Looking to the Future in Domains of Microservices, Cloud and Edge Computing: *Osmotic Computing*

PROF. MASSIMO VILLARI

HEAD OF FUTURE COMPUTING RES LAB (FCRLAB) UNIVERSITÀ DI MESSINA

http://fcrlab.unime.it/



Our Present is ALREADY Part of Our Future

Our past determines our present, and our present is what shapes our future.

- unknown



Pisa - 24/FEB/17

PROF. MASSIMO VILLARI PHD



Outline: Part I

Where we are:

- Cloud Computing
- FOG Computing
- Edge Computing

• IoT

•

• Big Data: it is the driver



Outline: Part I

Where we are, in particular:

- MicroServices and Containerization
- Social Platforms
- Serverless
- APIs for any taste:
 - RESTFul
 - CoAP



Outline: Part II

Consolidated Activities:

- Cloud Computing:
- FOG Computing
- IoT
- Big Data



Outline: Part II

Where are we going??:

- Current Trends
- New Devices
- Needs
- •

• A new concept: Osmotic Computing



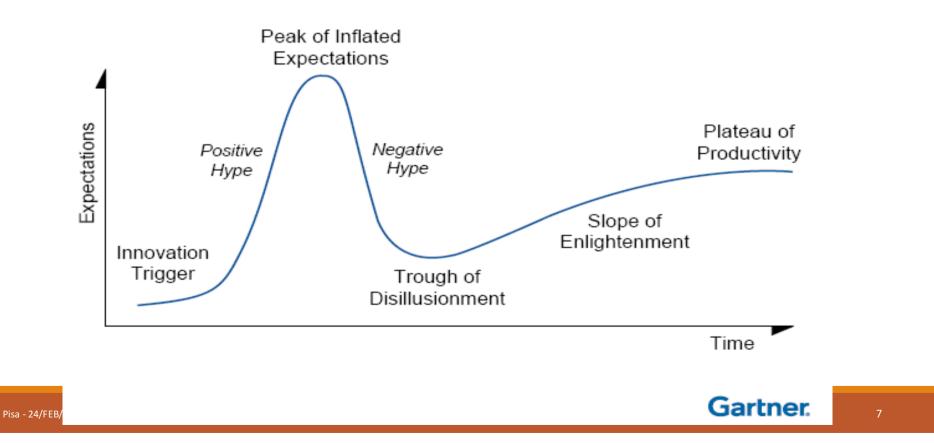
PART I

Pisa - 24/FEB/17

PROF. MASSIMO VILLARI PHD

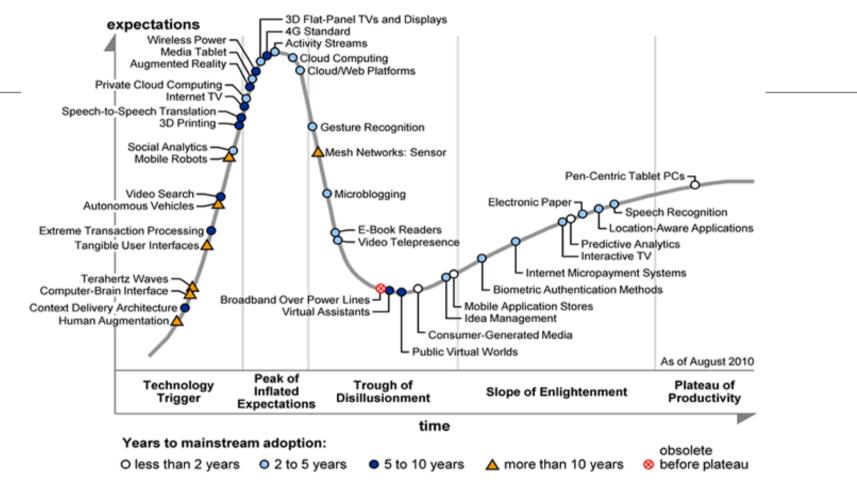


Simple graphical approach to get a better understanding: *Gartner hype cycle*





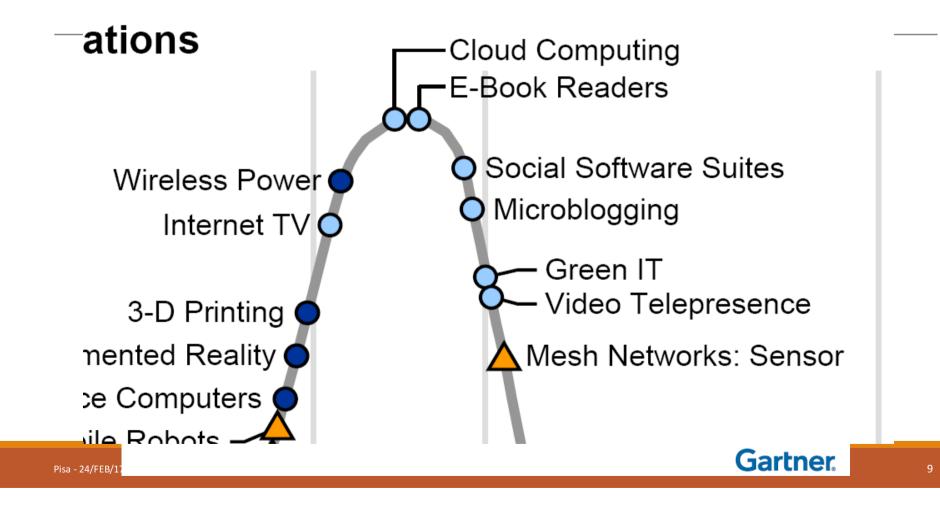
Hype Cycle: 2010



PROF. MASSIMOVILLARI PHD



Inside the Hype Cycle: What's Hot and What's Not in 2009

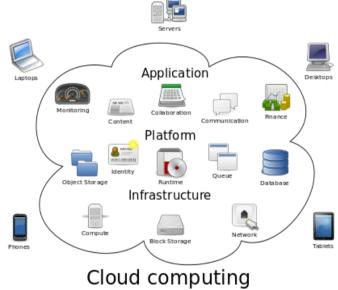




Cloud Computing on the wiki

Cloud computing

is a type of Internet-based computing that provides shared computer processing resources and data to computers and other devices on demand. It is a model for enabling ubiquitous, on-demand access to a shared pool of configurable computing resources (e.g., computer networks, servers, storage, applications and services), which can be rapidly provisioned and released with minimal management effort. Cloud computing and storage solutions provide users and enterprises with various capabilities to store and process their data in either privately owned, or third-party data centers that may be located far from the user-ranging in distance from across a city to across the world. Cloud computing relies on sharing of resources to achieve coherence and economy of scale, similar to a utility (like the electricity grid) over an electricity network.





Why Cloud? THE WALL STREET JOURNAL USINESS & FINANCE Europe Ups the Ante In Battle for Internet European official says new actions needed to "shock" digital world Firms 's FCC In France: "Les European official says new actions needed to 'shock' digital world. Gafa" Google, Apple, Facebook and Amazon Assets an proposal—while lacking fic recommendations—adds heft to long-simmering de-Online privacy also has been a flash point between U.S. firms and European officials, reflecting the deeply different approaches to data protection. European officials look at the protection of personal data as a fundamental right, whereas the U.S. puts more emphasis on freedom of expression and currently regu-lates privacy mostly as a consumer-protection issue Online privacy also has been in Paris and Berlin for rules n. ads in Paris and Berlin for Tuest to be, rein in the growing influence of a cadre of largely American tech firms. In France, officials frequently refer to "les Gafa," an acronym for Google, Apple, Facebook and Ama-These Silicon Valley companies Google, Apple, Facebook and Anar zon, when discussing the power of big Internet companies. Over the summer, a French advi-sory board proposed a new concept called "platform neutrality" that inprotection issue. have hide Those two approaches have been most dramatically opposed in Eu-rope's new "right to be forgotten," established in a court decision earcalled "platform neutrality" that in-cludes ideas such as imposing in-teroperability among web platforms to let users easily switch services. Four German ministers recently echoed those calls in a letter to the **European Commission**, the EU's ex-ecutive arm, that also supported the idea of stronger antitrust powers for the digital era. "We need powers of sanctions " revenue and lier this year, that gives individuals

eavily in the

€11.5 billion

"We need powers of sanctions," Germany's state secretary for the economy, Matthias Machnig, said

Thursday in an appearance with Axelle Lemaire, France's deputy minister for digital affairs. "Our ex-

pean privacy regulators on Wednes

day issued new guidelines saying

search engines should apply the rul-ing to all of their websites in partic-

ular websites such as Google.com

global reach,

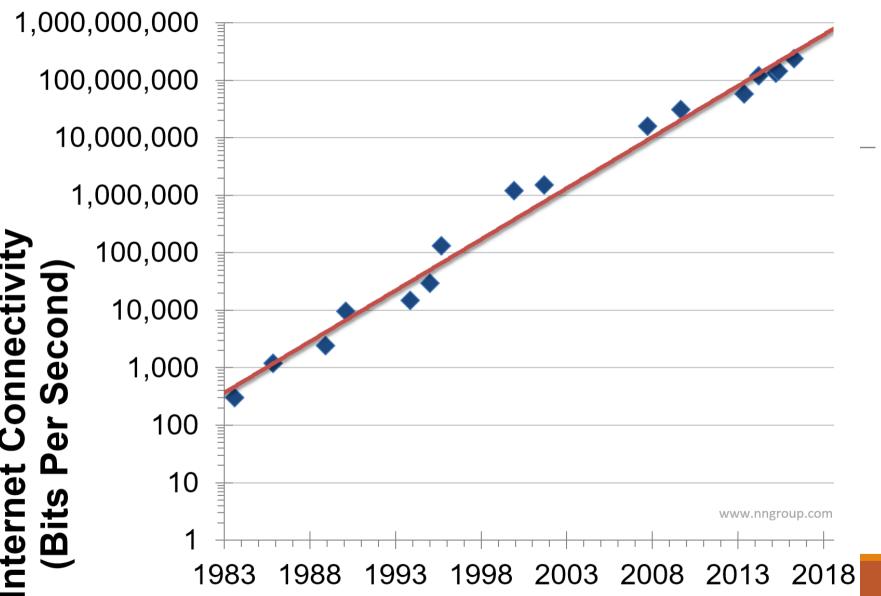
corporate

Prof Massimo Villari - Bratislava Oct 15th

but they pay little



nternet Connectivity (Bits Per Second)







FOG Computing on the Wiki

Fog computing can be perceived both in large cloud systems and big data structures, making reference to the growing difficulties in accessing information objectively.

