

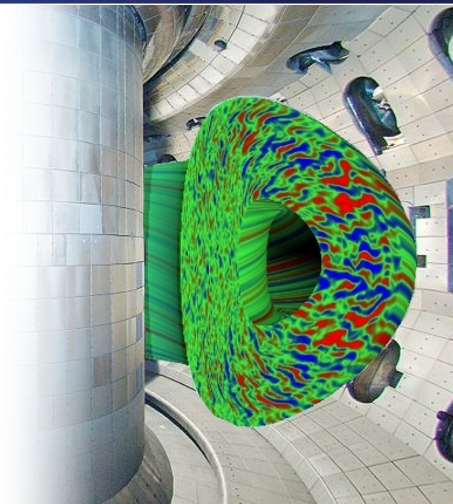
# Advancing Fusion Science with CGYRO using GPU-Based Leadership Systems

by  
**J. Candy<sup>1</sup>**, I. Sfiligoi<sup>2</sup> and E. Belli<sup>1</sup>.

<sup>1</sup>General Atomics, San Diego, CA

<sup>2</sup>San Diego Supercomputer Center, San Diego CA

Presented at  
**GTC 2019**  
**San Jose, CA**  
**18-21 March 2019**  
**ID: S9202**



# Sincere thanks to

- Chris Holland (UCSD)
- Orso Meneghini, Sterling Smith, Ron Waltz, Gary Staebler (GA)
- Nathan Howard, Alessandro Marinoni (MIT)
- Walter Guttenfelder, Brian Grierson (PPPL)
- George Fann (ORNL)
- Klaus Hallatschek (IPP, Germany)

# OUTLINE

## ① Who is General Atomics?

# OUTLINE

- ① Who is General Atomics?
- ② The case for **fusion energy**



# OUTLINE

- ① Who is General Atomics?
- ② The case for **fusion energy**
- ③ Mathematical formulation and GPU-based numerical solution

# OUTLINE

- ① Who is General Atomics?
- ② The case for **fusion energy**
- ③ Mathematical formulation and GPU-based numerical solution
- ④ Simulation of turbulent energy loss in a tokamak plasma

# OUTLINE

- ① Who is General Atomics?
- ② The case for **fusion energy**
- ③ Mathematical formulation and GPU-based numerical solution
- ④ Simulation of turbulent energy loss in a tokamak plasma
- ⑤ GPU performance: development and results

# Who is General Atomics?

# Who is General Atomics?

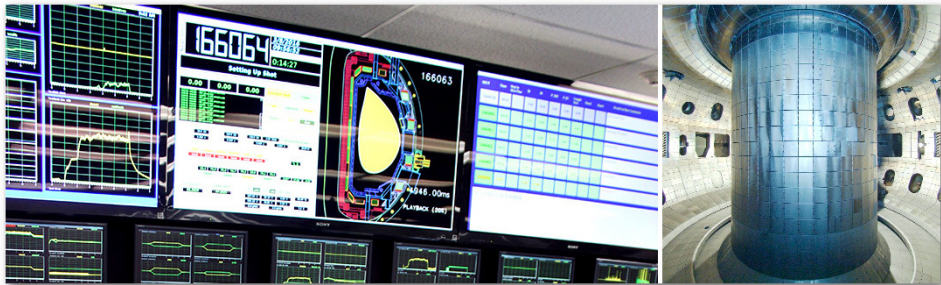
- ① **General Atomics** (GA) is a private contractor in San Diego

