Introduction to Cloud Computing

Antonio Brogi

Department of Computer Science
University of Pisa

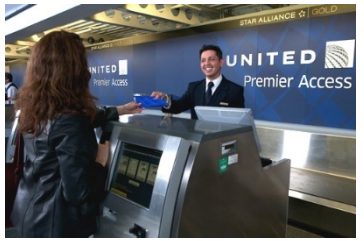

Service-based economy

Everything as a service

Economy trend: from “*goods*” to “*services*”

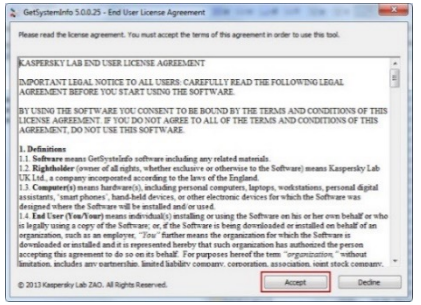
Service contracts

Customers do not (want to) know how the service they use is actually implemented



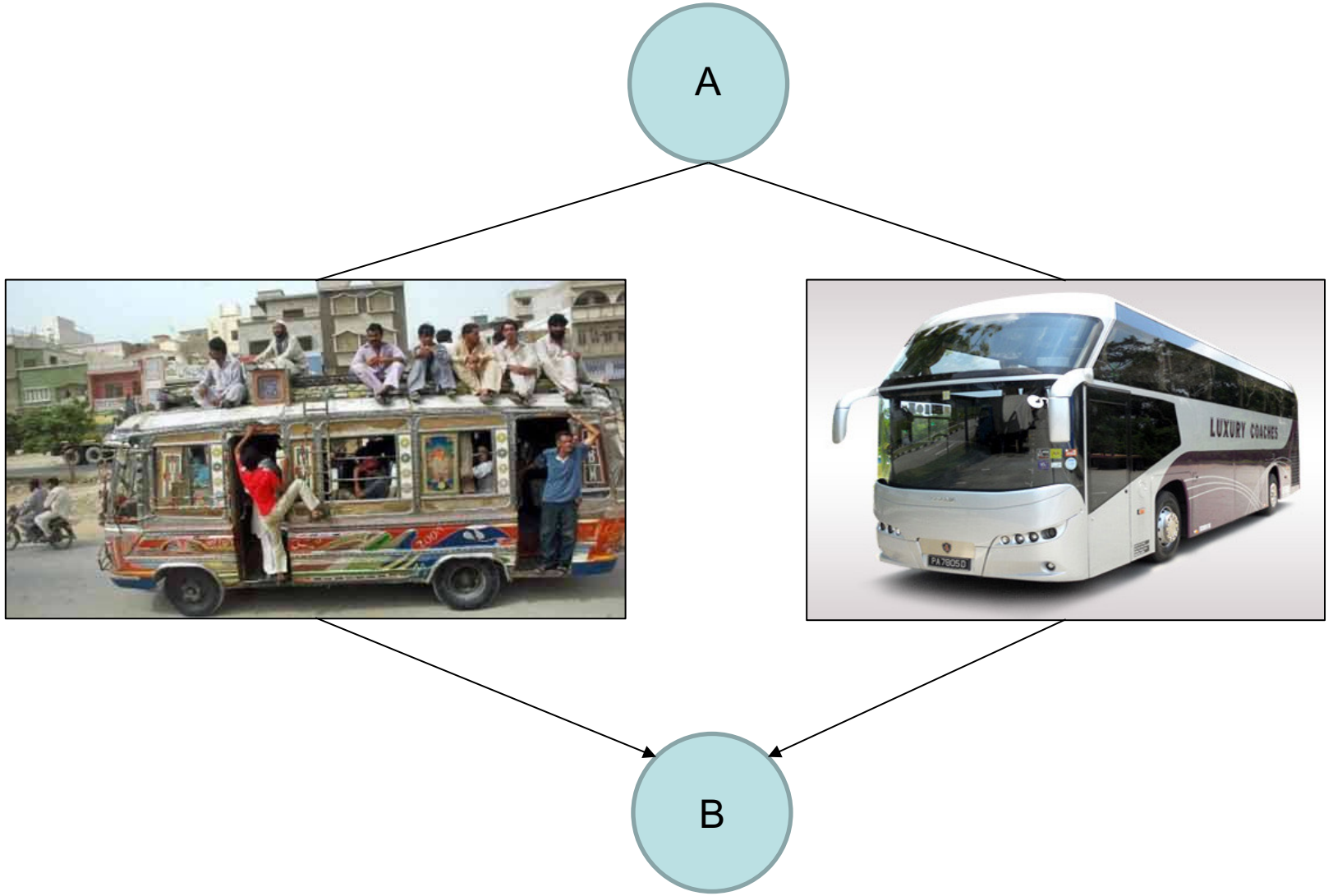
Customers (should) choose whether to use a service or not on the basis of the *contract* that is (or should) exposed by the service provider

… but … do customers *read* contracts at all?




Service-based economy
QoS & SLAs

QoS matters, though



QoS matters, though

To store all the data of my school, I have chosen the storage service offered by the new local software company *StoreForLess.*

We will save 50 USD per year if compared for instance with using Amazon S3!

