

# 京东弹性计算实践

---

京东/云平台/系统技术部 何小锋

[www.jd.com](http://www.jd.com)



1

**京东弹性计算之路**

2

弹性计算架构

3

弹性调度策略及算法

4

弹性计算应用场景

5

数据驱动的精细化运营



01

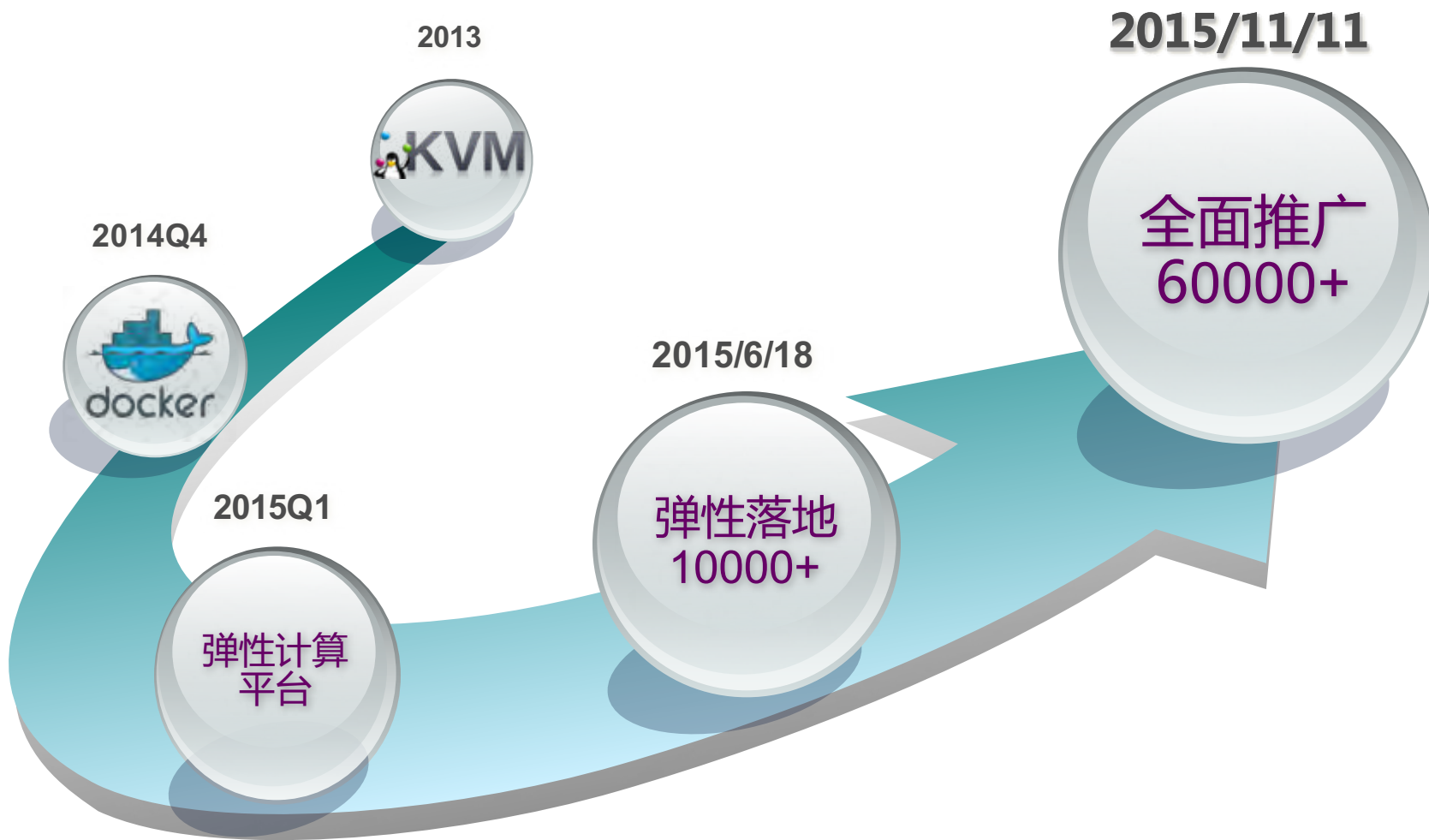
稳定性

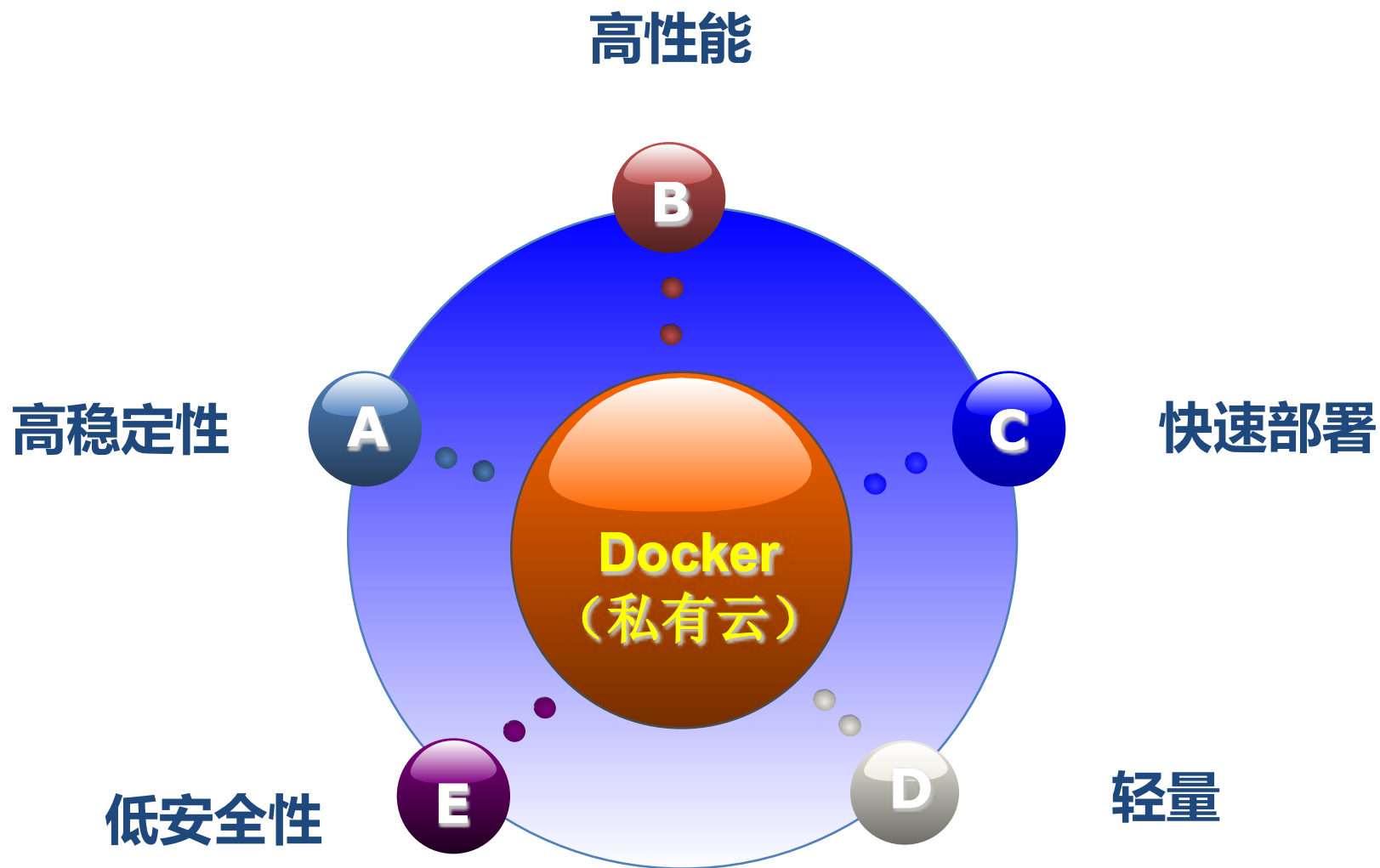
