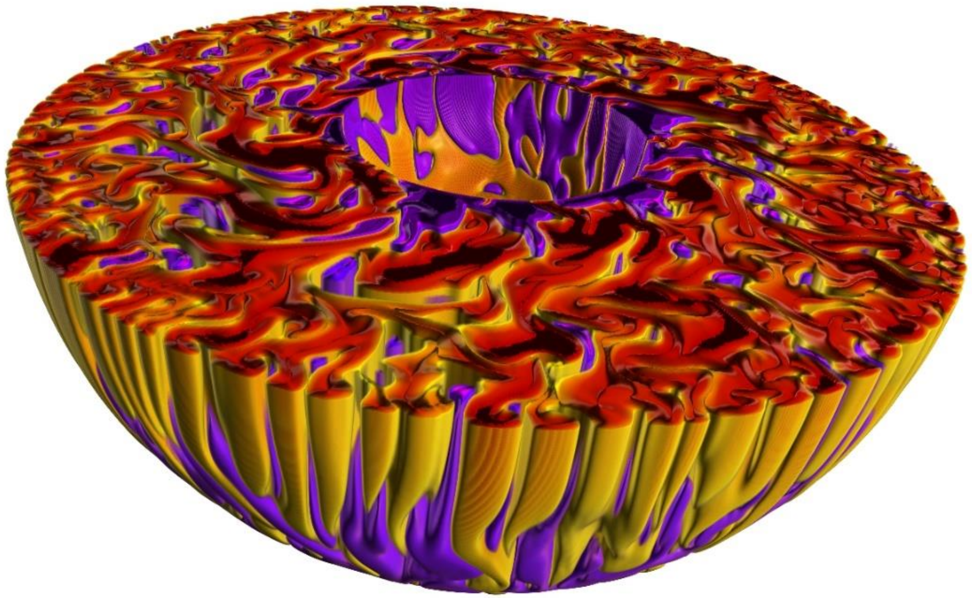
Rayleigh: Some Details on the Numerics



Nick Featherstone

Southwest Research Institute

Department of Solar and Heliospheric Physics,

Solar System Science and Exploration Division

Rayleigh Solves: The Boussinesq MHD Equations

$$\left[\frac{\partial \boldsymbol{v}}{\partial t} + \boldsymbol{v}\cdot\boldsymbol{\nabla}\boldsymbol{v} + \frac{2}{E}\hat{\boldsymbol{z}}\times\boldsymbol{v}\right] = \frac{Ra}{Pr}\left(\frac{r}{r_o}\right)^n \Theta\,\hat{\boldsymbol{r}} - \frac{1}{E}\boldsymbol{\nabla}P + \frac{1}{E\,Pm}(\boldsymbol{\nabla}\times\boldsymbol{B})\times\boldsymbol{B} + \boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{D}}$$

$$\left[\frac{\partial \Theta}{\partial t} + \boldsymbol{v}\cdot\boldsymbol{\nabla}\Theta\right] = \frac{1}{Pr}\boldsymbol{\nabla}\cdot\left[\tilde{\kappa}(r)\boldsymbol{\nabla}\Theta\right]$$

$$\frac{\partial \boldsymbol{B}}{\partial t} = \boldsymbol{\nabla}\times\left[\boldsymbol{v}\times\boldsymbol{B} - \frac{1}{Pm}\tilde{\eta}(r)\boldsymbol{\nabla}\times\boldsymbol{B}\right]$$

$$\mathcal{D}_{ij} = 2\tilde{\nu}(r)e_{ij}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{v} = 0$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{B} = 0$$

Rayleigh Solves: The Anelastic MHD Equations

$$\hat{\rho}(r)\left[\frac{\partial \boldsymbol{v}}{\partial t} + \boldsymbol{v}\cdot\boldsymbol{\nabla}\boldsymbol{v} + 2\Omega_0\hat{\boldsymbol{z}}\times\boldsymbol{v}\right] = \frac{\hat{\rho}(r)}{c_P}g(r)\Theta\,\hat{\boldsymbol{r}} + \hat{\rho}(r)\boldsymbol{\nabla}\left(\frac{P}{\hat{\rho}(r)}\right) + \frac{1}{4\pi}(\boldsymbol{\nabla}\times\boldsymbol{B})\times\boldsymbol{B} + \boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{D}}$$

$$\hat{\rho}(r)\,\hat{T}(r)\left[\frac{\partial\Theta}{\partial t} + \boldsymbol{v}\cdot\boldsymbol{\nabla}\Theta + v_r\frac{d\hat{S}}{dr}\right] = \boldsymbol{\nabla}\cdot\left[\hat{\rho}(r)\,\hat{T}(r)\,\kappa(r)\,\boldsymbol{\nabla}\Theta\right] + Q(r) + \Phi(r,\theta,\phi) + \frac{\eta(r)}{4\pi}\left[\boldsymbol{\nabla}\times\boldsymbol{B}\right]^2$$

$$\frac{\partial\boldsymbol{B}}{\partial t} = \boldsymbol{\nabla}\times\left[\,\boldsymbol{v}\times\boldsymbol{B} - \eta(r)\boldsymbol{\nabla}\times\boldsymbol{B}\,\right]$$

$$\mathcal{D}_{ij} = 2\hat{\rho}(r)\,\nu(r)\left[e_{ij} - \frac{1}{3}(\boldsymbol{\nabla}\cdot\boldsymbol{v})\delta_{ij}\right]$$

$$\Phi(r,\theta,\phi) = 2\,\hat{\rho}(r)\nu(r)\left[e_{ij}e_{ij} - \frac{1}{3}(\boldsymbol{\nabla}\cdot\boldsymbol{v})^2\right]$$

$$\boldsymbol{\nabla}\cdot\left[\hat{\rho}(r)\,\boldsymbol{v}\right] = 0$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{B} = 0.$$

