

CADC/CANFAR

An Integrated Science Platform for JCMT Users

JJ Kavelaars
JJ.Kavelaars@nrc-cnrc.gc.ca
JCMT Archive Scientist

Canadian Astronomy Data Centre

- National facility for open access
- Telescope collections:
 - Multiple missions, facilities and wavelengths
 - Pointed and survey observations
 - 12 telescopes
 - 6 advanced data collections
- Services
 - Archive services
 - Data curation
 - Community projects
- Many international collaborations
- Development and operations hub for CANFAR





Advanced Search

Search Results Error ADQL Help

Search Reset

Click on ? for explanations

Observation Constraints

- ▶ Observation ID
- ▶ P.I. Name
- ▶ Proposal ID
- ▶ Proposal Title
- ▶ Proposal Keywords
- ▶ Data Release Date

Science and Calibration data

Spatial Constraints

- ▶ Target
- ▶ Pixel Scale
- ☐ Do Spatial Cutout

Temporal Constraints

- ▶ Observation Date
- ▶ Integration Time
- ▶ Time Span

Spectral Constraints

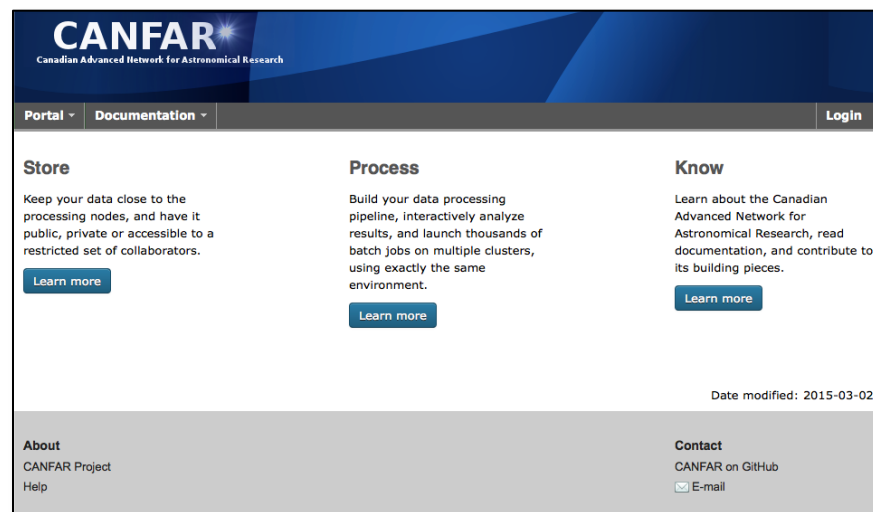
- ▶ Spectral Coverage
- ▶ Spectral Sampling
- ▶ Resolving Power
- ▶ Bandpass Width
- ▶ Rest-frame Energy
- ☐ Do Spectral Cutout

Additional Constraints						
Band	Collection	Instrument	Filter	Calibration Level	Data Type	Observation Type
All (7) Gamma-ray Infrared Millimeter Optical Radio UV Unknown	All (4) JCMT JCMTLS CGPS BLAST	All (30) FTS2-SCUBA-2 HARP-ACSIS POL-HARP-ACSIS POL-RXA3-ACSIS POL2-FTS2-SCUBA-2 POL2-SCUBA-2 RXA3-ACSIS RXA3M-ACSIS RXWB-ACSIS RXWD2-ACSIS SCUBA-2	All (24) 0.35MB 0.35um 0.45MB 0.45um 0.75um 0.85um 1.3um 1.4um 1350um 2.0um 2000um	All (5) (3) Product (2) Calibrated (1) Raw Standard (0) Raw Instrumental Unknown	All (5) catalog cube image Other spectrum	All (12) dream flatfield focus grid jiggle noise null pointing scan setup skydip

Search Reset

Canadian Advanced Network for Astronomical Research

- A cloud ecosystem for data intensive astronomy
- A platform supporting many virtual organisations
- User services
 - Store and share data
 - Create and share VMs
 - Run VMs close to data
 - Interactive for data exploration
 - Persistent for SaaS
 - Batch processing in Virtual Clusters
- Federated research cloud resources
 - Compute Canada
- Integrated:
 - Authentication and authorization
 - Access to telescope data
 - Access to user storage
- In operation since 2011



University
of Victoria



University of
British Columbia

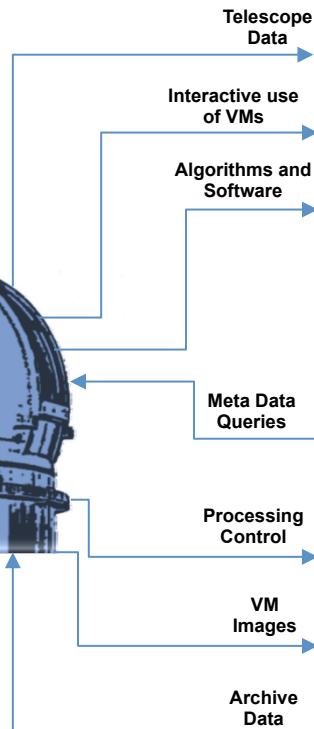
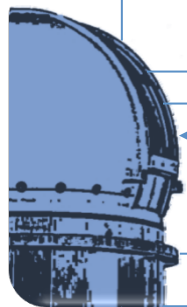
canarie



compute + calcul
CANADA



TELESCOPE CLIENT



SERVICES

CANFAR/CADC

Compute Canada Victoria Compute Canada Saskatoon NRC Herzberg Victoria Compute Canada Sherbrooke