02

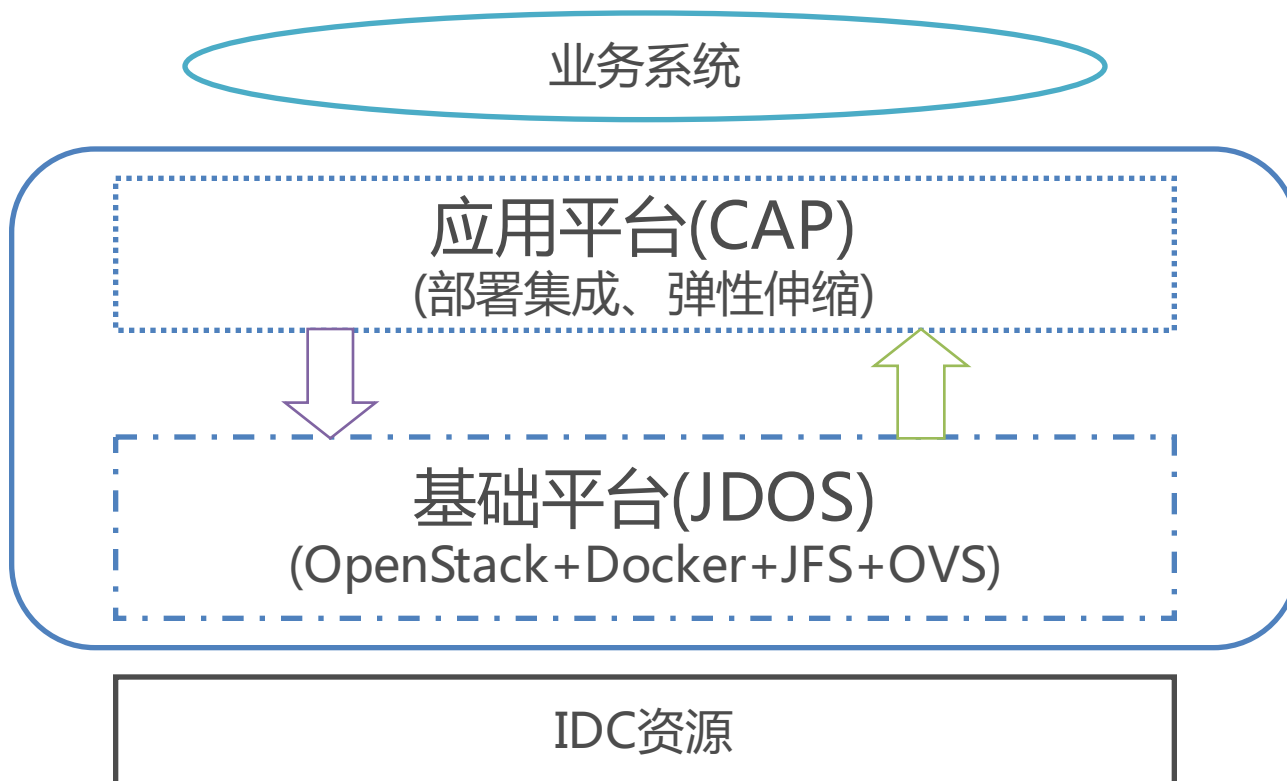
性能

03

用户习惯



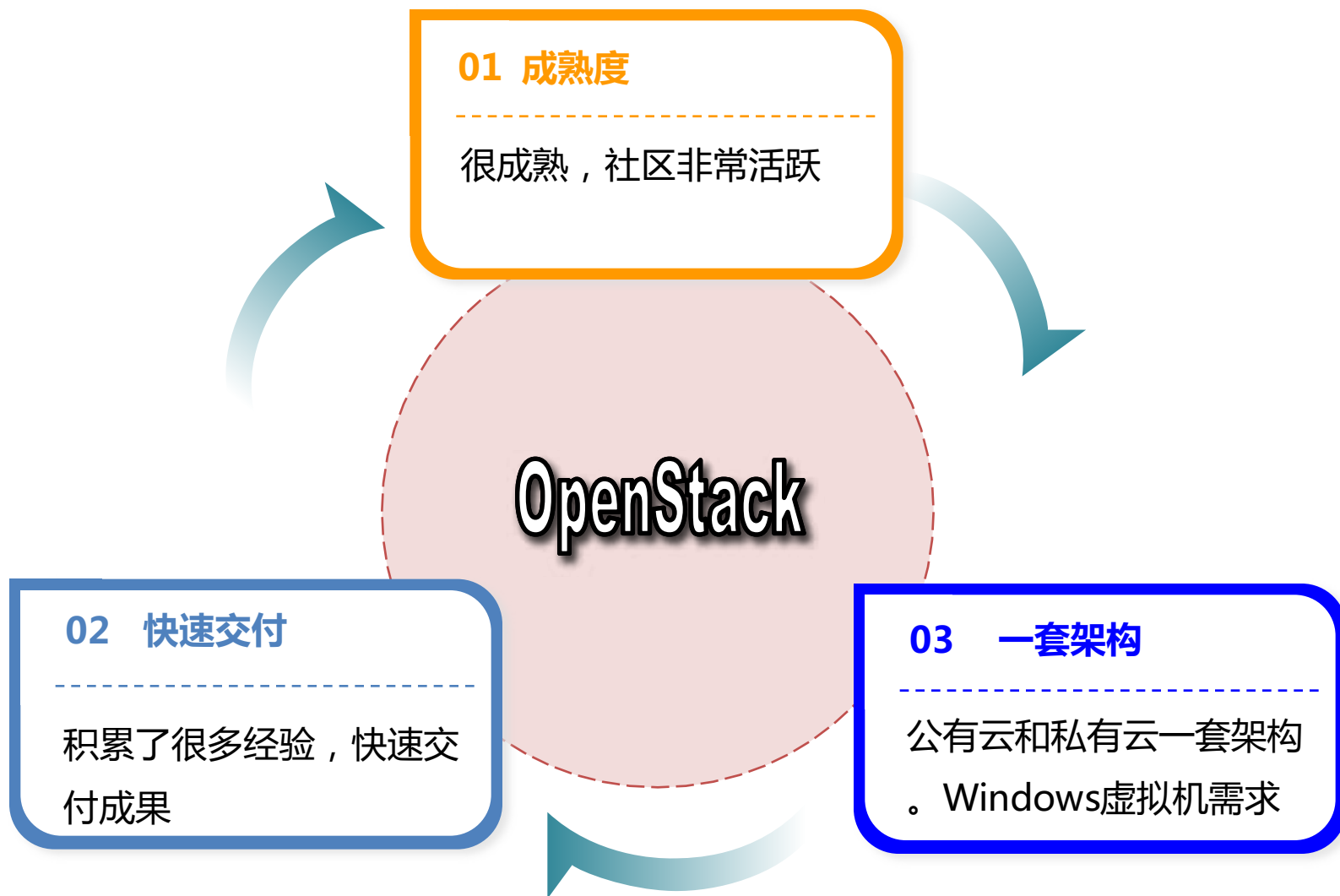




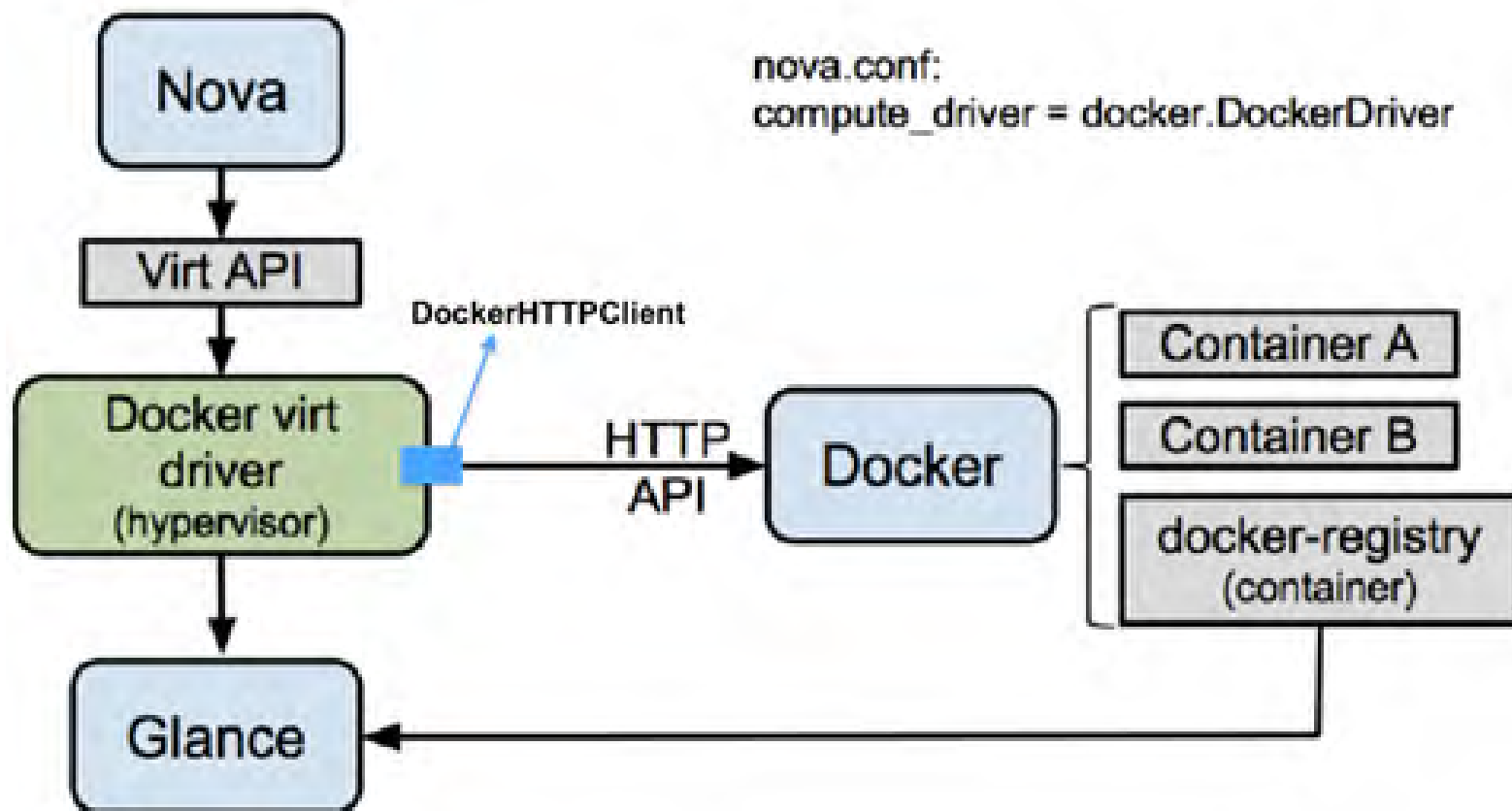
弹性计算平台 = JDOS ( JD Datacenter OS ) + CAP ( Cloud Application Platform ) 。

