Lecture 8:
Parallel Programming
Case Studies

Parallel Computer Architecture and Programming
CMU 15-418/15-618, Spring 2015
Tunes

The Lumineers

Ho Hey

(The Lumineers)

“Parallelism always seems to dominate the discussion, and I really felt someone needed to put the importance of data locality into song.”

- Wesley Shultz.
Demo: contention
Example: two students make appointments to talk about course material (at 3pm and at 4:30pm)

- Operation to perform: a student getting help with a question
- Execution resource: Professor Kayvon
- Steps in operation:
  1. Student walks from Gates Cafe to Kayvon’s office (5 minutes) = 
  2. Student waits in line (??) = 
  3. Student get question answered with brilliant answer (5 minutes) = 

Time

Student 1
(appt @ 3pm)

Time cost to student: 10 minutes

Student 2
(appt @ 4pm)

Time cost to student: 10 minutes

2:55pm 3pm 3:05pm ... 4:25pm 4:30pm 4:35pm
Office hours from 3-3:20pm (no appointments)

Time cost to student:
- Student 1: 10 minutes
- Student 2: 23 minutes

Problem: contention for shared resource results in longer overall operation times (and likely higher cost to students)
Example: memory system contention in CUDA

```c
#define THREADS_PER_BLK 128

__global__ void my_cuda_program(int N, float* input, float* output)
{
    __shared__ float local_data[THREADS_PER_BLK];
    int index = blockIdx.x * blockDim.x + threadIdx.x;

    // COOPERATIVELY LOAD DATA HERE
    local_data[threadIdx.x] = input[index];

    // WAIT FOR ALL LOADS TO COMPLETE
    __syncthreads();

    // DO WORK HERE ..
    //
}
```

In general, a good rule of thumb when CUDA programming is to make sure you size your thread blocks so that the GPU can fit a couple of thread blocks worth of work per core of the GPU. (This allows threads from one thread block can cover latencies from threads in another block assigned to the same core.)
Today: case studies!

- Five parallel application examples
  - Ocean simulation
  - Galaxy simulation (Barnes-Hut algorithm)
  - Parallel scan
  - Data-parallel segmented scan (Bonus material!)
  - Ray tracing (Bonus material!)

- Will be describing key aspects of the implementations
  - Focus on: optimization techniques, analysis of workload characteristics
Assumption: shared address space

- For the purposes of today’s lecture I encourage you to think about the example applications in the context of a large NUMA shared address space machine.
  (single address space, but each processor can access a local region of the address space more quickly)

- But issues we discuss certainly also arise in a distributed address space setting.
Simulation of Ocean Currents
(grid-based solver)

Example taken from: Culler, Singh, and Gupta, Chapter 3
Simulation of ocean currents

- Discretize ocean volume into slices represented as 2D grids
- Discretize time evolution of ocean: $\Delta t$
- High accuracy simulation requires small $\Delta t$ and high resolution grids

Figure credit: Culler, Singh, and Gupta
First question to ask: where are the dependencies?

Dependencies in one time step of ocean simulation

Parallelism within a grid (data-parallelism) and across operations on the different grids. The implementation only leverages data-parallelism. (simplicity)
Another question to ask: what are the critical working sets?

1. Local neighborhood for cell
2. Three rows of processor’s local partition of grid
3. Processor’s local partition of grid
Ocean implementation details

Recall shared-memory implementation discussed in previous classes:

- **Decomposition:**
  - Spatial partitioning of grid: each processor receives 2D tile of grid

- **Assignment**
  - Static assignment of tiles to processors

- **Synchronization**
  - Barriers (each grid computation is a phase)
  - Locks for mutual exclusion when updating shared variables (atomic update of ‘diff’)

---

CMU 15-418, Spring 2015
Recall: two layouts of 2D grid to memory

(Blue lines indicate consecutive memory addresses)

2D, row-major array layout

4D array layout (block-major)
Observations:

- Static assignment is sufficient (approximately equal busy time per thread)
- 4D blocking of grid reduces time spent on communication (reflected on graph as data wait time)
- Synchronization cost is largely due to waiting at barriers
Galaxy Evolution using Barnes Hut

Image credit: http://www.lsw.uni-heidelberg.de/users/mcamenzi/images/Universe_Box.gif

