

# CSE 444: Database Internals

## Section 7:

### Transactions – Recovery with ARIES

# ARIES

- A popular protocol for UNDO-REDO logging
- **Steal (like UNDO)**
  - Changes by uncommitted transactions can be written to disk when a dirty page is flushed
- **No-force (like REDO)**
  - Changes by committed transactions may not have been written to disk
- **Write-ahead logging:**
  - Any changes to a database object is first recorded in the log, and the log is written to disk, before the change to database object is written to disk
  - A record of every change to the database is available while recovering from a crash

# ARIES Data Structures

**Dirty page table**

| pageID | recLSN |
|--------|--------|
|        |        |
|        |        |
|        |        |

LSN

101

**Transaction table**

| transID | lastLSN | status |
|---------|---------|--------|
|         |         |        |
|         |         |        |

**Log**

| prevLSN | tID | pID | Log entry | Type | undoNextLSN |
|---------|-----|-----|-----------|------|-------------|
|         |     |     |           |      |             |
|         |     |     |           |      |             |
|         |     |     |           |      |             |
|         |     |     |           |      |             |

# Log Record “Types”

- **Update:** easy
- **Commit:** log-tail forced-written to disk, up to & including commit (note that still no-force, the actual modified pages may not be written, and much smaller cost)
- **Abort:** abort type log record is written + undo is initiated for this transaction
- **End:** when a transaction is aborted or committed, some additional actions are performed, after that an end record is written
- **CLR:** Undoing updates (during abort or recovery from crash), for every update record undone, write a CLR (Compensation Log Record)

Example.

1.  $T_{1000}$  changes the value of **A** from “abc” to “def” on page P500
2.  $T_{2000}$  changes the value of **B** from “hij” to “klm” on page P600
3.  $T_{2000}$  changes the value of **D** from “mnp” to “qrs” on page P500
4.  $T_{1000}$  changes the value of **C** from “tuv” to “wxy” on page P505
5.  $T_{2000}$  commits and the end log record is written
6.  $T_{1000}$  changes the value of **E** from “pq” to “rs” on page P700
7. P600 is flushed to disk
8. **Crash!!**

Example is adopted from Ramakrishnan-Gehrke book

# ARIES Data Structures

Dirty page table

| pageID | recLSN |
|--------|--------|
|        |        |
|        |        |
|        |        |

LSN

101

Log

| prevLSN | tID | pID | Log entry | Type | undoNextLSN |
|---------|-----|-----|-----------|------|-------------|
|         |     |     |           |      |             |
|         |     |     |           |      |             |
|         |     |     |           |      |             |
|         |     |     |           |      |             |

Transaction table

| transID | lastLSN | status |
|---------|---------|--------|
|         |         |        |
|         |         |        |

Buffer Pool

|                                                    |                                          |
|----------------------------------------------------|------------------------------------------|
| <b>P500</b><br>PageLSN= -<br><br>A = abc   D = mnp | <b>P600</b><br>PageLSN= -<br><br>B = hij |
| <b>P505</b><br>PageLSN= -<br>C = tuv               | <b>P700</b><br>PageLSN= -<br>E = pq      |

Disk

|                                                    |                                          |
|----------------------------------------------------|------------------------------------------|
| <b>P500</b><br>PageLSN= -<br><br>A = abc   D = mnp | <b>P600</b><br>PageLSN= -<br><br>B = hij |
| <b>P505</b><br>PageLSN= -<br>C = tuv               | <b>P700</b><br>PageLSN= -<br>E = pq      |

## First operation:

1.  $T_{1000}$  changes the value of **A** from “abc” to “def” on **page P500**?

### Dirty page table

| pageID | recLSN |
|--------|--------|
|        |        |
|        |        |

LSN

101

### Log

| prevLSN | tID | pID | Log entry | Type | undoNextLSN |
|---------|-----|-----|-----------|------|-------------|
|         |     |     |           |      |             |
|         |     |     |           |      |             |
|         |     |     |           |      |             |
|         |     |     |           |      |             |

### Transaction table

| transID | lastLSN | status |
|---------|---------|--------|
|         |         |        |

### Buffer Pool

|                                                    |                                          |
|----------------------------------------------------|------------------------------------------|
| <b>P500</b><br>PageLSN= -<br><br>A = abc   D = mnp | <b>P600</b><br>PageLSN= -<br><br>B = hij |
| <b>P505</b><br>PageLSN= -<br><br>C = tuv           | <b>P700</b><br>PageLSN= -<br><br>E = pq  |

### Disk

|                                                    |                                          |
|----------------------------------------------------|------------------------------------------|
| <b>P500</b><br>PageLSN= -<br><br>A = abc   D = mnp | <b>P600</b><br>PageLSN= -<br><br>B = hij |
| <b>P505</b><br>PageLSN= -<br><br>C = tuv           | <b>P700</b><br>PageLSN= -<br><br>E = pq  |

## Changes

1.  $T_{1000}$  changes the value of **A** from “abc” to “def” on **page P500**

### Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
|        |        |
|        |        |

LSN  
101

### Log

| prevLSN | tID   | pID  | Log entry                 | Type   | undoNextLSN |
|---------|-------|------|---------------------------|--------|-------------|
| -       | T1000 | P500 | Write A<br>“abc” -> “def” | Update | -           |
|         |       |      |                           |        |             |
|         |       |      |                           |        |             |
|         |       |      |                           |        |             |

### Transaction table

| transID    | lastLSN | status  |
|------------|---------|---------|
| $T_{1000}$ | 101     | Running |
|            |         |         |

### Buffer Pool

|                                                    |                                          |
|----------------------------------------------------|------------------------------------------|
| <b>P500</b><br>PageLSN= 101<br><br>A = def D = mnp | <b>P600</b><br>PageLSN= -<br><br>B = hij |
| <b>P505</b><br>PageLSN= -<br>C = tuv               | <b>P700</b><br>PageLSN= -<br>E = pq      |

### Disk

|                                                  |                                          |
|--------------------------------------------------|------------------------------------------|
| <b>P500</b><br>PageLSN= -<br><br>A = abc D = mnp | <b>P600</b><br>PageLSN= -<br><br>B = hij |
| <b>P505</b><br>PageLSN= -<br>C = tuv             | <b>P700</b><br>PageLSN= -<br>E = pq      |



Next:

2.  $T_{2000}$  changes the value of **B** from “hij” to “klm” on **page P600**?

### Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
|        |        |
|        |        |

LSN  
101

### Log

| prevLSN | tID   | pID  | Log entry                 | Type   | undoNextLSN |
|---------|-------|------|---------------------------|--------|-------------|
| -       | T1000 | P500 | Write A<br>“abc” -> “def” | Update | -           |
|         |       |      |                           |        |             |
|         |       |      |                           |        |             |
|         |       |      |                           |        |             |

### Transaction table

| transID    | lastLSN | status  |
|------------|---------|---------|
| $T_{1000}$ | 101     | Running |
|            |         |         |

### Buffer Pool

|                                                    |                                          |
|----------------------------------------------------|------------------------------------------|
| <b>P500</b><br>PageLSN= 101<br><br>A = def D = mnp | <b>P600</b><br>PageLSN= -<br><br>B = hij |
| <b>P505</b><br>PageLSN= -<br>C = tuv               | <b>P700</b><br>PageLSN= -<br>E = pq      |

### Disk

|                                                  |                                          |
|--------------------------------------------------|------------------------------------------|
| <b>P500</b><br>PageLSN= -<br><br>A = abc D = mnp | <b>P600</b><br>PageLSN= -<br><br>B = hij |
| <b>P505</b><br>PageLSN= -<br>C = tuv             | <b>P700</b><br>PageLSN= -<br>E = pq      |

## Changes:

2.  $T_{2000}$  changes the value of **B** from “hij” to “klm” on **page P600**?

