AIST Managed Cloud Environment

—a PI's Perspective

2016 Earth Science Technology Forum (ESTF2016) , Annapolis, MD

Wednesday, June 15, 2016

Hook Hua
Jet Propulsion Laboratory / California Institute of Technology

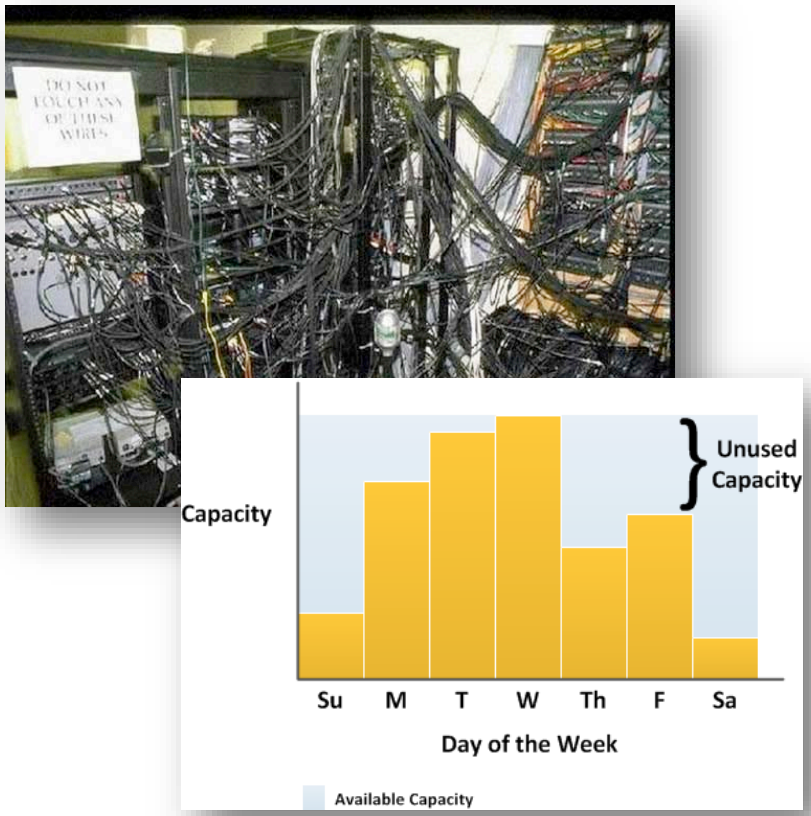
© 2016 California Institute of Technology. Government sponsorship acknowledged.

Reference herein to any specific commercial product, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement by the United States Government or the Jet Propulsion Laboratory, California Institute of Technology.

Benefits

- How the cloud offerings of AIST would specifically help support your research?
- How it can help AIST projects *get more computing resources beyond what they typically would get if they go with the traditional hardware procurement* route.

