

Computing Infrastructure: [Course Code]

[Sofia Martellozzo]

Academic Year 2021-2022

Contents

- 1 Computing Infrastructure 3**
 - 1.1 Introduction 3
- 2 Data WareHouse 4**
 - 2.1 Introduction 4
 - 2.2 From Data Centers to Warehouse-scale computers 4
 - 2.3 Architectural Overview of WSCs 6

1 Computing Infrastructure

1.1 Introduction

Definition 1. Computing Infrastructure: *Technological infrastructure that provides hardware and software for computation to other systems and services.*

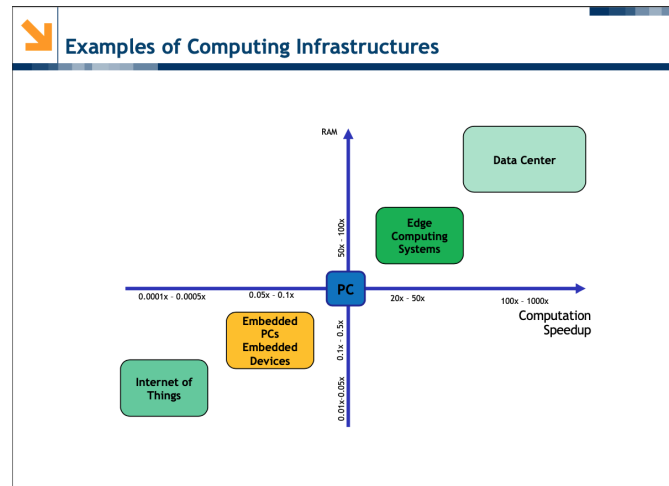


Figure 1: Computing Infrastructure

Advantages:

- Lower IT costs
- High performance
- Instant software updates
- “Unlimited” storage capacity
- Increased data reliability
- Universal document access
- Device Independence

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 (1% overall worldwide total energy consumption due to datacenters)
- Latency in making decision

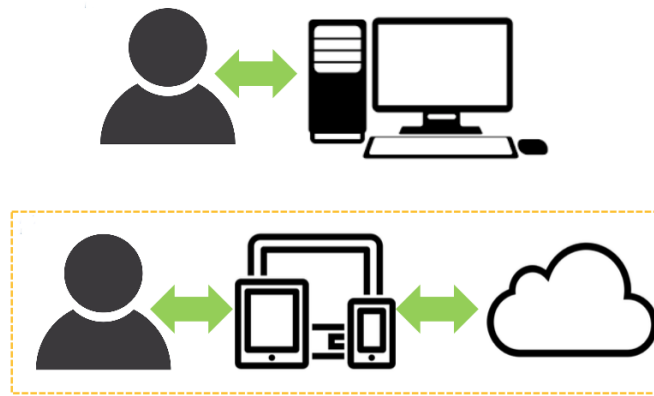
Amortized Cost	Component	Sub-Components
~45%	Servers	CPU, memory, disk
~25%	Infrastructure	UPS, cooling, power distribution
~15%	Power draw	Electrical utility costs
~15%	Network	Switches, links, transit

2 Data WareHouse

2.1 Introduction

In the last few decades, computing and storage have moved from PC- like clients to smaller, often mobile, devices, combined with large internet services.

Traditional enterprises are also shifting to Cloud computing.



User experience improvements

- Ease of management (no configuration or backups needed)
- Ubiquity of access

Advantages to vendors

- Software-as-a-service allows faster application development (easier to make changes and improvements)
- Improvements and fixes in the software are easier inside their data centers (instead of updating many millions of clients with peculiar hardware and software configurations)
- The hardware deployment is restricted to a few well-tested configurations.

Server-side computing allows

- Faster introduction of new hardware devices (e.g., HW accelerators or new hardware platforms)
- Many application services can run at a low cost per user.

Some workloads require so much computing capability that they are a more natural fit in datacenter (and not in client-side computing).

A couple of examples (Search services (web, images, and so on), Machine and Deep Learning (GPT-3)).

2.2 From Data Centers to Warehouse-scale computers

Data centers = is a place in which there are many servers

Wharehouse = is a type of Data Center, it works as a computer.

