SC23
Denver, CO | i am hpc.

SPEChpc 2021 Benchmarks on Ice Lake and Sapphire Rapids InfiniBand Clusters: A Performance and Energy Case Study

Ayesha Afzal, Georg Hager, Gerhard Wellein

Erlangen National High Performance Computing Center (NHR@FAU)
Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

NHR FAU

FAU Friedrich-Alexander-Universität Erlangen-Nürnberg

Overview

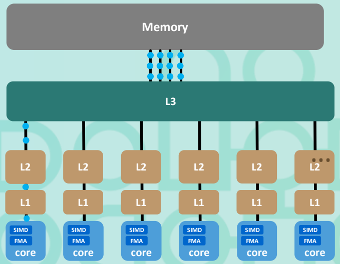
DOI: 10.1145/3624062.3624197

Overview
SPEChpc suite,
hardware and software

Overview

DOI: 10.1145/3624062.3624197

Overview
SPEChpc suite,
hardware and software



Single-ccNUMA domain and single-node analysis

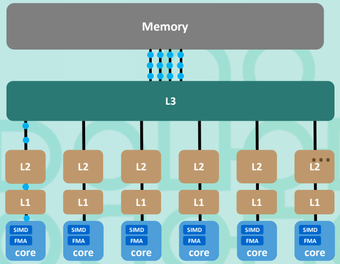
tiny data set

Multi-node analysis

small data set

Overview

DOI: 10.1145/3624062.3624197



Overview
SPEChpc suite,
hardware and software

Single-ccNUMA domain and single-node analysis

tiny data set

Multi-node analysis

small data set

Performance

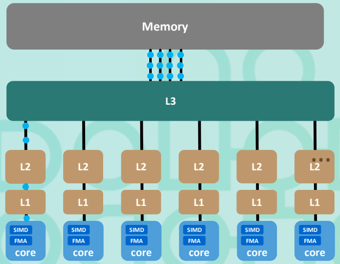
Power

Energy

Overview

DOI: 10.1145/3624062.3624197

Overview
SPEChpc suite,
hardware and software



Single-ccNUMA domain and single-node analysis

tiny data set

- Performance
- Power
- Energy

Multi-node analysis

small data set

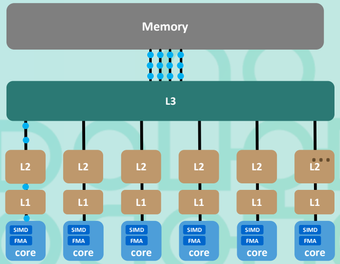
- Strong scaling performance
- Scaling impact on power
- Scaling impact on energy

Overview

DOI: 10.1145/3624062.3624197

Overview

Overview
SPEChpc suite,
hardware and software



- Single-ccNUMA domain and single-node analysis
  tiny data set
  - Performance
  - Power
  - Energy
- Multi-node analysis
  small data set
  - Strong scaling performance
  - Scaling impact on power
  - Scaling impact on energy

Key resource metrics to pinpoint bottlenecks and scaling issues

1. *Speedup*
2. *Memory Bandwidth*
3. *Data transfer volume*
4. *Scalability*

DOI: 10.1145/3624062.3624197

Overview of the SPEChpc suite and the benchmarking hardware

SEChpc 2021 Benchmarks

Benchmark	
lbm	Lattice-Boltzmann Method D2Q37 2D CFD solver
soma	Monte-Carlo acceleration for soft coarse grained polymers
TeaLeaf	Linear heat conduction equation on a 2D regular grid using a 5-point stencil with implicit solvers
CloverLeaf	Solving compressible Euler equations on a 2D Cartesian grid using an explicit second-order accurate method
MiniSweep	A successor to the well-known Sweep3D benchmark
Pot3D	Computing potential field solutions using a preconditioned CG sparse solver for the Laplace equation in 3D spherical coordinates
SPH-Exa	Smoothed Particle Hydrodynamics, a meshless Lagrangian method
HPGMGFV	Finite-volume-based High Performance Geometric Multigrid solving variable-coefficient elliptic problems on Cartesian grids
(mini)Weather	A traditional finite-volume control flow

Hardware overview

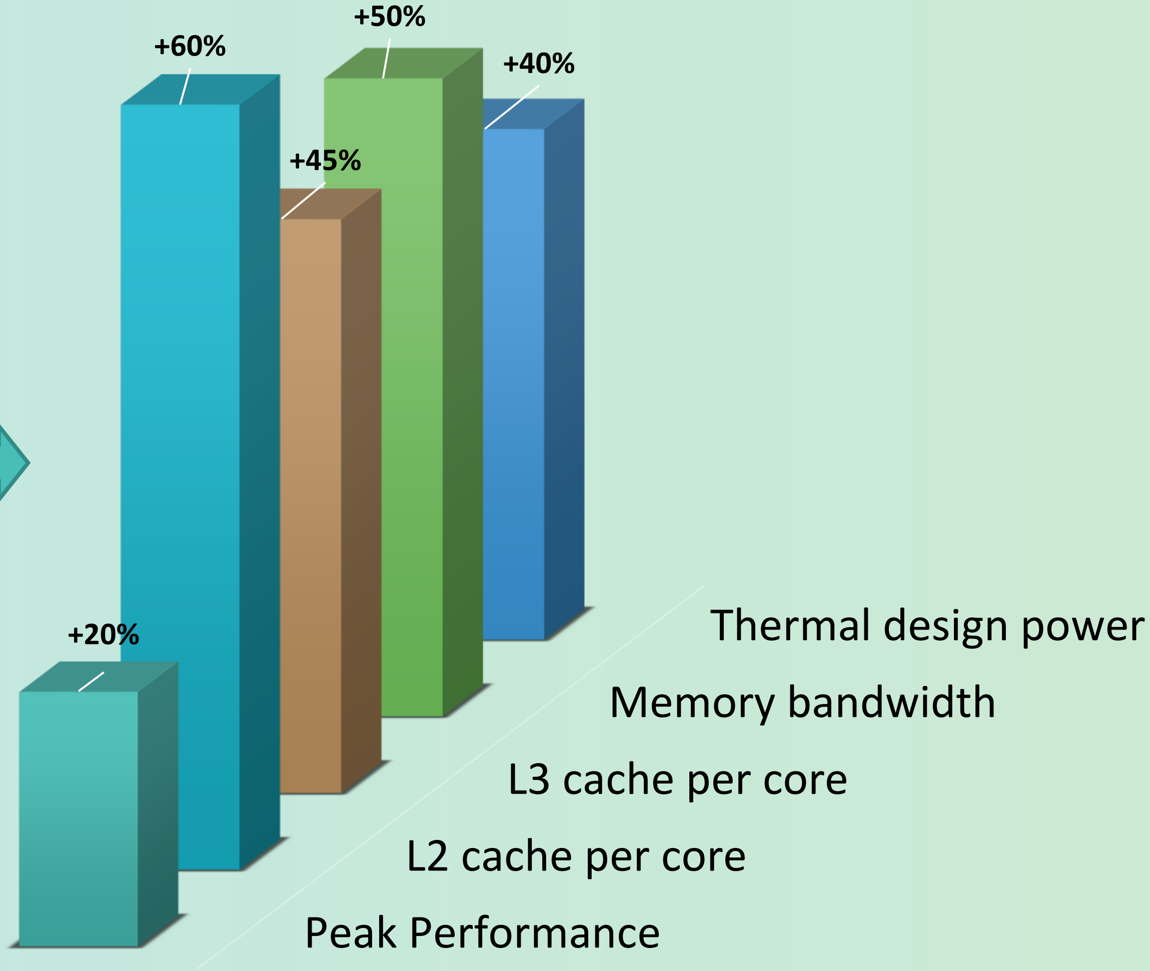
	Systems	Ice Lake cluster	Sapphire Rapids cluster
Micro-architecture	Intel Processor	Xeon Ice Lake	Xeon Sapphire Rapids
	Processor Model	Platinum 8360Y	Platinum 8470
	Base clock speed	2.4 GHz	2.0 GHz
	Physical cores per node	72	104
	ccNUMA domains per node	4	8
	Sockets per node	2	2
	Per-core L1/L2 cache	48 KiB (L1) + 1.25 MiB (L2)	48 KiB (L1) + 2 MiB (L2)
	Shared LLC	54 MiB (L3)	105 MiB (L3)
	Memory per node	4×64 GiB	8×128 GiB
	Socket memory type	8 channels DDR4-3200	8 channels DDR5-4800
	Theor. socket memory bandwidth	2×102.4 GB/s	4×76.8 GB/s
	Thermal design power	250 W	350 W

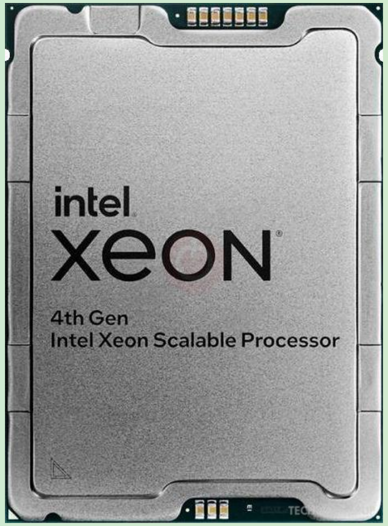
Hardware overview

	Systems	Ice Lake cluster	Sapphire Rapids cluster
Micro-architecture	Intel Processor	Xeon Ice Lake	Xeon Sapphire Rapids
	Processor Model	Platinum 8360Y	Platinum 8470
	Base clock speed	2.4 GHz	2.0 GHz
	Physical cores per node	72	104
	ccNUMA domains per node	4	8
	Sockets per node	2	2
	Per-core L1/L2 cache	48 KiB (L1) + 1.25 MiB (L2)	48 KiB (L1) + 2 MiB (L2)
	Shared LLC	54 MiB (L3)	105 MiB (L3)
	Memory per node	4×64 GiB	8×128 GiB
	Socket memory type	8 channels DDR4-3200	8 channels DDR5-4800
	Theor. socket memory bandwidth	2×102.4 GB/s	4×76.8 GB/s
	Thermal design power	250 W	350 W

Ice Lake vs. Sapphire Rapids

Ice Lake cluster

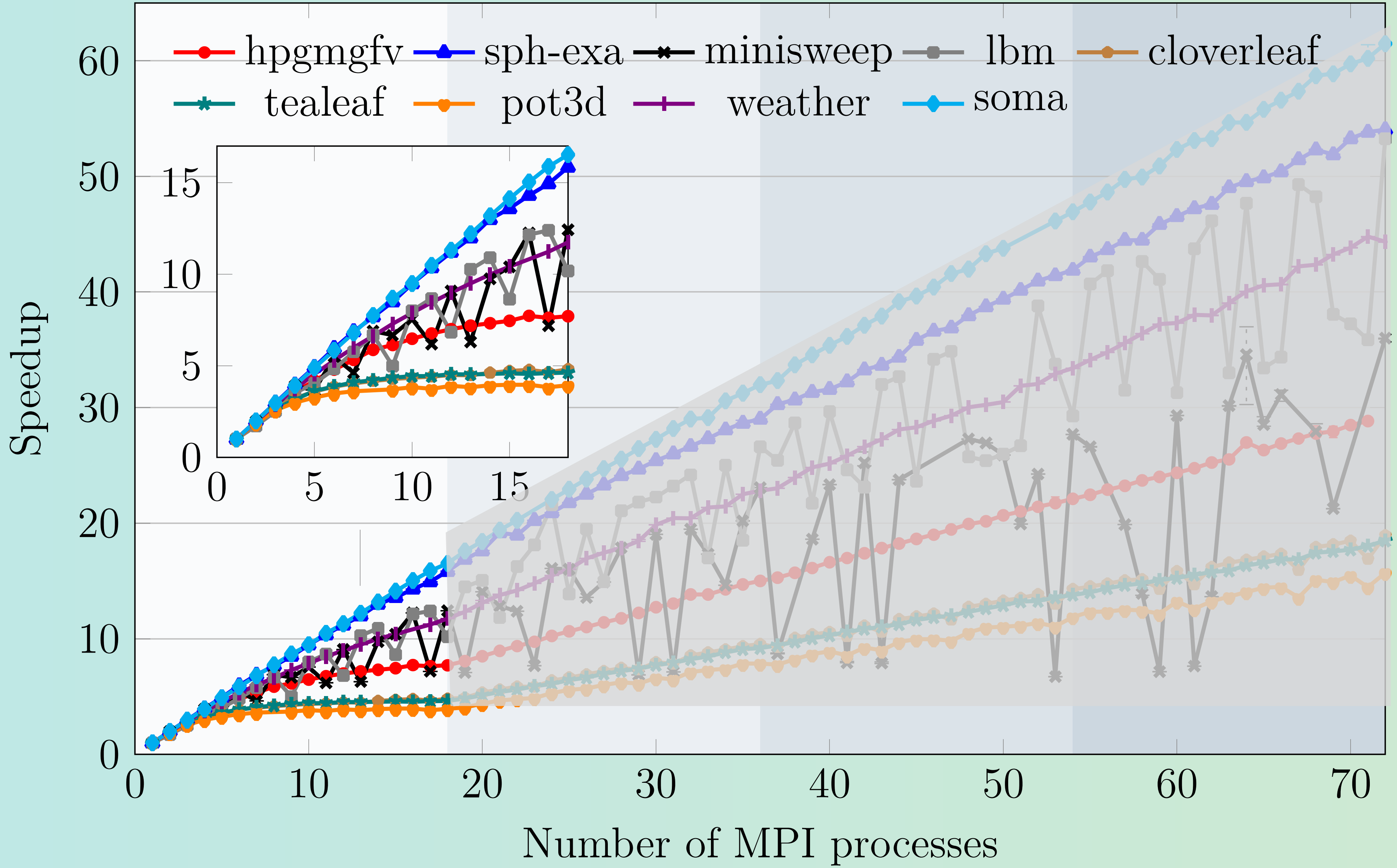




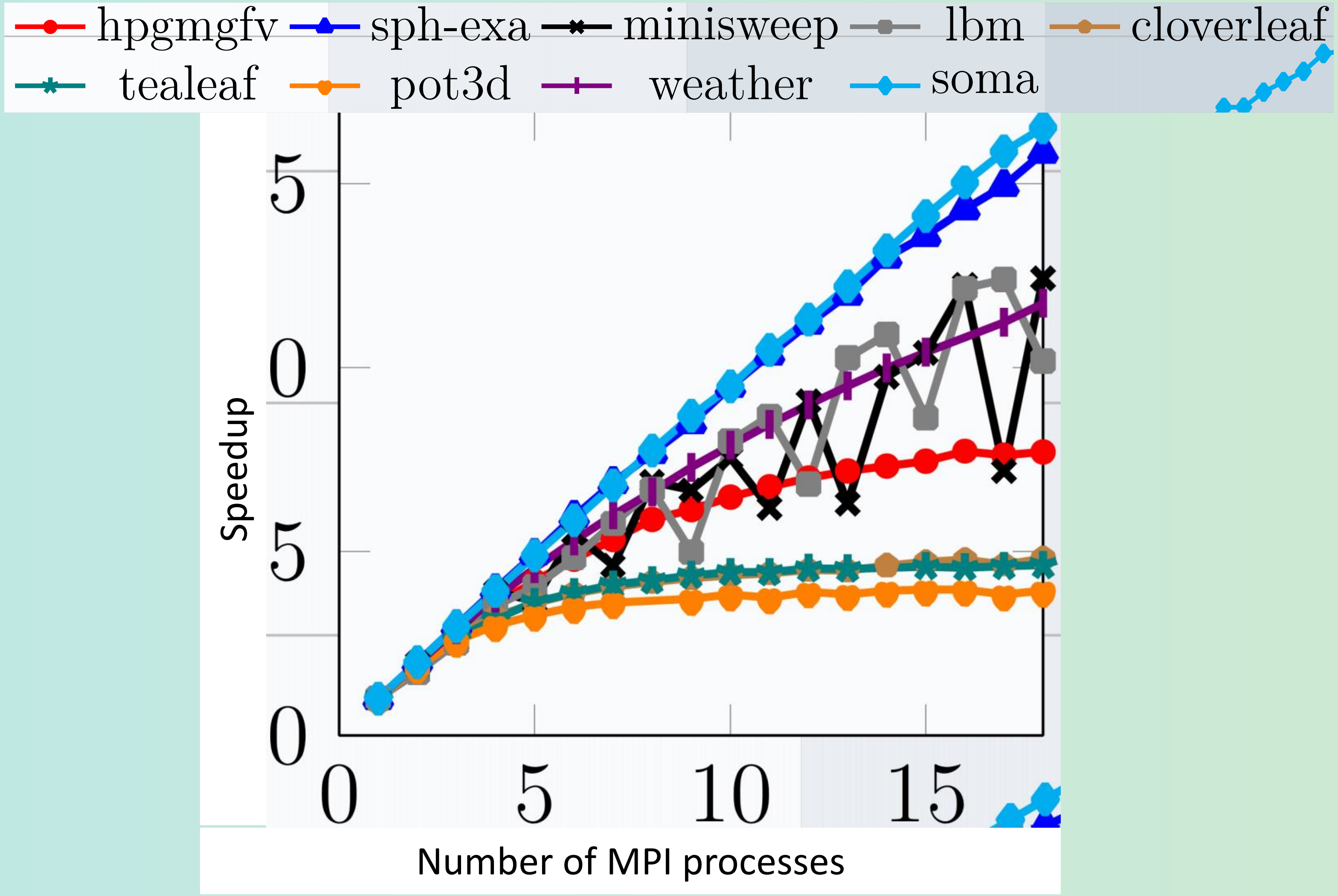
Sapphire Rapids cluster

Node-level Performance Analysis

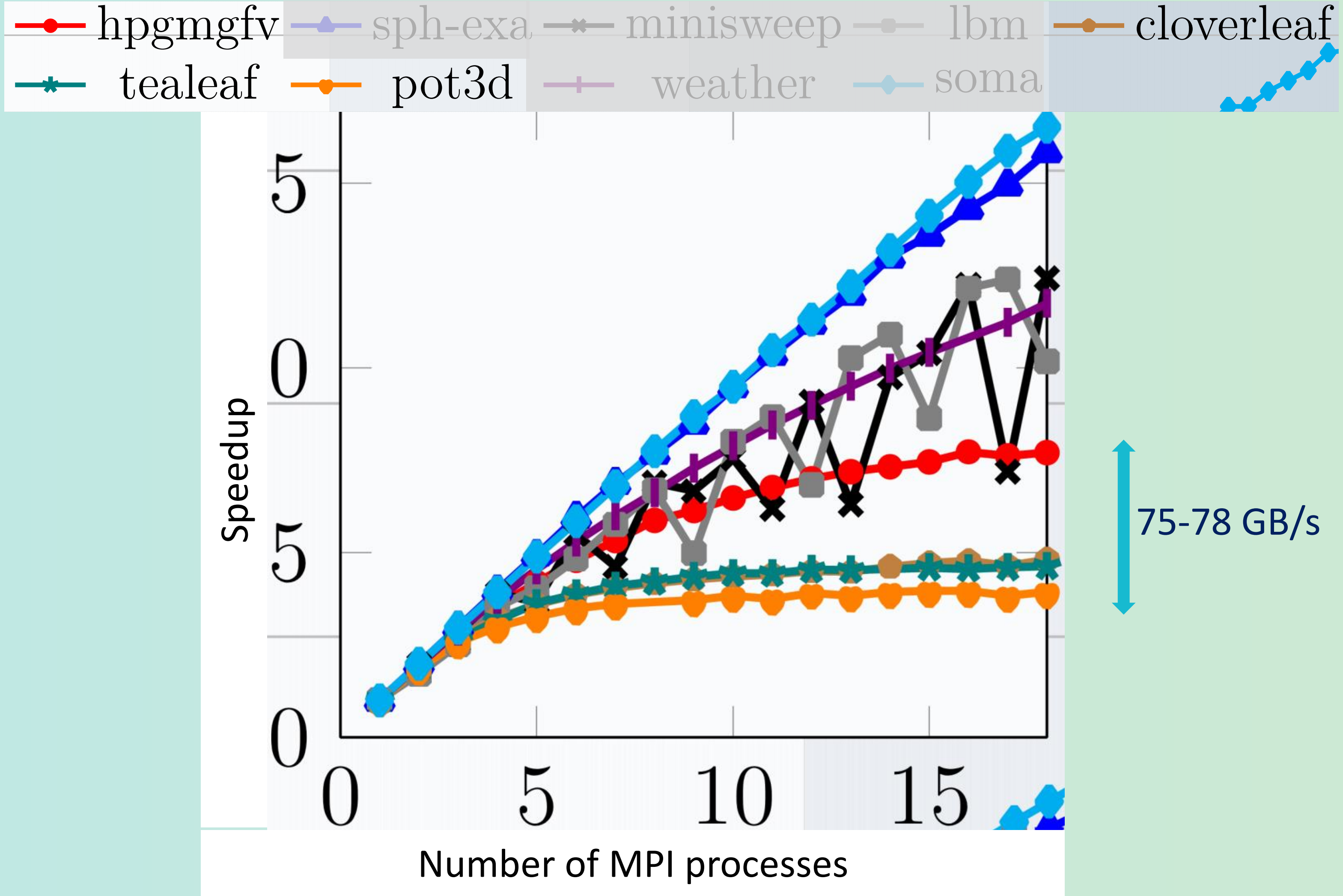
Speedup: Ice Lake



Speedup: Ice Lake



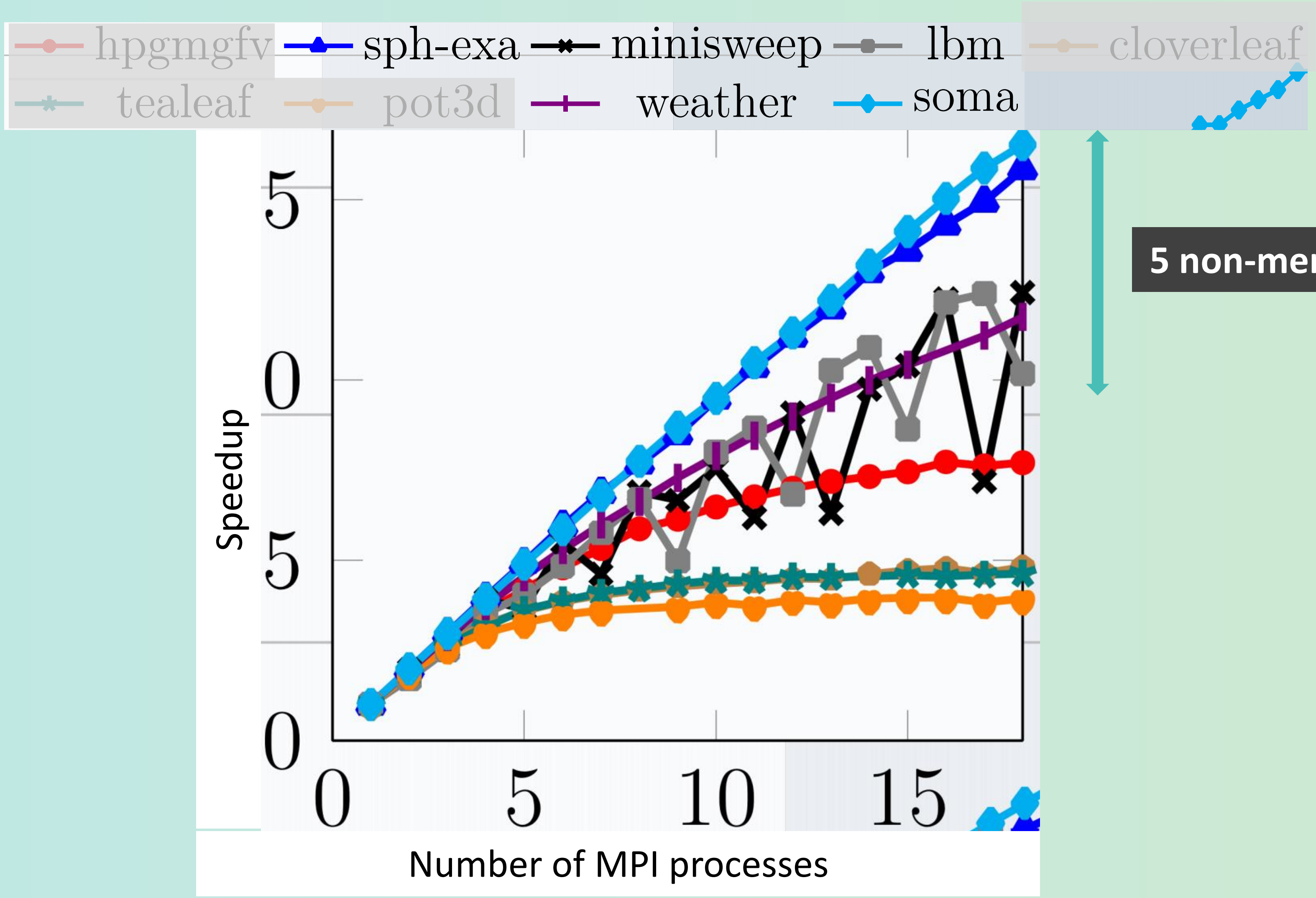
Speedup: Ice Lake



4 memory-bound codes

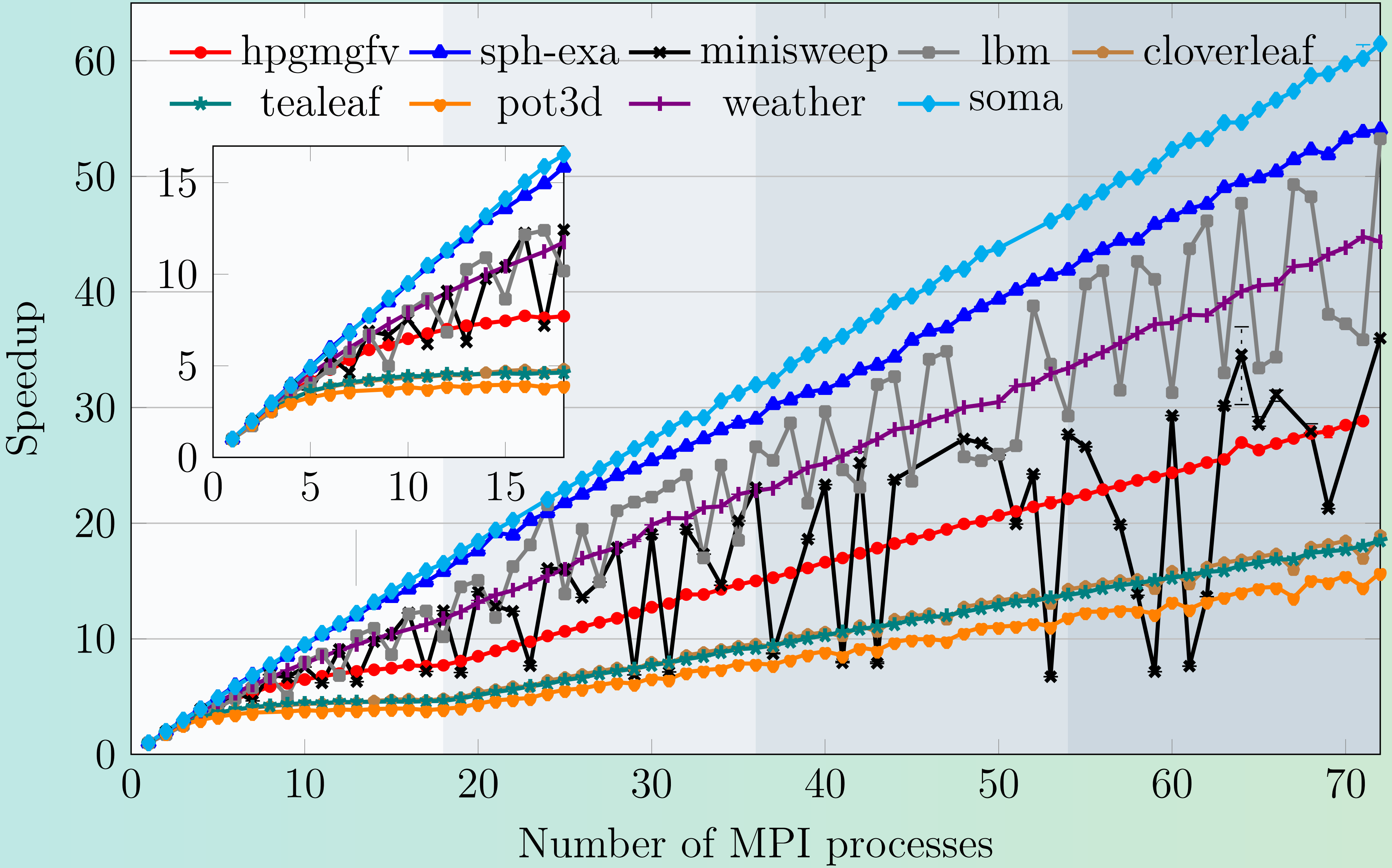
speedup approaching a limit across the cores of a ccNUMA domain

