



Gdevops

全球敏捷运维峰会

Oracle x86化企业建设方案

演讲人：罗春

Content/内容

1

数据库基础架构

2

高可用数据库容灾方案

3

海量数据库自动化运维

影响数据库性能的要三要素

计算

数据运算能力-CPU主频和核数决定

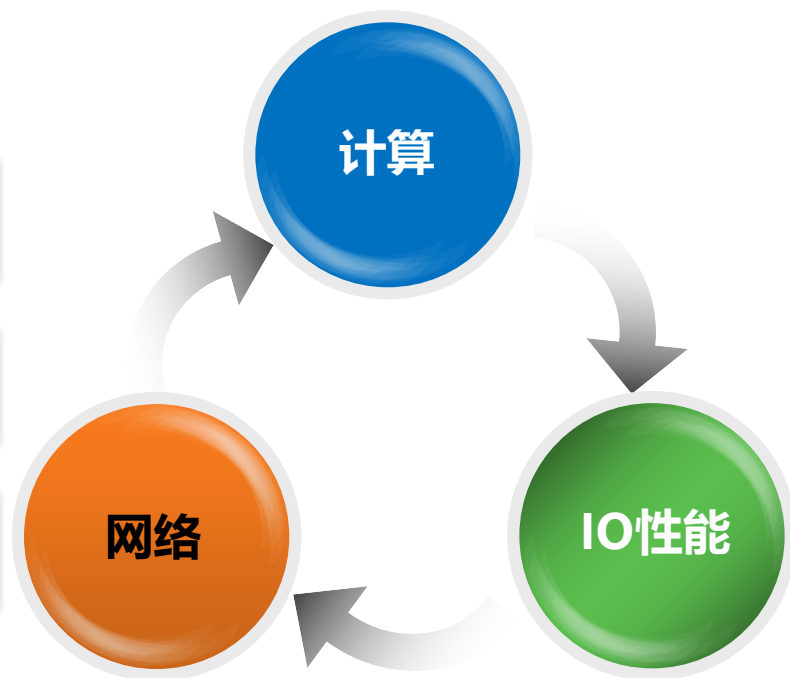
网络

数据从存储介质跑到CPU的关键路径—网络传输、南北桥桥接, CPU Cache

IO性能

IO设备自身的性能极限--IO延迟与IO吞吐量。

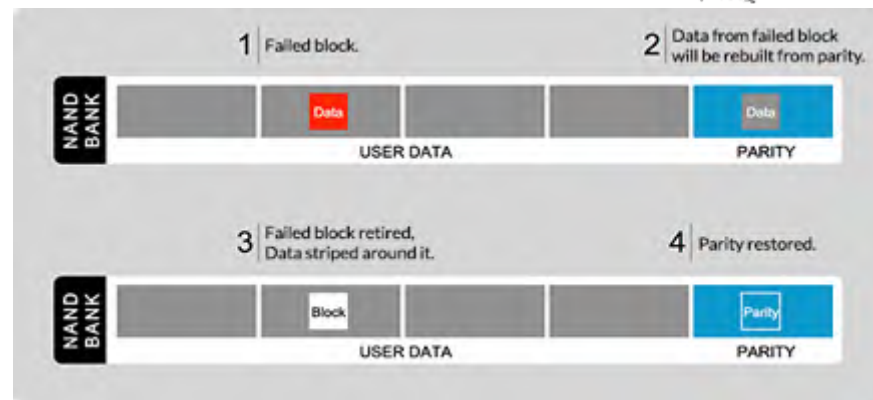
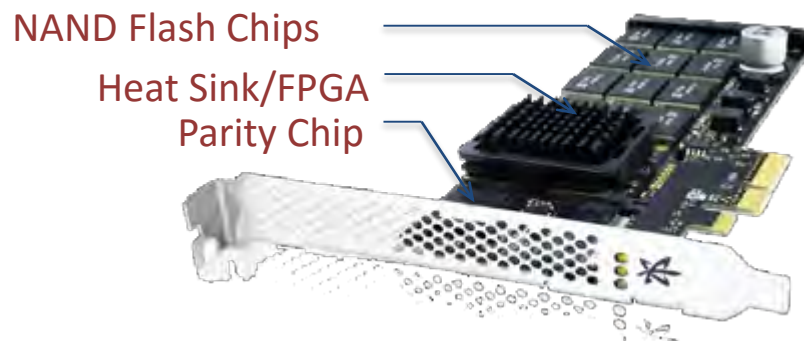
三管齐下, 极致提升数据库整体性能.



