Cloud-Based Data Processing

Introduction

Jana Giceva

About me

Jana Giceva
Chair for Database Systems
Boltzmannstr. 3, Office: 02.11.043
jana.giceva@cit.tum.de

Academic Background:

- 2011 – 2017 PhD in Computer Science at *ETH Zurich* (topic: DB/OS co-design)
- 2017 – 2019 Lecturer in Department of Computing at *Imperial College London*
- Since 2020 Assistant Professor for Database Systems at *TUM*

Connections with Industry:

- Held roles with *Oracle Labs* and *Microsoft Research* in the USA in 2013 and 2014
- PhD Fellowship from *Google* in 2014
- Early Career Faculty Award from *VMware* in 2019
- Collaborating with *SAP* (joint PhD project on Elastic Compute)

What this course is about

- ***Learn*** **how to design scalable and efficient cloud-native systems**
  - Understand the demands of **novel** (data) **workloads** and the **challenges** at **scale**
  - Get to know the **internals** of modern **data centers** and emerging technologies and trends
  - Learn the **fundamental principles** for building **scalable system** software

- ***Build*** **a cloud-native multi-tier data processing system:**
  - Work across multiple layers of the stack: storage, synchronization, caching, compute, etc.
  - Tailor the system for given workload requirements
  - **Think** in terms **of performance**, **scalability**, **fault tolerance**, **elasticity**, **high availability**, **cost**, **privacy**, etc.
  - Use modern cloud constructs like containers or serverless functions.

- ***Apply*** **the knowledge with hands-on work:**
  - Modular homework assignments
  - Project work

