



数据分析大数据最佳配置建议

程从超

英特尔云创新中心架构师

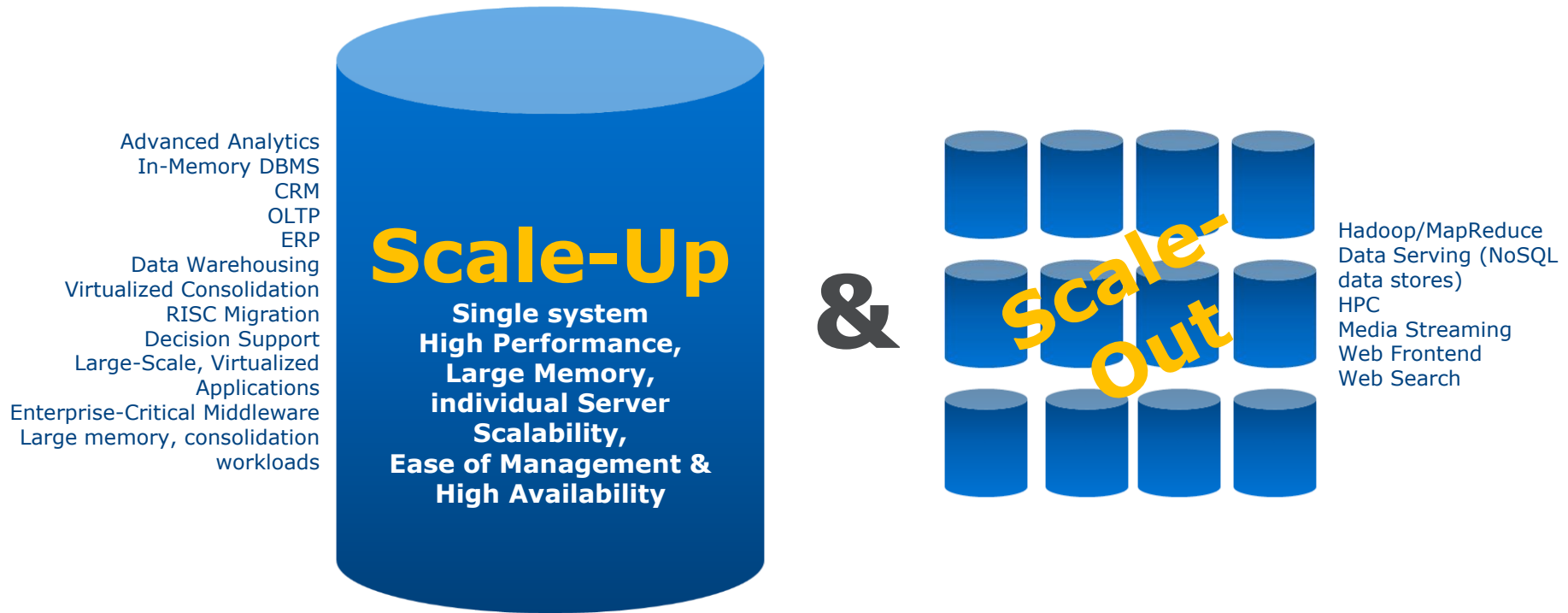


议程

- Scale up和scale out的定义
- 传统关系型数据库配置
- MPP DB数据库配置
- 内存数据库配置
- Hadoop典型配置

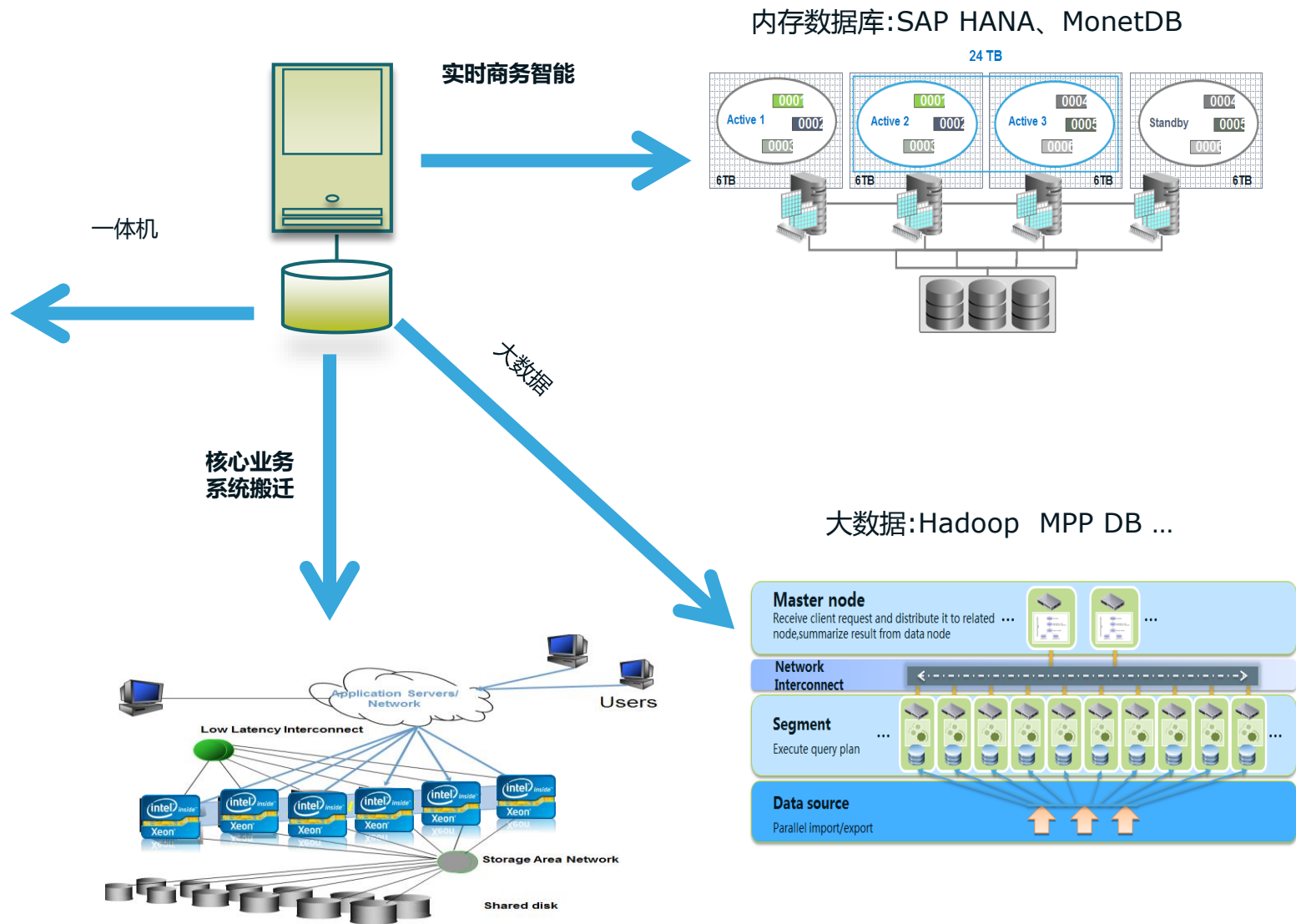
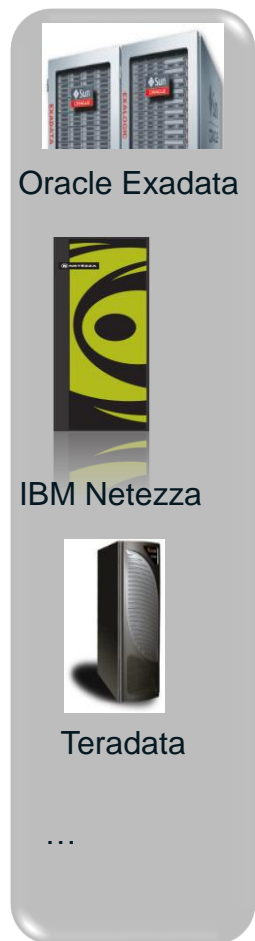


Scale-Up/Scale-out



纵向扩展架构适合于实时/高级的分析和数据驱动的负载类型

那些是scale-up/哪些是scale out?



