HYBRID CLOUD ECONOMICS

EXECUTIVE SUMMARY

The emergence of public cloud computing has transformed enterprise IT and public cloud services continue to evolve and grow at a rapid pace. Public cloud offerings are rich and extensive, providing certain advantages to enterprise customers. However, public cloud service complexities make it difficult to select the optimal cloud service and difficult to compare the cost of public clouds with privately owned and operated data centers ("private clouds") or hybrid clouds (aka, "multicloud").

In this comprehensive study, our key findings are that applications with high data transfer requirements and/or GPU compute requirements are much more expensive to run on a public cloud versus a private cloud, and applications with less intensive compute requirements are more cost-effective in a public cloud. More specifically, the following types of use cases are most cost-effective running on the public cloud:

- Use cases with modest compute, memory, and network requirements are very cost-effective on the public cloud
- Use cases that require intermittent access to compute resources such that compute resources are idle a large percentage of time are very cost- effective on public cloud
- Use cases that run in batch mode and can be scheduled are very cost effective
- Use cases that use spot instances are most cost-effective

Report Highlights

- Public Cloud: AWS and Azure public virtual machine services.
- Private Cloud: Privately owned and operated datacenter with compute, storage, and network infrastructure.
- Hybrid Cloud: a combination of public cloud and private cloud services, aka, multicloud.
- Compute services with modest CPU requirements are more cost-effective on the public cloud.
- Network transport expenses are extremely high in public clouds as compared to private clouds: Applications with high data transport requirements are 8 to 15 times more expensive on the public cloud than private cloud.

Sponsored by JUNIPER NETWORKS

Conversely, the following types of use cases are more cost-effective running on a private cloud:

- Use cases with high network traffic (NFV, video, gaming, etc.) are much less expensive in a private cloud, a fact that is clearly driving edge cloud buildouts by service providers and other entities
- Use cases requiring GPU computing, such as AI/ML workloads, are much less expensive in a private cloud
- Use cases with consistently high compute requirements are generally less expensive in a private cloud

We also find that cost, scalability, and flexibility can be optimized with a hybrid cloud deployment, that is, running different workloads in different clouds. Depending on the applications and use cases, we have found up to a 91% cost savings in a hybrid cloud. Put another way, it is all about the right cloud for the job. As a result, it is critical that enterprises and web scalers use a unified software stack to facilitate workload portability between public and private clouds.

This paper provides an overview of public and private cloud economics and a detailed comparison of monthly expenses for nine different use cases and applications running on AWS, Azure, private, and hybrid clouds. While our economic model covers the majority of the CapEx and OpEx expenses in private data centers it does not model every possible configuration and operations activity carried out. The paper also excludes any volume discounts that may be extended to heavy users of public cloud services. Furthermore, the model does not consider the many new edge cloud business models that are emerging. This is a topic for a future study.

Sponsored by JUNIPER NETWORKS

Table of Contents

INTRODUCTION

Public cloud computing has driven dramatic changes in IT and enterprise digitization strategies over the last decade. Public cloud services are extensive, comprehensive, and continually growing. However, public cloud services are also extremely complex. Key benefits to cloud services include:

- Rich service offerings with many options for different use cases
- Customers are no longer responsible for building and maintaining hardware and software infrastructures to run applications
- Cloud services offer high availability and performance and geo-redundancy
- Cloud services are flexible and scalable

There are also downsides to public cloud offerings due to this complexity:

- Despite being thought of as simple, cloud services are designed for engineers and generally require engineering knowledge to configure, deploy, and consume
- Engineers are focused on the easiest way to solve problems and typically not focused on cost/benefits of alternative approaches
- It is exceedingly difficult to compare the relative cost of various public cloud alternatives to each other and more difficult to compare the cost of public cloud alternatives to private cloud
- The most complex problem is how to optimize hybrid cloud so that total cost is minimized while services, availability, and performance are maximized

The focus of this paper is a deep dive into the economics of public, private, and hybrid cloud virtual machine (VM) services. Although it is relatively easy to use available cloud calculators to estimate the price of a particular service, it is quite difficult to compare the cost of different applications and use cases between public and private clouds. This is the motivation behind our model which provides a detailed comparison of AWS, Azure, private cloud, and hybrid cloud expenses for a combination of use cases.

