FLATIRON INSTITUTE
Scientific Computing Core

# BENCHMARKING GROMACS ON OOKAMI

## Ookami User Group meeting

**Géraud Krawezik**
02/10/2022

# The Benchmark

- GROMACS
  - Molecular Dynamics code written in MPI + OpenMP
  - Real life, mature scientific tool, with ARM optimizations (SVE)
  - Does not need a lot of memory per core
  - Already have results on x86_64 clusters, and GPUs
- Benchmark executed through MDBenchmark
  - 15 minutes runs
  - Metric: nanoseconds of simulation per day (ns/day)
  - Easy to change run-time conditions: nodes, ranks, threads
  - Formatted results

FLATIRON INSTITUTE
Scientific Computing Core

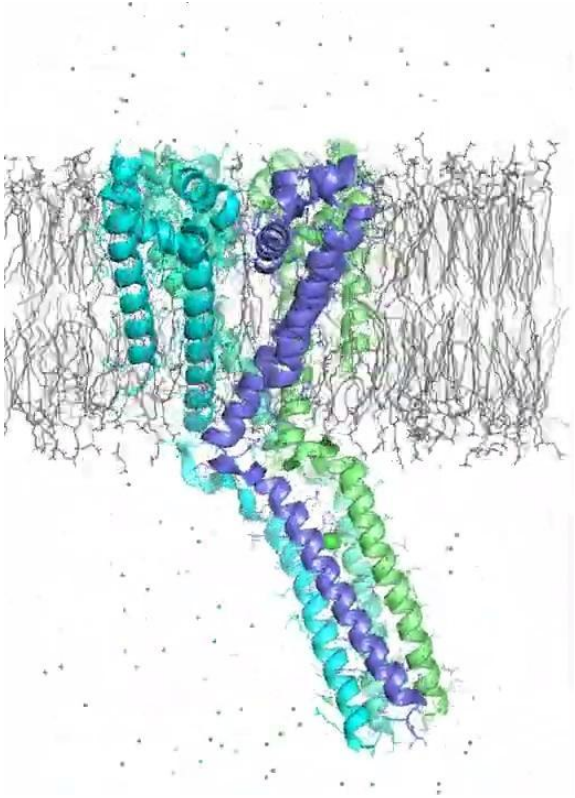
# Version Used, Build Details

- 2021.3: used in production at FI
  - Lots of performance numbers with it
  - Since that version, improvements for large # ranks on x86_64
- However, does not work with all compilers
  - Only GCC works out of the box with: -msve-vector-bits=512
  - Fujitsu/Clang and ARM work with patch
- FFTW: Fujitsu SVE implementation

# The Systems (aka inputs)

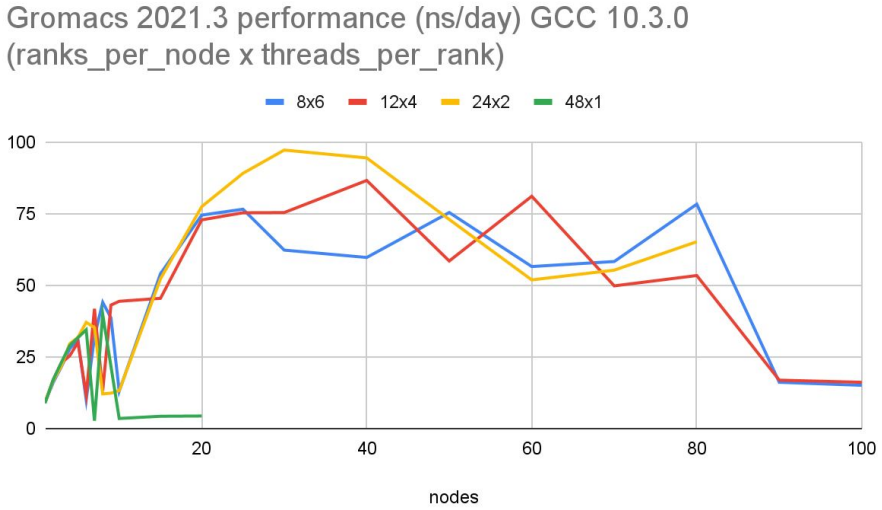
- **Ion channel** (main benchmark)
  - 150k atoms
  - Courtesy Sonya Hanson

