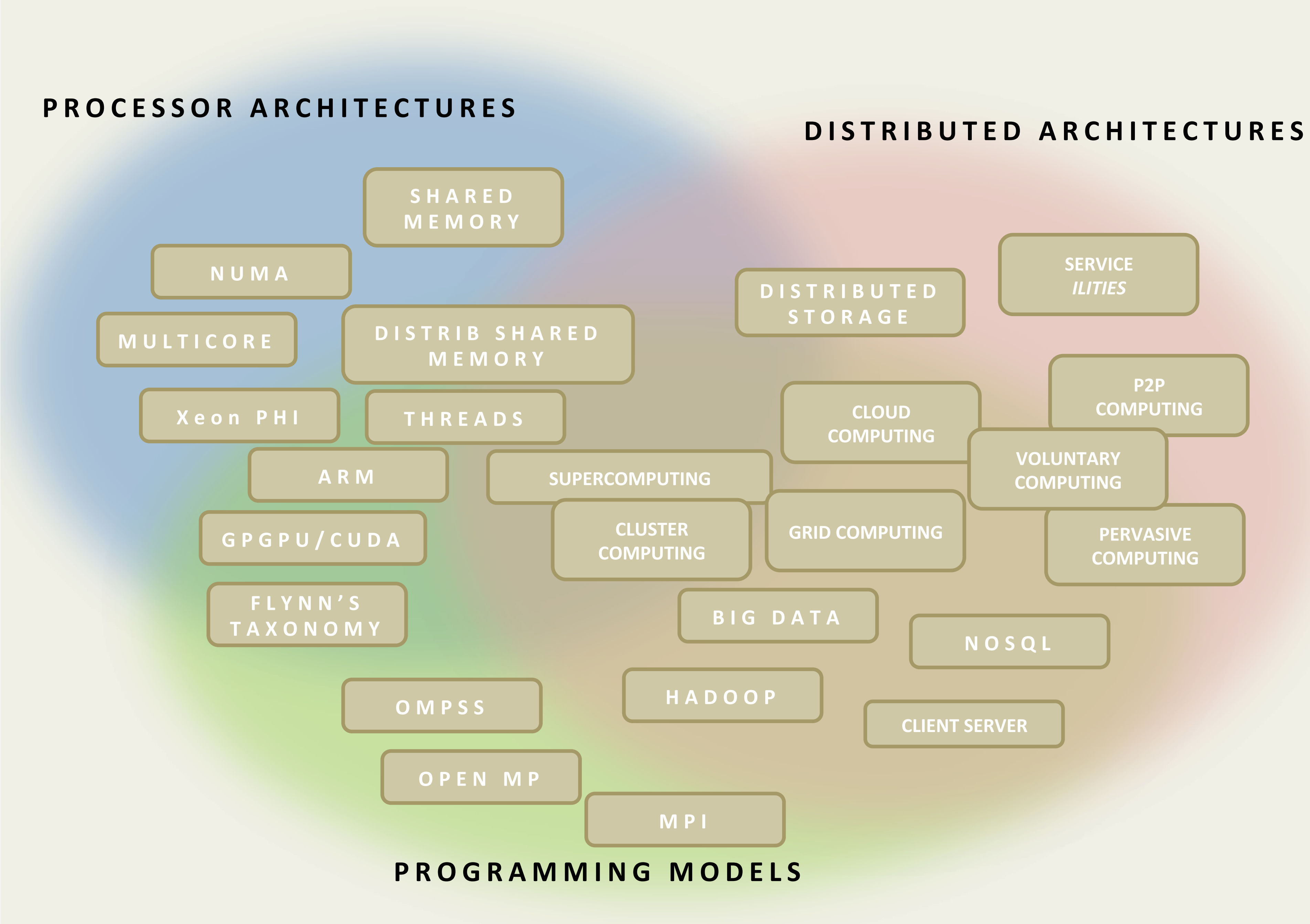
INTRODUCTION

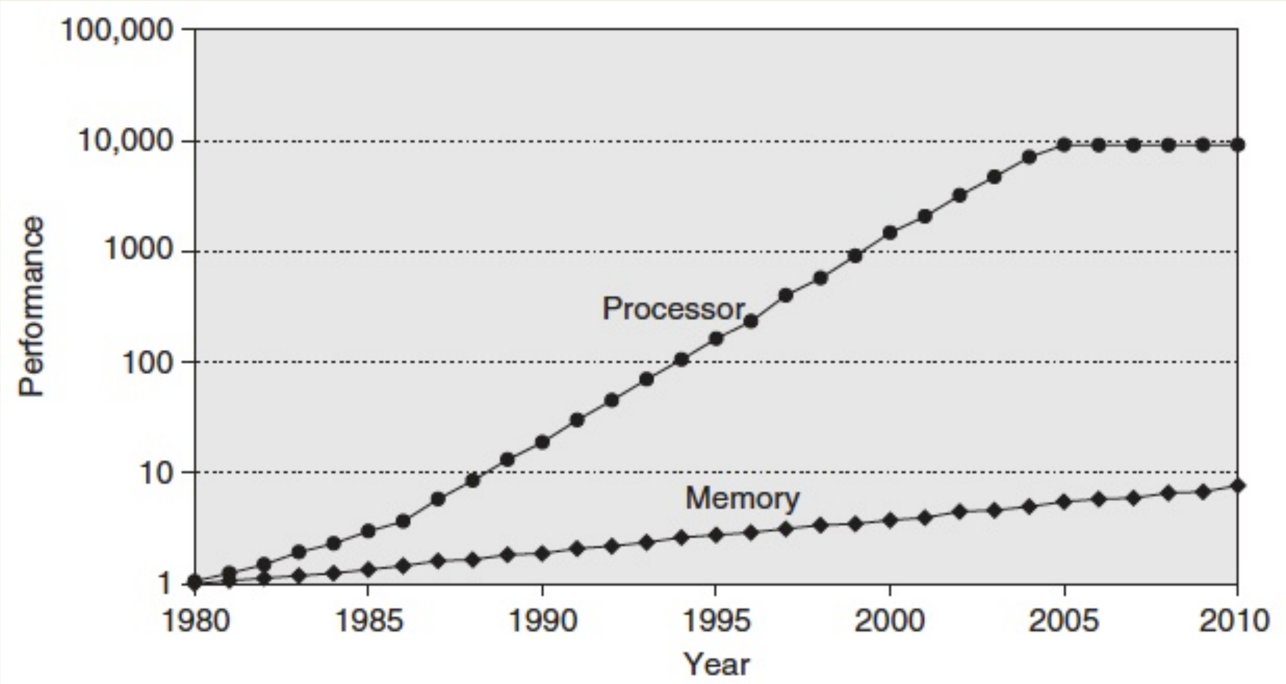
DISTRIBUTED SYSTEMS AND PARALLEL COMPUTING

