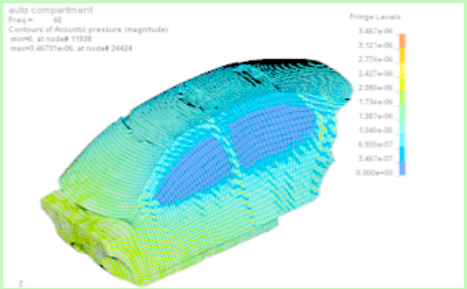
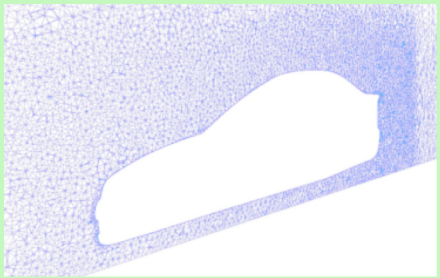
Implicit Analysis of Jet Engine Models on Thousands of Processors

LSTC
Livermore Software Technology Corp.

July 11, 2019

LSTC	Cray	NCSA	Rolls-Royce
Cleve Ashcraft, Roger Grimes	Jef Dawson	Erman Guleryuz	James Ong
Bob Lucas, Francois-Henry Rouet	Ting-Ting Zhu	Seid Koric	Todd Simons





Lego Porsche

Heise / ADAC https://www.youtube.com/watch?v=J2ECwX7Dl0c

Marco Theile, et.al., European LS-DYNA Conference, May 14, 2019

Outline

- Industrial Grand Challenge
- Tale of woe
- New algorithms and software
- Results
- Plans

Rolls-Royce Representative Engine Model (REM)

- Ambitious goals

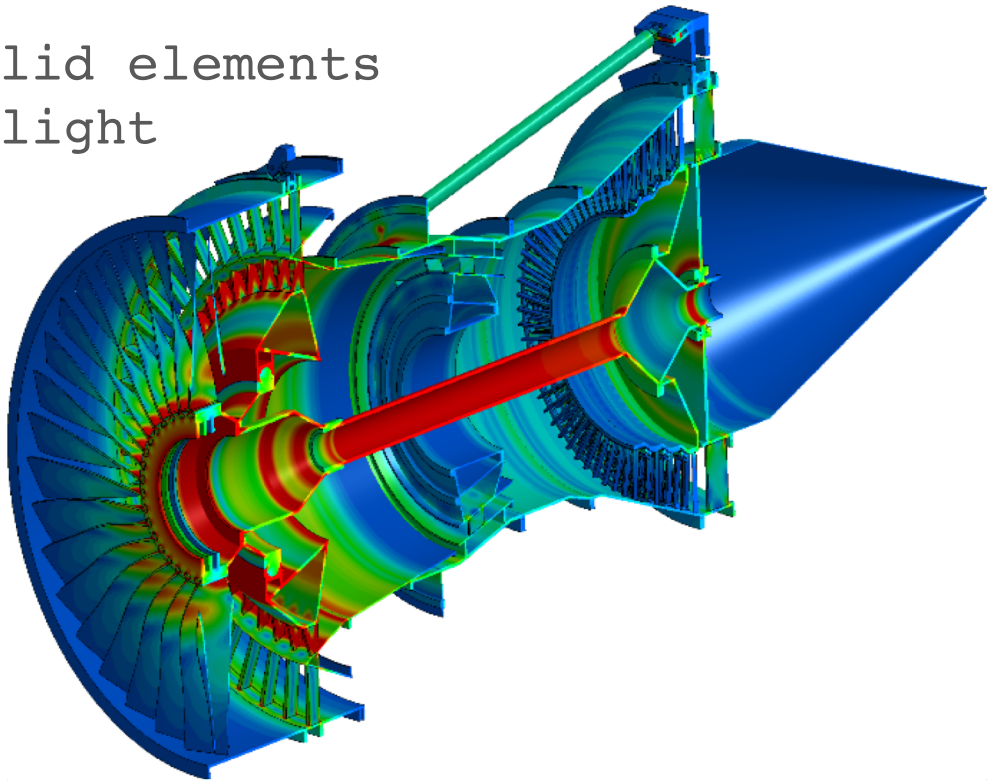
  Greater efficiency
  Virtual certification
  Digital twins

- Biggest implicit model known to LSTC

  Large version has 66M solid elements
  2000 sec. of simulated flight

- Initial experiments

  Explicit: 49 days
  Implicit: 19 days



Rolls-Royce found a bigger computer

- University of Illinois Urbana-Champaign

  National Center for Supercomputing Applications (NCSA)
  Industrial HPC program

- Blue Waters

```
Funded by the US National Science Foundation & Illinois
28,864 node Cray XE/XK (16 AMD cores, 64 GBytes std.)
    4,228 augmented with GPUs
7.1 Pflop/s (w/o GPUs)
```

And another …

- Oak Ridge National Laboratory

  Oak Ridge Leadership Computing Facility (OLCF)
  INCITE award

- Titan

  Funded by US Dept. of Energy Office of Science
  18,688 node Cray XK (8 AMD cores, 32 Gbytes, K20X GPU)
  27 Pflop/s (incl. GPUs)

Last year's tale of woe

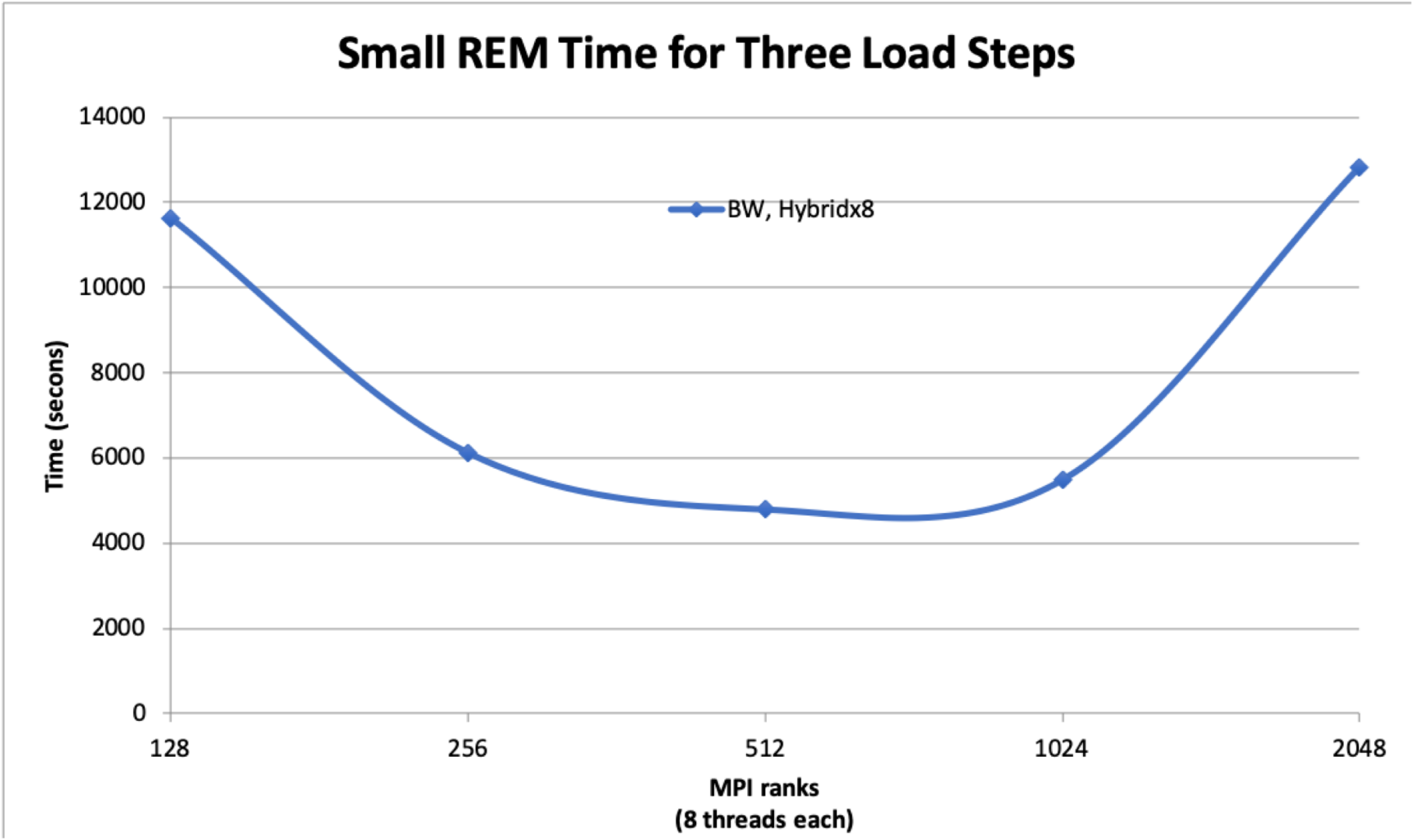
- Existential Problems

  64 GB isn't enough for processor zero to parse the input files
  METIS_NodeWND crashed
  Fortran I/O was 99.9999999% reliable
  Rank exceeded the period of a BCSLIB random number generator