# Who is General Atomics?

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

# Who is General Atomics?

- 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**



# Founded on July 18, 1955 (photo 1957)

## The General Atomic Division of General Dynamics





# Laboratory formally dedicated on June 25th, 1959

John Jay Hopkins Laboratory for Pure and Applied Science

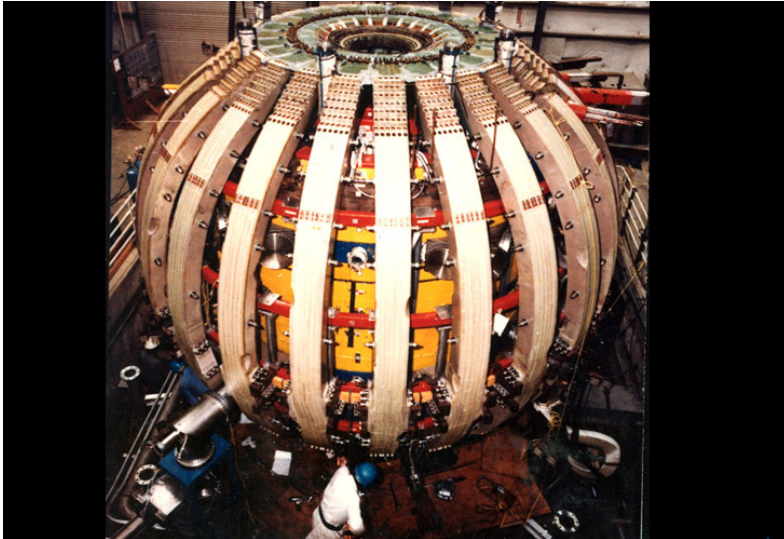


# Present-day Campus (2019)

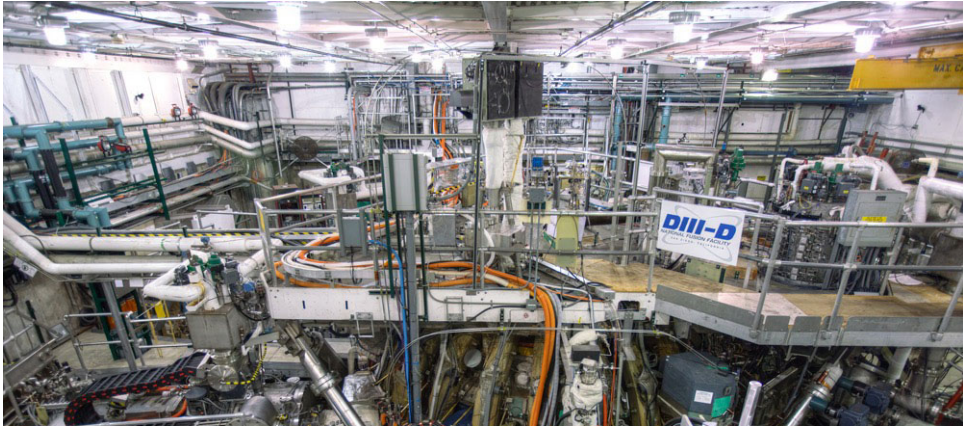
Retains feel of early architecture



# Doublet III (1974)



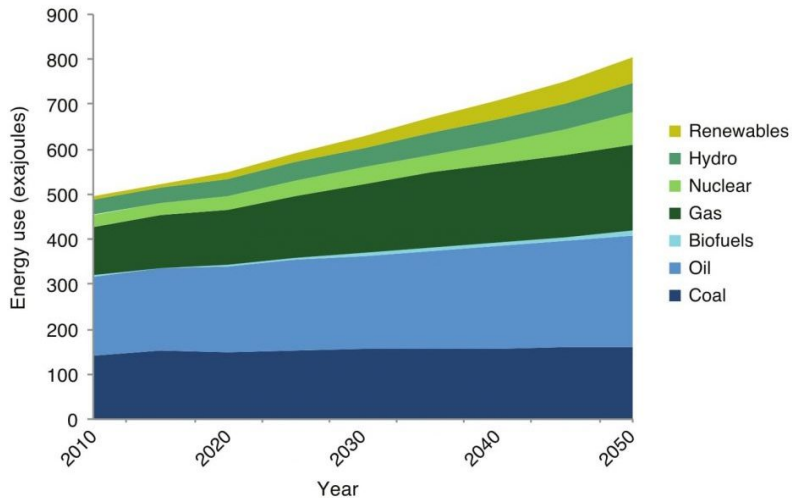
# DIII-D (Present day)



# The case for fusion energy

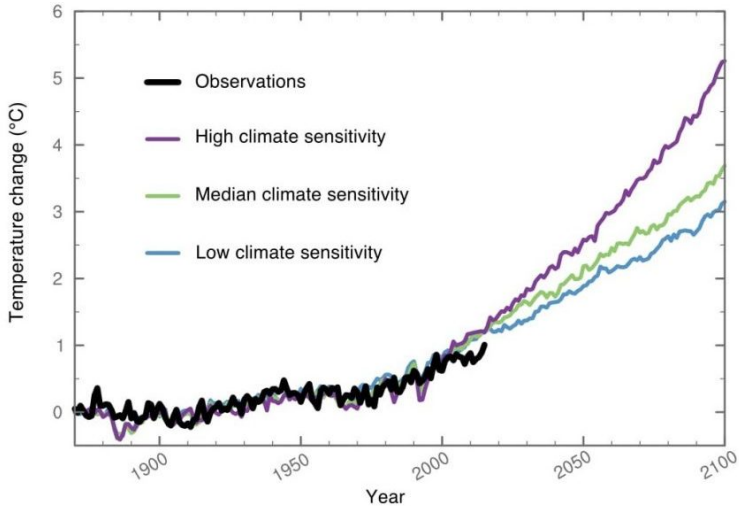
# Energy Use by Technology and Year

[energy.mit.edu/news/limiting-global-warming-aggressive-measures-needed](https://energy.mit.edu/news/limiting-global-warming-aggressive-measures-needed)

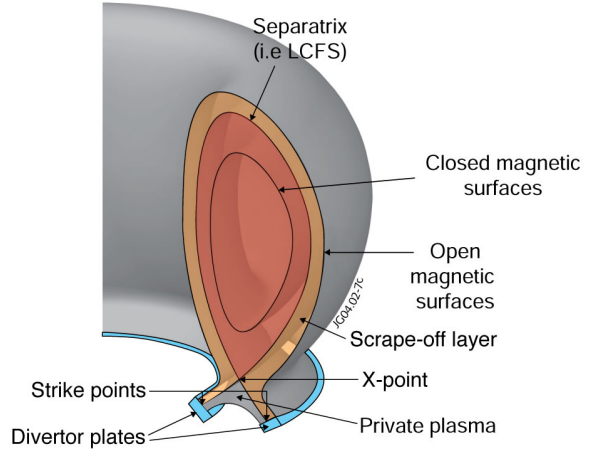
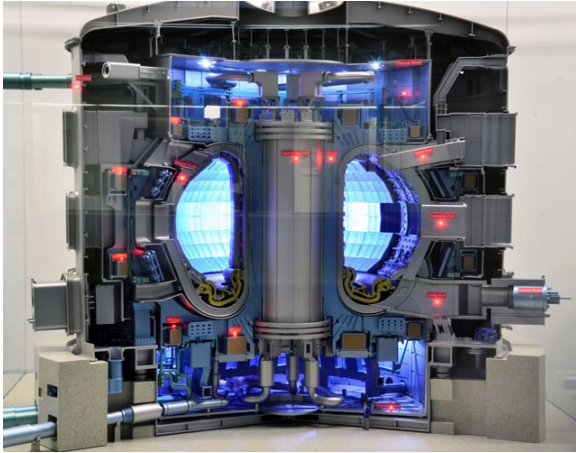


# Surface Temperature Anomaly

[energy.mit.edu/news/limiting-global-warming-aggressive-measures-needed](https://energy.mit.edu/news/limiting-global-warming-aggressive-measures-needed)

