

OPEN SOURCE TOOLS FOR POINT CLOUD PROCESSING, STORAGE, SUBSETTING, AND VISUALIZATION

Kevin Beam

Professional Research Assistant

National Snow and Ice Data Center (NSIDC)

~OR~

GETTING DATA BY FLYING AROUND AND SHOOTING LASERS





University of Colorado
Boulder





Manage and distribute scientific data

Create tools for data access

Support data users

Perform scientific research

Educate the public about the cryosphere

WHAT WE'LL TALK ABOUT

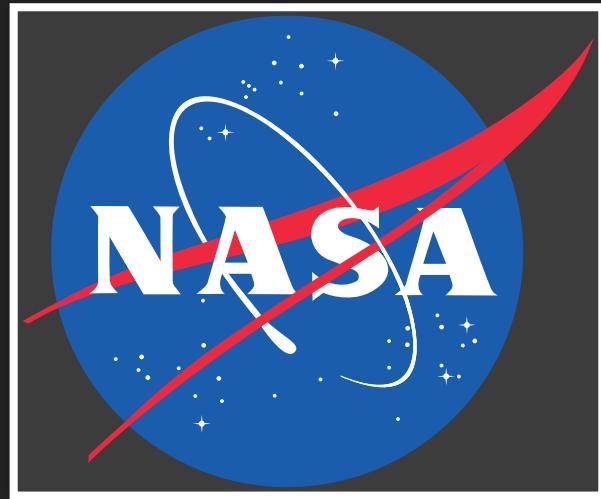
Operation IceBridge and LIDAR

Processing Point Cloud Data

Storing It Efficiently

Subsetting

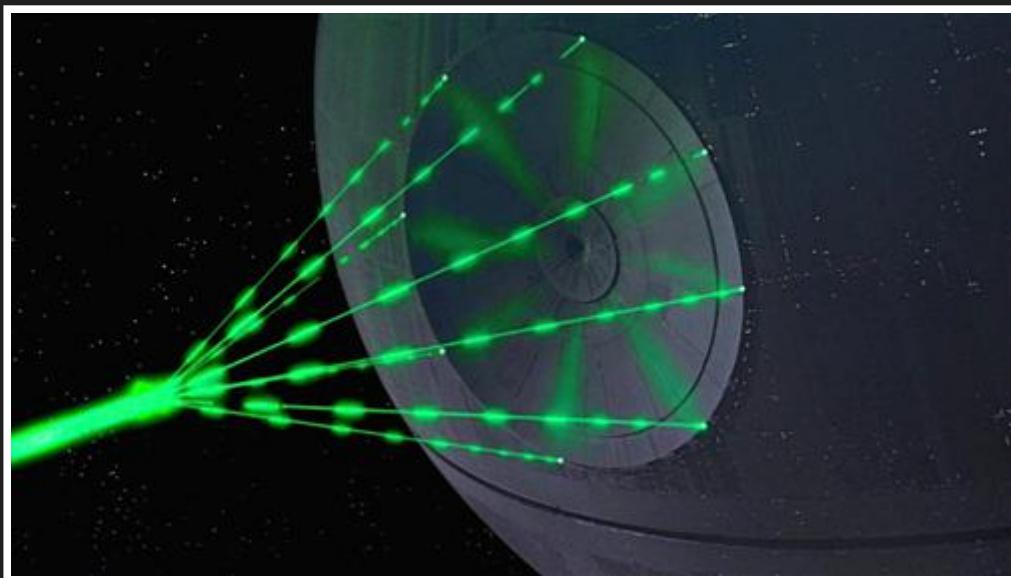
Visualization



OPERATION ICEBRIDGE

ICESat

2003 ...