- Sequential bottlenecks

  Input processing
  Reordering
  Symbolic factorization
  Linear constraints

Initial implicit results on Blue Waters



Addressing the sequential bottlenecks

- Input Processing

  Off-line pre-decomposition helps

- Reordering

  LS-GPart nested dissection

- Symbolic factorization

  Designed to follow a nested dissection reordering
  Independent symbolic factorization of domains and separators

- Linear constraints

  Improved the communication
  True parallel solution is work-in-progress

Addressing the linear solver bottlenecks

- Input Processing

  Off-line pre-decomposition helps

- Reordering

  LS-GPart nested dissection

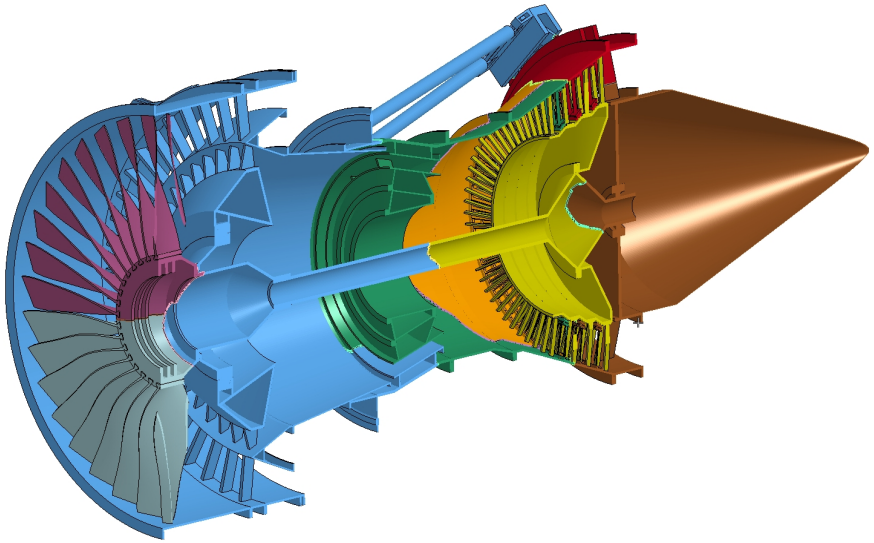
- Symbolic factorization

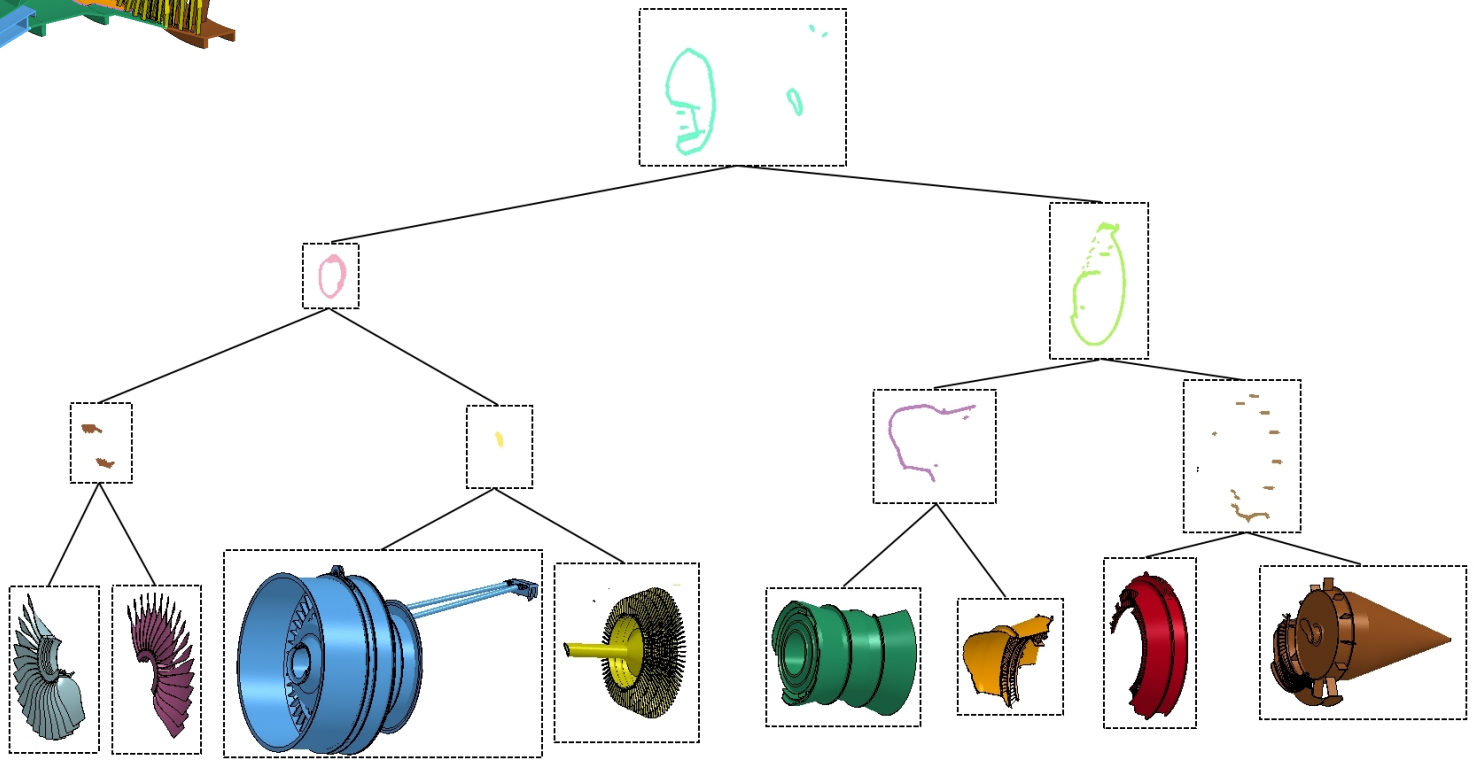
  Designed to follow a nested dissection reordering
  Parallel symbolic factorization of domains and separators