Raúl RAMOS POLLÁN

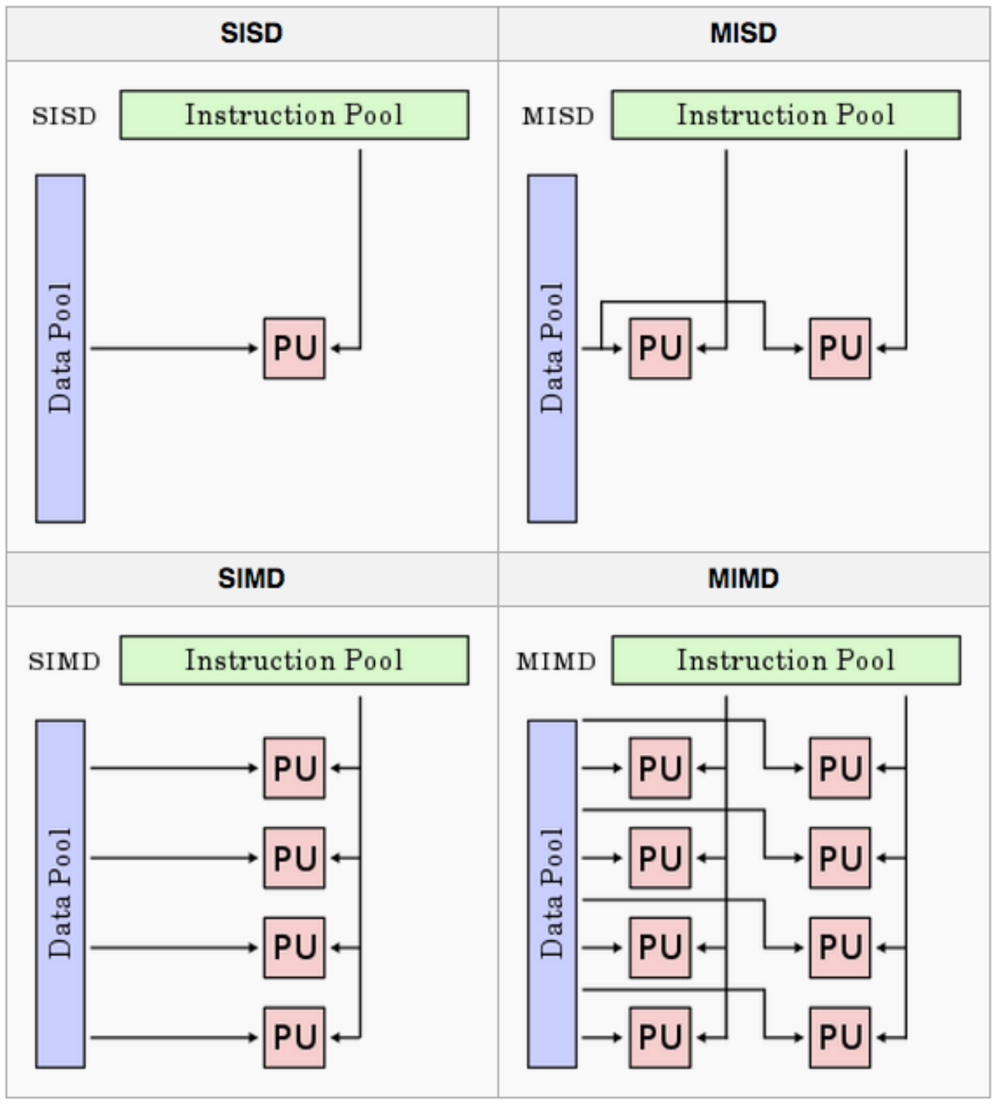

PROCESSOR ARCHITECTURES
DISTRIBUTED ARCHITECTURES
PROGRAMMING MODELS
SHARED MEMORY
NUMA
MULTICORE
DISTRIB SHARED MEMORY
Xeon PHI
THREADS
ARM
GPGPU/CUDA
FLYNN'S TAXONOMY
SUPERCOMPUTING
CLUSTER COMPUTING
DISTRIBUTED STORAGE
SERVICE ILITIES
CLOUD COMPUTING
P2P COMPUTING
VOLUNTARY COMPUTING
GRID COMPUTING
PERVASIVE COMPUTING
BIG DATA
NOSQL
HADOOP
CLIENT SERVER
OMPSS
OPEN MP
MPI

