Terabyte Bandwidth Initiative
Architectural Considerations for Next-Generation Memory Systems

Craig Hampel
Rambus Fellow

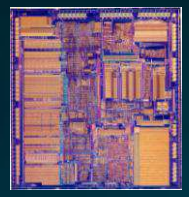
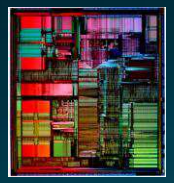
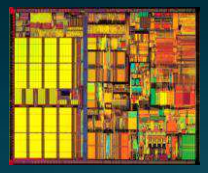
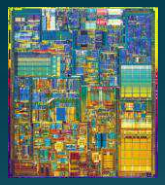
Rambus®

Outline

- Memory impact of processor and computing trends
- Throughput memory trends
- Terabyte Bandwidth Initiative summary

Important Trends in Processor Design

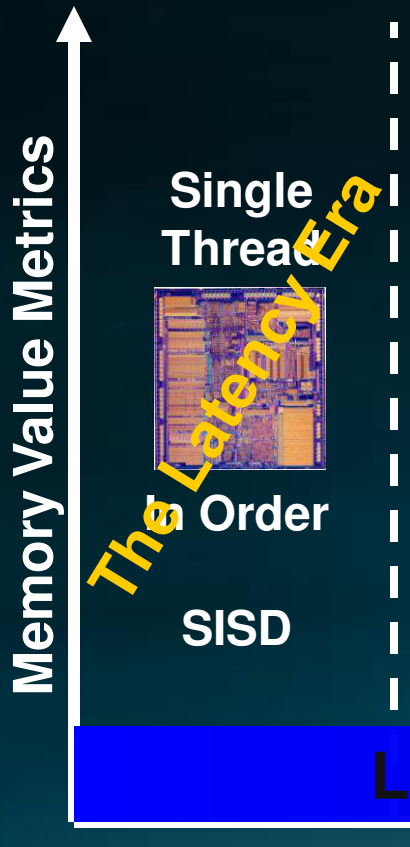
- Transistor densities continuing to increase

	386™ DX 1985	Pentium® 1994	Pentium® III 1999	Pentium® 4 with HT 2002	Pentium® 4 EE 2004
					
Transistors	275K	3.3M	9.5M	55M	178M
Clock Speed	33MHz	75MHz	600MHz	3GHz	3.46GHz
Power	2W	8W	20W	84W	111W

- Significant changes in processor architecture

How has this changed memory architecture?

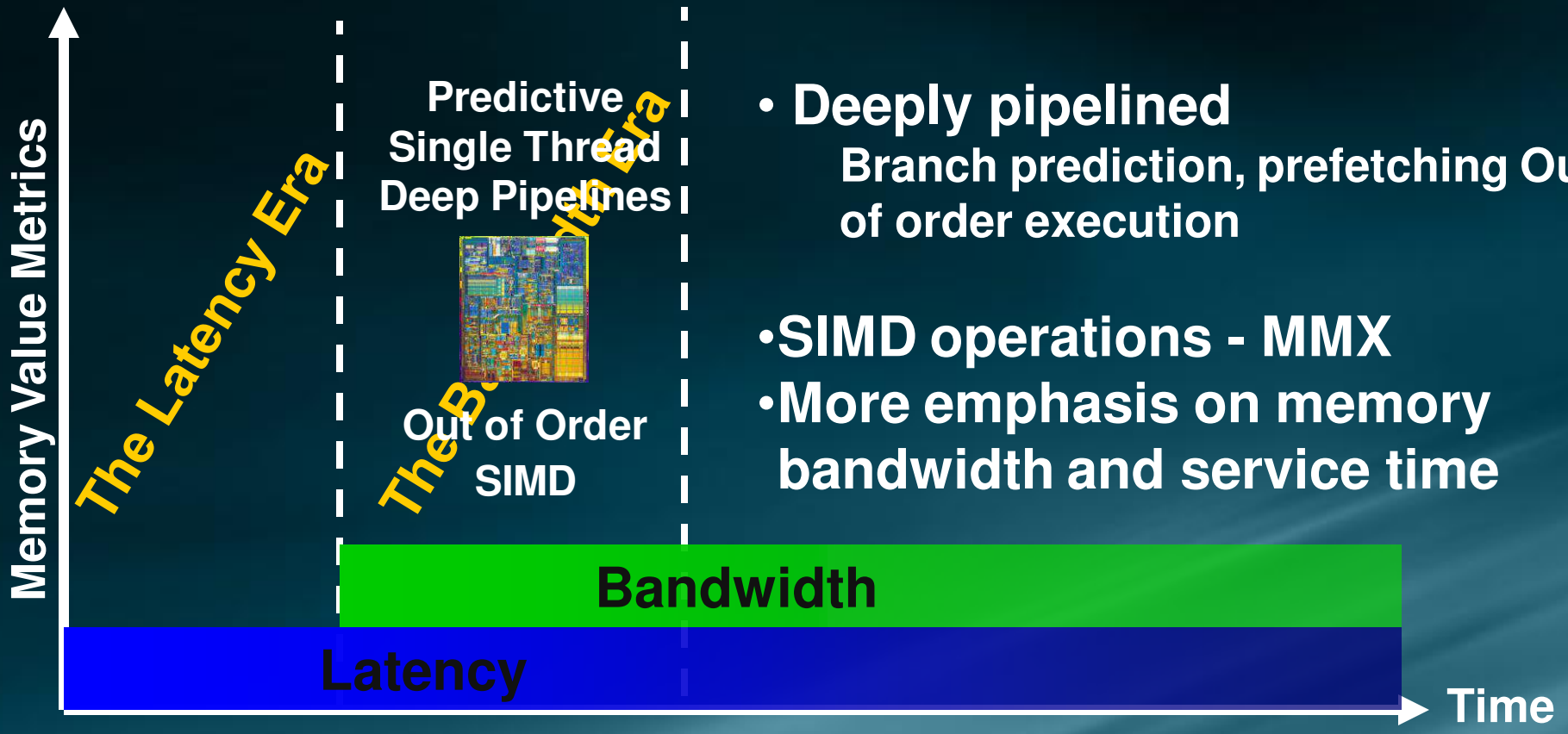
The Evolution of Memory Architecture



- Single in order execution thread
- Caches applied to reduce latency
- Most memory reads stall CPU execution

Memory architecture is heavily influenced by CPU architecture

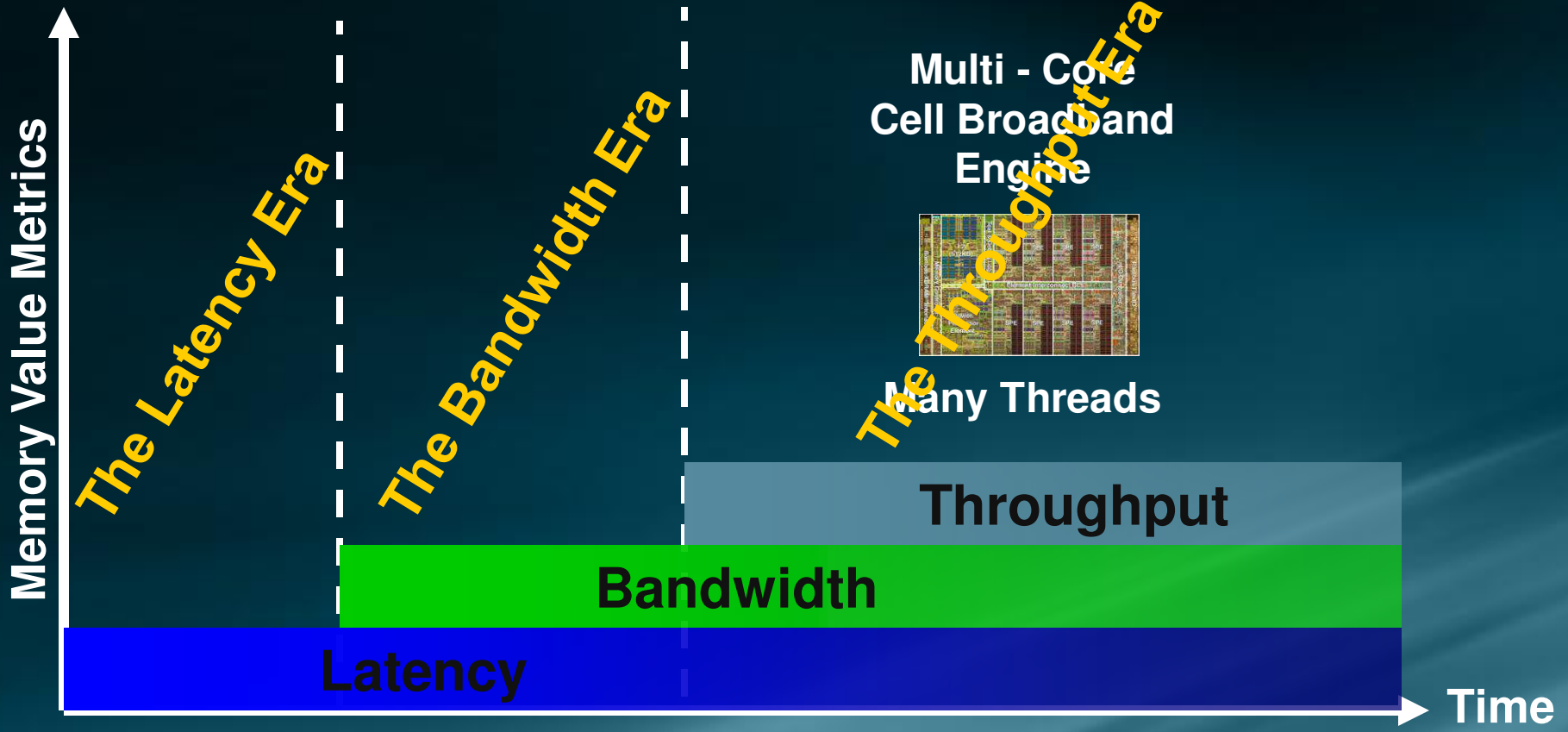
The Evolution of Memory Architecture



- **Deeply pipelined**
  - Branch prediction, prefetching Out of order execution
- **SIMD operations - MMX**
- **More emphasis on memory bandwidth and service time**