On-premise hardware procurements



Managed cloud approach

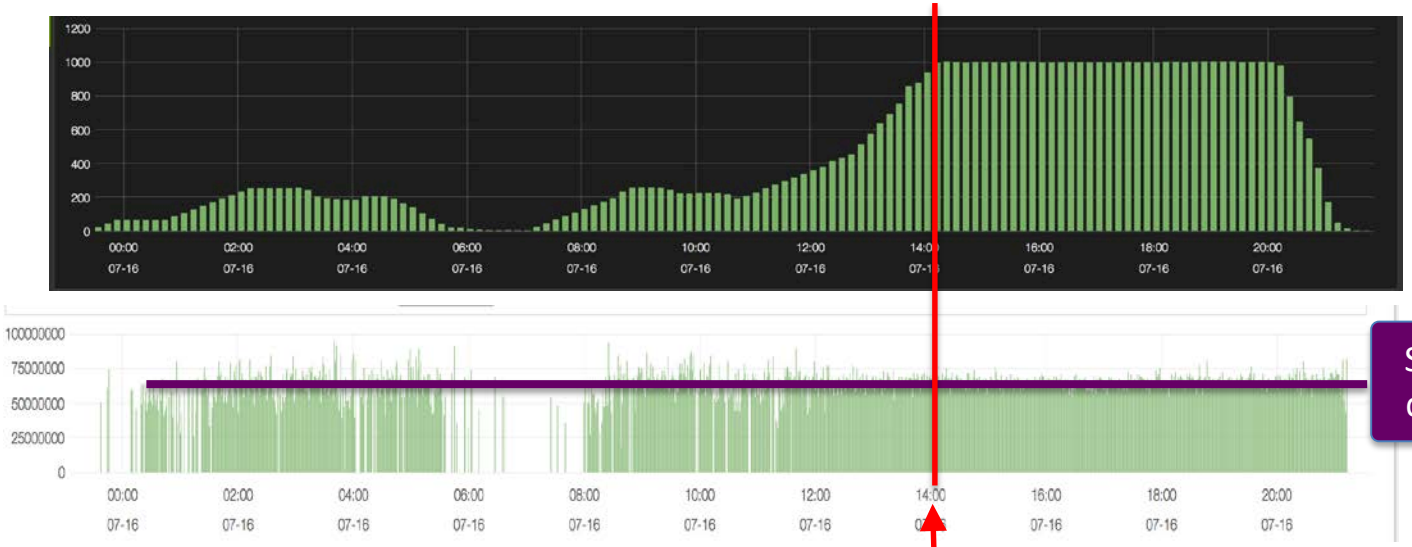
- Capital-expenditure Free
  - *"Pay as you go"*
- Agility and elasticity
  - Reacting to changing requirements.
- Virtualization
  - Management, deployability, reproducibility
- Scalable performance
  - High data volume and high data throughout
- Cloud data center efficiencies
- Resource management, including billing
- IT security compliance
  - Security updates, compliance, monitoring,

Example: OCO-2 Science Data System Adapted to AWS

- Science requirements change. Can science data systems react quickly enough to accommodate science demands?
- OCO-2 L2 Full Physics processing operational in AWS
  - Auto scaling of compute resources
  - Demonstrated capability of high data throughput rates and elastic compute/storage

compute nodes over time

Per node transfer rate over time



Scalable internal data throughput

@ 32,000 processors

Initial Lessons Learned

- What is the "user experience" in terms of any potential complexity?
- *"Trimming the fat"*
- Cost models
  - Costs differences among compute instance types and storage types
  - Cost differences between regions
  - Egress!
- Geographical locality
  - Where is compute and storage?
  - Moving data and processing.
- *Benchmark, benchmark, benchmark, ..*
  - "*Cost per unit*" yields non-obvious results

prefered minimun job duration (m) :	50

	test runs					prefered job size		on-demand (US-West-1)		spot (US-West-1)		on-demand (US-West-2)		spot (US-West-2)	
instance	soundings per job	concurrent processors	# test runs	mean runtime (s)	mean runtime per sounding per processor (m)	# soundings	# soundings (evenly distributed processors)	mean costs per job	$ per sounding	mean costs per job	$ per sounding	mean costs per job	$ per sounding	mean costs per job	$ per sounding
m2.4xlarge	32	8	42	1382	5.8	69	7	$0.4223	0.01320	$0.0250	0.00078	$0.4138	0.01293	$0.0384	0.00120
cc2.8xlarge	32	32	14365	615	10.3	156	16					$0.3417	0.01068	$0.0462	0.00144
m3.2xlarge															
c3.8xlarge	32	32	126	461	7.7	208	22	$0.2448	0.00765	$0.2151	0.00672	$0.2151	0.00672	$0.3073	0.00960
r3.8xlarge	32	32	131	489	8.2	196	22	$0.4238	0.01324	$0.0543	0.00170	$0.3803	0.01189	$0.3803	0.01189