CPU + RAM PERFORMANCE

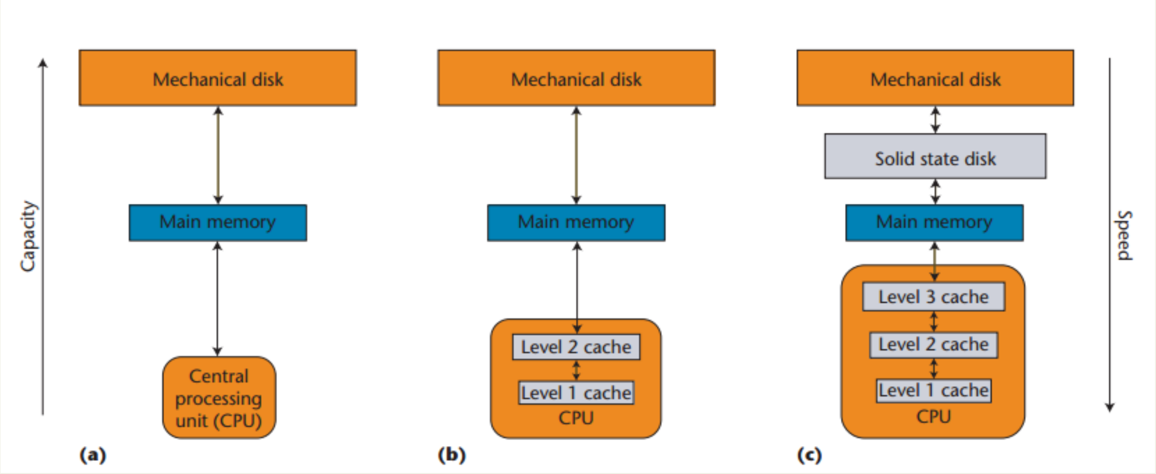
IT’S ALL ABOUT THE BANDWIDTH!!!!

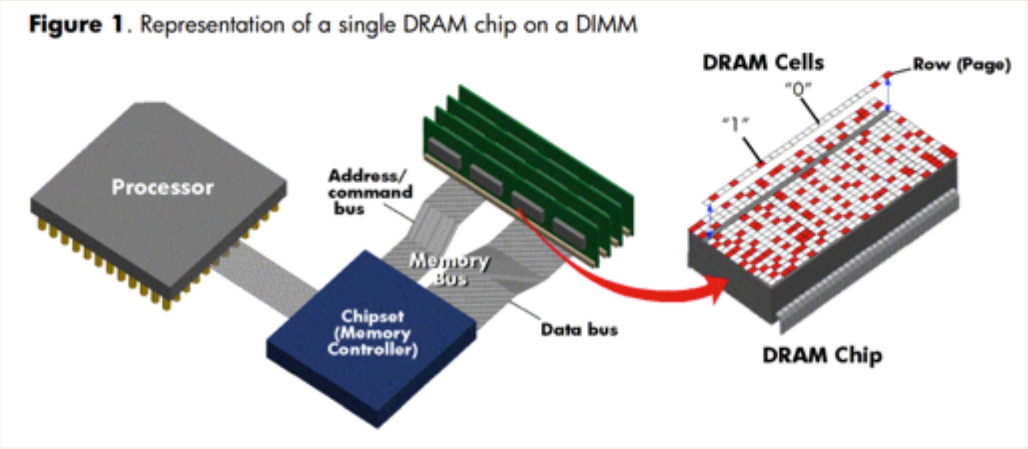


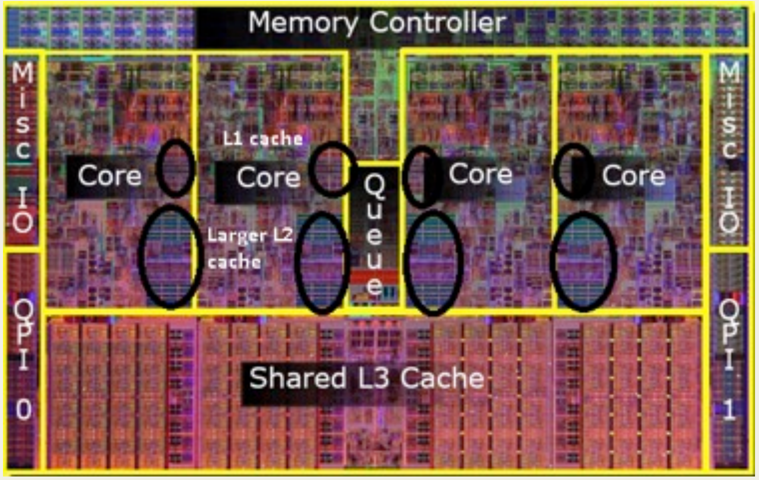
FLYNN'S TAXONOMY



NUMA ARCHITECTURES



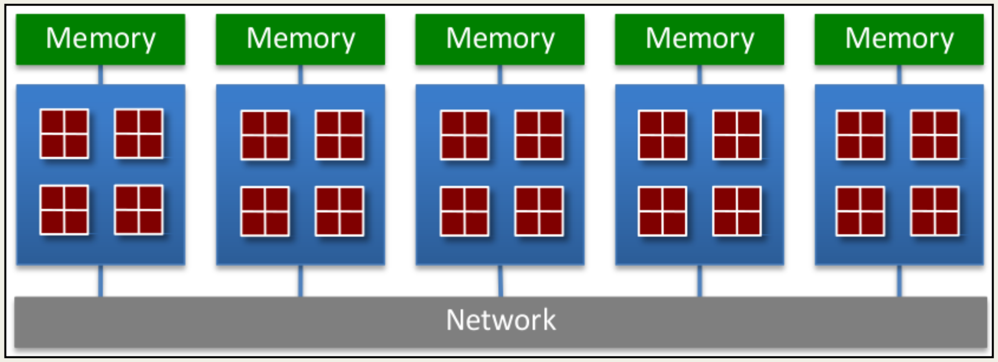




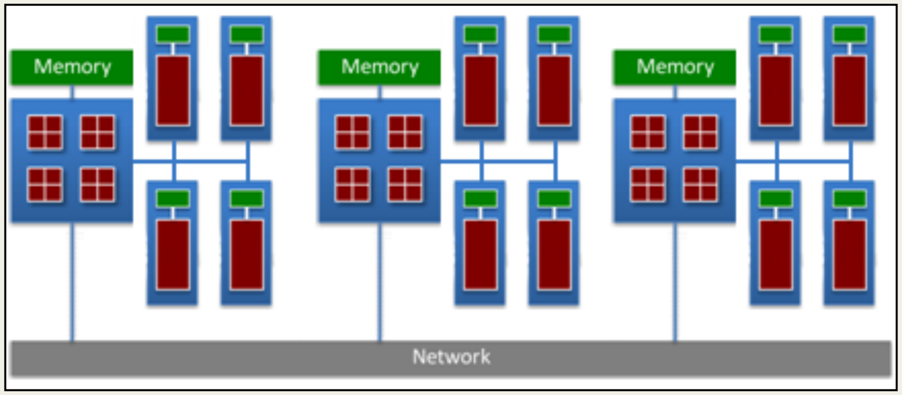
PROCESSORS ARCHITECTURES

response to physical limits on clock speed on 2005: multicore + coprocessors + GPUs