- Linear constraints

  Improved the communication
  True parallel solution is work-in-progress

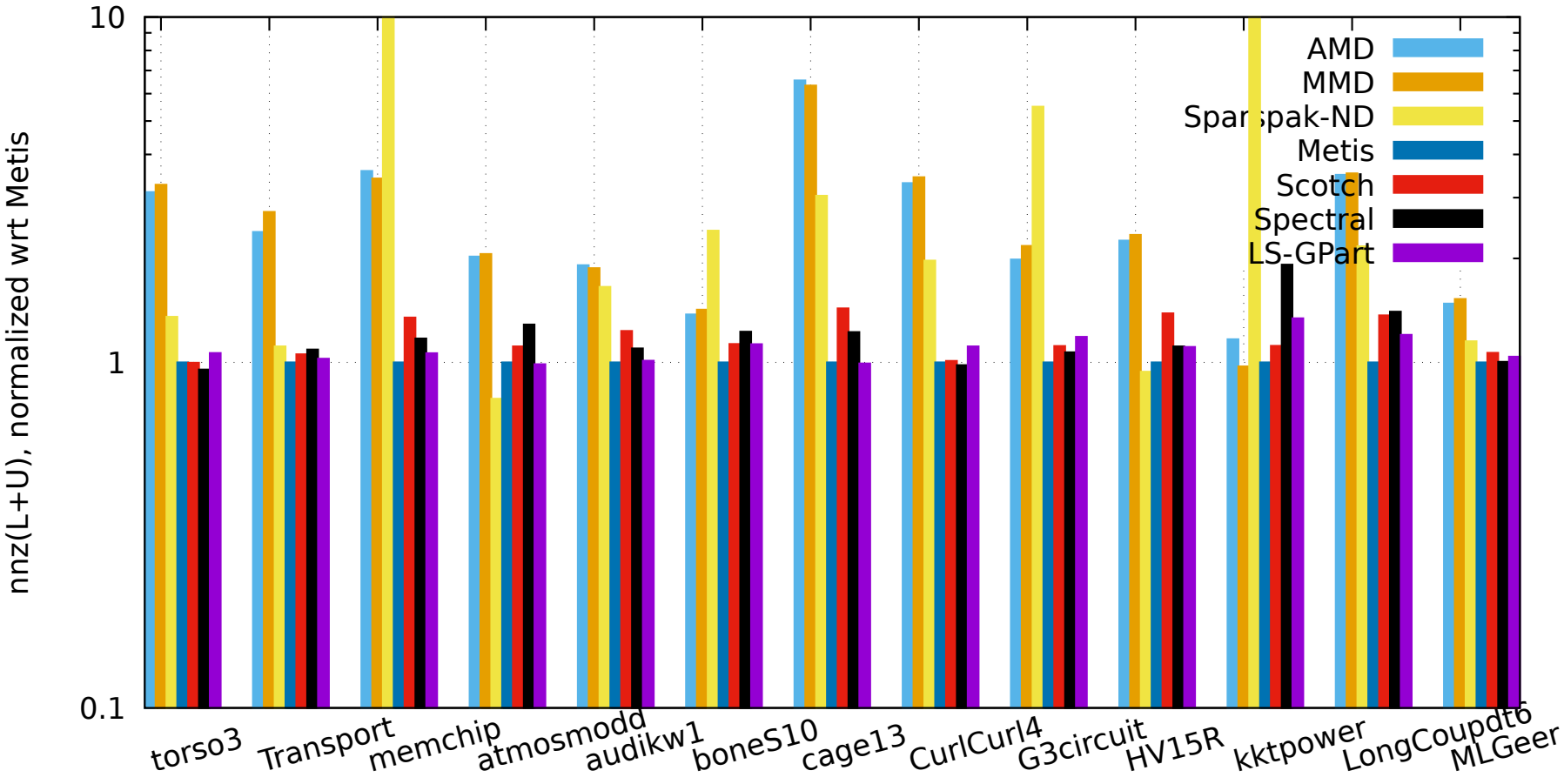
Nested dissection of REM for eight processors





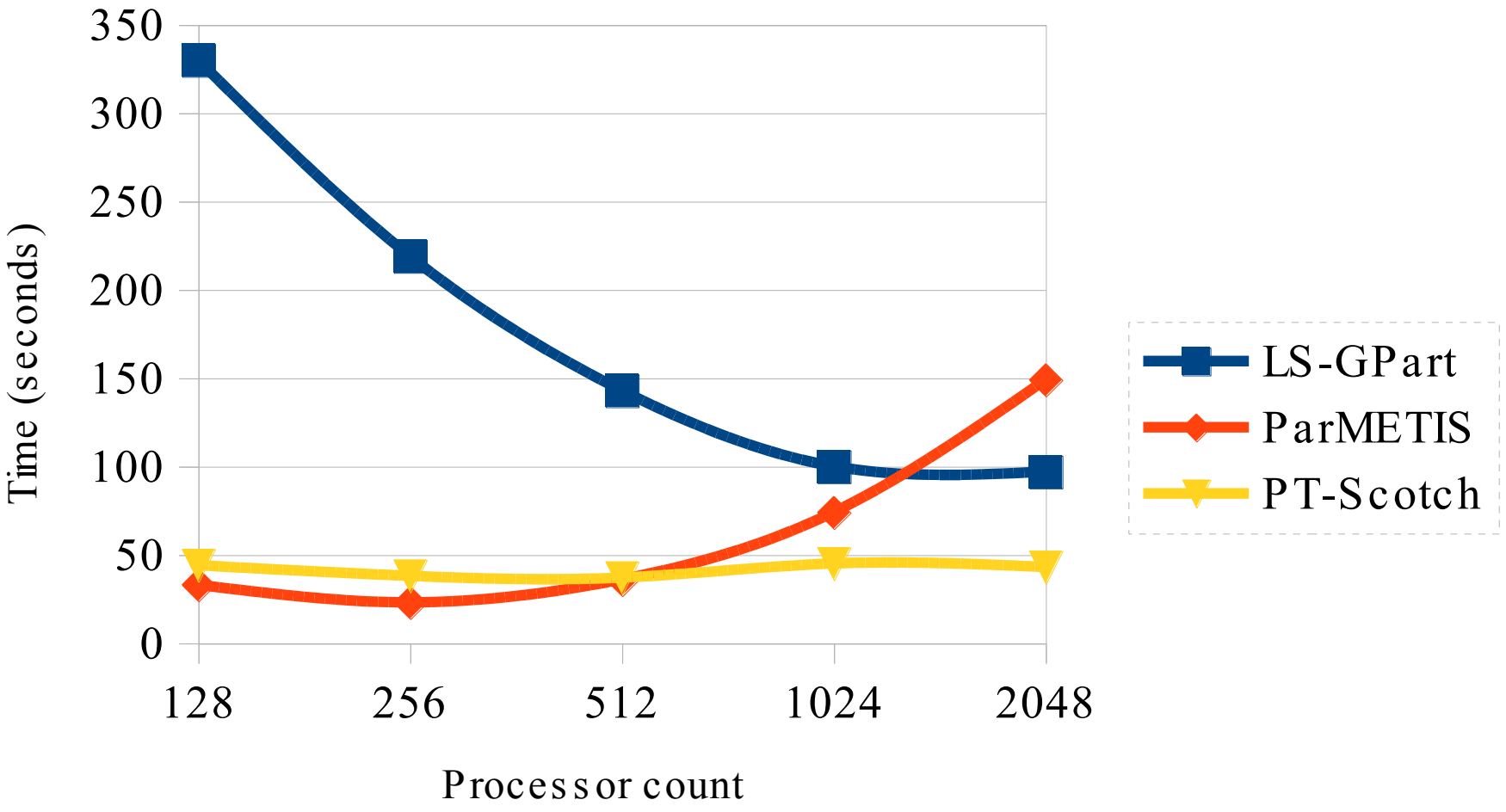
Disclaimer: No jet engines were harmed during the making of this presentation ☺

LS-GPart ordering quality



LS-GPart added to reordering comparison presented in “Preconditioning using Rank-structured Sparse Matrix Factorization”, Ghysels, et.al., SIAM PP 2018

