Understanding the Cost of Computing in the Cloud: Comparing the cost of 3 different Cloud Configurations in Private vs Public environments !

Sandeep Vuzzini

“Private Cloud Configuration I”

As per the given requirement in Configuration I of the project, Hadoop/Spark Cluster must meet below specs:

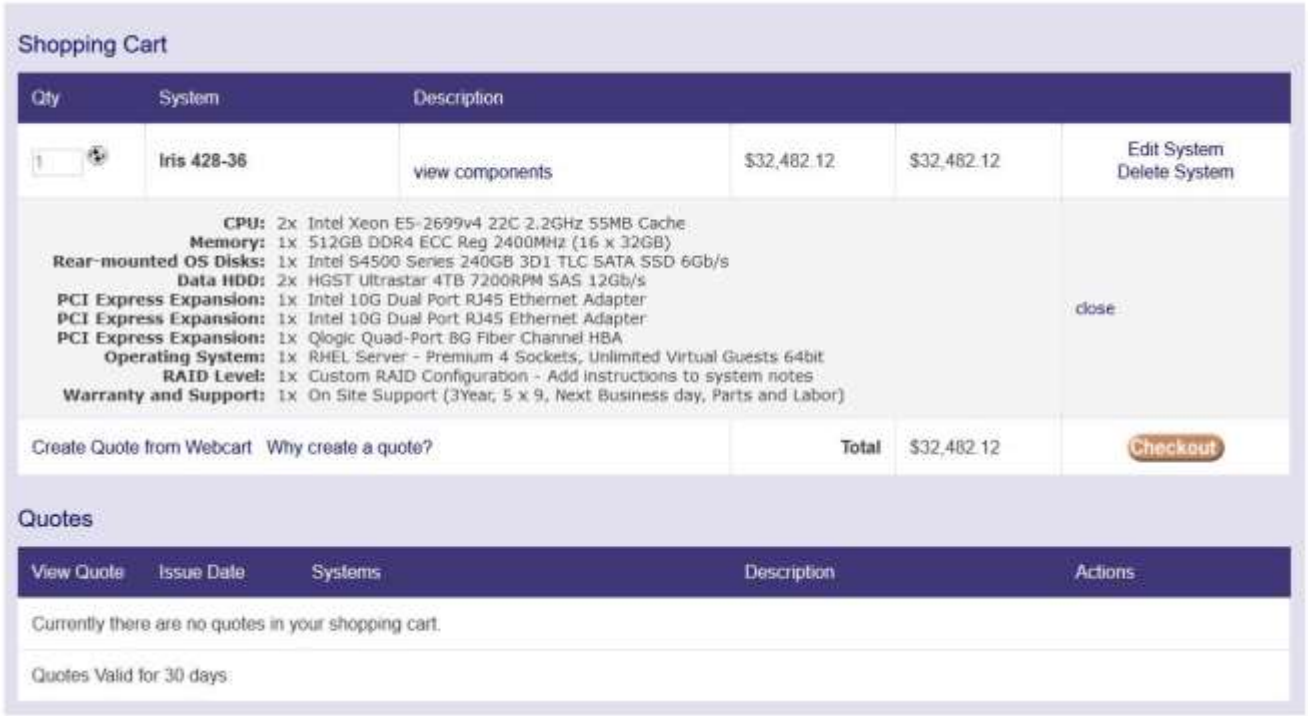
1) 32,000 cores
2) 256 TB Memory
3) 50 PB HDD
4) 10 Gb/s Ethernet Fat-Tree network
5) 100 PB additional storage distributed across the servers with enough capacity of 100 GB/sec throughput
6) Each VM must be equivalent to d2.8xlarge()

We have chosen the below **compute server** for configuration 1:

Server Model and Specs:

Model: Iris 428-36

Specs:



Shopping Cart

Qty	System	Description			
1	Iris 428-36	view components	$32,482.12	$32,482.12	Edit System Delete System
	CPU: 2x Intel Xeon E5-2699v4 22C 2.2GHz 55MB Cache **Memory:** 1x 512GB DDR4 ECC Reg 2400MHz (16 x 32GB) **Rear-mounted OS Disks:** 1x Intel S4500 Series 240GB 3D1 TLC SATA SSD 6Gb/s **Data HDD:** 2x HGST Ultrastar 4TB 7200RPM SAS 12Gb/s **PCI Express Expansion:** 1x Intel 10G Dual Port RJ45 Ethernet Adapter **PCI Express Expansion:** 1x Intel 10G Dual Port RJ45 Ethernet Adapter **PCI Express Expansion:** 1x Qlogic Quad-Port 8G Fiber Channel HBA **Operating System:** 1x RHEL Server - Premium 4 Sockets, Unlimited Virtual Guests 64bit **RAID Level:** 1x Custom RAID Configuration - Add instructions to system notes **Warranty and Support:** 1x On Site Support (3Year, 5 x 9, Next Business day, Parts and Labor)				close
Create Quote from Webcart Why create a quote?			**Total**	$32,482.12	Checkout

Quotes

View Quote	Issue Date	Systems	Description	Actions
Currently there are no quotes in your shopping cart.				
Quotes Valid for 30 days				

To meet the number of cores requirement and memory requirement on the entire Hadoop/Spark cluster, we need 728 U of the similarly configured servers (specs mentioned above for each Unit)

Configuration 1 table for Private Cloud

Configuration 1	Description	Price per Item	Quantity	Total Price	Notes
Computer Servers	Iris 428-36 (4U Rack Mount Servers)	$32,482	728	$23,646,983	
Network Switches	Cisco Catalyst 4948E-F - Switch - L3 - managed - 48 x 10/100/1000 + 4 x 10 Gigabit SFP+ - rack-mountable	$13,920	34	$473,280	Our Datacenter used Top of the Rack switches, we would procure two network switches for each rack (each switch has 48 ports). Purpose of two switches per rack is used to support Fat tree config of 23 upstream ports and 23 downstream ports in addition to input and output ports.
Network Cables	Cat 7 Ethernet Cable Flat 10 FT- Ealona Shielded (SSTP) Lan Patch Cable 10 Gigabit 600Mhz Gold Plated RJ45 Connectors Networking Cable for Switch Router	$8	1500	$11,985	We planned to procure two network cables for each Unit of Compute Server and Storage Server
Racks	Data Center Rack Cabinet Enclosures (48 U)	$2,459	17	$41,803	As these are 48 U rack, we plan to mount 44 U (Server) in each rack, for mounting 728 servers we need 17 x 48U racks
Storage Servers	Petarack	$299,000	83	$24,817,000	Petarack (1.25 PB server) costs $299,000 each, we planned to procure 83 Storage rack unit servers as this sum up to 103 PB capacity.
Electric Power	1280 watts /Unit or Server/hr	$0.13	45442500	$5,721,210.75	Quantity Calculation = (728 Compute Servers + 83 Storage Servers + 34 Network Switches) *(1280*24*365*5) kWh
Cooling	APC InRow RP Chilled Water 460-480V 60Hz	$19,345.99	10	$193,459.90	An Unit of Chilled APC cooler will be used between two racks for allowing air flow!
Administration	Man power + Other admin stuff	$120,000	12	1440000	
TOTAL				**$56,345,722**	

We have chosen the below **storage server** for configuration 1 to meet the 100 PB distributed storage shared across the entire cloud:

Model: Petarack Storage Server (We used the same storage server model for 3 private cloud configurations)

Specs:

Components

Dual 2U Head Units for High Availability
2 x 4U Dual Controller JBODs
Expandable to 8 x 4U Dual Controller JBODs
Industry Standard APC Rackmount Cabinet
LCD/KVM Control Panel

Capacity

1,800 TB
(2 JBOD with 180 x 10TB Drives)

Supported RAID Levels

Stripe, Mirror, RAID-Z1 (Single parity),
RAID-(Z2) (Double parity)

Operating System

Enterprise-Class NexentaStor 4.0, 128-bit ZFS
Platform OpenSolaris kernel with the
GNU/Debian user interface

Head Unit Processor

Dual Intel Xeon E5 processor 2600 v4 series
Broadwell-EP Microarchitecture

Storage IO

Dual 10GbE iSCSI Ethernet Built-in
Fibre Channel Target Optional

Warranty

5-Year Unlimited

* For cooling solution, we used APC InRow RP Chilled Water 460-480V 60Hz cooling rack solution. We shall add the screenshot of the cart in reference.

