Exploring the advantages of cloud native, easily consumable, scalable formats for downstream scientific exploitation of point cloud data

Trevor Skaggs, Matthew Hanson, Dan Pilone

Element 84

About the Project

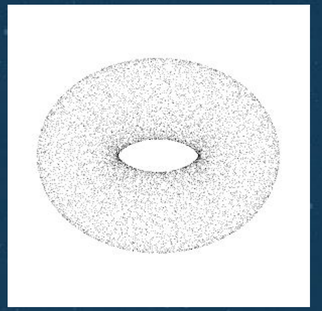
- Project Objective: To perform a proof of concept (small scale) transformation of ICESAT2 ATL06 product into a cloud native, easily usable, and scalable format for downstream scientific exploitation.
- Work is being performed under the NASA ACCESS Grant Program that works towards increasing the scalability and accessibility of all large NASA satellite datasets.

Who is Element 84?

- Element 84 is a company that aims to help organizations and government agencies (NASA, USGS, etc) solve problems using big remote sensing, life sciences, and transportation datasets.
- Trevor Skaggs - Senior Geospatial Engineer
- Matthew Hanson - Senior Software Engineer
- Dan Pilone - CEO/CTO

What is a Pointcloud?

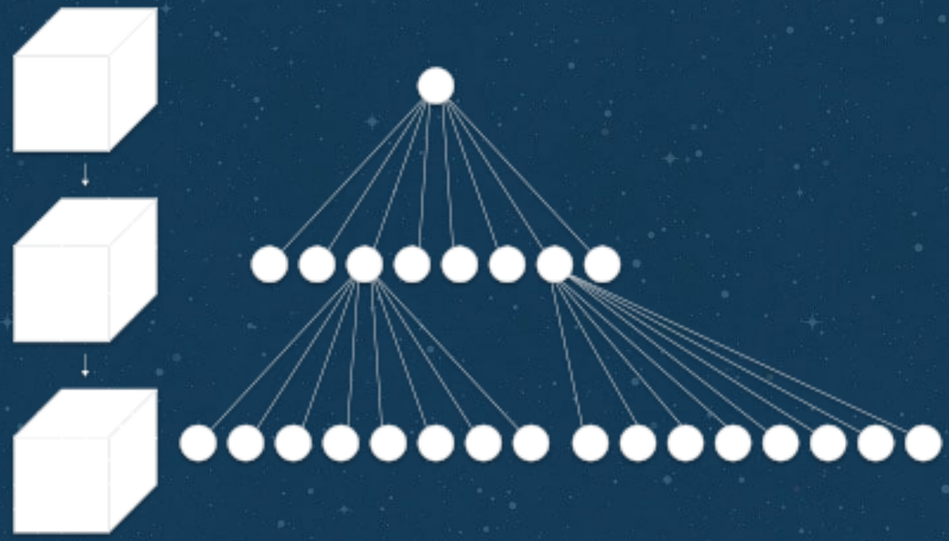
- Point clouds are a collection of points that represent a 3D shape or feature. Each point has its own set of X, Y and Z coordinates and in some cases additional attributes.
- Point clouds are most often created by methods used in photogrammetry or remote sensing. Remote sensing is a way of collecting data of the Earth by use of satellites or aircrafts.



Source: https://community.safe.com/s/article/what-is-a-point-cloud-what-is-lidar

What is Entwine Point Tiles (EPT)?

- Simple and flexible octree-based storage format for point cloud data.
- Encoding-agnostic
- Flexible attribute schema
- Lossless



