CI CoE PILOT

Cyberinfrastructure Center of Excellence Pilot

Funded by the National Science Foundation Grant #1842042

Ewa Deelman, USC (PI)

Co-PIs:
Anirban Mandal, RENCI

Jarek Nabrzyski, Notre Dame University

Valerio Pascucci and **Rob Ricci**, University of Utah

USC Viterbi School of Engineering *Information Sciences Institute* · University of Notre Dame · The University of Utah · Center for Applied Cybersecurity Research · renci

Overview

- CiCOE Pilot project introduction and goals
- Data Life Cycle (DLC)
- DLC case study - IceCube Large Facility (LF)
- DLC taxonomy across four LFs
- Disaster Recovery (DR) in the context of DLC
- DR planning case study – IceCube LF
- Lessons learned from initial engagements
- **Feedback from participants**

CI CoE PILOT

Recognizing the importance of CI in Large Facilities

Manish Parashar (PI and Chair), Rutgers University and OOI
Stuart Anderson, LIGO
Ewa Deelman, USC
Valerio Pascucci, University of Utah
Donald Petravick, LSST
Ellen M. Rathje, NHERI

September 2017 Workshop report at http://facilitiesci.org/

- **Establish a center of excellence** (following a model similar to the NSF-funded Center for Trustworthy Scientific Cyberinfrastructure, CTSC) as a resource providing expertise in CI technologies and effective practices related to large-scale facilities as they conceptualize, start up, and operate.
- Foster the creation of a facilities' CI community and establish mechanisms and resources to enable **the community to interact, collaborate, and share**.

USC Viterbi School of Engineering Information Sciences Institute
University of Notre Dame
The University of Utah
Center for Applied Cybersecurity Research
renci

CI CoE Pilot Project Goals

Award #1842042

Develop a model and a plan for a Cyberinfrastructure Center of Excellence

- Platform for knowledge sharing and community building
- Key partner for the establishment and improvement of Large Facilities with advanced CI architecture designs
- Grounded in re-use of dependable CI tools and solutions
- Forum for discussions about CI sustainability and workforce development and training
- Pilot a study for a CI CoE through close engagement with NEON and further engagement with other LFs and large CI projects.

Data Life Cycle (DLC) and Disaster Recovery (DR) for Large Facilities

Team members

CICoE Pilot: Anirban Mandal, Laura Christopherson, Erik Scott, Ilya Baldin, Paul Ruth (RENCI)

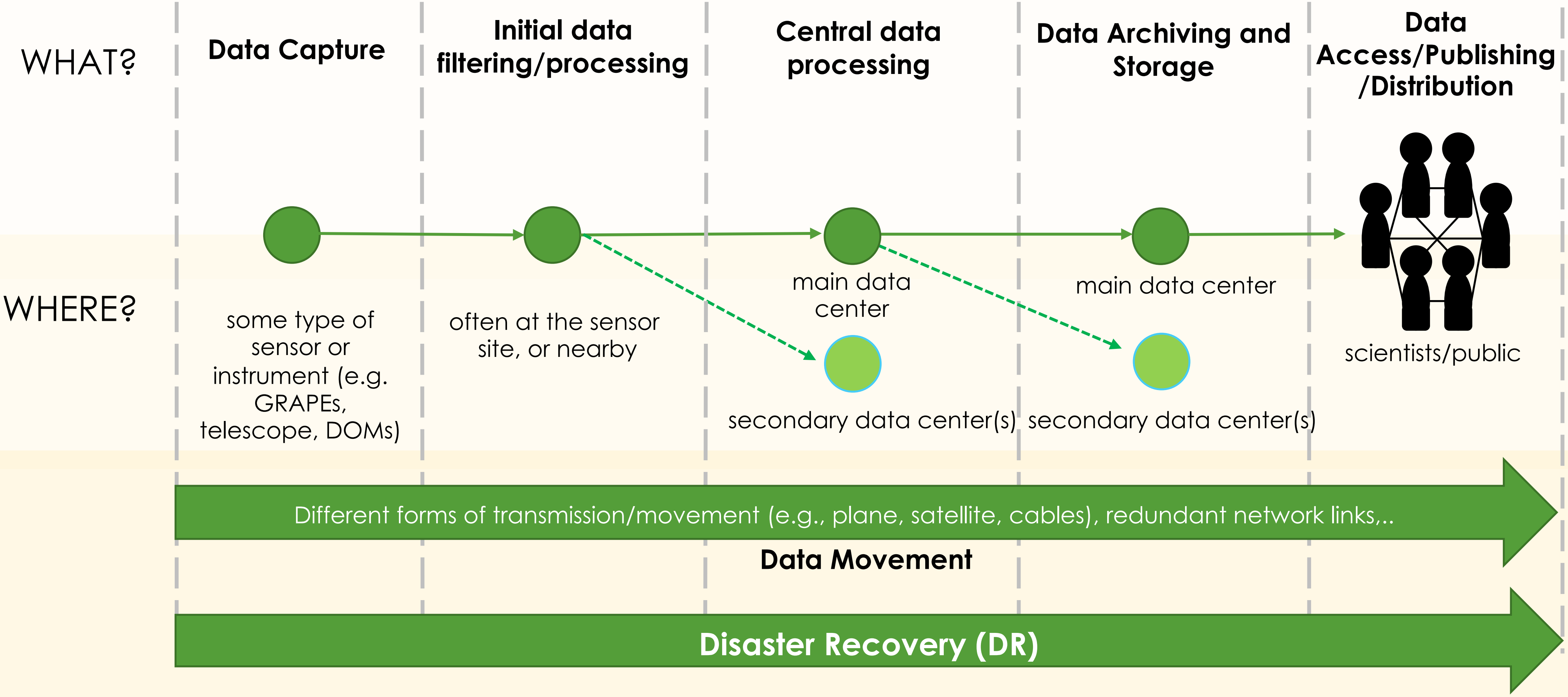
NEON: Philip Harvey, Steve Jacobs, Tom Gulbransen (NEON Large Facility, Boulder)
IceCube: Benedikt Riedel (Wisconsin IceCube Particle Astrophysics Center)

Data Life Cycle and DR for LFs: Goals

- Understand and document the best practices and solutions for data life cycle (DLC) and disaster recovery (DR) plans for Large Facilities (LFs).
- Develop DLC model, taxonomy and DR planning models
  - a generalizable DLC model for LFs based on engagements with LFs: NEON, IceCube, and others.
  - a taxonomy of CI services, architectures, and functionalities that support the different DLC stages.
  - effective process guides for DR planning for LFs in the context of DLC stages.