The trends toward server-side computing and widespread internet services created a new class of computing systems:

Definition 2. warehouse-scale computers (WSCs) *The massive scale of the software infrastructure, data repositories, and hardware platform.*

- *is an internet service (= service provided by inyneternet)*
- *may consist of tens or more individual programs (= not a single program, a collection of them that together create/provide the service)*
- *such programs interact to implement complex end-user services such as email, search, maps or machine learning.*

Data centers are buildings where multiple servers and communication units are co-located because of their common environmental requirements and physical security needs, and for ease of maintenance.

In **Traditional Data Center**: typically host a large number of relatively small- or medium-sized applications, each applications is running on a dedicated hardware infrastructure that is de-coupled and protected from other systems in the same facility, applications tend not to communicate each other. Those data centers host hardware and software for multiple organizational units or even different companies.

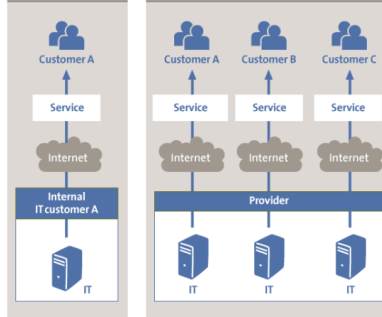


Figure 2: Traditional Data Center

WSCs belong to a single organization, use a relatively homogeneous hardware and system software platform (=, easier to manage and cheaper, but has limitations on functionalities), and share a common systems management layer (such as Google, Facebook, Alibaba, Amazon, Dropbox...).

(you have 1 services that you want to provide to a 'huge' amount of customers)

Run a smaller number of very large applications (or internet services).

The common resource management infrastructure allows significant deployment flexibility.

The requirements of:

- homogeneity
- single-organization control
- cost efficiency

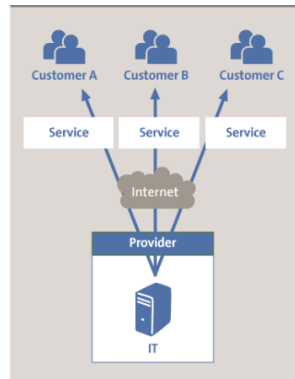


Figure 3: Warehouse-Scale Computers

Initially designed for online data-intensive web workloads, WSCs also now power public clouds computing systems (e.g., Amazon, Google, Microsoft). Such public clouds do run many small applications, like a traditional data center. (All of these applications rely on Virtual Machines (or Containers), and they access large, common services for block or database storage, load balancing, and so on, fitting very well with the WSC model).

These are not just a collection of servers:

The software running on these systems executes on clusters of hundreds to thousands of individual servers (far beyond

a single machine or a single rack)

+

The machine is itself this large cluster or aggregation of servers and needs to be considered as a single computing unit. (=, scale up in terms of performance)

Several datacenters: Multiple Data Center located far apart (placed near point of interest) =, becomes important also the **PRIVACY**: data of a country must remain in it.

Multiple data centers are (often) replicas of the same service (to reduce user **latency** and improve serving **throughput**).

A request is typically fully processed within one data center.

Availability: Services provided through WSCs must guarantee high availability, typically aiming for at least 99.99% uptime (i.e., one-hour downtime per year). Achieving such fault-free operation is difficult when a large collection of hardware and system software is involved.

WSC workloads must be designed to gracefully tolerate large numbers of component faults with little or no impact on service level performance and availability.

2.3 Architectural Overview of WSCs

Hardware implementation of WSCs might differ significantly each other; However, the architectural organization of these systems is relatively stable.

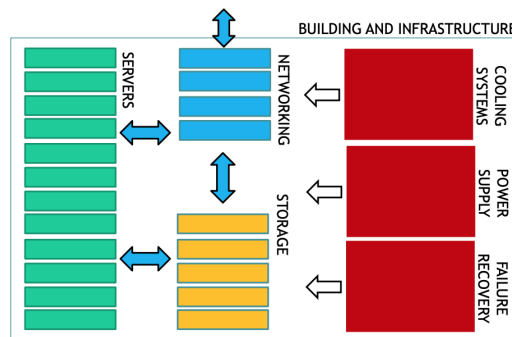


Figure 4: Warehouse-Scale Computers Overview