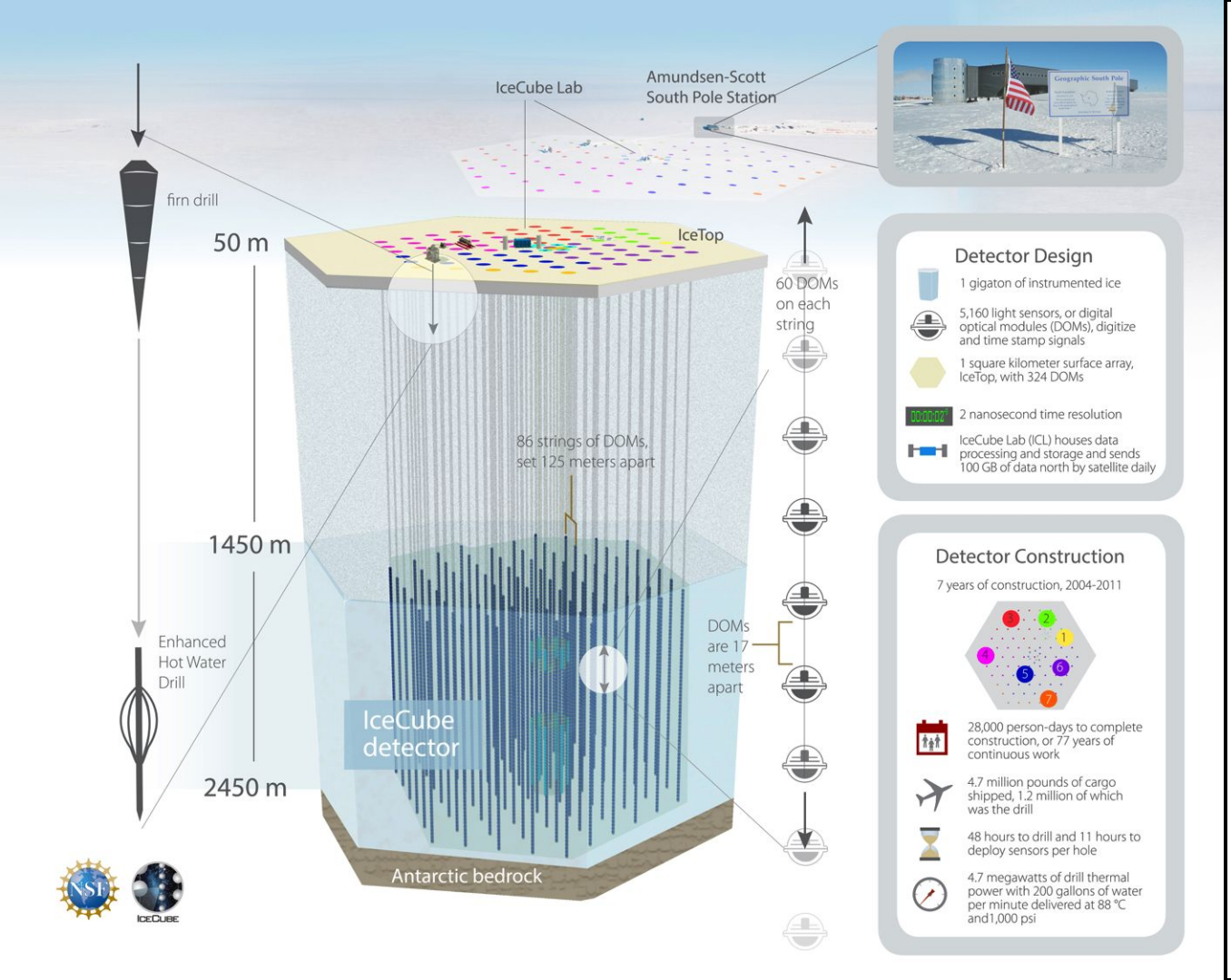
IceCube Data Management

Rucio Community Workshop
March 2018, CERN

G. Merino, P. Meade



Neutrino observatory at the geographic South Pole

5483 Digital Optical Modules (DOMs) in 1km^3 of ice

7 years after completion the detector continues to operate nominally.

