

Welcome

JD.COM 京东



京东技术 开放日

JD Technology OpenDay

交流 · 分享 · 融合 · 创新

大流量高稳定交易系统

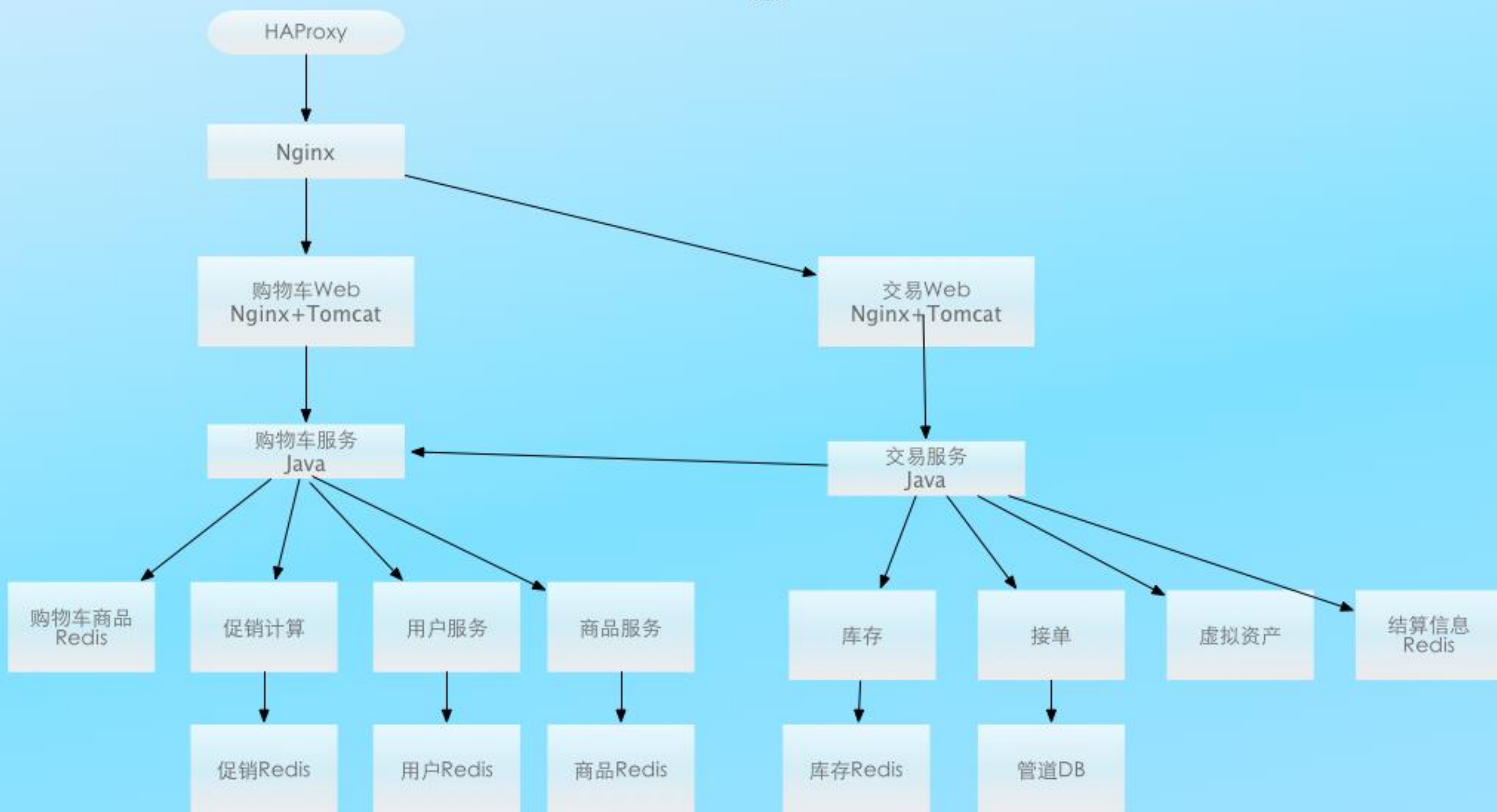
王晓钟 xiaohuook@gmail.com

十大原则

- 原则一：按照业务切分系统
- 原则二：灾备切换
- 原则三：水平扩展-分流
- 原则四：超预期流量-限流
- 原则五：保护系统-降级
- 原则六：提高性能-读写分离
- 原则七：唯快不破-异步化
- 原则八：动态，静态流量分离
- 原则九：线上压力测试-发现瓶颈
- 原则十：预案

