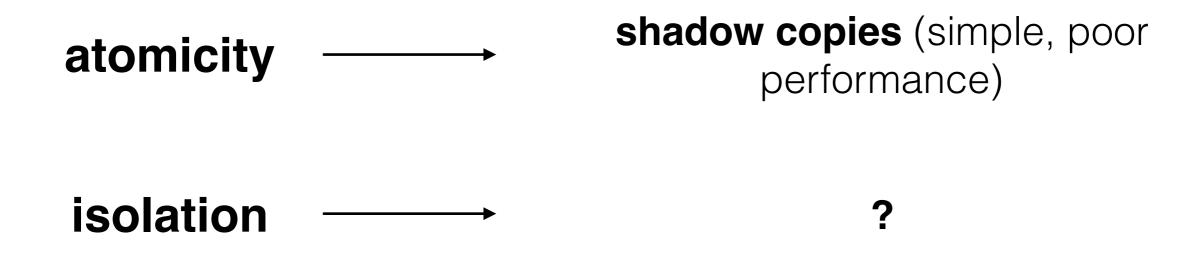
6.033 Spring 2017

Lecture #16

Atomicity via Write-ahead logging

goal: build reliable systems from unreliable components the abstraction that makes that easier is

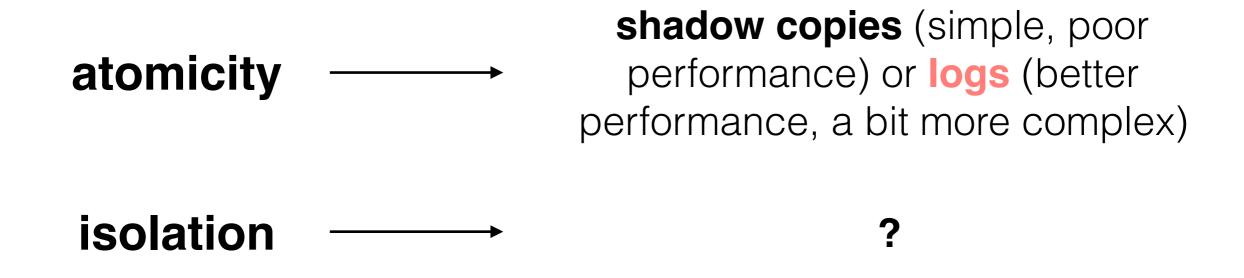
transactions, which provide atomicity and isolation, while not hindering performance



eventually, we also want transaction-based systems to be **distributed**: to run across multiple machines

goal: build reliable systems from unreliable components the abstraction that makes that easier is

transactions, which provide atomicity and isolation, while not hindering performance



eventually, we also want transaction-based systems to be **distributed**: to run across multiple machines

using shadow copies to abort on error

```
transfer(bankfile, account_a, account_b, amount):
   bank = read_accounts(bankfile)
   bank[account_a] = bank[account_a] - amount
   bank[account_b] = bank[account_b] + amount
   if bank[account_a] < 0:
      print "Not enough funds"
   else:
      write_accounts("tmp_bankfile")
      rename(tmp_bankfile, bankfile)</pre>
```

with transaction syntax

```
transfer(account_a, account_b, amount):
    begin
    write(account_a, read(account_a) - amount)
    write(account_b, read(account_b) + amount)
    if read(account_a) < 0: // not enough funds
        abort
    else:
        commit</pre>
```

```
begin // T1
A = 100
B = 50
commit // A=100; B=50
begin // T2
A = A - 20
B = B + 20
commit // A=80; B=70
begin // T3
A = A+30
crash! 💥
```

problem: after crash, A=110,
but T3 never committed

we need a way to revert to A's previous committed value

-	+		 			 	}	H
TID	T1	T1	T1	T2	T2	T2	Т3	
	UPDATE	UPDATE	COMMIT	UPDATE	UPDATE	COMMIT	UPDATE	
OLD	A=0	B=0		A=100	B=50		A=80	ĺ
NEW	A=100	B=50		A=80	B=70		A=110	

```
begin // T1
A = 100
B = 50
commit // A=100; B=50
begin // T2
A = A - 20
B = B+20
commit // A=80; B=70
begin // T3
A = A+30
```

