6.033 Spring 2021

Lecture #17: Isolation
what do we want from isolation, and how do we get it?
our goal is to build **reliable systems from unreliable components**. We want to build systems that serve many clients, store a lot of data, perform well, all while keeping availability high.

**transactions** — which provide **atomicity** and **isolation** — make it easier for us to reason about failures.

Our job in lecture is to understand how a system *implements* these two abstractions. How do our systems guarantee atomicity? How do they guarantee isolation?

**Atomicity**: provided by **logging**, which gives better performance than shadow copies* at the cost of some added complexity.

**Isolation**: we don’t really have this yet. (Coarse-grained locks perform poorly; fine-grained locks are difficult to reason about.)

*Shadow copies are used in some systems.
our goal is to build reliable systems from unreliable components. we want to build systems that serve many clients, store a lot of data, perform well, all while keeping availability high.

transactions — which provide atomicity and isolation — make it easier for us to reason about failures.

our job in lecture is to understand how a system implements these two abstractions. how do our systems guarantee atomicity? how do they guarantee isolation?

atomicity: provided by logging, which gives better performance than shadow copies* at the cost of some added complexity.

isolation: provided by two-phase locking.

* shadow copies are used in some systems.
**goal:** run transactions T1, T2, .., TN concurrently, and have it “appear” as if they ran sequentially

when we run two transactions concurrently, we’ll always run the steps of a single transaction in order (e.g., T1.1 before T1.2). But we might interleave steps of T2 in between steps of T1.

**naive approach:** actually run them sequentially, via (perhaps) a single global lock
**goal:** run transactions **T1**, **T2** concurrently, and have it “appear” as if they ran sequentially

<table>
<thead>
<tr>
<th>T1</th>
<th>begin</th>
<th>T1.1 read(x)</th>
<th>T1.2 tmp = read(y)</th>
<th>T1.3 write(y, tmp+10)</th>
<th>commit</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2</td>
<td>begin</td>
<td>T2.1 write(x, 20)</td>
<td>T2.2 write(y, 30)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(commit)

(assume x, y initialized to zero)

result: x=20; y=30

let’s look at a few different schedules of **T1** and **T2** (this is not an exhaustive list)

<table>
<thead>
<tr>
<th>T2.1 write(x, 20)</th>
<th>T1.1 read(x)</th>
<th>T2.1 write(x, 20)</th>
<th>T1.1 read(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1.1 read(x)</td>
<td>T2.1 write(x, 20)</td>
<td>T1.2 tmp = read(y)</td>
<td>T2.1 write(x, 20)</td>
</tr>
<tr>
<td>T2.2 write(y, 30)</td>
<td>T1.2 tmp = read(y)</td>
<td>T2.2 write(y, 30)</td>
<td>T1.2 tmp = read(y)</td>
</tr>
<tr>
<td>T1.2 tmp = read(y)</td>
<td>T1.3 write(y, 30)</td>
<td>T1.3 write(y, 30)</td>
<td>T1.3 write(y, tmp+10)</td>
</tr>
<tr>
<td>T1.3 write(y, tmp+10)</td>
<td>T2.1 write(x, 20)</td>
<td>T2.2 write(y, 30)</td>
<td>T2.2 write(y, 30)</td>
</tr>
</tbody>
</table>

result: x=20; y=40

result: x=20; y=40

result: x=20; y=10

it seems like the middle schedule is out; **x=20; y=10** is not possible in either of our serialized schedules
**goal:** run transactions **T1**, **T2** concurrently, and have it “appear” as if they ran sequentially

<table>
<thead>
<tr>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>begin</td>
<td>begin</td>
</tr>
<tr>
<td>T1.1 read(x)</td>
<td>T2.1 write(x, 20)</td>
</tr>
<tr>
<td>T1.2 tmp = read(y)</td>
<td>T2.2 write(y, 30)</td>
</tr>
<tr>
<td>T1.3 write(y, tmp+10)</td>
<td>commit</td>
</tr>
</tbody>
</table>

