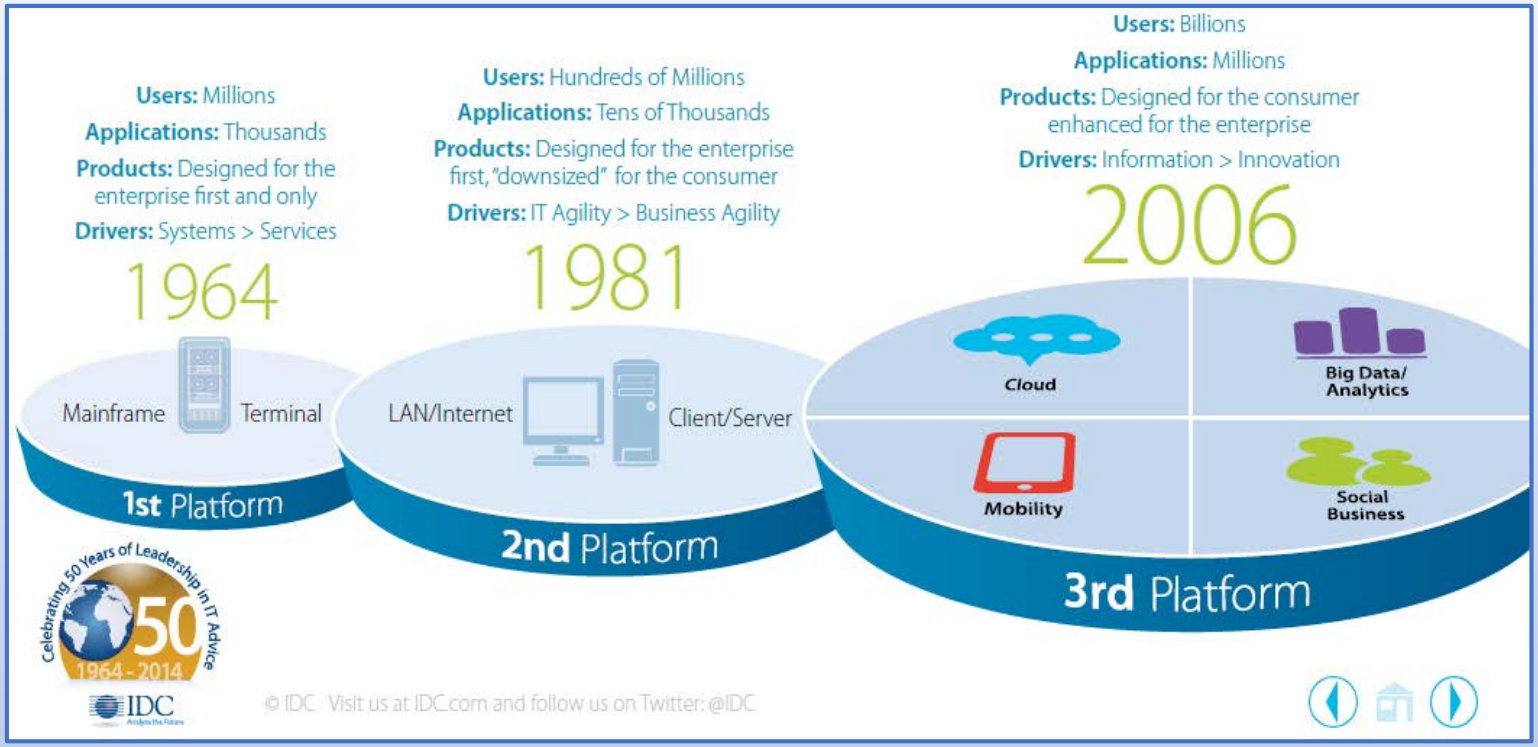
A Cloud based Solution in Hydrographic Data Processing: The Shift to a Web Centric Software Platform

US Hydro 2017

Daniel Wright, *Process Researcher* & Charles Wright, *Principal Software Engineer*



- *2017 projected worldwide spending on public cloud services and infrastructure $122.5 billion, an increase of 24.4% over 2016*

- *2015-2020 forecast - 21.5% compound annual growth rate (CAGR) nearly seven times the rate of overall IT spending growth*

- *By 2020 software as a service (SaaS) versus traditional software deployment will be over 25%, packaged software will shrink to 10% of new enterprise installations*

IDC Research, Inc.

