

Data collection

On-board the RV Investigator

Steven Van Graas | Seagoing Data Analyst
18 July 2017

INFORMATION & DATA CENTRE | OCEANS & ATMOSPHERE
www.csiro.au



Overview

- RV Investigator
- Data streams - diversity of sources
- Aggregating data – making it useful
- Data presentation and availability





The RV Investigator

- Delivered September 2014
- Length: 94m
- Range 10,000 NM
- 40 science berths, 20 crew



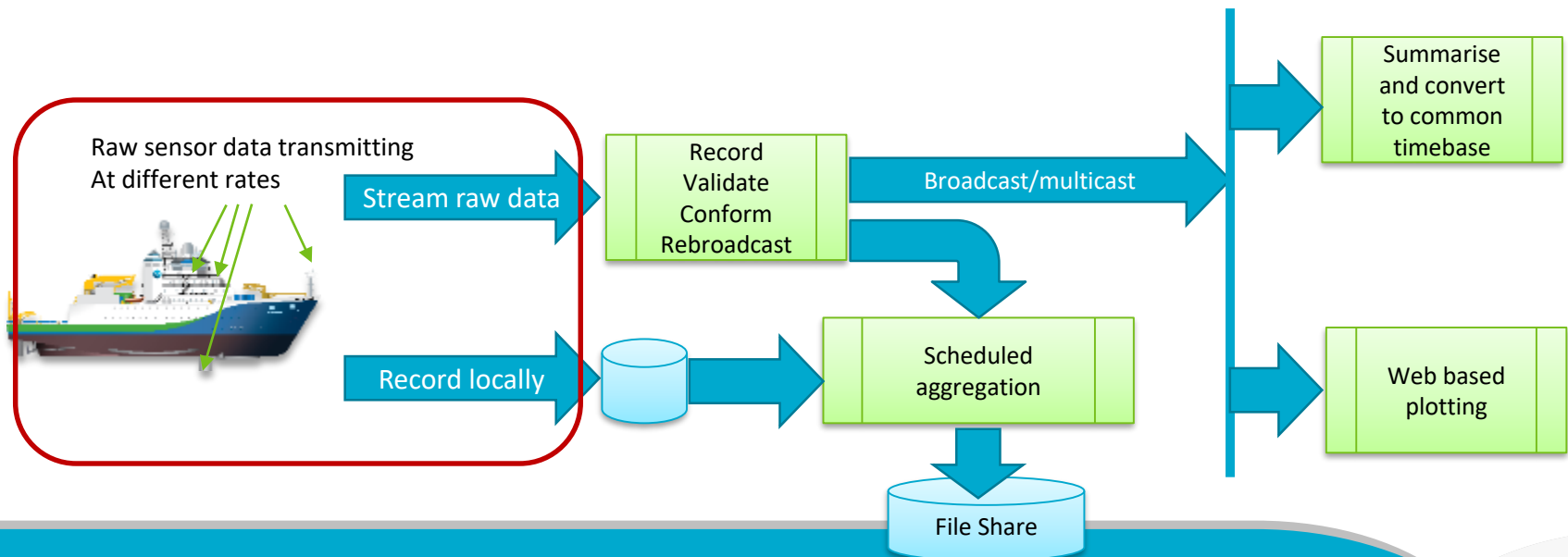
Supporting the Investigator

- Technical Support
 - Data Acquisition and Processing (DAP)
 - Seagoing Instrumentation Team(SIT)
 - Geophysical Surveying and Mapping (GSM)
 - Hydrochemistry Team
- Equipment and Instrumentation
 - Continuously monitoring sensors
 - Meteorology, aerosols, underway seawater
 - Towed and lowered packages: CTD, Triaxus, towed video
 - Sounders: Single and Multibeam, ADCP
 - Satellite internet and communications platforms