CI CoE PILOT

Data Life Cycle for LFs



USC Viterbi School of Engineering Information Sciences Institute

UNIVERSITY OF NOTRE DAME

THE UNIVERSITY OF UTAH

CENTER FOR APPLIED CYBERSECURITY RESEARCH

renci

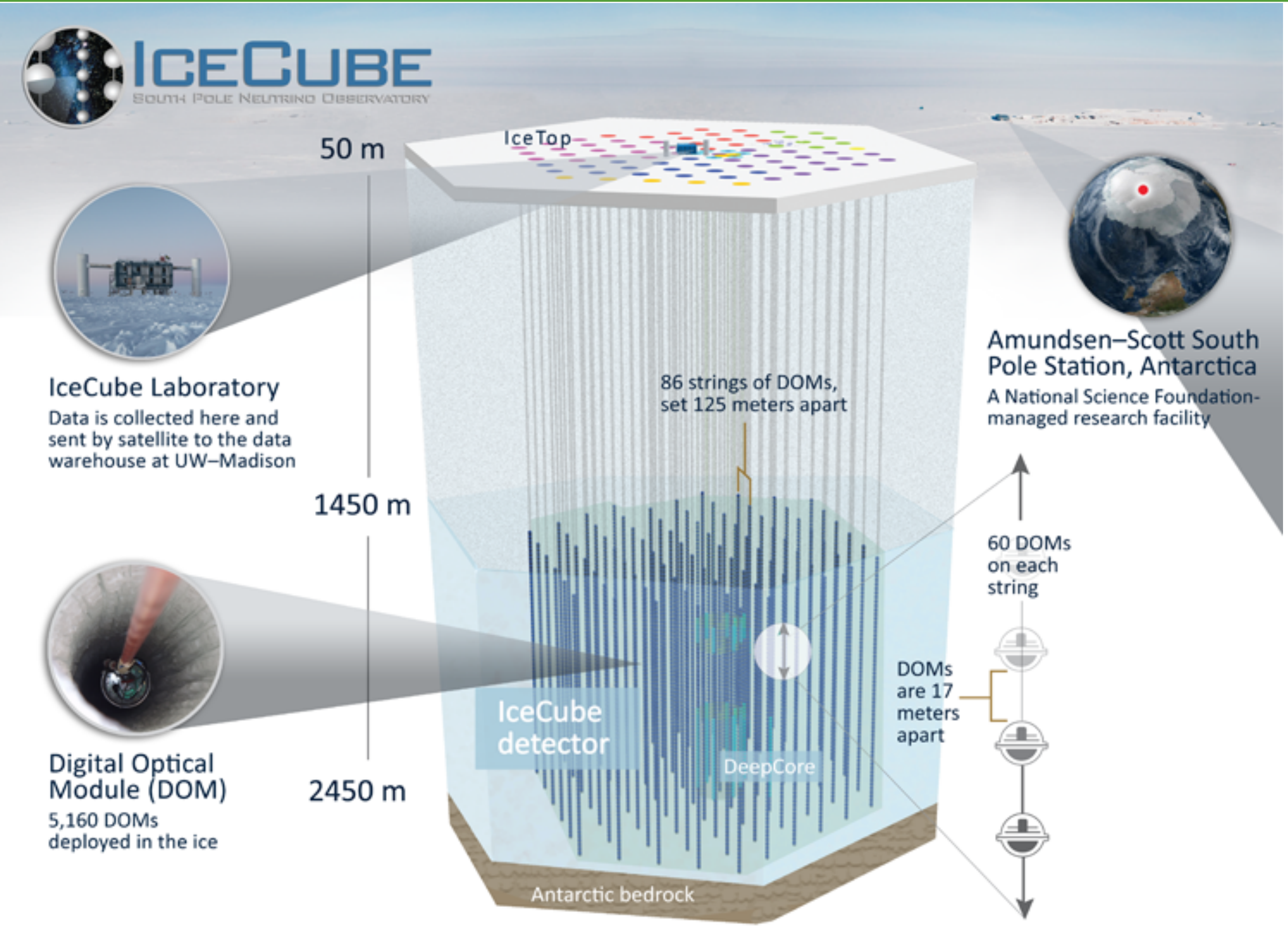
DLC case study – IceCube Large Facility

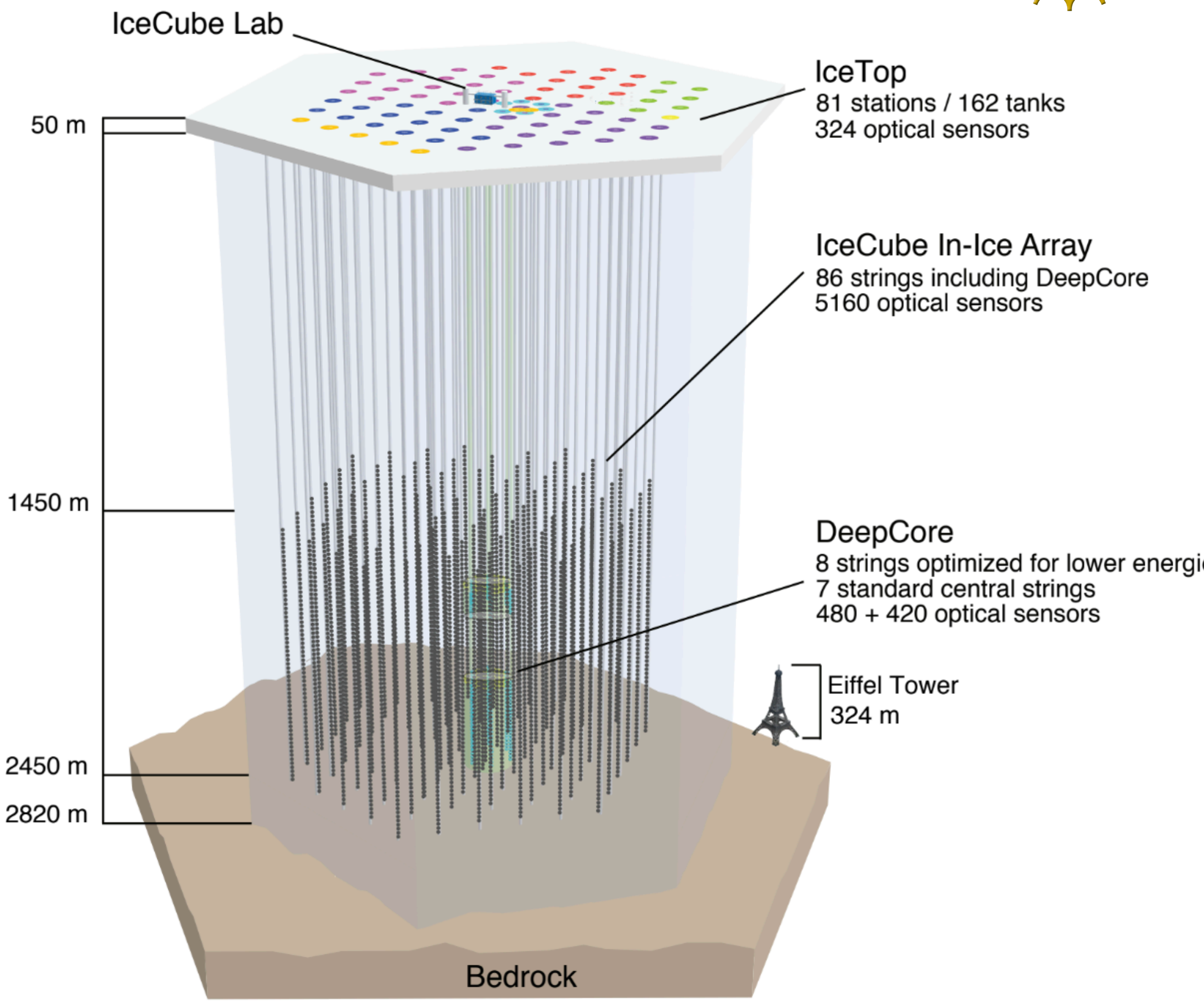
IceCube materials courtesy: Dr. Benedikt Riedel, Wisconsin IceCube Particle Astrophysics Center

IceCube: About the Data

- Data represents
  - hits
  - events – time period of interest with fixed read out window
  - metadata and secondary streams (e.g., time calibration, monitoring)
- 4 types of data
  - *PFRAW* - full data set originating at the South Pole (~3TB/day)
  - *PFFILT Level 1* - ~100 GB/day of PFRAW that is filtered and send to UW-Madison
  - *Level* 2 - Level 1 data that has directional reconstructions and is "science ready"
  - *Level* 3 - Level 2 data that has been reduced, with extra reconstructions applied, by a particular science working group

IceCube Data Life Cycle: Data Capture





IceCube Data Life Cycle: Initial Processing/Filtering

Initial filtering, processing: South Pole

- Data is received by DOMHubs and IceCube Lab (surface of South Pole) - ~500 core filtering cluster; ~100 machines for detector readout; Hits are output as events.
- Internal PnF system selects events based on their usefulness for a particular analysis. It also creates event metadata and reduces data volume before it is transmitted away from the South Pole.
- **Alert production** is an important step that happens at the South Pole.