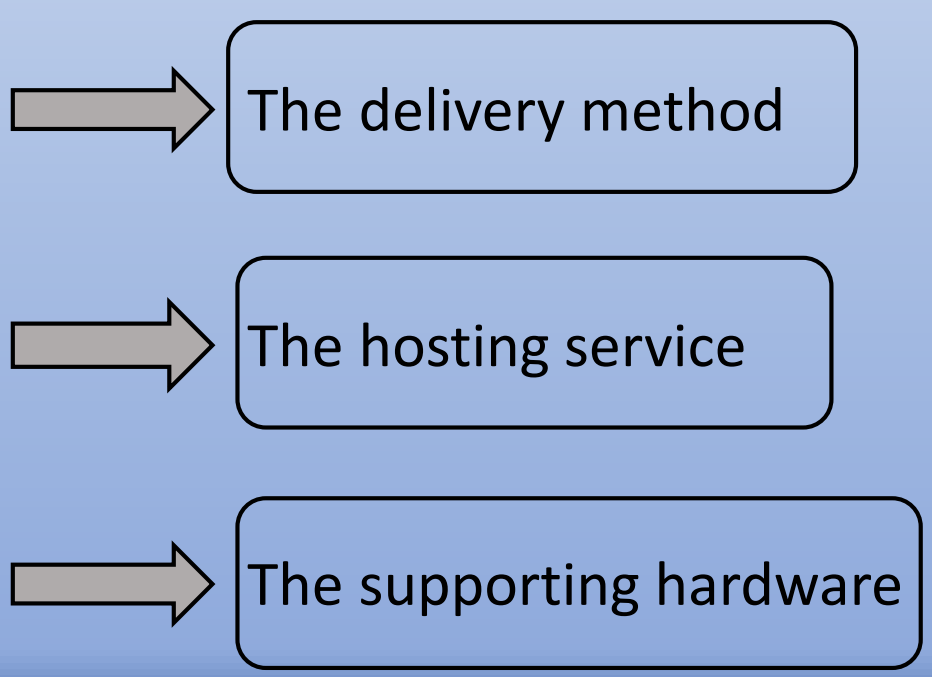
NIST
National Institute of Standards and Technology
U.S. Department of Commerce

Definition of Cloud Computing

(Special Publication 800-145, 2011)

"…a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources….and that can be rapidly provisioned and released with minimal management effort or service provider interaction"

Service Models for Cloud Computing (NIST)
Software as a Service (SaaS) – Provider's applications running on a cloud infrastructure. The applications are accessible through a thin client interface, such as a web browser.
Platform as a Service (PaaS) – For deploying created or acquired applications onto the cloud infrastructure using provider tools and services.
Infrastructure as a Service (IaaS) – provisioning of processing, storage, networks, and other fundamental computing resources.



NIST
National Institute of Standards and Technology
U.S. Department of Commerce

Definition of Cloud Computing

(Special Publication 800-145, 2011)

Essential Characteristics of Cloud Computing (NIST)
On-demand self-service **– the ability to automatically provision computing resources.**
Broad network access **- available over the network and accessed through standard mechanisms (e.g., mobile phones, tablets, laptops, and workstations)**
Resource pooling **- The provider's computing resources are pooled to serve multiple consumers**
Rapid elasticity **- Capabilities can be elastically provisioned and released automatically, to scale rapidly commensurate with demand.**
Measured service **- Resource usage can be metered, monitored, controlled, and reported, providing transparency for both the provider and consumer**

Elasticity

Provision and release resources automatically.

Scaling Up

Increase CPU capacity, memory, disk space, virtual machines (VMs), domains, certificates, staging slots.

Scaling Out

Number of instances running an application, i.e. # VM's running at any given time.