Motivation

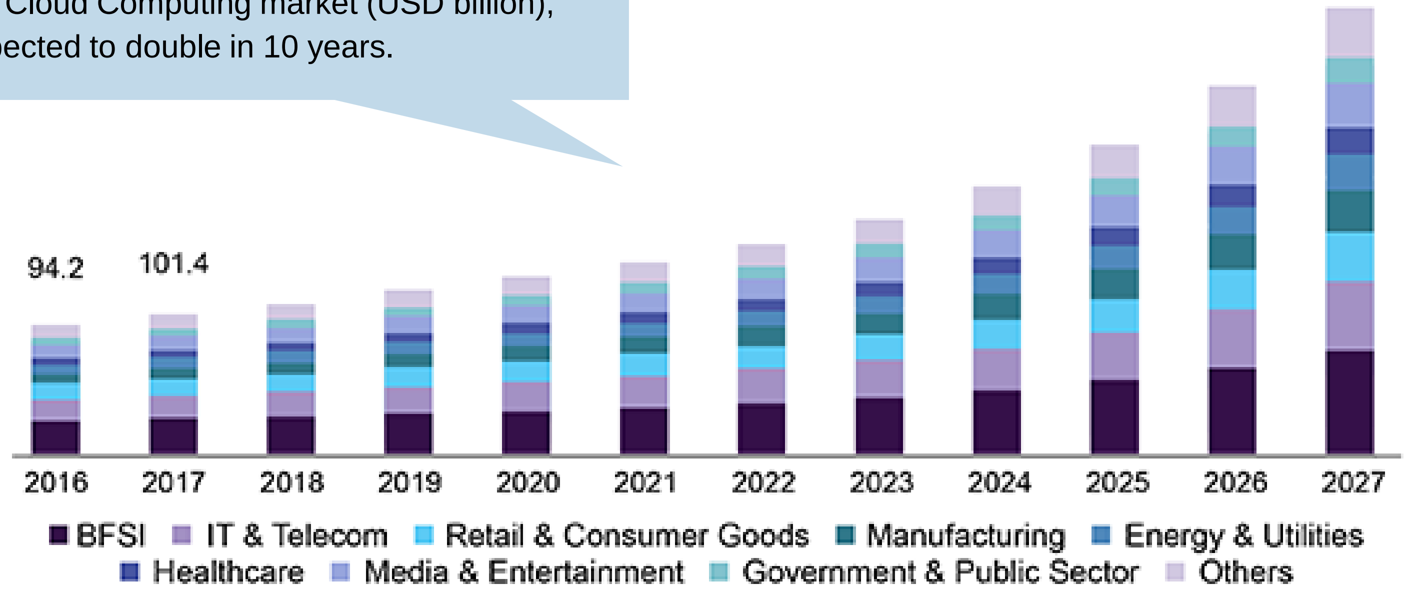
Motivation

- Why should we care about the cloud?
- What impact does the cloud have on system development?
- Why should we focus on data-processing in particular?

Why is Cloud important?

- The internet has around **4.5 billion users** today, and the number is still growing
- **Digitalization** of society and the **Cloud transform** whole **industries**
- **25%** increase in cloud usage during the pandemic (src: Gartner 2022)

US Cloud Computing market (USD billion), expected to double in 10 years.



https://www.grandviewresearch.com/industry-analysis/cloud-computing-industry

https://99firms.com/blog/google-search-statistics

How the Cloud impacts technology development?

- Cloud helps in fast dissemination of new technologies
- Easy, fast and cheap exposure to new trends available for everyone

- **Accelerators**

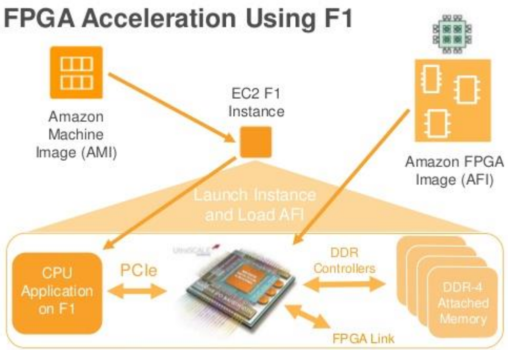
  EC2 offers instances with the latest GPUs, custom ML inference ASICs or FPGA, TPUs

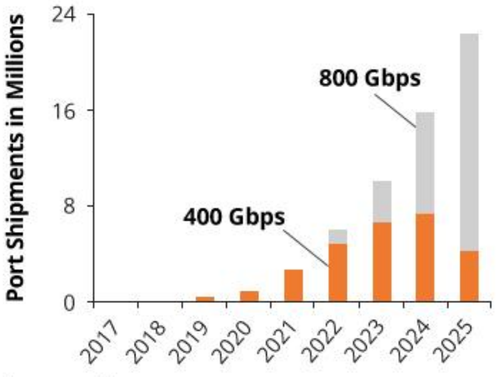
Fast network interconnects

M6in.metal already offers 128 cores, 512 GiB memory and 200Gbps network (cost varies)

Latest storage technologies

Microsoft's revolutionary glass storage with [Project Silica]() or Holographic storage (HSD)





*source: Dell'Oro Group

Cloud providers control the full stack

- **Influence the hardware landscape**
  – Innovation from novel chip design, to new switches and network fabrics, incl. storage technologies
- **Control the full software stack**
  – they can change or customize it (OS, virtualization, containers, etc.)
- **Introduce or popularize new programming methodologies and paradigms**
  – Map-Reduce, actor-based programming models, micro-services and serverless, etc.
- **Revolutionize how we approach application design and implementation**
  – Scale, elasticity, cost, privacy, etc.

How are things different at scale?

As reported by Google (slides from Jeff Dean) in 2010:

Focus is more on meeting the SLOs (service-level objectives) with respect to:

- Performance (latency)
- High availability
- Efficiency
- Elasticity

Most complexity is absorbed by the cloud system software infrastructure

The Joys of Real Hardware

Typical first year for a new cluster:

~1 network rewiring (rolling ~5% of machines down over 2-day span)

~20 rack failures (40-80 machines instantly disappear, 1-6 hours to get back)

~5 racks go wonky (40-80 machines see 50% packetloss)

~8 network maintenances (4 might cause ~30-minute random connectivity losses)

~12 router reloads (takes out DNS and external vips for a couple minutes)

~3 router failures (have to immediately pull traffic for an hour)

~dozens of minor 30-second blips for dns

~1000 individual machine failures

~thousands of hard drive failures

slow disks, bad memory, misconfigured machines, flaky machines, etc.

