

数据来源：数据库产品上市商用时间



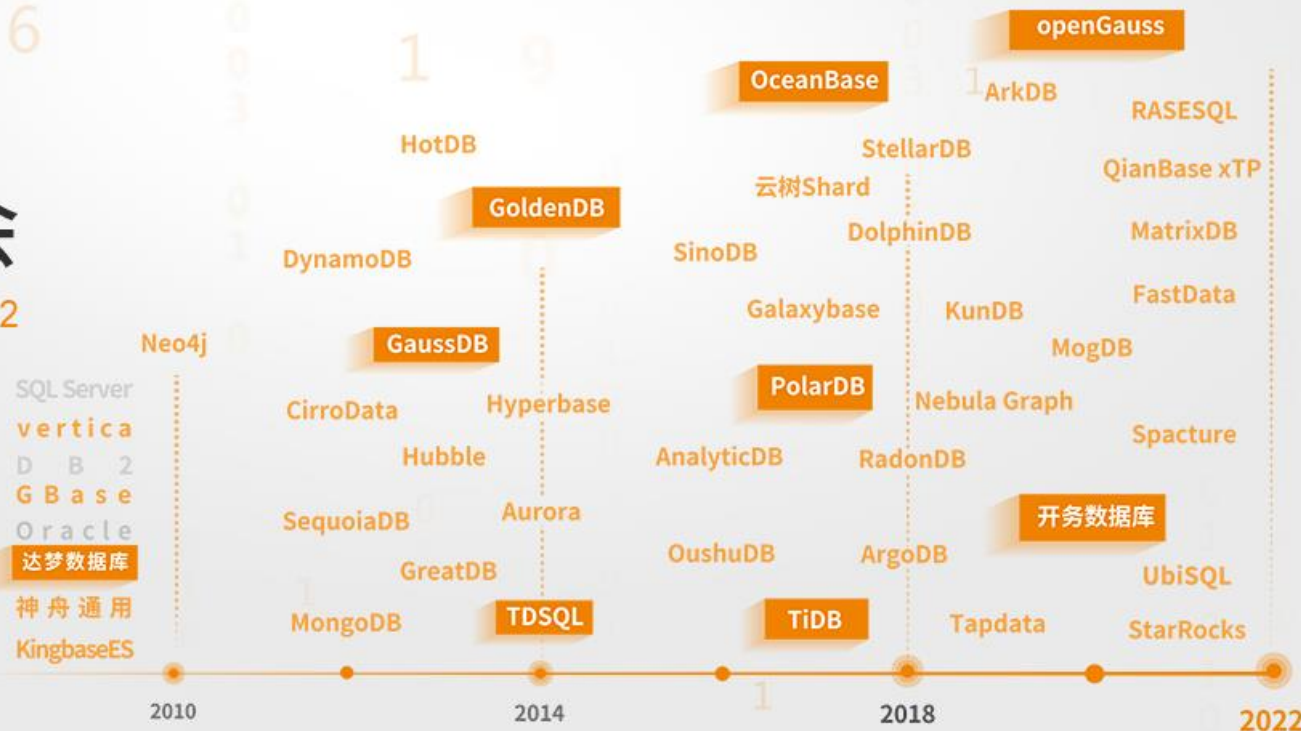
# 第十三届中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2022

