The Supercomputer Company

Quantum-ESPRESSO Performance on Cray Systems

Roberto Ansaloni, Cray Italy
Carlo Cavazzoni, CINECA
Giovanni Erbacci, CINECA

Agenda

- Quantum-ESPRESSO suite
- FPMD basic structures
- Porting and optimization on Cray systems
- FFT transpose: MPI vs CAF
- H2O-256 benchmark
- Performance analysis
- Conclusion and future work

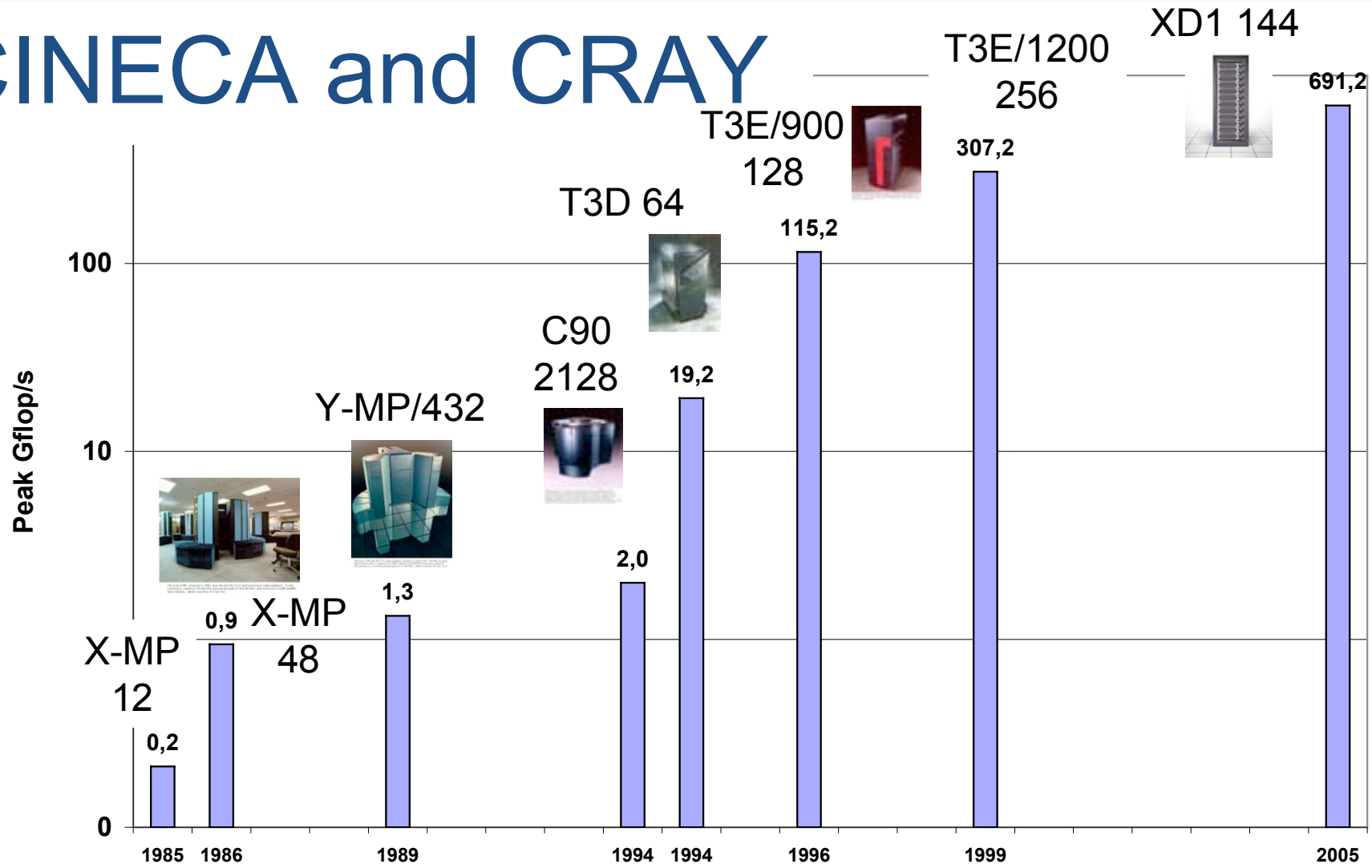
CINECA

- Consortium of 24 Italian Universities and CNR
- Founded in 1969
- Under the control of Ministry of University and Scientific Research
- 6 Sector Managers, 15 Group coordinator, 250 Employees
- Main activities: Research and Services for universities, private enterprises and Ministry

PETROGLYPHS TO PETAFLOPS
CUG 2005

Web site: http://www.cineca.it

CINECA and CRAY



Quantum-ESPRESSO

opEn-Source Package for Research in Electronic Structure, Simulation, and Optimization

- Merge of several pre-existing packages
  - existing since long time
  - used by many people around the world
- Implementation of techniques hindered by the lack of available software
  - Car-Parrinello Molecular Dynamics
  - linear response, ultrasoft pseudopotentials
- Suitable for easy addition of new developments
  - modular structure, easy to maintain and extend
  - collaborative free/open-source license