IceCube Data Life Cycle: Central Processing/Filtering

Central processing: UW-Madison processes what is sent from the South Pole to a "science ready level" or level 3.

- UW-Madison 6500 core, 300 GPU cluster, ~10 PB storage
- Additional downstream processing happens using a mix of resources: DESY, OSG, IceCube Grid (campus clusters, contributed resources, etc.), XSEDE allocations, DOE resources (e.g. NERSC)
- Increased demand for GPU resources
- PyGlidein + HTCondor based distributed computing middleware
- Exploring cloud resources for CPU, GPU, ML

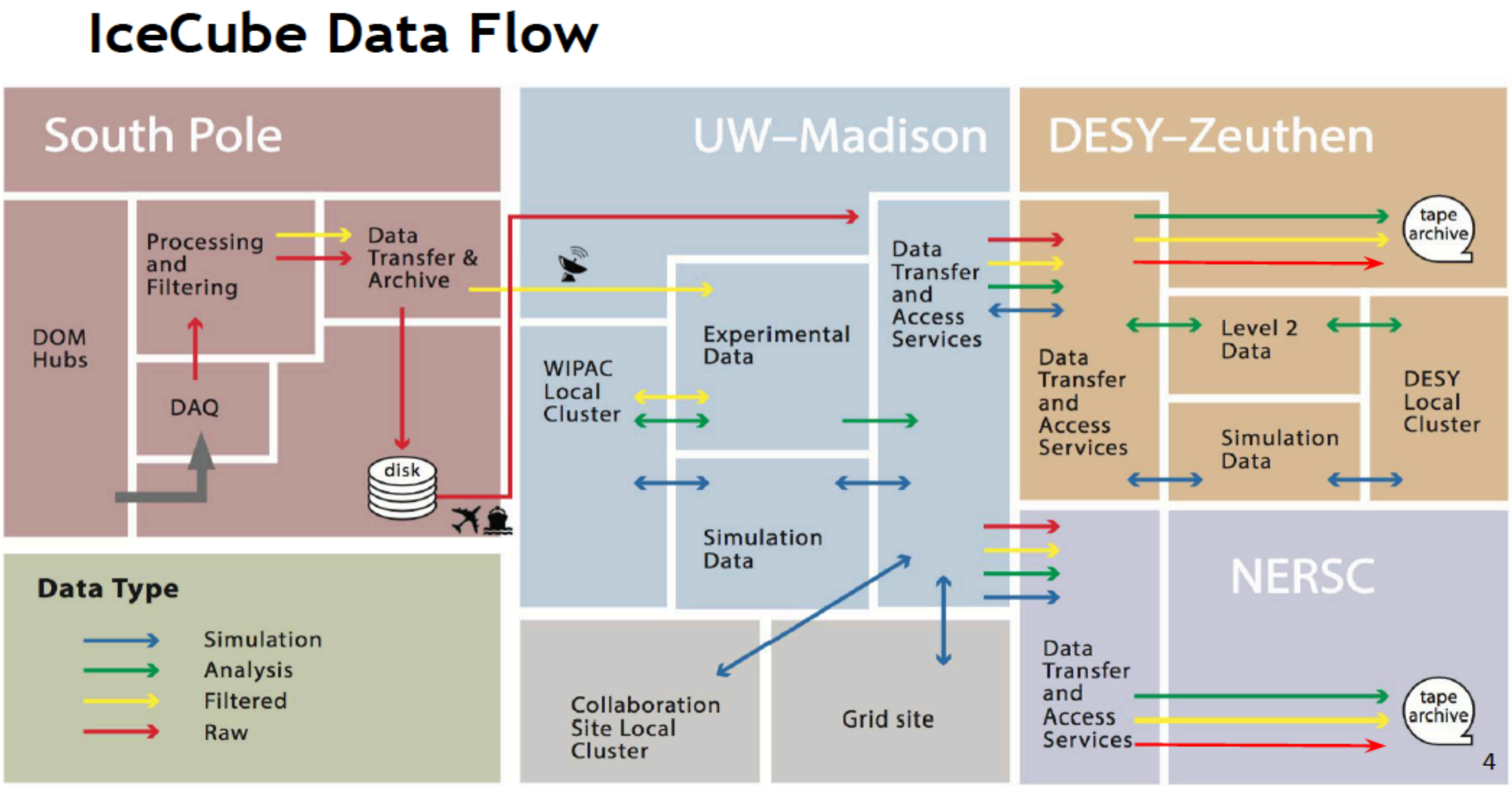
IceCube Data Life Cycle: Movement

1. Hits at DOMs
2. Sent to DOMHubs
3. Sent to Data Acquisition System (DAQ) - hits are output to events (PFRAW)
4. Sent to Processing and Filtering System (PnF) - PFRAW made ready for analyses
5. Sent to South Pole Station JADE for archival storage to disk (PFRAW and PFFILT/Level 1)
6. JADE transmits via satellite to UW-Madison (PFFILT)
7. PFFILT sent to DESY and PFRAW sent to NERSC for additional tape backups

In addition, Alerts are sent out using GCN (Gamma Ray Coordination Network - operated by NASA) or Astronomical telegrams along with initial estimate of PFRAW data sample via satellite link to UW.

- Limited bandwidth of ~125 GB/day from South Pole to UW; 3TB/day raw data is filtered down to ~80GB/day and transmitted via satellite from South Pole Station to UW
- Once a year, raw data from the South Pole is sent via plane and disks to UW-Madison
- UW connected to SciDMZ through Starlight-ESNet for connection to DOE facilities
- Leverages GridFTP for data transfers from UW-Madison to DESY/NERSC/OSG

IceCube Data Life Cycle: Movement



IceCube Data Life Cycle: Archiving

JADE (archival system) exists in ~3 locations

- South Pole JADE - writes 2 copies to disk (3 TB/day)
- JADE North (UW) - warehouses the data to disk (~200 TB/yr)
- JADE Long Term Archive (LTA) in DESY – keeps replicas of Level 1 and 2 data

NERSC archives PFRAW

Dissemination of Alerts

- Alerts happen at the South Pole during Level 1 processing.
- Alerting systems detect events and then an immediate alert is sent out using GCN (Gamma Ray Coordination Network - operated by NASA) or Astronomical telegrams along with initial estimate/small portion of PFRAW data sample via satellite link to UW.
- When a full PFFILT data set is available at UW later, a refinement of the first alert is sent.

