Computing Infrastructures - Notes

260236

March 2024

Preface

Every theory section in these notes has been taken from two sources:

- The Datacenter as a Computer: Designing Warehouse-Scale Machines, Third Edition. [1]
- Quantitative System Performance: Computer System Analysis Using Queueing Network Models. [2]

About:

GitHub repository

Contents

1 Hardware Infrastructures

1.1 System-level

1.1.1 Computing Infrastructures and Data Center Architectures

There's no single definition of a Data Center, but it can be summarized as follows.

Definition 1

Data Centers are buildings where multiple servers and communication gear are co-located because of their common environmental requirements and physical security needs, and for ease of maintenance. [1]

Definition 2

A **Computing Infrastructure** (or IT Infrastructure) is a technological infrastructure that provides hardware and software for computation to other systems and services.

A number of computing infrastructures exist:

- **Cloud** offers virtualized computing, storage and network resources with highly-elastic capacity.
- **Edge Servers** are on-premises hardware resources that perform more compute-intensive data processing.

  In other words, an edge server is a piece of hardware that performs data computation at the end (or "edge") of a network. Like a regular server, an edge server can provide compute, networking, and storage functions.[1]
- **IoT** and **AI-enabled Edge Sensors** are hardware devices where the data acquisition and partial processing can be performed at the edge of the network.

[1] More info here.

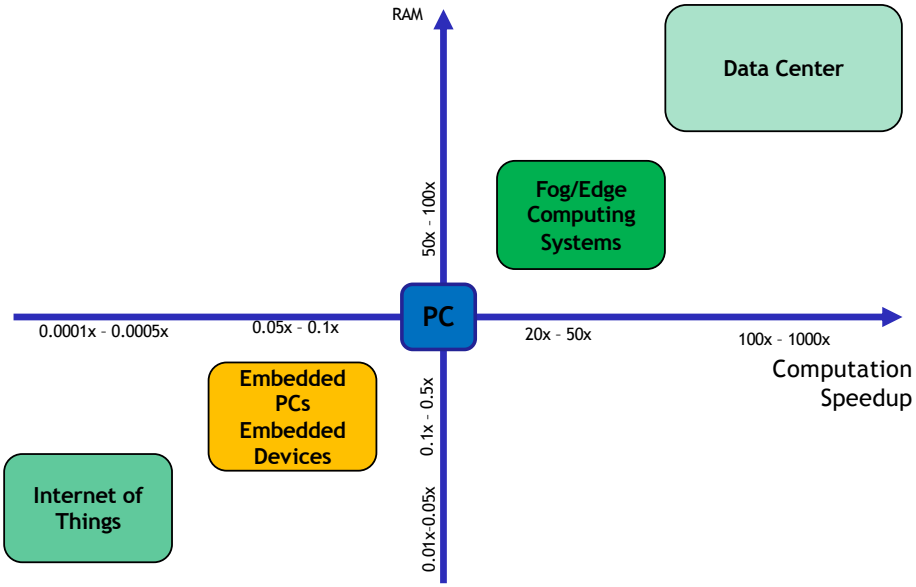


Figure 1: An **example** of Computing Infrastructures. [4]

The **Computing Continuum**, a novel paradigm that extends beyond the current silos of cloud and edge computing, can enable the seamless and dynamic deployment of applications across diverse infrastructures. [3]

The Computing Continuum

IoT/Edge — Fog — HPC/Cloud/Instrument

Size	Nano	Micro	Milli	Server	Fog	Campus	Facility
Example	Adafruit Trinket	Particle.io Boron	Array of Things	Linux Box	Co-located Blades	1000-node cluster	Datacenter
Memory	0.5K	256K	8GB	32GB	256G	32TB	16PB
Network	BLE	WiFi/LTE	WiFi/LTE	1 GigE	10GigE	40GigE	N*100GigE
Cost	$5	$30	$600	$3K	$50K	$2M	$1000M

Figure 2: The Computing Continuum. [4]

In the following pages, we analyze the computing infrastructures mentioned in the previous example.

Data Centers

The definition of a Data Centers can be found on page 4.

✔ Data Centers Advantages

- **Lower IT costs.**
- **High Performance.**
- **Instant software updates.**
- **"Unlimited" storage capacity.**
- **Increased data reliability.**
- **Universal data access.**
- **Device Independence.**

👎 Data Centers Disadvantages

- **Require a constant internet connection.**
- **Do not work well with low-speed connections.**
- **Hardware Features might be limited.**
- **Privacy and security issues.**
- **High power Consumption.**
- **Latency in taking decision.**

Internet-of-Things (IoT)

An **Internet of Things (IoT) device is any everyday object embedded with sensors, software, and internet connectivity**.

