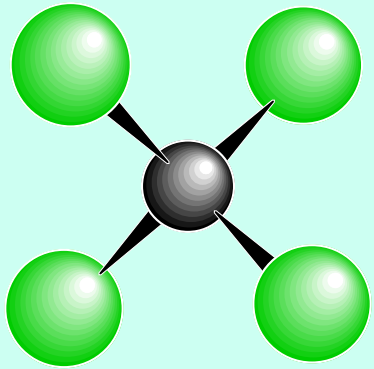
Massive Parallelism - The Hardware for Computational Chemistry?



Martyn F. Guest, CCLRC Daresbury Laboratory
m.f.guest@daresbury.ac.uk

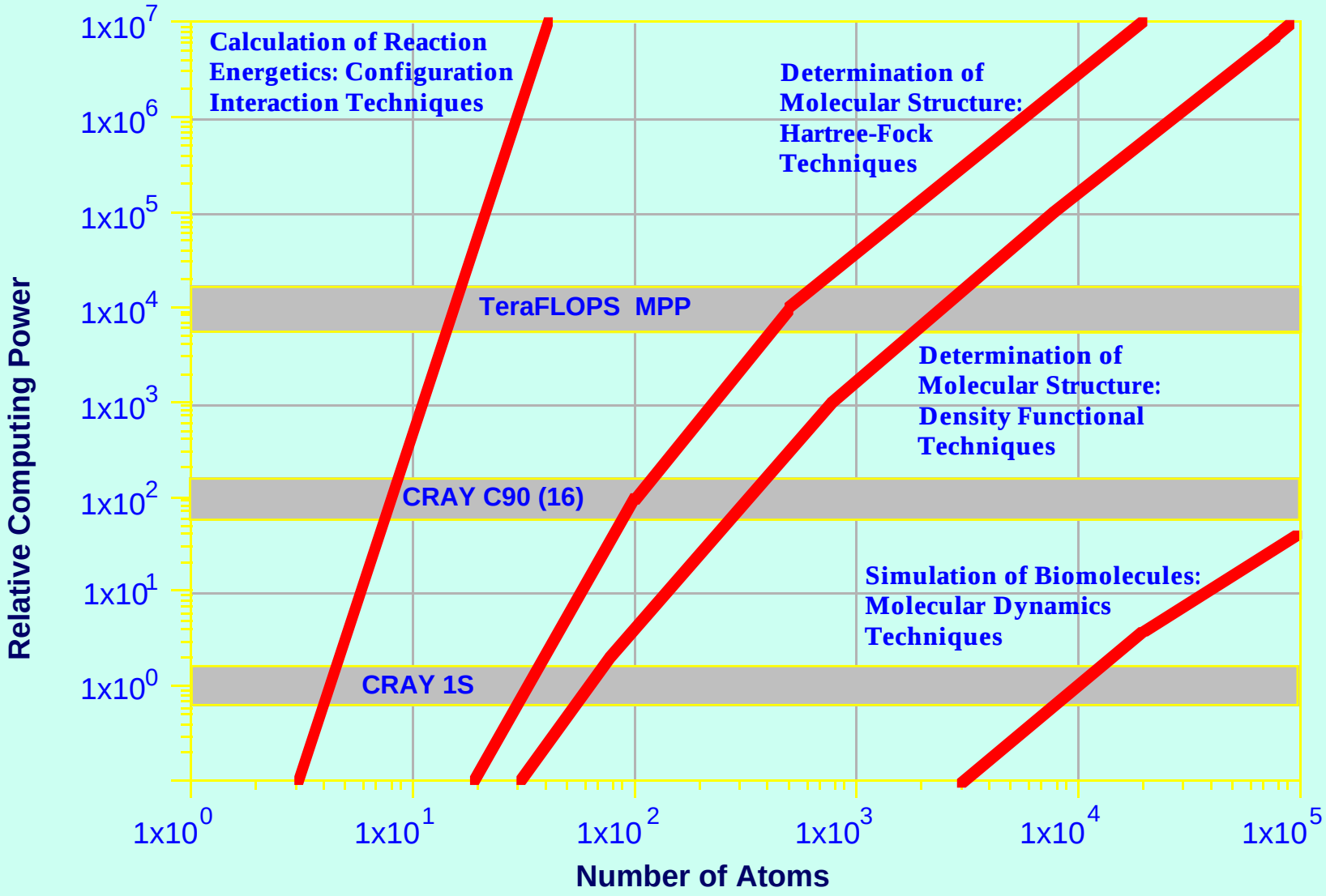
and

Giuseppe Vitillaro, University of Perugia
peppe@unipg.it

Outline

- **Issues in HPC and Computational Chemistry**
  - Background
  - Scalability, Distributed Data Structures and NUMA
  - "Grand Challenge" vs. Throughput requirements
- **Grand Challenge Requirements**
  - Distributed data structures and Global Arrays (GAs)
  - NWChem & the EMSL(Environmental Molecular Sciences Laboratory)
- **Minimising time-to-solution and Throughput**
  - GAMESS-UK; replicated data and GA's (limitations on current MPPs)
  - Direct-SCF, MP2 gradients and Density Functional Theory (DFT)
- **Cost-effectiveness of MPPs:**
  - Single-node performance
  - UNIX workstations vs. PC's (Pentium Pro)
- **Integral bottleneck and O(N) developments**
  - Multipole approximations and CMM
- **Summary**

Scaling of Molecular Computations



MPPs & Outstanding Problems in Computational Chemistry

- High latencies and low bandwidths
- Availability of parallel linear algebra libraries
- Availability of compilers which generate efficient code for linear algebra constructs
- Poor stability and usability of new MPPs
- Lack of standards, *e.g.*, in message-passing and languages (MPI and HPF)
- Exploiting continuing advances in computational chemistry and computer H/W and S/W
- Lack of scalable I/O

MPP Technology

	Current	Next Generation
LFK (node Mflop)	300-400	1000-1200
Communications:		
Bandwidth (Mbytes)	100-300	1000
Latency (usec)	10-30	3-10

Scalable Parallel Algorithms

An algorithm is scalable if it is possible to maintain constant time to solution by using more processors for larger problems with an acceptable level of efficiency.

- **Accustomed to cpu efficiency –** $\varepsilon(P,N,\ldots) = \dfrac{T(1,N,\ldots)}{P \cdot T(P,N,\ldots)}$
  - **Should consider all machine resources – memory, disk, …**
- **Only scalable algorithms are cost-effective**
  - **E.g., distributed vs. replicated data**
  - **500-1000 node MPP vs. 50 X 16-processor SMP**

Hartree-Fock SCF Model

- **Wavefunction:** $\Psi = A\phi_1\phi_2\cdots\phi_N$
- **Expansion of one-electron orbitals:** $\phi_j = \sum_{\mu=1}^{n} \chi_\mu C_{\mu j}$
- **Scales as** n^3 - n^4 **(with integral screening)**

$$f_{\mu\nu} = h_{\mu\nu} + \sum_{i}^{N/2} \sum_{\lambda\sigma} C_{\lambda i} C^{*}_{\sigma i} \left[2(\mu\nu|\sigma\lambda) - (\mu\lambda|\sigma\nu) \right]$$

