



Metadata Requirements and Management for VOS/SOOP – META- T Pilot Project

Derrick Snowden
NOAA/Climate Observation Division
META-T PP

Hester Viola, Mathieu Belbeoch, Bob Keely, Bill Burnett, Cathy Chen

Outline

- Meta-T Goals
- Workflow and Project History
- Metadata Timeline
- Example for VOS data flow
- Future needs from SOT

META-T Goals

Originally a request from OOPC to collect manage and distribute Metadata from JCOMM observing systems to ensure both real time applicability and long term preservation of data

Implication was that JCOMM data streams were not fully satisfying climate requirements

META-T Goals

Instrument/Platform

- Sensor and Platform characteristics including calibration, installation configuration, reporting characteristics etc.

Lineage or Provenance

- Sequence of events leading up to the data as it exists on your desktop

Collection/Distribution

- Sources and types of data, including algorithms, that characterize a collection (e.g. icoads)

META-T Process

- Start with user group surveys to help define requirements
- Assemble the responses and group the metadata elements into subgroups depending on when they are needed in time
- Assign the metadata needed in real time to the GTS group → BUFR templates
- Begin to draft VOS/XBT BUFR templates
- ...

How to get the effort moving again?

- For each data stream, analyze the steps and determine what information we have access to/and when
- Look at the gaps relative to the initial requirements
- Augment the GTS message (BUFR) to contain as much metadata as is available at the time of distribution
- Create a process that links the delayed mode metadata with the real time observation

META-T Categories

- GTS Distribution
- Anything not coming from the ship must be available in real time on a server

Real Time with observation

1

- Less time sensitive
- Should be available as soon as it is known in a central repository
- Delayed mode updates to Pub47 for example

