

Prednáška II

Logovanie a algoritmy pre obnovu

Obsah prednášky

- Logovanie a obnova dát

Pripomenutie

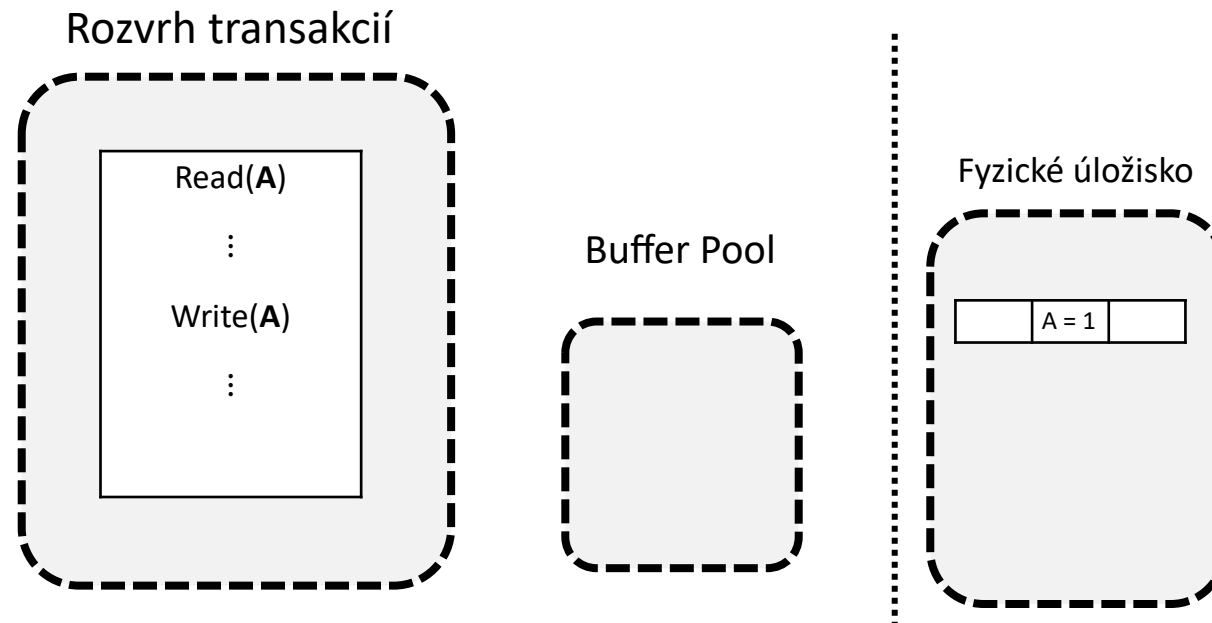
- **ACID**

- Atomicity
- Consistency
- Isolation
- Durability

Logovanie a obnova

Pripomenutie - načítanie a ukladanie dát

- Volatívna vs nevolatívna pamäť



Ako prebieha načítanie a ukladanie informácií

Rozvrh transakcií

Je potrebné načítať hodnotu **A** v rámci transakcie

Read(**A**)

⋮

Write(**A**)

⋮

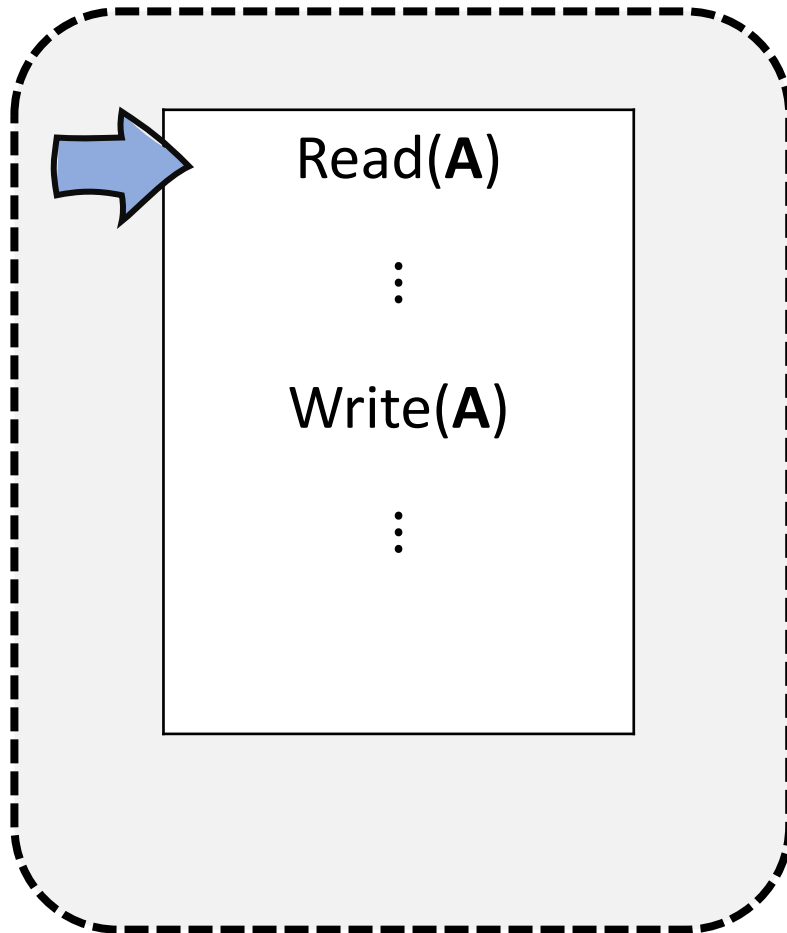
Buffer Pool

Fyzické úložisko

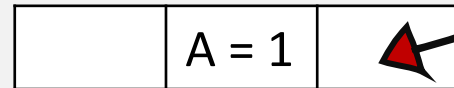
	A = 1	
--	-------	--

Ako prebieha načítanie a ukladanie informácií

Rozvrh transakcií

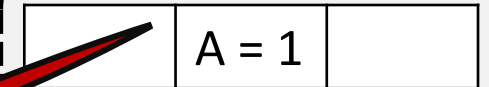


Buffer Pool



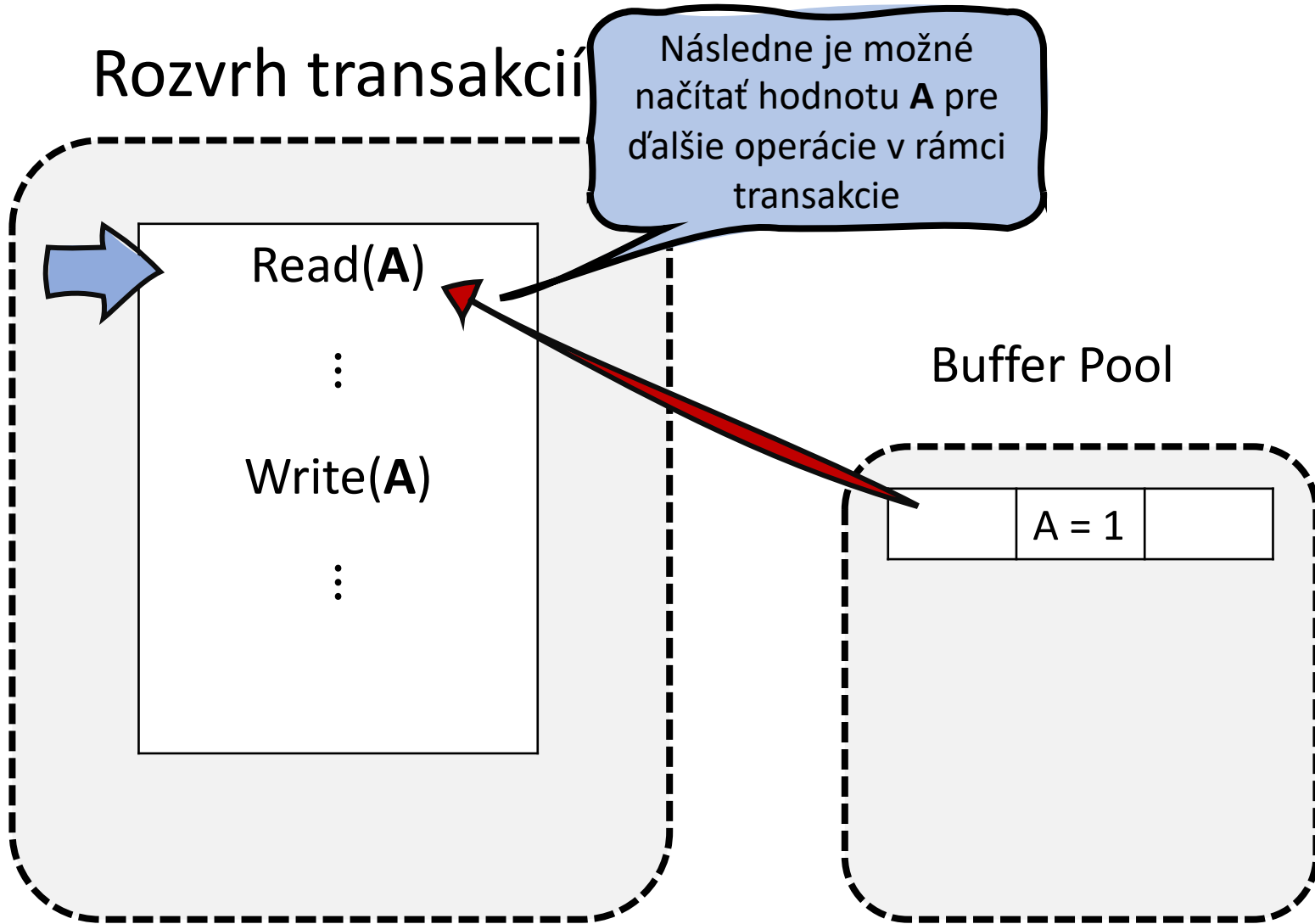
Stránka (Page) je
skopírovaná z disku
do hlavnej pamäte

Fyzické úložisko

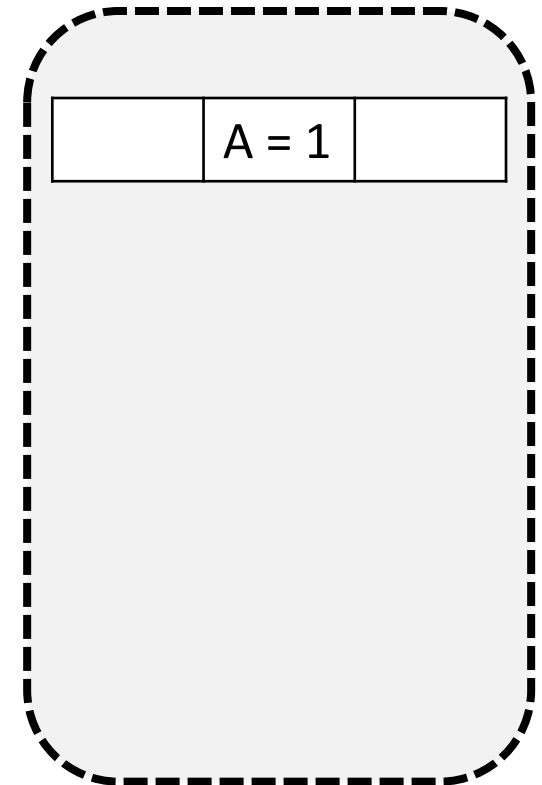


Ako prebieha načítanie a ukladanie informácií

Rozvrh transakcií

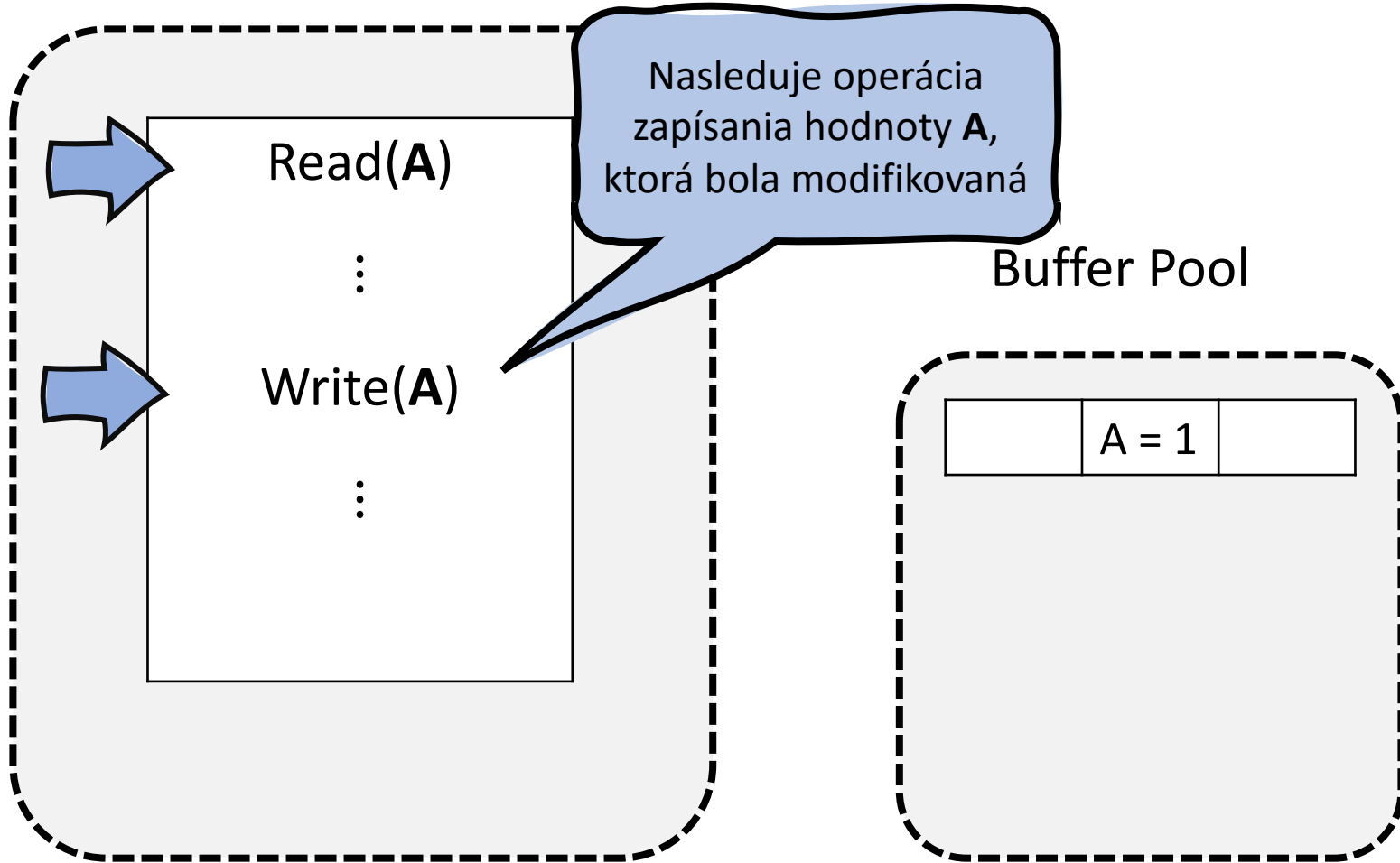


Fyzické úložisko

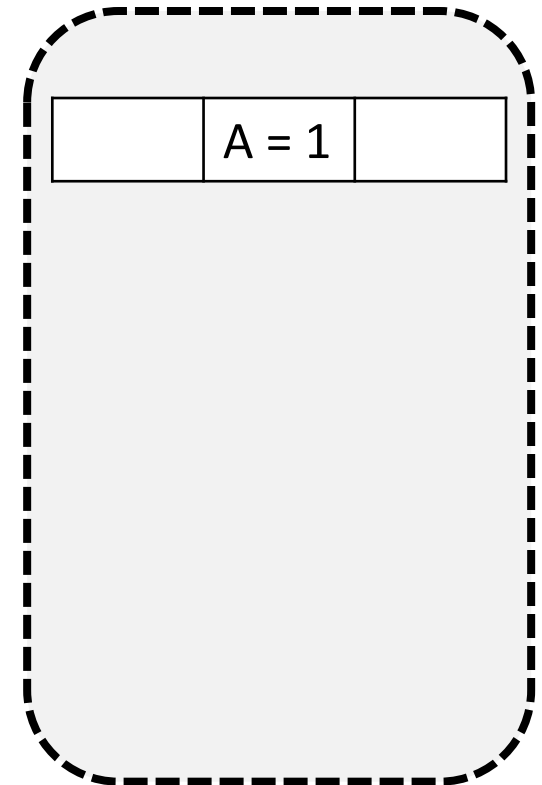


Ako prebieha načítanie a ukladanie informácií

Rozvrh transakcií

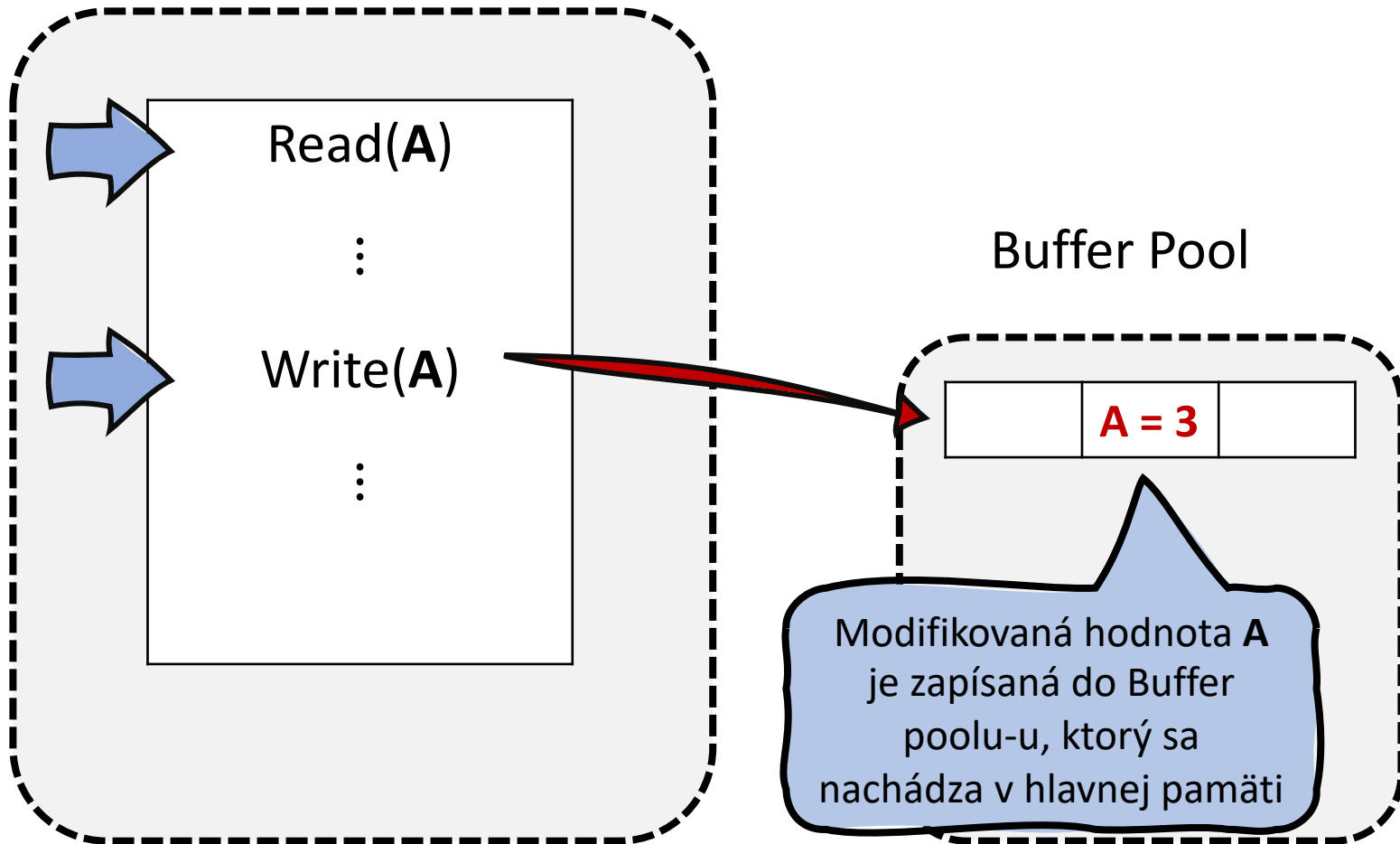


Fyzické úložisko

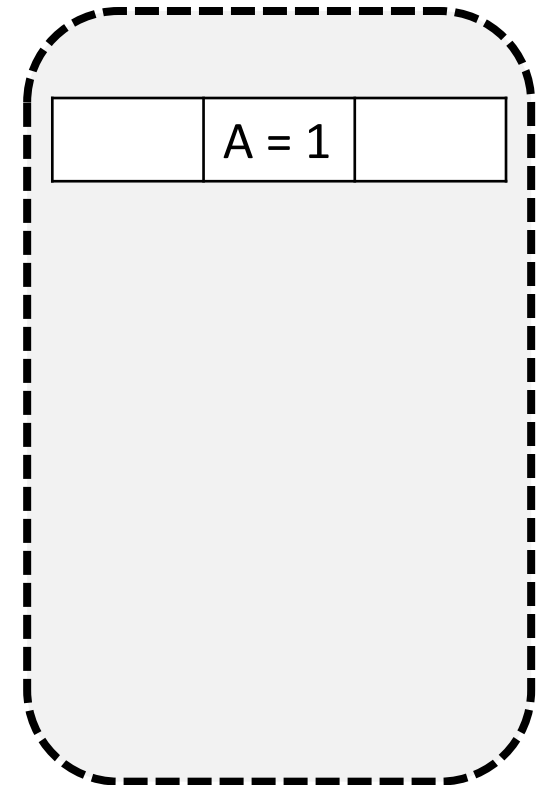


Ako prebieha načítanie a ukladanie informácií

Rozvrh transakcií

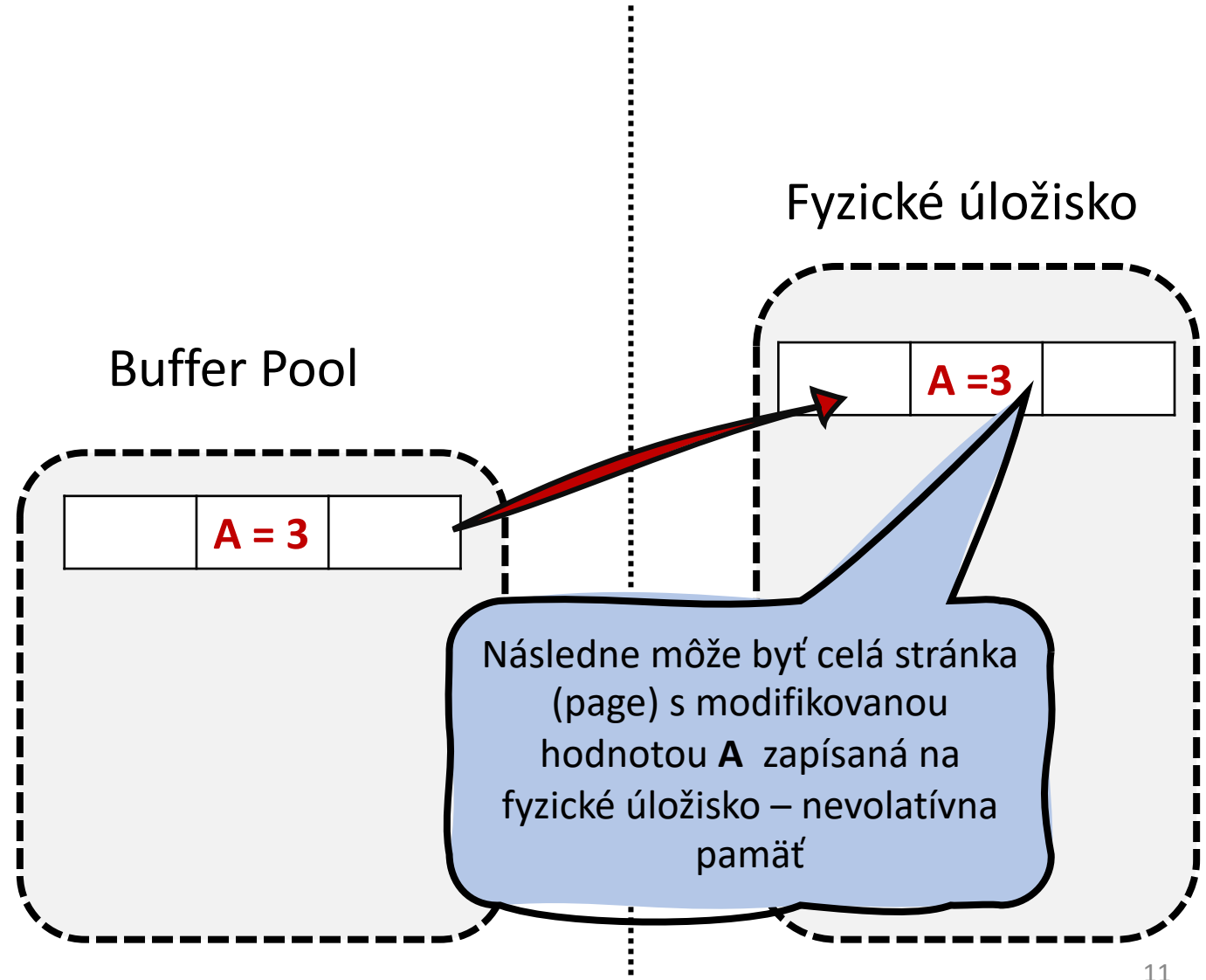
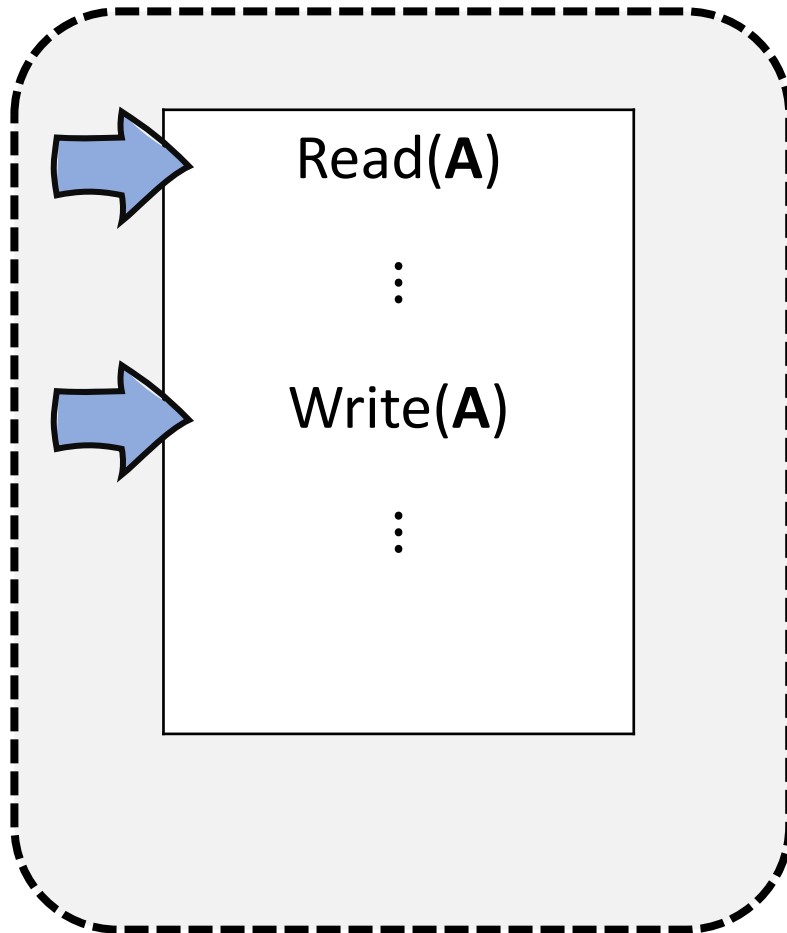


Fyzické úložisko



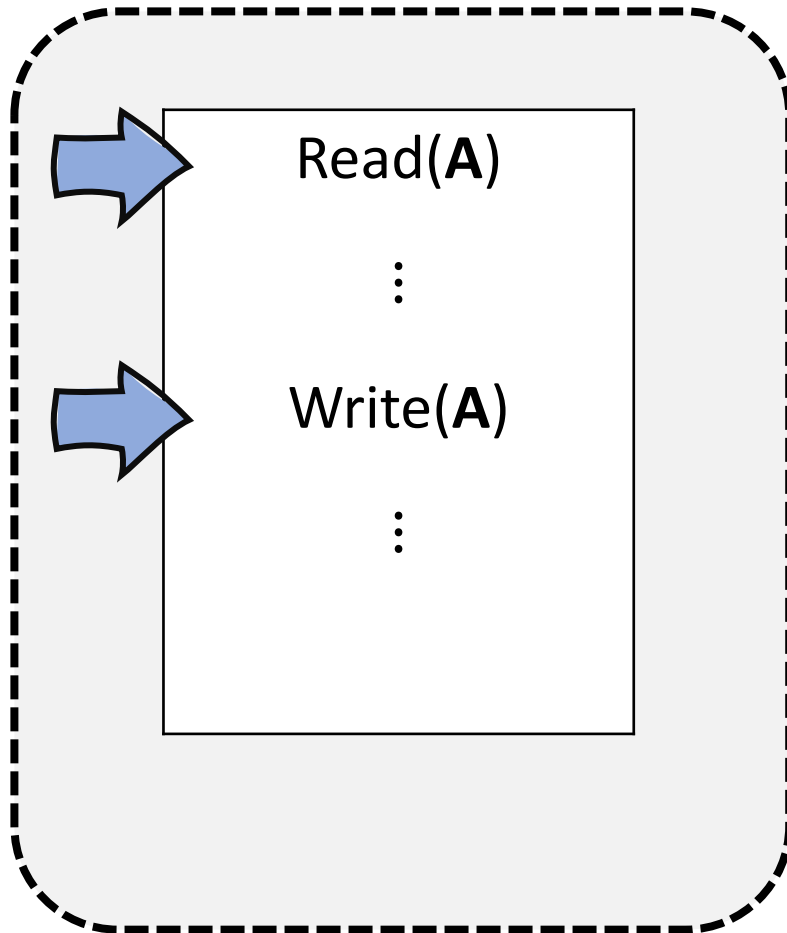
Ako prebieha načítanie a ukladanie informácií

Rozvrh transakcií



Ako prebieha načítanie a ukladanie informácií

Rozvrh transakcií

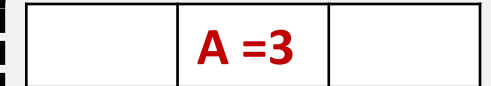


Buffer Pool



Stránka môže zostať v buffer pool-e, alebo môže byť odstránená v závislosti od spravovania Buffer Pool-u

Fyzické úložisko



Hodnota je uložená a v prípade výpadku je modifikovaná hodnota **A** zachovaná

Obnova databázy po zlyhaní

- Algoritmy pre obnovu (Recovery Algorithm) predstavujú mechanizmy, ktorých cieľom je zabezpečiť, že databáza bude:
 - konzistentná
 - atomicitu transakcií
 - a tiež trvácnosť (durability) databázy
- Celkovo algoritmy obnovy môžeme rozdeliť na dve časti:
 - Operácie, ktoré sa vykonávajú počas bežného spracovania transakcií – cieľom je zabezpečenie informácií, ktoré je možné použiť následne pri obnove, keď nastane zlyhanie
 - Operácie, ktoré sa vykonávajú po zlyhaní (vypádku) pre obnovenie DB pre zachovanie atomicity, konzistentnosti a trvácnosti (durability)

Motivácia - prečo riešiť obnovu dát?