Serial Numerics

Horizontal "Discretization"

Spherical Harmonics: FFTW + DGEMM (Legendre)

Radial "Discretization"

Chebyshev Polynomials: Colocation scheme

Time-Stepping: Hybrid Crank-Nicolson/Adams-Bashforth

Direct Matrix Solve: LAPack LU Decomposition routines

Recalculate matrices when Δt changes

Pseudospectral Method: Quick Overview

Spectral Configuration

$$u_x^k$$

$$\frac{\partial u_x^k}{\partial x} = iku_x^k$$

$$\left[u_x \frac{\partial u_x}{\partial x}\right]^k$$

Inverse FFT

FFT

Physical-Space Configuration

$$u_x$$

$$\frac{\partial u_x}{\partial x}$$

$$u_x \frac{\partial u_x}{\partial x}$$

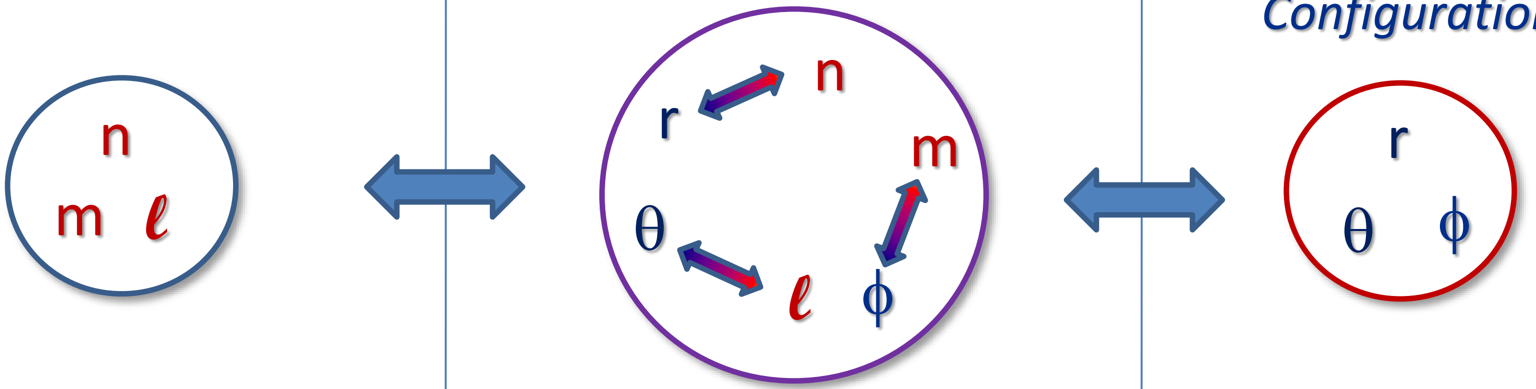
- Time-stepping: spectral configuration
- Derivatives: spectral configuration
  - Exponential convergence as ngrid grows
- Nonlinear products: physical configuration

Conceptual View of a Pseudo-Spectral Approach

Spectral Configuration

Hybrid Configuration

Physical Space Configuration



- Linear terms

- Transforms
- Derivatives

- Nonlinear Terms

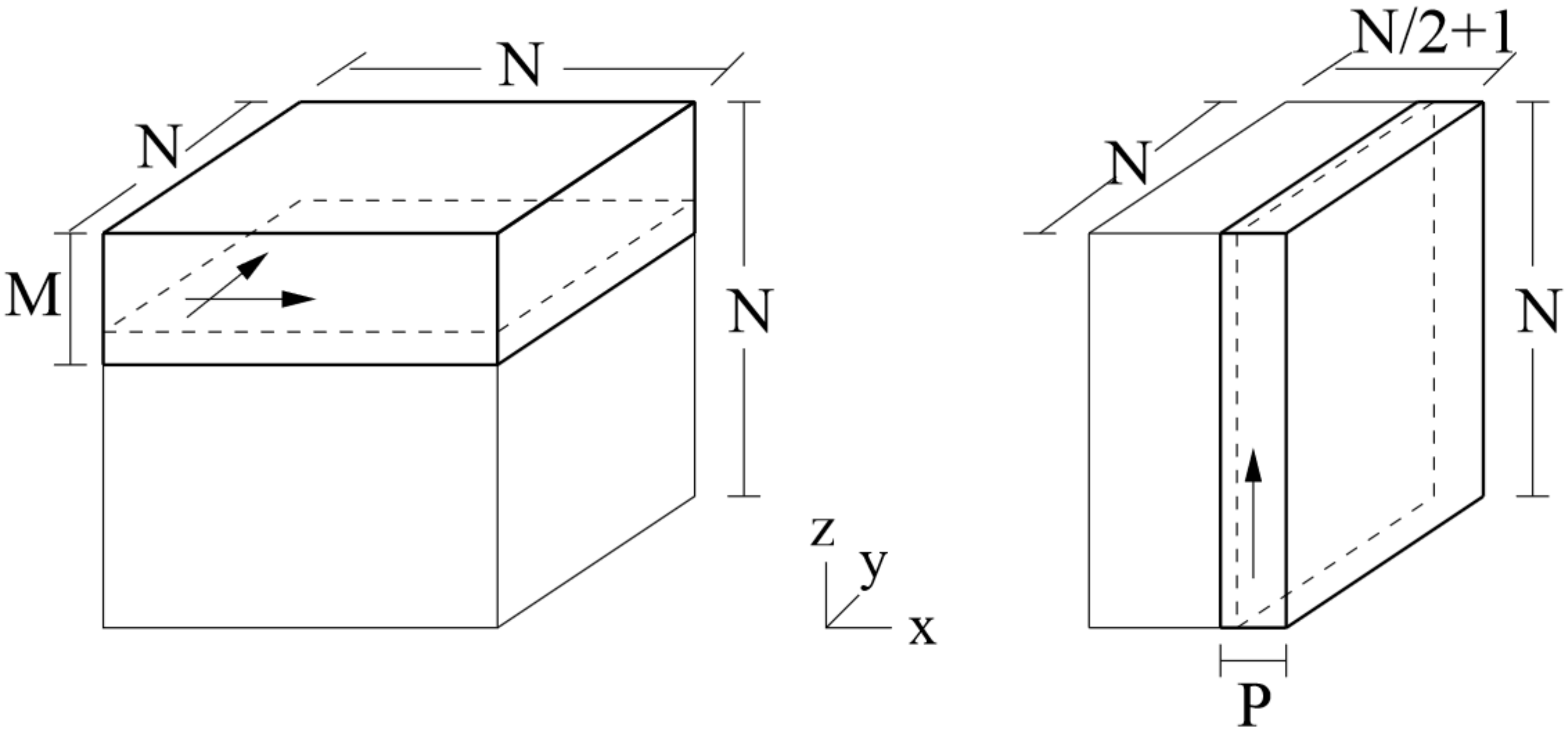
Movement between configurations requires:

