IOWA

CS5630

Foundation (5): Security and the Cloud

Prof. Supreeth Shastri
Computer Science
The University of Iowa

Research Paper: Review and Presentation

17 teams

30 mins **presentation** + **15** mins **discussion**

- We will start the student-led presentations from 9/28
- Look for a teammate (e.g., on Slack)
- Select a paper that you find interesting (or propose one)
- Slots for presentation will be opened tonight
- Please ask all your questions/concerns

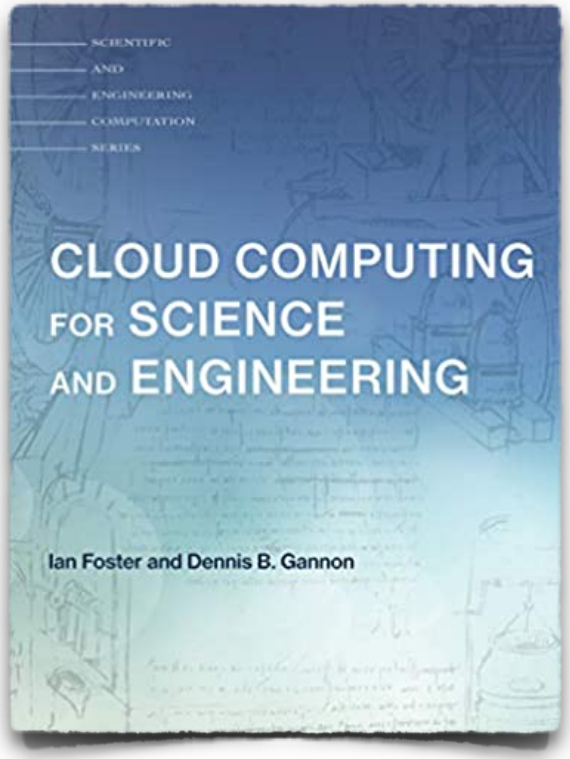
Tasks	Weight
Paper review	25%
Presentation	50%
Handling Q&A	25%

Remember that this activity makes up 25% of your final grade

Lecture goals

An exploration of security and privacy of data and computing on the cloud

- *Shared model of cloud security*
- *Security from a cloud user perspective*
- *Case study: GDPR and the cloud*



Chapter 15

aws

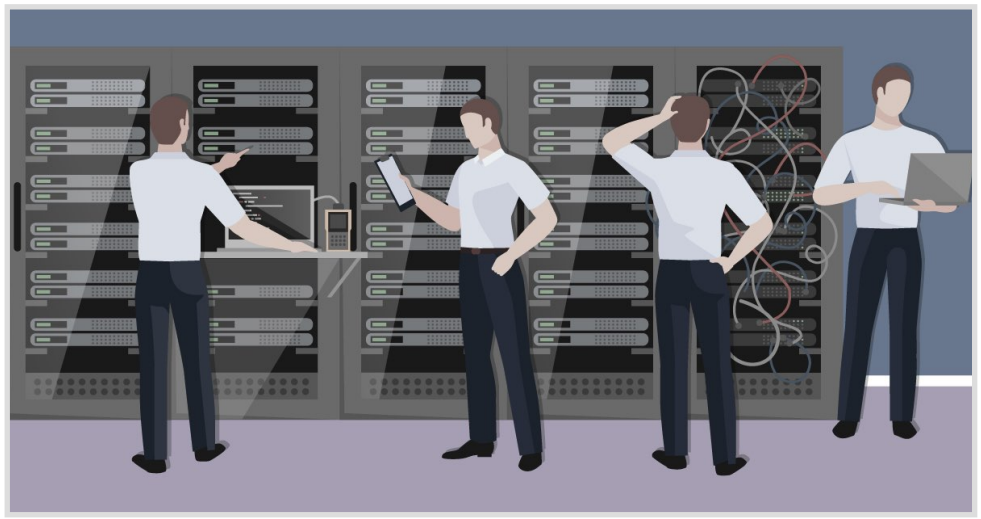
Google Cloud

IOWA

Which of the following offers better security and privacy?

Software and data hosted
on-premises

Software and data residing
on ***the cloud***



Thinking about security in the cloud

The ultimate security is your understanding of the reality — Stanley Judd

The Internet

Any time your computing infrastructure is exposed to the Internet, security is a concern.

