XAAS: ACCELERATION AS A SERVICE TO ENABLE PRODUCTIVE HIGH-PERFORMANCE CLOUD COMPUTING
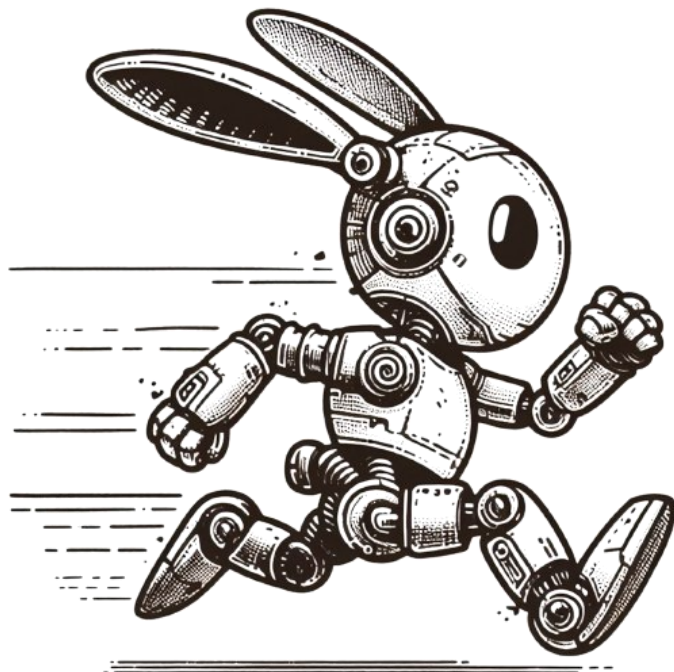
Panelists:
- Dan Ernst (NVIDIA)
- Ian Foster (Argonne National Laboratory)
- Torsten Hoefler (ETH Zurich)
- Thomas C. Schulthess (CSCS)

Moderated by Marcin Copik (ETH Zurich)

What is Axeleration as a Service?

Paper

Theme Article: Converged Computing: A Best-of-Both Worlds of HPC and Cloud

XaaS: Acceleration as a Service to Enable Productive High-Performance Cloud Computing

Torsten Hoefler, *ETH Zurich & Swiss National Supercomputing Centre (CSCS), Switzerland*