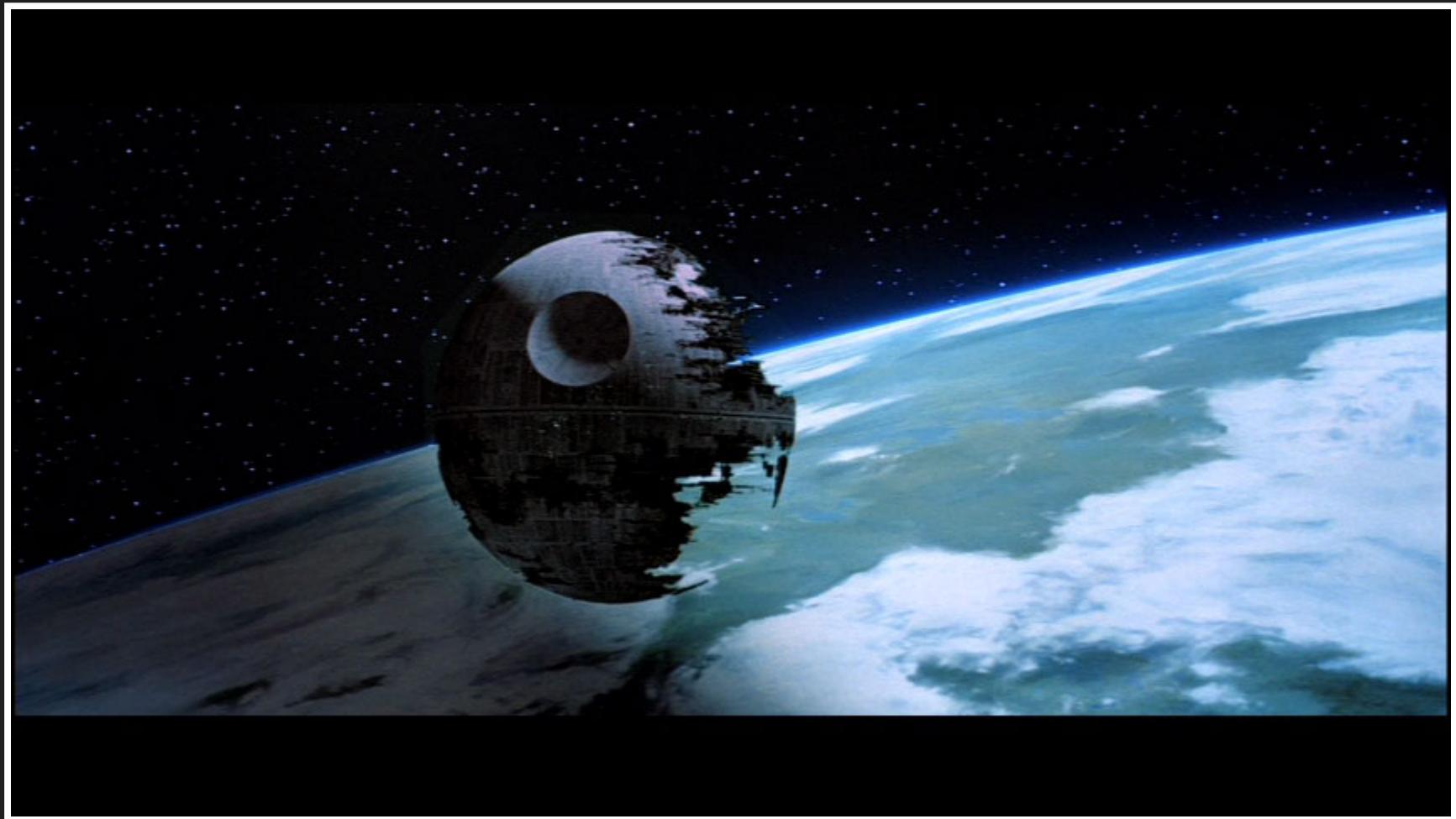
ICESat

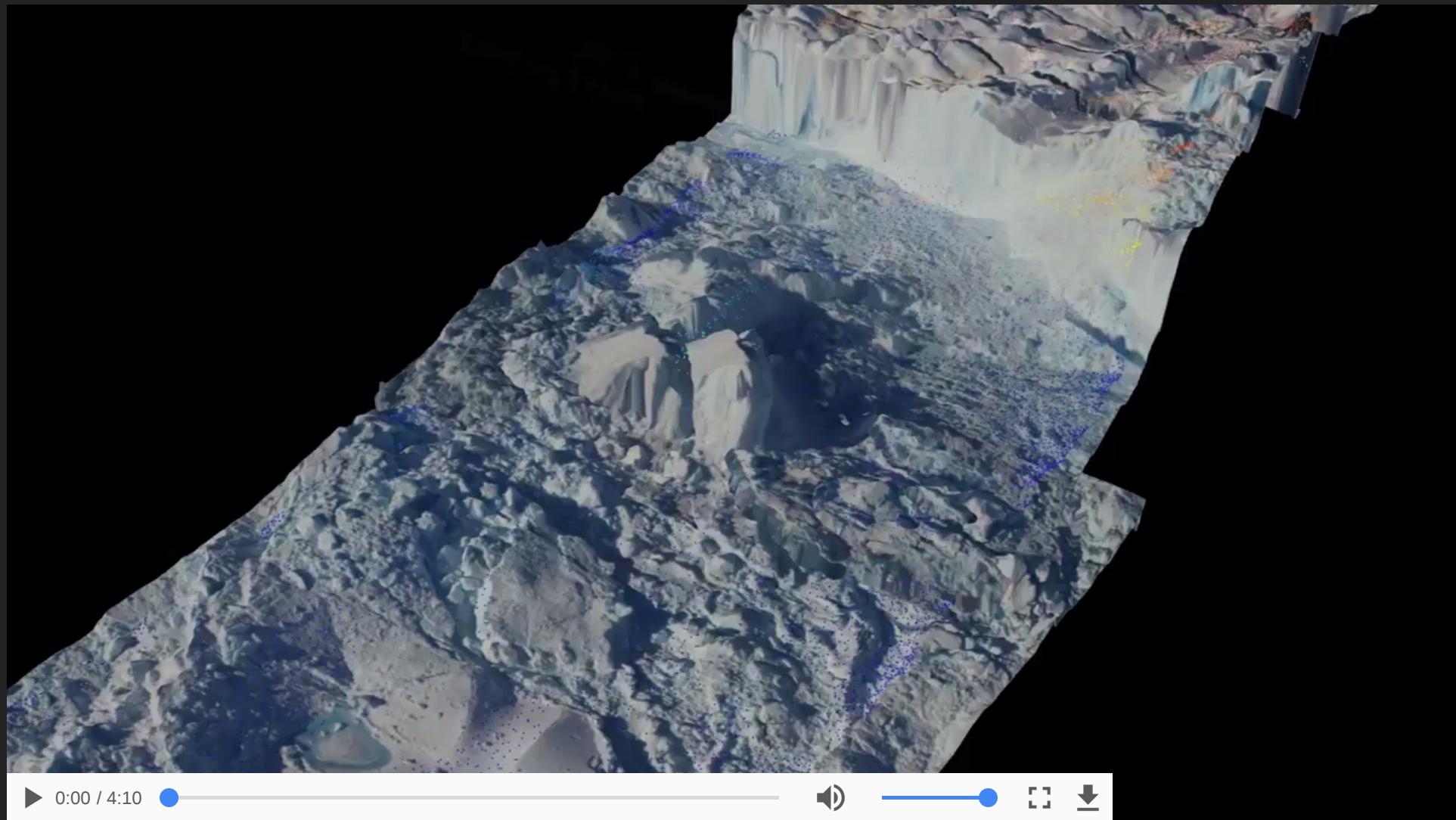
... 2009



ICESat-2

2018+



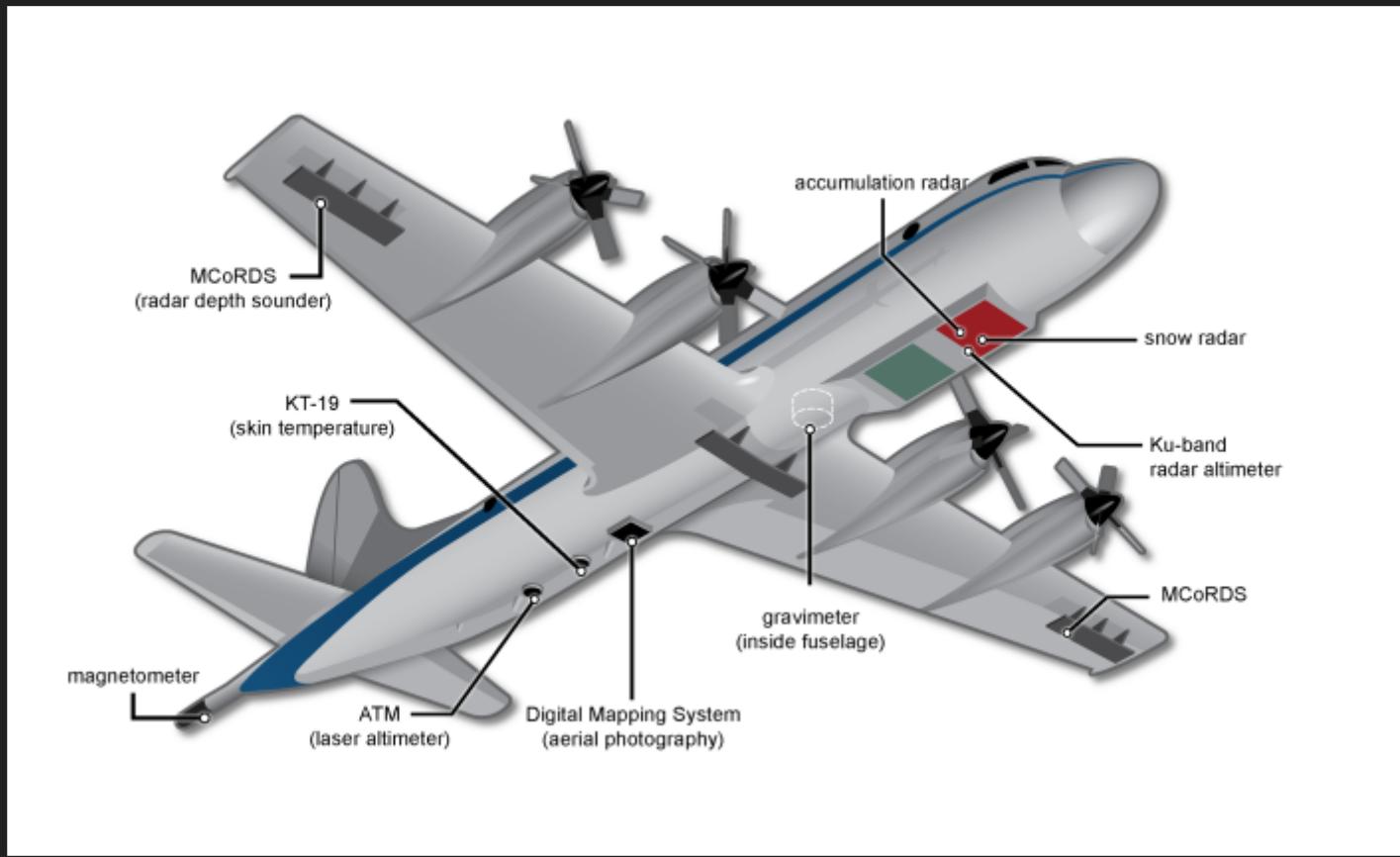


▶ 0:00 / 4:10

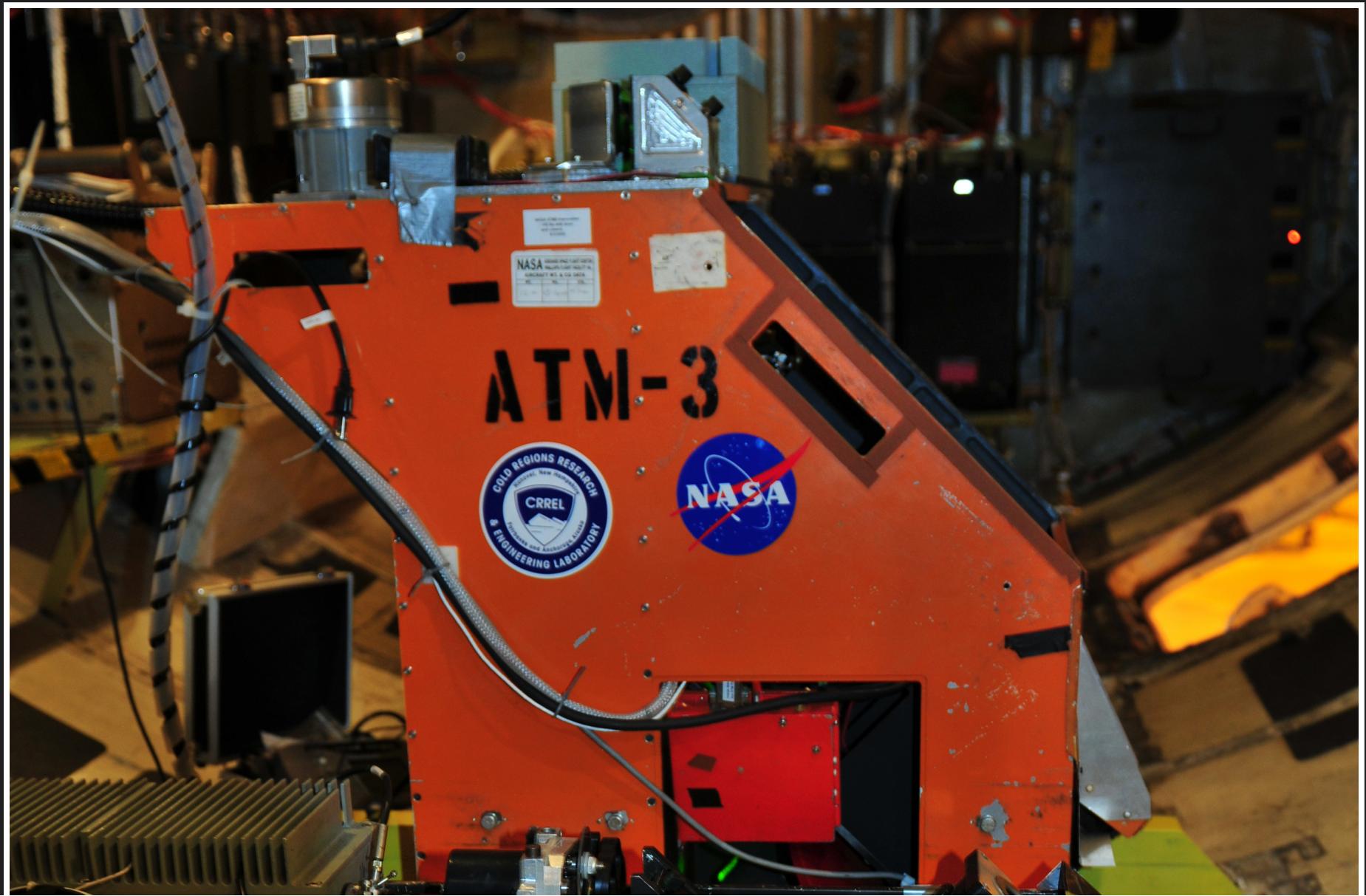