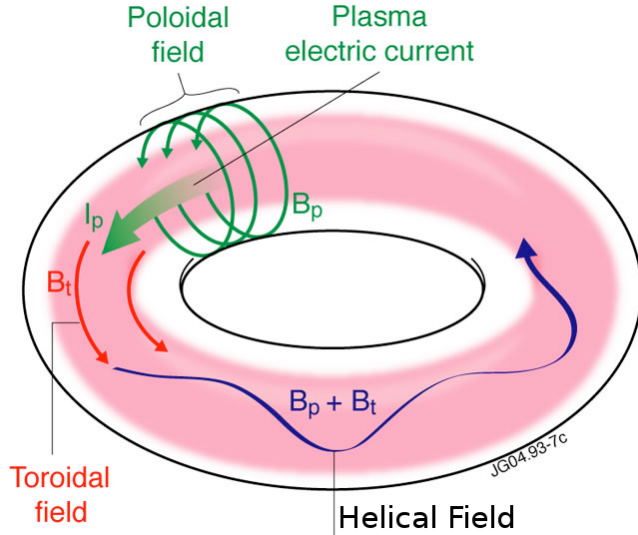


# Plasma theory in closed fieldline region well-understood

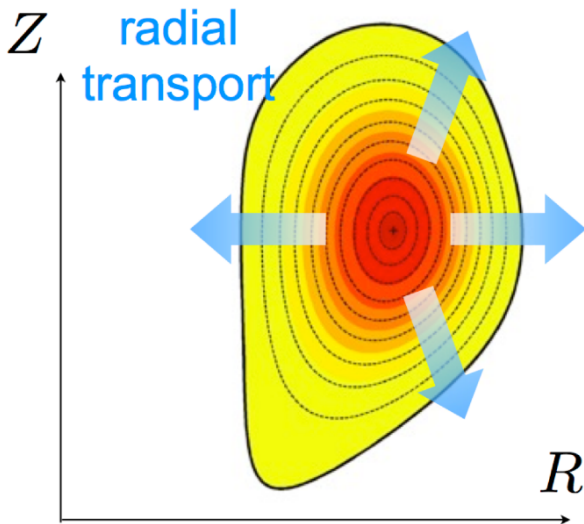




# Helical field perfectly confines plasma (almost)



# There is a small amount of radial energy/particle loss



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

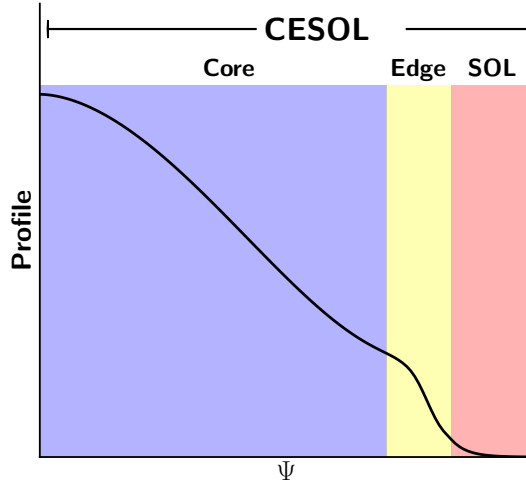
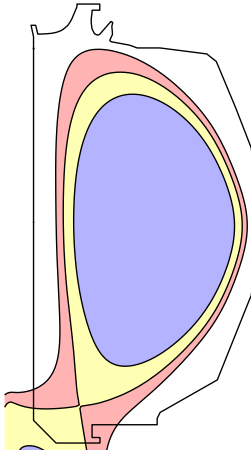
**flux**  $\Gamma \sim v(\rho/a)^2$

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

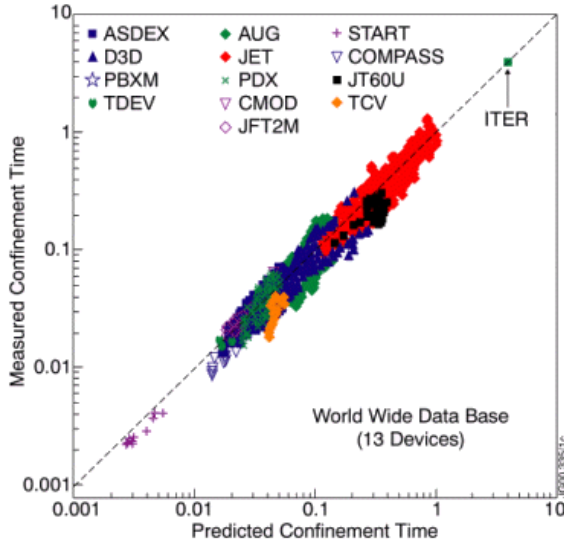
- $a$  = torus radius
- $\rho$  = particle orbit size
- $v$  = particle velocity