Configuration 1 table for public cloud

Number of EC 2 Instances (d2.8xlarge)	Standard 1-year term upfront pricing per Instance	Standard 1-year term upfront pricing * 5 (as we need to purchase for a term of 5 year) per Instance	Total cost for procuring 1050 VM's of d2.8xlarge AWS config specs for a term of 5 years
1050	$23,616	$118,080	$123,984,000

S3 Storage (Quantity in GB)	Standard Storage Pricing/month/GB	Standard Storage Pricing/60 months/GB	Total cost of S3 Storage for 100 PB over a time span of 5 years
100000000	$0.02	$1.20	$120,000,000.00

Total pricing of AWS EC2 Instance and S3 storage
$243,984,000.00

NOTE: We have considered as 1 vCPU = 1 physical core (referred from KB article https://kb.vmware.com/kb/1010184) for calculating the number of VM's required to be deployed in public cloud (this is done across all public cloud configurations for 3 scenarios)

Observation:

Public Cloud costs approximately 4.3 times more than utilizing private cloud.

"**Private Cloud Configuration II**"

As per the given requirement in Configuration II of the project, support 1 million virtual machine (VM) where each VM requires:

2-core, 15 GB RAM, 32 GB SSD storage, and 1 Gb/s Fat-Tree network

In public cloud, each VM should be equivalent to **r3.large instances** (The above specified configuration is same as the configuration of **r3.large instance** in AWS public cloud)

In addition to the compute resource, a **10 PB distributed storage shared across the entire cloud** should be procured, with enough capacity for 10GB/sec throughput!

To support 1 million VM's, we need below config specs:

(a) Total number of cores required: 2 * 10,00,000 = 20,00,000 cores

(b) Total amount of memory: 15 GB * 10,00,000 → 150,00,000 GB → 15,000 TB → 15 PB
(c) Total amount of SSD storage → 32 GB * 10,00,000 → 320,00,000 → 32,000 TB → 32 PB
(d) 1 Gb/sec Fat-Tree Network.

We have chosen the below **compute server** for configuration 2:

Server Model and Specs:

Model: Iris 4882

Specs:

Not Logged In View a generated quote Login or Signup

Shopping Cart

Qty	System	Description			
1	Iris 4882	view components	$171,239.60	$171,239.60	Edit System Delete System

CPU:	2x	Intel Xeon E5-2699v4 22C 2.2GHz 55MB Cache
Memory:	1x	512GB DDR4 ECC Reg 2400MHz (16 x 32GB)
SATA HDD/SSD:	3x	Intel S4600 Series 240GB 3D1 TLC SATA SSD 6Gb/s
Operating System:	1x	RHEL Server - Premium 4 Sockets, Unlimited Virtual Guests 64bit
CPU:	2x	Intel Xeon E5-2699v4 22C 2.2GHz 55MB Cache
Memory:	1x	512GB DDR4 ECC Reg 2400MHz (16 x 32GB)
SATA HDD/SSD:	1x	Intel S4600 Series 240GB 3D1 TLC SATA SSD 6Gb/s
Operating System:	1x	RHEL Server - Premium 4 Sockets, Unlimited Virtual Guests 64bit
CPU:	2x	Intel Xeon E5-2699v4 22C 2.2GHz 55MB Cache
Memory:	1x	512GB DDR4 ECC Reg 2133MHz (8 x 64GB)
SATA HDD/SSD:	1x	Seagate Exos 7E2000 1TB SATA 6Gb/s
Operating System:	1x	RHEL Server - Premium 4 Sockets, Unlimited Virtual Guests 64bit
CPU:	2x	Intel Xeon E5-2603v4 6C 1.7GHz 15MB Cache
Memory:	1x	16GB DDR4 ECC Reg 2400MHz (2 x 8GB)
NVMe SSD:	3x	Intel S4600 Series 240GB 3D1 TLC SATA SSD 6Gb/s
Operating System:	1x	RHEL Server - Premium 4 Sockets, Unlimited Virtual Guests 64bit
CPU:	2x	Intel Xeon E5-2699v4 22C 2.2GHz 55MB Cache
Memory:	1x	512GB DDR4 ECC Reg 2133MHz (8 x 64GB)
SATA HDD/SSD:	3x	Intel S4600 Series 240GB 3D1 TLC SATA SSD 6Gb/s
Operating System:	1x	RHEL Server - Premium 4 Sockets, Unlimited Virtual Guests 64bit
CPU:	2x	Intel Xeon E5-2699v4 22C 2.2GHz 55MB Cache
Memory:	1x	256GB DDR4 ECC Reg 2133MHz (4 x 64GB)
SATA HDD/SSD:	3x	Intel S4600 Series 240GB 3D1 TLC SATA SSD 6Gb/s
Operating System:	1x	RHEL Server - Premium 4 Sockets, Unlimited Virtual Guests 64bit
CPU:	2x	Intel Xeon E5-2699v4 22C 2.2GHz 55MB Cache
Memory:	1x	256GB DDR4 ECC Reg 2133MHz (4 x 64GB)
SATA HDD/SSD:	3x	Intel S4600 Series 240GB 3D1 TLC SATA SSD 6Gb/s
Operating System:	1x	RHEL Server - Premium 4 Sockets, Unlimited Virtual Guests 64bit
CPU:	2x	Intel Xeon E5-2699v4 22C 2.2GHz 55MB Cache
Memory:	1x	256GB DDR4 ECC Reg 2133MHz (4 x 64GB)
SATA HDD/SSD:	1x	Intel S4600 Series 240GB 3D1 TLC SATA SSD 6Gb/s
Operating System:	1x	RHEL Server - Premium 4 Sockets, Unlimited Virtual Guests 64bit
Warranty and Support:	1x	On Site Support (3Year, 5 x 9, Next Business day, Parts and Labor)