- Cieľ dosiahnuť ACID
 - v rámci algoritmu pre obnovu dát po zlyhaní vieme dosiahnuť atomicitu, konzistenciu a trvácnosť - izoláciu nám riešia protokoly pre riadenie súbežnosti napr. 2-phase locking (2PL)

Príklad

- Máme rozvrh transakcií, ktorý obsahuje transakciu **T1**
- Transakcia **T1** robí operácie nad objektom **A** (čítanie a zápis)
- Následne urobí COMMIT

Motivácia - príklad

Rozvrh transakcií

T1

BEGIN

Read(A)

⋮

Write(A)

COMMIT

Buffer Pool

A = 3

Hodnota **A** bola
modifikovaná

Fyzické úložisko

A = 1

Motivácia - príklad

Rozvrh transakcií

T1

BEGIN

Read(A)

⋮

Write(A)

COMMIT

Buffer Pool

A = 3

Hodnota **A** bola
modifikovaná

Je hodnota **A** v tomto čase
zapísaná na fyzické
úložisko ?

fyzické úložisko

A = 1

Motivácia - príklad

Rozvrh transakcií

T1

BEGIN

Read(A)

⋮

Write(A)

COMMIT

Buffer Pool

A = 3

Hodnota **A** bola
modifikovaná

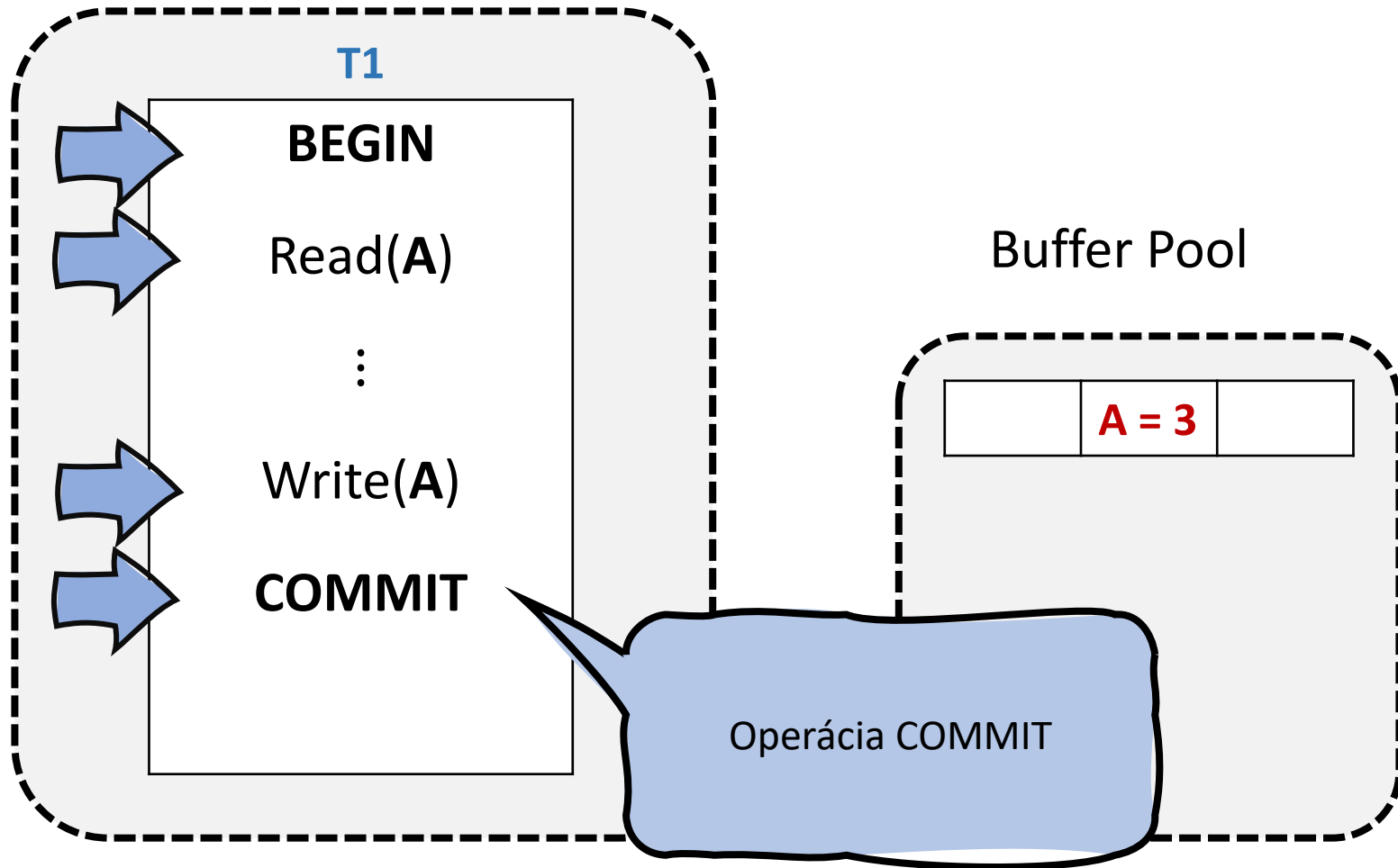
Môže, ale tiež nemusí
závisí od spravovania
Buffera a mechanizmu
logovania a obnovy

zické úložisko

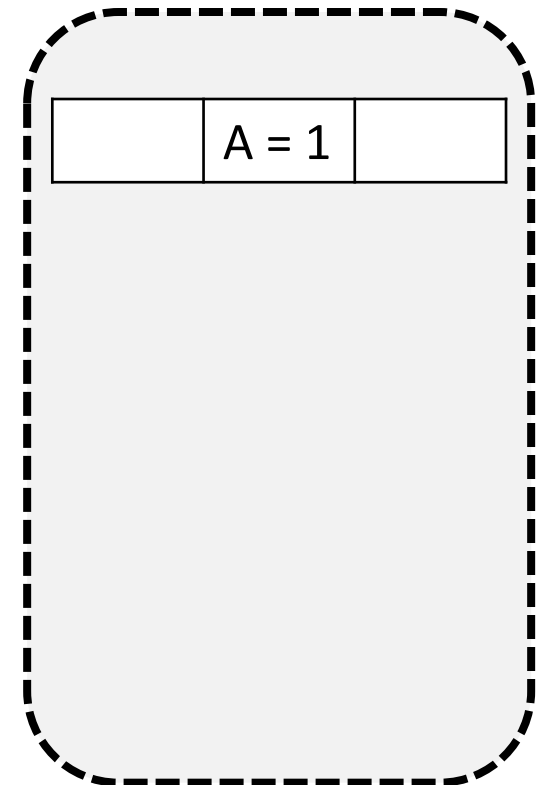
A = 1

Motivácia - príklad

Rozvrh transakcií

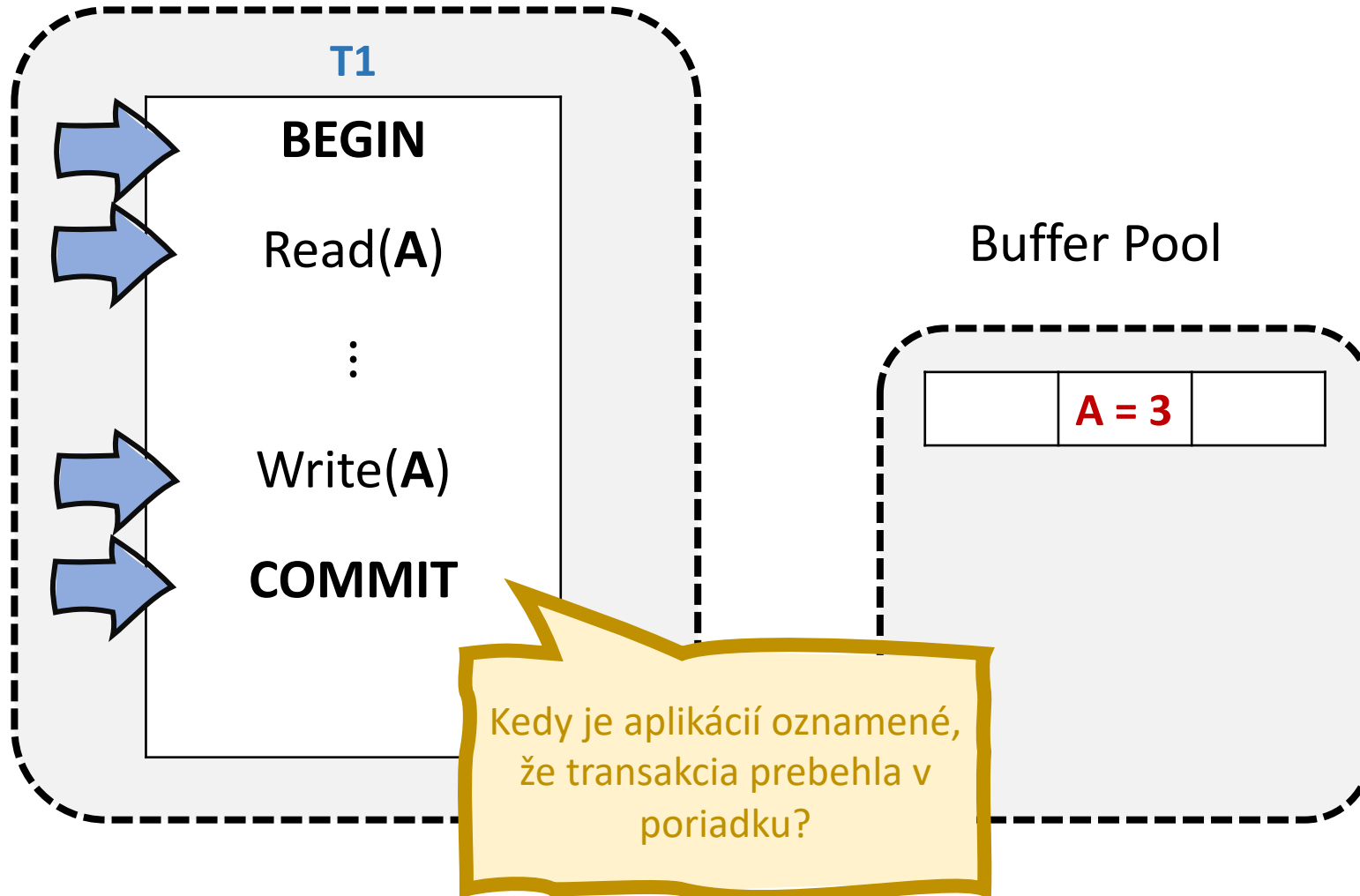


Fyzické úložisko

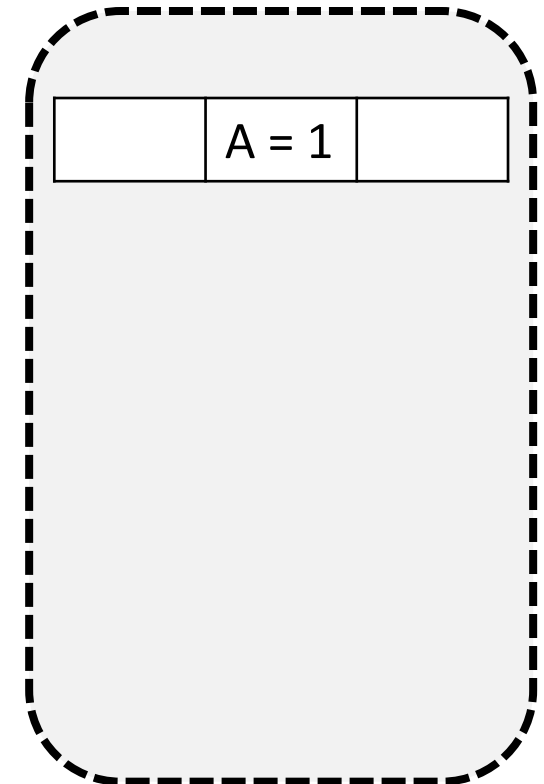


Motivácia - príklad

Rozvrh transakcií

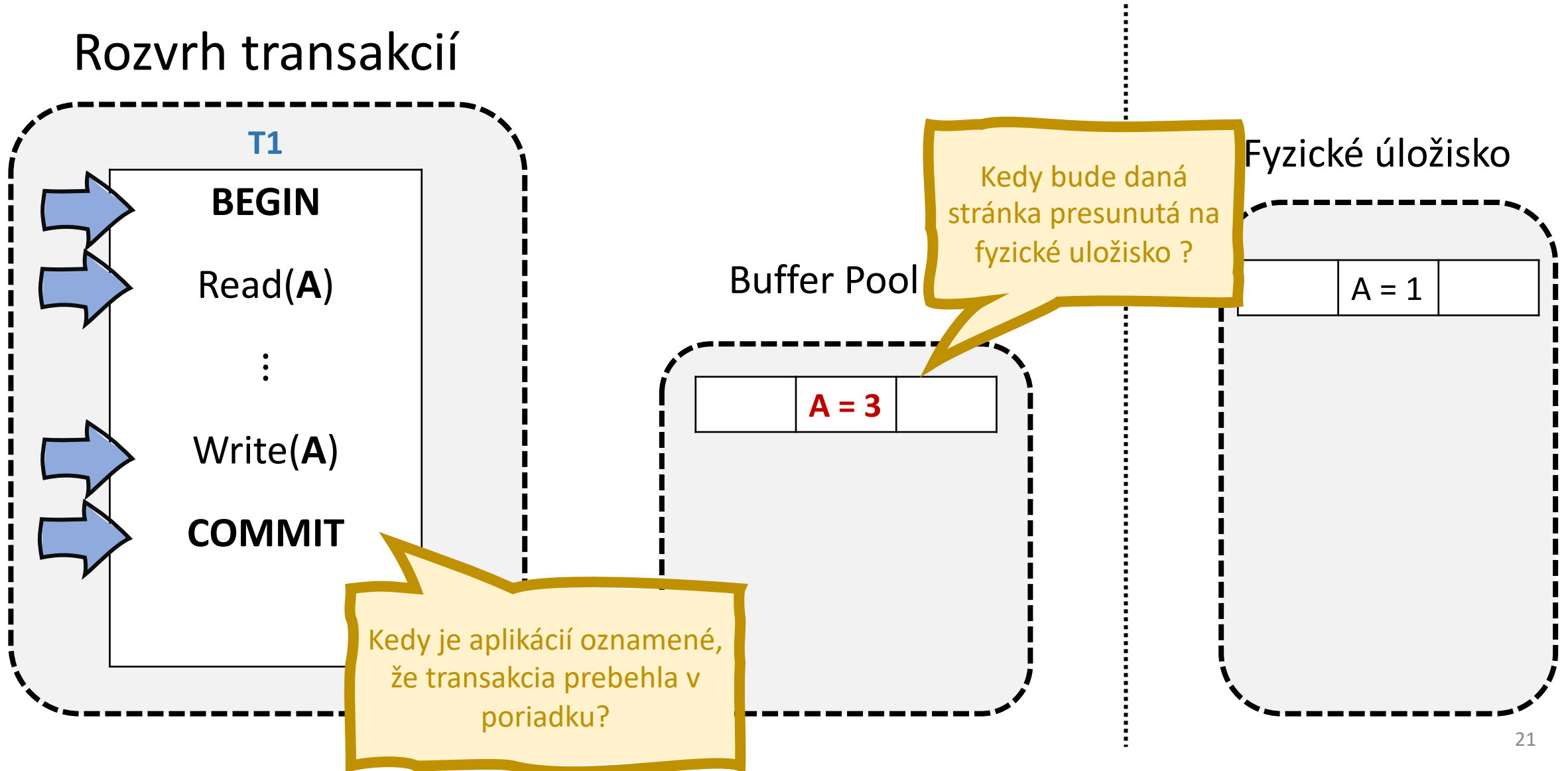


Fyzické úložisko



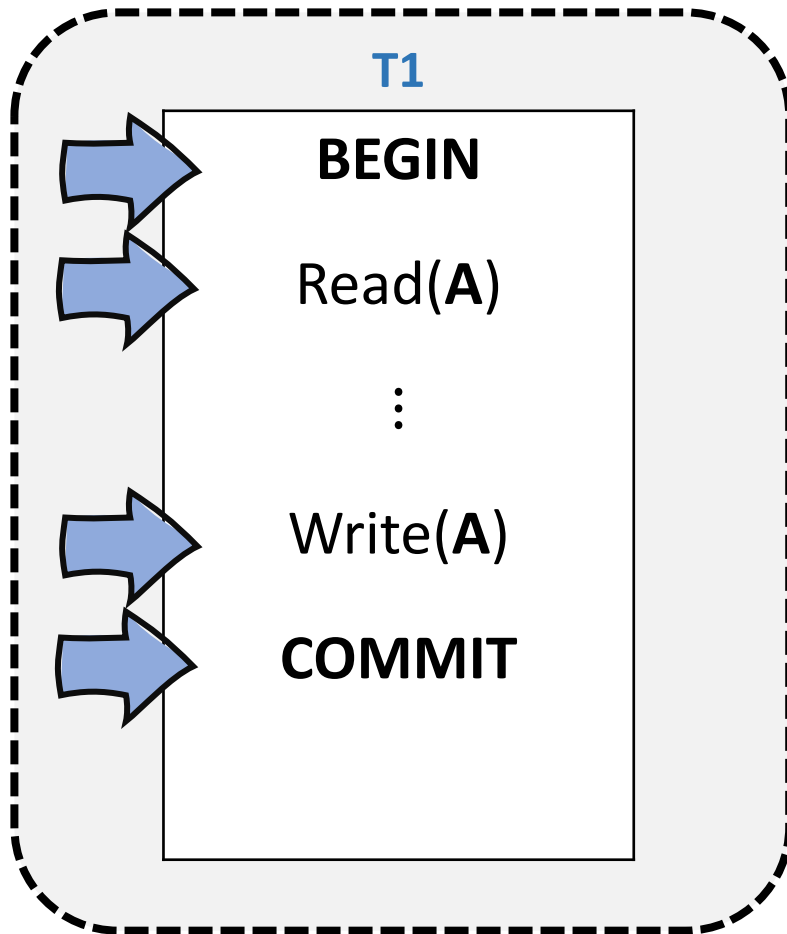
Motivácia - príklad

Rozvrh transakcií

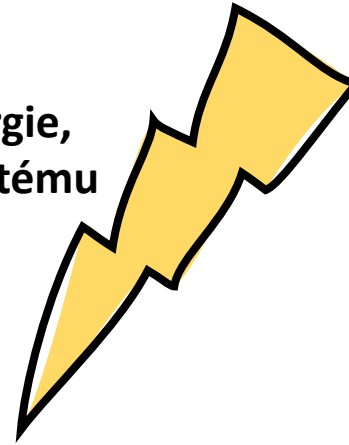


Motivácia - príklad

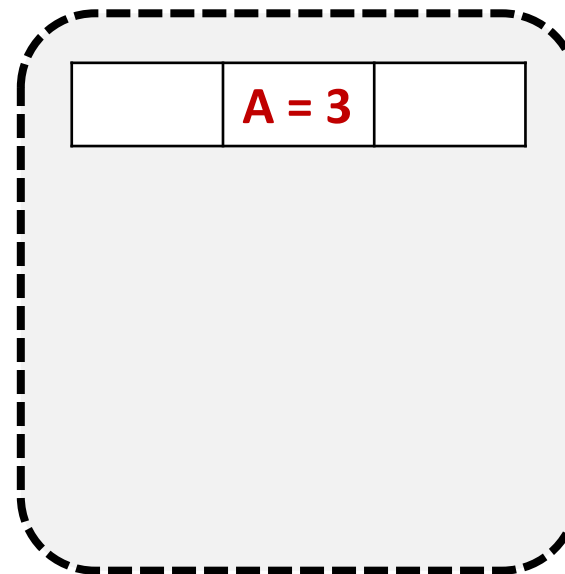
Rozvrh transakcií



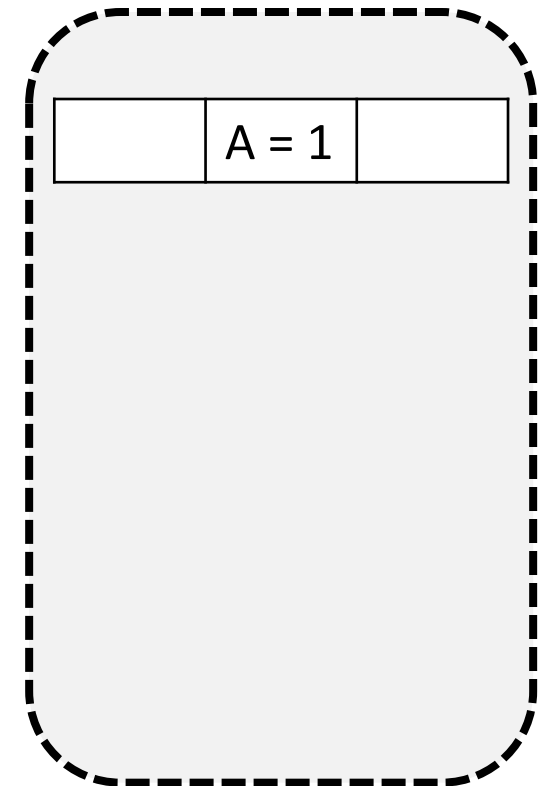
Výpadok – el. energie,
útok, padnutie systému



Buffer Pool



Fyzické úložisko



Typy zlyhaní

- Rozlišujeme typy zlyhaní
 - Zlyhanie transakcie
 - Zlyhanie systému
 - Zlyhanie úložiska (Storage)

Zlyhanie transakcie

- Zlyhanie transakcie môžeme rozdeliť na dve časti:
 - **Logické chyby** – transakcia nemôže byť dokončená v dôsledku nejakej vnútornej chyby napr. táto chyba vyvolá obmedzenie v DB
 - **Chyba vnútorného stavu** (Internal state error) – DBMS musí ukončiť/zrušiť transakciu v dôsledku chyby napr. vznikne deadlock

Zlyhanie systému

- **Softvérové zlyhanie** – nedostatky v implementácií DBMS napr. delenie nulou, s ktorým sa nepočítalo pri implementácií
- **Hardvérové zlyhanie**
 - Počítač, ktorý hostuje DBMS padne napr. výpadok prúdu
 - Predpokladá sa, že výpadok nemá vplyv na nevolatívnu pamäť – nie je poškodená (Fail-stop assumption)

Zlyhanie úložiska (storage)

- **Neopraviteľná hardvérová chyba**
 - zlyhanie disku napr. poškodenie sektorov, hlavy disku atď.
 - je predpoklad, že dané zlyhanie v rámci úložiska je detegovateľné napr. pomocou checksum
- DBMS sa nevie zotaviť z danej chyby – DB musí byť obnovená z nejakej zálohy

Zápis dát do nevolatívnej pamäte

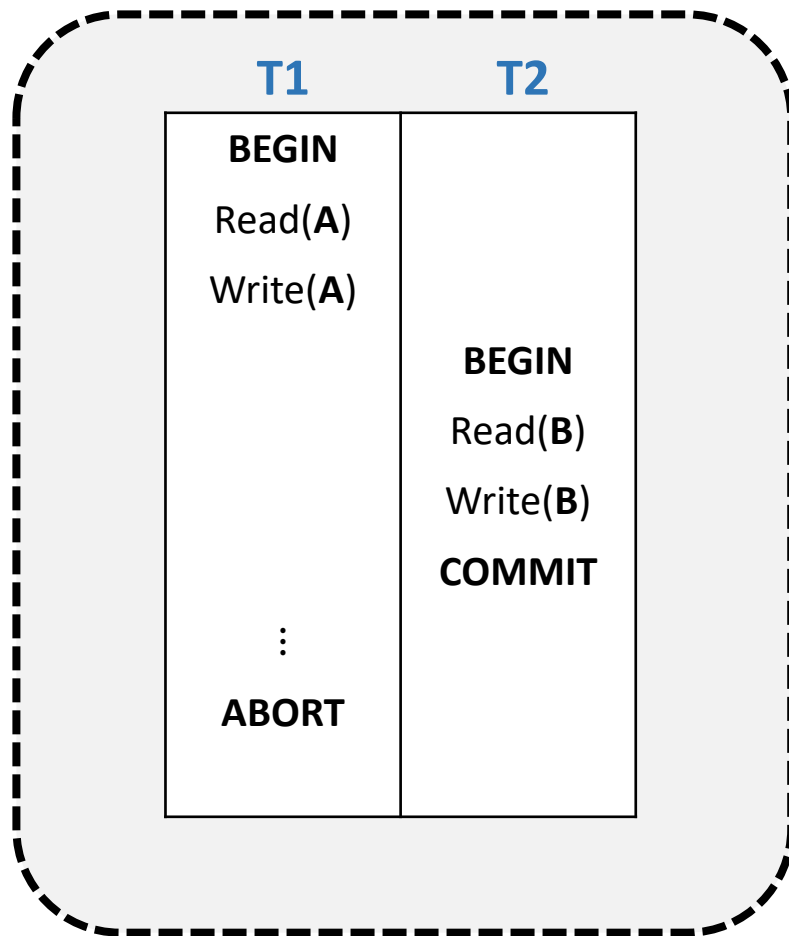
- Zápisanie hodnoty môže prebehnúť dvomi spôsobmi:
 - Okamžitým zápisom dát - Immediate data modification
 - Dáta môžu byť zapísané na disk aj pred tým ako transakcia dosiahne COMMIT
 - Je nutné robiť UNDO aj REDO operácie
 - Oneskoreným zápisom dát - Deferred data modification
 - Dáta sú zapísané na disk až keď bol dosiahnutý COMMIT
 - Nie je nutné robiť UNDO operácie, postačujú REDO operácie
 - Potreba cache
- Tieto dva spôsoby súvisia s obnovovacími algoritmami a logovaním
 - preto spomenutie operácií UNDO a REDO

Manažment - Buffer pool

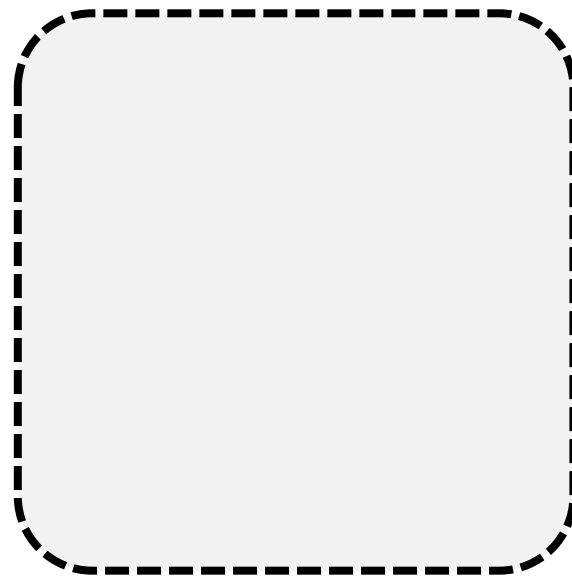
- Rozhoduje o tom, kedy dôjde k zápisu modifikovanej stránky na disk
 - niektoré operácie môžu vyžadovať okamžitý zápis informácií, niektoré nie
- V prípade plnej pamäte rozhoduje, ktoré stránky opustia pamäť a budú zapísané na disk
 - používajú sa algoritmy známe z OS (napr. LRU, FIFO, atď.)
- Súvisí s tým politika manažmentu buffer
 - Steal policy
 - Force policy

