



第九届中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2018

高性能K-V数据库 Aerospike

刘喆@精硕科技

DTCC
2018

2018.05.10 – 12 北京国际会议中心



IT168.com

ChinaUnix

ITPUB

✧ 简介

✧ Aerospike 指北

✧ Aerospike best practice

✧ Aerospike在Admaster应用

是为简介

- 刘喆
 - 2010 年百度 hadoop
 - 2011 年人民搜索 运维部
 - 2013 年精硕科技 基础架构
- 精硕科技
 - 广告
 - 营销全流程
 - 计划
 - 投放
 - 度量
 - 社交
 - 人群画像



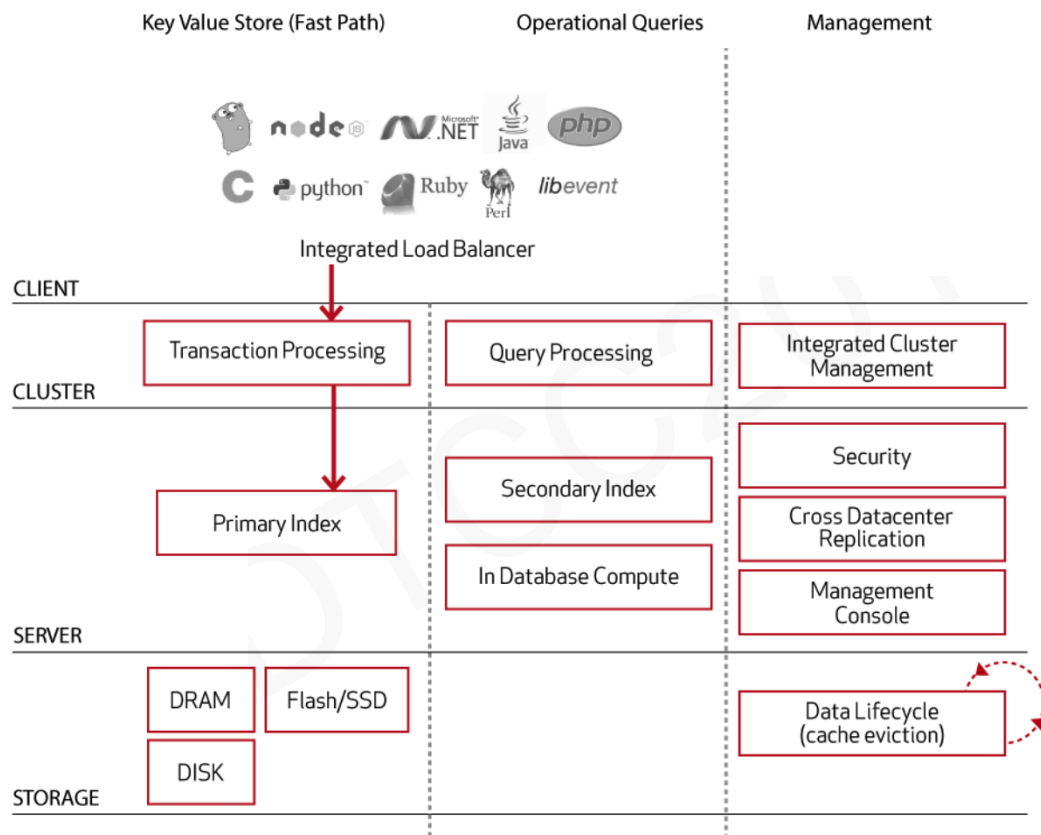
Aerospike指北之概述

- Aerospike是一款分布式的NoSQL数据库和k-v存储数据库。在具备高性能的同时，具备强一致性和鲁棒性，且具备以下优点：
 - ✓ 分布式架构，sharding冗余策略。
 - ✓ Indexes in RAM模式，并支持数据存储RAM+DISK模式。
 - ✓ 极易水平扩展和集群sharding自动balance功能。
 - ✓ 强大的UDF功能和secondary index支持。
 - ✓ 自定义高水位和自动data eviction策略。
 - ✓ 强大的集群高并发支撑。

Aerospike指北 之 术语

- **Clients:** 提供数据读写api，通过TCP/IP协议链接。并具备确认数据存储位置和集群节点变化等功能。
- **Cluster:** 通过无共享架构实现高可靠存储和数据的自动故障恢复、副本策略和跨集群同步。
- **Storage:** 数据存储支持内存、SSD和传统机械磁盘。

Aerospike指北之体系结构



Aerospike指北 之 术语

- **NameSpace**: Aerospike中最上层的数据存储空间。可以类比为传统关系型数据库中的database级别。
- **Set**: 存在于namespace中, 用于存储记录的一个逻辑空间。并继承namespace的配置策略。可类比为传统数据库中的table。
- **Record**: 实际的存储记录, 可类比为传统数据库中的行。一个record属于特定的set和namespace。并通过key来作为唯一标识。
 - ✓ **Key**: 是record中的唯一标识, 并通过key的hash值来确定存储位置。
 - ✓ **Bins**: 类似于传统数据库中的字段, 与key形成映射。每个bin可拆分为binName、binValue。



