High performance computing and data management requirements related to computer simulations

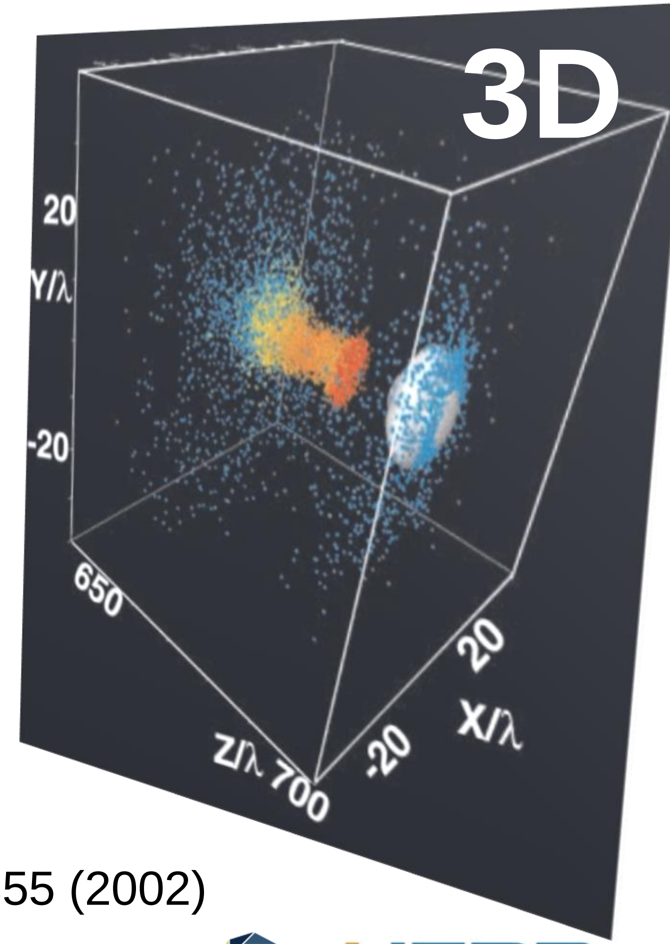
Michael Bussmann

HZDR
HELMHOLTZ
ZENTRUM DRESDEN
ROSSENDORF

Mitglied der Helmholtz-Gemeinschaft

The role of simulations in plasma-based accelerators

- Inventing new schemes for acceleration, injection and control of beam
- Post-hoc analysis of experimental findings
- Parameter studies
- Quantitative Predictions of Beam Parameters

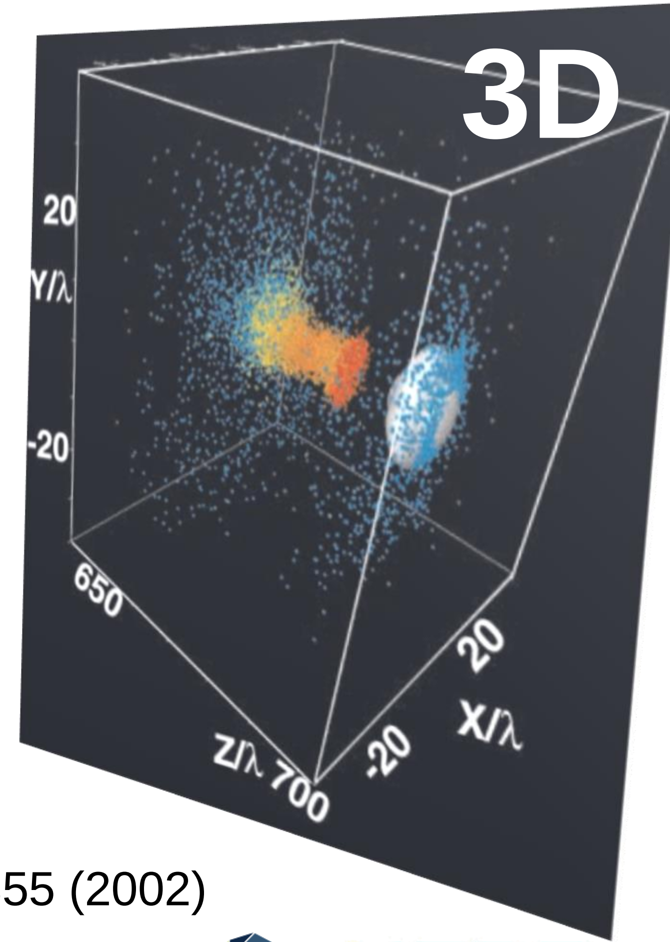


A. Pukhov, J. Meyer-Ter-Vehn, Appl. Phys B 74, 355 (2002)

Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de

The role of simulations in plasma-based accelerators

- Inventing new schemes for acceleration, injection and control of beam
- Post-hoc analysis of experimental findings
- Parameter studies
- Quantitative Predictions of Beam Parameters



A. Pukhov, J. Meyer-Ter-Vehn, Appl. Phys B 74, 355 (2002)

Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

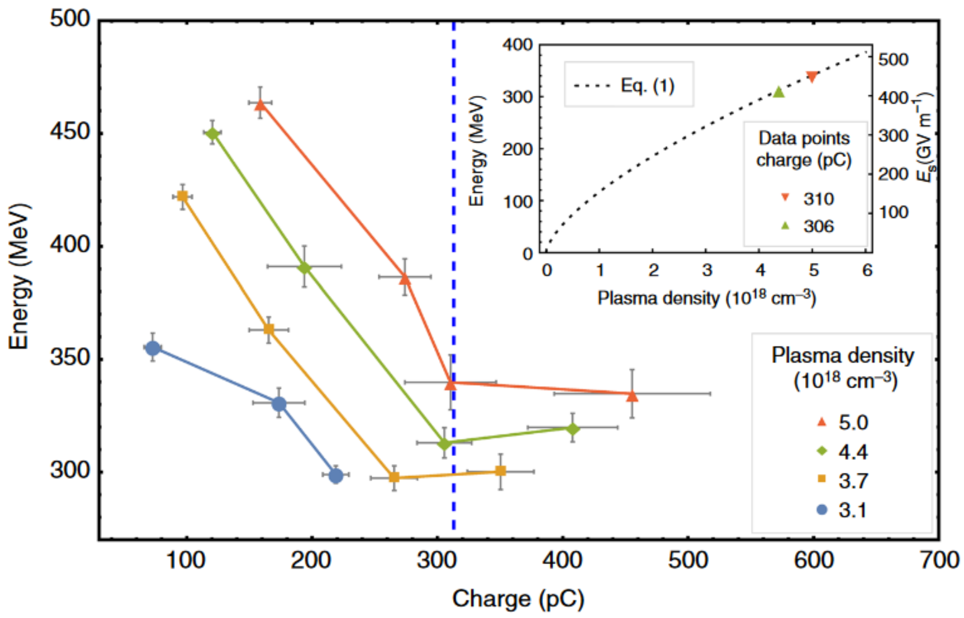
Do we want & need predictive capabilities?

- Finding a range of parameters for good, stable operation is doable
- For prospective studies, you need some assurance that the code is o.k.
- For planning applications, you should know your code is o.k.

