science + computing

a Bull business

HPC Security

aus der Serie

Dinge mit Containern



Dr. Holger Gantikow

Bull science + computing

04.07.2026

Tübix 2026

Tübix › Tübinger Linuxtag

© Bull SAS

science + computing

a Bull business

0
whoami

1
HPC 101

2
Jupiter@Jülich

3
Securing HPC – supported by Containers

4
Road Ahead

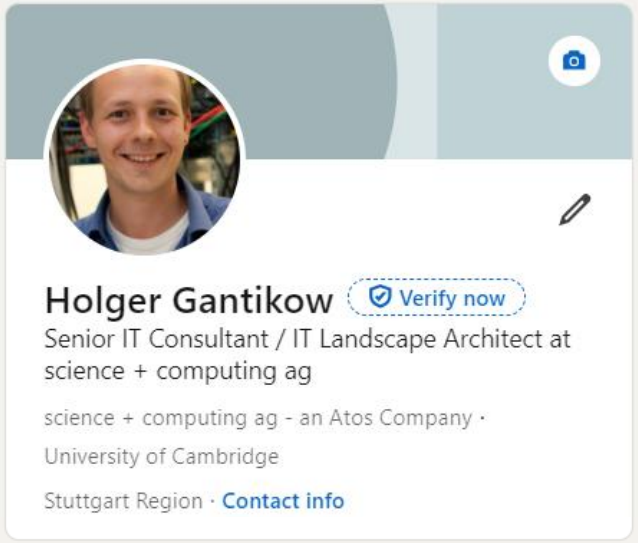
science + computing

a Bull business

0 whoami

About me

Linux, Container, Security, HPC, Cloud, Automotive. Part time Nerd.





Experience

science + computing ag - an Atos Company
15 yrs 8 mos

Senior IT Consultant / CAE IT Landscape Architect
Full-time
Jul 2022 - Present · 1 yr 11 mos
Baden-Württemberg, Germany

Senior Systems Engineer
Apr 2009 - Oct 2022 · 13 yrs 7 mos

System Engineer
2009 - Oct 2022 · 13 yrs 10 mos

Graduand
Oct 2008 - Mar 2009 · 6 mos

Diploma Thesis: "Virtualisierung im Kontext von Hochverfügbarkeit" - "Virtualization in the context of High Availability"

Intern
Karlsruhe Institute of Technology (KIT) / University of Karlsruhe (TH)
Aug 2008 - Sep 2008 · 2 mos

Research on optimization of computing workflow using Sun Grid Engine (SGE) for MCNPX calculations.

Education

University of Cambridge
GOLD for Technology Leaders, Department of Engineering, Institute for Manufacturing (IfM)
Sep 2022

Successfully completed the Atos Group Talent program "GOLD for Technology Leaders", with the Cambridge University's Institute for Manufacturing (IfM) and Paderborn University's Software Innovation Campus (SICP)

Paderborn University
GOLD for Technology Leaders, Software Innovation Campus (SICP)
Sep 2022

Successfully completed the Atos Group Talent program "GOLD for Technology Leaders", with the Cambridge University's Institute for Manufacturing (IfM) and Paderborn University's Software Innovation Campus (SICP)

Hochschule Furtwangen
Dipl. Inform. (FH), Coding, HPC, Clustering, Unix stuff :-)
2003 - 2009

science + computing - Quick Facts

Focus on technical & scientific computing with 35 years of expertise

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

founded in 1989
as a university spin-off

science + computing

dedicated unit for
high-end & business critical
IT-Services

FY 2022

€ 44M
external revenue

Atos
EVIDEN
Bull

an Atos / Eviden company
> 31.03.26 A Bull company


Tübingen


Berlin


Munich


Düsseldorf


Timișoara


Braşov

science + computing

a Bull business

offers **IT consulting and services** for the efficient use of complex compute environments in **research, development and simulation**.

Our customers include **manufacturers and suppliers** in the automotive and aerospace sectors, as well as **scientific research institutes**.

We help our customers **focus on their core business objectives** by providing an **efficient, cost-effective and resilient** IT landscape for product development and engineering.



Bull at a glance

Leading global technology player with strong European sovereign capabilities to help you master your AI destiny

- High-performance computing, AI & quantum
- Enterprise AI
- Enterprise computing and network
- End-to-end strategic services and product-related services
- R&D (Europe, India, Americas)
- Factory of the future (France)

3 000 engineers and experts

32 countries of operation

13% of revenue invested in R&D

1,600+ patents

~€720M FY25 revenue

#1 HPC player in Europe, India and South America (2025)

#1, #2 and #3 supercomputers in the Green500 (2025)

#2 in Europe with 140+ quantum patents, pioneering since 2016

300+ data scientists

250+ engineers in industrial HPC

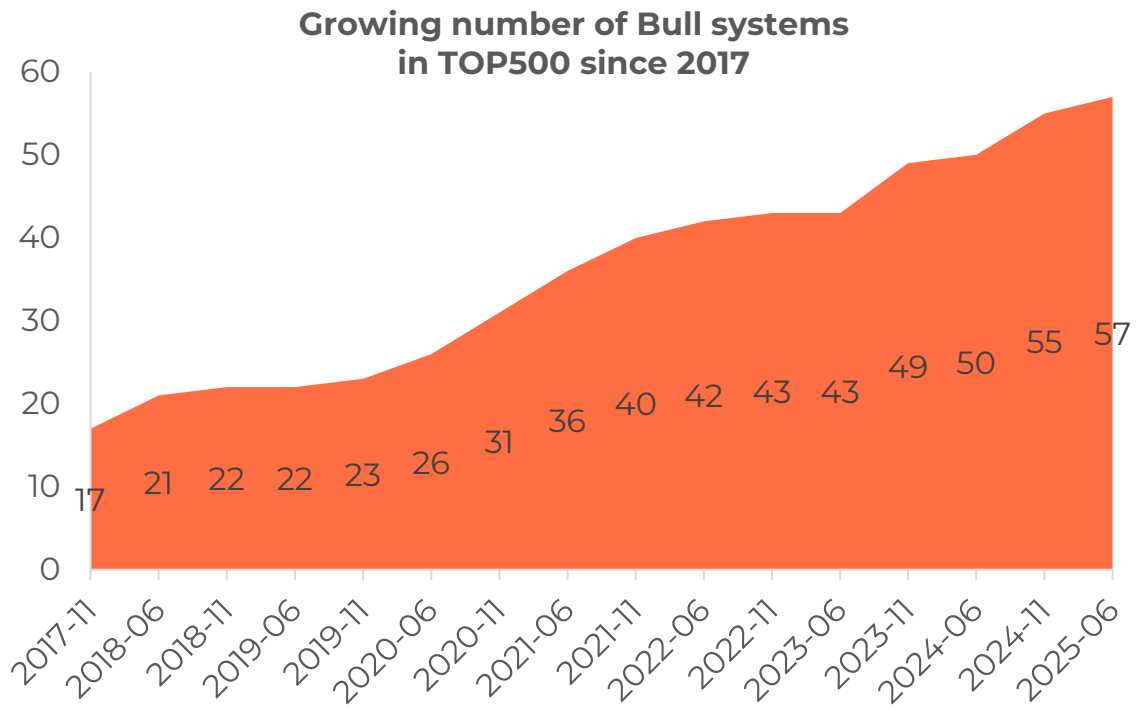
10+ world-breaking results for enterprise servers SAP benchmarks

Long Term Shareholder securing long term competitiveness and intensive investment to be the leading high-performance computing and AI European technology provider

science + computing
a Bull business

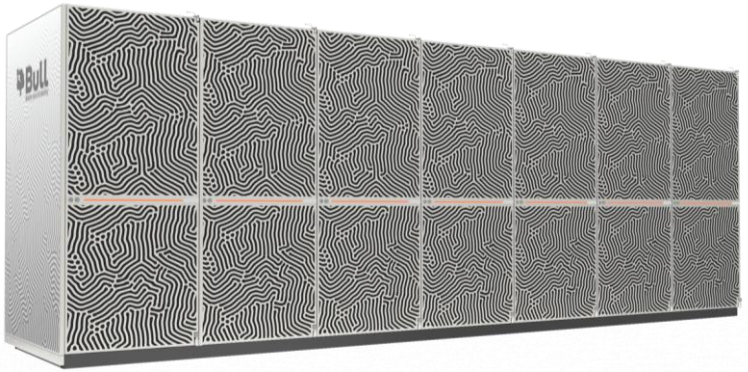
TOP500 – June 2026 – 65th list

History



- 55 as “Bull” manufacturer
- +1 as “ParTec/Bull” manufacturer (JEDI)
- +1 as “NVIDIA” (DCAI Gefion)

- **Bull is #3** in number of systems in TOP500 (WW)
- **Bull is #1** in number of systems in Europe (49)
- **Bull is #1** in number of systems in South America (5)
- **Bull is #1** in number of systems in India (3)
- **3 systems** in TOP20
- **Bull is #1, #2 and #3 in Green500 !!**
  - 5% more energy efficient than competition !



science + computing
a Bull business

science + computing

a Bull business

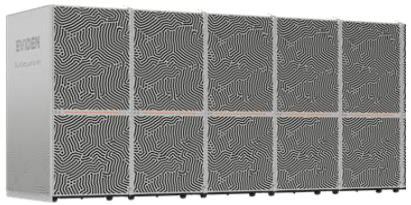
1 HPC 101

HPC = High Performance Computing

What do I need this for?

Well, talking about cars... Size of your problem?





Supercomputer, “HPC”





Gaming PC / Workstation





Laptop





Smartphone / Tablet

Thanks to Suryanarayanan Natarajan for the inspiration ☺

Central Components

Fitted to the max, supporting various use cases

- Processors
  - AMD, Intel, nvidia, Sipearl, …
- Accelerators
  - GPUs, FPGAs, TPUs
- Storage
  - Fast shared storage
  - Parallel File Systems (Lustre, GPFS, …)

Hardware
Compute + Storage

Network
“Interconnect”

- Low Latency, very high Bandwidth
  - Infiniband
  - RoCE, Omnipath, BXI, …

- Cooling
  - Airbased, Direct Liquid Cooling
- Power Supply
- Location, Weight Support of Floor

Infrastructure
(Facility, Cooling)

Software
OS
Workload Manager, Frameworks for parallelisation

- Linux
- Slurm, LSF, PBS, *Grid*, …
- MPI (n Nodes), OpenMP (1 Node)

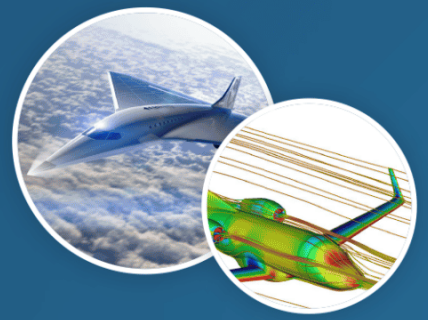
Ausgewählte Beispiele

Aufstrebend

COMPUTATIONAL SCIENCE & ENGINEERING

Building Blocks

Aerospace

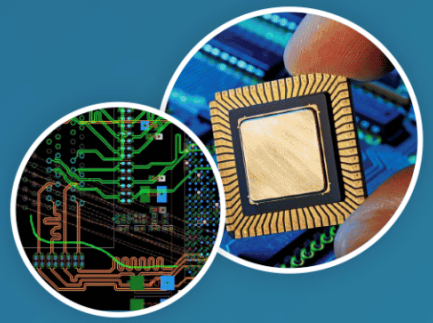


Propulsion

Aerodynamics

Acoustics

Semiconductor & EDA

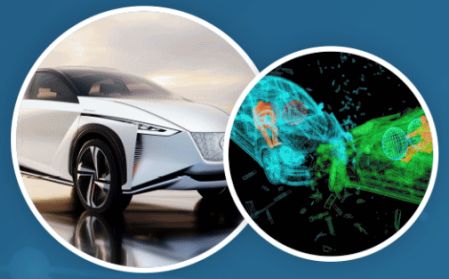


Design verification

Chip Layout

Classification

Automotive



Crash safety

Advanced Driver Assistance Systems

Combustion

Earth Science & Energy

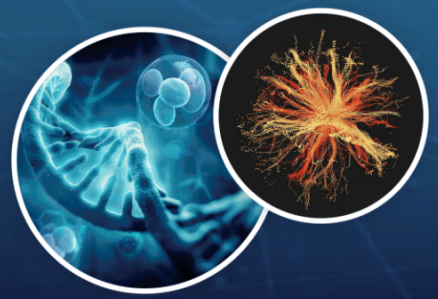


Climate & Weather Modeling

Drill & Pipeline Safety

Seismic & Reservoir

Life Sciences



Medical Devices

Metabolomics

Drug discovery

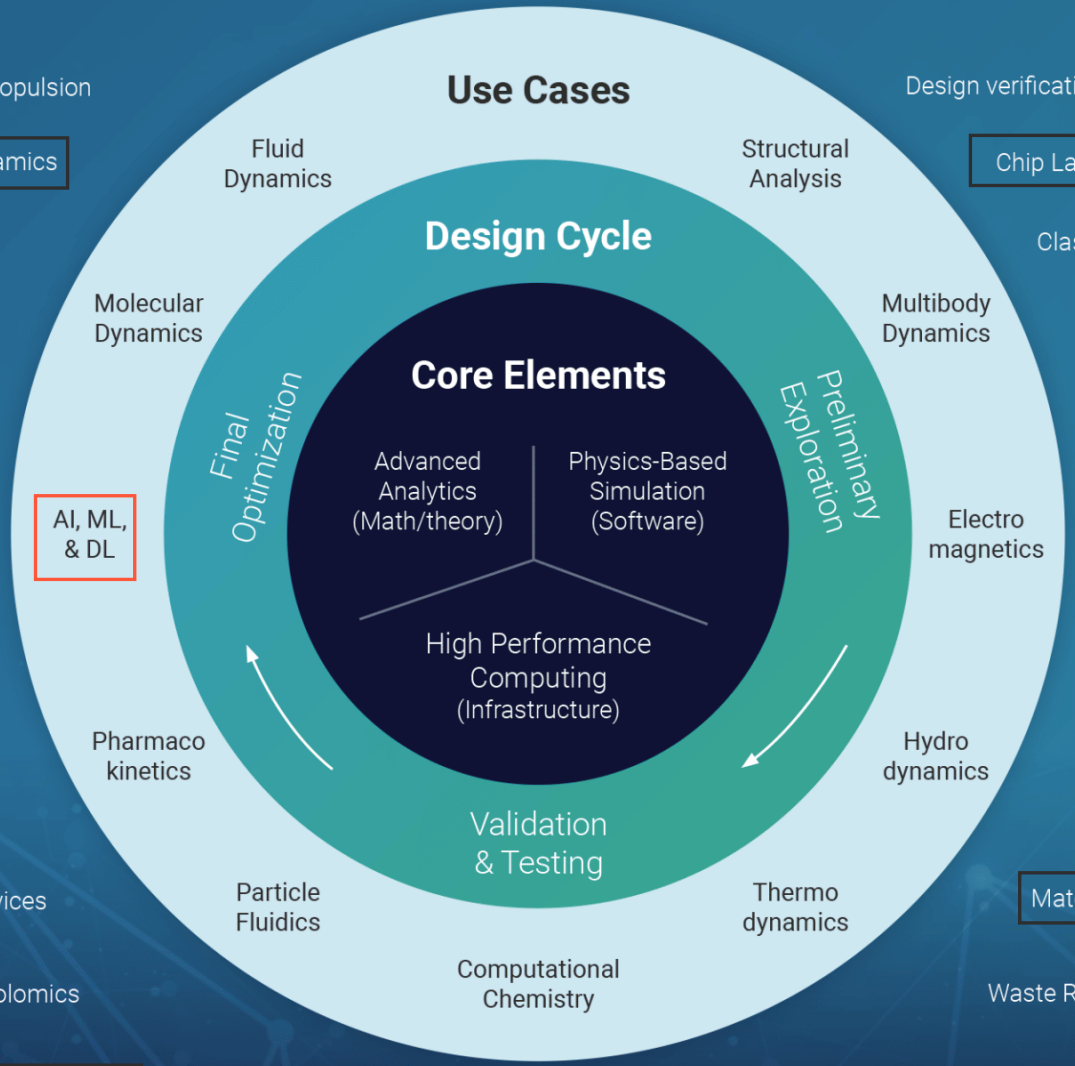
Manufacturing



Materials Testing

Waste Reduction

Factory Optimization



rescale

Source: https://rescale.com/cloud-hpc/

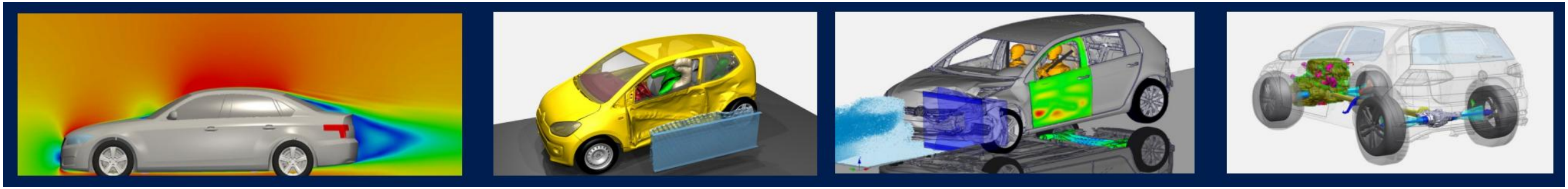
Computer Aided Engineering (CAE)