PUBLIC CLOUD VIRTUAL MACHINE OFFERING

Public cloud VM offerings are complex. There are many options that have great impact on the price of the service. In many cases it is not clear which options should be configured and how the cost of options compares with private or hybrid cloud alternatives. The following are some of the parameters that are used to specify VMs.

Type of VM

The most fundamental parameter that must be selected is the type of VM. Some of the key types are:

- General Purpose
- Compute Optimized
- Memory Optimized
- GPU Instances
- Storage Optimized

Each type of VM has multiple subtypes. For example, the AWS general purpose instances have multiple subinstances: a1, t3, t3a, t2, m6g, m5, m5a, m5ad, m5d, m5dn, m5n, m4. The high-level VM types are clearly defined. General purpose versus compute optimized are clearly targeted at different applications and use cases. However, it is not clear what the targets of the sub-instances are. For example, when should one use m5 versus m5ad? Both instances have similar configurations for vCPU and memory, but they have different price points. Choosing the correct sub-instances of a VM is not trivial and the impact on expenses can be significant.

vCPU and Memory Requirements

Regardless of the type of VM, the user must specify the vCPU and memory requirements. As applications scale up, VMs will require greater vCPU and memory resources.

VM Usage

Pricing of VMs is highly dependent on the requirements for VM usage. The key types of VM usage are:

- On-Demand Instances
- Reserved Instances
- Spot Instances

If VMs are needed at unpredictable times then on-demand instances must be used. VMs will spin up whenever they are needed and might run 24X7. If VMs are required at specific times (for example, batch processes that might run at night) or if VMs must run all the time, then reserved instances can be used. Reserved instances typically have about a 38% discount from on-demand instances. Alternatively, if applications can run at flexible times when general public cloud demand is low then spot instances can be used, which can save 90% of the VM expense.

Region

There are many different regions that also effect the service pricing. Regions need to be specified for both compute instances and data transfer pricing. Regions can be important when it comes to latency-sensitive services, compliance, and data sovereignty issues.

Operating System

The operating system (OS) must also be specified, and it also has an impact on service pricing. Examples of alternatives are CentOS Linux, RHEL, SLES, and windows. The lowest cost versions are usually CentOS, which is a free OS. Microsoft Azure also has low pricing for Windows because it is a Microsoft product.

Data Transfer Requirements

Cloud services charge for data transfer, and the pricing is dynamic based on the number of GBytes of data transferred in a month. For applications with large data transfer requirements this cost is significant and must be estimated in advance. It is possible to negotiate pricing for applications that require high levels of network traffic.

Term of Contract

There are multiple options for terms of contracts that include one year, three year and payment options, which cover all upfront, partial upfront or no upfront. These contract terms also have impact on service pricing.

PRIVATE CLOUD ARCHITECTURE AND ASSUMPTIONS

To compare public and private cloud expenses we developed a private cloud model (**Error! Reference source not found.**). The private cloud model includes both hardware and software required to provide VMs to end users. The key hardware components considered in the private cloud are:

- General Purpose Servers
- GPU optimized Servers
- Leaf Switches
- Spine Switches
- Data Center Interconnect (DCI) Routers

The cloud software components are:

- OpenStack
- Service Orchestration Software

Servers are connected to redundant leaf switches using 10GE connections. Leaf switches are connected to spine switches using redundant 100GE connections. All traffic to and from the internet moves through redundant DCI routers using multiple 100GE connections. Virtual machines are provided by OpenStack, and end-to-end service orchestration is provided by the orchestration software.

