

ICT Infrastructure - Appunti

Francesco Lorenzoni

February 2024

Contents

- 1 Introduction 7**
 - 1.1 Course map 7
- 2 Datacenter 9**
 - 2.1 Structure 9
 - 2.2 Power Management 9
 - 2.3 Cooling 10
 - 2.4 CRAC 10
 - 2.5 Inrow Cooling 11

Course info

Don't be shy to send multiple emails, prof. Cisternino receives many emails, and he knows he can't reply to each one. He is okay to be contacted through teams using the @ symbol to "mention" him.

He designed the UniPi datacenters.

"Italy is more about the multiple micro businesses than the few existing industries"

Exam

The exam is **oral**.

Prof. Cisternino expects students to get the full picture, and understand key concepts, not to remember everything—which still wouldn't be bad ☺—.

Chapter 1

Introduction

Prof. Cisternino dropped a lot of measures in terms of Watts, Dollars, Gigabits and so on.

He mentioned with emphasis the problem of energy consumption. To give an idea, a single rack of a datacenter designed ~ 10 years ago, absorbs up to $15kW$. The datacenter in *San Piero a Grado* is made up of 60 racks. It is not meant to provide the maximum energy possible for all racks simultaneously, but it still helps to get an idea of how things work in similar contexts.

1.1 Course map

1. Elements
 - i. Datacenters
 - (a) Power
 - (b) Cooling
 - ii. Cabling
 - iii. Networking
 - iv. Storage
 - v. Compute
 - vi. Virtualization
 - (a) Hypervisor
 - (b) Containers
2. Cloud
 - i. Reference architecture
 - ii. Resilience
 - iii. Security
 - iv. Legal aspects
 - (a) GDPR
 - (b) Security frameworks
 - v. Procurement aspects
 - vi. Operations
 - i.e. Keep the system up and running while upgrading the system

Chapter 2

Datacenter

10 years ago datacenters were no more than a room with some computers, air conditioners and some plugs to power up the devices. Later on, customers started asking server vendors to include in the servers utilities to allow an *automated datacenter management*. Thus the trend moved towards Software Defined Datacenter, which currently is the only possible way to deploy a Datacenter.

An **Active Datacenter** allows for internet storage (?)

A Datacenter should be **future-proof**: servers may be replaced, but updating a whole datacenter is at least a 1-year project.

2.1 Structure

Racks are made of —~ 42— *units*.

Besides server themselves, there is a **cooling system**. The first issue is the how to provide cool air. Then there is also how to define an evacuation plan, which must take into account dust.

However also the **floor** is not to be neglected.

- ◇ *Floating floor* or Ground floor
 - “A “floating floor” in a data center, also known as a “raised floor”, is a type of construction used in data centers to create a void between the actual concrete floor and the floor tiles where the servers and other equipment are located¹². This space is typically used for routing cables and for air circulation, which helps with cooling the equipment¹.”¹
- ◇ *Resistance* usually around $1 \frac{\text{ton}}{\text{m}^2}$

For example, in San Piero A Grado, there was a power cabin receiving current from three lines. Now the whole power management components are in a container outside the building placed close to the facility.

Cables are not super-resistant to current. A lot of current passing through a copper wire will *exhaust* both the wire and the components receiving such current; hence the current should also be balanced among different cables, to avoid exhausting some components before the others.

A UPS —first of all— stabilizes the output current.

In theory $1V * 1A = 1W$, but in reality, performing such conversion something gets lost, so we have

$$I * V * \cos\phi = W$$

2.2 Power Management

Electric panels (aka *switchboards*) allow segmenting the power supply in the various zones of the datacenter.

PDUs stands for *Power Distribution Units*, and allow to distribute power for a server units in a rack. Typically, for each PDU there is another one, providing redundancy and thus resilience/robustness.

Definition 2.1 (PUE) *Power Usage Effectiveness* **PUE** measures the efficiency of a Datacenter.

$$PUE = \frac{\text{Total energy}}{\text{ICT energy}}$$

¹ChatGPT 4.0 - Generated

The reason for improving Datacenter design is to lower the PUE; basically to save money, but also for “green-environment” concerns.

But when should PUE be measured?

The PUE in January is very different from the one in August, so generally it is calculated as the average of one year.

Note that a poorly designed datacenter placed in Siberia with -20° may have a lower PUE than a datacenter in Italy, for instance.

In particulare geographical zones with high temperature variations over the year (e.g. in Italy the temperature variates from 40 to 50 Celsius degrees), are strongly unrecommended to build datacenters in. A counterintuitive example is the desert, where the temperature is very high during the day and very low during the night, but in general the temperature over the year is **stable**; allowing for defining physical processes exploiting such stability. Also the oceans have a very stable temperature; not on the surface, but deep down it is very stable.

2.3 Cooling

Note that racks are always placed back-to-back, because the front requires cool air, and the back outputs hot air.

2.4 CRAC



Figure 2.1: CRAC/CRAH cooling architecture

Chillers take hot air from above and push cool air in the bottom. Then air pushed under the floating floor wants to exit, and does so going through the grates placed in front of the racks. The racks suck the cool air in front and output hot air from the back.

There are two **drawbacks**:

1. It is difficult to confine and keep separated hot air and cool air. The mixup between the two leads to cooling inefficiency, thus energy and money waste
2. In case a rack has a workload heavier than others and thus requires more cooling air, the chiller must provide more cool air to all the racks in the same row; this makes this architecture particularly inefficient for datacenter which have heterogeneous workloads.

2.5 Inrow Cooling

The fan “towers” are called *inrow cooling*.

The first advantage is that it allows for heterogeneous cooling in the datacenter. Secondly, a fan outputs hot air directly where another fan expects it to be. This allows to confine hot air and to avoid wasting energy in outputting air and sucking it.

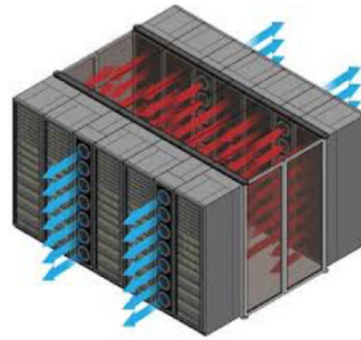


Figure 2.2: Inrow cooling architecture