CprE 381: Computer Organization and Assembly Level Programming

Multi-core and GPU Architectures

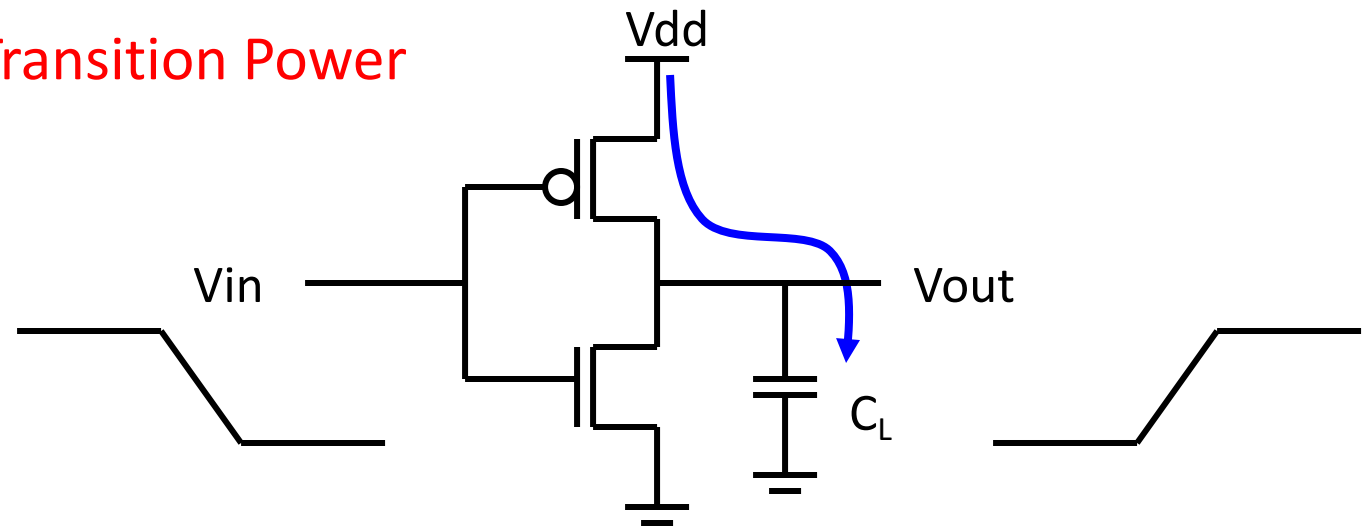
Henry Duwe
Electrical and Computer Engineering
Iowa State University

Administrative

- HW11 due on Mon April 29
  - Real cache exploration
  - Final HW
- Part 4 due in lab next week
  - WARNING: No extensions!!!
- Final Exam
  - When: Mon May 6 at 7:30am
  - Where: Marston 2155 (everyone)
  - What: Control hazards and data forwarding through HW security

