HPC Benchmarking

Vladimir Marjanovic, Colin W. Glass, Jose Gracia

Goals of TOP500 HPC Benchmarking

- Ranking supercomputers
- One number all properties

HPC Benchmarking Approaches

- Two approaches:
  - single application (kernel) :

    HPL, HPCG, **HPGMG**
  - many applications (kernels):

    HPGMG, NAS benchmark, HPC Challenge, MPI SPEC 2007 etc

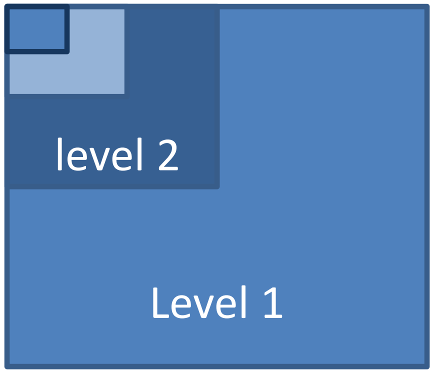
Real App Representation

- HPL - direct solver $O(n^3)$
- HPCG, HPGMG(FV, FE) – iterative solver O(n)
- HPL is n^2 times slower than HPCG and HPGMG

HPC Reference Codes

	MPI	OpenMP	External lib	Metric
HPL	yes	BLAS MT	BLAS	GFLOP/s
HPCG	yes	SpMV	no	GFLOP/s
HPGMG-FV	yes	yes	no	DoF/s
HPGMG-FE	yes	no	PETSC	GFLOP/s

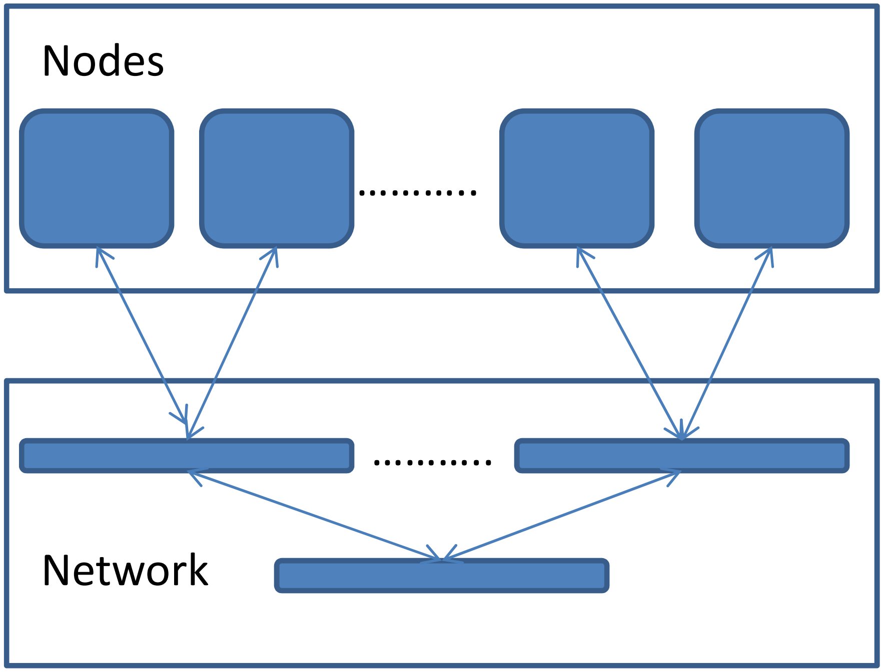
- Well written balanced codes
- HPCG – MG only 3 levels
- HPGMG – MG many levels



Important System Properties

- Compute power (Flop/s)
- Main Memory BW (B/s)