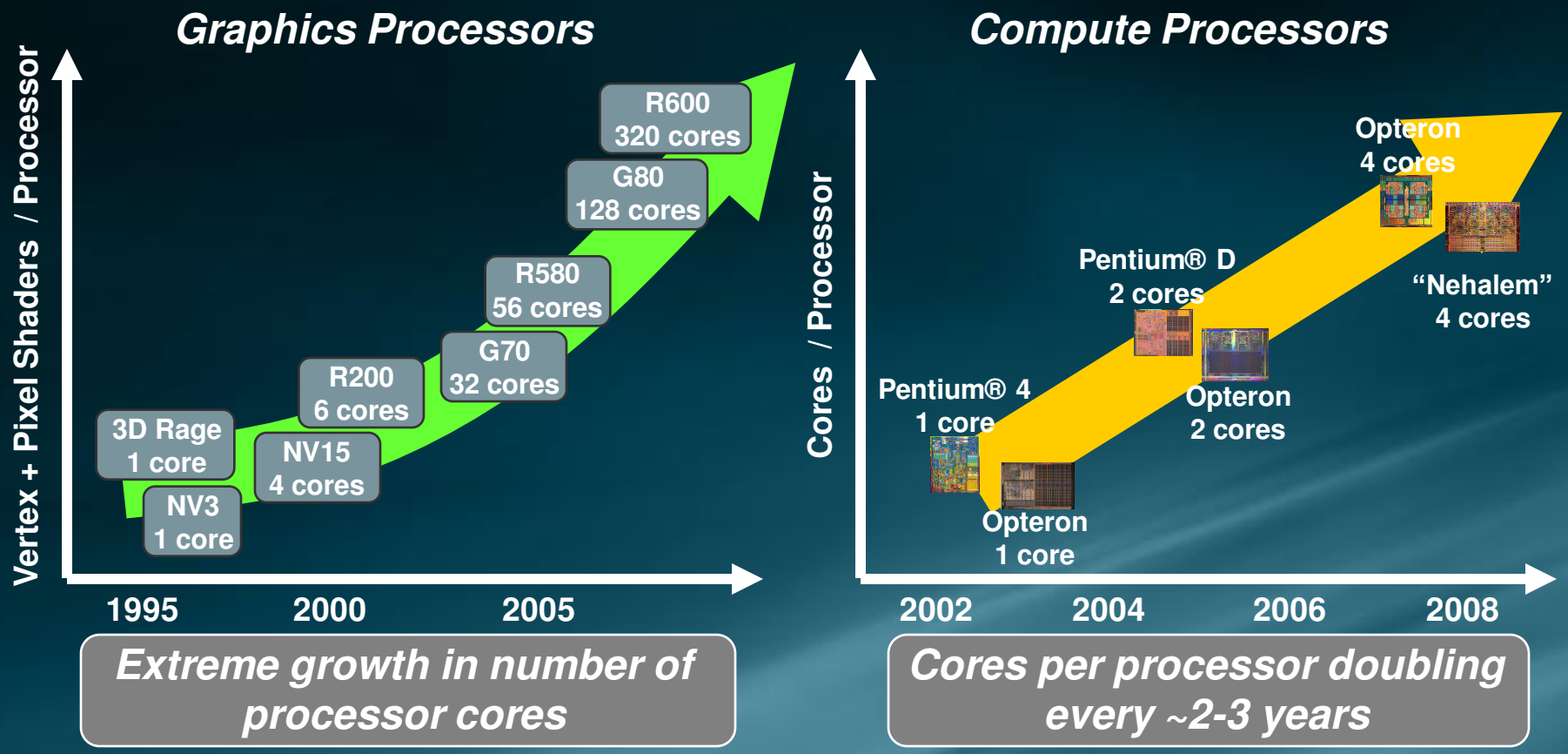
Memory architecture is heavily influenced by CPU architecture

The Evolution of Memory Architecture



Multi core processors motivate The Throughput Era

Dramatic Growth in Number of Graphics and Compute Processor Cores

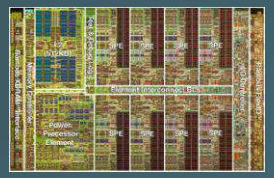


Rising core counts placing increasing demands on memory performance

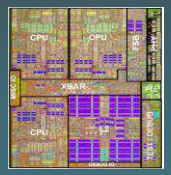
Multi-Core Architectures Adopted Across Computing Markets

Game Consoles

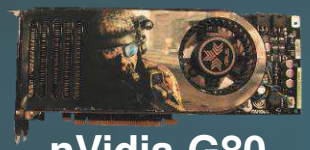
Cell Broadband Engine (9 cores)



Xbox 360 CPU (3 cores)



Graphics

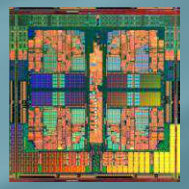


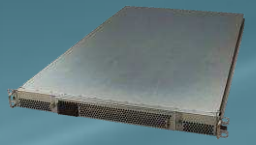
nVidia G80
128 Shader Processors

ATI R600
320 Stream Processing Units

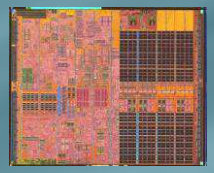
Computing

AMD Opteron (4 cores)



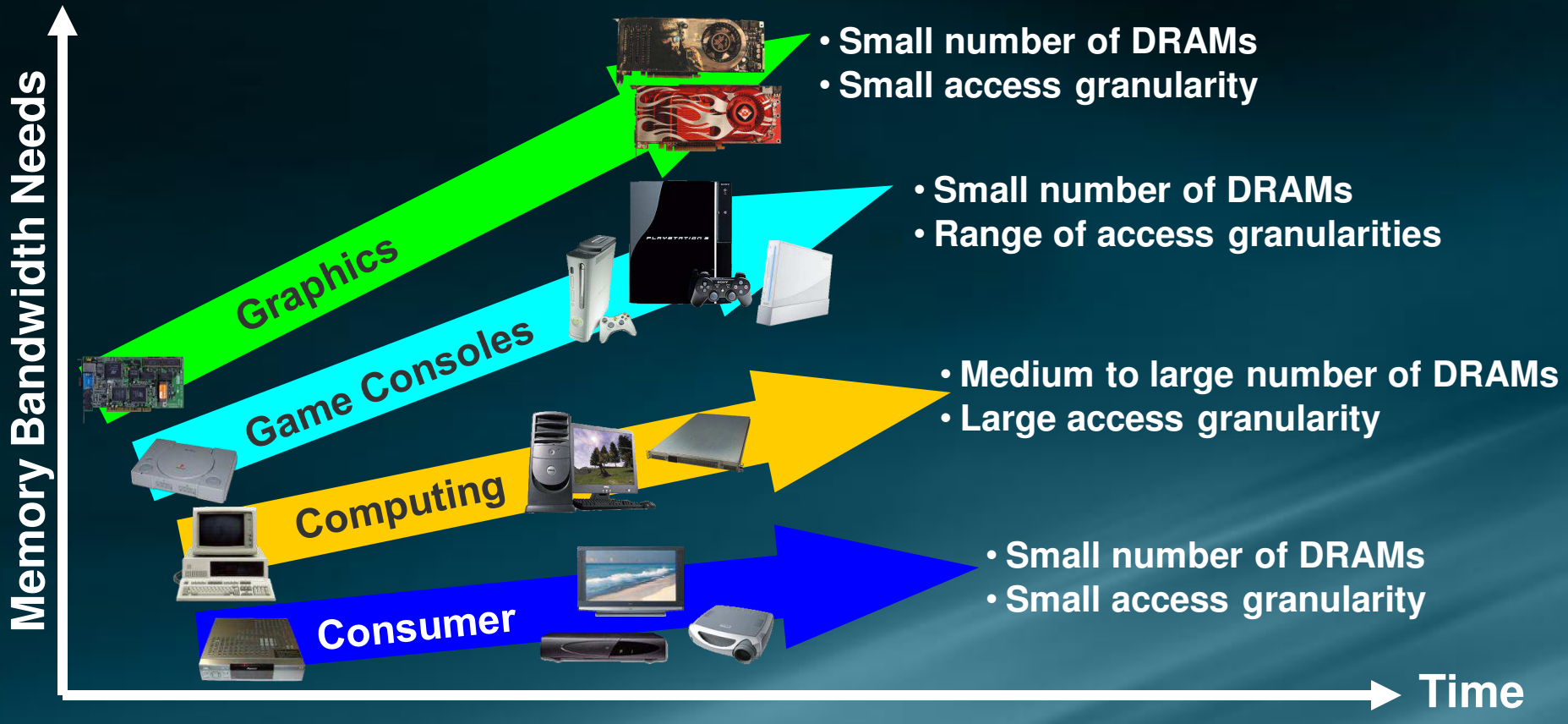


Intel Pentium D (2 cores)



Graphics and Game Consoles have greatest demands for throughput memory

Increasing Number of Cores Driving Memory Bandwidth Needs Higher



Bandwidth needs growing across markets, but differences in capacity and access granularity

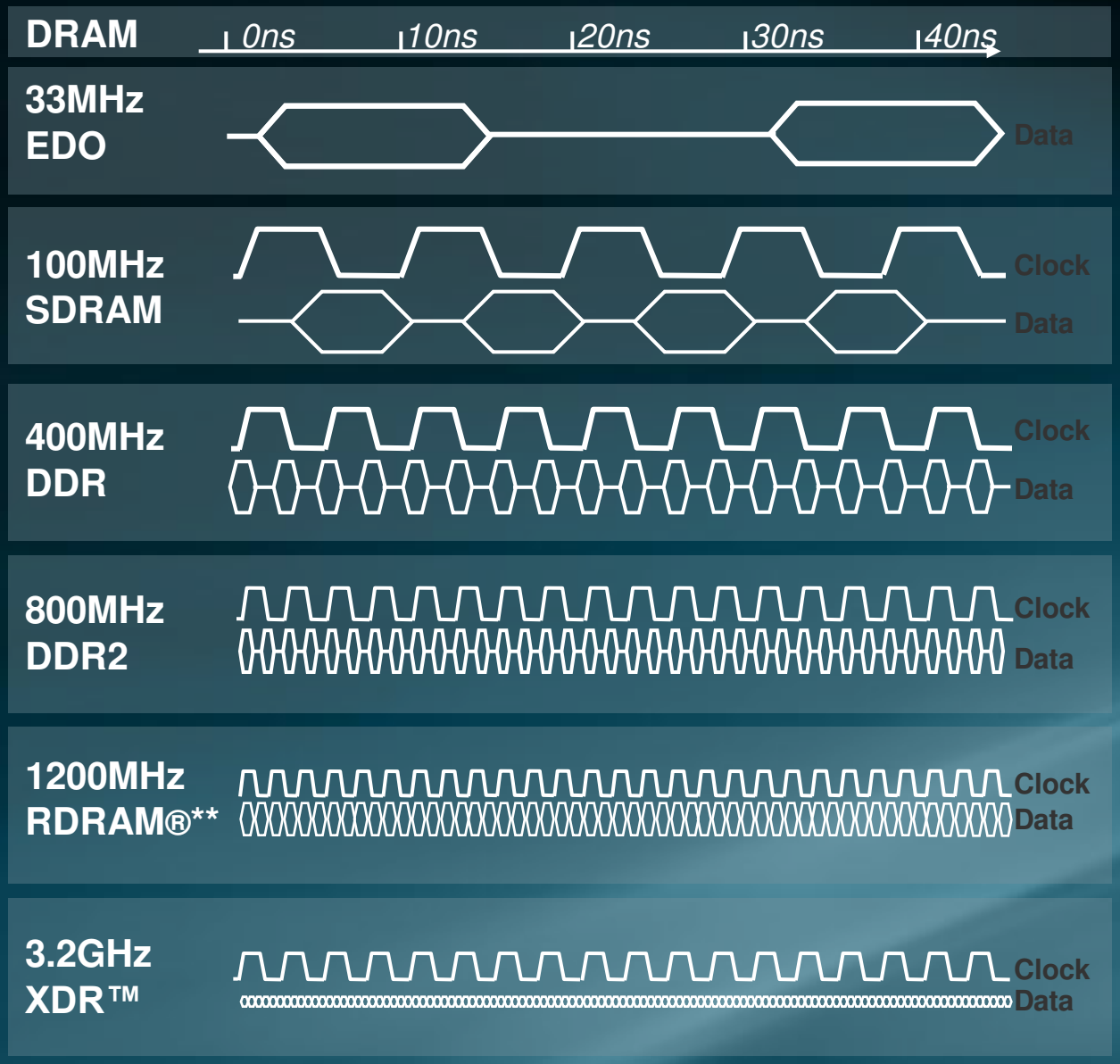
Outline

- Memory impact of processor and computing trends
- Throughput memory trends
- Terabyte Bandwidth Initiative summary