IceCube Data Life Cycle: Access/Publishing/Distribution

3 forms of **data access** eligibility:

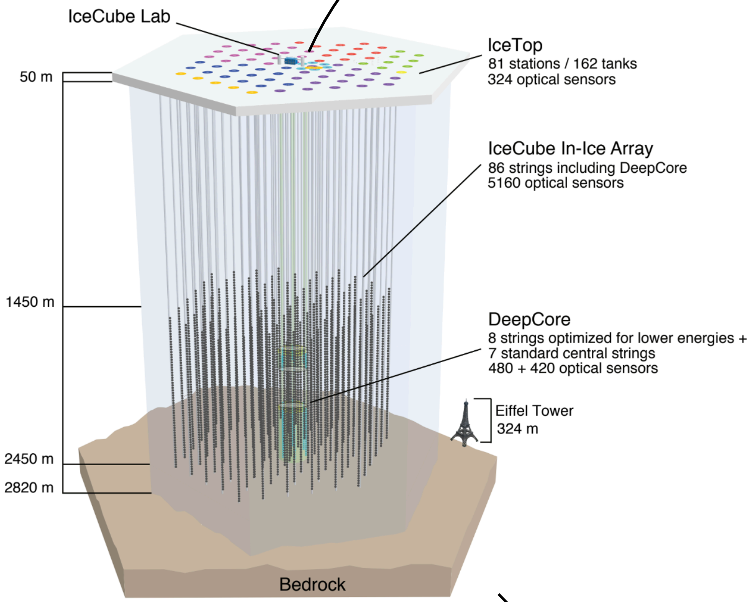
- Be a member of the IceCube Collaboration
- Be an "associate member" which means one applies for use of the data for a particular purpose but is not required to fulfill collaboration obligations
- By anyone using the public data pages on the web

Planned enhancements for data organization, management, access, and data catalog

- Xrootd-based solution, Ceph/www

Data is released to members and associates. When the data has been analyzed and those analyses published, it becomes available for release to others.

IceCube Logical Architecture



IceCube Lab
~500 core filtering cluster
~100 machines for detector readout
2 copies of all data archived

fiber

PFFILT
~125 GB/day bandwidth
100GB bundles

Amundsen-Scott

All PFRAW
disk
1/year

UW-Madison

6500 core, 300 GPU cluster

~10 PB storage

SciDMZ through Starlight

ESNET for connection to DOE facilities

Grid FTP

DESY

Replicates Level 1 and 2 from UW; Tape backup

½ of Level 2 processing

Grid: 384 cores, ~6300 HEPSpec O6

Storage: 360 TB dCache, 150 TB Lustre

Local CPU: 1000 cores, ~13,800 HEPSpec06

Grid FTP

NERSC

Replicates Level 2 and 3 from UW; Tape backup

OSG, XSEDE, DOE Facilities, Campus clusters

Other Processing

IceCube materials courtesy: Dr. Benedikt Riedel, Wisconsin IceCube Particle Astrophysics Center

DLC Taxonomy for LFs

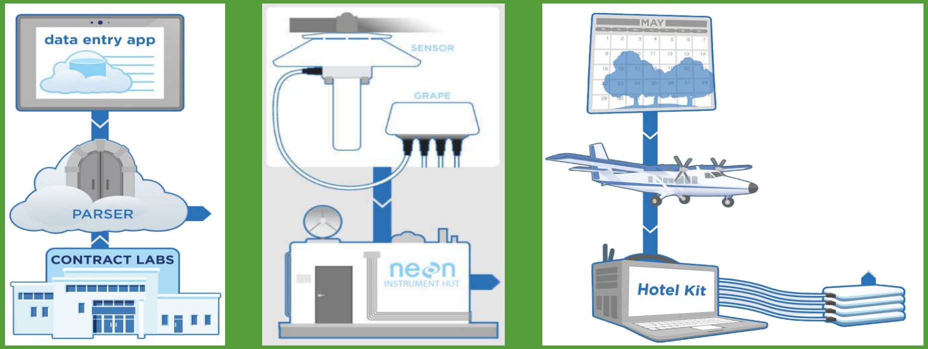
NEON materials courtesy: Tom Gulbransen, Battelle
OOI materials courtesy: Dr. Ivan Rodero, Rutgers Discovery Informatics Institute
IceCube materials courtesy: Dr. Benedikt Riedel, Wisconsin IceCube Particle Astrophysics Center

About the Data

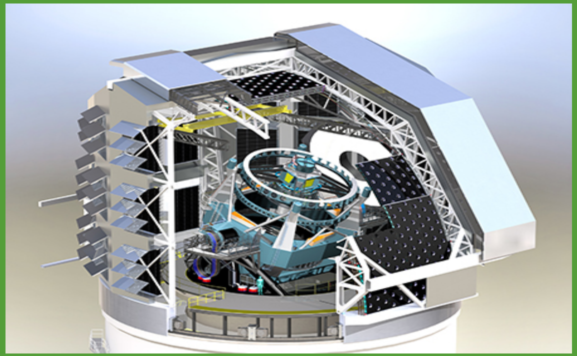
NEON	OOI	IceCube	LSST
6 levels of data • L0: raw data • L0': raw data formatted for external use • L1: unit conversions, time averaged, calibrated, quality checked • L2: temporally interpolated • L3: spatially interpolated • L4: integrated data products such as indices or fluxes. 3 types of data • observational samples (OS) • instrumented systems (IS) • aerial observatory platform (AOP)	Data comes from • Cabled sensors (150kbps, .6 TB/yr) • Uncabled sensors (440kbps, 70 TB/yr) • Video (1.95 Gbps, 7.7 PB/yr)	Data includes: • hits • events • metadata and secondary streams (e.g., time calibration, monitoring) 4 types: • PFRAW - full data set originating at the South Pole (~3TB/day) • PFFILT Level 1 - ~100 GB/day of PFRAW that is filtered and send to UW-Madison • Level 2 - Level 1 data that has directional reconstructions and is "science ready" • Level 3 - Level 2 data that has been reduced, with extra reconstructions applied, by a particular working group	3 main types: • *Prompt/nightly* including alerts released every 60 seconds and raw images released every 24 hours • *Data Release* - prompt/nightly data that has been processed/filtered, released annually • *User Generated - Data Release* that has been used by project teams and may be released to others as Data Release

