



In the future, most HPC processors
will be... ?

David Rich
Director Marketing, AMD
President, HyperTransport Consortium

Agenda

What I'm supposed to talk about;

- AMD quad core: Barcelona
- Accelerated Computing
 - Stream
 - Torrenza

What I want to talk about;

- Which way are we really going with “peta silicon”?

Disclaimer: After slide 6, it's just my opinion...

AMD Quad-Core Processor Architecture

A Closer Look at Barcelona...



Comprehensive Upgrades for SSE128

Can quadruple floating-point capabilities

Enhancements to AMD Virtualization™ technology (AMD-V™)

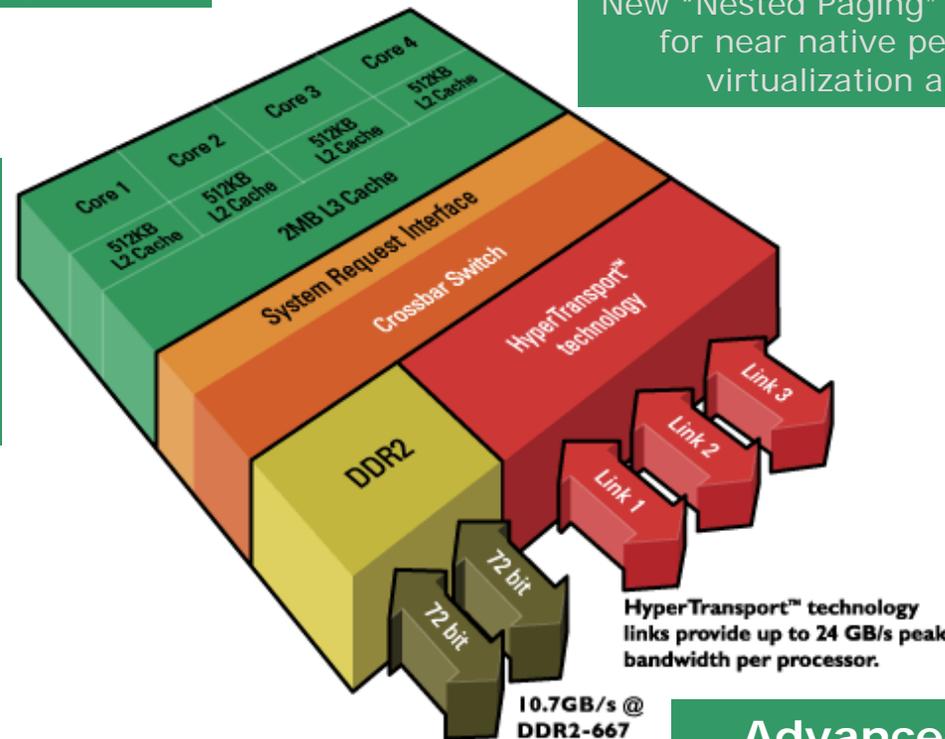
New "Nested Paging" feature designed for near native performance on virtualization applications

New Highly Efficient Cache Structure including a shared L3

Balance of dedicated and shared cache for optimum Quad-Core performance

Enhanced CPU Cores

Benefits applications by improving overall efficiency and performance of cores



HyperTransport™ technology links provide up to 24 GB/s peak bandwidth per processor.

