Feb. 22-23, 2011, Houston, TX
Exascale data management, Analysis, and Visualization

Needs for Extreme Scale DM, Analysis and Visualization in Fusion Particle Code XGC*†

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

[1]Courant Institute of Mathematical Sciences, NYU; and PPPL
[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.

Acknowledgement

We thank the **GTC** and **GEM** teams for working together in the CPES project to produce the new code XGC full-f code in magnetic separatrix geometry. We also thank the analytic theory, computer science, performance engineering and applied mathematics teams in working together with the numerical physicists as one united team to make the XGC project successful.

Center for Plasma Edge Simulation
Chang (NYU)

- **Advisory Committee**: M.Greenwald, chair; J. Hesthaven; H. Wilson
- **Consultant**: P. Diamond; W. Lee

Physics — Chang (NYU)

- **Kinetics** — Parker (Colorado U)
  - Theory: Hahm(PPPL), Hinton, W. Lee
  - Comp: Lin(UC Irvine), Chen, Ku, Park, Xiao
- **Neutral/Wall** — Stotler (PPPL)
  - Neutrals: Stotler(PPPL)
  - PSI & Atomic Physics: Schultz (ORNL), Pindzola, Krstić, Reinhold
- **MHD** — Sugiyama (MIT)
  - Theory: Strauss (NYU, consulting)
  - Comp: Sugiyama (MIT), Pankin
- **V&V** — Kritz (Lehigh U)
  - Validation: Bateman (Lehigh), E. Feibush
  - Verification: Weitzner (NYU)
- **Coupling Integration** — Samtaney (PPPL)

CS and Applied Math — Shoshani (LBNL), Keyes (Columbia U)

- **DM/Viz** — Klasky (ORNL)
  - Code Integration: Parahsar (Rutgers)
  - Workflow: Shoshani (LBNL), Podhorszki
  - Visualization: Klasky (ORNL), Tchoua
  - Data Access: Schwan (Georgia Tech), M. Wolf
- **Algorithms** — Greengard (NYU)
  - Linear Solvers: D'Azevedo (ORNL), Adams
  - Other Algorithms: Zorin(NYU)
- **Performance** — Worley (ORNL)
  - Parallel Algorithms: Ethier (PPPL)
  - Application specific: Cummings (Caltech)

NOTES:

Each participants are listed according to the main work assignment, but they also participate in other group activities.

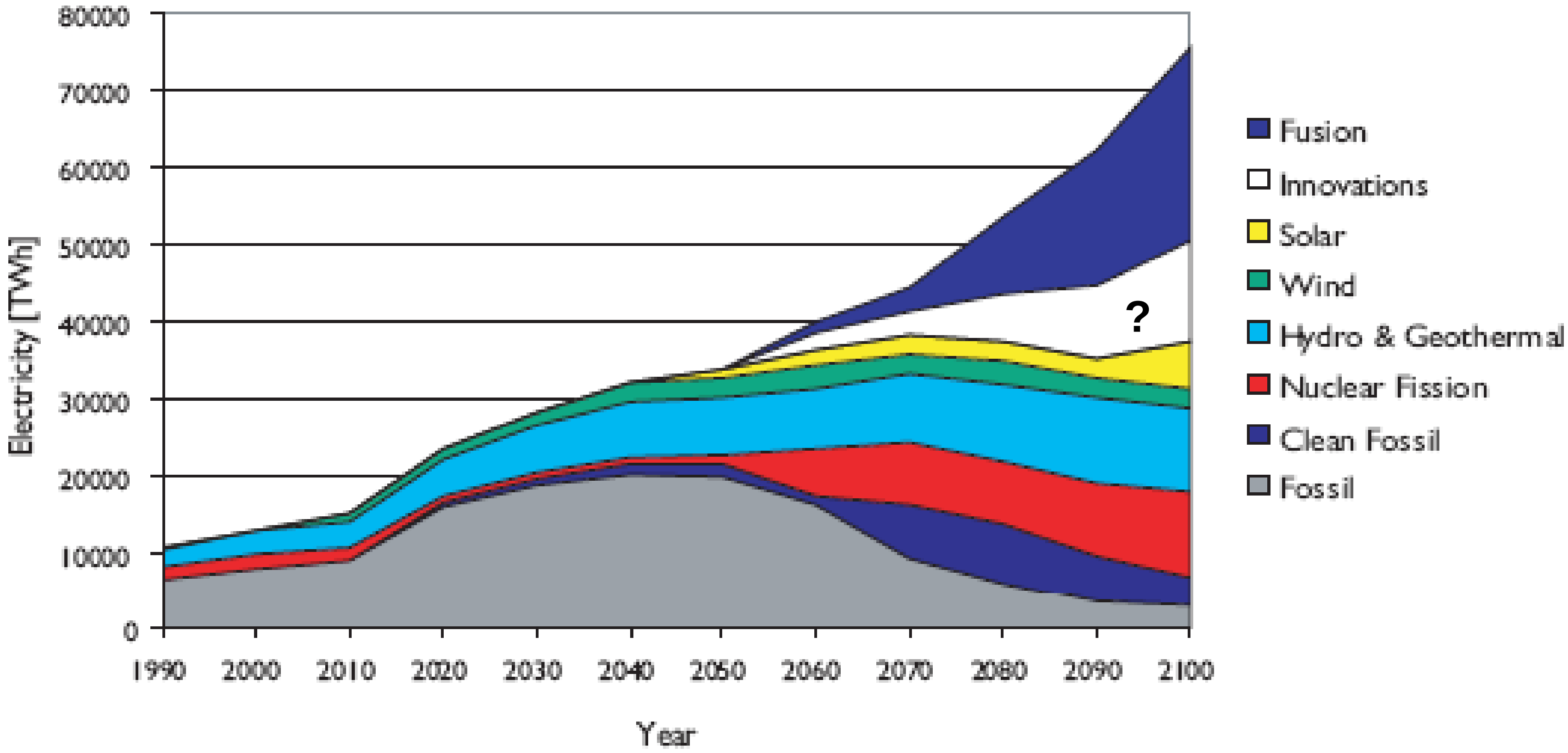
(9 Universities, 3 National labs, and 1 company)

Unpaid collaborators e.g., Samatova, Ma, etc

Outline

- A short introduction to fusion plasma simulation
- Full function particle-in-cell code (XGC1)
- A peak at scientific discovery made by XGC1
- Large I/O problem, rescued by Adios
- Analysis bottleneck imposed by turbulence data I/O
- Simulation bottleneck imposed by the capacity of HPC
  - Electron scale turbulence and ITER size plasma
- What can we expect at 10PF, 100 PF and 1 EF?
- Exa-scale dream
- Conclusion

Energy scenario for the world electricity up to the year 2100
(Source: the Research Institute of Innovative Technology for the Earth, Tokyo)



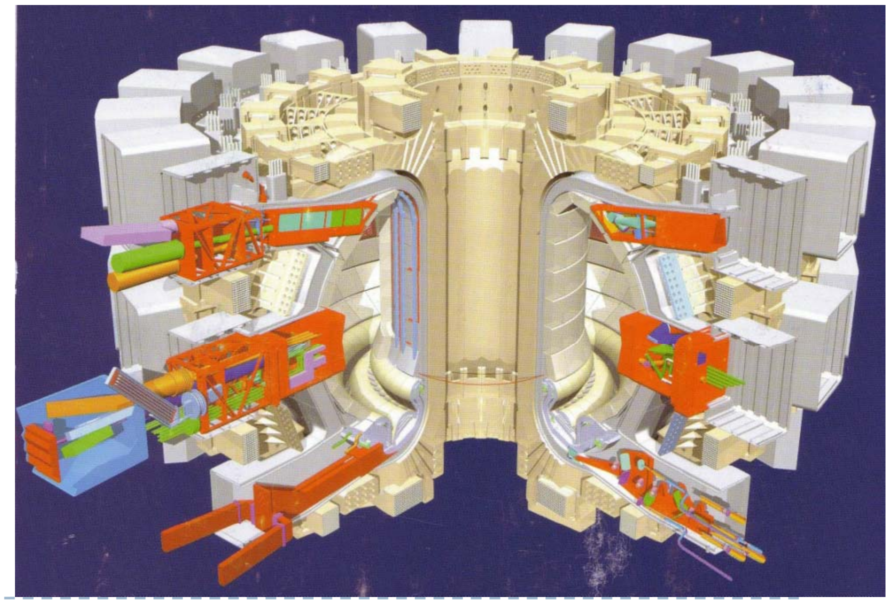
There is a limit on how fast human can build the nuclear and fusion reactors.
With all the energy resources being developed, we are still short significantly.
We have a serious problem!

Fusion Energy of Stars in a Magnetic Donut Reactor

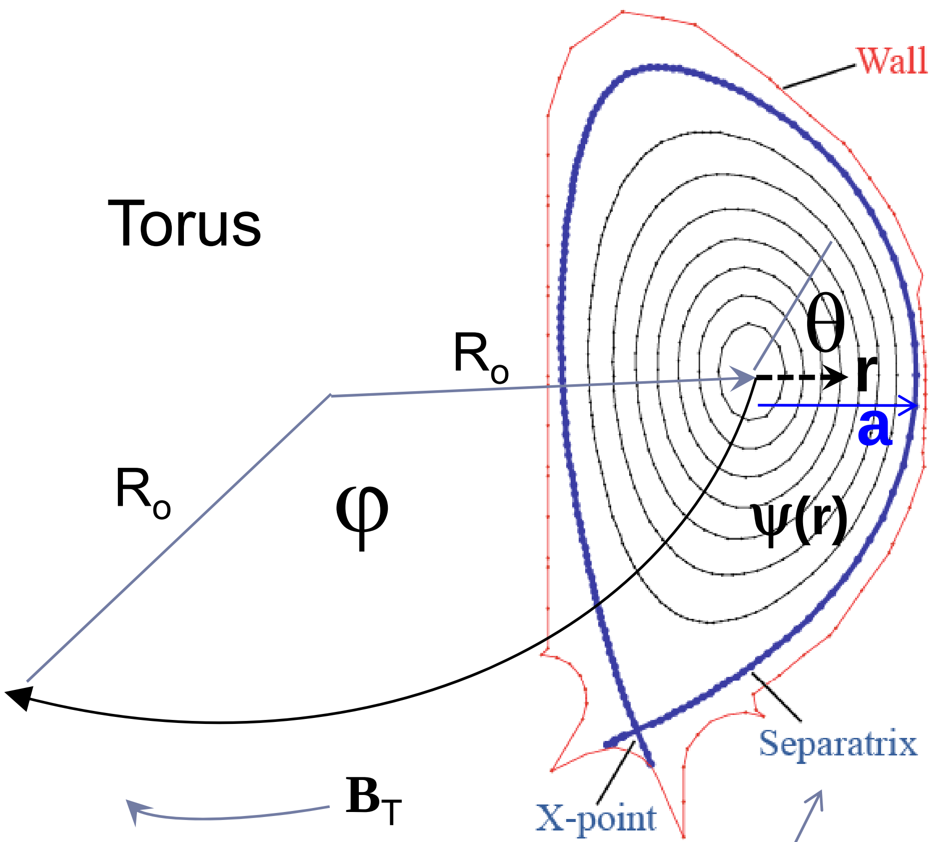
- Tokamak fusion plasma is a gaseous system of hot charged particles (D^+, T^+ and e^-), immersed in strong donut-shaped magnetic field with separatrix and X-point (diverted geometry), surrounded by material wall.
- When D^+ and T^+ ions are >10 keV, they fuse together to form α particles and release energetic14 MeV neutrons ($E=\Delta m\ c^2$) for electricity generation.