## 数据智能 价值创新



线上直播 | 2022/12/14-16



# 可计算存储在数据库应用 场景的实践

ScaleFlux 解决方案 梅庆

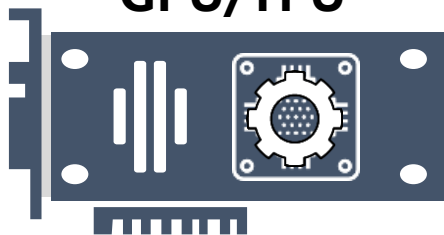
# 可计算存储诞生背景

- 数据存储产能增长跟不上数据的增长速度，数据存储成本在增加。
- 单机 CPU 算力跟不上数据存储的增长速度，计算向专用设备转移。



## 异构计算

GPU/TPU

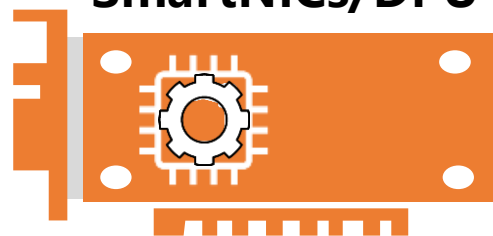


智能加速



## 互联网络

SmartNICs/DPU

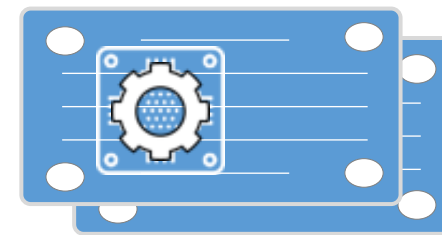


网络带宽能力显著增加  
10 → 100-400Gb/s



## 闪存

Accelerated SSD



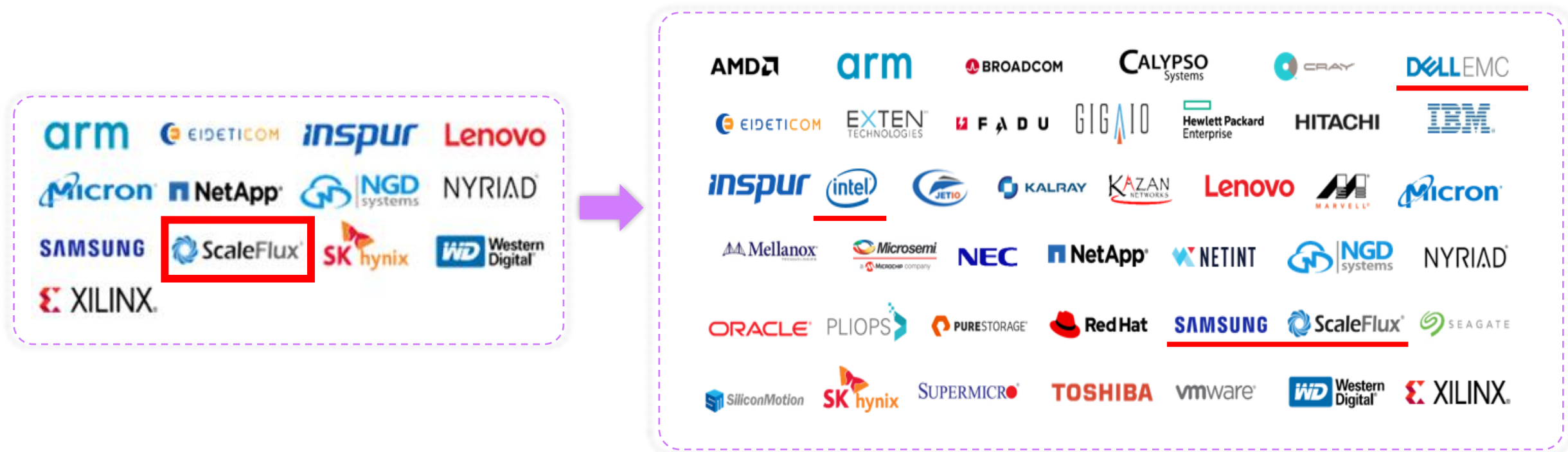
存储性能和容量激增

# 可计算存储工作组

DTCC 2022

第十三届中国数据库技术大会  
DATABASE TECHNOLOGY CONFERENCE CHINA 2022

- 2018年SNIA（存储和网络工业协会）成立了可计算存储工作组，ScaleFlux为创始成员之一
  - 该组织主要致力于可计算存储标准及编程模型的规范定义。
- ScaleFlux：创立于 2014 年，从事可计算存储的研究和生产销售，是大规模部署可计算存储的领导者。