交易系统概述

B28



业务拆分原则

一 按照业务逻辑切分系统

网站系统： 首页， 列表， 频道， 单品， 搜索

交易系统： 价格， 购物车， 结算， 支付， 订单中心

基础服务： 促销， 库存， 商品， 用户

订单系统： 管道， OFW， 订单中间件， 生产系统

灾备切换原则

灾备切换原则

一 多机房部署

机房内部出现大面积故障和机房入口流量故障

1 入口流量切换策略

1) DNS域名：较慢，机房入口处问题+机房内部，有地域性

2) LVS+HAProxy 直接路由：较快，机房内部大面积故障

二 机房内部切换

调用者客户端切换

1 Java应用切换策略：底层服务框架，VIP，内部DNS

2 缓存切换策略

3 数据库切换策略：客户端，DB自身

灾备切换原则

三 数据一致性保证：

1 程序双写：灵活，集群间相互切换快。

a 同步双写

数据一致性高，影响性能，不适合多个集群。

b 异步双写

数据一致性差，性能高，适合多个集群。

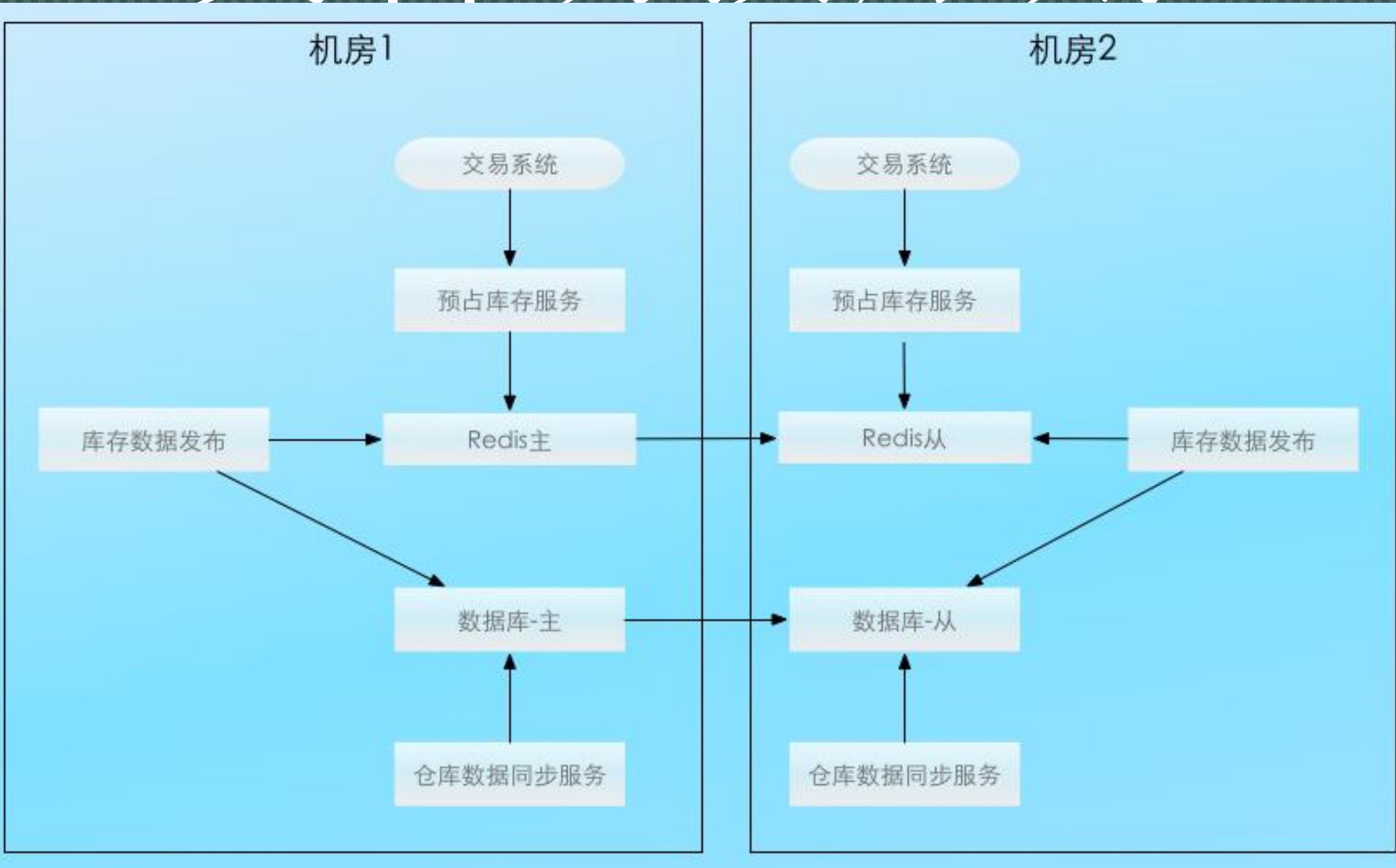
举例：购物车系统。

2 底层存储数据复制

优点：数据一致性高

缺点：主从切换不灵活，复制节点出问题后，恢复慢。

灾备切换原则



第一步：交易流量机房切换

第二步：数据库停写

第三步：数据库停复制

第四步：Redis停复制

第五步：逆向恢复

灾备切换原则

四 灾备集群

1 空闲灾备集群

热点缓存，管理成本

2 小流量灾备集群

增加事故风险

3 全部切换-部分切换

限流原则

限流原则：影响到用户体验，系统健康，谨慎使用

一 区分正常流量和超预期流量

超预期流量：恶意，秒杀

二 限流策略

1 Nginx层限流：

a 自主研发的模块

b 几个规则：账户，IP，系统调用逻辑

2 Web应用限流

3 业务应用系统限流：

1) 读少限，写多限