This results in a lack of quality of the obtained content.

The effects of fog computing on cloud computing and big data systems may vary; yet, a common aspect that can be extracted is a limitation in accurate content distribution, an issue that has been tackled with the creation of metrics that attempt to improve accuracy.

Fog networking consists of a **control plane and a data plane**. For example, on the data plane, fog computing enables computing services to reside at the edge of the network as opposed to servers in a data-center.

PROF. MASSIMOVILLARI PHD



FOG Computing on the Wiki

Compared to cloud computing, fog computing emphasizes proximity to end-users and client objectives, dense geographical distribution and local resource pooling, latency reduction for quality of service (QoS) and edge analytics/stream mining, resulting in superior user-experience and redundancy in case of failure.

Fog networking supports the Internet of Things (IoT) concept, in which most of the devices used by humans on a daily basis will be connected to each other. Examples include phones, wearable health monitoring devices, connected vehicle and augmented reality using devices such as the Google Glass.

ISO/IEC 20248 provides a method whereby the data of objects identified by edge computing using Automated Identification Data Carriers [AIDC], a barcode and/or RFID tag, can be read, interpreted, verified and made available into the "Fog" and on the "Edge" even when the AIDC tag has moved on.

Pisa - 24/FEB/17

PROF. MASSIMOVILLARI PHD



Open FOG Architecture:

- Security;
- Scalability;
- Open; Autonomy;
- Programmability;
- RAS (Reliability, Availability, and Serviceability);
- Agility; and
- Hierarchy.



Edge Computing on the Wiki

Edge computing pushes applications, data and computing power (services) away from centralized points to the logical extremes of a network. Edge computing replicates fragments of information across distributed networks of web servers, which may be vast. As a topological paradigm, edge computing is also referred to as **mesh computing, peer-to-peer computing, autonomic (self-healing)** <u>computing, grid computing</u>, and other names implying non-centralized, nodeless availability.

To ensure acceptable performance of widely dispersed distributed services, large organizations typically implement edge computing by deploying Web <u>server farms</u> with <u>clustering</u>. Previously available only to very large corporate and government organizations, technology advancement and cost reduction for large-scale implementations have made the technology available to small and medium-sized businesses.

The target end-user is any Internet client making use of commercial Internet application services.

Edge computing imposes certain limitations on the choices of technology platforms, applications or services, all of which need to be specifically developed or configured for edge computing.



Edge Computing on the Wiki

Edge computing has many advantages:

Edge application services significantly decrease the data volume that must be moved, the consequent traffic, and the distance the data must go, thereby reducing transmission costs, shrinking latency, and improving <u>quality of service</u> (QoS).

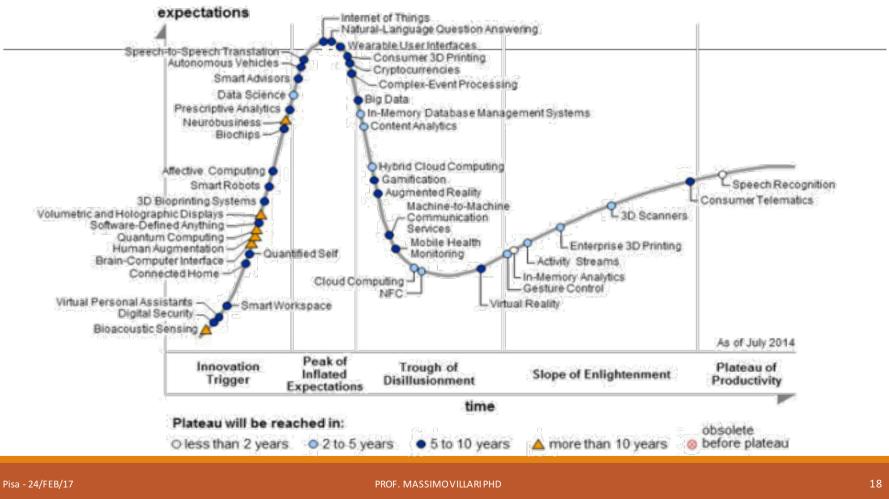
Edge computing eliminates, or at least de-emphasizes, the core computing environment, limiting or removing a major bottleneck and a potential point of failure.

Security is also improved as encrypted data moves further in, toward the network core. As it approaches the enterprise, the data is checked as it passes through protected firewalls and other security points, where viruses, compromised data, and active hackers can be caught early on.

Finally, the ability to "virtualize" (i.e., logically group CPU capabilities on an as-needed, real-time basis) extends <u>scalability</u>. The edge computing market is generally based on a "**charge for network services**" **model**, and it could be argued that typical customers for edge services are organizations desiring linear scale of business application performance to the growth of, e.g., a subscriber base.

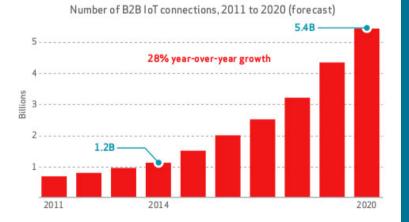


Hype Cycle: 2014





Internet of Things



IIoT DEVELOPERS WANTED

Experts have been throwing out numbers and estimates for awhile now about the future of the Industrial Internet of Things, usually focusing on projected total revenue or total number of connected devices. VisionMobile has released a report that estimates the actual people that will get involved in the massive growth of Internet of Things.

This latest statistic details how there the workforce could sport up to 1.5 million IoTspecific developers by 2016, and an estimated 4.5 million by 2020. That's a huge influx of professionals into a relatively new industry, and there's no doubt that companies will have to adapt and nurture the techcentric communities that currently exist within their walls to keep up with the growing demands and advancing technology of the IoT. Developers Split by Primary Audience Worldwide Developer Population 8.8mm Enterprise

Not sure

Other

Source: VisionMobil

CENTRIC DIGITAL

Digital Trends Daily

8.7mm

1.5mm

Mobile IoT

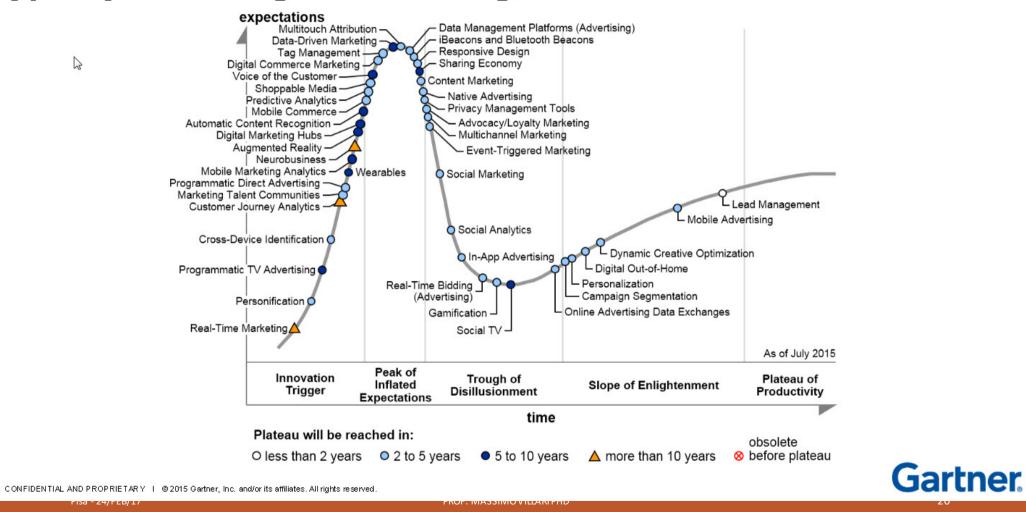
Other software

Source: Evans Data



5

Hype Cycle for Digital Marketing, 2015

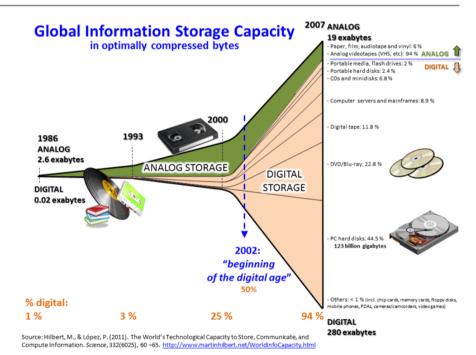




Data

Big Data Challenges:

include <u>capture</u>, <u>storage</u>, <u>analysis</u>, <u>data curation</u>, search, <u>sharing</u>, <u>transfer</u>, <u>visualization</u>, <u>querying</u>, updating and <u>information privacy</u>. The term "big data" often refers simply to the use of <u>predictive</u> <u>analytics</u>, <u>user behavior analytics</u>, or certain other advanced data analytics methods that extract value from data, and seldom to a particular size of data set.



PROF. MASSIMO VILLARI PHD



Microservice on the Wiki

In a microservices architecture, services should have a small granularity and the protocols should be lightweight. A central microservices property that appears in multiple definitions is that services should be independently deployable.The benefit of distributing different responsibilities of the system into different smaller services is that it enhances the cohesion and decreases the coupling. This makes it easier to change and add functions and qualities to the system at any time. It also allows the architecture of an individual service to emerge through continuous refactoring, and hence reduces the need for a big up-front design and allows for releasing software early and continuously.