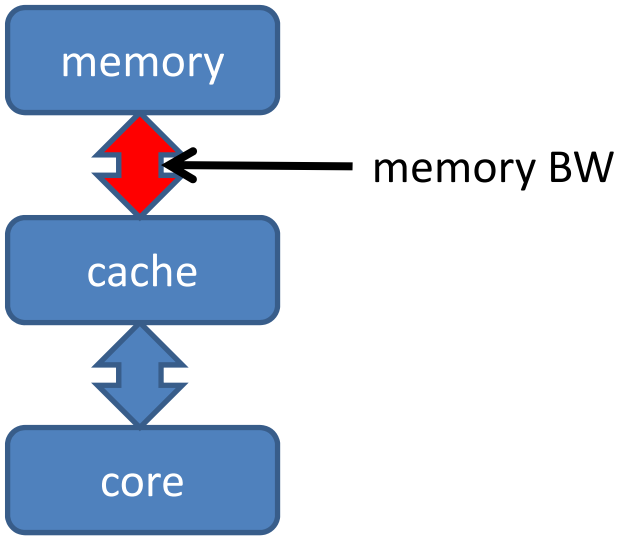
- IC BW(B/s)
- IC latency(s)
- IC message rate



One number = f(properties)

Memory/Compute bound – Byte/FLOP 1/2

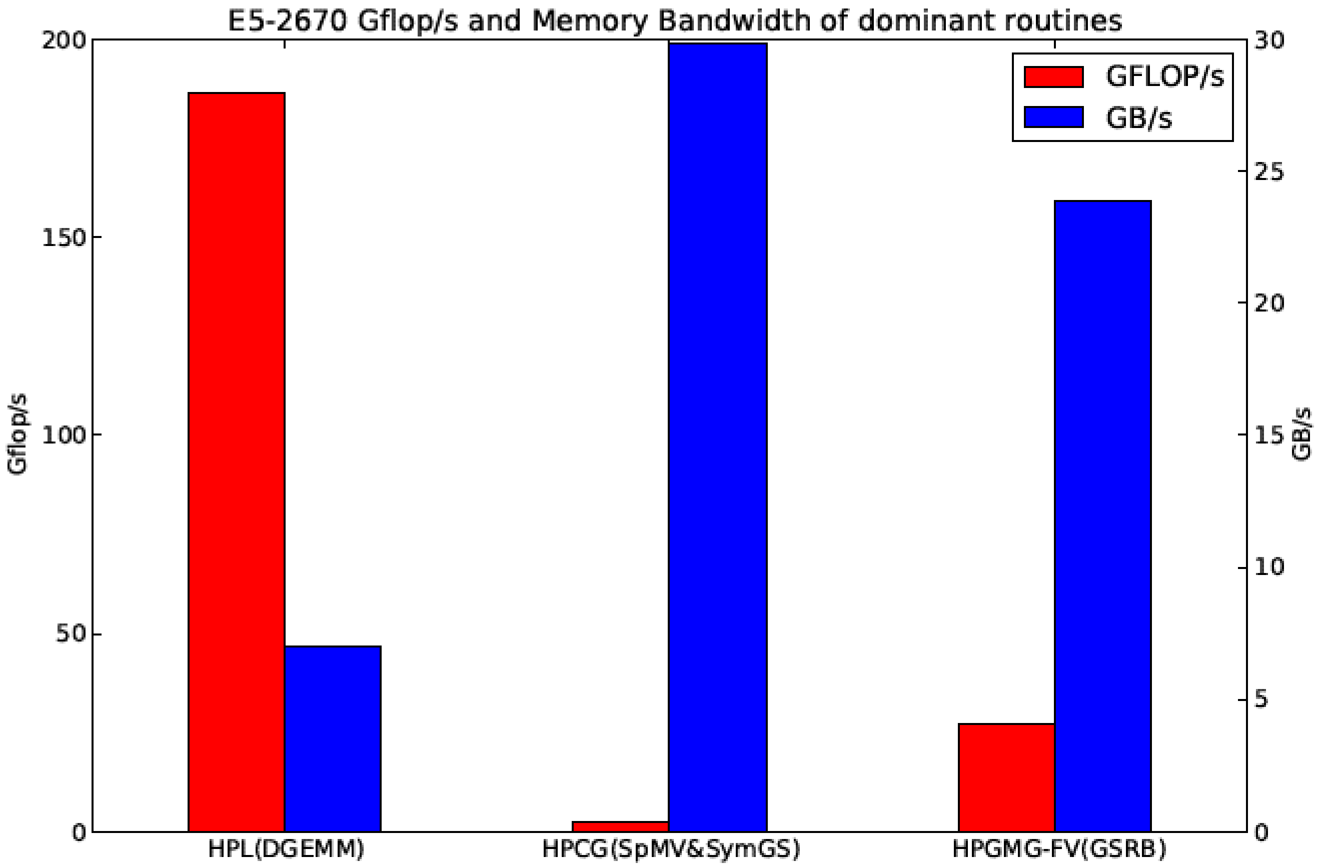
benchmark	kernel	Byte/FLOP
HPL	DGEMM	12/n = f(n)
HPCG	SpMV, SYMGS	> 4
HPGMG	GSRB	>1