J.P. Couperous et al.,
Nat. Comm. 8(1), 487 (2017)

S. Katsouleas et al.,
Part. Accel. 22, 81 (1987)

M. Tzoufras et al.,
PRL 101,145002 (2008)



Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

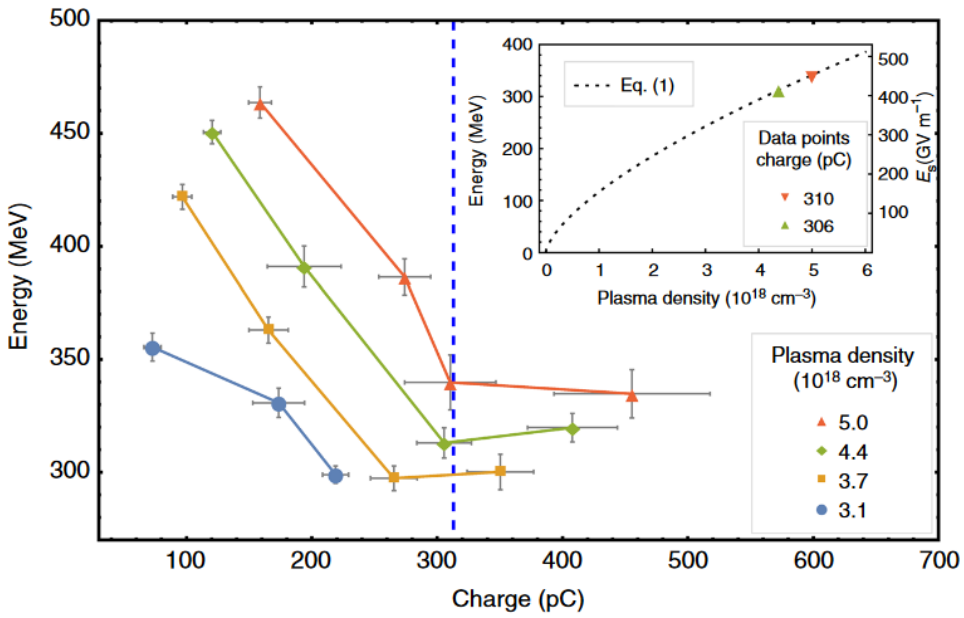
Do we want & need predictive capabilities?

- Finding a range of parameters for good, stable operation is doable
- For prospective studies, you need some assurance that the code is o.k.
- For planning applications, you should know your code is o.k.

J.P. Couperous et al.,
Nat. Comm. 8(1), 487 (2017)

S. Katsouleas et al.,
Part. Accel. 22, 81 (1987)

M. Tzoufras et al.,
PRL 101,145002 (2008)



Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

We need and have a broad selection of tricks & tools (not just PIC!)

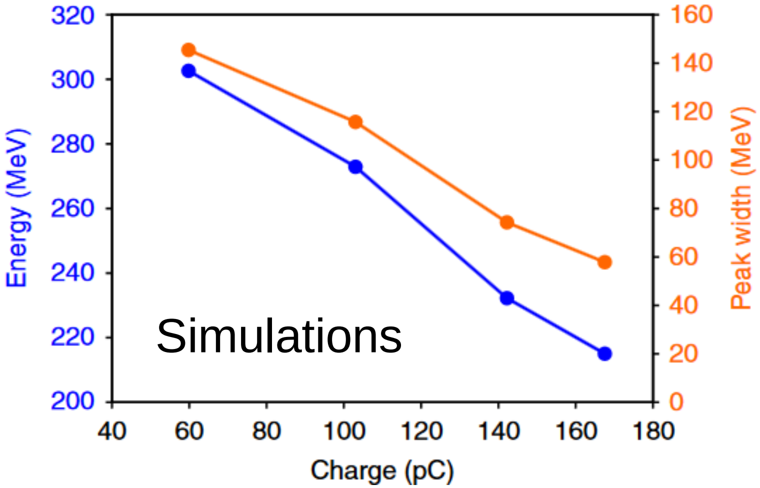
- Pen, paper, intuition & experience
- Hybrid + mixed fluid / kinetic schemes, electrostatic, MD/MC, etc.
- Reduced dimensions, boosted frames, ML / reduced models

Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

Why predictive simulations are hard

- Key experimental parameters are not well known & hard to reproduce
- Variations in input parameters (e.g. initial phase space distribution) & choice of models (e.g. Maxwell Solvers) influence results
- Simulations predict fundamental plasma properties (e.g. field strength & density) that are hard to measure directly
- Vary schemes + parameter ranges

J.P. Couperous et al.,
Nat. Comm. 8(1), 487 (2017))
80 3D Simulations and a broken filter



Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

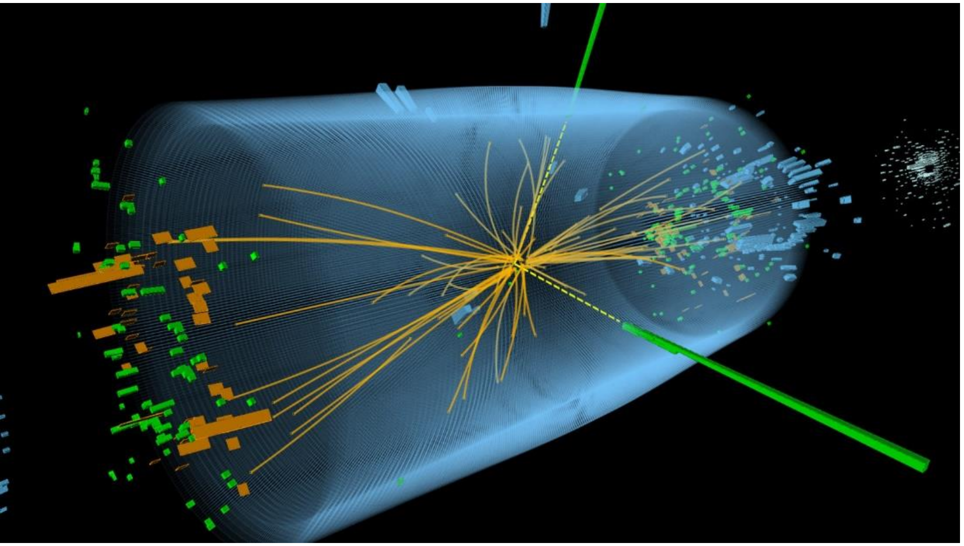
Don‘t trust ANY code (especially the one written by yourself!)

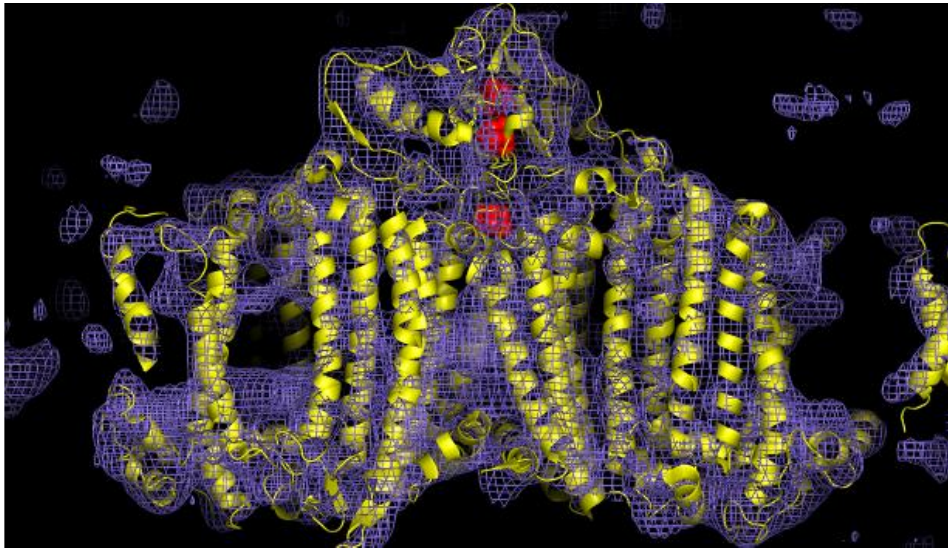
- Not all methods are implemented in every code
- Some methods cannot be implemented together
- Some methods interact spuriously with others
- Some methods have limited applicability
- Some methods are approximate
- Some methods are hard to parallelize

