

High performance computing and data management requirements related to computer simulations

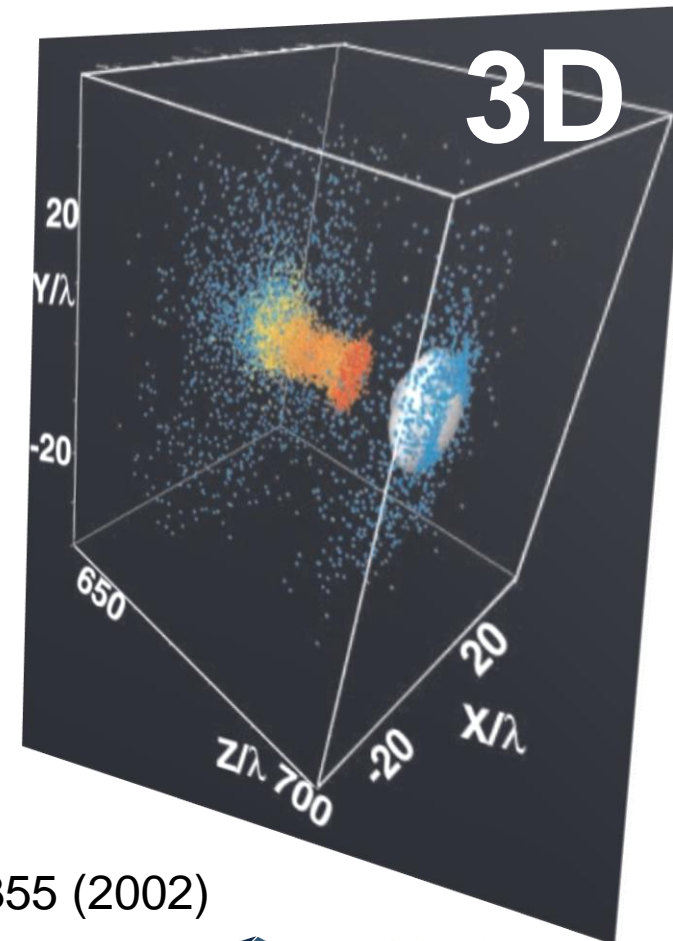
Michael Bussmann

HZDR

 **HELMHOLTZ**
ZENTRUM DRESDEN
ROSSENDORF

The role of simulations in plasma-based accelerators

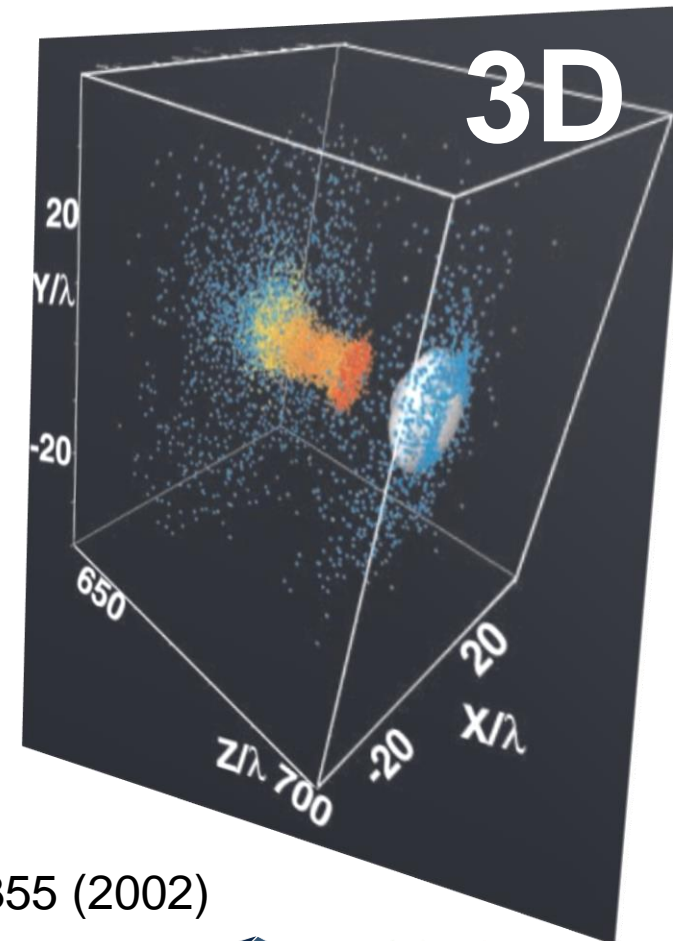
- Inventing new schemes for acceleration, injection and control of beam
- Post-hoc analysis of experimental findings
- Parameter studies
- Quantitative Predictions of Beam Parameters



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

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)

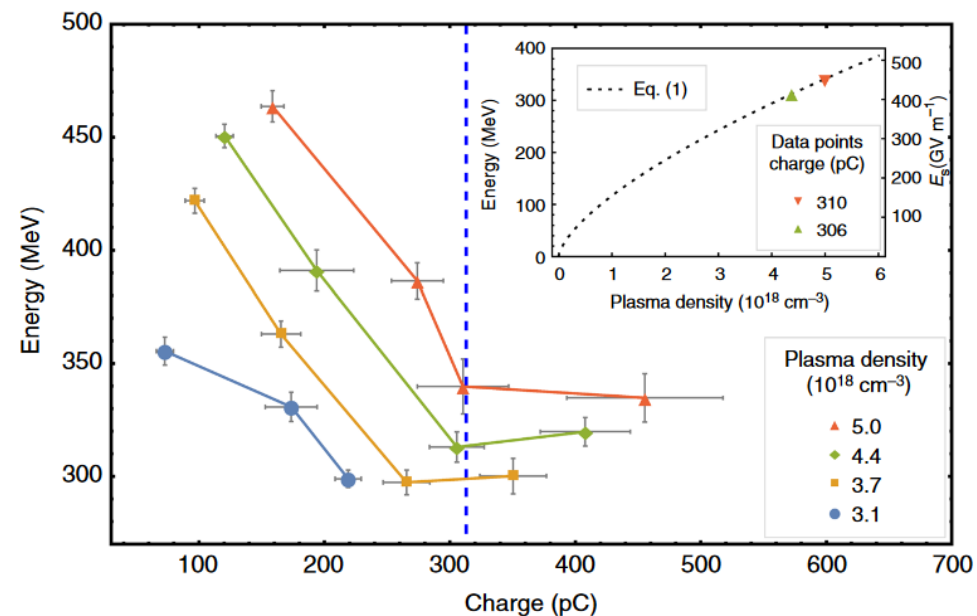
Do we want & need predictive capabilities?

- 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)



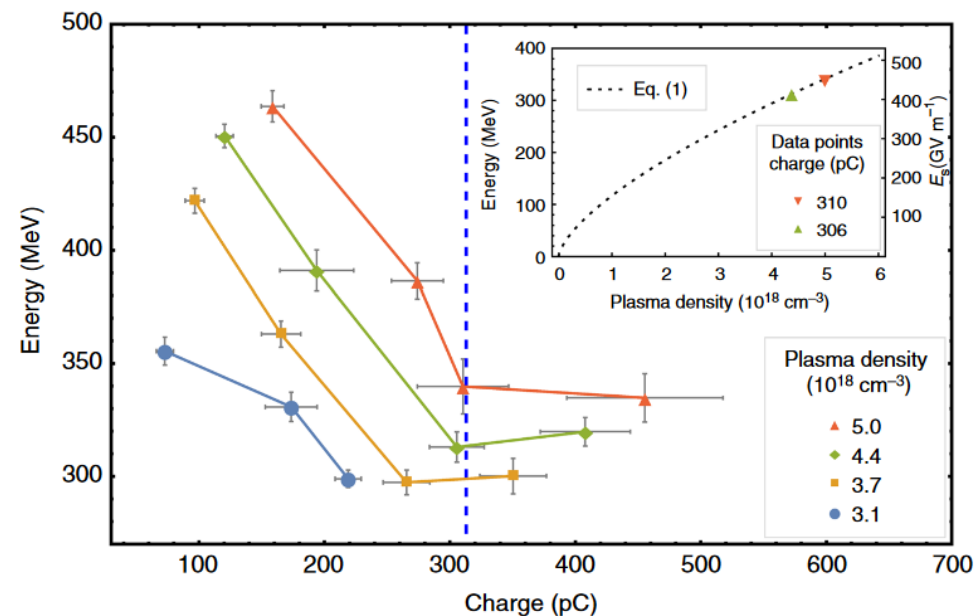
Do we want & need predictive capabilities?