- More reliable than expected
- Only 32 of 5483 DOMs have failed

In Apr 2016: NSF renewed UW-Madison M&O contract until 2021

THE ICECUBE COLLABORATION

49 institutions
12 countries
>300 researchers

AUSTRALIA
University of Adelaide

BELGIUM
Université libre de Bruxelles
Universiteit Gent
Vrije Universiteit Brussel

CANADA
SNOLAB
University of Alberta–Edmonton

DENMARK
University of Copenhagen

GERMANY
Deutsches Elektronen-Synchrotron
ECAP, Universität Erlangen-Nürnberg
Humboldt–Universität zu Berlin
Ruhr-Universität Bochum
RWTH Aachen University
Technische Universität Dortmund
Technische Universität München
Universität Mainz
Universität Wuppertal
Westfälische Wilhelms-Universität Münster

JAPAN
Chiba University

NEW ZEALAND
University of Canterbury

REPUBLIC OF KOREA
Sungkyunkwan University

SWEDEN
Stockholms Universitet
Uppsala Universitet

SWITZERLAND
Université de Genève

UNITED KINGDOM
University of Oxford

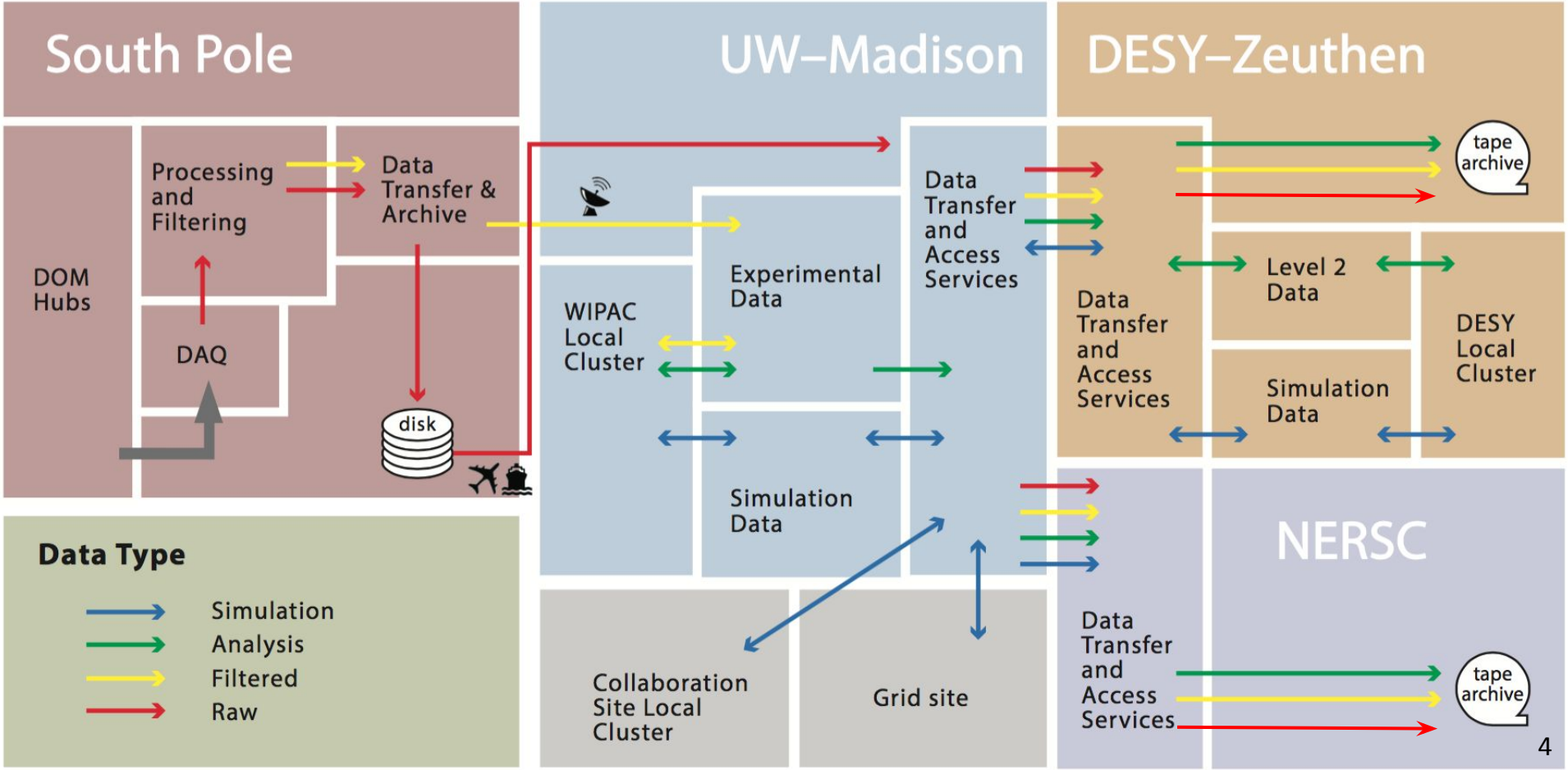
UNITED STATES
Clark Atlanta University
Drexel University
Georgia Institute of Technology
Lawrence Berkeley National Lab
Marquette University
Massachusetts Institute of Technology
Michigan State University
Ohio State University
Pennsylvania State University
South Dakota School of Mines and Technology
Southern University and A&M College
Stony Brook University
University of Alabama
University of Alaska Anchorage
University of California, Berkeley
University of California, Irvine
University of California, Los Angeles
University of Delaware
University of Kansas
University of Maryland
University of Rochester
University of Texas at Arlington
University of Wisconsin–Madison
University of Wisconsin–River Falls
Yale University

