Workload Limits on Cloud tolerance of Capacity Variation

Chaojie Zhang (Microsoft) and Rajini Wijayawardana, Siqi Li, Andrew A Chien (University of Chicago)
http://zccloud.cs.uchicago.edu/
March 28, 2023

Workshop on Scheduling Variable Capacity Resources for Sustainability
Paris, France
First, an 8-year story about Sustainable, Zero-carbon computing...
Example: MISO Wind Site – Stranded Power

- EcoMax (potential) vs Delivered [2015: Jan-April], 5 minute intervals
Stranded Power is Usable, Long Intervals

Best Single Site: MISO [YC2016]

- Temporal and spatial structure was largely unstudied
- Per-site opportunity study

Commercial Site: Lancium, 2020


- Exploit excess renewable power, compute when they are available
- Zero-marginal carbon power
- Help the Grid absorb more renewables

Zero-Carbon Cloud eliminates carbon footprint by exploiting zero-marginal carbon power. (scope 2).
Exploiting Excess power for HPC

- Trace: 1-year workload from Argonne LCF
- Compare – Performance (turnaround time), Total-cost of Ownership
ZCCloud Datacenters can be Cost-effective

- Savings in power and physical infrastructure cost (TCO) vs. Reduced duty factor
- For flexible workloads, can already be Cost effective. Even more so in regions with high power costs.

Computing as a Dispatchable Load (compliant with generation!)

2015 Research Idea (Zero-carbon Cloud)

- Compute with Stranded Power (curtailment, negative price)
- Zero-carbon - Useful, economic, low-reliability computing
- HPC Better TCO, Peak, ...
- Good for the Grid decarbonization

http://zccloud.cs.uchicago.edu/

2022 Commercial Reality
Lancium, Inc. + others

- 1.5GW of Dispatchable Load (>95% availability),
- Ex. 1GW off July 2022 (heat wave)
- Large Flexible Load Task Force (LFLTF), ERCOT
  - + other states

http://www.lancium.com/
Data Center Integration and Grid RPS

- How you integrate DC’s matters: +7% RPS (~18 TWh)
- Best: Dispatchable loads; Worst: DC+Windfarm

Outline

- Workload shifting: Why hasn’t it happened at Scale? What are the limits?
- Real Cloud Workloads
- Variable Capacity Study
- Why is load so inflexible?
- Observations and Discussion
Workload Shifting

• Compute Temporal Load Shifting
• Electrical Load shifting to when renewable energy is available, but also when there is less competitive load
• Other grids are More complicated
Real Cloud Workloads

- Microsoft Azure cloud VM workload
  - Low, core parallelism
  - Varied runtimes
  - ...
- Google BorgTNG Workload
  - GCU’s (normalized)
  - Varied Runtimes
  - ...

![Azure workload diagram](image)

![Borg workload diagram](image)
Grid-driven Capacity (best for Carbon)

- Common Power Grid Researcher Proposal
- Problem: Creates high rates of Capacity Change (Average Hourly Change)
Variable Capacity Study

- **Trace Driven Scheduling Study**
  - Cloud Workloads (Google BorgTNG, Microsoft Azure)
  - Synthetic Capacity Variation (Random walk, varying “range”)
    - Focus on dynamic range

- **Metrics**
  - Goodput (productive throughput)
  - Job Terminations
  - Job Latency (including restarts)

---

Zhang and Chien, “Scheduling Challenges for Variable Capacity Resources”, JSSPP, July 2021
Goodput drops sharply with Capacity Variation

- Azure 3% at 5%, and 40% at 60% Dynamic range
- BorgTNG 18% at 5%, and 60% at 60% Dynamic range
- Jobs are terminated, work lost, when capacity decreases
  - Dependences, Short/Long Job Fractions
Job failures increase dramatically with dynamic range of variation

Long jobs, and short (dependences)
Drilling Down – Job Disruptions

- At higher levels of variation, Long Jobs are heavily Disrupted
Long Job Heavy Workloads

- Cloud workloads dominated by long-running VMs
  - Azure and Google are similar (slight differences)
  - Amazon assumed to be similar
Observations

- Cloud workloads have limited ability to tolerate variation
  - Need ability to control, shape, and plan capacity (real costs)
    - Grid unlikely to be accepted as controller for this
- Economics of High Utilization limits Shifting opportunity
  - Ex. Borg () -> BorgTNG() -> ?

=> What are realistic limits on benefits for temporal shifting?

- Caveat: trace-based simulation methodology, doesn’t describe flexibility, uncertainty
Discussion

- New Resource Consumption Models in Workloads
  - Not long running VM’s
  - Serverless, Spot, Harvest/Transient VM’s -> none are designed for sustainability
  - => New compute resource models

- Other workloads? (eg. HPC)
  - Batch, re-shapeable, flexible resource requirements

- Cooperation Models
  - Two-way communication
  - Cooperation: Resource managers and Power Grid dispatch
  - Markets, Fairness, and Price/Profit

- Multinationals, Multiple Clouds?
Zero-Carbon Compute Project (http://zccloud.cs.uchicago.edu)

- Adaptive Datacenters, HPC & Cloud Computing, Economics-Opex/Capex
  - Chaojie Zhang and Andrew A Chien, "Scheduling Challenges for Variable Capacity Resources", Workshop on Job Scheduling for Parallel Processing (JSSPP), July 2021

- Helping the Grid (renewable absorption)

- Characterization of Growing "Stranded power", MISO, CAISO, ERCOT

- Thanks to NSF, VMWare, Google, Intel, Samsung, others for support.
BACKUP