close

Create Quote from Webcart Why create a quote? **Total** $171,239.60 Checkout

Quotes

To meet the number of cores requirement and memory requirement to support 1 million VM's, we need 5682 x 8U of the similarly configured servers (specs mentioned above for each Unit)

Configuration 2 table for private cloud configuration

Configuration 2	Description	Price per Item	Quantity	Total Price	Notes
Computer Servers	Iris 4882 (Eight compute nodes in 4U)	$171,239.60	5682	$972,983,407.20	
Network Switches	Cisco Catalyst 4948E-F - Switch - L3 - managed - 48 x 10/100/1000 + 4 x 10 Gigabit SFP+ - rack-mountable	$13,319.99	260	$3,463,197.40	Our Datacenter used Top of the Rack switches, we would procure two network switches for each rack (each switch has 48 ports). Purpose of two switches per rack is used to support Fat tree config of 23 upstream ports and 23 downstream ports in addition to input and output ports.
Network Cables	Cat 7 Ethernet Cable Flat 10 FT-Ealona Shielded (SSTP) Lan Patch Cable 10 Gigabit 600Mhz Gold Plated RJ45 Connectors Networking Cable for Switch Router	$8	11464	$91,712	We planned to procure two network cables for each Unit of Compute Server and Storage Server
Racks	Data Center Rack Cabinet Enclosures (48 U)	$2,458.98	129	$317,208.42	Number of rack unit is calculated based on number of unit or servers divided by 48U
Storage Servers	Petarack (1 PB of Raw data storage, expanded to 4.3 PB) in one 42 U rack	$299,000	2	$598,000	Petarack (1.25 PB server) costs $299,000 each, we planned to procure 83 Storage rack unit servers as this sums upto 103 PB capacity.
Electric Power	2000W Gold-level High-efficiency Redundant Power Supply/Unit or Server/hr	$0.13	3981858000	$517,641,540.00	Power consumption cost is calculated based on kWh; Quantity Calculation = (45455 Compute Nodes + 2 Storage Servers) *(2000*24*365*5)Wh
Cooling	APC InRow RP Chilled Water 460-480V 60Hz	$19,345.99	65	$1,257,489.35	We are planning to use single APC InRow Chilled Water 460-480V 60Hz rack in between two-48U server
Administration	Man power + Another admin stuff	$150,000	30	4500000	

TOTAL				**$1,500,852,554.37**	

We have chosen the same **storage server** (PetaRack) specs for configuration 2 to meet the 10 PB distributed storage shared across the entire cloud, but with lesser capacity this time.

Configuration 2 table for public cloud

Number of EC 2 Instances (r3.large)	Standard 1-year term all upfront pricing per Instance	Standard all upfront pricing for a term of 5 years per Instance	Total cost for procuring 1 million VM's of r3.large AWS config specs for a term of 5 years
10,00,000	$772	$3,860	$3,860,000,000

S3 Storage (Quantity in GB)	Standard Storage Pricing/month/GB	Standard Storage Pricing/60 months/GB	Total cost of S3 Storage for 10 PB over a time span of 5 years
10000000	$0.02	$1.20	$12,000,000.00

Total pricing of AWS S3 Storage and EC2 Instance
$3,872,000,000.00

Observation:

Public Cloud for configuration 2 costs approximately 3 times more than utilizing private cloud for configuration 2 specs.

