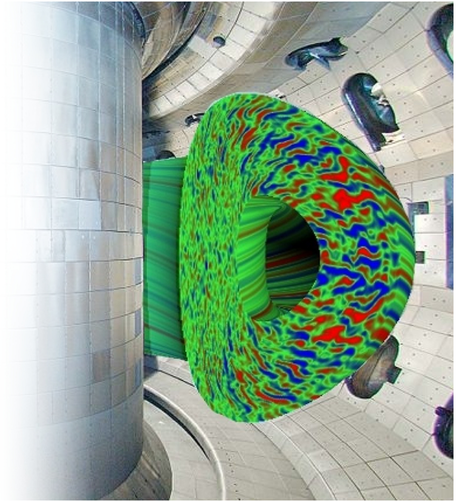
GPU-Acceleration of Plasma Turbulence Simulations for Fusion Energy

by
J. Candy[1], I. Sfiligoi[1], E. Belli[1], N. Howard[2], C. Holland[3]

[1]**General Atomics, San Diego, CA**
[2]**MIT, Cambridge, MA**
[3]**UCSD, San Diego CA**



Presented at

SC18/NVIDIA
Dallas, TX
11-16 Nov 2018

Background and Motivation

1. **General Atomics** (GA) is a private contractor in San Diego

Background and Motivation

1. **General Atomics** (GA) is a private contractor in San Diego
2. The GA **Magnetic Fusion** division does DOE-funded research

Background and Motivation

1. **General Atomics** (GA) is a private contractor in San Diego
2. The GA **Magnetic Fusion** division does DOE-funded research
3. Hosts **DIII-D National Fusion Facility**

Background and Motivation

1. **General Atomics** (GA) is a private contractor in San Diego
2. The GA **Magnetic Fusion** division does DOE-funded research
3. Hosts **DIII-D National Fusion Facility**

4. THIS TALK: GPU-based plasma turbulence simulation using **gyrokinetic model**

Important locations for CGYRO

Source code

`github.com/gafusion/gacode`

DOI

`www.osti.gov/doecode/biblio/20298`

User Documentation

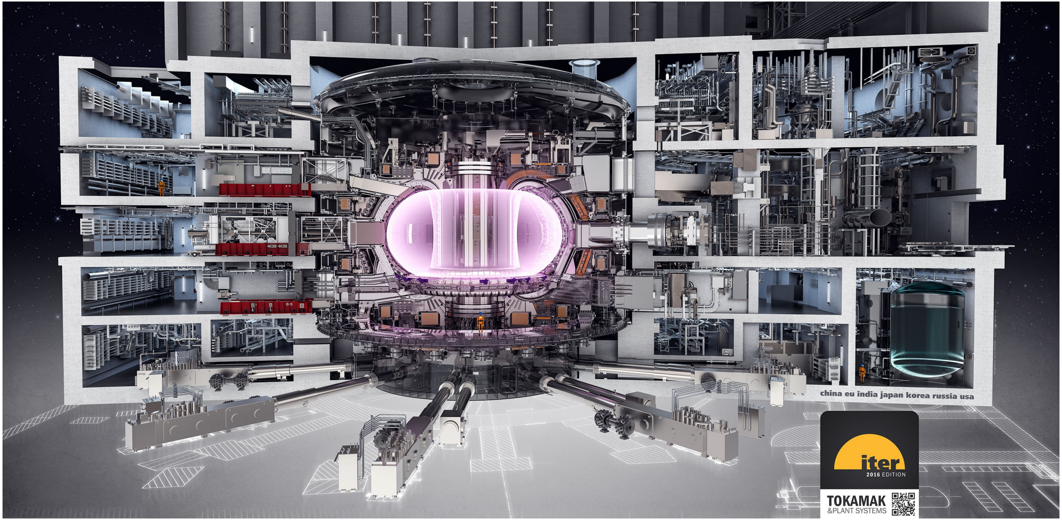
`gafusion.github.io/doc`

Documentary Video (for GYRO)

`www.youtube.com/watch?v=RLI6QW2x4Lg`

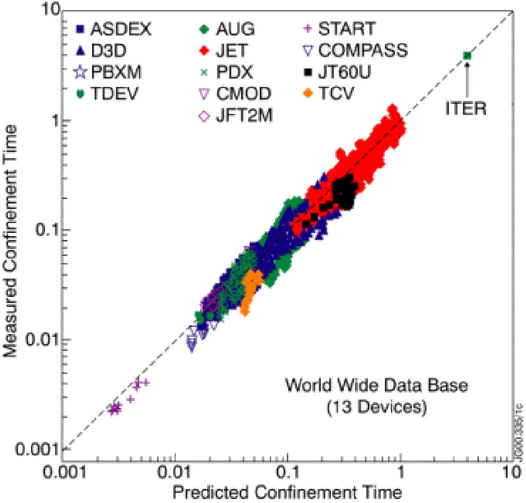
ITER Facility (35 nations) under construction in France

GOAL: Simulate turbulent plasma in core (magenta) region



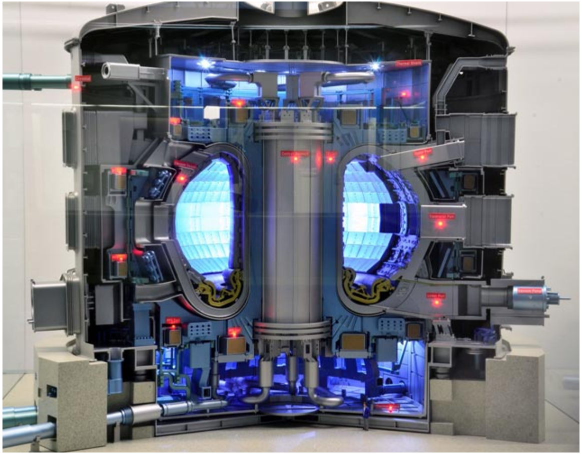
Why such a large facility?

