Extreme scale computing for integrated multi-scale physics in CPES*†

C.S. Chang[1] for the CPES[2] Team

[1]Princeton Plasma Physics Laboratory
[2]SciDAC Proto-FSP Center for Plasma Edge Simulation

*Work supported by US DOE OASCR and OFES.
†Computing time provided by OLCF INCITE and NERSC ERCAP awards.

Outline

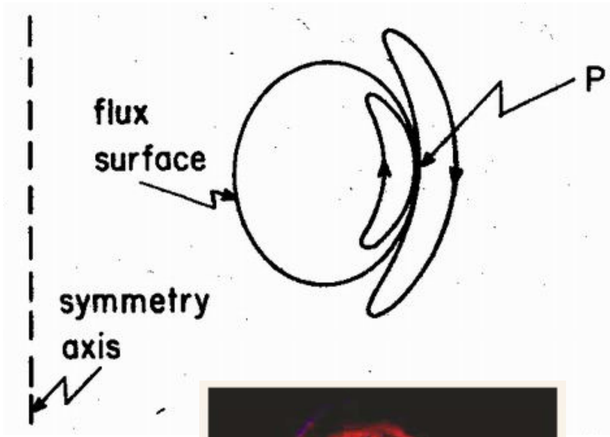
- Introduction of extreme scale simulation of multi physics in CPES
- Extreme scale elements: XGC full-f kinetic codes for the first-principles multi-physics in the whole volume
- Different challenges for extreme scale integration
- EFFIS for small to extreme scale integration
- Conclusion and discussion

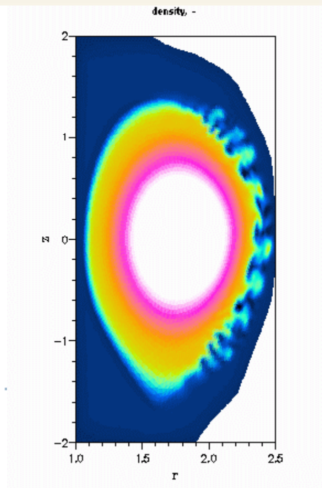
Extreme scale simulation of multi physics in CPES

- **Peta scale multi-physics, XGC1:** full-f kinetic, whole volume, turbulence + neoclassical
- **0.1 peta scale, XGC0+DEGAS2:** full-f kinetic transport modeling in whole volume,
  - Neoclassical + neutral + impurities + atomic + modeled anomalous transport
  - Poisson + Ampere for 3D B-perturbation
- **Tera scale, M3D (& NIMROD):** nonlinear MHD
- Small scale codes

  1) **Grid transformation (M3D_omp)**
  2) **Magnetic equilibrium reconstruction (FlowM3D)**
  3) **Linear ELM criterion (Elite)**
- **Integration tools: Adios and Kepler**

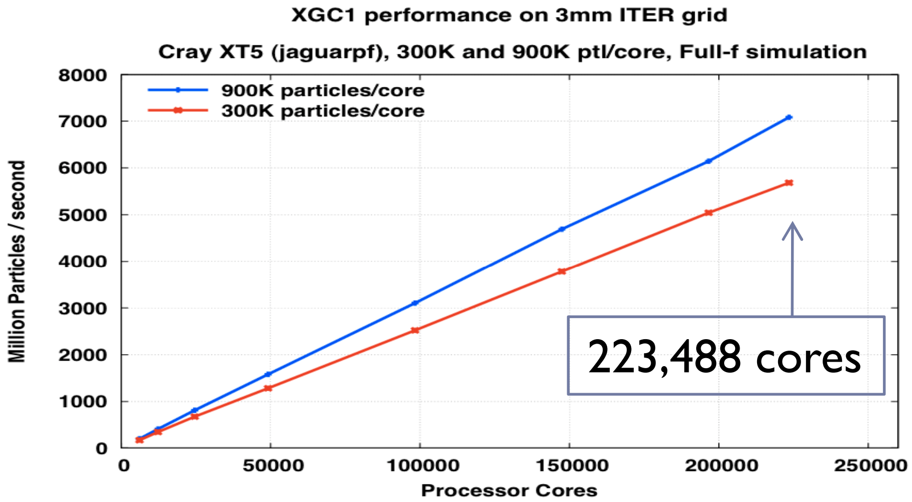
  1) **In-memory and file based couplings, simultaneously**
  2) **In-situ job control (parameter injection)**
  3) **Data analysis and visualization on remote eSiMon dashboard**