Current status

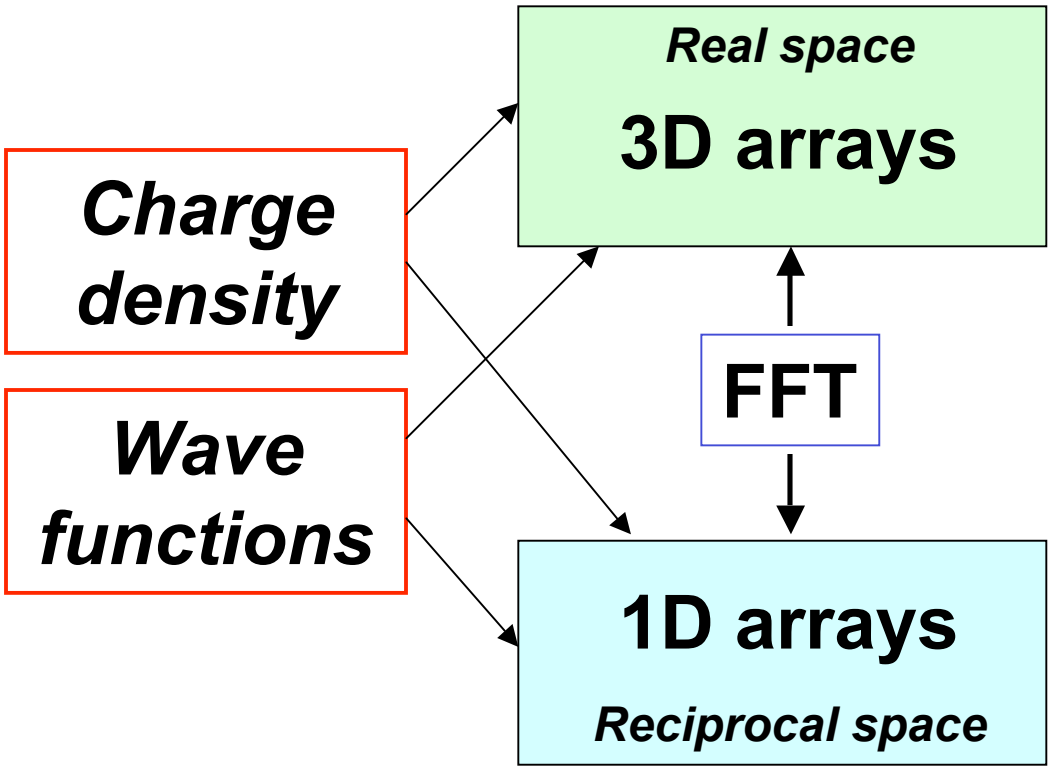
- **PWscf** (Trieste, Lausanne)
  - self-consistent electronic structure,structural relaxation, dynamics, linear-response, postprocessing
- **CP** (Lausanne, Princeton)
  - variable-cell Car-Parrinello molecular dynamics
  - ultra-soft pseudopotentials
- **FPMD** (Bologna, Trieste)
  - variable-cell Car-Parrinello molecular dynamics
  - norm-conserving pseudopotentials
- The three packages share
  - common installation method, formats, basic code
  - the package is maintained as a single CVS tree

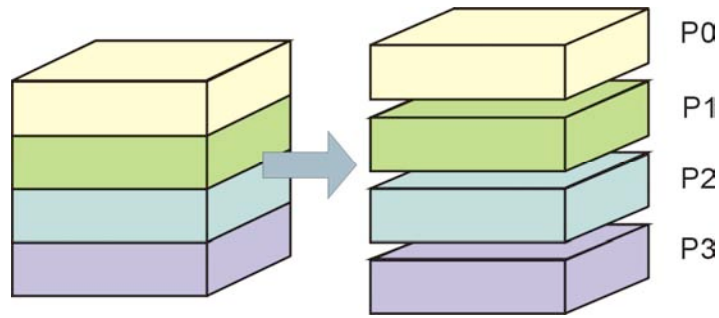
Web site: http://www.democritos.it/scientific.php

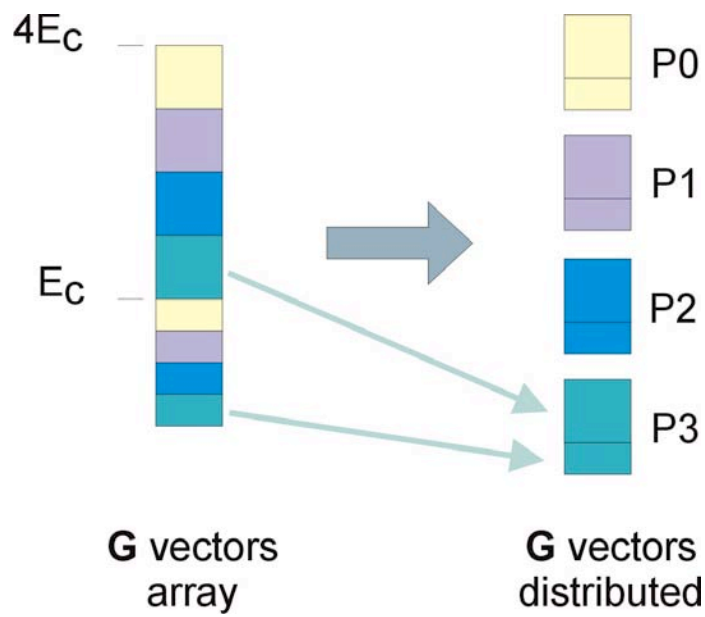
FPMD code

- Car-Parrinello variable-cell molecular dynamics
  - Carlo Cavazzoni (CINECA) and others
- Main features
  - Verlet dynamics with mass preconditioning
  - Various electronic and ionic minimization schemes
  - Temperature control with Nose'-Hoover thermostat
  - Constrained dynamics
  - And more… look at the web site
- Limitations
  - no Ultra-Soft Pseudopotentials