Advantages of a cloud based processing platform

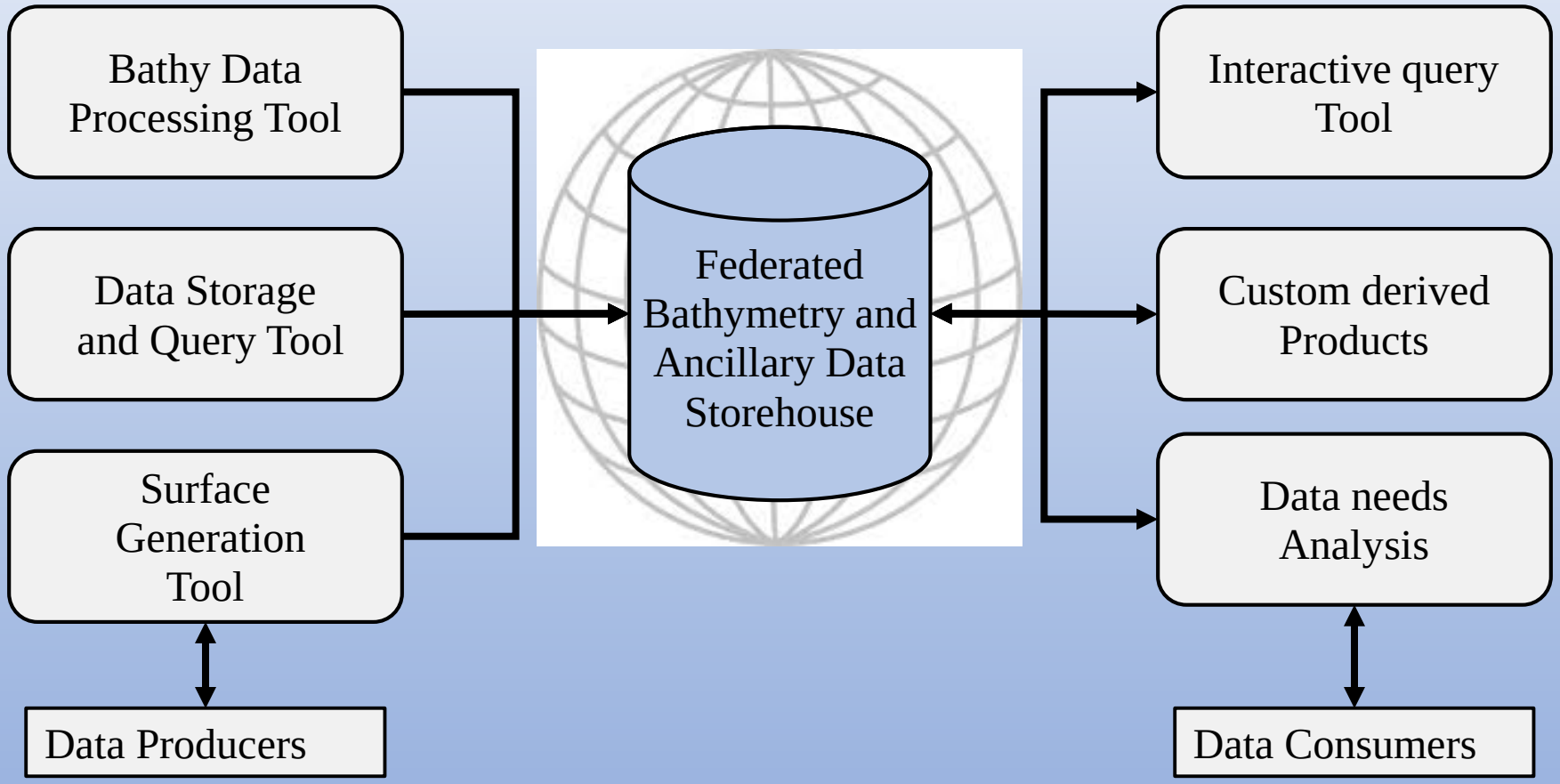
- Faster and more responsive product development cycle
- Elastic and scalable to user demand
- Improved security for some organizations
- High service availability (mostly)
- Automated backup and redundancy of data storage
- Document and data version control issues reduced
- Very low capital expenditures
- Pay-for-Use or Subscription service model
- Reduced environmental footprint

Cautions!

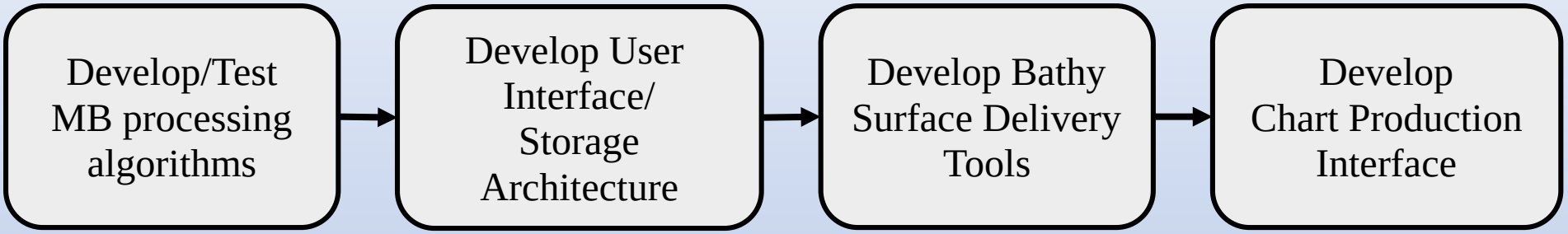
Potential pitfalls of cloud based processing

- Internet dependent – no internet, no service.
- Security - very low probability of attack, but still possible.
- Vendor Lock In - difficult to migrate services from one vendor to another.
- Inherent liability of working with data intended for navigation.
- Lack of specialized tools and features - they must be built to run on a cloud based application.
- Regulatory and Compliance issues – be sure this meets your organization/country specific requirements.