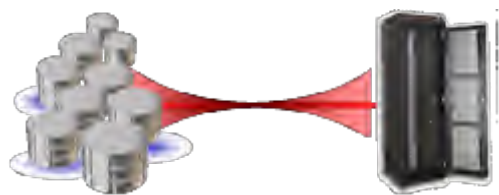
革命性的高速Flash IO设备

- 专利技术的磨损均衡 (wear-leveling) 算法
- 容量超配 (Over Provisioning)
 - Reserve space for handling individual cell/LEB death
 - Reserve space is adjustable if higher write performance is needed
- 更高密度的ECC算法进一步提升数据的可靠性以及设备的使用寿命
- N+1 存储块级的数据保护 (think RAID protection on card)
- 意外掉电情况下数据完全不丢失 (OEM认证)

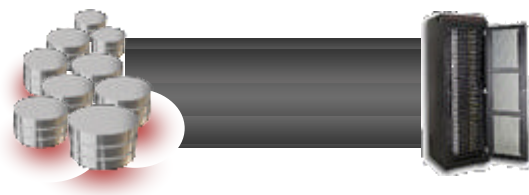
注：所列性能参数来自一线品牌厂商阶段性公开数据，仅供参考



数据传输架构的变革



传统架构 FC通道



高速 RDMA通道

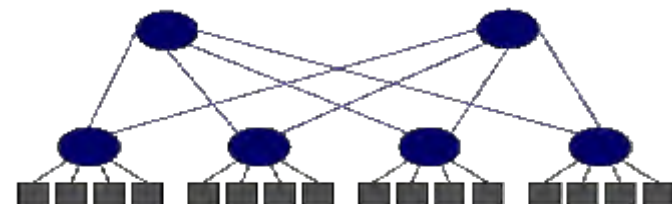
高性能、低延时的Infiniband网络互联，打破数据传输瓶颈

•更高的带宽：FC 8-16Gb/S → IB 40-56Gb/S

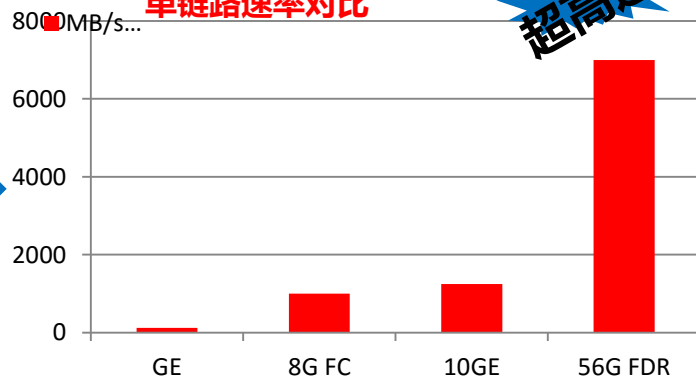
•更快的数据传输： →  1000X

Infiniband网络特点：

- 40-56Gbps InfiniBand，超高速互联。
- 近似无阻塞通信网络，数据交换无瓶颈
- ns级通信时延，计算存储信息及时传递
- 无损网络QOS，数据传送无丢失
- 主备端口多平面通信，冗余通信无忧