Service Level Agreements

Google Compute Engine SLA

The Covered Service will provide a Monthly Uptime Percentage to Customer as follows (the "Service Level Objective" or "SLO"):

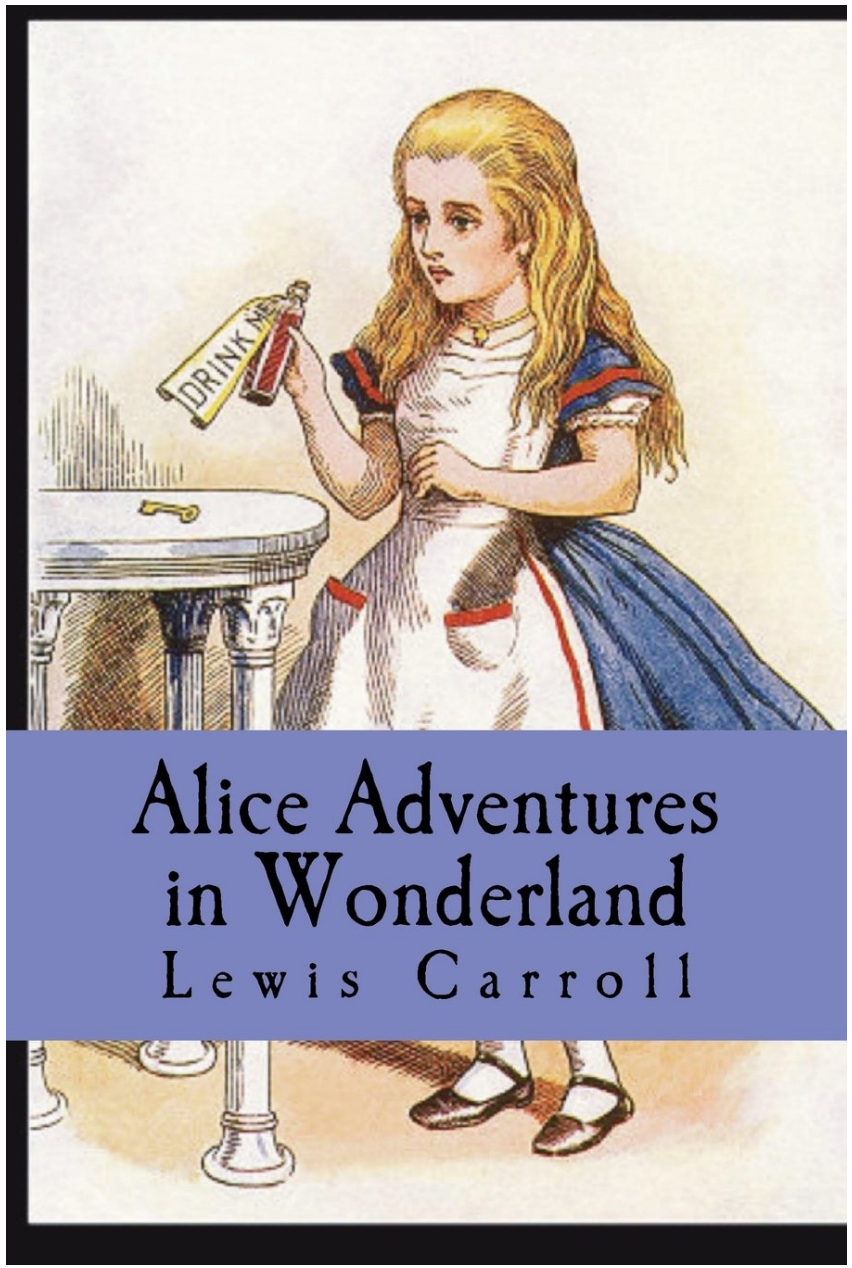
Covered Service	Monthly Uptime Percentage
A Single Instance	>= 99.5%

If Google does not meet the SLO, and if Customer meets its obligations under this SLA, Customer will be eligible to receive the Financial Credits described below.

Monthly Uptime Percentage	Percentage of monthly bill for a Single Instance in the Region that did not meet SLO that will be credited to Customer's future monthly bills
95.00% - < 99.50%	10%
90.00% - < 95.00%	25%
< 90.00%	100%

Q: How much credit can I get if my virtual machine instance was not reachable from 8am to 5:30pm on April 1st and 2nd?
(I tried to connect to it every 30', each time for 5' - no success)

SLA ≠



Google Compute Engine SLA

Monthly Uptime Percentage	Percentage of monthly bill for a Single Instance in the Region that did not meet SLO that will be credited to Customer's future monthly bills
95.00% - < 99.50%	10%
90.00% - < 95.00%	25%
< 90.00%	100%

Q: How much credit can I get if my virtual machine instance was not reachable from 8am to 5:30pm on April 1st and 2nd?
(I tried to connect to it every 30', each time for 5' - no success)

"Monthly Uptime Percentage" means total number of minutes in a month, minus the number of minutes of Downtime suffered from all Downtime Periods in a month, divided by the total number of minutes in a month.

"Downtime Period" means a period of one or more consecutive minutes of Downtime. Partial minutes or intermittent Downtime for a period of less than one minute will not count towards any Downtime Periods.

Customer Must Request Financial Credit In order to receive any of the Financial Credits described above, Customer must notify Google technical support within 60 days from the time Customer becomes eligible to receive a Financial Credit. Customer must also provide Google with server log files showing loss of external connectivity errors and the date and time those errors occurred. If Customer does not comply with these requirements, Customer will forfeit its right to receive a Financial Credit.

$$\frac{(30 \times 24 \times 60) - (2 \times 20 \times 5)}{30 \times 24 \times 60} = \frac{43000}{43200} = 0{,}995370370$$