Speedup: Ice Lake

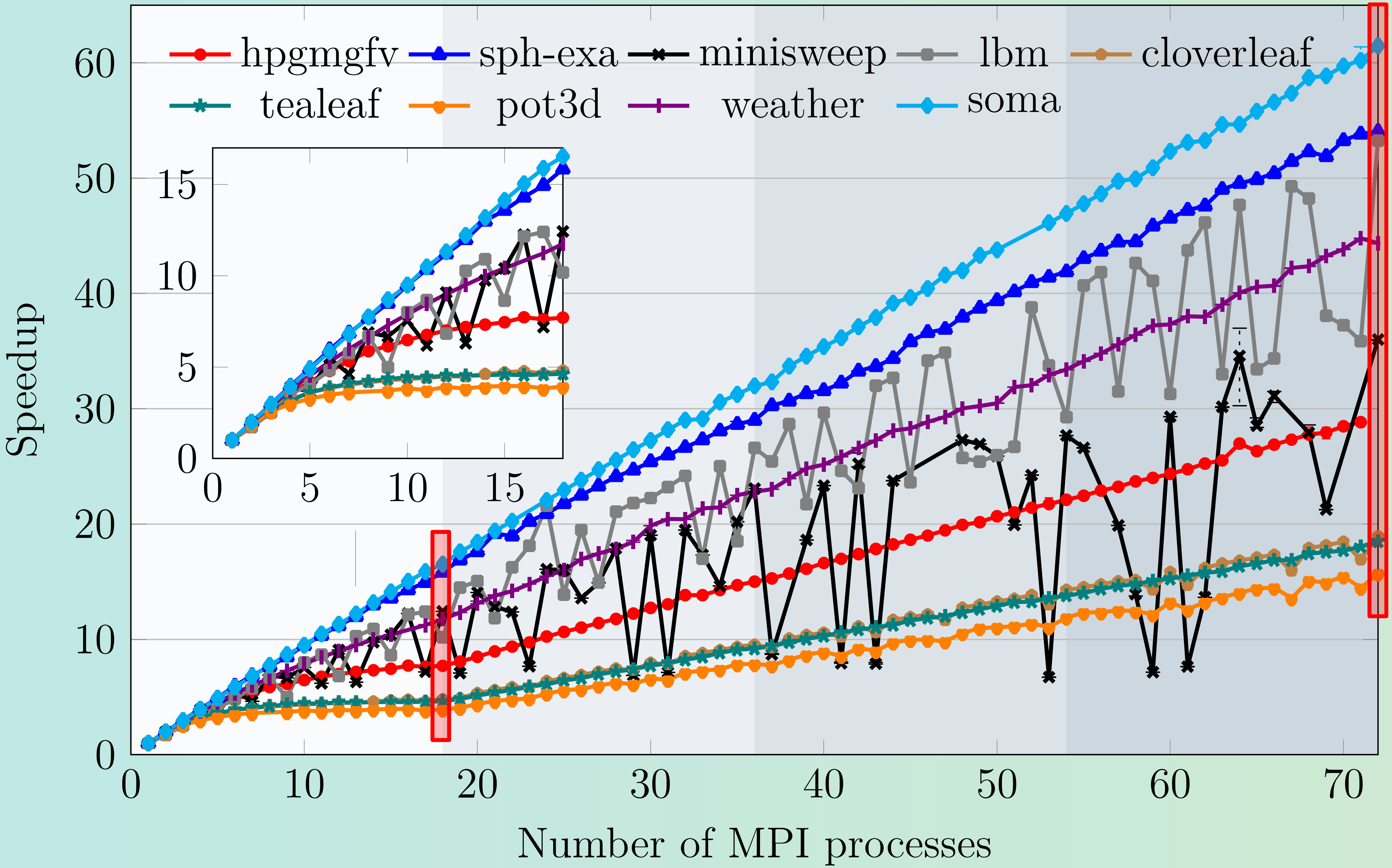


5 non-memory-bound codes

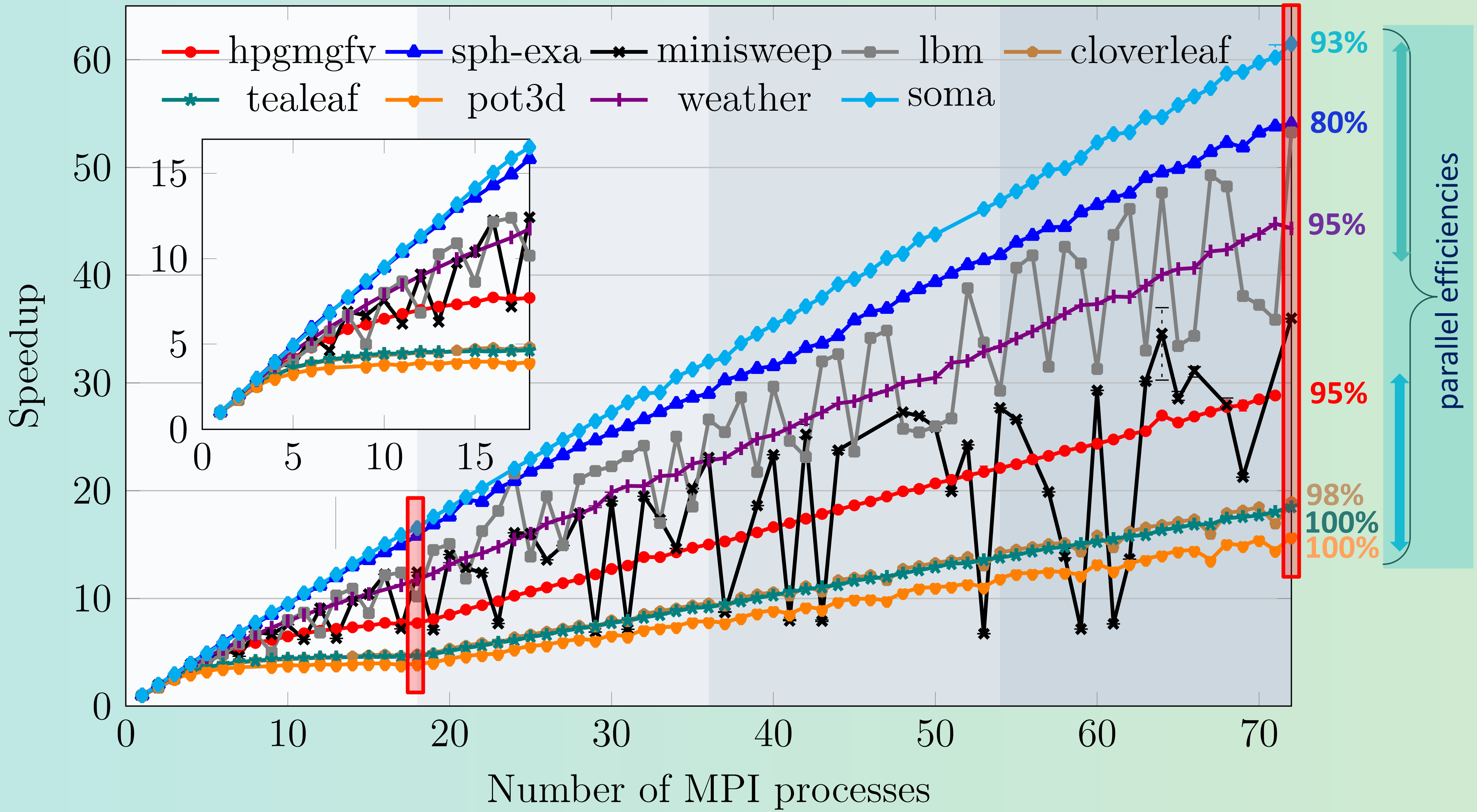
Speedup: Ice Lake



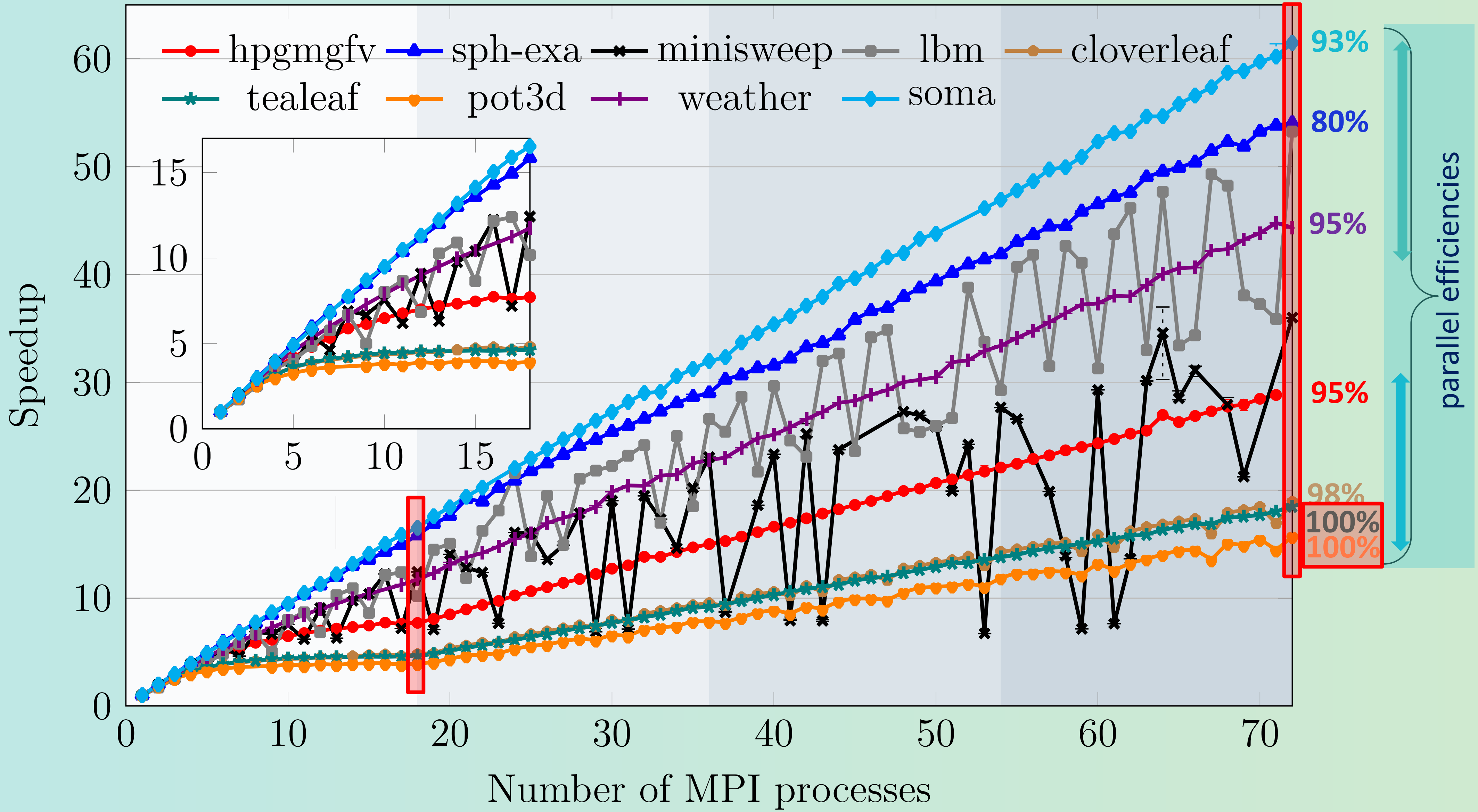
Speedup: Ice Lake



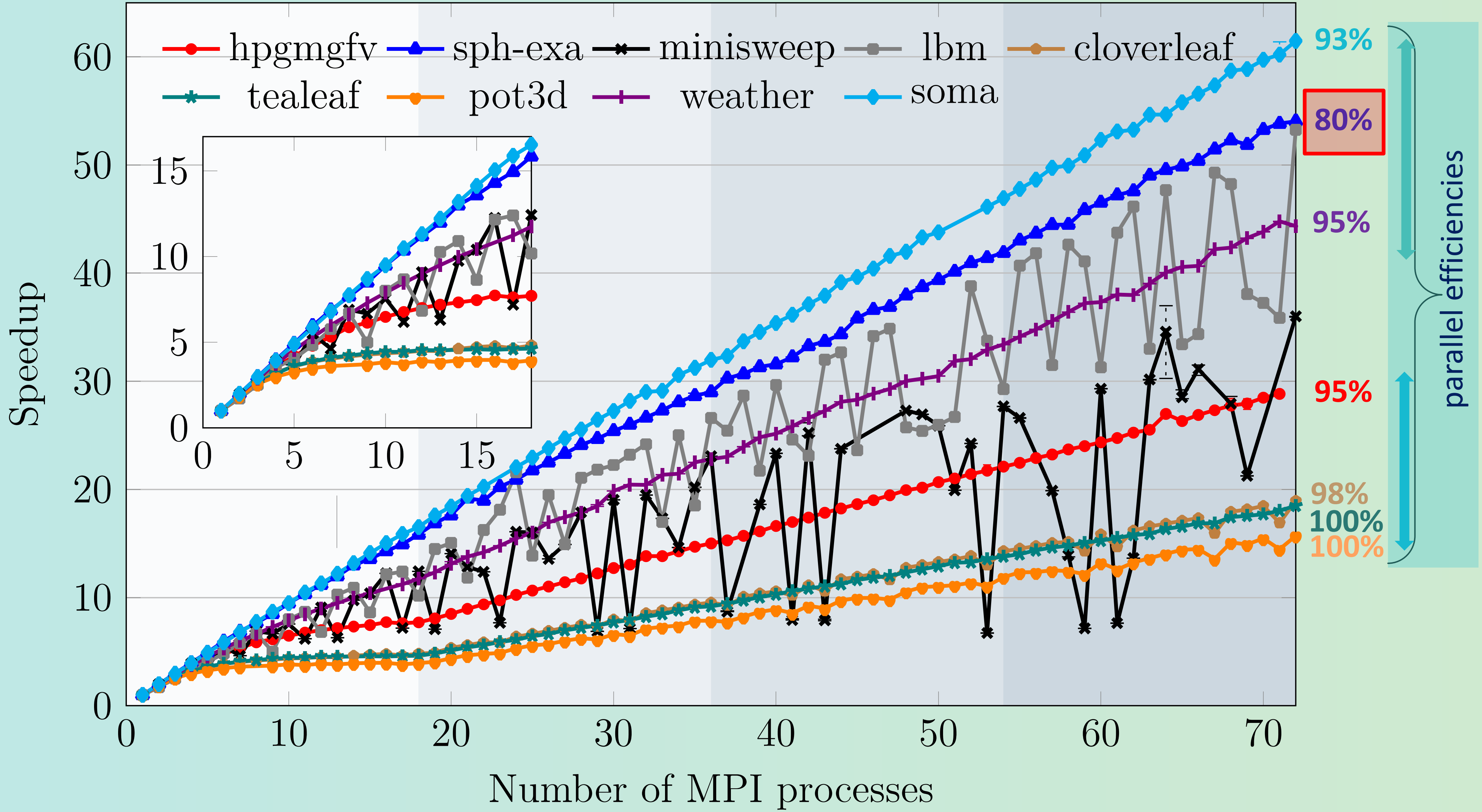
Speedup: Ice Lake



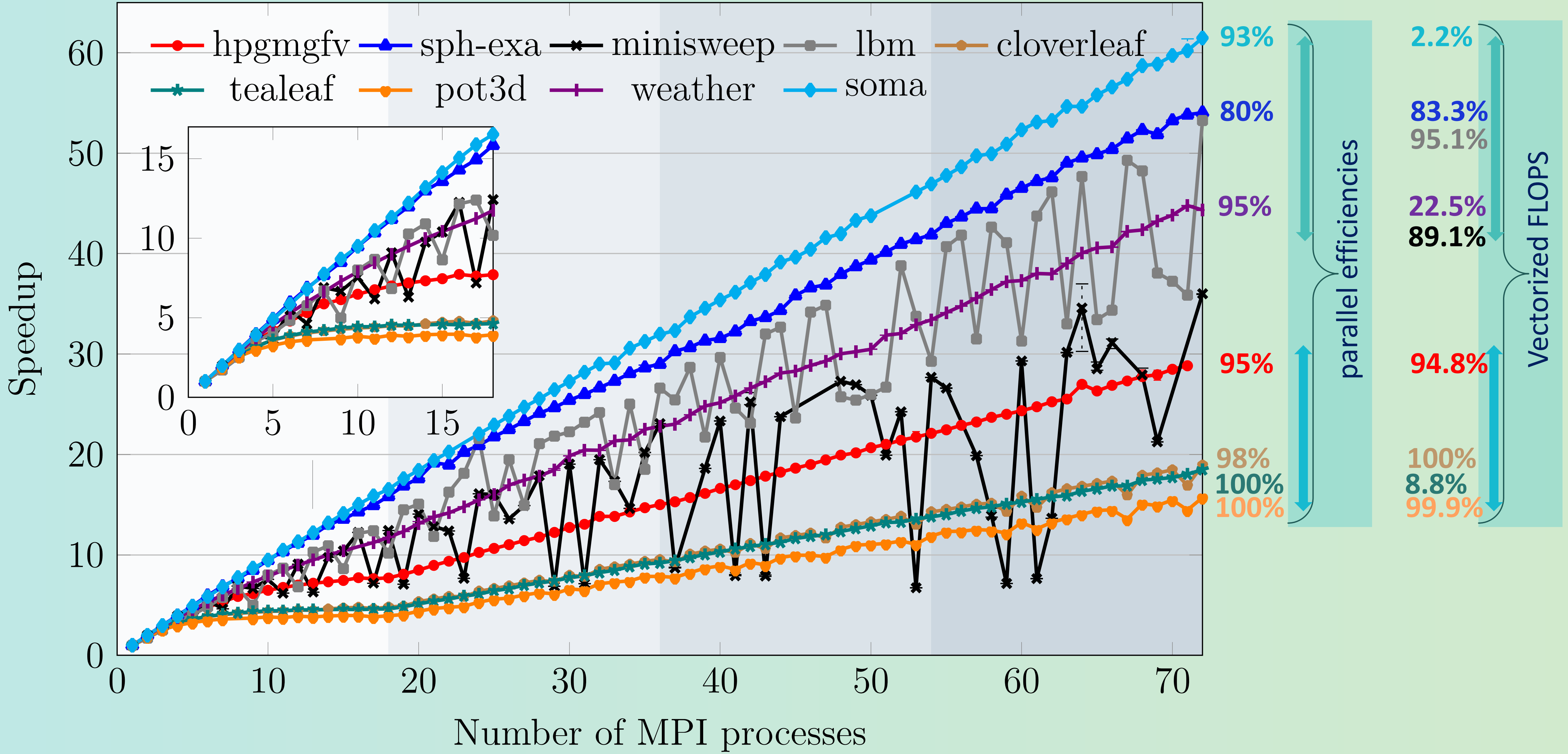
Speedup: Ice Lake



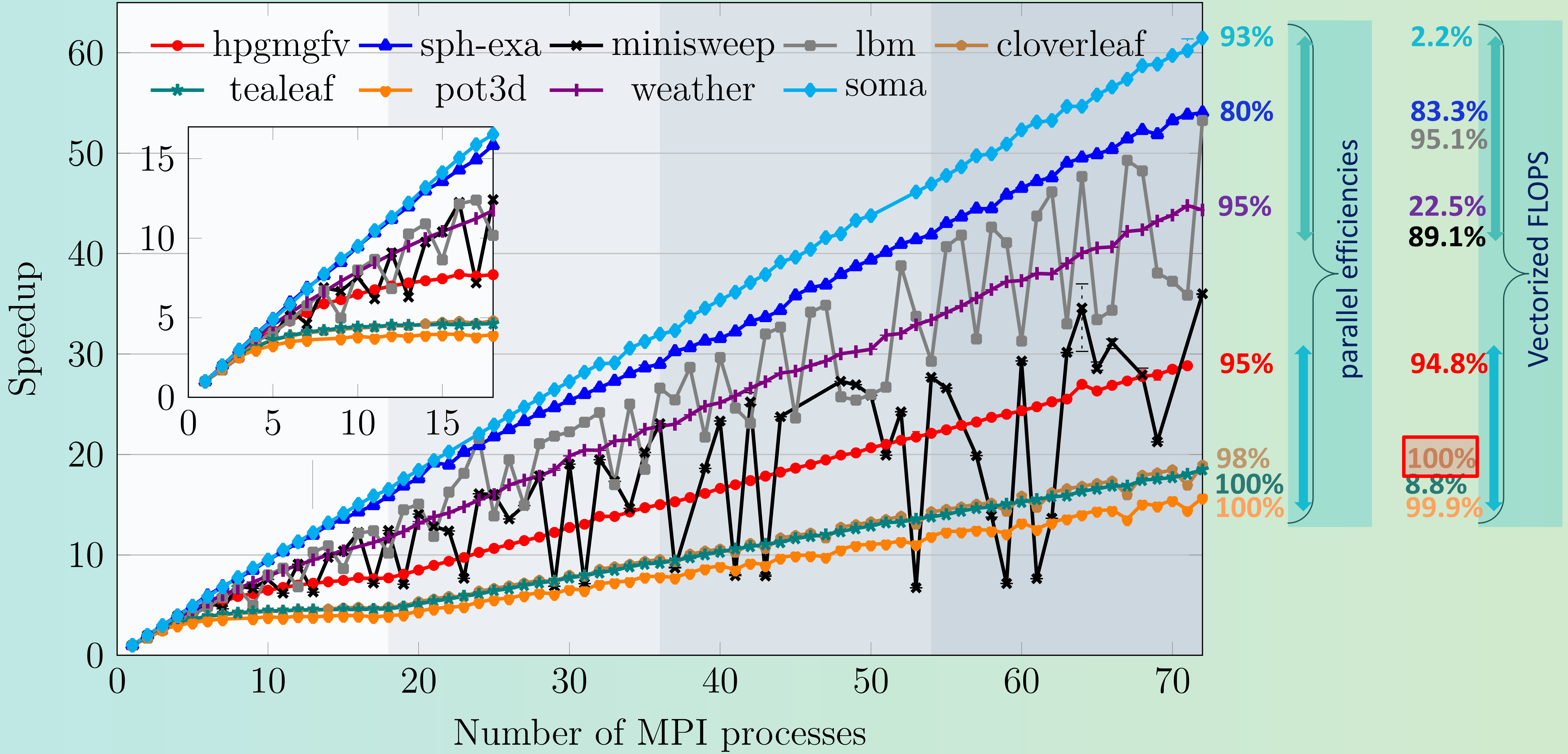
Speedup: Ice Lake



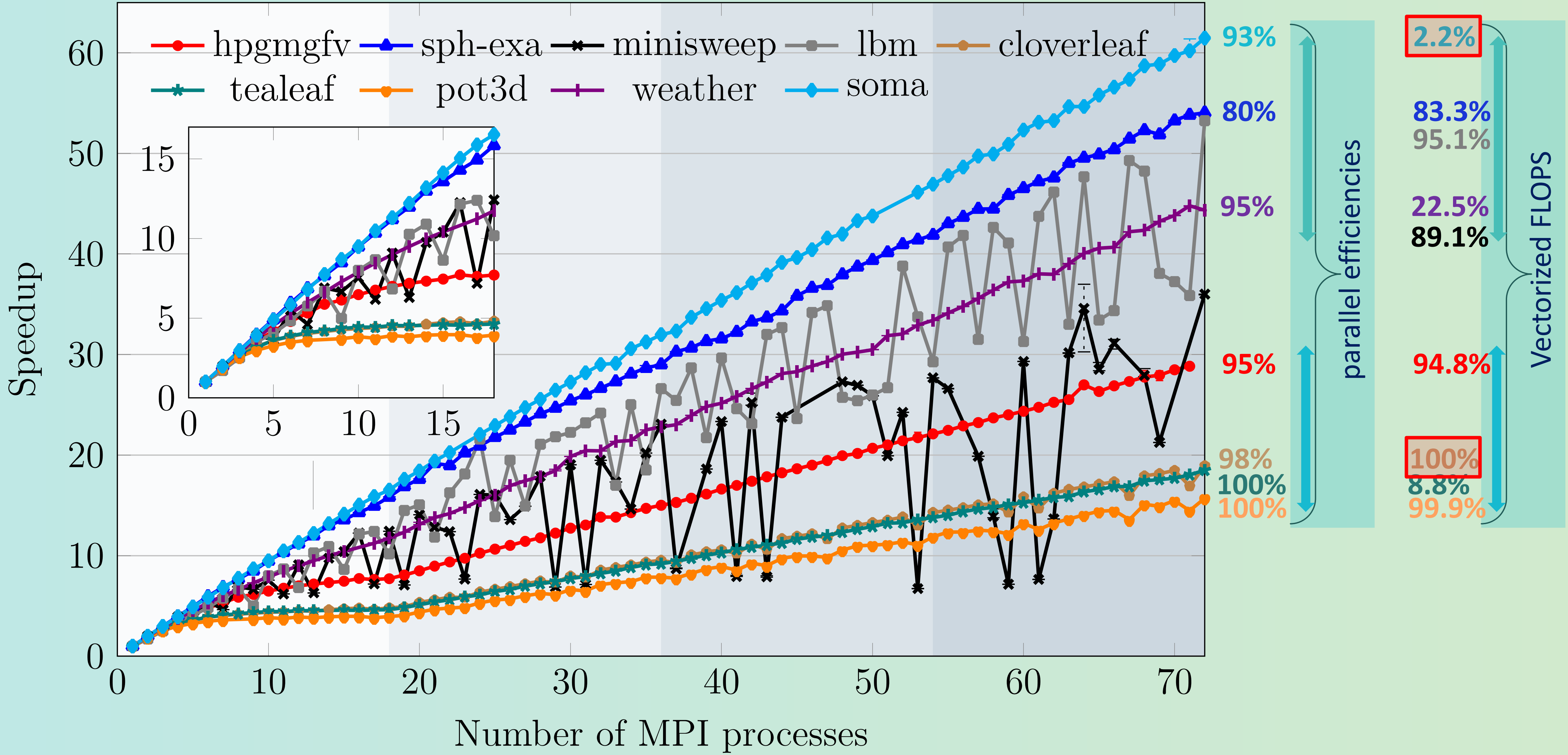
Speedup: Ice Lake



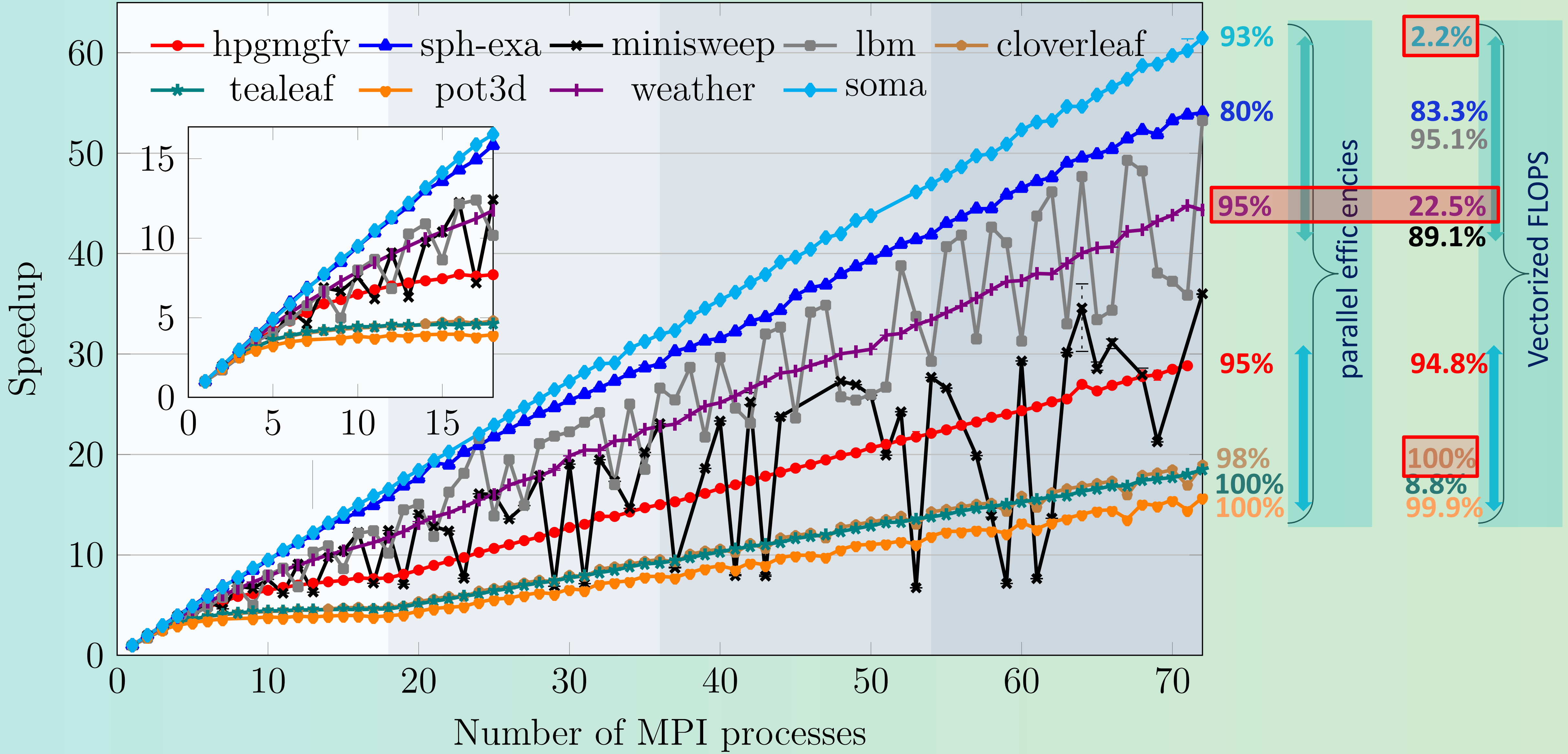
Speedup: Ice Lake



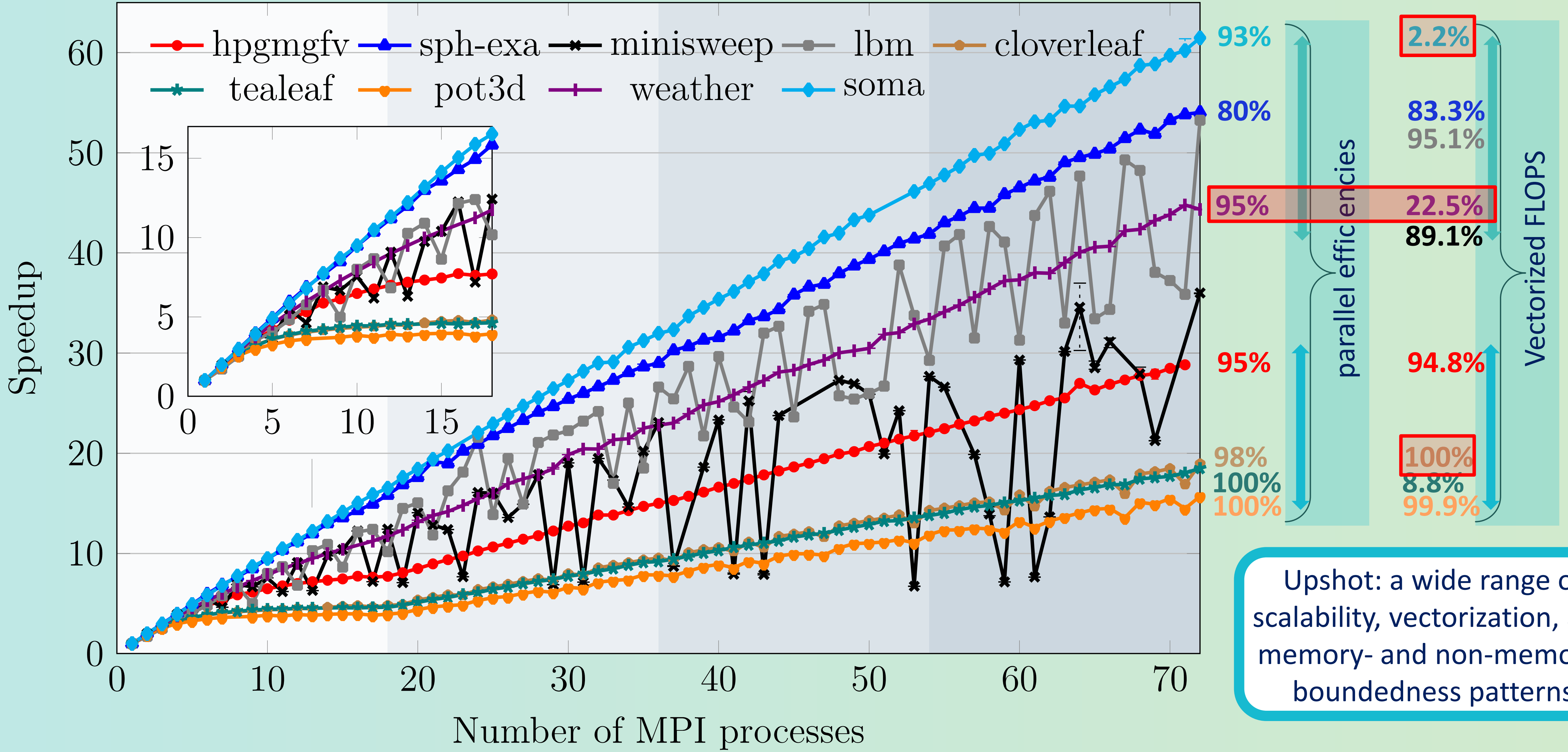
Speedup: Ice Lake



Speedup: Ice Lake



Speedup: Ice Lake



Upshot: a wide range of scalability, vectorization, and memory- and non-memory-boundedness patterns