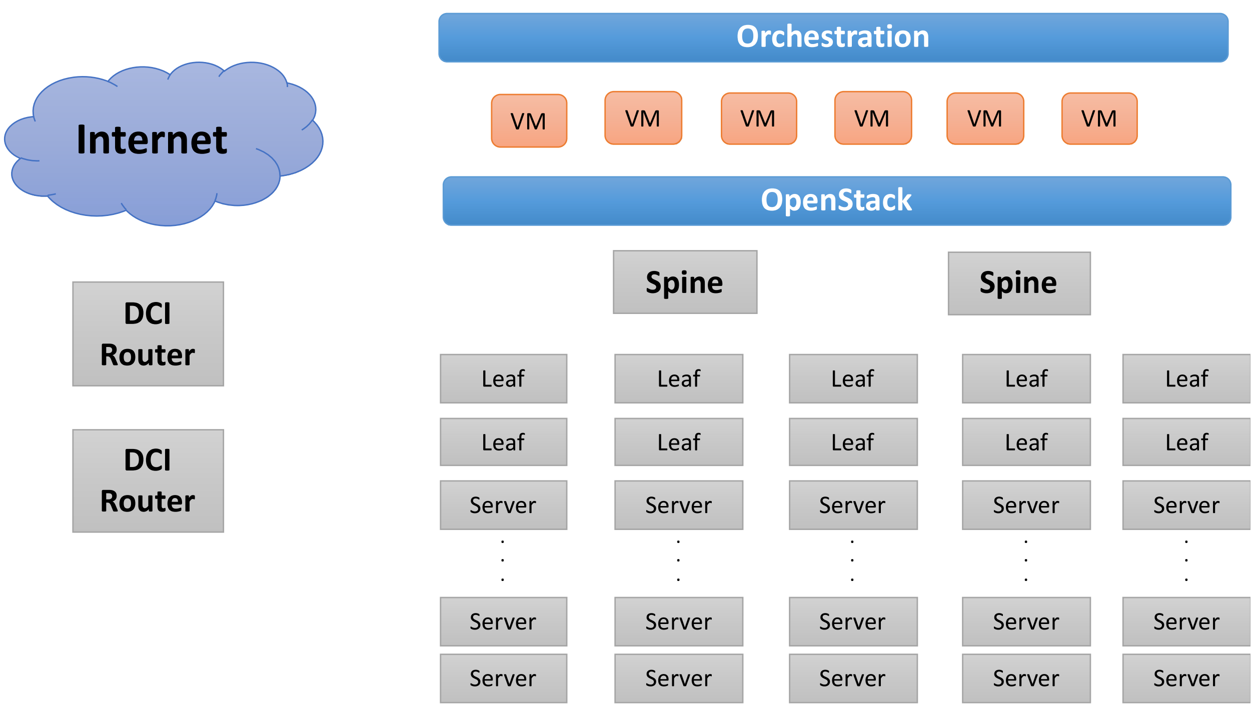

Figure 1. Traffic to and from the Internet

The key assumptions for capital expense (capex) are specified in **Error! Reference source not found.** and operations expense (opex) assumptions are in **Error! Reference source not found.**. Capex expenses are one-time expenses for purchasing hardware. Opex includes a combination of annual vendor support expenses, annual software subscriptions, facilities, power, and labor expenses. Some of the general assumptions for private cloud are presented in **Error! Reference source not found.**.

Capex Assumptions	Cost
General Purpose Server	$15,000
GPU Optimized Server	$25,000
Spine Switches per Server	$220
Racks & ToRs per Server	$1,060

Table 1. Key Assumptions for Capex

Opex Assumptions	Cost	Description
Server Vendor Support	$ 750	Annual Vendor Support Contract
GPU Server Vendor Support	$1,250	Annual Vendor Support Contract
Spine Switch Vendor Support	$11	Annual Vendor Support Contract
Racks & ToR Vendor Support	$53	Annual Vendor Support Contract
OpenStack per Server	$ 800	Annual OpenStack Subscription
Data Center SDN per Server	$500	Annual SDN Software Subscription
Service Orchestration per Server	$4,000	Annual Orchestration Software Subscription
Facilities Expense per Server	$432	Cost of All Facilities: Real Estate, HVAC, & Management
Labor Expense per Server	$75	IT Labor Expenses for Managing Networks & Servers
Power Expense per Server	$25	Annual Power Expense per Server
Cooling Expense per Server	$18	Annual Cooling Expense per Server