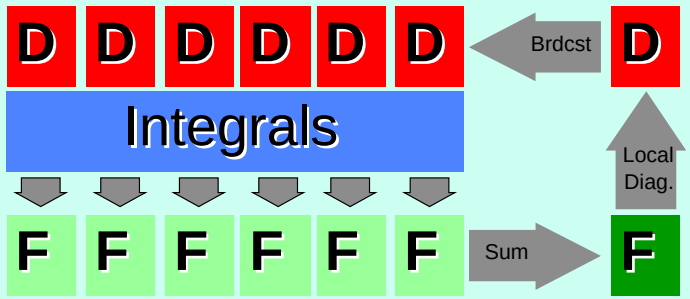
fock matrix

kinetic energy + nuclear attraction

orbital expansion coefficients

two-electron integrals

Replicated Data SCF



HPCC - Tools

Message Passing **Memory Allocation** **Global Arrays** **Math Libraries**

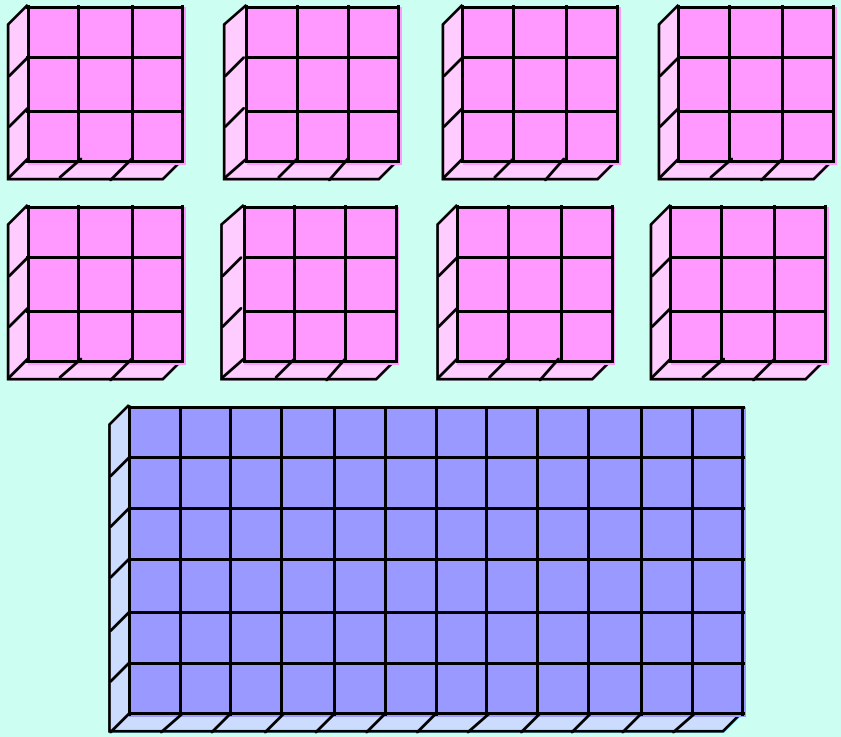
- Efficient and portable “shared memory” interface
- Scalable
- Adheres to standards or predicted standards
- Insulates chemistry programmer from CS intricacies
- Placed in the public domain to encourage adoption

These tools, developed at PNNL over the past few years, are the basis for the success of the NWChem project and the wealth of functionality coming on line.

ftp.pnl.gov/pub/permanent

Global Arrays

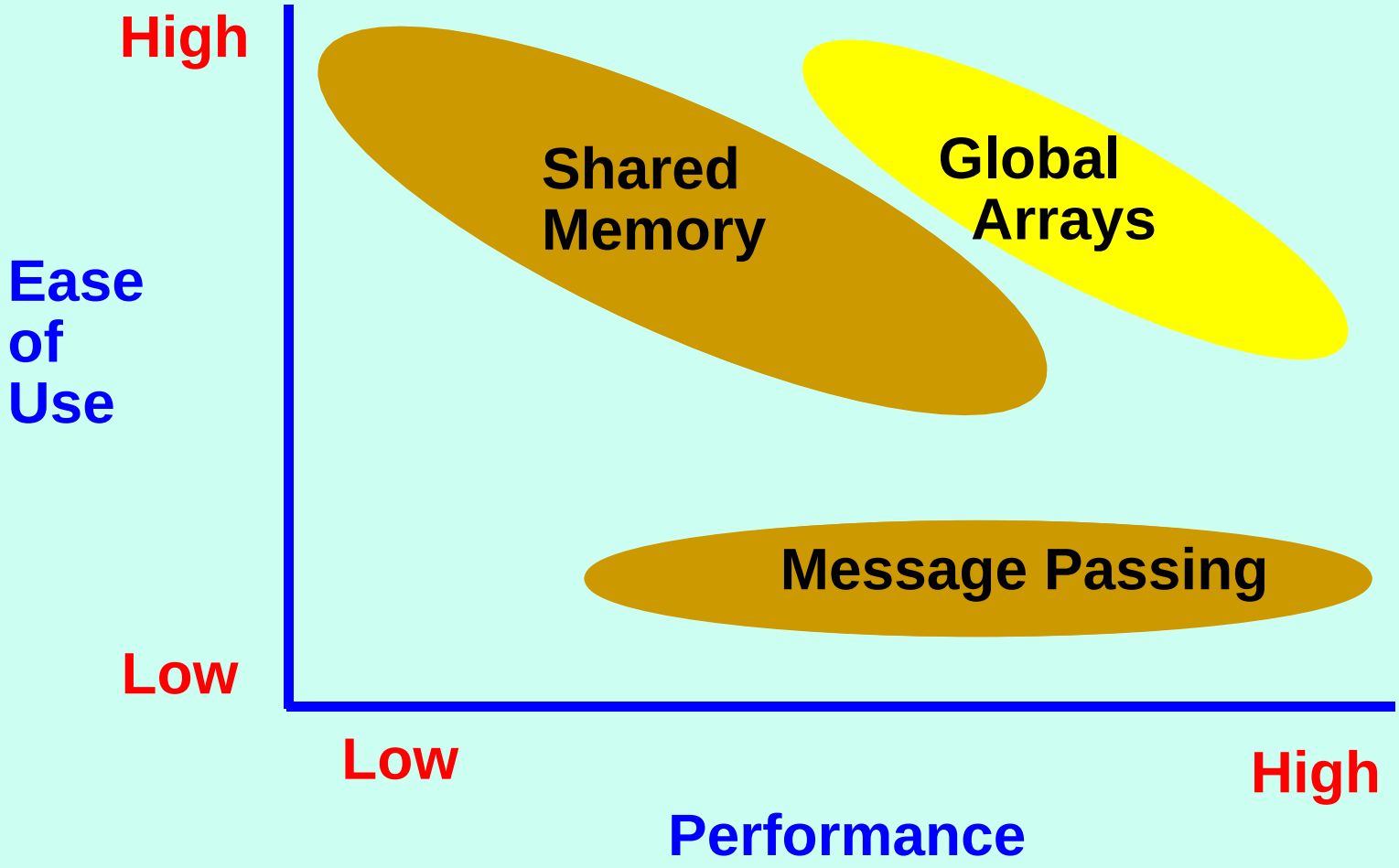
Physically distributed subblocks



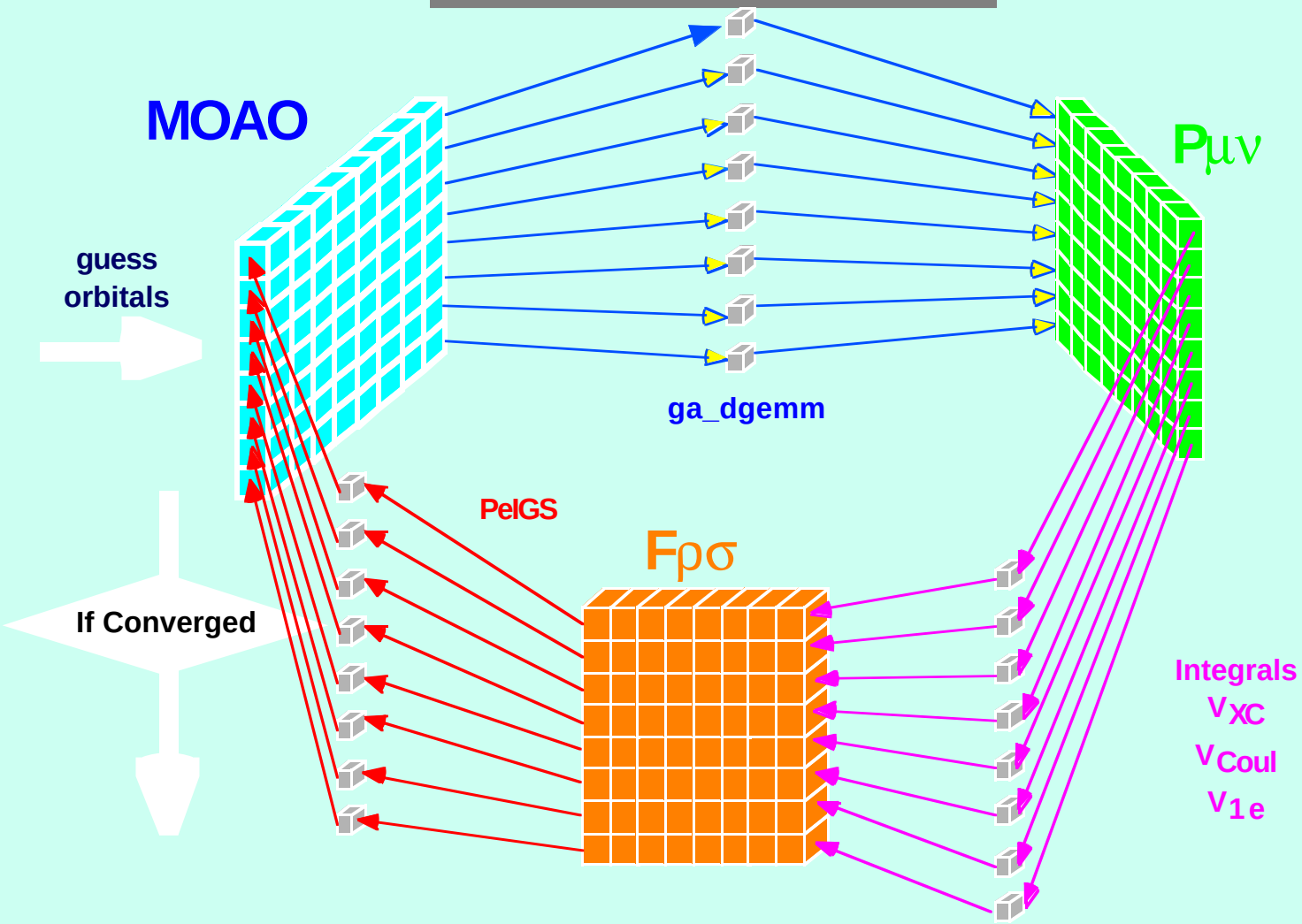
Single logical data object

- *One-sided* asynchronous remote access
- Fast local access
- MIMD and data-parallel modes
- Logical array addressing - topology independent
- BLAS and linear algebra functionality
- Layered over message passing primitives or shared-memory facilities

