Exploring Scalable Systems

Eroeffnung des CoE for HyperTransport an der Uni Mannheim

Dr. John R. Busch
Research Director
Computer Architecture and Analysis
Sun Microsystems Research Laboratories
February 14, 2007
Sun Labs Research Strategy

• Applied research aligned with Sun's business
• High risk / high return component in a distributed innovation company
• Innovate, demonstrate, transfer
  • Top talent from multiple sources
  • Small teams to prototype innovative technologies (e.g. Java), but not the only source of innovation at Sun
  • Transfer our knowledge, prototypes, people and projects into Sun products and processes, open-source and standards initiatives
In Search of Excellence
Lessons from America's Best-Run Companies

Thomas J. Peters and Robert H. Waterman Jr.

• Harper and Row, 1982
Best practices synthesized + correlated
Recipe

- Managing ambiguity and paradox
- A bias for action
- Close to the customer
- Autonomy and entrepreneurship
- Productivity through people
- Hands-on, value-driven
- Stick to the knitting
- Simple form, lean staff
Status of the 43 "Excellent" Companies Five Years Later*

1. Allen Bradley (Rockwell)
2. Disney
3. Boeing
4. DEC
5. Emerson
6. Frito-Lay
7. IBM
8. Intel
9. Johnson & Johnson
10. Mars
11. Maytag
12. McDonald's
13. Merck
14. Wal-Mart

1. Bristol-Myers
2. Delta
3. Dow
4. DuPont
5. Hughes (GM)
6. Levi Strauss
7. Marriott
8. Procter & Gamble
9. Standard Oil Amoco
10. 3M

1. Amdahl
2. Bechtel
3. Caterpillar
4. Dana
5. Hewlett-Packard
6. Kodak
7. Raychem
8. Schlumberger
9. Texas Instruments
10. Tupperware (Dart)
11. Wang
12. Atari
13. Chesapeake-Pond's
14. Avon
15. Data General
16. Fluor
17. Kmart
18. National Semiconductor
19. Revlon

* Percent of Total

Excellent
Solid but Loss of Leadership
Weakened Position
Troubled
THE INNOVATOR'S SOLUTION
Creating and Sustaining Successful Growth

Clayton M. Christensen
Author of The Innovator's Dilemma
Michael E. Raynor

The Third Dimension of the Disruptive Innovation Model

**Sustaining Strategy**
- Bring a better product into an established market

**Low-End Disruption**
- Address underserved customers with a lower-cost business model

**New-Market Disruption**
- Compete against nonconsumption

- **Performance**
- **Time**
- **Different Measure of Performance**
- **Nonconsumers or Nonconsuming Occasions**
Innovation Imperative

• Sustaining profitable growth
  • <10% established companies do it
  • Stall → hard to recover

• Causality: radical sustaining innovation drives growth
  • Differentiate: starting a new trajectory
  • Along metrics of key value to customers
  • Key corporate asset

• Key Elements: Resources, Processes, Values
Who Says Elephants Can't Dance?

National Bestseller

Learn Gerstner's Management Secrets

A *BusinessWeek* Best Book of 2002

Inside IBM's Historic Turnaround

Louis V. Gerstner, Jr.

IBM Stall in 1990s

• Gerstner: Key success factors
  • Expected
    • Vision, strategy, focus, marketing, financial
  • Reality
    • Product definition process
      – Building the wrong things
  • Culture
    – “NO!”, NIH
Experimentation Matters

Unlocking the Potential of New Technologies for Innovation

STEFAN H. THOMKE

Designers Combining Manual Renderings, CAS, and Clay Models

(a) Designer preparing hand sketches for brainstorming design concepts.
(b) Craftsman working on clay model of concept car.
(c) Designers using computer-aided styling (CAS) for group brainstorming.
Last 10 Years @ BMW

- Concept -> deliver: cut by 50% (7->3 years)
  - Executive mandate (after incubated)
  - Improved quality, performance, customer acceptance

- Enablers
  - Integrated teams
  - Computerization of car design and crash test

- Increased parallelism through front-loaded processes
  - Computer simulations iterate and test ideas rapidly
  - More experimentation but fewer prototypes
  - Required changing sequential culture
  - Integrating new and traditional technologies
THE STRUCTURE OF SCIENTIFIC REVOLUTIONS

THOMAS S. KUHN

THIRD EDITION
“Paradigm Shifts”

• The scientific revolutions in history were driven by new instruments, coupled with models and experimentation
  – Telescope
  – Microscope
  – Particle accelerator
  – ...