Microsoft Services Agreement

6. Service Availability

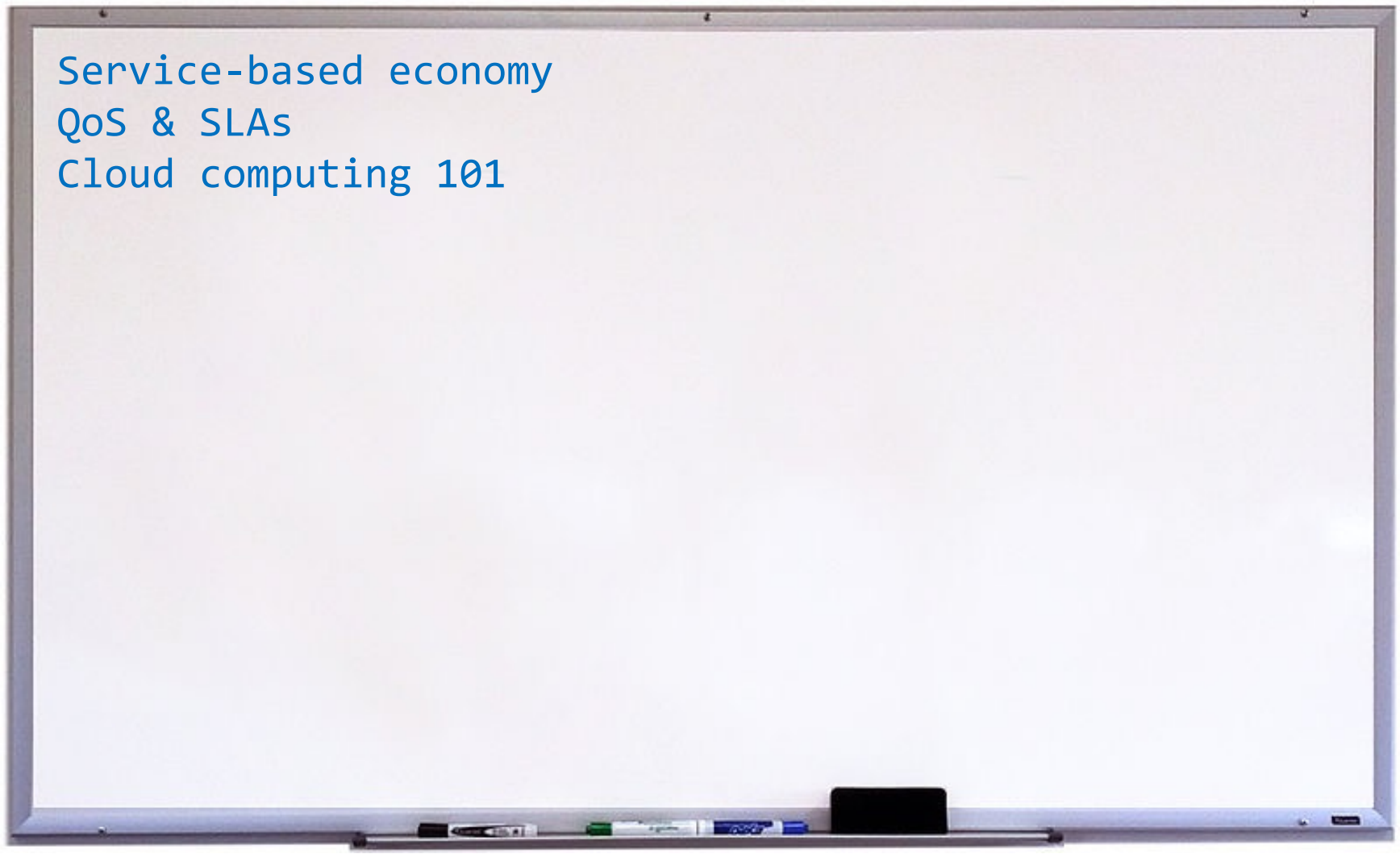
[…] We strive to keep the Services up and running; however, all online services suffer occasional disruptions and outages, and Microsoft is not liable for any disruption or loss you may suffer as a result. In the event of an outage, you may not be able to retrieve Your Content or Data that you've stored. We recommend that you regularly backup Your Content and Data that you store on the Services or store using Third-Party Apps and Services.

11. **Warranties.**

a. MICROSOFT, AND OUR AFFILIATES, RESELLERS, DISTRIBUTORS, AND VENDORS, MAKE NO WARRANTIES, EXPRESS OR IMPLIED, GUARANTEES OR CONDITIONS WITH RESPECT TO YOUR USE OF THE SERVICES. YOU UNDERSTAND THAT USE OF THE SERVICES IS AT YOUR OWN RISK AND THAT WE PROVIDE THE SERVICES ON AN "AS IS" BASIS "WITH ALL FAULTS" AND "AS AVAILABLE." MICROSOFT DOESN'T GUARANTEE THE ACCURACY OR TIMELINESS OF THE SERVICES. YOU MAY HAVE CERTAIN RIGHTS UNDER YOUR LOCAL LAW. NOTHING IN THESE TERMS IS INTENDED TO AFFECT THOSE RIGHTS, IF THEY ARE APPLICABLE. YOU ACKNOWLEDGE THAT COMPUTER AND TELECOMMUNICATIONS SYSTEMS ARE NOT FAULT-FREE AND OCCASIONAL PERIODS OF DOWNTIME OCCUR. WE DO NOT GUARANTEE THE SERVICES WILL BE UNINTERRUPTED, TIMELY, SECURE, OR ERROR-FREE OR THAT CONTENT LOSS WON'T OCCUR, NOR DO WE GUARANTEE ANY CONNECTION TO OR TRANSMISSION FROM COMPUTER NETWORKS.

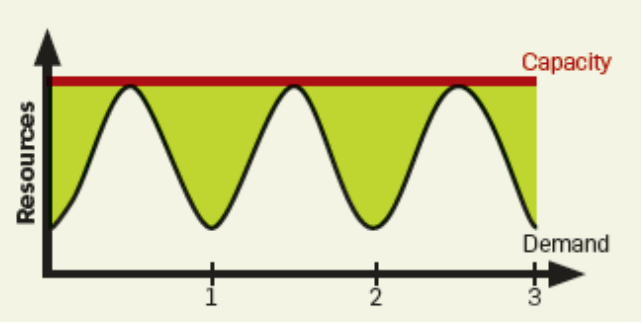
[…]

https://www.microsoft.com/en/servicesagreement/


Service-based economy
QoS & SLAs
Cloud computing 101

Estimating service demand

Service demand changes with time, nobody knows how



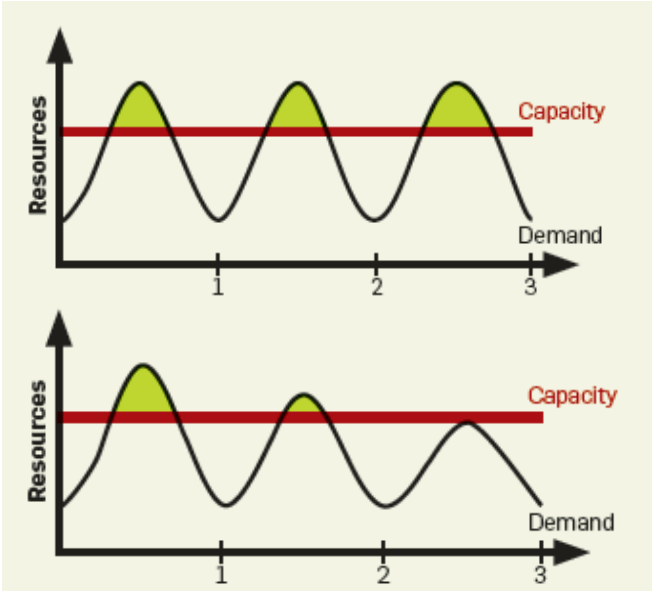
Overprovisioning

Ensuring in advance provisioning for expected demand peaks (due to diurnal or seasonal patterns or unexpected demand bursts) leads to wasting resources (if prediction is correct - even worse if spike is overestimated)