# Tokamak physics spans multiple space/timescales

## Core-edge-SOL (CESOL) region coupling

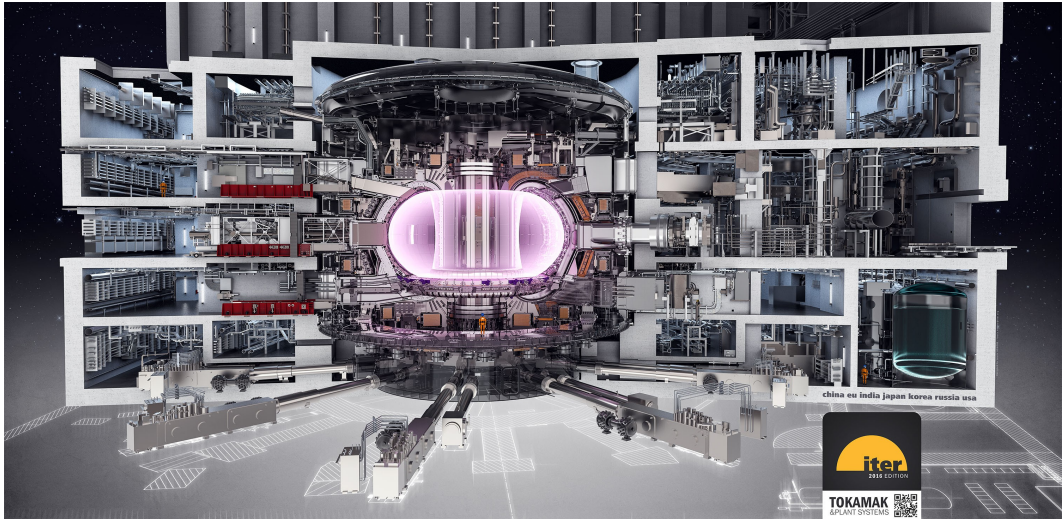


# Tokamak confinement improves with LARGE PLASMA VOLUME



# ITER Facility (35 nations) under construction in France

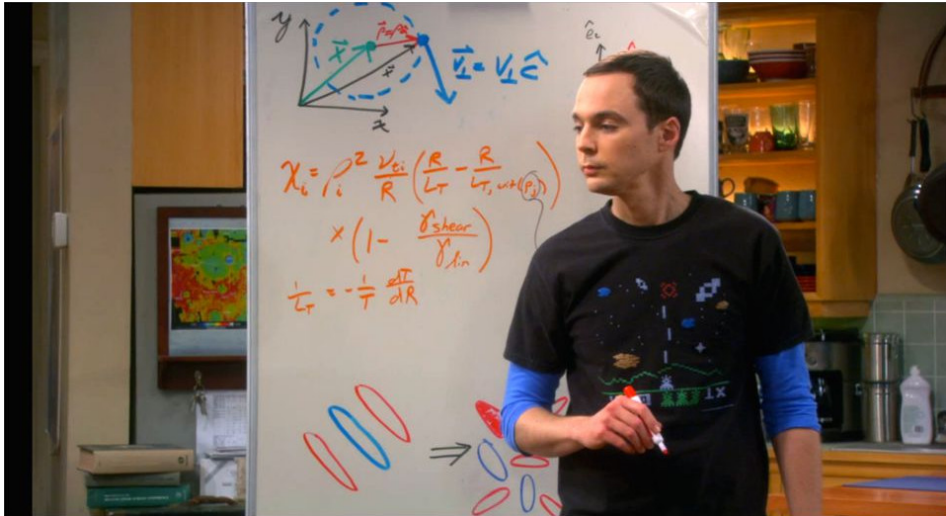
GOAL: Simulate turbulent plasma in core (magenta) region



# Mathematical formulation and GPU-based numerical solution

# Gyrokinetic Theory for Magnetized Plasma

## The Cooper/Kripke Inversion



# Gyrokinetic equation for plasma species $a$

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

$$\frac{\partial \tilde{h}_a}{\partial \tau} - i\Omega_s X \tilde{h}_a - i(\Omega_\theta + \Omega_\xi + \Omega_d) \tilde{H}_a - i\Omega_* \tilde{\Psi}_a + \Omega_{\text{NL}}(\tilde{h}_a, \tilde{\Psi}_a) = \mathcal{C}_a$$

Symbol definitions

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



# Gyrokinetic equation for plasma species $a$

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

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

## Symbol definitions

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

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

# Electromagnetic GK-Maxwell Equations

Coupling to fields is a MAJOR complication!

$$\begin{aligned} \left( k_{\perp}^2 \lambda_D^2 + \sum_a z_a^2 \frac{T_e}{T_a} \int d^3v \frac{f_{0a}}{n_e} \right) \delta\tilde{\Phi} &= \sum_a z_a \int d^3v \frac{f_{0a}}{n_e} J_0(\gamma_a) \tilde{H}_a \\ \frac{2}{\beta_{e,\text{unit}}} k_{\perp}^2 \rho_s^2 \delta\tilde{A}_{\parallel} &= \sum_a z_a \int d^3v \frac{f_{0a}}{n_e} \frac{v_{\parallel}}{c_s} J_0(\gamma_a) \tilde{H}_a \\ -\frac{2}{\beta_{e,\text{unit}}} \frac{B}{B_{\text{unit}}} \delta\tilde{B}_{\parallel} &= \sum_a \int d^3v \frac{f_{0a}}{n_e} \frac{m_a v_{\perp}^2}{T_e} \frac{J_1(\gamma_a)}{\gamma_a} \tilde{H}_a \end{aligned}$$

# Gyrokinetic equation for plasma species $a$

Typically, deuterium, some carbon, and electrons

$$\frac{\partial \tilde{h}_a}{\partial \tau} - i \Omega_s X \tilde{h}_a - i (\Omega_\theta + \Omega_\xi + \Omega_d) \tilde{H}_a - i \Omega_* \tilde{\Psi}_a + \Omega_{\text{NL}}(\tilde{h}_a, \tilde{\Psi}_a) = \mathcal{C}_a$$

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

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

# Gyrokinetic equation for plasma species $a$

Typically, deuterium, some carbon, and electrons

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

## Streaming

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

# Gyrokinetic equation for plasma species $a$

Typically, deuterium, some carbon, and electrons

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

## Trapping

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

# Gyrokinetic equation for plasma species $a$

Typically, deuterium, some carbon, and electrons

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

## Drift motion

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

# Gyrokinetic equation for plasma species $a$

Typically, deuterium, some carbon, and electrons

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

## Gradient drive

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

# Gyrokinetic equation for plasma species $a$

Typically, deuterium, some carbon, and electrons

$$\frac{\partial \tilde{h}_a}{\partial \tau} - i\Omega_s X \tilde{h}_a - i(\Omega_\theta + \Omega_\xi + \Omega_d) \tilde{H}_a - i\Omega_* \tilde{\Psi}_a + \Omega_{\text{NL}}(\tilde{h}_a, \tilde{\Psi}_a) = \mathcal{C}_a$$

## Nonlinearity

$$\Omega_{\text{NL}}(\tilde{h}_a, \tilde{\Psi}_a) = \frac{ac_s}{\Omega_{cD}} \sum_{\mathbf{k}'_\perp + \mathbf{k}''_\perp = \mathbf{k}_\perp} (\mathbf{b} \cdot \mathbf{k}'_\perp \times \mathbf{k}''_\perp) \tilde{\Psi}_a(\mathbf{k}'_\perp) \tilde{h}_a(\mathbf{k}''_\perp)$$