Dynamic Power

Transition Power

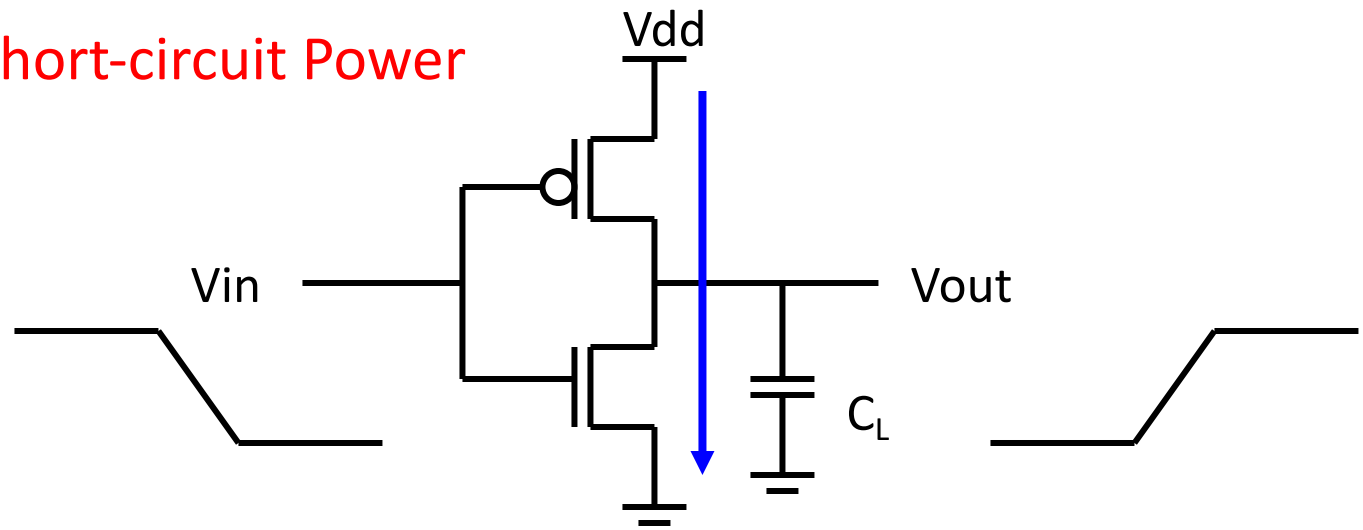


$$\text{Energy/transition} = C_L * V_{DD}^2 * P_{0/1 \rightarrow 1/0}$$

$$\text{Power} = C_L * V_{DD}^2 * f$$

Dynamic Power

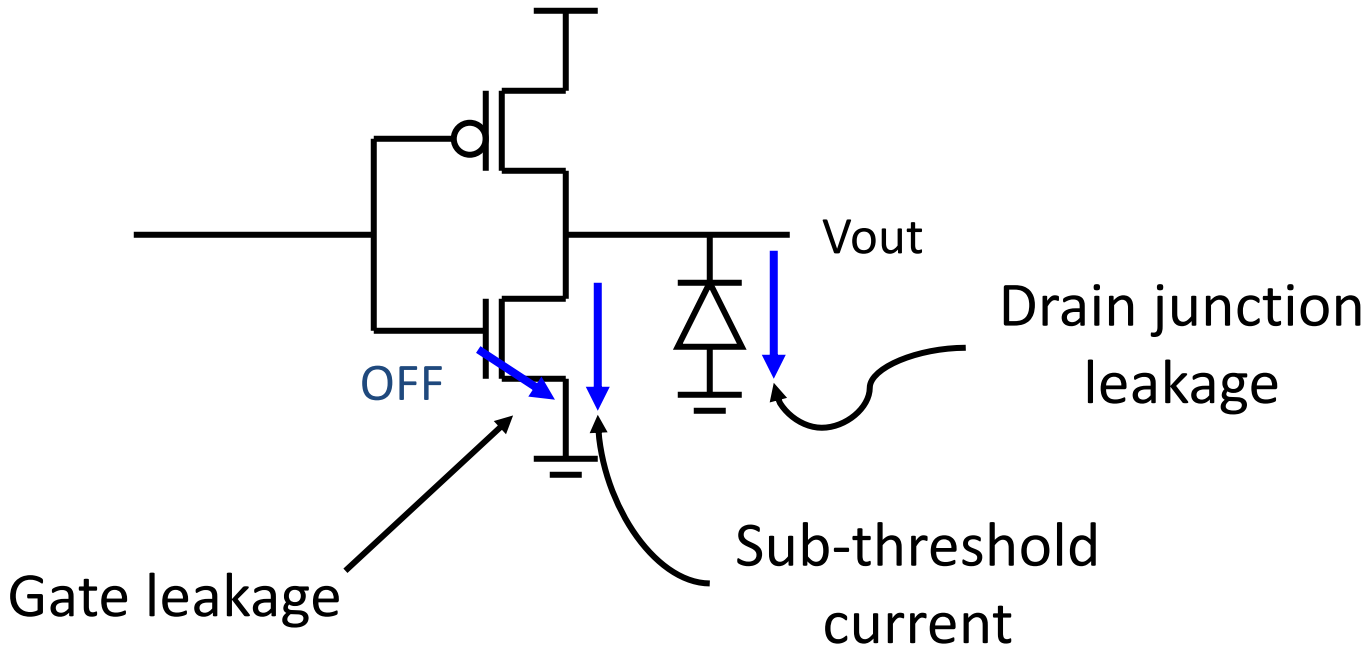
Short-circuit Power



Energy/transition = $t_{sc} * V_{DD} * I_{peak} * P_{0/1 \rightarrow 1/0}$

