Concepts and Models of Parallel and Data-centric Programming

Foundations I

Lecture, Summer 2020

Dr. Christian Terboven <terboven@itc.rwth-aachen.de>

i12 High Performance Computing
RWTH AACHEN UNIVERSITY

Outline

0. Organization
1. Foundations
2. Shared Memory
3. GPU Programming
4. Bulk-Synchronous Parallelism
5. Message Passing
6. Distributed Shared Memory
7. Parallel Algorithms
8. Parallel I/O
9. MapReduce
10. Apache Spark

a. Cluster Architecture
b. Convergence of HPC and Big Data
c. Parallel Programming Teasers
d. Harsh Realities

Foundations

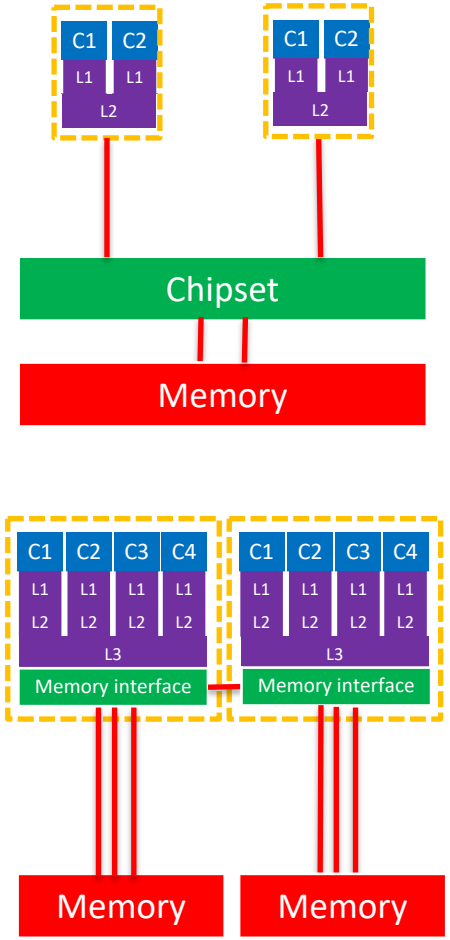
Cluster-Architecture

How would you draw a computer?

- Early multicore design
  - Uniform Memory Architecture (UMA)
  - Flat Memory design

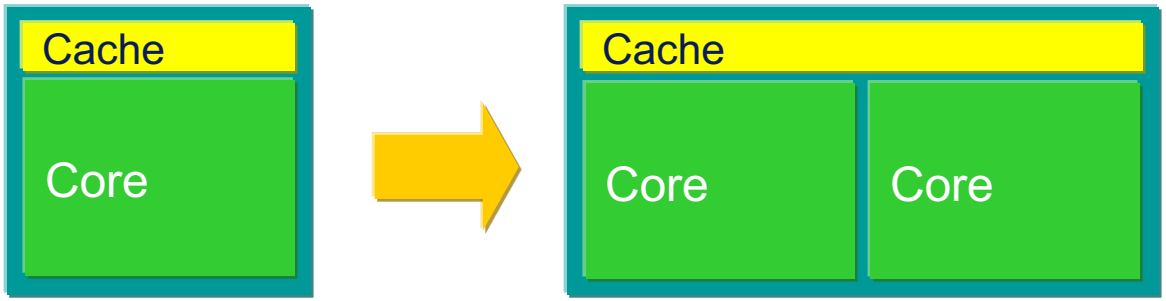
- “A shared-memory parallel computer is a system in which a number of running CPUs work on a common, shared physical address space” (Wellein & Hager)

- Recent multicore design
  - ccNUMA (Cache Coherent Non-Uniform Memory Architecture)
  - Memory Interface + HT/QPI provides inter-socket connectivity



Multi-Core Processors

- A multi-core processor integrates two or more independent computing cores on a single die
- Rule of thumb: Reduction of 1% voltage and 1% frequency reduces the power consumption by 3% and the performance by 0.66%.