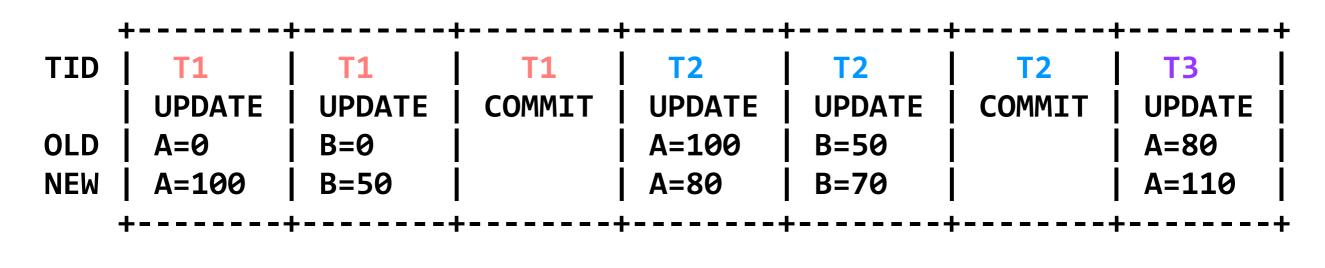
```
TID
                                       T2
               UPDATE
       UPDATE
                       COMMIT
                              UPDATE
                                      UPDATE
                                              COMMIT
                                                      UPDATE
                              A=100
  OLD
                                      B=50
       A=0
               B=0
                                                      A = 80
  NEW
       A = 100
               B=50
                              A = 80
                                      B = 70
read(log, var):
  commits = \{\}
  // scan backwards
  for record r in log[len(log) - 1] .. log[0]:
    // keep track of commits
    if r.type == commit:
       commits.add(r.tid)
    // find var's last committed value
    if r.type == update and
        r.tid in commits and
        r.var == var:
         return r.new value
```

```
TID
                               T2
                                       T2
       UPDATE
               UPDATE
                       COMMIT
                              UPDATE
                                      UPDATE
                                              COMMTT
                                                      UPDATE
              B=0
  OLD
                              A = 100
                                     I B=50
       A=0
                                                      A = 80
               B=50
  NEW
       A = 100
                              A = 80
                                      B = 70
read(log, var):
  commits = \{\}
  // scan backwards
  for record r in log[len(log) - 1] .. log[0]:
    // keep track of commits
    if r.type == commit:
       commits.add(r.tid)
    // find var's last committed value
    if r.type == update and
        (r.tid in commits or r.tid == current tid) and
        r.var == var:
         return r.new value
```

_	 	+			+		+ <i>-</i>	+
TID	T1	T1	T1	T2	T2	T2	Т3	
	UPDATE	UPDATE	COMMIT	UPDATE	UPDATE	COMMIT	UPDATE	
OLD	A=0	B=0		A=100	B=50		A=80	Ī
NEW	A=100	B=50		A=80	B=70		A=110	Ī
_	L	L	L	L	L	L	L	_

performance?

problem: reads are slow



cell storage

A 110

B 70

read(var):

return cell_read(var)

write(var, value):

log.append(current_tid, update, var, read(var), value)
cell write(var, value)

```
TID
                              T1
                                       T2
                                                  T2
                                                              T2
                                                                       T3
       UPDATE
                 UPDATE
                            COMMIT
                                      UPDATE
                                                 UPDATE
                                                           COMMIT
                                                                      UPDATE
      A=0
                                      A=100
OLD
                 B=0
                                                B=50
                                                                      A = 80
NEW
       A = 100
                 B=50
                                      A = 80
                                                 B = 70
                                                                      A = 110
     cell storage
                         A 110
                                    В
                                        70
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
      if r.type == commit:
        commits.add(r.tid)
      if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```

```
TID
                                      T2
                                                T2
                                                            T2
      UPDATE
                 UPDATE
                           COMMIT
                                               UPDATE
                                                          COMMIT
                                     UPDATE
                                     A=100
                                               B=50
OLD
      A=0
                 B=0
                                                                    A = 80
NEW
      A = 100
                 B=50
                                     A = 80
                                               B = 70
                                                                    A = 110
                                                         commits = \{\}
     cell storage
                        A 110
                                   В
                                       70
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
     if r.type == commit:
        commits.add(r.tid)
     if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```

```
TID
                                      T2
                                                T2
                                                            T2
                                                         COMMIT
      UPDATE
                 UPDATE
                           COMMIT
                                               UPDATE
                                     UPDATE
                                     A=100
OLD
      A=0
                                               B=50
                                                                    A = 80
                 B=0
NEW
      A = 100
                 B=50
                                     A = 80
                                               B = 70
                                                                    A = 110
                                                         commits = \{\}
     cell storage
                        A 110
                                   В
                                       70
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
        if r.type == commit:
        commits.add(r.tid)
        if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```

```
TID
                            T1
                                     T2
                                               T2
                                                          T2
      UPDATE
                UPDATE
                          COMMIT
                                    UPDATE
                                              UPDATE
                                                        COMMIT
                                                                  UPDATE
                                    A=100
OLD
      A=0
                B=0
                                              B=50
                                                                  A = 80
NEW
      A = 100
                B=50
                                    A = 80
                                              B=70
                                                                  A = 110
                                                        commits = \{\}
    cell storage
                        A 110
                                   В
                                      70
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
      if r.type == commit:
        commits.add(r.tid)
      if r.type == update and r.tid not in commits:
        cell write(r.var, r.old val) // undo
```

```
TID
                                      T2
                                                T2
                                                           T2
                                                         COMMIT
      UPDATE
                UPDATE
                          COMMIT
                                    UPDATE
                                               UPDATE
                                    A=100
OLD
      A=0
                                              B=50
                                                                   A = 80
                B=0
NEW
      A = 100
                B=50
                                    A = 80
                                               B=70
                                                                   A = 110
                                                        commits = \{\}
     cell storage
                        A 110
                                   В
                                       70
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
     if r.type == commit:
        commits.add(r.tid)
     if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```

```
TID
                            T1
                                     T2
                                               T2
                                                           T2
                                                        COMMIT
      UPDATE
                UPDATE
                          COMMIT
                                    UPDATE
                                              UPDATE
                                    A=100
OLD
      A=0
                                              B=50
                                                                  A = 80
                B=0
NEW
      A = 100
                B=50
                                    A = 80
                                              B=70
                                                                  A = 110
                                                        commits = \{\}
                                      70
    cell storage
                        A 80
                                   B
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
     if r.type == commit:
        commits.add(r.tid)
     if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```

