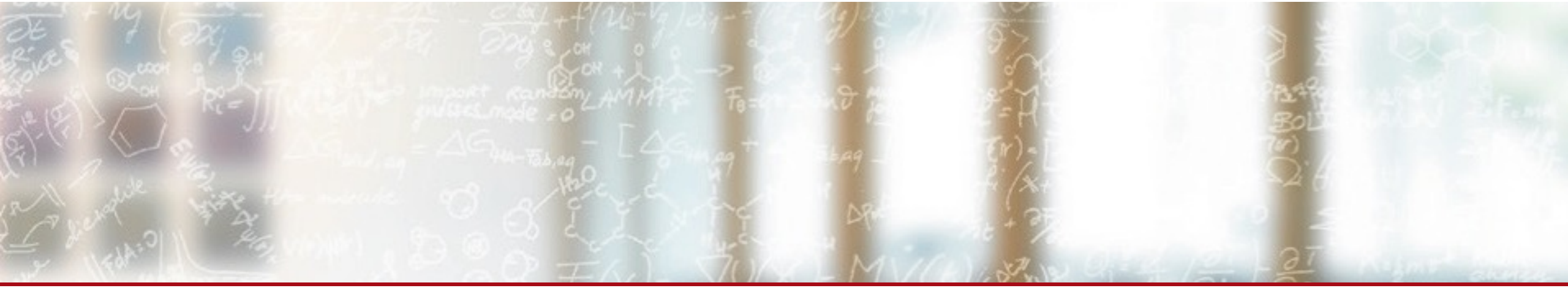
CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

ETH zürich



The future of extreme compute and data

Swiss SKA Days 2022

Maxime Martinasso, CSCS

Oct 3rd, 2022

Extreme compute

- Today's world compute capability?
  - EFlops sum of top 10 systems (~0.5 EFlops/year):
    - Jun 2022 – 2.25 EFlops
    - Nov 2018 – 0.5 EFlops
  - Not listed machines: how many exaflops machines in China or in the industry?

- Technology
  - We are in the silicon era with the use of accelerators
  - Changing era?
    - Nanotechnology, carbon nanotubes instead of silicon transistor
    - Photonics, photons instead of electrons
    - When for large-scale manufacturing?
  - Quantum computing: might be a specific accelerator?

Extreme compute challenges

For most countries sustaining a unique 1-Eflops machine is today's limit

- Power efficiency barrier
  - 1 Eflops ~= 20 MW
    - Large hydropower plant: ~30-50 MW
    - Nuclear plant: ~500-1000 MW
  - Reduce power needs, moving large HPC systems to colder countries
- Money challenge
  - Frontier 1.1 EFlops = $600 Million to build
  - Operational cost over the supercomputer lifetime – 20-30% of building cost
  - Variability of power cost
- Usability and programming
  - Build programming abstraction in scientific software
    - DSL / specialized libraries / hardware-specific back-end
    - Refactoring existing codes takes many years of development
  - Performance engineering

Extreme data

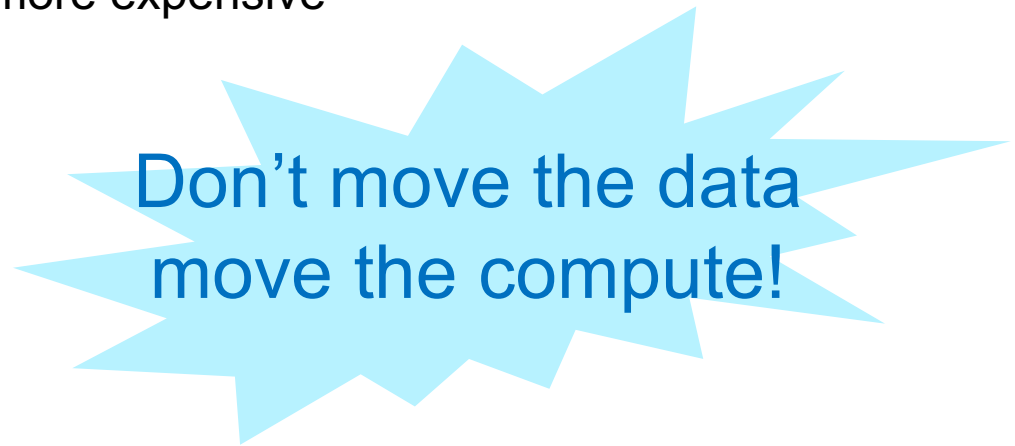
- Scientific data explosion
  - New scientific instrument, more precise
  - International collaboration, distributed instruments, aggregation of data
  - Scientific challenges for finer resolution simulations