Driving digital product development – not only of cars

CAE is a **key consumer of Enterprise HPC** resources and leverages these for a broad range of engineering tasks

Such environments **differ from typical research lab**:

- **Very limited** User-Defined Software Stacks (**UDSS**) / Bring Your Own * (BYO) – changing with AI/ML workloads
  - Instead: High share of Independent Software Vendor **(ISV) code** – with a variety of lifecycle challenges
    - **Wide variety** of applications and versions
    - Need to **be operational over a long period of time** (compliance, product liability topics, OS releases, …)
  - **Maintained by Ops** team of the environment, deployed on site-/env-wide installation shares
- Actual simulation is part of a usually **complex workflows**
  - **Complexity** of this typically is hidden from the end user, globally defined, beyond the influence of the individual engineer



Source: Gantikow, Mayer: Containers in Automotive - Journey and Road Ahead, HPCW@ISC High Performance 2024

science + computing

a Bull business



Source: JUPITER: Managing massive AI and high performance computing capabilities in the exascale era, Jean-Olivier Gerphagnon, Red Hat Summit 2024

2 By Example: Jupiter@Forschungszentrum Jülich

Exascale?

{Kilo,Mega,Giga,Tera,Peta,Exa,Zetta,Yotta}-FLOP – 10^3, 10^6, 10^9, 10^12, 10^15, 10^18, 10^21, 10^24 Floating Point Operations per Second

The arrival of Exascale in Europe

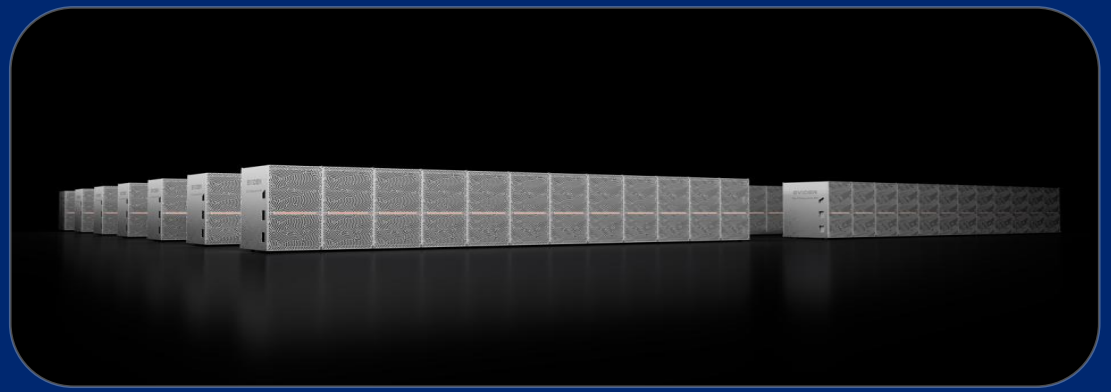
JUPITER



Exascale - 1 Quintillion calculations per second

1,000, 000, 000, 000, 000, 000 calculation/sec (10^{18})

If every person on the planet did ***1*** calculation/sec, it would take the ***4 years*** (with no eating or sleeping) to do the same amount of work as JUPITER can do in just ***1*** second!



JUPITER ascending

JUPITER

Joint **U**ndertaking **P**ioneer for **I**nnovative and **T**ransformative **E**xascale **R**esearch

BullSequana XH3000

EVIDEN



Booster
GraceHopper SuperChip

NVIDIA



General
RHEA1 CPU

SiPEARL



Network
InfiniBand NDR

NVIDIA



AI – 93 ExaFLOPs (FP8)
HPC – 1.0 ExaFLOPs (FP64)



24,000 NVIDIA SuperChips
(GraceHopper200)

1.2PB/s Ag. Bandwidth

29PB + 300PB Raw storage
(ExaFlash + ExaStore[1])

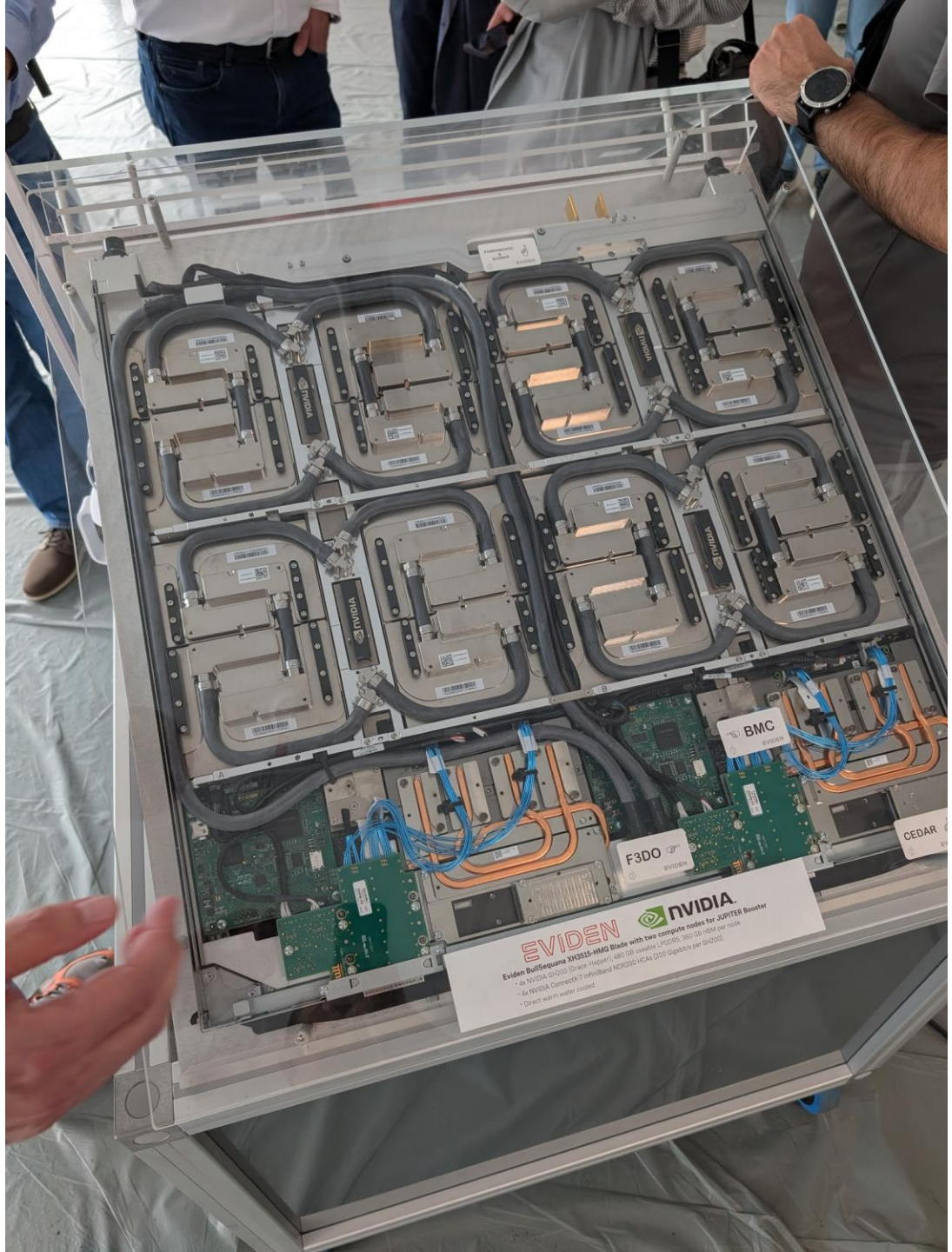
~18 MW Peak Power

[1] ExaStore is a separated project but will be used by JUPITER system

Hardware

Blades

For earch 1U 2x Compute Nodes
For each node 4x GH200,
2x Nvidia Quantum-2 IB,
480 GB LPDDR5
360 GB HBM



JUPITER – Software Environment I



	Technology	Challenges	Provider
Operating System	Linux	Security, Performance, Stability, HW support	
Management Storage	Ceph	Multi-use, Performance, Scalable	IBM
Management Plane	Kubernetes	Scalable, 0 downtime, Flexible, Open	kubernetes
Configuration Management	Ansible	Standard, Easy to extend, Open	
Boot Image(s) Management	ImageBuilder	ARM / x86 support, Tracking, Integration	
Container(s)	UBI Universal Binary Images	Standard, Security, Consistency	

JUPITER – Software Environment II



	Technology	Challenges	Provider
Resource Manager	Slurm	Scalable, Known, API	SchedMD The Slurm Company; ParaStation PSSLURM
Parallel Storage	Storage Scale System (GPFS)	Performance, Scalable, Data security	IBM
MPI Runtime	Message Passing Interface	Stable, GPU-support, Performance, Bug-free	ParaStation MPI; EVIDEN Open MPI
GPU Support	CUDA, HPC SDK	Memory management, Performance, Integration	NVIDIA; Red Hat
Monitoring & Logging	Prometheus + Thanos, Syslog + Fluentd	Usable, Scalable, Handle data storm	Prometheus; Thanos; fluentd
Reference Database	Data Center Information Management	Automation, API, Coherent	netbox

Jupiter ascending



From

JUPITER

Joint **U**ndertaking **P**ioneer for **I**nnovative and **T**ransformative **E**xascale **R**esearch

to



Jupiter ascending

JUPITER

Joint **U**ndertaking **P**ioneer for **I**nnovative and **T**ransformative **E**xascale **R**esearch



From

to



Jupiter ascending

JUPITER

Joint **U**ndertaking **P**ioneer for **I**nnovative and **T**ransformative **E**xascale **R**esearch



Jupiter ascending

Modular Approach

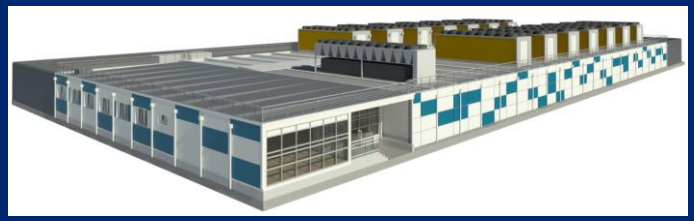
JUPITER

Joint **U**ndertaking **P**ioneer for **I**nnovative and **T**ransformative **E**xascale **R**esearch









Jülich Modular Data Center (MDC)

Jupiter ascending

Modular Approach

EVIDEN

Direct Liquid Cooling

Direct liquid cooling with up to 40°C warm water cooling for improved energy efficiency.

Heat Reuse

Potential for reusing heat generated in the MDC within Jülich Research Center.

End of life recycle

MDC is easier for materials recycling, contributing to sustainability efforts.

JUPITER

MODULAR DATA CENTER





Jupiter ascending

Modular Approach

JUPITER

MODULAR DATA CENTER





Quick facts:

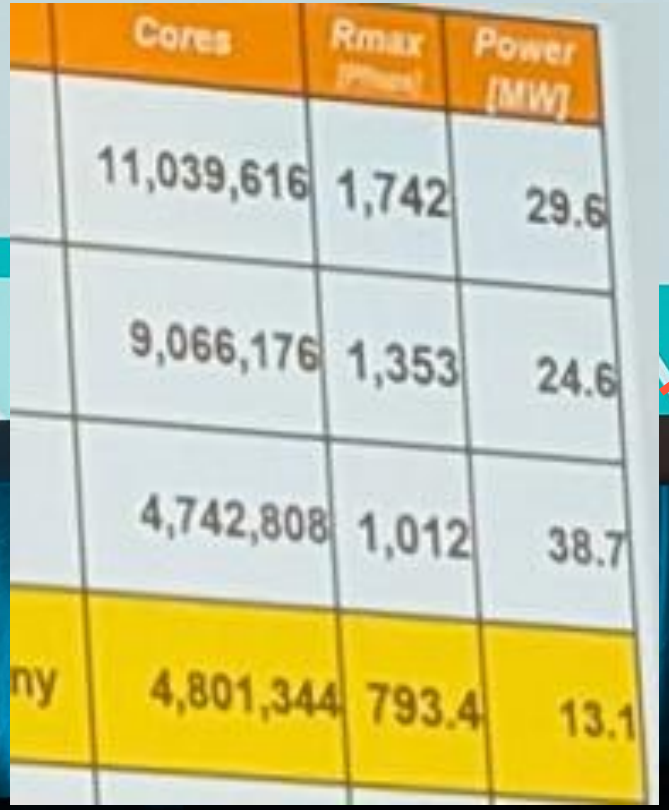
> **2,300 m²** data center

> **50** pre-built and interchangeable modules

Up to **30 MW** power capacity



#	Site	Manufacturer	TOP10 Computer of the TOP500	Country	Cores	Rmax [Pflops]	Power [MW]
1	Lawrence Livermore National Laboratory	HPE	El Capitan HPE Cray EX255a, AMD EPYC 24C 1.8GHz, Instinct MI300A, Slingshot-1	USA	11,039,616	1,742	29.6
2	Oak Ridge National Laboratory	HPE	Frontier HPE Cray EX235a, AMD EPYC 64C 2.0GHz, Instinct MI250X, Slingshot-1	USA	9,066,176	1,353	24.6
3	Argonne National Laboratory	Intel	Aurora HPE Cray EX/Intel Exascale Compute Blad Xeon Max 9470, Data Center GPU Max, Slingshot-1	USA	4,742,808	1,012	38.7
4	EuroHPC / FZJ	EVIDEN	JUPITER Booste BullSequana XH3000, NVIDIA GH200 Superchip, InfiniBand NDR	Germany	4,801,344	793.4	13.1
5	Microsoft Azure	Microsoft	Eagle Microsoft NDv5, Xeon Platinum 8480C, NVIDIA H100, Infiniband NDR	USA	1,123,200	561.2	
6	Eni S.p.A. Center for Computational Science	HPE	HPC6 HPE Cray EX235a, AMD EPYC 64C 2.0GHz, Instinct MI250X, Slingshot-11	Italy	3,143,520	477.9	8.5
7	RIKEN Center for Computational Science	Fujitsu	Fugaku Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D	Japan	7,630,848	442.0	29.9
8	Swiss National Supercomputing Centre (CSCS)	HPE	Alps HPE Cray EX254n, NVIDIA Grace 72C 3.1GHz, GH200, Slingshot-11	Switzerland	2,121,600	434.9	7.1
9	EuroHPC / CSC	HPE	LUMI HPE Cray EX235a, AMD EPYC 64C 2.0GHz, Instinct MI250X, Slingshot-11	Finland	2,752,704	379.7	7.1
10	EuroHPC / CINECA	EVIDEN	Leonardo Atos BullSequana XH2000, Xeon 32C 2.6GHz, NVIDIA A100, HDR Infiniband	Italy	1,824,768	241.2	7.5