The world has begun construction of an experimental fusion reactor ITER in France

ITER partners: USA, Russia, EU, Japan, China, Korea, India



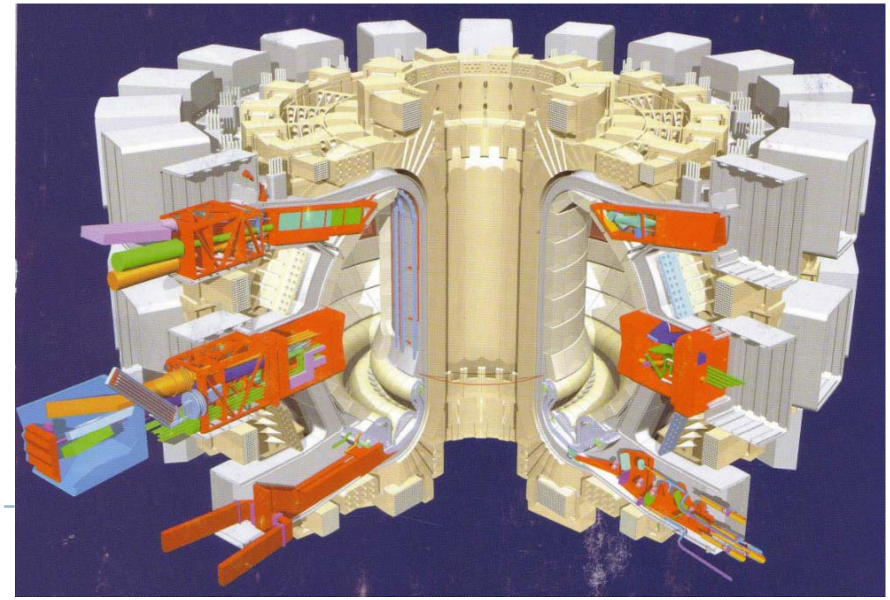
Tokamak geometry



Poloidal cross-section (poloidal plane) at a constant toroidal angle

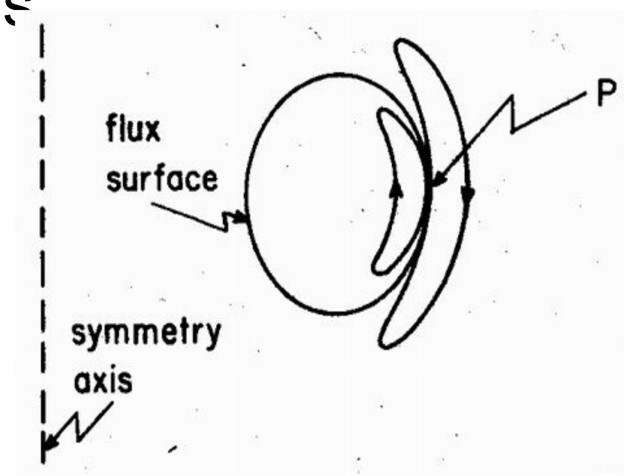
Poloidal magnetic flux label is a minor-radial coordinate
$\psi(r)$: 1 at r/a=1, 0 at r/a=0

Difficult to simulate: XGC is unique.



First-principles kinetic fusion plasma simulation

- Tokamak plasma: multi-scale, multiphys
  1) **Classical collisional transport enhanced by large scale banana drift motions from curved B field (neoclassical transport)**
  2) **Micro-scale turbulences and experimental time-scale loss of plasma**
  3) **Neutral transport and atomic physics**
  4) **Multi species**
  5) Abrupt macro-scale instabilities and instantaneous loss of plasma



- Assuming that the "5) abrupt macro-scale instabilities" are controlled in a fusion reactor

  → **A full gyrokinetic simulation can perform 1) - 4) in multi-scale, multiphysics in realistic diverted geometry in contact with material wall.**

  → **Guide ITER and fusion reactor design and operation**

  → **Requires extreme scale HPC.**

Gyrokinetic simulation of fusion plasma

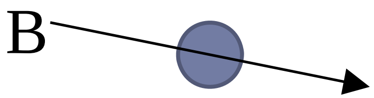
- **The original plasma kinetic equation in its full capacity demands exa-scale HPC, projected from today's PIC technology and hardware/software.**

  Vlasov eq. with Fokker-Planck v-space collision operator C(f)

  $\boldsymbol{\partial f/\partial t + v \bullet \nabla_x f + e(E+vxB) \bullet \nabla_v f = C_v(f),}$ 6D (3D **r** + 3D **v**) , and

  Maxwell's equation In 3D r-space.

- **Gyrokinetic:** Reduce 6D (x,y,z,v_x,v_y,v_z) to 5D $(x,y,z,v_{||},v_{\perp})$ by assuming that the gyrofrequency is much faster and that the gyroradius is much shorter than the space-time scales of interest.



  - Enables first principles simulation at << exa-scale, enhancing science/numerical fidelity as HPC capacity increases.

Gyrokinetic codes for fusion plasma

- **Two complementary approaches**
  - **Continuum:** solves **5D** PDE system on **5D Eulerian** grid
    - Pro: The noise issue does not enter.
    - Difficult for large-scale parallel computing (CFL limit, memory intensive) Currently, optimized for ≤1,000 processor cores
  - **Particle-in-cell:** solves statistically distributed marker particle dynamics in 5D **Lagrangian** position-velocity space. Solves Maxwell's equations on **3D** position grid
    - Pro: Well-suited for leadership-class computing (larger device or higher resolution physics → more grid→ more # particles → more # cores)
    - **3D** grid instead of **5D,** and multiple domain decomposition, significantly reduces memory requirement (≤0.3 Gb per core in XGC1, going down further).
    - Con: Statistical particle noise 1/Sqrt(N) → Smoothing or large enough N, with convergence and sensitivity studies, is needed.

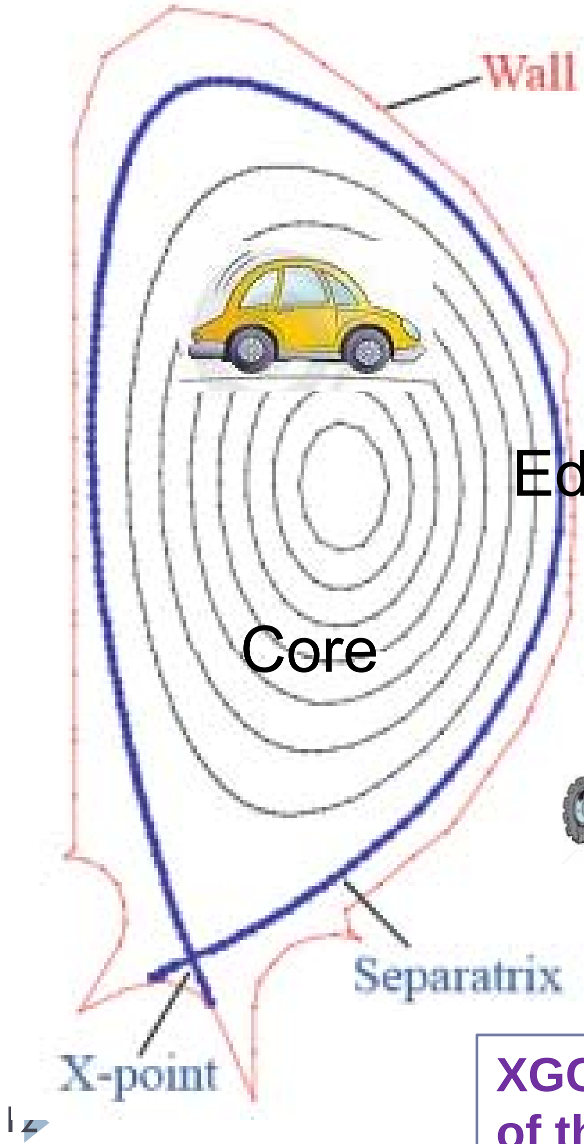
Perturbation approach vs full function approach

- **Perturbation approach: developed for economy when computing was not so powerful.**
  - $\delta f = f_{full} - f_0$, assuming conserved thermodynamic equilibrium system
  - δf codes only calculate perturbed turbulence physics
  - Assumes that f_o is unaffected by δf and vice versa, hence loses the multi-scale interactions between f_0 and δf
  - Cannot handle a non-equilibrium plasma in contact with wall. Thus, whole volume simulation is not possible.
- **Full function approach: full-f for complete physics**
  - Mean and perturbed physics are simulated together in multi-scale
  - Can handle non-equilibrium plasma in contact with **wall**
  - Requires ~100x more HPC power than δ**f** approach

Why not partner up with extreme-scale HPC development plan, and take the full-function approach?

→ **XGC project on extreme-scale HPC:** unique in the world fusion program

Magnetic separatrix prohibits not only δf approach, but also a shortcut coordinate system→ HPC*HPC



Whole volume simulation should include open-field **edge**: Many experimental evidences exist for critical non-local core-edge interactions.

Magnetic separatrix and magnetic axis are singular surfaces for core codes which use the easy-to-handle "magnetic" coordinate system.

