



DURACLOUD™

Building community clouds to support access to scholarship

Michele Kimpton

CEO, DuraSpace

Jonathan Markow

CSO, DuraSpace



Who are we?



We are committed to providing open source technologies and services that promote durable, persistent access to the scholarly record.

Our Technologies



Community challenges

- Libraries being asked to steward increasing amounts of digital content
- Difficult to easily provision storage, without advanced planning and capital
- Limited technical resources
- Need more flexible and scalable solutions dependent on type of content they are managing

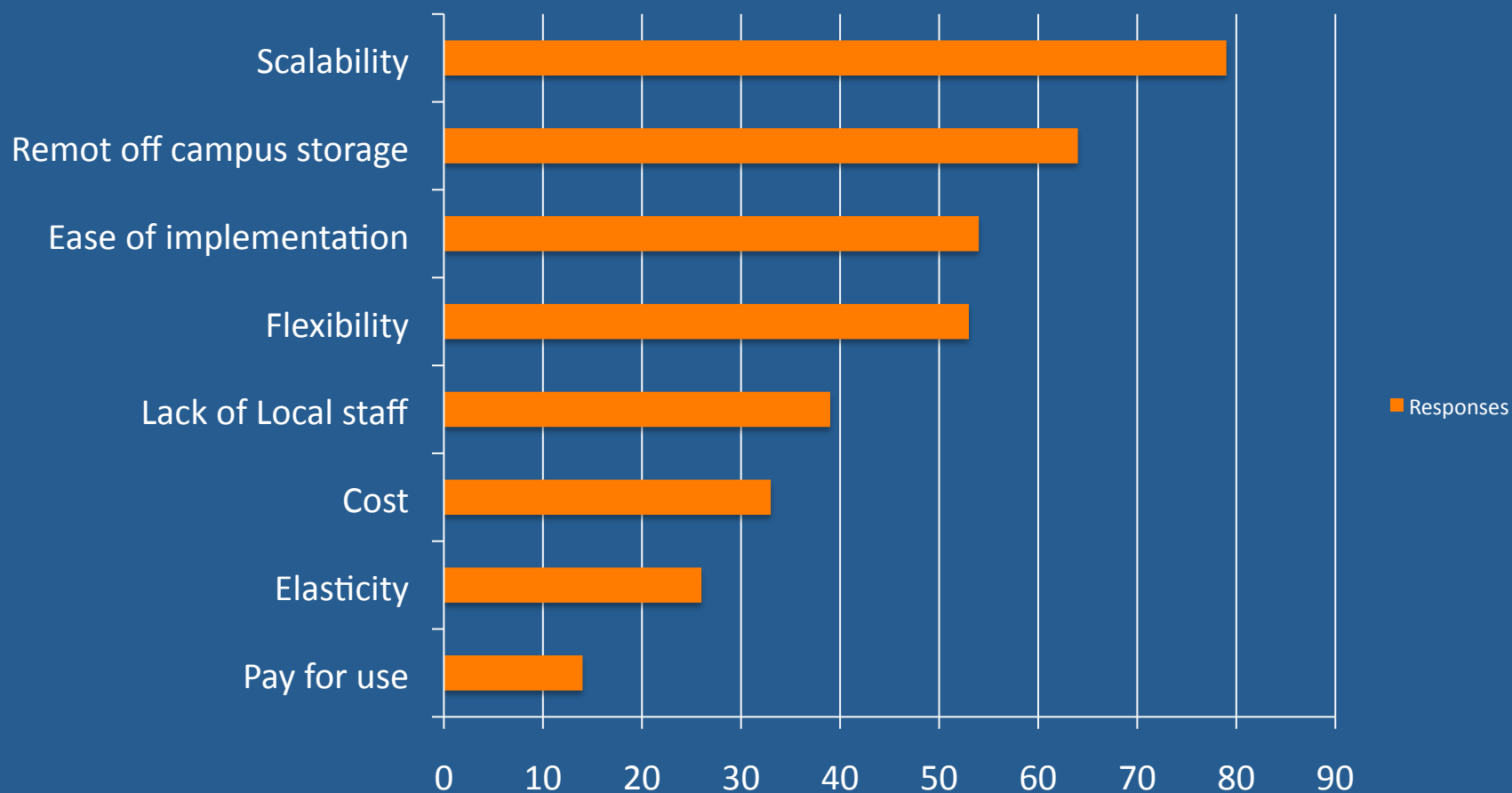
Why cloud?