10.7GB/s @
DDR2-667

**Quad-Core
AMD Opteron™
Processor Design** for Socket F (I207)

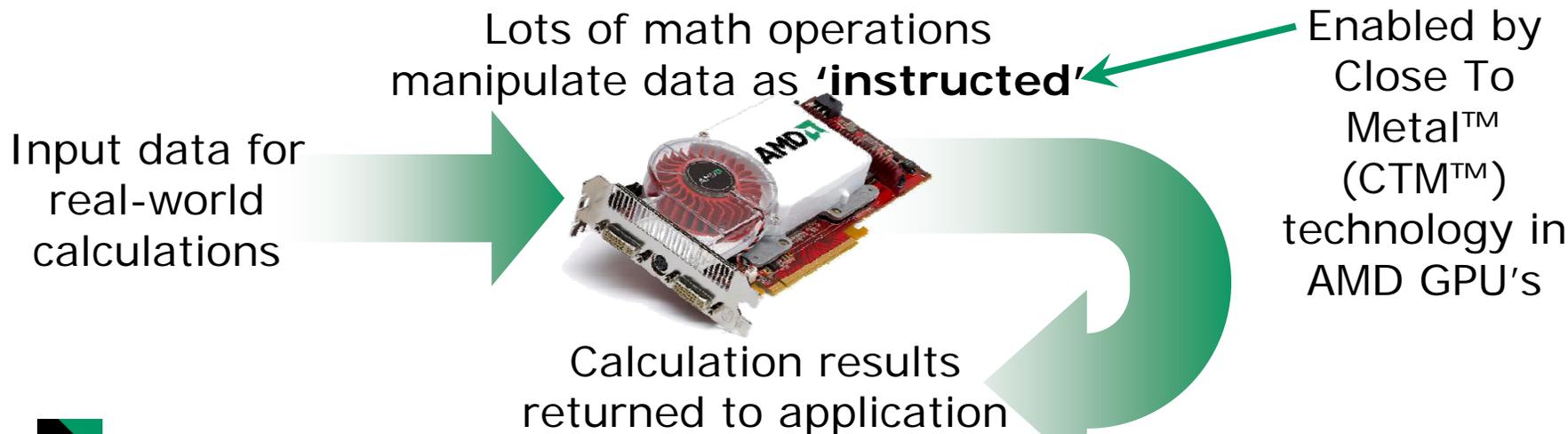
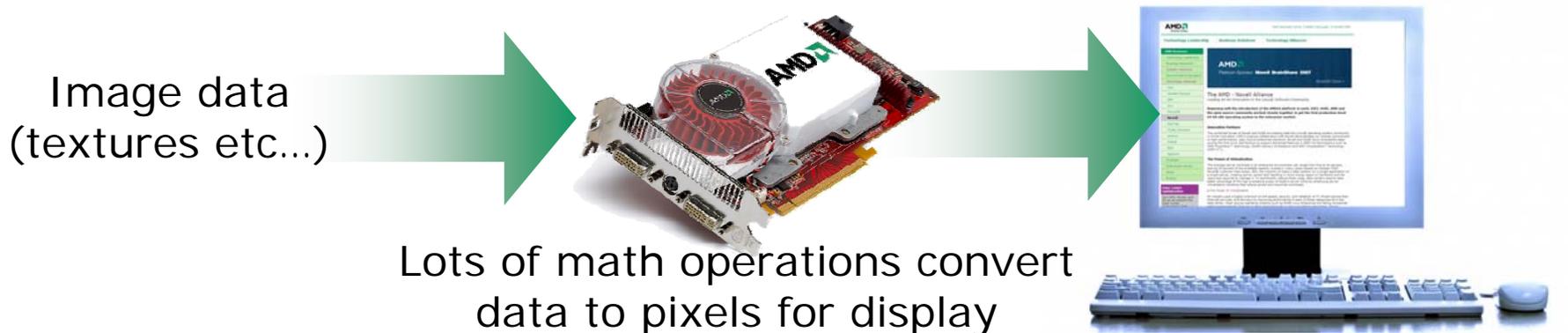
Advanced Power Management

Provides granular power management resulting in improved power efficiency

What is Stream Computing ... ?



Leverage of massively parallel processors, generally used for 3D graphics applications, to solve real-world problems.





Torrenza Today



- Direct Connect Accelerators in sockets or slots deliver superior performance without bridge chips
- 100s of GFlops to solve complex math

- Familiar programming interfaces speed time to implementation

Partners:

- Altera
- Celoxica
- DRC
- Xilinx
- XtremeData

- Stream
- ASICs
- FPGA

Compute Offload

- Application Libraries
- Compilers
- Hardware Interfaces

App Programming

Partners:

- CTM
- Celoxica
- OpenFPGA
- Peakstream
- Rapidmind

Torrenza

Partners:

- Bay Microsystems
- Commex
- NetLogic
- Qlogic
- RMI
- Tarari
- Woven

- IB
- XML
- iSCSI
- 10Gb E
- Search
- Storage
- Security

I/O

IT Infra

- Scale Up
- Virtualization

Partners:

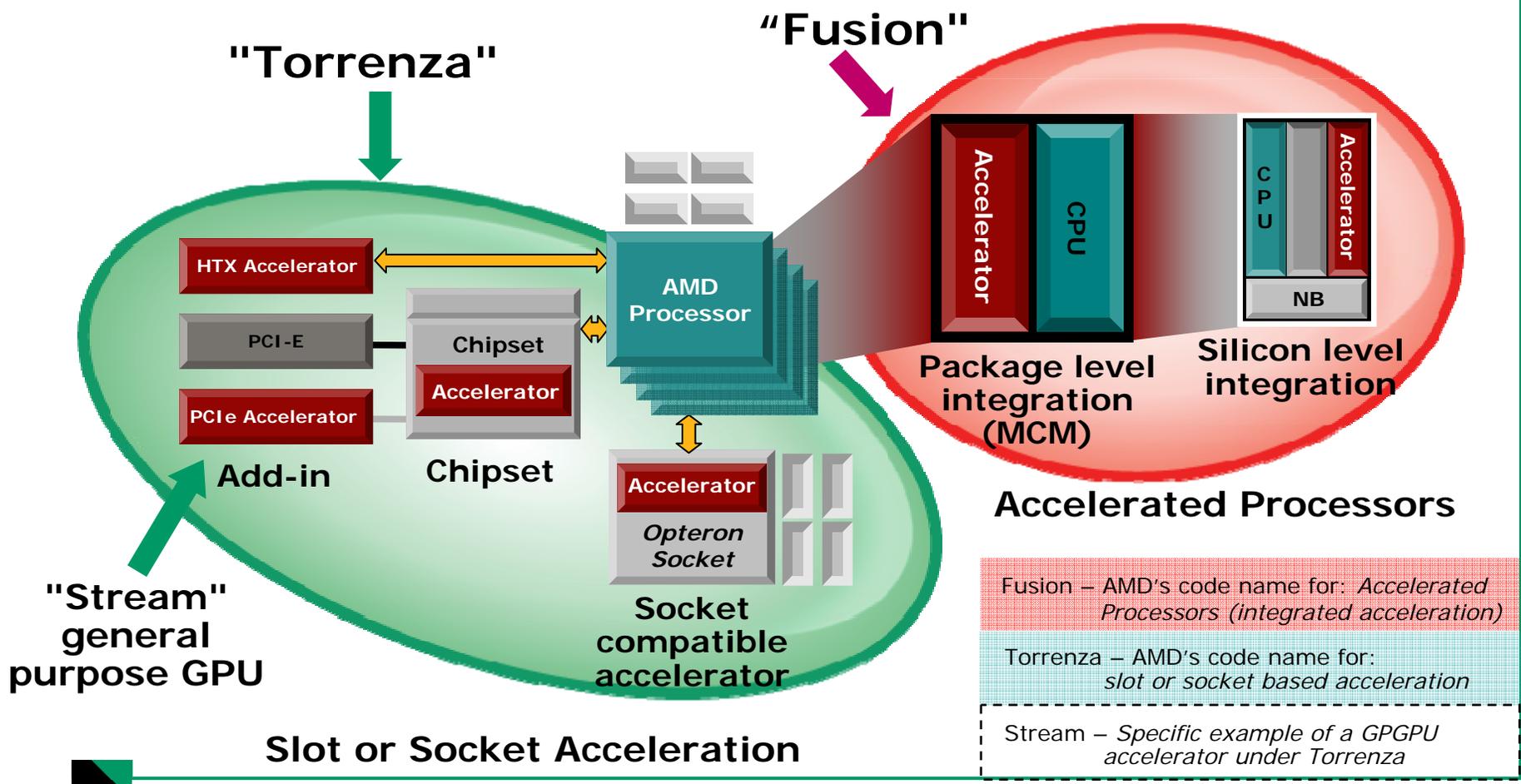
- 3Leaf Systems
- Liquid Computing
- Mannheim
- Panta Systems

- Specialized Direct Connect devices for high-throughput, low-latency processing

- cHT and HT provide peer level interfaces to build systems from commodity building blocks

