

The background features a white canvas with scattered, semi-transparent blue and pink geometric shapes, including triangles and squares. Several vibrant, multi-colored lines (pink, blue, green, orange) cut across the scene from the corners towards the center, creating a dynamic, modern aesthetic.

1169 TECH

Apache Ignite 内存计算

2017.06.23

李照宇 青岛·海尔

Agenda

- Apache Ignite 项目简介
- Ignite场景和组件
- 为数据分析加速
- Q&A

Apache Ignite 项目

apache / ignite
mirrored from git://git.apache.org/ignite.git

Watch 144

★ Unstar 1,006

🍴 Fork 593

<> Code

Pull requests 548

📁 Projects 0

Insights ▾

Contributors

Commits

Code frequency

Punch card

Network

Members

Dependents

Feb 16, 2014 – Jun 23, 2017

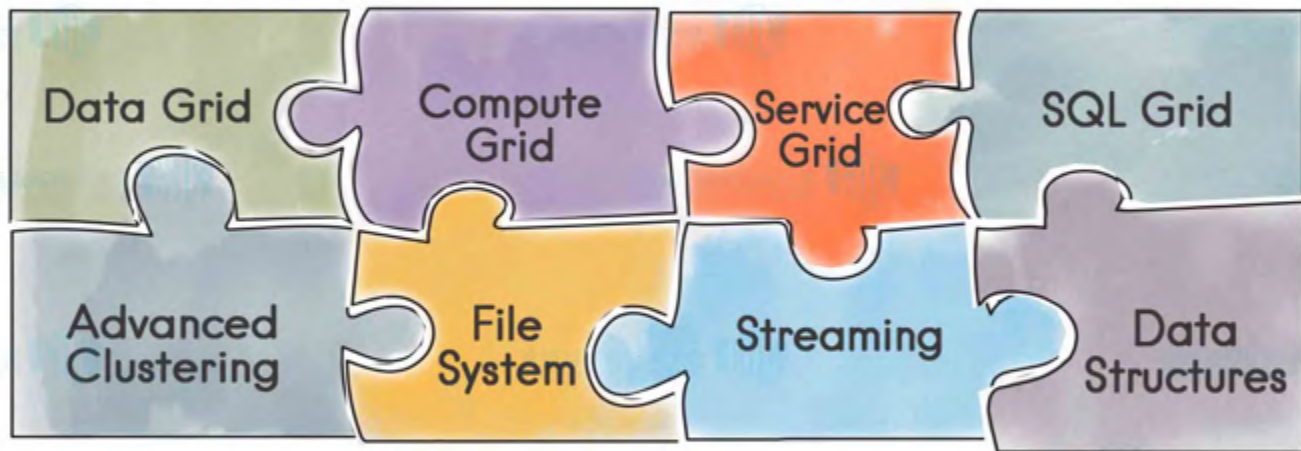
Contributions: Commits ▾

Contributions to master, excluding merge commits



● Java 72.9% ● C# 12.3% ● C++ 6.6% ● JavaScript 2.8% ● Scala 1.7% ● Shell 1.5% ● Other 2.2%

IMDF



高性能的分布式内存计算平台，可以用于处理各种大规模数据集



server

server

server

server

server

server

SQL

NoSql

Hadoop /
BI



server

server

server

server

server

server

caching?

computing?

searching?

message?