- 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)



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

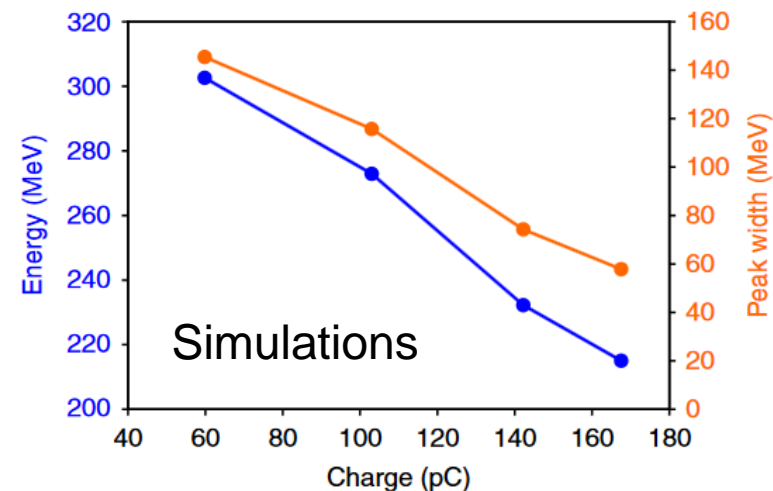
- Pen, paper, intuition & experience
- Hybrid + mixed fluid / kinetic schemes, electrostatic, MD/MC, etc.
- Reduced dimensions, boosted frames, ML / reduced models



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



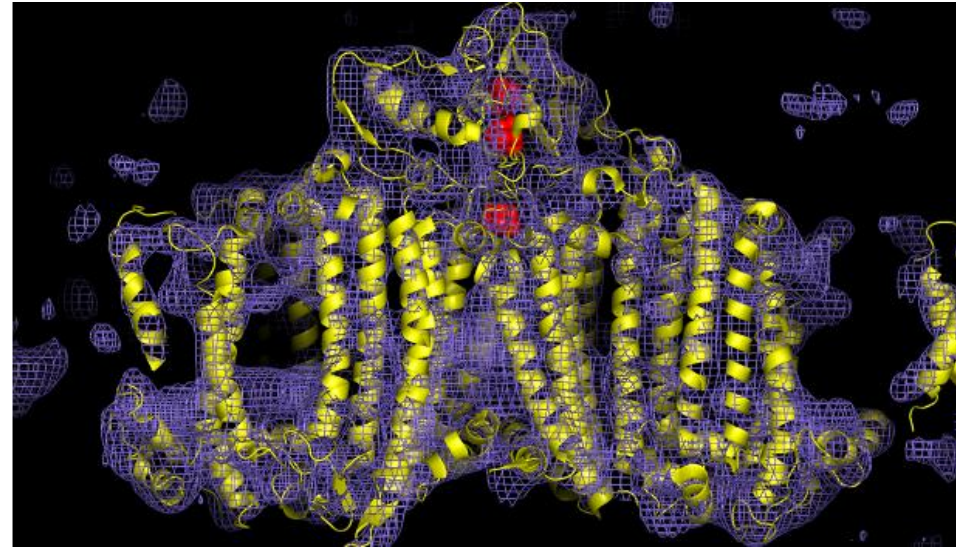
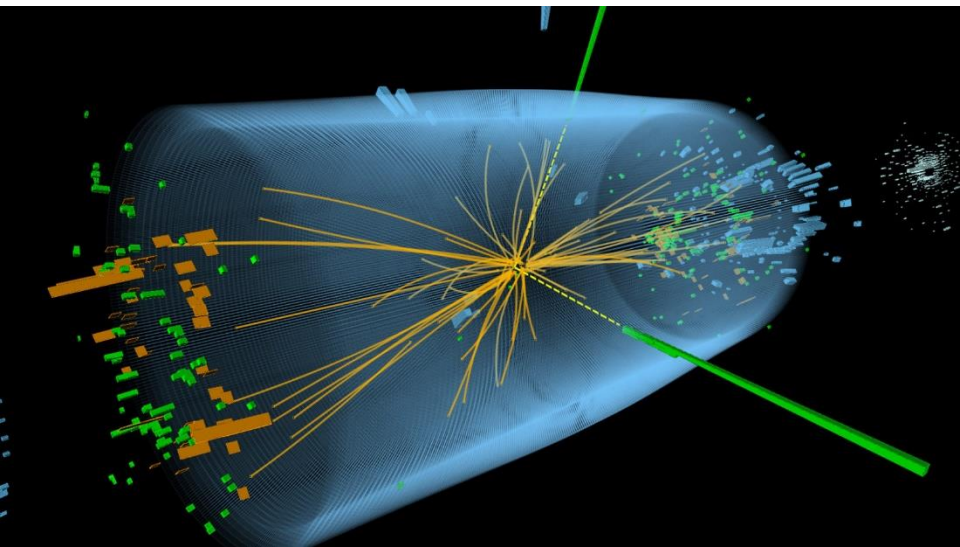
Don't trust ANY code (especially the one written by yourself!)