(commit)

(assume x, y initialized to zero)

let’s look at a few different schedules of **T1** and **T2** (this is not an exhaustive list)

- **T1.1 read(x)**
- **T1.2 tmp = read(y)**
- **T1.3 write(y, tmp+10)**
- **commit**

- **T2.1 write(x, 20)**
- **T2.2 write(y, 30)**

result: x=20; y=40

but take a closer look at the third schedule; in the first step, **T1.1 reads x=0**, and in the fourth step, **T1.2 reads y=30**. those two reads together aren’t possible in a sequential schedule. **is that okay?**
there are many ways for multiple transactions to “appear” to have been run in sequence; we say there are different notions of **serializability**. what type of serializability you want depends on what your application needs.
conflicts: two operations conflict if they operate on the same object and at least one of them is a write

<table>
<thead>
<tr>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>begin</td>
<td>begin</td>
</tr>
<tr>
<td>T1.1</td>
<td>T2.1 write(x, 20)</td>
</tr>
<tr>
<td>T1.2</td>
<td>T2.2 write(y, 30)</td>
</tr>
<tr>
<td>T1.3</td>
<td>commit</td>
</tr>
<tr>
<td>tmp = read(y)</td>
<td>write(y, 30)</td>
</tr>
<tr>
<td>write(y, tmp+10)</td>
<td>write(y, 30)</td>
</tr>
</tbody>
</table>

(assume x, y initialized to zero)

in any schedule, two conflicting operations A and B will have an order: either A is executed before B, or B is executed before A. we'll call this the order of the conflict (in that schedule).

order of conflicts

| T1.1 read(x) | T2.1 write(x, 20) |
| T1.2 tmp = read(y) | T2.2 write(y, 30) |
| T1.3 write(y, tmp+10) | T2.2 write(y, 30) |
**conflicts:** two operations conflict if they operate on the same object and at least one of them is a write

<table>
<thead>
<tr>
<th>T1</th>
<th>T2</th>
</tr>
</thead>
<tbody>
<tr>
<td>begin</td>
<td>begin</td>
</tr>
<tr>
<td>T1.1 read(x)</td>
<td>T2.1 write(x, 20)</td>
</tr>
<tr>
<td>T1.2 tmp = read(y)</td>
<td>T2.2 write(y, 30)</td>
</tr>
<tr>
<td>T1.3 write(y, tmp+10)</td>
<td>commit</td>
</tr>
</tbody>
</table>

(assume x, y initialized to zero)

in any schedule, two conflicting operations A and B will have an order: either A is executed before B, or B is executed before A. we'll call this the **order** of the conflict (in that schedule).

<table>
<thead>
<tr>
<th>T1.1 read(x)</th>
<th>T1.1 -&gt; T2.1</th>
<th>T2.1 write(x, 20)</th>
<th>T2.1 -&gt; T1.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1.2 tmp = read(y)</td>
<td>T1.2 -&gt; T2.2</td>
<td>T2.2 write(y, 30)</td>
<td>T2.2 -&gt; T1.2</td>
</tr>
<tr>
<td>T1.3 write(y, tmp+10)</td>
<td>T1.3 -&gt; T2.2</td>
<td>T1.1 read(x)</td>
<td>T1.1 -&gt; T2.1</td>
</tr>
<tr>
<td>T2.1 write(x, 20)</td>
<td></td>
<td>T1.2 tmp = read(y)</td>
<td></td>
</tr>
<tr>
<td>T2.2 write(y, 30)</td>
<td></td>
<td>T1.3 write(y, tmp+10)</td>
<td></td>
</tr>
</tbody>
</table>

notice that, if we execute **T1** and **T2** serially, then in the ordering of the conflicts we see either *all* of **T1**'s operations occurring first, or *all* of **T2**'s operations occurring first.
**conflicts:** two operations conflict if they operate on the same object and at least one of them is a write operation.