LS-GPart performance

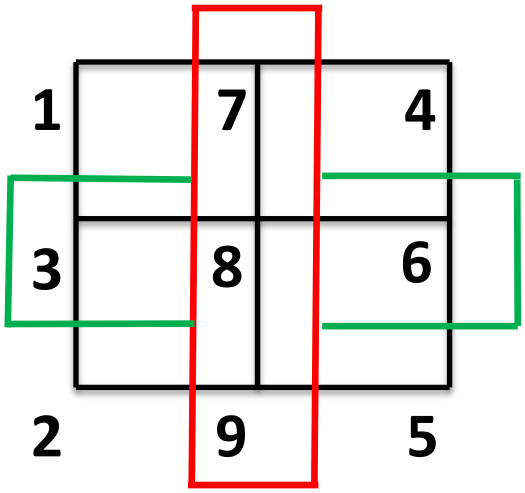


Parallel symbolic factorization

- Assume a nested dissection ordering

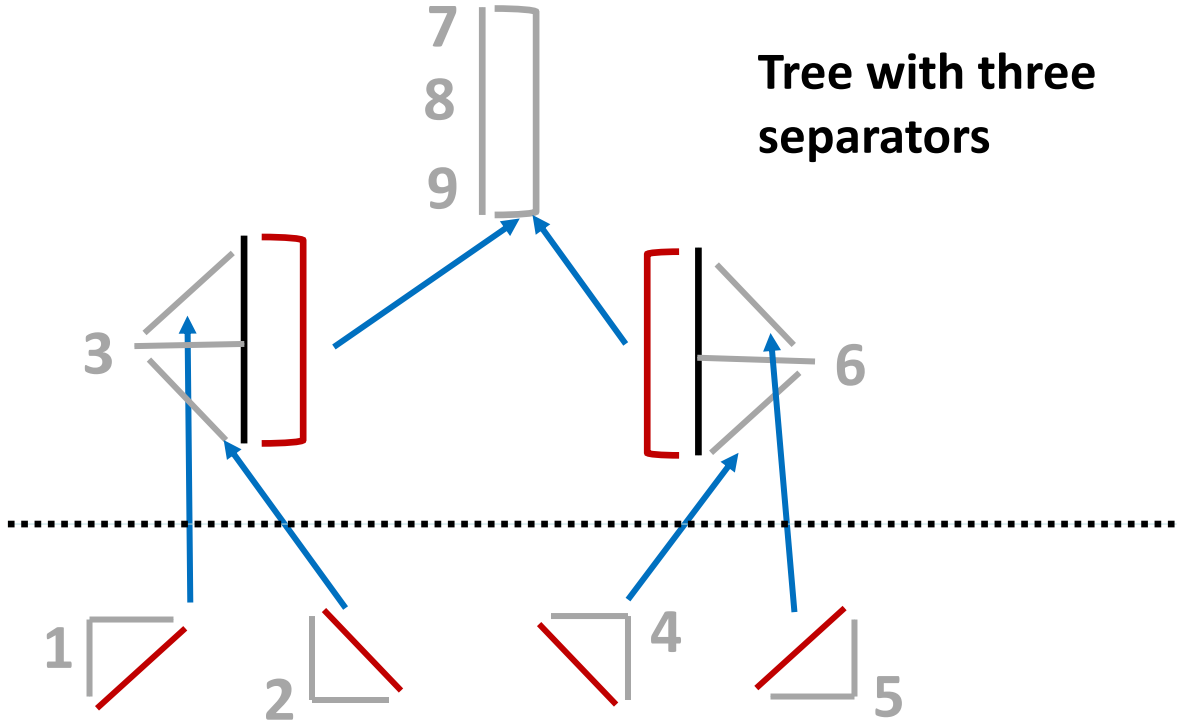
  Symbolic factorization of domains performed independently