大数据是整体方案不是某个产品

企业服务总线

批量数据共享

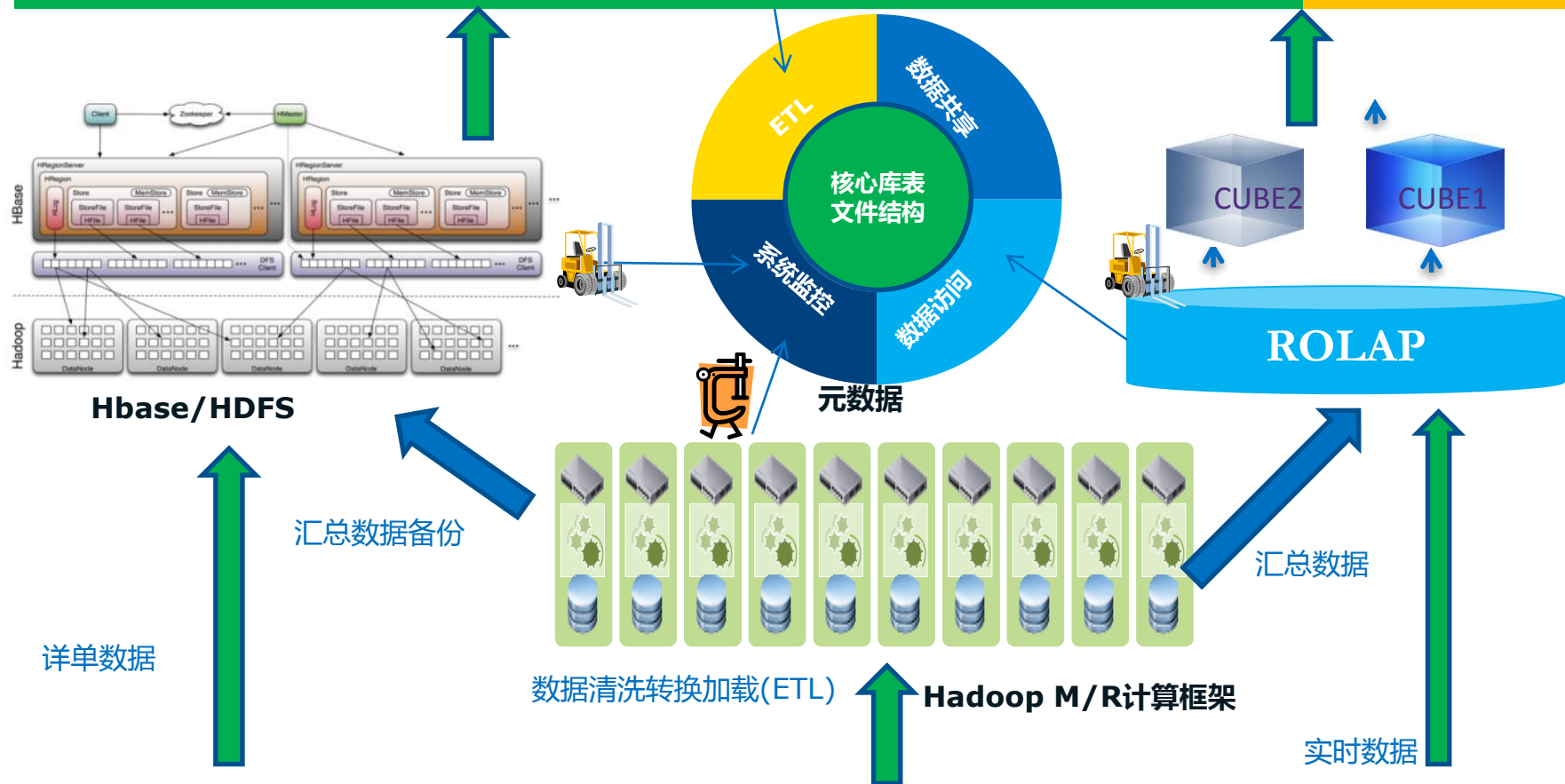
即席查询

异步任务提交

移动客户端

数据访问框架(Spring)

传统BI 工具



非结构化数据

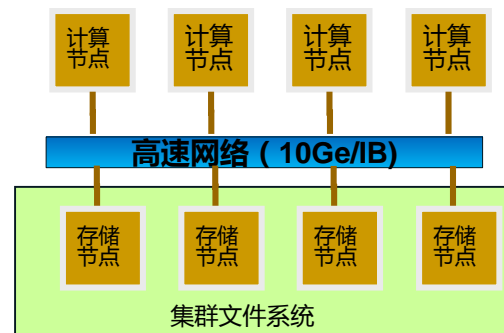
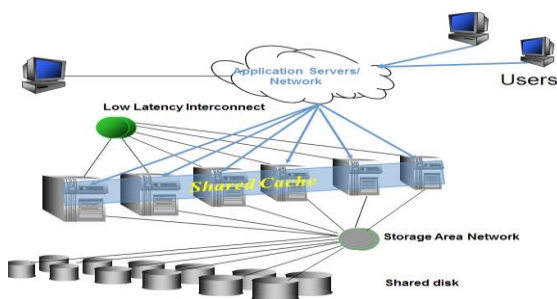
结构化数据



议程

- Scale up和scale out的定义
- **传统关系型数据库配置**
- MPP DB数据库配置
- 内存数据库配置
- Hadoop典型配置

传统关系数据库演进1:share-everything+SDS



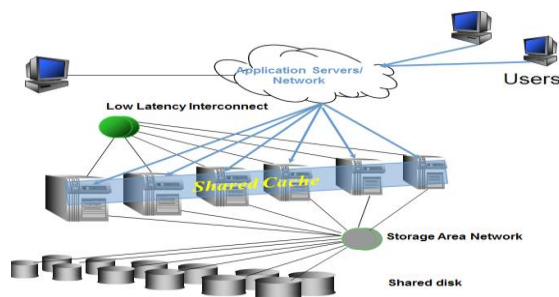
传统的小机+阵列

- 成本高
- 扩展性差

计算节点采用4路服务器+存储节点
采用2路服务器+SSD+高速网络

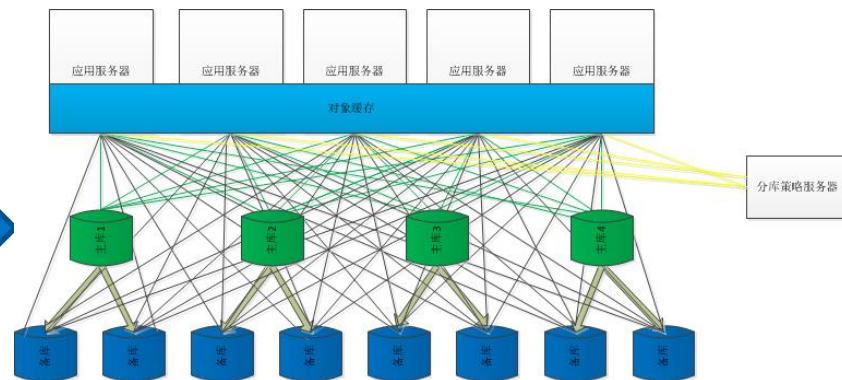
- 性价比高
- 扩展性好

传统关系数据库演进2:sharding



传统的小机+阵列

- 成本高
- 扩展性差

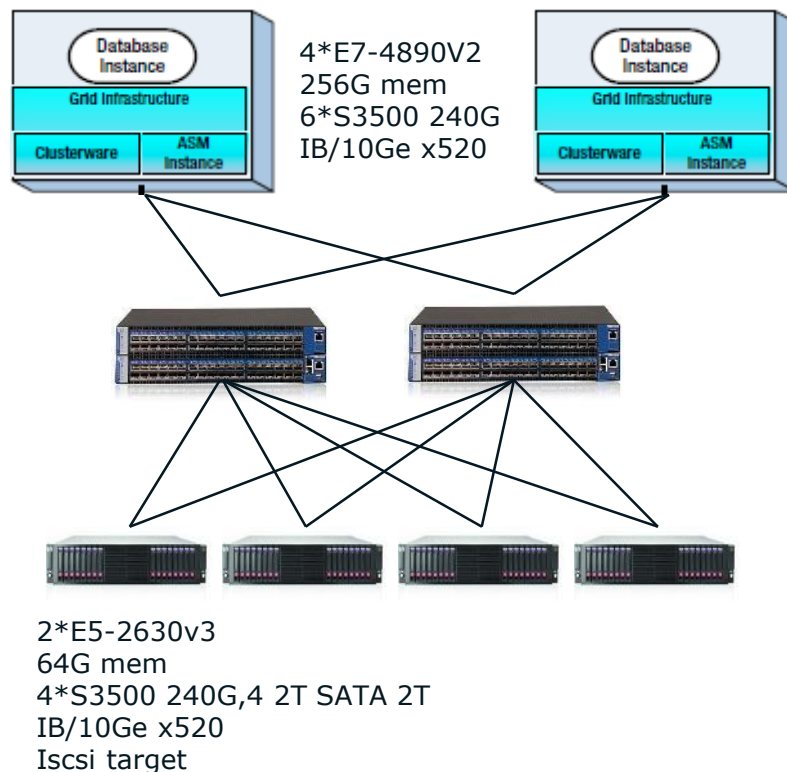


• Share-nothing架构

- 性价比高
- 扩展性好
- 原有应用需要重新设计架构
- 扩容时数据重分布复杂
- 跨库交易复杂

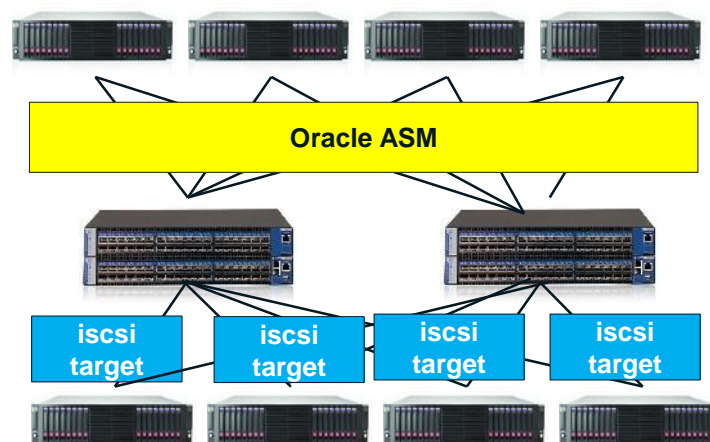
Share-everything的典型配置一

- 计算节点和存储节点分开
- 计算采用4路服务器，CPU越高越好，每节点内存G=核数×6-12
- 存储采用两路服务器，CPU采用中端E5-2640V3+即可，存储节点根据系统的性能词啊用NVMe PCI-E SSD/SATA SSD/SAS/SATA,或者分层存储。
- 存储运行分布式文件系统
- 网络最好使用Infiniband，10GE也可以，但最好在应用层做负载分担



Share everything的典型配置二

- 计算节点和存储节点合二为一，每台服务器即是计算节点，也是存储节点，节点CPU主频越高越好E5-2699V3，每节点内存G=核数×6-12
- 所有服务器运行分布式文件系统
- 最好用infiniband做数据网络和心跳，10Ge也可，但需要应用层做负载分担。
- 存储根据需要可以采用PCI-E ssd，sata ssd或者sas，也可以混插支持分层存储。