Marcin Copik, *ETH Zurich, Switzerland*

Pete Beckman, *Argonne National Laboratory, USA*

Andrew Jones, *Microsoft, United Kingdom*

Ian Foster, *Argonne National Laboratory, USA*

Manish Parashar, *Utah University, USA*

Daniel Reed, *Utah University, USA*

Matthias Troyer, *Microsoft, USA*

Thomas Schulthess, *Swiss National Supercomputing Centre (CSCS), Switzerland*

Dan Ernst, *NVIDIA, USA*
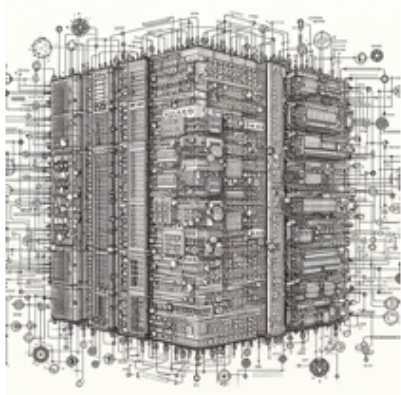
Jack Dongarra, *University of Tennessee, USA*
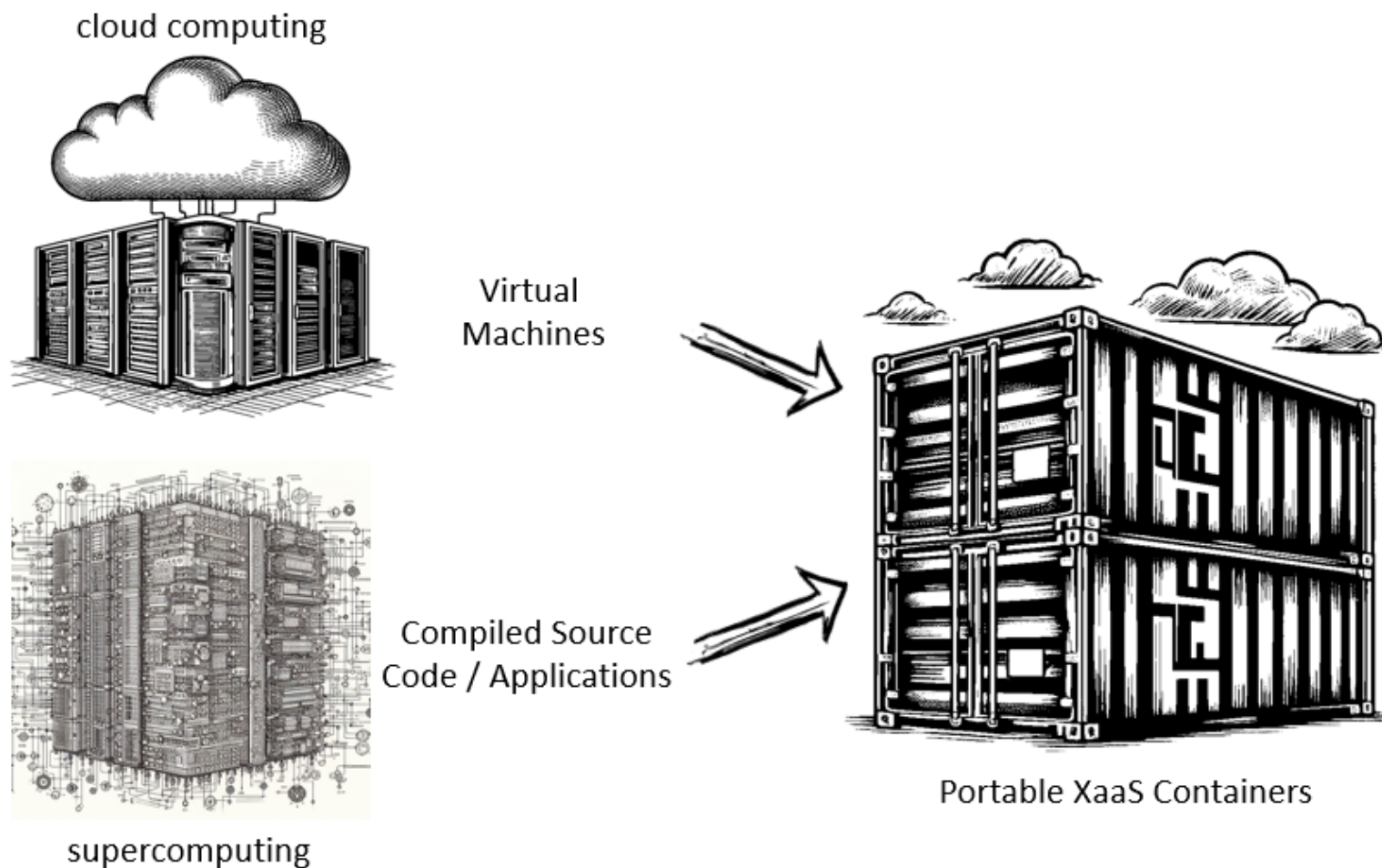
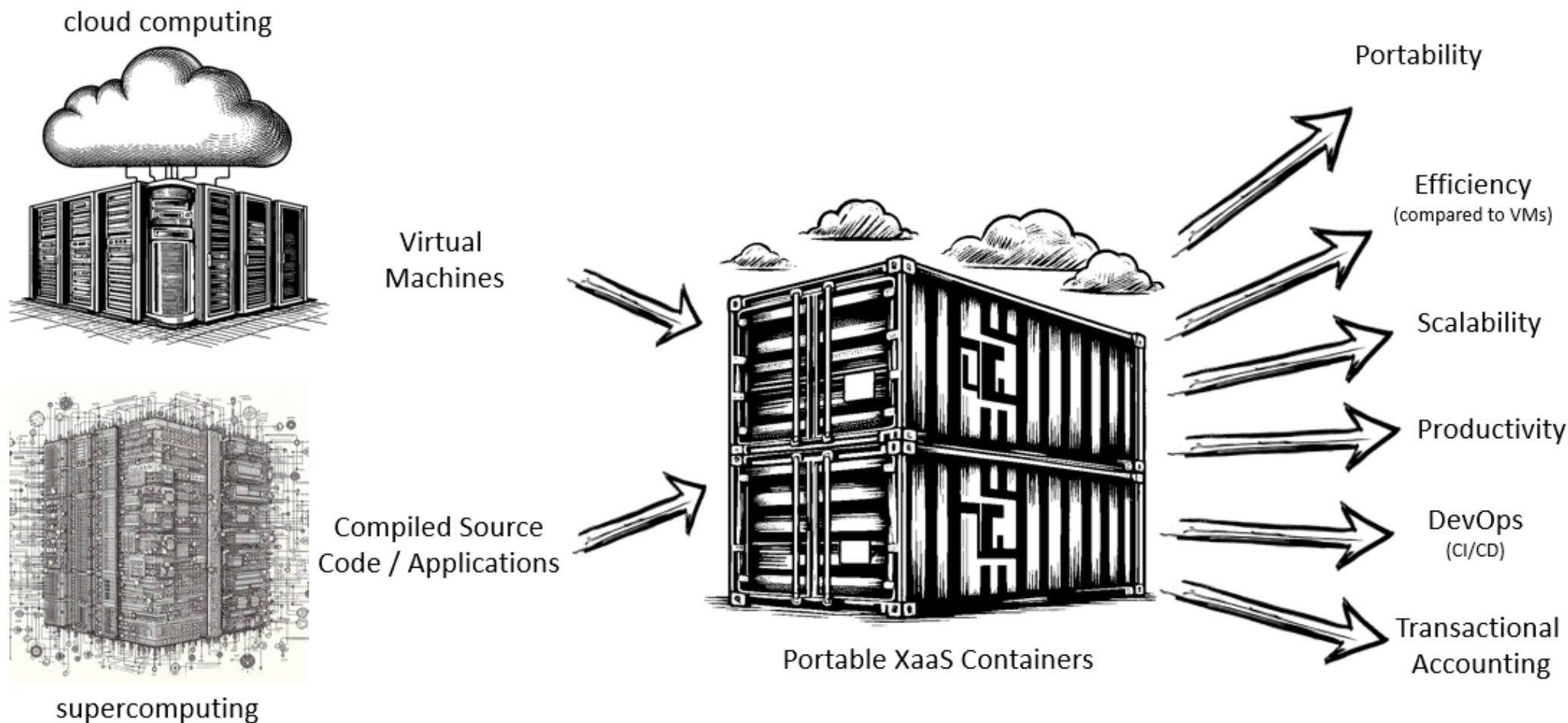# Performance Portable Containers