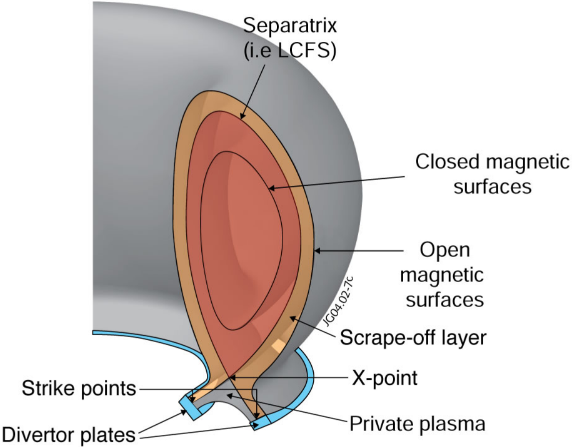
Tokamak confinement improves with LARGE PLASMA VOLUME



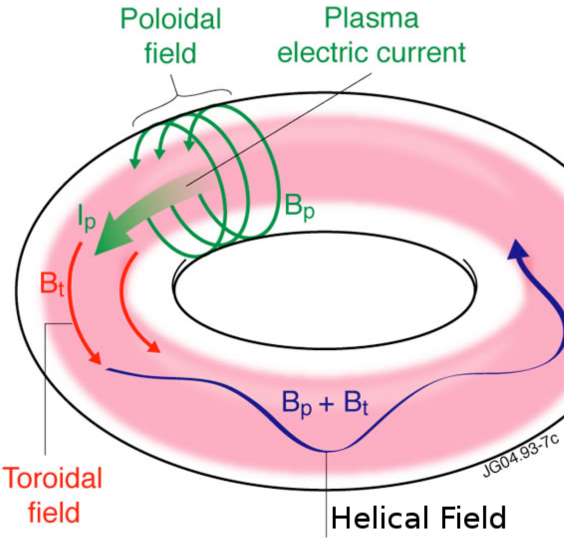
GENERAL ATOMICS

Plasma theory in closed fieldline region well-understood

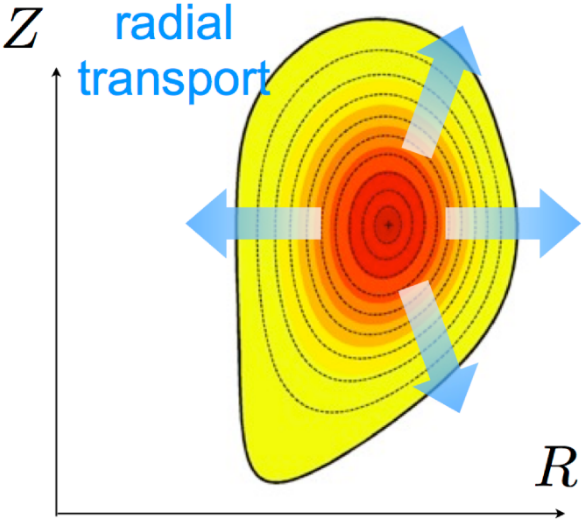




Helical field perfectly confines plasma (almost)



There is a small amount of radial energy/particle loss



- Collisions (1970s): $\Gamma_{\text{collision}}$
- Turbulence (1980s): $\Gamma_{\text{turbulence}}$
- Both exhibit **gyroBohm scaling**

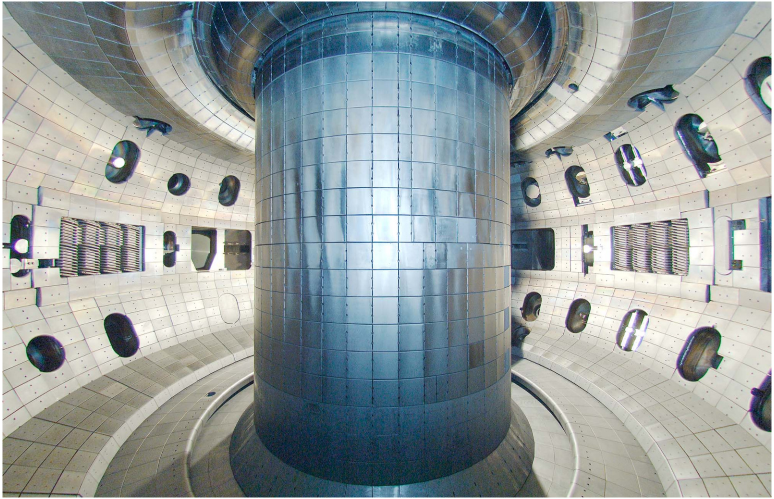
$$\textbf{flux} \quad \Gamma \sim v(\rho/a)^2$$

$$\textbf{confinement time} \quad \tau = \frac{a}{\Gamma} \sim \boxed{\frac{a^3}{v\rho^2}}$$