2*E5-2699v3
64G mem
4*S3500 240G,4 2T SATA 2T
IB/10Ge x520
Iscsi target

英特尔® 至强™ 处理器E7 V3产品家族

领先的企业计算和内存计算平台

卓越
性能

6倍

多达

英特尔事务同步扩展
(TSX) 指令集助力实现
多达 6 倍性能提升¹

更大
规模

18核

多达

每个 CPU 多达 18 个
内核，并支持
四路以上配置

海量
内存

12TB

多达

支持八路配置及
DDR3 和 DDR4 内存

高度
可信

稳定
可靠

英特尔 AVX2 和英特
尔可靠运行技术
(Intel Run Sure
Technology)

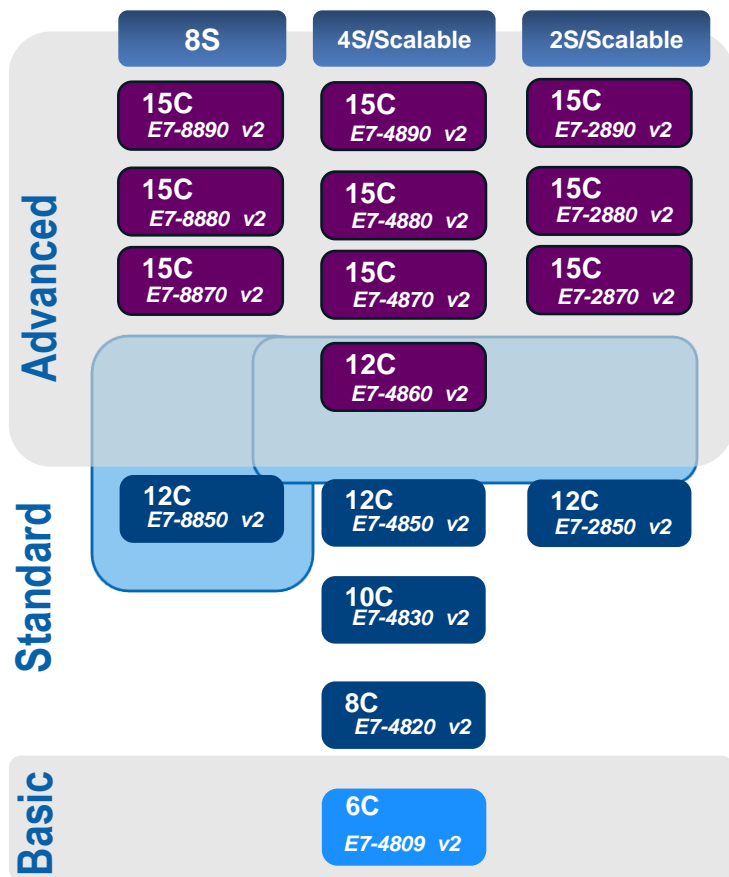
在性能检测过程中涉及的软件及其性能只有在英特尔微处理器的架构下方能得到优化。诸如 *SSmark* 和 *MobileMark* 等性能测试均系基于使用特定计算机系统、组件、软件、操作系统及功能。上述任何要素的变动都有可能导致测试结果的变化。请参考其它信息及性能测试（包括结合其它产品使用时的运行性能）以对目标产品进行全面评估。如欲了解更多信息，请访问：<https://www.intel.com/performance>。

¹ 多达 6 倍的性能提升基于 *SAP HANA*™ *1SP7* 内部在线交易处理 (OLTP) 工作负载，比较了英特尔® 至强® 处理器 *E7-8870 v3*（配有英特尔® TSX，得分 *87,677 tpm*）和英特尔® 至强® 处理器 *E7-4870 v2*（得分 *14,327 tpm*）。

