

Responsive Storage

LUG 2017

Ryan Chard

Rick Wagner

Argonne National Lab &
University of Chicago



THE UNIVERSITY OF
CHICAGO





Globus Labs: Responsive Storage



Rachana
Ananthakrishnan



Ben Blaiszik



Francesco
de Carlo



Kyle Chard



Ryan Chard



Ian Foster



Raj
Kettimuthu



Dula
Parkinson



Arnab Paul



Jim Pruyne



Steve Tuecke

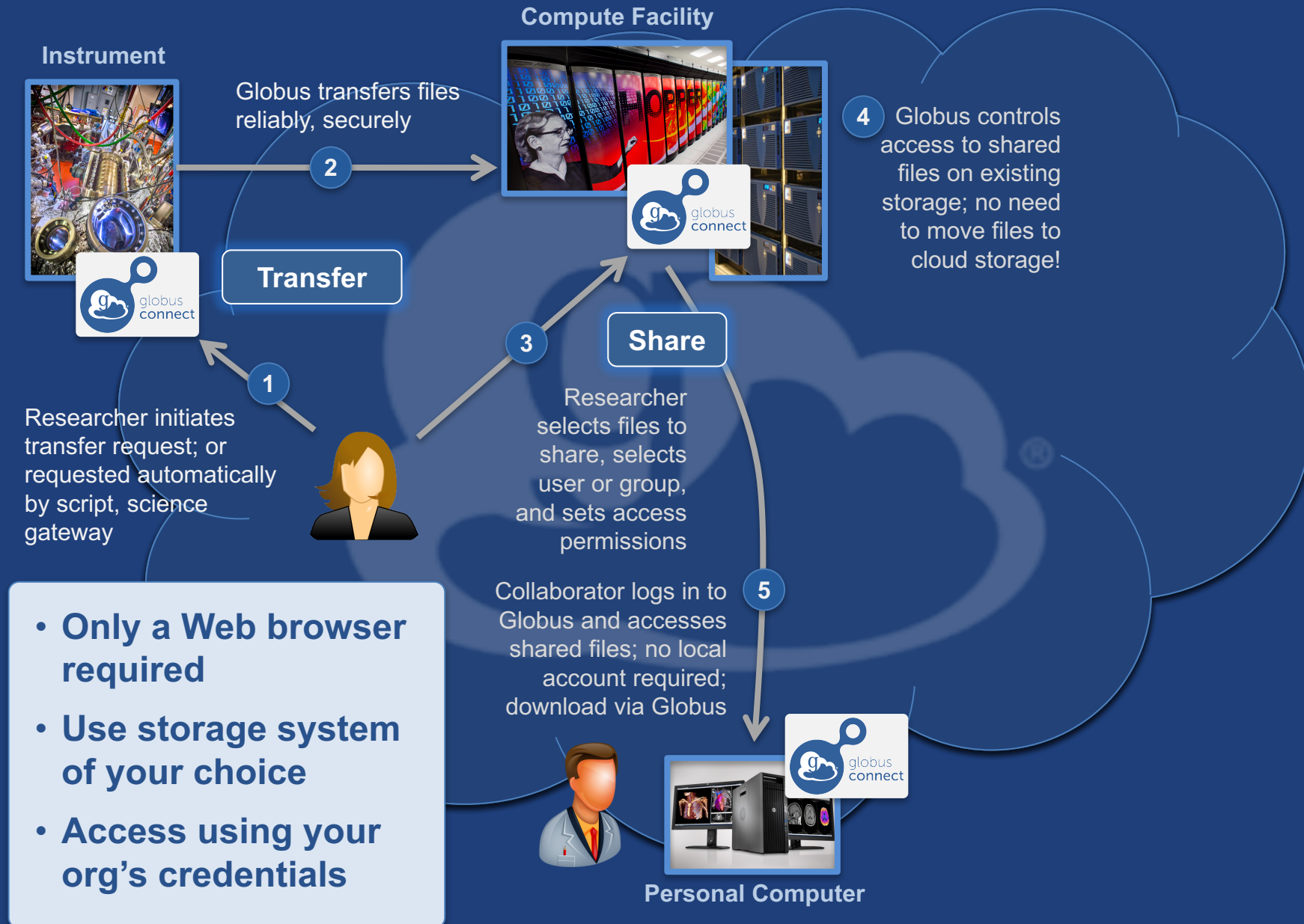


Rick Wagner

Argonne National Laboratory, Lawrence Berkeley National Laboratory, and other institutions



What is Globus?