How GA Compares to Other Scalable Models



SCF Module

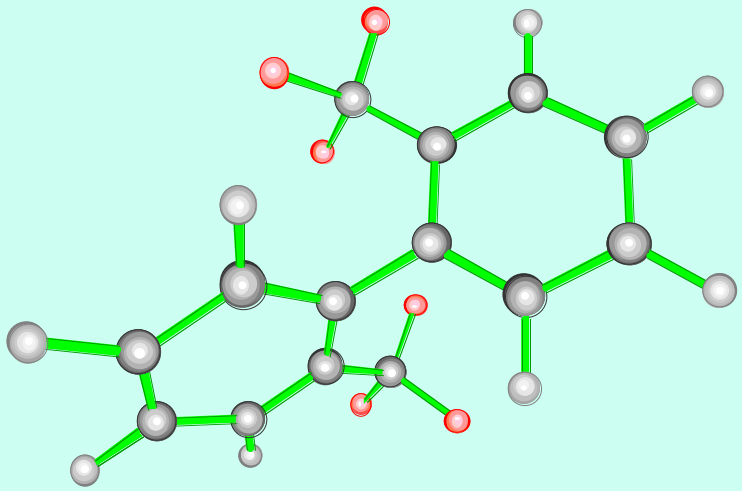

MOAO
guess
orbitals
Pμν
ga_dgemm
PelGS
Fρσ
If Converged
Integrals
VXC
VCoul
V1e

HF Scalability for In-core Calculations

- $(C_6H_4(CF_3))_2$ – 324 function DZP basis
- Representative of many routine calculations – at the limit of in-core capability
- Massive parallelism and in-core make this *interactive*
- Demonstrates that NWChem will fully exploit faster integrals

Total SCF time (10^{-5} in density)

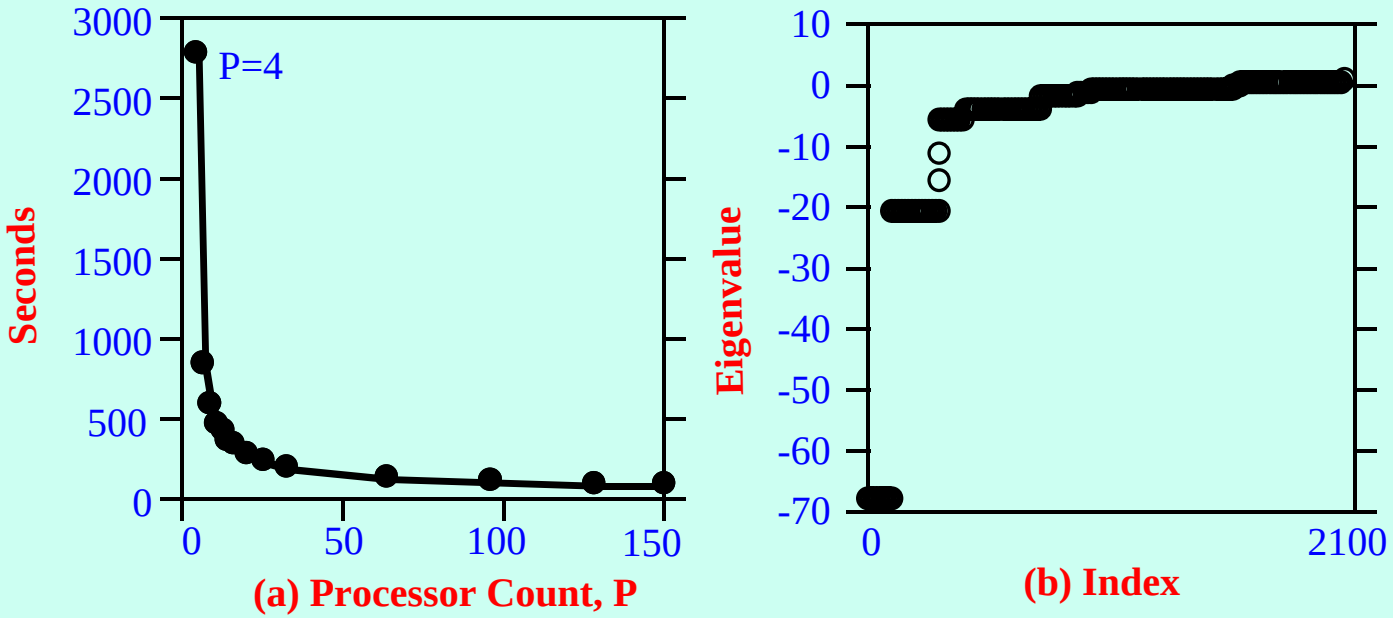
# CPU	Time / minutes
T3D (128)	10.1
T3D (256)	5.6
HPCS-1 (432)	1.0 (est.)



CLRC

Distributed Matrices & Linear Algebra - PeIGS

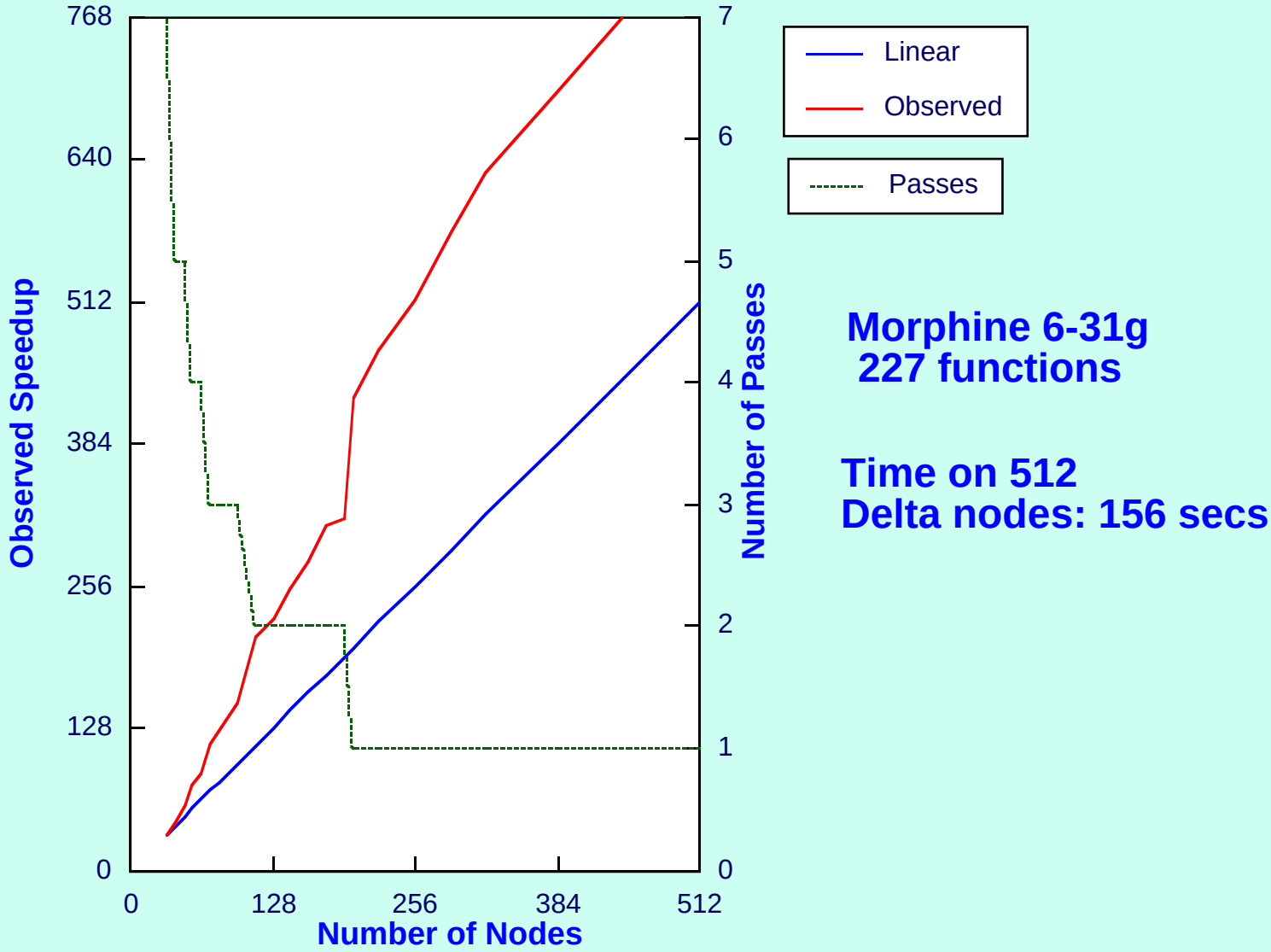
- Standard & generalized dense real symmetric eigenproblems on parallel computers.
- Produces highly accurate and orthogonal eigenvectors, using repeated inverse iteration and reorthogonalization against unconverged iterates.
- Scalable data structures and algorithms: implemented using message passing, interface to global arrays



Scalability (a) for the determination of all eigenpairs of a 2053 by 2053 matrix whose spectrum is displayed in (b).