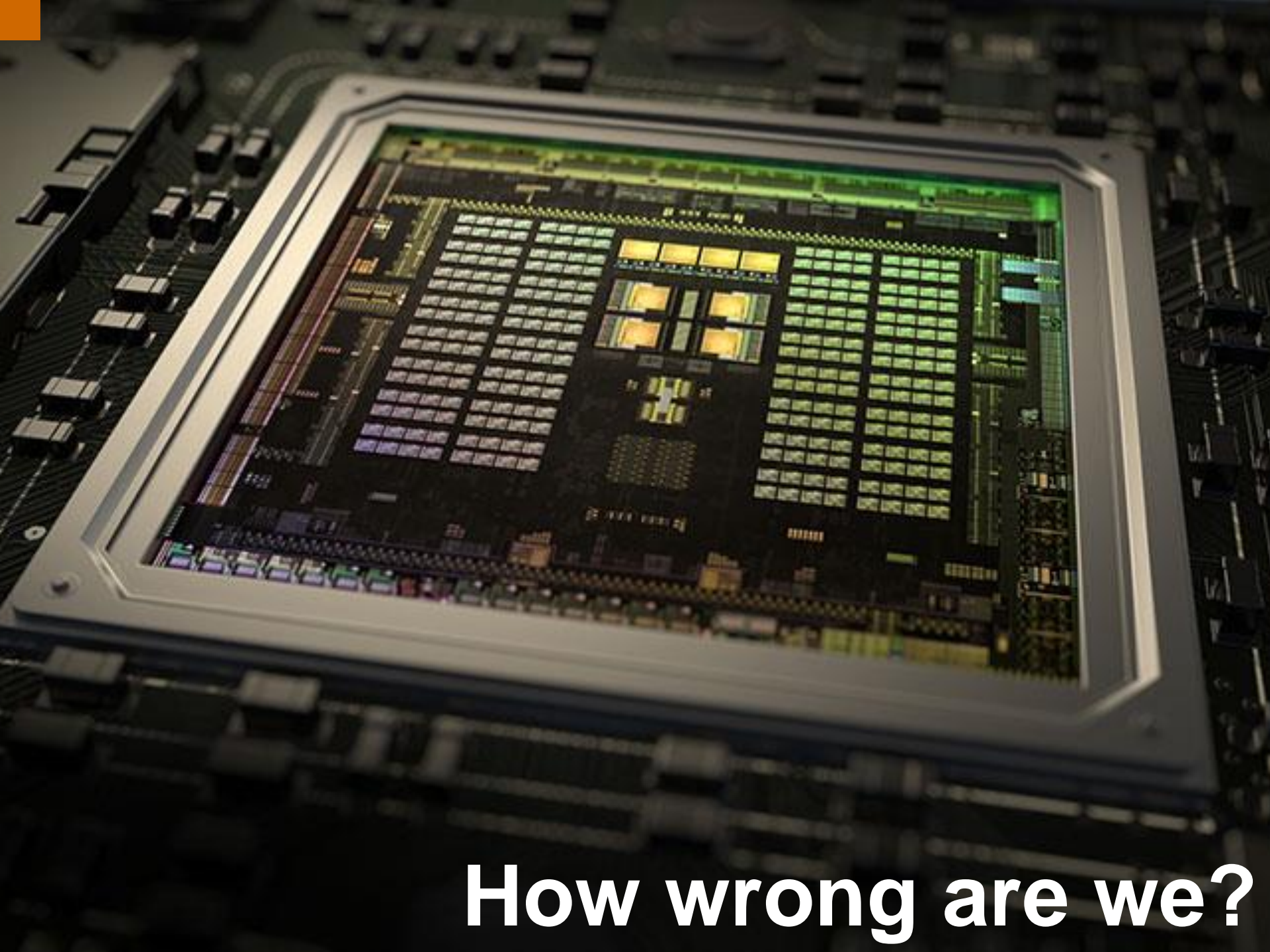
- Not all methods are implemented in every code
- Some methods cannot be implemented together
- Some methods interact spuriously with others
- Some methods have limited applicability
- Some methods are approximate
- Some methods are hard to parallelize

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

Start to End Simulations



NOVEL FUNDAMENTAL RESEARCH
COMPACT EUROPEAN PLASMA
ACCELERATOR WITH SUPERIOR
BEAM QUALITY

Simulations on 10 different schemes

Start to End simulations

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
    // ...

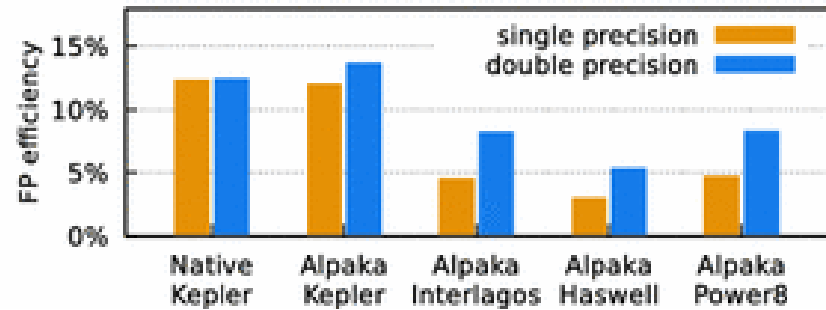
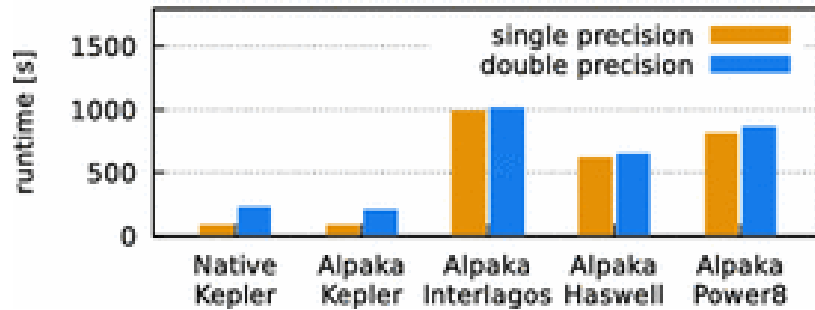
#else
    // Sequential CPU implementation
    // ...