As soon as possible at central server

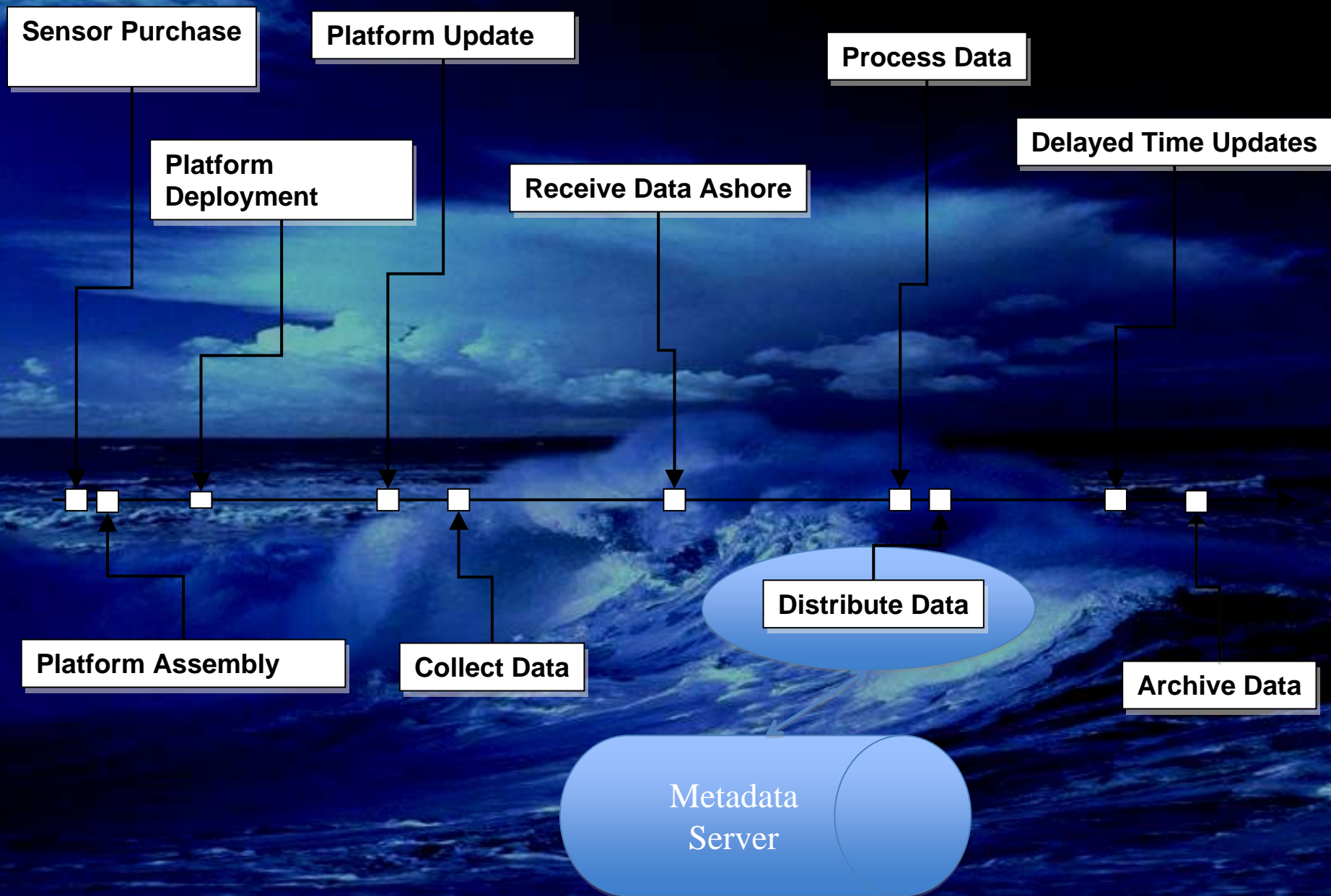
2

- Any other information that is needed to ensure the long