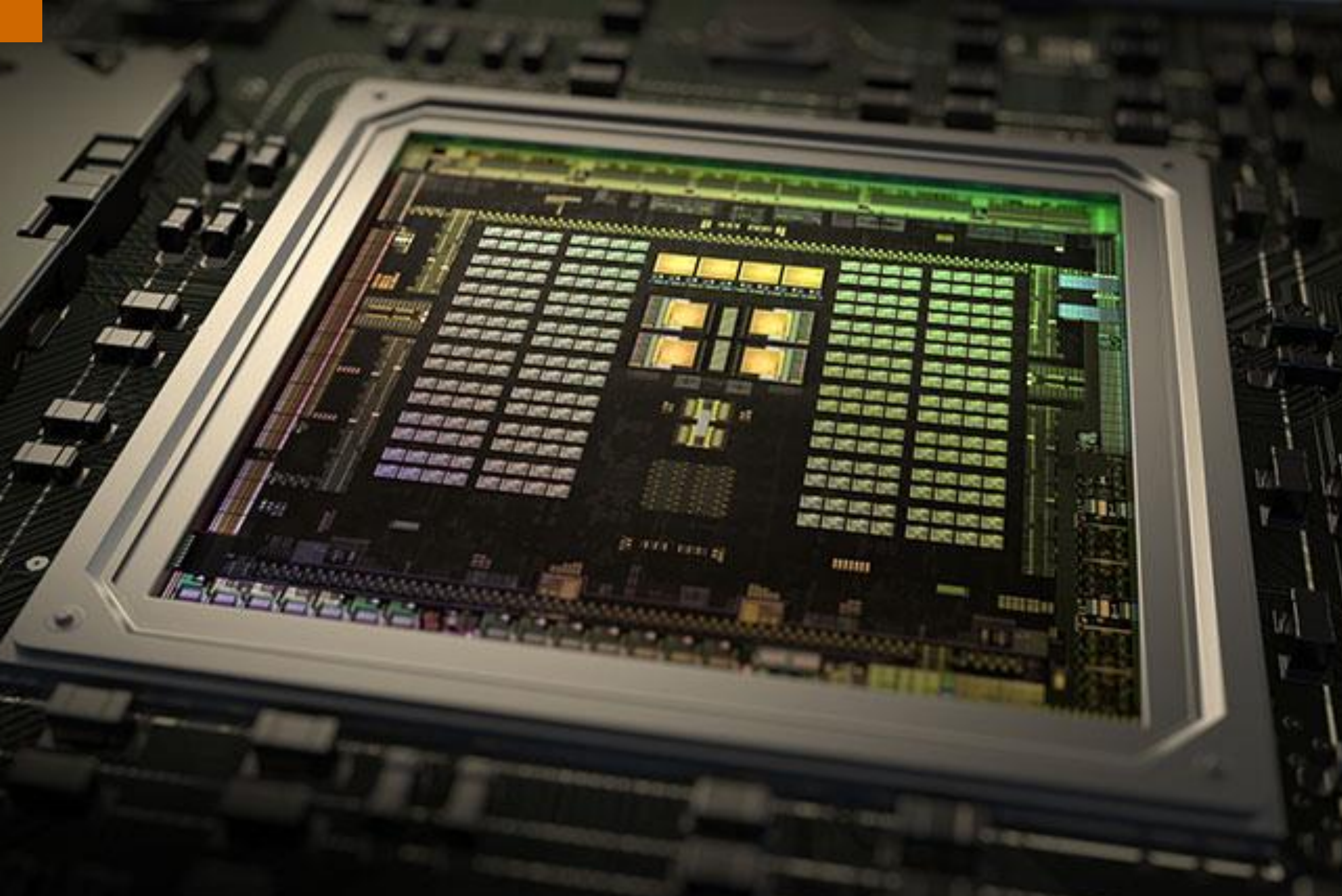
Parameters are defined by applications & users!

- Accelerators must work routinely and reliably
- Accelerators must fulfill user‘s needs

ENGINEERING!







How wrong are we?

Simulation Codes today (totally exaggerated!)

Simulation

Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

Start to End Simulations

EuPRAXIA

NOVEL FUNDAMENTAL RESEARCH

COMPACT EUROPEAN PLASMA ACCELERATOR WITH SUPERIOR BEAM QUALITY

Simulations on 10 different schemes

Start to End simulations

Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

Towards Exascale Computing (Capability + Capacity)

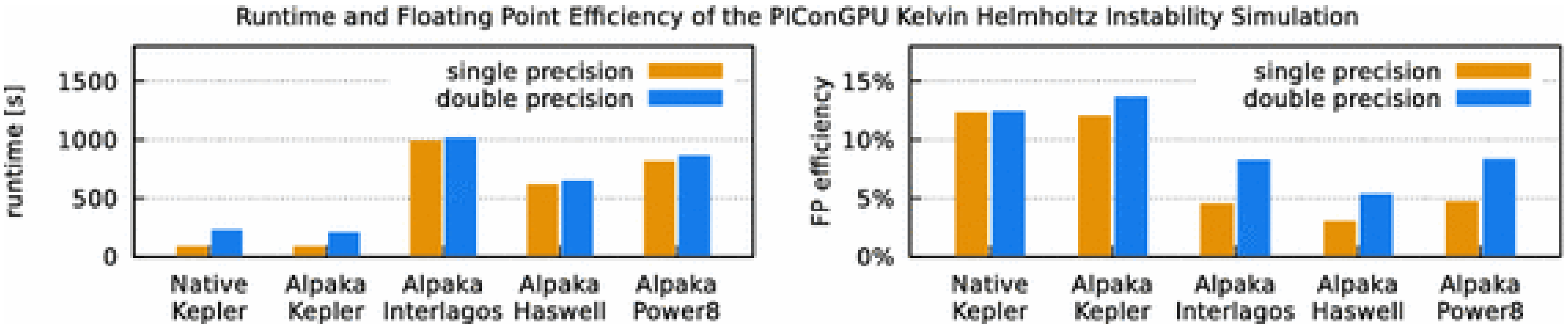
Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	**Summit** - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM DOE/SC/Oak Ridge National Laboratory United States	2,282,544	122,300.0	187,659.3	8,806
2	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
3	**Sierra** - IBM Power System S922LC, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM DOE/NNSA/LLNL United States	1,572,480	71,610.0	119,193.6	
4	**Tianhe-2A** - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000 , NUDT National Super Computer Center in Guangzhou China	4,981,760	61,444.5	100,678.7	18,482
5	**AI Bridging Cloud Infrastructure (ABCI)** - PRIMERGY CX2550 M4, Xeon Gold 6148 20C 2.4GHz, NVIDIA Tesla V100 SXM2, Infiniband EDR , Fujitsu National Institute of Advanced Industrial Science and Technology (AIST) Japan	391,680	19,880.0	32,576.6	1,649
6	**Piz Daint** - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 , Cray Inc. Swiss National Supercomputing Centre (CSCS) Switzerland	361,760	19,590.0	25,326.3	2,272
7	**Titan** - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x , Cray Inc. DOE/SC/Oak Ridge National Laboratory United States	560,640	17,590.0	27,112.5	8,209
8	**Sequoia** - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom , IBM DOE/NNSA/LLNL United States	1,572,864	17,173.2	20,132.7	7,890

```
#ifdef CUDA_ENABLE
   // CUDA Kernel implementation
   // …

#elif OPENMP_ENABLE
   // OpenMP implementation
   // …

#else
   // Sequential CPU implementation
   // …

#endif
```

Start to End simulations

Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

Towards Exascale Computing (Capability + Capacity)



E. Zenker et al., Lecture Notes in Computer Science 9945, 293 (2002)

PIConGPU

Central Algorithms

Domain Specific Language

Performance Portability Layer

Scalability (Load Balancing)