2) 限流标准来自压力测试 交易-库存

4 DB限流

分流原则

分流原则

一 水平扩展 无状态应用简单，有状态难

二 系统峰值引流：

秒杀系统：

1数据完全独立的部署 热点数据少

2应用独立部署

3定制化的策略

三 非重要业务分流到单独集群

购物车的库存状态，地址

结算页的运费险，Promise标志

读写分离原则

读写分离原则

一 整体系统的读写分离

二 读性能提升

增加缓存

1 内存

2 远程

3 热点 100万单

三 写性能提升

1 异步化

管道服务异步写DB

2 数据分片

redis, DB

降级原则

降级原则

一 系统级降级

- 1 结算页-成功页-订单中心-订单详情
- 2 下游系统降级 一次拆分暂停

二 业务功能模块降级：

1 业务逻辑

实时价格更新不及时

2 页面降级

我的京东首页 动态降级到静态

3 非关键业务屏蔽

购物车库存状态

4 远程服务降级到本地缓存 运费

动静分离原则

动静分离原则

一 静态流量

1 CDN

2 单独部署

二 动态流量

异步化原则

异步化原则

一 页面异步化
购物车，结算页

二 写逻辑异步化
促销，库存

线上压测原则

线上压测

一 读逻辑

二 写逻辑

三 覆盖范围

四 性能指标

1 硬件：cpu，内存，硬盘，网络

2 软件：吞吐量，并发数，服务响应时长（客户端/服务端）

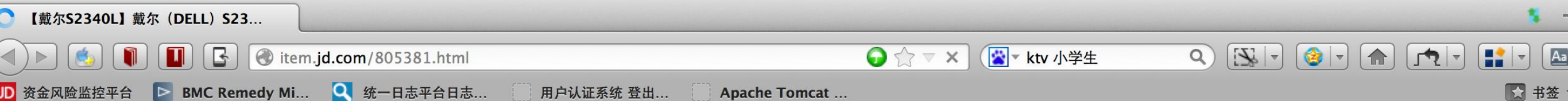
五工具

1 tcpcopy

2 压测工具

3 恶意流量

系统举例-实时价格



全部商品分类

首页

服装城

食品

团购

夺宝岛

闪购

金融

电脑、办公 > 电脑配件 > 显示器 > 戴尔 (DELL) > 戴尔S2340L



戴尔 (DELL) S2340L 23英寸宽屏IPS镜面无边框液晶显示器 (带HDMI)