FPMD data structure



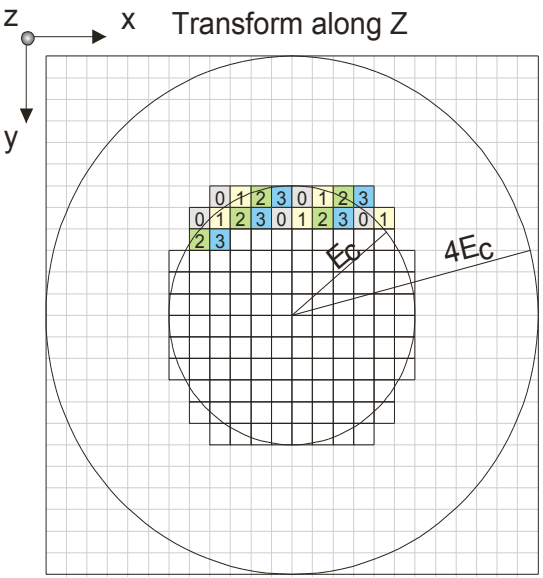




FPMD FFT implementation

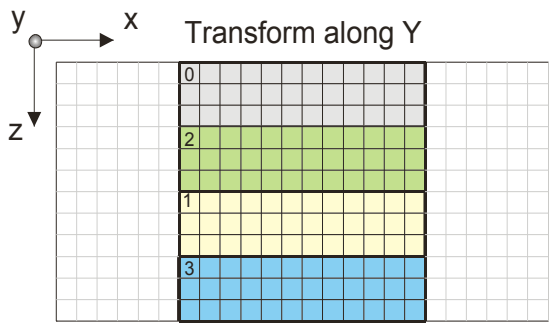
Reciprocal space
1D sticks

Real space
3D array slices

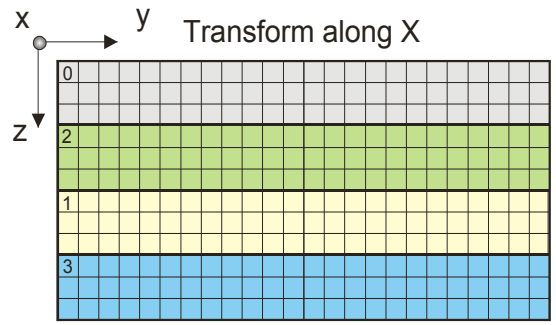


~ Nx Ny / 5 FFT along z

Parallel Transpose



Nx Nz / 2 FFT along y



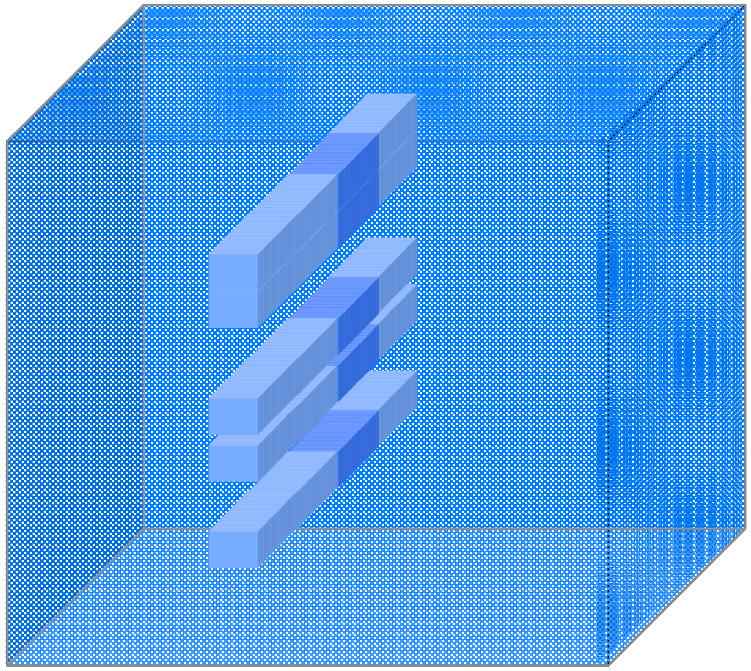
Ny Nz FFT along x

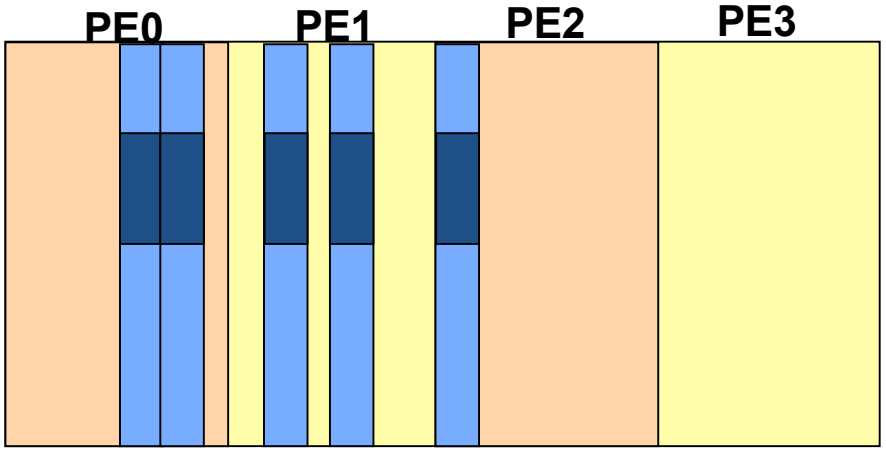
Optimization on Cray systems

- Cray Programming Environments: no porting issues
  - Cray X1: Cray PE, SciLib
  - Cray XT3, XD1: PGI compilers, ACML, FFTW
- Cray X1 optimizations
  - Critical loops vectorization
  - SciLib FFTs
  - Co-array (CAF) implementation of FFT Transpose
- Cray XD1 and XT3 optimizations
  - No optimization so far
  - XT3 will improve performance with optimized Portals