# ScaleFlux 的可计算存储产品 SFX3000

- 基于标准企业级NVMe SSD，内置**计算加速引擎**，在盘内实现数据加速处理的新型企业级NVMe SSD



- 行业标准形态

- 典型企业级特性

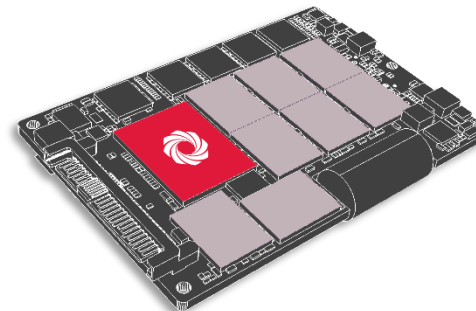
- 标准NVMe驱动
- 掉电保护
- 端到端的数据保护
- IO原子写

- 硬件压缩/解压缩

- 加解密

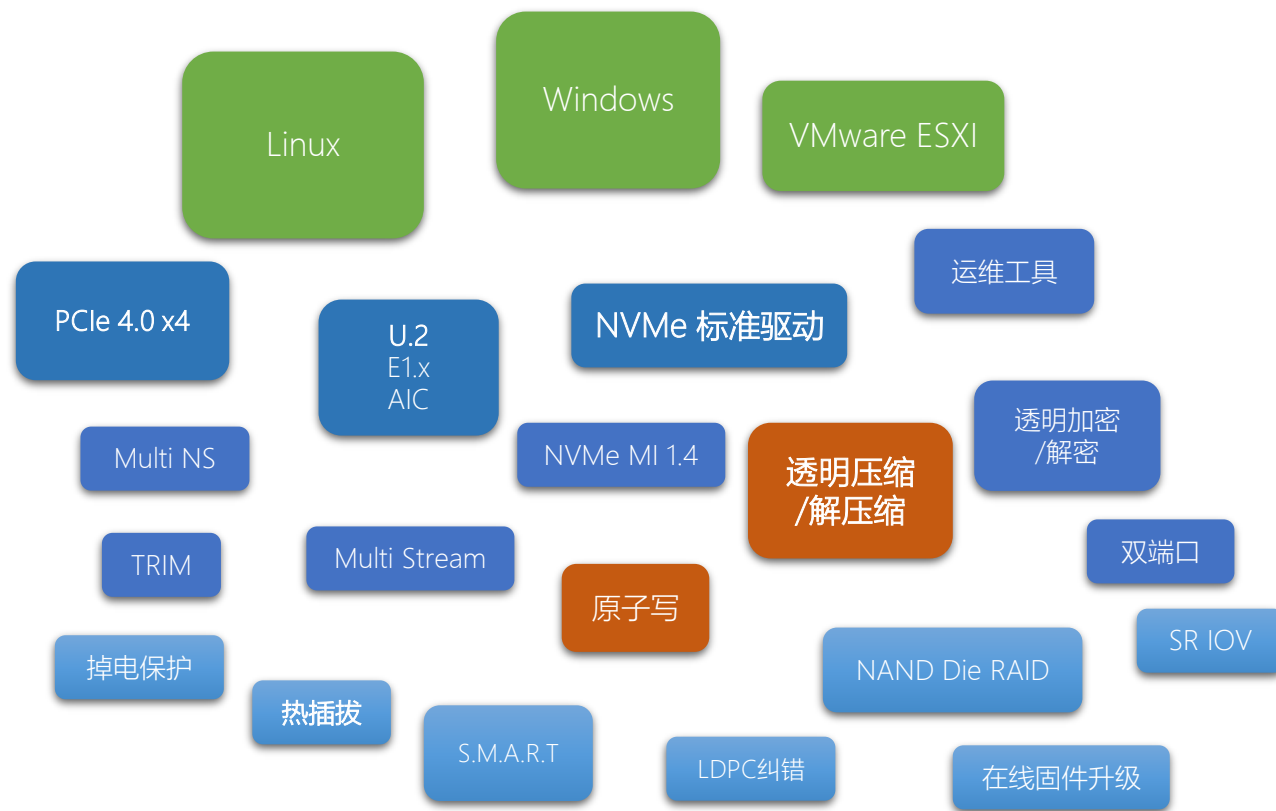
## 可计算存储设备

Computational Storage Device,  
**CSD**



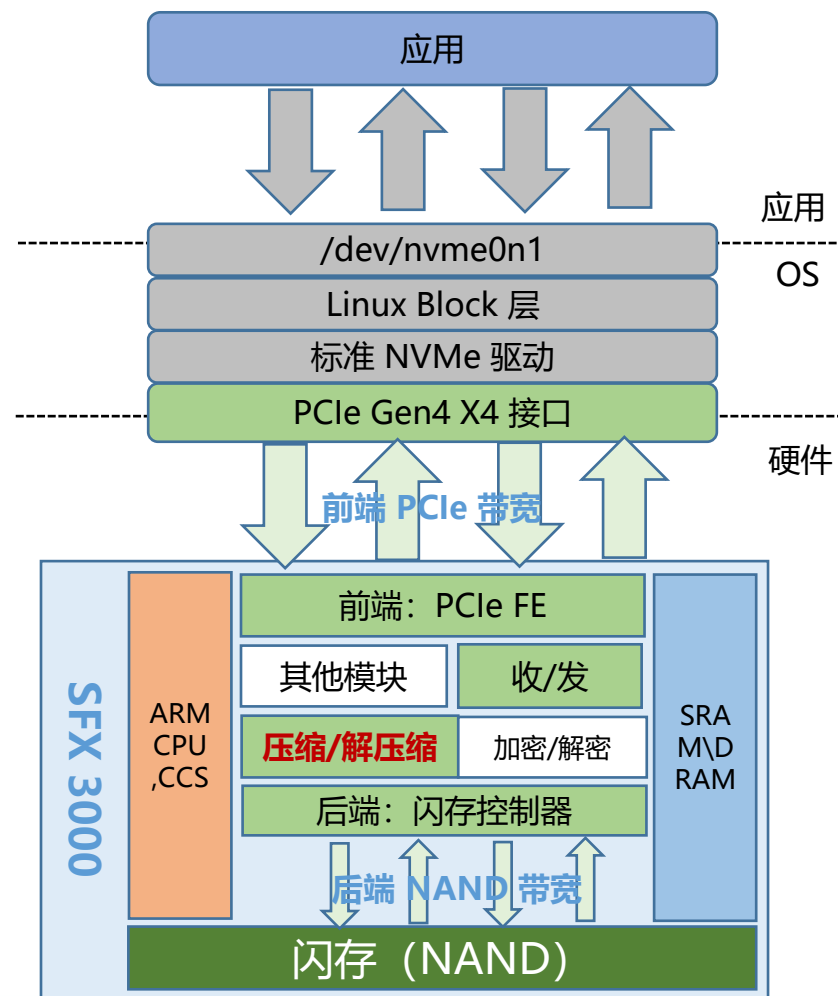
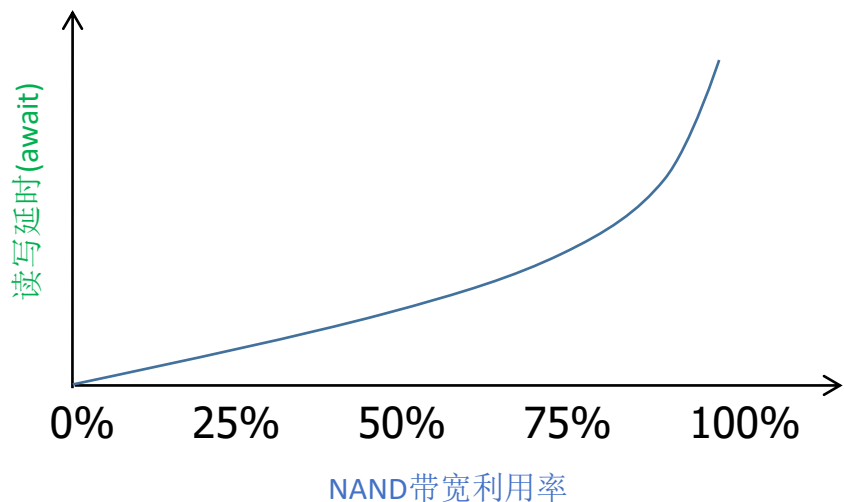
# SFX3000 : 企业级NVMe SSD

- 支持Linux / Windows / Vmware
- Gen4 产品 , 热插拔, 支持 U.2 接口
- 标准 NVMe 协议, 支持 S.M.A.R.T
- 支持透明压缩、加密、原子写
- 支持 Multi NS、Multi Stream



# CSD 透明压缩提升SSD读写性能原理

- CSD 透明压缩在SSD内部 ASIC 里，实时压缩。
- 压缩算法：zlib (level 6)。
- 透明压缩降低后端闪存 (NAND) 带宽利用率。
- SSD 读写延时主要由后端 NAND 带宽利用率决定 (正比关系)。