- $a =$ torus radius
- $\rho =$ particle orbit size
- $v =$ particle velocity

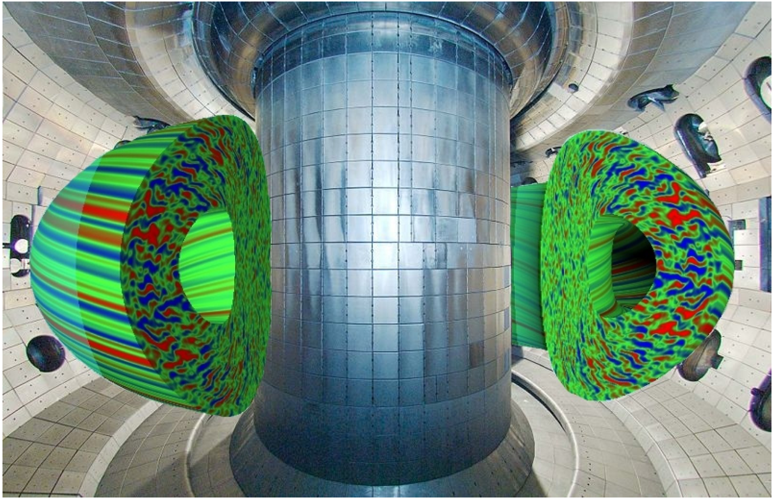
CGYRO computes the turbulent flux

DIII-D Tokamak at General Atomics in San Diego, CA



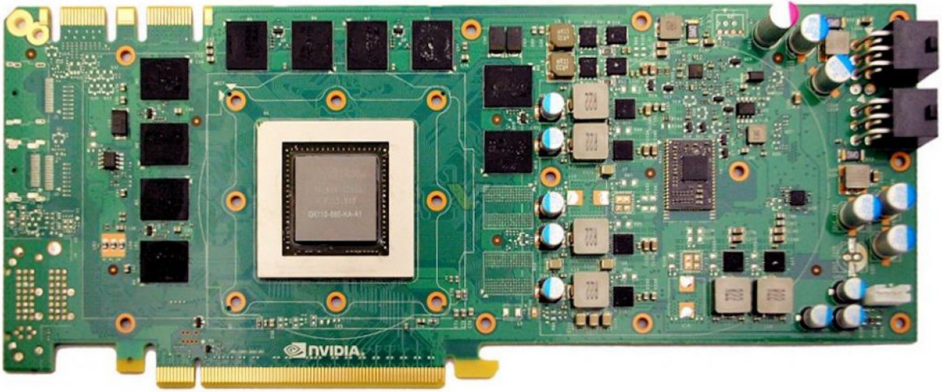
CGYRO computes the turbulent flux

DIII-D Tokamak at General Atomics in San Diego, CA



CGYRO fully ported to GPU

NCCS TITAN (Oak Ridge, TN) – **K20x**



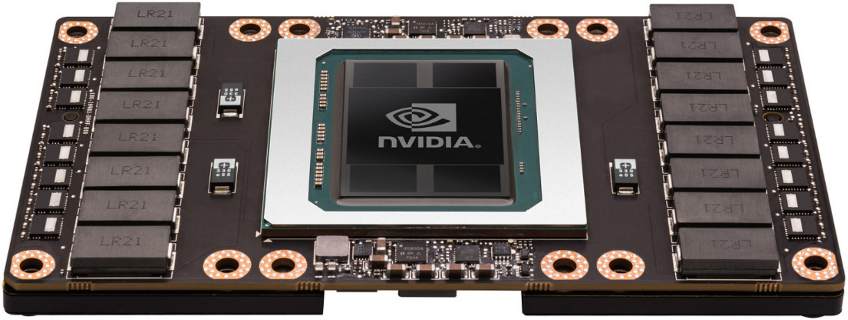
CGYRO fully ported to GPU

NCCS TITAN (Oak Ridge, TN) – **K20x**

GENERAL ATOMICS

CGYRO fully ported to GPU

CSCS PIZ DAINT (Lugano, Switzerland) – **P100**

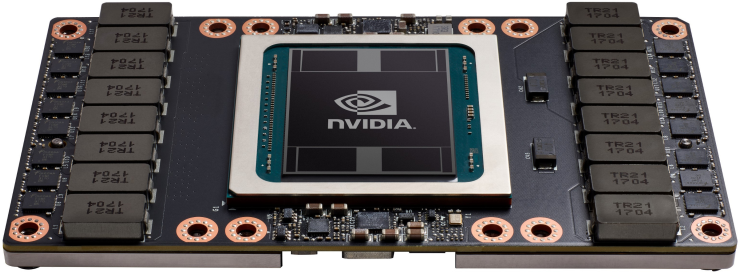


CGYRO fully ported to GPU

CSCS PIZ DAINT (Lugano, Switzerland) – **P100**

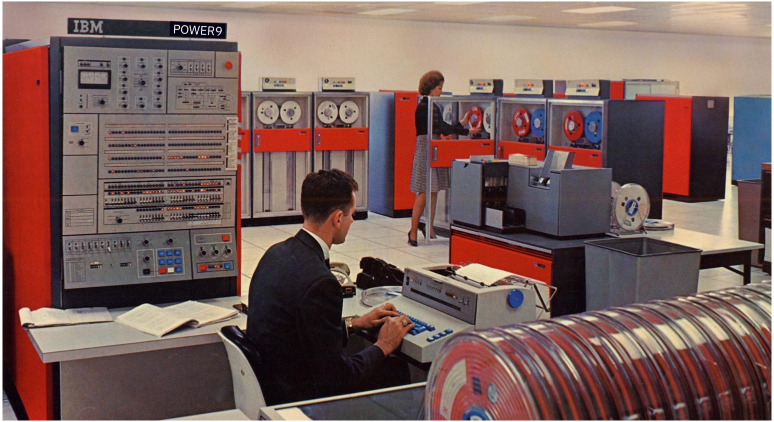
CGYRO fully ported to GPU

General Atomics Power9 (San Diego, CA) – **V100**



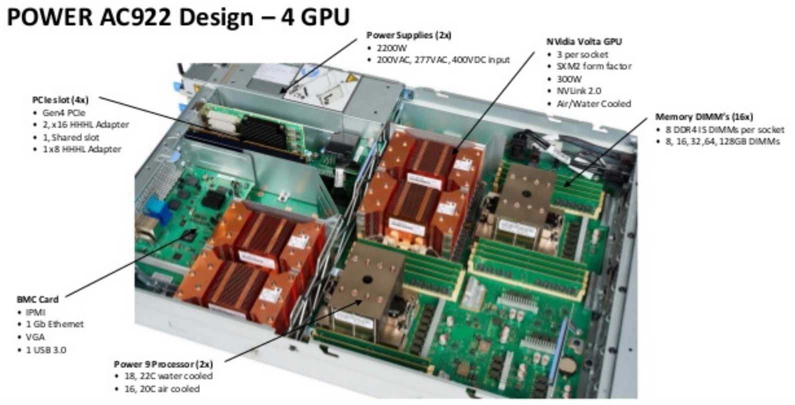
CGYRO fully ported to GPU

General Atomics Power9 (San Diego, CA) – **V100**



CGYRO fully ported to GPU

General Atomics Power9 (San Diego, CA) – **V100**



GENERAL ATOMICS

History of Energy Research at GA

General Atomics – June 25th, 1959

GENERAL ATOMICS

Gyrokinetic equation for plasma species a

Typically: $a = (\text{deuterium, carbon, electron})$

$$\boxed{\frac{\partial \widetilde{h}_a}{\partial \tau}} - i\Omega_s X \boxed{\widetilde{h}_a} - i\left(\Omega_\theta + \Omega_\xi + \Omega_\mathrm{d}\right) \boxed{\widetilde{H}_a} - i\Omega_* \boxed{\widetilde{\Psi}_a} + \Omega_\mathrm{NL}\left(\boxed{\widetilde{h}_a}, \boxed{\widetilde{\Psi}_a}\right) = \mathcal{C}_a$$

Symbol definitions

particles $\qquad \boxed{\widetilde{H}_a} = \widetilde{h}_a + \dfrac{z_a T_e}{T_a} \widetilde{\Psi}_a$

GENERAL ATOMICS

Gyrokinetic equation for plasma species a

Typically: $a = ($deuterium, carbon, electron$)$

$$\boxed{\frac{\partial \widetilde{h}_a}{\partial \tau}} - i\Omega_s X \boxed{\widetilde{h}_a} - i\left(\Omega_\theta + \Omega_\xi + \Omega_{\mathrm{d}}\right)\boxed{\widetilde{H}_a} - i\Omega_* \boxed{\widetilde{\Psi}_a} + \Omega_{\mathrm{NL}}\left(\boxed{\widetilde{h}_a}, \boxed{\widetilde{\Psi}_a}\right) = \mathcal{C}_a$$

