The Machine Changes …
Software Stacks

Omer Barkol
HP Labs
June 1, 2015
Forward-looking statements

This is a rolling (up to three year) Roadmap and is subject to change without notice.

This document contains forward looking statements regarding future operations, product development, product capabilities and availability dates. This information is subject to substantial uncertainties and is subject to change at any time without prior notification. Statements contained in this document concerning these matters only reflect Hewlett Packard's predictions and / or expectations as of the date of this document and actual results and future plans of Hewlett-Packard may differ significantly as a result of, among other things, changes in product strategy resulting from technological, internal corporate, market and other changes. This is not a commitment to deliver any material, code or functionality and should not be relied upon in making purchasing decisions.
What are the questions you can’t ask today?

Magnitude of the data

Flash flooding of legacy data bases

Unable to secure

Time consuming data integration

Real-time insight needed

Insufficient resources

The end of cheap hardware
The Past 60 Years

1950s

1960s

1970s

1980s

1990s

2000s

Today
The Machine
Special purpose cores

Photonics

Massive memory pool

The Machine
Memristor for NVRAM

Massive memory pool
Imagine if a computer ran at human speed …

- **Processor cycle**: 1 second

Time to retrieve a byte from …

- **SRAM**: 5 seconds
- **DRAM**: 2 minutes
- **Flash**: 1 day
- **Hard drive**: 2 months
- **Tape**: 1,000 years!

Massive memory pool

**Computer**

- **CPU**:
  - CPU registers
  - Level 1 cache
  - Level 2 cache
  - Level 3 cache
- **Main memory**
- **Flash accelerator**
- **SSD**
- **Local disk**
- **Network drive array**
- **Network backup**
- **Archive**

**Network**

- **Universal memory**
- **Network backup**
- **Volatile**
- **Non-volatile**

* actually an entire computer system with its own hierarchy
Data

Machine OS

Million-node management

Exabyte-scale algorithms

Analytics and visualization

Ultra-efficient hardware

Security built-in from silicon upwards
Implications of The Machine on programming

- Programming Simplicity
- New Opportunities & Challenges
- Performance
Opportunities & Challenges: Memory Management

• Very large shared NVM for both memory and storage
• Fast NVM load/store access thru photonic memory interconnect (<1µs latency)
• Many hardware threads per SoC
• Significant amount of fast local DRAM
• Persistent memory, but volatile caches with minimal ISA support for persistence
• No global cache coherence
Memory-Centric OS

- "Shared memory" with no cache coherency
- Accessible persistent storage through standard load/store path
- Heterogeneous and decentralized compute

Shared nothing

Shared something

Shared everything
Opportunities & Challenges: Security

PROTECT

DETECT

RECOVER
Programming Simplicity: No-IO

tag1 = sanitize(tag);
query = format("SELECT * FROM posts WHERE tag=%s ORDER BY date LIMIT %d", tag1, n);
if (cached(query)) {
    rows = deserialize(readFromCache(query));
} else {
    rows = readFromDB(query);
    cache(serialize(rows));
}
results = new list();
for (row in rows) {
    post = readPostFromRow(row)
    post.setComments(getPostComments(post.id));
    results.add(post);
}
return results;

return nameToTag.get(tag).getPosts().sort(byDate).sublist(0, n);
## No-IO Qualitas Corpus

100 major open-source Java applications

### Parse, extract, tag, collect

### Table: Issue, % out of projects, % effected methods, Notes

<table>
<thead>
<tr>
<th>Issue</th>
<th>% out of projects</th>
<th>% effected methods</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Storage</td>
<td>Most</td>
<td>30-40%</td>
<td>Mainly XML files and SQL stores</td>
</tr>
<tr>
<td>User Interface</td>
<td>Most</td>
<td>30% (MVC and web)</td>
<td></td>
</tr>
<tr>
<td>Exception Handling</td>
<td>60%</td>
<td>20%</td>
<td>Mostly I/O</td>
</tr>
<tr>
<td>Multithreading</td>
<td>50%</td>
<td>20%</td>
<td>Synchronization x10 than inter-thread communication rich concurrency package is rare</td>
</tr>
<tr>
<td>String processing</td>
<td>40%</td>
<td>15%</td>
<td>Regular expressions are commonly</td>
</tr>
</tbody>
</table>
Implications of The Machine on Analytics

Democratize Advanced Analytics

New Opportunities

Support State-of-the-Art

Performance
Examples for Applications

- Graph analytics
- Search space optimization
- Deep learning
- DMC …
DMC – Distributed Mesh Computing
“Adapting old programs to fit new machines usually means adapting new machines to behave like old ones.”

Alan J. Perlis, 1982
Thank you

Contact: omer.barkol@hp.com