HPC Challenge Awards 2010 Class2 XcalableMP Submission

Jinpil Lee, Masahiro Nakao, Mitsuhisa Sato
University of Tsukuba

Submission Overview

- XcalableMP
  - Language and model, proposed by XMP spec WG
  - Fortran + Language extension
  - OpenMP-like directives, co-array syntax
  - C version compiler is used for the submission
- Submission
  - 3 HPCC benchmarks: Linpack, FFT, RandomAccess
  - 2 other kernals: Laplace, NPB-CG
- Omni XcalableMP Compiler 0.5 (alpha) is released
  - download at www.xcalablemp.org

Evaluation Platform

T2K OpenSuperComputer Tsukuba System (using 1 ~ 64 nodes)		
CPU	AMD Opteron Quad-core 8800 2.3Ghz	x 4 sockets (16 cores)
MEMORY	32GB	
NETWORK	Infiniband DDR (4 rails)	
OS	Linux kernel 2.6.18 x86 64	
MPI	MVAPICH2 1.2	

Cray XT5 (using 1 ~ 16 nodes)		
CPU	AMD Opteron Shanghai 2.7Ghx	x 2 sockets (8 codes)
MEMORY	32GB	
NETWORK	SeaStar 3D-Torus interconnect	
OS	CLE 3.1 pre-release (based on Linux 2.6.27.48)	
MPI	Cray MPT(Message Passing Toolkit) 5.1 (based on MPICH2)	

Parallelizing Laplace

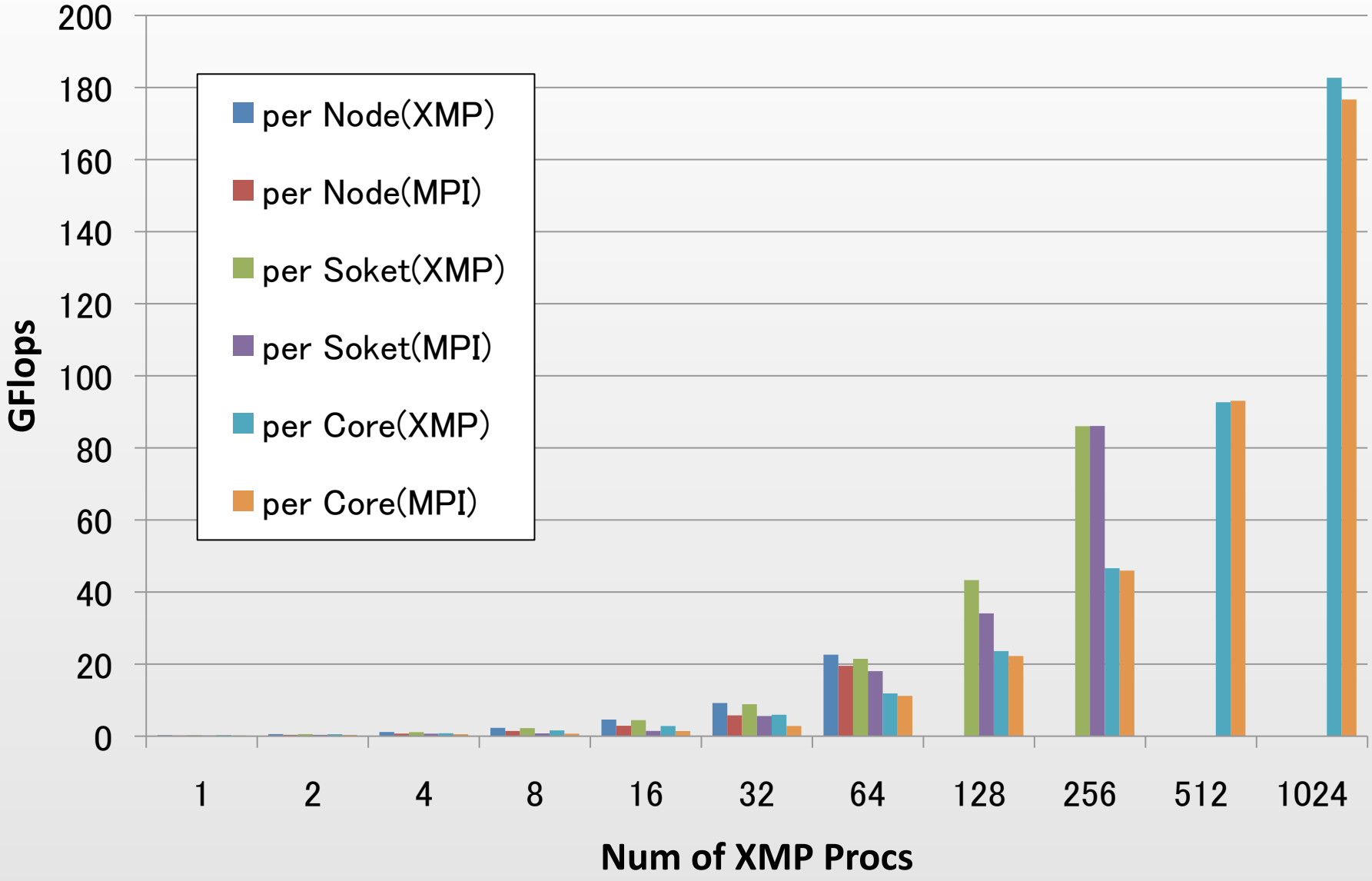
```
double u[XSIZE][YSIZE], uu[XSIZE][YSIZE] ;
#pragma xmp nodes p(NPCOL, NPROW)
#pragma xmp template t(0:XSIZE−1, 0:YSIZE−1)
#pragma xmp distribute t(block, block) onto p
#pragma xmp align u[j][i] with t(i, j)
#pragma xmp align uu[j][i] with t(i, j)
#pragma xmp shadow uu[1:1][1:1]
. . .
for (k = 0 ; k < niter; k++) {
#pragma xmp reflect uu
#pragma xmp loop (x, y) on t(y, x)
  for (x = 1 ; x < XSIZE−1; x++)
    for (y = 1 ; y < YSIZE−1; y++)
      u [x][y] = ( uu [x−1][y] + uu [x +1][y]
                 + uu [x][y−1] + uu [x][y + 1]) / 4.0;
```

