Assignment I – Computer Architecture

1.a) Research on the characteristics of CPUs, GPUs, and TPUs. What purpose have they been designed for? Which data science problems benefit most?

**CPUs (Central Processing Unit)**

Characteristics:

* Optimized for general-purpose computing.
* Fewer Cores (typically up to 64 cores in high-end server CPUs).
* High single threaded performance.

Designed Purpose:

They are designed for handling general computing tasks like running, operating system, applications and background processes.

Data Science Benefits:

Basic data science problems like data parsing, data preprocessing, and small scale machine learning tasks.

**GPUs (Graphic Processing Unit)**

Characteristics

* Optimized for parallel processing.
* Initially designed for rendering graphics but evolved into general-purpose parallel processors.
* Thousands of smaller cores.

Designed Purpose:

They are designed for graphics rendering, video editing and increasingly for general-purpose parallel computing tasks.

Data Science Benefits:

Highly beneficial for tasks that can be parallelized, such as training deep learning models, image and video processing.

**TPUs (Tensor Processing Unit)**

Characteristics:

* Optimized for tensor operations.
* Matrix multiplication is the core operation.

Designed Purpose:  
Accelerating machine learning workloads, especially deep learning models.

Data Science Benefits:  
Extremely beneficial for training large scale deep learning models and other tensor intensive computations.

\

1.b) Choose a current model each (one CPU,one GPU, and one TPU) and compare

* Performance
* Power consumption
* Cost

Current Model Comparison

**CPU: Intel core i9-12900k**

Performance:

* Multi-core: 16 Cores
* Single Threaded performance is excellent.

Power Consumption:

* TDP: 125W

Cost:

* Around $500 – $600

**GPU: NVIDIA GeForce RTX 3090**

Performance:

* CUDA Cores: 10496
* Memory: 24 GB GDDR6X

Power Consumption:

* TDP: 350W

Cost:

* Around $1500 - $2000

**TPU: Google Cloud TPU v4**

Performance:

* Matrix Multiply Units: 2048

Power Consumption:

* Not directly comparable due to it’s cloud nature

Cost:

* Around $32 per TPU per hour

References:

1. Chat gpt – 3
2. Inter core i9 -12900K: <https://www.intel.com/content/www/us/en/homepage.html>
3. NVIDIA GeForce RTX 3090: <https://www.nvidia.com>
4. Google Cloud TPU v4: <https://cloud.google.com/tpu>

Assignment III - Cost of Cloud Computing

For a data science project you need the following IaaS ressources

* Virtual machine with at least 4 vCPUs & 8 GB RAM (no special instance needed)
* Object storage for up to 100TB of data (standard)
* Hosted in Europe

1.a)  How much will this setting cost you per year? Choose one of the providers presented in the lecture.

**Cost per year on AWS:**

Virtual Machine:

For a VM with atleast **4 vCPUs** and **8 GB RAM**

Amazon EC2 Pricing:

* Instance type: t3.xlarge
* Price: $0.3328 per hour

We assume VM runs 24/7 for a year:

Cost per Day: 0.3328 \* 24 = $7,9872

Cost per Month: 7,9872 \* 30 = $239,616

Cost per Year = 12 \* 239,616 = $2875.392

A screenshot of a computer

Description automatically generated

Object Storage:

For **100TB** of standard object storage on AWS, let’s use Amazon S3.

Amazon S3 Pricing:

* Standard Storage Price:
  + - * First 50 TB/Month = $0.0245 per GB
      * Next 450 TB/Month = 0.0235 per GB

Cost per Month(100TB) = 50\* 1000\* 0.0245 + 50 \*1000\* 0.0235 = $2,400

Cost Per Year = 2,400\* 12 = $28,800

A screenshot of a computer

Description automatically generated

**Total Cost per Year** on AWS: 2875.392 + 28,800 = **$31,675.39**

1.b) You put your project in operation and monitor a mean value of 50TB/month data transfer (outgoing). How much will that cost you per month?

For outgoing transfer of **50TB/Month** on AWS:

Amazon EC2 Data Transfer Pricing:

* First 10 TB/Month: $0.09 per GB
* Next 40 TB/Month: $0.085 per GB

Cost per Month(50TB) = 10 \* 1000 \* 0.09 + 40 \* 1000 \* 0.085 = 900 + 3400 = **$4,300**

**A screenshot of a computer

Description automatically generated**