Data Life Cycle: Data Capture

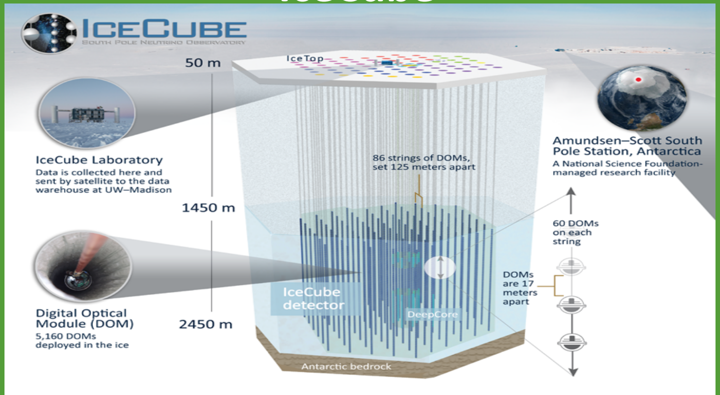
NEON



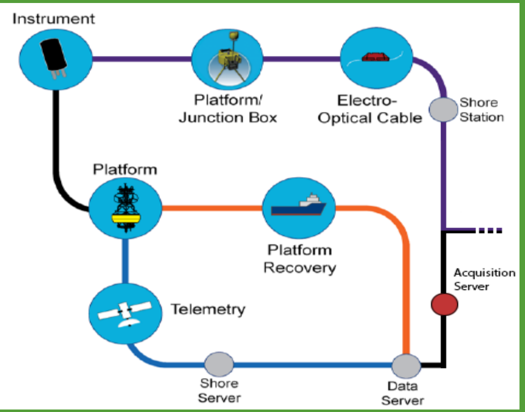
LSST



IceCube



OOI



Data Life Cycle: Processing/Filtering

	NEON	OOI	IceCube	LSST
Initial filtering, processing At/near Data Capture Location	• Eddy covariance may be applied to IS data at field sites.	Drivers/parsers receive data from the acquisition pipeline shown in the previous slide.	South Pole: • Data is received by DOMHubs and IceCube Lab (surface of South Pole) - ~500 core filtering cluster. Hits are output as events. • Internal PnF system selects events based on their usefulness for a particular analysis. It also creates metadata and reduces data volume before it is transmitted away from the South Pole. • Alert production is an important step that happens here.	Base Facility, La Serena Chile • real time alert generation • initial detector cross-talk correction • metadata creation
Central processing At/near Data Center Location(s)	Iron Mountain Data Center (DEN-1) in Denver • Apache Airflow (possibly use Pachyderm in future). All dockerized. • Oracle PDR (OS data), Elastic Cloud Storage (IS data), Common Object Storage (AOP data) • All transitions pipelines run in the data center. • Heavy processing largely takes place in the VMWare resource cluster with metadata/data pulled from ECS and SC9000 SAN • 476 total cores, 8.5 TB RAM (total)	Rutgers uFrame databases perform quality control, construct alerts/alarms, create calibration info, format the data, and generate metadata. Beyond this, P/F typically happens on demand when a user requests it.	UW-Madison processes what is sent from the South Pole to a "science ready level" or level 3, and coordinates simulation production. • 6500 core, 300 GPU cluster • ~10 PB storage Additional downstream processing happens using a mix of resources: DESY, OSG, IceCube Grid (campus clusters, contributed resources, etc.), XSEDE allocations, DOE resources (e.g. NERSC) Increased demand for GPU PyGlidein + HTCondor based distributed computing middleware Exploring cloud resources for GPU, ML	NCSA • *Nightly/Alert* production • 50% of *Data Release* production • *Calibration* production • moving object France • the other 50% of *Data Release production*

Data Life Cycle: Movement

NEON

OS data
1. Collected on Fulcrum tablets.
2. Tablets physically connected/uploaded to the network and parsed in pipeline parser.
3. Loaded into the Processed Data Repository (PDR).
4. Made available in the Data Portal.

IS data
1. Collected via Smart Sensors (SS)or GRAPES.
2. Potentially logged, buffered, processed on site.
3. Uploaded to the network.
4. Data stored and processed in Elastic Cloud Storage (ECS). Metadata stored in Processed Data Repository (PDR).
5. Made available in the Data Portal.

AOP data
1. Collected via hotel kits on the plane.
2. Data is saved on a server in the plane.
3. Server's drives are removed and physically transported to common object storage in Elastic Cloud Storage (ECS).
4. Processed in ECS.
5. Made available in the Data Portal.

OOI

Cabled data
1. Sent to a platform/junction box, then to an electro-optical cable
2. Sent to a shore station
3. Parsed
4. Sent to the uFrame databases for other forms of processing (e.g., quality control, calibration) Algorithms are applied in uFrame
5. Made available to users

Physically recovered and telemetered data
1. From the instrument to the platform
2. Recovered (e.g., by ship) or sent via sattelite to a shore server.
3. Then to an acquisition server
4. Parsed
5. uFrame framework and databases for more processing. Algorithms in uFrame
6. Made available to users

IceCube

1. Hits at DOMs
2. Sent to DOMHubs
3. Sent to Data Acquisition System (DAQ) - hits are output to events (PFRAW)
4. Sent to Processing and Filtering System (PnF) - PFRAW made ready for analyses
5. Sent to South Pole JADE for archival storage to disk (PFFILT/Level 1)
6. JADE transmits via satellite to UW-Madison (PFFILT)
7. PFFILT sent to DESY and PFRAW sent to NERSC for additional tape backups

* Alerts sent via GCN or Astronomical telegrams along with small portion of PFRAW via sattelite link
* Limited bandwidth of ~125 GB/day; 3TB/day raw data is filtered down to ~80GB/day and transmitted via satellite from South Pole Station to UW.
* Once a year data from the South Pole (triggered) is sent via plane and disks to UW-Madison ~125 GB/day from South Pole to UW
* UW connected to SciDMZ through Starlight-ESNet for connection to DOE facilities.
* Leverages GridFTP for data transfers from UW to DESY/NERSC/OSG

LSST

1. Telescope
2. Base Facility in La Serena, Chile
3. NCSA, Illinois
4. France
5. Data Access Centers (DACs) (e.g. Base Facility, NCSA)

600 Gbps from Summit to Base
2 x 100 Gbps from Base to NCSA

Data Life Cycle: Archiving

NEON	OOI	IceCube	LSST

DEN-1 in Denver is the primary location for ingest, storage, transition and publishing.
• Data may be archived in several kinds of stores: file-oriented, relational databases, NoSQL database
• Automated data movement in and out of storage arrays
• Long-term source code control, archive, and version management
• Software: Oracle DB for PDR, elastic cloud storage, Drupal content management system
• Hardware:
    • 104 cores assigned to replicated Oracle installations, two database portals, and bulk movement of data from the Oracle database environment into the bulk object storage
    • Dell EMC SC9000 SAN general purpose storage
    • 2 Dell EMC Elastic Cloud Storage replicates the data between each other
• Partners (e.g., DataONE) keep sync-ed caches of NEON data
	Copies of data are kept by the two data centers and each partner keeps a copy of the data as well.
	• Portland, OR - Cassandra cluster (~50TB) for raw, SAN (~2PB) for formatted/"gold", NAS for user access (~500TB); 126 dual-socket nodes; uFrame, VMware; 100GB network backbone
	• Rutgers, NJ - Same as Portland except also includes a tape archive (~18PB)
	Overall capacity: 25PB storage, 126 servers (dual Xeon)
	
	Connected via redundant Internet 2 links (10GB). Portland mirrors Rutgers. Both store and provide computing infrastructure for the data.
	
	All sites are linked via VPN appliances.
	