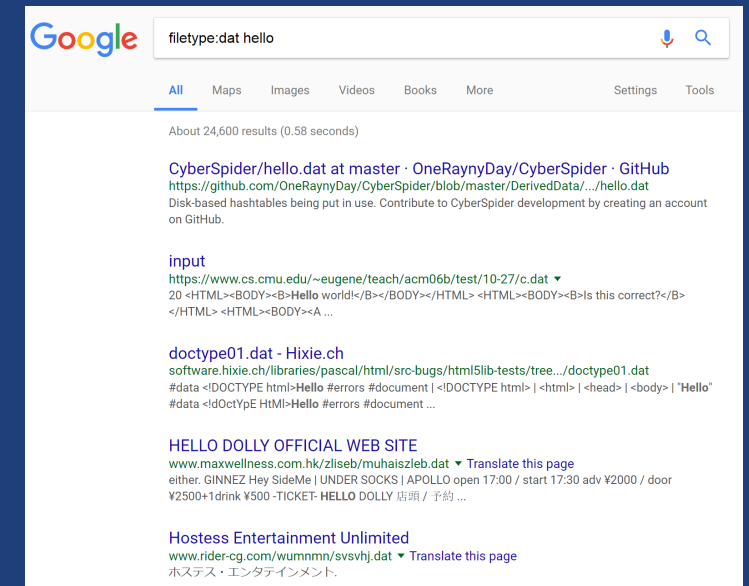
- Only a Web browser required
- Use storage system of your choice
- Access using your org's credentials



Use Case: Where's My Data?

We want to catalog every file created by a user

- Identify file types
- Extract relevant metadata
- Enable search across all my files



RIPPLE: A Responsive Storage Solution

- Reliable event-driven data management automation
- Designed for users
 - Non-technical users define simple trigger/action policies
 - Developers can leverage as an automation framework
- Filesystem agnostic (`inotify`, `kqueue`, `Changelogs`, etc.)
 - Works on both edge and leadership platforms
- Coupled with Globus endpoint on machines mounting the filesystem

