Appendix M - Program Management Lifecycle Considerations

Introduction

When considering the costs to implement, enhance, or expand a national elevation program, in addition to the data acquisition costs that have been considered in the Benefit Cost Analyses (BCA), the program will need to be supported by a robust infrastructure that provides program management and lifecycle data management support. This would include activities such as governance, outreach, contract oversight, data management (e.g., data inventory, data quality validation, data archive, etc.), data processing to prepare derivative products, data provisioning to make products available to the public in ways that support their needs, and research and development to identify and enhance technologies that can best serve a national elevation program. All of these activities will need to be supported by personnel and an information technology (IT) infrastructure that includes hardware, software, networks, and commercial cloud resources, and an IT strategy that can process and deliver the volume of data resulting from a national elevation program. Table M.1 depicts some of the major elements of the lifecycle management of a national elevation program for which costs should be considered.

Table M.1. Lifecycle cost elements of a national 3D elevation program elevation program

3D Nation Lifecycle Management Cost Elements
Program planning, outreach and governance including the development and maintenance of standards and specifications
Contract management, planning, and management of data acquisition partnerships among various federal, state, and local agencies
Data validation of all acquired data and products
3D Nation data research and development (e.g., remote sensing technologies, including airborne, spaceborne, uncrewed systems, etc.)
3D Nation data production, maintenance and management beyond what is contracted to private sector
3D Nation data/database management and delivery
IT systems support staff for 3D Nation data
IT systems for 3D Nation (servers and cloud)

Figure M.1 depicts some of the major elements of program development that are needed to establish a national 3D elevation program.

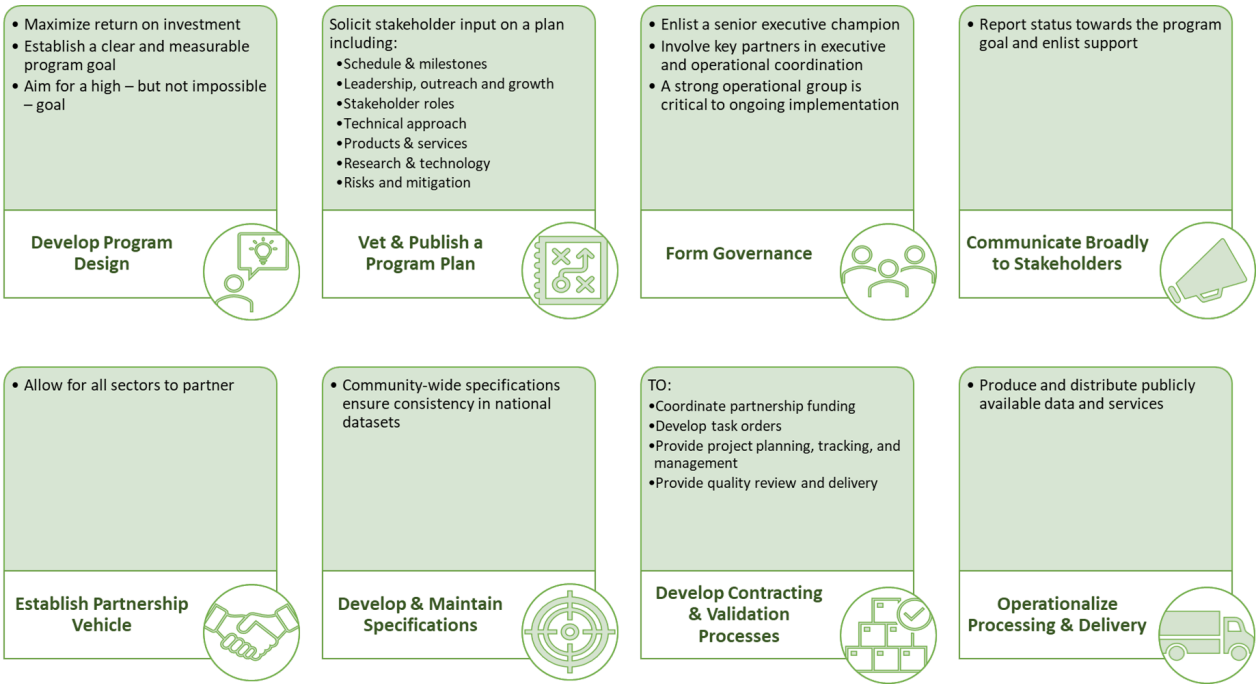

Figure M.1. Major program development elements needed for a national 3D elevation program

Once the initial national elevation program is developed, an institutional program management framework is needed. The detailed program management components may vary in order and by agency but the basic components include:

1. **Program Governance** – A strong governance framework is critical to ongoing implementation of a national elevation program.

    a. *Executive & Operational* - Key partners must be involved in the executive and operational coordination of a national elevation program. Interagency coordination is needed to establish and update the policies and procedures of a national elevation program.

    b. *Standards & Specifications* - Community-wide specifications ensure consistency in national datasets. As technologies change and programs evolve, standards and specifications must adapt, and the changes must be coordinated with the wider elevation community

2. **Outreach –** Continual outreach to partners and stakeholders is required to ensure a national elevation program continues to be responsive to the changing needs of users, to raise awareness of the program among agency geospatial leadership, and to promote best practices for interagency collaboration.

3. **Data Acquisition Coordination –** Coordinating the data-collection priorities of various stakeholders into annual acquisition planning is an ongoing task throughout the lifecycle of a national elevation program. Coordinating elevation data acquisition across stakeholders reduces the cost of acquisition through economies of scale, minimizes duplication of effort, and helps identify funding partners where agency priorities intersect. Individual data collection projects require coordination of funding partnerships, funding agreements, documentation of project requirements via task orders, and management of the contracts through the data acquisition and delivery process. Significant coordination with partners is required to ensure that the resulting data meet the needs of all.

4. **Processing and Delivery –** A vital element of a national elevation program is that the data and products are shared with the public. The Program must continually adapt its delivery mechanisms to satisfy changes in user requirements for data access, changes in technology, and the need for customized/non-standardized product generation needs. This continuous improvement of data processing capabilities and the data distribution methods must be coordinated with stakeholders as part of operational governance

5. **Research –** As technologies for collecting/processing and/or delivering elevation data change, a national elevation program must be able to assess the impact of such changes to the wider elevation community and the national elevation program itself. Research should include continuous improvement developments on operations, delivery and maximizing cost savings for producing and disseminating data

The technology and institutional framework will also need to be able to support the activities associated with collecting elevation data as acquisition occurs on an ongoing basis according to the program collection schedule. As these data are collected, they will need to be processed for inclusion in a national repository and served to users as the data become available.

When reviewing the current state of elevation data processing and future needs of a national elevation program, a simplified data management workflow such as the one shown and described below was considered. This workflow is depicted in Figure M.2 and explained further below.

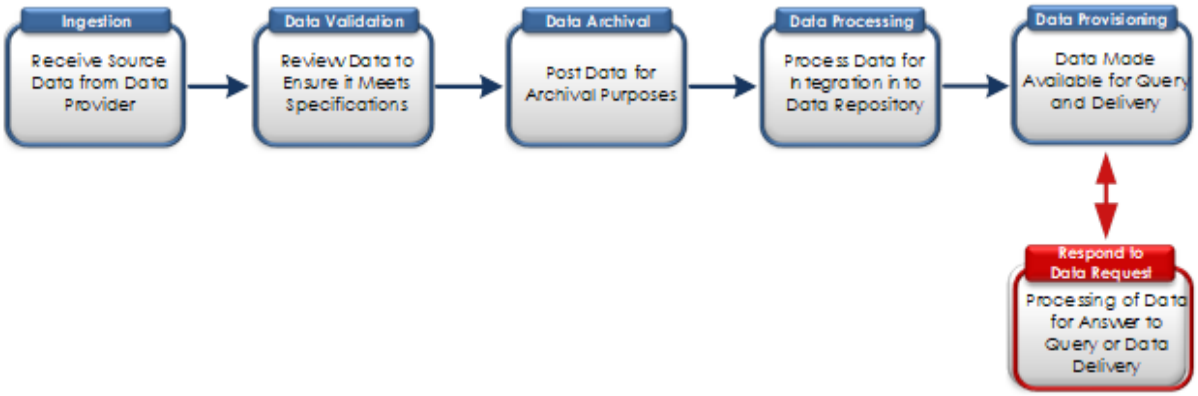

Figure M.2. Basic workflow for managing elevation data

Although the detailed steps in this process may vary in order and by agency, this workflow includes the basic data management procedures. These include:

1. **Ingestion –** Data need to be transmitted by the data producers to the entity responsible for managing these data in the context of a national program. The size of the data generally dictates the data delivery method, although online delivery is generally preferable. Data transfer through the cloud is optimal, especially if the managing entity can coordinate an efficient data transfer from the data producer directly in the cloud.
2. **Data Validation –** Data validation is required to determine if the incoming data are suitable for integration into the larger national data repository. This assessment will vary based on the program requirements but can include an independent review of data collection contractor deliverables to ensure that the data meet contracted specifications, or alternatively a less complete review to ensure that the deliverable is formatted appropriately for integration, with duties of full quality control left to another entity.
3. **Data Archive –** Once the data have been deemed suitable for integration, the source data must be archived into a formal repository. This archive should contain at least two copies in two physically distinct, geographically separated locations that allow for redundancy in situations where one copy becomes unusable. Depending on the program requirements, this archive may include the original data received from the source provider, the data modified as part of the data processing task (below), or both.
4. **Data Processing –** Some processing of data may be needed to reformat the source data for integration into the national repository or to create derivative products. The amount of processing may be influenced by the consistency of data coming from source providers. In addition, this will also be influenced by the products that are developed as part of the national program (such as Digital Elevation Models [DEM], contours, etc.), which will require processing to transform the source data into these products. Due to the potential of using a variety of technologies to produce elevation data for topography and bathymetry, the data processing component may also require the merge of disparate and temporal datasets with varying resolution, point density and accuracies.
5. **Data Visualization and Provisioning –** The source data that become part of the elevation product will need to be made available to users for query and download. In general, cloud hosting and web technology is used to allow these users to determine the data that are available and request downloads for subset areas of the data. Additional functionality may be available to the user from the web, including more robust queries (e.g., determine the elevation of a point or line), viewing the data products in a web browser map, integrating the data into other web or desktop mapping products with web services, and doing user customized processing of the data to create new data products.

Each of these tasks will drive the needs of the technology infrastructure and strategies to ensure that the capacity exists to effectively store data, process data, and serve the data to acquisition partners and the public. Most of these tasks have already been adopted by one or more of the main

custodians of national elevation data, discussed below, whose experience helped shape the needs further elaborated below.

Platforms

When looking at the needs and potential future expansion requirements for provisioning nationwide elevation data to the public, the following entities currently serving elevation data were examined.

- U.S. Geological Survey (USGS) The National Map (TNM)
- National Oceanic and Atmospheric Administration (NOAA) Digital Coast
- NOAA National Centers for Environmental Information (NCEI)
- NOAA Office of Coast Survey (OCS)

Input on the infrastructure that supports each of these platforms was provided by USGS and NOAA as noted below in the sections on each of the individual platforms.

The National Map

The National Map is a collaborative effort of USGS and other Federal, state, and local agencies to improve and deliver topographic and other information for the U.S. The purpose of the effort is to provide "…a seamless, continuously maintained set of public domain geographic base information that will serve as a foundation for integrating, sharing, and using other data easily and consistently."

The National Map is part of the USGS National Geospatial Program. The geographic information available in TNM includes orthoimagery (aerial photographs), elevation, geographic names, hydrography, governmental unit boundaries, transportation, selected structures, and land cover. The National Map is accessible via the Web, as products and services, and as downloadable data. Its uses range from recreation to scientific analysis to emergency response.

The National Map serves the geospatial community by providing high quality, integrated geospatial data and products and services including new generation digital topographic maps. The National Hydrography Dataset and Watershed Boundary Dataset comprise the hydrography components of TNM.

Elevation Data Products

Source Data Products

Source data products include lidar point clouds, source (original) resolution DEMs from which the 3D Elevation Program (3DEP) standard DEM datasets were produced, and additional data types produced from Interferometric Synthetic Aperture (IfSAR) collections.

- **Lidar point cloud:** These data are the foundational data for 3DEP in the conterminous U.S., Hawaii and territories and contain the original three-dimensional information from which the DEM products are derived. Most of the data collected in 2014 and later meet 3DEP specifications for Quality Level 2 (QL2) nominal pulse spacing and vertical accuracy, and data collected prior to 2014 often do not meet the QL2 specification. Distinctions in nominal pulse spacing are provided in the lidar point cloud status graphics

and in the download platform; however, other qualities such as vertical accuracy must be examined to determine if the data meet particular 3DEP quality level specifications.

- **Source resolution DEMs:** These are the original bare earth DEMs derived by contractors from the source lidar point cloud and breakline datasets. Source DEMs processed by the USGS after January 2015 are provided where the original DEM horizontal resolution or projection differ from the 3DEP standard DEM datasets.
- **IfSAR orthorectified radar intensity image:** These images (spatial resolutions vary), available only in Alaska, are derived from radar reflectance intensity recordings detected by the IfSAR sensor.
- **IfSAR digital surface model:** These 5-meter images, available only in Alaska, are the primary IfSAR product. Digital surface models provide elevation values of landscape features on the earth's surface. This topographic product contains the height of the highest surface on the ground including vegetation, man-made structures, and bare earth.
- **IfSAR digital terrain model:** These 5-meter images, available only in Alaska, provide elevation values of the underlying terrain of the earth's surface. This topographic product reflects the height of bare earth where the elevations of vegetation and man-made features have been removed.

USGS has seen an increased interest in lidar in Alaska and anticipates that will increase over time. Collection of elevation data in Alaska is complicated by environmental conditions and the ruggedness of the terrain. Collecting lidar in Alaska will increase the cost to mobilize and store data.

Standard DEMs

Standard DEMs represent the topographic surface of the earth and contain flattened water surfaces. Each DEM dataset is identified by its horizontal spatial resolution and is produced to a consistent set of specifications. Standard DEMs are characterized either as project-based or seamless. Project-based DEMs are available for the full areal extents of projects when produced from light detection and ranging (lidar), or as one-degree blocks with overedge when produced from IfSAR. Seamless DEMs are produced by blending only the highest quality project data into a continuous terrain surface for the U.S. These data are distributed in tiles that can be merged to support analysis across large geographic areas. Raster data are currently stored and served in Cloud-optimized GeoTIFF format.

Project-based DEMs

- **1-meter** – This dataset was introduced in 2015 with limited coverage of the U.S. but will be expanding as new DEMs from 3DEP QL2 or better lidar data are acquired. Coordinates are in the zone-appropriate Universal Transverse Mercator projection.
- **1/9 arc-second** – This dataset covers about 25 percent of the conterminous U.S. and is produced from 3-meter or finer spatial resolution bare earth DEMs acquired by the USGS prior to January 2015. Horizontal coordinates are referenced to geographic coordinates (longitude, latitude). The 1/9 arc-second dataset will no longer be updated with newly acquired DEMs; however, it will continue to be distributed.

- **5-meter** – This dataset is comprised of 5-meter IfSAR-derived DEMs (3DEP QL5) over Alaska only. Horizontal coordinates are referenced to Albers Equal Area Conical projection.

Seamless DEMs

- **1/3 arc-second** – This is the finest spatial resolution seamless DEM dataset for the U.S. with full coverage of the 48 conterminous states, Hawaii, and U.S. territories. Alaska coverage is partially available now and is being expanded to statewide coverage as part of the Alaska Mapping Initiative. Ground spacing is approximately 10 meters north/south, though pixel dimensions vary with geographic location.
- **1 arc-second** – A coarser spatial resolution seamless dataset providing complete coverage over the conterminous U.S. and partial coverage of Alaska. Most of Canada and Mexico are also covered by the 1 arc-second dataset. Ground spacing is approximately 30 meters north/south,, though pixel dimensions vary with geographic location.
- **2 arc-second** – This seamless dataset is the coarsest resolution seamless dataset available and covers only Alaska. Although ground spacing is approximately 60 meters north-south, east-west spacing can vary from 35 meters in southern Alaska to 20 meters on the North Shore.

USGS anticipates that the interest in seamless DEMs at finer spatial resolutions will increase over time. Creation of finer spatial resolution seamless DEMs, such a seamless 1-meter DEM, will increase costs to generate and store data.

Elevation Data Access

USGS provides users with multiple ways to interact with elevation data products, including online viewers or visualization tools, data explorers that allow users to find available data, and download tools.

3DEP Lidar Explorer

The LidarExplorer is a cloud-based tool that was created primarily to enable identification of lidar projects having 3D visualization enabled - giving users a mechanism to visualize the lidar through a web-based 3D viewer. Additionally, it provides users integrated access to project metadata without the need to separately download the spatial metadata. Downloads for the DEM products and Source products are provided in such a way to easily compare coverage and download sizes based on a user-defined Area of Interest (AOI).

Finally, LidarExplorer provides a mechanism for deriving products from lidar using cloud processing capabilities that take advantage of the publicly available lidar data and the Point Data Abstraction Library.