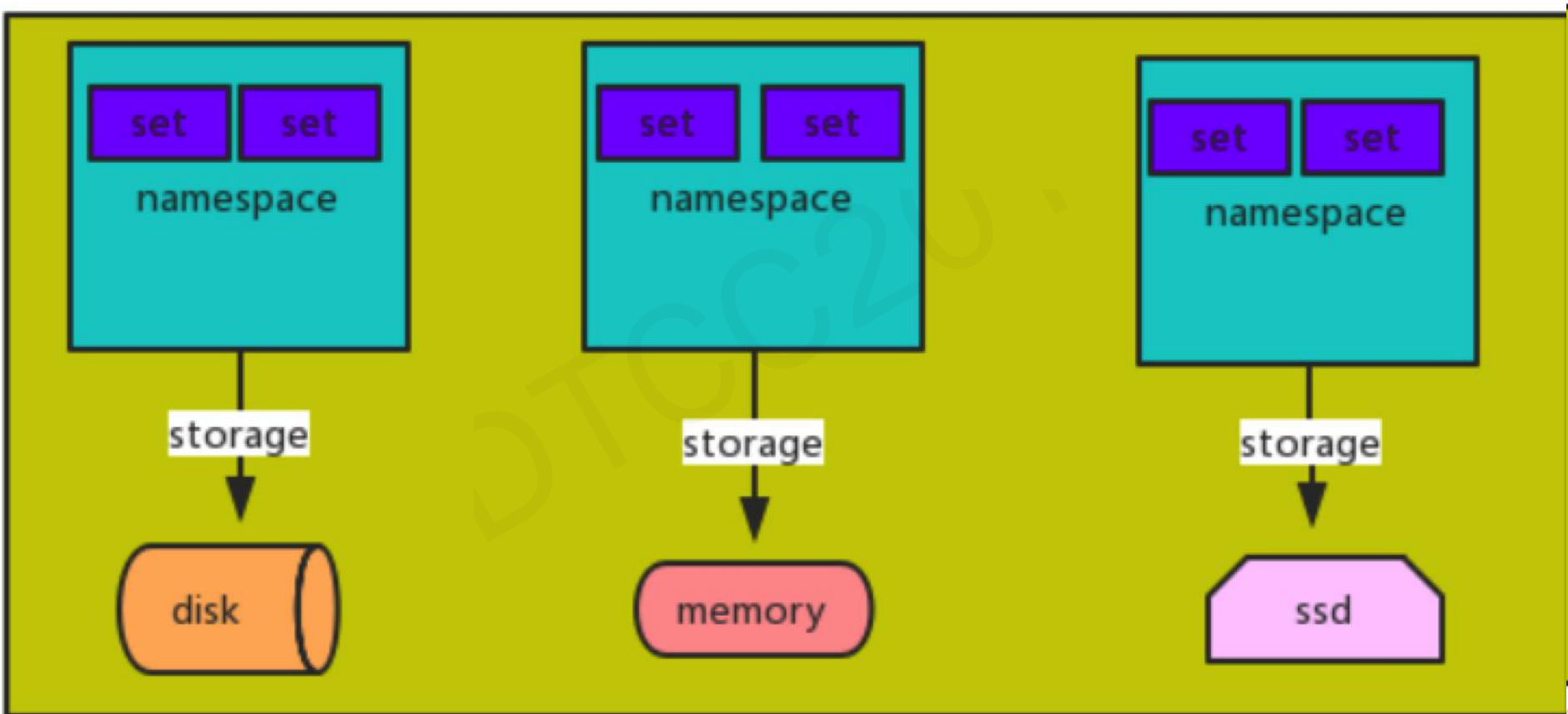
Aerospike指北 之 术语

- **Primary Index:** 在Aerospike中，primary index是整合了hash table和tree结构并平均分布在每个server上。
- **Secondary Index:** 是建立在bin的列中，可类比于传统数据库中列索引的概念。跟primary index相同，也是index in RAM存储，并且只包含本节点的索引和数据引用。
- **Hybrid Storage:** aerospike集群中索引和数据在没个节点中存储，在物理存储选择可以有内存、SSD、普通磁盘等模式可以选择，并可以混合搭配，aerospike提供了不通场景下数据清除和碎片整理等优化策略。



Aerospike存储结构

cluster



Aerospike指北 之 安装

- Aerospike Server: aerospike server端安装，组成aerospike集群。需要规划namespace及其存储相关策略。
- Aerospike AMC: 用于可视化监控aerospike集群，包括实时QPS、节点内存和磁盘使用率等等。
- Aerospike Tool: 主要包含aerospike shell工具。
- Aerospike Client: 提供各种语言的Client Library。

Aerospike指北 之 安装

➤ Aeorpsike集群监控界面:



Aerospike指北 之 安装

➤ Aeorpsike Tool:

```
aql> select * from data2.r_0 where pk = 'cctest188'
+-----+
| by_cc |
+-----+
| 188   |
+-----+
1 row in set (0.002 secs)
```