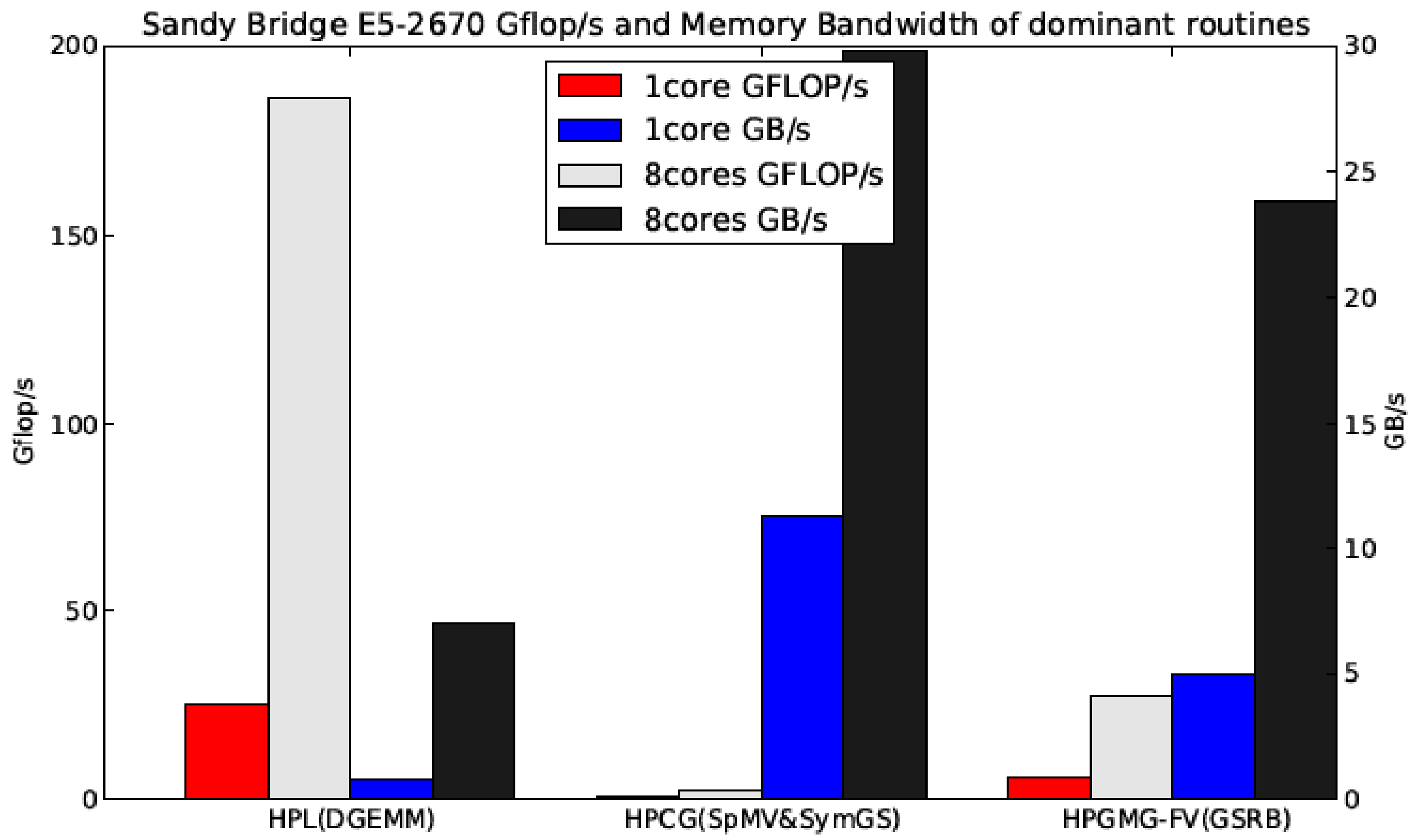
- Modern hardware ≈ 0.3 Byte/Flop
  e.g E2680v3 has 0.14 Byte/Flop
  NEC SX-ACE 1Byte/Flop