CLRC

Parallel Performance of 4-Index Transformation



Morphine 6-31g
227 functions

Time on 512
Delta nodes: 156 secs

GAMESS-UK

GAMESS-UK is the general purpose ab initio molecular electronic structure program for performing SCF- and MCSCF-gradient calculations, together with a variety of techniques for post Hartree Fock calculations.

The program is derived from the original GAMESS code, obtained from Michel Dupuis in 1981 (then at the NRCC), and has been extensively modified and enhanced over the past decade.

This work has included contributions from numerous authors*, and has been conducted largely at the CCLRC Daresbury Laboratory, under the auspices of the UK's Collaborative Computational Project No. 1 (CCP1).

Other major sources that have assisted in the on-going development and support include academic funding agencies in the Netherlands, and ICI plc.

Current install base 100+ academic licences and 6 industrial sites.

** M.F. Guest, J.H. van Lenthe, J. Kendrick, K. Schoeffel, P. Sherwood and R.J. Harrison, with contributions from R.D. Amos, R.J. Buenker, M. Dupuis, N.C. Handy, I.H. Hillier, P.J. Knowles, V. Bonacic-Koutecky, W. von Niessen, V.R. Saunders and A. Stone*

CLRC

GAMESS-UK: Summary of Features

- RHF, ROHF, GVB, UHF and MCSCF / CASSCF Wavefunctions
- Post Hartree-Fock Methods: MP2, MP3, MRDCI, CCSD(T), FULL-CI
- Methods for ionised (Green's Functions - OVGF & TDA) and excited states (RPA, MCLR etc)
- ECP Calculations using both Non-Local and Local Pseudopotentials
- Analytic first and second derivatives for the determination of HF and MP2 force constants and geometry optimisations
- Molecular Properties: dipole, quadrupole and higher multipoles, polarisabilities, electrostatic potentials and electric field gradients etc.
- DFT Calculations (energies and gradients, coulomb fitting etc)
- Graphical Analysis of Molecular Orbitals and Properties (Cerius2 / SDK)
- Optimised for a variety of platforms including parallel architectures

Available Hardware Platforms

Cray YMP-C90, J90 and T90

NEC SX-4, Fujitsu VPP300

IBM Power, Power2 and P2SC RS/6000-based models

Hewlett Packard 9000 Series Model 7xx, J200 and C160 (PA8000)

DEC (DEC Station 5000 series and Alpha AXP EV4 and EV5 family)

Silicon Graphics (Indigo^2 R4400, R8000, R5000 & R10000 PC, Origin 2000)

Sun (SparcStation 10 and 20, HyperSPARC and Ultra-SPARC-1 and -2).

PARALLEL VERSIONS

SCF energies, 1st and 2nd derivatives, MP2 energies and gradients, and DFT.

are available for the CRAY T3D, T3E , IBM SP2, Intel iPSC/860 and workstation clusters .

New Platforms: Hitachi SR2201 (Pallas), NEC SX-4, Fujitsu VPP300

GAMESS-UK Software Developments

- **Software Development Tools**

  - Message Passing, Memory Allocation
  - Global Arrays (PNNLToolkit & Visualizer)
  - Distributed Matrices & Linear Algebra

- **GAMESS-UK & Parallel Computational Chemistry**

  -Common Application Components
  - Self Consistent Field Hartree Fock + gradients
  - SCF second derivatives
  - Direct 4-Index Transformation
  - Moller Plesset Perturbation Theory (MP2 gradients)
  - CCSD(T)
  - Density Functional Theory (DFT)

  - Molecular Dynamics (DL_POLY)
  - Combined QM /MM Theory (Chemsh)

Interactive Molecular Modelling through Parallelism: IMMP

IMMP represented a 2 year project within the European Fine Chemicals Industry to:

Develop, exploit and integrate the power of Massively Parallel Processing (MPP) technology across a wide spectrum of industrial applications in Computational Chemistry

The IMMP consortium comprises a number of academic and industrial partners. Parallel program development involves 3 major chemistry application codes judged to benefit most from current (and foreseen) developments in computer technology.

- **GAMESS-UK : Ab-initio quantum chemistry (Daresbury Laboratory);**
- **ADF - Density Functional treatment of molecules and polymers (Amsterdam);**
- **VAMP - Semi-Empirical Molecular Orbital theory for molecules (T. Clark, Erlangen)**

– Representative members of the European chemical industry.

- **ICI (Wilton, UK), Dr John Kendrick**
- **Zeneca, (Blackley, UK), Dr John Charlton**

– The Commercial Software House, Oxford Molecular Ltd (OML, Oxford).

Direct SCF Calculations

Energies and geometries of medium to large ground-state molecules at Hartree-Fock Level:

- Integrals recomputed or held in core rather than stored on disk

- Memory (scales as n^2, n=no. of basis functions), leads to limitations on workstations (approx. 2000 basis functions on 128 Mbyte machines)

- Cost of Fock Matrix build scales as n^4 (small systems) to *ca* $n^{2.5}$ asymptotically. Cost of matrix operations within SCF sheme n^3 (matrix diagonalisation, MMO etc.)

- Integral generation highly parallel (DLB); $O(n^3)$ operations potential bottleneck on MPP machines

- O(n) techniques for integral evaluation; matrix OPs the bottleneck.

Applications of GA Tools within GAMESS-UK

- **SCF and DFT**
  - Distributed Data: Random access to blocks of F/D
  - Replicated data: Caching of I/O; restart & checkpoint files
  - Linear algebra (via PeIGS)
- **SCF second derivatives**
  - Distribution of <vvoo> and <vovo> integrals
- **MP2 gradients**
  - Distribution of <vvoo> and <vovo> integrals
  - Simplification of transpositions
- **Coupled-Cluster**
  - Sharing of amplitudes and integrals

GAMESS-UK - Functionality and Benchmarking

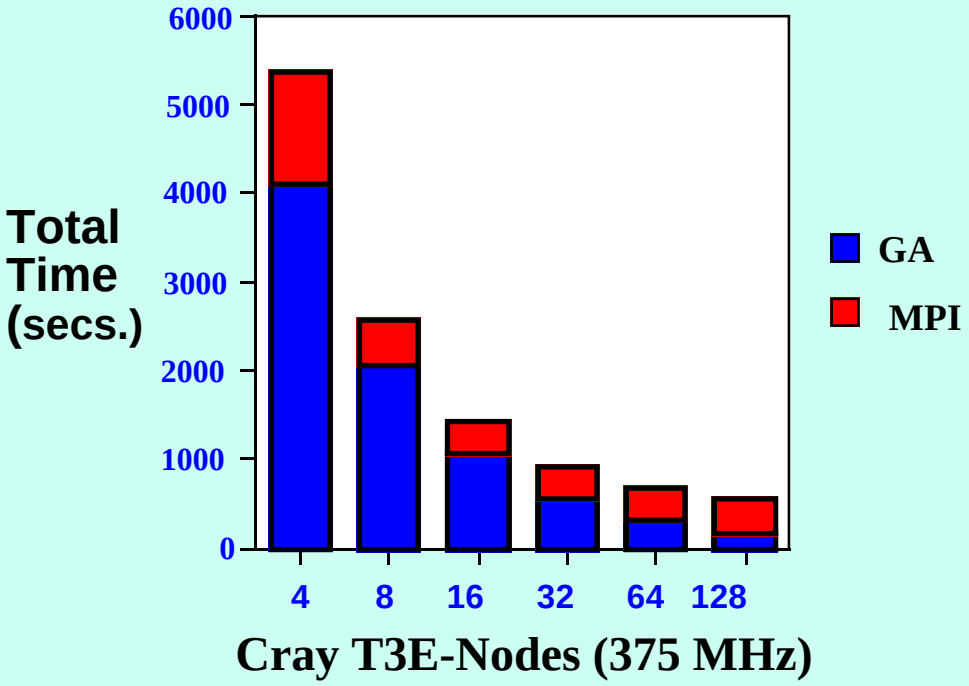
Parallel Functionality

- SCF Energy and Gradient
  - Replicated Data model
  - Dynamic Load Balancing
- MP2 Energy and Gradient
  - Global memory storage
- SCF Second Derivatives

