You are hired by a start-up company who is considering using cloud computing instead of building its own infrastructure

(Isn’t a private cloud same as building own infrastructure

What is the meaning when you say cloud computing vs own infrastructure?

What other infrastructure is being referred to here?)

Cloud computing software stack at the Iaas layer will be used

But not sure whether the computing resources need to be rented from a public cloud on-demand or a private cloud should be purchased

You are tasked to find the cost breakdown of a private cloud, and compare that to what Amazon would charge

Amazon instance types and prices

<http://aws.amazon.com/ec2/instance-types/>

<http://aws.amazon.com/ec2/pricing/>

stick to Linux on-demand pricing

Since there is a need to estimate the cost of the hardware when building a private cloud

Can find prices of hardware in the following link

Dell - <http://www.dell.com/p/enterprise-products.aspx?c=ae&l=en&s=bsd&~ck=mn>

AcmeMicro - <http://www.acmemicro.com>

PogoLinux - <http://www.pogolinux.com>

(good sources for server h/w)

You must include a printout of your shoping cart in your final writeup report for this assignment; include this as an appendix at the end of your report.

You are to estimate the cost of different configurations for 3 different set of requirements: conf1 conf2 and conf3

you may assume a 5-year amortization cost

you will have to factor in things other than hardware, such as cooling, power, administration costs, network infrastructure (e.g. switches); you can assume 1 system administrator is needed for every 1000 servers

Strategy

1) Can first list down the h/w cost for each component needed

2) find out the performance achieved by combining the hardware in different ways

a) consider the performance of each component

b) how it works when combined

3) must compare both Amazon on-demand service and private cloud

**Cloud computing course**

**Establishing the cloud computing environment**

**1st project to know how each individual item perform so that you can get to know how to combine them to a bigger cluster depending on the requirement**

**2nd project is to introduce how data analytics on cloud is being performed**

**3rd ?**

**Project**

**Uses 1st knowledge and knowledge about cloud computing environment to build a cloud infrastructure cost**

**And provides a comparison how our cost with amazon cost**

**How does amazon achieve that can also be seen by the project**

1) Can first list down the h/w cost for each component needed

Private cloud

PowerEdge Tower server

PowerEdge Rack server

Modular infrastructure

Amazon

ECU unit of compute performance

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Name** | **Core(vCPU)** | **Mem(GiB)** | **Hdd/ssd(GB)** | **n/w** | **Dist storage** | **Total cost** |
| **D2.8x large** | **36** | **244** | **48000 HDD** |  | **132,120,576$** | **1,208,880$** |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

**Forming a cluster under public cloud Amazon EC2 instance with S3 distributed storage**

**1) can we use multiples of d2.8xlarge instance**

**2) can we use some other instance and make up a d2.8x large instance**

**3) how to calculate 100gb/s throughput**

**They can focus on the data storage space and costs, and they can ignore the bandwidth costs of accessing data on S3. (But for their private cloud, they need to make sure they will have enough storage and network bandwidth to deliver 1000PB of storage and 100GB/sec of I/O performance.)**

**4) should the calculation be based on the number of instances got by calculating using individual components like (HDD/MEM)**

**5) how to be sure of the fat n/w 10gb/sec ethernet being used in our cluster**

**6) calculation of 100PB distributed storage using S3 strategy valid?**

**7) how to calculate admin cost**

**9) what do you mean by each VM**

**10)suppose we use different instance cluster .. will we be able to segregate a vm to form d2 configuration**

**Config 1**

**Number of d2.8x large instances required to full fill below requirement**

|  |  |
| --- | --- |
| Config 1 | Description |
| Compute Servers | 32 k cores |
| Memory | 256 TB |
| Disk | 50 PB HDD |
| Network | 10 Gb/sec |
| Distributed storage | 100 PB |