<table>
<thead>
<tr>
<th>Schedule</th>
<th>Operations</th>
<th>Order of Conflicts</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>read(x)</td>
<td>T2.1 -&gt; T1.1</td>
</tr>
<tr>
<td>T1</td>
<td>tmp = read(y)</td>
<td>T2.2 -&gt; T1.2</td>
</tr>
<tr>
<td>T1</td>
<td>write(y, tmp+10)</td>
<td>T2.2 -&gt; T1.3</td>
</tr>
</tbody>
</table>

(assume x, y initialized to zero)

in any schedule, two conflicting operations A and B will have an order: either A is executed before B, or B is executed before A. we’ll call this the **order** of the conflict (in that schedule).

<table>
<thead>
<tr>
<th>Schedule</th>
<th>Operations</th>
<th>Order of Conflicts</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2</td>
<td>write(x, 20)</td>
<td>T1.1 -&gt; T2.1</td>
</tr>
<tr>
<td>T2</td>
<td>write(y, 30)</td>
<td>T2.2 -&gt; T1.2</td>
</tr>
<tr>
<td>T2</td>
<td>tmp = read(y)</td>
<td>T1.2 -&gt; T1.3</td>
</tr>
<tr>
<td>T2</td>
<td>write(y, tmp+10)</td>
<td>T1.3 -&gt; T2.2</td>
</tr>
</tbody>
</table>

on the left schedule, the order of conflicts is the same as if we had run **T2** entirely before **T1**; on the right schedule, the order of conflicts isn’t the same as either serial schedule.
a schedule is **conflict serializable** if the order of all of its conflicts is the same as the order of the conflicts in some sequential schedule.

(assume x, y initialized to zero)

**T1**

begin
T1.1 read(x)
T1.2 tmp = read(y)
T1.3 write(y, tmp+10)
commit

**T2**

begin
T2.1 write(x, 20)
T2.2 write(y, 30)
commit

**conflicts**

T1.1 read(x) and T2.1 write(x, 20)
T1.2 tmp = read(y) and T2.2 write(y, 30)
T1.3 write(y, tmp+10) and T2.2 write(y, 30)

**order of conflicts**

T2.1 write(x, 20) -> T1.1
T2.2 write(y, 30) -> T1.2
T1.3 write(y, tmp+10) -> T2.2

**order of conflicts**

T1.1 read(x) -> T2.1
T2.1 write(x, 20) -> T1.2
T2.2 write(y, 30) -> T1.3
T1.2 tmp = read(y) -> T2.2
T1.3 write(y, tmp+10) -> T2.2

this schedule is **conflict serializable**

this schedule is **not** conflict serializable
a schedule is **conflict serializable** if the order of all of its conflicts is the same as the order of the conflicts in some sequential schedule.

\[
\begin{align*}
\text{T1} & \begin{align*}
\text{begin} \\
\text{T1.1} & \text{read(x)} \\
\text{T1.2} & \text{tmp = read(y)} \\
\text{T1.3} & \text{write(y, tmp+10)} \\
\text{commit}
\end{align*} \\
\text{T2} & \begin{align*}
\text{begin} \\
\text{T2.1} & \text{write(x, 20)} \\
\text{T2.2} & \text{write(y, 30)} \\
\text{commit}
\end{align*}
\end{align*}
\]

(assume x, y initialized to zero)

\[
\begin{align*}
\text{conflicts} & \\
\text{T1.1} & \text{read(x)} \text{ and } \text{T2.1} \text{ write(x, 20)} \\
\text{T1.2} & \text{tmp = read(y)} \text{ and } \text{T2.2} \text{ write(y, 30)} \\
\text{T1.3} & \text{write(y, tmp+10)} \text{ and } \text{T2.2} \text{ write(y, 30)}
\end{align*}
\]

we can express the order of conflicts more succinctly with a **conflict graph**: there is an edge from \( T_i \) to \( T_j \) if and only if \( T_i \) and \( T_j \) have a conflict between them and the first step in the conflict occurs in \( T_i \).

