Comparison among JFRS-1, Plasma Simulator and Fugaku

Y. Todo (NIFS)

IFERC Workshop on GPUs Programming

(December 16, 2020, Zoom)

Outline

- JFRS-1 at IFERC-CSC (Cray XC-50)
- Plasma Simulator "Raijin" at NIFS (NEC SX-Aurora TSUBASA)
- Fugaku at RIKEN (Fujitsu)
- Hybrid simulation code MEGA for energetic particles and MHD
- Optimization of MEGA
- MEGA Performance on PS and Fugaku
- Comparison of performance among JFRS-1, PS, and Fugaku

JFRS-1 at IFERC-CSC

Computer	Cray XC-50
Total Nodes	1370
CPU	2 x Intel Xeon Gold 6148 (2.4GHz, 20 cores/CPU)
Peak Performance	4.2 PF (3.072 TF/node x 1370)
Total Memory	256 TiB (=192GiB/node x 1370)
Interconnect network	Aries interconnect
Data storage	27 PB
Top 500 List (Nov. 2020)	119th in Top500 (61st June 2018)

Plasma Simulator “Raijin” at NIFS

PLASMA SIMULATOR

Computer	NEC SX-Aurora TSUBASA A412-8
Total Nodes (Vector Hosts)	540
Vector Engines (VEs)	4320 (8 VEs/VH x 540)
Peak Performance	10.5 PF (2.433 TF/VE x 4320)
Total Memory	202 TiB (=48 GiB/VE x 4320, HBM2)
Interconnect network	InfiniBand HDR200 x 2 800 GB/s (node-node, bi-direction)
Data storage	32.1 PB
Top 500 List (Nov. 2020)	33rd in Top500
	10th in HPCG

Vector Engine (VE) of NEC SX-Aurora TSUBASA

- VE is a vector processor
- PCI Express card
- High memory bandwidth ~1 TB/s
  (8 cores and 6 HBM2 memory modules)
- VE is now distributed also to HPC system integrators:
  Visual Technology (Japan)
  Colfax International (USA)

  https://jpn.nec.com/press/202011/20201119_01.html

Fugaku at RIKEN

Computer	Supercomputer Fugaku (Fujitsu FX1000)
Total Nodes	158,976
CPU	1 x A64FX (Armv8.2-A SVE 512bit, 2.2GHz, 48 cores/CPU)
Peak Performance	537.212 PF (3.379 TF/CPU x 158,976)
Total Memory	4.85 PiB (=32 GiB/node x 158,976, HBM2)
Interconnect network	Tofu Interconnect D
Data storage	-
Top 500 List (Nov. 2020)	1st in Top500
	1st in HPCG

https://www.r-ccs.riken.jp/en/fugaku/project/outline

MEGA code

– Hybrid code with GK energetic particles (EP) + full magnetohydrodynamics (MHD)
– EP and MHD are coupled through EP current density in MHD momentum equation (current coupling model)
– 4th order finite difference for MHD + 4th Runge-Kutta for time integration
– Parallelized with MPI + OpenMP
– 3D domain decomposition (R, ϕ, z) + particle decomposition

[Y. Todo and T. Sato, Phys. Plasmas (1998)]

Optimization of MEGA

- Particle-in-cell simulation for plasmas
  - particle location and velocity $\mathbf{x}_n$, $\mathbf{v}_n$ → ρ($\mathbf{x}$), j($\mathbf{x}$), P($\mathbf{x}$) on grids
  - electromagnetic fields on grids E($\mathbf{x}$), B($\mathbf{x}$) → fields on particles E($\mathbf{x}_n$), B($\mathbf{x}_n$)
- Memory access is a key for high performance computing: sort of particles with respect to grids (1 sort / 10 time steps)
- Vectorization for Plasma Simulator, SIMD for JFRS-1 and Fugaku

subroutine "push" to move and accelerate particles is most demanding

- this subroutine consists of the five steps
- push1: collect field arrays into one array:
  Er(i,j,k), Ez(i,j,k), Ephi(i,j,k), … → FLD(M,i,j,k) (M=1,…, 30)
- push2: find the grids where each particle is located and the interpolation coefficients
- push3: interpolate fields on particle:
  FLD(M,i,j,k) → FLP(M, n)
- push4: calculate Δx_n, Δv_n, … for each particle using FLP(M, n)
- push5: originally a part of push4, divided to avoid "if" bifurcation

subroutine “density” to distribute particles to grids Plasma Simulator (SX-Aurora TSUBASA)

```
Integer, parameter::nblkd=255
real(8)::dns(lr,lz,lphi), wkdns(nblkd,lr,lz,lphi)

dns = 0.0d0
wkdns = 0.0d0

do m = 1, npr, nblkd
  nn = nblkd
  if(npr – m + 1 .le. nblkd) nn = npr – m + 1
do mm = 1, nn
  n = m + mm – 1
  i = …
  j = …
  k = …
  wkdns(mm, i, j, k) = …
end do
end do

do k = 1, lphi
do j = 1, lz
do i = 1, lr
  do mm = 1, nblkd
    dns(i,j,k) = dns(i,j,k) + wkdns(mm,i,j,k)
  end do
end do
end do
end do
```