STORAGE MANAGEMENT
2.6 PETABYTES

META DATA MANAGEMENT
8.5 TERABYTES



Meta Data

PROCESSING MANAGEMENT
400 CORES-YEARS



Compute Canada Victoria Compute Canada Calgary Compute Canada Sherbrooke

UNIVERSITY RESEARCHER CLIENT

CADC

Archive Data

Meta Data Queries

User Data

VM Images

VM Control

Interactive use of VMs

Processing Control

VM Service Creation and Deployment



SERVICES

CANFAR

Key Data Activities

- Data engineering
- Operations and user support
- Software development
- Software integration
- Data processing
- Data management
- User web services
- User web interfaces

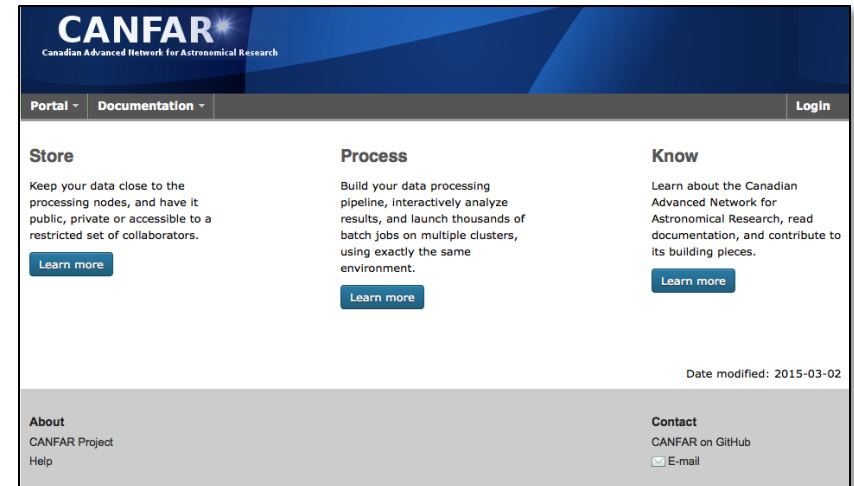
University researchers and telescope staff have privileges to upload data, create VMs and install science applications, run interactive VM sessions, submit batch processing jobs to VMs, share their VMs, control the life-cycle for their VMs, offer software-as-a-service applications in their VMs.

Definition: VM – Virtual Machine

	Data In		Data Out	
	# of files	Terabytes	# of files	Terabytes
Peak per day	2,169,190	8.0	648,093	16.8
Avg per day	130,952	0.4	99,253	2.6

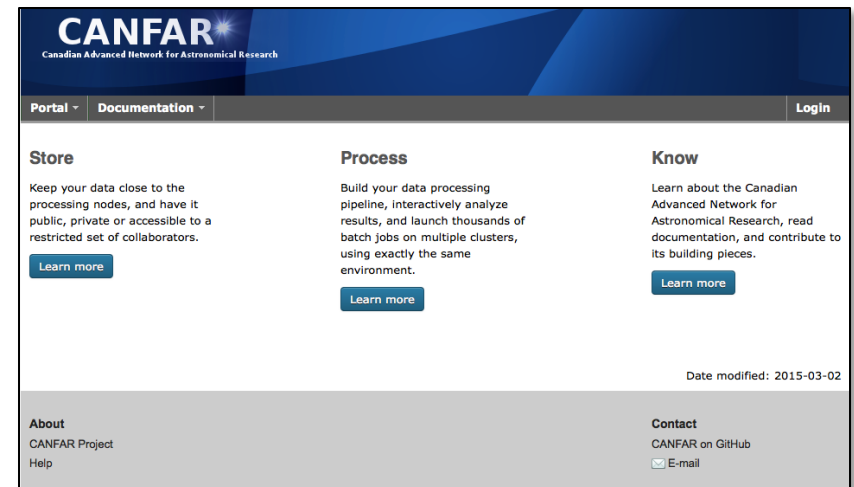
CANFAR/CADC 2014

- Size:
 - 233M files (932M files)
 - 597 TiB (2.3 PiB)
- Users
 - Authenticated access: 762
 - Anonymous access: 7,544
 - Registered: 7,018
- Data handled in the last year
 - TiB: 1,106
 - Files: 91M
- Batch processing
 - 488 Core-years
 - 2.7M jobs



CADC JCMT ACCESS IN 2015

- Size:
 - 16.1 M files
 - 213 TiB
- Users
 - Authenticated access: 158
 - Anonymous access: N/A