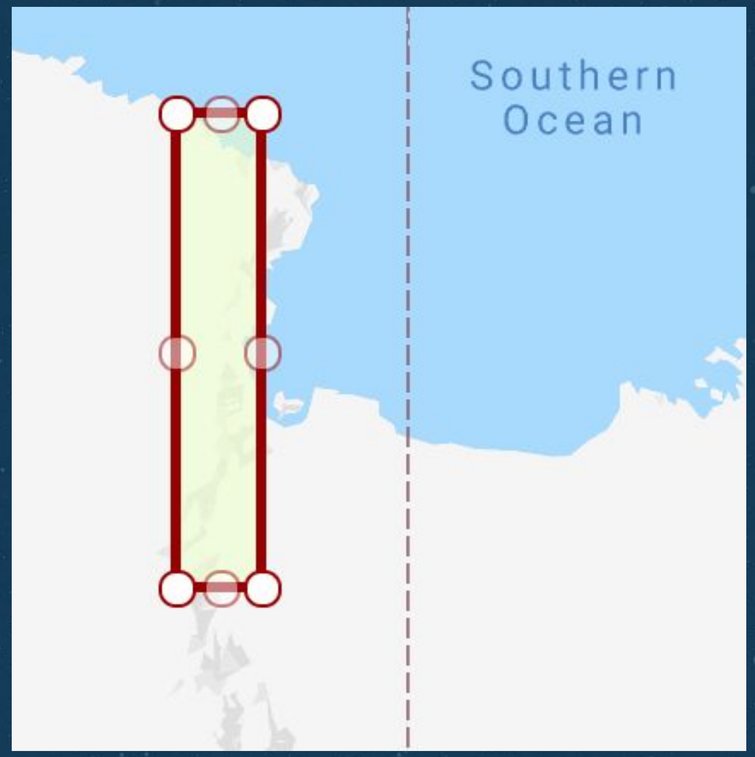
https://en.wikipedia.org/wiki/Octree#/media/File:Octree2.svg

- Supporting Groups: https://entwine.io/, https://hobu.co/

EPT Processing Overview

Sample ATL06 Data

- Bounding Box: 156,-81,162,-69
- Start Time: 2018-10-14T00:00:00Z
- End Time: 2020-05-21T18:52:30Z
- Data across Cycles 1-6



Sample ATL06 Data (cont)

- Standard Attributes Mapping
  - X: [beam]/land_ice_segments/longitude
  - Y: [beam]/land_ice_segments/latitude
  - Z: [beam]/land_ice_segments/h_li
  - GpsTime: [beam]/land_ice_segments/delta_time
- Custom Attributes:
  - atl06_quality_summary, h_li_sigma, segment_id, sigma_geo_h, CycleNumber, FileId (PointSourceId), BeamId (ReturnNumber)*

* We are using Return Number as a proxy for the beam id, the mappings are as follows:

- gt1l > 0, gt1r > 1, gt2l > 2, gt2r > 3, gt3l > 4, gt3r > 5

Sample ATL06 Data (cont) **

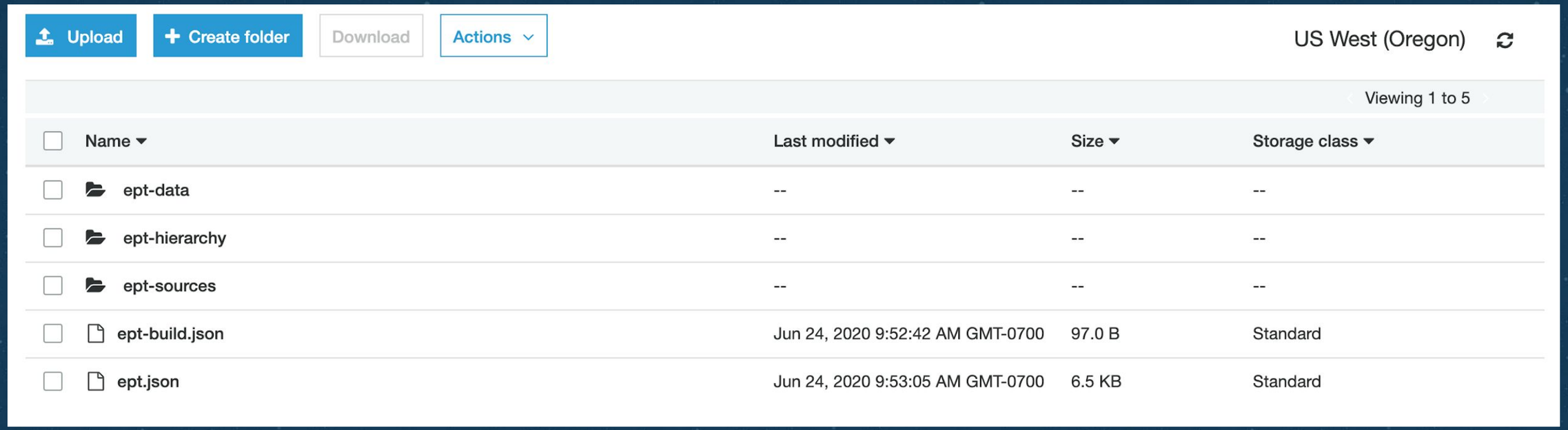
- Raw HDF5 Files
  - 1208 Total Files, 88 GB
- LAZ Files (intermediate product)
  - 1208 Total Files, 4.5 GB*
- EPT Files
  - 13,127 Total Files, 6.4 GB*

- 6.4 GB / 88 GB = ~7.25% of original storage size!