	Cores	Rmax [Pflops]	Power [MW]
	11,039,616	1,742	29.6
	9,066,176	1,353	24.6
	4,742,808	1,012	38.7
ny	4,801,344	793.4	13.1

NG THE DOTS ↘



Use Cases on Jupiter

By example:

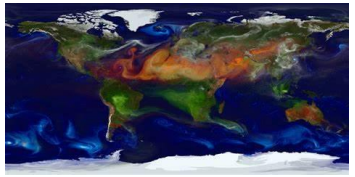


Advanced Climate science

- Creating more complex climate and weather models
- Increasing resolution

Leading to:

- The ability to better predict severe weather events with more time and precision
- Better earth system models



Large Language Models

- Training for Generative AI, including LLMs
- Use of AI in HPC Simulations

Leading to:

- More accurate Generative AI tools
- Model trained on European data sets



Digital Twins

- Digital Twins for Manufacturing
- Human brain project

Leading to:

- Improved manufacturing techniques
- Innovative materials
- Personalized and targeted medicines

Use Cases on Jupiter

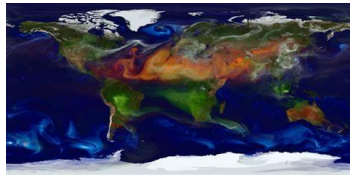
By example:

Advanced Climate science

- Creating more complex climate and weather models
- Increasing resolution

Leading to:

- The ability to better predict severe weather events with more time and precision
- Better earth system models



1. AI as an HPC workload
2. AI’surrogates to accelerate simulations
3. AI to enhance overall efficiency
   (predictive maintenance, scheduling, etc.)

- More ac… Generative AI tools
- Model t…d on Europ… sets

Digital Twins

- Digital Twins for Manufacturing
- Human brain project

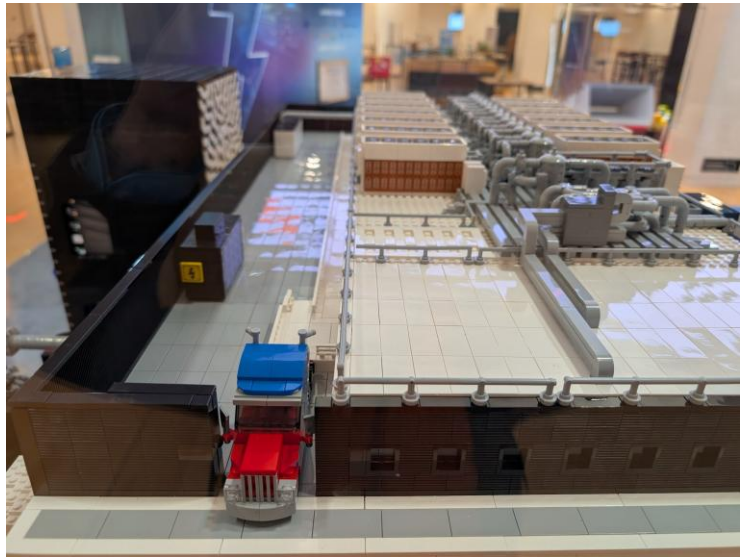
Leading to:

- Improved manufacturing techniques
- Innovative materials
- Personalized and targeted medicines



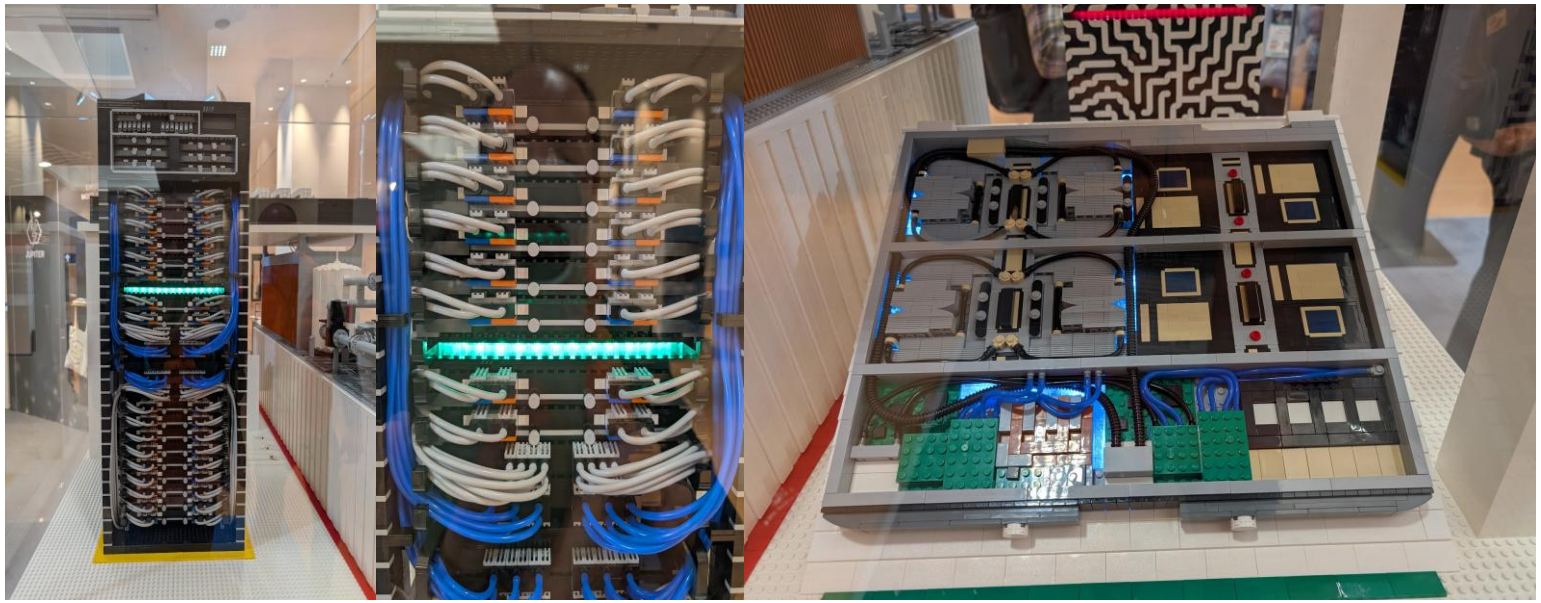
Lego goes Exascale

ISC2025, Jülich Booth









science + computing
a Bull business

June 2026: One year later…

Only the decorative flash is still missing ;-)



TOP 500 & Green 500

June 26 – 65th list

Rank	System	Cores	Rmax (PFlop/s)	Rpeak (PFlop/s)	Power (kW)
1	**LineShine** - LingKun, LX2 304C 1.55GHz, LingQi, Kylin OS, Shenzhen Cloud Computing Center Co., Ltd. National Supercomputing Centre in Shenzhen (NSCS) China	13,789,440	2,198.40	2,735.82	42,220
2	**El Capitan** - HPE Cray EX255a, AMD 4th Gen EPYC 24C 1.8GHz, AMD Instinct MI300A, Slingshot-11, TOSS, HPE DOE/NNSA/LLNL United States	11,340,000	1,809.00	2,821.10	29,685
3	**Frontier** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE Cray OS, HPE DOE/SC/Oak Ridge National Laboratory United States	9,066,176	1,353.00	2,055.72	24,607
4	**Aurora** - HPE Cray EX - Intel Exascale Compute Blade, Xeon CPU Max 9470 52C 2.4GHz, Intel Data Center GPU Max, Slingshot-11, Intel DOE/SC/Argonne National Laboratory United States	9,264,128	1,012.00	1,980.01	38,698
5	**JUPITER Booster** - BullSequana XH3000, GH Superchip 72C 3GHz, NVIDIA GH200 Superchip, Quad-Rail NVIDIA InfiniBand NDR200, RedHat Enterprise Linux, Bull EuroHPC/FZJ Germany	4,801,344	1,000.00	1,226.28	15,794



Green500 Data

Rank	TOP500 Rank	System	Cores	Rmax (PFlop/s)	Power (kW)	Energy Efficiency (GFlops/watts)
1	445	**KAIROS** - BullSequana XH3000, GH Superchip 72C 3GHz, NVIDIA GH200 Superchip, Quad-Rail NVIDIA InfiniBand NDR200, RedHat Enterprise Linux, Bull CALMIP / University of Toulouse - CNRS France	13,056	3.05	46	73.282
2	192	**ROMEO-2025** - BullSequana XH3000, Grace Hopper Superchip 72C 3GHz, NVIDIA GH200 Superchip, Quad-Rail NVIDIA InfiniBand NDR200, Red Hat Enterprise Linux, Bull ROMEO HPC Center - Champagne-Ardenne France	47,328	9.86	160	70.912
3	250	**Levante GPU extension** - BullSequana XH3000, GH Superchip 72C 3GHz, NVIDIA GH200 Superchip, Quad-Rail NVIDIA InfiniBand NDR200, RedHat Enterprise Linux, Bull DKRZ - Deutsches Klimarechenzentrum Germany	35,904	6.75	110	69.426

Additional Content on Jupiter

Quite entertaining and informative!



JUPITER – Europas schnellster Supercomputer erklärt

Forschungszentrum Jülich
3590 Abonnenten
Abonnieren
136
Teilen

5.773 Aufrufe 10.06.2025
JUPITER am Forschungszentrum Jülich ist der erste Supercomputer in Europa, der mehr als 1 ExaFLOP/s erreichen soll. Das entspricht einer Trillion Berechnungen pro Sekunde – eine „1" mit 18 Nullen. Für diesen Rechner der Spitzenklasse wurde am Jülich Supercomputing Centre eigens ein neuartiges, innovatives Rechenzentrum errichtet. Ausgestattet mit 24.000 Hochleistungs-Chips von NVIDIA belegt JUPITER aktuell den 4. Platz auf der TOP500-Liste der schnellsten Supercomputer – und zählt zu den leistungsfähigsten Systemen für KI überhaupt. Mit seiner enormen Rechenkraft soll JUPITER wissenschaftliche Simulationen auf ein neues Level heben und bahnbrechende Fortschritte beim Einsatz von KI ermöglichen.

https://www.youtube.com/watch?v=vVUFIrY-_xg



Europas schnellster Computer in Betrieb

tagesschau
1,97 Mio. Abonnenten
Abonnieren
2380
Teilen

ARD ist Teil des deutschen öffentlich-rechtlichen Rundfunks. Wikipedia

97.083 Aufrufe vor 1 Monat #nachrichten #tagesschau #supercomputer
Der schnellste Computer Europas ist im Forschungszentrum Jülich in Nordrhein-Westfalen in Betrieb gegangen. Er kann gigantische Datenmengen verarbeiten, etwa zur Berechnung von Klimamodellen oder zum Füttern von Anwendungen künstlicher Intelligenz. Gigantisch ist auch sein Stromverbrauch, obwohl der Rechner schon bes ...mehr

https://www.youtube.com/watch?v=zTxWXUvEx7Q



Jupiter: Europas KI auf dem Vormarsch | nano

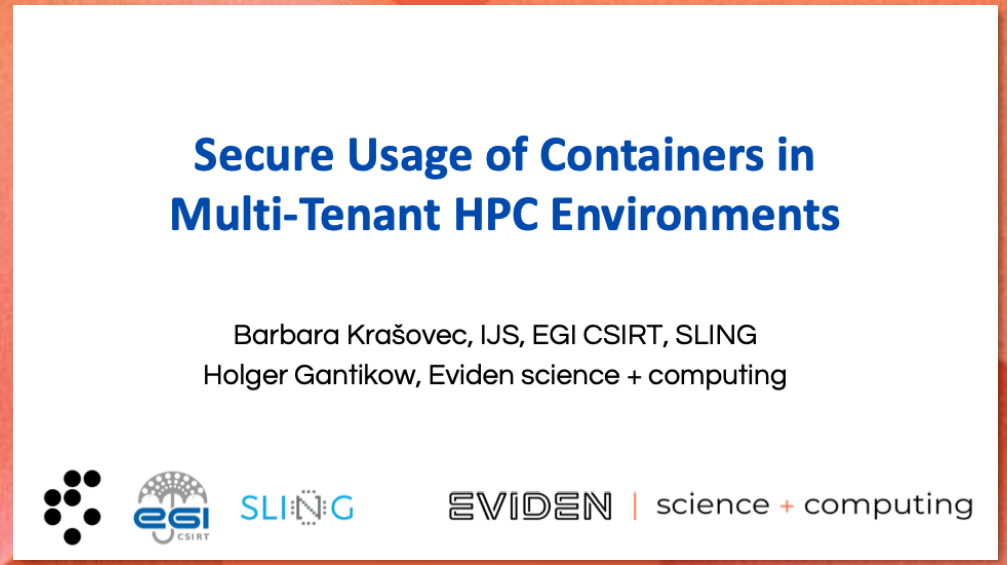
ZDFheute Nachrichten
1,68 Mio. Abonnenten
Abonnieren
2179
Teilen

ZDF ist Teil des deutschen öffentlich-rechtlichen Rundfunks. Wikipedia

114.099 Aufrufe vor 3 Monaten #IT #AI #KI
Wissenschaftler, Unternehmer und Politik sind sich einig: es muss eine europäische KI her. Im Forschungszentrum Jülich, das von Bund und Land NRW betrieben wird, steht Europas schnellster Supercomputer. Die KI-Fabrik soll den Brückenschlag zwischen Forschung und Entwicklung und der Anwendung in Unternehmen erleichtern und bes ...mehr

https://www.youtube.com/watch?v=sy0EUiNu9Xo

EVIDEN



Based on joint talk @ISC HPC 2025 – High Performance Container Workshop

3 Securing HPC – supported by Containers

The S in HPC stands for Security

Status Quo at many sites

- Users on the system are trusted
- Users have direct SSH access to the HPC environment / nodes

- Login nodes are protected and monitored, as they are accessible from the external network
- Zero trust is not implemented
- Defense-in-depth principle implementation is rare

- Clusters are heterogeneous and (growing need to) accept users from different disciplines, organizations, countries...

