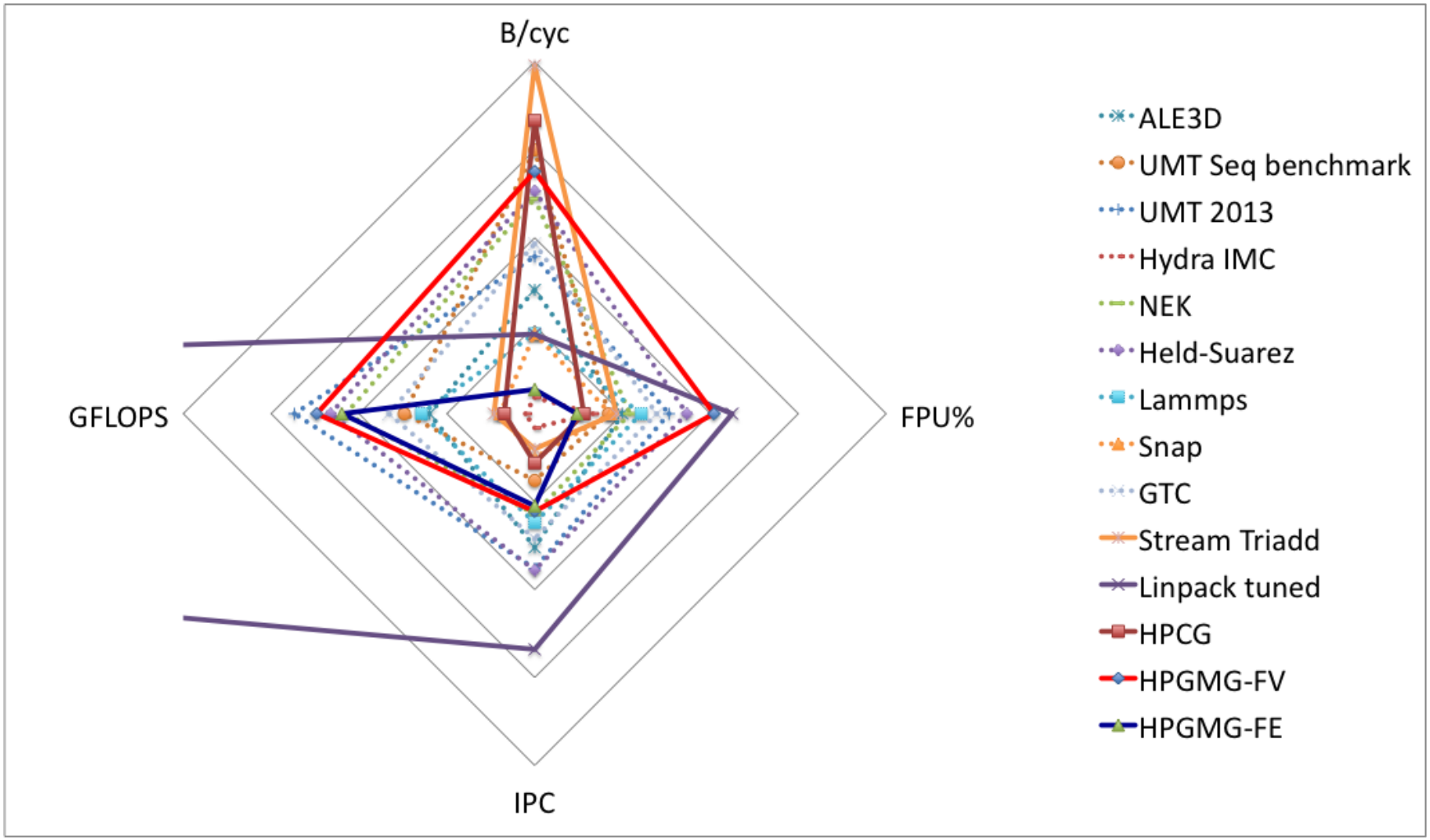
HPGMG: a new benchmarking proposal

- `https://hpgmg.org`, hpgmg-forum@hpgmg.org mailing list
- SC14 BoF: Wednesday, Nov 19, 12:15pm to 1:15pm
- Mark Adams, Sam Williams (finite-volume), myself (finite-element), John Shalf, Brian Van Straalen, Erich Strohmeier, Rich Vuduc
- Implementations