Data Acquisition

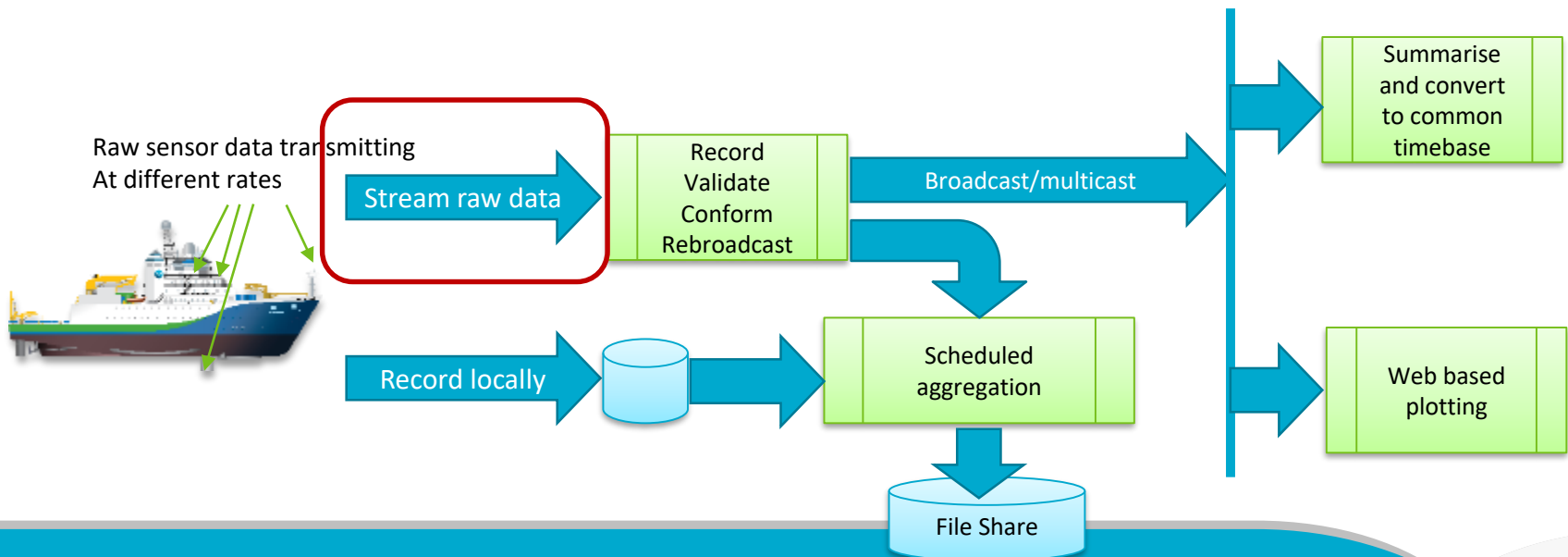
- Data Sources and Acquisition

- Over 100 fixed sensors, plus deployed equipment, temporary and container labs
 - Streaming data: Serial, UDP, multicast, TCP
 - Locally recorded