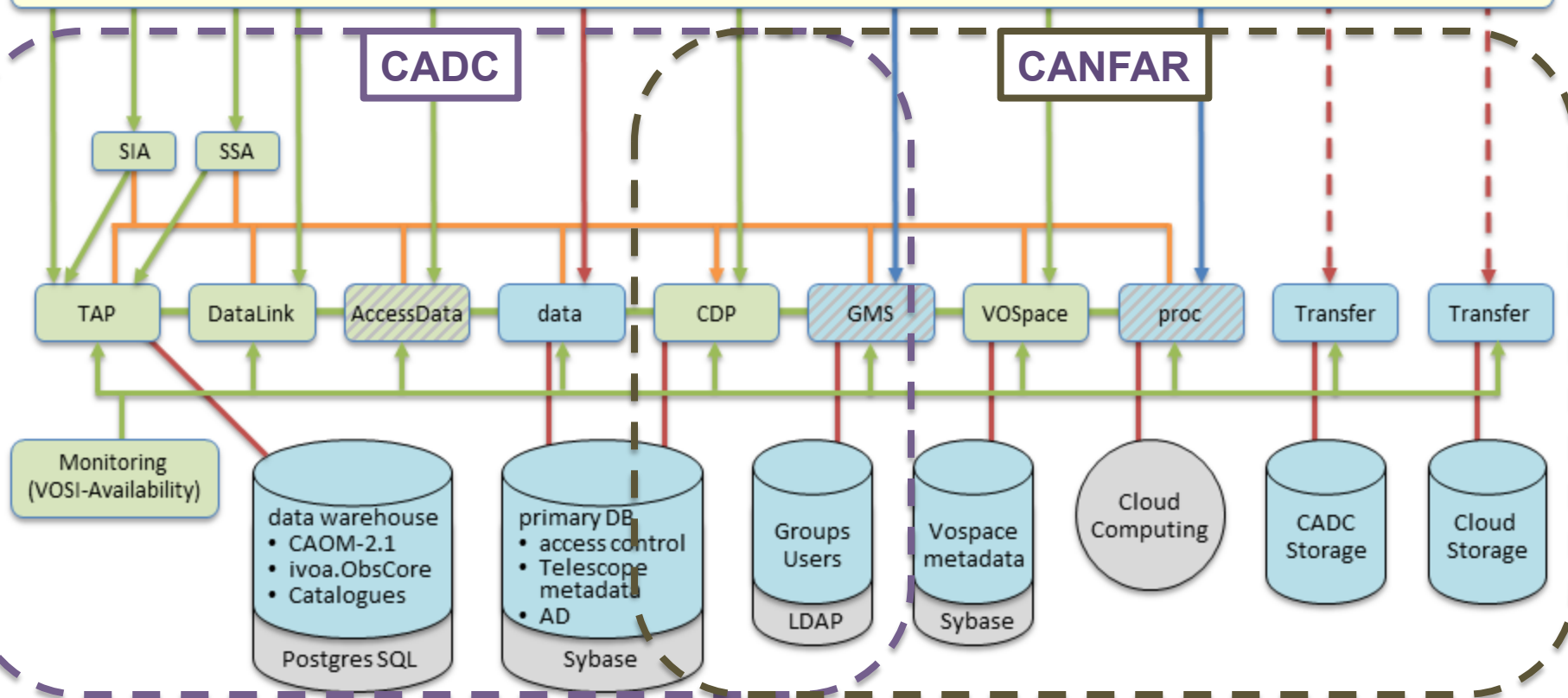
International Virtual Observatory Standards

- ADQL 2.0
- CDP 1.0
- DALI 1.0
- DataLink 1.0
- ObsCore 1.0
- SIA 1.0
- SimpleDALRegExt 1.0
- SSO 1.01
- TAP 1.0
- TAPRegExt 1.0
- UWS 1.0
- VODataService 1.1
- VOResource 1.03
- VOSI 1.0
- VOSpace 2.0
- VOTable 1.1, 1.2, 1.3
- RegistryInterfaces 1.0

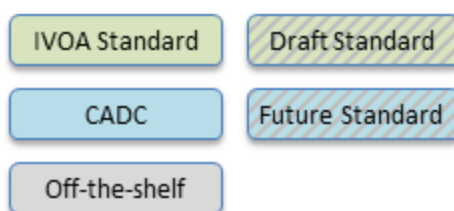
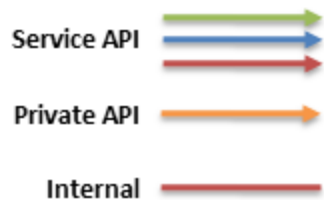
Users: client applications, browser-based applications, scripts & tools, etc

