Distributed Systems

Cloud & the Internet of Things

Björn Franke
University of Edinburgh, 2016

OVERVIEW

- Cloud Computing vs. Distributed Computing
  - Examples
  - SAAS, PAAS, IAAS
  - Goals, Types, Characteristics
- Internet of Things
  - Applications
  - Architecture, Characteristics
  - Scalability, Security, Privacy

WHAT IS DISTRIBUTED COMPUTING?

DEFINITIONS

- Appear as a single system (location transparency)
- Internally connected to several nodes which perform the designated computing tasks
- "software system in which components located on networked computers communicate and coordinate their actions by passing messages."

WHAT IS DISTRIBUTED COMPUTING?

DISTRIBUTED COMPUTING SYSTEM EXAMPLES

- World Wide Web
- Social Media Giant Facebook
- Hadoop's Distributed File System (HDFS)
- ATM
- Cloud Network Systems (Specialised form of Distributed Computing Systems)
- Google Bots, Google Web Server, Indexing Server

WHAT IS DISTRIBUTED COMPUTING?

MASTER/SLAVE ARCHITECTURE

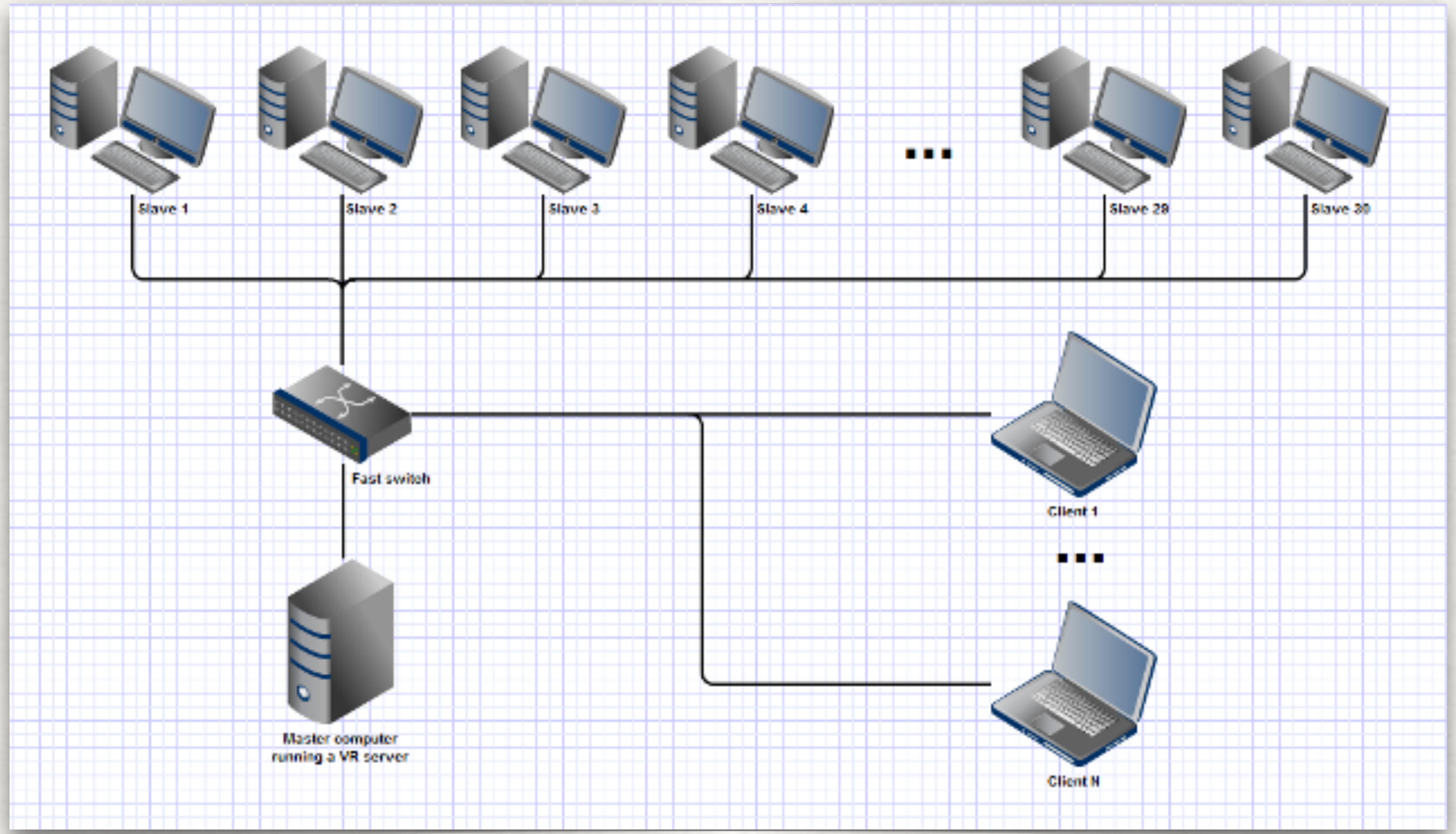


Image Credit :researchgate.net

WHAT IS DISTRIBUTED COMPUTING?