Data Distribution

Data Sync

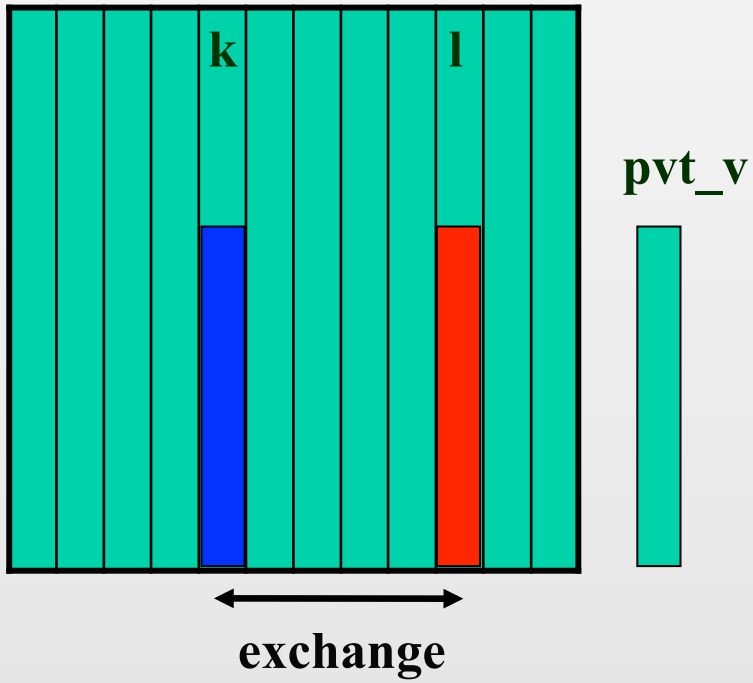
Work Sharing

Performance of Laplace (T2K-Tsukuba)



Parallelizing Linpack

- **1-dimensional Cyclic Distribution**
- **Selecting Pivot and Exchange**
  - needs internode communication
  - vector exchange using gmove

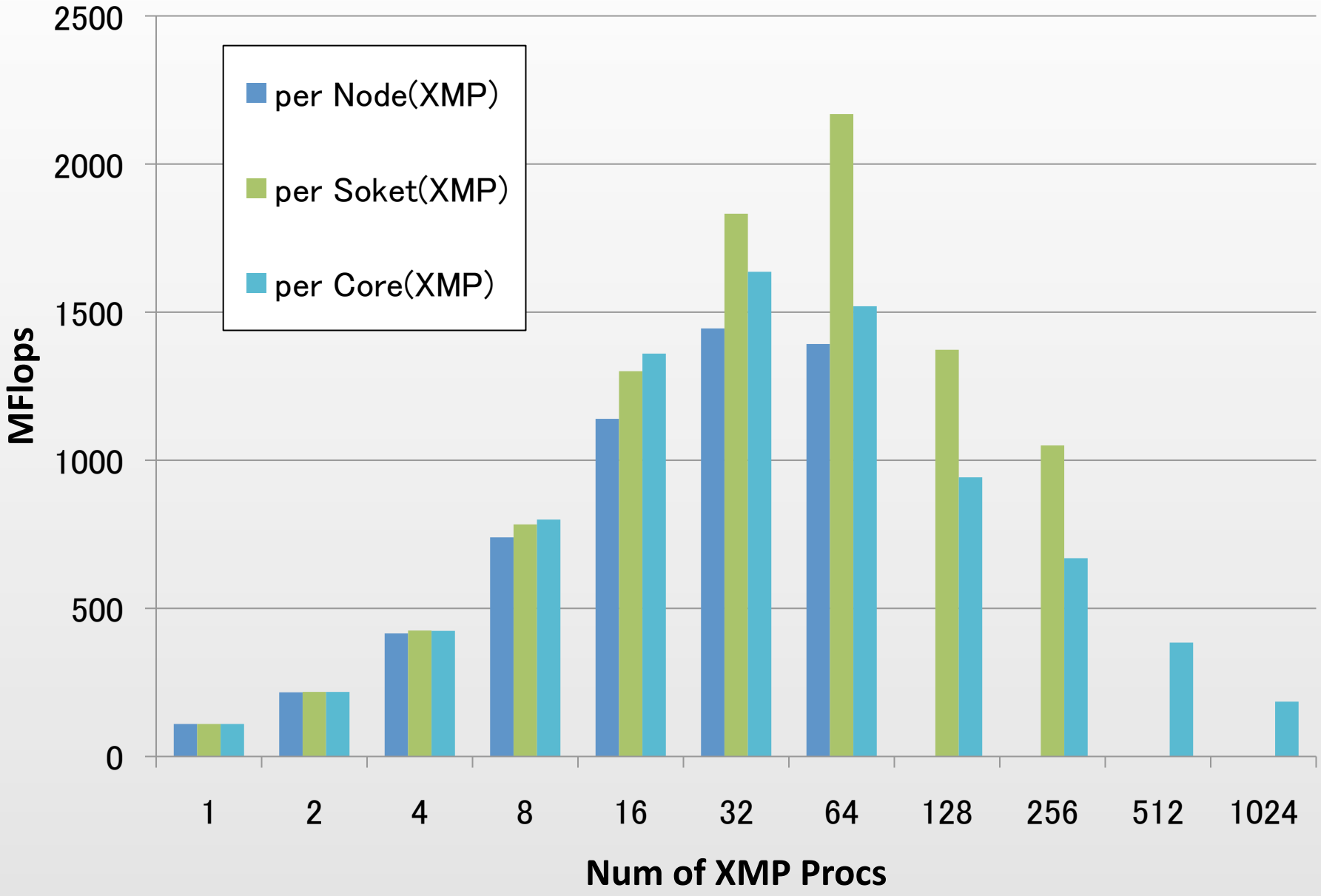


dgefa function:

```
#pragma xmp gmove
   pvt_v[k:n-1] = a[k:n-1][l];

   if (l != k) {
#pragma xmp gmove
     a[k:n-1][l] = a[k:n-1][k];
#pragma xmp gmove
     a[k:n-1][k] = pvt_v[k:n-1];
   }
```

