Cloud, Serverless Computing, HPC

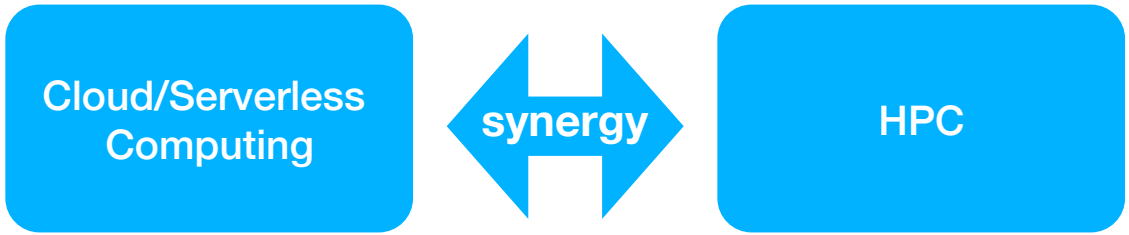
Kirk Cameron (VT) & Xipeng Shen (NCSU)

Participants (24+)

Andrew Chien
Anshul Gandhi
Barath Raghavan
Bo Wu
Eric Rozne
Jason Liu
Kirk Cameron
Kyle Chard
Lena Mashayekny
Liqiang Wang
Mattan Erez
Ming Zhao
Mor Harchol-Balter
Palden Lama
Peter Dinda
Richard Han
Satish Pwi
Timothy Zhu
Xian-He Sun
Xipeng Shen
Xuechen Zhang
Yong Chen
Zhiling Lan
Ziliang Zong

and more….

Focus



* Divergent design goals

Cloud/*Serverless*: Automate manageability (abstractions), focus on throughput
HPC: Control as much as possible, performance-centric; dependence among tasks; port and tweak legacy codes

* Commonalities

Large data needs; need for abstractions without sacrificing performance
Users want compute on demand at maximum throughput or turn around time
Aggressive adoption of new technologies

Major Challenges

- Increasing concerns on performance, resilience, security, reliability, & privacy [Cloud]
  - Abstractions, fine-grained function units, workload evolvement
- Guarantee of end-to-end latency and latency prediction dynamic scaling [Cloud]
- How to leverage emerging memory & increased architecture heterogeneity
- Data/app placement

Opportunities & Synergy

- For Cloud/Serverless Computing
  - More flexible scheduling (fine-grained functions)
  - Performance optimizations from HPC (parallel computing, locality, compilers & libraries, autotuning, performance modeling)
  - High speed connections and remote accesses
- For HPC
  - Resource management and scheduling techniques
  - Synergistic efforts on heterogeneity and data analytics
- Support HPC work on Cloud?

Suggestions

- More communications between the two communities
- Synergistic ecosystem for HPC and Cloud
- NSF programs in HPC and facilitate the development of this synergy
- Joint efforts between funding agencies and Cloud companies