FFT transpose

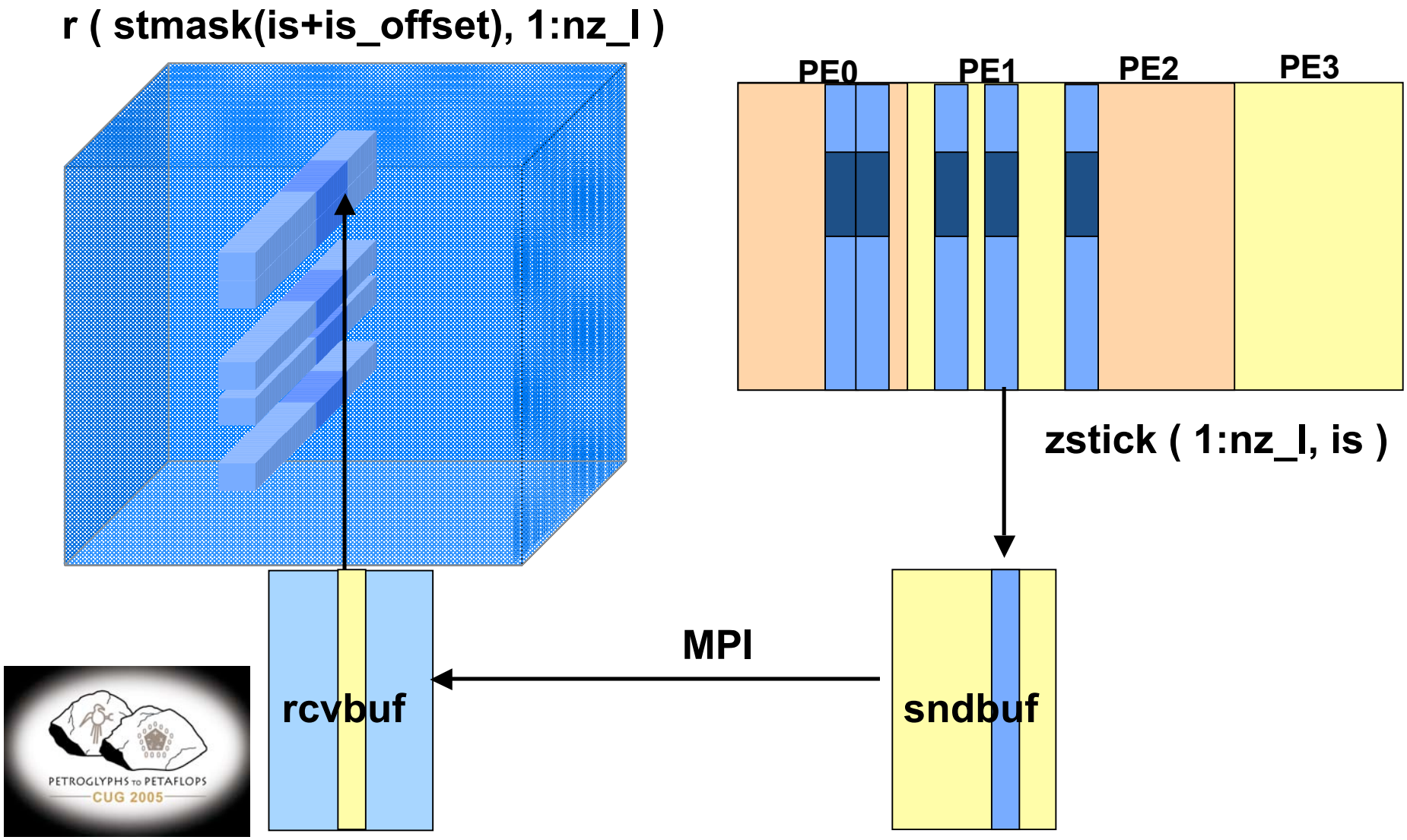
r (stmask(is+is_offset), 1:nz_l)





zstick (1:nz_l, is)