### Dirty page table

| pageID      | recLSN     |
|-------------|------------|
| P500        | 101        |
| <b>P600</b> | <b>102</b> |
|             |            |

### Transaction table

| transID                      | lastLSN    | status         |
|------------------------------|------------|----------------|
| $T_{1000}$                   | 101        | Running        |
| <b><math>T_{2000}</math></b> | <b>102</b> | <b>Running</b> |

### Log

LSN

101

102

| prevLSN | tID                          | pID         | Log entry                                  | Type          | undoNextLSN |
|---------|------------------------------|-------------|--------------------------------------------|---------------|-------------|
| -       | $T_{1000}$                   | P500        | Write A<br>“abc” -> “def”                  | Update        | -           |
| -       | <b><math>T_{2000}</math></b> | <b>P600</b> | <b>Write B</b><br><b>“hij” -&gt; “klm”</b> | <b>Update</b> |             |
|         |                              |             |                                            |               |             |
|         |                              |             |                                            |               |             |

### Buffer Pool

|                                                    |                                                          |
|----------------------------------------------------|----------------------------------------------------------|
| <b>P500</b><br>PageLSN= 101<br><br>A = def D = mnp | <b>P600</b><br><b>PageLSN= 102</b><br><br><b>B = klm</b> |
| <b>P505</b><br>PageLSN= -<br><br>C = tuv           | <b>P700</b><br>PageLSN= -<br><br>E = pq                  |

### Disk

|                                                  |                                          |
|--------------------------------------------------|------------------------------------------|
| <b>P500</b><br>PageLSN= -<br><br>A = abc D = mnp | <b>P600</b><br>PageLSN= -<br><br>B = hij |
| <b>P505</b><br>PageLSN= -<br><br>C = tuv         | <b>P700</b><br>PageLSN= -<br><br>E = pq  |

Next:

3.  $T_{2000}$  changes the value of **D** from “mnp” to “qrs” on **page P500**?

### Dirty page table

| pageID      | recLSN     |
|-------------|------------|
| P500        | 101        |
| <b>P600</b> | <b>102</b> |
|             |            |

### Transaction table

| transID                      | lastLSN    | status         |
|------------------------------|------------|----------------|
| $T_{1000}$                   | 101        | Running        |
| <b><math>T_{2000}</math></b> | <b>102</b> | <b>Running</b> |

### Log

LSN

101

102

| prevLSN | tID                          | pID         | Log entry                                  | Type          | undoNextLSN |
|---------|------------------------------|-------------|--------------------------------------------|---------------|-------------|
| -       | $T_{1000}$                   | P500        | Write A<br>“abc” -> “def”                  | Update        | -           |
| -       | <b><math>T_{2000}</math></b> | <b>P600</b> | <b>Write B</b><br><b>“hij” -&gt; “klm”</b> | <b>Update</b> |             |
|         |                              |             |                                            |               |             |
|         |                              |             |                                            |               |             |

### Buffer Pool

|                                                    |                                                          |
|----------------------------------------------------|----------------------------------------------------------|
| <b>P500</b><br>PageLSN= 101<br><br>A = def D = mnp | <b>P600</b><br><b>PageLSN= 102</b><br><br><b>B = klm</b> |
| <b>P505</b><br>PageLSN= -<br><br>C = tuv           | <b>P700</b><br>PageLSN= -<br><br>E = pq                  |

### Disk

|                                                  |                                          |
|--------------------------------------------------|------------------------------------------|
| <b>P500</b><br>PageLSN= -<br><br>A = abc D = mnp | <b>P600</b><br>PageLSN= -<br><br>B = hij |
| <b>P505</b><br>PageLSN= -<br><br>C = tuv         | <b>P700</b><br>PageLSN= -<br><br>E = pq  |

## Changes:

3.  $T_{2000}$  changes the value of **D** from “mnp” to “qrs” on page **P500**

### Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P600   | 102    |
|        |        |

### Transaction table

| transID    | lastLSN | status  |
|------------|---------|---------|
| $T_{1000}$ | 101     | Running |
| $T_{2000}$ | 103     | Running |

### Log

LSN

101

102

103

| prevLSN | tID        | pID  | Log entry                 | Type   | undoNextLSN |
|---------|------------|------|---------------------------|--------|-------------|
| -       | $T_{1000}$ | P500 | Write A<br>“abc” -> “def” | Update | -           |
| -       | $T_{2000}$ | P600 | Write B<br>“hij” -> “klm” | Update | -           |
| 102     | $T_{2000}$ | P500 | Write D<br>“mnp” -> “qrs” | Update | -           |
|         |            |      |                           |        |             |

### Buffer Pool

|                                                       |                                        |
|-------------------------------------------------------|----------------------------------------|
| <b>P500</b><br>PageLSN= 103<br>A = def <b>D = qrs</b> | <b>P600</b><br>PageLSN= 102<br>B = klm |
| <b>P505</b><br>PageLSN= -<br>C = tuv                  | <b>P700</b><br>PageLSN= -<br>E = pq    |

### Disk

|                                              |                                      |
|----------------------------------------------|--------------------------------------|
| <b>P500</b><br>PageLSN= -<br>A = abc D = mnp | <b>P600</b><br>PageLSN= -<br>B = hij |
| <b>P505</b><br>PageLSN= -<br>C = tuv         | <b>P700</b><br>PageLSN= -<br>E = pq  |

Next:

4.  $T_{1000}$  changes the value of **C** from “tuv” to “wxy” on page P505?

### Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P600   | 102    |
|        |        |

LSN

101

102

103

### Log

| prevLSN | tID               | pID  | Log entry                 | Type   | undoNextLSN |
|---------|-------------------|------|---------------------------|--------|-------------|
| -       | T1000             | P500 | Write A<br>“abc” -> “def” | Update | -           |
| -       | T <sub>2000</sub> | P600 | Write B<br>“hij” -> “klm” | Update | -           |
| 102     | T <sub>2000</sub> | P500 | Write D<br>“mnp” -> “qrs” | Update | -           |
|         |                   |      |                           |        |             |

### Transaction table

| transID           | lastLSN | status  |
|-------------------|---------|---------|
| T <sub>1000</sub> | 101     | Running |
| T <sub>2000</sub> | 103     | Running |

### Buffer Pool

|                                                       |                                        |
|-------------------------------------------------------|----------------------------------------|
| <b>P500</b><br>PageLSN= 103<br>A = def <b>D = qrs</b> | <b>P600</b><br>PageLSN= 102<br>B = klm |
| <b>P505</b><br>PageLSN= -<br>C = tuv                  | <b>P700</b><br>PageLSN= -<br>E = pq    |

### Disk

|                                              |                                      |
|----------------------------------------------|--------------------------------------|
| <b>P500</b><br>PageLSN= -<br>A = abc D = mnp | <b>P600</b><br>PageLSN= -<br>B = hij |
| <b>P505</b><br>PageLSN= -<br>C = tuv         | <b>P700</b><br>PageLSN= -<br>E = pq  |

## Changes:

4.  $T_{1000}$  changes the value of **C** from “tuv” to “wxy” on **page P505**

### Dirty page table

| pageID      | recLSN     |
|-------------|------------|
| P500        | 101        |
| P600        | 102        |
| <b>P505</b> | <b>104</b> |

### Transaction table

| transID    | lastLSN    | status         |
|------------|------------|----------------|
| $T_{1000}$ | <b>104</b> | <b>Running</b> |
| $T_{2000}$ | 103        | Running        |

### Log

| LSN | prevLSN    | tID                          | pID         | Log entry                                  | Type          | undoNextLSN |
|-----|------------|------------------------------|-------------|--------------------------------------------|---------------|-------------|
|     | -          | $T_{1000}$                   | P500        | Write A<br>“abc” -> “def”                  | Update        | -           |
| 101 | -          | $T_{2000}$                   | P600        | Write B<br>“hij” -> “klm”                  | Update        | -           |
| 102 | -          | $T_{2000}$                   | P500        | Write D<br>“mnp” -> “qrs”                  | Update        | -           |
| 103 | 102        | $T_{2000}$                   | P500        | Write D<br>“mnp” -> “qrs”                  | Update        | -           |
| 104 | <b>101</b> | <b><math>T_{1000}</math></b> | <b>P505</b> | <b>Write C</b><br><b>“tuv” -&gt; “wxy”</b> | <b>Update</b> | -           |

### Buffer Pool

|                                                      |                                        |
|------------------------------------------------------|----------------------------------------|
| <b>P500</b><br>PageLSN= 103<br>A = def D = qrs       | <b>P600</b><br>PageLSN= 102<br>B = klm |
| <b>P505</b><br><b>PageLSN= 104</b><br><b>C = wxy</b> | <b>P700</b><br>PageLSN= -<br>E = pq    |

### Disk

|                                              |                                      |
|----------------------------------------------|--------------------------------------|
| <b>P500</b><br>PageLSN= -<br>A = abc D = mnp | <b>P600</b><br>PageLSN= -<br>B = hij |
| <b>P505</b><br>PageLSN= -<br>C = tuv         | <b>P700</b><br>PageLSN= -<br>E = pq  |

Next:

5.  $T_{2000}$  commits and the end log record is written?

