



OST

Eastern Switzerland
University of Applied Sciences

Distributed Systems (DSy)

Introduction – Fault Tolerance

Thomas Bocek

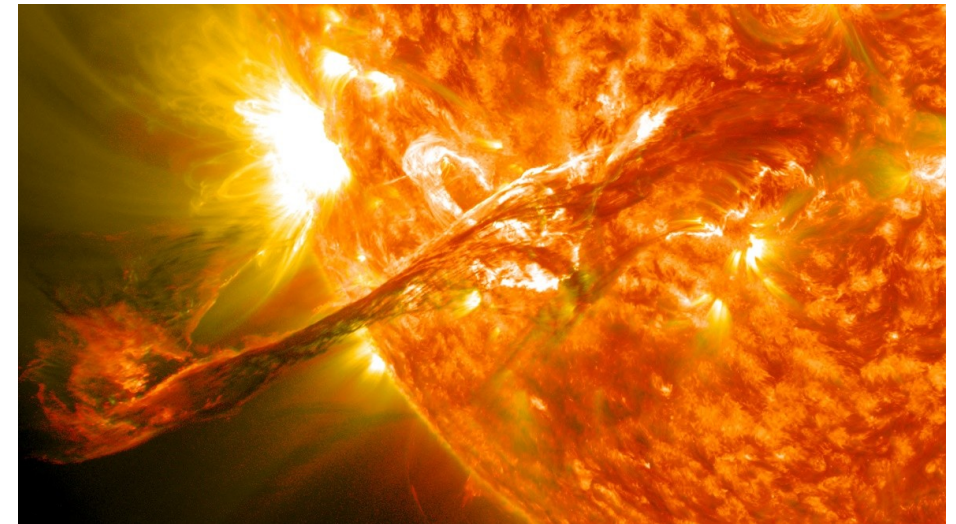
22.02.2025

Learning Goals

- Distributed systems add complexity. Avoid complexity!
- Why do we need distributed systems?
 - 1) Scaling (if one machine is not enough)
 - 2) Location (to move closer to the user)
 - 3) Fault-tolerance (HW will fail eventually)

Distributed Systems Motivation

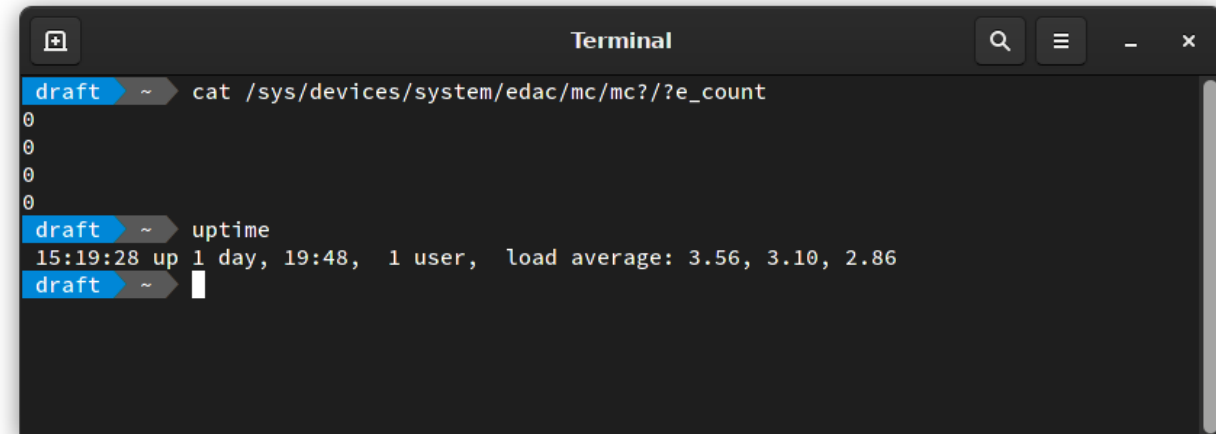
- Why Distributed Systems – Fault-tolerance
 - Any hardware will crash eventually
- Random bit flips in memory
 - 1990: “Computers typically experience about one cosmic-ray-induced error per 256 megabytes of RAM per month”
 - Google study 2009: more than 8% of DIMMs affected by errors per year
 - 2007: 44 reported memory errors (41 ECC and 3 double bit) on ~1300 nodes during a period of about 3 months
- Source: Bad Pin Connections, Incorrect RAM Timings, Clock Issues, RAM Design Flaws, CPU/RAM/Motherboard Integrated Logic Defects, DRAM Cell Amplification Errors, Cosmic Rays [link]
 - Cosmic rays
 - Solar flares, Coronal mass ejection, Solar proton events, Background radiation
- Cosmic rays “may” be blamed for an electronic voting error in Belgium (2003)
 - Bit flip in electronic voting machine
 - Added 4096 extra votes to one candidate
 - Candidate more votes than were possible



https://en.wikipedia.org/wiki/Solar_flare

Distributed Systems Motivation

- Influencing factors
 - Sensitivity of each transistor, number of transistors on the microchip, altitude, floor level
- Mars Rover?
 - Cassini reported 280 bitflip/day [link] – max 890 due to solar proton event - TMR with ~300MB RAM
 - Radiation-tolerant FPGAs → TMR
- Error-correcting code memory
 - Uses TMR or Hamming Code, correct 1 bitflip / detect 2 bitflips
 - Used for Servers, not (yet) used for consumer products – good idea?
- Double bit-flips unlikely?
 - Jaguar super computer with 360TB ECC RAM
 - Double bitflip → happened every 24h
- Check your HW
 - DDR5 has ECC built in
 - on-die vs. traditional ECC [link]



```
Terminal
draft ~ cat /sys/devices/system/edac/mc/mc0/e_count
0
0
0
0
0
draft ~ uptime
15:19:28 up 1 day, 19:48, 1 user, load average: 3.56, 3.10, 2.86
draft ~
```