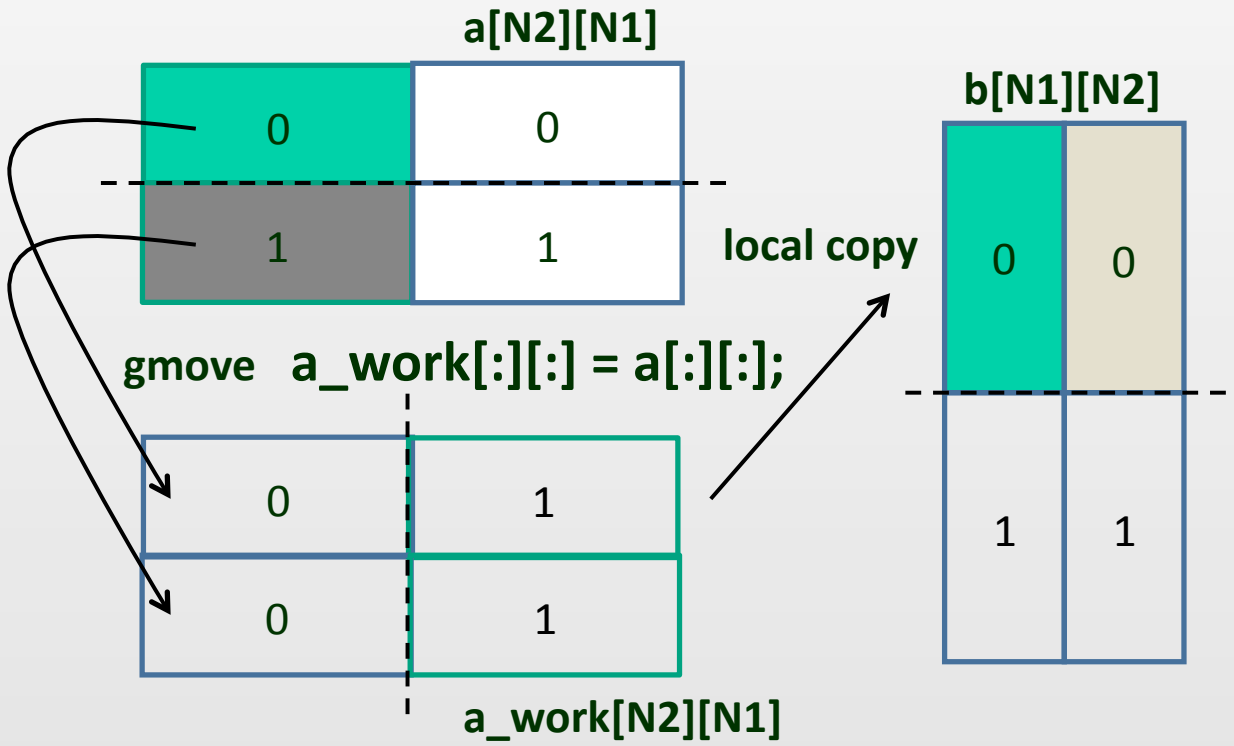
Performance of Linpack (T2K-Tsukuba)



Parallelizing FFT

- **six-step FFT algorithm**
  - needs three matrix transpose
- **1-dimensional block distribution**
  - transpose needs internode comm
  - implemented by gmove

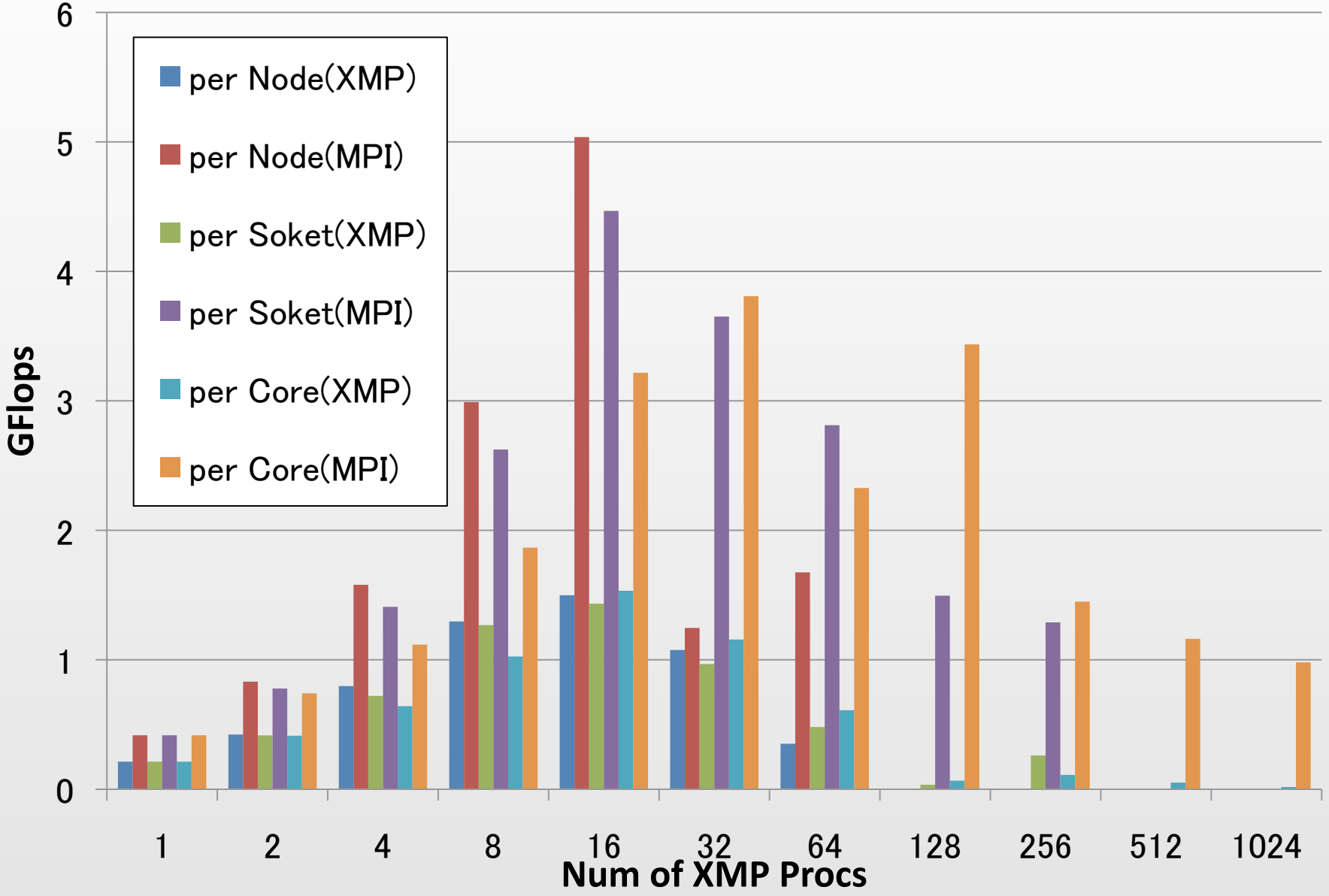
```
for (i = 0; i < N1; i++)
  for (j = 0; j < N2; j++)
    b[i][j] = a[j][i];
```