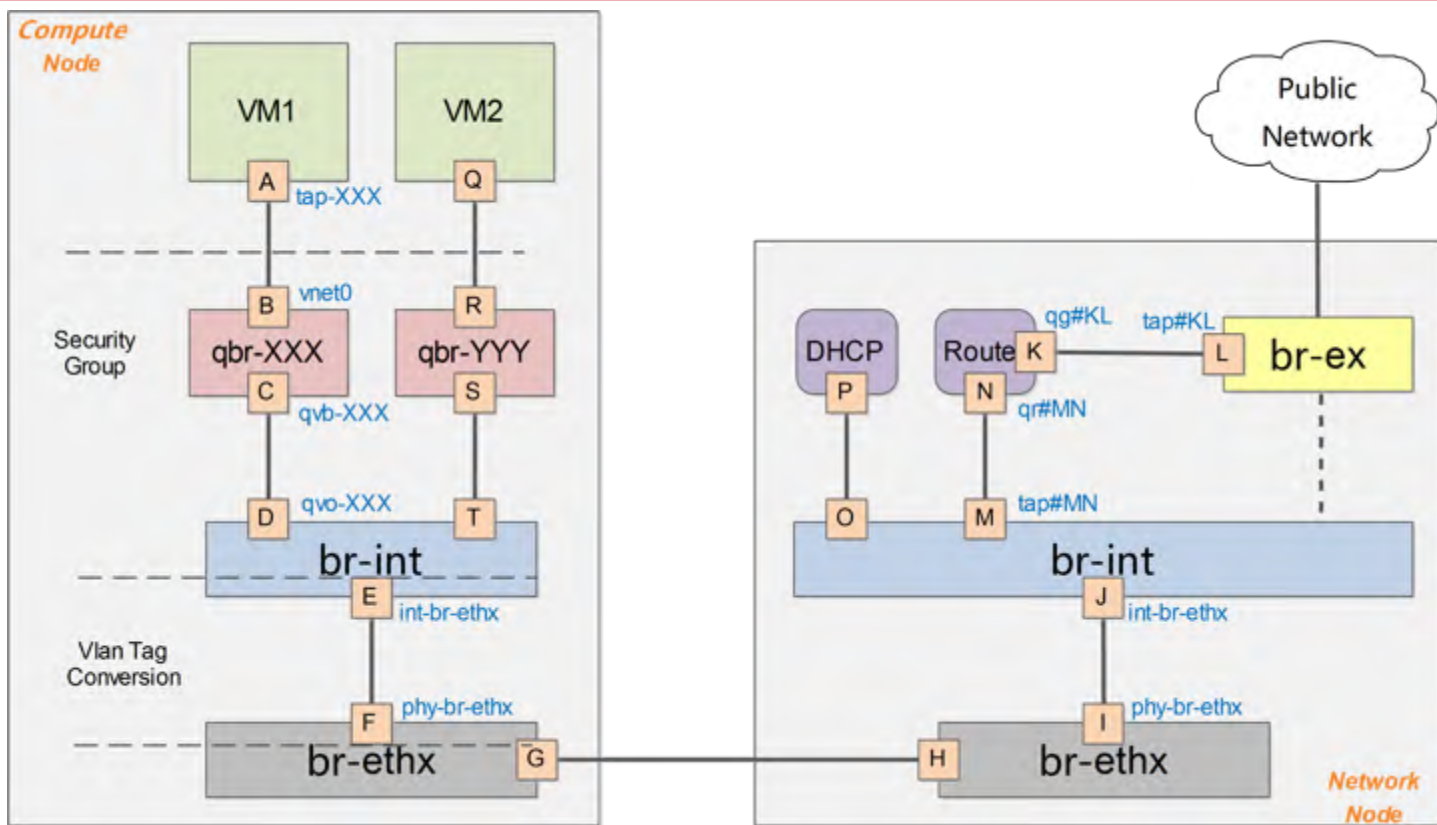
■ JDOS实现基础设施（网络，物理机，存储）的资源管理、容器的生命周期管理、监控指标采集；

