Query routing technique with Pgpool-II

(Pgpool-II的查询路由技术)

PgPool Global Development Group
(PgPool 全球开发组)

Tatsuo Ishii

石井 達夫

China PostgreSQL Association

Translated by: China PostgreSQL Association
由中国PostgreSQL分会协助翻译

SRA OSS, INC.

About me

- Tatsuo Ishii
  - Live in Japan
  - Pgpool-II Community Lead
  - PostgreSQL Committer
  - Love to read sci-fi books
    喜欢阅读科幻书籍

“The three body problem”
by Liu Cixin
刘慈欣的《三体》

Copyright(c) 2020 SRA OSS, Inc. Japan

What is Pgpool-II?

Pgpool-II 是什么？

- A proxy/middle ware for PostgreSQL
  （PostgreSQL的代理/中间件）
- Distributed under the same OSS license as PostgreSQL
  （使用与PostgreSQL相同的OSS许可证分发）

对现有PostgreSQL数据库应用程序透明

Transparency to existing PostgreSQL DB applications

Automatic failover Failback
自动故障转移
故障回复

In memory query result cache
内存中缓存查询结果

Avoid single point of failure
避免单点故障

Query routing
查询路由

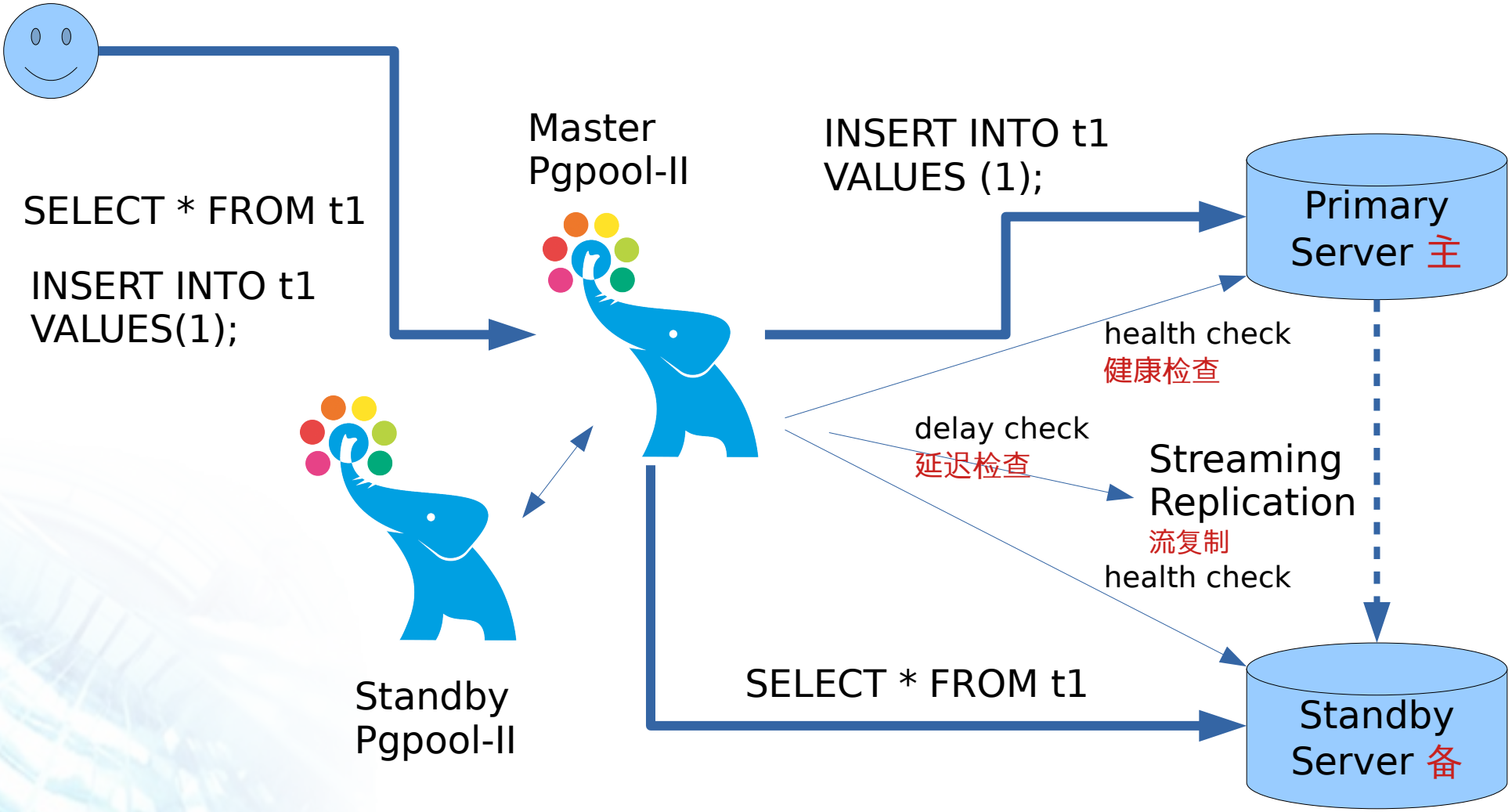
Read query load balancing
读取查询负载平衡

Copyright(c) 2020 SRA OSS, Inc. Japan

SRA OSS, INC.

Architecture of Pgpool-II

（Pgpool-II架构）



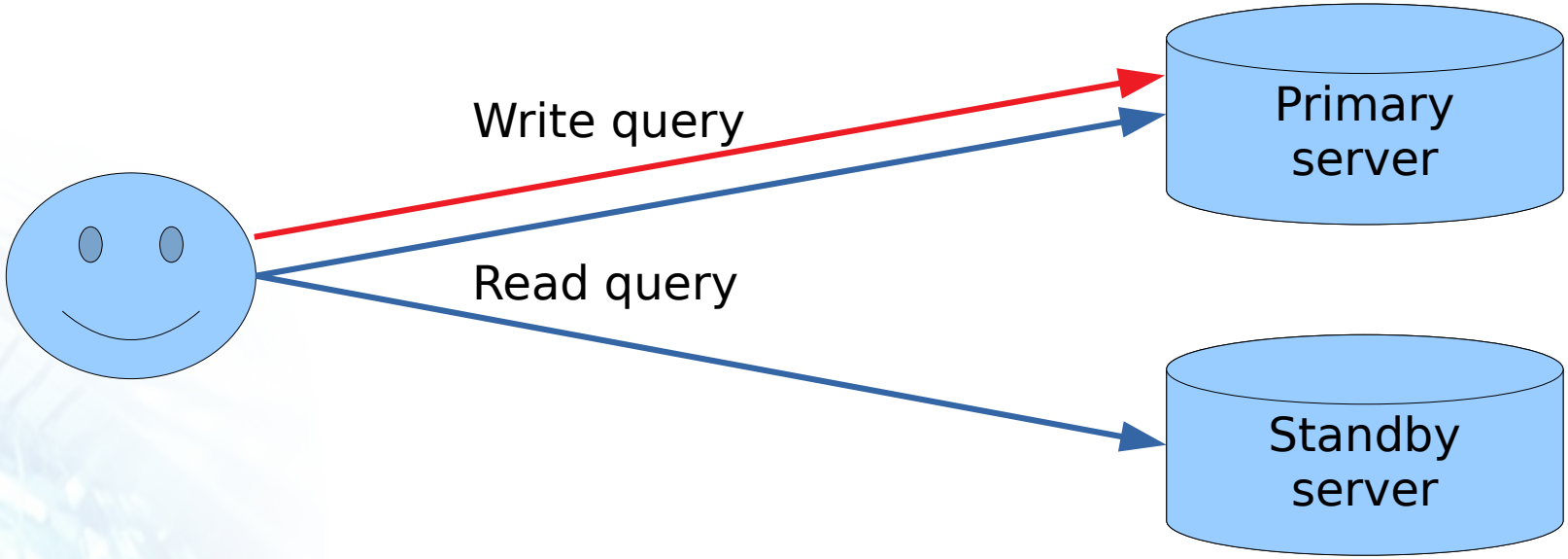
Copyright(c) 2020 SRA OSS, Inc. Japan

SRA OSS, INC.