* This size will increase if we decide to add additional attributes.
** Please see “Cycle Data Breakdown” slide at end of deck for breakdown of HDF/LAZ/EPT files/size by cycle.

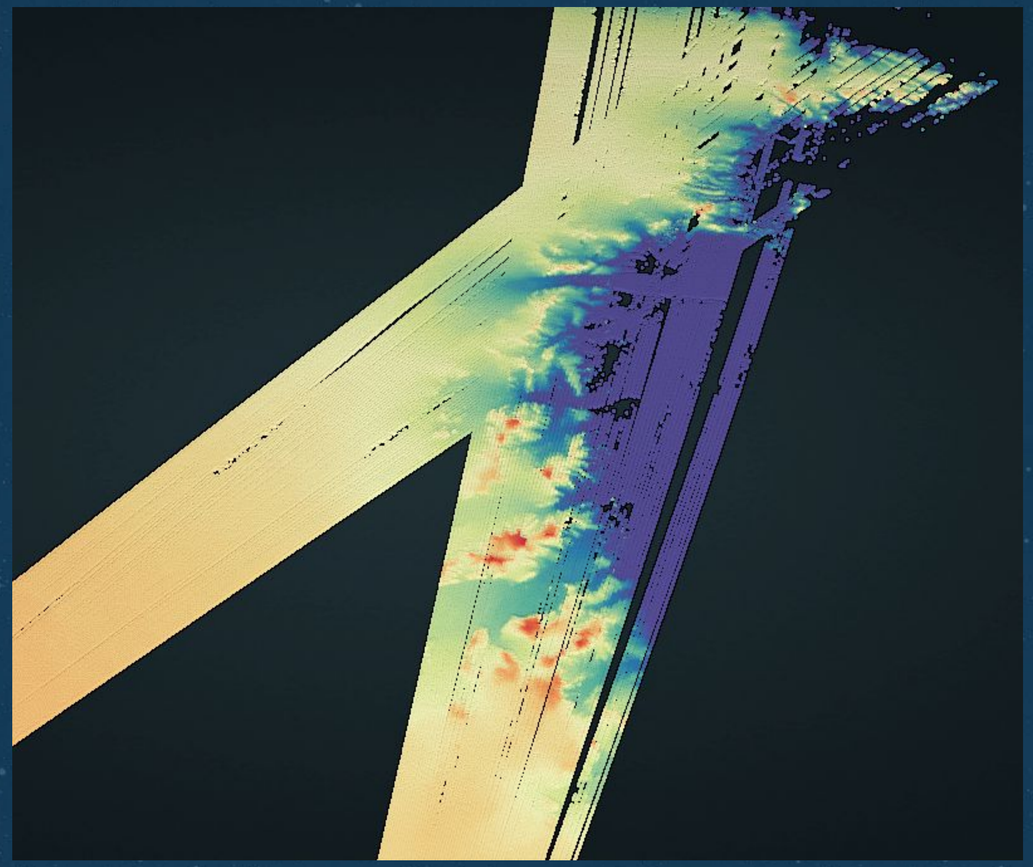
Sample ATL06 Data - Cycle 1 End Point

- Upon uploading to S3, the user is then able to access each end point directly (and without a server) via PDAL as a EPT Reader Stage
- Cycle 1 End Point Url:
  - https://s3-us-west-2.amazonaws.com/access-icesat2-entwine/1