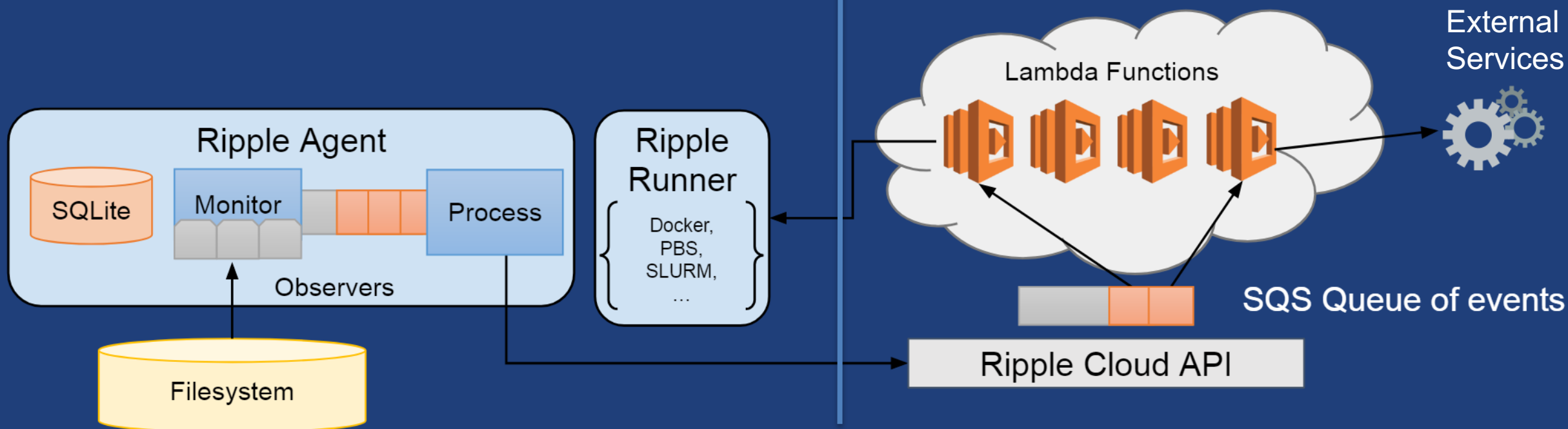
RIPPLE Architecture

Ripple Agent

- Sits locally on machine (edge or HPC)
- Detects & filters filesystem events (POSIX, Lustre)
- Facilitates execution of actions via a job runner

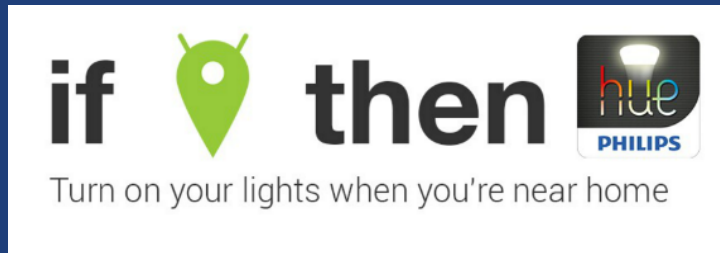
Ripple Cloud Service

- Stores triggers & actions
- Lambda functions process events
- Orchestrates execution of actions
- Records and manages execution of flows



RIPPLE Rules

- If-Trigger-Then-Action



- Register trigger conditions on an endpoint
- Associate an action to be executed in response
- Chain simple rules to construct complex data flows

```
"trigger" : {  
  "monitor" : "lustre"  
  "event" : "01CREAT",  
  "directory": "/mnt/scratch/",  
  "match": ".*.h5",  
  "endpoint" : "a782c1ac..."  
},  
"action" : {  
  "service" : "singularity"  
  "exec" : "python ingest.py $pathname",  
  "target_match" : "",  
  "target_replace" : ""  
  "endpoint" : "5e67e24e...",  
}
```



Monitoring Lustre Changelogs

Register as a changelog user on the MDS:

```
$ lctl --device scratch-MDT0000 changelog_register  
scratch-MDT0000: Registered changelog userid 'cl1'
```

Create a file and check logs on worker:

```
$ touch test.txt  
$ lfs changelog scratch-MDT0000  
1 01CREAT 17:46:34 2017.05.29 0x0 t=[0x200000401:0x1:0x0] \  
p=[0x200000007:0x1:0x0] test.txt  
2 11CLOSE 17:46:34 2017.05.29 0x42 t=[0x200000401:0x1:0x0]
```