**Calculations:**

**1.Instances Required**

**Each d2 instance**

**36 cores, 244 GB of Memory and 24\* 2000 HDD at the cost of $5.52 per hour**

**For the configuration mentioned in the above table. We require a minimum of 1**

**CPU**

**32K cores**

**1 d2 instance CPU’s = 36 vCPU**

**vCPU to cores = vCPU/2**

**= 18 cores**

**Instances required for 32 K core = 32000/18**

**= 1778 d2.8xlarge instances --------> a**

**Memory**

**256TB in GB = 262,144 GB**

**1 d2 instance Disk capacity = 244 GiB**

**244 GiB in GB = 244\* 1.074 = 262.06 GB**

**Instances required for 256TB = 262,144 /262.06**

**= 1,001 d2.8xlarge instances --------> b**

**Disk**

**50PB in GB = 52,428,800 GB**

**1 d2 instance Disk capacity = 48,000 GB**

**Instances required for 50PB = 52,428,800/48000**

**= 1,093 d2.8xlarge instances --------> c**

**Total number of d2.8xlarge instances required = max ( a, b, c)**

**= 1,778 instances**

**2.Distributed Storage**

**S3 standard storage pricing per month = $0.021 per GB**

**100 PB in GB = 104,857,600 GB**

**Storage cost for 5 years = 0.021 \* 104,857,600 \* 12 \* 5**

**= $ 132,120,576**

**Cost calculation**

**1 D2.8xlarge instance cost**

**36 cores, 244 GB of Memory and 24\* 2000 HDD at the cost of $5.52 per hour**

**1,093 instances cost per hour = 5.52 \* 1778**

**= $9,814.56**

**1,093 instances cost for 5 years = $9,814.56 \* 24 \* 365 \* 5**

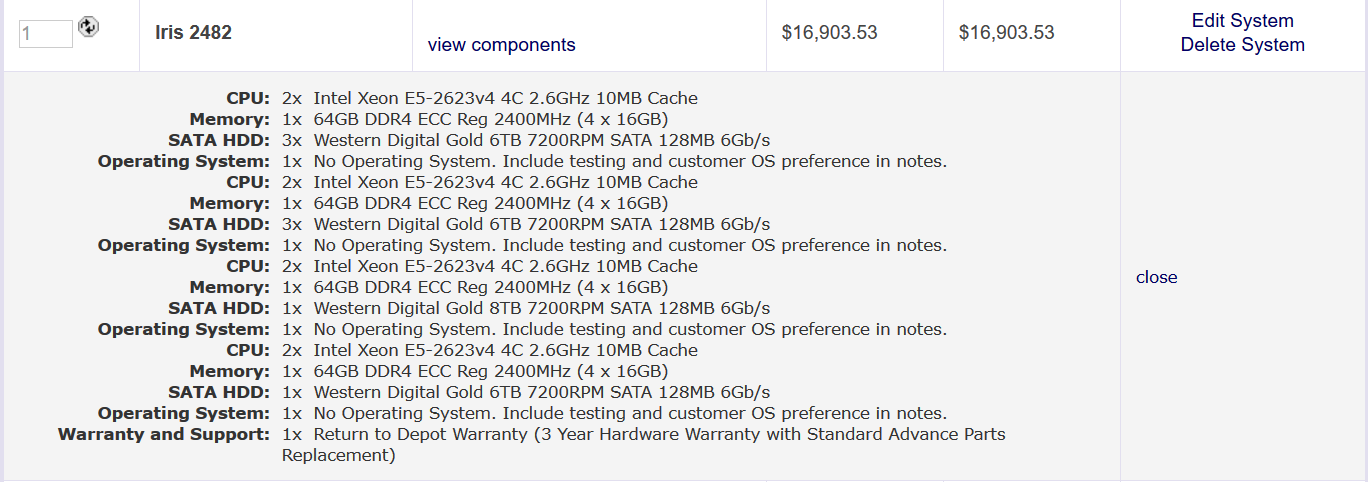
**= $429,877,728**

**Total cost = distributed storage cost + d2.8xlarge instance cost**

**= $ 132,120,576 + $ 429,877,728**

**= $ 561,998,304**

**Private cloud**



**d2.8xlarge 1 instance – $16,903.53**

**instances \* 1000 – 16,903,530**

**the rack has core - 2\*4, 2\*4, 2\*4, 2\*4 – 32 cores**

**memory - 1\*64, 1\*64, 1\*64, 1\*64 – 256 GB**

**Disk - 3\*6, 3\*6, 1\*6, 1\*8 – 50 TB**

Network cost with Dell z9000 switches in Fat tree configuration with 10GbE