通过aql可以方便的通过sql进行数据检索。

Aerospike指北 之 运行部署

- Admaster中集群部署方式采用ansible进行自动化部署，并所有服务进程利用supervisor进行运行监控，所有服务随机启动。
- Aerospike心跳通信方式采用mesh，数据存储模式分为纯内存和内存+SSD模式。
- 根据业务数据级别综合参考，设置副本数和存储模式。

Aerospike指北 之 运行部署

Admaster集群各namespace分配样例：

	Namespace	Master Objects	Replica Objects	Rep'l'n Factor	Avail%	Disk	RAM	Expired Objects	Evicted Objects
+	blueair Per-node Details	10,000,000	20,000,000	3	N/A	N/A	4.06 GB 55.94 GB	0	0
+	buzz Per-node Details	10,000,000	20,000,000	3	N/A	N/A	1.46 GB 58.54 GB	0	0
+	data1 Per-node Details	10,000,000	20,000,000	2	99%	0 B 750.00 GB	0 B 75.00 GB	0	0
+	data2 Per-node Details	10,000,000	20,000,000	2	N/A	N/A	17.39 GB 42.61 GB	0	0
+	dataservice Per-node Details	10,000,000	20,000,000	2	98%	217.68 MB 749.79 GB	2.25 MB 75.00 GB	0	0
+	dedup Per-node Details	10,000,000	20,000,000	2	88%	62.13 GB 777.87 GB	14.92 GB 105.08 GB	0	0
+	superspider Per-node Details	10,000,000	20,000,000	2	95%	96.49 GB 2.84 TB	2.69 GB 72.31 GB	0	0
+	track Per-node Details	10,000,000	20,000,000	2	93%	62.77 GB 1.40 TB	14.59 GB 210.41 GB	0	0



Aerospike指北 之Development

- Aerospike目前支持Java、C#、C、Python、Go等主流的开发语言，均具备较高的运行性能。
- 支持Lua语言的UDF开发，实现在Aerospike server端的相关计算，减少网络开销和客户端计算资源。
- Aerospike支持多种connector，包括Spring Data，hadoop，kafka等。

Aerospike指北 之redis比较

- 高并发支持: Aerospike对高并发支持更好, 集群中每个节点可支撑2w+连接, QPS可达600W+。相比redis较高
- 副本策略支持: Aerospike支持可自定义的副本策略, 保障高可用。Redis目前不支持副本。
- 数据TTL支持: Aerospike支持k-v级别的tt支持, 并支持自定义高水位后的数据自动驱逐策略, 与redis相同。
- 数据存储: Aerospike支持多模式数据存储, 可同时支持MEM+DISK模式, redis为MEM模式。Aerospike停机重启时数据可自动加载。

Aerospike best practice

➤ deployment:

- ✓ 自定义high water pct和stop writes pct相关
- ✓ Namespace storage engine
- ✓ Mesh heartbeat
- ✓ Tcp连接数相关

➤ Development:

- ✓ Sync client和Async client
- ✓ Max connect per node
- ✓ Batch multiple operations
- ✓ policy



Aerospike在Amdaster的应用案例1

- Track扩充日志缓存
 - ✓ 黑名单数据实时查询
 - ✓ Poi信息计算查询
 - ✓ 任务信息缓存
 - ✓ 广告位信息查询

DTCC2018

Aerospike在Amdaster的应用案例2

➤ 大爬虫数据缓存

- ✓ 抓取任务结果存储（包括页面文档信息）
- ✓ 抓取任务元信息存储
- ✓ 抓取任务状态查询
- ✓ 抓取任务去重

DTCC2018

THANKS





讲师申请

联系电话（微信号）：18612470168

关注“ITPUB”更多
技术干货等你来拿~

与百度外卖、京东、魅族等先后合作系列分享活动



让学习更简单

微学堂是以ChinaUnix、ITPUB所组建的微信群为载体，定期邀请嘉宾对热点话题、技术难题、新产品发布等进行移动端的在线直播活动。

截至目前，累计举办活动期数60+，参与人次40000+。

ITPUB学院

ITPUB学院是盛拓传媒IT168企业事业部（ITPUB）旗下
企业级在线学习咨询平台
历经18年技术社区平台发展
汇聚5000万技术用户
紧随企业一线IT技术需求
打造全方式技术培训与技术咨询服务
提供包括企业应用方案培训咨询（包括企业内训）
个人实战技能培训（包括认证培训）
在内的全方位IT技术培训咨询服务

ITPUB学院讲师均来自于企业
一些工程师、架构师、技术经理和CTO
大会演讲专家1800+
社区版主和博客专家500+

培训特色

无限次免费播放
随时随地在线观看
碎片化时间集中学习
聚焦知识点详细解读
讲师在线答疑
强大的技术人脉圈

八大课程体系

基础架构设计与建设
大数据平台
应用架构设计与开发
系统运维与数据库
传统企业数字化转型
人工智能
区块链
移动开发与SEO



联系我们

联系人：黄老师
电话：010-59127187
邮箱：edu@itpub.net
网址：edu.itpub.net
培训微信号：18500940168