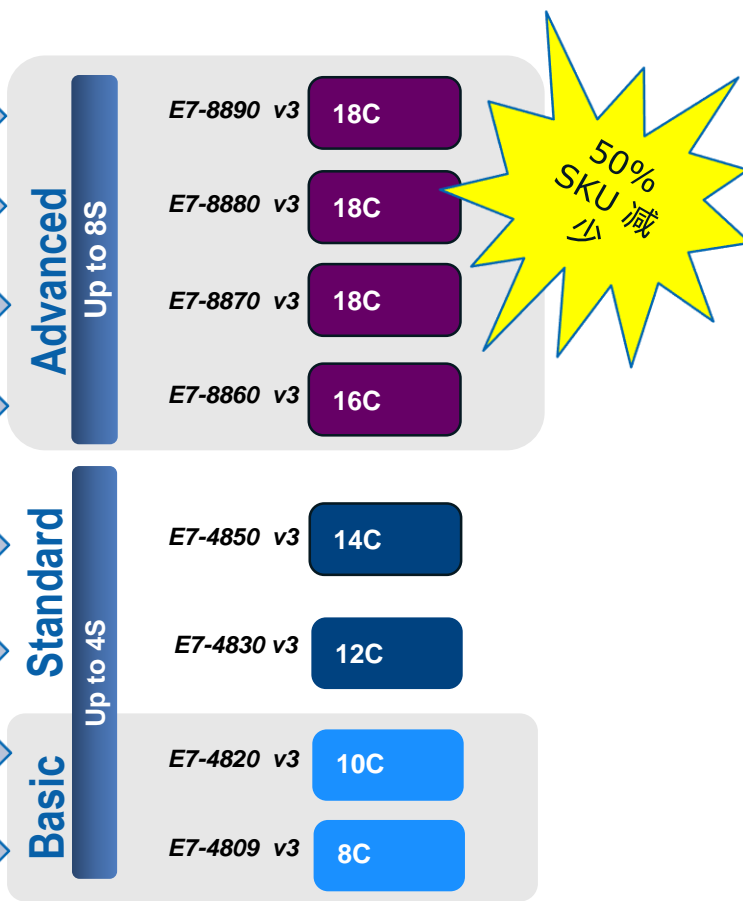


E7V3产品全系列

Intel® Xeon® Processor E7 v2 General Purpose SKUs



Intel® Xeon® Processor E7 v3 General Purpose SKUs



进一步简化了客户的验证和选择复杂性

英特尔®至强™ E5-2600 v3系列处理器产品升级指南

至强™ E5-2600 v2

性能提升¹

至强™ E5-2600 v3

**行业优化型
(E5-2600 v3)**

35-45 MB 三级缓存
9.6 GT/s QPI1.1
DDR4-2133
睿频加速
超线程

53%¹

高级型号

25-30 MB 三级缓存
9.6 GT/s QPI1.1
DDR4-2133
睿频加速
超线程

10%¹

17%¹

标准型号

15-20 MB 三级缓存
8.0 GT/s QPI1.1
DDR4-1866
睿频加速;超线程

30%¹

66%¹

基本型号

15 MB 三级缓存
6.4 GT/s QPI1.1
DDR4- 1600

E5-2697 v2 12C 130W
2.7G/30M/3.5 Turbo
E5-2695 v2 12C 130W
2.4 G/30M/3.2 Turbo

44%, 33%

30%, 26%

26%, 22%

21%, 20%

E5-2699 v3 18C 145W
2.3G/45M/3.6G Turbo

E5-2698 v3 16C 135W
2.3G/40M/3.6G Turbo

E5-2697 v3 14C 145W
2.6G/35M/3.6G Turbo

E5-2695 v3 14C 120W
2.3G/35M/3.0G Turbo

E5-2690 v2 10C 130W
3.0G/25M/3.6 Turbo

24%, 23%

E5-2680 v2 10C 115W
2.8G/25M/3.6 Turbo

24%, 21%

E5-2670 v2 10C 115W
2.5G/25M/3.3 Turbo

23%, 24%

E5-2660 v2 10C 95W
2.2G/25M/3.0 Turbo

22%, 20%

E5-2650 v2 8C 95W
2.6G/20M/3.4 Turbo

24%, 24%

E5-2690 v3 12C 135W
2.6G/30M/3.5G Turbo

E5-2680 v3 12C 120W
2.5G/30M/3.3G Turbo

E5-2670 v3 12C 120W
2.3G/30M/3.1G Turbo

E5-2660 v3 10C 105W
2.6G/25M/3.3G Turbo

E5-2650 v3 10C 105W
2.3G/25M/3.0G Turbo

E5-2640 v2 8C 95W
2.0G/20M/2.5 Turbo

34%, 28%

E5-2630 v2 6C 80W
2.6G/15M/3.1 Turbo

36%, 31%

E5-2620 v2 6C 80W
2.1G/15M/2.6 Turbo

22%, 22%

E5-2640 v3 8C 90W
2.6G/20M/3.4G Turbo

E5-2630 v3 8C 85W
2.4G/20M/3.2G Turbo

E5-2620 v3 6C 85W
2.4G/15M/3.2G Turbo

E5-2609 v2 4C 80W
2.5G/10M

26%, 24%

E5-2603 v2 4C 80W
1.8G/10M

47%, 39%

E5-2609 v3 6C 85W
1.9G/15M

E5-2603 v3 6C 85W
1.6G/15M

*上图所示处理器升级指南并不涉及任何价格的对应关系及建议

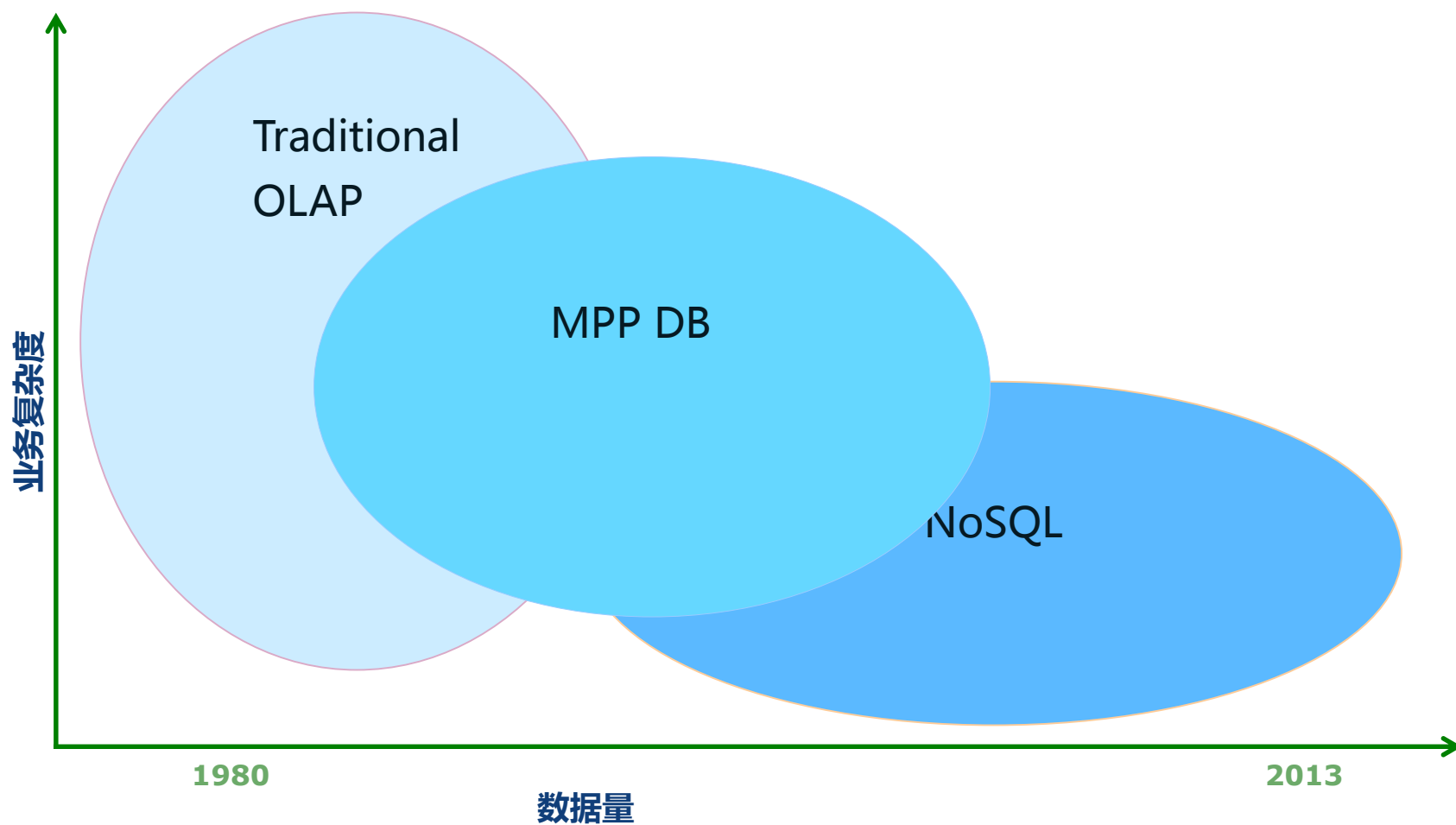
SPECint_rate_base2006%, SPECfp_rate_base2006%



议程

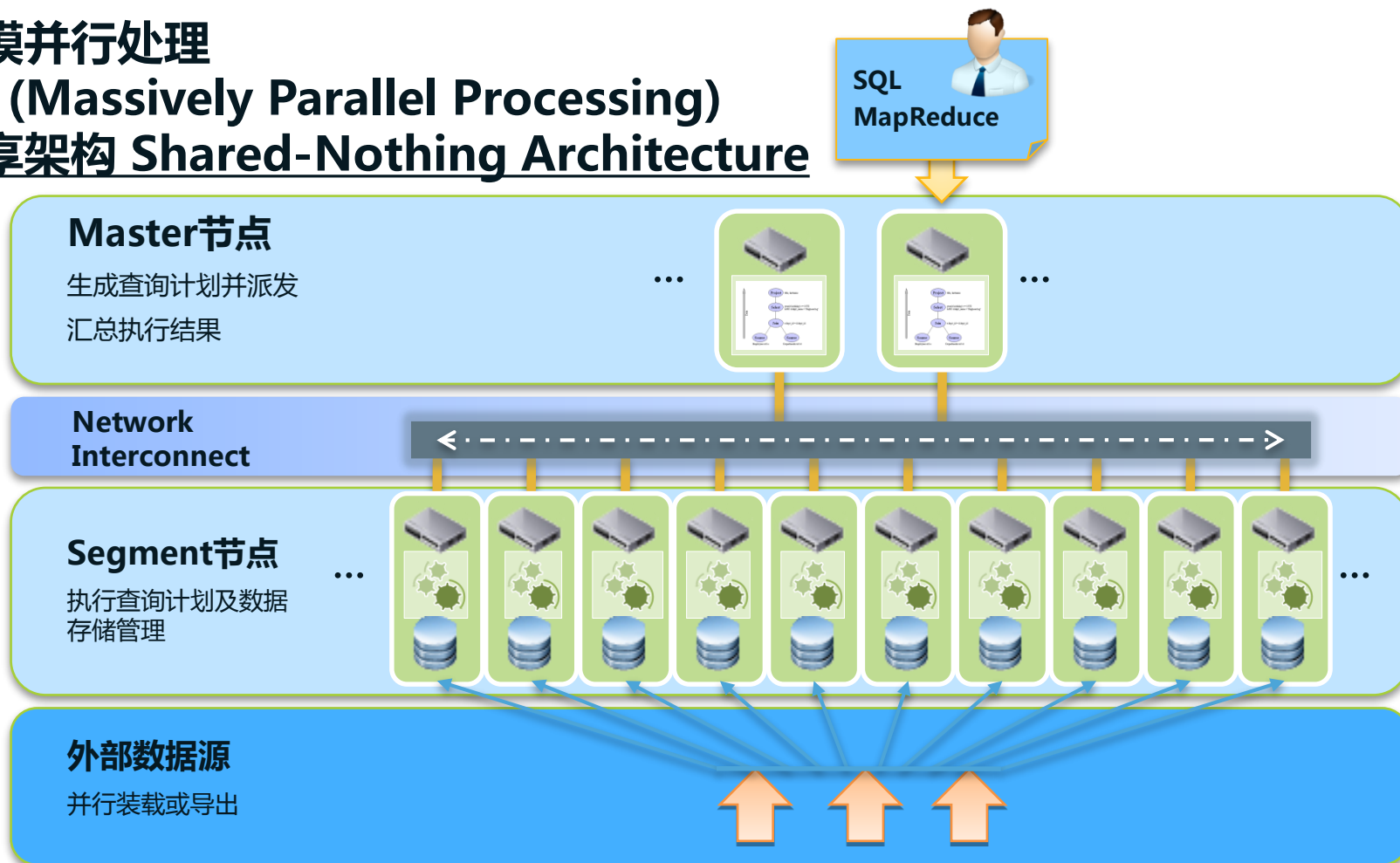
- Scale up和scale out的定义
- 传统关系型数据库配置
- **MPP DB数据库配置**
- 内存数据库配置
- Hadoop典型配置

当前的数据处理技术



MPP DB的架构形态

大规模并行处理
MPP (Massively Parallel Processing)
无共享架构 Shared-Nothing Architecture



配置建议

	计算型	存储型	综合
CPU	E5-2699V3 越高越好	E5-2640V3 标准型即可	E5-2660V3
内存	CPU核数*12	CPU核数×6	CPU核数×8
网络	10Ge	10Ge	10Ge
硬盘	SSD/SAS	SATA/SAS	SAS/SATA

存储可以采用分级的方式来区分热数据/温数据和冷数据

议程

- Scale up和scale out的定义
- 传统关系型数据库配置
- MPP DB数据库配置
- **内存数据库配置**
- Hadoop典型配置

内存数据计算的分类

通用内存数据库：遵循sql92接口标准，scale-up架构

- SAP HANA
- Oracle Timesten
- Altibase
- eXtremeDB
- SolidDB
- FastDB
- 华为 IMDB

内存数据网格：大部分为key/value方式，scale-out架构

- Oracle Coherence
- Pivotal Gemfire
- Spark/shark
- Redis
- ...