```
TID
                                      T2
                                                T2
      UPDATE
                UPDATE
                          COMMIT
                                     UPDATE
                                               UPDATE
                                                         COMMIT
                                                                   UPDATE
                                     A=100
OLD
      A=0
                                               B=50
                B=0
                                                                   A = 80
NEW
      A = 100
                B=50
                                     A = 80
                                               B=70
                                                                   A = 110
                                                        commits = \{\}
     cell storage
                        A 80
                                       70
                                   B
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
     if r.type == commit:
        commits.add(r.tid)
     if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```

```
TID
                            T1
                                     T2
                                               T2
                                                        COMMIT
      UPDATE
                UPDATE
                          COMMIT
                                    UPDATE
                                              UPDATE
                                                                  UPDATE
                                    A=100
OLD
      A=0
                                              B=50
                B=0
                                                                  A = 80
NEW
      A = 100
                B=50
                                    A = 80
                                              B=70
                                                                  A = 110
                                                        commits = \{\}
    cell storage
                        A 80
                                      70
                                   B
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
        if r.type == commit:
        commits.add(r.tid)
        if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```

```
TID
                            T1
                                     T2
                                               T2
                                                        COMMIT
      UPDATE
                UPDATE
                          COMMIT
                                    UPDATE
                                              UPDATE
                                                                  UPDATE
                                    A=100
OLD
      A=0
                                              B=50
                B=0
                                                                  A = 80
NEW
      A = 100
                B=50
                                    A = 80
                                              B=70
                                                                  A = 110
                                                        commits = \{\}
    cell storage
                        A 80
                                      70
                                   B
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
      if r.type == commit:
        commits.add(r.tid)
      if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```

```
TID
                            T1
                                     T2
                                               T2
                                                        COMMIT
      UPDATE
                UPDATE
                          COMMIT
                                    UPDATE
                                              UPDATE
                                                                  UPDATE
                                    A=100
OLD
      A=0
                                              B=50
                B=0
                                                                  A = 80
NEW
      A = 100
                B=50
                                    A = 80
                                              B=70
                                                                  A = 110
                                                      commits = \{T2\}
    cell storage
                       A 80
                                  В
                                      70
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
      if r.type == commit:
        commits.add(r.tid)
      if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```

```
TID
                            T1
                                     T2
                                               T2
                                                        COMMIT
      UPDATE
                UPDATE
                          COMMIT
                                    UPDATE
                                              UPDATE
                                                                  UPDATE
                                    A=100
OLD
      A=0
                                              B=50
                B=0
                                                                  A = 80
NEW
      A = 100
                B=50
                                    A = 80
                                              B=70
                                                                  A = 110
                                                      commits = \{T2\}
    cell storage
                       A 80
                                      70
                                  B
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
      if r.type == commit:
        commits.add(r.tid)
      if r.type == update and r.tid not in commits:
        cell write(r.var, r.old val) // undo
```

```
TID
                                      T2
                                                T2
      UPDATE
                 UPDATE
                           COMMIT
                                     UPDATE
                                               UPDATE
                                                         COMMIT
                                                                   UPDATE
                                     A=100
OLD
      A=0
                 B=0
                                               B=50
                                                                   A = 80
NEW
      A = 100
                 B=50
                                     A = 80
                                               B = 70
                                                                   A = 110
                                                       commits = \{T2\}
     cell storage
                        A 80
                                       70
                                   B
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
     if r.type == commit:
        commits.add(r.tid)
     if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```

```
TID
                                      T2
                                                           T2
                                               UPDATE
      UPDATE
                UPDATE
                          COMMIT
                                                         COMMIT
                                    UPDATE
                                                                   UPDATE
                                    A=100
                                               B=50
OLD
                B=0
                                                                   A = 80
      A=0
NEW
      A = 100
                B=50
                                    A = 80
                                               B=70
                                                                   A = 110
                                                       commits = \{T2\}
     cell storage
                        A 80
                                       70
                                   B
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
     if r.type == commit:
        commits.add(r.tid)
     if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```

```
TID
                                     T2
                                                          T2
                                                                   T3
                                              UPDATE
      UPDATE
                UPDATE
                          COMMIT
                                                        COMMIT
                                    UPDATE
                                                                  UPDATE
                                    A=100
                                              B=50
OLD
      A=0
                                                                  A = 80
                B=0
                                                                  A=110
NEW
      A = 100
                B=50
                                    A = 80
                                              B=70
                                                      commits = \{T2\}
    cell storage
                        A 80
                                      70
                                  B
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
      if r.type == commit:
        commits.add(r.tid)
      if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```

```
TID
                            T1
                                     T2
                                                          T2
                                                                   T3
                                              UPDATE
      UPDATE
                UPDATE
                          COMMIT
                                                        COMMIT
                                    UPDATE
                                                                  UPDATE
                                    A=100
                                              B=50
OLD
      A=0
                                                                  A = 80
                B=0
NEW
      A = 100
                B=50
                                    A = 80
                                              B=70
                                                                  A = 110
                                                      commits = \{T2\}
    cell storage
                        A 80
                                      70
                                   B
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
      if r.type == commit:
        commits.add(r.tid)
      if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```

