

京东云

用云最佳实践白皮书

在京东云上构建高质量业务系统

京东云

2019-8-30

1	概述	1
1.1	京东云介绍	1
1.1.1	京东云概况.....	1
1.1.2	京东云主要优势.....	1
1.2	云计算平台特点	2
1.3	高质量业务系统	4
2	构建高可用的系统	5
2.1	简介	5
2.1.1	可用性.....	5
2.1.2	高可用主要技术.....	6
2.1.3	京东云高可用架构支撑.....	7
2.2	计算高可用方案	8
2.2.1	云主机热迁移机制.....	8
2.2.2	云平台内高可用方案.....	8
2.2.3	跨云高可用方案.....	10
2.3	数据级高可用方案	11
2.3.1	关系数据库服务 RDS	11
2.3.2	NoSQL 数据库 MongoDB.....	11
2.3.3	NewSQL 数据库 TiDB	12
2.4	网络高可用方案	13
2.4.1	高可用私有网络.....	13
2.4.2	公网链路高可用.....	14
2.4.3	网络专线高可用.....	15
2.5	高可用系统最佳实践	15
2.5.1	利用高可用组构建核心业务系统.....	15
2.5.2	采用云平台 PaaS 产品构建高可用系统.....	15
2.5.3	跨可用区构建高速的高可用网络.....	16
2.5.4	采用两地三可用区部署实现异地容灾.....	16
2.5.5	利用脆弱测试发现系统故障单点.....	16
2.5.6	基于专线的混合云架构实现稳定系统.....	16
3	构建可扩展的系统	16
3.1	系统扩展方法	16
3.1.1	水平扩展.....	17
3.1.2	垂直扩展.....	17
3.2	京东云主要产品可扩展性	17
3.2.1	可扩展架构.....	17
3.2.2	弹性计算扩展.....	18
3.2.3	网络扩展.....	18
3.2.4	云数据库与缓存扩展.....	19
3.2.5	互联网中间件扩展.....	19
3.3	高可扩展性系统最佳实践	20
3.3.1	采用无状态应用.....	20
3.3.2	采用微服务架构.....	20
3.3.3	采用弹性伸缩.....	20

3.3.4	创建弹性网络.....	20
4	构建高性能系统	21
4.1	简介	21
4.1.1	性能概述.....	21
4.1.2	性能提升方法.....	21
4.2	京东云主要产品性能	22
4.2.1	弹性计算.....	22
4.2.2	网络.....	22
4.2.3	数据库与缓存.....	23
4.3	全链路性能测试	23
4.3.1	测试场景.....	23
4.3.2	测试指标.....	23
4.3.3	测试方法.....	24
4.4	性能优化最佳实践	25
4.4.1	云主机性能优化.....	25
4.4.2	网络性能优化.....	25
4.4.3	存储系统性能优化.....	26
4.4.4	云数据库与缓存性能优化.....	26
4.4.5	互联网中间件性能优化.....	26
4.4.6	应用系统性能优化.....	26
4.4.7	应用架构性能优化.....	26
4.4.8	操作系统性能优化.....	27
5	构建高安全性的系统	27
5.1	简介	27
5.1.1	安全威胁与防护.....	27
5.1.2	京东云安全保障.....	28
5.2	网络安全防护	29
5.2.1	网络安全威胁.....	29
5.2.2	网络安全防护方案.....	29
5.3	管理和运行环境安全防护	30
5.3.1	管理和运行环境安全威胁.....	30
5.3.2	管理和运行环境安全防护方案.....	30
5.4	云主机安全防护	31
5.4.1	云主机安全威胁.....	31
5.4.2	云主机安全防护方案.....	31
5.5	应用系统安全防护	31
5.5.1	应用系统安全威胁.....	31
5.5.2	应用系统安全防护方案.....	32
5.6	数据安全防护	32
5.6.1	数据安全威胁.....	32
5.6.2	数据安全防护方案.....	33
5.7	安全服务保障	33
5.7.1	基线检测服务.....	33
5.7.2	漏洞扫描服务.....	33
5.7.3	渗透测试服务.....	33

5.7.4	应急响应服务.....	34
6	构建易运维的系统	34
6.1	简介	34
6.1.1	运维.....	34
6.1.2	运维指标.....	34
6.2	系统管理	35
6.2.1	资源管理.....	35
6.2.2	账户管理.....	36
6.3	系统监控	36
6.3.1	云监控.....	36
6.3.2	基于 DevOps 的监控	37
6.3.3	基于云计算管理后台的监控.....	38
6.3.4	监控服务.....	38
6.4	DevOps.....	38
6.4.1	介绍.....	38
6.4.2	功能.....	38
6.4.3	优势.....	39
6.5	故障应急响应	39
7	构建高性价比系统	40
7.1	简介	40
7.2	时间成本削减	40
7.2.1	采购、建设时间成本.....	40
7.2.2	研发时间成本.....	41
7.2.3	维护时间成本.....	41
7.2.4	战略转型时间成本.....	41
7.3	人力成本削减	41
7.3.1	研发人力成本.....	41
7.3.2	维护人力成本.....	41
7.4	资源成本削减	42
7.4.1	基础设施.....	42
7.4.2	硬件资源.....	42
7.4.3	软件资源.....	42
7.5	绩效提升	42
7.5.1	业务扩展与转型.....	42
7.5.2	融入生态体系.....	42
8	六高服务	42
8.1	高可用优化服务	43
8.2	扩展性优化服务	43
8.3	性能优化服务	43
8.4	安全优化服务	43
8.5	运维优化服务	43
8.6	降本增效服务	43
9	总结	44
10	引用	44

1 概述

1.1 京东云介绍

1.1.1 京东云概况

京东云(JD Cloud)是京东集团旗下的全平台云计算综合服务提供商，拥有全球领先的云计算技术和丰富的云计算解决方案经验。为用户提供从 IaaS、PaaS 到 SaaS 的全栈式服务(Full Stack)，从 IDC 业务、云计算业务到综合业务的全频道服务(Full Spectrum)，以及包含公有云、私有云、混合云、专有云在内的全场景服务(Full Services)和跨行业的全生态云服务(Full Ecosystem)。同时，京东云依托京东集团在云计算、大数据、物联网和移动互联网应用等多方面的长期业务实践和技术积淀，形成了从基础平台搭建、业务咨询规划，到业务平台建设及运营等全产业链的云生态格局，为用户提供一站式全方位的云计算解决方案。

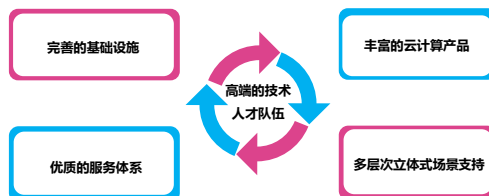


京东云拥有全球最大规模之一的 Docker 集群，中国最大的 GPU 集群之一；华北北京、华东上海、华东宿迁、华南广州 4 区域 10 可用区布局；99.95%的月度服务级别 SLA；专有云平台 JD Stack 完全兼容公有云，发挥云平台高可用、易扩展、低运维成本优势。京东云以一流的技术和可信赖的运维能力，向全社会提供安全、专业、稳定、便捷的云计算服务。

当前京东云具有可信云服务认证、支付卡行业数据安全标准认证 (PCI DSS)、ISO27001 信息安全管理体系国际认证、ISO9001 质量管理体系认证、信息系统等保三级安全认证、CSA STAR 云安全认证、C STAR 云计算安全评估认证等数十项资质认证。

1.1.2 京东云主要优势

京东云经过多年的发展，在不断的技术积累与创新下，形成了如下五大优势：



高端的技术人才队伍

京东云基于京东集团近 20 年的互联网技术积累，组建了一支成熟稳定的以高端技术人

才为核心主力的人才队伍。在以申元庆先生为首的一批云计算行业领军人物的带领下，京东云技术人才队伍不断拼搏创新，实现业界领先的技术实力和服务能力。

完善的基础设施

京东云在覆盖全国的 4 个地域建设 10 个可用区，每个可用区都由设施先进、功能完善的 IDC 组成。地域之间通过超高带宽的骨干网络连接，形成巨大的网络数据传输优势。基于稳定可靠的基础设施，京东云以一流的技术和运维能力向用户提供安全、专业、稳定、便捷的云计算服务。

丰富的云计算产品

经过多年的努力，基于京东云对整个行业的深入理解，当前已经为全社会贡献出 16 大类共 220 多款产品和服务。并且京东云还在不断提升产品的品质，丰富产品的种类，通过人工智能、大数据和物联网等行业领先的高技术产品，为用户业务的快速发展助力赋能。

多层次立体式场景支持

京东云通过功能丰富的标准产品提供 IaaS、PaaS、SaaS 等全栈式云计算服务，能够很好的支持公有云、私有云、混合云、专有云等多种场景，能够对外提供 IDC 业务、云计算业务、综合业务等全频段业务服务，还能够针对用户需求提供快速的定制化开发，充分满足用户对云计算的全方位需求。

优质的服务体系

京东云基于用户服务和通用技术服务等基础服务，以促进用户成功为理念，建立了金牌服务、优质架构服务、应急服务、迁移服务、系统优化服务等高技术价值服务，形成了基础稳固、技术先进、用户满意的层次化服务体系。通过配备高于业界水准的技术服务人力资源，使得用户服务更加专注、贴心，技术保障更加有力。

基于以上的五大优势，用户选择京东云之后，会立即获得多种保障，经过科学合理的系统设计和优化，将会对用户的业务发展有非常大的支撑和推动作用，使用户获得云计算技术带来的多种优势。

1.2 云计算平台特点

云计算技术的快速发展为信息技术带来很多新的机遇，用户在云计算技术发展的浪潮中获得了前所未有的益处。云计算平台是支持云计算技术的基石，具有鲜明的特点，不但包含很多优势，还包含一些风险。

采用云计算平台的主要优势有：

基础设施稳固。稳固的基础设施是云计算平台的基础。公有云云计算平台，往往需要搭建覆盖整个国家甚至是全球的稳固基础设施。设施先进、规模庞大的基础设施支撑了上述的丰富资源和丰富场景。

产品丰富易用、资源易管理。云计算平台能够为用户提供非常丰富的 IaaS、PaaS 和 SaaS 产品。云计算平台能够选择最实用最有价值的产品提供给用户，而且会屏蔽用户无需关心的技术细节和构建、配置过程，使得用户能够更加轻松地使用产品实现业务功能。提供非常丰富的资源，从 CPU 和内存形成的计算能力，到磁盘、SSD 硬盘提供的大容量存储能力，再到可以连接多个网络运营商的基础网络能力。还能提供研发平台、中间件、应用软件等丰富的软件资源。提供功能强大的云上 IT 资源的管理平台，用户能够通过管理平台轻松地管理成百上千不同种类的 IT 资源，极大地降低了大型系统的维护难度。

场景丰富。云计算平台能够提供公有云、私有云、混合云、专有云等多种场景，用户可以根据自身的业务需求选择适合自身的场景部署业务系统。

技术迭代快。信息技术飞速发展，用户需求不断变化，顶级技术厂商和众多开源社区、研究机构共同促进着技术的快速迭代，云计算平台也通过不断的更新自己的技术和产品，将最先进的有价值的新技术快速融入到平台能力中，为用户跟上快速的技术发展创造可能。

数据处理能力强。庞大的基础设施和丰富的 IT 资源使云计算平台具备了前所未有的处理能力。只有凭借云计算平台才能破解上亿用户集中访问网络站点的超级难题。

服务便捷。云计算平台不仅提供产品输出，还能为用户提供非常丰富的服务资源，帮助用户更容易和科学地构造和运营云上的业务系统。

使用成本低。在相同的系统效能下，云计算平台能够帮助用户最大限度地减少时间成本、人力成本和费用成本，从而有效降低用户研发和使用系统的整体成本。



尽管云计算平台具有以上很多优点，由于它也是一种大型的分布式信息系统，会具有以下一些风险：

全局范围内机器故障经常发生。由于每台服务器都有一定的故障概率，无论是由几十台

服务器组成的私有云，还是由几十万台服务器组成的公有云，尽管每台服务器的故障概率很小，当服务器规模增大时机器故障就会更容易出现。由于云计算平台是大型的分布式系统，机器故障在整个平台范围内会经常出现。

性能弹性大。云计算平台提供非常丰富的 IT 资源和产品，规格和种类都很多，当建立较为复杂的业务系统时，资源规格若没有很好的匹配，则会很大程度影响整体系统的性能。系统的性能弹性很大，需要通过有效手段进行优化。

安全威胁多。公有云云计算平台处于公网环境中，而公网环境中的安全形势往往比较严峻，遭受多种安全威胁。有效的安全保护措施对云上系统的安全稳定运行极为重要。

技能要求高。云计算平台提供的产品往往具有技术先进性和新颖性，若想更好的利用这些先进的技术则需要用户的技术人员具备较高的技术能力。若用户的技术人员没有相关的技术能力，则会对业务系统的高效运行带来一定的风险。

以上是云计算平台固有的主要优势和风险，本文旨在通过从多个角度提供解决方案，帮助京东云上的用户更好的发挥云计算平台的优势，并降低云计算平台的风险，使用户的业务系统有更科学合理的架构和更好的系统质量。