- HPCG & HPGMG kernels are memory bound on modern hardware

Memory/Compute bound – Byte/FLOP 2/2



HPL, HPCG, HPGMG core&node



HPCG:

- Gflops 2.75x
- BW 2.63x

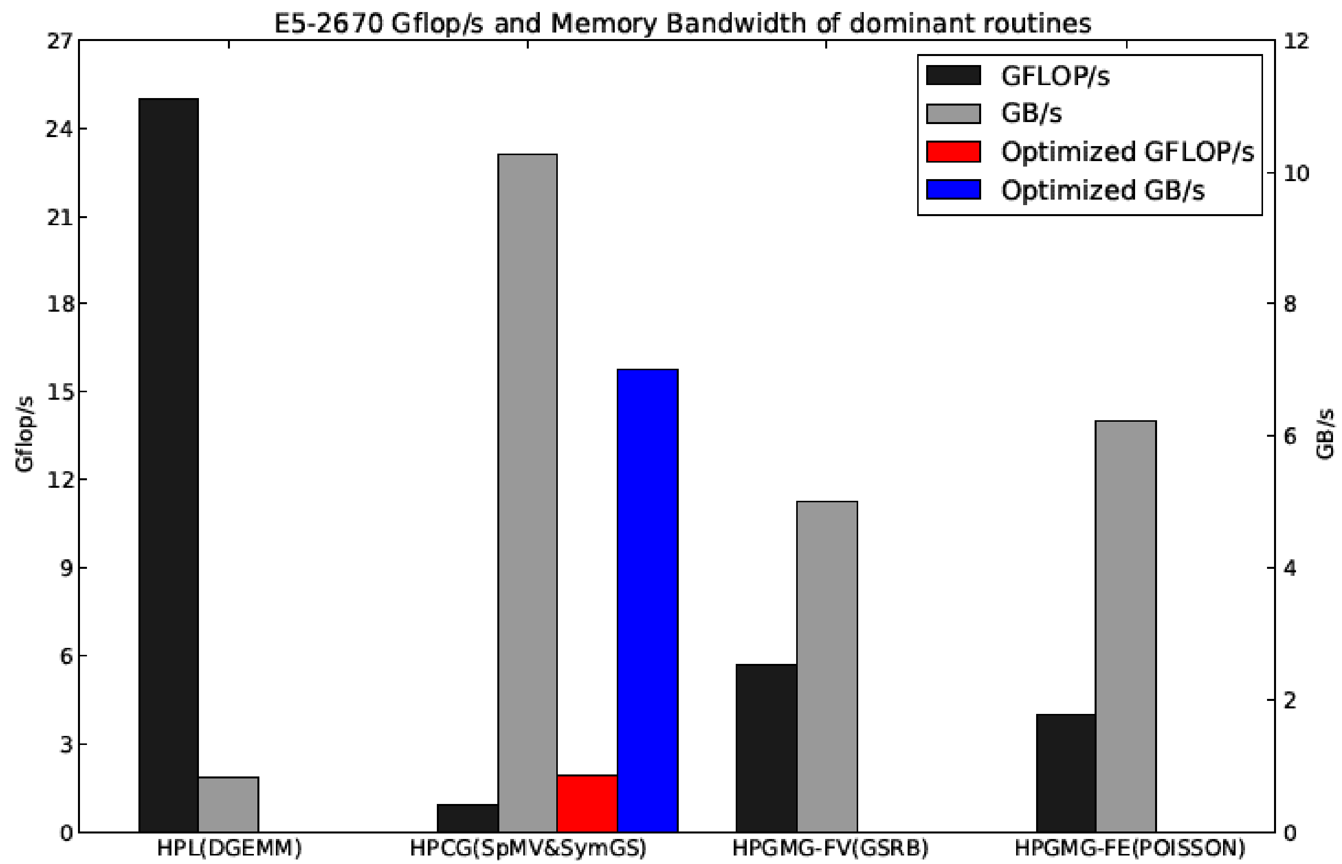
HPGMG:

- Gflops 4.78x
- BW 4.77x

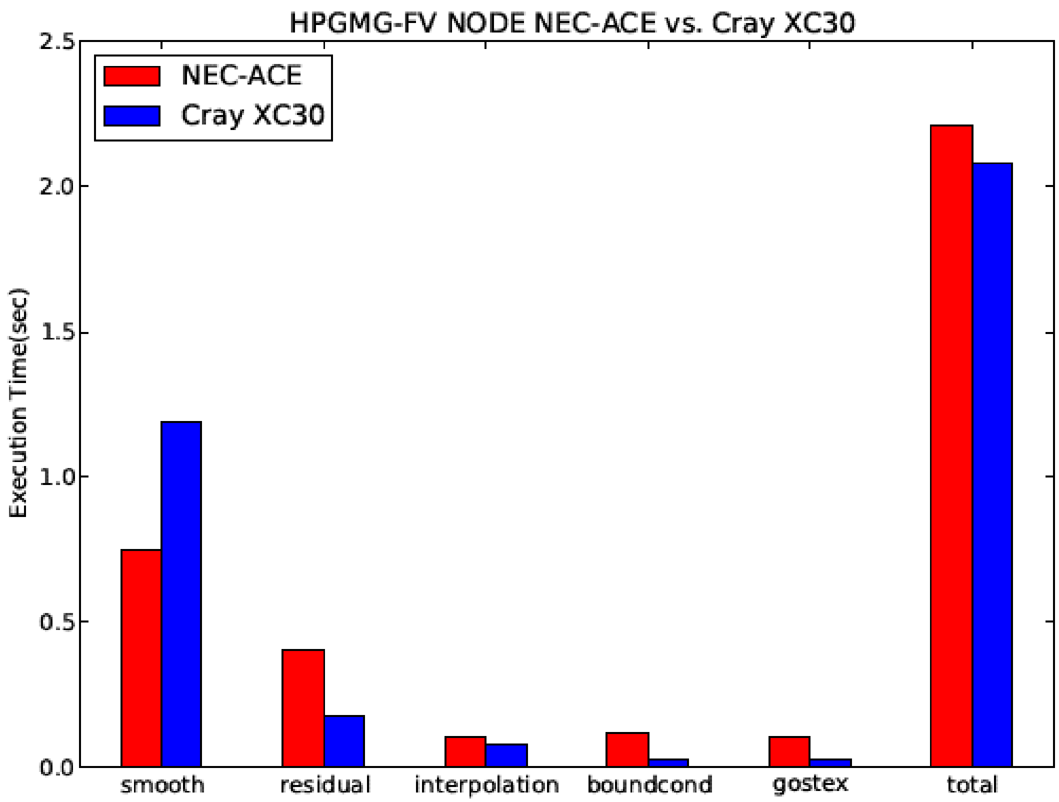
HPL

- ~8x

HPCG Intel Optimized Version



HPGMG-FV NEC-ACE versus CRAY XC30

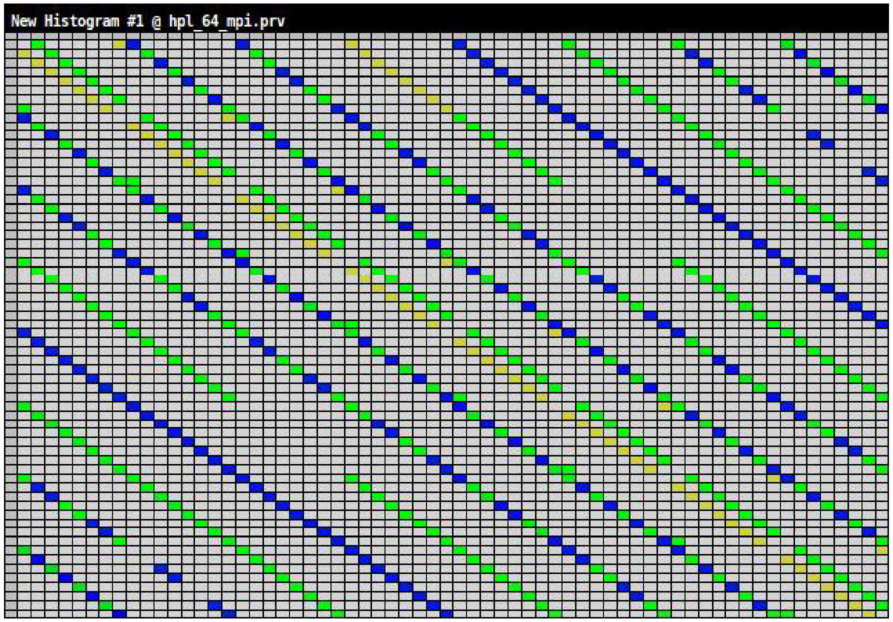


Ivy Bridge 5576

NEC SX-ACE

- 4 cores for 256GFlops peak performance
- 1 GHz