\[
\begin{align*}
\text{T2.1} & \text{write(x, 20)} \\
\text{T1.1} & \text{read(x)} \\
\text{T2.2} & \text{write(y, 30)} \\
\text{T1.2} & \text{tmp = read(y)} \\
\text{T1.3} & \text{write(y, tmp+10)}
\end{align*}
\]

\[
\begin{align*}
\text{conflict graph} & \\
\text{T2} & \rightarrow \text{T1}
\end{align*}
\]

\[
\begin{align*}
\text{T1.1} & \text{read(x)} \\
\text{T2.1} & \text{write(x, 20)} \\
\text{T2.2} & \text{write(y, 30)} \\
\text{T1.2} & \text{tmp = read(y)} \\
\text{T1.3} & \text{write(y, tmp+10)}
\end{align*}
\]

\[
\begin{align*}
\text{conflict graph} & \\
\text{T2} & \leftrightarrow \text{T1}
\end{align*}
\]

this schedule is conflict serializable

this schedule is **not** conflict serializable

\[
\begin{align*}
\text{a schedule is conflict serializable if and only if it has an acyclic conflict graph}
\end{align*}
\]

Katrina LaCurts | lacurts@mit.edu | 6.033 2021
A schedule is **conflict serializable** if the order of all of its conflicts is the same as the order of the conflicts in some sequential schedule.

<table>
<thead>
<tr>
<th>T1 start</th>
<th>T2 start</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>T1.1</strong> read(x)</td>
<td><strong>T2.1</strong> write(x, 20)</td>
</tr>
<tr>
<td><strong>T1.2</strong> tmp = read(y)</td>
<td><strong>T2.2</strong> write(y, 30)</td>
</tr>
<tr>
<td><strong>T1.3</strong> write(y, tmp+10)</td>
<td>commit</td>
</tr>
<tr>
<td>commit</td>
<td></td>
</tr>
</tbody>
</table>

(assume x, y initialized to zero)

**conflicts**

- **T1.1** read(x) and **T2.1** write(x, 20)
- **T1.2** tmp = read(y) and **T2.2** write(y, 30)
- **T1.3** write(y, tmp+10) and **T2.2** write(y, 30)

**our goal (in lecture) is to run transactions concurrently, but to produce a schedule that is conflict serializable**

how does a system do that? one way might be to generate all possible schedules and check their conflict graphs, and run one of the schedules with an acyclic conflict graph, but this will take some time
two-phase locking (2PL)

1. each shared variable has a lock

2. before **any** operation on a variable, the transaction must acquire the corresponding lock

3. after a transaction releases a lock, it may **not** acquire any other locks

---

2PL still gives us options for where we place the locks

```
T1
begin
acquire(x.lock)
acquire(y.lock)
T1.1 read(x)
T1.2 tmp = read(y)
T1.3 write(y, tmp+10)
release(x.lock)
release(y.lock)
commit
```
**two-phase locking (2PL)**

1. each shared variable has a lock

2. before **any** operation on a variable, the transaction must acquire the corresponding lock

3. after a transaction releases a lock, it may **not** acquire any other locks

2PL still gives us options for where we place the locks

```
T1
begin
acquire(x.lock)
T1.1 read(x)
acquire(y.lock)
T1.2 tmp = read(y)
T1.3 write(y, tmp+10)
release(x.lock)
release(y.lock)
commit
```
two-phase locking (2PL)

1. each shared variable has a lock

2. before any operation on a variable, the transaction must acquire the corresponding lock

3. after a transaction releases a lock, it may not acquire any other locks

2PL still gives us options for where we place the locks

T1
begin
acquire(x.lock)
T1.1 read(x)
acquire(y.lock)
release(x.lock)
T1.2 tmp = read(y)
T1.3 write(y, tmp+10)
release(y.lock)
commit
two-phase locking (2PL)