  Finite Volume memory bandwidth intensive, simple data dependencies

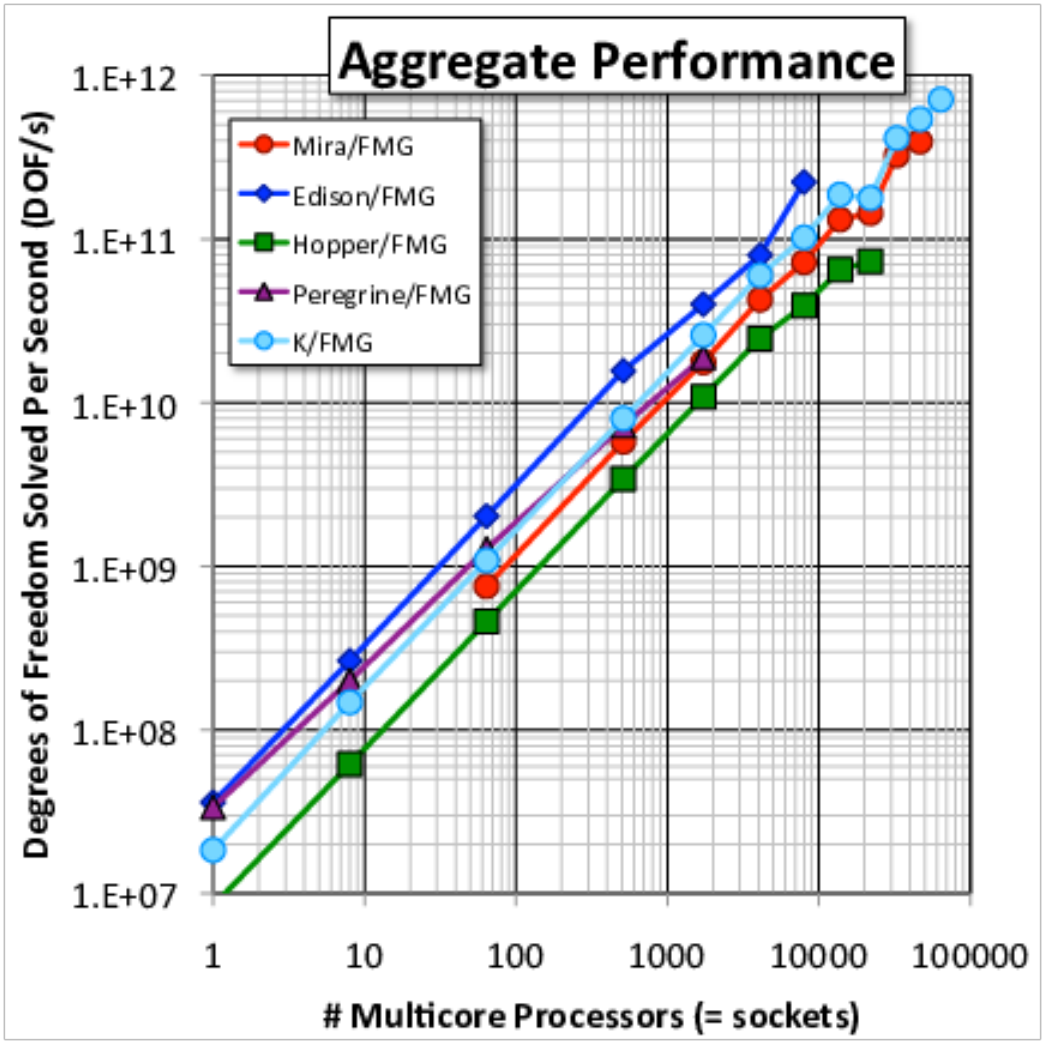
  Finite Element compute- and cache-intensive, vectorizes, overlapping writes
- Full multigrid, well-defined, scale-free problem
- Goal: necessary and sufficient
  - Every feature stressed by benchmark should be necessary for an important application
  - Good performance on the benchmark should be sufficient for good performance on most applications

