OKEANOS EXPLORER

EX1304: NORTHEAST U.S. CANYONS EXPLORATION
DATA MANAGEMENT PLAN

Document Purpose
This document is an addendum to the overarching Okeanos Explorer FY13 Data Management Plan (EX_FY13_DMP.pdf) and is specific to the EX-13-04 mission entitled "Northeast U.S. Canyons Exploration" For more detailed information on the data management effort for the Okeanos Explorer in FY13, please refer to that document.

General Description of the Data to be Managed
EX1304 will employ the new 6,000 meter ROV and the camera sled Seirios to explore some targeted areas in the Northeast U.S. Canyons over two legs: July 8-25 and July 31-August 17. 24-hour operations, including multibeam mapping, ROV operations, and CTD/rosette deployment will be conducted. All data will be fully processed according to OER standard onboard procedures and will be archived with the NOAA National Data Centers. Data management procedures are fully documented in the data management plan for the *Okeanos Explorer* for the FY13 field season (EX_FY13_DMP.pdf)

- Name of Dataset
  - : "EX1304: Northeast U.S. Canyons Exploration"
- Mission Specific Keywords:
  - Place Specific:
    - New England Seamount Chain
    - Western North Atlantic Ocean
    - Mytilus Seamount
    - Block Canyon
    - Alvin Canyon
    - Atlantis Canyon
    - Hydrographer Canyon
    - Veatch Canyon
    - Oceanographer Canyon
    - Lydonia Canyon
    - Powell Canyon
    - Munson Canyon
    - Nygren Canyon
    - Heezen Canyon
    - Chebacco Canyon

  - Theme Specific:
    - Multibeam
    - Multibeam sonar
    - Multi-beam sonar
    - Sub-bottom profile
    - Mapping survey
    - Multibeam backscatter

1

- Water column backscatter
- Singlebeam sonar
- Singe beam sonar
- Single-beam sonar
- Sub-bottom profile mapping
- ACUMEN
- Atlantic Canyons Undersea Mapping Expedition
- Extended Continental Shelf
- ECS
- Continental shelf mapping
- USGS Marine Hazards
- Gas seeps
- Landslide features
- Deep sea corals
- Underwater cultural heritage
- Deep Discoverer
- Seirios

- Summary Description:

EX-13-04 efforts complement and continue the 2012 Atlantic Canyons Undersea Mapping Expeditions (ACUMEN). The 2012 series of five ACUMEN cruises (NOAA Ships *Okeanos Explorer, Ferdinand Hassler and Henry B. Bigelow*), and two *Okeanos Explorer* cruises since then gathered baseline information on deep water canyons off the northeastern U.S. seaboard, mapping along the continental shelf and slope from Virginia to the northeastern boundary of the U.S. Exclusive Economic Zone (EEZ). These mapping operations provide the basis for preliminary target selection, and enable EX-13-04 Legs I and II to commence the next steps in systematic ocean exploration, investigating deep water areas in and along the northeast canyons off the U.S. East Coast.

- Temporal Bounds:
    - Leg I: July 8 – July 25, 2013
    - Leg II: July 31 – August 17, 2013
- Spatial Bounds:
    - Northern: 41.5
    - Southern: 39
    - Western: -72
    - Eastern: -66
- Data Type Collections for Preservation/Stewardship:
    - Underwater Video – capture of full-resolution underwater video in several resolutions from the 6,000 meter ROV and the Seirios camera sled
    - Multibeam Bathymetry – continuous collection during the duration of the expedition.
    - Bottom Backscatter – continuous collection during the entire duration of the expedition
    - Water Column Backscatter – continuous collection during the entire duration of the expedition
    - Scientific Computing System (SCS) output – continuous collection of navigational, meteorological, integrated oceanographic sensor data
    - CTD / Rosette – casts will be conducted as needed to guide science operations

