

云生态

Cloud Ecosystem

特刊

第9期

Vol.09



{ 2015 中国公有云服务 } 发展报告

打造中国最优质的云生态媒体

Geekbang | 极客邦科技

InfoQ



08 | 团队建设篇

18 | 产品研发篇



33 | 服务运营篇

50 | 用户体验篇

67 | 其他讨论

云生态特刊

本期主编 魏 星

本期主笔 蒋清野

发 行人 霍泰稳

联系我们

提供反馈 feedback@cn.infoq.com

商务合作 sales@cn.infoq.com

内容合作 editors@cn.infoq.com

卷首语

一份“过期”的调查报告

初识蒋清野

在 2012 年 12 月 8 日 OpenStack 中国行（天津站）的活动上，我跟蒋清野有了第一次的接触。彼时的他布衣布鞋，颇有“一蓑烟雨任平生”的感觉，登台分享时先趋步向前，鞠一个九十度的大躬，台下听众多是读着他博客长大的后生。蒋清野十分低调内敛，讲起话来思维缜密、声音轻柔。他自我介绍时用“农民、程序员、技术翻译”这样的字眼。蒋清野曾先后在美国导航与控制有限公司、Sun 公司、北京交通大学软件学院、天涯在线、Eucalyptus Systems 等多家单位任职，负责多个不同方面的研发与管理工作，也曾与人联合创办 ezCloud 是国内云计算领域的先驱。两年多前，他再度求学，目前在悉尼大学读博。

2015 年 12 月，在 InfoQ 中国负责云计算领域的我策划年终盘点，读到蒋清野的《浅谈“中国”语境下的公有云发展》，有感于他文中的观点，决心请他

操刀盘点一下国内公有云的发展状况。选择蒋清野有两个原因，一来是因为他本人是云计算老兵，在这个领域比我这个媒体人更专业；二来是因为他本人旅居海外超过两年，从旁观者的角度对国内的公有云环境观察和描绘起来会更客观。

“做点对别人有用的东西”

起初我策划的是一篇几千字的文章，在和蒋清野沟通的过程中，产生了一些观点的碰撞。他告诉我说，我们要做就做点对别人有用的东西。

就这样，我把事情搞大了。

我首先选出了几家 2015 年国内公有云服务商代表，然后起草保密协议，与各家相关的负责人联系，沟通我做这件事的初衷。最后除了阿里云拒绝了我的邀约外，其余各家在保密协议的基础上，相关负责人与我和蒋清野分别进行了电话访谈。

所有访谈结束后时间逼近了 2016 年 5 月，原定 6 月底交稿，由于蒋清野在学业上的繁忙，时间步入了 7 月。在初稿完成后，蒋清野告诉我说，从内容来看，这篇报告怕是对国内公有云行业有消极的影响。在了解情况后，我也不得不开始作这方面的考虑。正当我迟疑的时候，2016 年 8 月 1 日，AWS 宣布落地中国。

蒋清野再次联系我，由于他同时也是 AWS 悉尼技术支持中心的员工，怕这层身份会引起更大的误会，意欲雪藏这篇报告。这让我陷入了更大的两难境地——自 6 月以来，接受访谈的公有云服务厂商几度询问我访谈文章的进度。我一时踌躇不决并跟几个厂商知会了当时的情况，有些厂商不免担心起来——似乎以蒋清野此时的身份来发布这样的报告有些欠妥。

我怎么看这份报告

在经过一番艰难的三方沟通后，我决定放出这篇报告。

我首先说服蒋清野打消顾虑。我们获取到的数据和内容足以客观公正（这里我想提一件小事，为了撰写报告蒋清野在国内几家公有云服务商开通了帐号并进行了测试，由此产生了一笔费用。当我提出要给他报销时候，他拒绝了，说这样会有失公正。）；目前蒋清野在 AWS 悉尼技术支持中心工作，与国内市场相关度