FFT transpose - MPI

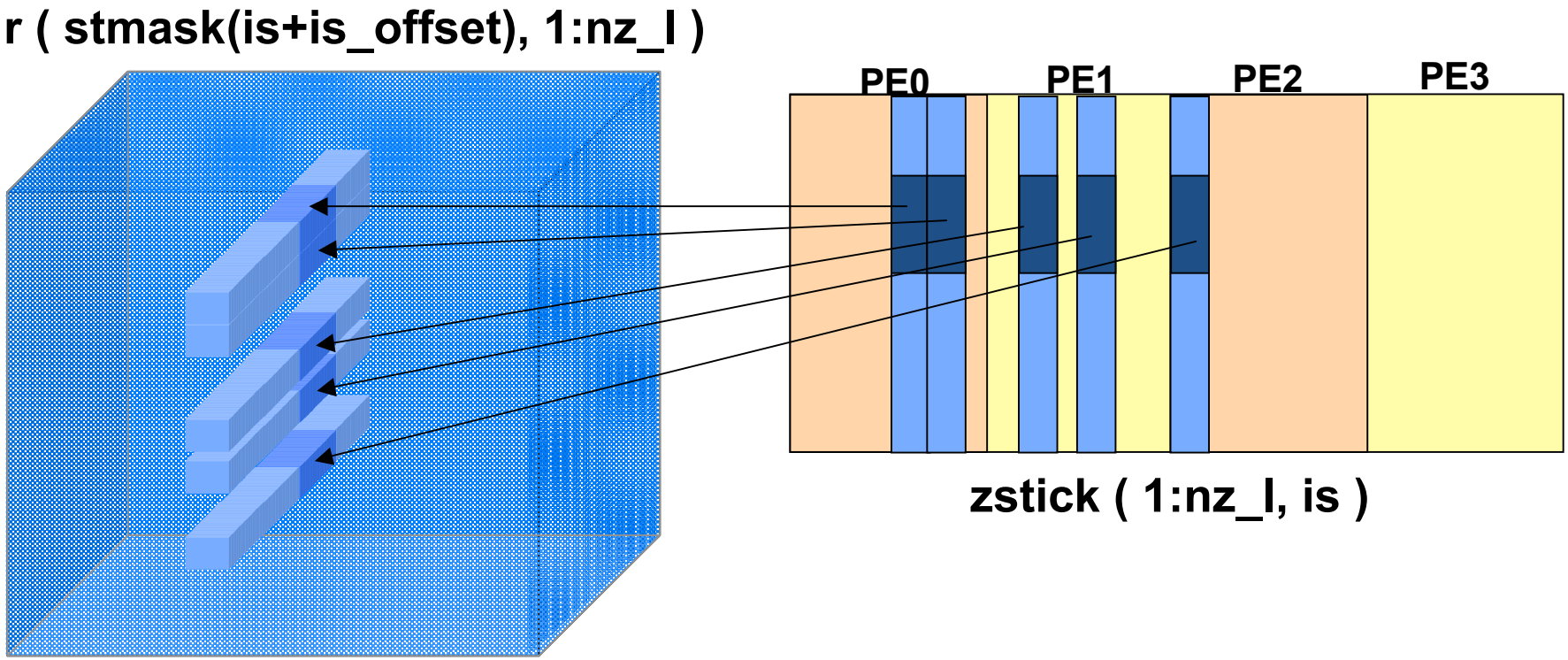


FFT transpose – MPI code

```
DO ipz = 1, npz
   itag = mype + 1 + npz * ( ipz - 1 )
     CALL mpi_irecv(rcvbuf(1,ipz), nbuf, MPI_DOUBLE_COMPLEX,
  END DO
DO ipz = 1, npz
   k_start = ( ipz - 1 )  * nz_l + 1 ;  k_end   = k_start  + nz_l - 1      offset  = - k_start + 1
   DO is = 1, ns_l
      DO k = k_start , k_end
           sndbuf(k + offset, ipz) = zstick( k + (is-1)*ldz )
         END DO
      offset = offset + nz_l
   END DO
   itag = ipz + npz * mype
     CALL mpi_isend( sndbuf(1,ipz), nbuf, MPI_DOUBLE_COMPLEX,
  END DO
111      CONTINUE
      DO IPZ = 1, NPZ
        CALL mpi_test(irhand(ipz),rtest(ipz),istatus(1,ipz), ierr)
```


```
IF( rtest(ipz) .AND. .NOT. rdone(ipz) ) THEN    offset = 0
is_offset = dfft%iss( ipz )
 DO is = 1, ns_lp - 1
   mc1 = stmask( is   + is_offset )
     DO k = 1 , nz_l
       r( mc1 + (k-1)*ldx*ldy ) = rcvbuf(k + offset, ipz)
```

FFT transpose: CAF direct copy



FFT transpose - CAF code

```
DO ipz = 1, npz
   k_start = ( me - 1 )  * dfft%npp( ipz )
   is_offset = dfft%iss( ipz )
   DO is = 1, ns_lp
      mc1 = stmask( is   + is_offset )
      DO k = 1 , nz_l
     r( mc1 + (k-1)*ldx*ldy ) =
        pzstick[ipz]%p(k_start+k+(is-1)*ldz)
      END DO
   END DO
END DO
```

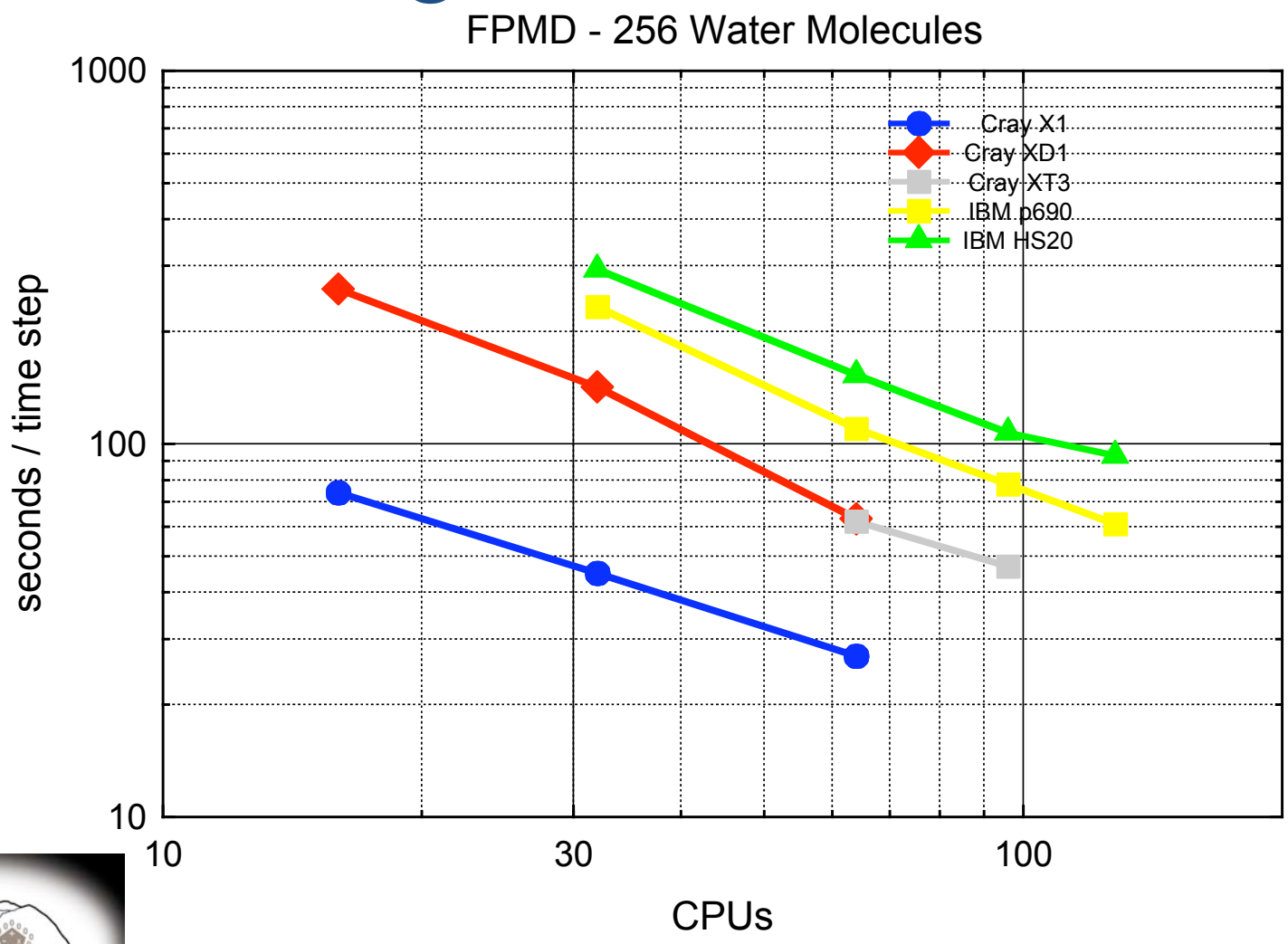
Benchmark testcase – H2O 256

- Water, 256 molecules
- FFT Grid: 220 x 220 x 220
- Electronic states: 1024
- Plane waves (wave functions): 513171
- Plane waves (charge density): 4105867
- Number of atoms: 768
- Total memory: 32 GBytes

