Take-home

• Load balancing should come first
• In the regions with acceptable efficiency, single-node execution is most important
• A single-node performance model guides the way to optimal node performance
• Reducing resource consumption is never a bad idea (memory, cache, working set,...)
• In the end, care about optimizations for massive parallelism
  – Comm/comp overlap
  – Sync reduction
  – Eliminate global operations
  – MPI fine-tuning
  – ...

Single/Multi-node optimization - Howto

- Check basic issues:
  - Do you use the latest/most performant compiler/MPI library/... on your platform? The right optimization flags?
- Check load imbalance/communication issues
  - Tools that may help: Scalasca, Vampir,
- Vectorize loops
  - Done by compiler
  - Or help the compiler (pragmas,...)
  - Tools that may help: MAQAO, DECAN, LIKWID,
- Improve spatial/temporal locality
  - Might require restructuring of data to avoid strided access
- And much more → tutorial needed
Thanks for your attention.
Any more questions?

Bettina Krammer, UVSQ/ECR, bettina.krammer@uvsq.fr
Georg Hager, Jan Treibig, Gerhard Wellein, RRZE, {firstname.lastname@rrze.uni-erlangen.de}
Anthony Scemama, CNRS, scemama@irsamc.ups-tlse.fr