term preservation of the data at an archive center

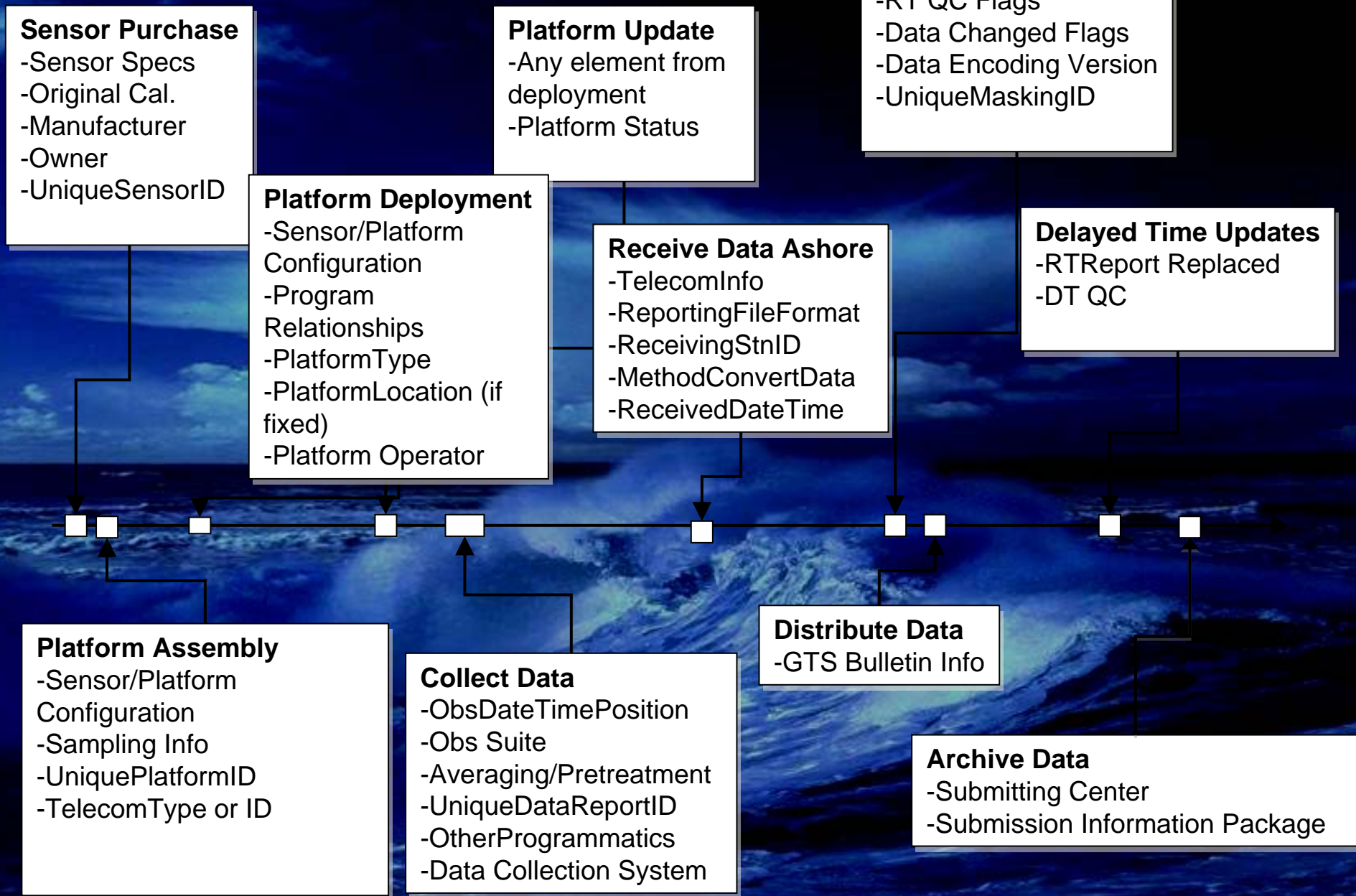
Delayed mode as part of the Archive package

