




2016 杭州·云栖大会
THE COMPUTING CONFERENCE

Greenplum上云与优化

— ApsaraDB for Greenplum介绍

2016
The Computing Conference



主办单位:  杭州

 Alibaba Group
阿里巴巴集团

战略合作伙伴: 

张广舟 (明虚)
阿里云高级专家



目 录
content

- ApsaraDB for GP的定位
- ApsaraDB for GP的内核优化
- 未来的规划



ApsaraDB for GP的定位



ApsaraDB for GP的定位

GP的优势？

与其他技术的对比？

为什么上云？

ApsaraDB for GP的定位

MPP + 列存压缩

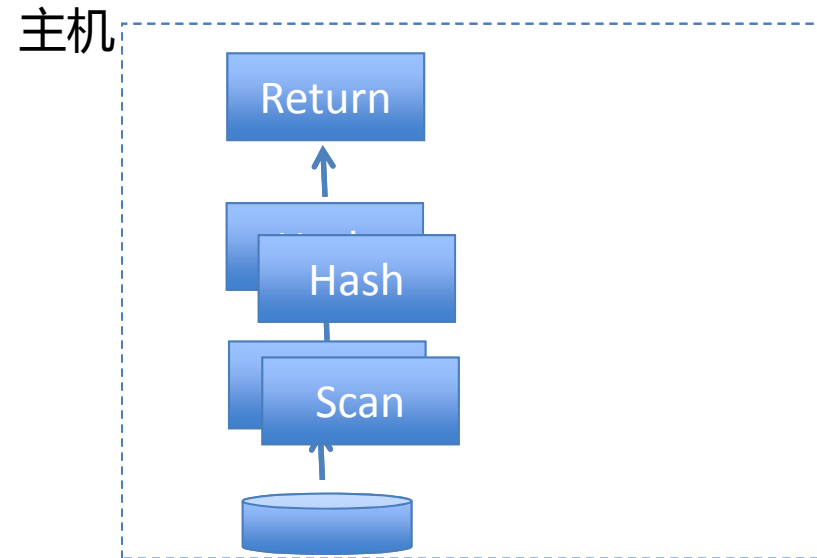
ApsaraDB for GP = 复杂SQL + 查询优化器

本地高效存储 + 高速网络 + 预置稳定资源

= 简单、高效解决大数据分析需求

GP vs. RDS? MPP处理举例

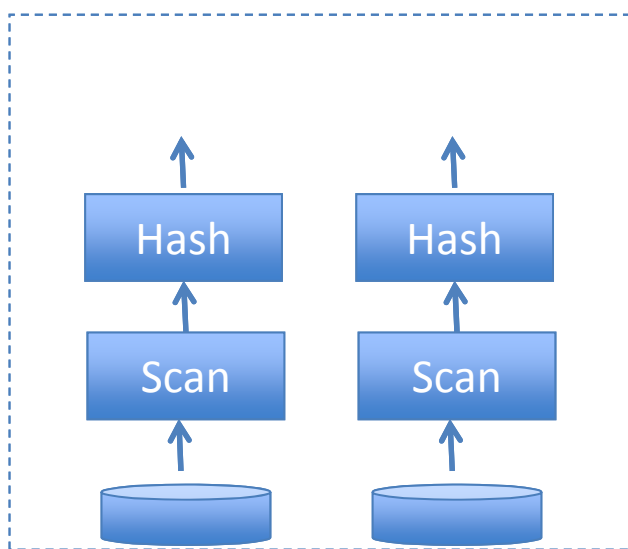
Select count(*) from customer group by city



