

Chapter 9. Log manager

PostgreSQL WAL

Jongbaeg Lee

hundredbag@gmail.com

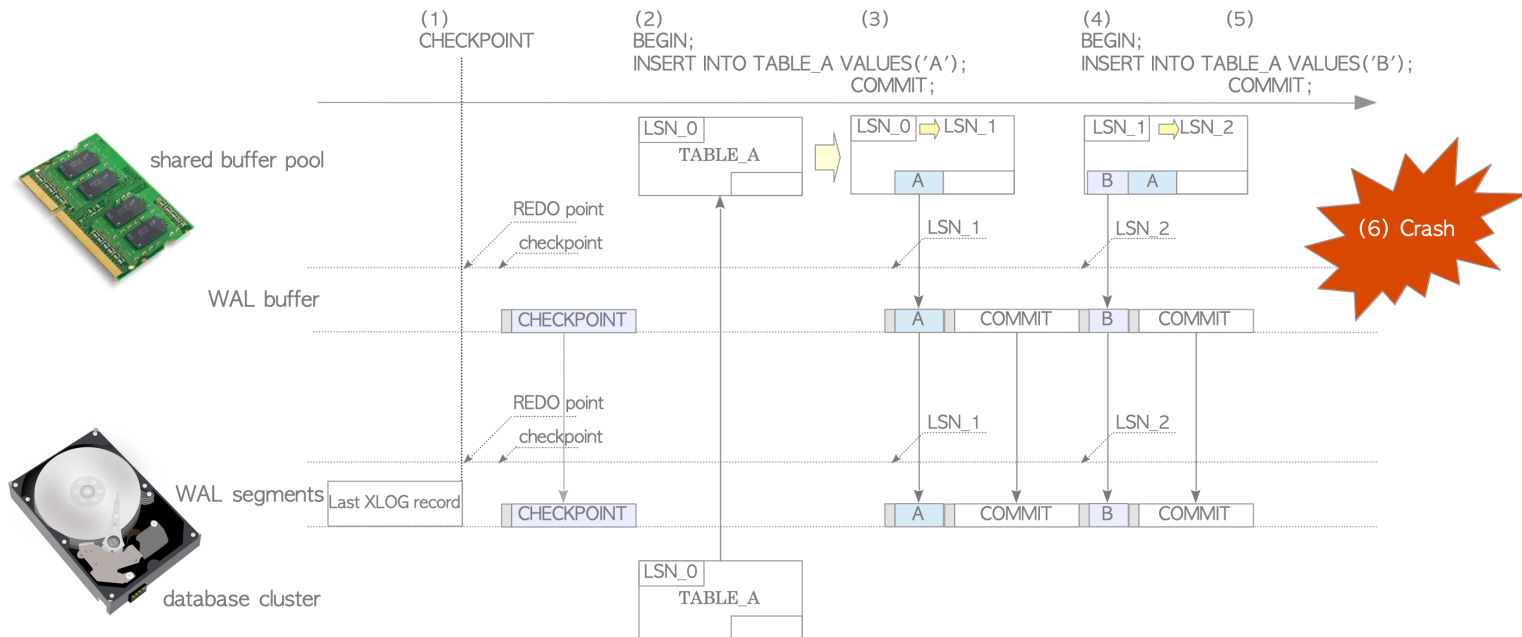


VLDB
Lab.