FUNDING AGENCIES

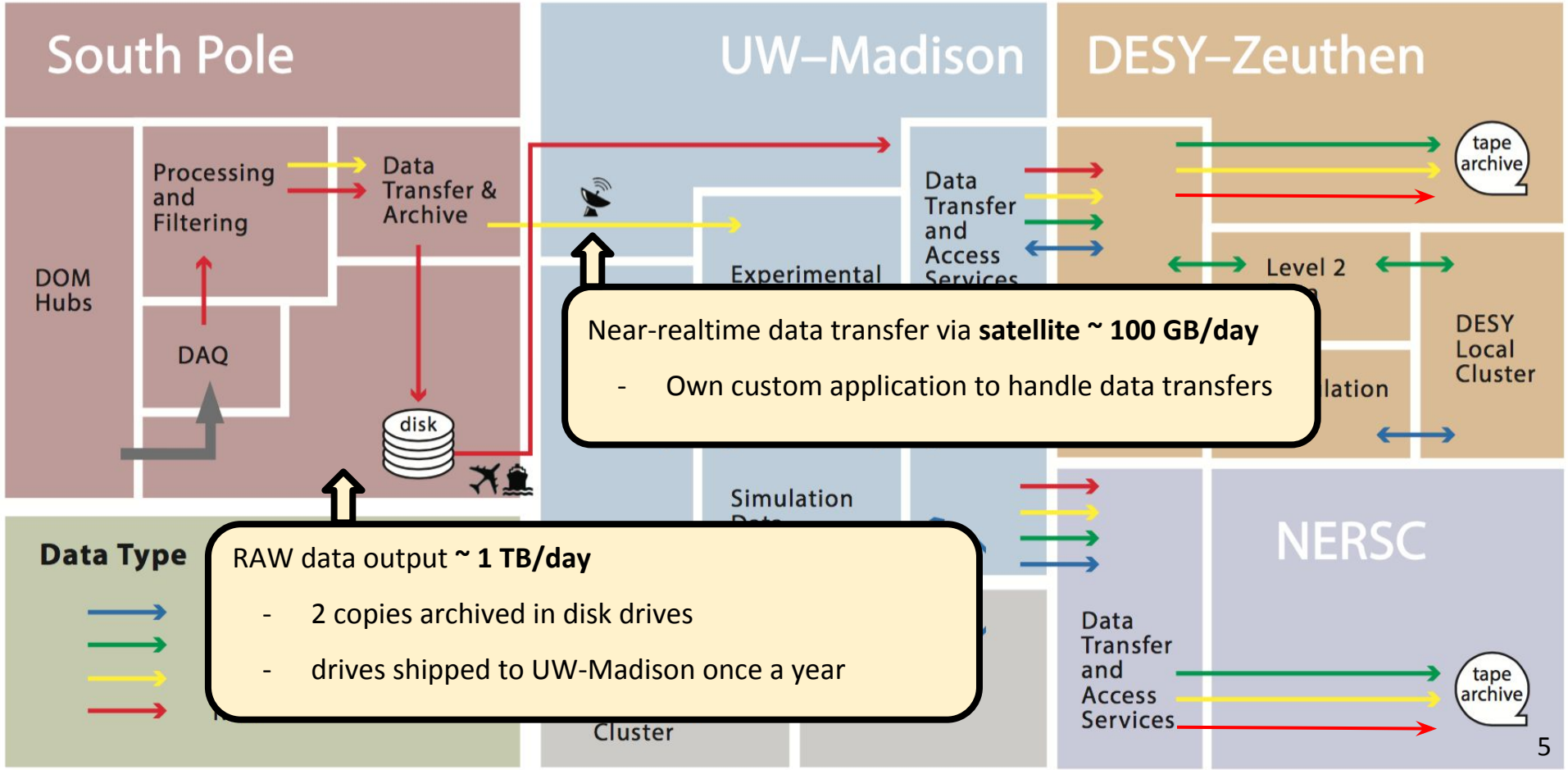
Fonds de la Recherche Scientifique (FRS-FNRS)
Fonds Wetenschappelijk Onderzoek-Vlaanderen (FWO-Vlaanderen)
Federal Ministry of Education and Research (BMBF)
German Research Foundation (DFG)
Deutsches Elektronen-Synchrotron (DESY)
Japan Society for the Promotion of Science (JSPS)
Knut and Alice Wallenberg Foundation
Swedish Polar Research Secretariat
The Swedish Research Council (VR)
University of Wisconsin Alumni Research Foundation (WARF)
US National Science Foundation (NSF)

