Ceph Feature Request

from the DKIST Data Center

Add a service backed by tape that is analogous to AWS Glacier

Overview of the DKIST Project

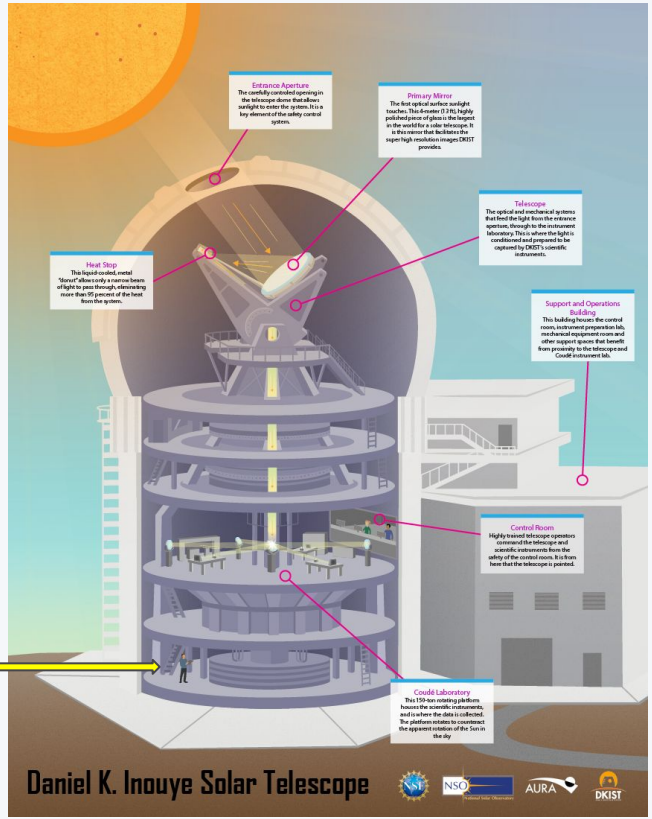
Telescope Overview

DKIST

- is an acronym for Daniel K. Inouye Solar Telescope
- is a four-meter solar telescope on the island of Maui, Hawai'i
- is the largest solar telescope ever constructed
- has a cutting-edge instrument suite
- gathers unprecedented data and images from the Sun's surface to its lower atmosphere
- observes fundamental solar magnetic and plasma processes
- provides scientists data to investigate
    - the structure and dynamics of magnetic fields
    - magnetic field interactions with the embedding plasma
    - transport of energy through the Sun
    - flares and coronal mass ejections

Telescope Overview

person for scale

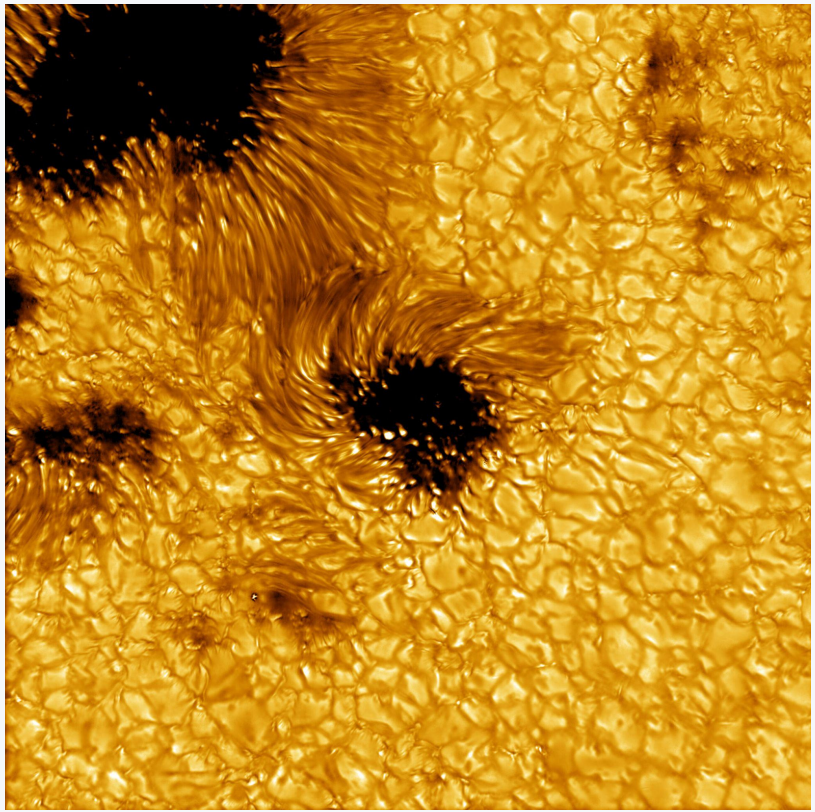


Data Center Overview

DKIST Data Center

- is located in Boulder, Colorado
- consists of application services running in a private cloud
- receives data sent via globus from the telescope into an S3 bucket in Ceph
- ingests received data that meets requirements into another S3 bucket in Ceph and archives a copy offsite to AWS Deep Glacier
- programmatically calibrates ingested data via pipelines and stores calibrated data in Ceph
- makes calibrated data available for search and download from Ceph via a data portal
- provides a set of user tools to assist navigating and downloading large datasets