Example taken from: Culler, Singh, and Gupta, Chapter 3
Galaxy evolution

Barnes-Hut algorithm

- Represent galaxy as a collection of N particles (think: particle = star)
- Compute forces on each particle due to gravity
  - Gravity has infinite extent: naive algorithm is $O(N^2)$ since all particles interact with all others
  - Magnitude of gravitational force falls off with distance (so algorithm approximates forces from far away stars)
  - Result is an $O(N \log N)$ algorithm for computing gravitational forces between all stars
Barnes-Hut tree

- Interior nodes store center of mass + aggregate mass of all child bodies
- To compute forces on each body, traverse tree... accumulating forces from all other bodies
  - Compute forces using aggregate interior node if $L/D < \Theta$, else descend to children
- Expected number of nodes touched $\sim \lg N / \Theta^2$
Barnes-Hut application structure

for each time step in simulation:
  build tree structure
  compute (aggregate mass, center-of-mass) for interior nodes
for each particle:
  traverse tree to accumulate gravitational forces
  update particle position based on gravitational forces

- Challenges:
  - Amount of work per body, and communication pattern of work is non-uniform
    (it depends on the local density of bodies)
  - The bodies move: so costs and communication patterns change over time
  - Irregular, fine-grained computation

- But, there is a lot of locality in the computation (bodies that are near in space
  require similar data to compute forces -- should co-locate these computations!)
Work assignment

- **Challenge:**
  - Equal number of bodies per processor ≠ equal work per processor
  - Want equal work per processor AND assignment should preserve locality

- **Observation:** spatial distribution of bodies evolves slowly

- **Use semi-static assignment**
  - Each time step, for each body, record number of interactions with other bodies (the application profiles itself)
    - Cheap to compute. Just increment local per-body counters
    - Use values to periodically recompute assignment
Assignment using cost zones

- Leverage locality inherent in tree

- Compute total work estimate $W$ for all bodies (computed by summing per-body costs)

- Each processor is assigned $W/P$ of the total work ($P =$ num processors)

- Each processor performs depth-first (post-order) traversal of tree (accumulates work seen so far)

- Processor $P_i$ responsible for processing bodies corresponding to work: $iW/P$ to $(i+1)W/P$

- Each processor can independently compute its assignment of bodies. (The only synchronization required is the sum reduction to compute total amount of work = $W$)

Figure credit: Culler, Singh, and Gupta
Barnes-Hut: working sets

- Working set 1: data needed to compute forces between body-body (or body-node) pairs
- Working set 2: data encountered in an entire tree traversal
  - Expected number of nodes touched for one body: \( \sim \lg N / \Theta^2 \)
  - Computation has high locality: consecutively processed bodies are nearby, so processing touches almost exactly the same nodes!
Barnes-hut: data distribution

- Cost zones technique computes a good work assignment. What about data distribution?

- Difficult to distribute data
  - Work assignment changes with time: would have to dynamically redistribute all simulation data
  - Data accessed at fine granularity (single tree node)

- Luckily: high temporal locality
  - Bodies assigned to same processor are nearby in space, so tree nodes accessed during force computations are very similar.
  - Data for traversal already in cache (Barnes-Hut benefits from large caches, smaller cache line size)

- Result: Unlike OCEAN, data distribution in Barnes-Hut does not significantly impact performance
  - Use static distribution (interleave particles throughout the machine)

Figure credit: Culler and Singh, “Parallel Computer Architecture”
Load balance is good even with static assignment because of random assignment
- On average, each processor does approximately the same amount of work

But random assignment yields poor locality
- Significant amount of inherent communication
- Significant amount of artifactual communication (fine-grained accesses to tree nodes)

Common tension: work balance vs. locality (cost-zones get us both!)
(similar to work balance vs. synchronization trade-offs in “work distribution” lecture)
Summary

- Today so far: two examples of parallel program optimization
- Key issues when discussing the applications
  - How to balance the work?
  - How to exploit locality inherent in the problem?
  - What synchronization is necessary?
Parallel Scan
Data-parallel scan

let \( a = [a_0, a_1, a_2, a_3, \ldots, a_{n-1}] \)

let \( \oplus \) be an associative binary operator with identity element \( I \)

\[
\text{scan}\_\text{inclusive}(\oplus, a) = [a_0, a_0 \oplus a_1, a_0 \oplus a_1 \oplus a_2, \ldots] \\
\text{scan}\_\text{exclusive}(\oplus, a) = [I, a_0, a_0 \oplus a_1, \ldots]
\]