Underprovisioning

If spike is underestimated then underprovisioning may accidentally turn away excess users

Cost of underprovisioning more difficult to measure, but as serious as cost of overprovisioning - not only do rejected users generate zero revenue, they may never come back




Seemingly infinite resources,
available on demand

NIST National Institute of Standards and Technology U.S. Department of Commerce

2. The NIST Definition of Cloud Computing

Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

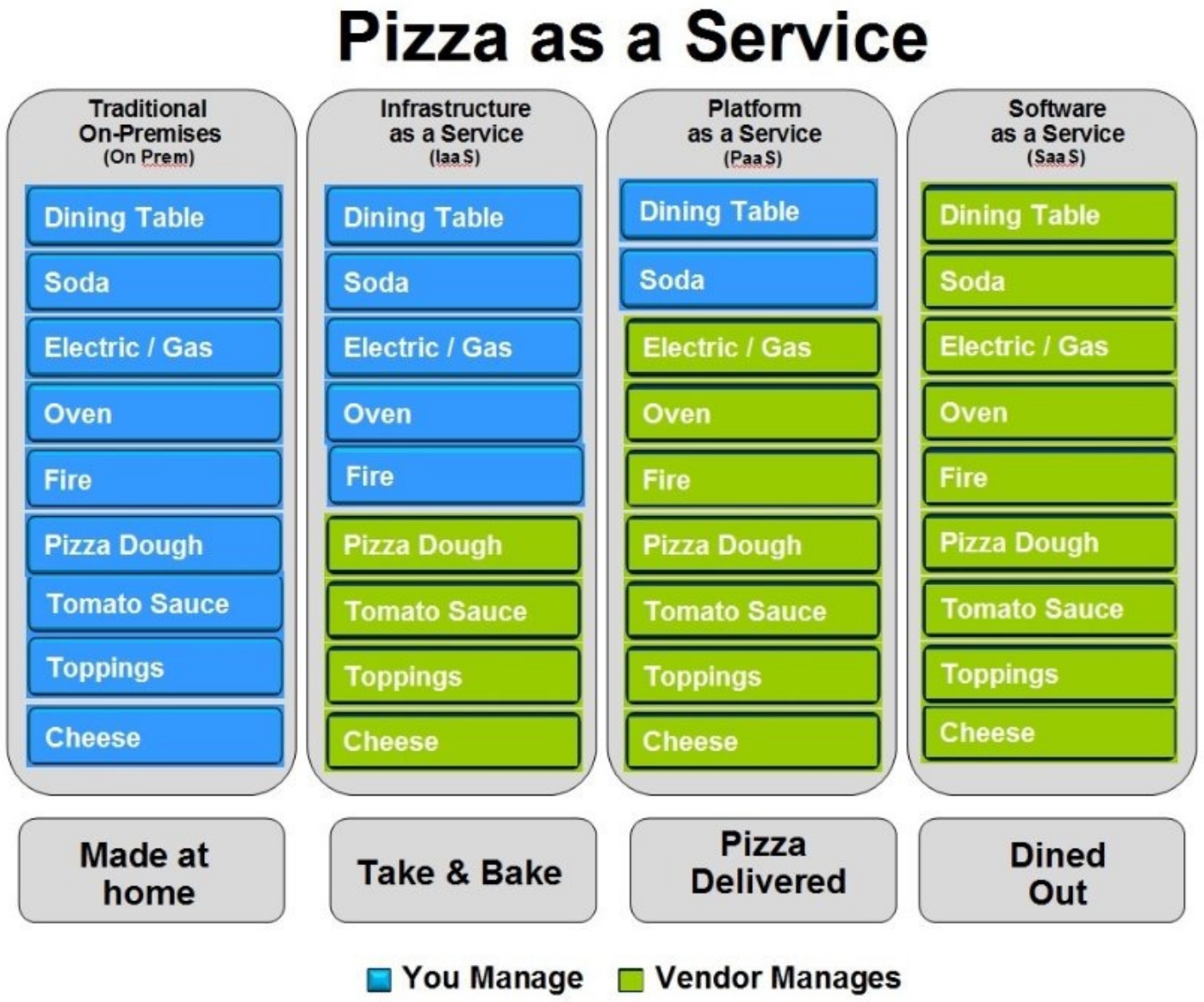
Key ideas

- Efficient pooling of on-demand, self-managed virtual infrastructures, consumed as services
- Delivery of dynamically scalable, virtualised resources over the Internet to multiple clients
- Decoupling delivery of computing services from underlying technology

Economics

- Elimination of up-front commitment by cloud users (converting CapEx to OpEx)
- Pay-per-use
  - customers just love this! ❤️
  - even if pay-per-use more expensive, cost is outweighed by economical benefits of ***elasticity*** and ***transference of risk***

Service models



[Albert Barron - IBM Software]