单链路速率对比



超高速

关键业务支撑 - x86分布式集群基础架构

服务器x86化

- 标准x86 Linux服务器
- 高性价比

分布式存储

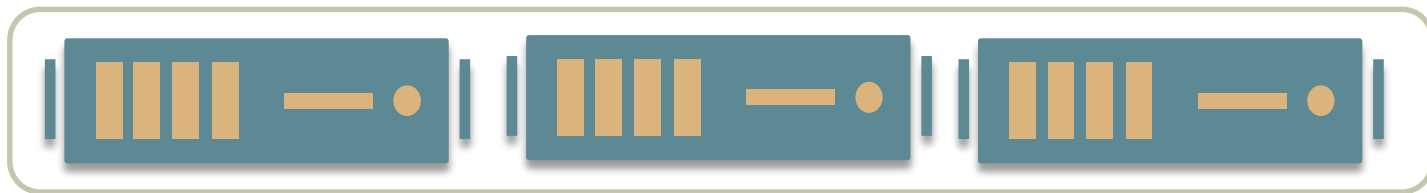
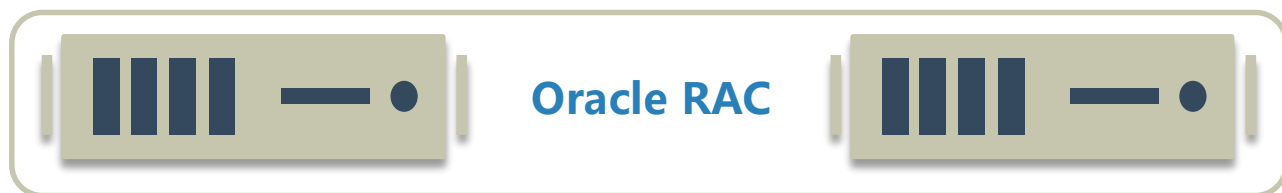
- x86分布式存储
- 消除单点故障与性能瓶颈