cloud computing

Virtual
Machines

Compiled Source
Code / Applications

supercomputing

Performance Portable Containers
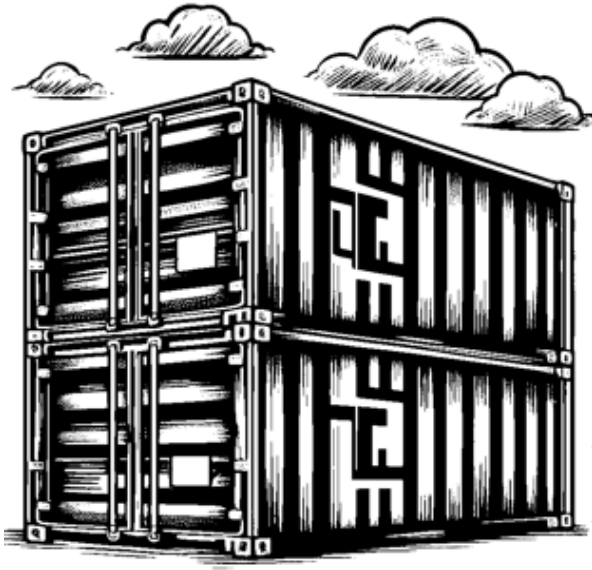
Performance Portable Containers

cloud computing

Virtual Machines

Compiled Source Code / Applications

supercomputing

Portable XaaS Containers

Portability

Efficiency (compared to VMs)