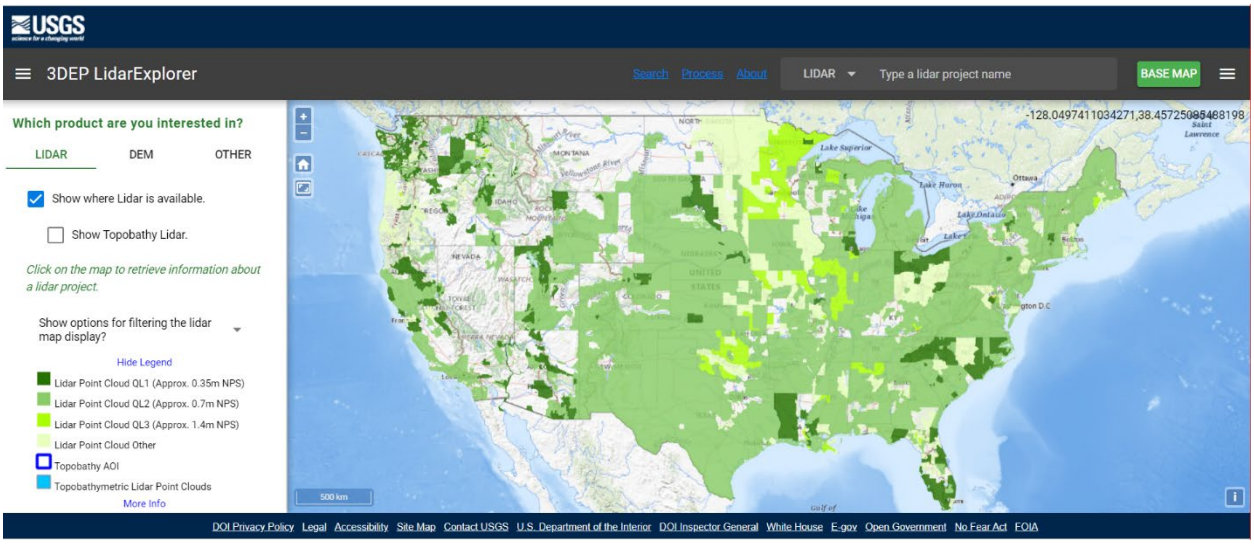

Figure M.3. USGS 3DEP Lidar Explorer user interface

The 3DEP Bare Earth DEM Dynamic Service and Demo Viewer

The 3DEP Bare Earth DEM dynamic service allows users to access multiple-resolutions of 3DEP data available in TNM. In addition, Open Geospatial Consortium Web Map Service and Web Coverage Service interfaces are enabled to support direct access to the DEM data in GIS software for display and analysis.

The DEM demo viewer creates multi-resolution visualizations DEMs on-the-fly generated dynamically from the USGS 3DEP elevation data. Users may explore a variety of derivative representations including: aspect; contours; various hillshades and slope maps.

Figure M.4. USGS 3DEP Viewer with tinted hillshade example

The National Map Downloader

The National Map Downloader is a map-based dataset locator. It displays all available TNM products including elevation. The user can choose among the available datasets and download the desired data.

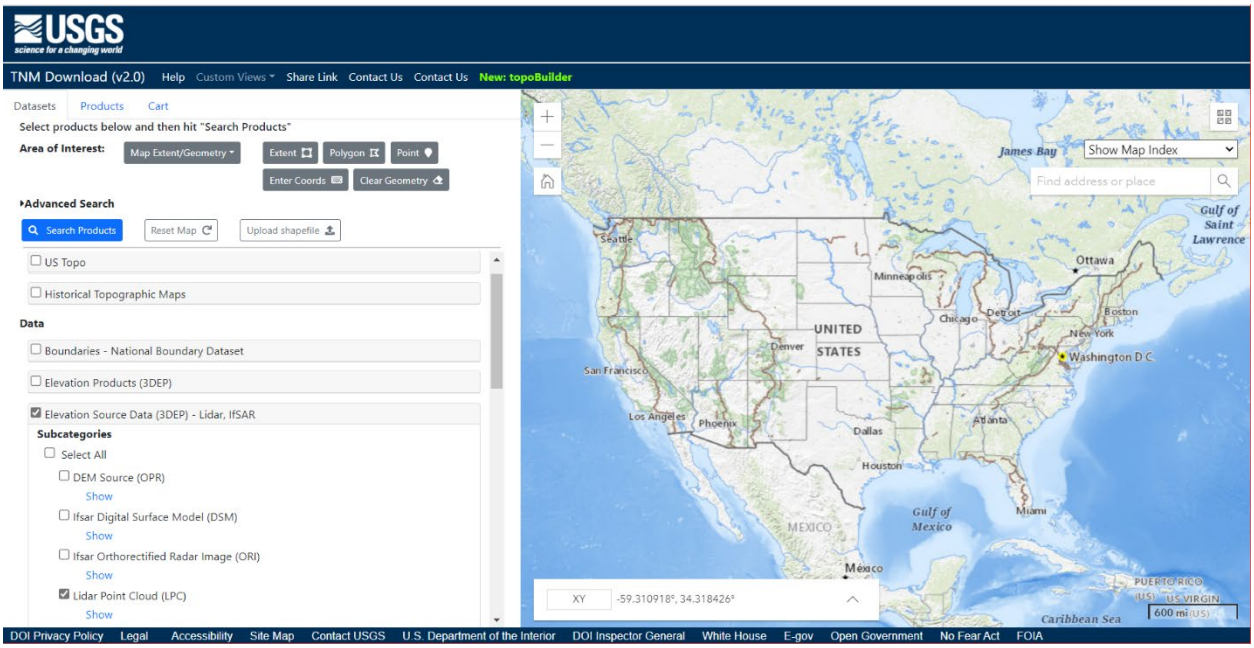

Figure M.5. USGS TNM Downloader user interface

USGS/Entwine Lidar Visualization

The USGS Entwine Lidar Visualization Tool allows users to choose a project area and visualize lidar data online using opensource Potree or Cesium visualization tools.

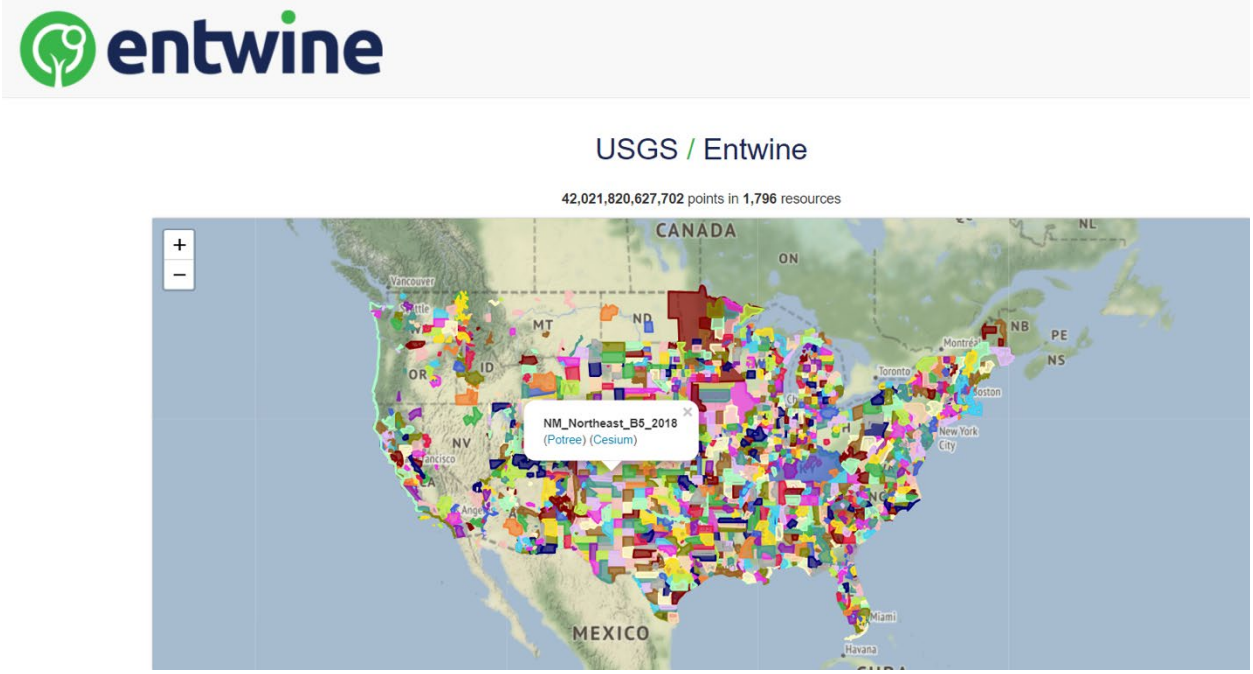

Figure M.6. USGS Entwine user interface

Staged Products Directory

The Staged Products Directory, viewable through the USGS Cloud Browse tool provides access to TNM datasets including elevation data in various file formats stored in Amazon S3 cloud. It functions in the same way as a cloud-based File Transfer Protocol (FTP) directory. Users can download data by named tiles. USGS's Rockyweb provides the same functionality as Cloud Browse.