Manažment - buffer pool príklad

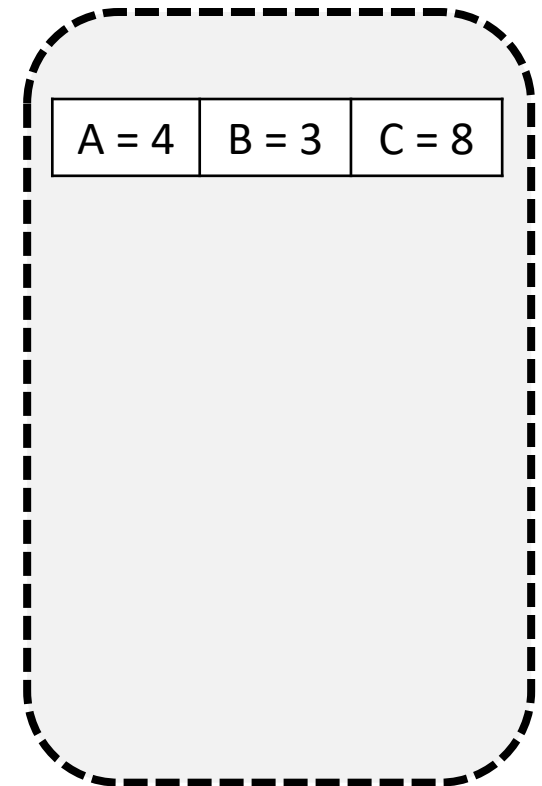
Rozvrh transakcií



Buffer Pool

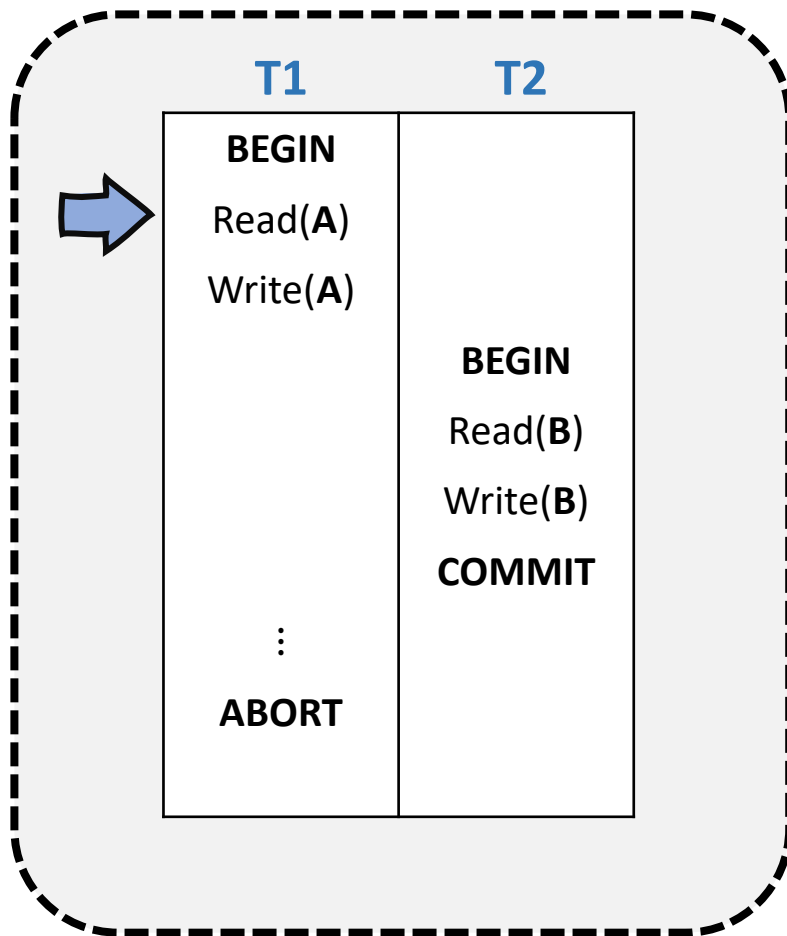


Fyzické úložisko

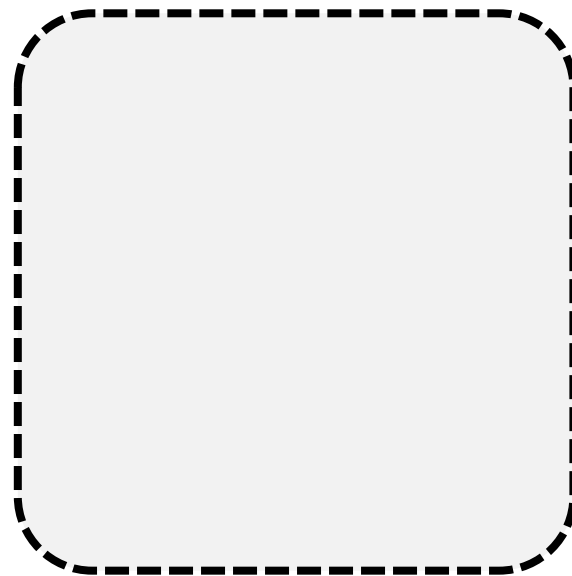


Manažment - buffer pool príklad

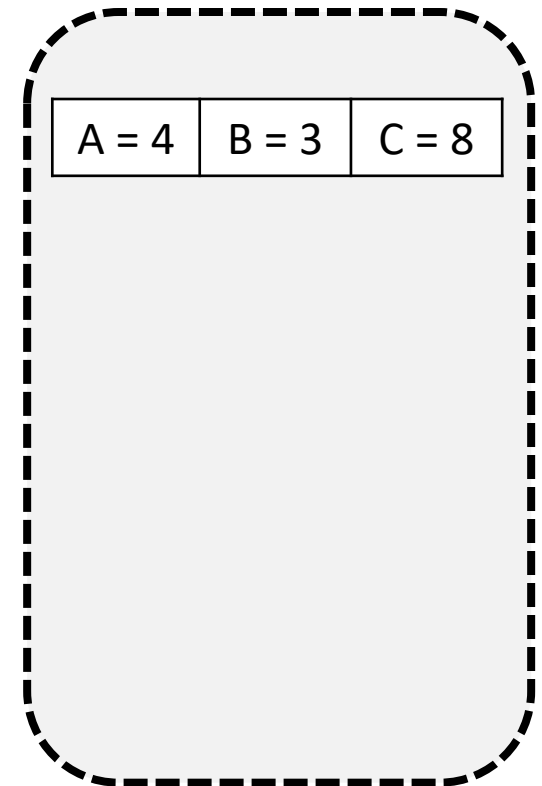
Rozvrh transakcií



Buffer Pool

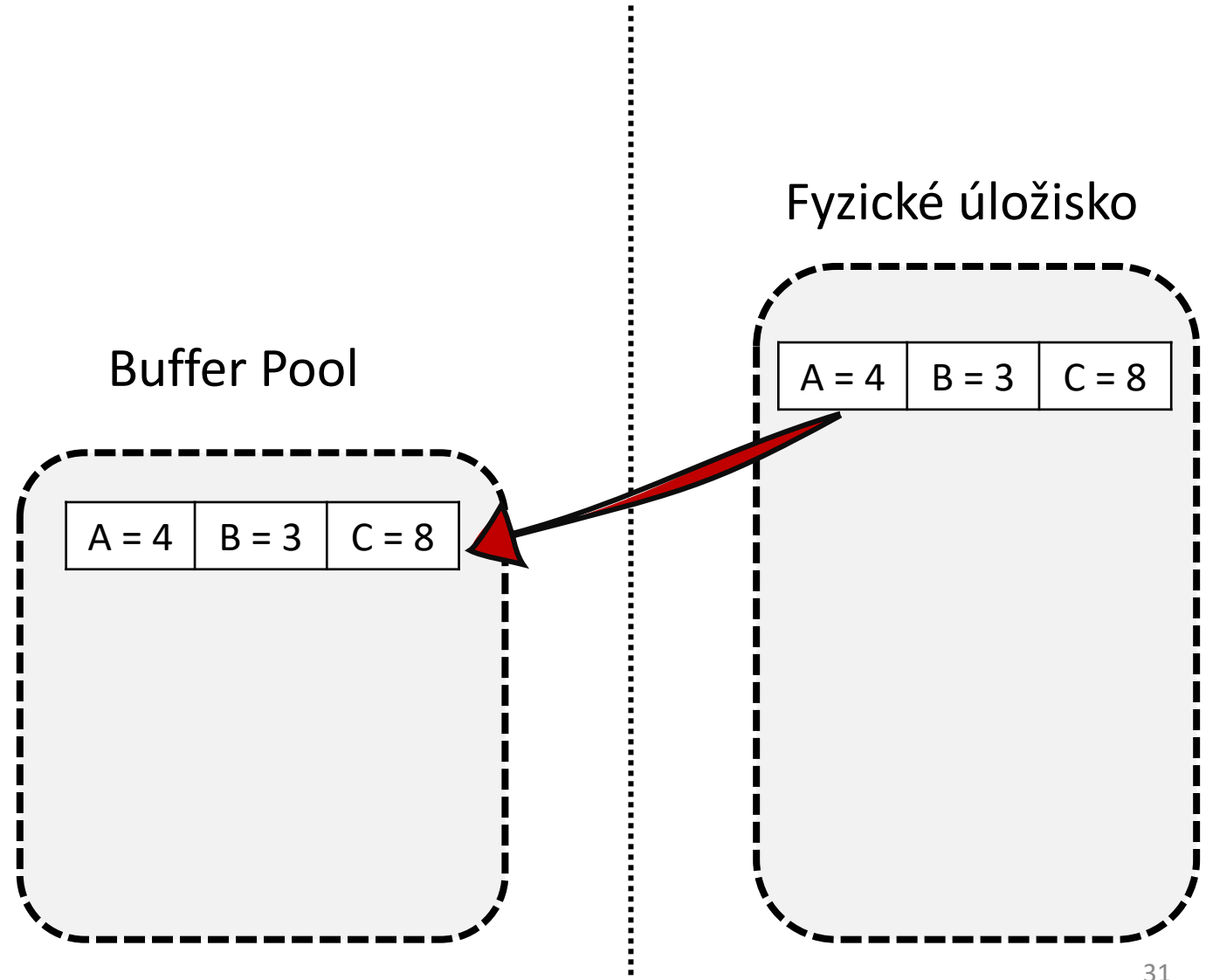
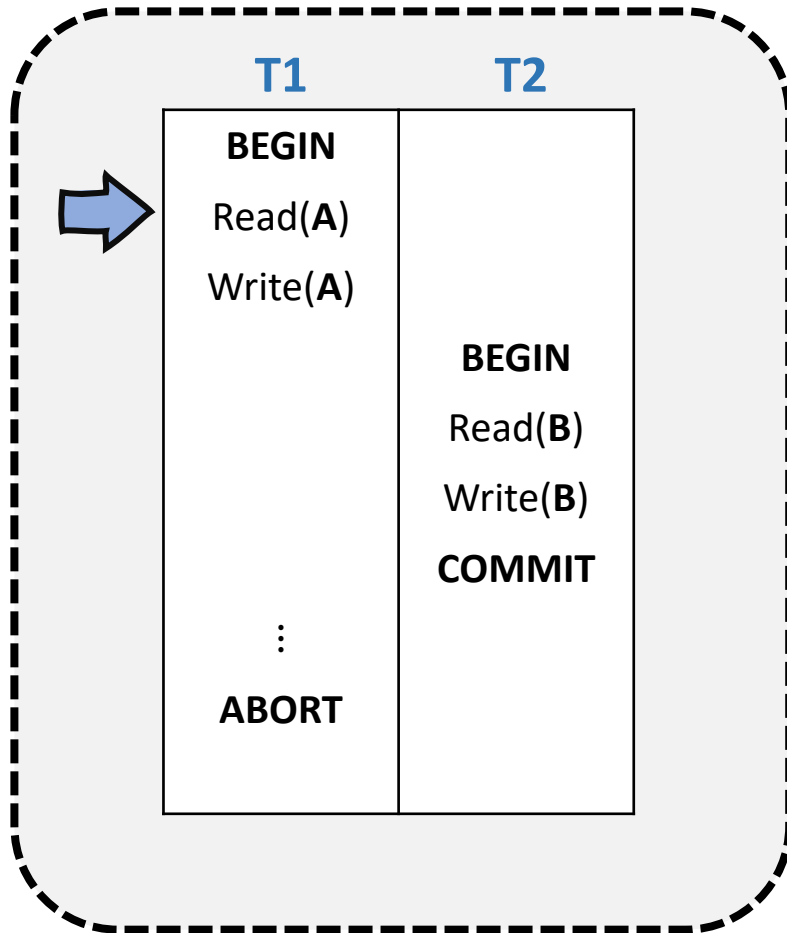