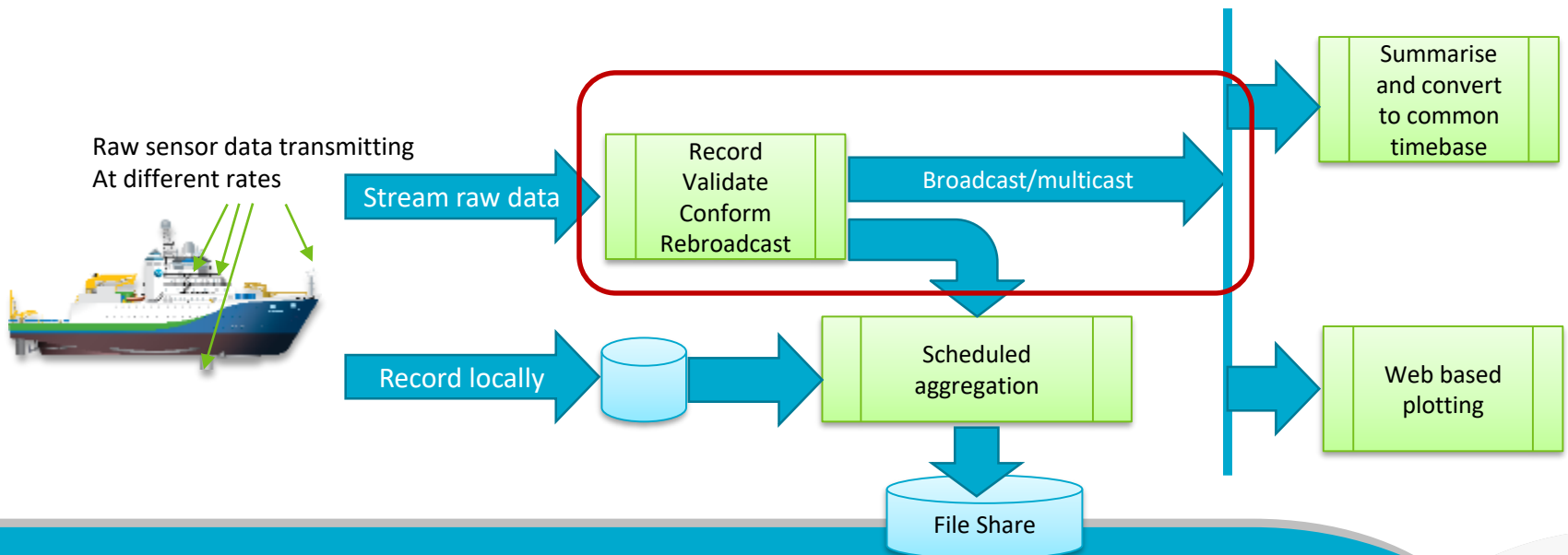
Data Acquisition

- Data Sources and Acquisition
 - Streamed data
 - Generalise to network: UDP/multicast
 - Control
 - 0194 boards



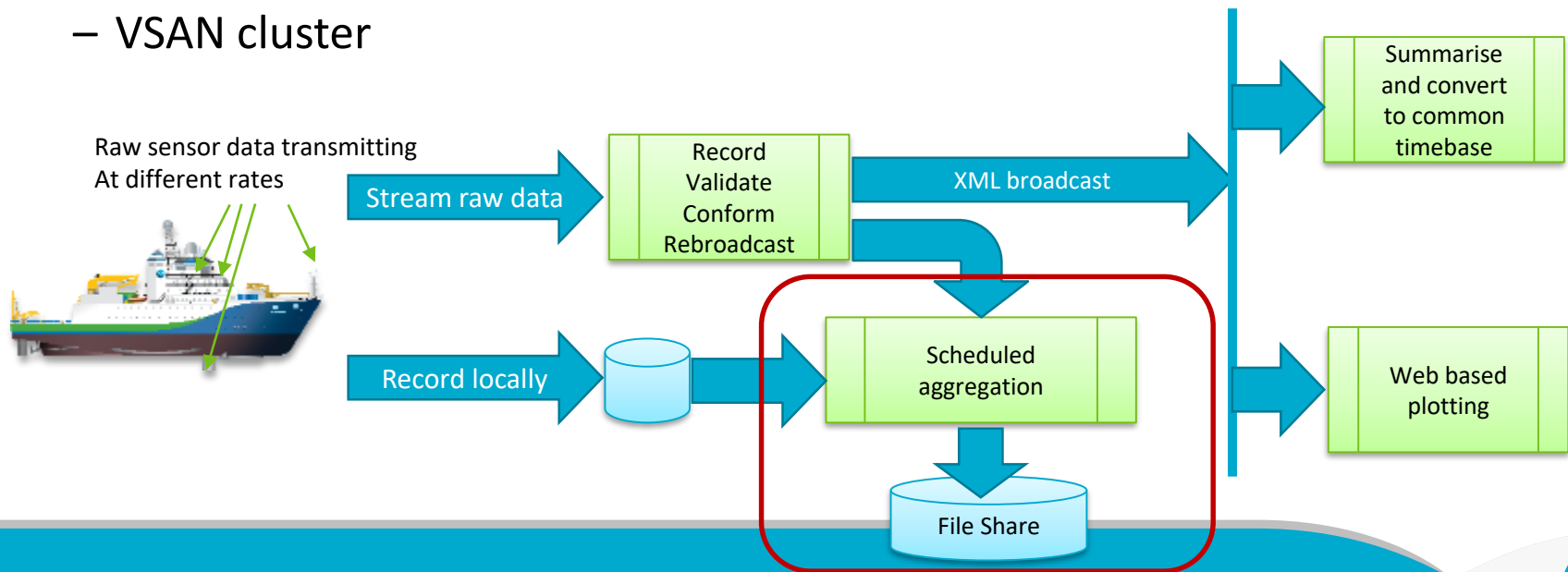
Data Acquisition

- Network streamed data recorded via TECHSAS
 - C++, Ifremer, customised in house
 - Validation and alerts
 - Converting raw data to standardised format with metadata (NetCDF files)
 - Tailored output broadcast/multicast