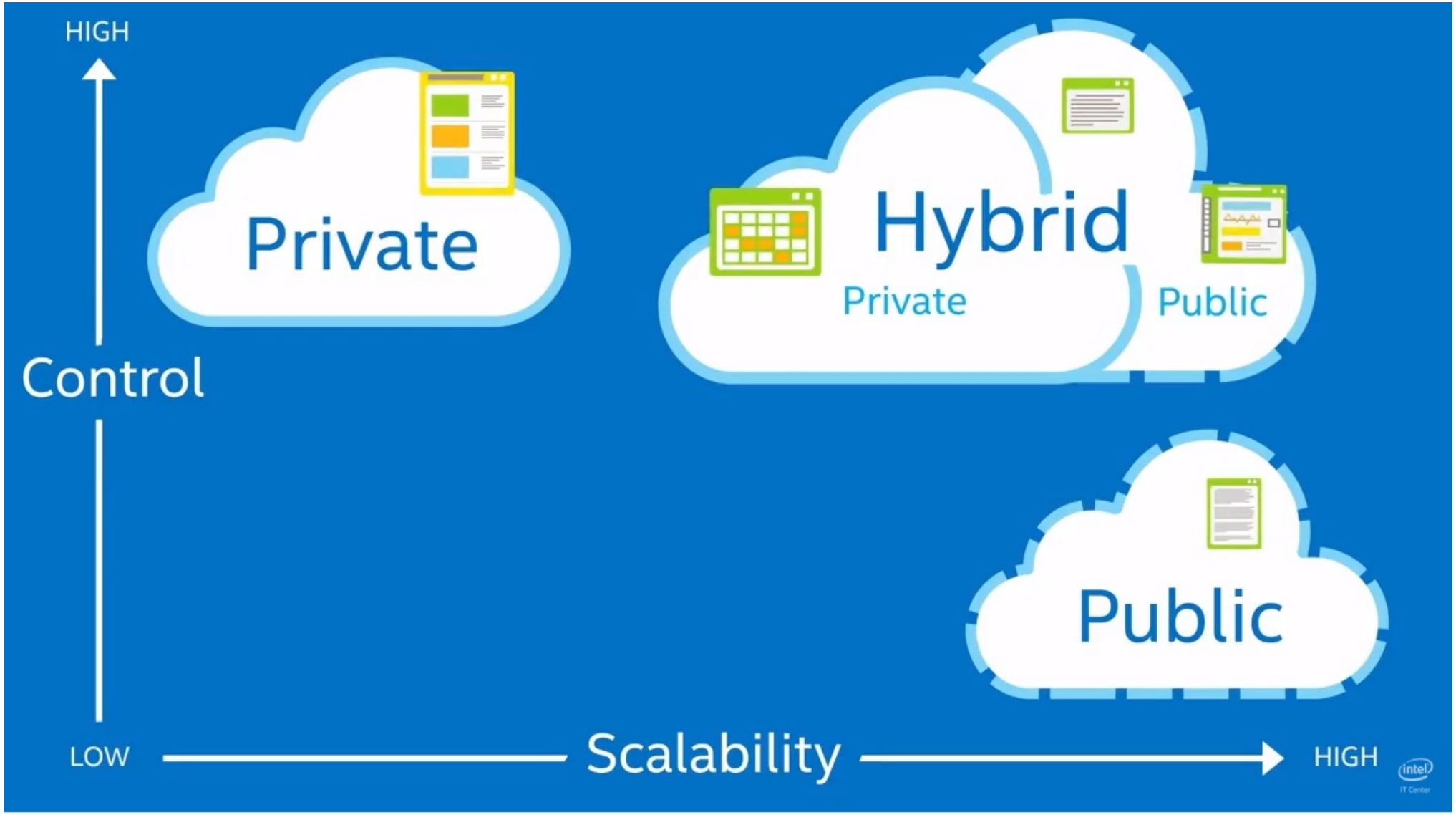
Service models

- SaaS provides software on-demand for use, accessible via thin clients or APIs
- SaaS provider manages infrastructure + OS + app
- Client responsible for nothing
- Example: salesforce.com

- PaaS provides whole platform as a service (VMs, OS, services, SDKs,…)
- PaaS provider manages infrastructure + OS + enabling SW
- Client responsible for installing and managing app
- Examples: Heroku, Azure, GAE

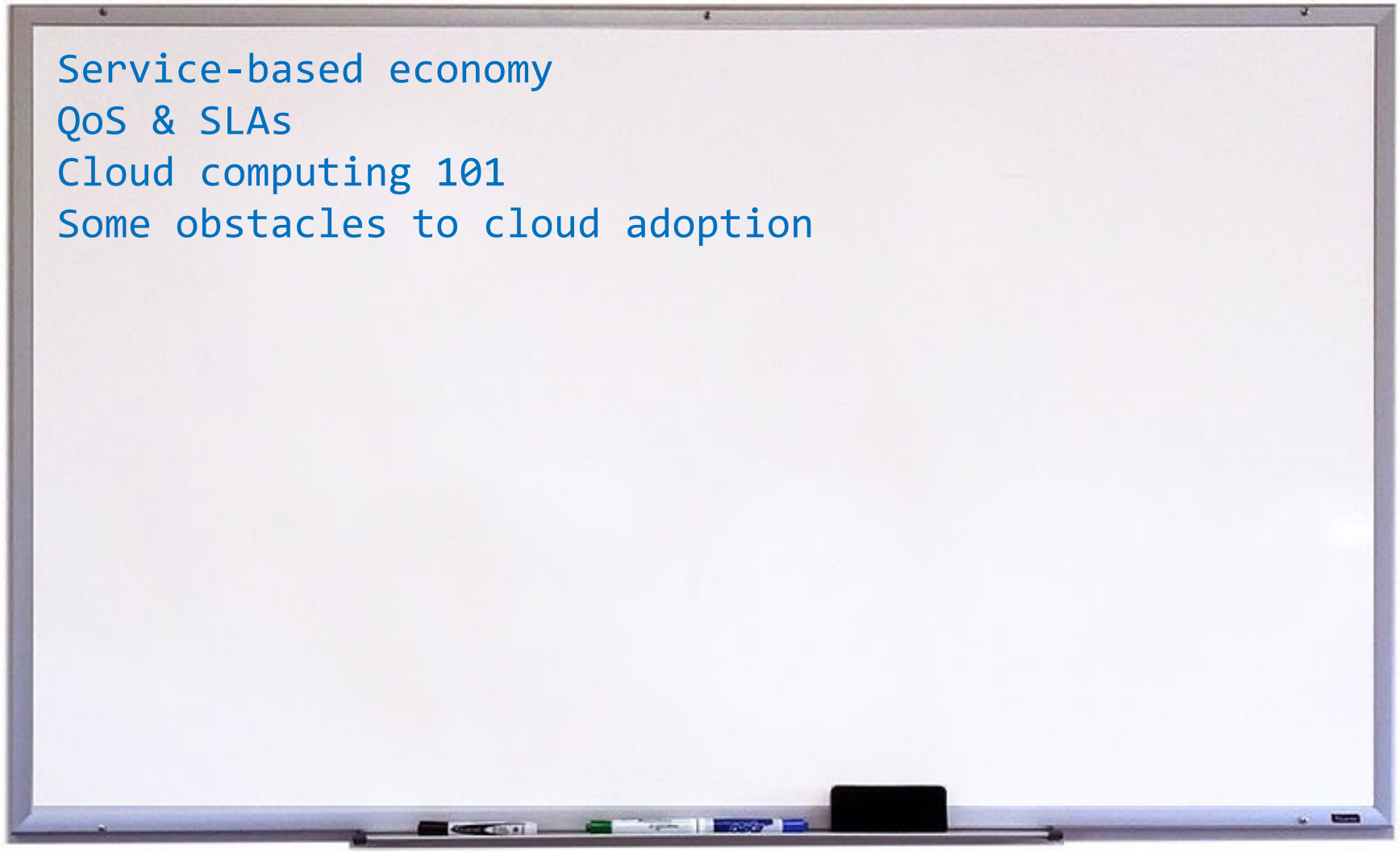
- IaaS provides (virtualized) servers, storage, networking
- IaaS provider manages all infrastructure
- Client responsible for all other aspects of the deployment (e.g., OS, app)
- Example: EC2, S3

