Ateneo de Manila University

Building the AGILA HPCS

Ateneo High Performance Computing Group
11 December 2000
http://www.math.admu.edu.ph/ahpc/
william.s.yu@ieee.org
Section I

Introduction
Is It Primetime for Parallel Computing?

★ increase in performance of mass-market commodity microprocessors
★ declining costs of high speed network hardware
★ scientific computing applications that can benefit from clustering:
   ★ Artificial Intelligence Models
   ★ Weather and Climate Modelling
   ★ Computational Fluid Dynamics
★ commercial applications that can benefit from clustering:
   ★ Email Systems
   ★ Proxy Systems
   ★ Web Services
What Makes an Application Parallelizable?

★ Is the problem inherently parallel?
  ★ Cellular Automata
  ★ Neural Networks
  ★ Genetic Algorithms

★ Is the Problem Data-Parallel?

★ What is the granularity of the problem?
  ★ Fine Grain - no way!
  ★ Medium Grain
  ★ Coarse Grain
Amdahl’s Law

\[
\frac{R(P)}{R(1)} = \frac{(T_s + T_p)}{(T_s + (P \times T_{is}) + \left(\frac{T_p}{P}\right) + T_{ip})}
\]

★ Amdahl’s Law strictly limits the amount of speedup

★ serial components

★ communications overhead

★ optimistic estimation
Nature of the Problem: Some Questions

★ Is it compute intensive or communications intensive?

★ How much memory does the application need?

★ Does it benefit from the presence of a local disk?

★ How big should the cluster be?
Section II

Design
## Hardware Components for the AGILA HPCS

<table>
<thead>
<tr>
<th>ITEM</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node System: CPUs(Athlon 650MHz w/ 128 MB SDRAM, 6.4 GB HD, Intel EEPro100 NIC)</td>
<td>7</td>
</tr>
<tr>
<td>Server System: (Athlon 650MHz w/ 256MB SDRAM, 10.4 GB HD, (2) Intel EEPro100 NIC Monitor, Trident 3D AGP, 1.44 Floppy, Keyboard</td>
<td>1</td>
</tr>
<tr>
<td>Intel 410T Switch</td>
<td>1 x 24 port</td>
</tr>
<tr>
<td>APC Smart UPS 3000VA</td>
<td>1 (20 CPUs)</td>
</tr>
<tr>
<td>Cabling and others</td>
<td>1 box and 1 pack</td>
</tr>
</tbody>
</table>

Table 1: Hardware Components of the AGILA HPCS
AGILA HPCS Design
AGILA HPCS Software Setup

<table>
<thead>
<tr>
<th>Software Used</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Linux Distribution</td>
<td>Redhat Linux 6.2 (Zoot)</td>
</tr>
<tr>
<td>Linux Kernel</td>
<td>Linux Kernel v. 2.2.17</td>
</tr>
<tr>
<td>Compilers</td>
<td>GCC</td>
</tr>
<tr>
<td>Message Passing Libraries</td>
<td>PVM 3.4.3, MPI-MPICH 1.2.1, MPI-LAM 6.3.3b48</td>
</tr>
<tr>
<td>Scientific Libraries</td>
<td>SCALAPACK, FFTW, PETSc</td>
</tr>
</tbody>
</table>

Table 2: AGILA HPCS Software Setup
## Proposed Additional 8-nodes

<table>
<thead>
<tr>
<th>ITEM</th>
<th>COST (pesos)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node System:</td>
<td></td>
</tr>
<tr>
<td>CPUs (Athlon 750MHz</td>
<td>240,000</td>
</tr>
<tr>
<td>w/ 128 MB SDRAM,</td>
<td></td>
</tr>
<tr>
<td>10.4 GB HD,</td>
<td></td>
</tr>
<tr>
<td>Intel EEPro100 NIC)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>240,000</td>
</tr>
</tbody>
</table>

Table 3: Proposed Additional 8-nodes
Conclusion

★ not all problems are parallelizable
★ clusters are built to solve problems
★ Beowulfery is not simply about putting computers together
★ a well planned beowulf is a good thing (TM)