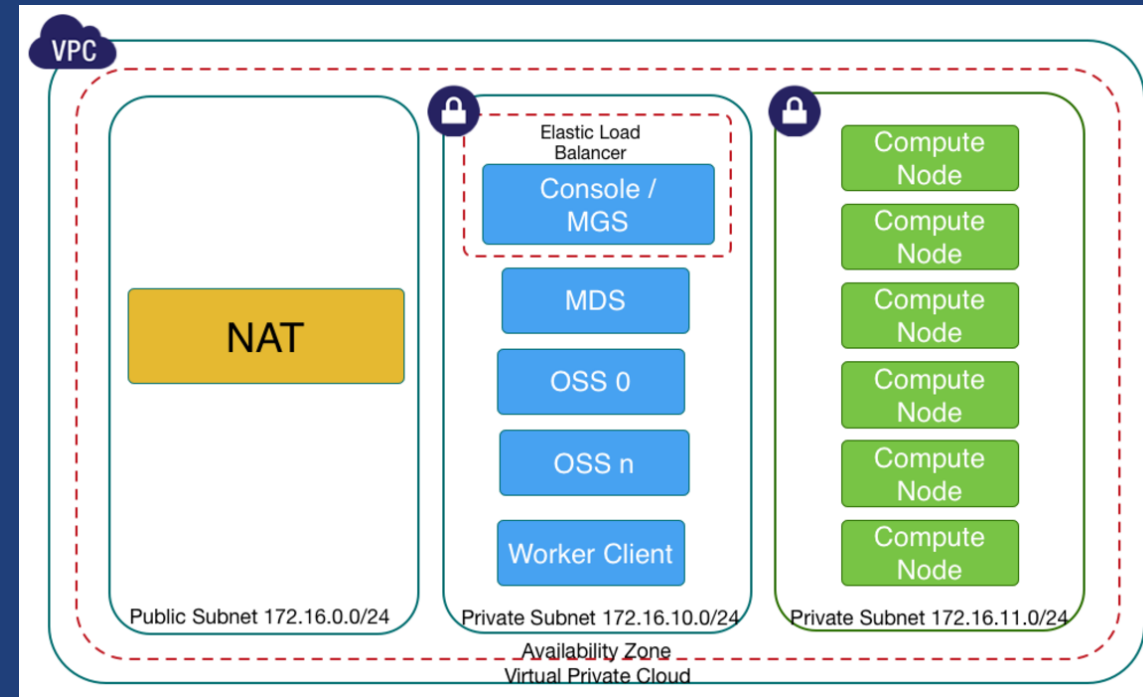
Use fid2path to resolve path and enable trigger matching:

```
$ lfs fid2path /mnt/scratch 0x200000401:0x1:0x0  
/mnt/scratch/test.txt
```



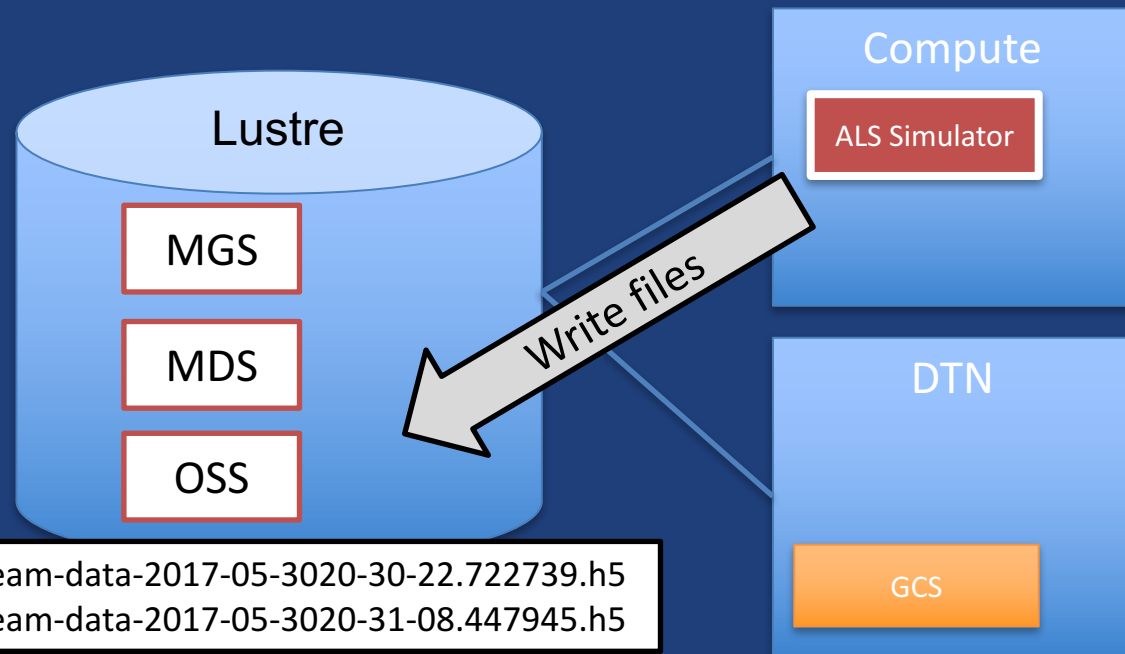

RIPPLE “Where’s My Data” testbed

- Intel Cloud Edition for Lustre
- CloudFormation deployment
- Cloud Edition v1.3.1
 - t2.micro instances
 - lustre: 2.8.0
 - kernel: patchless_client
 - build: jenkins-arch=x86_64, build_type=server, distro=el7, ib_stack=inkernel-12--PRISTINE-3.10.0-327.3.1.el7_lustre.x86_64



