PostgreSQL的大版本升级方案

嘉宾：高云龙
公司：去哪儿网

- Why Upgrade Major Version
- Traditional Method
- How PgQ do it
- Conclusion && Future

Why Upgrade Major Version

- Major enhancements

Version	New Features
9.3	add materialized views support event triggers
9.4	add jsonb add alter system add logical decoding
9.5	support upsert add BRIN
9.6	parallel execution of sequential scans allows multiple synchronous standby servers
10	Logical replication using publish/subscribe Declarative table partitioning Improved query parallelism

https://www.postgresql.org/docs/current/static/release.html

- End Of Life (EOL) dates

Version	Current minor	Supported	First release date	EOL date
10	10.0	Yes	October 2017	October 2022
9.6	9.6.5	Yes	September 2016	September 2021
9.5	9.5.9	Yes	January 2016	January 2021
9.4	9.4.14	Yes	December 2014	December 2019
9.3	9.3.19	Yes	September 2013	September 2018
9.2	9.2.23	No	September 2012	September 2017

Traditional Method

	advantage	disadvantage
pg_upgrade	单DB Node速度快(物理copy且可开多threads)	pg的make及initdb 参数必须一致或兼容 必须停止Instance的运行 无法同时做Cluster的standby 某些第三方Extension需要额外处理
pg_dumpall	pg的make及Instance 的initdb参数可不同 可继续提供只读服务 同时做Cluster的Standby 可灵活处理Extension，保证升级后兼容性	DB Instance不可写 单DB Node速度慢

既要99.999的服务，又要保证升级的兼容性
鱼和熊掌是否可兼得？？？

让我思考会人生

How PgQ do it

What is PgQ

Skytools

It is a package of tools in use in Skype for replication and failover

PgQ

It is a queuing system written in PL/pgSQL,Python and C code
It is split into 3 layers:Producers, Ticker, Consumers

Londiste

It is a Replication tool written in Python, using PgQ as event transport

Walmgr

It is a script to dose runtime WAL archive,initial backup and restore

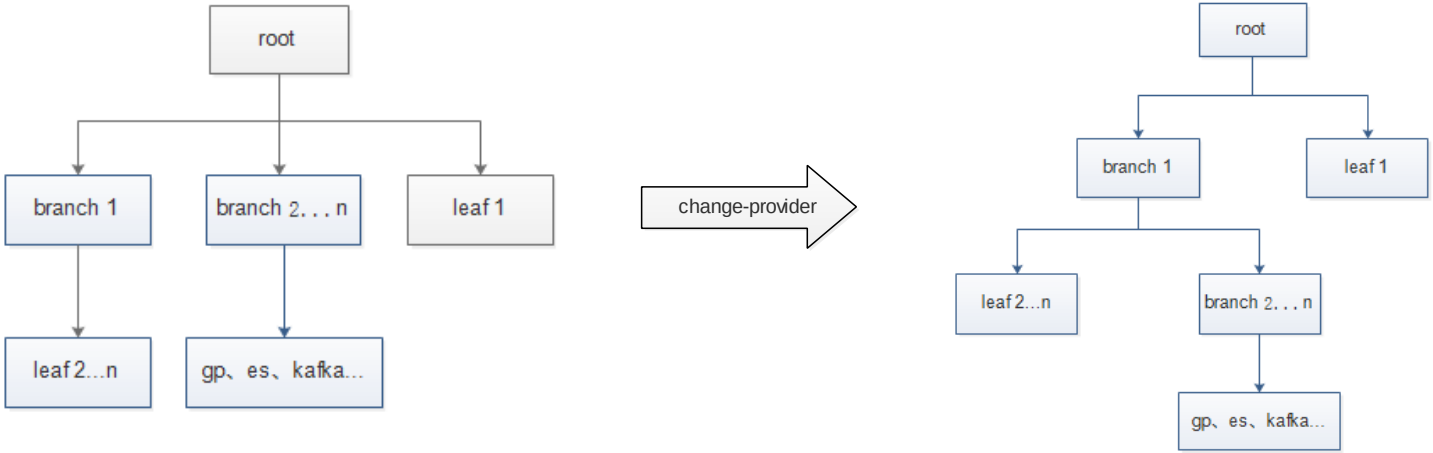
PgQ Cluster Node

root
整个同步集群的唯一数据写入入口，向下游推送event

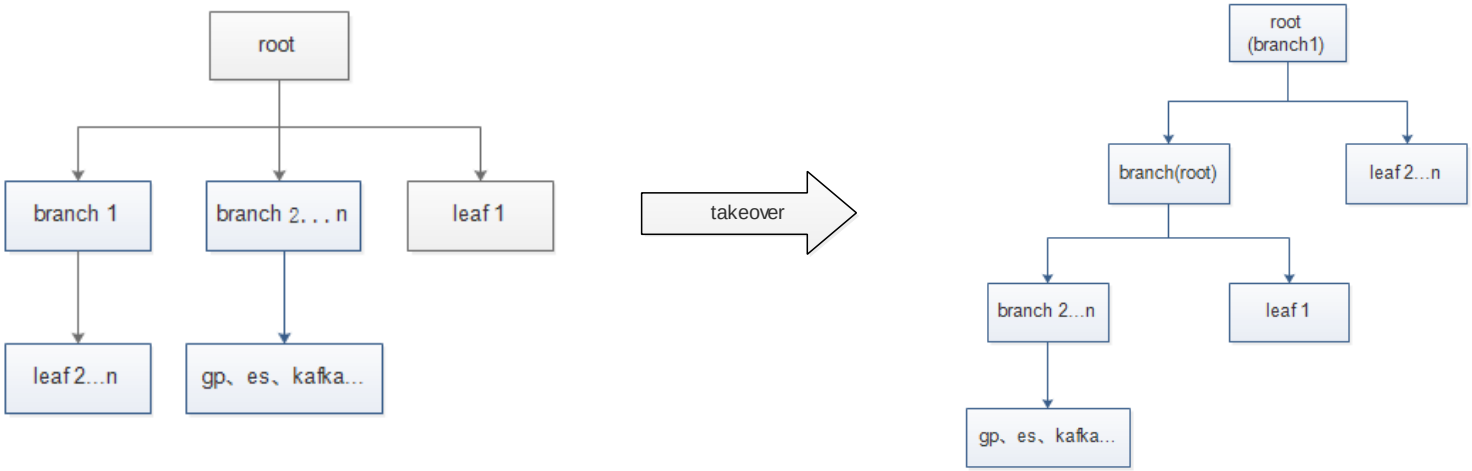
branch
既可接收上游推送来的event写入到自己的DB 实例，又可将event推送到下游节点

leaf
接收上游推送的event并写入到自己的节点数据库

Change Node



Change Root



Upgrade Operating process

- 确定PgQ及PostgreSQL版本
- 配置PgQ参数文件
  pgq 配置文件 ：base_connstr、database_list
  node配置文件：pgq_queue_name、job_name、db
- 添加同步对象
  table必须有Primary Key(或unique + not null)
- 切换 root 节点
  londiste3 branch.ini takeover root_node
- 重启高版本PostgreSQL DB实例

Conclusion

- 可秒级切换
- 可保证Upgrade前后PostgreSQL及诸多Extension的兼容性
- 可支持灵活的回滚操作，数据不丢失

Future

- 提交patch，更好的支持pg9.6及pg10.0
- 两条腿走路，实践基于Logical Decoding的方案
  dbsync (https://github.com/aliyun/rds_dbsync)
  pglogical (https://www.2ndquadrant.com/en/resources/pglogical/)
  pgrepup (http://gasparin.net/)

Thanks!