BENEFITS OF DISTRIBUTED COMPUTING

- Provide a better price/performance ratio when compared to a centralised computer
- More computational power than centralised computing systems
- Provide incremental growth so that organisations can add software and computation power in increments as and when business needs

WHAT IS THE CLOUD?

CLOUD COMPUTING EXAMPLES

- **YouTube** is the best example of cloud storage which hosts millions of user uploaded video files.

- **Picasa** and **Flickr** host millions of digital photographs allowing their users to create photo albums online by uploading pictures to their service's servers.

- **Google Docs** is another best example of cloud computing that allows users to upload presentations, word documents and spreadsheets to their data servers. Google Docs allows users edit files and publish their documents for other users to read or make edit

WHAT IS THE CLOUD?

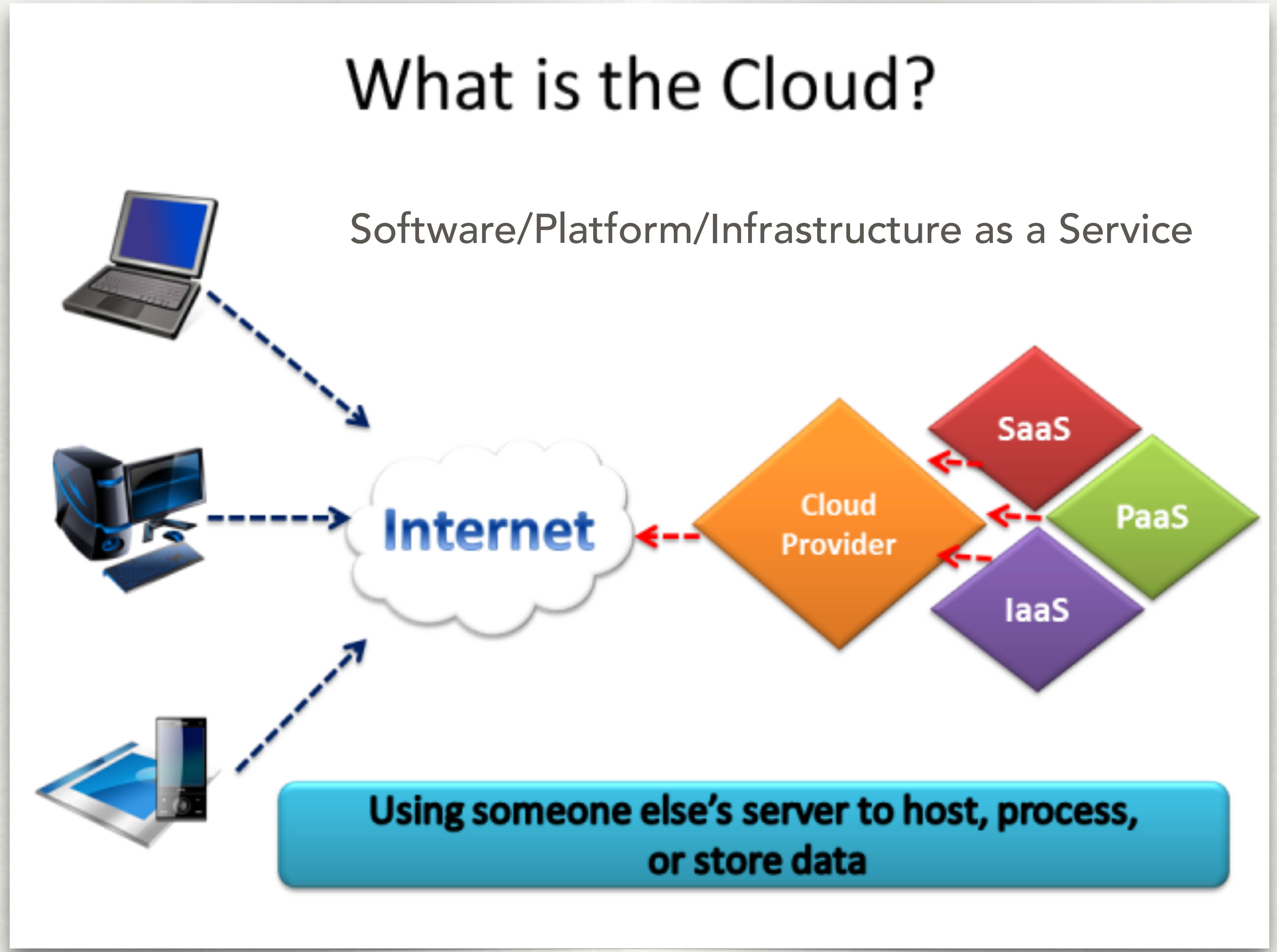


Image Credit: imscindiana.com

WHAT IS THE CLOUD?

SAAS/PAAS/IAAS

- **Applications** are hosted by a vendor or service provider and made available to customers over a network, typically the Internet, e.g. Google Docs
- **Delivery models** for SaaS
  - **Hosted application management (hosted AM):** provider hosts commercially available software for customers and delivers it over the Web
  - **Software on demand:** provider gives customers network-based access to a single copy of an application created specifically for SaaS distribution
- **Benefits:** easier administration, automatic updates and patch management, compatibility: All users will have the same version of software, easier collaboration, for the same reason, global accessibility.

WHAT IS THE CLOUD?