Cloud computing, by definition, is resources offered over the Internet as a service.

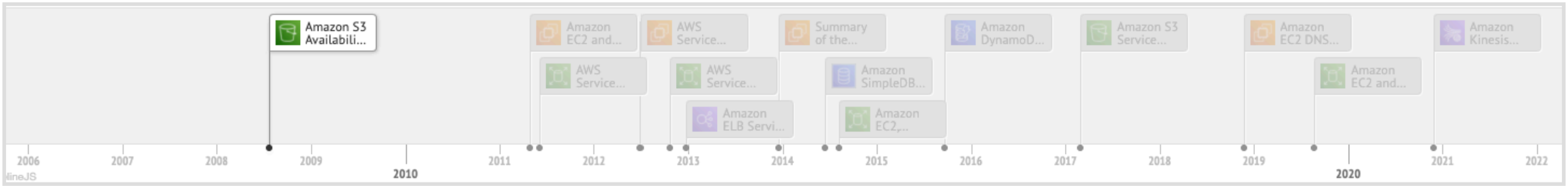
Security “in” the cloud vs. security “of” the cloud

Cloud providers have expert security teams, whose job is protect their hardware and software.

However, just because Amazon EC2 is safe does not necessarily mean your apps and data are!

One of the biggest security risk in cloud computing is:
someone breaking into your personal computer and then accessing your cloud resources from there

How secure is the cloud?