# CSD透明压缩抑制SSD写放大

- 透明压缩降低数据占用的 NAND 容量。
- 透明压缩降低 GC带来的数据搬迁，抑制写放大。





# 透明压缩对SSD功能的促进作用

■ 当数据的压缩比增加时，CSD 指标

- ✓ 写放大 (WAF) 下降
- ✓ 寿命 (DWPD) 增加
- ✓ 随机读写性能 增加

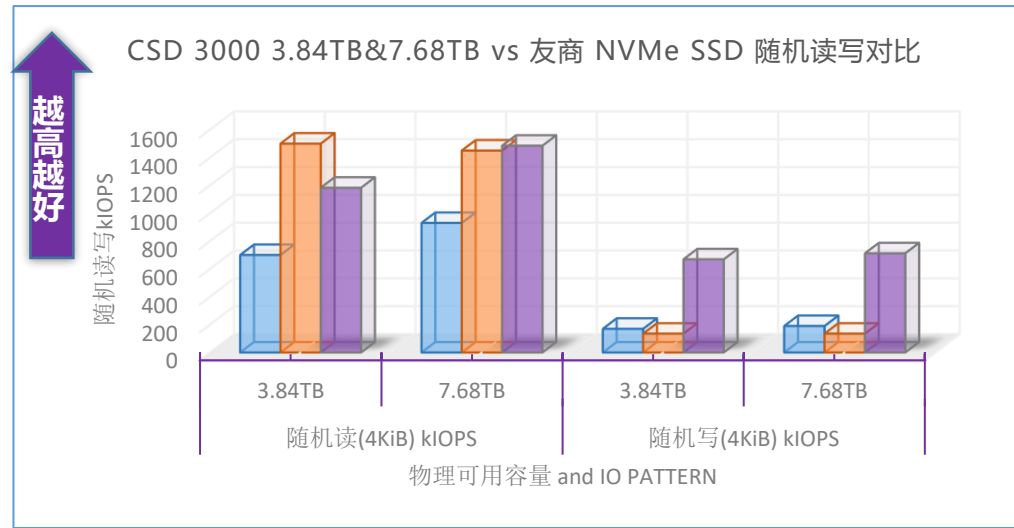
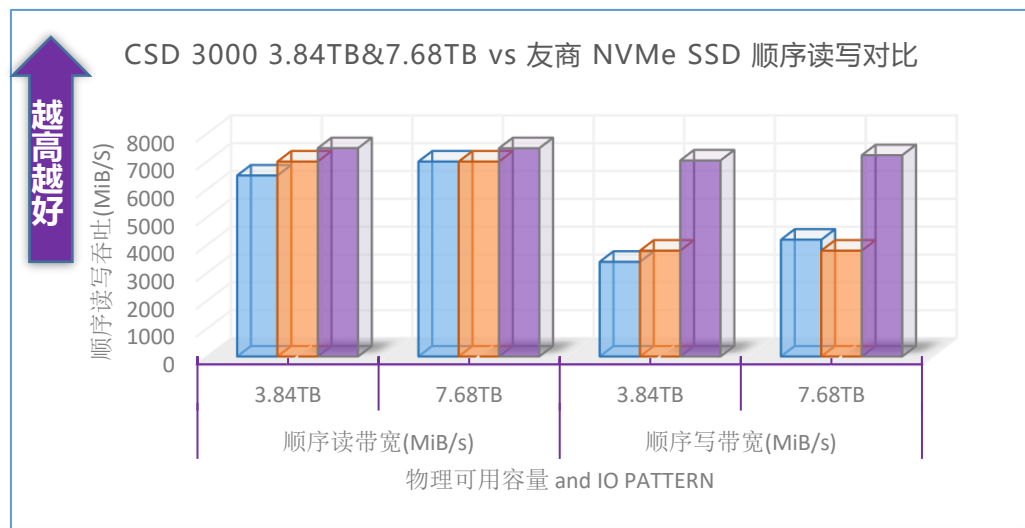


# CSD透明压缩提升顺序/随机读写性能

## ■ CSD 优势

- ✓ 顺序读写最大带宽更高
- ✓ 随机写和混合读写IOPS更高
- ✓ 随机写延时更低、稳定

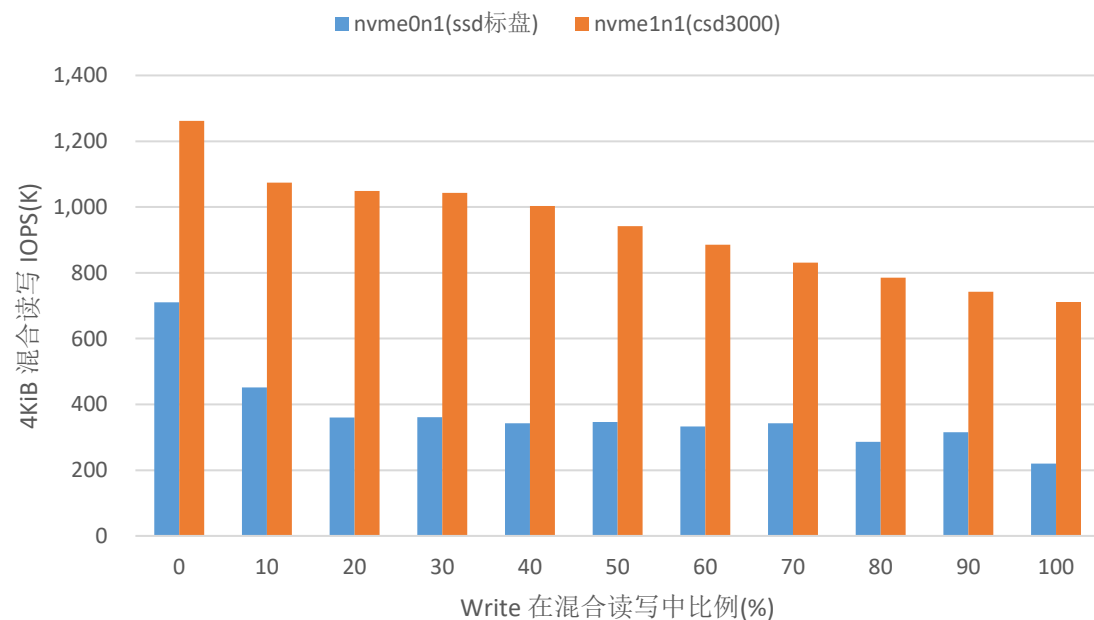
注：紫色背景的是 SFX3000，数据压缩比 2.0.