1. each shared variable has a lock

2. before any operation on a variable, the transaction must acquire the corresponding lock

3. after a transaction releases a lock, it may not acquire any other locks

2PL still gives us options for where we place the locks

T1
begin
acquire(x.lock)
T1.1 read(x)
release(x.lock)
acquire(y.lock)
T1.2 tmp = read(y)
T1.3 write(y, tmp+10)
release(y.lock)
commit

we can't do this; it breaks the third rule of 2PL
two-phase locking (2PL)

1. each shared variable has a lock

2. before any operation on a variable, the transaction must acquire the corresponding lock

3. after a transaction releases a lock, it may not acquire any other locks

often we release locks after commit, which is technically strict two-phase locking

2PL still gives us options for where we place the locks

T1
begin
acquire(x.lock)
T1.1 read(x)
acquire(y.lock)
T1.2 tmp = read(y)
T1.3 write(y, tmp+10)
commit
release(x.lock)
release(y.lock)

T2
begin
acquire(x.lock)
T2.1 write(x, 20)
acquire(y.lock)
T2.2 write(y, 30)
commit
release(x.lock)
release(y.lock)

notice that with this approach to 2PL, we will effectively force these two transactions to run serially. we’ll address that in a few slides!

there are some lingering issues related to possible deadlocks and performance; we’ll deal with those, but let’s first try to understand why 2PL produces a conflict-serializable schedule
two-phase locking (2PL)

1. each shared variable has a lock

2. before any operation on a variable, the transaction must acquire the corresponding lock

3. after a transaction releases a lock, it may not acquire any other locks

2PL produces a conflict-serializable schedule
(equivalently, 2PL produces a conflict graph without a cycle)

**proof:** suppose not. then a cycle exists in the conflict graph
to cause the conflict, each pair of conflicting transactions must have some shared variable that they conflict on

\[ T_1 \xrightarrow{x_1} T_2 \xrightarrow{x_2} T_3 \xrightarrow{x_3} \ldots \xrightarrow{x_{k-1}} T_k \]

in the schedule, each pair of transactions needs to acquire a lock on their shared variable

\[ T_1 \text{ acquires } x_1.\text{lock} \]
\[ T_2 \text{ acquires } x_1.\text{lock} \]
\[ T_2 \text{ acquires } x_2.\text{lock} \]
\[ T_3 \text{ acquires } x_2.\text{lock} \]
\[ \ldots \]
\[ T_k \text{ acquires } x_k.\text{lock} \]
\[ T_1 \text{ acquires } x_k.\text{lock} \]

in order for the schedule to progress, \( T_1 \) must have released its lock on \( x_1 \) before \( T_2 \) acquired it
two-phase locking (2PL)

1. each shared variable has a lock

2. before any operation on a variable, the transaction must acquire the corresponding lock

3. after a transaction releases a lock, it may not acquire any other locks

2PL produces a conflict-serializable schedule
\[(\text{equivalently, 2PL produces a conflict graph without a cycle})\]

\textbf{proof:} suppose not. then a cycle exists in the conflict graph

\begin{align*}
T_1 \xrightarrow{x_1} T_2 \xrightarrow{x_2} T_3 \xrightarrow{x_3} \ldots \xrightarrow{x_{k-1}} T_k \\
\end{align*}

\text{to cause the conflict, each pair of conflicting transactions must have some shared variable that they conflict on}

\begin{align*}
T_1 &\text{acquires } x_1.\text{lock} \\
T_1 &\text{releases } x_1.\text{lock} \\
T_2 &\text{acquires } x_1.\text{lock} \\
T_2 &\text{acquires } x_2.\text{lock} \\
T_3 &\text{acquires } x_2.\text{lock} \\
\ldots \\
T_k &\text{acquires } x_k.\text{lock} \\
T_1 &\text{acquires } x_k.\text{lock} \\
\end{align*}