Long distance links: wild dogs, sharks, dead horses, drunken hunters, etc.

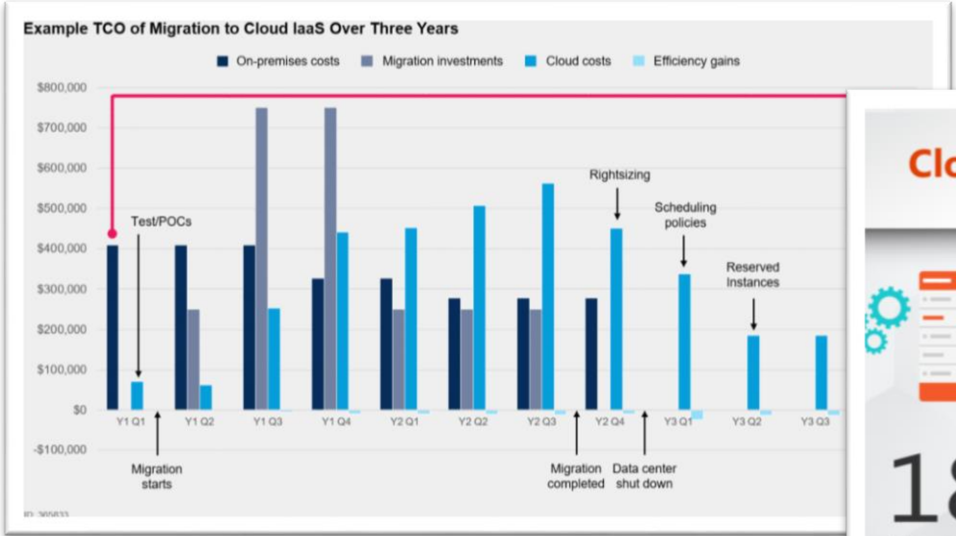
Reliability/availability must come from software!

Google

https://static.googleusercontent.com/media/research.google.com/en//people/jeff/Stanford-DL-Nov-2010.pdf

But it is not just scale!

- Incentives highly driven by reduction of cost
- Skeptics primarily worried about cloud's privacy and security.

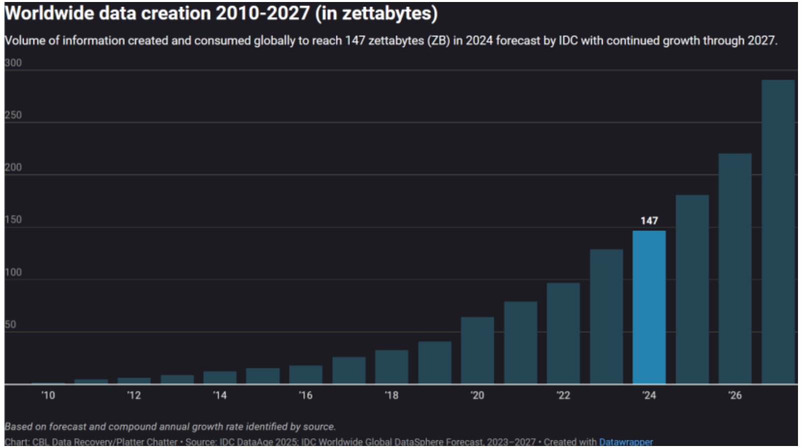


https://blogs.gartner.com/marco-meinardi/2018/11/30/public-cloud-cheaper-than-running-your-data-center/

