

Towards Cloud Research Support

**Understand user requirements by
consultation and survey**

Nuyun "Nellie" Zhang, Ph.D.

Research Scientist

Team Lead for Software and Collaboration Support at PACE

Georgia Institute of Technology

nuyun@gatech.edu

Motivation

- ▶ Public clouds have great potential to advance research, but there are gaps that need to be addressed in order to make this happen
- ▶ Understanding the gaps to better supporting research in public cloud at higher education

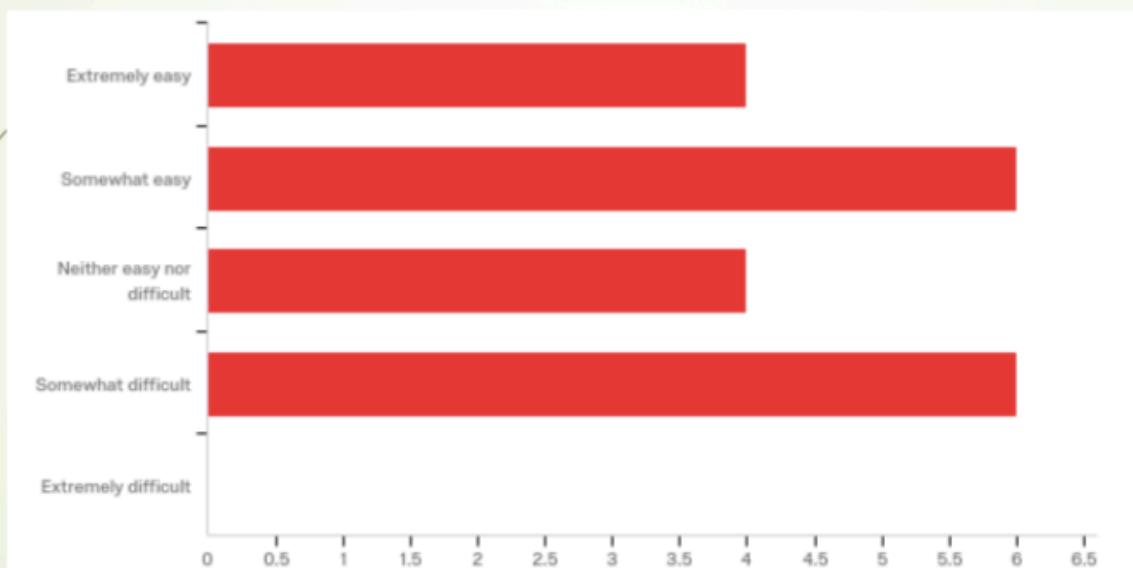
Challenge: Limited resources - IT staff

- ▶ Method
 - ▶ Survey by emails
 - ▶ Consultations
 - ▶ Meet with university/department level administrators