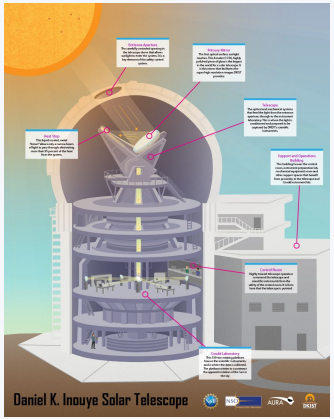
Early example of DKIST imaging



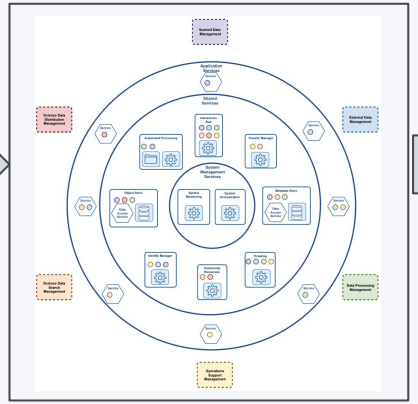
Data leading to this image were acquired with the Visible Broadband Imager blue channel at a wavelength of 450 nanometers.

Data were acquired from observations on May 11, 2021 (not from the first science observation).

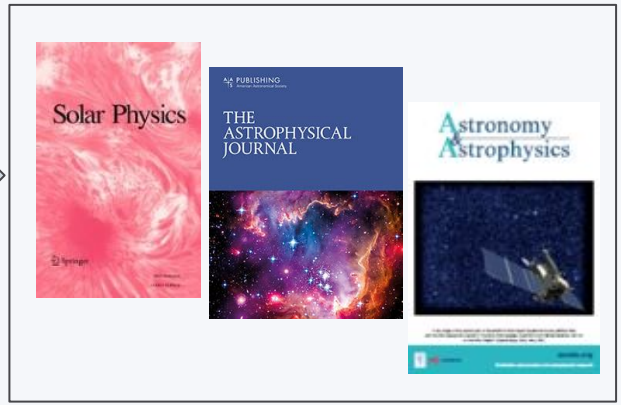
Vastly oversimplified view of DKIST “sunlight to science”



globus
data transfer



globus
data transfer



Telescope
Maui

- *Capture sunlight*
- *Generate data*

Data Center
Boulder

- *Get telescope data*
- *Archive telescope data*
- *Calibrate telescope data*
- *Archive calibrated data*
- *Provide search of calibrated data*
- *Provide access to calibrated data*

Scientists
Worldwide

- *Search calibrated data via portal and User Tools*
- *Get calibrated data via portal and globus*
- *Do science*
- *Publish*

Feature Request

Add a service backed by tape that is analogous to AWS Glacier

Key advantages and use cases for tape with Ceph

Key advantages

- potential to reduce storage cost
- reduce space utilization
- reduce power consumption
- reduce cooling load

Main use cases

- Long-term archive
  - long-term storage of datasets characterized by
    - infrequent access
    - large size (30+ GB)
    - acceptable availability for download in hours or days
  - *tape lifespan depends on environmental conditions, hardware, and use*
- Ransomware protection

Additional use cases

- Disaster Recovery
- Backup and restore

Current options for using tape with Ceph

- Third party S3 API implementation
  - Spectra Logic Black Pearl in front of tape library
    - *drawback: did not provide required API functionality*
- Third party application
  - Spectra Logic StorCycle in front of Black Pearl and tape library
    - can either leave object in Ceph or remove it from Ceph
    - *drawback: manual process breaks design goal of full automation*
    - *drawback: level of effort to integrate exceeds available capacity of team*
  - Nodeum
    - *Marketing case study claims integration with ceph, have not investigated offering*
- Intercept S3 API in Ceph
  - Tsolo Blue, Thomas Bennett and Martin Slabber presented at Cephalocon 2023

Desired service features for Ceph tape storage

native integration

- Ceph manages the operations of tape infrastructure needed for its purposes
- `ceph` command includes category for tape

transparent media transitions

- maintain RGW client's object name as an object transitions between types of storage
- automate strategy to avoid single point of failure and assure integrity of data on tape
- programmatically manage S3 requests for objects stored on tape

programmatic migration of tape infrastructure

- provide a method to automate workflows to migrate to new versions of tapes, tape libraries, and associated software for managing tape infrastructure lifecycle

capability to air gap service's hardware and software

- to protect against ransomware attacks

Conceptual workflow to set up Ceph tape storage

1. create pool that will use tape for storage

```
ceph osd pool create default.rgw.glacier.data <pg_num> <protection>
```

Unclear how pg_num and protection would be applied to tape storage

2. add new storage class to zonegroup

```
$ radosgw-admin zonegroup placement add \
      --rgw-zonegroup default \
      --placement-id default-placement \
      --storage-class GLACIER
```

3. configure zone placement for the storage class

```
$ radosgw-admin zone placement add \
     --rgw-zone default \
     --placement-id default-placement \
     --storage-class GLACIER \
     --data-pool default.rgw.glacier.data
```

4. configure S3 lifecycle

Examples on web are 2 to 3 years old and use s3cmd with xml rules file.
Request: `ceph` commands to set, get, modify, delete S3 lifecycle

Conceptual placement with Ceph tape storage

```
$ radosgw-admin zone placement get --placement-id default-placement
{
    "index_pool": "default.rgw.buckets.index",
    "storage_classes": {
        "STANDARD": {
            "data_pool": "default.rgw.buckets.data"
        },
        "TAPE": {
            "data_pool": "default.rgw.tape.data"
        }
    },
    "data_extra_pool": "default.rgw.buckets.non-ec",
    "index_type": 0
}
```