Benchmarking Platforms

- IBM SP-2
  - 64/128 Mb node memory, local disc
  - DL (16 nodes), Cornell (512 nodes)
- SG PowerChallengeArray
  - 16 nodes share 2-4 Gb memory
  - EPCA (64 nodes)
- Parsytec GC
  - 64 Mb node memory, two PC601 procs./node
  - Paderborn (64 nodes)
- Workstation Clusters
  - HP9000 series, IBM RS6000

GA vs MPI-based Performance of GAMESS-UK



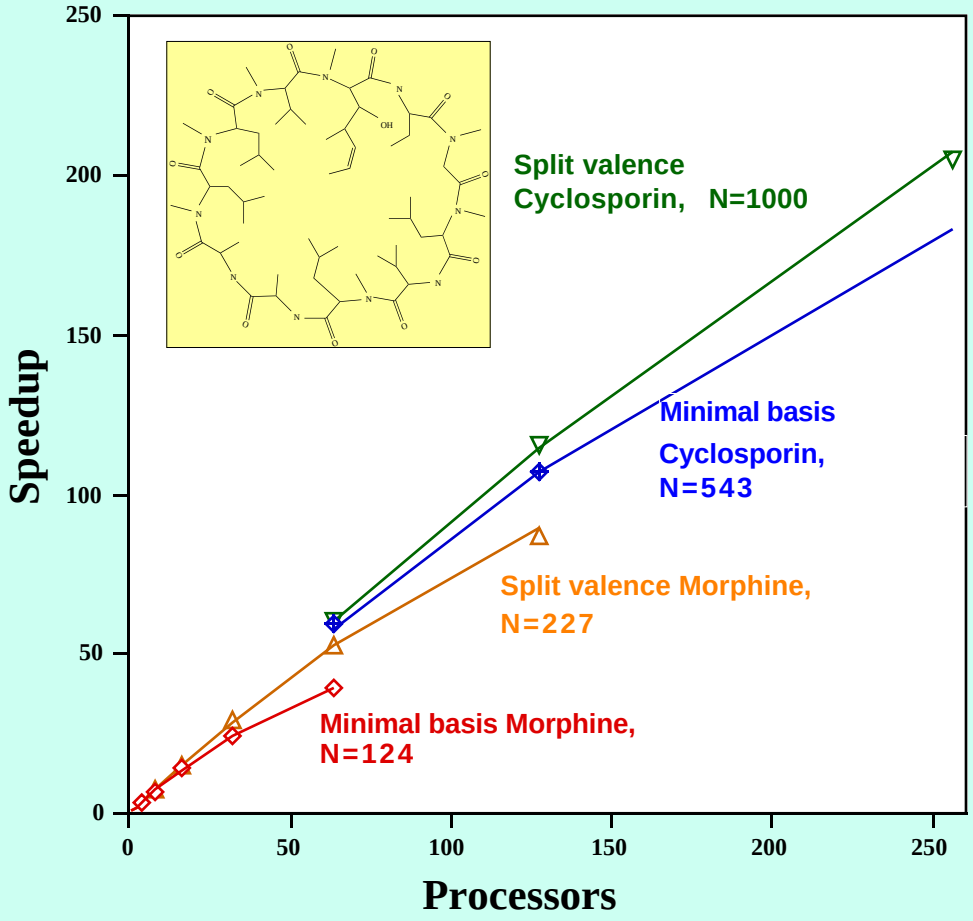
Direct-SCF 6-31G on Morphine**

410 GTOs, 9 SCF Iterations
128-node Cray T3E :
GA-code 170 secs.
MPI-code 543 secs.

Workstation Code:
Indigo2 R10000 (175 MHz)

3.73 hours

GAMESS-UK: Direct SCF Performance (Cray T3D)



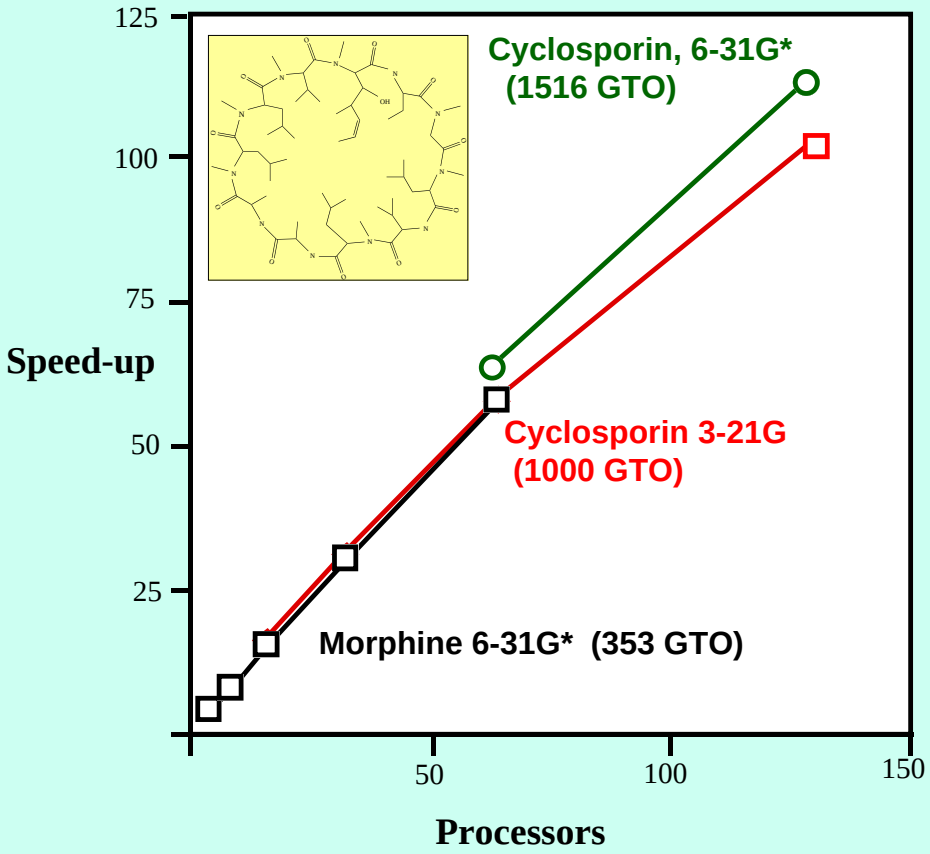
New Features:

1. Parallel Linear Algebra: PeiGS etc.
2. GA tools to cache I/O

Total Times
Cyclosporin (Split Valence)

256 node T3D	40 minutes
DEC AXP 3000/700	89 hours

GAMESS-UK: Direct SCF Performance (Cray T3E)



Cray T3E 128-128 / 375 MHz

Konrad-Zuse-Centrum fuer Informationstechnik, Berlin

New Features:

1. Parallel Linear Algebra: PeiGS etc.
2. GA tools to cache I/O

Total Times	T3E Nodes	Minutes
Morphine (6-31G*)	64	5
Cyclosporin (3-21G)	128	16
	256 node T3D	40
	(SGI R10000 / 175	19.7 hours)
Cyclosporin (6-31G*)	128	62

Large-Scale Correlated Calculations

"Compact" Systems

- Moderately sized molecule/ large basis set.
- Dense integrals and excitation amplitudes.
- Treat with RI-MP2 or more traditional approaches.

"Extended" Systems

- Large molecule/moderate basis.
- Localizable MOs, sparse integrals & amplitudes.
- Use AO-based approach to take advantage of sparsity.