  Symbolic factorization of separators involves communication



Reordered mesh

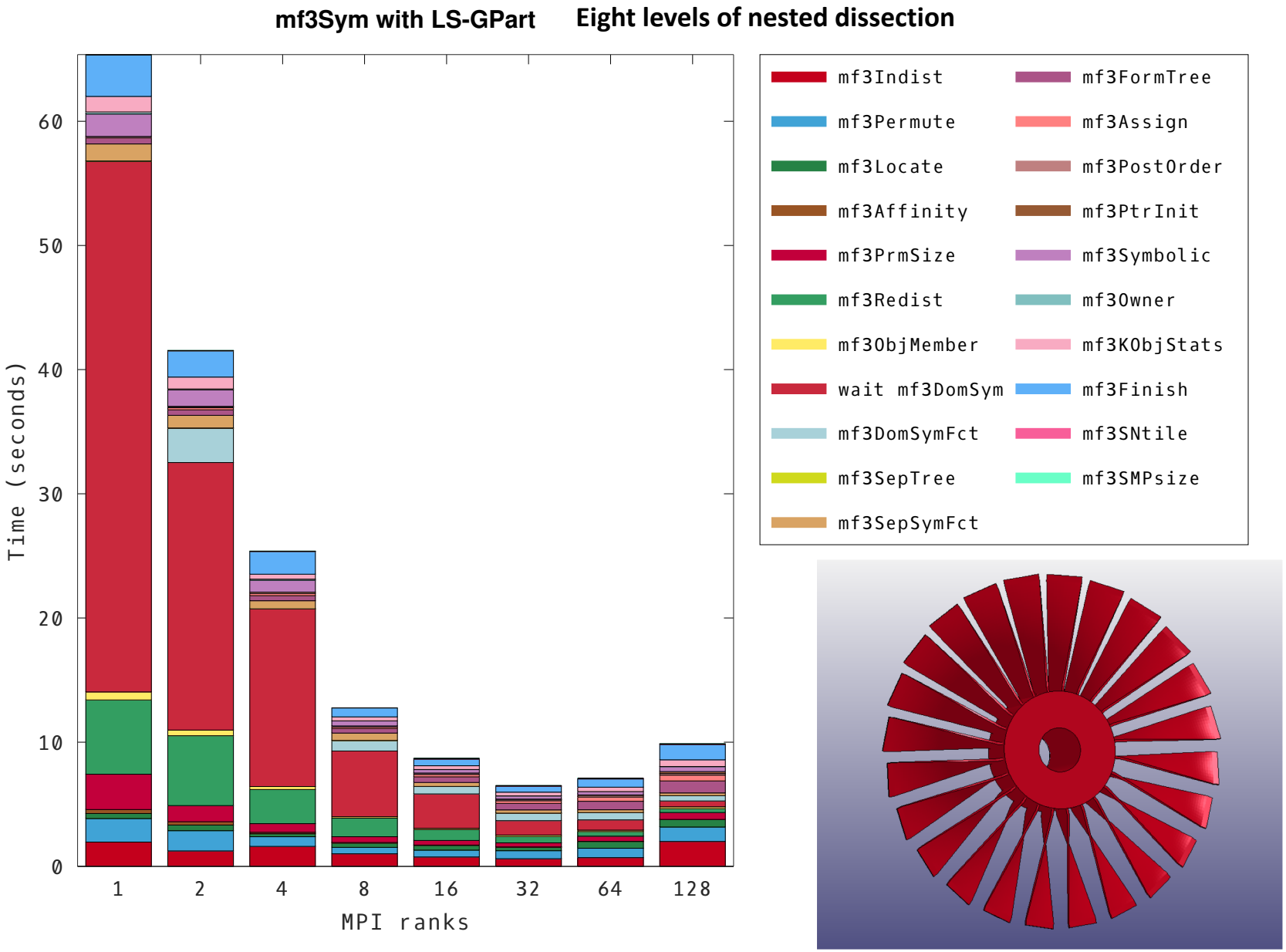
With separators



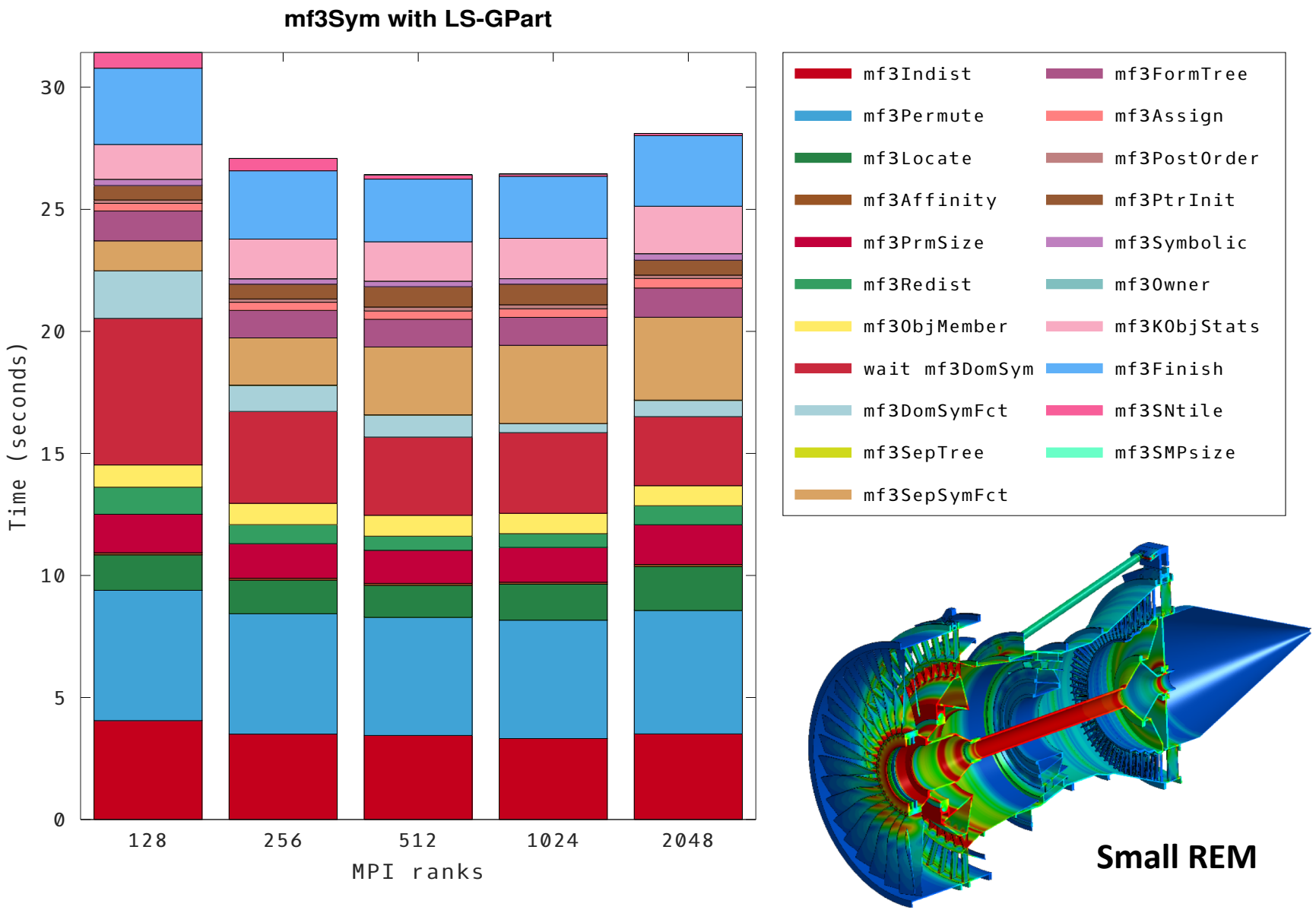
Tree with three separators

Four independent domains

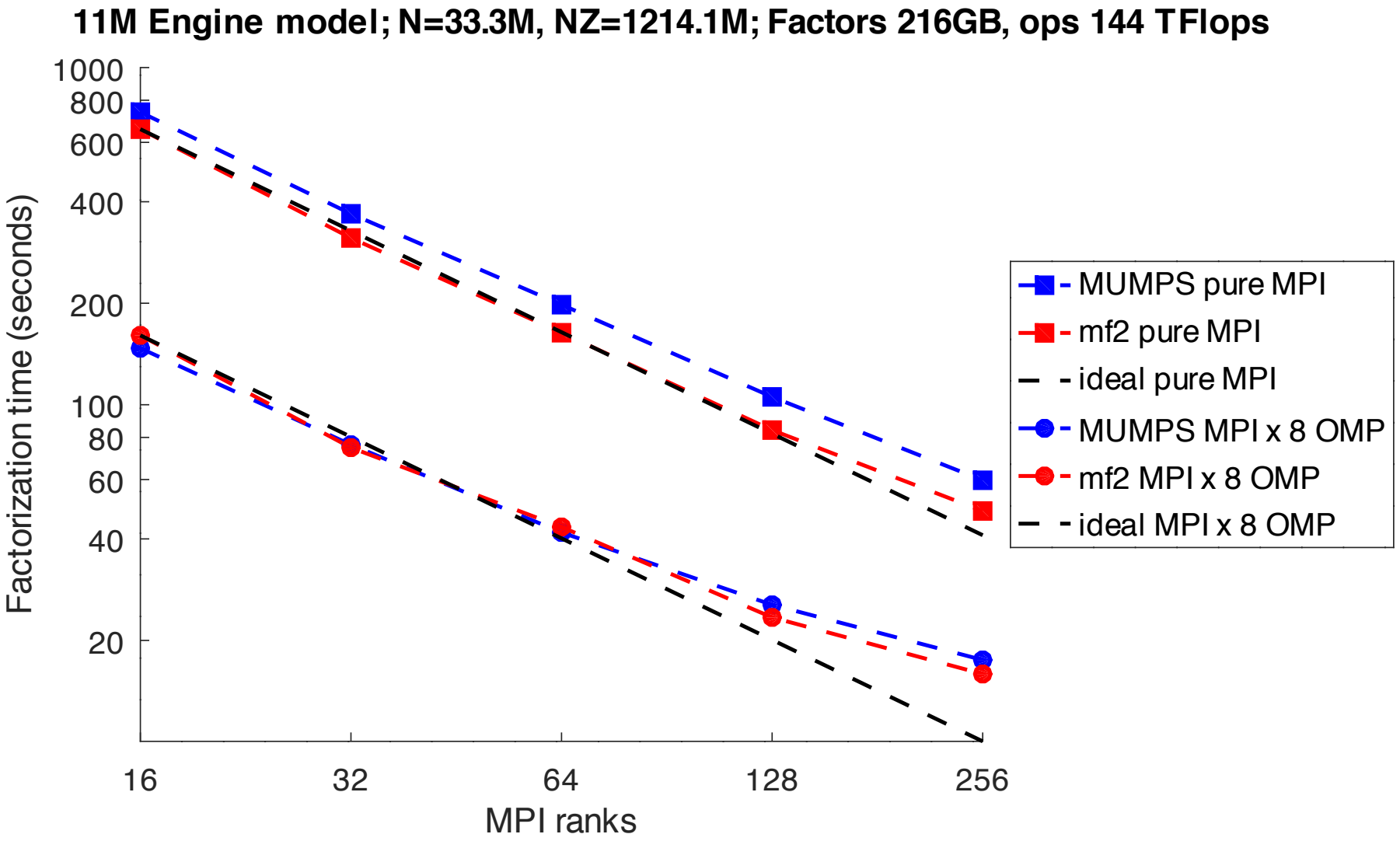
Parallel symbolic factorization (Sept 2018)



