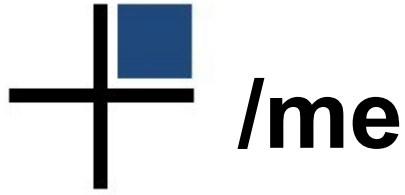


PostgreSQL Clustering

Petr Jelínek

petr@2ndquadrant.com



PostgreSQL Developer & Consultant @
2ndQuadrant

PostgreSQL contributor 10+ let
Co-maintainer PgBouncer, BDR, pglogical



Škálování

Vertikální

- Maximální využití jednoho serveru
- Lepší HW

Horizontální

- Rozložení výkonu na více serverů
- Více HW



Škálování

OLTP – čtení

OLTP – zápis

OLAP – velké databáze a query

Latence



Vertikální škálování

OLTP

- Lock fastpath (9.2)
- Atomické operace (9.5)
- Lockfree buffer replacement (9.5)



Vertikální škálování

OLAP

- Paralelní query (9.6)
- Sampling (9.5)
- Brin (9.5)
- Optimalizace sortu, atd (9.5)



Horizontální škálování

Hot Standby

PL/Proxy

Greenplum

Postgres-XL

CitusDB

BDR



Hot Standby

Stabilní

Integrované do Postgresu

- Jednoduchá instalace
- Široká podpora externích nástrojů

Jen čtení

- V 9.6 paralelní scany



Hot Standby

Primary

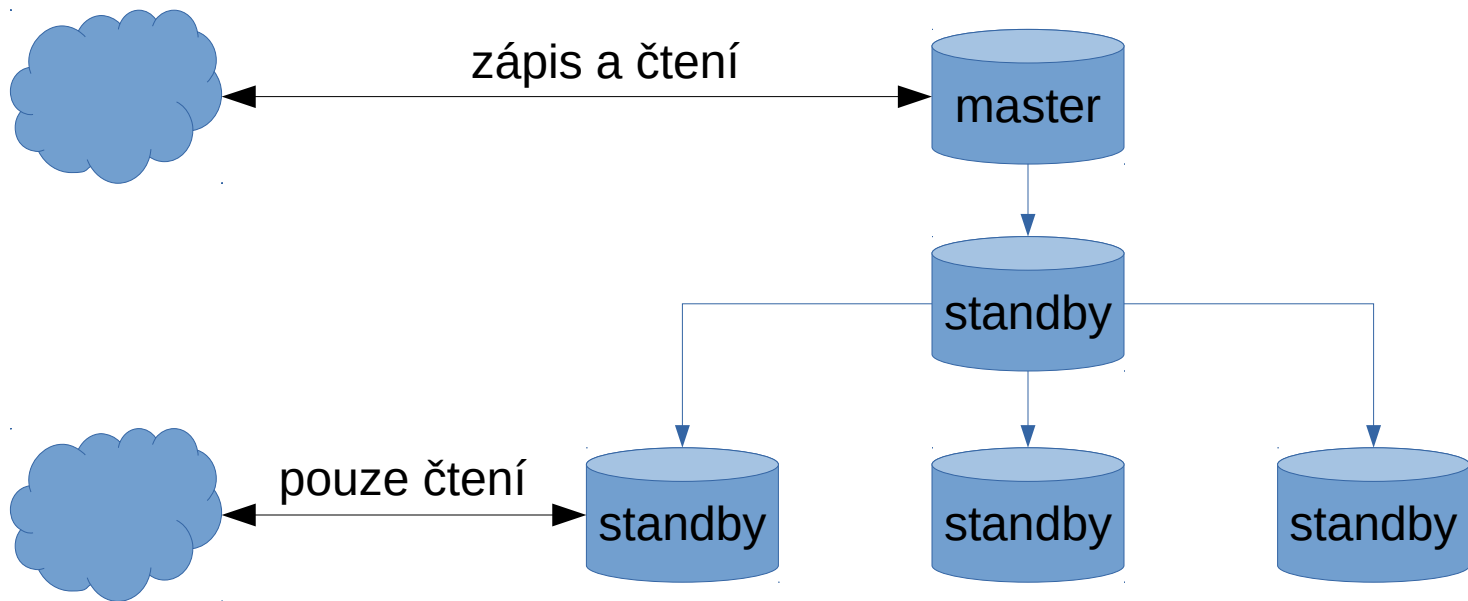
- Zápis i čtení
- Žádná metadata o clusteru*

Standby

- Pouze čtení
- Čte a aplikuje transakční log z primary
- Přesná kopie primary



Hot Standby





Skype

Open-source extenze (BSD)

