Scaling the hardware zoo:

user-side timings of common quantum chemistry routines in *Gaussian 16*

Stepas Toliautas | Vilnius University, Faculty of Physics | *stepas.toliautas@ff.vu.lt*

Hardware: VU HPC, Saulėtekis location[1]

Vendor: Bull / Atos | *All node types contain 2 CPUs with hardware threads enabled

CPU type	Base freq., GHz	Released/ deployed	Cores / HT*	RAM, GB	Nodes
Intel Xeon X5650	2,66	2010 / 2011	**12** (24)	96	80
Intel Xeon E5-2670	2,60	2012 / 2014	**16** (32)	128	54
Intel Xeon Gold 6130	2,10	2017 / 2020	**32** (64)	384	141

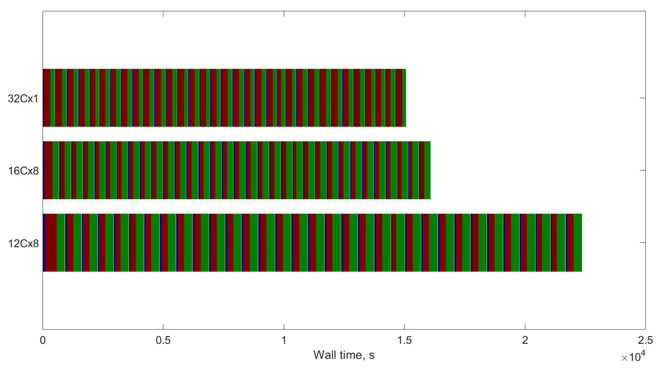
Software: *Gaussian 16*[2]

Shared-memory (within node) and **cluster** (between nodes) parallel
- Each node type has CPU-optimized binaries
- Single-node jobs use separate executables

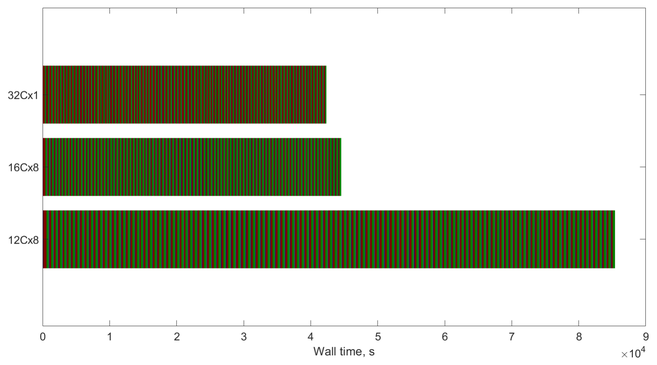
Direct comparison: scanning the same **ES** path

L502 | **L914** | **L1002**

First 500 links – 15,0k / 16,0k / 22,4k wall seconds



All (1400) links – 42,3k / 44,5k / 62,7k (85,4k*) wall seconds



* Because of differences in numerical stability, 12-core job took 21 additional optimization steps (27 → 58) on a single (complicated) point of **ES** path, which accounted for 32 % extra work (66 → 87 steps)

Timings and strong scaling

12C | **16C** | **32C**

Top: **ES** optimization step
- High spread of values, even after re-scaling by system size
- Acceptable scaling (25–50 % efficiency through 8 nodes)
- 4–8 12C/16C nodes approximate 1 32C node in computation time

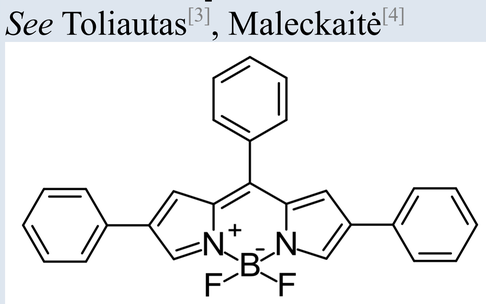
Bottom: **GS** vibrational frequencies
- L703 does both Hessian update (fast) and full derivatives (expensive)
- L1002 (CPHF) scales poorly with smaller (double-zeta) basis sets
- 32C node looks superior, even compared to older 8-node clusters

Outlook

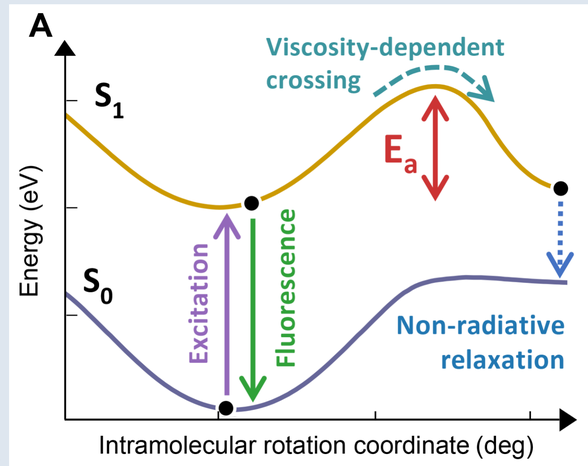
Legacy nodes account for:
- 48 % of currently available resources
- 10 % computing capacity for *complex* quantum chemistry jobs
- 20 % capacity for easy/ short/ exploratory computations
- …still enabled by good software flexibility/ support!

Scientific problem

See Toliautas[3], Maleckaitė[4]



BP-PH



Solution steps
- Optimize ground-state (**GS**, S_0) structure
- Optimize excited-state (**ES**, S_1) structures
- Scan adiabatic **ES** path (by multiple opts.)

