



Microsoft  
Research

# Multicore: Hey, Wait a Minute!




Dan Reed  
reed@microsoft.com

Scalable and Multicore Computing Strategist

## Presentation Outline

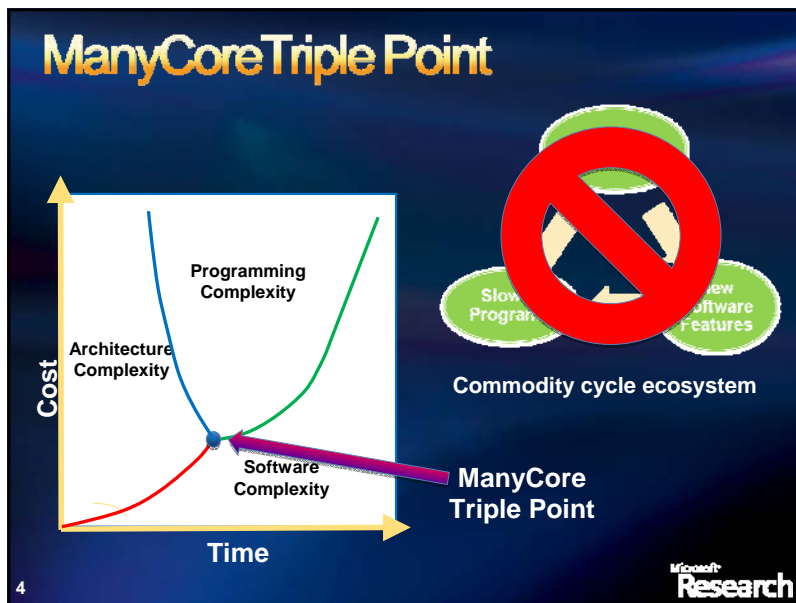
- Context and research
  - Sapir-Whorf and the future
- Multicore futures
  - Homogeneity/heterogeneity
- Programmability
  - Raising abstraction levels
- The future is ours*
  - To shape or misshape*



*Fortitudine Vincimus!*

3

Microsoft  
Research

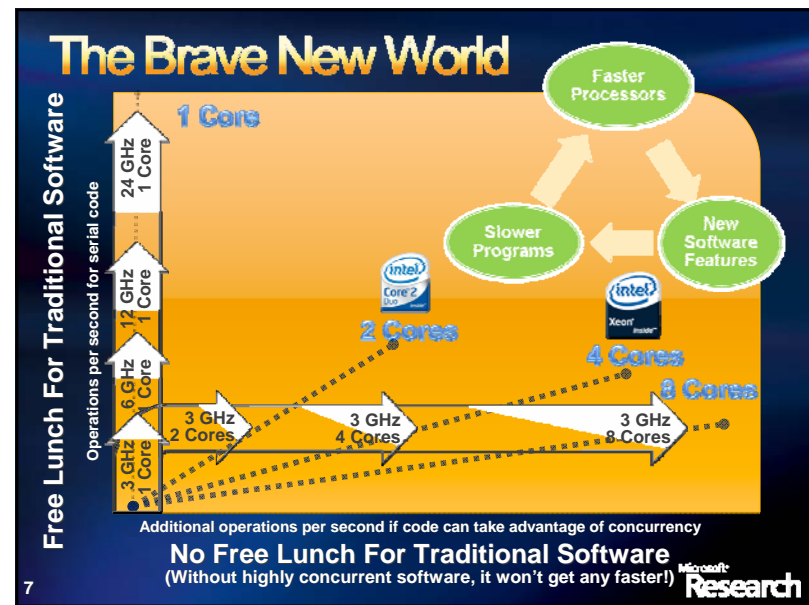
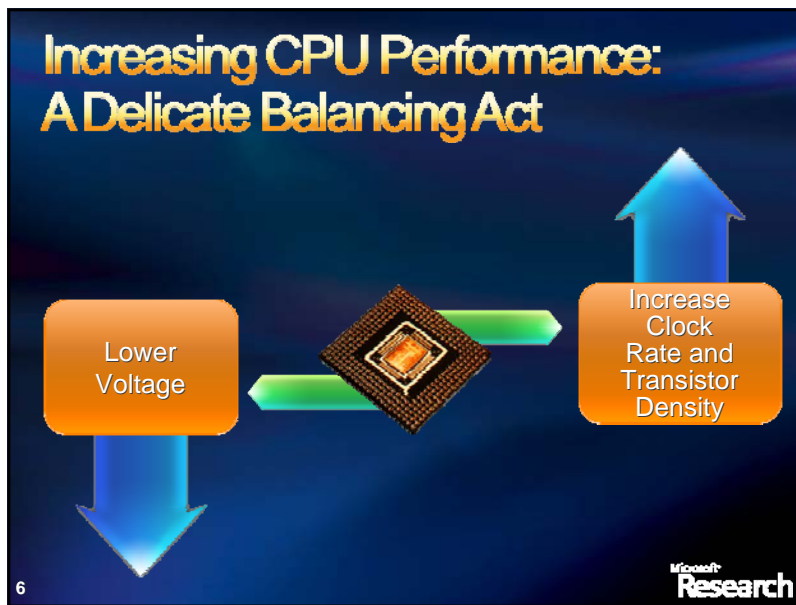