Symbol definitions

particles $\boxed{\widetilde{H}_a} = \widetilde{h}_a + \frac{z_a T_e}{T_a}\widetilde{\Psi}_a$

fields $\boxed{\widetilde{\Psi}_a} = J_0(\gamma_a)\left(\delta\widetilde{\phi} - \frac{v_\parallel}{c}\delta\widetilde{A}_\parallel\right) + \frac{v_\perp^2}{\Omega_{ca} c}\frac{J_1(\gamma_a)}{\gamma_a}\delta\widetilde{B}_\parallel$

Electromagnetic GK-Maxwell Equations

Coupling to fields is a MAJOR complication!

$$\left(k_\perp^2 \lambda_D^2 + \sum_a z_a^2 \frac{T_e}{T_a} \int d^3v \frac{f_{0a}}{n_e}\right) \boxed{\delta\widetilde{\phi}} = \sum_a z_a \int d^3v \, \frac{f_{0a}}{n_e} J_0(\gamma_a) \, \boxed{\widetilde{H}_a}$$

$$\frac{2}{\beta_{e,\mathrm{unit}}} k_\perp^2 \rho_s^2 \, \boxed{\delta\widetilde{A}_\parallel} = \sum_a z_a \int d^3v \, \frac{f_{0a}}{n_e} \frac{v_\parallel}{c_s} J_0(\gamma_a) \, \boxed{\widetilde{H}_a}$$

$$-\frac{2}{\beta_{e,\mathrm{unit}}} \frac{B}{B_\mathrm{unit}} \, \boxed{\delta\widetilde{B}_\parallel} = \sum_a \int d^3v \, \frac{f_{0a}}{n_e} \frac{m_a \mathrm{v}_\perp^2}{T_e} \frac{J_1(\gamma_a)}{\gamma_a} \, \boxed{\widetilde{H}_a}$$

Gyrokinetic equation for plasma species a

Typically, deuterium, some carbon, and electrons

$$\frac{\partial \widetilde{h}_a}{\partial \tau} - i\,\boxed{\Omega_s X \widetilde{h}_a} - i\,(\Omega_\theta + \Omega_\xi + \Omega_{\rm d})\,\widetilde{H}_a - i\Omega_* \widetilde{\Psi}_a + \Omega_{\rm NL}(\widetilde{h}_a, \widetilde{\Psi}_a) = \mathcal{C}_a$$

$\mathbf{E} \times \mathbf{B}$ flow

$$-i\Omega_s = -i\,\frac{k_\theta L}{2\pi}\,\frac{a}{c_s}\gamma_E$$

```
$acc parallel loop
```


GENERAL ATOMICS

Gyrokinetic equation for plasma species a

Typically, deuterium, some carbon, and electrons

$$\frac{\partial \widetilde{h}_a}{\partial \tau} - i\Omega_s X \widetilde{h}_a - i\left(\boxed{\Omega_\theta} + \Omega_\xi + \Omega_\mathrm{d}\right)\widetilde{H}_a - i\Omega_* \widetilde{\Psi}_a + \Omega_\mathrm{NL}(\widetilde{h}_a, \widetilde{\Psi}_a) = \mathcal{C}_a$$

Streaming

$$-i\Omega_\theta = \frac{v_\parallel}{w_s}\frac{\partial}{\partial\theta}$$

```
$acc parallel loop
```


