



Postgres中国技术大会 2017 ( PG 大象会 )

# PPT 分享



Postgre中国用户会

## ◆ QQ交流群

PostgreSQL专业2群: 100910388

PostgreSQL专业3群: 150657323

PostgreSQL专业4群: 461170054

◆ 文档翻译群: 309292849

◆ 欢迎投稿: [press@postgres.cn](mailto:press@postgres.cn)

## ◆ 微信公众号



官方微信公众号

## ◆ 新浪微博



官方微博

## ◆ 官方网站

[www.postgres.cn](http://www.postgres.cn)

## ◆ FaceBook

China PostgreSQL User Group

## ◆ Twitter

China PostgreSQL User Group



# 2017

## Postgresql 版本升级方案演进

postgresql / upgrade / multi solutions

Ping An Technology was founded in 2008 as a wholly-owned subsidiary of Ping An Group. Headquartered in Shenzhen, the company has branches in Beijing, Shanghai, Chengdu and Nanjing, and employs more than 4,000 IT professionals and management experts.



2017 Postgres大象会



平安科技  
PING AN TECHNOLOGY

石勇虎，平安科技数据库技术部高级数据库工程师，有多年数据库运维，开发经验，目前主要致力于postgresql架构，优化，监控，版本升级等工作。





# 目录 Catalog

1 升级背景  
Background.

2 方法论与本土化  
Multi solutions and practice.

3 总结  
Summary.



2017 Postgres大象会



平安科技  
PING AN TECHNOLOGY

# 01.

## 升级背景

Background in Ping An.



2017 Postgres大象会





**1** 9.4  
2015年初PG9.4.1开始投产，年底已达到400+实例

**2** 9.5  
2016年引入PG9.5，并且，同年PG实例数突破1000+，成为国内最大的PG用户

**3** 9.6  
2017年引入PG9.6，到目前为止已达到1500+实例数，平均每天都有1-2新建库的需求。

**4** 10.X  
2018年初会引入PG10版本



### 升级必要性

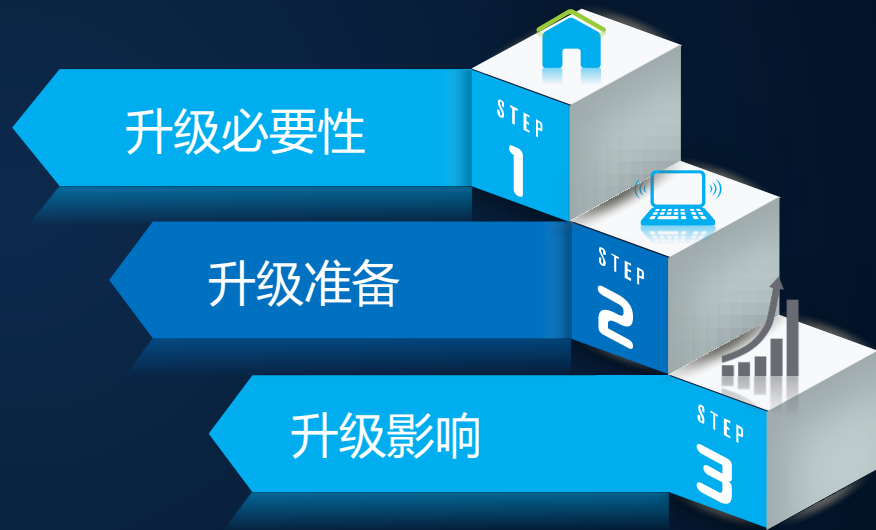
- 1 新功能增加，性能提升，提高用户体验。（dba推动升级或者用户要求升级）  
多个版本并行，增加管理复杂度。  
版本生命周期。