MULTICORE CLUSTER



MULTICORE MULTICOPROCESSOR CLUSTER



TOP 500

Rank	Site	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	National Super Computer Center in Guangzhou China	**Tianhe-2 (MilkyWay-2)** - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P NUDT	3,120,000	33,862.7	54,902.4	17,808
2	DOE/SC/Oak Ridge National Laboratory United States	**Titan** - Cray XK7 , Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x Cray Inc.	560,640	17,590.0	27,112.5	8,209
3	DOE/NNSA/LLNL United States	**Sequoia** - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom IBM	1,572,864	17,173.2	20,132.7	7,890
4	RIKEN Advanced Institute for Computational Science (AICS) Japan	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect Fujitsu	705,024	10,510.0	11,280.4	12,660
5	DOE/SC/Argonne National Laboratory United States	**Mira** - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM	786,432	8,586.6	10,066.3	3,945
6	Swiss National Supercomputing Centre (CSCS) Switzerland	**Piz Daint** - Cray XC30, Xeon E5-2670 8C 2.600GHz, Aries interconnect , NVIDIA K20x Cray Inc.	115,984	6,271.0	7,788.9	2,325
7	Texas Advanced Computing Center/Univ. of Texas United States	**Stampede** - PowerEdge C8220, Xeon E5-2680 8C 2.700GHz, Infiniband FDR, Intel Xeon Phi SE10P Dell	462,462	5,168.1	8,520.1	4,510
8	Forschungszentrum Juelich (FZJ) Germany	**JUQUEEN** - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect IBM	458,752	5,008.9	5,872.0	2,301
9	DOE/NNSA/LLNL United States	**Vulcan** - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect IBM	393,216	4,293.3	5,033.2	1,972
10	Leibniz Rechenzentrum Germany	**SuperMUC** - iDataPlex DX360M4, Xeon E5-2680 8C 2.70GHz, Infiniband FDR IBM	147,456	2,897.0	3,185.1	3,423