GENERAL ATOMICS

Gyrokinetic equation for plasma species a

Typically, deuterium, some carbon, and electrons

$$\frac{\partial \widetilde{h}_a}{\partial \tau} - i\Omega_s X \widetilde{h}_a - i\left(\Omega_\theta + \boxed{\Omega_\xi} + \Omega_{\mathrm{d}}\right)\widetilde{H}_a - i\Omega_* \widetilde{\Psi}_a + \Omega_{\mathrm{NL}}(\widetilde{h}_a, \widetilde{\Psi}_a) = \mathcal{C}_a$$

Trapping

$$\begin{aligned} -i\Omega_\xi = &- \frac{v_{ta}}{\mathrm{w}_s}\frac{u_a}{\sqrt{2}}\left(1-\xi^2\right)\frac{\partial \ln B}{\partial \theta}\frac{\partial}{\partial \xi} \\ &- \frac{1}{2u_a}\frac{\partial \lambda_a}{\partial \theta}\left[\frac{v_\parallel}{\mathrm{w}_s}\frac{\partial}{\partial u_a} + \frac{\sqrt{2}v_{ta}}{\mathrm{w}_s}\left(1-\xi^2\right)\frac{\partial}{\partial \xi}\right] \end{aligned}$$

Fold into collision operator

Gyrokinetic equation for plasma species a

Typically, deuterium, some carbon, and electrons

$$\frac{\partial \widetilde{h}_a}{\partial \tau} - i\Omega_s X \widetilde{h}_a - i\left(\Omega_\theta + \Omega_\xi + \boxed{\Omega_\mathrm{d}}\right)\widetilde{H}_a - i\Omega_* \widetilde{\Psi}_a + \Omega_\mathrm{NL}(\widetilde{h}_a, \widetilde{\Psi}_a) = \mathcal{C}_a$$

Drift motion

$$\begin{aligned} -i\Omega_\mathrm{d} = {} & a\frac{v_{ta}}{c_s}\mathbf{b} \times \left[u_a^2\left(1+\xi^2\right)\frac{\nabla B}{B} + u_a^2\xi^2\frac{8\pi}{B^2}\left(\nabla p\right)_\mathrm{eff}\right] \cdot i\mathbf{k}_\perp \rho_a \\ & + M_a\frac{2av_\parallel}{c_s R_0}\mathbf{b} \times \left(\frac{R}{\mathcal{J}_\psi B}\frac{\partial R}{\partial \theta}\nabla\varphi - \frac{B_t}{B}\nabla R\right) \cdot i\mathbf{k}_\perp \rho_a \\ & + \frac{a}{c_s}\mathbf{b} \times \left(-\frac{v_{ta}}{T_a}\mathbf{F}_c + \frac{c}{B}\nabla\Phi_*\right) \cdot i\mathbf{k}_\perp \rho_a \end{aligned}$$

Fold into streaming (diagonal)

Gyrokinetic equation for plasma species a

Typically, deuterium, some carbon, and electrons

$$\frac{\partial \widetilde{h}_a}{\partial \tau} - i\Omega_s X \widetilde{h}_a - i\left(\Omega_\theta + \Omega_\xi + \Omega_\mathrm{d}\right)\widetilde{H}_a - i\,\Omega_* \widetilde{\Psi}_a + \Omega_\mathrm{NL}(\widetilde{h}_a, \widetilde{\Psi}_a) = \mathcal{C}_a$$

Gradient drive

$$\begin{aligned}-i\Omega_* = & \left[\frac{a}{L_{na}} + \frac{a}{L_{Ta}}\left(u_a^2 - \frac{3}{2}\right) + \gamma_p v_\parallel \frac{a}{v_{ta}^2}\frac{RB_t}{R_0 B}\right] i k_\theta \rho_s \\ & + \left\{\frac{a}{L_{Ta}}\left[\frac{z_a e}{T_a}\Phi_* - \frac{M_a^2}{2R_0^2}\left(R^2 - R(\theta_0)^2\right)\right]\right. \\ & \left. + M_a^2 \frac{aR(\theta_0)}{R_0^2}\frac{dR(\theta_0)}{dr} + M_a \gamma_p \frac{a}{v_{ta}R_0^2}\left(R^2 - R(\theta_0)^2\right)\right\} i k_\theta \rho_s\end{aligned}$$

Fold into streaming (diagonal)

Gyrokinetic equation for plasma species a

Typically, deuterium, some carbon, and electrons

$$\frac{\partial \widetilde{h}_a}{\partial \tau} - i\Omega_s X \widetilde{h}_a - i\left(\Omega_\theta + \Omega_\xi + \Omega_{\mathrm{d}}\right)\widetilde{H}_a - i\Omega_* \widetilde{\Psi}_a + \boxed{\Omega_{\mathrm{NL}}(\widetilde{h}_a, \widetilde{\Psi}_a)} = \mathcal{C}_a$$

Nonlinearity

$$\Omega_{\mathrm{NL}}(\widetilde{h}_a, \widetilde{\Psi}_a) = \frac{a c_s}{\Omega_{cD}} \sum_{\mathbf{k}'_\perp + \mathbf{k}''_\perp = \mathbf{k}_\perp} \left(\mathbf{b} \cdot \mathbf{k}'_\perp \times \mathbf{k}''_\perp\right) \widetilde{\Psi}_a(\mathbf{k}'_\perp) \widetilde{h}_a(\mathbf{k}''_\perp)$$

cuFFT

Gyrokinetic equation for plasma species a

Typically, deuterium, some carbon, and electrons

$$\frac{\partial \widetilde{h}_a}{\partial \tau} - i\Omega_s X \widetilde{h}_a - i\left(\Omega_\theta + \Omega_\xi + \Omega_\mathrm{d}\right)\widetilde{H}_a - i\Omega_* \widetilde{\Psi}_a + \Omega_\mathrm{NL}(\widetilde{h}_a, \widetilde{\Psi}_a) = \mathcal{C}_a$$