Stabilní

Feature complete – jen maintenance

Bez podpory

Primárně OLTP – čtení i zápis

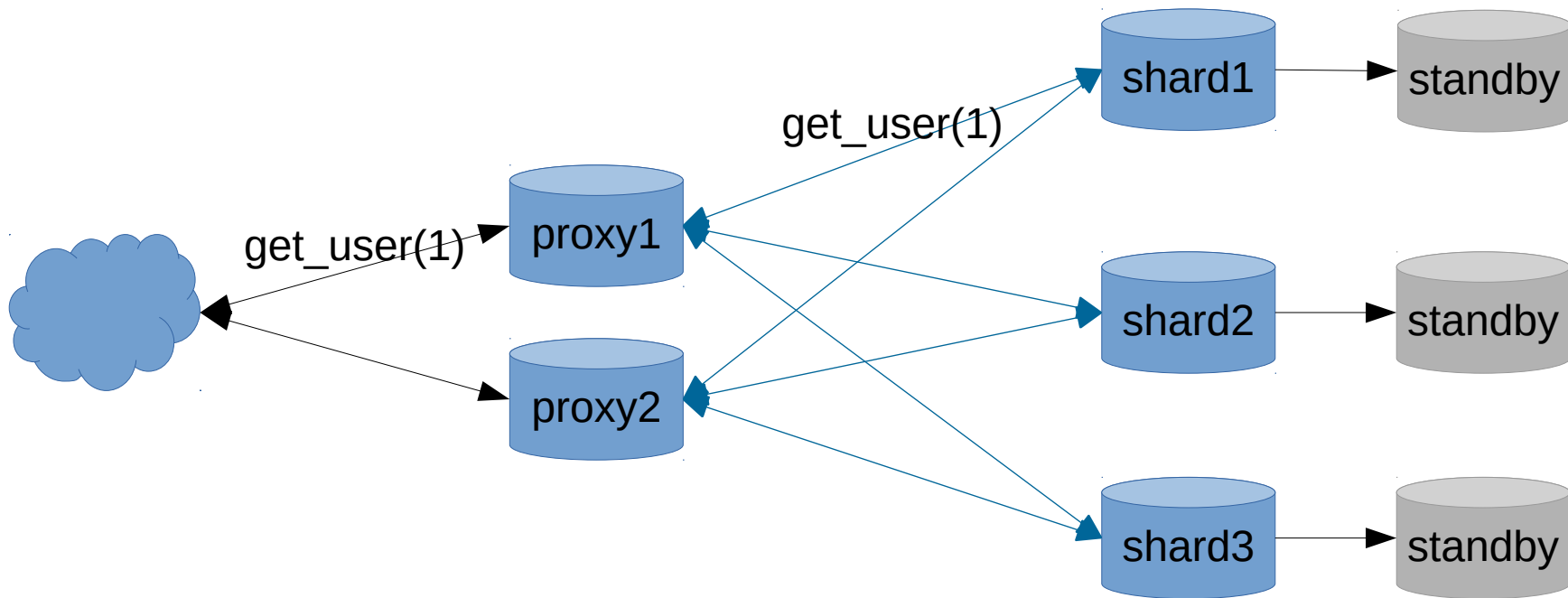


Map-reduce

- Funkce
- Není transparentní
- Nemá globální konzistenci
- Vlastní jednoduchý jazyk

<http://shards.io>

PL/Proxy





Pivotal

Fork Postgresu (~8.2)

Stabilní

Open-source od loňského podzimu (Apache)

MPP databáze

- OLAP



Historicky větší odklon od Postgresu

- Nový optimizer
- Vlastní tooling
- Vlastní reimplementace post 8.2 vlastností
- Tabulky optimalizované na INSERT