Multi-Core Throughput Processors Drive Demand for High Throughput Memory

- Balance CPU, memory, I/O systems to ensure good performance
- Multi-core processors pushing memory capacities and bandwidths higher
  - Memory capacity and bandwidth are key drivers for system performance
  - Cores need data to process
  - Efficiency of memory system becoming important
- Graphics, game console, compute, and consumer system bandwidth needs growing

Memory Data Rates Have Increased by 100X



- Synchronous Clocking
- Programmable Read Latency
- Programmable Write Latency
- Variable Block Size

- DLL or PLL on a DRAM
- Source Synchronous Strobe
- Dual Edge Clocking
- Core Prefetch

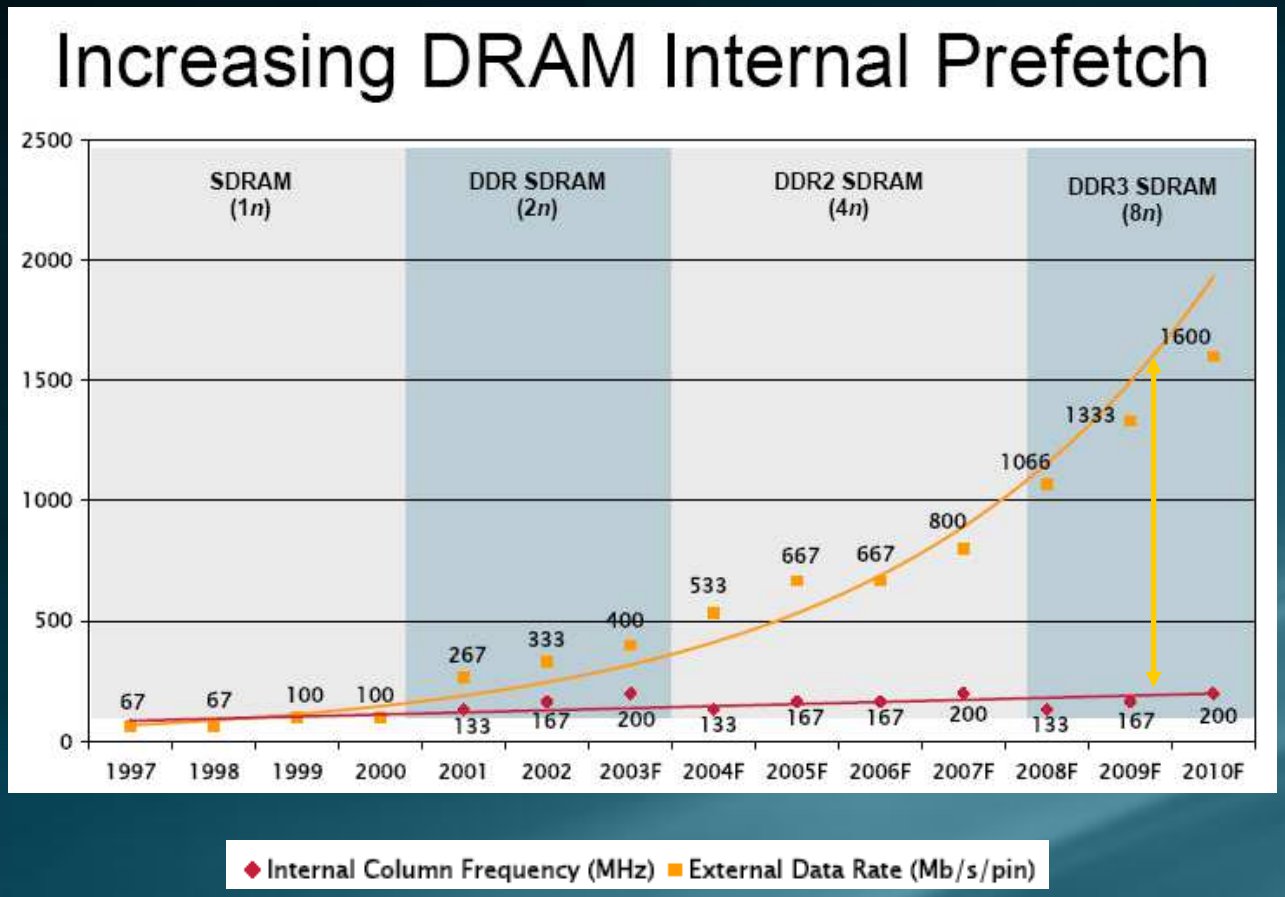
- On-Die Termination
- Calibrated Drive Strengths

- Clock Duty Cycle Correction/ Jitter Reduction

- Fly-By Command/Address
- FlexPhase Timing Deskew
- Differential Signaling
- Dynamic Point-to-point

* Direct RDRAM Memory
* This list is not exhaustive

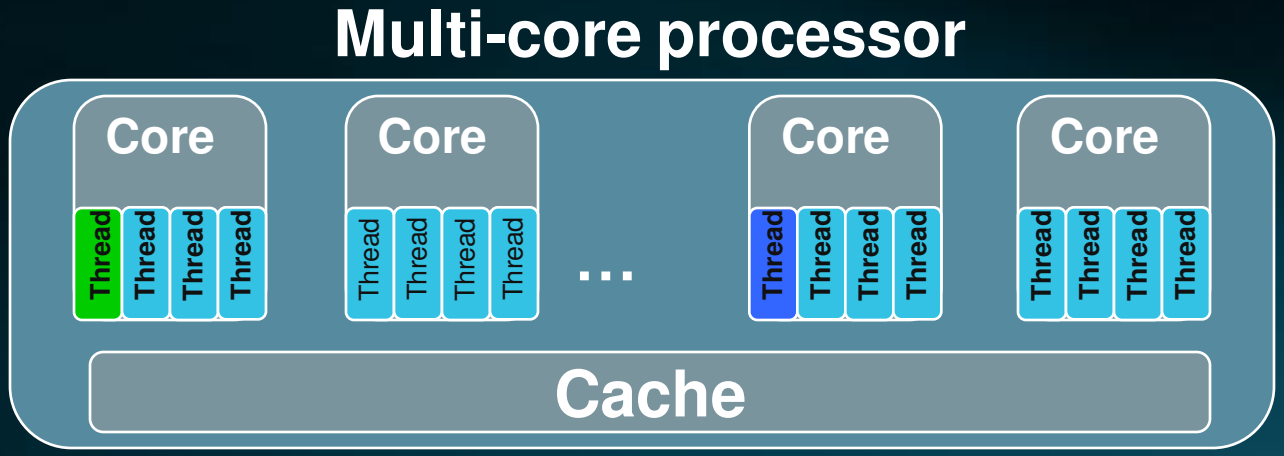
DRAM Column Access Trend



- Column prefetch is utilized to increase bandwidth without increasing column frequency
- Data width multiplied by prefetch determines column access granularity

Source: Micron

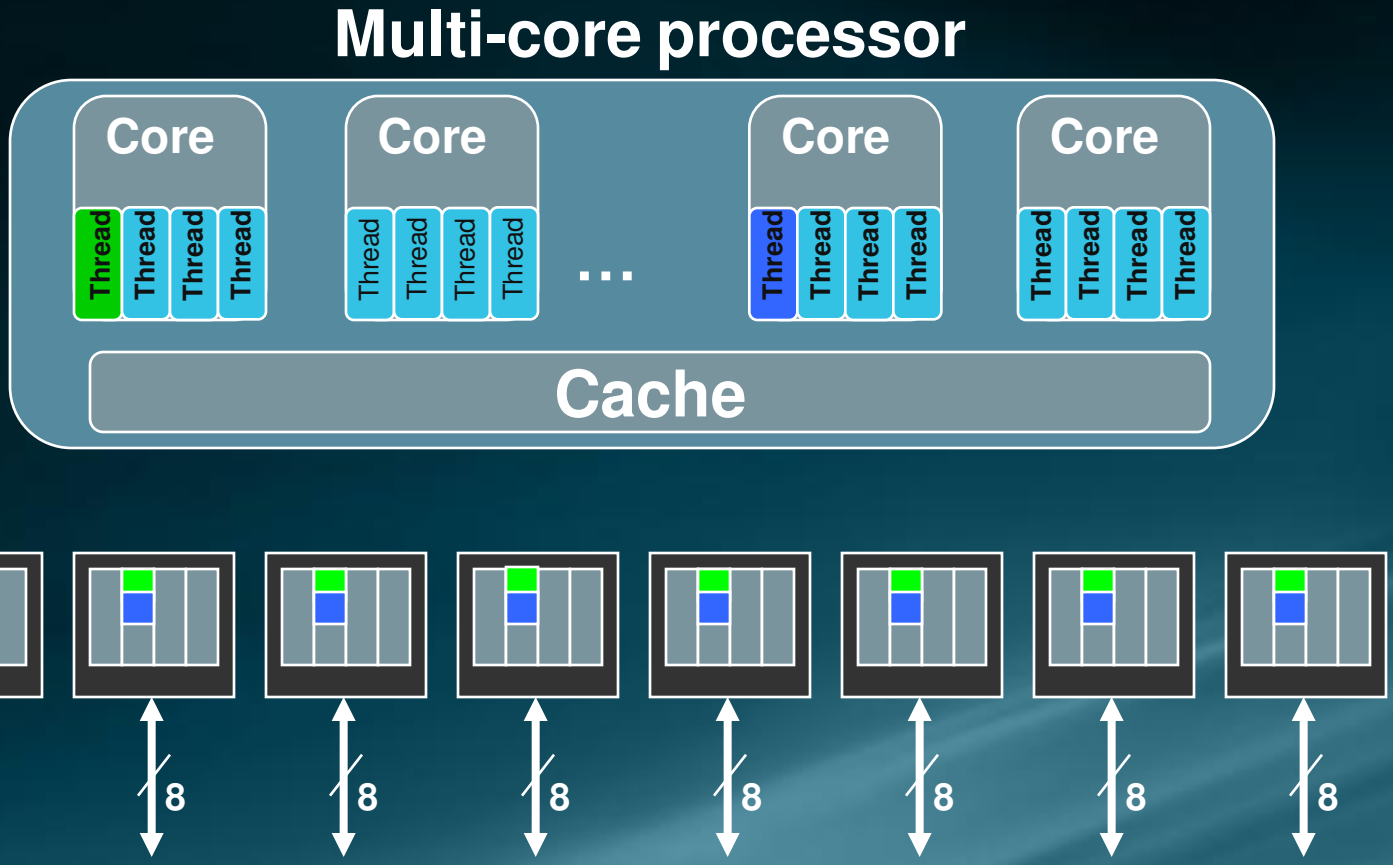
DRAMS as a Parallel Resource



- Architectural philosophy of thread level parallel
  - Maintain parallel resources whenever possible
  - Avoid unnecessary synchronization/arbitration
  - Optimize for throughput
  - Optimize for efficiency - power