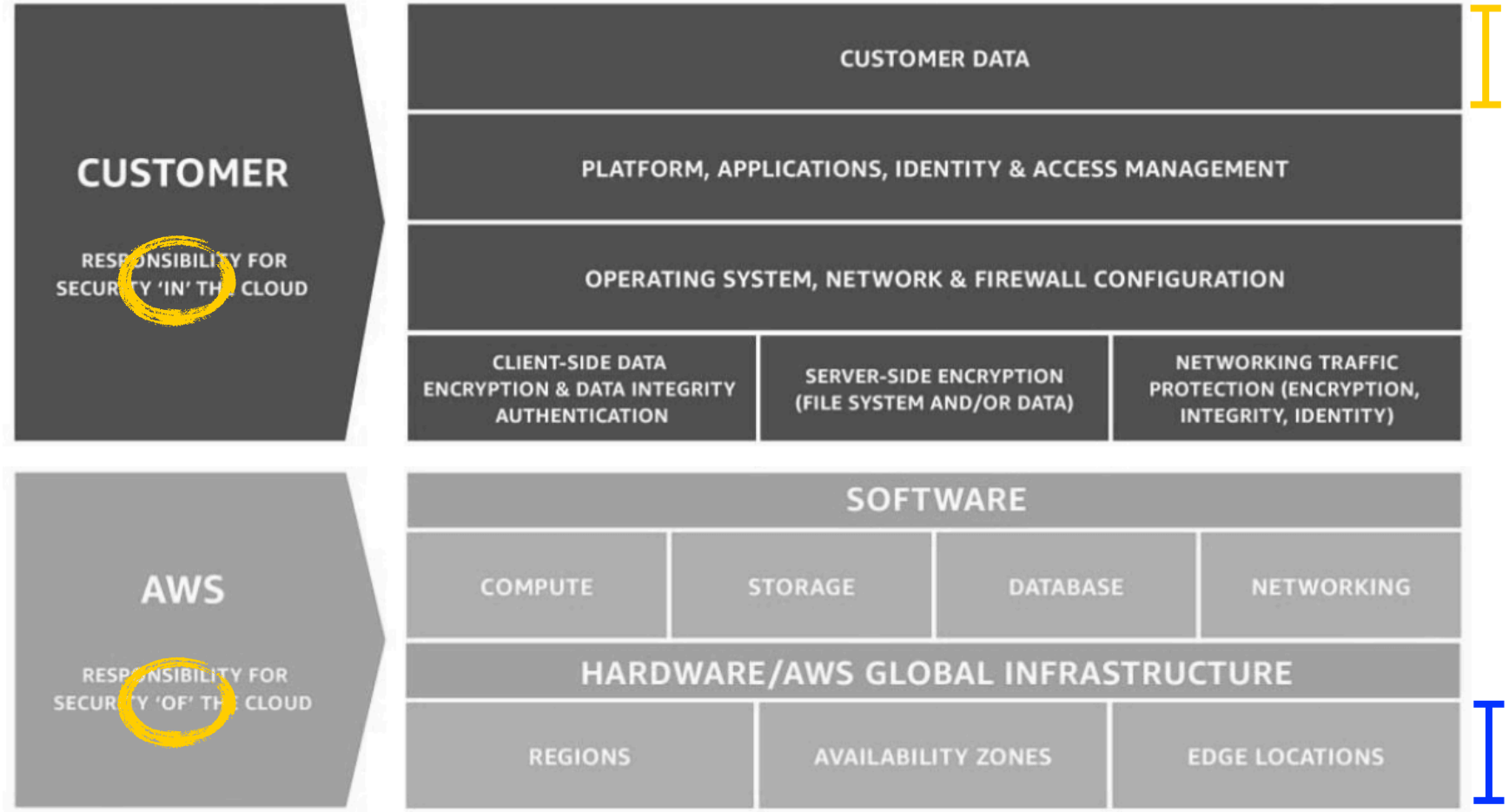
Major outages in AWS

- 14 outages/system failures in all its existence (2006 till now)
- All but one outages were fixed within the same day!
- Explore more: https://awsmaniac.com/aws-outages/

US-East-1 (Virginia) outage of 2012

- Storm on 6/29 night —> a spike shut down the electric grid at 7:24pm —> AWS backup generator failed to start in 1/10 datacenters —> UPS kicked in but lasted till 8:04pm; then servers began shutting down —> onsite personnel fixed the backup generator at 8:14 pm—> full datacenter back online at 8:24pm
- Majority of EC2 and EBS servers fully operational by 12:25am; ~10% of EBS had in-flight data during failure —> consistency checks/restoration took 2 additional days (all I/O paused for those disks)
- RDS that relied on affected EBS disks remained unavailable for ~2 days

AWS Shared Responsibility Model



- *What data to store?*
- *Where to store it?*
- *How to store it?*
- *Who to give access to?*

- *Are the buildings safe?*
- *Is electric grid functioning?*
- *Did solar storms affect magnetic storage?*

Azure Shared Responsibility Model

	SaaS	PaaS	IaaS	On prem
Data classification & accountability	You	You	You	You
Client & endpoint protection	Shared	You	You	You
Identity & access management	Shared	Shared	You	You
Application-level controls	Provider	Shared	You	You
Network controls	Provider	Provider	Shared	You
Host infrastructure	Provider	Provider	Shared	You
Physical security	Provider	Provider	Provider	You

- **Responsibilities change depending on the service model.** *For e.g., applying latest OS/software patches is Azure responsibility in SaaS whereas it is customer's responsibility in IaaS.*
- **Separation of mechanism vs. policy.** *For e.g., Azure provides tools for access control; customer policies govern how they're configured.*

Security from a cloud user perspective

1. Fine-grained access control and usage monitoring

- Cloud customers get to control what every user (within their org) can and cannot do on the cloud
  - AWS provides *Identity and Access Management (IAM)*
- Customers have the ability to monitor resource usage — both by users and by applications
  - *AWS CloudTrail* provides a detailed history of all APIs invoked and events/triggers observed

2. Securing data in the cloud

- When data is transferred in to and out of the cloud (i.e., data in transit) ==> use *Secure Socket Layer (SSL)*
- when data is stored in the cloud (i.e., data at rest) ==> use *client-side encryption* or *server-side encryption*

```
# Upload the file 'test.jpg' into the newly created bucket
s3.Object('datacont', 'test.jpg').put(
    Body=open('/home/mydata/test.jpg', 'rb'),
    ServerSideEncryption='AES256')
```

