

Todo

- Upload slides
- Competition slide (not other groups)
- Verification slide (emphasis benchmark)
- Graph cleanup
- Risk (maybe, not asked for)

libmsr

Tools and Components for Memory and Power Profiling

Barry Rountree (LLNL-CASC)

Martin Schulz (LLNL-CASC)

David K. Lowenthal (UArizona)

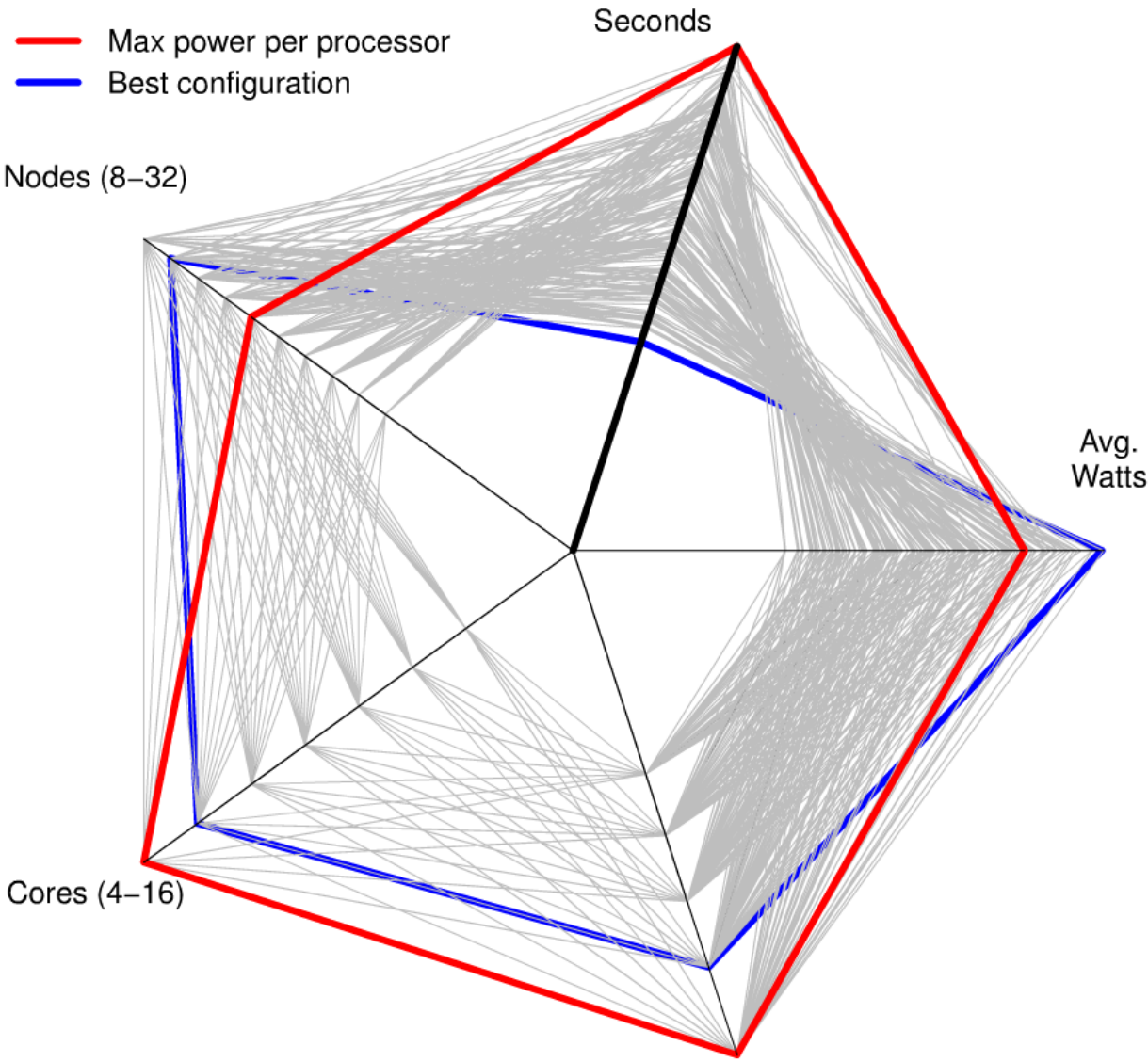
Peter Bailey (UArizona)

Tapasya Patki (UArizona)

20MW

sp-mz, 4500W power bound

- Max power per processor
- Best configuration



Processor Power Bound (Watts) (51, 65, 80, 95, 115)

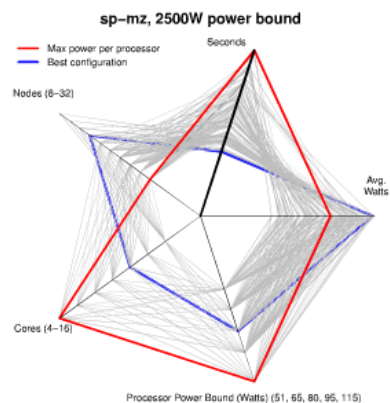
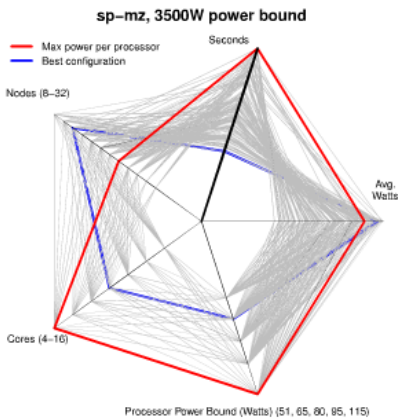
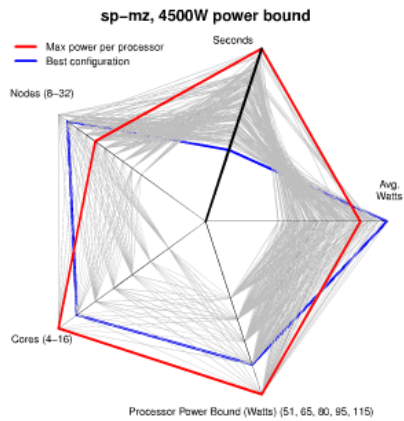
/dev/cpu/X/msr

Intel's Running Average Power Limit (RAPL)

Energy measurement (DRAM, CPU)
Power caps (CPU, maybe DRAM)