\text{in the schedule, each pair of transactions needs to acquire a lock on their shared variable}

\text{in order for the schedule to progress, } T_1 \text{ must have released its lock on } x_1 \text{ before } T_2 \text{ acquired it}

\textbf{contradiction:} this is not a valid 2PL schedule
two-phase locking (2PL)

1. each shared variable has a lock

2. before any operation on a variable, the transaction must acquire the corresponding lock

3. after a transaction releases a lock, it may not acquire any other locks

problem: 2PL can result in deadlock

T1
begin
acquire(x.lock)
T1.1 read(x)
acquire(y.lock)
T1.2 tmp = read(y)
T1.3 write(y, tmp+10)
commit
release(x.lock)
release(y.lock)

T2
begin
acquire(y.lock)
T2.1 write(y, 30)
acquire(x.lock)
T2.2 write(x, 20)
commit
release(x.lock)
release(y.lock)

for example, suppose T2 wrote to y before x

one solution to this problem is a global ordering on locks; but we hate that! a better solution is to take advantage of atomicity and abort one of the transactions
two-phase locking (2PL) with reader-/writer- locks

1. each shared variable has two locks: one for reading, one for writing

2. before any operation on a variable, the transaction must acquire the appropriate lock

3. multiple transactions can hold reader locks for the same variable at once; a transaction can only hold a writer lock for a variable if there are no other locks held for that variable

4. after a transaction releases a lock, it may not acquire any other locks

problem: performance

T1
begin
acquire(x.reader_lock)
T1.1 read(x)
acquire(y.reader_lock)
T1.2 tmp = read(y)
acquire(y.writer_lock)
T1.3 write(y, tmp+10)
commit
release(x.reader_lock)
release(y.reader_lock)

T2
begin
acquire(x.writer_lock)
T2.1 write(x, 20)
acquire(y.writer_lock)
T2.2 write(y, 30)
commit
release(x.writer_lock)
release(y.writer_lock)
two-phase locking (2PL) with reader-/writer- locks

1. each shared variable has two locks: one for reading, one for writing

2. before any operation on a variable, the transaction must acquire the appropriate lock

3. multiple transactions can hold reader locks for the same variable at once; a transaction can only hold a writer lock for a variable if there are no other locks held for that variable

4. after a transaction releases a lock, it may not acquire any other locks

problem: performance

T1
begin
acquire(x.reader_lock)
T1.1 read(x)
acquire(y.reader_lock)
T1.2 tmp = read(y)
acquire(y.writer_lock)
release(x.reader_lock)
release(y.reader_lock)
T1.3 write(y, tmp+10)
commit
release(y.writer_lock)

T2
begin
acquire(x.writer_lock)
T2.1 write(x, 20)
acquire(y.writer_lock)
T2.2 write(y, 30)
commit
release(x.writer_lock)
release(y.writer_lock)

we will often see reader locks be released before the commit
our goal is to build **reliable systems from unreliable components**. we want to build systems that serve many clients, store a lot of data, perform well, all while keeping availability high.

transactions — which provide **atomicity** and **isolation** — make it easier for us to reason about failures.

our job in lecture is to understand how a system *implements* these two abstractions. how do our systems guarantee atomicity? how do they guarantee isolation?

**atomicity:** provided by **logging**, which gives better performance than shadow copies* at the cost of some added complexity  

**isolation:** provided by **two-phase locking**

* shadow copies are used in some systems
different types of **serializability** allow us to specify precise what we want when we run transactions in parallel. **conflict-serializability** is a relatively strict form of serializability.

**two-phase locking** allows us to generate conflict-serializable schedules. we can improve its performance by allowing concurrent reads via reader- and writer- locks.

2PL does not produce every possible conflict-serializable schedule — that's okay! the claim is only that the schedules it does produce are conflict-serializable