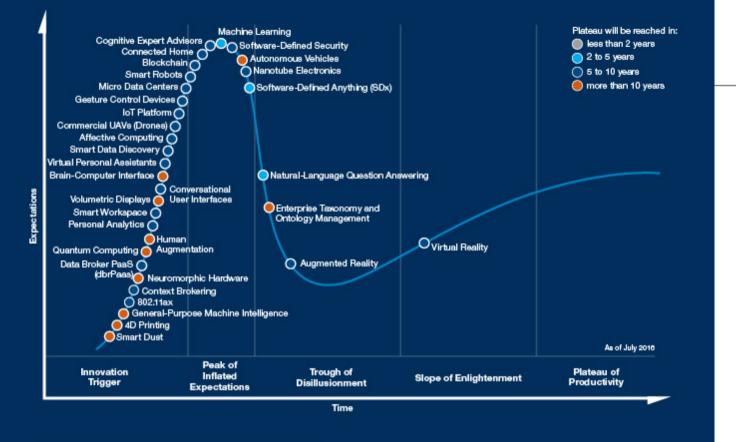
PART II

Pisa - 24/FEB/17

PROF. MASSIMOVILLARI PHD



Gartner Hype Cycle for Emerging Technologies, 2016



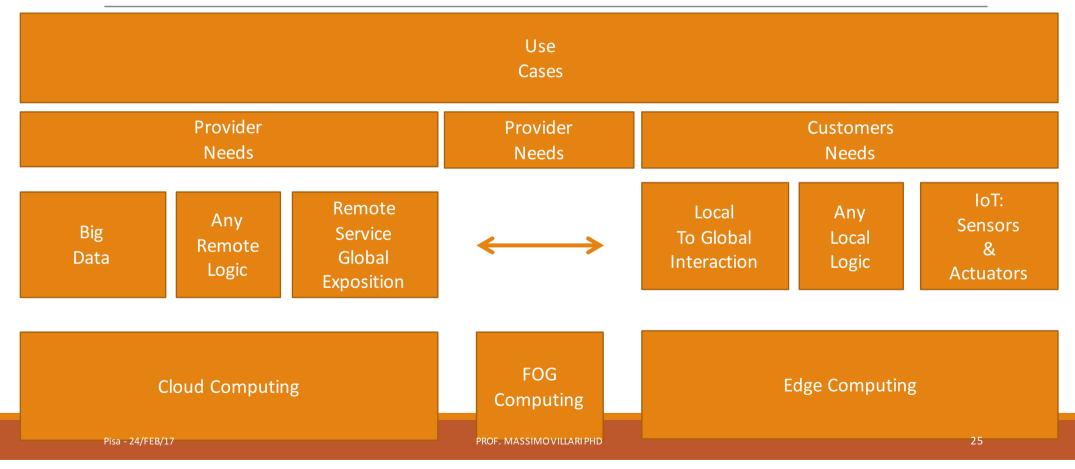
gartner.com/SmarterWithGartner

Source: Gartner © 2016 Gartner, Inc. and/or its affiliates. All rights reserved

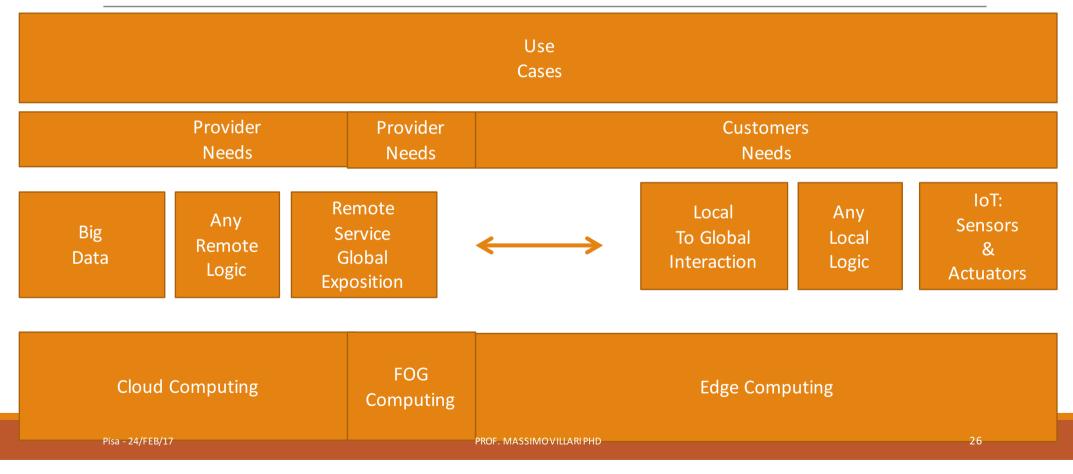


Pisa - 24/FEB/17

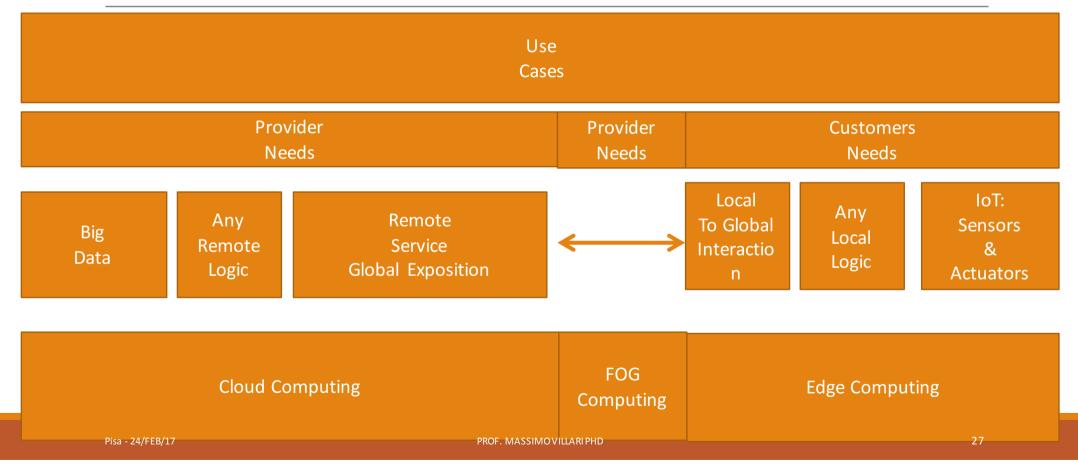




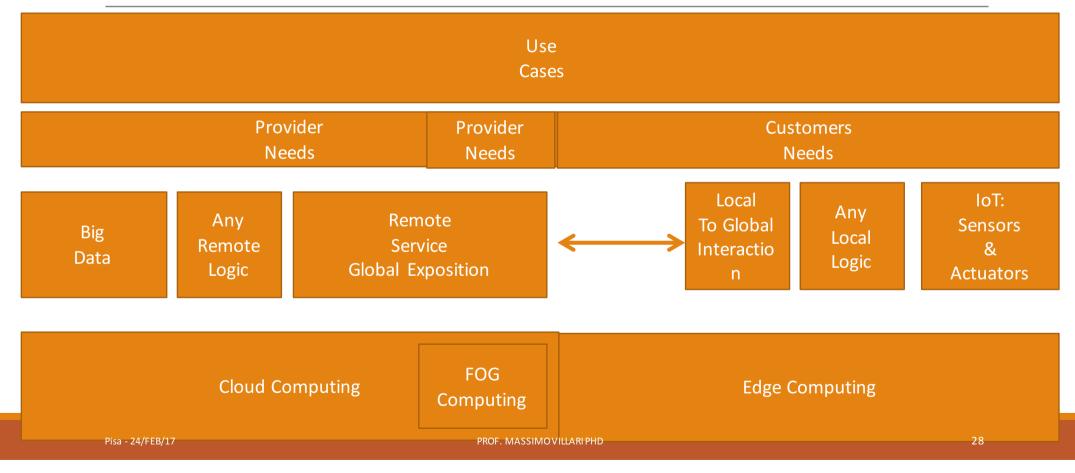




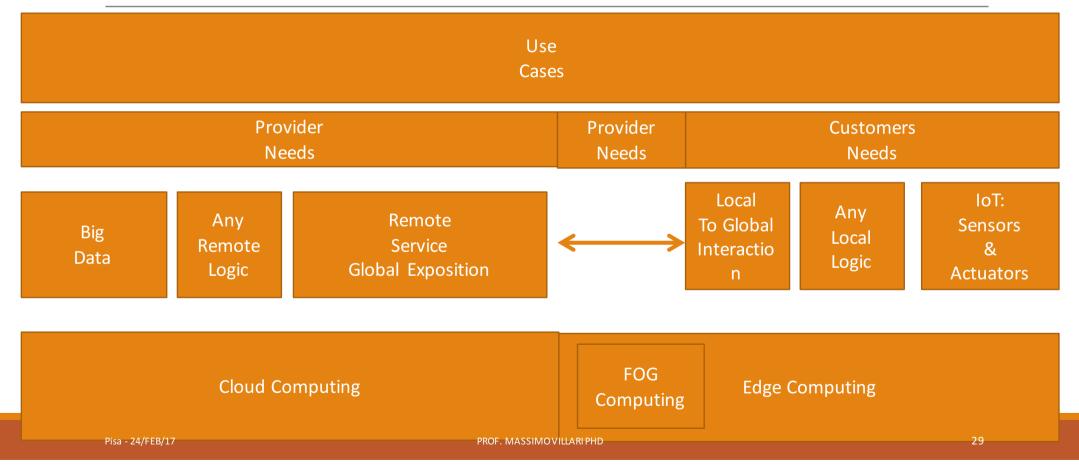


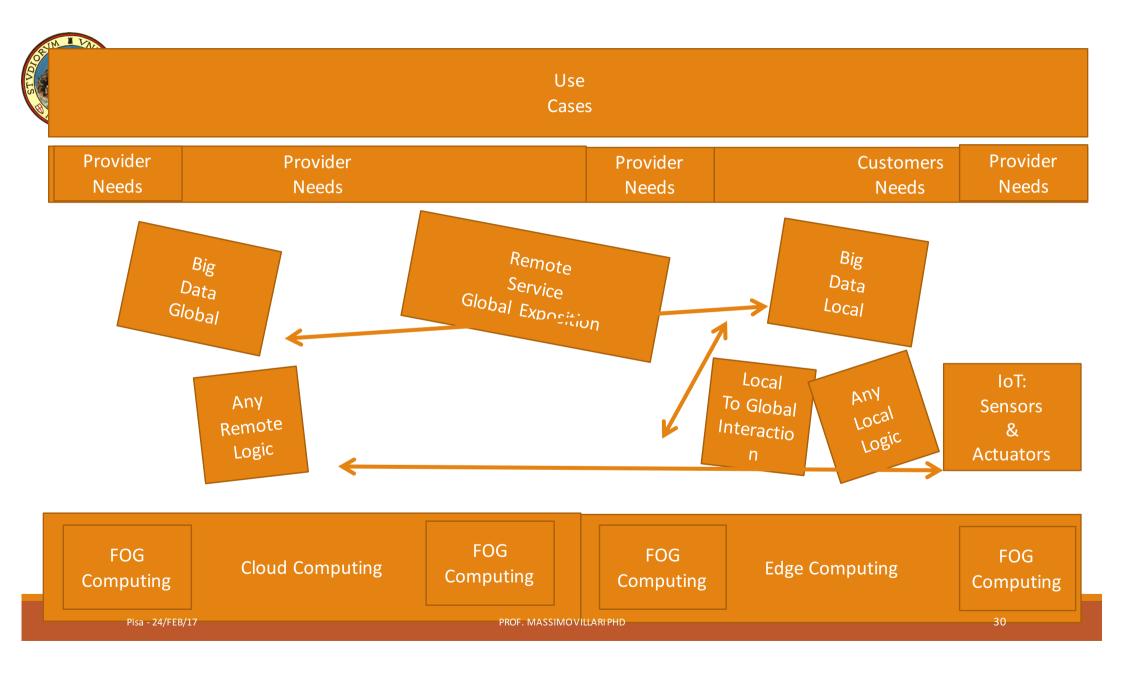






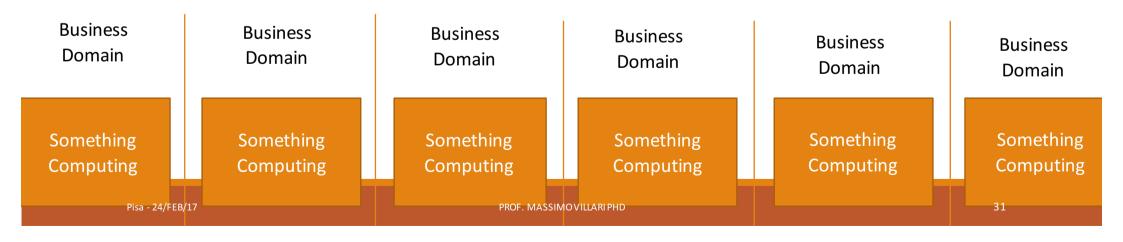






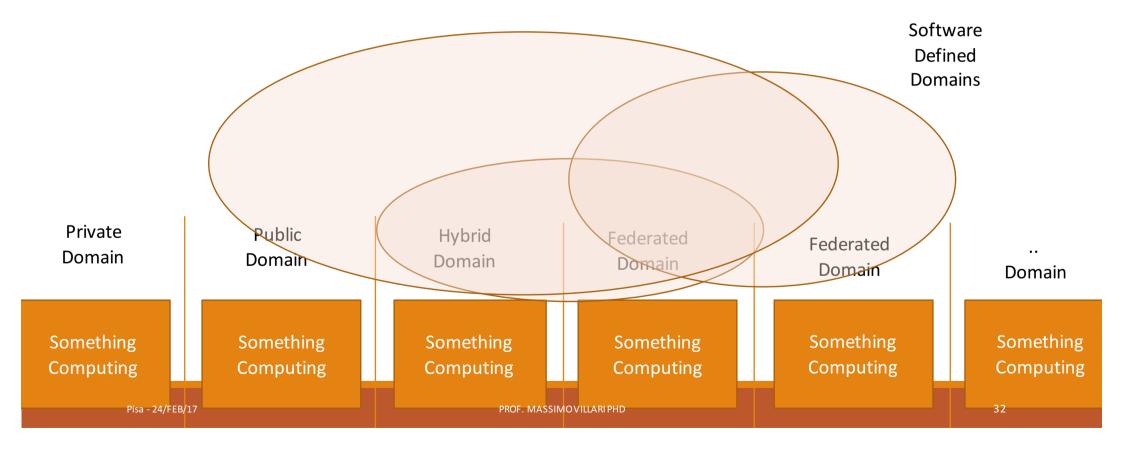


Something Computing -> Software Defined Anything





Something Computing -> Software Defined Anything





Something Computing -> Software Defined Anything

In concrete:

- Define a Set of APIs
- Migrability of Services (Micro is better)
- Bounds for Security (No Breaches-> if [yes] delete service
- A priory agreements



Looking at Devices

Smarter?:

- Even more Abstracted
- Even more Capable (Proc. / Storage / Net.)
- Even more Multiple Cores



What devices?: MPU or MCU MPUs offer more functionality and faster time to market, while MC<u>Us provide a smaller, more cost-efficient solution.</u>

MPU:

MPUs generally use open-source operating systems such as Linux and Android, although high-reliability applications may require proprietary operating systems like those from Green Hills Software and Wind River. These operating systems