Computer System Industry

• Challenge: Radical sustaining system solution innovation
  • Significant competitive improvements
  • Timely
  • Can't rely on functional design/prototype/merge cycle
    – Limits design space, takes too long

• Opportunity: explore alternative system architectures through workload characterization and modeling
  • Before committing to product development cycle
System Exploration Thru Architecture Analysis

- Technology Trends
- System Modeling
- Detailed Workload Characterization
- System Architecture Exploration (chip, system, o/s, apps, storage, network)
- System-wide optimized design
- OS and Apps Implementation
- System Implementation
- Chip Implementation
Benefits of Methodology

• Detailed component models and high-level system models
• Rapid turn around
• Explore design space
• Early refinement
• Trade-off analysis
• Analysis of competitive alternatives
• Drive design decisions on current and future instances
## Exploiting Workload Characteristics and Technology Trends for Customer Values

<table>
<thead>
<tr>
<th>Market</th>
<th>Workload Characteristics</th>
<th>Seed Ideas Exploiting Technology Trends</th>
<th>Key Customer Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMT Node</td>
<td>High thread level parallelism, low instruction level parallelism, high data sharing, very sensitive to latency</td>
<td>Lots of simple cores integrated onto a chip with shared L2 + mem ctrls; high b/w chip to chip connection for smp expansion</td>
<td>Price/Performance, Availability</td>
</tr>
<tr>
<td>Internet Application Servers</td>
<td>Shared session state, sensitive to interconnect latency, low network bandwidth, moderate end-to-end recovery</td>
<td>CMT nodes using low latency interconnect, switch-based failure recovery and load balancing</td>
<td>Scalable cluster performance, TCO, QOS session recovery</td>
</tr>
<tr>
<td>Shared Memory Scaled Clusters</td>
<td>Data and processing for single problem scales to many machines, very sensitive to interconnect latency and bandwidth</td>
<td>Low latency, high bandwidth interconnect fabric + low cost CMP nodes</td>
<td>Productivity, Performance, Cost for HPC and Commercial Apps</td>
</tr>
</tbody>
</table>

*Note: different system markets have different workload characteristics, customer value metrics, and system architecture and technology opportunities*
Scalable Coherent CMT Exploration
Example 1999/2000
2006: Sun's CMT Niagara: SunFire T1000/T2000

• 5X throughput in 1/4 space and power
• Fastest ever Sun product ramp

"...the machines put Sun at the cutting edge of one of the chip industry's biggest trends... multi-core systems."

Sun Fire T2000: Most Innovative Server of 2005
"Changes the math for threaded performance and power consumption"
Architecture Trade-off Studies: Example Alternative Application Server Architectures

Lower is better

Modeled what BCI would measure.

Trade2 – 200 users, 1600 TPS
Design Study: Example
Session Size Variation
CPU ms/ trans

Network Activity: Orig Arch
Network Activity: Alternate Arch
Shared Memory Programmability Benefit

Reduction Ratio Versus Benchmark

<table>
<thead>
<tr>
<th>Benchmark</th>
<th>Ratio (Orig NCSL/High Prog NCSL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAB</td>
<td>5</td>
</tr>
<tr>
<td>SSCA2</td>
<td>3</td>
</tr>
<tr>
<td>GTC</td>
<td>3</td>
</tr>
<tr>
<td>CG</td>
<td>10</td>
</tr>
<tr>
<td>BT</td>
<td>7</td>
</tr>
<tr>
<td>MG</td>
<td>11</td>
</tr>
<tr>
<td>sPPM</td>
<td>9</td>
</tr>
</tbody>
</table>
Cost: cluster vs SMP

Cost vs Node Count graph showing the cost comparison between cluster and SMP systems. The graph indicates that the cost of the cluster system increases linearly with the node count, while the SMP system has a higher and more rapid increase in cost at lower node counts.
Sun Labs + Uni Mannheim CAG Collaboration
Research Platform for Clustered Systems and Applications
Architectural Analysis Methodology
Global Shared Memory Systems
Clustered Application Modelling
Example : GTC

System Level Projections: Scalability and Performance
Takeaways: Role of System Analysis in Innovation and Computer System Architecture

• Key ingredient of radical innovation and growth
  – Enabler = leading edge workload characterization and modeling capabilities

• Integral to collaborative exploration
  – Cross-polinate, up-front, experiment and learn

• Provides critical context:
  – Components, systems, value metrics, competitiveness

• On-going leverage of new analysis capabilities
  – to subsequent design, capacity planning, operational deployment phases
Beyond HT 3.0: Collaborative Exploration

• Concepts
  • Topologies
  • Adaptive Routing
  • Congestion Management
  • Fault Tolerance
  • QOS

• Exploration
  • Analyze strawmen for scalable processor to processor interconnect fabric
    • Component and System Level
    • Limitations and Opportunities
    • Competitive Context
  • Integral to our learning, experimentation, innovation
Publications


Publications


• I. Gluhovsky and J. Bonebakker. "Defining Relevant Distances Between Server Workloads", submitted to ACM Transactions on Computer Systems
System Architecture Exploration (chip, system, s/w, network, storage)

Technology Trends
System Modeling
Detailed Workload Characterization

Build The Right Stuff