80% Intel

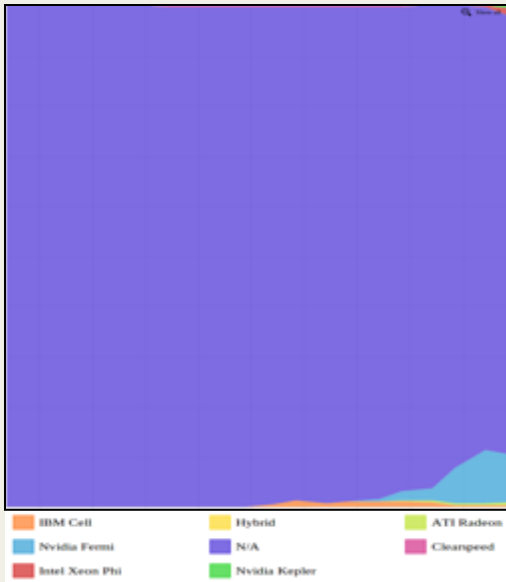
80% Infiniband GigabitEthernet

96% Linux

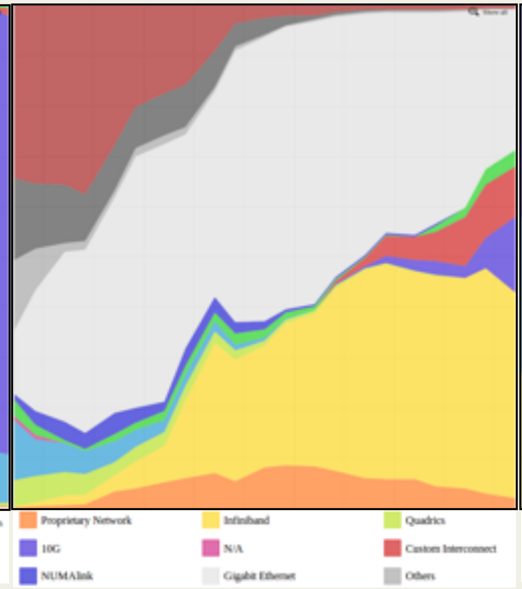
85% Cluster architecture

TOP 500 2003-2013

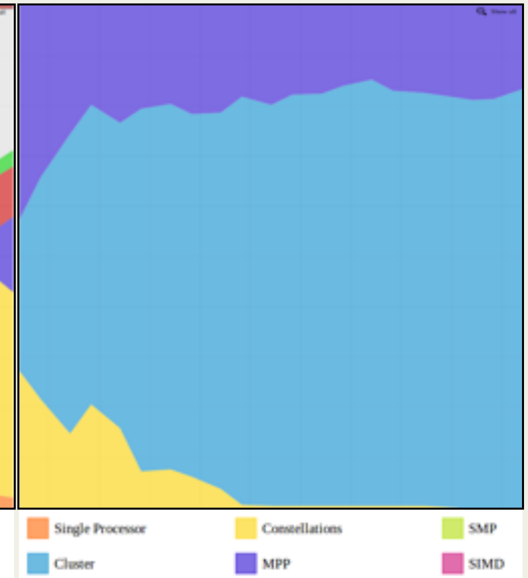
COPROCESSOR



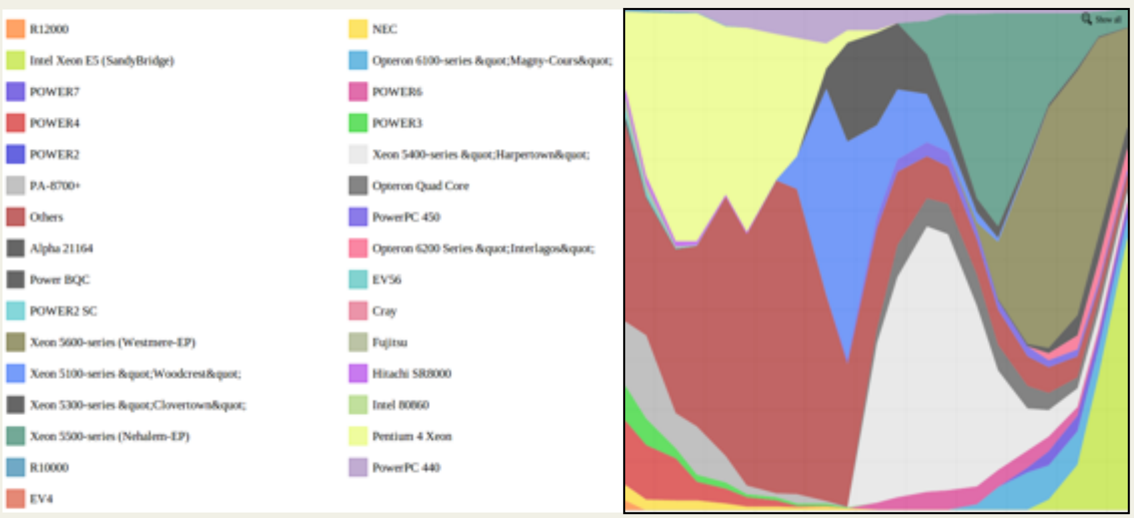
INTERCONNECT FAMILY



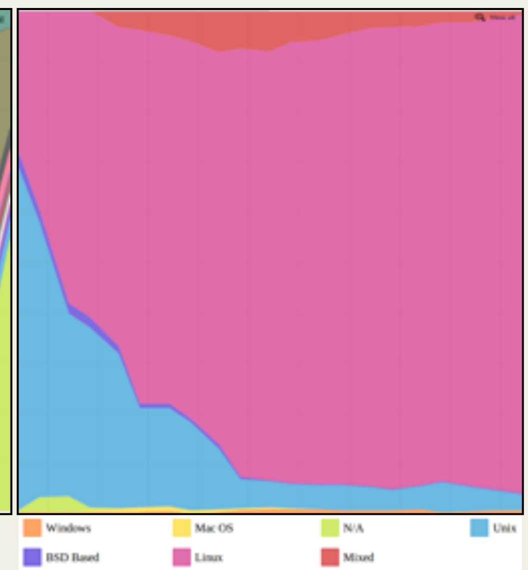
ARCHITECTURE



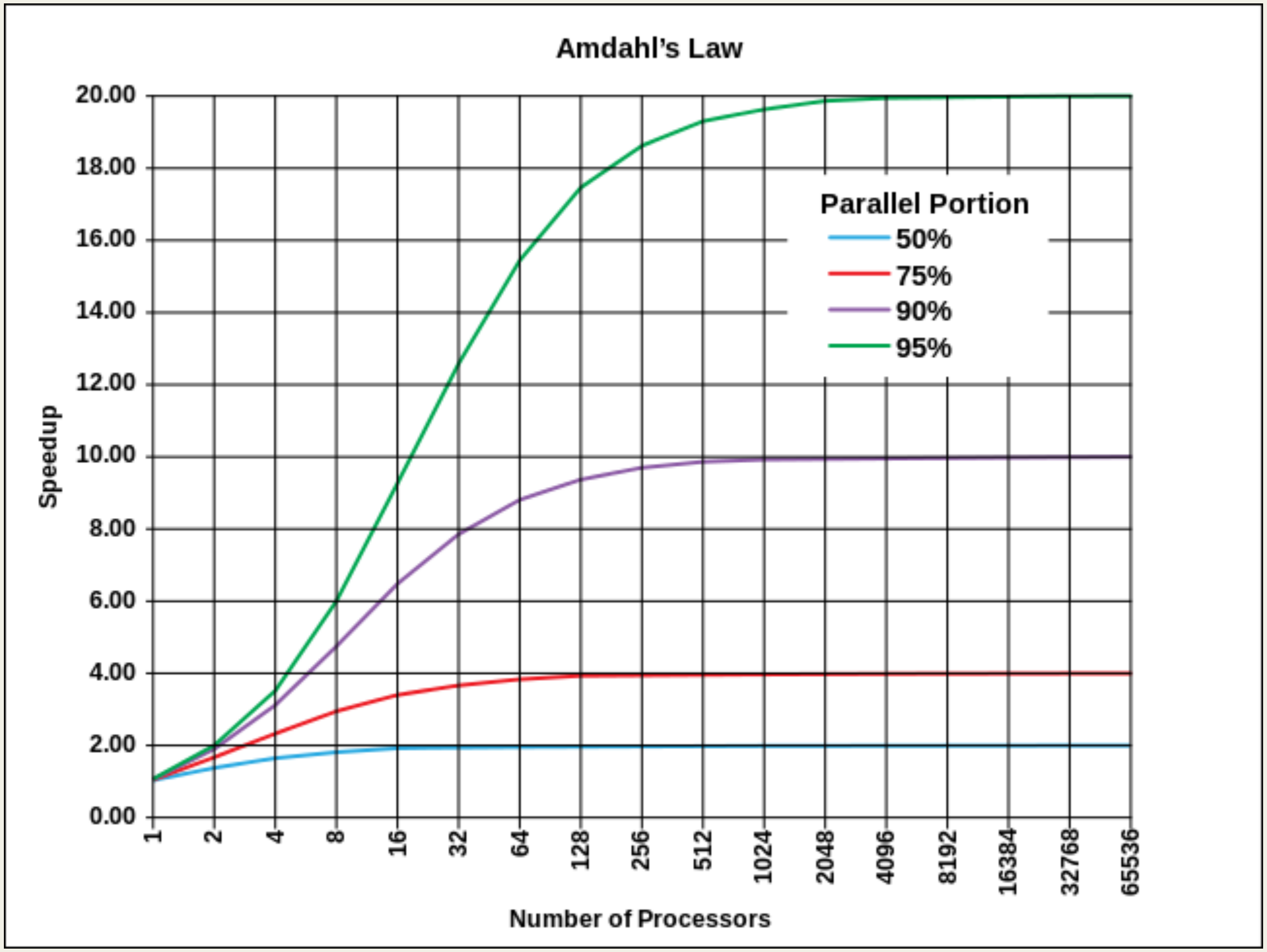
PROCESSOR



OS FAMILY