Performance on Plasma Simulator (SX-Aurora TSUBASA)

- grid points (256, 256, 64), 1024 particles/grid
- 256 nodes (2048 VE, 16384 cores) with flat MPI parallelization

Subroutine	Elapse time [%]	GFLOPS / core	Vector Operation Ratio [%]	Average Vector Length [%]	Performance / Peak [%]
Total	100.0	19.882	99.62	251.6	**6.5**
push	48.9	35.441	99.77	255.7	11.6
density	21.0	7.849	99.89	254.2	2.6
t_integration	11.8	3.066	99.26	256.0	1.0

Performance on Fugaku

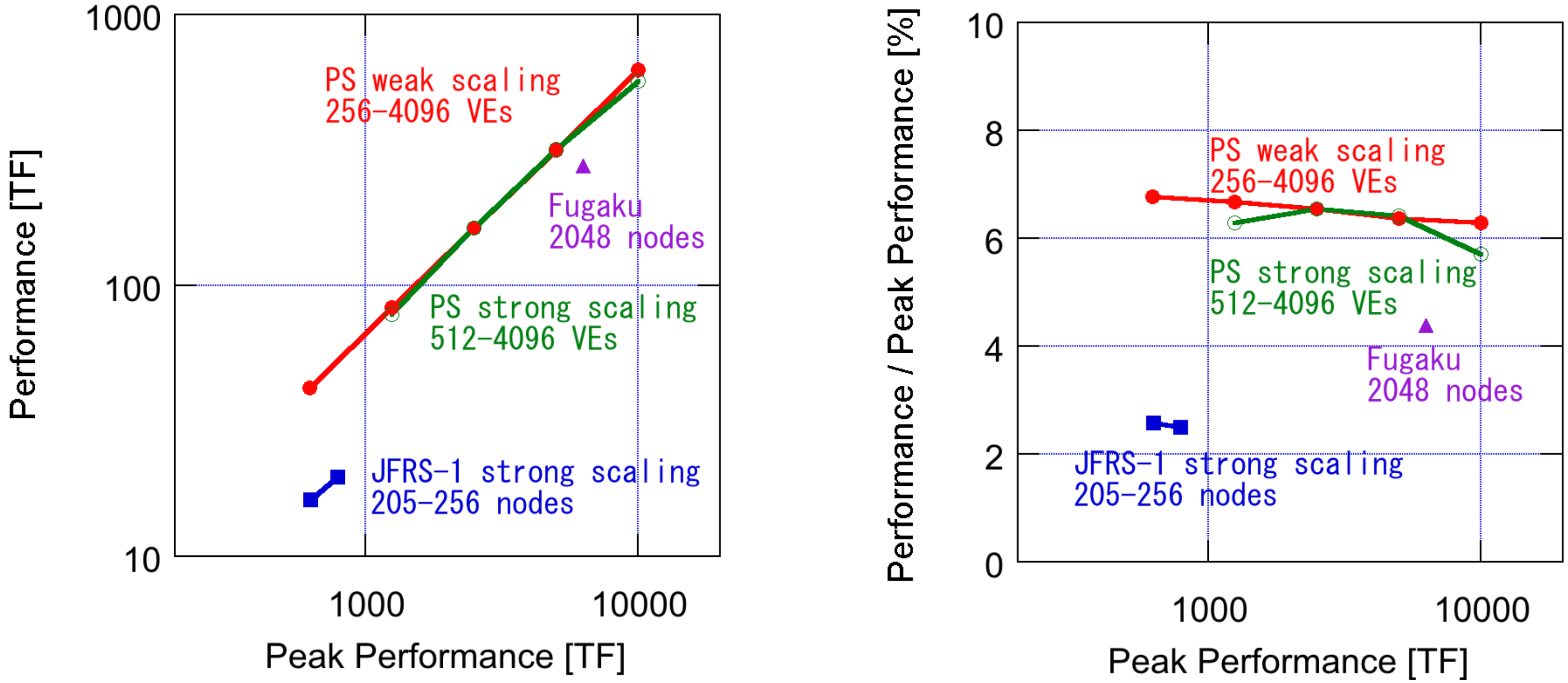
- grid points (256, 256, 64), 1024 particles/grid
- 2048 nodes (98304 cores) with 8192 MPI + 12 OpenMP threads

Subroutine	Elapse time [%]	GFLOPS / core	Scalable Vector Extension* Operation Ratio [%]	Performance / Peak [%]
Total	100.0	2.8086	92.34	**4.4**
push	75.7	3.3630	98.52	5.3
density	13.4	1.8125	27.32	2.8
t_integration	9.2	0.5906	100.00	0.93