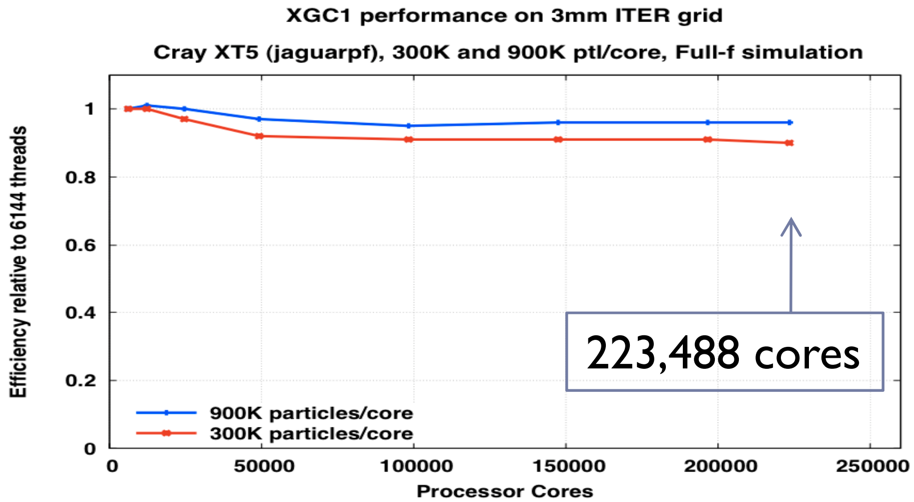




XGC1 Scales efficiently to the maximal number of Jaguar cores, and utilizes peta-scale routinely