#endif
```

Start to End simulations

Towards Exascale Computing (Capability + Capacity)

Runtime and Floating Point Efficiency of the PICongPU Kelvin Helmholtz Instability Simulation



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



Central Algorithms

Domain Specific Language

Performance Portability Layer

Scalability (Load Balancing)

Start to End simulations

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”

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.**”*



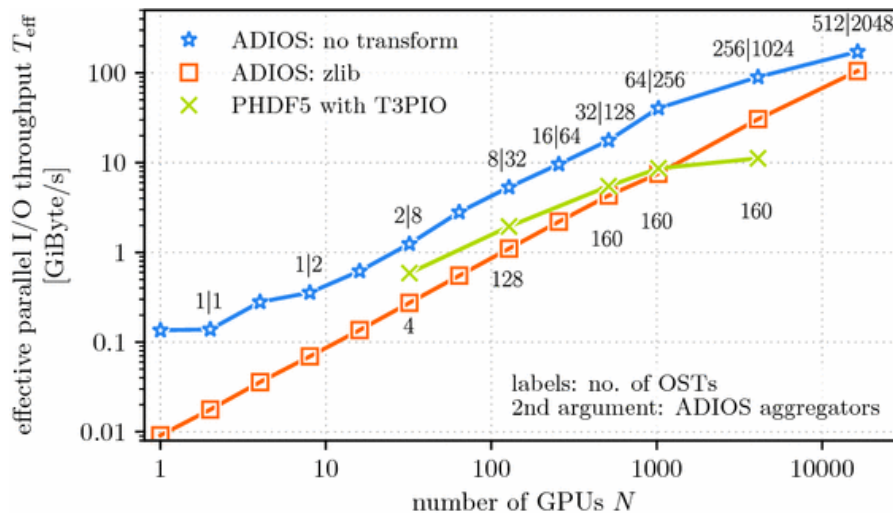
Central Algorithms

Domain Specific Language

Performance Portability Layer

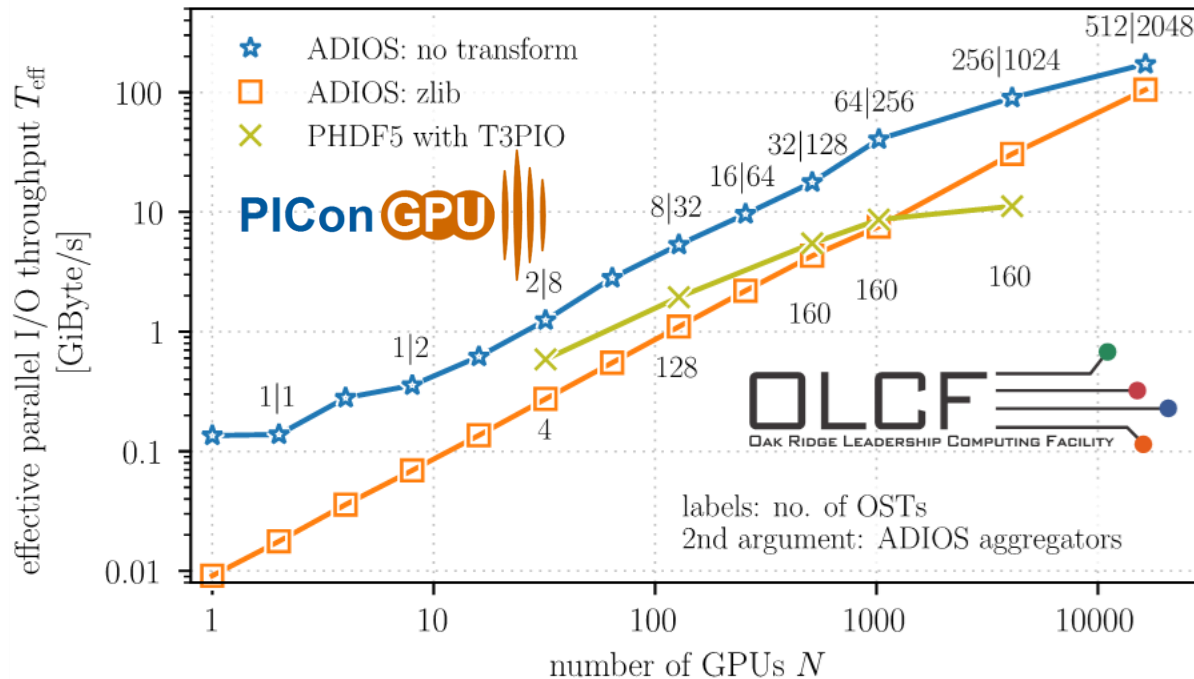
Scalability (Load Balancing)

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



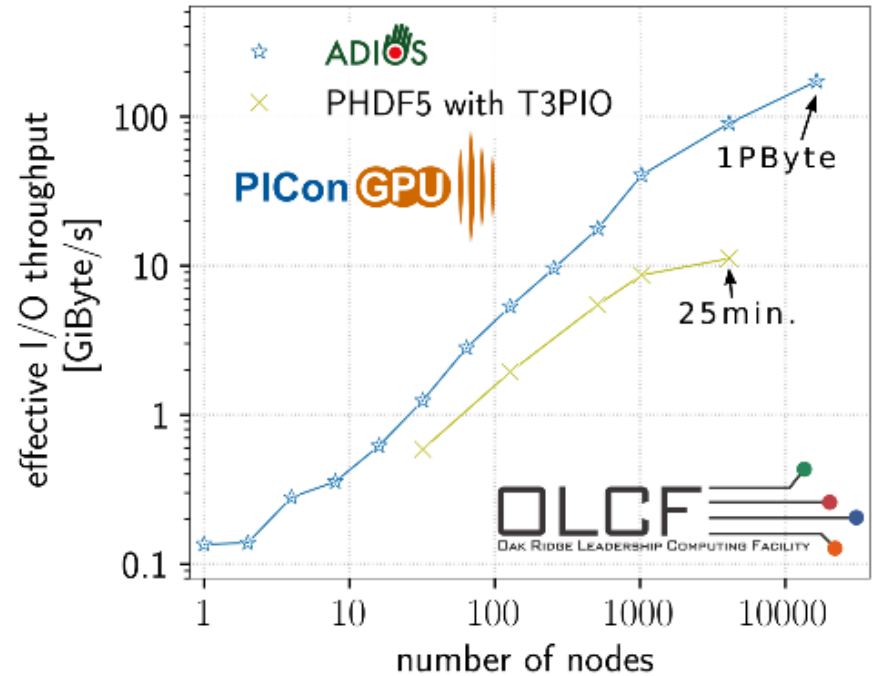
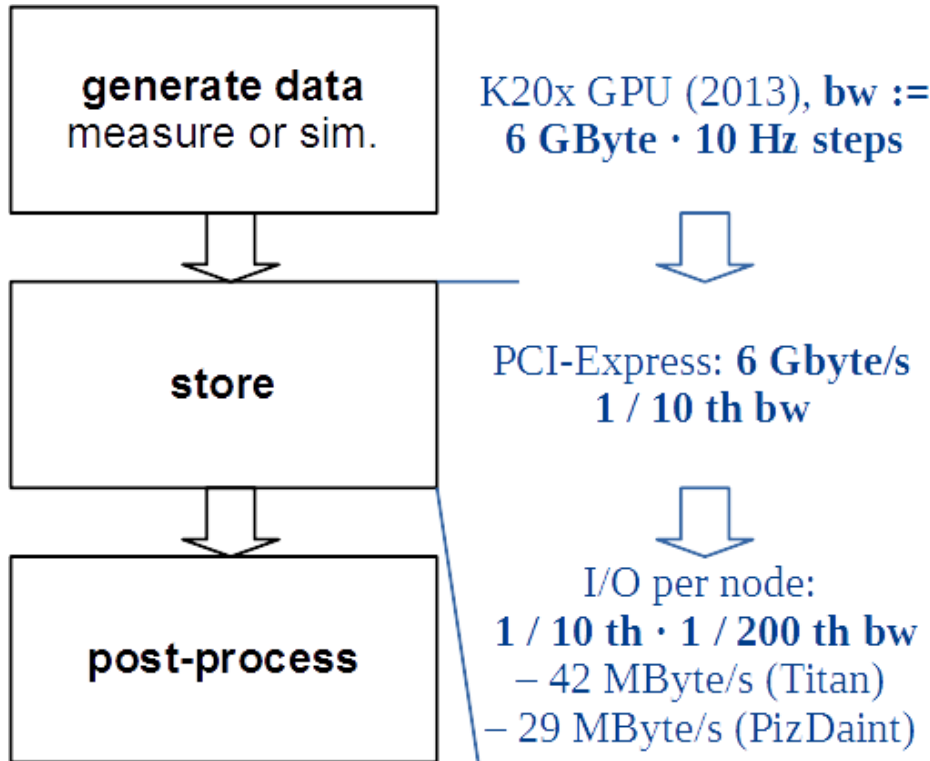
Start to End simulations

Exascale I/O for everybody



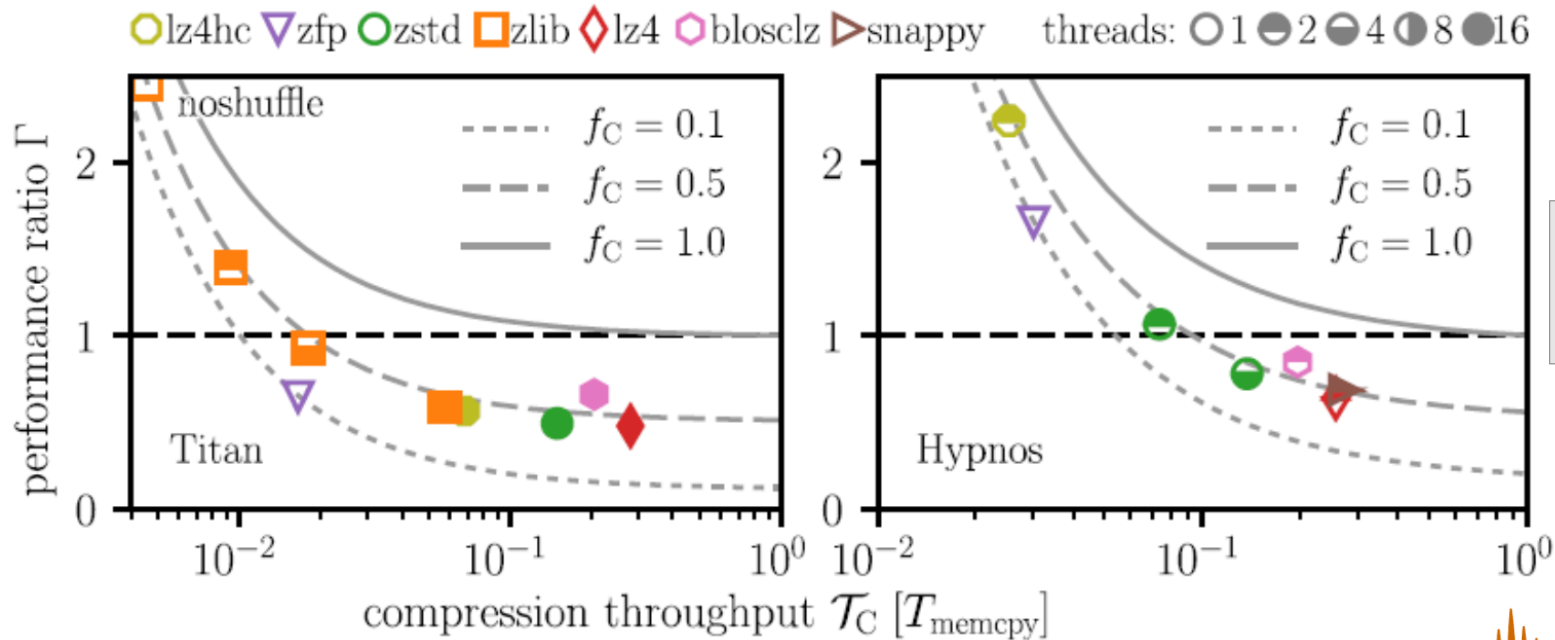
- 1 TB/s theoretical throughput
- 250 GB/s reached
- 14 MB/s per node
- Exclusive access
- How is it accessed?
- How long may it live?
- 1 Pb I/O in 1 hour

Writing to disk won't work anymore



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

Data reduction won't help, except when you throw stuff away



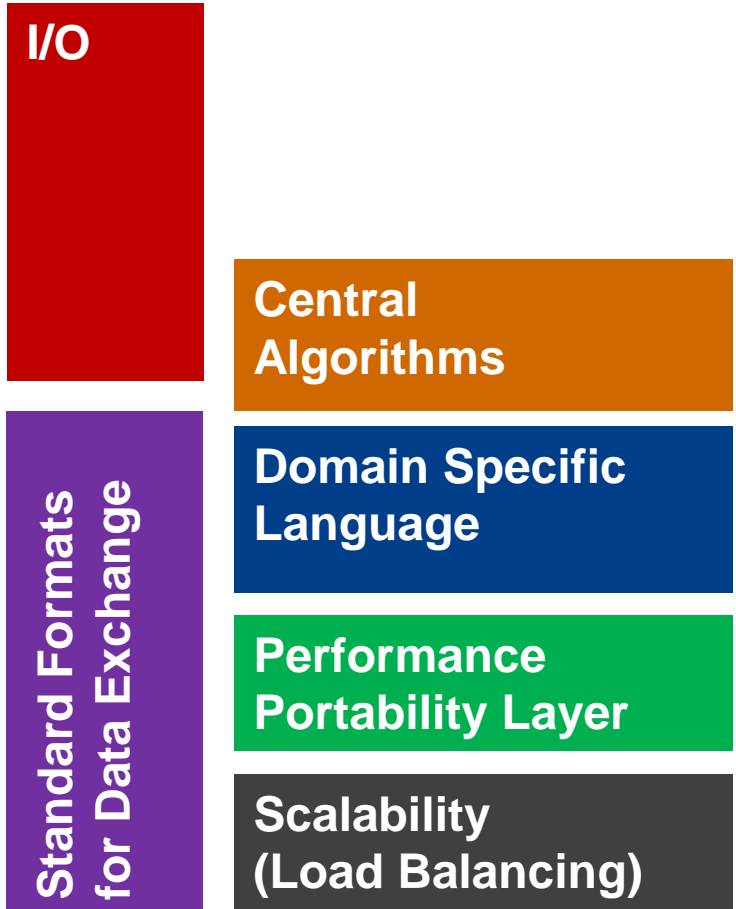
$$\Gamma \equiv \frac{t_{I/O}^{\text{reduced}}}{t_{I/O}^{\text{unreduced}}}$$



Open, self-descriptive data exchange between applications

Share data between