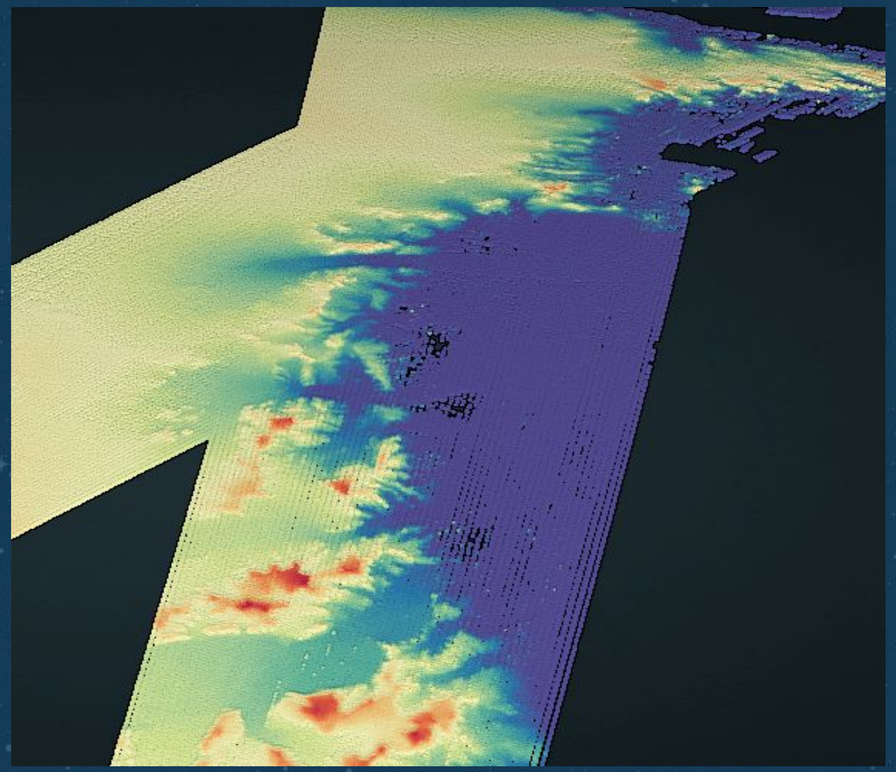
Sample ATL06 Data Visualized

Cycle 1 Only, Colorized by Elevation



https://potree.entwine.io/data/view.html?r=%22https://s3-us-west-2.amazonaws.com/access-icesat2-entwine/1%22