16 leaf switches =16 \* $**862**.62 = 13,801.92

8 spine switches = 8 \* $**862**.62 = 6,900.96

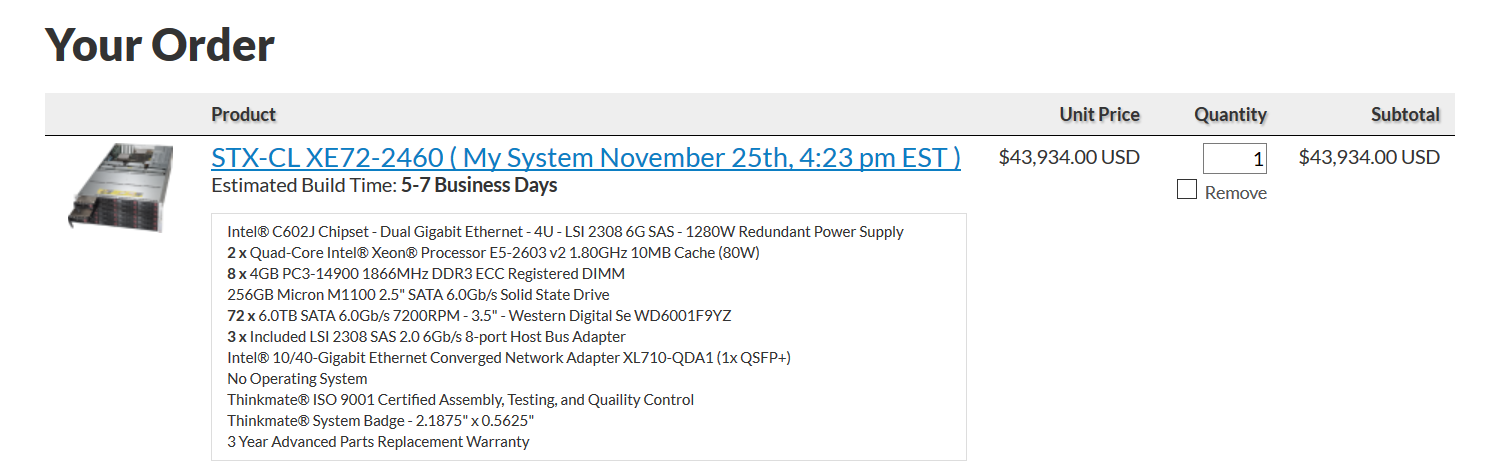
leaf to spine cable = 16 \* 16 \* 590 = 151,040

( # leaf switches \* # spine switches \* # of connection per switch \* cost of cable

leaf to node cable = 16\*4 splitter( 16 ports splitted to 4 each) = 16 \*16 \* 720 = 184,320

( # leaf switch port \* cost of splitter cable)

**Total Network cost = 356,062.88**



**Distributed storage calculation**

**Each node = 72 \* 6 TB HDD SATA= 432 TB/node**

**Instances required for 100PB = 100 \* 1024 /432 = 237 instances**

**Total cost = instances \* cost per instance = 237 \* 43,934 = 10,412,358**

**Performance**

**100GB/sec**

**Each instance throughput = 72 \* transfer rate per HDD =72 \* 171MB/s = 12312 = 12 GB/s**

**237 instances = 237 \* 12 GB/s = 2,844 GB/s**

Network cost with Dell z9000 switches in Fat tree configuration with 10GbE

31 leaf switches =31 \* $**862**.62 = 26,741.22

16 spine switches = 16 \* $**862**.62 = 13,801.92

leaf to spine cable = 31 \* 16 \* 590 = 292,640

( # leaf switches \* # spine switches \* # of connection per switch \* cost of cable

leaf to node cable = 15 cables = 15 \* 16 \* 590 = 141,600

( # leaf switch port \* cost of splitter cable)

**Total Network cost = 474,783.14**

**Total cost = $27,790,671.14**

**TODO:**

**calculate private cloud cost for 5 years 24/7**

**include system admin cost**

**cooling cost**

**power cost**

**utilization for 5 years**

**Config 2**

**Number of r3.large instances required to full fill below requirement**

|  |  |
| --- | --- |
| Config 2 | Description |
| Compute Servers | 1 Million \* 2 cores |
| Memory | 1 Million \* 15 GB RAM |
| Disk | 1 Million \* 32 GB SSD |
| Network | 1 Million \* 1GB/sec |
| Distributed storage | 10 PB |