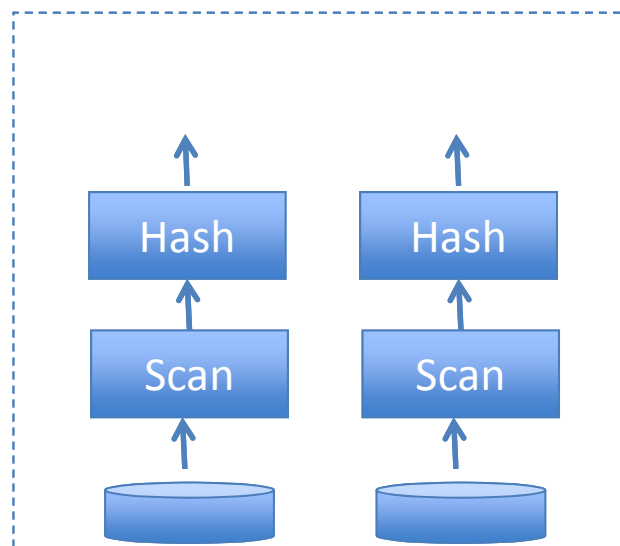
GP vs. RDS? MPP处理举例

Select count(*) from customer group by city

主机

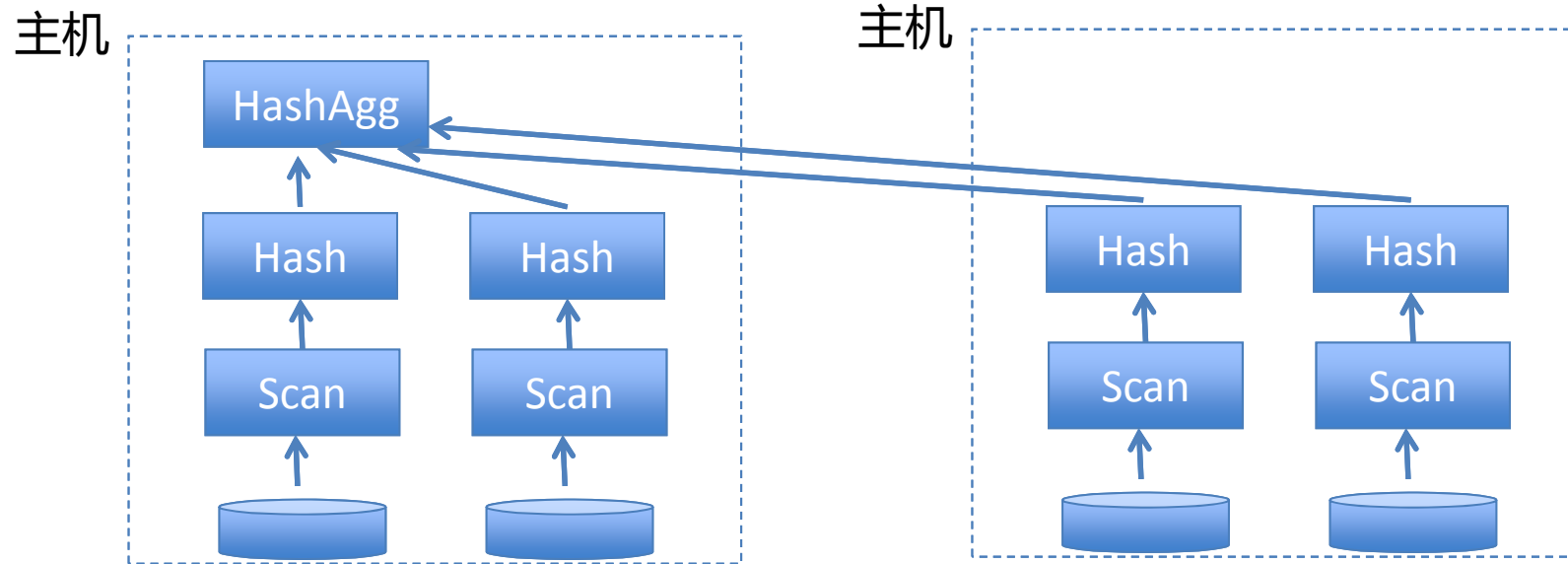


主机



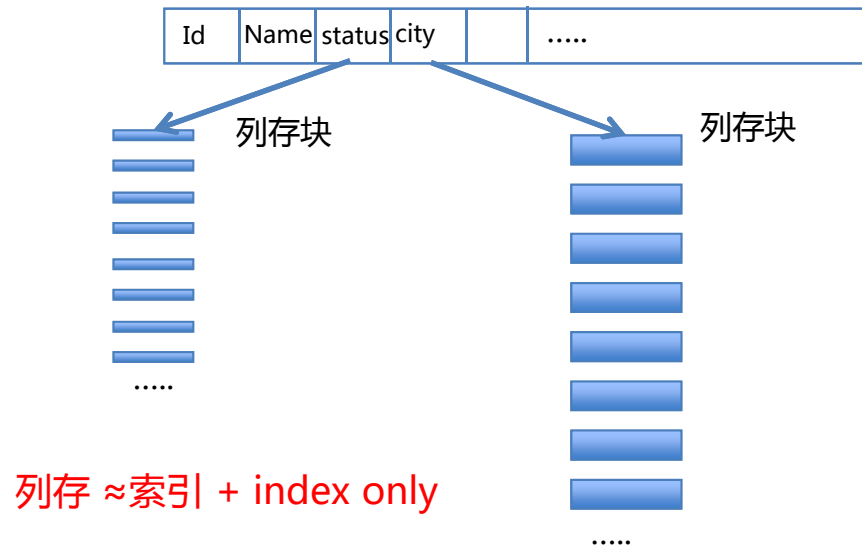
GP vs. RDS? MPP处理举例

Select count(*) from customer group by city



GP vs. RDS? 列存与压缩原理举例

Select count(*) from customer where status = valid group by city



GP vs. Hadoop?