Thus, all the other US gyrokinetic codes stay in the core

- at a safe distance inside the magnetic separatrix surface.
- At a safe distance from magnetic axis
- using **delta-f** perturbed simulation

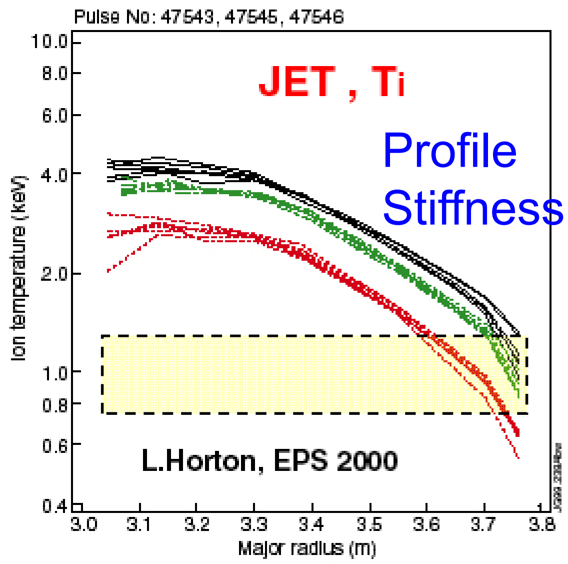
Removal of "magnetic" coordinate system → expensive

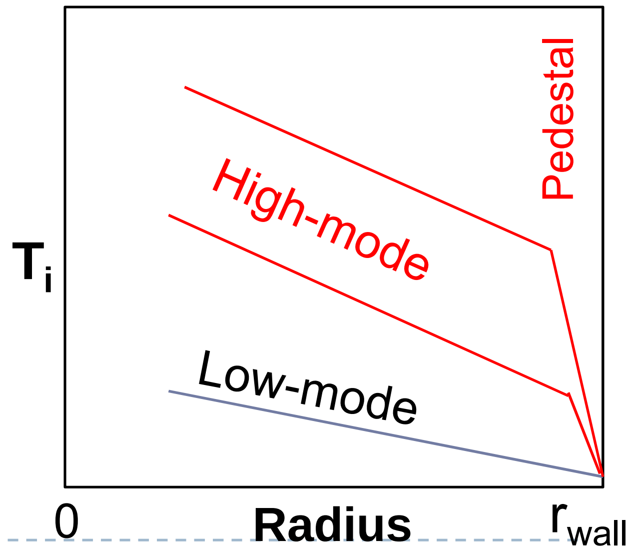
XGC on cylindrical grid is the only kinetic code capable of the whole-volume simulation in full-function.

Good edge confinement leads to hot core plasma

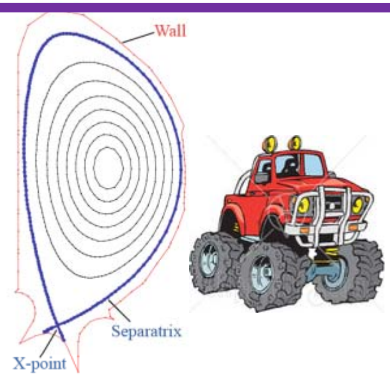
- Plasma near material wall must stay cold (~100eV)
- Plasma in the central core must be hot enough for successful fusion (>10 keV)
- Temperature-slope is limited by turbulence
  - T_i is too low in fusion core if in L-mode (≤1980)
- ITER assumes H-mode pedestal
  - **Strong core-heating is necessary**
  - **Short propagation time (<< τ_{conf}) of the edge→core confinement properties**
  - **Stiff T_i profile**
- This physics must be understood

Whole-volume full-function kinetic simulation is needed.→ **Requires extreme-scale HPC → XGC project**





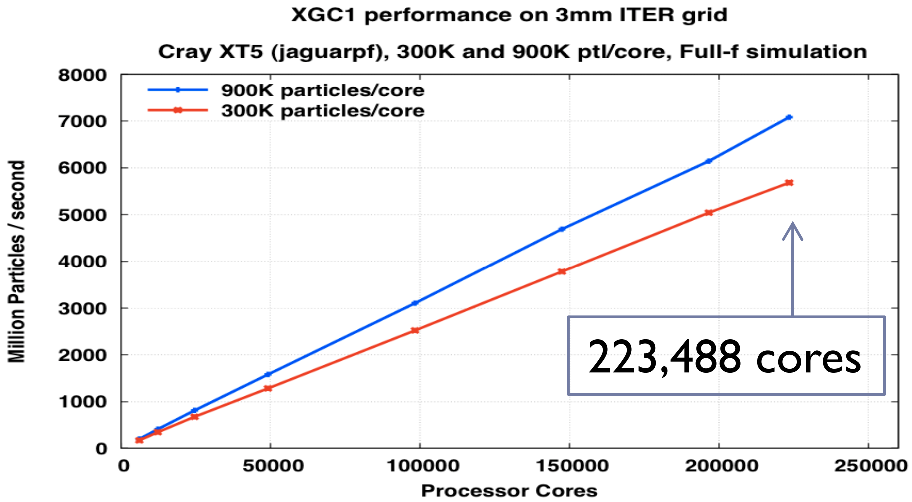
XGC1 gyrokinetic PIC code (50M Incite hrs at OLCF)
(Currently on 100K-220K cores, aimed for 1 wall-clock day)

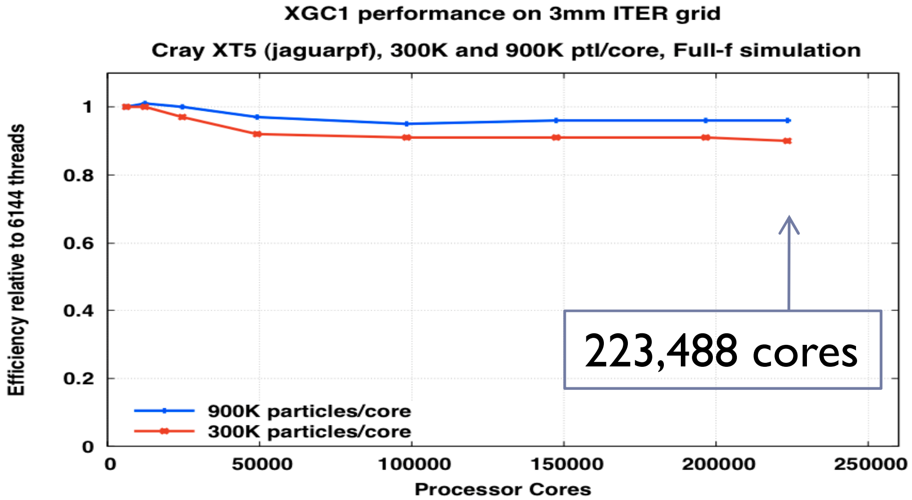


- **XGC1: full-function, X-point included Gyrokinetic Code in realistic tokamak geometry across magnetic separatrix surface**
- **Spatial simulation domain:** whole tokamak plasma volume with realistic tokamak edge geometry and Dirichlet wall boundary condition (grounded wall).
- **Unstructured triangular grid. Particles advance in cylindrical coordinate. Field solver on B-following grid.**
- **Capability in hand: Electrostatic ion turbulence dynamics without scale-separation from mean plasma dynamics, with heat source and conserving Coulomb collisions**
  - full-f ions and adiabatic background electrons
  - delta-f ions (for verification against other delta-f simulations)
  - Limited full-f electron capability (requires more powerful HPC)
- **Memory localization: $\leq$ 0.3 Gbytes per core, going down further**
- **GPU hybrid is under investigation**
- **Capability under development**: **Full-function electromagnetic turbulence**
  - Current electromagetic capability is in delta-f
- **Developed initially using δf GTC technology, enhanced by δf GEM technology**.

XGC1 Scales efficiently to the maximal number of Jaguar cores

12 cores per node, 2 MPI processes per node





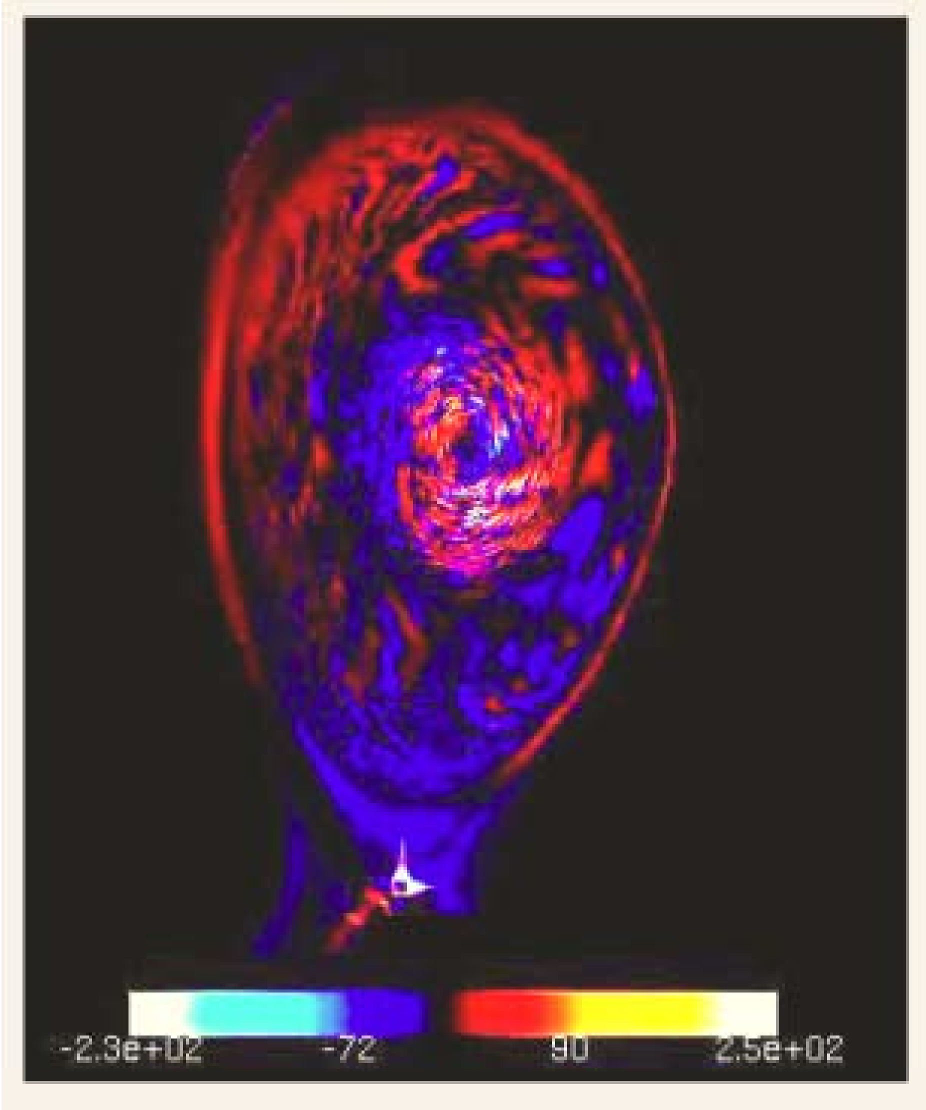
- **900K particles per thread problem is more computationally intensive than 300K problem, which leads to ~20% higher particle push rate.**
- **Performance scaling is excellent for both problems.**