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



Start to End simulations

Open Source META DATA format for Plasma simulation codes

HDFView 2.9 interface showing the metadata for the 'electrons' group. The tree structure is: example.h5 > data > 0 > fields > rho > B > x, y, z; particles > electrons. The metadata for 'electrons' is:

- electrons (63328, 4)
- Group size = 5
- Number of attributes = 6
- currentDeposition = Esirkepov
- longName = My first electron species
- particleInterpolation = Trilinear
- particlePush = Boris
- particleShape = 3.0
- particleSmoothing = none

HDFView 2.9 interface showing the metadata for the 'charge' group. The tree structure is: example.h5 > data > 0 > fields > rho > B > x, y, z; particles > electrons > charge. The metadata for 'charge' is:

- charge (64360, 4)
- Group size = 0
- Number of attributes = 3
- unitDimension = 0.0,0.0,1.0,1.0,0.0,0.0,0.0,0.0
- unitSI = 1.60217657E-19
- value = -1.0

GitHub repository page for openPMD / openPMD-standard. The page shows the version 2.0.0, updated 8 days ago. It lists several pull requests in the Proposed, Accepted, and Review stages.

- Proposed (9):**
 - #195 positionOffset optional (major change)
 - #182 Rename position and positionOffset (question)
 - #181 gridGlobalOffset: allow int (major change)
 - #162 `particleName/`: Clarify it is arbitrary (revision change)
 - #160 Make "time", "dt" and "timeUnitSI" (Automated as To do)