其他：

- MonetDB(Vectorwise)

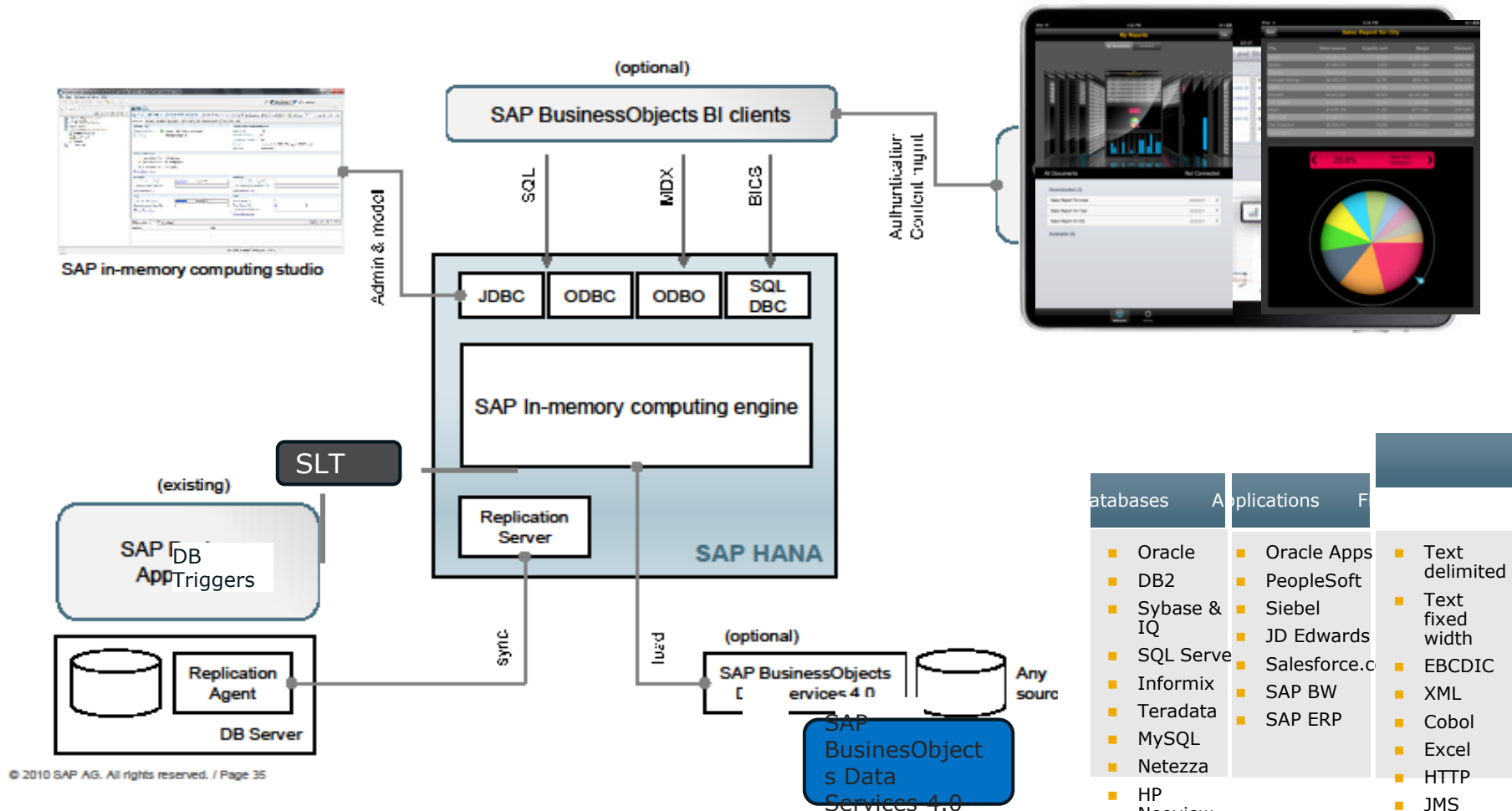
内存数据库和传统数据库的比较

- Scale-up架构居多，受制于单台机器的内存限制。
- 两种应用模式：
 - 单独作为数据存储
 - 作为后台关系型数据库的缓存

	内存数据库	传统数据库
架构	内存为中心	磁盘为中心
实施方式	可以嵌入应用	单独的服务器
I O 操作	无	有
响应速度	微秒到毫秒	毫秒级
数据容量	较少	超大数据量
数据特性	短暂	持久

SAP HANA

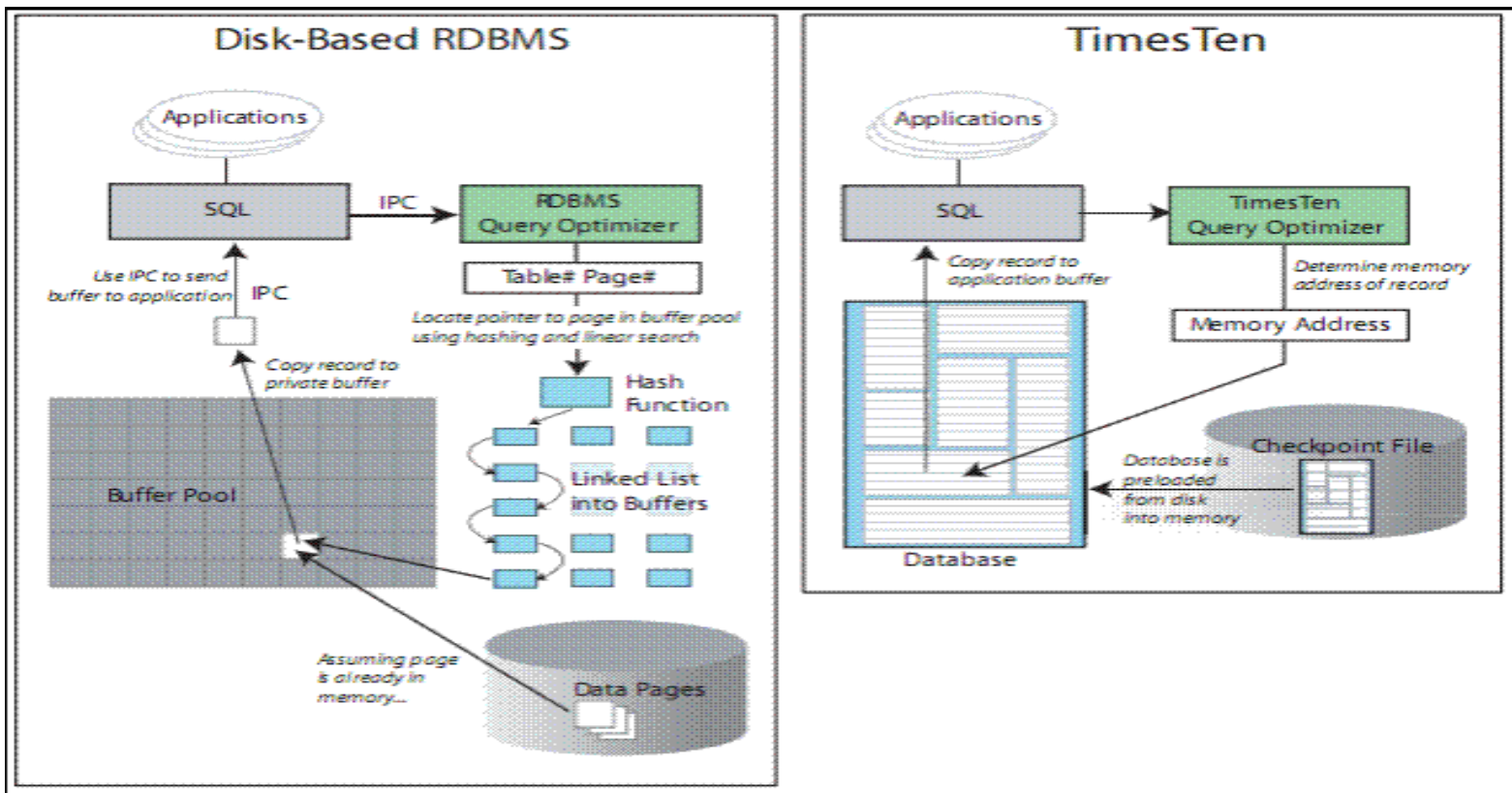
- 当前市场上呼声最大的内存数据库
- 特别适合SAP ERP加速



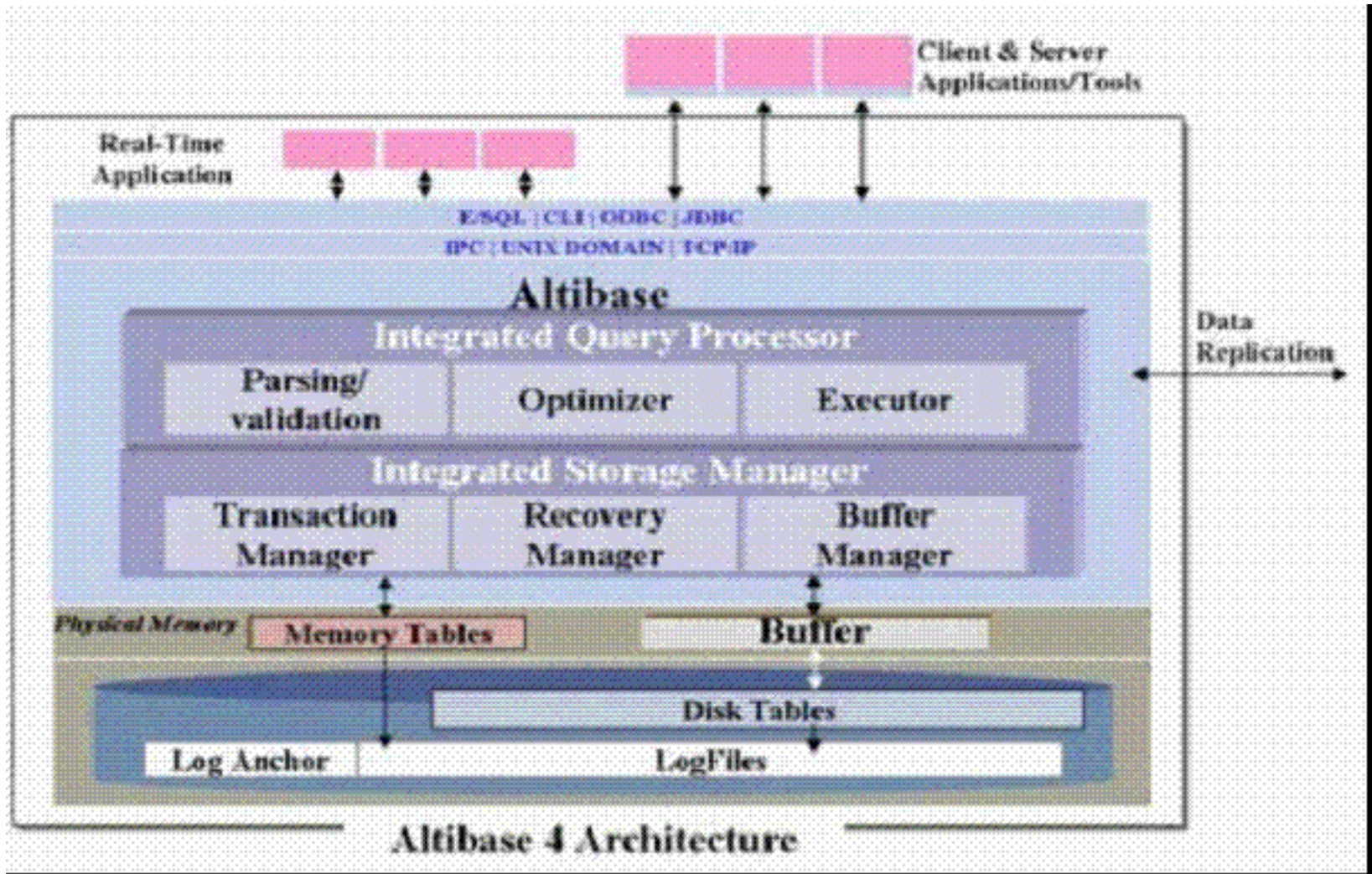
Databases	Applications	Formats
<ul style="list-style-type: none"> Oracle DB2 Sybase & IQ SQL Server Informix Teradata MySQL Netezza HP Neoview ODBC 	<ul style="list-style-type: none"> Oracle Apps PeopleSoft Siebel JD Edwards Salesforce.com SAP BW SAP ERP 	<ul style="list-style-type: none"> Text delimited Text fixed width EBCDIC XML Cobol Excel HTTP JMS SOAP