### Dirty page table

| pageID      | recLSN     |
|-------------|------------|
| P500        | 101        |
| P600        | 102        |
| <b>P505</b> | <b>104</b> |

### Transaction table

| transID    | lastLSN | status  |
|------------|---------|---------|
| $T_{1000}$ | 104     | Running |
| $T_{2000}$ | 103     | Running |

LSN

101

102

103

104

### Log

| prevLSN | tlD        | pID  | Log entry                 | Type   | undoNextLSN |
|---------|------------|------|---------------------------|--------|-------------|
| -       | $T_{1000}$ | P500 | Write A<br>"abc" -> "def" | Update | -           |
| -       | $T_{2000}$ | P600 | Write B<br>"hij" -> "klm" | Update | -           |
| 102     | $T_{2000}$ | P500 | Write D<br>"mnp" -> "qrs" | Update | -           |
| 101     | $T_{1000}$ | P505 | Write C<br>"tuv" -> "wxy" | Update | -           |

### Buffer Pool

|                                                |                                        |
|------------------------------------------------|----------------------------------------|
| <b>P500</b><br>PageLSN= 103<br>A = def D = qrs | <b>P600</b><br>PageLSN= 102<br>B = klm |
| <b>P505</b><br><b>PageLSN= 104</b><br>C = wxy  | <b>P700</b><br>PageLSN= -<br>E = pq    |

### Disk

|                                              |                                      |
|----------------------------------------------|--------------------------------------|
| <b>P500</b><br>PageLSN= -<br>A = abc D = mnp | <b>P600</b><br>PageLSN= -<br>B = hij |
| <b>P505</b><br>PageLSN= -<br>C = tuv         | <b>P700</b><br>PageLSN= -<br>E = pq  |

# Changes:

5.  $T_{2000}$  commits and the end log record is written --- step 1

## Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P600   | 102    |
| P505   | 104    |

## Transaction table

| transID    | lastLSN | status    |
|------------|---------|-----------|
| $T_{1000}$ | 104     | Running   |
| $T_{2000}$ | 103     | Committed |

LSN

101

102

103

104

105

106

| prevLSN | tID        | pID  | Log entry                 | Type   | undoNextLSN |
|---------|------------|------|---------------------------|--------|-------------|
| -       | $T_{1000}$ | P500 | Write A<br>"abc" -> "def" | Update | -           |
| -       | $T_{2000}$ | P600 | Write B<br>"hij" -> "klm" | Update | -           |
| 102     | $T_{2000}$ | P500 | Write D<br>"mnp" -> "qrs" | Update | -           |
| 101     | $T_{1000}$ | P505 | Write C<br>"tuv" -> "wxy" | Update | -           |
| 103     | $T_{2000}$ |      |                           | Commit |             |
| 105     | $T_{2000}$ |      |                           | End    |             |

## Buffer Pool

|                                                |                                        |
|------------------------------------------------|----------------------------------------|
| <b>P500</b><br>PageLSN= 103<br>A = def D = qrs | <b>P600</b><br>PageLSN= 102<br>B = klm |
| <b>P505</b><br>PageLSN= 104<br>C = wxy         | <b>P700</b><br>PageLSN= -<br>E = pq    |

## Disk

|                                              |                                      |
|----------------------------------------------|--------------------------------------|
| <b>P500</b><br>PageLSN= -<br>A = abc D = mnp | <b>P600</b><br>PageLSN= -<br>B = hij |
| <b>P505</b><br>PageLSN= -<br>C = tuv         | <b>P700</b><br>PageLSN= -<br>E = pq  |



# Changes:

5. **T<sub>2000</sub>** commits and the end log record is written --- step 2

## Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P600   | 102    |
| P505   | 104    |

T2000 removed from transaction table

## Transaction table

| transID                     | lastLSN        | status               |
|-----------------------------|----------------|----------------------|
| T <sub>1000</sub>           | 104            | Running              |
| <del>T<sub>2000</sub></del> | <del>102</del> | <del>Committed</del> |

LSN

| prevLSN | tID               | pID  | Log entry                 | Type   | undoNextLSN |
|---------|-------------------|------|---------------------------|--------|-------------|
| -       | T <sub>1000</sub> | P500 | Write A<br>"abc" -> "def" | Update | -           |
| -       | T <sub>1000</sub> | P600 | Write B<br>"hij" -> "klm" | Update | -           |
| -       | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs" | Update | -           |
| 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy" | Update | -           |
| 103     | T <sub>2000</sub> |      |                           | Commit |             |
| 105     | T <sub>2000</sub> |      |                           | End    |             |

## Buffer Pool

|                                                |                                        |
|------------------------------------------------|----------------------------------------|
| <b>P500</b><br>PageLSN= 103<br>A = def D = qrs | <b>P600</b><br>PageLSN= 102<br>B = klm |
| <b>P505</b><br>PageLSN= 104<br>C = wxy         | <b>P700</b><br>PageLSN= -<br>E = pq    |

## Disk

|                                              |                                      |
|----------------------------------------------|--------------------------------------|
| <b>P500</b><br>PageLSN= -<br>A = abc D = mnp | <b>P600</b><br>PageLSN= -<br>B = hij |
| <b>P505</b><br>PageLSN= -<br>C = tuv         | <b>P700</b><br>PageLSN= -<br>E = pq  |

# Changes:

5.  $T_{2000}$  commits and the end log record is written --- step 2

Log

## Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P600   | 102    |
| P505   | 104    |

## Transaction table

| transID    | lastLSN | status    |
|------------|---------|-----------|
| $T_{1000}$ | 104     | Running   |
| $T_{2000}$ | 103     | Committed |

| LSN | prevLSN | tID        | pID  | Log entry                 | Type   | undoNextLSN |
|-----|---------|------------|------|---------------------------|--------|-------------|
| 101 | -       | $T_{1000}$ | P500 | Write A<br>"abc" -> "def" | Update | -           |
| 102 | -       | $T_{2000}$ | P600 | Write B<br>"hij" -> "klm" | Update | -           |
| 103 | 102     | $T_{2000}$ | P500 | Write D<br>"mnp" -> "qrs" | Update | -           |
| 104 | 101     | $T_{1000}$ | P505 | Write C<br>"tuv" -> "wxy" | Update | -           |
| 105 | 103     | $T_{2000}$ |      |                           | Commit |             |
| 106 | 105     | $T_{2000}$ |      |                           | End    |             |

## Buffer Pool

|                                                |                                        |
|------------------------------------------------|----------------------------------------|
| <b>P500</b><br>PageLSN= 103<br>A = def D = qrs | <b>P600</b><br>PageLSN= 102<br>B = klm |
| <b>P505</b><br>PageLSN= 104<br>C = wxy         | <b>P700</b><br>PageLSN= -<br>E = pq    |

## Disk

Log written to disk  
Note: no force = not the dirty  
pages changed by  $T_{2000}$ !

|                                      |                                     |
|--------------------------------------|-------------------------------------|
| A = abc D = mnp                      | B = hij                             |
| <b>P505</b><br>PageLSN= -<br>C = tuv | <b>P700</b><br>PageLSN= -<br>E = pq |

Whenever a transaction commits,  
log is flushed to the disk == the log-tail is written to disk

**NOTE:**

1. The “Commit” record is required to be flushed (i.e. all logs up to and including that commit record)
2. The “End” record is not required to be flushed, in this case we are only assuming that it has been flushed as well (so that we have a good example while doing recovery 😊)

Next:

6.  $T_{1000}$  changes the value of **E** from “pq” to “rs” on **page P700**?  
Log

### Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P600   | 102    |
| P505   | 104    |

### Transaction table

| transID    | lastLSN | status  |
|------------|---------|---------|
| $T_{1000}$ | 104     | Running |
|            |         |         |

| LSN | prevLSN | tID        | pID  | Log entry                 | Type   | undoNextLSN |
|-----|---------|------------|------|---------------------------|--------|-------------|
| 101 | -       | $T_{1000}$ | P500 | Write A<br>“abc” -> “def” | Update | -           |
| 102 | -       | $T_{2000}$ | P600 | Write B<br>“hij” -> “klm” | Update | -           |
| 103 | 102     | $T_{2000}$ | P500 | Write D<br>“mnp” -> “qrs” | Update | -           |
| 104 | 101     | $T_{1000}$ | P505 | Write C<br>“tuv” -> “wxy” | Update | -           |
| 105 | 103     | $T_{2000}$ |      |                           | Commit |             |
| 106 | 105     | $T_{2000}$ |      |                           | End    |             |