Start to End simulations

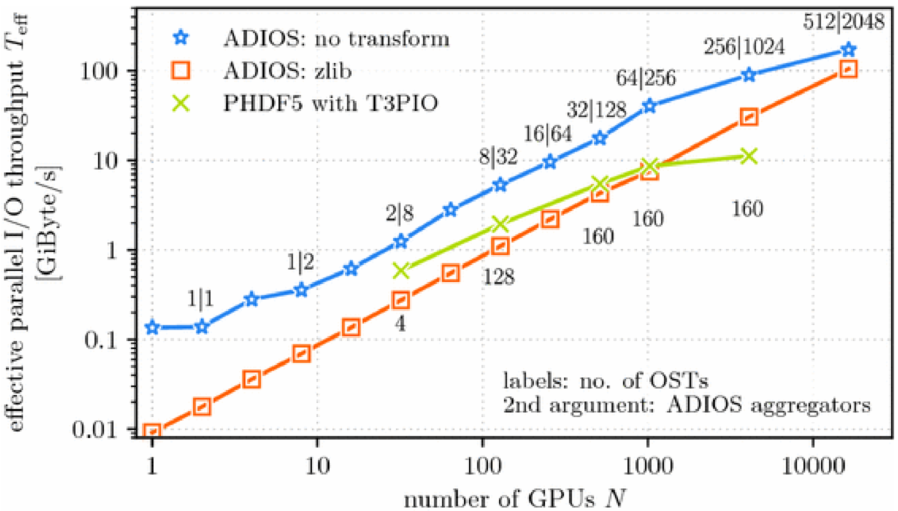
Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de

Mitglied der Helmholtz-Gemeinschaft

Exascale I/O

"The TNG simulations produced more than 500 Terabyte […] The full analysis will keep the participating scientists busy for many years to come"

A. Huebl et al., High Performance Computing 10524, 15 (2017)



I/O

*"Overall , this is an outstanding proposal. The HPC resource request are appropriate. **The PIs should try to reduce the data requirements and try to find a solution that is technically possible for CSCS**."*

Central Algorithms

Domain Specific Language

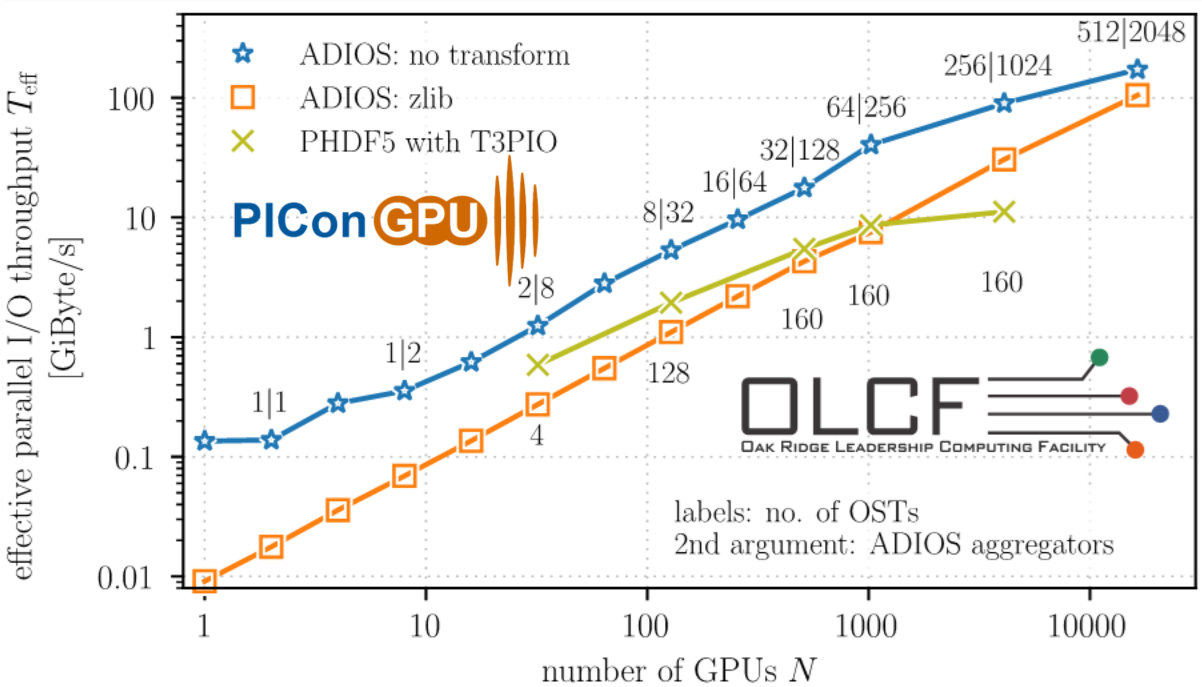
Performance Portability Layer

Scalability (Load Balancing)

Start to End simulations

Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

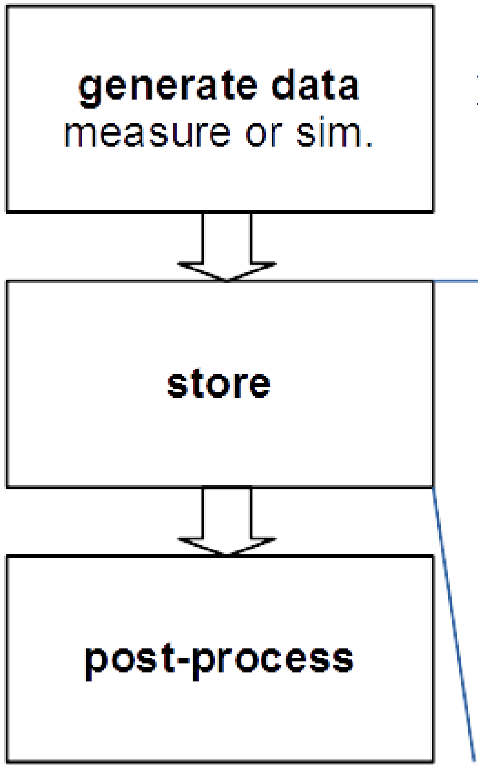
Exascale I/O for everybody



- 1 TB/s theoretical throughput
- 250 GB/s reached
- 14 MB/s per node
- Exclusive access
- How is it accesed?
- How long may it live?
- 1 Pb I/O in 1 hour

Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

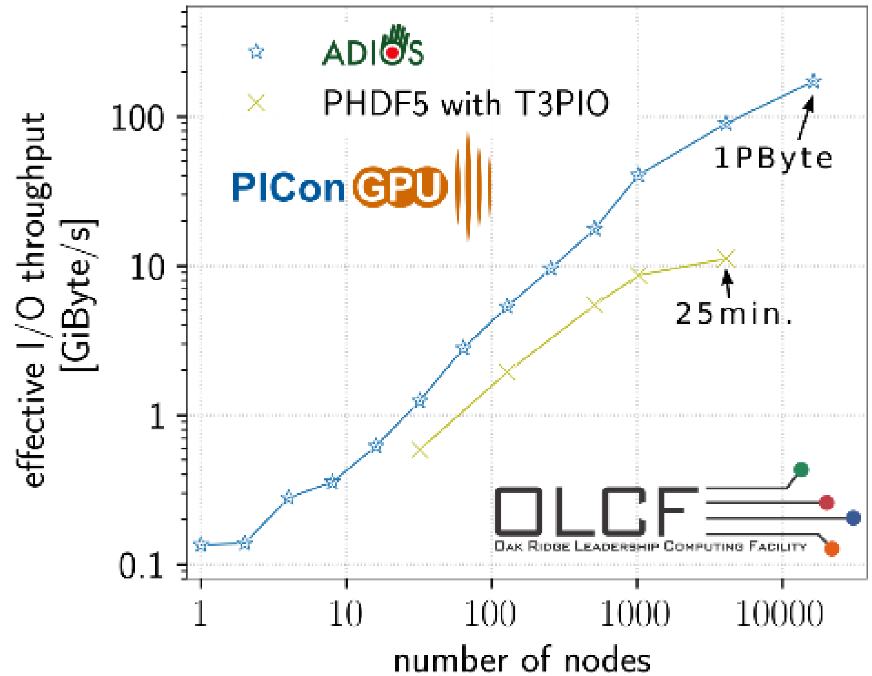
Writing to disk won‘t work anymore



K20x GPU (2013), **bw := 6 GByte · 10 Hz steps**

PCI-Express: **6 Gbyte/s**
1 / 10 th bw

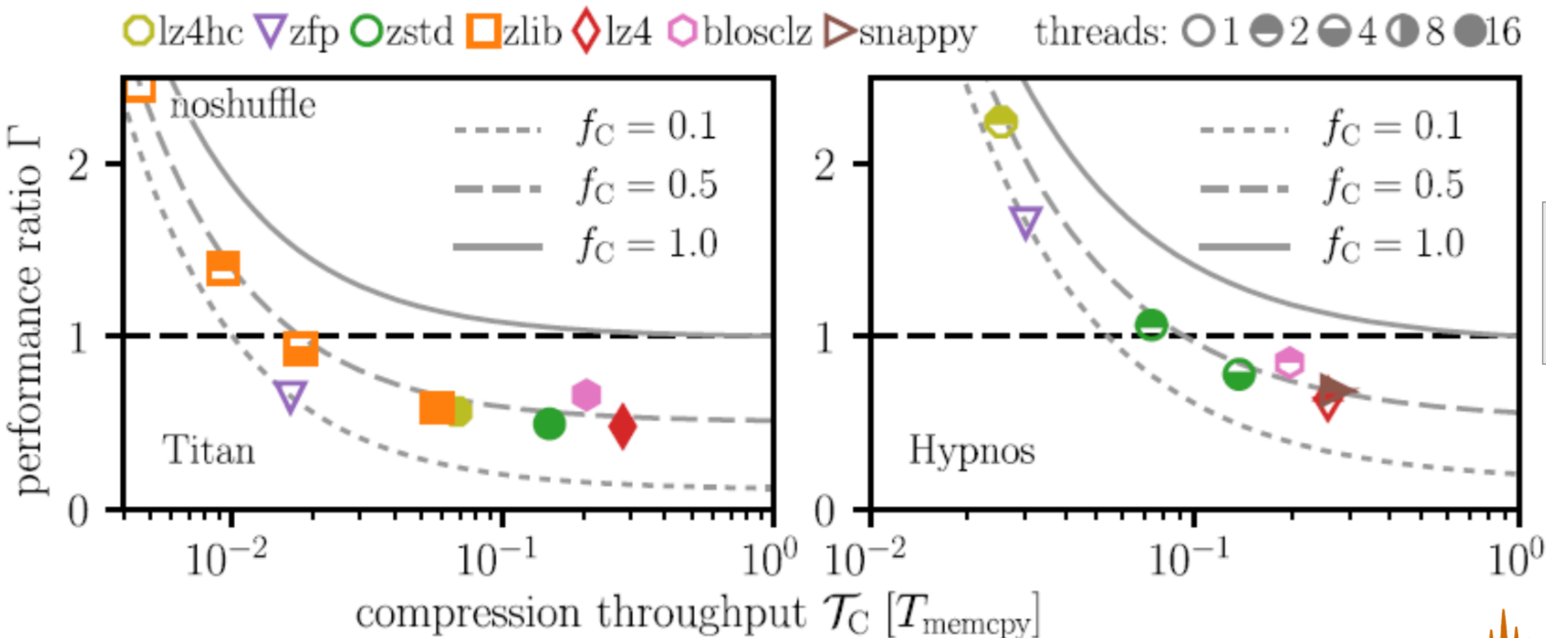
I/O per node:
1 / 10 th · 1 / 200 th bw
– 42 MByte/s (Titan)
– 29 MByte/s (PizDaint)



Summit (ORNL, 2018): ratio **4x “worse” - gap of 10^4**

Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de

Data reduction won‘t help, except when you throw stuff away



$$\Gamma \equiv \frac{t_{\text{I/O}}^{\text{reduced}}}{t_{\text{I/O}}^{\text{unreduced}}}$$

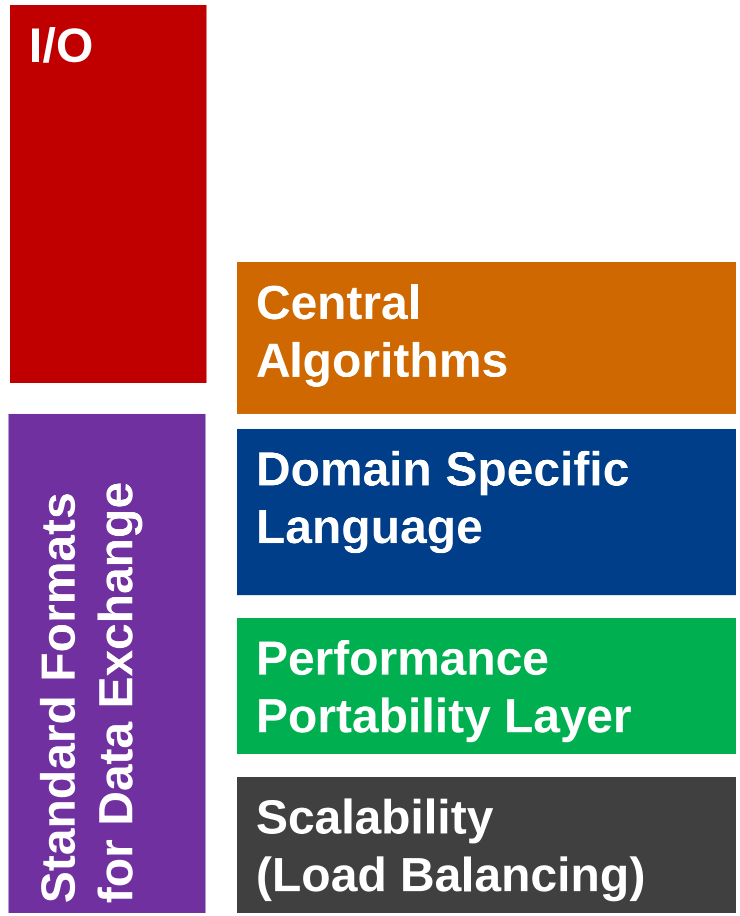
PIConGPU

Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de

Open, self-descriptive data exchange between applications

Share data between

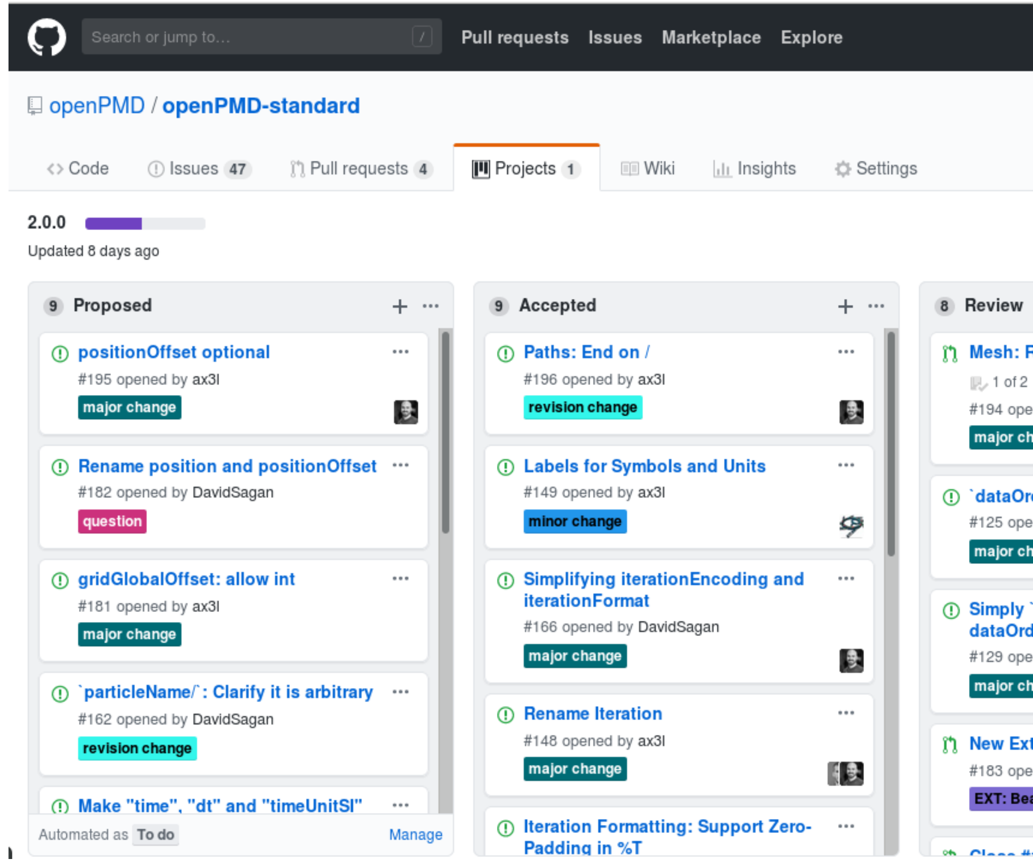
- PIC Codes
- Accelerator Codes
- FEL Codes
- Analysis Tools
- Visualization Tools
- …