Table 14-2. RAPL MSR Interfaces and RAPL Domains

Domain	Power Limit (Offset 0)	Energy Status (Offset 1)	Policy (Offset 2)	Perf Status (Offset 3)	Power Info (Offset 4)
PKG	MSR_PKG_POWER_LIMIT	MSR_PKG_ENERGY_STATUS	RESERVED	MSR_PKG_RAPL_PERF_STATUS	MSR_PKG_POWER_INFO
DRAM	MSR_DRAM_POWER_LIMIT	MSR_DRAM_ENERGY_STATUS	RESERVED	MSR_DRAM_RAPL_PERF_STATUS	MSR_DRAM_POWER_INFO
PP0	MSR_PP0_POWER_LIMIT	MSR_PP0_ENERGY_STATUS	MSR_PP0_POLICY	RESERVED	RESERVED
PP1	MSR_PP1_POWER_LIMIT	MSR_PP1_ENERGY_STATUS	MSR_PP1_POLICY	RESERVED	RESERVED



Requires

msr kernel module

creates /dev files

root

rw perms on /dev file

one-time setup

trusted users

root exploits exist

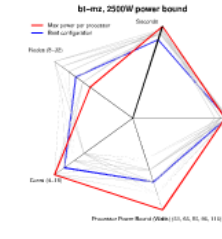
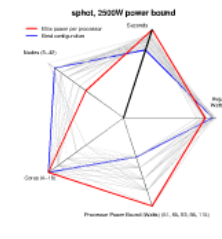
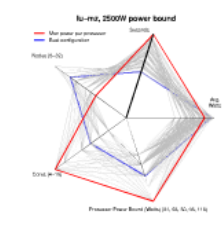
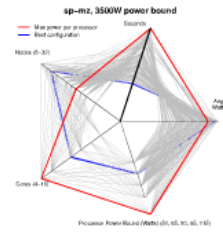
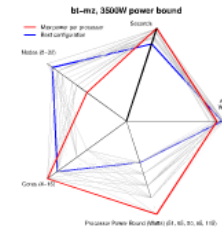
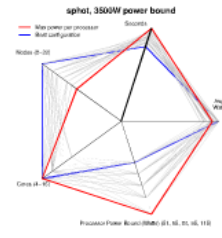
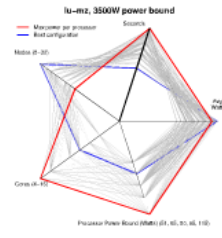
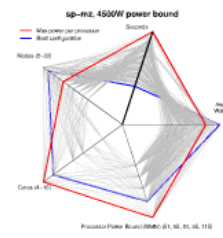
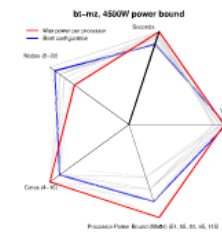
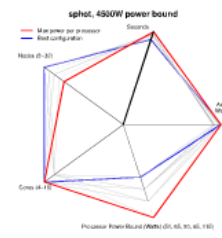
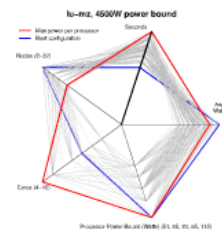
cap_rawio

BUG (we're working on it)