# Gyrokinetic equation for plasma species $a$

Typically, deuterium, some carbon, and electrons

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

## Cross-species collision operator

$$\mathcal{C}_a = \sum_b C_{ab}^L(\tilde{H}_a, \tilde{H}_b)$$

$$\begin{aligned} C_{ab}^L(\tilde{H}_a, \tilde{H}_b) = & \frac{v_{ab}^D}{2} \frac{\partial}{\partial \xi} (1 - \xi^2) \frac{\partial \tilde{H}_a}{\partial \xi} + \frac{1}{v^2} \frac{\partial}{\partial v} \left[ \frac{v_{ab}^\parallel}{2} \left( v^4 \frac{\partial \tilde{H}_a}{\partial v} + \frac{m_a}{T_b} v^5 \tilde{H}_a \right) \right] \\ & - \tilde{H}_a k_\perp^2 \rho_a^2 \frac{v^2}{4v_{ta}^2} \left[ v_{ab}^D (1 + \xi^2) + v_{ab}^\parallel (1 - \xi^2) \right] + R_{\text{mom}}(\tilde{H}_b) + R_{\text{ene}}(\tilde{H}_b) \end{aligned}$$

# Sonic Transport Fluxes

These are inputs to an independent TRANSPORT CODE

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

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

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

# What do we solve for

5-dimensional distribution for every plasma species

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

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

The **spatial** coordinates are

$k_x \longrightarrow$  radial wavenumbers

$k_y \longrightarrow$  binormal wavenumbers

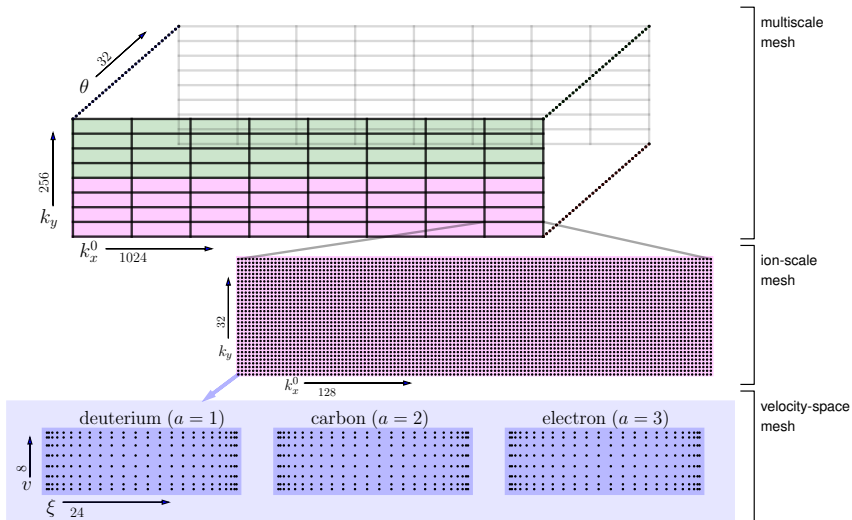
$\theta \longrightarrow$  field-line coordinate

The **velocity-space** coordinates are

$\xi = v_{\parallel}/v \longrightarrow$  cosine of the pitch angle  $\in [-1, 1]$

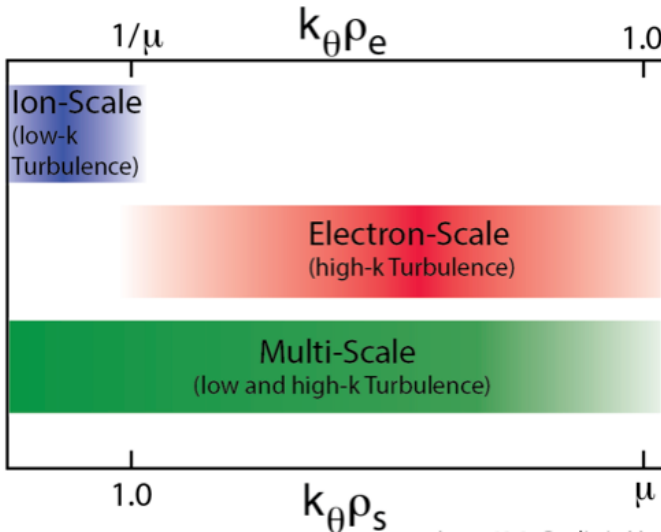
$v \longrightarrow$  speed  $\in [0, \infty]$ .