MP2 Theory

- Cheapest correlated method
- Accuracy: for first row compounds optimised geometries are within a few percent of experiment

$$E = E^{SCF} + E^{MP2} \qquad E^{MP2} = \sum_{ijab} A_{ij}^{ab} (ia|jb)$$

$$A_{ij}^{ab} = \frac{2(ia|jb) - (ib|ja)}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b} \qquad (pq|rs) = \sum_{\mu\nu\lambda\sigma} C_{\mu_p} C_{\nu_q} C_{\lambda_r} C_{\sigma_s} (\mu\nu|\lambda\sigma)$$

- Full transformation costs n^5
- Partial transformation (VOVO subset) costs $n_O n^4$

MP2 Gradient Algorithms

Serial

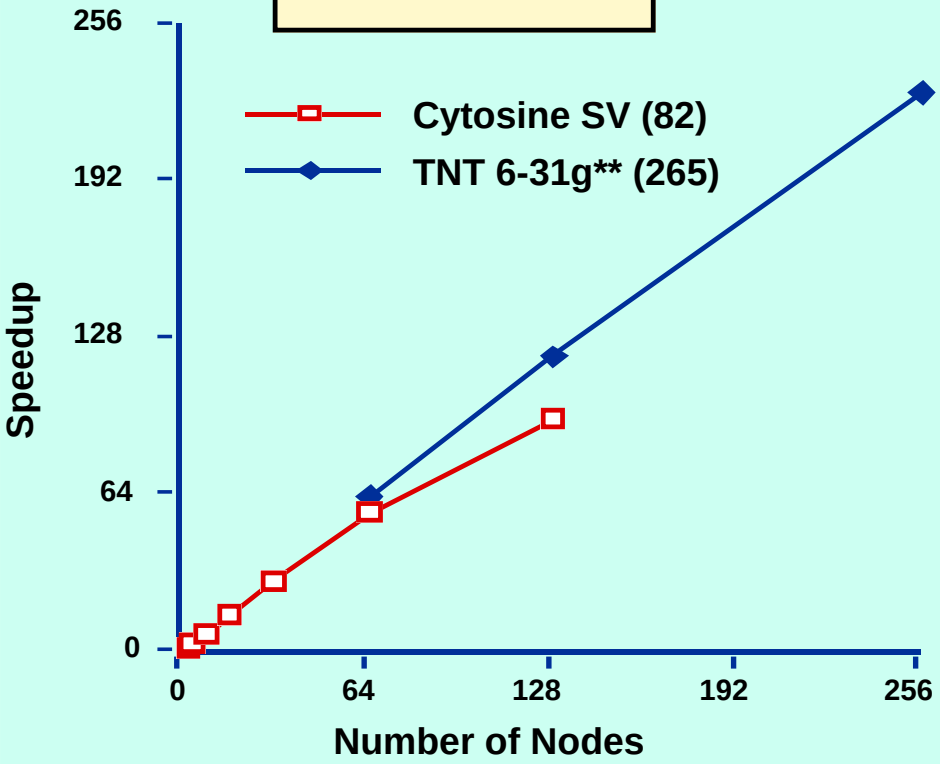
- **Conventional**
  - integrals written to disc
  - integrals read back, transformed, written out, re-sorted, and so on
  - I/O demands are heavy
- **Direct/Semi-direct (Frisch, Head-Gordon and Pople, Haase and Ahlrichs)**
  - replace all/some I/O with batched integral recomputation

Parallel

- **Poor I/O-to-compute performance of MPPs**
  - direct approach
- **Current MPPs have large global memories**
- **Store subset of the MO integrals**
  - reduce number of integral recomputations
  - increase communication overhead
- **Subset includes VOVO, VVOO, VOOO, OOOO**
  - VVVO-class too large
  - compute VVVO-terms separate step

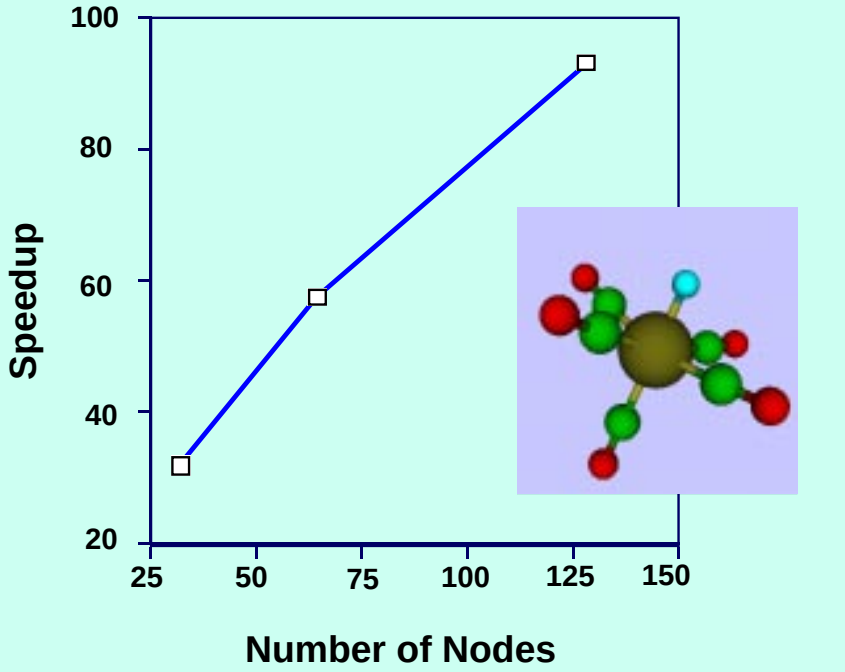
GAMESS-UK: Speedup of MP2 Gradient on the Cray-T3D and Cray-T3E

Cray T3D



Cray T3E

$Mn(CO)_5H$ - MP2 geometry opt.:
TZVP+f-basis (217 GTOs);
25 mins. on 128 nodes



GAMESS-UK: Gaussian DFT Module

- **Developed under auspices of CCP1 as modular code**

  - Variety of functionals (LDA, BLYP, B3LYP etc)

- ***Exchange Correlation Module:***

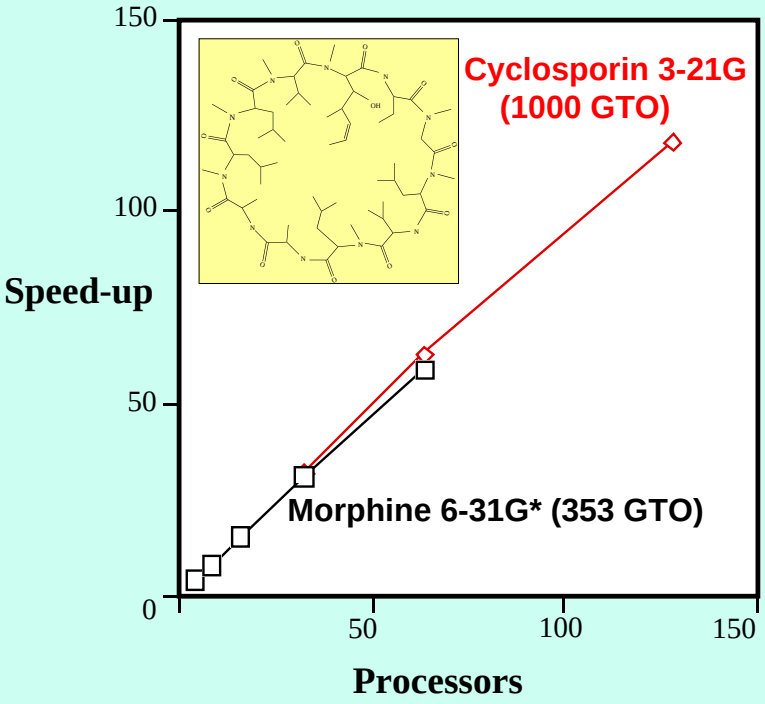
  - Numerical grid-based technology. Radial (Euler Maclaurin) and Angular parts (Gauss Legendre, Lebedev, SG1 grid etc.). Weight schemes (Becke, Frisch etc.)
  - Extensive use of screening (density matrix and points); scaling O($N^{1.5}$) in series of water clusters

- ***Coulomb Module:***

  - Dunlap auxiliary gaussian fitting method (screening on AO shells)
  - Multipole developments (stepping stone toward CMM/FMM)

    - Coulomb problem split into bi- and mono-electronic region
    - Shell-associated auxiliary gaussian used to compute shell penetration factor (code from CRYSTAL)

GAMESS-UK: DFT-BLYP Performance (Cray T3E)



Cray T3E 128-128 / 375 MHz

Konrad-Zuse-Centrum fuer Informationstechnik, Berlin

- CCP1-DFT project (1994-1997)
- full featured *replicated-data* DFT code, with focus on large molecules; modular development for incorporation into any QC package;
- developed and tested within GAMESS-UK;
- multipole code (CRYSTAL) now functional