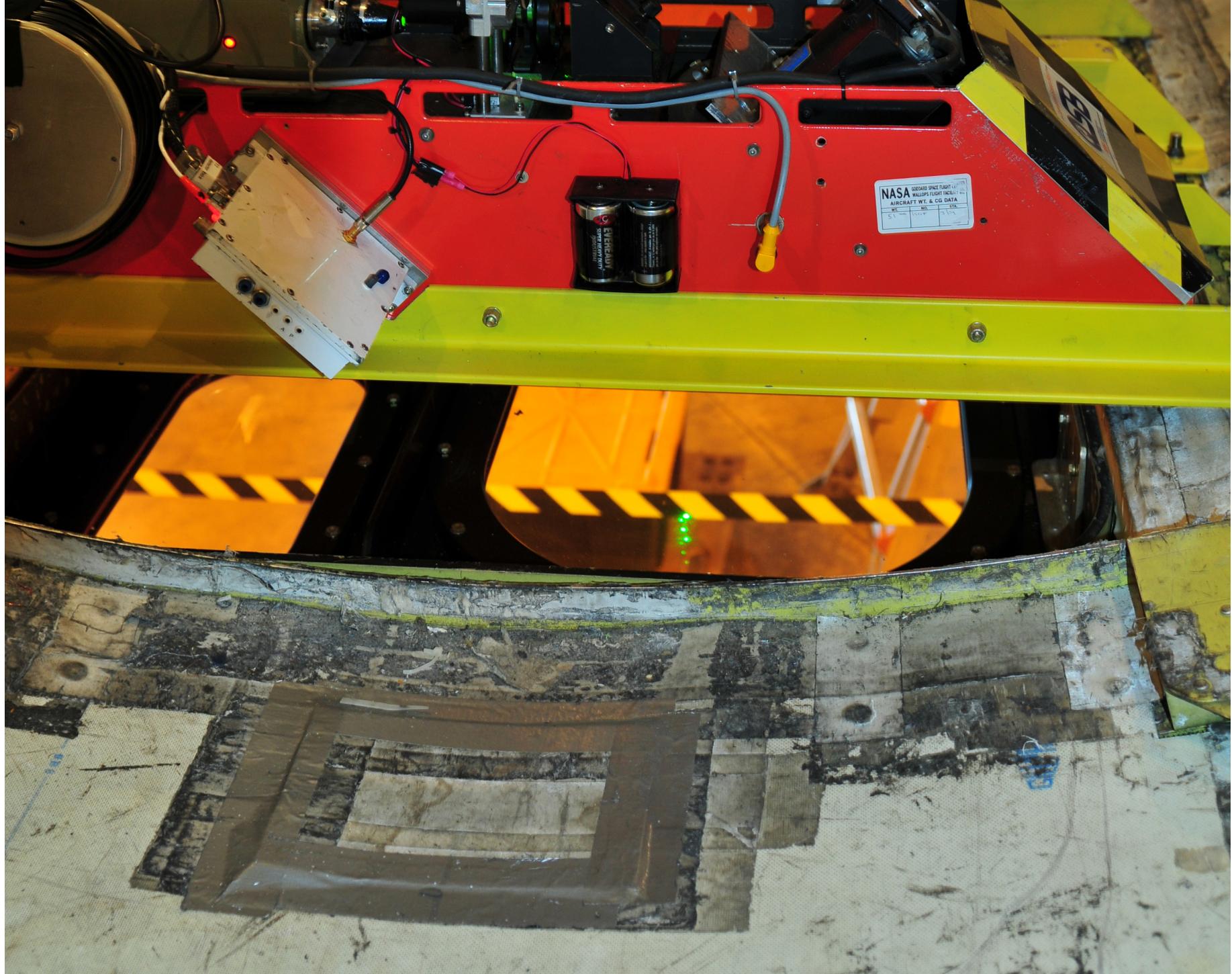
—





Airborne Topographic Mapper (ATM)





MONTE CARLO

Laser Vegetation Imaging Sensor (LVIS)



THE DATASETS

LVIS1B

- 2009 through 2015
- Two distinct datasets: v1 and v2
- v1: Big-endian binary format
- v2: HDF5 format
- Data contains seconds of day, not date/time
- Longitude needs to be shifted to -180 to 180
- Three sets of lat, lon, elevations to pick from!

ATM1B

- 1993 through 2017
- Three distinct datasets: pre-IceBridge, v1, and v2
- Pre-IceBridge:
 - Qfit binary format
 - Three distinct formats
 - Both big and little endian!
- v1: Qfit Binary format
- v2: HDF5
- Data contains GPS seconds, not UTC date/time
- Longitude needs to be shifted to -180 to 180

Using the data is complex!

But now: which files do you need?

- When?
- Where?

ICEBRIDGE PORTAL

Solves *some* of these problems...

