NASA Archives and Data Stewardship in the Cloud with Cumulus

2019 Winter ESIP

Lauren Frederick
Cumulus Lead
lauren@element84.com

This work was supported by NASA/GSFC under Raytheon Co. contract number NNG15HZ39C.
This document does not contain technology or Technical Data controlled under either the U.S. International Traffic in Arms Regulations or the U.S. Export Administration Regulations.
The Problem
More than just storage
Dealing with Ingest Failures

Sync Granule → Process → Move granule → Post to CMR

Handle Failure → Report Failure
Failures to Trigger Ingest
Kinesis Failures
Data Provenance
Diagram showing interactions between different segments of a system:

- **Service Users**
- **Data Users**
- **Data Services**
- **Data Consumers**
- **NASA / EOSDIS Services**
- **DAAC Operators**
- **Data Providers**

These components are interconnected, indicating a flow of interaction or data exchange.
Operator Dashboard

CUMULUS Dashboard

Updates

1 Errors
203 Collections
41 Granules
5,313 Executions
13 Ingest Rules
STEWARDSHIP

SECURITY & COMPLIANCE

BACKUP & RECOVERY

SIMPLE ACCESS

ERROR HANDLING

PROVENANCE

SERVICES & METRICS
Granule Reconciliation

Reconciliation report

<table>
<thead>
<tr>
<th>Created</th>
<th>2018-11-01T19:39:11.633Z</th>
<th>Status</th>
<th>SUCCESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>OK file count</td>
<td>172</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Files only in DynamoDB (579)

<table>
<thead>
<tr>
<th>GranuleId</th>
<th>Filename</th>
<th>Bucket</th>
<th>S3 Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOD09GQ.A9917737.KFw8Br.006.5000644002066</td>
<td>MOD09GQ.A9917737.KFw8Br.006.5000644002066.hdf</td>
<td>If-protected</td>
<td>Link</td>
</tr>
<tr>
<td>MOD09GQ.A0007560.zuzuRq.006.0290148655705</td>
<td>MOD09GQ.A0007560.zuzuRq.006.0290148655705.hdf.met</td>
<td>If-private</td>
<td>Link</td>
</tr>
</tbody>
</table>

Files only in S3 (1231)

<table>
<thead>
<tr>
<th>Filename</th>
<th>Bucket</th>
<th>S3 Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>TMI_correction_rain.txt</td>
<td>cumulus-data-shared</td>
<td>Link</td>
</tr>
<tr>
<td>MOD09GQ.A0023605._gOahZ.006.8702444238062.hdf.met</td>
<td>If-private</td>
<td>Link</td>
</tr>
</tbody>
</table>
Validation

- Unit test and stack deployment
  - # 1872.2  🏁  eslint
  - # 1872.3  🏁  Unit tests except API
  - # 1872.4  🏁  API unit tests
  - # 1872.5  🏁  Deploy Integration Test Stack

- Integration tests
  - # 1872.6  🏁  Run integration tests

- Cleanup integration tests
  - # 1872.7  🏁  Cleanup integration tests
Come Find Us

Check us out on github!
https://github.com/nasa/cumulus

Read up
https://nasa.github.io/cumulus/

Use our packages
https://www.npmjs.com/org/cumulus
Questions?
Thank You!

Lauren Frederick

lauren@element84.com
This work was supported by NASA/GSFC under Raytheon Co. contract number NNG15HZ39C.

Raytheon

in partnership with