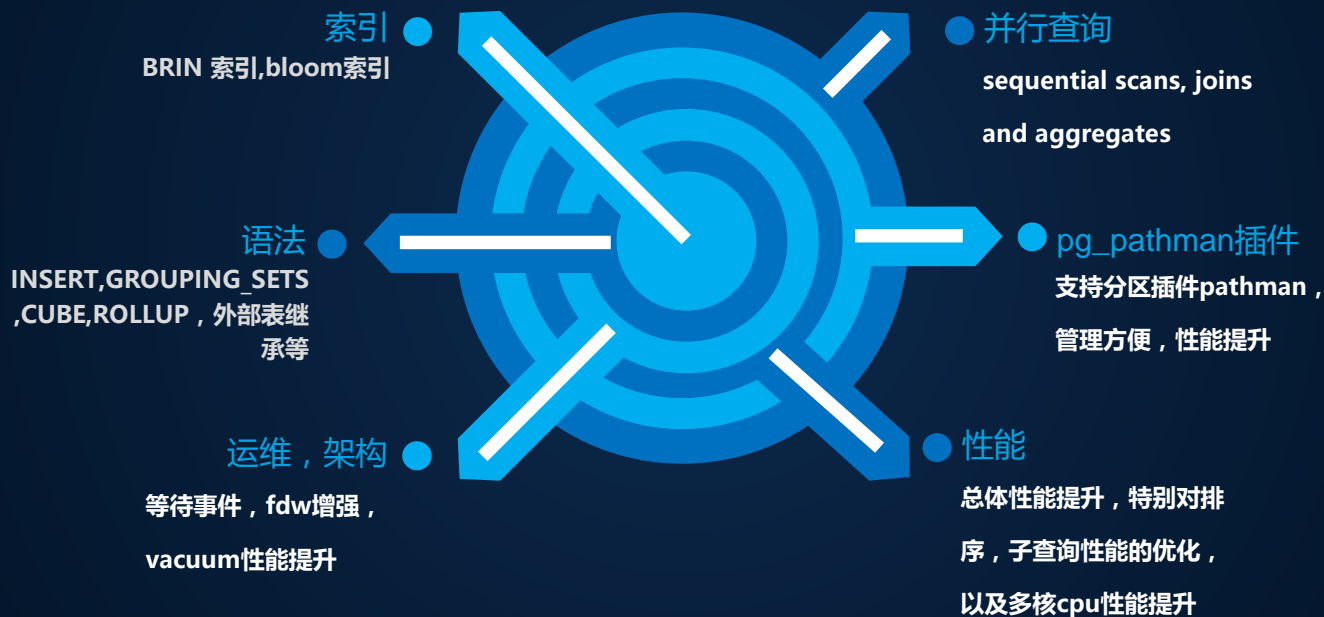
### 升级准备

- 2 升级方案评估，升级环境准备，功能测试，性能测试，失败回退等

### 升级影响

- 3 升级所涉及应用系统，升级过程中应用中断对业务的影响，升级带来的收益。







# 02.

## 方法论与本土化

Find the most suitable method for Ping An.



2017 Postgres大象会



平安科技  
PING AN TECHNOLOGY

## 升级列表

| No | 工作项（升级前）         | 明细  |
|----|------------------|---|
| 1  | 确定升级方案           | 确定升级方案，升级流程及时间点，完成升级文档                    |
| 2  | 升级开发测试环境         | 升级开发测试环境，升级后发版时需要源测试环境和升级后测试环境同步进行        |
| 3  | 回归测试             | 开发测试人员完成回归测试                              |
| 4  | 性能测试             | DBA完成性能测试                                 |
| 5  | 同城环境准备           | 同城环境准备（已有同城，需要检查同城环境同步是否正常），需要主机组分配主机     |
| 6  | 同城环境准备           | 同城环境准备（已有同城，需要检查同城环境同步是否正常），存储组分配挂载/data卷 |
| 7  | temp卷准备（或搭建逻辑复制） | 生产备机挂载temp卷，用于存放逻辑导出文件                    |
| 8  | 备份               | 升级前dba需要完成一次pg_rman全备                     |
| No | 工作项（升级过程中）       | 明细  |
| 1  | 停生产              | 运营停应用，dba停生产库                             |
| 2  | 实施升级             | dump/restore,pg_upgrade,pglogical         |
| 3  | 验证               | dba验证数据一致性                                |
| 4  | 应用验证             | 通知运营验证应用                                  |
| No | 工作项（失败回退）        | 明细  |
| 1  | vip切换至同城         | 将vip切换至同城（或者源生产），由同城（源生产）直接提供服务           |



## 性能测试

- 1.通过pgawr获取top sql，logana获取绑定变量值
- 2.在9.4和9.6测试环境分别进行回放，并通过logana工具解析日志获取执行计划
- 3.使用report工具生成测试报告