If operator is \( + \), then \( \text{scan}\_\text{inclusive}(+, a) \) is a prefix sum

\[
\text{prefix}\_\text{sum}(a) = [a_0, a_0 + a_1, a_0 + a_1 + a_2, \ldots]
\]
Data-parallel inclusive scan
(Just subtract original vector to get exclusive scan result: not shown)

Work: $O(N \lg N)$
Inefficient compared to sequential algorithm!
Span: $O(\lg N)$
Work-efficient parallel exclusive scan (O(N) work)

<table>
<thead>
<tr>
<th>a₀</th>
<th>a₁</th>
<th>a₂</th>
<th>a₃</th>
<th>a₄</th>
<th>a₅</th>
<th>a₆</th>
<th>a₇</th>
<th>a₈</th>
<th>a₉</th>
<th>a₁₀</th>
<th>a₁₁</th>
<th>a₁₂</th>
<th>a₁₃</th>
<th>a₁₄</th>
<th>a₁₅</th>
</tr>
</thead>
<tbody>
<tr>
<td>a₀</td>
<td>a₀₋₁</td>
<td>a₂</td>
<td>a₂₋₃</td>
<td>a₄</td>
<td>a₄₋₅</td>
<td>a₆</td>
<td>a₆₋₇</td>
<td>a₈</td>
<td>a₈₋₉</td>
<td>a₁₀</td>
<td>a₁₀₋₁₁</td>
<td>a₁₂</td>
<td>a₁₂₋₃</td>
<td>a₁₄</td>
<td>a₁₄₋₁₅</td>
</tr>
<tr>
<td>a₀</td>
<td>a₀₋₁</td>
<td>a₂</td>
<td>a₀₋₃</td>
<td>a₄</td>
<td>a₄₋₅</td>
<td>a₆</td>
<td>a₄₋₇</td>
<td>a₈</td>
<td>a₈₋₉</td>
<td>a₁₀</td>
<td>a₈₋₁₁</td>
<td>a₁₂</td>
<td>a₁₂₋₃</td>
<td>a₁₄</td>
<td>a₁₂₋₁₅</td>
</tr>
<tr>
<td>a₀</td>
<td>a₀₋₁</td>
<td>a₂</td>
<td>a₀₋₃</td>
<td>a₄</td>
<td>a₄₋₅</td>
<td>a₆</td>
<td>a₀₋₇</td>
<td>a₈</td>
<td>a₈₋₉</td>
<td>a₁₀</td>
<td>a₈₋₁₁</td>
<td>a₁₂</td>
<td>a₁₂₋₃</td>
<td>a₁₄</td>
<td>a₈₋₁₅</td>
</tr>
<tr>
<td>a₀</td>
<td>a₀₋₁</td>
<td>a₂</td>
<td>a₀₋₃</td>
<td>a₄</td>
<td>a₄₋₅</td>
<td>a₆</td>
<td>a₀₋₇</td>
<td>a₈</td>
<td>a₈₋₉</td>
<td>a₁₀</td>
<td>a₈₋₁₁</td>
<td>a₁₂</td>
<td>a₁₂₋₃</td>
<td>a₁₄</td>
<td>a₀₋₇</td>
</tr>
<tr>
<td>a₀</td>
<td>a₀₋₁</td>
<td>a₂</td>
<td>a₀₋₃</td>
<td>a₄</td>
<td>a₄₋₅</td>
<td>a₆</td>
<td>a₀₋₃</td>
<td>a₈</td>
<td>a₈₋₉</td>
<td>a₁₀</td>
<td>a₀₋₇</td>
<td>a₁₂</td>
<td>a₁₂₋₃</td>
<td>a₁₄</td>
<td>a₀₋₁₁</td>
</tr>
<tr>
<td>a₀</td>
<td>a₀₋₁</td>
<td>a₂</td>
<td>a₀₋₁</td>
<td>a₄</td>
<td>a₄₋₅</td>
<td>a₆</td>
<td>a₀₋₅</td>
<td>a₈</td>
<td>a₀₋₇</td>
<td>a₁₀</td>
<td>a₀₋₉</td>
<td>a₁₂</td>
<td>a₀₋₁₁</td>
<td>a₁₄</td>
<td>a₀₋₁₃</td>
</tr>
<tr>
<td>a₀</td>