Data Acquisition

- Aggregation and Archiving
 - Locally recorded data collected by Delta
 - Python, web-based (flask), in house
 - Scheduled retrieval from network shares
 - Generates manifest of all data verified with MD5 hashes
 - Manages storage configuration upon start and end of voyage
 - VSAN cluster





<input type="checkbox"/>	Job 101	PENDING	underway merger data \\underwaymerger\data	1. underway merger data current <ul style="list-style-type: none"> Path: \\science\voyages\current\underway Files: 16803 Size 66.55 GB 	1. 1 files copied - 320.03 KB	0.5	00:43:21	COPY
<input type="checkbox"/>	Job 39	PENDING	OS75 ADCP \\ADCP-ACQ-OS75\ADCP Data\current	1. ADCP OS75 Science Share <ul style="list-style-type: none"> Path: \\science\voyages\current\adcp\os75 Files: 206 Size 109.18 MB 	1. 15 files copied - 359.49 KB	0.5	00:00:20	COPY
<input type="checkbox"/>	Job 56	PENDING	LADCP CTD machine Data \\ctd\Data\LADCP	1. LADCP CTD current voyage <ul style="list-style-type: none"> Path: \\science\voyages\Current\adcp\ladcp Files: 1 Size 578 B 	1. 0 files copied - 0 B	0.25	00:01:52	COPY
<input type="checkbox"/>	Job 49	PENDING	ops-gp1 ELOG \\ops-gp1\logbooks	1. ELOG current voyage <ul style="list-style-type: none"> Path: \\science\voyages\current\science\ELOG\logbooks Files: 32 Size 114.97 KB 	1. 1 files copied - 644 B	0.5	00:05:20	COPY
<input type="checkbox"/>	Job 2	PENDING	ek60 2nd data \\ek60\F Drive Data\in2015_v01	1. current voyage EK60 <ul style="list-style-type: none"> Path: \\science\voyages\current\sounders\ek60 Files: 1247 Size 42.14 GB 	1. 0 files copied - 0 B	0.5	00:11:20	COPY
<input type="checkbox"/>	Job 40	PENDING	ADCP 150 current \\ADCP-ACQ-OS150\ADCP Data\current	1. ADCP 150khz Science Share <ul style="list-style-type: none"> Path: \\science\voyages\current\adcp\os150 Files: 1183 Size 2.97 GB 	1. 2 files copied - 2.48 MB	0.5	00:26:20	COPY
<input type="checkbox"/>	Job 41	PENDING	PCO2 \\150.229.233.130\out	1. current voyage PCO2 <ul style="list-style-type: none"> Path: \\science\voyages\current\pco2 Files: 17 Size 1.46 MB 2. PCO2 outgoing <ul style="list-style-type: none"> Path: \\dc1\Public\outgoing_data\pco2 Files: 23 Size 2.12 MB 	1. 0 files copied - 0 B 2. 0 files copied - 0 B	0.5	00:32:20	COPY

