



DTCC

2017第八届中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2017

金融行业数据库存储和加速

徐王锦

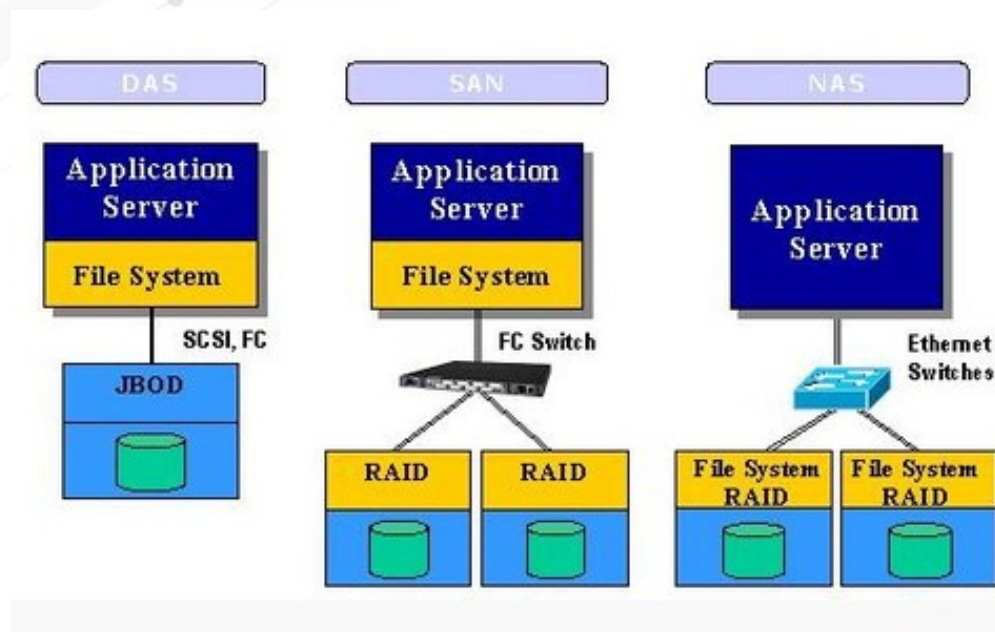
数据存储方式

- EMC、HP、Dell Storage
- SSD、Pcie
- SSD 在MPP（Greenplum）中加速
- Pcie 在RDB（Sqlserver、Mysql HA、LB）使用



传统存储—存储方式

- Nas
- San
- Das



传统存储—存储方式

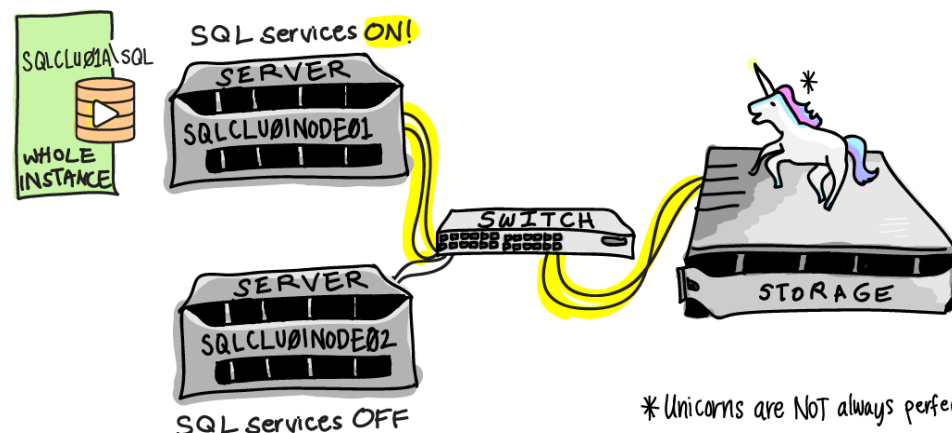
	DAS	NAS	FC-SAN	IP-SAN
成本	低	较低	高	较高
数据传输速度	快	慢	极快	较快
扩展性	无扩展性	较低	易于扩展	最易扩展
服务器访问存储方式	直接访问存储数据块	以文件方式访问	直接访问存储数据块	直接访问存储数据块
服务器系统性能开销	低	较低	低	较高
安全性	高	低	高	低
是否集中管理存储	否	是	是	是
备份效率	低	较低	高	较高
网络传输协议	无	TCP/IP	Fibre Channel	TCP/IP