2

- o XBT – casts will be conducted at an interval defined by prevailing oceanographic conditions, but not to exceed 6 hours. Casts will collect water temperature at depth for sound velocity calculations to maintain multibeam data quality
  - o Knudsen CHIRP 3260 –sub-bottom profiler data collected between 1000 and 1800 each day
  - o EK60 – single beam sonar for water column features during the entire duration of the expedition
- Data Product/Product Collections for Preservation/Stewardship:
  - o Gridded bathymetry (.txt)
  - o Gridded bathymetric image (.tif)
  - o Fledermaus gridded bathymetry imagery (.sd)
  - o Fledermaus gridded backscatter imagery (.sd)
  - o Fledermaus multi-layer image file (.scene)
  - o Google Earth gridded bathymetry (.kml)
  - o ArcView gridded bathymetry (.asc)
  - o SCS data output in NetCDF
  - o Highlight Videos
  - o Final Mapping Summary document
  - o Final Cruise Summary document
- Volume of Data Expected
  - o The volume of data expected from this cruise is approximately 6 TB per leg including full-resolution video. Without the full-resolution video, the total volume of data expected is 650 GB.
- Personally Identifiable Information (PII) concerns
  - o No PII will be included in these data.

Points of Contact
- Overall Point of Contact (POC) for the data:
  - o Expedition Coordinator: Kelley Elliott, Kelley.Elliott@noaa.gov
  - o Data Acquisition: EX Mapping Team: oar.oer.exmappingteam@noaa.gov
  - o Data Management: OER Data Management Team (oer.info.mgmt@noaa.gov)
- Responsible for Data Quality:
  - o Seafloor mapping and water column data:
    EX Mapping Team: oar.oer.exmappingteam@noaa.gov
  - o SCS data: Office of Marine and Aviation Operations (OMAO): Lt. Laura Gallant, Okeanos Explorer Operations Officer (Ops.Explorer@noaa.gov)
- Responsible for data documentation and metadata activities:
  - o National Coastal Data Development Center (NCDDC); OER Data Management Team (oer.info.mgmt@noaa.gov)
- Responsible for the data storage and data disaster recovery activities:
  - o NOAA National Data Centers; National Oceanographic Data Center (NODC), National Geophysical Data Center (NGDC), NOAA Central Library (NCL)
- Responsible for ensuring adherence to this data management plan, including resources are made available to implement the DMP:
  - o Data Acquisition: Kelley Elliott, Expedition Coordinator
  - o Data Acquisition: EX Mapping Team, oar.oer.exmappingteam@noaa.gov
  - o Data Acquisition: Lt. Laura Gallant, OMAO, Okeanos Explorer Operations Officer
  - o Data Management: OER Data Management Team

3

Data Stewardship
- What quality control procedures will be employed?
  - Quality control procedures for the data from the Kongsberg EM302 is handled at UNH CCOM/JHC. Raw (level-0) bathymetry files are cleaned/edited into new data files (level-1) and converted to a variety of products (level-2).
  - Data from sensors monitored through the SCS are archived in their native format and are not quality controlled.
  - Data from XBT firings are archived in their native format and are not quality controlled.
- What is the overall lifecycle of the data from collection or acquisition to making it available to customer?
  - All ship data from this mission is expected to be archived and accessible within 60-90 days post-mission.
  - METOC data from the SCS are converted in a post-mission model into archive-ready compressed NetCDF3 format and stored within the NCDDC THREDDS open-access server.
  - CTD data from casts are processed in a post-mission model and converted into archive-ready compressed NetCDF3 format and stored within the NCDDC THREDDS open-access server.

Data Documentation
- An ISO format metadata record to document the mission will be generated during pre-cruise planning and published in an OER catalog for public discovery and access. Documentation templates will be provided for post-mission products with references back to the overall mission metadata documents. Data collections and products will be documented with ISO or FGDC CSDGM metadata and published at the appropriate NOAA Data Center.
- ISO 19115-2 Geographic Information with Extensions for Imagery and Gridded Data will be the metadata standard employed.

Data Sharing
- All data recorded, observed, generated or otherwise produced on the *Okeanos Explorer* are considered non-proprietary and will be made available to the public as soon as possible after a period of due diligence in performing quality assurance and data documentation procedures.

Initial Data Storage and Protection
- Data are recorded and stored on NOAA shipboard systems compliant with NOAA IT procedures. Data are moved from ship to shore using a variety of standard, documented data custody transfer procedures. Data are transferred to NOAA data centers using digital and physical data transfer models depending upon data volume.

Long-Term Archiving and Preservation
- Data from this mission will be preserved and stewarded through the NOAA National Data Centers. Refer to the *Okeanos Explorer* FY13 Data Management Plan (EX_FY13_DMP.pdf) for detailed descriptions of the processes, procedures, and partners involved in this collaborative process. Appendix A has an excerpt from EX_FY13_DMP.pdf) that illustrates the data and product pipelines that will be employed for this mission.