Email survey

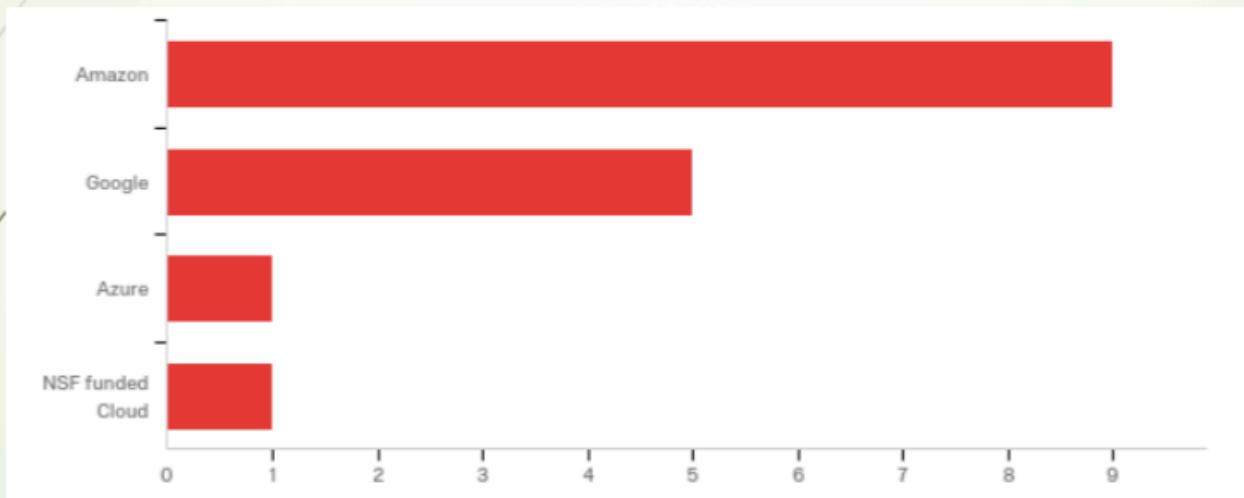
- Target: Georgia Institute of Technology - PACE cluster users
- 5 questions, anonymous, summer 2019
- 20 valid responses
- NSF grants: key word has "Cloud" - 18