12 cores per node, 2 MPI processes per node



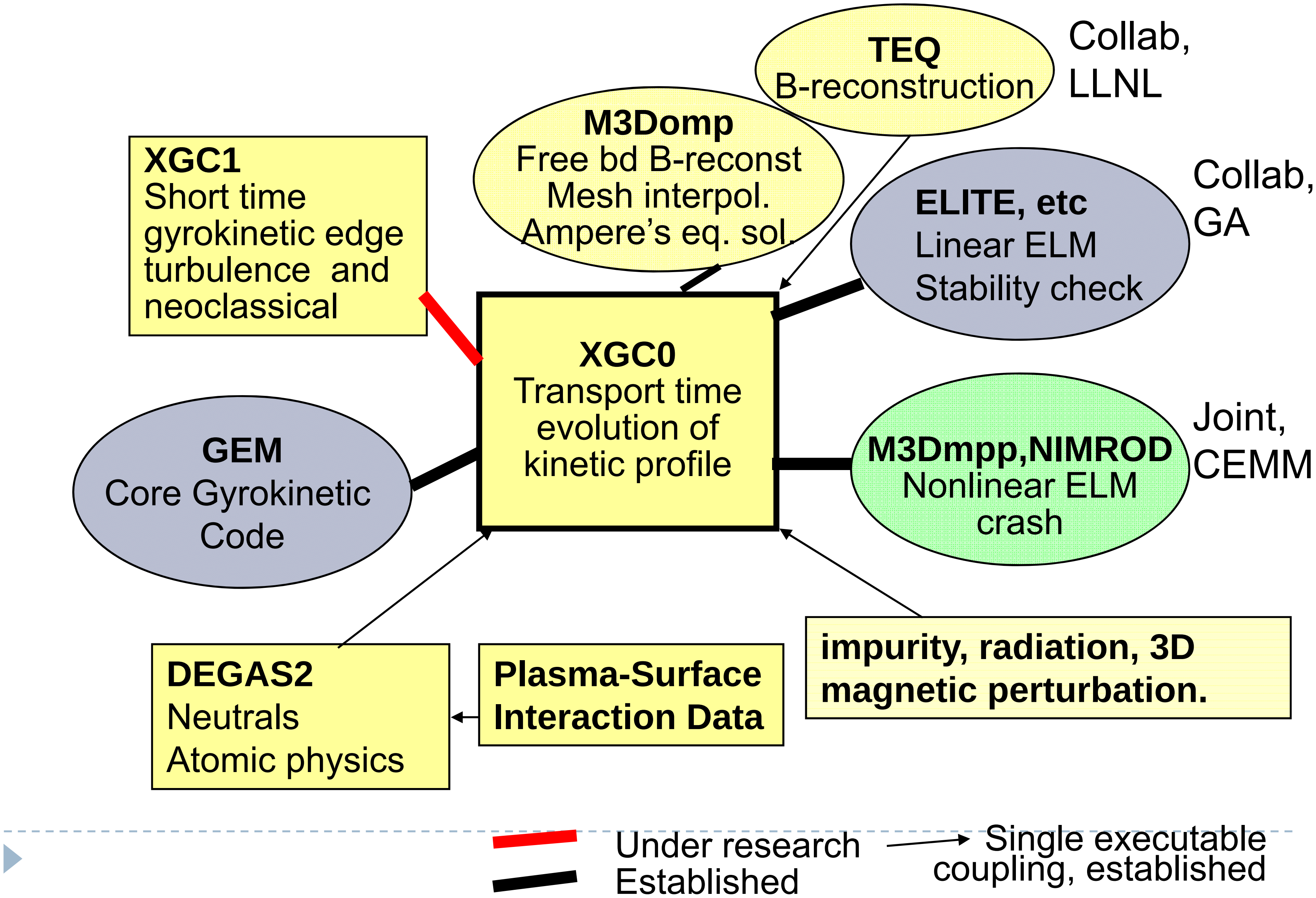


- **900K particles per thread problem is more computationally intensive than 300K problem, which leads to ~20% higher particle push rate.**
- **Performance scaling is excellent for both problems.**

CY10 Average Job Size and Utilization by Job Size Bins (Jan. 1 – June 27, 2010)

	Average Job Size in Cores	Utilization in Core-Hours
Jobs requesting <20% of the available resources	3,079	8,446,978
Jobs requesting between 20% and 60% of the available resources	66,474	3,042,575
Jobs requesting >60% of the available resources	170,304	14,311,232

Multi-physics Integration on EFFIS
(End-to-end Framework for Fusion Integrated Simulation)

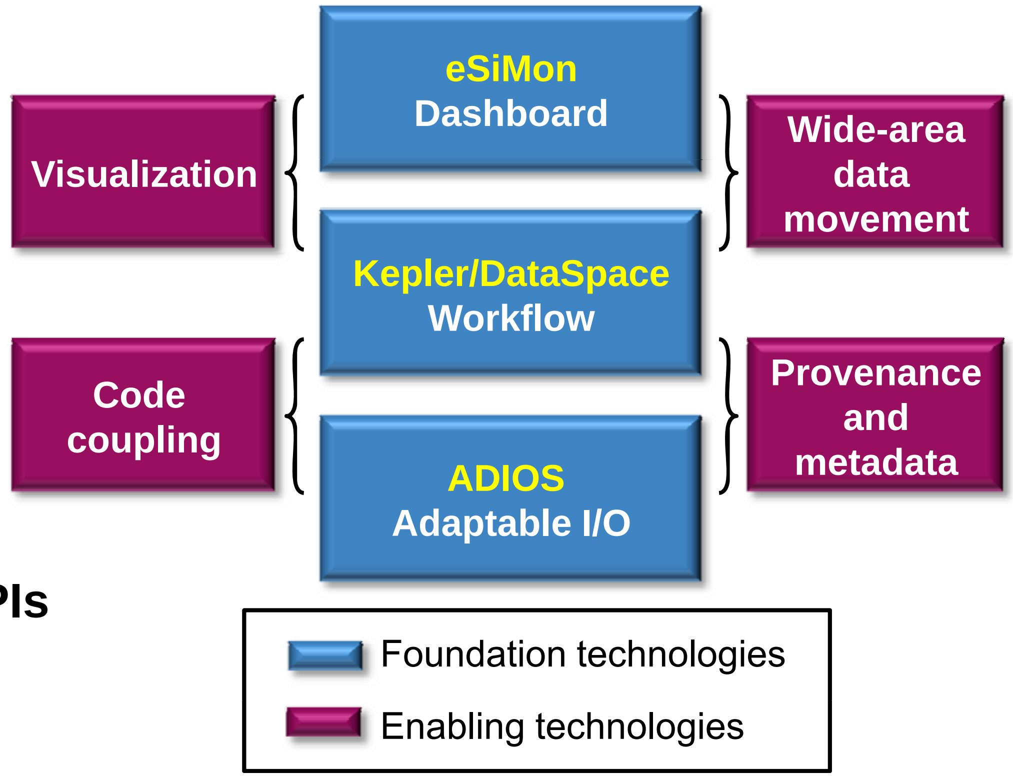


Extreme scale integration faces different challenges

- There are advantages, if adequately used:
  - Code couplings can be established in memory without leaving the HPC platform, while still keeping independence of each code executables.
  - Coupled simulation can finish in short time (ideallly, $\propto$ 1/flop speed)
- Requires large restart file I/O:
  - Before ADIOS, 2Tb XGC1 restart file took > 1 hour for every hour of run on 196,608 process cores (using parallel HDF5).
  - Adios (Adaptive I/O) in EFFIS: ~40GB/s: takes ~ 1m
- Smaller memory size per process core:
  - XGC is now down to <0.3Gb/core
- Expensive data movement
  - Data localization in XGC
- Framework should be able to support extreme scale computing and coupling: EFFIS is developed for this purpose.
- Inhomogeneous computing will be inevitable
- Fault tolerance needs to be built-in.