Power = $t_{sc} * V_{DD} * I_{peak} * f$

Static Power



Independent of switching

Total Power

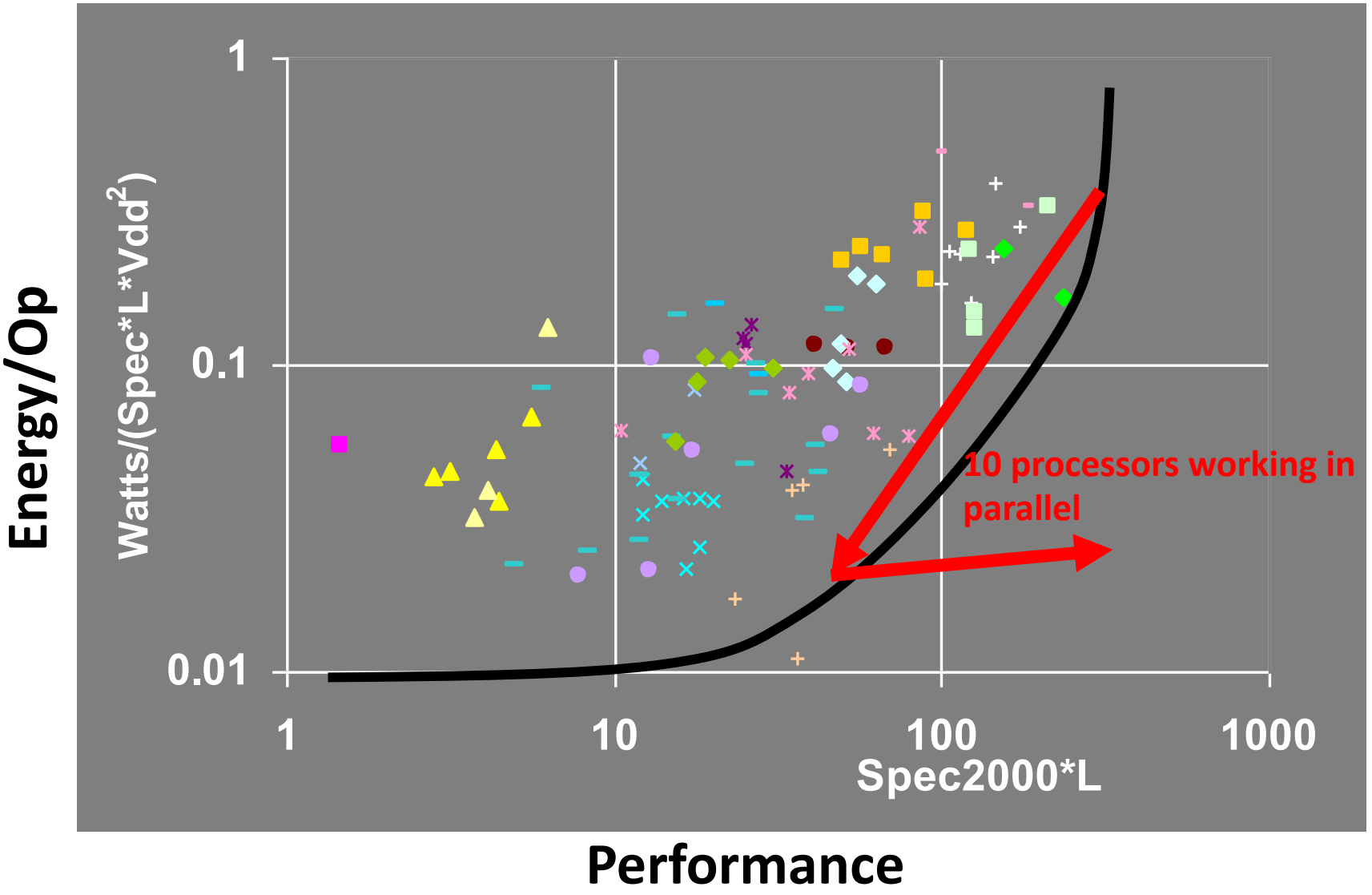
$$P = \alpha f C_L V_{DD}^2 + V_{DD} I_{peak} (P_{0\rightarrow1} + P_{1\rightarrow0}) + V_{DD} I_{leak}$$

Dynamic power (≈ 40 - 70% today and decreasing relatively)