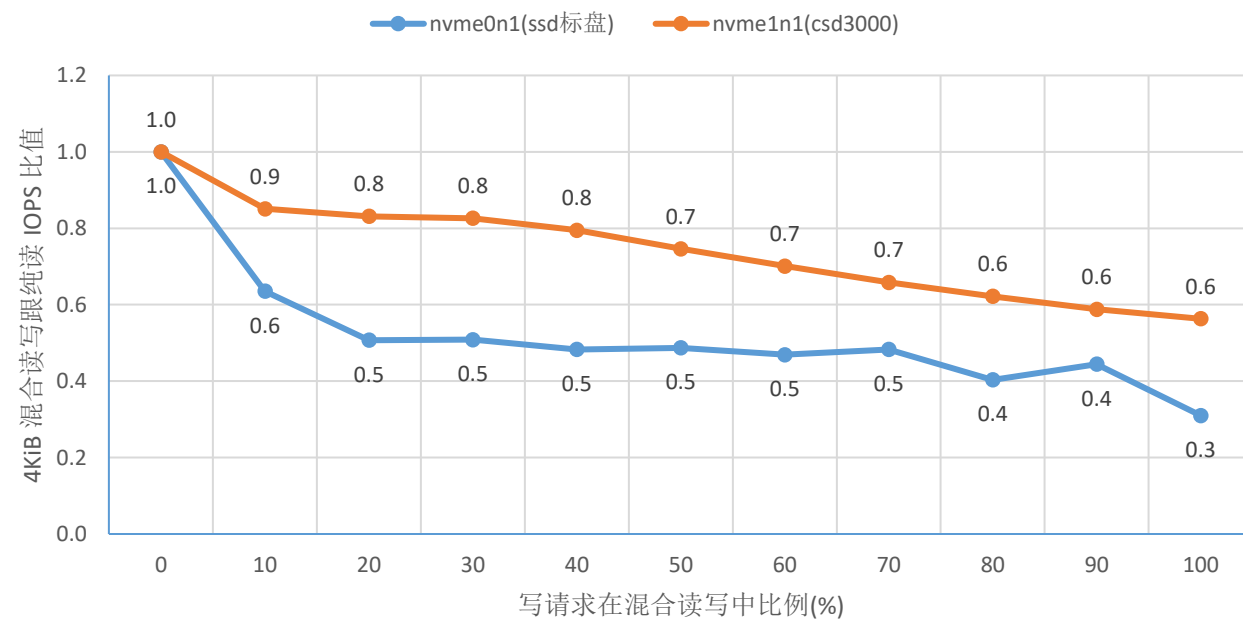
# CSD透明压缩提升混合读写性能

- CSD 随机读写混合的性能高于 SSD 标盘。
- SSD 随机读写混合的性能会随着写比例的增加而下降，CSD 下降速度要低于 SSD 标盘。

4KiB混合读写IOPS(K)跟写请求比例关系



4KiB混合读写跟纯读IOPS比值随写请求比例变化关系



# CSD 在数据库场景价值



- ✓ 更低延迟
- ✓ 3倍 QPS
- ✓ 2倍存储容量/单价



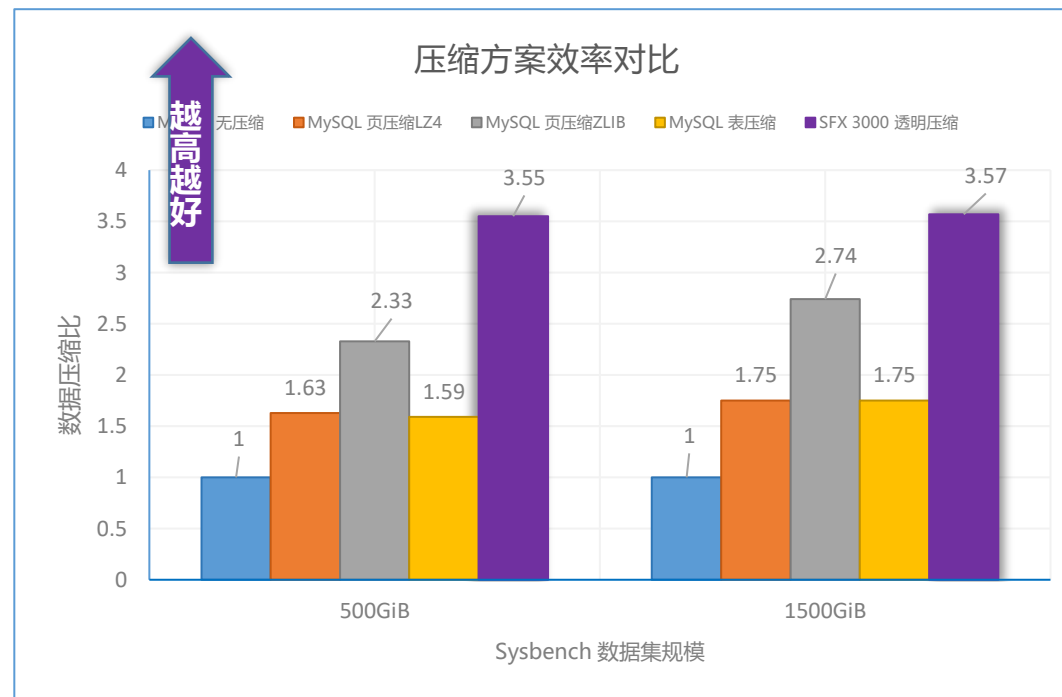
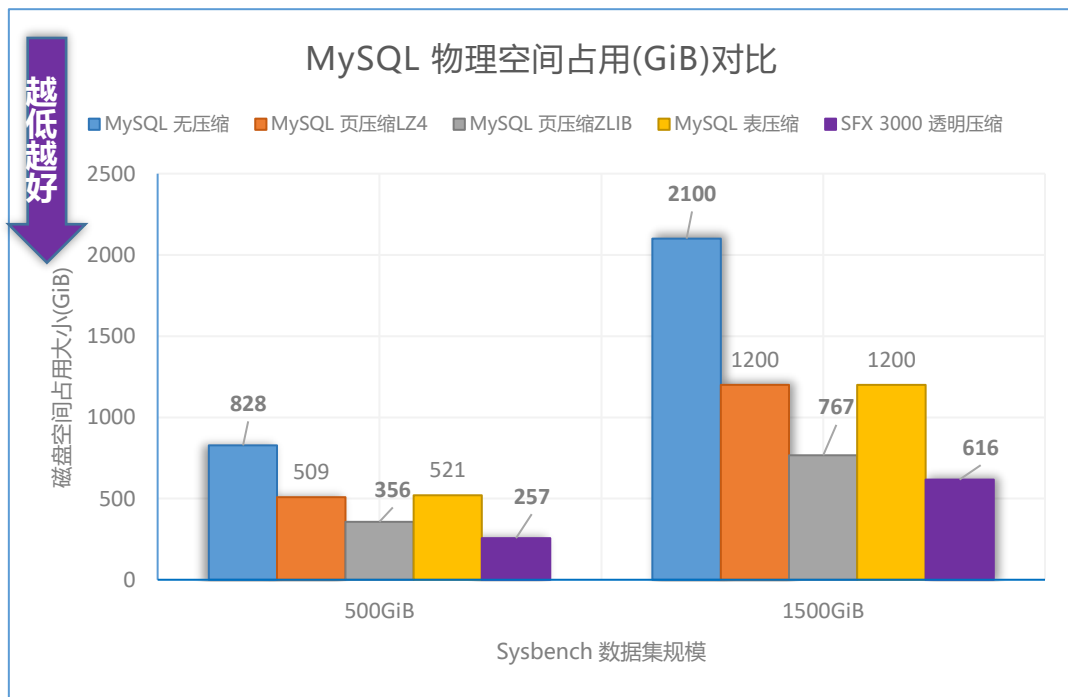
- ✓ 更低延迟
- ✓ 1.6~2倍 QPS
- ✓ 2倍存储容量/单价



- ✓ 更低延迟