### Sapir-Whorf. Context and Research

- Sapir-Whorf Hypothesis (SWH)
  - language influences the habitual thought of its speakers
- Computing analog
  - available systems shape research agendas
- Consider some past examples
  - VAX 11/780 and UNIX
  - workstations and Ethernet
  - PCs and web
  - **inexpensive clusters**
    - **our HPC sweet spot**
- Today's examples
  - multicore, clouds and services ...

5

Microsoft Research

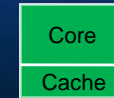


## We've All Followed This Path ...

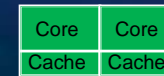
Given the clear and pressing need for improved computer performance, there are several means of achieving this end. *In the simplest approach, current computer architectures are reimplemented using faster device technologies.* Although this approach will always be exploited, physical, technological and economic limitations make it incapable of providing all the needed computational power. Instead, *parallelism must be exploited to obtain truly significant performance improvements.* November 1987



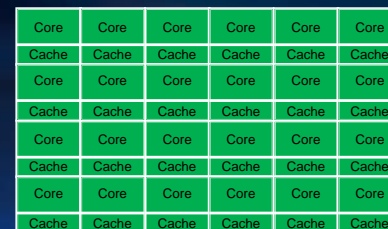
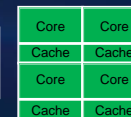
## One, Two, Three, Many? ...



Single Thread



Single/Multiple Thread Balance




Serious Parallelism and Optimization


Acknowledgment: Tim Mattson, Intel




## Think Chocolates, Not Cookies

- Sugar cookies
  - Similar, modulo process variation
  - You must eat lots to be satisfied
- Designer chocolates
  - Diversity is a feature
  - Forrest Gump was right




10 

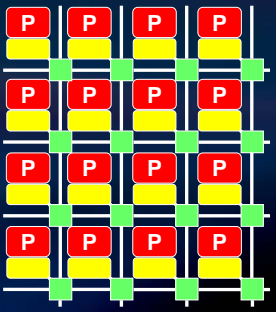
## Multicores Options




- On-chip possibilities – let's explore
  - PIM and mixed processes
  - Network protocol processing
  - Optical interconnect (MEMS or other)
  - Crypto-processing
  - DSP/image processing
  - Rendering
- On-chip/near-chip communications
  - Thinned wafers and chip stacking
  - Buses or bus networks
  - Shared memory
  - Point-to-point interconnect
  - Coherent/non-coherent caches

pose interesting challenges

- Power and structure



11 

## The Other Axis: Core Complexity

- Remember Amdahl's Law

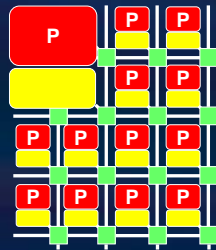
$$\text{Speedup} = (S + (1-S)/N)^{-1}$$

- Multicore implications?

- Symmetric or asymmetric cores
- Legacy and new code
- Programming heterogeneity

- Some very nice work by Mark Hill

- "Amdahl's Law in the Multicore Era," M. D. Hill and M. R. Marty, *IEEE Computer*, July 2008



G. Amdahl, "Validity of the Single Processor Approach to Achieving Large-Scale Computing Capabilities," *AFIPS Conference Proceedings*, pp. 483-485, 1967

12

Microsoft  
Research

## Architectural Futures

- Replication of tweaked cores
  - interconnect (it really matters)
  - mix of core types
  - heterogeneity and programmability

- Or, more radical ideas ...

- Other issues ...