Source Code of FFT

```
#pragma xmp align a_work[*][i] with t1(i)
#pragma xmp align a[i][*] with t2(i)
#pragma xmp align b[i][*] with t1(i)
. . .
#pragma xmp gmove
  a_work[:][:] = a[:][:];                    // collect elmts
#pragma xmp loop on t1(i)
  for(i = 0; i < N1; i++)
    for(j = 0; j < N2; j++)
      c_assgn(b[i][j], a_work[j][i]);        // local copy
#pragma xmp loop on t1(i)
  for(i = 0; i < N1; i++)
    HPCC_fft235(b[i], work, w2, N2, ip2);    // 1-dim FFT
```

Performance of FFT (T2K-Tsukuba)



Parallelizing NPB-CG

```
#pragma xmp template t(0:N-1,0:N-1)
#pragma xmp distribute t(block, block) on p
#pragma xmp align a[j][i] to t(i,j)
#pragma xmp align p[i] to t(i,*)
#pragma xmp align w[j] to t(*,j)
….
for(){
…
#pragma xmp loop on t(*, j)
for (j = 1; j <= lastrow-firstrow+1; j++) {
    sum = 0.0;
    for (k = rowstr[j]; k < rowstr[j+1]; k++) {
          sum = sum + a[k]*p[colidx[k]];
    }
    w[j] = sum;
}
#pragma xmp reduction(+:w) on p(*, :)
#pragma xmp gmove
q[:] = w[:];
….
….. Update p with q ….
}
```

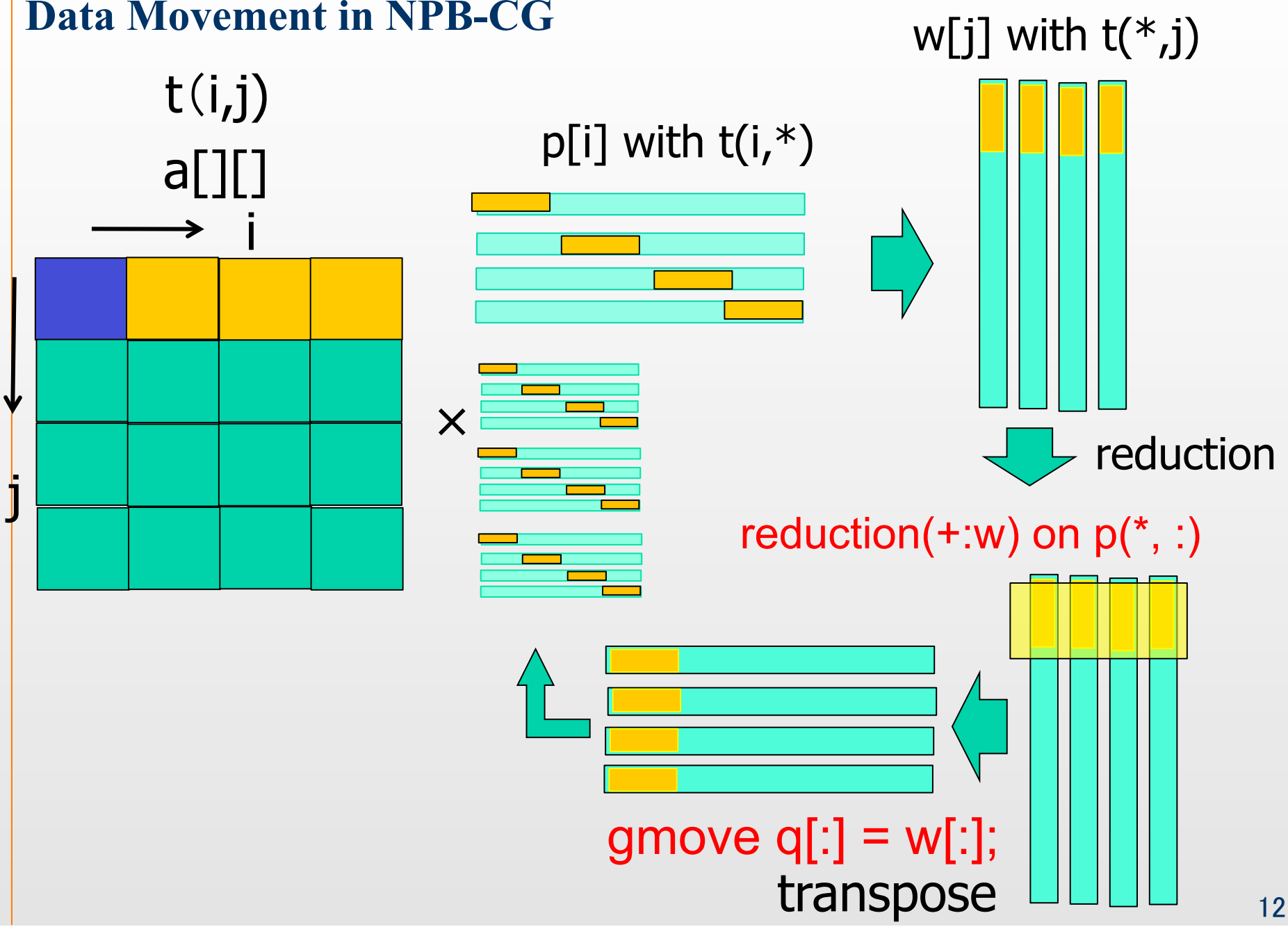
p[], q[], and w[] are distributed arrays.