    (Center for Computational Biology @FI)
- Lignocellulose
  - 1M atoms
  - Courtesy PRACE (EU HPC partnership)
  - Developed to test 10k+ core machines
  - Cannot run on “all-GPU” mode



# Results: GCC 10.3.0 + OpenMPI

- Varying the number of MPI ranks per node
- Fill with OpenMP threads
- Goal = using all 48 cores
- Holes= no automatic data decomposition found for the number of ranks




Best performance: **97 ns/day**

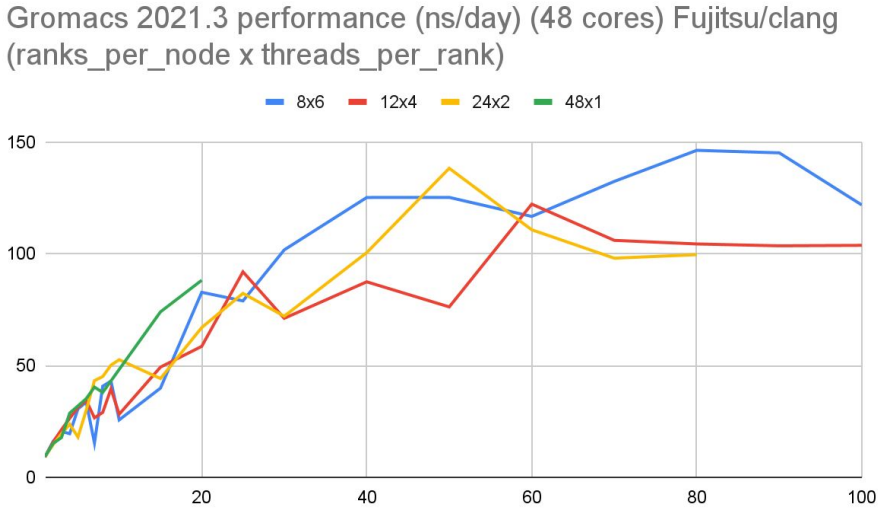
FLATIRON INSTITUTE
Scientific Computing Core

# Results: Fujitsu/clang + MPI

- Fujitsu + Clang compiler
- Fujitsu “trad” does not work
- Code has to be patched
- Courtesy: Gilles Gouaillardet (RIST)