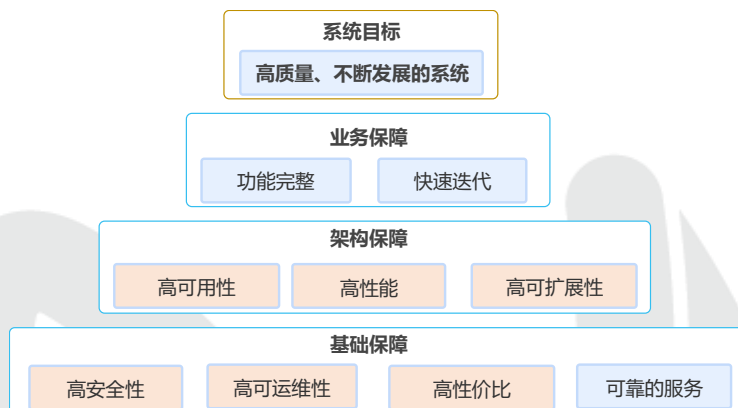
1.3 高质量业务系统

基于上述云计算平台的优势和风险，用户搭建高质量的云上业务系统需要遵循下面的主要原则：

- **基于分布式系统架构设计。**云计算平台是天然的分布式系统的友好平台，业务系统基于分布式系统架构进行设计和实现有利于发挥云计算平台的优势和降低风险。
- **高可用。**针对用户业务系统特点，对京东云上目标系统合理设计，保证用户业务系统部署在京东云上实现高可用，业务系统能够连续稳定运行和输出。
- **可扩展。**对用户业务系统不同时段性能和资源需求进行分析，对用户业务未来发展规划进行梳理，通过合理的设计，使用户业务系统具有满足需求的可扩展性。
- **性能达标。**对用户原有系统性能需求进行分析，设计合理的云上系统架构和选用适当的云资源规格，保障用户业务系统在京东云上运行性能达到目标。
- **安全可靠。**针对云上业务系统可能受到的安全威胁，利用云平台提供的安全能力，对系统进行充分的安全性设计，有效保护系统安全。
- **易运维。**保障用户在京东云上的业务系统和云资源能够被有效的运维管理和监控，确保及时发现系统故障，帮助用户对云上资源进行方便的管理。

- **成本低**。在满足用户业务系统的性能、安全性、高可用等前提下，帮助用户通过技术手段和管理手段降低系统迁移后的运行和维护成本，实现降本增效的目标。

为了提高用户业务系统的质量，促进用户业务的不断发展，京东云基于上述主要原则为用户提供高质量云上系统解决方案。如图所示。



通过基础保障，帮助用户实现高安全性、高可运维性，通过控制合理的成本开销实现系统的高性价比，并通过优质可靠的服务解决用户的疑难问题。通过架构保障，大幅提升用户业务系统的可用性、性能和可扩展性。京东云通过以上两层重点保障，帮助用户最终实现高质量的云上业务系统，促进业务系统的不断发展。

在下文中，简称具有“高可用性、高可扩展性、高性能、高安全性、高可运维性、高性价比”等特征的系统为“六高”系统。

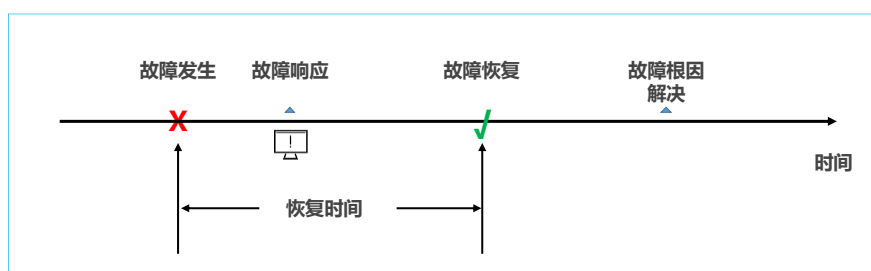
2 构建高可用的系统

2.1 简介

2.1.1 可用性

对于 IT 系统，可用性是指在指定的一段时间范围内系统可以被访问和可以被使用的能力。若系统在指定的时间范围内因故障或某些操作导致不能被访问和被使用，则系统的可用性就会下降。对于需要连续运行并提供服务的系统，可用性是非常重要的系统质量好坏的评判指标，尤其对互联网企业和关系民生、安全等重要 IT 系统具有高可用性极为重要。提升云上的系统的可用性是云厂商和用户 IT 部门的共同责任，因为云厂商提供云上资源，而用户需要通过良好的设计将这些资源进行科学合理的组合，形成质量较好的系统。

衡量系统可用性的指标有很多，其中最重要的是“恢复时间目标”（Recovery Time Objective, RTO），是指信息系统从灾难状态恢复到可运行状态所需的时间，用来衡量容灾系统的业务恢复能力。



京东云对用户提供的可用性 SLA 主要有（其中服务周期按月计算）：

- 云主机、原生容器、RDS（关系数据库服务）提供的服务可用性不低于 99.95%。
- 对象存储提供的服务可用性不低于 99.9%。
- 负载均衡、CDN 提供的单实例服务可用性不低于 99.9%。

2.1.2 高可用主要技术

2.1.2.1 容错机制

通过引入容错机制可提升系统可用性。主要的容错机制包括侦测、自动切换、自动恢复。

侦测机制是利用在计算节点外部部署的监听程序侦测该计算节点是否正常运行。通过侦测机制可以了解计算节点的活性，并可以根据计算节点的状态进行系统调整，防止由单个计算节点的故障影响整个系统的正常运行。

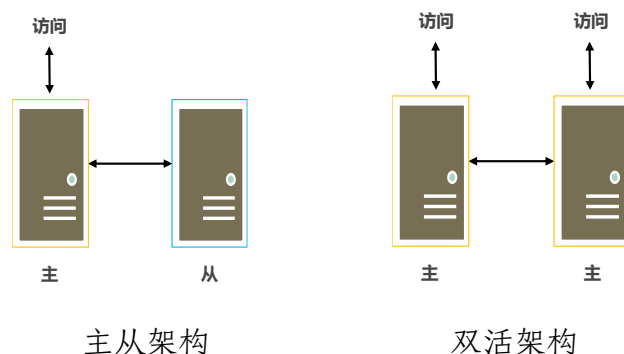
自动切换机制，是在侦测阶段发现计算节点故障后，利用其它的计算节点来接替发生故障计算节点处理的工作，使业务能正常运行。

自动恢复机制，是完成自动切换之后，通过离线或在线的方式让故障计算节点恢复到可正常工作的状态，再根据故障处理的策略选择将业务恢复到原计算节点上或者让原计算节点作为正在工作的计算节点的备用计算节点。

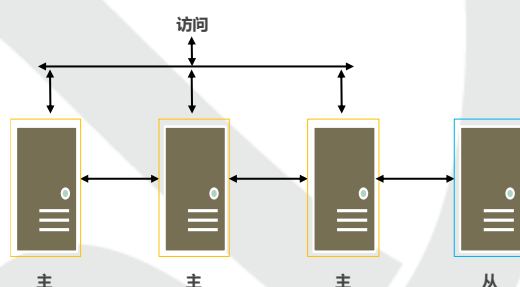
2.1.2.2 高可用架构

基于容错机制，实现高可用系统的云上架构主要有主从架构、双活架构、集群架构等。

主从架构是一种非对称的容错架构，分为主计算节点和从计算节点。主计算节点用来处理业务请求，从计算节点具有和主计算节点相同的能力但不处理业务请求，而一般情况下只侦测主计算节点的活性。当主计算节点宕机时，从计算节点接替主计算节点进行工作。主从架构分为主从热备和主从冷备两种。主从热备架构中从计算节点与主计算节点同时运行。主从冷备架构中从计算节点不与主计算节点同时运行，当主节点出现故障停机时，需要手动启动从计算节点。主从架构中，若从节点的个数大于等于 2，则称为副本集架构。通过增加从节点个数可实现更高的可用性，但管理复杂性也会相应增加。



双活架构是一种对称的容错架构，分为两个等同的计算节点。两个计算节点都处理业务请求，并互相侦测对方的活性，若有一方出现故障则另一方会将故障一方的任务接手处理。当故障计算节点恢复后，两个计算节点再度同时处理业务请求。



集群架构是以上两种架构的扩展，采用多个计算节点处理业务请求，当某一计算节点出现故障后，集群中选择一个或多个计算节点代替故障计算节点处理业务请求。

2.1.3 京东云高可用架构支撑

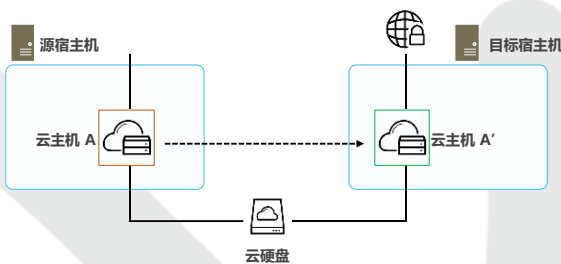


京东云基于长期的技术积累，构建了完整的能够支撑高可用架构的技术体系。基于京东云稳固的基础设施保障，利用有效的容错机制和可靠的架构基础，研发出能够支撑用户高可用业务应用系统的计算、数据、网络等技术先进的云计算产品。

2.2 计算高可用方案

2.2.1 云主机热迁移机制

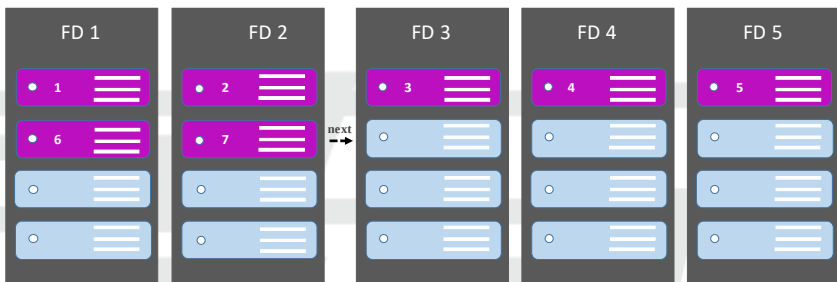
京东云的云主机采用热迁移机制实现高可用。热迁移，又叫动态迁移、实时迁移，即将整个虚拟机的运行状态快速完整同步到不同硬件平台上，并切换网络配置指向迁移后实例。迁移过程中虚拟机保持平滑运行，用户无感知。京东云借助热迁移方案实现单服务器 99.95% 可用性。



2.2.2 云平台内高可用方案

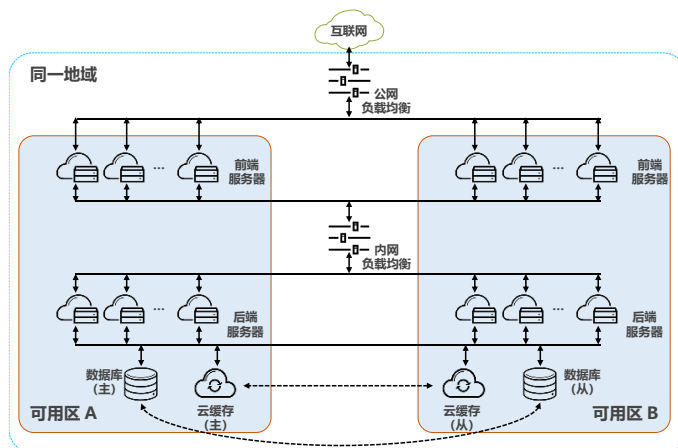
2.2.2.1 跨故障域实现高可用

故障域是数据中心内相互隔离的机架组。京东云上可以通过选择高可用组实现跨故障域的高可用。高可用组是京东云提供的云主机逻辑集合，高可用组内的云主机分散部署在相互隔离的物理资源上。单个高可用组在单个可用区内默认分配 5 个故障域。高可用组支持在同一可用区内跨不同故障域的机柜，也支持跨可用区。当出现硬件故障或定时维护时只会影响部分云主机，用户的业务仍可用。



2.2.2.2 跨可用区实现高可用

可用区是同一个地域下电力及网络互相独立的数据中心。不同可用区之间的可用性是相互完全独立，互不影响。因此用户业务系统的服务器可以通过跨可用区部署很大程度提升可用性指标。京东云同一地域内不同可用区之间相隔数十千米，为了让用户业务在多个可用区部署后能流畅运行，京东云提供可用区之间的高速网络连接，保障一个地域内的多个可用区之间的网络延时小于 1.5ms。跨可用区高可用方案为用户的关键应用提供高等级的容错和容灾能力，如图所示 VM 集群可用性可达 99.99%。