Orca优化器

SQL Runtime

本地存储

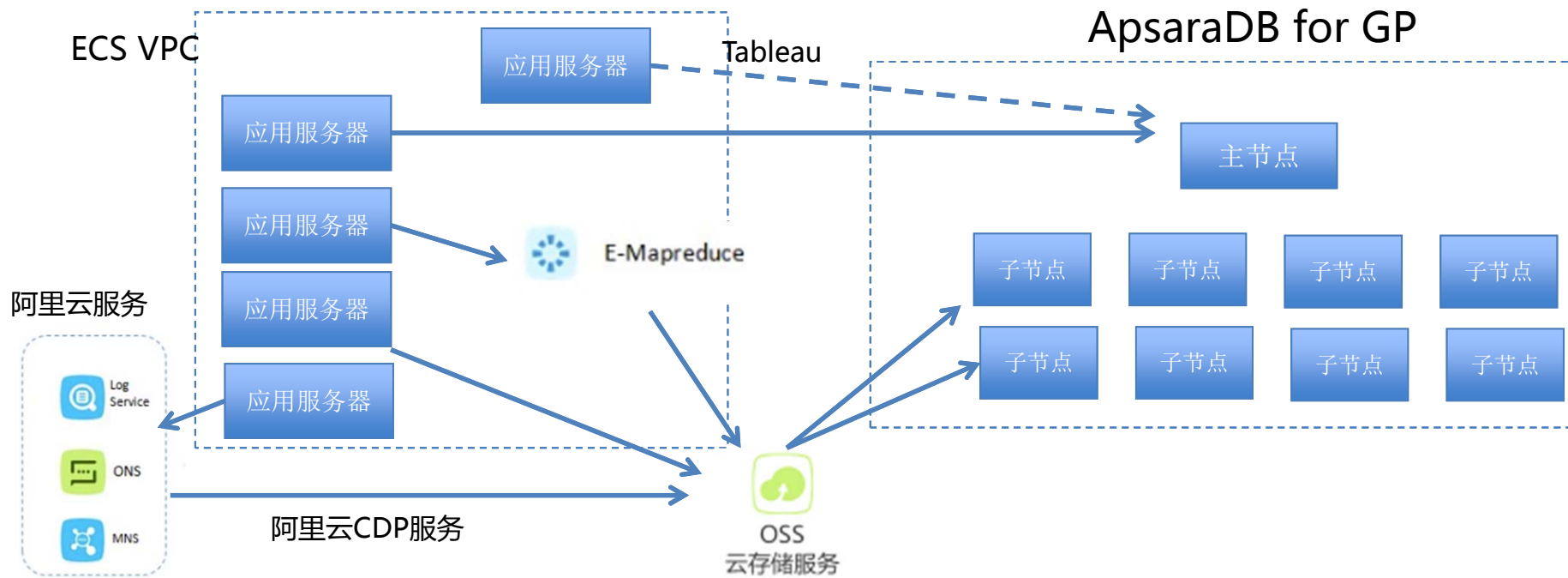
> 5-30倍的性能优势

ApsaraDB for GP vs. AWS Redshift?

“有史以来卖的最好的云服务”

对比项目	ApsaraDB for Greenplum(SSD/SATA)	AWS RedShift
外部表	支持OSS外部表	无此设计
地理信息支持	支持 (自带PostGIS)	不支持
分区表支持	支持	不支持
数据类型	支持所有PostgreSQL 9.0以下类型	支持11个PostgreSQL类型
横向扩容	支持	支持, 需要停机 (数分钟)
索引	支持 (B-tree、Bitmap)	不支持