4500W

3500W

2500W



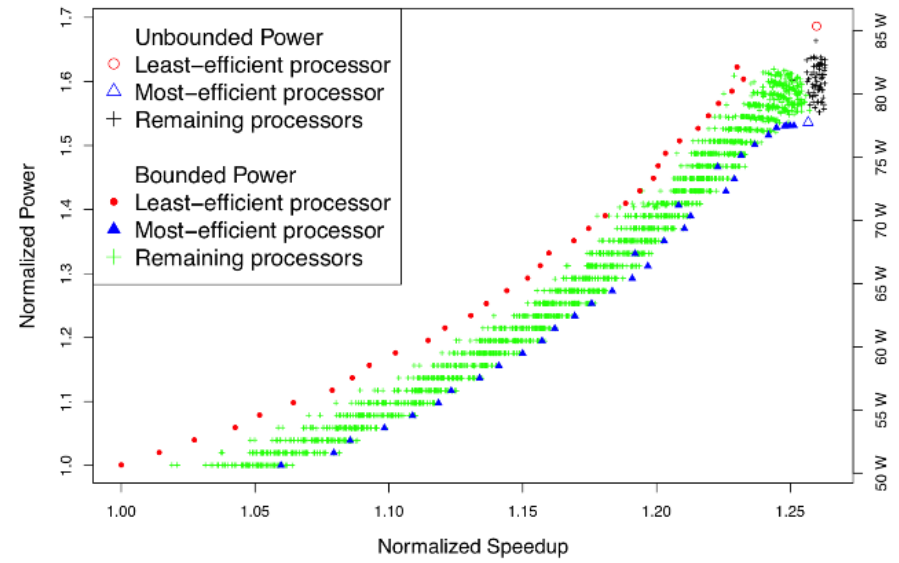
lu-mz

sphot

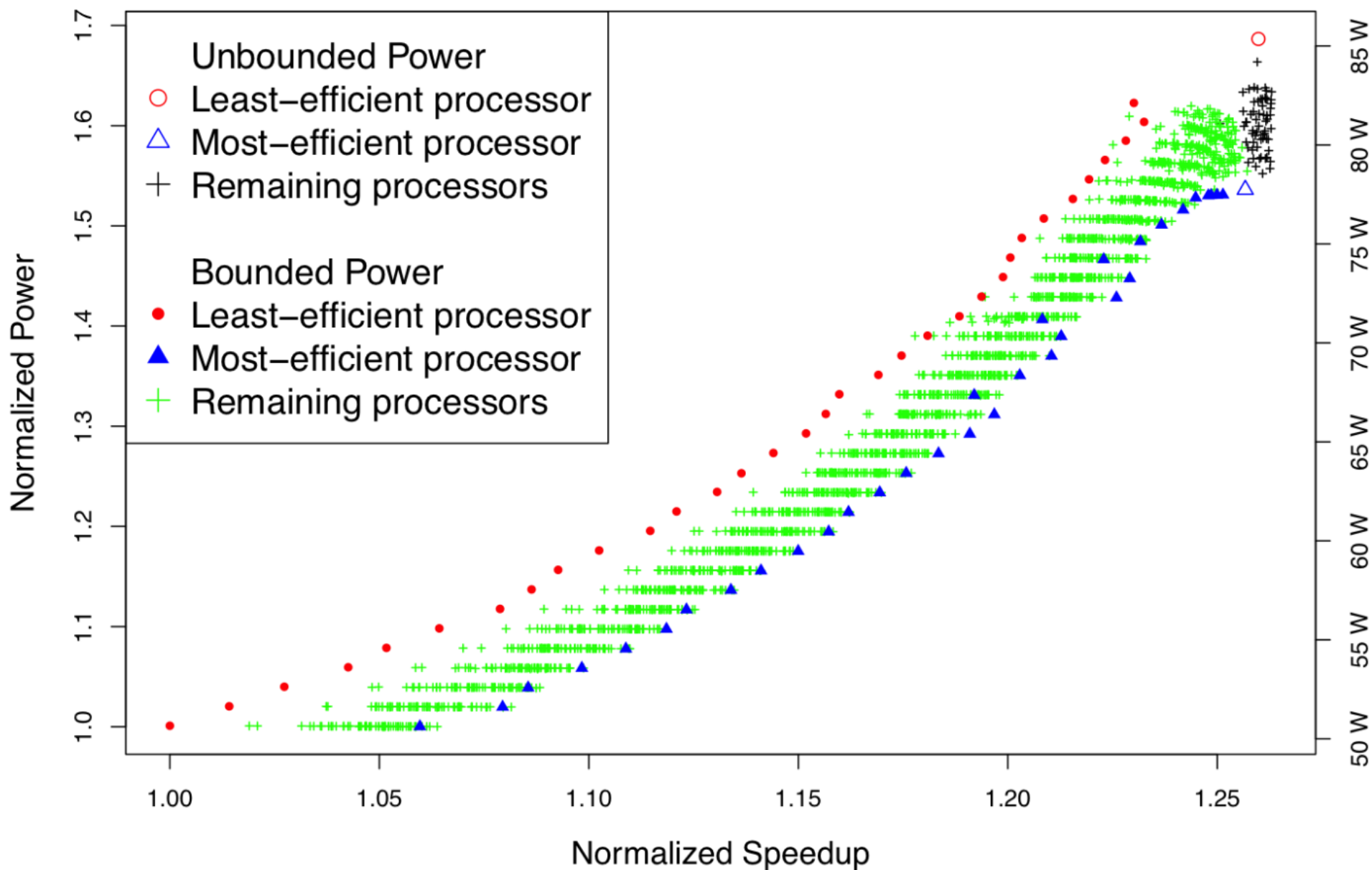
bt-mz

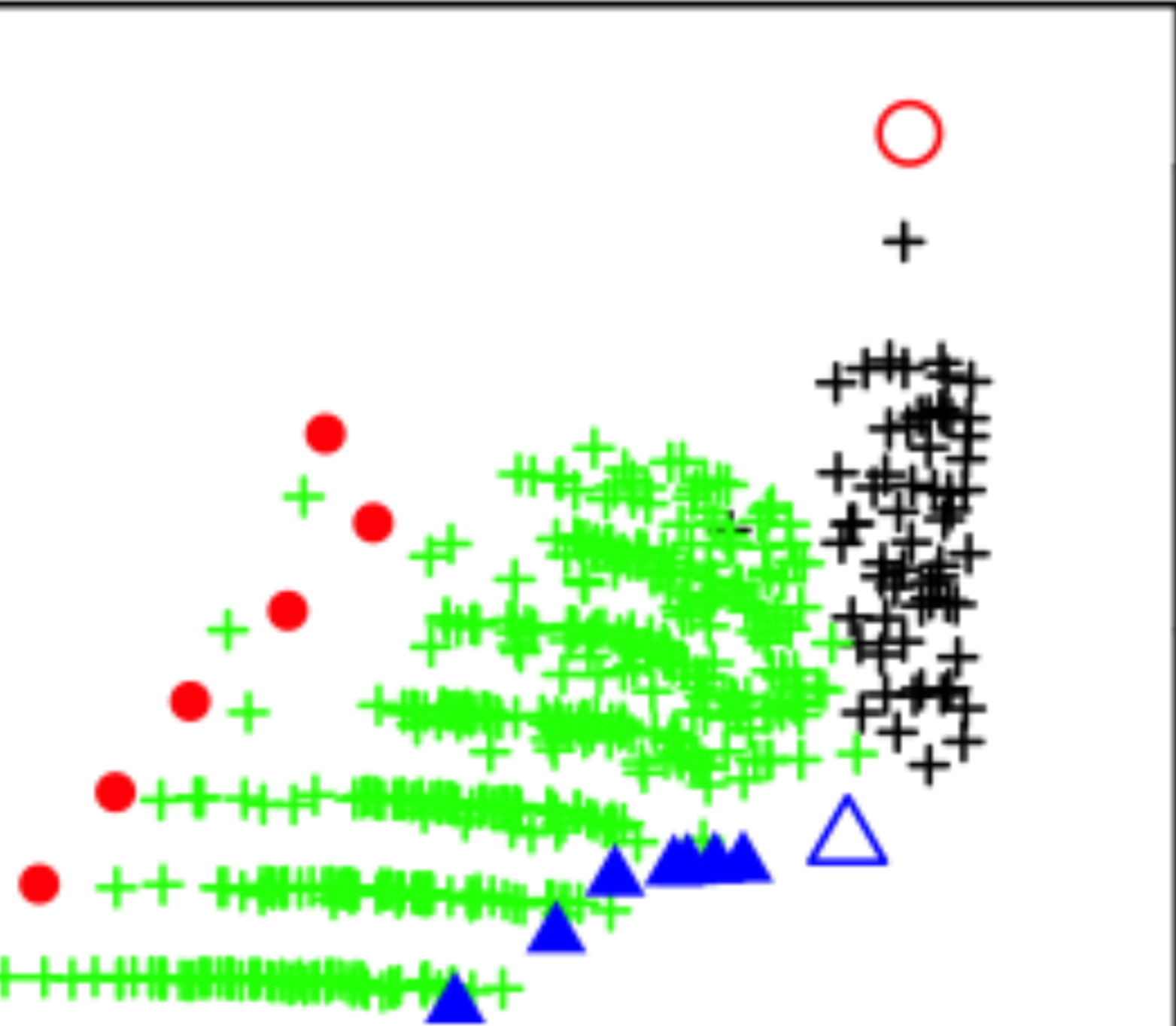
sp-mz

rzzin, mg.C.8 single processor over 64 processors with unbounded and 50 power bound