- 16 vectore pipes per core
- 32 memory channels for 256GB/s bandwidth, one core or sharable

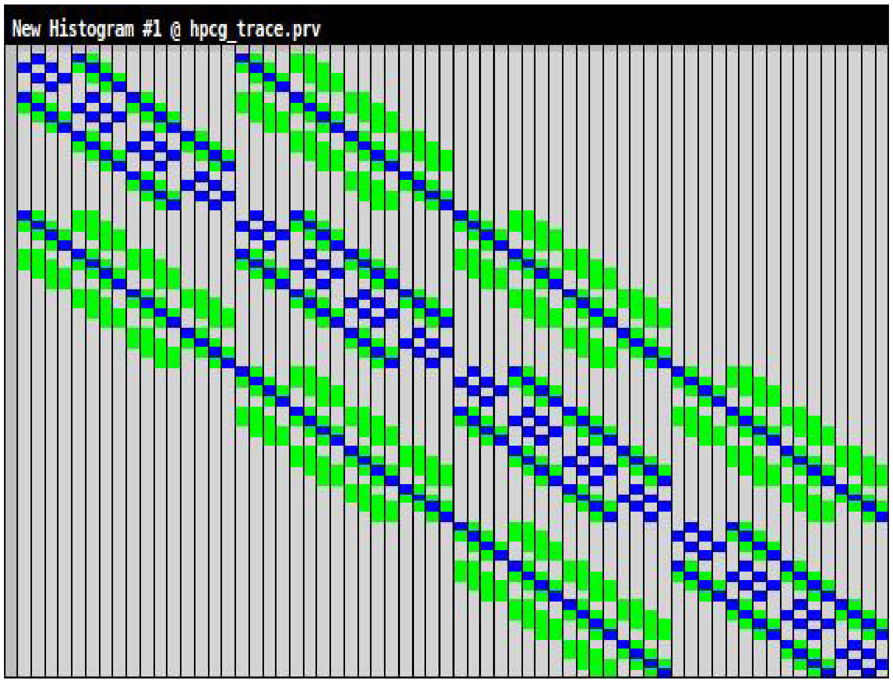
Communication HPL



- MPI communicators
- whole range of message sizes (few bytes < size < 100 MB)
- large number of small messages (fact > 256)