```
TID
                                      T2
                                                           T2
                                               UPDATE
      UPDATE
                UPDATE
                          COMMIT
                                                         COMMIT
                                    UPDATE
                                                                   UPDATE
                                    A=100
                                               B=50
OLD
                B=0
                                                                   A = 80
      A=0
NEW
      A = 100
                B=50
                                    A = 80
                                               B=70
                                                                   A = 110
                                                       commits = \{T2\}
     cell storage
                        A 80
                                       70
                                   B
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
     if r.type == commit:
        commits.add(r.tid)
     if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```

```
TID
                                                T2
                                                            T2
      UPDATE
                 UPDATE
                                               UPDATE
                                                          COMMIT
                           COMMIT
                                                                    UPDATE
                                     A=100
                                               B=50
OLD
                 B=0
                                                                    A = 80
      A=0
NEW
      A = 100
                 B=50
                                     A = 80
                                               B = 70
                                                                    A = 110
                                                        commits = \{T2\}
     cell storage
                        A 80
                                       70
                                   B
```

```
recover(log):
  commits = {}
  for record r in log[len(log)-1] .. log[0]:
    if r.type == commit:
       commits.add(r.tid)
    if r.type == update and r.tid not in commits:
       cell_write(r.var, r.old_val) // undo
```

```
TID
                                                T2
                                                           T2
      UPDATE
                UPDATE
                                               UPDATE
                                                         COMMIT
                          COMMIT
                                                                   UPDATE
                                     A=100
                                              B=50
OLD
                B=0
                                                                   A = 80
      A=0
NEW
      A = 100
                B=50
                                     A = 80
                                               B=70
                                                                   A = 110
                                                       commits = \{T2\}
     cell storage
                        A 80
                                       70
                                   B
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
        if r.type == commit:
        commits.add(r.tid)
        if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```

```
TID
                                                T2
                                                           T2
                                                                    T3
                          COMMIT
      UPDATE
                UPDATE
                                               UPDATE
                                                         COMMIT
                                                                   UPDATE
                                     A=100
                                               B=50
OLD
      A=0
                B=0
                                                                   A = 80
NEW
      A = 100
                B=50
                                     A = 80
                                               B=70
                                                                   A = 110
                                                       commits = \{T2\}
                        A 80
                                       70
     cell storage
                                   B
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
      if r.type == commit:
        commits.add(r.tid)
      if r.type == update and r.tid not in commits:
        cell write(r.var, r.old val) // undo
```

```
TID
                                                T2
                                                           T2
      UPDATE
                UPDATE
                                               UPDATE
                                                         COMMIT
                          COMMIT
                                                                   UPDATE
                                     A=100
                                               B=50
OLD
                B=0
                                                                   A = 80
      A=0
NEW
      A = 100
                B=50
                                     A = 80
                                               B=70
                                                                   A = 110
                                                       commits = \{T2\}
     cell storage
                        A 80
                                       70
                                   B
```

```
recover(log):
  commits = {}
  for record r in log[len(log)-1] .. log[0]:
    if r.type == commit:
       commits.add(r.tid)
    if r.type == update and r.tid not in commits:
       cell_write(r.var, r.old_val) // undo
```

```
TID
                                                T2
                                                           T2
                           COMMIT
                                              UPDATE
      UPDATE
                UPDATE
                                                         COMMIT
                                    UPDATE
                                                                   UPDATE
                                    A=100
                                              B=50
OLD
                B=0
                                                                   A = 80
      A=0
NEW
      A = 100
                B=50
                                    A = 80
                                              B=70
                                                                   A = 110
                                                       commits = \{T2\}
    cell storage
                        A 80
                                       70
                                   B
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
     if r.type == commit:
        commits.add(r.tid)
     if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```

```
TID
                                                T2
                                                           T2
                                                                    T3
                           COMMIT
                                               UPDATE
      UPDATE
                UPDATE
                                                         COMMIT
                                     UPDATE
                                                                   UPDATE
                                     A=100
OLD
      A=0
                B=0
                                              B=50
                                                                   A = 80
NEW
      A = 100
                B=50
                                     A = 80
                                               B=70
                                                                   A = 110
                                                       commits = \{T2\}
     cell storage
                        A 80
                                       70
                                   B
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
      if r.type == commit:
        commits.add(r.tid)
      if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```