Data Management Objectives
The DMT's common objectives for this mission are:
- Ensure the near real-time update of the *Okeanos Atlas* with
    - Ship track and hourly observations received via email.
    - Daily logs pulled from URI through RSS feeds and links to related images on oceanexplorer.noaa.gov website.
    - Daily cumulative bathymetric image overlays received via URI SRS.
- Execute multibeam and oceanographic data pipelines according to the FY13 DMP (EX_FY13_DMP.pdf).
- Develop ISO metadata for collection-level and dataset-level records collected from the ship (multibeam, singlebeam sonar, sub-bottom profiler, XBT, CTD, EX METOC,)

Specific mission objectives include:
- Ensure that SOP and standard naming conventions are understood and being followed by ship personnel.
- Review and correct the list of camera names/codes/locations in the standard naming conventions documentation
- Ensure that the full-resolution video capture procedures are fully functioning and are capturing the video at the best possible resolution
- Develop a consistent, automated method for ensuring that video products are being generated
- Determine the status of the SCS data capture of the CTD sensor data on the submersibles and resolve any issues
- If additional sensors are added to the CTD, develop SOP to properly capture data in SCS

Expedition Principals for Data Management
Webb Pinner, OER Telepresence, EX Data and Information Lead, Webb.Pinner@noaa.gov
Brendan Reser, NCDDC, Okeanos Explorer Program Data Manager, Brendan.Reser@noaa.gov
Sharon Mesick, NCDDC, Federal Program Manager, Data Management IPT Chair,
    Sharon.Mesick@noaa.gov
Susan Gottfried, NCDDC, OER Data Management Coordinator, Susan.Gottfried@noaa.gov
Andy Navard, NCDDC, Okeanos Atlas Developer, Andrew.Navard@noaa.gov
Evan Robertson, NGDC, Geophysical Data Officer, Evan.Robertson@noaa.gov
Tom Ryan, NODC, Oceanographic Data Officer, Thomas.Ryan@noaa.gov
Anna Fiolek, NCL, Multimedia Librarian, Anna.Fiolek@noaa.gov

5

Appendix A: Data and Product Pipelines (excerpt from EX_FY13_DMP.pdf)

A.   *Oceanographic/Meteorological/Navigational Data Archive Pipeline*

Data from hull-mounted and off-board oceanographic and meteorological (METOC) sensors; integrated oceanographic sensors from the submersibles; and navigational instrumentation on both the vessel and its submersibles are monitored through the ship's Scientific Computer System (SCS). Some of these data will be used in a near real-time mode to update the *Okeanos Atlas*. All of these data will be archived at the National Oceanographic Data Center (NODC) Marine Data Stewardship Division (MDSD) in Silver Spring, MD. A cruise-level and several collection level metadata records describing the data inventory to be archived at the NODC/MDSD will be included with the data submission.