<td>a₀₋₁</td>
<td>a₂</td>
<td>a₀₋₁</td>
<td>a₄</td>
<td>a₀₋₃</td>
<td>a₆</td>
<td>a₀₋₅</td>
<td>a₈</td>
<td>a₀₋₇</td>
<td>a₁₀</td>
<td>a₀₋₉</td>
<td>a₁₂</td>
<td>a₀₋₁₁</td>
<td>a₁₄</td>
<td>a₀₋₁₃</td>
</tr>
<tr>
<td>a₀</td>
<td>a₀₋₁</td>
<td>a₂</td>
<td>a₀₋₁</td>
<td>a₄</td>
<td>a₀₋₃</td>
<td>a₆</td>
<td>a₀₋₅</td>
<td>a₈</td>
<td>a₀₋₇</td>
<td>a₁₀</td>
<td>a₀₋₉</td>
<td>a₁₂</td>
<td>a₀₋₁₁</td>
<td>a₁₄</td>
<td>a₀₋₁₃</td>
</tr>
<tr>
<td>a₀</td>
<td>a₀₋₁</td>
<td>a₂</td>
<td>a₀₋₁</td>
<td>a₄</td>
<td>a₀₋₃</td>
<td>a₆</td>
<td>a₀₋₅</td>
<td>a₈</td>
<td>a₀₋₇</td>
<td>a₁₀</td>
<td>a₀₋₉</td>
<td>a₁₂</td>
<td>a₀₋₁₁</td>
<td>a₁₄</td>
<td>a₀₋₁₃</td>
</tr>
<tr>
<td>a₀</td>
<td>a₀₋₁</td>
<td>a₂</td>
<td>a₀₋₁</td>
<td>a₄</td>
<td>a₀₋₃</td>
<td>a₆</td>
<td>a₀₋₅</td>
<td>a₈</td>
<td>a₀₋₇</td>
<td>a₁₀</td>
<td>a₀₋₉</td>
<td>a₁₂</td>
<td>a₀₋₁₁</td>
<td>a₁₄</td>
<td>a₀₋₁₃</td>
</tr>
<tr>
<td>a₀</td>
<td>a₀₋₁</td>
<td>a₂</td>
<td>a₀₋₁</td>
<td>a₄</td>
<td>a₀₋₃</td>
<td>a₆</td>
<td>a₀₋₅</td>
<td>a₈</td>
<td>a₀₋₇</td>
<td>a₁₀</td>
<td>a₀₋₉</td>
<td>a₁₂</td>
<td>a₀₋₁₁</td>
<td>a₁₄</td>
<td>a₀₋₁₃</td>
</tr>
<tr>
<td>a₀</td>
<td>a₀₋₁</td>
<td>a₂</td>
<td>a₀₋₁</td>
<td>a₄</td>
<td>a₀₋₃</td>
<td>a₆</td>
<td>a₀₋₅</td>
<td>a₈</td>
<td>a₀₋₇</td>
<td>a₁₀</td>
<td>a₀₋₉</td>
<td>a₁₂</td>
<td>a₀₋₁₁</td>
<td>a₁₄</td>
<td>a₀₋₁₃</td>
</tr>
<tr>
<td>a₀</td>
<td>a₀₋₁</td>
<td>a₂</td>
<td>a₀₋₁</td>
<td>a₄</td>
<td>a₀₋₃</td>
<td>a₆</td>
<td>a₀₋₅</td>
<td>a₈</td>
<td>a₀₋₇</td>
<td>a₁₀</td>
<td>a₀₋₉</td>
<td>a₁₂</td>
<td>a₀₋₁₁</td>
<td>a₁₄</td>
<td>a₀₋₁₃</td>
</tr>
<tr>
<td>a₀</td>
<td>a₀₋₁</td>
<td>a₂</td>
<td>a₀₋₁</td>
<td>a₄</td>
<td>a₀₋₃</td>
<td>a₆</td>
<td>a₀₋₅</td>
<td>a₈</td>
<td>a₀₋₇</td>
<td>a₁₀</td>
<td>a₀₋₉</td>
<td>a₁₂</td>
<td>a₀₋₁₁</td>
<td>a₁₄</td>
<td>a₀₋₁₃</td>
</tr>
<tr>
<td>a₀</td>
<td>a₀₋₁</td>
<td>a₂</td>
<td>a₀₋₁</td>
<td>a₄</td>
<td>a₀₋₃</td>
<td>a₆</td>
<td>a₀₋₅</td>
<td>a₈</td>
<td>a₀₋₇</td>
<td>a₁₀</td>
<td>a₀₋₉</td>
<td>a₁₂</td>
<td>a₀₋₁₁</td>
<td>a₁₄</td>
<td>a₀₋₁₃</td>
</tr>
</tbody>
</table>