CADC

CANFAR

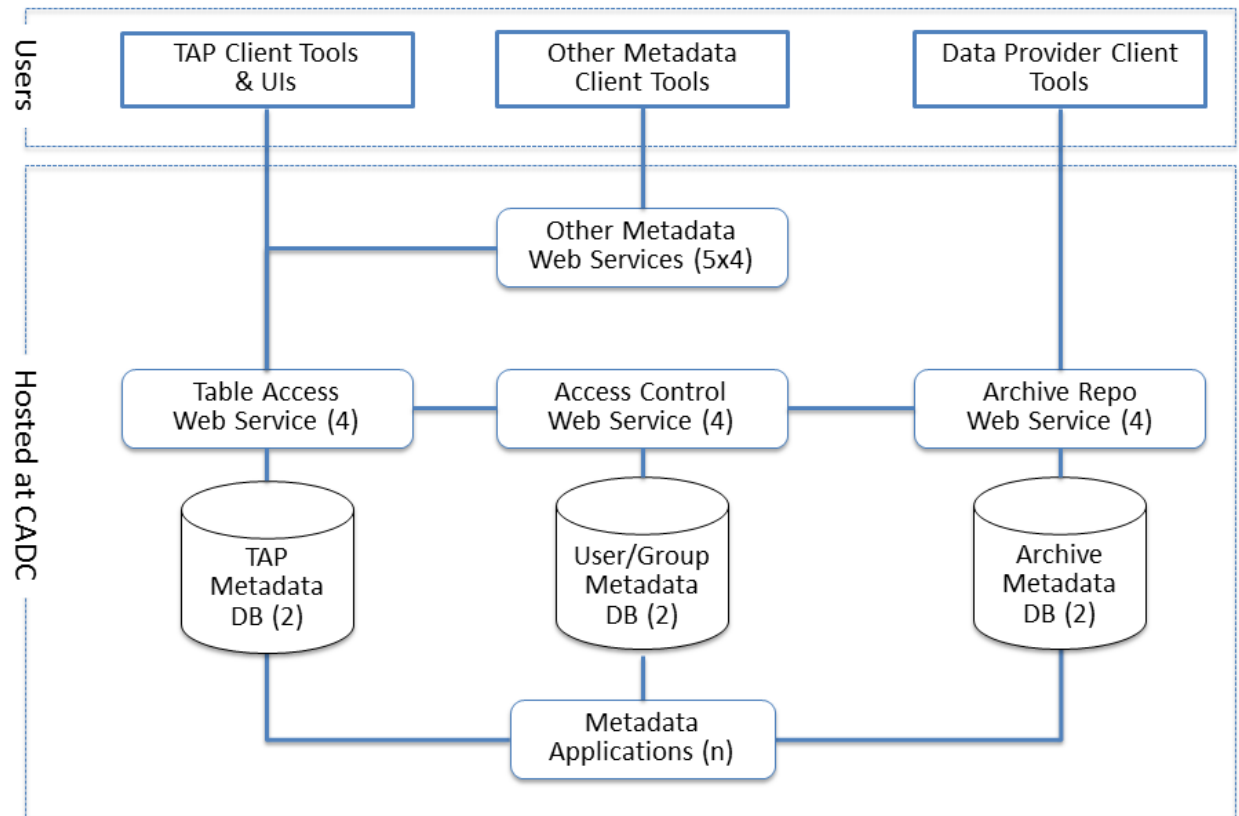


Legend



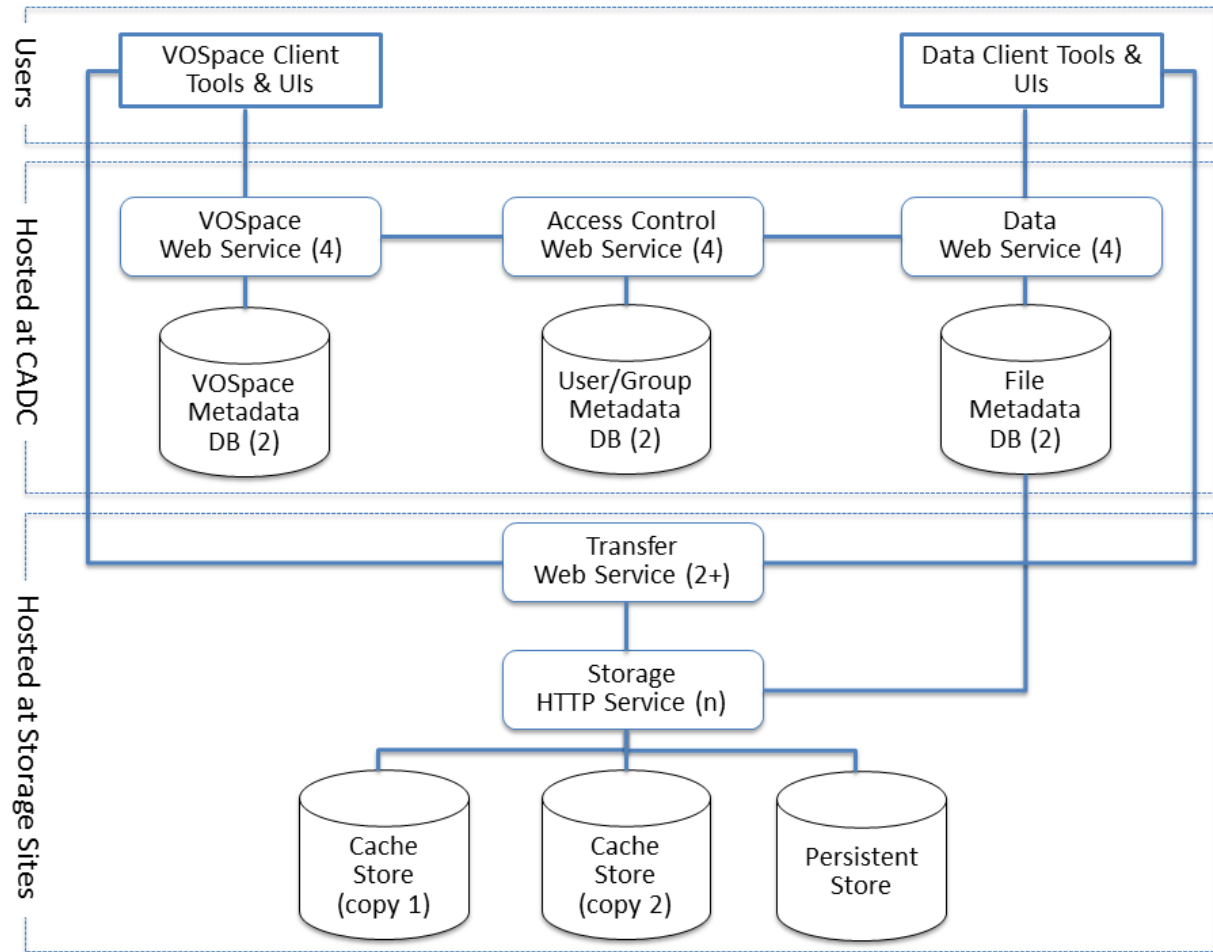
Publishing Metadata

- Supported by:
 - python and java libraries
 - java application
- Web service for persistence and retrieval
- [Google Code repository](#)
- Tools for user contributed publishing