Fyzické úložisko



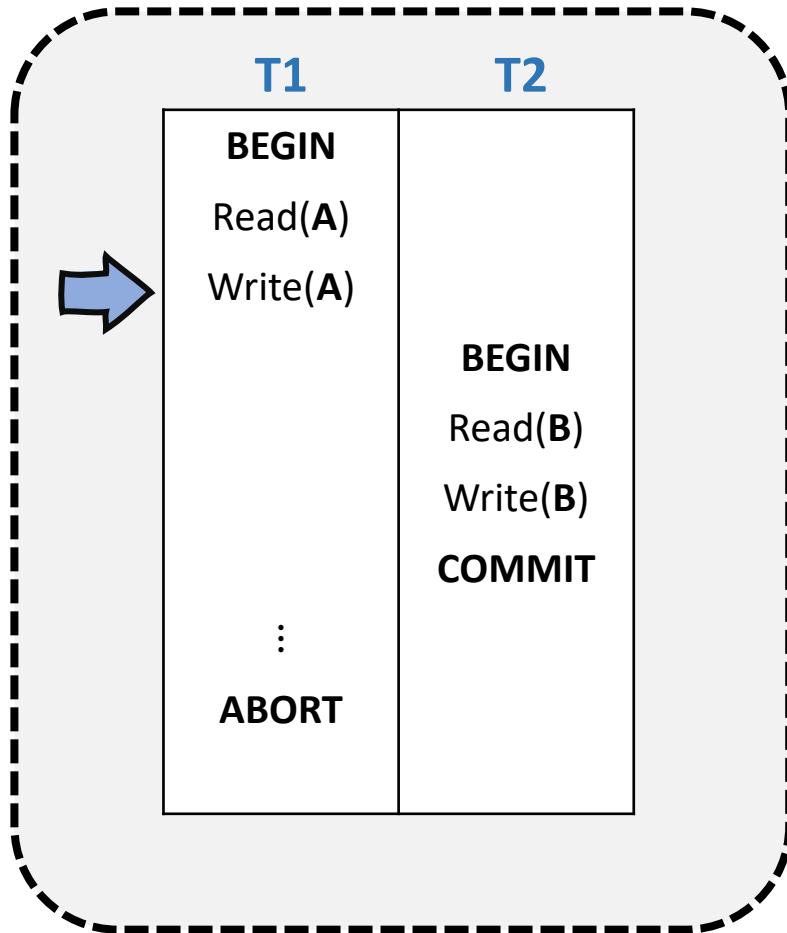
Manažment - buffer pool príklad

Rozvrh transakcií

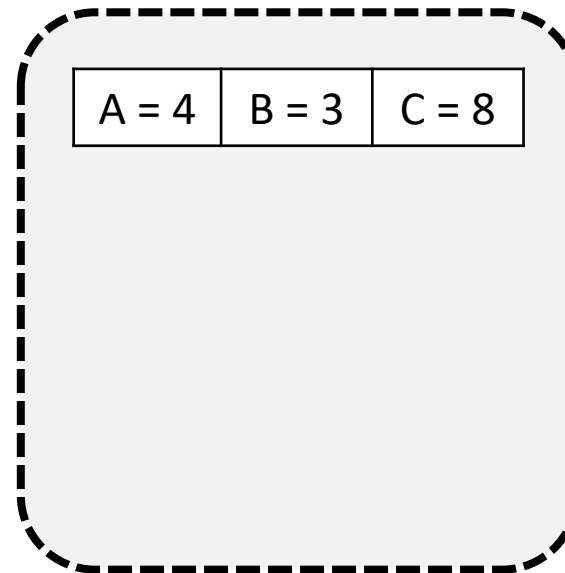


Manažment - buffer pool príklad

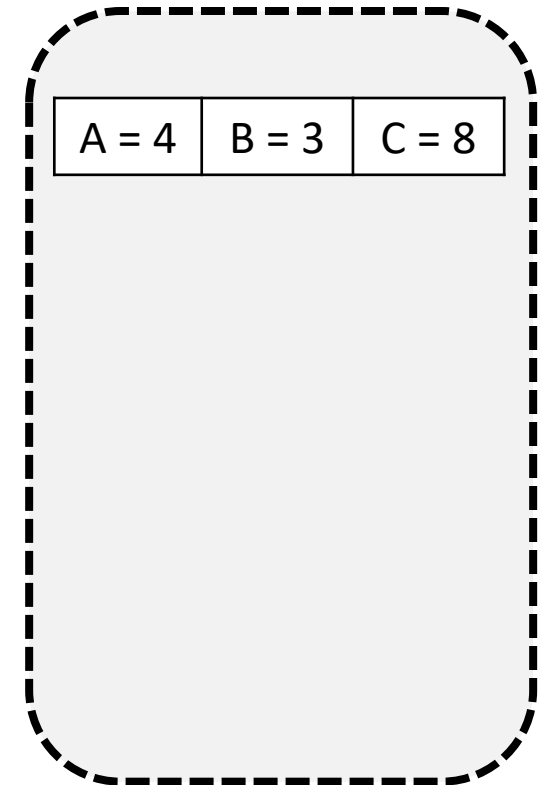
Rozvrh transakcií



Buffer Pool

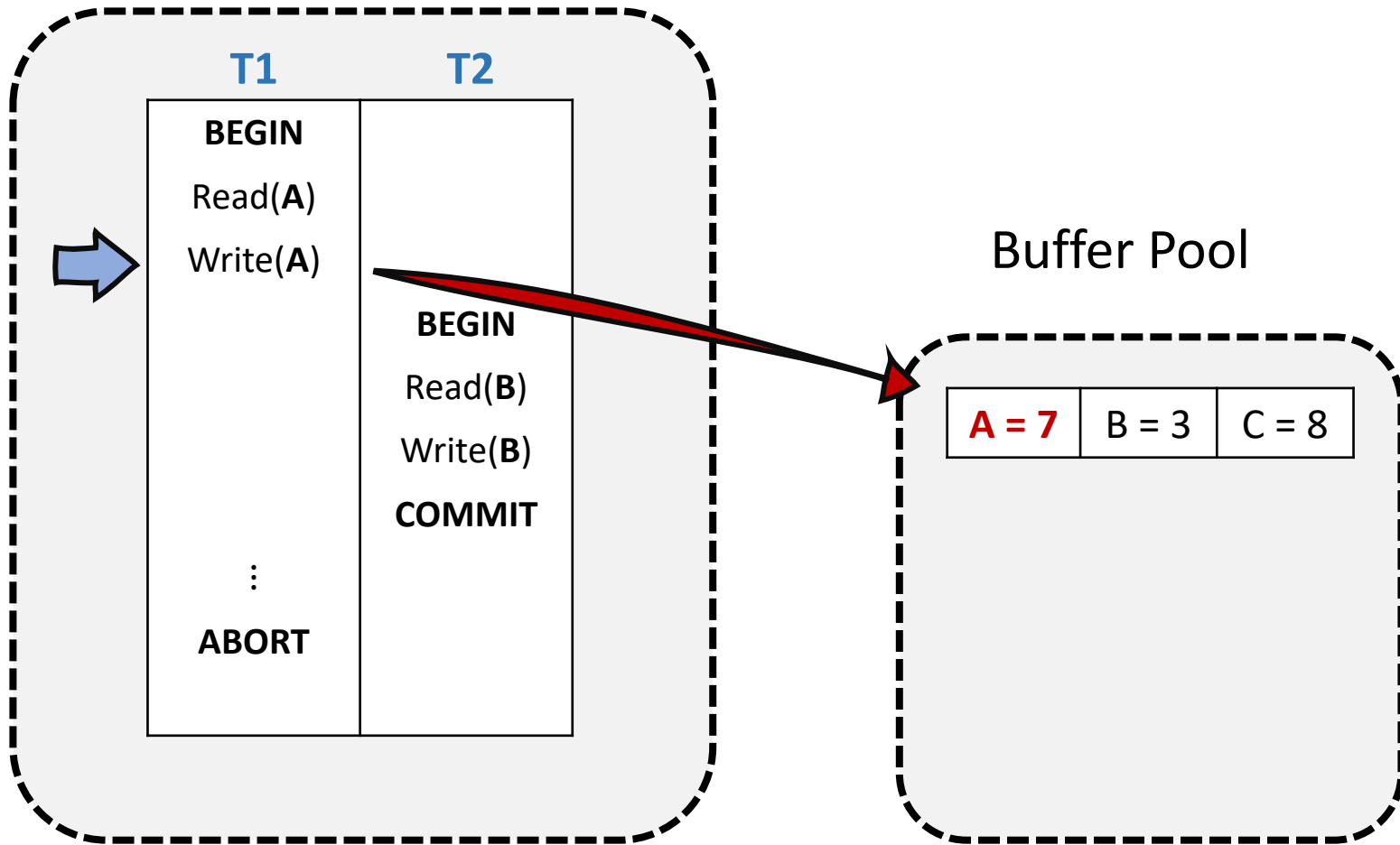


Fyzické úložisko

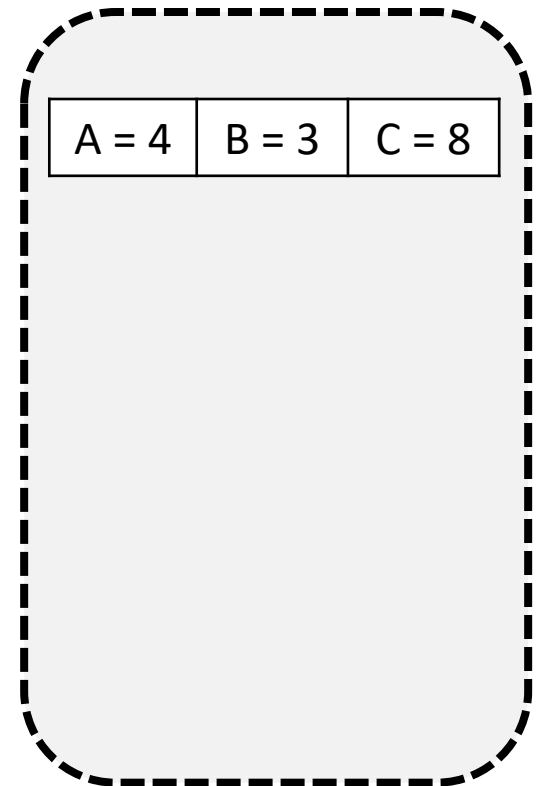


Manažment - buffer pool príklad

Rozvrh transakcií

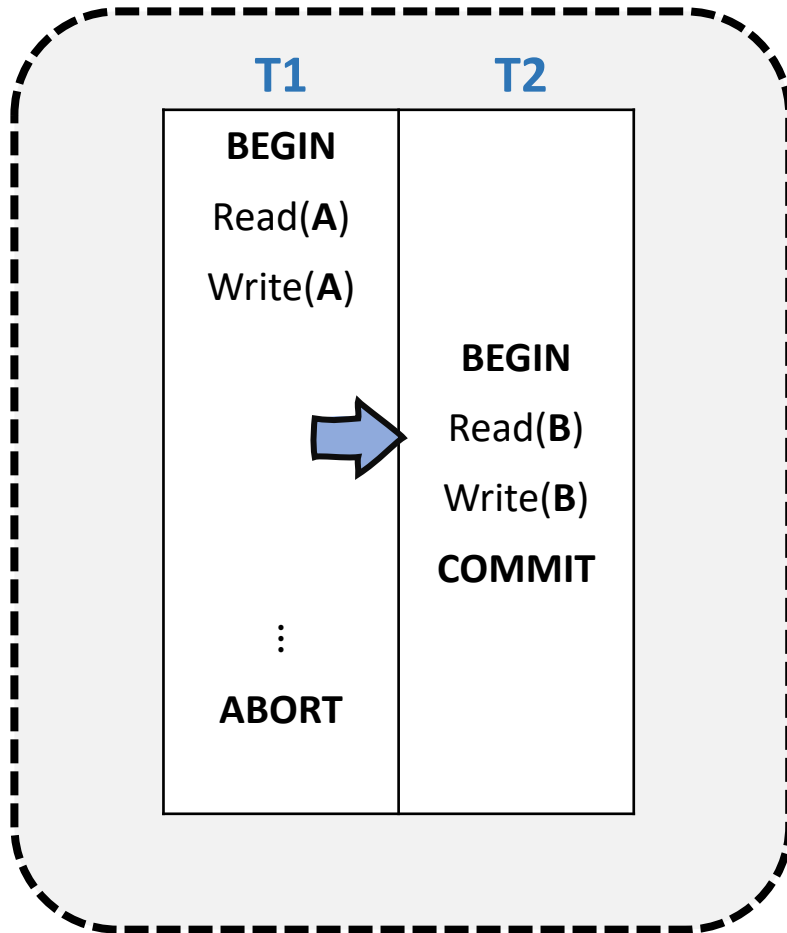


Fyzické úložisko

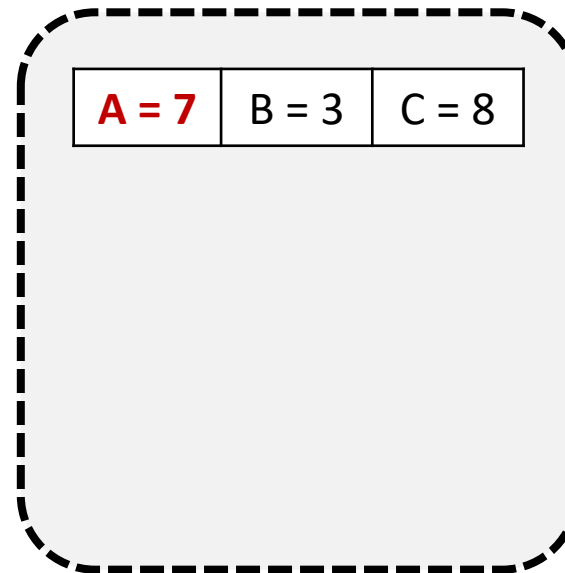


Manažment - buffer pool príklad

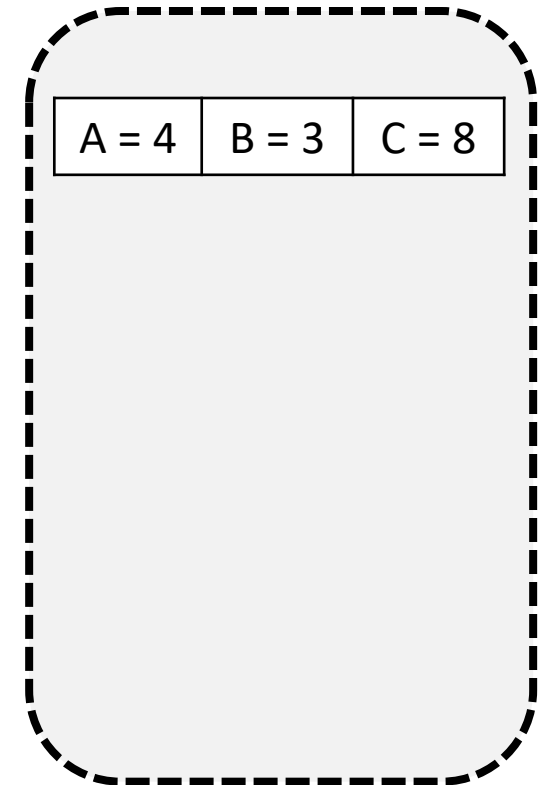
Rozvrh transakcií



Buffer Pool

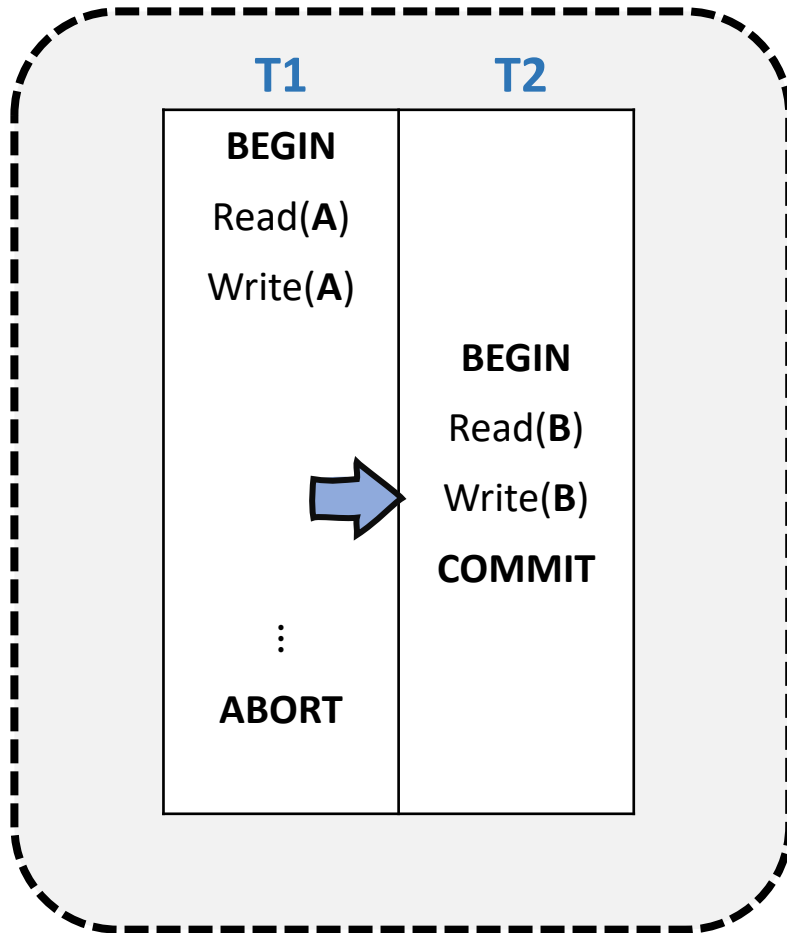


Fyzické úložisko

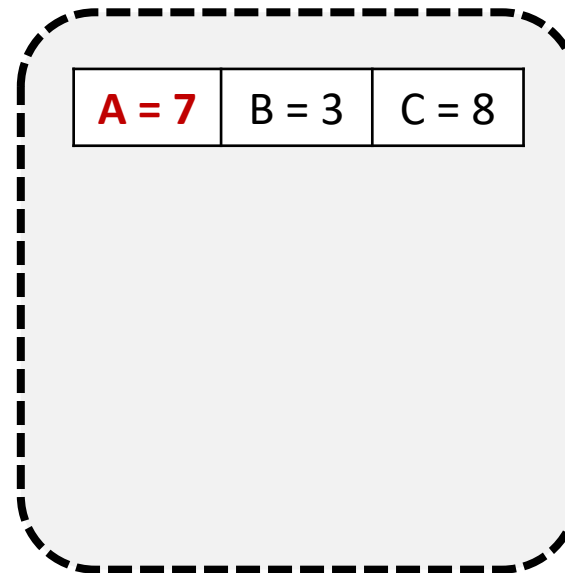


Manažment - buffer pool príklad

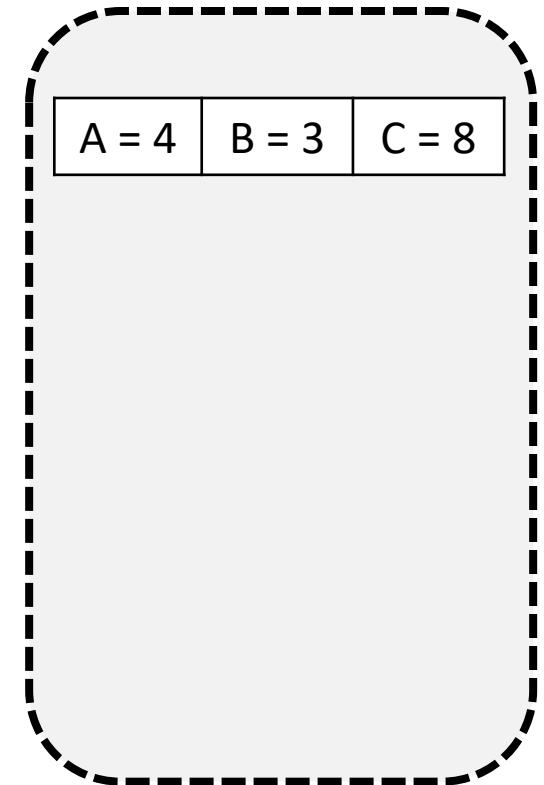
Rozvrh transakcií



Buffer Pool

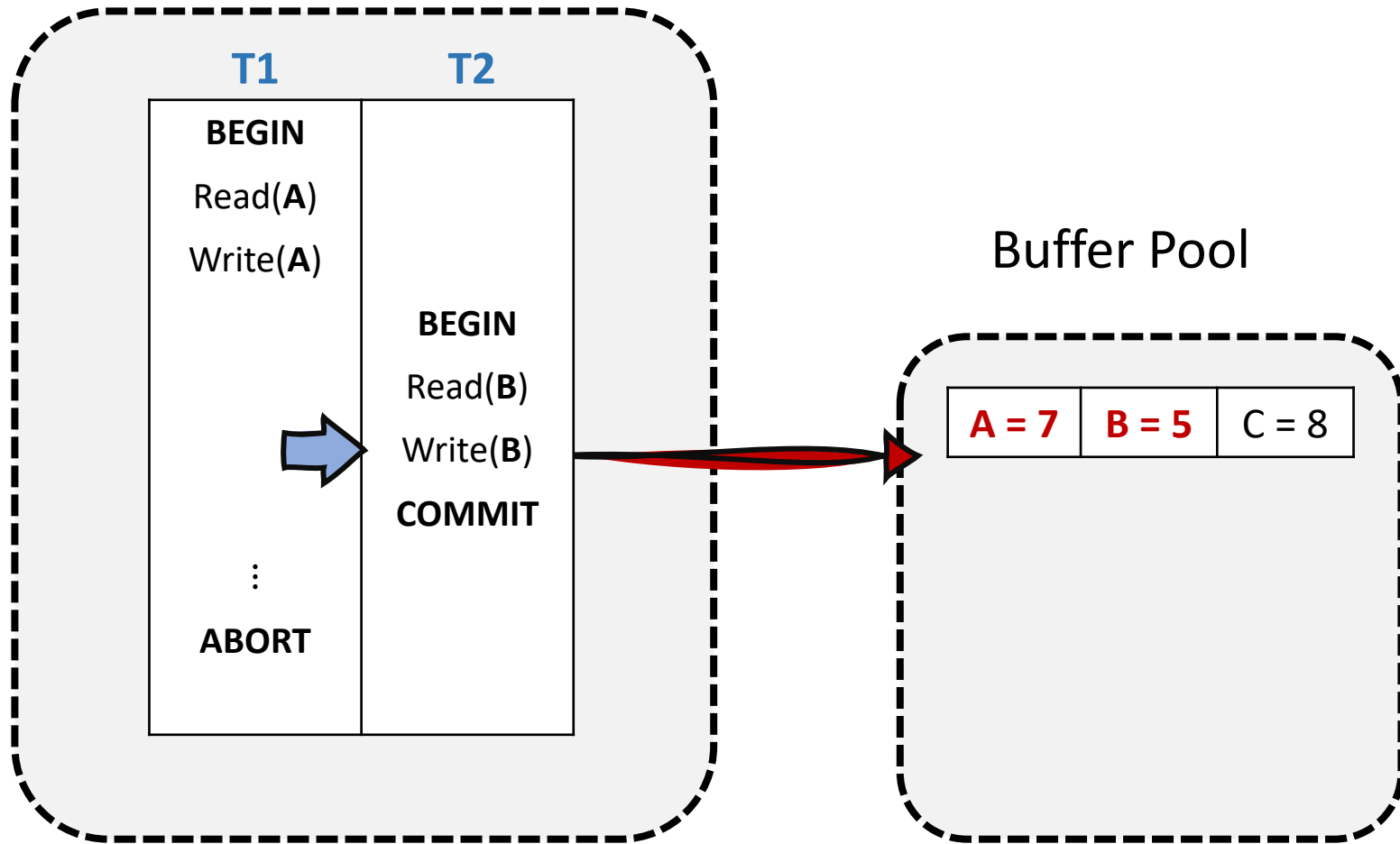


Fyzické úložisko

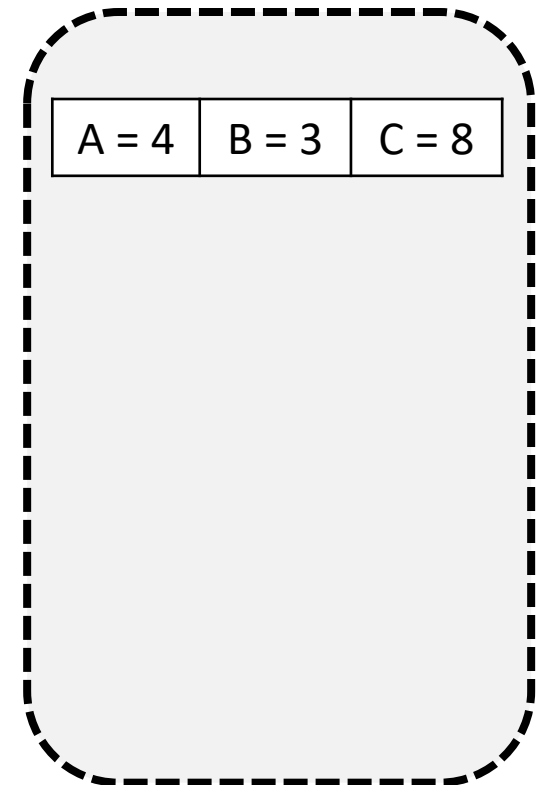


Manažment - buffer pool príklad

Rozvrh transakcií

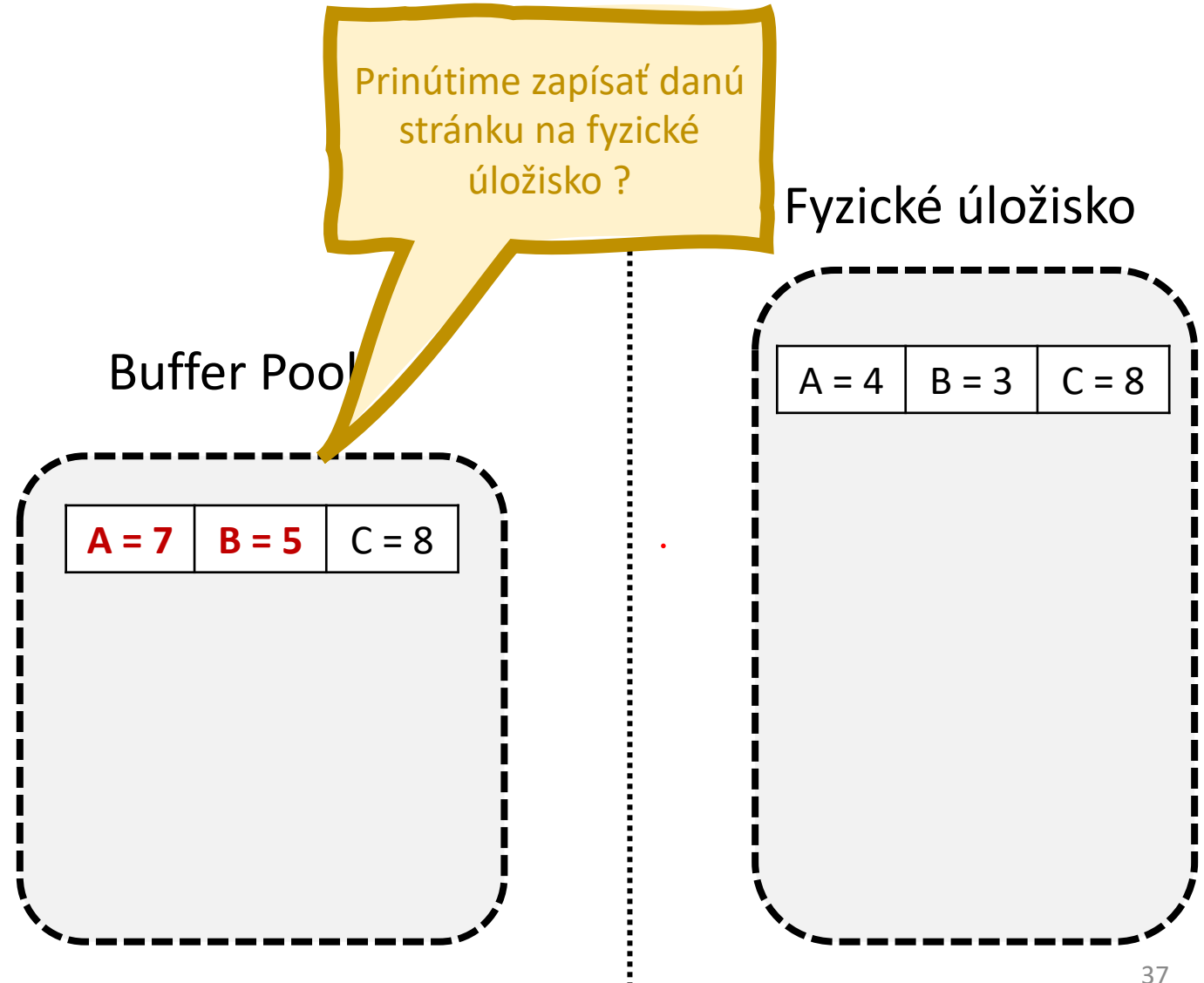
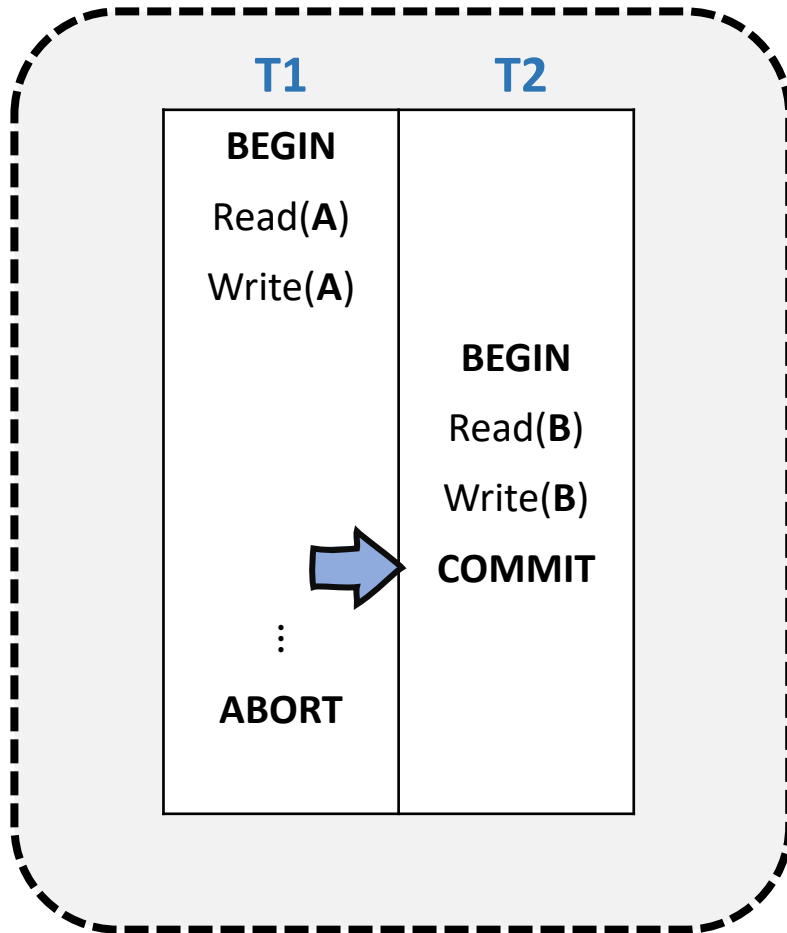