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

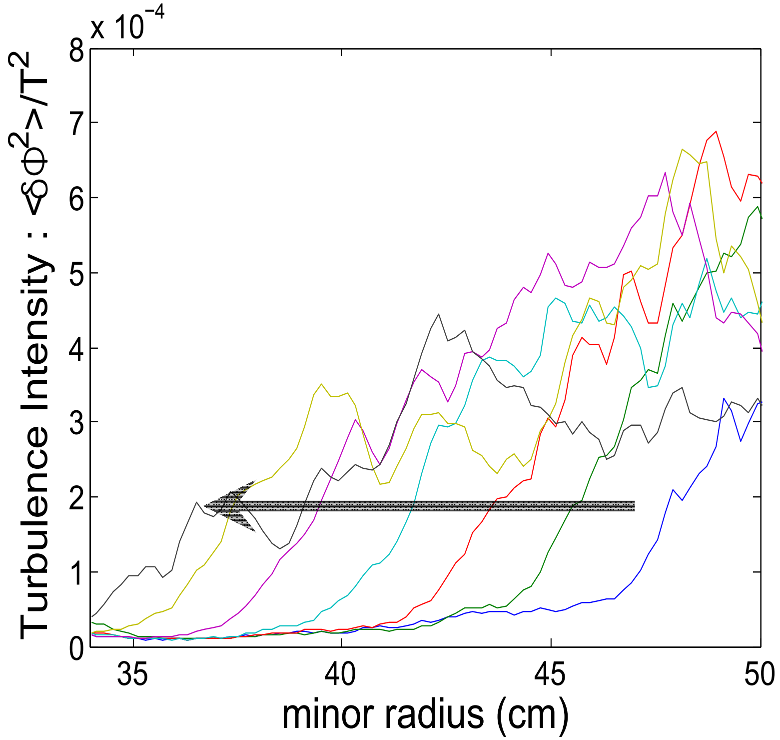
	Average Job Size in Cores	Utilization in Core-Hours
Jobs requesting <20% of the available resources	3,079	8,446,978
Jobs requesting between 20% and 60% of the available resources	66,474	3,042,575
Jobs requesting >60% of the available resources	170,304	14,311,232

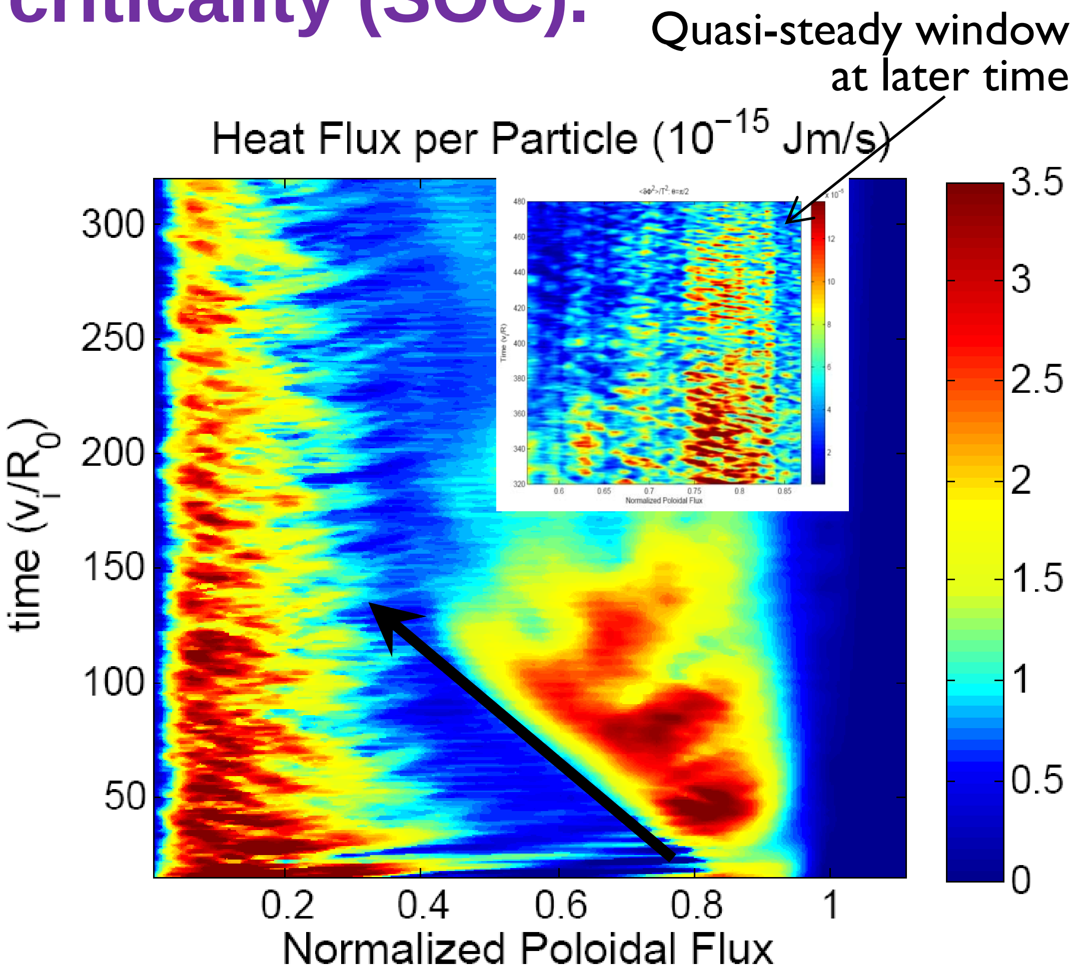
Whole-Volume, full-f ITG Simulation for DIII-D

- ITG (Ion Temperature Gradient) driven turbulence is the most robust and fundamental micro-turbulence in a tokamak plasma.
- Includes diverted edge geometry and magnetic axis
- Realistic Dirichlet BD condition Φ=0 on conducting wall.
- Heat source in the central core
- This type of simulation is possible only on extreme HPCs → needs to push the edge of future HPC
- Several new scientific discoveries have already emerged.



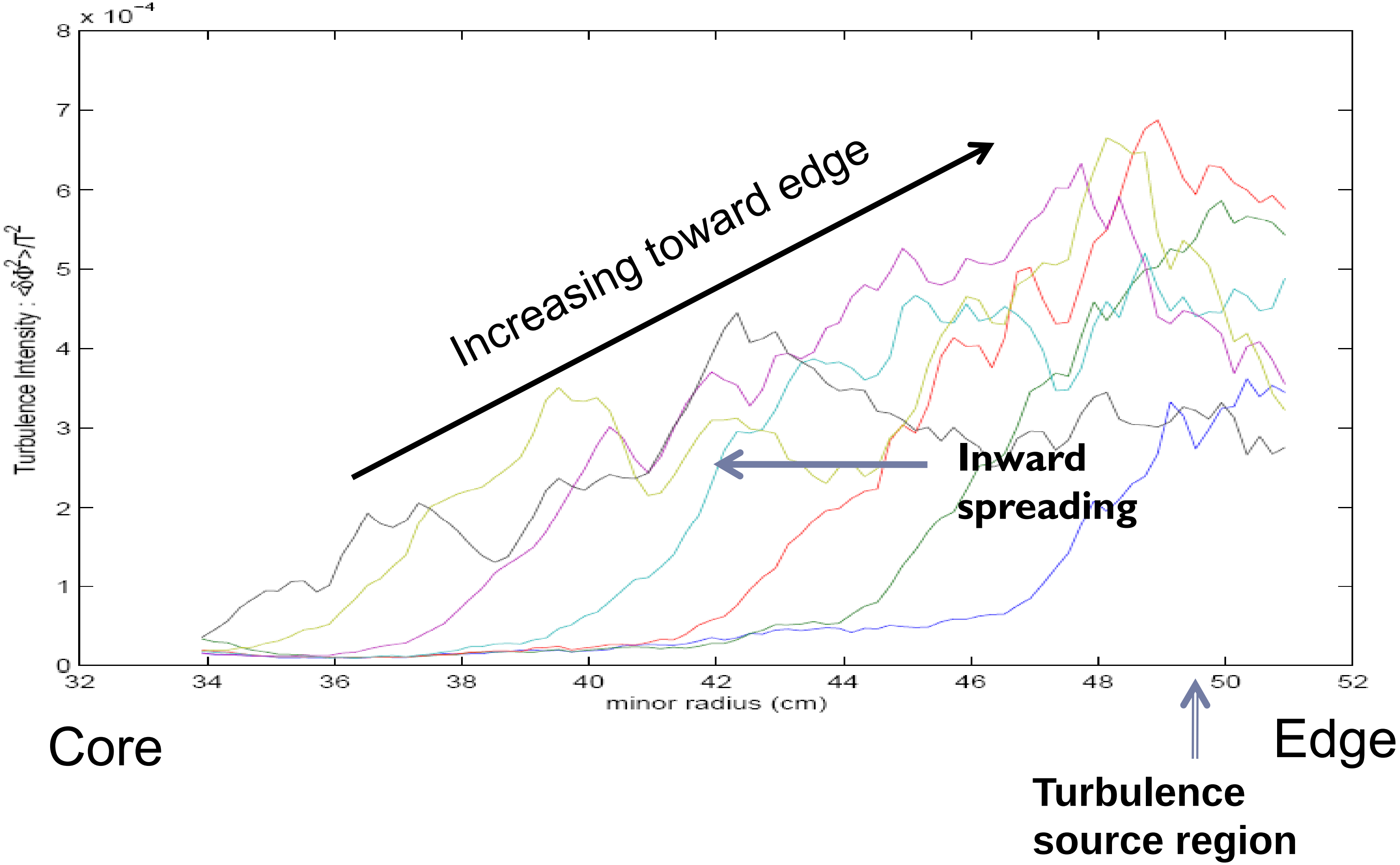
Edge turbulence propagates deep into the core and self-organizes the global temperature profile to criticality (SOC).