Continuum of Solutions

Accelerated Computing



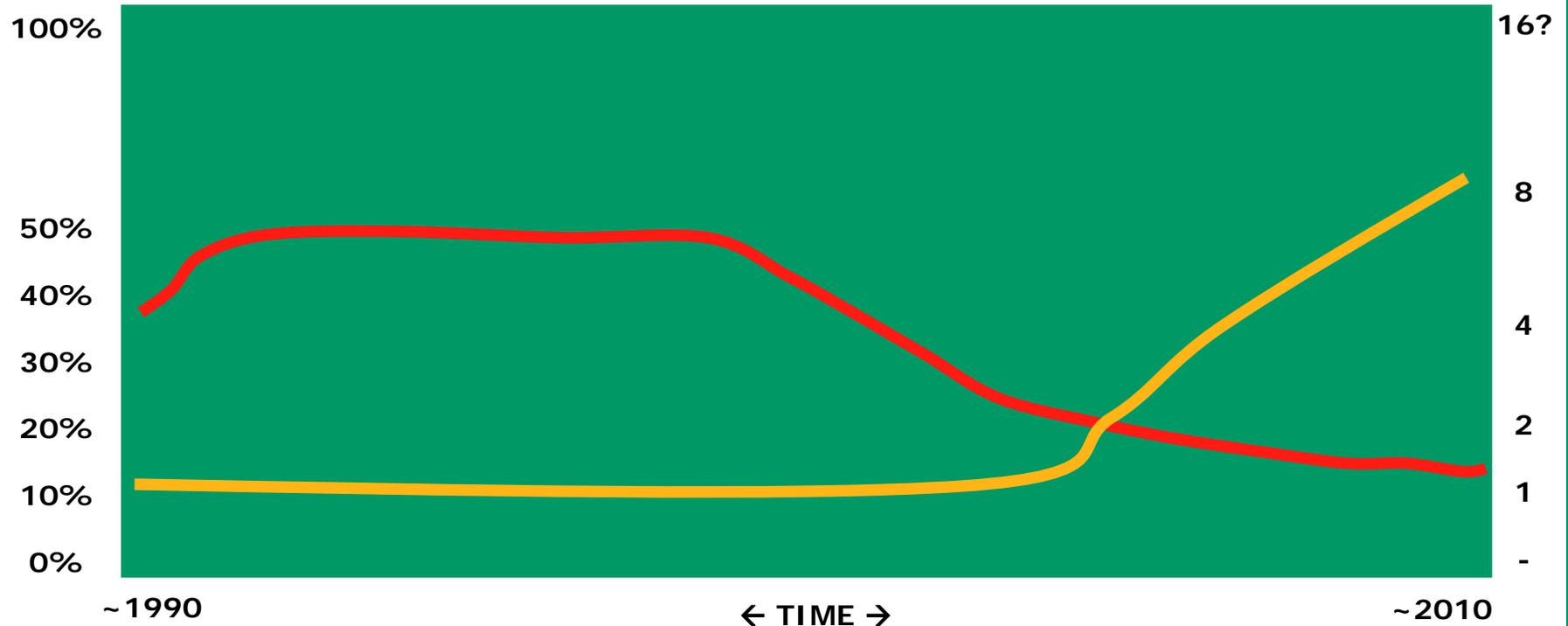
Slot or Socket Acceleration

Trends / What can you do with CMOS?