ICECUBE
South Pole Neutrino Observatory
icecube.wisc.edu

IceCube Data Flow



IceCube Data Flow



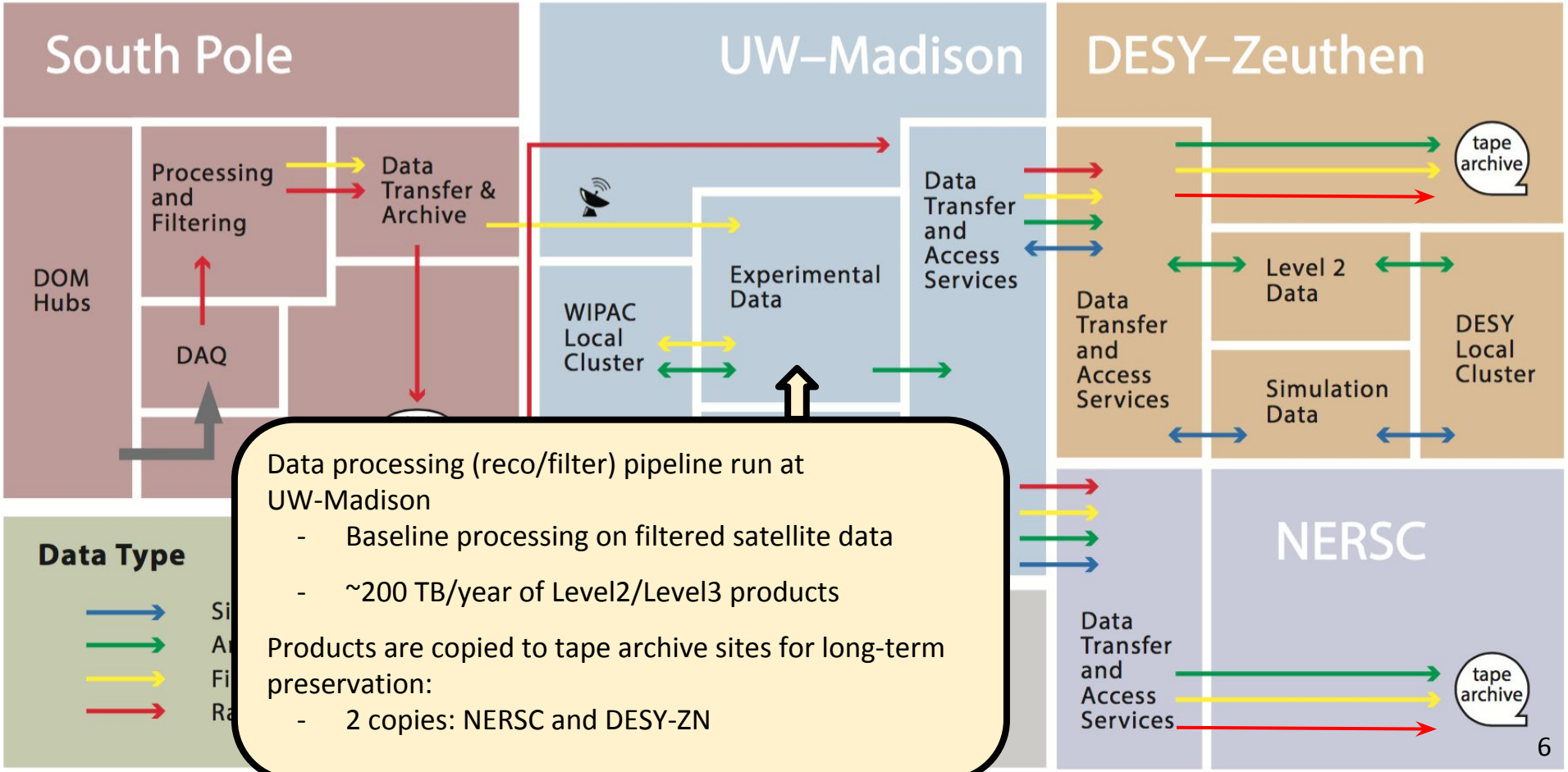
Near-realtime data transfer via **satellite ~ 100 GB/day**