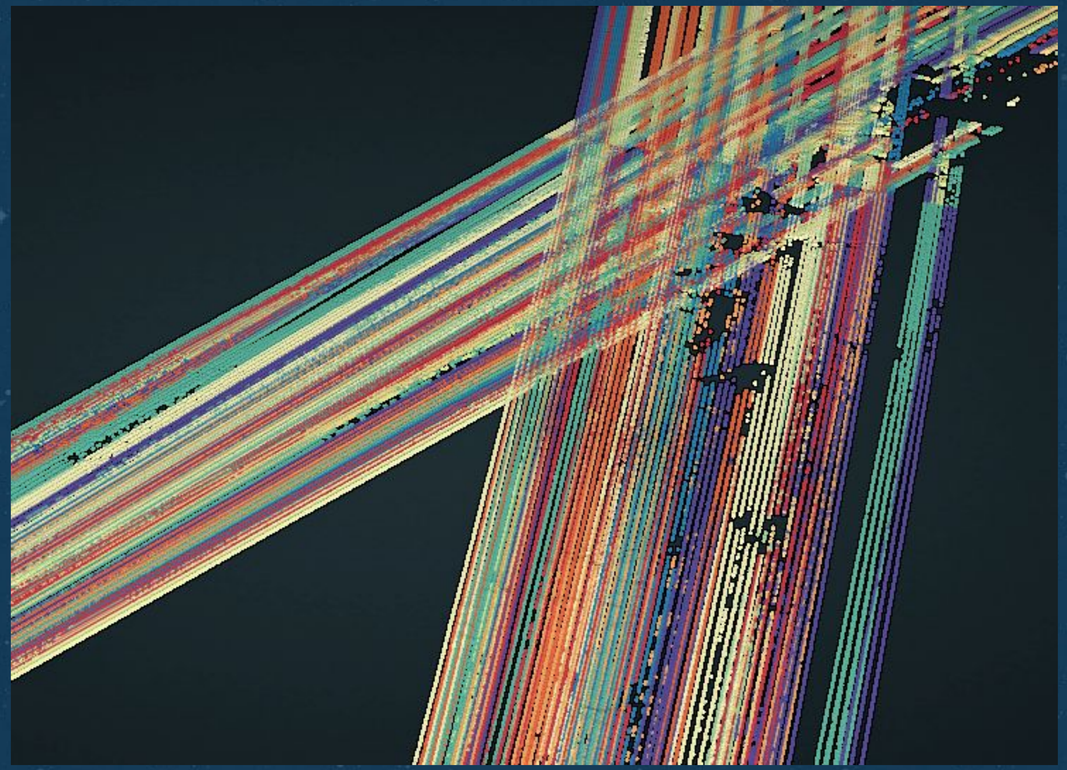
Cycles 1-6, Colorized by Elevation



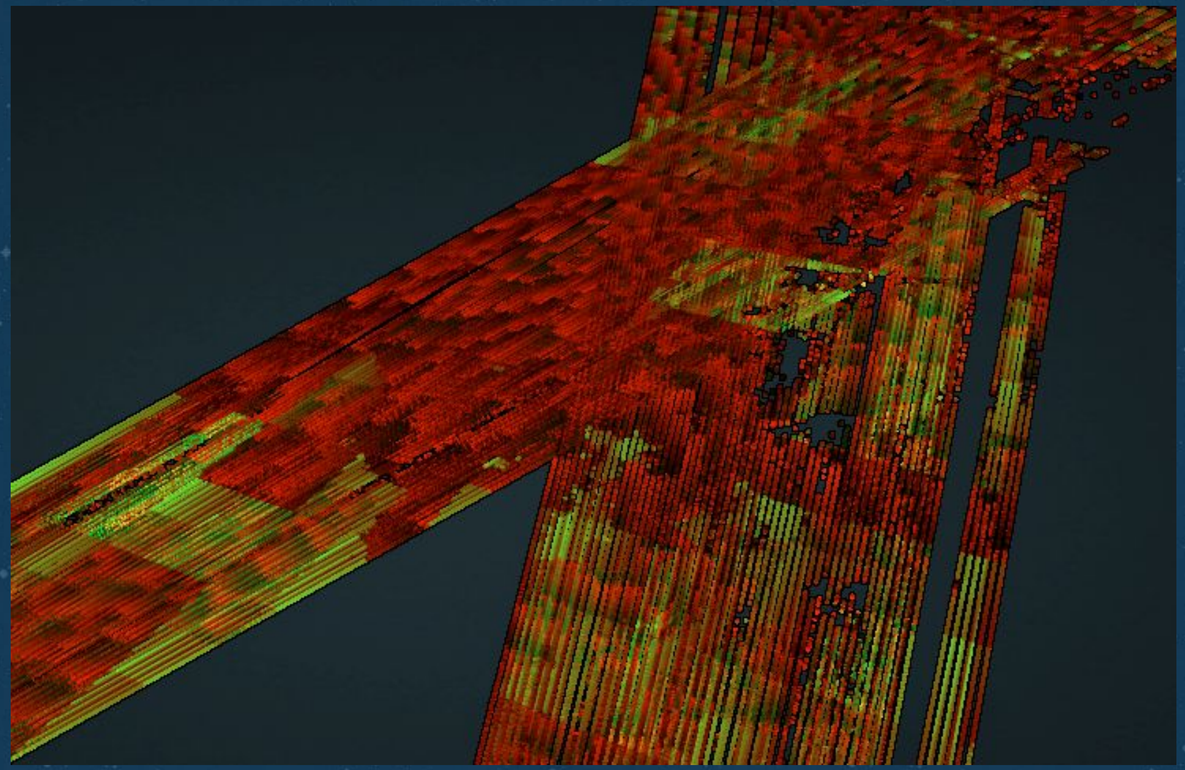
https://potree.entwine.io/data/view.html?r=[%22https://s3-us-west-2.amazonaws.com/access-icesat2-entwine/1%22,%22https://s3-us-west-2.amazonaws.com/access-icesat2-entwine/2%22,%22https://s3-us-west-2.amazonaws.com/access-icesat2-entwine/3%22,%22https://s3-us-west-2.amazonaws.com/access-icesat2-entwine/4%22,%22https://s3-us-west-2.amazonaws.com/access-icesat2-entwine/5%22,%22https://s3-us-west-2.amazonaws.com/access-icesat2-entwine/6%22]

Sample ATL06 Data Visualized

Cycle 1 Only, Colorized by PointSourceID (File ID)



