Extreme Scale Computational Science Challenges in Fusion Energy Research

William M. Tang

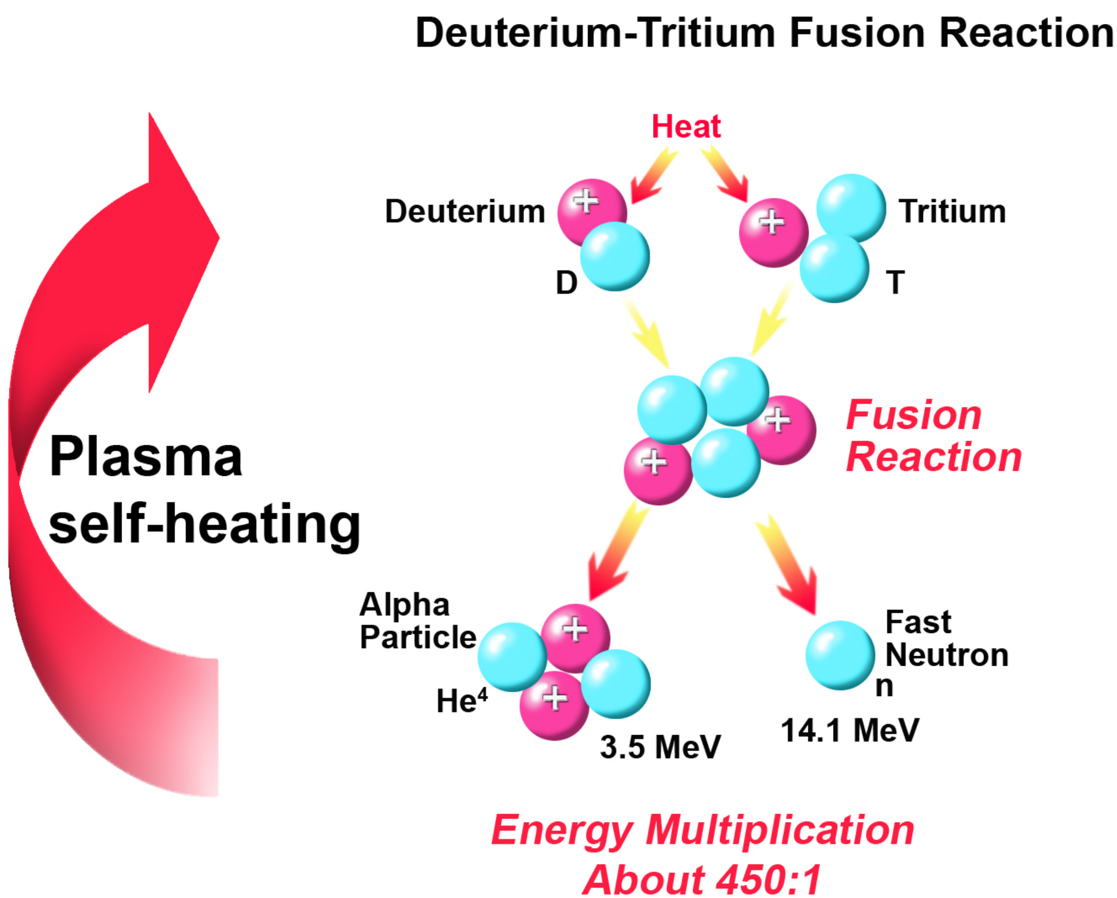
Princeton University, Plasma Physics Laboratory
Princeton, NJ USA

International Advanced Research 2012 Workshop on HPC (HPC 2012)

Cetraro, Italy

June 25 - 29, 2012

Fusion Energy: *Burning plasmas are self-heated and self-organized systems*

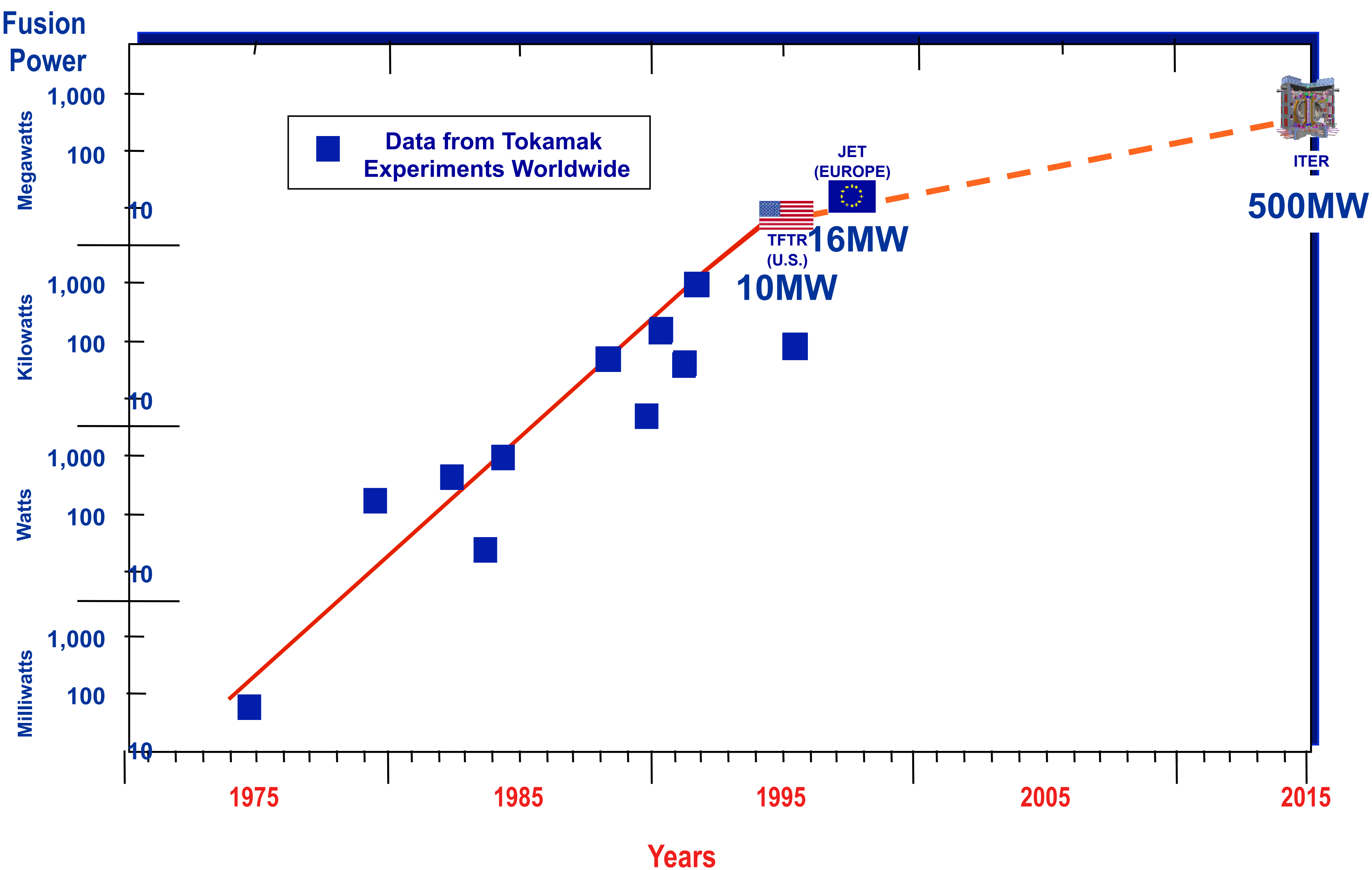


$$D^{+} + T^{+} \rightarrow {}^{4}He^{++}\ (3.5\ MeV) + n^{0}\ (14.1\ MeV)$$

Fusion: an Attractive Energy Source

- Abundant fuel, available to all nations
  - Deuterium and lithium easily available for millions of years
- Environmental advantages
  - No carbon emissions, short-lived radioactivity
- Cannot "blow up or melt down," resistant to terrorist attack
  - Less than a minute's worth of fuel in the chamber
- Low risk of nuclear materials proliferation
  - No fissile materials required
- Compact relative to solar, wind and biomass
  - Modest land usage
- Not subject to daily, seasonal or regional weather variation; no requirement for local CO_2 sequestration
  - Not limited in its application by need for large-scale energy storage nor for long-distance energy transmission
- Fusion is complementary to other attractive energy sources