Cloud Based Tool Kit



Development Phases



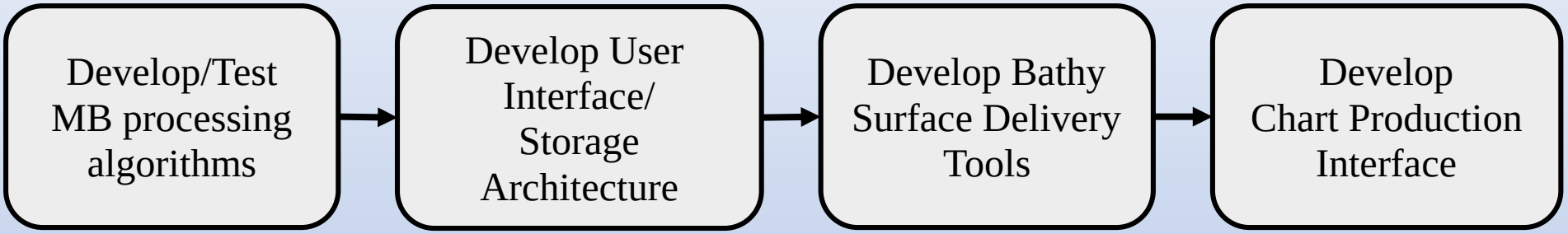
Stage 1: Develop/Test Bathy Processing Engine

- Optimize code - highest speed/lowest cost per transaction
- Iterative development cycles for continuous improvements
- Processing tools built specifically on and for cloud architecture

Stage 2: Develop User Interface and Storage Architecture

- Create simplified User Interface that is platform agnostic
- Emphasize empirical analytical tools versus visualization tools
- Provide Storage as a Service within a Federated database structure

Development Phases



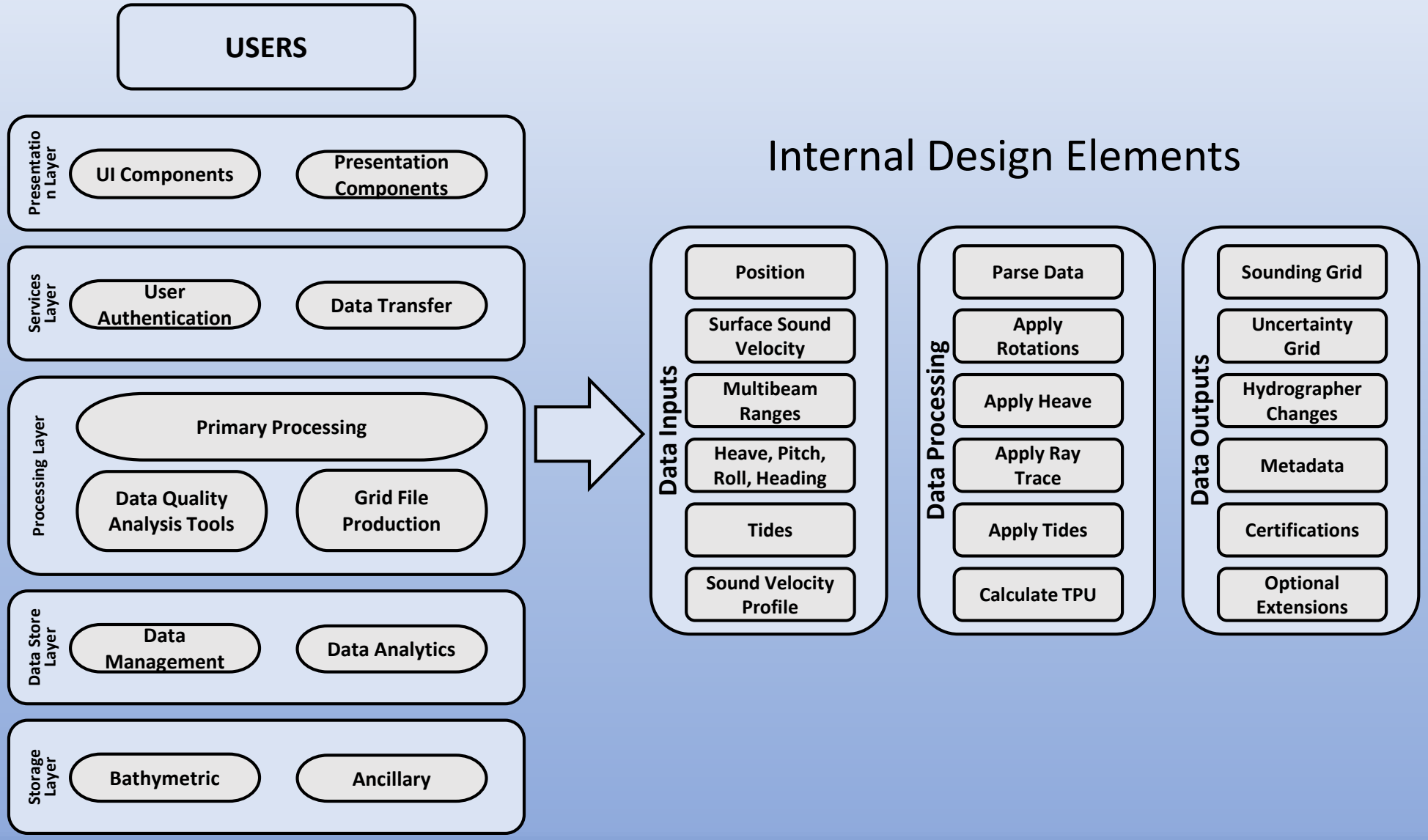
Stage 3: Bathymetric Surface Delivery Tools

- Customer defined products: S-101, SSDM, OGC WMS
- Utilize all available resources for a given area: current data, existing bags, Satellite Derived Bathymetry, legacy (smooth sheets), etc.
- Determine gridding methods appropriate for desired output: tiling schemas, single resolution, variable resolution, hybrid, etc.