Large scale integration framework must handle, in real time, the large scale I/O data, coupling data, data movement, analysis, and visualization

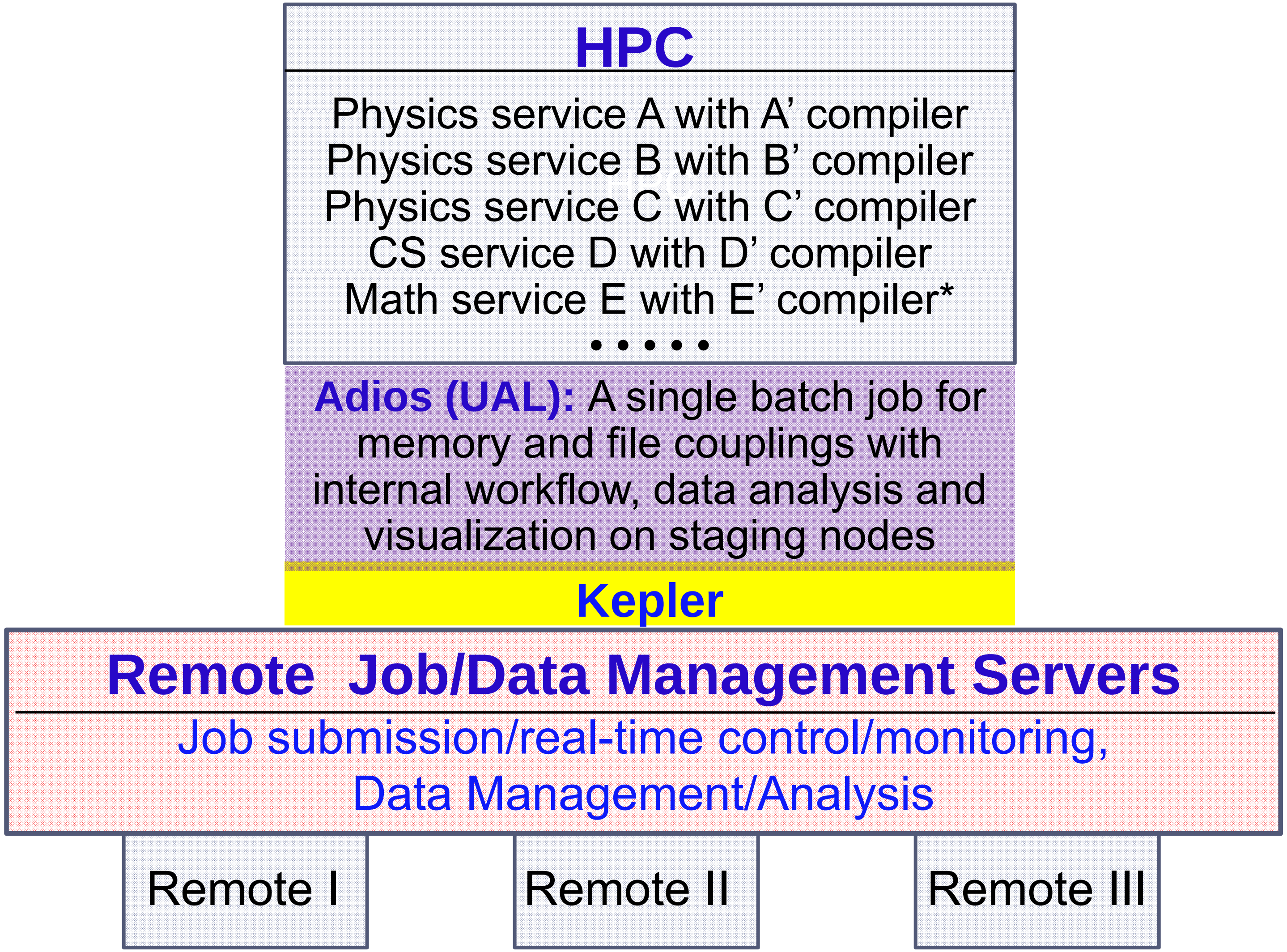
→ EFFIS



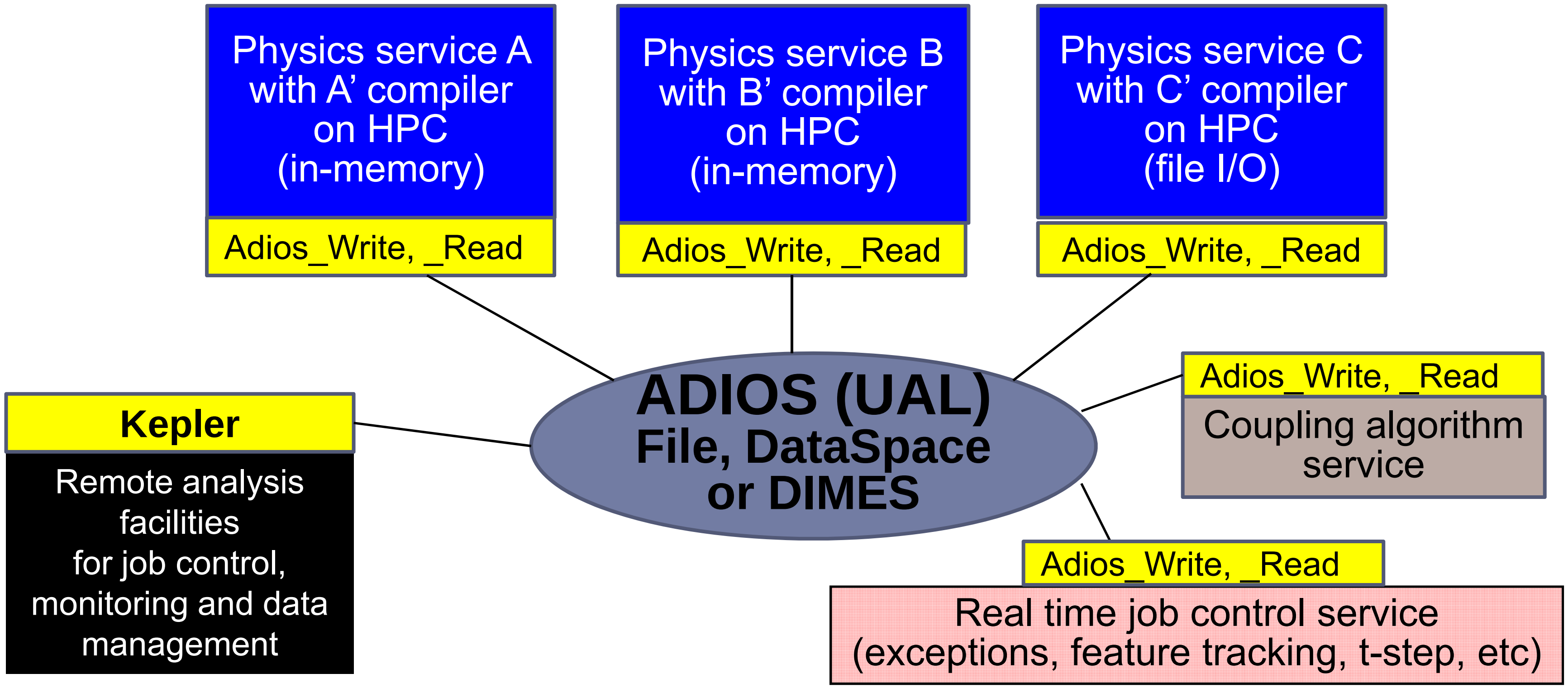
Need to lean only a dozen ADIO APIs

Adios_open	Adios_f_open
Adios_close	Adios_g_open
Adios_init	Adios_read_var
Adios_finalize	Adios_g_close
	Adios_f_close

EFFIS Design in Service Oriented Architecture
(End-to-end Framework for Fusion Integrated Simulation)



EFFIS framework is a convenient tool-set for SOA code integration (Example)



Fast I/O & data movement tools

Conclusion and discussion

- Development of extreme scale integration tools "EFFIS" for multi-physics simulation is moving along in CPES.
  - Take advantage of extreme scale computers: Perform as much multi-physics in full-f kinetic code.
  - However, still requires integration of separate executables for experimental time scale simulation including MHD and RF physics: spatio-temporal multi-scale.
  - Our current effort is more emphasized on ADIOS than Kepler
    - In-memory coupling operation in DataSpace
    - Real-time job control capability is in operation (first version, parameter injection)
- Challenges at large
  - Data size is becoming extreme. More advanced data management, analysis, and visualization tools are needed.
  - Fault tolerance is a universal issue, but heavier for integrated simulations. Proper tools needed in service oriented architecture.
  - Smart use of inhomogeneous computing
  - More …