Fyzické úložisko



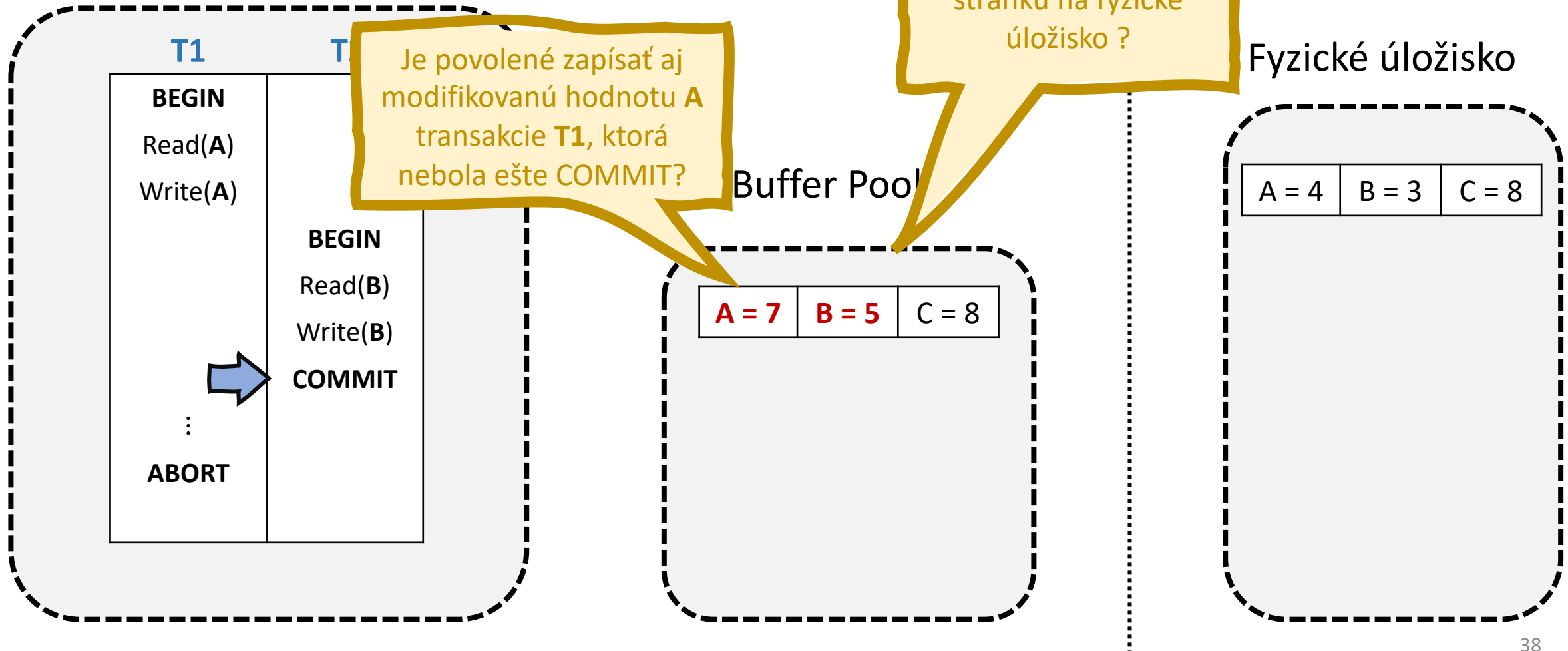
Manažment - buffer pool príklad

Rozvrh transakcií



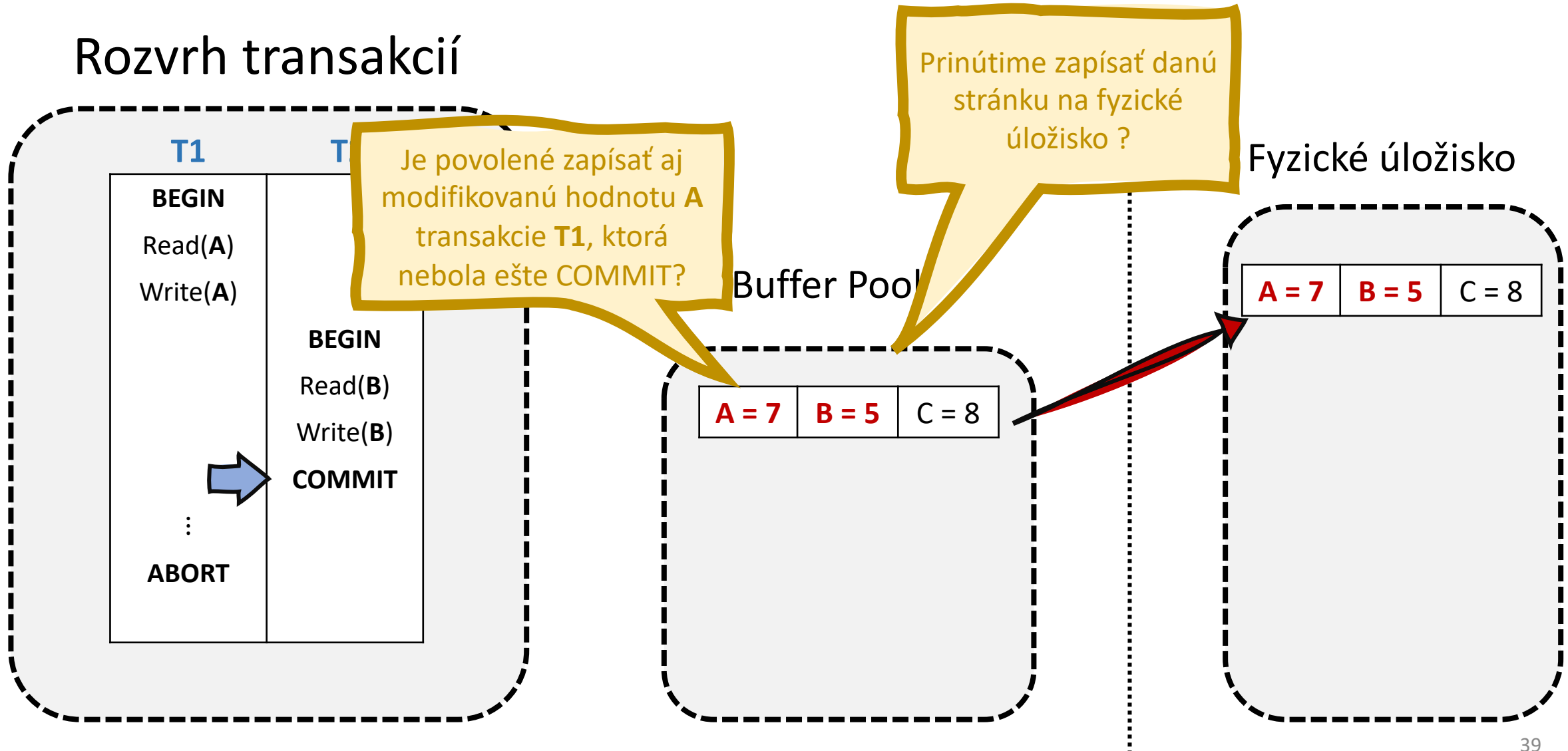
Manažment - buffer pool príklad

Rozvrh transakcií



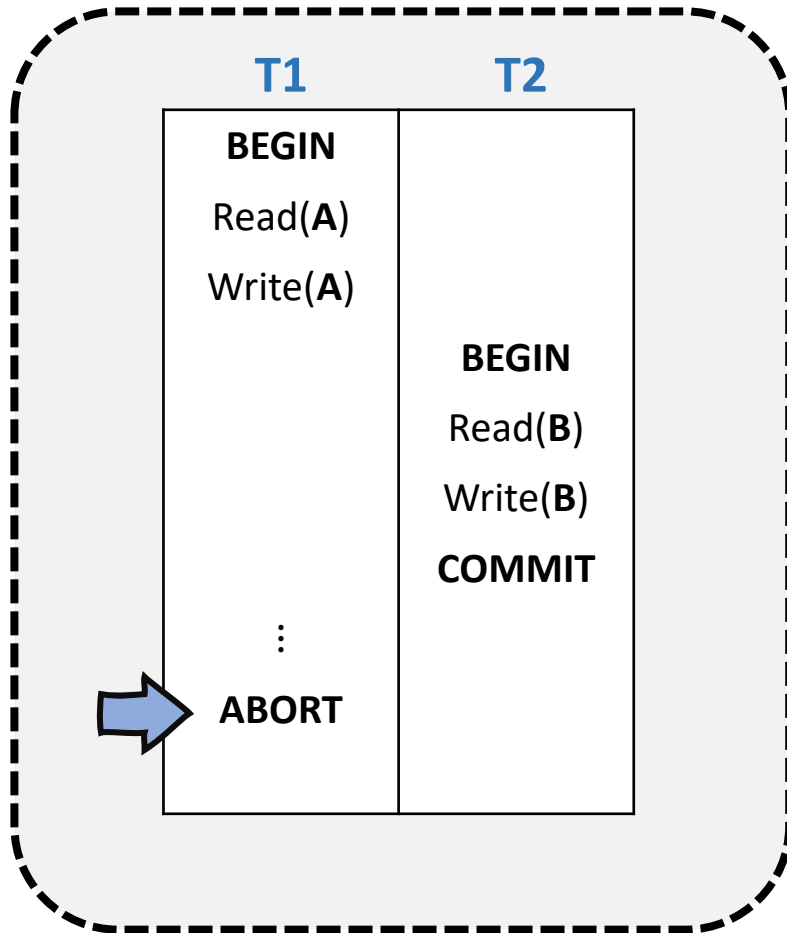
Manažment - buffer pool príklad

Rozvrh transakcií

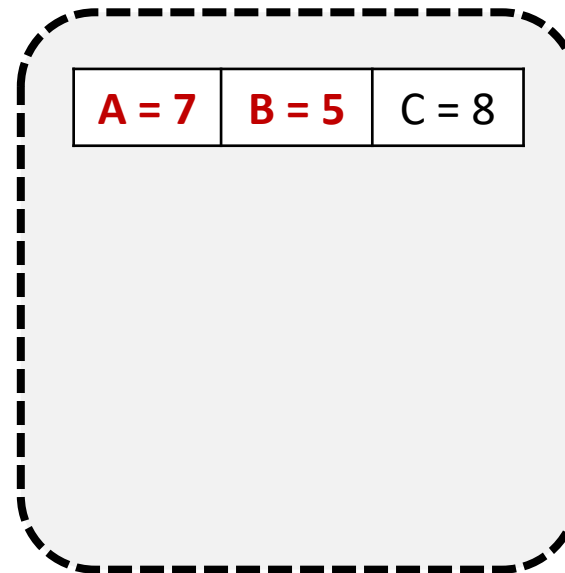


Manažment - buffer pool príklad

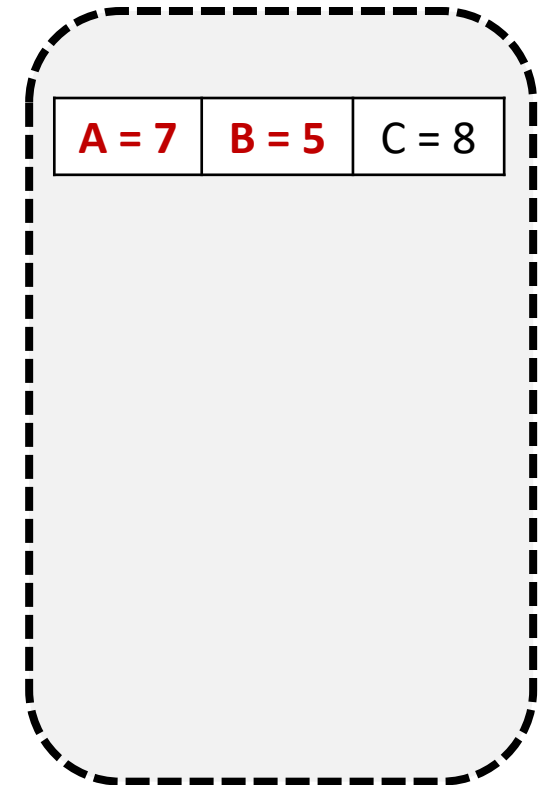
Rozvrh transakcií



Buffer Pool

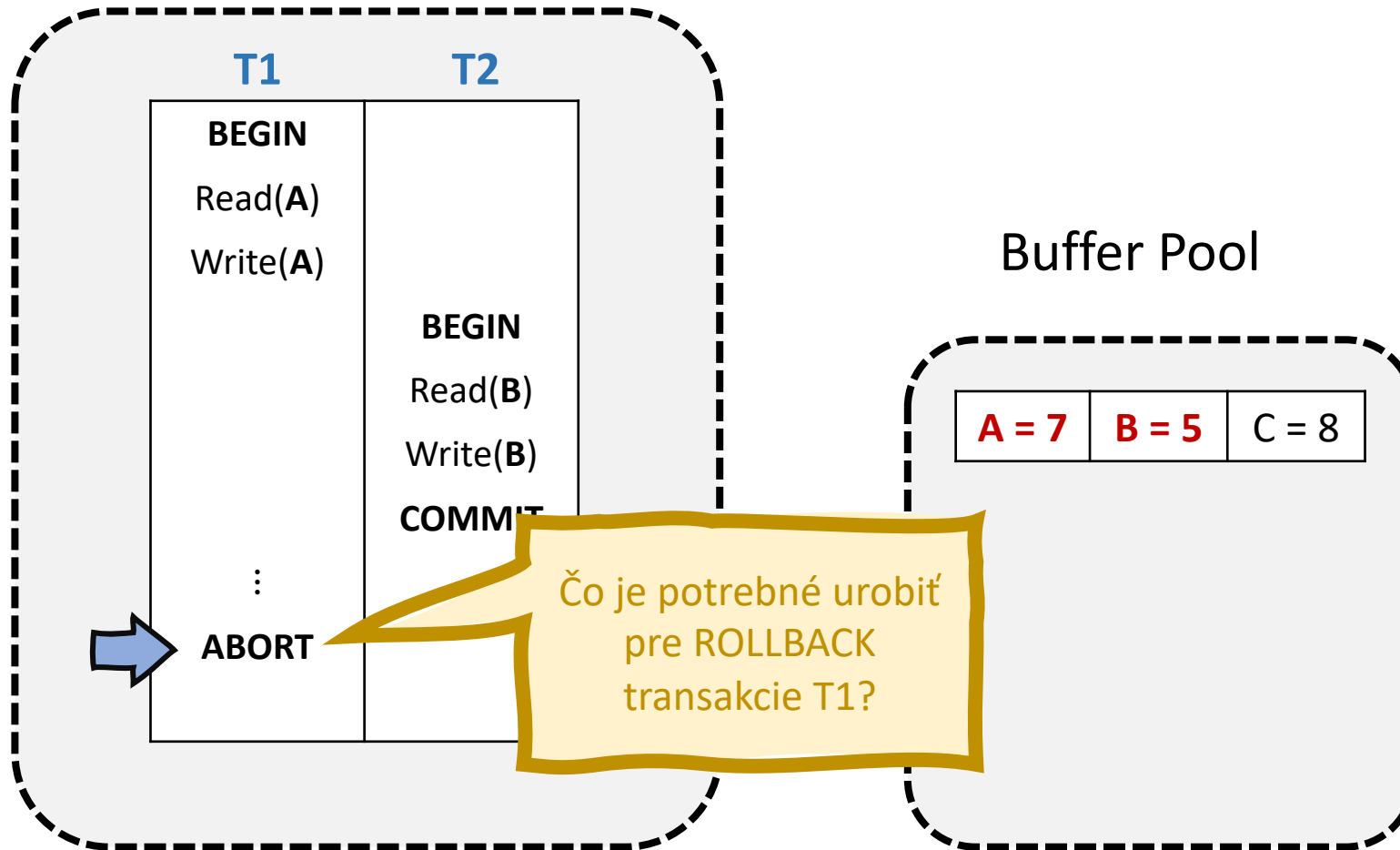


Fyzické úložisko

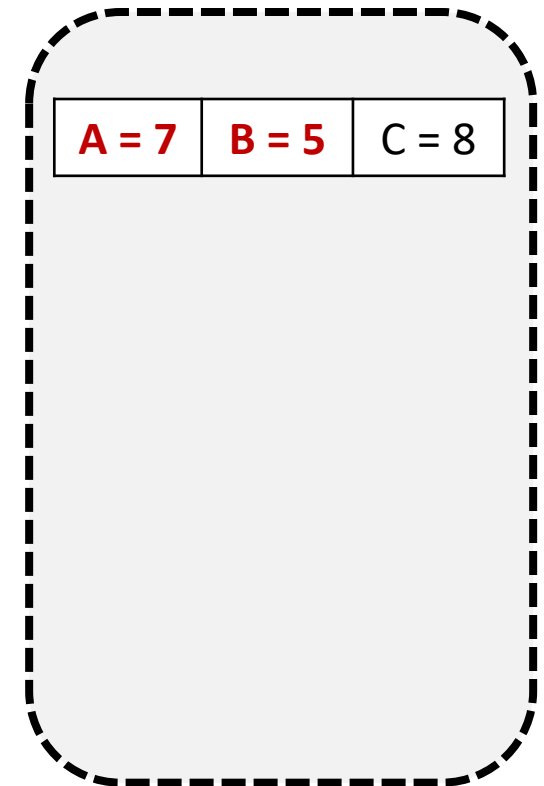


Manažment - buffer pool príklad

Rozvrh transakcií

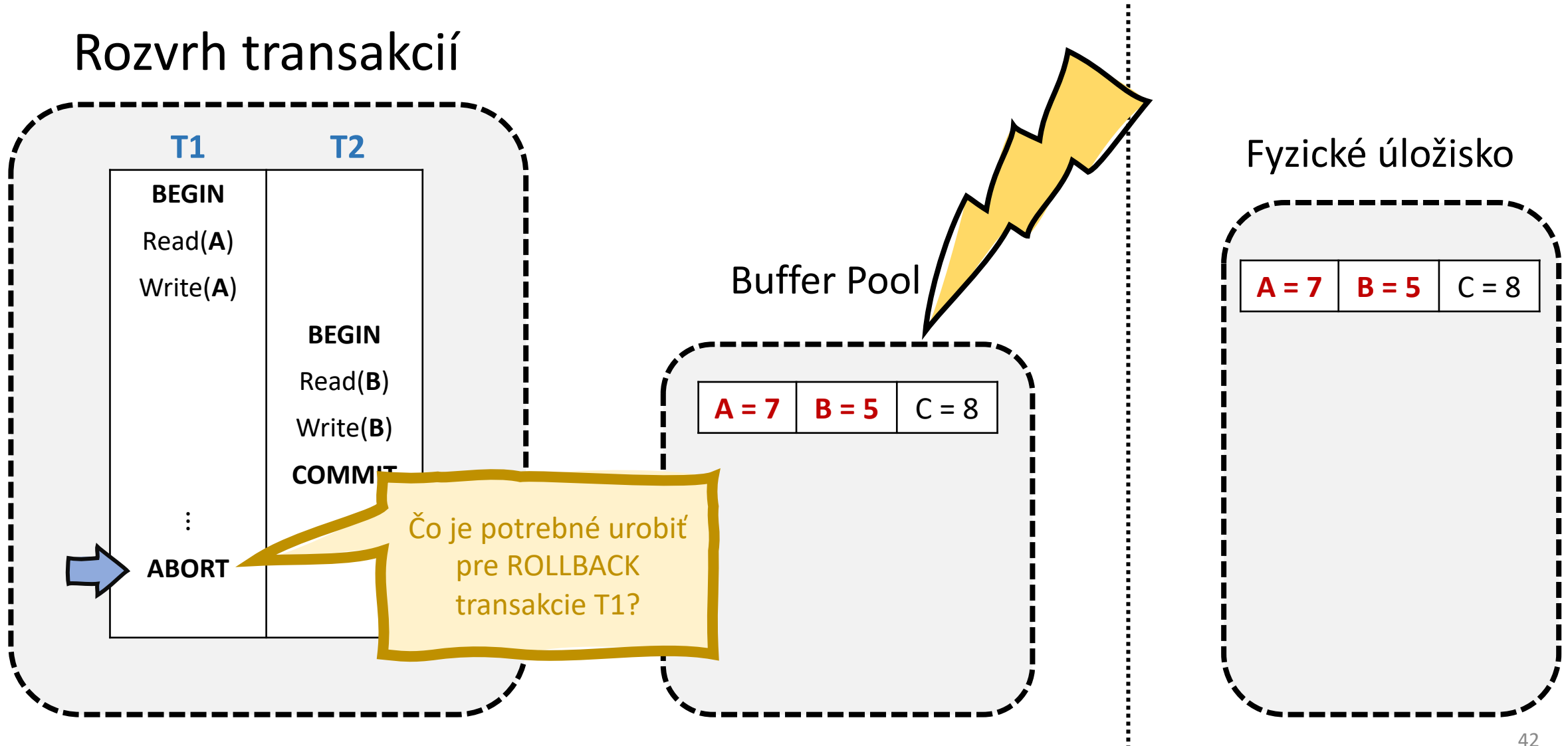


Fyzické úložisko



Manažment - buffer pool príklad

Rozvrh transakcií



Fyzické úložisko

A = 7	B = 5	C = 8
-------	-------	-------

Steal policy

- **STEAL** nastáva vtedy, keď dochádza k prepísaniu hodnoty v nevolatívnej pamäti hodnotou transakcie, ktorá nie je COMMIT.
 - dochádza k prepísaniu hodnoty COMMIT-ovanej transakcie
- **STEAL** – je povolené
- **NO-STEAL** – nie je povolené

Force policy

- **FORCE** všetky zmenené hodnoty sú zapísané do nevolatívnej pamäte pred tým než je samotný COMMIT dovolený (oznámený aplikácií)
 - to znamená, že aktualizované hodnoty sú zapísané do nevolatívnej pamäte a až následne je oznámený commit
- **FORCE** – je vyžadované uskutočňovanie FORCE
- **NO-FORCE** – nie je vyžadované uskutočňovanie FORCE

Porovnanie Steal a Force policy

Výkonosť systému

	No-steal	Steal
No-Force	-	Rýchle
Force	Pomalé	-

Rýchlosť obnovy

	No-steal	Steal
No-Force	-	Pomalé
Force	Rýchle	-

- Skoro každý DBMS používa **NO-FORCE + STEAL**

Operácie UNDO a REDO

- **UNDO** operácia - proces, ktorý odstraňuje efekt na DB nekompletných transakcií alebo zrušených (Abort)
 - vracia hodnoty do pôvodného stavu ako boli pred začatím transakcie
- **REDO** operácia – proces, ktorý opätovne zavádza zmeny vykonané COMMIT transakciami (Committed)
 - kvôli výpadku nie sú uchované v DB ale je ich možné obnoviť
- To ako DBMS podporuje tieto operácie závisí od implementácie spravovania buffer pool-u (napr NO-STEAL + FORCE)

Porovnanie Steal a Force policy

Výkonosť systému

	No-steal	Steal
No-Force	-	Rýchle
Force	Pomalé	-

Rýchlosť obnovy

	No-steal	Steal
No-Force	-	Pomalé
Force	Rýchle	-

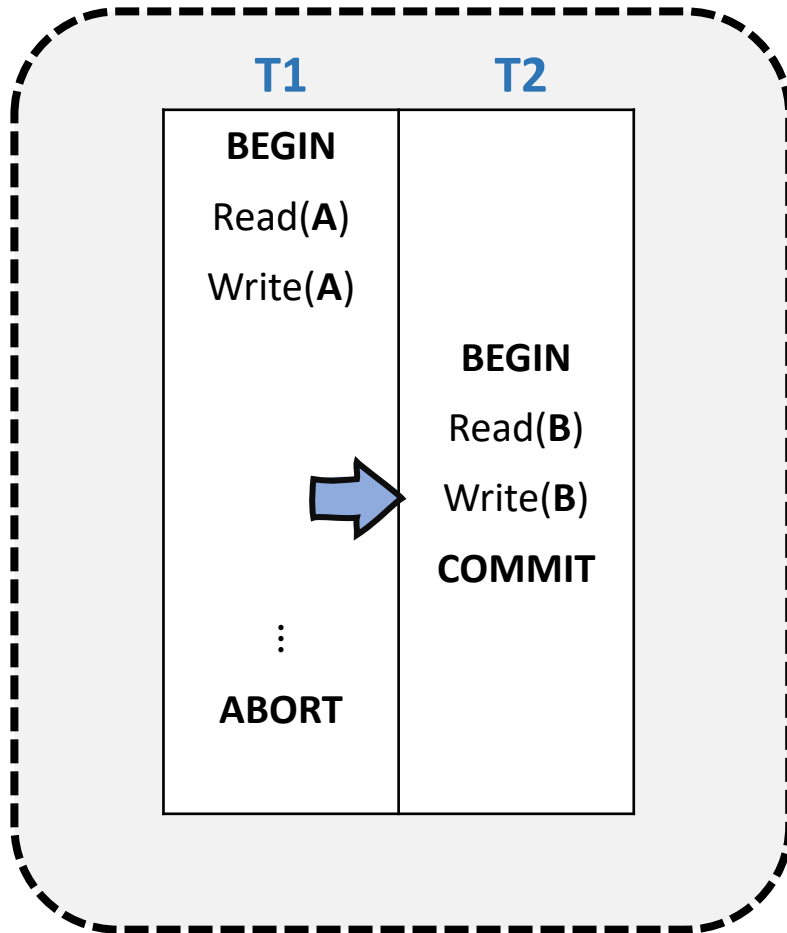
UNDO + REDO

NO UNDO + NO REDO

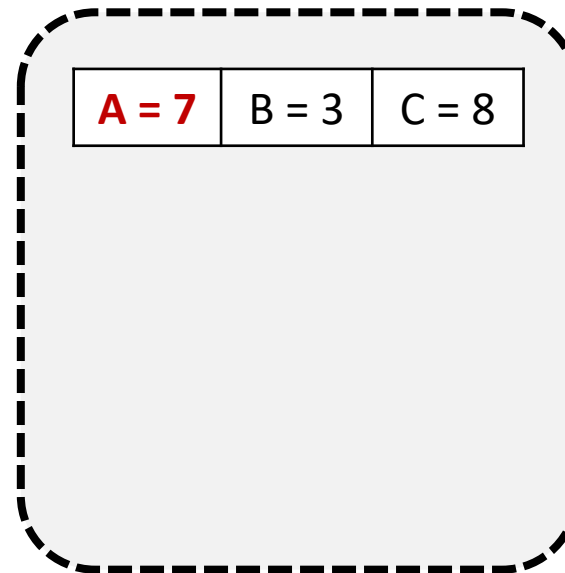
- Skoro každý DBMS používa **NO-FORCE + STEAL**