p[i], q[i] with t(i, *)
w[i] with t(*, i)

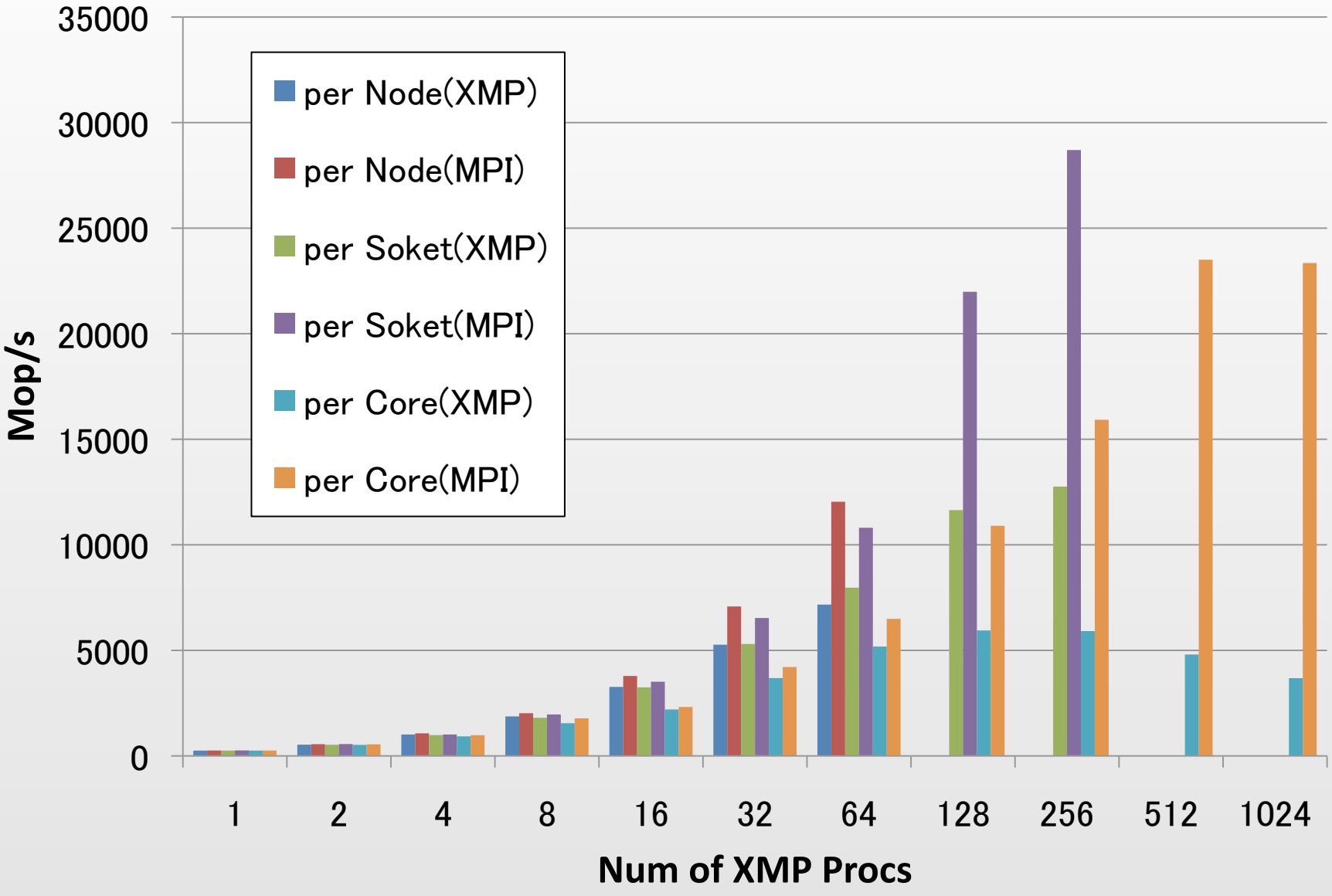
Reduction operation on replicated array

copy arrays with different distributions

Data Movement in NPB-CG



Performance of NPB-CG (T2K-Tsukuba)