Deployment models



Public vs. private vs. hybrid cloud

https://www.youtube.com/watch?v=h6BzHP-oIKs


Service-based economy
QoS & SLAs
Cloud computing 101
Some obstacles to cloud adoption

Some obstacles to cloud adoption

Hi Boss! Why don’t we move
all company data on the cloud?

Some obstacles to cloud adoption

Data confidentiality

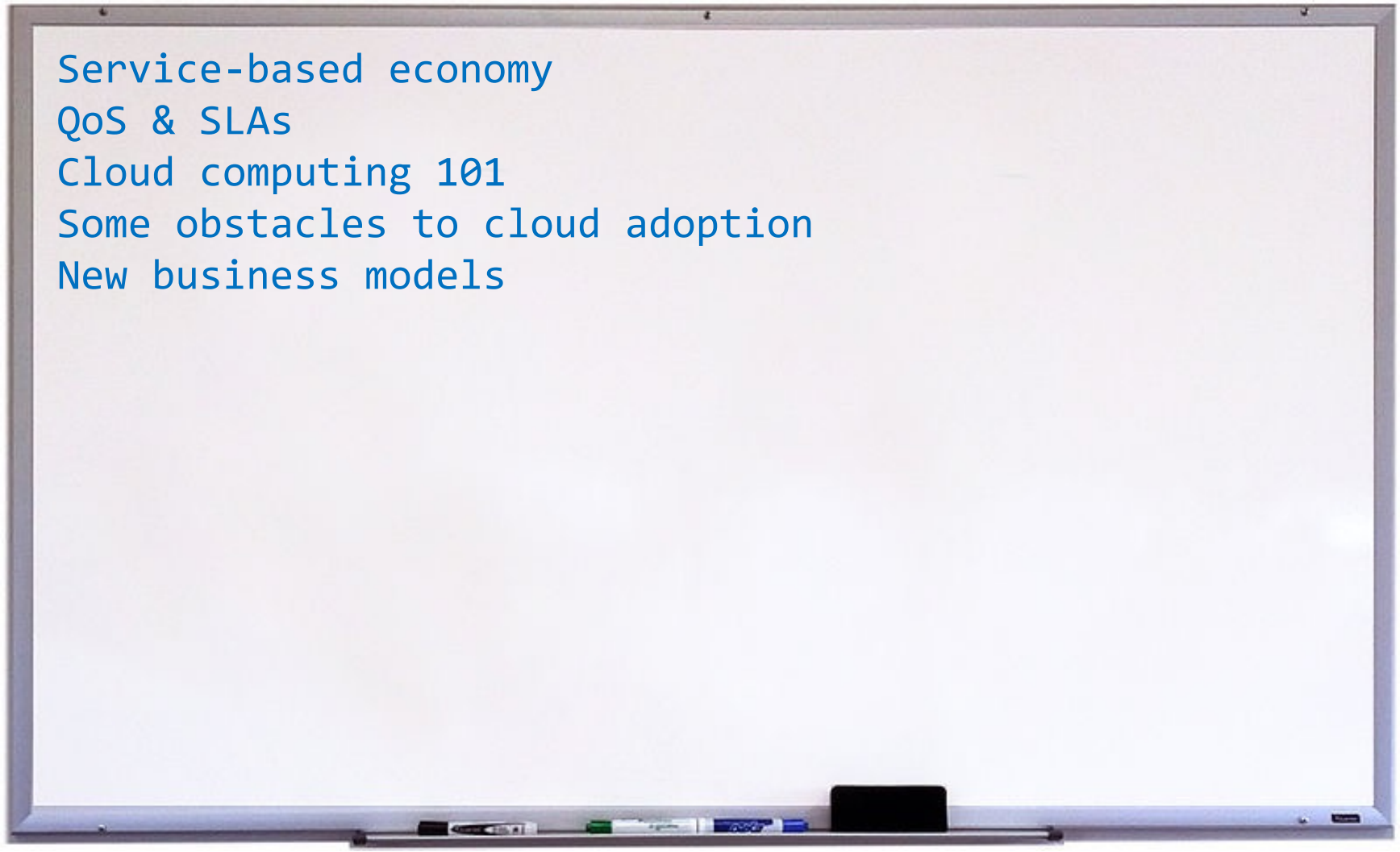
- Where will our data be stored, concretely?
- Will data privacy and integrity be guaranteed? How?
- How can we know whether a problem occurred?