No steal + Force

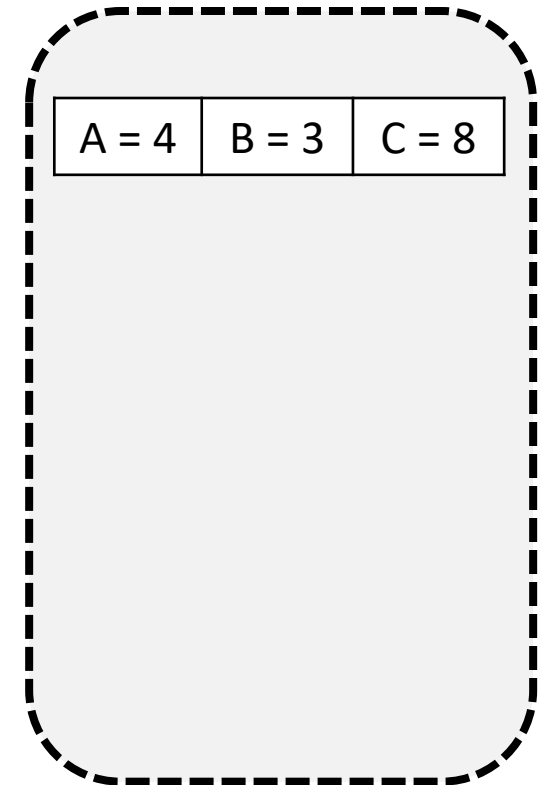
Rozvrh transakcií



Buffer Pool

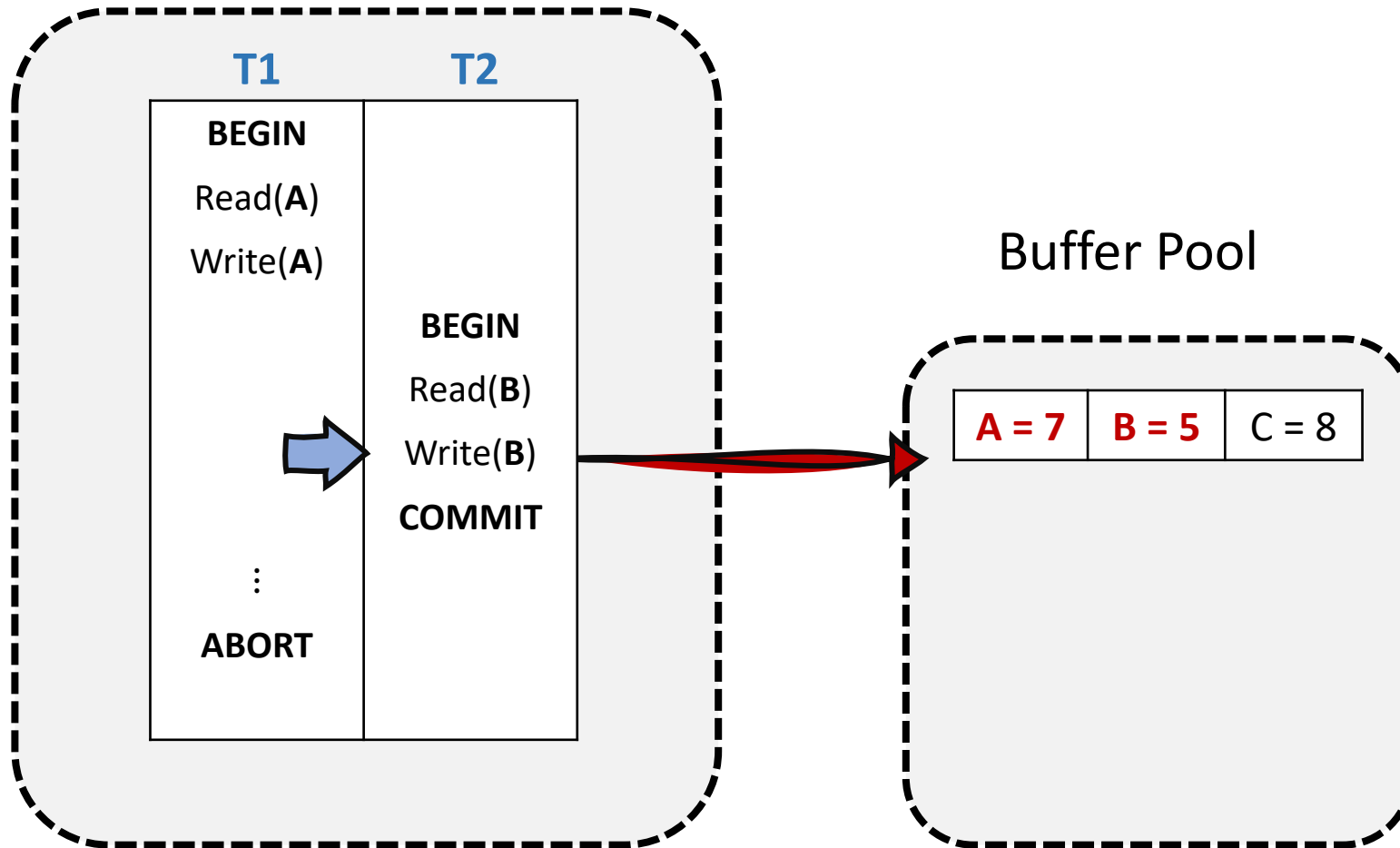


Fyzické úložisko

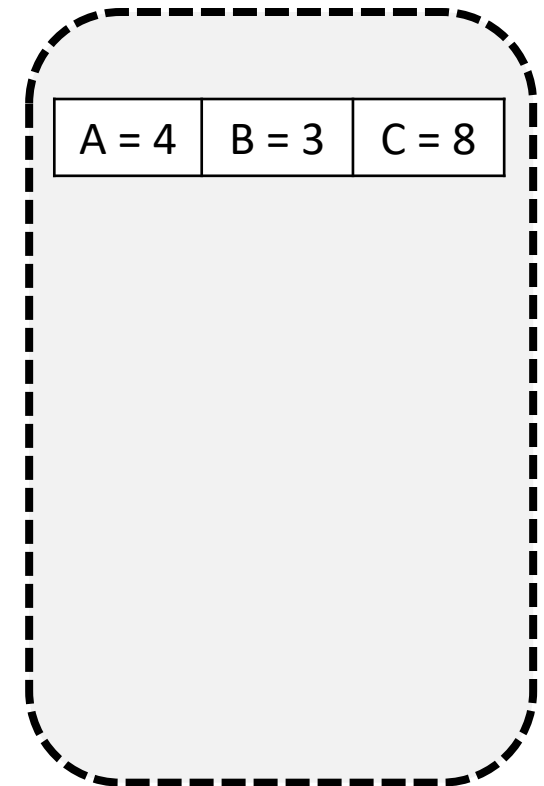


No steal + Force

Rozvrh transakcií

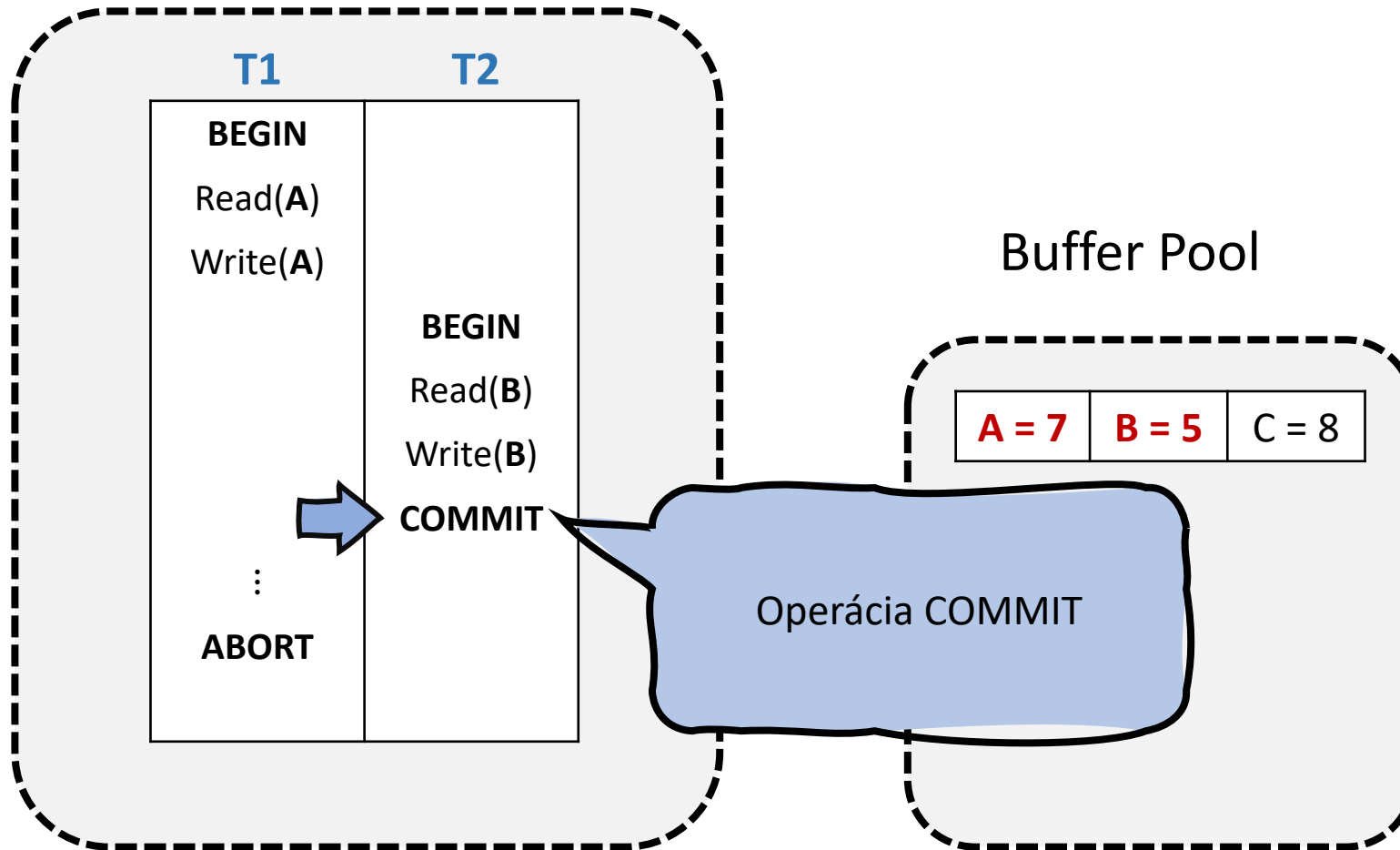


Fyzické úložisko

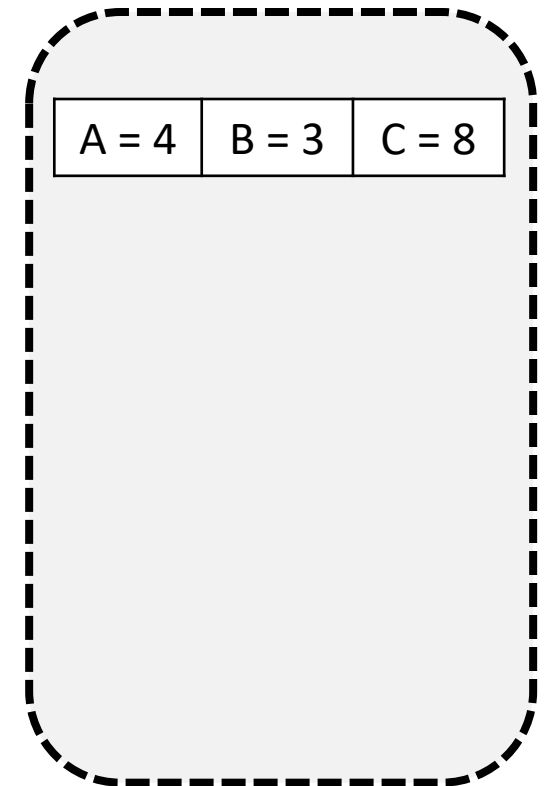


No steal + Force

Rozvrh transakcií

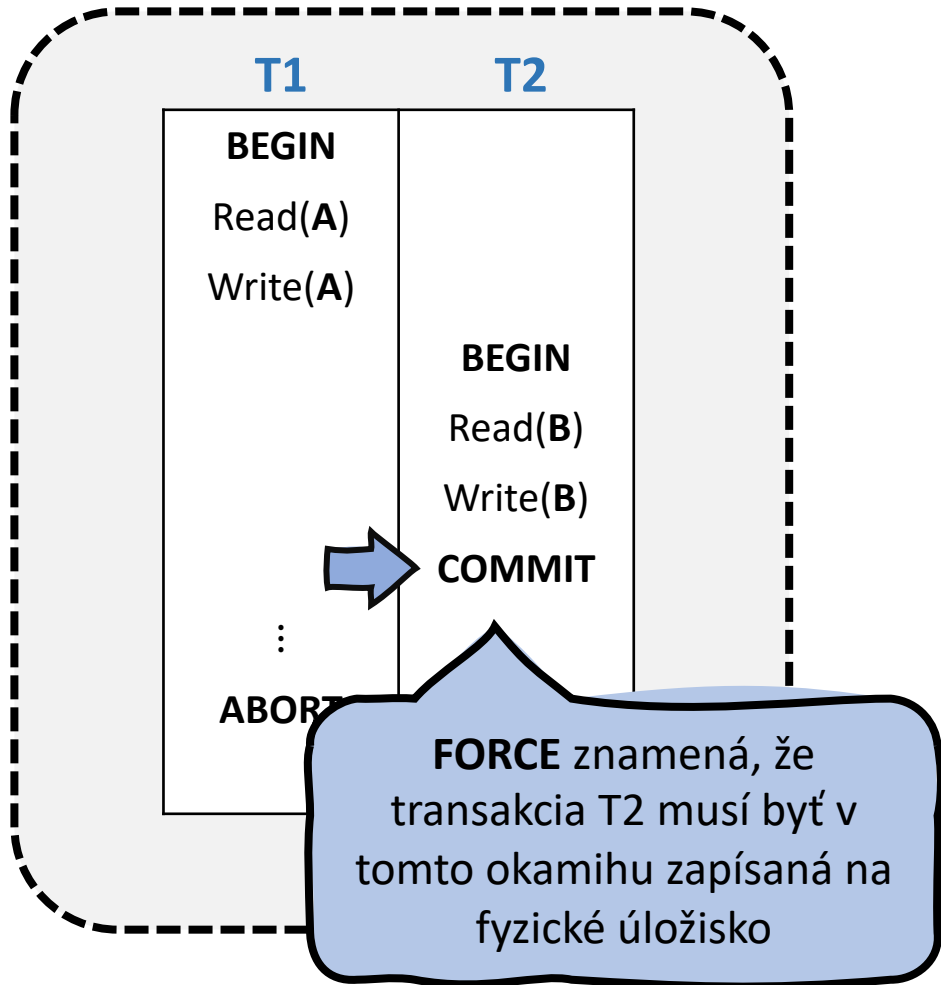


Fyzické úložisko

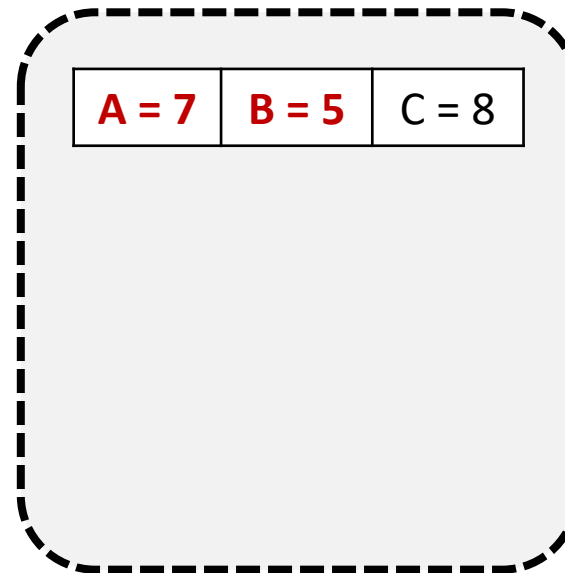


No steal + Force

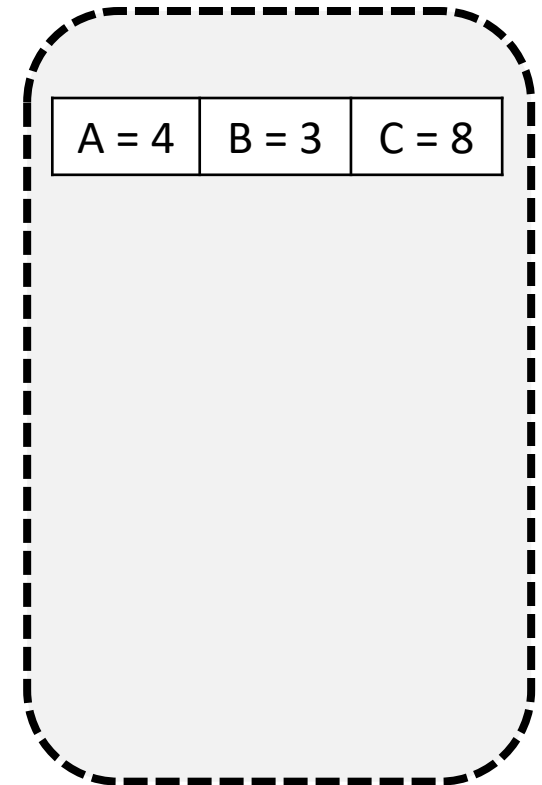
Rozvrh transakcií



Buffer Pool

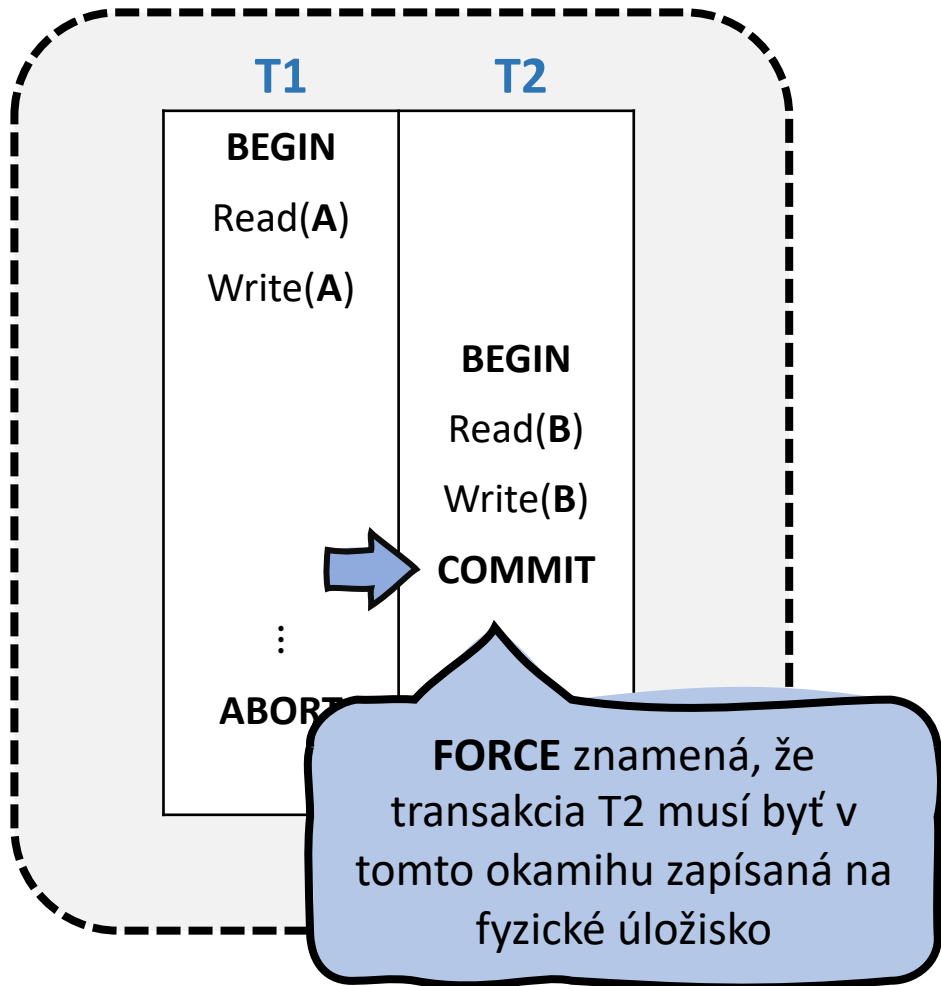


Fyzické úložisko



No steal + Force

Rozvrh transakcií



NO STEAL nám neumožňuje zapísanie hodnoty **A**, ktorá ešte nejde byť COMMIT

Buffer Pool

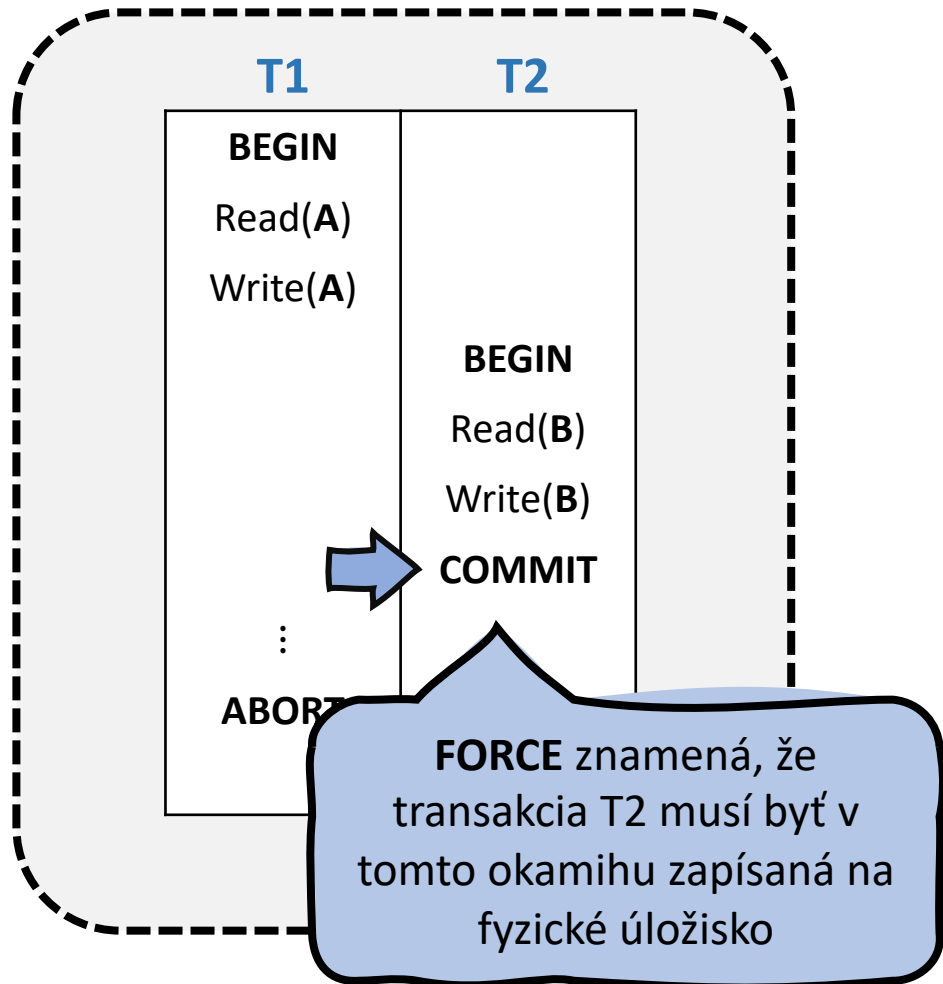
A = 7 **B = 5** C = 8

Fyzické úložisko

A = 4 B = 3 C = 8

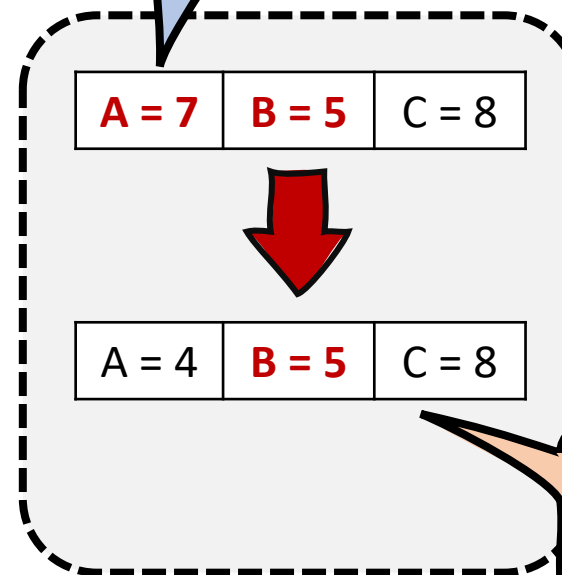
No steal + Force

Rozvrh transakcií



NO STEAL nám neumožňuje zapísanie hodnoty A, ktorá ešte nejde byť COMMIT

Buffer Pool



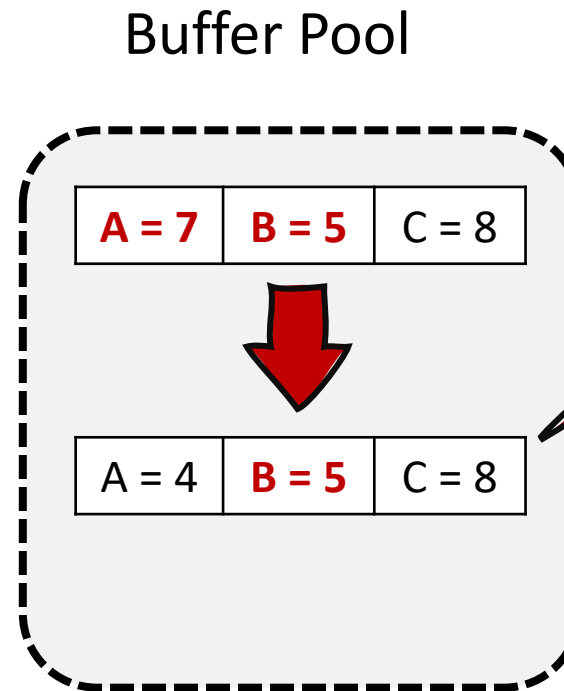
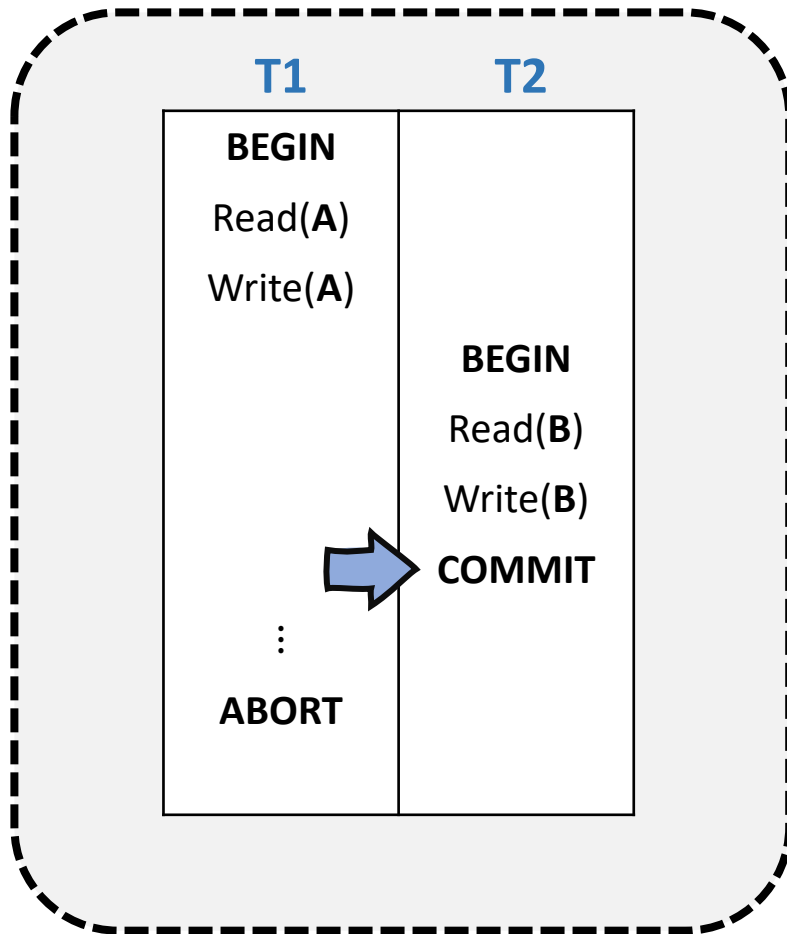
Je potrebné vytvoriť kópiu stránky

Fyzické úložisko

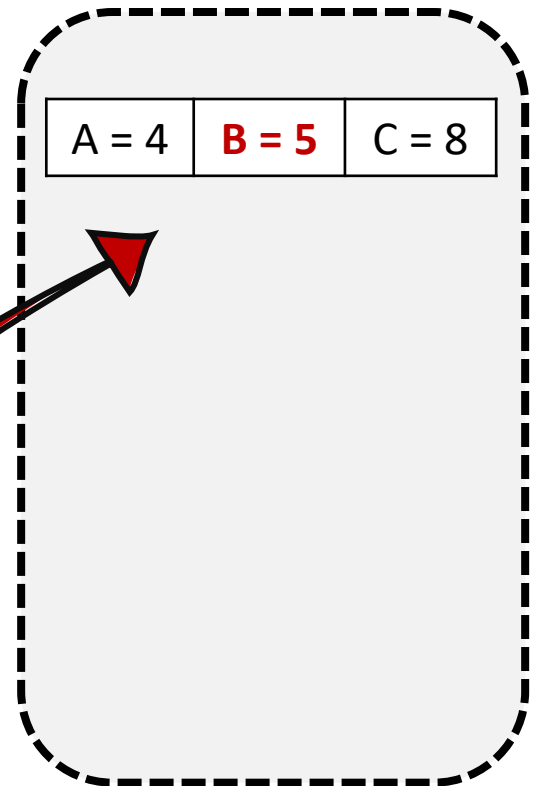
A = 4	B = 3	C = 8
-------	-------	-------

No steal + Force

Rozvrh transakcií

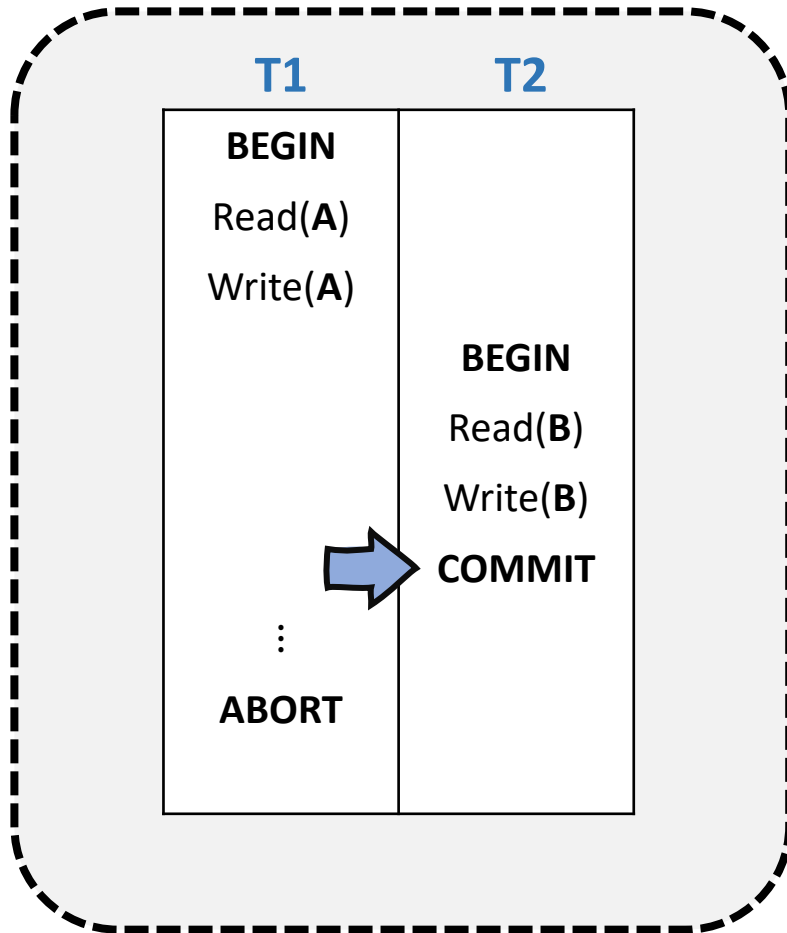


Fyzické úložisko

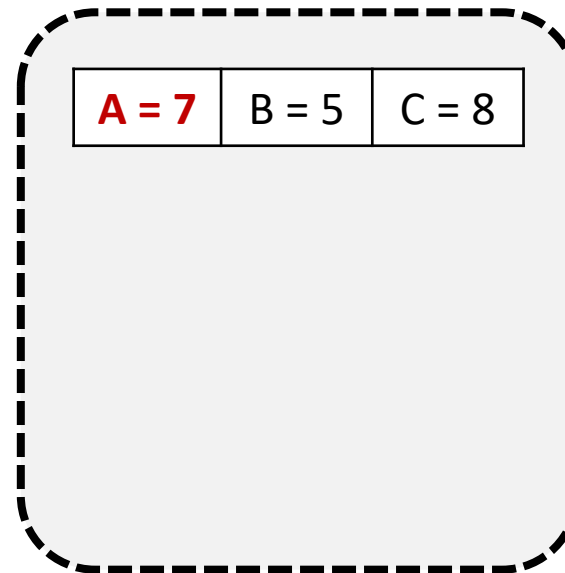


No steal + Force

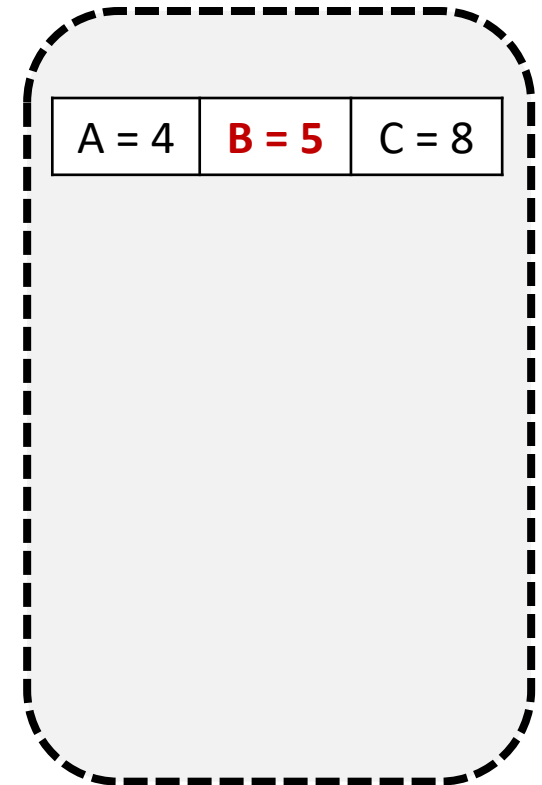
Rozvrh transakcií



Buffer Pool

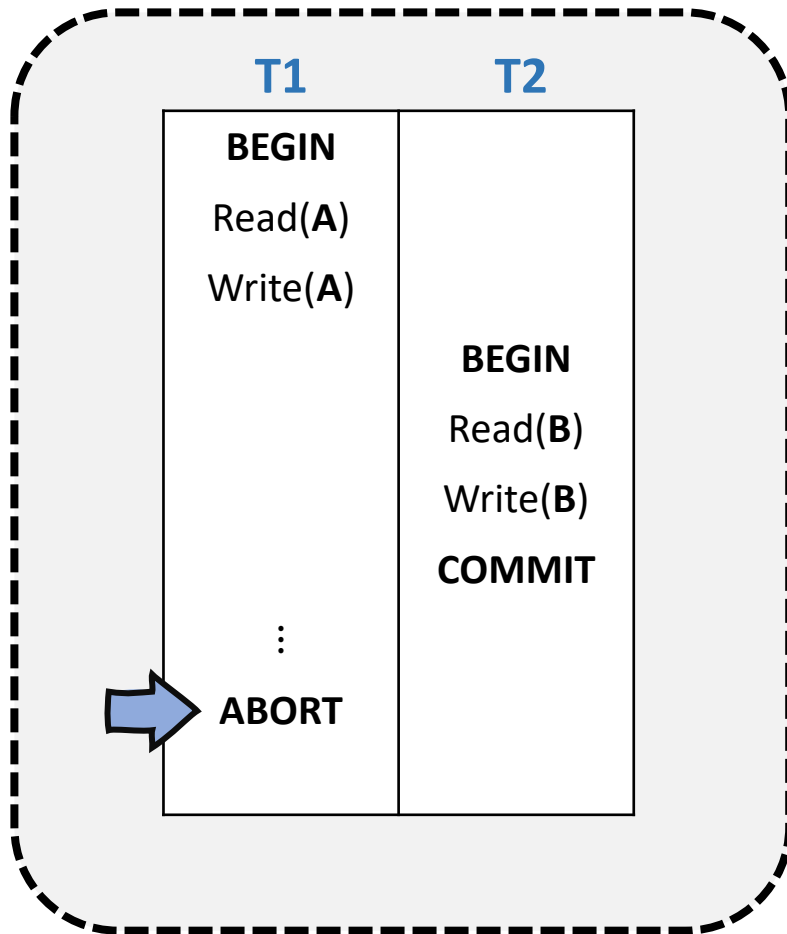


Fyzické úložisko

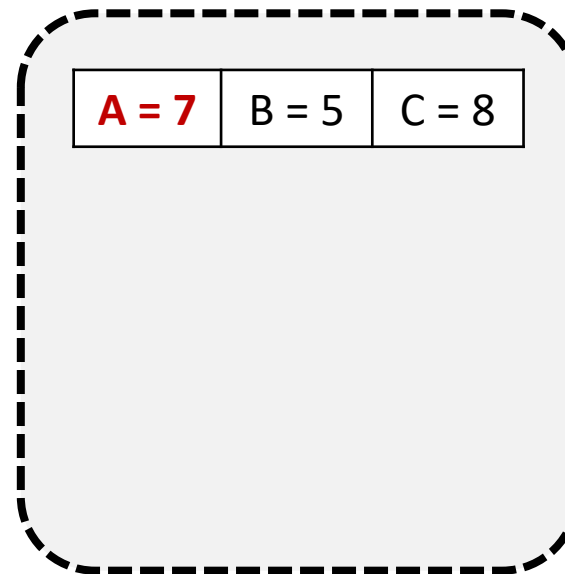


No steal + Force

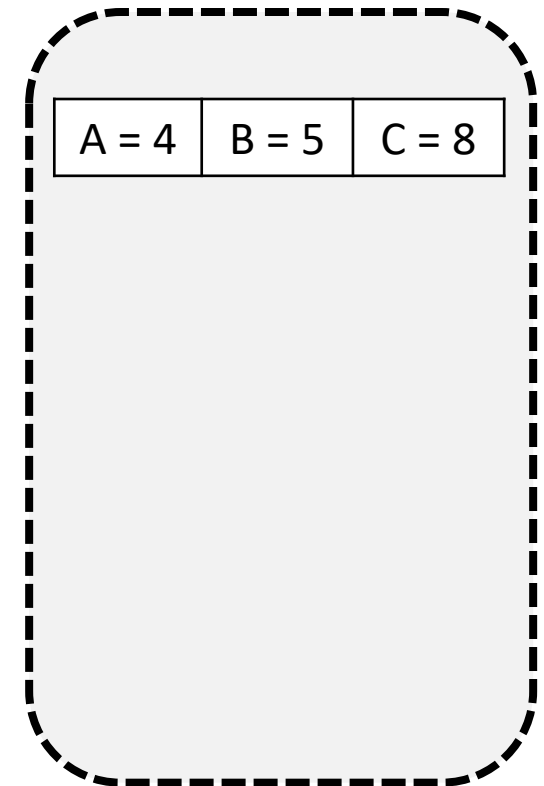
Rozvrh transakcií



Buffer Pool

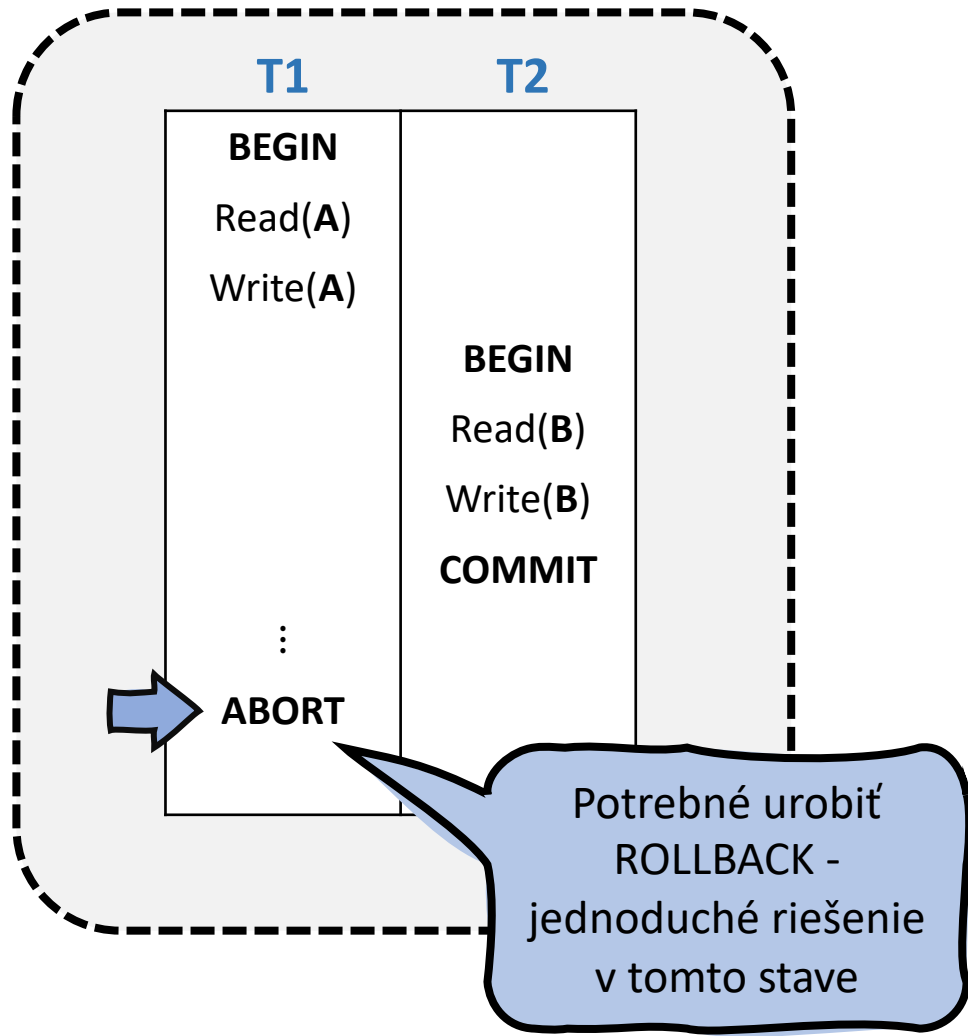


Fyzické úložisko

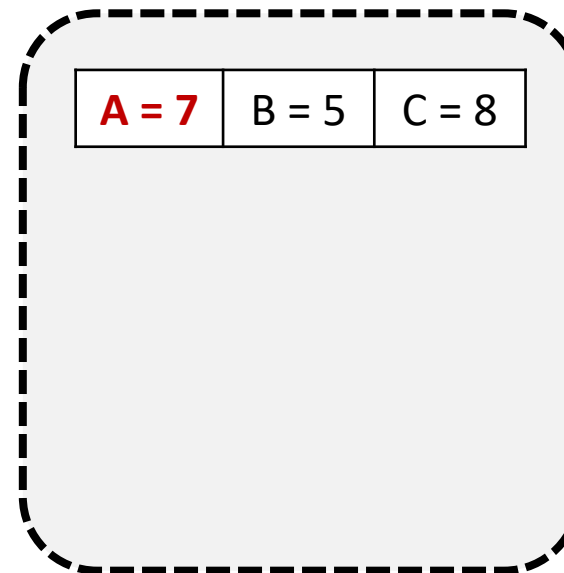


No steal + Force

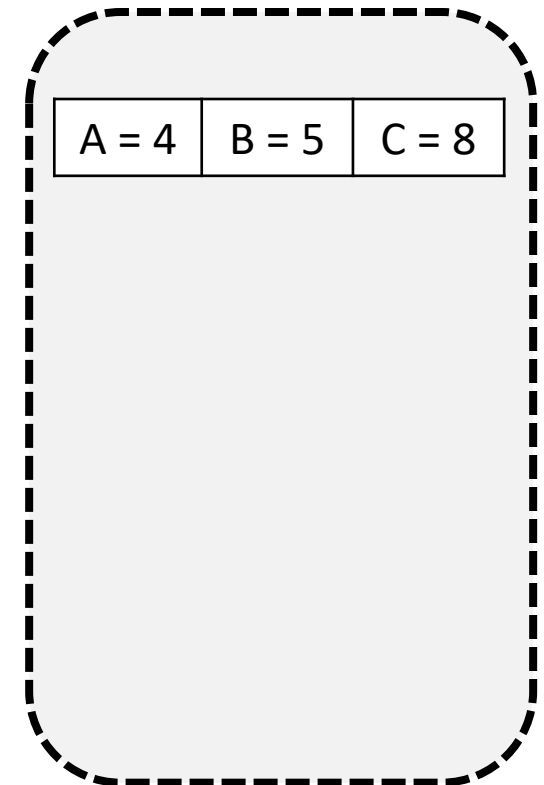
Rozvrh transakcií



Buffer Pool



Fyzické úložisko



No steal + Force

- Ľahká implementácia
 - v prípade zrušenia transakcie (ABORT) nie je potrebné robiť UNDO operácie, pretože zmeny nie sú zapísane na disku
 - v prípade COMMITu nie je potrebné robiť REDO operácie, pretože zmeny sú zapísané na disku
- Príkladom No steal + force policy je **Shadow Paging**

Atomicita a durability

- **Atomicita**

- zaručuje nám to, že transakcia je buď vykonaná úplne alebo vôbec

- **Durability**

- zaručuje nám, že dáta prežijú aj prípadný výpadok

- Vieme ich dosiahnuť pomocou dvoch prístupov

- **Logovanie**

- log file

- **Shadow-paging**

Shadow-Paging

- Obsahuje dve kópie stromu Master a Shadow
 - Koreň (root) ukazuje na Master
 - Aktualizácie sú uskutočnené v rámci Shadow
- Pre uplatnenie zmien je potrebné aktualizovať ROOT, ktorý bude ukazovať na shadow zapísane shadow stránky --> stanu sa master

Logovanie

- Logovací súbor (log file) predstavuje sekvenčný súbor, ktorý zaznamenáva zmeny uskutočnené jednotlivými transakciami
 - predpokladá sa, že log file je uložený na stabilnom úložisku (nevolatívna pamäť)
 - garantovanie atomicity diskových prístupov je dôležitá pre log file
 - musí obsahovať všetky potrebné informácie pre uskutočnenie prípadných UNDO a REDO operácií, ktoré súvisia s obnovou DB
- Log file je uložený separátne od súborov s dátami
- Zmeny v log file sa zapíšu okamžite na disk
- Filozofia **WRITE-AHEAD**
 - pred tým ako DBMS zapíše zmeny v DB na disk **musia** byť tieto zmeny zapísane do log file na disku (nestačí aby boli zapísane len v pamäti)
- Buffer pool policy: STEAL + NO-FORCE

Write-Ahead Logging (WAL) protokol

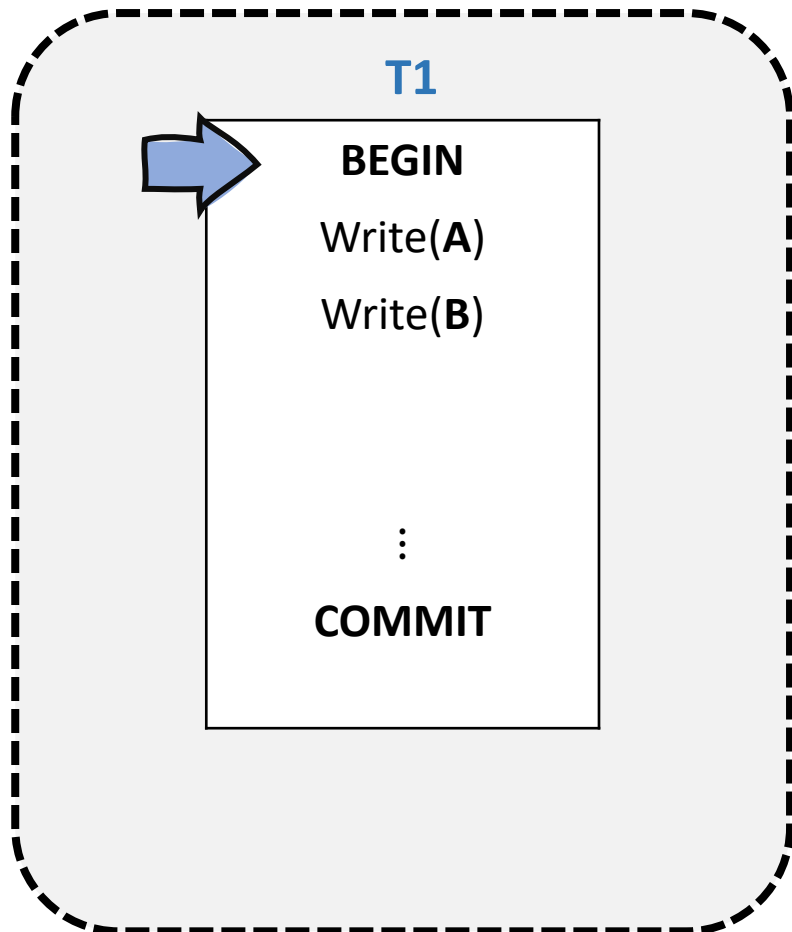
- Jednotlivé logy transakcií sú ukladané DBMS do volatívnej pamäte
- Všetky log záznamy týkajúce sa aktualizovanej stránky sú zapísane do nevolatívnej pamäte **pred tým** ako je samotná stránka zapísaná do nevolatívnej pamäte
- Transakcia nie je považovaná za COMMIT, pokiaľ jej všetky log záznamy nie sú zapísané do nevolatívnej pamäte.
 - informácia pre aplikáciu, že transakcia je COMMIT nastáva až po zapísaní všetkých log záznamov – je potrebné garantovať, že dané dáta sú skutočne aktualizované

Write-Ahead Logging - file

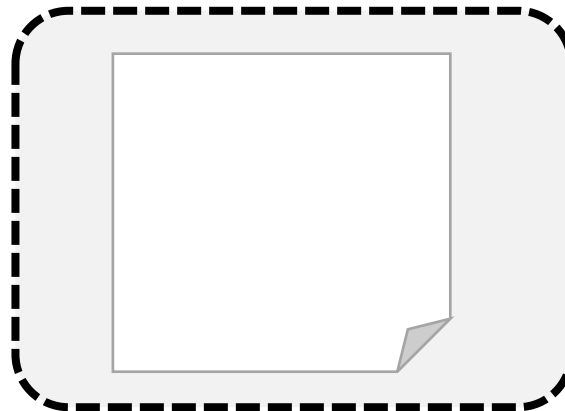
- Záznam log file musí obsahovať
 - identifikáciu transakcie
 - Id objektu, ktorého sa daná zmena týka
 - Hodnotu pred zmenou
 - Hodnotu po zmene
- V rámci záznamov musí byť označené:
 - kedy daná transakcia začala <BEGIN> alebo <START>
 - COMMIT – transakcie <COMMIT>
- Môže obsahovať ďalšie hodnoty záznamy ako napr. CHECKPOINT, DUMP atď.