Transforms: $O(N^2)$ and $O(N \log N)$ … expensive but accurate

Transposes: All-to-Alls … limit scalability

Transpose: GLOBAL rearrangement of data

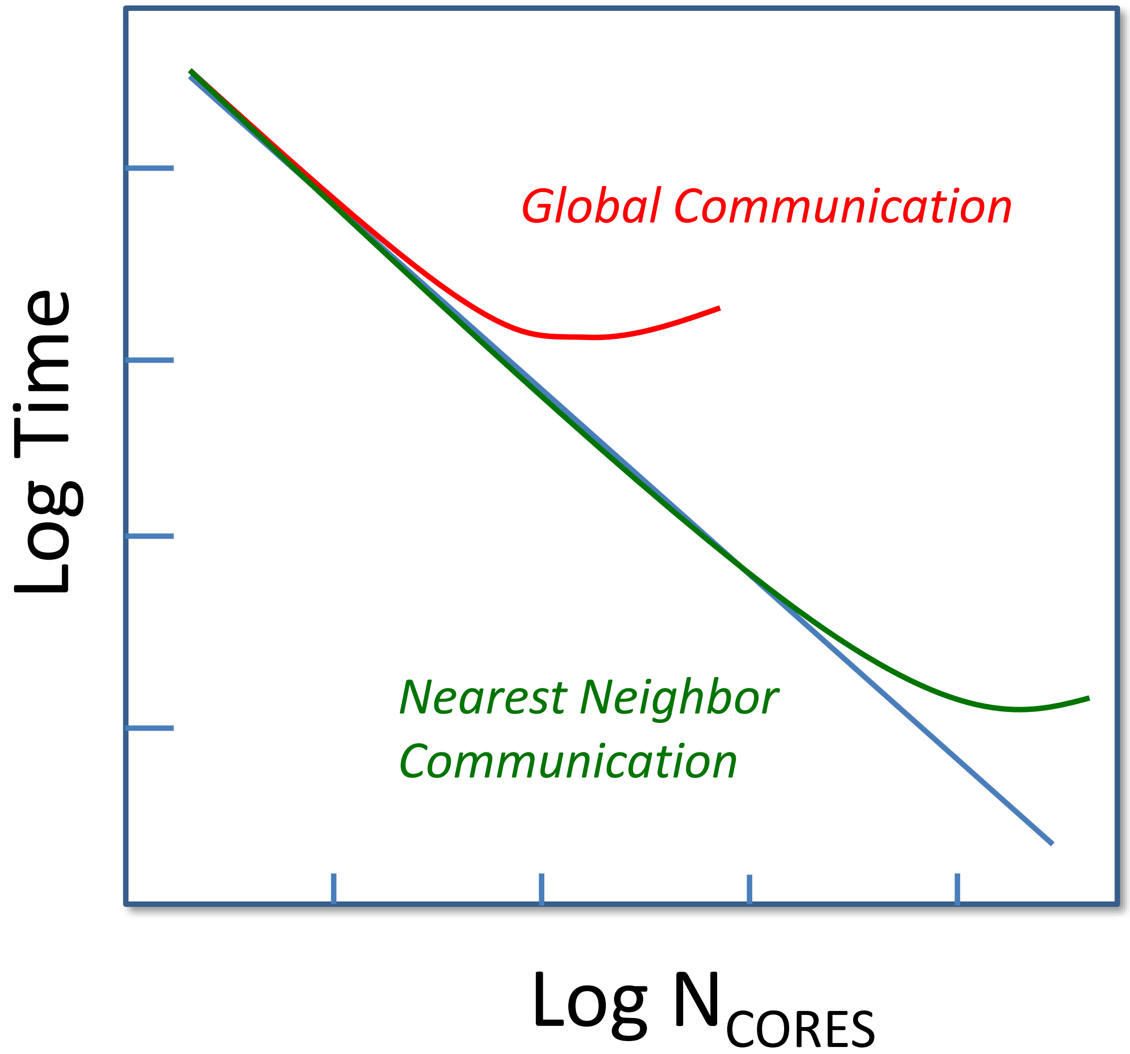
Why a Transpose?