What is query routing? Why we need it?

什么是查询路由？我们为什么需要它？

- PostgreSQL's streaming replication has two types of servers: **primary server** (accepts r/w query) and **standby server** (only accepts read query)
  PostgreSQL的流复制有两种类型的服务器：主服务器（接受读/写查询）和备用服务器（仅接受读查询）
- We need to route the write query (**query routing**) to only primary
  我们需要将写查询（查询路由）仅路由到主数据库
- Read queries can be routed to either primary or standby according to the requirement
  可以根据需要将读取查询路由到主查询或备用查询

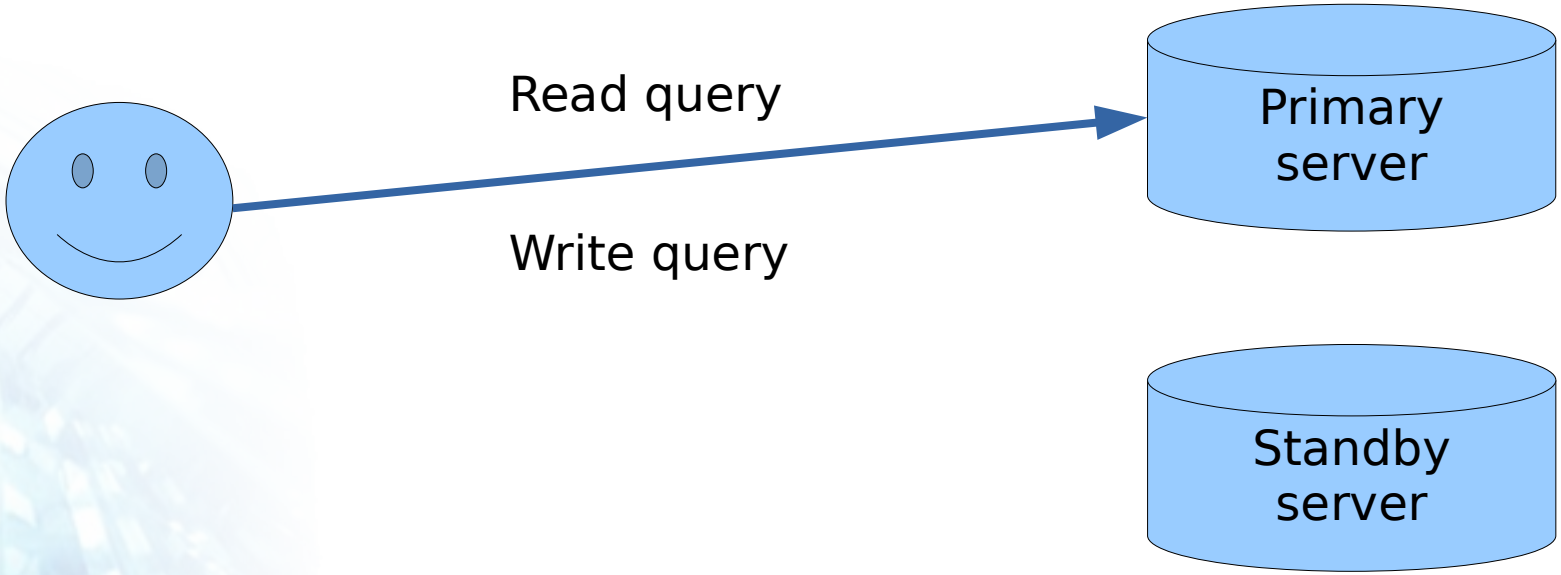


Copyright(c) 2020 SRA OSS, Inc. Japan

Pros and Cons of query routing methods

查询路由方法的优缺点

- Route everything to the primary
  将所有请求都路由到主数据库
  - ✓ Simple and easy 简单、容易
  - ✓ Load of the primary could become high 主服务器的负载可能会变高
  - ✓ Wasting resource of standby servers 备用服务器的资源浪费

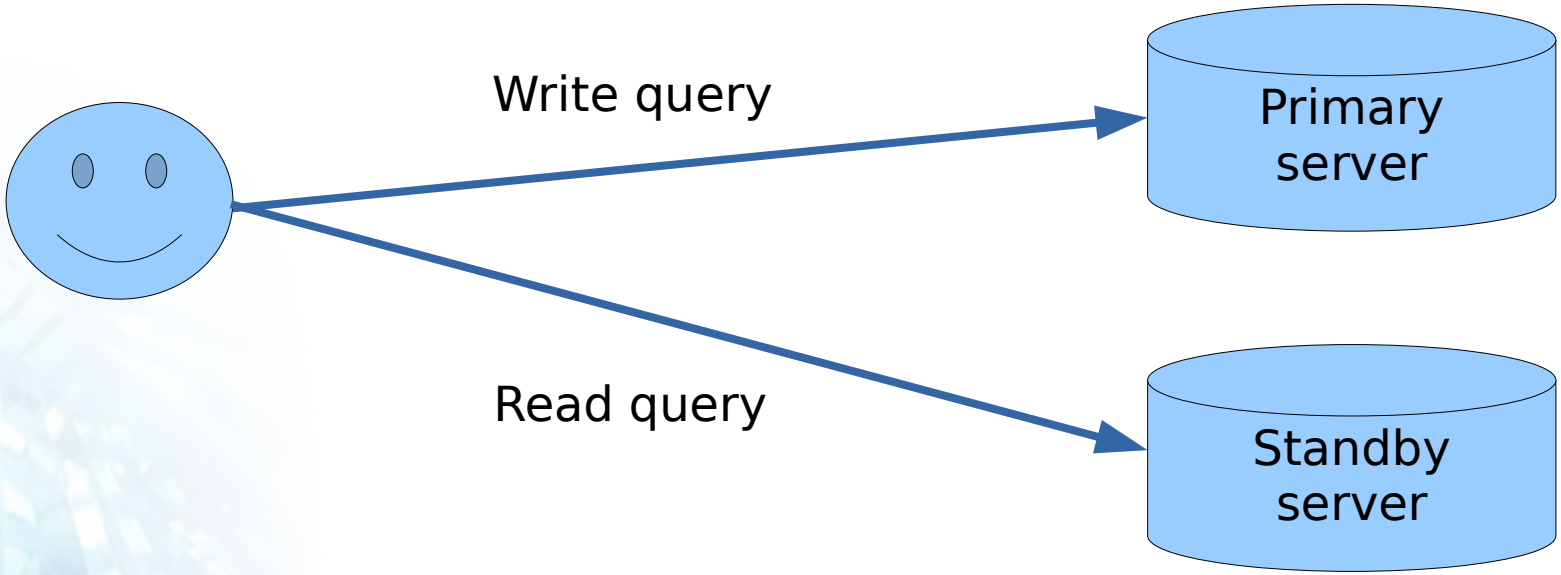


Copyright(c) 2020 SRA OSS, Inc. Japan

Pros and Cons of query routing methods

查询路由方法的优缺点

- Route write query to the primary while route read query to the standby
  将写查询路由到主数据库，将读查询路由到备用数据库
  - Can utilize resource of both primary and standby servers
    可以有效利用主服务器和备用服务器的资源
  - Possible to lower the load of the primary server
    可以降低主服务器的负载
  - Not easy task. Probably too complex for applications
    绝非易事，对于应用程序来说可能太复杂了



