

General AI/ML

Unit 3: Development Environment
Familiarization



3.1.1

Introduction to Cloud

Cloud vs Local Computing

Cloud Computing

- Cloud computing is computing on a server (a computer) that is maintained by someone else
- Users can access software, storage, and processing power on-demand from such servers over the internet, without the need for direct active management by the user
- Allows us to access a vast amount of computing power and storage on a pay-as-you-go basis

Local / On-Premises Computing

- Running applications and storing data on computers or servers that are physically located in-house
- Offers greater control over data and infrastructure
- Reduced dependency on external providers

Cloud Providers

- Amazon Web Services, Microsoft Azure, and Google Cloud Platform are three major hyperscale cloud providers, each their own unique strengths
- In this competition, we will be working with Google Cloud Platform

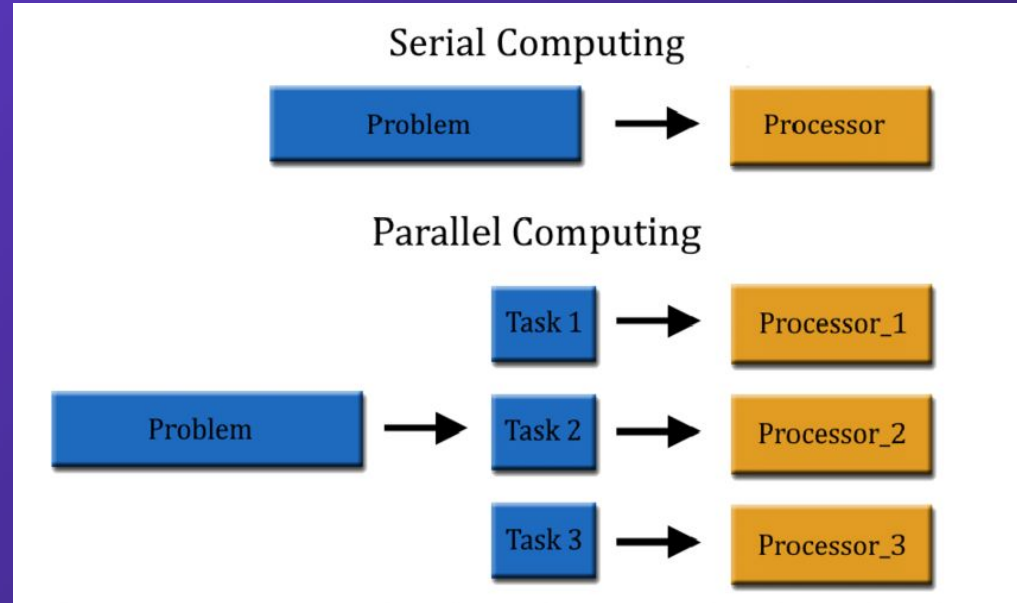


Why Cloud Computing?

- **Cost:** Powerful hardware is expensive. Compared to buying dedicated hardware, cloud computing reduces upfront hardware costs by paying only for what you use
- **Scalability:** Easier to scale computing resources up or down based on demand
- **Reliability:** AI & computing workloads are less likely to be interrupted by hardware failures or downtime compared to a self-managed local setup

Why Cloud Computing for AI?

- Parallelization
 - The ability to perform multiple computations simultaneously by dividing tasks across multiple processors or cores.
- Cloud platforms offer access to specialized hardware for deep learning tasks (GPUs, TPUs), optimized for parallel processing of AI algorithms.



Key Considerations for Choosing Cloud

- Scalability: If your AI needs fluctuate, the cloud's flexibility is a major advantage
- Hardware needs: Cloud provides access to clusters of servers and specialized hardware for parallel processing
- Data sensitivity: Highly sensitive or regulated data may need on-premises storage
- Budget: Consider potential short and long-term costs of cloud services