Scalability

Productivity

DevOps (CI/CD)
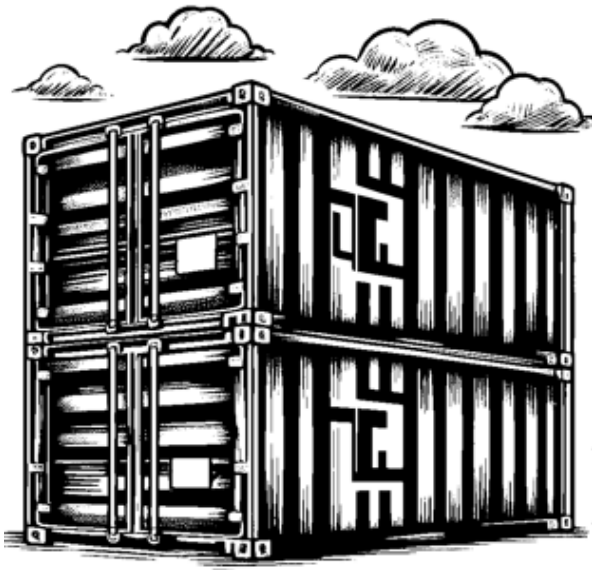
Transactional Accounting

Three Building Blocks of XaaS

Portable XaaS Containers

Three Building Blocks of XaaS

Portable XaaS Containers

High-Performance I/O

- Data storage managed by system provider
- Communication between XaaS containers
- Specialization to network fabric

Three Building Blocks of XaaS

Portable XaaS Containers

High-Performance I/O

- Data storage managed by system provider
- Communication between XaaS containers
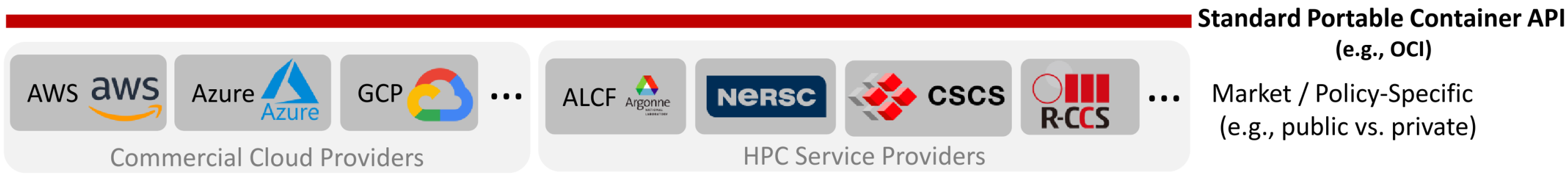- Specialization to network fabric

Scheduling & Invocations

- Flexible scheduling to reduce wait times
- Balance between batch and interactive workloads
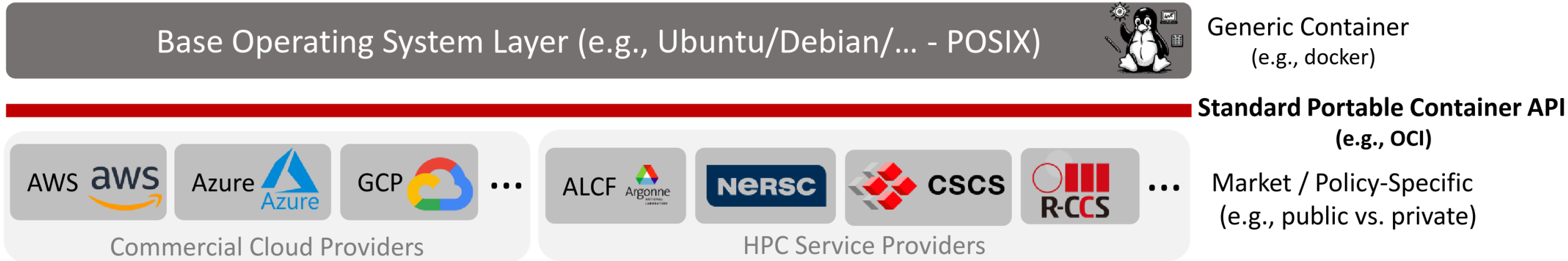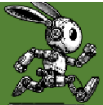- High-level control plane API, e.g., REST

XaaS System Architecture

Commercial Cloud Providers: AWS, Azure, GCP ...

HPC Service Providers: ALCF (Argonne National Laboratory), NeRSC, CSCS, R-CCS ...