- Accepted (9):**
 - #196 Paths: End on / (revision change)
 - #149 Labels for Symbols and Units (minor change)
 - #166 Simplifying iterationEncoding and iterationFormat (major change)
 - #148 Rename Iteration (major change)
 - #147 Iteration Formatting: Support Zero-Padding in %T (major change)
- Review (8):**
 - #194 Mesh: F (major change)
 - #125 dataOr (major change)
 - #129 Simply dataOr (major change)
 - #183 New EXT: Bea (EXT: Bea)



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

openPMD Eco-System



github.com/openPMD/openPMD-projects

openPMD standard (1.0.0, 1.0.1, 1.1.0)
the underlying file markup and definition
A Huebl et al., doi: 10.5281/zenodo.33624

base standard	extensions
<i>general description</i>	<i>domain-specific</i>
e.g. ED-PIC, SpeciesType, BeamPhysics	



native data tools

HDF5, ADIOS1/2, NetCDF, ...
e.g. h5ls, h5repack, h5dump, bpdump

writers & converters

simulations, frameworks, measurements
e.g. PIconGPU, Warp, SIMEX_Platform

HDF Compass

HDF5 & ADIOS file explorer
open and explore file trees

readers

coupled simulations, post-processing frameworks, ...
e.g. SIMEX_Platform, Visit, yt-project, openPMD-viewer

openPMD-updater

update to new standard
edit in- or new file

openPMD-api

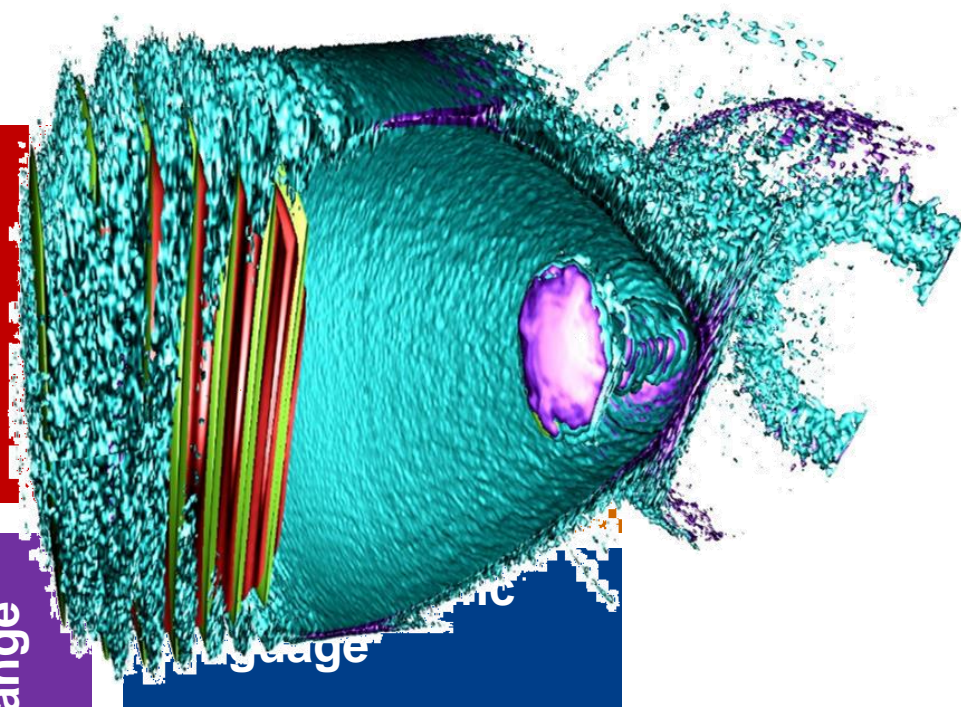
I/O library abstraction
file format agnostic

data repositories

exchange and long-time archival
e.g. Zenodo, RODARE (HZDR)



