EE482S
Lecture 2
Discussion of 2 Papers
Conclusion of Introductory Material

April 9, 2002

William J. Dally
Computer Systems Laboratory
Stanford University
billd@csl.stanford.edu

Today's Class Meeting

- Discuss two papers
  - Imagine: Media Processing with Streams
    - Khailany et al., IEEE Micro, March-April 2001
  - Polygon Rendering on a Stream Architecture
    - Owens et al., Eurographics HWWS, 2000.
- Conclude introductory discussion on Stream Architecture
  - What is a stream processor

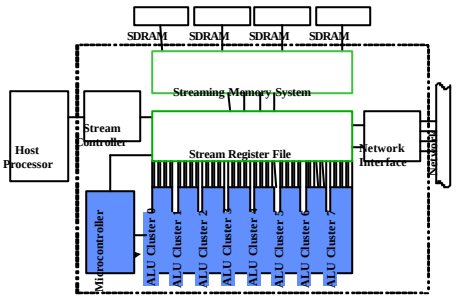
Discussion of Imagine Paper

Discussion of Polygon Rendering Paper

What is a Stream Processor?

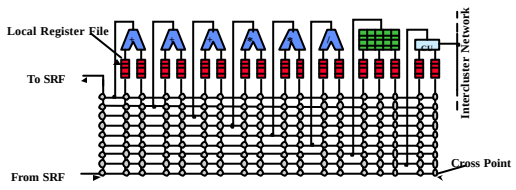
- A processor that is optimized to execute a stream program
- Features include
  - Exploit parallelism
    - TLP with multiple processors
    - DLP with multiple clusters within each processor
    - ILP with multiple ALUs within each cluster
  - Exploit locality with a bandwidth hierarchy
    - Kernel locality within each cluster
    - Producer-consumer locality within each processor
- Many different possible architectures

The Imagine Stream Processor

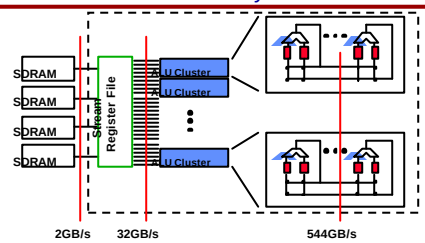


Imagine Stream Processor

Arithmetic Clusters

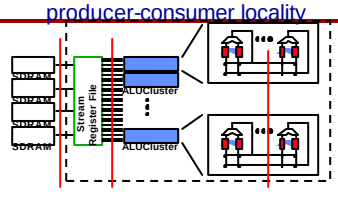


A Bandwidth Hierarchy exploits locality and concurrency



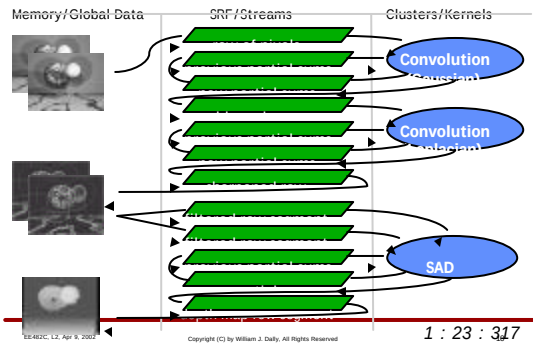
- VLIW clusters with shared control
- 41.2 32-bit floating-point operations per word of memory BW

A Bandwidth Hierarchy exploits kernel and producer-consumer locality



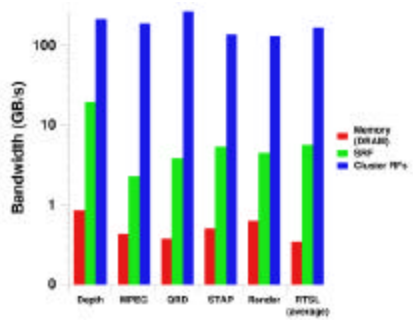
	Memory BW	Global RF BW	Local RF BW
Depth Extractor	. GB/s	. GB/s	. GB/s
MPEG Encoder	. GB/s	. GB/s	. GB/s
Polygon Rendering	. GB/s	. GB/s	. GB/s
QR Decomposition	. GB/s	. GB/s	. GB/s