SAAS/PAAS/IAAS

- Cloud provider delivers **hardware and software tools** - usually those needed for **application development** - to its users as a service, e.g. Microsoft Azure

- Provider hosts the hardware and software on its own infrastructure.

- Frees users from having to install in-house hardware and software to develop or run a new application.

- Business relies on PaaS providers for key services, such as Java development or application hosting.

- PaaS provider supports all the underlying computing and software; users only need to log in and start using the platform – usually through a Web browser interface.

WHAT IS THE CLOUD?

SAAS/PAAS/IAAS

- Provides **virtualised computing resources** over the Internet, e.g. Amazon EC2 but also Microsoft Azure
- Third-party provider hosts hardware, software, servers, storage and other infrastructure components on behalf of its users
- Hosts users' applications and handle tasks including system maintenance, backup and resiliency planning
- Automation of administrative tasks, dynamic scaling, desktop virtualisation and policy-based services

WHAT IS THE CLOUD?

BENEFITS OF CLOUD COMPUTING

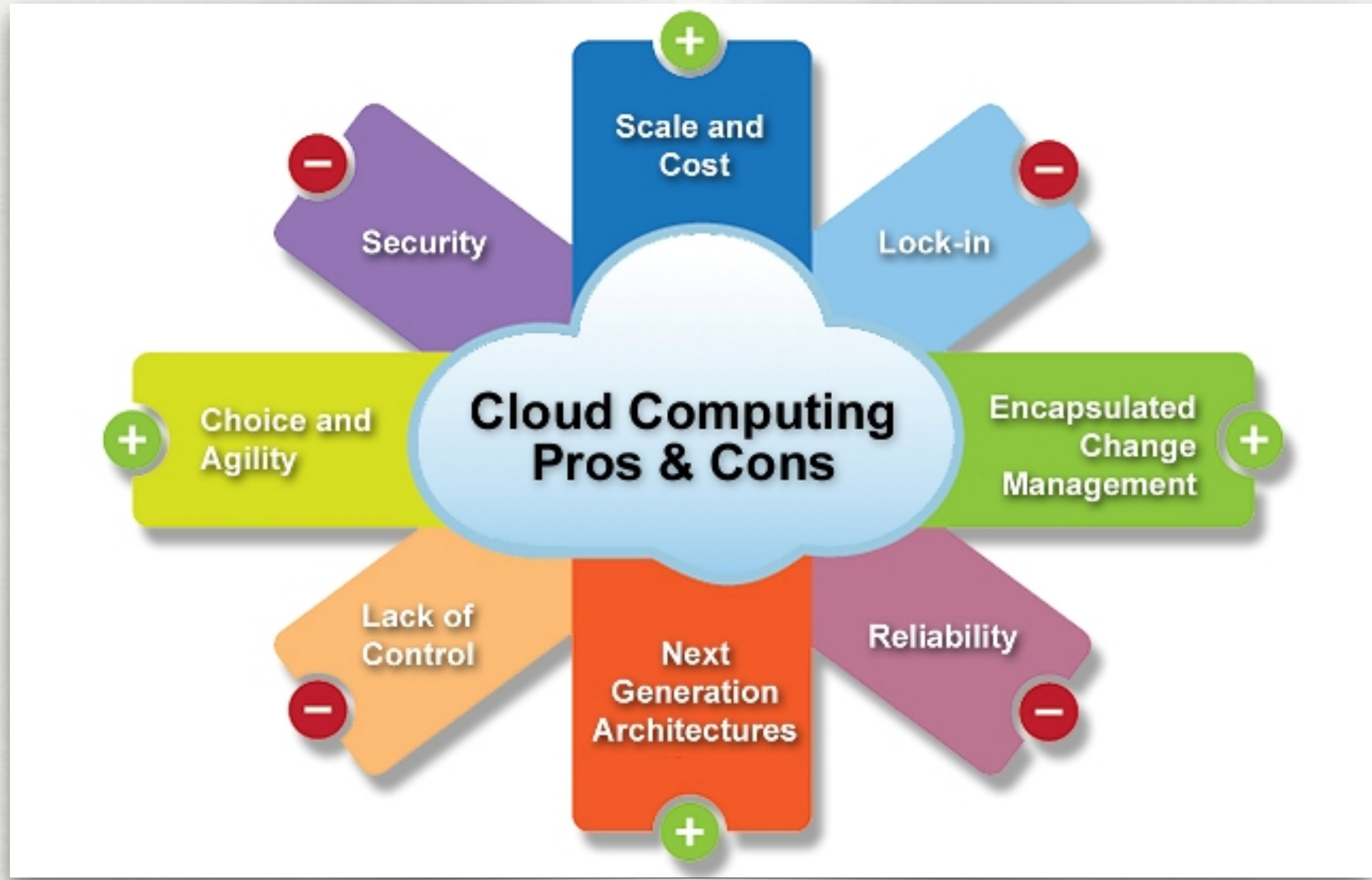


Image Credit :belden.com

CLOUD COMPUTING VS. DISTRIBUTED COMPUTING

GOALS

- **Distributed Computing**
  - Collaborative resource sharing by connecting users and resources
  - Administrative scalability (number of domains in administration), size scalability (number of processes and users), and geographical scalability (maximum distance between the nodes in the distributed system)
- **Cloud Computing**
  - Delivering services or applications on demand
  - Increased scalability, transparency, security, monitoring and management
  - Services are delivered with transparency not considering the physical implementation within the Cloud