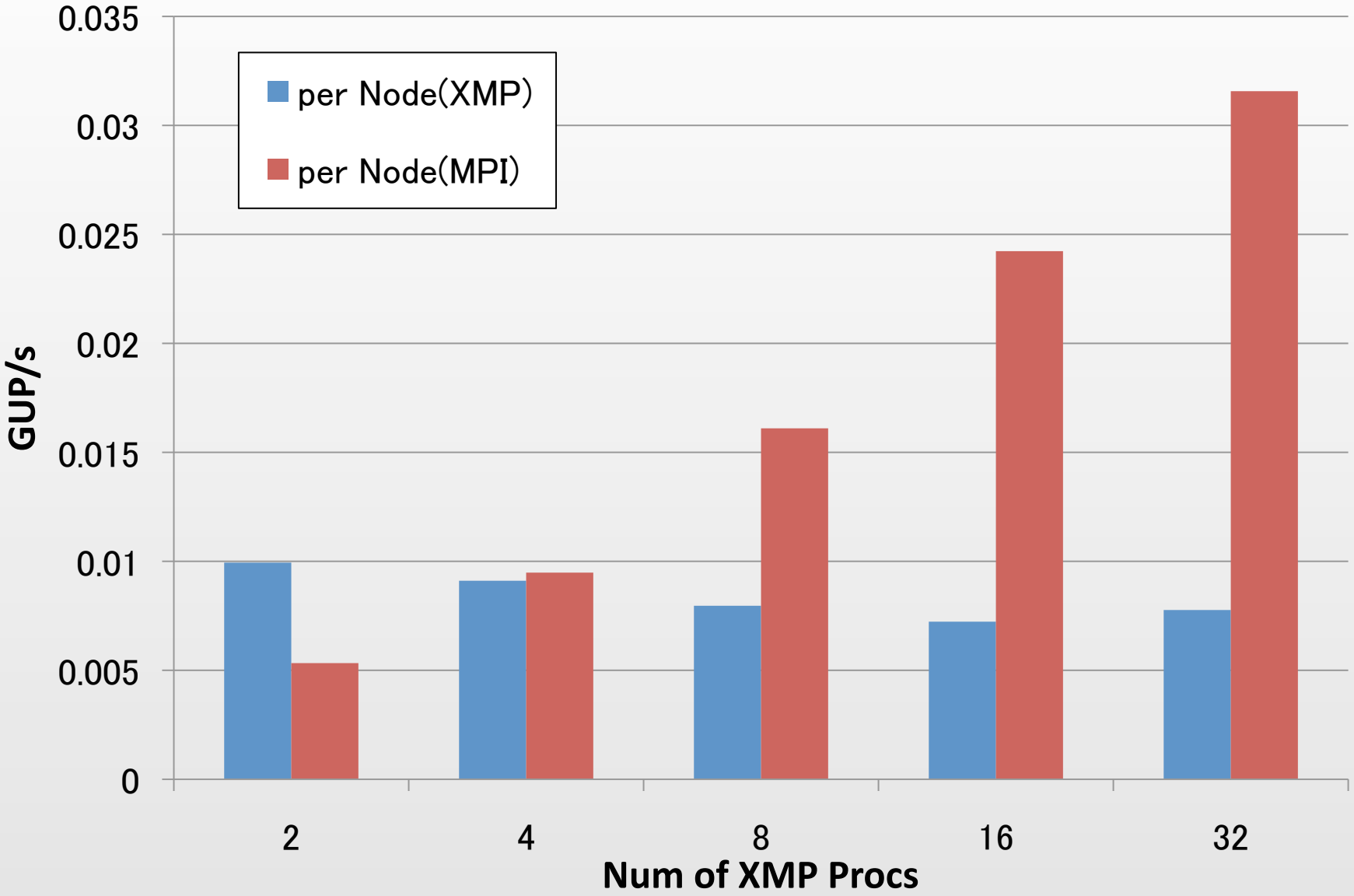


Parallelizing RandomAccess

- Local View Programming with Co-array

```
#define SIZE TABLE_SIZE/PROCS
u64Int Table[SIZE] ;
#pragma xmp nodes p(PROCS)
#pragma xmp coarray Table [PROCS]
. . .
for (i = 0; i < SIZE; i++) Table[i] = b + i ;
. . .
for (i = 0; i < NUPDATE; i++) {
  temp = (temp << 1) ^ ((s64Int)temp < 0 ? POLY : 0);
  Table[temp%SIZE]:[(temp%TABLE_SIZE)/SIZE] ^= temp;
}
#pragma xmp barrier
```

Performance of RandomAccess (T2K-Tsukuba)



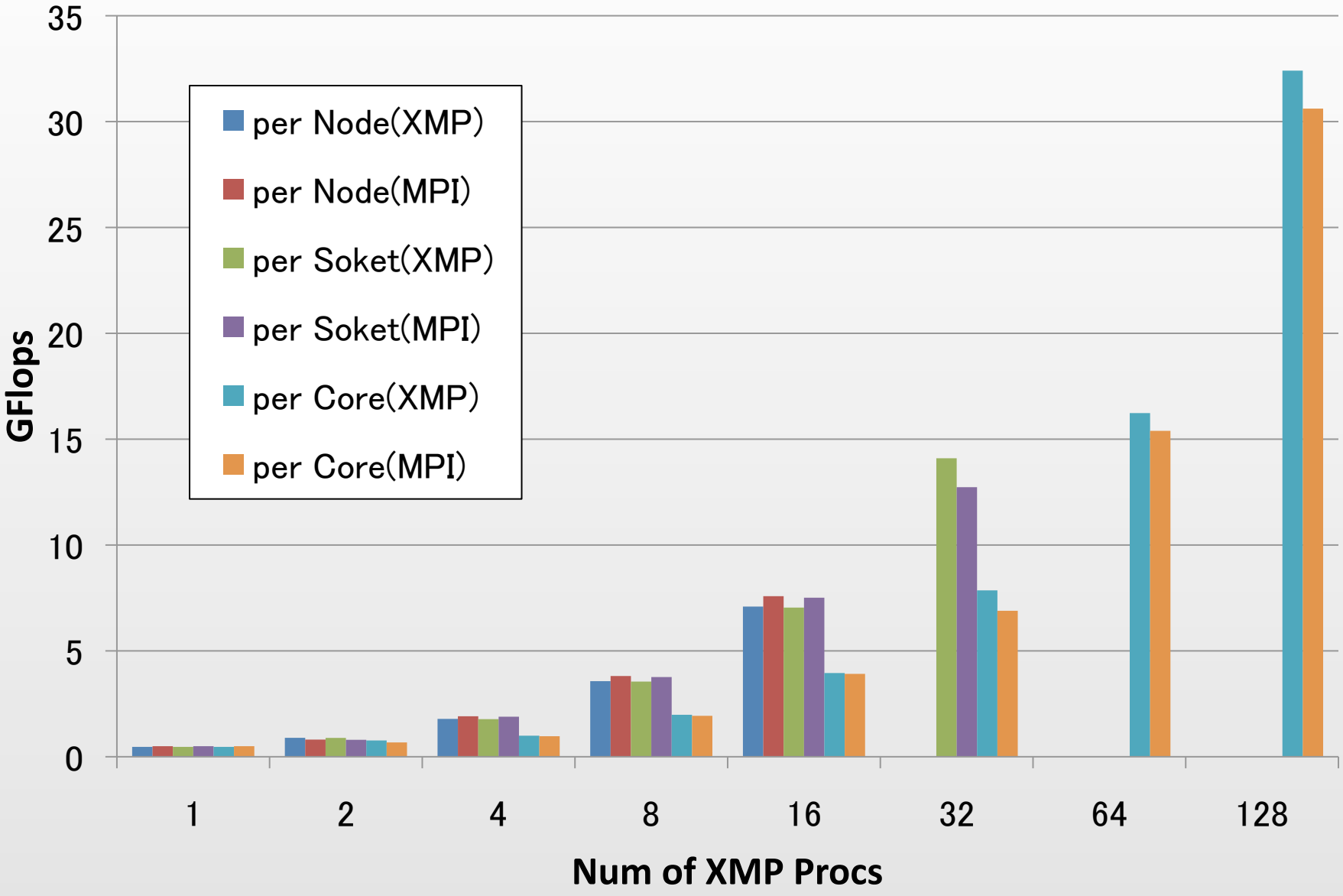
Thanks for your attention

for more information. . .