Think of a 3-D FFT

Adapted from D.O. Go’mez et al. 2005

The Problem with Spectral Methods



Strong Scalability

Ideal Scaling:

$$\text{Time} \propto \frac{1}{N_{\text{CORES}}}$$

How Long Should an All-to-All Take?

Time = *Initiation Time* + *Transmission Time*

Local Problem Size:	P
Number of MPI Ranks:	N
Single Message Initiation Time:	I
Bandwidth:	B

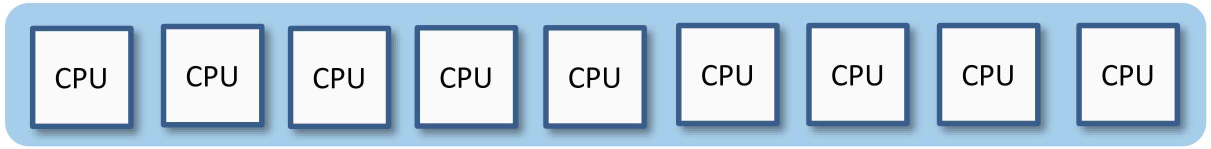
$$\textit{Transmission Time} = N \times \overbrace{(P/N)}^{\text{Message Size}} / B = P/B \quad \ldots \quad \textit{constant}$$

$$\textit{Initiation Time} = N \times I \quad \ldots \quad \textit{growing}$$

Try to limit message count

Mitigation Strategy #1: 2-D Domain Decomposition

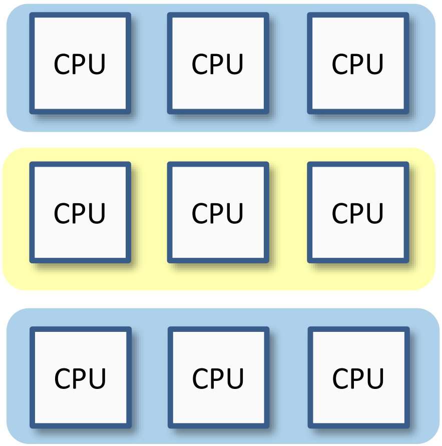
1-D Domain Decomposition



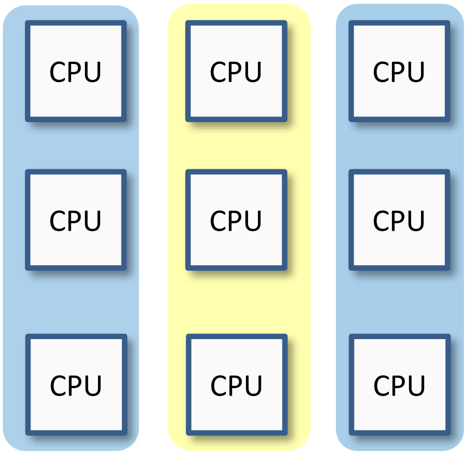
$$\text{Time} = \frac{P}{B} + NI$$

One Large All-to-All

2-D Domain decomposition



2 Passes



$$\text{Time} = \frac{2P}{B} + 2I\sqrt{N}$$

Higher Max N_{CORES}

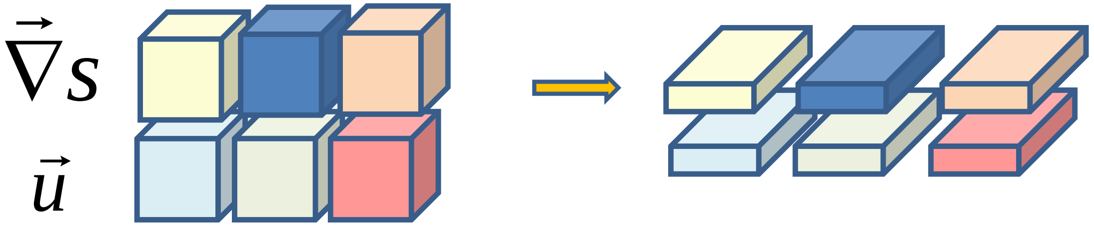
Mitigation Strategy #2: Collect the Collectives

Example: Entropy Advection

Obvious approach: *transpose each field*

Six Transposes

$$T = \frac{6P}{B} + 6NI$$



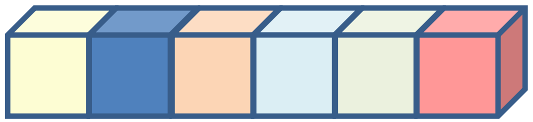
Spectral space

Physical space

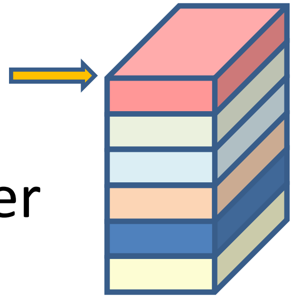
Alternative approach: *transpose a buffer*

Single Transpose

$$T = \frac{6P}{B} + \underbrace{NI}_{\textit{Lower init time...}}$$



Place each field in buffer



Transpose the buffer

Rayleigh Parallelization: Load Balancing

Triangular Truncation:

$$0 \le \ell \le \ell_{\max} \qquad 0 \le m \le \ell$$

Uniform resolution on the Sphere.