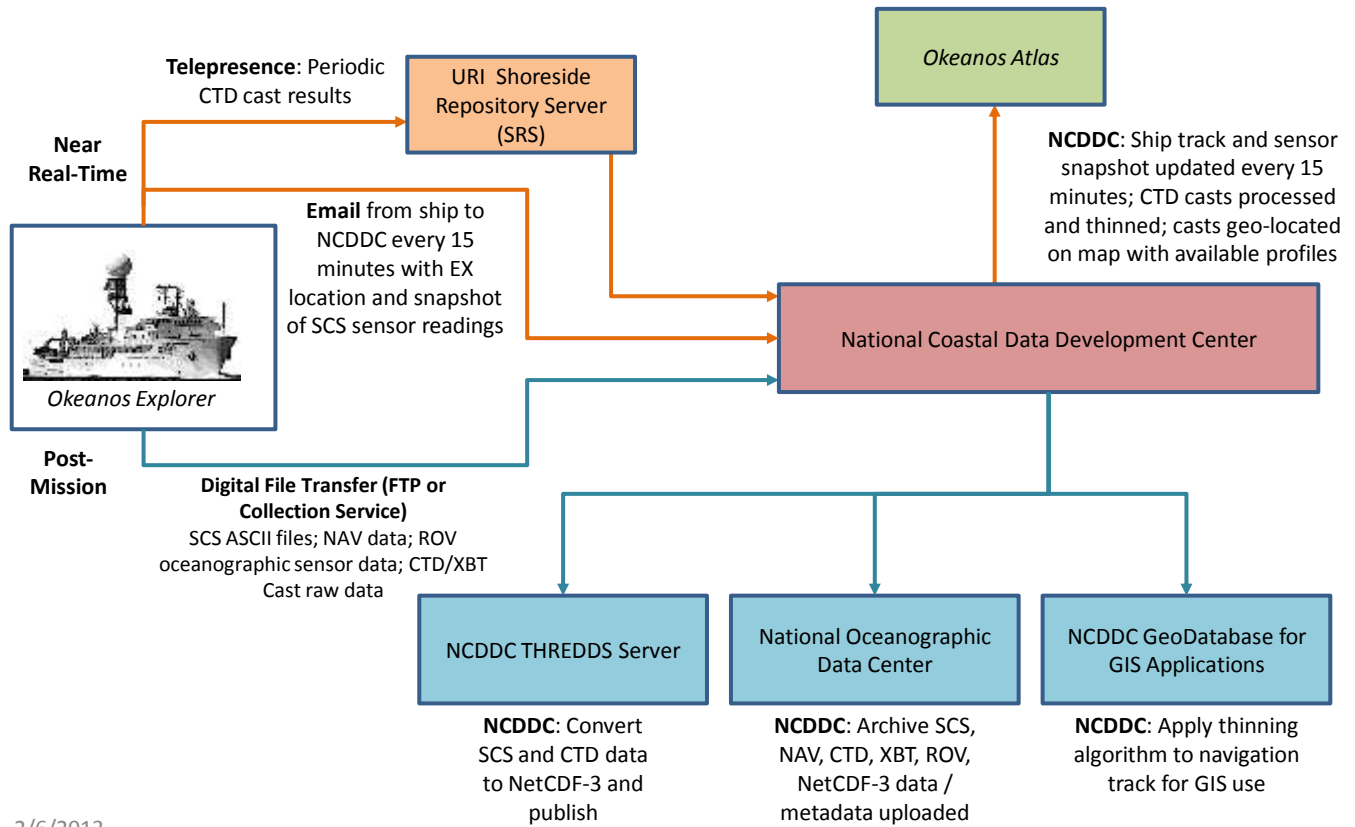## Oceanographic/Meteorological/Navigational Data/Products Pipeline

Fig 4: Okeanos Explorer Oceanographic Data Pipeline

Near Real-Time:

At periodic (currently fifteen minutes) intervals, an email from the ship to NCDDC is delivered with the ship's position and a snapshot of the SCS sensor suite.

As CTD casts are deployed, the results of the cast are included in the periodic synchronizations to the SRS.

The GIS team at NCDDC processes:
- CTD cast data into thinned profiles for comparison to World Ocean Atlas historical profiles in the same region and month. The thinned profiles are geo-located on the *Okeanos Atlas*. The corresponding temperature profile plot from the World Ocean Atlas is added for comparison.
- Ship track and sensor snapshot readings are geo-located on the *Okeanos Atlas*.

Post-Mission

All SCS data, including navigation and CTD/XBT cast data are delivered to NCDDC either via ftp or through a Collection Service.

SCS navigation data are used to apply a thinning algorithm and return an optimized thinned navigation track, which is added to the GeoDatabase for GIS applications.

Using the SCS configuration file, a header line is appended to each SCS ASCII data file.

All of the SCS data files are used to generate an archive-ready compressed NetCDF-3 formatted file.

The CTD Cast raw data are used to generate a second NetCDF-3 formatted file.

ncISO metadata records are generated for the NetCDF-3 files, and FGDC CSDGM metadata records are generated for the SCS ASCII files, the NAV data set, and the CTD and XBT data sets.

All data sets and the corresponding metadata are uploaded to the National Oceanographic Data Center (NODC), where they will be accessioned and archived.

The NetCDF3 file will be ingested into an NCDDC hosted Thematic Real-time Environmental Distributed Data Services (THREDDS) server for user discoverability and access.