Master server

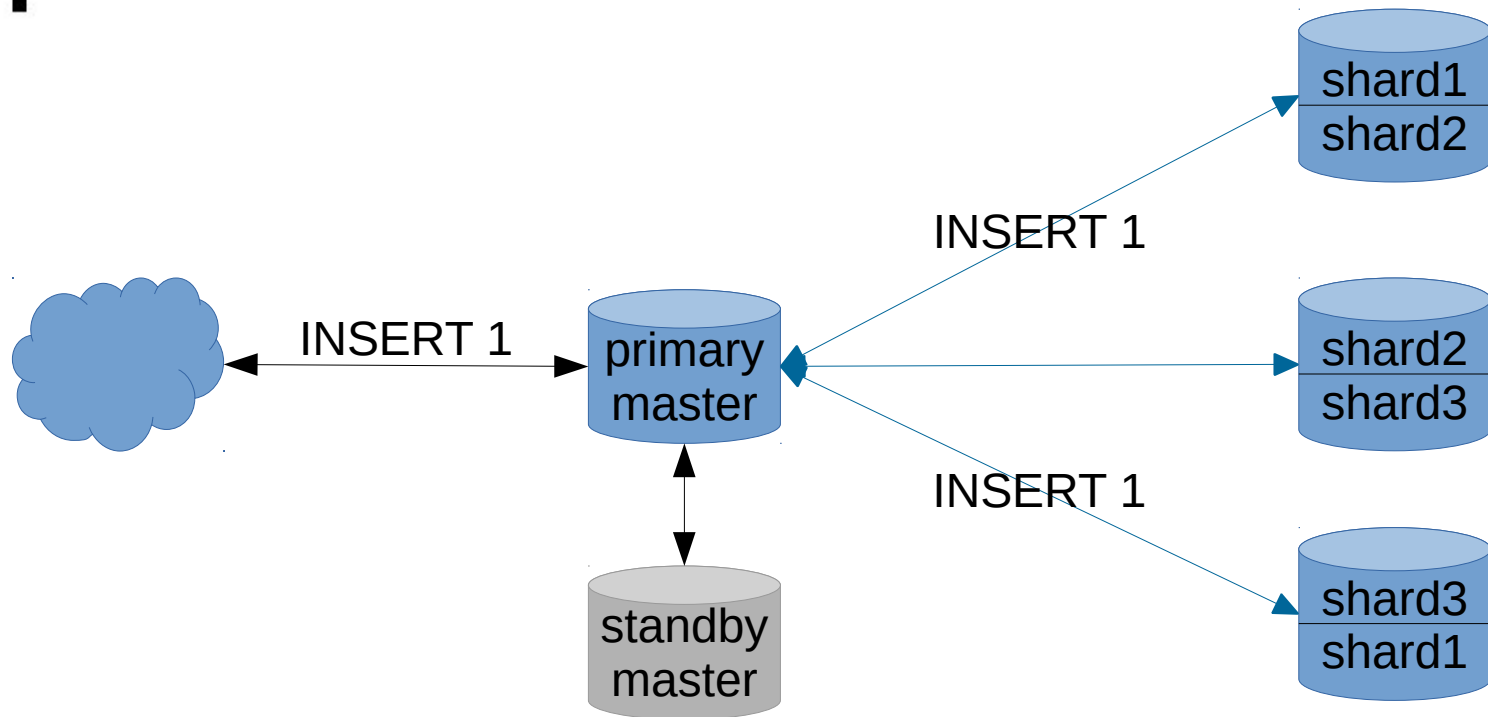
- Metadata
- Distributed query planner

Segment server

- Několik logických shardů
- Distribuované tabulky



Greenplum





CitiusDB

CitusData

Fork Postgresu + estenze (9.4)

Brzy pouze estenze (9.4+9.5)

Brzy open-source varianta (9.4+9.5)