Natural Load Balancing :

Keep all ℓ's for each m in processor.

Pair high and low m modes.

Max Angular NCPUs = $\dfrac{\ell_{\max}+1}{2}$

Example
Mode Distribution

$$\ell_{\max} = 5$$

Distributing m's is awkward

Fold the triangle

m = 0	m = 1	m = 2	m = 3	m = 4	m = 5
ℓ	ℓ	ℓ	ℓ	ℓ	ℓ
0	1	2	3	4	5
1	2	3	4	5	
2	3	4	5		
3	4	5			
4	5				
5					

Rayleigh Parallelization: Load Balancing

Triangular Truncation:

$$0 \leq \ell \leq \ell_{\max} \qquad 0 \leq m \leq \ell$$

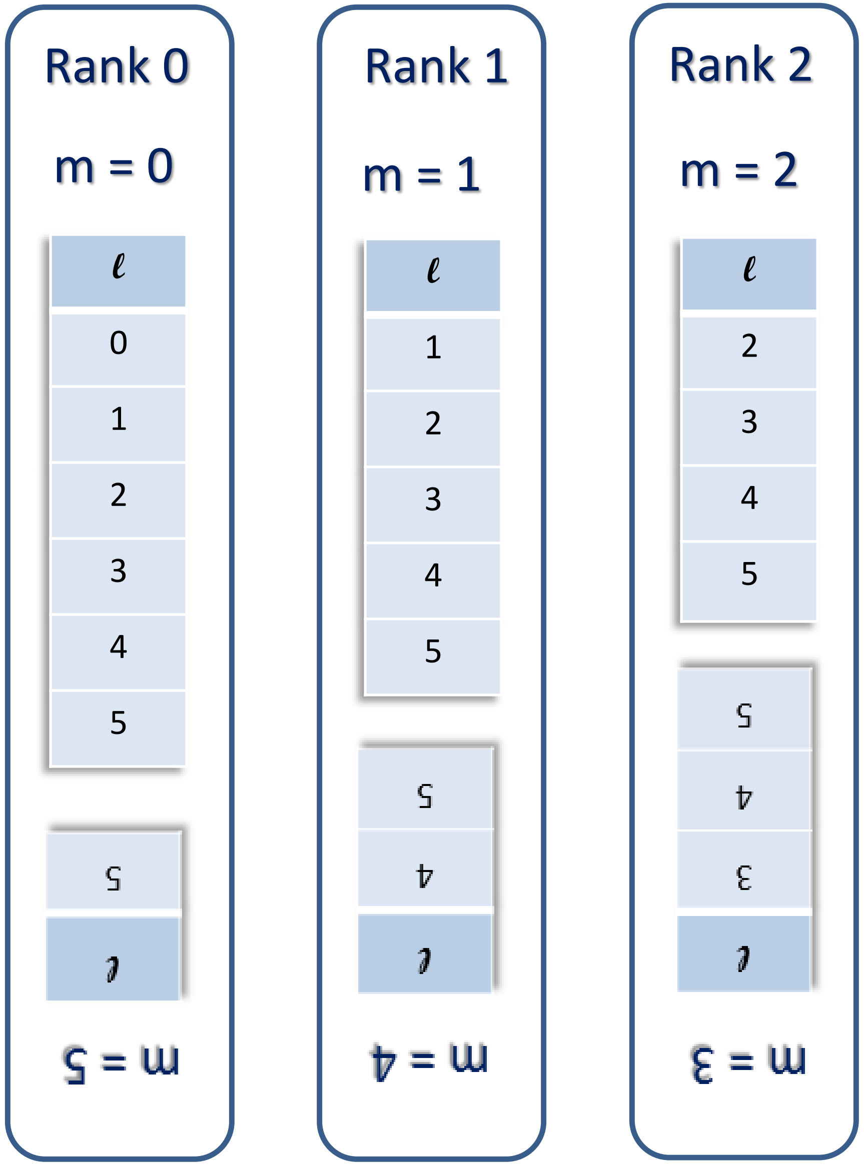
Uniform resolution on the Sphere.