Security from a cloud user perspective

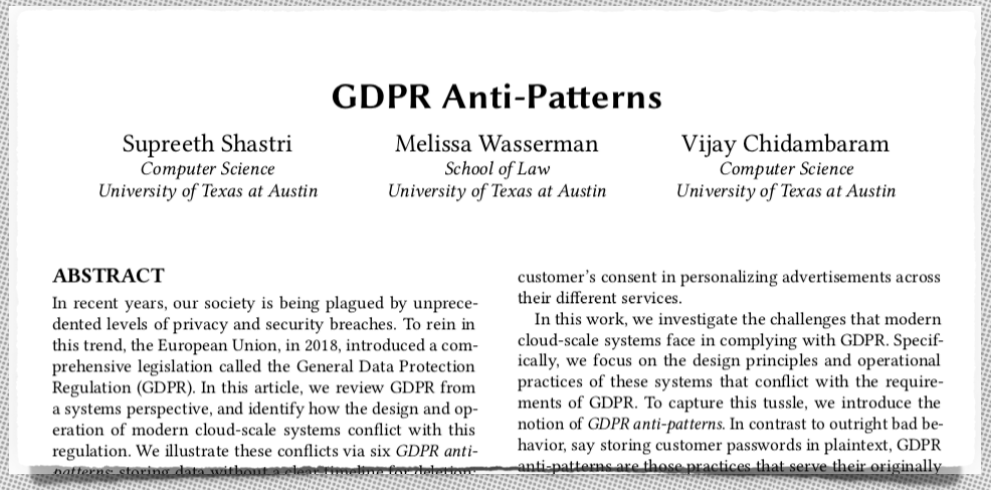
3. Securing compute in the cloud

- Poisoned VM image that allow attackers to gain control of your server
  - Use VM/container images from *trusted sources* (e.g., AWS, Ubuntu etc)
- Illegal access to running VMs
  - *IAM* for access control; *SSH public-key* for secure remote logins
- Intercepting communications between VMs
  - Configure a *Virtual Private Network (VPN)*

GDPR and the Cloud

New rules on storing and processing personal data

Do you need to do anything differently when you store and process someone else's (personal) data on the cloud?



GDPR Anti-Patterns

Supreeth Shastri
Computer Science
University of Texas at Austin

Melissa Wasserman
School of Law
University of Texas at Austin

Vijay Chidambaram
Computer Science
University of Texas at Austin

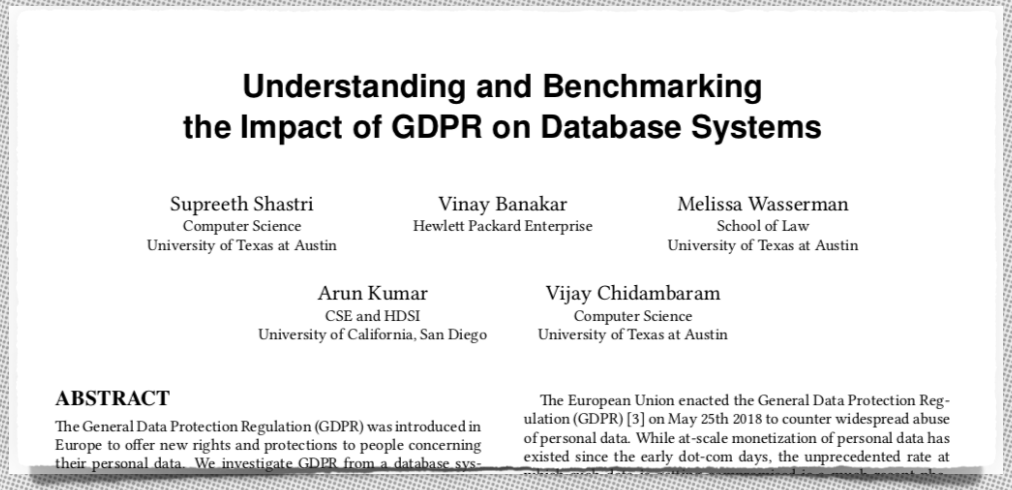
ABSTRACT

In recent years, our society is being plagued by unprecedented levels of privacy and security breaches. To rein in this trend, the European Union, in 2018, introduced a comprehensive legislation called the General Data Protection Regulation (GDPR). In this article, we review GDPR from a systems perspective, and identify how the design and operation of modern cloud-scale systems conflict with this regulation. We illustrate these conflicts via six *GDPR anti-patterns*: storing data without acknowledging for deletion

customer's consent in personalizing advertisements across their different services.

