3 Phases of ARIES

1. Analysis
   - First log record
   - Earliest change applied by a loser transaction
   - Earliest possibly dirty page

2. Redo
   - Log file

3. Undo
   - Last log record

Notice: for the moment no checkpointing!
### Transaction Table TT

<table>
<thead>
<tr>
<th>TaID</th>
<th>lastLSN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>567</td>
</tr>
<tr>
<td>2</td>
<td>42</td>
</tr>
<tr>
<td>7</td>
<td>67</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
</tr>
</tbody>
</table>

**TaID:** key of the transaction

**lastLSN:** LSN of the most recent log record seen for this transaction, i.e. the latest change done by this transaction

---

### Dirty Page Table DPT

<table>
<thead>
<tr>
<th>pageID</th>
<th>recoveryLSN</th>
</tr>
</thead>
<tbody>
<tr>
<td>42</td>
<td>567</td>
</tr>
<tr>
<td>46</td>
<td>568</td>
</tr>
<tr>
<td>77</td>
<td>34</td>
</tr>
<tr>
<td>3</td>
<td>42</td>
</tr>
</tbody>
</table>
### Transaction Table TT

<table>
<thead>
<tr>
<th>TaID</th>
<th>lastLSN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>567</td>
</tr>
<tr>
<td>2</td>
<td>42</td>
</tr>
<tr>
<td>7</td>
<td>67</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
</tr>
</tbody>
</table>

**TaID**: key of the transaction  

**lastLSN**: LSN of the most recent log record seen for this transaction, i.e. the **latest** change done by this transaction

### Dirty Page Table DPT

<table>
<thead>
<tr>
<th>pageID</th>
<th>recoveryLSN</th>
</tr>
</thead>
<tbody>
<tr>
<td>42</td>
<td>567</td>
</tr>
<tr>
<td>46</td>
<td>568</td>
</tr>
<tr>
<td>77</td>
<td>34</td>
</tr>
<tr>
<td>3</td>
<td>42</td>
</tr>
</tbody>
</table>

**pageID**: key of a page  

**recoveryLSN**: LSN of **first** log record that made this page dirty, i.e. the **earliest** change done to this page

**assumptions**: page may contain changes from losers and winners, i.e. two concurrent transactions may operate **on the same page**

---

**Diagram Notes**:
- p42 → commits
- s6 ← 26
- tuple-wise lock
- NJ - buffer
- STEAL
- NO-FORCE
The Log Records in Detail

All Log Records:

\[ \text{prevLSN, TaID, type} \]

\[
\begin{array}{c|c|c}
\text{prevLSN} & \text{TaID} & \text{type} \\
420 & 42 & \\
\end{array}
\]

prevLSN: (Previous Log Sequence Number): link to the previous LSN for this TA (if not first entry)

TaID: foreign key to unique transaction ID
The Log Records in Detail

All Log Records:

\[\text{prevLSN, TaID, type}\]

prevLSN: (Previous Log Sequence Number): link to the previous LSN for this TA (if not first entry)

TaID: foreign key to unique transaction ID

type: kind of log record; either: update, commit or compensation (for undoing updates)
The Log Records in Detail: Update Log Records

All Log Records:

[prevLSN, TaID, type]

Update Log Record:

[prevLSN, TaID, "update", pageID, redo info, undo info]
The Log Records in Detail: Update Log Records

All Log Records:

[prevLSN, TaID, type]

Update Log Record:

[prevLSN, TaID, "update", pageID, redo info, undo info]

pageID: foreign key to pageID
The Log Records in Detail: Update Log Records

All Log Records:

[prevLSN, TID, type]

Update Log Record:

[prevLSN, TID, "update", pageID, redo info, undo info]

*pageID*: foreign key to pageID

*redo info*: information how to redo the change reflected by this log record (physiological)
The Log Records in Detail: Update Log Records

All Log Records:

[prevLSN, TaID, type]

Update Log Record:

[prevLSN, TaID, "update", pageID, redo info, undo info]

For the moment

ARIES: redo: physiological
    undo: (logical) physiological

pageID: foreign key to pageID

redo info: information how to redo the change reflected by this log record (physiological)

undo info: information how to undo the change reflected by this log record (logical)
Example Log File

LSN: [prevLSN, TaID, type]

LSN: [prevLSN, TaID, “update“, pageID, redo info, undo info]
Example Log File

LSN: [prevLSN, TaID, type]
LSN: [prevLSN, TaID, "update", pageID, redo info, undo info]

```
log

1: [-, 1, "update", 42, a+=1, a-=1]
```

<table>
<thead>
<tr>
<th>TT</th>
<th>TaID</th>
<th>lastLSN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DPT</th>
<th>pageID</th>
<th>recoveryLSN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>42</td>
<td>1</td>
</tr>
</tbody>
</table>
Example Log File

LSN: [prevLSN, TaID, type]
LSN: [prevLSN, TaID, “update“, pageID, redo info, undo info]

log
1: [-, 1, “update“, 42, a+=1, a-=1]
2: [-, 2, “update“, 42, b+=3, b-=3]