CLOUD COMPUTING VS. DISTRIBUTED COMPUTING

TYPES OF DISTRIBUTED COMPUTING

- **Distributed Information Systems**
  - Distribute information across different servers through various communication models like RMI and RPC.
- **Distributed Pervasive Systems**
  - Consist of embedded computer devices such as portable ECG monitors, wireless cameras, PDA's, sensors and mobile devices. Distributed Pervasive systems are identified by their instability when compared to more "traditional" distributed systems.
- **Distributed Computing Systems**
  - Computers connected within a network communicate through message passing to keep a track of their actions.

CLOUD COMPUTING VS. DISTRIBUTED COMPUTING

TYPES OF CLOUD COMPUTING

- **Private Cloud:** Cloud infrastructure dedicated to a particular IT organisation for it to host applications so that it can have complete control over the data without any fear of security breach.

- **Public Cloud:** Cloud infrastructure hosted by service providers and made available to the public. Customers have no control or visibility about the infrastructure. For example, Google and Microsoft own and operate their own their public cloud infrastructure by providing access to the public through Internet.

- **Community Cloud:** A multi-tenant cloud infrastructure where the cloud is shared by several IT organisations.

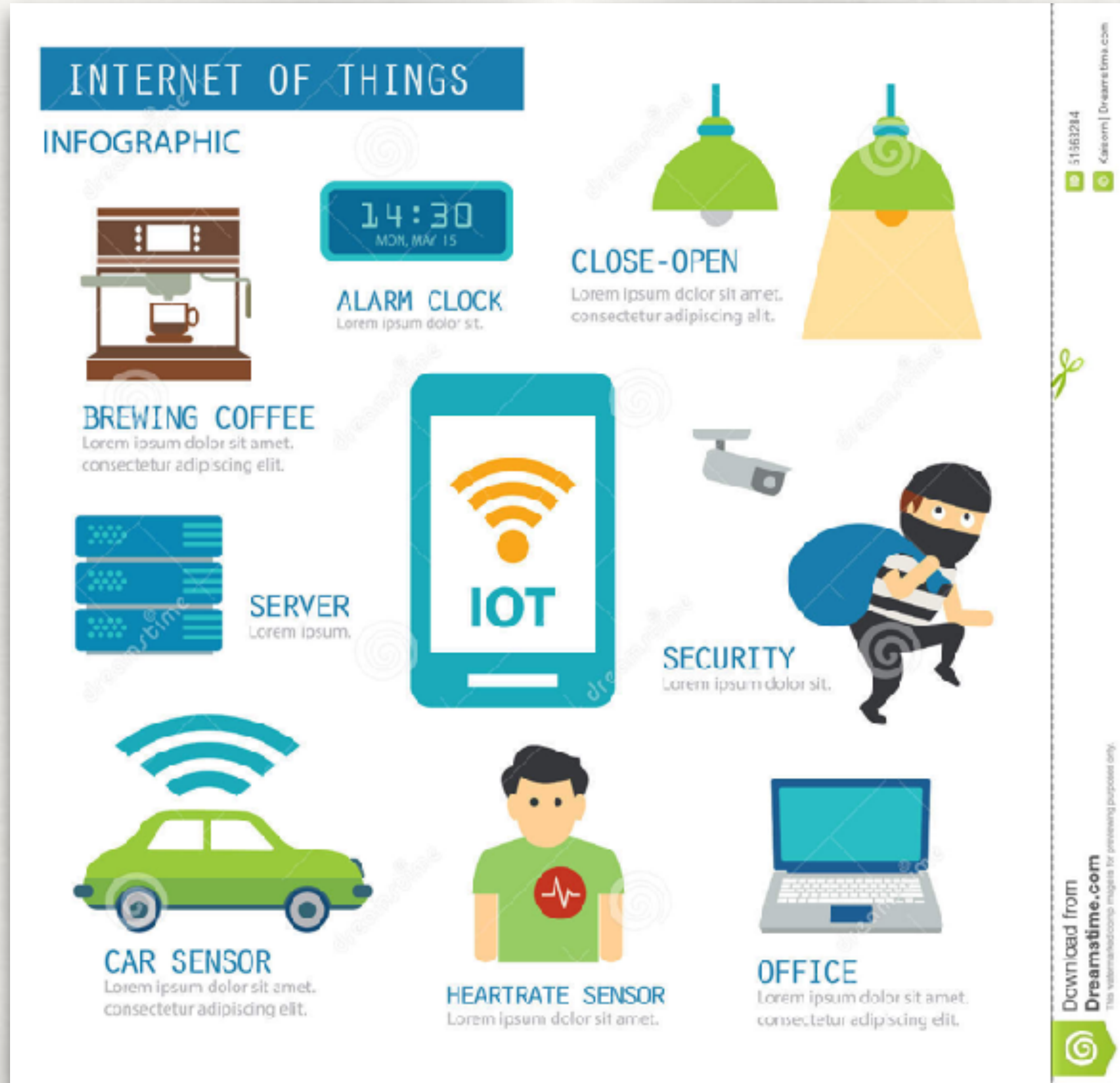
- **Hybrid Cloud:** A combination or 2 or more different types of the above mentioned clouds (Private, Public and Community) forms the Hybrid cloud infrastructure where each cloud remains as a single entity but all the clouds are combined to provide the advantage of multiple deployment models.