■ CAP负责应用治理、部署、监控报警、资源利用率统计、手动和自动的弹性伸缩。

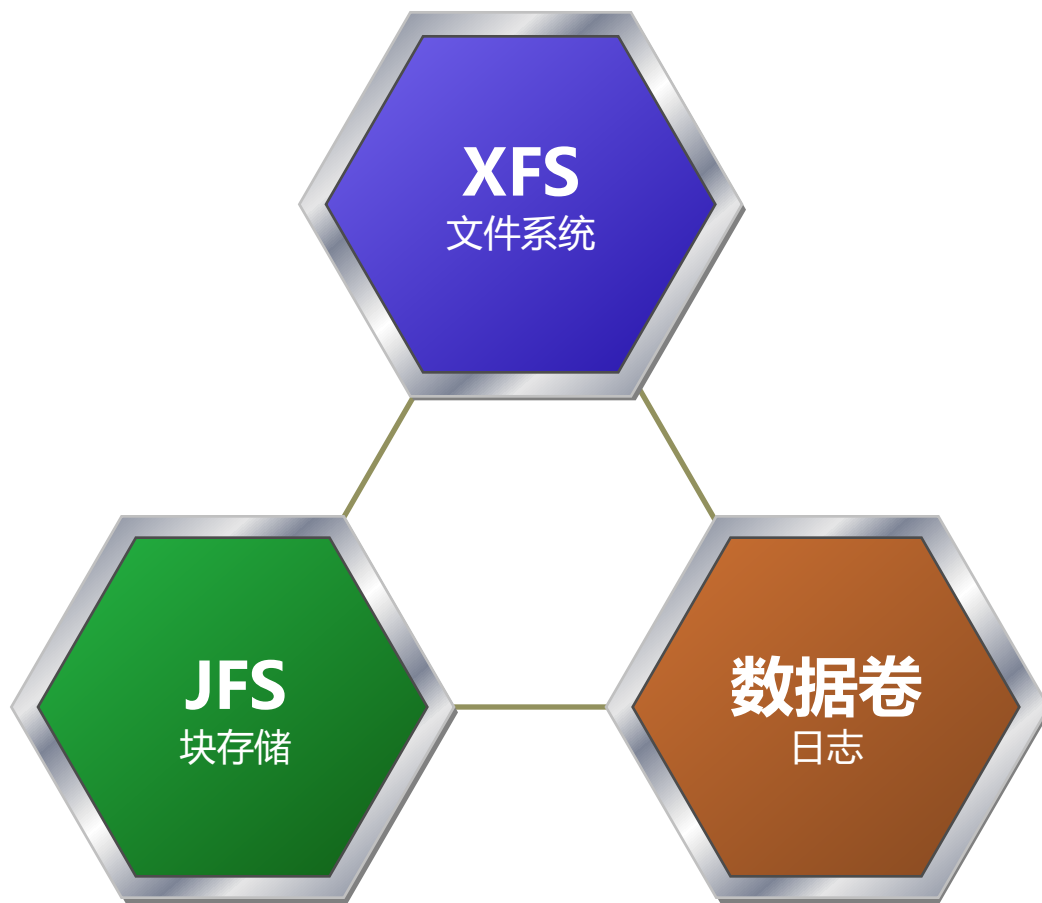


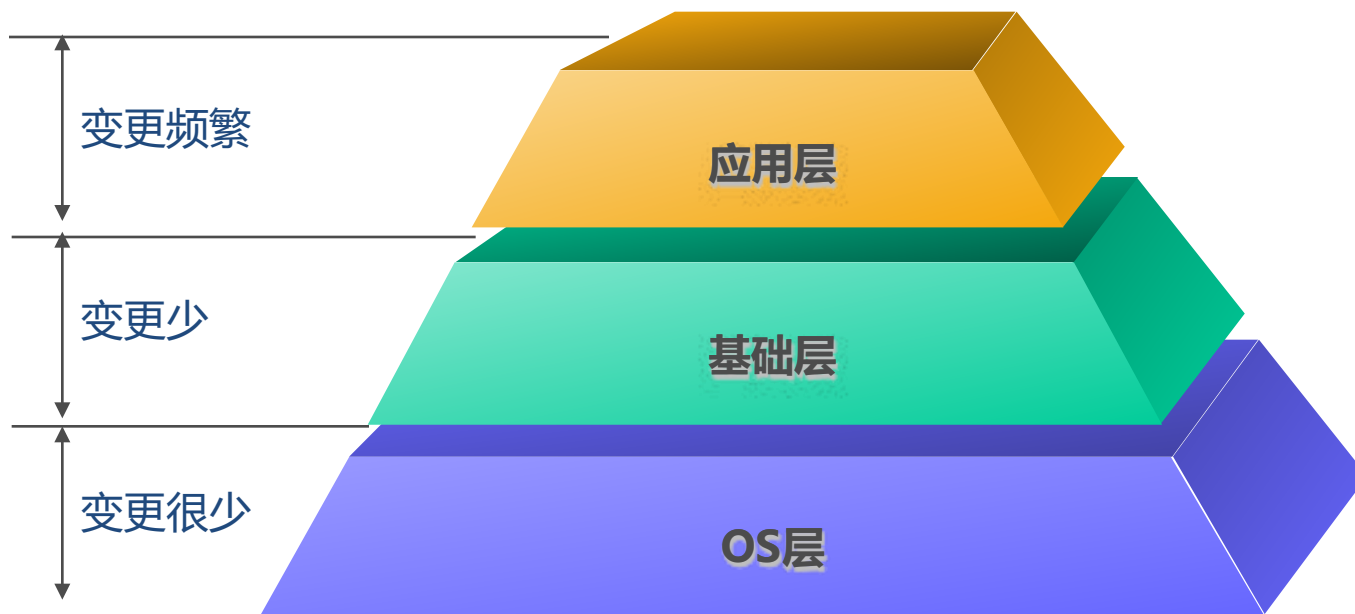


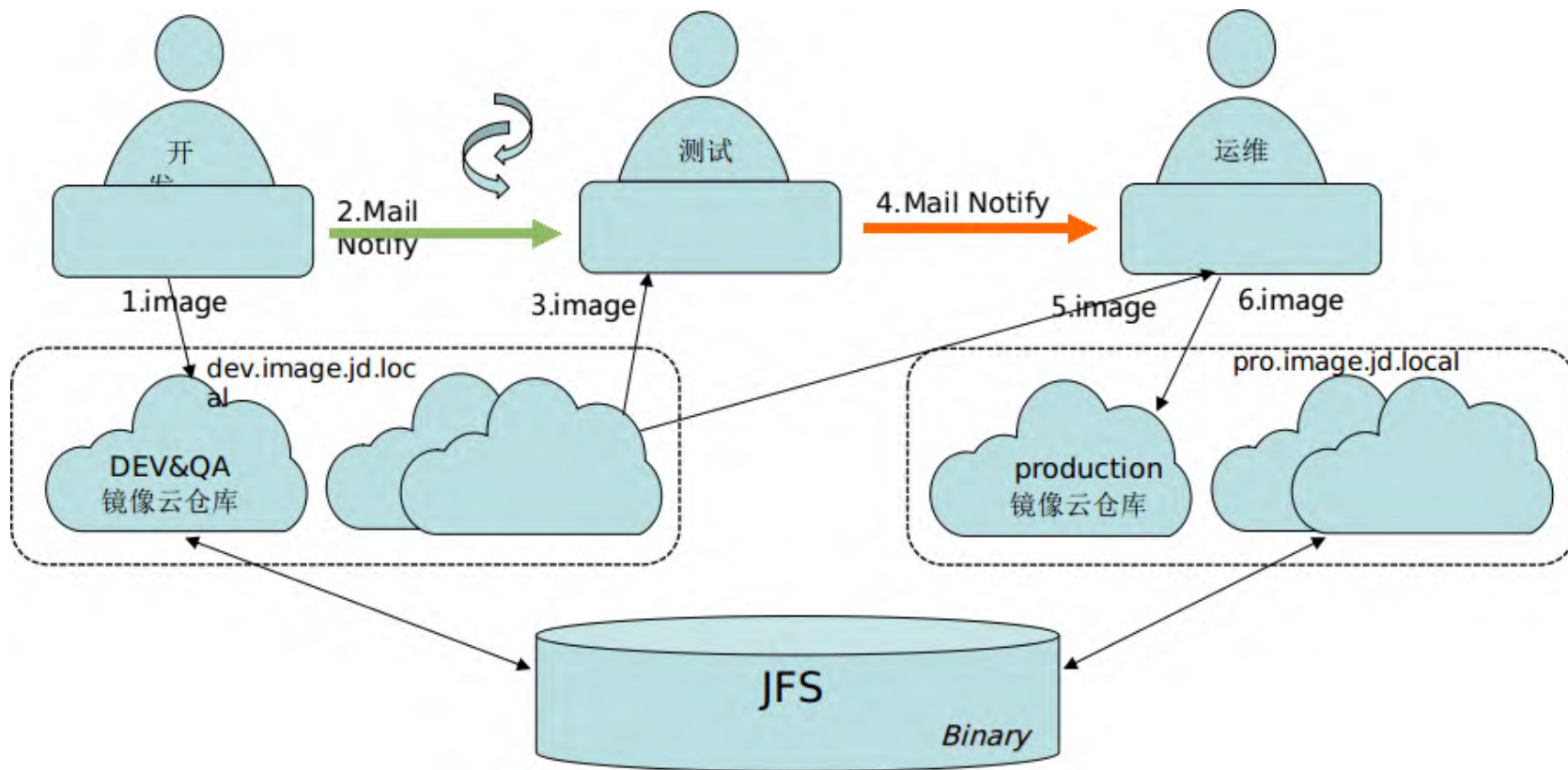