### Buffer Pool

|                                                      |                                            |
|------------------------------------------------------|--------------------------------------------|
| <b>P500</b><br>PageLSN= 103<br><br>A = def   D = qrs | <b>P600</b><br>PageLSN= 102<br><br>B = klm |
| <b>P505</b><br>PageLSN= 104<br><br>C = wxy           | <b>P700</b><br>PageLSN= -<br><br>E = pq    |

### Disk

|                                                    |                                          |
|----------------------------------------------------|------------------------------------------|
| <b>P500</b><br>PageLSN= -<br><br>A = abc   D = mnp | <b>P600</b><br>PageLSN= -<br><br>B = hij |
| <b>P505</b><br>PageLSN= -<br><br>C = tuv           | <b>P700</b><br>PageLSN= -<br><br>E = pq  |

# Changes:

6.  $T_{1000}$  changes the value of E from "pq" to "rs" on page P700

## Dirty page table

| pageID      | recLSN     |
|-------------|------------|
| P500        | 101        |
| P600        | 102        |
| P505        | 104        |
| <b>P700</b> | <b>107</b> |

## Transaction table

| transID    | lastLSN    | status         |
|------------|------------|----------------|
| $T_{1000}$ | <b>107</b> | <b>Running</b> |
|            |            |                |

LSN

|            | prevLSN    | tID        | pID         | Log entry                 | Type          | undoNextLSN |
|------------|------------|------------|-------------|---------------------------|---------------|-------------|
| 101        | -          | $T_{1000}$ | P500        | Write A<br>"abc" -> "def" | Update        | -           |
| 102        | -          | $T_{2000}$ | P600        | Write B<br>"hij" -> "klm" | Update        | -           |
| 103        | 102        | $T_{2000}$ | P500        | Write D<br>"mnp" -> "qrs" | Update        | -           |
| 104        | 101        | $T_{1000}$ | P505        | Write C<br>"tuv" -> "wxy" | Update        | -           |
| 105        | 103        | $T_{2000}$ |             |                           | Commit        |             |
| 106        | 105        | $T_{2000}$ |             |                           | End           |             |
| <b>107</b> | <b>104</b> | $T_{1000}$ | <b>P700</b> | Write E<br>"pq" -> "rs"   | <b>Update</b> | -           |

## Buffer Pool

|                                                    |                                                         |
|----------------------------------------------------|---------------------------------------------------------|
| <b>P500</b><br>PageLSN= 103<br><br>A = def D = qrs | <b>P600</b><br>PageLSN= 102<br><br>B = klm              |
| <b>P505</b><br>PageLSN= 104<br><br>C = wxy         | <b>P700</b><br><b>PageLSN= 107</b><br><br><b>E = rs</b> |

## Disk

|                                                  |                                          |
|--------------------------------------------------|------------------------------------------|
| <b>P500</b><br>PageLSN= -<br><br>A = abc D = mnp | <b>P600</b><br>PageLSN= -<br><br>B = hij |
| <b>P505</b><br>PageLSN= -<br><br>C = tuv         | <b>P700</b><br>PageLSN= -<br><br>E = pq  |

Next:

## 7. Page P600 is flushed to disk?

### Dirty page table

| pageID      | recLSN     |
|-------------|------------|
| P500        | 101        |
| P600        | 102        |
| P505        | 104        |
| <b>P700</b> | <b>107</b> |

### Transaction table

| transID                 | lastLSN    | status         |
|-------------------------|------------|----------------|
| <b>T<sub>1000</sub></b> | <b>107</b> | <b>Running</b> |
|                         |            |                |

LSN

|     | prevLSN | tID               | pID  | Log entry                 | Type   | undoNextLSN |
|-----|---------|-------------------|------|---------------------------|--------|-------------|
| 101 | -       | T <sub>1000</sub> | P500 | Write A<br>"abc" -> "def" | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>"hij" -> "klm" | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs" | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy" | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                           | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                           | End    |             |
| 107 | 104     | T <sub>1000</sub> | P700 | Write E<br>"pq" -> "rs"   | Update | -           |

### Buffer Pool

|                                                    |                                                     |
|----------------------------------------------------|-----------------------------------------------------|
| <b>P500</b><br>PageLSN= 103<br><br>A = def D = qrs | <b>P600</b><br>PageLSN= 102<br><br>B = klm          |
| <b>P505</b><br>PageLSN= 104<br><br>C = wxy         | <b>P700</b><br><b>PageLSN= 107</b><br><b>E = rs</b> |

### Disk

|                                                  |                                          |
|--------------------------------------------------|------------------------------------------|
| <b>P500</b><br>PageLSN= -<br><br>A = abc D = mnp | <b>P600</b><br>PageLSN= -<br><br>B = hij |
| <b>P505</b><br>PageLSN= -<br><br>C = tuv         | <b>P700</b><br>PageLSN= -<br><br>E = pq  |

Next:

Pink = On disk

White = In memory

## 7. Page P600 is flushed to disk – Step 1

### Dirty page table

| pageID          | recLSN         |
|-----------------|----------------|
| P500            | 101            |
| <del>P600</del> | <del>102</del> |
| P505            | 104            |
| P700            | 107            |

### Transaction table

| transID           | lastLSN | status  |
|-------------------|---------|---------|
| T <sub>1000</sub> | 107     | Running |
|                   |         |         |

| LSN | prevLSN | tlD               | plD  | Log entry                 | Type   | undoNextLSN |
|-----|---------|-------------------|------|---------------------------|--------|-------------|
| 101 | -       | T <sub>1000</sub> | P500 | Write A<br>“abc” -> “def” | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>“hij” -> “klm” | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>“mnp” -> “qrs” | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>“tuv” -> “wxy” | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                           | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                           | End    |             |
| 107 | 104     | T <sub>1000</sub> | P700 | Write E<br>“pq” -> “rs”   | Update | -           |

### Buffer Pool

|                                                |                                        |
|------------------------------------------------|----------------------------------------|
| <b>P500</b><br>PageLSN= 103<br>A = def D = qrs | <b>P600</b><br>PageLSN= 102<br>B = klm |
| <b>P505</b><br>PageLSN= 104<br>C = wxy         | <b>P700</b><br>PageLSN= 107<br>E = rs  |

### Disk

|                                              |                                        |
|----------------------------------------------|----------------------------------------|
| <b>P500</b><br>PageLSN= -<br>A = abc D = mnp | <b>P600</b><br>PageLSN= 102<br>B = klm |
| <b>P505</b><br>PageLSN= -<br>C = tuv         | <b>P700</b><br>PageLSN= -<br>E = pq    |

Next:

## 7. Page P600 is flushed to disk – Step 2

### Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P505   | 104    |
| P700   | 107    |

### Transaction table

| transID           | lastLSN | status  |
|-------------------|---------|---------|
| T <sub>1000</sub> | 107     | Running |
|                   |         |         |

| LSN | prevLSN | tlD               | plD  | Log entry                 | Type   | undoNextLSN |
|-----|---------|-------------------|------|---------------------------|--------|-------------|
| 101 | -       | T <sub>1000</sub> | P500 | Write A<br>“abc” -> “def” | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>“hij” -> “klm” | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>“mnp” -> “qrs” | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>“tuv” -> “wxy” | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                           | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                           | End    |             |
| 107 | 104     | T <sub>1000</sub> | P700 | Write E<br>“pq” -> “rs”   | Update | -           |

### Buffer Pool

|                                                |                                        |
|------------------------------------------------|----------------------------------------|
| <b>P500</b><br>PageLSN= 103<br>A = def D = qrs | <b>P600</b><br>PageLSN= 102<br>B = klm |
| <b>P505</b><br>PageLSN= 104<br>C = wxy         | <b>P700</b><br>PageLSN= 107<br>E = rs  |

### Disk

|                                              |                                                      |
|----------------------------------------------|------------------------------------------------------|
| <b>P500</b><br>PageLSN= -<br>A = abc D = mnp | <b>P600</b><br><b>PageLSN= 102</b><br><b>B = klm</b> |
| <b>P505</b><br>PageLSN= -<br>C = tuv         | <b>P700</b><br>PageLSN= -<br>E = pq                  |



## “Flush page P600 to disk”

1. Write current content of P600 to disk (along with pageLSN)
2. Remove from Dirty Page table

### **NOTE:** Write Ahead Log

1. All LSNs changing that page must be written to disk
2. In this case it is okay, since the last log record involves P700 while P600 is being flushed
3. When a page is written, we need to ensure that all log records up to the lastLSN of the last transaction that ever wrote to that page are on disk