高性能技术

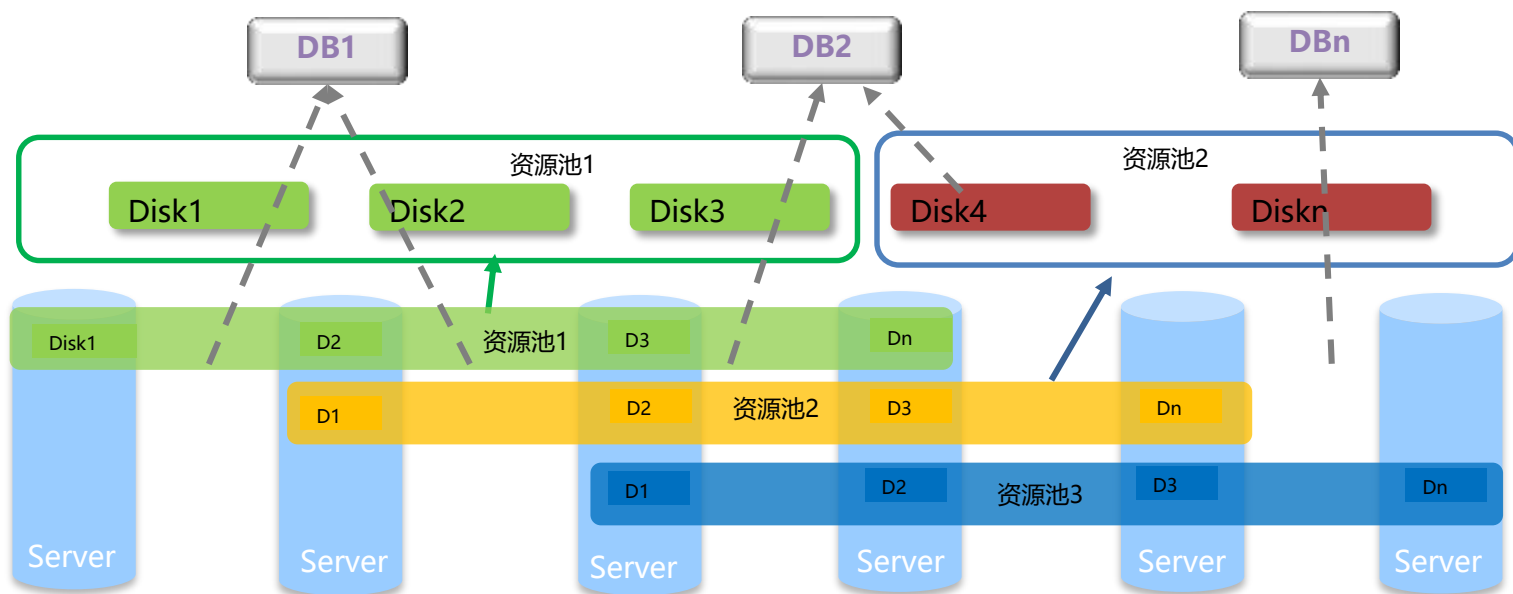
- IB网络, 闪存IO
- 基于RDMA的软件技术

全冗余架构

- 计算, 存储, 网络冗余
- 跨机器的分布式镜像



分布式存储池-x86数据库存储资源池



资源池: 类似于SAN的RAID组概念, 与RAID相比, 其优点是:

- 存储池化: 存储节点形成存储池, 按需划分存储资源
- 存储共享: 到lun级别的存储资源共享, 允许lun在多个数据库间切换使用

券商手机APP-上亿级交易资产查询



实时资产



市场行情

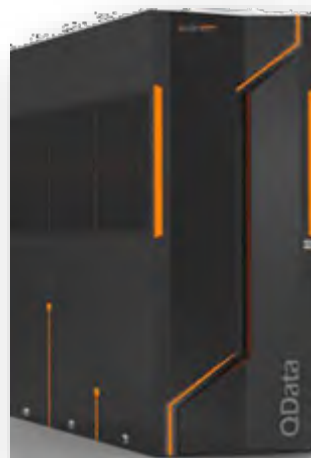


委托交易



股票资讯

券商手机APP-上亿级交易资产查询



小型机存储架构 核心交易库

X86分布式集群 同步查询库

数据库压力

- 15000查询SQL/秒
- 50000 IO请求/秒
- CPU利用率40%

业务压力

- 1000万涨乐乐友
- 交易数据实时同步
- 高峰期段1.2亿次
乐友资产刷新

Content/内容

1

数据库基础架构

2

高可用数据库容灾方案

3

海量数据库自动化运维

x86化面临的备份问题

- **存在数据丢失风险**

丢失最近一次备份之后的数据

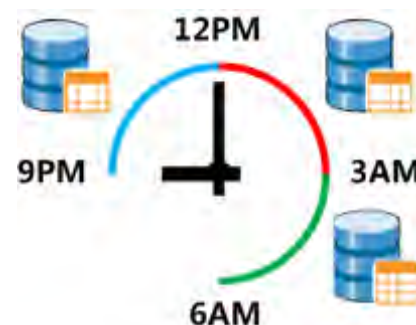


- **恢复有效性无法验证**

介质可靠性/恢复操作过程/恢复的时间



- 备份窗口紧张
避开业务高峰
错开使用设备



- **备份系统的鸡肋**

日常情况下备份起不到作用



数据库与虚拟化存储结合的CDP技术

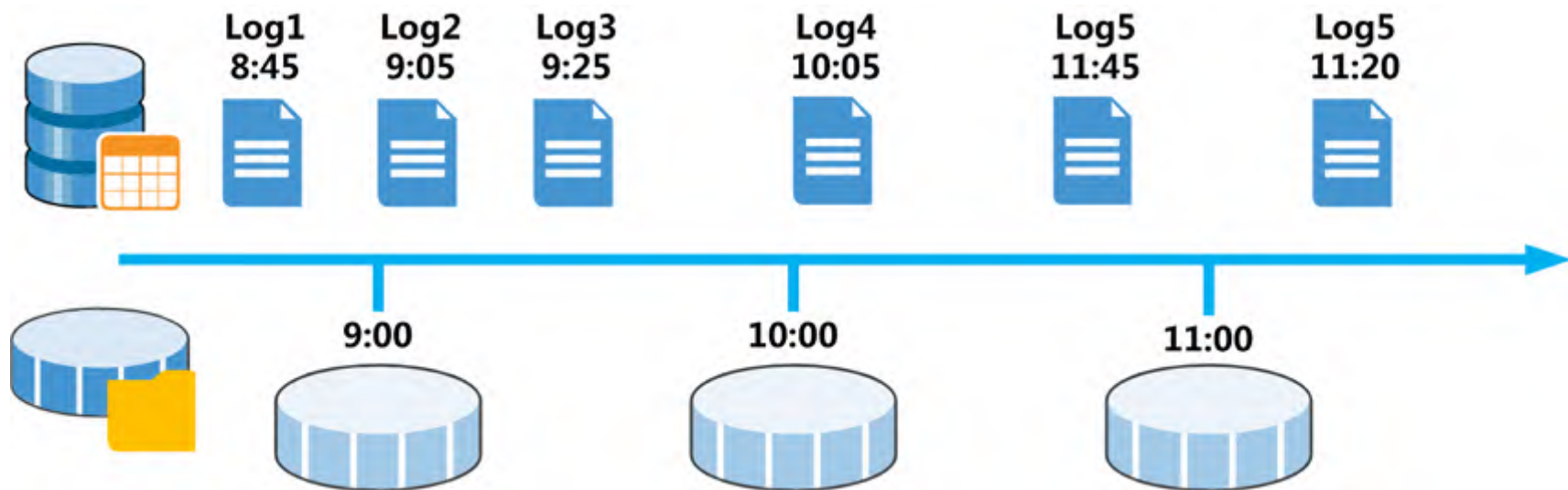
100%可用的数据库持续备份

- 基于数据库复制技术 (Oracle DataGuard/MySQL Replication)
- 结合快照，随时保证数据有完整的备份



数据库的时光轴

存储层定时快照技术+数据库日志



回溯查询

存储层定时快照技术结合指定时间点日志应用

SQL>drop table account

10:30



• 定时快照



10:00



10:00



+



Log4
10:05

+



Log5
10:05-10:29



10:29



Exp &
Imp



~150s



RMAN

>30min*

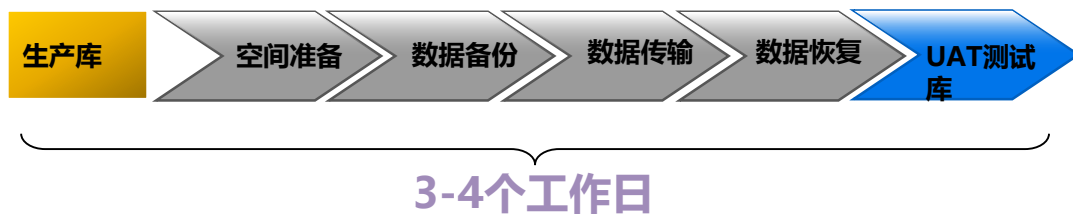


准生产测试环境准备的苦恼

1 测试环境空间不足

2 数据导出需求反复

3 定期的生产同步



数据中心数据层双活定义

双中心同时读写
真正意义的双活

故障无干预
平滑切换，自动接管

性能可接受
支持业务性能激增

分布式集群同城异地读写双活架构



Content/内容

1

数据库基础架构

2

高可用数据库容灾方案

3

海量数据库自动化运维

海量数据库管理运维的困境

随着数据库x86化的不断建设，数据库规模越来越大，带给管理人员的运维成本随之增加，日常DBA经常面临各种问题



数据库一键式自动巡检

The image shows a screenshot of the QMonitor interface. On the left, there is a sidebar with several categories of checks, each with a red box around its title:

- 系统OS配置检查** (System OS Configuration Check):
 - 主机状态 (Host Status)
 - 虚拟内存使用状况 (Virtual Memory Usage Status)
 - 内核 (Kernel)
 - Gird 用户limit (Gird User Limit)
 - 内存CPU使用状况 (Memory CPU Usage Status)
 - OS系 (OS System)
- 数据库配置** (Database Configuration):
 - 数据库基本信息 (Database Basic Information)
 - 数据库实例信息 (Database Instance Information)
 - Oracle (Oracle)
 - 15天日志归档量 (15-day Log Archiving Volume)
 - 4天Undo消耗量 (4-day Undo Consumption)
 - 数据 (Data)
 - 日志切换频率 (Log Switch Frequency)
 - 数据库失效对象 (Database Invalid Objects)
 - 数据 (Data)
- 数据库性能** (Database Performance):
 - 数据库性能 (Database Performance)
- GRID集群与ASM** (GRID Cluster and ASM):
 - GRID软件及PSU (GRID Software and PSU)
 - GRID ASM核心配置参数 (GRID ASM Core Configuration Parameters)
 - GRID (GRID)

On the right, a modal window titled "选择实例应用模板配置" (Select Instance Application Template Configuration) is open. It features a dropdown menu with the following options:

- Select...
- DBMAN : PRIMARY
- standby10g : PHYSICAL STANDBY
- RAC : PRIMARY
- PRIMARY : PRIMARY
- WXH : PRIMARY

集成一体化装机

```
WOQU Cobbler System | woqutech.com
```

```
(local)
```

```
Memory-os
```

```
Hyper-RHEL72-x86_64
```

```
QBackup-RHEL72-x86_64
```

```
MySQL-5.6-rhel66-x86_64
```

```
Oracle-10.2.0.5-rhel5.8-x86_64
```

```
Oracle-11.1.0.7-rhel6.6-x86_64
```

```
Oracle-11.2.0.3-rhel6.6-x86_64
```

```
Oracle-11.2.0.4-rhel6.6-x86_64
```

```
Oracle-12C-rhel6.8-x86_64
```

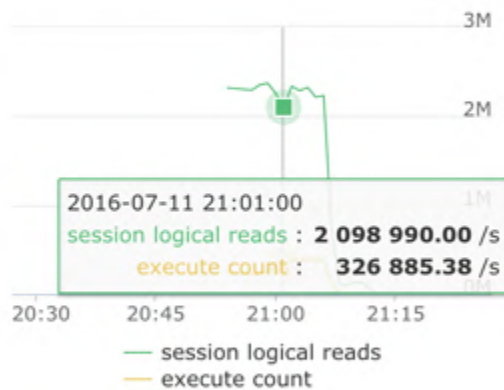
自定义屏幕性能仪表盘

实例首页 > 仪表盘 > 核心库性能仪表盘

核心库CPU: CPU IDLE



核心CRM: RAC : rac1



核心CRM: RAC : rac2



添加监控图表

Oracle SCN HeadRoom监控分析

QMonitor

- 实例列表
- 实时告警
- 智能诊断
 - 巡检报告
 - SCN 诊断
- 仪表盘
- 系统管理
 - 概览
 - 用户
 - 角色
 - 管理授权
 - 配置参数
 - 批量导入
 - 告警模板管理
 - 配置邮件服务