- Comparable situation at many industry users
  - (VPNs, 2FA, but risk of insider attacks)
  - Fear of “insider attacks”

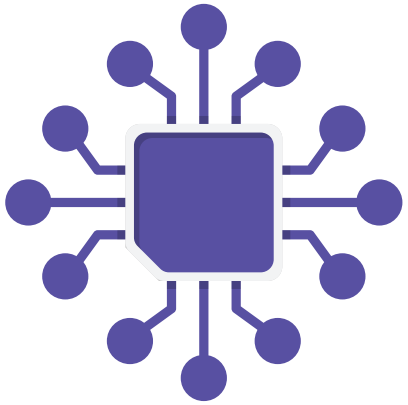
The 'castle and moat' approach to system security is inadequate for modern threats.
In today's cyber landscape, attackers easily bypass perimeter defenses, rendering this traditional model ineffective. A layered, proactive, zero-trust approach is needed — it provides more robust protection against sophisticated threats.



User Software

Adoption of "non-traditional workloads" makes support for containerized workloads mandatory

See CAE context before – if relying on non ISV codes, these effects multiply



- High complexity of installation
- Multiple dependencies
- A lot of legacy software
- Updates and changes on the host system impact the user software
- Performance issues due to software upgrades
- Reproducibility is required
- Portability is needed



Convenient solution: containers
Greatly helps with BYO* / UDSS, also way to improve security

But as always, there's catch:
Using containers solves some challenges and brings new ones...
+ "wait, this is new, we do things differently" mentality ;-)

Causes for Lack of Trust in Containers

Or well… "Hesitation"

- **User namespaces** and potential kernel attack surface
  - A number of vulnerabilities in namespaces in the past few years, in most cases the HPC environment wasn't impacted,
  - Kernel should be patched.

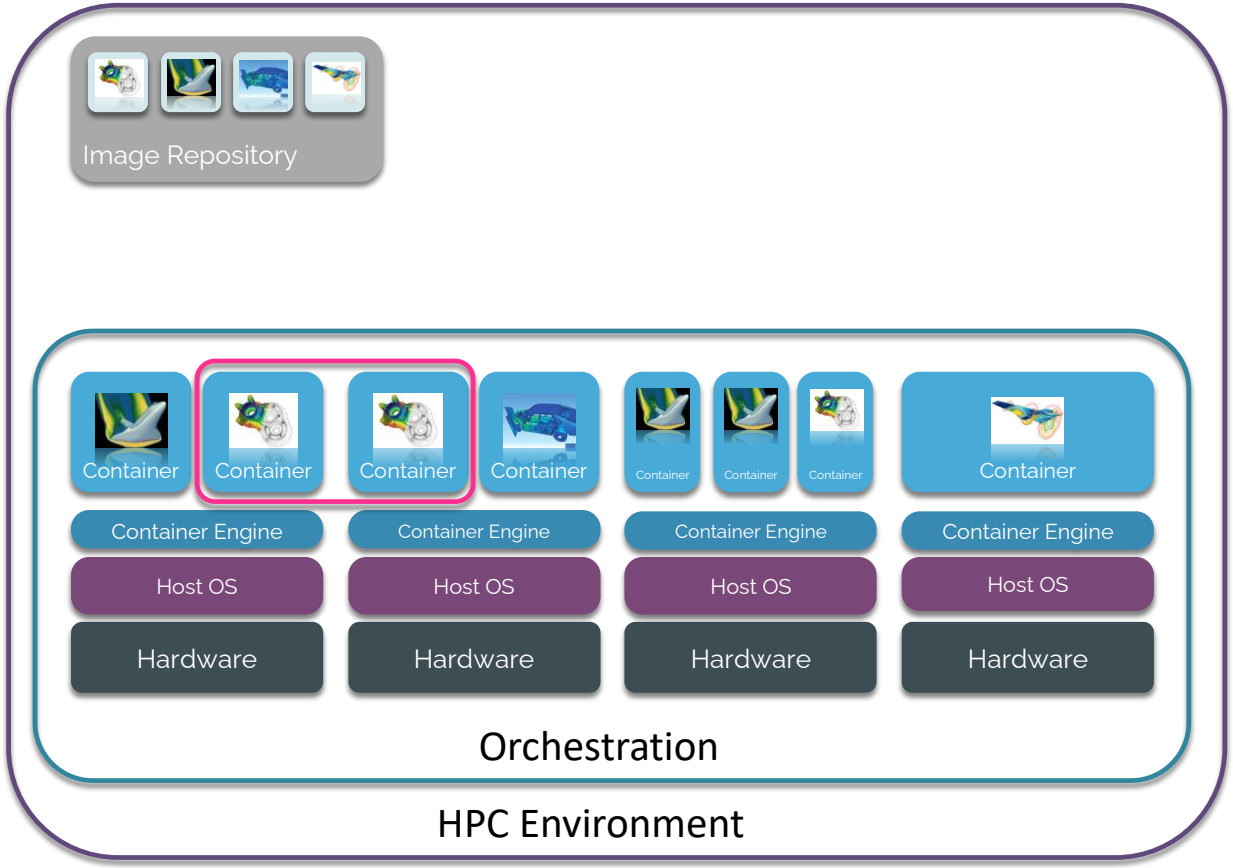
- **Isolation limitations**: containers != Virtual machines
  - Containers are processes, no hypervisor between the host and the container
  - Often (per job) host-level isolation, may change in the future (many core CPUs + cohabitation, …)

- **Supply chain risks** – how to verify the images and what's inside?
  - Containers sometimes considered black bloxes – what's inside? Which version?
  - Is a vulnerability "relevant"?
  - How to control and monitor the workload running in the container?

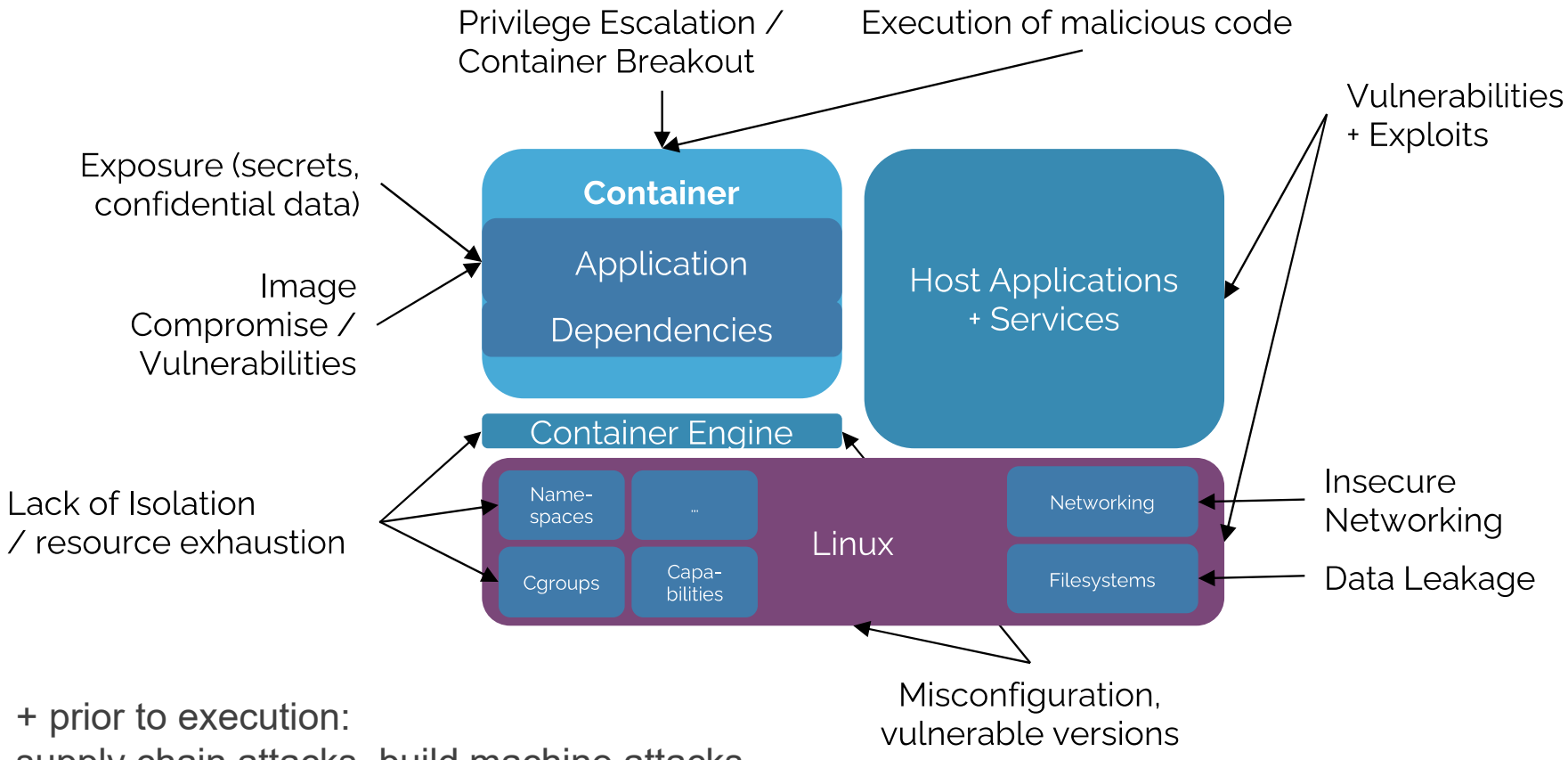
- **Performance and compatibility** considerations (GPU, MPI)
  - Typically, direct resource access required (GPU, Infiniband, …)

Note:
All these supply chain related topics already apply without containers, dealing with them however gets easier in containerized operation.
Ideally: SW@Cluster = SW@Host + Containers

What do we need to protect?



Typical Attack Vectors



+ prior to execution:
supply chain attacks, build machine attacks

Isolation

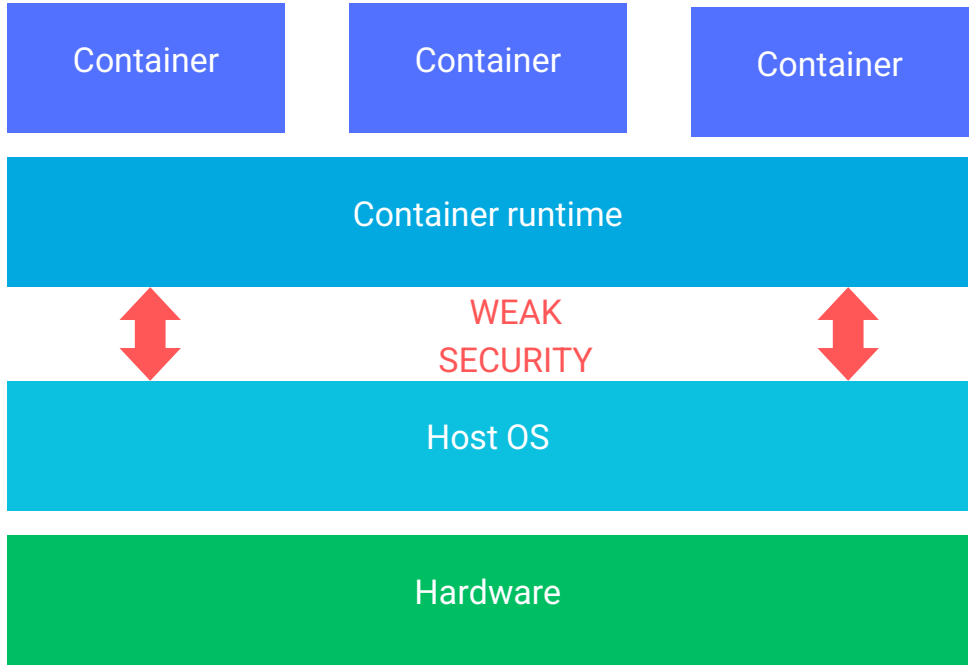
Basics

- Cgroups
  - Control resource usage.
  - Avoid over usage (DoS) by individuals
  - Dismiss the noisy neighbour effect.

- Namespaces
  - Independent view of global resources such as Process IDs
  - Apptainer by default uses user namespaces (since version 1.1) and mount namespaces, additional isolation can be done by using IPC, network, PID, and UTS namespaces for further isolation)
  - Reduce attack surface

Shared OS kernel, often considered
as weak boundary between the host and the container.

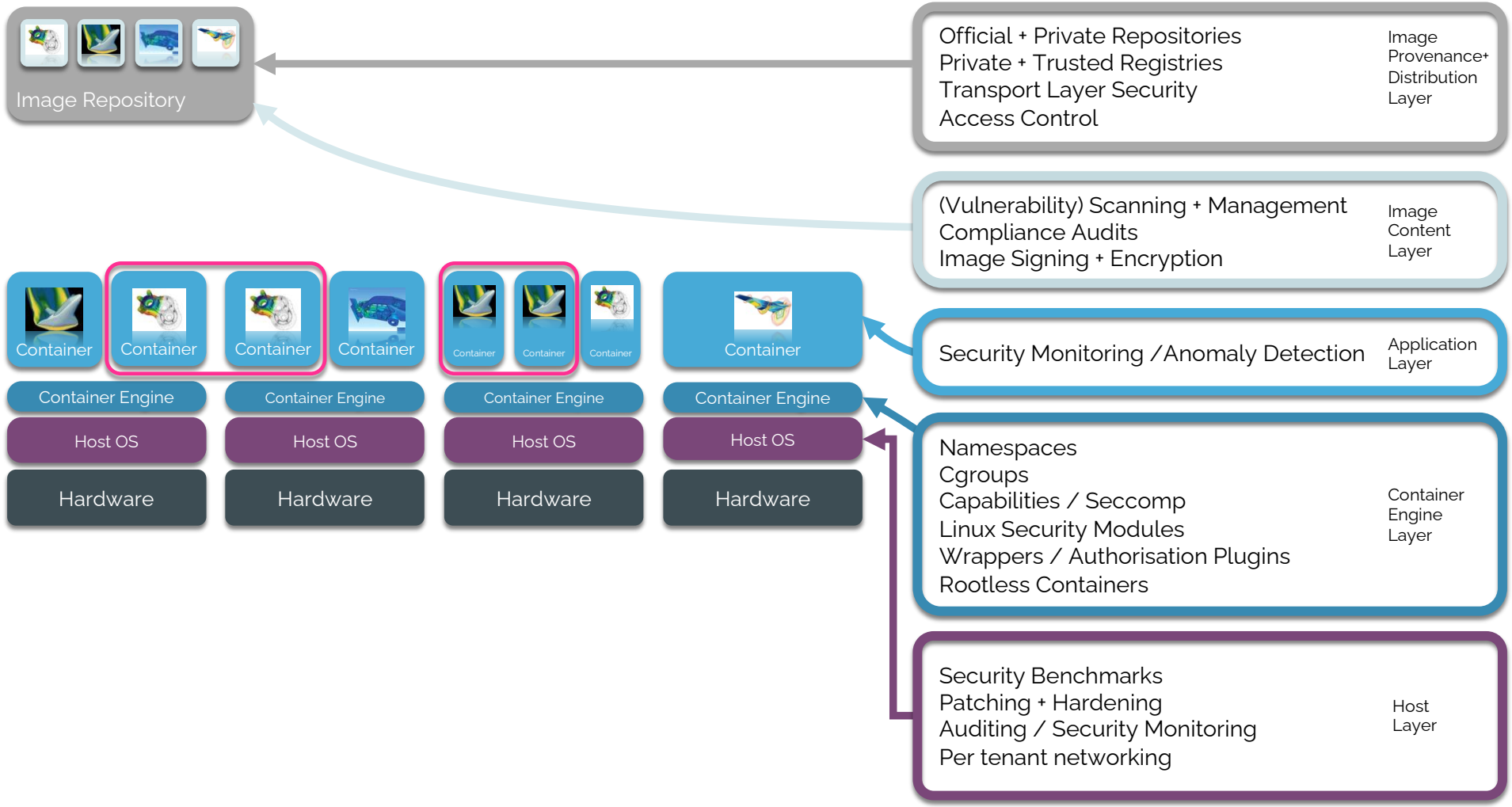
But there is way more applicable to secure containers…