# Visual representation of computational mesh



# CGYRO optimized for challenging multiscale turbulence

COMPLETE REDESIGN of world-renowned GYRO code



# Simulation of turbulent energy loss in a tokamak plasma

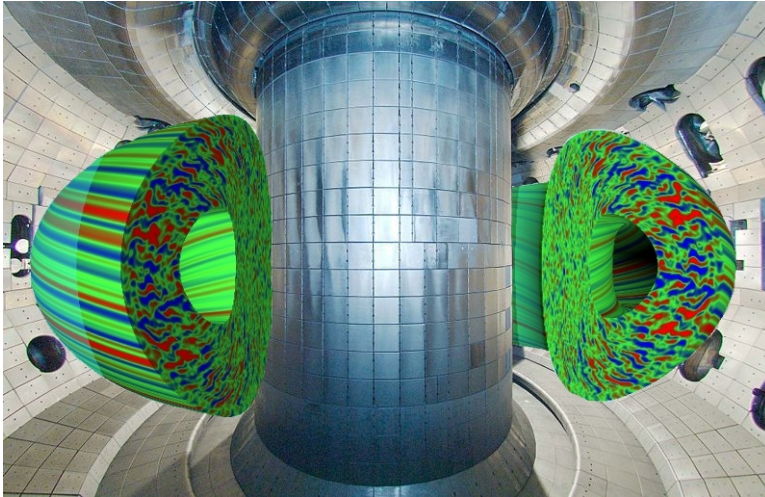
# CGYRO computes the turbulent flux

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



# CGYRO computes the turbulent flux

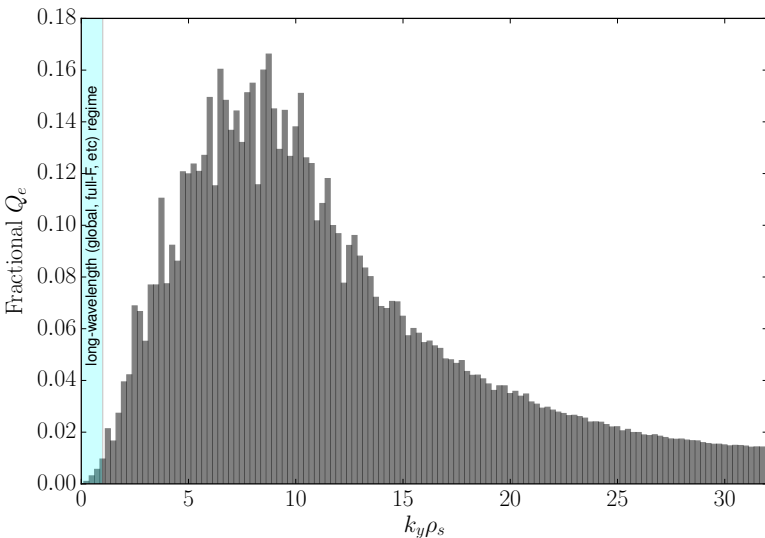
## DIII-D Tokamak at General Atomics in San Diego, CA





# Multiscale DIII-D Simulation at $r/a = 0.92$

ITER baseline discharge (Haskey, Grierson) 164988



## Resolution

$$k_x \rho_s \leq 124.0, \quad k_y \rho_s \leq 31.8$$

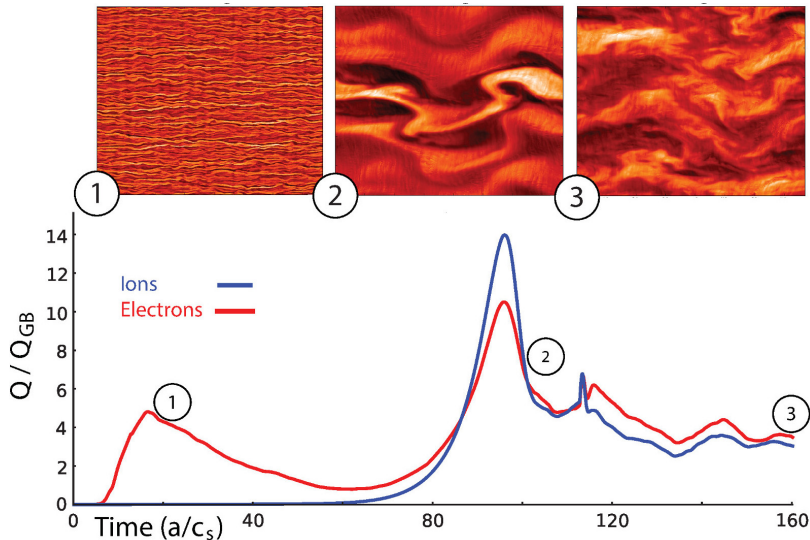
## Time

9 hrs on 32K cores

	$Q_i/Q_{GB}$	$Q_e/Q_{GB}$
pwrbal	2.5	8.2
NEO	2.7	0.0
CGYRO	0.0	8.0

# Simulation underway on Titan (NCCS)

4986 nodes = 4986 Tesla K20X GPUs



# Important locations for CGYRO

## Source code

[github.com/gafusion/gacode](https://github.com/gafusion/gacode)

## DOI

[www.osti.gov/doecode/biblio/20298](http://www.osti.gov/doecode/biblio/20298)

## User Documentation

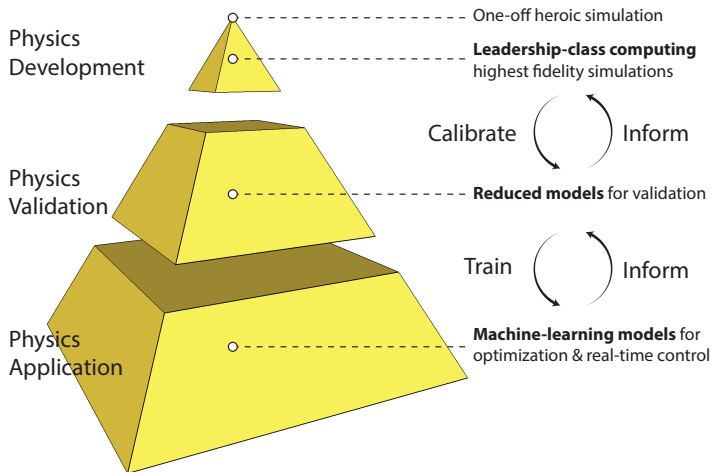
[gafusion.github.io/doc](https://gafusion.github.io/doc)

## Documentary Video (for GYRO)

[www.youtube.com/watch?v=RLI6QW2x4Lg](https://www.youtube.com/watch?v=RLI6QW2x4Lg)

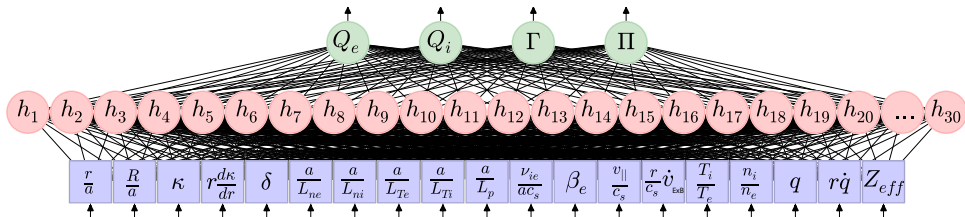
# Fidelity Hierarchy (Pyramid)

Range of models all the way up to leadership codes



# Create TGLF-NN neural net from TGLF reduced model

- **23 inputs**  $\rightarrow$  **4 outputs**
- Each dataset has 500K cases from 2300 multi-machine discharges
- Trained with TENSORFLOW
- Must be retrained as TGLF model is updated
- TGLF itself derived from **HPC CGYRO simulation**