include a library of drivers, for example for codecs, Ethernet, USB, etc. Most MPU providers create and maintain both Linux and Android releases for their evaluation boards, often supporting them with large software teams. Full-fledged operating systems make the development of software much more simple and lend themselves to "plug and play" hardware.

MCU:

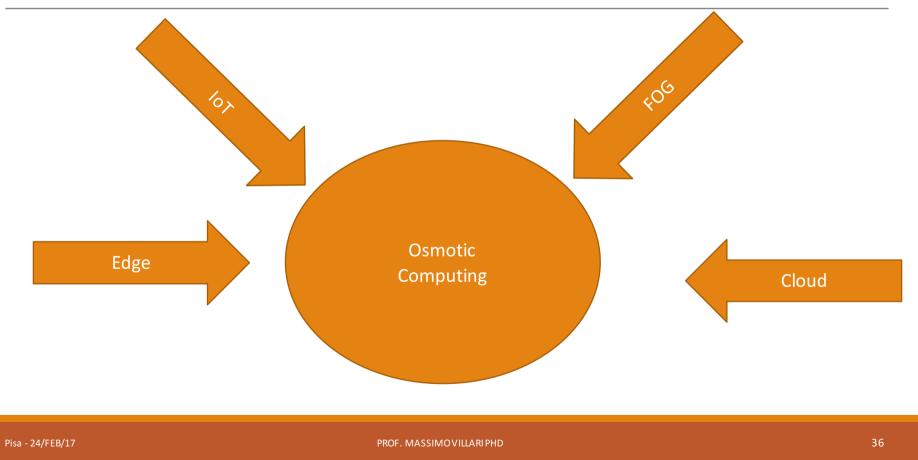
An MCU delivers the right balance of cost, size, efficiency, and reliability. Vendors leverage modular platforms and value pricing to create multiple part numbers from one die. As a result, designers have a range of options in terms of memory size, pin count, and peripherals to develop the most cost-effective solution. MCUs are sometimes referred to as SoCs because of the large amount of built-in functionality. They typically integrate reset functionality, low-voltage inhibit, clock sources, interrupts, and on-chip regulators, to name a few.

The CPU cores used in MCUs are designed explicitly to deliver the very low interrupt latency and deterministic code execution required for motor control in real-time applications like fans, compressors, washing machines, etc. The Cortex-M4, for example, has a worst-case interrupt latency of 12 cycles and uses a three-stage pipeline.

Pisa - 24/FEB/17

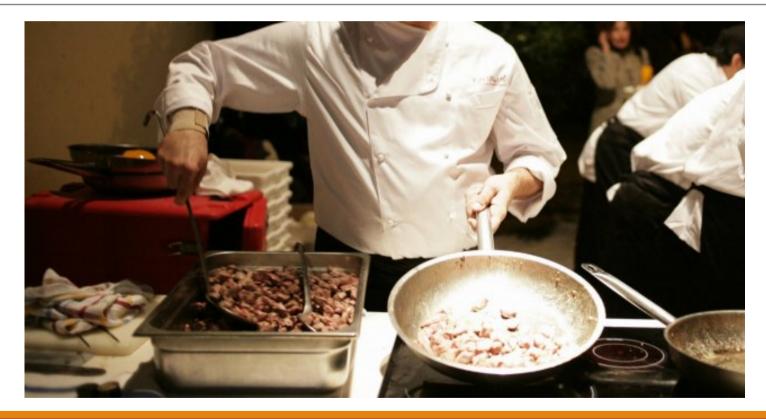


Looking at the Overall Picture: E2E





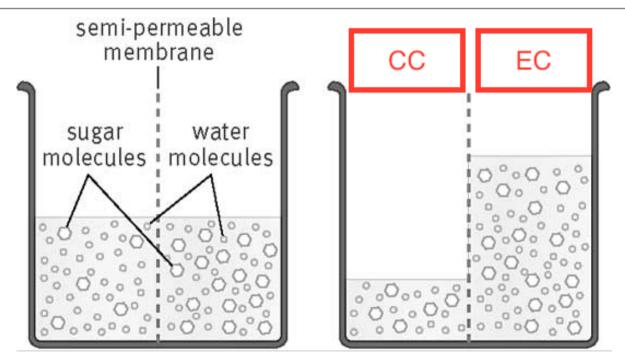
Time for Cooking all Ingredients Together