- process variation and cores
- performance and reliability
- dynamic power management



Time domain (sec)	Mechanism	Delay impact (%)
$1 \times 10^{-11}$	Lithography node	20%
$1 \times 10^0$	Electromigration	8%
$1 \times 10^0$	Hot electron effect	5%
$1 \times 10^0$	NBTI	15%
$1 \times 10^0$	Chip electrical mean variation	18%
$1 \times 10^0$	Across-chip $V_{DD}$ variation	15%
$1 \times 10^0$	Self heating/temperature	12%
$1 \times 10^0$	SCI history effect	10%
$1 \times 10^{-10}$	Supply voltage	17%
$1 \times 10^{-10}$	Line-to-line coupling	10%
$1 \times 10^{-11}$	Residual S/D charge	6%

Source: Semiconductor International

- Where is our architectural vision?



- Where are the new ideas?

13


Microsoft  
Research




## BEE3: Architectural Experimentation

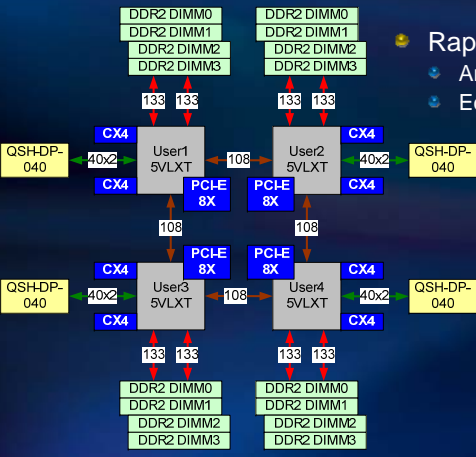
Chuck Thacker/John Davis  
MSR SVC




- Several highly-connected FPGAs
  - Multicore architecture research
  - Hardware/software co-evaluation

14 <http://research.microsoft.com/projects/BEE3> 

## BEE3 Architecture



- Rapid multicore prototyping
  - Architectural exploration
  - Ecosystem exploration

15 <http://research.microsoft.com/projects/BEE3> 





## The Siren Call ...

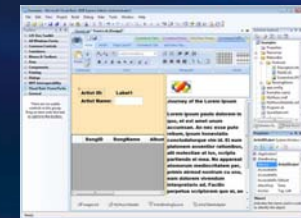


- We've seen parts of this movie before
  - Vector processors, systolic arrays, attached processors
- Success requires optimizing for efficiency
  - data movement, computation and software costs
- Efficient exploitation, in two senses
  - achieved application performance
    - holistic assessment, not just application kernels
  - high human productivity
    - extant software base, available tools
- **We must raise the abstraction level ...**
  - **Managed code and SLAs**
  - **Performance and failures**



## Programming Groups

- Three developer groups
  - Heroes
  - Mainstream
  - Entry/Novice
- Each with differing needs
- Heroes
  - "Neurosurgery?"
  - *No problem, hand me that screwdriver."*
- Mainstream
  - The typical computing graduate
- Entry/novice
  - Think of Visual Basic developers



## Breaking Through The Brick Wall

### Industry Giants Try to Break Computing's Dead End

NY Times March 19, 2008 John Markoff

Intel and Microsoft said Tuesday that they planned to finance two groups of university researchers to start over and design a new generation of computing systems intended to break the industry out of a technological cul-de-sac that threatens to end decades of performance increases in computers.

## Microsoft/Intel UPCRCs

- Two academic research centers
  - UC-Berkeley and UIUC
- Jointly funded by Intel and Microsoft
  - \$20M over five years
  - Matching funds from each institution
- Rationale
  - Long-term approaches to parallel computing
  - Integrated thinking – applications to architectures



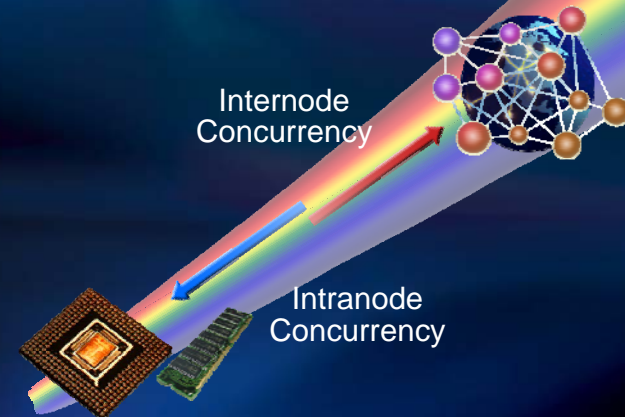
## Programming Heterogeneity

- Homogeneous parallelism remains challenging
  - We've been working on this for over forty years
- Heterogeneity poses additional challenges
- Coordinating heterogeneity
  - Multiple programs (ugh)
  - Automated mapping
  - Synthesized specifications
- **Raise the abstraction level**
  - *Coordination languages*
  - *Managed software*
  - *Specification not implementation*

