History and Future Projections for MPP Computers

Swiss SPEEDUP Society Meeting

Erik P. DeBenedictis

Agenda

• Past
  – Beginnings of Massively Parallel Computers (MPPs) as machines to solve problems

• Present
  – DOE Scaling Rules

• Future
  – Scaling to a Petaflop
  – Extreme Future
    – Exotic methods needed

Past: 1981

• Application: Quantum Chromodynamics Simulation (QCD) Simulation
  – Simulate 3d physical region by assigning a submesh to each MPP processor

Past: 1981

• Design Rules
  – for n points in each dimension
    – n^3 memory
    – n timesteps
    – n^4 computation
    – 3n^2 communication

• Implementation
  – 50 KFLOPS

CPU

Memory
Scaling QCD

- Story 1
  - Finite Difference Scaling Each Generation
  - Each point replaced by eight points
  - Time step halves
  - 8x memory/16x CPU
  - Memory \( \propto \) FLOPS

Present: Other Applications

- Sandia Applications
  - generally related to simulating 3d objects
  - (that is, not commercial databases or code-breaking)
- Examples
  - Finite Elements
  - Dense/Sparse Matrices
  - Particle-in-Cell
  - Bioinformatics

DOE Balance Criteria

<table>
<thead>
<tr>
<th></th>
<th>SNL (Tomkins)</th>
<th>LLNL (Seager)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IO Bytes/sec/FLOPS</td>
<td>2</td>
<td>.1/12=.01</td>
</tr>
<tr>
<td>IO Bytes/sec/FLOPS total</td>
<td>2</td>
<td>.1</td>
</tr>
<tr>
<td>Memory Bytes/FLOPS</td>
<td>1.0 ( \left( \frac{\text{Perf.}}{\text{TFLOP}} \right)^{25} )</td>
<td>1.0 ( \left( \frac{\text{Perf.}}{\text{TFLOP}} \right)^{25} )</td>
</tr>
<tr>
<td>Memory Bytes/sec/FLOPS</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

Other Balance Criteria

<table>
<thead>
<tr>
<th></th>
<th>SNL (Tomkins)</th>
<th>LLNL (Seager)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum (SMP) Processors</td>
<td>20,000</td>
<td>N/A</td>
</tr>
<tr>
<td>Interconnect Topology</td>
<td>3D Mesh</td>
<td>N/A</td>
</tr>
<tr>
<td>Bisection Bandwidth Bytes/FLOPS</td>
<td>.075</td>
<td>.05</td>
</tr>
<tr>
<td>Floorspace, Power, etc.</td>
<td>Etc.</td>
<td>Etc.</td>
</tr>
</tbody>
</table>
Comparison

• Sandia Approach
  – Have problem
  – Buy or build machine to solve problem

• Beowulf and Grid Approach
  – Use existing hardware as machine
  – See what problems can be solved

Future: Technology Scaling

• Moore’s Law
  – Concise

• SIA Semiconductor Roadmap (right)
  – Detailed
  – Applies to industry, not necessarily Universities and DOE

• Industry Data
  – Covered by NDA
  – Business projections (!)

Future: Petaflops Planner

• Observation
  – Scalability rules are sufficiently concrete to permit basic layout of MPPs in future years

• Result
  – Developed program to “floorplan” MPPs of different designs in future
  – Multiple uses of program can show limits of scalability and trends

Example Input

• Factors
  – Performance target
  – Balance factors
  – Time period
Example Output

- Qualitative Features
  - Trends in layout, power, etc.
- Quantitative Features
  - Chip size, # wires, clock rate, etc.
  - Power
  - Cost

Cost & Pin Driver Speed Story

Apparent Problem
- Failure of Moore’s law due to communications and memory bandwidth

Manifestation
- ALL available pins in use
- Still not enough

External Speed

- Problem
  - Moore’s law failed (temporarily) due to non-scalability of bandwidth across pins
  - Pad density fixed
  - Speed to drive a wire across a PC board fixed by power and s/n ratio

- Solution
  - Dual wire differential transmission line
  - Low voltage
  - Meets internal clock rates

Wiring Density

- Wiring a Problem?
  - LLNL balance factors OK
  - Sandia data-intensive balance factors become unwireable with standard packaging

- Maximum Circuit Board Edge Connector Density
  - About 70 Twisted Pairs Per Inch
- Area Volume Rule
  - 3-d mesh has $3n^3$ connections but only $6n^2$ surface
  - Doesn’t solve problem
**Shish Kabob Packaging**

- Hand crank??
- Water-cooled Base
- X Wiring on Base
- Z Wiring via Flex PC Board
- Y Wiring via Flex PC Board

**Power**

- Problem: Projections Show Microprocessor Die Dissipation Growing to 1 KW and Beyond
  - Not so for PIMs
- Potential Solutions
  - Power saving circuit design
  - Heat-capable packaging
  - Cut performance estimates
  - Ignore problem

**Power Dilemma**

- View 1
  - Microprocessors have been increasing in performance by 100%/year recently and the industry will make sure the trend continues
- View 2
  - Microprocessor power consumption has been growing exponentially, and will exceed packaging capability in the next couple generations

**Technology Becoming More Exotic**

- Dimensionality of Network Ought To Fit Dimensionality of Universe
  - 3d mesh
- Wires Become Transmission Lines
  - Dual conductor
  - 30 GHz
  - Very expensive
- Processor Pipelining Becomes Too Deep
  - hundreds of pipeline stages
  - affects commercial processors as well
- Power Dissipation Will Affect Architecture
  - used to just “calculate heat sink size”
Conclusions & Far Future

• Sandia-Balanced MPPs Will Scale to 1-10 Petaflops
  – Same breadth of applications
  – Same programming methods
  – CMOS technology

• Uncertainty Beyond 10 Petaflops
  – Wiped out by non-scalability of the “speed of light”
  – New technology
    • optical computing
    • quantum
    • biological
  – New prog. methods
    • neural networks