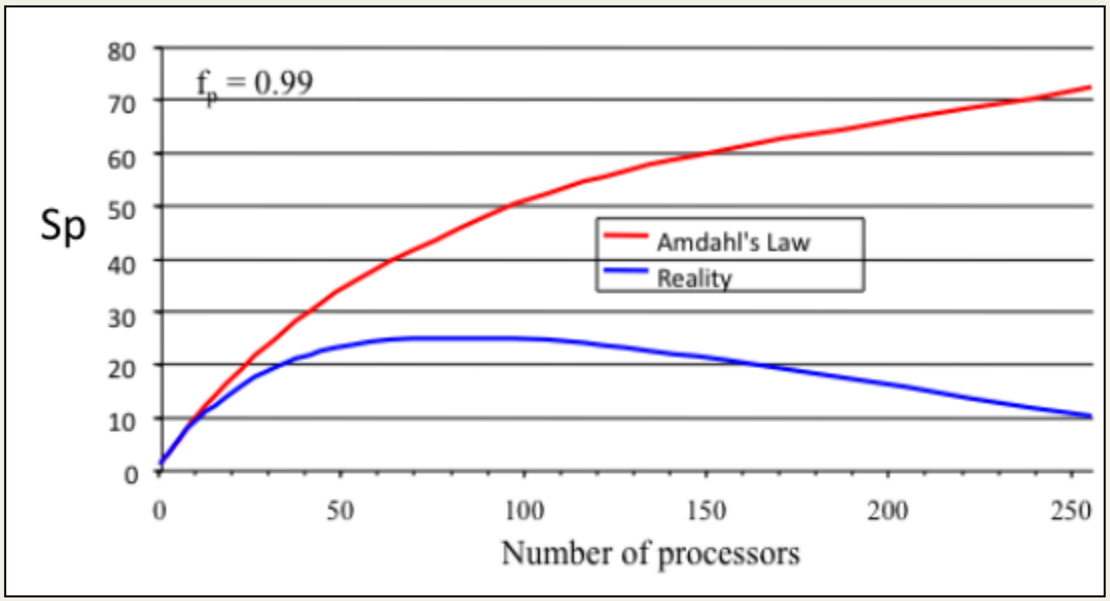
AMDAHL’S LAW


Amdahl’s Law
Speedup
20.00
18.00
16.00
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
1
2
4
8
16
32
64
128
256
512
1024
2048
4096
8192
16384
32768
65536
Number of Processors
Parallel Portion
50%
75%
90%
95%

AMDAHL'S LAW AND REALITY

Loss of parallelism due to:

- Load balancing (synch/wait among computing units)
- Scheduling (shared processors and/or memory)
- I/O latencies

