

[Cloud Computing Panels](#) [1]

Submitted by cyang3 on Sun, 2013-12-01 09:53 Thursday, January 9, 2014 - 15:30 to 17:00

Event: [Winter Meeting 2014](#) [2]

Session Type: [Panel](#) [3]

Collaboration Area: [Cloud Computing](#) [4]

[Decisions](#) [5]

[Discovery](#) [6]

[Earth Science Collaboratory](#) [7]

[Geospatial](#) [8]

[Information Technology and Interoperability](#) [9]

[Products and Services](#) [10]

Abstract/Agenda:

The cloud computing panel is organized to share the latest development of cloud computing from different organizations. This panel will discuss:

1. Latest available cloud services.
2. Challenges and solutions for Earth science and how cloud can help.
3. The new initiatives in cloud computing.
4. Cloud computing research, development and applications

Panelists:

- Deirdre Byrne (NOAA)
- Emily Law (ESDSWG and JPL)
- Kevin Murphy (EOSDIS)
- Vaughn Noga (EPA)
- Karen Petraska, NASA HQ

Notes:

4 panelists introduced themselves to the group. Panelists represent academic, government, public and private sectors.

Karen Petraska (NASA HQ), Deirdre Byrne (NOAA), David Mintz (EPA), Kevin Murphy (NASA/EOSDIS)

Ideas from the panel:

"Peering" with other cloud computing clusters

Examples of how to make data matter with cloud computing?

Dave:

- Have datasets in the cloud ready to be queried at any time
- How can EPA function better as regulatory agency using cloud computing?
- Regulations and security policies that need to be overcome/addressed to fully utilize cloud computing.

Deirdre:

- Very heterogeneous dataset (different instruments, etc)
- For climate study need long term data
- 5 Petabytes of satellite data alone on tapes. 50Pb 5 years from now
- Archival storage should be storage but needs to be accessible for case and climate studies

Cloud Computing Panels

Published on Commons (<http://commons.esipfed.org>)

Kevin:

- Cloud will improve data access abilities to foster completion of interdisciplinary goals
- Cloud will reduce replication of information
- Bring data streams together, even merging with other agency datasets
- Make data interoperable in the cloud so perhaps we can offset costs of data processing

Karen:

- Cloud will lead to new discoveries that wouldn't be possible with current computing power a single group/agency could buy
- Generate much publicly available data but hard for international folks to access
- Cloud computing could make US data available more easily abroad
- Difficult to run efficient data centers with changes in best practices
- Could make startup of data processing easier by making processing tools available in the cloud as well

What tools/techniques/etc will be embraced going forward?

Deirdre:

- Elastic computing (scalable)
- Object storage
- In-house object storage with increased response time to get away from caching
- Account monitoring and metrics

Dave:

- Easy access to data. Easy to explore relationships across multiple datasets
- Not cost effective to have a data server for every purpose
- Less downtime (patches, hardware, software upgrades)
- Allows for tech folks to refine processing instead of relying on a diverse array of contractors who often keep code in a 'black box'

Kevin

- Work together with cloud computing groups to make it easier to upload/harvest information
- Bring in industry partners to aid in putting data into the cloud
- ESIP's clusters could harness the cloud to accommodate testbed advanced visualization

Karen

- How to manage enterprise framework?
- Start with one (Amazon) initially
- Allow for more robust monitoring, tracking, billing, etc

Audience Qs:

1) What are trade space providers are looking at when looking to switch to a commercial cloud?

- Heavily prioritize not owning infrastructure: Ask 'can data go in a cloud?' 'Why/why not?'
- Commercial cloud: commercial cloud or commercial/private cloud
- Look at lifetime cost of physical hardware vs cloud. Found 6 years was cutoff
- Distributing data: Large cost if using commercial cloud to distribute large amounts of data every day
- Consider security of the data (Satellite data vs satellite command and control)
- Location of audience compared to location of data (in house or one to many)
- Stewardship and security
- Who owns the data
- User needs

2) When moving into clouds, can we get away from on-site computing capacity?

- Reliability: Amazon has gone down for hours before. Need to have a hybrid system (e.g. research (cloud) vs operations (on-site))
- COOP: multiple Amazon sites but with increased cost
- Comparing cost: on-site 3-5 year capital investments vs flat budget
- Hybrid infrastructure is very common for cloud applications
- People want a backup/alternatives. Multiple cloud providers?

Cloud Computing Panels

Published on Commons (<http://commons.esipfed.org>)

3) James from OpenDAP

Example: SnapChat infrastructure started in the cloud and has remained there. Netflix is completely in the cloud. There is a perceived fear of giving up physical hardware. 'What am I missing?'

- Cannot confuse business models of passing increasing cost onto customers with providing free access to data with cost fixed by gov't agency budget
- True cost of datacenter could be reduced through consolidation
- Specialized websites and data centers would be more difficult to transfer to cloud and thus not cost effective

4) Security and reliability. With clouds spread around, must have extreme bandwidth to transfer data between centers and TIC (trusted internet connections)

- Government has a brand to protect!
- Distribution requirements will fluctuate

5) Data in cloud is available to users to create higher level products. How do we get data into cloud and enable researchers to use the data once in the cloud?