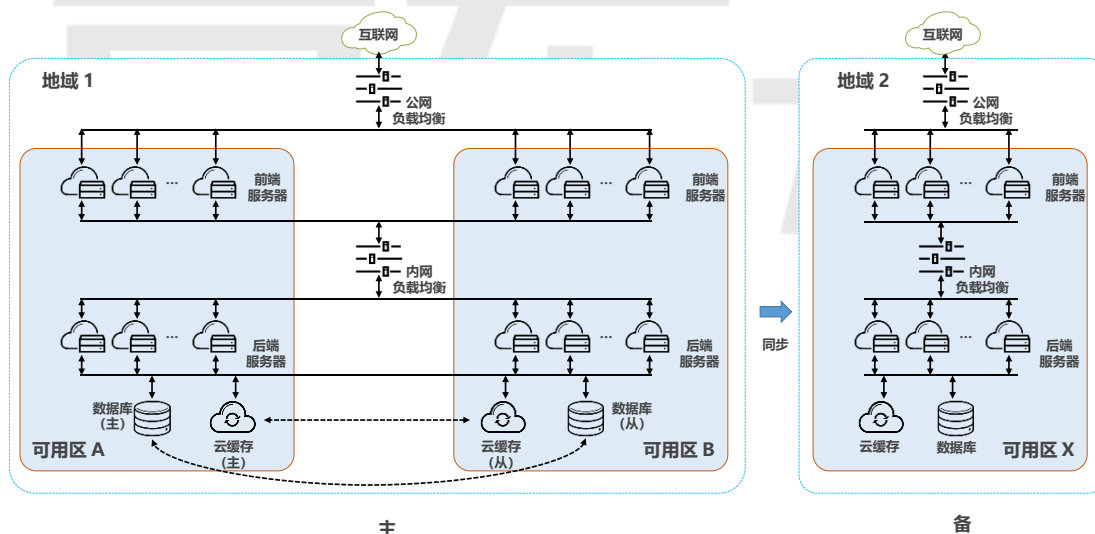


2.2.2.3 跨地域实现高可用

一个地域内的云资源部署在同一个城市或地区内的数据中心内。在不同的地域提供云资源可以有效降低云上业务到达最终用户的网络延时，并能更好的满足不同地域的数据及服务法规要求。通过跨地域实现业务系统高可用，最大的优势是能够防止如地震、军事打击等导致的某个城市或地区内的所有数据中心不可用，而跨故障域和跨可用区的高可用方案都达不到这个要求。利用公有云多地域的能力可以实现高可用的典型部署场景包括两地三可用区部署、异地多活等。

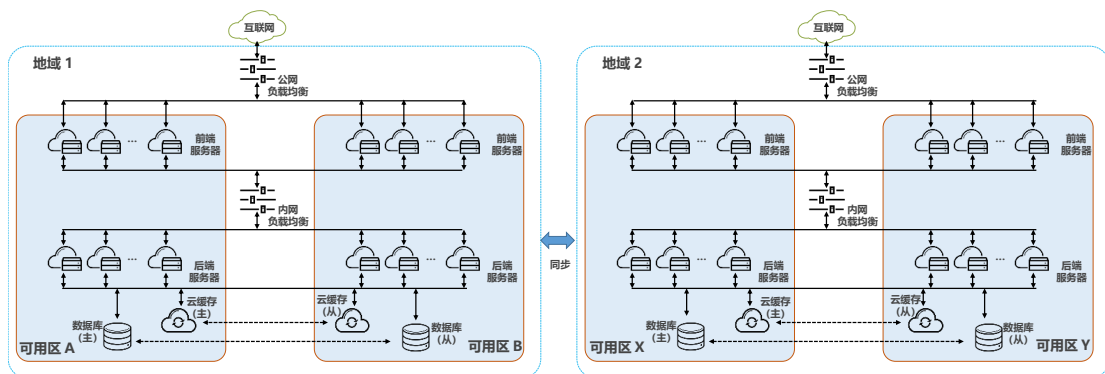
2.2.2.3.1 跨地域主备方案

在两地三可用区部署方案中，用户的业务系统部署在两个地域上的三个可用区内，每个可用区内部署相同的资源和系统，在同一个地域上的两个可用区内的系统同时运行并进行数据同步，另一个地域上的系统是冗余备份的系统，不对外提供服务，数据从地域 Region1 传输到地域 Region2 的备份系统中。倘若某一地域内系统完全不可用，则通过另一个地域的系统提供服务即可，业务仍然可以正常运行。



2.2.2.3.2 跨地域双活方案

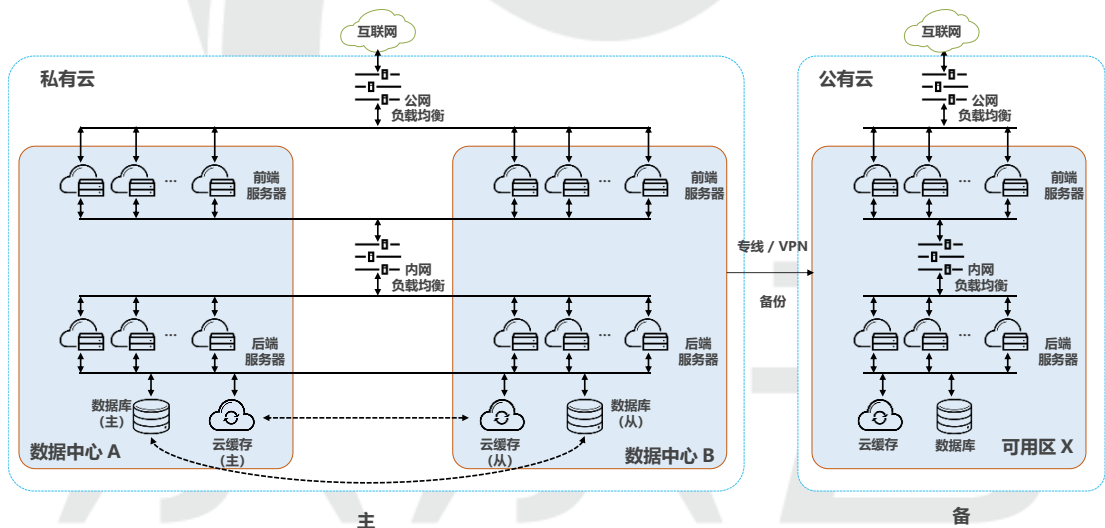
如下图所示，异地多活方案中，多套相同的业务系统部署在两个地域中，每个地域中的业务系统都对外提供服务，两个地域多个可用区的数据会进行同步。



2.2.3 跨云高可用方案

2.2.3.1 混合云高可用方案

混合云是由京东云公有云和用户私有云一起组成的云计算平台，私有云和公有云之间通过安全的高速网络连接。混合云上的高可用架构类似于两地三中心方案。如图所示，在私有云上通过建立多个故障域或多个可用区实现跨故障域或跨可用区的高可用，同时与公有云在不同的地域建立冗余备份系统。



2.2.3.2 多云高可用方案

所谓多云是指同时利用不同的云厂商提供的云计算平台搭建用户业务系统。基于便于管理、数据安全和网络高效传输等技术考虑，若无特殊需求，在搭建高可用系统时不建议用户采用多于 2 个云厂商的云计算平台。多云高可用架构与跨地域高可用方案基本相同，主要是将同一个云计算平台的两个地域更换为在两个地域上的两个云计算平台。

2.3 数据级高可用方案

2.3.1 关系数据库服务 RDS

京东云 RDS（Relational Database Service）产品包括 MySQL、Percona、MariaDB、SQL Server 和 PostgreSQL 等应用最广泛的数据库。所有 RDS 产品的部署方式均默认为主从热备的高可用架构，其中数据库主节点提供读写服务，从节点用于灾备发生时的主从热切换，平时不提供服务。单个 RDS 实例的服务可用性不低于 99.95%。

用户在创建 RDS 实例时，可选择将数据库主从节点部署在同一可用区或分开部署在两个可用区。若主从实例部署在同一可用区，则依托京东云高可用组实现跨故障域高可用。若主从实例多可用区部署，主从实例将分开部署在两个可用区中，通过京东云骨干网络实现内网互通。建议用户选择多可用区部署，以获得最佳数据库服务可用性。RDS 数据库单可用区部署架构如下图（左），跨可用区部署架构如下图（右）所示。



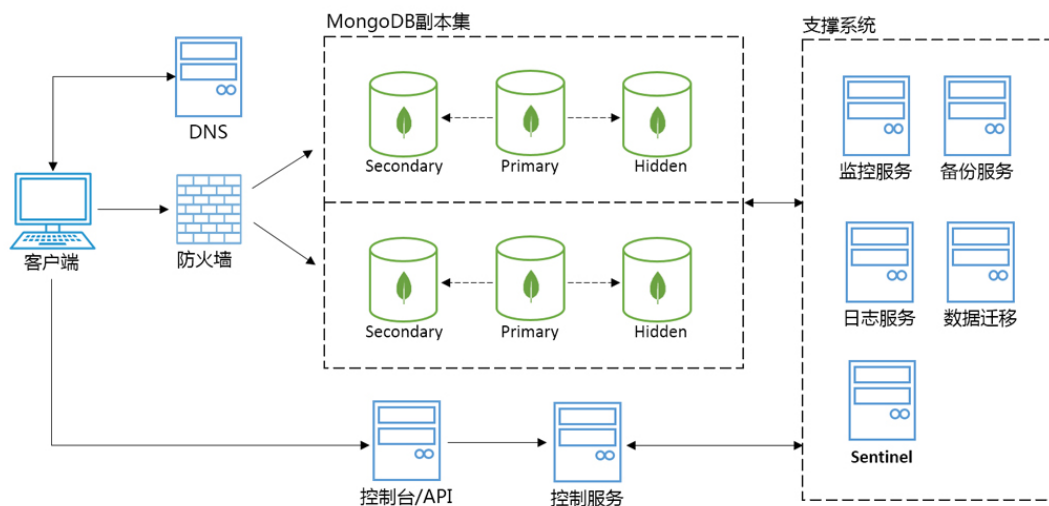
为实现故障发生时主从自动切换，京东云提供 RDS 实例监控模块。监控模块负责监控后端主从节点的健康状态，当主节点不可用时，监控模块将虚拟 IP 地址（VIP）切换至节点，切换时间小于 60 秒。

数据库主从节点间通过半同步方式进行数据同步。支持半同步数据复制策略是为了同时获得最佳的性能和数据一致性保障。异步复制虽然对数据库性能影响最小，但主从节点的数据一致性难以保证；而同步复制虽然保证了主从节点数据的强一致性，但数据库性能会受到较大影响。

京东云 RDS 产品默认每天自动执行备份操作，备份时间可由用户设定。备份文件保存在京东云对象存储 OSS 中，并保存 7 天。利用京东云 OSS 的三副本存储架构，备份数据获得 99.999999999%（11 个 9）的持久性，确保了备份数据不会丢失。

2.3.2 NoSQL 数据库 MongoDB

京东云 MongoDB 默认提供包含三节点副本集的高可用架构，支持故障自动切换。三副本节点包括 Primary、Secondary 和 Hidden 等三个节点。其中 Primary 节点和 Secondary 节点对用户可见，Hidden 节点对用户不可见。单个 MongoDB 实例的服务可用性不低于 99.95%。



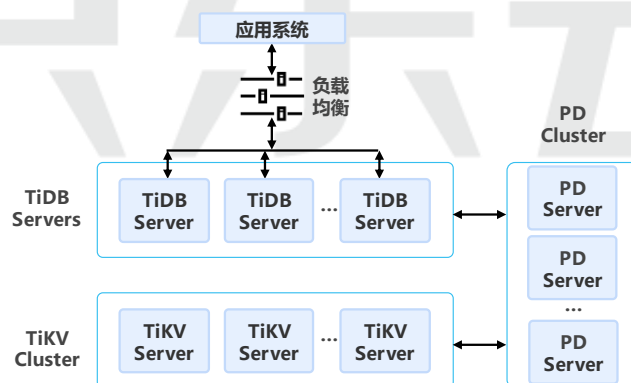
Primary 节点负责处理用户的读和写操作请求。Secondary 和 Hidden 节点通过从 Primary 节点的 oplog 拉取操作日志同步数据。Secondary 节点可以处理用户的读操作请求。

支撑系统负责 MongoDB 副本集的监控、备份和记录日志。Sentinel 服务会监控副本的健康状态，当发现某一副本不可用时，会根据高可用策略自动进行主备切换，以保证服务的高可用性。

用户在创建 MongoDB 实例时，可选择单可用区或多可用区部署。若选择单可用区，则三个节点均会部署在同一可用区中。若选择多可用区部署，则 Primary 和 Secondary 节点会部署在同一可用区，Hidden 节点部署在另一可用区。建议用户选择多可用区部署，以提升系统的可用性。

京东云 MongoDB 产品默认每天自动执行备份操作，备份时间可由用户设定。与 RDS 相同，备份文件保存在京东云对象存储 OSS 中。

2.3.3 NewSQL 数据库 TiDB



京东云 TiDB 是京东云联合 PingCAP 基于国内开源 NewSQL 数据库 TiDB 打造的一款同时支持 OLTP 和 OLAP 两种场景的分布式云数据库产品，支持自动水平伸缩、强

一致的分布式事务，兼容 MySQL 协议，极大减少迁移、使用成本。TiDB 系统架构图如上图所示。

京东云 TiDB 集群主要包含 TiDB Servers、PD Cluster 和 TiKV Cluster 等三个组件。TiDB Servers 负责处理 SQL 请求。TiKV Server 负责存储数据，提供事务 Key-Value 存储引擎。PD Server 负责整个集群的管理，存储集群的元信息，对 TiKV 集群进行调度和负载均衡。

高可用是 TiDB 的突出特点，各个组件都能容忍部分实例的失效。TiDB Servers 实例是无状态的，建议用户至少部署两个实例，通过接入前端负载均衡对外提供服务。PD Cluster 集群的节点间通过 Raft 协议保持数据一致性。非 Leader 节点失效不会影响集群可用性。若 Leader 节点失效，为保证可用性，则集群会选出新的 Raft Leader 并自动恢复服务。TiKV 集群也通过 Raft 协议保证数据一致性，并通过 PD 集群进行负载均衡。非 Leader 节点失效时，不影响服务。若 Leader 节点失效，则服务中断并等待重新选举 Leader 节点进行恢复。