3

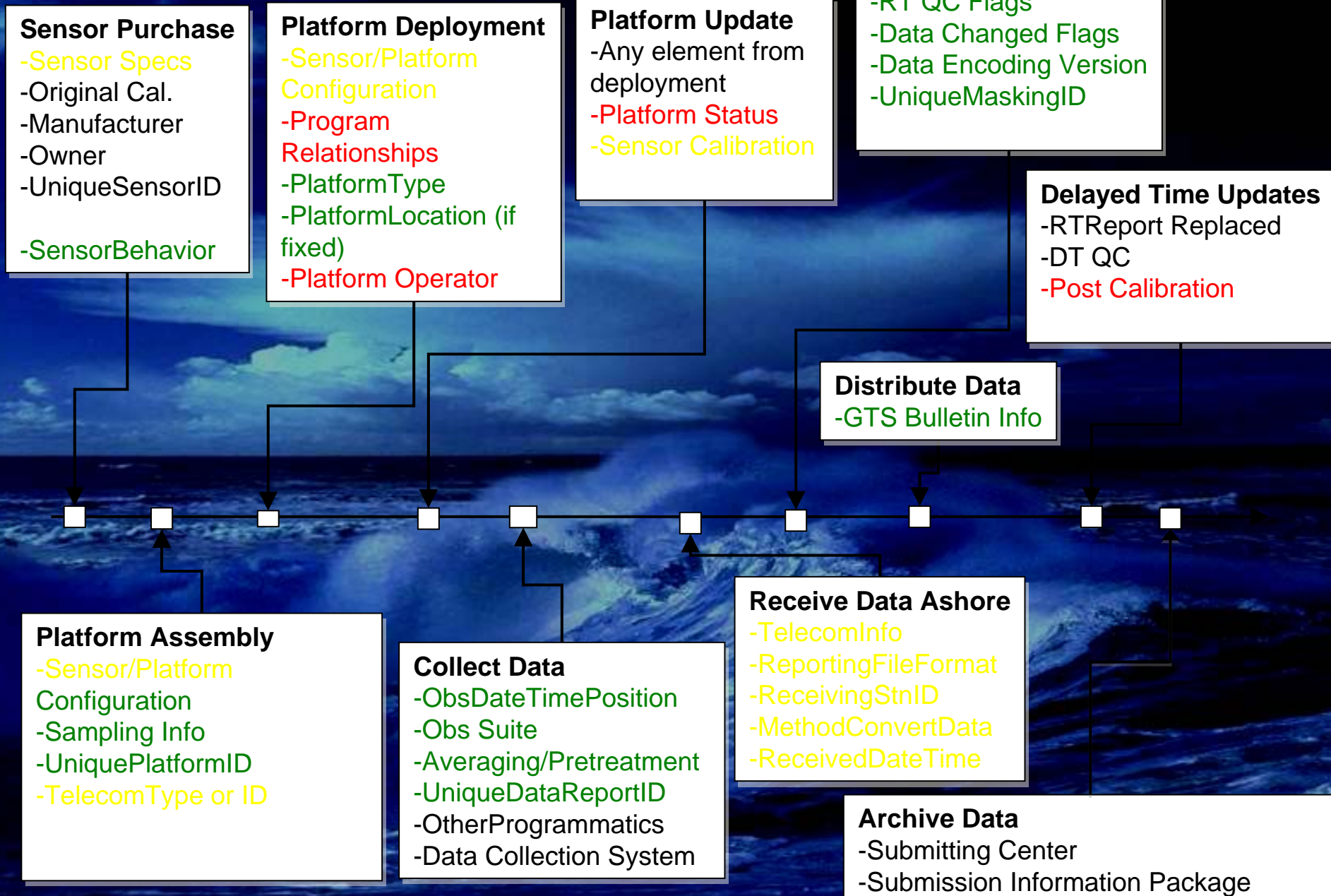
Sensor/Platform View of Data Lifecycle



Sensor/Platform View of Data Lifecycle



Sensor/Platform View of Data Lifecycle



Example: VOS – There is extended metadata available

- Ships have telecoms and elogbooks
- Ships have telecoms and no elogbooks
- Ships have elogbooks and no telecoms
- Ships have neither

All imply
mods to
eLogbook
software
and
operational
procedures

Cost implications...?

Generic Data Flow

GOAL

Automated Reports
from the Ship
(satcoms)

Manual Data entry
from eLogbook

Manual Data entry
from Paper or
Portable PC

Metadata
Server

GTS
Distribution

Network
Monitoring

DAC/Archiv
e

Scenario 1: Ship has telecoms and eLogbook software

All metadata with each obs

- High transmission cost
- Low risk of information loss

Send periodic Admin message

- Flexible and able to be done frequently enough for MD to be relevant
- Requires a strategy to match up with observations somewhere down the path
- SEAS/TurboWin/ObsJMA must be modified

eLogbook summary file retrieved by PMO

- File size isn't an issue
- Potentially large MD latency (Need more input from SOT here)
- SEAS/TurboWin/ObsJMA must be modified