<table>
<thead>
<tr>
<th>TT</th>
<th>TaID</th>
<th>lastLSN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DPT</th>
<th>pageID</th>
<th>recoveryLSN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>42</td>
<td>1</td>
</tr>
</tbody>
</table>
Example Log File

LSN: [prevLSN, TaID, type]
LSN: [prevLSN, TaID, “update“, pageID, redo info, undo info]

all
“update“

<table>
<thead>
<tr>
<th>DB buffer</th>
<th>log</th>
<th>TT</th>
<th>DPT</th>
</tr>
</thead>
<tbody>
<tr>
<td>page 42</td>
<td></td>
<td>TaID</td>
<td>lastLSN</td>
</tr>
<tr>
<td></td>
<td>1:</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2:</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>page 46</td>
<td>3:</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3:</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

LSN=2
a=78
b=58

LSN=3
c=24
Example Log File

LSN: [prevLSN, TaID, type]
LSN: [prevLSN, TaID, "update", pageID, redo info, undo info]

DB buffer
page 42
LSN=4
a=78
b=59

log

1: [-, 1, "update", 42, a+=1, a-=1]
2: [-, 2, "update", 42, b+=3, b-=3]
3: [2, 2, "update", 46, c+=2, c-=2]
4: [1, 1, "update", 42, b+=1, b-=1]

TT
<table>
<thead>
<tr>
<th>TaID</th>
<th>lastLSN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

DPT
<table>
<thead>
<tr>
<th>pageID</th>
<th>recoveryLSN</th>
</tr>
</thead>
<tbody>
<tr>
<td>42</td>
<td>1</td>
</tr>
<tr>
<td>46</td>
<td>3</td>
</tr>
</tbody>
</table>
Example Log File

LSN: [prevLSN, TaID, type]

LSN: [prevLSN, TaID, "update", pageID, redo info, undo info]

all

"update"

<table>
<thead>
<tr>
<th>TT</th>
</tr>
</thead>
<tbody>
<tr>
<td>TaID</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DPT</th>
</tr>
</thead>
<tbody>
<tr>
<td>pageID</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>42</td>
</tr>
<tr>
<td>46</td>
</tr>
</tbody>
</table>
The Log Records in Detail: Compensation Log Records

All Log Records:

[prevLSN, TalD, type]

Update Log Record:

[prevLSN, TalD, "update", pageID, redo info, undo info]

Compensation Log Record (CLR, for undoing updates):

[prevLSN, TalD, "compensation", redoTheUndo info, undoNextLSN]

redoTheUndo info: information how to redo the undo (reflected by the log record creating this CLR)
The Log Records in Detail: Compensation Log Records

All Log Records:

[prevLSN, TaID, type]

Update Log Record:

[prevLSN, TaID, “update“, pageID, redo info, undo info]

Compensation Log Record (CLR, for undoing updates):

[prevLSN, TaID, “compensation“, pageID, redoTheUndo info, undoNextLSN]

redoTheUndo info: information how to redo the undo (reflected by the log record creating this CLR)

undoNextLSN: link to the next log record to be undone for this transaction

i.e., if a log record LR wrote this CLR, we set CLR.undoNextLSN := LR.prevLSN
Example Log File

LSN: [prevLSN, TaID, type]
LSN: [prevLSN, TaID, "update", pageID, redo info, undo info]
LSN: [prevLSN, TaID, "compensation", redoTheUndo info, undoNextLSN]

DB buffer
page 42
LSN=4
a=78
b=59

log

1: [-, 1, "update", 42, a+=1, a-=1]
2: [-, 2, "update", 42, b+=3, b-=3]
3: [2, 2, "update", 46, c+=2, c-=2]
4: [1, 1, "update", 42, b+=1, b-=1]
5: [3, 2, "commit"]

TT
<table>
<thead>
<tr>
<th>TaID</th>
<th>lastLSN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

DPT
<table>
<thead>
<tr>
<th>pageID</th>
<th>recoveryLSN</th>
</tr>
</thead>
<tbody>
<tr>
<td>42</td>
<td>1</td>
</tr>
<tr>
<td>46</td>
<td>3</td>
</tr>
</tbody>
</table>
Example Log File

LSN: [prevLSN, TaID, type]  
all
LSN: [prevLSN, TaID, “update“, pageID, redo info, undo info]  
“update“  
LSN: [prevLSN, TaID, “compensation“, redoTheUndo info, undoNextLSN]  
“compensation“

DB buffer  
log

page 42  
LSN=6  
a=78  
b=58

page 46  
LSN=3  
c=24

1: [-, 1, “update“, 42, a+=1, a-=1]
2: [-, 2, “update“, 42, b+=3, b-=3]
3: [2, 2, “update“, 46, c+=2, c-=2]
4: [1, 1, “update“, 42, b+=1, b-=1]
5: [3, 2, “commit“]
6: [4, 1, “compensation“, 42, b-=1, 1]

TT
<table>
<thead>
<tr>
<th>TaID</th>
<th>lastLSN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
</tr>
</tbody>
</table>