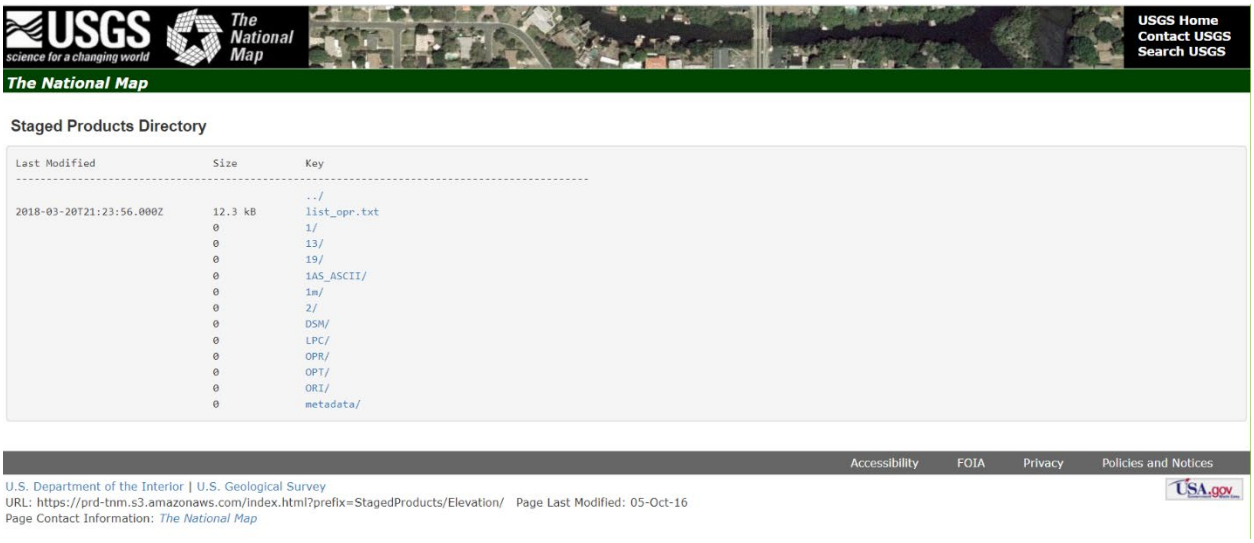

Figure M.7. USGS Staged Products Directory

Other

Elevation data can also be accessed through several major collaboration or private sector sites. Examples include:

- Open Topography
- National Geospatial-Intelligence Agency/U.S. Army Corps of Engineers (USACE) Geospatial Repository and Data Management System (GRiD)
- Microsoft Planetary Computer
- Google Earth Engine

Elevation Data Infrastructure

USGS provided the following information about the current IT infrastructure that supports delivery of 3DEP data to the public via TNM.

Data Storage Volume

The current storage volume is approximately 550 terrabytes (TB) of LAZ data, which represents approximately 3.6 million square miles of lidar point cloud data. The average storage size based on current collection methods is computed for LAS files. The approximate compression ratio is 1:5 for LAZ to LAS.

Average Storage Size per 1,000 sq. mi. LAS by Quality Level based on actual collected data:

- QL0 – 2.25 TB/1,000 sq. mi.
- QL1 – 2.25 TB/1,000 sq. mi.
- QL2 – 0.75 TB/1,000 sq. mi.
- Avg – 1.02 TB/1,000 sq. mi.

Current 3DEP processes use both LAS and LAZ depending on the specific process step. A breakdown of the square mileage of data by Quality Level derived from the 3DEP Work unit Extent Spatial Metadata as of February 2022 is shown in Table M.2.

Table M.2. Square mileage of existing 3DEP data holdings by Quality Level

Quality Level	Linear Mode Lidar (sq. mi.)	Topobathy Lidar (sq. mi.)	IfSAR (sq. mi.)	Bathy Lidar (sq. mi.)	Geiger Mode Lidar (sq. mi.)	Topobathy Model (sq. mi.)	Single Photon Lidar (sq. mi.)	Total (sq. mi.)
QL 0	136	0	0	0	0	0	0	136
QL1	187,634	1,640	0	116	27,527	0	2,785	219,702
QL2	1,685,761	341	0	0	2,547	0	0	1,688,650
QL3	193,329	0	0	0	0	0	0	193,329
QL5	0	0	746,890	0	0	0	0	746,890
Other	703,206	0	0	0	0	310	0	703,517
Null	0	0	0	0	0	0	0	1,351
Total	2,770,066	1,981	746,890	116	30,074	310	2,785	3,553,573

Forecasts for changes in the future should be worked out in relation to the potential program scenarios from the study. USGS anticipates that the interest in finer resolutions of data (e.g., QL1

or better) will increase over time. Finer spatial resolution elevation data will require more capacity to review/validate, process, store and deliver.

Data format(s)

The data are currently stored in LAZ format. The data are converted to LAS during the data validation review process. When appropriate, the Entwine Point Tile (EPT) format is used for visualization and direct cloud processing.

System Architecture

An overview of the cloud strategy for delivery and publication is provided in Figure M.8.

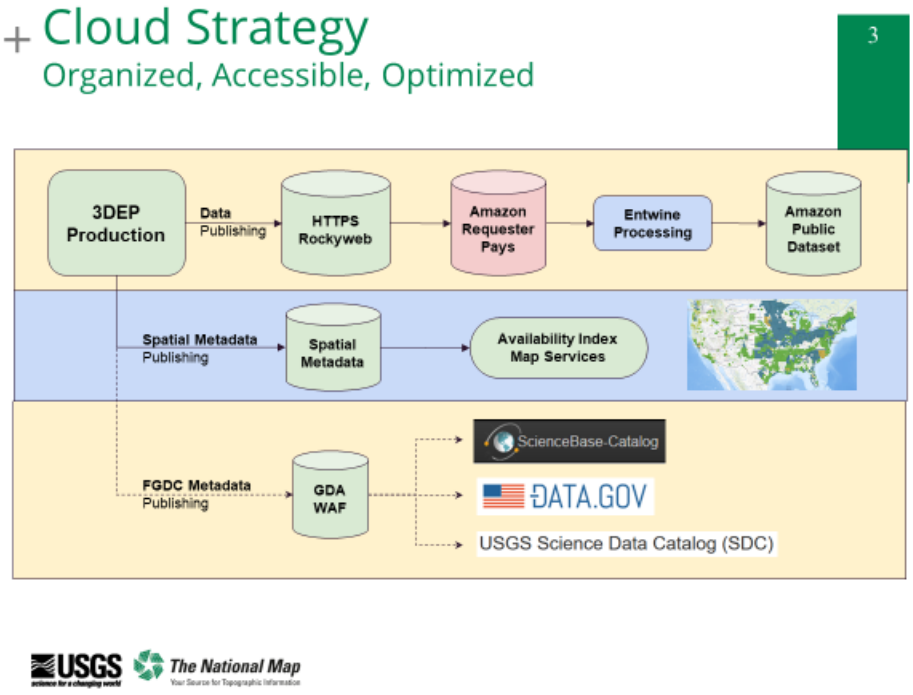

Figure M.8. USGS cloud strategy for delivery and publication of 3DEP data

Servers

The Data Validation / Review unit uses 12 virtual Microsoft Windows 2016 servers with 20 cores (2.2 gigahertz [GHz]) 128 gigabytes (GB) of Random Access Memory (RAM), and one virtual Graphic Processing Unit (GPU) each. Total of 240 cores and 1.5 TB of RAM.

The data production unit utilizes eight virtual Ubuntu 20.04 Long Term Support servers with 20 cores (2.2 GHz) and 128 GB of RAM each. Total of 160 cores and 1 TB of RAM.

Data Archive

When source data are received from acquisition contractors, the data are usually kept on local storage with local backup for up to 150 days (90-day initial review and two 30-day correction cycles if required). Once the source data have been reviewed and accepted they are handed off to USGS's internal product generation system. Once the source DEM, lidar point cloud, and metadata

production steps are complete the source data are copied to the Amazon S3 Elevation Archive bucket. The next day the data are moved to Amazon Deep Glacier for archiving. USGS archives anything that was delivered by its contractors as well as any work done in house by USGS on the project. This can vary per project but usually consists of the DEMs, LAZ, project reports (contractor and National Geospatial Technical Operations Center Data Validation created), any ancillary metadata (XML files, maximum surface height rasters, intensity, swath separation rasters, checkpoints/ground control, breaklines), error geopackages, map catalogs (DEM and LAS), etc.