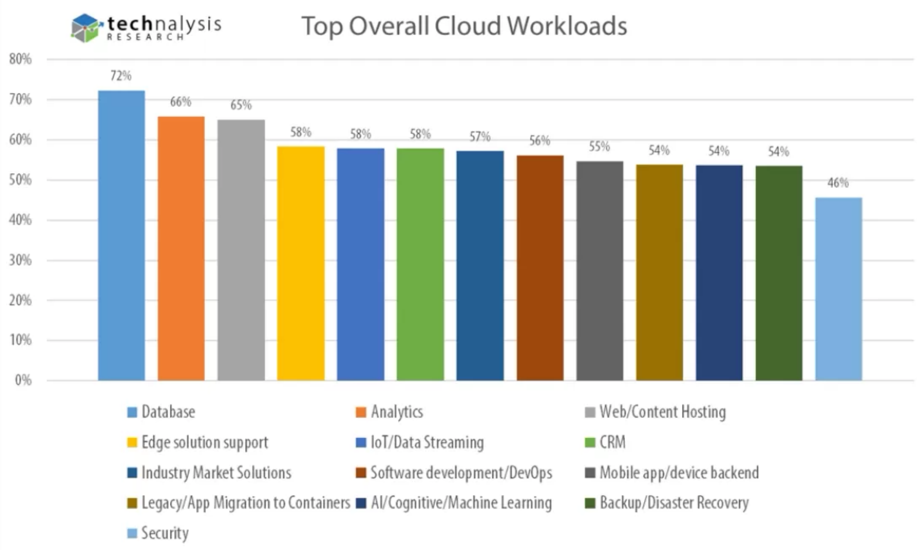
https://dzone.com/articles/data-security-an-integral-aspect-of-cloud-computin

Why focus on data-processing?

- Surge in data volumes produced and consumed
  - IDC now estimates 291 ZB by 2027
- Data-processing is still the dominant workload:
  - Databases, analytics, streaming, etc.



https://www.datawrapper.de/_/Huqfa/
Src: IDC DataAge 2025; Global DataSphere Forecase 2023-2027



https://www.techspot.com/news/83646-companies-taking-advantage-different-cloud-options-putting-different.html

Course administrivia

Course content

- **Data centers and cloud computing**
- **Distributed data basics (partitioning, replication, fault-tolerance, consistency, consensus)**
- **Design principles for cloud-based applications**
- **Design and build scalable systems for the cloud:**
  - Covering **storage, query, and transaction processing.**
- **Trends, emerging technologies and their impact on the future of cloud-systems**
  - Hardware and accelerators, resource disaggregation, software-defined networking/storage

Special focus on **state-of-the-art systems** that are **used in production**

Course Organization

Lecture:

- **In-person** lectures on **Thursdays 2-4pm (Galileo 8120.EG.001)**
  – **Slides** uploaded on course web-page and moodle (by Thursday noon).
  – Old lecture video recordings from WS 20/21 available on moodle.
- Course **website**: https://db.cit.tum.de/teaching/ws2425/clouddataprocessing/
- Please check regularly for updates

Tutorials:

- **In-person** tutorials after the lectures
- **Thursdays 4-5pm (Galileo 8120.EG.001)** – not recorded
- **TAs** for the course are
  Michalis Georgoulakis (michalis.georgoulakis@tum.de**)**
  Tobias Götz (goetzt@in.tum.de**)**
- First session: today for introduction, Q&A session and general set-up
- Consider that **exercise material** is **part** of the **course content**!

Assignments and Project

- The main goal of the course is **critical thinking** and analyzing the main **design decisions** behind **scalable systems** and understanding what it takes to build them.

- **The assignments will give you a range of different skillsets:**
  1. Analysis on different design decisions on how to build a data processing system in the cloud
  2. Measurement study on existing cloud services, system design and back-of-the-envelope calc.
  3. Hands-on implementation of a data processing task that uses the cloud services you benchmarked.