- ✓ 2倍存储容量/单价
- ✓ 3-8倍耐久性
- ✓ 1.2~2倍 TPS

# CSD透明压缩降低数据库实际存储成本

- CSD 透明压缩效率高于MySQL的压缩降本方案。
- 高压缩比的好处是 CSD可以扩容降成本。



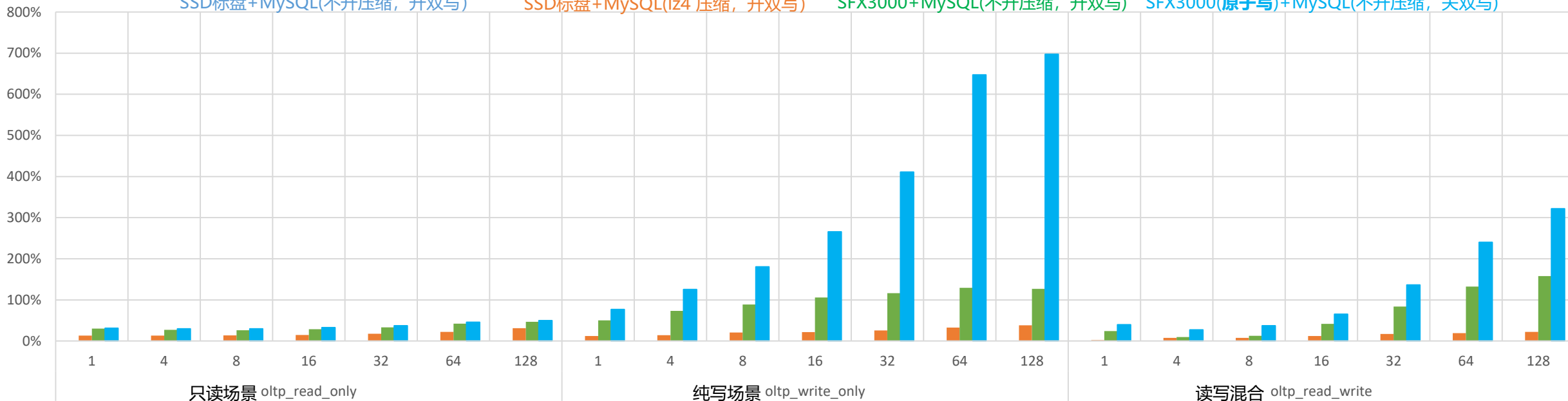


# CSD透明压缩提升MySQL读写性能

- CSD 透明压缩下 OLTP 的性能高于 MySQL 不开压缩和开压缩下的性能。
- 使用SFX3000 原子写代替 MySQL的双写缓冲（Double-Write Buffer）方案，可以极大提升OLTP的写性能。

1500GiB数据集-友商 TLC 3.84TB Gen4 vs CSD 3000 3.84TB TLC Gen4 压缩方案 - 3.0 compress ratio - QPS(%)