Oracle Timesten

- 内存优化的关系数据库；
- 支持服务器间的实时数据复制，以获得高可用性和负载共享；
- 应用程序层中的 Oracle 数据创建实时、可更新的高速缓存。



Altibase

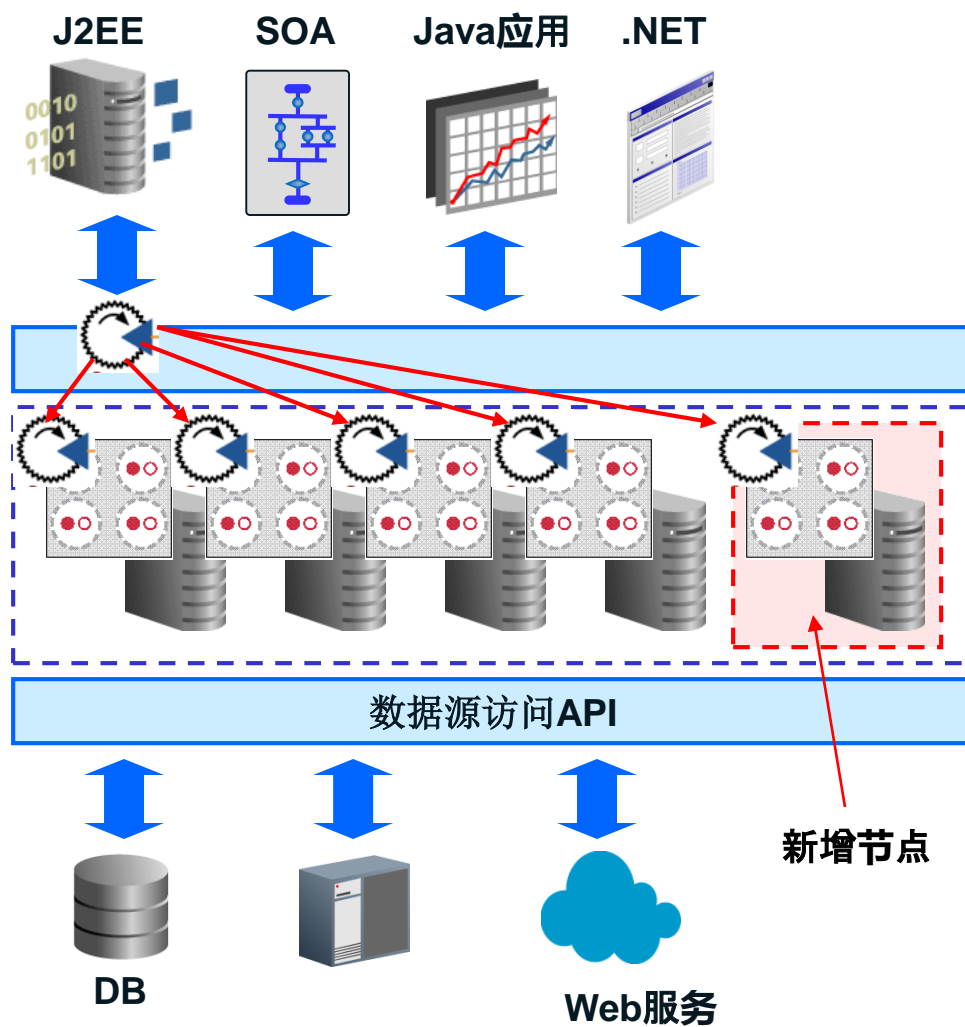


内存数据库使用分析

- 适用于主数据存储
- 和应用服务器部署在一起，起到应用加速的作用
- 一般都是active/standby模式，系统scale-out能力有限
- 为保证数据安全和事务，需要在磁盘保存内存映象和日志，磁盘的IOPS对内存数据库的影响巨大，建议用SSD盘或者PCI-E卡。
- 内存比较昂贵，很多厂商采用压缩技术，用计算能力换取存储能力，对CPU的依赖较高。
- 大部分的内存数据库为了保证容量和计算速度都采用4路及以上机器。

内存数据网格

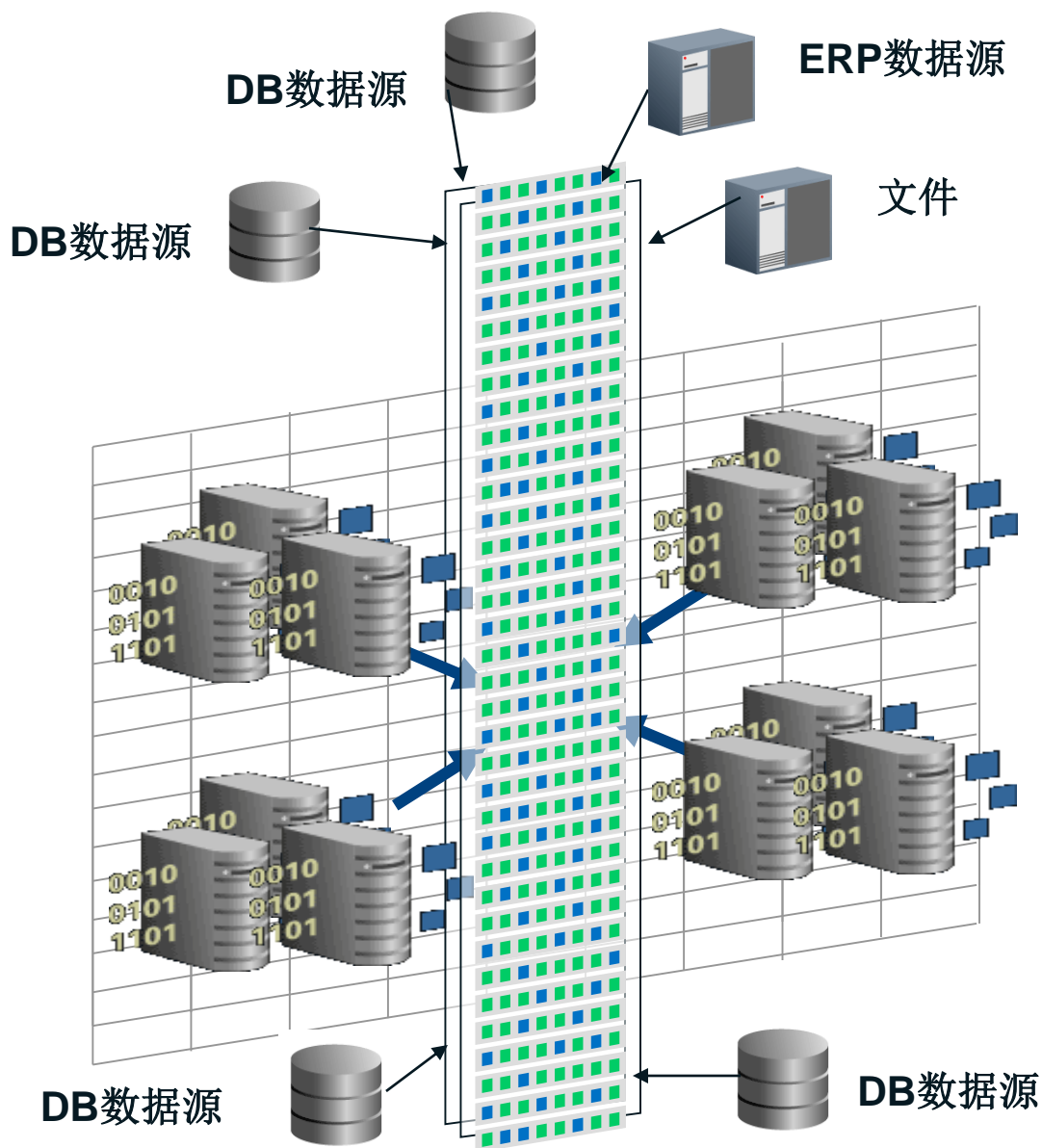
- 从高性能计算衍生出来的一种数据存储和计算方式，其实现方式是将从多台机器的内存通过高速网络连接起来，作为一个统一的**fusion cache**进行数据存储和数据计算，数据以**key/value**的方式作为对象存储在内存中
- 适用于计算量大，同时实时性高要求高的场景。
- 系统的可扩展性非常好，几乎没有机器台数的限制。
- 可以充分利用企业已有的计算资源，组成一个数据网格，通过数据网格技术发挥出 $1+1>2$ 的计算能力。
- 数据网格的数据源接口是透明的，可以通过**JDBC/sql**等接口或者适配器访问各种数据源。