CMU 15-418, Spring 2015
Work efficient exclusive scan algorithm

Up-sweep:

for d=0 to (log₂n - 1) do
  forall k=0 to n-1 by 2^{d+1} do
    a[k + 2^{d+1} - 1] = a[k + 2^d - 1] + a[k + 2^{d+1} - 1]

Down-sweep:

x[n-1] = 0
for d=(log₂n - 1) down to 0 do
  forall k=0 to n-1 by 2^{d+1} do
    tmp = a[k + 2^d - 1]
    a[k + 2^d - 1] = a[k + 2^{d+1} - 1]
    a[k + 2^{d+1} - 1] = tmp + a[k + 2^{d+1} - 1]

Work: O(N) (but what is the constant?)
Span: O(lg N) (but what is the constant?)
Locality: ??
Scan implementation: two cores
Exclusive scan: two processor implementation

<table>
<thead>
<tr>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
<th>$a_4$</th>
<th>$a_5$</th>
<th>$a_6$</th>
<th>$a_7$</th>
<th>$a_8$</th>
<th>$a_9$</th>
<th>$a_{10}$</th>
<th>$a_{11}$</th>
<th>$a_{12}$</th>
<th>$a_{13}$</th>
<th>$a_{14}$</th>
<th>$a_{15}$</th>
</tr>
</thead>
</table>

**P1**

Sequential scan on elements [0-7]

Let base = $a_{0-7}$

($a_{0-6}$ already computed as part of exclusive scan, just add $a_7$)

Add base to elements $a_8$ thru $a_{8-11}$

**P2**

Sequential scan on elements [8-15]

Add base to elements $a_{8-12}$ thru $a_{8-15}$

Work: $O(N)$ (but constant is now only 1.5)

Data-access:

- Very high spatial locality (contiguous memory access)
- P1's access to $a_8$ through $a_{8-11}$ may be more costly on large NUMA system, but on small-scale system access likely same cost as from P2
Exclusive scan: SIMD implementation (in CUDA)

Example: perform exclusive scan on 32-element array: SPMD program, assume 32-wide SIMD execution

When `scan_warp` is run by a group of 32 CUDA threads, each thread returns the exclusive scan result for element `idx` (note: upon completion `ptr[]` stores inclusive scan result)

```cpp
template<class OP, class T>
__device__ T scan_warp(volatile T *ptr, const unsigned int idx)
{
    const unsigned int lane = idx & 31; // index of thread in warp (0..31)

    if (lane >= 1) ptr[idx] = OP::apply(ptr[idx - 1], ptr[idx]);
    if (lane >= 2) ptr[idx] = OP::apply(ptr[idx - 2], ptr[idx]);
    if (lane >= 4) ptr[idx] = OP::apply(ptr[idx - 4], ptr[idx]);
    if (lane >= 8) ptr[idx] = OP::apply(ptr[idx - 8], ptr[idx]);
    if (lane >= 16) ptr[idx] = OP::apply(ptr[idx - 16], ptr[idx]);

    return (lane>0) ? ptr[idx-1] : OP::identity();
}
```

Work: ??
Exclusive scan: SIMD implementation (in CUDA)

template<class OP, class T>
__device__ T scan_warp(volatile T *ptr, const unsigned int idx)
{
    const unsigned int lane = idx & 31; // index of thread in warp (0..31)

    if (lane >= 1) ptr[idx] = OP::apply(ptr[idx - 1], ptr[idx]);
    if (lane >= 2) ptr[idx] = OP::apply(ptr[idx - 2], ptr[idx]);
    if (lane >= 4) ptr[idx] = OP::apply(ptr[idx - 4], ptr[idx]);
    if (lane >= 8) ptr[idx] = OP::apply(ptr[idx - 8], ptr[idx]);
    if (lane >= 16) ptr[idx] = OP::apply(ptr[idx - 16], ptr[idx]);

    return (lane>0) ? ptr[idx-1] : OP::identity();
}

Work: $N \log(N)$

Work-efficient formulation of scan is not beneficial in this context because it results in low SIMD utilization. It would require more than 2x the number of instructions as the implementation above!
Building scan on larger array