- You can then **apply them** for your **project** in the last 5 weeks of the course.

Assessment and Exam

- Bonus: extra points for the final exam

- Maximum bonus: 11 points
  - Homework assignments: up to 6 points
  - Project: up to 5 points

- Passing criteria:
  - Exam needs to be passed so the bonus points can be accounted for
  - For the homework assignments – details later in the tutorial session

- Written exam
  - 90min written (90 points)
  - No retake offered

Course Set-up

Let's make the course as interactive as possible

- During the lecture and tutorials, please speak-up, ask questions and discuss!
- Also engage in asynchronous discussions on Mattermost
- Send the TAs questions you want to be addressed during the tutorial sessions

The material we discuss is relevant in practice:

- We will provide examples
- You will achieve the maximum fun factor if you do the project work

- We will have a **few guest speakers** (also from **industry**)
  - Details to be announced later in class.

Course material

This is not a standard course – it is state of the art (bleeding edge) systems and research

- There is **no real textbook** for this course, but a good overview of the principles behind **building scalable systems** are covered in:
  - "Designing Data-Intensive Applications" by Martin Kleppmann
  - "Azure Application Architecture Guide" by Microsoft
  - "Architecting for the Cloud" by AWS [v.2024]
- More on **hardware-** and **software-virtualization** is covered in:
  - "Hardware and Software Support for Virtualization" by Ed Bougnon, Jason Nieh, and Dan Tsafrir.

- The **lecture slides** are available **online**
- Most **material** that we are going to cover is **taken out of research papers:**
  - The references to those papers (all good, easy and fun! to read) will be given as we go.
  - Relevant conferences: ACM/USENIX SOSP/OSDI, ACM SOCC, USENIX ATC, NSDI, ACM EuroSys, ACM SIGMOD, VLDB, ACM SIGCOMM, IEEE ICDE, ACM CoNEXT, etc.

Cloud-based application design

Challenges

Distributed Computing Challenges

Scalability

- Independent parallel processing of sub-requests or tasks
- E.g., adding more servers permits serving more concurrent requests

Fault Tolerance

- Must mask failures and recover from hardware and software failures
- Must replicate data and service for redundancy

High Availability

- Service must operate 24/7

Consistency

- Data stored / produced by multiple services must lead to consistent results

Performance

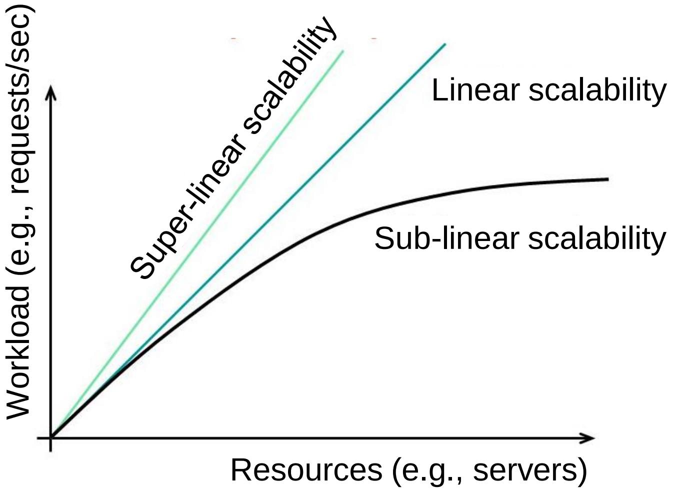
- Predictable low-latency processing with high throughput

Scalability matters

Ideally, adding N more servers should support N more users!