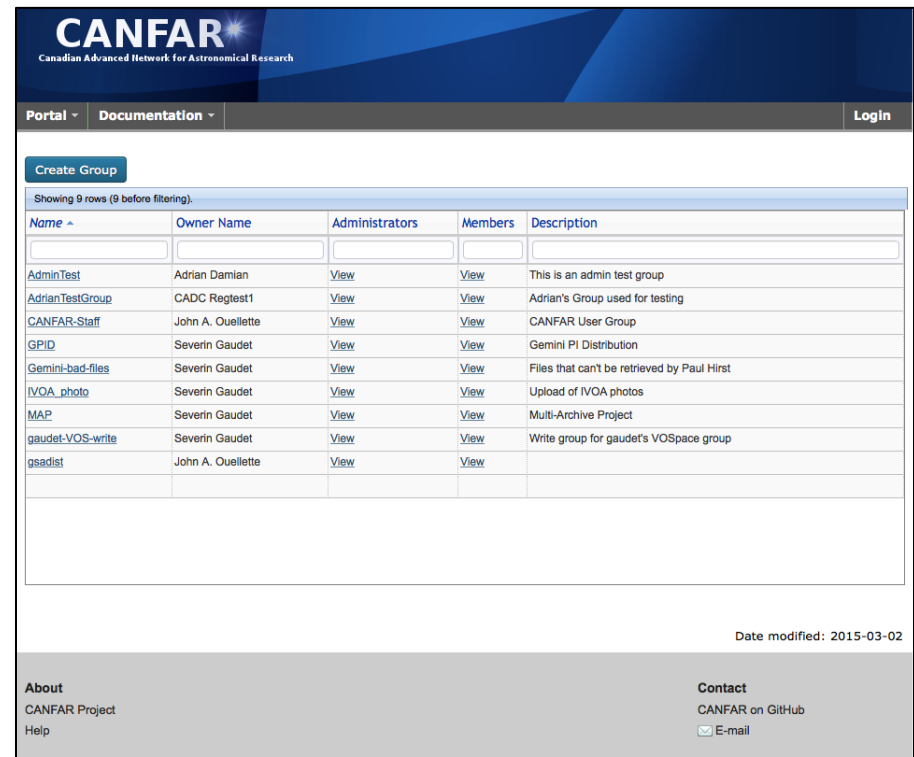
Data services for users and telescopes

- Storage web services using several distributed storage resources
- Optimization and QoS strategies not user nor provider dependent
- Same system for both archive and users
- File in/File out



CANFAR Access Control and Group Management

- Project, team or user managed
- Processing, storage, querying
- Group Management Service
 - PI can assign data access rights to others
 - Same groups provide access to VOSpace storage areas, can be for projects or individuals



The screenshot shows the CANFAR web interface for group management. The header includes the CANFAR logo and the tagline "Canadian Advanced Network for Astronomical Research". Navigation tabs for "Portal" and "Documentation" are present, along with a "Login" link. A "Create Group" button is located above a table. The table displays a list of groups with columns for Name, Owner Name, Administrators, Members, and Description. Below the table, there are links for "About" and "Contact".

CANFAR
Canadian Advanced Network for Astronomical Research

Portal Documentation Login

Create Group

Showing 9 rows (9 before filtering).

Name ^	Owner Name	Administrators	Members	Description
AdminTest	Adrian Damian	View	View	This is an admin test group
AdrianTestGroup	CADC Regtest1	View	View	Adrian's Group used for testing
CANFAR-Staff	John A. Ouellette	View	View	CANFAR User Group
GPID	Severin Gaudet	View	View	Gemini PI Distribution
Gemini-bad-files	Severin Gaudet	View	View	Files that can't be retrieved by Paul Hirst
IVOA_photo	Severin Gaudet	View	View	Upload of IVOA photos
MAP	Severin Gaudet	View	View	Multi-Archive Project
gaudet-VOS-write	Severin Gaudet	View	View	Write group for gaudet's VOSpace group
gsadist	John A. Ouellette	View	View	

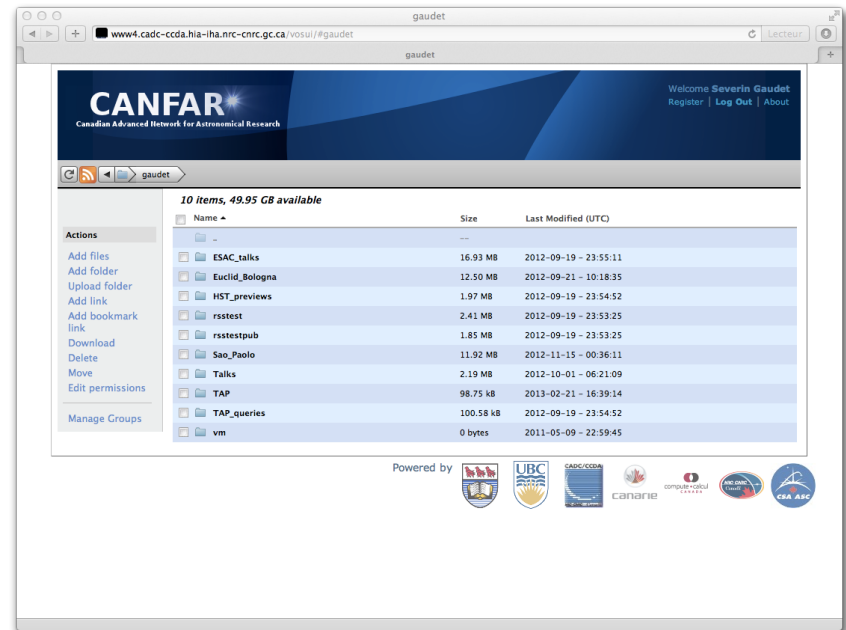
Date modified: 2015-03-02

About
CANFAR Project
Help

Contact
CANFAR on GitHub
[E-mail](#)