# GPU performance: development and results

# CGYRO: Roadmap for efficient GPU implementation

- ① Numerical algorithms selected to allow intensive threading/acceleration
  - Nonlinearity (nl) = FFT
  - Collisions (coll) = Matrix-vector multiply

# CGYRO: Roadmap for efficient GPU implementation

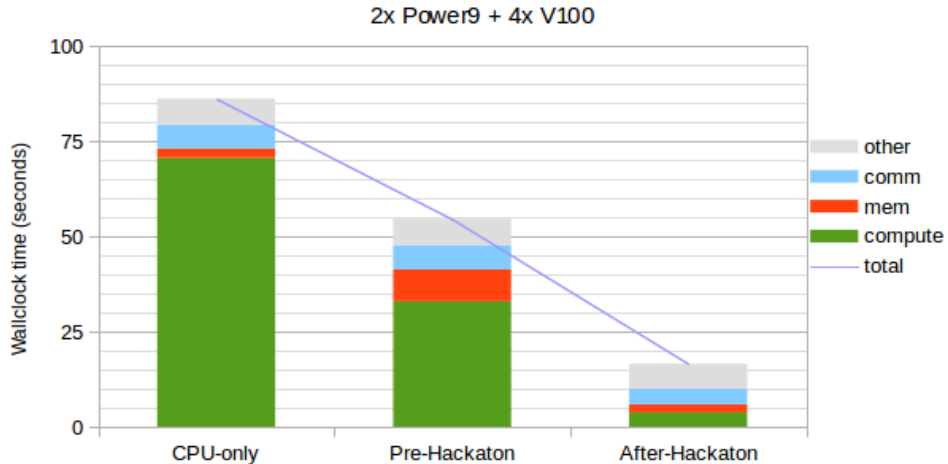
- ① Numerical algorithms selected to allow intensive threading/acceleration
  - Nonlinearity (nl) = FFT
  - Collisions (coll) = Matrix-vector multiply
- ② Key kernels have threaded (default) and accelerated variations
  - Smart loop order and good memory management keeps kernels similar



# CGYRO: Roadmap for efficient GPU implementation

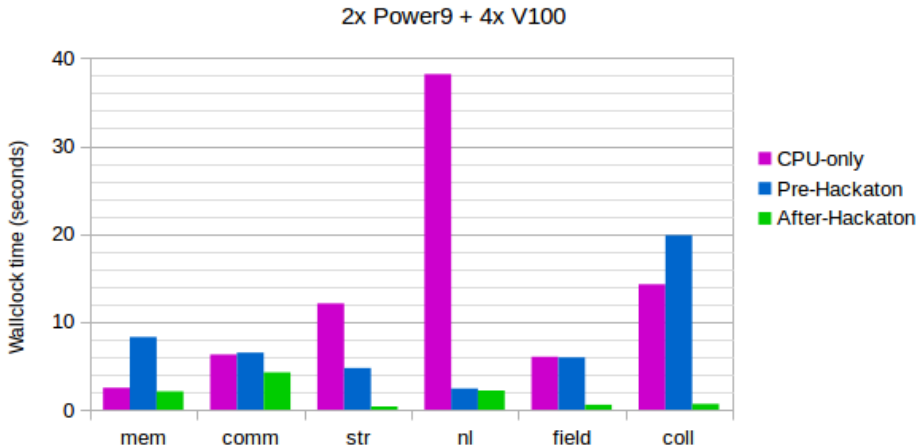
- ① Numerical algorithms selected to allow intensive threading/acceleration
  - Nonlinearity (nl) = FFT
  - Collisions (coll) = Matrix-vector multiply
- ② Key kernels have threaded (default) and accelerated variations
  - Smart loop order and good memory management keeps kernels similar
- ③ Implemented **GPU-aware MPI** (utilizes GPUDirect and GPU-Infiniband RDMA)

# Initial thought was that nonlinearity (nl) would dominate



# Acceleration of nl exposed cost of other kernels

Titan K20 GPU too small to store collision matrix



# CGYRO: Roadmap for efficient GPU implementation

- ① Numerical algorithms selected to allow intensive threading/acceleration
  - Nonlinearity (nl) = FFT
  - Collisions (coll) = Matrix-vector multiply
- ② Key kernels have threaded (default) and accelerated variations
  - Smart loop order and good memory management keeps kernels similar
- ③ Implemented **GPU-aware MPI** (utilizes GPUDirect and GPU-Infiniband RDMA)

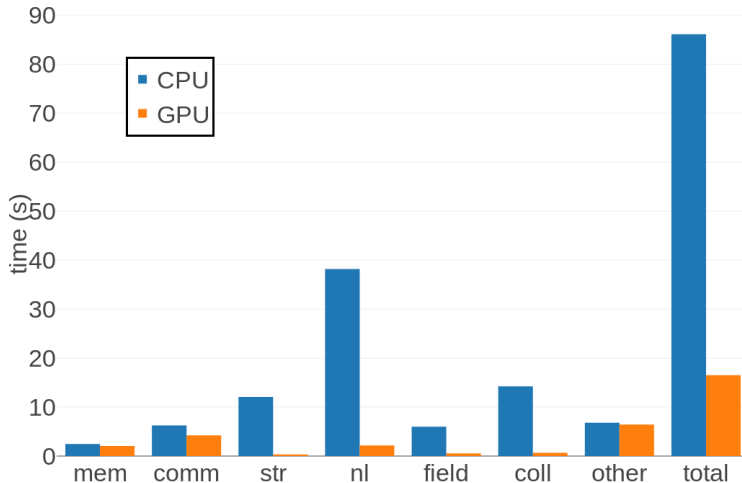
# CGYRO: Roadmap for efficient GPU implementation

```
!$acc loop seq
    do ivp=1,nv
        cvec_re = real(cvec(ivp))
        cvec_im = aimag(cvec(ivp))
!$acc loop vector
        do iv=1,nv
            cval = cmat(iv,ivp,ic_loc)
            bvec(iv) = bvec(iv) + cmplx(cval*cvec_re,cval*cvec_im)
        enddo
    enddo
```