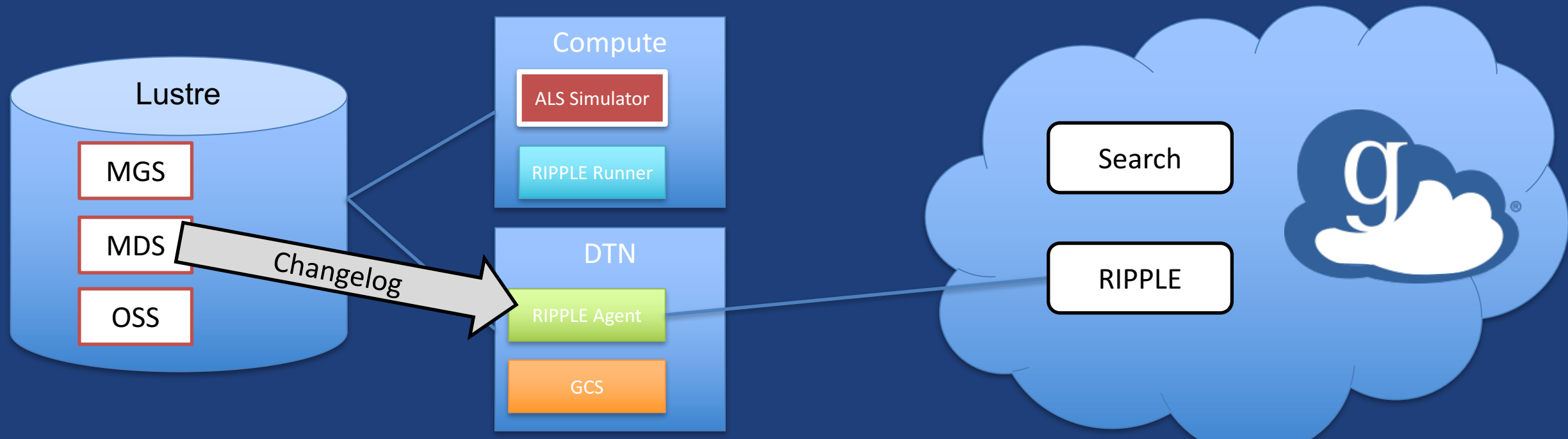
“Where’s My Data?” Use Case

- Five AWS instances: MGS, MDS, OSS, and two workers (DTN/Compute)
- ALS simulator generates “beamline” .h5 files