```
TID
                                                T2
                                                           T2
                          COMMIT
                                              UPDATE
      UPDATE
                UPDATE
                                                         COMMIT
                                    UPDATE
                                                                   UPDATE
                                    A=100
                                              B=50
OLD
      A=0
                B=0
                                                                   A = 80
NEW
      A = 100
                B=50
                                    A = 80
                                              B=70
                                                                   A = 110
                                                       commits = \{T2\}
    cell storage
                        A 80
                                       70
                                   B
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
      if r.type == commit:
        commits.add(r.tid)
      if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```

```
TID
                                              T2
                                                         T2
                          COMMIT
                                             UPDATE
      UPDATE
                UPDATE
                                                       COMMIT
                                    UPDATE
                                                                 UPDATE
                                    A=100
OLD
                B=0
                                             B=50
                                                                 A = 80
      A=0
NEW
      A = 100
                B=50
                                    A=80
                                             B=70
                                                                 A = 110
                                                commits = \{T2,
    cell storage
                       A 80
                                  В
                                      70
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
     if r.type == commit:
        commits.add(r.tid)
     if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```

```
TID
                                               T2
                                                          T2
                          COMMIT
                                              UPDATE
      UPDATE
                UPDATE
                                                        COMMIT
                                    UPDATE
                                                                 UPDATE
                                    A=100
                                             B=50
OLD
                B=0
                                                                 A = 80
      A=0
NEW
      A = 100
                B=50
                                    A=80
                                              B = 70
                                                                 A = 110
                                                 commits = \{T2, T1\}
                       A 80
                                  В
                                      70
    cell storage
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
      if r.type == commit:
        commits.add(r.tid)
      if r.type == update and r.tid not in commits:
        cell write(r.var, r.old_val) // undo
```

```
TID
                                               T2
                                                          T2
                                              UPDATE
                                                        COMMIT
      UPDATE
                UPDATE
                          COMMIT
                                    UPDATE
                                                                  UPDATE
                                    A=100
                                              B=50
OLD
                B=0
                                                                  A = 80
      A=0
NEW
      A = 100
                B=50
                                    A = 80
                                              B=70
                                                                  A = 110
                                                 commits = \{T2, T1\}
    cell storage
                       A 80
                                  В
                                      70
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
     if r.type == commit:
        commits.add(r.tid)
     if r.type == update and r.tid not in commits:
        cell write(r.var, r.old val) // undo
```

```
TID
                                     T2
                                               T2
                                                          T2
                                              UPDATE
                                                        COMMIT
      UPDATE
                UPDATE
                          COMMIT
                                    UPDATE
                                                                  UPDATE
                                    A=100
OLD
      A=0
                B=0
                                              B=50
                                                                 A = 80
NEW
      A = 100
                B=50
                                    A = 80
                                              B=70
                                                                 A = 110
                                                 commits = \{T2, T1\}
    cell storage
                       A 80
                                  В
                                      70
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
      if r.type == commit:
        commits.add(r.tid)
      if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```

```
TID
                                    T2
                                              T2
                                                         T2
                                             UPDATE
      UPDATE
                UPDATE
                          COMMIT
                                                       COMMIT
                                   UPDATE
                                                                 UPDATE
                                   A=100
OLD
      A=0
                B=0
                                             B=50
                                                                 A = 80
                                                                 A=110
NEW
      A = 100
                B=50
                                   A = 80
                                             B=70
                                                commits = \{T2, T1\}
    cell storage
                       A 80
                                  В
                                     70
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
      if r.type == commit:
        commits.add(r.tid)
      if r.type == update and r.tid not in commits:
        cell write(r.var, r.old val) // undo
```

```
TID
                                    T2
                                              T2
                                                         T2
      UPDATE
                UPDATE
                          COMMIT
                                             UPDATE
                                                       COMMIT
                                   UPDATE
                                                                 UPDATE
                                   A=100
OLD
      A=0
                B=0
                                             B=50
                                                                 A = 80
                B=50
NEW
      A = 100
                                   A=80
                                             B=70
                                                                 A = 110
                                                commits = \{T2, T1\}
    cell storage
                       A 80
                                  В
                                     70
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
      if r.type == commit:
        commits.add(r.tid)
      if r.type == update and r.tid not in commits:
        cell write(r.var, r.old_val) // undo
```

```
TID
                                     T2
                                               T2
                                                          T2
                                              UPDATE
                                                        COMMIT
      UPDATE
                UPDATE
                          COMMIT
                                    UPDATE
                                                                  UPDATE
                                    A=100
OLD
      A=0
                B=0
                                              B=50
                                                                 A = 80
NEW
      A = 100
                B=50
                                    A = 80
                                              B=70
                                                                 A = 110
                                                 commits = \{T2, T1\}
    cell storage
                       A 80
                                  В
                                      70
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
      if r.type == commit:
        commits.add(r.tid)
      if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```

```
TID
                                    T2
                                              T2
                                                         T2
                                             UPDATE
                                                       COMMIT
      UPDATE
                UPDATE
                          COMMIT
                                   UPDATE
                                                                 UPDATE
                                   A=100
OLD
                B=0
                                             B=50
                                                                 A = 80
      A=0
NEW
      A=100
                B=50
                                   A = 80
                                             B=70
                                                                 A = 110
                                                commits = \{T2, T1\}
    cell storage
                       A 80
                                  В
                                     70
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
      if r.type == commit:
        commits.add(r.tid)
      if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```

```
TID
                                    T2
                                              T2
                                                        T2
                                             UPDATE
                                                      COMMIT
      UPDATE
                UPDATE
                         COMMIT
                                   UPDATE
                                                                UPDATE
                                   A=100
OLD
                B=0
                                            B=50
                                                                A = 80
      A=0
                                                                A=110
NEW
      A=100
                B=50
                                   A = 80
                                             B=70
                                                commits = \{T2, T1\}
    cell storage
                       A 80
                                 В
                                     70
```