- Example of dataset sizes

Insitution	Scientific domain	Workflows	Data size
European Centre for Medium-Range Weather Forecasts (ECMWF)	Weather Forecast	Ensemble forecasts, data assimilation,...	12 PB/year
Paul Scherrer Institute (PSI)	Synchrotron imaging	X-ray spectroscopy, high resolution microscopy,...	10-20 PB/year
CERN	High-energy particle collider	LHC Run 2	1PB/day

Extreme data challenges

- Technology
  - Transfer rate: still orders of magnitude lower than needed
    - 120 TB over internet = 111 days
    - 120 TB sent as discs by Fedex = 1 day
    - Need to build dedicated link with multiple 100Gbps to connect data producers to consumers
    - Transfer only what is relevant for the consumers
  - Storing/presisting PB of data
    - Tapes is still the solution, poor access time
    - Keeping data accessible in near-compute storage is more expensive



Versatile software-defined cluster (vCluster)

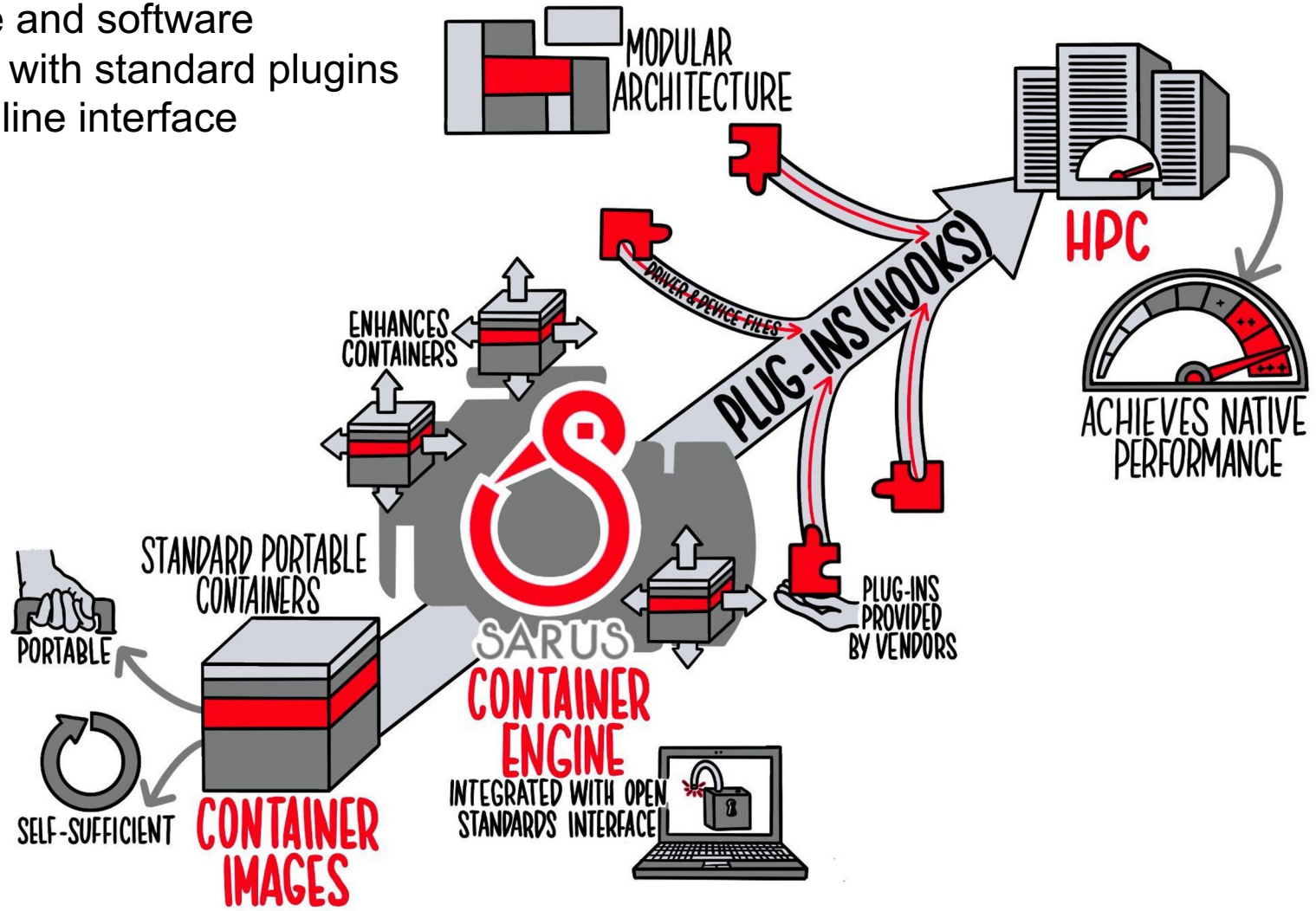
In development

- Deploy an identical platform on several system technologies
  - The platform is defined as a set of config files
    - Each service is a config file (IAM, batch scheduling, data managers, accounting,…)
  - Abstraction layer to be system independant
  - Provide an engine to deploy such software-defined platforms
  - Use Infrastructure-as-Code to implement specific instance of the engine back-end
- Transfer scientific software stack and ensure performance
  - Use containers to store the software stack
  - Use OCI Hooks to expose host HPC-capable stack in the container

- Enable programmatic and web access to the platform
  - Create a RESTful API to access services on the machine
  - Build community portal / scientific workflows using that API