OLTP i OLAP



Master node

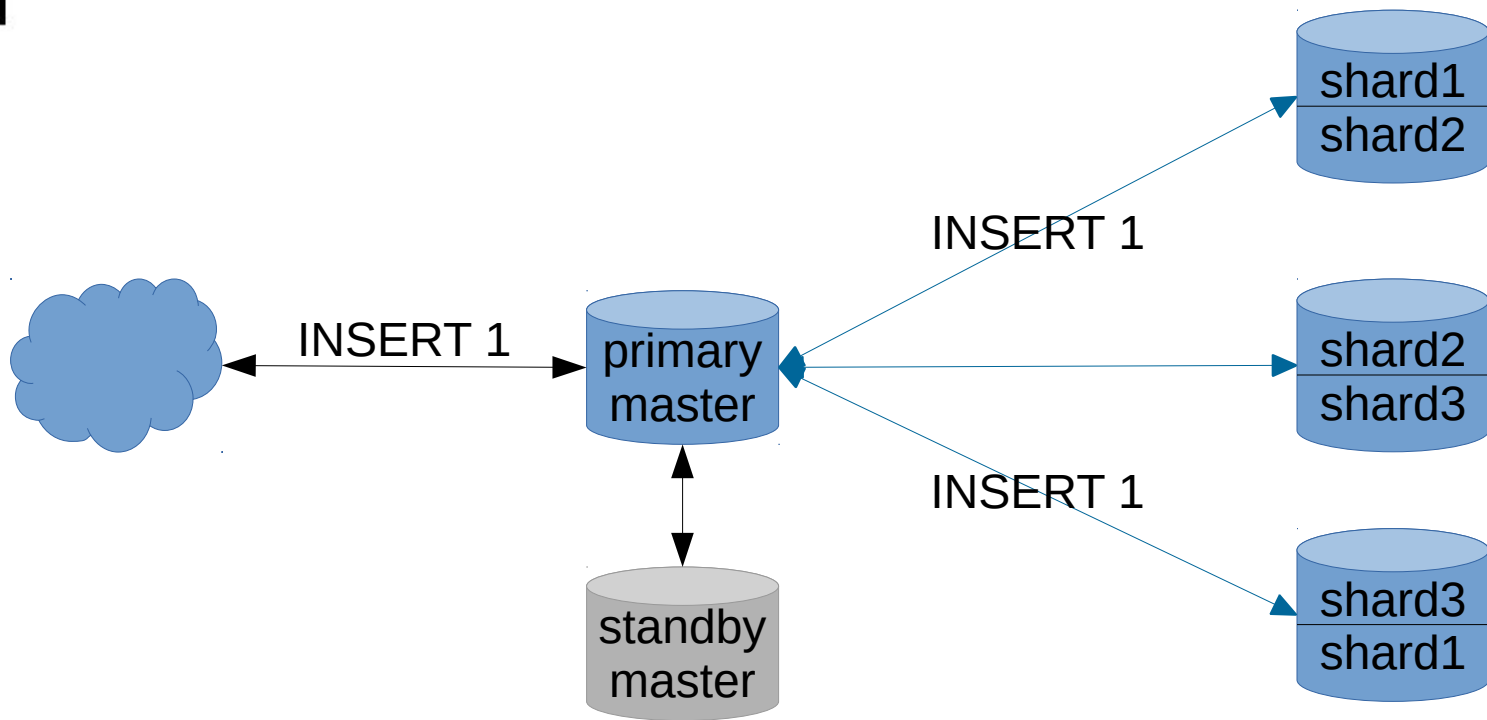
- Metadata
- Distributed query planner pro čtení
- pg_shard pro distribuci dat

Worker nodes

- Mnoho logických shardů
- Distribuované tabulky



CitusDB





Postgres-XL

Fork Postgresu (9.2)

Vychází z Postgres-XC, StormDB

Open-source (MPL)