Rate of single thread performance growth

Number of processor cores in commodity x86

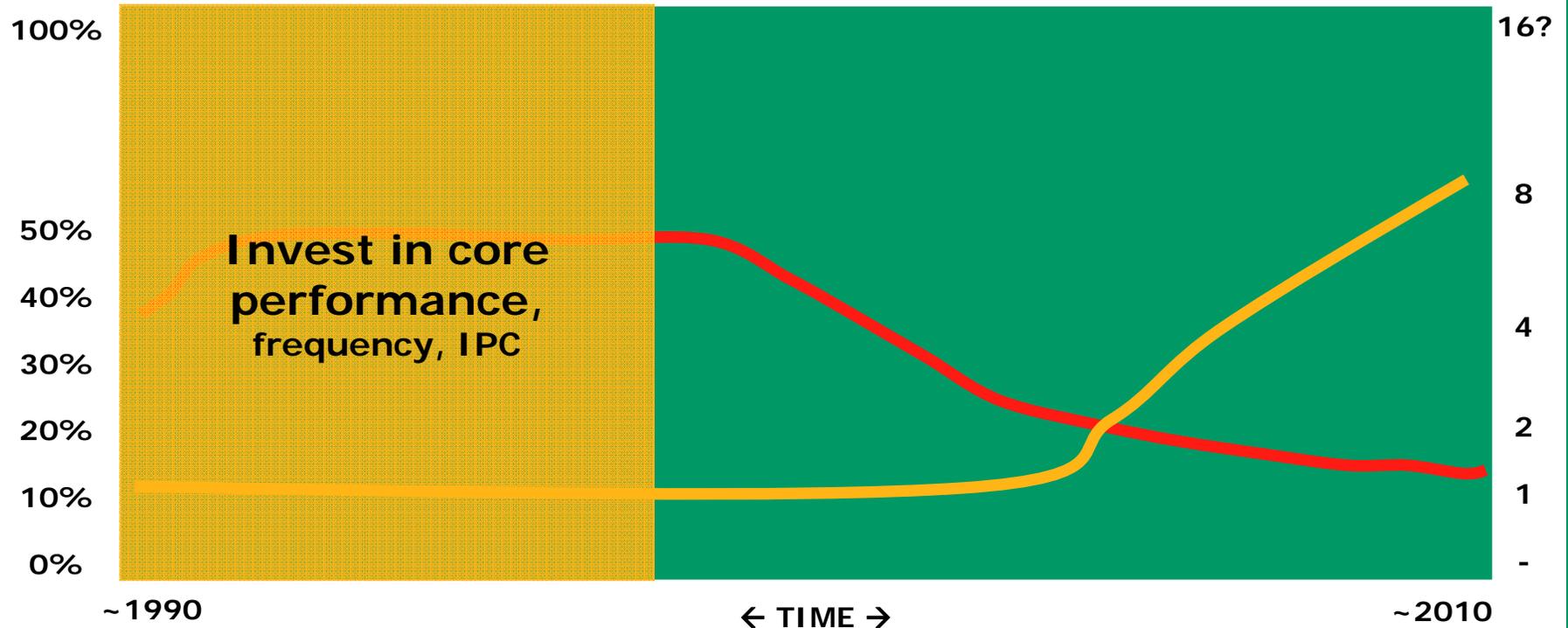


Graph is illustrative only

Trends / What can you do with CMOS?

Rate of single thread performance growth

Number of processor cores in commodity x86

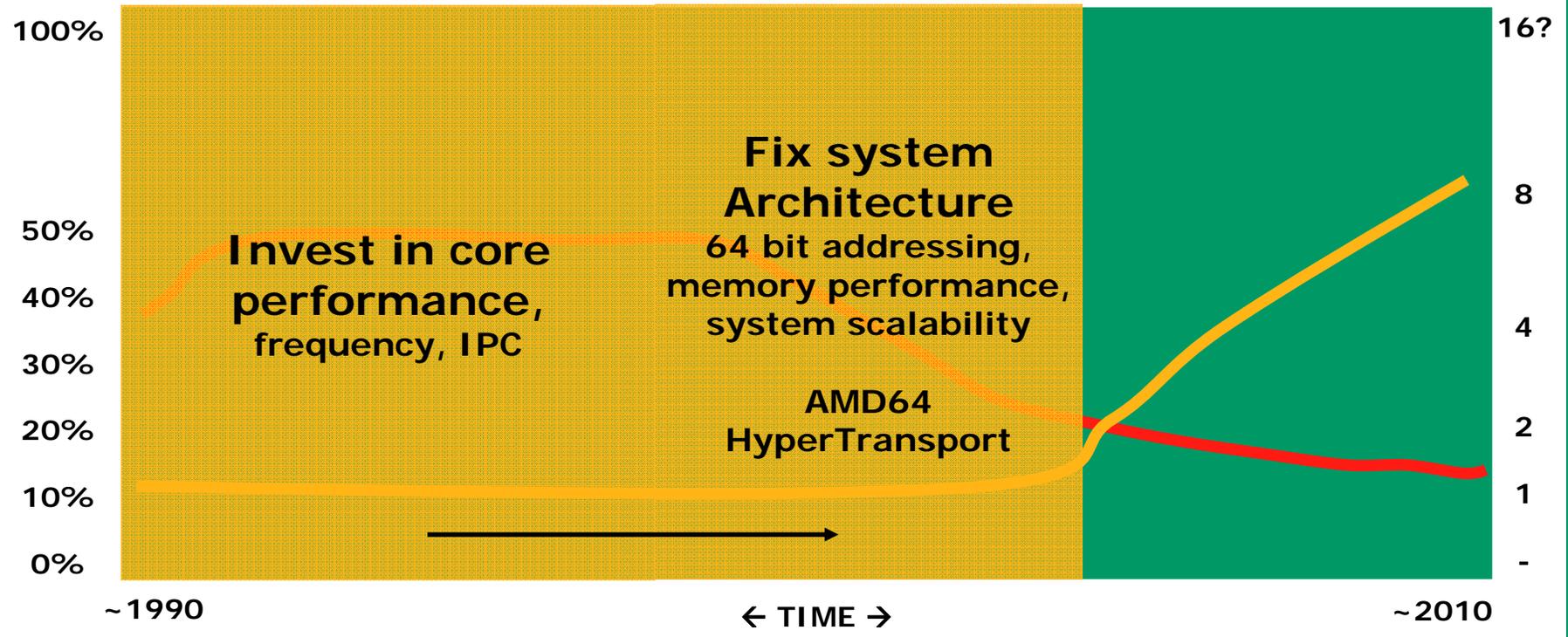


Graph is illustrative only

Trends / What can you do with CMOS?