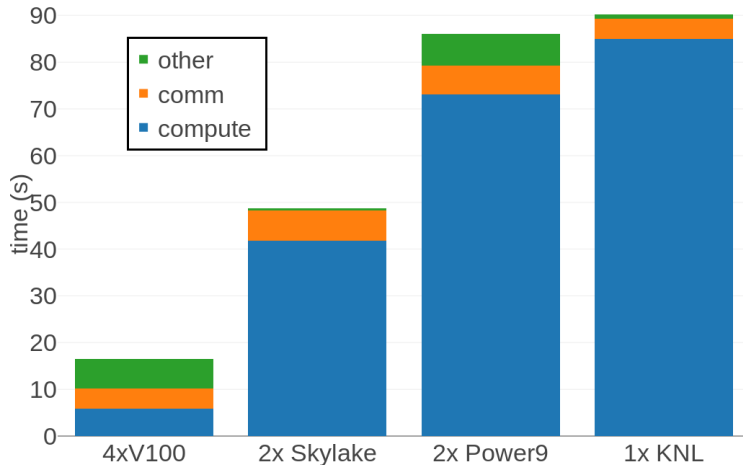
# CGYRO: Roadmap for efficient GPU implementation

```
#ifdef DISABLE_GPUDIRECT_MPI
!$acc update host(fsendr)
#else
!$acc host_data use_device(fsendr,f)
#endif
    call MPI_ALLTOALL(fsendr,nsend,MPI_DOUBLE_COMPLEX, &
                      f,      nsend,MPI_DOUBLE_COMPLEX,lib_comm,ierr)
#ifdef DISABLE_GPUDIRECT_MPI
!$acc update device(f)
#else
!$acc end host_data
#endif
```

# Power9 (CPU) versus Power9 + 4X V100 (GPU)



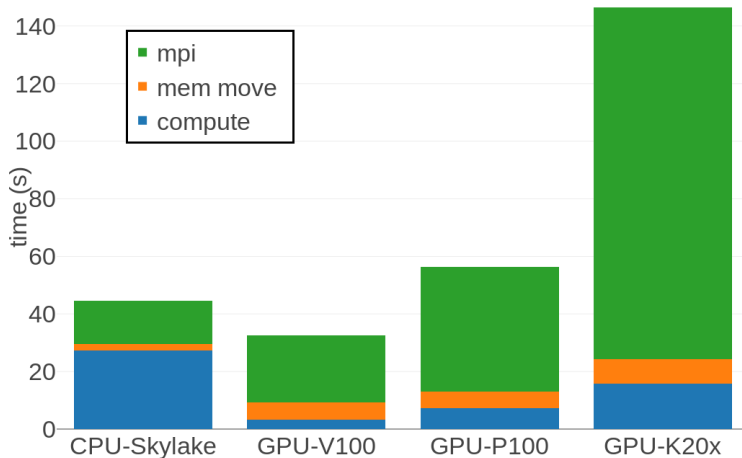
# CPU systems versus 4X V100





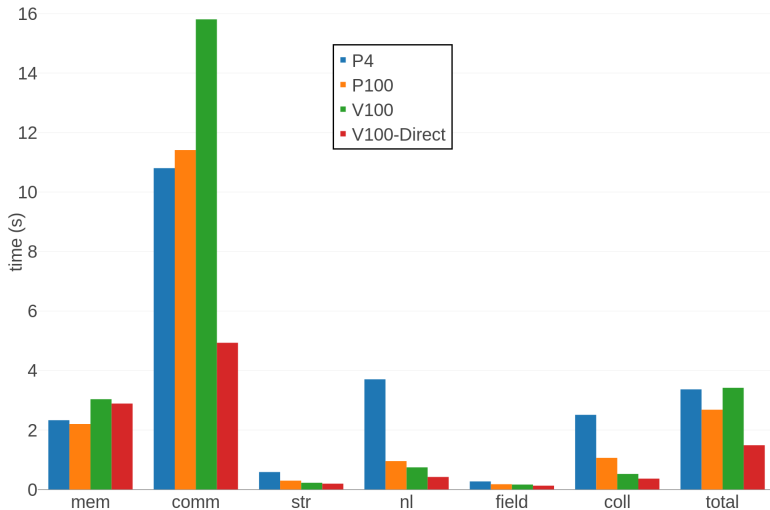
# GPU type comparison

Stampede2, GA, Piz Daint, Titan

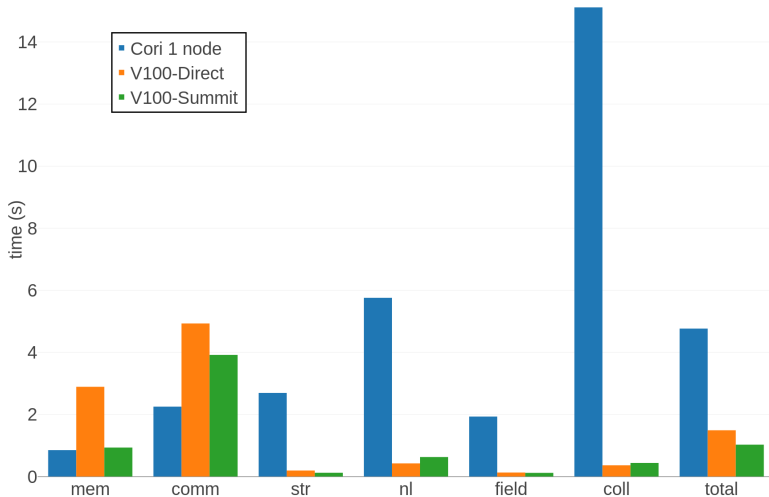


# Google Cloud Partition Comparison

Santa Fe (last week)



# Cloud V100 compared to Summit and Cori



# OUTLINE

- ① History of General Atomics?
- ② The case for fusion energy
- ③ Mathematical formulation and GPU-based numerical solution
- ④ Simulation of turbulent energy loss in a tokamak plasma
- ⑤ GPU performance: development and results

# 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.