Brzy 9.5

MPP

- OLTP i OLAP



Postgres-XL

GTM

- Globální konzistentní snapshoty

Coordinator

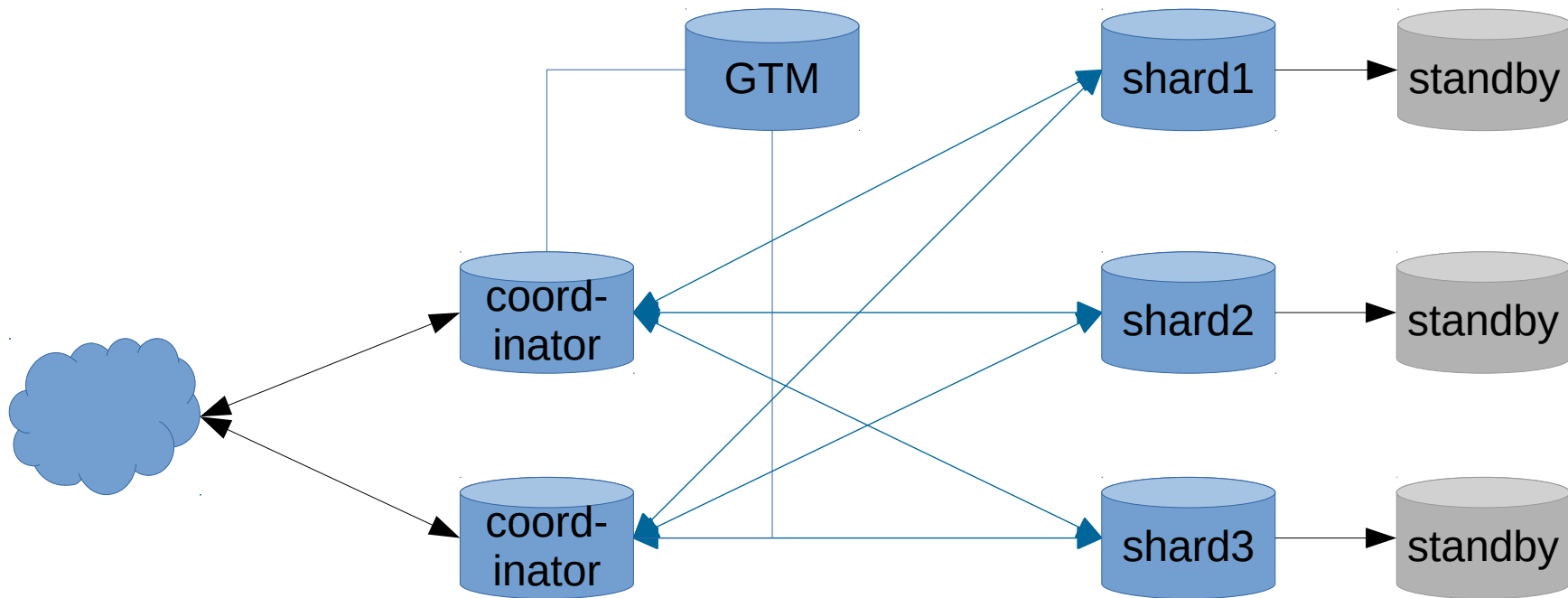
- Metadata
- Distributed query planner

Datanode

- Distribuované a replikované tabulky (star schema)
- Fyzické shardy



Postgres-XL





2ndQuadrant

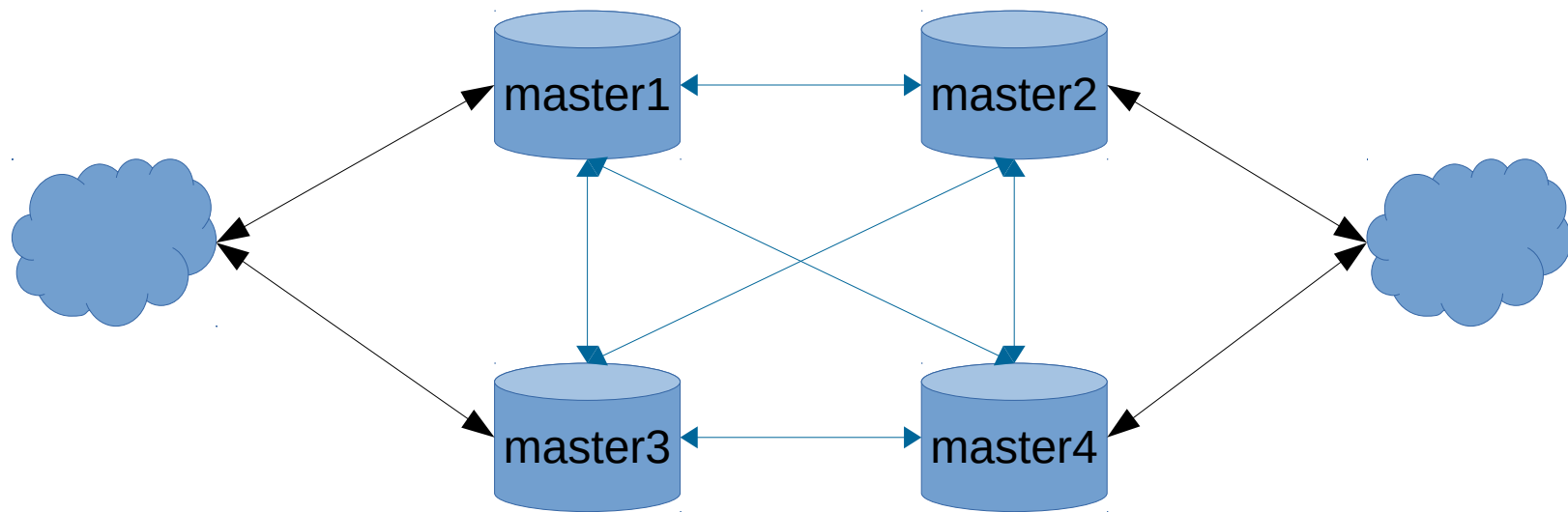
Fork Postgresu (9.4) s cílem integrovat jednotlivé komponenty zpět do PostgreSQL

- Logical decoding, replication slots, replica identity
- Background workers, event triggers
- Commit timestamps, replication origins



Asynchronní multimaster

- Každý server má všechna data
- Každý server přijímá zápis
- Optimistická detekce a řešení konfliktů
- Eventually consistent
- Primárně zaměřeno na optimalizaci latence





Dotazy?
We are hiring