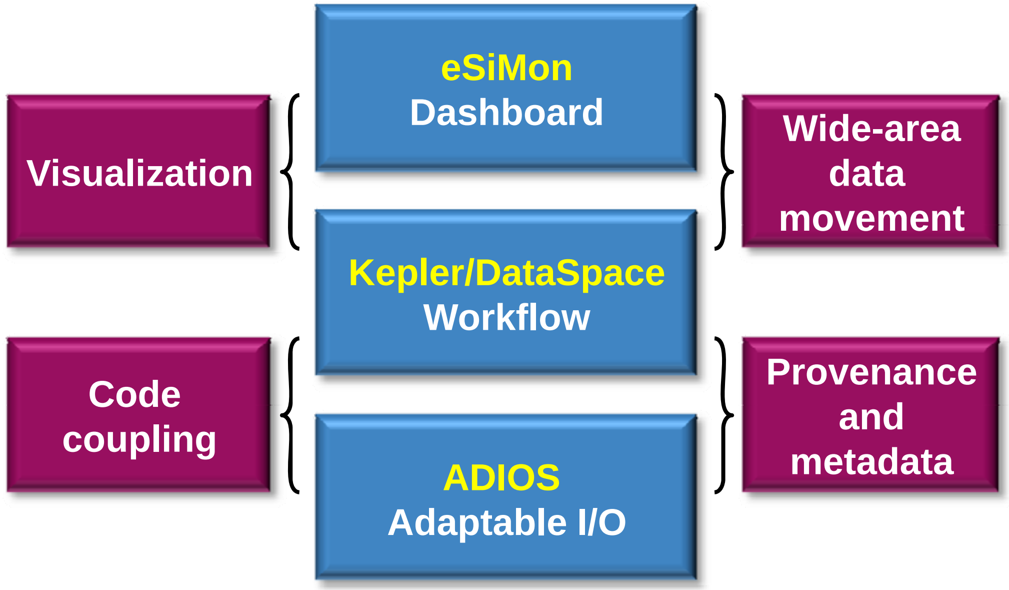
As edge turbulence arrives, local turbulence is aroused/modified and induces adequate heat flux to yield self-organized criticality.

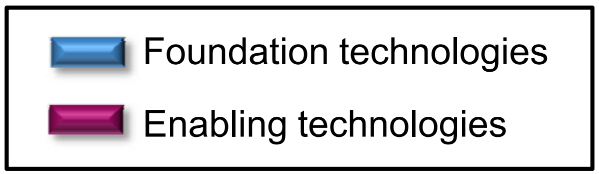
Self-organized $\delta\Phi^2/T^2$ increases toward the edge.
Seen in experiments, but unexplained for 30 years.



XGC1 deals with large scale I/O, code-coupling, analysis, and visualization

→ Our CS team responded and developed EFFIS



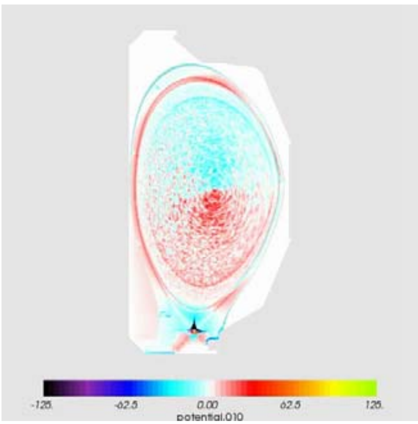


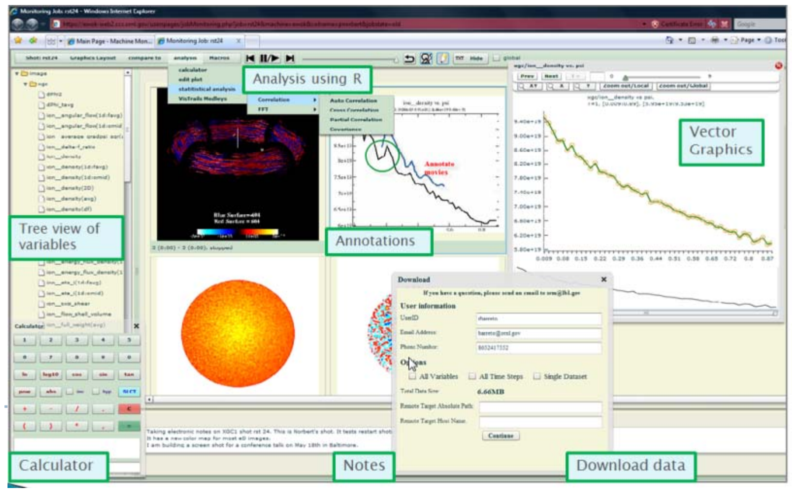
Need to lean only a dozen ADIO APIs

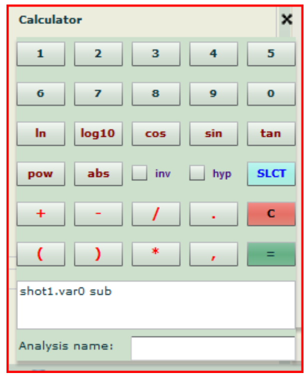
Adios_open	Adios_f_open
Adios_close	Adios_g_open
Adios_init	Adios_read_var
Adios_finalize	Adios_g_close
	Adios_f_close

Example 1: Peta-scale XGC1 research is enabled by data management and visualization in EFFIS

- Before ADIOS, 2Tb restart file was taking > 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 for 2Tb restart file
- Before EFFIS, the job originator and the collaborators had to wait until the long simulation was finished and/or the large size data was moved before they could monitor/analyze the result.
  - With **Adios I/O, Kepler workflow**, **DataMover-lite**, and **eSimMon Dashboard** in **EFFIS**, the job originator and the collaborators can monitor/analyze the data in real time on their laptops..
  - Still needs further development



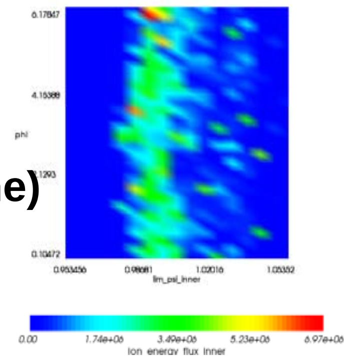




Example 2: Weakly coupled Kinetic-MHD simulation for pedestal-ELM cycle enabled by automated EFFIS

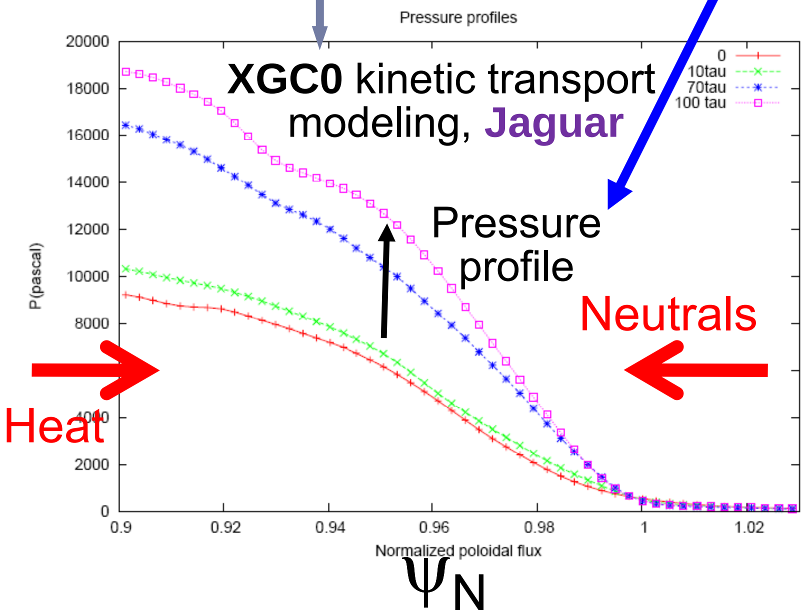
Linear stability check (Binary **Elite**)
(**file coupling**)

Divertor Heat Load
(2010 OFES Milestone)

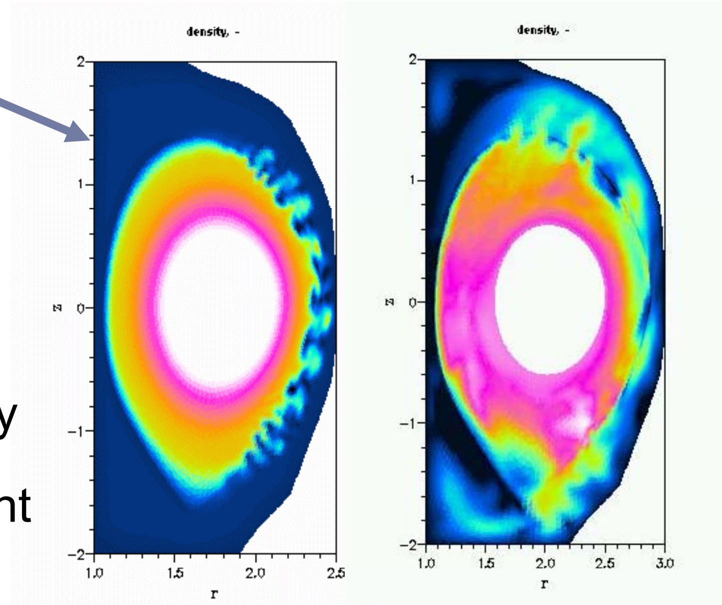


B-reconstruction and mesh interpolation by **M3D-OMP**
(**file coupling**)

ELM crash in extended **M3D-MPP**,



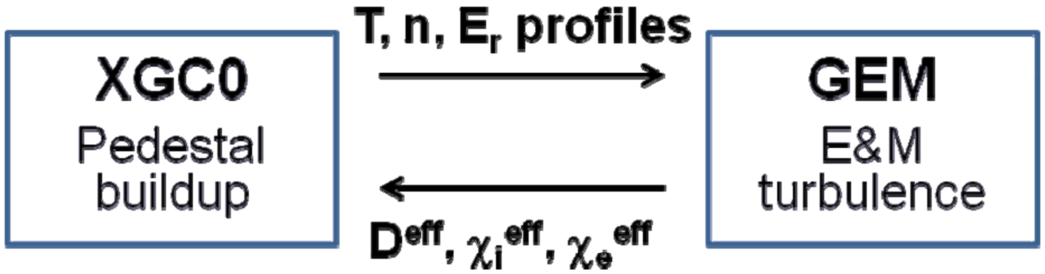
ADIOS
In memory coupling of turbulent B and E