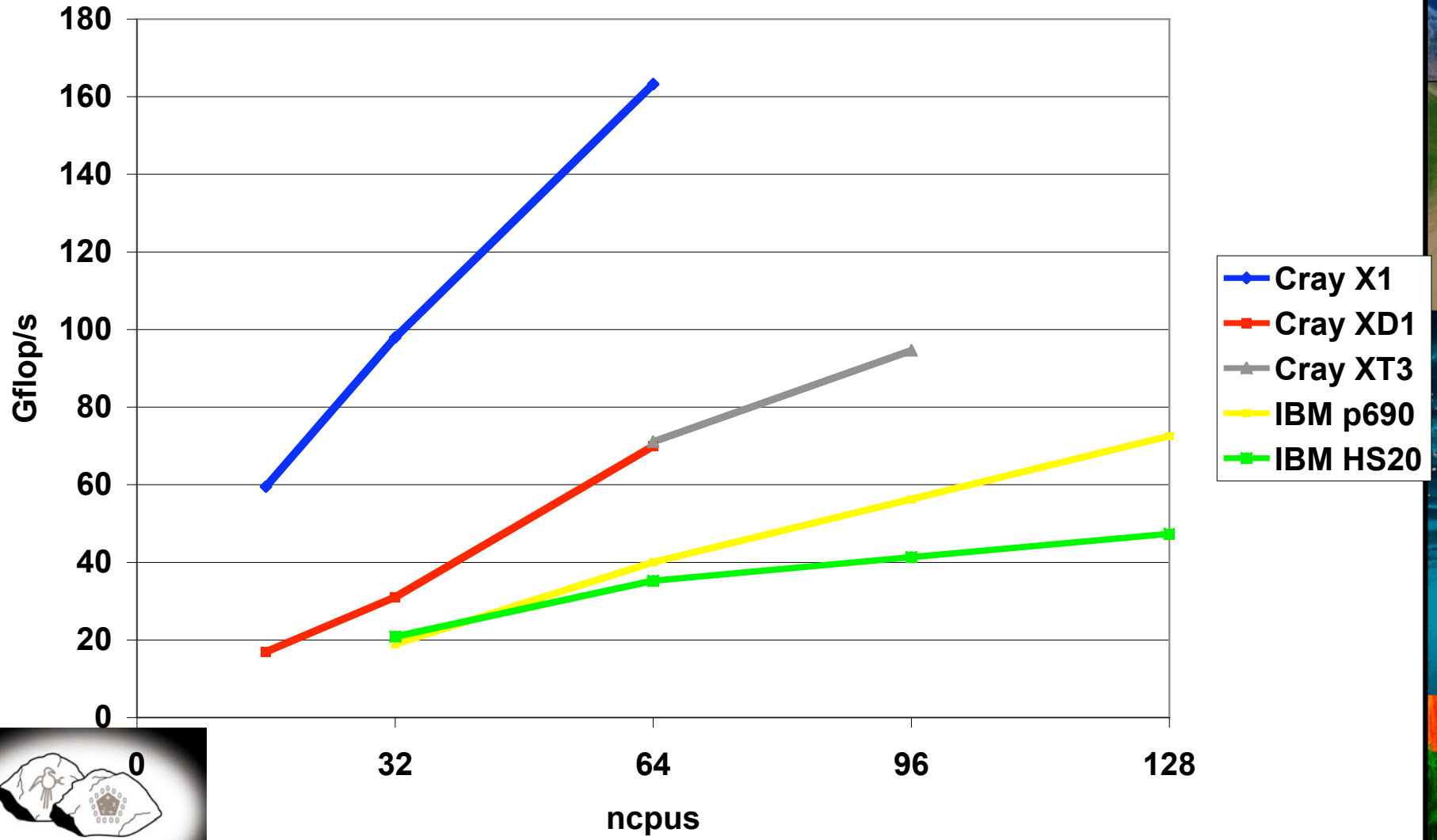
Systems evaluated

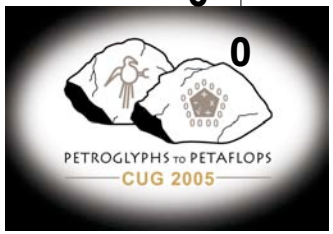
System	Processor	Peak cpu Gflop/s	Interconnect	Ncpus
Cray X1	Cray MSP 800 MHz	12.8	Custom Cray	16 – 64
Cray XD1	Opteron 2.4 GHz	4.8	Cray RapidArray	16 – 64
Cray XT3	Opteron 2.4 GHz	4.8	Cray SeaStar	64 - 96
IBM p690	IBM Power4 1.3 GHz	5.2	IBM HPS Federation	32 - 128
IBM HS20	Intel Xeon 3.06 GHz	6.1	Myrinet LAM C-D	32 - 128