2.4 网络高可用方案

2.4.1 高可用私有网络

2.4.1.1 南北向网络流量实现高可用

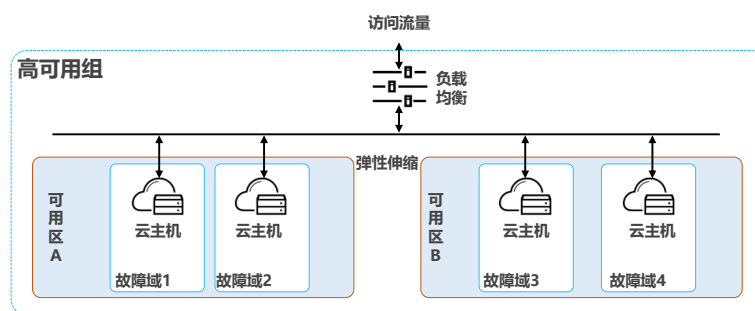
基于京东云负载均衡产品可以实现南北向网络高可用。京东云负载均衡 (LB) 产品可实现将并发流量分发到多个后端实例，不但能够调整资源利用率，还能消除单实例故障对系统的影响，提高系统可用性。京东云提供应用负载均衡、网络负载均衡、分布式网络负载均衡等三种负载均衡产品。

单可用区资源冗余机制

负载均衡实例在集群上部署，采用双活架构，单可用区下至少提供 2 个资源实例。如一个负载均衡实例故障，可自动将故障实例处理的流量转发至另一个负载均衡实例，保障业务正常运行。

多可用区部署

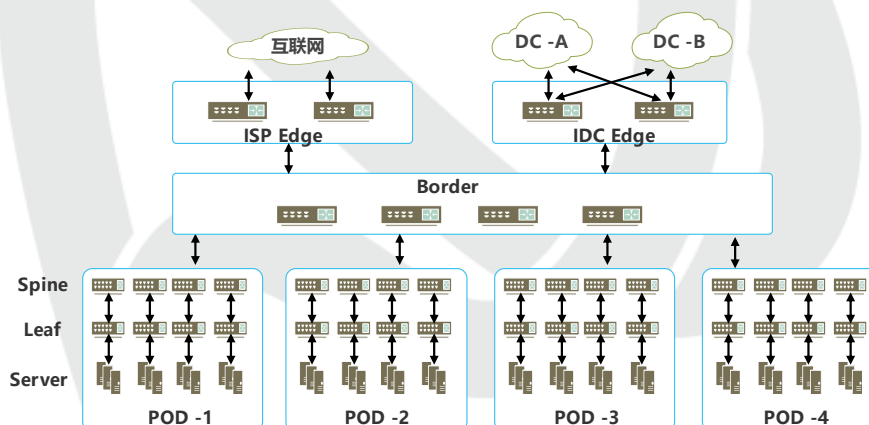
可将负载均衡实例部署在多个可用区，一旦可用区机房或资源故障，可将流量自动分发至其它可用区，实现系统跨可用区高可用。



高可用组

负载均衡支持绑定高可用组，系统可根据业务流量和服务器实例负载情况动态调整服务器数量，实现资源的合理分配，保证业务正常运行。

2.4.1.2 东西向网络流量实现高可用



京东云 VPC (私有网络) 的内部网络设备节点都为双活或多活架构，因此所有东西向网络流量的链路均能实现高可用，任一个 VR 或 TOR 路由器的失效均不会影响链路连通性。

京东云 VPC 还支持跨可用区部署，不同可用区之间的内网东西向流量通过核心交换机房与其它可用区连接。在同一地域内包含两个核心交换机房，每个可用区与核心交换机房均为双上联架构，实现链路高可用。

2.4.2 公网链路高可用

弹性公网 IP

弹性公网 IP 是可以独立申请的公网 IP 地址，支持与云主机、容器、负载均衡、NFV 实例等资源进行动态绑定和解绑。当实例出现故障，通过手动配置，弹性公网 IP 能够漂移到另一实例上，实现故障快速响应和屏蔽。

京东云采用业内领先的双活 vRouter (Virtual Router, 虚拟路由器) 高可用技术，支持高并发连接、超大流量负载以及流量突发等场景，实现链路冗余和网络高可用。

资源冗余

京东云每个地域部署两个核心交换机房，每个核心交换机房与多家主要网络运营商骨干网络相连接，实现了公网链路高可用。每个可用区的公网链路均同时与所在地域的两个核心机房连接，连接均采用双上联架构，保证了可用区到核心交换机房链路的高可用。同时，京东云支持动态 BGP IP 资源。当地域内某网络运营商链路不可用时，可将 IP 地址通过其它网络运营商宣告，实现跨网络运营商的网络高可用。

SD-WAN

京东云可支持用户通过 SD-WAN 技术在京东云上搭建满足高可用的公网链路。借助于 SD-WAN 技术，终端用户可以通过多条网络运营商链路接入到京东云上的服务器，任何一家网络运营商链路的故障都不会影响用户业务系统的正常运行。

2.4.3 网络专线高可用

京东云专线服务是一款高速、安全、稳定的网络接入服务，实现京东云网络与用户的 IDC 网络环境进行内网通信、数据备份及跨机房容灾。

京东云专线服务的所有组件全部采用和支持高可用架构设计。支持客户在同地域接入两根以上的物理链路，以保障物理链路的高可用性。边界网关默认支持双活模式创建，并且边界网关会自动关联到该用户同地域下的全部 VPC。专线通道和托管通道默认支持双活模式创建。从而实现网络专线端到端全路径的双通道高可用。

2.5 高可用系统最佳实践

2.5.1 利用高可用组构建核心业务系统

利用京东云提供的高可用组，能够帮助用户快速实现高可用架构，还能简化用户业务系统的设计，降低运维难度。高可用组产品提供了组内单元在数据中心内横跨多个故障域均衡部署的机制，云主机实例分散部署在相互隔离的物理资源上，当出现硬件故障或实施定时维护时只有部分云主机实例被影响，不影响业务可用性。高可用组还支持跨可用区部署云主机实例，具有更高可用性。

2.5.2 采用云平台 PaaS 产品构建高可用系统

京东云向用户开放的云计算平台 PaaS 产品都经过高可用设计，并向用户提供 SLA 保障。一些用户出于成本、熟悉程度、改造难度等方面的考虑，往往采用自行搭建的数据库、中间件等实现业务系统。由于高可用系统的设计和实现难度较高，用户自建的系统常常不能实现高可用，为系统带来巨大的运行隐患。京东云的 PaaS 产品经过长期的应用检

验，可用性有很好的保障。

2.5.3 跨可用区构建高速的高可用网络

同地域的不同可用区之间网络质量优异，并且通过系统跨可用区可以满足大多数用户的高可用需求。因此用户可以充分利用京东云可用区间高带宽网络的优势构建高可用业务系统。用户可以在同一地域内跨多个可用区创建统一的 VPC 私有网络，在 VPC 内实现内网互联。若系统包含多个 VPC，则可采用 VPC 对等连接进行互联。对等连接不依赖某个独立硬件，因而不存在单点故障或带宽瓶颈。

2.5.4 采用两地三可用区部署实现异地容灾

对可用性和可靠性要求高的用户可采用基于云计算平台的跨两地三可用区的部署架构，实现异地容灾的同时还能确保系统运行的高可用性。用户可以选择在三个可用区内同时部署相同的完整的业务系统。在同一个可用区内系统可以采用高可用组的部署方式来提升可用性。在三个可用区中，使其中两个可用区内部署的独立系统作为对外提供服务的主业务系统，另一个系统作为备用系统，通过数据同步机制保障三个可用区内的系统数据一致。当任意一个可用区或地域发生灾难导致系统故障时，都能保证至少有一个可用区内的业务系统能够正常对外提供服务。

2.5.5 利用脆弱测试发现系统故障单点

某些业务系统设计复杂，各模块间关联关系复杂，在设计和部署时不容易确定隐藏的系统故障单点。因此，需要对用户的业务系统进行全面的脆弱性测试，提前定位系统容易出现关联故障之处，排查故障单点。利用脆弱测试可以帮助用户大大改进业务系统可用性。

2.5.6 基于专线的混合云架构实现稳定系统

混合云架构为用户提供数据和业务安全性的同时可以实现系统的高可用，利用专线连接不同云平台，不但能实现数据高速传输，还能提高可用性。对网络连通性要求高的用户可以采用两个专线提供商搭建的独立专线。若以其中一条专线为主传输数据，则需定期检查备用专线的连通情况。

3 构建可扩展的系统

3.1 系统扩展方法

可扩展性是指云计算平台上的资源可以处理增加或减少的使用需求的能力，反映了资源在弹性范围内变化能够适应工作负载波动的能力。在京东云上搭建业务系统，实现系统可

扩展的目标是让系统运行过程中能够利用云计算平台资源实现理想性能的同时降低资源成本开销，同时使系统能够更好适应业务发展的需要，能够平稳处理突发流量等极端情况。

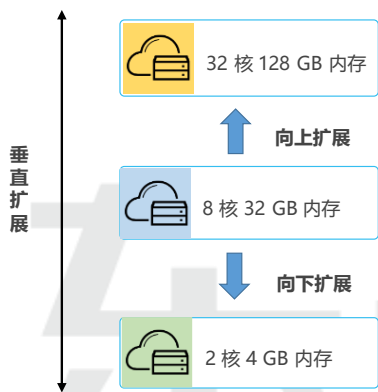
3.1.1 水平扩展

云计算平台上资源的水平扩展包括资源的向外扩展和向内扩展。水平分配新的资源即为向外扩展。水平释放冗余的资源即为向内扩展。对云计算平台上的资源进行水平扩展简单易行，并结合高可用架构设计方法，跨故障域、跨可用区或跨地域对资源进行向外扩展，可以使系统具有更高的可用性。



3.1.2 垂直扩展

云计算平台上资源的垂直扩展包括资源的向上扩展和向下扩展。垂直提升资源的容量或能力等规格即为向上扩展。垂直降低资源的容量或能力等规格即为向下扩展。由于在云计算平台上对资源进行垂直扩展时需要先停止资源实例，需要用用户业务系统能够承受一段时间的资源实例不可用。



3.2 京东云主要产品可扩展性

3.2.1 可扩展架构

京东云基于不断的技术探索和技术经验积累，构建了完整的能够支撑高可扩展性的技术体系。基于京东云稳固的基础设施保障，利用先进的虚拟化平台技术和功能强大的资源管理平台，研发出能够支撑用户实现资源可扩展的计算、存储、网络等云计算产品，助力客户实现具有高可扩展性的业务系统。



3.2.2 弹性计算扩展

京东云弹性计算产品的扩展能力如下表所示。

名称	可扩展性描述
云主机	水平扩展：通过控制台、OpenAPI 实现水平扩展。采用弹性伸缩产品实现基于业务负载的自动水平扩展。 垂直扩展：通过调整实例规格有效调节计算能力。
云硬盘	水平扩展：支持用户同一个地域创建 40 块云硬盘，可用工单提升配额。单个云主机最多挂载 8 块云硬盘。 垂直扩展：容量扩展方面，单个云硬盘容量 20GB 到 4TB。为保证用户原有数据可用，仅支持向上扩展。能力扩展方面，可通过更换云硬盘类型实现垂直能力调整。提供容量型 HDD、通用型 SSD、性能型 SSD 三种类型，性能逐个递增。
原生容器	水平扩展：用户一个地域内可创建 20 个原生容器实例，支持通过工单方式提升配额。 垂直扩展：用户可利用容器镜像创建不同规格原始容器实例实现垂直扩展。但不支持已有实例规格调整。
Kubernetes 集群	支持 Kubernetes 集群扩缩容。通过 Kubernetes 集群服务，借助容器快速启动、快速部署、海量并发的能力，可快速启动成百上千台机器，适应业务突发需求。

3.2.3 网络扩展

京东云网络产品的扩展能力如下表所示。

名称	可扩展性描述
私有网络	同地域不同可用区之间扩展：同地域可用区之间通过骨干网实现互连，同 VPC 内可跨可用区部署云资源，实现可用区之间的资源扩展。 可用区与用户 IDC 之间扩展：用户网络可通过 VPN 或专线的方式与京东云 VPC 进行互连，用户可将部分业务系统扩展到京东云 VPC 内。 利用弹性网卡扩展主机网口：每台云主机可以绑定多块弹性网卡，绑定数量依据云主机实例规格而定。多块弹性网卡可使云主机接入不同网络。
弹性公网 IP	支持带宽调整，用户可根据业务流量变化随时更改带宽，修改即时生效。
负载均衡	水平扩展：通过负载均衡用户可根据业务发展的需要增减云主机实例个数来调节应用系统服务能力。 垂直扩展：京东云提供应用负载均衡、网络负载均衡、分布式网络负载均衡等三款产品，用户可通过选用不同类型产品实现业务系统的性能目标。
CDN	支持将 CDN 边缘节点部署在境外，扩展用户的海外业务。
专线	支持通过多条物理连接进行 ECMP 扩充带宽上限。

3.2.4 云数据库与缓存扩展

京东云云数据库与缓存产品的扩展能力如下表所示。

名称	可扩展性描述
云数据库 RDS	水平扩展：可通过增减 RDS 实例的个数实现扩展，或者选用 DRDS 实现分库分表，增强数据库处理能力。 垂直扩展：可通过更改 RDS 实例的规格实现扩展。
云数据库 MongoDB	通过更改实例的规格实现处理能力的垂直扩展。
云数据库 TiDB	通过增加节点数量适应业务的数据增长，TiDB 根据存储、网络等因素，动态进行负载均衡调整，以保证更优的读写性能。
云缓存 Redis	支持主从版架构和集群版架构，主从版和集群版之间能够灵活变配，水平扩展节点数量。支持在线扩容，不影响用户线上业务。
云缓存 Memcached	存储容量在线扩容，不中断服务，不影响用户业务。