Cross-species collision operator

$$\mathcal{C}_a = \sum_b C^L_{ab}\left(\widetilde{H}_a, \widetilde{H}_b\right)$$

$$C^L_{ab}(\widetilde{H}_a, \widetilde{H}_b) = \frac{\nu^D_{ab}}{2}\frac{\partial}{\partial \xi}\left(1-\xi^2\right)\frac{\partial \widetilde{H}_a}{\partial \xi} + \frac{1}{\mathrm{v}^2}\frac{\partial}{\partial \mathrm{v}}\left[\frac{\nu^\parallel_{ab}}{2}\left(\mathrm{v}^4\frac{\partial \widetilde{H}_a}{\partial \mathrm{v}} + \frac{m_a}{T_b}\mathrm{v}^5 \widetilde{H}_a\right)\right]$$

$$-\widetilde{H}_a k_\perp^2 \rho_a^2 \frac{\mathrm{v}^2}{4v_{ta}^2}\left[\nu^D_{ab}\left(1+\xi^2\right) + \nu^\parallel_{ab}\left(1-\xi^2\right)\right] + R_\mathrm{mom}(\widetilde{H}_b) + R_\mathrm{ene}(\widetilde{H}_b)$$

```
$acc parallel loop
```

Sonic Transport Fluxes

These are inputs to an independent TRANSPORT CODE

$$\text{particle flux} \quad \boxed{\Gamma_a} = \sum_{\mathbf{k}_\perp} \left\langle \int d^3v \, \widetilde{H}_a^* c_{1a} \widetilde{\Psi}_a \right\rangle$$

$$\text{energy flux} \quad \boxed{Q_a} = \sum_{\mathbf{k}_\perp} \left\langle \int d^3v \, \widetilde{H}_a^* c_{2a} \widetilde{\Psi}_a \right\rangle$$

$$\text{momentum flux} \quad \boxed{\Pi_a} = \sum_{\mathbf{k}_\perp} \left\langle \int d^3v \, \widetilde{H}_a^* c_{3a} \widetilde{\Psi}_a \right\rangle$$

What do we solve for

5-dimensional distribution for every plasma species

Six-dimensional array (mapped into internal 2D array in CGYRO)

$$H_a(\underbrace{k_x, k_y, \theta, \xi, \mathrm{v}}_{\text{5D mesh}}, t)$$

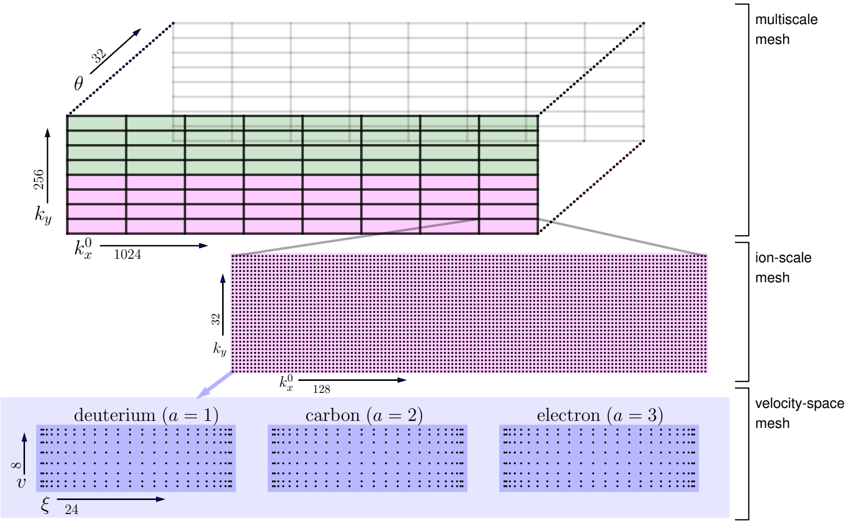
The **spatial coordinates** are

$$\begin{aligned} k_x &\longrightarrow \text{radial wavenumbers} \\ k_y &\longrightarrow \text{binormal wavenumbers} \\ \theta &\longrightarrow \text{field-line coordinate} \end{aligned}$$

The **velocity-space coordinates** are

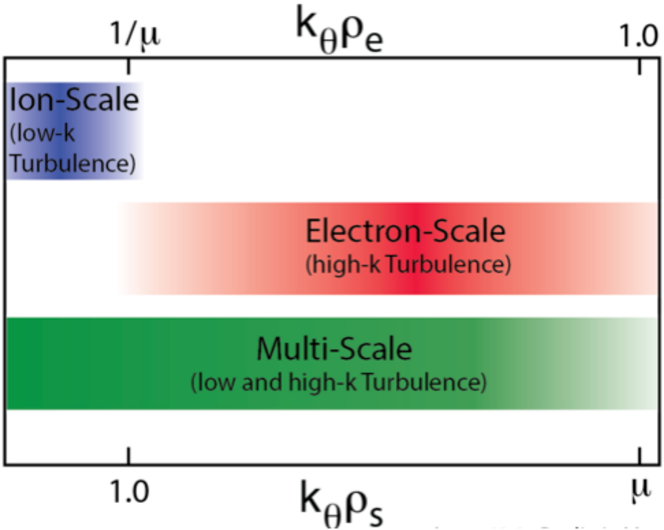
$$\begin{aligned} \xi = v_\parallel/\mathrm{v} &\longrightarrow \text{cosine of the pitch angle} \in [-1, 1] \\ \mathrm{v} &\longrightarrow \text{speed} \in [0, \infty] \ . \end{aligned}$$