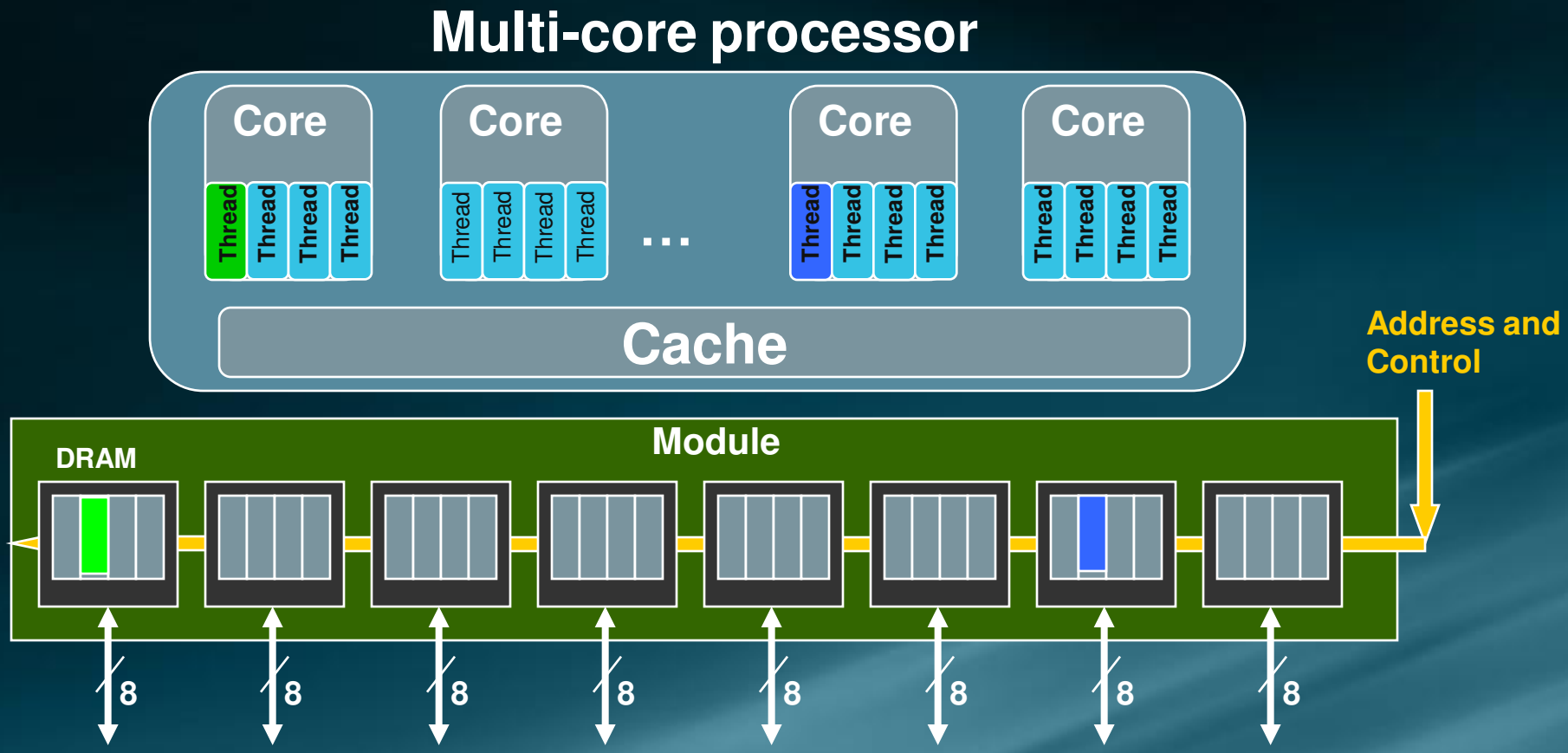
What if we extend these priorities to main memory?

Traditional Mapping of DRAM



Does this make sense for modern processors?

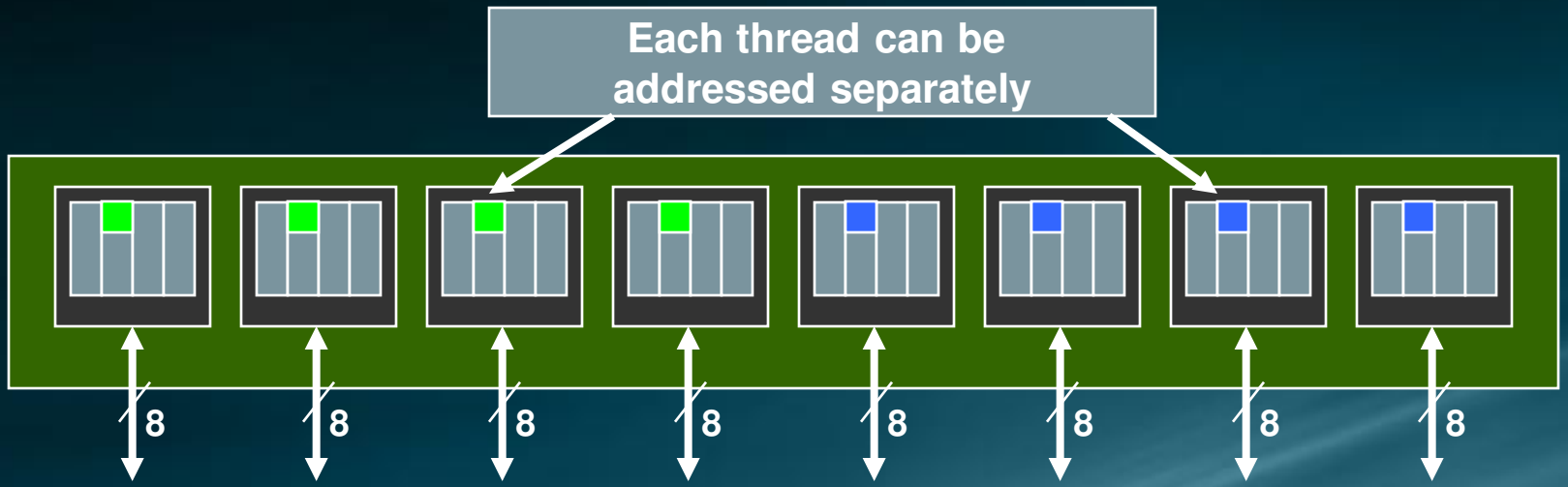
Alternative Mapping of DRAM as a Parallel Resource



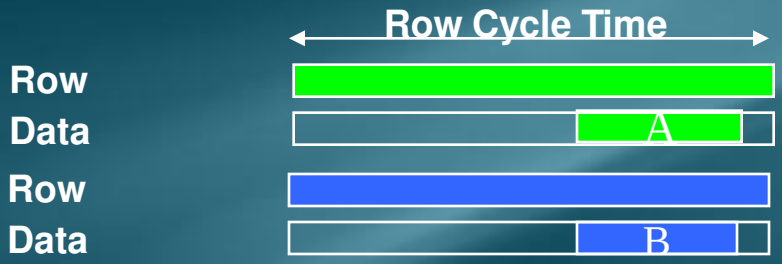
Can we maintain more parallelism with current DRAMs and modules? … Module Threading

Module Threading

- Each Rank can support two (or more) independent threads
  - Control can be time multiplexed with CS



- Effectively doubles the number of banks and decreases bank utilization by 50%



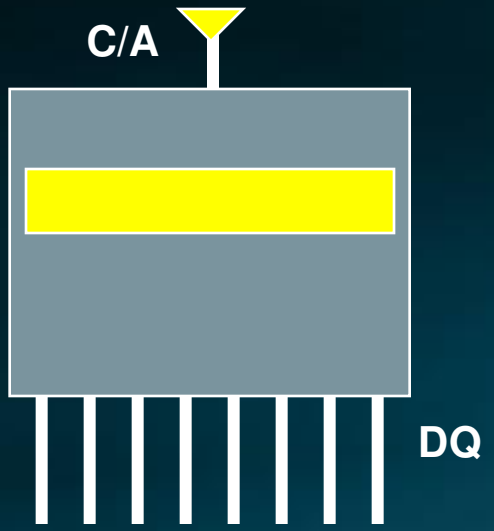
Example Benefits of Module-threading

- DDR3 without module threading (8n pre-fetch)
  - Row granularity 128 Bytes/module
  - Column granularity 64 Bytes/module
- Threading allows module to operate at:
  - Lower power per transaction (½ Row power)
  - Finer access granularity
  - Higher performance

Threading Factor	128 Byte transfer - BW efficiency	Power per Transaction
Single	54%	1.000
Dual	75%	0.750 – 0.875

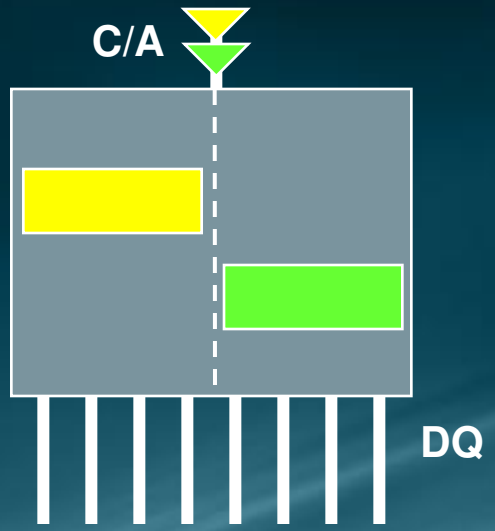
Micro-Threading in High Bandwidth Memory Systems

Traditional DRAM Read Access



- Initiate one transaction at a time
- Return one data packet

Micro-Threaded DRAM Read Accesses



- Initiate two transactions at a time
- Return two data packets, each over different DQ resource
- Access granularity cut in half, high bandwidth achieved

Multi-Core Throughput Processors Drive Demand for High Throughput Memory

- Processor architecture drives memory architecture
  - The Latency Era
  - The Bandwidth Era
  - The Throughput Era
- Multi-core processors pushing memory capacities and throughput higher
  - Memory capacity and bandwidth are key drivers for system performance
  - Cores need data to process
  - Efficiency of memory system becoming important
- Rambus Module and DRAM threading increase DRAM concurrency to improve throughput

Outline

- Memory impact of processor and computing trends
- Throughput memory trends

Terabyte Bandwidth Initiative Summary

- Example 1TB/sec memory system
- *Terabyte Bandwidth Initiative* technology elements
- Test chip and technology status

Rambus Terabyte Bandwidth Initiative

An *initiative* focused on memory signaling technologies for next-generation Terabyte bandwidth memory systems

- Goal: Deliver *1 TB/s memory bandwidth* to a single System-on-Chip with *16Gbps data rates per data link*
- Suitable for low-cost, high volume, high performance system manufacturing
- Ideal for next-generation multi-core, gaming, and graphics applications
- *Technology initiative*, not a product announcement