Massively scalable compute and storage offered as a service



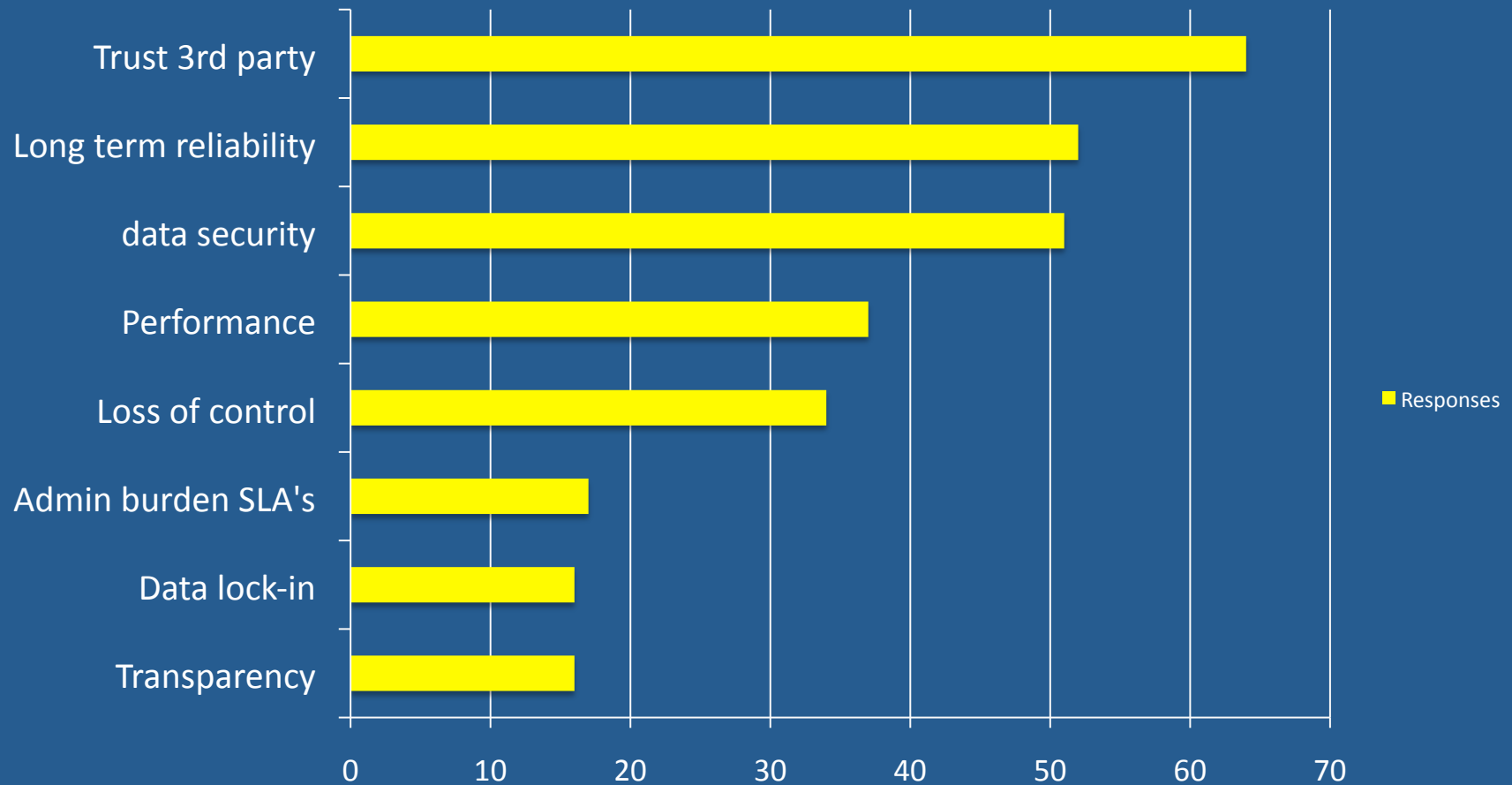
Survey Higher Ed, 145 responses

Key Benefits



Survey Higher Ed, 145 responses

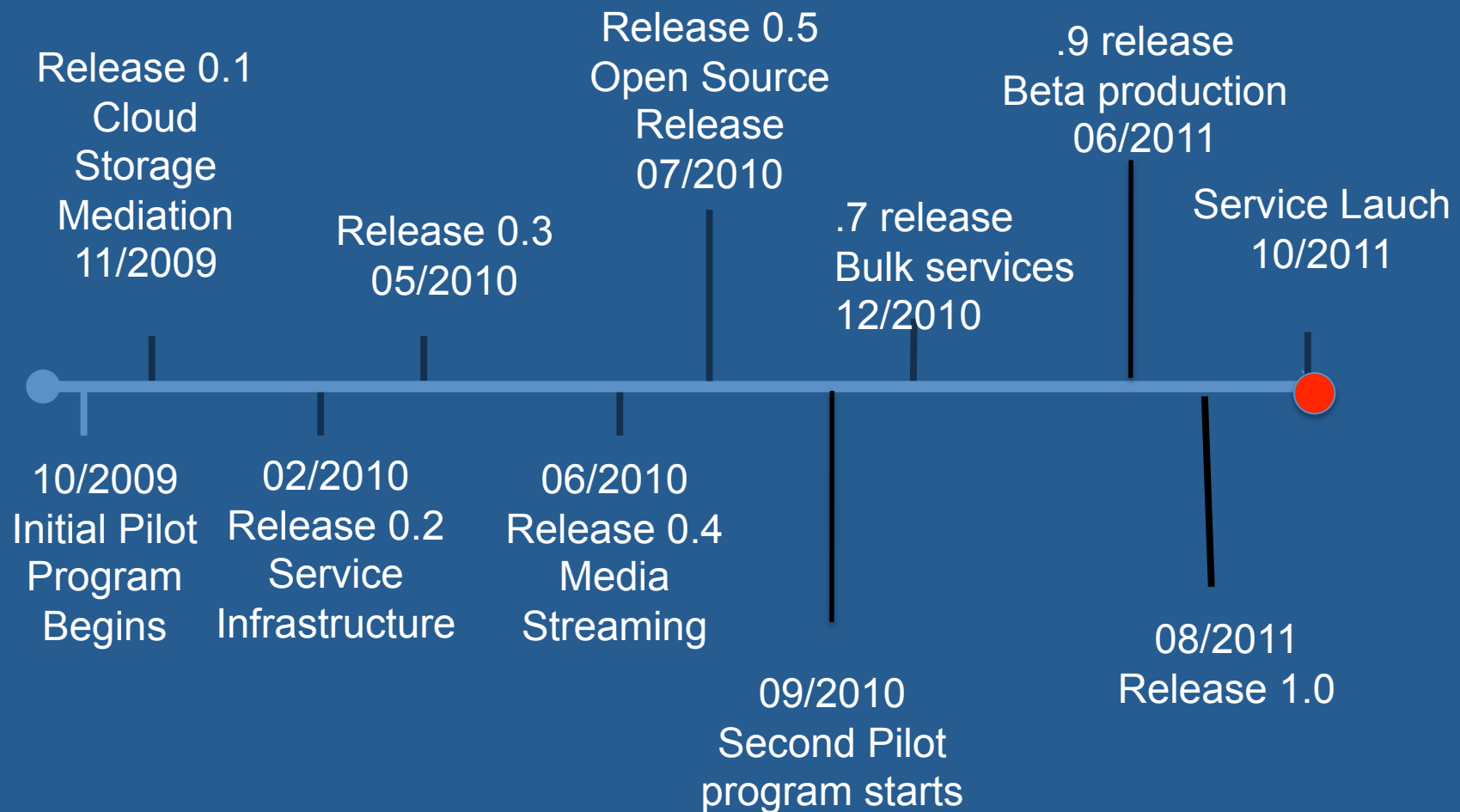
Key Challenges



Advantages of a community cloud solution

- Develop cloud expertise within our own community
- Mitigate the risks of working with a single cloud vendor
- Develop solutions to solve needs of our community
- Aggregate demand across our community to come up with best cloud infrastructure and best pricing
- Consolidate efforts across community to produce a diverse and robust platform and service