较小；保密协议正是为了让受访者告诉我们一些“真实的信息”，鉴于中文媒体领域有价值文字越来越少，我们既然做了，投鼠忌器不放出来反而更容易遭人揣测。

我相信这篇报告对国内公有云行业是有所影响的，但这个影响并不是几篇文章、几家媒体能够引导和左右的。我相信市场的发展变化必然遵循其客观规律，即便如此，对本篇报告乃至媒体主观导向和公正与否的质疑依然会存在，索性我坦坦荡荡地放出来。

“过期”报告的价值

我自诩是“一个码字的”，近半年来多次说过，“一篇文章最大的价值是什么？不是宣扬了某个真理、道出了某种正确；而是给人带来思考，并让人因此有所得。”

的确，此时发布一篇题为《2015 年 ××× 报告》的文章太过时了。但我这么做有两方面的考虑，首先是内容本身是对 2015 年公有云服务商的访谈，对受访企业、对行业的所谓影响此时已经可以降到最低；其次是我们不是 Gartner 之类的机构，不善于预测未来，而以史为鉴是中国人的传统，希望国内公有云服务商以及广大从业者能继往开来。

报告全文含“团队建设、产品研发、服务运营、用户体验、其他讨论”共 5 个章节、将分成 7 篇文章刊出。

最后，我向参与本报告访谈的蒋清野、黄允松及青云团队、季昕华及 UCloud 团队、李爽及美团云团队、钱广杰及盛大云团队、沈志华及又拍云团队、王慧星及腾讯云团队、许式伟及七牛云团队、朱桦及金山云团队致以诚挚的谢意。

魏星

背景介绍

2015 年 12 月，InfoQ 的编辑魏星邀请作者撰写一篇关于中国公有云服务发展状况的文章。因为作者个人对公有云这个领域一直抱有很大的兴趣，便贸然答应了下来。在这篇文章的准备过程中，作者系统地阅读了国内较为知名的几份云计算白皮书^[1, 2, 3]。作者发现这些报告大都高瞻远瞩提纲挈领，缺乏对具体的公有云服务提供商的描述，未能让读者一窥国内公有云服务发展之真实面貌。在 InfoQ 的协调下，作者与国内多家公有云服务提供商的主要负责人进行了电话访谈，围绕团队建设、产品研发、服务运营这三个问题进行了讨论。除此之外，作者也在本文所探讨的所有公有云上都注册了账号，从用户体验的角度进行了一些小规模测试。这篇文章的目的，便是从团队建设、产品研发、服务运营、用户体验等四个方面对中国的公有云服务发展状况做一个简要的综述。

根据美国国家标准技术研究院（NIST）的定义^[4]，云计算在服务模型上可以划分为软件即服务（SaaS）、平台即服务（PaaS）和设施即服务（IaaS），在发布模型上又可以划分为私有云、社区云、公有云和混合云。需要说明的是，随着云计算技术的发展，如上所述服务模型和发布模型之间的界限也日趋模糊。在本文的范畴内，“公有云”一词泛指面向公众开放服务的平台即服务和设施即服务。除此之外，各种名义的私有云（Private Cloud）、专有云（Dedicated

Cloud)、托管云(Managed Cloud)均未包括在本文的范畴之中。

本文中“团队建设”、“产品研发”、“服务运营”三个小节的数据来源有两个。一个是云服务提供商主动发布的新闻资讯,另一个是作者与云服务提供商的主要负责人之间的电话访谈。作者与黄允松(青云)、季昕华(UCloud)、李爽(美团云)、钱广杰(盛大云)、沈志华(又拍云)、王慧星(腾讯云)、许式伟(七牛云)、朱桦(金山云)等业内专家(按姓氏拼音排序)的访谈,是由InfoQ方面统一协调安排的,在此作者深表感谢。这个三个小节的内容,在定稿之前均经过受访者及其公关/市场团队的确认,反映的是云服务提供商自身的观点和思路。在审稿阶段,青云撤回了与作者进行访谈时所发表的一切言论;出于保护商业机密的考虑,阿里云拒绝了作者的访谈邀请。因此,如上三个小节未能包括青云和阿里云的观点。