Progress in Magnetic Fusion Energy (MFE) Research



ITER Goal: *Demonstration of the Scientific and Technological Feasibility of Fusion Power*

- ITER *is a dramatic next-step for Magnetic Fusion Energy (MFE)*:

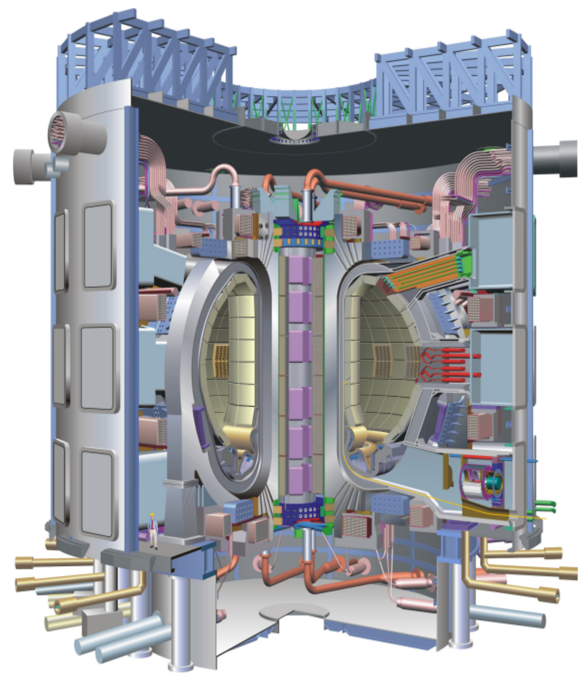
  -- Today: 10 MW(th) for 1 second with gain ~1
  -- ITER: 500 MW(th) for >400 seconds with gain >10

- Many technologies used in ITER will be same as required in power plant *but additional R&D will be needed*
  -- "DEMO": 2500 MW(th) continuous with gain >25, in a device of similar size and field as ITER

- Ongoing R&D programs worldwide [experiments, theory, computation, and technology] essential to provide growing knowledge base for ITER operation targeted for ~ 2020

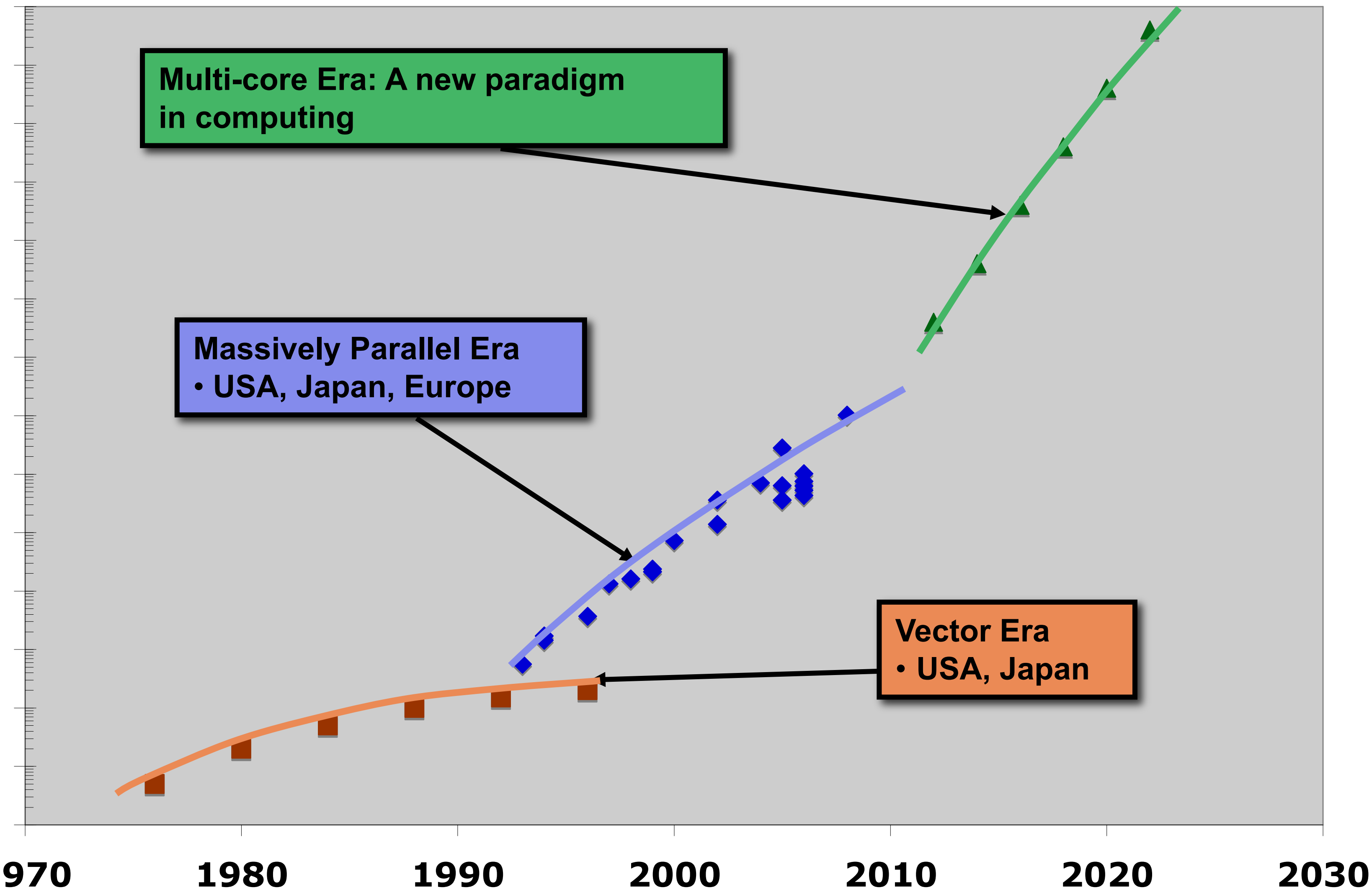
➔ *Realistic HPC-enabled simulations required to cost-effectively "steer" & harvest key information from expensive (~$1M/long-pulse) shots*

ITER



FES Needs to be Prepared to Exploit *Local Concurrency* to Take Advantage of Most Powerful Supercomputing Systems in 21st Century

(e.g., U.S.'s Blue-Gene-Q & Titan, Japan's Fujitsu-K, China's Tianhe-1A, ….)



Modern HPC can *Transform* Many Domain Applications Areas in Science (including FES) & in Industry

- ***Practical Considerations:*** **[achieving "buy-in" from general scientific community]**
  - **Need to distinguish between *"voracious"* (*more of same - just bigger & faster*) *vs.* *"transformational"* (*achievement of major new levels of scientific understanding*)**
  - **Need to improve *significantly*** on ***experimental validation*** **together with *verification & uncertainty quantification* to enhance realistic predictive capability**
- ***Associated Extreme Scale Computing Challenges:***

■ ***Hardware complexity***: Heterogenous multicore (e.g., gpu+cpu => OLCF's "Titan"), power management, memory, communications, storage, …

■ ***Software challenges***: Operating systems, I/O and file systems, and coding/algorithmic & solver needs in the face of increased computer architecture complexity … must deal with local concurrency (MPI + threads, CUDA, etc. → rewriting code focused on data movement over arithmetic)

- References:

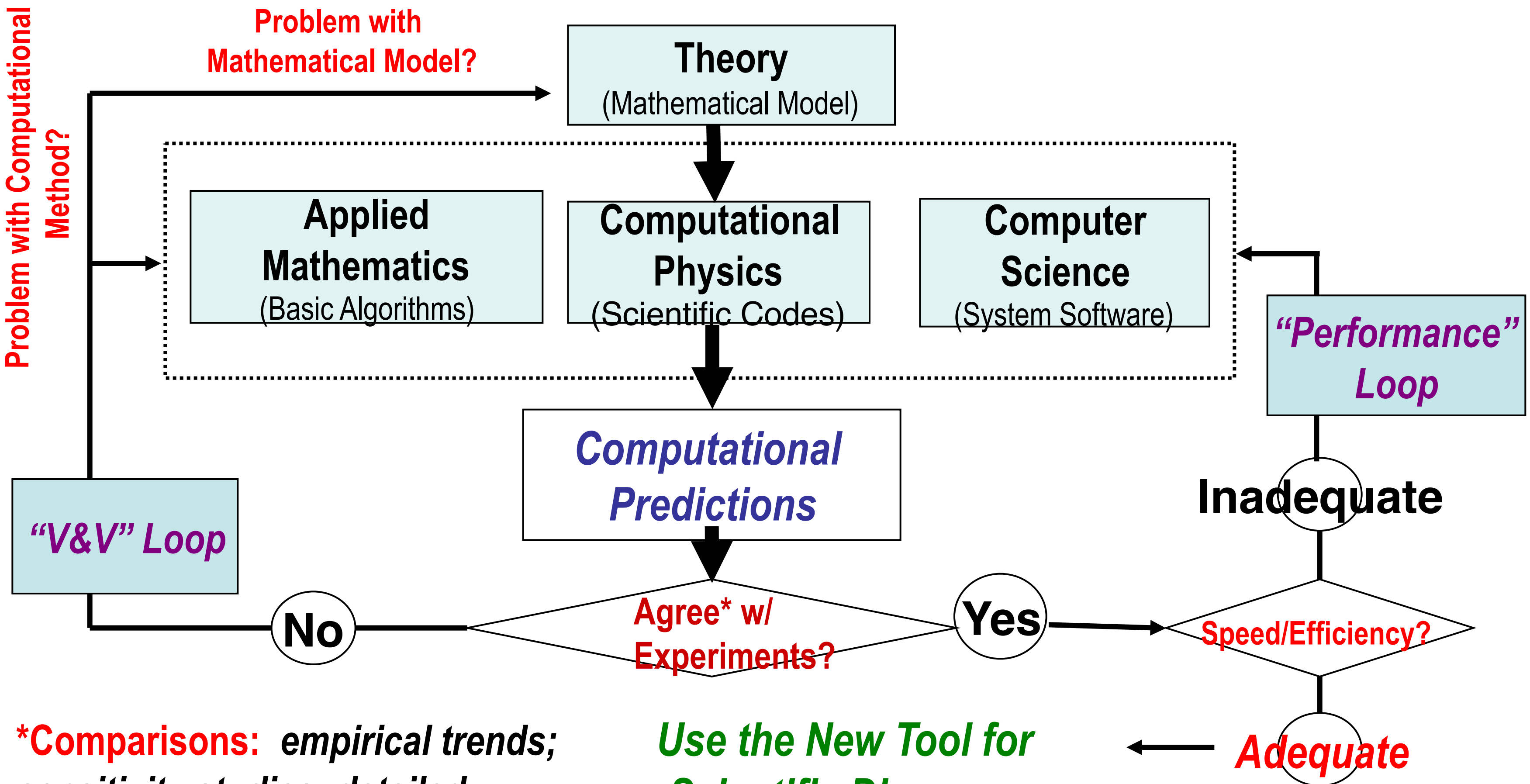
W. Tang, D. Keyes, et al., ***"Scientific Grand Challenges: Fusion Energy Sciences and the Role of Computing at the Extreme Scale,"*** PNNL-19404, 212pp (March, 2009). http://www.er.doe.gov/ascr/ProgramDocuments/Docs/FusionReport.pdf

R. Rosner, et al., ***"Opportunities & Challenges of Exascale Computing"*** – DoE Advanced Scientific Computing Advisory Committee Report (November, 2010).

G8 Exascale Software Projects (2011- 2014)

- “Enabling Climate Simulation @ Extreme Scale” (ECS) – US, Japan, France, Canada, Spain
- “Climate Analytics on Distributed Exascale Data Archives” (ExArch) UK, US, France, Germany, Canada, Italy
- “Icosahedral-Grid Models for Exascale Earth System Simulations” (ICOMEX) – Japan, UK, France, Germany, Russia
- “Nuclear Fusion Simulations @ Exascale” (NuFuSE) – UK, US, Germany, Japan, France, Russia
- “Modeling Earthquakes and Earth's Interior based on Exascale Simulations of Seismic Wave Propagation” (Seismic Imaging) – US, Canada, France
- “Using Next-Generation Computers & Algorithms for Modeling Dynamics of Large Bio-molecular Systems” (INGENIOUS) -- Japan, UK, France, Germany, Russia

Advanced Scientific Codes --- *"a measure of the state of understanding of natural and engineered systems"* (T. Dunning, 1st SciDAC Director)

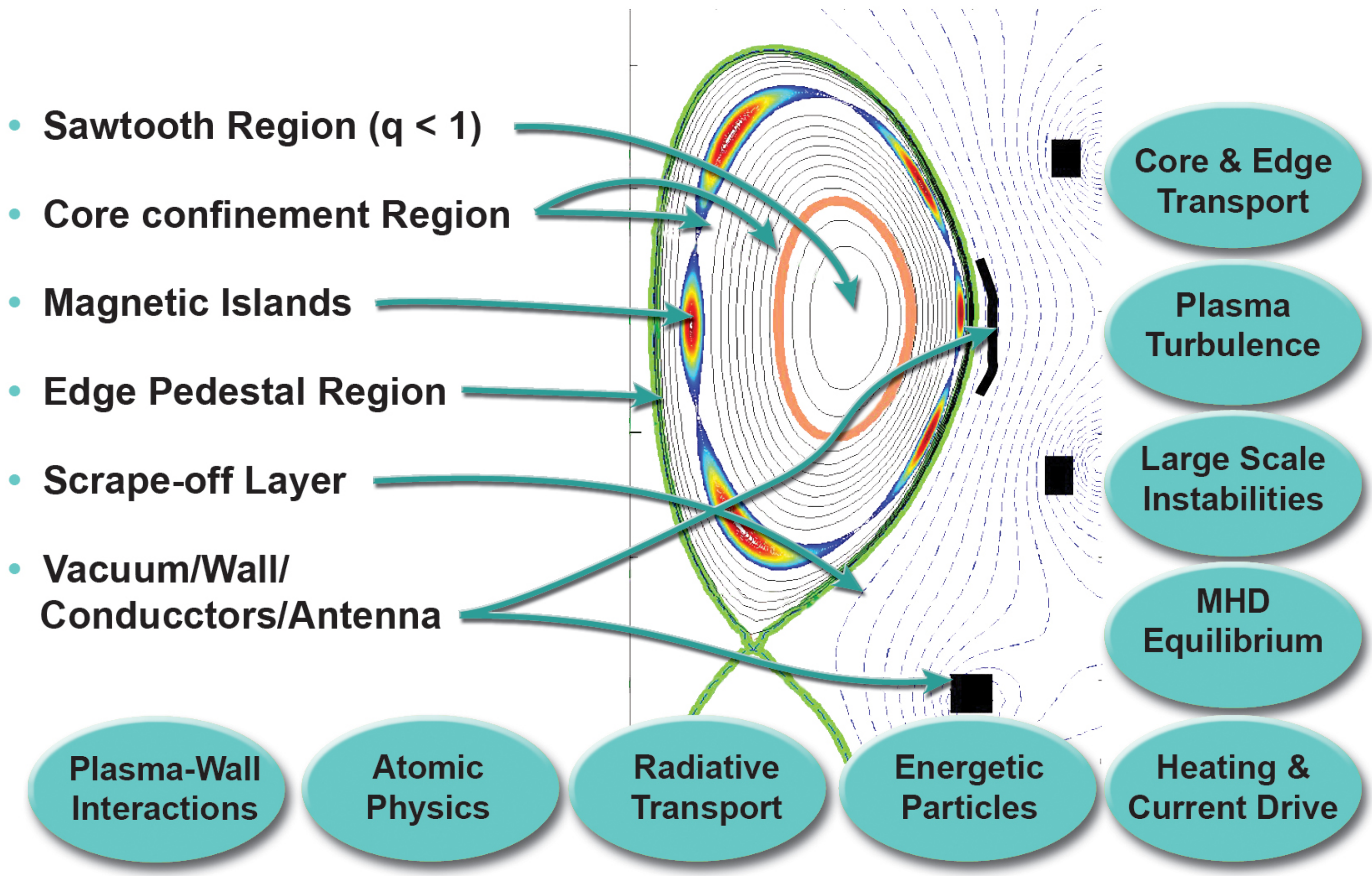


*Comparisons: *empirical trends; sensitivity studies; detailed structure (spectra, correlation functions, …)*