CLOUD COMPUTING VS. DISTRIBUTED COMPUTING

CHARACTERISTICS

- **Distributed Computing**
  - Tasks distributed amongst different computers for computational functions to be performed at the same time using RPC
- **Cloud Computing**
  - On-demand network model is used to provide access to shared pool of configurable computing resources

INTERNET OF THINGS



- Network of physical objects or "things" embedded with electronics, software, sensors, and network connectivity
- Enables these objects to collect and exchange data
- Allows objects to be sensed and controlled remotely across existing network infrastructure
- Each thing uniquely identifiable and able to interoperate within the existing Internet infrastructure

INTERNET OF THINGS

APPLICATIONS

- **Media:** Measure, collect and analyse a variety of behavioural statistics (combination of analytics for conversion tracking with behavioural targeting).
- **Environmental monitoring:** air/water quality, atmospheric/soil conditions, movements of wildlife and their habitats, earthquake/tsunami early-warning systems
- **Infrastructure management:** Monitoring and controlling operations of bridges, railway tracks, on- and offshore- wind-farms
- **Manufacturing:** Control and management of manufacturing equipment, asset and situation management, or manufacturing process control
- **Energy management:** Integrated into all forms of energy consuming devices (switches, power outlets, bulbs, televisions, etc.). Communicate with the utility supply company in order to effectively balance power generation and energy usage.
- Medical and healthcare systems, Building and home automation, Transportation…

INTERNET OF THINGS

UNIQUE ADDRESSABILITY OF THINGS

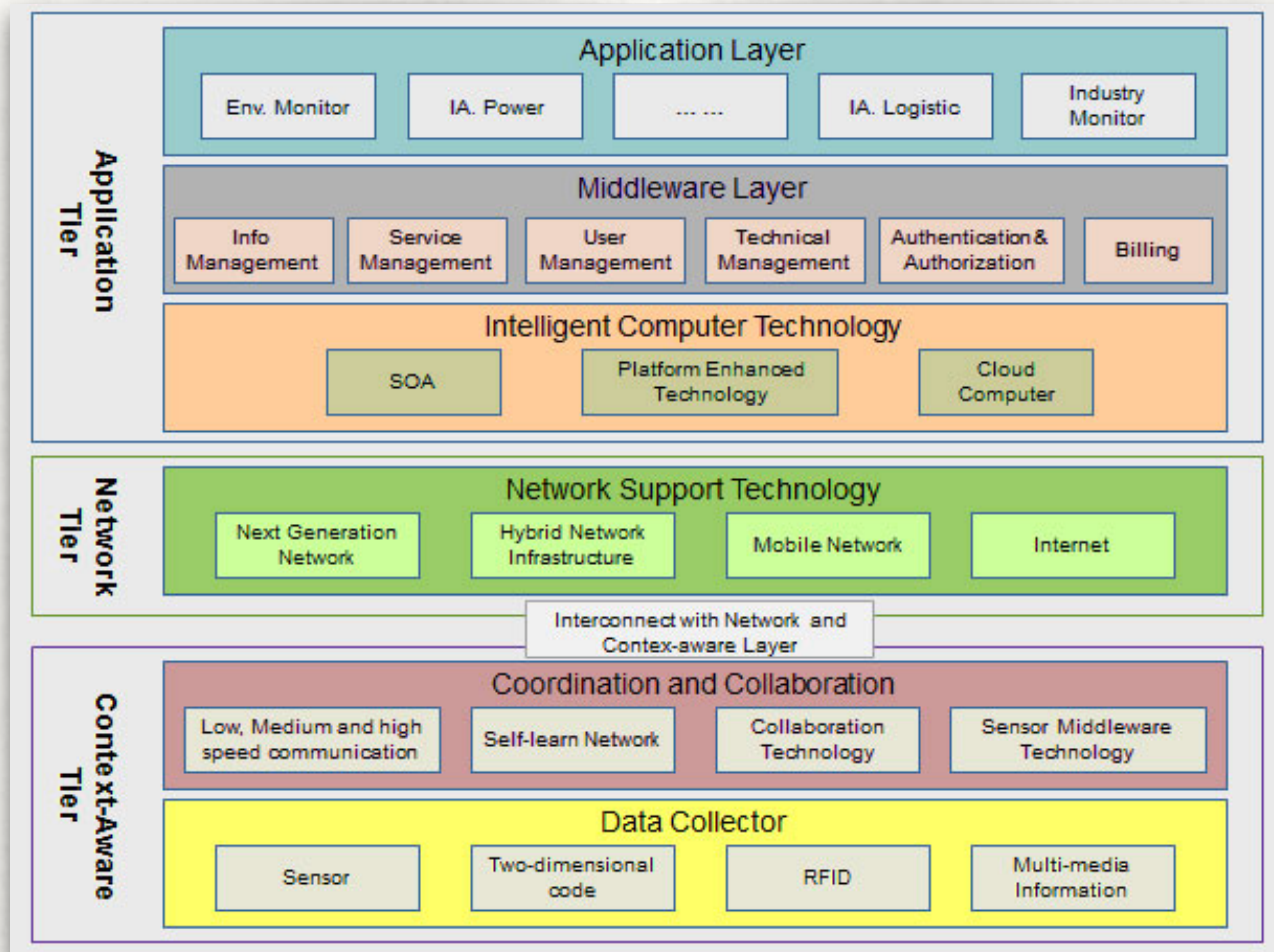
- Original idea: **Auto-ID Center** based on RFID-tags and unique identification through the Electronic Product Code