3.2.5 互联网中间件扩展

京东云互联网中间件产品的扩展能力如下表所示。

名称	可扩展性描述
云搜索 Elasticsearch	支持在线弹性扩容，灵活实现按需配置，动态扩容集群以满足业务增长需求。提供 6 种可选主机规格，实现节点运算能力垂直扩展。
消息队列 JCQ	Topic 数量可弹性扩展，集群规模可弹性扩缩。
分布式服务框架 JDSF	基于京东云分布式服务框架，可以很容易地将自身业务迁移到微服务框架上。采用微服务架构是系统实现良好可扩展性的最佳方法之一。

3.3 高可扩展性系统最佳实践

3.3.1 采用无状态应用

无状态应用无需在云主机和云硬盘中保存应用系统的会话状态等状态信息，也不用存储数据库和日志等数据信息。因此在云主机中部署无状态应用后，通过增加云主机的个数，即可达到对应用系统进行水平扩展的目的。系统需要存储的信息可以通过网络在某处集中存储，云主机中运行的服务可以通过网络获取相关信息。

3.3.2 采用微服务架构

通过微服务架构设计，可将复杂的业务逻辑切分成功能专一、耦合度低、高度自治的一组服务。每个服务都可以进行独立扩展，因此系统扩展非常灵活。微服务系统可以通过增加独立的服务实现整体扩展。用户采用京东云 Kubernetes 集群或分布式服务框架 JDSF 都能快速实现微服务架构业务系统。

3.3.3 采用弹性伸缩

弹性伸缩（Auto Scaling）是根据客户的业务需求和伸缩策略，云平台为客户自动调整资源规模的产品。在业务高峰期，会自动增加云主机实例的数量，以保证业务性能不受影响；当业务需求较低时，则会减少云主机实例数量，以节省成本。

3.3.4 创建弹性网络

通过在云主机上绑定弹性网卡实现将云主机接入不同子网。采用弹性网卡可以实现业务流量分离、帮助系统承载多种业务，并且在网络高可用等场景提供有力支撑。

4 构建高性能系统

4.1 简介

4.1.1 性能概述

系统的性能是在指定的时间范围内系统处理任务的速度。在云计算平台上，不同规格的资源实例能够对外输出不同的处理能力，因而影响系统的性能。京东云通过在多个地域向用户提供同一产品的多种规格实例，帮助用户利用丰富的资源构建满足自身业务系统性能需求的系统。

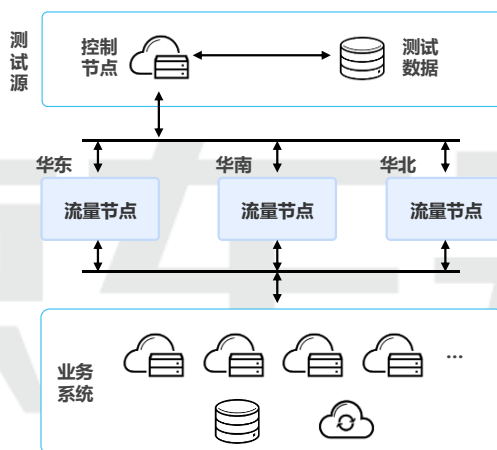
4.1.2 性能提升方法

4.1.2.1 利用云计算平台产品构建高性能系统

京东云提供一系列高性能 IaaS 和 PaaS 产品，为用户实现高性能业务系统奠定基础。京东云产品指标都按业内最高标准进行设计，包括 CDN 响应时间、云硬盘吞吐率、私有网络带宽、负载均衡请求处理速率等指标都达到业内先进水平。用户可以利用京东云产品构建符合性能目标要求的系统。

4.1.2.2 通过性能测试提升系统性能

要构建高性能业务系统，需要进行全面的性能测试，找到系统性能瓶颈并进行性能优化。



基于用户业务系统设计目标，可以对系统进行基于真实业务场景的大流量测试。通过模拟极端业务压力，测试系统的负载承受能力和稳定性。部分系统问题，只有在系统持续运行一定时间后才会表现，因此需要在系统上线前进行长时间高负载性能测试。通过性能测试和评估，可以掌握云资源和业务承载力的关系，并通过对云主机、云硬盘、数据库、带宽等资源数量和规格的调整，实现资源最优配置。

4.2 京东云主要产品性能

4.2.1 弹性计算

京东云弹性计算产品的性能如下表所示。

名称	性能描述	优势
云主机	提供通用型、计算优化标准型、内存优化标准型、高频计算优化型、GPU 标准型等 5 种类型，每种类型配有多个不同的 CPU 和内存规格，满足各种业务的性能需求。	处理能力可弹性伸缩。
云硬盘	提供容量型 HDD、通用型 SSD、性能型 SSD 等 3 种类型，性能逐个增强，每种类型都提供多种容量选择。	适应不同业务系统存储性能需求。
原生容器	提供通用型、计算优化型、内存优化型、高频计算型四种实例规格，适应不同性能需求。最小实例配置支持 1 核 1 GB 内存、最大实例配置支持 72 核 576 GB 内存。	性能损耗小。

4.2.2 网络

京东云网络产品的性能如下表所示。

名称	性能描述	优势
私有网络	京东云对私有网络内云产品间的网络传输没有性能限制，内网速率取决于云主机网卡性能及网卡队列数。	内网速度极快。
弹性公网 IP	最大带宽可达用户购买带宽的 150%。用户带宽若小于 100 Mbps 则流入 IP 的带宽上限是 100 Mbps，只有大于 100 Mbps 时上下行流量才对等。	性价比高。
应用负载均衡	单可用区最大连接数 200 万，多可用区 400 万；每秒新建连接数在单和多可用区分别达 2 万和 4 万。	简单易用。
网络负载均衡	最大连接数在 8 实例集群下达 1.28 亿；每秒新建连接数 400 万。	处理能力强。
分布式网络负载均衡	最大连接数和每秒新建连接数都无限制，性能取决于用户公网 IP 带宽。	极致性能。
CDN	CDN 全网节点数 600+，全网带宽输出能力 10 Tbps+，命中率达 95%以上。	覆盖面广、性价比高。
专线	支持 1000Base-T、1000Base-LX、10GBase-LR 三种接入物理端口，支持 RJ45 电口和 LC 模式光口，单物理连接最大支持 10 Gbps 带宽，可通过多条物理连接进行 ECMP 扩充带宽上限。	接入便捷快速，安全可靠。

4.2.3 数据库与缓存

京东云数据库与缓存产品的性能如下表所示。

名称	性能描述	优势
RDS	支持 MySQL, Percona, MariaDB, SQL Server, PostgreSQL 五种数据库类型, 每种类型数据库都提供最大 16 核 64 GB 内存规格。MySQL 支持最大连接数 1.6 万, 支持添加最多 8 个只读实例, 水平扩展数据库性能。	应用广泛、适应性强。
DRDS	最大支持 2 个 16 核 64GB 内存的 RDS MySQL 节点。	处理能力强。
MongoDB	存储引擎为高性能 WiredTiger, 支持最大 16 核 64GB 内存规格, 最大连接数 1.6 万, 最大 IOPS 达 1.6 万。	应用简单, 功能强大。
TiDB	京东云 TiDB 公测阶段其各项性能指标均远超 MySQL。	极致性能。
Redis	主从版支持最大 64 GB 缓存容量, 最大连接数 1 万, 内网带宽上限 48 MBps; 集群版支持最大 512 GB 缓存容量, 最大连接数 320 万, 内网带宽上限 1.5 GBps。	配置灵活。
Memcached	支持最大 32GB 缓存, 最大连接数 1 万。	简单易用。

4.3 全链路性能测试

4.3.1 测试场景

下面几个场景下一般需要进行全链路性能测试:

- 新系统上线前, 要进行负载能力评估, 找到瓶颈并优化;
- 业务高峰前, 预估流量并考察系统压力承受能力;
- 进行系统容量规划时, 为确定资源用量提供决策依据;
- 系统主要模块更新时, 对涉及到的链路进行测试。

4.3.2 测试指标

衡量云上系统的性能主要指标有:

- 系统响应时间。指服务器接收到任务请求之后处理完成任务并通过网络回复给用户任务处理结果的完整时间。
- 系统吞吐量。指服务器在单位时间内接收并完成发来的任务请求数量。

- 系统并发量。指系统在同一时间并发处理同一类任务请求的数量。
- 系统错误率。指系统在负载情况下，请求失败的概率。

4.3.3 测试方法

4.3.3.1 搭建测试环境

计算环境隔离。选一部分线上运行的云主机放入隔离环境中，并配置只接收压测流量。隔离环境不需要改动业务系统，云主机规格保持与线上规格基本一致。测试完成后隔离环境的资源还原到线上集群中，不影响线上集群容量。

基础数据构造。构造测试数据以线上生产数据为基础，借助数据处理工具进行数据抽取和同步，并对关键数据进行脱敏处理，形成测试基础数据池。利用基础数据池中的数据，并基于业务场景进行构造，最后形成大量可用的压测数据。

数据隔离。测试中，需对压测流量进行标记，并且这些标记能够随中间件协议的调用关系进行传递，应用系统可根据标记识别压测流量。

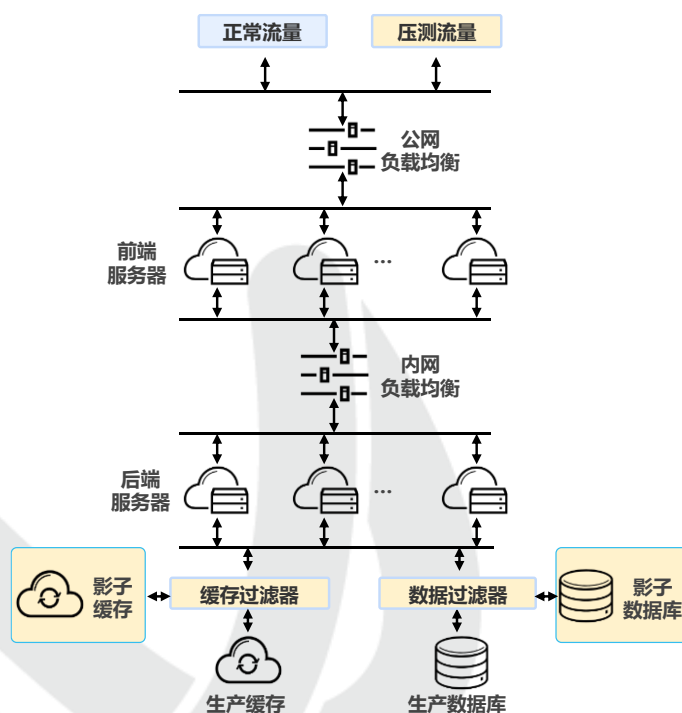
4.3.3.2 实施测试

关键节点。全链路性能测试需要测试的链路中关键节点包括：用户请求模块、数据链路模块、后端业务系统服务程序、数据库和中间件、运行库等。

模拟单用户动作。在测试环境中，设计处理单个用户业务请求的测试方案，对单个用户发出请求后数据流经过的所有关键节点的处理时间进行记录。当对业务进行加压时，只发出一个用户的业务请求，但阶梯式提升请求的频率和数据处理量。单用户模式测试可以定位最耗时的关键节点。

模拟集合用户动作。设计处理集体用户业务请求的并发测试方案，阶梯式提升并发用户数、每个用户的请求频率、数据处理量，对数据流经过的所有关键节点的处理性能进行记录，定位性能相对较低的关键节点。

业务加压。在对业务进行加压时，有包含和不包含用户思考时间两种加压方式。不包含用户思考时间情况下，能够定位极限情况下的系统性能瓶颈。在模拟用户思考时间的情



况下，能够真实反映生产环境中最可能出现的系统性能瓶颈。

性能结果收集与分析。对性能数据输出格式进行统一定义，每条数据都应包含测试条件、测试指标等足够的信息。为了减小结果输出对测试的影响，测试中可将中间结果输出到内存中，适时将结果写入硬盘文件中。性能数据分析应客观，对推断出的任何结论都应标明支撑结论的数据与得到数据的测试环境与测试方法。

系统更新与迭代测试。利用测试数据和分析结论定位到影响系统性能的关键节点，针对其功能和性能特点进行性能改进。性能优化的主要方法有：

- 通过减少冗余操作、重复操作，增加并发和并行处理等方法提升应用程序层的性能。
- 对调用数据库、中间件等的业务处理逻辑进行流程优化，减少不必要的数据操作，将分散数据合并一起处理。
- 通过调整参数和配置等方法优化数据库、中间件的性能。
- 对运行库、操作系统等运行环境的参数和版本进行调整。
- 通过水平和垂直扩展提升计算实例、网络、存储等组件的性能。
- 通过变更计算实例、网络、存储等组件的类型提升性能。

在系统进行调整后，再次重复性能测试，确定系统改进是否确实提升系统性能。

4.4 性能优化最佳实践

4.4.1 云主机性能优化

对云主机进行性能优化是提升用户系统性能的最佳方法之一，主要方法是进行垂直扩展和水平扩展。进行垂直扩展时，可选用更高规格的云主机，或选择云物理服务器替代云主机。若用户选择云物理服务器，需注意其网络与云主机网络在内网是隔离的。进行水平扩展时，用户可在高可用架构下选择多个云主机实例或利用弹性伸缩产品实现性能提升。