But, **linear scalability** is **hard** to achieve:

- Overheads + synchronization
- Load-imbalances create hot-spots
  (e.g., due to popular content, poor hash function)
- Amdahl's law → a straggler slows everything down



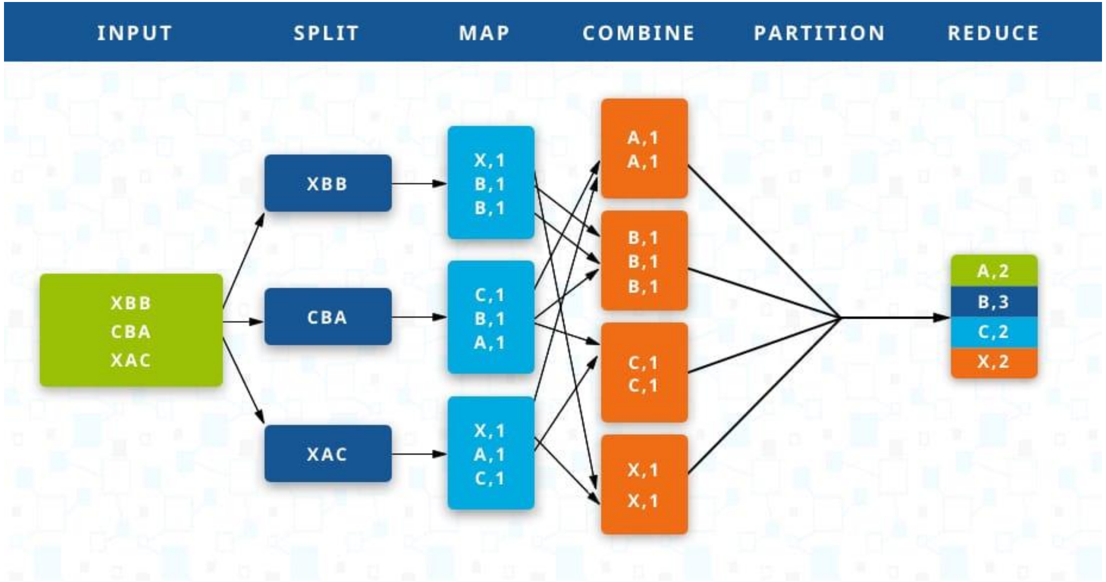
Therefore, one needs to **partition both data** and **compute.**

Scaling computation

How do data-intensive applications scale?

- Enable task-parallel or data-parallel processing
- Frontend does the aggregation of (select top-k documents)
- Back-ends provide partial responses

e.g., Map-Reduce



Img src: https://www.talend.com/de/resources/what-is-mapreduce/

Fault tolerance

- **Think of failure as the common case.**

- **Full redundancy** is too **expensive** → **use failure recovery**.
  - Impossible to build redundant systems at scale
  - Rather reduce the cost of failure recovery

- Failure recovery: **replication** or **re-computation**
  - Which one is better, depends on the respective costs

- **Replication**:
  - Need to replicate data and service
  - Introduces the consistency issues

- **Re-computation**
  - Easy for stateless services
  - Remember data lineage for compute jobs

High availability

- **Downtime** → **bad** customer **experience**, and **loss** in **revenue**.
- According to Gartner, a **minute** of IT **downtime** costs companies **$5’600 on average**.

Cloud service providers offer **service level agreements (SLAs)** to their clients.
A **commitment/contract** for the **quality** of the **service** (e.g., availability, performance, etc.)

Translating downtime for a typical SLA for availability:

- **99.9%** (“three nines”) availability means **8.77 hours downtime per year** → close to $3 million.
- **99.99%** (“four nines”) availability means **52.6 minutes downtime per year** → close to $300’000.

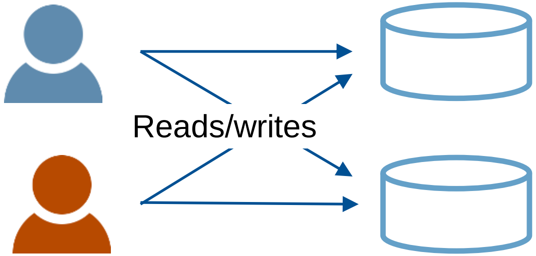
For a **high available** service one needs to design and:

- Eliminate **single point of failure** by adding redundancy in the system.
- Have a **reliable crossover.**
- Have an efficient way to **monitor** and **detect failures** when they occur.