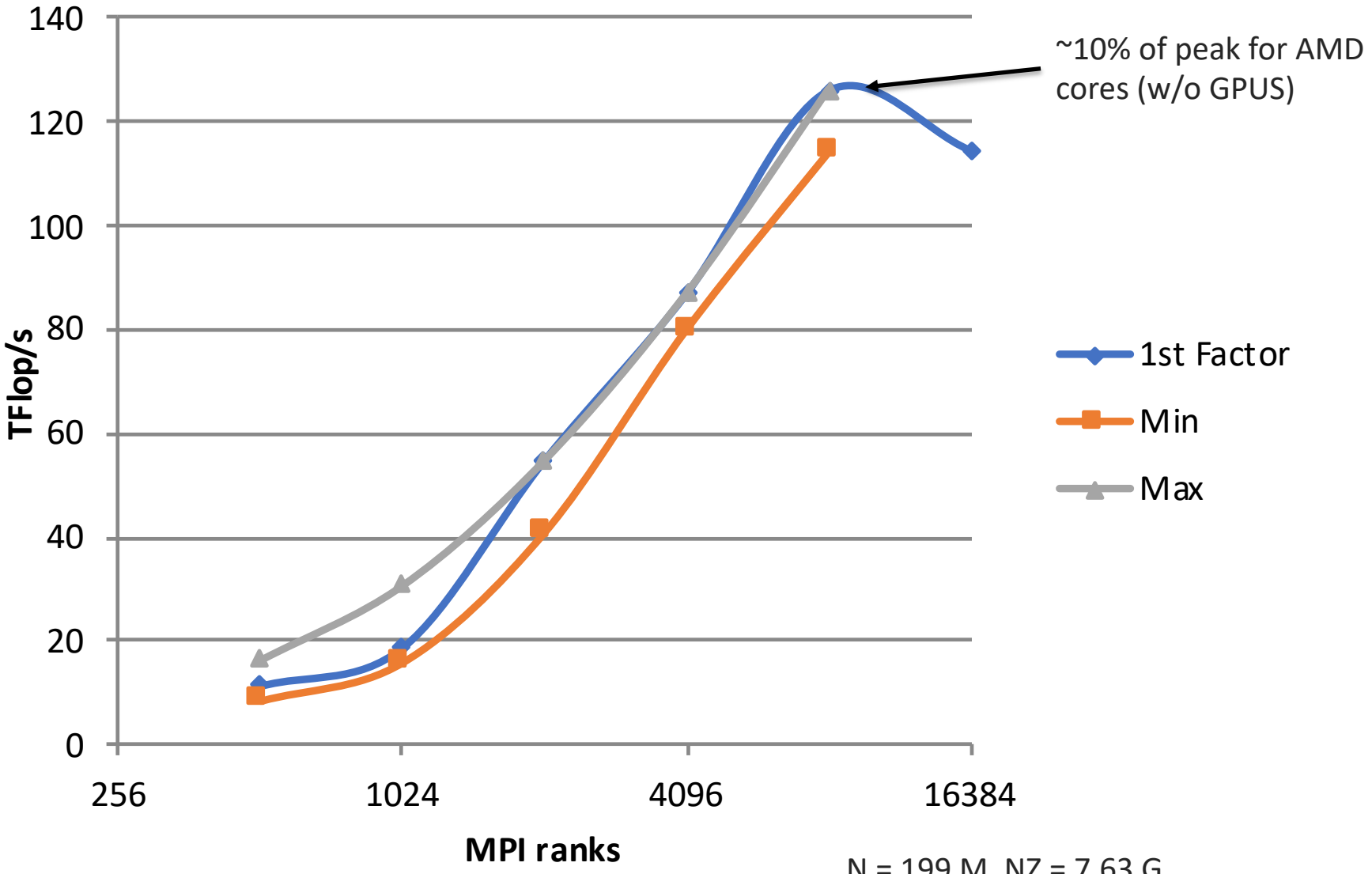
Parallel symbolic factorization (Mar 2019)



Multifrontal factorization timing (engine parts)

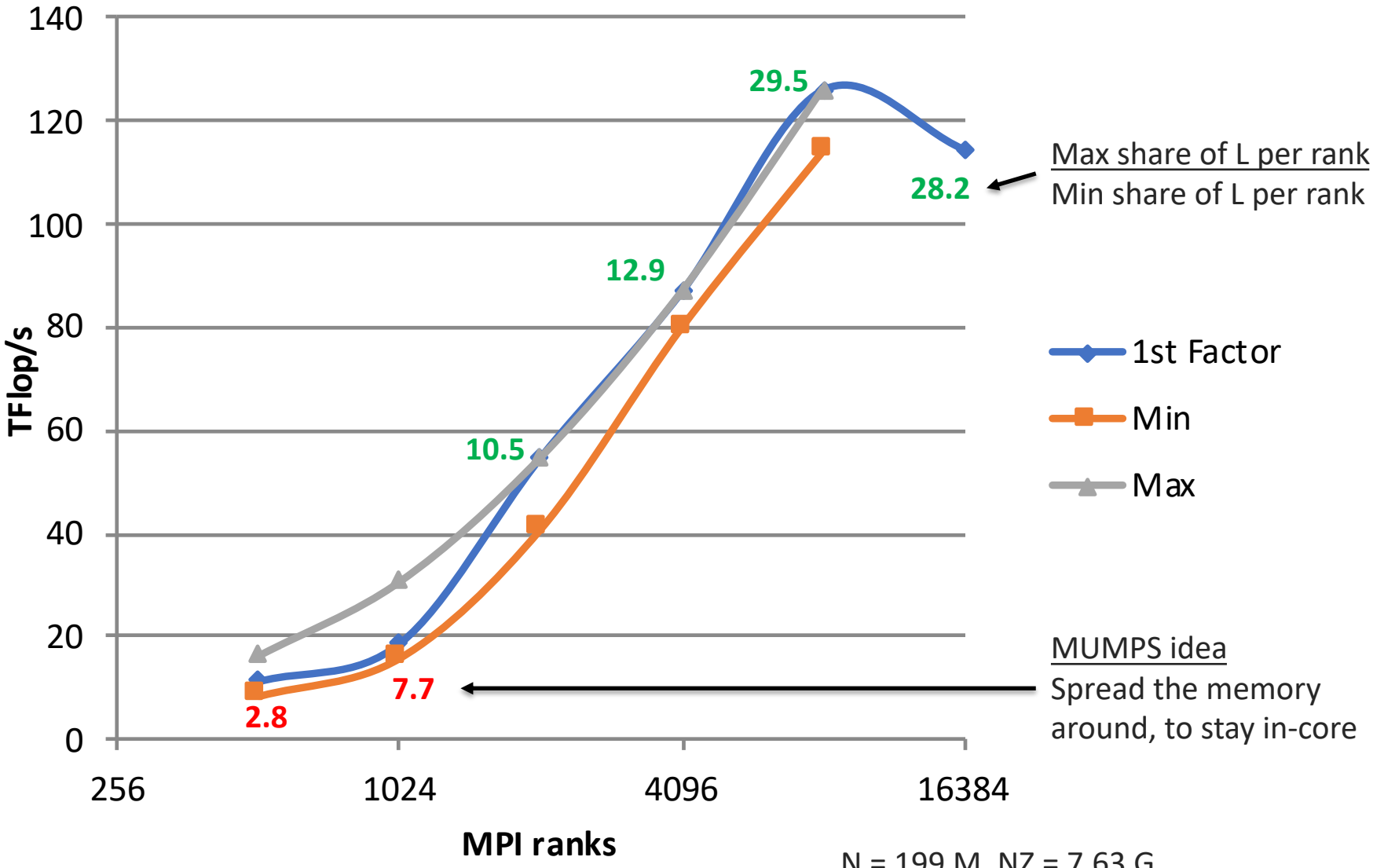


Multifrontal factorization performance (LREM)



N = 199 M, NZ = 7.63 G
Factors = 479 G, Ops = 8.0 – 8.1 P

Multifrontal factorization performance (LREM)



N = 199 M, NZ = 7.63 G
Factors = 479 G, Ops = 8.0 – 8.1 P

Factorization discussion

- Static load balance could be better

  Attempts to minimize the run time of a performance model
  Calibrated on Clement's 16-core Sandybridge workstation

