Flat Budget and Hardware Costs

Chris Hollowell <hollowec@bnl.gov>
US-ATLAS Tier1
Scientific Data and Computing Center (SDCC)
Brookhaven National Laboratory

WLCG Workshop - 11/9/2022

@BrookhavenLab

Rising Hardware Costs

- Inflation, volatile demand, and continued supply chain issues have resulted in increasing hardware costs this year

https://www.ign.com/articles/intel-will-raise-the-price-of-its-cpus-in-2022

https://www.hardwaretimes.com/nvidia-amd-bend-to-tsmcs-price-hikes-affecting-the-rtx-40-ryzen-radeon-7000-chips-report/

Arista CEO details supply chain woes, mulls price hike

https://www.networkworld.com/article/3658997/arista-ceo-details-supply-chain-woes-mulls-price-hike.html

AMD is raising the price of its EPYC CPUs, but favorite customers escape the hike

https://www.techradar.com/news/amd-is-raising-the-price-of-its-epyc-cpus-but-favorite-customers-escape-the-hike

Brookhaven National Laboratory

Hardware Cost Increases at BNL

- We have experienced these cost increases firsthand
  - Network equipment costs among the worst affected
    - We use Arista network gear at BNL/SDCC, but this is not the only manufacturer affected
      - The US-ATLAS T2 (F. Luehring) manager reported significant (up to +27%) increases for some Juniper equipment cost as well
  - Recent price increases for BNL purchases (approximate):

Hardware	**Cost Change**
Network switches	+15% since winter 2022
HTC compute nodes	+10% since last year's purchase
A100 GPUs	+20% since fall 2021
Disk storage	+3% since last year's purchase
Tape storage	No significant change

Hardware Delays at BNL

- We have also experienced significant delays in the delivery of some hardware due to lingering supply chain issues
  - Again, network equipment among the worst affected
    - Arista ToR switches we purchase are now estimated to take up to a year for delivery
      - BNL is now purchasing switches well in advance of systems that will be connected to them
    - Infiniband components are estimated to be delayed 6 months
    - The US-ATLAS T2s reported 6-12 months for Juniper switch deliveries in some cases
  - Recent additional delays in equipment delivery time at BNL:

Hardware	**Delivery Time Change**
Network switches	+10 months
HTC compute nodes	+1 month
HPC compute nodes	+6 months (due to IB components)
Disk storage	+2 months
Tape storage	No significant change

Hardware Delays at BNL (Cont.)

- Why has networking equipment been so adversely affected by increases/delays?
  - Cost increases and delays are prevalent in almost all IT hardware, but networking seems particularly affected
    - High demand, lingering production effects from the pandemic
    - It appears semiconductor manufacturers are prioritizing the production of high-profit/complex silicon like new consumer CPUs/GPUs instead of the typically simpler/less-expensive ICs used in network equipment
      - Also explains why there continues to be a significant chip shortage in the auto industry

- In FY21/22, our primary compute node vendor, Supermicro, indicated there were worse supply chain issues for AMD than Intel
  - This was in part due to the fact that Intel uses its own fabs for CPUs, and thus has greater control of its supply chain
    - TSMC manufactures AMD silicon, and thus AMD must compete with Apple, Nvidia and others (including Intel for GPUs) using TSMC's fabs
  - One reason we chose to purchase Intel-based systems in FY21/22
  - Issue seems to be improving: a number of EPYC Milan models have become more readily available in recent months