DuraCloud Project History



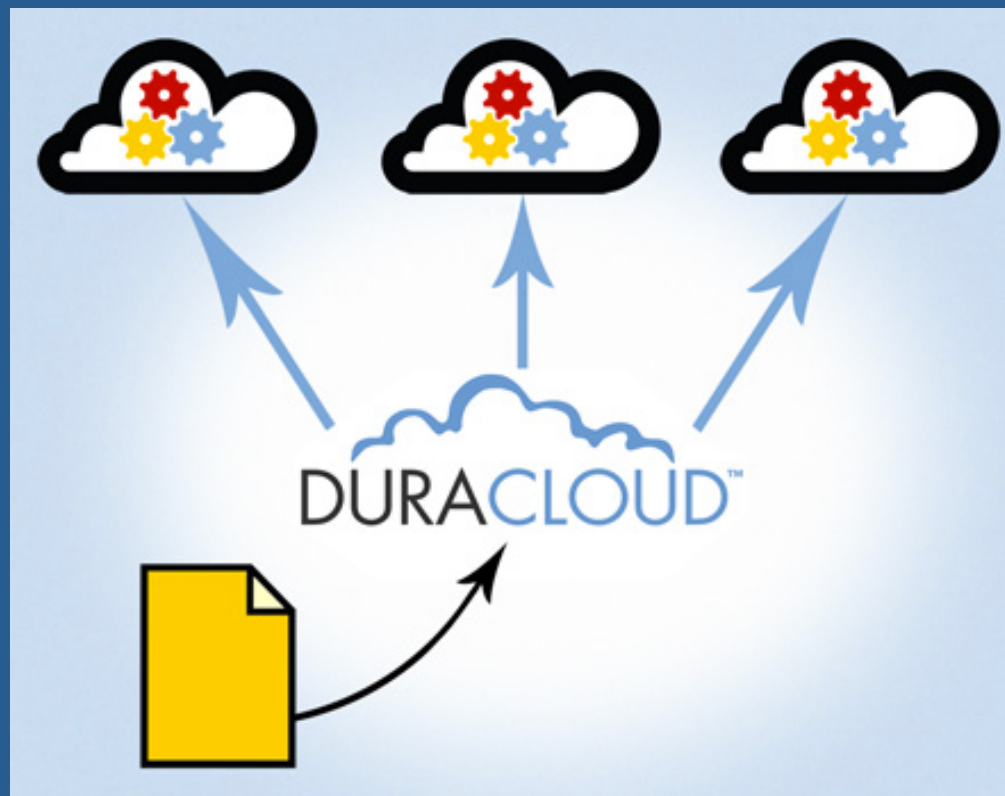
Pilot Partners

| University | Use Case | Repository |
|------------------------|------------------------------------|----------------------|
| Rice U | Preservation | DSpace, meta archive |
| Hamilton College | Access/international collaboration | Fedora |
| Northwestern U | Preservation books, audio, image | Fedora |
| U of PEI | Image access | Fedora/Islandora |
| ICPSR | Preservation | Fedora |
| IUPUI | Preservation | DSpace, Content DM |
| Rhodes College | Image Access | DSpace |
| North Carolina State U | Preservation | DSpace |
| CARL | Preservation and Services | Fedora |
| MIT | Preservation | Dspace |
| Columbia | Preservation and Services | IA, Fedora |

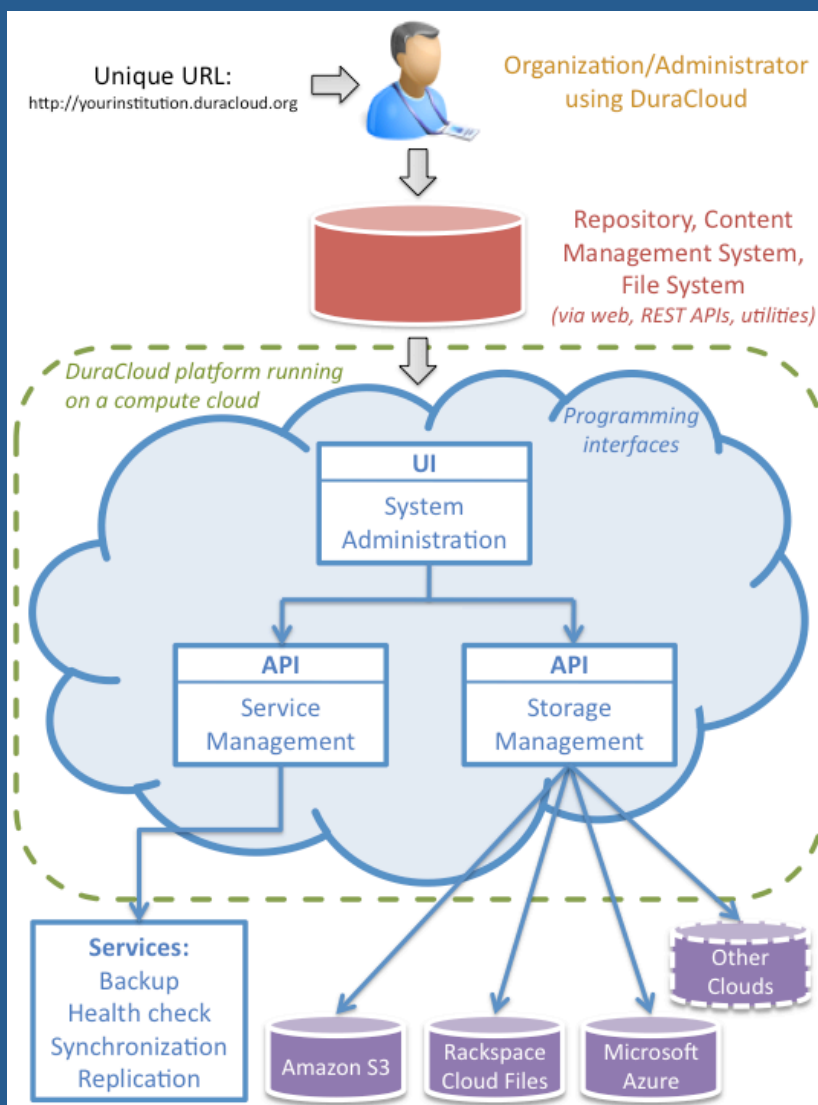
What is DuraCloud?