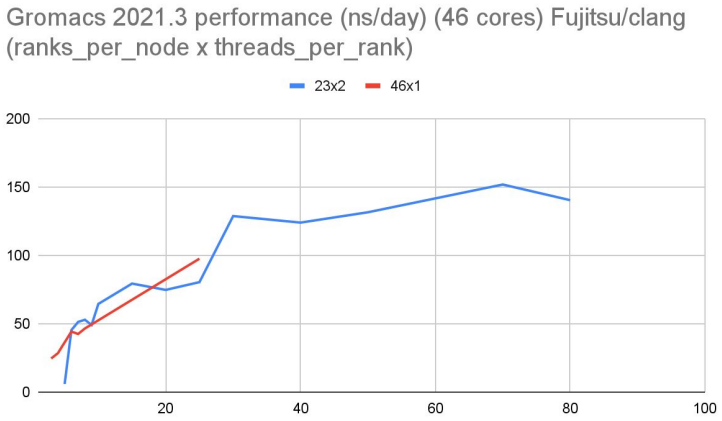
Best performance: **146 ns/day**

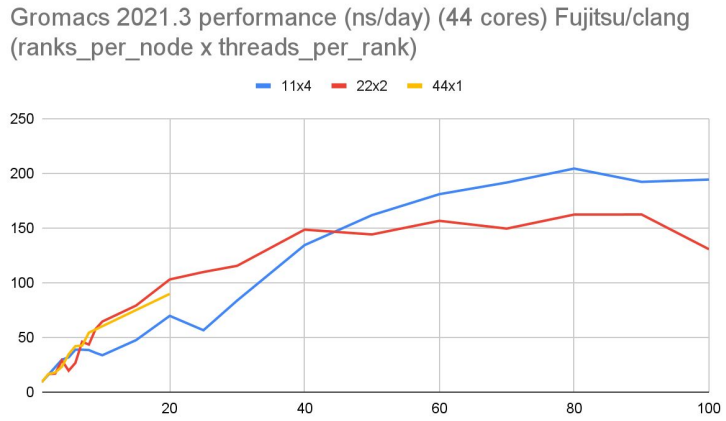
Fujitsu/clang gives better performance than GCC 10.3.0




FLATIRON INSTITUTE
Scientific Computing Core

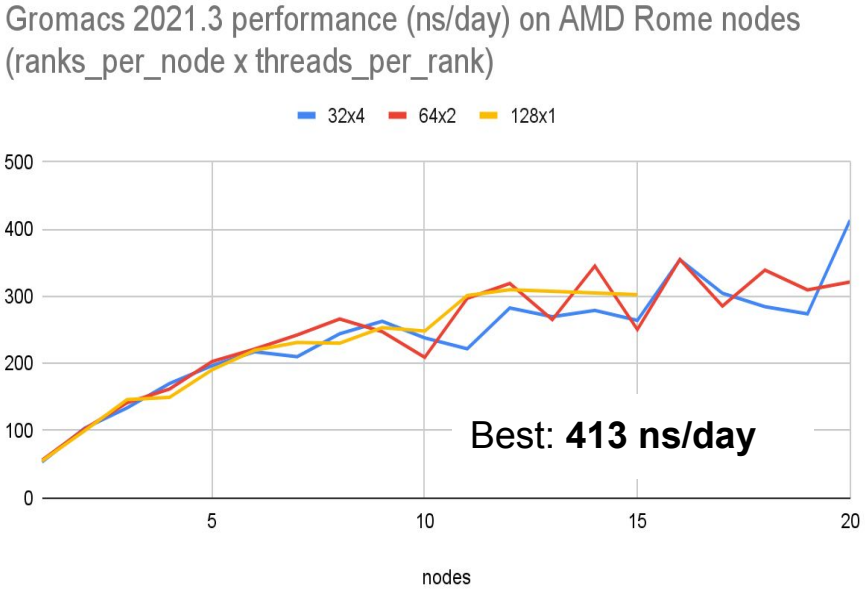
# Leaving some cores for the OS? (Fujitsu/clang)