Table 2. Key Opex Assumptions

General Assumptions	Values
vCPU per Core	2
Cores per Server	44
Power per Server (Watts)	500
Power Expense (Kwatthr)	$0.07
Cooling as Percentage of Power	70%
Average Server Utilization	70%

Table 3. General Assumptions for Private Cloud

CLOUD ECONOMICS FOR DIFFERENT USE CASES

Our modeling and analysis has found that there are major differences in public and private cloud expenses for different use cases. In general, the following types of use cases are most cost-effective running on the public cloud:

- Use cases with modest compute, memory, and network requirements are very cost-effective on the public cloud
- Use cases that require intermittent access to compute resources such that compute resources are idle a large percentage of time are very cost-effective on public cloud
- Use cases that run in batch mode and can be scheduled are very cost-effective
- Use cases that use spot instances are most cost-effective

Alternatively, the following types of use cases are more cost-effective running on a private cloud:

- Use cases with high network traffic (NFV, video, gaming, etc.) are much less expensive in a private cloud
- Use cases requiring GPU computing are much less expensive in a private cloud
- Use cases with consistently high compute requirements are generally less expensive in a private cloud

Some enterprises use public cloud for development and testing of applications because of the better agility of public clouds and then run the applications on a private cloud production environment to retain more control. The reason for this is that the development use case is very efficient to run on the public cloud; the production use case is too costly to run on the public cloud. For this reason it is critical that enterprise and web scalers use a unified software stack that allows portability between public and private clouds to provide this flexibility.

To analyze cost differences between public and private cloud we developed a detailed model that compares the cost of AWS, Azure, and private cloud for many types of use cases. Use cases can be defined by the end user of the model. For the purposes of this paper we compared a set of nine use cases, which are defined in **Error! Reference source not found.**. All comparisons were in US regions, used reserved instances (which are less expensive than on-demand instances), and used one-year contracts. For each use case we assumed that 100 VMs of the type defined were running. This allows an adequate comparison of applications running at scale between a public and private cloud deployment. The private cloud expenses used the model and assumptions defined in the previous section. This model compares the cost of public and private clouds over one year. It should be noted that over a longer period (3−5 years) the cost savings for some of the use cases will be more significant.

Use Case	Description	VM Type	vCPU/Memory	Data Xfer
Cost-Effective Compute	Scale-out workloads such as web servers, containerized microservices, caching fleets, & distributed data stores, & development environments	General Purpose	2 vCPU & 4GB	0.02 TB/Month
Web Sites & Apps	Web sites & web applications hosting general-purpose web applications and low-priority business applications	General Purpose	2 vCPU & 8GB	0.02 TB/Month
Mid-size Data Processing	Small & mid-size databases, data processing tasks that require additional memory, caching fleets, & for running	General Purpose	8 vCPU & 32GB	0.1 TB/Month

	backend servers for SAP, Microsoft SharePoint, cluster computing, & other enterprise applications			
Windows App Servers	Microsoft Windows application & database servers. Data processing tasks requirement memory, caching, & medium level performance	General Purpose	8 vCPU & 32GB	0.1 TB/Month
Multiplayer Gaming	High-performance multiplayer gaming & video encoding, compute optimized VM	Compute Optimized	36vCPU & 72GB	250 TB/Month
High-Performance Batch	Compute optimized high-performance batch processing for scientific modeling, ML training, data analytics, & other compute intensive applications	Compute Optimized	36vCPU & 72GB	0.5 TB/Month
Network VNF	Network VNF such as a virtual router, firewall or packet core. Compute intensive and network traffic intensive.	Compute Optimized	36vCPU & 72GB	4000 TB/Month
AI/ML	Machine learning, high-performance data analytics, scientific computing	GPU Instances	36vCPU & 488GB	0.5 TB/Month
High performance DB	Memory-intensive applications such as high- performance databases, distributed web scale in-memory caches, mid-size in-memory databases, real-time big data analytics, & other enterprise applications	Memory Optimized	96vCPU & 768GB	0.5 TB/Month