传统存储—San/Das适用场景

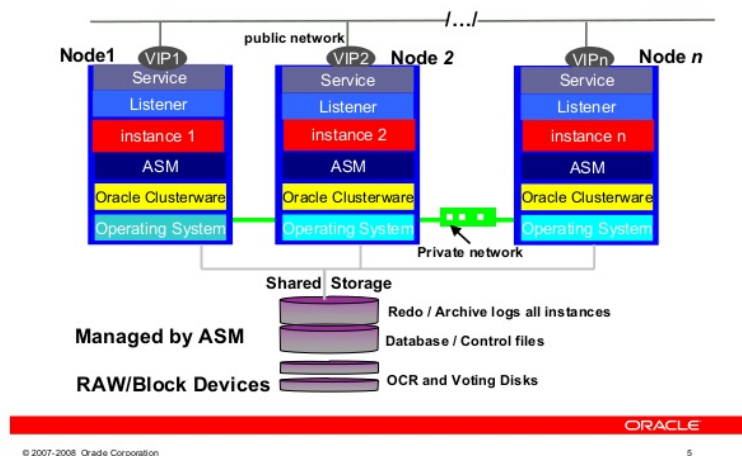
归纳为:

- 交易数据
- 高安全、高性能
- HA

SQLCLUSTER01: ONE SQL SERVER INSTANCE, TWO NODES



Oracle RAC Architecture

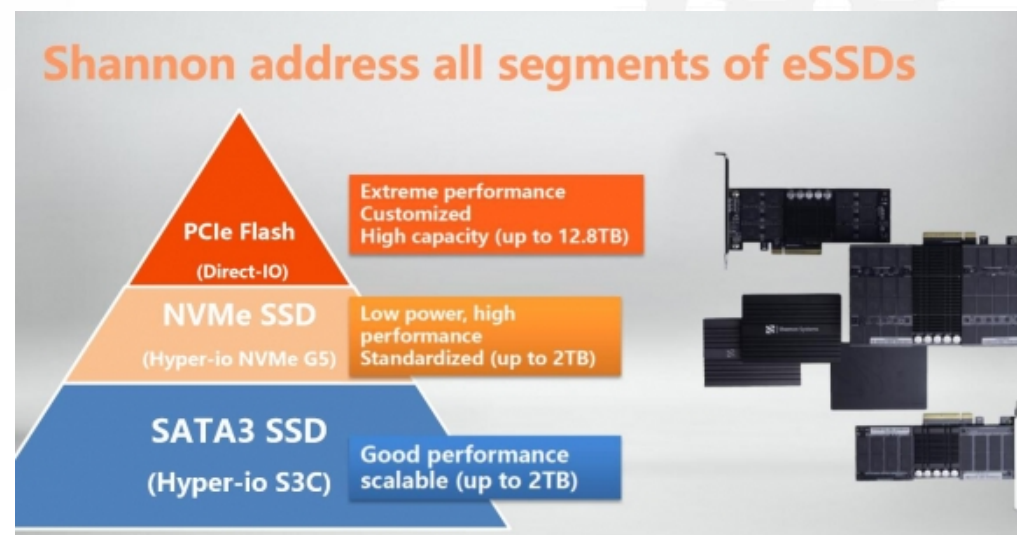


传统存储—优缺点

- 成本高
- 横向扩展有一定局限性
- IO 读取受限于控制器
- 存储本身单点


存储方式—PCIE、SSD

- 相比存储更灵活，性价比高
- 横向扩展灵活
- Pcie、Nvme、Sata3
- 无单点故障
- 单节点QPS/TPS 2w-4w 成本在5w左右
- Mysql MHA、Sqlserver Alwayson cluster