FPMD timings

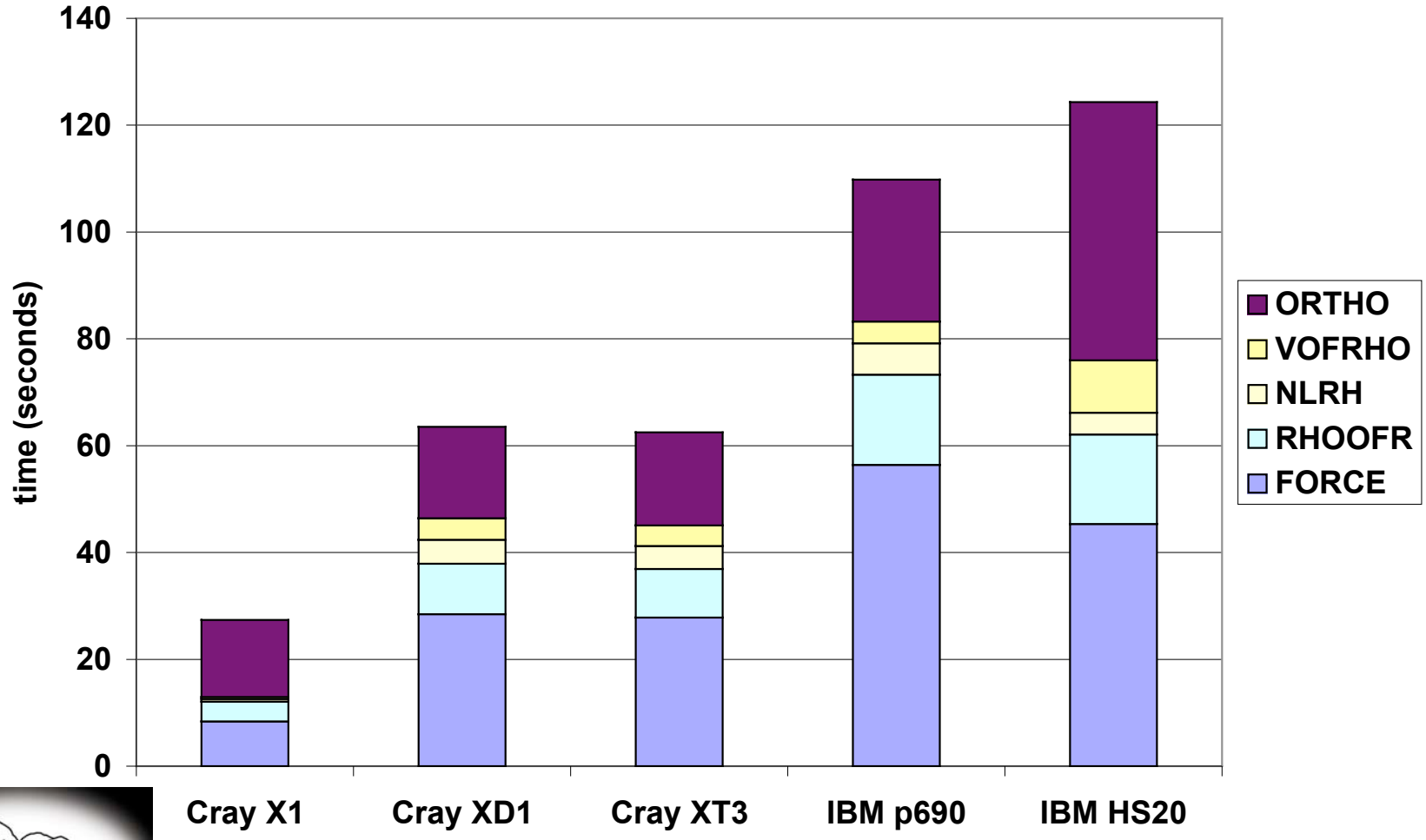


FPMD performance

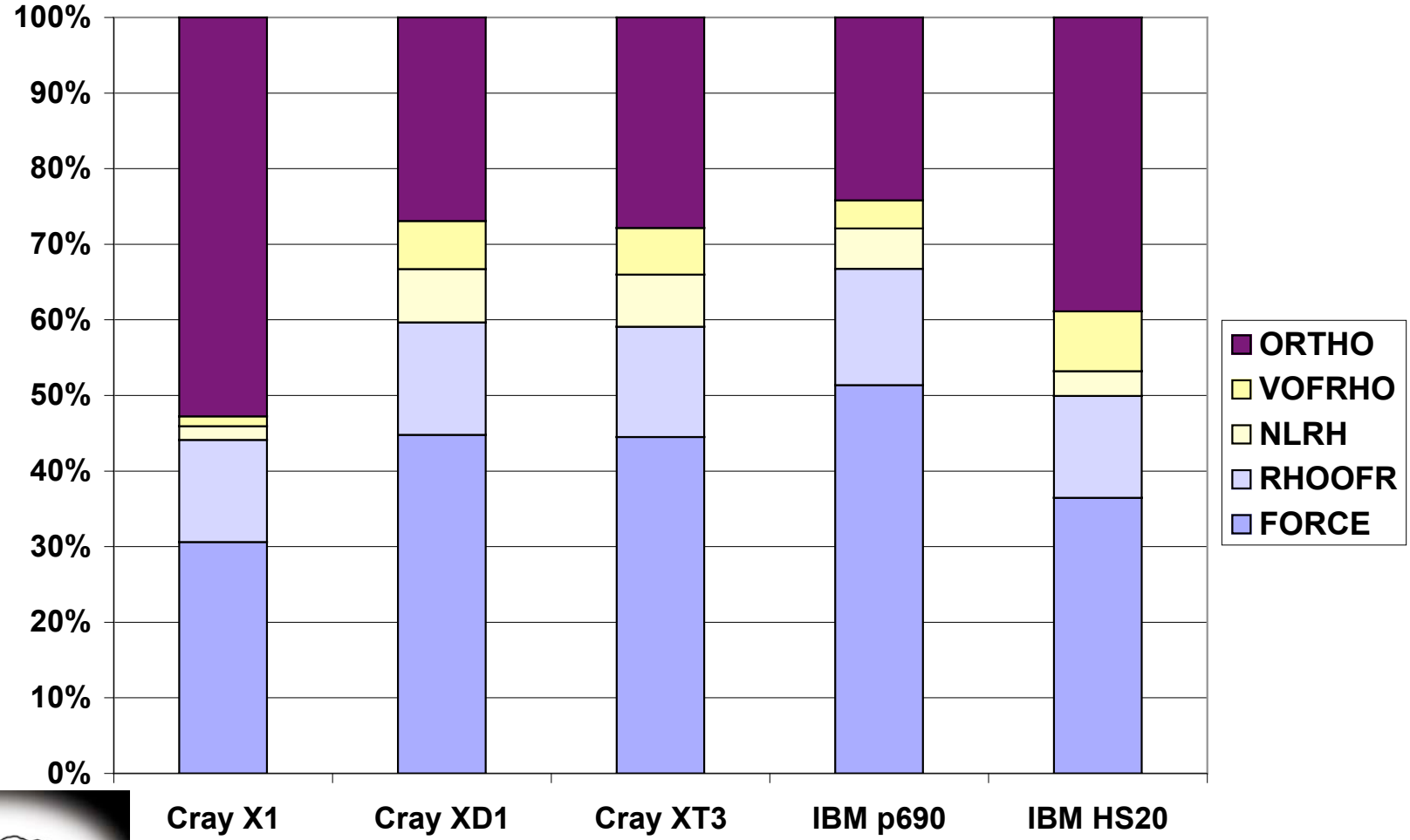




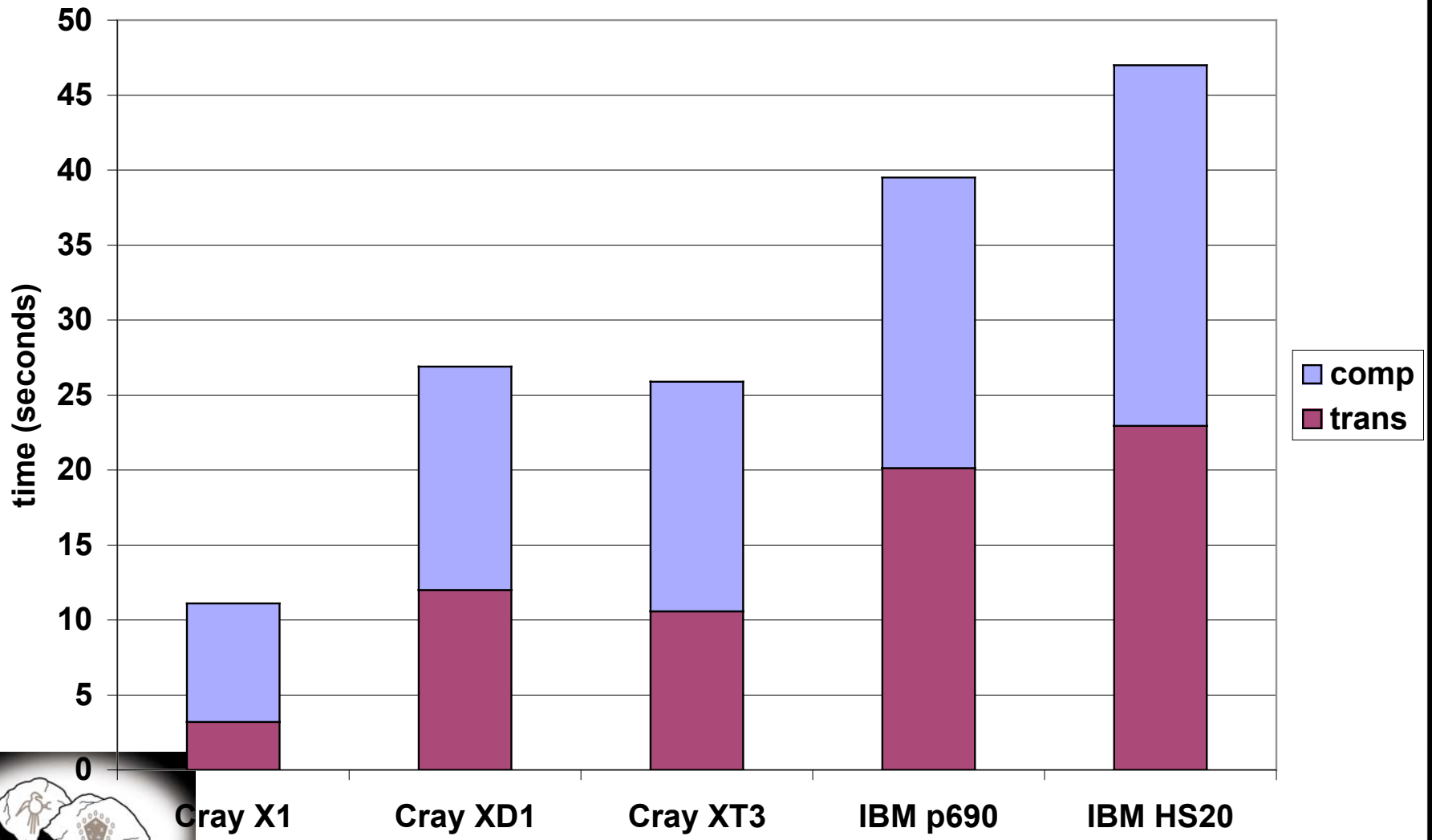
Work distribution at 64 cpus



Work distribution at 64 cpus (%)



FFT timings at 64 cpus



Conclusions

- FPMD shows good portability and scalability
- Optimization on Cray X1 has required some effort
  - SciLib routines, CAF-based communication
- Cray X1 is the fastest system
  - Fastest FFT: 2.3 times faster than XT3
  - Unoptimized ORTHO section affects scalability
  - Vector length reduction affects scalability
- Cray XT3 and XD1 are the fastest scalar MPP
  - FFT computational part ~ 25% faster than SP4
  - FFT transpose ~ 2x faster than SP4

In spite of peak performance

Future work

- Extend the work to the other codes in the suite
  - Exploit shared code
  - Run benchmarks
- Cray XT3 and XD1 optimizations
  - PGI compiler optimization
  - ACML FFT routines
  - SHMEM based FFT transpose