Table 4. Use Cases

The results of the analysis are summarized in **Error! Reference source not found.**[1], which shows the total monthly expenses for AWS, Azure, and our private cloud solution previously defined. For the private cloud solution we calculate monthly expenses using depreciated capex and annual opex. The table is color coded to show where public cloud is more expensive and less expensive than private cloud. Red indicates that public cloud is significantly more expensive, yellow indicates public cloud is slightly more expensive, and green indicates public cloud is less expensive than private cloud.

There are two components of public cloud expenses:

- CPU expenses
- Network transport expenses

[1] Very large customers can negotiate significant discounts on public cloud pricing. In this case the costs of the public cloud solutions would be lower than the costs calculated with the publicly available pricing.

These expenses are displayed in **Error! Reference source not found.** and **Error! Reference source not found.**. It is clear that network transport expenses are significantly higher in the public cloud versus the private cloud. Therefore, any use case that has a significant amount of network traffic is much more expensive on the public cloud versus the private cloud (**Error! Reference source not found.**). These high costs are certainly driving the buildouts of edge clouds by service providers and other entities. It is important to keep in mind that heavy users likely negotiate volume discounts with public cloud providers to influence the economic decision. It is also true that GPU instances are more expensive on the public cloud than the private cloud (**Error! Reference source not found.**) in the AI/ML use case. Another more obvious result is that users that need a Windows OS are better off using Azure. The differences are the percentage difference between AWS, Azure, and the private cloud. For example, for cost-effective compute AWS is 61% of the cost of private cloud, whereas for network VNF AWS is 1580% (or 15 times) the cost of private cloud.

	Total Monthly Expense			Public vs Private Difference	
Applications	**AWS**	**AZURE**	**Private Cloud**	**AWS**	**AZURE**
Cost Effective Compute	$ 2,447	$ 2,254	$ 4,034	61%	56%
Web Sites & Apps	$ 4,313	$ 5,638	$ 4,034	107%	140%
Mid size Data Processing	$ 18,556	$ 17,174	$ 13,116	141%	131%
Windows App Servers	$ 34,984	$ 17,174	$ 13,116	267%	131%
Multiplayer Gaming	$ 818,299	$ 837,005	$ 103,501	791%	809%
High Performance Batch	$ 16,928	$ 18,901	$ 59,533	28%	32%
Network VNF	$ 12,077,138	$ 12,087,005	$ 764,347	1580%	1581%
AI/ML	$ 322,908	$ 619,015	$ 69,091	467%	896%
High performance DB	$ 243,627	$ 295,325	$ 157,265	155%	188%

Table 5. Total Monthly Expenses

	CPU Monthly Expense			Public vs Private Difference	
Applications	**AWS**	**AZURE**	**Private Cloud**	**AWS**	**AZURE**
Cost Effective Compute	$ 2,277	$ 2,088	$ 4,030	56%	52%
Web Sites & Apps	$ 4,143	$ 5,472	$ 4,030	103%	136%
Mid size Data Processing	$ 17,856	$ 16,474	$ 13,098	136%	126%
Windows App Servers	$ 34,284	$ 16,474	$ 13,098	262%	126%
Multiplayer Gaming	$ 68,299	$ 87,005	$ 59,445	115%	146%
High Performance Batch	$ 15,428	$ 17,401	$ 59,445	26%	29%
Network VNF	$ 77,138	$ 87,005	$ 59,445	130%	146%
AI/ML	$ 321,408	$ 617,515	$ 69,003	466%	895%
High performance DB	$ 242,127	$ 293,825	$ 157,177	154%	187%