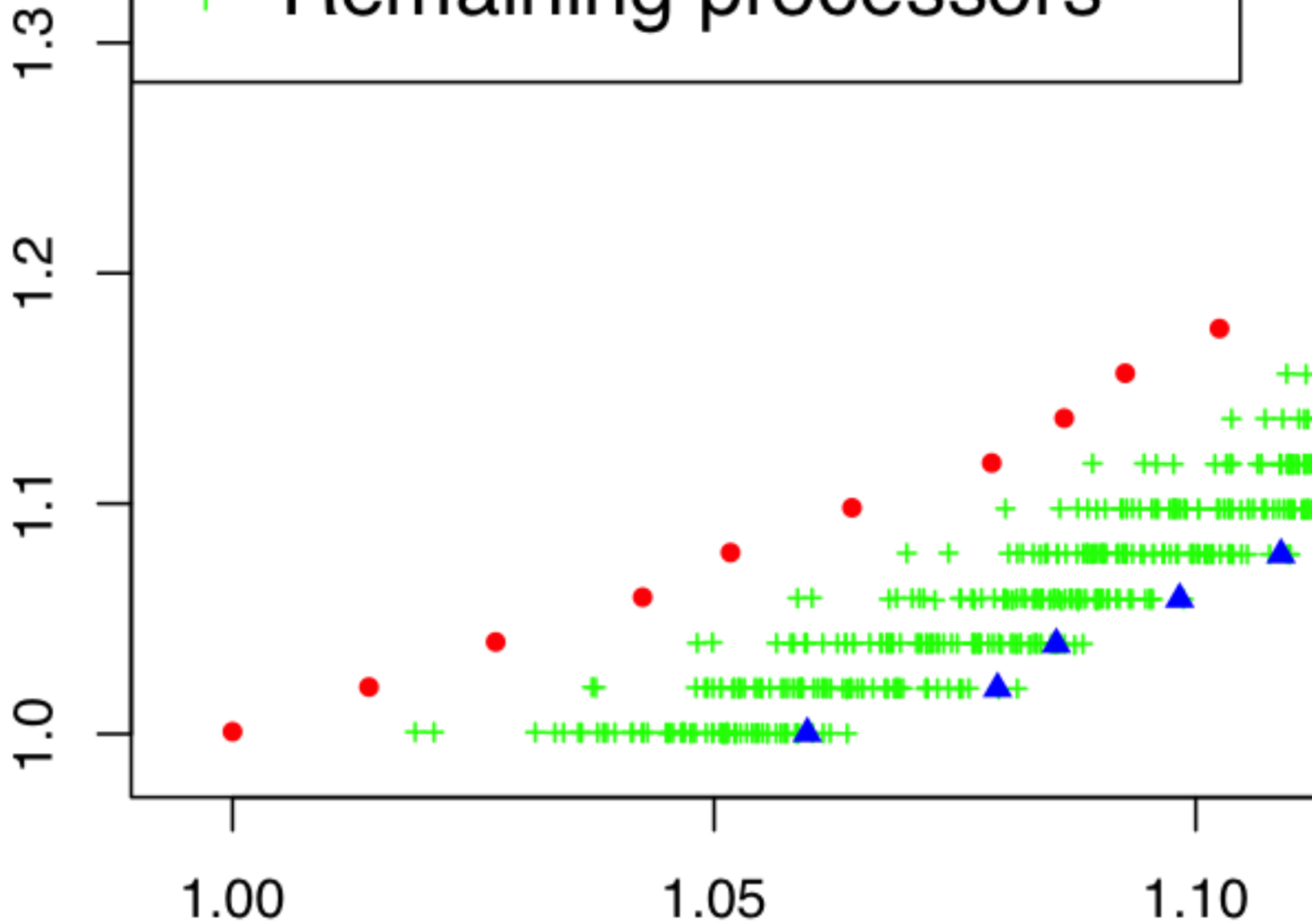
rzzin, mg.C.8 single processor over 64 processors with unbounded and 50 power bound