**(Log is always written to disk in order, i.e. we can never skip some log entries in between!)**



8. CRASH!!

8. Crash!! ---- These are gone from memory

Dirty page table

|      |
|------|
| P700 |
|------|

Transaction table

| transID | lastLSN | status |
|---------|---------|--------|
|         |         | ing    |

| LSN | prevLSN | tlD               | plD  | Log entry                 | Type   | undoNextLSN |
|-----|---------|-------------------|------|---------------------------|--------|-------------|
| 101 | -       | T1000             | P500 | Write A<br>"abc" -> "def" | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>"hij" -> "klm" | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs" | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy" | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                           | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                           | End    |             |
| 107 |         |                   |      |                           |        |             |

Buffer Pool

|                                        |
|----------------------------------------|
| <b>P500</b><br>PageLSN= 103<br>A = def |
| <b>P505</b><br>PageLSN= 107<br>C = wxy |

Disk

|                                              |                                        |
|----------------------------------------------|----------------------------------------|
| <b>P500</b><br>PageLSN= -<br>A = abc D = mnp | <b>P600</b><br>PageLSN= 102<br>B = klm |
| <b>P505</b><br>PageLSN= -<br>C = tuv         | <b>P700</b><br>PageLSN= -<br>E = pq    |

# Three Recovery Phases of ARIES

- Analysis
  - Reconstructs Dirty page table (for Redo) + Transaction Table (for Undo, active transactions at crash)
- Redo
  - Restores the database state at the time of crash by repeating Updates + CLR
- Undo
  - Undoes the actions of uncommitted transactions
  - Only updates can be undone, No CLR is ever undone!

# Recall: ARIES Method Illustration

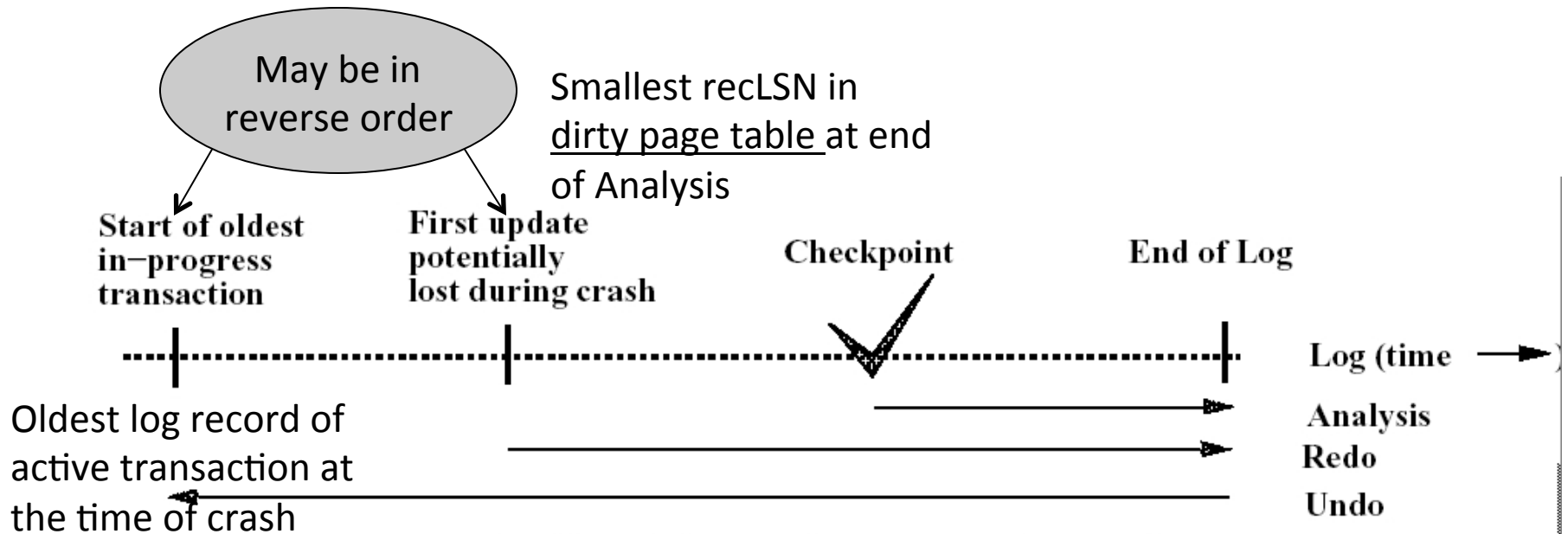


Figure 3: The Three Passes of ARIES Restart

(Trace back using  
transaction Table)

[Franklin97]

# About Checkpointing

- This example has no checkpointing == Checkpointing at the beginning (like hw5 !)
- Analysis phase in the recovery starts with empty Dirty Page table and empty Transaction Table
  - Otherwise the latest copies of these tables have to be read from disk from the last checkpoint

# Analysis Phase

- Reconstruct (conservatively) Dirty Page Table and Transaction Table
- Read the log from the last checkpoint
- Actual pages are not used

# Analysis Phase

## Log

### Dirty page table

| pageID | recLSN |
|--------|--------|
|        |        |
|        |        |
|        |        |
|        |        |

### Transaction table

| transID | lastLSN | status |
|---------|---------|--------|
|         |         |        |
|         |         |        |
|         |         |        |

| LSN | prevLSN | tID               | pID  | Log entry                 | Type   | undoNextLSN |
|-----|---------|-------------------|------|---------------------------|--------|-------------|
| 101 | -       | T1000             | P500 | Write A<br>"abc" -> "def" | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>"hij" -> "klm" | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs" | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy" | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                           | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                           | End    |             |
|     |         |                   |      |                           |        |             |

Disk

### Buffer Pool

|  |  |
|--|--|
|  |  |
|  |  |

Will not show the buffer pool from the next slide

#### P500

PageLSN= -

A = abc D = mnp

#### P600

PageLSN= 102

B = klm

#### P505

PageLSN= -

C = tuv

#### P700

PageLSN= -

E = pq



# Analysis Phase

## Log

### Dirty page table

| pageID      | recLSN     |
|-------------|------------|
| <b>P500</b> | <b>101</b> |
|             |            |
|             |            |
|             |            |

### Transaction table

| transID      | lastLSN    | status               |
|--------------|------------|----------------------|
| <b>T1000</b> | <b>101</b> | <b>U=</b><br>Unknown |
|              |            |                      |
|              |            |                      |

| LSN | prevLSN | tID               | pID  | Log entry                 | Type   | undoNextLSN |
|-----|---------|-------------------|------|---------------------------|--------|-------------|
| 101 | -       | T1000             | P500 | Write A<br>"abc" -> "def" | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>"hij" -> "klm" | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs" | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy" | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                           | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                           | End    |             |
|     |         |                   |      |                           |        |             |

|                           |                            |
|---------------------------|----------------------------|
| <b>P500</b><br>PageLSN= - | <b>P600</b><br>PageLSN=102 |
| A = abc D = mnp           | B = klm                    |
| <b>P505</b><br>PageLSN= - | <b>P700</b><br>PageLSN= -  |
| C = tuv                   | E = pq                     |

**Disk**

In hw5, you can also write "Running/In progress" instead of "Unknown"

# Analysis Phase

## Log

### Dirty page table

| pageID      | recLSN     |
|-------------|------------|
| P500        | 101        |
| <b>P600</b> | <b>102</b> |
|             |            |
|             |            |

### Transaction table

| transID      | lastLSN    | status   |
|--------------|------------|----------|
| T1000        | 101        | U        |
| <b>T2000</b> | <b>102</b> | <b>U</b> |

| LSN | prevLSN | tID               | pID  | Log entry                 | Type   | undoNextLSN |
|-----|---------|-------------------|------|---------------------------|--------|-------------|
| 101 | -       | T1000             | P500 | Write A<br>"abc" -> "def" | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>"hij" -> "klm" | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs" | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy" | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                           | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                           | End    |             |
|     |         |                   |      |                           |        |             |

|                           |                            |
|---------------------------|----------------------------|
| <b>P500</b><br>PageLSN= - | <b>P600</b><br>PageLSN=102 |
| A = abc D = mnp           | B = klm                    |
| <b>P505</b><br>PageLSN= - | <b>P700</b><br>PageLSN= -  |
| C = tuv                   | E = pq                     |

**Disk**

# Analysis Phase

## Log

### Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P600   | 102    |
|        |        |
|        |        |