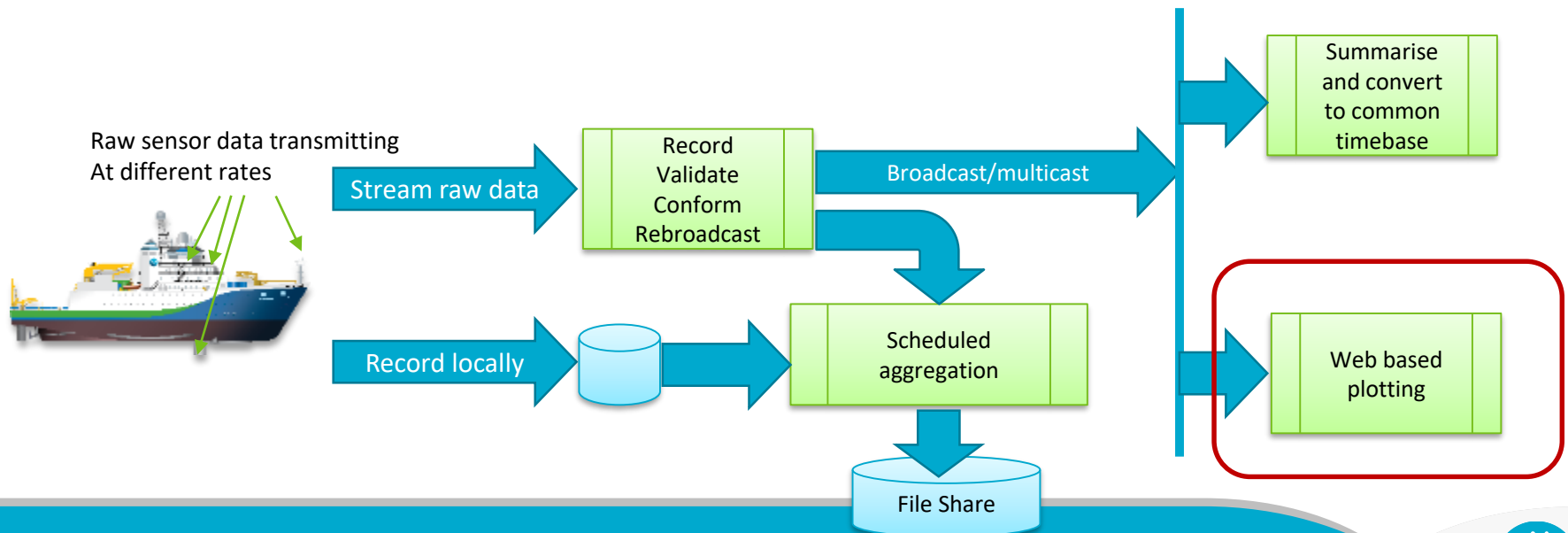
1. [current voyage triaxus](#)