SQL

NoSql

Hadoop /
BI

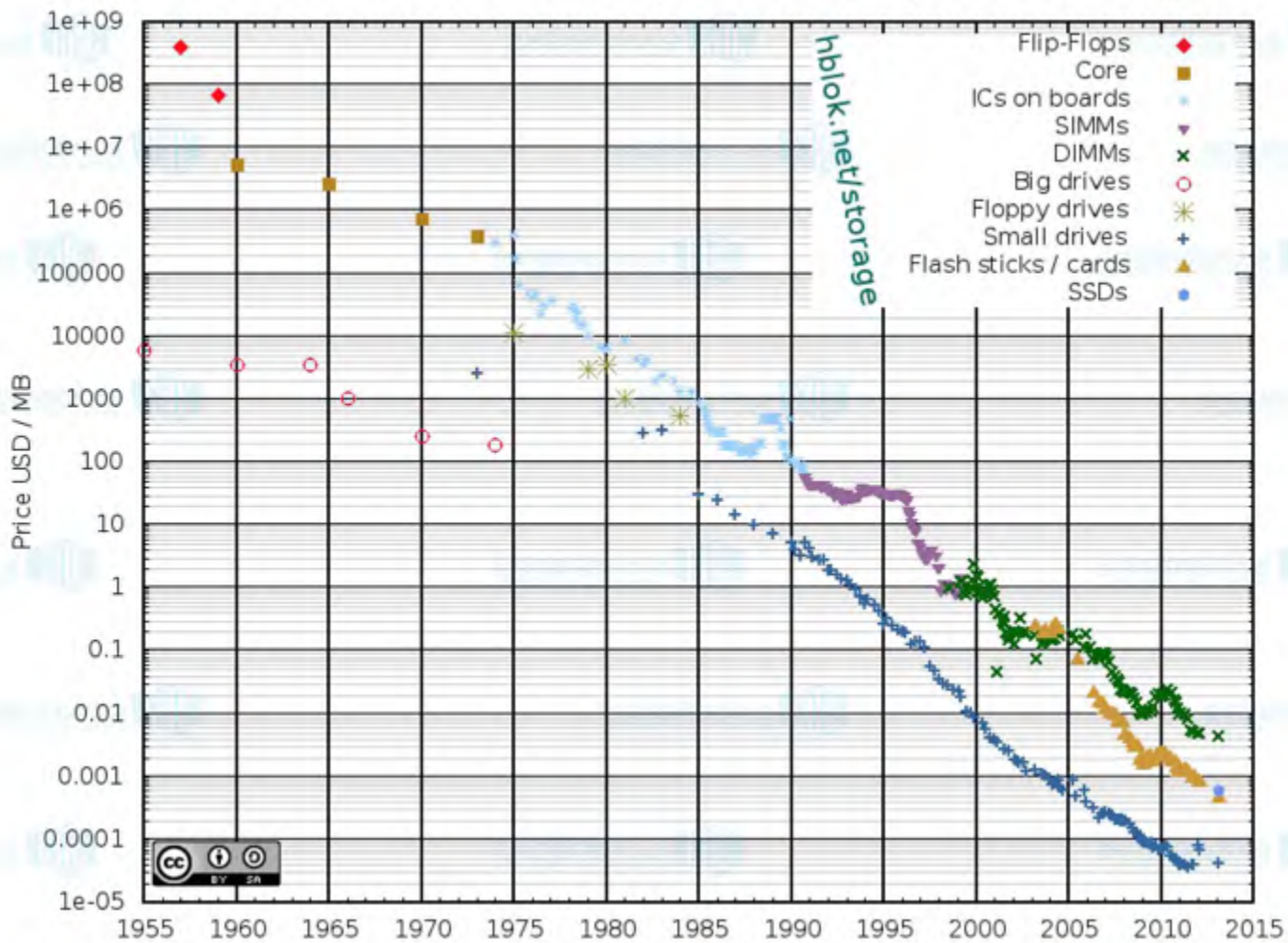
CUSTOMERS WHO BOUGHT THIS ITEM:



ALSO BOUGHT:



Historical Cost of Computer Memory and Storage



” Ram is the new disk, and disk is the new tape

Gartner.



