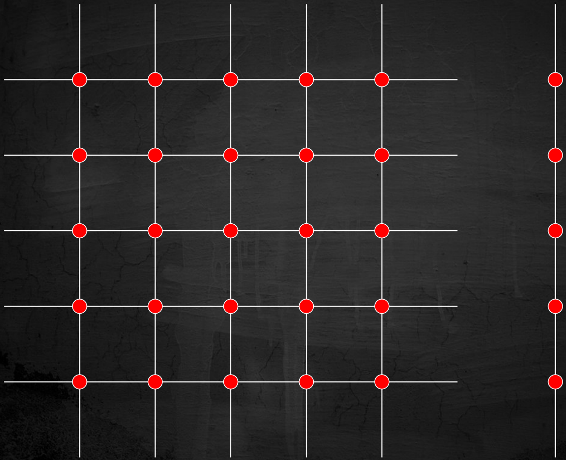
Approximating a Multi-Grid Solver

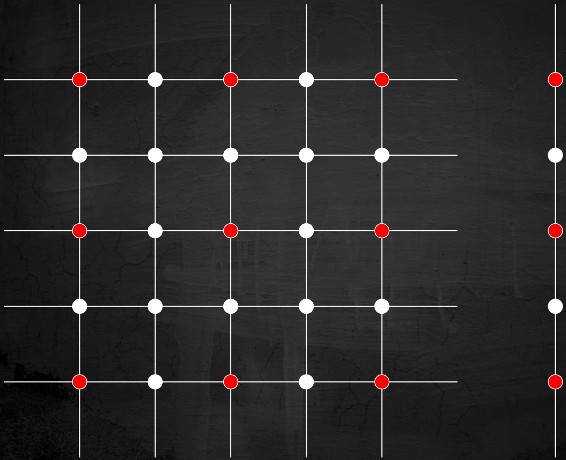
Valentin Le Fevre,
Leonardo A. BAUTISTA GOMEZ, Osman Unsal, Marc Casas

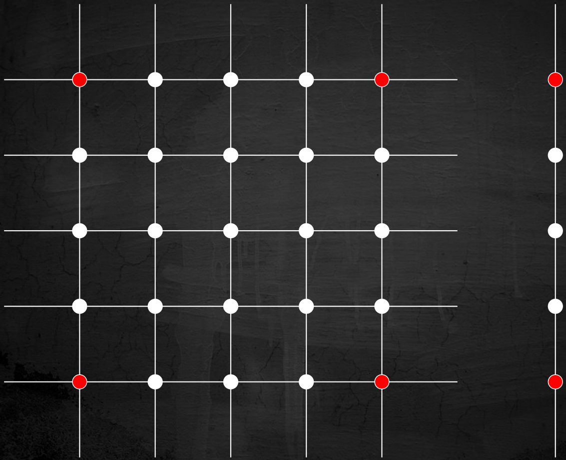
ENS Lyon, Barcelona Supercomputing Center (BSC)

November 12, 2018

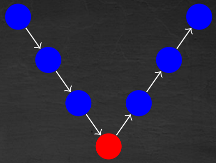
Introduction







4096 points
1024 points
256 points
64 points



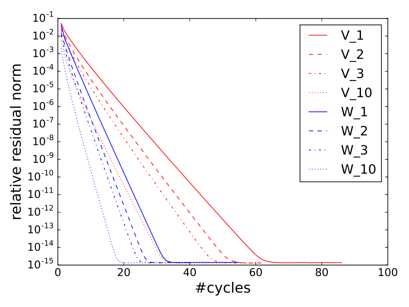
- Different level of coarseness.
- Faster than classical methods.
- Accuracy is limited by the hardware.

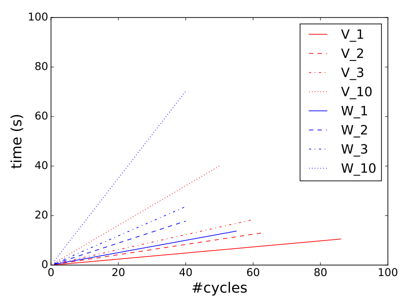
Trade-off between **precision** and **performance**.