Namespace	Description	Controller	Description
pid	Process ID	blkio	Access to block devices
net	Network Interfaces, Routing Tables, …	cpu	CPU time
ipc	Semaphores, Shared Memory, Message Queues	devices	Device access
mnt	Root and Filesystem Mounts	memory	Memory usage
uts	Hostname, Domainname	net_cls	Packet classification
user	UserID and GroupID	net_prio	Packet priority

Applicable Security Mechanisms

Let's try to break it down...



science + computing

a Bull business

A) Site Operator Perspective

Protect the environment from the user

Main Threats in this Scenario

- Supply chain security: untrusted images, outdated dependencies, malicious image as source for malware distribution
- Privilege escalation and container escape
- Resource contention and performance degradation (noisy neighbour effect)
- Data leakage and espionage

- **Key Question: Who is running what?**

science + computing

a Bull business

Prior to execution

Supply-Chain Attacks

Why does this matter? Just to name a few…

2020 Solarwinds

- Russian state-sponsored hackers gained access to SolarWinds' build system and inserted trojanized code
- Long-term espionage access.
- Updates delivered to ~18,000 organisations

2021 Ua-parser-js

- Node Package Manager (NPM) hijacked, cryptominer injected into popular JavaScript Library
- 2,5+ millions relied (directly / indirectly) on this parser

2025 tj-actions/changedfiles

- Attackers gained access to repo + retroactively modified existing versions tags
- Pipelines pinned to these previously trusted versions executed compromised code
- Scanned the host environment

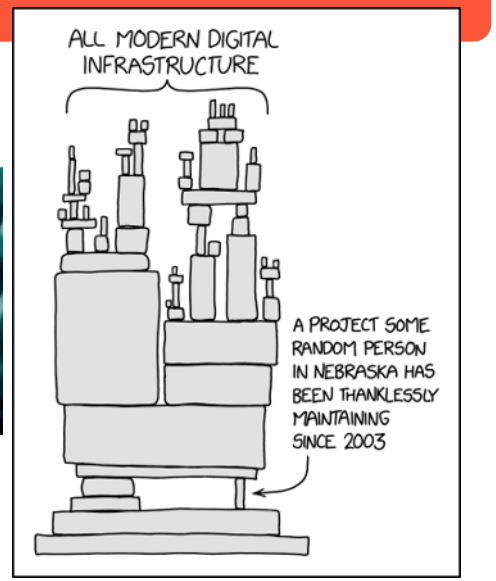
2021 Log4shell

- RCE in ubiqitious logging library

Log4j

- I guess we all remember…





2026 Trivy

- New achievement unlocked: security tool as attack vector

Note:
Many vulnerabilities buried deep in the dependency tree. Hard to find – hard to patch. HPC envs often rely on compiled software – doesn't make scanning for these easier…

science + computing
a Bull business

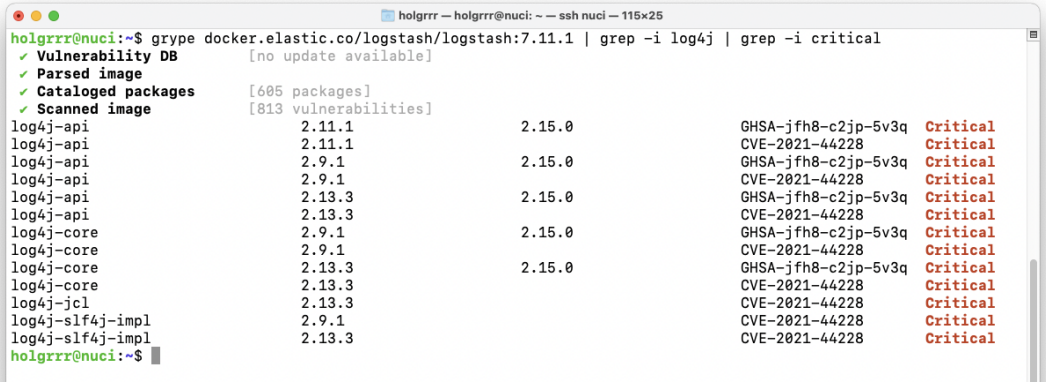
Attacks selection inspired by Barbara's 2026 update talk at the High Performance Containers Workshop at ISC2026

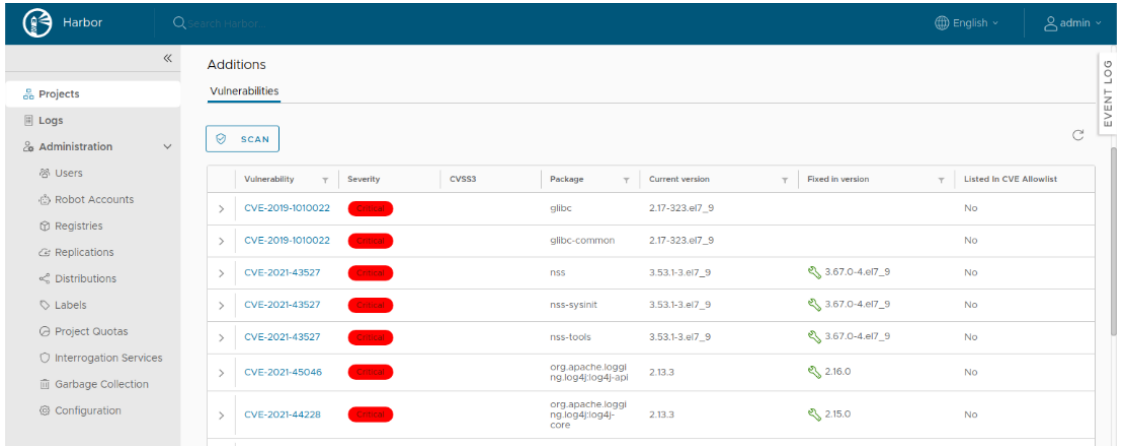
Vulnerabilities, SBOM

Still, containers ease identifying vulnerabilities

- Standalone image scanners / registry integration
  - Representatives: Clair, Syft / Grype, Trivy
  - Native support for OCI images
  - Singularity Image Format possible
    (directory scan of loopback mount)
  - Often extensible by plugins
  - Limitation: custom added code
- Image exploration with Dive (layer content, …)

- Scanning simple, next steps complicated
  - Deactivate image?
  - Rebuild image?
  - Use in strongly isolated environment?
  - Does the vulnerability affect me at all?

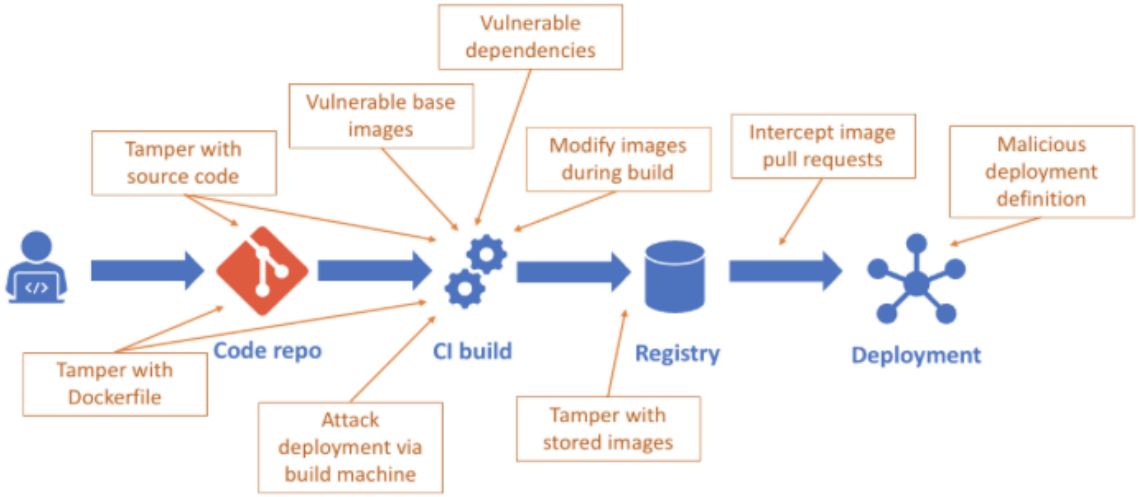




science + computing
a Bull business

Secure Image Provisioning

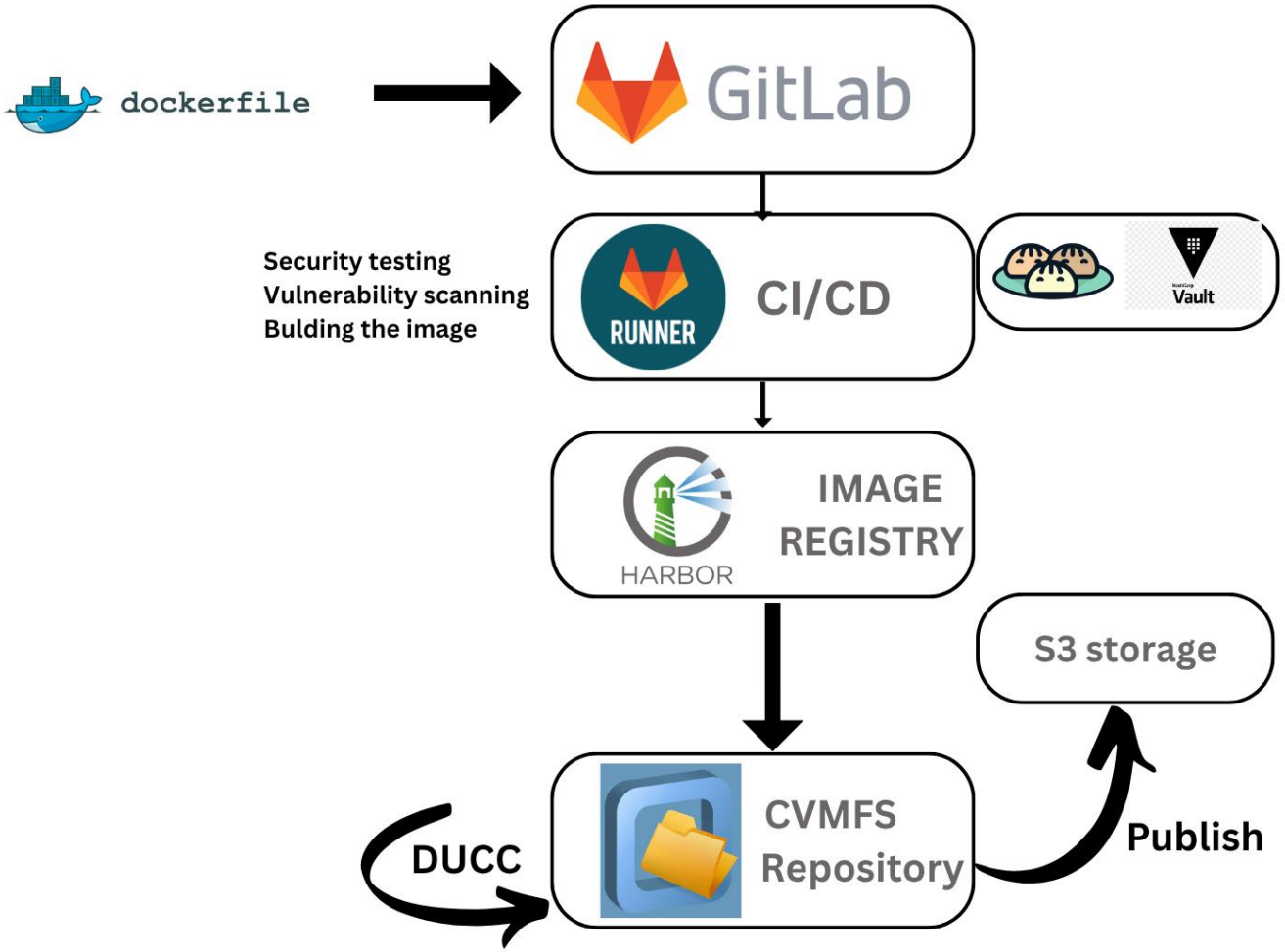
- Secure image storage: registry
  - Should also be “Single Source of Truth”
  - harden registry host, **prune stale images, publish only images that pass security tests**, use TLS to transfer images to the registry
- Use automatic deployment of images with security tests and vulnerability scanning
- Remote build machine



Source: Liz Rice: Container Security, O'Reilly, 2020

Automatic Deployment of Images to Private Registry

Example



science + computing

a Bull business

Security Controls to protect the Host OS and the Container Runtime

Linux Kernel Capabilities

Coarse customisability approach

Linux capabilities split root privilege into multiple capabilities/privileges that can be granted to processes

- Apptainer/Singularity by default use CAP_SYS_ADMIN, CAP_MKNOD, CAP_SETUID, CAP_SETGID, CAP_DAC_OVERRIDE and CAP_CHOWN

- Capabilities can be added to user

(see https://docs.sylabs.io/guides/3.0/admin-guide/configfiles.html)

```
nsc-login1 ~# apptainer capability add --user=barbara CAP_NET_RAW
WARNING: Adding 'CAP_NET_RAW' capability will likely allow user barbara to escalate
privilege on the host
WARNING: Use 'apptainer capability drop --user barbara CAP_NET_RAW' to reverse this
action if necessary

nsc-login1 ~# apptainer capability list barbara
barbara [user]: CAP_NET_RAW
```

Seccomp Profiles

Fine-grained approach

```
SECCOMP
template: https://github.com/apptainer/singularity/blob/master/etc/seccomp-profiles/default.json

singularity exec --security seccomp:/path/to/seccomp.json my_container.sif

are seccomp profiles supported in the kernel
# grep SECCOMP /boot/config-$(uname -r)
CONFIG_HAVE_ARCH_SECCOMP=y
CONFIG_HAVE_ARCH_SECCOMP_FILTER=y
CONFIG_SECCOMP=y
CONFIG_SECCOMP_FILTER=y
# CONFIG_SECCOMP_CACHE_DEBUG is not set

strace -e trace=socket,connect,accept,bind,listen sendto recvfrom sendmsg recvmsg ping 8.8.8.8

nsc-login1 ~# dmesg -T| grep seccomp
nsc-login1 ~# journalctl -k | grep seccomp
```

Fine-grained control over which system calls are allowed or denied

Restrictions to the exact need of the process

Blacklist commands

Challenging to create them manually, tools available to automate profile creation and monitor syscall activities: *strace, Tracee, seccomp-tools, Sysdig Falco, oci-seccomp-bpf-hook, etc.*

```
--security="seccomp:/usr/local/etc/singularity/seccomp-profiles/default.json"
--security="apparmor:/usr/bin/man"
--security="selinux:context"
--security="uid:1000"
--security="gid:1000"
--security="gid:1000:1:0" (multiple gids, first is always the primary group)
```