Market / Policy-Specific (e.g., public vs. private)

XaaS System Architecture

Standard Portable Container API
(e.g., OCI)

Market / Policy-Specific
(e.g., public vs. private)

AWS | Azure | GCP | •••

Commercial Cloud Providers

ALCF | NeRSC | CSCS | R-CCS | •••

HPC Service Providers

XaaS System Architecture

Base Operating System Layer (e.g., Ubuntu/Debian/… - POSIX)

Generic Container
(e.g., docker)

Standard Portable Container API
(e.g., OCI)

AWS | Azure | GCP | •••

Commercial Cloud Providers

ALCF | NeRSC | CSCS | R-CCS | •••

HPC Service Providers

Market / Policy-Specific
(e.g., public vs. private)

XaaS System Architecture

| XaaS Accelerated Compute, Communication, I/O libraries, and APIs | System-Specific (e.g., HPE, Intel) |

| Base Operating System Layer (e.g., Ubuntu/Debian/... - POSIX) | Generic Container (e.g., docker) |

Standard Portable Container API
(e.g., OCI)

Commercial Cloud Providers — HPC Service Providers
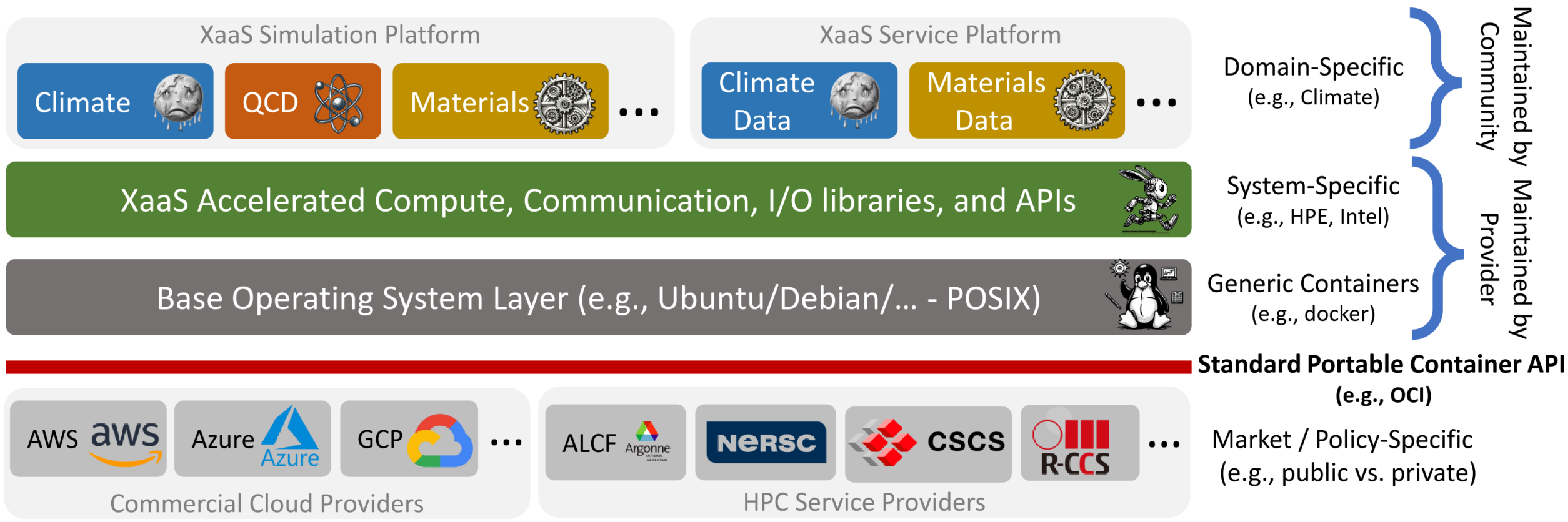
Market / Policy-Specific (e.g., public vs. private)