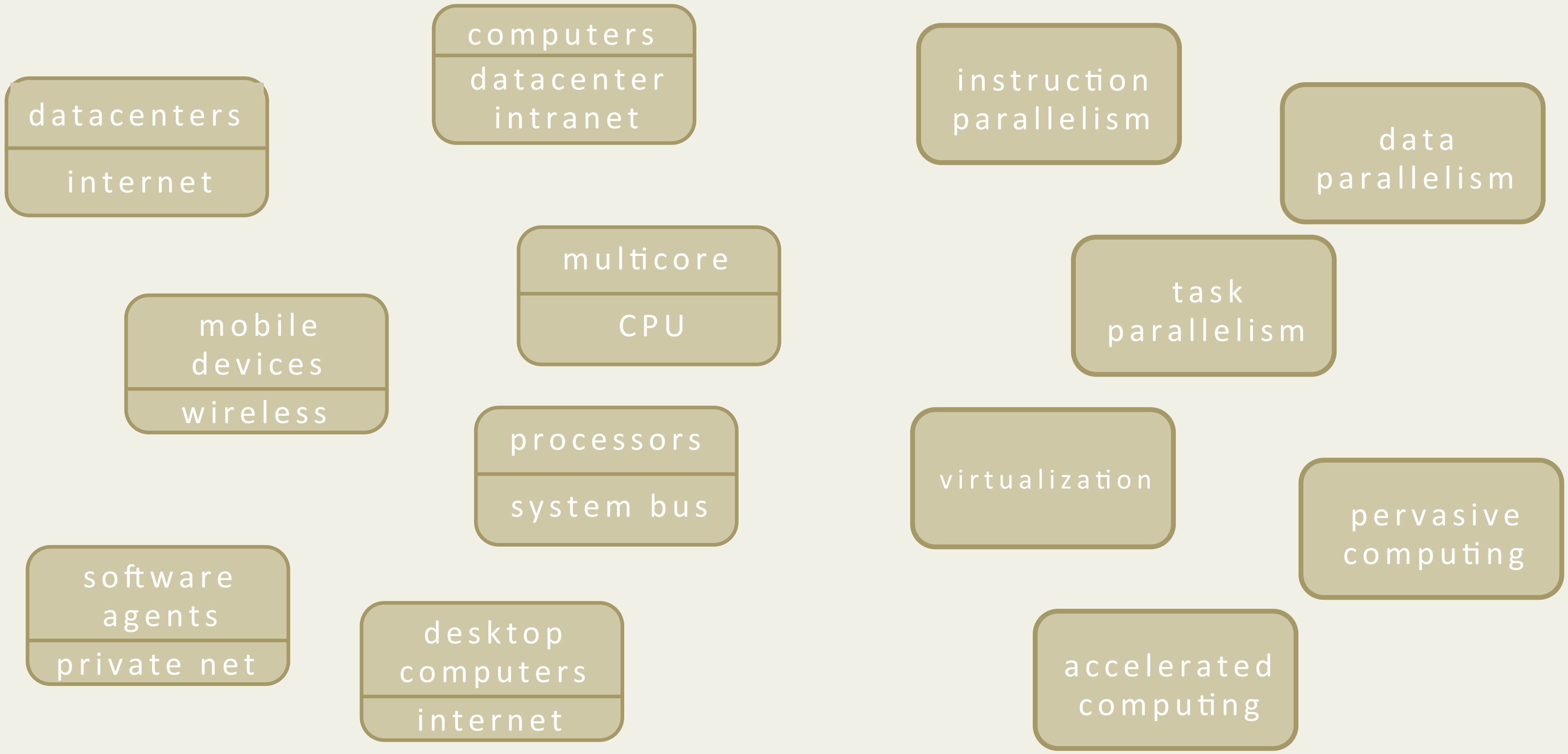


DISTRIBUTED SYSTEMS

Computer resources linked by some interconnect

PARALLEL COMPUTING

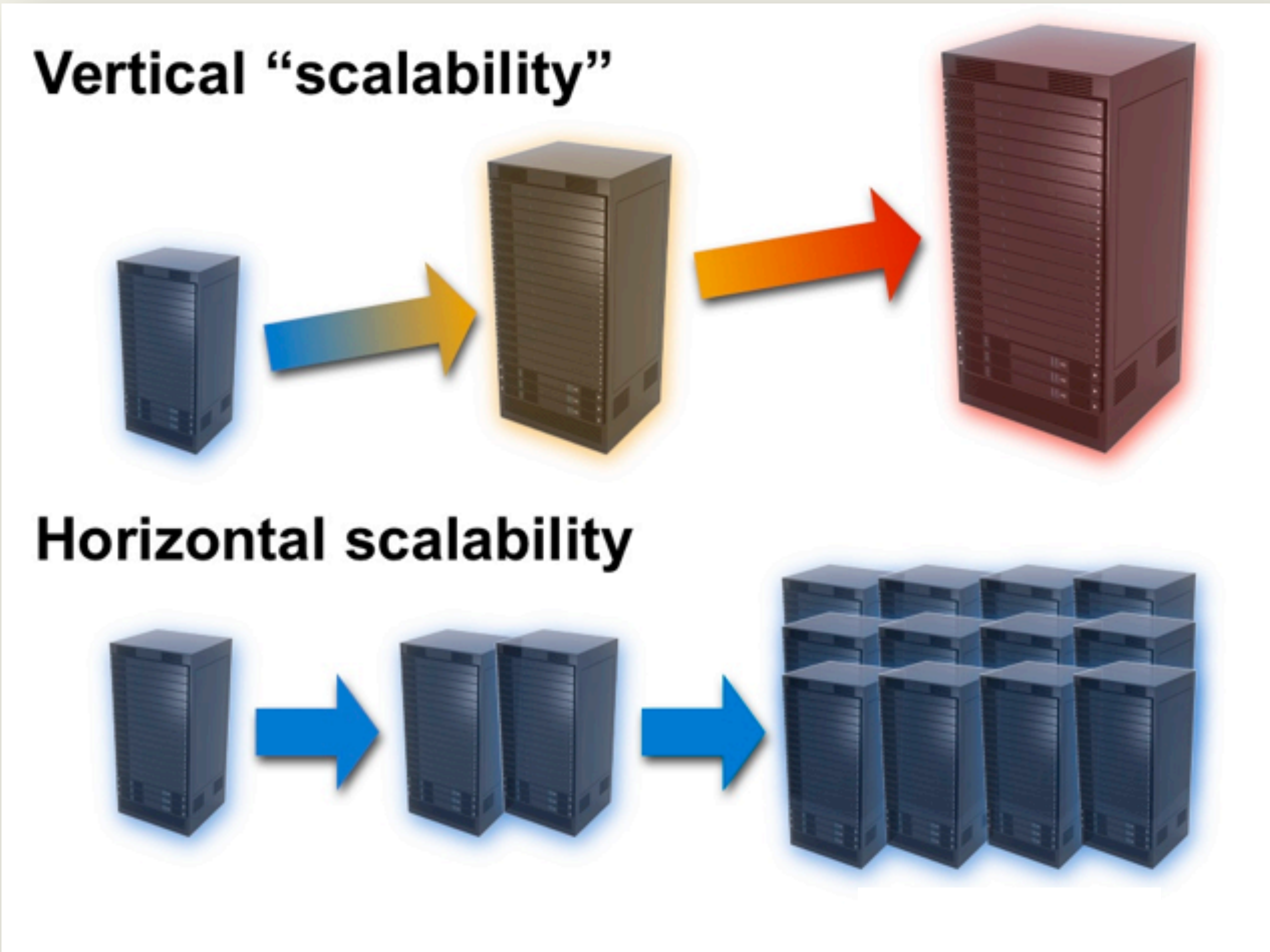
Methods and architectures for executing calculations simultaneously



distributed systems as a kind of parallel computing

parallel computing over distributed systems

SCALABILITY



scale up
traditional DBs

scale out
noSQL

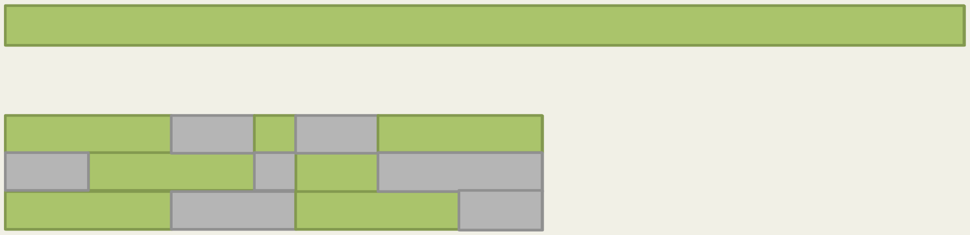
seek "triviality" for appropriate sw+hw architectures