- Own custom application to handle data transfers

RAW data output **~ 1 TB/day**

- 2 copies archived in disk drives
- drives shipped to UW-Madison once a year

IceCube Data Flow



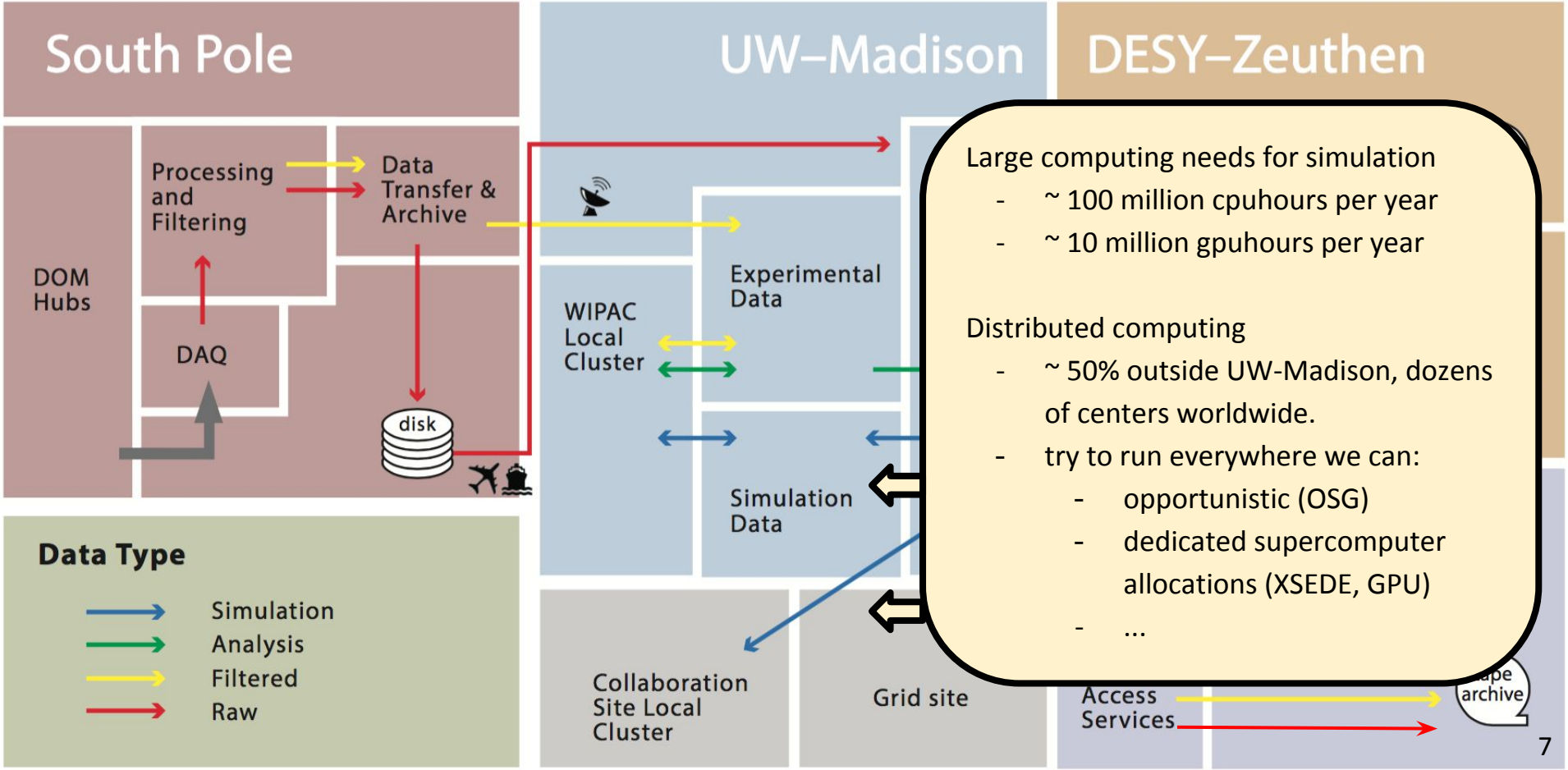
Data processing (reco/filter) pipeline run at UW-Madison

- Baseline processing on filtered satellite data
- ~200 TB/year of Level2/Level3 products

Products are copied to tape archive sites for long-term preservation:

- 2 copies: NERSC and DESY-ZN

IceCube Data Flow



Large computing needs for simulation

- ~ 100 million cpuhours per year
- ~ 10 million gpuhours per year

Distributed computing

- ~ 50% outside UW-Madison, dozens of centers worldwide.
- try to run everywhere we can:
  - opportunistic (OSG)
  - dedicated supercomputer allocations (XSEDE, GPU)
  - ...

