Analytics and Storage Class Memory

Doug Voigt
May 1, 2913
Hadoop vs. Massively Parallel Processing DB

**Hadoop**
- Schema-less Model
- Human query optimization
- Ability to create complex data flows with multiple inputs and outputs
- Parallelize many analytic functions
- Adept at unstructured classification and feature extraction

**MPP**
- Performance
- Machine Query Optimization
- Mature workload management
- High concurrency interactive query processing
Today

- Hadoop
  - Server
  - Storage

- MPP
  - Server
  - Storage

2020

- Hadoop and MPP
  - Server
  - SCM
  - Storage

- New Storage Class Memory (SCM) Technology
Storage Class Memory

**Computer memory with:**
- Access attributes similar to RAM (latency, bandwidth, addressability, etc)
- Cost/bit less than DRAM
- Data retention even when not powered.

<table>
<thead>
<tr>
<th>Medium</th>
<th>Capacity</th>
<th>Performance</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRAM</td>
<td>Low</td>
<td>Very High</td>
<td>$$$$</td>
</tr>
<tr>
<td>SCM</td>
<td>Low</td>
<td>Very High</td>
<td>$$</td>
</tr>
<tr>
<td>NAND</td>
<td>Medium</td>
<td>High</td>
<td>$$</td>
</tr>
<tr>
<td>Disk</td>
<td>High</td>
<td>Medium</td>
<td>$$</td>
</tr>
<tr>
<td>Tape</td>
<td>Very High</td>
<td>Low</td>
<td>$</td>
</tr>
</tbody>
</table>

**Opportunity to:**
- Fundamentally change software and hardware paradigms
- Define the next server & storage market leaders
- Productize high value high margin products
Data Access Revolution

**Opportunities:**
- Virtually free durability
  - No interface translation
  - Low write latencies
- Reuse, share durable data

**Challenges:**
- How do we keep persistent data consistent?
- What’s the programming complexity?

**New language and library constructs**

**New programming models for memory speed persistence**

**Extended API’s aligned with existing applications**

Models requiring more flexibility:

- General programming
- Persistent data structures
- Logging
- HPC-style checkpointing
- Database
- OS (e.g. filesystems)
- Fast block device

*NVM in main memory will drive major renovation of the entire storage and memory stack*
Thank You