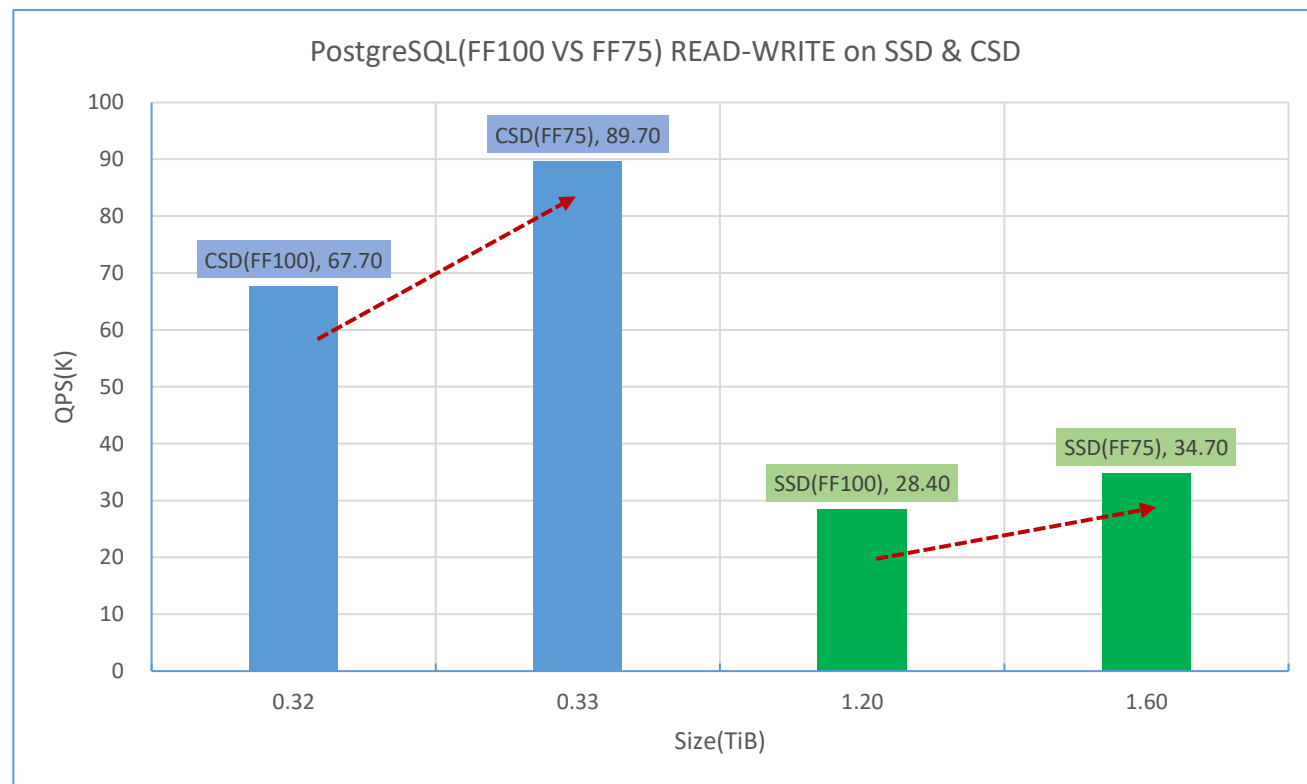
■ 1500G\_nocompress\_dw\_友商i5510\_4T-QPS  
SSD标盘+MySQL(不开压缩, 开双写) ■ 1500G\_page\_lz4\_dw\_友商i5510\_4T-QPS  
SSD标盘+MySQL(lz4 压缩, 开双写) ■ 1500G\_transparent\_dw\_tlc\_4T-QPS  
SFX3000+MySQL(不开压缩, 开双写) ■ 1500G\_transparent\_nodw\_atw\_tlc\_4T-QPS  
SFX3000(原子写)+MySQL(不开压缩, 关双写)



# CSD 透明压缩降低PG空间成本和提升PG性能

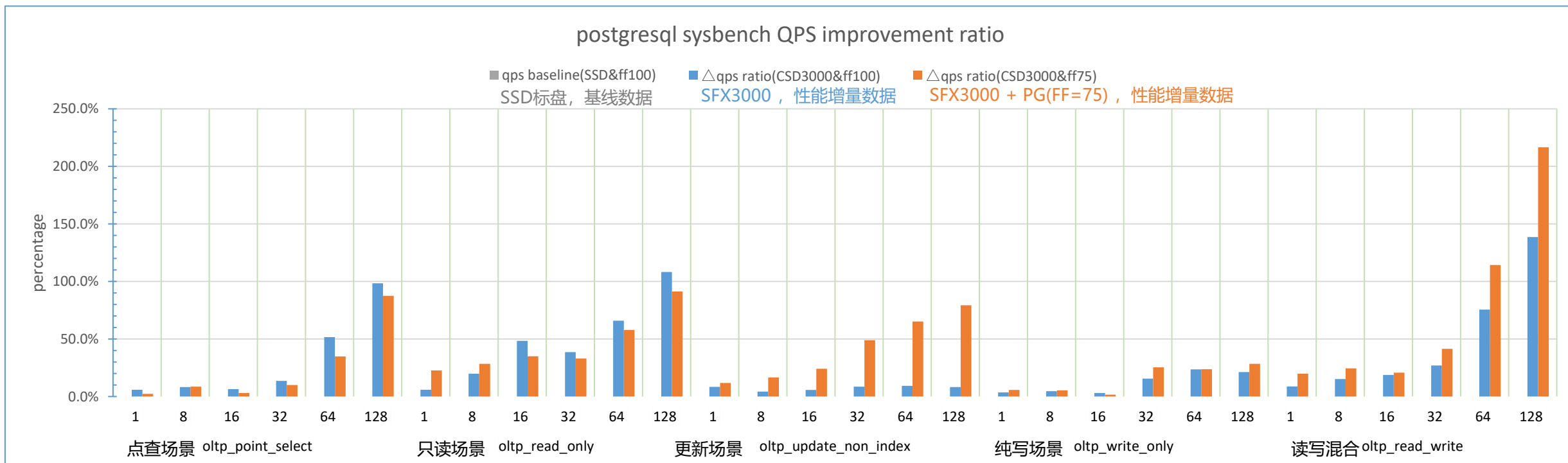
- CSD 透明压缩可以降低 PostgreSQL 的数据存储成本 并提升 OLTP 读写性能。
- 降低 PostgreSQL 的填充因子 (FillFactor, 简称FF, 默认100) , 可以实现提升写性能的同时还不浪费存储空间。