4.4.2 网络性能优化

在京东云上实现网络性能优化有多种方案可选。首先，用户可通过提升网络的带宽提升数据传输能力。第二，用户可选择将不必要外网访问的云主机移到内网环境中，利用京东云提供的高速内网进行数据传输。第三，对于应用负载均衡、网络负载均衡等网络设施，可以通过增加实例个数实现网络处理能力的提升。或者选择分布式网络负载均衡产品提升性能。第四，若用户的业务系统向全国多个地区提供服务，则可以选择采用京东云 CDN（内容分发网络）提升终端用户访问系统的网络性能。

4.4.3 存储系统性能优化

用户可通过选择更高规格的产品实现存储系统性能提升，如将容量型 HDD 云盘升级为通用型 SSD 云盘。若升级不能达到性能目标，则可进一步升级为性能型 SSD 云盘，在 IOPS 和吞吐量等指标上都能够得到大幅提升。若通过改变云硬盘类型还不能达到性能要求，则可以在系统中采用内存型云主机产品，并将一部分内存作为存储系统的缓存。

用户利用云数据库存储和处理数据时，可以通过选择更高规格的数据库实例实现性能提升。若性能依然不能达到要求，则可以选择云缓存产品提升数据存储和处理性能。

4.4.4 云数据库与缓存性能优化

在京东云上，云数据库同时支持进行垂直扩展和水平扩展。当需要对云数据库进行垂直扩展时，可以选择规格更高的数据库实例，从而大幅提升数据库的处理能力。也可以选用分布式数据库或采用多个数据库实例的方式实现水平扩展。选择不同的性能扩展方式的标准是不能因为性能扩展影响系统业务的正确性和稳定性。对云缓存进行性能提升可以选择采用垂直扩展的方式，通过提升云缓存的容量规格提升性能。

4.4.5 互联网中间件性能优化

对消息队列 JMQ 和云搜索 Elasticsearch 等互联网中间件进行性能优化的主要方法是进行垂直扩展，即提升实例的规格，进而提升实例的处理性能。对云搜索 Elasticsearch，还可以选择水平扩展的方式提升性能，通过增加集群节点的数量，提升集群的处理能力。若有需要，也可以选择同时进行水平和垂直扩展，以达到更高的数据处理能力。

4.4.6 应用系统性能优化

对应用系统程序进行性能优化是提升系统性能最直接方案，这也往往是最佳方案。若系统比较复杂，那么在开发过程中很难对全局的性能有最准确的把控，因此需要在系统搭建完成后进行统一的性能测试和性能优化。性能优化的最主要方法是逐步消除系统的性能瓶颈。当某个系统最大的性能瓶颈被消除后，则重点对次大的性能瓶颈进行处理。经过多轮的瓶颈消除，系统应用层将逐步收敛到接近最佳性能状态。

4.4.7 应用架构性能优化

当对应用系统的业务逻辑进行全面优化之后，系统性能依然不能满足需求或不能达到预定的目标，则需要对应用系统的架构进行优化和调整。最主要的方法是使架构能够实现水平扩展，通过增加线程、进程、实例等方式提升系统性能。为了更好的利用云计算平台优势，应用系统应该采用并发处理架构，支持水平扩展。

4.4.8 操作系统性能优化

有时，对操作系统依赖较高的业务系统需要选择对操作系统性能进行优化。在云计算平台上，如无特别需要，不建议用户更改京东云提供的官方操作系统镜像的内核或主要模块。因为往往对操作系统的优化并不比对实例进行水平或垂直扩展更好，而且一旦用户对系统进行了较大的改动，则可能遇到稳定性和可用性方面的问题。

5 构建高安全性的系统

5.1 简介

5.1.1 安全威胁与防护

5.1.1.1 云上系统主要安全威胁

云安全联盟发布的“十二大云安全威胁”阐述了用户用云过程中面临的主要安全威胁。其中，对于用户构建云上业务系统，主要将面临如下安全威胁：

数据泄露。它是 2018 年最严峻的安全威胁，Facebook 和国外政府部门都因数据泄露造成巨大损失。若无保护措施，云平台上业务系统也面临数据泄露风险。

数据丢失。硬件故障在云计算平台上较为常见，需要依赖更好的技术手段防止由于部分硬件的故障导致用户的数据丢失。

未授权的访问。未被授权的情况下，攻击者可能通过某种技术方法访问云平台上应受保护的用户系统和资源。实现未授权的访问是进一步攻击和破坏的基础。

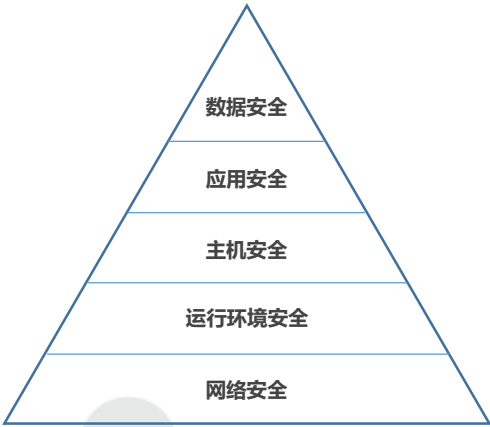
系统漏洞利用。攻击者可以采用技术手段利用系统漏洞实现系统突破，进而进行更大的破坏和威胁。用户需采取必要措施防止系统漏洞被攻击者利用。

渗透攻击。公网上的业务系统容易遭受渗透攻击，因此用户需借助安全防护措施更好地保护系统。

分布式拒绝服务 (DDoS) 攻击。DDoS 攻击是公网服务器容易遭受的攻击。据中国电信云堤和绿盟科技公布的数据，2018 年监测到 14.8 万次 DDoS 攻击。因此几乎时时刻刻都在发生 DDoS 攻击。用户需要采用必要的抵抗 DDoS 措施。

5.1.1.2 安全防护模型

根据系统受到威胁和被攻击的层次,将云上系统安全划分为五层结构,从外向内分别是网络安全、运行环境安全、主机安全、应用安全和数据安全。京东云重点在每个安全层上建立安全防护机制,从而全面的保障用户系统安全。



5.1.2 京东云安全保障

5.1.2.1 云安全架构



京东云遵循业界先进的安全标准保障用户系统和数据全生命周期安全,采取平台安全保护措施、安全管理和信息安全技术等手段进行全面安全体系建设。在云平台的基础安全防护方面,京东云从数据中心自身的安全容灾,到设备的选型和测试、产品和平台的研发设计、数据和网络的访问及控制,经过了严格设计和全面测试,最终为用户提供安全、可靠、稳定易用的云服务。在安全产品和服务方面,京东云为用户提供网络安全、运行环境安全、服务器主机安全、应用系统安全、数据安全等多层次立体式的安全产品集,再加上合作伙伴提供的安全服务,为用户在京东云上提供全面的安全保障。

5.1.2.2 京东云安全资质

京东云的部分主要安全资质如下表所示。

资质名称	主要作用
中国信息通信研究院可信云服务认证	标志着京东云成为国家认可的安全、可信的云服务商。
支付卡行业数据安全标准认证(PCI DSS)	全球最严格金融数据安全认证之一，包含近 300 项标准审核。标志着京东云在开拓金融领域业务方面更具保障性及竞争力。
ISO27001 信息安全管理体系国际认证	标志着京东云信息安全管理能力达到国际先进水平。
公安部信息系统安全三级等保认证	标志着京东云符合国家在信息系统安全方面的技术和管理要求，能够应对信息安全威胁。
云服务企业信用评级 AAA 级认证	标志着京东云提供的服务质量高度可信，并能够引领行业信誉。
CSA STAR 云安全认证	标志着京东云安全管理和技术能力达到国际最先进的云安全标准要求，安全水平值得信赖。
C STAR 云计算安全评估认证	标志着京东云的安全管理和技术能力达到中国地区最高水准。

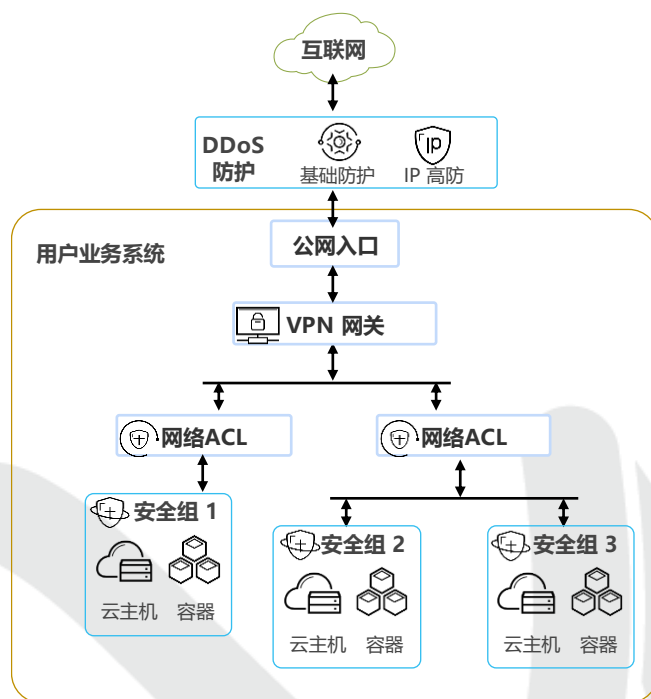
5.2 网络安全防护

5.2.1 网络安全威胁

在公网环境中，用户公网 IP 容易遭受 DDoS 攻击，并且传输的关键数据容易被窃取。在子网环境中，用户会受到非法的南北向或东西向的子网异常流量的威胁。在云主机和容器中，用户会被攻击者扫描开放的端口。以上攻击可能会对用户造成重大威胁和损失。

5.2.2 网络安全防护方案

针对网络分层的特点，京东云采用多层安全防护机制，在网络中全部关键位置部署安全产品来确保用户的公网 IP 资源和 VPC 内部网络资源安全。



为了防止用户公网 IP 遭受 DDoS 攻击并造成重大损失，京东云提供多个级别的 DDoS 安全防护机制，包括 DDoS 基础防护、DDoS 防护包和 IP 高防，这三种防护措施对 DDoS 的防护能力逐个加强。在 VPC 内，用户可通过部署 VPN 网关，建立安全可靠的数据加密通道，防止用户的关键数据被攻击者窃取。在用户子网中，通过部署网络 ACL 来控制子网南北向和东西向的网络异常流量。在用户云主机和容器上，通过建立安全组，防止不应开放的网络端口被攻击；通过主机安全登录机制，利用 SSH 秘钥方法，实现主机登录的高安全性。

5.3 管理和运行环境安全防护

5.3.1 管理和运行环境安全威胁

针对账号管理系统和运维管理系统的主要威胁包括攻击者的攻击和合法用户的危险操作。攻击者一般采用破解管理员口令的方法突破运维系统，而后在未授权的情况下对系统资源进行非法访问，对系统进行破坏或者利用系统攻击其它用户。合法用户有时有意或无意的进行一些危险操作，若系统没有准确识别和记录，则对系统可能造成严重危害。

5.3.2 管理和运行环境安全防护方案

为保障用户系统运行环境的安全，京东云采用多维度安全技术和服务。



IAM 系统支持用户设定主账号和子账号。主账号可以对子账号的资源管理和访问权限进行设置和管理，通过最小权限原则，使子账号仅能对被授权访问的资源进行操作，减小了用户整体系统的风险。

京东云提供的云堡垒机为用户运维系统提供用户身份认证、用户访问授权和人员操作日志审计等功能，使用户的系统可被安全的管理，并且操作事故可追溯。

京东云提供企业版的态势感知安全保障，利用大数据和安全技术的融合，全方位分析用户系统中计算、存储、网络相关安全数据，将用户系统受到的威胁和可能遭受的风险进行清晰的展示，为用户管控安全风险提供强大保障。

5.4 云主机安全防护

5.4.1 云主机安全威胁

攻击者通过暴力破解、漏洞攻击等方法侵入云主机，从而获得主机的控制权，进而对系统进行破坏或数据窃取。攻击者也可能利用病毒或木马针对云主机进行攻击，严重时会造成系统重大损失。

5.4.2 云主机安全防护方案

京东云提供主机安全 V2 产品，为用户提供有效的云主机和云上容器的安全保护，能够检查并防护大部分攻击者针对云主机的渗透和破坏。

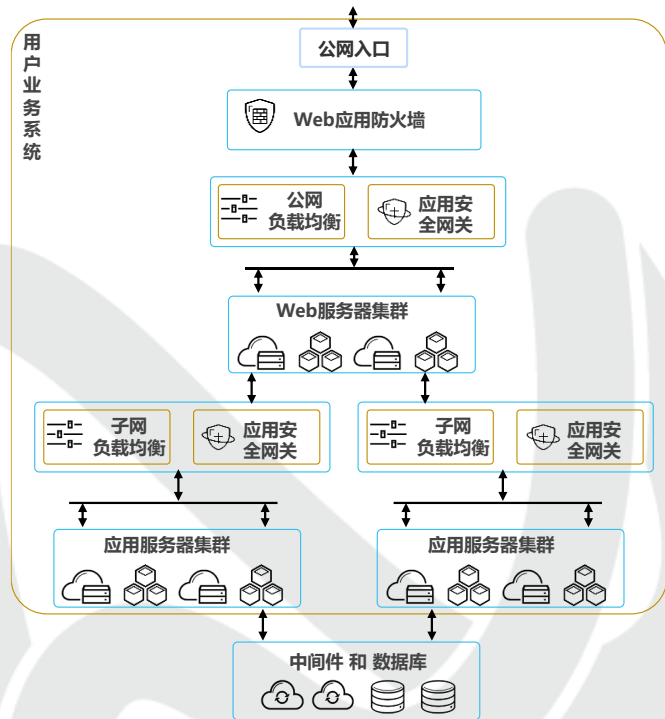
5.5 应用系统安全防护

5.5.1 应用系统安全威胁

针对用户应用系统的主要安全威胁种类繁多，包括针对用户 Web 服务和业务应用服务的 SQL 注入、命令执行攻击、敏感文件探测、恶意扫描攻击、后门文件攻击、XML 注入攻击、目录遍历攻击、跨站脚本攻击等多种攻击，一旦攻击成功，用户业务系统将受到严重威胁。