Visual representation of computational mesh



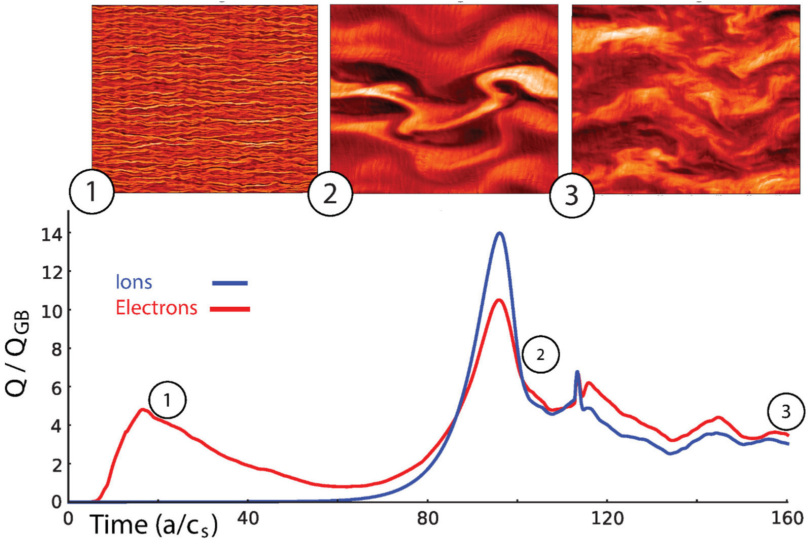
CGYRO optimized for challenging multiscale turbulence

COMPLETE REDESIGN of world-renowned GYRO code



Simulation underway on Titan (NCCS)

4986 nodes = 4986 Tesla K20X GPUs



GENERAL ATOMICS

Recent aggressive GPU optimization

Significant progress at Boulder Hackathon (Summer 2018)

- Huge thanks to **Craig Tierney** and **Brent Leback** for guidance!

Recent aggressive GPU optimization

Significant progress at Boulder Hackathon (Summer 2018)

- Huge thanks to **Craig Tierney** and **Brent Leback** for guidance!
- CGYRO design **anticipated** aggressive thread/GPU utilization
  1. Huge nonlinear convolution (Poisson bracket) via FFT
  2. Large nested loops remain after MPI distribution

Recent aggressive GPU optimization

Significant progress at Boulder Hackathon (Summer 2018)

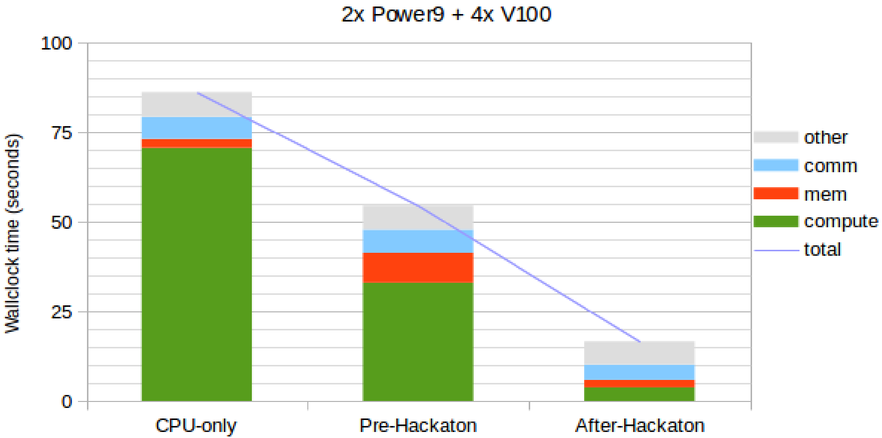
- Huge thanks to **Craig Tierney** and **Brent Leback** for guidance!
- CGYRO design **anticipated** aggressive thread/GPU utilization
  1. Huge nonlinear convolution (Poisson bracket) via FFT
  2. Large nested loops remain after MPI distribution
- Took **full advantage** of GPUs with minimal changes to code logic
  1. Existing FFTW code was ported directly to **cuFFT**
  2. Nested loops accelerated by **OpenACC** without restructuring or invasive changes
  3. Implemented **GPU-aware MPI** (utilizes GPUDirect and GPU-Infiniband RDMA)

CGYRO kernels

Kernel	Data dependence	Dominant operation	GPU approach
`str`	$k_x, \theta, [k_y]_2, [\xi, \mathrm{v}, a]_1$	loop	**OpenACC**
`field`	Same as `str`	loop	**OpenACC**
`coll`	$[k_x, \theta]_1, [k_y]_2, \xi, \mathrm{v}, a$	mat-vec multiply	**OpenACC**
`nl`	$k_x, k_y, [\theta, [\xi, \mathrm{v}, a]_1]_2$	FFT	**cuFFT**

Scaling: CGYRO nl01

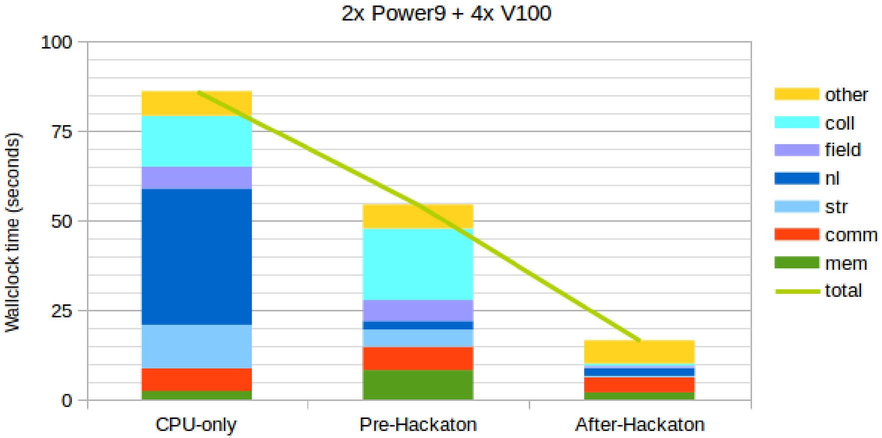
V100-GPU Performance improvement over time



GENERAL ATOMICS

Scaling: CGYRO nl01 (individual kernels)