Use the New Tool for Scientific Discovery

(Repeat cycle as new phenomena encountered)

Elements of an MFE Integrated Model ➔ *Complex Multi-scale, Multi-physics Processes*



• W.Tang, D. Keyes, et al., ***"Scientific Grand Challenges: Fusion Energy Sciences and the Role of Computing at the Extreme Scale,"*** PNNL-19404, 212pp (March, 2009). http://www.er.doe.gov/ascr/ProgramDocuments/Docs/FusionReport.pdf

Microturbulence in Fusion Plasmas – Mission Importance: *Fusion reactor size & cost determined by balance between loss processes & self-heating rates*

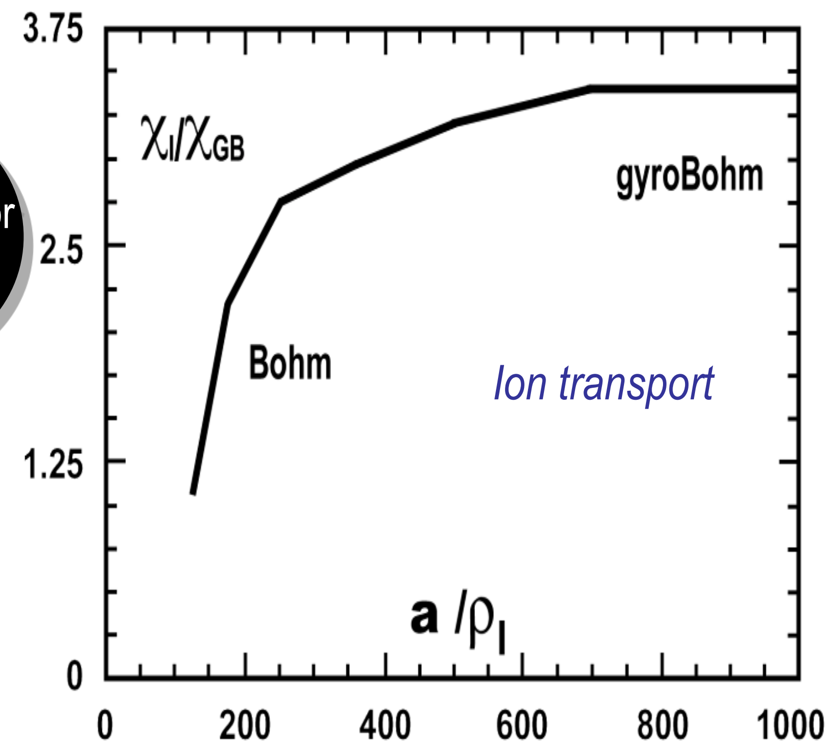
- ***"Scientific Discovery"*** - *Transition to favorable scaling of confinement produced in simulations for ITER-size plasmas*
  - a/ρ_i = 400 (JET, largest present lab experiment) through
  - a/ρ_i = 1000 (ITER, ignition experiment)

Good news for ITER!

- *Multi-TF simulations* using GTC global PIC code [e.g., Z. Lin, et al, Science, *281, 1835 (1998), PRL (2002)*] deployed a billion particles, 125M spatial grid points; 7000 time steps at NERSC ➔ *1st ITER-scale simulation with ion gyroradius resolution*

- ***Understanding*** physics of favorable plasma size scaling trend demands *much greater computational resources* + improved algorithms [radial domain decomposition, hybrid (MPI+Open MP) language, ..] & modern diagnositics

  -- *current Early Science Applications (ESA) GTC-P project on ALCF*



➔ ***Excellent Scalability of Global PIC Codes on LCF's enables advanced physics simulations to improve understanding***

- Global PIC code development ***for GPU and other low memory/core environments*** actively pursued
  [e.g. -- ***SC2011 Paper on GPU version of GTC; 2011 Beijing Exascale CoDesign Workshop → GTC on Tianhe-1A, China***]

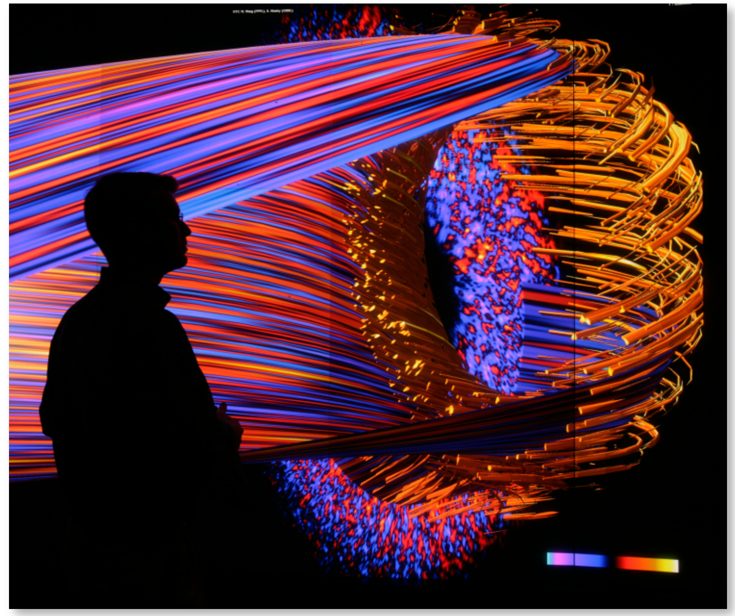
Demonstrated GTC Capability: Faster Computer → Achievement of Improved Fusion Energy Physics Insights

GTC simulation	Computer name	PE# used	Speed (TF)	Particle #	Time steps	Physics Discovery (*Publication*)
1998	Cray T3E NERSC	10^2	10^{-1}	10^8	10^4	Ion turbulence zonal flow (*Science, 1998*)
2002	IBM SP NERSC	10^3	10^0	10^9	10^4	Ion transport scaling (*PRL, 2002*)
2007	Cray XT3/4 ORNL	10^4	10^2	10^{10}	10^5	Electron turbulence (*PRL, 2007*); EP transport (*PRL, 2008*)
2009	Jaguar/Cray XT5 ORNL	10^5	10^3	10^{11}	10^5	Electron transport scaling (*PRL, 2009*); EP-driven MHD modes *(Pub?)*
2012 (current)	Cray XT5 → Titan ORNL Tianhe-1A (China)	10^5	10^4	10^{12}	10^5	Kinetic-MHD; Turbulence + EP + MHD
2018 (future)	Path to Exascale HPC Resources	TBD	10^6	10^{13}	10^6	Turbulence + EP + MHD + RF

* GTC first FES code delivering production run simulations @ TF in 2002 and PF in 2009

Petascale-capability enables ***multi-scale simulations*** providing new insights into nature of plasma turbulence