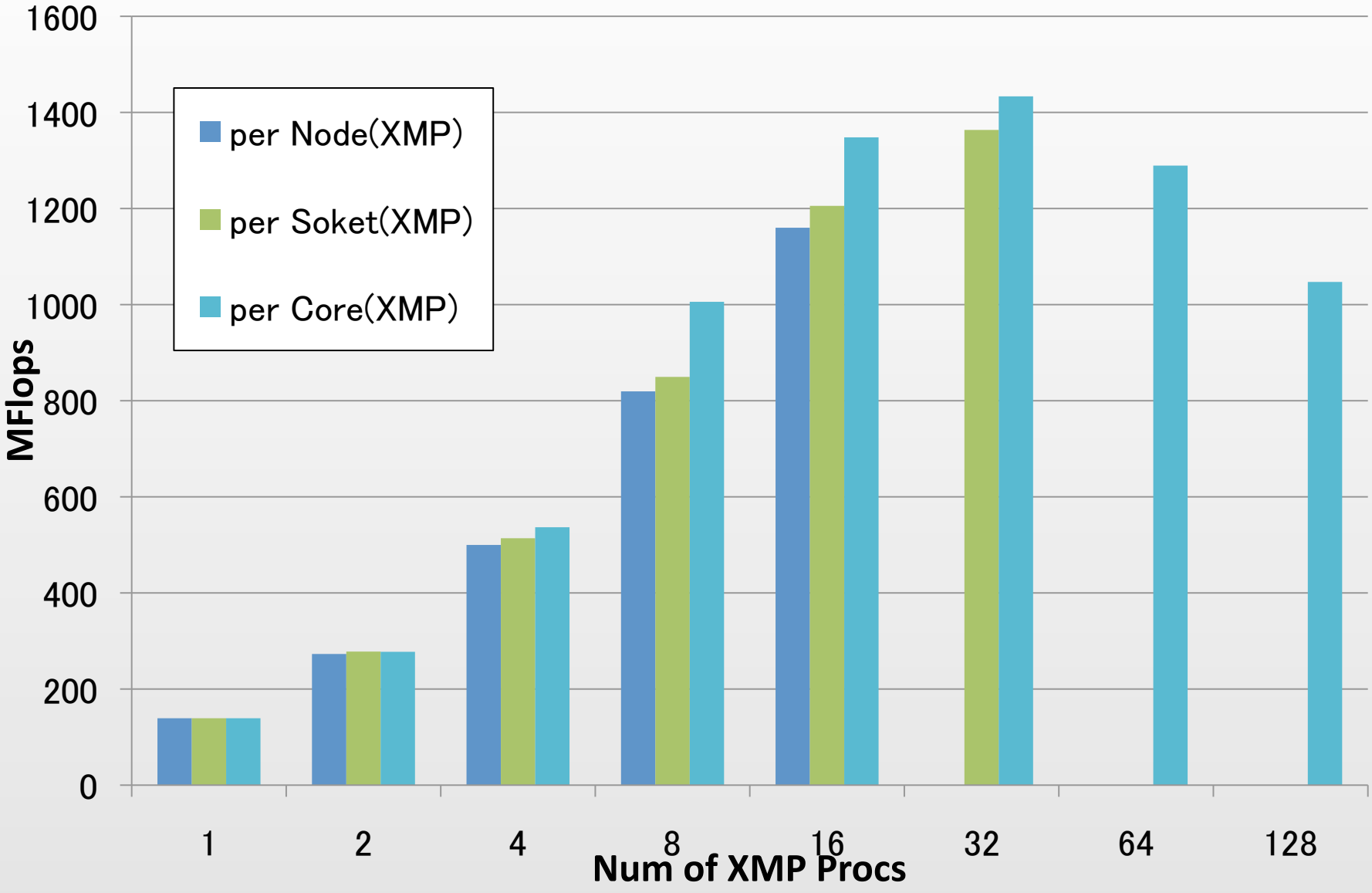
visit #3213 Center for Computational Sciences, University of Tsukuba