	SAN, Cassandra, and VMWare are backed up to tape.
		JADE (archival system) exists in ~3 locations
		• South Pole JADE - writes 2 copies to disk (3 TB/day)
		• JADE North (UW) - warehouses the data to disk (~200 TB/yr)
		• JADE Long Term Archive (LTA) in DESY – keeps replicas of Level 1 and 1 data
		
		NERSC archives PFRAW data
			Long term copies are stored in 3 locations:
			• Base Facility
			• NCSA
			• France

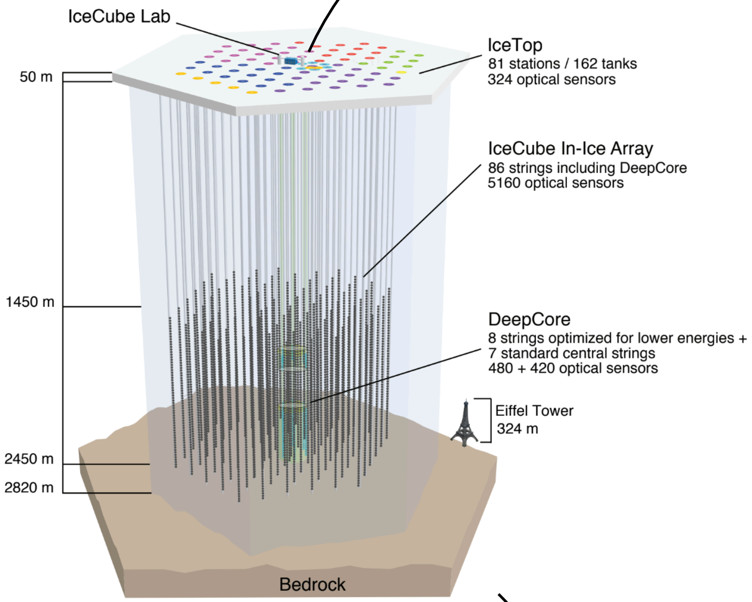
Data Life Cycle: Access/Publishing/Distribution

NEON	OOI	IceCube	LSST
• Data portal - Drupal-based website allows querying the collection • API - Web services interface for external users • Disk transport (e.g., AOP data) • Integrated connections to partner organizations (e.g., AERONET, AmeriFlux, BOLD, MG-RAST, PhenoCam) • 130 cores and 2.5 TB of RAM (aggregate) provisioned for access via web portals, database/data mart connectivity, and data export • Currently moving from a 1Gps pipeline to a 5 Gps to allow higher speed access to data. This was slated to complete in June 2019.	Access means: • OOI Data Portal • THREDDS server • Raw data on the website • ERDDAP server Large datasets are available in NetCDF Data products can be generated on the fly both synchronously and asynchronously Data is served via the NAS	Alert dissemination • Alerts happen at the South Pole during Level 1 processing. • Alerting systems detect events and then an immediate alert is sent out using GCN (Gamma Ray Coordination Network - operated by NASA) or Astronomical telegrams along with initial estimate/small portion of PFRAW data sample via satellite link to UW. • When a full PFFILT data set is available at UW later, a refinement of the first alert is sent. 3 forms of eligibility for access: • Be a member of the Collaboration • Be an "associate member" which means you apply for use of the data for a particular purpose but you are not required to fulfill Collaboration obligations • By anyone and go through the public data pages on the web * Planned enhancements for data organization, management, access, and data catalog (xrootd, ceph/www) * Data is released to members and associates. When the data has been analyzed and those analyses published, it becomes available for release to others.	Access means: • At a Data Access Center (currently at La Serena and NCSA) • Using the Science Platform (can be used from home but fewer affordances than if used at a DAC) • APIs

IceCube Logical Architecture

PFFILT
~125 GB/day bandwidth
100GB bundles

fiber



Amundsen-Scott

IceCube Lab
~500 core filtering cluster
~100 machines for detector readout
2 copies of all data archived

All PFRAW
disk
1/year

UW-Madison

6500 core, 300 GPU cluster

~10 PB storage

SciDMZ through Starlight

ESNET for connection to DOE facilities

Grid FTP

DESY

Replicates Level 1 and 2 from UW; Tape backup

½ of Level 2 processing

Grid: 384 cores, ~6300 HEPSpec O6

Storage: 360 TB dCache, 150 TB Lustre

Local CPU: 1000 cores, ~13,800 HEPSpec06

Grid FTP

NERSC

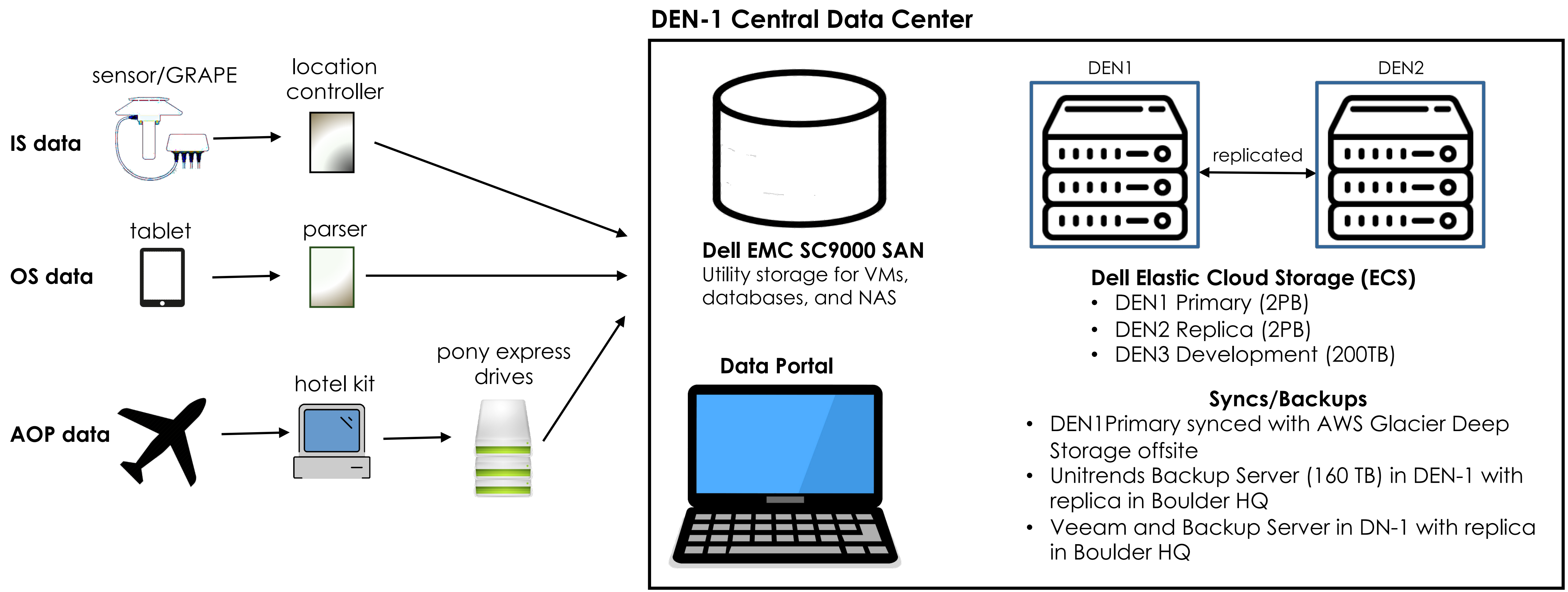
Replicates Level 2 and 3 from UW; Tape backup