DPT
<table>
<thead>
<tr>
<th>pageID</th>
<th>recoveryLSN</th>
</tr>
</thead>
<tbody>
<tr>
<td>42</td>
<td>1</td>
</tr>
<tr>
<td>46</td>
<td>3</td>
</tr>
</tbody>
</table>
Example Log File

LSN: [prevLSN, TaID, type]
LSN: [prevLSN, TaID, "update", pageID, redo info, undo info]
LSN: [prevLSN, TaID, "compensation", redoTheUndo info, undoNextLSN]

DB buffer page 42

| LSN=7 | a=77 | b=58 |

log

1: [-, 1, "update", 42, a+=1, a-=1]
2: [-, 2, "update", 42, b+=3, b-=3]
3: [2, 2, "update", 46, c+=2, c-=2]
4: [1, 1, "update", 42, b+=1, b-=1]
5: [3, 2, "commit"]
6: [4, 1, "compensation", 42, b-=1, 1]
7: [6, 1, "compensation", 42, a-=1, -]

TT

<table>
<thead>
<tr>
<th>TaID</th>
<th>lastLSN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
</tr>
</tbody>
</table>

DPT

<table>
<thead>
<tr>
<th>pageID</th>
<th>recoveryLSN</th>
</tr>
</thead>
<tbody>
<tr>
<td>42</td>
<td>1</td>
</tr>
<tr>
<td>46</td>
<td>3</td>
</tr>
</tbody>
</table>
Example Log File

LSN: [prevLSN, TaID, type]
LSN: [prevLSN, TaID, “update“, pageID, redo info, undo info]
LSN: [prevLSN, TaID, “compensation“, redoTheUndo info, undoNextLSN]

DB buffer
page 42
LSN=7
a=77
b=58

page 46
LSN=3
c=24

log

1: [-, 1, “update“, 42, a+=1, a-=1]
2: [-, 2, “update“, 42, b+=3, b-=3]
3: [2, 2, “update“, 46, c+=2, c-=2]
4: [1, 1, “update“, 42, b+=1, b-=1]
5: [3, 2, “commit“]
6: [4, 1, “compensation“, 42, b-=1, 1]
7: [6, 1, “compensation“, 42, a-=1, -]

TT
TaID  lastLSN  full

DPT
pageID  recoveryLSN
42       1
46       3
Starting Points

\[ \text{minDirtyPageLSN} := \text{SELECT min(recoveryLSN) FROM DPT}; \]

- firstLSN: earliest change possibly applied by a loser transaction
- minDirtyPageLSN: earliest possibly dirty page

---

**log file**

**Analysis**

---

**Redo**

---

**Undo**

---

notice: for the moment no checkpointing!

note: relative starting points of the phases may differ
Fuzzy Checkpoint

\[ \text{minDirtyPageLSN} := \text{SELECT min}(\text{recoveryLSN}) \text{ FROM DPT}; \]

- **firstLSN**: earliest change possibly applied by a loser transaction
- **minDirtyPageLSN**: earliest possibly dirty page = DPT plus TT last checkpoint
- **last log record**: last log record

---

**log file**

- **Undo**
- **Redo**
- **Analysis**

*note: relative starting points of the phases may differ*
Background Write Thread

\[ \text{minDirtyPageLSN} := \text{SELECT} \ \text{min} \text{(recoveryLSN)} \ \text{FROM} \ \text{DPT}; \]

- **firstLSN**: earliest change possibly applied by a loser transaction
- **minDirtyPageLSN**: earliest possibly dirty page
- **= DPT plus TT last checkpoint**: last log record

---

**log file**

- **Undo**
- **Redo**
- **Analysis**

Note: relative starting points of the phases may differ
Short Running Transactions

\[ \text{minDirtyPageLSN} := \text{SELECT min(recoveryLSN) FROM DPT}; \]

- \( \text{minDirtyPageLSN} \)
- \( \text{firstLSN} \)
- \( \text{earliest possibly dirty page} \)
- \( \text{earliest change possibly applied by a loser transaction} \)
- \( \text{DPT plus TT last checkpoint} \)

- **Analysis**
- **Redo**
- **Undo**

note: relative starting points of the phases may differ
Fuzzy Checkpointing Details

$LSN = 423$

$\frac{\sqrt{423}}{TT + DPT}$

= DPT plus TT as of begin_checkpoint

$log file$

(1.) begin_checkpoint

(2.) end_checkpoint

(3.) flush position of begin_checkpoint to extra place on stable storage

$log file$
Tasks of the 3 Phases in ARIES

Analysis Phase:
construct DPT and TT to reflect state at the time of the crash
compute minDirtyPageLSN (where to start Redo phase)

Redo Phase: repeat history from a different view
repeat all actions: including those from loser transactions, including CLRs
Tasks of the 3 Phases in ARIES

Analysis Phase:

construct DPT and TT to reflect state at the time of the crash
compute minDirtyPageLSN (where to start Redo phase)

Redo Phase:

repeat all actions: including those from loser transactions, including CLRls
restore database to where it was at the time of the crash

Undo Phase:

undo actions of all loser transactions
log undos to CLRls