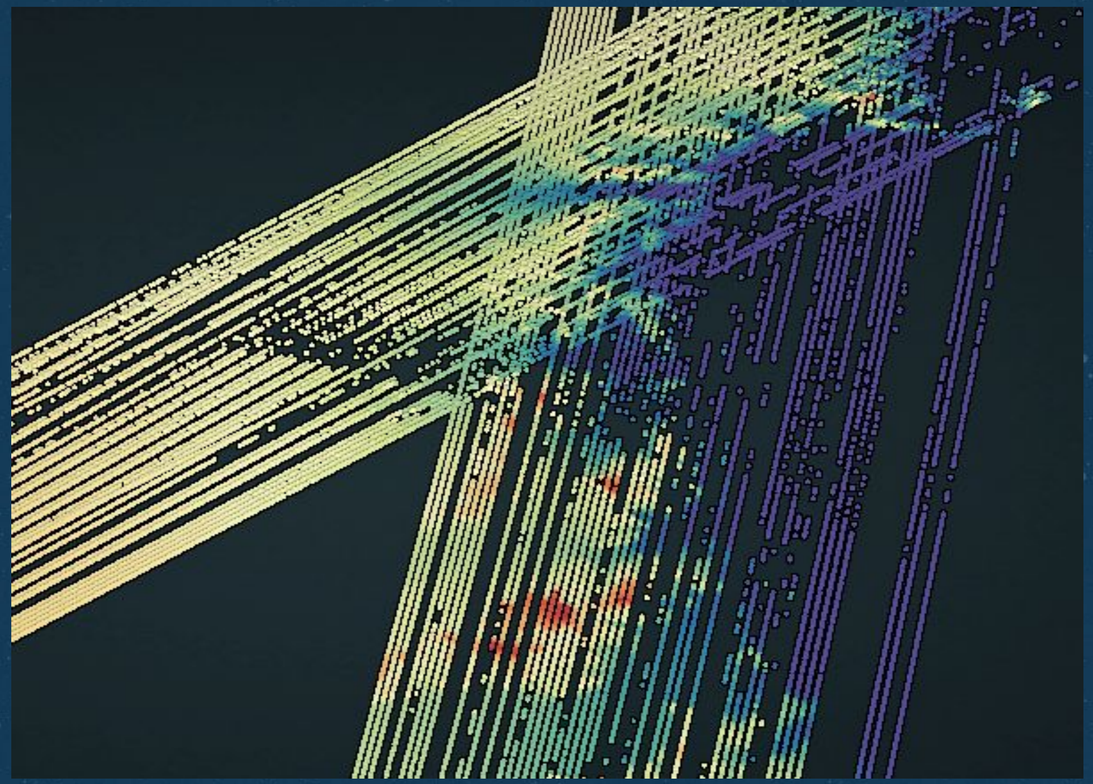
Cycle 1 Only, Colorized by Index



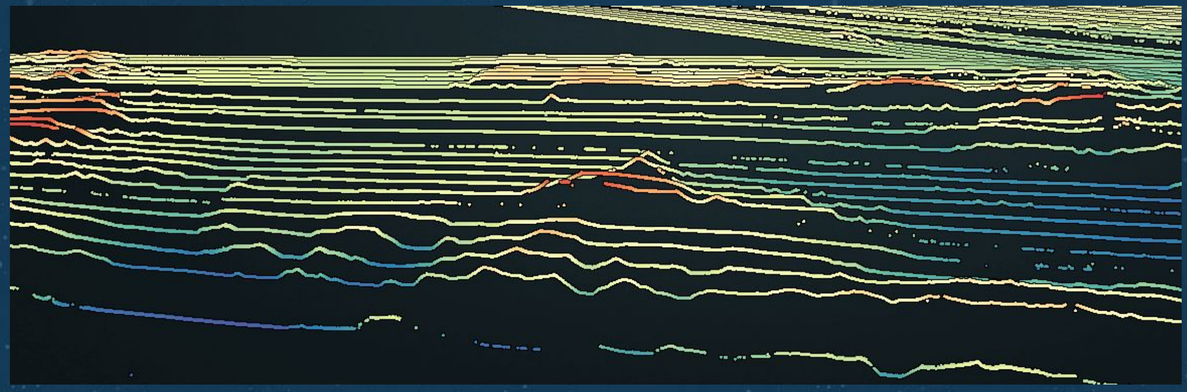
https://potree.entwine.io/data/view.html?r=%22https://s3-us-west-2.amazonaws.com/access-icesat2-entwine/1%22

Sample ATL06 Data Visualized

Cycle 1 Only, Colorized by Elevation, Filtered to gt1l Beam Only



Cycle 1, Colorized by Elevation, Filtered to gt1l/gt1r Beam Pair



https://potree.entwine.io/data/view.html?r=%22https://s3-us-west-2.amazonaws.com/access-icesat2-entwine/1%22

Thank you

Element 84

Cycle Data Breakdown

Cycle	# HDF5	HDF5 Size (MB)	# LAZ	LAZ Size (MB)	# EPT	EPT Size (MB)
1	168	13000	168	638	1,810	869
2	226	16000	226	841	2,406	1,200
3	237	17000	237	918	2,685	1,300
4	154	12000	154	604	1,774	841
5	240	18000	240	928	2,664	1,300
6	183	12000	183	635	1,788	878
Total	1208	88000	1,208	4,564	13,127	6,388