- What can happen: e.g., expr segfaults

Distributed Systems Motivation

- HDD break [[link](#)], SSDs wear out
 - SSDs consist of NAND cells with a limited lifetime
 - An SSD disk has spare NAND that are used when cells break
 - `smartctl -a /dev/xyz`
- SLC, MLC, TLC, QLC
 - SLC: 10'000 – 100'000 write/erase cycles
 - MLC: 10'000 – write/erase cycles
 - TLC: 1'000 – write/erase cycles
 - QLC < 1'000 – write/erase cycles
- 100% → no spare used, My old laptop was at 92%
 - When value is down at 0% disk capacity degrades
- E.g., Samsung 4TB drive uses QLC [[link](#)]
 - Write 100 times the same 4kb file, and cells are broken?
- Wear leveling: distribute write and erase operations across all memory cells
 - If wear leveling goes wrong: Samsung 990 Pro [[link](#)]
- Caching with SLC → files / cells that are frequently changed, store on SLC, once they don't change that often move to MLC/TLC/QLC

```
== START OF SMART DATA SECTION ==
SMART overall-health self-assessment test result: PASSED

SMART/Health Information (NVMe Log 0x02)
Critical Warning:          0x00
Temperature:              49 Celsius
Available Spare:           100%
Available Spare Threshold: 10%
Percentage Used:           17%
Data Units Written:        764,591,703 [391 TB]
Host Read Commands:        7,756,128,132
Host Write Commands:       4,042,594,169
Controller Busy Time:      17,710
Power Cycles:              462
Power On Hours:            23,008
```

Distributed Systems Motivation

- Random bit flips in memory
 - Bitsquatting: DNS Hijacking without exploitation (2015)
 - Register names with single bit error, e.g,

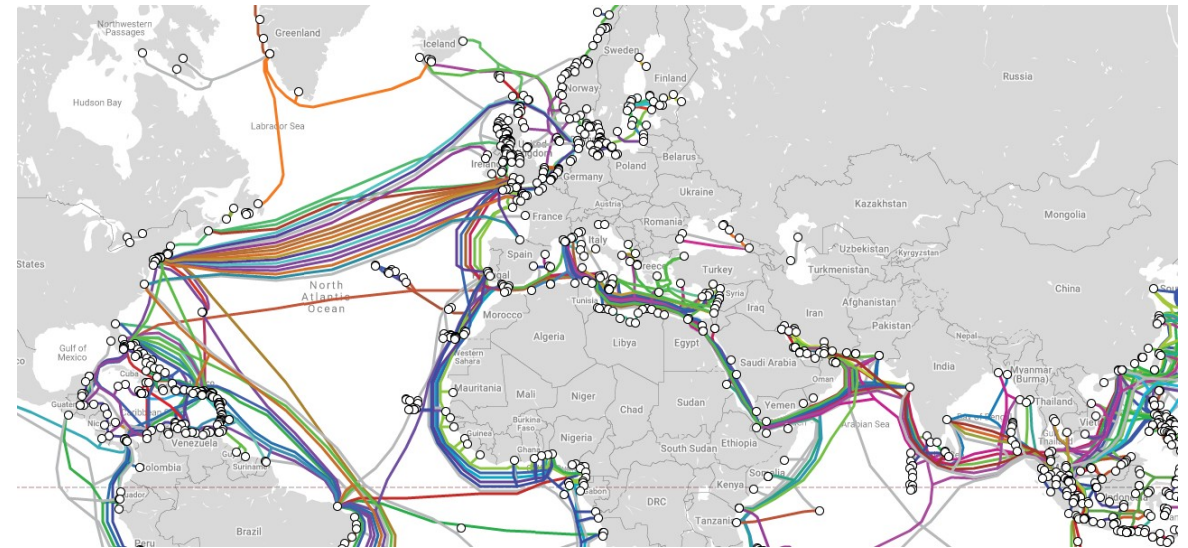
Bitsquat Domain	Original Domain
ikamai.net	akamai.net
aeazon.com	amazon.com
a-azon.com	amazon.com
amazgn.com	amazon.com
microsmft.com	microsoft.com
micrgsoft.com	microsoft.com

- Idea: if bitflip happens, it may happen for DNS names in your memory
 - Early tests by Artem Dinaburg: “59 unique IPs per day made HTTP requests to my 32 bitsquat domains”
- Key findings
 - Most users from China (more bitflips on Chinese machines?)

Fault Tolerance

- Network outages happens [[link](#)]
- 27.12.2024: Finland Detains Russia-Linked Vessel Over Damaged Undersea Power Cable in Baltic Sea [[link](#)]
- 06.01.2025: Taiwan Reportedly Claims China-Linked Ship Damaged One of Its Submarine Cables [[link](#)]
- 14.01.2025: When America's Allies Have Undersea Cables Severed, Suspicion Falls on Russia and China [[link](#)]
- 20.01.2025: Baltic Subsea Cable Damage Was Accidental, Not Sabotage – US and European Officials [[link](#)]

- [Submarine Cable Map](#)



Distributed Systems Motivation

- Hardware will fail eventually
 - Bit flips in memory can occur
 - SSDs will wear out over time
 - Network cables can be damaged
 - Not a question of "if" but "when"
- Distributed Systems provide a solution
 - Multiple machines provide redundancy
 - When one machine fails, others can take over
- Load can be redistributed among remaining machines
- System continues to function despite individual failures
- Trade-off consideration
 - Distributed systems add complexity
 - Use only when benefits outweigh the added complexity
 - Redundancy and fault tolerance must justify the complexity