Availability

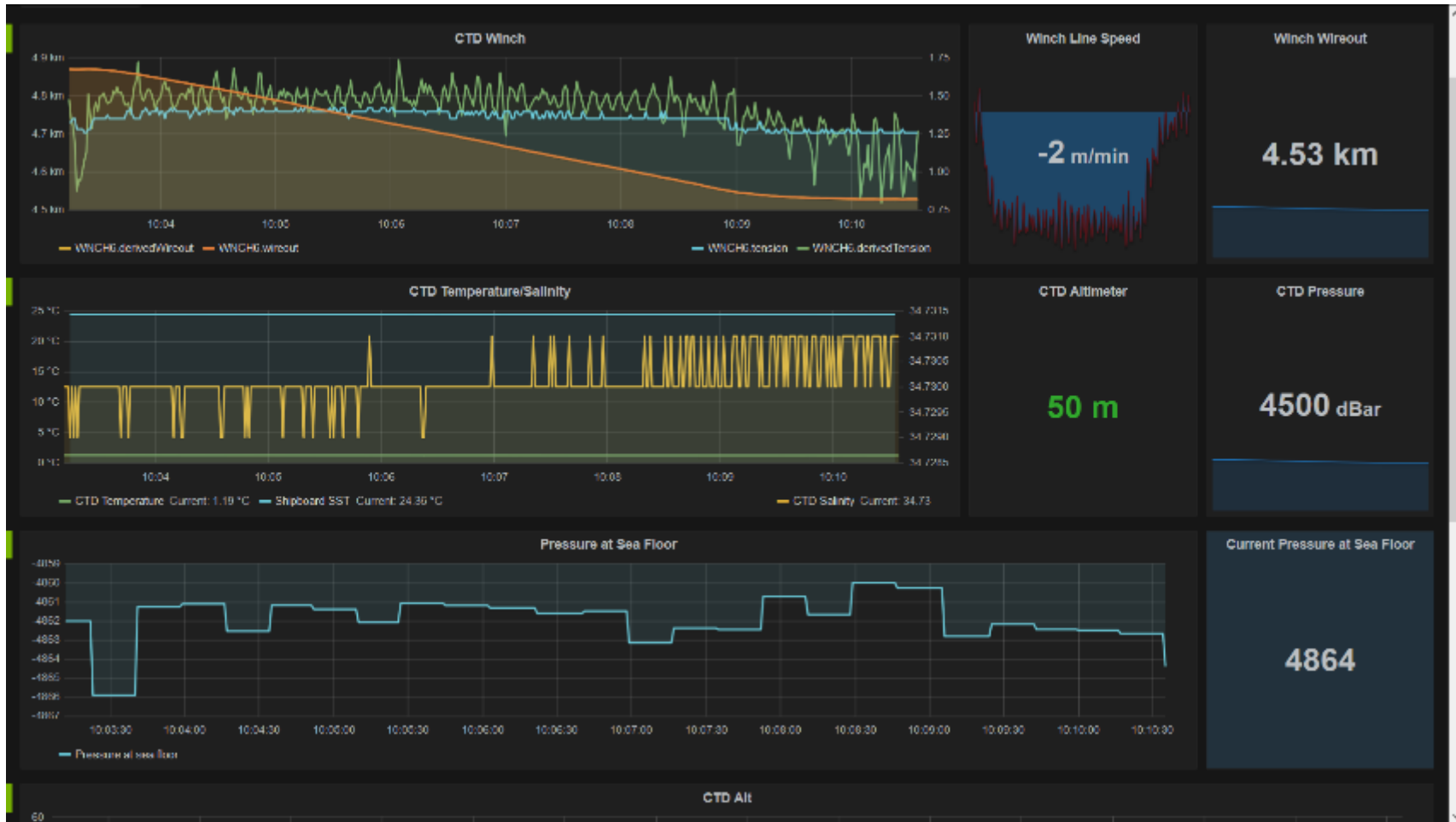
- Presentation

- InfluxDB and Grafana

- open source – time series analytics software
 - Web based plotting and basic data extraction (csv)
 - Python collection scripts allow for data manipulation