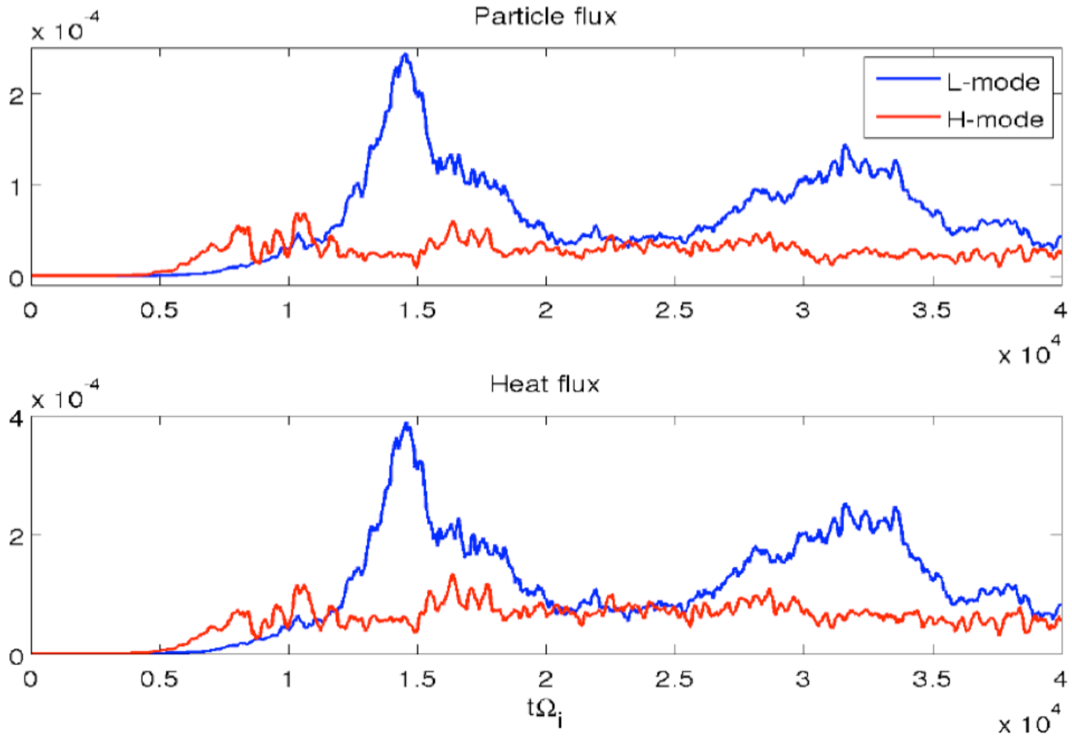
T=76
saturation

T = 496
relaxation

Example 3: Weak coupling between particle and particle codes (XGC0 and GEM) for electromagnetic turbulence transport in the edge pedestal evolution



GEM imports a radial section of XGC0 plasma inside the magnetic separatrix surface



Particle and ion heat fluxes in L- and H-mode XGC0 plasmas with a DIII-D Experimental magnetic equilibrium.

Example 4: Strong coupling between particle and PDE codes to study RMP penetration and plasma evolution

XGC0 evolves plasma profiles under new δB and calculates new perturbed current $\delta J_{||}$.

$\delta J_{||}$

Solver code calculates new 3D δB from perturbed current $\delta J_{||}$ evaluated in XGC0

$\delta\vec{B}$

- Solve two coupled systems

  $\delta j_{||}/B = F(\delta\psi)$:Vlasov-Poisson system (XGC0)

  $\Delta^{*}\delta\psi = \mu_0\, \mathrm{I}\, \delta j_{||}/B$: Perturbed magnetic field solver.

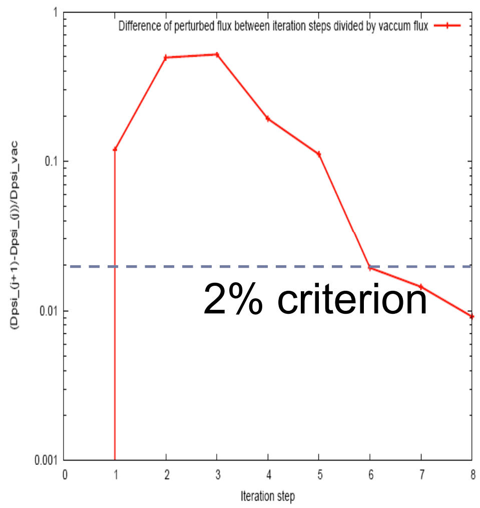
- Use damped iteration scheme

  $\delta\psi_{k+1,(m,n)}(r_i) = \delta\psi_{k,(m,n)}(r_i) + s_{(m,n)}\, \Delta\psi_{k+1,(m,n)}(r_i)$

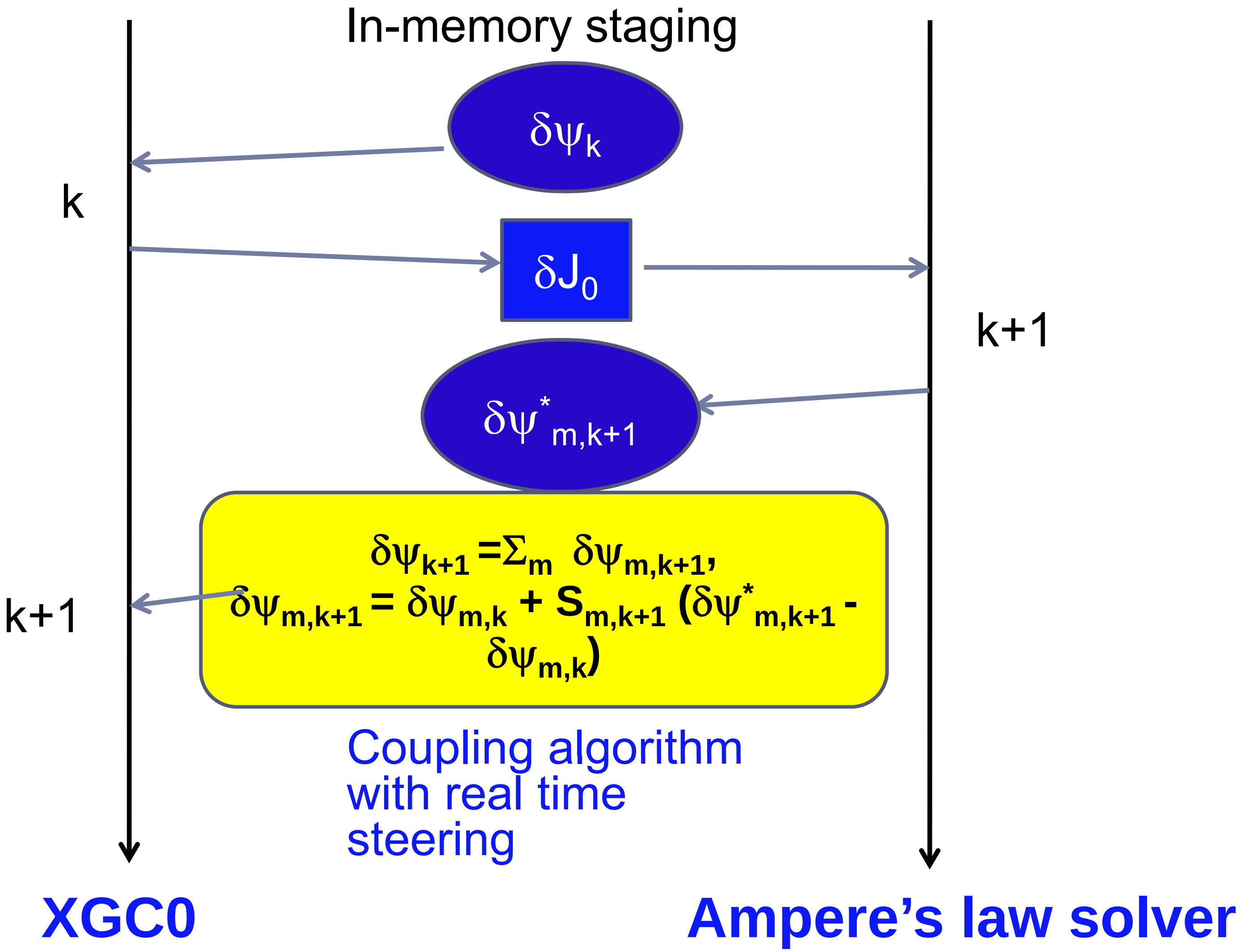
  $s_{(m,n)} = \mathrm{Min}_{r,m}\, [1, \alpha\, \mathrm{Min}\, (|\delta\psi_{k,(m,n)} / \Delta\psi_{k+1,(m,n)}|)]$

  $\Delta\psi_{k,(m,n)}$ is the correction of $\psi_{k,(m,n)}$ at the k-th iteration step.

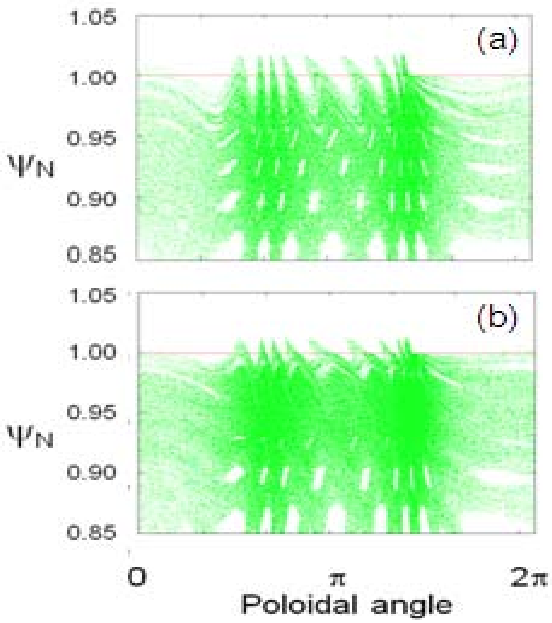
- Converged solution with the criterion $\Delta\psi/\ \delta\psi_{vacuum}$<2% is obtained in 7 iterations for the case studied here.

