Consistency Models

• Contract between a distributed system and the applications that run on it

• A consistency model is a set of guarantees made by the distributed system
Linearizability

• All replicas execute operations in some total order

• That total order preserves the real-time ordering between operations
  • If operation A completes before operation B begins, then A is ordered before B in real-time
  • If neither A nor B completes before the other begins, then there is no real-time order
    • (But there must be some total order)
Real-Time Ordering Examples

\[ P_A \vdash w(x=1) \]
\[ P_B \vdash w(x=2) \]
\[ P_C \vdash w(x=3) \]
\[ P_D \vdash w(x=4) \vdash w(x=5) \]
\[ P_E \vdash w(x=6) \]
Linearizable?

\[
\begin{align*}
P_A & \quad \models w(x=1) \\
P_B & \models w(x=2) \\
P_C & \models w(x=3) \\
P_D & \models w(x=4) \quad \models w(x=5) \\
P_E & \models w(x=6) \\
P_F & \models r(x)=1 \quad \models r(x)=2 \quad \models r(x)=3 \quad \models r(x)=6 \quad \models r(x)=5
\end{align*}
\]
Linearizable?

\[ \begin{align*}
\mathcal{P}_A & \quad \vdash w(x=1) \\
\mathcal{P}_B & \quad \vdash w(x=2) \\
\mathcal{P}_C & \quad \vdash w(x=3) \\
\mathcal{P}_D & \quad \vdash w(x=4) \quad \vdash w(x=5) \\
\mathcal{P}_E & \quad \vdash w(x=6) \\
\mathcal{P}_G & \quad \vdash r(x)=1 \quad \vdash r(x)=2 \quad \vdash r(x)=5 \quad \vdash r(x)=6 \quad \vdash r(x)=5 \quad X
\end{align*} \]
Linearizable?

\[ P_A \models w(x=1) \]
\[ P_B \models w(x=2) \]
\[ P_C \models w(x=3) \]
\[ P_D \models w(x=4) \quad \models w(x=5) \]
\[ P_E \models w(x=6) \]

\[ P_H \models r(x)=1 \quad \models r(x)=4 \quad \models r(x)=2 \quad \models r(x)=3 \quad \models r(x)=6 \]
Linearizable?

\[ P_A \vdash w(x=1) \]
\[ P_B \vdash w(x=2) \]
\[ P_C \vdash w(x=3) \]
\[ P_D \vdash w(x=4) \quad \vdash w(x=5) \]
\[ P_E \vdash w(x=6) \]

\[ P_I \vdash r(x)=1 \quad \vdash r(x)=4 \quad \vdash r(x)=5 \quad \vdash r(x)=6 \quad \vdash r(x)=3 \quad X \]
Linearizability == “Appears to be a Single Machine”

• Single machine processes requests one by one in the order it receives them
  • Will receive requests ordered by real-time in that order
  • Will receive all requests in some order

• Atomic Multicast, Viewstamped Replication, Paxos, and RAFT provide Linearizability

• Single machine processing incoming requests one at a time also provide Linearizability 😊
Linearizability is Ideal?

• Hides the complexity of the underlying distributed system from applications!
  • Easier to write applications
  • Easier to write correct applications

• But, performance trade-offs
Stronger vs Weaker Consistency

• Stronger consistency models
  + Easier to write applications
  - More guarantees for the system to ensure
    Results in performance tradeoffs

• Weaker consistency models
  - Harder to write applications
  + Fewer guarantees for the system to ensure
Strictly Stronger Consistency

• A consistency model $A$ is strictly stronger than $B$ if it allows a strict subset of the behaviors of $B$
  • Guarantees are strictly stronger
Sequential Consistency

• All replicas execute operations in some total order

• That total order preserves the process ordering between operations
  • If process P issues operation A before operation B, then A is order before B by the process order
  • If operations A and B and done by different processes then there is no process order between them
    • (But there must be some total order)
Sequential Consistency ≈
“Appears to be a Single Machine”

• Single machine processes requests one by one in the order it receives them
  • Will receive requests ordered by process order in that order
  • Will receive all requests in some order
Linearizability is strictly stronger than Sequential Consistency

- Linearizability: ∃ total order + real-time ordering
- Sequential: ∃ total order + process ordering
  - Process ordering ⊆ Real-time ordering
Sequential But Not Linearizable

A

w(alicePwd=0x5f4)

B

r(alicePwd)=0x81d

One week later
Consistency Hierarchy

- Linearizability
- Sequential Consistency
- Causal+ Consistency
- Eventual Consistency

Examples:
- e.g., RAFT
- e.g., Bayou
- e.g., Dynamo
Causal+ Consistency

• Partially orders all operations, does not totally order them
  • Does not look like a single machine