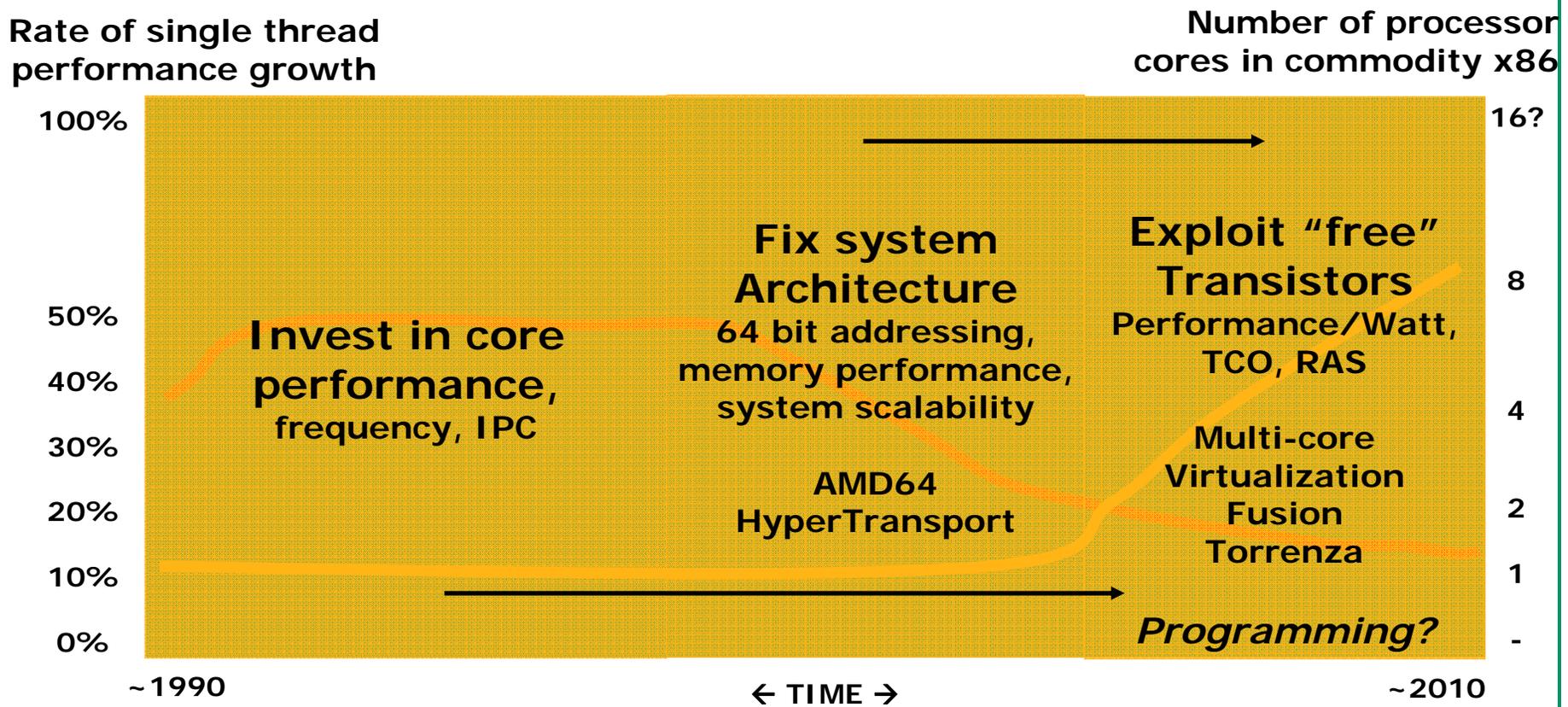
Rate of single thread performance growth

Number of processor cores in commodity x86



Graph is illustrative only

Trends / What can you do with CMOS?



Graph is illustrative only

One big question for AMD...