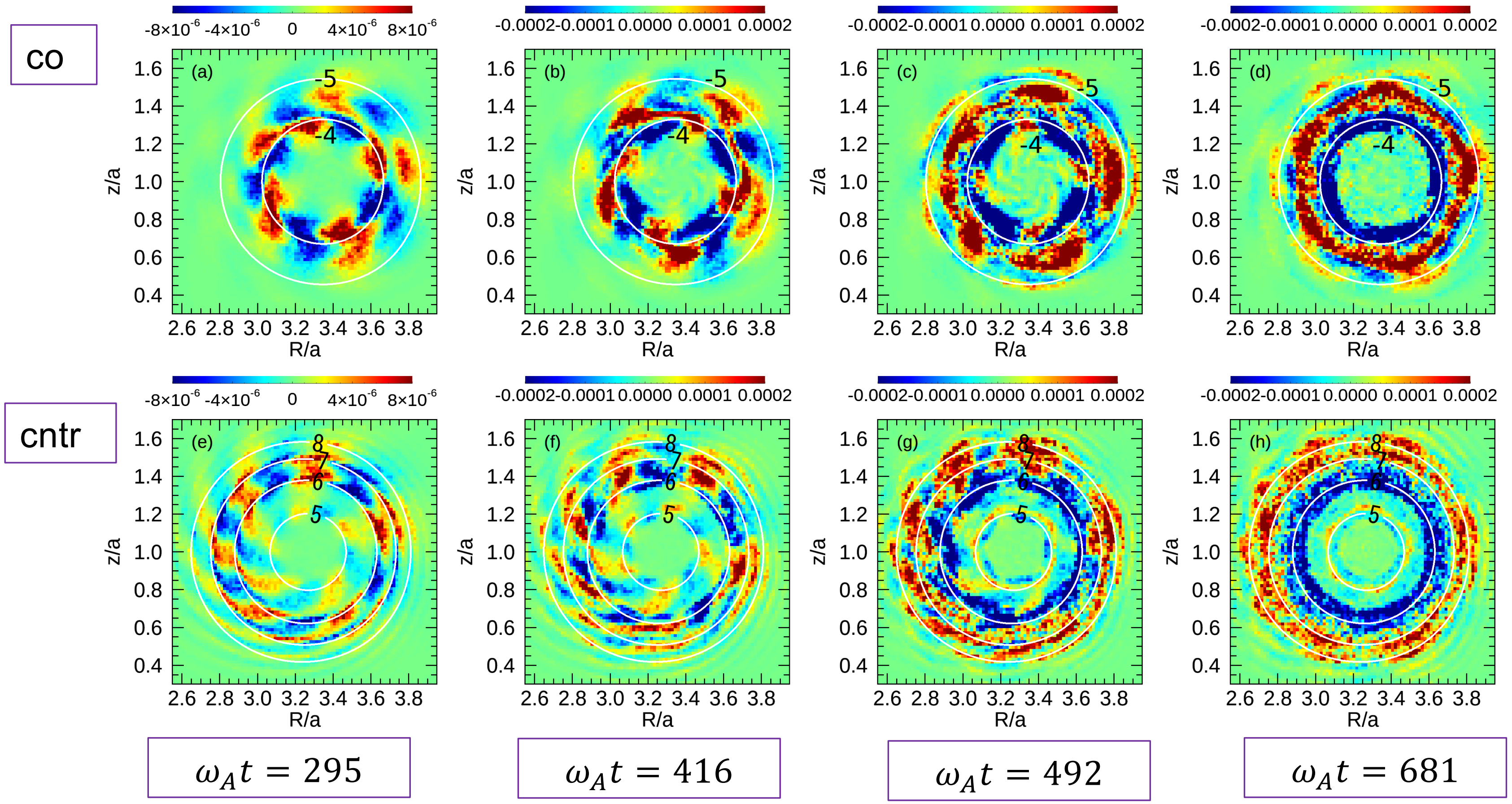
* https://www.fujitsu.com/global/Images/armv8-a-scalable-vector-extension-for-post-k.pdf

Scaling on JFRS-1, Plasma Simulator, and Fugaku

- grid points (256, 256, 64), 1024 particles/grid
- For weak scaling on Plasma simulator, the problem size is scaled starting from 1024 VEs.



Time evolution of distribution function fluctuation $\delta f(R, z)$ along $E' = const.$



δf on an (R, z) plane along

$$E' = E - \left(\frac{\omega}{n}\right) P_\varphi = const.$$

Resonance condition:

$$\omega - n\omega_\varphi - L\omega_\vartheta = 0$$

The poloidal mode number of df agrees with L=> resonance

The distribution function is flattened by the particle trapping

Y. Todo et al., submitted to PPCF

Summary

- Performance of MEGA is compared among JFRS-1, Plasma Simulator, and Fugaku.
- Good performance of MEGA is demonstrated for both strong and weak scaling on Plasma Simulator.
- MEGA achieved 6.3 – 6.7% of the peak performance for 256 – 4096 VEs on Plasma Simulator.
- Performance of MEGA on JFRS-1 and Fugaku is 2.5% and 4.4%, respectively.