Current Download Web Traffic

USGS reports that the average download web traffic generated per month on USGS sites is as follows. Data are also accessed through several major collaboration or private sector sites such as Open Topography and GRiD.

- USGS-lidar (October to January 2022): 90 TB/month (Requester Pays)
- USGS-lidar-public (April to September 2021): 200 TB/month (Public Dataset)
- Rockyweb (October to January 2022): 100 TB/month (Free Lidar Download)
- S3 (October to January 2022): 31 TB/month (Free Non-Lidar Download)

Digital Coast

The Digital Coast is managed by NOAA's Office for Coastal Management (OCM) and was first released in 2007. The Digital Coast was developed to meet the unique needs of the coastal management community. The website provides not only coastal data, but also the tools, training, and information needed to make these data truly useful. Content comes from many sources, all of which are vetted by NOAA.

Datasets range from economic data to satellite imagery. The site contains visualization tools, predictive tools, and tools that make data easier to find and use. Training courses are available online or can be brought to the user's location. Information is also organized by focus area or topic.

The elevation data managed by the Digital Coast include coastal topographic and topobathymetric lidar.

Elevation Data Products

Coastal Topographic Lidar

The coastal topographic lidar repository within the Digital Coast contains U.S. topographic data gathered by many groups using a variety of lidar sensors. Using the custom download for up to 1.5 billion points, users may search and select output by year, area, data provider, elevation product, projection, datum, and format. Output types include point clouds, digital elevation models, and contours. Requests are submitted to an automated system for custom processing, and a Universal Resource Locator link is emailed to the user upon completion. For larger requests, such as full counties, the point cloud is available in LAZ format and geographic coordinates via a bulk download link. Datasets with "DEM" in the name will have a bulk download link to the data in their original projection and tiling.

Coastal Topobathymetric Lidar

Coastal topobathymetric lidar are sourced from NOAA, USGS, and the USACE. These elevation data are collected using a green laser for bathymetry and a green laser or near-infrared lidar for topography. Depth of maximum retrievals is dependent on water clarity and the instrument used. USACE data are derived from the Joint Airborne Lidar Bathymetry Technical Center of Expertise and are coordinated through the Interagency Working Group on Ocean and Coastal Mapping (IWG-OCM).

Data collected up to and including 2013 are primarily along the ocean-facing sandy coastline and the Great Lakes. Data for bays, estuaries, and behind barrier islands are found in collections from 2014 and beyond. Custom processing through the Digital Coast Data Access Viewer (DAV) is recommended for requests under 1.5 billion points. For larger requests, such as entire datasets, the point cloud is available via https in LAZ format and geographic coordinates.

Elevation Data Access

Digital Access Viewer

The Digital Coast provides access to the data through the DAV tool. The DAV allows a user to search for and download lidar, imagery, and land cover data - hosted by the NOAA OCM - by user-defined geography. Once found, the data can be downloaded with an easy-to-use checkout interface. Users can select from multiple file formats, projections, and datums. An email provides a link to the customized data, while the original dataset is available through a link within the viewer.

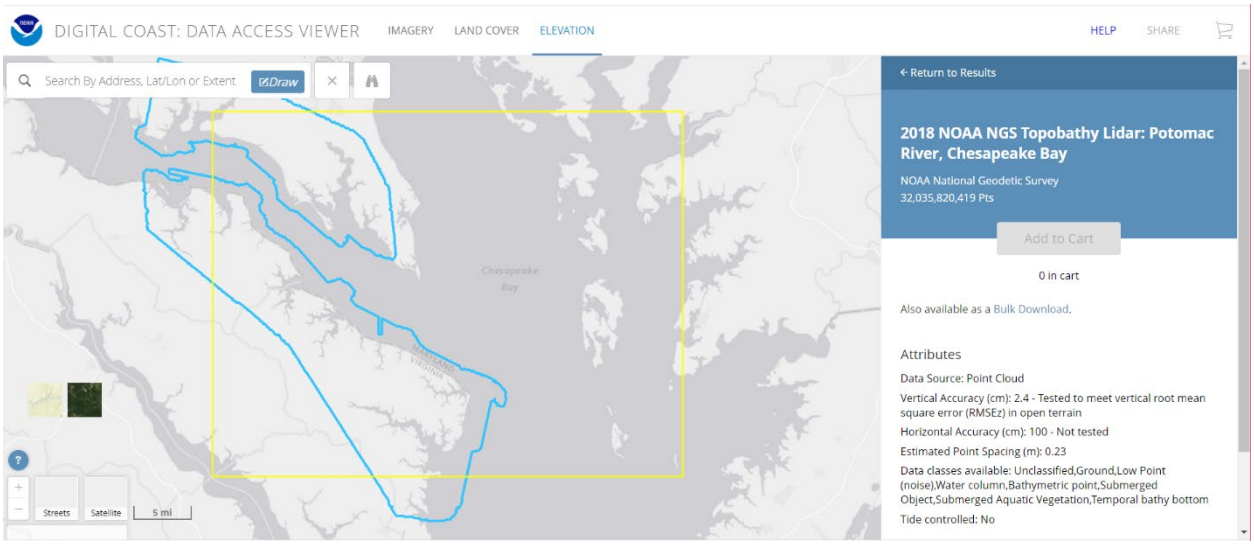

Figure M.9. Digital Coast DAV user interface

Elevation Data Infrastructure

Information in this section was provided by NOAA's National Ocean Service (NOS).

Data Storage Volume

While the Digital Coast does also serve up USGS topographic lidar data for coastal areas, this section is focused on the topobathymetric lidar data holdings of the Digital Coast.

A rough breakdown of the topobathymetric lidar data holdings in the Digital Coast is as follows.

- 60% QL1B topobathymetric lidar
- 33% QL2B topobathymetric lidar
- 7% Other

The data currently cover about 5.7 million square kilometers using web Mercator. If overlaps are dissolved and then transformed to equal area Behrmann projection, the data cover about 2.5 million square kilometers.

NOAA anticipates that the percentage of QL1B data will only increase over time. Looking at the trend since 2007, NOAA sees the volume of data doubling approximately every 2.5 years. This has been fairly consistent and it does not appear that it will drop; if anything, it will go up as the percentage of QL1B data increases.

Data Format(s)

Data are currently stored in LAZ format. The data are sorted by time to improve compression.

System Architecture

The data are stored in LAZ files. Each dataset is in its own directory and the substructure is dataset dependent (e.g., there might be subdirectories by county or block or no subdirectories). Supplemental information such as reports are kept with the datasets in a subfolder. Storage is spinning disk on-premises with a backup in the Azure cloud. A relational database is used to track each dataset. Information is stored at the dataset level (e.g., point spacing, vertical accuracy, name, etc.) as well as the file level (bounding boxes, vertical range, classes and the range for each class). Provisioning is done in multiple ways: 1) Digital Coast DAV for custom provisioning; 2) Simple static file download by dataset; and 3) Using EPTs for streaming points. The custom provisioning allows users to derive DEMs and contours from the point clouds and change projections and datums.

The static data are in an Amazon S3 bucket for cloud computing (no access or egress cost). Apps like Potree and QGIS can access the EPTs.

Data Archive

The data are archived and retained forever.

Servers

There are currently seven Linux machines with a total of 144 cores used for processing data and for the DAV custom download. All are well over five years old. Most are running with Intel Xeon X5570 chips at 2.93 GHz.

Current Download Web Traffic

Downloads from the static http/ftp site is typically over 30 TB a month, though that number is probably over a year old. Custom DAV requests are about 5,000 per month for lidar[1]. Statistics for the Entwine data are not yet available.

[1] See https://coast.noaa.gov/dataviewer_stats/

National Centers for Environmental Information

The NCEI, managed by NOAA, is the Nation's leading authority for environmental data, and manages one of the largest archives of atmospheric, coastal, geophysical, and oceanic research in the world. The NCEI contributes to NOAA's National Environmental Satellite, Data, and Information Service mission by developing new products and services that span the science disciplines and enable better data discovery. The NCEI is the long-term archive for U.S. coastal lidar data distributed publicly through the NOAA OCM.

The NCEI supports projects like the Weather Research and Forecasting Innovation Act and the NOAA Blue Economy Initiative. Its stewardship practices maximize the organization's investment in environmental research, converting scientific insights into dynamic, usable information that inform strategy and decision making in government, academia, and the private sector.

The NCEI archive and backup copy contains more than 37 petabytes of data, equivalent to about 400 million filing cabinets filled with documents. The NCEI offers users access to over 26,000 datasets and products.

The NCEI and the International Hydrographic Organization (IHO) Data Center for Digital Bathymetry (DCDB) archive and share depth data acquired by hydrographic, oceanographic, and industry vessels and platforms during surveys or while on passage. These data, which are used in several national and international mapping bathymetry projects, are free to the public with no restrictions.