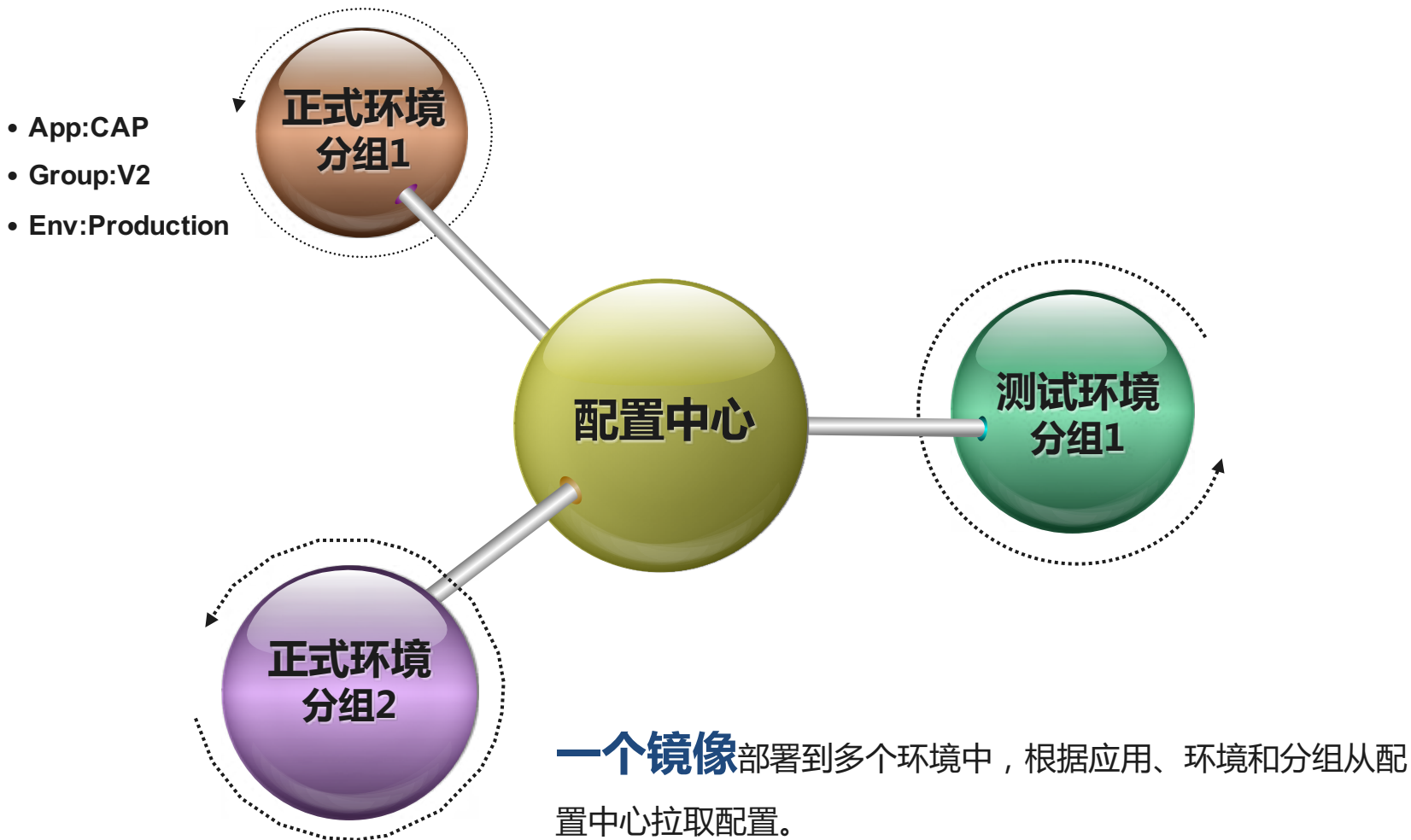


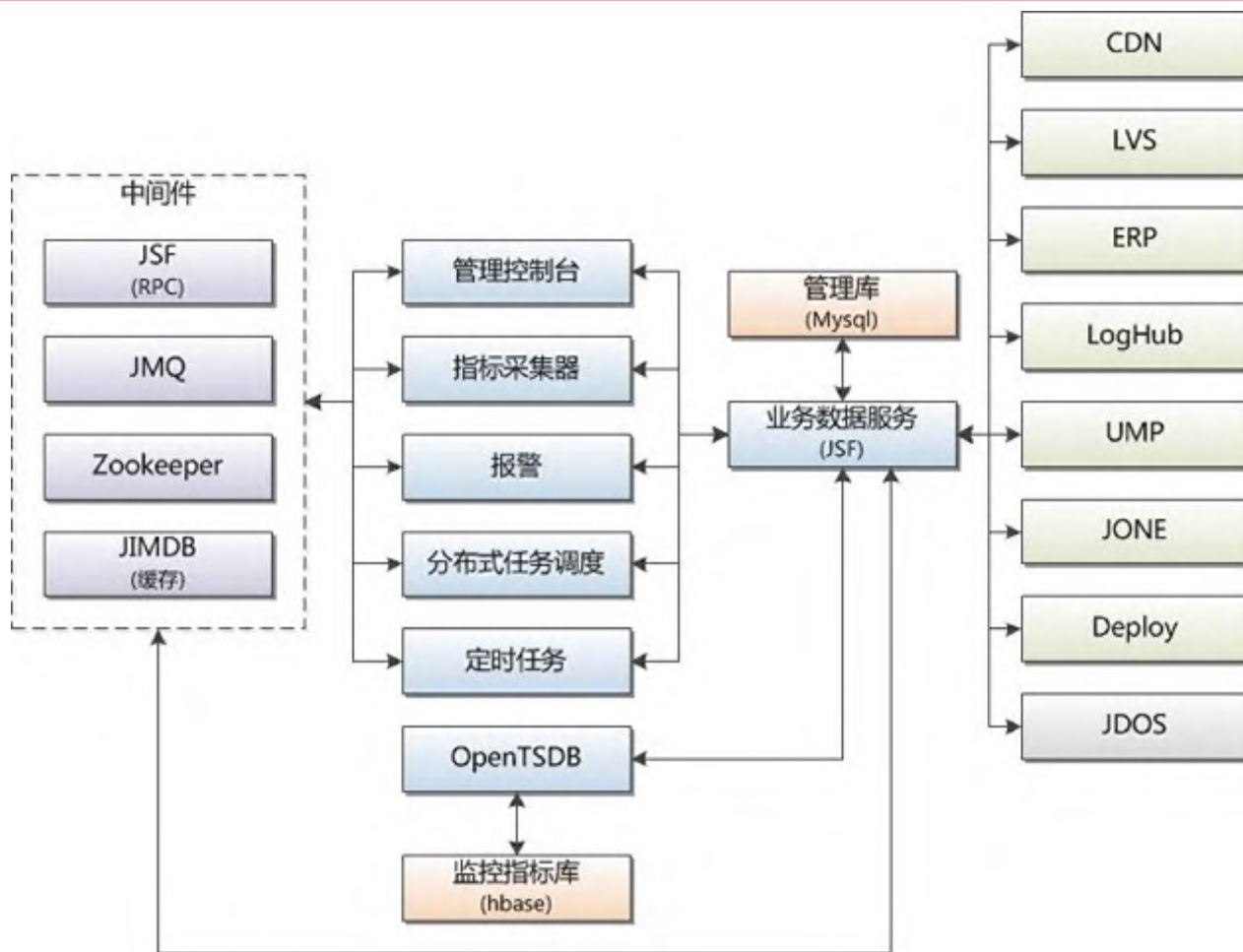
- 为了兼容现在的基础设施系统，满足用户习惯，每个容器都有独立的IP。
- 禁用了Docker网络，采用Neutron集成OVS；
- 优化OVS转发层，提升网络小包延迟，适用于微服务调用；