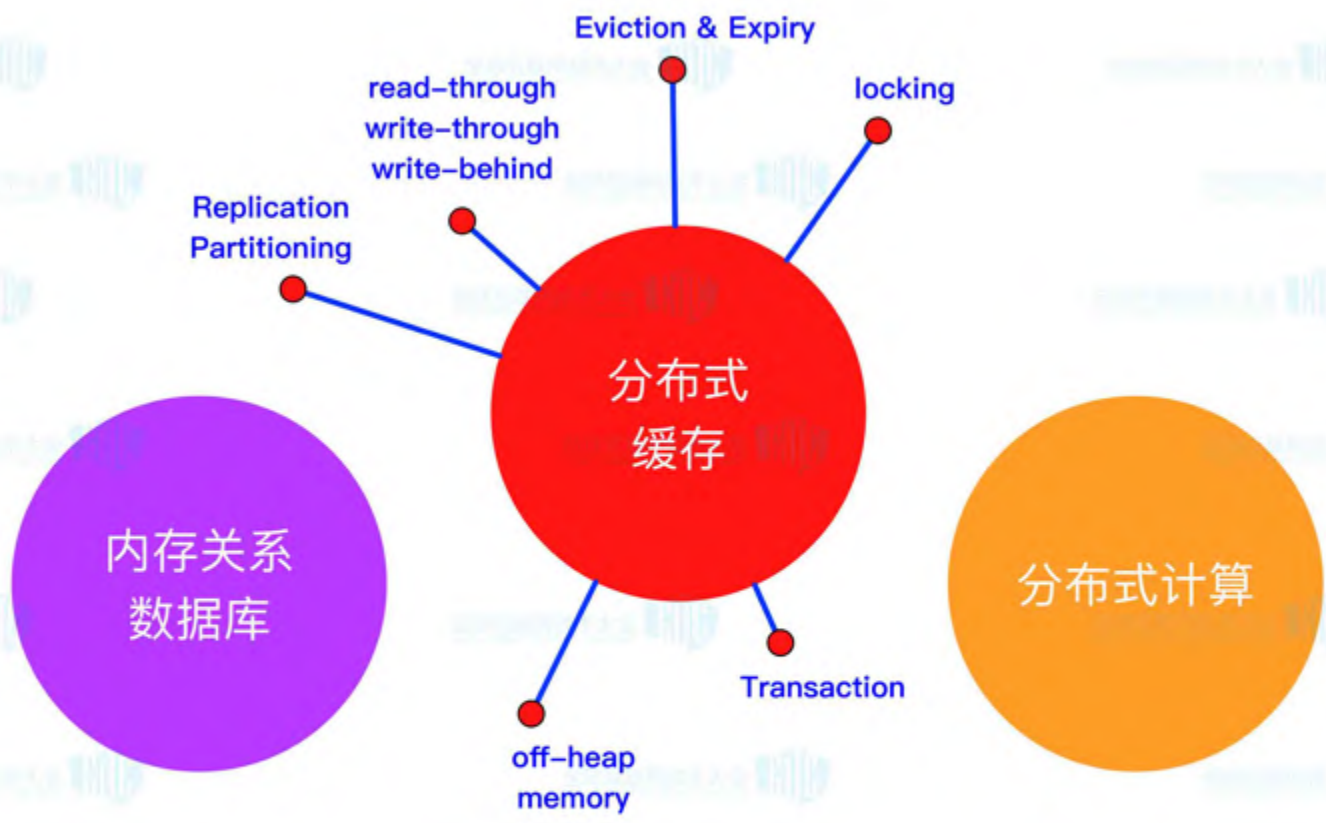
- API call <> OS IO<> 控制器 <> 硬盘
- 延时：毫秒级 ($10e-3$)

- API call <> 指针运算（内存访问）
- 延时：纳秒级或者微秒级 ($10e-6, 10e-9$)