Elevation Data Products

Coastal Lidar Archive

The coastal lidar archive includes data from both topographic and bathymetric lidar surveys along U.S. coasts. Data in the archive span from the mid-1990s to the present and are collected using a variety of different lidar systems. The extent of individual lidar surveys varies, ranging from shoreline strips to full county coverage. Many different organizations have contributed to the lidar data collection, including federal, state and local partners. Data are available for all U.S. coastal states and include multiple U.S. territories such as Puerto Rico, U.S. Virgin Islands, Guam, and American Samoa. The NCEI is the long-term archive for U.S. coastal lidar data distributed publicly through the NOAA OCM.

Coastal DEMs

Coastal DEMs help researchers and decision makers understand and predict environmental changes that affect coastal regions. DEM data are used in a wide range of critical monitoring activities, including coastal process modeling (tsunami inundation, storm surge, sea-level rise, contaminant dispersal, etc.), ecosystem management, habitat research, coastal and marine spatial planning, hazard mitigation, and community preparedness.

Estuarine Bathymetric Digital Elevation Models

The NOS Estuarine Bathymetric DEMs are gridded bathymetry datasets interpolated from 150 years' worth of hydrographic survey data collected by the former NOS Special Projects Office.

The initiative produced datasets for 70 estuaries in the conterminous United States with sufficient data coverage to support detailed bathymetric processing.

Great Lakes Bathymetry

The Great Lakes Bathymetry collection was compiled to rescue lake floor geological and geophysical data for public use. It contains bathymetry and detailed topographic maps for each of the Great Lakes derived from more than 100 years' worth of soundings collected by USACE (before 1970), the NOAA NOS (after 1970), the Canadian Hydrographic Service, and other organizations. These data are an important resource for researchers and local communities concerned with Great Lakes science, pollution, coastal erosion, climate changes, ecosystem threats, and fishing industry health.

U.S. Extended Continental Shelf Data

The mission of the U.S. Extended Continental Shelf (ECS) Project is to establish the full extent of the U.S. continental shelf, consistent with international law. Data collected for the project include bathymetric, subbottom, gravity, magnetic, seismic, and geologic sample data from the U.S. coastal waters to the deep ocean. United States ECS project data are in the public domain.

National Ocean Service Hydrographic Survey Data

The NCEI maintains the NOS hydrographic data and product archive, as well as the Hydrographic Survey Metadata Database (HSMD). Data and products acquired by the OCS provide coverage of U.S. coastal waters and the exclusive economic zone.

The HSMD contains International Organization for Standardization metadata records and can be used to search and display NOAA data and products. Metadata for historical foreign surveys is limited or non-existent.

Survey products include BAG files, descriptive reports, smooth sheet images, survey data images, textual gridded data, and geo-referenced side-scan sonar mosaics, Federal Geographic Data Committee/Remote System Explorer metadata, and survey statistics.

Elevation Data Access

The NCEI provides access to the data through several different online tools.

- NCEI Bathymetric Data Viewer – This interactive map allows users to search and access bathymetric data, including multi-beam, single-beam, lidar, and crowdsourced bathymetry data. See Figure M.10.
- IHO Bathymetric Data Viewer – This tool provides access to DCDB's bathymetric data holdings, as well as the spatial extent of data archived at other repositories.
- Marine Trackline Geophysical Data Search – This tool allows users to browse the NCEI archive via a text search interface for water column sonar, multi-beam bathymetry, trackline geophysical and NOS hydrographic data.
- AutoGrid – This application can be used to create grids and maps from multi-beam bathymetric data.
- Grid Extract – This interactive map gives access to combined bathymetric and topographic data, including the Earth Topography and Bathymetry Dataset Global Relief Model.

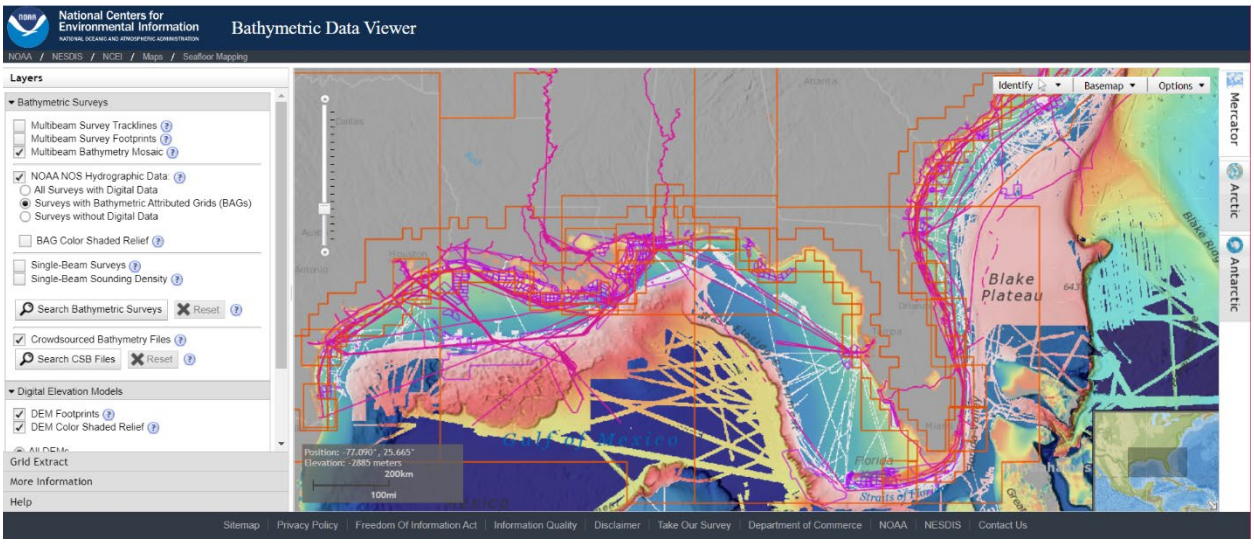

Figure M.10. NCEI Bathymetric Data Viewer user interface

Elevation Data Infrastructure

Information in this section was provided by the NOAA Bathymetry Data Manager for the NCEI.

Data Storage Volume

A raw data project with sensor data and shipboard files necessary to recreate the project is generally stored for each survey. A corresponding processed data project may also be stored for future reference and potential reuse in navigation safety projects or other scientific applications.

NOAA reports the following data are stored in the NCEI.

- **Multi-beam**: Approximately 27 TB compressed and 40 TB uncompressed volume within the U.S. Exclusive Economic Zone (EEZ). Data in the multi-beam archive can include raw, processed, products, and ancillary data; however, the vast majority is raw data.
- **Single-beam**: Approximately 20 GB uncompressed and 7 GB compressed (not including ancillary data). All survey data are processed (Level_01).
- **Crowdsourced Bathymetry**: ~25GB
- **NOS Hydrography**: Total volume on archive is 1.18 PB; ~25% raw data, ~75% processed, supporting data and products

These data cover the following extents.

- **Multi-beam**: 2,235,000 sq. mi surface area within U.S. EEZ, 2,194 surveys
- **Single-beam**: 8.88 million miles, 42.4 million data points, 3,545 surveys
- **Crowdsourced Bathymetry**: Not sure how to calculate this at this time
- **NOS Hydrography**: Over 625,000 linear nautical miles

NOAA anticipates the following growth in storage volume over time.

- **Multi-beam**: Multi-beam data are expected to grow by 15% Year Over Year growth. NOAA is planning for 15% growth; however, this may increase due to the expected implementation of the National Strategy for Ocean Mapping, Exploring, and Characterizing the United States Exclusive Economic Zone and mapping campaigns like Seabed 2030.
- **Single-beam**: Single-beam data are expected to grow by 15% Year Over Year growth.
- **Crowdsourced Bathymetry**: Future growth is unknown at this point.
- **NOS Hydrography**: 15% Year Over Year growth is expected for the immediate future.

Data Format(s)

The following data formats are supported.

- **Multi-beam:** Data are primarily stored in their native sonar format such as .all or .s7k for raw, and .gsf for processed data.
- **Single-beam:** Data are primarily stored in MGD77T format[2], a non-proprietary ASCII format.
- **Crowdsourced Bathymetry:** Data are primarily stored in non-proprietary ASCII format.
- **NOS Hydrography:** Data are primarily stored in BAG and TIFF formats.

[2] Information about the Marine Geophysical Data Exchange format (MGD77T) can be found at: https://www.ngdc.noaa.gov/mgg/dat/geodas/docs/mgd77.pdf

System Architecture

- **Multi-beam:** Figure M.11 illustrates the NCEI multi-beam system architecture (MABLE)[3].