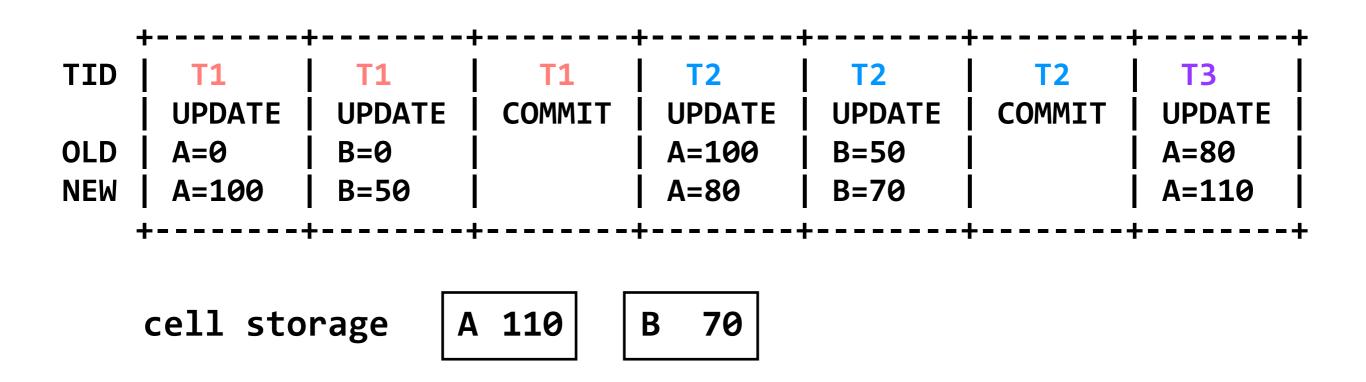
```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
        if r.type == commit:
        commits.add(r.tid)
        if r.type == update and r.tid not in commits:
        cell write(r.var, r.old_val) // undo
```

```
TID
                                    T2
                                              T2
                                                         T2
                          COMMIT
                                             UPDATE
      UPDATE
                UPDATE
                                   UPDATE
                                                       COMMIT
                                                                 UPDATE
                                   A=100
OLD
      A=0
                B=0
                                             B=50
                                                                 A = 80
NEW
      A=100
                B=50
                                   A = 80
                                             B=70
                                                                 A = 110
                                                commits = \{T2, T1\}
    cell storage
                       A 80
                                     70
                                  B
```

```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
      if r.type == commit:
        commits.add(r.tid)
      if r.type == update and r.tid not in commits:
        cell write(r.var, r.old_val) // undo
```

```
TID
                                    T2
                                              T2
                                                         T2
                                             UPDATE
                                                       COMMIT
      UPDATE
                UPDATE
                          COMMIT
                                   UPDATE
                                                                 UPDATE
                                   A=100
OLD
                B=0
                                             B=50
                                                                 A = 80
      A=0
NEW
      A=100
                B=50
                                   A = 80
                                             B=70
                                                                 A = 110
                                                commits = \{T2, T1\}
    cell storage
                       A 80
                                  В
                                     70
```

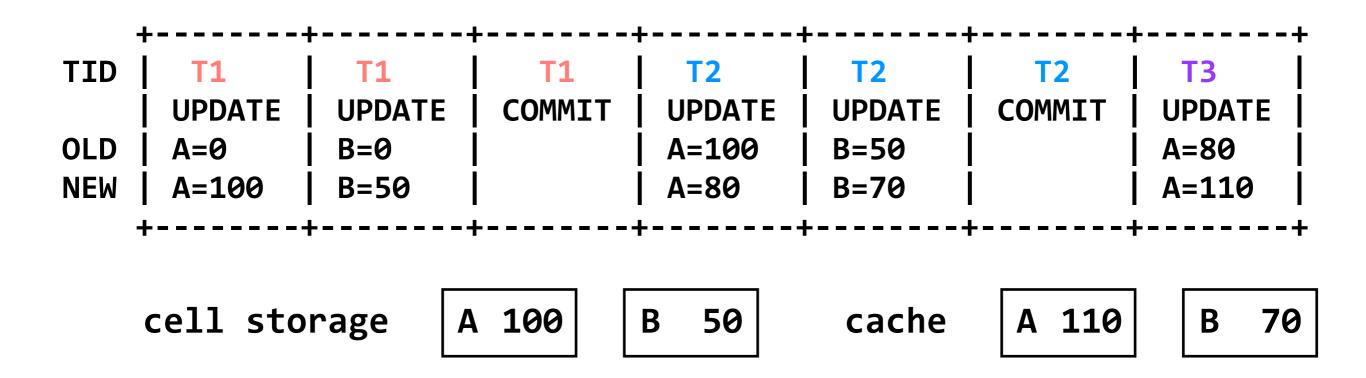
```
recover(log):
   commits = {}
   for record r in log[len(log)-1] .. log[0]:
      if r.type == commit:
        commits.add(r.tid)
      if r.type == update and r.tid not in commits:
        cell_write(r.var, r.old_val) // undo
```



performance?