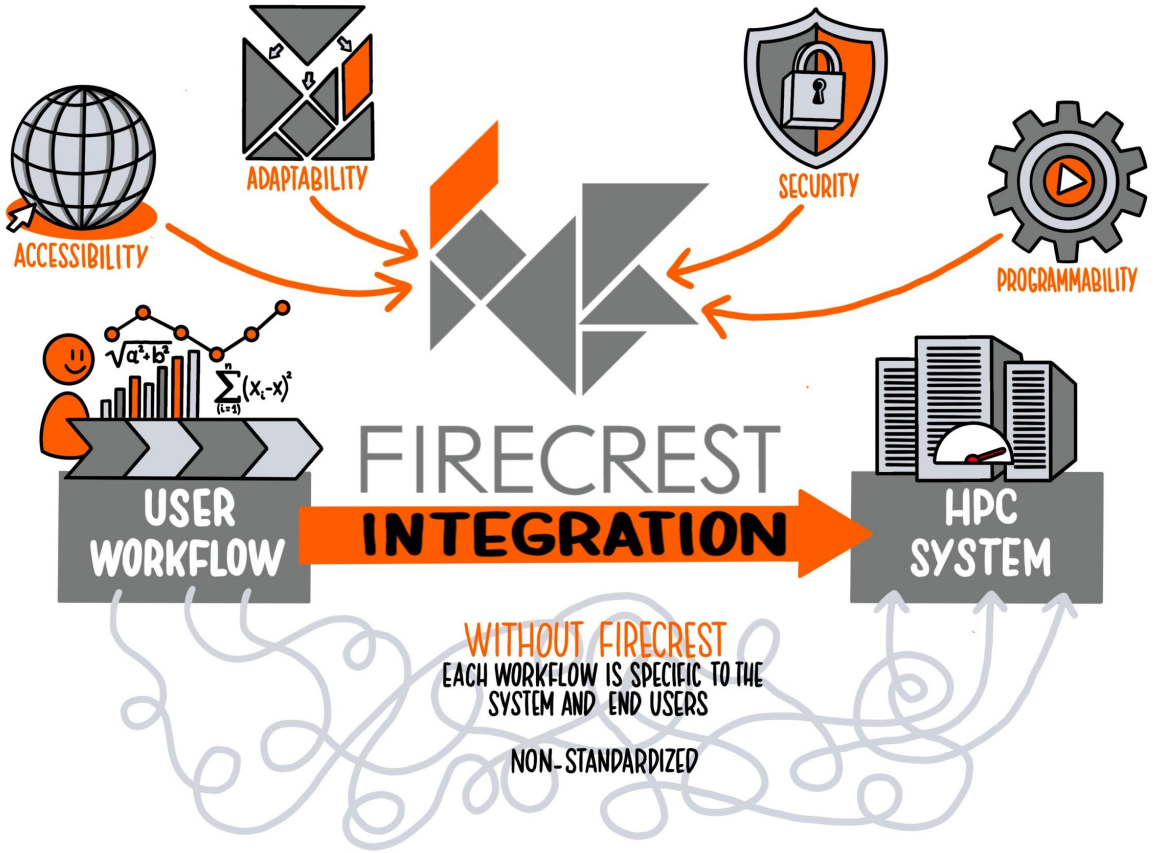
Sarus concept

- Combines container portability with native HPC performance
- Integrates with HPC infrastructure and software
- Customizes containers at runtime with standard plugins
- Provides a Docker-like command line interface



FirecREST concepts

- Accessibility: opens web access to HPC from any device
- Adaptability: modular design to support diverse HPC ecosystems
- Security: provides multiple authorization control layers
- Programmability: uses standard interface for a simplified automation