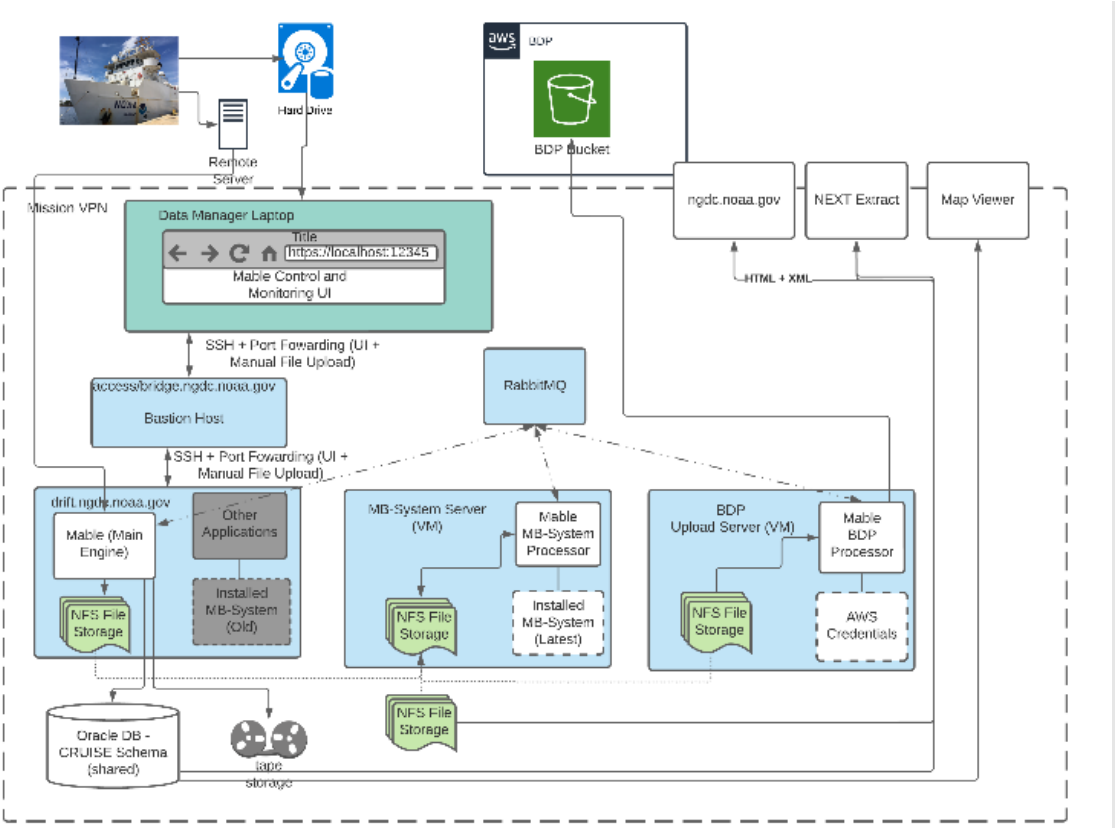

Figure M.11. NCEI multi-beam system architecture (MABLE)

- **Single-beam**: The single-beam system architecture (TrackPipe) uses Oracle metadata, ancillary data catalog, and data point store databases. StoreNext Tape Archive. Amazon Web Services QA/QC and assimilation software. Spinning Disk ancillary data storage for delivery.

[3]Image source: https://lucid.app/publicSegments/view/4b9b21ce-23d8-429a-b611-22451ae345d0/image.png

- **Crowdsourced Bathymetry**: Figure M.12 illustrates the NCEI crowd-sourced bathymetry system architecture (CROWBAR)[4]:

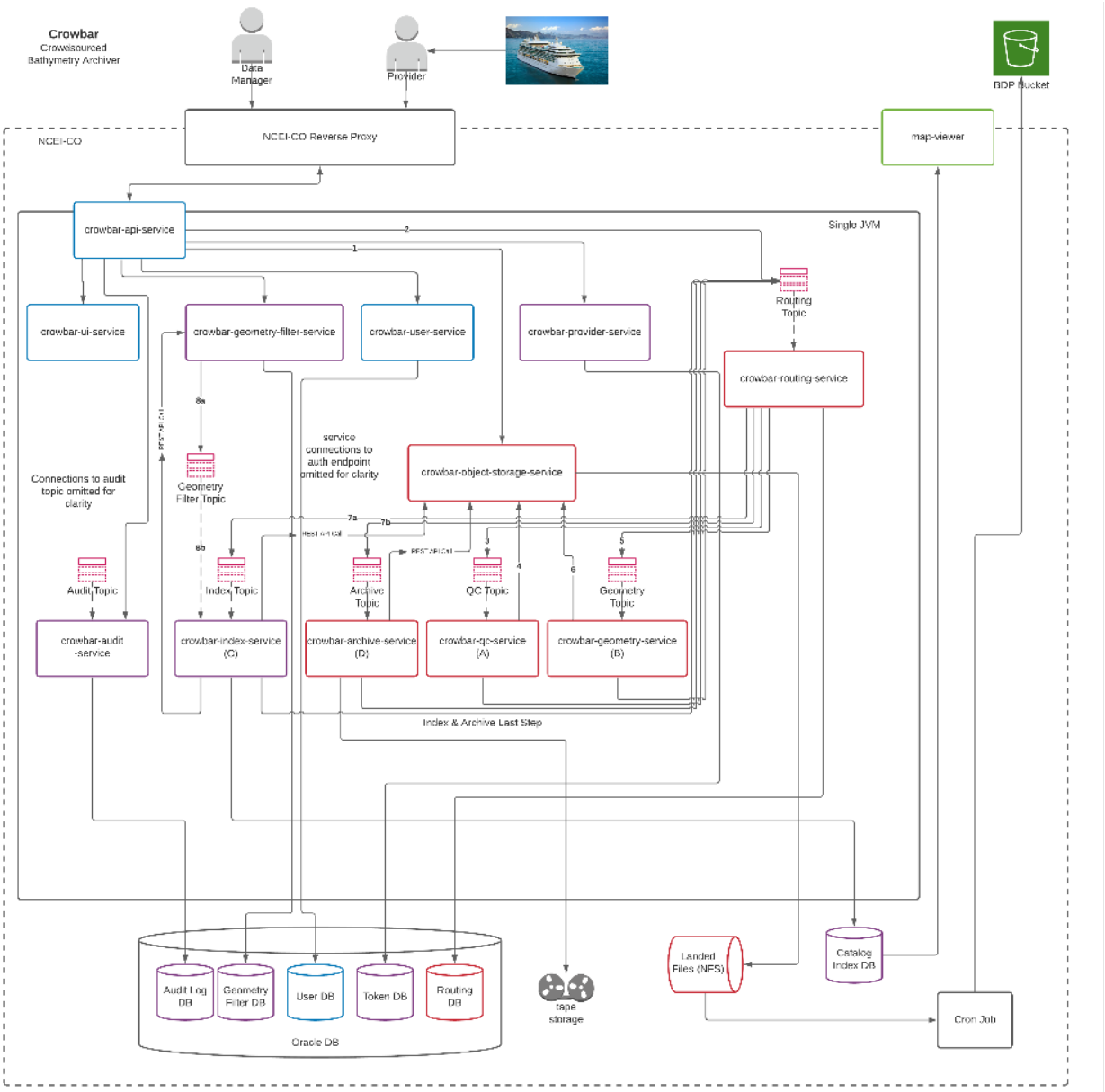

Figure M.12. NCEI crowd-sourced bathymetry system architecture (CROWBAR)

[4] Image source: https://lucid.app/publicSegments/view/f806e00d-c3b4-493b-81bf-8d63f92d2641/image.png

The following applications are used to access and distribute the data.

- **Multi-beam**: NCEI Data Extract System (NEXT) (NCEI-CO data delivery tool), Bathymetric Data Viewer
- **Single-beam**: NEXT (NCEI-CO data delivery tool), Marine Trackline Geophysical Data Search
- **NOS Hydrography**: Bathymetric Data Viewer

Data Archive

All geophysical data including bathymetry are archived for 75 years (this applies to all datasets).

Current Download Web Traffic

The NCEI does not have formal download metrics but looking at the NEXT daily query logs the following rough estimates for each of the datasets are provided. Numbers of queries and data volumes delivered can vary by orders of magnitude per month.

- **Multi-beam**: ~30 queries/day or ~900 queries/downloads per month
- **Single-beam**: Approximately ~12 queries/day or 400 queries/month (from a few megabytes to multiple gigabytes).
- **NOS Hydrography**: ~20 queries/day or 600 queries/month. There is not a reliable system to provide these metrics.

Office of Coast Survey

The Office of Coast Survey maintains the nation's nautical charts and publications for U.S. coasts and the Great Lakes. The Office of Coast Survey creates and maintains over a thousand charts covering 95,000 miles of shoreline and 3.4 million square nautical miles of water. It supports the 1.3 billion metric tons of cargo valued at $1.8 trillion that comes in and out of U.S. ports every year.