Data Management: JADE

In-house developed application to handle data files: transfer, archive, bookkeeping, etc.

South Pole JADE: consumption from DAQ, packaging, archive and satellite transfer

JADE North: Indexing and warehousing satellite data to the disk at UW-Madison

JADE LTA (Long Term Archive):

- Indexing and bundling (>500GB zip bundles) archival data
- Transfer to tape archives: UW-Madison → DESY-ZN, NERSC
  - Currently using Globus for handling GridFTP transfers to tape archives

Data retention/archival policies

Data type	Subtype	Growth (TB/yr)	DESY-ZN tape	NERSC tape	Years on disk at UW-Madison
Experimental	RAW	320		yes	2
	SuperDST	70	yes	yes	2
	Filtered	40	yes	yes	2
	Level2	100	yes	yes	3
	Level3	100		yes	10
Simulation	Level2	400			3
	Level3	100		yes	10
	Photon tables	8			5

The total file count since detector started is ~500M (not counting private user files)

Current data model is very centralized: only 4 Storage Elements

- Primary Disk copy at UW-Madison (main processing and analysis facility)
- Disk storage at DESY-ZN with a replica of L2,L3 data and some simulation data
- Two archival (Tape) sites: DESY-ZN and NERSC

IceCube files

IceCube data is organized in files. Each file contains a number of “frames” (~events ~triggers).

One RAW file for instance contains ~2.5 minutes of detector data (~1.2 GB)

Data taking is organized in “runs”: run changes every 8 hours, or if detector configuration changes.

Some experimental files examples:

```
/data/exp/IceCube/2016/filtered/level2/0101/Run00127347/Level2_IC86.2015_data_Run00127347_Subrun00000147.i3.bz2

/data/exp/IceCube/2016/unbiased/PFRaw/0601/PFRaw_PhysicsFiltering_Run00127988_Subrun00000000_00000012.tar.gz
```

File Metadata

The path/filename encodes a number of file metadata fields:

/data/exp/IceCube/2016/filtered/level2/0101/Run00127347/Level2_IC86.2015_data_Run00127347_Subrun00000147.i3.bz2

```
data_type = experimental
create_date = 2016-01-01T20:34:54.38Z
processing_level = level2
run_number = 127347
season = 2015
```

Users often discover data by browsing the filesystem.

File Metadata

There are several additional metadata fields that we want to keep record of.

These are currently in bookkeeping DBs internal to the various s/w systems that handle files.

We are in the process of building a “file metadata catalog” that will be the core repository for this information.

`logical_name`	`start_datetime`	`software`
`file_size`	`end_datetime`	`validation.vali`
`checksum`	`first_event`	`dated`
`content_status`	`last_event`	`validation.date`
		`...`

Workflows

Two different methods used by two different groups:

- Organized simulation production and processing
  - Using in-house developed framework called IceProd
  - Mainly “sits on top of HTCondor” for job submission, but handles dataset prioritization, bookkeeping, etc.
  - Few users (~2-5) driving most of the data “writes”

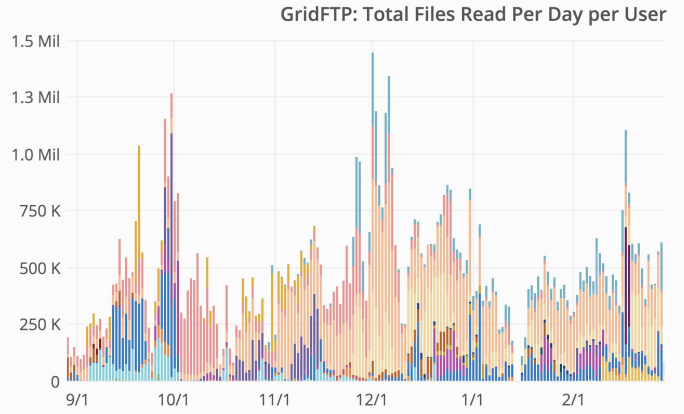
- User analysis
  - Mostly use HTCondor DagMan for managing workflows
  - Mostly reading the “official” data produced by the 1st group, and writing private analysis user output (no bookkeeping)