Dropbox Security Bug Made Passwords Optional

Business continuity / service availability

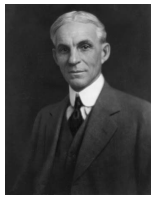
- What if cloud provider goes out of business?
- CS mantra: “no single point of failure”...

Vendor lock-in ?


Service-based economy
QoS & SLAs
Cloud computing 101
Some obstacles to cloud adoption
New business models

New business models

If I had asked people what they wanted, they would have said “faster horses”

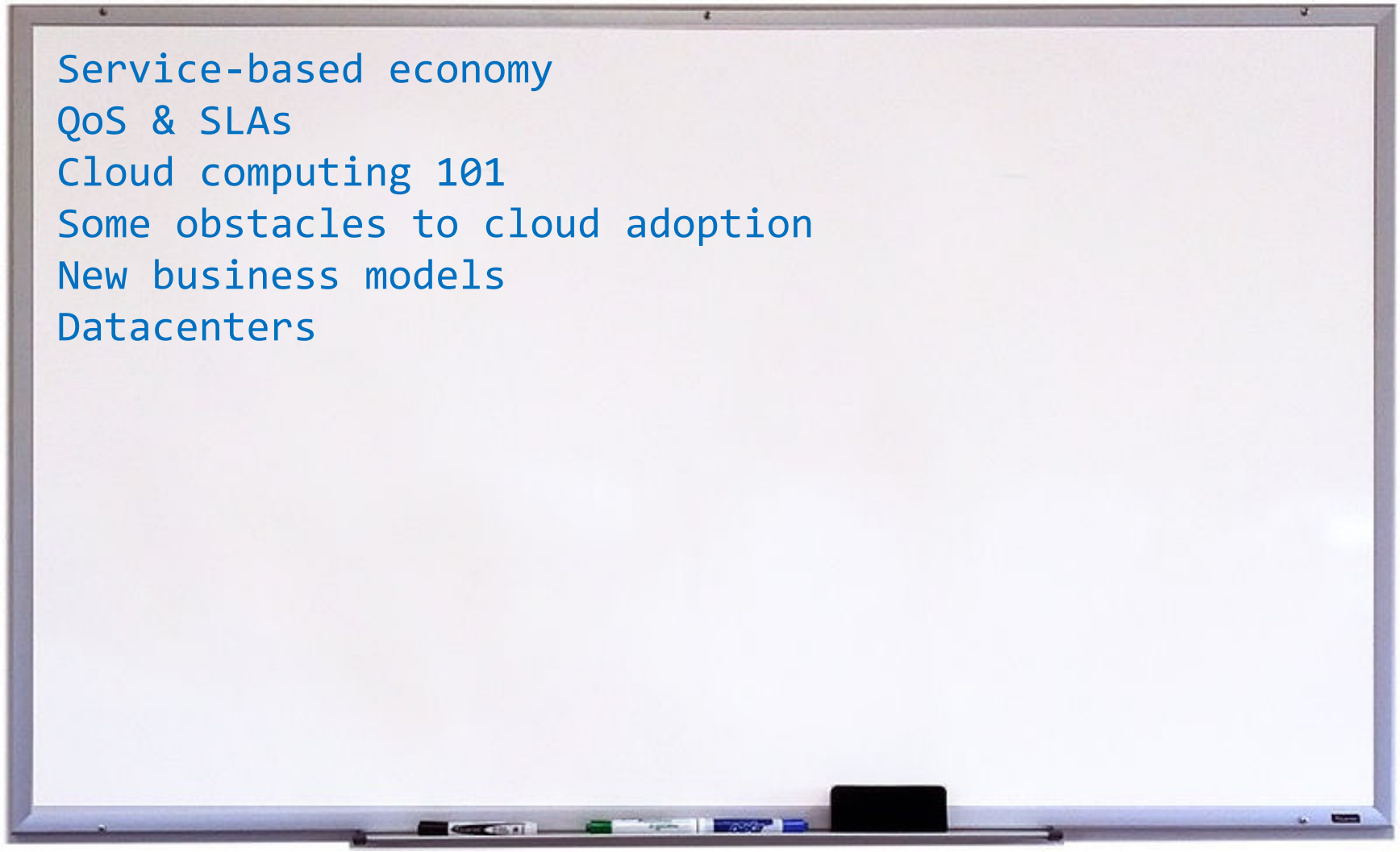


Henry Ford

What if … we provide free music to everybody? — Spotify — Freemium

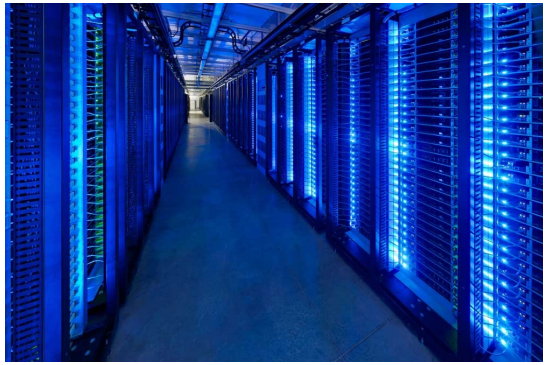
What if … we provide free storage to everybody? — Freemium

What if … we provide a free search engine to everybody? — Google — Customised advertising

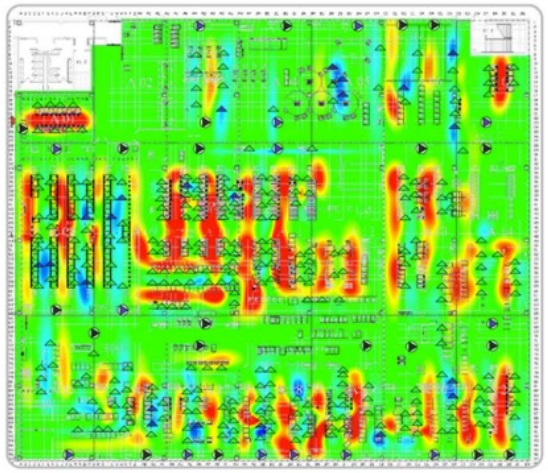

Service-based economy
QoS & SLAs
Cloud computing 101
Some obstacles to cloud adoption
New business models
Datacenters