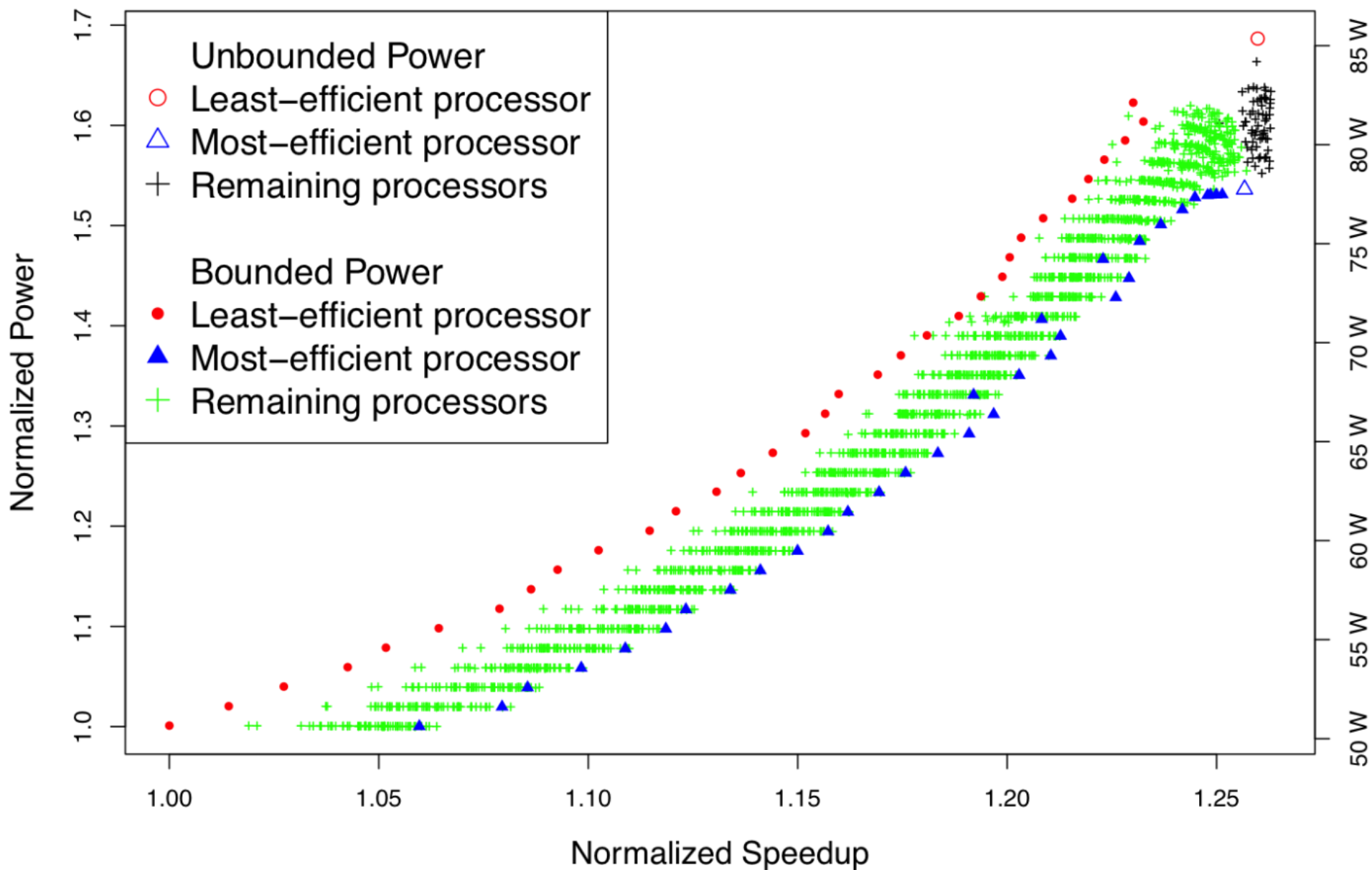


N 80 W 85 W

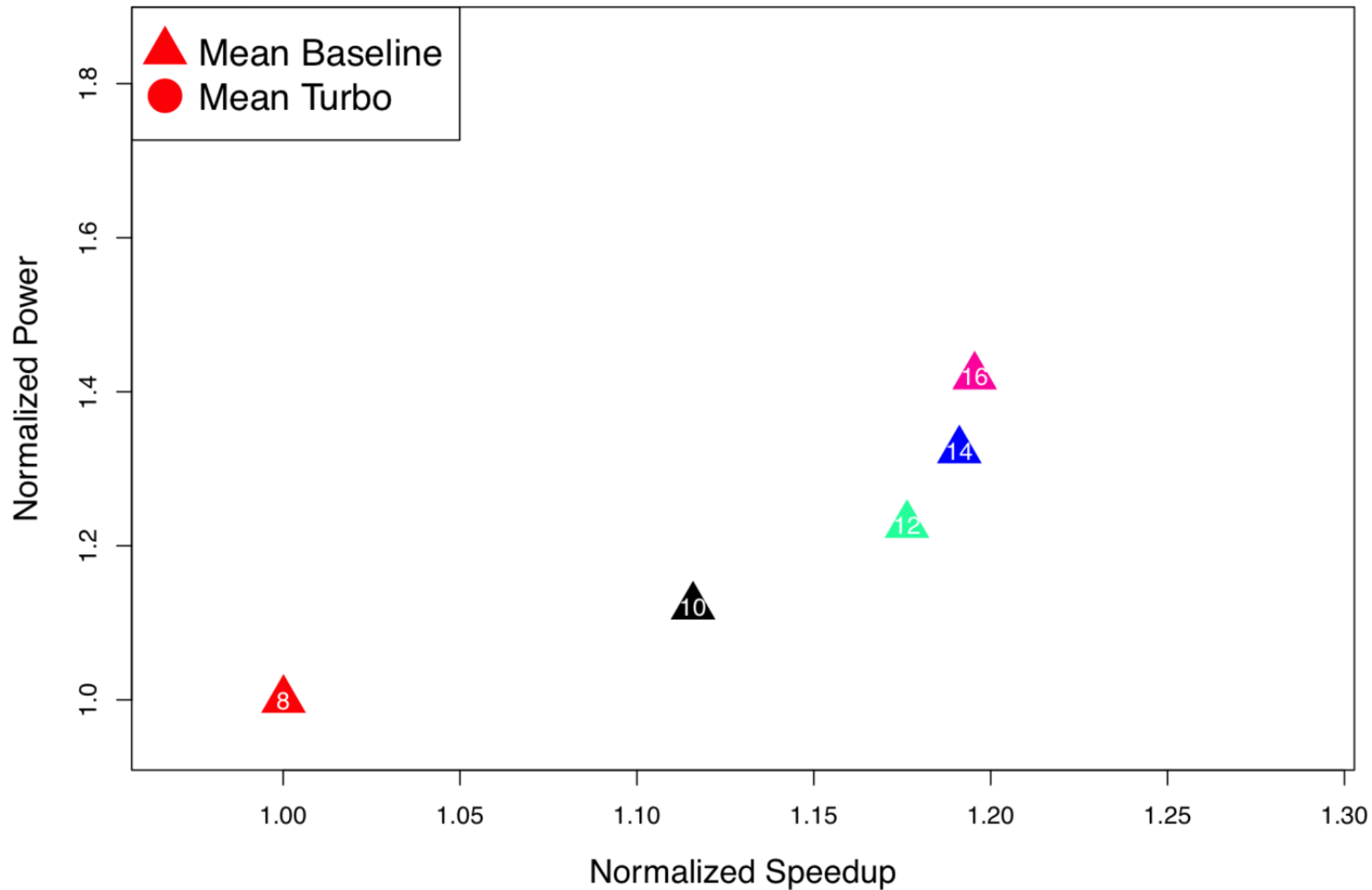
+ Remaining processors



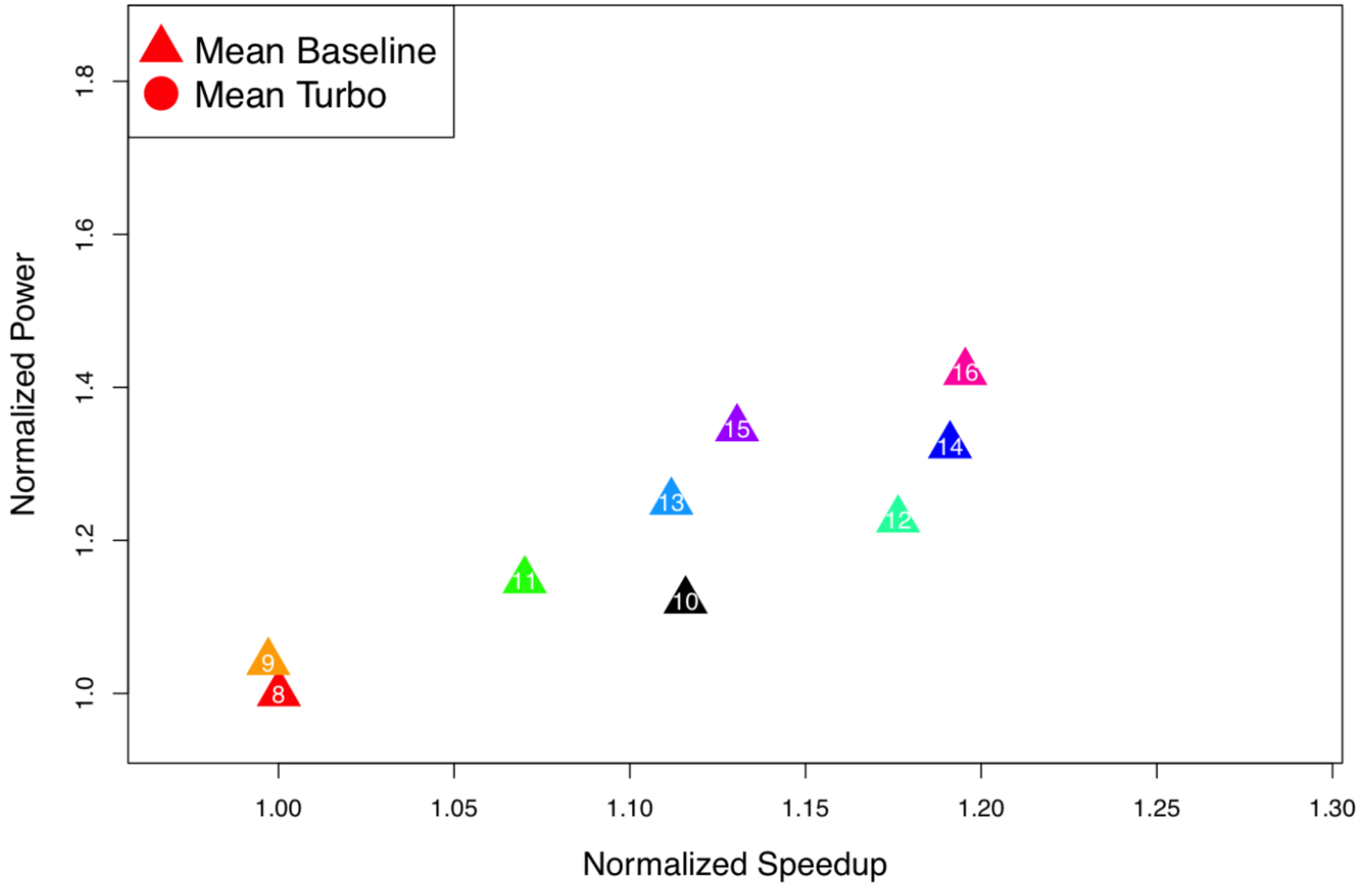