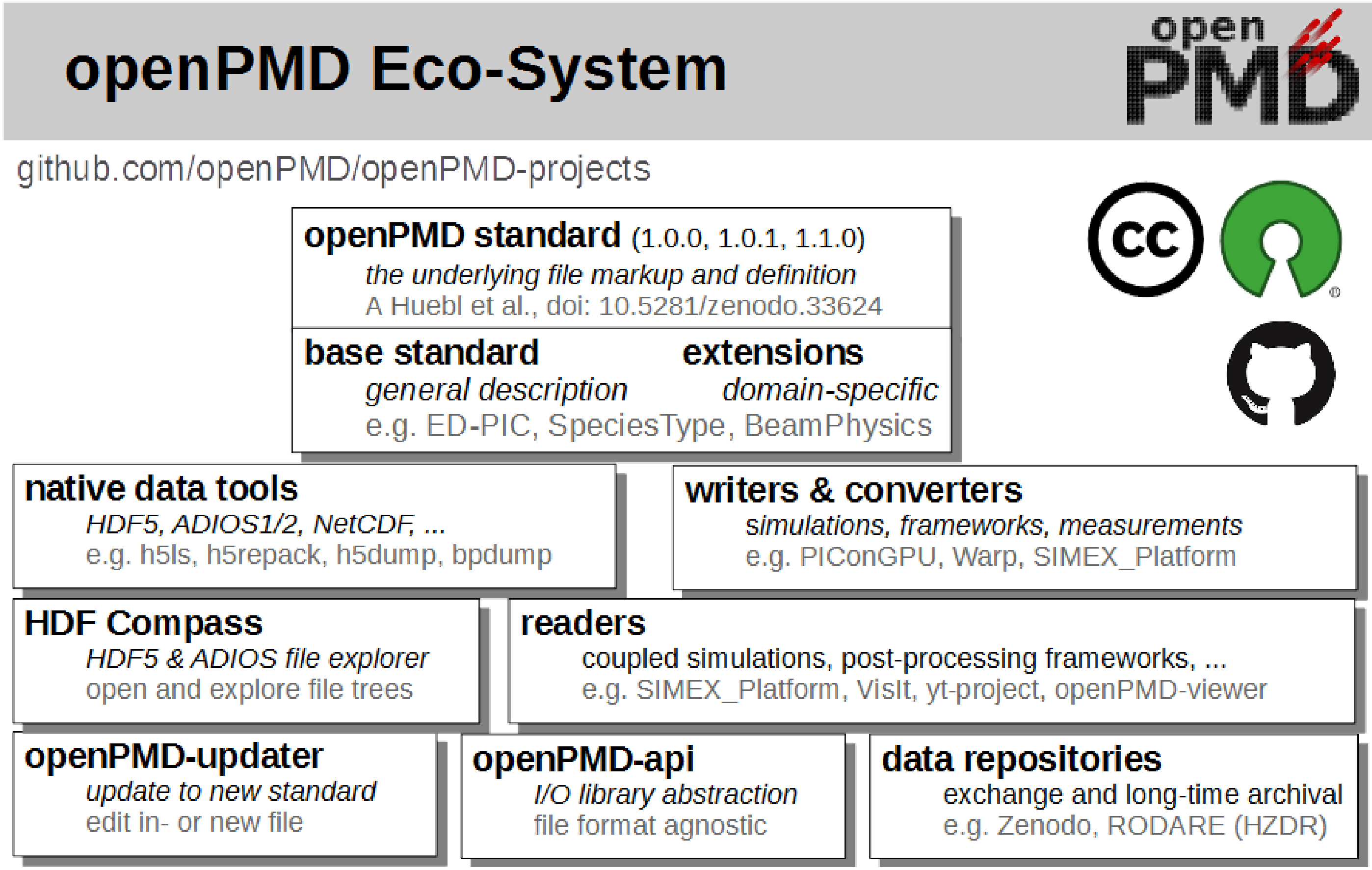
Start to End simulations

Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de

Open Source META DATA format for Plasma simulation codes



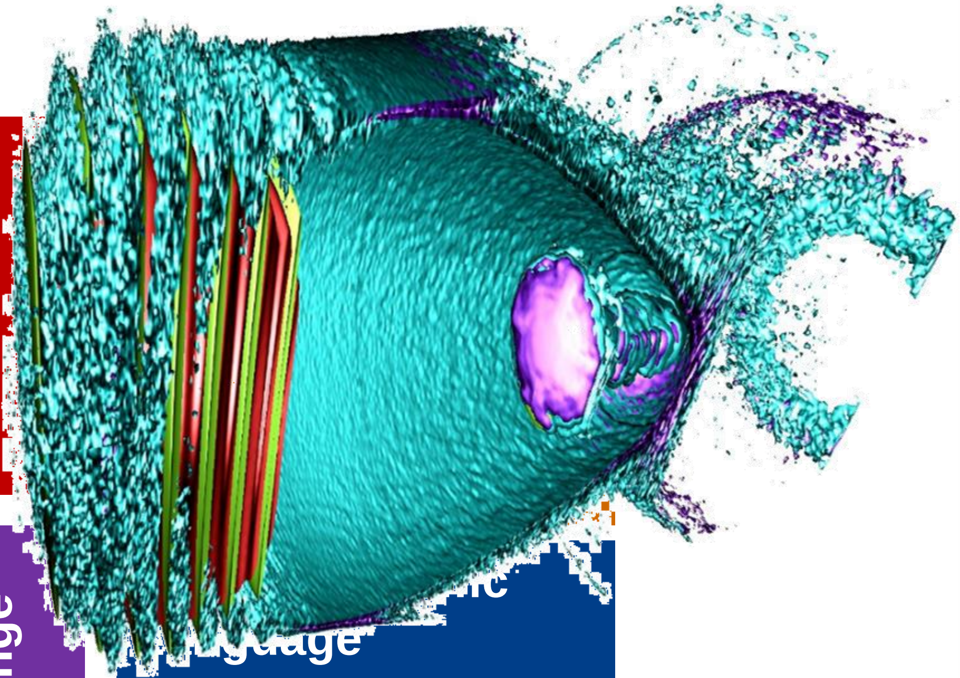
An Ecosystem of Tools, soon including Exascale I/O Capabilities