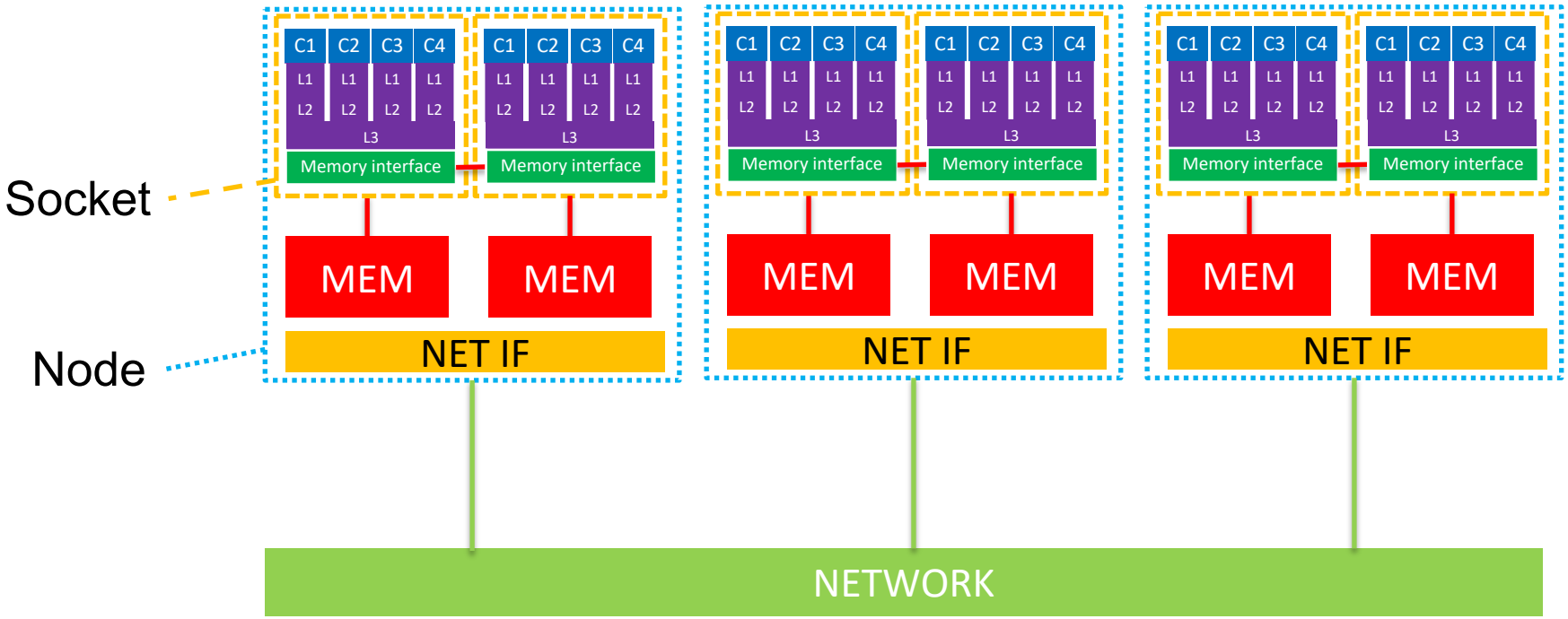


Voltage	=	1	Voltage	=	-15%
Freq	=	1	Freq	=	-15%
Area	=	1	Area	=	2
Power	=	1	Power	=	~1
Perf	=	1	Perf	=	~1.8

(Based on slides from Shekhar Borkar, Intel Corp.)

How would you draw a Cluster?