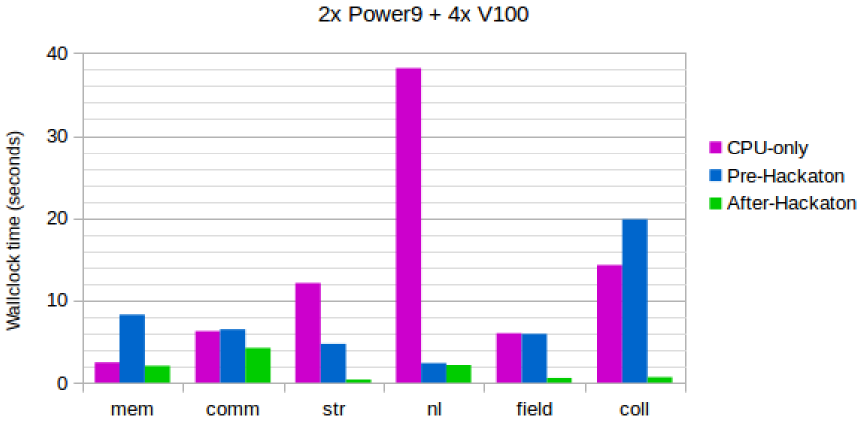
V100-GPU Performance improvement over time



GENERAL ATOMICS

Scaling: CGYRO nl01 (individual kernels)

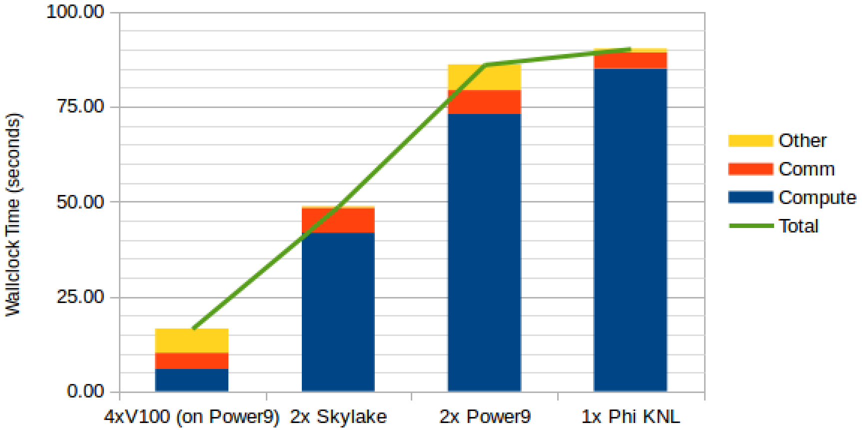
V100-GPU Performance improvement over time



GENERAL ATOMICS

Scaling: CGYRO nl01

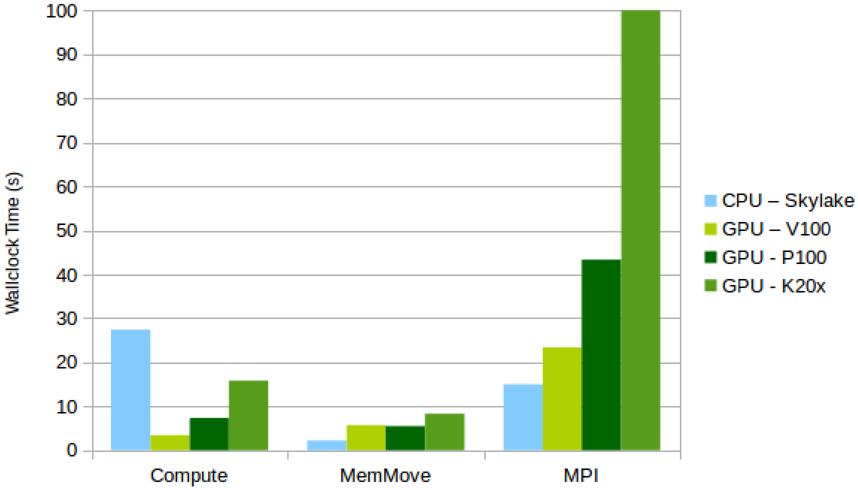
GPU versus Skylake and KNL



GENERAL ATOMICS

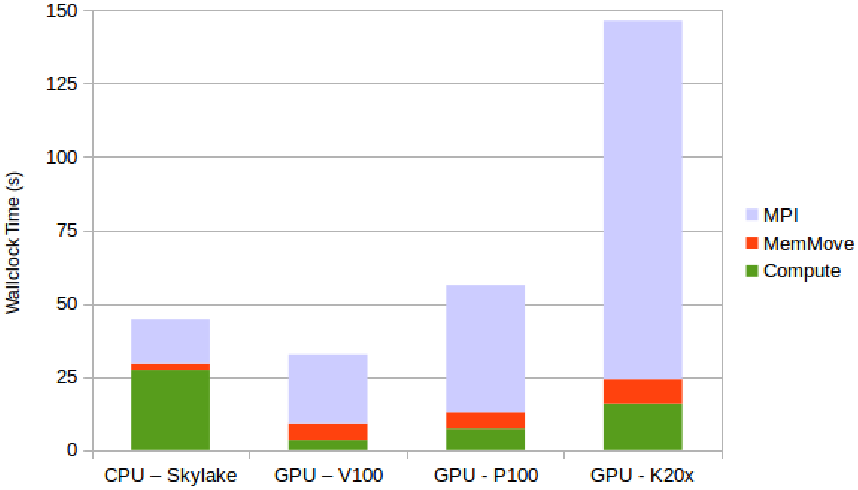
Scaling: CGYRO nl03 – much larger case

Skylake versus 3 different GPUs



Scaling: CGYRO `nl03` – much larger case

Skylake versus 3 different GPUs



Disclaimer

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, or those of the European Commission.