内存关系
数据库

分布式
缓存

分布式计算



Eviction & Expiry

read-through
write-through
write-behind

locking

Replication
Partitioning

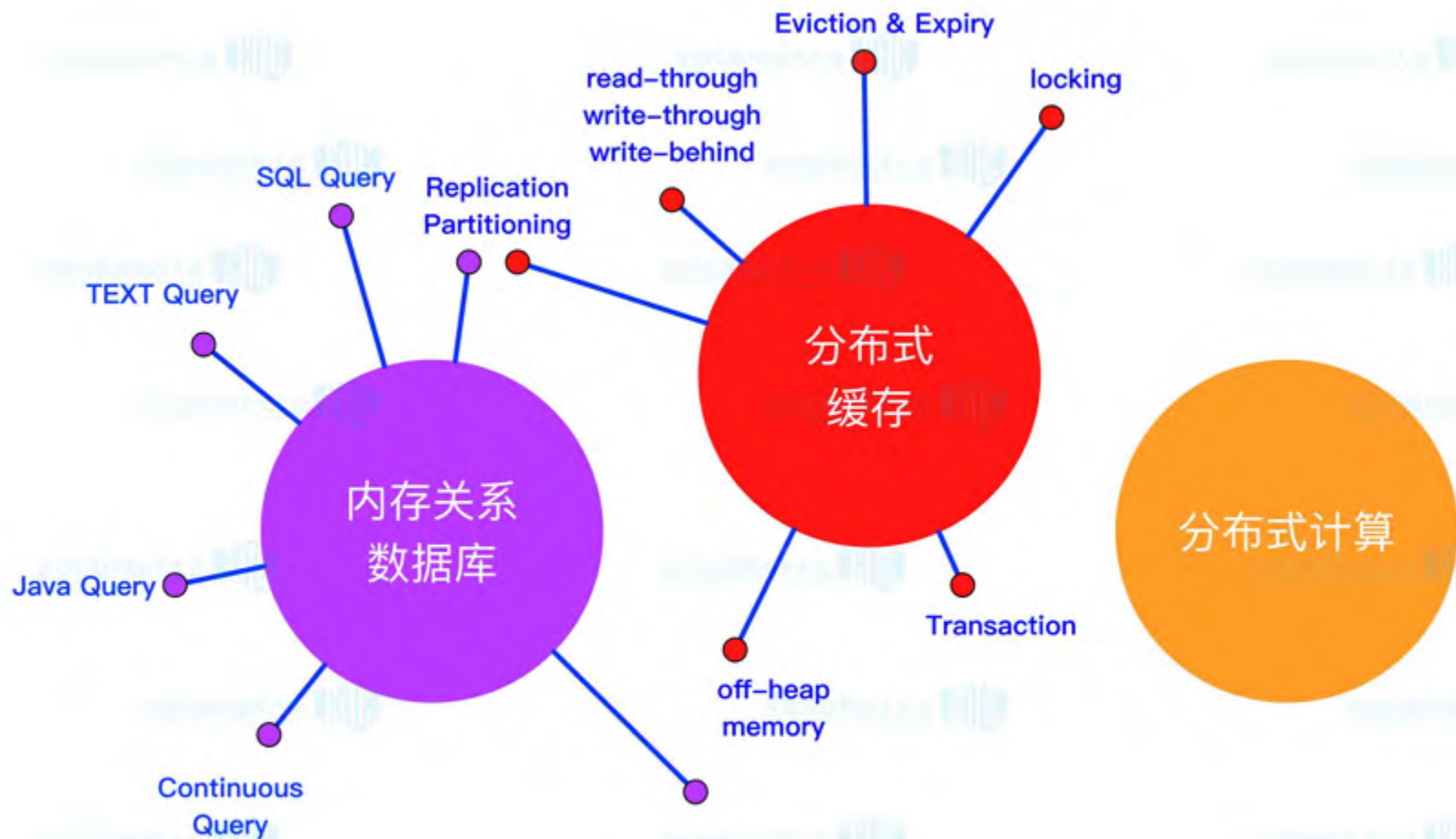
分布式
缓存