rzzin, mg.C.8 single processor over 64 processors with unbounded and 50 power bound



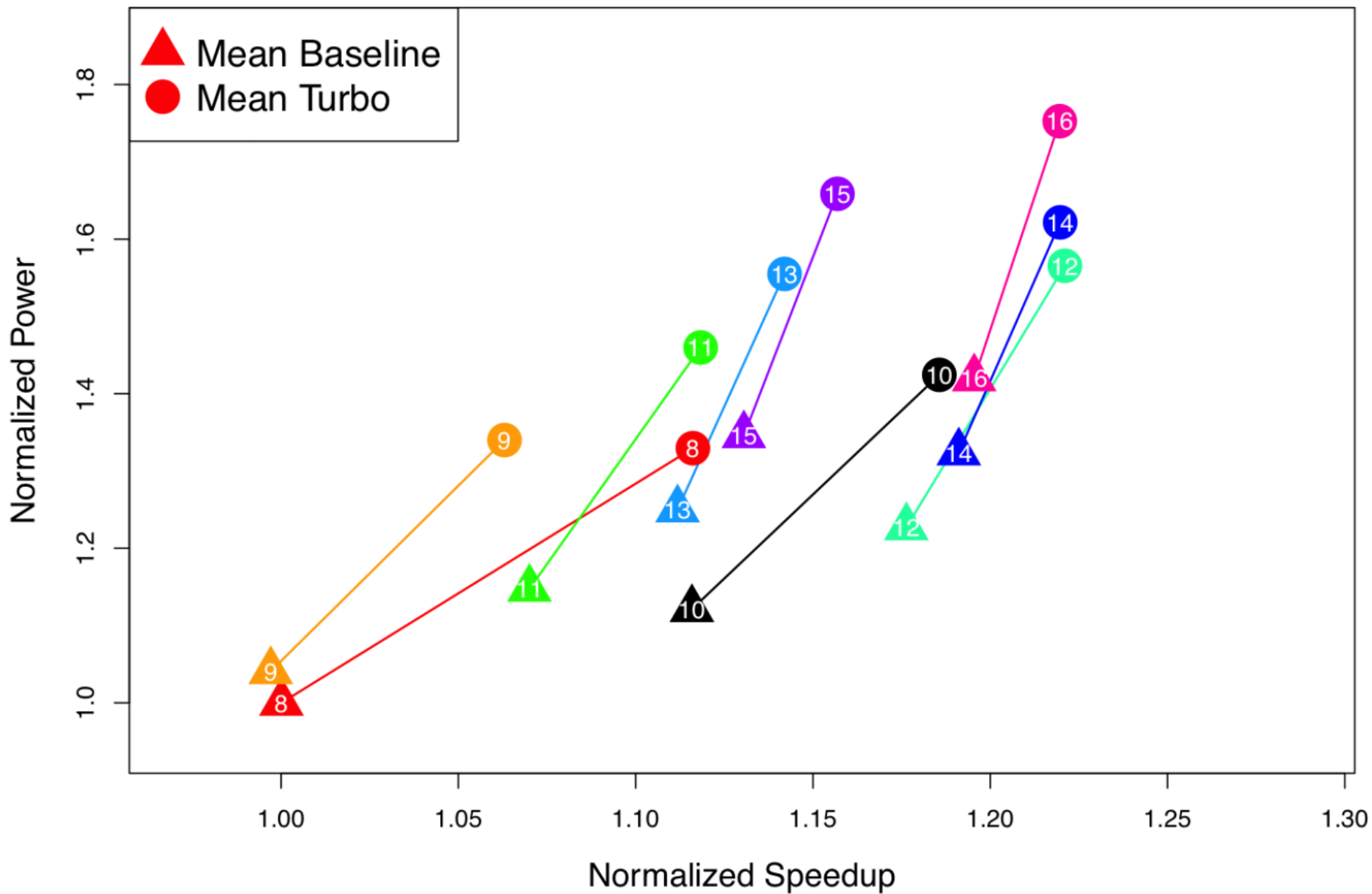
Turbo and Baseline, Single-Node mg.D, 8-16 OMP Threads, Scatter, ~175 Cab Nodes