#1321 T2K Open Supercomputer Alliance

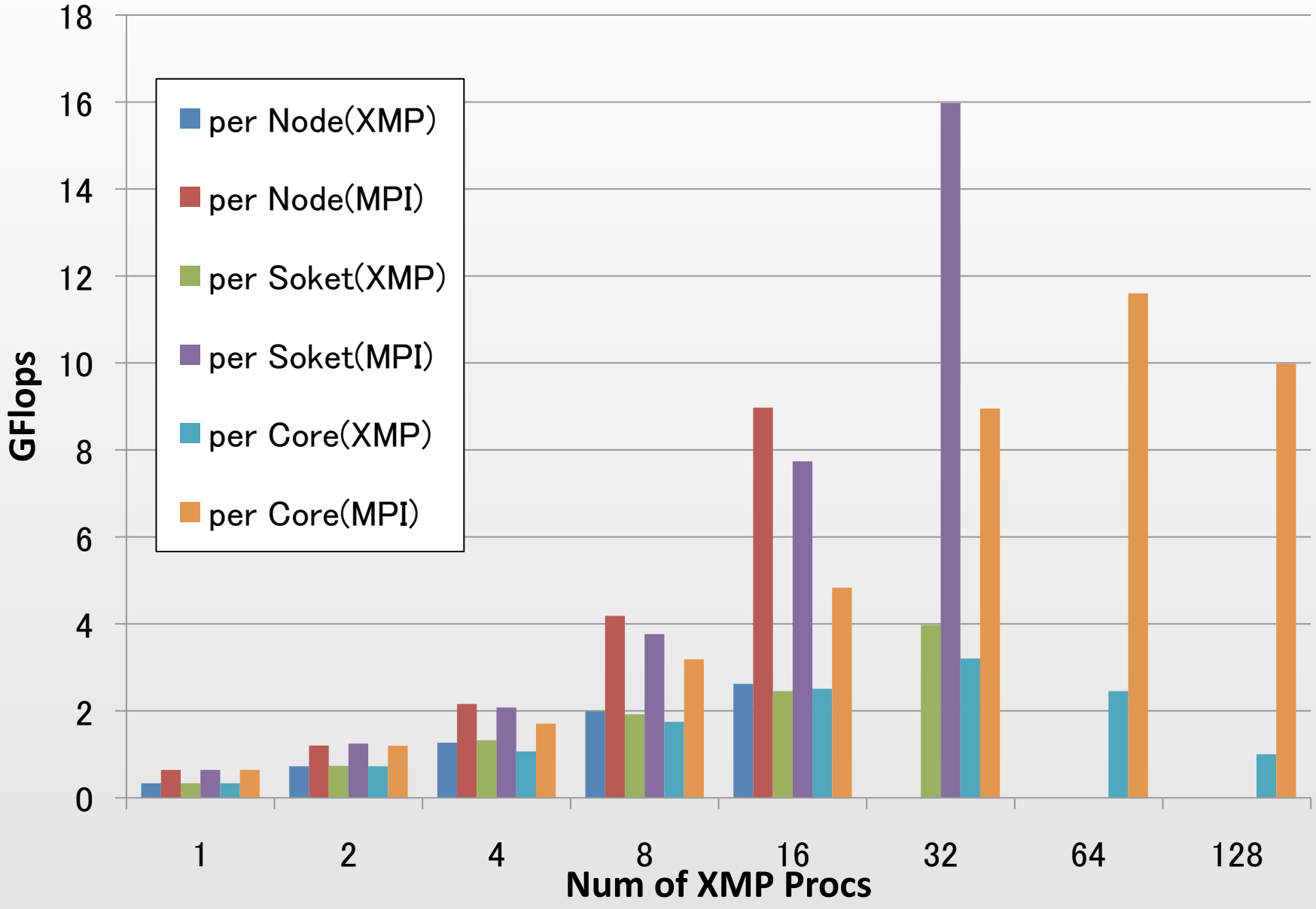
Performance of Laplace (Cray XT5)



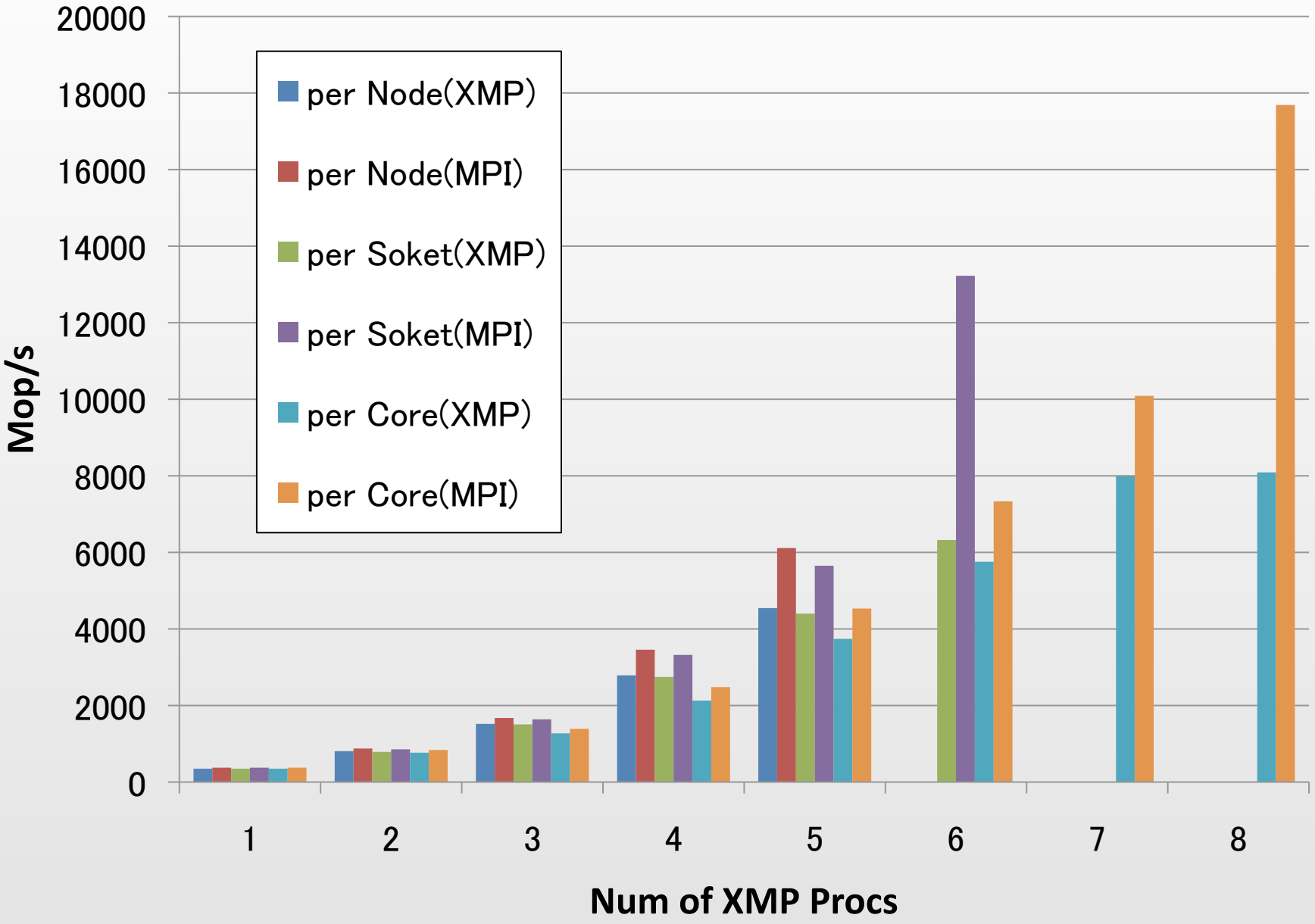
Performance of Linpack (CRAY XT5)



Performance of FFT (Cray XT5)



Performance of NPB-CG (Cray XT5)



Performance of RandomAccess (Cray XT5)