XaaS System Architecture

XaaS System Architecture

XaaS in Practice
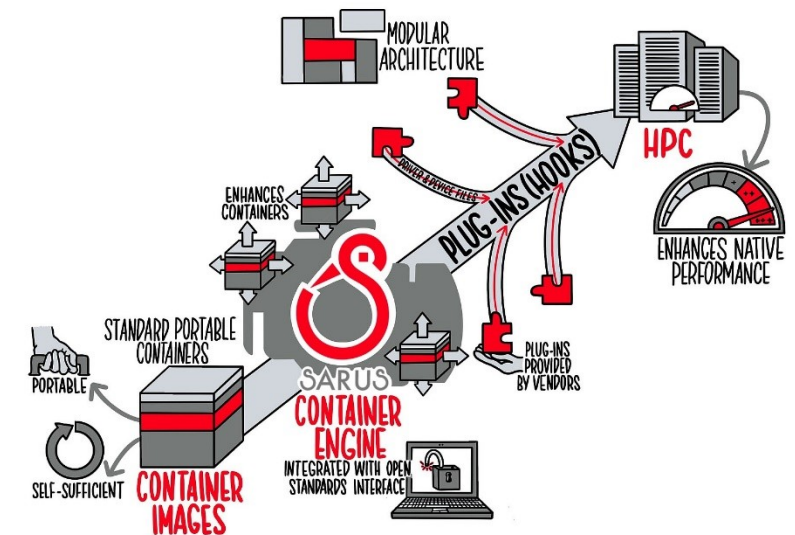
Spack

Full Build & Specialization

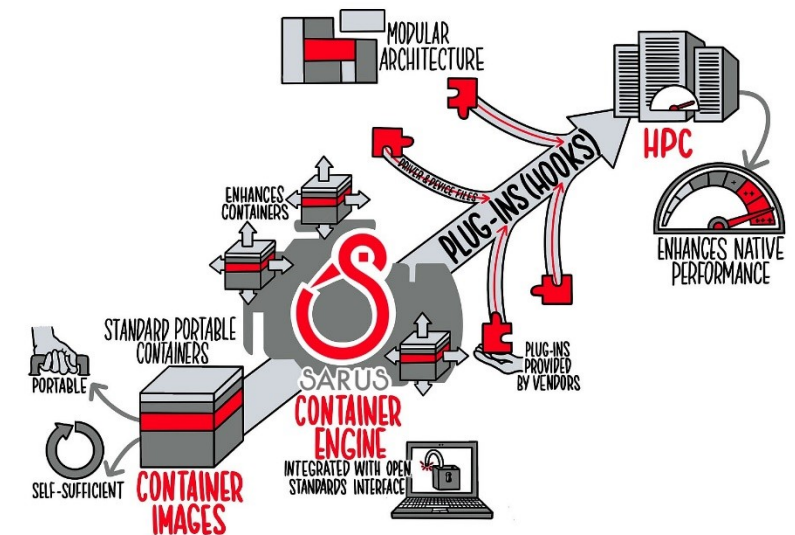# XaaS in Practice

Spack

Full Build & Specialization

Flexible Library Hooks

XaaS in Practice

Spack

Full Build & Specialization

More Productivity

Flexible Library Hooks
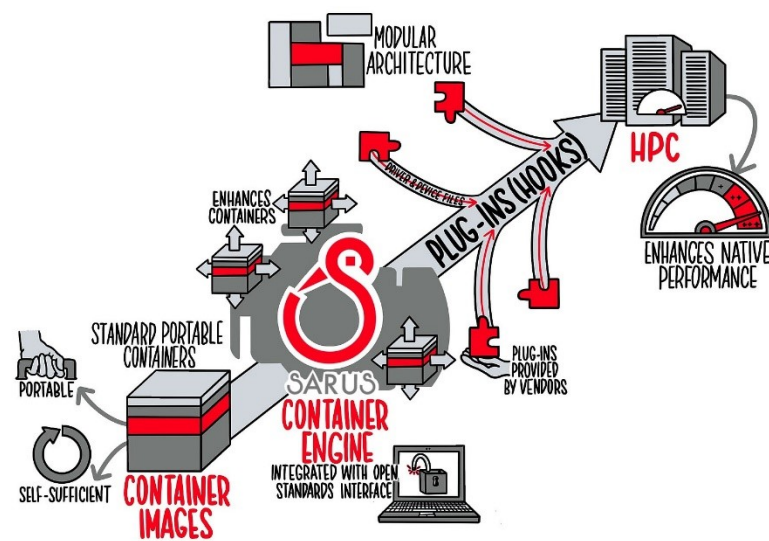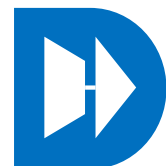
XaaS in Practice