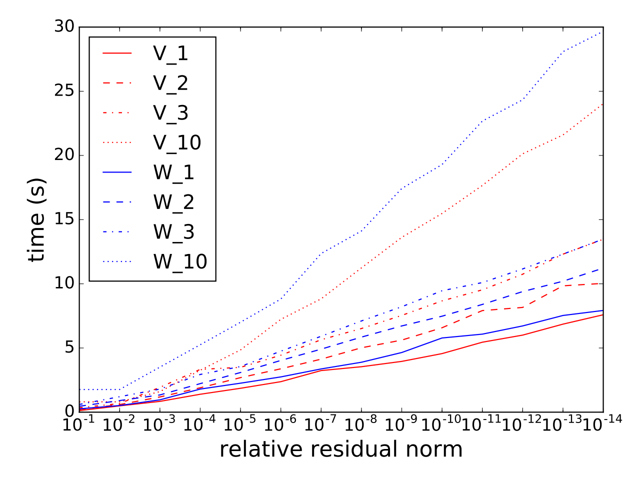
- No exact result exists (search queries)
- Faulty hardware (Fast adders)
- Memory without ECC
- Branching to avoid useless computations
- Skip steps in loops
- Precision of a floating-point value

Approximating the Algorithm

- Add more iterations at each level.
- Add more complex cycles.





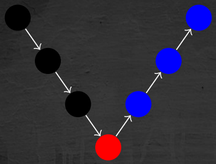

V_1
V_2
V_3
V_10
W_1
W_2
W_3
W_10
time (s)
relative residual norm
0
5
10
15
20
25
30
10^{-1} 10^{-2} 10^{-3} 10^{-4} 10^{-5} 10^{-6} 10^{-7} 10^{-8} 10^{-9} 10^{-10} 10^{-11} 10^{-12} 10^{-13} 10^{-14}

Level	Matrix size	Non-zero elements	Relax (down)	Relax (up)	Restriction	Interpolation
1	512,000	4,042,520	20 ms	20 ms	15 ms	-
2	256,000	6,475,239	20 ms	25 ms	12 ms	4 ms
3	58,893	2,000,513	8 ms	8 ms	3 ms	2 ms
4	14,285	788,509	2 ms	2 ms	1 ms	0.7 ms
5	4,238	386,333	1 ms	1 ms	0.5 ms	0.2 ms
6	609	53,493	$<$ 0.1 ms	$<$ 0.1 ms	$<$ 0.1 ms	$<$ 0.1 ms
7	69	2,873	$<$ 0.1 ms	$<$ 0.1 ms	$<$ 0.1 ms	$<$ 0.1 ms
8	2	4	$<$ 0.1 ms	-	-	$<$ 0.1 ms

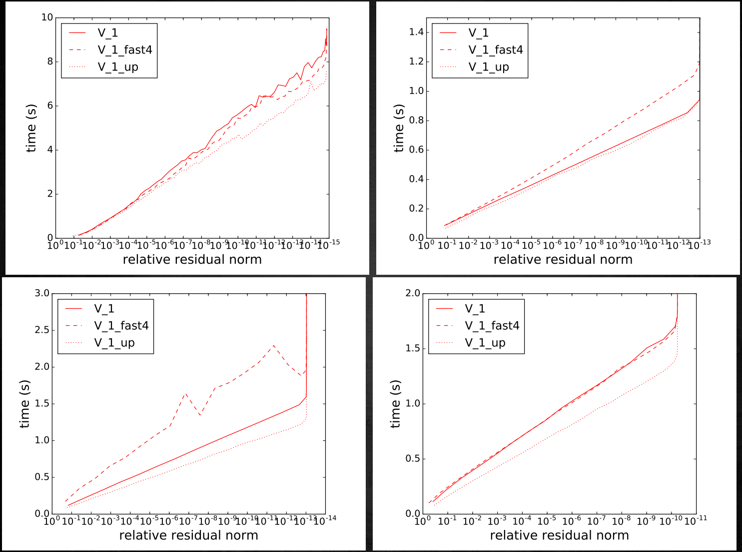
Table: Time breakdown of a V-cycle with α = 1.

$\Rightarrow$ Relaxations represent $\approx$ 66% of the total cost of a V-cycle.

Relaxations only when going up in the V-cycle.

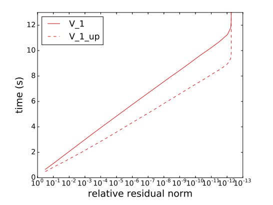


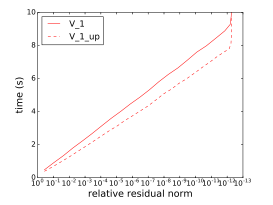
- Blue: relaxation.
- Red: exact solve.
- Black: nothing.