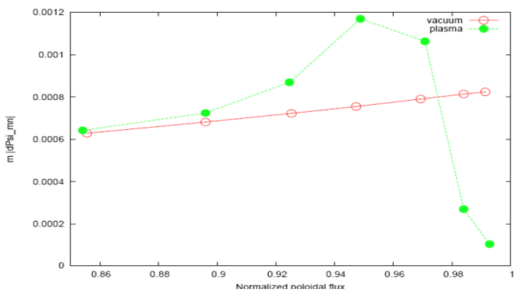


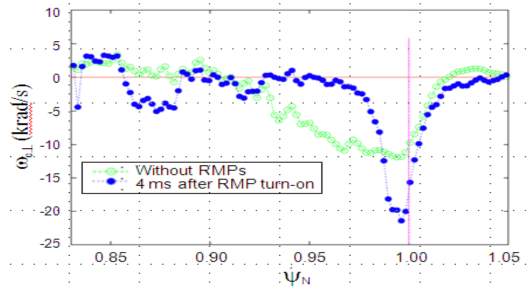
Strong Coupling: Damped Iteration Solution on EFFIS (Adios/DataSpace)



Example 4 cont'd: Strong coupling between particle and PDE codes to study RMP penetration and pedestal evolution



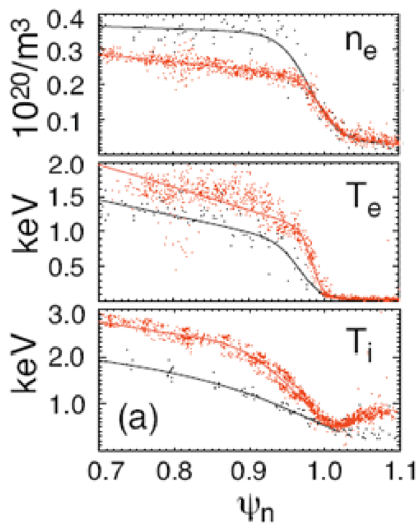


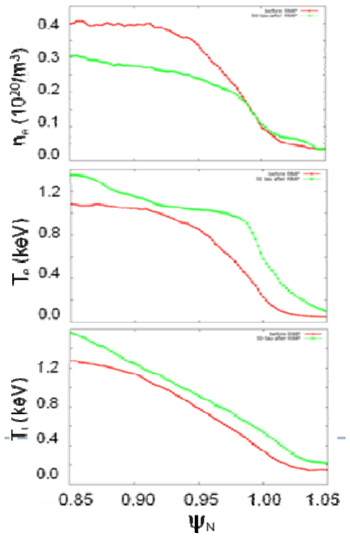


Amplitude of the resonant perturbation components

Electron perpendicular rotation speed (angular frequency) before and 4 ms after the RMP turn-on.

Predictions





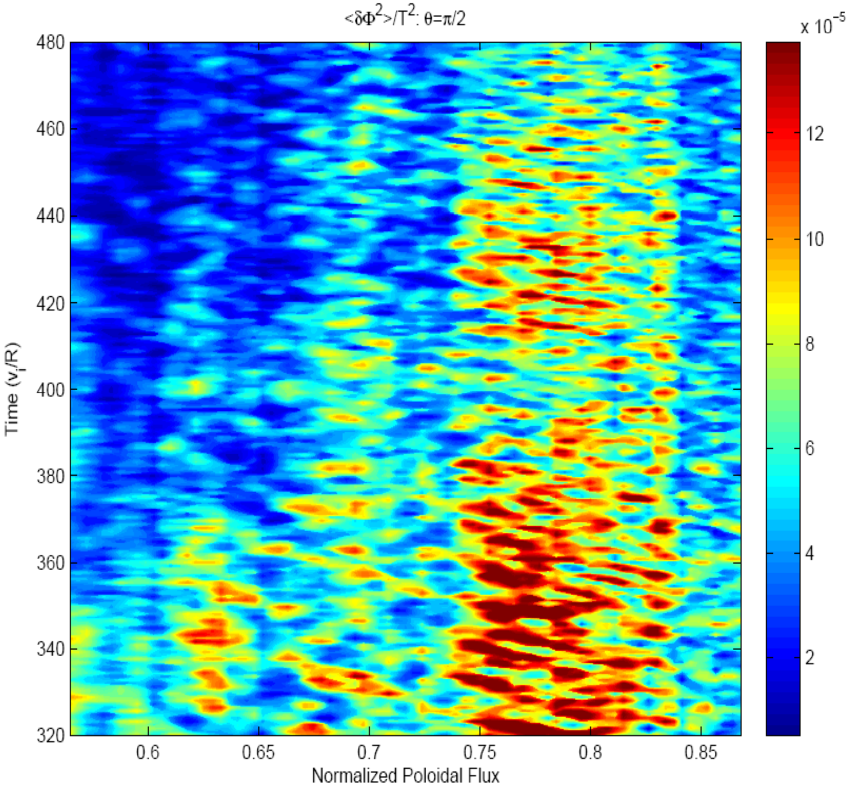
Left: Experiment:
Black is before and red is after the RMP turn-on.
Right: Simulation (4ms after RMP turn-on)
Red is before and green is after the RMP turn-on.

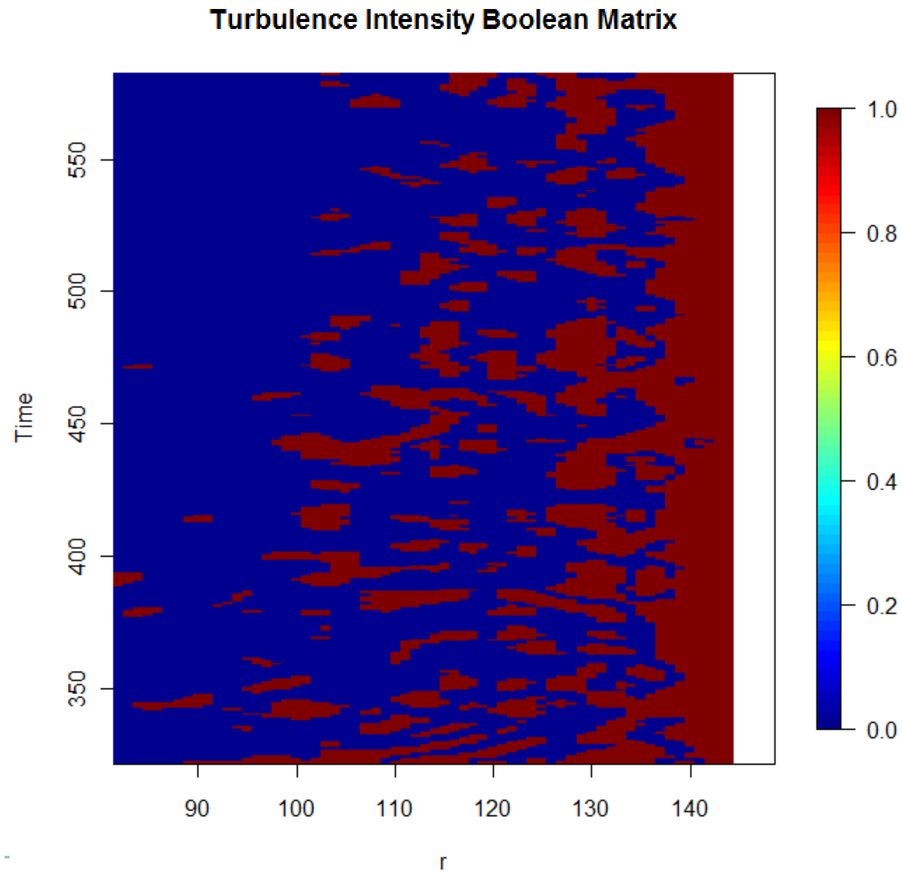
Validation

Example 5: Data analysis for scientific discovery:
We need to search for unknown discovery out of extreme scale data.

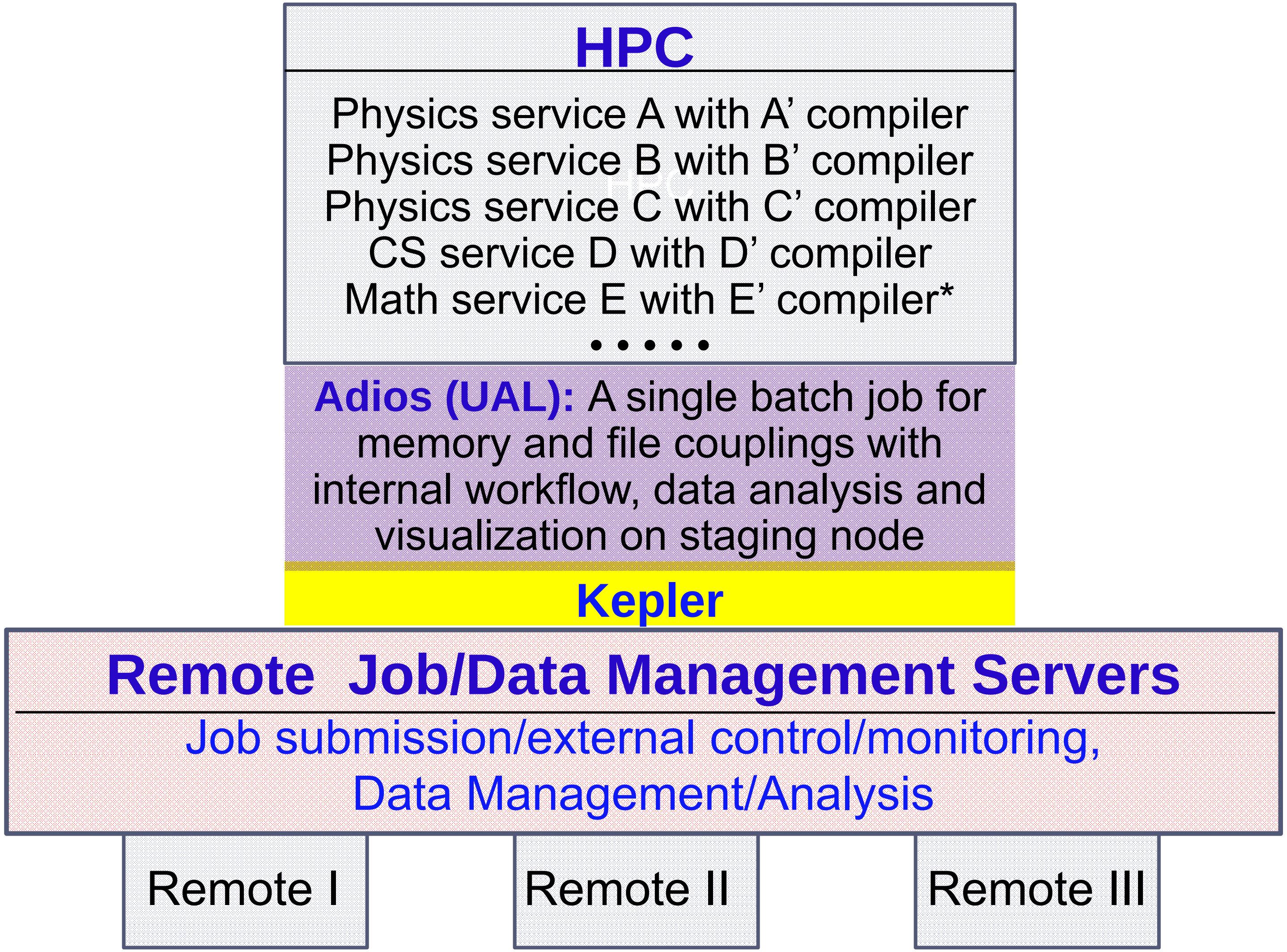
Boolean matrix representation of the dynamical turbulence intensity (N. Samatova)