Amazon

Google

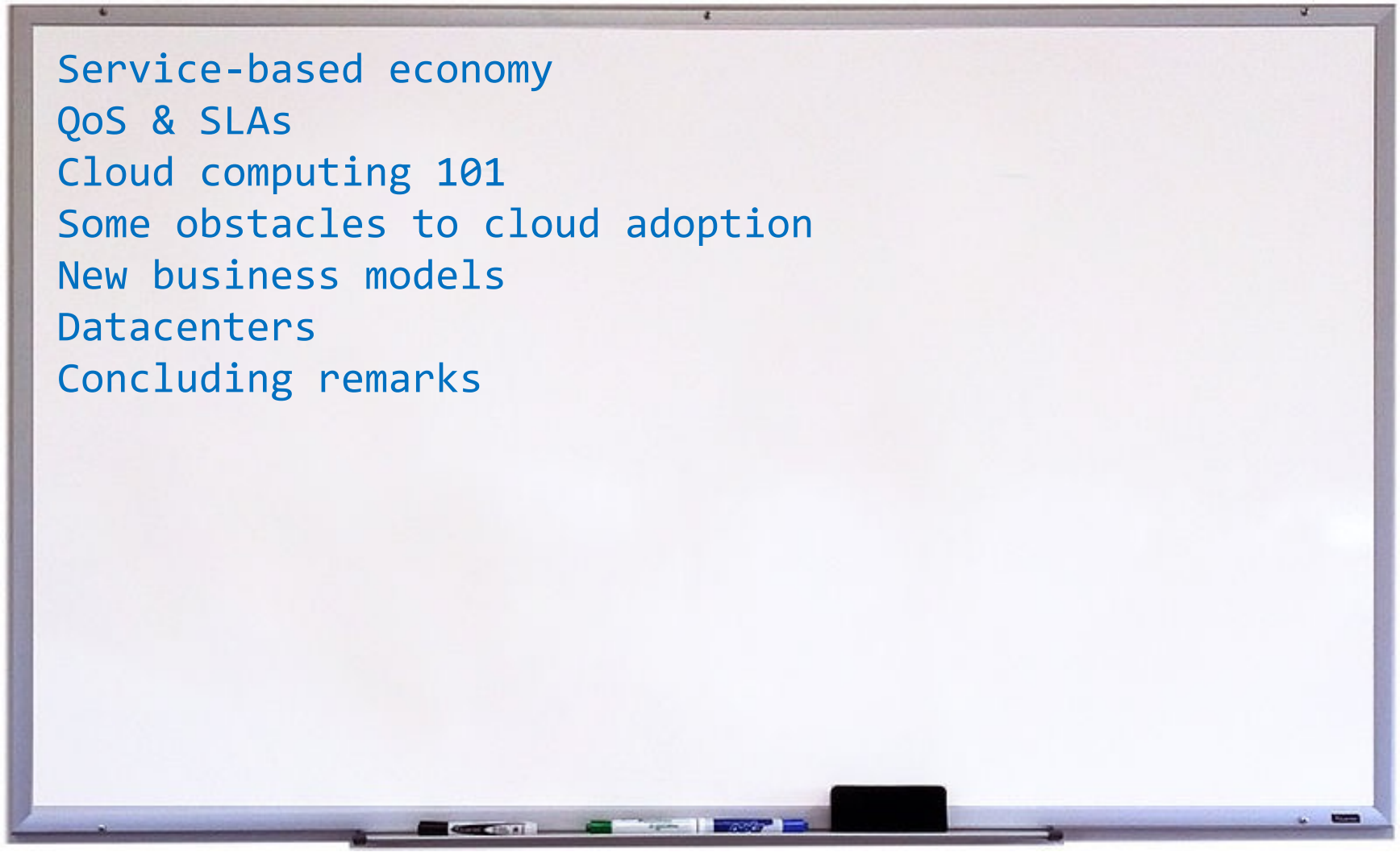


Facebook

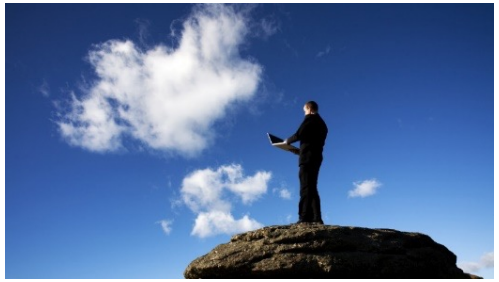


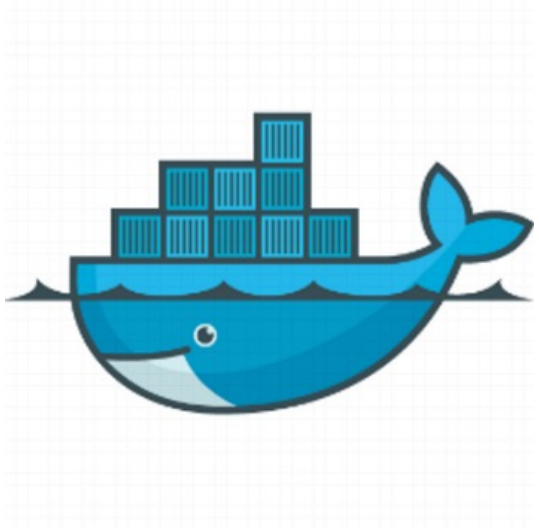
Cooling

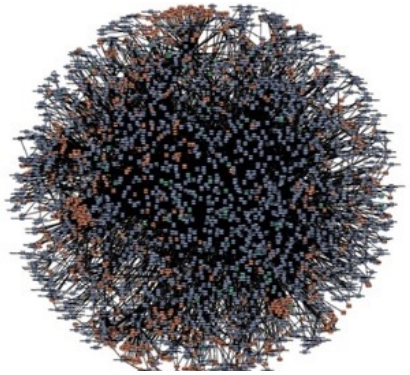
DCIM


Service-based economy
QoS & SLAs
Cloud computing 101
Some obstacles to cloud adoption
New business models
Datacenters
Concluding remarks

Cloud computing is here to stay







(PS: Why the name «cloud»?)