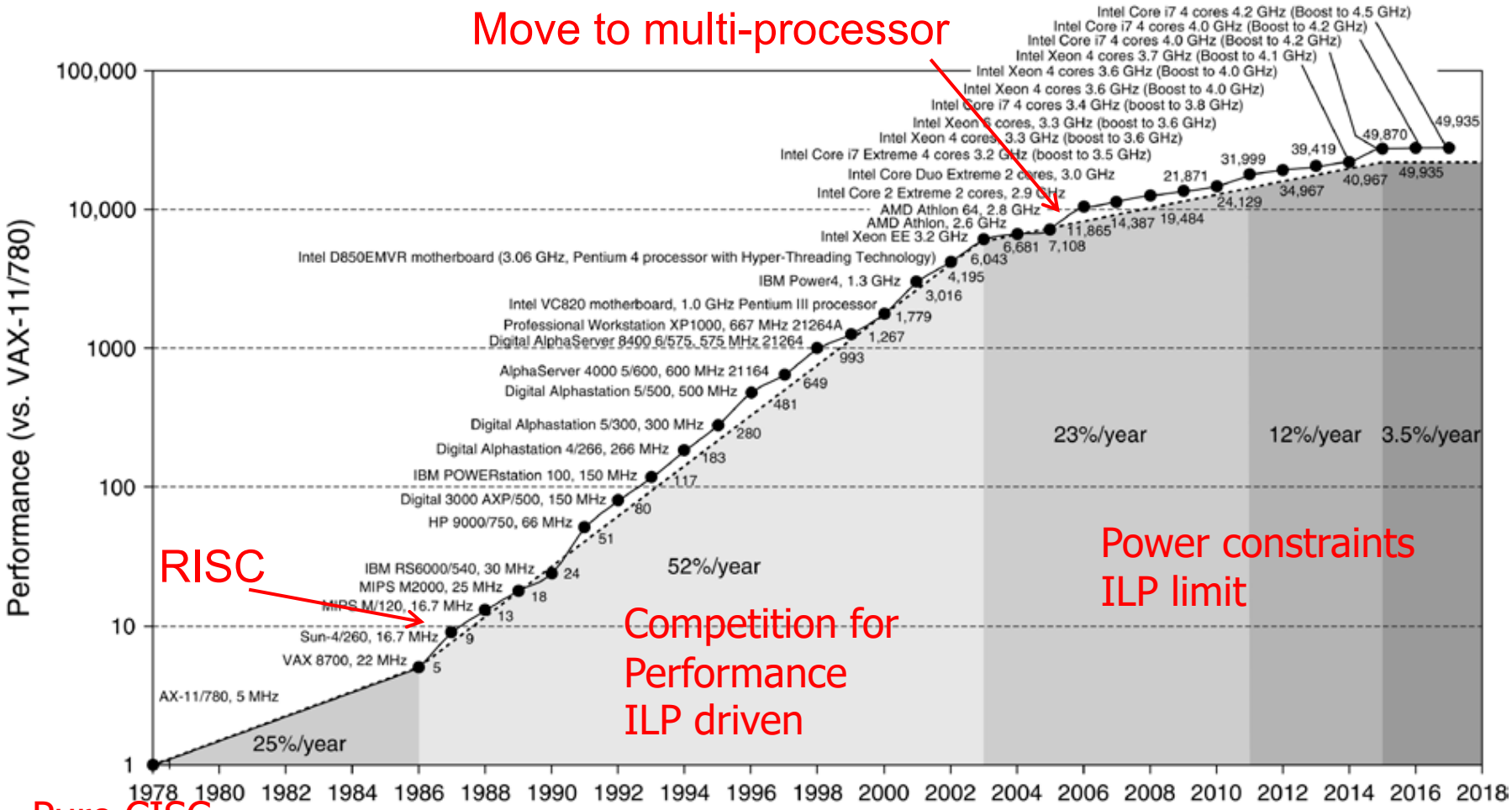
Short-circuit power (≈ 10 % today and decreasing absolutely)

Leakage power (≈ 20 – 50 % today and increasing relatively)

Existing Processors



Computer Architecture Today



Multiprocessors

- Multicore microprocessors
  - More than one processor per chip
- Requires explicitly parallel programming
  - Compare with instruction level parallelism
    - Hardware executes multiple instructions at once
    - Hidden from the programmer
  - Hard to do
    - Programming for performance
    - Load balancing
    - Optimizing communication and synchronization

Instruction and Data Streams

- A classification

		Data Streams	
		Single	Multiple
Instruction Streams	Single	**SISD**: Intel Pentium 4	**SIMD**: SSE instructions of x86
	Multiple	**MISD**: No examples today	**MIMD**: Intel Xeon e5345

- SPMD: Single Program Multiple Data
  - A parallel program on a MIMD computer
  - Conditional code for different processors

Instruction and Data Streams

- A classification

	Data Streams

In-class Assessment!