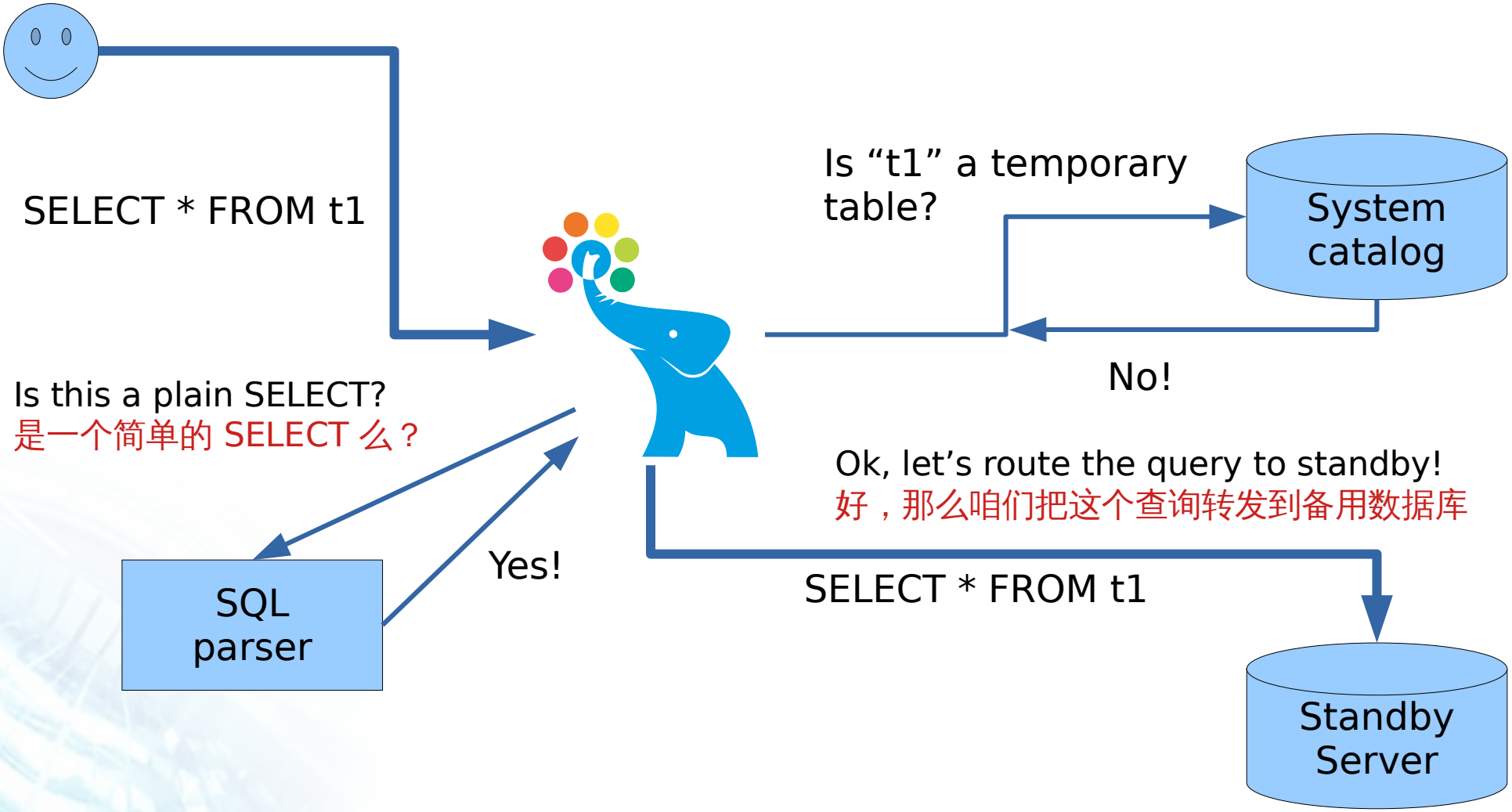
Copyright(c) 2020 SRA OSS, Inc. Japan

Why query routing is difficult?
为什么查询路由很困难？

- Some SELECT cannot be executed on standby 有些SELECT不能在备机查询
  - Accessing temporary tables 访问临时表
  - Calling functions doing writes 调用写函数
  - SELECT ... FOR UPDATE
- SERIALIZABLE transactions cannot be executed on standby
  SERIALIZABLE事务无法在备用数据库上执行
- LISTEN/NOTIFY cannot be executed on standby
- Many more corner cases...

Copyright(c) 2020 SRA OSS, Inc. Japan

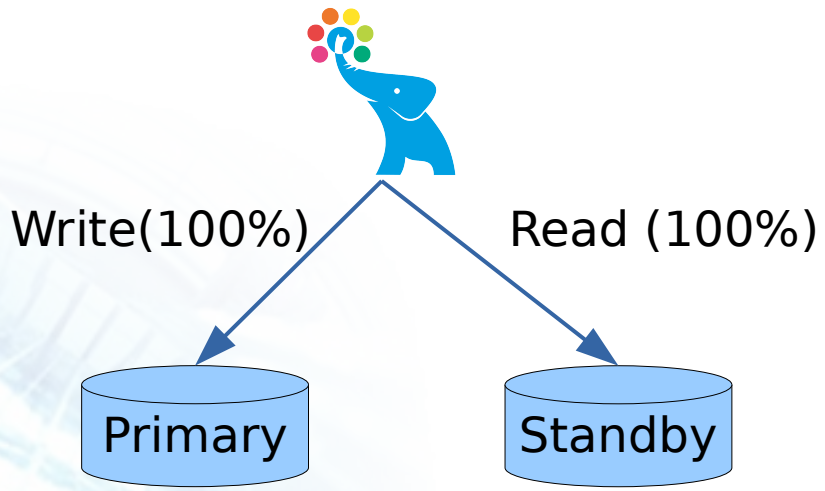
How Pgpool-II handles query routing?
Pgpool-II是如何处理查询路由的？



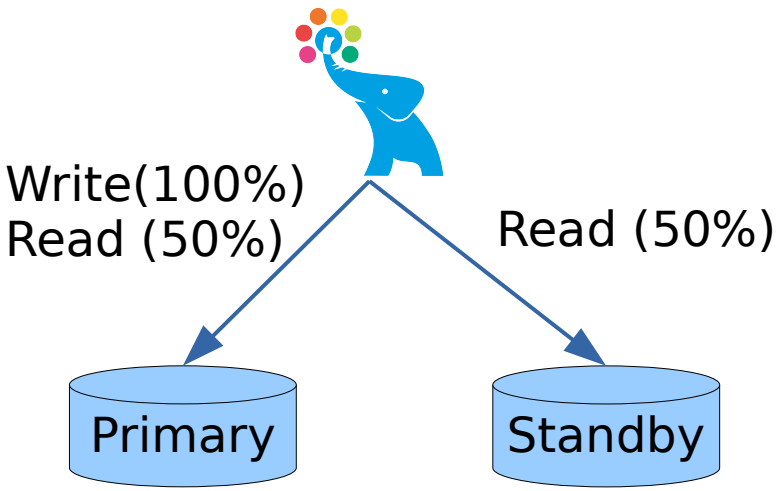
Copyright(c) 2020 SRA OSS, Inc. Japan
SRA OSS, INC.