Combinations

- Allow for national implementations that combine the above

Scenario 2-4: All involve manual update

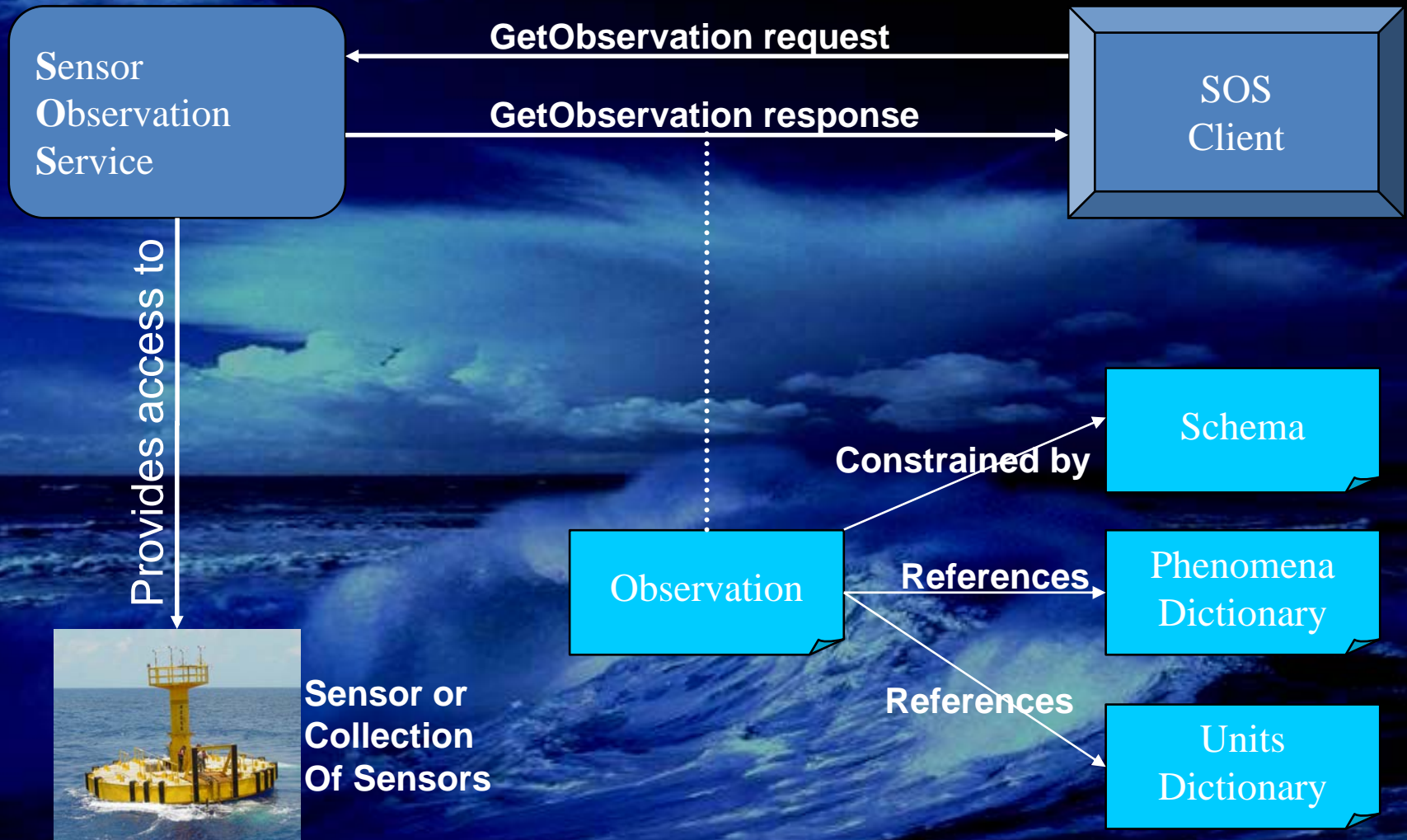
Update via eLogbook

- Allows for more standardization
- Data entry through same tool as obs tool

Update via paper forms or hand carried laptop

- e.g. Pub47 XML Generator
- Requires most manual entry and effort by PMO

SOS Concept



Is there anything new here?

- For the VOS in particular, is there anything in the META-T project that is not simply a reframing of the same issues surrounding updating and managing the Pub47?
- Minor things such as formatting, Pub47 XML vs. Standard Metadata Description Schema (SensorML, ISO)
- Distribution and management timeline?
- Using Pub47 in real time?
- Integrating Pub47 type metadata collection schemes across other ship based systems.

Conclusions and Recommendations

- Need input from TSG community and better interaction with the VOS on actual data flow
- Need an interaction with eLogbook developers to see how we might augment the metadata collection
- For VOS, where can we collect Pub47 submissions more frequently than quarterly? Direct access to Esurfmar or similar? Submission to JCOMMOPS?
- For SOOP, need to define a Pub47 analog.
- Minor developments to be expected at NDBC, but mostly this is an unfunded effort, VOLUNTEERS
- NMDIS in China is pursuing similar efforts for the ODAS data stream.