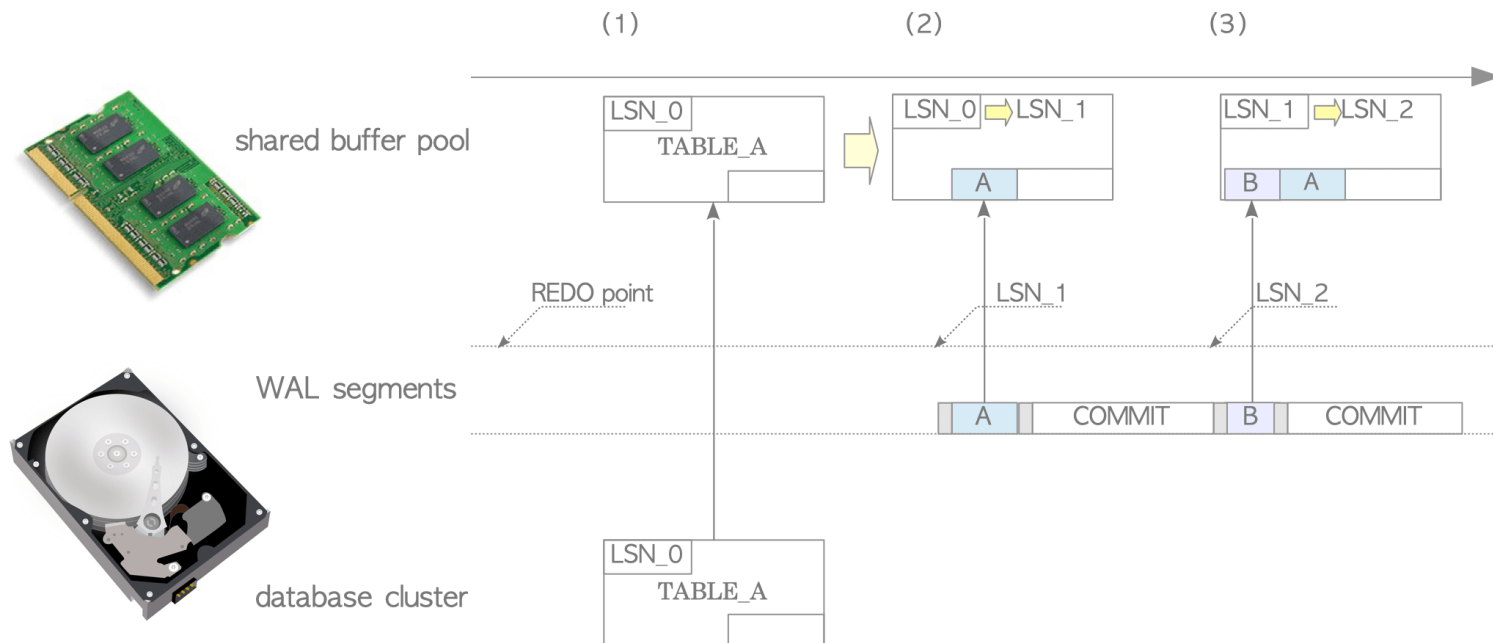
Write Ahead Logging

- PostgreSQL writes all modifications as history data into a storage
 - They are called as **XLOG record** or **WAL data**
- XLOG record (WAL data)
 - XLOG records are written into the in-memory **WAL buffer**
 - They are immediately written into a **WAL segment file** on the storage when a transaction commits/aborts
 - **LSN (Log Sequence Number)** of XLOG record represents the location where its record is written on the transaction log
 - **Redo point** : the location to write the XLOG record at the moment when the latest **checkpoint** is started