How does the turbulence intensity behave to cause the non-local core-edge interaction and global SOC?: feature tracking and statistical study





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



Current, and growing problem, in data storage and analysis in XGC1 → new paradigm

- Current turbulence data from DIII-D device simulation is >100 Mbytes per time step. 100K time steps make the data file size to > 10 Tbytes
- More complete physics analysis demands the storage of all the particle data, >1 TB per step

  → >>1 PB total, approach 1 EB in the near future.
- Application scientists cannot spend "forever" to analyze a data file
  - Write out only small fraction of the data, in multiple files
  - However, scientific discovery often demands us to analyze unsaved data: unknown nonlinear multiscale physics
  - Re-run the expensive code. How many times?
- Moving toward smart (and in situ) data storage, management, analysis and movement.

Scientific discoveries expected in XGC on tomorrow's computing systems (per one day run)

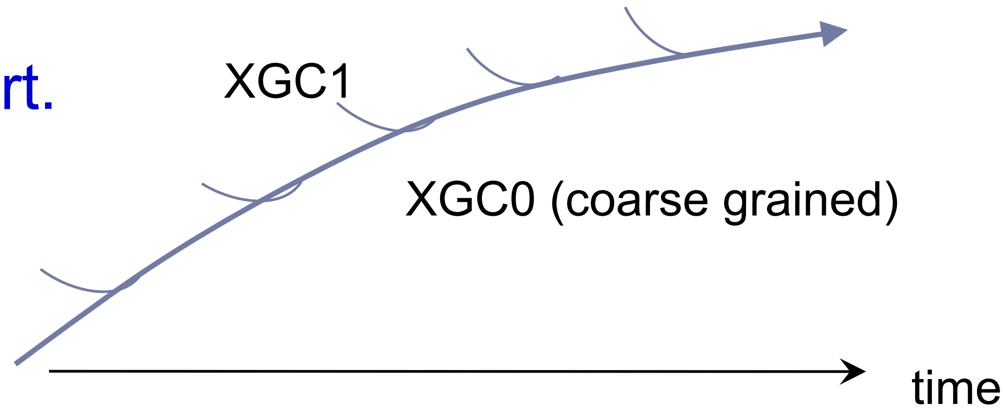
- **Research on these topic should start NOW!**
- **10 PF**
  - **Multi-scale gyrokinetic simulation of whole-volume DIII-D device to experimental time scale including ion scale electromagnetic turbulence (ITG, TEM, etc; full-f ions and delta-f electrons)**
  - **Multiscale gyrokinetic simulation of whole-volume ITER device to turbulence saturation time scale including ion scale electromagnetic turbulence**
- **100 PF**
  - **Virtual gyrokinetic DIII-D device, including electron scale electromagnetic turbulence**
  - **Multi-scale gyrokinetic simulation of whole-volume ITER device to experimental time scale including ion scale electromagnetic turbulence**
- **1 EF: dream come true!**
  - **Virtual gyrokinetic ITER device: Multi-scale simulation of whole-volume ITER device to experimental time scale, including electron scale electromagnetic turbulence**
  - **Full 6D multiscale electromagnetic simulation without gyrokinetic approximation, including kinetic and MHD phenomena**
- **>1 EF**
  - **Full 6D kinetic virtual ITER device to experimental time scale: Live in the dream!**

Large scale multi-scale coupling in XGC on tomorrow's computing systems → Experimental time scale

Mathematically tight kinetic-kinetic coupling in XGC, to extend the first-principles full-f simulation to experimental time scale, is where the development on **extreme scale in-memory data management, analysis, and visualization** (as well as the state-of-the-art applied math) are desperately in need.

- Lifting to coarse grained system for experimental scale time-marching
- Restricting to fine grained system for microscopic level fidelity

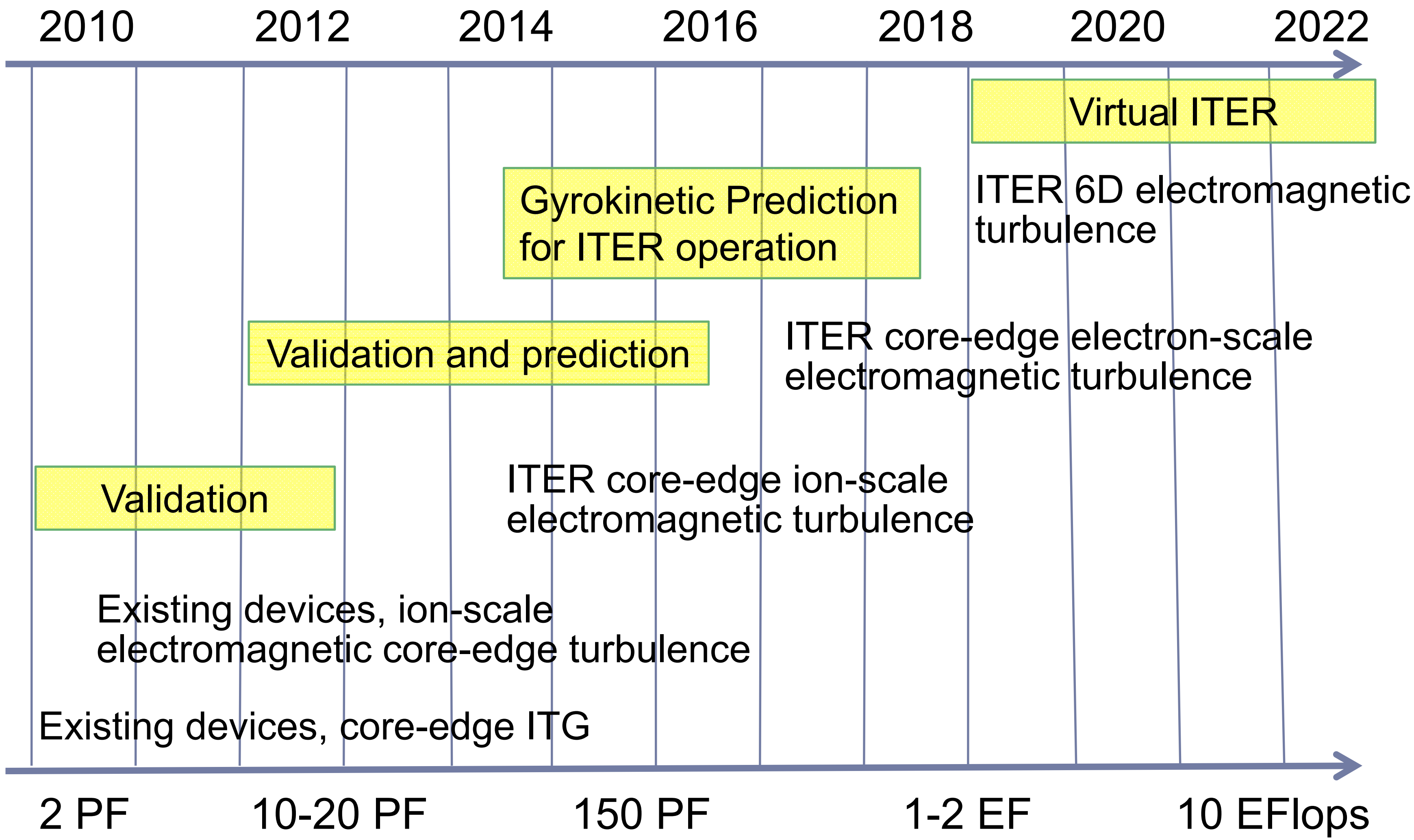
We need to be smart.



Looking forward to exascale dream

- Lack of computing power has forced us to reduce the 6D Vlasov plasma system into the 5D gyrokinetic system, restricting the kinetic simulation validity to << gyrofrequency and ≥ gyroradius
- On exa-scale HPC, the dream of a 6D whole-volume tokamak simulation (either PIC or FMM) can be realized, but highly challenging: requires close co-design with computer science and applied mathematics.
  - Implicit time-marching to avoid CFL trouble with Alfven waves
  - Localization of the data and computation
  - Efficient in-memory data staging and data analysis
  - Resiliency and fault tolerance?
  - Concurrency issue: dynamics load balancing
  - Flexibility to unknown new hardware and programming models
  - WAN data movement
  - Data storage
  - And more
- Happy that DM, analysis and vis are going to get solved ~_~

XGC1 Roadmap to Exascale (1 day run target)



Conclusion

- XGC is a new generation fusion particle code, efficiently scaling to the maximal number of Jaguarpf cores.
  - Unlike other existing gyrokinetic codes, **XGC1** simulates the **whole-volume tokamak** in realistic diverted magnetic field geometry in full-function (as opposed to the perturbative delta-f).
  - XGC1 simulates **background and turbulence dynamics together**, in multiscale.
- For higher fidelity modeling in XGC and experimental time telescoping, more extreme scale HPC is needed.
  - Data size is becoming extreme. We need advanced data management, analysis, and visualization.
- If an **exascale** HPC is available in the future, fusion particle code can make a **quantum jump into the formidable 6D tokamak physics simulation**. An efficient co-design is a necessity.