Node-level Power Analysis

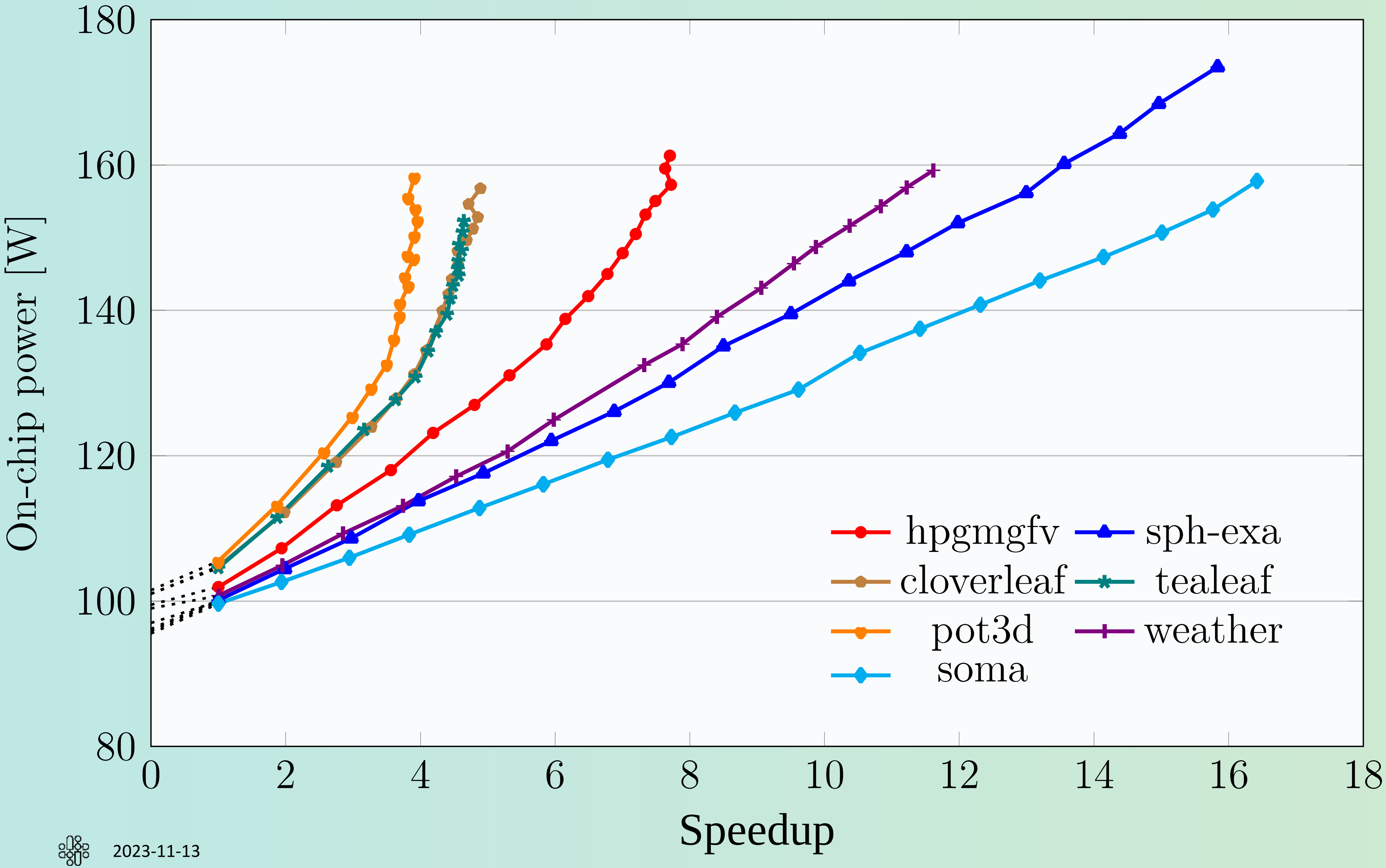
Node-level Power Analysis

- What impact does non-ideal scalability have on the power dissipated by “hot” and “cold” codes?

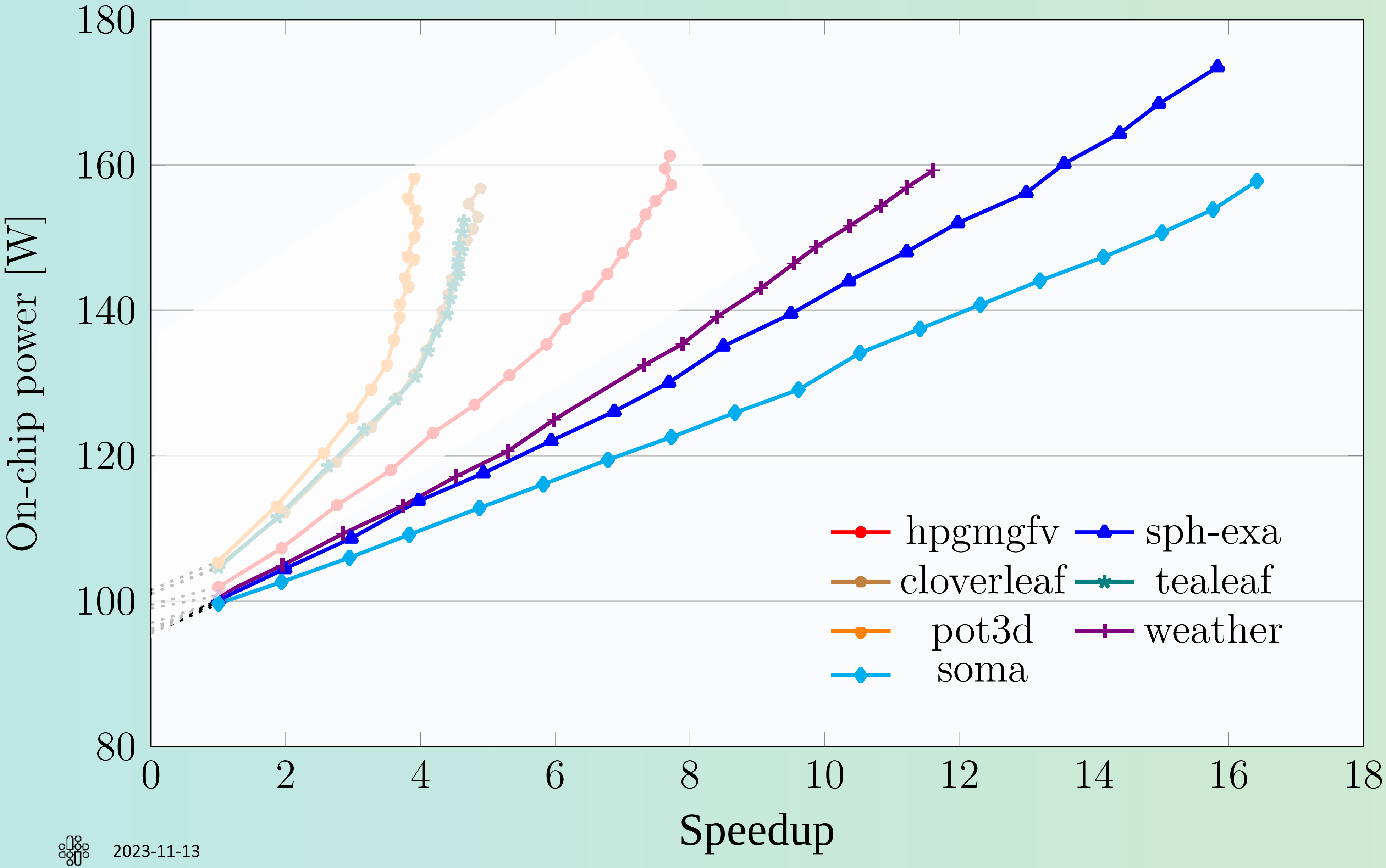
Node-level Power Analysis

- What impact does non-ideal scalability have on the power dissipated by “hot” and “cold” codes?
- How significant is the zero-core baseline power?

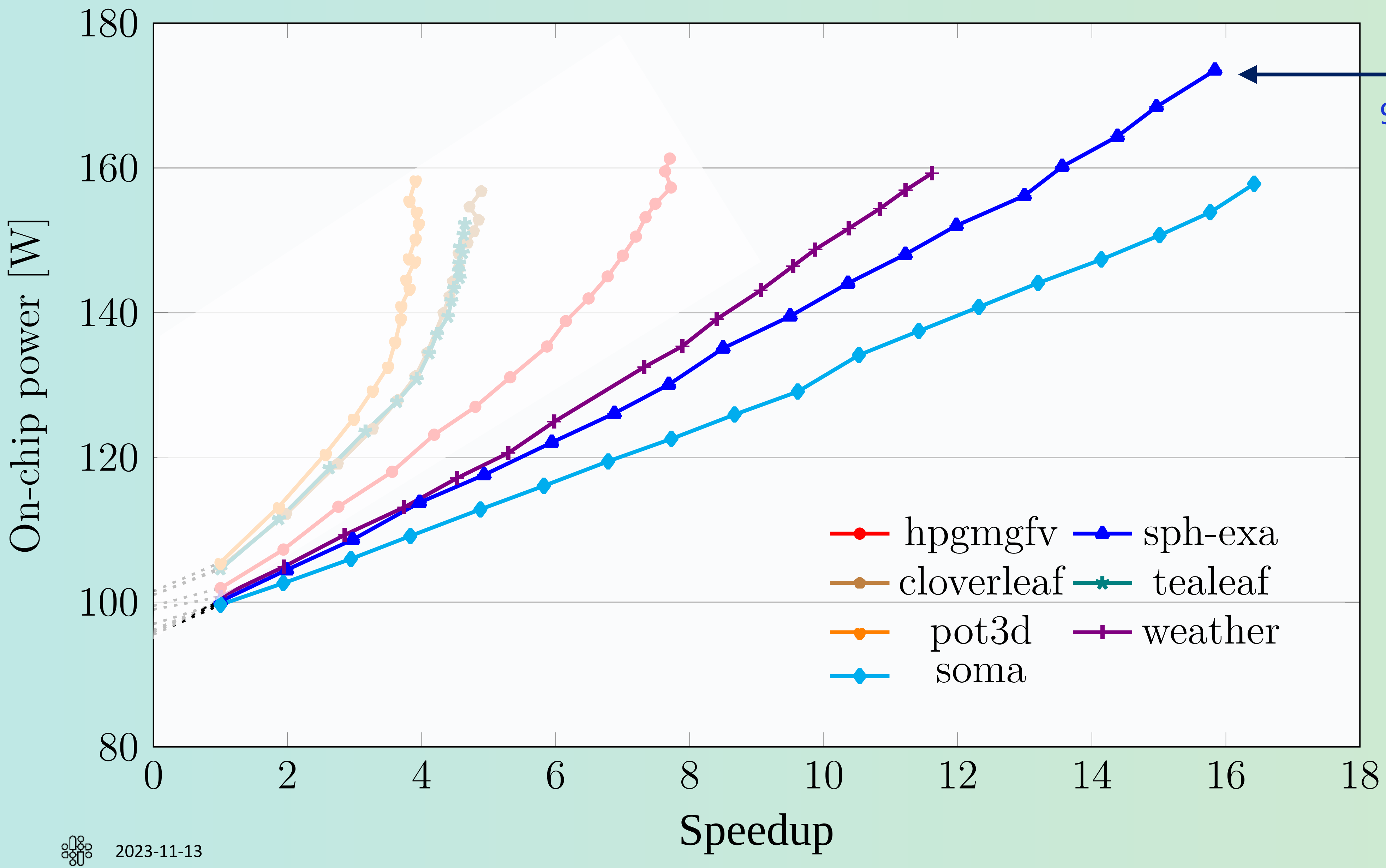
Node Power: Ice Lake (1 ccNUMA domain)



Node Power: Ice Lake (1 ccNUMA domain)

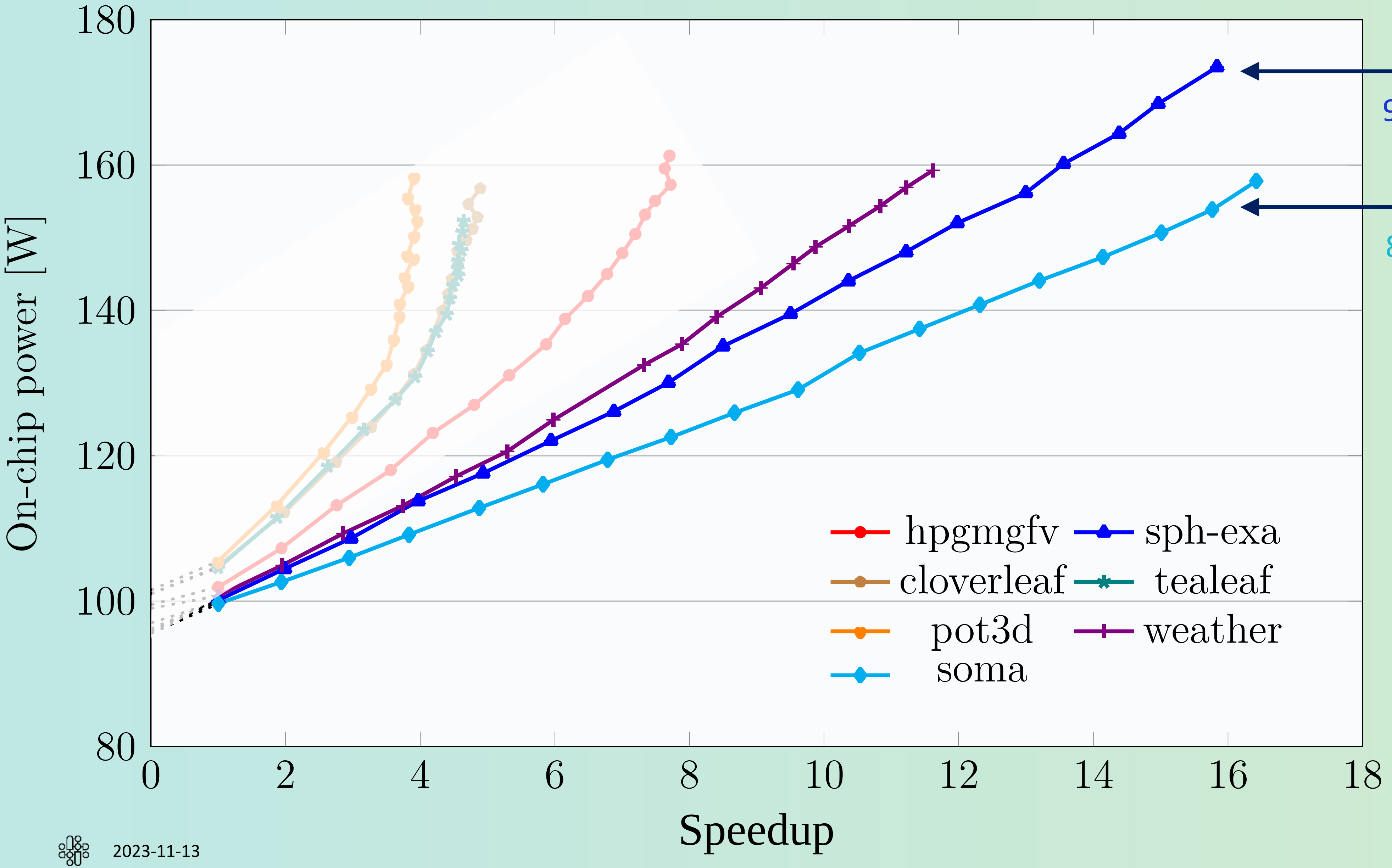


Node Power: Ice Lake (1 ccNUMA domain)



hottest benchmark
98% of socket TDP (244 W)

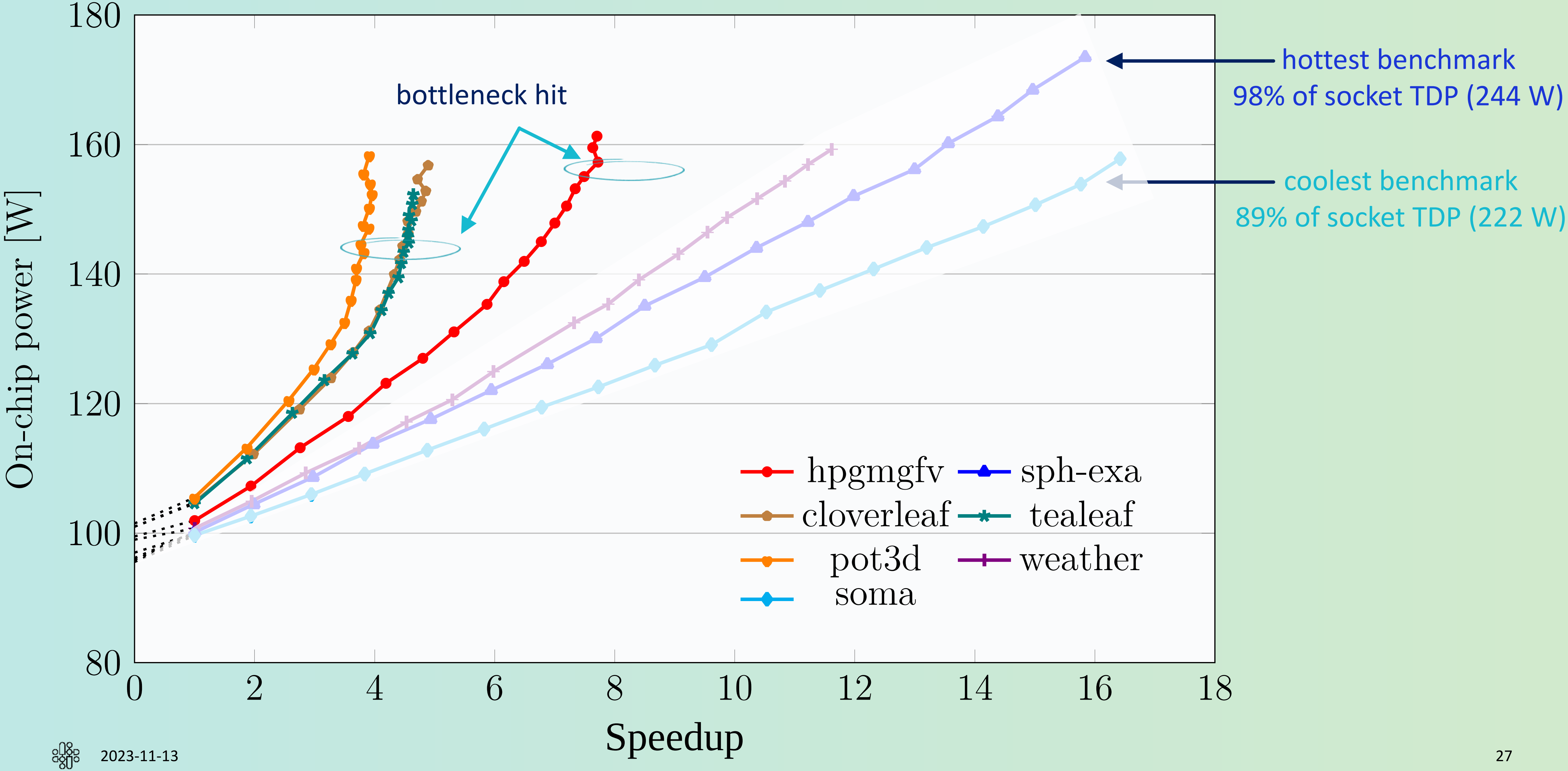
Node Power: Ice Lake (1 ccNUMA domain)



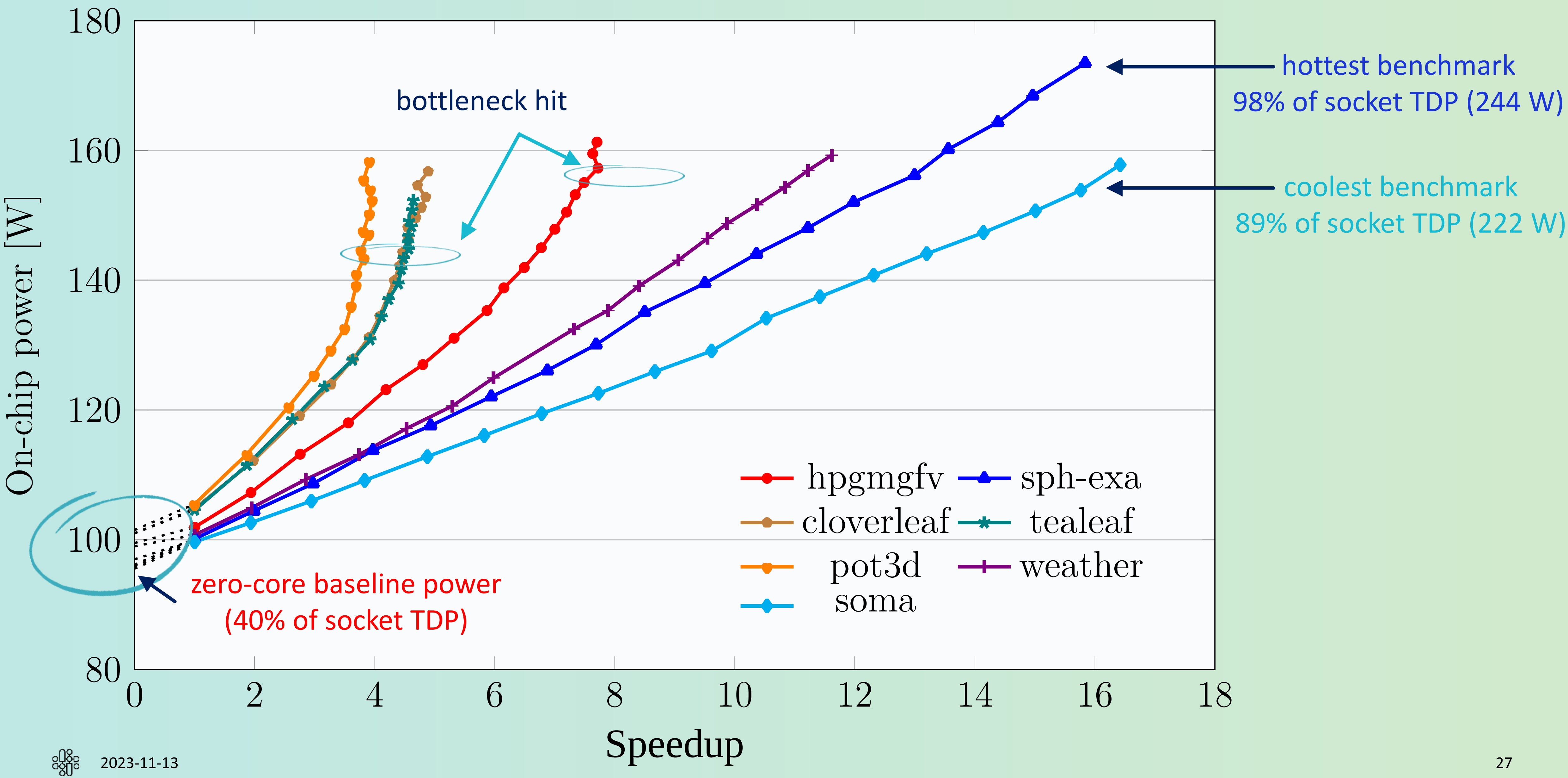
hottest benchmark
98% of socket TDP (244 W)

coolest benchmark
89% of socket TDP (222 W)

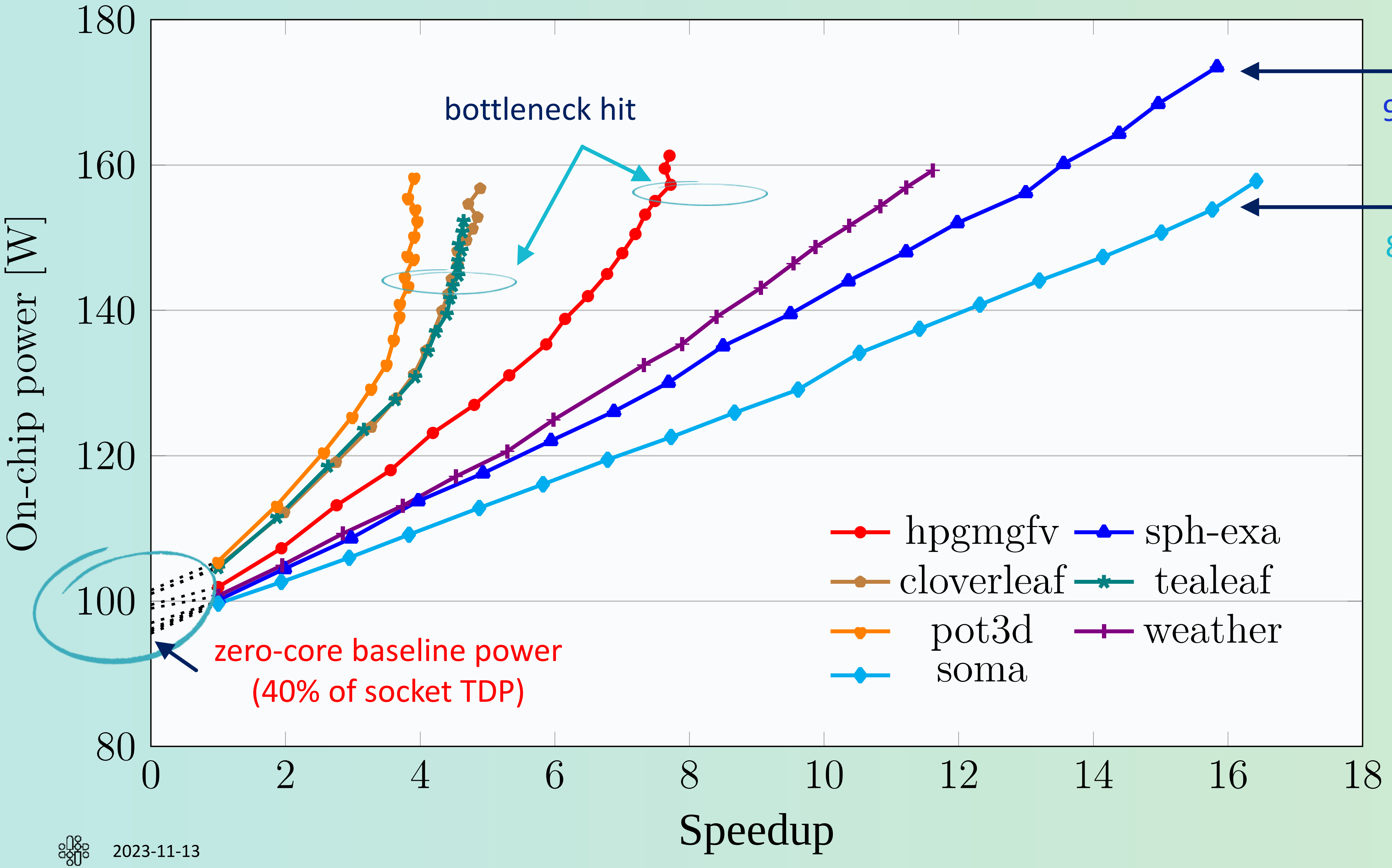
Node Power: Ice Lake (1 ccNUMA domain)