"**Private Cloud Configuration III**"

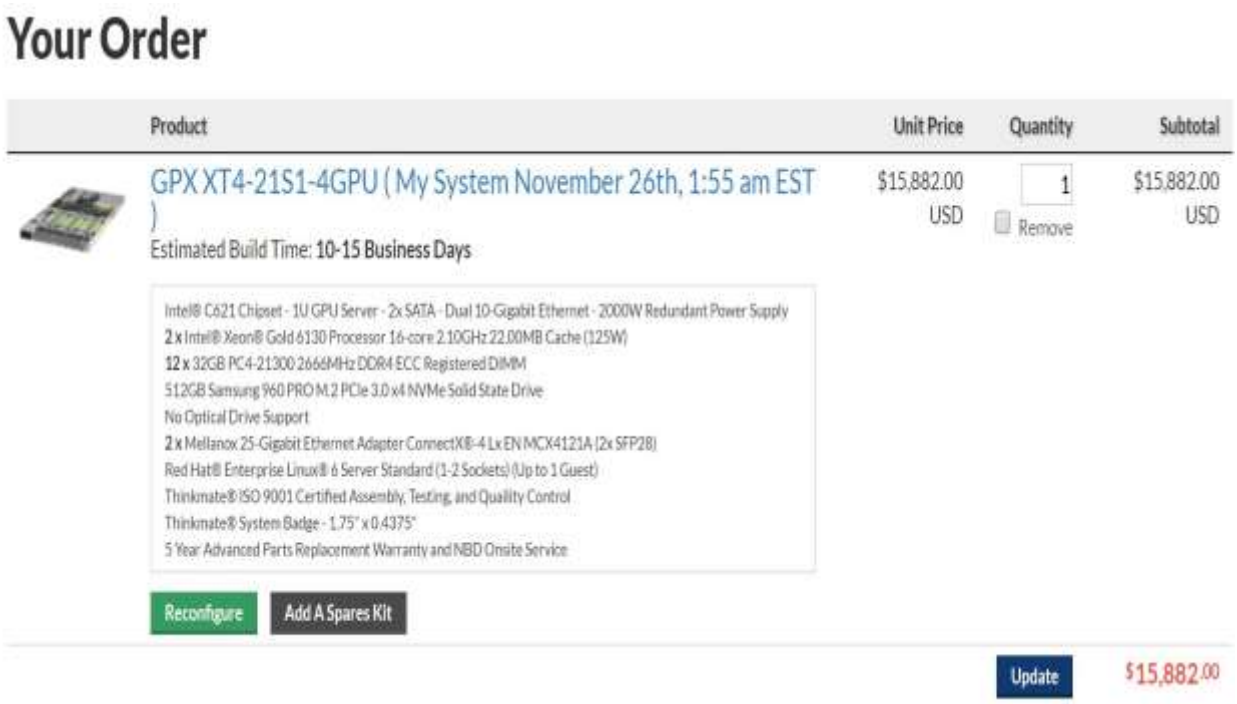
As per the given requirement in configuration 3, we must design a private cloud infrastructure to support deep learning with 1 exaflops of mixed precision performance, to meet this specs each VM must to equivalent to p3.16xlarge instances (which consists of 8 NVIDIA V100 GPU's, and allocate 8-cores per GPU (64 cores per node) with 8 GB of memory per core (512 GB per node); the network to use is at least 10 Gb/s per GPU (100 Gb/s) and must be organized in a Fat tree network.

We have chosen the below **GPU server** for configuration 3:

Server Model and Specs:

Model: GPX XT4-21S1-4GPU

Specs:



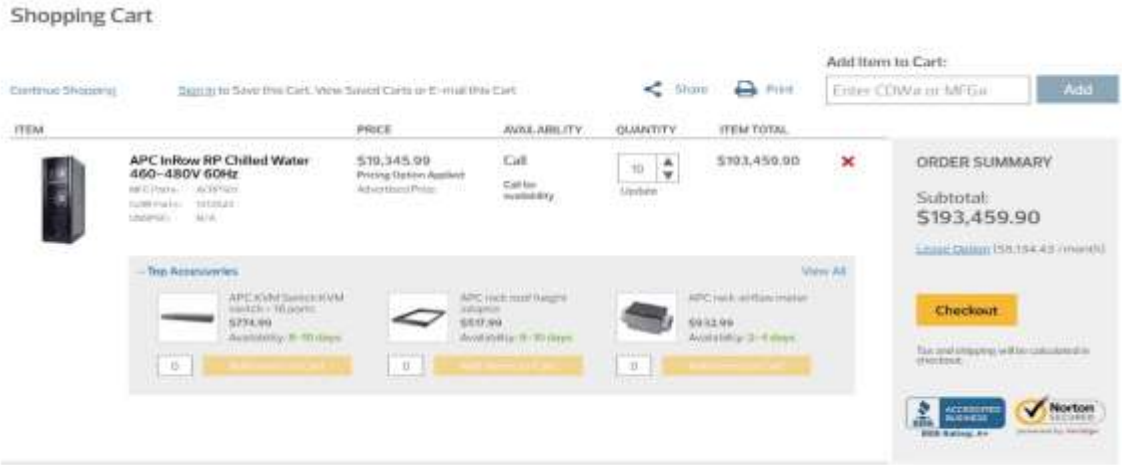
Screenshot of Shopping cart of NVIDIA GPU V100 Server for a single unit.

To meet the 1 exaflops of mixed precision performance requirement, we require 8000 NVIDIA V100 GPU's (which deliver 125 TFLOPS of performance each).

As per our server specs above, it has 4 NVIDIA V100 GPU's, so we need 2000 of similar specs server.

We have chosen the same **storage server** specs for configuration 2 to meet the 10 PB distributed storage shared across the entire cloud, but with lesser capacity this time.

Cooling specs are also met by the similar equipment as used for Configuration 1 and Configuration 2:



Screenshot of Shopping cart of APC In-Row Single cooling for a single unit.

Configuration 3 table for private cloud configuration

Configuration 3	Description	Price per Item	Quantity	Total Price	Notes
Computer Servers	GPX XT4-21S1-4GPU	$15,882.00	2000	$31,764,000.00	
Network Switches	Cisco Catalyst 4948E-F - Switch - L3 - managed - 48 x 10/100/1000 + 4 x 10 Gigabit SFP+ - rack-mountable	$13,920	92	$1,280,639.08	Our Datacenter used Top of the Rack switches, we would procure two network switches for each rack (each switch has 48 ports). Purpose of two switches per rack is used to support Fat tree config of 23 upstream ports and 23 downstream ports in addition to input and output ports.
Network Cables	Cat 7 Ethernet Cable Flat 10 FT-Ealona Shielded (SSTP) Lan Patch Cable 10 Gigabit 600Mhz Gold Plated RJ45 Connectors Networking Cable for Switch Router	$8	402	$3,212	We planned to procure two network cables for each Unit of Compute Server and Storage Server
Racks	Data Center Rack Cabinet Enclosures (48 U)	$2,459	46	$113,114.00	Number of rack unit is calculated based on number of unit or

					servers divided by 48U
Storage Servers	Petarack	$299,000	1	$299,000	Petarack (1 PB server) costs $299,000 each, we planned to procure 1 Storage rack unit servers as this sums upto 1 PB capacity.
Electric Power	2000W Gold-level High-efficiency Redundant Power Supply/Unit or Server/hr	$0.13	175287600	$22,787,388.00	Power consumption cost is calculated based on kWh; Quantity Calculation = (2000 Compute Nodes + 1 Storage Servers) *(2000*24*365*5)Wh
Cooling	APC InRow RP Chilled Water 460-480V 60Hz	$19,345.99	23	$444,957.77	We are planning to use single APC InRow Chilled Water 460-480V 60Hz rack in between two-48U server
Administration	Man power + other administration work	$150,000	8	1200000	
TOTAL				**$57,892,310.83**	

Configuration 3 table for public cloud configuration

Number of EC 2 Instances (p3.16xlarge)	Reserved instance monthly pricing per Instance	Reserved instance pricing for a term of 5 years per Instance	Total cost for procuring 1 million VM's of p3.16xlarge AWS config specs for a term of 5 years
1000	$12,197	$731,820	$731,820,000

S3 Storage (Quantity in GB)	Standard Storage Pricing/month/GB	Standard Storage Pricing/60 months/GB	Total cost of S3 Storage for 1 PB over a time span of 5 years
1000000	$0.02	$1.20	$1,200,000.00