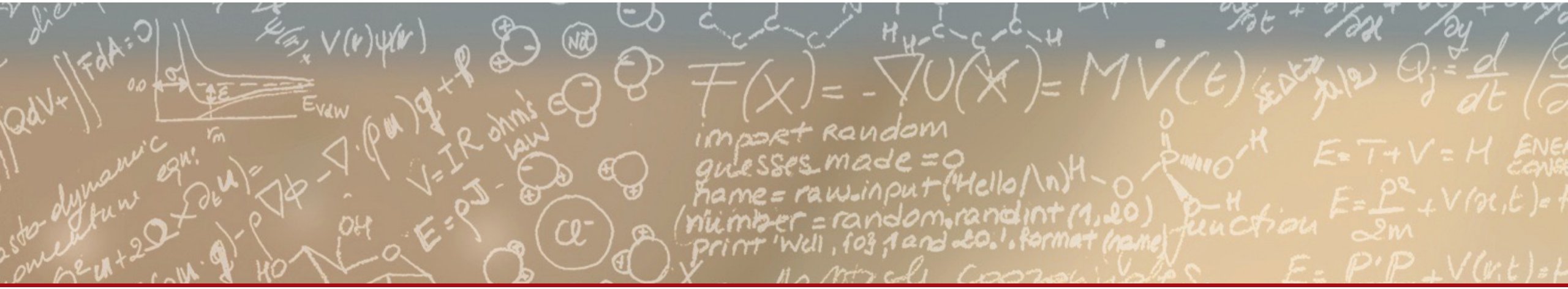


Next steps

- CSCS is commited in the development of the vCluster solution
  - vCluster v1.0, Sarus v2.0, FirecREST v2.0
  - Alps will exclusively use the vCluster technology

- Identify use cases and partners
  - With EPFL/SCITAS looking into the abstraction layer of the vCluster
  - With PSI enhancing vCluster to enable IaaS, tenant-managed vCluster

- Promote and disseminate the vCluster technology
  - Find new use cases and opportunities



Thank you for your attention.