1. How easy is it to obtain the resources that you need for doing research (computing, network, storage)

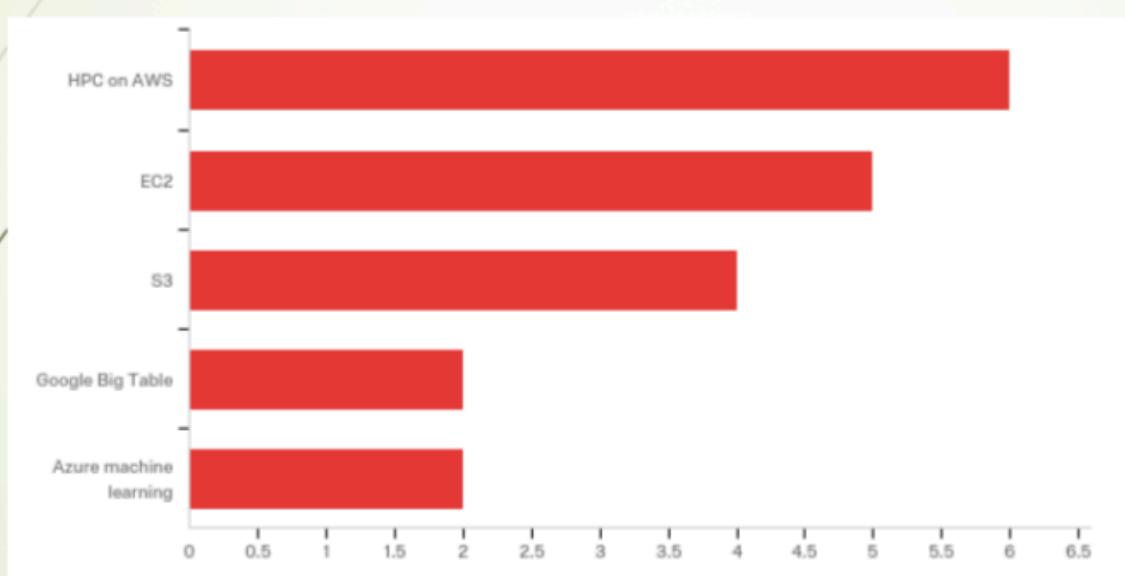


5

2. Have you used any Cloud resources for your research

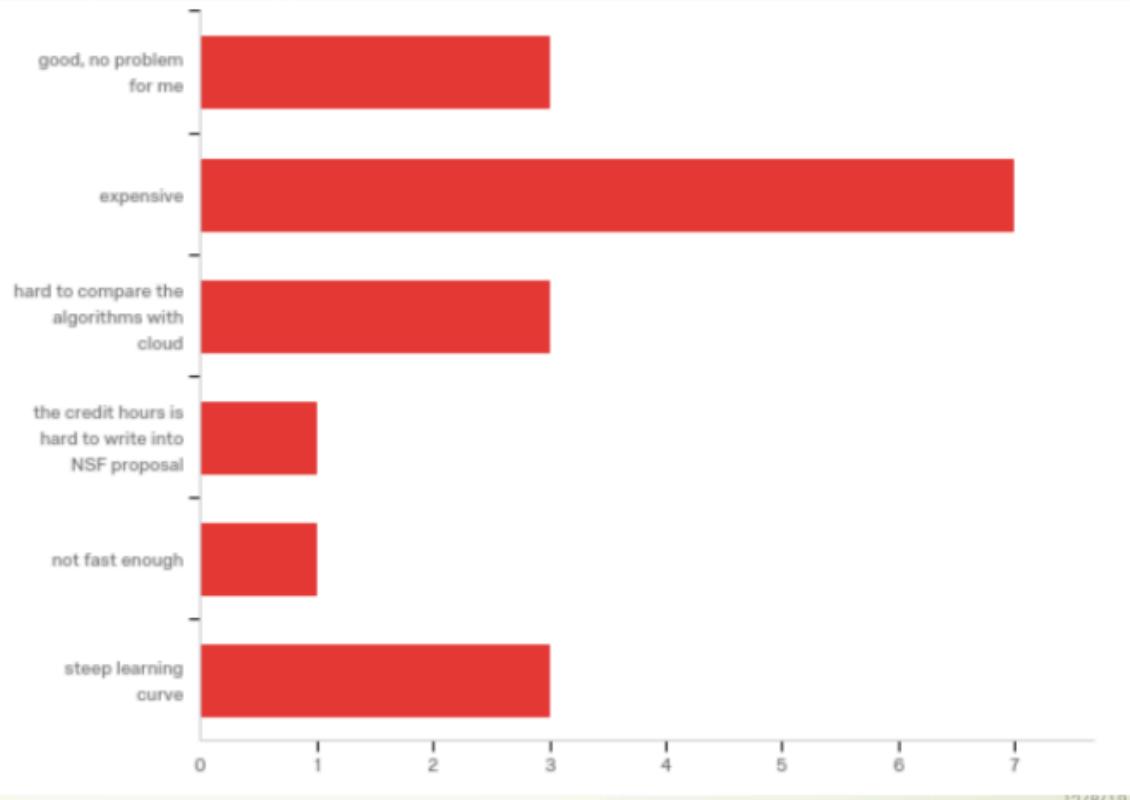


3. What specific services you have used?



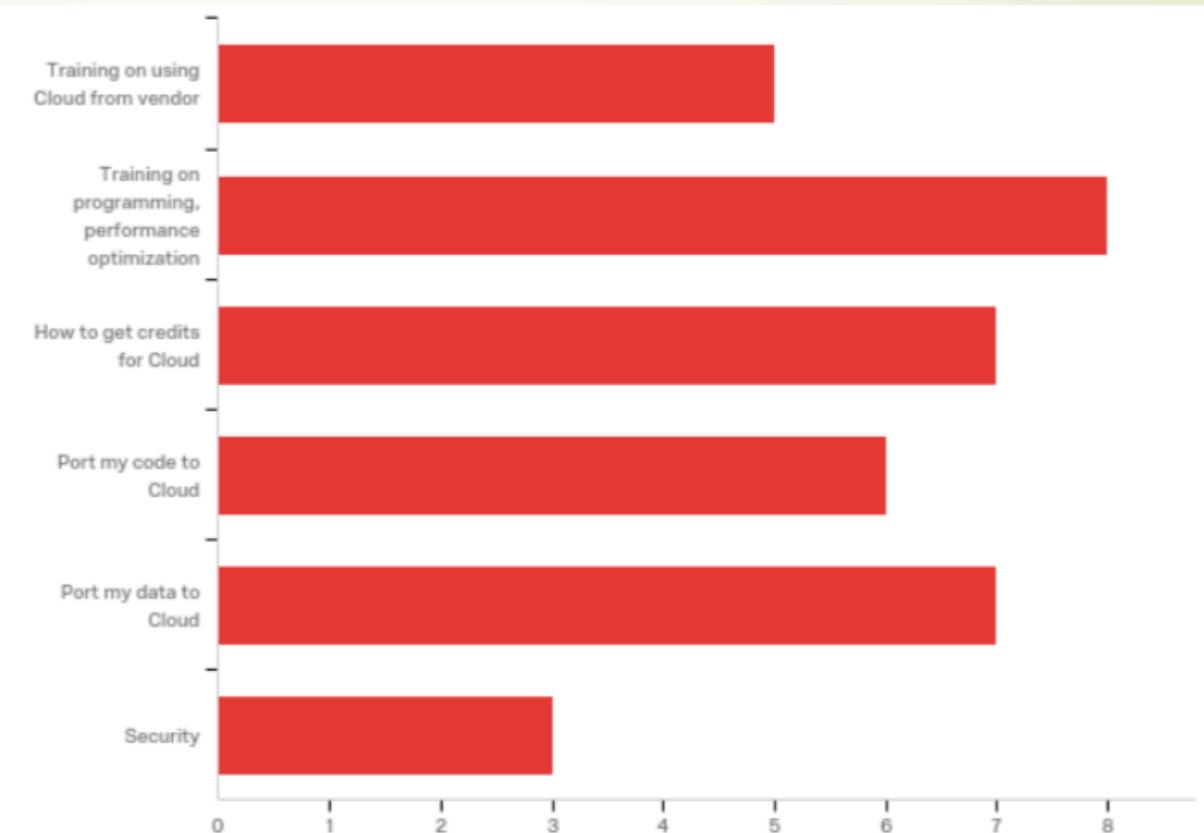
4. What are your experiences with them?

7



8

5. What support would you like to get to better your research with Cloud?



Consultations

- ▶ Biweekly HPC user consultations for 6 months
- ▶ Special consultations
 - ▶ NSF South Big Data Hub PI
 - ▶ ECE professor with 10 years experience of Cloud
 - ▶ Associate Vice Provost for Research Computing
 - ▶ CS professor
 - ▶ Physics professor

Gaps in Adopting Cloud for Research

- ▶ Users are not system admins
- ▶ Big data transfer is slow
- ▶ Real-time response is difficult
- ▶ Performance is unpredictable when model is trained on the Cloud
- ▶ If PIs buy physical machines, it is much easier to write the cost into proposals
- ▶ Bias in renting Cloud providers' services
- ▶ Change the source code of software in the Cloud is unnatural
- ▶ Cloud is like a black box
- ▶ Very steep learning curve in some cases
- ▶ Different platforms need different code porting workflow and extra code
- ▶ No dedicated IT staff to support Cloud researchers

What researchers like about cloud

- ▶ High-throughput computing (HTC) application
- ▶ Data sharing with remote collaborators
- ▶ Avoid restricted resources on on-premise clusters
- ▶ New experimental hardware
- ▶ Belief in advanced technology

Discussion and Future Work

- Cloud is good for domain scientists other than computer science or computer engineering researchers to use it
 - Optimal cost for getting a result for an experiment
 - Not a transparent and controllable environment
- Training/support from vendors are very important
 - IT usually does not have special roles for Cloud research support
- Future work
 - Consultations
 - Research computing team and other IT team work together
 - Knowledge transfer to students
 - Workshops in Cloud, code and data porting,
 - More projects experience to understand the procedure and best practices.

Thank you

Thank you Brian Voss, Craig Stewart,
IU PTI, GT PACE, Internet2, Microsoft

Questions: nuyun@gatech.edu