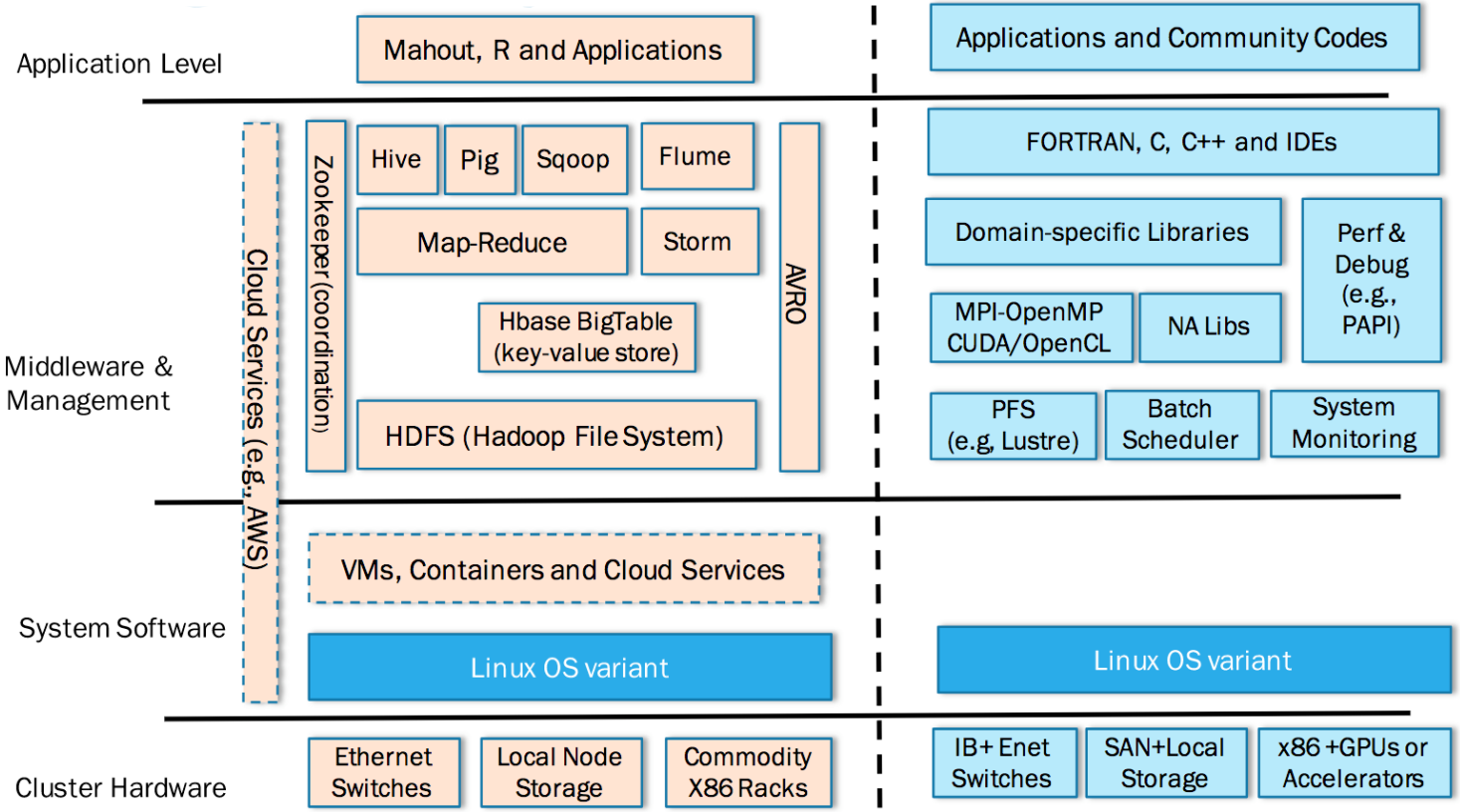
- All large-scale parallel computing systems are neither purely of the shared nor the distributed memory type. They are a mixture of both.



Convergence of HPC and Big Data (Infrastructure)

Infrastructure & Ecosystems

- US National Strategic Computing Initiative:
  - Increasing coherence between the technology base used for modeling and simulation and that used for data analytic computing.