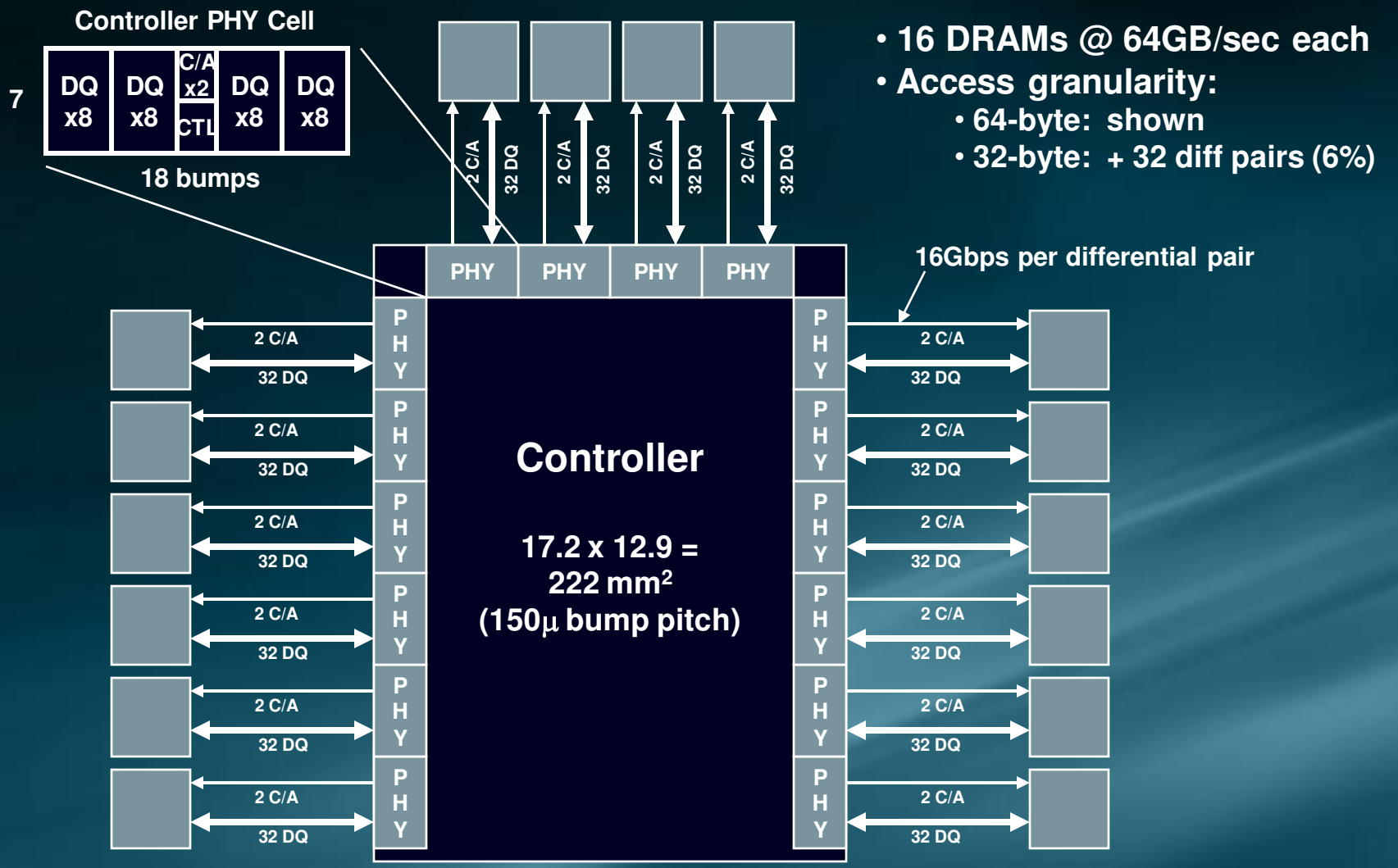
Key Challenges for 1TB/sec+ Systems

- **Maximize link bandwidth**
  - 16 Gbps per differential pair and beyond
- **Maximize link density at controller**
  - Leverage packaging trends:

Parameter	Trend
Flip chip bump pitch	180µ going down to 150µ
BGA ball pitch	1.0mm going down to 0.8mm
Package body size	42.5mm and going up
Substrate thickness	400µ-800µ going down to 200µ

- **Efficiently utilize available links**
  - Per-byte masks, strobes, clocks, DBI, etc. consume valuable pins
  - Use narrow command/address channels to reduce overhead
- **Improve power efficiency**

Example 1TB/sec Memory System



- 16 DRAMs @ 64GB/sec each
- Access granularity:
  - 64-byte: shown
  - 32-byte: + 32 diff pairs (6%)

Terabyte Bandwidth Initiative Highlighted Innovations

Fully Differential Memory Architecture (FDMA)

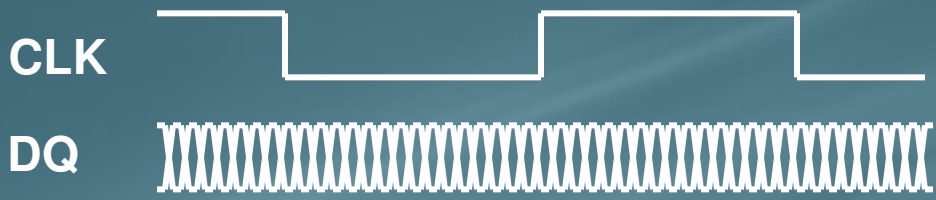
- Differential signaling for Command/Address and DQ links
- Enables 16Gbps signaling rates

FlexLink™ C/A

- Full-speed, flexible, scalable point-to-point Command/Address link
- Enables different access granularities in different applications

32X Data Rate

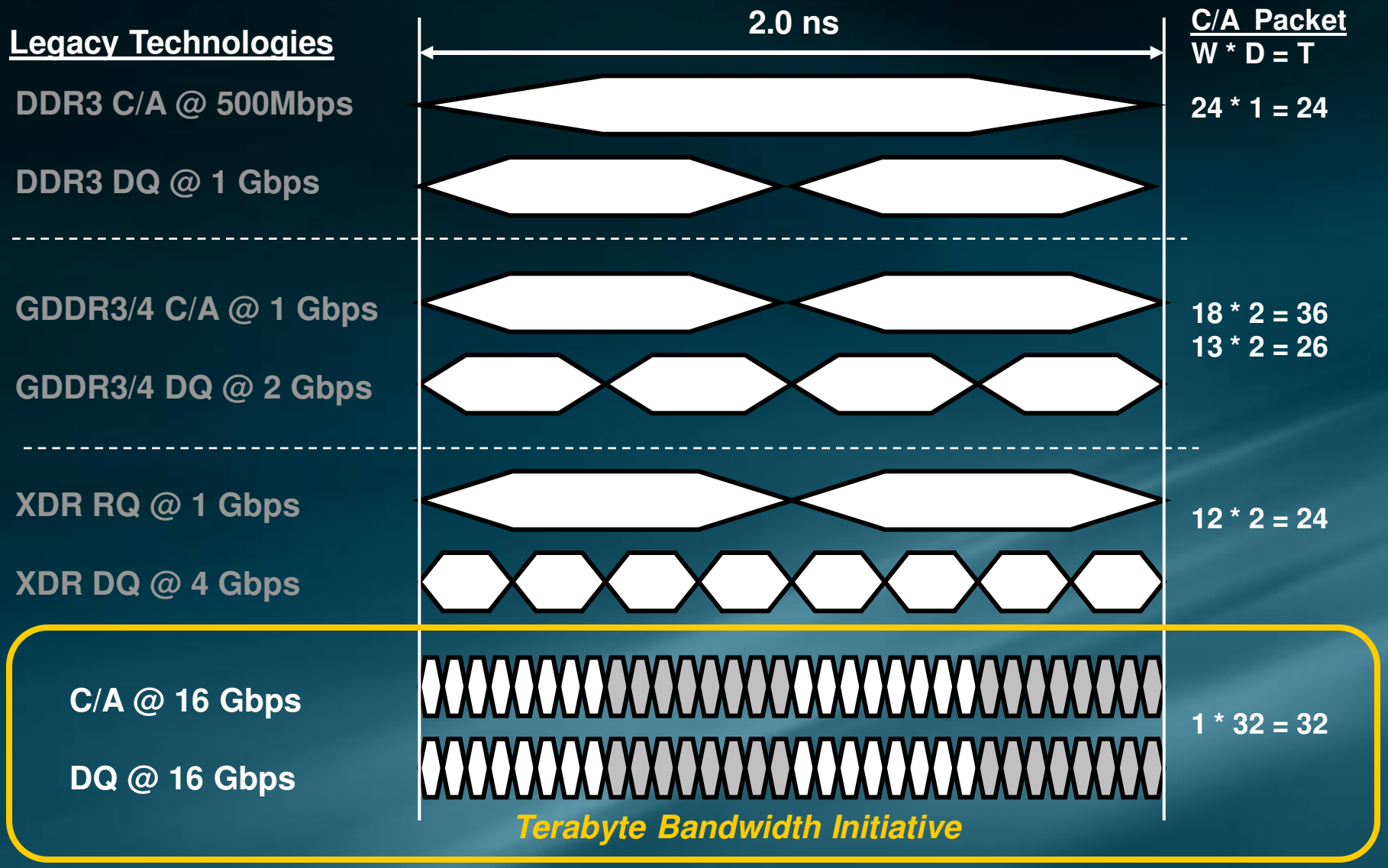
- 32 data bits per reference clock cycle
- 16Gbps data rate with 500MHz clock



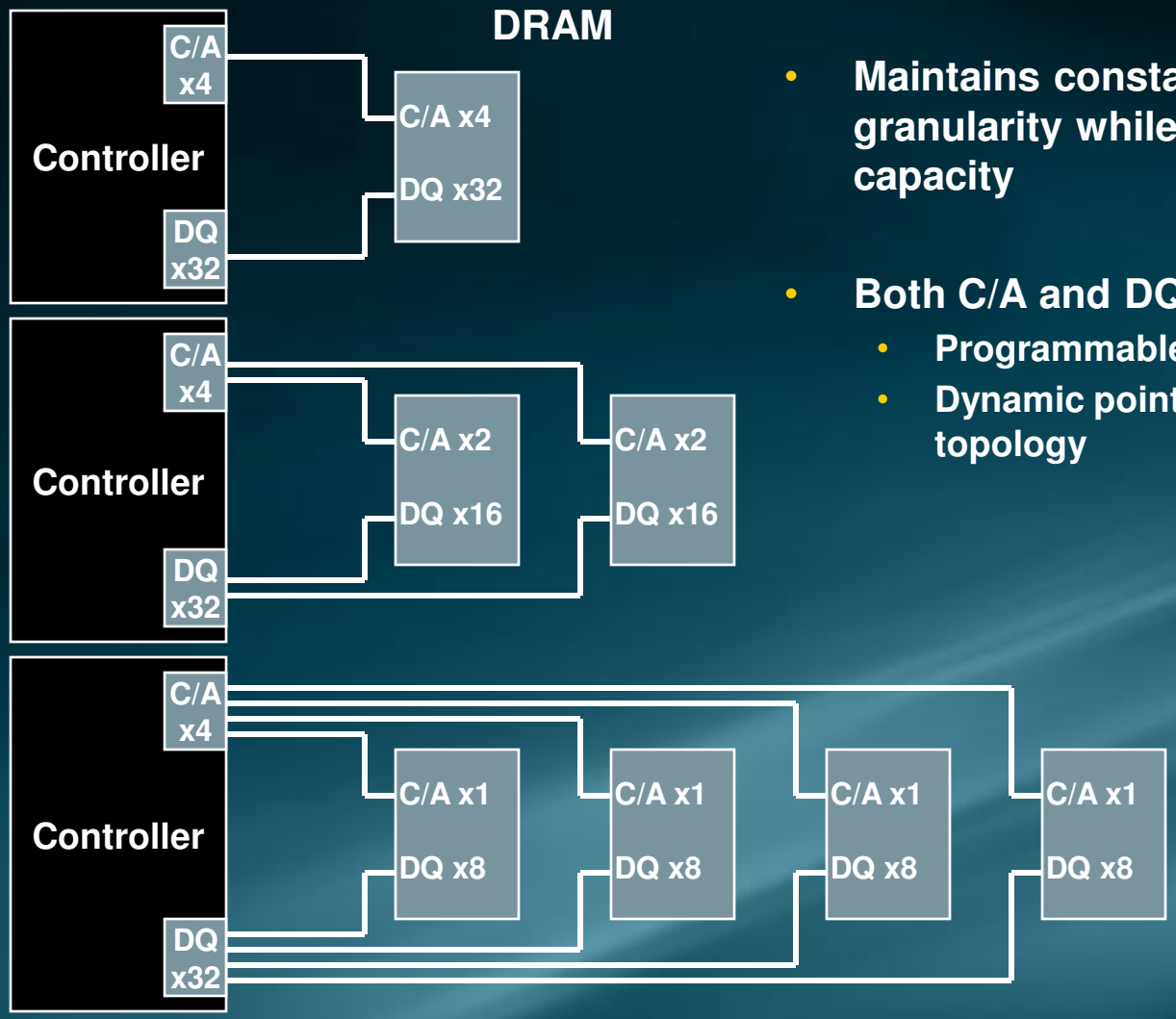
FlexLink™ C/A

- Full command/address channel within a single link
  - Differential point-to-point connection
  - Very pin efficient … Operates at DQ rate
  - FlexPhase™ calibration
- Flexible and Scalable
  - Fully concurrently addressed memory
  - Capacity
  - Access granularity
- 32-bit request packet
  - Addressability through 16Gb generation
  - 2ns timing granularity @ 16Gbps