Scientific Workflows, Online Analysis & Visualization, SaaS, …

Memory Staging & Workflows

I/O



ISAAC

Offline + Online Analysis

In-situ Visualization

Standard Formats for Data Exchange

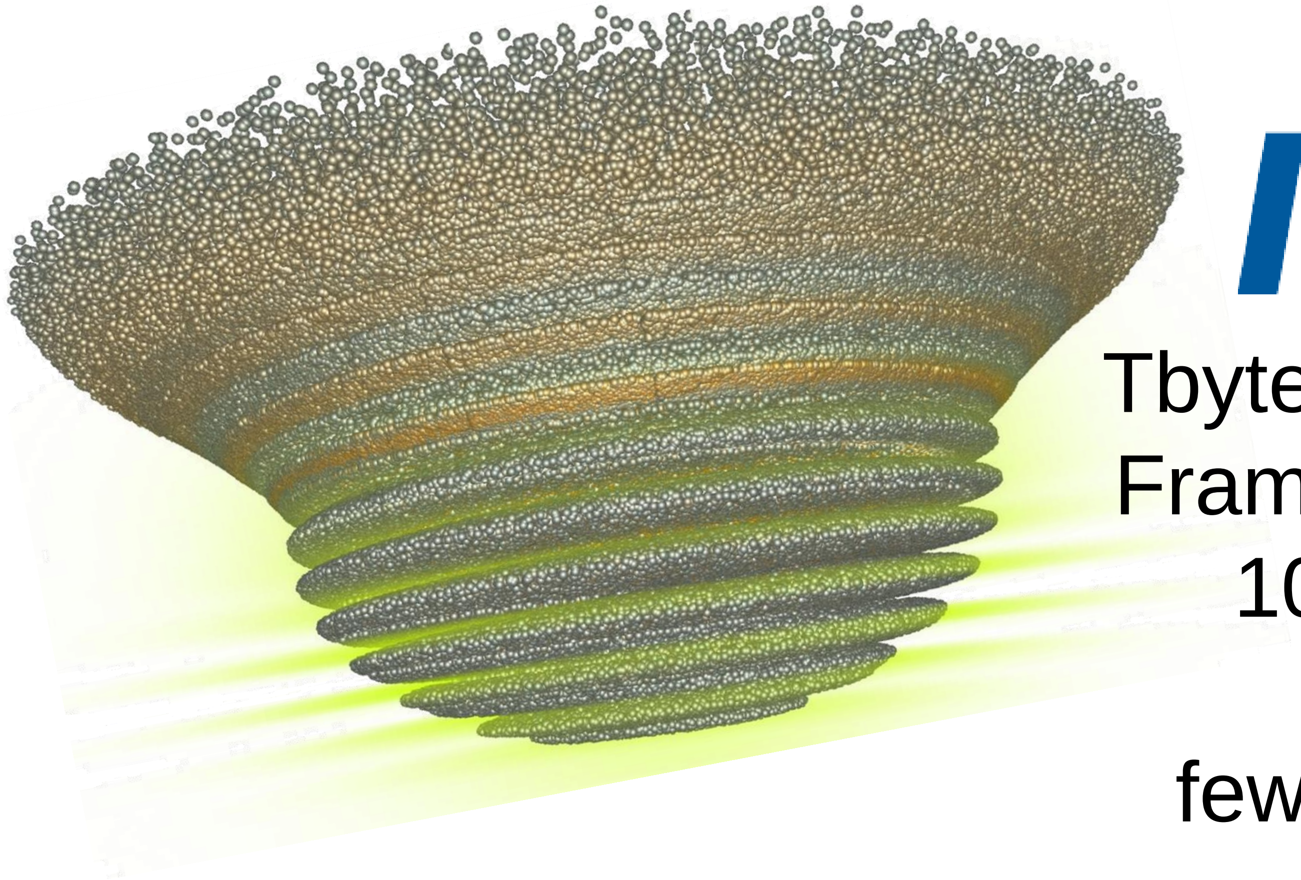
Performance Portability Layer

Scalability (Load Balancing)

Start to End simulations

Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

Scientific Workflows, Online Analysis & Visualization, SaaS, …

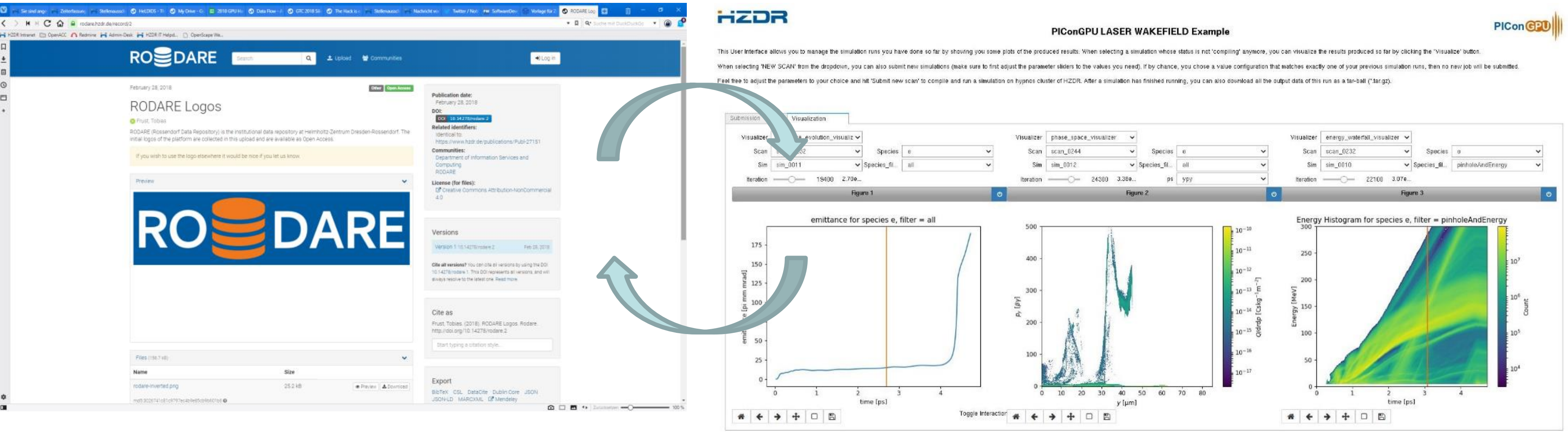


ISAAC

Tbyte / s throughput
Frames per second
10^{11} particles
10^{9} cells
few % overhead

Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

Scientific Workflows, Online Analysis & Visualization, SaaS, …



1) Simulation as a Service campaigns with in-situ exploration + analysis
2) Data + Metadata management with citation
3) Reproducibility, Accounting, User Community Access

Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de

Simulation as a Service / Data Repositories in the Cloud (Usability)

Simulation as a Service

Parameter Studies

Open Data in the Cloud

Combination with other Codes

Memory Staging & Workflows

I/O

Central Algorithms

Offline + Online Analysis

In-situ Visualization

Standard Formats for Data Exchange

Domain Specific Language

Performance Portability Layer

Scalability (Load Balancing)

Start to End simulations

Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

Common Inputs (Experimental Data?, MC?, Methods?)

Common Input to Simulations (LBNL, DESY, MdIS, …)

Simulation as a Service

Parameter Studies

Open Data in the Cloud

Combination with other Codes

Memory Staging & Workflows

I/O

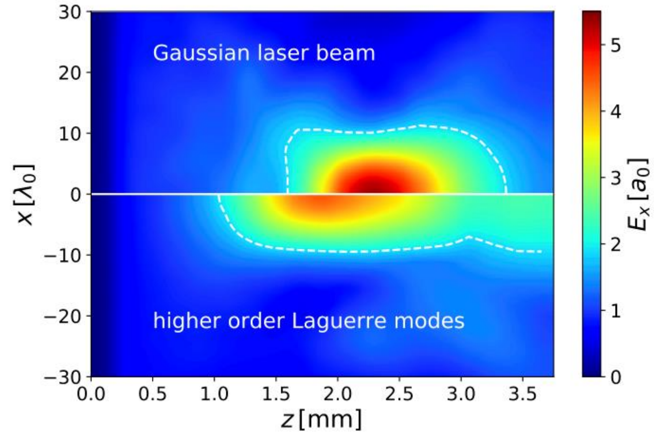
Central Algorithn

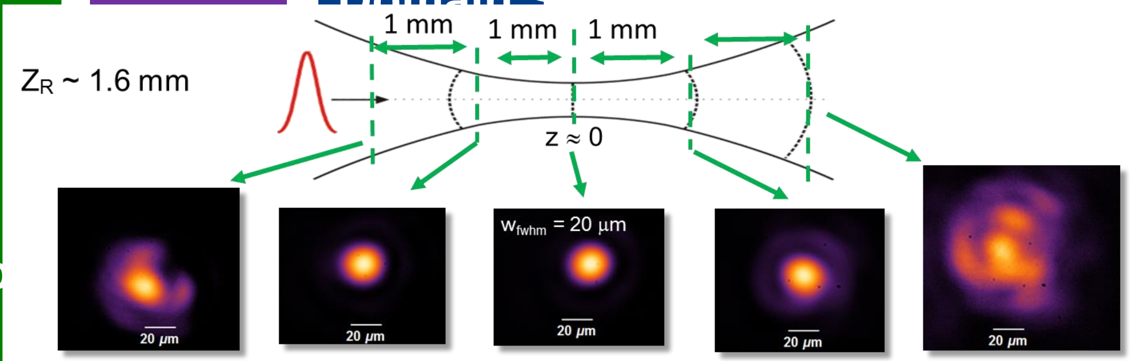
Domain S

Offline + Online Analysis

In-situ Visualizatio

(Load Balancing)