新增节点

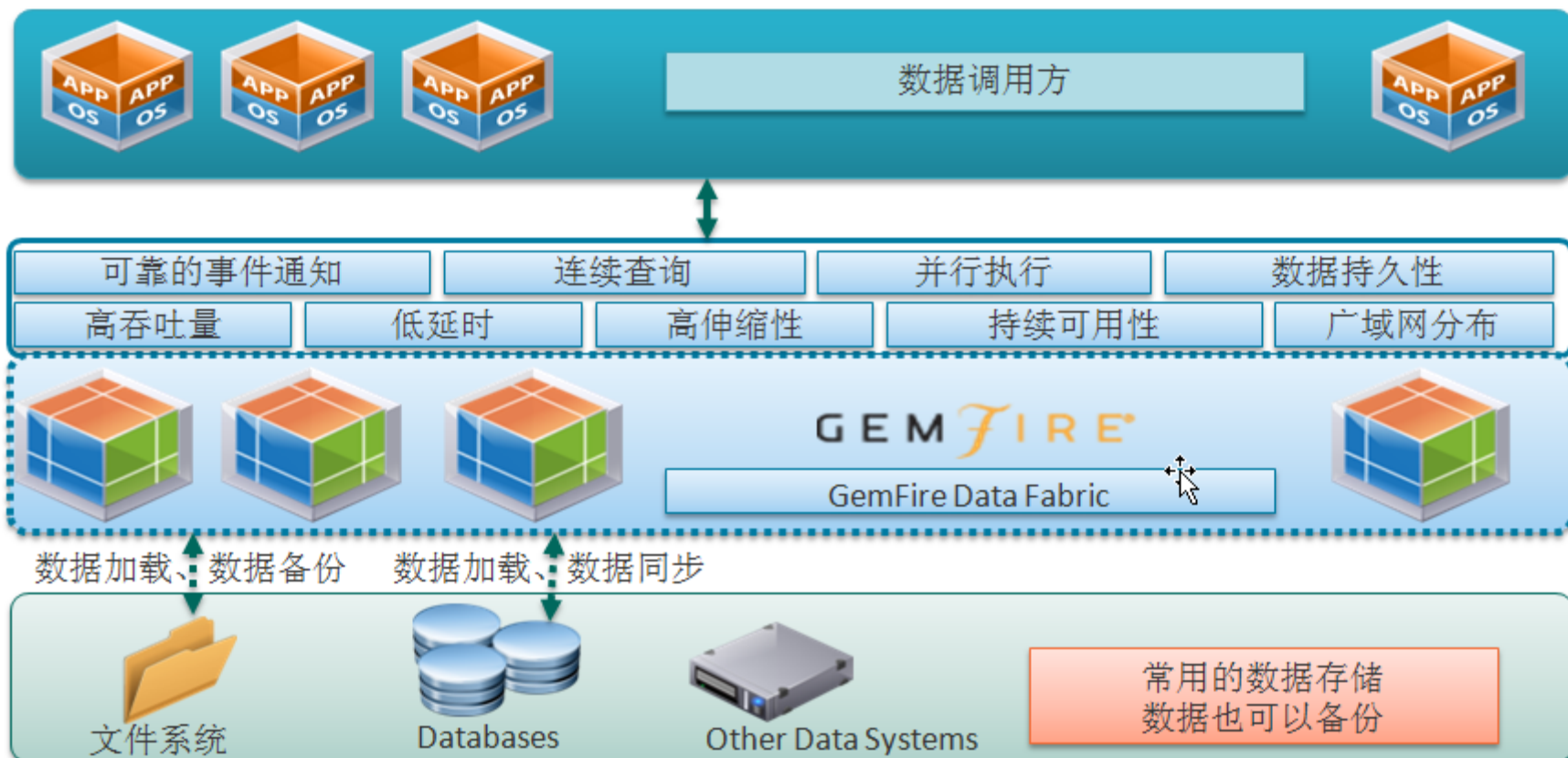
内存数据网格实现方式

- 多台服务器组成数据网格，将各种数据源的信息都抽取到数据网格中。
- 当需要计算时，所有的计算都在数据网格中进行，一方面数据都已经在内存中，不需要访问数据库，另一方面可以充分利用数据网格中服务器的计算能力。



Pivotal GemFire架构

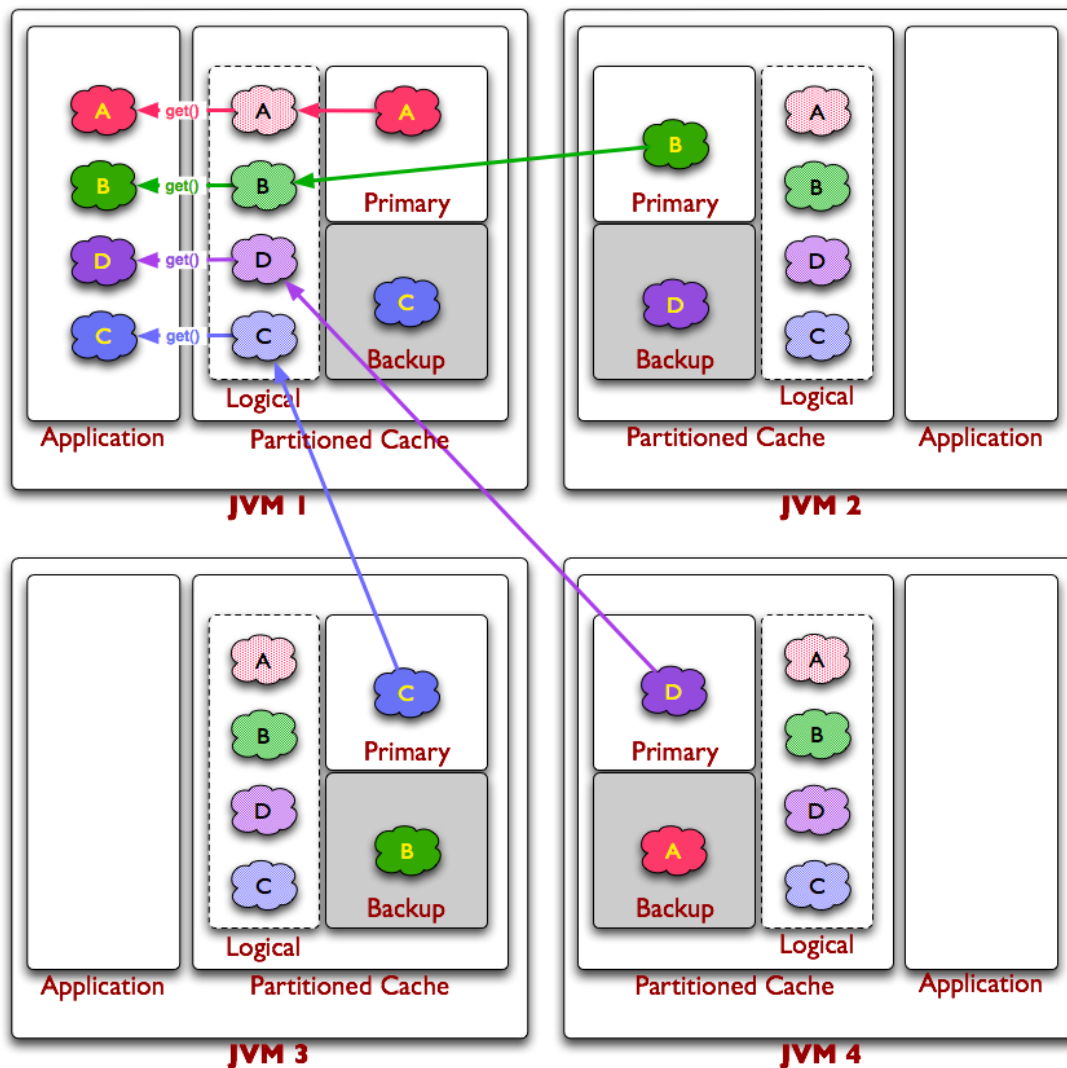
GemFire功能架构



Pivotal

数据网格数据分布

- 数据在集群成员中分布和备份
- 对开发人员透明
- 任何成员可以访问所有数据
- 所有数据位置都是已知的---不需要寻找和注册表!
- 系统的**scale-out**性能非常好，可以扩展到上千个节点



12306售票系统架构



讨论

- 内存数据库仍然以面向传统的关系型数据库的，适用于数据量不大，但读写要求很高的OLTP应用，HA一般采用active/standby方式，采用scale-up架构。
- 内存数据网格适用于偏向计算和前端海量并发的场景，扩展性好，对网络的依赖比较大，系统采用scale-out架构。
- 列式内存数据仓库如HANA, MonetDB(Vectorwise)是新兴的一种OLAP查询方式,但目前仍然以scale-up为主。