Pisa - 24/FEB/17

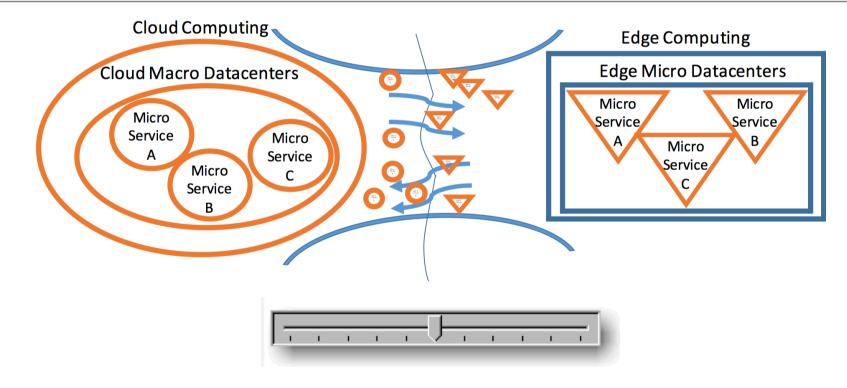


Osmosis Process



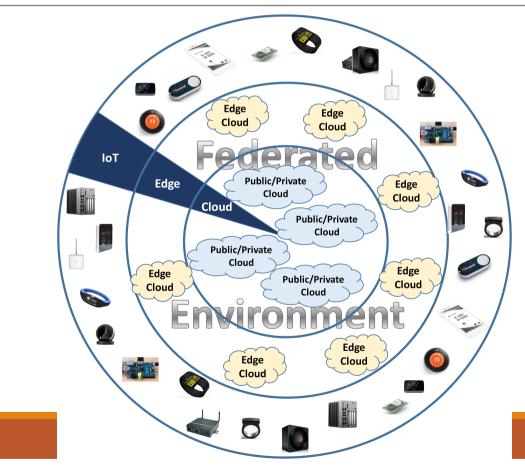


Osmotic Computing Concept

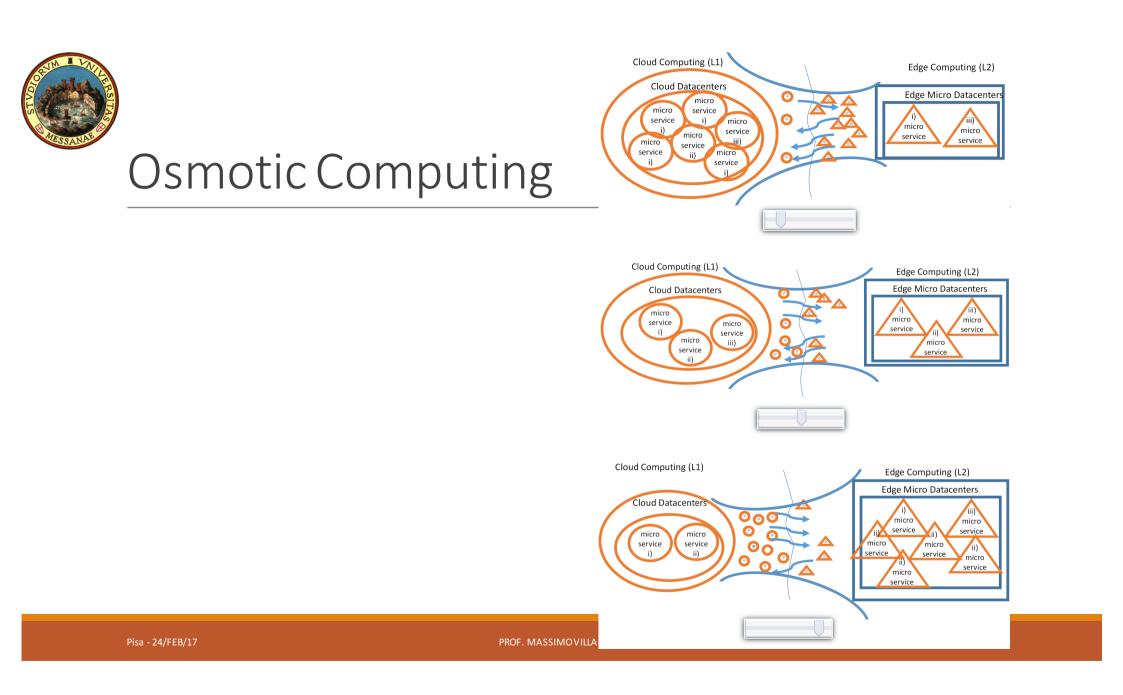




Osmotic Computing Federated Scenario

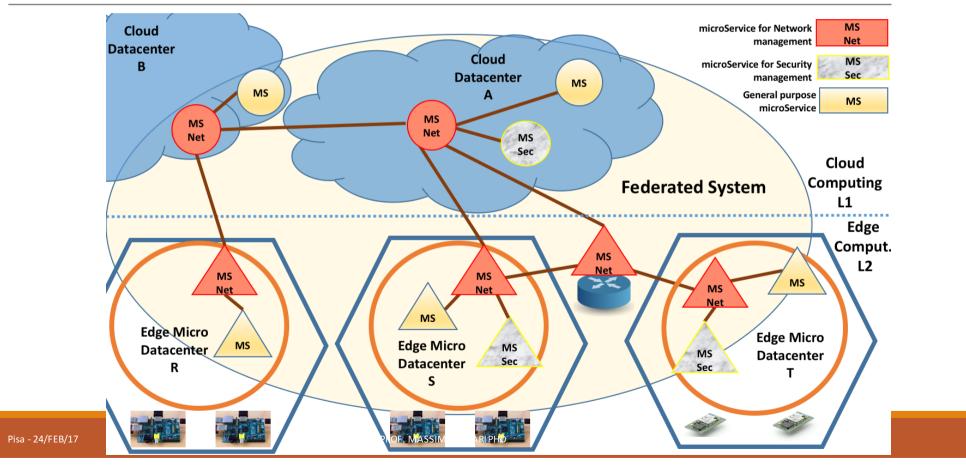


Pisa - 24/FEB/17

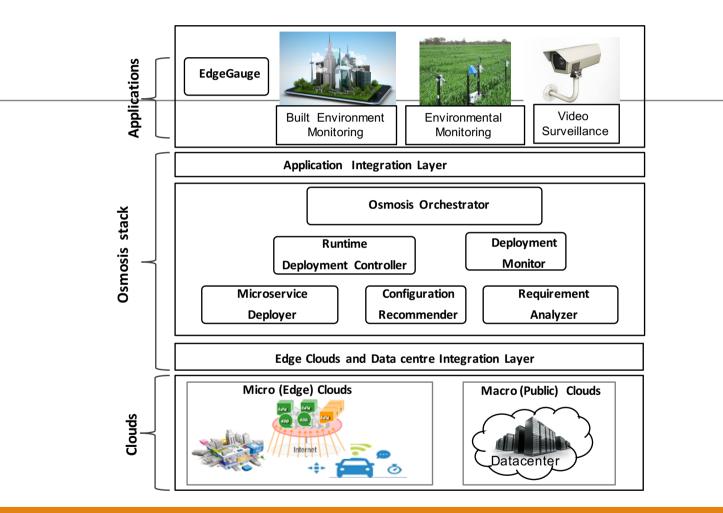




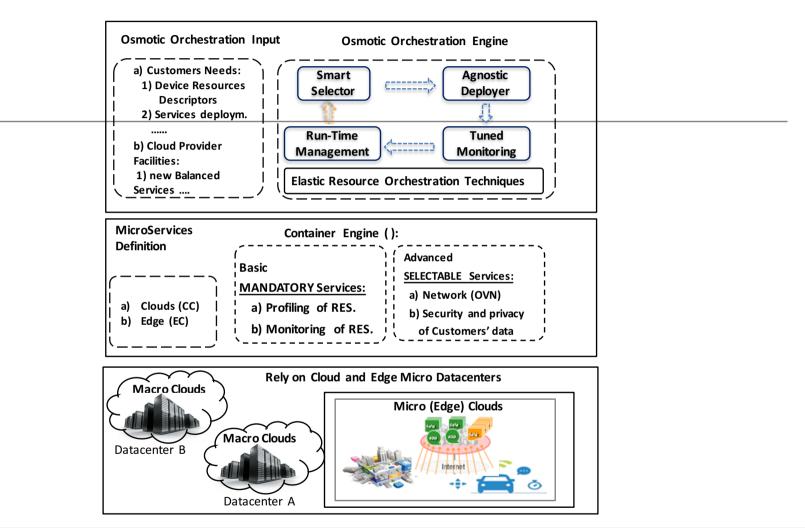
Osmotic Computing













IEEE Cloud Computing > Volume: 3 Issue: 6 (Dec 2016): Osmotic Computing:Massimo Villari ; Maria Fazio ; Schahram Dustdar ; Omer Rana ; Rajiv Ranjan A New Paradigm for Edge/CloudA. Runtime Microservice Deployment

- B. Microservice configuration
- C. Microservice Networking
- D. Microservice Security
- E. Edge Computing
- F. Microservice workload contention and interference evaluation
- G. Monitoring
- H. Microservice orchestration and elasticity control

Pisa - 24/FEB/17



Runtime Microservice Deployment:

An Osmotic Computing framework should provide a microservice Engine, allowing users and developers to deploy containers running microservices on IoT and Edge devices, enabling microservice execution and deployment. The innovation delivered by Osmotic Computing will facilitate the creation of a market of virtual IoT based applications. Software adaptation and versioning mechanisms will allow Edge Cloud providers to deploy microservices consisting of a heterogeneous pool of physical devices. Benefits of Osmotic Computing include deployment of distributed IoT oriented microservices, software consolidation, and service optimization.



Microservice configuration:

In Osmotic Computing developing holistic decision-making frameworks that automate configuration selection across microservices and resources in Cloud and Edge datacenters to meet QoS constraints is necessary. To this end, novel decision-making techniques based on multi-criteria optimization (e.g., Genetic Algorithms) and multi-criteria decision making (e.g., Analytic Network Process) techniques should be investigated.



Microservice Networking:

Osmotic Computing is based on an abstraction of networks that spawn from Cloud to Edge and vice versa for improving the performance of the communication among microservices. The network here represents an enabler that allows us to **dynamically adjust the overall microservices behavior according to user requirements.** The network management advances in Osmotic Computing should include the development of an **interoperability layer** for remote orchestration of heterogeneous Edge devices, for example, exploiting **Software Defined Networking (SDN) and Network Function Virtualization (NFV) capabilities**, accessible through an API.



Edge Computing

The approach suggests the need to combine **"mobile offloading" with "data centre offloading"** i.e., we off-load computation initially carried out within a datacenter to a mobile device. **This "reverse" off-loading** enables computation to be undertaken closer to the phenomenon being measured (overcoming latency and data transfer costs). The Osmotic Computing approach is therefore focused on **understanding the types of microservices which would be more relevant to execute at the Edge**, rather than within a datacenter environment, and vice versa.



Microservice workload contention and interference evaluation

The co-deployed microservices on Cloud or Edge datacenters can lead to **contention problems which will affect QoS**. During deployment of microservices, orchestration techniques must consider which microservices should be combined on a datacenter resource, to minimize resource contention due to workload interference. **Workload resource consumption and QoS are not additive**, so understanding the nature of their composition is critical to deciding which microservices can be deployed together.

Research in Osmotic Computing should be focus on novel **microservice consolidation techniques** that can **dynamically detect and resolve resource contention** via microservice performance characterization, workload prioritization and coordinated deployment.



Monitoring

(i) monitor and instrument data (workload input and QoS metrics, disruptive event) across microservices, Cloud datacenter, intransit network, and Edge datacenter

(ii) detect root causes of QoS violations and failures across the infrastructure based on workload and QoS metrics logs.

Researchers should investigate scalable methods (based on self-balanced trees) to

monitor QoS and security metrics across multiple-levels of Osmotic Computing including microservices, Cloud datacenters and Edge micro-datacenters.