Example: 128-element scan using four-warp thread block

max length 32 SIMD scan warp 0

length 32 SIMD scan warp 0

length 32 SIMD scan warp 1

length 32 SIMD scan warp 2

length 32 SIMD scan warp 3

add base[0] warp 1

add base[1] warp 2

add base[2] warp 3

base:

\[ a_{0-31} a_{32-63} a_{64-95} a_{96-127} \]
Multi-threaded, SIMD implementation

Example: cooperating threads in a CUDA thread block perform scan

We provided similar code in assignment 2.
Code assumes length of array given by *ptr* is same as number of threads per block.

```cpp
template<class OP, class T>
__device__ void scan_block(volatile T *ptr, const unsigned int idx)
{
    const unsigned int lane = idx & 31;     // index of thread in warp (0..31)
    const unsigned int warp_id = idx >> 5;  // warp index in block

    T val = scan_warp<OP,T>(ptr, idx);     // Step 1. per-warp partial scan
    // (Performed by all threads in block,
    // with threads in same warp communicating
    // through shared memory buffer ‘ptr’)

    if (lane == 31) ptr[warp_id] = ptr[idx];     // Step 2. thread 31 in each warp copy
    __syncthreads();  // partial-scan bases in per-block
                       // shared mem

    if (warp_id == 0) scan_warp<OP, T>(ptr, idx);  // Step 3. scan to accumulate bases
    __syncthreads();  // (only performed by warp 0)

    if (warp_id > 0)
        val = OP::apply(ptr[warp_id-1], val);     // Step 4. apply bases to all elements
    __syncthreads();  // (performed by all threads in block)

    ptr[idx] = val;
}
```
Building a larger scan

Example: one million element scan (1024 elements per block)

Exceeding 1 million elements requires partitioning phase two into multiple blocks
Scan implementation

- **Parallelism**
  - Scan algorithm features $O(N)$ parallel work
  - But efficient implementations only leverage as much parallelism as required to make good utilization of the machine
    - Reduce work and reduce communication/synchronization

- **Locality**
  - Multi-level implementation to match memory hierarchy
    (CUDA example: per-block implementation carried out in local memory)

- **Heterogeneity: different strategy at different machine levels**
  - CUDA example: Different algorithm for intra-warp scan than inter-thread scan
  - Low core count CPU example: based largely on sequential scan
Parallel Segmented Scan
Segmented scan

- Generalization of scan
- Simultaneously perform scans on arbitrary contiguous partitions of input collection

```plaintext
let a = [[1,2],[6],[1,2,3,4]]
let ⊕ = +
segmented_scan_exclusive(⊕, a) = [[0,1], [0], [0,1,3,6]]
```

Assume simple “head-flag” representation:

```plaintext
a = [[1,2,3],[4,5,6,7,8]]
flag: 0 0 0 1 0 0 0 0
data: 1 2 3 4 5 6 7 8
```
Work-efficient segmented scan

Up-sweep:
for d=0 to (log₂n - 1) do
  forall k=0 to n-1 by 2^{d+1} do
    if flag[k + 2^{d+1} - 1] == 0:
      data[k + 2^{d+1} - 1] = data[k + 2^{d} - 1] + data[k + 2^{d+1} - 1]
      flag[k + 2^{d+1} - 1] = flag[k + 2^{d} - 1] || flag[k + 2^{d+1} - 1]

Down-sweep:
data[n-1] = 0
for d=(log₂n - 1) down to 0 do
  forall k=0 to n-1 by 2^{d+1} do
    tmp = data[k + 2^{d} - 1]
    data[k + 2^{d} - 1] = data[k + 2^{d+1} - 1]
    if flag_original[k + 2^{d}] == 1:   // must maintain copy of original flags
      data[k + 2^{d+1} - 1] = 0
    else if flag[k + 2^{d} - 1] == 1:
      data[k + 2^{d+1} - 1] = tmp
    else:
      data[k + 2^{d+1} - 1] = tmp + data[k + 2^{d+1} - 1]
    flag[k + 2^{d} - 1] = 0
Segmented scan
Sparse matrix multiplication example

\[
\begin{bmatrix}
  y_0 \\
  y_1 \\
  y_2 \\
  \vdots \\
  y_{n-1}
\end{bmatrix}
= 
\begin{bmatrix}
  3 & 0 & 1 & \cdots & 0 \\
  0 & 2 & 0 & \cdots & 0 \\
  0 & 0 & 4 & \cdots & 0 \\
  \vdots \\
  0 & 2 & 6 & \cdots & 8
\end{bmatrix}
\begin{bmatrix}
  x_0 \\
  x_1 \\
  x_2 \\
  \vdots \\
  x_{n-1}
\end{bmatrix}
\]