Total pricing of AWS S3 Storage and EC2 Instance

$733,020,000.00

Observation:

Public cloud for configuration 3 specs costs approximately 12 times more than utilizing private cloud for configuration 3 specs

---

Table 2: Summary table comparing the 3 configurations between the public and private cloud; your cost of power, cooling, and administration should be to cover 5 years of costs

Table 2	Configuration 1	Configuration 2	Configuration 3
Public Cloud (including EC2 and S3) Cost over 5 years, 24/7 operation, with 100% usage	$243,984,000	$3,872,000,000	$733,020,000
Private Cloud cost over 5 years, 24/7 operation, with 100% usage	$56,345,722	$1,500,852,554.37	$57,892,310.83
What utilization must be achieved with the private cloud to make the private cloud option more attractive than the public cloud?	23 %	25 %	12.5 %

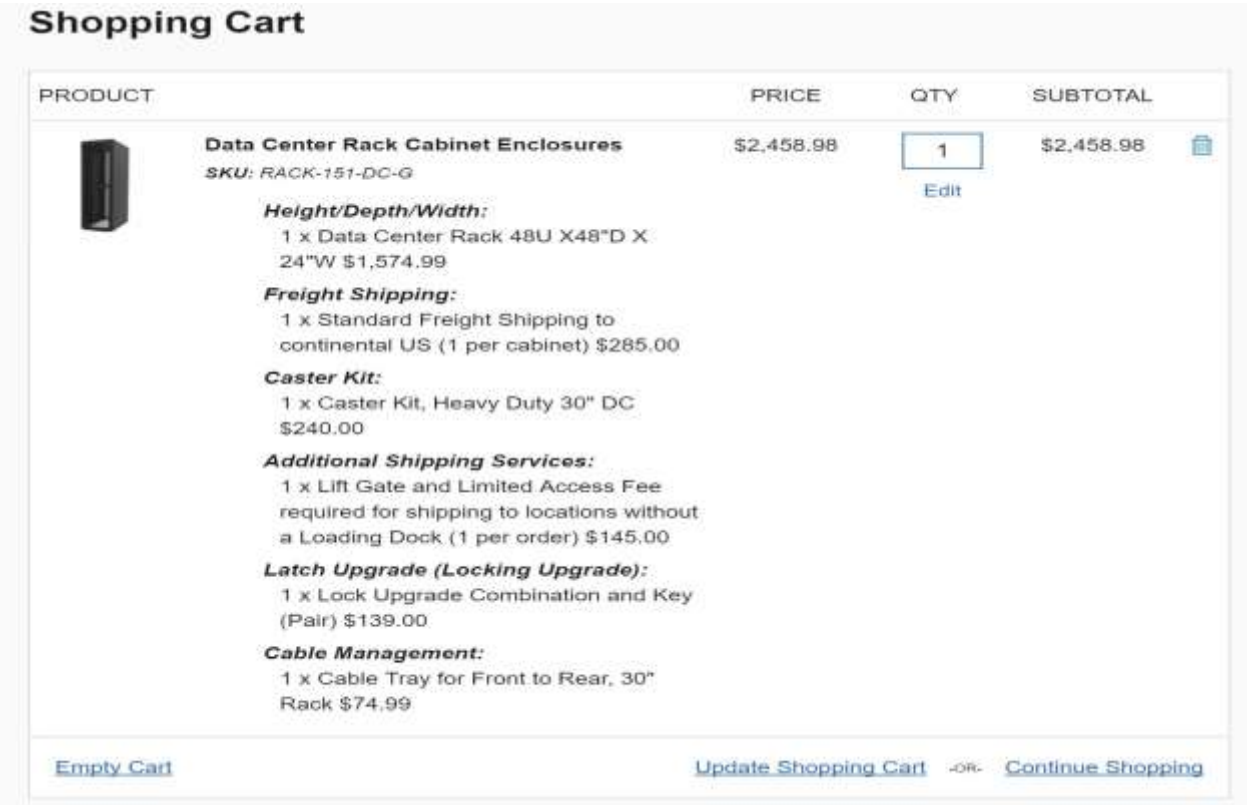
NOTE: The utilization is calculated based on the discussion with professor in class. In theory, the specified percentage of utilization (mentioned in the table above) must be achieved with the private cloud to make the private cloud option more attractive than the public cloud.