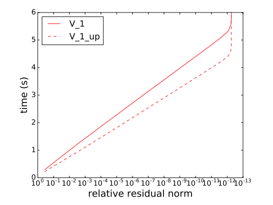


*a) Unstructured-anisotropy *b) 3D Laplace 7-pt (5,5,5) *c) 3D Laplace 27-pt *d) PDE Dirichlet

a) 3x3x3 (27), b) 6x6x1 (36), c) 4x4x4 (64)



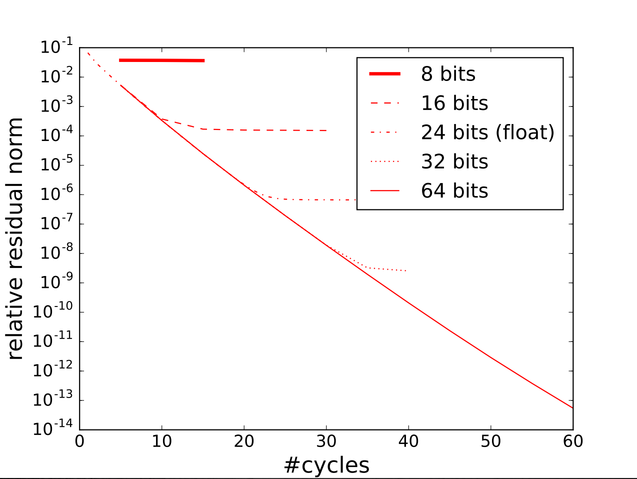




Between **7% and 28%** of improvement to reach max accuracy.

Approximating the Data

- Reducing Precision increases performance and energy-efficiency.
- Multiple-precision floating-point computations (MPFR).
- Rewrite MG algorithm with MPFR (arbitrary precision).
- Hardware limitations for performance and energy measurements.
- Is it possible to reach the same accuracy with less precision?

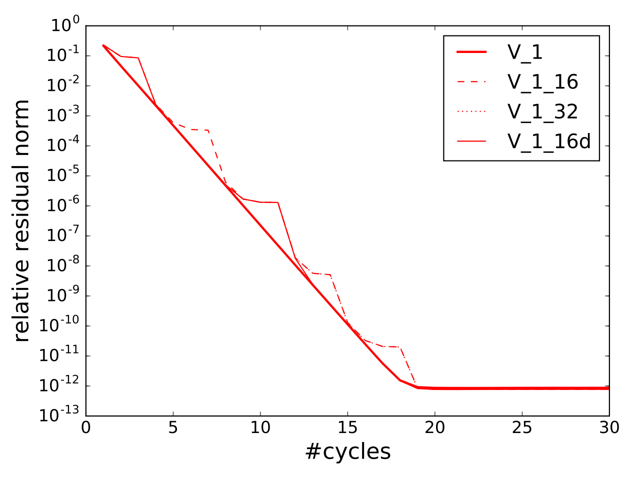


Accuracy reached with different precisions

- $b \leftarrow 64$.
- While $\texttt{nb\_iters} < \texttt{max\_iter}$ **and** $\texttt{rel\_res\_norm} > \texttt{tolerance}$
  1. Do a cycle at precision b.
  2. Compute new_rel_res_norm.
  3. rel_res_norm $\leftarrow$ new_rel_res_norm.
  4. nb_iters $\leftarrow$ nb_iters+1.

Define t a threshold to update precision b.

- $b \leftarrow 16$.
- While `nb_iters` $<$ `max_iter` **and** `rel_res_norm` $>$ `tolerance`
    1. Do a cycle at precision b.
    2. Compute new_rel_res_norm.
    3. **If *new_rel_res_norm* $>$ $t \times$ *rel_res_norm* Then** $b \leftarrow$ **UPDATE**(b)**.**
    4. rel_res_norm $\leftarrow$ new_rel_res_norm.
    5. nb_iters $\leftarrow$ nb_iters+1.



Adaptive precision with a precision threshold of 0.8.

- MPFR is 20 to 30 times slower
- How to estimate performance benefits?