【下单立减100元】HDMI高清接口镜面无边,京东独家畅享高清画感! 实物无往不胜~! S时尚家居系列1000-100,500-50店庆狂欢周

京东价: **¥1199.00** (降价通知)

商品编号: 805381

促销信息: **满减** 满1000.0减100.0, 满500.0减50.0 [详情 >>](#)

商品评分: **★★★★★** (已有8583人评价)

配送至: 北京朝阳区四环到五环之间 **有货**, 23:00前完成下单, 预计明日 (06月25日) 送达

服务: 由 京东 发货并提供售后服务。 支持: **夜间配** **货到付款** **自提** **免运费**

选择版本: S系列21.5寸IPS镜面屏 S系列23寸IPS镜面屏 S系列24寸MVA镜面屏

S系列27寸IPS镜面屏 **S2340M升级版23寸IPS镜面屏HDMI**

购买数量: **1**

京东服务: **延长保修1年** ¥239.00 **上门装机** ¥128.00 [详情](#)

已选择“**S2340M升级版23寸IPS镜面屏HDMI**”

加入购物车

轻松购

白条/分期

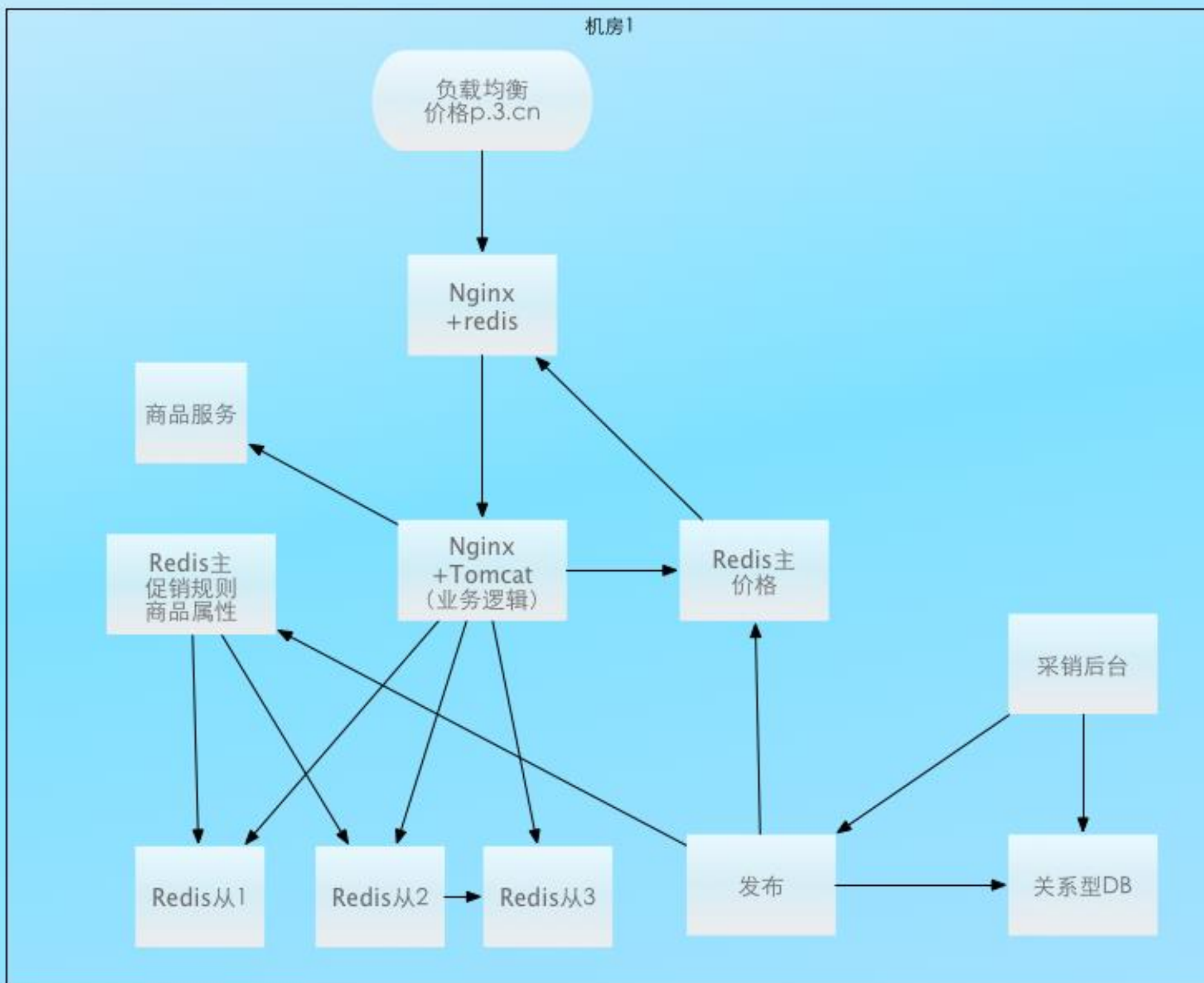
+ 加关注

客户端首次下单
送5元京券 [去手机购买](#)

调查问卷

返回顶部

系统举例-实时价格



特点1：纯动态，无法进行CDN缓存

特点2：读写调用量大

特点3：读写比高

系统举例-实时价格

一 业务逻辑概述

京东的实时价格由采销人员后台设置，京东价+促销信息（类型，开始，结束时间）。用户在前台商品详情页访问。京东价由采销设置的京东价+促销规则（直降等）计算而出。

二 写逻辑流程

- 1 采销后台更新价格，写关系型DB，通过异步任务通知发布程序。
- 2 发布程序更新促销信息主redis数据，更新价格主redis的过期时间戳。

三 读逻辑流程

- 1 Nginx直接读取本地redis中的价格数据，无过期则直接返回用户。
- 2 过期或没有命中则回源访问价格服务，取最新数据返回用户。

四 回源写逻辑

- 1 Java价格服务读取促销信息主redis，计算价格返回用户，同时异步写价格主redis集群。