Main routines (*Gaussian* links)

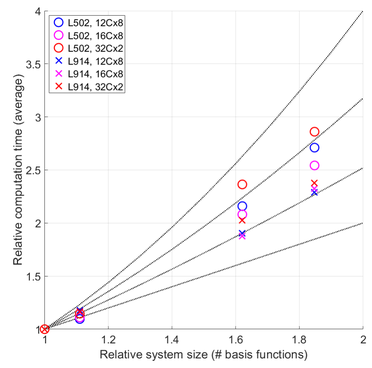
L502	DFT self-consistent field (SCF) energy	*8–15 iters./step*
L701, L1101	1-electron integral derivatives	
L703, L1110	2-electron integral derivatives	
L914	TD-DFT excited-state energies	*9–11 iters./step*
L1002	Coupled-perturbed SCF (vibrations / displacements)	*8 / 24 iters/step*

Job types
- **GS** optimization + vib: L(502, *701, 703*)×**N**, L(1101, 1110, 1002, 701, 703)×**2**
  - Single-point **ES**: L(502, 914, 1002)
- **ES** optimization/ scanning: L(502, 914, 1002*, 701, 703*)×**N**

System size scaling

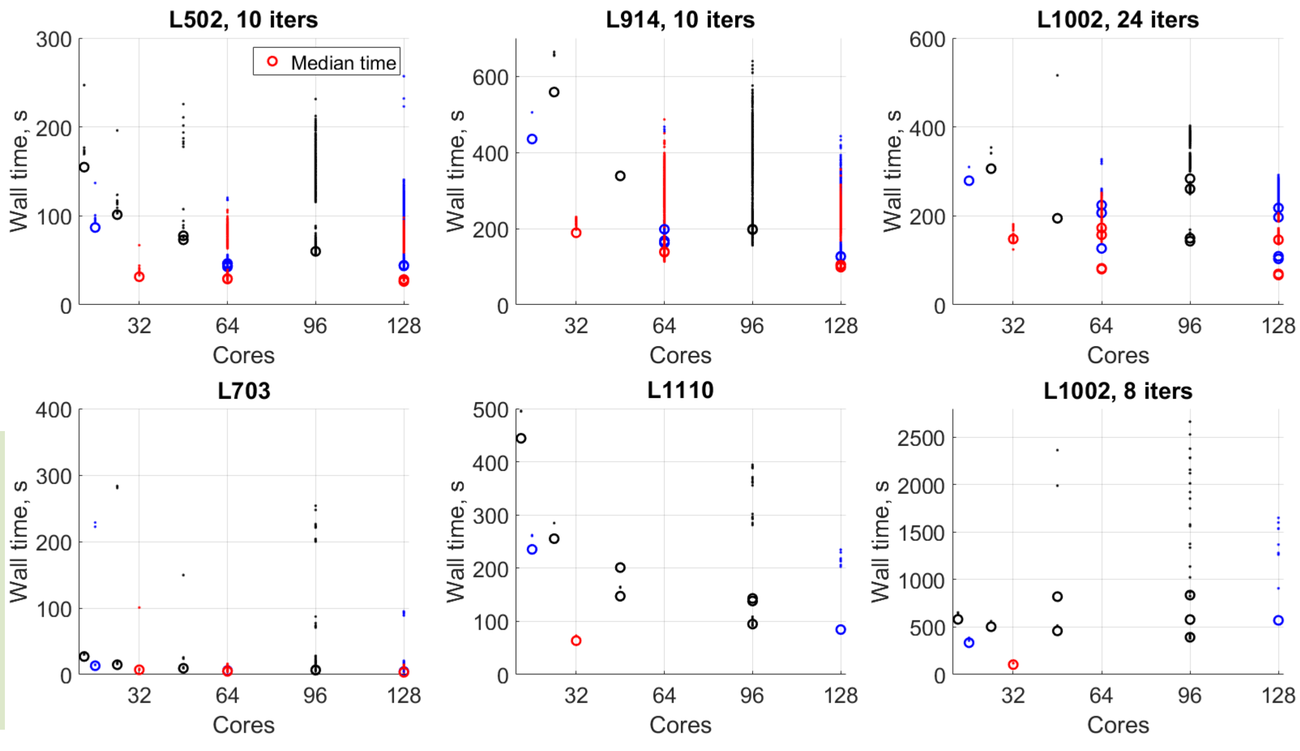
Computational costs strongly depend on basis set size ***M***
- Theory[5]: **GS** → M^2–M^3, **ES** → M^4–M^{10}

Practice: limit not reached for small/ medium systems
- 335–619 reference basis functions
- 10^1–10^3 s / optimization step
- Power factor M^K, K = 1,3–1,7



Job set: 1404 core-days, 68 days of run-time

# BF	12Cx1	16Cx1	32Cx1	12Cx2	32Cx2	12Cx4	16Cx4	32Cx4	12Cx8	16Cx8
335	0	0	0	0	9	0	2	1	16	2
372	1	2	3	4	5	1	2	0	7	5
543	0	0	0	0	5	0	0	1	13	11
619	0	0	0	0	5	2	3	15	17	18



References

[1] *sauletekis.ff.vu.lt/en/info/kontaktai/*
[2] *Gaussian* 16, Rev. **C.01**, M. J. Frisch et al., Gaussian, Inc., Wallingford CT (2016)
[3] S. Toliautas *et al.*, *Chem. Eur. J.* **25**, 10342–10349 (2019)
[4] K. Maleckaitė *et al.*, *Molecules* **27**, 23:1–14 (2022)
[5] P.-F. Loos *et al.*, *J. Phys. Chem. Lett.* **11**, 2374–2383 (2020)