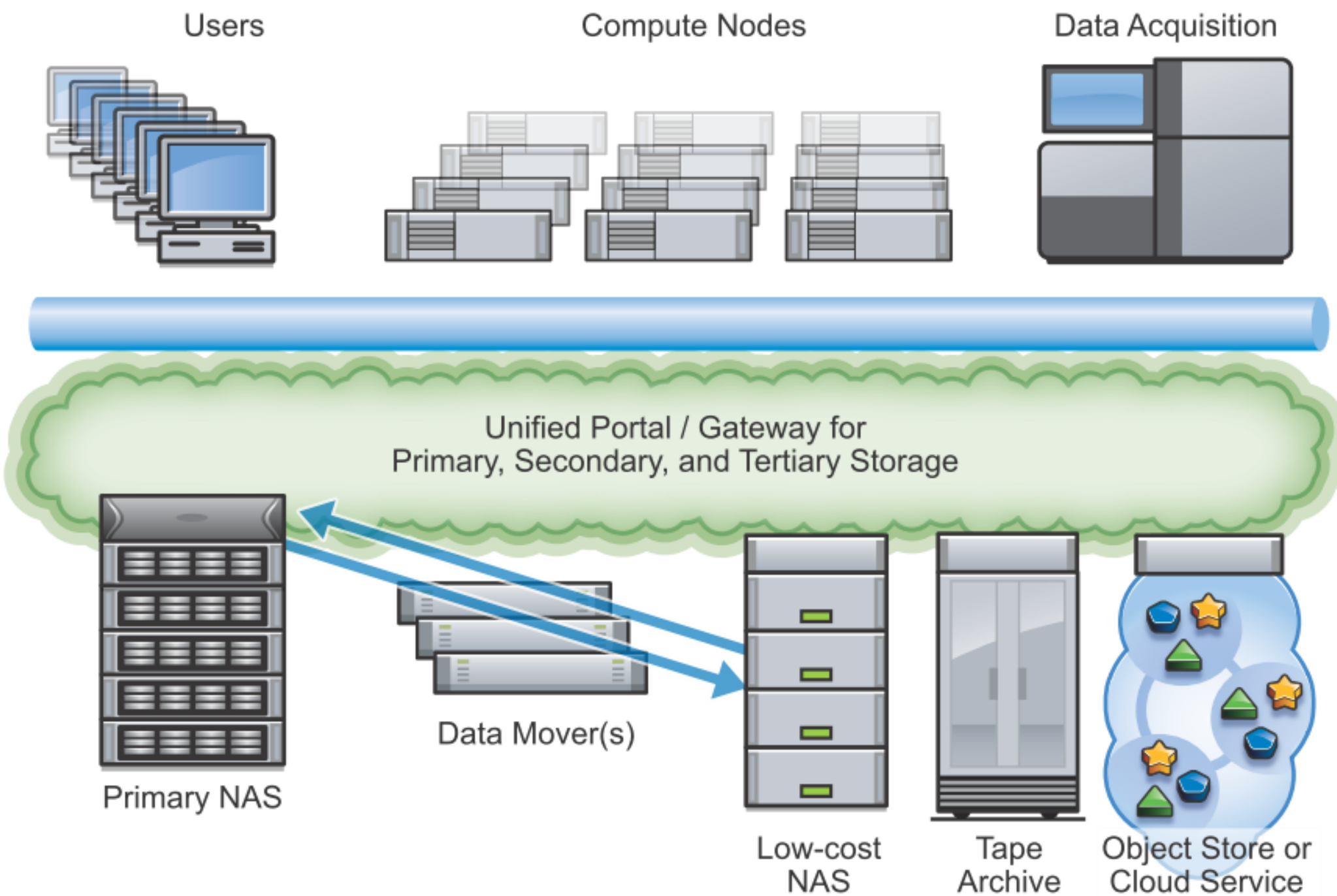
Data Management "Stack"



Data Management Stack with Middleware



All in One: Backups, Archiving, Curation, Storage Management



A “Virtual” Global File System