5.5.2 应用系统安全防护方案

针对应用系统的网络攻击主要采用 Web 应用防火墙和应用安全网关两款安全产品进行保护。



在公网环境下，可以采用 Web 应用防火墙来拦截针对应用系统的攻击，用户可以选择 Web 应用防火墙的高级版、企业版和旗舰版，这几个版本的 Web 应用防火墙的安全能力逐个增强。

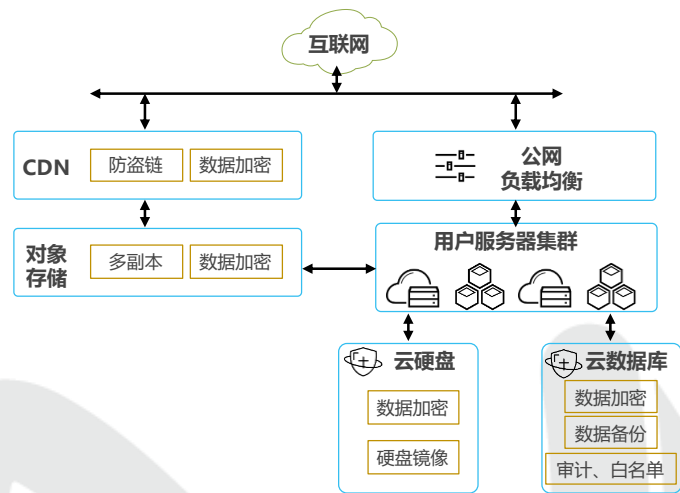
用户可选择在公网负载均衡实例上部署应用安全网关，实现对应用系统网络攻击的保护。应用安全网关有检测和拦截两种模式，用户在需要较高安全保障的情况下需要启用拦截模式来进行安全保护。在子网环境下，可以选择应用安全网关对子网中的应用系统进行网络攻击保护。

5.6 数据安全防护

5.6.1 数据安全威胁

数据安全威胁种类繁多。攻击者会采用盗链或盗取网络数据的方式损害 CDN 用户的权益。攻击者会从对象存储上窃取用户的数据。用户的数据也可能由于对象存储设备的故障导致数据丢失。用户在云硬盘上存储的数据也可能和存储在对象存储中一样被攻击者窃取或由于设备故障导致数据损失。用户存储在云数据库上的数据也存在安全威胁。数据可能由于数据库的故障丢失或不一致。另外非法用户可能对数据库中存储的数据进行窃取。还可能存在合法用户执行危险指令导致的数据库损失或数据库资源耗尽。

5.6.2 数据安全防护方案



在京东云的 CDN 网络上，采用防盗链技术防止攻击者盗取用户源站数据，并且通过视频加密等技术防止用户的视频资源在存储过程中被盗取。在京东云对象存储上，通过建立三个数据副本的方式防止数据损失，通过提供默认加密功能，防止明文数据被攻击者获取。在京东云的云硬盘上，通过建立三个副本的方式为用户提供极高的数据可用性，并在创建云硬盘实例时提供可选的数据加密功能，通过标准的加密算法保护用户明文数据安全。在云数据库中，提供数据加密、数据备份、安全审计和白名单等安全机制保护数据安全¹。

5.7 安全服务保障

5.7.1 基线检测服务

京东云安全基线检测服务为客户提供丰富细致的安全脆弱性现状指数与分析。通过全局基线检测让客户掌握信息系统整体安全脆弱性状况。依据脆弱性检测结果与客户业务模式特点，提供有针对性的安全修补建议，防止安全隐患再次被利用而产生的安全危害。

5.7.2 漏洞扫描服务

漏洞扫描服务是京东云安全团队经过多年深入研究当前各类留下攻击手段（如 SQL 注入漏洞、跨站脚本攻击等）的经验累积。本地检测和远程检测技术相结合，对用户的信息系统进行全面的、深入的漏洞检测，采用状态检测技术，速度更快，结果更加准确。

5.7.3 渗透测试服务

渗透测试服务对现有系统不造成任何损害的前提下，以第三方视角对信息系统进行检

¹ 注：更详细的数据安全方案详见《京东云数据安全白皮书》第四章京东云数据安全实践

查，通过模拟黑客入侵的技术手段对用户指定系统进行全面深入的健壮性测试，发现系统中潜在的弱点，减少云上业务系统的风险隐患，让用户了解系统中安全漏洞可被利用的情况，并针对风险隐患提供修补建议。

5.7.4 应急响应服务

应急响应服务是当用户遭遇网络攻击、木马病毒、数据窃取等黑客入侵事件时，京东云能够提供及时的事件止损、事件分析、系统加固和事件溯源的技术支撑，降低安全事件的影响范围并抑制损失。

6 构建易运维的系统

6.1 简介

6.1.1 运维

运维是在系统上线前进行部署和系统上线之后对系统中所包含的软硬件资源进行管理、监控、维护和故障应急响应等工作。可运维性是指用户对系统实施运维工作时的便利程度和实施的成效。

- 运维管理是对用户云上资源和业务系统进行统一管理，对资源进行合理的规划、部署和配置，对云平台和业务系统的管理账户进行安全管理。
- 运维监控是对京东云上的资源和用户自有的 IT 资源进行监控，掌握资源的运行状态、资源效率、故障情况，对故障进行报警。
- 系统维护是对云上资源和用户 IT 资源进行有计划的维护操作，保障整个系统的网络畅通，应对突增流量，保障系统的稳定运行、安全可靠。
- 故障应急响应是在云计算平台或用户系统出现故障时进行的快速响应动作和操作，包括通知技术人员、快速排除故障、恢复业务运行。后续查找故障原因并进行解决。

6.1.2 运维指标

系统的可运维性是对系统正常运行和更新进行运营和维护的能力，包括对资源的管理、维护、监控和故障应急响应。系统的可运维性对系统持续正常输出服务能力具有非常重要的作用，是影响业务运转和发展的关键要素。

系统的可运维性较难进行定量的衡量，因此一般采用定性的衡量方法，即通过一些组合的方法评估可运维性的好坏程度。衡量可运维性好坏主要可以根据以下几个方面来判断：

-
- 系统部署的工具化和脚本化程度。
 - 系统资源和服务监控的广度和深度。
 - 系统对故障的响应和恢复能力。
 - 系统迭代和更新的能力。

6.2 系统管理

6.2.1 资源管理

6.2.1.1 京东云控制台

京东云控制台是用户管理云上资源的便捷工具，用户可以通过网页登录到控制台，并对资源进行创建、配置、删除等多种操作。控制台通过三级目录方式展现资源，用户可在对应目录下选择资源并对实例进行操作。控制台不仅能够方便地管理云上资源，还支持账号管理、费用管理、工单处理、信息搜索等多种功能。

6.2.1.2 资源管理 API

京东云 OpenAPI 是将京东云所有资源的管理能力通过 API 的方式提供给用户和合作伙伴使用。OpenAPI 是控制台的有效补充，方便用户更灵活管理资源。京东云 OpenAPI 使用 Restful 接口风格，调用 OpenAPI 需要包含如下信息：请求协议，请求方式，请求地址，请求路径，请求头，请求参数，请求体。推荐直接使用京东云 SDK 接入京东云 OpenAPI。京东云 SDK 支持的程序开发环境包括 Java、Python、Go、Node.js、PHP、.Net、C++ 等。

6.2.1.3 标签管理

标签是京东云用户管理资源实例的工具。用户可以通过为资源绑定标签来对资源进行分类，进而更有效地管理资源。用户可根据需要对跨地域、跨产品类型的资源进行多维度分类和管理，从不同维度对资源进行统计和批量管理。

- 资源分类统计。帮助用户快速、准确地统计当前账户下使用的资源。
- 资源筛选。用户可通过标签快速筛选某类资源，并对这类资源批量管理。

目前京东云标签管理支持的资源类型有：云主机、云硬盘、SQL Server 数据库、MySQL、Percona、MariaDB、PostgreSQL、MongoDB、分布式关系型数据库 DRDS、公网 IP、云搜索 Elasticsearch、消息队列 JCQ、云缓存 Memcached 和云缓存 Redis。

6.2.2 账户管理

账户管理。用户可以通过京东云账户管理模块注册京东云账号、设置云计算平台资源的管理员联系方式、设置操作保护和验证机制，并对 Access Key ID 和 Access Key Secret 进行查看和管理，设置虚拟 MFA（多因子认证）认证方式。

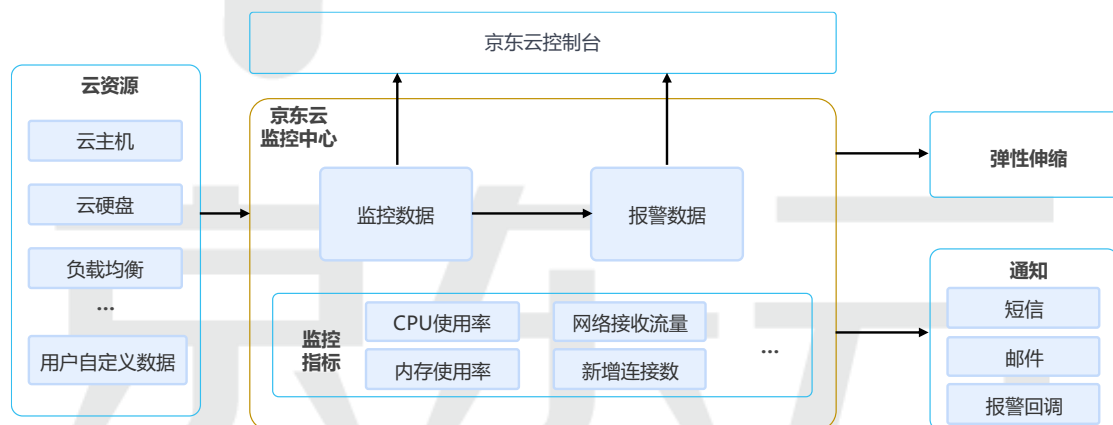
费用管理。用户可通过费用管理模块管理京东云上的费用。支持线上充值和线下汇款。能够便捷地查看使用云资源的消费情况，并管理代金券、发票和合同。

消息模块。用户可以通过消息模块方便地查看和管理京东云向用户发送的账户、产品和安全等三类重要消息。用户可以在控制台上接收到全部消息，也可以设置通过邮箱和手机接收指定类型的消息。

工单系统。用户可以通过京东云控制台的工单系统向京东云客服人员发送用云过程中遇到的问题、咨询用云的方法、开通某些新的功能或者提出投诉或建议。

6.3 系统监控

京东云云监控（Monitoring）是对用户云资源进行监控和报警的服务，展现各项监控指标并对指标设置报警。云监控可通过短信、邮件等方式发送报警通知。通过监控，方便用户了解在京东云上的资源使用情况、性能和运行情况；通过报警，用户可以及时应对，保障系统稳定运行。



6.3.1 云监控

6.3.1.1 监控场景

- **日常业务监控管理。**用户需要对自身的业务进行监控，可以通过所购云服务详情页的监控页面和云监控模块的监控页面对资源的运行状态和使用情况进行监控，同时可以根据云资源的特性设置报警规则。
- **发现异常。**通过设置报警规则，当监控指标达到预设的阈值时，则会向用户设置的

手机号发送短信告警或向电子邮箱中发送邮件告警。

- **查询异常。**用户可根据报警管理中的报警历史发生的时间，进入云产品的监控页或云监控页面查询异常情况。

6.3.1.2 监控类型

- **资源监控。**云监控产品为用户提供云主机、云硬盘、公网 IP、CDN、RDS 等 26 类云资源共 174 项指标的监控数据，为用户全面了解资源运行情况提供强大支撑。
- **站点监控。**站点监控是一款针对互联网网络状况进行探测的监控类产品。用户可使用该产品，设置需要探测的站点地址，选择京东云分布于全国各地的不同运营商网络的探测节点，对站点地址发送类似真实用户访问的请求，从而得到站点地址的可用性和响应时间。
- **可用性监控。**可用性监控是将用户服务器作为探测源，对指定路径或端口发起探测，根据响应时间及返回的状态码判断该路径或端口是否正常、是否可用。支持设置报警规则，当触发报警规则时，可短信和邮件通知用户。
- **自定义监控。**自定义监控功能为用户提供上报监控数据的接口，方便用户将采集的时序数据上报到云监控。支持 OpenAPI 和命令行工具 CLI 方式上报。
- **监控报警。**京东云监控报警功能帮助用户基于预设的报警模板设置符合用户业务特点的报警项。用户可以设置资源、站点、可用性、自定义等监控类别中监控项的报警触发规则。