OSG, XSEDE, DOE Facilities, Campus clusters

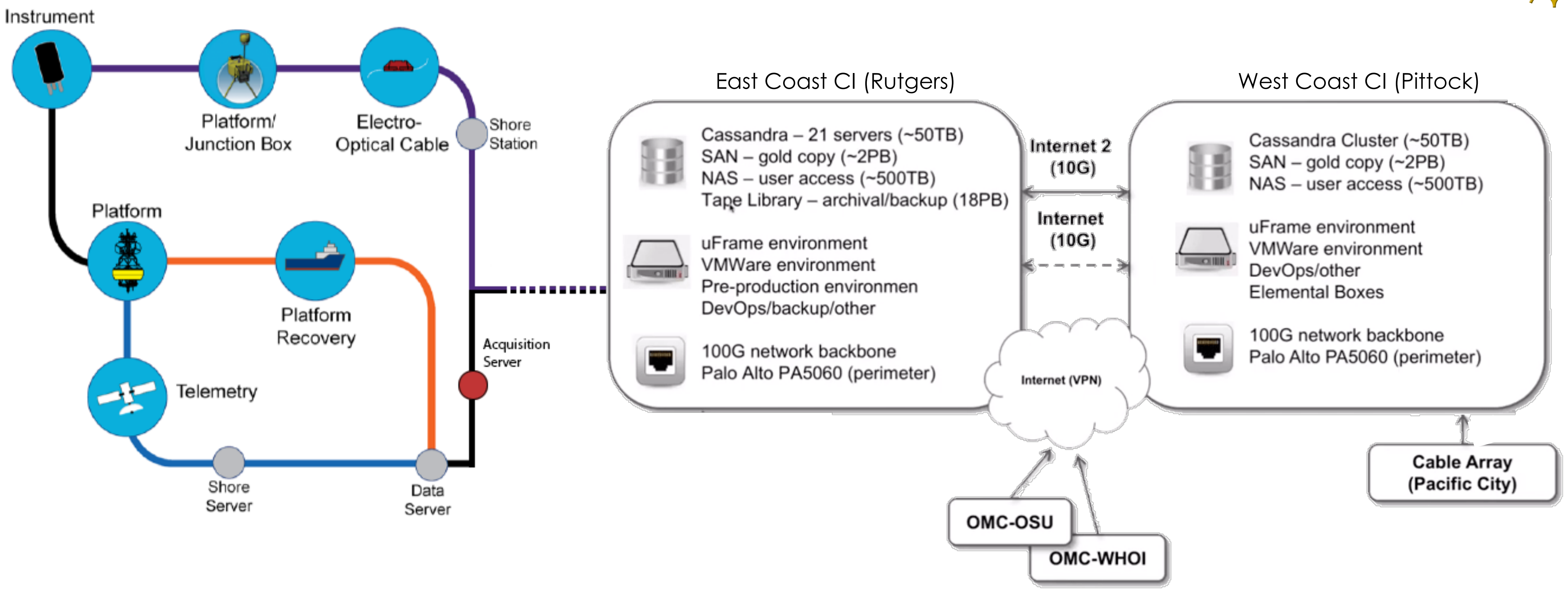
Other Processing

IceCube materials courtesy: Dr. Benedikt Riedel, Wisconsin IceCube Particle Astrophysics Center

NEON Logical Data Flows



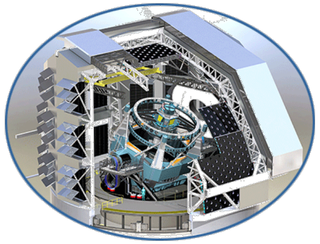
OOI Logical Architecture



* OOI materials courtesy: Dr. Ivan Rodero, Rutgers Discovery Informatics Institute

LSST Logical Architecture

Summit Facility Cerro Pachon Telescope



600 Gbps →

Base Facility, La Serena (Data Access Center)

Processing
- real time alert generation
- initial detector cross-talk correction
- metadata creation

Computing: ~ 2400-3000 cores
File storage: ~ 4 PB (user file workspace)
DB storage: ~ 3 PB (user DB storage)

Data retention of Nightly and Data Release
Copy 1 of data

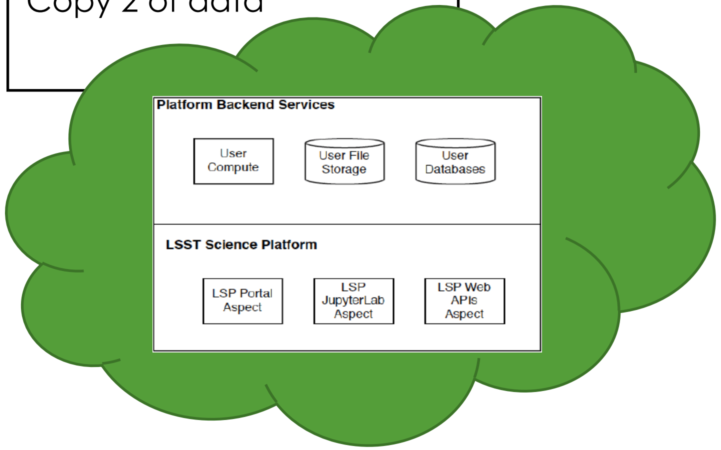
2x 100 Gbps →

NCSA (Data Access Center)

Processing
- *Night/Alert* production
- *50% of Data Release*
- *Calibration* production
- moving object

Computing: 18,000 cores
File storage: ~ 24 PB
DB storage: ~30 PB

Copy 2 of data



→

CC-IN2P3, France

Processing
- the other 50% of *Data Release*

Copy 3 of data

Disaster Recovery (DR) in the context of DLC

Existing Disaster Recovery (DR) Strategies for Some LFs across the Data Life Cycle

	Data Capture	Processing/ Filtering	Data Movement	Data Archiving and Storage	Data Access
NEON	GRAPES can buffer up to 1 month of data. Smart sensors have limited buffering capability. Replace GRAPE if fails.	No failover for compute - onsite or offsite.	If 2nd data center is built, we might see some replication there (TBD).	Replication (cloud) using AWS S3 Glacier Deep Archive for backup of ECS. ECS is replicated on site at DEN-1 Backups performed by different appliances depending on the type of storage. Plans for 2nd data center in Wyoming (for just data replication).	No existing DR strategies. Fail overs? Availability guarantees? SLA?
OOI	Replacement? How long is this data kept in retrieval locations (e.g., Pacific City)? How much is buffered or cached?	West Coast isn't used for processing but could be. Has plans for failover.	Redundant network links between East and West.	West Coast replicates data from East Coast. No automatic failover, but has plans.	Failovers planned for user access. Availability guarantees? SLA?
IceCube	Replace with a spare.	Good separation between remote and central processing. Distributed processing provides resilience.	Different ways to transmit - plane, sattelite, Internet. GridFTP to both DESY and NERSC.	At least 4 copies of data in different locations: 1 copy kept at the South Pole, 1 each in UW, DESY and NERSC.	No formal SLA. Varies based on the caching solution (e.g. xrootd).
LSST	Base Facility has a copy of data. Significant buffering planned for anticipated network failures.	Multiple facilities do processing of different types. No failover or redundant processing capabilities.	Redundant connection from BASE to NCSA. Protection against network failures for Summit to Base and from Base to NCSA	3 copies of data reside in different places: Base facility in Chile, NCSA, CC-IN2P3 (France)	Different means of access including different Data Access Centers, via web, via APIs. Availability guarantees? SLA?
Strategies for LFs	• Caching/buffering • Backup copies • Replace with a spare	• Failover compute sites?	• Plans for failover • Redundant connections	• Data replication • Backup services	• Automatic failovers for data access • Multiple data access points