**Can one product satisfy both
commercial and scientific workloads?**

Today = “Sea change” or Blip?

Reduction in x86 single thread performance growth leaves an opening for new technologies

- FPGA, GPGPU, Accelerators (not taken seriously five years ago...)

x86 processor vendors talk about integration of new features; multi-core, “many-core”, Fusion, ...

Forces;

- Volume, software support
- Delivered application specific performance
 - Perhaps through abandonment of full custom design?
- Virtualization helps HPC
 - Because it makes memory performance matter to the rest of the market 😊

My PetaFLOP Machine is...

Limits; Cost, Power, Space, Time, Weight

Performance Attributes; Peak FLOPs, Total Memory, Memory per "SMP", Memory BW/(node, processor, core, thread), I/O Latency, Storage capacity, time to load dataset, MTBF...

Functionality; Runs x, y, z compiler, ISV app, libraries...

Ratios; Memory BW / FLOP, I/O per FLOP, "Performance"/Watt

Benchmarks; Standard, tender specific...

Too much data...!

Can we create groups to simplify?

For example;

- Embarrassingly Parallel “EP” Cluster
 - Hardware designed with cost in mind
 - Including power, operations...
- Super Cluster
 - Upgrades interconnect, likely more memory per node
 - Does it include accelerators?
 - GPGPU, FPGA, Cell, Clearspeed...
- Grand Challenge
 - System tuned to problem attributes
 - Likely not supported solely through unit sales

Note; could be purchased as small or large systems...

More Comments...

\$



<p>EP Clusters Highest Volume Lowest Cost</p> <p><i>HPC community has more to offer here...</i></p>	<p>Share's attributes with general purpose datacenter implementation Easy to justify feature investment Likely must be made from high volume product</p>
<p>Super Clusters Incremental improvements Volume justifies SW availability</p>	<p>Improvements are evaluated against market size and delivered benefit May show benefits with specific general purpose applications</p>
<p>Grand Challenge Contains novel technology Narrow applicability</p>	<p>Buyer and seller clear on "project" nature of the relationship Limited SW support likely AMD supports with open interfaces...</p>
<p>? Data Centric ?</p>	<p>What is a system with 1TB of memory per processor?</p>

Potentially Controversial Conclusion (mine, not AMD's)

Grand challenge machines will continue to have custom technologies solving critical problems...

But for the “mass HPC market” (some time in the future):

- Volume processor vendors will respond with:
 - Lots of flops in commodity products
 - Some way to fix memory bandwidth limits
 - A small number of critical new features (what are they?)
 - Relatively stable platform on which to build and tune software
- Software will adapt
 - Always has...
 - Efficiencies might not be great, but if cost/performance keeps going down, so what?
 - (and I used to be a software guy!)

Note: I can argue other outcomes, but only over drinks!

Disclaimer

Certain statements in this presentation, including but not limited to any projections regarding proliferation of AMD technology and the features and performance of such technology, may be considered “forward-looking.” Such forward-looking statements are based on current expectations, and, accordingly, entail various risks and uncertainties.

Risks that could cause actual results to differ materially from such forward-looking statements include potential software or hardware-related issues or conflicts, delays in the development of new products or technologies, problems with overall system or individual component performance, lack of production capacity, failure of solution providers to provide necessary support for AMD’s products, and other foreseen or unforeseen risks. We therefore cannot provide any assurance that such forward-looking statements will materialize. We assume no obligation to update or revise any forward-looking statement, whether as a result of new information, future events or any other reason. Additional information concerning risks and uncertainties affecting our business and other factors that could cause actual results to differ materially from any forward-looking statement is contained in our filings with Canadian and U.S. securities regulatory authorities, and are available at www.sedar.com and www.sec.gov.

Trademark Attribution

AMD, the AMD Arrow logo, AMD Athlon, AMD Opteron, AMD Virtualization, AMD-V, Close to Metal, CTM, ATI, the ATI logo, and combinations thereof are trademarks of Advanced Micro Devices, Inc. HyperTransport is a licensed trademark of the HyperTransport Technology Consortium. Microsoft and Windows are registered trademarks of Microsoft Corporation in the United States and/or other jurisdiction. All Novell marks referenced in this presentation are trademarks or registered trademarks of Novell, Inc. in the United States and other countries. Other names are for informational purposes only and may be trademarks of their respective owners.

©2007 Advanced Micro Devices, Inc. All rights reserved.

AMD

Smarter Choice