**Cores = 11,364 (44 core 4 node 176)**

**Memory = 3663**

**Disk = 32Mil / 23347.2 = 1371**

**Each r3.large instance**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Name** | **vCPU** | **Mem(GiB)** | **HDD/SSD(GB)** | **cost** |
| **R3.large** | **2** | **15** | **1\*32SSD** | **$0.166/Hour** |

**Calculations:**

**1.Instances Required**

**For the configuration mentioned in first table**

**CPU**

**2 million cores**

**1 r3.large instance CPU’s = 2 vCPU**

**vCPU to cores = vCPU/2**

**= 1 core**

**Instances required for 2Million core = 2,000,000/1**

**= 2,000,000 r3.large instances --------> a**

**Memory**

**15 Million GB**

**1 r3.large instance Disk capacity = 15 GiB**

**15 GiB in GB = 15 \* 1.074 = 16.11 GB**

**Instances required for 15 Million GB = 15000000/16.11**

**= 931,099 r3.large instances --------> b**

**Disk**

**32 Million GB**

**1 r3.large instance Disk capacity = 32 GB**

**Instances required for 32 Million GB = 32000000/32**

**= 1,000,000 r3.large instances --------> c**

**Total number of r3.large instances required = max ( a, b, c)**

**= 2,000,000 instances**

**2.Distributed Storage**

**S3 standard storage pricing per month = $0.021 per GB**

**10 PB in GB = 10485760 GB**

**Storage cost for 5 years = 0.021 \* 10485760 \* 12 \* 5**

**= $ 13,212,057.6**

**3.Cost calculation**

**1 r3.large instance cost**

**36 cores, 244 GB of Memory and 24\* 2000 HDD at the cost of $5.52 per hour**

**2,000,000 instances cost per hour = 0.166 \* 2,000,000**

**= $332,000**

**2,000,000 instances cost for 5 years = $332,000 \* 24 \* 365 \* 5**

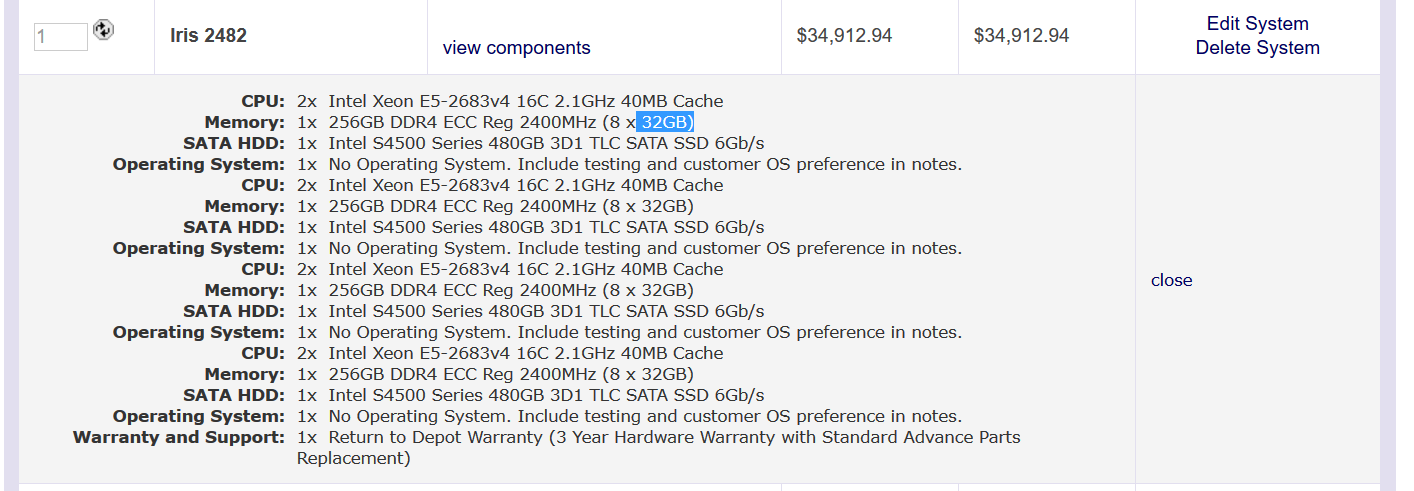