	物理容量(TiB)	读写性能QPS (K)	压缩比
SSD(FF100)	1.20	28.4	1.00
SSD(FF75)	1.60	34.7	1.00
CSD(FF100)	0.32	67.7	3.91
CSD(FF75)	0.33	89.7	4.58



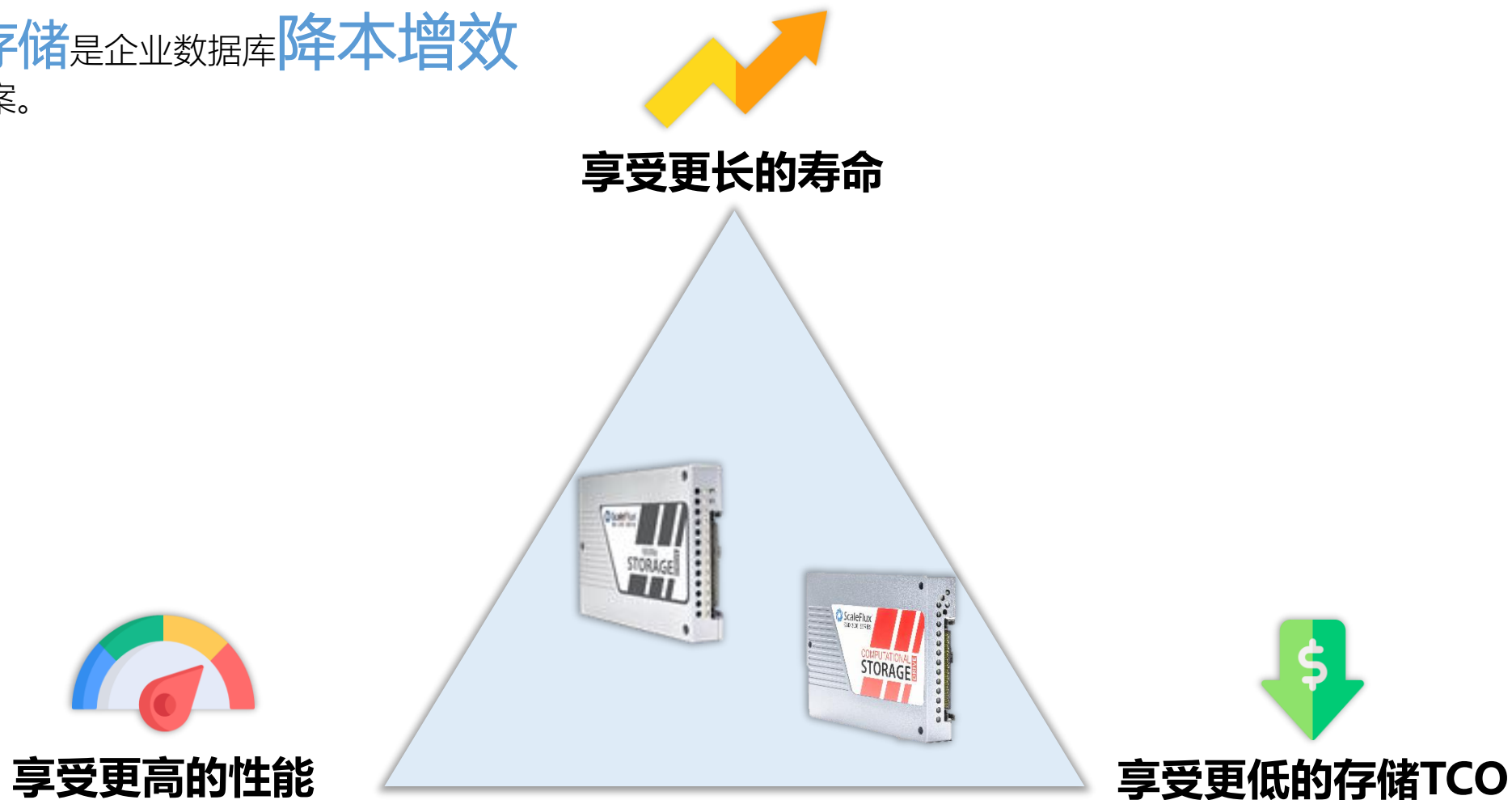
# CSD 透明压缩提升PG读写性能

- CSD 透明压缩可以降低 PostgreSQL 的数据存储成本 并提升 OLTP 读写性能。
- 降低 PostgreSQL 的填充因子 (FillFactor) , 可以实现提升写性能的同时还不浪费存储空间。



# 总结

- 可计算存储是企业数据库**降本增效**的强有力方案。



■ 欢迎扫码，入群交流





# THANKS

SQL Server  
vertica  
D B 2  
G B a s e  
O r a c l e  
达梦数据库  
神舟通用  
KingbaseES

2010

2014

2018

openGauss  
OceanBase  
ArkDB  
RASESQL  
HotDB  
StellarDB  
QianBase xTP  
云树Shard  
GoldenDB  
DolphinDB  
MatrixDB  
DynamoDB  
SinoDB  
FastData  
Galaxybase  
KunDB  
GDB  
GaussDB  
PolarDB  
Kunlun  
Spacture  
SequoiaDB  
OushuDB  
ArgoDB  
开务数据库  
GreatDB  
MongoDB  
TDSQL  
TiDB  
Tapdata  
StarRocks  
UbiSQL