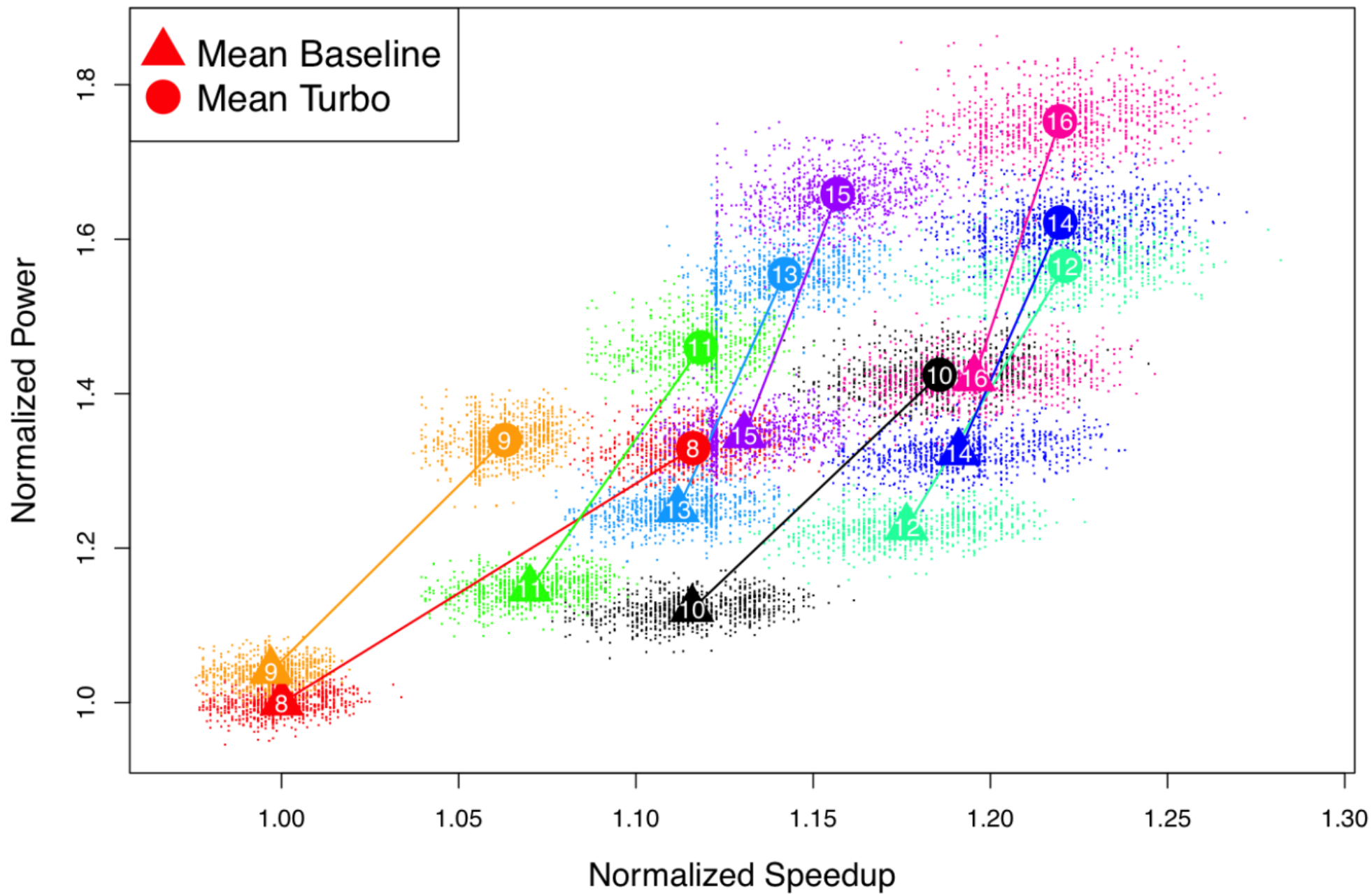
Turbo and Baseline, Single-Node mg.D, 8-16 OMP Threads, Scatter, ~175 Cab Nodes