Digital archiving solution run on cloud infrastructure

Multiple copies in multiple clouds



DuraCloud Platform



Use cases explored

Digital archiving



Video streaming



Data management



Image serving



Collaboration

What partners are saying

- “The ***ease-of-use*** of DuraCloud is its most impressive feature.”
~Columbia



- “DuraCloud delivers the ***benefits*** of a diverse network of storage locations, but ***without the overhead*** of managing different vendors.”
~ICPSR



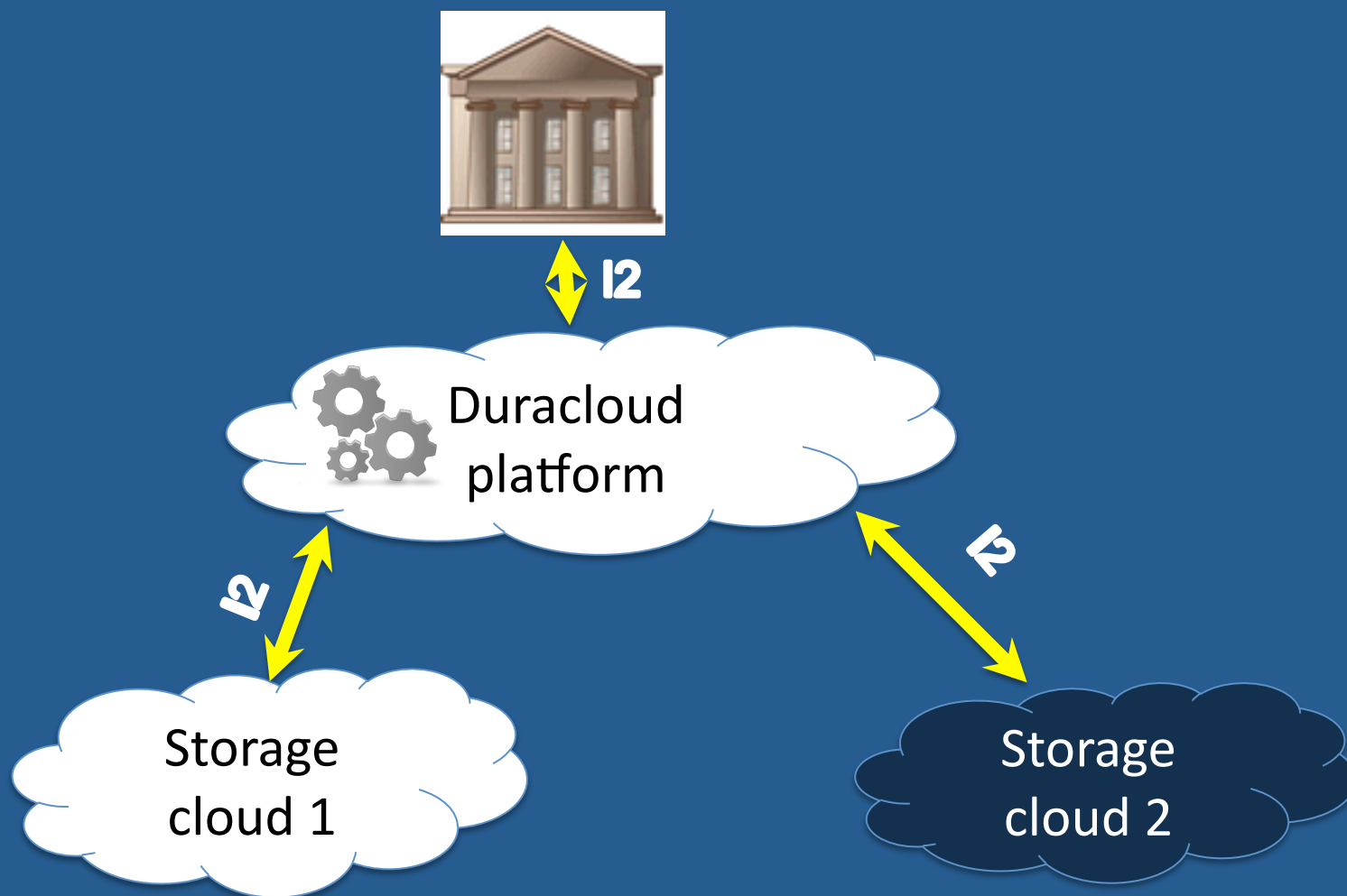
- DuraCloud provides a robust infrastructure to support digital preservation of content in our institutional repository, along with services that enhance the delivery and integrity of the content."
~Geneva Henry, Rice University