[Demo](#)

But here's what we'd *really* like to be able to do:

- Specify a date/time range
- Specify a spatial area
- Get one format: HDF5
- Don't make me think about endianness!
- Everything in UTC date/time
- Normalized latitude & longitude

VALKYRIE

Simple RESTful subsetting service

Valkyrie API

Created by NSIDC User Services

See more at <https://nsidc.org/about/contact.html>

[Contact the developer](#)

default

Show/Hide | List Operations | Expand Operations

GET

/

GET

/ATM1B

GET

/ILVIS2

[BASE URL: , API VERSION: 0.0.0]

NOW HOW DO WE IMPLEMENT THIS
BEHIND THE SCENES?

WOULDN'T IT BE NICE IF WE COULD JUST
PUT THE DATA IN A DATABASE AND
WRITE A SIMPLE QUERY?

We can with:

PostgreSQL

PostGIS

pgpointcloud

WHY PGPOINTCLOUD?

Couldn't we store each point as a row in a table?

We could...up to a point... (<-- pun)

To scale up to billions of points, we need to optimize

PGPOINTCLOUD CONCEPTS

- pgPointCloud extends PostgreSQL and PostGIS
- A point can contain arbitrary attributes
- A point's attributes are described by a schema
- A point is stored in a new PG type
- Points are stored in groups called patches
- Patches are compressed bundles of points
- Patches typically contain hundreds or thousands of points

COOL. WHAT CAN WE *DO* WITH IT?

Write simple SQL queries!

- Geospatial conditions
- Temporal conditions

Show Me!

A polygon

```
'SRID=4326;POLYGON((  
    -49.699 69.081,  
    -48.333 69.142,  
    -48.280 69.270,  
    -49.846 69.209,  
    -49.699 69.081))'::geometry
```

A date/time range

```
SELECT  
    PC_Get(PC_Explode(  
        PC_FilterBetween(pa, 'DateTime', 291817682, 291817682.01))  
FROM  
    atm1b  
WHERE  
    rc_compute_range_for_a_patch(pa, 'DateTime') \  
    && NUMRANGE(291817682, 291817682.01);
```

HOW DO WE GET DATA INTO THE DATABASE?

We need to:

- Condition the data
- Insert these Point and Patch things...

PDAL TO THE RESCUE

- PDAL: Point Data Abstraction Library
- As GDAL handles raster and vector data, PDAL handles point data

WHAT DOES THAT MEAN?

PDAL Lets You Create a Pipeline From:

- Readers for a wide variety of formats
- Processing nodes or "filters"
- Writers for various file formats
- Including...
- Pgpointcloud

SO WE MADE A SIMPLE PIPELINE

```
Files ->
-> Python
-> Better Files
-> PDAL Reader
-> PDAL Chipper
-> PDAL Writer
-> PostGIS Database
```

```
PIPELINE = {
    "pipeline": [{"  
        "type": "readers.text",  
        "filename": "tmpfile",  
        "separator": ","  
    }, {  
        "type": "filters.chipper",  
        "capacity": 400  
    }, {  
        "type": "writers.pgpointcloud",  
        "connection": DB_STR,  
        "table": "atm1b",  
        "overwrite": "false",  
        "pcid": "1"  
    }]  
}
```

```
$ wget "http://VALKYRIE.nsidc.org/ATM1B?\n  polygon=\n    -51.599741,71.582102,\n    -52.046274,71.816551,\n    -50.007416,72.303672,\n    -49.114916,71.792674,\n    -51.599741,71.582102&\n  time_range=\n    2011-03-31T00:00:00,\n    2011-03-31T23:59:59"
```

VISUALIZATION

Can we see this point cloud you speak of?

ENTWINE

- "Point cloud indexing for massive datasets"
- "Designed to conquer datasets of hundreds of billions of points"
- Most Importantly...
- It can ceate Cesium 3D Tiles

CESIUMJS



Let's take a look...

Press F11 to exit full screen



▶ 0:00 / 5:40



Other Tools

- Greyhound
- speck.ly
- plas.io
- potree

The experts:

Howard Butler

Connor Manning

Hobu, Inc.

Thank you!

Questions?