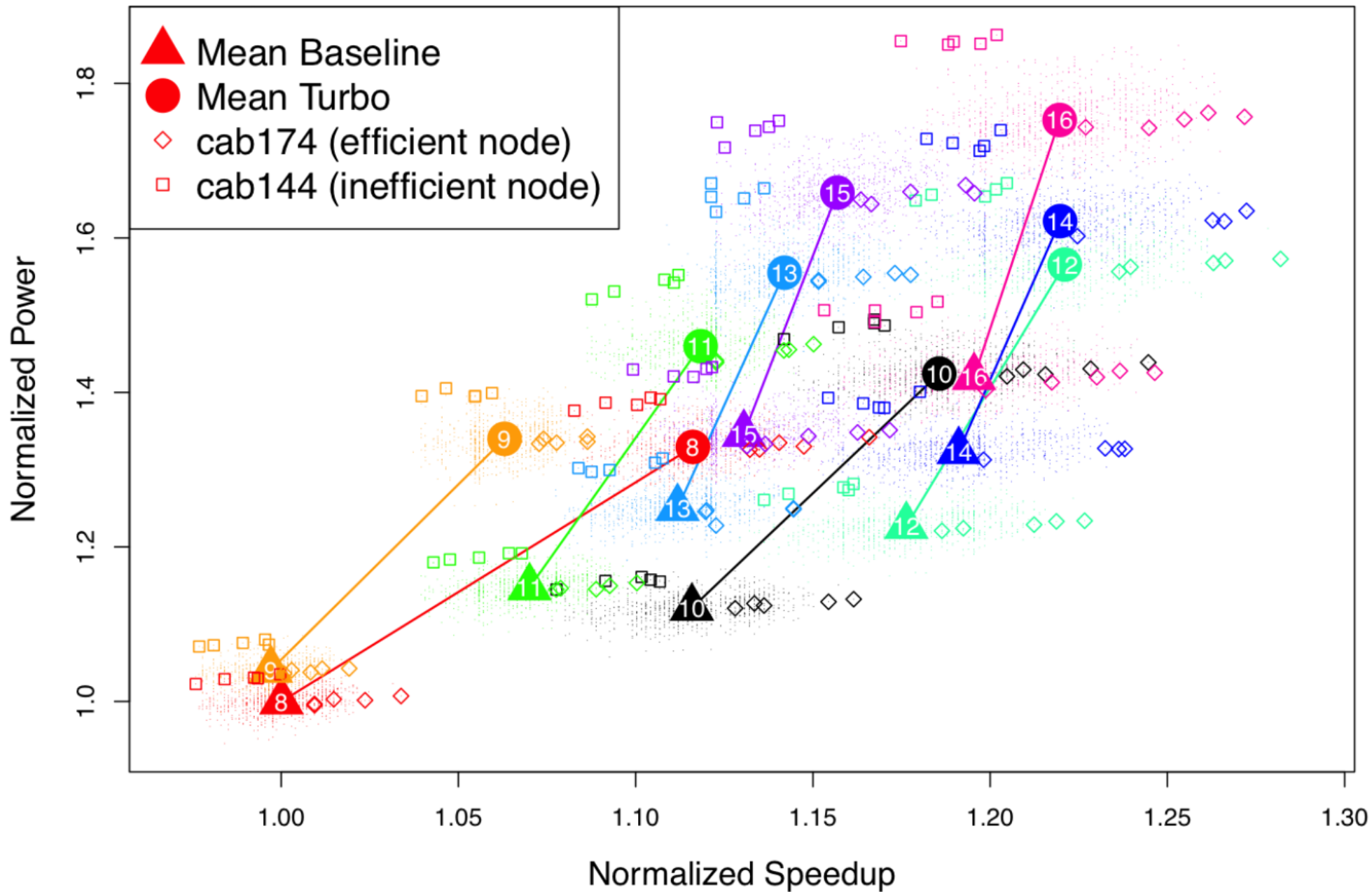
Turbo and Baseline, Single-Node mg.D, 8-16 OMP Threads, Scatter, ~175 Cab Nodes

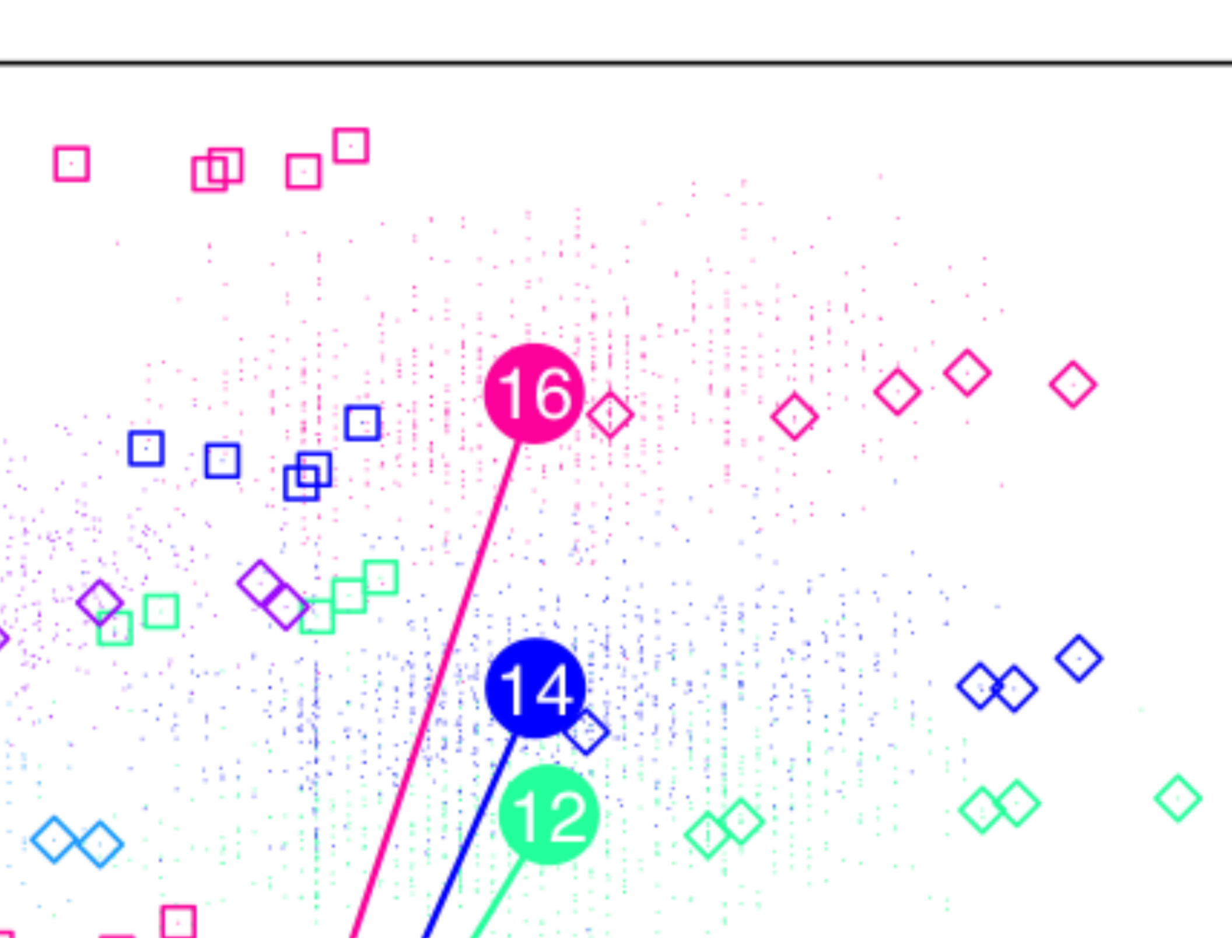


Turbo and Baseline, Single-Node mg.D, 8-16 OMP Threads, Scatter, ~175 Cab Nodes



Turbo and Baseline, Single-Node mg.D, 8-16 OMP Threads, Scatter, ~175 Cab Nodes





Linpack Socekt0 Temperatures
3 Linpack Tasks & 1 Read Task

Current features

RAPL

Thermal

PEBS

Turbo

Clocks

CPUID

Future plans

Uncore counters

Clock gating

Challenges

Development of safe_msr kernel module

Maintenance of out-of-kernel module