**= $14,541,600,000**

**Total cost = distributed storage cost + r3.large instance cost**

**= $13,212,057.6 + $14,541,600,000**

**= $ 14,554,812,057.6**

**Private cloud**



**60 = 1 rack – $34,912.94**

**1 million vm = 16,667 instances \* 34,912.94 = $ 581,893,970.98**

**the rack has core - 2\*16, 2\*16, 2\*16, 2\*16 – 128 cores**

**memory - 1\*256, 1\*256, 1\*256, 1\*256– 1024 GB**

**Disk - 1\*480, 1\*480, 1\*480, 1\*480 – 1920 GB**

**60 \* 1VM**

Network cost with Dell z9000 switches in Fat tree configuration with 10GbE

1042 leaf switches =1042 \* $**862**.62 = 898,850.04

521 spine switches = 521 \* $**862**.62 = 449,425.02

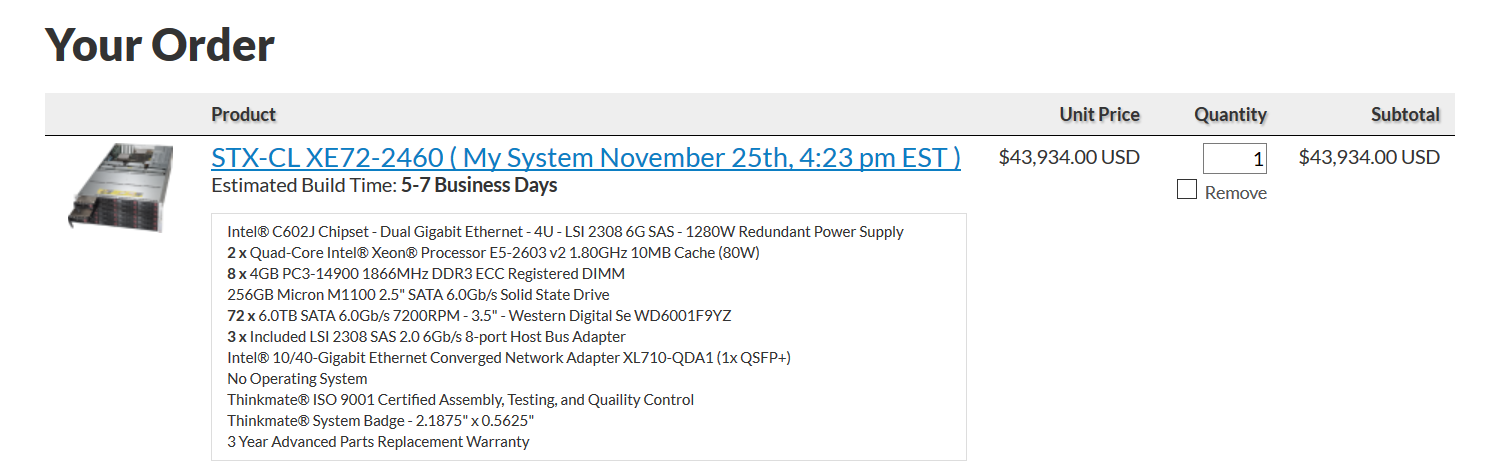
leaf to spine cable = 1042 \* 16 \* 590 = 9,836,480

( # leaf switches \* # spine switches \* # of connection per switch \* cost of cable

leaf to node cable = 1042\*4 splitter( 1042 ports splitted to 4 each) = 1042 \*16 \* 720 = 12,003,840

( # leaf switch port \* cost of splitter cable)

**Total Network cost = 23,188,595.06**



**Distributed storage calculation**

**Each node = 72 \* 6 TB HDD SATA= 432 TB/node**

**Instances required for 10PB = 10 \* 1024 /432 = 23.7 instances**

**Total cost = instances \* cost per instance = 24 \* 43,934 = 1,054,416**

**Performance**

**100GB/sec**

**Each instance throughput = 72 \* transfer rate per HDD =72 \* 171MB/s = 12312 = 12 GB/s**

**24 instances = 237 \* 12 GB/s = 2,844 GB/s**

Network cost with Dell z9000 switches in Fat tree configuration with 10GbE

1043 leaf switches =1043 \* $**862**.62 = 899712.66

522 spine switches = 522 \* $**862**.62 = 450287.64

leaf to spine cable = 1043 \* 16 \* 590 = 9845920

( # leaf switches \* # spine switches \* # of connection per switch \* cost of cable

leaf to node cable = 1043\*4 splitter( 1043 ports splitted to 4 each) = 1043 \*16 \* 720 = 12015360