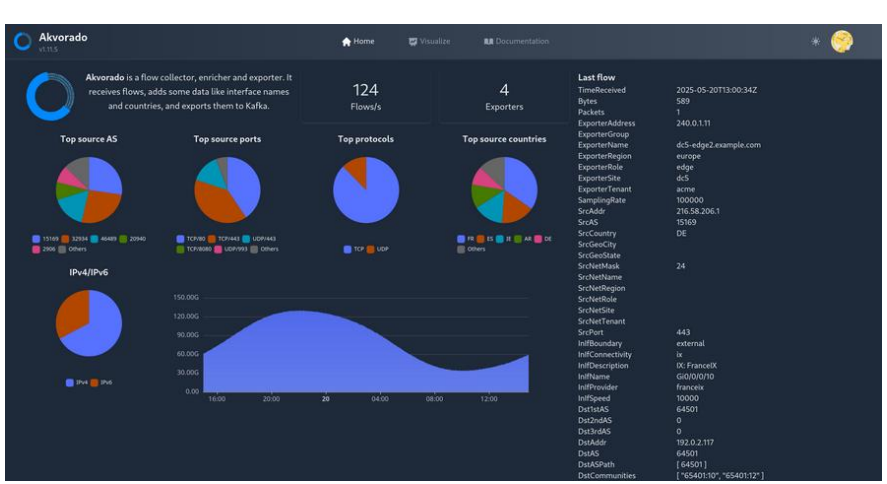
Security Monitoring / Auditing

“ Classics”

- **Enable** auditing (auditd, SElinux, AppArmor) – big difference to 20yo best practice ;)

Next steps

- Run active security monitoring on the logs (ELK, Greylog, OpenSearch, Loki)
- Use Role-based Access Control (RBAC)
- Monitor resource usage by Prometheus/Grafana
- Automated pentesting: Pakiti, Greenbone, Nessus, OpenVAS



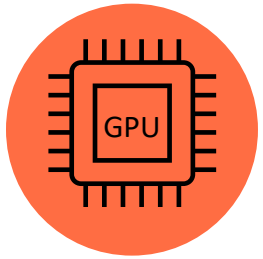
- Or take the actual workload more into account

Workload Monitoring



System Calls

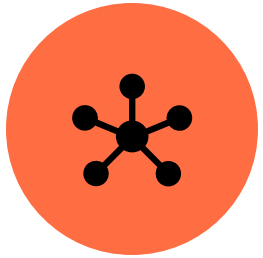
- Great level of details (call pattern, files used, …), collectable with various tools (strace, eBPF, Sysdig, …), on the host level - no ”agent” required (as opposed to VMs)
- Can be used to distinguish cryptomining / ransomware from typical workload



GPU Usage

- Different workloads utilize different functions of the GPU

P. Zou, A. Li, K. Barker, and R. Ge. Detecting Anomalous Computation with RNNs on GPU-Accelerated HPC Machines. ACM International Conference Proceeding Series, 2020. doi: 10.1145/3404397.3404435.



Infiniband Communication

- IB Integrated Performance Monitoring Logs to fingerprint the codes used

S. Peisert. Fingerprinting Communication and Computation on HPC Machines. (LBNL-3483E), 2010. http://www.cs.ucdavis.edu/~peisert/research/2010-Peisert-TR-LBNL-3483E-FingerprintingHPC.pdf



User Monitoring

- Analysis of multitude of log files capturing the user’s full behavior on a site.
- “Meta Data” + entered commands, screen output + X

G. Zhu, Y. Zeng, and M. Guo. A Security Analysis Method for Supercomputing Users’ Behavior. Proceedings - 4th IEEE International Conference on Cyber Security and Cloud Computing, CSCloud 2017 and 3rd IEEE International Conference of Scalable and Smart Cloud, SSC 2017, doi: 10.1109/CSCloud.2017.19.

Note: Closely consider legal aspects. Especially when considering detailed User Monitoring.

science + computing

a Bull business

B) User Perspective

Protect the user from the environment

Why does this matter? Shared usage, sometimes lack in operator trust

Example: Alice Recoque - Europas 2. Exascale System – big interest from industry users

More than a single supercomputer Alice Recoque will expand Europe's AI and research capabilities while ensuring energy efficiency and sovereignty

Paris, France – November 18, 2025 – Eviden, the Atos Group product brand leading in advanced computing, and AMD (NASDAQ: AMD) announced their selection to build Alice Recoque, a next-generation supercomputer to support the need for scientific computing (HPC) and artificial intelligence (AI), serving as an AI Factory. Alice Recoque will be France's first and Europe's second Exascale supercomputer, a machine to expand Europe's AI and research capabilities while ensuring energy efficiency and sovereignty.

This project, led by France GENCI, operated by CEA, will have the capacity to surpass one Exaflop per second (i.e. a billion billion calculations per second) for scientific simulation. This extreme performance, equivalent to more calculations than humanity could accomplish in four years of relentless mental computation, will mark a fiftyfold an increase in the computing capacity compared to the previous system, while only multiplying by 5 the electrical power.

This project, representing an overall cost of 554 million euros over 5 years of operation, is funded by EuroHPC JU, with budget stemming from the Digital Europe Programme (DEP), and by the Jules Verne consortium, led by France through GENCI and CEA with the participation of Netherlands with SURF and Greece with GRNET.

Alice Recoque will be installed within France's CEA's Very Large Computing Center (TGCC), which already hosts Eviden-built systems such as GENCI's Joliot-Curie machine and CEA's Topaze machine.

An AI-HPC factory to tackle Europe's most pressing AI challenges

Alice Recoque will tackle Europe's most pressing societal, scientific, and industrial challenges by combining traditional High-Performance Computing (HPC) and Artificial Intelligence (AI) workloads from large-scale simulations, data analysis, and AI models.

Alice Recoque covers the entire computing lifecycle, integrating cutting-edge hardware and advanced AI software to deliver scalable high-impact solutions. This project is a live implementation of the strong collaboration and commitment from Eviden and AMD to accelerate research and industrialization of HPC-AI use cases, with a major investment in both human and technological resources.

The acceleration partition of the Alice Recoque system will be powered by next-gen AMD EPYC™ CPUs, codenamed "Venice," AMD Instinct™ MI430X GPUs—a new MI400 Series accelerator engineered for sovereign AI and scientific computing— and AMD FPGAs. In addition, Alice Recoque will welcome a scalar compute partition based on the European Rhea2 processor developed by SiPearl. The entire system will be interconnected by Eviden's network solution (BXIv3) into its newest BullSequana XH3500 platform, along with storage solutions provided by DDN.

This powerful ecosystem will enhance climate modeling, accelerate innovation in materials and energy, enable digital twins for personalized medicine, and support next-gen European foundational AI models. It will also address the vast amount of data generated by scientific instruments such as telescopes, satellites as well as IoT devices, and AI applications, driving breakthroughs across multiple domains.

Such large scale systems highly interesting for industry users – which have high interest in increased isolation – more later

Source: https://www.bull.com/en/press-releases/alice-recoque

Main Threats in this Scenario

- Data leak, unauthorised access to data
- Data loss

- Runtime problems: containers run in privileged mode, misconfigurations, sensitive directories bind-mounted to container
- Privilege escalation, (lack of) isolation from other users
- Supply chain risks: usage of malicious/infected images, software vulnerabilities

- Key aspects:
  - How to protect the images and private registries?
  - How to ensure privacy and security of the data?

Use verified / trusted images

- **Sign images** (supported in Singularity 3.0) with PGP key and verify the signature: access tokens can be used.
- Use **automatic deployment** (CI/CD)
- Avoid unnecessary code or packages
- Remove development tools from the image
- Use multi-stage build
- **Do not use arbitrary third-party base images**
- Perform **a secret and vulnerability scan** of the images

```
$ singularity keys newpair
Enter your name (e.g., John Doe) : Barbara Krasovec
Enter your email address (e.g., john.doe@example.com) : barbara.krasovec@ijs.si
Enter optional comment (e.g., development keys) : Containers NSC key
Enter a passphrase :
Retype your passphrase :
Generating Entity and OpenPGP Key Pair... done
barbarak@sampo:~$ singularity key list
Public key listing (/home/barbarak/.apptainer/keys/pgp-public):

0)  User:              Barbara Krasovec (containers NSC key) <barbara.krasovec@ijs.si>
 Creation time:     2025-05-09 08:07:07 +0200 CEST
 Fingerprint:       2F1B3AA3C55D9722FF049BA0A6DA71F99B0920CD
 Length (in bits):  4096

$ singularity keys push 2F1B3AA3C55D9722FF049BA0A6DA71F99B0920CD
INFO:    Key server response: Upload successful. This is a new key, a welcome email has been sent.
public key `2F1B3AA3C55D9722FF049BA0A6DA71F99B0920CD' pushed to server successfully

$ singularity keys search barbara.krasovec@ijs.si
Showing 1 results

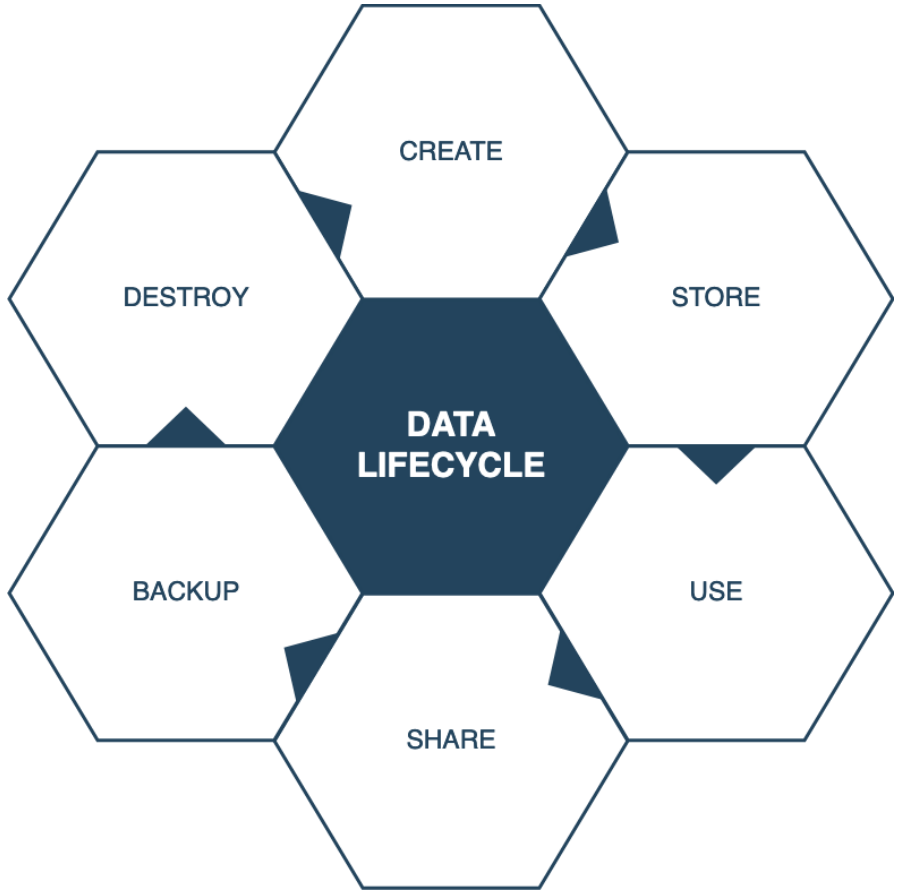
KEY ID    BITS  NAME/EMAIL
9B0920CD  4096  Barbara Krasovec (containers NSC key) <barbara.krasovec@ijs.si

$ singularity sign singularity-defs/mpi-mellanox.sif
INFO:    Signing image with PGP key material
Enter key passphrase :
INFO:    Signature created and applied to image 'singularity-defs/mpi-mellanox.sif'

$ singularity verify singularity-defs/mpi-mellanox.sif
INFO:    Verifying image with PGP key material
[LOCAL]   Signing entity: Barbara Krasovec (containers NSC key) <barbara.krasovec@ijs.si>
[LOCAL]   Fingerprint: 2F1B3AA3C55D9722FF049BA0A6DA71F99B0920CD
Objects verified:
ID  |GROUP   |LINK    |TYPE
------------------------------------------------
1   |1       |NONE    |Def.FILE
2   |1       |NONE    |JSON.Generic
3   |1       |NONE    |JSON.Generic
4   |1       |NONE    |FS
INFO:    Verified signature(s) from image 'singularity-defs/mpi-mellanox.sif'
```

"Big Data"

- No (sensitive or other) data should be in the container
  - General rule, same for AI/ML training data and other input data
- Encryption is often not possible (too much data)
- Security is limited: access controls, RBAC, and token-based access.
- The whole data lifecycle should be addressed.

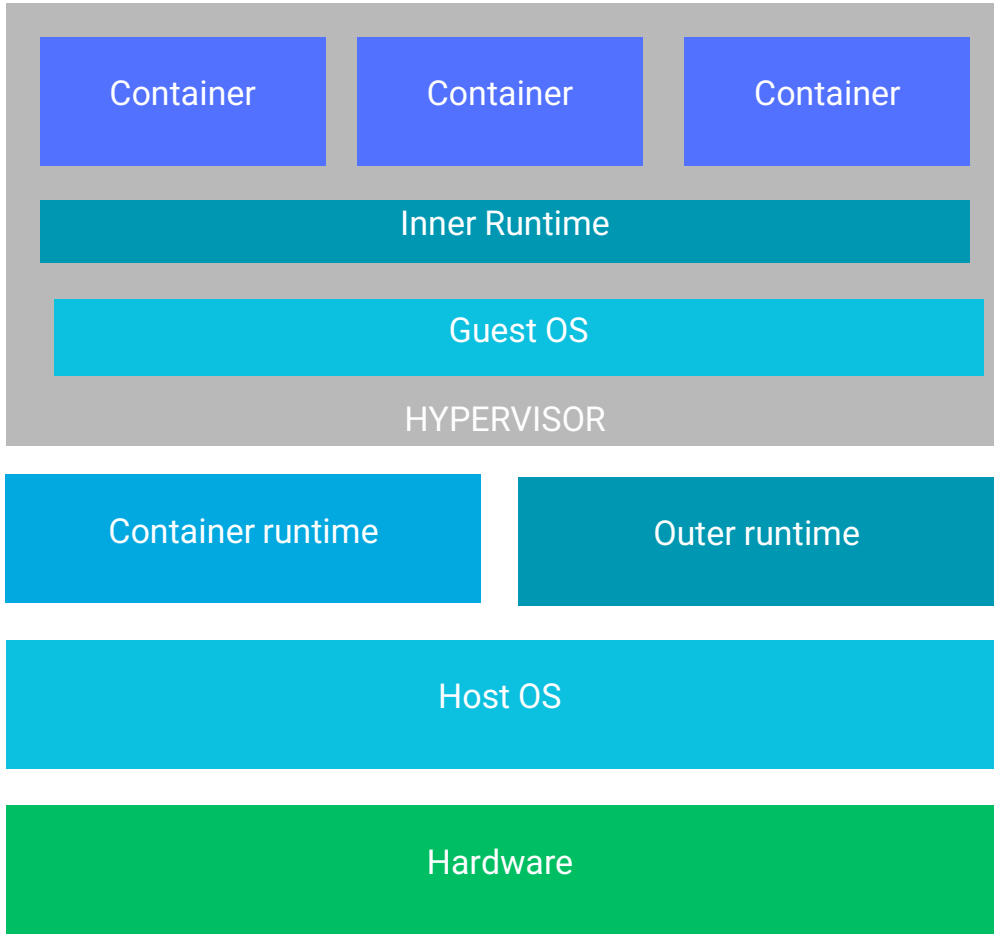


What about data in memory?