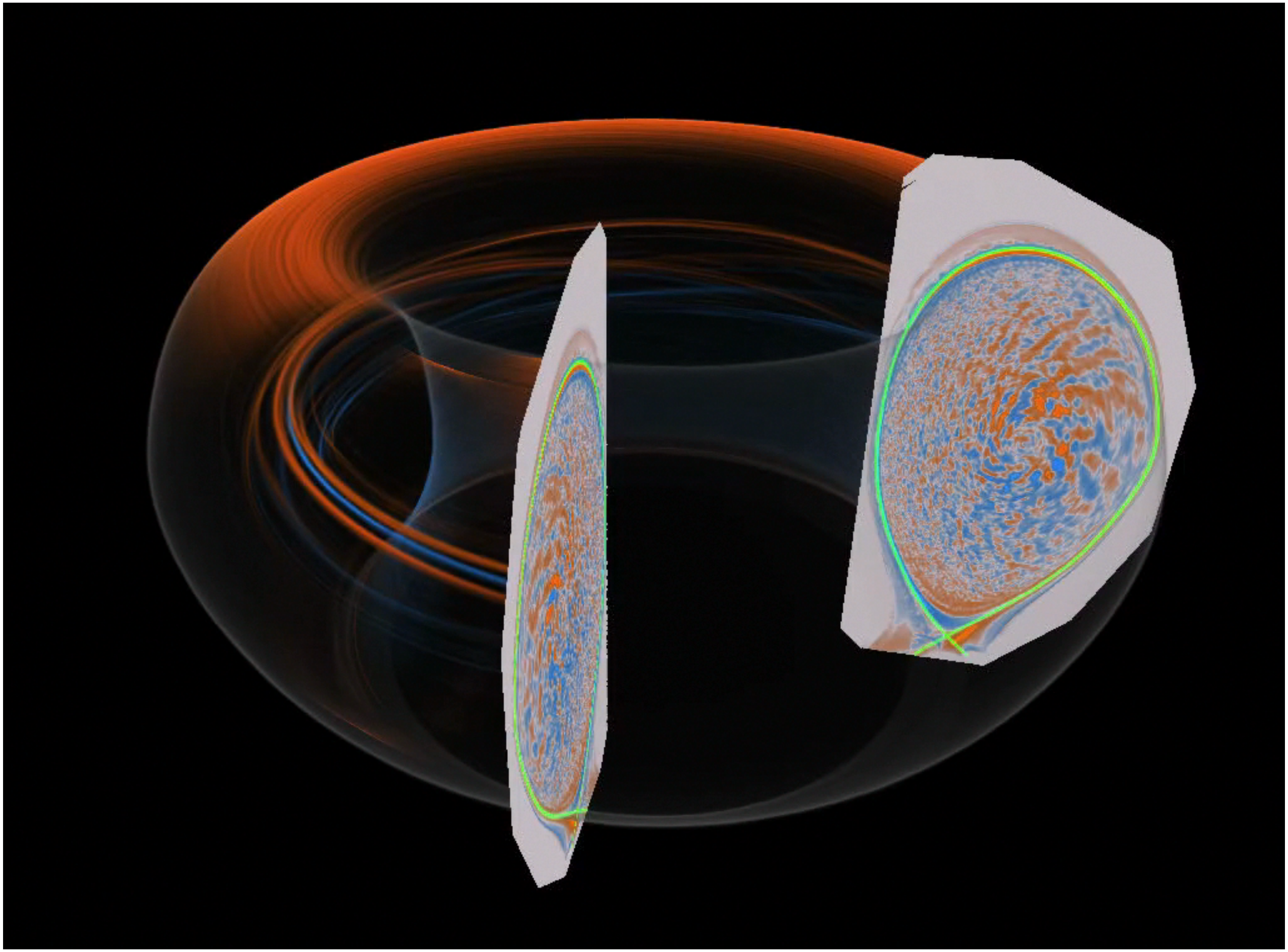
- Multi-scale simulations accounting for fully global 3D geometric complexity of problem ***spanning micro and meso scales*** have been carried out on ORNL's Jaguar LCF [GTS & XGC-1 PIC codes]
- Dynamics in complex Edge region integrated with Core Plasma in XGC1-code [C. S. Chang, et al.]
  - XGC-1 solves for total distribution function directly -- with source and sink + associated "noise challenges"
  - Demands access to modern petascale platforms for ***needed resolution***
  - Example -- Current petascale-level production runs with XGC-1 require 24M CPU hours (100,000 cores × 240 hours)



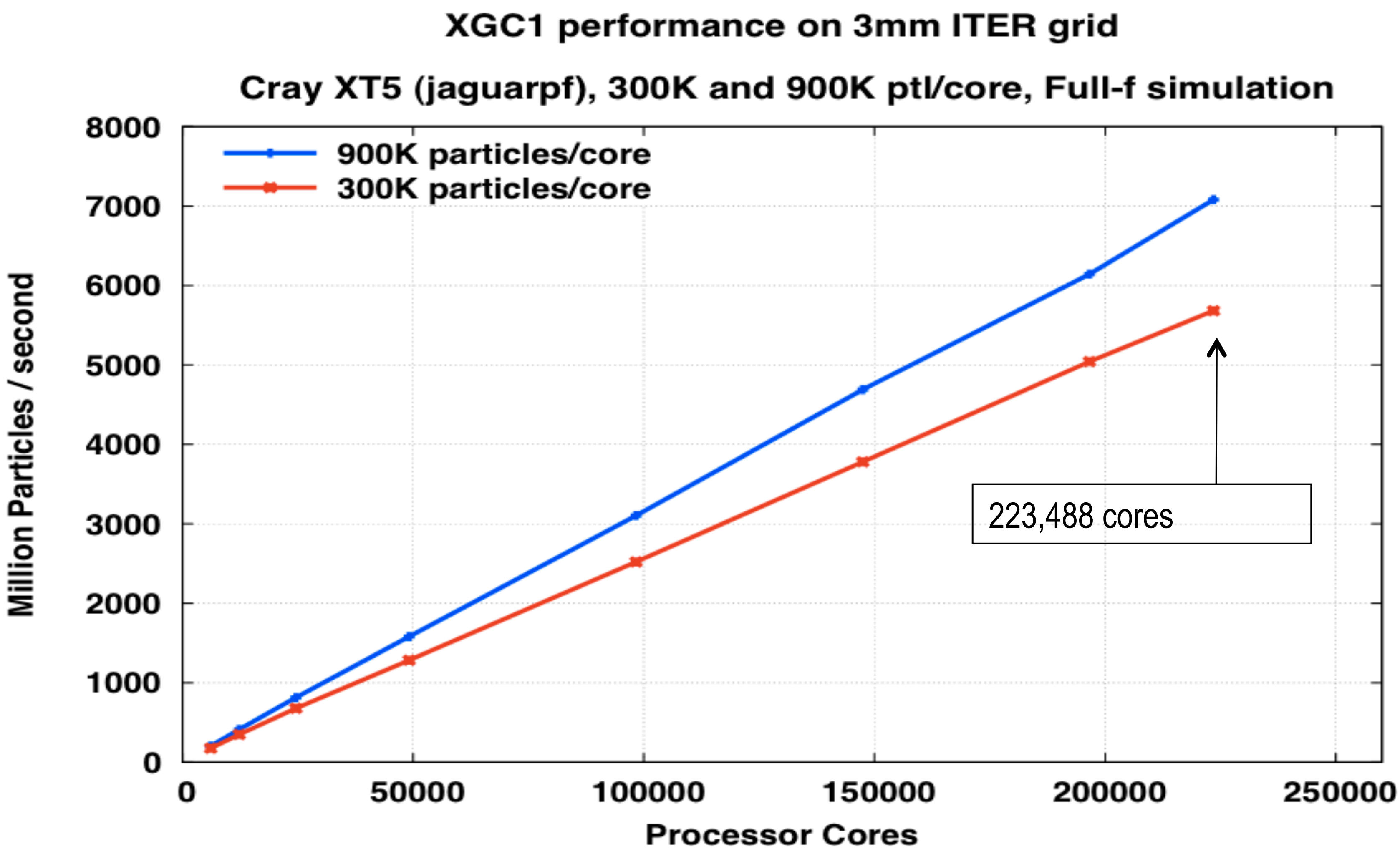
Key Impact:
Petascale computing power has accelerated progress in understanding heat losses caused by plasma turbulence

- Exascale-level production runs are needed to enable running codes with even higher physics fidelity and more comprehensive & realistic integrated dynamics

Modern 3D Visualization: Advanced PIC Simulations with XGC-1 Code on "Jaguar" OLCF [C.S. Chang, et al., SciDAC CPES Project]