Microservice orchestration and elasticity control

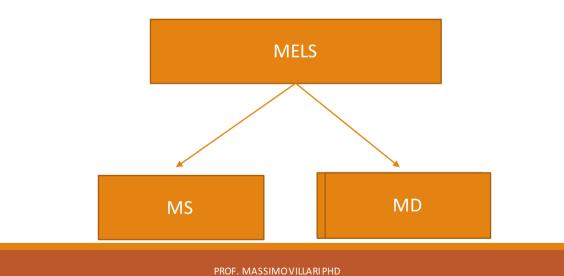
In Osmotic Computing, the traditional **notion of run-time control and reconfiguration** which only considers resources hosted in Cloud datacentes, to resources that are deployed and available at the Edge, should be **extended**. **Machine learning techniques for developing predictive models to forecast workload** input and performance metrics across multiple, colocated microservices on Cloud and Edge datacenter resources should be investigated. Additionally, **intelligent**, **QoS-aware**, **and contention-aware resource orchestration algorithms** should be developed based on the above models, **monitoring systems**, **and configuration selection** techniques.



MicroELementS (MELS) split into two main abstracted components:

- MS(MicroServices) and
- MD(MicroData).

Pisa - 24/FEB/17





It will appear Soon in CN4IoT/Springer - Osmotic Computing:

Basic Elements and Their Details

Main root of the hierarchy is represented from MicroELementS, whereas underneath there are MSs and MDs.

The leaf of hierarchy is represented by:

- MUS
- MOS
- MUD
- MOD

U for User and O for Operational.



It will appear Soon in CN4IoT/Springer - Osmotic Computing: Basic Elements and Their Details

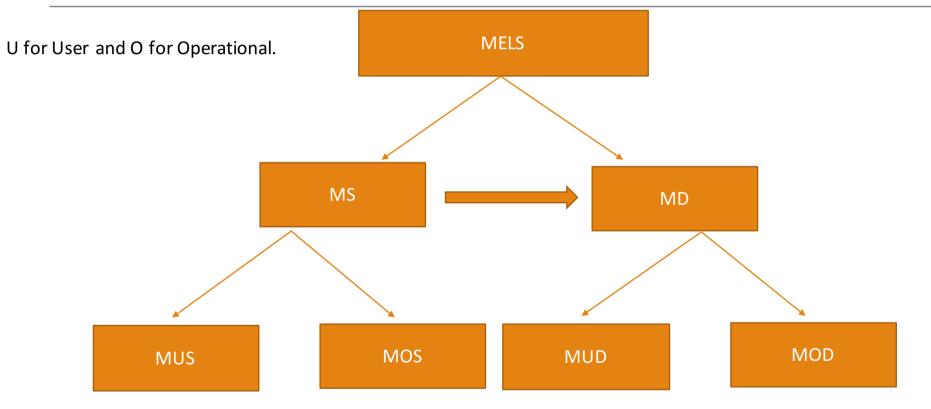
Main root of the hierarchy is represented from MicroELementS, whereas underneath there are MSs and MDs.

The leaf of hierarchy is represented by:

- MicroOperationalService (like an Operating System) and
- MicroUserService(like a user application on OS).
- MicroOperationalData(-> MS configuration; <- MS monitoring) and
- MicroUserData(-> User Data; <- User Data from IoT).

MD and MS are mobile, can be portable and cross-platform.







Name space for MELS

MS:nameUrl:tag_UUID.

- MUS:nameUrl:tag_UUID
- MOS:nameUrl:tag_UUID

MD:nameUrl:tag_UUID

- MOD:nameUrl:tag_UUID
- MOD:nameUrl:tag_UUID



MS:nameUrl:tag_UUID.MD:nameUrl:tag_UUID

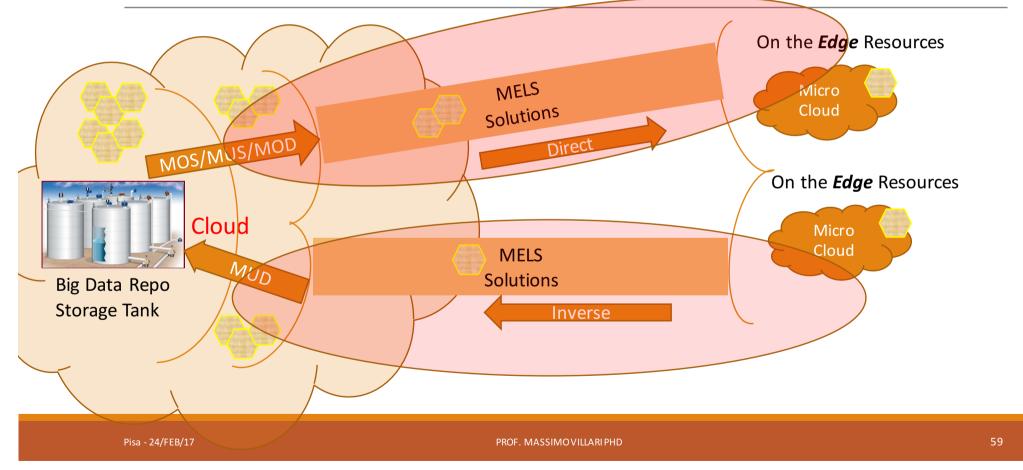


Osmosis equivalences:

- Solutions,
- Membranes (semipermeable and impermeable),
- Storage Tanks,
- Chambers, and
- Tubes



Osmosis: Solutions, Membranes (semipermeable and impermeable), Storage Tanks, Chambers, and Tubes





OC: Software Defined Membrane (permeability)

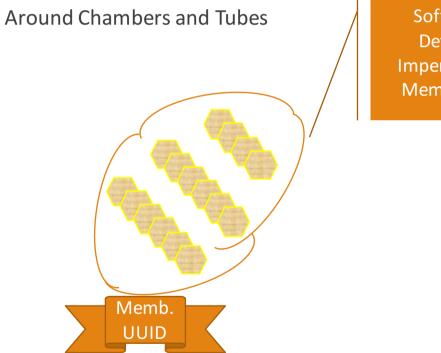
- Membranes As A Filter
- Membranes assembled from MS
- Membranes Assimilated at Gateways and/or Proxys
- Able to be permeable or impermeable

Software Defined Membranes



OC: Software Defined Pressure in Chambers and Tubes

Security Domain



Software Defined Impermeable Membranes

Pisa - 24/FEB/17



Big Security Issues in the EDGE $\ensuremath{\mathfrak{S}}$

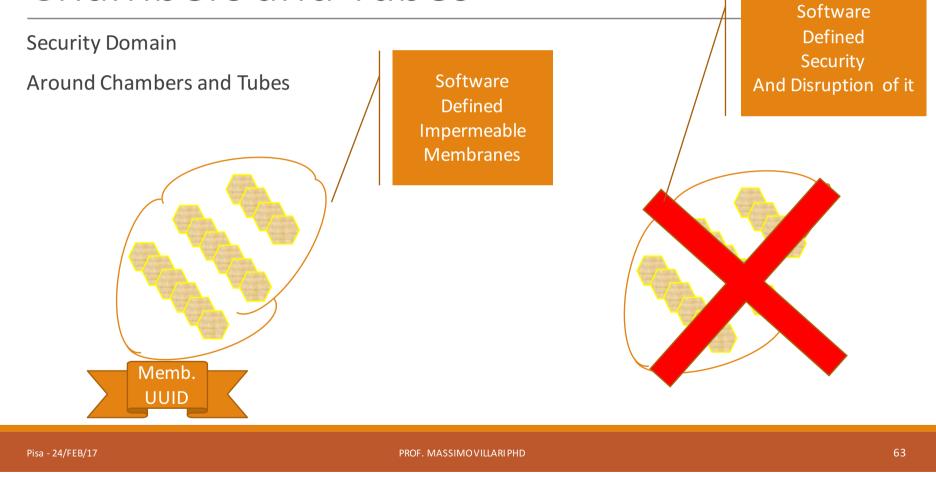
In November 2013, the owner of a smart TV made by LG Electronics discovered that <u>the</u> <u>device was collecting information</u> about his viewing habits, even when the "collection of watching info" feature was turned off. Worse, the TV also sent back to LG's servers the names of files stored on external media devices and even network shares.

2 more wireless baby monitors hacked: Hackers remotely spied on babies and parents

Two more wireless baby monitors were hacked. One family heard voices as the camera followed them about the room; the second mom was freaked out and scared as a hacker remotely controlled the camera to follow her movements.