WAL príklad

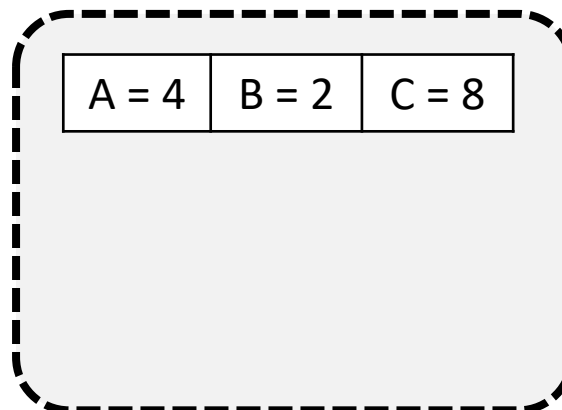
Rozvrh transakcií



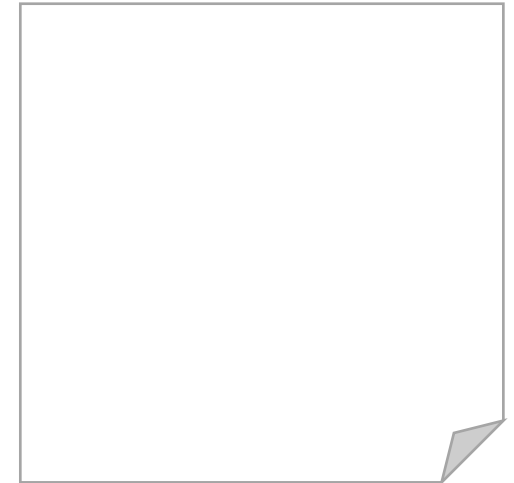
WAL buffer



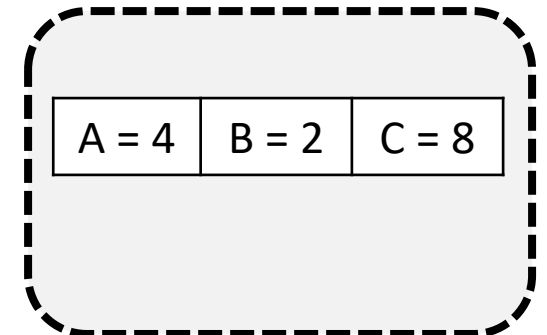
Buffer Pool



Fyzické úložisko

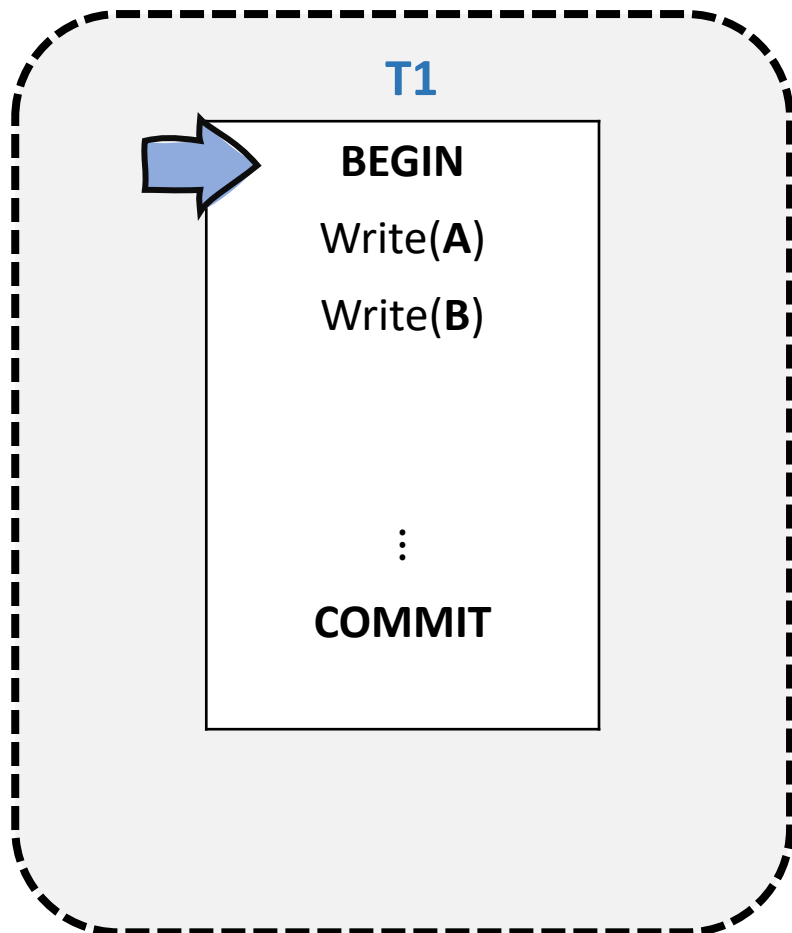


Databáza



WAL príklad

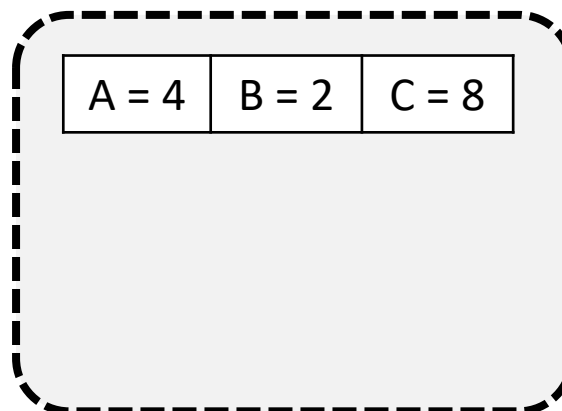
Rozvrh transakcií



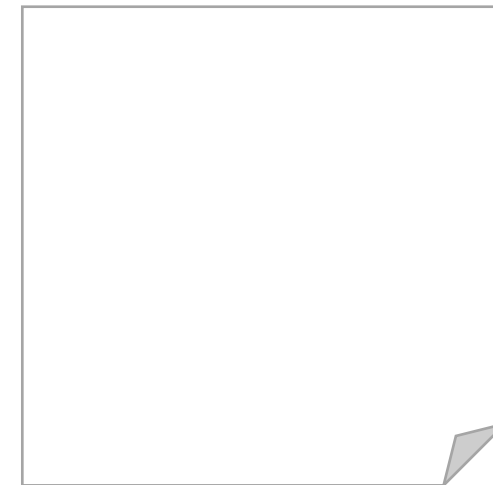
WAL buffer



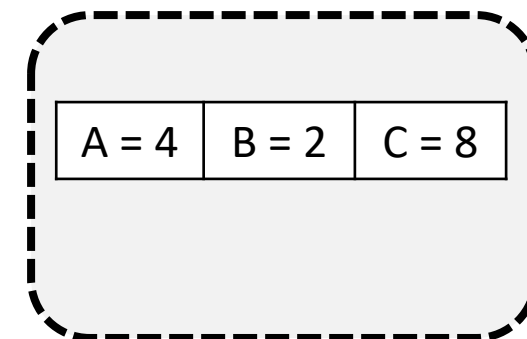
Buffer Pool



Fyzické úložisko

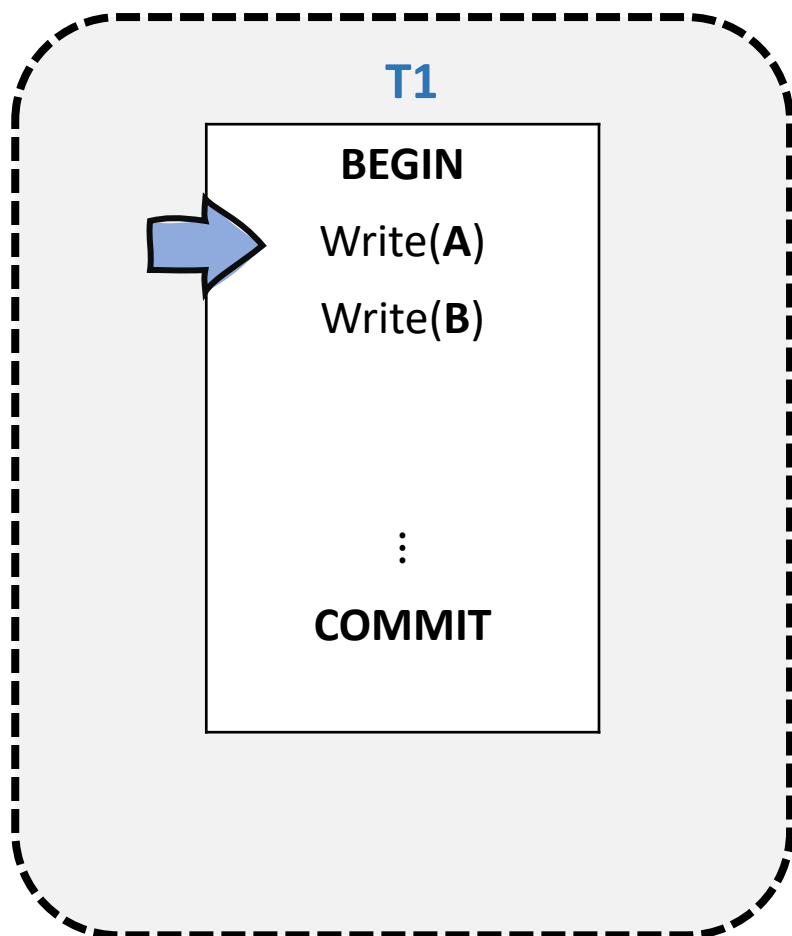


Databáza

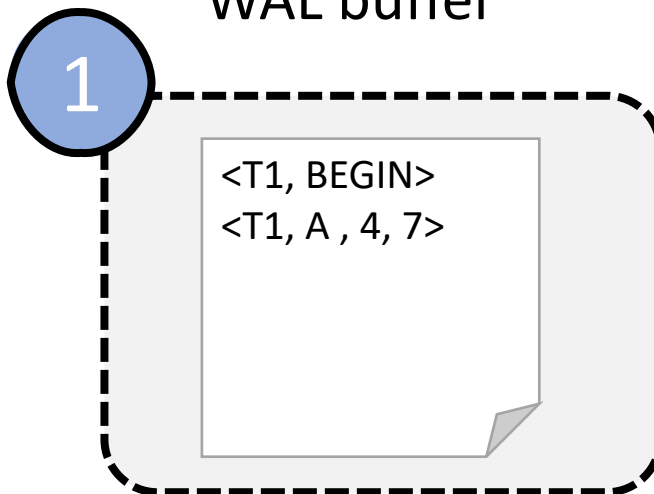


WAL príklad

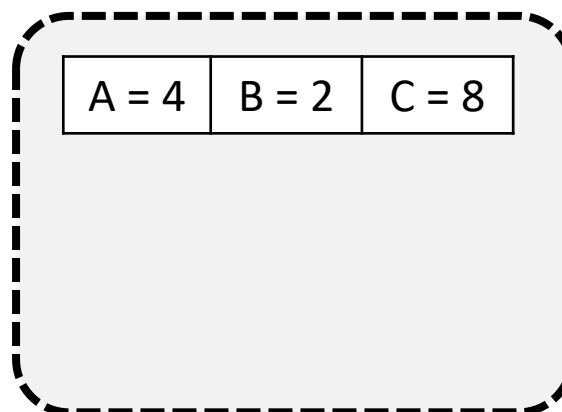
Rozvrh transakcií



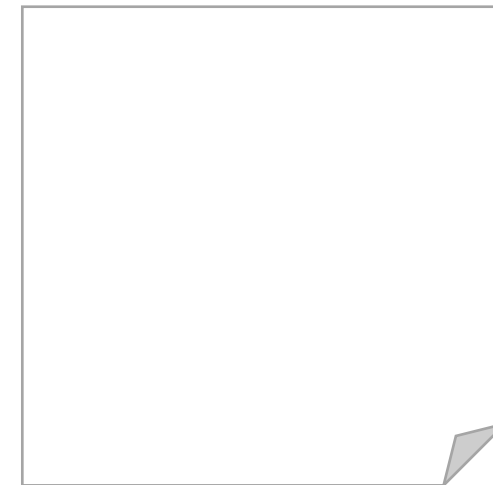
WAL buffer



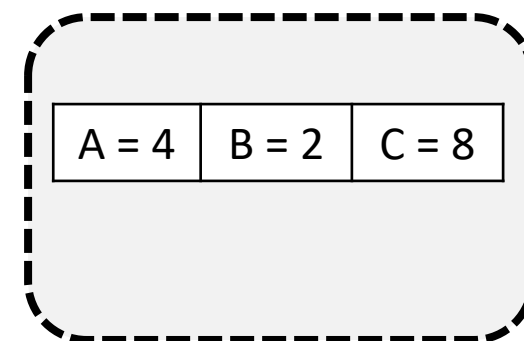
Buffer Pool



Fyzické úložisko

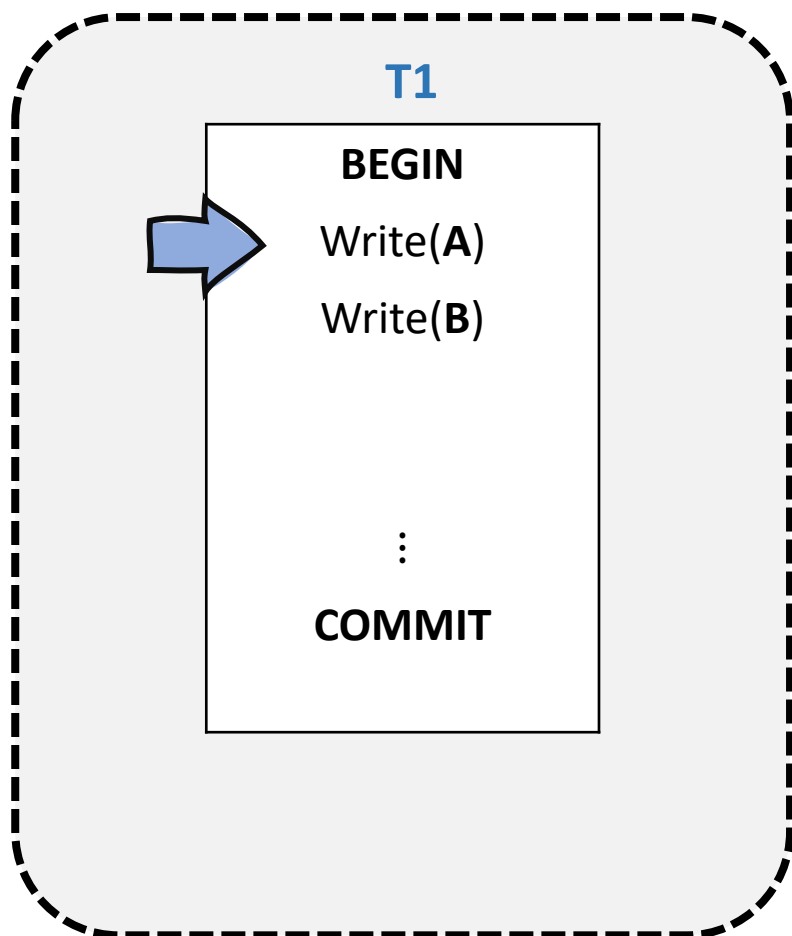


Databáza

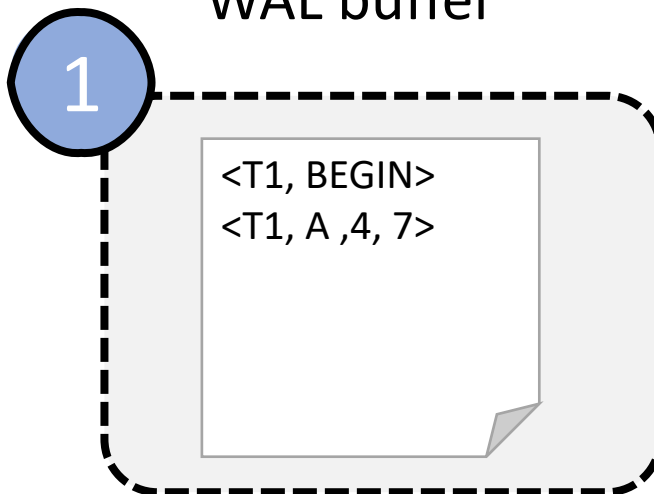


WAL príklad

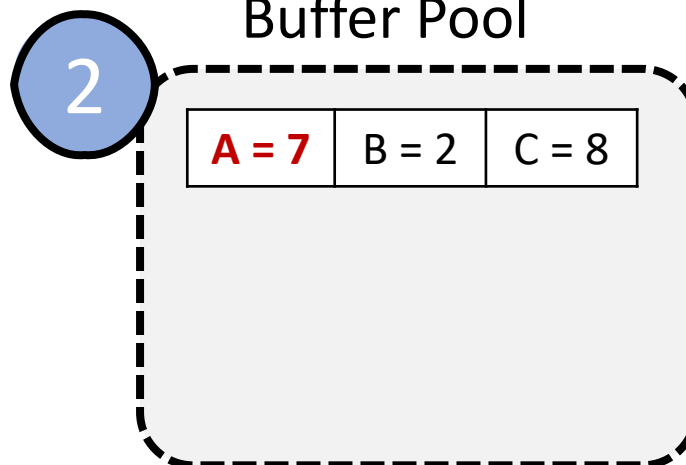
Rozvrh transakcií



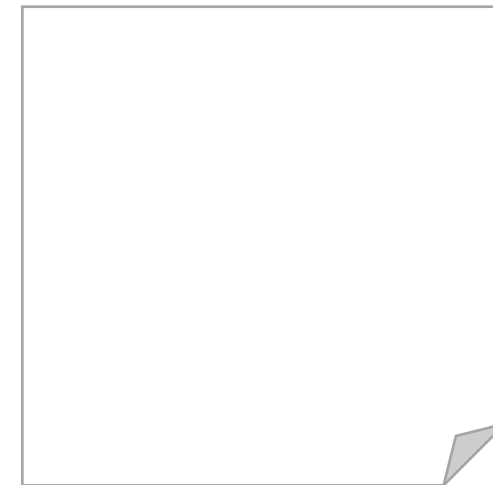
WAL buffer



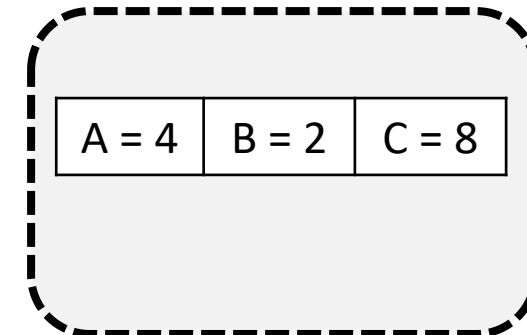
Buffer Pool



Fyzické úložisko

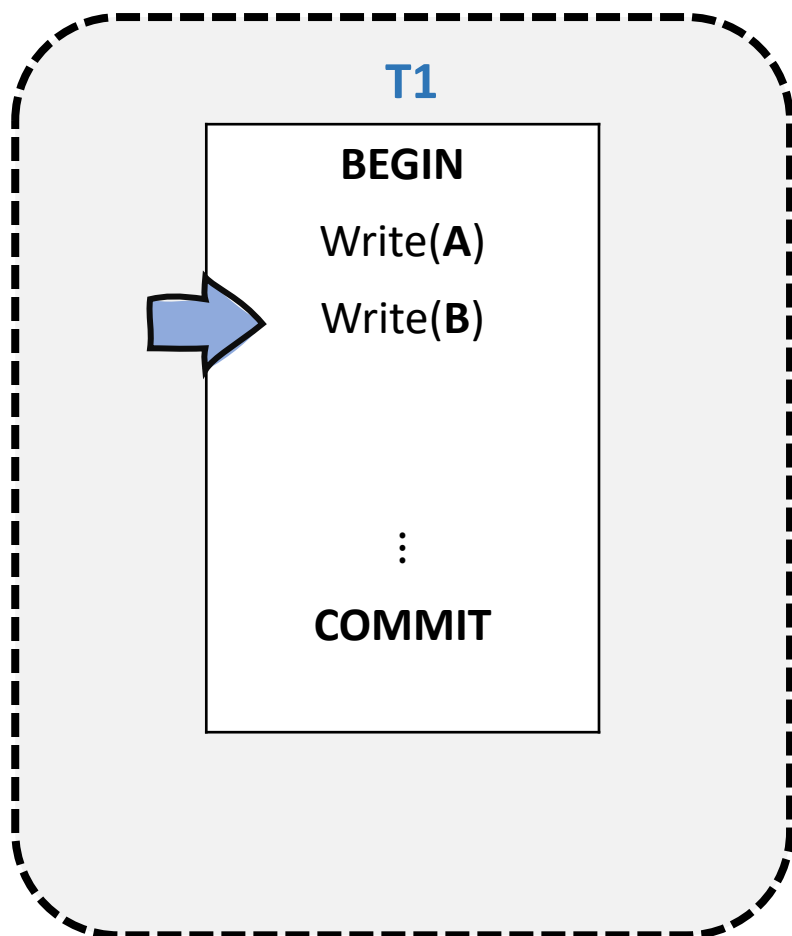


Databáza

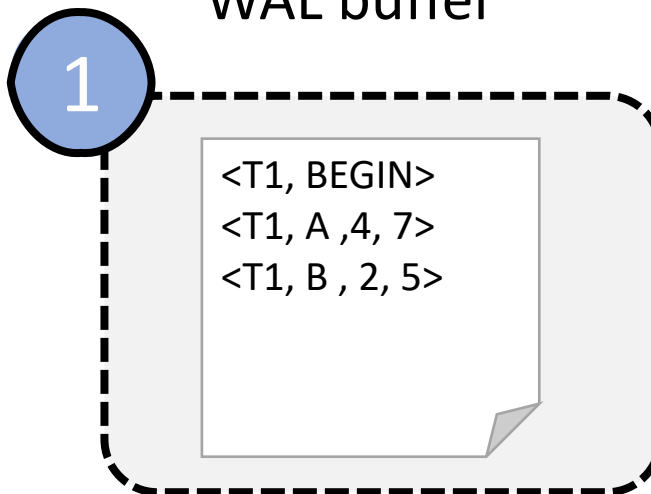


WAL príklad

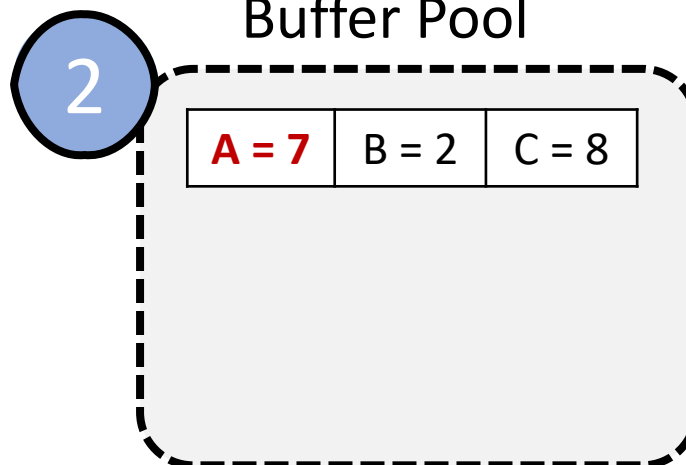
Rozvrh transakcií



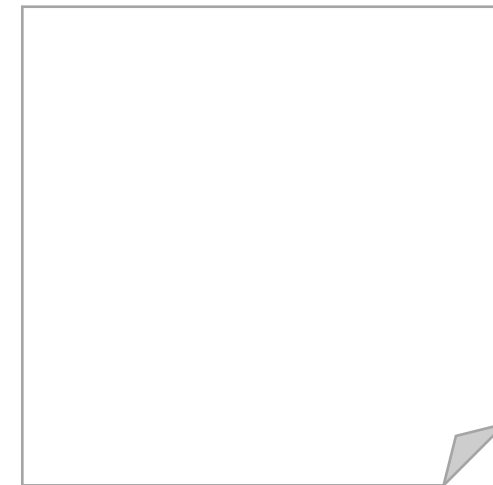
WAL buffer



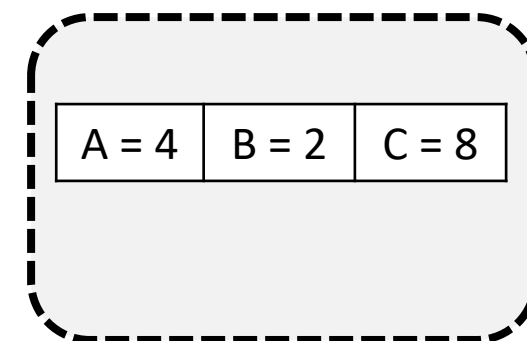
Buffer Pool



Fyzické úložisko

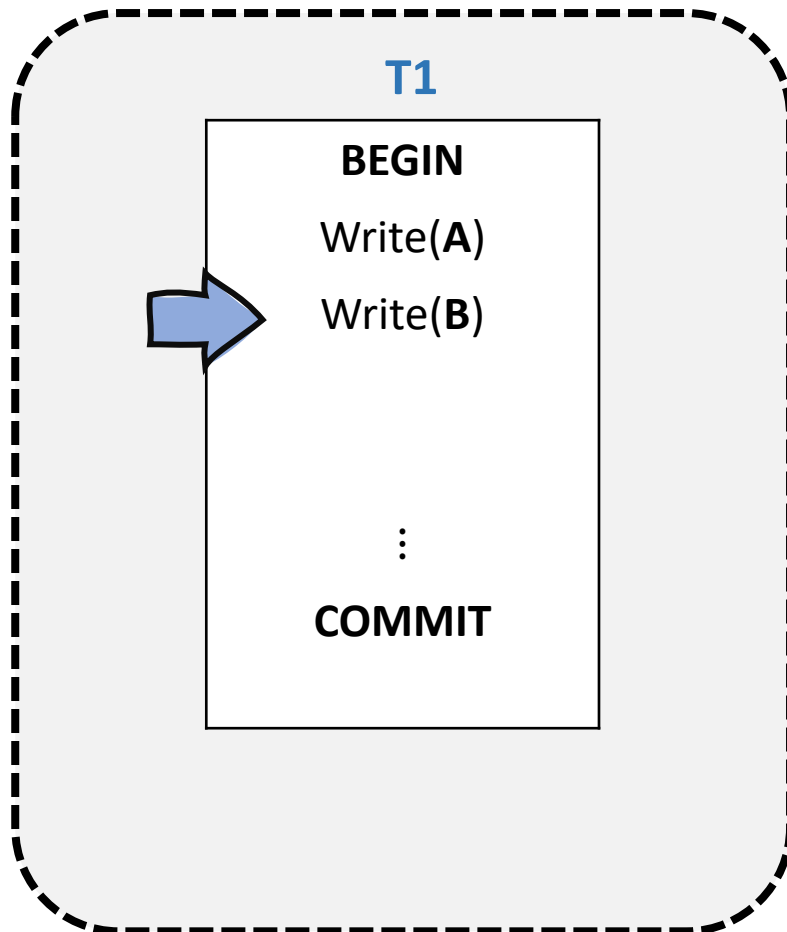


Databáza

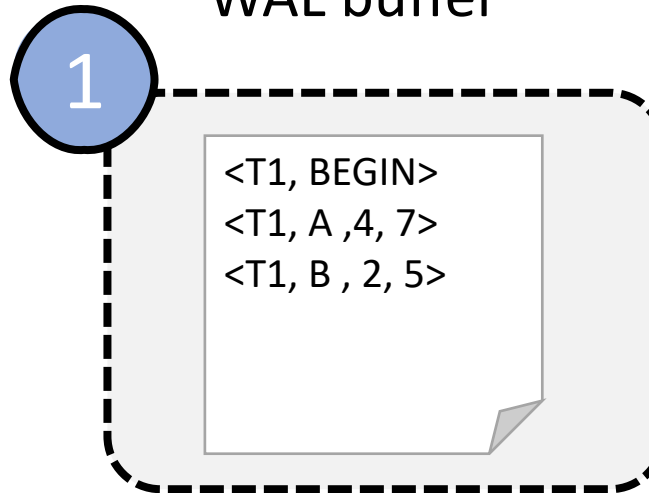


WAL príklad

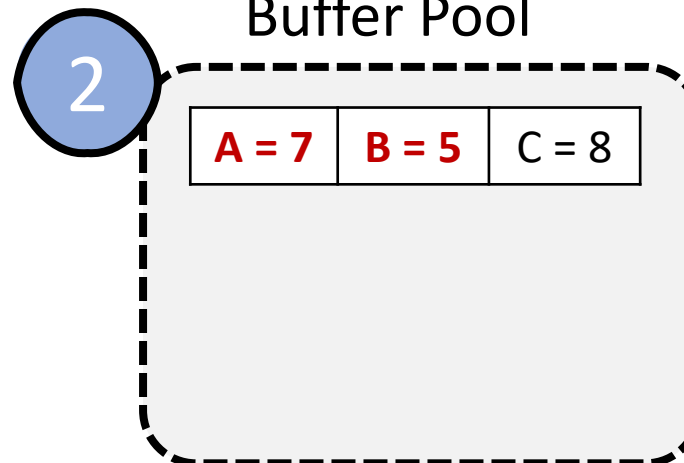
Rozvrh transakcií



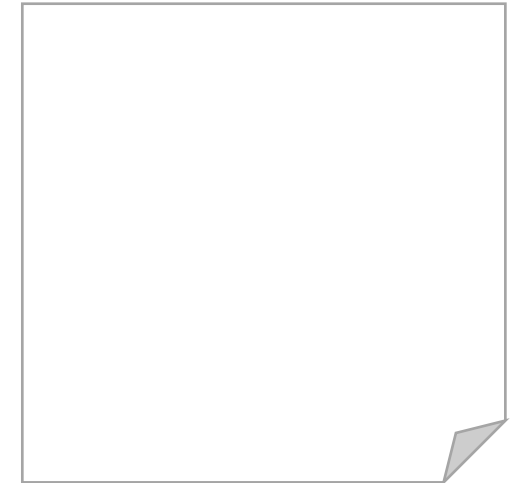
WAL buffer



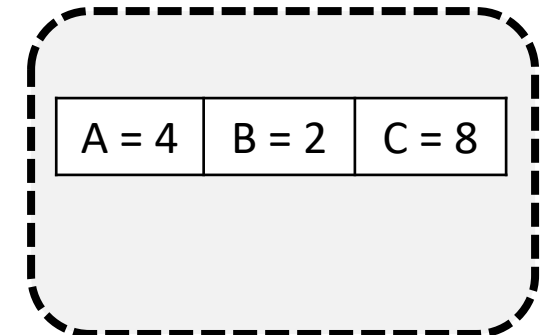
Buffer Pool



Fyzické úložisko

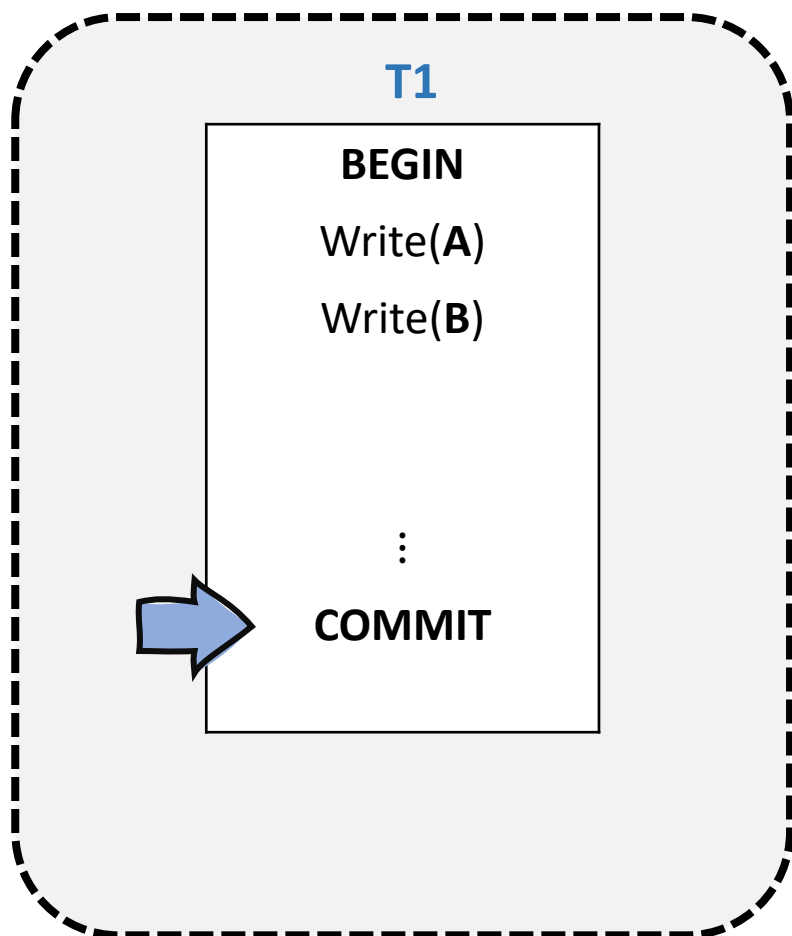


Databáza

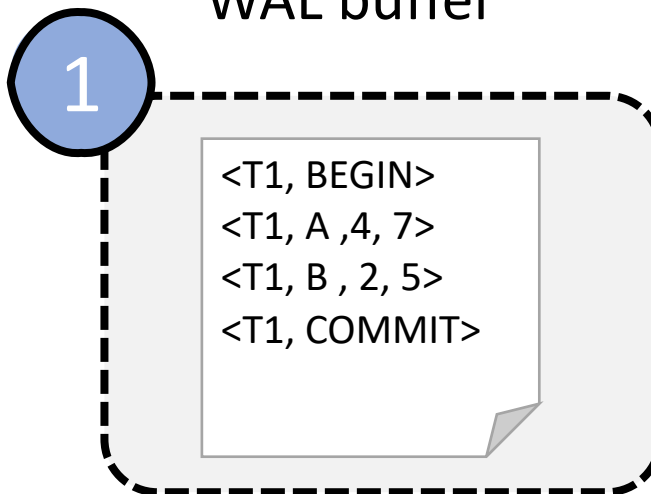


WAL príklad

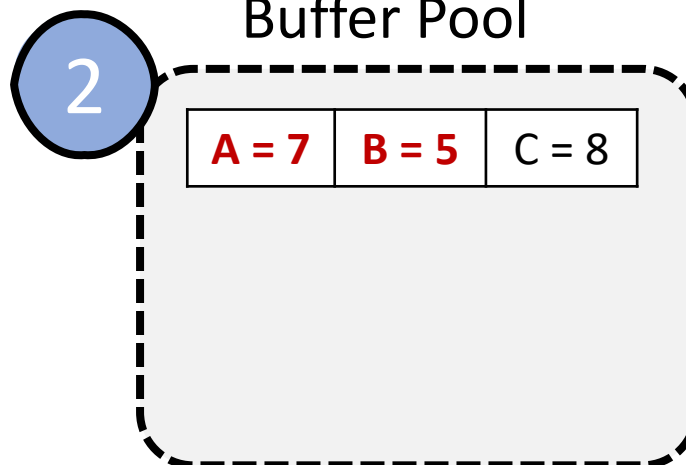
Rozvrh transakcií



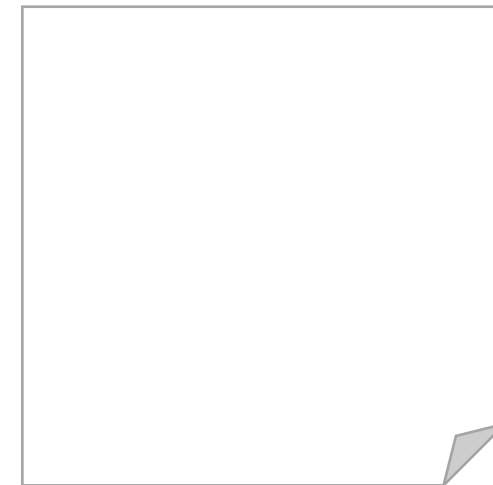
WAL buffer



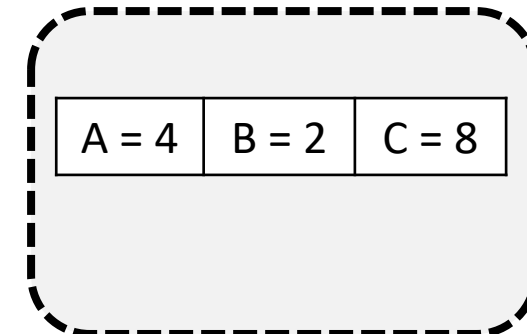
Buffer Pool



Fyzické úložisko

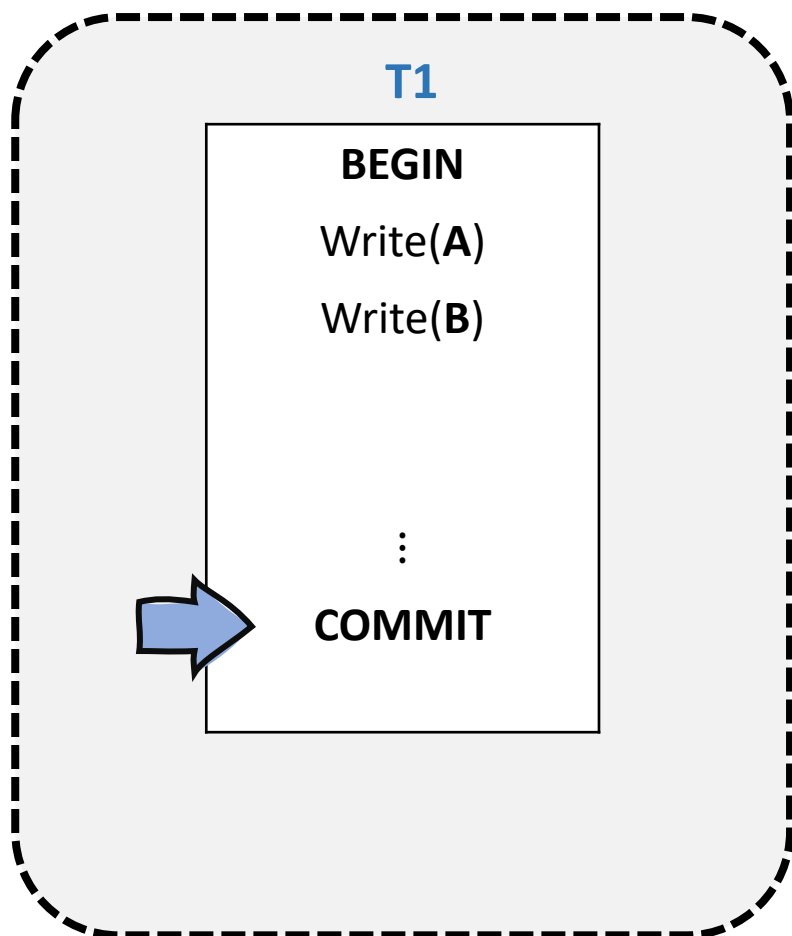


Databáza

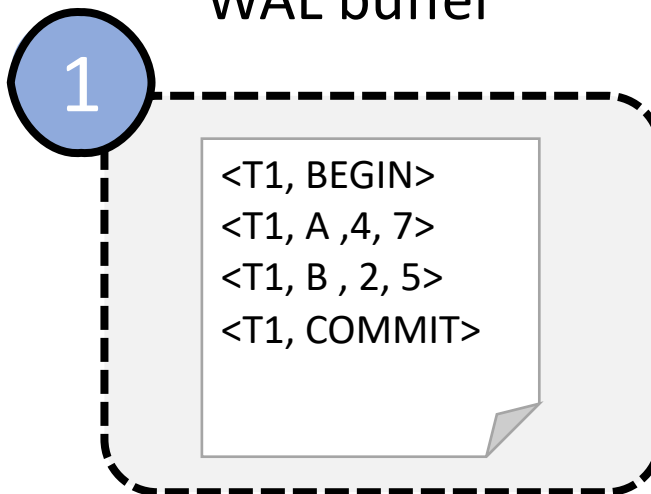


WAL príklad

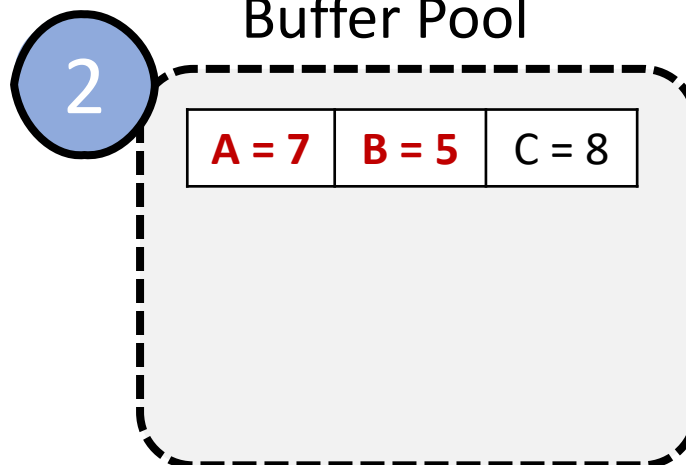
Rozvrh transakcií



WAL buffer



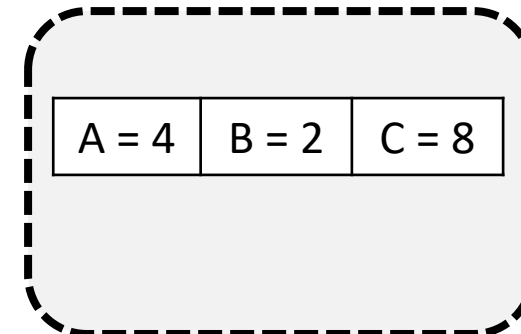
Buffer Pool



Fyzické úložisko

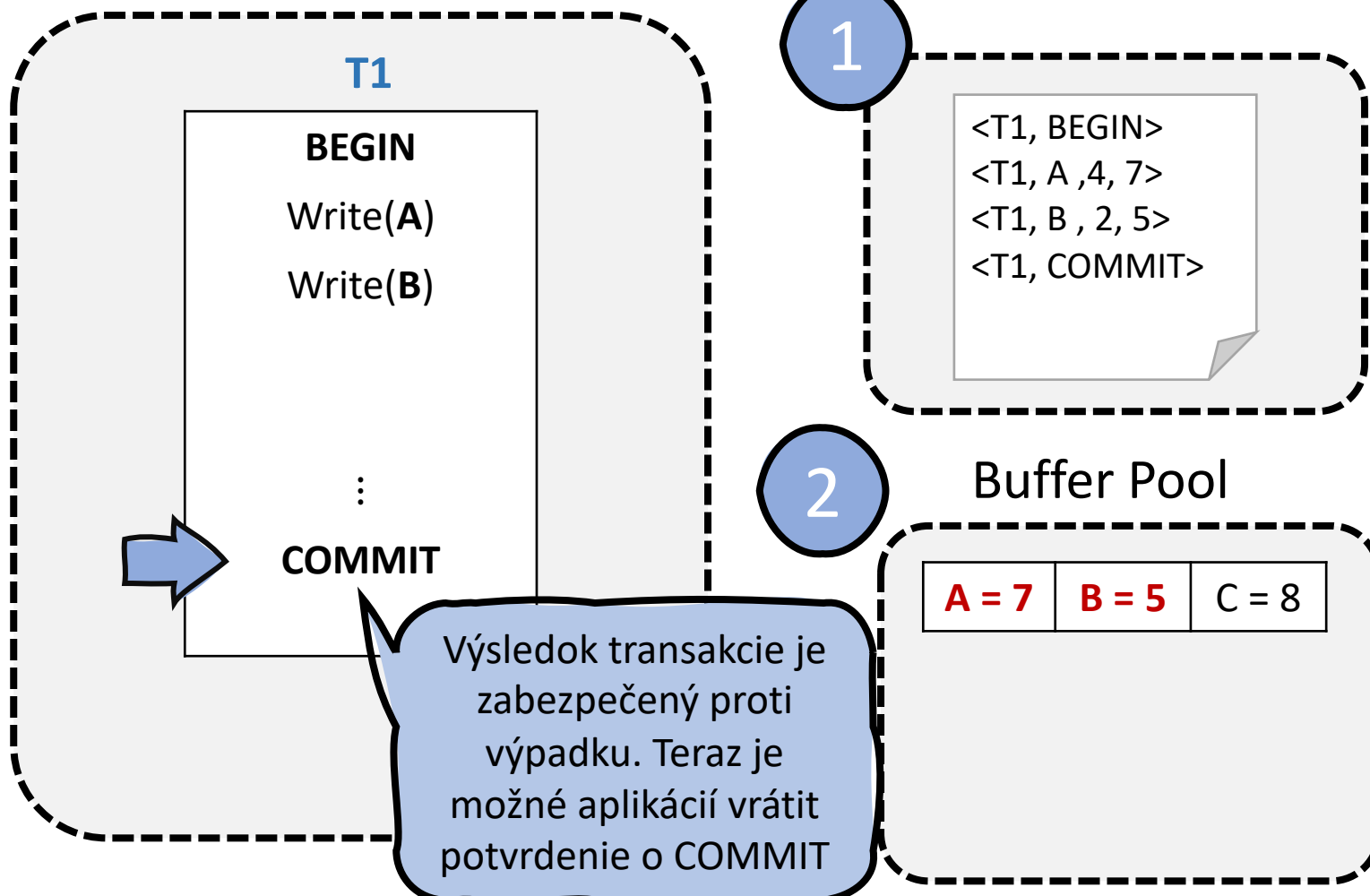
<T1, BEGIN>
<T1, A ,4, 7>
<T1, B , 2, 5>
<T1, COMMIT>

Databáza



WAL príklad

Rozvrh transakcií



Fyzické úložisko

<T1, BEGIN>
<T1, A ,4, 7>
<T1, B , 2, 5>
<T1, COMMIT>

Databáza

A = 4	B = 2	C = 8
-------	-------	-------

WAL implementácia

- Kedy má DBMS zapísať log file na disk ?
- Kedy má DBMS zapísať dirty záznamy na disk ?

WAL implementácia

- Kedy má DBMS zapísať log file na disk ?
 - keď transakcia robí COMMIT.
 - možnosť zápisu viacerých log záznamov na disk naraz – zrýchlenie výkonu systému
- Kedy má DBMS zapísať dirty záznamy na disk ?
 - Pri každej aktualizácii objektu alebo len raz pri COMMIT ?

Checkpoints

- Log file rastie donekonečná
 - nie ideálne riešenie do praxe
 - dlhý recovery time
- Po výpadku je potrebné prejsť celý log file a uskutočniť jednotlivé zmeny, ktoré vykonali jednotlivé transakcie
- Checkpoint
 - pomáhajú redukovať dĺžku log-filov a skracujú čas potrebných na obnovu systému

Checkpoint

- Uvažuje o atomicku procedúru, ktorá obsahuje nasledujúce časti:
 - Zapíše všetky log záznamy na disk (nevolatívna pamäť)
 - Zapíše všetky dáta na disk
 - zapíš **Checkpoint** do log file – môžu sa zapísať aj ďalšie informácie napr. zoznam aktívnych transakcií
 - skráť log-file – garbage collector
 - skratí sa o časť, ktorá sa nachádza pred najstaršou aktívnou transakciou – pred jej BEGIN

Algoritmus obnovy

- Je závislý od
 - použitého concurrency control protokolu
 - použitéj techniky zápisu na disk

Checkpoint - algoritmus obnovy

Všeobecná metóda obnovy sa uskutočňuje v nasledujúcich krokoch:

1. prechádzajú sa záznamy zostupne (od posledného po začiatok) a vytvára sa UNDO a REDO list. Zároveň sa uskutočňuje UNDO operácia