This allows to collect and exchange data with other devices and systems, typically over the internet, with limited need of process and store data.

Some **examples** are Arduino, STM32, ESP32, Particle Argon.

✔ **Internet-of-Things Advantages**

- **Highly Pervasive**.
- **Wireless connection**.
- **Battery Powered**.
- **Low costs**.
- **Sensing and actuating**.

👎 **Internet-of-Things Disadvantages**

- **Low computing ability**.
- **Constraints on energy**.
- **Constraints on memory (RAM/FLASH)**.
- **Difficulties in programming**.

Embedded (System) PCs

An **Embedded System** is a computer system, a combination of a computer processor, computer memory, and input/output peripheral devices, that has a dedicated function within a larger mechanical or electronic system.

A few **examples**: Odroid, Raspberry, jetson nano, Google Coral.

✔ **Embedded System Advantages**

- **Persuasive computing**.
- **High performance unit**.
- **Availability of development boards**.
- **Programmed as PC**.
- **Large community**.

👎 **Embedded System Disadvantages**

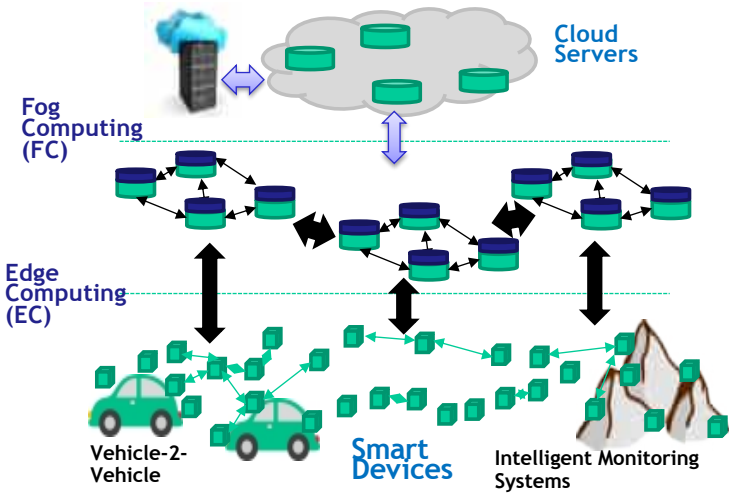
- **Pretty high power consumption**.
- **(Some) Hardware design has to be done**.

Edge/Fog Computing Systems

The key **difference** between **Fog Computing** and **Edge Computing** is associated with the location **where the data is processed**:

- In **edge computing**, the data is processed closest to the sensors.
- In **fog computing**, the computing is moved to processors linked to a local area network (IoT gateway).

Edge computing places the intelligence in the connected devices themselves, whereas, fog computing puts in the local area network.



✔ Fog/Edge Advantages

- **High computational capacity.**
- **Distributed computing.**
- **Privacy and security.**
- **Reduced Latency in making a decision.**

👎 Fog/Edge Disadvantages

- **Require a power connection.**
- **Require connection with the Cloud.**

Feature	Edge Computing	Fog Computing
Location	Directly on device or nearby device.	Intermediary devices between edge and cloud.
Processing Power	Limited due to device constraints, sending data to central server for analysis.	More powerful than edge devices. However, sending data to a central server for analysis.
Primary Function	Real-time decision-making, low latency. However, central server analyzing combined data and sending only relevant information further.	Pre-process and aggregate data, reduce bandwidth usage. However, central server analyzing combined data and sending only relevant information further.
Advantages	Low latency, reduced reliance on cloud, security for sensitive data.	Bandwidth efficiency, lower cloud costs, complex analysis capabilities.
Disadvantages	Limited processing power, single device focus.	Increased complexity, additional infrastructure cost.

Table 1: Differences between Edge and Fog Computing Systems.

References

[1] L.A. Barroso, U. Hölzle, and P. Ranganathan. *The Datacenter as a Computer: Designing Warehouse-Scale Machines, Third Edition*. Synthesis Lectures on Computer Architecture. Springer International Publishing, 2022.

[2] E.D. Lazowska. *Quantitative System Performance: Computer System Analysis Using Queueing Network Models*. Prentice-Hall, 1984.

[3] Jacopo Marino and Fulvio Risso. Is the computing continuum already here?, 2023.

[4] Gianluca Palermo. Lesson 1, computing infrastructures. Slides from the HPC-E master's degree course on Politecnico di Milano, 2024.

Index