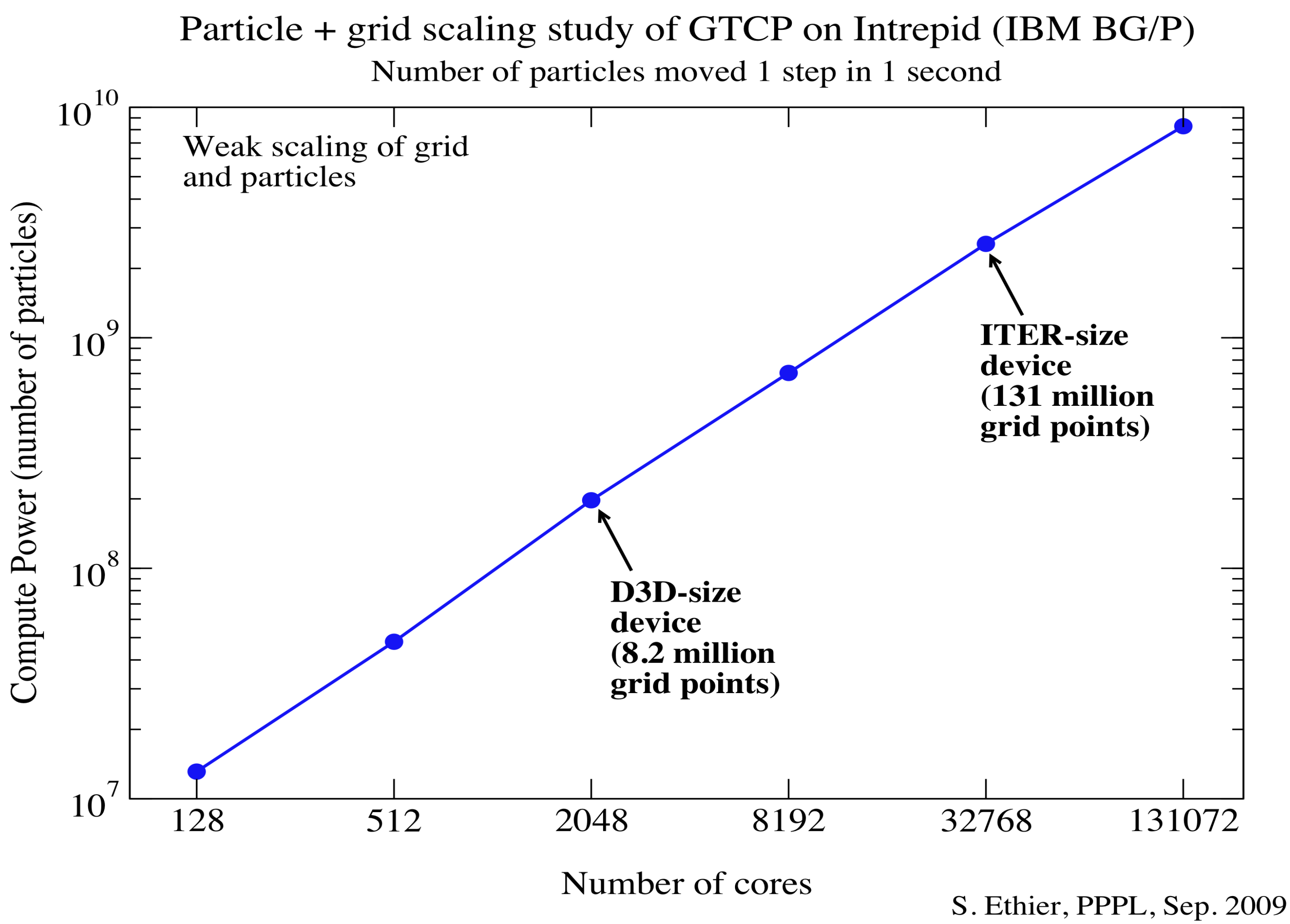


XGC1 Petascale Studies on Jaguar (OLCF)



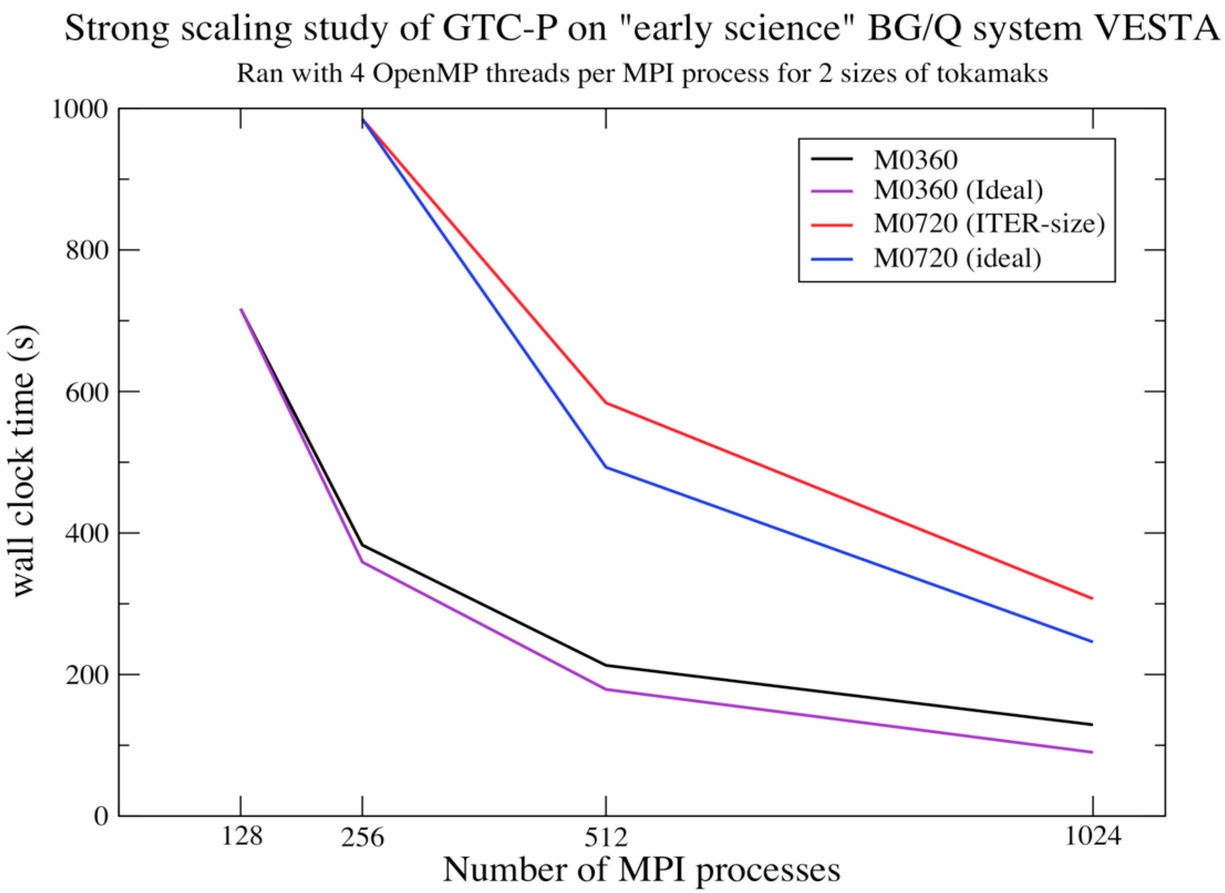
XGC1 scales efficiently all the way to full Jaguarpf capability (with MPI+ OpenMP) & routinely uses >70% capability

Weak Scaling Study GTC-P on IBM BG-P at ALCF



- Excellent scalability demonstrated [both grid size and # of particles increased proportionally with # of cores] → (*also on 294,912 cores, 72 racks @ JSC in Germany*)
- Plans in place for similar weak scaling collaborative studies on Fujitsu-K Machine in Japan

Strong Scaling Study of GTC-P in "Early Science" Project on Single-Rack IBM BG/Q "Vesta" System at ALCF



Excellent performance demonstrated –> ***recent results from Early Science ALCF Project show ~ order of magnitude improvement on new (multi-petaflop) IBM BG-Q ("Mira")***