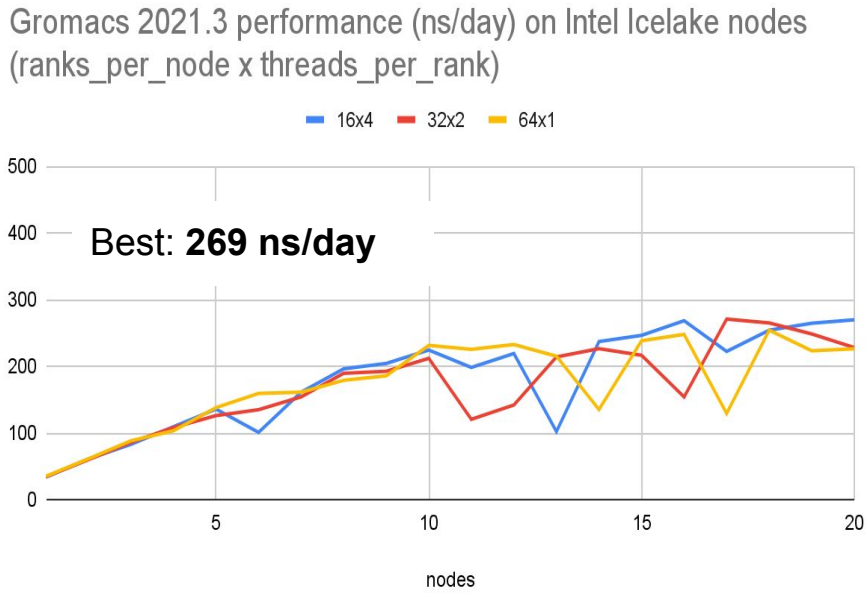




# Results on other machines



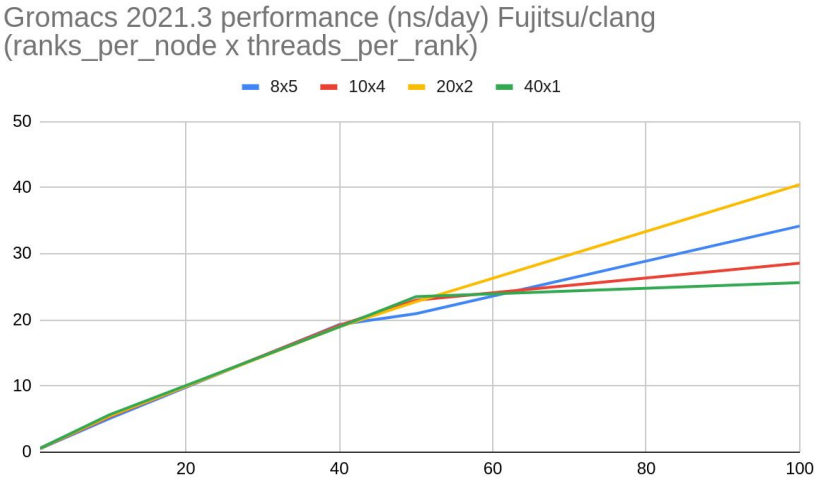

Single-node GPU performance:

- 4× NVIDIA V100-32GB: **150 ns/day**
- 4× NVIDIA A100-40GB: **178 ns/day**

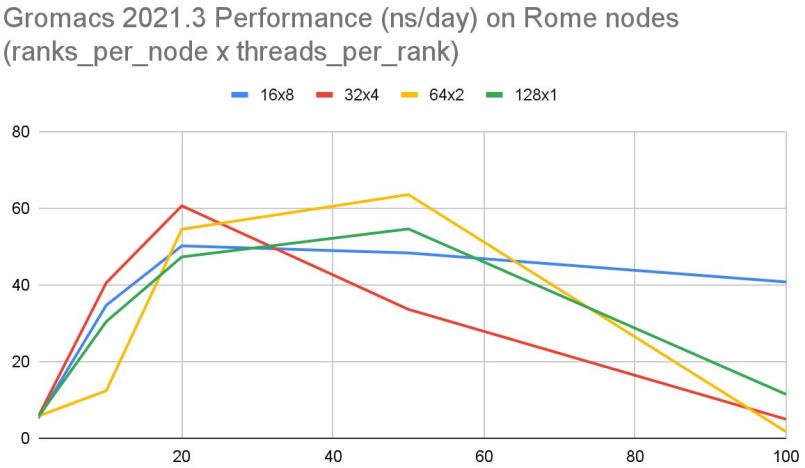
FLATIRON INSTITUTE
Scientific Computing Core

# Lignocellulose

## Ookami (Fujitsu/clang) 40 cores/node




## Rome (GCC) 128 cores/node




*Performance on Rome is more likely limited by collective operations: checking with a 48-core Cascadelake to confirm*

FLATIRON INSTITUTE
Scientific Computing Core

# Conclusions

- Gromacs works… but compiler selection is tricky (many choices, but few work)
- Problem size is very important:
  - medium system: 100 nodes comparable to 5 rome nodes
  - large system: things look a bit better, with better strong scaling
- Future work: ARM compiler, LLVM-14 once stable enough

*Thank you Eva, Tony, Robert, Catherine, Smeet for the help*

FLATIRON INSTITUTE
Scientific Computing Core

# Thank you

Any questions?