➢ stresses bandwidth, latency and message rate

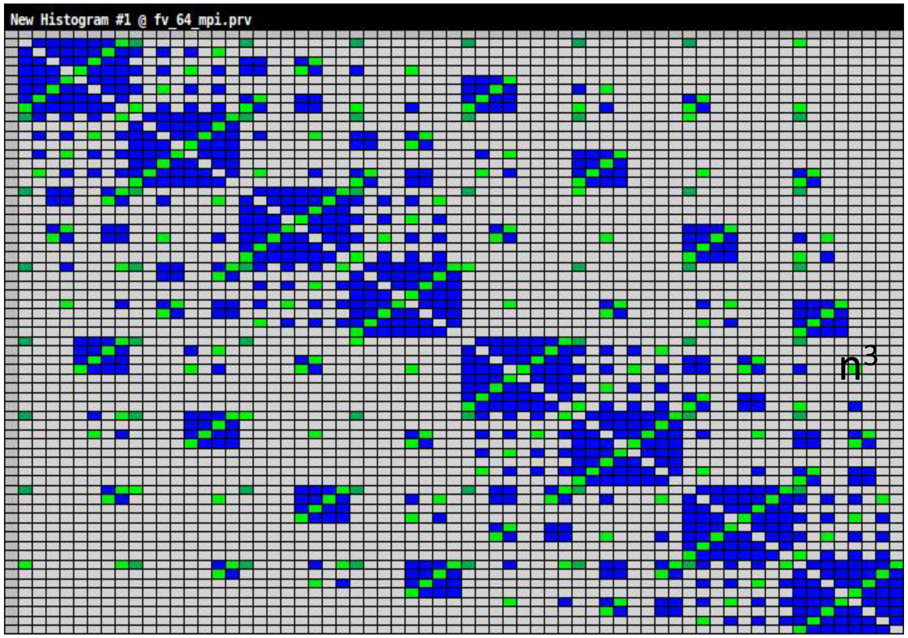
Communication HPCG



- small and large messages
  (few bytes < size < 512 KB (4 MB))
- number of messages
  (ex.halos max is 26)

➢ stresses latency

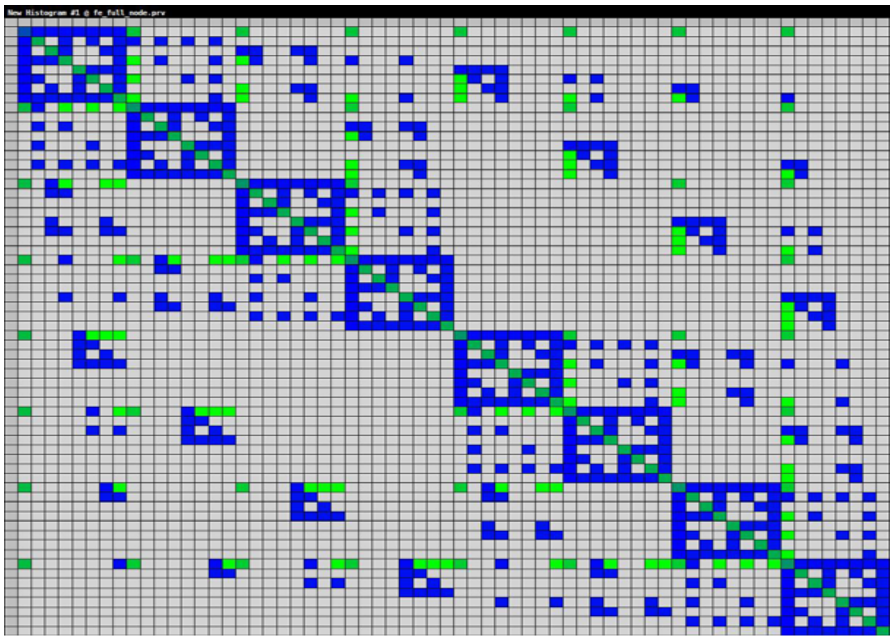
Communication HPGMG-FV



- MPI communicators
- small and medium messages few (bytes < size < 4 MB)
- large number of small messages (> 800 for mpi64)

➢ stresses latency and message rate

Communication HPGMG-FE



- MPI communicators
- small and medium messages few (bytes < size < 4 MB)
- large number of small messages

➢ stresses latency and message rate

HPL HPCG HPGMG Overview

	Representation	Compute Power	Memory BW	IC BW	IC Latency	IC Routing
HPL	X		X			
HPCG		X		X		X
HPGMG				X		

- Potential of overlapping of communication and computation
- What will happened in poor IC (Data Center, Cloud) ?

Arbitrary problem size issue

- HPL complexity: computational $O(n^3)$ communication $O(n^2)$
- HPCG,HPGMG complexity: computational $O(n)$ communication $O(n^{2/3})$

- Large problem size -> computation dominates communication
- Uniform behavior of routines or single dominant routine

- HPL - > DGEMM
- HPCG - > SpMV, SYMGS
- HPGMG -> GSRB

Conclusions

- Stresses memory bandwidth, compute unit and routing
- Representation of real applications
- Arbitrary problem size (weighting or integral as solution)
- Clear rules for benchmarking