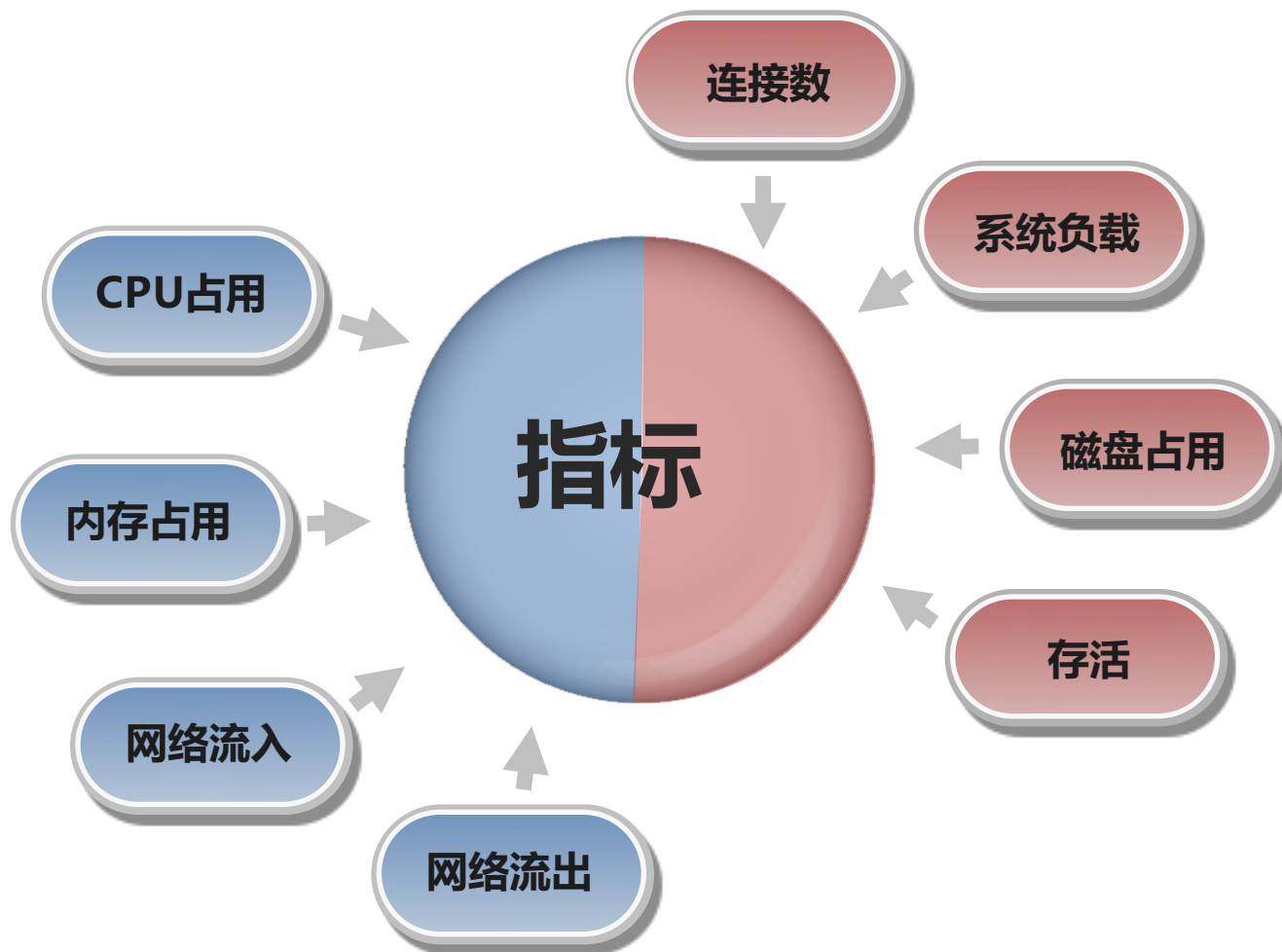




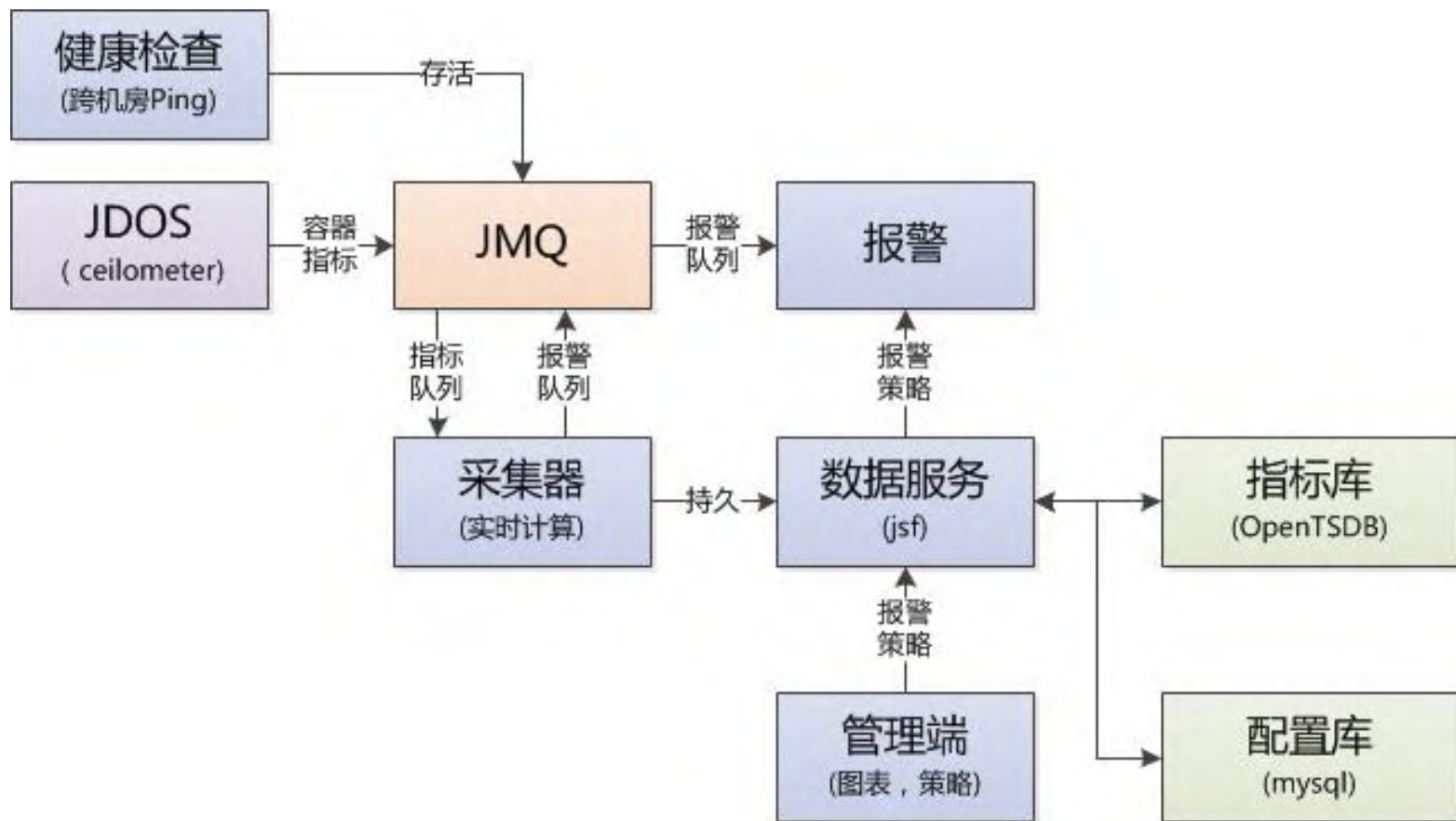




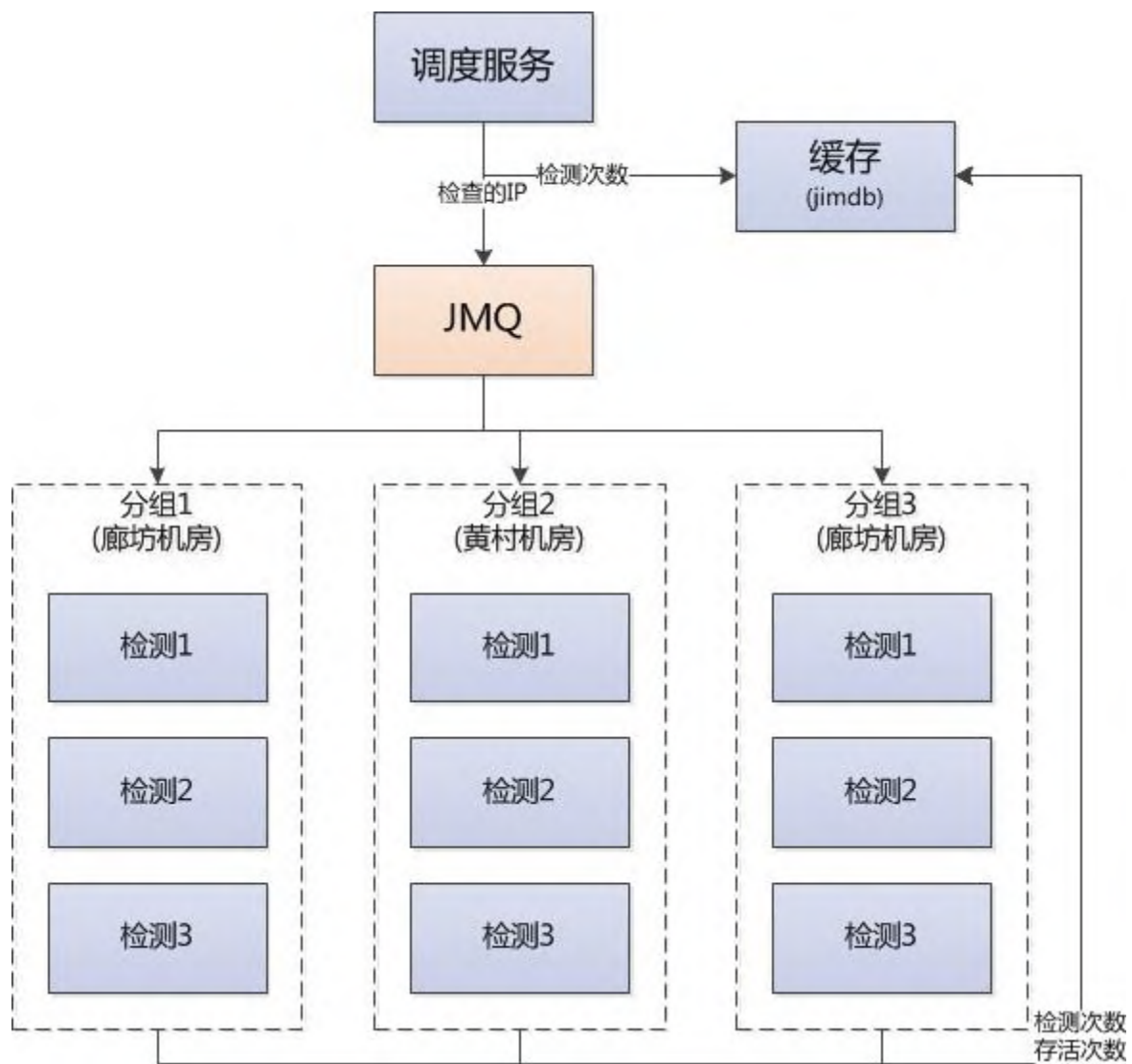
核心是一套 workflow，基于 Zookeeper 分布式调度引擎来实现。能动态注册发现节点；能控制单个节点并发任务数，失败重试次数，确保同一应用互斥任务串行执行。

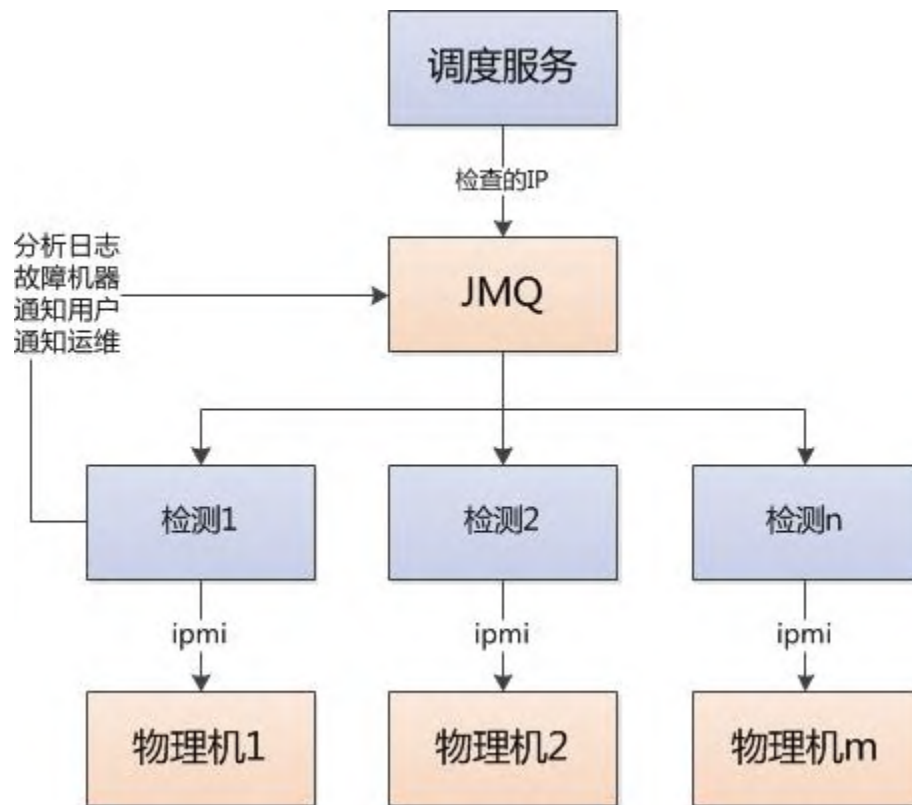






- 指标数据带有明显的时间特性，每日数据上亿，采用了成熟的OpenTSDB方案。
- 提供了从应用和实例多个维度查看负载情况，满足用户的需求。
- 可以对应用配置警策略，进行短信或邮件报警。