Node Power: Ice Lake (1 ccNUMA domain)



Node Power: Ice Lake (1 ccNUMA domain)

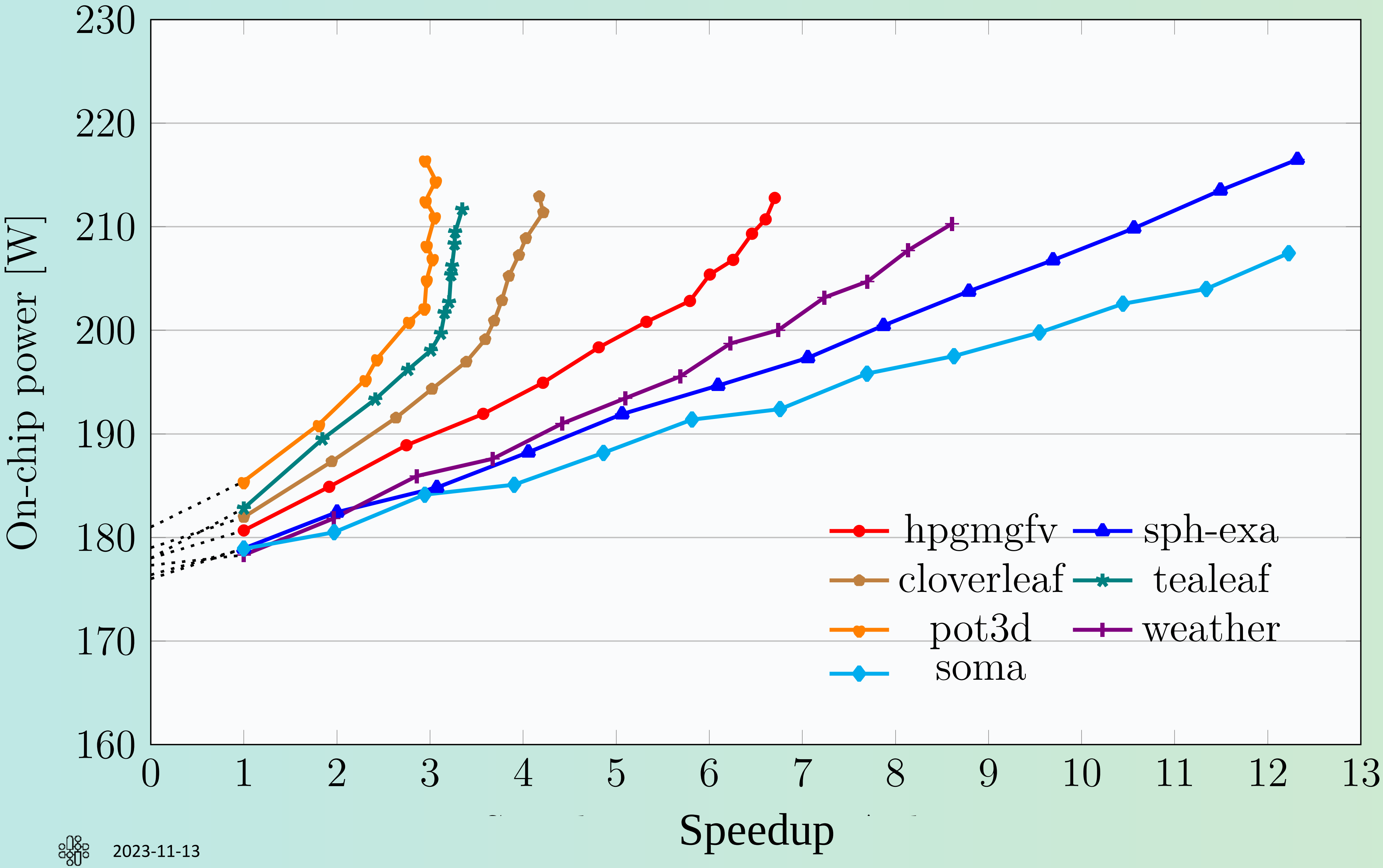


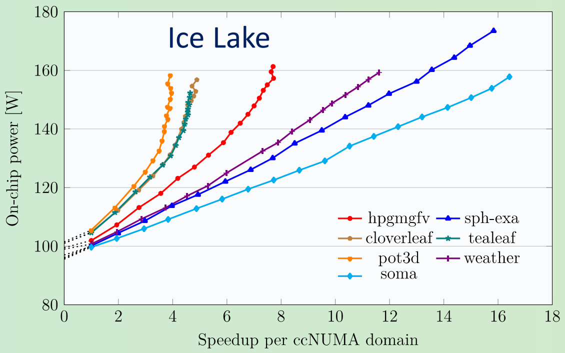
hottest benchmark
98% of socket TDP (244 W)

coolest benchmark
89% of socket TDP (222 W)

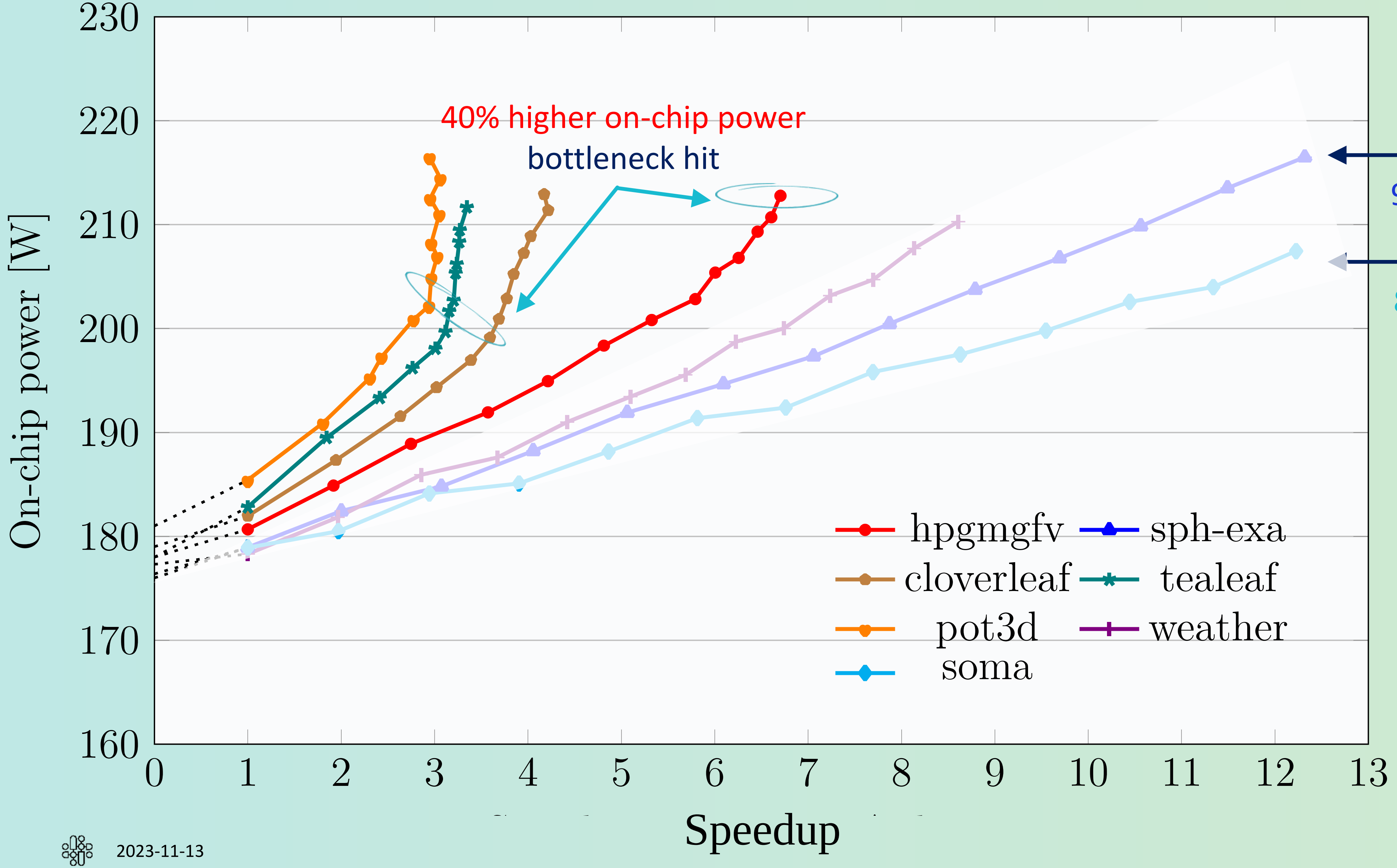
Intel Sandy Bridge (2012)
zero-core baseline power
<20% of 120 W socket TDP

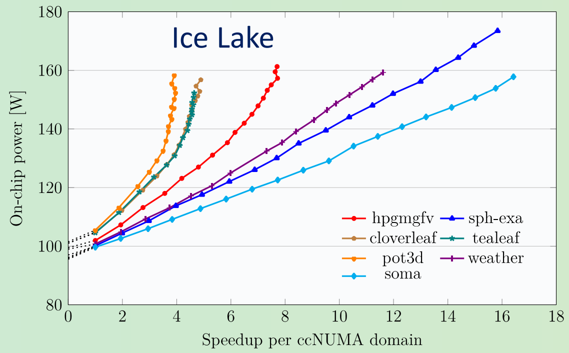
Node Power: Sapphire Rapids vs. Ice Lake



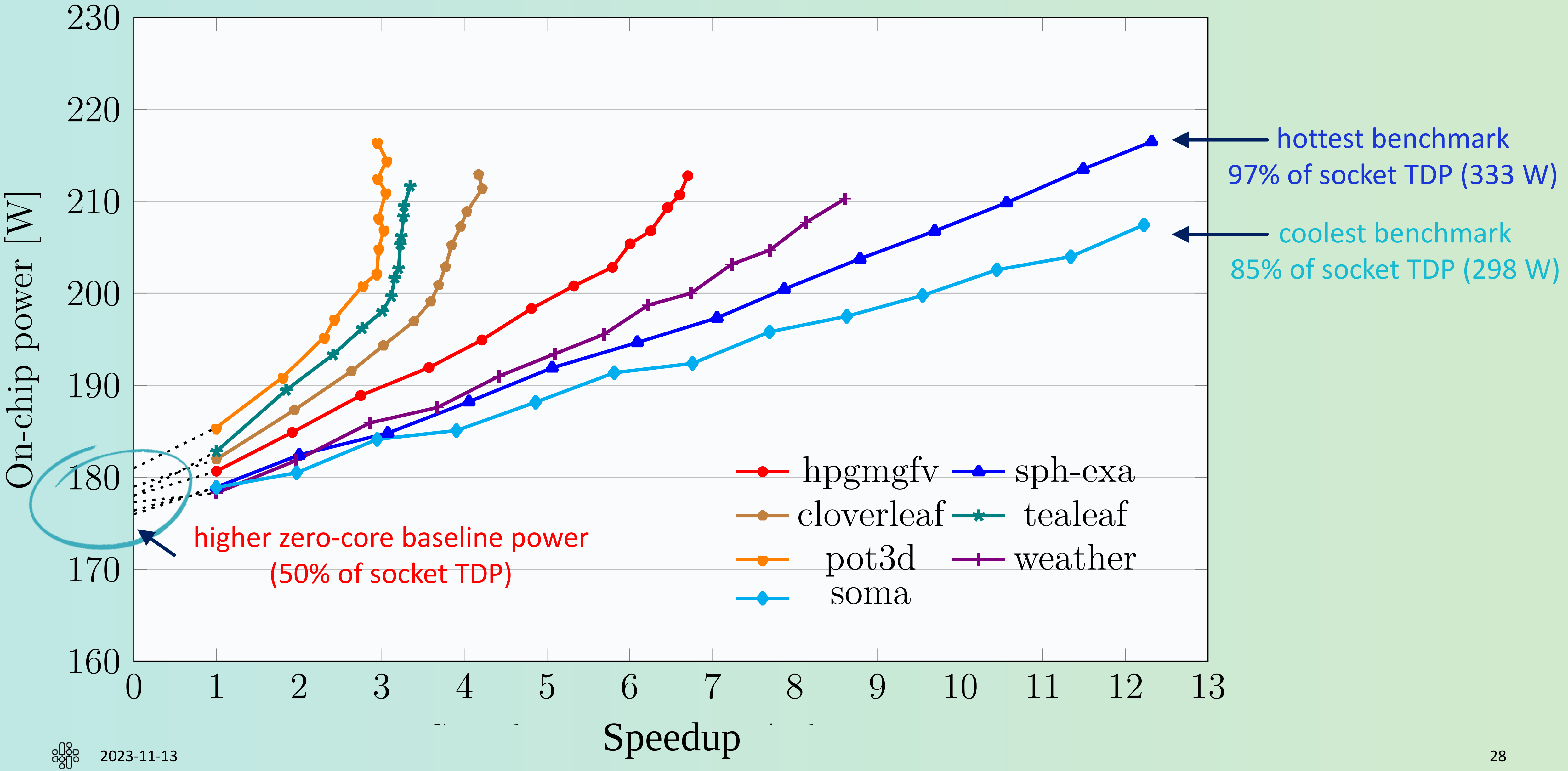


Node Power: Sapphire Rapids vs. Ice Lake





Node Power: Sapphire Rapids vs. Ice Lake



Node-level Energy Analysis

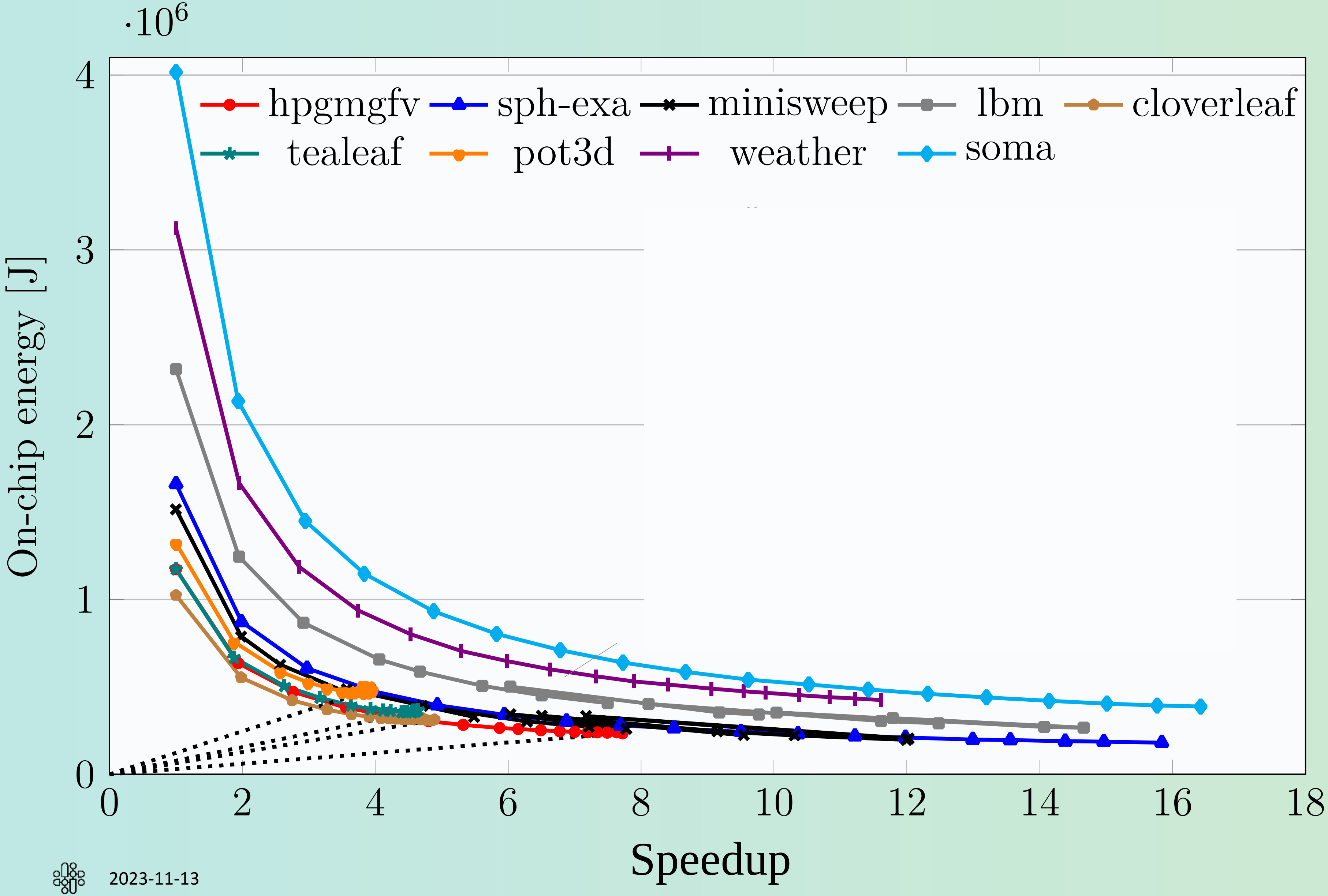
Node-level Energy Analysis

- What impact does non-ideal scalability have on the energy consumed by “hot” and “cold” codes?

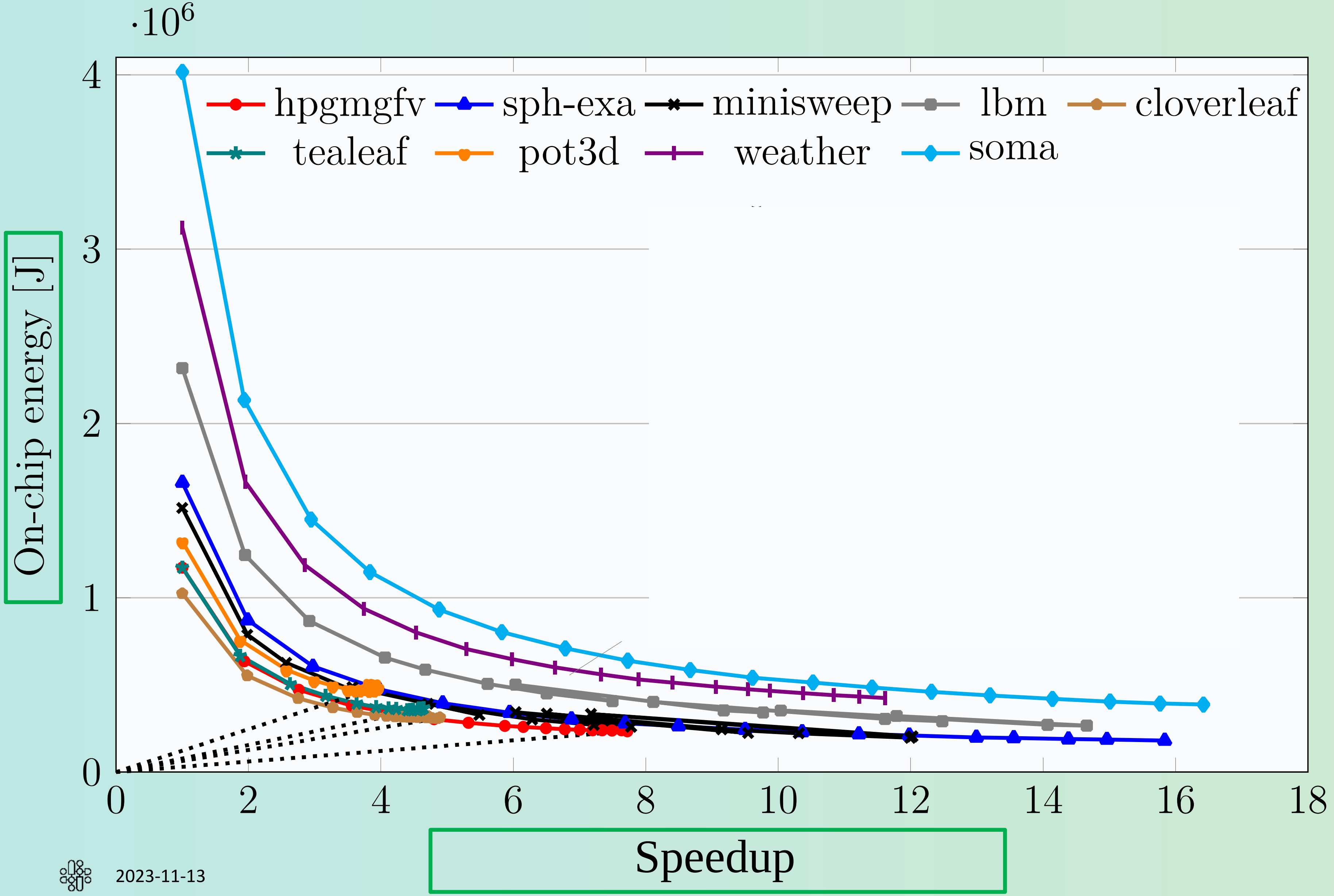
Node-level Energy Analysis

- What impact does non-ideal scalability have on the energy consumed by “hot” and “cold” codes?
- How significant is the energy-delay product?

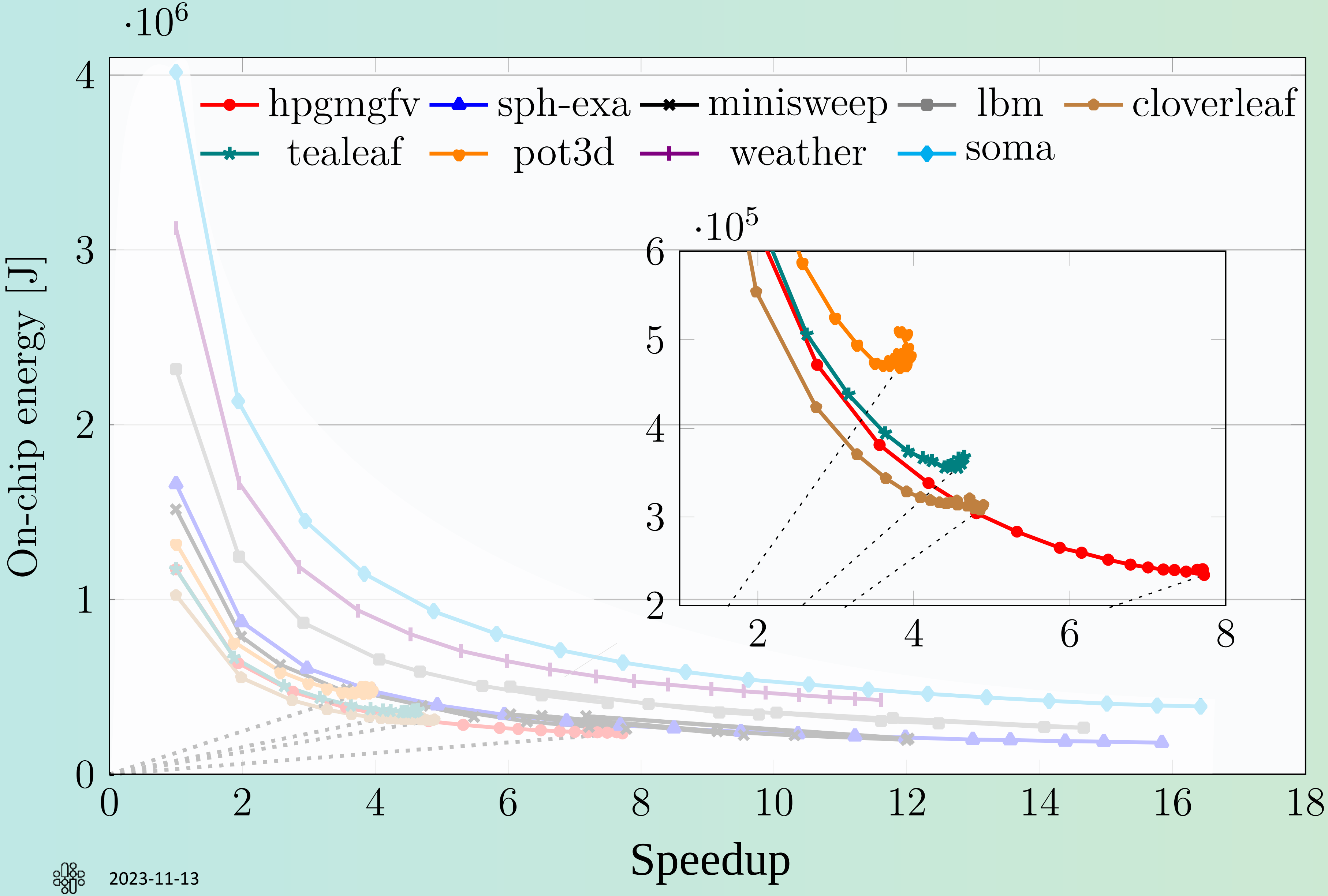
Energy to Solution: Ice Lake (1 NUMA domain)



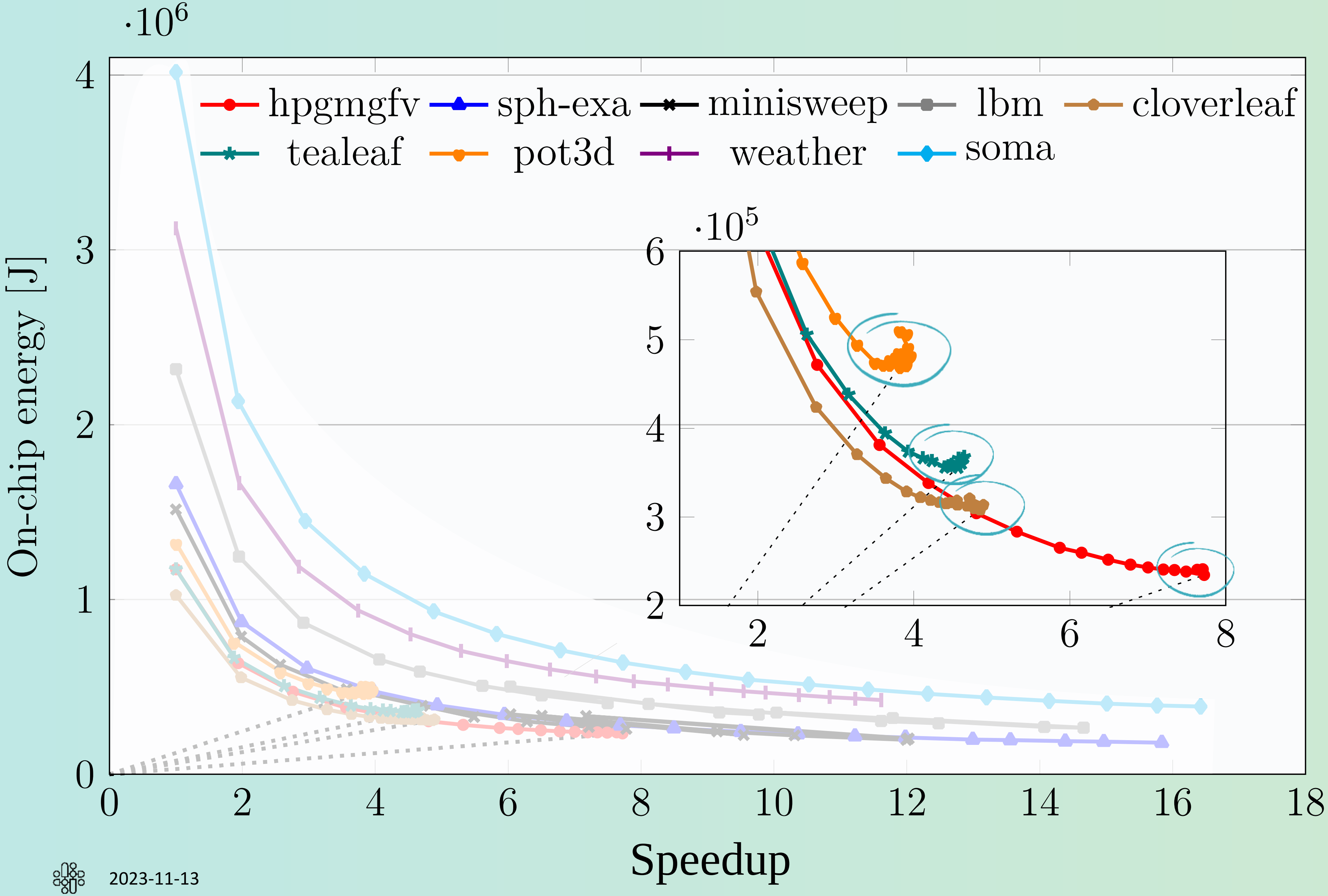
Energy to Solution: Ice Lake (1 NUMA domain)