- MPFR is 20 to 30 times slower
- How to estimate performance benefits?

$$Time(n, b) = a \cdot n^3 \cdot b^{\alpha} + c$$

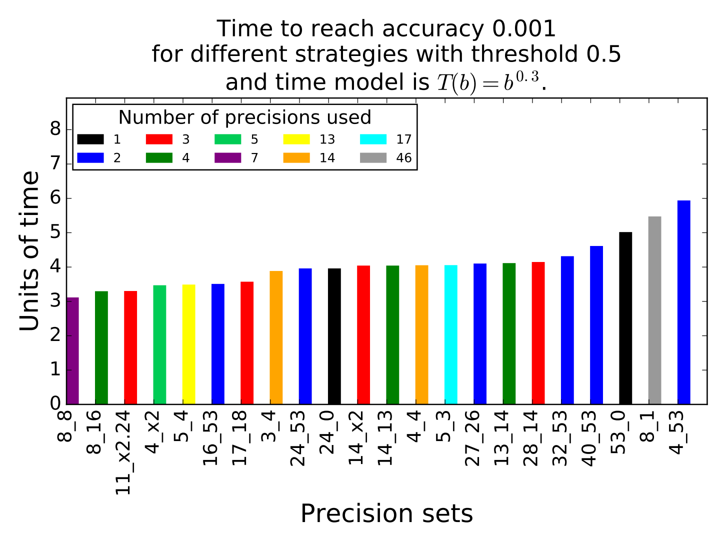
- n: size of the problem (3D grids).
- b: mantissa precision in bits.
- a, α, c: constants to determine.

- MPFR is 20 to 30 times slower
- How to estimate performance benefits?

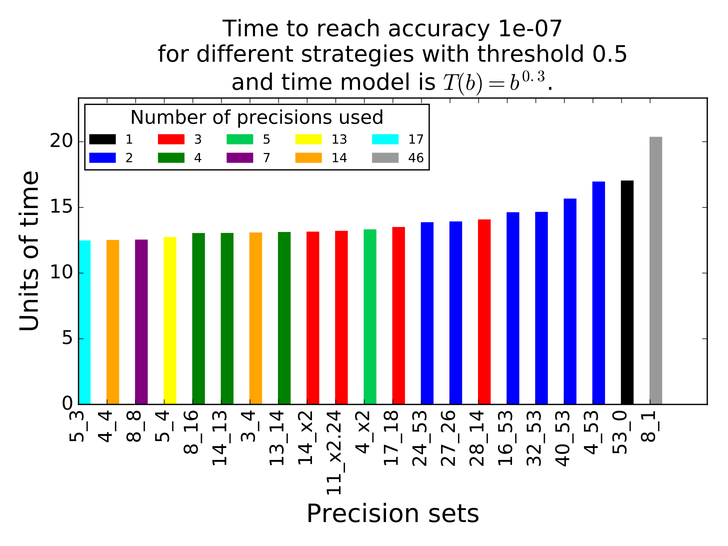
$$Time(n, b) = a \cdot n^3 \cdot b^{\alpha} + c$$

- n: size of the problem (3D grids).
- b: mantissa precision in bits.
- a, α, c: constants to determine.

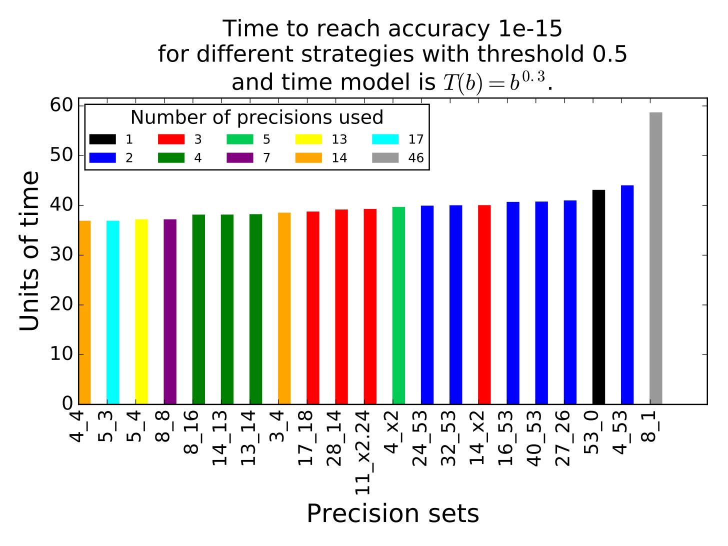
Find constant values using measurements and interpolation.



Compared to Double-precision: **34%** improvement.



Compared to Double-precision: **23%** improvement.



Compared to Double-precision: **9%** improvement.

Conclusions

- A faster cycle shape: the Up-cycle.
- A new adaptive precision algorithm for any MG solver.
- Up to 30% expected improvement on multi precision systems.
- Over 15% performance improvement with same accuracy.

- Change precision inside a cycle.
- Model (or measure) the gains in energy consumption.
- Link to silent data corruption or fussy logic.

Thank you for your attention.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 708566 (DURO). The European Commission is not liable for any use that might be made of the information contained therein. This work has been supported by the Spanish Government (Severo Ochoa grant SEV2015-0493)