In this work, we investigate the challenges that modern cloud-scale systems face in complying with GDPR. Specifically, we focus on the design principles and operational practices of these systems that conflict with the requirements of GDPR. To capture this tussle, we introduce the notion of *GDPR anti-patterns*. In contrast to outright bad behavior, say storing customer passwords in plaintext, GDPR anti-patterns are those practices that serve their originally

CACM 2021



Understanding and Benchmarking the Impact of GDPR on Database Systems

Supreeth Shastri
Computer Science
University of Texas at Austin

Vinay Banakar
Hewlett Packard Enterprise

Melissa Wasserman
School of Law
University of Texas at Austin

Arun Kumar
CSE and HDSI
University of California, San Diego

Vijay Chidambaram
Computer Science
University of Texas at Austin

ABSTRACT

The General Data Protection Regulation (GDPR) was introduced in Europe to offer new rights and protections to people concerning their personal data. We investigate GDPR from a database sys-

The European Union enacted the General Data Protection Regulation (GDPR) [3] on May 25th 2018 to counter widespread abuse of personal data. While at-scale monetization of personal data has existed since the early dot-com days, the unprecedented rate at

VLDB 2020

General Data Protection Regulation (GDPR)

2017 — May 2018 — 2019

1300+

Data breaches in 2017

Google

€50M

French Data Protection Authority, Jan 2019

BRITISH AIRWAYS

£183M

UK Data Protection Agency, Jun 2019

Marriott HOTELS & RESORTS

$123M

UK Data Protection Agency, Jun 2019

Public *Complaints*

144,376

EU-wide (Year 1)

General Data Protection Regulation (GDPR)

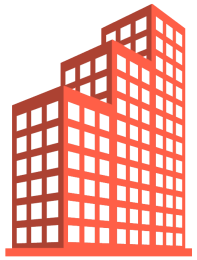
Privacy and protection of personal data is a fundamental right of natural persons

99 Legal Articles

Regulate the collection, processing, protection, transfer and deletion of personal data

Grants Rights to People

Grants all European people a right to protection and privacy of their personal data



Assigns Responsibilities to Companies

Those who collect and process personal data are solely responsible for its privacy and protection

Hefty Penalty

Max penalty of 4% of global revenue or €20 million, whichever is greater

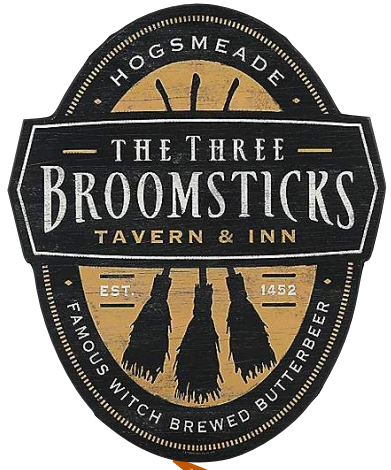
Personal Data

any information relating to an identified or identifiable natural person

GDPR §4(1)

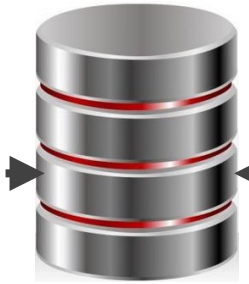
Prof. Albus Dumbledore

- ☐ *Has a phoenix as pet*
- ☐ *Drinks coffee at 8am*
- ☐ *Lectured at Iowa on Sep-9*



I have
eight rights!

Right to know, access, rectify, erase, object, port, restrict processing, and withdraw from automated processing



I have
responsibilities

To obtain consent, track data usage, keep it secure, notify breaches etc.

Store Data with a Timeline for Deletion

§ 5(1)(E): Storage Limitation

"[...] kept for no longer than is necessary for the purposes for which the personal data are processed [...]"

§17: Right To Be Forgotten

(1) The data subject shall have the right to obtain from the controller the erasure of personal data without undue delay [...]

GDPR-compliant datastore should:

Associate a **time-to-live** attribute with all data

Have support for **timely deletion** of data

Keep Record of Data Processing Activity

§ 30: Records of Processing Activities

(1) Each controller [...] shall maintain a record of processing activities under its responsibility.

§ 33: Notification of A Data Breach

(1) the controller shall without undue delay and not later than 72 hours after having become aware of it, notify [...] (3) The notification shall at least describe the nature of the personal breach."