- Most values in matrix are zero
  - Note: logical parallelization is across per-row dot products
  - But different amounts of work per row (complicates wide SIMD execution)

Example sparse storage format: compressed sparse row

values = \([ [3,1], [2], [4], ..., [2,6,8] ] \)

row_starts = \([0, 2, 3, 4, ...] \)

cols = \([ [0,2], [1], [2], ...., ] \)
Sparse matrix multiplication with scan

values = [ [3,1], [2], [4], [2,6,8] ]
cols = [ [0,2], [1], [2], [1,2,3] ]
row_starts = [0, 2, 3, 4]

\[
\begin{bmatrix}
  y_0 \\
  y_1 \\
  y_2 \\
  y_3 \\
\end{bmatrix}
= \begin{bmatrix}
  3 & 0 & 1 & 0 \\
  0 & 2 & 0 & 0 \\
  0 & 0 & 4 & 0 \\
  0 & 2 & 6 & 8 \\
\end{bmatrix}
\begin{bmatrix}
  x_0 \\
  x_1 \\
  x_2 \\
  x_3 \\
\end{bmatrix}
\]

1. Map over all non-zero values: \( \text{products}[i] = \text{values}[i] \times x[\text{cols}[i]] \)
   - \( \text{products} = [3x_0, x_2, 2x_1, 4x_2, 2x_1, 6x_2, 8x_3] \)
2. Create flags vector from row_starts: flags = [1,0,1,1,0,0]
3. **Inclusive** segmented-scan on (multiples, flags) using addition operator
   - \( [3x_0, 3x_0+x_2, 2x_1, 4x_2, 2x_1, 2x_1+6x_2, 2x_1+6x_2+8x_2] \)
4. Take last element in each segment:
   - \( y = [3x_0+x_2, 2x_1, 4x_2, 2x_1+6x_2+8x_2] \)
Scan/segmented scan summary

- **Scan**
  - Parallel implementation of (intuitively sequential application)
  - Theory: parallelism linear in number of elements
  - Practice: exploit locality, use only as much parallelism as necessary to fill the machine
    - Great example of applying different strategies at different levels of the machine

- **Segmented scan**
  - Express computation and operate on irregular data structures (e.g., list of lists) in a regular, data parallel way
Parallel Ray Tracing on SIMD Architectures

(since many students always ask about parallel ray tracing)
Ray tracing

Problem statement:
Given a “ray”, find closest intersection with scene geometry

Simplest ray tracer:
For each image pixel, shoot ray from camera through pixel into scene.
Color pixel according to first surface hit.
Accelerating ray-scene intersection

Preprocess scene to build data structure that accelerates finding “closest” geometry along ray

Idea: group objects with spatial proximity (like quad-tree in Barnes-Hut)

- Hierarchical grouping adapts to non-uniform density of scene objects

Bounding Volume Hierarchy (BVH)
(Binary tree organizing the scene)
Tracing a ray (accelerated using BVH)

// ClosestIsectInfo stores information about closest intersection found so far along a ray
struct ClosestIsectInfo {
    Primitive primitive; // primitive that is the closest intersection
    float distance; // distance along ray to closest intersection
};

void traceRay(Ray ray, BVHNode node, ClosestIsectInfo isectInfo) {
    if (!intersect(ray, node.bbox) ||
        (closest point on box is farther than hitInfo.distance))
        return;

    if (node.leaf) {
        for (each primitive in leaf node) {
            (does_isect, distance) = intersect(ray, primitive);
            if (does_isect && distance < isectInfo.distance) { // check if new isect is closest
                isectInfo.primitive = primitive;
                isectInfo.distance = distance;
            }
        }
    } else {
        traceRay(ray, node.leftChild, hitInfo);
        traceRay(ray, node.rightChild, hitInfo);
    }
}
Decomposition and assignment

- Spatial decomposition of image (2D blocks)
  - 2D blocks maximize spatial locality of rays
- Create many more tiles than processors (just like assignment 1, problem 2)
- Use simple work queue to dynamically assign tiles to processors
  - Cost to render a block is large, therefore synchronization overhead is trivial
Ray tracing using ray packets