- Revised idea: Objects have an **IP address** or **URI**

- **IPv6**: Able to address and communicate with devices attached to virtually all human-made objects because of the extremely large address space of the IPv6 protocol

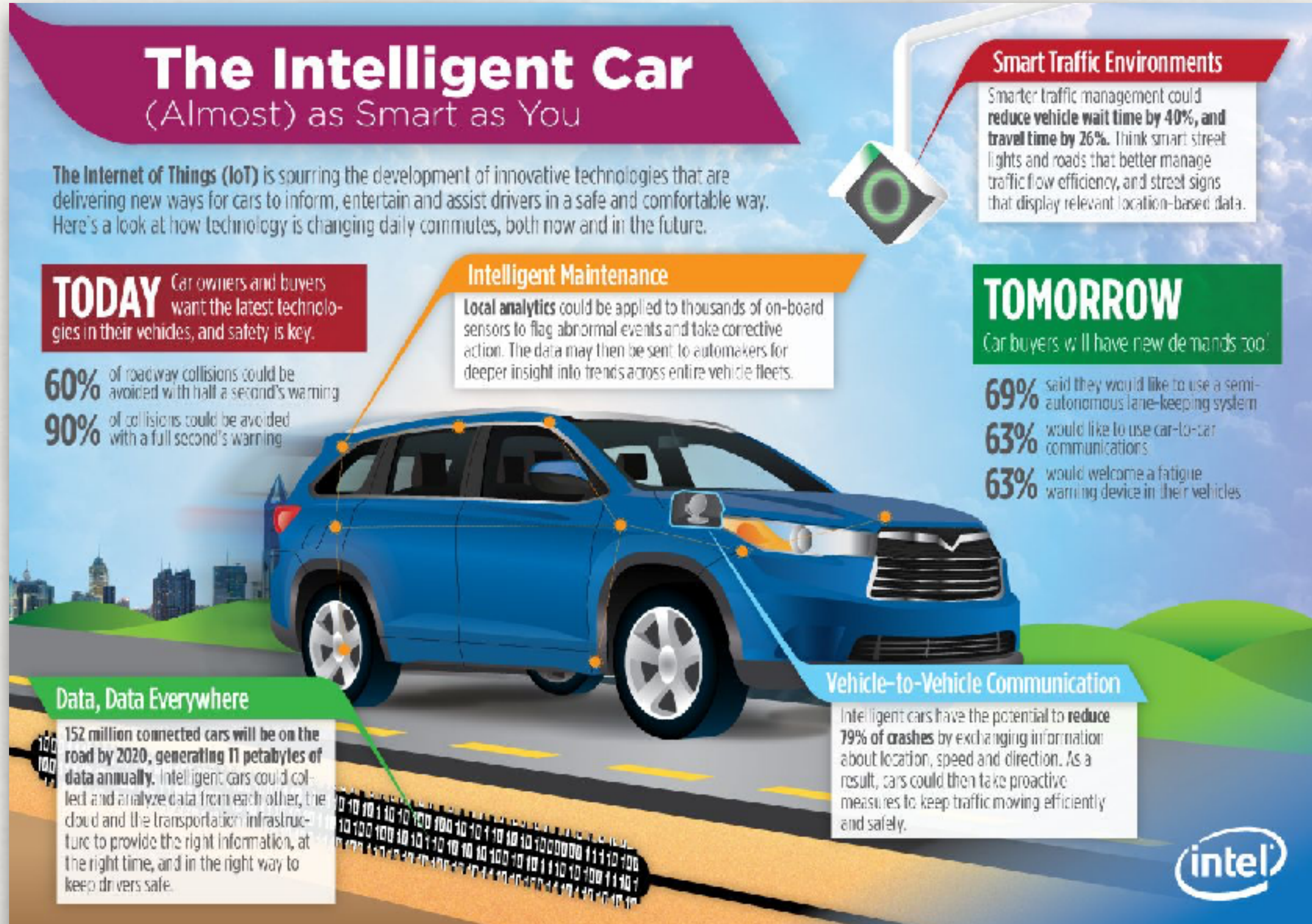
INTERNET OF THINGS

ARCHITECTURE AND LAYERS



INTERNET OF THINGS

EXAMPLE



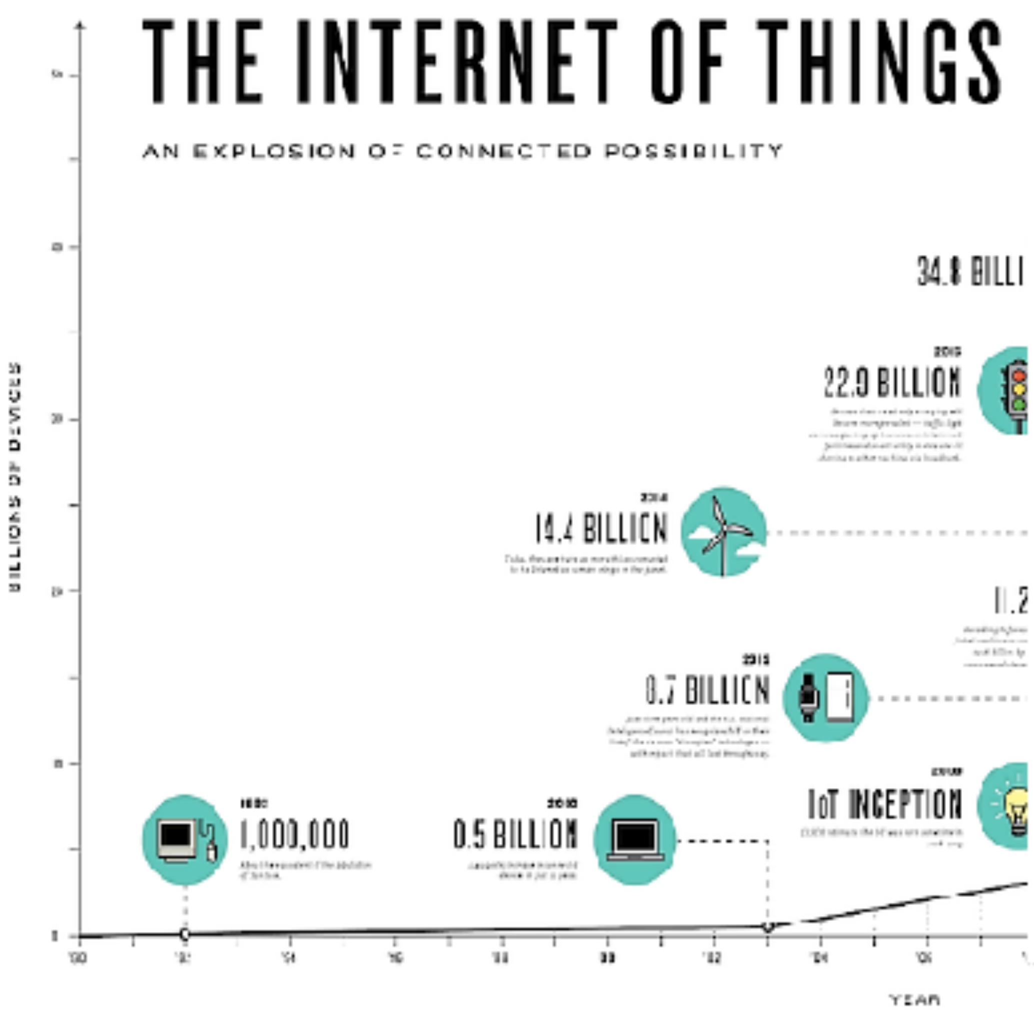
INTERNET OF THINGS

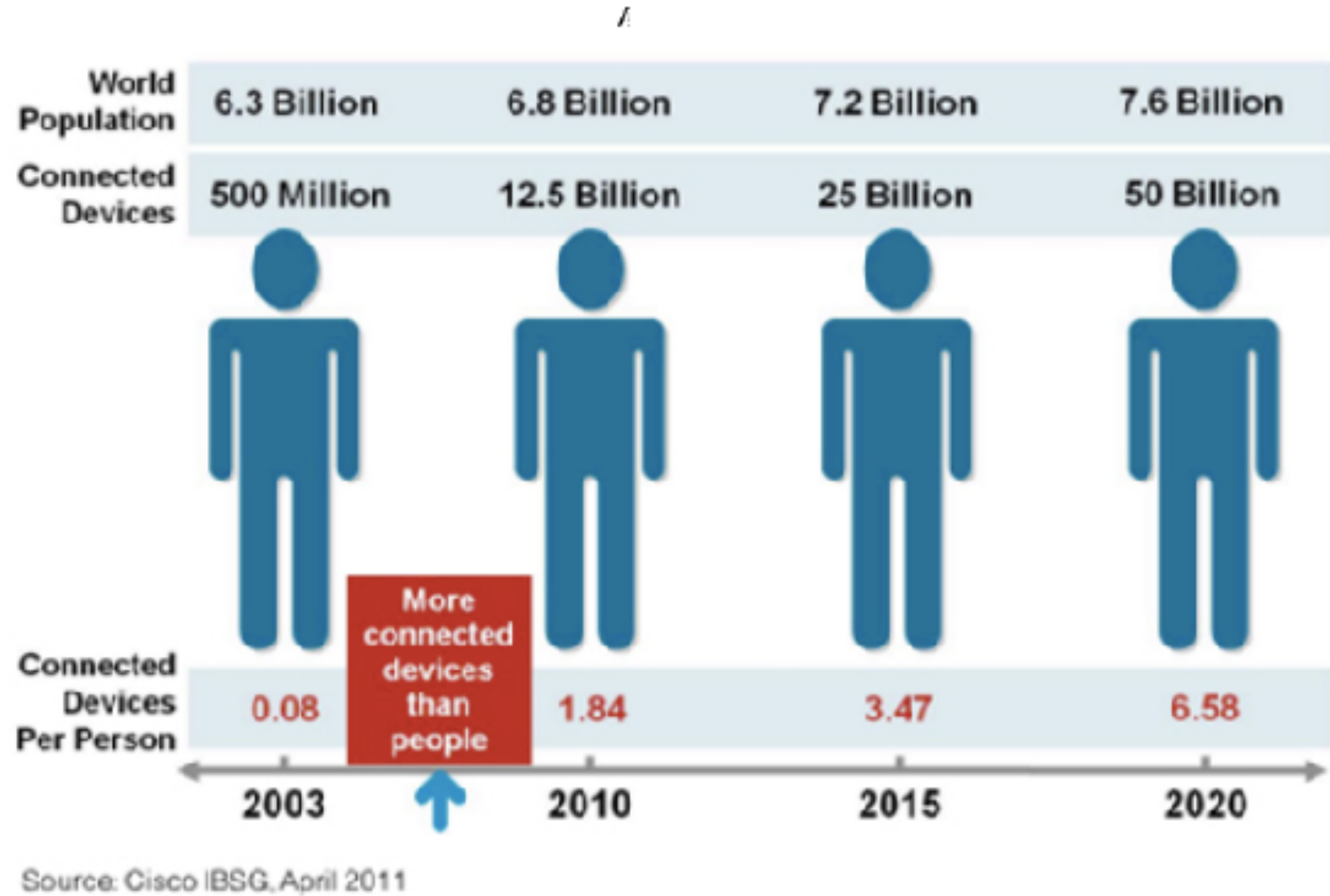
CHARACTERISTICS

- 1. Intelligence – Knowledge extraction from the generated data
- 2. Architecture – A hybrid architecture supporting many others
- 3. Complex system – A diverse set of dynamically changing objects
