Financial Case Study:
Use of Cloud Resources in HEP Computing

Christopher Hollowell, Jerome Lauret, Shigeki Misawa, Tejas Rao,
Alexandr Zaytsev
Scientific Data and Computing Center (SDCC)

CHEP 2023
May 9, 2023

@BrookhavenLab

Motivation

Financial viability is of paramount importance when looking to site the dominant services on which NP/HEP depend in the cloud

- Cost of basic compute resources
- Cost of associated storage resources

For these services, technical viability of cloud resources is of little value if it isn't cost competitive with on premises resources

However, the cloud may be an alternative for other on premises services and may provide access to capabilities that might not be viable on premises.

Complexities

Multiple factors complicate the calculation of costs in the cloud

- Type of service
  - Compute/storage/database/data transfer/identity management/batch/interactive
  - Community specific vs widely adopted services
- Service level requirements
  - HA/guaranteed/on demand/"spot"
- Quantitative understanding of service resource utilization
  - *EVERYTHING* is metered with Cloud services
- Imprecise mapping of capabilities
  - Identical hardware or service may not be available.
- Migration strategy
  - "Lift and shift" or re-engineer
  - "Ecosystem" integration

Storage Is Complicated in the Cloud

A sample of the plethora of choices and metered metrics to consider when putting data in the cloud

- Multiple types of storage (Block/File/Object/Parallel/Tape)
- Capacity ($/GB), IOPS ($/IO/sec) and BW charges ($/GB/sec/TB)
- Minimum storage duration
- Multiple QoS tiers per type @ varying $/op
- Multiple classes of API calls @ varying $/op
- Data retrieval charges ($/GB) and $/request
- Inter zone/region transfer charges ($/GB)
- Egress fees (transfers out of cloud) ($/GB)

Core Use Cases : On Premises vs Cloud

Cost of four services are examined

- Compute
  - High Performance Computing w/GPU
  - High Throughput Computing
- Storage
  - Home directory file system - IOPS intensive
  - High performance “scratch” - Emphasizing bandwidth/capacity

These services represent the bulk of on premises resources and costs [1]. They are also mostly unencumbered by dependence on services metered by the cloud vendors.

[1] Home directory file system is not a significant cost for the facility, but it is necessary for use of compute resources.

Baseline Assumptions (On Premises)

The cost analysis assumes the following

- 100% resource utilization
- Large scale on premises operations:
  - Multi-MW data center
  - ~100 or more racks of compute and storage equipment
  - Economies of scale substantially reduce support costs per node
- Data center construction costs not included.
  - They are typically not borne by the scientific programs
  - Most sites already have a data center
- All costs include overhead, electricity, and cooling and are based on procurement experience at BNL, which may be different for other organizations.

Costs incurred migrating to the cloud aren't included.

Brookhaven National Laboratory

Baseline Assumptions (Cloud)

The cost analysis assumes the following

- Non-preemptable (on demand) cloud pricing is used as
  - On-premises resources provide this level of service
  - “Base load” demand must be satisfied
- Network egress fees are not included for cloud as they will only increase cloud costs and complicate the calculation
- Cloud pricing is assumed to be stable over a 5 year period
- Cloud pricing assumes single availability zone
- All cost include procurement overhead

Costs of repatriation and cloud to cloud migration are not included, but must be considered to protect against vendor lock in.

High Performance Computing (HPC)

On premises node specification

- 2 x Xeon 6336Y (48 physical cores total)
- 1 TiB RAM
- 1 x 4TB SSD
- 200 Gbps Infiniband network
- 4 x Nvidia A100
- <u>5 year life + power + cooling</u>

Amazon EC2 (p4de.24xlarge)

- 96 vCPUs (Xeon P-8275CL)
- 8 x Nvidia A100
- 1 TiB RAM
- 8 x 1TB NVMe
- 400 Gbps ENA+EFA network
- 5 years @ 3 year reserve instance rate

Node equivalence

- <u>2 on premises nodes = 1 EC2 node</u>

EC2 is ~4x more expensive [1]

- EC2 is > 2.5x more when using Amazon On premises Amazon Spot pricing, even after ignoring cost of eviction with Spot instances (as of 4/2023)

[1] Not including labor costs as they are small on a per node basis for large datacenters

High Throughput Computing (HTC)

On premises node specification

- 2 x Xeon 6336Y (96 logical cores total)
- 384 GiB RAM
- 4 x 2 TB SSD
- 10 Gbps Ethernet network
- 5 year life + power + cooling