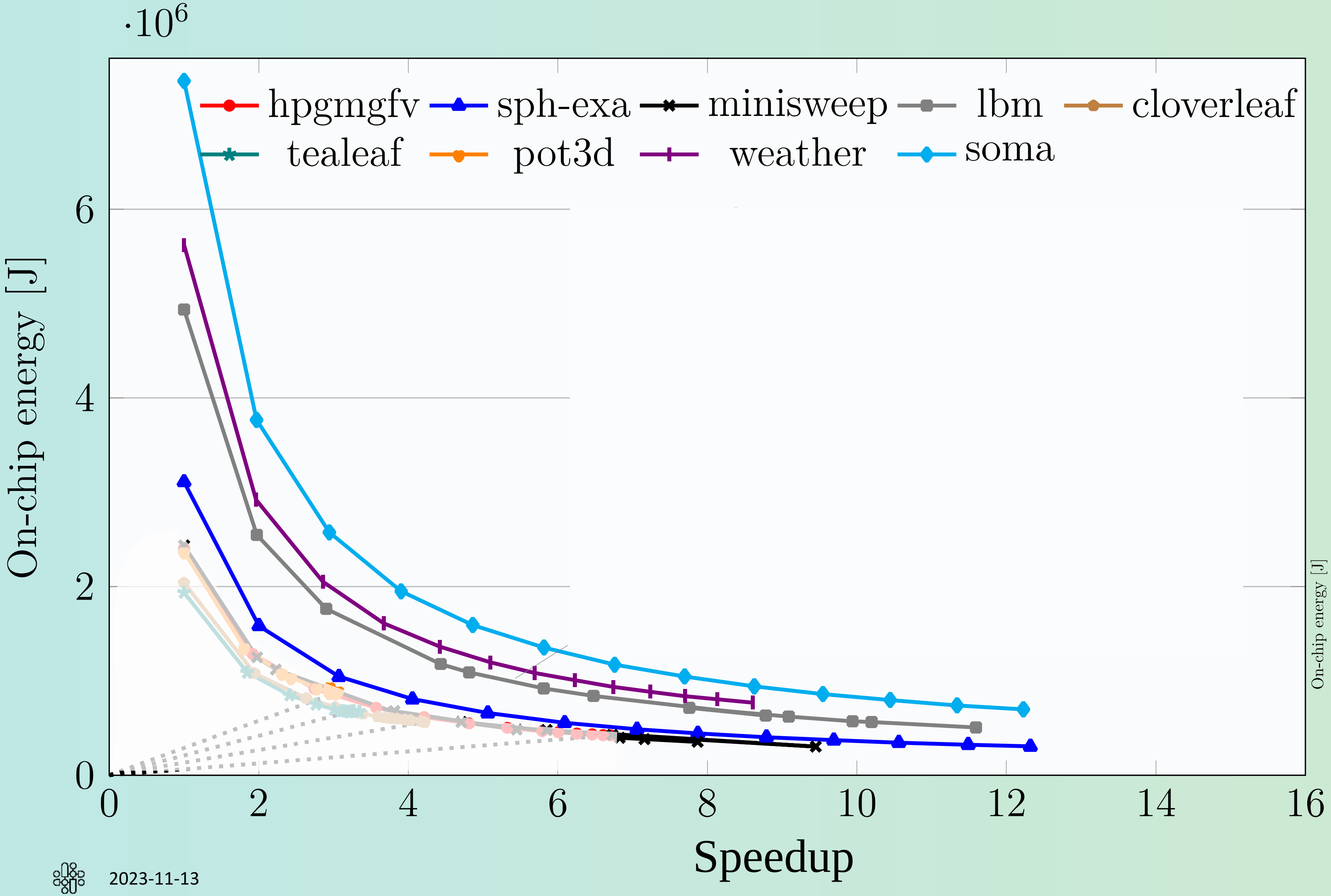
Energy to Solution: Ice Lake (1 NUMA domain)

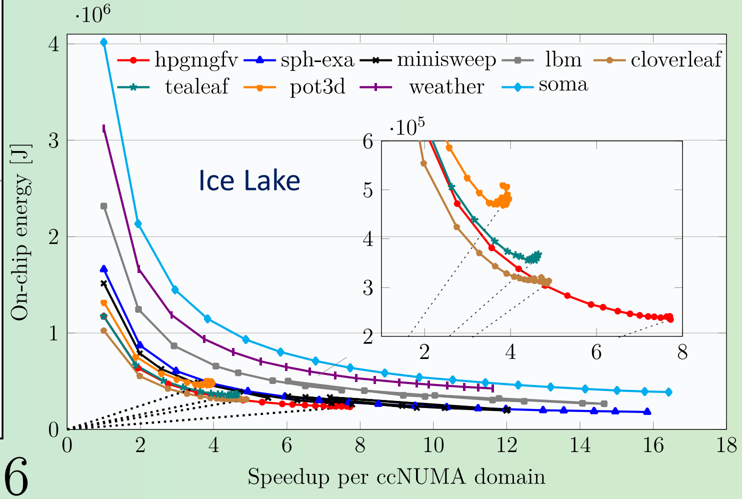


Energy to Solution: Ice Lake (1 NUMA domain)

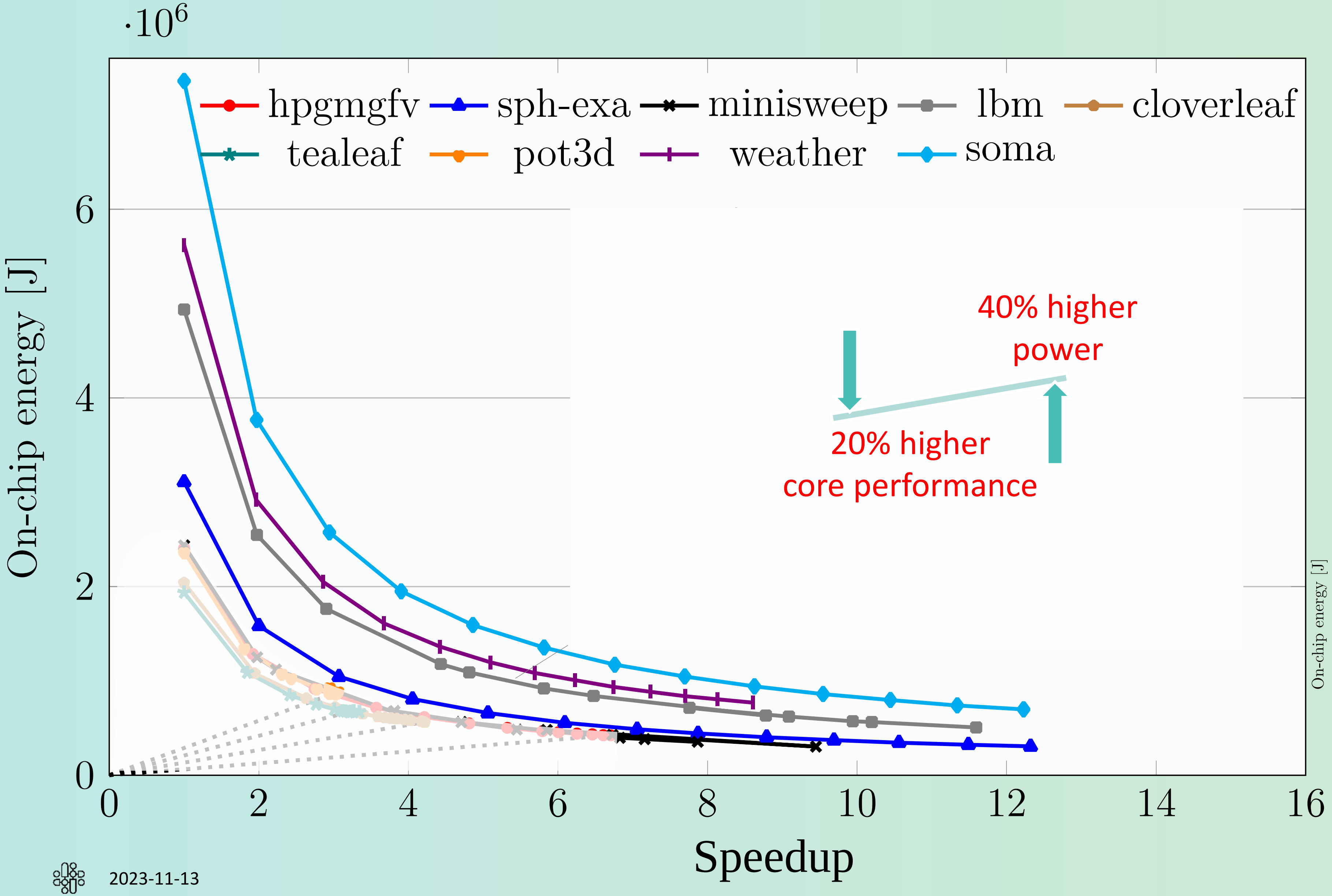


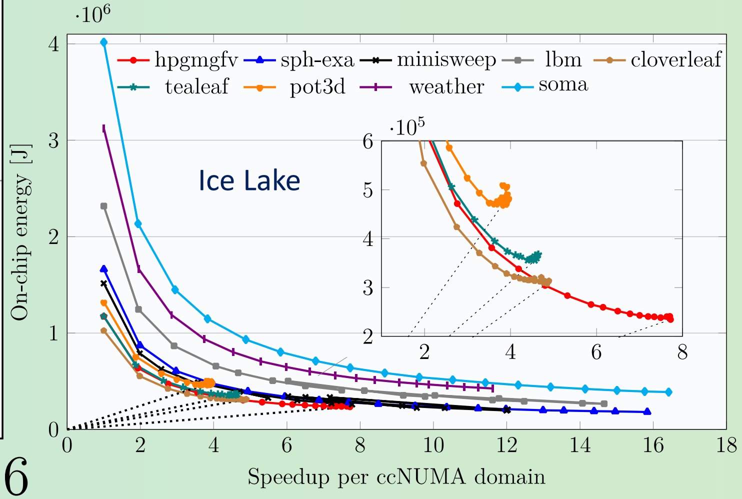
Energy: Sapphire Rapids vs. Ice Lake (1 NUMA domain)



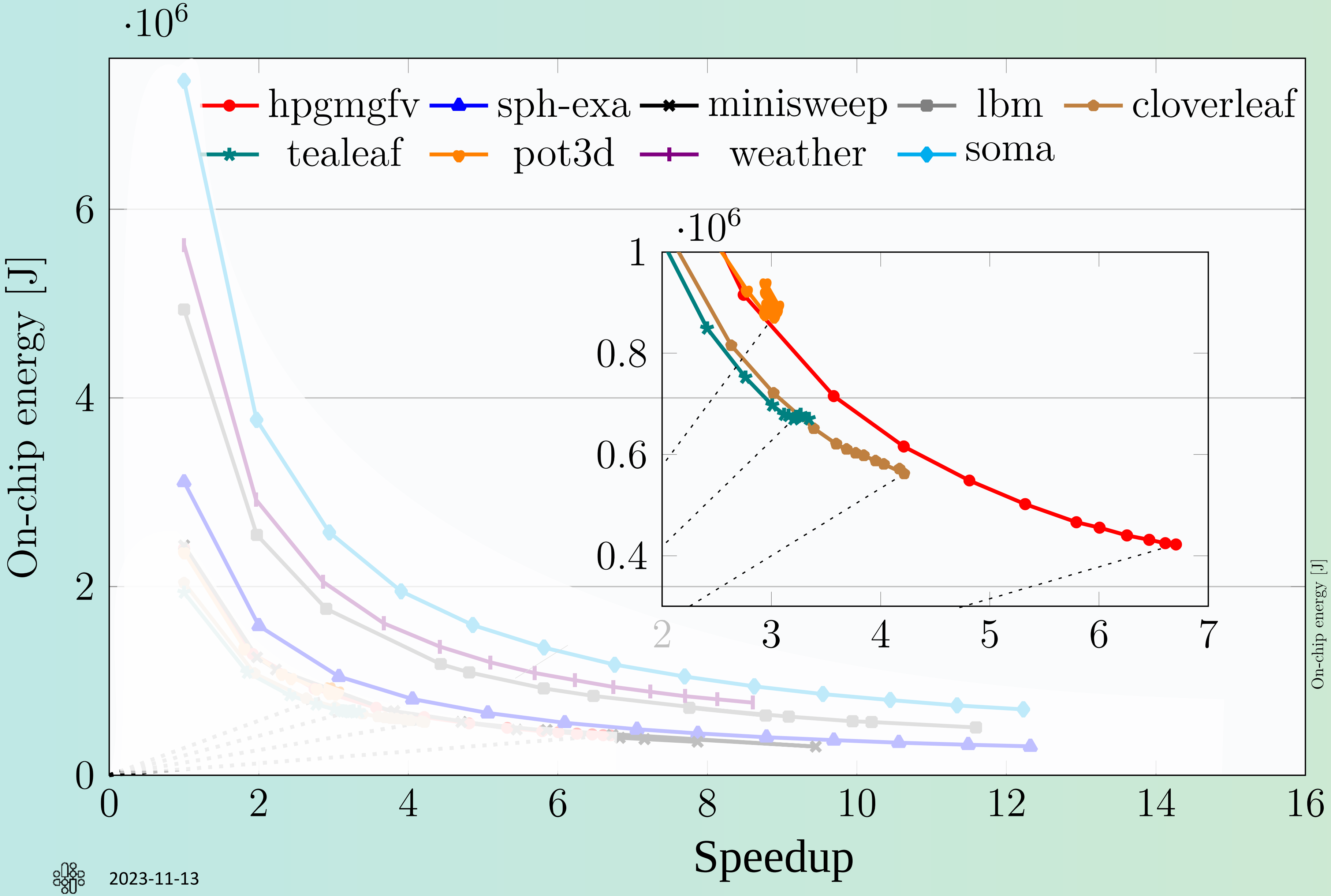


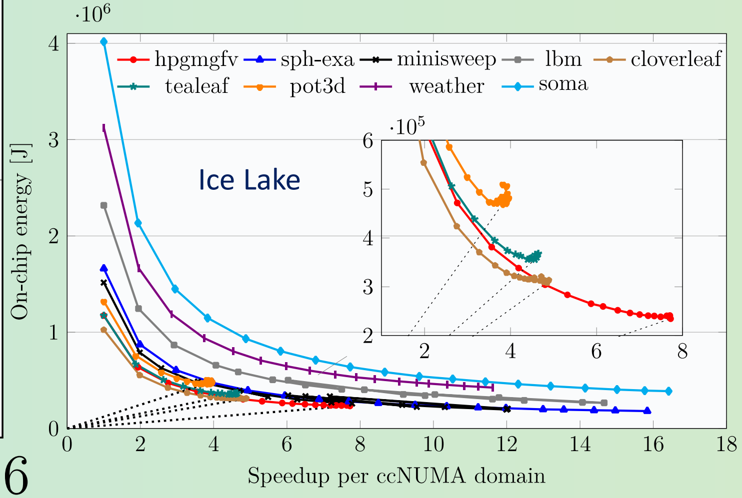
Energy: Sapphire Rapids vs. Ice Lake (1 NUMA domain)



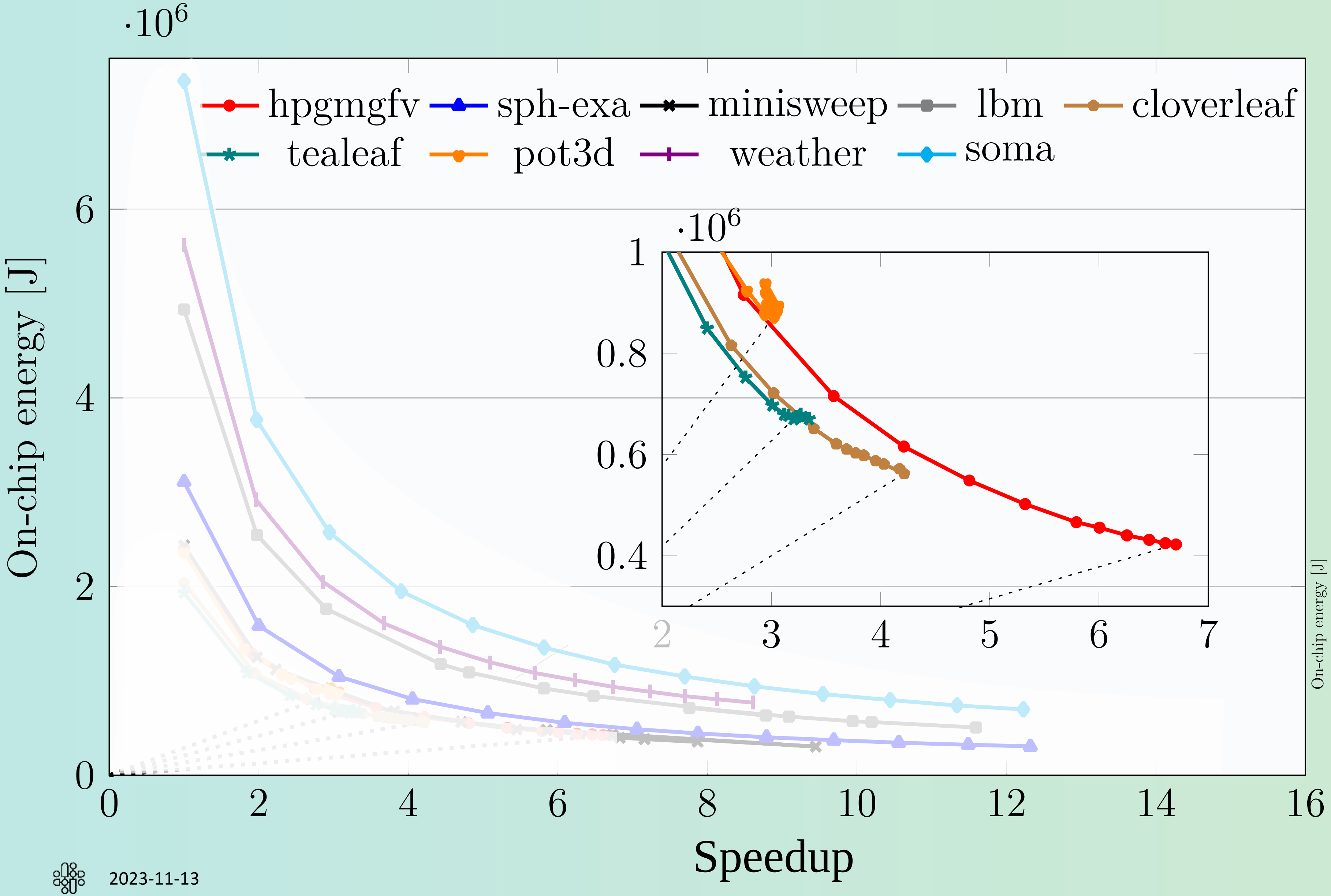


Energy: Sapphire Rapids vs. Ice Lake (1 NUMA domain)



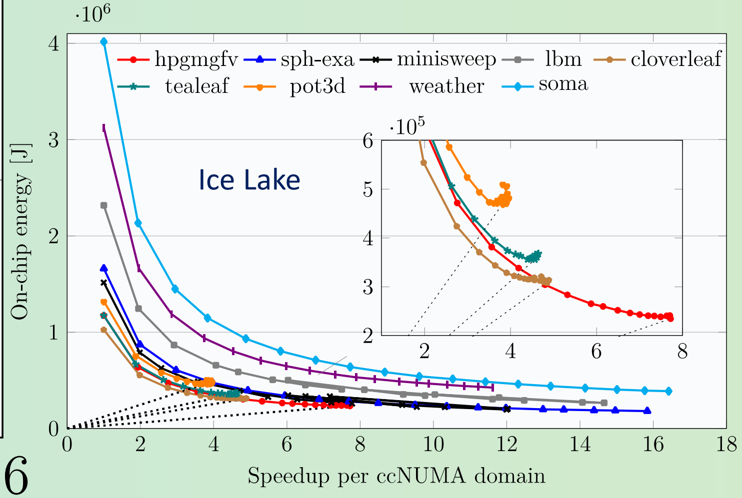


Energy: Sapphire Rapids vs. Ice Lake (1 NUMA domain)

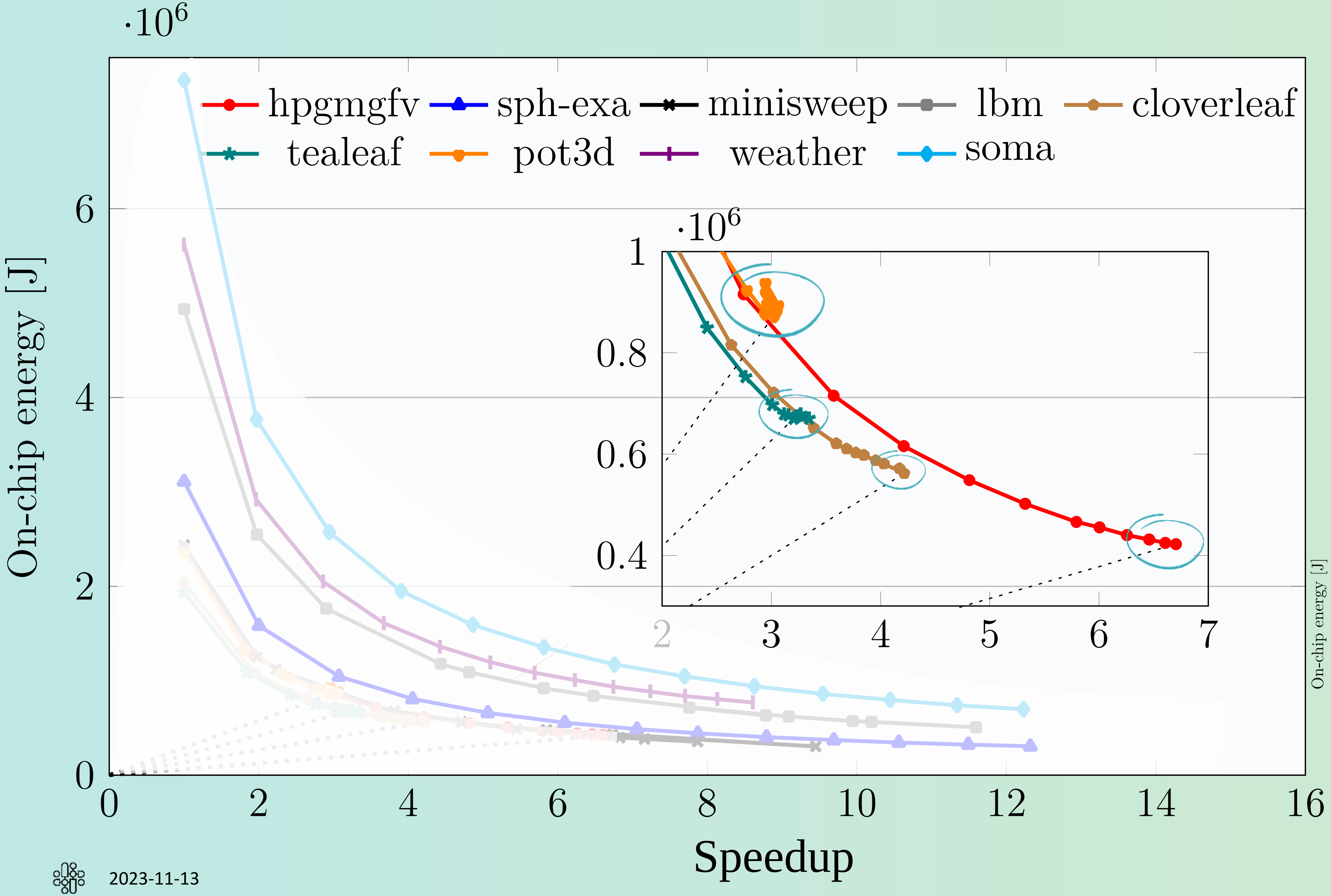


40% higher power

50% higher memory bandwidth

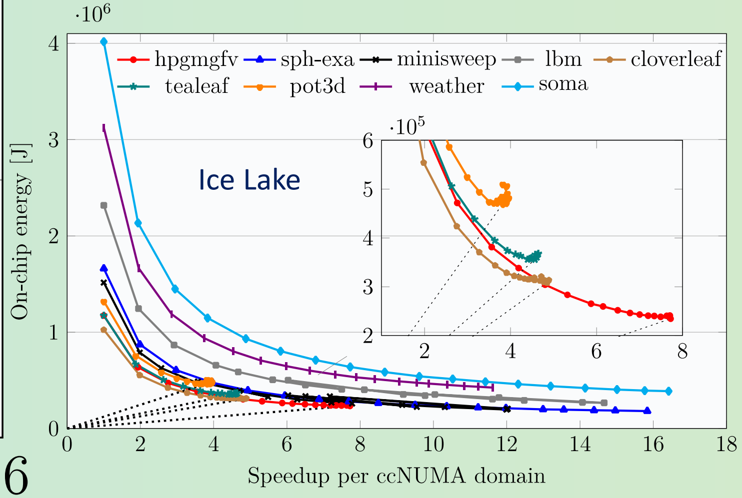


Energy: Sapphire Rapids vs. Ice Lake (1 NUMA domain)