Sandboxing

Different solutions available:

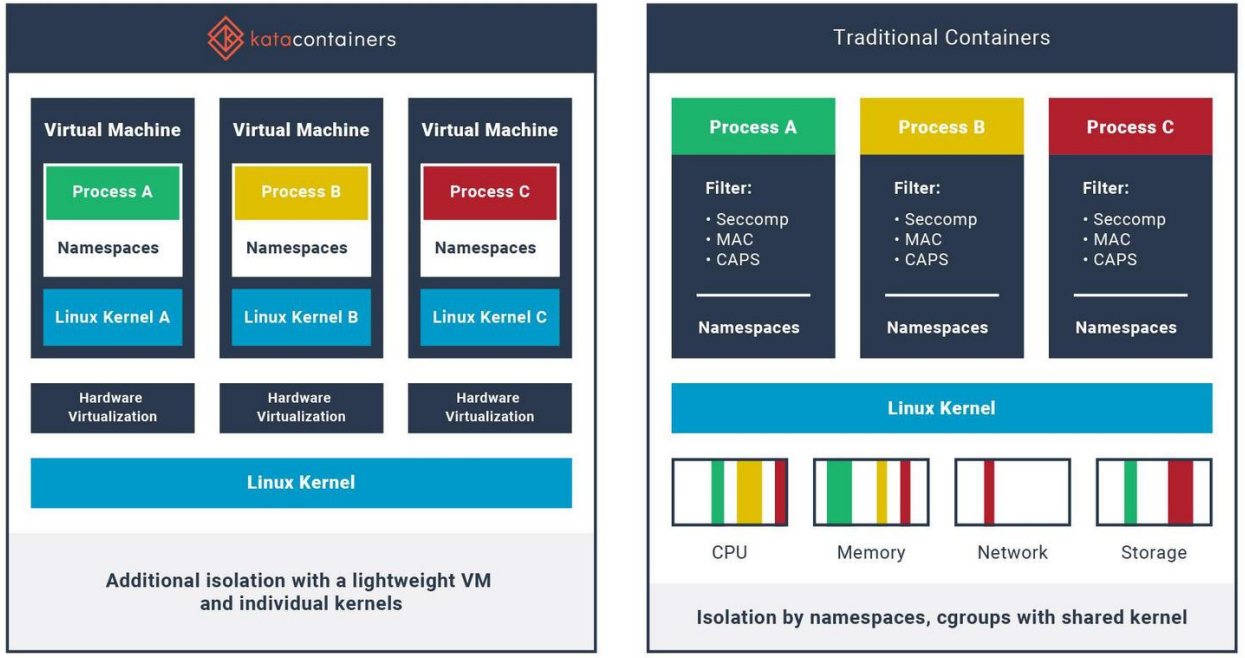
- **IBM Nabla**: unikernel as a process with reduced syscalls
- **gVisor**: user space kernel, implementation of a majority of syscalls, but no GPU support
- **Firecracker**: uses KVM, works as a microVM, no GPU support
- **Kata Containers**: microVM, supports multiple hypervisors, has support for GPU
- **Bubblewrap**: wrapper around namespaces and seccomp profiles
- **Sydbox**: uses seccomp profiles, namespaces, landlock, ptrace and MDWE
- **Firejail**: requires setuid → large attack surface



Kata Containers

Vs. traditional containers

- Feels like a container runtime, but is actually a lightweight containers under the hood
- Merged Intel Clear Containers with Hyper.sh RunV
- Integrates well with typical container environment
- Low performance overhead*
- Limitations apply (RDMA, …)

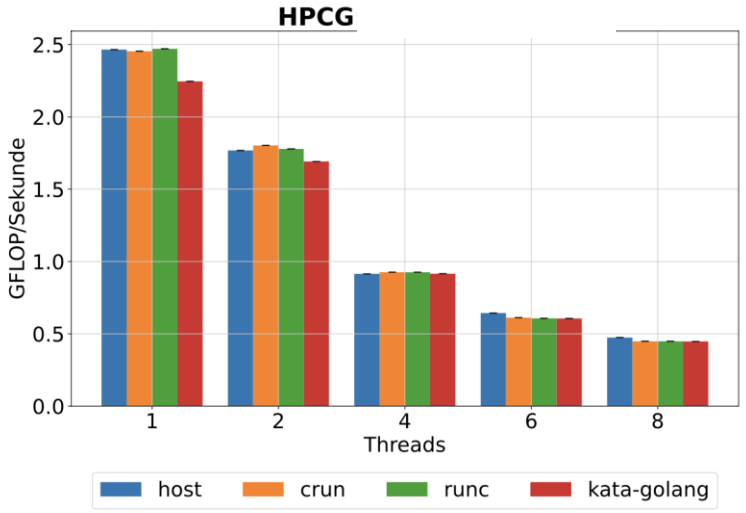


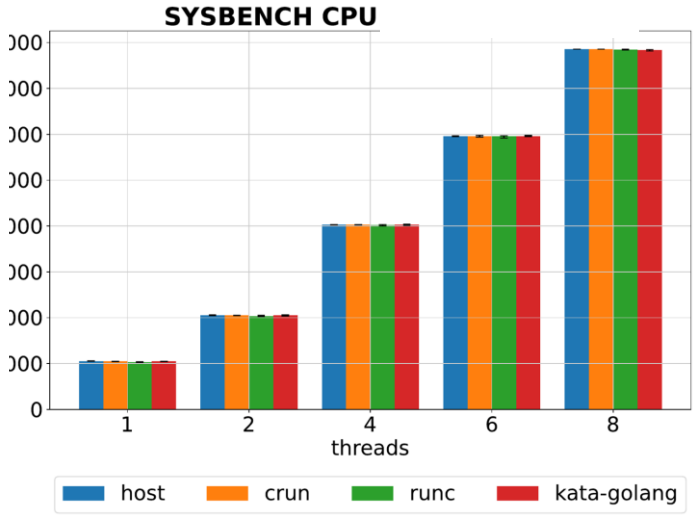
Source: https://katacontainers.io/learn/

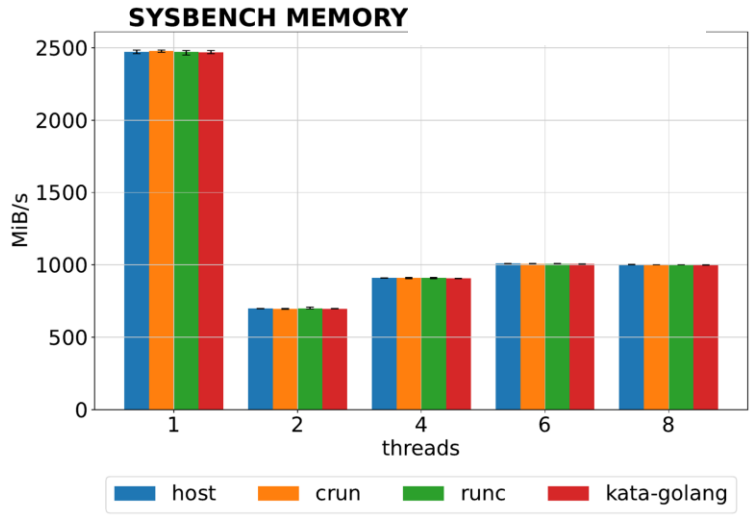
* Your mileage may vary, depending on the actual workload.

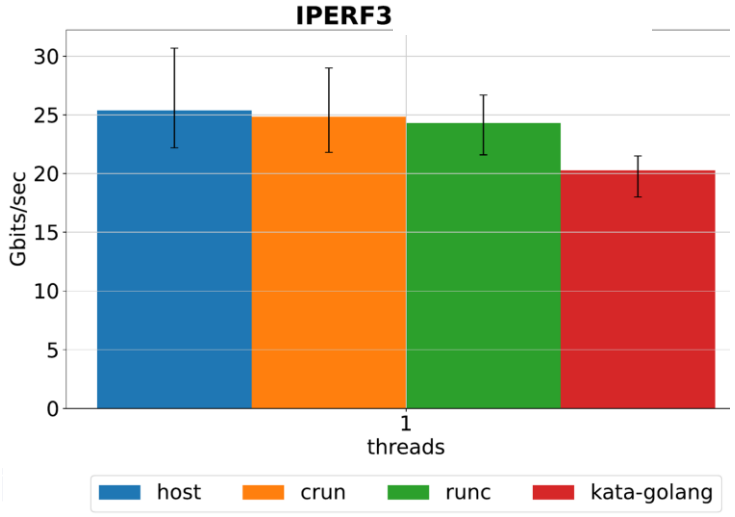
Experience with Kata Containers (Q1/2025)

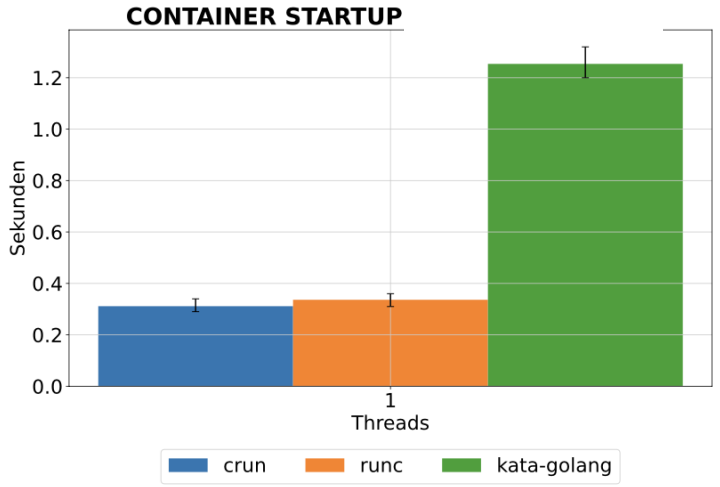
Bachelor Thesis Justin Dretvic, Hochschule Furtwangen, 2025











Results from evaluation on Bare Metal System
8 Cores (no HT), 16GB RAM
Ubuntu 24.04 LTS; Docker 27.3.1;
runc 1.2.2; crun 1.18.2; Kata Container 3.10.1

Among others: HPCG 3.1; Sysbench 1.0.20;iPerf3 3.16

Sysbench CPU below 2%
Sysbench Memory : below 2%
IPerf3 Network: up to 20%

science + computing

a Bull business

Options for processing confidential and sensitive data in use – aka *“the holy grail”*

Protecting *Data in Use*

Not yet fully solved



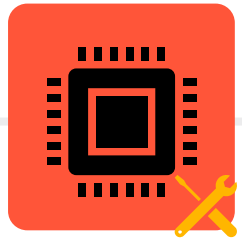
Data at Rest

Add your favorite encryption solution



Data in Transit

Add your favorite encryption solution



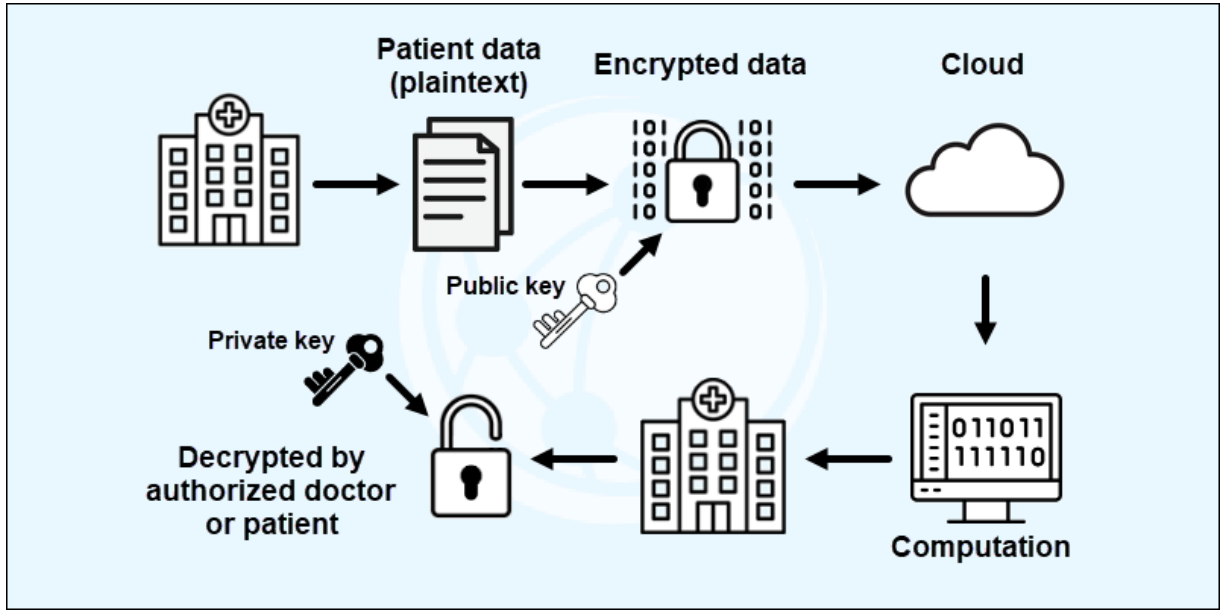
Data in Use

Not fully established, yet

Homomorphic Encryption

Holy grail - allows computations to be performed on encrypted data.

- Sensitive information can remain secure and unexposed while being processed.
- Different tools available, such as IBM's Fully Homomorphic Encryption (FHE) Toolkit for Linux, Microsoft SEAL, HElib, and PALISADE.
- Computationally very expensive and will impact performance (severely) .
- Not general purpose – potentially suited for special data (e.g. patient data)



Source: https://phoenixnap.com/kb/homomorphic-encryption

Confidential Computing

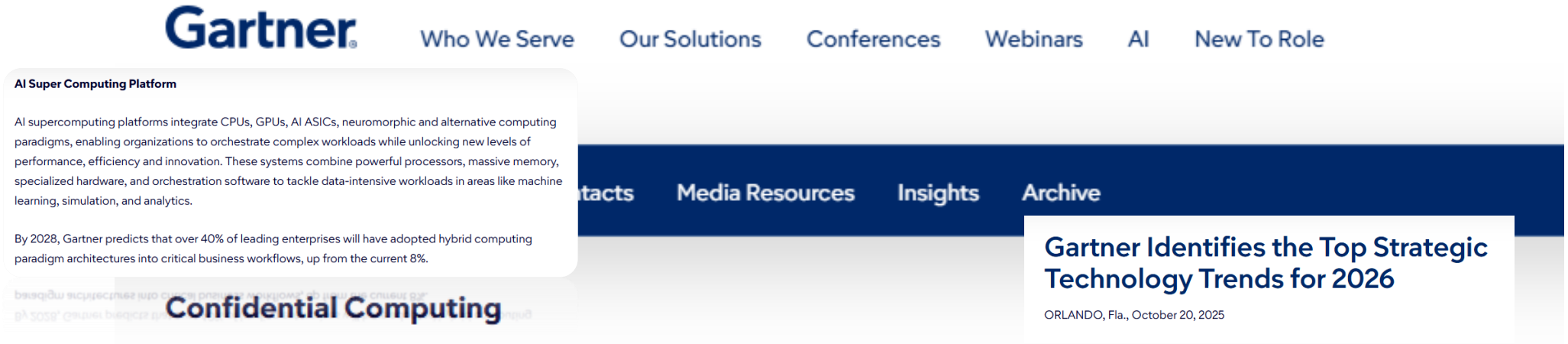
Hardware-based ("Trusted Execution Environments (TEEs)")

- Securing data using dedicated hardware components
- Most (all?) public cloud providers provide hardware encryption
- Trusted execution environments (TEEs) are isolated execution environment within a device's processor, (e.g., Intel SGX+TDX, AMD SEV, or Arm TrustZone) and are based on hardware encryption
- TEEs provide a secure enclave that isolates data and computations from the rest of the system
- To protect data in memory also from administrator on the system
- Trust moves to CPU vendor, using remote attestation
- Note: Performance overhead and other limitation to be expected (accelerators, interconnects, mutli-node workloads, …)

Note:
Dedicated talk at Tuebix 2026

Confidential Computing is Gaining Momentum

Gartner Strategic Technology Trends for 2026



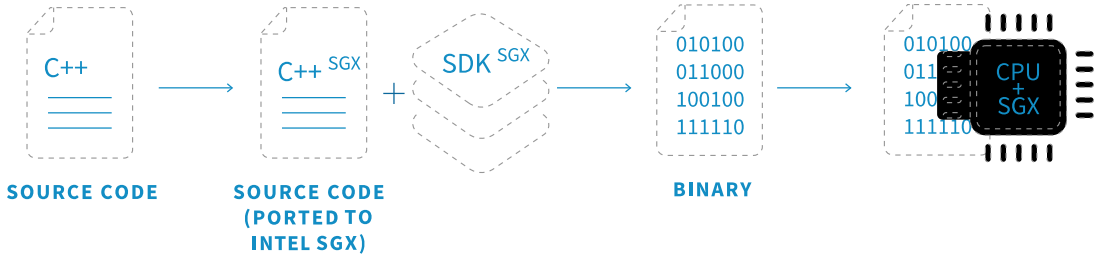
Confidential computing changes how organizations handle sensitive data. By isolating workloads inside hardware-based trusted execution environments (TEEs), it keeps content and workloads private even from infrastructure owners, cloud providers, or anyone with physical access to the hardware. This is especially valuable for regulated industries and global operations facing geopolitical and compliance risks and for cross-competitor collaboration.