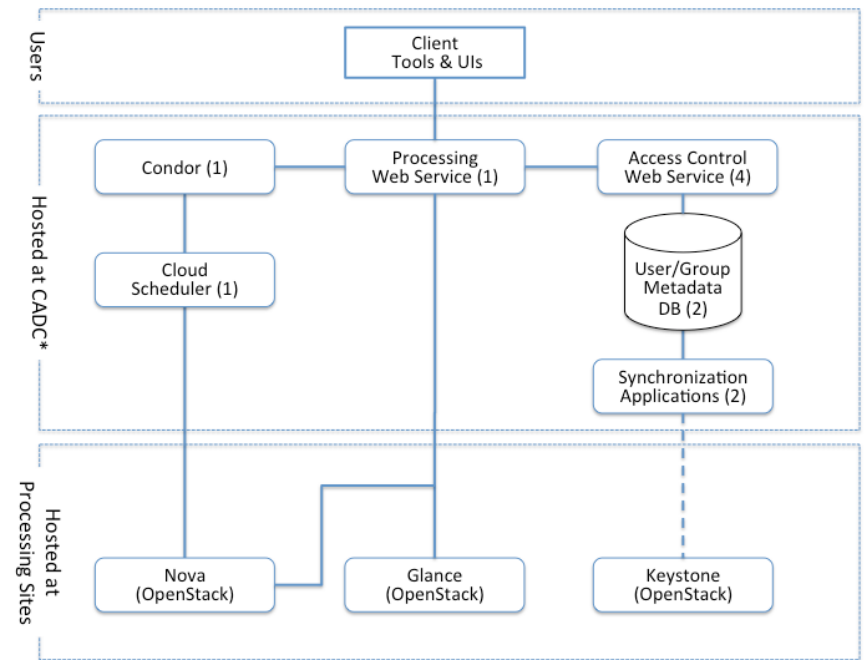
CANFAR VOSpace

- User collaboration storage
- Processing file persistence
- Asynchronous query input and output
- Browser UI
- Python clients
- Mountable file system: mountvofs
- <https://github.com/canfar/vos>
- Full access control
- Notifications via RSS feeds



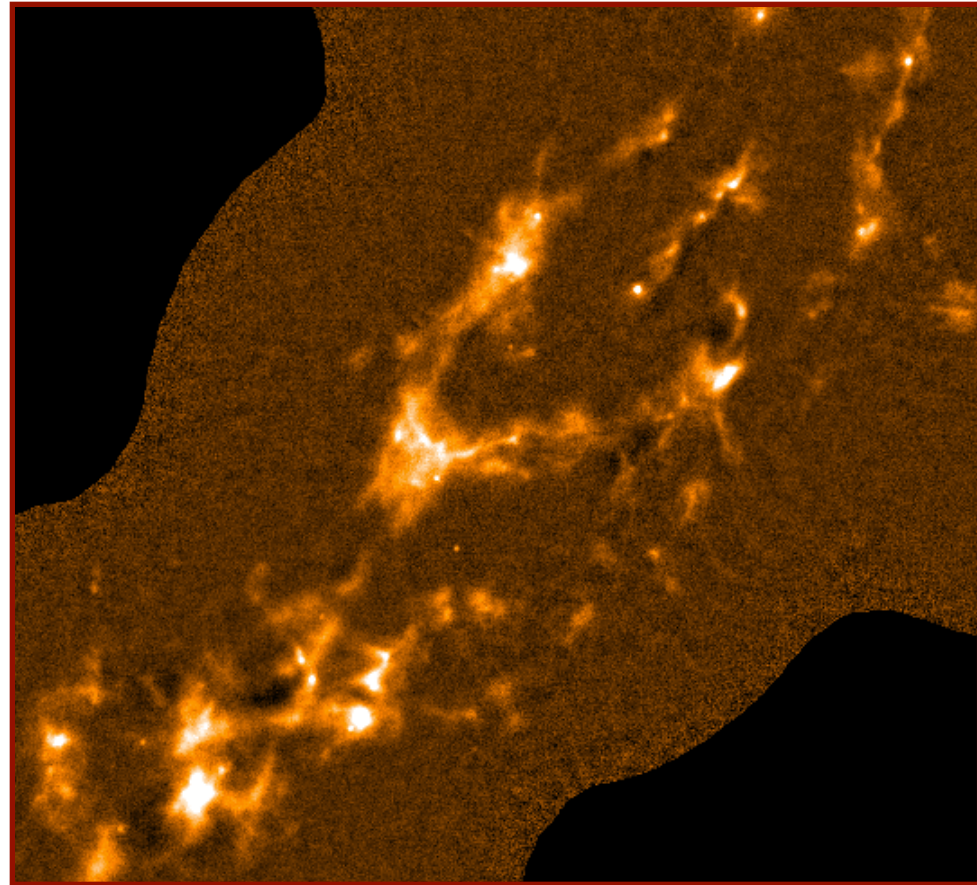
CANFAR Processing

- User create processing: VMs and containers
- Moving code to data
 - Processing resources co-located with archive and user storage
- Modes
 - Interactive
 - Persistent
 - Batch (Condor and Cloud Scheduler)
- Integrated access control
 - Sharing
 - Credential propagation for data access