Strehl-ratio > 0.9 (~84%energy within the beam waist)

Start to End simulations

Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de

A common interface for PIC codes (MDK)

Common Input to Simulations

Simulation as a Service

Parameter Studies

Open Data in the Cloud

Combination with other Codes

Memory Staging & Workflows

Offline + Online Analysis

In-situ Visualization

I/O

Standard Formats for Data Exchange

Common Interface

Central Algorithms

Domain Specific Language

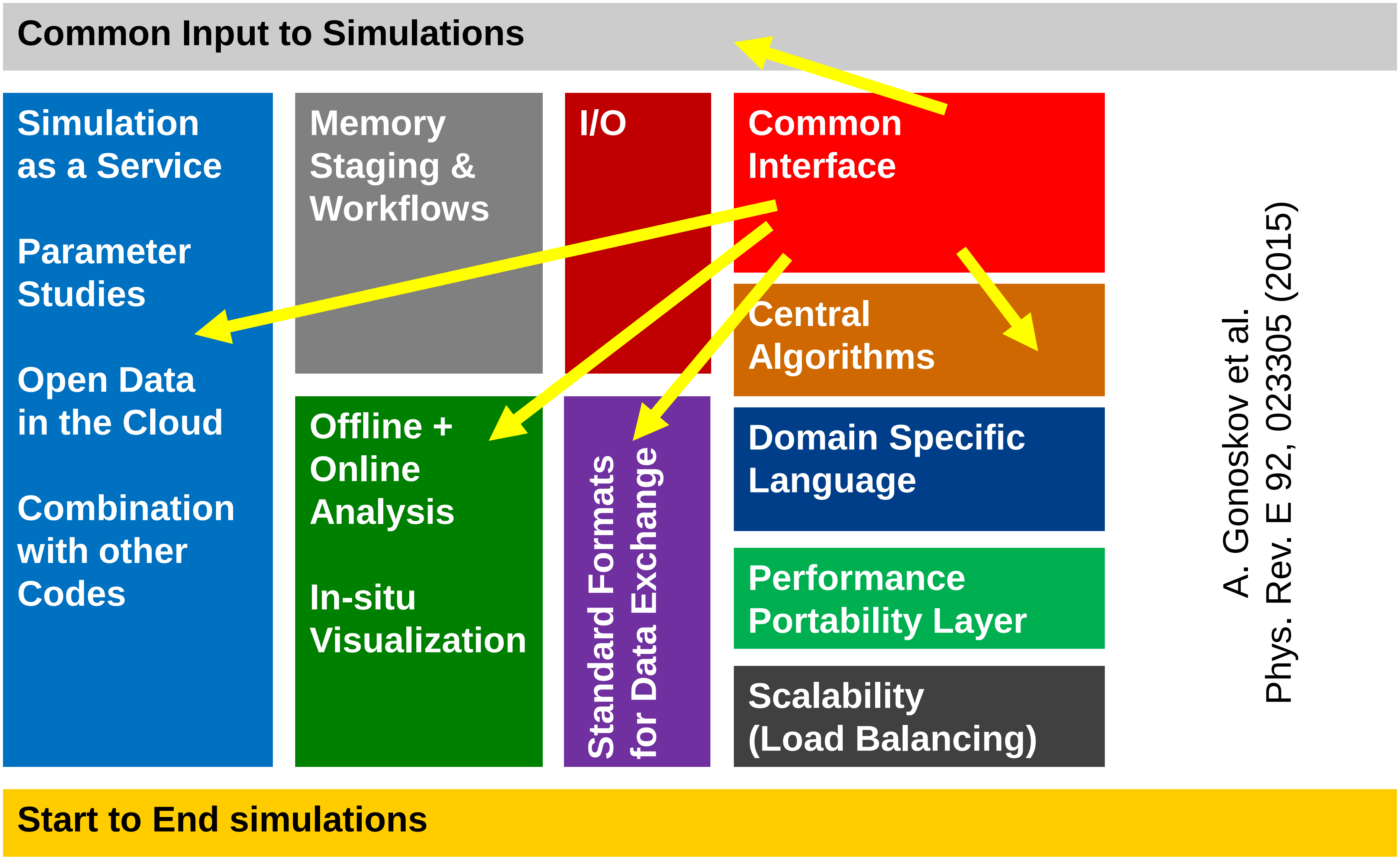
Performance Portability Layer

Scalability (Load Balancing)

A. Gonoskov et al.
Phys. Rev. E 92, 023305 (2015)

Start to End simulations

A common interface for PIC codes (MDK)



A. Gonoskov et al.
Phys. Rev. E 92, 023305 (2015)

Comparing to Experiments

Common Input to Simulations

Simulation as a Service

Parameter Studies

Open Data in the Cloud

Combination with other Codes

Memory Staging & Workflows

Offline + Online Analysis

In-situ Visualization

I/O

Standard Formats for Data Exchange

Common Interface

Central Algorithms

Domain Specific Language

Performance Portability Layer

Scalability (Load Balancing)

Synthetic Diagnostics

(passive, probing)

Start to End simulations

Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

Comparing to Experiments (Ionization as a Diagnostic)

Q: „Are you considering validating your codes?“

A: „We do not see the need for it,
we have done this already in 2008“

Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de
DRESDEN concept
HZDR
Mitglied der Helmholtz-Gemeinschaft

Synthetic Diagnostics

Common Input to Simulations

Simulation as a Service

Parameter Studies

Open Data in the Cloud

Combination with other Codes

Memory Staging & Workflows

Offline + Online Analysis

In-situ Visualization

I/O

Standard Formats for Data Exchange

Common Interface

Central Algorithms

Domain Specific Language

Performance Portability Layer

Scalability (Load Balancing)

Synthetic Diagnostics

(passive, probing)

Start to End simulations

Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

Simulation Codes today (totally exaggerated!)

Simulation

Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

Reliability, Reproducibility, Usability, Scalability, Open Science

Common Input to Simulations

Simulation as a Service

Parameter Studies

Open Data in the Cloud

Combination with other Codes

Memory Staging & Workflows

Offline + Online Analysis

In-situ Visualization

I/O

Standard Formats for Data Exchange

Common Interface

Central Algorithms

Domain Specific Language

Performance Portability Layer

Scalability (Load Balancing)

Synthetic Diagnostics

(passive, probing)

Start to End simulations

Michael Bussmann · Computational Radiation Physics · www.hzdr.de/crp
m.bussmann@hzdr.de
Mitglied der Helmholtz-Gemeinschaft

A community with codes is a community of community codes

- Dedicated, continuous Code Validation Programme (+ Experiment!)
  - Infrastructure for this is missing and must be created!
- SaaS & Open Data Repositories for the Community (Cloud Solutions)
  - Infrastructure for this is missing, but can be provided!
- Community access to Supercomputer Resources (PRACE, XSEDE)
  - Political support needed

PRACE XSEDE
Extreme Science and Engineering Discovery Environment

DRESDEN concept HZDR