Benefits of working with Community Partner, DuraSpace

- Trusted non profit from academic community
- Community source software
- We have developed expertise cloud technology
- We can aggregate volume across institutions and negotiate a better price and better terms
- We can monitor all cloud providers looking out for your best interests

Internet 2 implementation community cloud



Advantages

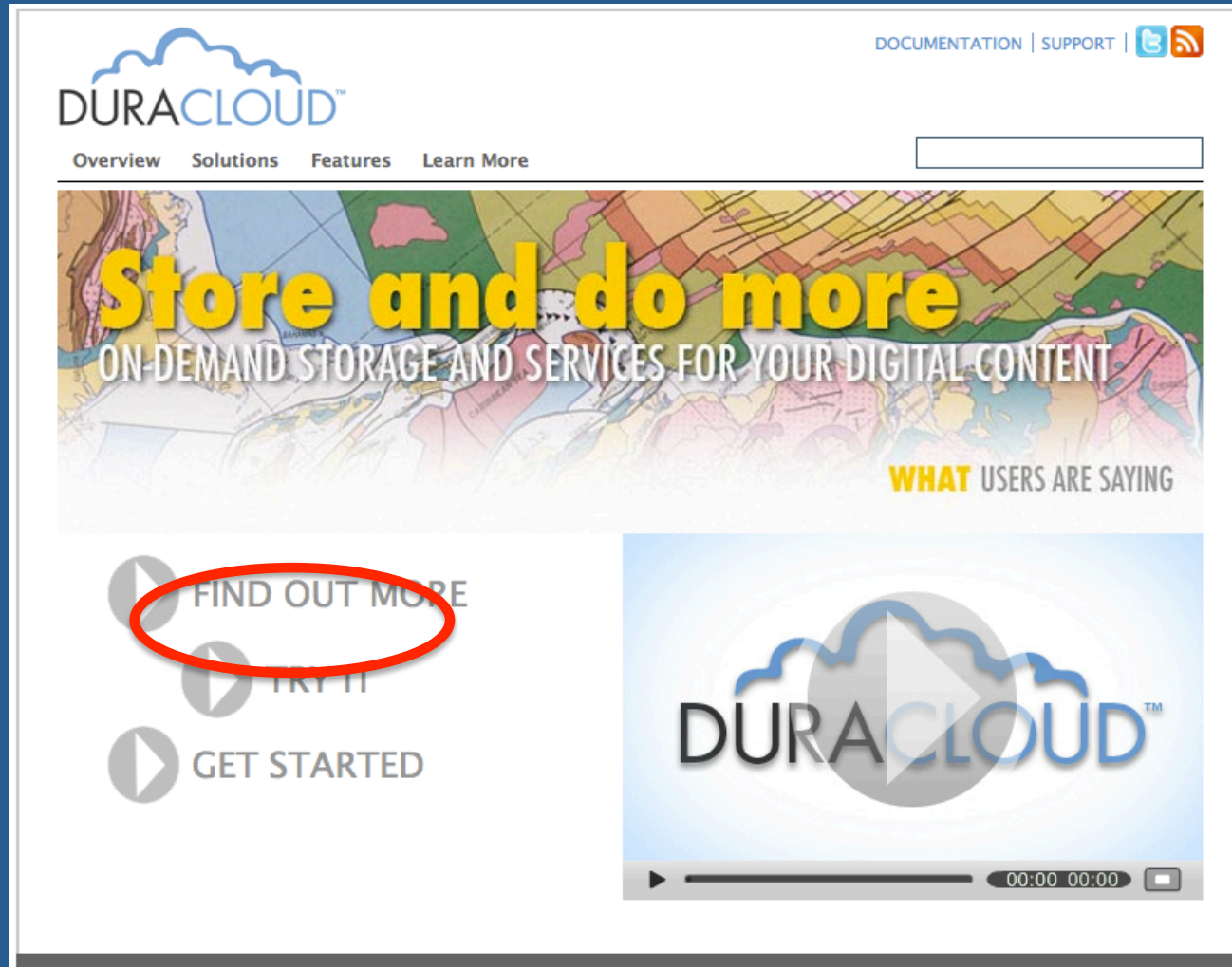
- Shibboleth authentication and integration
- No additional bandwidth charges
- Cloud providers are part of Internet 2 partner network
- Community managed and supported service