• Guarantees
  • For each process, ∃ an order of all writes + that process’s reads
  • Order respects the happens-before (⇒) ordering of operations
  • + replicas converge to the same state
    • Skip details, makes it stronger than eventual consistency
Causal Consistency

1. Writes that are potentially causally related must be seen by all processes in same order.

2. Concurrent writes may be seen in a different order on different processes.
   - Concurrent: Ops not causally related
Causal Consistency

1. Writes that are potentially causally related must be seen by all processes in the same order.

2. Concurrent writes may be seen in a different order on different processes.
   • Concurrent: Ops not causally related
Causal Consistency

<table>
<thead>
<tr>
<th>Operations</th>
<th>Concurrent?</th>
</tr>
</thead>
<tbody>
<tr>
<td>a, b</td>
<td>Y</td>
</tr>
<tr>
<td>b, f</td>
<td>N</td>
</tr>
<tr>
<td>c, f</td>
<td>Y</td>
</tr>
<tr>
<td>e, f</td>
<td>Y</td>
</tr>
<tr>
<td>e, g</td>
<td>N</td>
</tr>
<tr>
<td>a, c</td>
<td>Y</td>
</tr>
<tr>
<td>a, e</td>
<td>N</td>
</tr>
</tbody>
</table>

Physical time ↓
## Causal Consistency

<table>
<thead>
<tr>
<th>Operations</th>
<th>Concurrent?</th>
</tr>
</thead>
<tbody>
<tr>
<td>a, b</td>
<td>N</td>
</tr>
<tr>
<td>b, f</td>
<td>Y</td>
</tr>
<tr>
<td>c, f</td>
<td>Y</td>
</tr>
<tr>
<td>e, f</td>
<td>Y</td>
</tr>
<tr>
<td>e, g</td>
<td>N</td>
</tr>
<tr>
<td>a, c</td>
<td>Y</td>
</tr>
<tr>
<td>a, e</td>
<td>N</td>
</tr>
</tbody>
</table>

![Diagram showing operations and concurrency]

Physical time ↓
Causal+ But Not Sequential

\[ P_A \quad \models \quad w(x=1) \quad \models \quad r(y)=0 \quad \models \]

\[ P_B \quad \models \quad w(y=1) \quad \models \quad r(x)=0 \quad \models \]

√ Casual+

Happens Before Order

\[ w(x=1) \rightarrow r(y)=0 \]

\[ w(y=1) \rightarrow r(x)=0 \]

\[ P_A \text{ Order: } w(x=1), r(y)=0, w(y=1) \]

\[ P_B \text{ Order: } w(y=1), r(x)=0, w(x=1) \]

X Sequential

Process Ordering

\[ w(x=1) \rightarrow r(y)=0 \]

\[ w(y=1) \rightarrow r(x)=0 \]

No Total Order

\[ w(x=1) \rightarrow r(y)=0 \]

\[ w(y=1) \rightarrow r(x)=0 \]
Eventual But Not Causal+

\[ P_A \models w(x=1) \models w(y)=1 \models \]

\[ P_B \models r(y)=1 \models r(x)=0 \models \]

\checkmark Eventual

As long as \( P_B \) eventually would see \( r(x)=1 \) this is fine

\x Causal+

Happens Before Ordering

\[ w(x=1) \rightarrow w(y)=1 \]

\[ r(y)=1 \rightarrow r(x)=0 \]

No Order for \( P_B \)

\[ w(x=1) \rightarrow w(y)=1 \]

\[ r(y)=1 \rightarrow r(x)=0 \]
Consistency Hierarchy

- Linearizability: e.g., RAFT
- Sequential Consistency
- Causal+ Consistency: e.g., Bayou
- Eventual Consistency: e.g., Dynamo
Causal Consistency: Quiz

- Valid under causal consistency

- Why? \( x=3 \) and \( x=2 \) are concurrent
  - So all processes don’t (need to) see them in same order

- \( P_C \) and \( P_D \) read the values ‘1’ and ‘2’ in order as potentially causally related. No ‘causality’ for ‘3’.
Sequential Consistency: Quiz

- Invalid under sequential consistency

- Why? $P_C$ and $P_D$ see 2 and 3 in different order

- But fine for causal consistency
  - 2 and 3 are not causally related
Causal Consistency

$P_A \vdash w(x=1)$

$P_B \vdash r(x)=1 \quad \vdash w(x=2)$

$P_C \vdash r(x)=2 \quad \vdash r(x)=1$

$P_D \vdash r(x)=1 \quad \vdash r(x)=2$

$X \quad x=2$ happens after $x=1$
Causal Consistency

\( P_A \vdash w(x=a) \)
\( P_B \vdash w(x=b) \)
\( P_C \vdash r(x)=a \)
\( P_D \vdash r(x)=b \)

✓ \( P_B \) doesn’t read value of 1 before writing 2