Query routing and load balancing
查询路由与负载均衡

- Query routing just decides whether a query should be sent to primary or not
  查询路由只是决定是否将一个查询发送给主数据库
- On top of query routing, we can implement load balancing
  在查询路由基础上，我们可以实现负载均衡



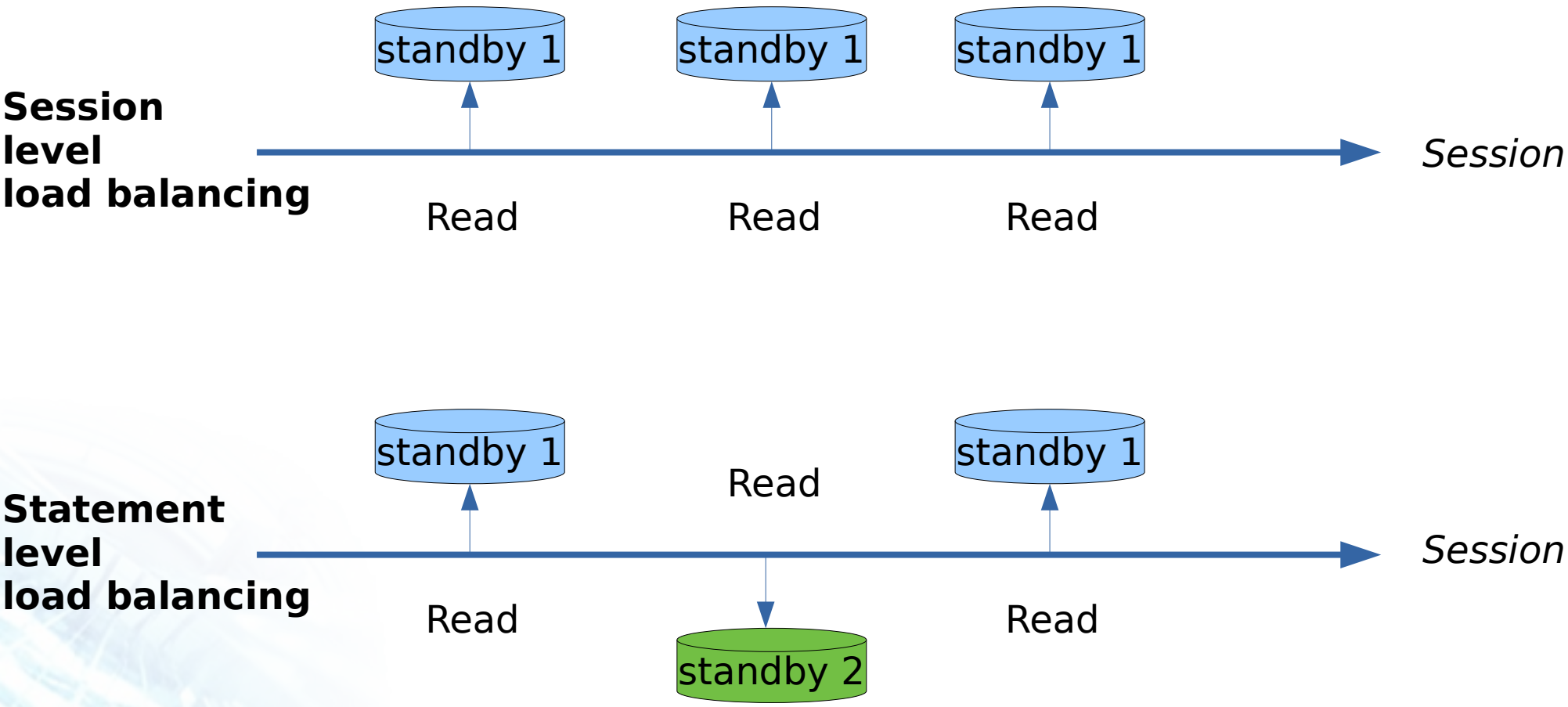
Query routing



Load balancing

Copyright(c) 2020 SRA OSS, Inc. Japan

Two methods of load balancing
两种负载均衡的方法



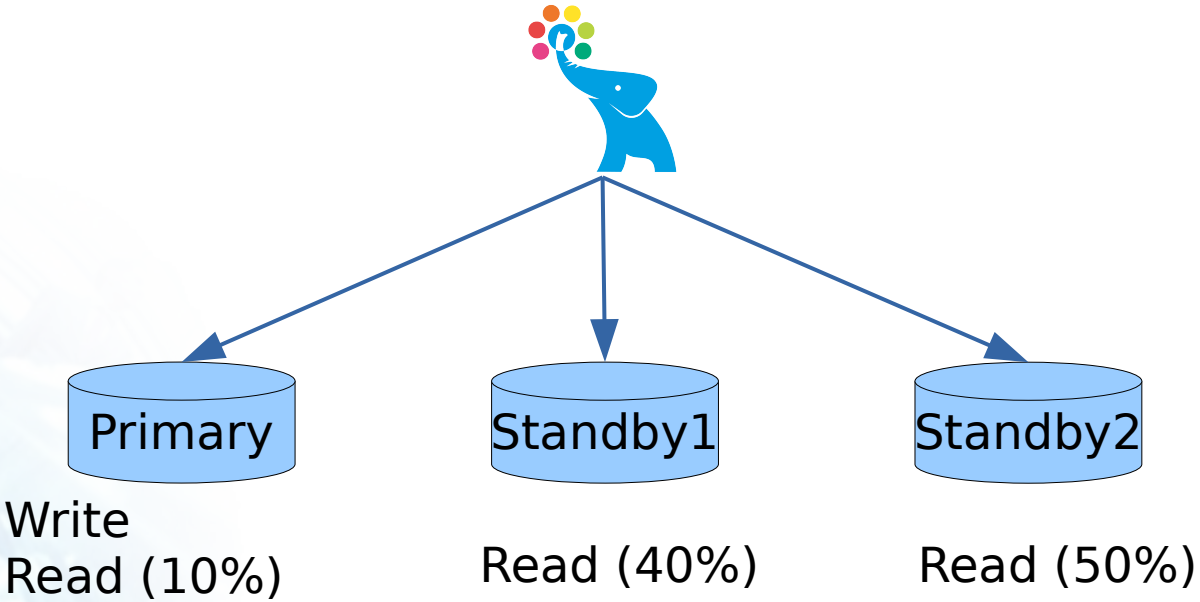
Copyright(c) 2020 SRA OSS, Inc. Japan

Basic load balancing

基本的负载均衡

- load_balance_mode = on
- backend_weight0 = 10
- backend_weight1 = 40
- backend_weight2 = 50