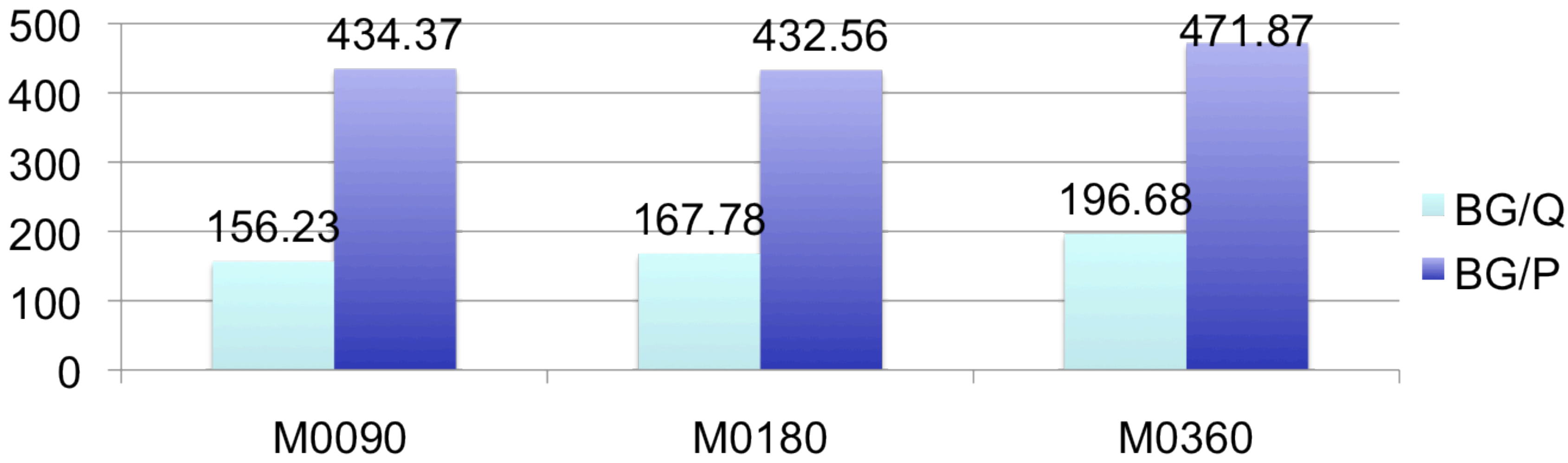


Figure: GTC-P performance comparison (seconds) on BG/P and BG/Q for M0090, M0180 and M0360 with particle per cell (ppc)=100 for 100 steps

M0090	Total # of nodes	Total # of cores	# of cores/ node	# of threads/ core	Time (s) for 100 steps	Speed up per core	Speed up Per node
BG/P	128	512	4	1	434.37	1.0	1.0
BG/Q	32	512	16	4	156.23	2.78	11.12
M0180							
BG/P	512	2048	4	1	432.56	1.0	1.0
BG/Q	128	2048	16	4	167.78	2.58	10.31
M0360							
BG/P	2048	8192	4	1	471.87	1.0	1.0
BG/Q	512	8192	16	4	196.68	2.40	9.60

Tables: experiment settings and performance results on BG/P and BG/Q

M0180 ppc=100	**Our test**	**ANL**	**IBM**
Speed up per node (Q/P ratio)	10.31	10.7	11.2

Table 2: Speed up per node comparison with ALCF and IBM results for M0180 problem with ppc=100 for 100 steps

GTC ON TIANHE-1A ➔ Particle-in-cell global kinetic turbulence code (GTC) running on CPU's only in scaling case study with GPU+CPU version under current active development

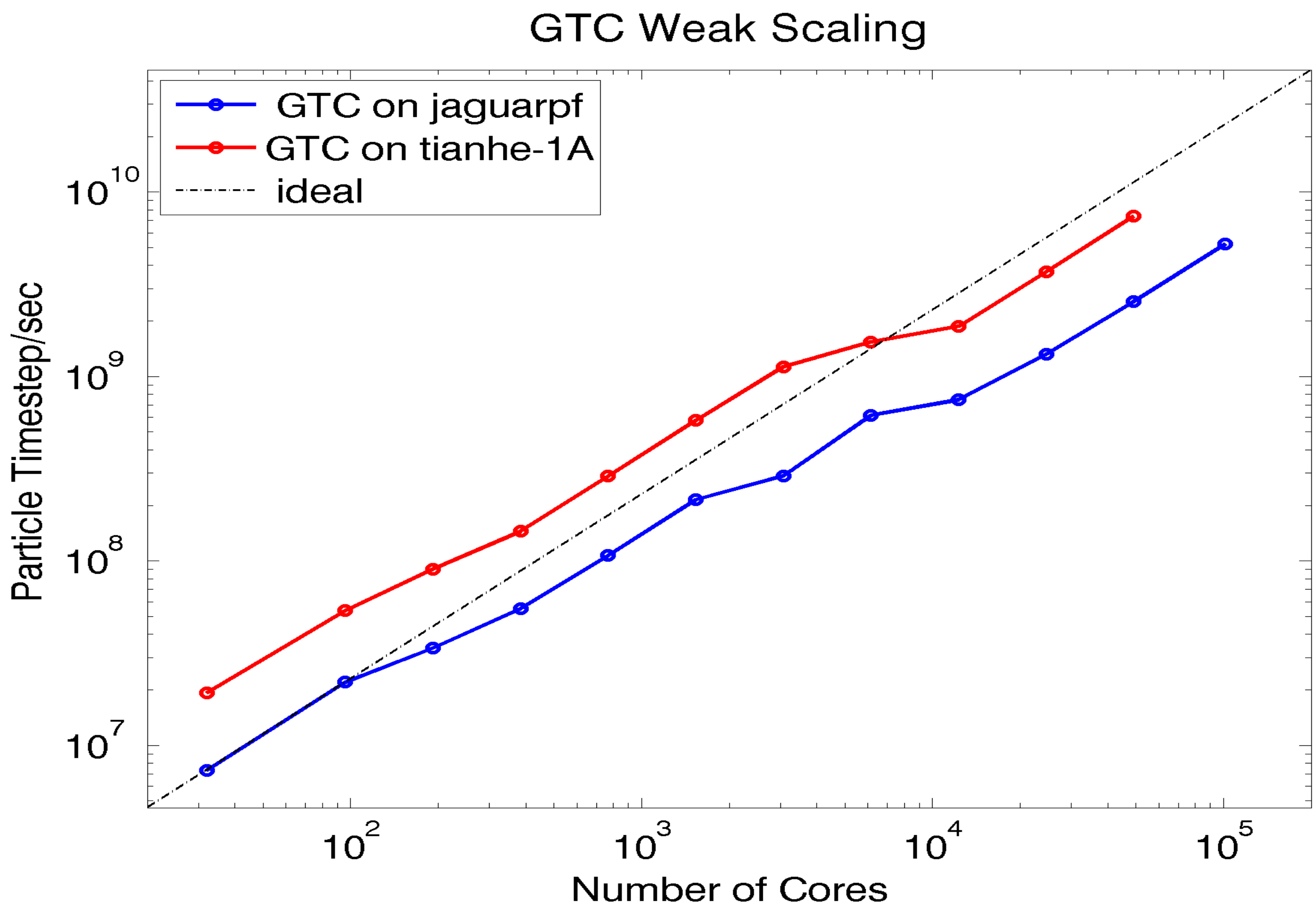
Observations on improved performance:

• Tianhe-1A (8 core nodes) & Jaguarpf (12 core nodes) ➔ improvement actually ~ 1.7

• Improvement due primarily to Intel processor & compiler performance on Tianhe-1A

• GTC's relative insensitivity to communication time ➔ little benefit from Tianhe-1A's better network

--



New GTC-GPU Code *(K. Ibrahim, LBNL; B. Wang, Princeton U; et al.)*

Introduced at SC2011:

K. Madduri, *K. Ibrahim*, S. Williams, E.J.Im, S. Ethier, J. Shalf, L. Oliker, *"Gyrokinetic Toroidal Simulations on Leading Multi- and Manycore HPC Systems"*

- Use current GTC version with demonstrated comprehensive physics
- Challenge: massive fine-grained parallelism and explicit memory transfers between multiple memory spaces within a compute node

- Approach: consider 3 main computational phases: *charge deposition, particle push and particle shift*

-- integrates three programming models [nVidia, Cuda, & OpenMP] within a node, and MPI between nodes

-- demonstrated excellent scaling behavior on NERSC Dirac test-bed

-- explored breaking the limit of Amdhal's law on speedup by parallelizing - using atomics - the charge deposition phase, which has iterations with loop-carried dependency → *Memory locality improves performance of most routines but degrades performance for atomics because of access conflicts*