Access Code: TaxTime

Note: sharing access code to those outside of classroom or using access code while outside of classroom is considered cheating

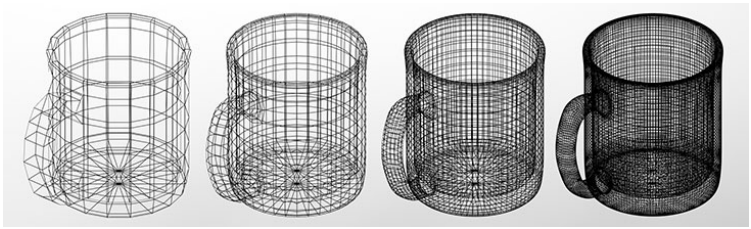
- SPMD: Single Program Multiple Data
  - A parallel program on a MIMD computer
  - Conditional code for different processors

History of GPUs

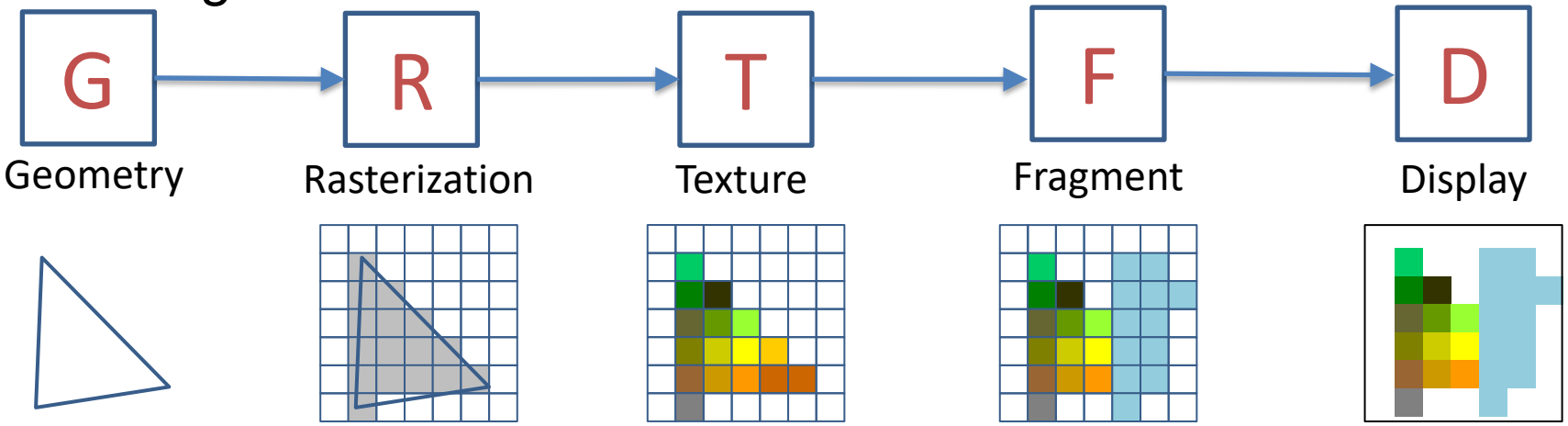
- Early video cards
  - Frame buffer memory with address generation for video output
- 3D graphics processing
  - Originally high-end computers (e.g., SGI)
  - Moore's Law $\Rightarrow$ lower cost, higher density
  - 3D graphics cards for PCs and game consoles
- Graphics Processing Units **(Trend: Accelerators!)**
  - Processors oriented to 3D graphics tasks
  - Vertex/pixel processing, shading, texture mapping, rasterization

Foundation: Computer Graphics

- Use of fundamental geometrical primitives (points, lines, triangles) to convert from 3D coordinates to a 2D screen