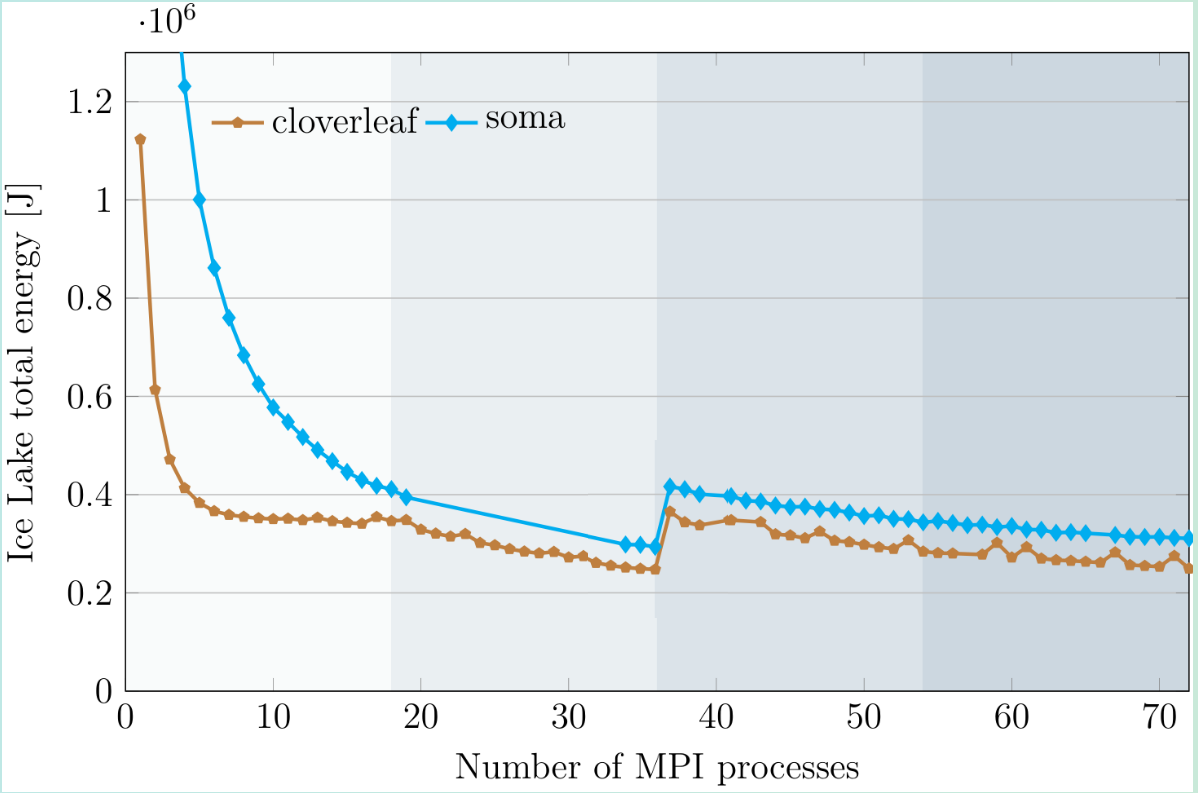
40% higher power

50% higher memory bandwidth

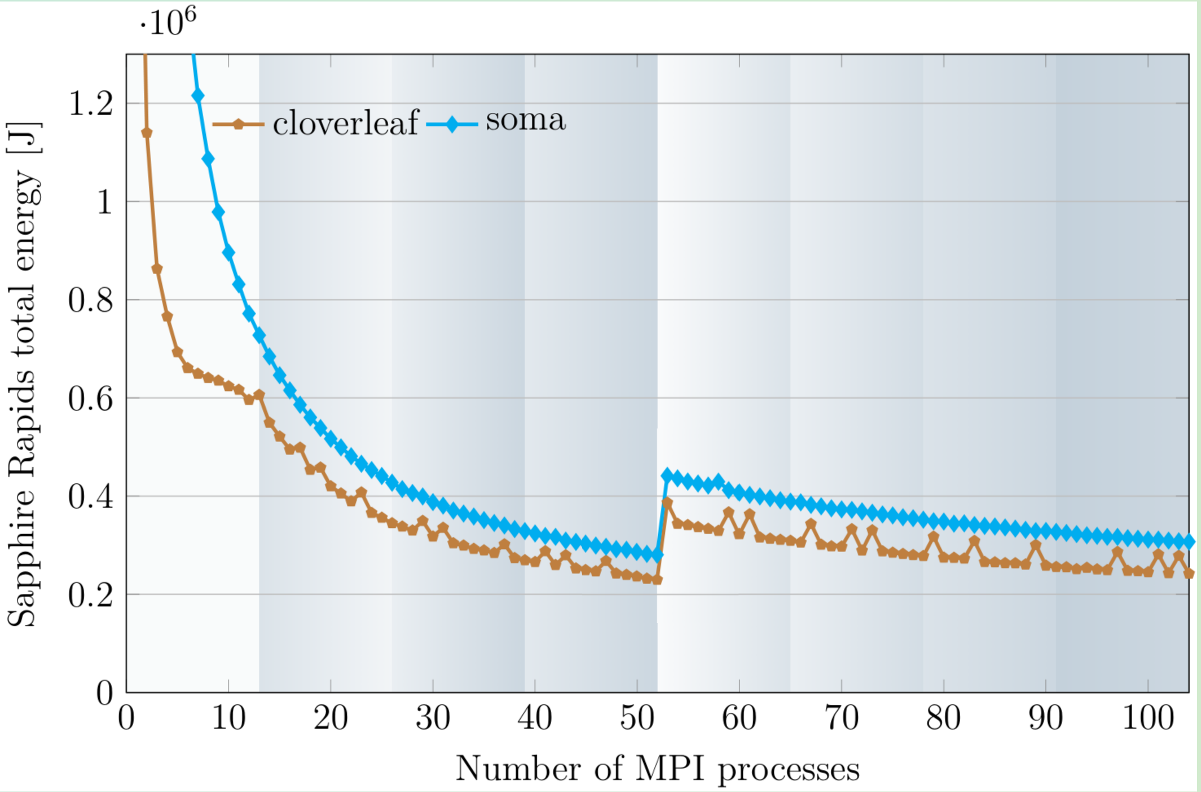


Full-node Energy to Solution ICX vs. SPR

Ice Lake

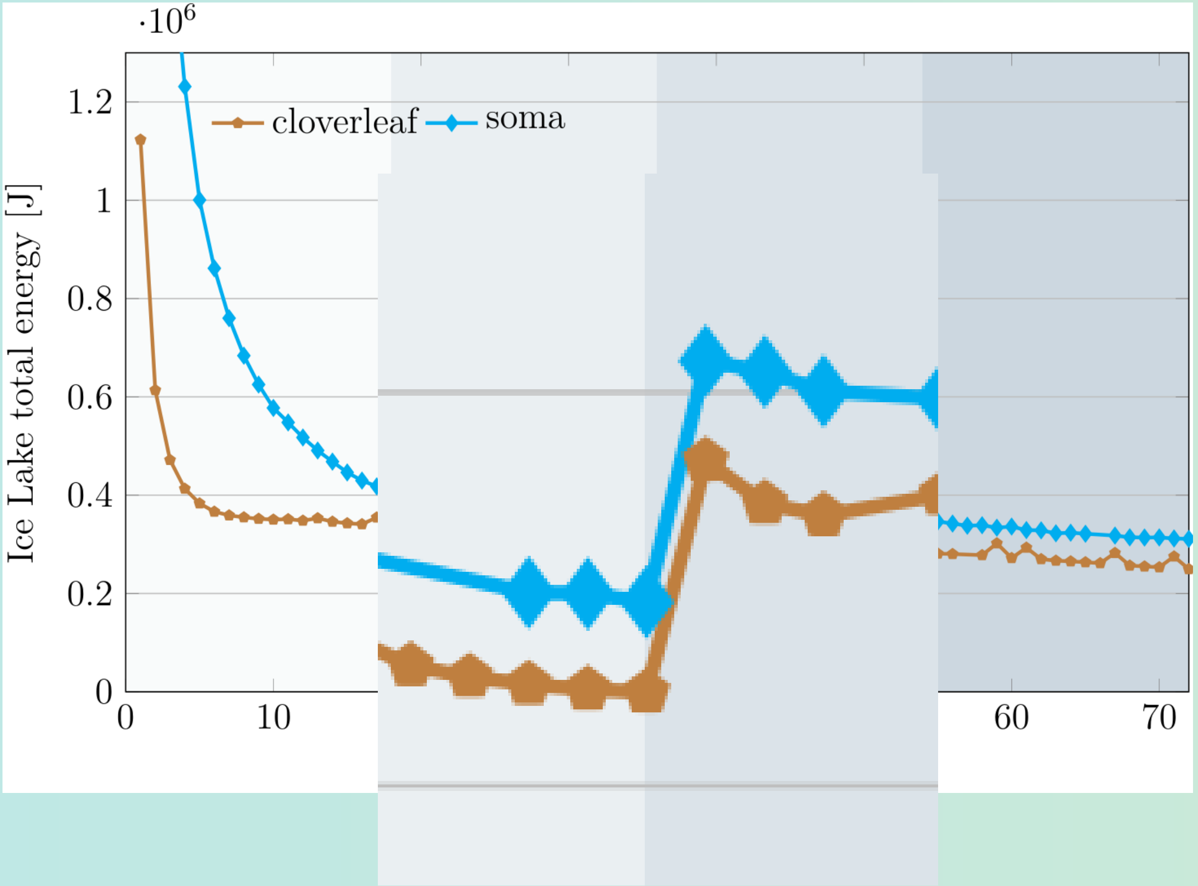


Sapphire Rapids

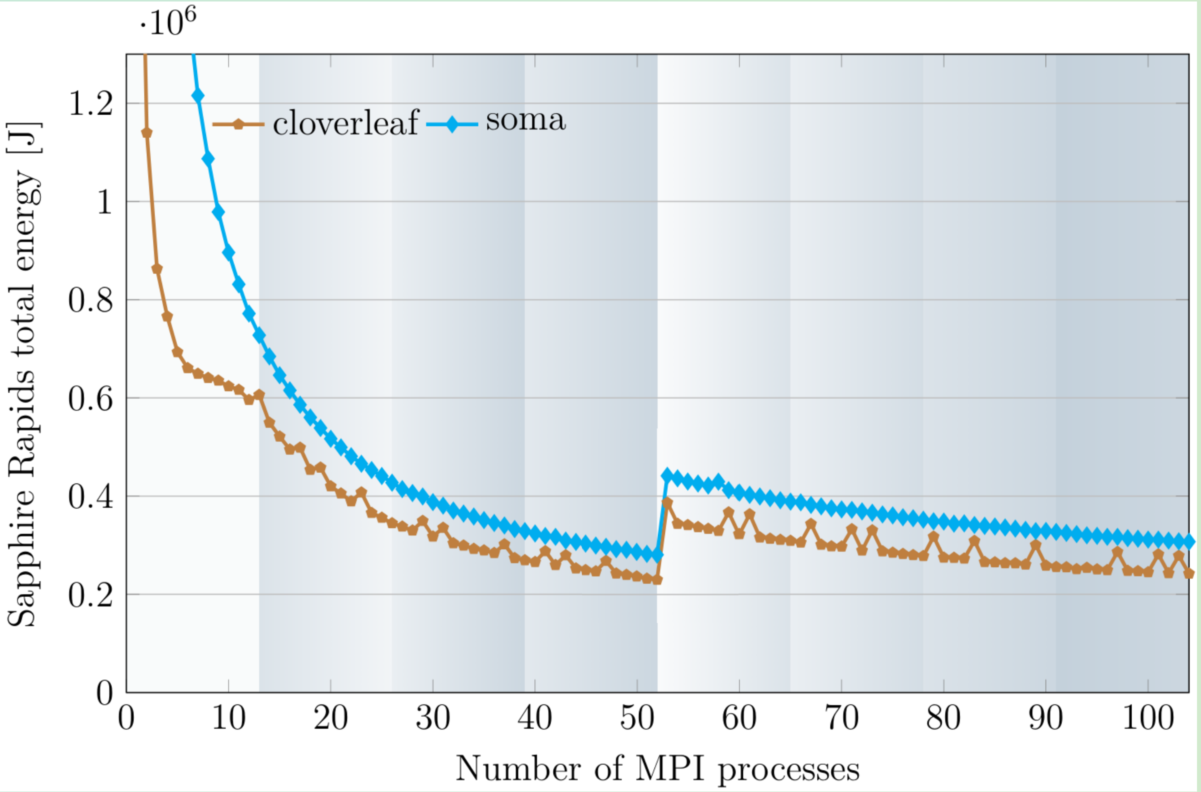


Full-node Energy to Solution ICX vs. SPR

Ice Lake

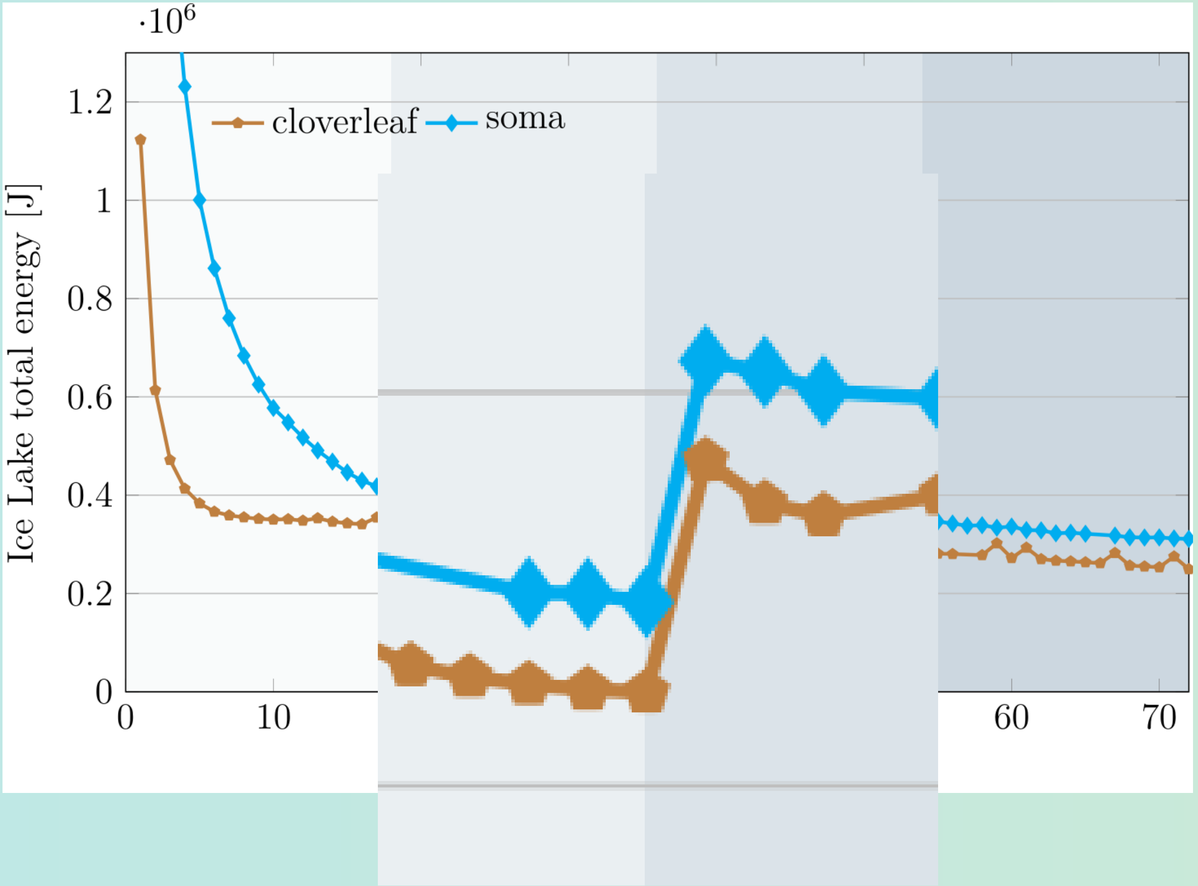


Sapphire Rapids

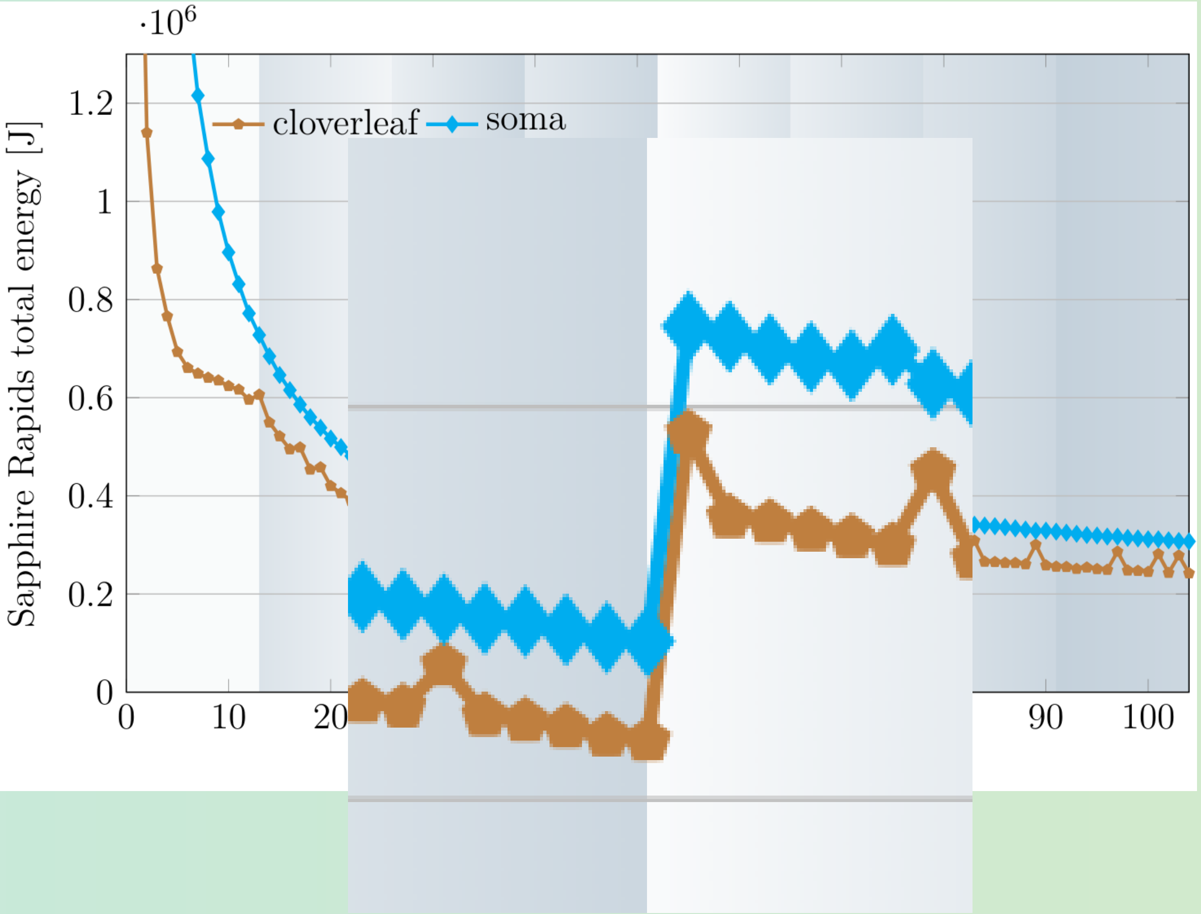


Full-node Energy to Solution ICX vs. SPR

Ice Lake

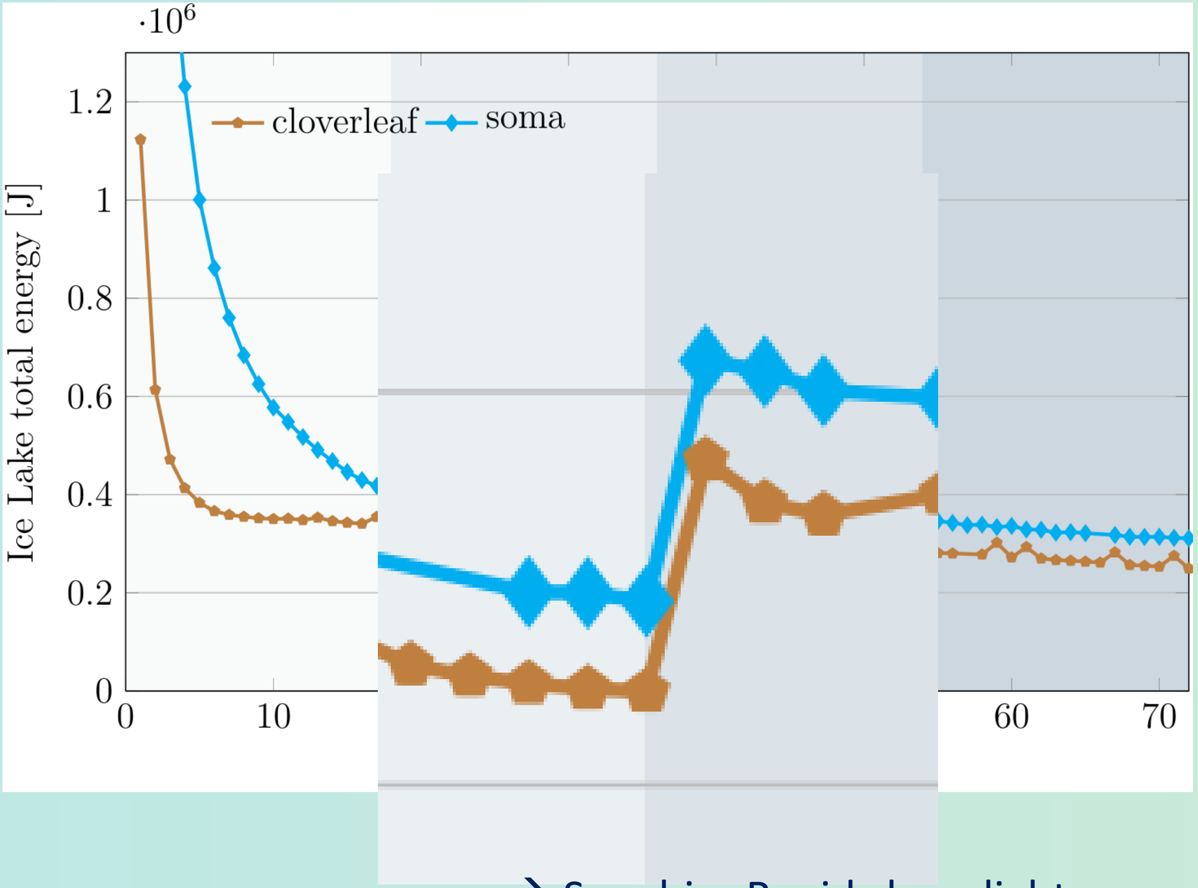


Sapphire Rapids

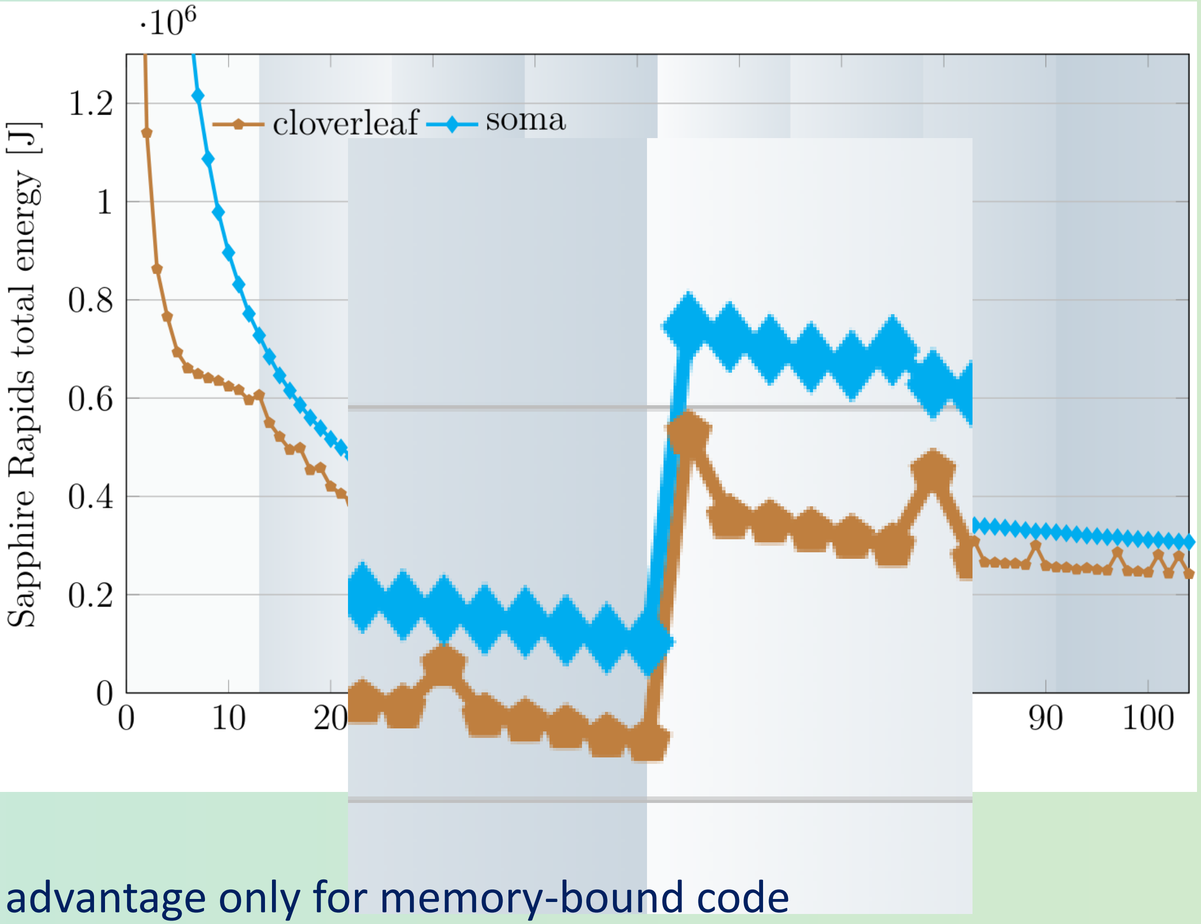


Full-node Energy to Solution ICX vs. SPR

Ice Lake



Sapphire Rapids



→ Sapphire Rapids has slight energy advantage only for memory-bound code

Multi-node Strong-scaling Performance Analysis

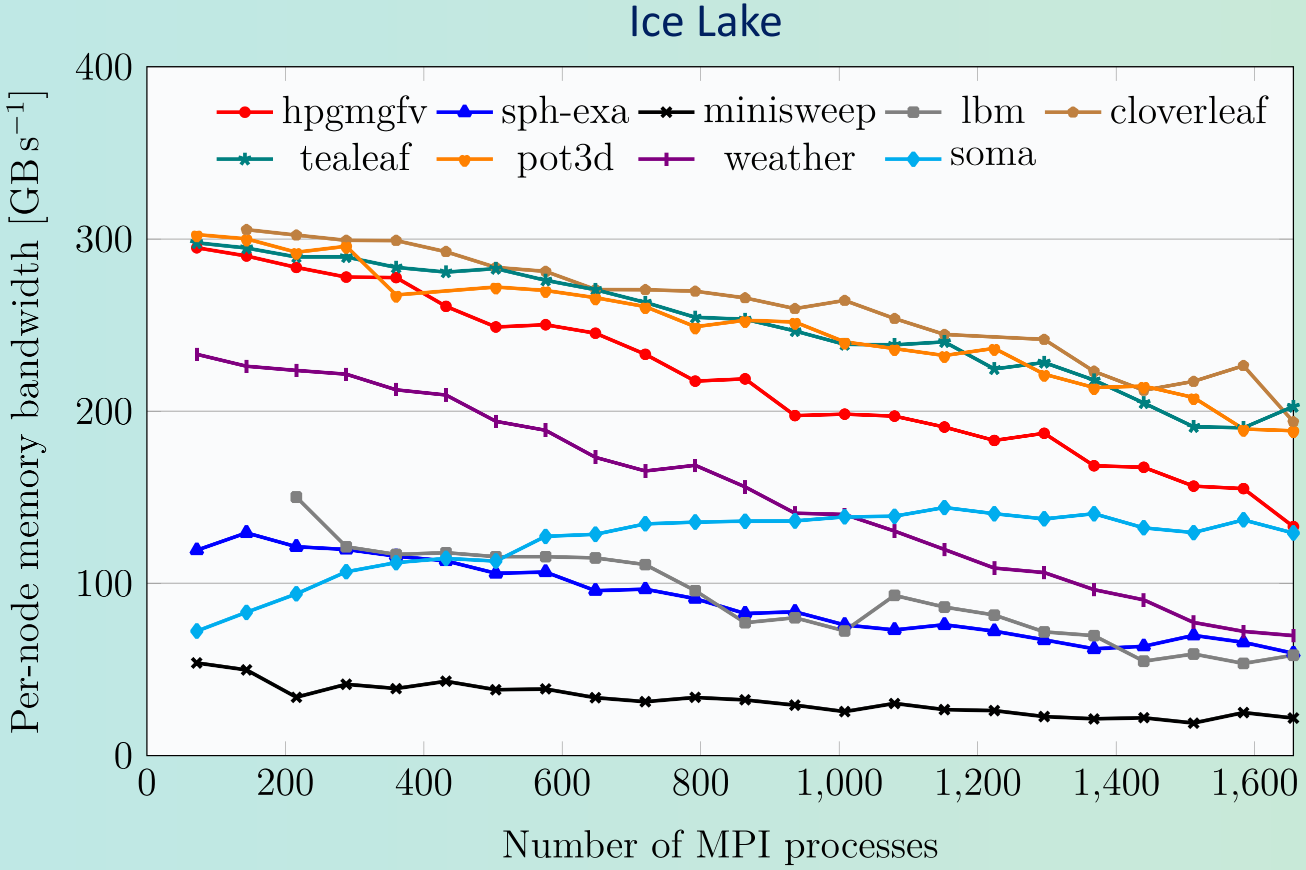
Multi-node Strong-scaling Performance Analysis

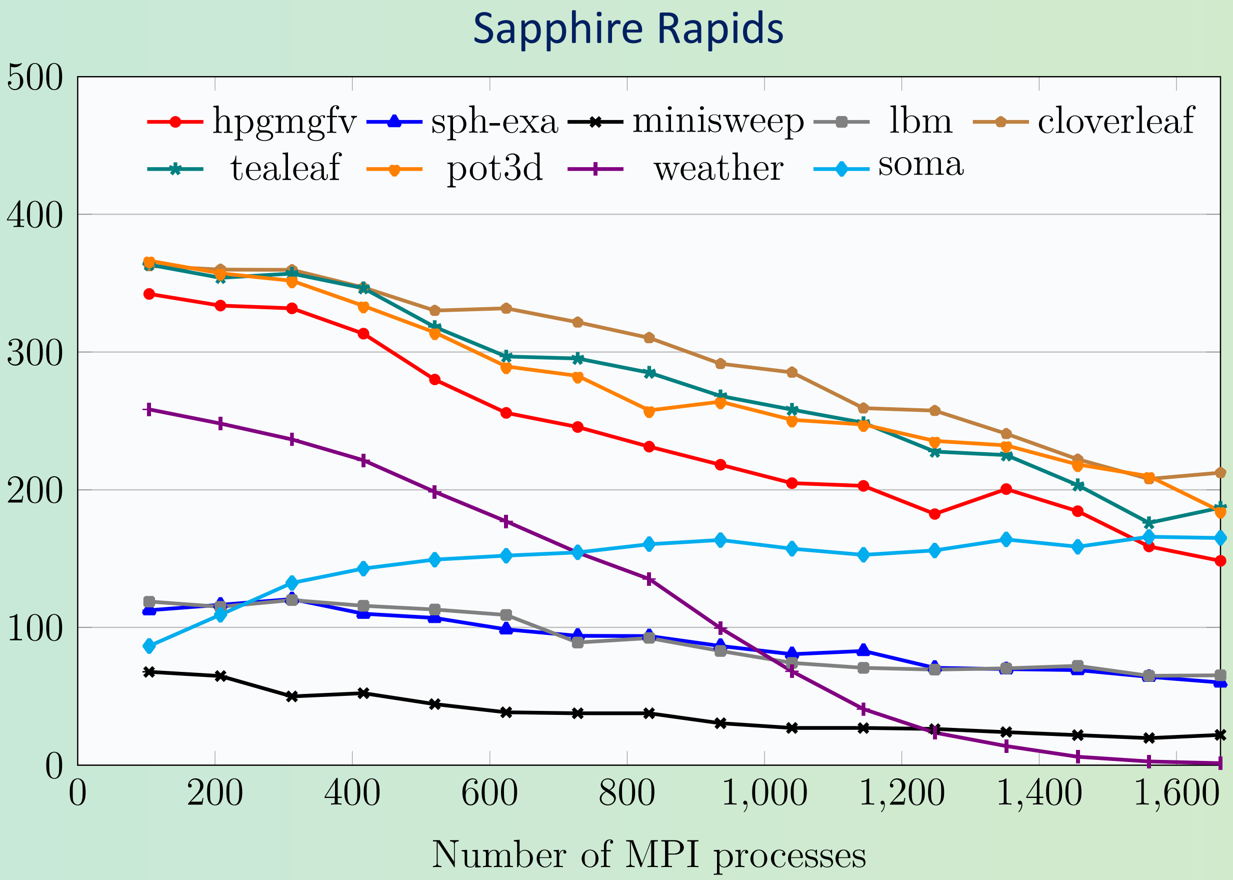
- How can scaling behavior be attributed in terms of the counteracting cause(s) of cache effects and communication overhead effects for “poor” and “super-linear” codes?

