MittOS: Supporting Millisecond Tail Tolerance with Fast Rejecting SLO-Aware OS Interface

Mingzhe Hao, Huaicheng Li, Michael Hao Tong, Chrisma Pakha, Riza O. Suminto, Cesar A. Stuardo, Andrew A. Chien, and Haryadi S. Gunawi

No Millisecond TT (Tail Tolerance)

Nowadays low and stable latency is a critical key to success of many services. Unfortunately, most NoSQL systems serve requests with millisecond-level SLOs, but none is tail tolerant at this granularity.

Ineffectiveness of Current TT Methods

<table>
<thead>
<tr>
<th>Def. TT</th>
<th>TO Val.</th>
<th>Fail-over</th>
<th>Clone</th>
<th>Hedged/ Tied</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cassandra</td>
<td>×</td>
<td>12s</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td>Couchbase</td>
<td>×</td>
<td>75s</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td>HBase</td>
<td>×</td>
<td>60s</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>MongoDB</td>
<td>×</td>
<td>30s</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td>Riak</td>
<td>×</td>
<td>10s</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td>Voldemort</td>
<td>×</td>
<td>5s</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table: Tail tolerance in NoSQL.

MittOS provides operating system support that helps data-parallel applications cut millisecond-level tail latencies.

- Accurately predicts the latency of an IO based on white-box knowledge of resource managements
- Promptly returns EBUSY when IO SLO cannot be met
- Allows the application to failover to less-busy node without waiting

MittOS’ Principle & Use-Case

We build MittOS within the storage stack:
- Disk: MittNOOP (noop scheduler) + MittCFQ (CFQ scheduler)
- SSD: MittSSD (Open-Channel SSD)
- Cache: MittCache (OS Cache)

MittOS’ no-wait approach helps reduce IO completion time up to 35% compared to existing approaches.

Future Work

- Automatic adoption of storage devices via ML/DL techniques
- Incorporating settings of certainty/confidence for SLO
- Providing hints for applications to setup appropriate SLO deadline
- Extension of MittOS’ principle to CPU, VM and runtime memory management, SMR drives, etc.