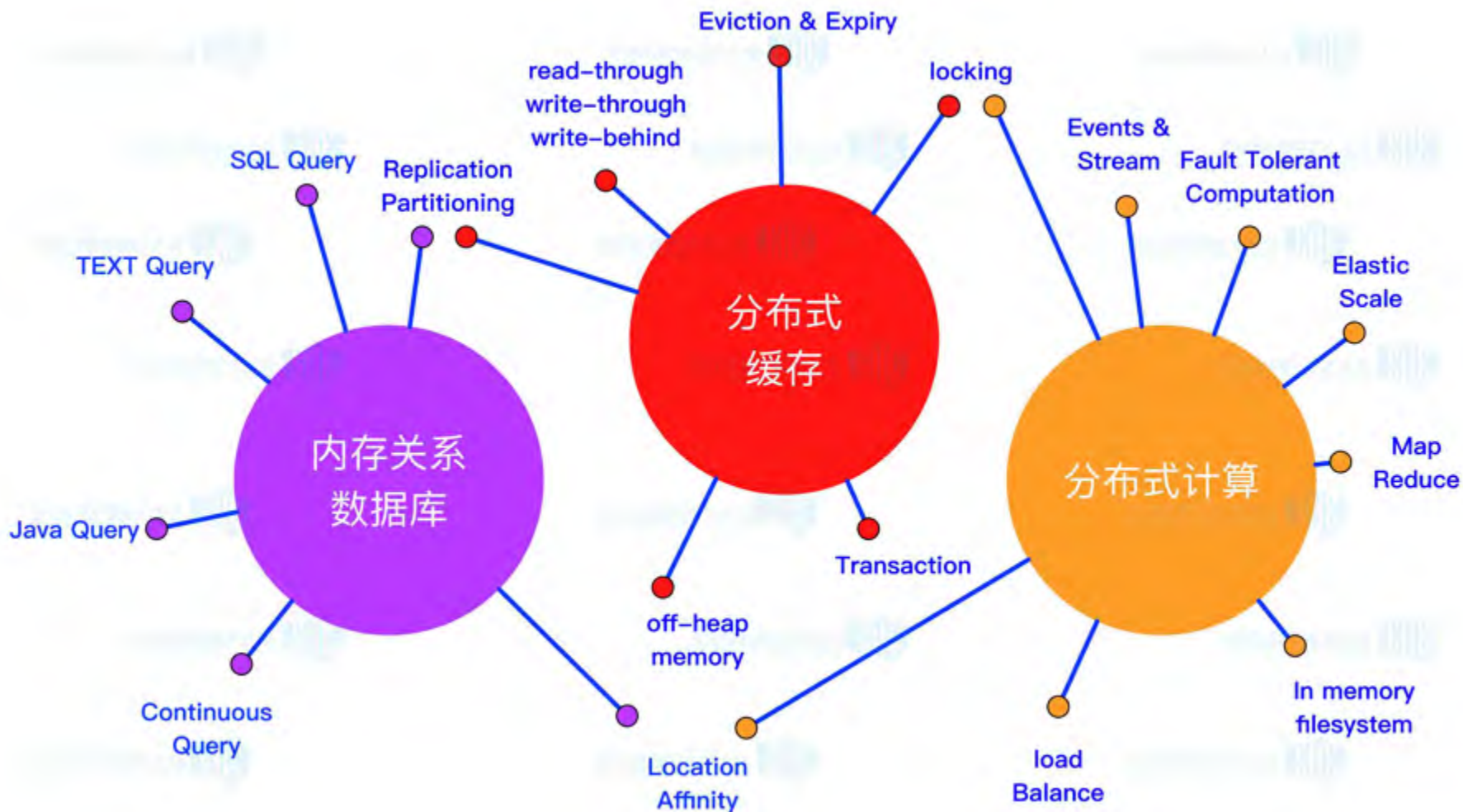
内存关系
数据库

分布式计算

Transaction

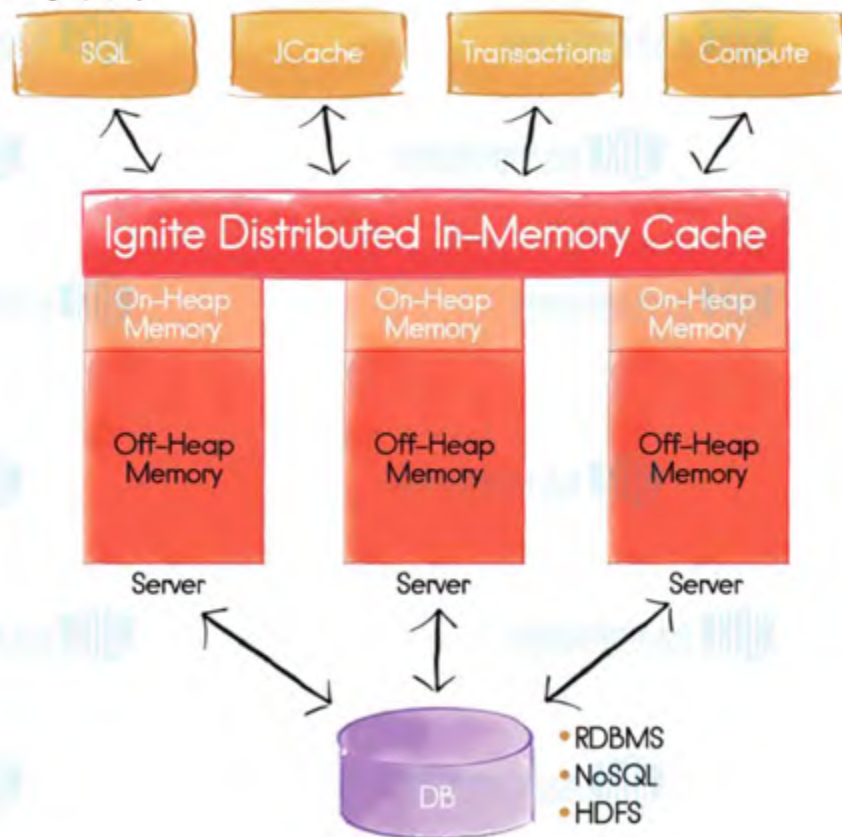
off-heap
memory



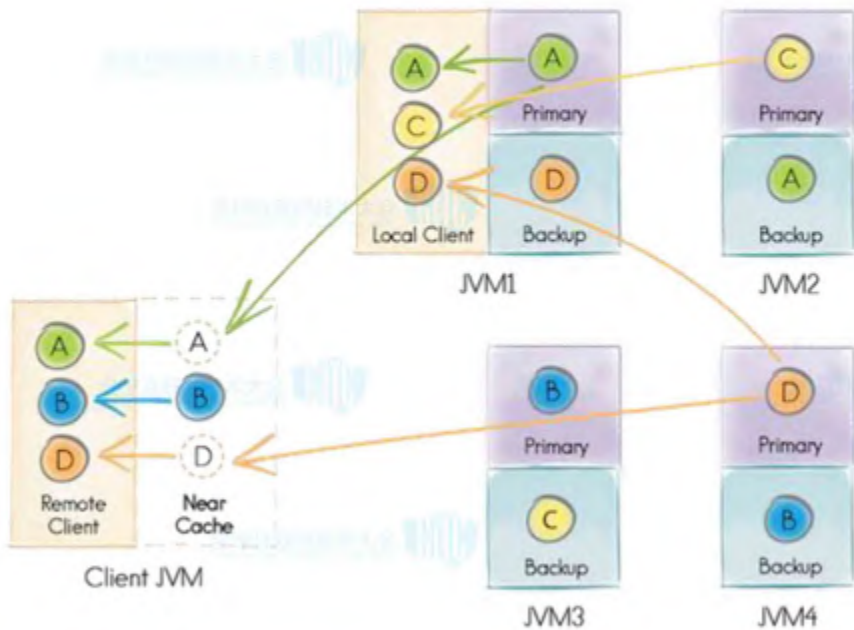


数据节点