22

Microsoft  
Research

## Concurrency Spectrum



23

Microsoft  
Research

## Experiences, Not Devices, Not Clients

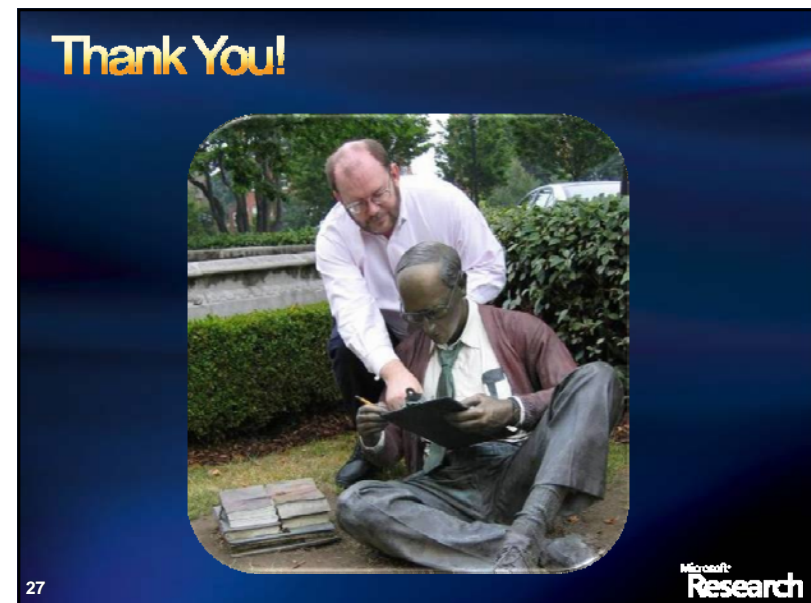
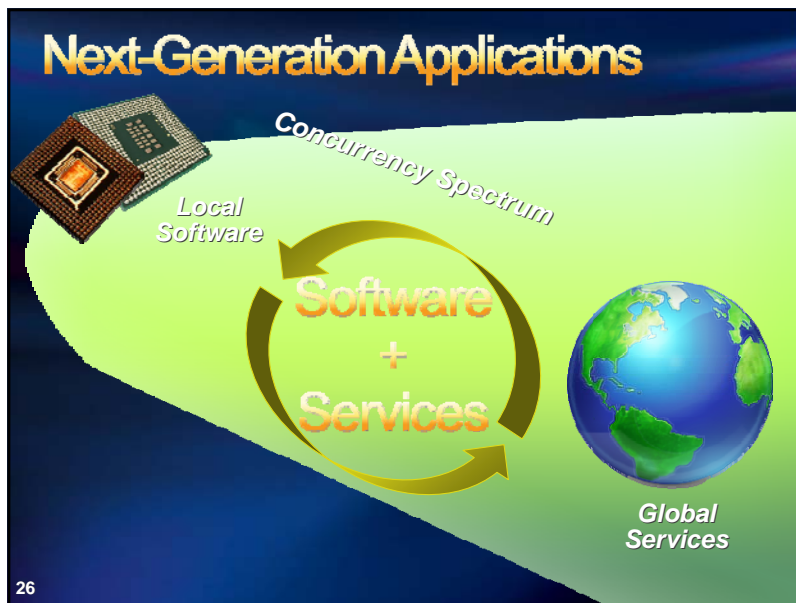
Clients + servers + sensors + actuators = experiences

24

## The Transductive Continuum

- Dynamic optimization
  - Power, computing and bandwidth
  - Code and transductive mobility
  - Edge devices and central facilities
- Context-aware information
  - The right information at the right time
    - Rich data management
- Peanuts Pigpen model
  - The cloud follows you *everywhere*

25



## Dan Reed (Microsoft)



© 2007 Microsoft Corporation. All rights reserved.

Microsoft, Windows, Windows Vista and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries. The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.