

Data Distribution Working Group

May 26 2021

Kevin Krieger¹, Kevin Sterne², Paul Breen³

SuperDARN Workshop

Virtual - SANSA

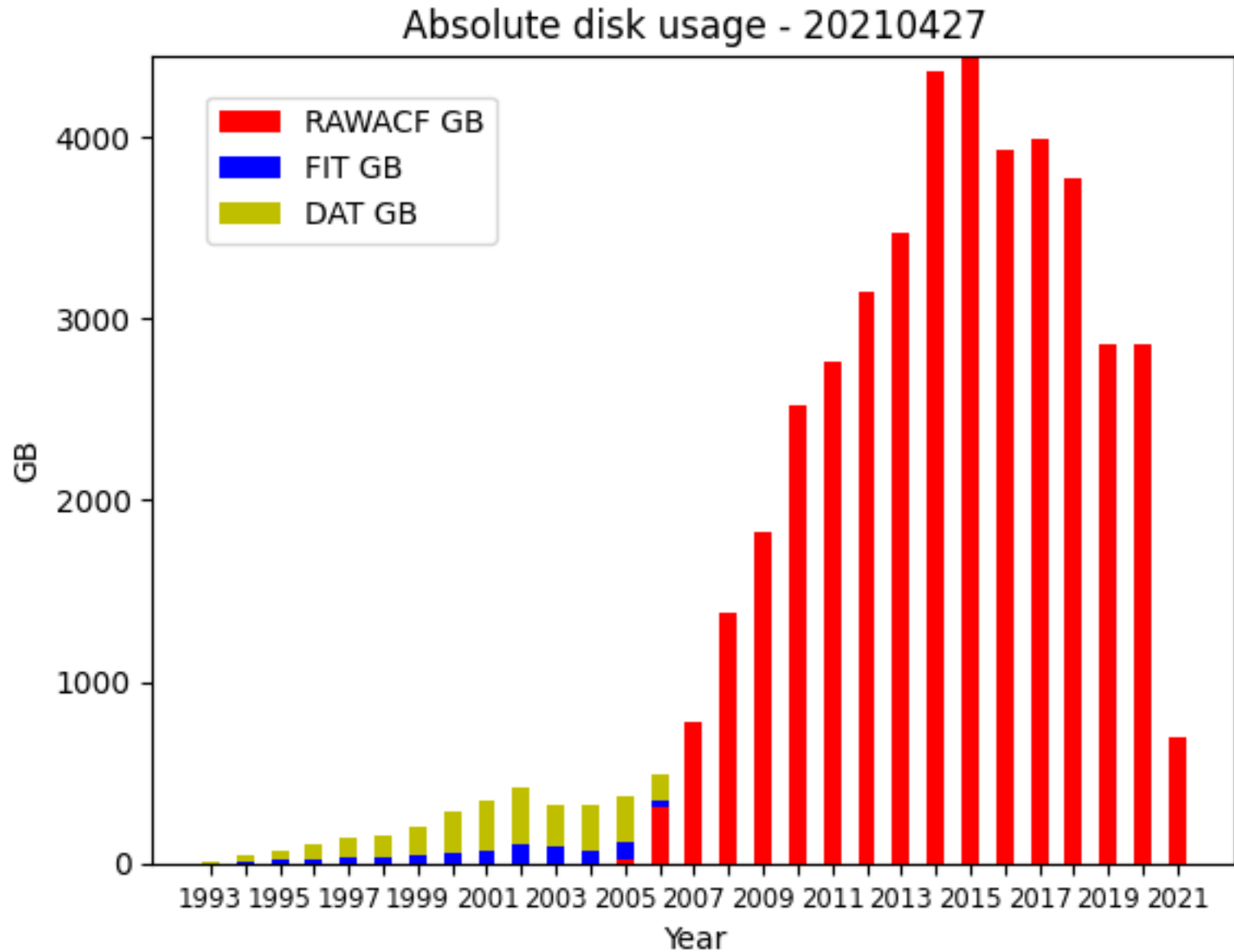


1 – University of Saskatchewan

2 – Virginia Tech

3 – British Antarctic Survey





Access to Data

- Globus mirror (endpoint: computecanada#cedar-dtn) –
 - **Contact Kevin Krieger – kevin.krieger at usask dot ca**
 - BAS mirror available via bslsuperdarn.nc.nerc-bas.ac.uk
 - contact **Paul Breen - pbree at bas dot ac dot uk**
- * New mirrors in the future: PRIC and NSSC



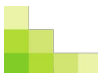
Major updates in the past year

rsync access provided to BAS to sync files from usask. Paul working to close the gap between Globus and BAS mirror differences

VT moved data servers to a temporary location in September 2020, they are expected to move to a more permanent home in the future, date TBD.

BAS started doing direct transfers from PI staging servers (see quarterly reports for a table of the progress)

Temporary mirror superdarn-cssdp.usask.ca decommissioned



Major updates in the past year

File naming convention finalized. Backlog of Borealis (pgr, cly, sas) data now being uploaded one month at a time

DDWG landing page transitioned from VT to SuperDARN Github
<https://github.com/SuperDARN/DDWG>

2194 DAT files retrieved from optical media in storage

A list of failed files (DMAP issues, BZ2 issues, corrupt files, etc) was collated and placed on the Globus mirror.

Master list of all previously updated files placed on the Globus mirror



Major updates in the past year

NSSC data mirror has copied 38TB of data (as of 20210513) from BAS, and ideally have all files copied by late May.

PRIC data mirror finished preliminaries, will be ideally open to community very soon



Outstanding Issues & Discussions

- Differences between mirrors – lack of human resources
-
- Explicit, direct and clear communication when changes in data distribution are made will be appreciated
-
- Data gaps resolution – is this still valuable?
-
- Data flowing from each institution to both VT and BAS for redundancy
-
- Delays/gaps in data distribution for some radars:
 - cly, han, jme, kap, pgr, sas, bpk, ker, sys, tig, unw



Future

- More mirrors coming online (PRIC & NSSC)
- <https://www.frdr.ca/repo/> for a place to publish 'finished' datasets
 - archiving for publications; version control; assigns DOI number to dataset to refer to in publication
- Compute Canada will be transitioning to NDRIIO (New Digital research Infrastructure Organization)