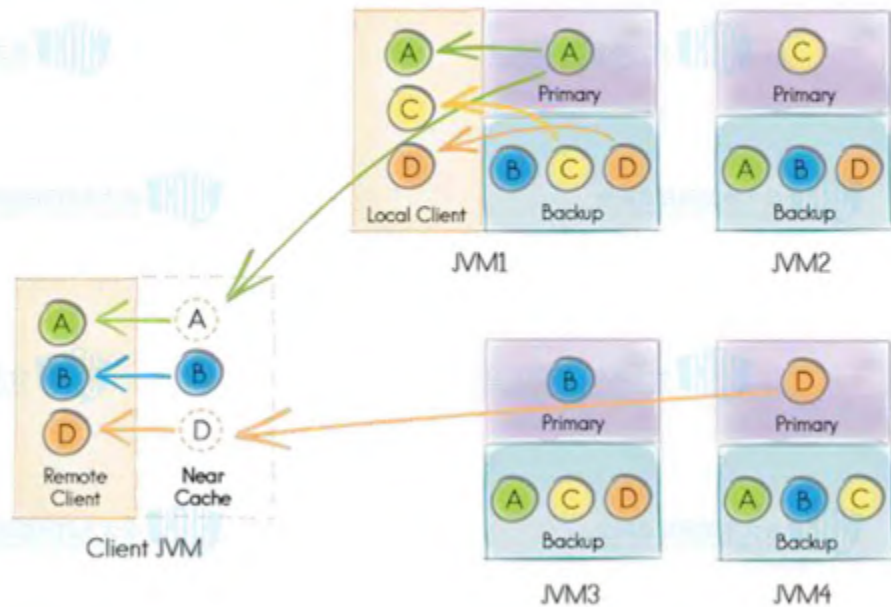
- 分布式k-v存储
- In-memory index
- On-heap/off-heap
- Tb级的任意格式的数据
- 自动 Failover
- 分布式ACID事务
- ANSI99 sql query
- JDBC Driver
- 可选的持久化存储



