International Climate Network Working Group (ICNWG)

Research Data Alliance 6th Plenary Meeting
Paris
September 25, 2015
Outline

• What is ICNWG?
  – Objectives, goals, collaborators

• Implementation and Methodology to Achieve WG goals
  – Network
  – Performance
  – Data transfers
  – Data Center Architectures

• Globus integration

• Current status

• Challenges and Future Plans
About ICNWG

• Background
  – Started as an Enlighten Your Research Global award in 2013
  – https://www.enlightenyourresearch.net
  – Formed in December 2013 as ESGF working group
  – http://icnwg.llnl.gov/

• WG Goal:
  – Improve data transfer performance between major climate data facilities to improve time-to-solution for climate research

• WG project was a lesson out of CMIP5 project:
  – Legacy transfer tools were not suited for moving large-scale data sets to/from climate repositories
  – Need to augment current tools (http/wget) to scale to CMIP6 (and other climate projects) data volumes
  – Technologies exist to significantly improve performance
    • Science DMZ
    • Globus
ICNWG Collaborators
WG Objectives

• Near-term
  – implement replication workflow using GridFTP
  – Increase network performance to 500MB/sec
  – Increase transfer performance to 500MB/sec (4Gbps) disk to disk—or 1 PB/month

• Longer-term
  – Increase transfer performance to 1GB/sec (8Gbps) disk to disk
  – Increase transfer performance 2GB/sec (16Gbps) disk to disk by June 2016 (stretch goal)—or 1 PB/week
WG Implementation

- Verify network path
WG Implementation

- Track the health of network connections

ESnet to ICNWG Site Packet Loss Testing

- Loss rate is <= 0.001
- Loss rate is >= 0.001
- Loss rate is >= 0.1
- Unable to retrieve data
- Check has not yet run

Average Loss is 0.004%
Average Loss is 0.000%
With loss, high performance beyond metro distances is essentially impossible.
WG Implementation

• Analyze end systems, infrastructure and tools
  – What data transfer protocols are they using?
  – Do they have separate servers for this?
  – Do they have a “Science DMZ”?
Science DMZ

- Border Router
  - perfSONAR
  - Clean, High-bandwidth WAN path
  - Site / Campus access to Science DMZ resources
- Science DMZ Switch/Router
  - Per-service security policy control points
- High performance Data Transfer Node with high-speed storage

Paths:
- High Latency WAN Path
- Low Latency LAN Path

ICNWG | International Climate Network Working Group
Data Transfer Tool

• Globus being integrated into ESGF
  – GridFTP-based data transfer protocol
  – Globus software pieces for ESGF are done
  – What remains is to integrate into deployed ESGF data nodes
  – Globus I/O nodes attach directly to filesystem, serve data directly
Current Status and Progress

- **Near-term**
  - Implement replication workflow using dedicated GridFTP servers (some sites complete)
  - Achieved network performance at 500MB/sec
  - Achieved transfer performance to 500MB/sec (4Gbps) disk to disk—or 1 PB/month

- **Site-by-site progress**
  - **LLNL:**
    - Science DMZ already implemented
    - Deployed 4 dedicated data transfer nodes
  - **NCI:**
    - Traffic rerouted to traverse US to get to Europe—more bandwidth available at 100Gbps
    - Science DMZ implemented
  - **BADC:**
    - Implemented Science DMZ and saw more than a 10% increase in data transfer performance
    - Deploying dedicated GridFTP servers
  - **DKRZ:**
    - Working on infrastructure and systems upgrade
  - **KNMI:**
    - Dutch network provided 10G upgrade
    - ESGF node to be deployed summer 2016
Challenges

- Storage performance is still the bottleneck

<table>
<thead>
<tr>
<th>globus-url-copy of 50G.dat (single 50GB file) to /dev/null</th>
</tr>
</thead>
<tbody>
<tr>
<td>anl-diskpt1.es.net</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>anl-diskpt1.es.net</td>
</tr>
<tr>
<td>bnl-diskpt1.es.net</td>
</tr>
<tr>
<td>cern-diskpt1.es.net</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>globus-url-copy of 50G.dat (single 50GB file)</th>
</tr>
</thead>
<tbody>
<tr>
<td>anl-diskpt1.es.net</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>anl-diskpt1.es.net</td>
</tr>
<tr>
<td>bnl-diskpt1.es.net</td>
</tr>
<tr>
<td>cern-diskpt1.es.net</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>globus-url-copy of 50G data set (651 directories, 1875 files total, 625 files each of size 10MB, 20MB, and 50MB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>anl-diskpt1.es.net</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>anl-diskpt1.es.net</td>
</tr>
<tr>
<td>bnl-diskpt1.es.net</td>
</tr>
<tr>
<td>cern-diskpt1.es.net</td>
</tr>
</tbody>
</table>
Future Directions

- Best practices for other ESGF data centers

1) Remote DTN conducts data search using ESGF portal
2) ESGF portal returns download script or Globus job with URLs pointing to Science DMZ DTNs
3) User runs download script on remote DTN, or submits Globus job for data transfer
4) Data is transferred directly from Science DMZ DTNs to remote DTN without further portal involvement
Future Directions

- Other organizations joining the working group:
  - IPSL, NOAA, NASA
Questions? Comments?

Research Data Alliance 6th Plenary Meeting
Paris
September 25, 2015