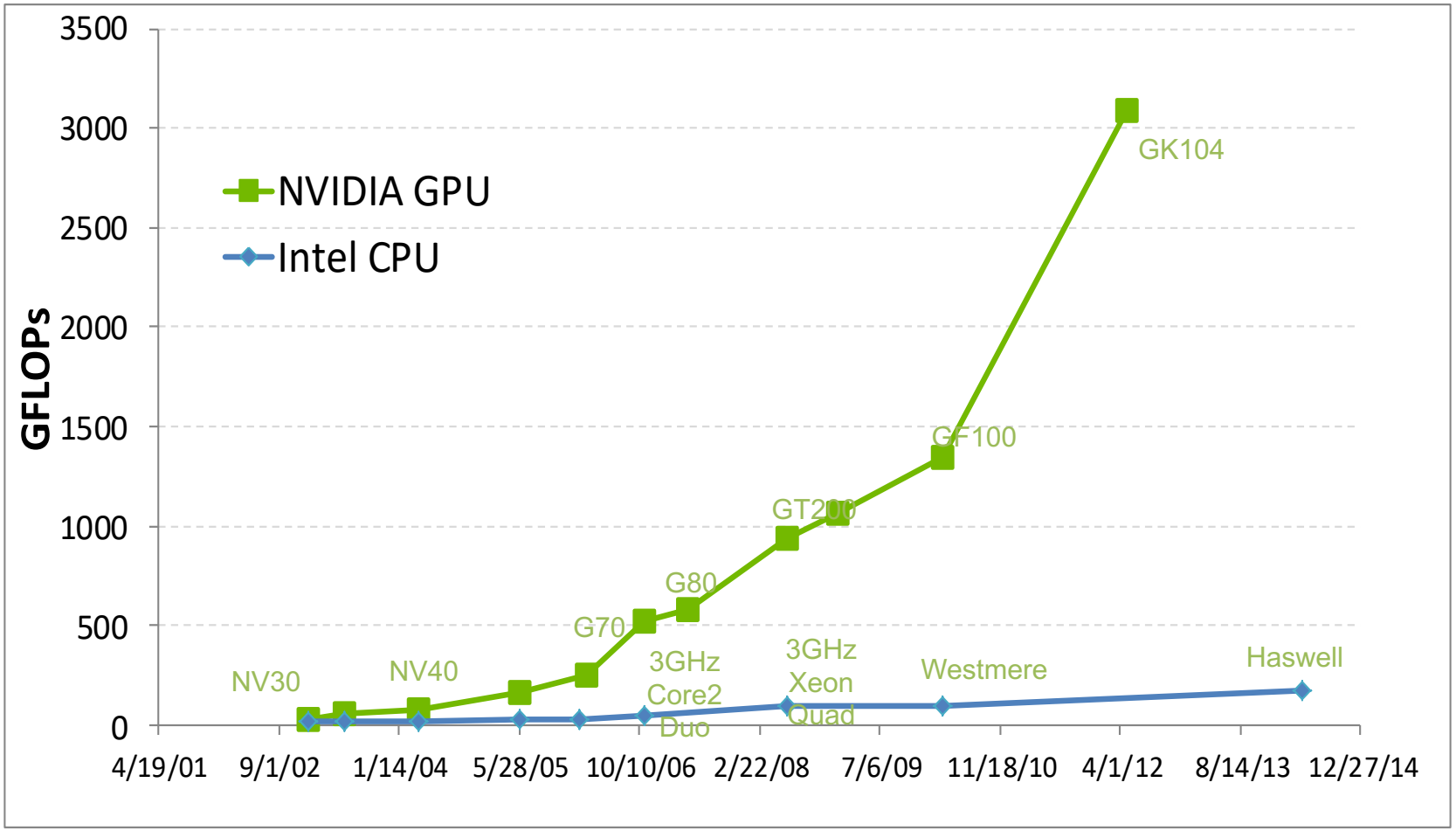
- Typically defined as a set of pipeline stages corresponding to rendering tasks:



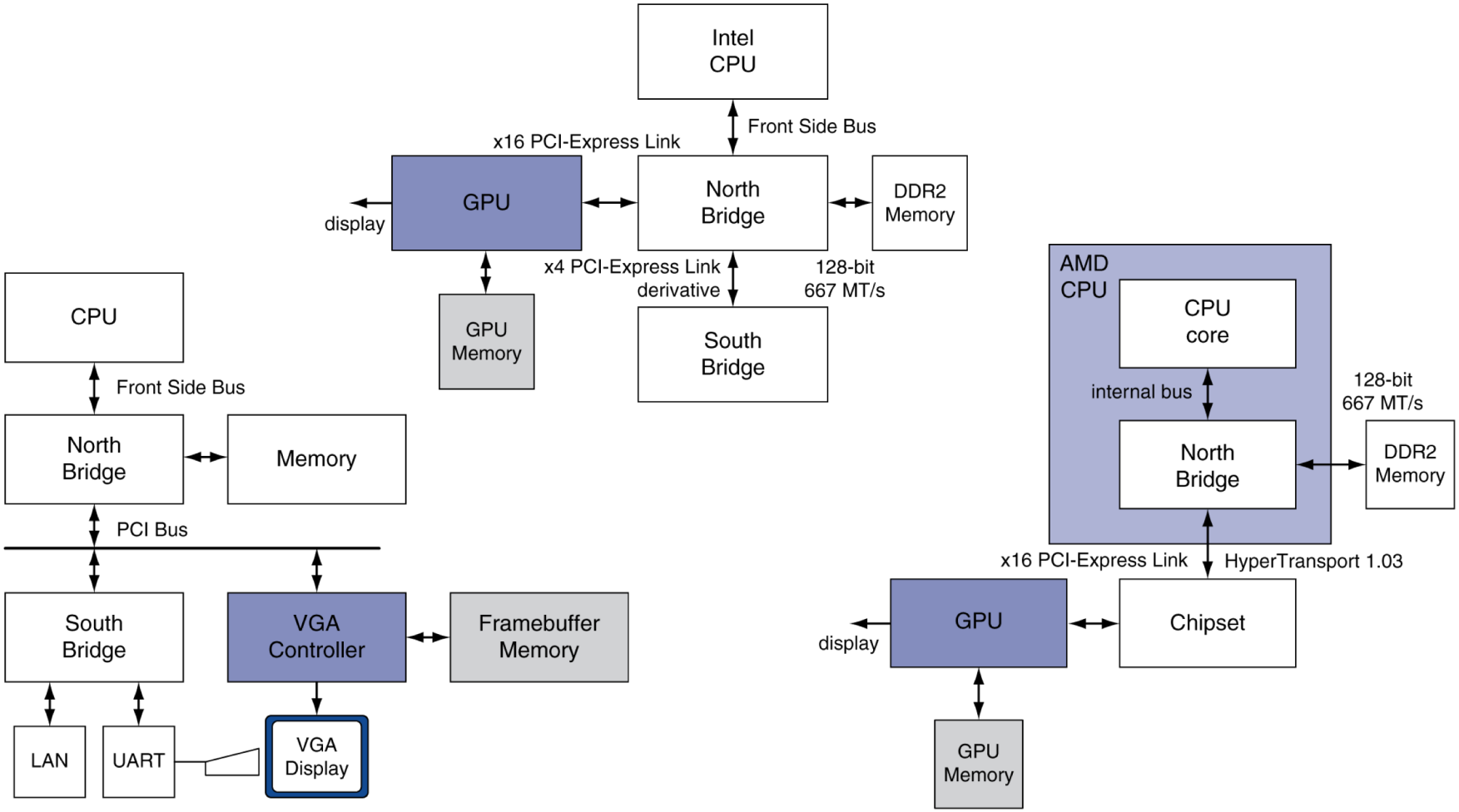
- Each rendering stage then corresponds to multiple (programmable) hardware pipeline stages

So Why Do We Care?

- A (once quiet) revolution and build-up



Graphics in the System



Acknowledgments

- These slides contain material developed and copyright by:
  - Joe Zambreno (Iowa State)
  - Akhilesh Tyagi (Iowa State)
  - David Patterson (UC Berkeley)
  - Mary Jane Irwin (Penn State)
  - Christos Kozyrakis (Stanford)
  - Onur Mutlu (Carnegie Mellon)
  - Krste Asanović (UC Berkeley)
  - Karu Sankaralingam (UW Madison)
  - Morgan Kaufmann