- How to mount data centers into the cloud?
- Must work with agency partners and other agencies
- Different cloud structures and need to reformat data when uploaded into cloud. Gov't needs to work with industry
- Specify cloud computing requirements for projects in proposals
- Somebody needs to tell partner agencies to sit down and figure out this cloud computing problem

6) Avoid creating new stovepipes within cloud. Avoid reinventing and building new wheels from scratch. Need a forum for interagency collaboration on a technical level regarding cloud computing. Does one exist?

- ESIP is the perfect for this. Agencies and are represented and can take findings from ESIP working group to respective agencies.
- Don't want to create a new group but use the existing community

General comments

- Moving data to cloud is no good without a use case for it.
- Connect information management with data folks
- Need to think about secondary uses of data. Look for opportunities
- Maybe look at restructuring data or choosing what data is collected
- Need to get people that use clouds to come up with use cases for how they are done

Final comments: Looking Ahead

Karen: Start providing cloud services this year putting engineering and science services into the cloud. Within 5 years, hope 75% of new starts will be in cloud and 100% of public data in cloud and 40% of legacy systems into cloud. Legacy is very slow and expensive to transfer.

Deirdre: Extra bump of effort to change code to use the cloud. Want to be smart and do science better but who will do that engineering and development? How agencies are supposed to transfer to cloud but still keep up day-to-day missions. CIO office needs to plan for it.

Dave: Increase awareness that individual agencies are not the only ones tackling problems. Need interagency collaboration.

Kevin: How services can be improved using elastic computing capability. Extend work that has already been done e.g. Giovanni application, metadata search and order applications. Cost is reengineering code base for use in the cloud. Price is becoming competitive compared to 3-5 years ago. Maybe see prices continue to drop in 3-5 years to make it more affordable and enticing to transfer. Ultimately have data and applications in the same place.

Session Leads:

Name: [Rick Martin](#) [11]

Organization(s): [SAIC](#) [12]

Name: [Phil Yang](#) [13]

Cloud Computing Panels

Published on Commons (<http://commons.esipfed.org>)

Organization(s): [GMU](#) [14]

Name: [Thomas Huang](#) [15]

Organization(s): [SWEET](#) [16]

Email: thomas.huang@jpl.nasa.gov [17]

Presenters:

Name: [Kevin Murphy](#) [18]

Organization(s): [NASA](#) [19]

Name: [Karen Petraska](#) [20]

Organization(s): [NASA HQ](#) [21]

Name: [Deirdre Byrne](#) [22]

Organization(s): [NOAA](#) [23]

Name: [David Mintz](#) [24]

Organization(s): [EPA](#) [25]

Email: Mintz.David@epamail.epa.gov [26]

Notes takers:

Name: [Kyle Nelson](#) [27]

Organization(s): [University of Wisconsin
Madison](#) [28]

Email: wxkylenelson@gmail.com [29]

Creative Common License: Creative Commons Attribution 3.0 License

Accepted:

Source URL: <http://commons.esipfed.org/node/1863>

Links:

[1] <http://commons.esipfed.org/node/1863>

[2] <http://commons.esipfed.org/taxonomy/term/1029>

[3] <http://commons.esipfed.org/session-type/panels>

[4] <http://commons.esipfed.org/collaboration-area/cloud-computing>

[5] <http://commons.esipfed.org/collaboration-area/decisions>

[6] <http://commons.esipfed.org/collaboration-area/discovery>

[7] <http://commons.esipfed.org/collaboration-area/earth-science-collaboratory>

[8] <http://commons.esipfed.org/collaboration-area/geospatial>

[9] <http://commons.esipfed.org/collaboration-area/information-technology-and-interoperability>

[10] <http://commons.esipfed.org/collaboration-area/products-and-services>

[11] <http://commons.esipfed.org/node/348>

[12] <http://commons.esipfed.org/taxonomy/term/250>

[13] <http://commons.esipfed.org/node/345>

[14] <http://commons.esipfed.org/taxonomy/term/213>

[15] <http://commons.esipfed.org/node/1407>

[16] <http://commons.esipfed.org/taxonomy/term/770>

[17] <mailto:thomas.huang@jpl.nasa.gov>

[18] <http://commons.esipfed.org/node/1530>

[19] <http://commons.esipfed.org/taxonomy/term/228>

[20] <http://commons.esipfed.org/node/2055>

[21] <http://commons.esipfed.org/taxonomy/term/1248>

[22] <http://commons.esipfed.org/node/2056>

[23] <http://commons.esipfed.org/taxonomy/term/242>

[24] <http://commons.esipfed.org/node/635>

[25] <http://commons.esipfed.org/taxonomy/term/206>

[26] <mailto:Mintz.David@epamail.epa.gov>

Cloud Computing Panels

Published on Commons (<http://commons.esipfed.org>)

[27] <http://commons.esipfed.org/node/1936>

[28] <http://commons.esipfed.org/taxonomy/term/222>

[29] <mailto:wxkylenelson@gmail.com>