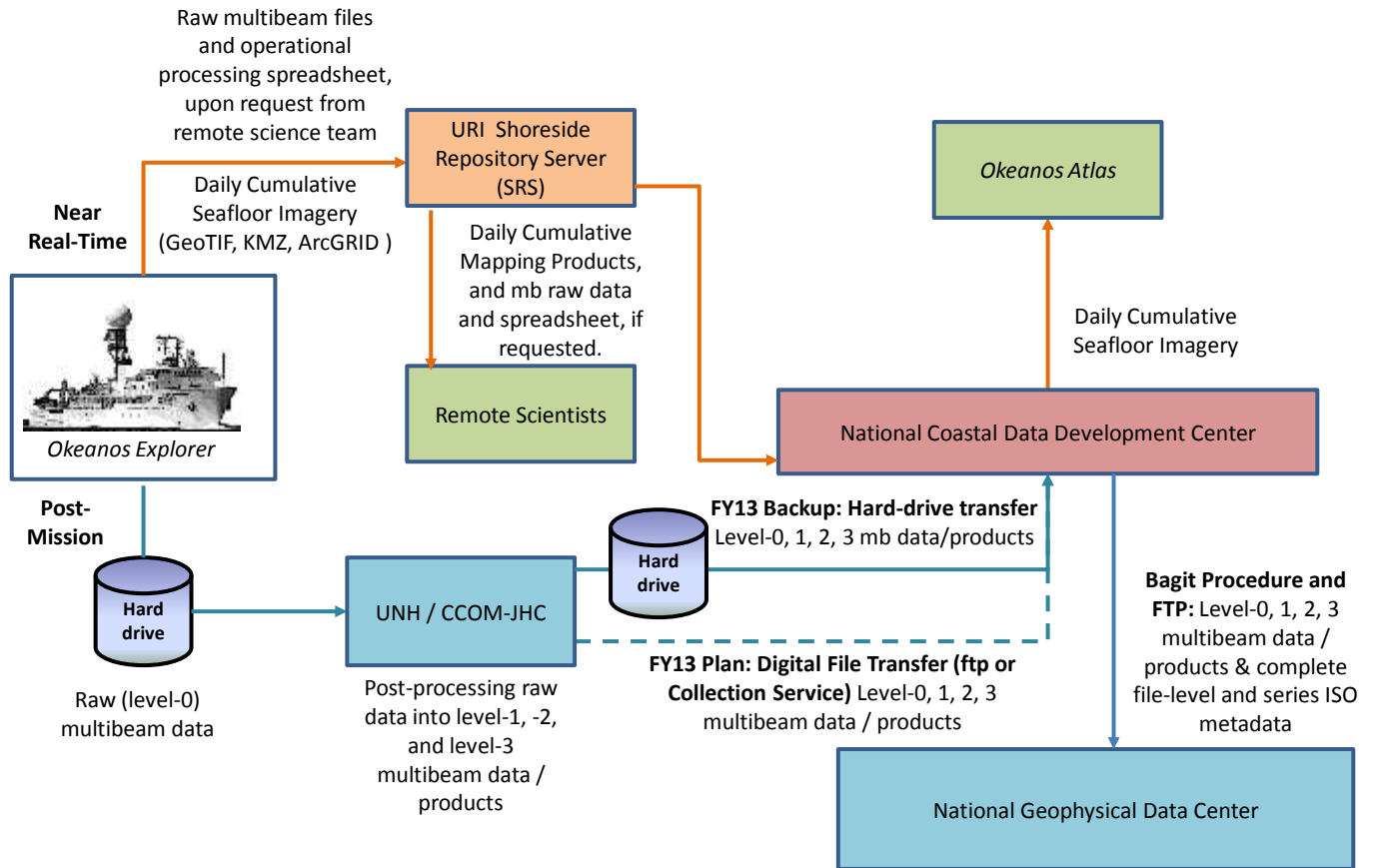| Data Class	Instrument	Data Type	Format	Metadata Granularity	Archive Center
OCN/ MET	All SCS monitored sensors	Meteorological and Oceanographic data sensors	ASCII	1 meta rec	NODC/MDSD
NAV	DGPS, CNAV	EX, ROV, and sled navigation	ASCII	1 meta rec	NODC/MDSD
ALL	All	Archive Ready	NetCDF-3	1 meta rec	NODC/MDSD

Table 3: Oceanographic/Meteorological/Navigational Metadata Granularity and Target Archive

7

B.   *Multibeam Survey Data Archive Pipeline*

The multibeam survey data collected by bottom-looking and complementary sensors, data from the calibration instruments, and the products generated after the data is returned to and post-processed at UNH will be archived at the NGDC. These data will be accompanied with a collection level metadata record for the NGDC as well as individual metadata records for each raw (level-0) file, each edited (level-1) file and each data product (level-2) and report (level-3) generated as a result. In addition, the submission to NGDC will include the following:
- raw (level-0) mapping survey and water column data files,
- CTD and/or XBT profile data used for calibration in multibeam survey,
- post-processed, quality assured, and edited (level-1) data files,
- specific data products (level-2) from the Fledermaus software, including cumulative GeoTIF images, gridded bathymetric files, KML files, KMZ images, .sd output files, and an ArcGrid format, and
- comprehensive mapping survey data summary (level-3) report.