Example
Mode Distribution

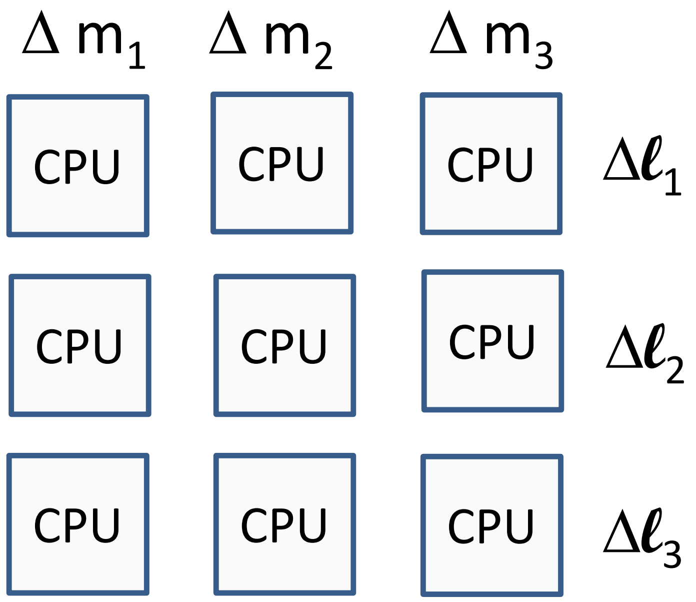
$$\ell_{\max} = 5$$

Distribute pairs of m-values

Each process gets same number of ℓ-values



Rayleigh Parallelization



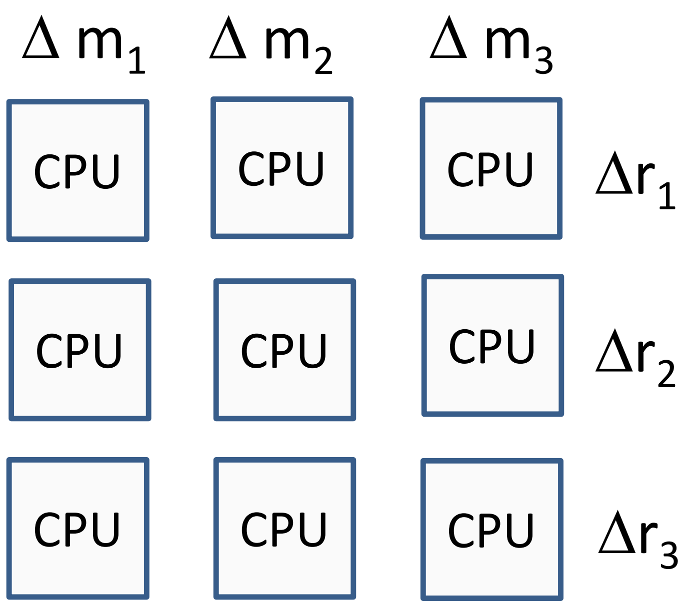
2-D Domain Decomposition

- Pure MPI
- Processes placed in columns and rows
- Three Configurations

Configuration 1

- ℓ-values distributed across rows
- m-values distributed across columns
- Radius in-processor
- Chebyshev Transforms
- Linear Solves

Rayleigh Parallelization



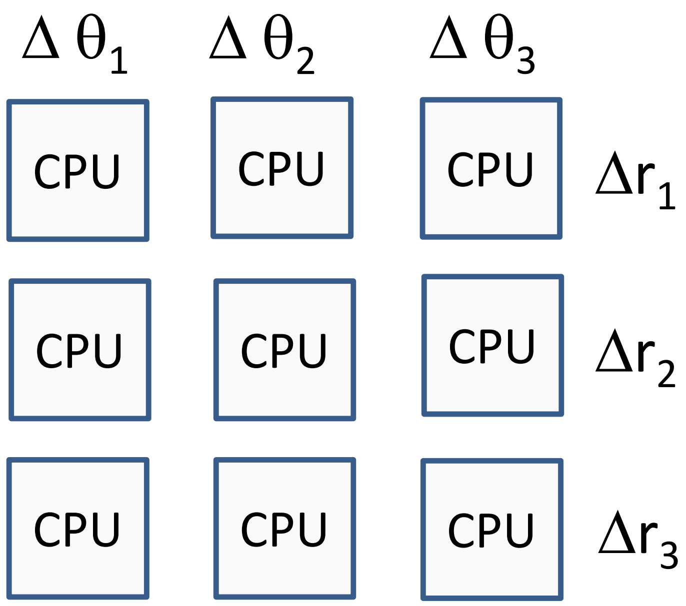
2-D Domain Decomposition

- Pure MPI
- Processes placed in columns and rows
- Three Configurations

Configuration 2

- Radial levels distributed across rows
- m-values distributed across columns
- ℓ / θ in-processor
- Legendre Transforms