Insertion Operations with WAL

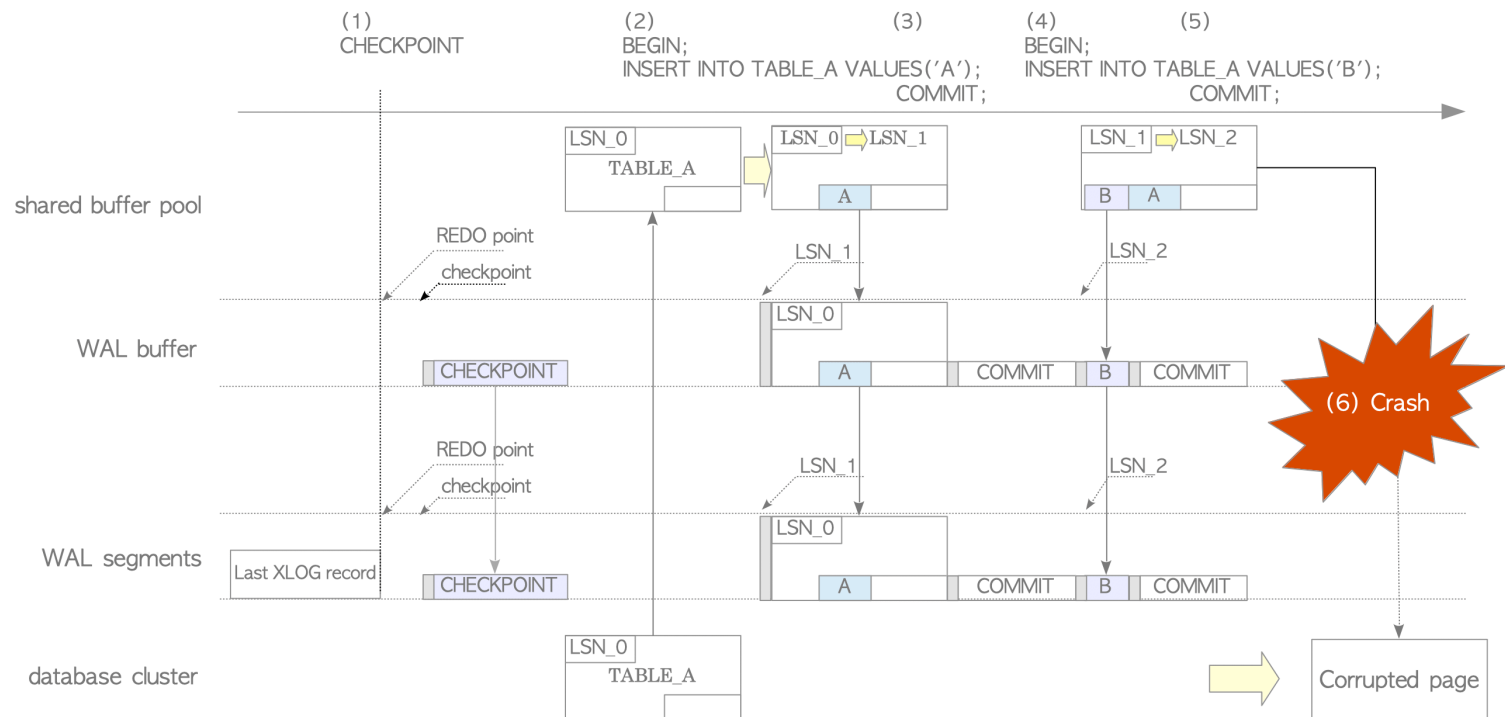


Database Recovery using WAL

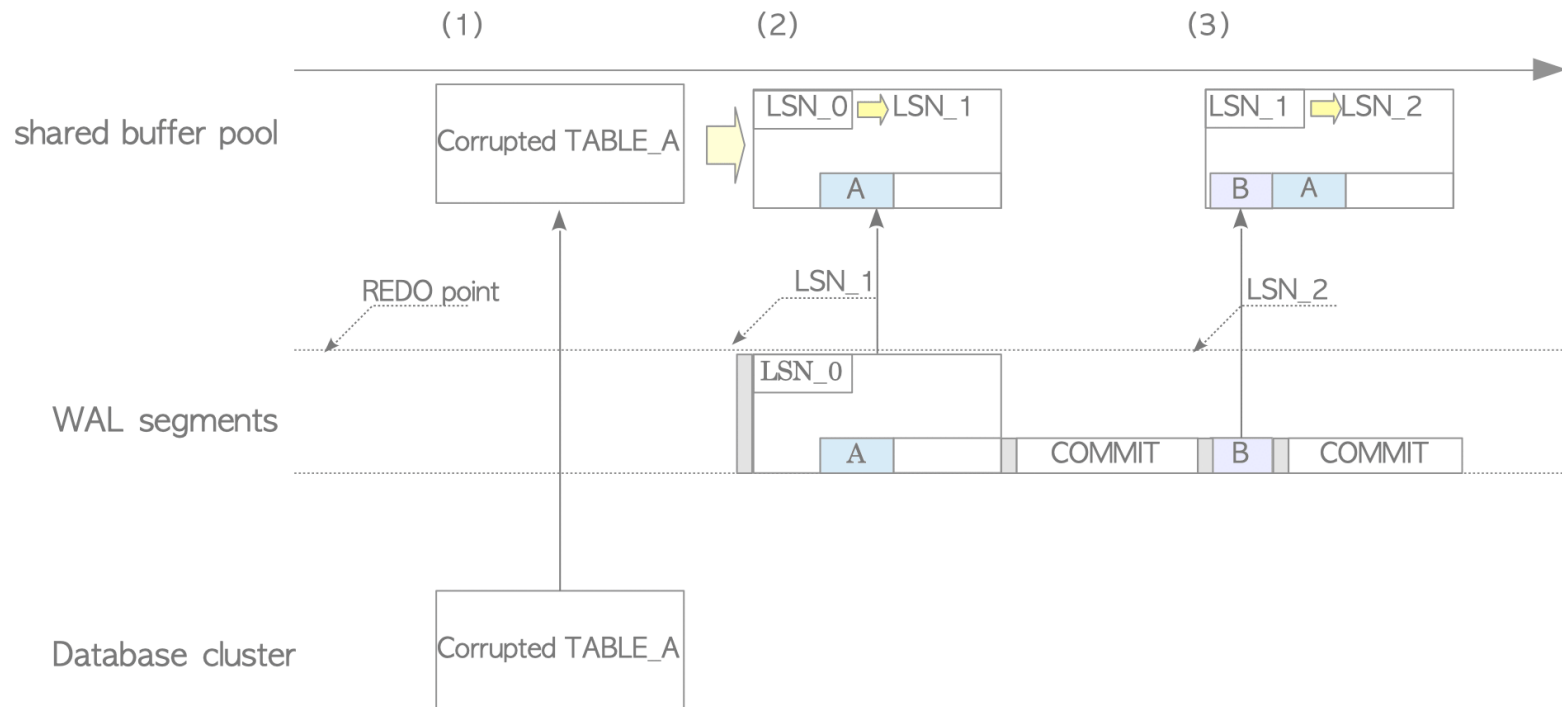


Full-Page Writes

- PostgreSQL writes a **pair of the header-data and the *entire page* as XLOG** record during the first change of each page after every checkpoint
- XLOG record containing the entire page is referred to as **backup block (or full-page image)**