Stage 4: Chart Production Interface

- Continuous analysis of data warehouse needs and resources
- On demand delivery of verified bathymetry products
- Integration to existing ENC production software with cloud based data delivery services

Architecture of a Cloud based Processing Platform



De-Coupled Design

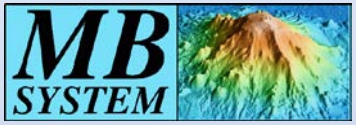
Inputs require specific formatting

- Designed for efficiency of transfer and processing.
- Does not require changing the architecture or design of the system.
- Only the data critical to the sounding solution gets into the system.
- Data input can be configured using common python tools
- Or…We can write code to ingest whatever data streams are present and provide feedback on quality, systematic errors or other indicators.

Outputs are standalone products

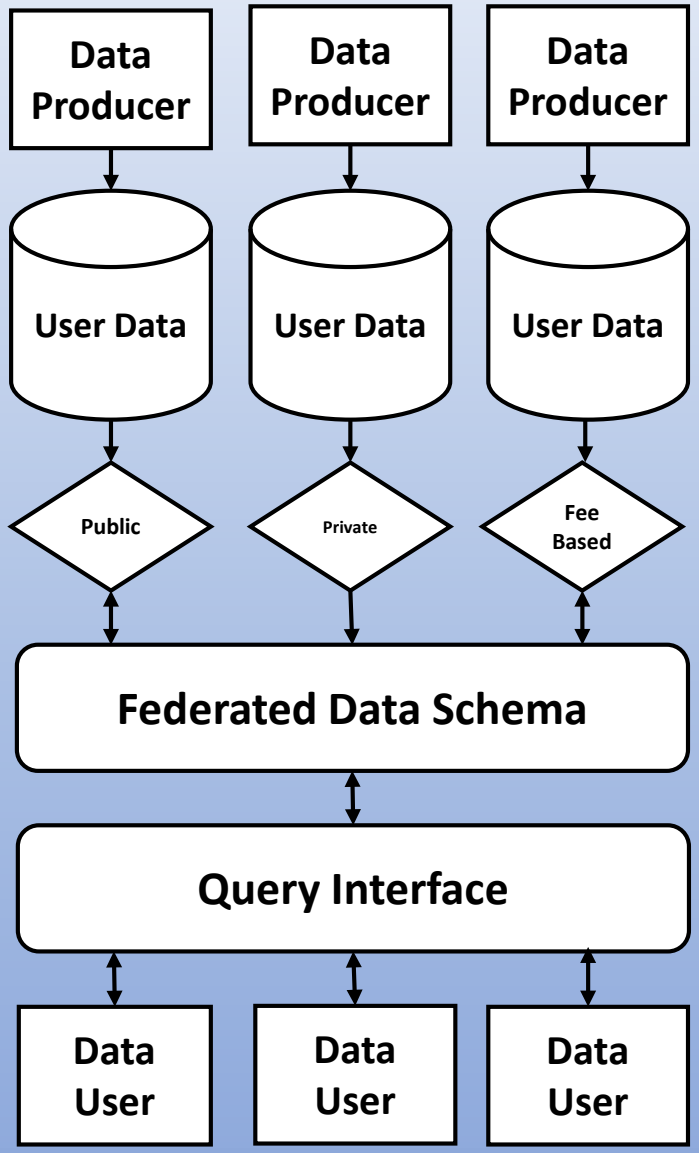
- Grids are intended as the final product offering.
- This follows the work of the ONSWG.
- BAG file format designed as a self contained record, the end user does not require a direct connection to the source.