Multi-node Strong-scaling Performance Analysis

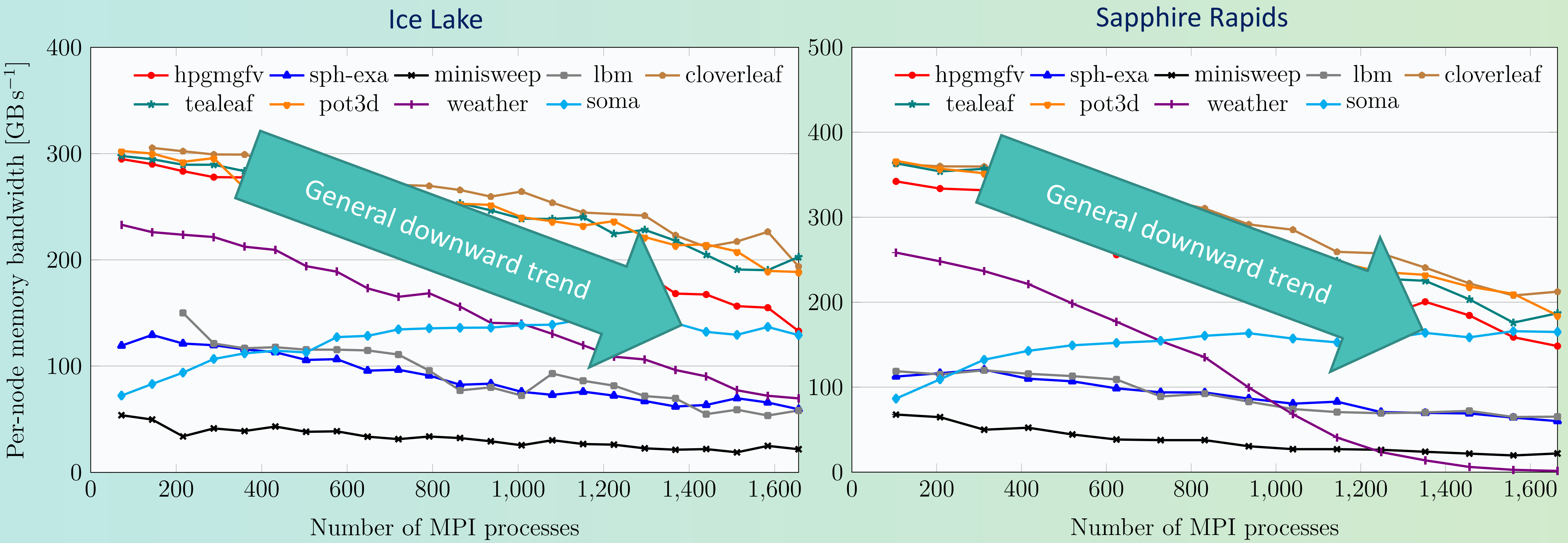
- How can scaling behavior be attributed in terms of the counteracting cause(s) of cache effects and communication overhead effects for “poor” and “super-linear” codes?
- How can fundamental resource metrics be used to uncover the roots of “unusual” scaling patterns?