Total Times	T3E Nodes	Minutes
Morphine	64	7
Cyclosporin (3-21G)	128	21

GAUSSIAN-94: B-LYP / DFT

IBM Power2 RS/6000 33 hours 48 minutes

GAMESS-UK: Single Processor BenchMarks

Number	Module	Basis (GTOs)	Details	Molecular Species
1.	SCF	STO-3G (124)		Morphine
2.	SCF	6-31G (154)		$C_6H_3(NO_2)_3$
3.	ECP Geometry Optimization	ECPDZ (70)		Na_7Mg^+
4.	Direct-SCF	6-31G (82)		Cytosine
5.	CASSCF Geometry Opt.	TZVP (52)	480 csf	H_2CO
6.	MCSCF	(5s3p2d/3s1p/f(O) (74)	5608	H_2CO
7.	Direct-CI	(5s3p2d/3s1p) (64)	3M/167194	H_2CO/H_2+CO
8.	Table-CI (26M/6R)	ECP (59)	2301815/4097	$TiCl_4$
9.	MP2 Geometry Optimization	6-31G* (70)		H_3SiNCO
10.	SCF Second Derivatives	6-31G (64)		C_5H_5N
11.	MP2 Second Derivatives	6-31G* (60)		C_4
12.	Direct-MP2	DZ+D(N) (76)		C_5H_5N

The GAMESS-UK Benchmark: Total CPU time (user and system) Elapsed time (minutes) and Efficiency (%) for Calculations 1 - 12

Machine	CPU Time User	CPU Time System	Elapsed Time	Efficiency (%)
SGI Origin 2000/195	20.1	3.1	27.3	85%
IBM RS/6000-P2SC/135	23.2	1.9	32.9	76%
DEC Alpha 600/5/333	23.4	7.6	34.4	91%
SGI PowerOnyx 10000/194	23.1	3.6	36.7	73%
HP 9000/C160-160	27.7	3.7	45.9	68%
SGI Indigo2 R10000/175	26.0	4.8	70.7(*)	44%
Sun Ultra-2/200	31.5	3.0	38.1	91%
DEC Alpha 2100/5/250	30.2	8.2	43.6	88%
SGI R8000 Power Onyx	48.1	5.6	55.2	97%
Sun Ultra-1/170	40.5	4.9	56.3	81%
IBM RS/6000-3CT	40.1	4.4	58.3	76%
SGI Indigo2 R8000	50.8	7.2	60.6	96%
IBM RS/6000-590	46.6	4.7	74.8	68%
DEC Alpha 600/5/266	30.6	6.4	62.2	59%
DEC Alpha 250/4/266	46.0	9.7	85.7	65%
DEC Alpha 8400/5/300	26.9	10.4	92.8*	40%
HP PA/9000-735/125	51.6	8.4	93.3	64%
HP PA/9000-755	64.9	8.8	98.2	75%
SGI Indy R5000	72.9	11.3	101.0	83%
DEC AXP/3000-700	53.4	11.5	105.4	62%

Machine	CPU Time User	CPU Time System	Elapsed Time	Efficiency (%)
DEC AXP/3000-600	72.1	14.1	107.8	80%
IBM RS/6000-370	78.7	8.4	108.8	80%
SGI Challenge L/150	93.3	10.3	110.6	94%
SGI R4400 Indigo2	94.3	10.0	120.6	86%
HP 9000/715-100	81.4	14.3	124.7	77%
Sun Ultra-1/140	45.3	5.7	139.8*	36%
HP PA/9000-715/80	93.7	15.9	140.1	78%
DEC AXP/3000/500	93.0	40.1	148.6	90%
HP PA/9000-750	111.1	17.4	140.2	92%
IBM RS/6000-550	117.6	11.2	146.4	88%
SGI Challenge L/100	140.8	17.1	178.0	89%
DEC AXP/3000-300	129.8	39.6	184.0	92%
Sun SS-20/HS21	82.5	16.2	202.7	49%
SGI R4000 Indigo	153.2	17.7	206.0	83%
Sun SPARCserver 1000	175.1	50.8	445.6	51%
Sun SPARCstation 10/41	201.4	21.0	260.4	85%
Sun SPARCstation 5/85	240.5	25.8	299.0	89%
Sun SPARCstation 10/30	251.4	31.7	306.3	92%
CONVEX C-3860	124.3	3.5	210.9	61%
Cray Y-MP/8128	56.8	6.0	188.4	33%
Cray YMP-C98/4256	37.9	2.0	41.2	97%

Performance of PC's in Computational Chemistry: GAMESS-UK Benchmarks

Machines Under Evaluation - H/W and O/S

Pentium 133 (FreeBSD)

Intel Pentium Motherboard "TUCSON" TC430HX
Intel Pentium 133Mhz
512 Kb L2 cache
64Mb EDO RAM
Adaptec 2940 (SCSI-2 Narrow) SCSI Adapter
IBM DORS-32160 2Gb disk
FreeBSD 2.2.2
f2c + pgcc 2.7.2p
Fortran Compiler flags: -O6 -frisc -mno-fp-ret-in-387

Pentium-Pro (FreeBSD)

Intel Pentium Pro MotherBoard "VENUS" VS440FX
Intel PentiumPro 200Mhz
256Kb Internal L2 cache
64Mb EDO RAM
Adaptec 2940 (SCSI-2 Narrow) SCSI Adapter
IBM DORS-32160 2Gb disk
FreeBSD 2.2.1
f2c + pgcc 2.7.2p
Fortran Compiler flags: -O6 -frisc -mno-fp-ret-in-3872

Pentium 166 (FreeBSD)

Amptron Pentium Motherboard
Intel Pentium 166Mhz
512 Kb L2 cache
64Mb RAM
Adaptec 2940 (SCSI-2 Narrow) SCSI Adapter
IBM DORS-32160 2Gb disk

FreeBSD 2.1.5
f2c + pgcc 2.7.2p
Fortran Compiler flags: -O6 -frisc -mno-fp-ret-in-387

Pentium Pro (Linux and Windows NT)

Intel Pentium Pro MotherBoard "VENUS" VS440FX
Intel PentiumPro 200Mhz
256Kb Internal L2 cache
128Mb EDO RAM
NCR (SCSI-2 Narrow) SCSI Adapter
SEAGATE Hawk ST15230N 4Gb disk

Linux RedHat 4.1
f2c + gcc 2.7.2.1
Fortran Compiler flags: -O

Windows NT 4.0
DIGITAL Visual Fortran (Beta Release)
Microsoft Visual C++ 4.2
Fortran Compiler flags: -optimize:5 -fpe:3 -debug:none

Performance of PC's in Computational Chemistry: GAMESS-UK Implementation Notes

- GAMESS-UK uses fortran functions that return integers packed in a 64 bit double precision numbers; this generates floating point exceptions on the Intel Pentium processors.
- This problem is present under all operating systems used: FreeBSD, Linux and Windows NT 4.0. These exceptions are probably wasting some CPU time under each of the environments.
- Some benefit found under FreeBSD from using the Pentium GNU cc compiler and an assembler coded math library that directly uses the Pentium floating point instructions to calculate sqrt, log, exp., trig. functions etc. (10% improvement in CPU time). This library is part of the operating system.
- The different timings between Linux and FreeBSD suggests that Linux is probably lacking the assembler code optimized math library.
- Performance difference between FreeBSD and WinNT PentiumPro probably arises from the better compiler on the latter. Note that pgcc has a "-mcpu=pentiumpro" option that may be of help in further optimizing on PentiumPro processors. We have still to test this option.

Performance of PC's in Computational Chemistry: GAMESS-UK Benchmark

Direct-SCF calculation on Cytosine (6-31G basis, 82 CGTOs). All the timings are obtained as an average of 3 runs from the csh time.

Machine	CPU Time (sec) user	system	Elapsed Time (sec)	Ratios vs AXP 600 / 333 CPU	Elapsed
Pentium 133 (FreeBSD)	537.4	0.8	540.2	6.23	6.21
Pentium 166 (FreeBSD)	463.1	0.7	472.5	5.37	5.43
Pentium Pro (Linux)	402.7	1.7	408.1	4.68	4.69
Pentium Pro (FreeBSD)	349.7	0.4	352.1	4.05	4.05
Pentium Pro (WinNT)			309.		3.56
Pentium Pro (WinNT+ Intel Reference Fortran)			*150*		*2.0*
IBM RS-6000/370	332.0	0.0	336.0	3.84	3.86
SGI R4400/250	238.4	0.8	241.8	2.77	2.78
IBM RS-6000.3CT	195.8	0.3	198.0	2.27	2.28
SUN UltraI/140	178.0	0.0	180.0	2.06	2.07
SUN Ultra2/200	128.0	0.0	129.0	1.48	1.48
DEC AXP 5-8400	100.6	0.9	101.4	1.17	1.17
IBM-P2SC/135	98.7	0.2	100.2	1.14	1.15
SGI R10000/175	93.8	0.6	191.4	1.09	2.20
HP PA8000/-C160	95.4	0.4	96.0	1.11	1.10
DEC AXP600/333	86.0	0.4	87.0	1.00	1.00
SGI R10000/194	83.2	0.4	84.0	0.97	0.97
SGI Origin-2000	83.9	0.3	85.2	0.97	0.98