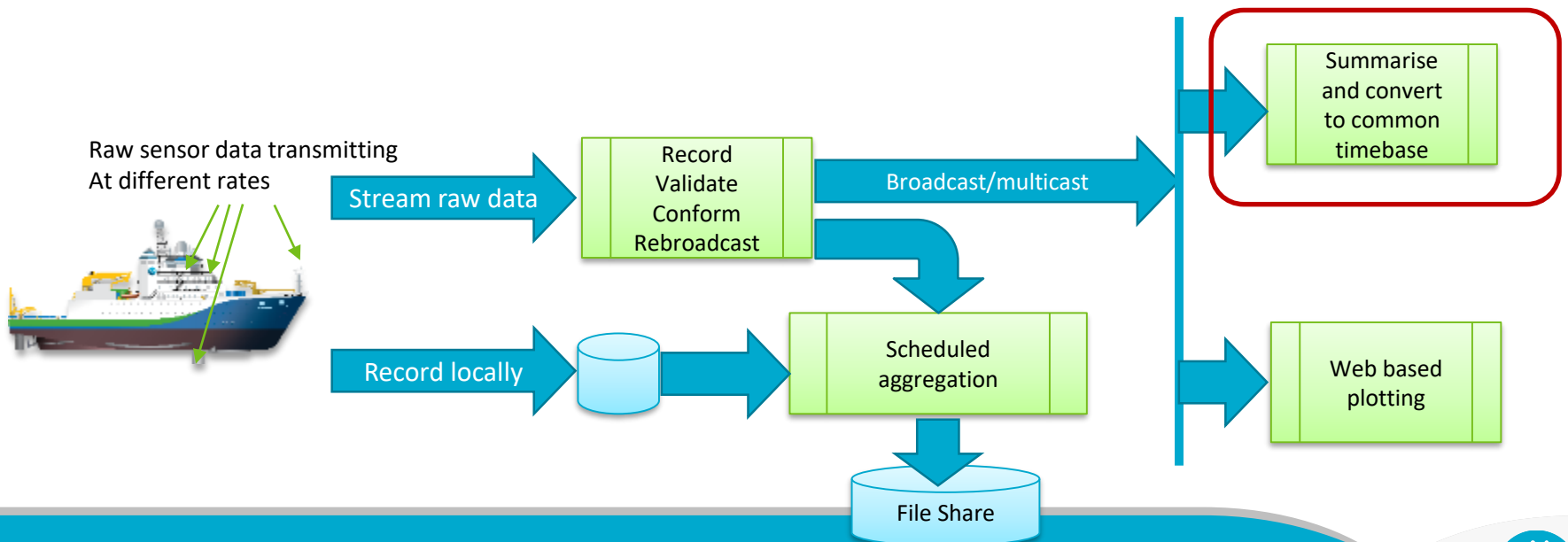


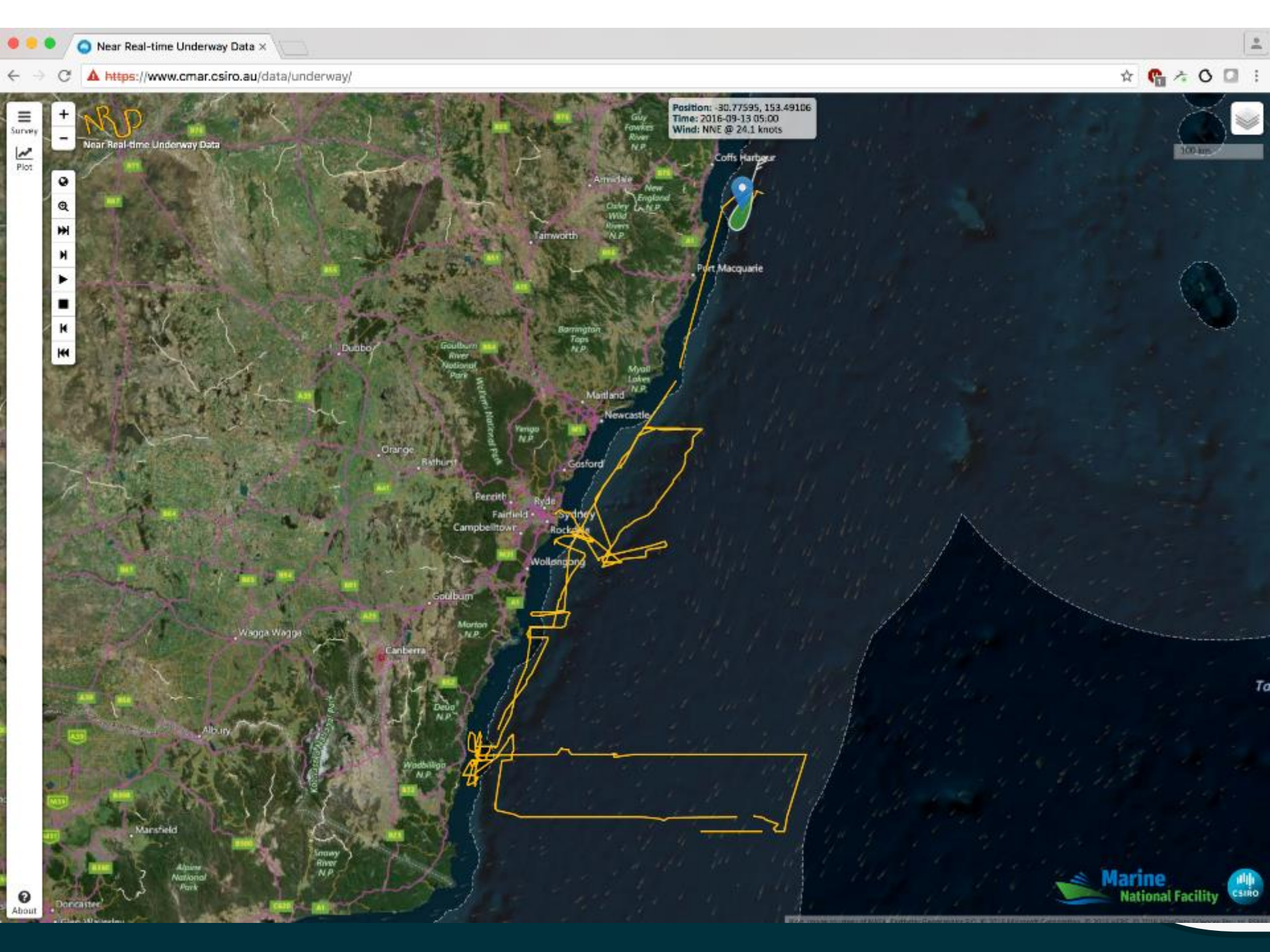
Grafana



Availability

- Summary data - UnderwayMerger
 - Java, in house
 - Converts to common time base (5 second) via interpolation (NetCDF)
 - Generates 60 second average data based on interpolation (CSV)
 - Feeds shore-based summary database (NRUD)





Summary

- Data aggregation and dissemination processes
- Evolving platform
- Higher availability of data and resources improving outcomes



Questions?

Thank you

CSIRO

Steven Van Graas
Seagoing Data Analyst

t +61 3 6232 5047

e steven.vangraas@csiro.au

OCEANS & ATMOSPHERE

www.csiro.au