| sqlID  | 9.4     |          | 9.6     |          | 执行计划是否改变 | 执行计划改变原因                   | 性能是否下降 |
|--|---------|----------|---------|----------|----------|----------------------------|--------|
|  | buffers | time/ms  | buffers | time/ms  |          |                            |        |
| <a href="#">168a4320fc13d6d1bf1a8a54df978d20</a> | 27      | 1400.387 | 20      | 1572.568 | 否        |                            | 是      |
| <a href="#">44dd33a10b61b9583189aa372a2596fc</a> | 614596  | 16087    | 97372   | 39575    | 是        | HashAggregate-》Sort Unique | 是      |
| <a href="#">34181f09de66160112a5209c482b336c</a> | 1569    | 10.327   | 2801    | 30.264   | 否        |                            | 是      |
| <a href="#">ccb76b7497a3347af3c139ada758e402</a> | 1272    | 1229921  | 1101    | 1933074  | 否        |                            | 是      |
| <a href="#">ee16278e2aff49bde125d6dc772b101e</a> | 1446    | 71       | 2745    | 133      | 否        |                            | 是      |
| <a href="#">5e41a2cec5b5496767a0ac018381bb57</a> | 706     | 205191   | 791     | 673384   | 否        |                            | 是      |
| <a href="#">80e17a2e63dccb4b6f69352bae6a1b11</a> | 1175952 | 500713   | 1504316 | 692358   | 否        |                            | 是      |
| <a href="#">49e971bd8c1abb4b6e229403be8800d5</a> | 93825   | 240      | 20840   | 264      | 否        |                            | 是      |
| <a href="#">b07d22f4b86c357f2d3ac6f7e6136f88</a> | 0       | 19192    | 0       | 36907    | 否        |                            | 是      |
| <a href="#">77362cca8cf7bab17c4f020c980e6ab</a>  | 0       | 26734.34 | 0       | 39474.71 | 否        |                            | 是      |
| <a href="#">d7624ab56f62d933a95968dba26c9068</a> | 4373375 | 564300   | 7230557 | 642701   | 是        |                            | 是      |
| <a href="#">3f05a9feed45030d39d4520b32f5922c</a> | 93825   | 386.928  | 0       | 3455.367 | 否        |                            | 是      |
| <a href="#">5b4250eb4e5808b8a185e80a339a4d65</a> | 510     | 0.781    | 0       | 113      | 否        |                            | 是      |
| <a href="#">bb4cdab9f2c11ba95f7ec30427abb414</a> | 0       | 181404.6 | 0       | 264118.2 | 否        |                            | 是      |
| <a href="#">428690b8932ebf8fb24eea2faa04c3b9</a> | 1805866 | 421829.8 | 2019517 | 1104113  | 否        |                            | 是      |