Specify
load balance
ratio

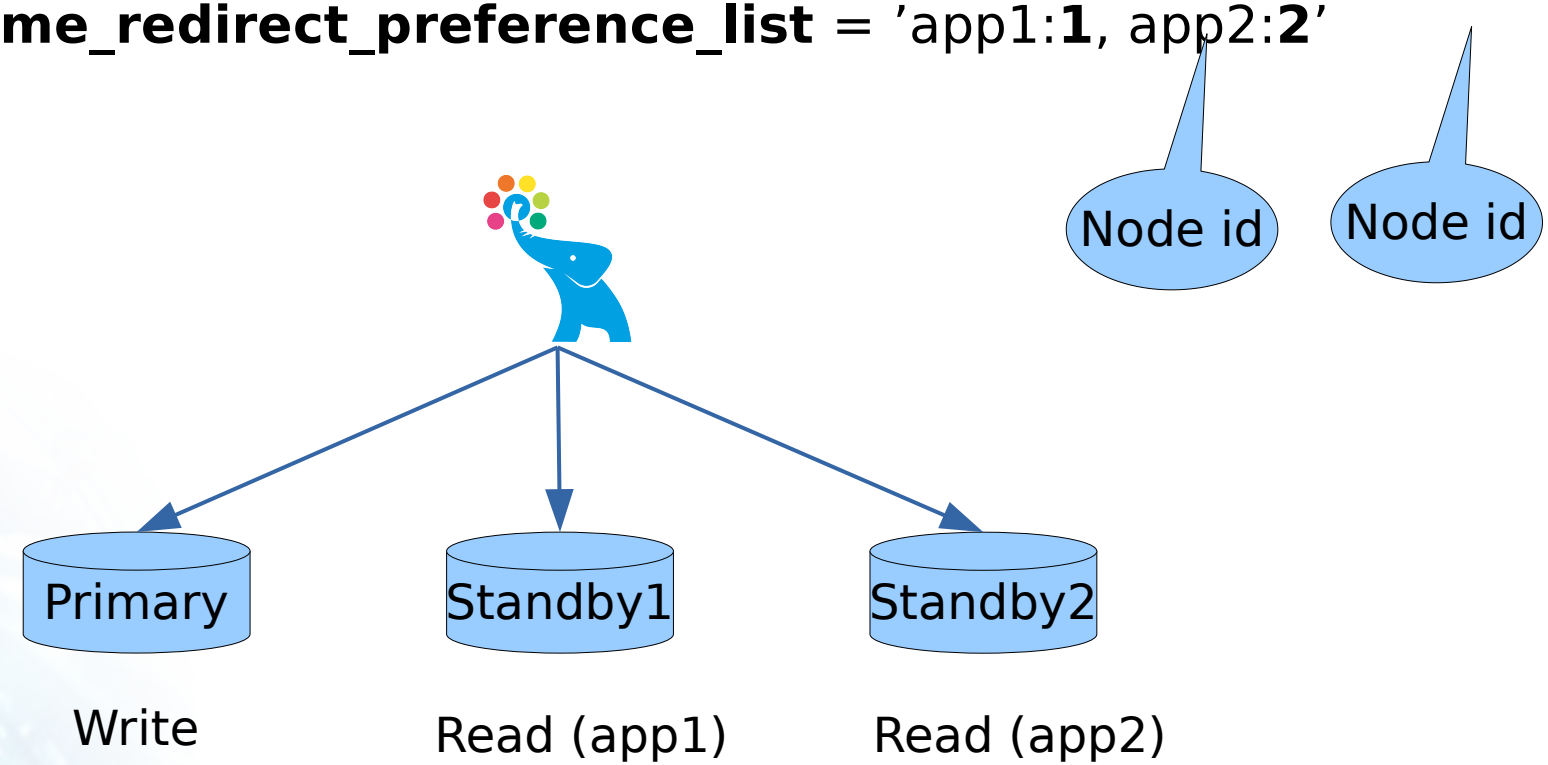


Copyright(c) 2020 SRA OSS, Inc. Japan

Advanced load balancing: load balance by application

高级负载均衡：通过应用程序实现

- Allow application to have dedicated standby server so that large applications **do not disturb each other**
  允许应用程序具有专用的备用服务器，以便大型应用程序不会互相干扰
- **app_name_redirect_preference_list** = ’app1:**1**, app2:**2**’

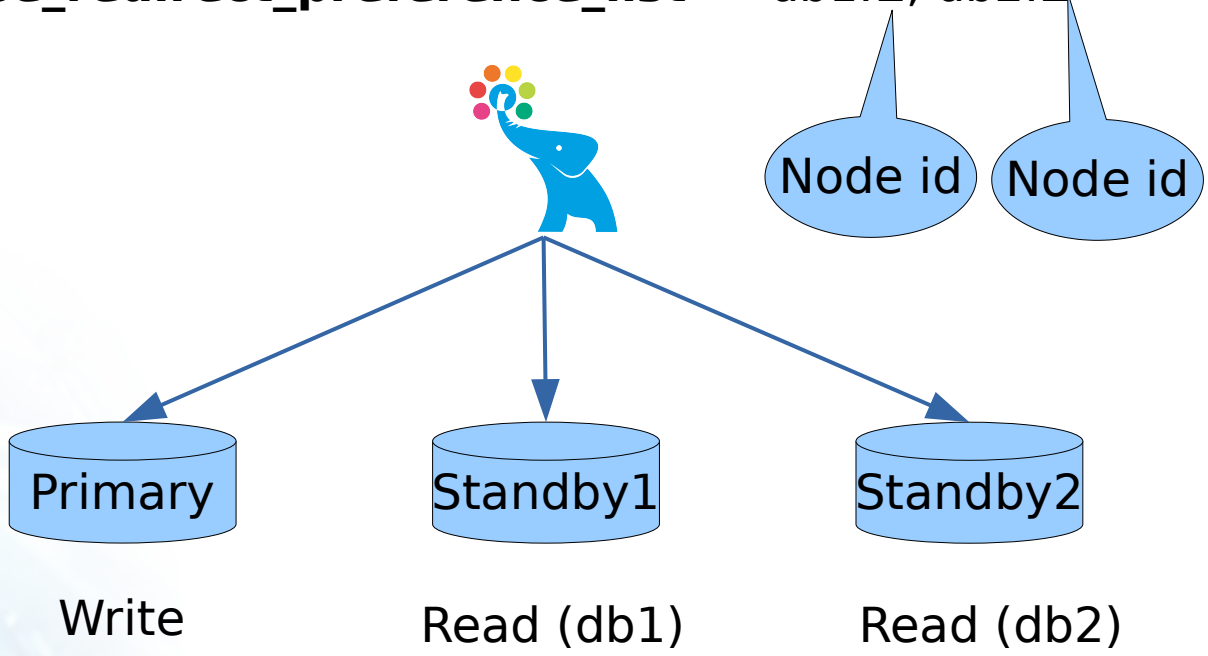


Copyright(c) 2020 SRA OSS, Inc. Japan

Advanced load balancing: load balance by database

高级负载均衡：通过数据库实现

- Allow database applications to have dedicated standby server
  允许数据库应用程序具有专用的备用服务器
- Applications using particular database **do not disturb each other**
  使用特定数据库的应用程序来确保不会互相干扰
- **database_redirect_preference_list** = ’db1:**1**, db2:**2**’