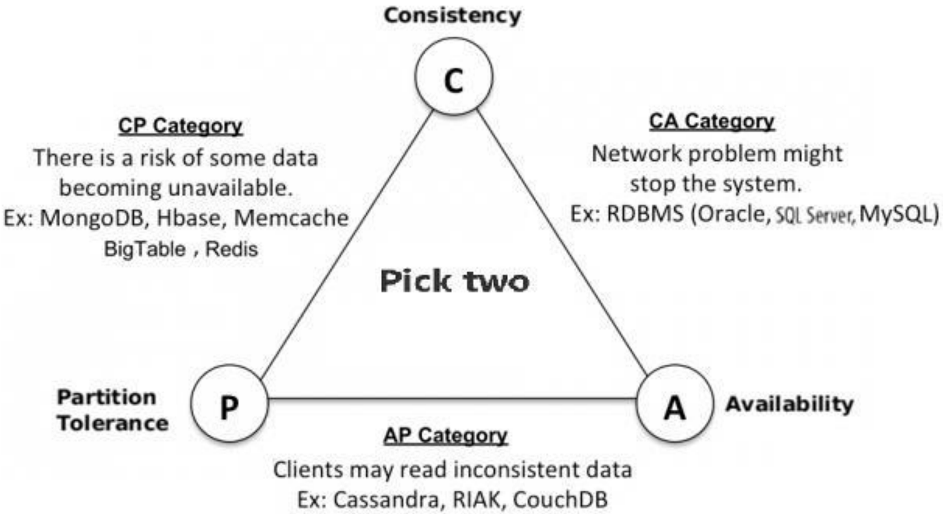
e.g., Amazon S3 offers 11 9s of availability of objects across multiple availability zones (AZs).

Consistency

- Many **applications** need **state replicated** across a wide area, for reliability, availability and low latency.



- **CAP Theorem:** *It is impossible for a distributed data store to simultaneously provide more than two out of the three guarantees*:
  - Consistency
  - Availability
  - Partition tolerance



Src: CAP Theorem by Eric Brewer, formally defined by Gilbert and Lynch

Consistency models

- **Two main choices:**
  - **Strongly consistent** operations (e.g., use Paxos, Raft, etc.)
    - Often at the cost of additional latency for the common case
  - **Inconsistent** operations
    - Better performance / availability, but applications are harder to write and reason about the model

- **Many applications aiming for high availability gravitated towards eventual consistency**
  - E.g., **Gmail**: marking a message as read is asynchronous, but sending a message needs to be a consistent operation
  - Order of posts in **LinkedIn** news feed? Access from multiple devices?
  - Count of song popularity in **Spotify**?

- But, modern data analytics (data lakes, training ML on PBs of data) require strong consistency
  https://www.allthingsdistributed.com/2021/04/s3-strong-consistency.html