### 物理升级

维护窗口时间短。  
但必须保证不同版本间编译参数一致

1

pg\_upgrade

### 逻辑升级

逻辑复制，切换时间短，支持升级后失败回退。  
前期准备工作较多，表必须有主键。

3

pg\_logical

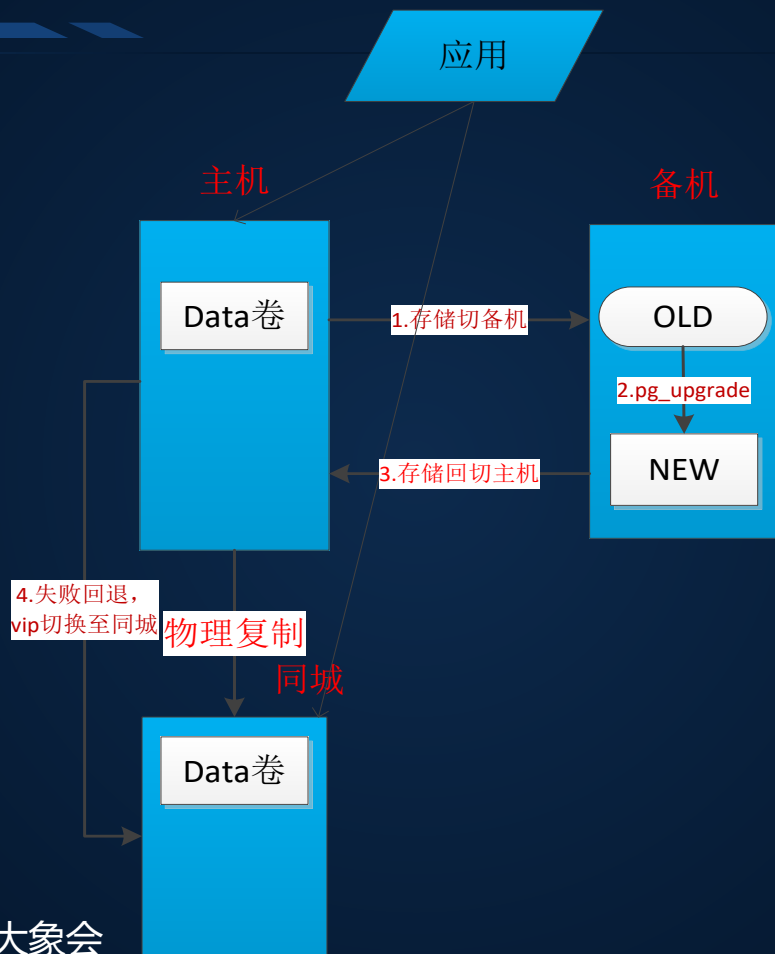
2

pg\_dump/pg\_restore

### 逻辑升级

消除deadtuple，升级准备期间支持应用发版  
数据需要逻辑导出，导入，维护窗口时间相对较长。









# 03.

## 总结

Summary



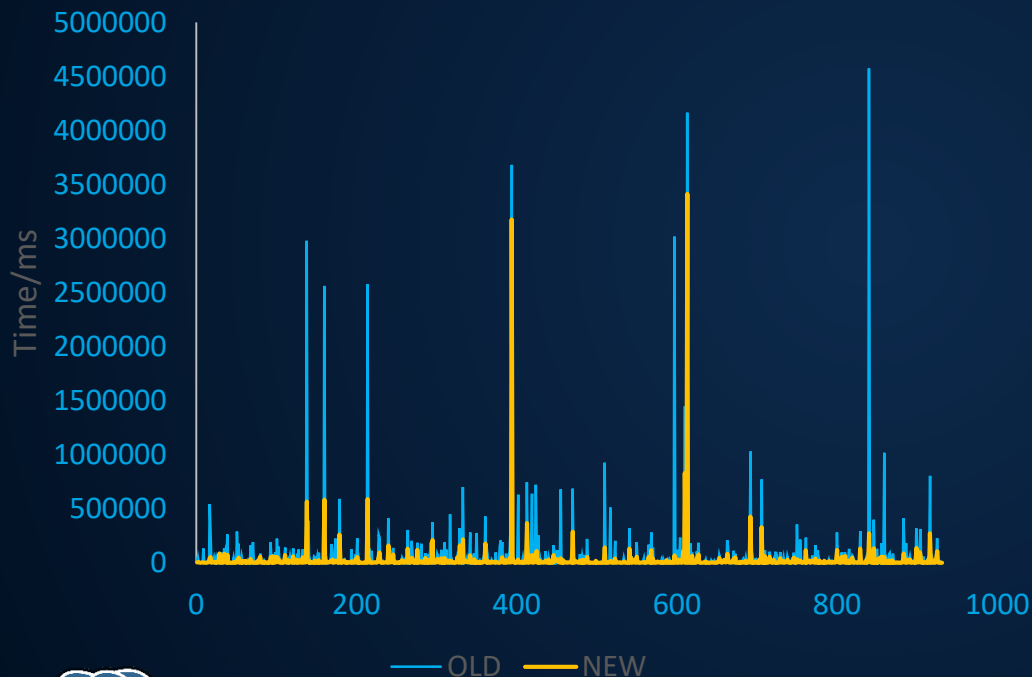
2017 Postgres大象会



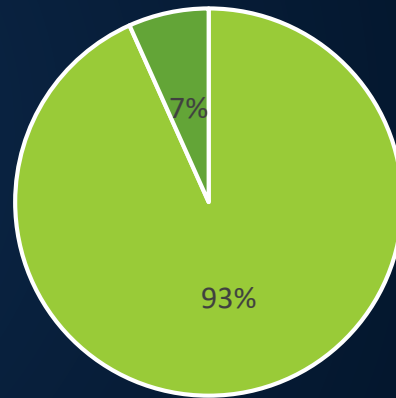
平安科技  
PING AN TECHNOLOGY



## 结果对比



### sql条数（时间大于1s）



■ 执行时间减少 ■ 执行时间增加



在9.6.2中， `group by` 使用 `GroupAggregate + Sort` 方式，而9.4.1使用的是 `HashAggregate`。在Oracle 10g以后，oracle也是使用 `HashAggregate` 的方式做 `group by` 的。

```
enable_sort (boolean)
```

```
enable_hashagg (boolean)
```



调整random\_page\_cost

基于成本优化器，随机的PAGE访问的开销seq\_page\_cost的多少倍，seq\_page\_cost默认是1，random\_page\_cost默认是4.

高端SAN,SSD建议random\_page\_cost调小到趋近于seq\_page\_cost甚至相等





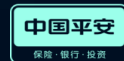
# 2017 演示完毕 感谢聆听

Thanks for listening

Technology drives finance.



2017 Postgres大象会



平安科技  
PING AN TECHNOLOGY