Column/Address Serialization Comparison



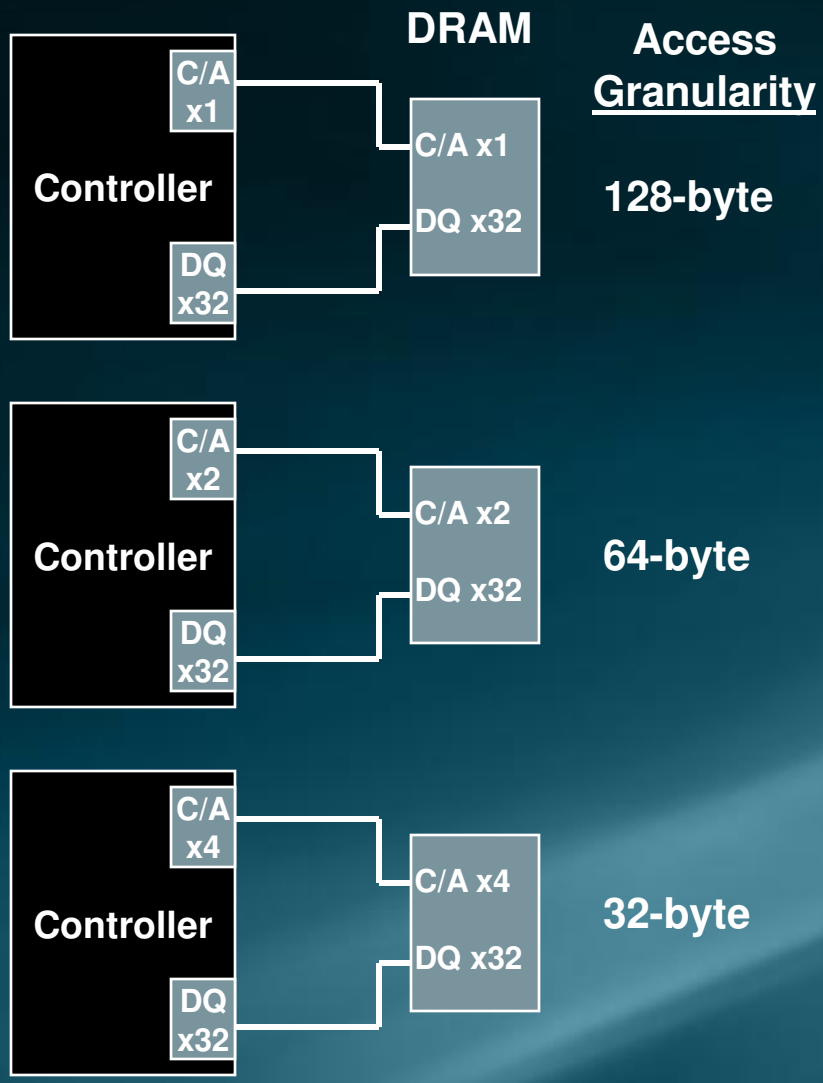
FlexLink™ C/A
Example Capacity Scalability



- Maintains constant access granularity while scaling capacity
- Both C/A and DQ support:
  - Programmable width
  - Dynamic point-to-point topology

32 byte access granularity shown

FlexLink™ C/A
Example Granularity Scalability



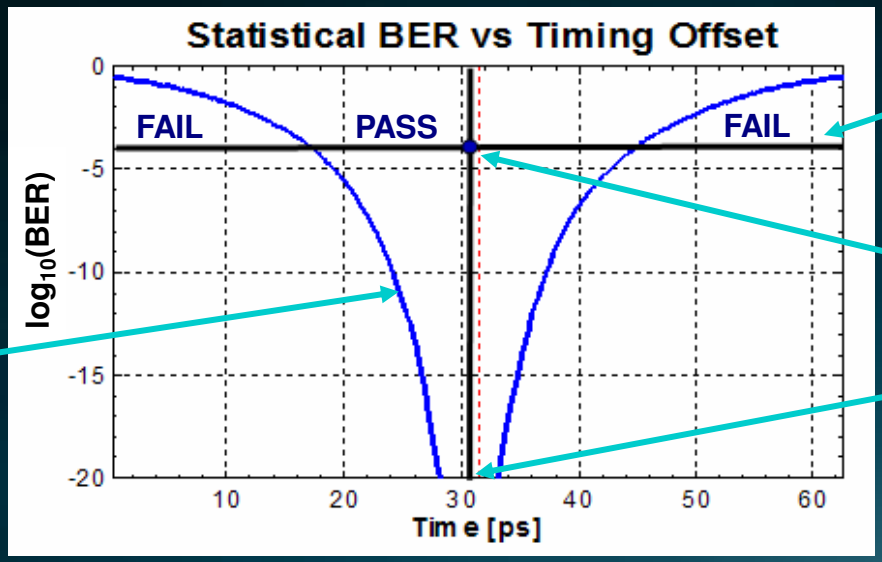
- Same DRAM for all examples
- Controller implements required number of C/A links

Jitter Reduction Technology

- Low phase noise reference clock generator
- Controller
  - LC PLL per byte
  - Ring-based PLL per 2 bit slices
  - AC-coupled CML-to-CMOS converter for clock distribution
- DRAM
  - LC PLL per byte with inductors in 3 metal layers
  - Regulated supplies for VCO & front-end circuits
  - AC-coupled CML-to-CMOS converter for clock distribution

Enhanced FlexPhase™ Calibration

Enhanced FlexPhase™ Calibration



Timing calibration performed at low enough BER to fully account for all DJ sources

Timing calibration center phase

Ideal center phase @ BER = 1E-20

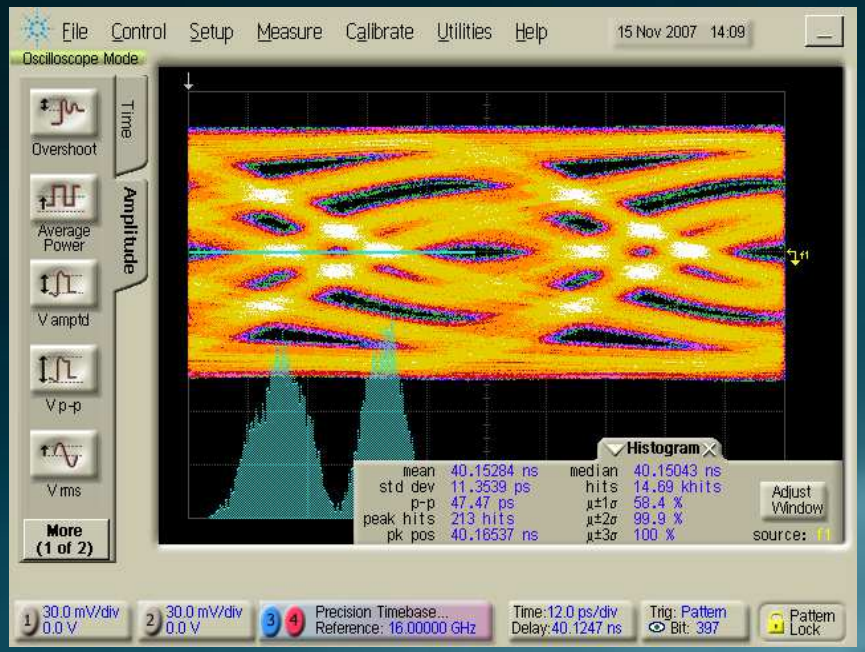
Further eye closure results from RJ only

- Improved linearity and resolution relative to existing XDR technology
  - < 1ps resolution
- Enhanced FlexPhase™ phase adjustment is also used to measure BER bathtub curve during characterization

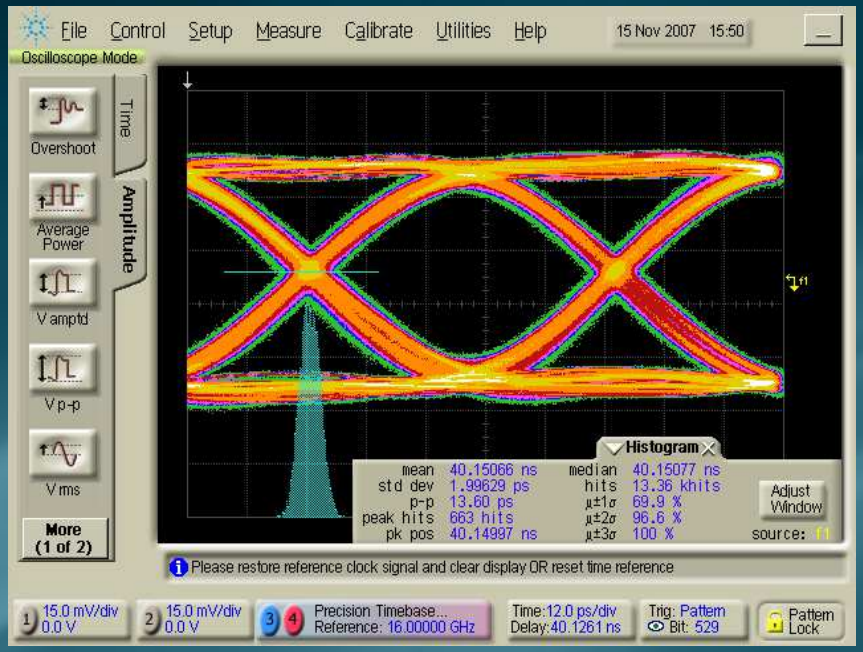
Equalization Enables Robust High Speed Signaling in Future Memory Systems

- Equalization required to reduce ISI at high data rates
- Asymmetric equalization reduces cost, complexity in DRAM

Controller unequalized 16Gbps TX eye



Controller equalized 16Gbps TX eye



Pattern is PRBS 2^11-1
3" FR4 + 12" cable to scope

Terabyte Bandwidth Initiative
Test Chip Primary Objectives

- Demonstrate viability of future controller & DRAM product PHY:

Metric	Target
Performance	16Gbps @ BER ≤ 1E-20
Controller Area	1TB/sec in ~17 x 13 = 221mm^2 die
DRAM Area	64GB/sec in < 800µ PHY height

- Build confidence that DRAM PHY can meet performance, power, and area objectives with a test chip built in an ASIC* process
  - Emulate target DRAM process and constraints to the extent possible

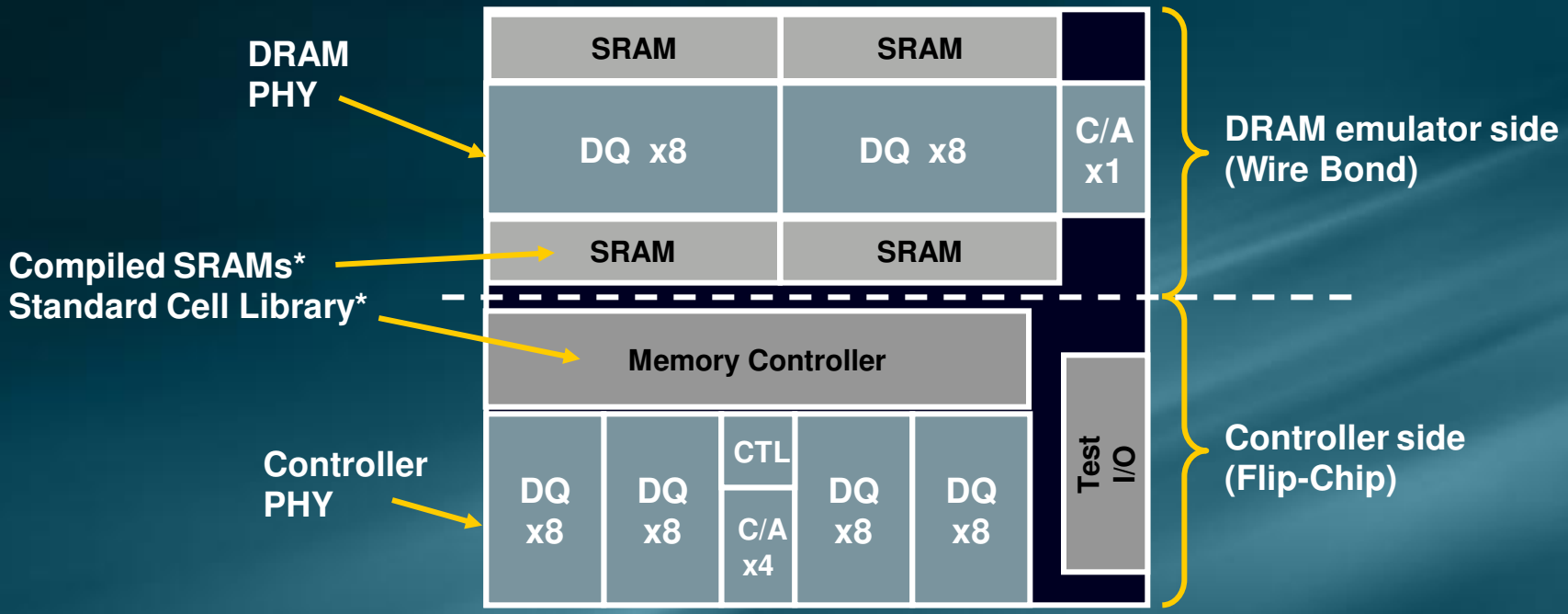
Parameter	Emulation Technique
F04 delay	High Vt transistors + reduced voltage
Metal layer count	Restrict usage to 3 metal layers
Metal sheet resistance	Stack metal layers to match target
Clock distribution	Use realistic floorplan & aspect ratio

- Evaluate circuit architecture tradeoffs
- Create and close system voltage/timing budget

* TSMC 65nm G+

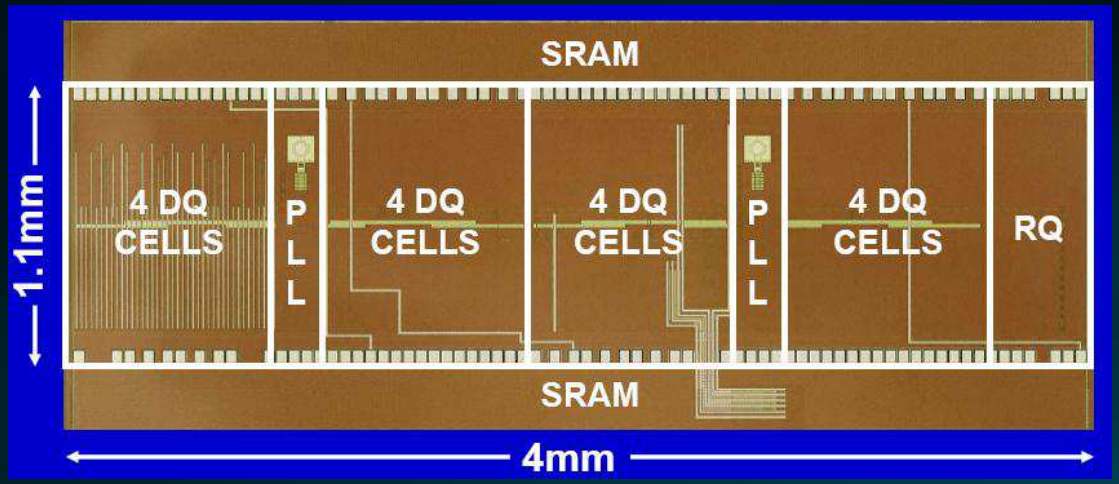
Terabyte Bandwidth Initiative
Test Chip Diagram

- Test chip is effectively two chips in one:
  - Controller PHY and test logic
  - DRAM Emulator (with small SRAM cores)



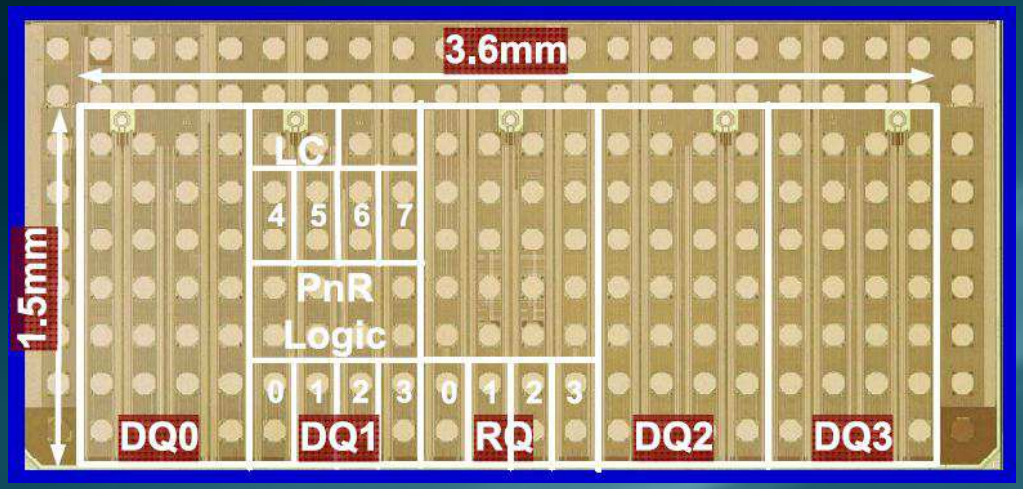
* Virage Logic (TSMC 65nm G+)

Test Chip PHY Photomicrographs



"Half of a x32 DRAM PHY"

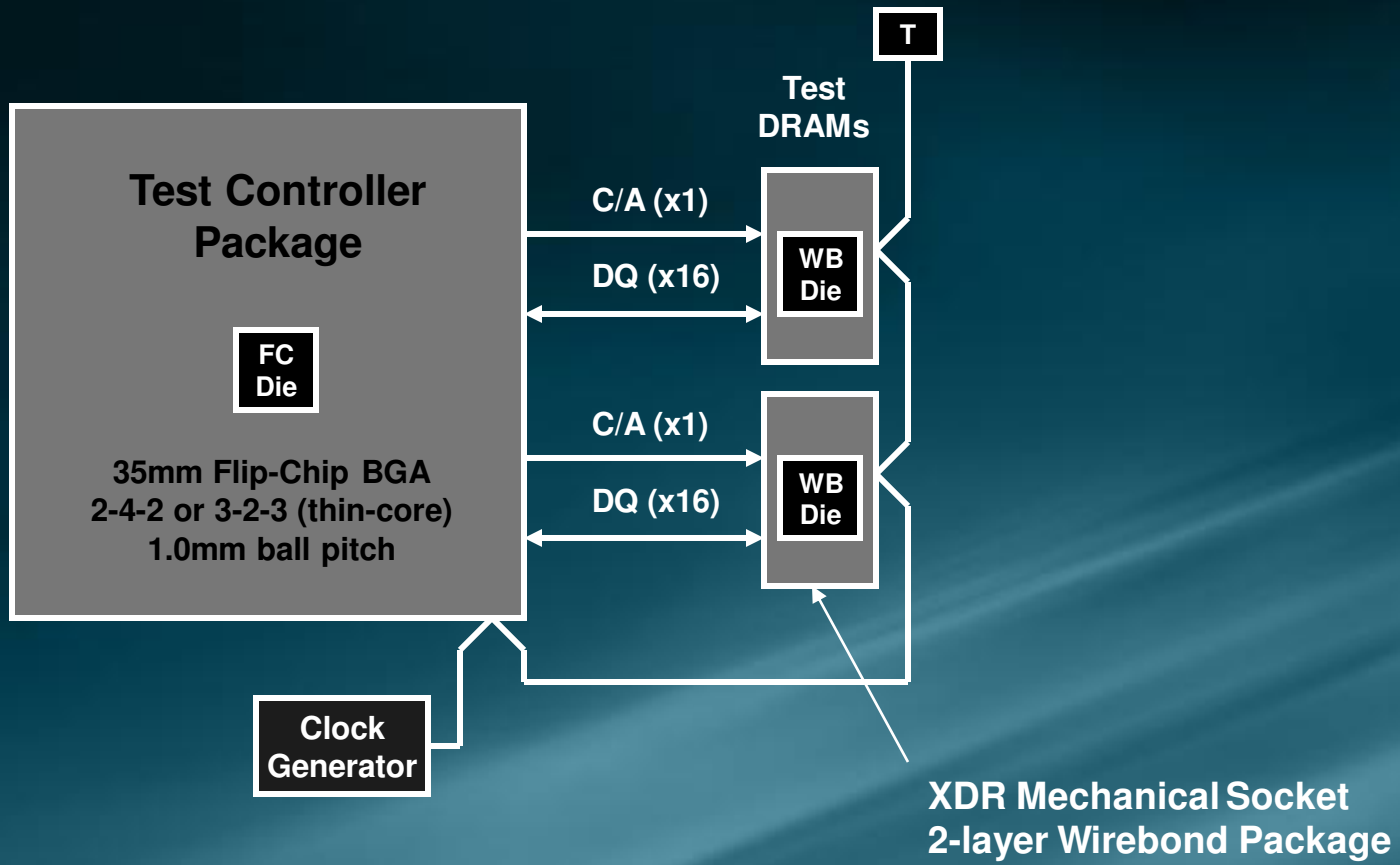
Source: N. Nguyen, et. al. "A 16Gb/s Differential I/O Cell with 380fs RJ in an Emulated 40nm DRAM Process", 2008 Symposium on VLSI Circuits