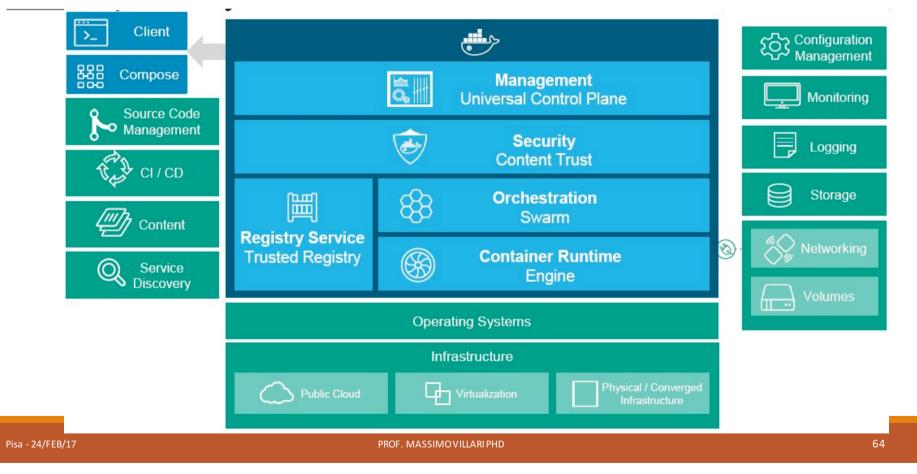


OC: Software Defined Pressure in Chambers and Tubes



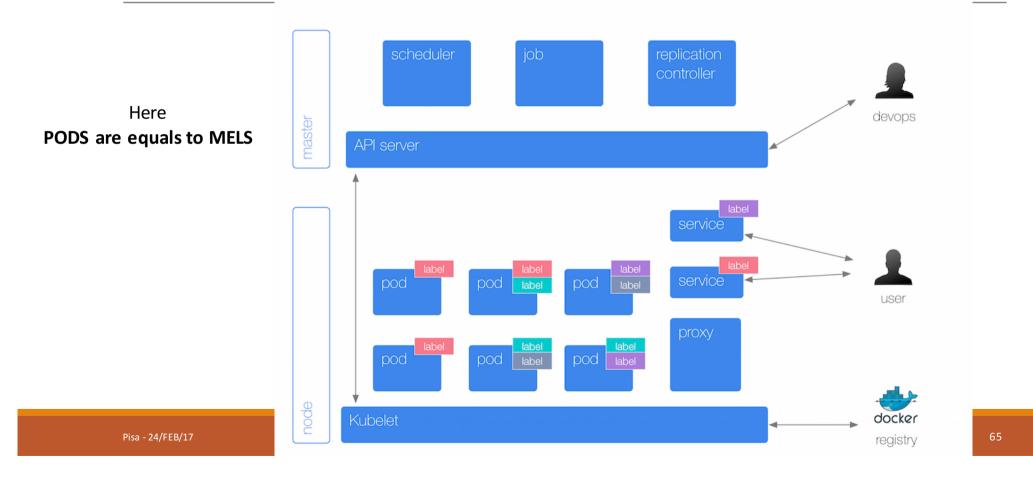


Osmotic NOT from Scratch: eg. containerization



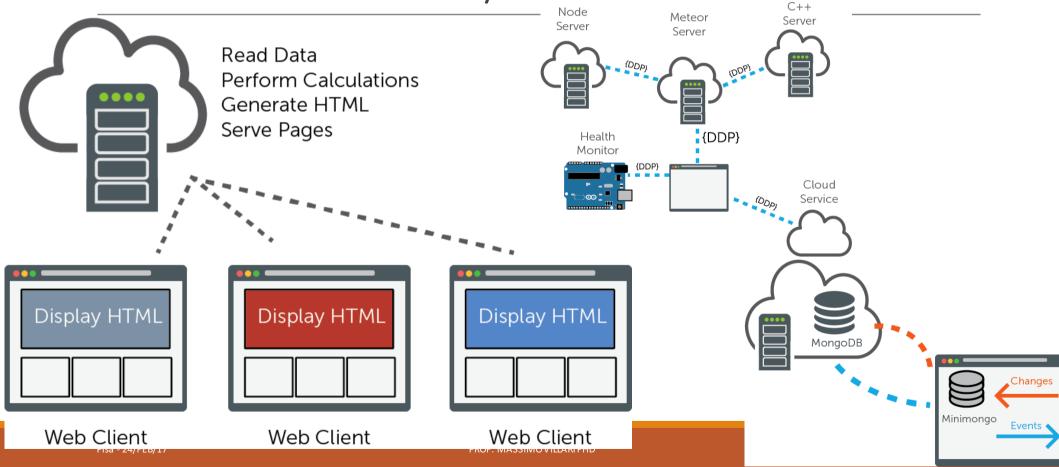


Osmotic NOT from Scratch: eg. containerization





Osmotic NOT from Scratch: eg. Javascript MeteorJS on MPU/MCU





Osmotic Computing: in Docker and JS Meteor or..

In 1- MSs (MOS and MUS) are equals to Container

 $\ln 1 - MD \operatorname{can} \operatorname{be:}$

- MOD in YAML for deploying Services
- MUD in Json for charactering Filesysyem layers in AUFS and LayaoutFS

i.e., Kubernates leverages the approach of more MOS for deploying the Agents

In 2- MS (MOS and MUS) are equals to Javascript code:

 $\ln 2 - MD \operatorname{can} be:$

• Json

* The TOSCA Simple Profile in YAML V1.1 is now out for a public review.

For details and links to the specification, see the announcement at <u>https://www.oasis-open.org/news/announcements/15-day-public-review-for-tosca-simple-profile-in-yaml-version-1-1-ends-march-2nd</u>

Pisa - 24/FEB/17



Osmotic Computing *which devices: mpu/mcu*





With Raspberry



A new Prototype with:

- Archlinux Read Only Mode
- Docker
- Kubernates and Hypecube
- Fl..

For the Future (looking at life cycles of Microservices) :

- PXE in Raspberry
- Filesystem and with Snapshot, TAGs and Versioning

0



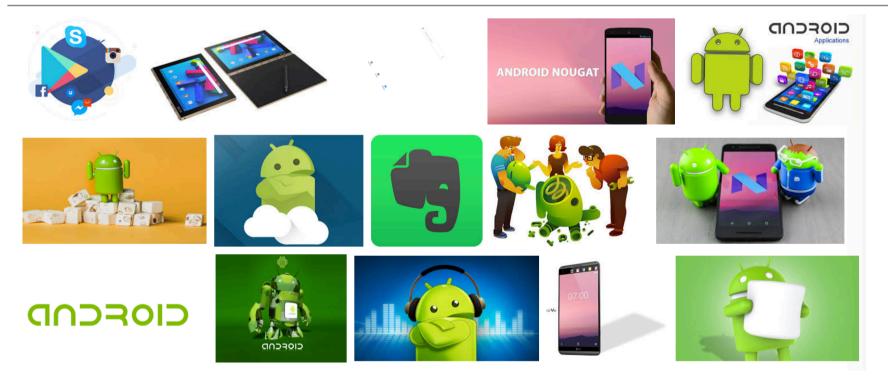
With ESP 8266: looking at platforms and protocols

A new Prototype with:

- LUA and Python Architectures
- Simple Code Injection
- CoAP Client and Server in Python



Osmotic Computing *which oth. devices??*





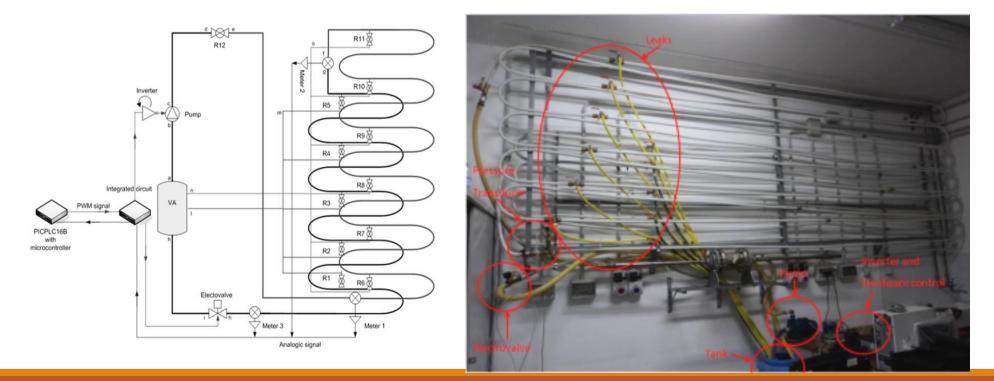
Osmotic Computing On the Cloud

Reusing the existing cloud infrastructure:

- Cloud Storage
- Cloud Processing
- NFV-SFC
- IAM
- NoSQL DB: eg., MongoDB
- APIs RESTFul
- ...
- VMs and Containers



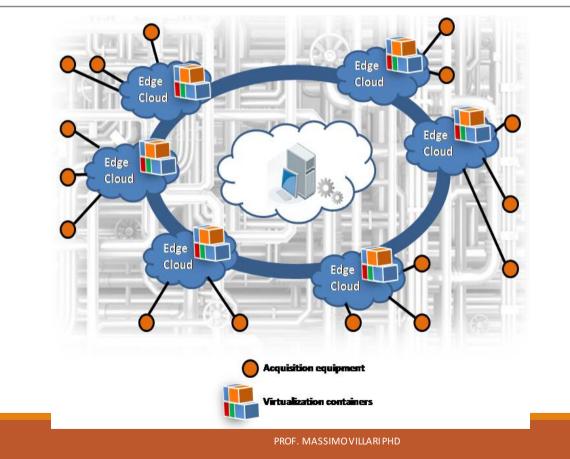
Osmotic Computing Use-Cases: Acquedotto





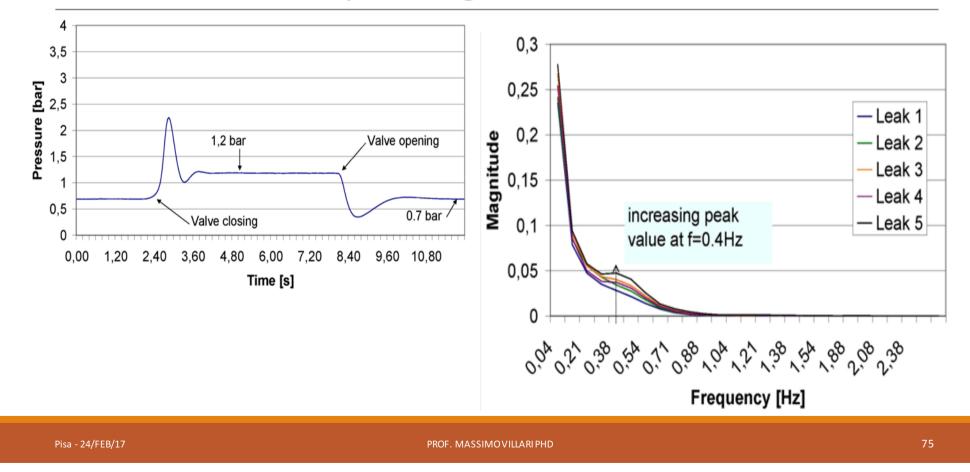
Pisa - 24/FEB/17

Osmotic Computing Use-Cases





Osmotic Computing Use-Cases: FFT





Hazard the applicability

User Interfaces (UI):

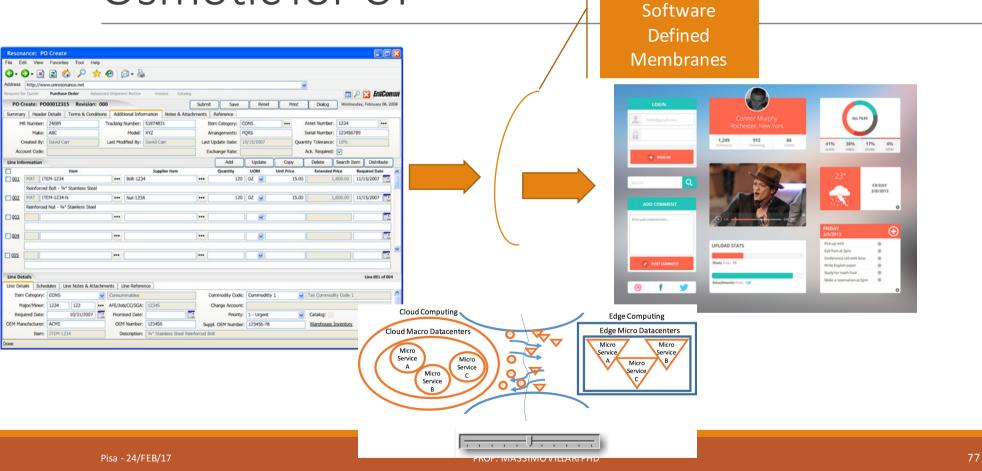
- Social ??
- Serverless ??