Full Build & Specialization

More Specialization

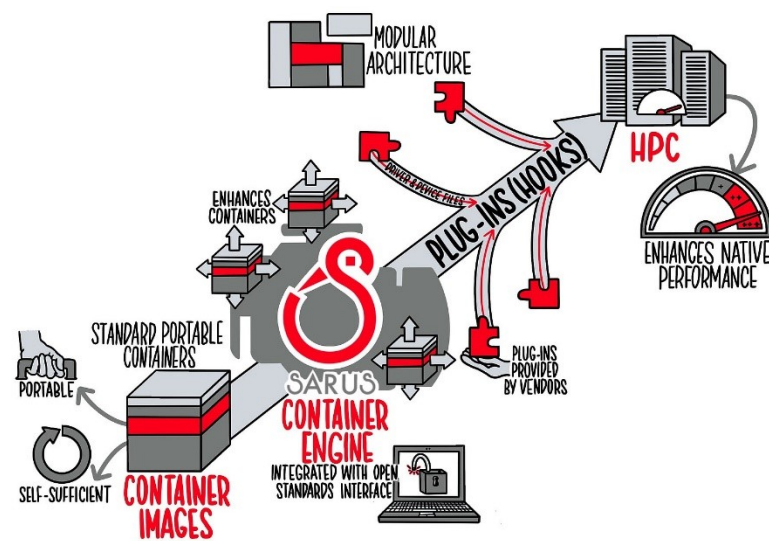More Productivity

Flexible Library Hooks

XaaS in Practice

Spack

Full Build & Specialization

DaCe

Partial Recompilation

More Specialization
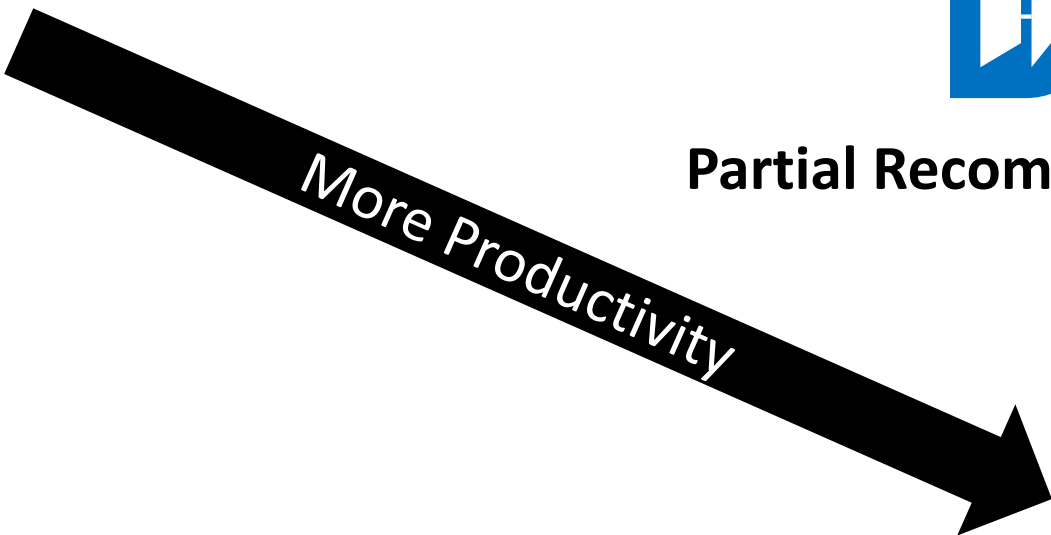
More Productivity

Flexible Library Hooks

What XaaS Can Offer Us?

What XaaS Can Offer Us?

Cloud
Productivity

What XaaS Can Offer Us?

Cloud
Productivity

Performance
Portability

What XaaS Can Offer Us?

Cloud Productivity

Performance Portability

Unified Interface

What XaaS Can Offer Us?

Cloud Productivity

Performance Portability

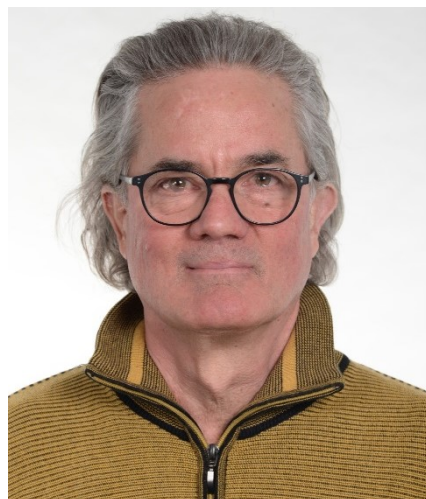**Unified Interface**

Flexible Scheduling

Q&A

Dan Ernst
NVIDIA

Ian Foster
Argonne National Laboratory

Torsten Hoefler
ETH Zurich

Thomas Schulthess
Swiss National Supercomputing
Centre (CSCS)

Marcin Copik
ETH Zurich