- 2 价格主redis集群同步数据到前置Nginx节点的从redis

系统举例-实时价格

水平扩展-分流原则：

前置Nginx+Redis的架构是无状态的，Redis中几乎存放了全部商品价格信息。可以水平扩展，瓶颈在于主redis的数据复制能力。

超预期流量-限流原则：

前置Nginx增加了限流规则：规则+限流数据。

保护系统-降级原则：通过开关控制回源，降级为不回源，业务影响为价格不更新，但交易系统的价格为最新价格。

提高性能-读写分离原则：

读写逻辑分成几个闭环，完全互不干扰。

唯快不破-异步化原则：

采销后台异步调用发布，发布异步写价格主redis，回源异步写价格主redis。

系统举例-实时价格

基础技术应用 Nginx+Redis+SSDB

nginx做为webserver和代理，擅长Http请求的header或body处理，京东把nginx定位为业务router和业务handler服务。

例子：预售秒杀

nginx，lua和redis三者部署在同一台服务器，减少避免网络开销，提升性能和稳定性，节约了服务器数量；对于单台redis可以存储全量的业务，每一台服务器都是一个完整的集群，扩展非常容易。

例子：实时价格

nginx，lua和ssdb三者部署在同一台服务器，解决了redis数据持久化的缺点

例子：广告词服务

实时价格的压力测试

线上压力测试和系统瓶颈

1 回源是个瓶颈：

回源原因：1) 促销生效时会触发回源，为保证数据一致性必须重新计算价格。

2) 商品没有命中，redis的内存容量有限。

应对措施：1) 降级业务逻辑，不回源。

2) 使用其他存储代替redis。

2 线上前置负载曾经是瓶颈

管理-预案

一 预案准备： 共1000+预案， 执行100+个。

二 预案执行： 时间， 人

1 发现问题

软件， 硬件， 人肉监控

2 定位问题

监控数据， 日志

3 解决问题

二 预案演练

1 线上演练

2 压力测试

3 平日积累

结束语

招聘广告：

- 1 定制轮子-技术攻关 2种架构师
- 2 提供大流量，高并发的平台
- 3 良好的技术氛围

联系方式：

王晓钟 新浪微博 xiaohuook@gmail.com