Copyright(c) 2020 SRA OSS, Inc. Japan

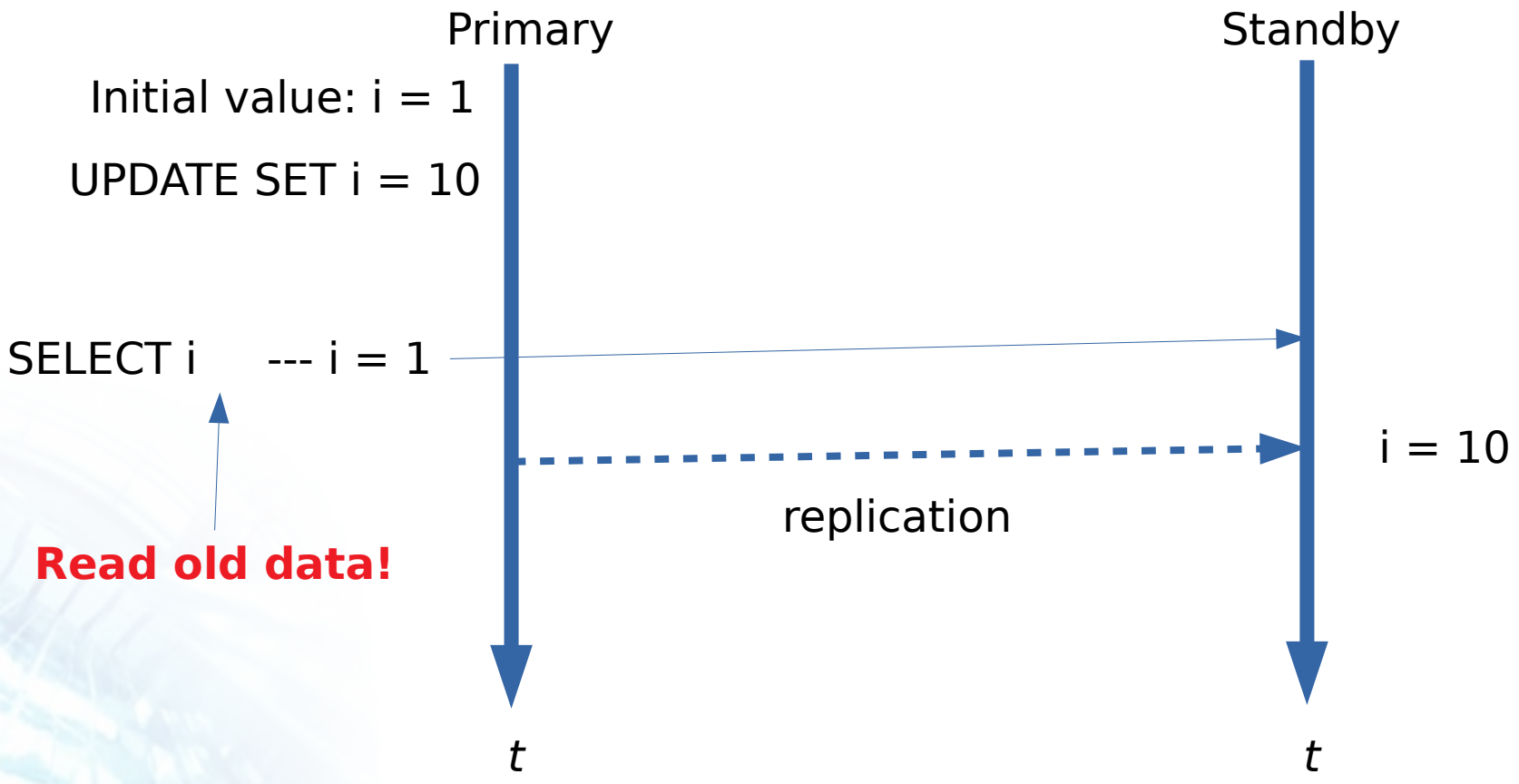
Routing particular query to primary
将特定查询路由到主数据库

- Route any SQL including string “my_table” to primary
  将任何包含字符串“ my_table”的SQL路由到主数据库
  - **black_query_pattern** = ’.*my_table.*’
- Specify functions which do writes
  指定执行写的函数
  - **black_function_list** = 'nextval,setval,lastval,currval'
  - ’SELECT nextval()’ will be routed to primary
- Insert **SQL comment** in front of SQL
  在SQL前面插入SQL注释
  - /*NO LOAD BALANCE*/SELECT * FROM t1

Copyright(c) 2020 SRA OSS, Inc. Japan
SRA OSS, INC.

Dealing with replication delay
处理复制延迟

- Load balanced query may read old data from standby
  负载均衡查询可能会从备用数据库读取旧的数据



Copyright(c) 2020 SRA OSS, Inc. Japan

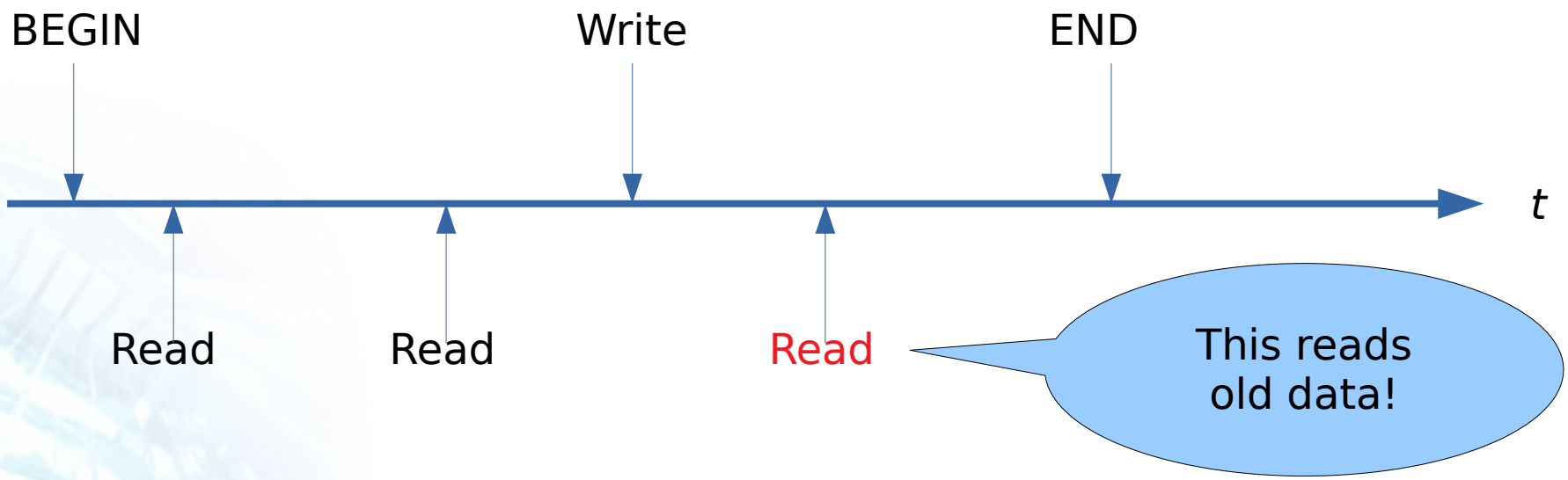
Dealing with replication delay
处理复制延迟

- **delay_threshold** parameter
  - Replication delay is continuously monitored by Pgpool-II
    由Pgpool-II负责持续监视复制延迟
  - If the **delay > delay_threshold**, then the standby node will not be chosen as the load balancing node (the target of load balance)
    如果delay> delay_threshold，则不会选择备用数据库作为负载均衡节点（负载均衡的目标）
  - You can use synchronous replication of PostgreSQL but it can **hurt the performance** of writing to primary
    您可以使用PostgreSQL的同步复制，但是会影响主数据库的写性能

Copyright(c) 2020 SRA OSS, Inc. Japan
SRA OSS, INC.