Multibeam Data/Products Pipeline

Raw multibeam files and operational processing spreadsheet, upon request from remote science team

Near Real-Time

Daily Cumulative Seafloor Imagery (GeoTIF, KMZ, ArcGRID)

URI Shoreside Repository Server (SRS)

Daily Cumulative Mapping Products, and mb raw data and spreadsheet, if requested.

Okeanos Explorer

Okeanos Atlas

Remote Scientists

Daily Cumulative Seafloor Imagery

National Coastal Data Development Center

Post-Mission

FY13 Backup: Hard-drive transfer Level-0, 1, 2, 3 mb data/products

Hard drive

Bagit Procedure and FTP: Level-0, 1, 2, 3 multibeam data / products & complete file-level and series ISO metadata

Hard drive

Raw (level-0) multibeam data

UNH / CCOM-JHC

Post-processing raw data into level-1, -2, and level-3 multibeam data / products

FY13 Plan: Digital File Transfer (ftp or Collection Service) Level-0, 1, 2, 3 multibeam data / products

National Geophysical Data Center

2/6/2013

Fig. 5: Okeanos Explorer Multibeam Data Pipeline

Near Real-Time

If the remote science team has requested that some raw multibeam data be transferred in near real-time to the SRS, the raw data and a current copy of the processing spreadsheet will be transmitted during the Rsync process.

As operational GeoTIFF images are created, these will also be transmitted to the SRS by the Rsync process.

The data management team at NCDDC will pull the GeoTIFF images and the operational bathymetry processing spreadsheet for near real-time metadata generation. Participating scientists wanting access to the raw multibeam in near real-time can pull the individual files with the metadata that provides operational and provisional processing steps and a disclaimer for non-QC status of the data.

Daily cumulative GeoTIFF images of the seafloor imagery will be geo-located on the Okeanos Atlas by the GIS team at NCDDC.

Post-Mission

All bottom-looking sensor data and complementary data (water column and sound velocity) are saved to a hard-drive. This hard-drive will be either brought back or shipped to the University of New Hampshire Center for Coastal and Ocean Mapping (UNH CCOM) for post-processing.

A full complement of multibeam data from a 30-day EX cruise on which the Kongsberg EM302 multibeam system runs continuously will produce 200-300 Gigabytes of raw multibeam (37.5% of total volume) and water column data (62.5% of total volume). At UNH, the mapping team will post-process the multibeam data through the following steps:

- The raw (level-0) data will be saved to the CCOM file servers, where they will be quality checked and post-processed.
- The edited level-0 data is saved as level-1 data files in a non-proprietary format – ASCII xyz files (cleaned not gridded).
- The post-processing steps used to produce the level-1 data will be documented.
- Level-2 products will be generated from the level-1 data files.
- The post-processing steps used to produce the level-2 data products will be documented.
- The level-1 data, level-2 products, post-processing steps, and working data processing spreadsheets will be copied to the hard drive in a new folder. A processing spreadsheet for FY13 will contain the temporal and spatial limits of each file and any supplemental information documenting problems or issues that affected the quality of the data in that file.

In FY13, an attempt to use an ftp protocol or collection service to transfer the multibeam data and products from UNH is planned. A normal hard-drive delivery will remain in effect as a backup until the digital file transfer process is sufficiently tested and becomes normal operations.

At NCDDC, all multibeam related files will be post-processed through metadata generation procedures. Metadata will be generated for each individual survey track file (level-0 and -1), for accompanying CTD/XBT profile data sets, for composite xyz files, KMZs, GeoTIFs, png images, and Fledermaus output (level-2), and a set of data products and reports (level-3). Finalized data/metadata will be compressed and bundled using the Bagit software and delivered to NGDC via ftp protocol.