### Transaction table

| transID      | lastLSN    | status   |
|--------------|------------|----------|
| T1000        | 101        | U        |
| <b>T2000</b> | <b>103</b> | <b>U</b> |

| LSN | prevLSN | tID               | pID  | Log entry                 | Type   | undoNextLSN |
|-----|---------|-------------------|------|---------------------------|--------|-------------|
| 101 | -       | T1000             | P500 | Write A<br>"abc" -> "def" | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>"hij" -> "klm" | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs" | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy" | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                           | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                           | End    |             |
|     |         |                   |      |                           |        |             |

|                           |                            |
|---------------------------|----------------------------|
| <b>P500</b><br>PageLSN= - | <b>P600</b><br>PageLSN=102 |
| A = abc D = mnp           | B = klm                    |
| <b>P505</b><br>PageLSN= - | <b>P700</b><br>PageLSN= -  |
| C = tuv                   | E = pq                     |

**Disk**

# Analysis Phase

## Log

### Dirty page table

| pageID      | recLSN     |
|-------------|------------|
| P500        | 101        |
| P600        | 102        |
| <b>P505</b> | <b>104</b> |
|             |            |

### Transaction table

| transID      | lastLSN    | status   |
|--------------|------------|----------|
| <b>T1000</b> | <b>104</b> | <b>U</b> |
| T2000        | 103        | U        |

| LSN | prevLSN | tID               | pID  | Log entry                 | Type   | undoNextLSN |
|-----|---------|-------------------|------|---------------------------|--------|-------------|
| 101 | -       | T1000             | P500 | Write A<br>"abc" -> "def" | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>"hij" -> "klm" | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs" | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy" | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                           | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                           | End    |             |
|     |         |                   |      |                           |        |             |

|                           |                            |
|---------------------------|----------------------------|
| <b>P500</b><br>PageLSN= - | <b>P600</b><br>PageLSN=102 |
| A = abc D = mnp           | B = klm                    |
| <b>P505</b><br>PageLSN= - | <b>P700</b><br>PageLSN= -  |
| C = tuv                   | E = pq                     |

**Disk**

# Analysis Phase

## Log

### Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P600   | 102    |
| P505   | 104    |
|        |        |

### Transaction table

| transID      | lastLSN    | status   |
|--------------|------------|----------|
| T1000        | 104        | U        |
| <b>T2000</b> | <b>105</b> | <b>C</b> |

| LSN | prevLSN | tID               | pID  | Log entry                 | Type   | undoNextLSN |
|-----|---------|-------------------|------|---------------------------|--------|-------------|
| 101 | -       | T1000             | P500 | Write A<br>"abc" -> "def" | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>"hij" -> "klm" | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs" | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy" | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                           | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                           | End    |             |
|     |         |                   |      |                           |        |             |

|                                              |                                       |
|----------------------------------------------|---------------------------------------|
| <b>P500</b><br>PageLSN= -<br>A = abc D = mnp | <b>P600</b><br>PageLSN=102<br>B = klm |
| <b>P505</b><br>PageLSN= -<br>C = tuv         | <b>P700</b><br>PageLSN= -<br>E = pq   |

**Disk**

Write A or Abort if you see an Abort log instead

# Analysis Phase

## Log

### Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P600   | 102    |
| P505   | 104    |
|        |        |

### Transaction table

| transID | lastLSN | status |
|---------|---------|--------|
| T1000   | 104     | U      |
| T2000   | 105     | C      |

| LSN | prevLSN | tID               | pID  | Log entry                 | Type   | undoNextLSN |
|-----|---------|-------------------|------|---------------------------|--------|-------------|
| 101 | -       | T1000             | P500 | Write A<br>"abc" -> "def" | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>"hij" -> "klm" | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs" | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy" | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                           | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                           | End    |             |
|     |         |                   |      |                           |        |             |

|                           |                            |
|---------------------------|----------------------------|
| <b>P500</b><br>PageLSN= - | <b>P600</b><br>PageLSN=102 |
| A = abc D = mnp           | B = klm                    |
| <b>P505</b><br>PageLSN= - | <b>P700</b><br>PageLSN= -  |
| C = tuv                   | E = pq                     |

**Disk**

Remove entry from Transaction Table if you see an End record (both for Aborted and Committed transactions)

# Analysis Phase

## Log

### Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P600   | 102    |
| P505   | 104    |
|        |        |

### Transaction table

| transID | lastLSN | status |
|---------|---------|--------|
| T1000   | 104     | U      |

| LSN | prevLSN | tID               | pID  | Log entry                 | Type   | undoNextLSN |
|-----|---------|-------------------|------|---------------------------|--------|-------------|
| 101 | -       | T1000             | P500 | Write A<br>"abc" -> "def" | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>"hij" -> "klm" | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs" | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy" | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                           | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                           | End    |             |
|     |         |                   |      |                           |        |             |

|                                              |                                       |
|----------------------------------------------|---------------------------------------|
| <b>P500</b><br>PageLSN= -<br>A = abc D = mnp | <b>P600</b><br>PageLSN=102<br>B = klm |
| <b>P505</b><br>PageLSN= -<br>C = tuv         | <b>P700</b><br>PageLSN= -<br>E = pq   |

**Disk**

Already written to disk,  
but reappears

# Compare with Dirty Table and Transaction Table right before Crash!!

## Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P505   | 104    |
| P700   | 107    |

## Transaction table

| transID           | lastLSN | status  |
|-------------------|---------|---------|
| T <sub>1000</sub> | 107     | Running |
|                   |         |         |

| LSN | prevLSN | tlD               | plD  | Log entry                 | Type   | undoNextLSN |
|-----|---------|-------------------|------|---------------------------|--------|-------------|
| 101 | -       | T <sub>1000</sub> | P500 | Write A<br>"abc" -> "def" | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>"hij" -> "klm" | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs" | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy" | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                           | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                           | End    |             |
| 107 | 104     | T <sub>1000</sub> | P700 | Write E<br>"pq" -> "rs"   | Update | -           |

## Buffer Pool

|                                                      |                                           |
|------------------------------------------------------|-------------------------------------------|
| <b>P500</b><br>PageLSN= 103<br><br>A = def   D = qrs |                                           |
| <b>P505</b><br>PageLSN= 104<br><br>C = tuv           | <b>P700</b><br>PageLSN= 107<br><br>E = rs |

## Disk

|                                                                                  |                                                   |
|----------------------------------------------------------------------------------|---------------------------------------------------|
| <b>P500</b><br><br>Lost update during crash,<br>but write ahead log, so<br>safe! | <b>P600</b><br><b>PageLSN= 102</b><br><br>B = klm |
| <b>P505</b><br>PageLSN= -<br><br>C = tuv                                         | <b>P700</b><br>PageLSN= -<br><br>E = pq           |



# REDO Phase

- “Repeating history” (including actions by transactions that would be aborted in the undo phase)
- Work with the Dirty Page Table
- Find the smallest recLSN in the dirty page table = FirstLSN
- Redo the “Update/CLR” action, unless (in this order)
  - Affected page is not in the dirty page table
  - Or, recLSN > LSN being checked
  - Or, pageLSN >= LSN being checked
- End/Commit/Abort LSNs are “skipped”
- **In HW, write “Redone” or “Skipped” for each LSN**

# Analysis Phase

## Log

### Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P600   | 102    |
| P505   | 104    |
|        |        |

### Transaction table

| transID | lastLSN | status |
|---------|---------|--------|
| T1000   | 104     | U      |

| LSN | prevLSN | tID               | pID  | Log entry                 | Type   | undoNextLSN |
|-----|---------|-------------------|------|---------------------------|--------|-------------|
| 101 | -       | T1000             | P500 | Write A<br>"abc" -> "def" | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>"hij" -> "klm" | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs" | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy" | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                           | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                           | End    |             |
|     |         |                   |      |                           |        |             |

|                           |                            |
|---------------------------|----------------------------|
| <b>P500</b><br>PageLSN= - | <b>P600</b><br>PageLSN=102 |
| A = abc D = mnp           | B = klm                    |
| <b>P505</b><br>PageLSN= - | <b>P700</b><br>PageLSN= -  |
| C = tuv                   | E = pq                     |

**Disk**

# REDO Phase: find firstLSN

Log

Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P600   | 102    |
| P505   | 104    |
|        |        |



LSN

| prevLSN | tID               | pID  | Log entry                 | Type   | undoNextLSN |
|---------|-------------------|------|---------------------------|--------|-------------|
| -       | T1000             | P500 | Write A<br>"abc" -> "def" | Update | -           |
| -       | T <sub>2000</sub> | P600 | Write B<br>"hij" -> "klm" | Update | -           |
| 102     | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs" | Update | -           |
| 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy" | Update | -           |
| 103     | T <sub>2000</sub> |      |                           | Commit |             |
| 105     | T <sub>2000</sub> |      |                           | End    |             |
|         |                   |      |                           |        |             |