( # leaf switch port \* cost of splitter cable)

**Total Network cost = 23,211,280.3**

**Total cost = $ 581,893,970.98 + 1,054,416 + 23,211,280.3**

**Total cost = $606,159,667.28**

**Config 3**

**Number of p3.16xlarge instances required to full fill below requirement**

|  |  |
| --- | --- |
| Config 3 | Description |
| Compute Servers | 1 exaflop mixed precision |
| (NVIDIA V100 GPUs) | 8 GPU/node(64cores/node) |
| Memory | 8GB/core(512GB/node) |
| Disk |  |
| Network | 10Gb/s per GPU(100Gb/s should work) |
| Distributed storage | 1 PB |

**Each p3.16xlarge instance**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Name** | **GPU’s per node** | **(vCPU)** | **Mem(GB)** | **HDD/SSD(GB)** | **Precision(mixed) / GPU** | **cost** |
| **Each p3.16xlarge** | **8** | **64** | **488** | **N/A** | **125 TF** | **24.48/hr** |

**2.Distributed Storage**

**S3 standard storage pricing per month = $0.021 per GB**

**1 PB in GB = 1,048,576 GB**

**Storage cost for 5 years = 0.021 \* 1048576 \* 12 \* 5**

**= $ 1,321,205.76**

**1 exaflop = 1,000,000 TF**

**Each instance perf = 8 \* 125 TF = 1000TF**

**Instances required = 1,000,000/1,000 = 1,000**

**Cost of total nodes = 1,000 \* 24.48 \* 24 \* 365 \* 5 = 1,072,224,000**

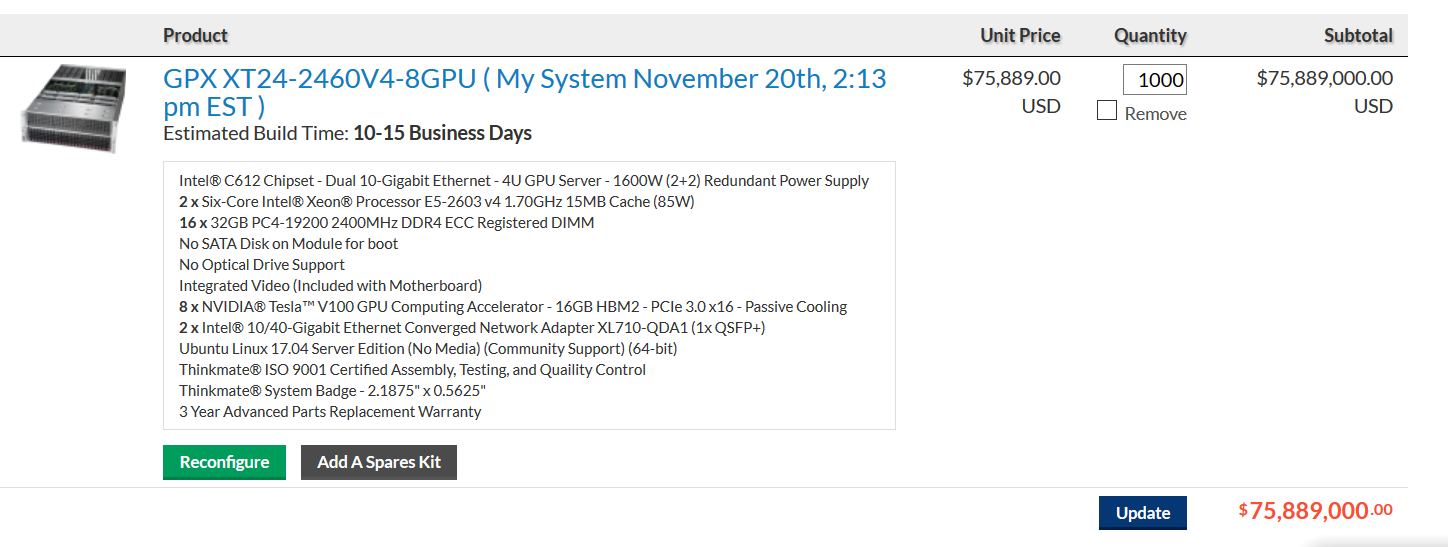