Scientific Workflows, Online Analysis & Visualization, SaaS, ...



Memory Staging & Workflows

I/O

Offline + Online Analysis
In-situ Visualization

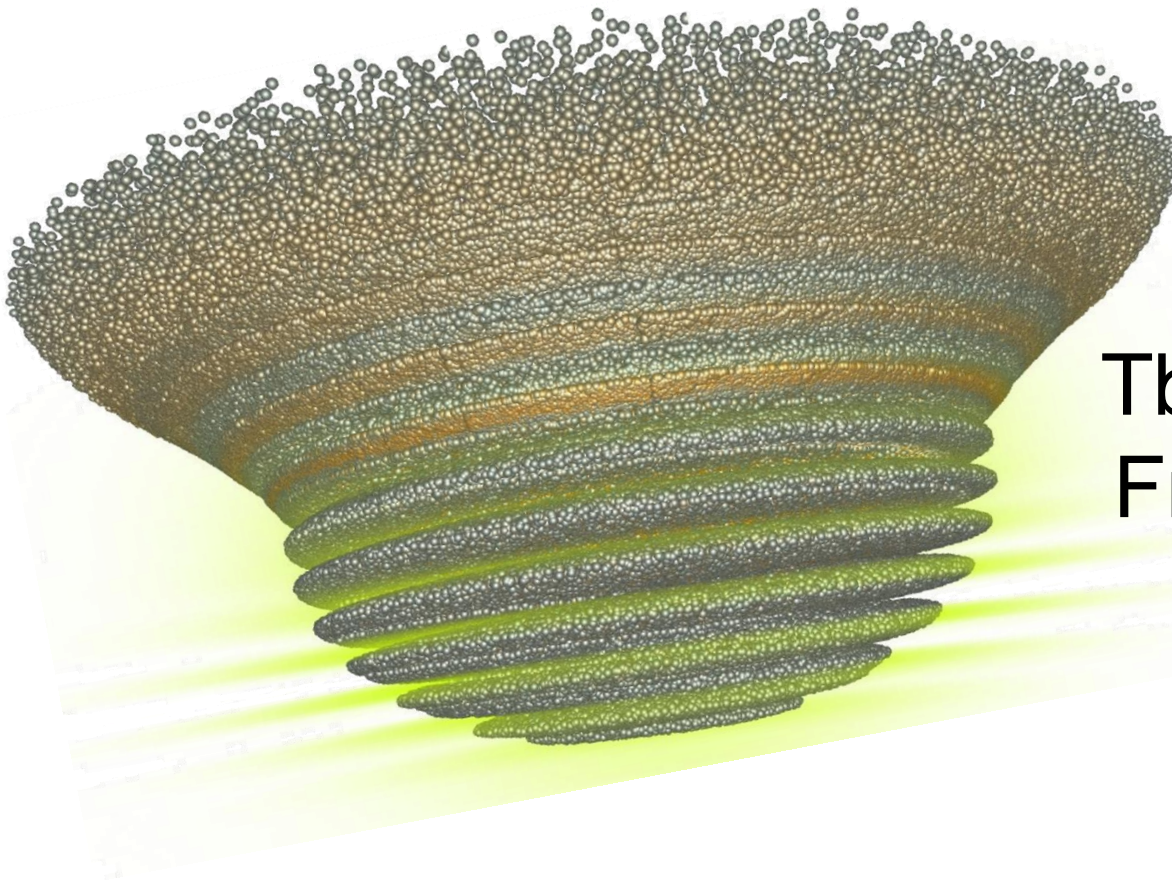
Standard Formats for Data Exchange

Performance Portability Layer

Scalability (Load Balancing)



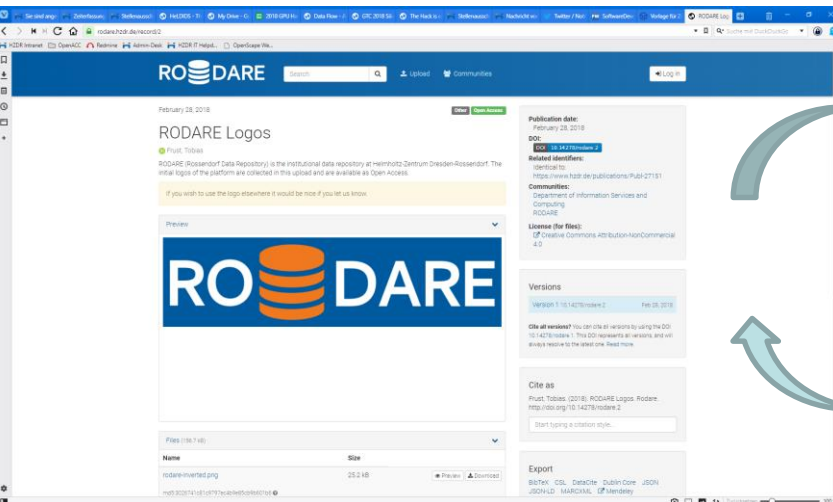
Start to End simulations



ISAAc

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

Scientific Workflows, Online Analysis & Visualization, SaaS, ...

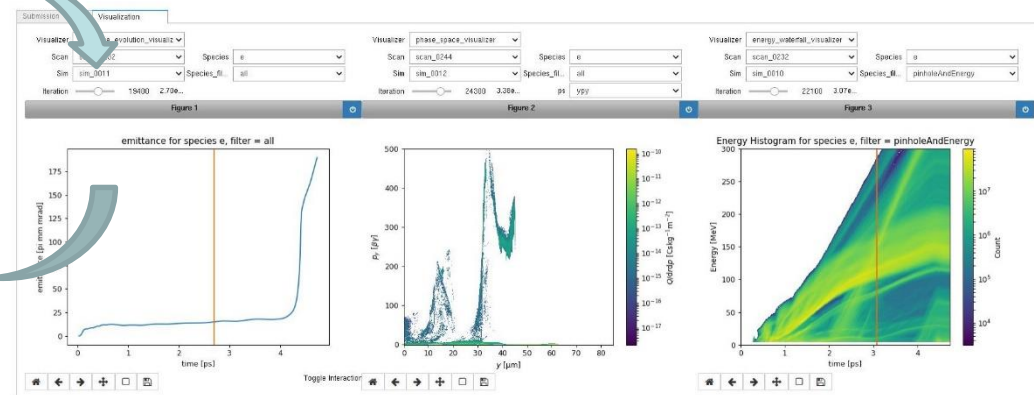


HZDR

PICongPU LASER WAKEFIELD Example

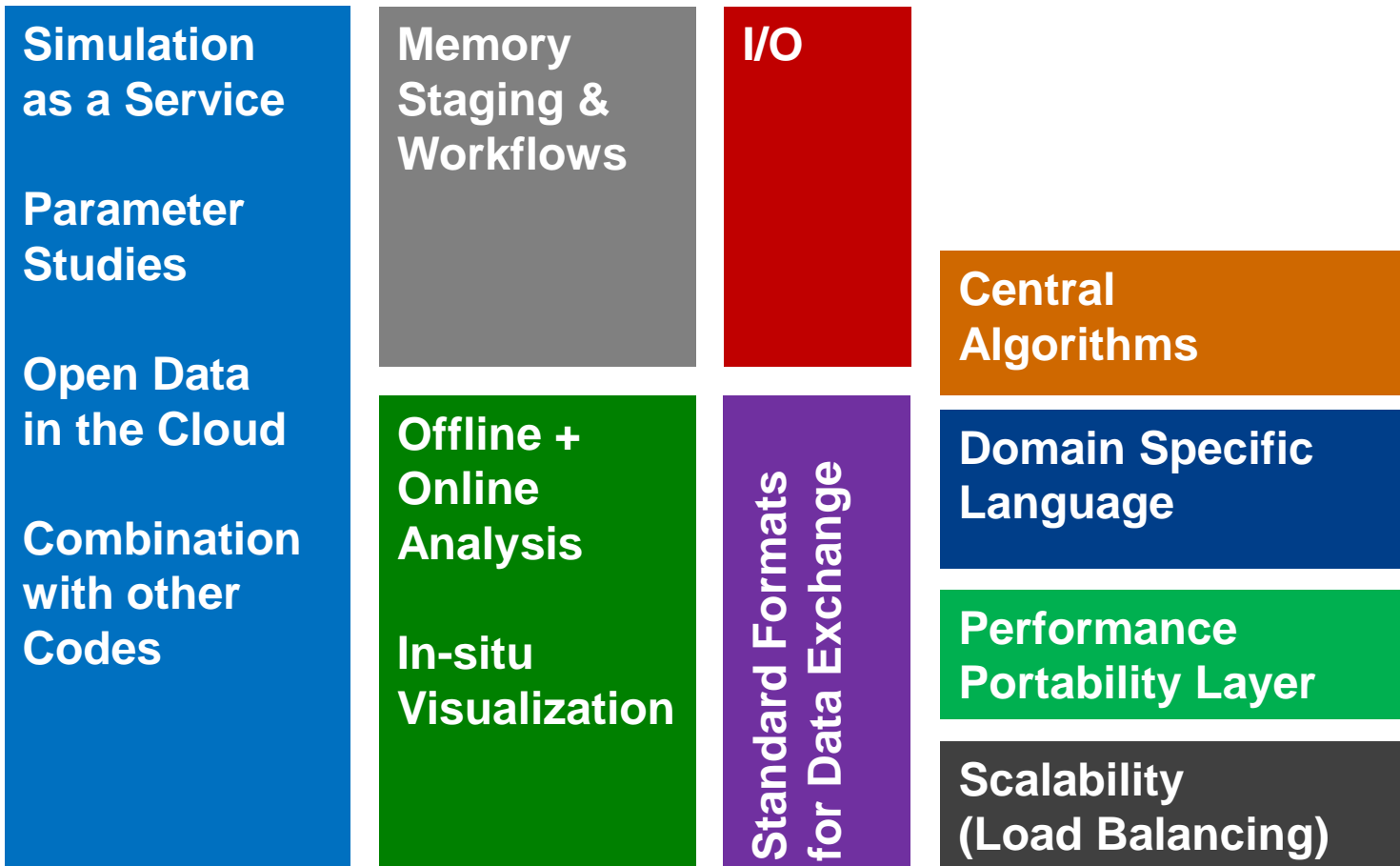
PICongPU

This User Interface allows you to manage the simulation runs you have done so far by showing you some plots of the produced results. When selecting a simulation whose status is not 'completing' anymore, you can visualize the results produced so far by clicking the 'Visualize' button. When selecting 'NEW SCAN' from the dropdown, you can also submit new simulations (make sure to first adjust the parameter sliders to the values you need). If by chance, you chose a value configuration that matches exactly one of your previous simulation runs, then no new job will be submitted. Feel free to adjust the parameters to your choice and hit 'Submit new scan' to compile and run a simulation on Hypos cluster of HZDR. After a simulation has finished running, you can also download all the output data of this run as a tar-ball (tar.gz).