Partitioned Cache



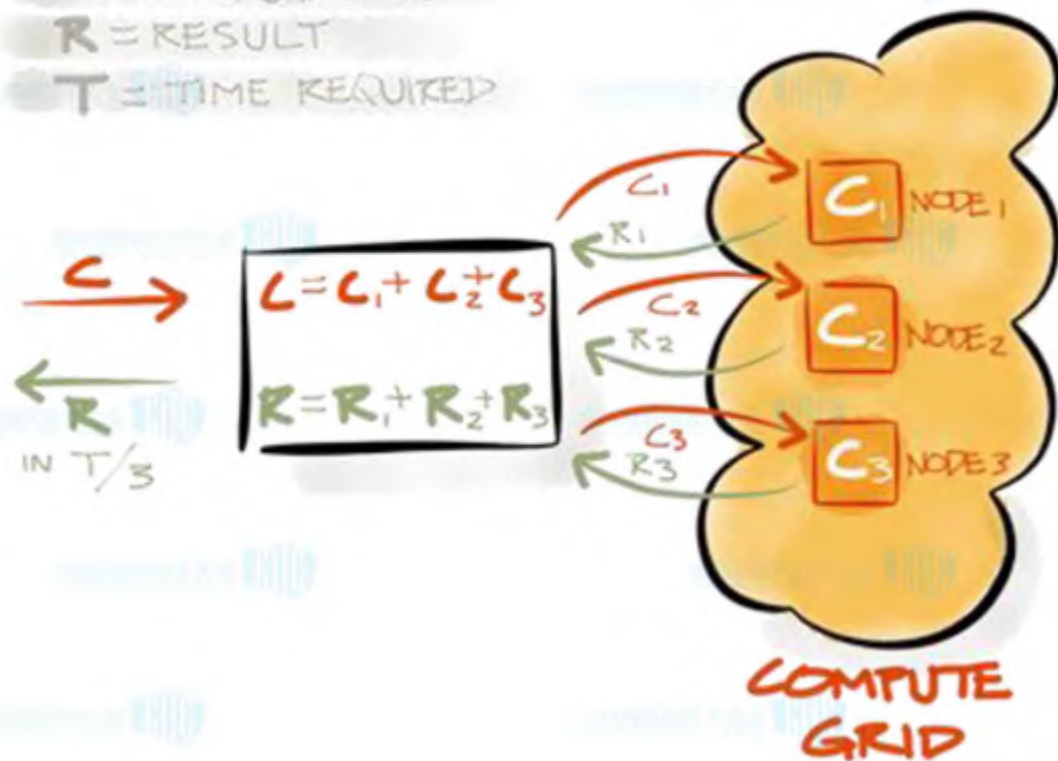
Replicated Cache



计算节点

- MapReduce
- 零部署
- 类似cron的定时任务
- State-checkpoint
- 负载均衡
- 自动Failover

C = COMPUTATION
 R = RESULT
 T = TIME REQUIRED

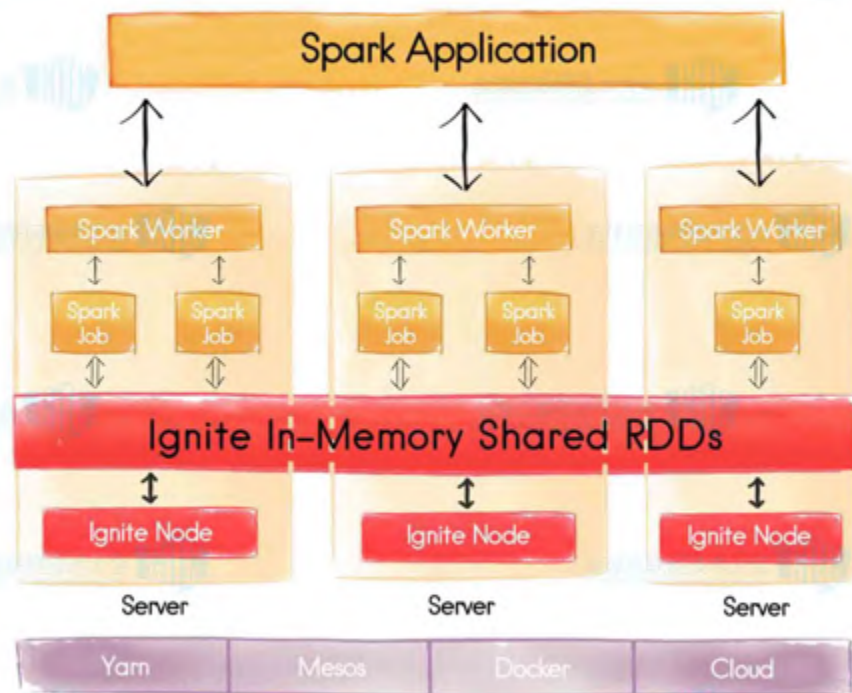


Ignite & SQL

- ANSI-99 SQL
- H2 引擎
- 内存索引(in/off memory)
- 分布式的集合函数、分组、排序
- 跨node的Join & Union
- 基于JVM的函数扩展(Java, Clojruе,Scala)

为spark 提速

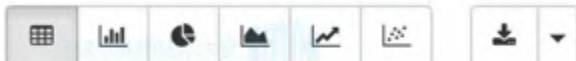
- Sharded in memory
RDD



<https://github.com/apacheignite/zeppelin-demo>



```
%ignite.ignitesql
SELECT o.name as Organization, avg(p.salary) as Salary
FROM Person p, "Organizations".Organization o
WHERE p.orgId = o.id
AND o.id IN (12553, 444134, 5173)
GROUP BY o.name
ORDER BY Salary ASC
```



organization

Organization444144

Organization5673

Organization12653

```
%sql
SELECT o.name as organization, avg(p.salary) as salary
FROM Person p, Organization o
WHERE p.orgId = o.id
AND o.id IN (12653, 444144, 5673)
GROUP BY o.name
ORDER BY salary DESC
```



organization

Organization444144

Organization5673

Organization12653

The background features a repeating pattern of a faint logo consisting of a stylized 'H' with vertical lines. Scattered throughout are various colorful geometric shapes: triangles, squares, and lines in shades of blue, green, orange, pink, and purple. Some shapes are larger and more prominent, while others are smaller and more subtle.

1169 TECH

谢谢观赏