议程

- Scale up和scale out的定义
- 传统关系型数据库配置
- MPP DB数据库配置
- 内存数据库配置
- **Hadoop典型配置**

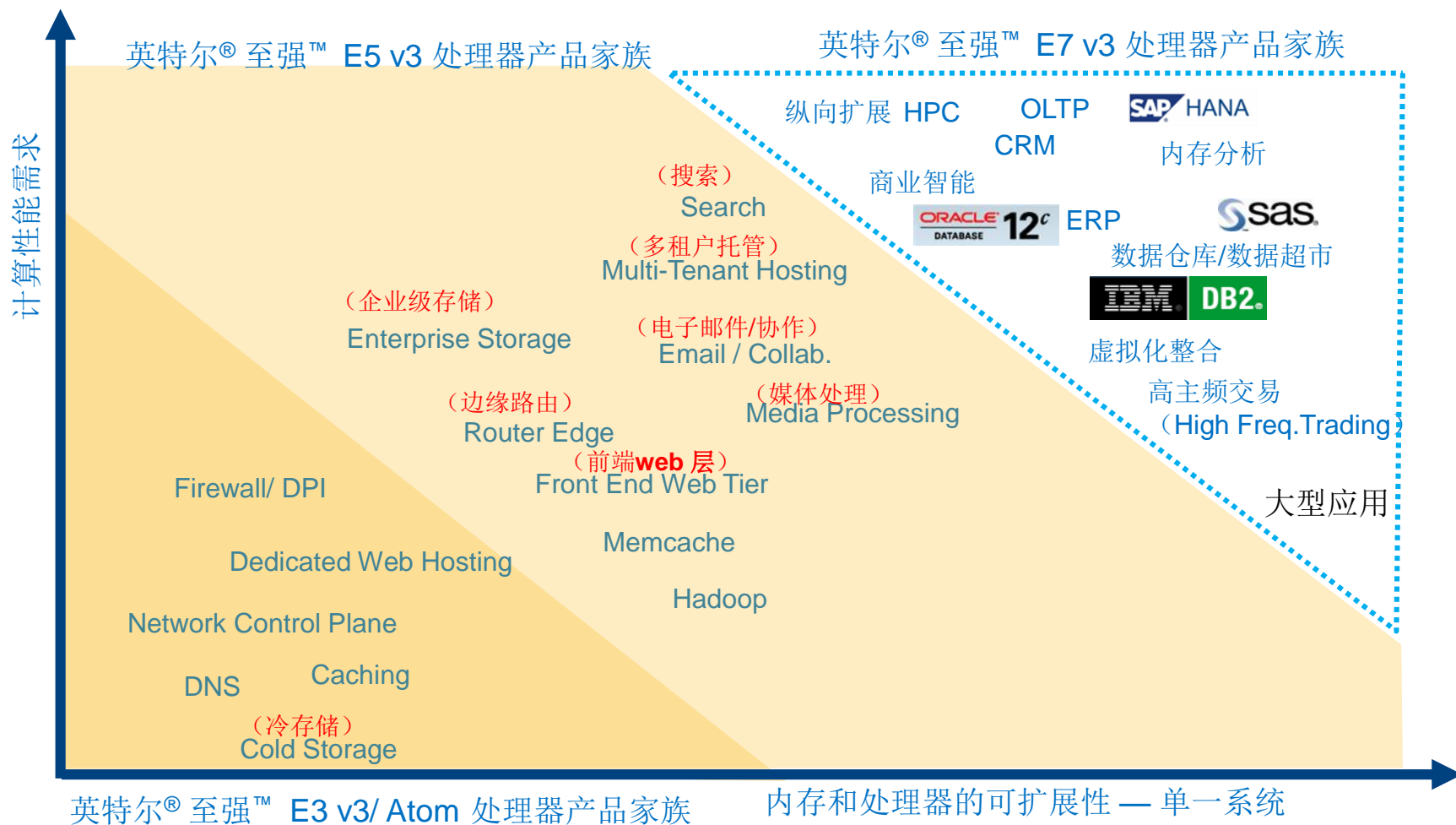
需要和应用结合

	CPU	网络	IO	内存
对象存储	2路E5(E3)	10GE	SATA	低
详单存储	2路E5-2630v3以上	10GE	SATA	中
数据处理	2路E5-2670v3以上	10GE	SSD/SAS	高

大数据系统的网络设计特别重要，保证跨柜的节点带宽
数据存储可以考虑分层：

- 热数据在内存
- 温数据在ssd
- 冷数据可以采用hdfs-raid存储

为您的负载找到最匹配的® 至强™ 处理器产品



* 其他的名称和品牌可能是其他所有者的资产。



英特尔存储产品概览



Intel® SSD DC P3700 Series



Intel® SSD DC P3600 Series



Intel® SSD DC P3500 Series

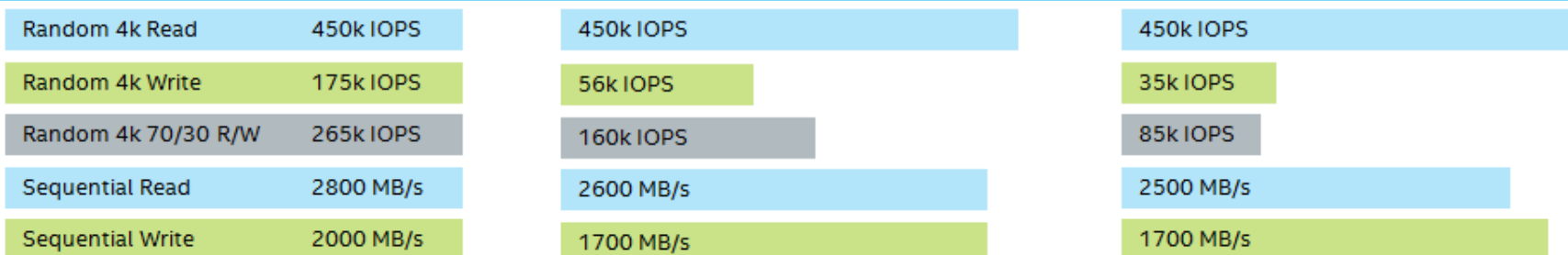
Capacity



Endurance



Performance



Sequential latency of 20µs

打造高效数据中心

服务器



网络



存储



最强性能
出色的每瓦性能



最高密度设计
最低功耗

至强提供对于计算、存储、网络理想的
通用平台

法律声明

- 在性能检测过程中涉及的软件及其性能只有在英特尔微处理器的架构下方能得到优化。诸如 **SYSmark** 和 **MobileMark** 等性能测试均系基于使用特定计算机系统、组件、软件、操作系统及功能。上述任何要素的变动都有可能测试导致测试结果的变化。请参考其他信息及性能测试（包括结合其他产品使用时的运行性能）以对目标产品进行全面评估。更多信息，敬请登录 <http://www.intel.com/performance>。
- 英特尔不对本文中引用的第三方基准数据或网站承担任何控制或审计的责任。您需要访问参考网站以确认所引用数据是否准确。
- SPEC**、**SPECint**、**SPECfp**、**SPECrate**、**SPECpower** 和 **SPECjbb** 是标准性能评估机构 (SPEC) 的注册商标。详情请参见 <http://www.spec.org>。
- 优化通知**
英特尔的编译器针对非英特尔微处理器的优化程度可能与英特尔微处理器相同（或不同）。这些优化包括 **SSE2**®、**SSE3** 和 **SSSE3** 指令集以及其它优化。对于在非英特尔制造的微处理器上进行的优化，英特尔不对相应的可用性、功能或有效性提供担保。此产品中依赖于处理器的优化仅适用于英特尔微处理器。某些不是专门面向英特尔微体系结构的优化保留专供英特尔微处理器使用。请参阅相应的产品用户和参考指南，以了解关于本通知涉及的特定指令集的更多信息。
- 英特尔处理器标号不作为衡量性能的标准。处理器标号仅用于区分同属一个产品家族的处理器特性，而不能够用于区分不同产品家族的处理器：详情敬请登录：
http://www.intel.com/products/processor_number
- 英特尔® 高级矢量扩展指令集（英特尔® **AVX**）*可为某些处理器操作提供更高的吞吐量。由于不同的处理器功耗特点，利用 **AVX** 指令可能会导致 a）一些部件运行的频率低于额定频率 b）一些采用英特尔® 睿频加速技术 2.0 的部件无法获得任何或最大的睿频频率。具体性能在不同的硬件、软件和系统配置上将有所差异，您可以访问以下网站了解更多信息：<http://www.intel.com/go/turbo> ○
- 任何计算机系统都无法提供绝对的安全性。需要使用支持的处理器、芯片组、固件和软件。请联系您的制造商或零售商了解更多信息。
- 任何计算机都无法提供绝对的安全性。您需要使用支持的计算机和软件。在某些情况下，您需要对服务提供商的服务进行订阅。请确保您所在的区域设立了服务提供商。因此导致的任何数据或系统的丢失或被盗或其它损失，英特尔将不负任何责任。请联系您的零售商进行确认。
- 英特尔® 虚拟化技术要求计算机系统搭载兼容的英特尔® 处理器、最新的基本输入输出系统（BIOS）及虚拟机监控器（VMM）。其运行、性能及其他表现取决于硬件及软件的配置。相关应用软件可能无法与所有的操作系统兼容。请咨询您的系统制造商。了解更多信息：<http://www.intel.com/go/virtualization> ○
- 英特尔公司 © 2014 年版权所有。所有权保留。英特尔、Intel 标识、Xeon、至强、Atom、凌动、Intel Core 和英特尔酷睿是英特尔公司在美国和其他国家(地区)的商标。
- 所有日期和产品仅用于规划目的，可随时更改，恕不另行通知。
- * 其他的名称和品牌可能是其他所有者的资产。