6.3.2 基于 DevOps 的监控

京东云 DevOps，是京东云基于自身自动化运维能力，构建的一套 DevOps 产品。京东云 DevOps 提供从基础资源到服务可用性、服务性能和业务逻辑的全链路智能监控服务，为用户提供全方位监控保障。智能监控主要包括：

- **数据采集。**通过采集插件获取需要监控的数据并使其标准化。支持机器监控、死机监控、端口监控、进程监控、日志监控、方法监控、组件监控、域名监控、自定义监控等 9 种监控类型。
- **数据可视化。**提供趋势图、仪表盘、单 IP 搜图三种不同可视化功能。
- **报警配置。**通过丰富的异常检测手段，做到精准报警。
- **智能分析。**具备多维度数据分析能力，覆盖故障发现、故障定位、故障恢复整个故障生命周期，用流型图整合报警、部署、服务树变更等事件，方便用户快速排

查和定位问题。

6.3.3 基于云计算管理后台的监控

京东云的云计算管理后台由于系统安全等原因并不向用户直接开放。对于使用京东云的VIP用户，京东云会派出技术专家进行保障。VIP用户可以委托京东云的技术专家基于云计算管理后台的监控功能实现对用户云上资源和业务系统活性的监控。

6.3.4 监控服务

6.3.4.1 监控指导服务

京东云的监控指导服务是基于京东云技术人员在监控领域的多年经验积累，为用户提供在京东云上实现更好监控效果的服务，帮助用户快速熟悉京东云提供的多种监控能力，提升用户负责业务系统运维的人员监控能力。

6.3.4.2 监控托管服务

京东云的监控托管服务是用户将系统监控的工作托管给京东云实施，京东云技术人员利用丰富的监控经验和管理后台的监控功能实现对用户云上资源和业务活性全方位监控。当发现故障时，京东云技术人员会进行更及时处理。

6.3.4.3 全景监控托管服务

京东云监控托管服务是用户将系统监控工作托管给京东云实施，京东云技术人员在用户云主机上部署专业监控代理程序，并基于监控代理程序的数据，利用丰富的监控经验对用户系统进行更专业深入的监控，并及时处理故障。

6.4 DevOps

6.4.1 介绍

京东云 DevOps，是京东云基于自身自动化运维能力，针对公有云、私有云和混合云的场景和特性，构建的一套 DevOps 产品，提供以 CI/CD 为核心，结合代码管理、自动化测试、智能监控等的一站式 DevOps 服务，与京东云资源无缝对接，助力用户快速实现高效运维管理。

6.4.2 功能

6.4.2.1 配置管理

提供配置管理数据库（CMDB）。CMDB 是 DevOps 编译构建、线上发布等功能的基

础，具有可视化的拓扑服务树。实现了基于服务树的角色权限管理，支持主机资源池管理和分配，同时提供功能丰富的 API 接口。

6.4.2.2 持续交付

一站式完成应用的生命周期管控，包括编译构建、程序包部署、启动、停止、自动扩缩容和应用下线等，实现对应用的全流程管理。支持自定义语言构建环境。支持快速扩缩容以应对业务爆发增长。

6.4.2.3 智能监控

提供从基础资源到服务可用性、服务性能和业务逻辑的全链路智能监控。

6.4.2.4 运维工具

提供海量且高效的运维操作平台。DevOps 提供高可用、安全高效的镜像广场服务，并提供脚本执行、文件分发等自动化服务，满足各种复杂运维场景，实现 Web 自动化运维，有效提高运维工作效率和可靠性。

6.4.3 优势

- **跨平台管理。**支持 VMware、OpenStack、物理机等统一接入和统一运维管理。
- **规模伸缩。**可管理任意规模应用，支持实例弹性扩展，应对业务爆发增长需求。
- **服务自动迁移容错。**云主机宕机时可保持内网 IP 不变，并自动迁移并部署应用，保障业务不掉线。系统自动监控服务健康状态，动态调整集群，实现故障自愈。
- **提升交付能力。**通过统一操作平台，搭建开发和运维团队之间的桥梁，提高应用编译、上线、部署效率。统一收集应用实例日志，能够快速查询、检索、下载，帮助快速定位问题，助力用户快速实现 DevOps，支撑用户业务快速稳定发展。

6.5 故障应急响应

为了保障京东云产品和服务突发异常导致用户业务系统也异常时能够快速应对故障并解决问题，京东云为重要客户提供优质的故障应急响应服务，目标是在异常情况下尽快恢复客户业务。京东云故障是指客户在使用京东云过程中，京东云平台对外、对内的基础设施和服务出现中断、性能下降、失去或降低原有功能、未能给客户提供稳定的服务等影响客户业务持续开展并需要采取紧急处置措施应对的事件。京东云根据客户业务的异常程度，将故障分为 S0 到 S4 共五个级别，并根据级别的不同采取响应的应对措施。

7 构建高性价比系统

7.1 简介

使用云计算平台相比于建设传统的 IT 信息系统能够极大的降低成本，这已成为业界共识。然而，即使用户采用云计算平台，必要的成本付出是实现业务系统和维护业务系统正常运行的基本保障，没有一定的成本付出也不会有预期的收益。因此，要实现高性价比目标并非仅靠压缩成本，而是在保障系统效能的前提下，通过科学合理的规划和操作，减少系统成本。用户采用京东云计算平台这个决策本身就能大幅减少新技术研发时间和人力成本，因此京东云已为用户实现更好的成本指标打下了很好的基础。

本文将系统功能、可用性、性能、安全性、可扩展性、可运维性等要素统称为系统效能。则成本是在指定的一段时间范围内，同等系统效能条件下的系统成本开销，主要包括时间、人力、资源成本等。衡量成本的主要指标包括：

时间成本。指完成业务需求并实现预定的系统效能的前提下，完成云上系统部署、测试和正常运行所需要的时间。时间越长则成本越高。

人力成本。指完成业务需求和进行系统管理、维护和升级所要付出的人力资源。在相同系统效用下，投入人员越多、人才能力级别越高则指标越高。

资源成本。指进行系统搭建和正常运行所需要的 IT 资源的开销。适当的资源成本开销是保障业务对外输出服务能力的基础，资源成本并非越低越好，而是需要控制在一个最佳的成本区间范围内。低于最佳成本区间会导致系统业务输出能力不足或面临较大风险，而高于最佳成本区间则会导致资源浪费。

资源利用率。指资源被应用的程度与其能提供的最大能力的比率。资源利用率过大会使系统的脆弱性增大，对突发流量失去适应能力；资源利用率过小会导致资源空置和浪费。

京东云能够为用户提供成本优化服务，利用京东云在控制成本上的多年经验和技術积累，帮助用户实现长期降本增效。

7.2 时间成本削减

7.2.1 采购、建设时间成本

在传统 IT 信息系统搭建过程中，设备采购和基础设施建设时间少则数月多则数年。若用户属于正在迅猛发展的行业，则几个月的时间开销将大大削弱用户在市场竞争中的优势。用户选择使用京东云的公有云平台搭建业务系统，则可以完全节省硬件设备采购、基

基础设施建设等时间成本，为时效性敏感的行业用户带来巨大优势。

当用户由于安全、可控等原因选择京东云的私有云、混合云部署业务系统时，私有云部分的设备采购和基础设施建设将带来一定的时间成本。由于京东云长期以来建立了优质的采购渠道和积累了丰富的基础设施建设经验，可以帮助用户大幅减少采购和建设时间成本。

7.2.2 研发时间成本

新技术不断更迭，当用户需要原业务系统适应新需求而采用新技术时，若选择自行研发，则将产生巨大研发时间成本。京东云基于京东集团在互联网行业领域多年的技术积累，提供的产品紧跟最新技术发展，并提供稳定可靠版本。因此，用户可以选择采用京东云新技术产品改进业务系统，可大大缩减研发时间成本、提升业务系统品质。

7.2.3 维护时间成本

大体量业务系统中，对种类和数量众多的 IT 资源进行维护会产生巨大工作量和耗费时间成本。通过京东云提供的高可运维性能力，用户在京东云控制台、OpenAPI、DevOps 等工具的帮助下可大大缩减系统维护时间成本。京东云 DevOps 产品还能使用户获得持续开发和持续更新的能力，促进业务快速发展。

7.2.4 战略转型时间成本

历史上很多成功的国际大型企业，为了不断适应新环境而适时进行战略转型，比如采用新技术、利用新平台并变换主营业务，从而保持长期竞争优势。京东云丰富的产品为用户实现快速战略转型奠定了基础。京东云能支持多种行业用户，这意味着用户可以借助京东云平台转换业务方向，实现快速战略转型。

7.3 人力成本削减

7.3.1 研发人力成本

人力成本是企业最大的成本支出之一。研发人员是高端技术人才，人员成本相对较高。用户采用京东云平台则可以节省大量研发基础平台的人力，技术人员可以专注于业务系统研发，从而大大降低系统研发的人力成本。

7.3.2 维护人力成本

系统基础设施和 IT 资源的维护工作难度大，需要专业技术人员。京东云提供控制台、OpenAPI、DevOps 等工具化的系统维护能力，不但可以压缩运维人力成本，还能减少运维过程中出错的几率。

7.4 资源成本削减

7.4.1 基础设施

采用京东云公有云平台，用户将大幅削减 IT 基础设施建设的资源成本。用户无需搭建和维护机房，无需重新建设网络设施，并且无需因合规性要求为基础设施申请资质认证。部分用户可通过搭建京东云的专线来实现网络的高效、安全通讯，其产生的成本不但可控还比较低。

7.4.2 硬件资源

硬件资源具有时效性。当硬件资源超过保质期，其故障率会大大提升，对业务系统稳定性和安全性造成巨大威胁。因此硬件资源需要按一定周期被更换，形成巨大连续性开销。用户选用京东云的公有云将完全避免硬件资源更新带来的资源成本。

7.4.3 软件资源

京东云能够支持用户采用海量的开源免费软件产品，也支持用户以合理的价格使用商业软件。用户可以根据自身的需求调整使用软件资源所付出的成本，以实现技术能力和软件成本之间的平衡。

7.5 绩效提升

7.5.1 业务扩展与转型

京东云能支撑各种行业、各种规模用户的业务系统。基于京东云，用户依据市场趋势和行业前景，能够实现快速业务扩展或转型，从而把握机遇并大幅提升业务绩效。

7.5.2 融入生态体系

京东云基于已有的数十万用户构建了一个非常庞大完整的生态体系。采用京东云的用户能够利用京东云生态体系提供的巨大商机实现业务的不断发展。

8 六高服务

京东云在帮助客户实现业务系统迁入京东云时，为保障客户业务系统具有高质量的架构，提供优质的系统优化服务。系统优化服务是建立在京东云基础服务、技术保障服务和云迁移基础服务之上的高端服务，包括高可用优化、扩展性优化、性能优化、安全优化、运维优化和降本增效等服务，亦称六高服务。

8.1 高可用优化服务

为用户提供改进业务系统可用性的架构技术指导，通过全面分析用户业务系统可能出现的影响系统业务连续性和有效性的风险隐患，基于京东云产品特点，为用户提供提升可用性的解决方案。

8.2 扩展性优化服务

基于京东云系统运维最佳实践，通过充分评估用户业务系统的运行特点和发展潜能，帮助用户设计关键节点的扩展方案，利用京东云产品提供的水平扩展和垂直扩展能力，实现用户业务系统的更好的可扩展性。

8.3 性能优化服务

基于京东云在系统性能优化方面的技术积累，提供业务系统全链路性能测试方案设计和实施服务，帮助用户快速定位系统性能瓶颈，通过调整京东云上用户系统的架构和使用的产品，提升用户业务系统的整体性能。

8.4 安全优化服务

包括基线检测、漏洞扫描、渗透测试和应急响应等服务。挖掘用户系统脆弱性状况，为用户提供脆弱性分析，进行全面、深入的漏洞检测。测试并发现系统中潜在弱点和风险隐患，制定有效解决方案。在用户系统遭遇入侵事件时，提供事件止损、分析、系统加固和事件溯源等技术支撑。

8.5 运维优化服务

包括监控指导服务、监控托管服务、全景监控托管服务。监控指导服务指导用户完成资源监控、业务监控、安全监控等重要监控项的配置指导。监控托管服务由京东云技术人员帮助用户进行资源监控和业务监控。全景监控托管服务实现更全面和进程级别细粒度资源监控和业务监控。

8.6 降本增效服务

帮助用户对京东云上整体系统的资源消耗和业务绩效贡献进行全面的梳理，定位不必要的冗余资源和不合理的架构和资源配置，为用户降低系统运行成本，提升系统资源的利用率提供技术支持。

9 总结

构建高质量的云上系统是一个高技术难度的系统工程。京东云基于长期的技术实践提出构建具有“高可用性、高可扩展性、高性能、高安全性、高可运维性、高性价比”等“六高”特征系统的方法，并基于技术积累分享相应的最佳实践，帮助用户在京东云上构建最好的系统。通过京东云提供的“六高服务”，用户可以便捷的获得系统基础设计和架构设计的强大支持，为业务的长期稳健发展奠定坚实的基础。

10 引用

- [1] “十二大云安全威胁”，云安全联盟。
- [2] 《云计算概念、技术与架构》，Thomas Erl 等。
- [3] 《计算机体系结构-量化研究方法》，John Hennessy 等。

京东云