Transaction table

| transID | lastLSN | status |
|---------|---------|--------|
| T1000   | 104     | U      |

|                                              |                                       |
|----------------------------------------------|---------------------------------------|
| <b>P500</b><br>PageLSN= -<br>A = abc D = mnp | <b>P600</b><br>PageLSN=102<br>B = klm |
| <b>P505</b><br>PageLSN= -<br>C = tuv         | <b>P700</b><br>PageLSN= -<br>E = pq   |

Disk

# REDO Phase

## Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P600   | 102    |
| P505   | 104    |
|        |        |

## Transaction table

| transID | lastLSN | status |
|---------|---------|--------|
| T1000   | 104     | U      |

## Buffer Pool

|                                    |  |
|------------------------------------|--|
| <b>P500</b><br>PageLSN= "-" to 101 |  |
| <b>A = def</b> D = mnp             |  |
|                                    |  |

## Log

| LSN | prevLSN | tID               | pID  | Log entry                 | Type   | undoNextLSN |
|-----|---------|-------------------|------|---------------------------|--------|-------------|
| 101 | -       | T1000             | P500 | Write A<br>"abc" -> "def" | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>"hij" -> "klm" | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs" | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy" | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                           | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                           | End    |             |
|     |         |                   |      |                           |        |             |

- Affected page is not in the dirty page table: **N**
- Else, recLSN > LSN being checked: **N**
- Else, pageLSN >= LSN being checked: **N**
- **REDO**

## Disk

|                           |                            |
|---------------------------|----------------------------|
| <b>P500</b><br>PageLSN= - | <b>P600</b><br>PageLSN=102 |
| A = abc D = mnp           | B = klm                    |
| <b>P505</b><br>PageLSN= - | <b>P700</b><br>PageLSN= -  |
| C = tuv                   | E = pq                     |

# REDO Phase

## Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P600   | 102    |
| P505   | 104    |
|        |        |

## Transaction table

| transID | lastLSN | status |
|---------|---------|--------|
| T1000   | 104     | U      |

## Buffer Pool

|                             |                            |
|-----------------------------|----------------------------|
| <b>P500</b><br>PageLSN= 101 | <b>P600</b><br>PageLSN=102 |
| A = def D = mnp             | <b>B = klm</b>             |
|                             |                            |

## Log

| LSN | prevLSN | tID               | pID  | Log entry                 | Type   | undoNextLSN |
|-----|---------|-------------------|------|---------------------------|--------|-------------|
| 101 | -       | T1000             | P500 | Write A<br>"abc" -> "def" | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>"hij" -> "klm" | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs" | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy" | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                           | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                           | End    |             |
|     |         |                   |      |                           |        |             |

- Affected page is not in the dirty page table: **N**
- Else, recLSN > LSN being checked: **N**
- Else, pageLSN >= LSN being checked: **Y**
- **NO REDO = SKIPPED**

## Disk

|                           |                            |
|---------------------------|----------------------------|
| <b>P500</b><br>PageLSN= - | <b>P600</b><br>PageLSN=102 |
| A = abc D = mnp           | B = klm                    |
| <b>P505</b><br>PageLSN= - | <b>P700</b><br>PageLSN= -  |
| C = tuv                   | E = pq                     |

# REDO Phase

## Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P600   | 102    |
| P505   | 104    |
|        |        |

## Transaction table

| transID | lastLSN | status |
|---------|---------|--------|
| T1000   | 104     | U      |

## Buffer Pool

|                                                                     |                                       |
|---------------------------------------------------------------------|---------------------------------------|
| <b>P500</b><br>PageLSN= 101 to <b>103</b><br>A = def <b>D = qrs</b> | <b>P600</b><br>PageLSN=102<br>B = klm |
|                                                                     |                                       |

## Log

| LSN | prevLSN | tID               | pID  | Log entry                 | Type   | undoNextLSN |
|-----|---------|-------------------|------|---------------------------|--------|-------------|
| 101 | -       | T1000             | P500 | Write A<br>"abc" -> "def" | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>"hij" -> "klm" | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs" | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy" | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                           | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                           | End    |             |
|     |         |                   |      |                           |        |             |

- Affected page is not in the dirty page table: **N**
- Else, recLSN > LSN being checked: **N**
- Else, pageLSN >= LSN being checked: **N**
- **REDO**

## Disk

|                                              |                                       |
|----------------------------------------------|---------------------------------------|
| <b>P500</b><br>PageLSN= -<br>A = abc D = mnp | <b>P600</b><br>PageLSN=102<br>B = klm |
| <b>P505</b><br>PageLSN= -<br>C = tuv         | <b>P700</b><br>PageLSN= -<br>E = pq   |

# REDO Phase

## Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P600   | 102    |
| P505   | 104    |
|        |        |

## Transaction table

| transID | lastLSN | status |
|---------|---------|--------|
| T1000   | 104     | U      |

## Buffer Pool

|                                                     |                                       |
|-----------------------------------------------------|---------------------------------------|
| <b>P500</b><br>PageLSN= 103<br>A = def D = qrs      | <b>P600</b><br>PageLSN=102<br>B = klm |
| <b>P505</b><br>PageLSN="-" to <b>104</b><br>C = wxy |                                       |

|                                              |                                       |
|----------------------------------------------|---------------------------------------|
| <b>P500</b><br>PageLSN= -<br>A = abc D = mnp | <b>P600</b><br>PageLSN=102<br>B = klm |
| <b>P505</b><br>PageLSN= -<br>C = tuv         | <b>P700</b><br>PageLSN= -<br>E = pq   |

## Log

| LSN | prevLSN | tID               | pID  | Log entry                 | Type   | undoNextLSN |
|-----|---------|-------------------|------|---------------------------|--------|-------------|
| 101 | -       | T1000             | P500 | Write A<br>"abc" -> "def" | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>"hij" -> "klm" | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs" | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy" | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                           | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                           | End    |             |
|     |         |                   |      |                           |        |             |

- Affected page is not in the dirty page table: **N**
- Else, recLSN > LSN being checked: **N**
- Else, pageLSN >= LSN being checked: **N**
- **REDO**

Disk

# UNDO Phase

- Work with the Transaction table in the analysis phase
  - “Loser transactions” must be undone
  - Changes during undo phases are written (CLR) so that it is not repeated at the time of repeated restarts
- Scan backward
- Maintain a set ToUndo
  - Initialize to lastLSNs of all “U” transactions at Transaction Table
  - undo the “largest LSN” in ToUndo at each step (the latest one in bottom-up order)



## Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P600   | 102    |
| P505   | 104    |
|        |        |

## Transaction table

| transID | lastLSN | status |
|---------|---------|--------|
| T1000   | 104     | U      |

## Buffer Pool

|                             |                            |
|-----------------------------|----------------------------|
| <b>P500</b><br>PageLSN= 103 | <b>P600</b><br>PageLSN=102 |
| A = def D = qrs             | B = klm                    |
| <b>P505</b><br>PageLSN= 104 |                            |
| C = wxy                     |                            |

|                           |                            |
|---------------------------|----------------------------|
| <b>P500</b><br>PageLSN= - | <b>P600</b><br>PageLSN=102 |
| A = abc D = mnp           | B = klm                    |
| <b>P505</b><br>PageLSN= - | <b>P700</b><br>PageLSN= -  |
| C = tuv                   | E = pq                     |

# UNDO Phase

## Log

| LSN | prevLSN | tID               | pID  | Log entry                 | Type   | undoNextLSN |
|-----|---------|-------------------|------|---------------------------|--------|-------------|
| 101 | -       | T1000             | P500 | Write A<br>"abc" -> "def" | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>"hij" -> "klm" | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs" | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy" | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                           | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                           | End    |             |
|     |         |                   |      |                           |        |             |

ToUNDO = {104}

Disk

## Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P600   | 102    |
| P505   | 104    |
|        |        |

## Transaction table

| transID | lastLSN | status |
|---------|---------|--------|
| T1000   | 104     | U      |

## Buffer Pool

|                                    |                            |
|------------------------------------|----------------------------|
| <b>P500</b><br>PageLSN= 103        | <b>P600</b><br>PageLSN=102 |
| A = def D = qrs                    | B = klm                    |
| <b>P505</b><br><b>PageLSN= 107</b> |                            |
| C =tuv                             |                            |

|                           |                            |
|---------------------------|----------------------------|
| <b>P500</b><br>PageLSN= - | <b>P600</b><br>PageLSN=102 |
| A = abc D = mnp           | B = klm                    |
| <b>P505</b><br>PageLSN= - | <b>P700</b><br>PageLSN= -  |
| C = tuv                   | E = pq                     |