Software Defined Osmosis for People

From Smart to Osmotic Cities

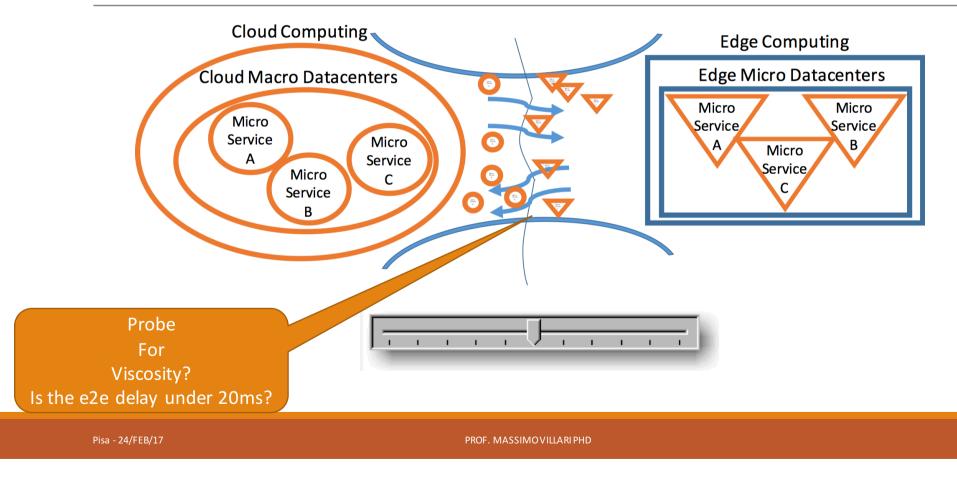


Osmotic for UI



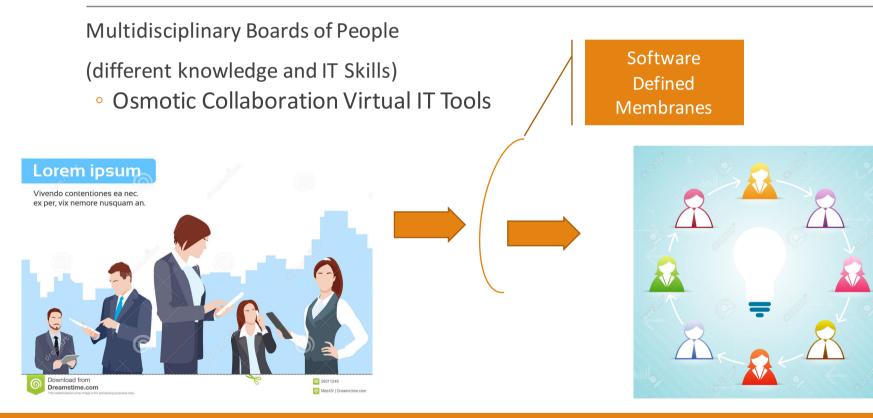


Osmotic Computing Concept





Software Defined Osmosis for People



Pisa - 24/FEB/17

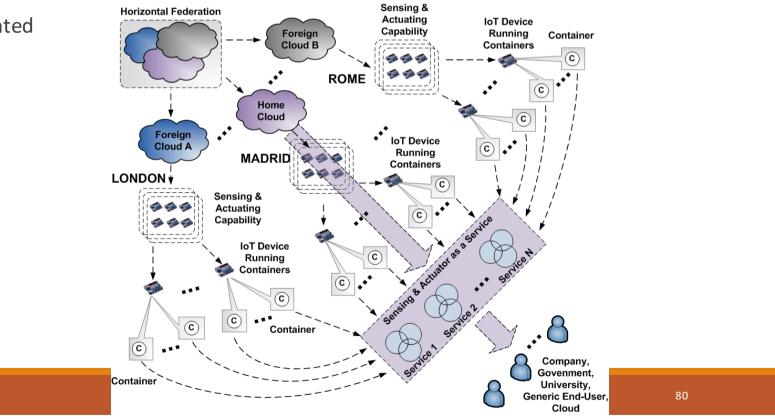


From Smart to Osmotic Cities

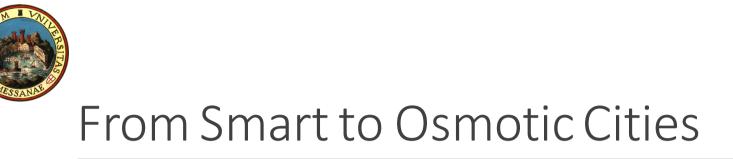
Natural Extension of Federation among Users and Service - Utility Providers

MicroElements Oriented

Osmotic Based



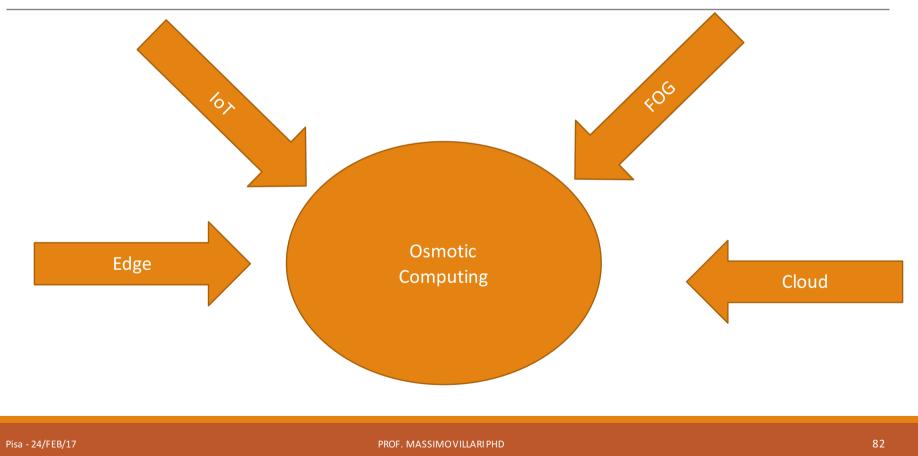
Pisa - 24/FEB/17



Where People and Devices Osmotically Behave

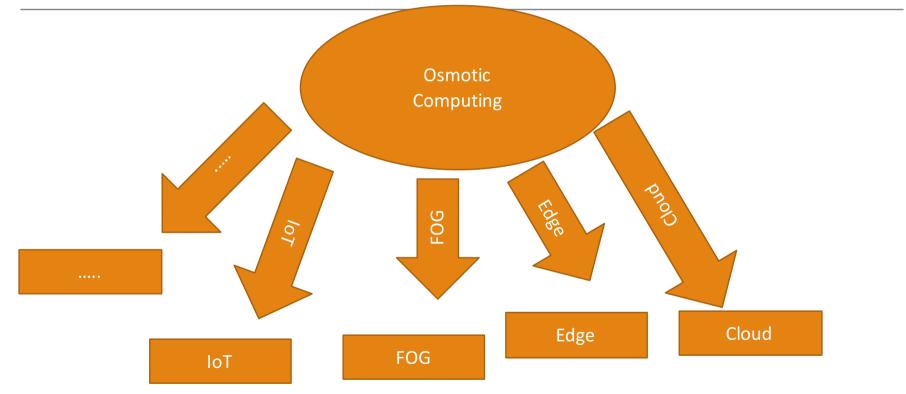


In Conclusion we can...

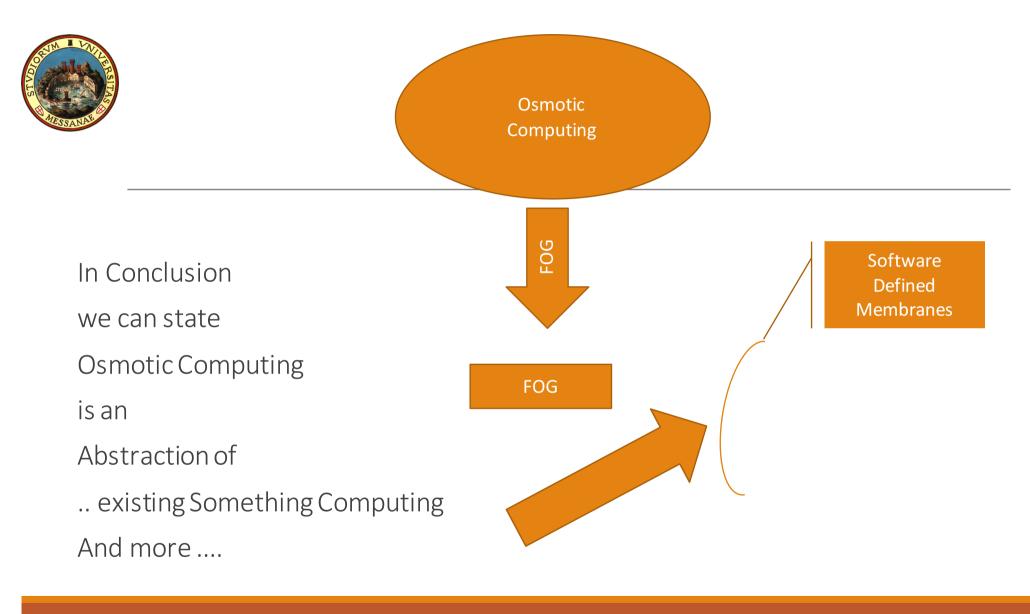




But we also can ...



Pisa - 24/FEB/17



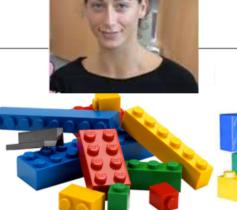


PROF. MASSIMO VILLARI

FCRLAB - UNIVERSITÀ DI MESSINA, ITALIA

MVILLARI@UNIME.IT









the sector



Prof Massimo Villari - Bratislava Oct 15th