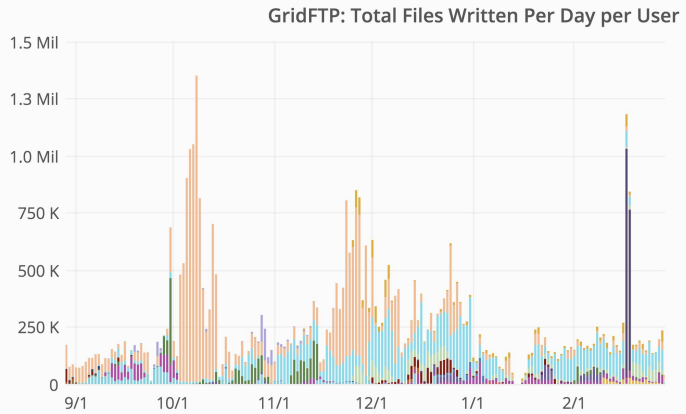
Data access

Jobs running in the central cluster at UW-Madison access the data directly (Lustre mounted filesystem)

Jobs running in the Grid (~50% of total) access data at UW-Madison via GridFTP

- x509 certs and VOMS are the technologies currently in use
- interested in contributing to/prototyping alternatives (e.g. SciTokens, etc)





Data sharing

Public data releases are still limited to targeted publications

- Published data varies from publication to publication: histogram bins, contour values, event lists …

A generic data release containing standard samples of reconstructed events is work in progress.

http://icecube.wisc.edu/science/data

Download Datasets

The pages below contain information about the data that were collected and links to the data files. We ask that you provide your contact information so that we may notify you if we revise the datasets in the future.

Dataset	Release Date
Measurement of atmospheric neutrino oscillations with three years of data from the full sky	2018 Feb 13
A combined maximum-likelihood analysis of the astrophysical neutrino flux	2016 Nov 15
Search for point sources with first year of IC86 data	2016 Nov 01
Search for sterile neutrinos with one year of IceCube data	2016 Jun 24
The 79-string IceCube search for dark matter	2016 Jan 05
Observation of Astrophysical Neutrinos in Four Years of IceCube Data	2015 Oct 21
Astrophysical muon neutrino flux in the northern sky with 2 years of IceCube data	2015 Aug 20
IceCube-59: Search for point sources using muon events	2015 Jun 19
Search for contained neutrino events at energies greater than 1 TeV in 2 years of data	2015 Feb 19
IceCube Oscillations: 3 years muon neutrino disappearance data	2015 Jan 27
Search for contained neutrino events at energies above 30 TeV in 2 years of data	2013 Nov 21
IceCube-40 String Data	2011 Sep 05
IceCube-22 Solar WIMP Data	2009 May 21
AMANDA 7 Year Data	2008 Sep 11

Data traceability

Q: How do you link your scientific publications to the data that you used?

- (Great question!) unfortunately, for now, we don’t
- This is a HOT topic inside the collaboration
  - We plan to develop tools for users to be able to register analysis datasets linked to publications and record provenance information
  - Needless to say that we are happy to share ideas and discuss with anyone out there doing this

Rucio Evaluation

We are interested in understanding the details of Rucio capabilities for data management

- The goal is to explore integration possibilities (avoid reinventing the wheel)

Some parts of our data management system (South Pole, satellite) are custom to the specific environment, but some others are doing very standard things

Our first goal is to make a limited scope evaluation by using Rucio to perform a task that we currently do with JADE LTA

- Replication of L2/L3 processed data from UW-Madison disk to DESY-ZN disk

With this exercise we hope to learn about the Rucio internals for then being able to propose potential further integrations with our system.

thanks

questions?