By 2029, Gartner predicts more than 75% of operations processed in untrusted infrastructure will be secured in-use by confidential computing.

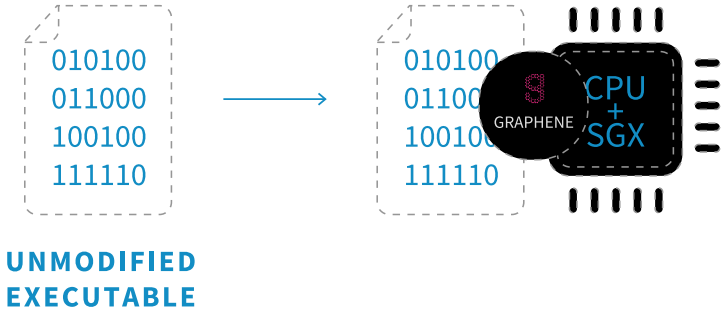
Source: https://www.gartner.com/en/newsroom/press-releases/2025-10-20-gartner-identifies-the-top-strategic-technology-trends-for-2026

Gramine + Gramine Shielded Containers

- Intel SGX is limited by need to modify code – problematic with ISV code
- Gramine [1] supports to run unmodified applications inside a secure enclave
- ”Developers provide a manifest file to configure the application environment and isolation policies, Gramine automatically does the rest.” [1]
- Gramine Shielded Containers (gsc) [2] “graminize” container images to run inside an Intel SGX enclave
- BUT: Expect severe performance hit



[1]



[1]

Sources:
[1] https://gramineproject.io/
[2] https://github.com/gramineproject/gsc

Confidential Computing + GPUs

- Support for confidential computing: **NVIDIA H100 (in 2022) and B100 (in 2024).**
- **H100** uses bounce buffers to exchange data between CPU and GPU, if the workload has a lot of communication between CPU and GPU, the performance overhead is significant (CPU limitations).
- **B100** supports TDISP/IDE [1], a new PCI security standard, all encryption is done on the PCI-E bus, you get full performance of the Blackwell architecture.

[1] https://pcisig.com/blog/ide-and-tdisp-overview-pcie%C2%AE-technology-security-features

Encryption of Network Traffic



Multi-tenant network

- **NVIDIA's BlueField-3** is a Data Processing Unit (DPU)
  - multi-tenant environments using virtual interfaces
  - offloading and accelerating software-defined networking functions
  - hardware-enforced isolation between tenants, ensuring that workloads remain secure and independent.
  - network traffic can be encrypted per tenant
  - reduces CPU utilisation

Summary: State of Confidential Containers in HPC

- Current HPC setups often do not support Confidential Computing yet
  - **Hardware Limitations**: Confidential containers rely on hardware-based encryption (e.g., AMD SEV, Intel SGX), which is not widely available in HPC clusters – might just not be enabled in UEFI
  - **Multi-Tenancy Issues**: Workloads often require direct access to shared hardware, conflicting with CoCo isolation

- **Performance Overhead**: Encryption and secure enclave technologies introduce latency - problem for HPC workloads
  - BUT might be acceptable for some workloads, become less severe with many core CPUs
    - previous multi-node workloads might now fit on single node
  - - RunC and Kata Containers have less performance overhead than gVisor
    - https://doi.org/10.1007/s10586-021-03517-8)
  - Overhead of SEV less severe than SGX (see https://doi.org/10.1109/SEED55351.2022.00018)

Future hybrid setups (HPC, AI, K8s, Cloud, …) will put increased emphasize on multitenancy and are expected to introduce support for Confidential Computing – still at a certain “cost”, in terms of feature limitations and performance overhead.

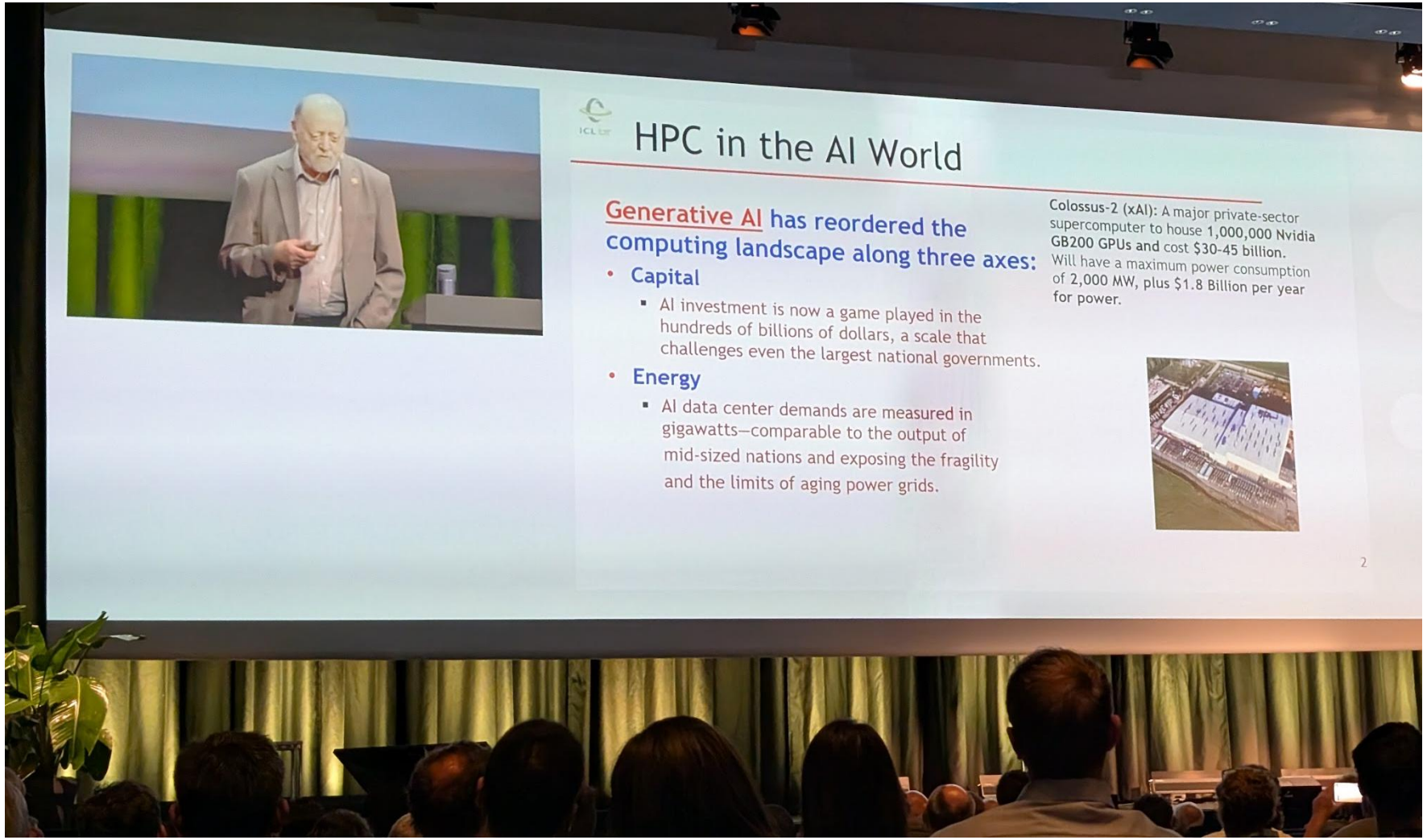
For now: Containers can already improve the security of an HPC environment – if done right

science + computing

a Bull business

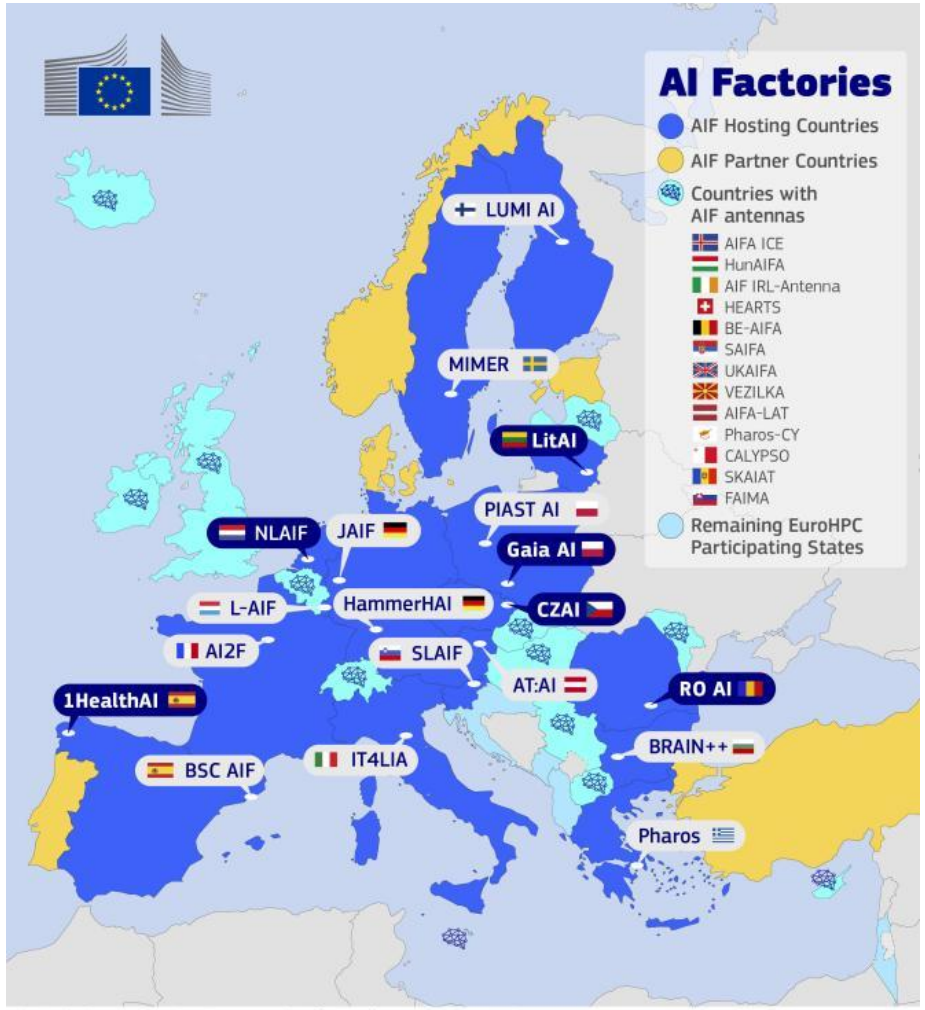
4 Road Ahead

It's an AI world, baby…



ISC2026 Closing Keynote: Jack Dongarra: HPC in Transition
https://www.youtube.com/watch?v=DTRCpQPD9Jw

AI Factories ahead + Gigafactories



Switzerland's participation is contingent upon the ratification of its accession to Horizon Europe.
The boundaries and names shown and the designations used on this map do not imply official endorsement or acceptance by the European Union.
This designation shall not be construed as recognition of a State of Palestine and is without prejudice to the individual positions of the Member States on this issue.
Administrative boundaries: © EuroGeographics © OpenStreetMap
Cartography: Eurostat – IMAGE, 05/2025

Keys facts and figures

€20 billion
Mobilising

Several
First AI Gigafactories

European AI
Crucial input towards a truly European AI ecosystem

Productivity
Strengthened innovation potential & productivity of EU industry

WAS SIND AI-GIGAFACTORIES?

Im Rahmen der europäischen Investitionsinitiative InvestAI plant die EU gemeinsam mit EuroHPC JU und den Mitgliedstaaten den Aufbau leistungsstarker Rechenzentren mit jeweils über 100.000 KI-Prozessoren. Diese Gigafabriken sollen zur Entwicklung komplexer KI-Modelle mit Milliarden von Parametern beitragen und Europa im globalen Wettbewerb – insbesondere gegenüber den USA und China – technologisch stärken.

VON AI FACTORIES ZU GIGAFACTORIES

Bereits bestehende AI Factories – darunter zwei in Deutschland – dienen der Entwicklung kleinerer KI-Anwendungen auf regionaler Ebene. Die geplanten Gigafactories hingegen sollen als zentrale Innovationszentren für europäische „Moonshot"-Projekte fungieren – vergleichbar mit internationalen Großprojekten wie dem US-amerikanischen Stargate.

science + computing
a Bull business

Sources:
https://digital-strategy.ec.europa.eu/en/policies/ai-factories
https://commission.europa.eu/topics/competitiveness/competitiveness-coordination-tool-projects/ai-gigafactories_en
https://www.elektronikforschung.de/service/aktuelles/aufruf-zur-interessenbekundung-aufbau-von-ai-gigafactories-in-europa

What does this imply?

Currently platforms are emerging that differ significantly from traditional HPC environments:

- User Software Stack: AI Workloads instead of numerical simulation
- Workload Deployment: Container-only
- Workload Management: K8s instead of Slurm
- Security “Model”: Multi-Tenancy instead of “Trust-First”

New conditions and requirements are increasingly bringing the topic of security to the forefront

- More open environments, opening up to additional user groups, workloads, …
- Even extending to the use of shared platforms with competitors
  - See Google paying $920 million a month for compute capacity at xAI data centers [1]

“Bring your own” is taking center stage—moving away from ISV code with “liability” and “admin-vetted” solutions, but also bringing greater responsibility for ensuring a secure environment

Also, an ever-growing threat/attack landscape:

- HPC environments becoming goal for attackers, potentially espionage
- Looots of exploits even only in 2026

Source:
[1] https://www.cnbc.com/2026/06/05/google-to-pay-spacex-920-million-a-month-for-xai-compute-capacity.html

science + computing

a Bull business

Thank you!
Questions?

For more information please contact:

Holger Gantikow

holger.gantikow@bull.com

Confidential information owned by Bull SAS, to be used by the recipient only. This document, or any part of it, may not be reproduced, copied, circulated and/or distributed nor quoted without prior written approval from Bull SAS.