Load balancing in a transaction
事务中的负载平衡

- Typically applications use an explicit transaction to perform set of tasks using multiple SQL
  通常，应用程序使用显式事务来使用多个认为执行一组任务
- Applications expect to read the newest data, but it is impossible to read the new data from standby server because the transaction is not committed yet
  应用程序希望读取最新数据，但是由于尚未提交事务，因此不可能从备用服务器读取新数据。
- But always reading from primary causes high load 但始终从主服务器读取会导致高负载
- **There's no "one-size-fits-all" solution for this problem**



Copyright(c) 2020 SRA OSS, Inc. Japan

Load balancing in a transaction
事务中的负载平衡

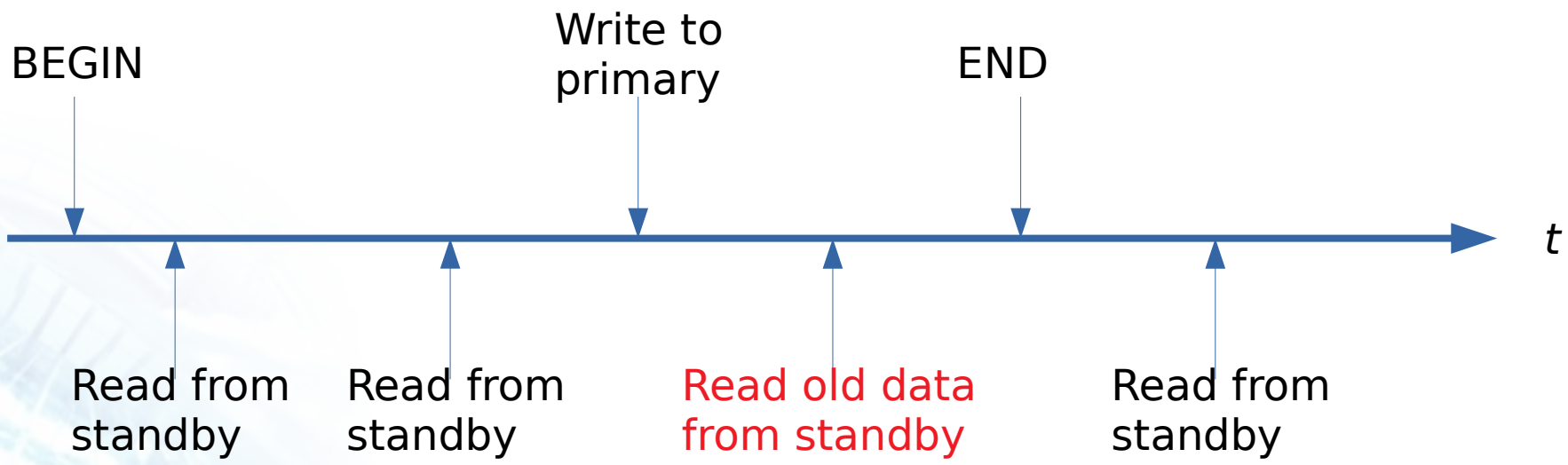
- **disable_load_balance_on_write** provides a way to control the balance between performance and reading new data
  disable_load_balance_on_write 提供了一种方法来控制性能与读取新数据之间的平衡

	Read new data	Performance
off	X	◎
transaction	△	○
trans_transaction	○	△
always	◎	X

Copyright(c) 2020 SRA OSS, Inc. Japan
SRA OSS, INC.

Load balancing in a transaction
事务中的负载平衡

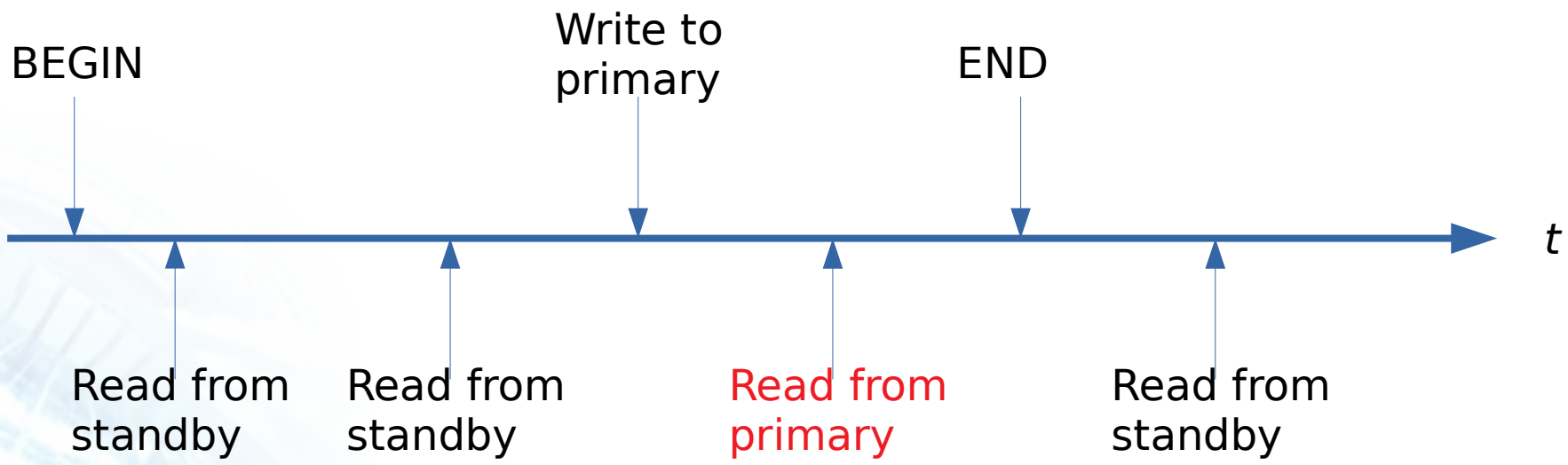
- **disable_load_balance_on_write = off**
  - Always allow to load balance = best performance
    总是允许负载平衡=最佳性能
  - May read old data from standby 会从备机读到老数据