**512GB /node = 488 GB/node \* 1.05**

**Inclusive of instance count from previous calculation = 1.05 \* 1,000 =1050**

**Cost of total nodes = 1,050 \* 24.48 \* 24 \* 365 \* 5 = 1,125,835,200**

**Total cost = distributed storage cost + compute cost = $ 1,321,205.76 + $1,125,835,200 = $1,127,156,405.76**

**Private cloud**



**1 exaflop = 1,000,000 TF**

**Each instance perf = 8 \* 125 TF = 1000TF**

**Instances required = 1,000,000/1,000 = 1,000**

**Cost of total nodes = 1,000 \* 75889 = $75,889,000**

Network cost with Dell z9000 switches in Fat tree configuration with 10GbE

# leaf = # instances / 16 port connection(each leaf) = 1000/16 = 63

# spine = # uplinks from each leaf /32 = 16 \* 63/32 = 32

Cost of 63 leaf switches =63 \* $**862**.62 = $54,345.06

32 spine switches = 32 \* $**862**.62 = $27,603.84

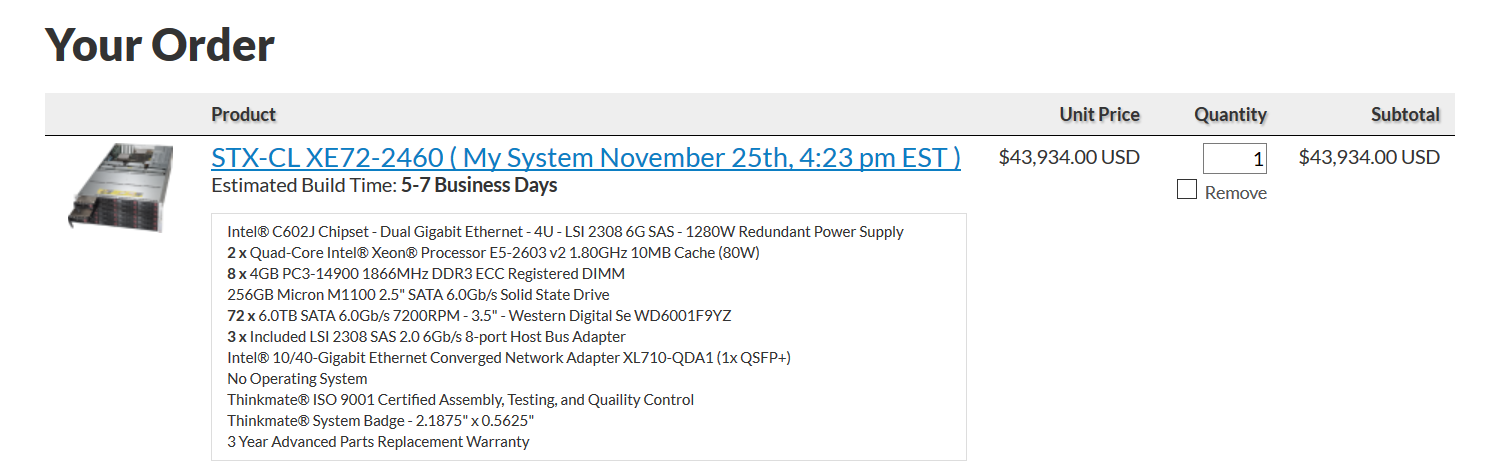
leaf to spine cable = 63 \* 16 \* 590 = 594,720

( # leaf switches \* # spine switches \* # of connection per switch \* cost of cable

leaf to node cable = 1 cable per port = 63 \* 16 \* 590 = $594,720

**Total Network cost = $ 1,271,388.9**

**Total Cost = compute cost + network cost = $75,889,000 + $ 1,271,388.9 + 131,802 = 77,292,190.9**



**Distributed storage calculation**

**Each node = 72 \* 6 TB HDD SATA= 432 TB/node**

**Instances required for 1PB = 1 \* 1024 /432 = 3 instances**

**Total cost = instances \* cost per instance = 3 \* 43,934 = 131,802**

**Performance**

**100GB/sec**

**Total Cost = compute cost + network cost = $75,889,000 + $ 1,271,388.9 + 131,802 = 77,292,190.9**