Developer Tools



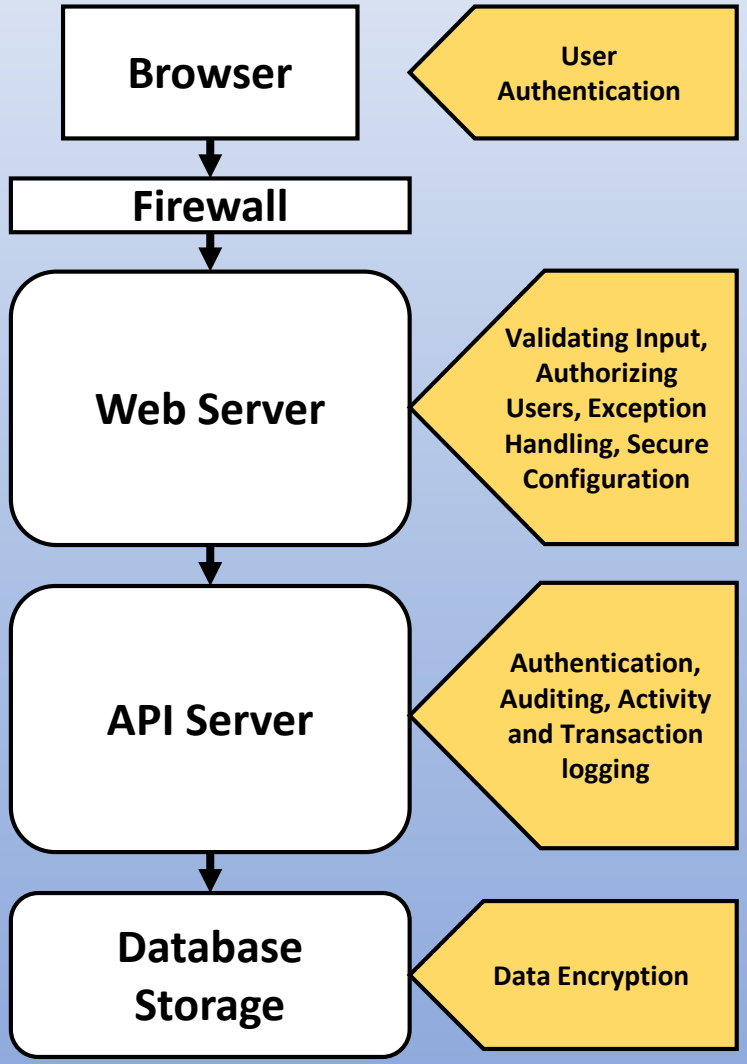
- Open Source components used wherever possible
- Allows transparency of the code and general acceptance
- Fosters a sense of ownership by the user community
- Positive feedback loop for continued product improvement
- Use a collaborative development platform (Microsoft Visual Studio) when code must be developed

A Federated Database



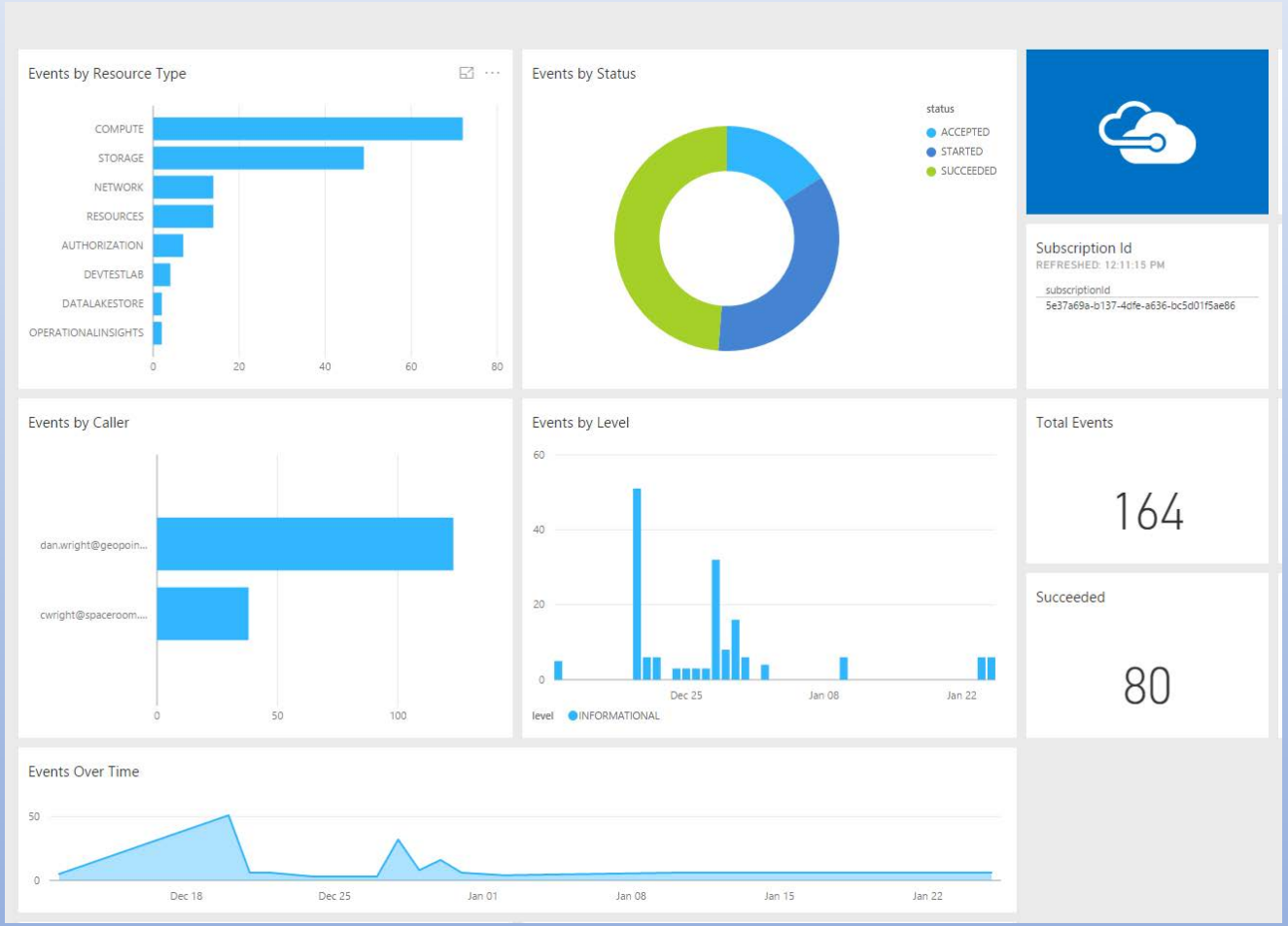
- Each data producer retains a unique entity
- Each constituent database remains autonomous
- Data producers opt to share or keep data private
- System acknowledges the existence of a particular data set via the index of BAG metadata
- Data retrieved via a common import/export schema
- Data Users can query and access authorized data for producing navigation products