存储方式—PCIE、SSD

Nvme、Sata3

	AHCI	
Uncacheable Register Reads Each consumes 2000 CPU cycles	4 per command 8000 cycles, ~ 2.5 μ s	0 per command
MSI-X and Interrupt Steering Ensures one core not IOPs bottleneck	No	Yes
Parallelism & Multiple Threads Ensures one core not IOPs bottleneck	Requires synchronization lock to issue command	No locking, doorbell register per Queue
Maximum Queue Depth Ensures one core not IOPs bottleneck	1 Queue 32 Commands per Q	64K Queues 64K Commands per Q
Efficiency for 4KB Commands 4KB critical in Client and Enterprise	Command parameters require two serialized host DRAM fetches	Command parameters in one 64B fetch

存储方式—PCIe、SSD

PCIe DIRECT IO

直接io 读写

- 直接io读取去掉了中间数据复制，减少了cpu和内存使用
- 适合高并发场景下RDB数据库应用

SSD 在MPP 中加速存储

归纳为:

- 行列混合存储
- 相对sql on Hadoop 针对500 column 返查询速度快
- 适合数仓层

多态存储

用户自定义数据存储格式

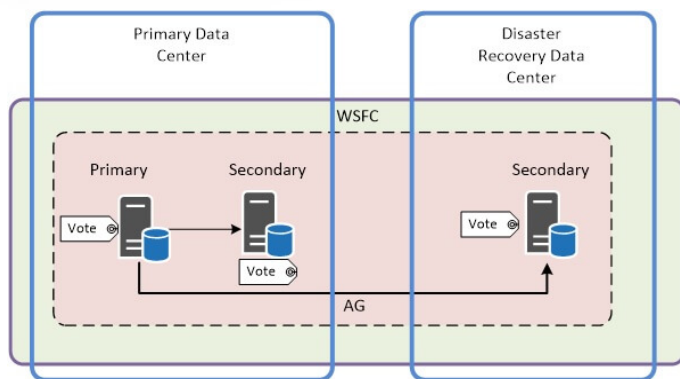
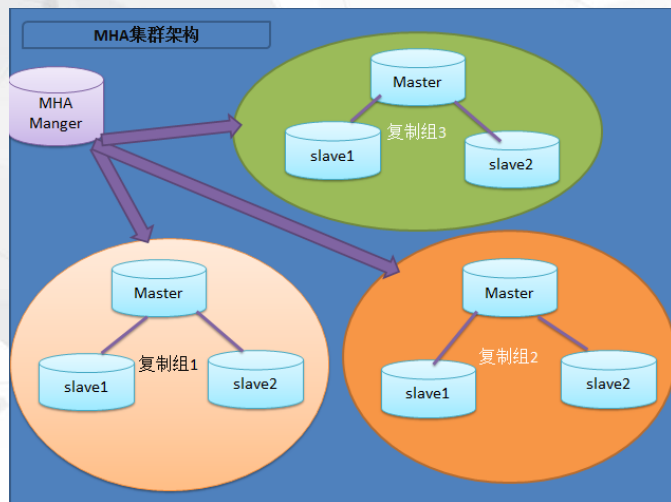


Pivotal

- Greenplum中的行列混存储合支持



Pcie 在RDB (HA/LB) 中应用



归纳为:

- Raid0 pcie 替代 raid10 SSD
- 1:10
- 成本降低50%
- 性能提升30%（长事务情况下）
- 通过DB本身HA 替换硬件磁盘上raid10



THANKS

SequeMedia
盛拓传媒

IT168.com

ITPUB

ChinaUnix