“Where’s My Data?” Use Case

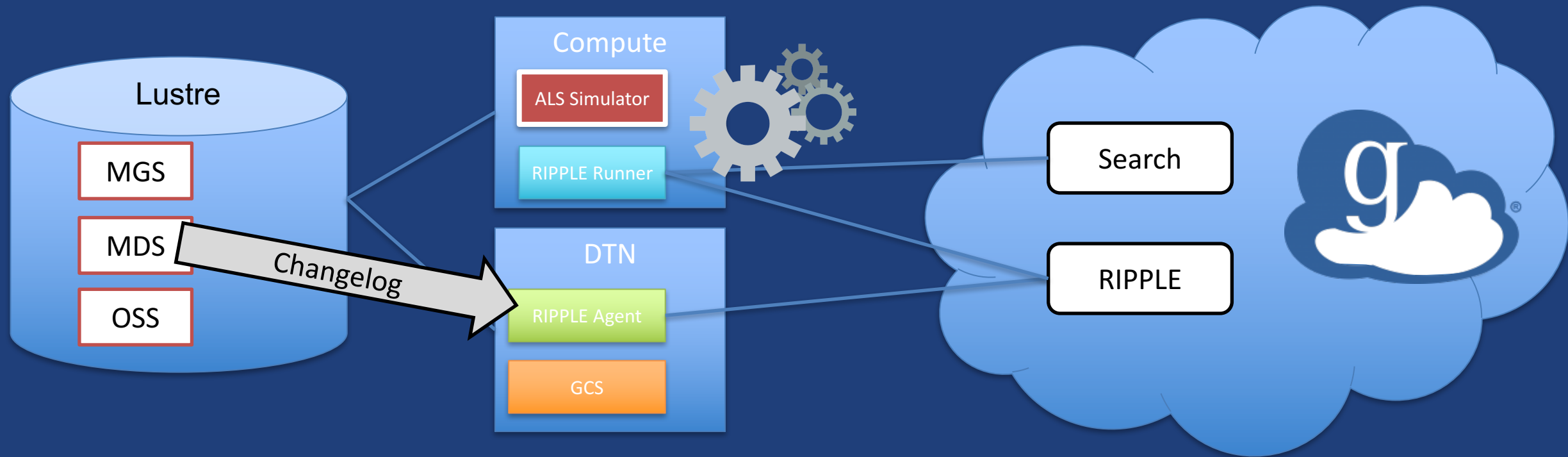
- RIPPLE is deployed on both workers.
- Detects ChangeLog events on the DTN
- Filters events locally and sends those that match rules to the cloud



```
218318 01CREAT 02:02:37 2017.05.31 0x0 t=[0x200000401:0x126d8:0x0] p=[0x200000007:0x1:0x0] beam-data...-h5
218319 11CLOSE 02:02:37 2017.05.31 0x242 t=[0x200000401:0x126d8:0x0]
```

“Where’s My Data?” Use Case

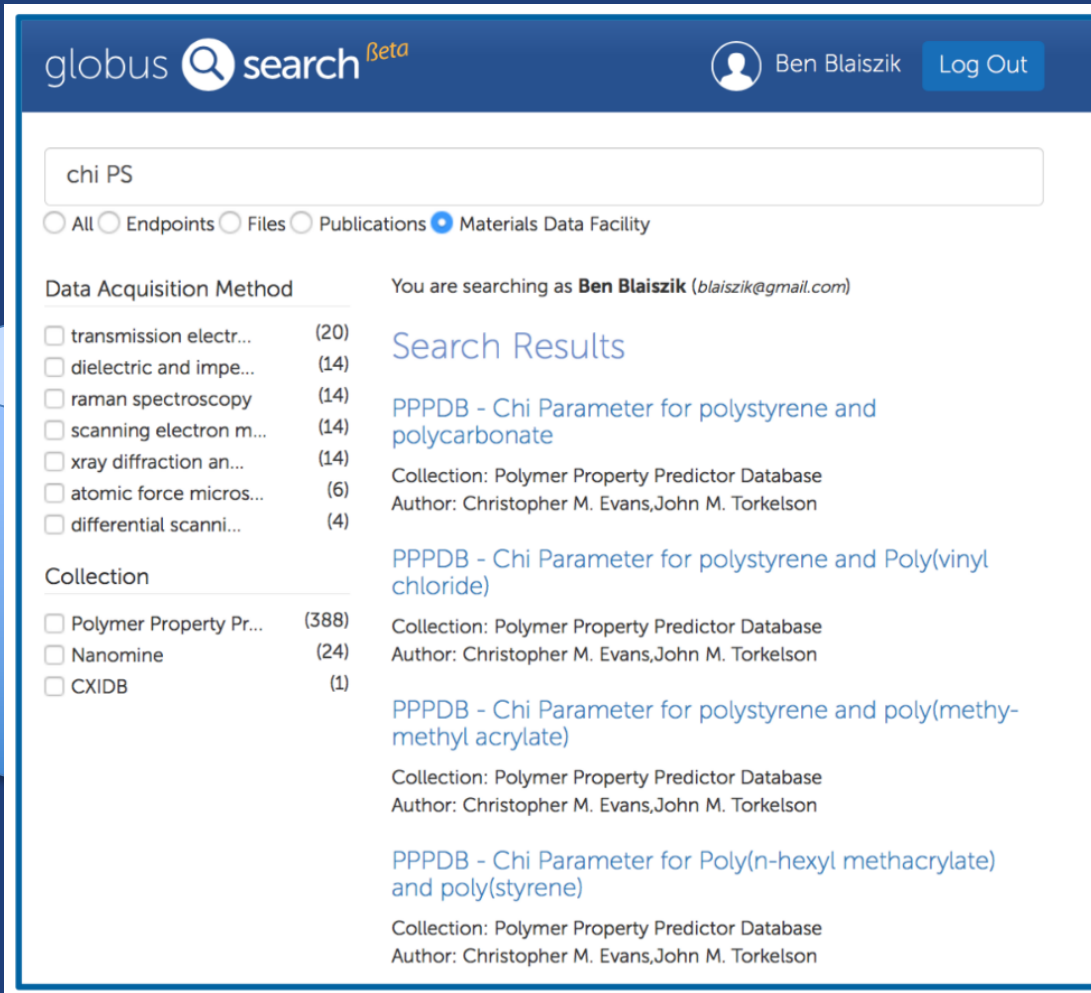
- Cloud service dispatches execution to the compute node
- Local runner uses Singularity* to: extract metadata, construct gmeta (json) blob, ingest gmeta with Globus Search API