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

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



Start to End simulations

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

Common Input to Simulations (LBNL, DESY, MdlS, ...)

Simulation as a Service

Parameter Studies

Open Data in the Cloud

Combination with other Codes

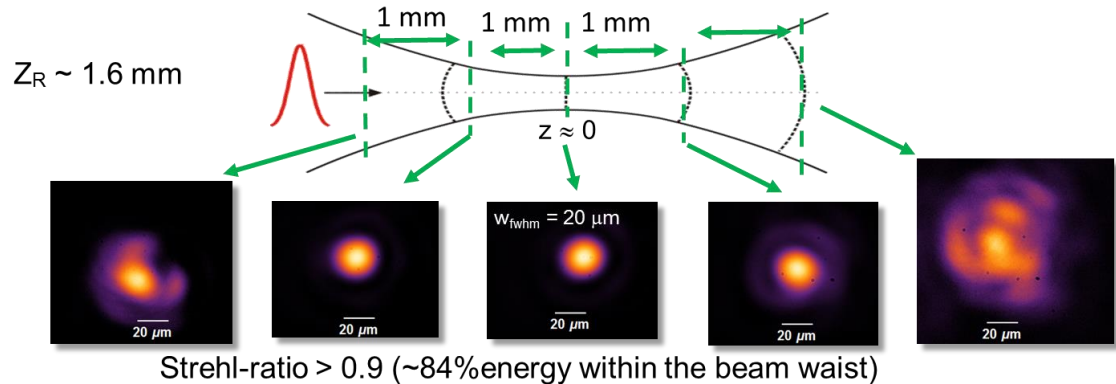
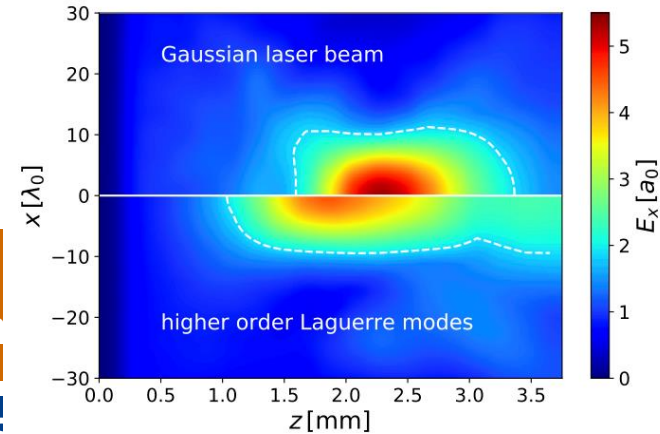
Memory Staging & Workflows

I/O

Central Algorithm

Domain S

Offline + Online Analysis
In-situ Visualization



Start to End simulations

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

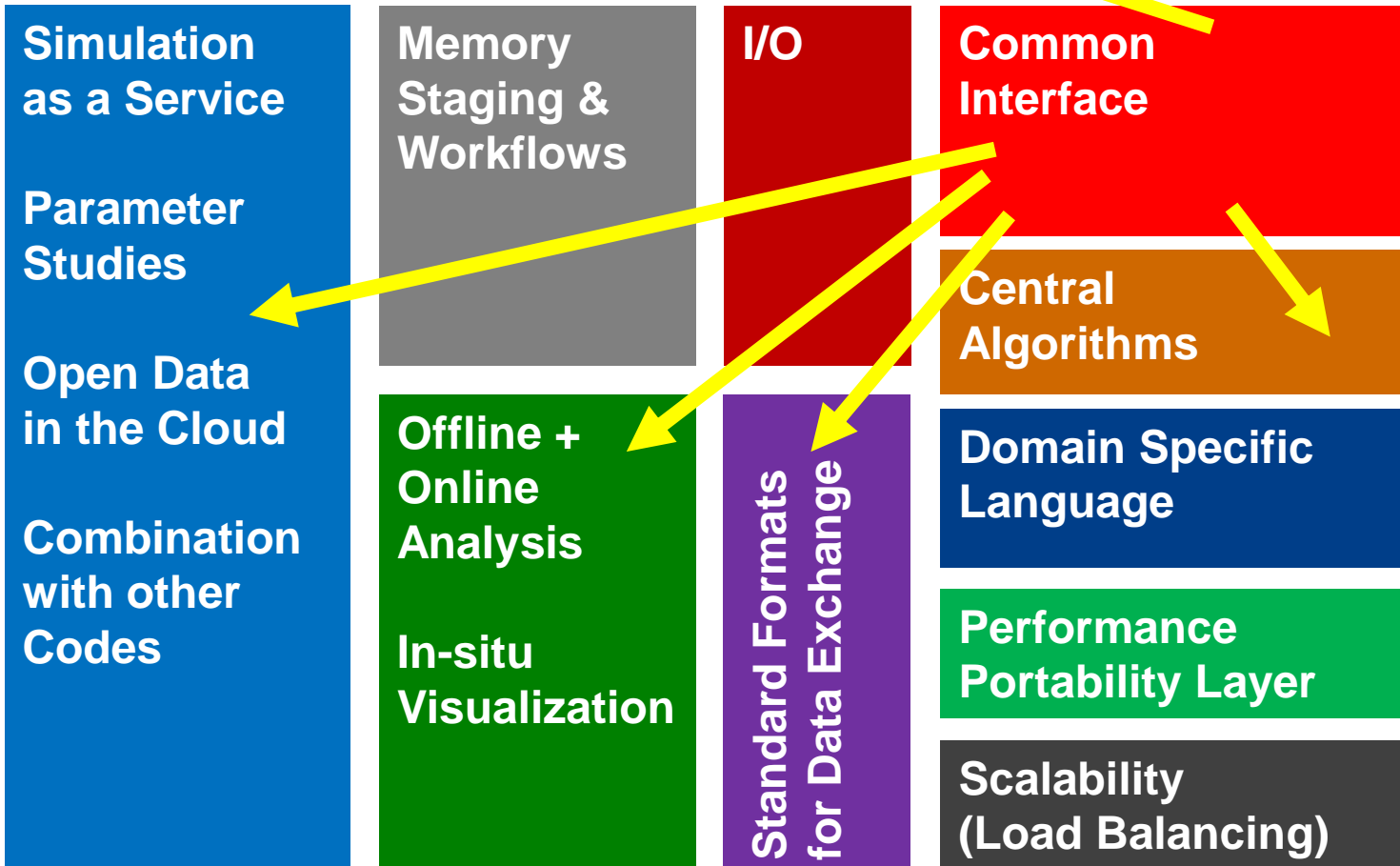
Scalability (Load Balancing)

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

Start to End simulations

A common interface for PIC codes (MDK)

Common Input to Simulations



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

Start to End simulations

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

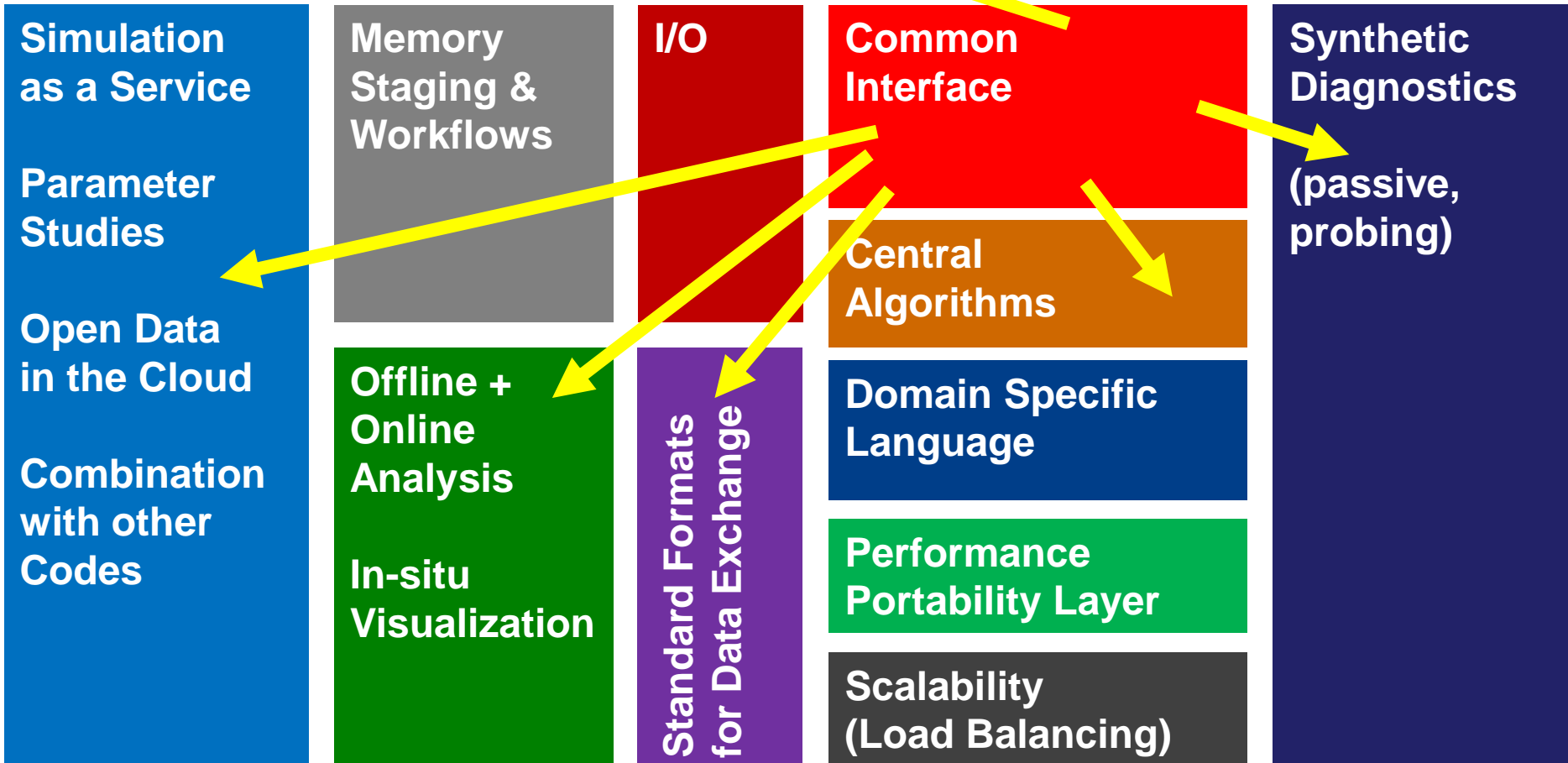
Comparing to Experiments (Ionization as a Diagnostic)

Q: *„Are you considering validating your codes?“*

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

Synthetic Diagnostics

Common Input to Simulations



Start to End simulations

Simulation Codes today (totally exaggerated!)

Simulation

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

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