- 4. Size considerations – Scalability
- 5. Time considerations – Billions of parallel and simultaneous events
- 6. Space considerations – Localisation
- 7. Everything as-a-service – Consuming resources as a service

INTERNET OF THINGS

SCALABILITY



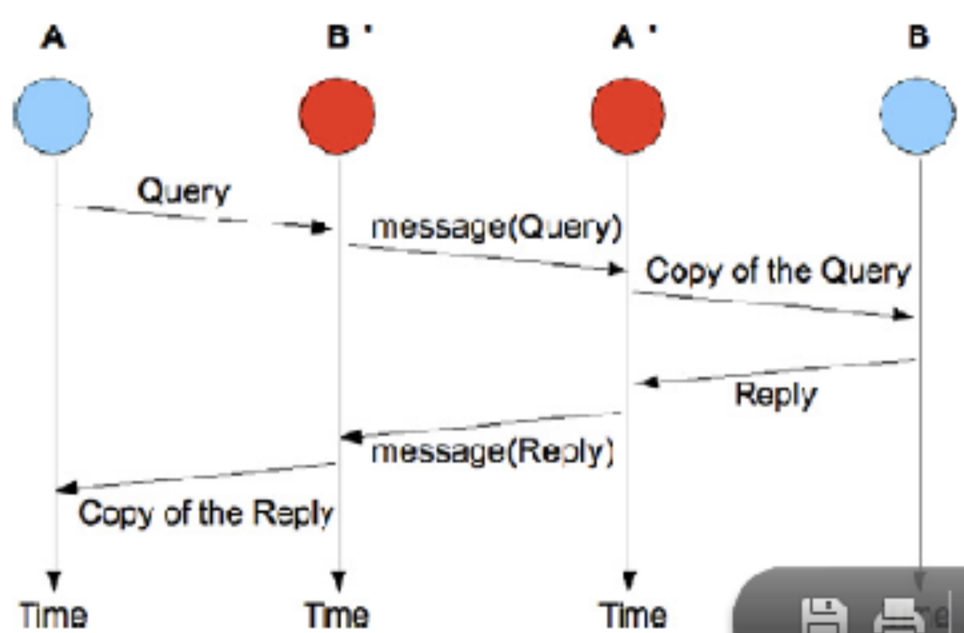


- Number of devices increasing exponentially
  - How can they uniquely be tagged/named?
  - How can the data generated by these devices be managed?

INTERNET OF THINGS

SECURITY

- The components spend most of the time unattended
  - It is easy to physically attack them
- IoT components are characterized by low capabilities in terms of both energy and computing resources
  - They can't implement complex schemes supporting security
- Authentication problem
  - Proxy attack, a.k.a. man in the middle attack problem



- Data integrity
  - Data should not be modified without the system detecting it
  - Attacks on the node
    - Memory protection
  - Attacks over the network
    - Keyed-Hash Message Auth. Code

INTERNET OF THINGS

PRIVACY

- How is it different than traditional privacy?
  - Legislative issues
  - Ethics issues
- Easy for a person to get involved in IoT even if he/she does not know
- Data can be stored indefinitely
- Current solutions are not enough
  - Encryption, pseudo-noise signal, privacy broker