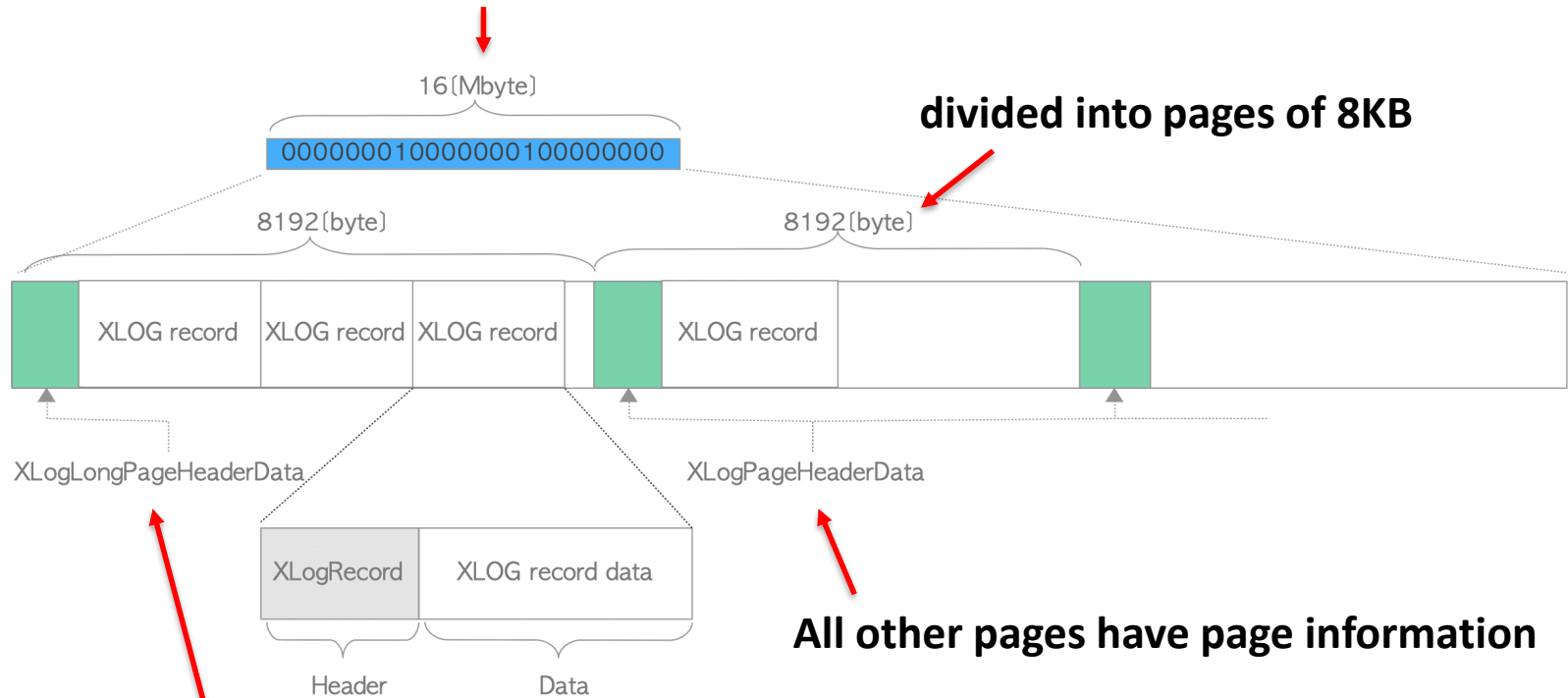


Database Recovery with Backup Block



Internal Layout of WAL Segment

a WAL segment (16MB) file



First page has a header-data

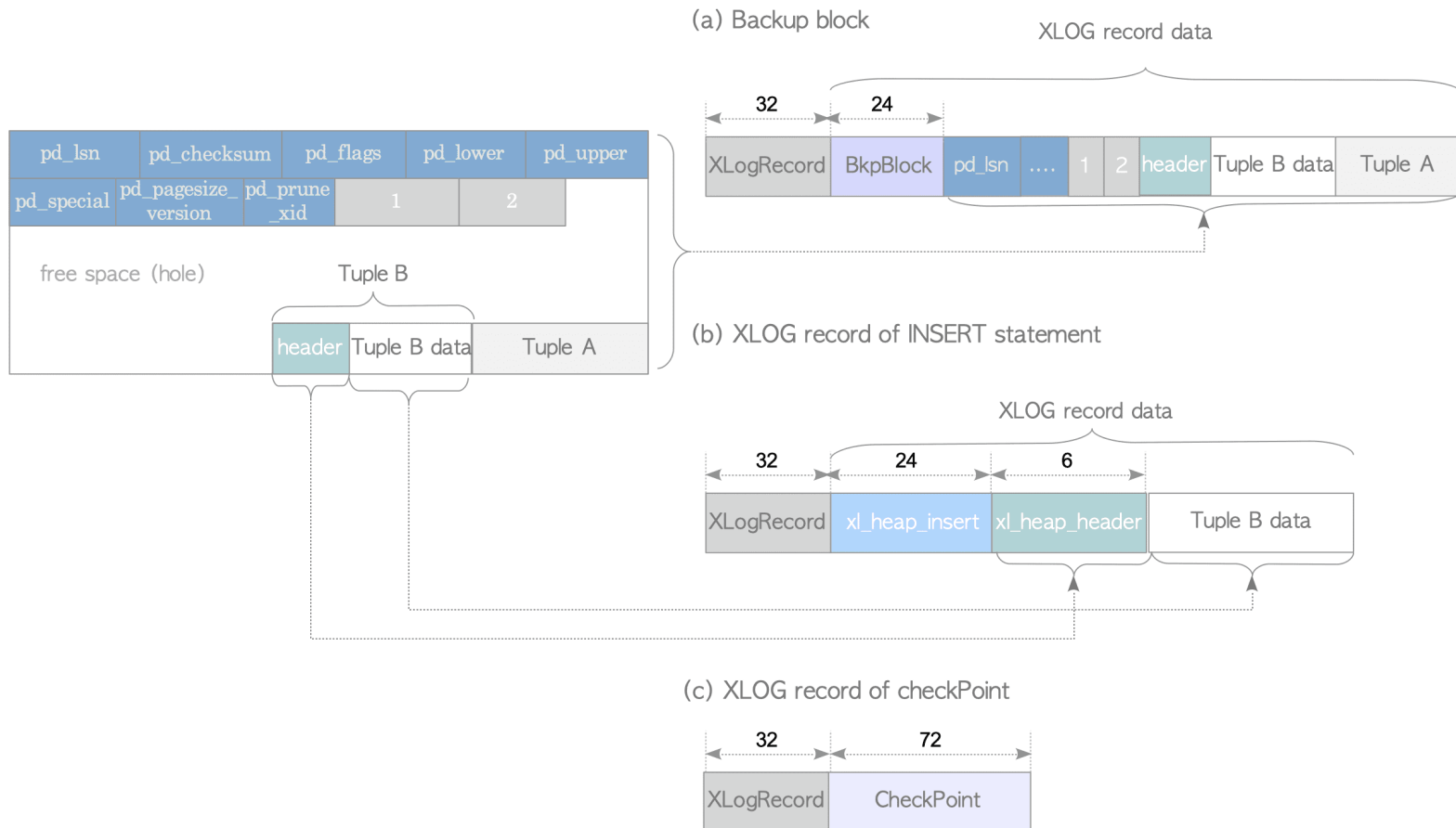
XLOG Record Layout

- XLOG Record Header

```
typedef struct XLogRecord
{
    uint32          xl_tot_len;    /* total len of entire record */
    TransactionId   xl_xid;        /* xact id */
    uint32          xl_len;        /* total len of rmgr data */
    uint8           xl_info;        /* flag bits, see below */