Security, Logging and Compliance



- User Authentication
  - Provides the means of managing and controlling user identity and access to the applications, environment and data.
  - Multi-factor authentication provides the highest level of security.
- Encrypted communications and processes
  - For uploading data between user devices and datacenters, and within datacenters
- Data at rest (in storage)
  - cloud service providers use encryption capabilities that meet the highest available standard (AES-256)

Security, Logging and Compliance



Real time event monitoring

- Customer usage stats
- Capacity Monitoring
- Usage Analytics

Customized User Dashboard

- Storage usage
- Processing efficiency
- Pay-Per-Use cost tracking

Security and Event Tracking

- Forensic research
- Hacking Prevention

Security, Logging and Compliance

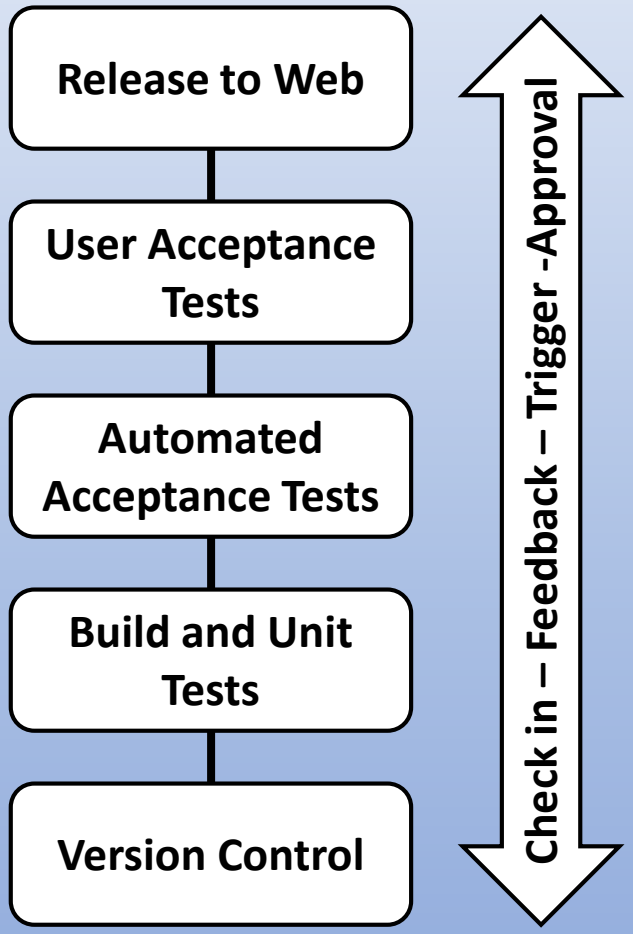
ISO/IEC 27001:2013 Information Security Management Standards

Federal Risk and Authorization Management Program (FedRAMP)

US Department of Defense (DoD) Level 5 Provisional Authorization

https://www.microsoft.com/en-us/trustcenter/compliance/complianceofferings

Continuous Integration, Delivery and Deployment



Continuous Integration (CI)

- Relies on a Source Control Repository
- Merges all developer working copies several times a day (avoids “integration hell”)
- Testing done at each stage of development and in parallel to live operations

Continuous Delivery (CD)

- Building, testing, and releasing software faster and more frequently
- Reduces cost, time, and risk of delivering changes by allowing for more incremental updates

Continuous Deployment

- Faster and more frequent releases, better customer feedback, improved reliability
- But…some users may not want updates during processing operations

Machine Learning and Quality Control

How can cloud processing help us transition from manual area (or swath) based quality control techniques to algorithmic?