*<http://singularity.lbl.gov>

“Where’s My Data?” Use Case

- Data is automatically indexed and can be found via Globus Search



The screenshot shows the Globus Search interface. At the top, there is a search bar containing the text "chi PS". Below the search bar, there are radio buttons for "All", "Endpoints", "Files", "Publications", and "Materials Data Facility", with "Materials Data Facility" selected. The interface is divided into two columns. The left column contains filters for "Data Acquisition Method" and "Collection". The right column shows the search results, including the title "Search Results" and several entries for "PPPDB - Chi Parameter for polystyrene and polycarbonate".

globus search ^{Beta} Ben Blaiszik Log Out

chi PS

All Endpoints Files Publications Materials Data Facility

Data Acquisition Method You are searching as Ben Blaiszik (blaiszik@gmail.com)

transmission electr... (20)
 dielectric and impe... (14)
 raman spectroscopy (14)
 scanning electron m... (14)
 xray diffraction an... (14)
 atomic force micros... (6)
 differential scanni... (4)

Collection

Polymer Property Pr... (388)
 Nanomine (24)
 CXIDB (1)

Search Results

PPPDB - Chi Parameter for polystyrene and polycarbonate
Collection: Polymer Property Predictor Database
Author: Christopher M. Evans,John M. Torkelson

PPPDB - Chi Parameter for polystyrene and Poly(vinyl chloride)
Collection: Polymer Property Predictor Database
Author: Christopher M. Evans,John M. Torkelson

PPPDB - Chi Parameter for polystyrene and poly(methyl-methyl acrylate)
Collection: Polymer Property Predictor Database
Author: Christopher M. Evans,John M. Torkelson

