Home-made High Performance Computing on Mac OS X

Ken Tabb
Neural Systems Group, University of Hertfordshire
Outline

• Problem domain
• Vectorising algorithms
• Mac OS X Performance Analysis Tools
• Neural Networks
• Clusters
Problem domain

- Detecting and tracking moving humans
- Computer Vision
- Neural Networks
AltıVec

Scalar maths is for wimps
What is AltiVec?

- a.k.a. Velocity Engine
- SIMD (Single Instruction, Multiple Data) co-processor
- Allows for mass computation
- 128-bit vector unit
- 1 AltiVec unit per G4 / G5 processor
- Instruction set available from C or Fortran
Edge detection
Edge detection
Edge detection
Edge detection

Max\[
\begin{bmatrix}
+1 & +1 & 0 \\
+1 & 0 & -1 \\
0 & -1 & -1 \\
\end{bmatrix}
, \begin{bmatrix}
0 & +1 & +1 \\
-1 & 0 & +1 \\
-1 & -1 & 0 \\
\end{bmatrix}
\]
Demo:
Vec-tastic
(Image processing with AltiVec)

UH
Is that as fast as it gets?

• No!

• Some maths was redundant

\[
\max \begin{pmatrix}
+1 & +1 & 0 \\
+1 & 0 & -1 \\
0 & -1 & -1 \\
\end{pmatrix}, \begin{pmatrix}
0 & +1 & +1 \\
-1 & 0 & +1 \\
-1 & -1 & 0 \\
\end{pmatrix}
\]

• Darwin load balances threads across idle CPUs
How do you know it’s optimised?

- You don’t
- Vectorising can be easy to apply, but difficult to optimise
- The obvious vector algorithm isn’t always the best
- You have to use performance diagnostic tools
Mac OS X
Performance Analysis tools
Xcode

- Based on GCC tools
- Zero-link
- Fix and continue
- Predictive compile
- Distributed build
<table>
<thead>
<tr>
<th>Tool</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>QuartzDebug</td>
<td>Window refreshing</td>
</tr>
<tr>
<td>OpenGL Profiler</td>
<td>Monitor OpenGL calls</td>
</tr>
<tr>
<td>MallocDebug</td>
<td>Locate memory leaks</td>
</tr>
<tr>
<td>ObjectAlloc</td>
<td>Monitor object allocation</td>
</tr>
<tr>
<td>Thread Viewer</td>
<td>Identify deadlocks / waits</td>
</tr>
<tr>
<td>Spin Control</td>
<td>Identify waits / GUI locks</td>
</tr>
<tr>
<td>Sampler</td>
<td>Monitor performance</td>
</tr>
<tr>
<td>Shark</td>
<td>Monitor performance</td>
</tr>
</tbody>
</table>
CHUD.framework

- Performance analysis API
- Allows integrated reports
- Highlights / logs relevant measurements
Neural Networks
A crash course
Neural Networks

- Sun Microsystems c1997: “The network is the computer”
Neural Networks

- Sun Microsystems c1997: “The network is the computer”
- Brain is full of processors (neurons)
- Average human brain:
  - $1,000,000,000,000$ ($10^{12}$) neurons
  - $1,000,000,000,000,000$ ($10^{15}$) neural connections
- A single neuron may have 1,000s or 10,000s of connections
Neural Networks

- Sun Microsystems c1997: “The network is the computer”
- Brain is full of processors (neurons)
- Average human brain:
  - 1,000,000,000,000 (10^{12}) neurons
  - 1,000,000,000,000,000 (10^{15}) neural connections
- A single neuron may have 1,000s or 10,000s of connections
Neural Networks
Neural Networks
Neural Networks
Neural Networks
Neural Networks

Input

Hidden

Output
Neural Networks

Input

Hidden

Output
Neural Networks

- AltiVec can be used for mass-calculation of inter-neuron signals
- There is a need to develop multiple networks at once
- This is going to take one computer a very long time
- What is a hard-up researcher to do?
What is Clustering?

- Some computing jobs are easy for today’s CPUs
- Some jobs aren’t so easy
Some computing jobs are easy for today’s CPUs

Some jobs aren’t so easy

Many hands make light work

What is Clustering?
Computational Clustering

- MPI (Message Passing Interface) used for communication between nodes
- Ethernet provides physical connection
- Code has to be re-written to use MPI
- MPI is platform-agnostic
  - MPI code re-compiles as is
- Mac OS X nodes integrate with other UNIX nodes
- Rendezvous used at UH for service discovery of nodes
Summary

• Vectorising algorithms can significantly increase computational speed
• Ethernet adds an unknown variable
• Algorithms may need customisation accordingly
• Rendezvous can be used for node discovery
• CHUD provides highly customisable performance analysis API
And finally...

Xgrid

Apple’s solution for parallel and distributed high performance computing
Q & A