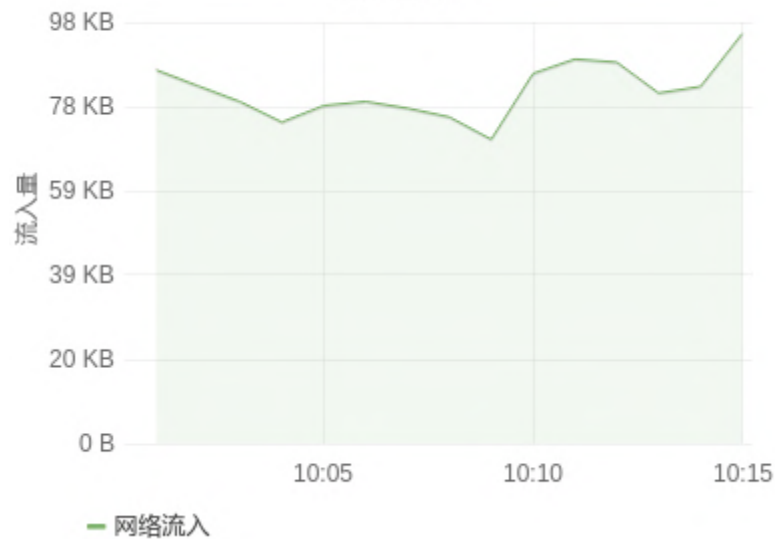




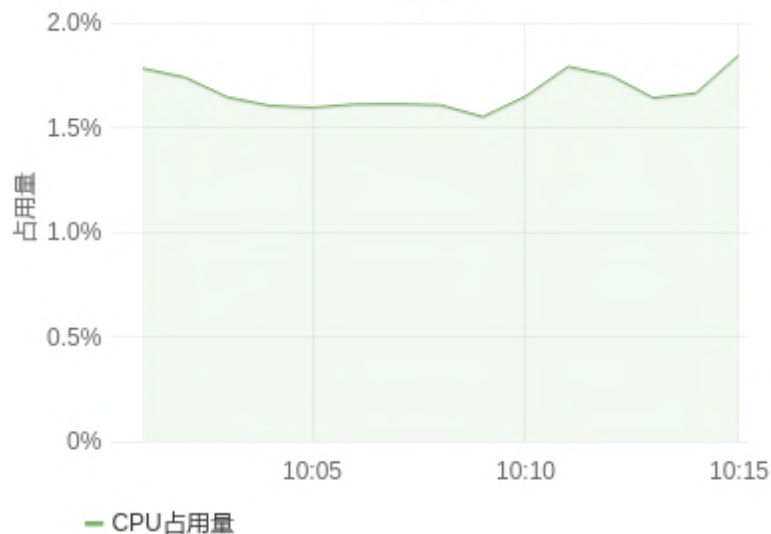
警告	严重
2	10

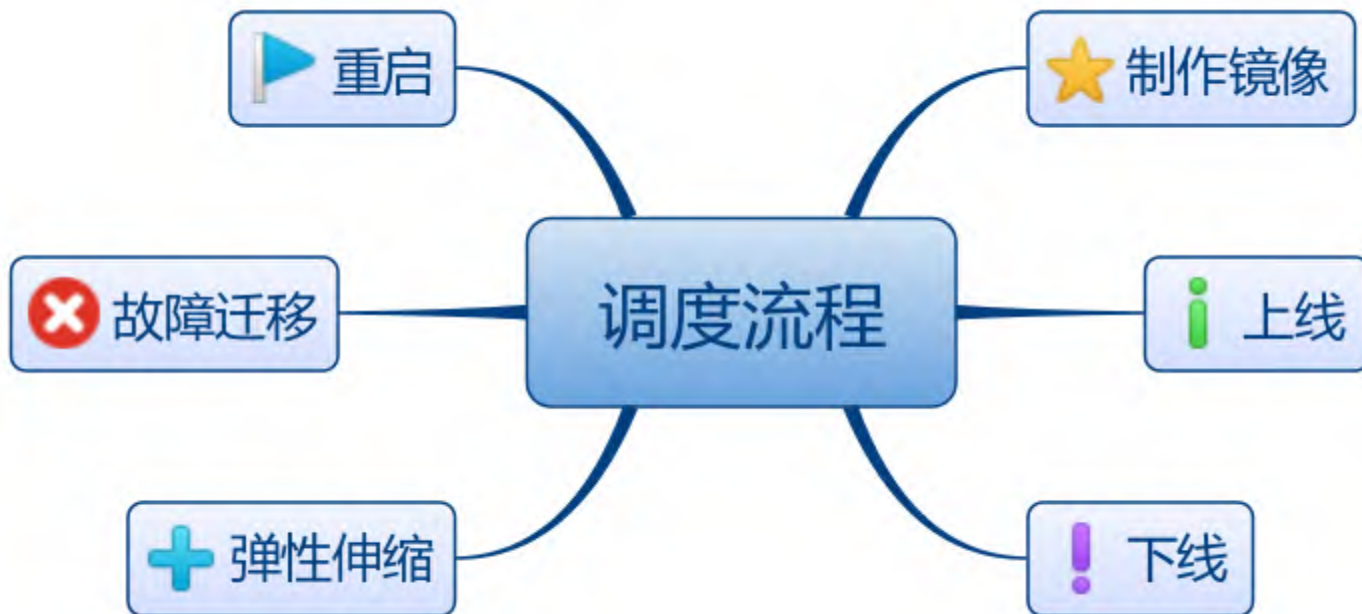
容器IP	主机IP	规格	机房	应用	部门	负责人	CPU	内存	入网	出网	TCP	磁盘	状态
172.19.118.42	172.19.164.39	超配...	永丰	xbopen	京东集团-...	njqiao...	<div style="width: 100%; height: 10px; background-color: green;"></div>	<div style="width: 100%; height: 10px; background-color: green;"></div>	<div style="width: 100%; height: 10px; background-color: green;"></div>	<div style="width: 100%; height: 10px; background-color: green;"></div>	<div style="width: 100%; height: 10px; background-color: green;"></div>	<div style="width: 100%; height: 10px; background-color: green;"></div>	存活
172.19.118.40	172.19.164.39	超配...	永丰	mqsto...	京东集团-...	bjchen...	<div style="width: 100%; height: 10px; background-color: green;"></div>	<div style="width: 100%; height: 10px; background-color: green;"></div>	<div style="width: 100%; height: 10px; background-color: green;"></div>	<div style="width: 100%; height: 10px; background-color: green;"></div>	<div style="width: 100%; height: 10px; background-color: green;"></div>	<div style="width: 100%; height: 10px; background-color: green;"></div>	存活
172.19.118.39	172.19.174.23	标配...	永丰	sk_gaj...	京东集团-...	bjtu	<div style="width: 100%; height: 10px; background-color: green;"></div>	<div style="width: 100%; height: 10px; background-color: green;"></div>	<div style="width: 100%; height: 10px; background-color: green;"></div>	<div style="width: 100%; height: 10px; background-color: green;"></div>	<div style="width: 100%; height: 10px; background-color: green;"></div>	<div style="width: 100%; height: 10px; background-color: green;"></div>	存活

