

Globální stav DS a distribuované transakce

Paralelní a distribuované systémy, přednáška 11

Tomáš Urbanec

Katedra informatiky PŘF UPOL

16.12.2025

Co nás čeká?

1. Globální stav

- Snapshoty a řezy
- Chandy-Lamport algoritmus
- Dijkstra-Scholten algoritmus

2. Distribuované transakce

- Motivace
- Jedno-, dvou- a třífázový commit

Globální stav

Globální stav

Základy

- Lokální stav procesu
 - hodnoty proměnných,
 - aktivované zdroje,
 - stav vykonávání programu.
- Globální stav DS
 - lokální stav všech procesů DS,
 - stavy všech kanálů mezi nimi.
- Řez: rozdělení události v systému na dvě skupiny
 - již se staly,
 - ještě se nestaly.
- Konzistence řezu $C: (a \rightarrow b \wedge b \in C) \Rightarrow a \in C$
- (tabule)
- Snapshot
 - $\approx$ globální stav DS v daném okamžiku
 - $=$ stav (konzistentního) řezu
- Posloupnost konzistentních řezů $\approx$ průběh výpočtu v DS

Chandy-Lamport algoritmus

Globální stav

- Výpočet snapshotu (zjištění stavu DS)
 1. Iniciátor uloží svůj stav pošle zprávu s markerem (snapshot request).
 2. Proces obdrží zprávu s markerem
 - 2.1 Úplně poprvé: uloží svůj stav, pošle jej iniciátorovi a nadále posílá marker s každou zprávou. Navíc začne sledovat stavy (zprávy) ostatních kanálů.
 - 2.2 Poprvé na daném kanálu: uloží stav kanálu a pošle jej iniciátorovi.
 3. Proces s markerem dostane zprávu bez markeru
 - musí být z okamžiku před snaphostem,
 - uloží ji do stavu kanálu.
- FIFO kanály (jinak problém)
- (tabule)

Dijkstra-Scholten algoritmus

Globální stav

- Detekce ukončení výpočtu.
- V obecné síti buduje kostru ($\approx$ volba leadera v ad-hoc síti).
- Uzel je aktivní, nebo skončil.
- Zprávy
 - SIGNAL: Výzva počítáš/nepočítáš?
 - ACK: Odpověď už nepočítám, nebo se už hlásím jinam.
- V síti bez cyklů lze modelovat pomocí deficitů kanálů, tj.
#SIGNAL - #ACK zpráv na vstupu/výstupu kanálu.
- (tabule)

Distribučované transakce

Distribuované transakce

Základy

- Transakce, která zahrnuje více uzlů v DS.
- Mezibankovní převod (různé systémy),
- zápis do více různých DB (i lokálně),
- . . . ,
- obecná změna více úložišť/systémů.
- Zaměříme se na distribuovaný commit.
 - ≈ Závěrečná fáze distribuované transakce.
 - Skutečné provedení změn na všech uzlech.
 - Musí být provedeno buď všude, nebo nikde.

Jednofázový commit

Distribučované transakce

- Naivní řešení.
- Procesy: koordinátor a následovníci.
- Koordinátor pošle všem následovníkům instrukci.
- Hotovo ...
- ... ale co když
 - uzel nejede?
 - zpráva nedorazí?
 - uzel neumí reagovat dle požadavků?
 - ...
- (tabule)

Dvoufázový commit

Distribuované transakce

- Rozumné řešení.
- Procesy: koordinátor a následovníci
- Koordinátor pošle všem VOTE_REQUEST.
 - Uzel je připraven → reaguje VOTE_COMMIT.
 - Uzel není připraven → reaguje VOTE_ABORT.
- Koordinátor dostal zpět
 - všechny VOTE_COMMIT → GLOBAL_COMMIT.
 - alespoň jeden VOTE_ABORT → GLOBAL_ABORT.
- Uzel dostal GLOBAL_COMMIT → provede commit.
- (tabule)

Dvofázový commit

Distribuované transakce

- Problémy?
 1. Koordinátor vypadne (následovník v READY)
 2. Následovník vypadne (koordinátor ve WAIT)
- Řešení:
 1. čekání na zotavení koordinátora, nebo domluva s ostatními.
 2. koordinátor čeká, nebo ABORT.
- Nejpoužívanější varianta.
- Např. PostgreSQL
 - PREPARE TRANSACTION,
 - COMMIT/ROLLBACK PREPARED.
 - Neplést s (klasickými) transakcemi.

Třífázový commit

Distribučované transakce

- Řeší problémy dvoufázového commitu.
 1. Ze žádného stavu nelze volit rovnou COMMIT nebo ABORT.
 2. Neexistuje stav, kde neznáme výsledek a lze přejít do COMMIT.
- Nový (mezi)stav PRECOMMIT a zpráva PREPARE-COMMIT.
- (tabule)
- Problémy?
 1. Kordinátor vypadne (následovník READY nebo PRECOMMIT)
 2. Následovník vypadne (koordinátor v PRECOMMIT)
- Řešení:
 1. čekání na zotavení koordinátora, nebo domluva s ostatními (většina READY → ABORT).
 2. koordinátor čeká, pak COMMIT (vyžaduje zotavení u následovníka)

Shrnutí

Shrnutí

- Pokrytá témata na webu.
- Spíše jsme klouzali po povrchu.
- Alespoň jednou otevřete knihu Van Steen & Tanenbaum!
- Pokud Vás něco zaujalo, otevřete i doplňkové zdroje.
- Slidy mohou dostat drobných úprav
 - korektura a formualce,
 - ne obsah.
- Na slidech není vše, co bylo řečeno. . .

Changelog