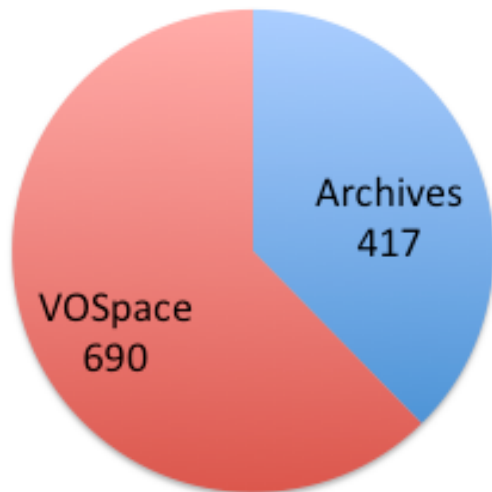
CANFAR Processing - Example

- Helen Kirk processing Gould Belt observations using CANFAR.
- Tested processing in interactive mode, added internal checks on processing success
- Batch processing on CANFAR using 120GB RAM / 32 core processors



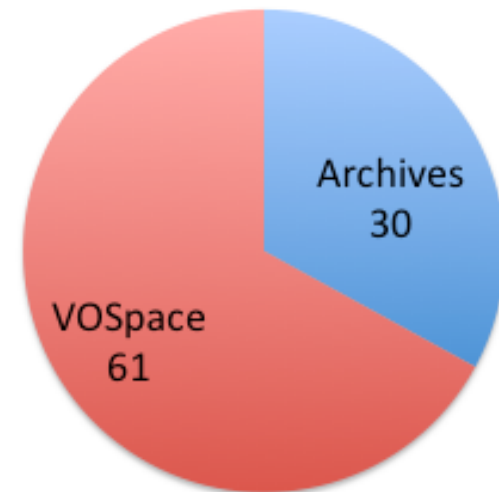
VOSpace usage in 2014

TiB



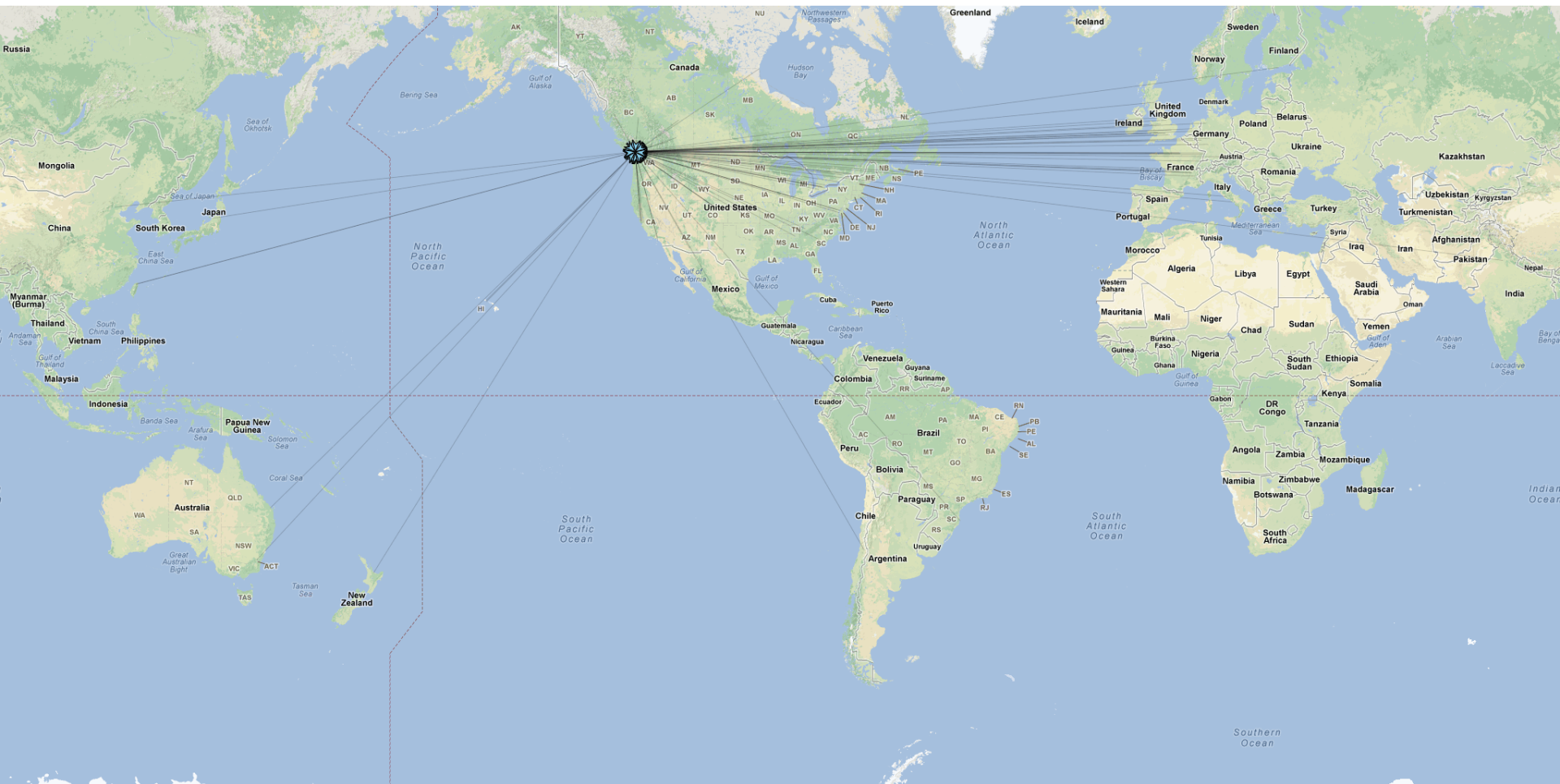
62% VOSpace
Average per week: 13.3
Peak week: 39.0