On the Roadmap

- Services
 - Shibboleth authentication
 - DSpace as a service application
 - Video management
 - Data curation and management
- Storage providers
 - UCSD cloud store, Chronopolis
 - SHI cloud, Net plus offering
 - Local Eucalyptus integration

Free DuraCloud Trial Accounts

www.duracloud.org



The screenshot shows the DuraCloud website homepage. At the top left is the DuraCloud logo. To its right are links for "DOCUMENTATION" and "SUPPORT", along with social media icons for Twitter and RSS. Below the logo is a navigation menu with "Overview", "Solutions", "Features", and "Learn More". A search bar is located to the right of the menu. The main banner features a colorful map background with the text "Store and do more" in large yellow letters, followed by "ON-DEMAND STORAGE AND SERVICES FOR YOUR DIGITAL CONTENT" in white. Below this, on the right, is the text "WHAT USERS ARE SAYING". On the left side of the banner, there are three circular buttons with play icons: "FIND OUT MORE" (circled in red), "TRY IT", and "GET STARTED". To the right of these buttons is a video player showing the DuraCloud logo with a large play button in the center. The video player has a progress bar at the bottom showing "00:00 00:00".

DuraCloud DfR

A DuraSpace Project Funded by
the Alfred P. Sloan Foundation

Why DuraCloud DfR?

- Protect vulnerable research data
- Enable archiving, access, and preservation
- Facilitate cooperation between researcher and institutional data managers
- Provide services to support the research process

Advisors and Participants

- Cornell
- George Washington
- Georgia Tech
- Harvard
- ICPSR
- Johns Hopkins
- MIT
- NCAR/UCAR
- Oregon State
- Rice
- Smithsonian
- U of Oregon
- U of Prince Edward Island
- U of Virginia
- Fluid Project
- DuraSpace

Top Five Priorities

- #1: Connect the operational and archival phases of the data management lifecycle.

Top Five Priorities

- #2: Create simple workflows across the data management lifecycle that automatically capture metadata and provenance.
- (...and create incentives for additional metadata creation)

Top Five Priorities

- #3: Ensure confidentiality, security, privacy, and predictability of data in the cloud. (Trust and Control)

Top Five Priorities

- #4: Automate basic metadata creation and “catalogue” creation.

Top Five Priorities

- #5: Create interoperability of operational systems, archiving solutions, and discovery systems used by specific research communities.



DfR Principles

- Open source, enterprise software solution
- Capture data close to the source
- Don't interfere with researchers' processes
- Provide incentives, added value for metadata creation
- Easy to use; workflows for collaboration, hand-off to institution