GDPR-compliant datastore should:

Associate an **audit trail** with all data

Implement support for **monitoring/logging** of all data accesses

Translating GDPR Articles into Systems-Level Attributes and Actions

We analyzed all the 99 articles of GDPR, both individually and collectively…

GDPR Metadata

Associate seven behavioral attributes with personal data

1	*Purpose*
2	*TTL*
3	*A…*
	…rigin of data
6	*Externally shared?*
7	*Use in automated decision-making?*

Storage overhead

GDPR Capabilities

Implement five features in the database system

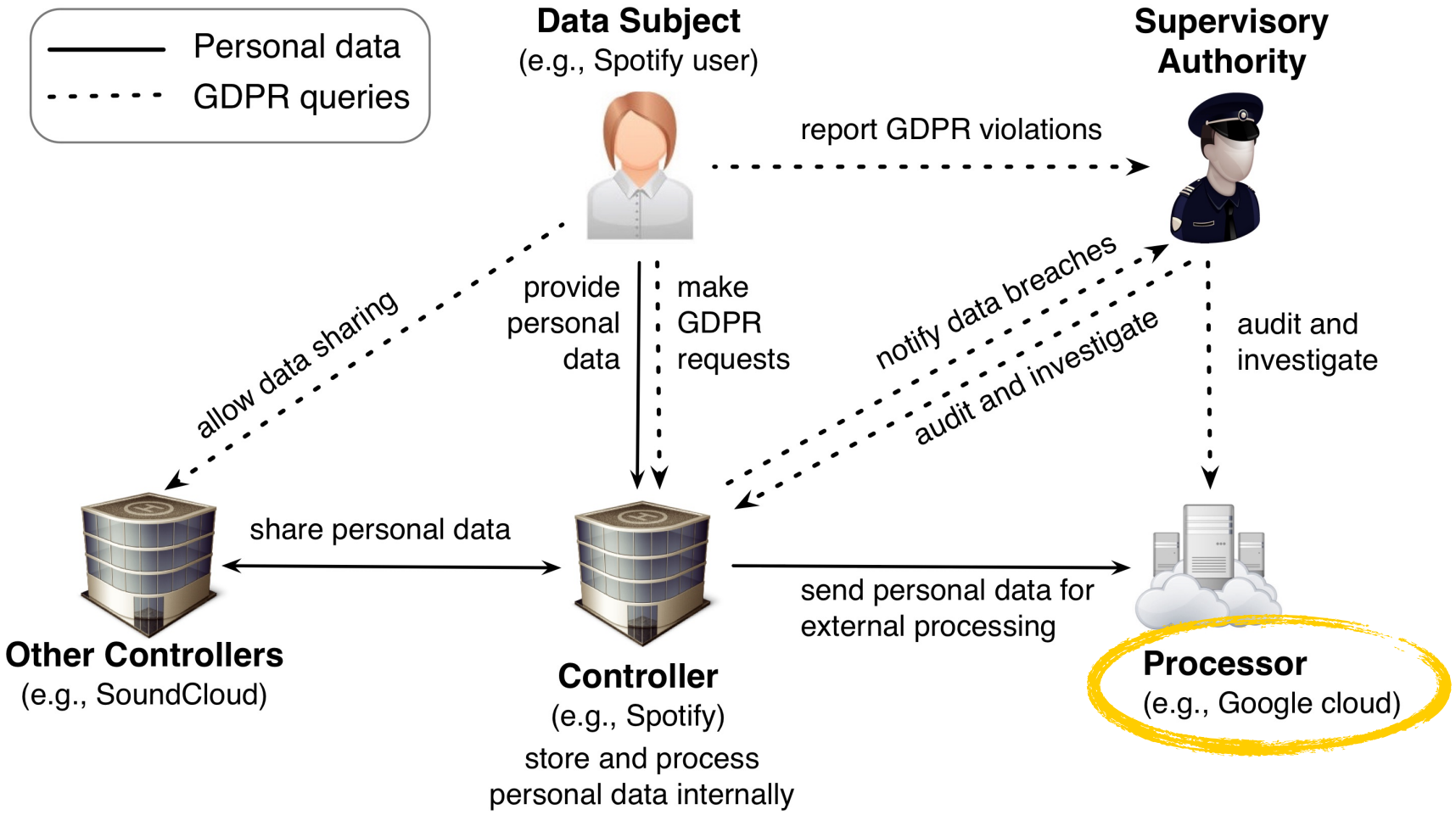
Access control

Timely deletion

Metadata-based querying

Performance overhead

How does this all map to cloud services?



Spot Quiz (ICON)