推荐应用架构



ApsaraDB for GP的定位

MPP + 列存压缩

ApsaraDB for GP = 复杂SQL + 查询优化器

本地高效存储 + 高速网络 + 预置稳定资源

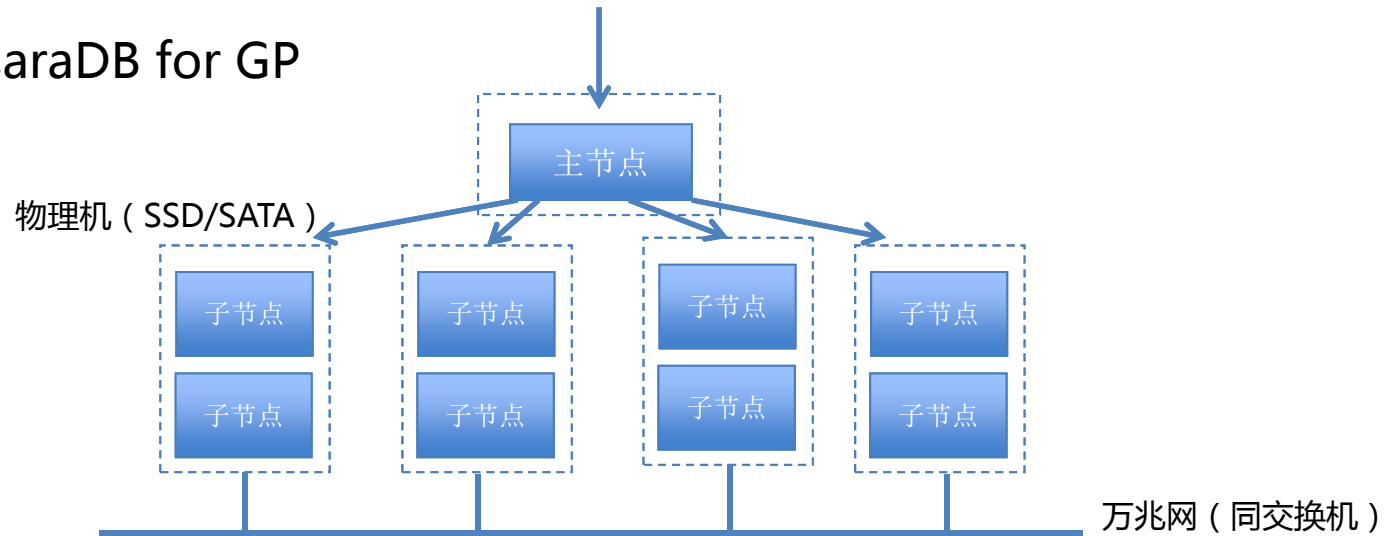
= 简单、高效解决大数据分析需求



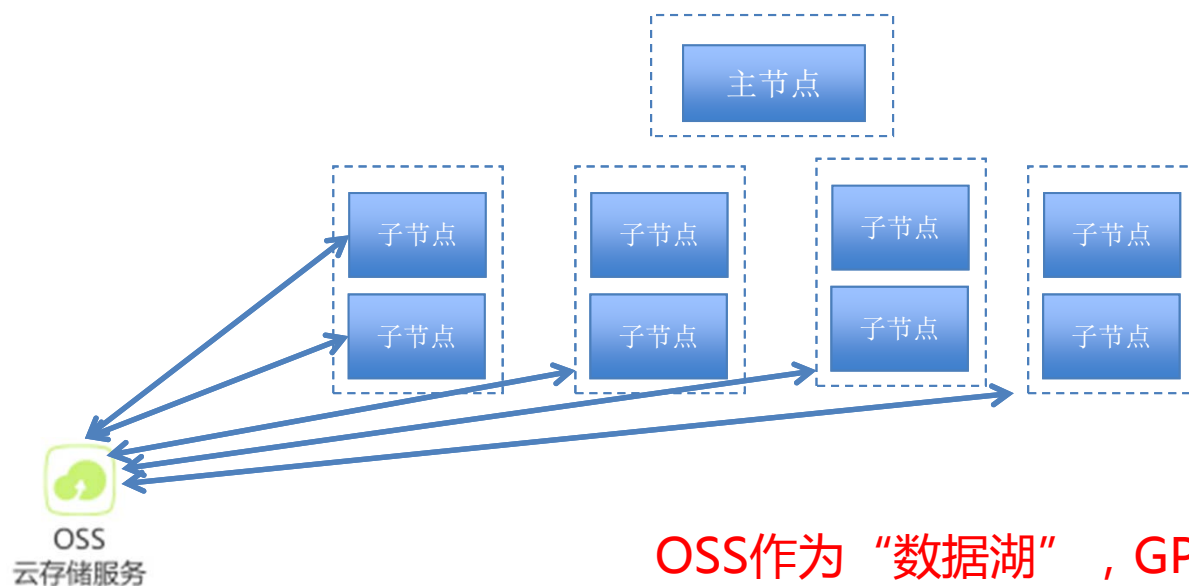
ApsaraDB for GP的内核优化

ApsaraDB for GP架构

ApsaraDB for GP



支持OSS外部表读写



OSS作为“数据湖”，GP作为分析引擎

```
Insert into T_GP  
Select my_udf(name)  
from T_OSS  
Where ts > 2016
```

```
Insert into T_OSS  
Select count(*) from T_GP  
Group by city
```




支持外部扩展已插件形式管理

支持插件创建的语法

CREATE EXTENSION

DROP EXTENSION

Patch已提交社区

HyperLogLog支持

```
create extension hll;
```

```
create table access_date (acc_date date unique, userids hll);
```

```
select #userids from access_date where acc_date=current_date;
```



目前支持的插件

MADLib

PL/Java

Pgcrypto

Fuzzstringmatch

PostGIS

Orca

hll

完善的JSON数据类型支持

Create table users(id int, info **json**)

Select info->name from users

Operator	Right Operand Type	Description	Example	Example Result
->	int	Get JSON array element (indexed from zero, negative integers count from the end)	'[{"a": "foo"}, {"b": "bar"}, {"c": "baz"}]'::json->2	{"c": "baz"}
->	text	Get JSON object field by key	'{"a": {"b": "foo"}}'::json->'a'	{"b": "foo"}
-->	int	Get JSON array element as text	'[1,2,3]'::json-->2	3
-->	text	Get JSON object field as text	'{"a":1,"b":2}'::json-->'b'	2
#>	text[]	Get JSON object at specified path	'{"a": {"b": {"c": "foo"}}}'::json#>'a,b'	{"c": "foo"}
#>	text[]	Get JSON object at specified path as text	'{"a": [1,2,3], "b": [4,5,6]}'::json#>>'a,2'	3

用户权限定制

为RDS用户提供一个特有的“superuser”角色

```
Create role my_user rds_superuser;
```

```
Alter role my_user nords_superuser;
```