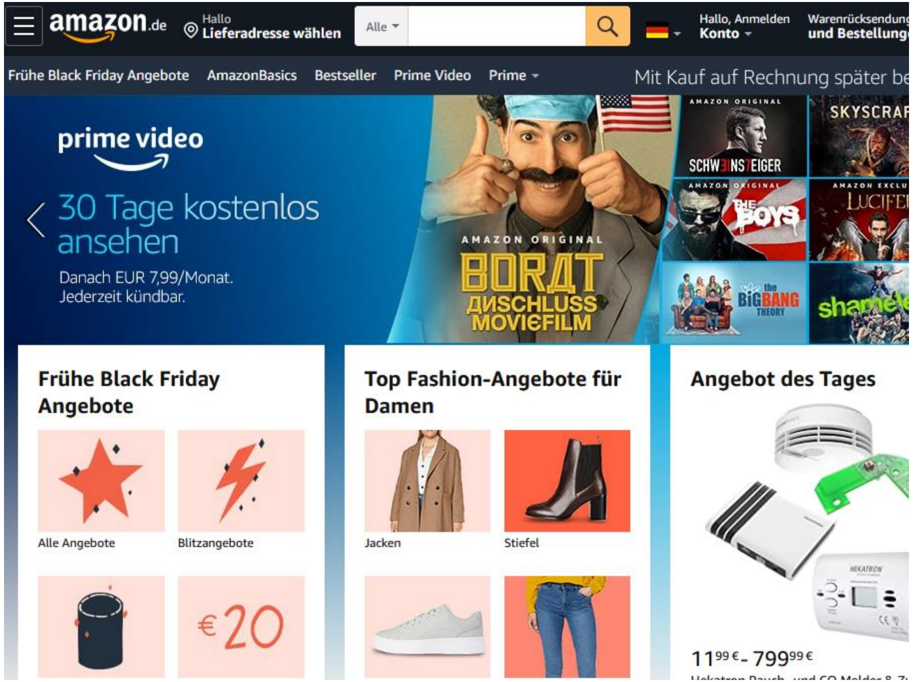
"Eventual Consistency Today: Limitations, Extensions, and Beyond" by Bailis and Ghodsi in ACM Queue vol. 11, issue 3, 2013

Performance matters

Online services (e.g., Facebook, Google search, Bing):

- Expected response time < 100ms

https://www.gigaspaces.com/blog/amazon-found-every-100ms-of-latency-cost-them-1-in-sales/



Performance affects revenue:

- Values reported 10 years ago
  - **Amazon**: every 100ms of latency costs them 1% in sales
  - **Google** found an extra 0.5 secs drops traffic by 20%
- Akamai in 2017 found that a 100ms delay in page load time results in 6% drop in sales
- Even more valid **today** in mobile web browsing/app responsiveness

The tail at scale

- At scale, looking at the average request latency is **not** enough.

- **Tail latency** = the last 0.X% of the request latency distribution graph.
  - e.g., we can take the slowest 1% response times or the 99%ile response time.

- Tail latency is **amplified** by **scale**, due to **fan-outs** for
  - **Micro-services**, **data partitions**

- Overall latency ≥ latency of the **slowest** component

- Servers with 1ms average, but 1sec 99%ile latency
  - 1 server: 1% of the requests take >= 1 sec
  - 100 servers: 63% of the requests take >= 1sec

"The Tail at Scale" by Jeffrey Dean and Luiz Andre Barroso in *Comm. Of the ACM, 2013*
Tail Latency Might Matter More than you Think – Marc Brooker blog.

The tail at scale

- Increased fan-out has a large impact on the latency distributions.
- At Google scale:
  - 10ms 99% percentile for any single request
  - The 99% percentile for all requests is 140ms and the 95% percentile is 70ms
    - Waiting for the slowest 5% of the requests accounts for half of the total 99% percentile latency.

Table 1. Individual-leaf-request finishing times for a large fan-out service tree (measured from root node of the tree).

	50%ile latency	95%ile latency	99%ile latency
One random leaf finishes	1ms	5ms	10ms
95% of all leaf requests finish	12ms	32ms	70ms
100% of all leaf requests finish	40ms	87ms	140ms

"The Tail at Scale" by Jeffrey Dean and Luiz Andre Barroso in *Comm. Of the ACM, 2013*

Distributed Computing Challenges (recap)

Scalability

- Being able to elastically scale (out and in) to meet the load demand is crucial.

Fault Tolerance

- Accept the reality that faults are common and build for quick detection and recovery.

High Availability

- Target multiple 9s availability to minimize costs for downtime.

Consistency

- Embracing **eventual consistency** for high availability is often preferred for many use-cases.

Performance

- Optimizing for tail latency is **important.**