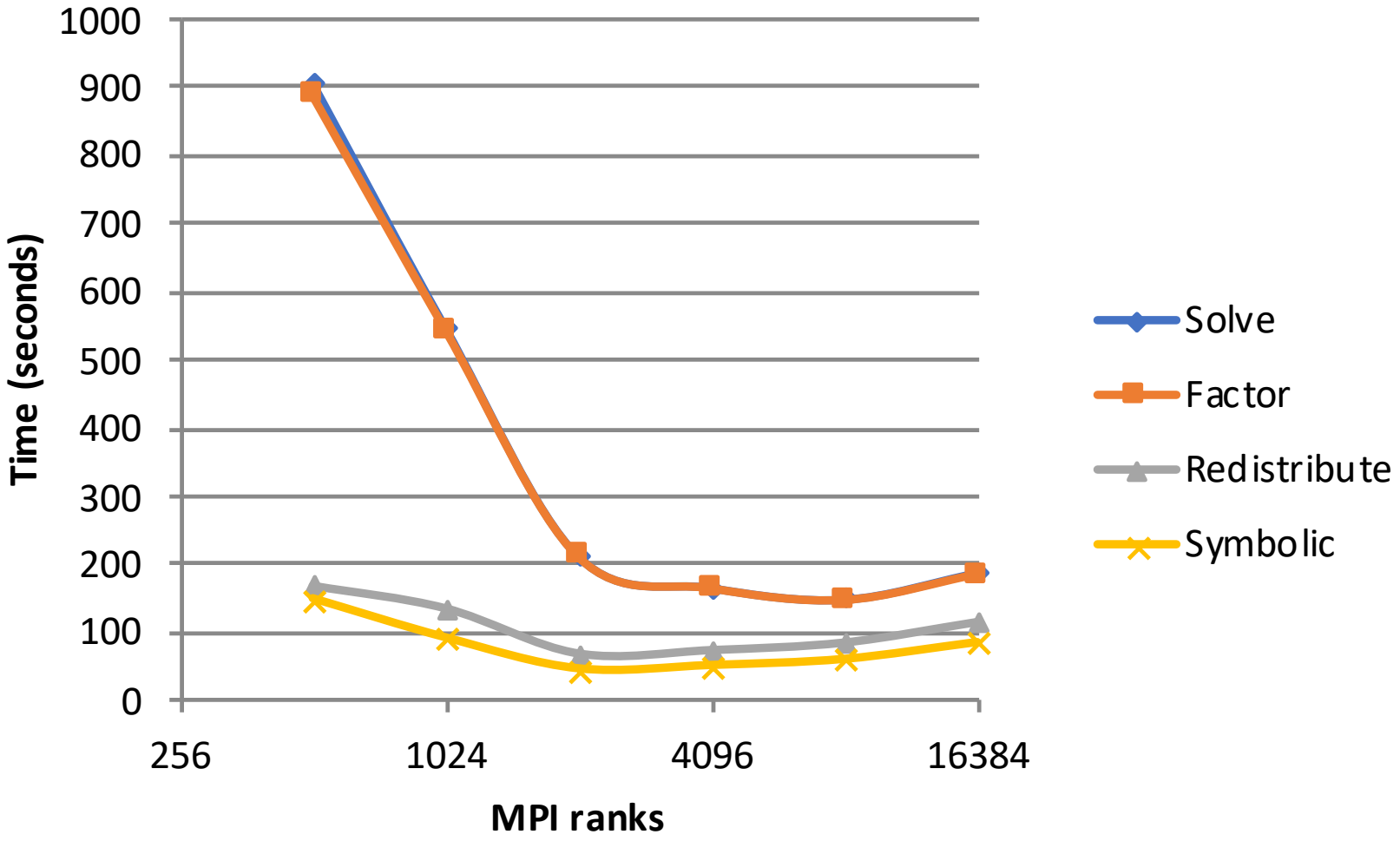
- Communication between frontal matrices (assembly)

  Weakness uncovered at scale on Titan
  Fine-grain, asynchronous communication
    MPI_ISEND & MPI_IRECV

- Communication within frontal matrices (factorization)

  With 16K MPI ranks, MPI_BCAST dwarfs DGEMM time
    Even with just one thread per rank