problem: read performance is now great, but writes got (a little bit) slower and recovery got (a lot) slower

```
TID
                                 T2
                                          T2
                                         UPDATE
       UPDATE
               UPDATE
                        COMMIT
                                UPDATE
                                                 COMMIT
                                                          UPDATE
                                A=100
 OLD
       A=0
                                         B=50
               B=0
                                                          A = 80
               B=50
                                                          A = 110
 NEW
       A = 100
                                A = 80
                                         B=70
     cell storage
                     A 110
                               В
                                  70
                                         cache
                                                  A 110
                                                            B
                                                               70
read(var):
  if var in cache:
    return cache[var]
  else:
    // may evict others from cache to cell storage
    cache[var] = cell read(var)
    return cache[var]
write(var, value):
  log.append(current_tid, update, var, read(var), value)
  cache[var] = value
```



suppose we flushed the cache after **T1** committed, but have not flushed it since then

```
TID
                         T1
                                 T2
                                         T2
                                                   T2
                                        UPDATE
      UPDATE
               UPDATE
                       COMMIT
                                UPDATE
                                                 COMMIT
                                A=100
OLD
      A=0
               B=0
                                        B=50
                                                          A = 80
NEW
      A = 100
               B=50
                                A=80
                                        B = 70
                                                          A = 110
     cell storage
                     A 100
                               В
                                  50
                                         cache
recover(log):
  commits = \{\}
  for record r in log[len(log)-1] .. log[0]:
    if r.type == commit:
       commits.add(r.tid)
    if r.type == update and r.tid not in commits:
       cell write(r.var, r.old val) // undo
```

```
TID
                         T1
                                 T2
                                         T2
                                                   T2
                                        UPDATE
      UPDATE
               UPDATE
                       COMMIT
                                UPDATE
                                                 COMMIT
                                A=100
OLD
      A=0
               B=0
                                        B=50
                                                         A = 80
NEW
      A = 100
               B=50
                                A=80
                                        B=70
                                                         A = 110
                                         cache
     cell storage
                     A 80
                              B
                                  50
recover(log):
  commits = \{\}
  for record r in log[len(log)-1] .. log[0]:
    if r.type == commit:
       commits.add(r.tid)
    if r.type == update and r.tid not in commits:
       cell write(r.var, r.old val) // undo
```

```
TID
                         T1
                                 T2
                                         T2
                                        UPDATE
                                                 COMMIT
      UPDATE
               UPDATE
                       COMMIT
                                UPDATE
                                                          UPDATE
                                A=100
OLD
      A=0
               B=0
                                        B=50
                                                          A = 80
NEW
      A = 100
               B=50
                                A=80
                                        B = 70
                                                          A = 110
     cell storage
                     A 80
                               B
                                  50
                                         cache
recover(log):
  commits = \{\}
  for record r in log[len(log)-1] .. log[0]:
    if r.type == commit:
       commits.add(r.tid)
    if r.type == update and r.tid not in commits:
       cell write(r.var, r.old val) // undo
```

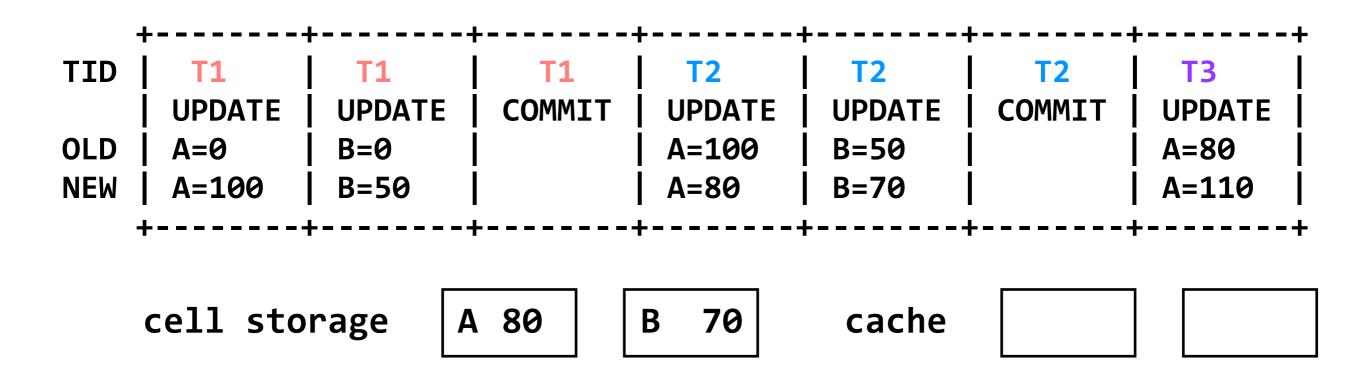
```
TID
                         T1
                                 T2
                                                   T2
                                                           T3
                                         UPDATE
      UPDATE
               UPDATE
                       COMMIT
                                                 COMMIT
                                UPDATE
                                                          UPDATE
                                A=100
                                         B=50
OLD
      A=0
               B=0
                                                          A = 80
NEW
      A = 100
               B=50
                                A = 80
                                         B=70
                                                          A = 110
                                         cache
     cell storage
                     A 80
                               B
                                  50
recover(log):
  commits = \{\}
  for record r in log[len(log)-1] .. log[0]:
    if r.type == commit:
       commits.add(r.tid)
    if r.type == update and r.tid not in commits:
       cell write(r.var, r.old val) // undo
```

```
TID
                        COMMIT
                                UPDATE
      UPDATE
               UPDATE
                                         UPDATE
                                                  COMMIT
                                                          UPDATE
                                         B=50
OLD
                                A = 100
      A=0
               B=0
                                                          A = 80
NEW
      A = 100
               B=50
                                A = 80
                                         B=70
                                                          A = 110
     cell storage
                     A 80
                               B
                                  50
                                          cache
recover(log):
  commits = \{\}
  for record r in log[len(log)-1] .. log[0]:
    if r.type == commit:
       commits.add(r.tid)
    if r.type == update and r.tid not in commits:
       cell write(r.var, r.old val) // undo
```

