From 20 Years of Beowulf
MPI & Cluster Computing

Aiichiro Nakano

Collaboratory for Advanced Computing & Simulations
Department of Computer Science
Department of Physics & Astronomy
Department of Chemical Engineering & Materials Science
Department of Biological Sciences
University of Southern California

Email: anakano@usc.edu

http://dl.acm.org/citation.cfm?id=2737911&CFID=717321656&CFTOKEN=85686577
http://www.crest.iu.edu/beowulf14
Beowulf Cluster

- Parallel computer made of commodity components
- Open software (Linux; network driver developed)
- Started by Thomas Sterling & Don Becker in ’94

(Left) Caltech’s Thomas Sterling (left) and John Salmon discuss the building of Naegling, one of the largest Beowulf computers. Each of its 120 processors performs 70 million floating-point operations per second (flops) across a range of applications. (Right) Don Becker, Goddard Space Flight Center, holds a personal computer network adapter. He has developed, and is constantly updating, software to drive nearly all adapters for use with the Linux operating system.

http://www.hq.nasa.gov/hpcc/insights/vol7/beowulf.htm
Digress: LIGO

- Early 70’s: Bill Hamilton started building a cryogenic gravitational-wave detector at LSU
- Mid 90’s: LIGO construction started

Gravitational-wave research in Louisiana

The Nobel Prize in Physics 2017 was divided, one half awarded to Rainer Weiss, the other half jointly to Barry C. Barish and Kip S. Thorne "for decisive contributions to the LIGO detector and the observation of gravitational waves".
20 Years of Beowulf and MPI

- Dominant parallel-computing paradigm for the past 20 years: Distributed processes communicating via message passing

- First Beowulf (’94)
- Evolving by embracing multicore & accelerators per computing node

- MPI 1 (’94)
- Evolving: Using Advanced MPI
Counter-Approach: Anton 2

- Unified on-chip & inter-node networks
- New algorithm: Gaussian series expansion of the Coulombic interaction (no FFT required)

D. E. Shaw et al., Hot Chips 2014
Massive Data Analytics

- Scalable data-analytics/machine-learning algorithms are critically needed, e.g., $O(N^2) \rightarrow O(N)$ pair statistics
- Seven computational giants
  1. Basic statistics
  2. Generalized $N$-body problem
  3. Graph-theoretic computations
  4. Linear algebraic computations
  5. Optimization
  6. Integration
  7. Alignment problems

National Research Council, Frontiers of Massive Data Analytics ('13)
https://www.nap.edu/catalog/18374/frontiers-in-massive-data-analysis
\(O(N^3)\) Linpack to \(O(N)\) HPCG

- High performance conjugate gradient (HPCG) proposed toward exaflop/s, but ...

<table>
<thead>
<tr>
<th>Site</th>
<th>Computer</th>
<th>Cores</th>
<th>HPL Rmax (Pflops)</th>
<th>HPL Rank</th>
<th>HPCG (Pflops)</th>
<th>HPCG/HPL</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSCC / Guangzhou</td>
<td>Tianhe-2 NUDT, Xeon 12C 2.2GHz + Intel Xeon Phi 57C + Custom</td>
<td>3,120,000</td>
<td>33.9</td>
<td>1</td>
<td>.580</td>
<td>1.7%</td>
</tr>
<tr>
<td>RIKEN Advanced Inst for Comp Sci</td>
<td>K computer Fujitsu SPARC64 VIIIfx 8C + Custom</td>
<td>705,024</td>
<td>10.5</td>
<td>4</td>
<td>.427</td>
<td>4.1%</td>
</tr>
<tr>
<td>DOE/OS Oak Ridge Nat Lab</td>
<td>Titan, Cray XK7 AMD 16C + Nvidia Kepler GPU 14C + Custom</td>
<td>560,640</td>
<td>17.6</td>
<td>2</td>
<td>.322</td>
<td>1.8%</td>
</tr>
<tr>
<td>DOE/OS Argonne Nat Lab</td>
<td>Mira BlueGene/Q, Power BQC 16C 1.60GHz + Custom</td>
<td>786,432</td>
<td>8.59</td>
<td>5</td>
<td>.101#</td>
<td>1.2%</td>
</tr>
<tr>
<td>Swiss CSCS</td>
<td>Piz Daint, Cray XC30, Xeon 8C + Nvidia Kepler 14C + Custom</td>
<td>115,984</td>
<td>6.27</td>
<td>6</td>
<td>.099</td>
<td>1.6%</td>
</tr>
</tbody>
</table>

J. Dongarra et al., ISC14
Smartphones as Exascale Nodes

- Building an exaflop/s computer from commodity components (again, but with smartphones this time?)

**Blades of Glory:** Mont-Blanc’s prototype contains 15 nodes made up of ARM-core processors.

*IEEE Spectrum* (May ’14)

Or Raspberry Pi?