Kiviat diagrams

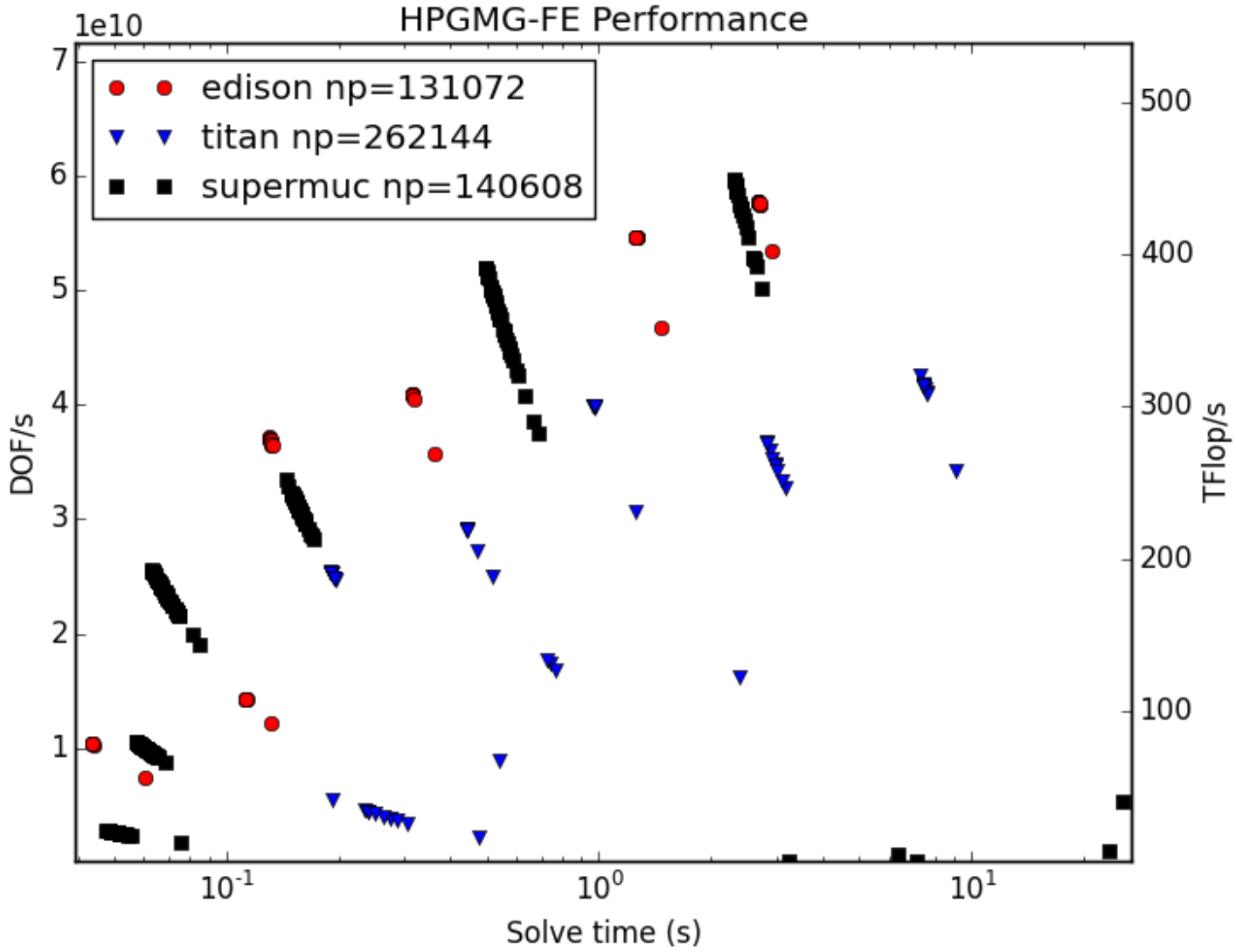


- c/o Ian Karlin and Bert Still (LLNL)

HPGMG distinguishes networks at 1M dofs/core

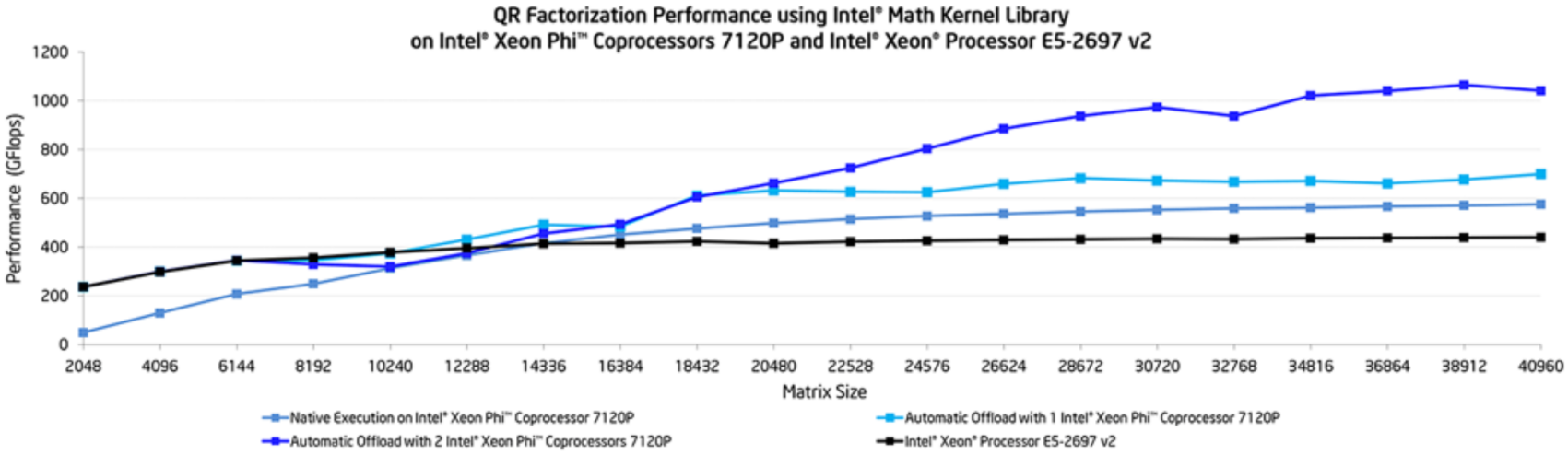


- Peregrine and Edison have identical node architecture
- Peregrine has 5:1 tapered IB, Edison has Aries dragonfly topology



- Turn-around time often not negotiable
  - policy, manufacturing, forecasting
- Users like predictable performance across a range of problem sizes
- Transient problems do not weak scale even if each step does

Where we are now: *QR* factorization with MKL on MIC



- Figure compares two CPU sockets (230W TDP) to one MIC (300W TDP plus host)
- Performance/Watt only breaks even at largest problem sizes
- Haswell-EP doubles performance within same power envelope
- $10^4 \times 10^4$ matrix takes 667 GFlops: about 2 seconds
- This is an $O(n^{3/2})$ operation on n data
- MIC cannot strong scale, no more energy efficient/cost effective
- “hard to program” versus “architecture ill-suited for problem”?