Disaster Recovery (DR) Effective Processes

- **Cross-cutting finding:** Although some DR strategies exist across some stages of the DLC for some LFs, DR hasn't been taken into account to the fullest extent it warrants when designing the CI architecture for LFs.
- There is a need for some careful consideration of **requirement analysis and planning for DR** as an effective process to be followed **before and after** a possible disaster.
- Developing an effective processes guide for planning for Disaster Recovery for LFs
  - **DR Planning Phase template** that Large Facilities can follow for planning for Disaster Recovery.
  - Based on federal guidance for developing an *Information System Contingency Plan (ISCP)* after doing a thorough *Business Impact Assessment (BIA)* – **NIST 800-34r1** (https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-34r1.pdf)

Disaster Recovery Effective Processes: ISCP

- **Information System Contingency Planning (ISCP) major steps**
  - **Overview:** Top-level view of CI DR needs of the LF and serves as a summary and context.
  - **Conduct Business Impact Analysis (BIA)**
    - **Mission, data life cycle, and recovery criticality:** CI for the entire data life cycle is identified and impact of disruption to those systems is determined along with ***outage impacts*** and ***estimated downtimes***.
    - **Resource requirements:** Thorough evaluation of resources required to restore systems and processes supporting the data life cycle and related interdependencies.
    - **Recovery priorities for system resources:** Link system resources to critical mission and business processes for LFs and establish priority levels for sequencing recovery activities and resources.
  - **Create Contingency Strategies**
    - (a) Backup and Recovery; (b) Backup methods and Offsite Storage; (c)Alternate Sites; (d) Equipment Replacement; (e) Cost Considerations; (f) Roles and Responsibilities;)g) Plan testing, training and exercises

We have created example DR Planning Phase templates for NEON and IceCube and engaged with them to validate and refine the template. We plan to engage with other LFs.

DR Planning case study – IceCube LF

DR Effective Processes: BIA: Mission, DLC, Criticality

- **Mission and data life cycle example (IceCube)**

Mission/Business Process	Description
Collect observational data for scientific inquiry	Operate an array of photodetectors buried in a cubic kilometer of ice at the south pole
Ingest data from sensors for further use	Operate networks at the south pole, satellite and physical data transport to UW-Madison, and bulk internet transfer to DESY-ZN and to NERSC (National Energy Research Scientific Computing Center). Notification of neutrino events to worldwide observatories, both radio and optical.
Process data for QA, events, and production of private and public datasets	Operate clusters of computers to analyze data for interesting events, QA, and calibration. Data reduction for low bandwidth satellite link.
Archive data for future use	Operate a curated repository of collected and processed data redundantly copied across three institutions.
Disseminate data	Provide **alerts** for interesting astrophysical events, i.e. during L1 analysis, if an event of interest is detected, an alert is sent out immediately. Operate an access portal to support querying the collection and accessing the stored and/or computed results. Provide well documented access methods and ensure that access methods can be added into the far future as needed.

DR Effective Processes: BIA: Mission, DLC, Criticality

- **Recovery Criticality – Outage Impacts example (IceCube)**

Outage Impacts

Impact category: {Severe, Moderate, Minimal}

Data Life Cycle Stage	Impact Category			
	Mission Impact	**Science Return**	**Cost**	**Impact**
Collect observational data for scientific inquiry	Severe	Severe	Severe	Severe
Ingest data from sensors for further use	Severe	Moderate	Moderate	Severe
Process Data for Multimessenger GCN Alerts	Severe	Severe	Moderate	Severe
Process data for QA, and production of private and public datasets	Moderate	Moderate	Minimal	Moderate
Archive data for future use	Severe	Severe	Severe	Severe
Disseminate data - Alerts	Severe	Severe	Minimal (short term)	Severe
Disseminate data – Level 2+ processed products	Moderate	Minimal	Minimal	Moderate

DR Effective Processes: BIA: Mission, DLC, Criticality

- **Recovery Criticality – Estimated Downtimes example (IceCube)**

Estimated Downtime

MTD: Maximum Tolerable Downtime
RTO: Recovery Time Objective
RPO: Recovery Point Objective

Data Lifecycle Stage	MTD	RTO	RPO
Collect sensor data for scientific inquiry – operation of the actual sensors	1 hour	1 hour	N/A
Ingest data from sensors for further use	Optical DOMs to IC Lab: 2 days; IC Lab to South Pole Lab: 6 months	1 day from DOM; 1 month from lab to station	1 day
Process data for QA, events, and production of private and public datasets	Events: 1 hour; All others: 1 month	Events: 30 minutes; Others: 4 days	Equal to recovery time
Archive sensor data for future use	6 months w/ no loss	1 hr	1 day
Disseminate data - alerts	6 hours	3 hours	None
Dissemination – Level 2+	1 year	1 month	1 day

DR Effective Processes: BIA: Resource Requirements

- **Resource requirements example (IceCube)**
  - For each step in the Data Lifecycle and the identified level of recovery criticality, describe resources used for day-to-day operation and the resources needed for recovery in the event of an outage.

System Resource/Component	Platform/OS/Version (as applicable)	Description
Collect observational data for scientific inquiry	DOMs (custom)	Sensors buried in the ice. Also power, IceCube Laboratory Facilities, staff
Ingest data from sensors for further use	Data communication to DOMs, network to South Pole Station, rest of world	Communications from DOMs, network to South Pole station, satellite and sneakernet to rest of world. JADE software for archive and distribution worldwide.
Process data for QA and for computed results	Server clusters, other CPU/GPU resources, HTCondor compute middleware, data distribution software	Datacenters at IceCube Lab and South Pole Station with special challenges. Data centers at UW-Madison, DESY, and NERSC. Distributed computing on OSG, XSEDE, DOE resources.
Archive sensor data for future use	JADE	Long term archive and archive management
Disseminate data	Web servers, data access/distribution middleware	Public-facing Portal, other data distribution methods (xrootd, ceph etc.)

DR Effective Processes: BIA: Recovery Priorities

- **Recovery priorities example (IceCube)**

Priority	System Resource/Component	Recovery Time Objective
1	***Collect environmental data for scientific inquiry***	1 hour
4	***Ingest data from sensors for further use***	Optical DOMs to IC Lab: 1 day; IC Lab to South Pole Lab: 2 months
5	***Process data for QA and for intermediate results***	Events: 30 minutes Others: 4 days
3	***Archive sensor data for future use***	1 hour
6	***Disseminate data – Level 2+***	1 month
2	***Disseminate Alerts***	3 hours

Data Life Cycle and DR Planning: Lessons Learned

- DLC is **ONE** way to **learn, reason and catalog the CI functionalities** at each stage of data operation for LFs.
- DLC abstraction helps reasoning about
  - What *services are offered* by each DLC stage ?
  - What *CI architectural elements* support each DLC stage ?
- There are both *fundamental commonalities and differences* across LFs for DLC.
- *Devil is in the details* for both for DLC and DR; Many a time, specific elements or types of data are prioritized.
- *Heterogeneity of data processing* – the set of processes handling the data differs according to the type of data.

Data Life Cycle and DR Planning: Lessons Learned

- Heterogeneity of tools and CI stacks.
- DR planning template was a good way to start thinking about disaster recovery – a *framework to document and prioritize CI architecture and operations*.
- DR planning template can be used by LFs to *quantitatively justify CI elements* for construction or future enhancements.
- *Effective communication between LF CI professionals and CiCOE Pilot team* was key to thorough understanding; Documentations and reports are necessary but not sufficient.

Thank you !!

Questions ?

Feedback