Code traces a collection of rays (a “packet”) through BVH at once

RayPacket
{
    Ray rays[PACKET_SIZE];
    bool active[PACKET_SIZE];
};

trace(RayPacket rays, BVHNode node, ClosestHitInfo packetHitInfo)
{
    if (!ANY_ACTIVE_intersect(rays, node.bbox) ||
        (closest point on box is farther than hitInfo.distance for all active rays))
        return;

    update packet active mask to de-activate rays that do not intersect node bbox (these
rays would not have visited children nodes in a non-packet implementation)

    if (node.leaf) {
        for (each primitive in node) {
            for (each ACTIVE ray r in packet) {
                (hit, distance) = intersect(ray, primitive);
                if (hit && distance < hitInfo.distance) {
                    hitInfo[r].primitive = primitive;
                    hitInfo[r].distance = distance;
                }
            }
        }
    } else {
        trace(rays, node.leftChild, hitInfo);
        trace(rays, node.rightChild, hitInfo);
    }
}
Ray tracing using ray packets

Blue = ray is active after node-box test

Note: $r_6$ intersects $F$, but does not pass node $F$ box test due to closest-so-far check.
Advantages of packets

- Enables SIMD execution of tree traversal logic for multiple rays in parallel
  - One vector lane per ray

- Amortize node data fetch (since all rays in packet visit node at same time)
  - Load BVH node once for all rays in packet
  - A reason to make packets much bigger than SIMD width!

- Amortize work
  - Packets are an algorithmic improvement: some computations can be shared across rays in a packet (this help sequential algorithm as well)
  - A reason to make packets much bigger than SIMD width!
Disadvantages of packets

- If any ray must visit a node in BVH, it drags all rays in the packet along with it!

- Implications:
  - Diminished SIMD execution efficiency during traversal due to divergence
  - Node traversal, intersection, etc. are now amortized over less than a full packet’s worth of rays
Ray packet tracing: “incoherent” rays

When rays in a packet are “incoherent” (travel in different directions), the benefit of packets can decrease significantly. In this example: packet visits all tree nodes. (All rays in packet visit all tree nodes)
Improving packet tracing with ray reordering

Idea: when packet utilization drops below threshold, resort rays and continue with smaller packet

- Increases SIMD utilization
- But loses amortization benefits of large packets

Example: 8-wide SIMD processor, 16-ray packets
2 SIMD instructions required to perform operation for all 16 rays in full packet

16-ray packet: 7 of 16 rays active
2 SIMD instructions do meaningful work for only 7 rays

Reorder rays
Recompute intervals/bounds for active rays

Continue tracing with 8-ray packet: 7 of 8 rays active
(Now: 1 SIMD instruction to performs operation for 7 active rays)
Giving up on packets

- Even with reordering, ray coherence during BVH traversal will diminish
  - Little benefit to packets (can even decrease performance compared to single ray code due to overhead of managing packet)

- Idea: exploit SIMD execution for parallel work within single ray-BVH intersection query
  - Interior: use wider-branching BVH (e.g., use degree four rather than binary BVH) (test single ray against multiple node bounding boxes in parallel)
  - Leaf: test single ray against multiple triangles in parallel
Ray tracing best practices on multi-core CPU

- **Trivial assignment of screen regions to CPU cores**
  - Use work queue to dynamically assign regions of screen to cores (some regions may be more expensive than others)

- **Effective use of SIMD execution is harder**
  - Use large packets for higher levels of BVH
    - Ray coherence always high at the top of the tree (visited by all rays)
  - Can switch to single ray (use SIMD for intra-ray parallelism) when packet utilization drops below threshold
  - Can use dynamic reordering to postpone time of switch
    - Reordering allows packets to provide benefit deeper into tree

- **Building a BVH in parallel is a tricky problem**
  - Multiple threads must cooperate to build a shared structure
  - Not addressed here
Summary

- Today we looked at several different parallel programs

- Key questions:
  - What are the dependencies?
  - What synchronization is necessary?
  - How to balance the work?
  - How to exploit locality inherent in the problem?

- Trends
  - Only need enough parallelism to keep all processing elements busy (e.g., data-parallel scan vs. simple multi-core scan)
  - Different parallelization strategies may be applicable under different workloads (packets vs. no packets) or different locations in the machine (different implementations of scan internal and external to warp)