Data Class	Instrument	Data Type	Format	Metadata Granularity	Archive Center
GEO	Kongsberg EM302 (30 kHz)	Multibeam Bathymetry, Bottom Backscatter, Water Column Backscatter (proprietary format read into MBSystem)	.all, .wcd (proprietary)	1 meta rec per .all file in Multibeam Data folder and subfolders	NGDC
GEO	Simrad EK60	Singlebeam (time,depth)	.txt, (ASCII), .raw (proprietary)	Included in the SCS feed	TBD
GEO	Knudsen CHIRP 3260 (3.5 kHz)	Sub-bottom profile	.sgy, .kea, .keb (proprietary)	1 meta rec = Subbottom Profile Data folder	NGDC
OCN	SeaBird SBE-911plus	CTD Cast	.hex, .con (Proprietary); .cnv, .hdr, .bl, .jpg (processed)	1 meta rec = CTD folder	NGDC
OCN	Sippican MK-21 eXpendable BathyThermograph (XBT)	XBT	.edf (ASCII), .rdf (proprietary)	1 meta rec = XBT folder	NGDC
OCN	RESON	Sound Velocity (m/s)	TBD	1 meta rec = RESON folder	NGDC
OCN	Calculated	Sound Velocity (m/s)	.asvp (ASCII)	1 meta rec = Profile_Data/SVP or Profile_Data/ASVP	NGDC

Table 4: Multibeam Survey Metadata Granularity and Target Archive

C.   Video Data Archive Pipeline

Low-resolution video segments will be archived at the NOAA Central Library (NCL) in Silver Spring, MD, a division of NODC. All available resolutions of the underwater video and their metadata will be temporarily stored in private and dedicated storage space on the NODC server and periodically backed up in a scheduled tape rotation.

Current Video Data/Products Pipeline
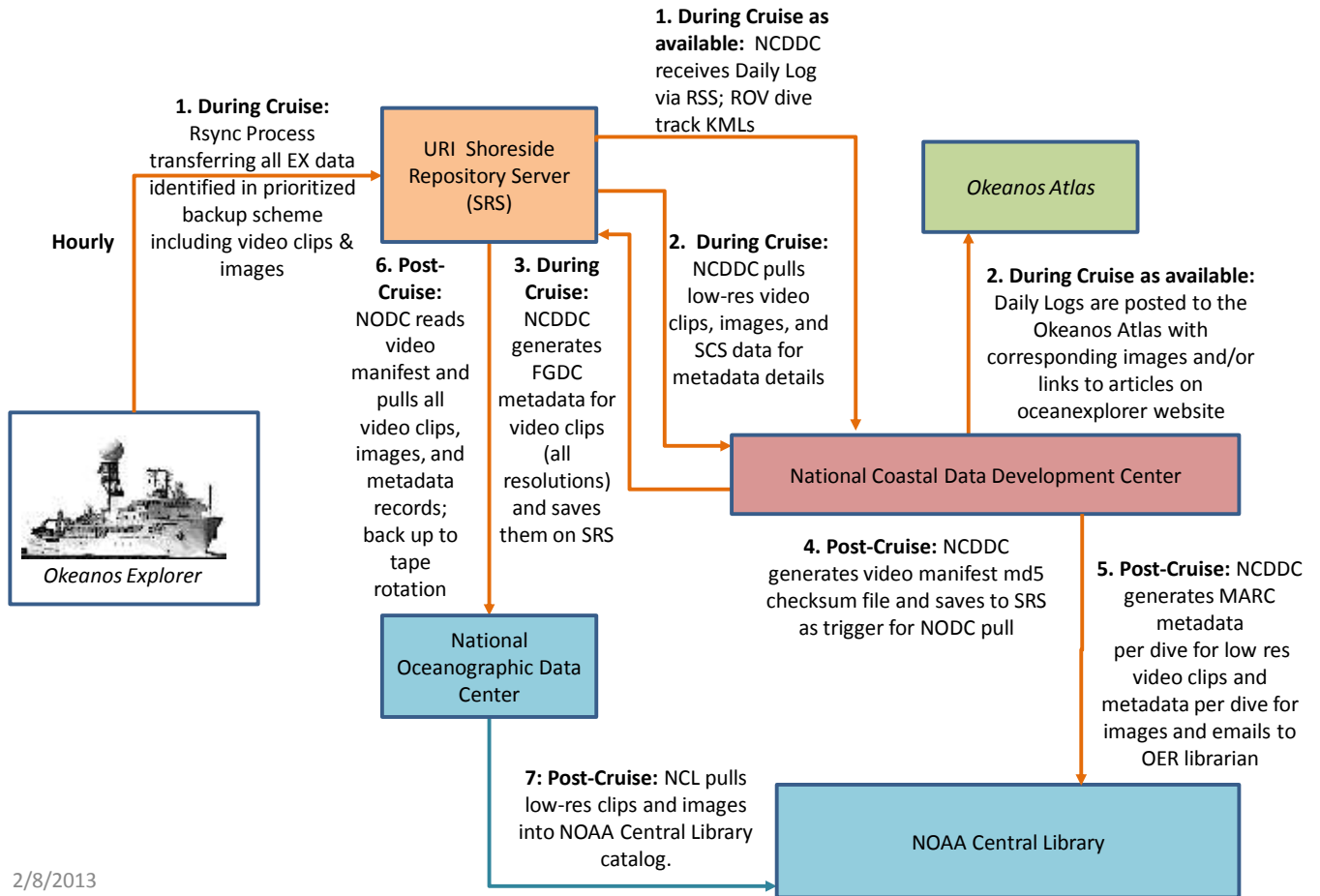
Fig 6: Okeanos Explorer Video Data Pipeline