admin

基础数据 智能分析

时间	IP-Port - 实例名	HEADROOM(h)	SCN 增量(/s)
2016-10-27 23:19:50	...	6793.30	0
2016-10-27 23:19:50	...	257325.57	0
2016-10-27 23:19:50	...	257321.76	2
2016-10-27 23:19:50	...	257324.01	0
2016-10-27 23:19:50	...	338.26	15
2016-10-27 23:19:50	...	338.55	20
2016-10-27 23:19:50	...	257327.12	0
2016-10-27 23:19:50	...	257326.28	0
2016-10-27 23:19:51	...	6793.27	38
2016-10-27 23:19:51	...	16298.78	1
2016-10-27 23:19:51	...	338.77	5
2016-10-27 23:19:51	...	338.77	14
2016-10-27 23:19:51	...	338.24	6
2016-10-27 23:19:51	...	338.24	16
2016-10-27 23:19:51	...	339.74	840
2016-10-27 23:19:51	...	340.14	21
2016-10-27 23:19:52	...	338.14	0
2016-10-27 23:19:52	...	338.26	5

HEADROOM(h)

SCN 增量(/s)

QMonitor

- 实例列表
- 实时告警
- 智能诊断
 - 巡检报告
 - SCN 诊断
- 仪表盘
- 系统管理
 - 概览
 - 用户

admin

基础数据 智能分析

起始时间: 2016-10-27 12:00:01 结束时间: 2016-10-28 11:59:59 跳变 HeadRoom(hour): 10 查询

跳变时间点	跳变数据库	HeadRoom 跳变(小时)	跳变 SCN 增长量(/s)	感染源	感染源级别
2016-10-27 22:59:41 2016-10-27 23:04:43	...	12	3002	...	1
2016-10-27 22:59:41 2016-10-27 23:04:43	...	12	3002	...	2
2016-10-27 22:59:41 2016-10-27 23:04:43	...	12	3002	...	3
2016-10-28 06:27:50 2016-10-28 06:32:52	...	48	1327	...	1
2016-10-28 06:27:50 2016-10-28 06:32:52	...	48	1327	...	2
2016-10-28 06:27:50 2016-10-28 06:32:52	...	48	1327	...	3
2016-10-28 06:27:50 2016-10-28 06:32:52	...	48	1327	...	4

HEADROOM增量(h)

数据库性能听诊器

核心一号库

```

Logical Reads      : 2919
Execute Count     : 3906
db block changes  : 2345
Redo Size         : 5,34
Parse Count(Total) : 234
Parse Count(Hard) : 23
User Commits     : 458
Redo writes       : 324
ConCurrent trans  : 30
    
```

TOP EVENT

```

library cache lock      10
shared pool latch       8
enq : index content     2
    
```

```

db file searead time : 2ms
db file scaread time : 4ms
log file par write time : 3ms
log file sync time   : 5ms
    
```

性能仪表

```

db file searead time
db file scaread time
log file par write time
log file sync time
    
```

```

Bytes sent
Bytes received
SQLNET r
    
```

TOP EVENT

```

library cache lock
shared pool latch
enq : index content
    
```

核心二号库

```

Logical Reads      : 3519
Execute Count     : 1302
db block changes  : 2479
Redo Size         : 523
Parse Count(Total) : 321
Parse Count(Hard) : 56
User Commits     : 530
Redo writes       : 270
ConCurrent trans  : 37
    
```

TOP EVENT

```

log file sync          15
shared pool latch       9
enq : row content       1
    
```

```

db file searead time : 1ms
db file scaread time : 3ms
log file par write time : 2ms
log file sync time   : 1ms
    
```

核心三号库

```

Logical Reads      : 2343
Execute Count     : 3094
db block changes  : 3498
Redo Size         : 219
Parse Count(Total) : 156
Parse Count(Hard) : 43
User Commits     : 254
Redo writes       : 234
ConCurrent trans  : 12
    
```

TOP EVENT

```

library cache lock      10
shared pool latch       8
enq : index content     2
    
```

```

db file searead time : 3ms
db file scaread time : 5ms
log file par write time : 2ms
log file sync time   : 6ms
    
```




G*devops*

全球敏捷运维峰会



THANK YOU !