Copyright(c) 2020 SRA OSS, Inc. Japan

Load balancing in a transaction
事务中的负载平衡

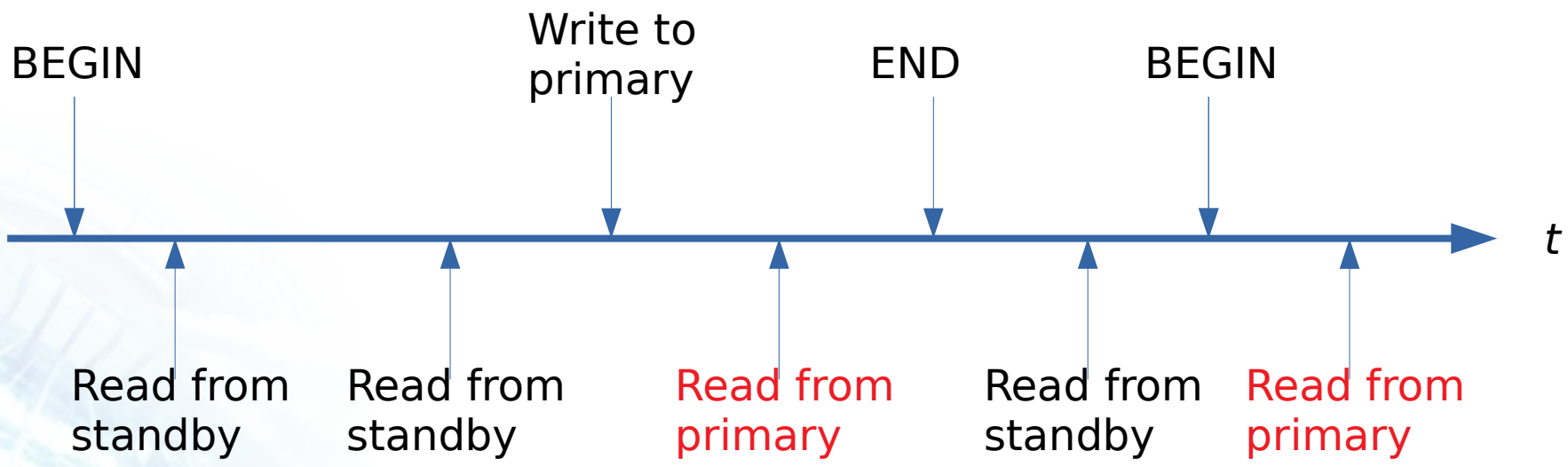
- **disable_load_balance_on_write = transaction**
  - Allow to load balance until write query issued 允许负载平衡，直到发出写查询为止
  - Always read new data if within a transaction 总是在事务中读取新数据
  - Good balance of performance and read new data 良好的性能平衡和读取新数据



Copyright(c) 2020 SRA OSS, Inc. Japan

Load balancing in a transaction
事务中的负载平衡

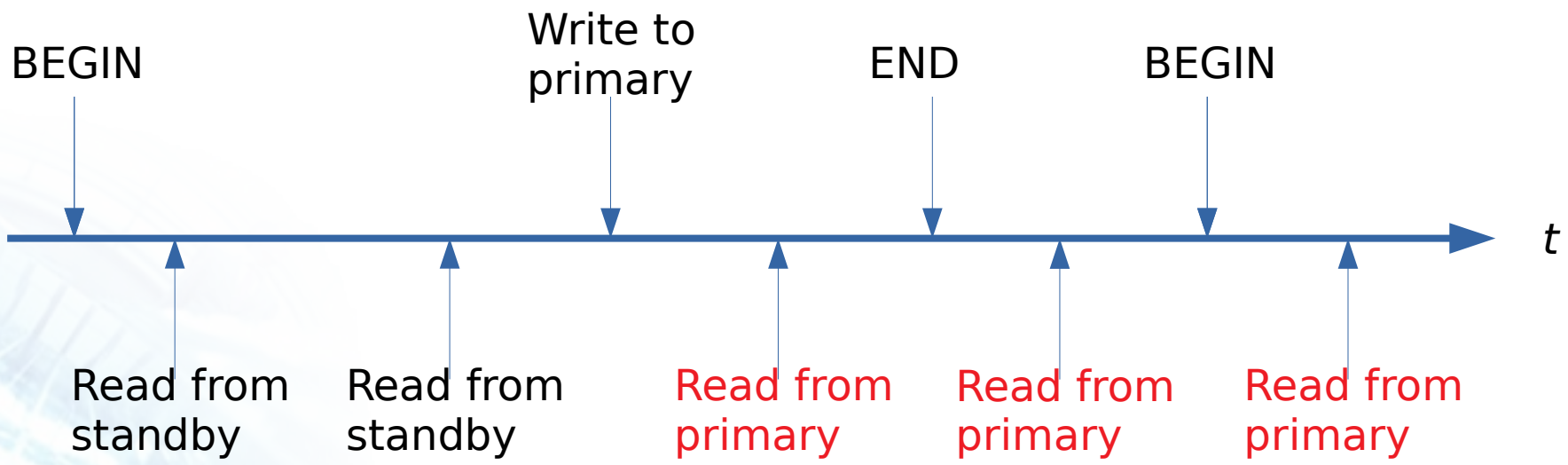
- **disable_load_balance_on_write = trans_transaction**
  - Even in the next transaction read from primary if a write query has been issued in the former transaction 即使在下一个事务中，如果在前一个事务中发出了写查询，则从主服务器读取该事务
  - Less chance to read old data than ‘transaction’ with more performance penalty
    与“transaction”相比，读取旧数据的机会更少，但会降低性能



Copyright(c) 2020 SRA OSS, Inc. Japan

Load balancing in a transaction

- **disable_load_balance_on_write = always**
  - Always read from primary once write query is issued 发出写查询后总是从主服务器读取
  - This is the safest mode with the largest performance penalty 这是最安全但性能损失最高的模式



Copyright(c) 2020 SRA OSS, Inc. Japan

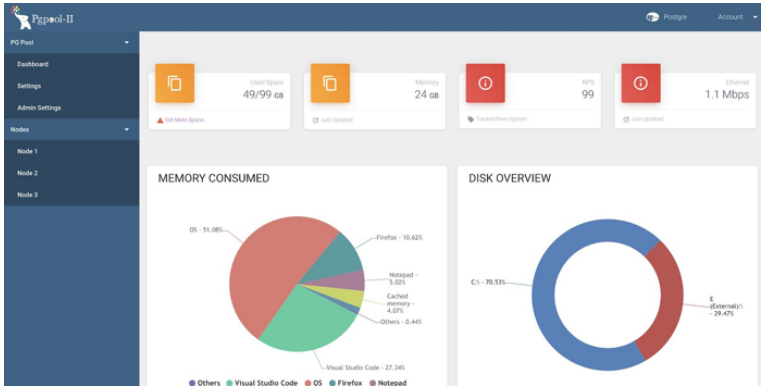
Summary 总结

- **Query routing and load balance are important techniques** for a clustering system using streaming replication
  查询路由和负载平衡是使用流复制的群集系统的重要技术
- Pgpool-II offers **intelligent ways to route queries.**
  Pgpool-II提供了路由查询的智能方法
- Pgpool-II also offers various parameters to control load balancing.
  Pgpool-II还提供各种参数来控制负载平衡
- Applications can use these parameters to utilize cluster resources with **minimum modifications to the application code**
  应用程序可以使用这些参数以及最少的程序代码修改来利用群集资源

Copyright(c) 2020 SRA OSS, Inc. Japan
SRA OSS, INC.

About next release 关于下一版本

Next release (4.2) will come in September or October
下一版本将会在今年9月或这10月面世



New features include 包括的新功能:

- New GUI tool 新的GUI工具
- Ready for PostgreSQL 13 支持PG13
- New replication mode: Snapshot isolation mode which guarantees atomic visibility 新的复制模式：快照隔离模式，可确保原子可见性
- Statistics of health check 健康检查统计
- Easy and consistent configuration of watchdog 看门狗的简单一致的配置
- Batch user registration 批量用户注册

Copyright(c) 2020 SRA OSS, Inc. Japan

References/URLs 参考链接

- Pgpool-II official home page 官网
  - https://pgpool.net
- Pgpool-II documentations 文档
  - https://www.pgpool.net/docs/latest/en/html/
- Community support 社区支持
  - https://pgpool.net/mediawiki/index.php/Mailing_lists
  - https://pgpool.net/mediawiki/index.php/Bug_tracking_system

Copyright(c) 2020 SRA OSS, Inc. Japan
SRA OSS, INC.

Thank you!

Copyright(c) 2020 SRA OSS, Inc. Japan