(Figure by Beckman & Reed, 2015)

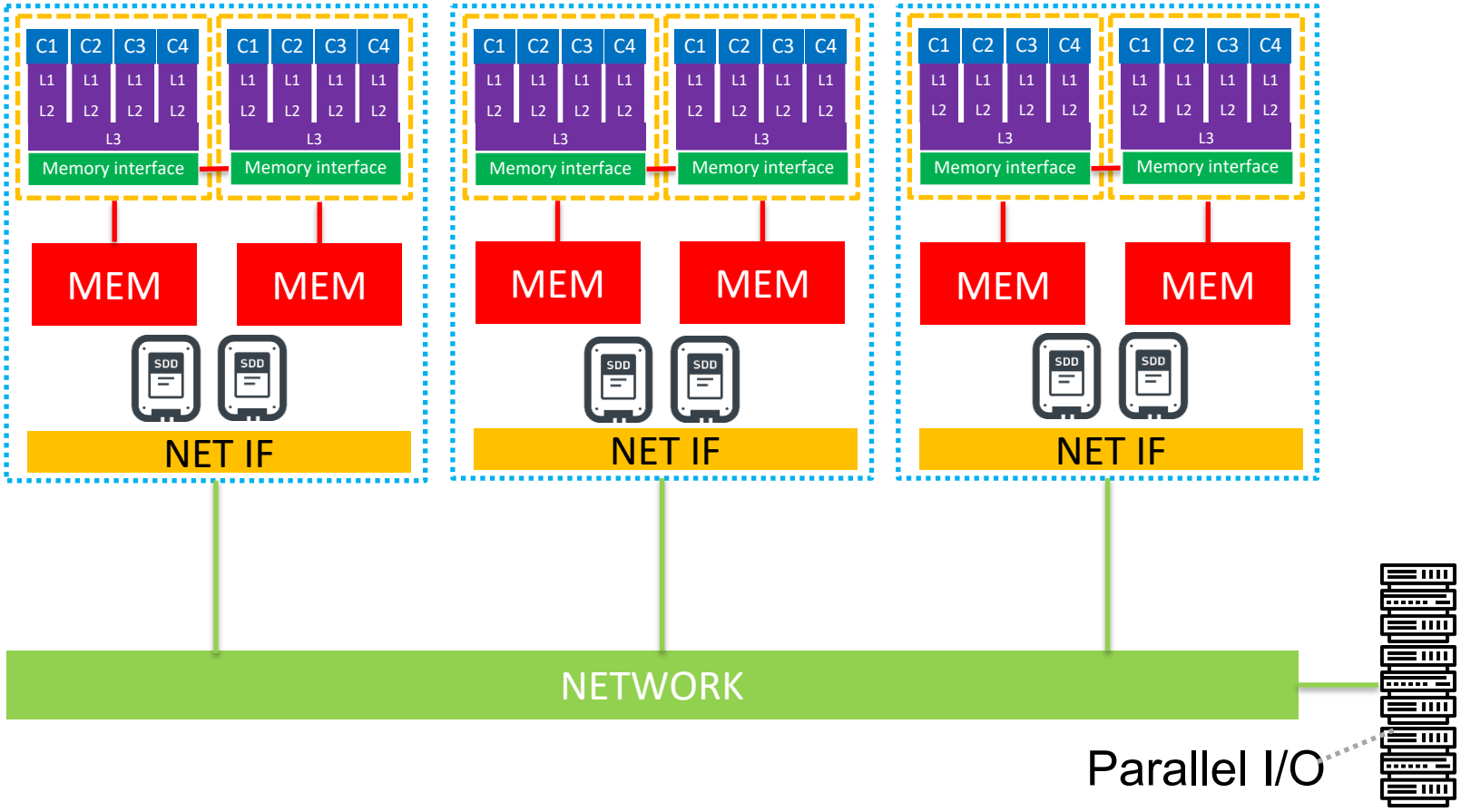
Convergence

- Motivation:
  - Science domains broadly become more intertwined
  - Large-scale data-centric computing

- Enabling technologies:
  - Docker and other container technologies
  - Integration of new I/O technologies

More Colors to the picture

- I/O is getting more attention
  - Network-attached parallel storage, local SSDs, non-volatile memory



What you have learnt

- A mile high view of Cluster architectures
  - Trend to multi-core and many-core
- Our view of how HPC and Big Data converge