Reference:

(1) Compute Servers purchase calculation done from PogoLinux - http://www.pogolinux.com/
(2) GPU Servers purchase calculation done from thinkmate site - https://www.thinkmate.com/
(3) Storage server purchase calculation done from Aberdeen site - https://www.aberdeeninc.com/petarack/
(4) Distributed and Cloud Computing book – by Kai Hwang

(5) Screenshots of purchasing cart

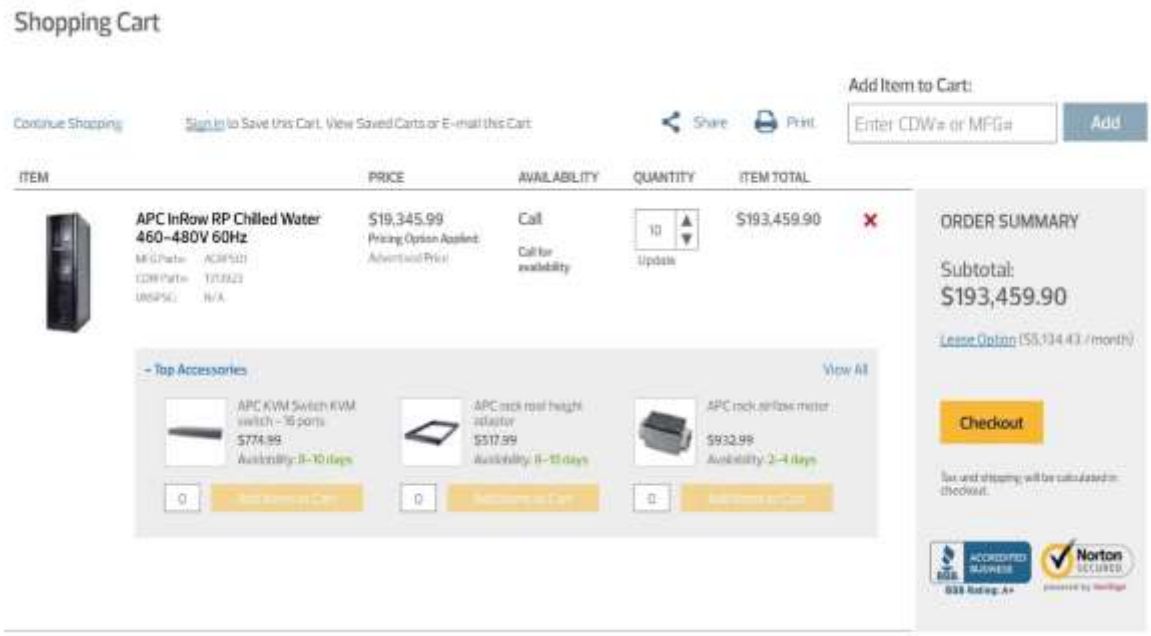
(a) Rack Unit – Single Unit Cost & Model



(b) Networking switch – Single Unit Cost & Model

(c) Networking Cable – Single Unit Cost & Model

(d) APC In-Row Cooling equipment – Single Unit Cost & Model



(e) d2.8xlarge AWS public cloud instance – Single Unit Cost

d2.8xlarge

STANDARD 1-YEAR TERM					
Payment Option	Upfront	Monthly*	Effective Hourly**	Savings over On-Demand	On-Demand Hourly
No Upfront	$0	$2347.68	$3.216	42%	$5.52 per Hour
Partial Upfront	$12048	$1004.48	$2.751	50%	
All Upfront	$23616	$0	$2.696	51%	

(f) r3.large AWS public cloud instance – Single Unit Cost

r3.large

STANDARD 1-YEAR TERM					
Payment Option	Upfront	Monthly*	Effective Hourly**	Savings over On-Demand	On-Demand Hourly
No Upfront	$0	$76.65	$0.105	37%	$0.166 per Hour
Partial Upfront	$514	$22.63	$0.090	46%	
All Upfront	$772	$0	$0.088	47%	

(g) p3.16x large AWS instance pricing is not listed explicitly on any Amazon websites, I referred this through a 3rd party site where it shows that reserved instance pricing for each month is $12,197.

(h) The specs for Compute Server and Storage Server (along with screenshot) for each configuration is already mentioned in individual configuration documentation.

----------End of Document----------