2. Po dosiahnutí záznam CHECKPOINT sa pokračuje zostupne až po prejdenie všetkých aktívnych transakcií v čase CHECKPOINTU
3. Spusti sa REDO fáza vzostupne od posledného záznamu najdenom v kroku 1

Checkpoint príklad

- Transakcie s COMMIT pred CHECKPOINT sú ignorované (T1)
- T2 a T3 neprebehol COMMIT pred CHECKPOINT a je preto treba urobiť
 - REDO operáciu pre T2 (COMMIT po CHECKPOINT)
 - UNDO pre operáciu T3 – nebol COMMIT pred výpadkom

WAL

```
<T1, BEGIN>
<T1, A ,4, 7>
<T1, COMMIT>
<T2, BEGIN>
<T2, B ,3, 8>
<T3, BEGIN>
<CHECKPOINT>
<T2, COMMIT>
<T3, A ,7, 8>
..
Výpadok
```

ARIES

- Algorithms for Recovery and Isolation Exploiting Semantics
- Vynájdenný v roku 1990 spoločnosťou IBM pre ich DB2 DBMS
- V súčasnosti sa využíva stále tento algoritmus aj keď nie úplne presne v rovnakej podobe

ARIES - fázy

- **Analýza** – čítanie WAL od posledného záznamu pre identifikovanie dirty pages a aktívnych transakcií v čase výpadku
- **REDO** – opakovanie akcií pre vrátenie stavu pred výpadkom.
- **UNDO** – vrátenie všetkých hodnôt transakcií, ktorým neprebehol COMMIT do času výpadku

Teoria obnoviteľnosti

- naspäť k rozvrhom z predchádzajúcej prednášky
- Seriovateľnosť rozvhov nie je dostačujúca pre možnosti obnovy

Príklad

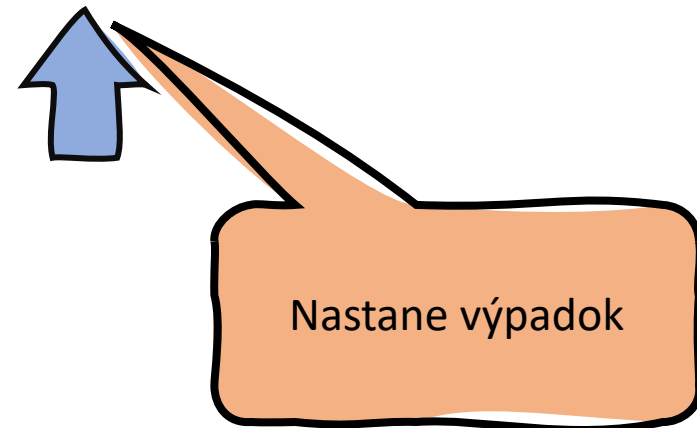
- Rozvrh S1: $r1(A)$, $w1(A)$, $r2(A)$, $w2(C)$, $c2$, $c1$

Teoria obnoviteľnosti

- naspäť k rozvrhom z predchádzajúcej prednášky
- Seriovateľnosť rozvhov nie je dostačujúca pre možnosti obnovy

Príklad

- Rozvrh S1: **r1(A), w1(A), r2(A), w2(C), c2, c1**



Teoria obnoviteľnosti

- Rozvrh S1: **rl(A), wl(A), r2(A), w2(C), c2, cl**
 - COMMIT transakcie T2 (označený ako C2) je navždy zapísaný v DB
 - T2 čítala hodnotu od T1, ktorá však nebola zapísana do DB v dôsledku výpadku
 - Čo ak hodnota C závisela od hodnoty A ?

Dirty read

- čítanie hodnoty zapísanej inou transakciou predtým ako prebehol COMMIT transakcie, ktorá modifikovala danú hodnotu

Obnoviteľný rozvrh

- Rozvrh je obnoviteľný (RC) práve vtedy, keď pre každú COMMIT transakciu T2 čítajúcu nekomitovanú hodnotu od T1 (označuje sa ako DIRTY READ) platí, že T1 commituje tiež a zároveň T1 COMMIT je v danom rozvrhu skôr ako T2 COMMIT.
- Predchádzajúci rozvrh S1 **nie je obnoviteľný**

Príklad

- rozvrh S2: w1(X) r2(X) w2(Y) c1 c2
 - je obnoviteľný
-
- Problém rozvrhu S2 je hroziaci kaskádový abort (cascading abort) v prípade, že T1 sa rozhodne pre ABORT, tak T2 bude musieť skončiť ako ABORT tiež (Ak ma byť tento rozvrh považovaný za obnoviteľný)

Obnoviteľný rozvrh

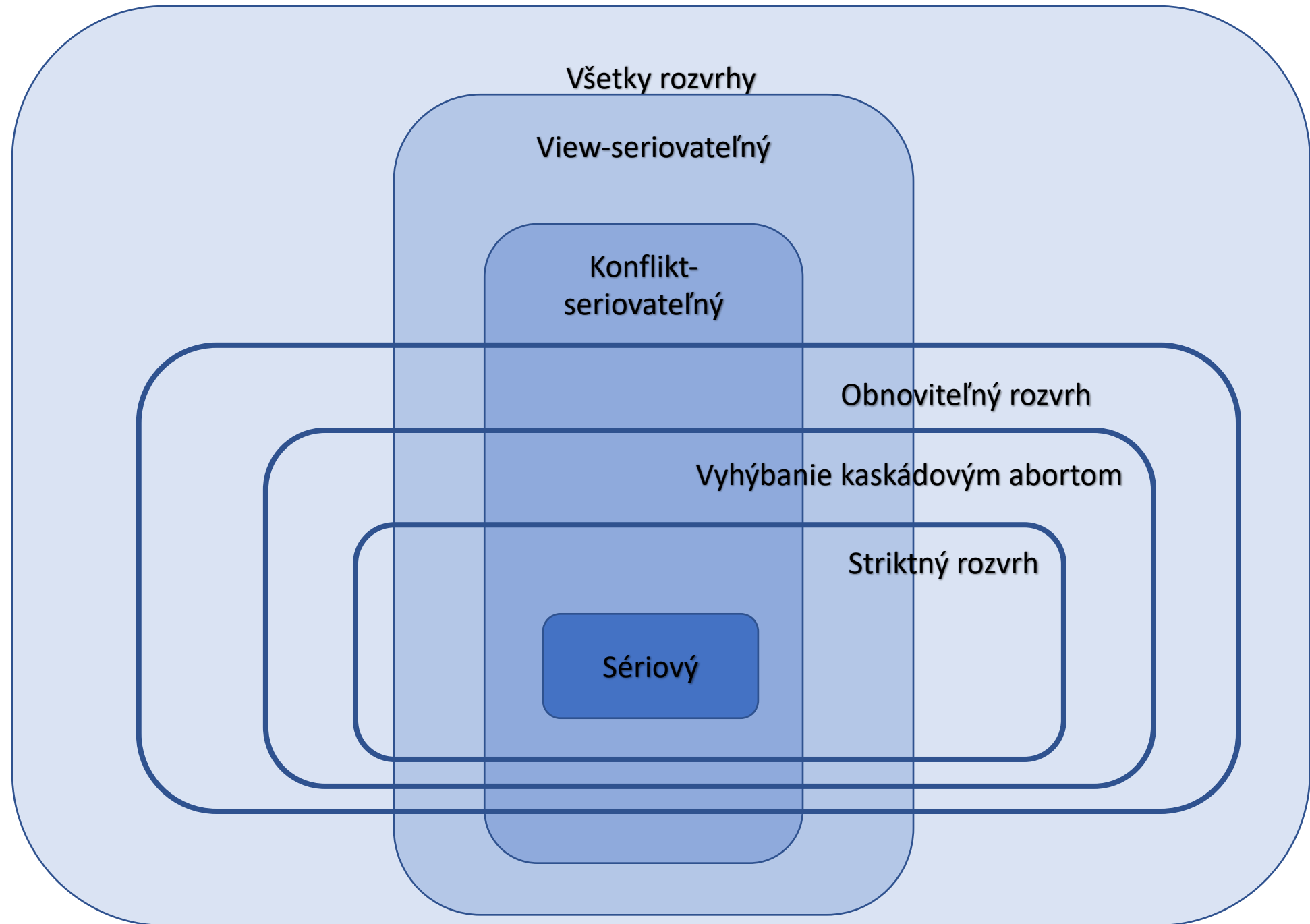
Príklad

- rozvrh S2: **w1(X), r2(X), w2(Y), c1, c2**
- je obnoviteľný
- Problém rozvrhu **S2** je hroziaci **kaskádový abort** (cascading abort) v prípade, že T1 sa rozhodne pre ABORT, tak T2 bude musieť skončiť ako ABORT tiež (Ak ma byť tento rozvrh považovaný za obnoviteľný)

Kaskadový abort - príklad

- Máme rozvrh S3:
 - **r1(x), w1(x), r2(x), w2(y), r3(y), w3(z), Abort1**
- Rozvrh, ktorý sa vyhýba kaskádovým abortom (Avoids Cascading Aborts), **práve vtedy keď neobsahuje dirty read**
- **Dirty read** – čítanie hodnoty zapísanej inou transakciou predtým ako prebehol COMMIT transakcie, ktorá modifikovala danú hodnotu

Rozvrhy



Strikne obnoviteľný rozvrh

- Striktný rozvrh – je taký rozvrh, ktorý neobsahuje **dirty reads ani dirty write**
- **Dirty write**
 - prepis hodnoty, ktorá nebola COMMIT inou transakciou
- Takýto rozvrh umožňuje rýchlejšiu obnovu dát – nie je nutné uskutočňovať REDO operáciu (nemusí sa druhýkrát prechádzať log)
- Takýto rozvrh vie generovať striktný 2PL

Zhrnutie

- Logovanie je potrebné pre zachovanie ACID a tiež pre možnosti obnovy dát pri výpadku
- Write-Ahead logovanie je vo väčšine prípadov výhodnejšie (Steal + No Force)
- CHECKPOINT – zrýchľuje obnovu a tiež skracuje log-file
- Pre obnovu sú potrebné operácie REDO a UNDO
 - v závislosti od spôsobu zapisovania