Elevation Data Products

The Office of Coast Survey conducts surveys of the ocean floor, creating detailed hydrographic maps of depths and features. These data are used to create raster and electronic nautical charts. The NCEI maintains the digital data archive for all hydrographic data of the coastal waters and exclusive economic zone of the U.S. and its territories collected by the OCS. The database provides hydrographic survey products which contain additional details of the ocean floor not shown on the nautical charts.

The Office of Coast Survey generates products from a wide range of data and is responsible for managing data across the various product pipelines. The Office of Coast Survey has built a data infrastructure that includes robust storage and dissemination architectures both on-premise and in the cloud. The Office of Coast Survey is committed to sharing data openly while ensuring that data are governed by a strict set of standards and requirements. Information on Hydrographic Survey and Charting Standards can be found at IHO.

Figure M.13 shows a breakdown of the source data workflow as it moves through the various offices in OCS. All source data go through Quality Assurance and Quality Control (QA/QC) at various offices as the data makes its way to multiple publicly accessible products.

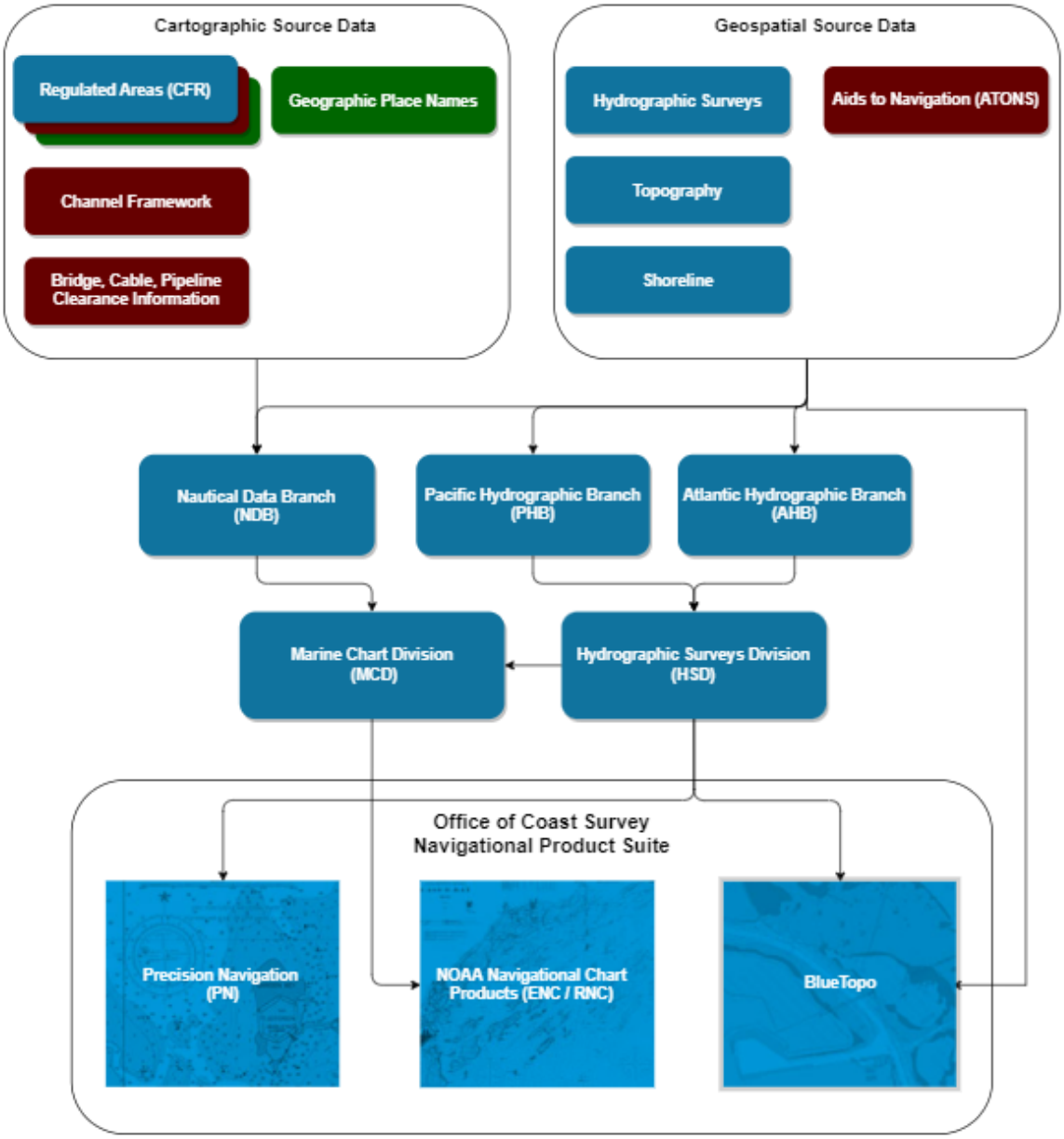

Figure M.13. Office of Coast Survey source data workflow

Elevation Data Access

BlueTopo

BlueTopo is a compilation of the nation's best available bathymetric data. In the same way that topographic map details the height of land, BlueTopo details the depth of lake beds and seafloor beneath navigationally significant U.S. waters. Created as part of the OCS nautical charting

mission and its National Bathymetric Source project, BlueTopo is curated bathymetric source data to provide a definitive nationwide model of the seafloor and the Great Lakes.

Historically, depth information on nautical charts has been drawn from many different, often sparse, bathymetric sources that can be difficult to access and update. The National Bathymetric Source project is part of a new OCS workflow to modernize the charting process and to provide a critical bathymetry resource to several customers, including the public. For example, NBS supports Precision Navigation by providing seamless, high-resolution bathymetric data to better equip mariners in making critical navigation decisions. BlueTopo is the public "not for navigation" compilation of bathymetry produced by the OCS.

Elevation Data Infrastructure

Information in this section was provided by NOAA's OCS IT Service Branch. The OCS IT Service Branch is responsible for the infrastructure which supports the mission of OCS. All OCS data are uploaded to the NCEI for archive and public release. Statistics for OCS data in the NCEI are included in the category listed as NOS Hydrography data.

Data Storage Volume

General storage of data is on three production HyperV servers. The following data storage figures must be doubled to account for backups:

- Marine Chart Division stores about 24 TB of data
- Hydrographic Surveys Division stores about 17 TB of data
- Navigation Services Division stores about 2 TB of data.

File Format(s)

Data maintained by OCS data are primarily BAGs and TIFFs.

System Architecture

Information was not able to be provided.

Data Archive

All data are archived via the NCEI.

Servers

General storage of data is currently on three production HyperV servers. OCS recently purchased four Dell PowerEdge R740xd2 servers with 240 TB of storage to accommodate on-premises backups. OCS data are uploaded to NCEI for archive and public release.

Current Download Web Traffic

All data downloads are done via the NCEI. The NCEI reports ~20 queries/day or 600 queries/month for OCS data that are included in the NOS Hydrography data category, although there is not a reliable system for reporting these metrics.

Conclusion

Based on the information provided by USGS and NOAA, it is evident that there is significant cost and management required for data storage and dissemination of high-resolution elevation data. Below is a bulletized list of concluding remarks.

- Acquisition costs do not include program and data lifecycle management cost elements.
- USGS and NOAA each already have robust program management and information technology support systems in place. However, these will likely need to be expanded to support the additional data generated by an expanded national elevation program.
- Data volumes are only going to increase as data density (i.e., Quality Level) and geographic area covered increase. Therefore, system architecture and storage systems will likely need to be expanded to support the increased volume of data. Costs will thus also increase.
- User appetite for data will also increase as more data become available and more users appreciate its value. Therefore, data provisioning systems may need to be expanded or modernized further.
- Program management, acquisition planning and coordination, and contract management requirements will increase as an expanded national elevation program is implemented.
- A general rule of thumb is that QA/QC costs run about 10-15% of data acquisition costs. Lifecycle management costs (including QA/QC costs) are likely to be on the order of 20-25% of acquisition costs and are likely to increase somewhat as data volumes increase over the life of a program. (The 2012 USGS National Enhanced Elevation Assessment report estimated 6-10% for IT infrastructure costs, not including QA/QC of contracted data acquisitions; the recent USGS 3D National Topography Model Call for Action: 3D Hydrography Program draft report estimated 8.8% for lifecycle management costs, again not including QA/QC of contracted data acquisitions.)
- Continued and frequent coordination between USGS, USACE, NOAA, and other federal partners regarding elevation data needs and acquisitions will remain important to reduce the costs for mobilization and demobilization, minimize duplication of efforts, and identify funding partnerships.
- It is important for the federal government to continue research and development efforts of their own as well as working with the private sector to identify and test new technologies that can reduce acquisition costs.