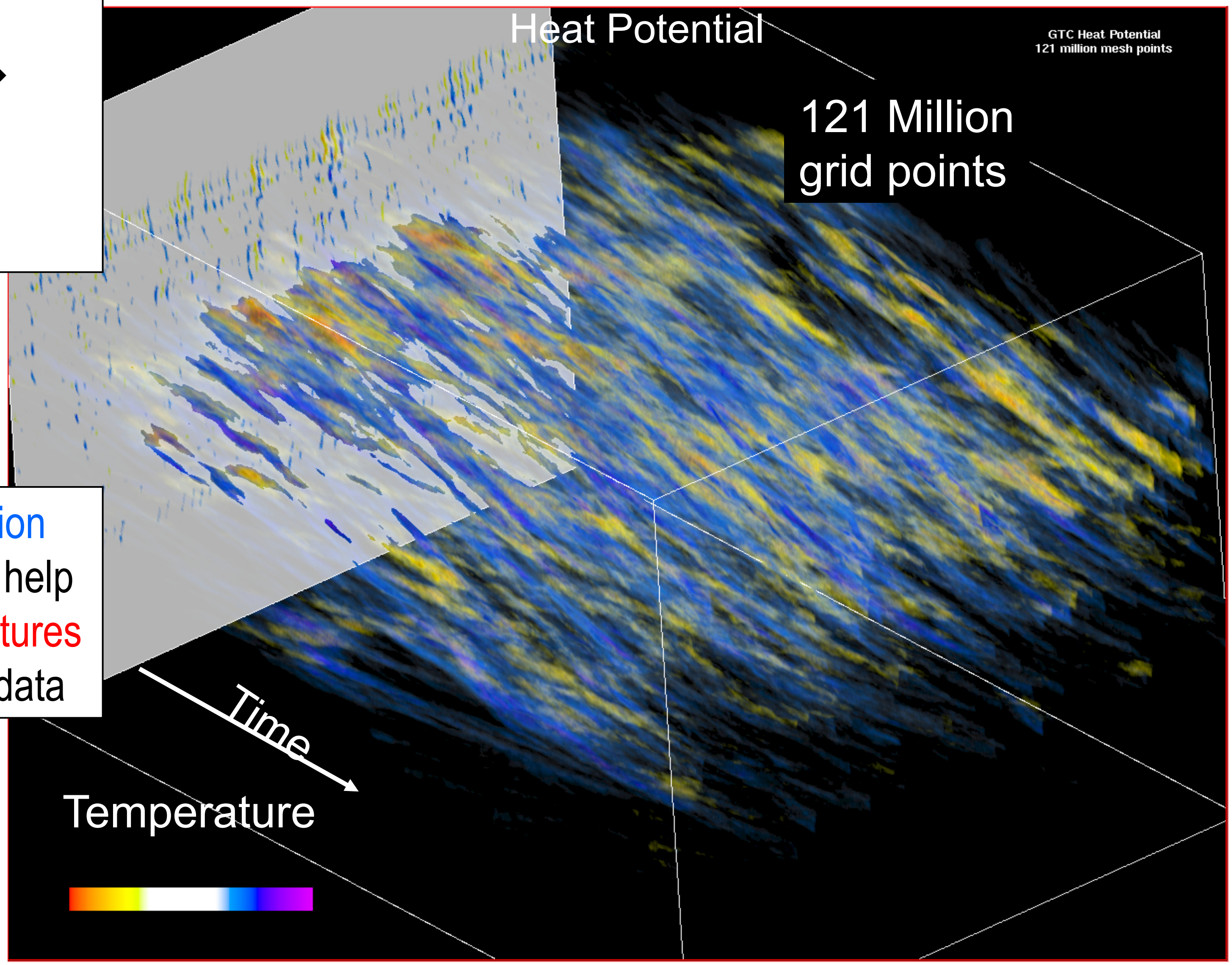
→ *Conflicting requirements for locality and conflict avoidance* make optimizing the performance on GPUs both interesting and challenging.

“Big Data” Challenges for FES Simulations & Experiments

Particle in Cell Turbulence Simulation

Multi-petabytes of data generated at LCF’s ➔ demands efficient new Data Management & Analysis Methods

New Multi-D Visualization Capabilities needed to help identify & track key features in complex simulated data



Data Management & Visualization Challenges

- Automated Workflow Environment:
  - Peta-bytes of data need to be moved automatically from simulations to analysis codes
  - Feature Detection/Tracking to harvest scientific information -- impossible to understand in timely way without new data mining techniques

- Parallel I/O Development and Support - define portable, efficient standard with interoperability between parallel and non-parallel I/O
  - Massively parallel I/O systems (e.g. “ADIOS” from ORNL) needed since storage capacity growing faster than bandwidth and access times
  - Feasibility of “Local I/O” future capabilities (e.g., M. Seager’s talk) of great interest

- Real-time visualization to enable “steering” of long-running simulations

FES DATA ANALYSIS CHALLENGES FOR ITER

- DATA TRANSFER FROM ITER TO US
  - Current estimates of data size is roughly 40 TB per shot for long-pulse shots of 400 seconds duration
    - -- would demand 100 GB/sec bandwidth
    - -- likely need to be able to parallelize at least a significant fraction of this data for streaming
  - Current estimates of time between shots is roughly 1600 seconds -- a rather limited period of time
    - -- I/O will be very stressed for:
    - (i) reading even a fraction of this amount of data from memory into CPUS & then writing back to disk
    - (ii) displaying of the information
    - realistic development of such capabilities is a major challenge
  - *Current capabilities not likely able to deal with future parallelism and streaming issues*

FES DATA ANALYSIS CHALLENGES FOR ITER (continued)

- LIKELY CHANGE IN PARADIGM: movement from current "data file paradigm" to "data streaming paradigm" to accommodate much larger data sets
  - analogous to looking at various frames of a movie while the movie is still being generated
  - advance image processing capabilities could enable end-users/physicists to examine/analyze information while shot in progress
- ASSOCIATED HARDWARE CHALLENGES
  - Most present-day computer systems do not have the memory (50 TB or so) needed to deal with large data collection
    - -- might lead to approach of examining one stream at a time or possibly processing one stream on one machine while simultaneously moving another stream
- ASSOCIATED SECURITY CHALLENGES
  - Users can access parts of data per shot but not allowed access to other associated information
  - Users need to add information/annotate shots & query off their own and other collaborators annotations
  - Important to keep connections "alive" for long periods & keeping the security channels open

HPC Challenges in Moving toward Exascale

Locality: *Need to improve data locality (e.g., by sorting particles according to their positions on grid)*

-- due to physical limitations, ***moving data between, and even within, modern microchips is more time-consuming than performing computations!***

-- scientific codes often use data structures that are easy to implement quickly but limit flexibility and scalability in the long run

Latency: *Need to explore **highly multi-threaded algorithms** to address memory latency*

Flops vs. Memory: *Need to utilize Flops (cheap) to better utilize Memory (limited & expensive to access)*

Advanced Architectures: *Need to deploy innovative algorithms within modern science codes on **low memory per node architectures** – **(e.g, BG/Q, Fujitsu-K, Tianhe-1A, & Titan)***

-- multi-threading within nodes, maximizing locality while minimizing communications

-- large future simulations *(e.g., PIC ➔ need to likely work with >10 billion grid points and over 100 trillion particles!!)*

Future Science Challenges and Opportunities

(1) ***Energy Goal*** in FES application domain is to increase availability of clean abundant energy by first moving to a ***burning plasma experiment*** -- the multi-billion dollar ***ITER*** facility located in France & involving the collaboration of 7 governments representing over half of world's population

-- ITER targets 500 MW for 400 seconds with gain > 10 to demonstrate ***technical feasibility of fusion energy*** & ***DEMO*** (***demonstration power plant***) will target 2500 MW with gain of 25

(2) ***HPC Goal*** is to harness increasing HPC power at the extreme scale to ensure timely progress on the scientific grand challenges in FES as described in DoE-SC report (2010) on ***"Scientific Grand Challenges: Fusion Energy Sciences and Computing at the Extreme Scale."***

(3) ***Experimental Validation Goal*** is to engage tokamaks worldwide to: (i) provide key data bases and (2) ***develop and deploy accurate new diagnostics*** to enable new physics insights – including ***realistic sensitivity studies to support uncertainty quantification***.

Overall "Path to Exascale" Goal in Fusion Energy Science:

Accelerate progress in delivering reliable integrated predictive capabilities – benefiting from access to HPC resources – from petascale to exascale & beyond -- together with appropriate data management and a vigorous verification, validation, & uncertainty quantification program