Amazon EC2 (m6id.24xlarge)

- 96 vCPUs Xeon 8375C
- 384 GiB RAM
- 4 x 1.425 TB NVMe
- 37.5 Gbps network

Node equivalence

- 1.2 on premises nodes = 1 EC2 node (based on HEPSPEC06)

EC2 is > 6x more expensive [1]

- EC2 > 5x more expensive @ Spot pricing rates

[1] Not including labor costs as they are small on a per node basis for large datacenters

Storage: Home Directories

Required by HPC and HTC. Characterized by:

- Small block access
- High concurrency, looks like random access
- POSIX-like access

On premises specifics

- 12:1 compression ratio [1]
- Assume 20% active data [2]
- SSD based NAS appliance

Compare with Amazon FSx NetApp ONTAP

- 5 year life cycle
- Includes labor, power, and overhead

On premises is > 2x more expensive than cloud

- Cost savings due to tiering in FSx
- Costs are roughly equivalent if 100% of the data is active (no tiering)
- But total cost small relative to compute and "scratch" storage costs

Brookhaven National Laboratory

[1] Actual measured ratio at the SDCC
[2] Reference # used by Amazon in cost estimate examples

Storage: “Scratch” File Systems

HPC/HTC working storage. Characterized by :

- High capacity (multi petabytes)
- Higher bandwidth access
  - Larger block I/O
  - More sequential access
- HDD backed hardware
- Parallel file system software, e.g., GPFS or Lustre.

Lustre HDD file system on premises and in the cloud [1]

- 5 year life cycle
- Includes labor, power, and overhead

Cloud cost > 6x higher than on premises

[1] Amazon FSx for Lustre 12MB/s/TiB

Comments (On Premises)

Comments on resources sited on premises

- Sufficient economies of scale and critical mass are needed to support services on premises
  - At scale, the cost of hardware and physical infrastructure support, are small relative to hardware acquisition and support costs.
  - Common set of services used by multiple experiments provide critical mass needed to economically support the services.
- Multiple services needed by supported experiments to make compute resources usable still remain on premises.
  - Identity services, data distribution, workflow management, database services, etc.
- Support efforts outside of direct hardware support, e.g., user support and support of services remain, independent of the location of equipment
- Support for the “condo” computing allows researchers in experiments to leverage on premises economies of scale by placing resources purchased with outside funds in the data center

Comments (Cloud)

Cloud provider strategy isn't a lowest cost provider strategy

- Low or no entry cost primarily used to entice new customers
- Differentiation through innovative and broad service offering, not through lowest cost
- Customer retention through data access and egress fees and unique service offering
- Some services only possible at "hyper" scale, including bursting capacity and "exotic" hardware at scale.
- Business model requires understanding the cost of everything and having sophisticated pricing strategies to maximize profit

Cloud Opportunities

Potential benefits of the cloud include

- Ability to rapidly scale resources to handle burst loads
  - But a process is needed to determine which jobs are worthy of utilizing the limited funds allocated to bursting into the cloud
  - Also, Open Science Grid (OSG) and similar organization allow users to access resources at other sites, when local resources are fully utilized, reducing this advantage of the cloud
- Ability to test drive new hardware and software at minimal cost
- Cloud vendors can create new technologies that become de facto standard that can be leveraged by the community (e.g., Kubernetes, S3)

Brookhaven National Laboratory

Cloud Complications

Migration to cloud comes with new issues

- Consequences of under or over utilization of resources
- Managing/monitoring charges to prevent “surprise” bills
- Unfettered developer access to a broad range of services increases probability of increased dependency on proprietary, metered services
- Attention to service portability (to other cloud providers or repatriation) needs to be maintained to avoid vendor lock in
- Cost for data “in motion” and “at rest”, as well as charges for API calls to storage systems makes choosing the appropriate storage service and calculating costs fraught with peril
- Egress fees complicates backout strategies if large volumes of data are moved from on premises storage to cloud based storage

Summary

Cloud hosting is significantly more expensive than hosting on premises for just the major resources required by NP/HEP

Cloud based resources will incur additional expenses, beyond just compute and storage, most notably network egress fees if all jobs run on premises are moved to the cloud

Calculation of costs in the cloud is complicated and are highly dependent on the services being moved, how the services are moved, and the data center that is making the move.

Mistakes in the enumeration of requirements can result in significant increases in costs in the cloud

- Inadvertent use of metered services
- Underestimation of usage of metered services