Files (M)

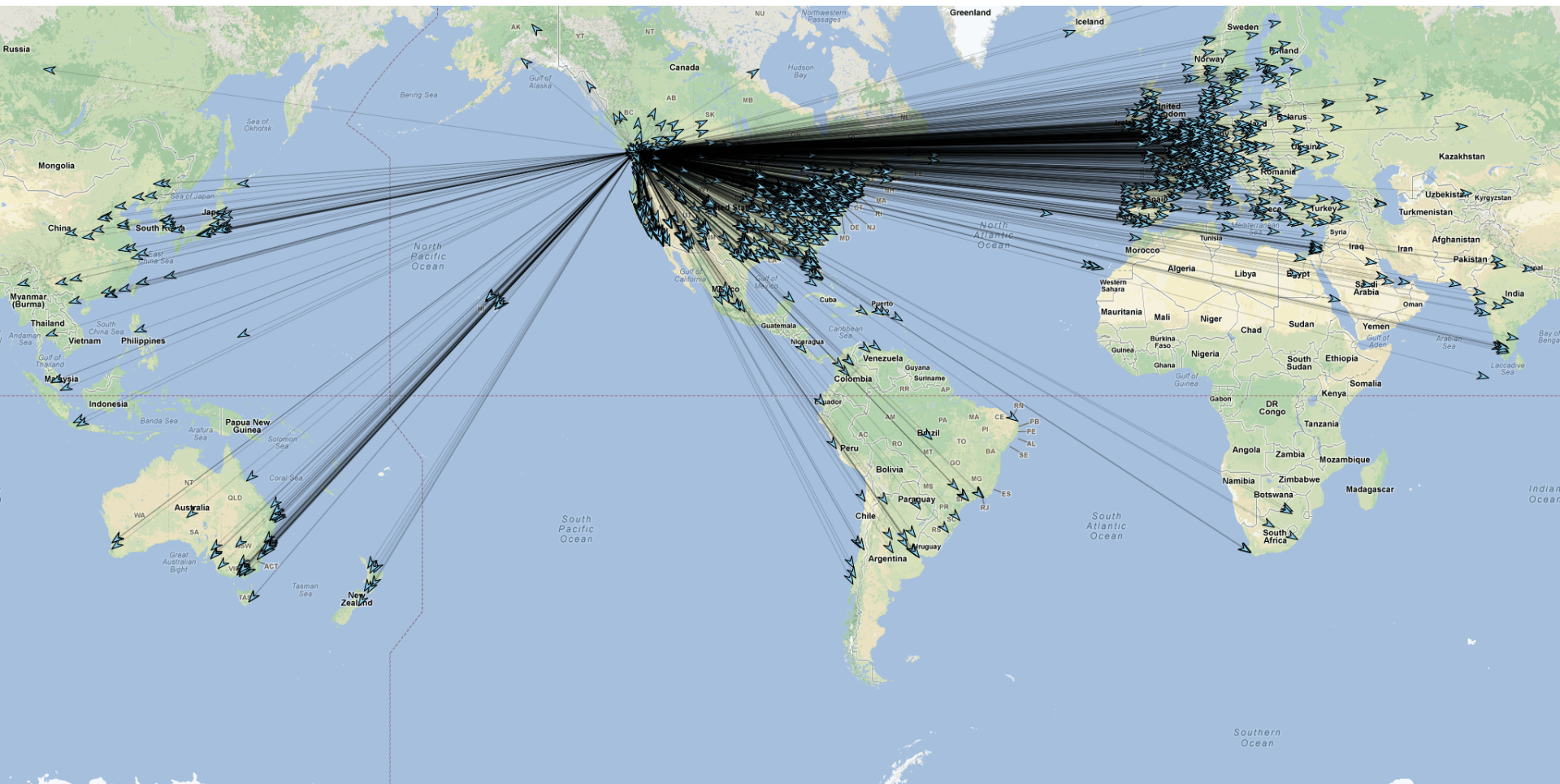


67% VOSpace
Average per week: 1.2M
Peak week: 11.7M

Geography of VOSpace PUTs



Geography of VOSpace GETs





CADC Home

[Advanced Search](#)

Telescope Data Products



[Gemini](#)



[CFHT](#)



[JCMT](#)



[HST](#)



[BLAST](#)



[MOST](#)



[DAO](#)



[MACHO](#)



[OMM](#)



[FUSE](#)



[UKIRT](#)

Advanced Data Products



[MegaPipe](#)



[HLA](#)



[IRIS](#)



[CGPS](#)



[CFHTLS](#)



[WIRwolf](#)

Services



[Meetings](#)



[Community](#)



[SSOIS](#)



[CANFAR](#) 

Date modified: 2014-04-28