Rayleigh Parallelization



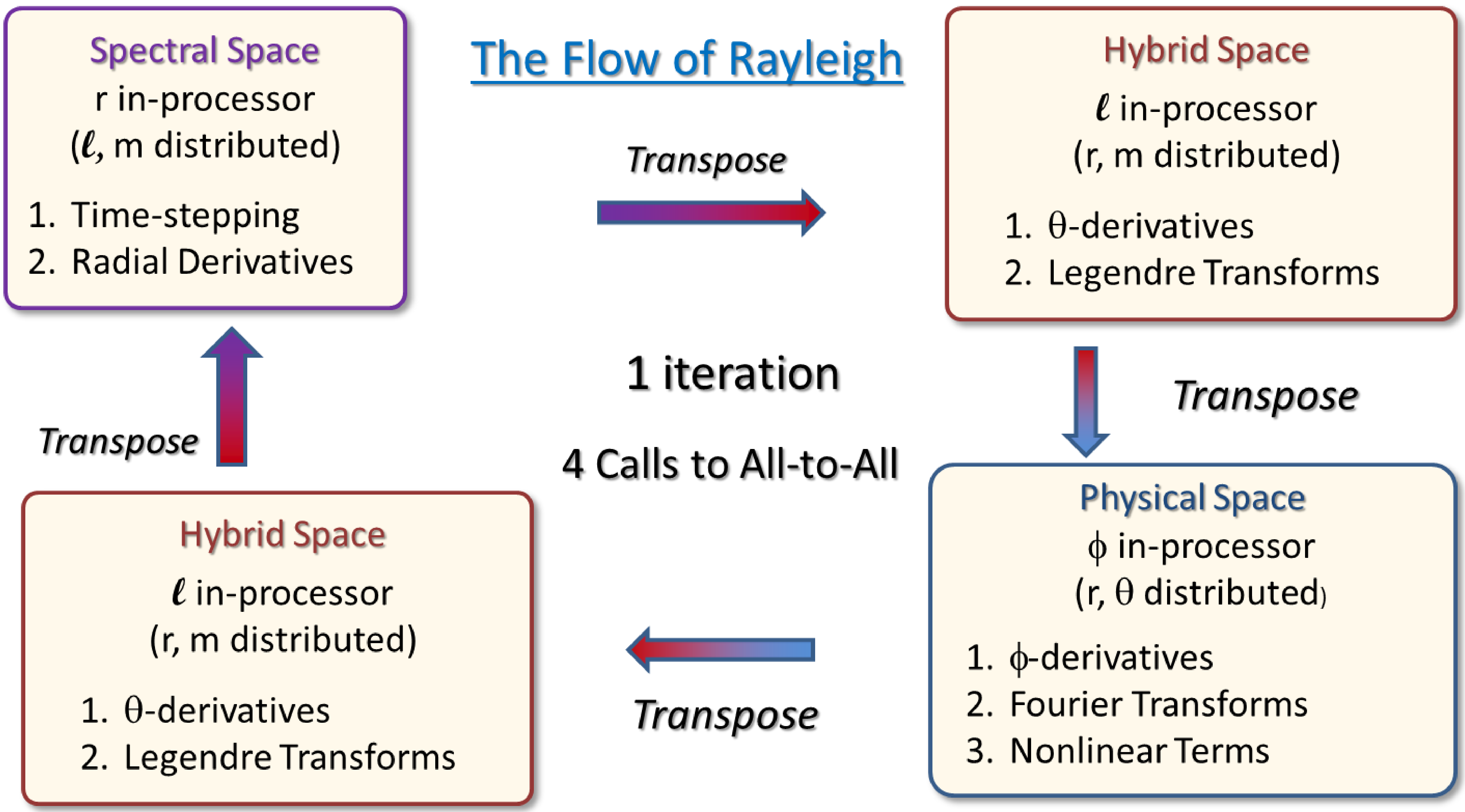
2-D Domain Decomposition

- Pure MPI
- Processes placed in columns and rows
- Three Configurations

Configuration 3

- Radial levels distributed across rows
- θ -values distributed across columns
- m/ϕ in-processor
- Fourier Transforms

The Flow of Rayleigh



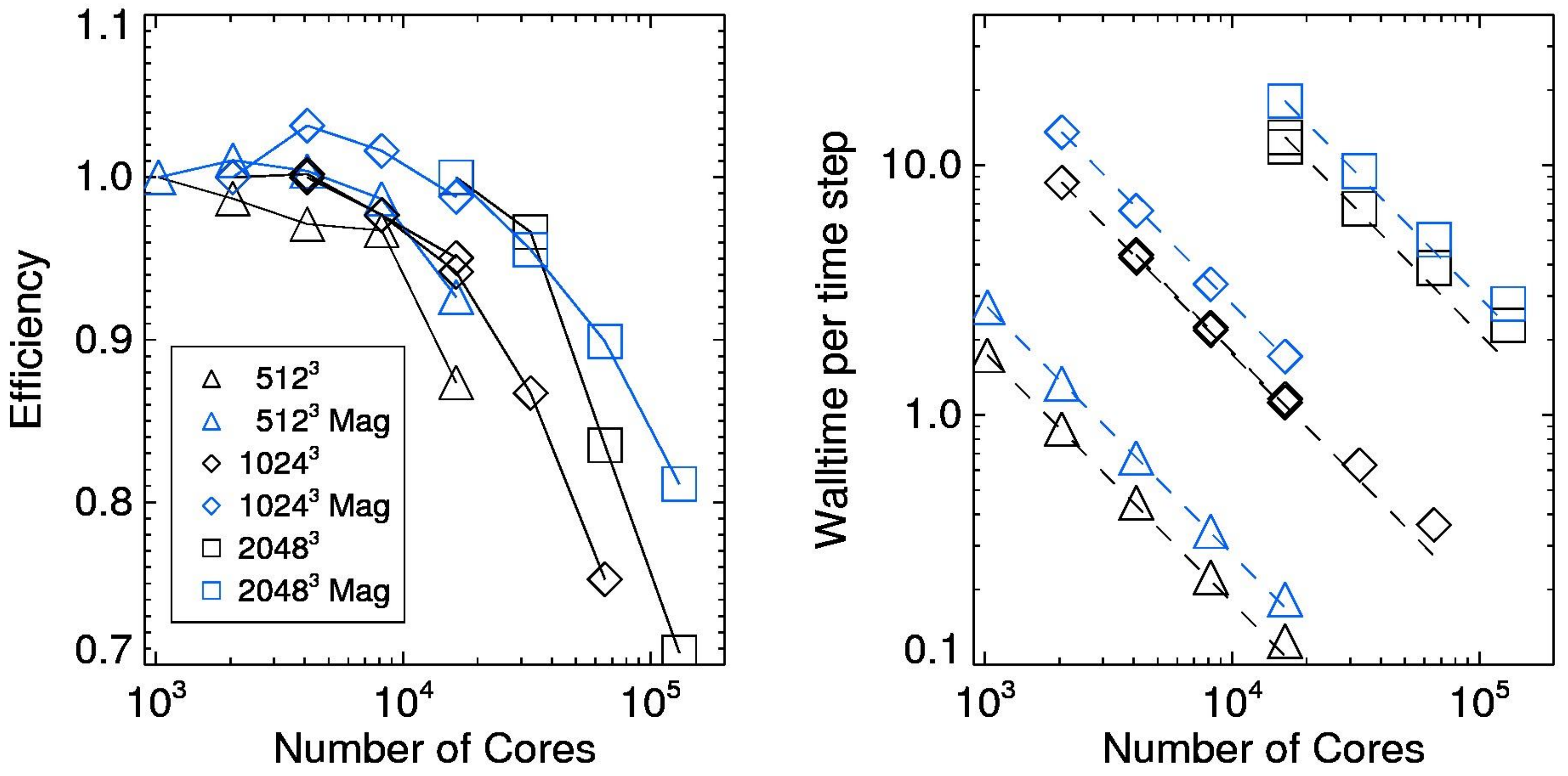
What about Memory?

