

High Speed Data Transfer Implementations at Geosciences Node

Dan Scholes

2/12/2018

Historically Large Data Transfers

- PDS Data Deliveries
 - Typically transferred via FTP or data brick
- Geo archive - user downloads
 - Use HTTP & FTP
- Orbital Data Explorer website (ODE)
 - Backend process - automated cart fulfillment of files hosted by other nodes
 - Done by HTTP
 - User downloads (cart & direct product downloads)
 - Done by HTTP & FTP

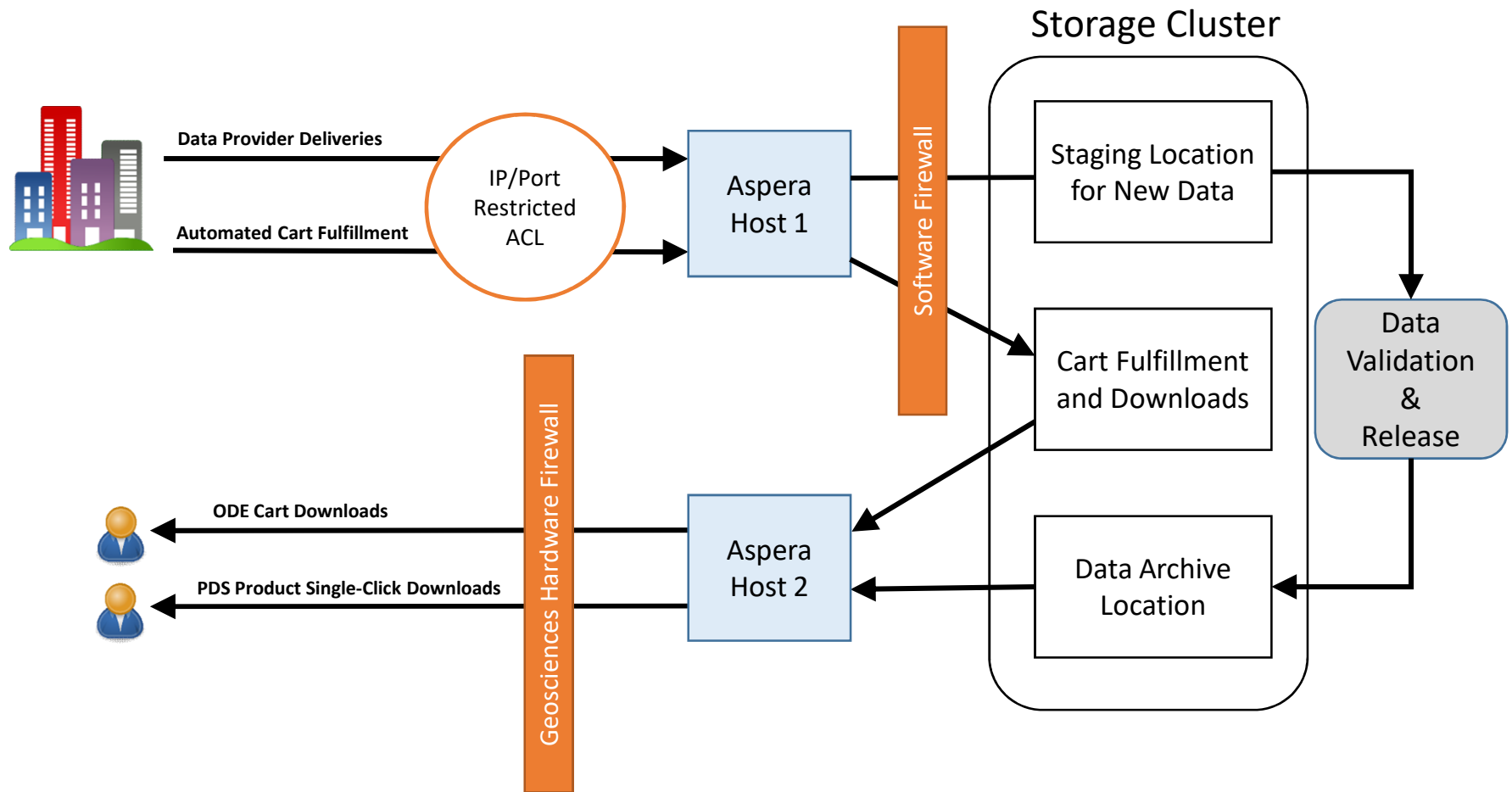
Washington University Research Network (WURN)

- Support on-campus and off-campus big data collaboration
 - High speed research network within the university (up to 40 Gbps)
 - High speed Internet2 access (10 Gbps)
- Licensing for high-speed data transfer software
- Geosciences Node participation since pilot program

Aspera

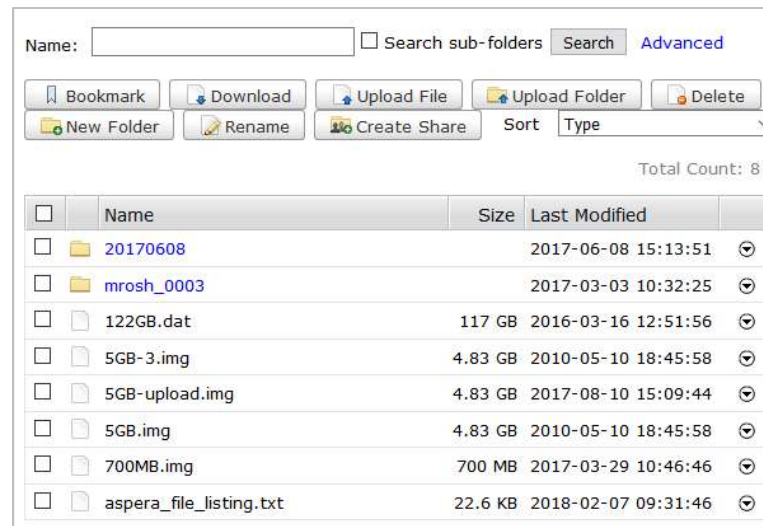
- High speed data transfer software platform
 - <http://asperasoft.com>
- Proprietary FASP (Fast Adaptive Secure Protocol) data transfer protocol
 - High latency tolerant
 - More effective use of bandwidth compared to standard HTTP/FTP
- Includes basic built-in Aspera web interface & web API
- Client application options
 - Command line
 - Web browser plug-in