# UNDO Phase

## Log

| LSN | prevLSN | tID               | pID  | Log entry                    | Type   | undoNextLSN |
|-----|---------|-------------------|------|------------------------------|--------|-------------|
| 101 | -       | T1000             | P500 | Write A<br>"abc" -> "def"    | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>"hij" -> "klm"    | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs"    | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy"    | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                              | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                              | End    |             |
| 107 |         | T <sub>1000</sub> |      | UndoT <sub>1000</sub> LSN104 | CLR    | 101         |

- A CLR is written
- PageLSN = LSN (CLR)
- Value of C is undone

ToUNDO = {101}

Disk

# CLR

- CLR is added so that no “Undo” action is undone
- If a CLR is encountered during UNDO phase, goes to the LSN in UndoNextLSN
- At most one CLR per update, limits log size

## Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P600   | 102    |
| P505   | 104    |
|        |        |

## Transaction table

| transID | lastLSN | status |
|---------|---------|--------|
| T1000   | 104     | U      |

## Buffer Pool

|                             |                            |
|-----------------------------|----------------------------|
| <b>P500</b><br>PageLSN= 103 | <b>P600</b><br>PageLSN=102 |
| A = def D = qrs             | B = klm                    |
| <b>P505</b><br>PageLSN= 107 |                            |
| C = tuv                     |                            |

|                           |                            |
|---------------------------|----------------------------|
| <b>P500</b><br>PageLSN= - | <b>P600</b><br>PageLSN=102 |
| A = abc D = mnp           | B = klm                    |
| <b>P505</b><br>PageLSN= - | <b>P700</b><br>PageLSN= -  |
| C = tuv                   | E = pq                     |

# UNDO Phase

## Log

| LSN | prevLSN | tID               | pID  | Log entry                    | Type   | undoNextLSN |
|-----|---------|-------------------|------|------------------------------|--------|-------------|
| 101 | -       | T1000             | P500 | Write A<br>"abc" -> "def"    | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>"hij" -> "klm"    | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs"    | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy"    | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                              | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                              | End    |             |
| 107 |         | T <sub>1000</sub> |      | UndoT <sub>1000</sub> LSN104 | CLR    | 101         |

Disk

ToUNDO = {101}

## Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P600   | 102    |
| P505   | 104    |
|        |        |

## Transaction table

| transID | lastLSN | status |
|---------|---------|--------|
| T1000   | 104     | U      |

## Buffer Pool

|                             |                            |
|-----------------------------|----------------------------|
| <b>P500</b><br>PageLSN= 108 | <b>P600</b><br>PageLSN=102 |
| A = abc D = qrs             | B = klm                    |
| <b>P505</b><br>PageLSN= 107 |                            |
| C = tuv                     |                            |

|                           |                            |
|---------------------------|----------------------------|
| <b>P500</b><br>PageLSN= - | <b>P600</b><br>PageLSN=102 |
| A = abc D = mnp           | B = klm                    |
| <b>P505</b><br>PageLSN= - | <b>P700</b><br>PageLSN= -  |
| C = tuv                   | E = pq                     |

# UNDO Phase

## Log

| LSN | prevLSN | tID               | pID  | Log entry                    | Type   | undoNextLSN |
|-----|---------|-------------------|------|------------------------------|--------|-------------|
| 101 | -       | T1000             | P500 | Write A<br>"abc" -> "def"    | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>"hij" -> "klm"    | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs"    | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy"    | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                              | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                              | End    |             |
| 107 |         | T <sub>1000</sub> |      | UndoT <sub>1000</sub> LSN104 | CLR    | 101         |
| 108 |         | T <sub>1000</sub> |      | UndoT <sub>1000</sub> LSN101 | CLR    | -           |

Disk

ToUNDO = {}

## Dirty page table

| pageID | recLSN |
|--------|--------|
| P500   | 101    |
| P600   | 102    |
| P505   | 104    |
|        |        |

## Transaction table

| transID | lastLSN | status |
|---------|---------|--------|
| T1000   | 104     | U      |

## Buffer Pool

|                                                |                                       |
|------------------------------------------------|---------------------------------------|
| <b>P500</b><br>PageLSN= 108<br>A = abc D = qrs | <b>P600</b><br>PageLSN=102<br>B = klm |
| <b>P505</b><br>PageLSN= 107<br>C = tuv         |                                       |

|                                              |                                       |
|----------------------------------------------|---------------------------------------|
| <b>P500</b><br>PageLSN= -<br>A = abc D = mnp | <b>P600</b><br>PageLSN=102<br>B = klm |
| <b>P505</b><br>PageLSN= -<br>C = tuv         | <b>P700</b><br>PageLSN= -<br>E = pq   |

# UNDO Phase

## Log

| LSN | prevLSN | tID               | pID  | Log entry                    | Type   | undoNextLSN |
|-----|---------|-------------------|------|------------------------------|--------|-------------|
| 101 | -       | T1000             | P500 | Write A<br>"abc" -> "def"    | Update | -           |
| 102 | -       | T <sub>2000</sub> | P600 | Write B<br>"hij" -> "klm"    | Update | -           |
| 103 | 102     | T <sub>2000</sub> | P500 | Write D<br>"mnp" -> "qrs"    | Update | -           |
| 104 | 101     | T <sub>1000</sub> | P505 | Write C<br>"tuv" -> "wxy"    | Update | -           |
| 105 | 103     | T <sub>2000</sub> |      |                              | Commit |             |
| 106 | 105     | T <sub>2000</sub> |      |                              | End    |             |
| 107 |         | T <sub>1000</sub> |      | UndoT <sub>1000</sub> LSN104 | CLR    | 101         |
| 108 |         | T <sub>1000</sub> |      | UndoT <sub>1000</sub> LSN101 | CLR    | -           |
| 109 |         | T <sub>1000</sub> |      |                              | End    | -           |

Disk

Write an END record  
(explicitly mentioned in hw 5 for  
the aborted transaction)

# What happens if T aborts?

1. Write an “abort” log record for T
  - Like “commit”
  - Also the status in Transaction table from “Running” to “Aborted”
- 2 Follow “prevLSN” to undo all updates by T
  - Like the “UNDO” phase
  - Undo page content in buffer pool, pageLSN changed to that LSN(CLR)
  - Write the CLR records
    - Log entry like “Undo T3 LSN5”
    - No prevLSN, but record undoNextLSN
    - Until undoNextLSN is null
3. Write an “End” log record for T

# Handling Crashes during Undo

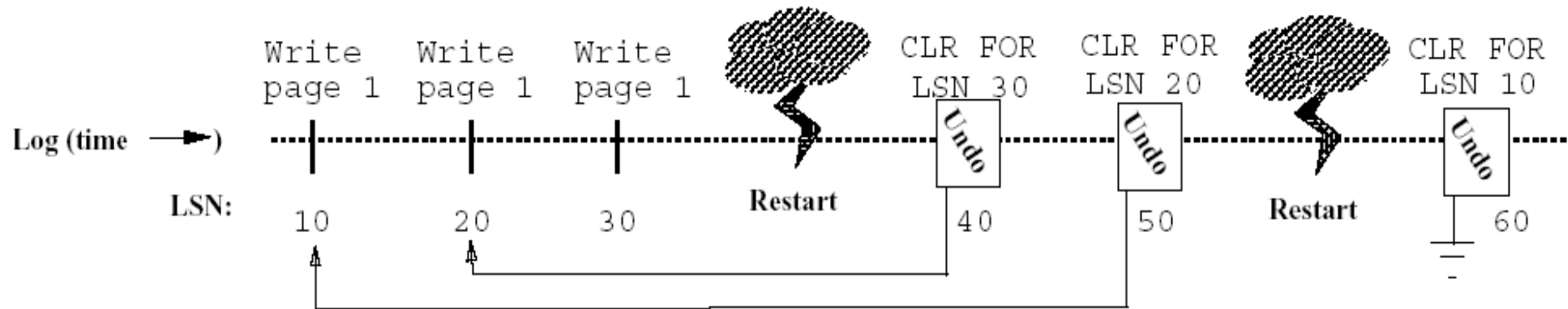


Figure 4: The Use of CLR's for UNDO

[Figure 4 from Franklin97]

In general, for every single crash (even for crash during Analysis/Redo/Undo phases), start again with Analysis

If some CLR records are written to disk during an UNDO phase, then a crash happens (e.g. here LSN 40, 50 are written to disk before the second crash), then the next UNDO phase will skip undoing those CLR's.