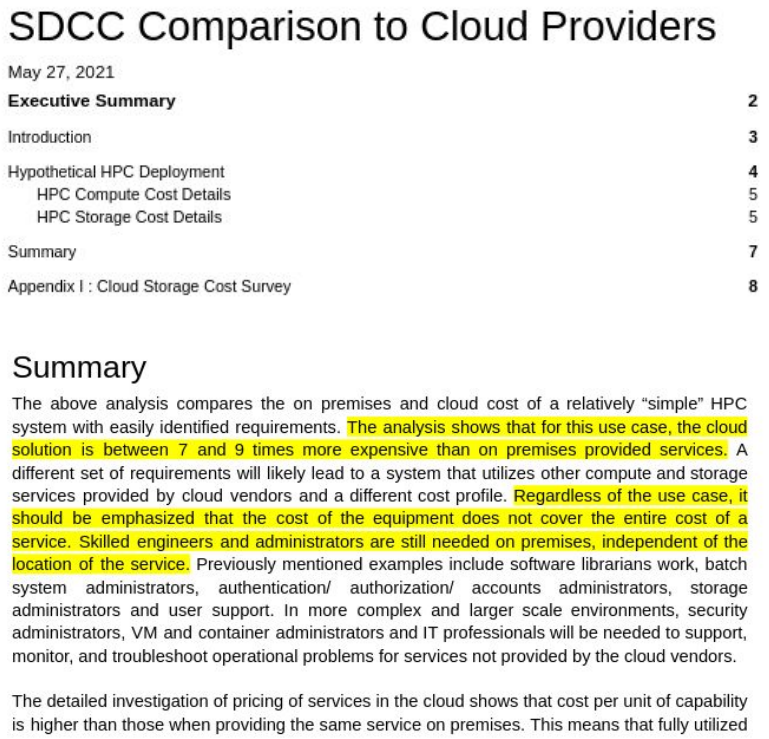
Is the Cloud an Answer to Rising Hardware Costs?

- At least for older instance types we've examined, it seems AWS pricing has not considerably increased in the past year
  - Could the cloud be a potential answer to the rising hardware costs we've seen?

- In 2021 BNL SDCC staff (A. Zaytsev, C. Hollowell, S. Misawa) developed datacenter and cloud computing cost models to allow comparison with the cost of on-premises equipment purchase/operation
  - Used in determining if a replacement for our IC (Institutional Cluster) HPC Cluster and storage should be purchased, or if cloud resources should be utilized instead
    - While primarily a comparison in the HPC context, HTC scenarios were also considered
  - Developing a cost model for on-premises, and cloud computing IT resources is a complex multi-variable problem

- ***The conclusion was that for a 100% utilized system, the cost of locating an equivalent HPC cluster and associated storage in the cloud would be 7-9 times more expensive over a 5 year period, even with considerably reduced DOE Google pricing***
  - **HTC compute was similarly ~10x the cost for 100% utilized resources**
  - **The cloud is still an attractive option for providing burstable resources to meet peak demand**

- It is important to note that cloud computing options do not significantly reduce staffing costs: high-level system/devops administrators/engineers are still required to run the system



SDCC Comparison to Cloud Providers

May 27, 2021

Summary

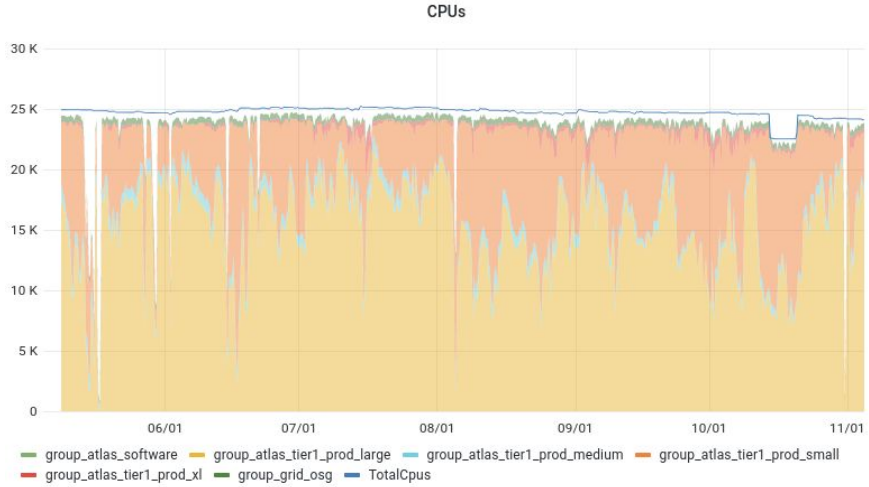
The above analysis compares the on premises and cloud cost of a relatively "simple" HPC system with easily identified requirements. The analysis shows that for this use case, the cloud solution is between 7 and 9 times more expensive than on premises provided services. A different set of requirements will likely lead to a system that utilizes other compute and storage services provided by cloud vendors and a different cost profile. Regardless of the use case, it should be emphasized that the cost of the equipment does not cover the entire cost of a service. Skilled engineers and administrators are still needed on premises, independent of the location of the service. Previously mentioned examples include software librarians work, batch system administrators, authentication/ authorization/ accounts administrators, storage administrators and user support. In more complex and larger scale environments, security administrators, VM and container administrators and IT professionals will be needed to support, monitor, and troubleshoot operational problems for services not provided by the cloud vendors.