为该用户放开部分superuser的权限

- 查看其它用户数据

- 查看所有连接信息

- 杀连接

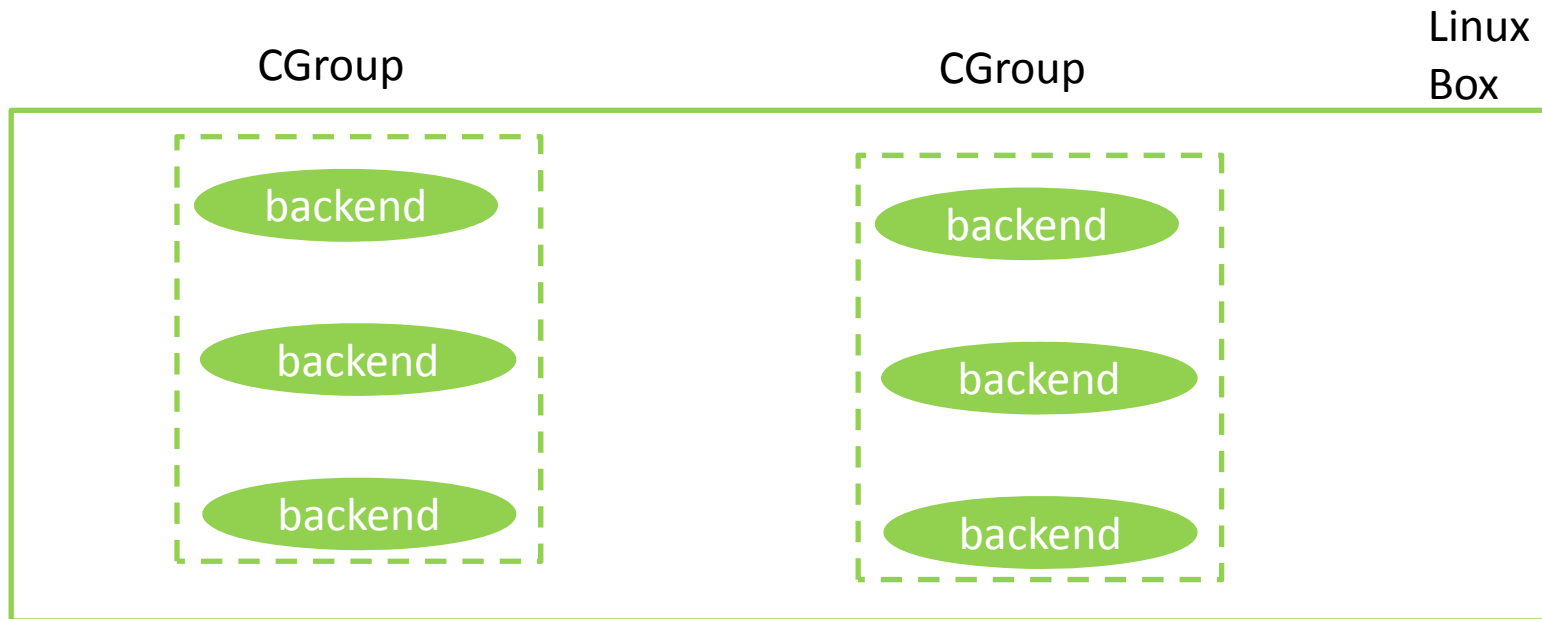
- 创建和删除插件

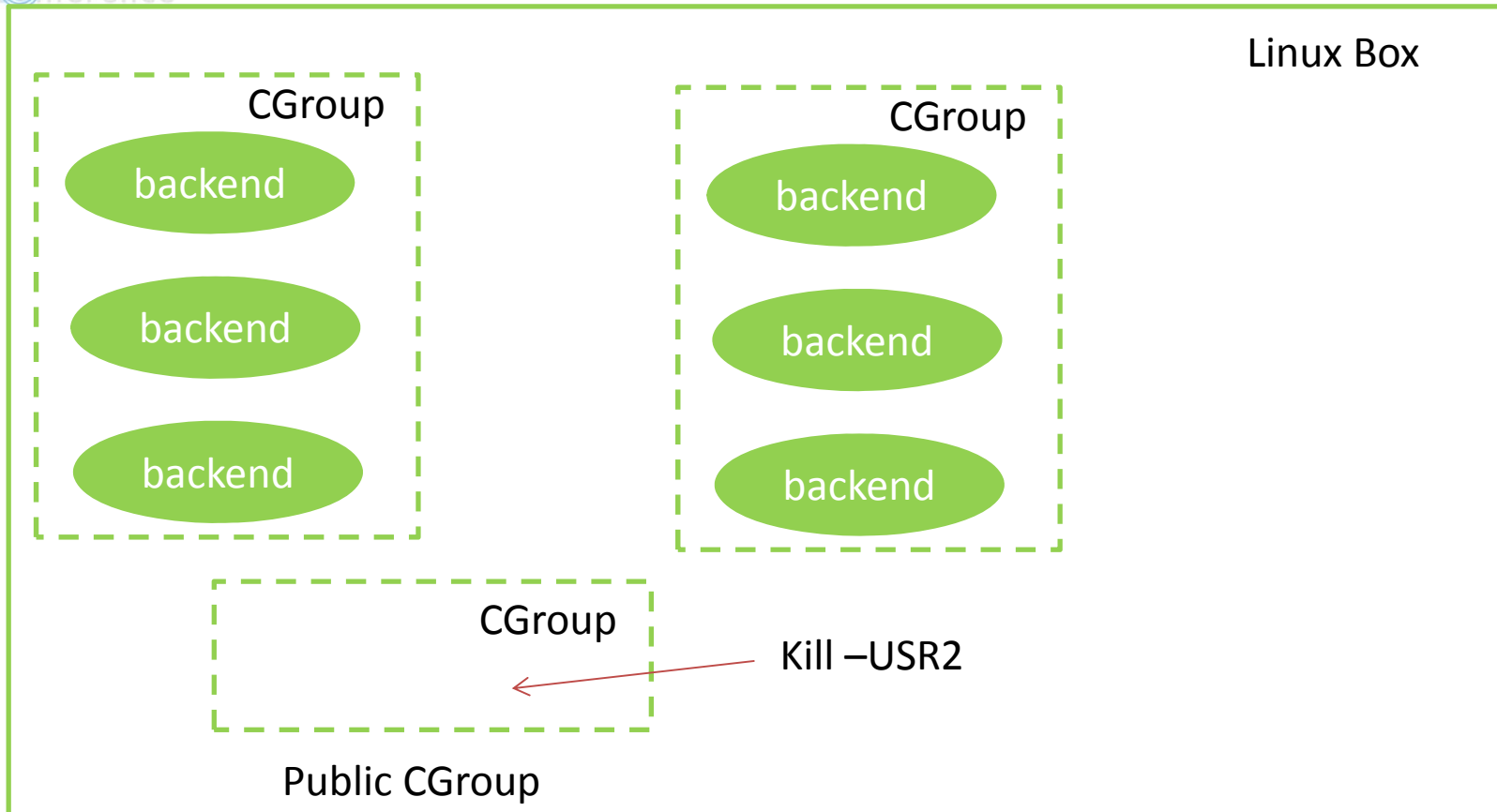
解决OOM问题

实例的OOM有时很频繁，同时OOM很难提前监控

我们的办法

利用外部脚本监控cgroup中的内存统计
发生内存水位较高时，将实例移入公共
cgroup；同时发出cancel query信号给内核
水位下降时移回实例的cgroup







未来规划

满足客户需求！

列存优化 – 某些场景下提高数倍性能

CPU优化 – 对GP执行器的静态编译优化

流式备份 – 实时的物理备份

社区参与 – 成为社区积极的贡献者

20 The
16 Computing
Conference
THANKS

 2016 杭州·云栖大会
THE COMPUTING CONFERENCE