    RmgrId          xl_rmid;        /* resource manager for this record */
    /* 2 bytes of padding here, initialize to zero */
    XLogRecPtr      xl_prev;        /* ptr to previous record in log */
    pg_crc32        xl_crc;        /* CRC for this record */
} XLogRecord;
```


XLOG Record Layout

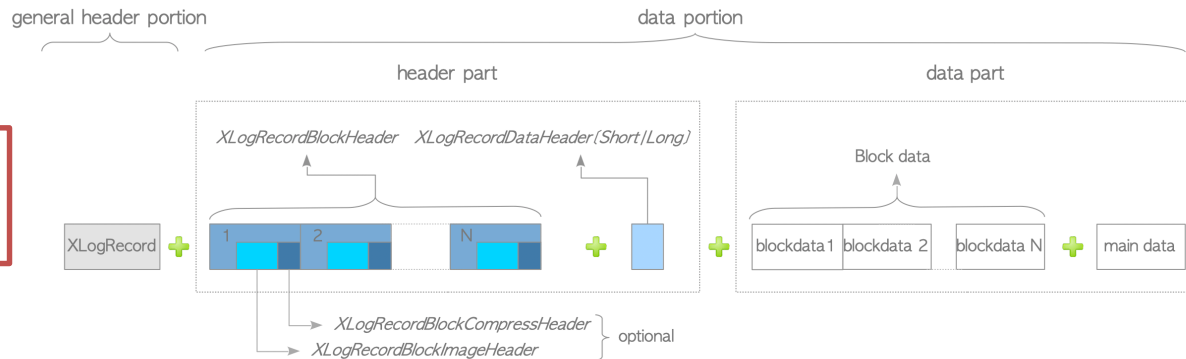
- XLOG Record Data Portion (version 9.4 or earlier)



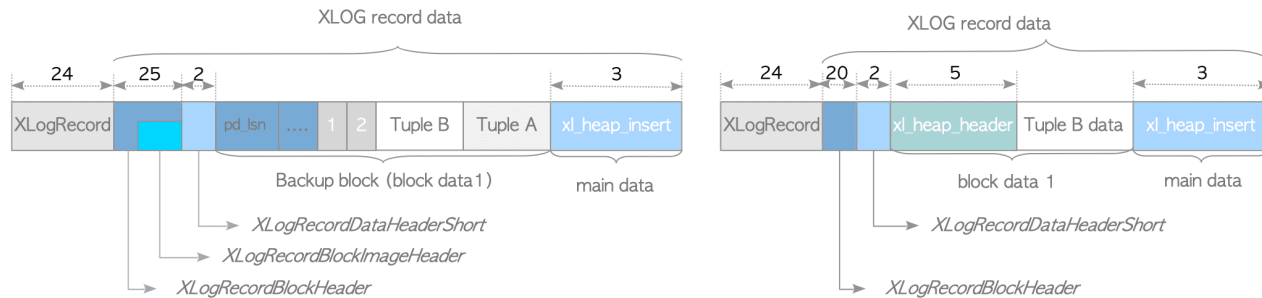
XLOG Record Layout

- XLOG Record Data Portion (version 9.5 or later)

Common XLOG record format



Examples of XLOG records



(c) XLOG record of CHECKPOINT action