Memory bandwidth: Ice Lake and Sapphire Rapids



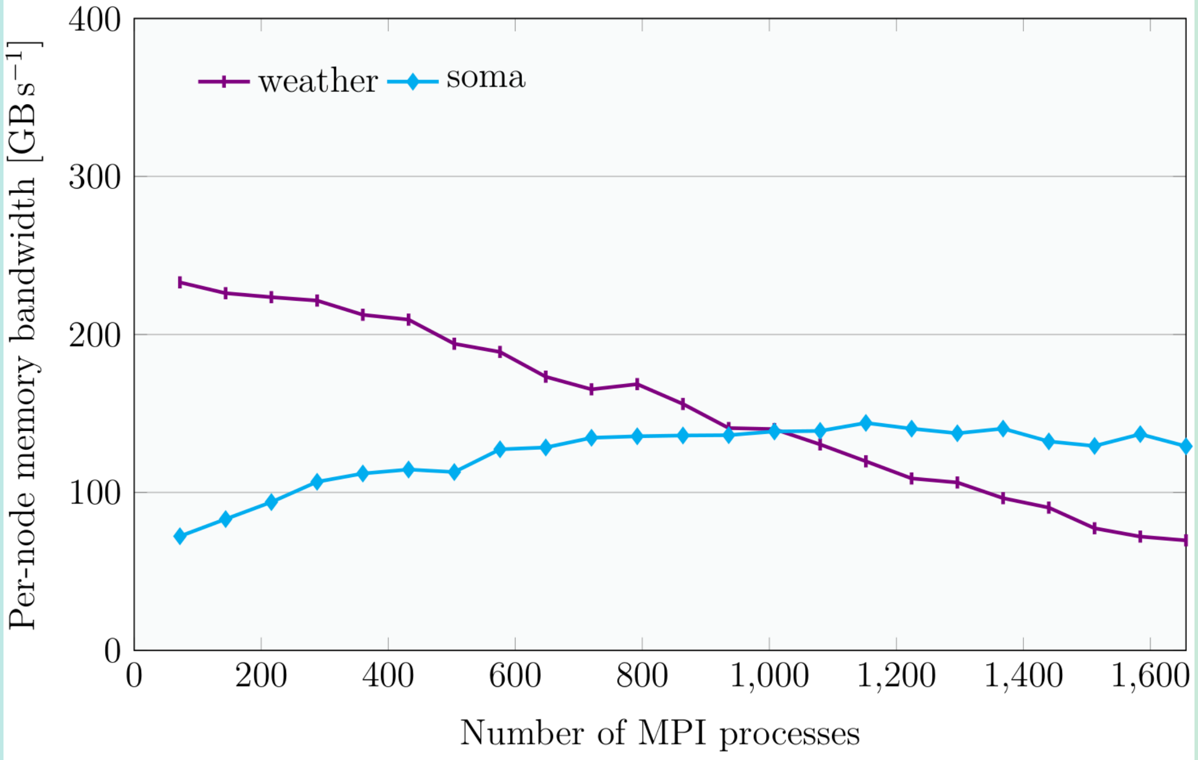


Memory bandwidth: Ice Lake and Sapphire Rapids

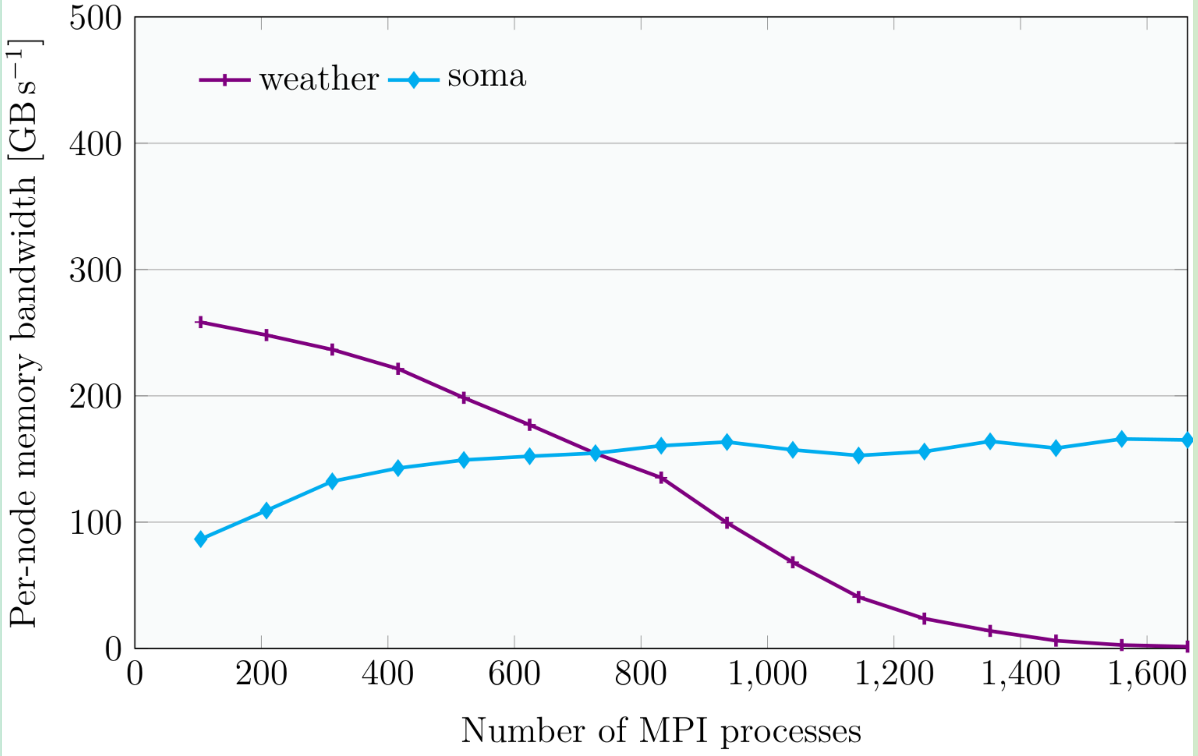


Memory bandwidth: weather and soma

Ice Lake

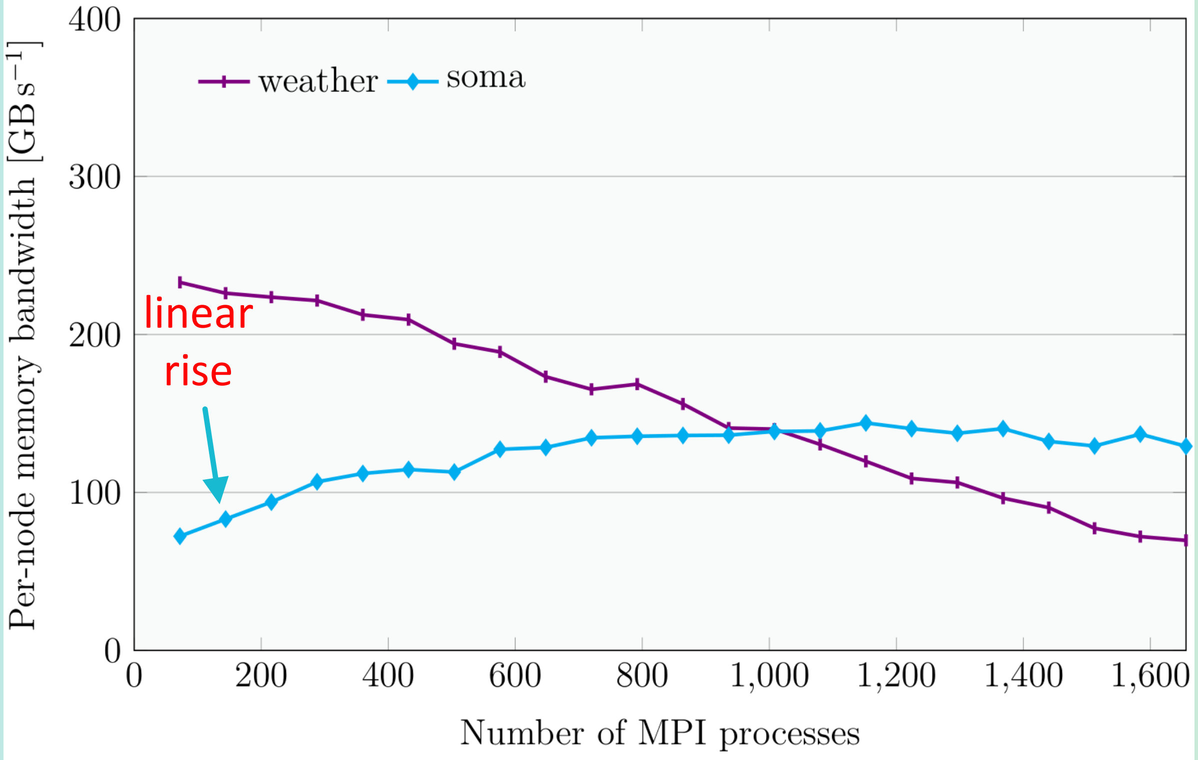


Sapphire Rapids

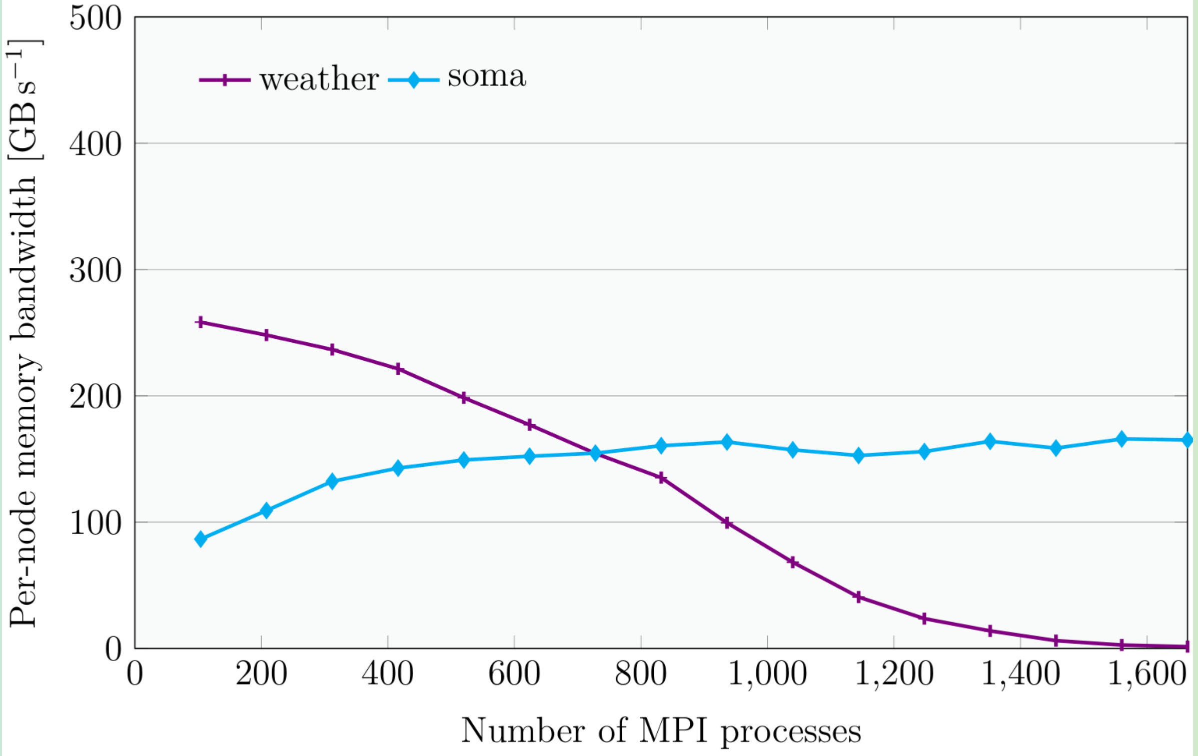


Memory bandwidth: weather and soma

Ice Lake

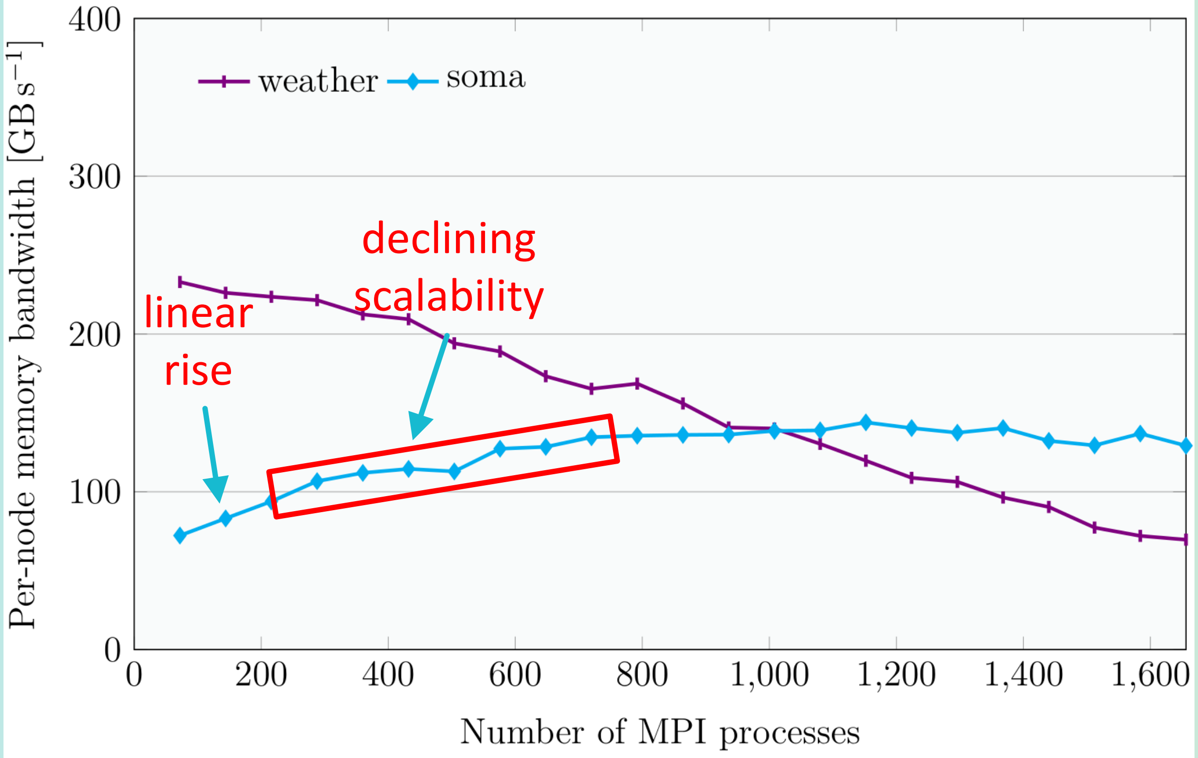


Sapphire Rapids

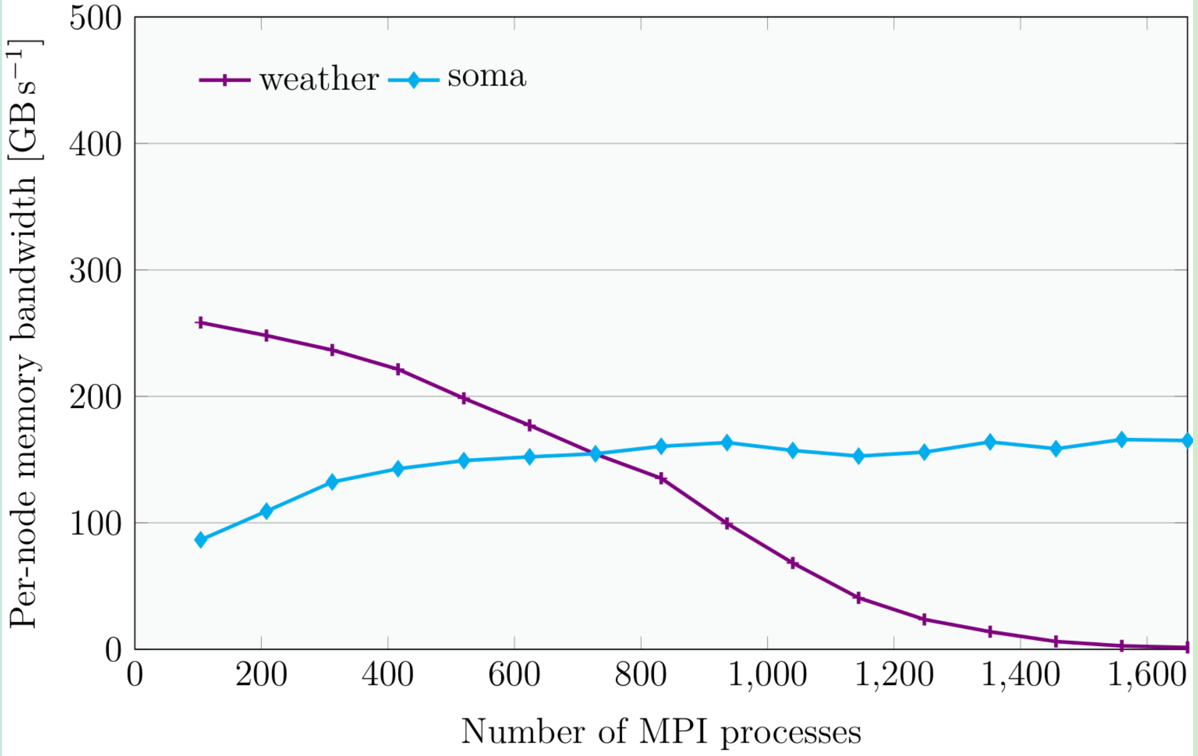


Memory bandwidth: weather and soma

Ice Lake

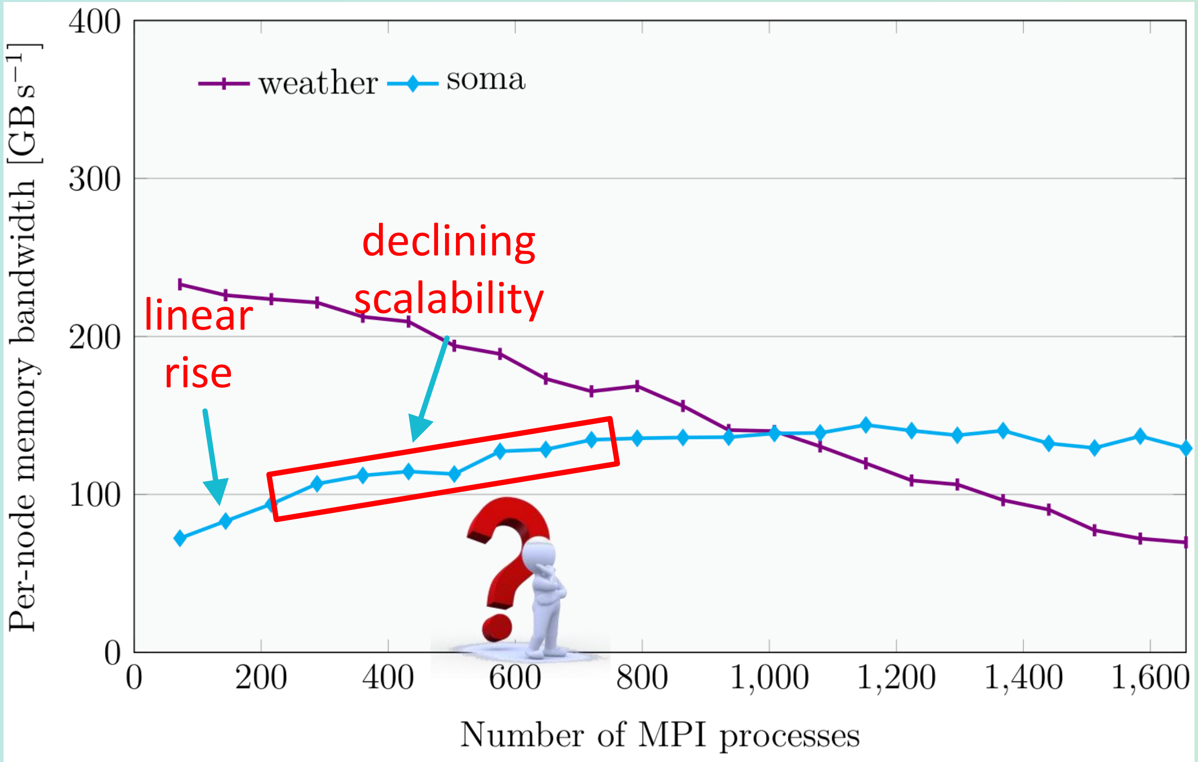


Sapphire Rapids

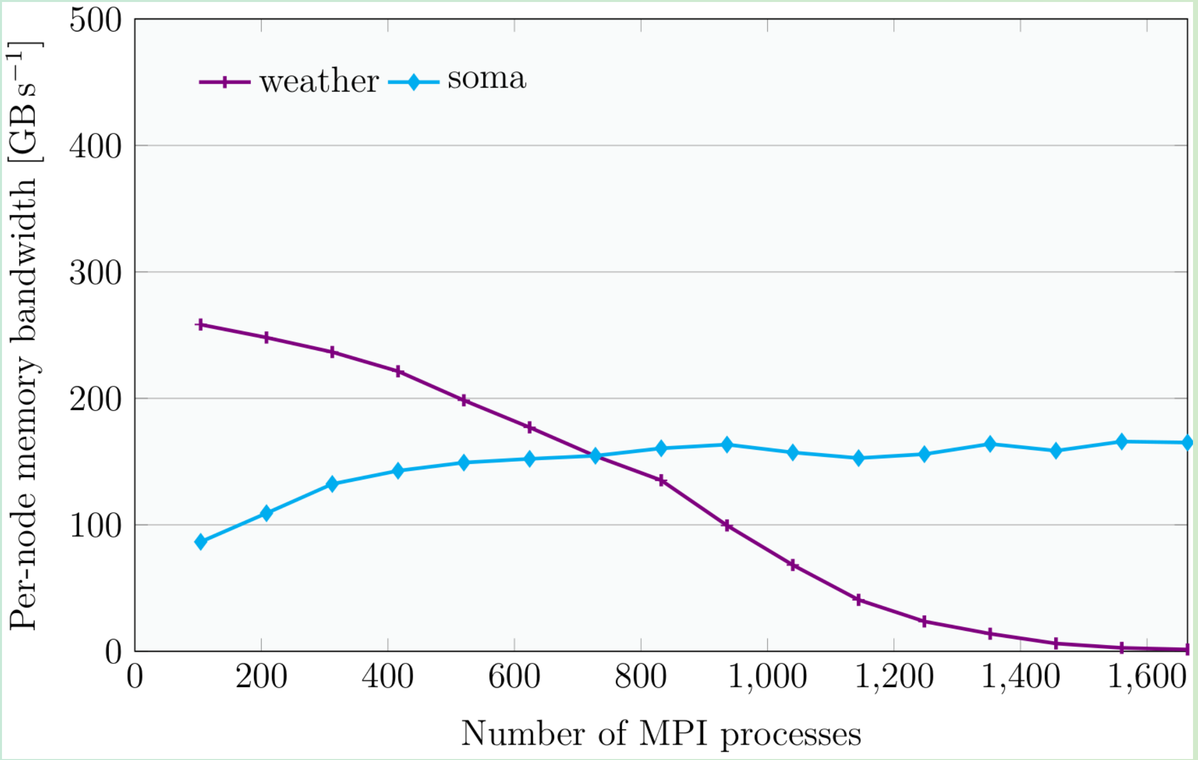


Memory bandwidth: weather and soma

Ice Lake

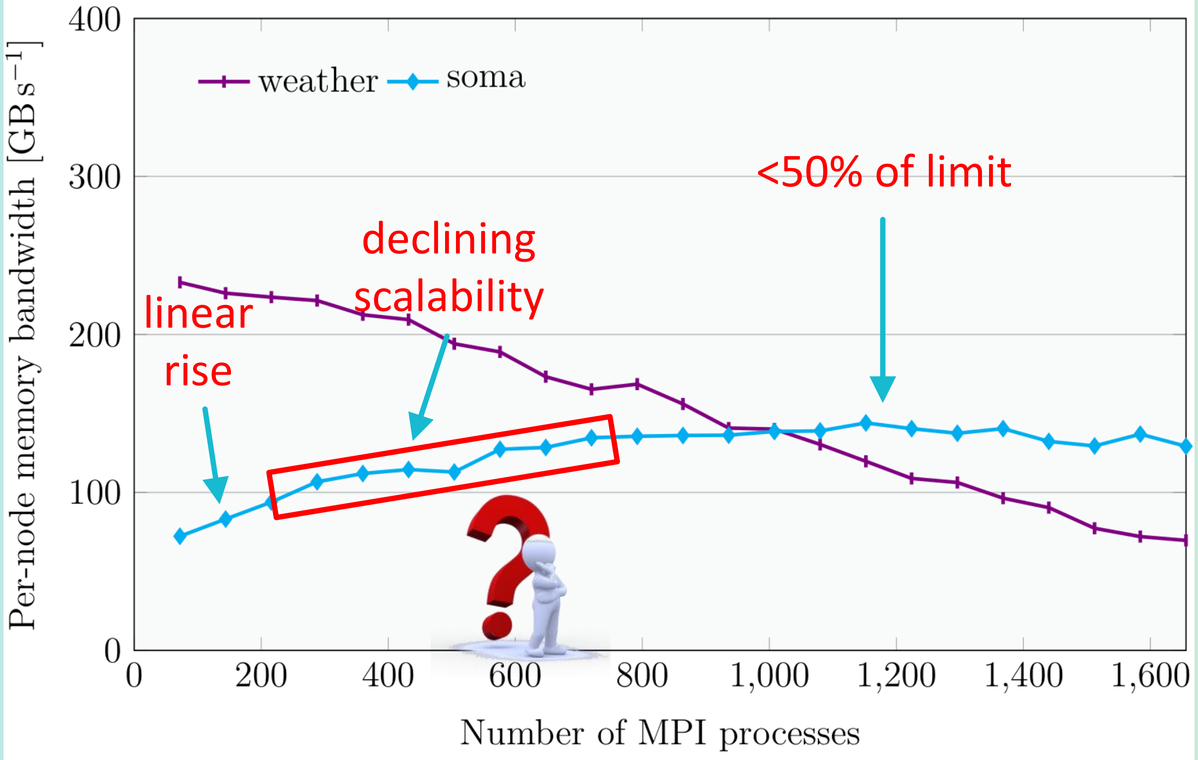


Sapphire Rapids

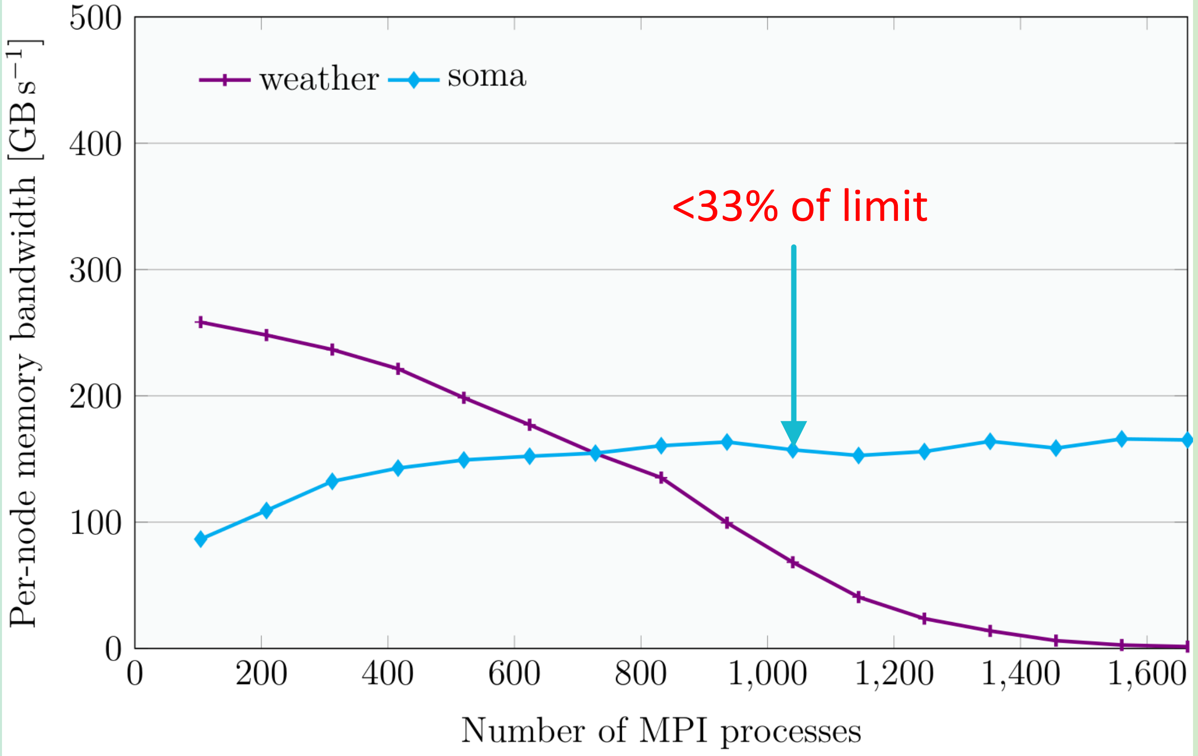


Memory bandwidth: weather and soma

Ice Lake

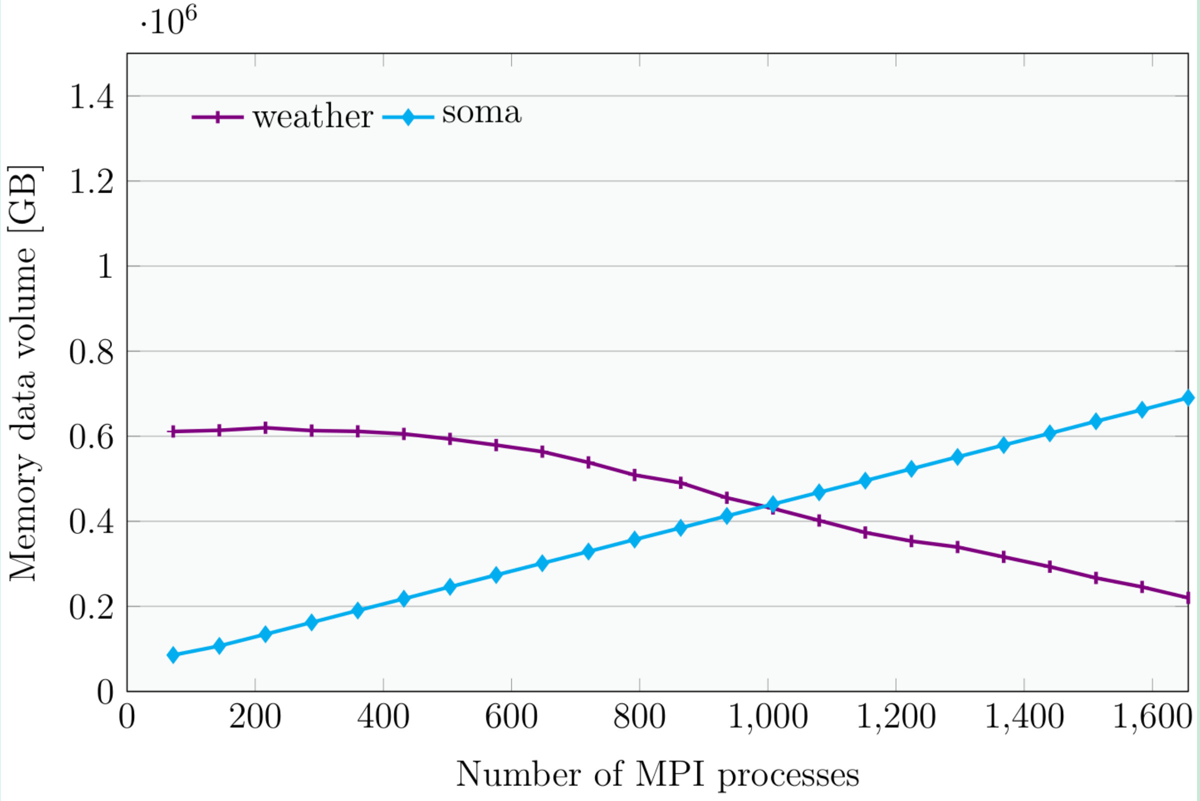


Sapphire Rapids

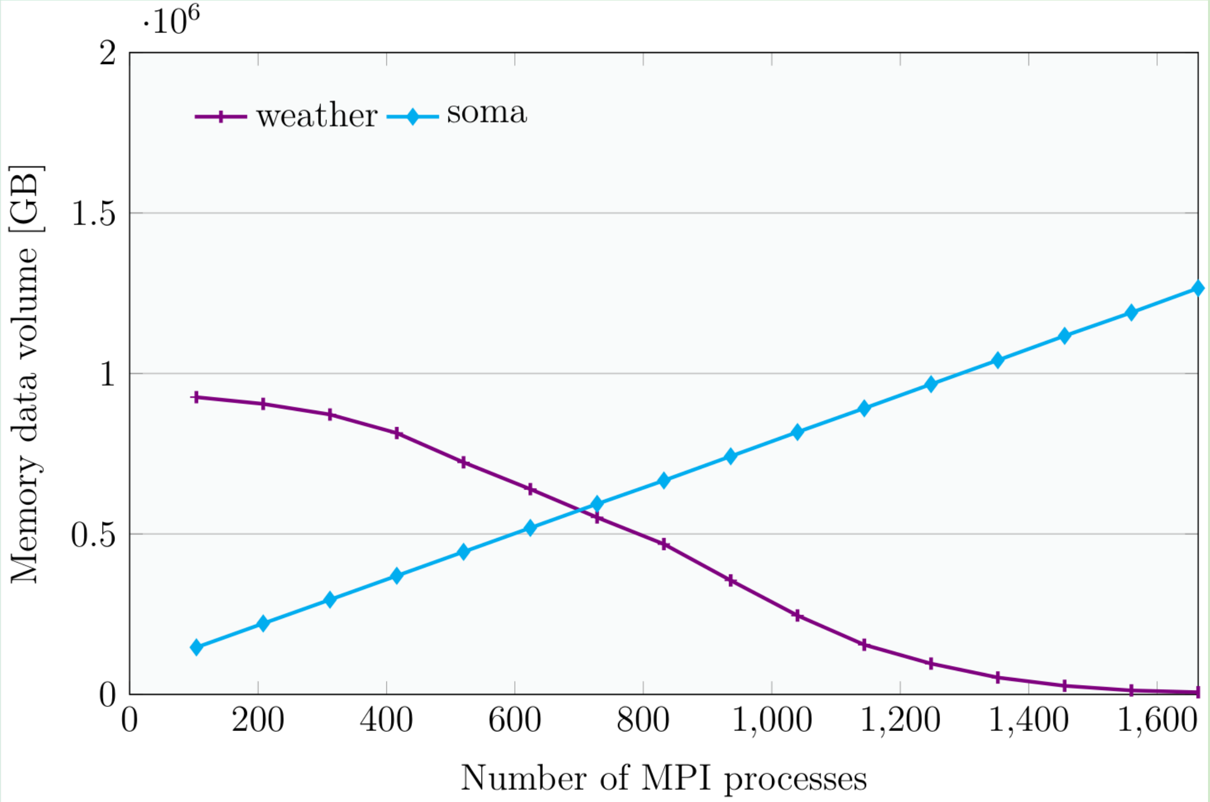


Memory data volume: Ice Lake and Sapphire Rapids cluster

Ice Lake

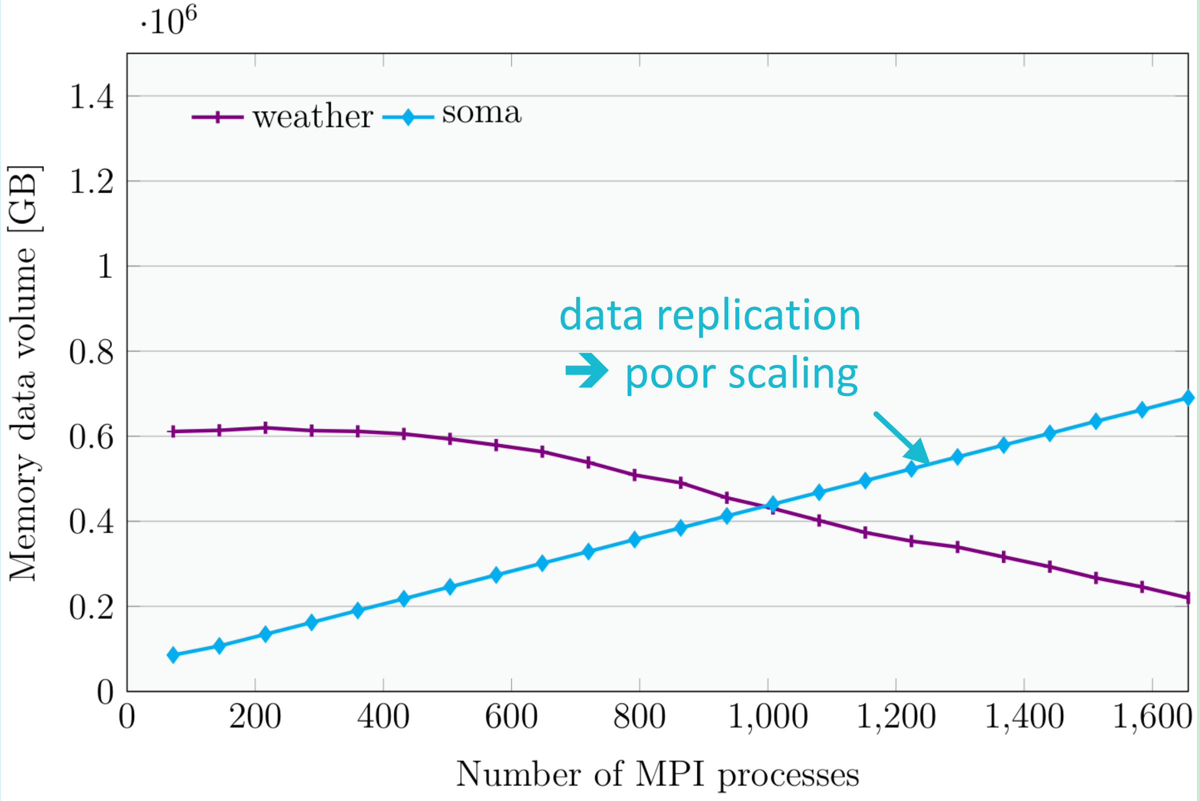


Sapphire Rapids

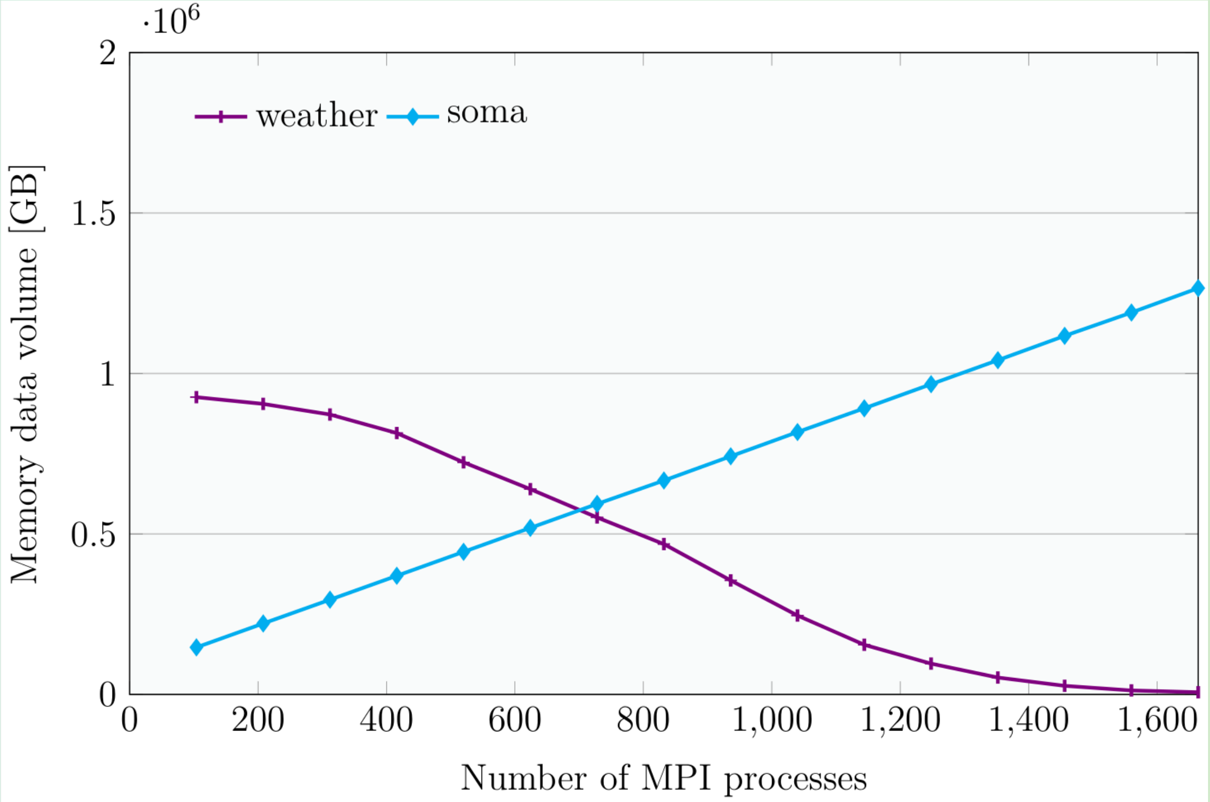


Memory data volume: Ice Lake and Sapphire Rapids cluster

Ice Lake

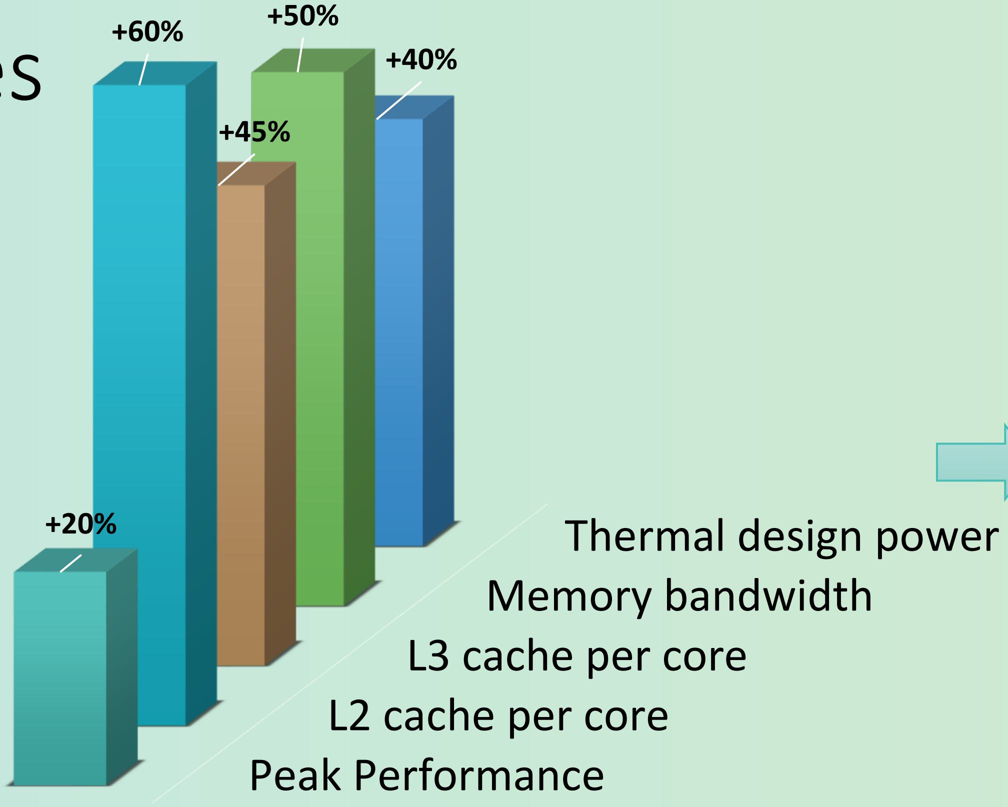


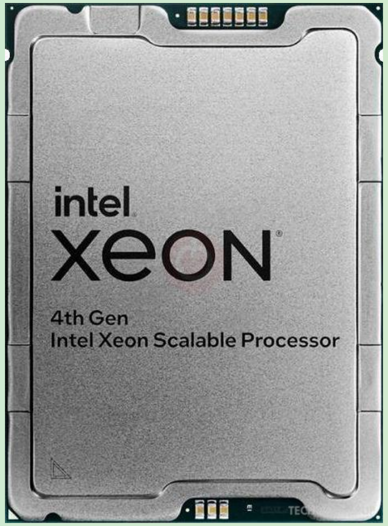
Sapphire Rapids



Take-home messages

Ice Lake

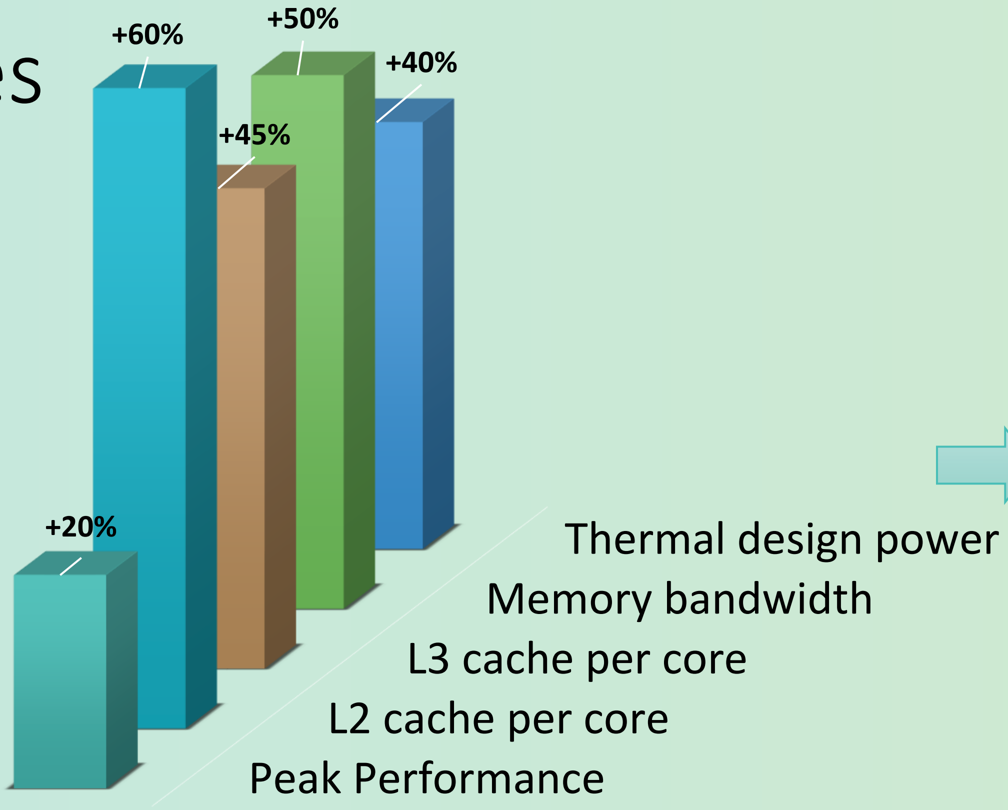


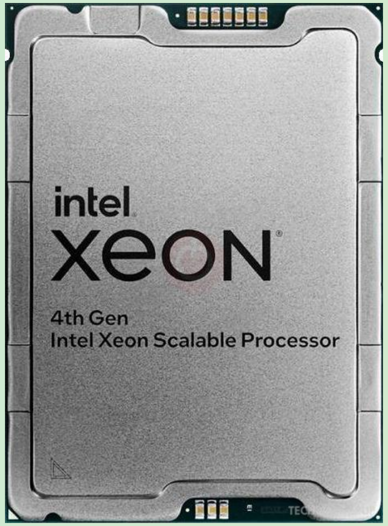


Sapphire Rapids

Take-home messages

Ice Lake



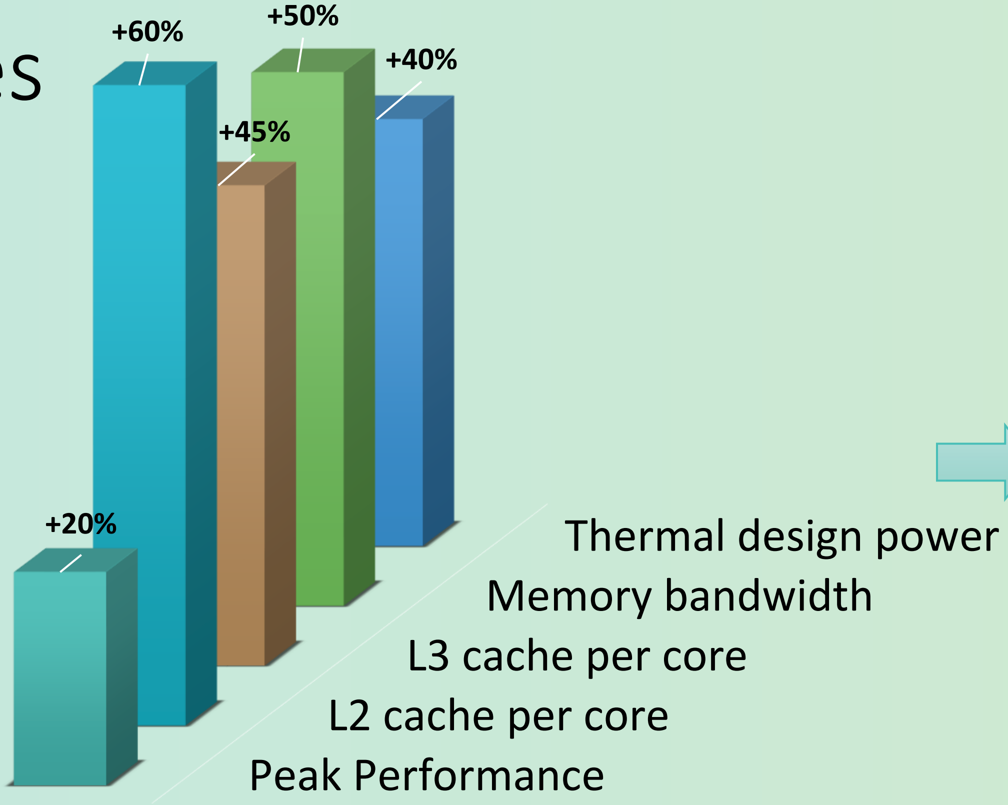


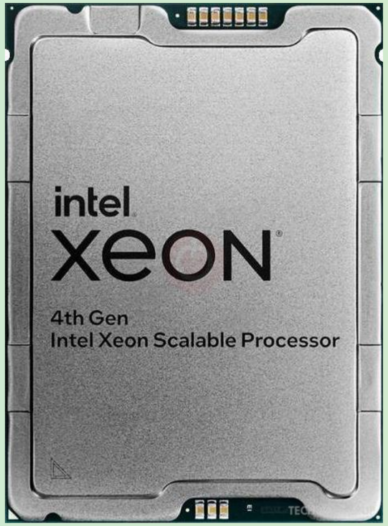
Sapphire Rapids

- SPEChpc 2021 is representative of many HPC workloads

Take-home messages

Ice Lake



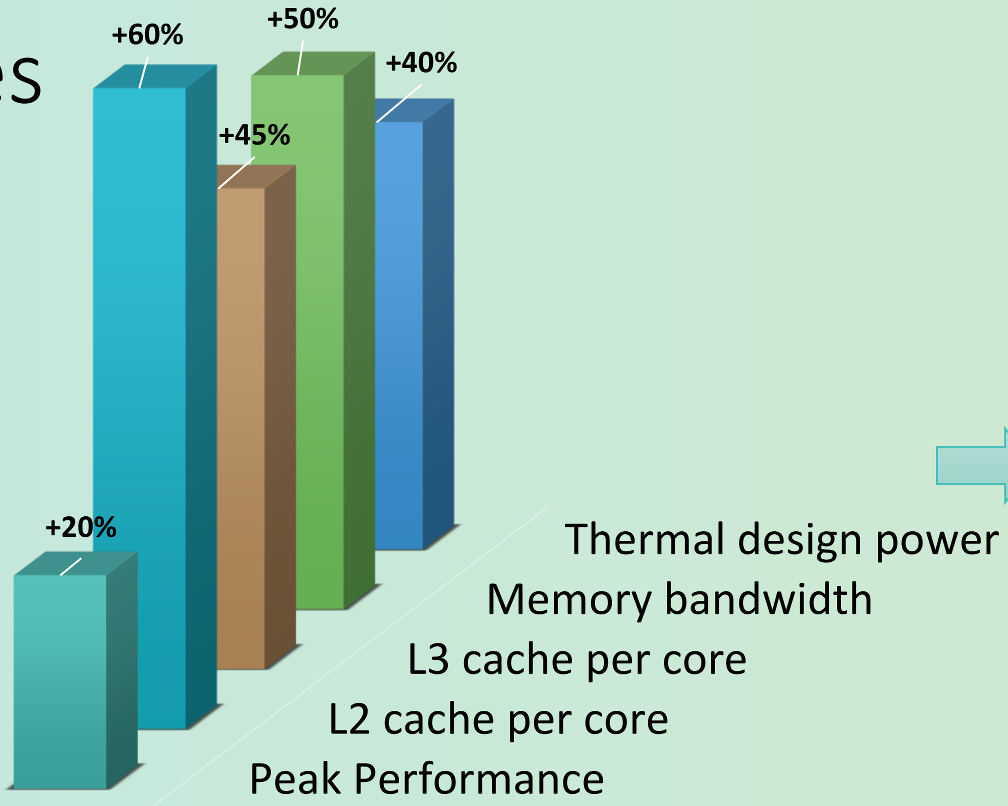


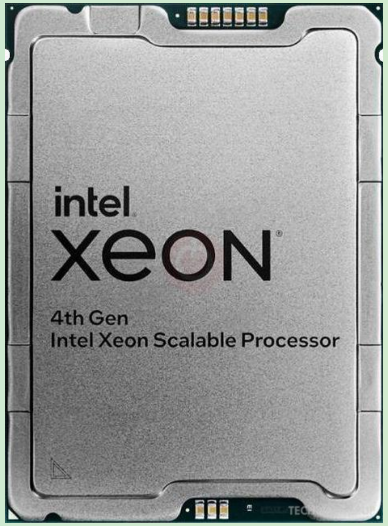
Sapphire Rapids

- SPEChpc 2021 is representative of many HPC workloads
- Only memory-bound code shows an energy advantage of Sapphire Rapids vs. Ice Lake

Take-home messages

Ice Lake





Sapphire Rapids

- SPEChpc 2021 is representative of many HPC workloads
- Only memory-bound code shows an energy advantage of Sapphire Rapids vs. Ice Lake
- Due to the dominance of chip idle power and code scaling characteristics in modern Intel server CPUs, race-to-idle is the most important strategy for E_{min} and EDP_{min}

GitHub

RRZE-HPC/PMBS23-AD

Zenodo

DOI: 10.5281/zenodo.8338037

Detailed Data Artifact Appendix

SC23 Denver, CO | i am hpc.

Title

SPEChpc 2021 Benchmarks
on Ice Lake and Sapphire Rapids InfiniBand Clusters:
A Performance and Energy Case Study

Contact

Ayesha Afzal
ayesha.afzal@fau.de

KONWIHR

OMI4papps

NHR FAU

FAU Friedrich-Alexander-Universität Erlangen-Nürnberg