**Rapid Increase in Modern Web Services**

- Web services
- Microservices

**Stringent SLOs + Moore’s law decline**

**Customized hardware**

Rapid increase in μservices -> greater need for custom hardware

---

**Are Custom Platforms Always Needed?**

- Customized platforms -> expensive
- Data centers prefer hardware resource fungibility
- Low testing overhead

Dire need for limited CPU SKUs that support a variety of μservices

---

**Performance of Commodity Servers**

- Use commodity hardware for procurement efficiency & scalability?

Key FB μservices occupy a large portion of the data center

---

**Contributions**

- Comprehensive characterization of Facebook’s microservices
  - System-level & architectural bottlenecks
  - Reveals enormous bottleneck diversity across microservices
- Concept of “soft” server SKUs
  - Tuning coarse-grained OS & hardware configuration knobs
- μSKU
  - Automates soft-SKU search & configuration via production A/B tests
  - Deploys soft SKUs on production microservices

**7.2% perf. boost on production μservices + no extra hardware**

---

**Soft SKU: Optimizing Server Architectures for Microservice Diversity @Scale**

Akshitha Sriraman**, Abhishek Dhanotia*, Thomas F. Wenisch**

University of Michigan**, Facebook**

---

**Facebook μServices’ Characterization**

<table>
<thead>
<tr>
<th>μService</th>
<th>Throughput (QPS)</th>
<th>Response latency</th>
<th>Pathlength</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web</td>
<td>O(100)</td>
<td>O(ms)</td>
<td>O(10^4)</td>
</tr>
<tr>
<td>Feed1</td>
<td>O(1000)</td>
<td>O(ms)</td>
<td>O(10^4)</td>
</tr>
<tr>
<td>Feed2</td>
<td>O(10)</td>
<td>O(s)</td>
<td>O(10^4)</td>
</tr>
<tr>
<td>Ads1</td>
<td>O(10)</td>
<td>O(ms)</td>
<td>O(10^4)</td>
</tr>
<tr>
<td>Ads2</td>
<td>O(100)</td>
<td>O(ms)</td>
<td>O(10^4)</td>
</tr>
<tr>
<td>Cache1</td>
<td>O(100K)</td>
<td>O(µs)</td>
<td>O(10^4)</td>
</tr>
<tr>
<td>Cache2</td>
<td>O(100K)</td>
<td>O(µs)</td>
<td>O(10^4)</td>
</tr>
</tbody>
</table>

Great diversity in bottlenecks  Use custom SKUs?  Prohibitively expensive

Can we achieve perf. efficiency without building custom SKUs?

---

**“Soft” SKUs: Best of Both Worlds**

- Tune coarse HW & OS knobs on commodity HW
- Performance efficiency
- Procurement efficiency & scalability

**μSKU: Soft SKU Design & Deployment**

---

**Soft SKU Performance**

Soft SKU can achieve ~7.2% throughput improvement on production systems with no extra hardware requirement