all other updates were committed; **B**'s value won't ever be changed

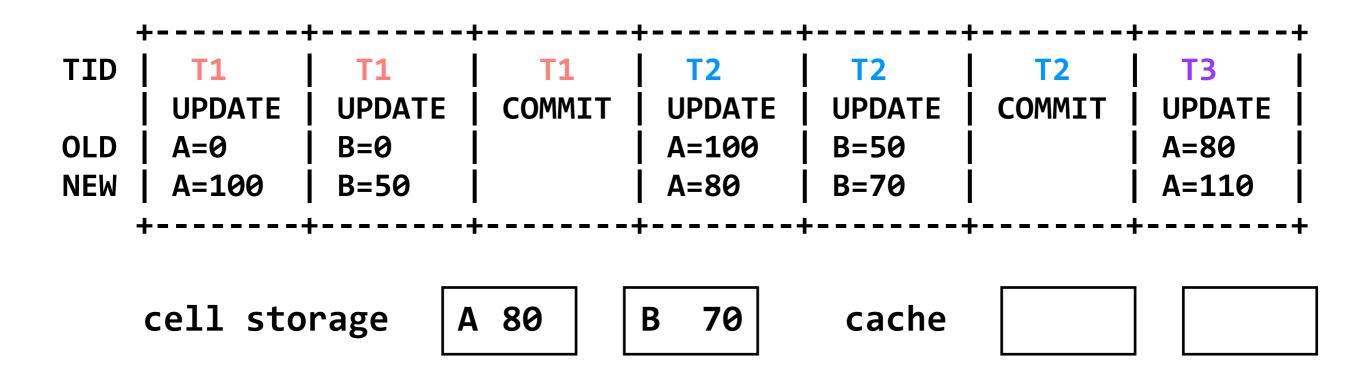
```
TID
                                        T2
              UPDATE
      UPDATE
                      COMMIT
                                       UPDATE
                               UPDATE
                                               COMMIT
                                                        UPDATE
OLD
              B=0
                               A = 100
                                      B=50
      A=0
                                                        A = 80
NEW
      A = 100
              B=50
                               A = 80
                                       B=70
                                                        A = 110
    cell storage
                             В
                                50
                                        cache
                    A 80
recover(log):
  commits = \{\}
  for record r in log[len(log)-1] .. log[0]:
    if r.type == commit:
      commits.add(r.tid)
    if r.type == update and r.tid not in commits:
      cell write(r.var, r.old val) // undo
  for record r in log[0] .. log[len(log)-1]:
    if r.type == update and r.tid in commits:
      cell write(r.var, r.new value) // redo
```

```
TID
                                        T2
              UPDATE
      UPDATE
                       COMMIT
                                       UPDATE
                               UPDATE
                                                COMMIT
                                                        UPDATE
OLD
              B=0
                               A = 100
                                       B=50
      A=0
                                                        A = 80
NEW
      A = 100
              B=50
                               A = 80
                                       B = 70
                                                        A = 110
    cell storage
                              В
                                 70
                                        cache
                    A 80
recover(log):
  commits = \{\}
  for record r in log[len(log)-1] .. log[0]:
    if r.type == commit:
      commits.add(r.tid)
    if r.type == update and r.tid not in commits:
      cell write(r.var, r.old val) // undo
  for record r in log[0] .. log[len(log)-1]:
    if r.type == update and r.tid in commits:
      cell write(r.var, r.new value) // redo
```



performance?

problem: recovery is still slow



performance?

solution: write checkpoints and truncate the log

- (Write-ahead) logs provide atomicity with better performance than shadow copies. The primary benefit is making small appends for each update, rather than copying and entire file over for every change.
- Cell storage is used with the log to improve readperformance, and caches and truncation can be used to improve write- and recovery-performance.