Architecture

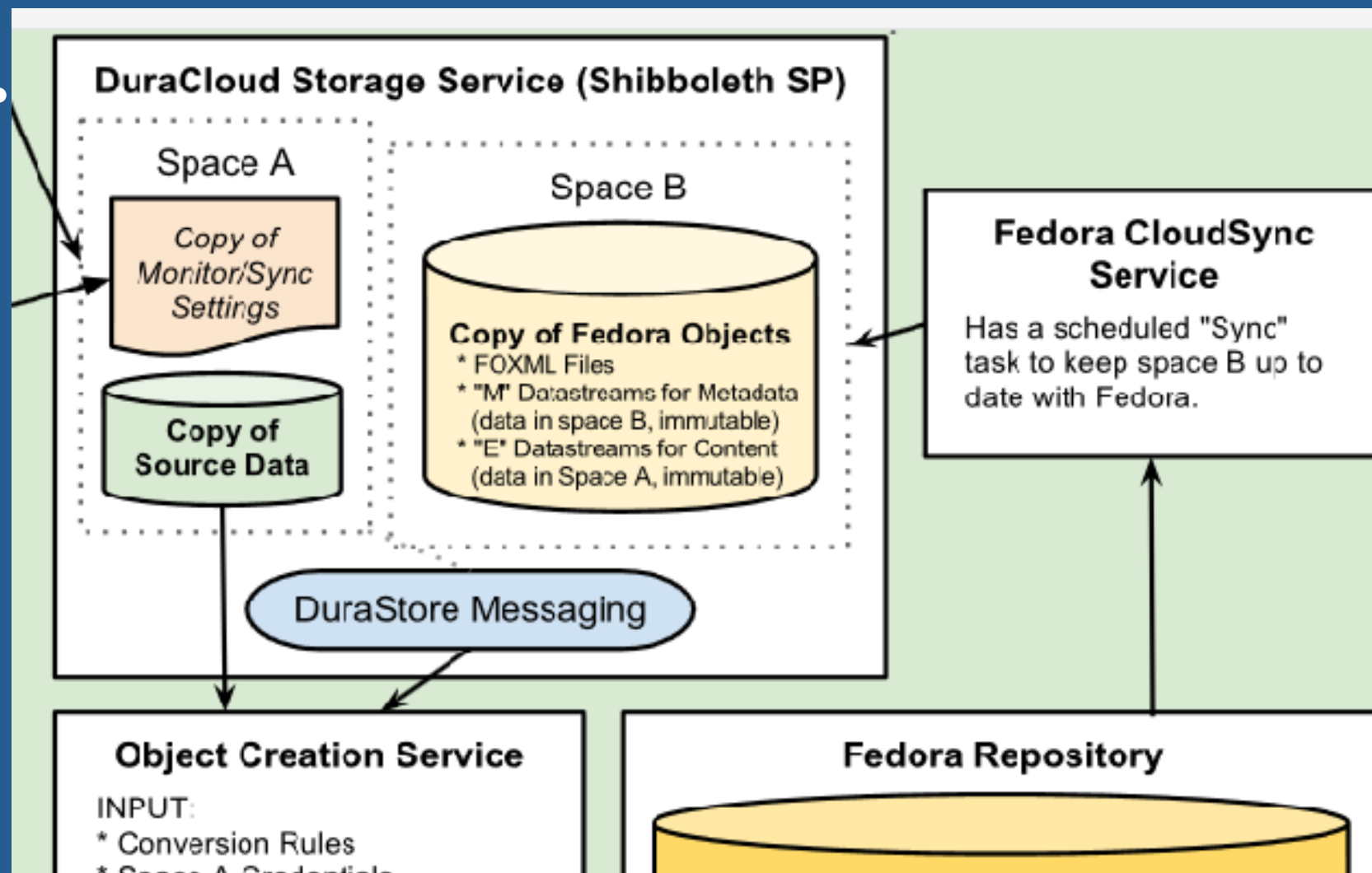
- Take advantage of open source reuse wherever possible, e.g.:
 - Authentication (Enterprise SSO, SAML2)
 - Cloud-based repository components
 - Encryption, service bus, messaging
 - DMP tools, persistent identifiers, citations, publishing references,
 - etc.

Components

- Authentication:
 - Pluggable
 - Shibboleth support in 1.0 release



Components



User eXperience

- Reuse existing tools for visualization, manipulation of research data
- Create seamless experience for researcher
- Utilize best practices for interaction design





Partnerships

- Smithsonian Institution Office of Research Information Services – UI for the management and visualization of research project data
- Internet2/InCommon (in discussion)

Project Schedule

- Iterative development
- Evolving prototypes
- User Focus
- Expanding group of participants
- First production release at end of 2012
(Four interim releases throughout the year)



Questions

DuraSpace: www.duraspace.org

DuraCloud: www.duracloud.org

Michele Kimpton
Chief Executive Officer
mkimpton@duraspacespace.org

Jonathan Markow
Chief Strategy Officer
jmarkow@duraspacespace.org