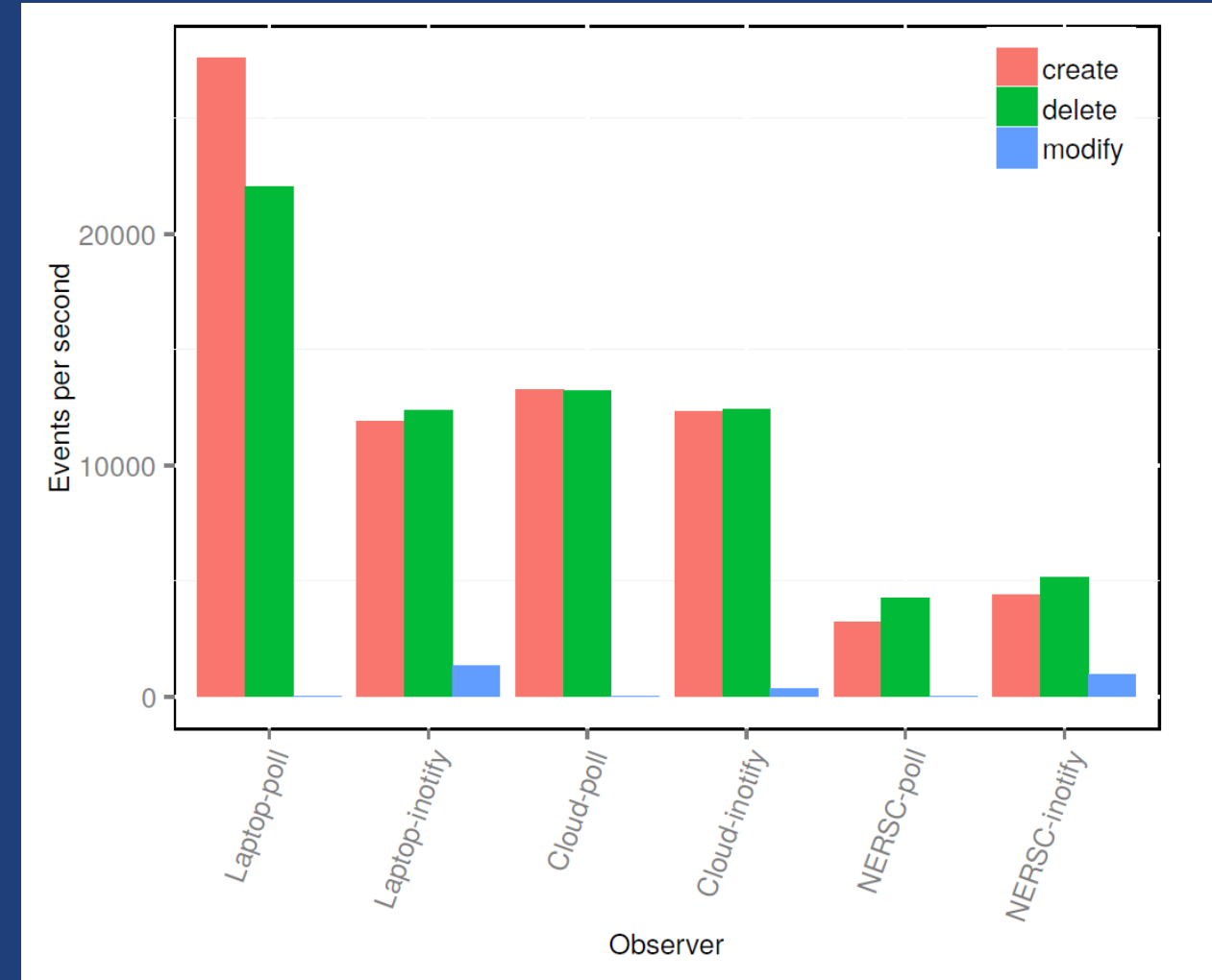
PPPDB - Chi Parameter for Poly(n-hexyl methacrylate) and poly(styrene)
Collection: Polymer Property Predictor Database
Author: Christopher M. Evans,John M. Torkelson

```
GlobusHTTPResponse({
  'count': 3,
  '@datatype': 'GSearchResult',
  'offset': 0,
  'total': 3,
  '@version': '2016-11-09',
  'gmeta': [{
    'content': [{
      'http://ripple.globus.org/#machine': 'bl832data',
      'http://ripple.globus.org/#name':
        '/mnt/scratch/beam-data....h5',
      'http://ripple.globus.org/#source': u'ALS'}}
    ]
  }
})
```



Performance evaluation

- Deployed RIPPLE Agent on:
 - laptop (XPS15)
 - AWS instance (c4.xlarge)
 - Edison login node
 - AWS Lustre (free edition – t2.micro)
- Detects and processes over 10,000 events per second
- On AWS t2.micro MDS & client ~700 events per second





Future work

- A scalable, high performance, Lustre event monitor for leadership computing resources
 - Hierarchical event reporting with ZeroMQ to monitor large stores in near real-time
 - How about configurable inotify for Lustre?
- Developing an event-driven trigger/action programming model
- Extending RIPPLE's capabilities and adding job runners
- Incorporating workflow management
- Testing at ALCF