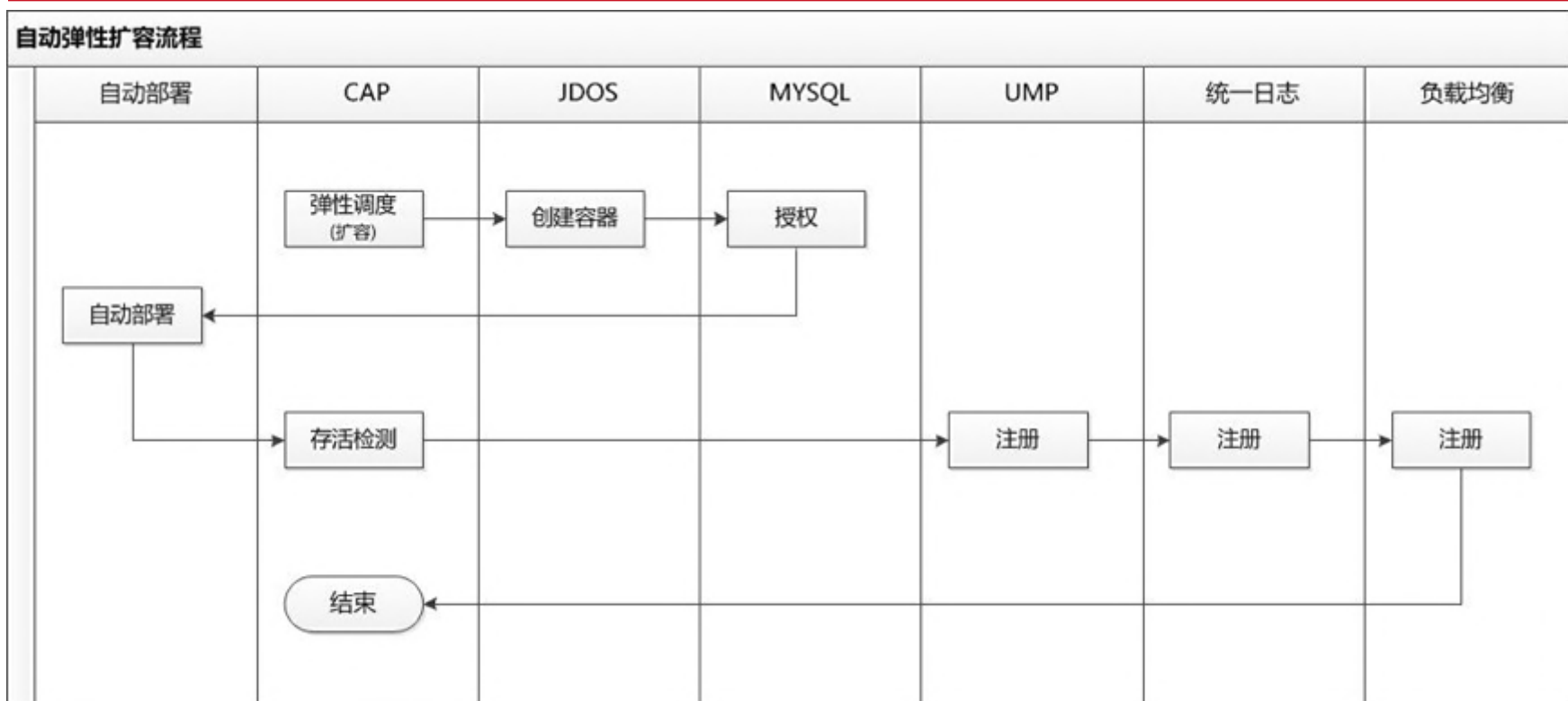
### 网络流入量



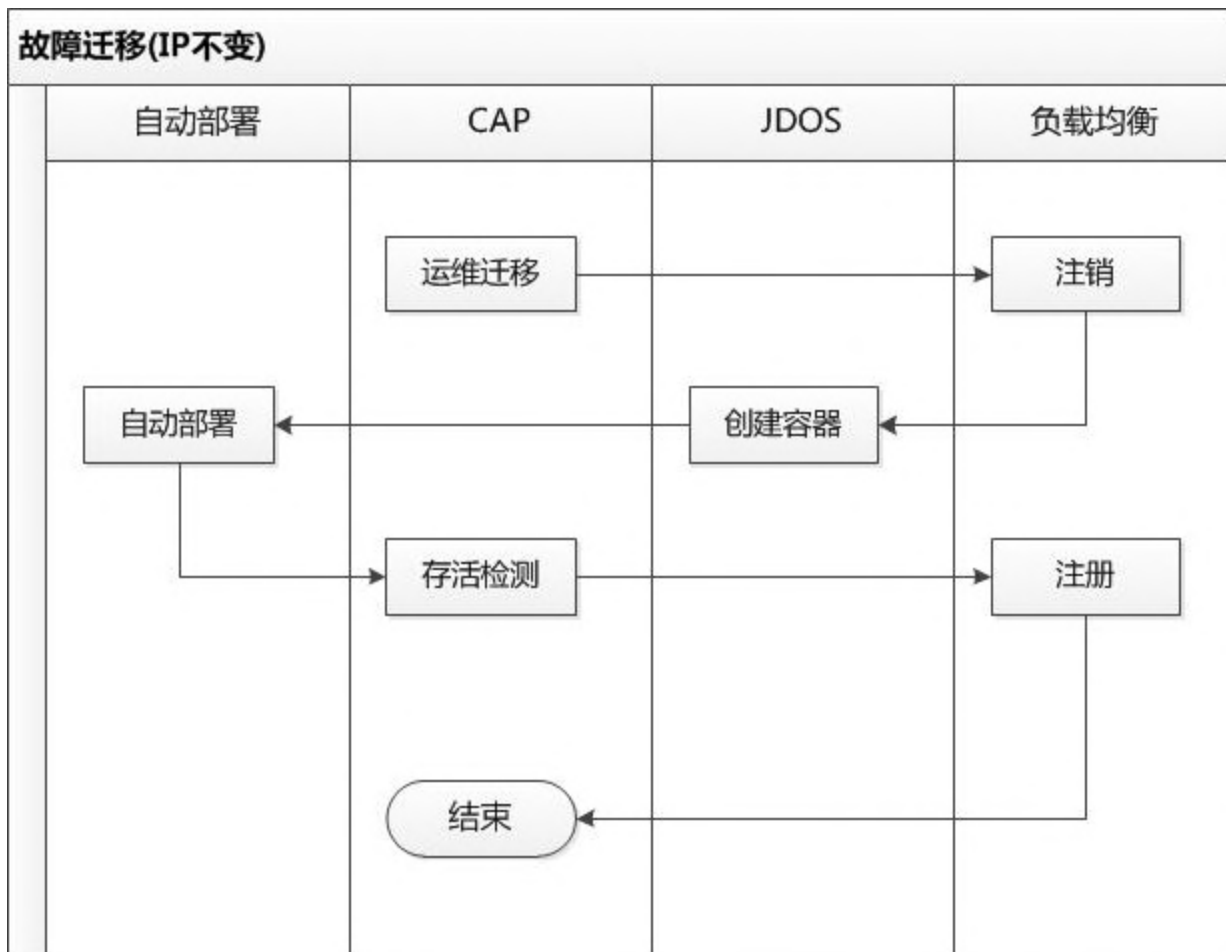
### CPU占用量







应用在启动之前可能需要数据库授权，启动之后需要挂载VIP，注册统一监控和统一日志。如何能自动发现应用的注册信息，采用了模版方式。应用先申请一个容器，手工注册这些信息，后续的扩容会以该容器为模版来进行自动注册



当遇到容器或物理机故障，需要进行快速的迁移，迁移后的容器需要保持原有的IP，避免还要重新申请授权。



- 调度单元是应用分组在一个机房的实例。
- 根据应用分组在指定机房的整体负载情况，预测下一时刻负载来进行弹性。





NGINX

JSF

Worker



- 双11全面使用弹性云来备战，线上应用在新机房都部署在容器上；
- 核心应用如：网站，交易，订单履约，配送，售后，无线，拍拍，金融，O2O等等平稳运行在容器上

应用名称	代码	部门	域名	负责人	IP	主机	分组	数据中心	ZONE	规格	使用率	CPU	内存	入网	出网	TCP	磁盘
弹性计...	ca...	京...	ca...	bjhe...	17...	17...	正式...	永丰	nova	[4核/8G...	30.66%	8.06%	30.66%	56844...	508128	25	13%
弹性计...	ca...	京...	ca...	bjhe...	17...	17...	正式...	永丰	nova	[4核/8G...	17.8%	2.45%	17.8%	7849.16	14896...	28	15%
弹性计...	ca...	京...	ca...	bjhe...	17...	17...	正式...	永丰	nova	[4核/8G...	17.32%	2.24%	17.32%	8197.73	19968...	28	32%
弹性计...	ca...	京...	ca...	bjhe...	17...	17...	正式...	永丰	nova	[4核/8G...	17.94%	2.18%	17.94%	11292...	12842...	28	15%
弹性计...	ca...	京...	ca...	bjhe...	17...	17...	正式...	永丰	nova	[4核/8G...	17.22%	2.1%	17.22%	7098.3	11652...	29	51%
弹性计...	ca...	京...	ca...	bjhe...	17...	17...	正式...	永丰	nova	[4核/8G...	40.34%	5.77%	40.34%	96237...	15652...	113	54%
弹性计...	ca...	京...	ca...	bjhe...	17...	17...	正式...	永丰	nova	[4核/8G...	48.47%	25.76%	48.47%	44550...	36142...	81	36%
弹性计...	ca...	京...	ca...	bjhe...	17...	17...	正式...	永丰	nova	[4核/8G...	73.63%	73.63%	50.34%	13050...	83189...	128	81%
jone_test	jon...	京...	jon...	bjlifa...	17...	17...	ces	永丰	nova	[4核/8G...	4.44%	1.27%	4.44%	6343.83	10390...	4	32%
弹性计...	ca...	京...	ca...	bjhe...	17...	17...	正式...	永丰	nova	[4核/8G...	17.07%	2.54%	17.07%	21600...	18464...	96	13%

- 以一小时为单位，计算容器的资源最大使用率；

应用代码	应用名称	部门	负责人	容器数量	资源使用率	CPU	内存	入网	出网	TCP	磁盘
foundati...	我的京东...	京东集团-...	付煜昆	1	45.4%	6.14%	45.4%	3557.76	3925.86	50	4%
order_to_...	订单kafk...	京东集团-...	颜伟彦	3	45.43%	45.43%	32.53%	273258.96	228278.43	43	97%
ss.jd.local	库存状态...	京东集团-...	张徐根	203	45.46%	15.17%	45.46%	10326547.83	3374014.81	6568	47%
app_heal...	业务系统...	京东集团-...	王远	2	45.57%	45.57%	11.73%	332119.5	171913.98	798	20%
npmon.ip...	ERP内配...	京东集团-...	邹庆林	3	45.78%	8.32%	45.78%	230708.65	366175.61	13	5%
p.m.beta	交易系统...	京东集团-...	周正	2	45.91%	20.69%	45.91%	36230.74	31230.91	155	13%
ivcvat.fm	资质、增票	京东集团-...	张子敬	48	46.09%	17.87%	46.09%	122393.51	11548.33	225	6%
pop-mar...	促销中心	京东集团-...	王代起	46	46.19%	9.55%	46.19%	20192713.9	4306318.71	3095	81%
peking	peking解...	京东集团-...	杨震	4	46.2%	46.2%	26.76%	301652.28	1152110.06	67	39%
pricecallb...	实时价格...	京东集团-...	彭晨	1	46.35%	3.13%	46.35%	6681.01	10558.26	228	8%

- 根据应用和容器的关系，统计应用资源使用率；

部门	容器数量	资源使用率	CPU	内存	入网	出网	TCP	磁盘
京东集团...	427	21.93%	8.91%	21.93%	3258337.13	3059869.06	2276	25%
京东集团...	75	18.87%	9%	18.87%	113675463....	63143689.08	436	77%
京东集团...	519	22.46%	15.51%	22.46%	1276289.66	981562.81	2288	88%
京东集团...	428	21.02%	12.39%	21.02%	63329789.73	6239280.21	1548	89%
京东集团...	14	17.29%	10.17%	17.29%	2553944.4	384858.66	107	65%
京东集团...	277	22.62%	2.57%	22.62%	5598788.73	3458102.3	2038	74%
京东集团...	160	20.94%	11.09%	20.94%	3196754.2	1221896.83	601	33%
京东集团...	380	31.78%	16.66%	31.78%	80524159.4	25227123.41	1766	91%
京东集团...	1793	19.31%	7.23%	19.31%	21229138.43	31931908.78	9760	100%
京东集团...	228	11.92%	4.68%	11.92%	822996.3	500376.78	1862	82%

- 根据负责人、部门、应用和容器的关系，统计部门资源使用率；

- 批量快照水平扩容；
- 批量水平扩容；
- 批量水平缩容；
- 批量垂直搜索；

应用	<input type="text"/>
数据中心	全部 ▼
分组	选择分组 ▼
现有规格	选择规格 ▼
更新规格	选择规格 ▼
是否强制	<input checked="" type="radio"/> 否 <input type="radio"/> 是

警告！强制执行缩容将导致内存和磁盘也会减少，可能导致应用不可用！

- 定期巡检应用容器部署情况，邮件报告；

## 超载 01

- 单个机房部署过多
- 单个交换机部署过多
- 单个物理机部署过多

## 未部署 02

- 申请的容器没有使用

## 规格不一致 03

- 容器规格不均匀，可能造成流量负载不均匀

- 无状态，同时对磁盘IO要求不高的应用，很适合部署到弹性云；
- 微服务应用由于能自动服务注册发现，辅助均衡，非常适合部署到弹性云
- 推荐万兆网络和网卡，避免网络共享出现资源竞争；
- 稳定的操作系统版本；
- 推荐高配置物理机，合理得CPU和内存比，便于充分利用资源；
- 采购高质量的交换机和物理机；

谢谢!

---

[www.jd.com](http://www.jd.com)