Geosciences Node Aspera Setup



Using Aspera for Data Provider Deliveries to Geo Node

- Improved convenience and speed
- Current average 600-700 Mbps
 - Transfer 1TB ~3 hours via Aspera, one week to ship on data brick
 - Current speed limitations have been on client-side, with no client-side optimization
- Example users
 - JHU APL - MRO CRISM
 - GSFC - LRO LOLA
 - UCLA - LRO Diviner



The screenshot shows a file manager interface with a search bar at the top, a toolbar with buttons for Bookmark, Download, Upload File, Upload Folder, Delete, New Folder, Rename, and Create Share, and a table of files and folders. The table has columns for Name, Size, and Last Modified. The total count of items is 8.

	Name	Size	Last Modified	
<input type="checkbox"/>	20170608		2017-06-08 15:13:51	⊕
<input type="checkbox"/>	mrosh_0003		2017-03-03 10:32:25	⊕
<input type="checkbox"/>	122GB.dat	117 GB	2016-03-16 12:51:56	⊕
<input type="checkbox"/>	5GB-3.img	4.83 GB	2010-05-10 18:45:58	⊕
<input type="checkbox"/>	5GB-upload.img	4.83 GB	2017-08-10 15:09:44	⊕
<input type="checkbox"/>	5GB.img	4.83 GB	2010-05-10 18:45:58	⊕
<input type="checkbox"/>	700MB.img	700 MB	2017-03-29 10:46:46	⊕
<input type="checkbox"/>	aspera_file_listing.txt	22.6 KB	2018-02-07 09:31:46	⊕

Aspera in ODE Automated Cart Fulfillment

- Remote script to upload requested files to Geo
 - Component of ODE cart request processing
 - ODE REST services provide list of requested files
- Current implementation
 - JPL Imaging Node - Thank you Jordan!
 - Over twenty fold improvement in transfer
 - Average 850 Mbps transfer rate
 - 10GB - previously 45 minutes, now < 2 minutes
- Pending implementation
 - LROC Data Node

ODE Product Level Download using Aspera

- Uses Aspera browser plug-in
- Single-click download of product files
 - Saved to a local directory matching product id

Aspera Connect Browser Plug-in Option - New Option
Only available for data hosted by the PDS Geosciences Node.
Note: Installing the browser plug-in will require restarting the browser.

[Download All Product & Derived Files with Aspera](#) [Aspera Download Help](#)

Downloading: frt00007fa4_07_sr165j_mtr3.img (37%)

PDS Product Files	Derived Files
Product Files & Labels KB	
frt00007fa4_07_if165j_mtr3.hdr <i>Product Data File</i>	15
frt00007fa4_07_if165j_mtr3.img <i>Product Data File</i>	1,065,238
frt00007fa4_07_if165j_mtr3.lbl <i>Product Label File</i>	

Transfers - Aspera Connect

frt00007fa4_07_if165j_mtr3.hdr (+94)
C:/Users/scholes/Downloads/FRT00007FA4_07_IF165J_MTR3/frt00007fa4_07_if165j_mtr3.img




Downloading 668.0 / 1,015.9 MB (575.88 Mbps) - 0:05 remaining
0:11 elapsed

ODE Cart Download Page using Aspera

- Uses Aspera browser plug-in
- Allows single-click download of entire cart request
 - Zip or other packaging not required
- Supports individual file download links

1 GB File Download		
Aspera	410 Mbps	18.6 sec
HTTP	36 Mbps	3.5 min
FTP	22.4 Mbps	5.7 min

Test run from JHU APL workstation

Aspera One Click - Download Entire Cart Request directory (sample)				
HTTP Link -		http://ode.rsl.wustl.edu/odecartrequest/sample		
FTP Link -		ftp://geoftp.wustl.edu/sample	Username: geoftp	Password: Odeuser1
Download Specific Directories and Files				
Browsing: sample/				
Type	Name	Aspera Download	HTTP Download	
	CH1M3_0004_part_0005.zip	Aspera Download Complete	Aspera	HTTP
	CH1M3_0004_part_0011.zip	Aspera Downloading (67%)	Aspera	HTTP
	CH1M3_0004_part_0010.zip		Aspera	HTTP

Next Steps

- Explore connectivity with additional data nodes
- Support high-speed ODE cart download from command line
- Support high-speed access to entire Geo archive

- Due to expense, evaluating licensing options
 - Wash U. negotiating with Aspera
 - Exploring other options (e.g. Globus)

Backup Slides

Geosciences Node's Aspera Setup

- Aspera host servers
 - Host 1 - Directly connected to WURN and Internet2
 - Secure - limited to IP address, specific users, and ports
 - Data provider access
 - Node to node data transfers (automated cart fulfillment)
 - Host 2 - Behind firewall, standard Internet connection
 - Website users - read only
- Aspera client applications
 - Command-line application
 - Data providers
 - Node to node data transfers (automated cart fulfillment)
 - Web browser plug-in
 - Data providers
 - Public web users