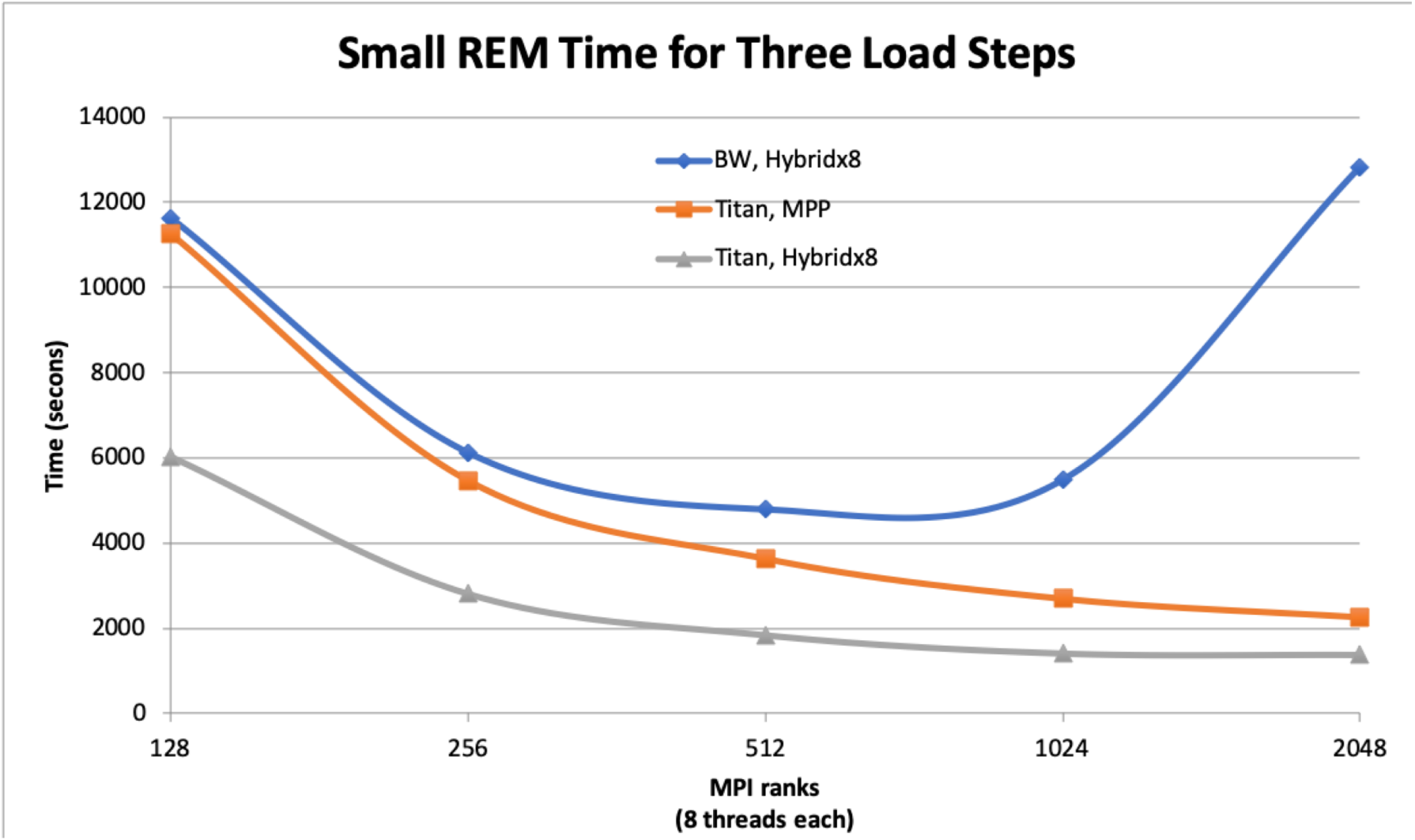
Factorization performance in context



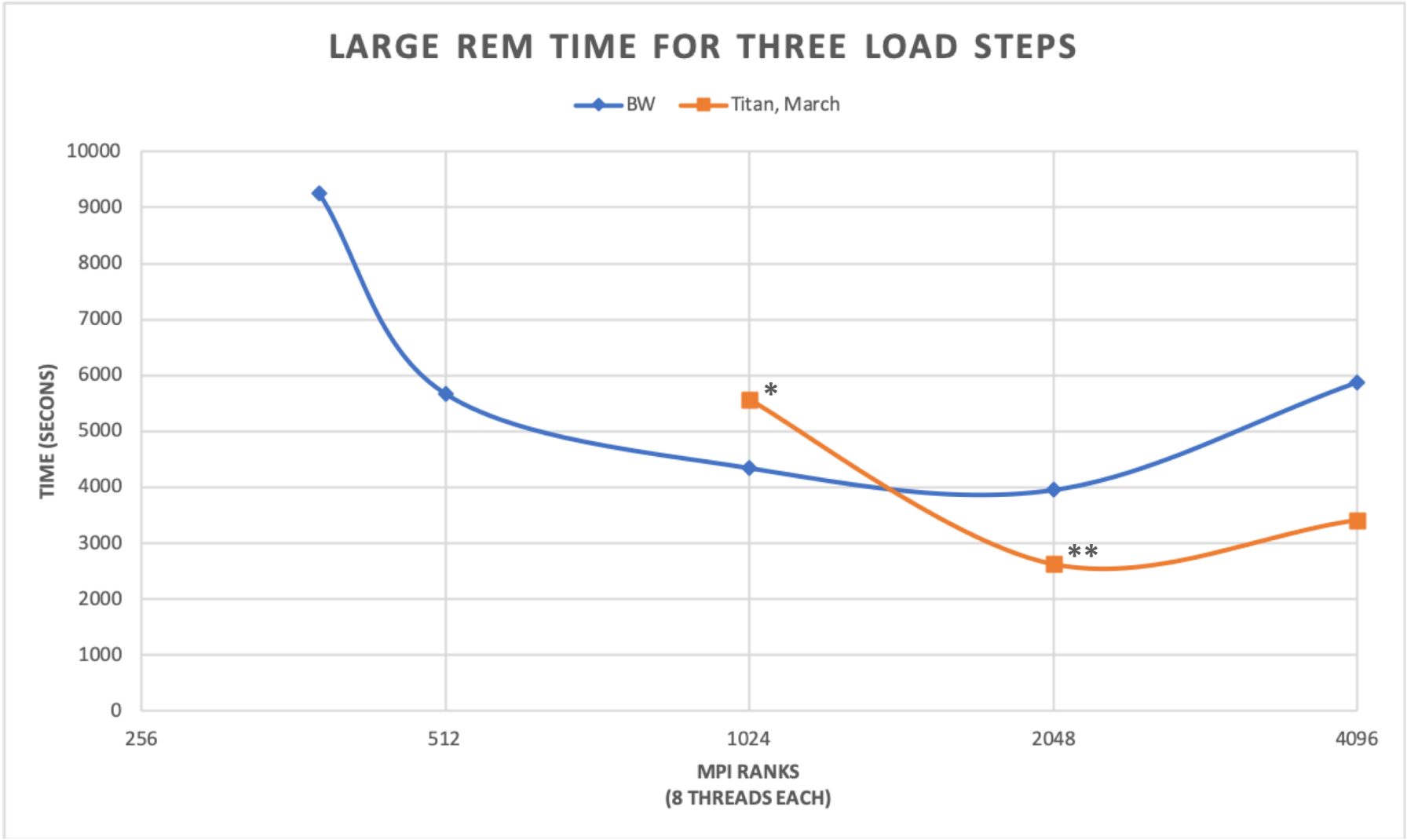
Curves are cumulative time

Factorization not the biggest problem at 128K cores

SREM after addressing the sequential bottlenecks



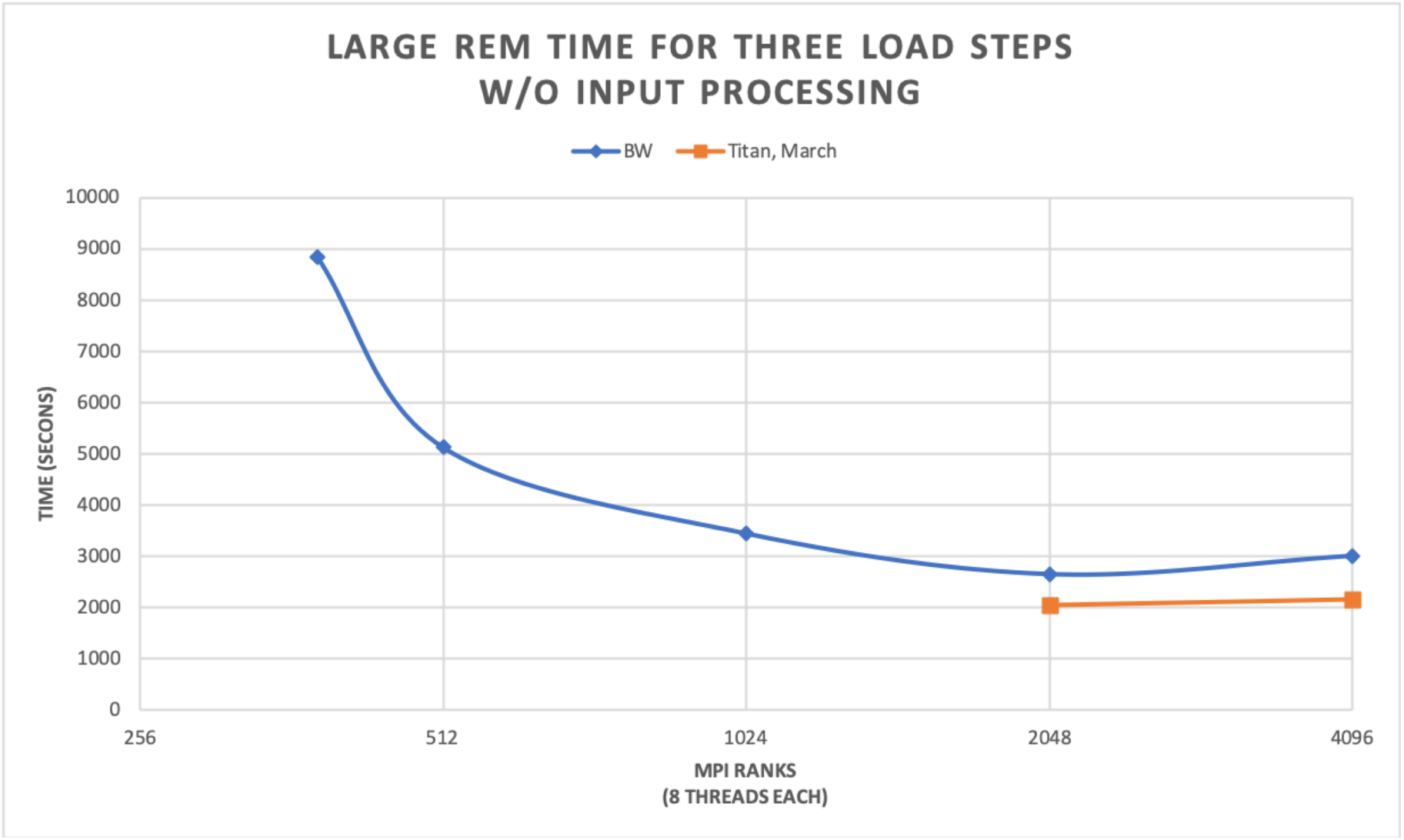
LREM parallel performance (three load steps)



*Load-balance distorted, Titan has only 32 GB memories

**We believe Titan's workload was less I/O and communication bound than Blue Waters'

LREM parallel performance after input processing



Practical impact

- Rolls-Royce's initial explicit run took seven weeks!

  A science project

- Now it runs overnight (12 hours)

  Practical turnaround time on local resources

- Still only a "representative" model, not a real one

Near-term plans for the linear algebra

- Linear constraint processing

  Complete rewrite in progress

- Reordering

  Should speedup on thousands of MPI ranks
  Should exploit multiple threads too

- Symbolic factorization

  Interface needs to be revisited

- Multifrontal solver

  Frontal matrix assembly
  Update performance model

Further plans

- Larger, more sophisticated models

  Perhaps a real engine model, spinning

  Loading representative of a real flight

    Taxi, take-off, climb, cruise, descend, land

- More physics

  Acoustics, thermal, fluids

  High-energy events

    Bird strike, blade-out containment

- Eigenvalues

  Thousands of them

  Extracted at multiple points in the simulated flight

LSTC
Livermore Software Technology Corp.

Thank you!