Producer-Consumer Locality in the Depth Extractor

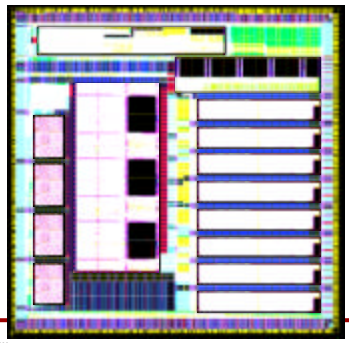


1 : 23 : 317

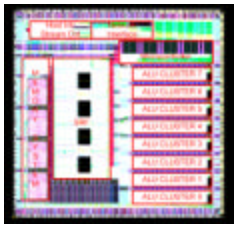
Bandwidth Demand of Applications

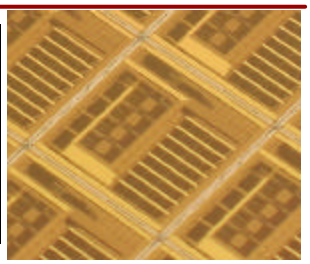


Die Plot



Die Photos

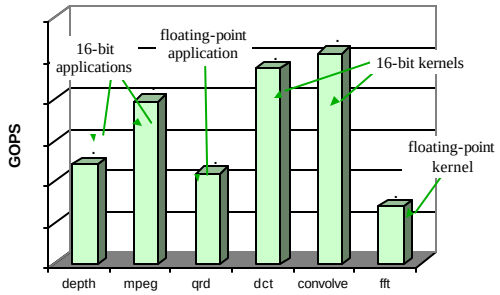




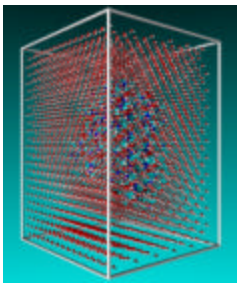
- 21 M transistors / TI 0.15µm 1.5V CMOS / 16mm x 16mm
- 300 MHz TTTT, hope for 400 MHz in lab
- Chips arrived 4/1/02, no fooling!

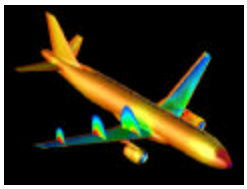
[Khailany et al., ICCD '02 (submitted)]

Performance demonstrated on signal and image processing

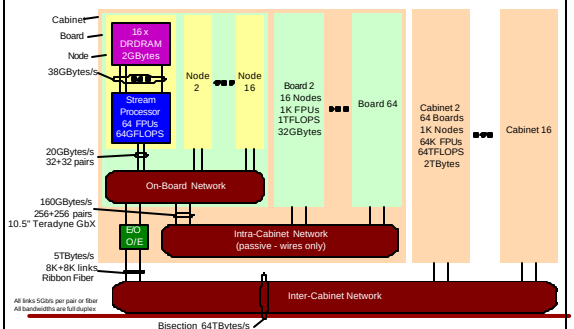


Initial studies indicate that it also applies to solving PDEs and ODEs

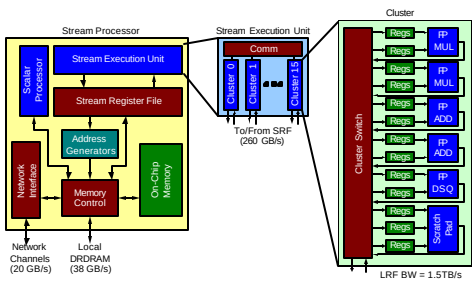




Architecture of a Streaming Supercomputer



Streaming processor



Rough per-node budget

Item	Cost	Per Node
Processor chip	200	200
Router chip	200	50
Memory chip	20	320
Board/Backplane	3000	188
Cabinet	50000	49
Power	1	50
Per-Node Cost		976
$/GFLOPS (64/node)		15
$/M-GUPS (250/node)		4

Preliminary numbers, parts cost only, no I/O included.

Many open problems

- A small sampling
- Software
  - Program transformation
  - Program mapping
  - Bandwidth optimization
  - Conditionals
  - Irregular data structures
- Hardware
  - Alternative stream models
  - Register organization
  - Bandwidth hierarchies
  - Memory organization
  - Short stream issues
  - ISA design
  - Cluster organization
  - Processor organization

Next Time

- Walk through a streaming application