Performance of PC's in Computational Chemistry: Machine Costs

DEC AXP 500 / 333		Pentium Pro (200 Mhz)	
Ethernet card	65	Ethernet card	65
Alpha station 500/333	3,303	PPRO (512k cache)	1,270
256 MByte RAM	2,182	256 MByte RAM	800
9 GByte Disk	1,100	9 GByte disk	810
OSF1/OS	-	NT/OS	-
Monitor	540	Monitor	540
CD-ROM	155	CD-ROM	155
Total (UK pounds)	7,345		3,640

Approximate Schemes for the Coulomb Potential

Conventional 2-electron integral scheme:

$$F_{\mu\nu} = F_{\mu\nu} + P_{\lambda\sigma} \langle \chi_\mu \chi_\nu \mid 1/r \mid \chi_\lambda \chi_\sigma \rangle$$

- **Formal scaling is $O(N^4)$ for small systems but $O(N^2)$ for large systems, if summation is restricted to $\chi_\mu\chi_\nu$ and $\chi_\lambda\chi_\sigma$ which satisfy an overlap test**

$$\langle \chi_\mu | \chi_\nu \rangle \quad > \quad \text{tol}$$

In practice screening tests are performed at the shell level by replacing χ_μ by an s function with the smallest exponent in the shell.

- **Approximate schemes are based on alternative representations of the density $P_{\lambda\sigma}\chi_\lambda\chi_\sigma$**

Fitted scheme: **Represent total density $\sum\sum P_{\lambda\sigma}\chi_\lambda\chi_\sigma$ in an auxillary basis**

Multipole scheme: **Partition $P_{\lambda\sigma}\chi_\lambda\chi_\sigma$ based on overlap of χ_λ with $\chi_\mu\chi_\nu$ and expand non-overlapping part as a multipolar expansion.**

Coulomb Fitting Scheme

- Expand total density using an auxilliary basis

$$\sum\sum P_{\lambda\sigma}\chi_\lambda\chi_\sigma \approx \sum c_r\chi_r$$

- Typically, $N_r \approx 3*N$. Fitting Coefficients determined variationally by Dunlap scheme,

$$c_q = \sum_r V_{qr}^{-1}\left\{\sum P_{\nu\mu}(\chi_\mu\chi_\nu||\chi_r) - \alpha\int\chi(r)\,dr\right\}$$

- Requires inversion of two-centre coulomb integrals of order N_r, $O(N^3)$

$$V_{qr} = (q||r)$$

- Coulomb matrix built using $O(N^2)$ 3-centre coulomb integrals

$$J_{\mu\nu} = \sum_r c_r(\chi_r||\chi_\mu\chi_\nu)$$

Multipole Approximations

- Determine shells (c) of basis functions χ_λ that do not penetrate overlap $\chi_\mu\chi_\nu$
- Construct multipole moments of the overlap distributions $\chi_\lambda\chi_\sigma$ for all basis functions σ about the site of shell c

$$\int \chi_\lambda \chi_\sigma X_{lm}(r-r_c)\,dr$$

- Contract with density matrix to form multipoles of the Mulliken charge distribution associated with shell c

$$\eta(\rho_c, r_c) = \sum\sum P_{\lambda\sigma} \int \chi_\lambda \chi_\sigma X_{lm}(r-r_c)\,dr$$

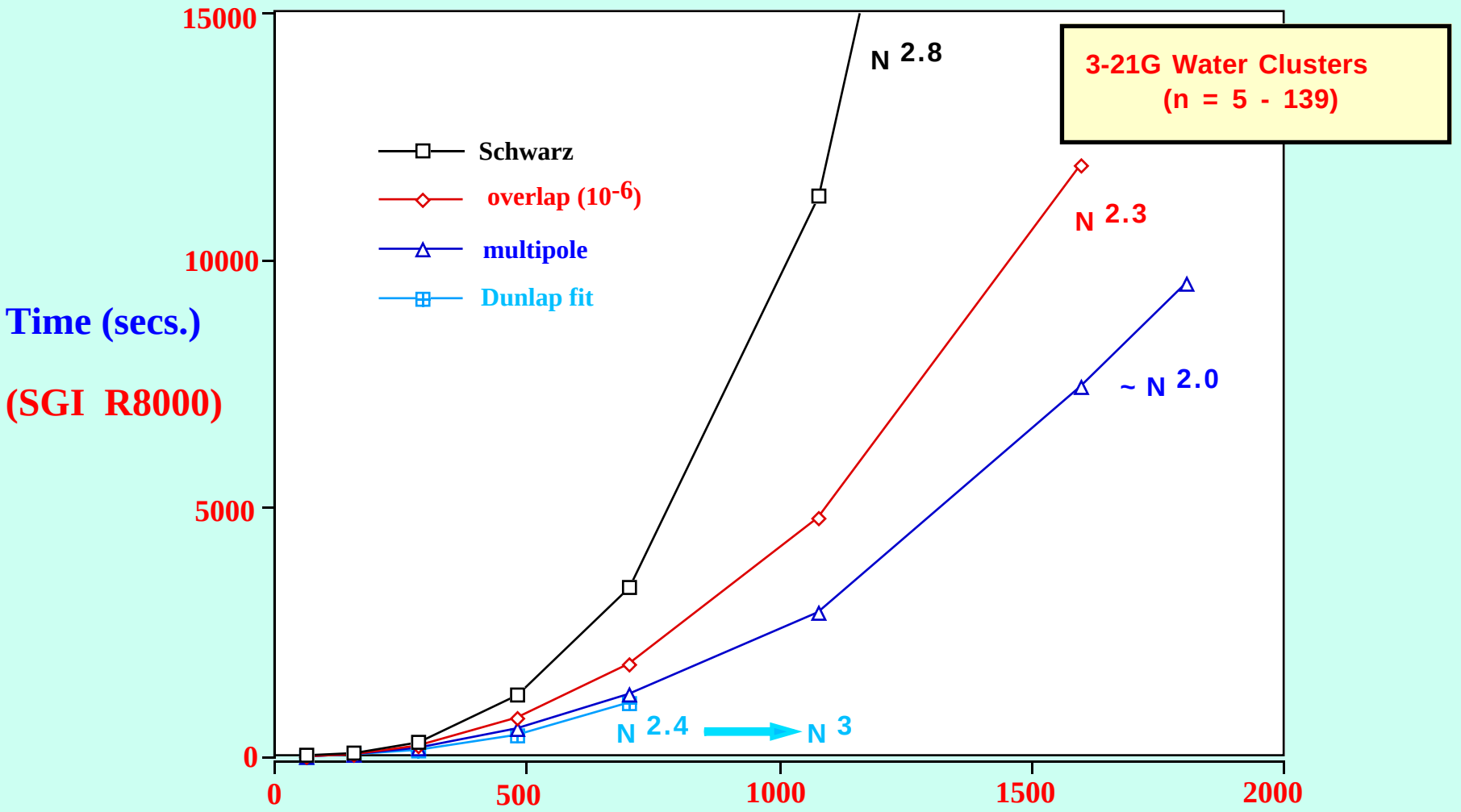
- Combined with the multipole field integrals

$$M_{lm\mu\nu} = \int \chi_\mu \chi_\nu Z_{lm}(r-r_c)\,dr$$

to generate a contribution to the fock matrix

$$J_{\mu\nu} = \sum\sum \eta(\rho_c, r_c)\, M_{lm\mu\nu}$$

Coulomb Evaluation Scaling



Coulomb Approximations - Summary

- **Fitting schemes - Dunlap scheme cost effective for "small" basis sets, eventually dominated by matrix inversion.**
- **Multipole Schemes for n > 1000 basis functions.**
- **Two-electron integrals formally O(N), but require efficient implementation of screening tests O(N^2). (currently dominant after *ca* 1200 basis functions).**
- **Multipole integrals are O(N^2) - optimise by contracting multipoles from shells on a given centre:**

  heirarchical schemes (CMM, FMM)

SUMMARY

- **Background issues in HPC and computational chemistry.**

- **Provided an outline of Grand Challenge and Throughput Requirements:**

NWChem and Grand Challenge Applications:

- **Full-featured electronic structure & MD code**
- **Distributed data Parallel implementations**

GAMESS-UK and Throughput Applications

- **Present capabilities - scalability within a replicated data approach (T3D / T3E performance figures)**
- **SCF, MP2-gradients and DFT performance on T3E.**
- **Integral Bottleneck and O(N) developments - Multipole approx.**

- **Cost-effectiveness of MPPs**

- **Unix workstations and PC's**
- **Balance lies in favour of GC applications, and not throughput ...**