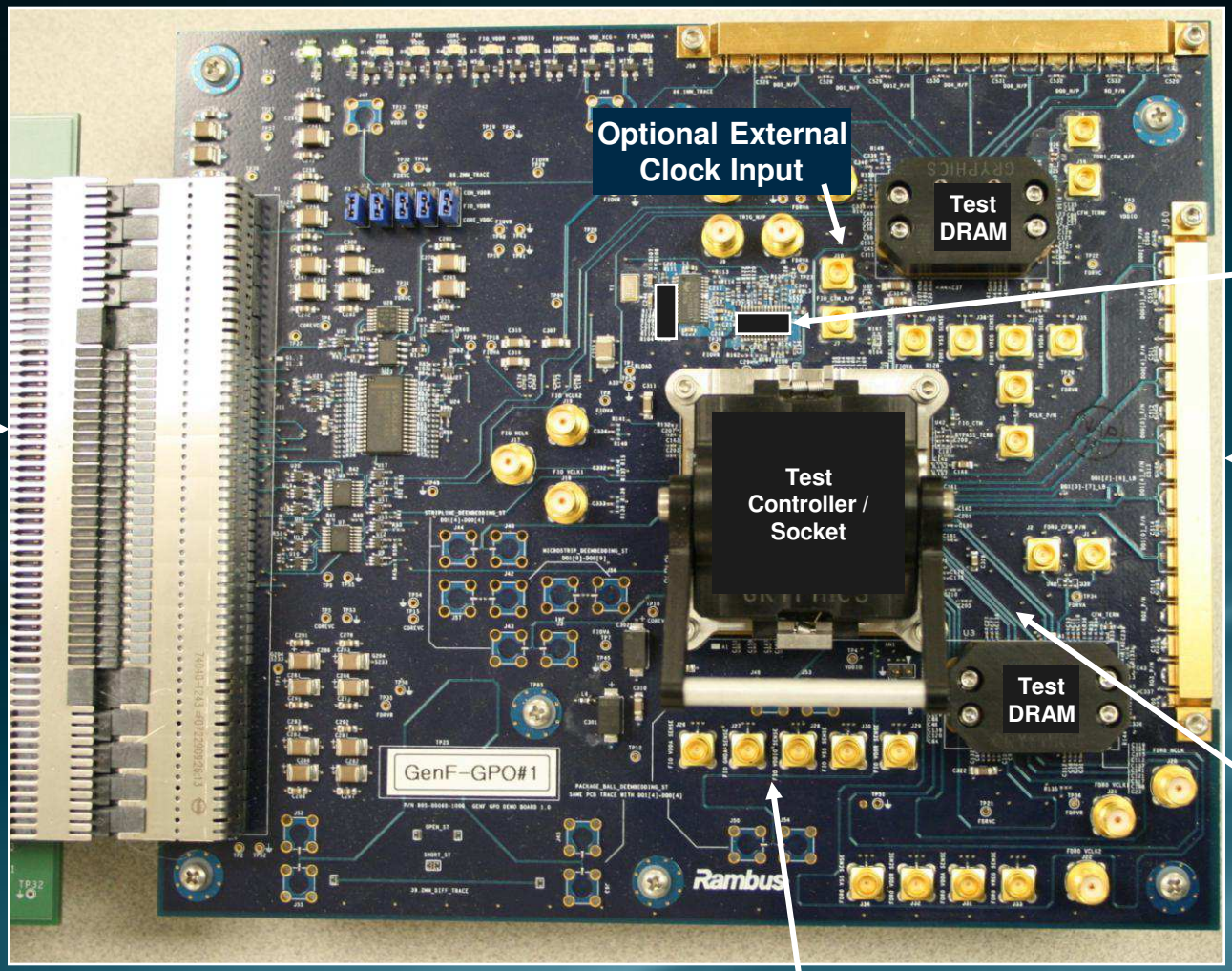
"x32 DQ + x4 RQ Controller PHY Cell"

Source: K. Chang, et. al. "A 16Gb/s/link, 64GB/s Bidirectional Asymmetric Memory Interface Cell", 2008 Symposium on VLSI Circuits

Terabyte Bandwidth Initiative
Demo Platform - Memory Subsystem



Terabyte Bandwidth Initiative Characterization Platform



Outline

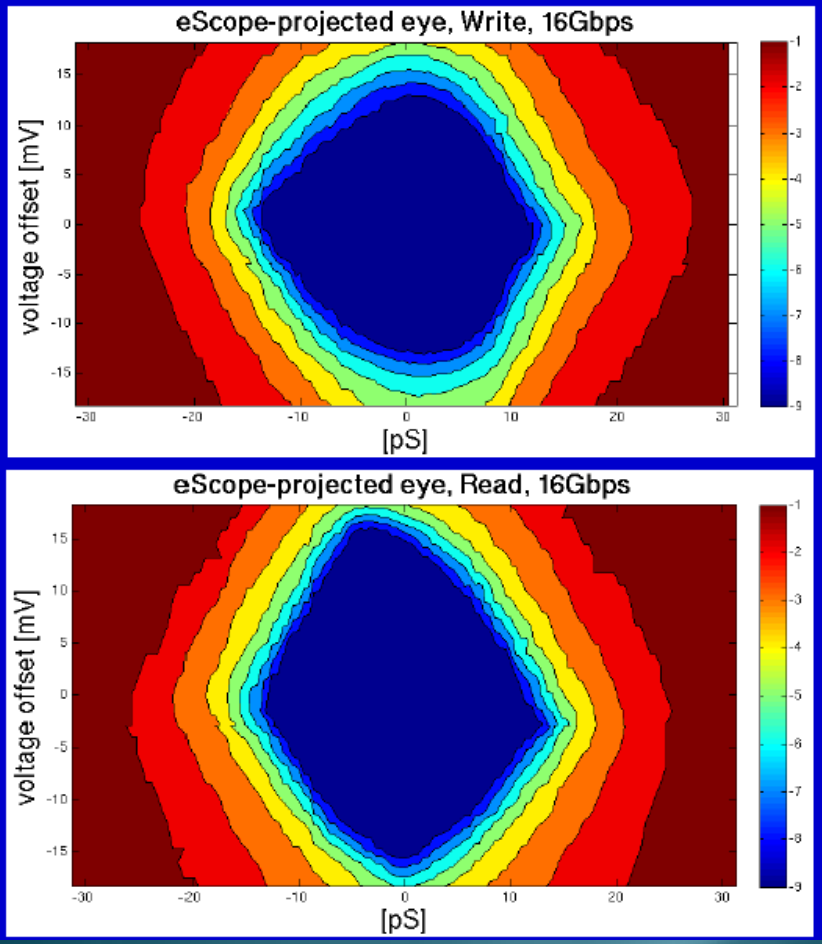
- Memory impact of processor and computing trends
- Throughput memory trends

Terabyte Bandwidth Initiative Summary

- Example 1TB/sec memory system
- *Terabyte Bandwidth Initiative* technology elements
- Test chip and technology status

eScope Receive Eye BER Contours

Receiver Eye Measurements – 16Gb/s



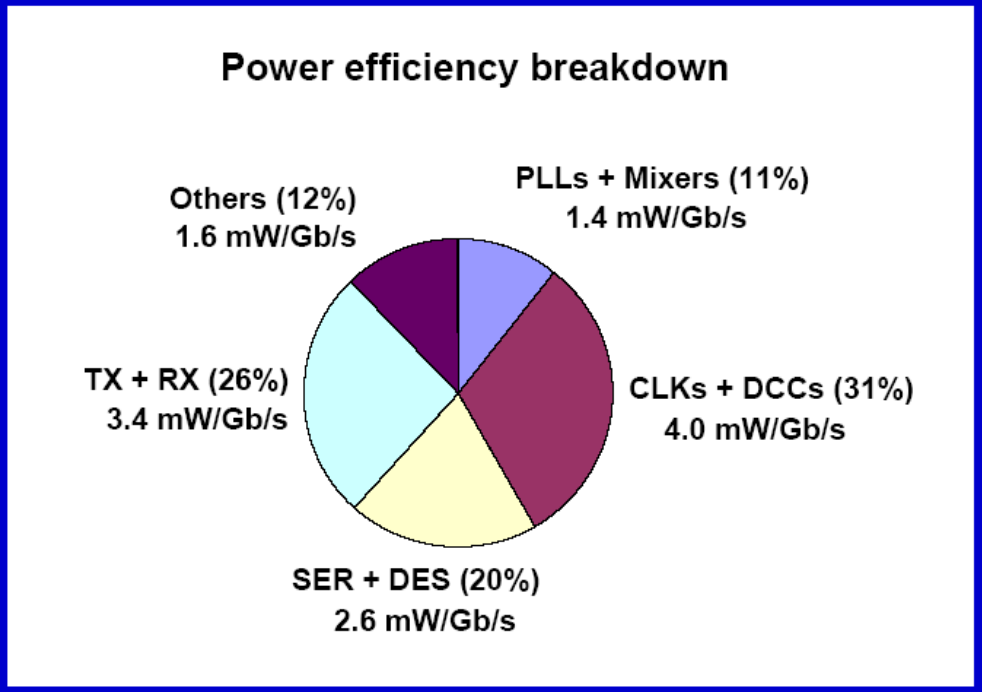
- Measured by using Tx twister to add offset on the channel for both read and write
- Half Tx swing
- Includes receiver effect
- Even with half swing, both directions show large timing margin (>0.3UI for BER = 10^{-9}) and voltage margin (>30mV)

Page 24

Source: K. Chang, et. al. "A 16Gb/s/link, 64GB/s Bidirectional Asymmetric Memory Interface Cell", 2008 Symposium on VLSI Circuits

Controller PHY Power Breakdown (65nm)

Power Breakdown



- 13mW/Gbs for entire interface (8mW/Gbps for DRAM)
- Clock, TX/RX, and high-speed SER/DES dominate

Page 27

Source: K. Chang, et. al. "A 16Gb/s/link, 64GB/s Bidirectional Asymmetric Memory Interface Cell", 2008 Symposium on VLSI Circuits

Summary

- 1TB/sec memory systems can be realized using these technology elements
- Fully Differential Memory Architecture removes existing C/A performance bottleneck
- FlexLink™ C/A allows:
  - Significantly reduced controller pin overhead
  - Scalable capacity
  - Flexible access granularity options
- Signaling rates of 16Gbps have been achieved by leveraging:
  - Jitter reduction technology
  - Asymmetric equalization
  - Enhanced FlexPhase™ calibration

16Gbps Memory signaling demonstrated

Rambus Terabyte Bandwidth Initiative Summary

An *initiative* focused on memory signaling technologies for next-generation Terabyte bandwidth memory systems

- Goal: Deliver *1 TB/s memory bandwidth* to a single System-on-Chip with *16Gbps data rates per DQ link* for next-generation multi-core, gaming, and graphics applications
- Suitable for low-cost, high volume, high performance system manufacturing
- Proven technology ingredients to enable a 1TB/s SOC
- For a more detailed description of the design:
  - N. Nguyen, et. al. "A 16Gb/s Differential I/O Cell with 380fs RJ in an Emulated 40nm DRAM Process", 2008 Symposium on VLSI Circuits
  - K. Chang, et. al. "A 16Gb/s/link, 64GB/s Bidirectional Asymmetric Memory Interface Cell",2008 Symposium on VLSI Circuits