recent technologies (virtualization, etc.) tend to favor the cost of scaling out

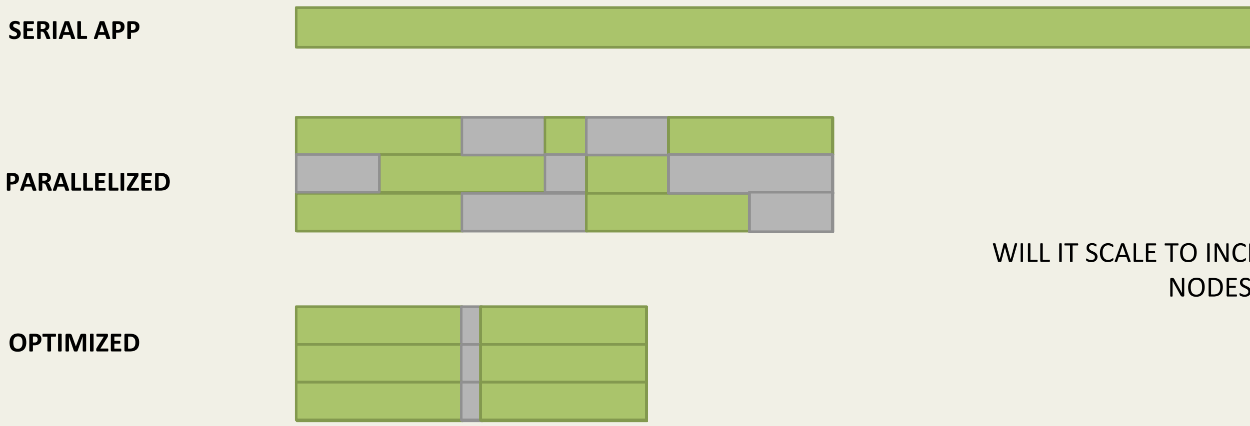
PROGRAMMING MODELS

SERIAL APP

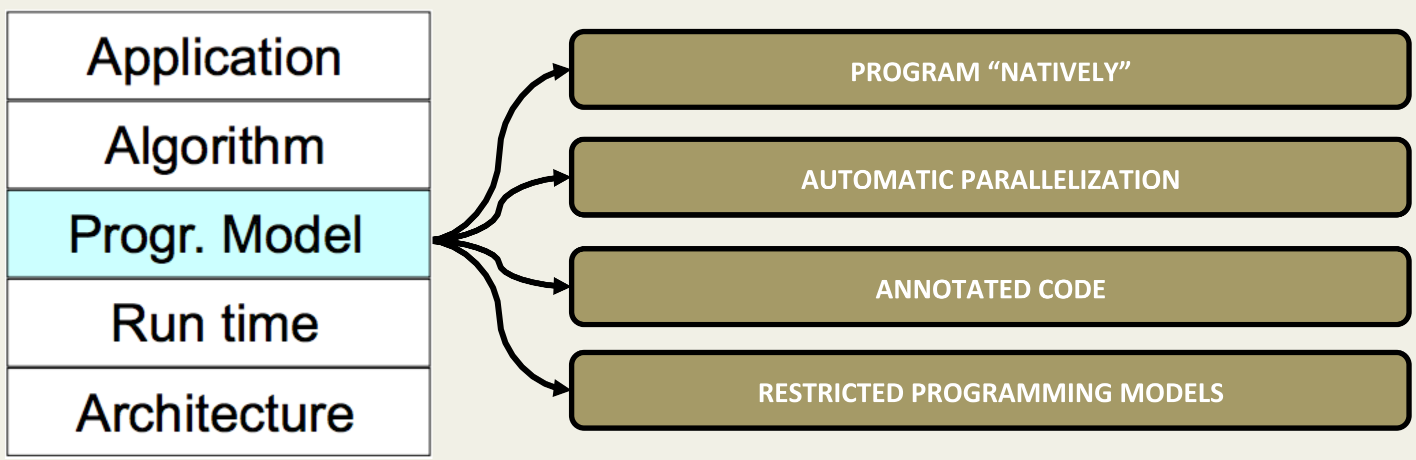
PARALLELIZED



PROGRAMMING MODELS



WILL IT SCALE TO INCREASING PARALLEL NODES ???



CHALLENGES AS OF TODAY

Managing ubiquity and heterogeneity: many computing-capable devices, infrastructures scattered, difficult to programme

Optimized parallel programming is hard, costly. Legacy applications, difficult to port

Interoperability.

Understand what architectures for what applications

Contribute to a culture of cost-benefit analysis → (is it worthwhile to double my computing infrastructure to increase my speedup from 9x to 13x?)

True horizontal scalability

Advances from both academy and industry → synergetic cultures

Power consumption footprint (green 500 top 3 nvidia nvidia xeon phi)

HOT ISSUES

DATA BANDWIDTH

Within chip

Within cluster

Take compute to data

Keep data close to compute

APP DEVELOPMENT

Within chip

Within cluster

Easy to use programming models

Realiable frameworks

MULTIDISCIPLINARITY

MULTIDISCIPLINARITY

The world of computing is flat, and anyone can do it. What will distinguish us from the rest of the world is **our ability to do it better and to exploit new architectures** we develop before those architectures become ubiquitous.

There is a clear and urgent need for a **new, modern approach to educating and training the next generation of researchers** in high performance computing specifically, and in modeling and simulation generally, for scientific discovery and engineering innovation.

Inadequate **education and training of the next generation of computational scientists** threatens global as well as U.S. growth of SBE&S […] unless we prepare researchers to develop and use the next generation of algorithms and computer architectures, we will not be able to exploit their game-changing capabilities.

There are clear and urgent opportunities for **industry-driven partnerships with universities** and national laboratories to hardwire scientific discovery and engineering innovation through SBE&S.

www.wtec.org/sbes 2009

GRACIAS