- Algorithms can be trained to discern between, and learn from, multiple sources of anomalous data to allows a set of solutions to be suggested, from which changes can be made to the data.
- The processing must be fast enough (a few seconds) to allow the inspection of multiple solutions interactively. Parallel processing can enable this function.
- This allows the user to experiment with ‘what if’ scenarios, such as:
  - Changes to patch-test solutions, beam angle or across track filters and CUBE variables.
  - Make observations of the effects of changes in real time, not on a single line but on the surface as a whole.
- HydrOffice QC Tools (CCOM/UNH) provides examples of algorithmic analysis in action: Flier Finder and Detect Holidays are two examples.

www.geopointsolutions.us

The website hosts examples of critical code for review and comment.

Test data sets will be provided for independent confirmation of results.

System Architecture and Design will be open for public comment and input.

Project goals and timelines will be posted.

Potential user community interest and involvement can be addressed.

Sources

J. Boers and L. Persson, “A View on Trends in Hydrographic Processing Software: In What Direction is Progress Headed?” *Hydro International*, vol. 20, no. 8, pp. 25–27, Nov./Dec. 2016.

Sales Force UK & Ireland Blog. “Why Move to the Cloud? 10 Benefits of Cloud Computing.” Posted Nov 17, 2016. [Online]. Available: https://www.salesforce.com/uk/blog/2015/11/why-move-to-the-cloud-10-benefits-of-cloud-computing.html

Peter Mell and Timothy Grance. (2011 , Sep .). Special Publication 800-145 The NIST Definition of Cloud Computing. National Institute of Standards and Technology [Online]. http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf

L. Leong, G. Petri, B. Gill and M. Dorosh, “Magic Quadrant for Cloud Infrastructure as a Service, Worldwide,” Gartner, Inc., Stamford, CT, Tech. Rep. G00278620, Aug. 2016. [Online] Available: https://www.gartner.com/doc/reprints?id=1-2G2O5FC&ct=150519

Microsoft Application Architecture Guide, 2nd Ed., Microsoft Corp., Redmond, WA, 2009, pp. 03-*05. Available [Online]* https://msdn.microsoft.com/en-us/library/ff650706.aspx

Calder B. et al., “The Open Navigation Surface Project” *International Hydrographic Review*, vol. 6, no. 2, pp. 1–10, Aug. 2005.

J. Coleman, private communication, Feb 2017.

Calder, Brian, "Parallel and Distributed Performance of a Depth Estimation Algorithm" (2013). Center for Coastal and Ocean Mapping. Paper 858. http://scholars.unh.edu/ccom/858

Description of Bathymetric Attributed Grid Object (BAG), Version 1.6, Open Navigation Surface Working Group et al., Jun. 2016. Available [Online] http://www.opennavsurf.org/

D. Heimbigner and D. McLeod, “A Federated Architecture for Information Management” *ACM Transactions on Office Information Systems*, vol. 3, no. 3, pp. 254–255, Jul. 1985.

2016 Future of Cloud Computing Survey, North Bridge Growth Equity Venture Partners, Waltham, MA. http://www.northbridge.com/cloud-computing

M. Leppanen et al, “The Highways and Country Roads to Continuous Deployment,” *IEEE Software*, vol. 32, no. 2, pp. 64–72, Mar/Apr. 2015.

C. LeBlanc, A. Roy, “CHS Atlantic Multibeam Processing Guide”, Canadian Hydrographic Service, Atlantic Region, Feb. 2013. http://www.charts.gc.ca/documents/data-gestion/guidelines-directrices/sg-ld-2013-eng.pdf

Field Procedures Manual, National Oceanic and Atmospheric Administration, Office of Coast Survey, pp. 154-155, April 2010 https://www.nauticalcharts.noaa.gov/hsd/docs/Field_Procedures_Manual_April_2010.pdf

M. Wilson, G. Masetti and B. Calder, “NOAA QC Tools: Origin, Development, and Future,” presented at the Canadian Hydrographic Conference, Halifax, Nova Scotia, CA, May 16-19, 2016.

Calder, B. and S. Smith, “A Time/Effort Comparison of Automatic and Manual Bathymetric Processing in Real-Time Mode”, Proc. US Hydro Conf. (Biloxi, MS), 2003

Technology evolves…..

Image Credit: Todd Royer

"If you're doing anything interesting in the world, you're going to have critics. If you can't tolerate critics, don't do anything new or interesting."

Jeff Bezos