The detailed investigation of pricing of services in the cloud shows that cost per unit of capability is higher than those when providing the same service on premises. This means that fully utilized

Brookhaven National Laboratory

Is the Cloud an Answer to Rising Hardware Costs? (Cont.)

- Given that ATLAS is ***very*** good at making near 100% use of its compute resources, allocating compute in the cloud does **not** appear to be a viable cost-saving mechanism in a flat-budget and increasing hardware cost scenario



- Can also be difficult to know exactly what you're getting from a performance perspective when you buy instances in the cloud
  - Descriptions of instance types often do not include an exact CPU model
    - Instead frequently indicate 2 or more models, or a family/generation of CPUs per instance type
      - CPU models in the same family can have very different performance characteristics and HS06/core ratings
    - Not having guarantees on exactly what CPU you're getting core allocations on can make cost/performance modeling and accounting difficult

- Egress charges can still be problematic
  - There are typically considerable charges for moving data out of the cloud
    - AWS: $0.08-0.12/GB
      - BNL dCache is serving 9 PB/month to the WAN for ATLAS: equivalent to ~$700k/month in egress if hosted at AWS
    - Although waivers have been made available in some cases for scientific research, particularly if direct network peering is established

Brookhaven National Laboratory

Alternatives In Flat Budget Scenarios

- Using US DOE leadership class HPC facilities to supplement computing shortfalls
  - Essentially free to use - requires winning a competitive allocation
  - But requires investment from the experiments in porting and optimizing their code to GPUs to make full use of the systems
    - Complicated by the fact that new exascale systems (Frontier/Aurora) are not using industry standard NVIDIA CUDA-compatible GPUs
      - Utilizing AMD and Intel GPUs instead
  - Potential issues with WAN connectivity for jobs at HPC centers, etc.

- Could alternative CPU architectures like ARM be more cost effective?
  - Again requires software engineering investment from the experiments to port their code
    - Work in this area has already been done in a number of cases

- One of the goals of the HEP Benchmarks Project and the HEPscore development effort is to be able to accurately measure and account for the performance of hardware accelerators and alternative CPU architectures

- Potential investment in code optimization?
  - Also requires additional significant software development effort from the experiments

Brookhaven National Laboratory

Conclusions

- Overall IT hardware costs have increased this year
  - Due to inflation, unpredictable demand, and lingering supply chain issues
  - Supply chain problems have also resulted in considerable equipment delivery delays
  - Several manufacturers have indicated they expect semiconductor supply chain issues will abate in 2023/2024
    - Unclear if costs will come down as issues resolve

- From cost modeling done by SDCC, it does not appear the cloud is a cost effective solution to supply steady compute or storage resources
  - HTC cloud compute costs ~10x more than local hardware if 100% utilized in a 5-year period
    - Egress charges may lead to additional cost
  - However, the cloud may be a viable option to meet short-term peak demands

- Some potential options in an increasing IT hardware cost and flat budget scenario include utilization of US Leadership Class HPC facilities, exploration of alternative architectures (ARM, hardware accelerators) and code optimization
  - All require significant software development effort