3 July 2013   EX1304_DMP.pdf

Near Real-Time

Video segments to be preserved will be marked and saved onboard the EX by onboard videographers through collaboration with the onboard and remote science team. These clips will be saved with embedded metadata – cruise ID, camera ID, date/time, lat/lon, and file name and saved in the Ship Board Repository Server (SBRS). These enhanced multimedia files will be transmitted via an automated process (outlined in Section VI-B) to the Shoreside Repository Server (SRS) and saved in web-streaming low-resolution quality and, if possible, a medium-to-high resolution. The files will be named using a strict naming convention outlined in the "Okeanos Shore-Side FTP Server Standard Operating Procedures" document and in Section VI-C of this document.

Low-resolution video clips and images will be downloaded by NCDDC from the SRS for metadata generation routines. Image and video files will have embedded metadata and the file name will also include fields for the metadata.

Daily logs generated by the Expedition Coordinator will also be pulled from the SRS as they become available. Dive tracks in kml format are pulled from the SRS as they become available.

Daily Logs and representative images and dive tracks and links to representative video clips are displayed on the Okeanos Atlas,

The embedded information and the file names of the downloaded low-res video clips will be used in the routines to produce the FGDC metadata for each. An FGDC metadata record will also be generated for the medium-to-high resolution counterpart to the clip, although that clip will not be downloaded from the SRS. The generated metadata records will be named similarly to the video clips they represent and all metadata will be uploaded to the SRS in the same folder with the video clips. A manifest file with md5 checksum values will be generated daily for all of the video clips and metadata records available on the SRS. The manifest file will be uploaded to the SRS.

Post-Mission

At the end of the mission, MARC metadata for each dive will be generated for video clips and framegrab images. All MARC metadata records are emailed to the NOAA Central Library for the mission catalog.

A final manifest file and md5checksum file are generated and uploaded to the SRS.

  NODC automated routines will be in place to recognize when the md5 checksum file is available for processing. Video clips and corresponding metadata will be saved to NODC dedicated storage space and backed up to tape until such time that a permanent solution to high-definition video archive is available.

  The NOAA Central Library will pull all of the low-res video clips from the NODC server and do a bulk ingest into their system, cataloging these clips by corresponding dive in their online video data management system (VDMS).

Data Class	Instrument	Data Type	Format	Metadata Granularity	Archive Center
MUL	ROV/Sled Cameras	Low-res video clips	h.264 low	1 MARC meta rec per each dive	NCL
MUL	ROV/Sled Cameras	Medium-res video clips	h.264 med	1 FGDC meta rec per each	NODC/MDSD (temporary hold)
MUL	ROV/Sled Cameras	Highlight Images	.jpg	1 MARC meta rec for the folder	NCL
MUL	ROV/Sled Cameras	Still images	.jpg	1 MARC meta rec for the folder	NCL
MUL	Topside Cameras	Still images	.jpg	1 MARC meta rec for the folder	NCL

Table 5: Video Metadata and Target Archive