Optimistic Replicated Two-Phase Commit (OR-2PC) is a new protocol for distributed transactional storage systems.

Conventional storage systems use:

- **Paxos/VR**
- **Concurrency Control**
- **Two-Phase Commit**

OR-2PC integrates these into a single protocol, taking advantage of cross-layer optimizations.

### The Protocol

1. Client sends **READ** requests to any replica in the partition responsible for the object and buffers all writes.
2. Partition replica returns its latest version of the object for **READ** requests.
3. Client selects a proposed timestamp for the transaction and sends **PREPARE** with the proposed timestamp to all participant replicas.
4. Participant replicas validate the transaction at the proposed timestamp and respond **PREPARE-OK** if validation succeeds.
5. Once \( f+1 \) replicas in each participant partition respond **PREPARE-OK**, the client sends **COMMIT** messages.
6. Participant replicas commit the transaction at the proposed timestamp.

### Evaluation

Compared to conventional systems, OR-2PC provides:

- **2x** improvement in latency
- **3x** improvement in throughput

**TXN-OCC** = 2PC + OCC + VR  
**TXN-LOCK** = 2PC + S2PL + VR  
**SPAN-OCC** = Spanner protocol + OCC  
**SPAN-LOCK** = Spanner protocol (S2PL)  
**TAPIR** = OR-2PC  
**QWSTORE** = Quorum Writes (Dynamo)

**Average Transaction Latency (microseconds)**

<table>
<thead>
<tr>
<th>System</th>
<th>TXN-OCC</th>
<th>TXN-LOCK</th>
<th>SPAN-OCC</th>
<th>SPAN-LOCK</th>
<th>TAPIR</th>
<th>QWSTORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latency</td>
<td>3653</td>
<td>2843</td>
<td>2792</td>
<td>2911</td>
<td>1432</td>
<td>1225</td>
</tr>
</tbody>
</table>

**Peak Throughput (Transactions/second)**

<table>
<thead>
<tr>
<th>System</th>
<th>TXN-OCC</th>
<th>TXN-LOCK</th>
<th>SPAN-OCC</th>
<th>SPAN-LOCK</th>
<th>TAPIR</th>
<th>QWSTORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Throughput</td>
<td>5482</td>
<td>6157</td>
<td>7209</td>
<td>7080</td>
<td>17500</td>
<td>71058</td>
</tr>
</tbody>
</table>