Table 6. Monthly CPU expenses

	Network Transport Monthly Expense			Public vs Private Difference	
Applications	**AWS**	**AZURE**	**Private Cloud**	**AWS**	**AZURE**
Cost Effective Compute	$ 170	$ 166	$ 4	4823%	4710%
Web Sites & Apps	$ 170	$ 166	$ 4	4823%	4710%
Mid size Data Processing	$ 700	$ 700	$ 18	3972%	3972%
Windows App Servers	$ 700	$ 700	$ 18	3972%	3972%
Multiplayer Gaming	$ 750,000	$ 750,000	$ 44,056	1702%	1702%
High Performance Batch	$ 1,500	$ 1,500	$ 88	1702%	1702%
Network VNF	$ 12,000,000	$ 12,000,000	$ 704,902	1702%	1702%
AI/ML	$ 1,500	$ 1,500	$ 88	1702%	1702%
High performance DB	$ 1,500	$ 1,500	$ 88	1702%	1702%

Table 7. Monthly network transport expenses

Hybrid Cloud

Although there are many factors driving the adoption of hybrid cloud, a key consideration should be the cost of running applications in the public cloud versus a private cloud. To compare public, private, and hybrid cloud we consider three scenarios:

1. All applications run on AWS public cloud
2. All applications run on private cloud
3. Some applications run on AWS and others run on the private cloud

For this example we have chosen to run the applications in the AWS red zone to the private cloud. Specifically, the applications that we run on the private cloud are:

- Windows App Servers
- Multiplayer Gaming
- Network VNF
- AI/ML

All other applications run on AWS. The results of this analysis are presented in Table 1. Both the private and hybrid cloud solutions are significantly less expensive than the AWS public cloud service. The biggest factors are the network transport expenses. The analysis shows the private cloud is slightly less expensive than the hybrid cloud, but the difference in cost is insignificant. The benefits of public cloud flexibility, availability, and scalability outweigh the small difference in expense. However, a purely public cloud solution is significantly more expensive. This example shows the financial benefits of an optimized hybrid cloud solution.

Monthly Expenses	Public (AWS)	Private	Hybrid	Hybrid Savings
Total	$13,554,122	$1,192,155	$1,249,842	91%
CPU	$796,382	$442,801	$495,237	38%
Network Transport	$12,757,740	$749,354	754,605	94%

Table 1. Applications that Run on AWS

CONCLUSION

Public, private, and hybrid clouds have received a great deal of attention over the last decade, but there has been little analysis of the relative expenses of these solutions. This paper demonstrates that use cases with high network transport requirements and GPU processing requirements are significantly more expensive to run on a public cloud than a private cloud. Our recommendation is to build hybrid clouds such that some applications run off the public cloud while others run in a private data center. Depending on the applications and use cases, we have found up to a 91% cost savings in a hybrid cloud.

This paper and the economic model developed by ACG was sponsored by Juniper. The model is extremely comprehensive with a large number of variables and use cases that can be configured to compare many different scenarios. This model can be used by cloud providers and enterprises to make decisions about workload placement in a hybrid cloud environment. Customers interested in using the model should contact their Juniper sales representative to request a consultation.

Peter Fetterolf, Ph. D (pfetterolf@acgcc.com) is CTO with ACG Research. His primary focus is developing business models for next generation networks which includes IP transport, SDN, NFV, vEPC, vRAN, and optical transport networks. He is also responsible for software development of the Business Analytics Engine (BAE) software network economic simulation tool.

ACG Research delivers information and communication technology market share/forecast reports, consulting services, and business case analysis services. Copyright © 2020 ACG Research. The copyright in this publication or the material on this website (including without limitation the text, computer code, artwork, photographs, images, music, audio material, video material and audio-visual material on this website) is owned by ACG Research. All Rights Reserved.