Assumes $N_r = ½_{\text{Ntheta}}$

N_{theta}	Max CPUs	10% Max CPUs
256	50 kB	500 kB
512	100 kB	1 MB
1024	200 kB	2 MB

Largest Buffer holds 20 Fields
Often fits into cache…

Rayleigh Performance



Mira (IBM Blue Gene/Q Argonne)

Most Visited Wavelets Linux Mint Community Forums Blog News Member Login General Informatio...

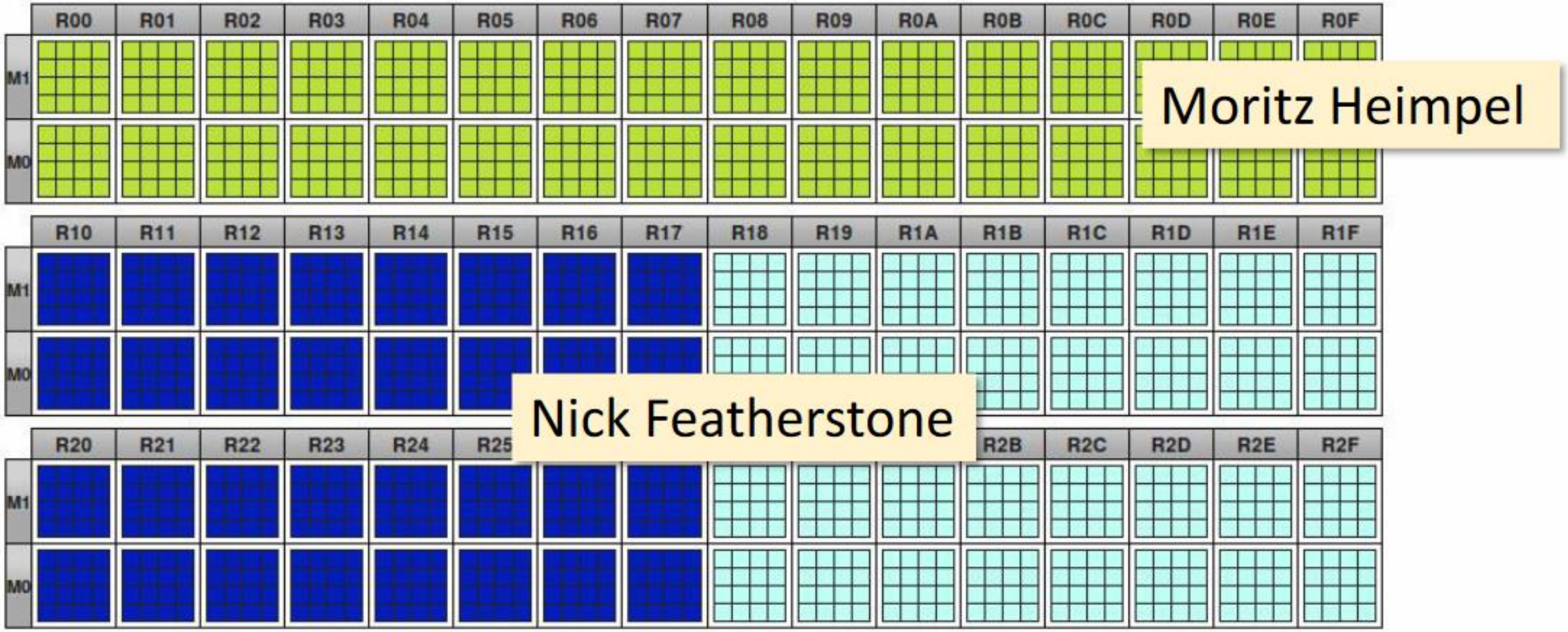
Argonne NATIONAL LABORATORY

Leadership Computing Facility

Mira Activity

786,432 cores -- Rayleigh

Home | Mira | Activity



Running Jobs | Queued Jobs | Reservations

Total Running Jobs: 3

Job Id	Project	Run Time	Walltime	Location	Queue	Nodes	Mode
900588	MagnetismHPC	16:51:55	1d 00:00:00	MIR-44000-7BFF1-16384	prod-capability	16384	script
[illegible]	MagnetismHPC	05:54:19	1d 00:00:00	MIR-04000-3BFF1-16384	prod-capability	16384	script
900636	MagnetismHPC	03:20:16	1d 00:00:00	MIR-00000-73FF1-16384	prod-capability	16384	script