“用户体验”和“其他讨论”这两个小节,是作者独立获得的数据以及由此引出的观点,在定稿之前未接受任何一家云服务提供商的审核。需要特别说明的是,如上所述云服务提供商的主要负责人接受作者的访谈并不代表他们认可作者在“用户体验”和“其他讨论”这两个小节中所报告的数据和观点。此外,作者本人也并不持有本文中所讨论的任何一家云服务提供商的内幕信息,作者独立获得的数据仅仅是基于作者所使用的测试方法得到的观测结果。受种种技术条件的限制,作者无法对这些数据的准确性进行背书,也无法对其误差范围进行估算。本文中报告的大部分数据是在2016年3月底之前获得的,这部分数据的获取时间在正文中不再特别说明;小部分数据是在2016年8月底获得的,这部分数据的获取时间在正文中会有特别说明。读者在引用本文所报告之数据时,应当考虑到数据的时效性。

本文中有多个小节对各个云服务提供商进行了逐一介绍。相关云服务提供商在这几个小节中出现的顺序是按照拼音字母次序排列的。

本文仅讨论中国本土的公有云服务提供商。Amazon Web Services(AWS)、Microsoft Azure、Google Cloud Platform等等进入或者未进入中国市场的外资企业不在本文的讨论范围之内。

2015年中国公有云服务发展报告

团队建设篇

在这个部分，我们以团队建设为主线，探讨不同公有云服务提供商的发展历程。我们关心的问题包括团队的历史变迁、融资状况、成长速度、人员流动、研发部门与产品部门的关系，等等。

金山云

公开的资料表明金山云成立于 2012 年初，2013 年 8 月完成 A 轮 2000 万美元融资，2015 年 3 月完成 6666 万美元的 B 轮融资。2016 年 2 月，完成 6000 万美元的 C 轮融资。2016 年 5 月，完成 C+ 轮融资 4890 万美元。实际上金山云的前身是珠海 WPS 项目孵化出来的金山快盘，同时运营 2B 和 2C 两类业务。2014 年 9 月，金山快盘个人版等 2C 类业务出售给迅雷，剩下的 2B 类业务发展为现在的金山云。

2014 年初，金山云的员工总数在 50 人左右，2014 年底发展到 120 人左右，业务也从单纯的企业云存储扩展到存储、计算和网络。2015 年 3 月，员工总数



达到 200 人，并且开始扩充市场和销售人员。2015 年 12 月，员工总数达到 560 人；2016 年 1 月，员工总数达到 600 余人。

金山云将自身定位为一个工程师文化驱动的团队。在面向国内同行进行招聘时，金山云有明确的导向性。与此同时，金山云认为一个团队的价值取决于几个关键的人，因此愿意高薪聘请行业内的顶尖专家。目前金山云研发人员的比重占 60% 左右，主力研发人员大部分来自微软、HP、Cisco、LinkedIn 等大型外企和 BAT。朱桦认为：“在云计算技术领域，我们国内并不差于国外，但是在创新方面存在畏惧心理。国内往往觉得做得跟 AWS 一样就差不多了，很少想要超越 AWS。国内在工程领域很强，但是创新有待加强。”

从 2014 到 2015 年，金山云的估值增长了 10 倍，团队的规模也增长了 10 倍。一个团队在如此之短的时间里发生如此巨大的变化，一定伴随着大量人才流动的情况发生。在某一阶段成为骨干的员工，在另一阶段可能成为阻碍的力量。在这个过程中，有一部分员工选择退出，是很正常的情况。

在离职员工的保密协议与同业竞争禁止协议方面，朱桦认为金山云的心态是开放的，到目前为止没有针对任何员工执行过保密协议或者同业竞争禁止协议。与此同时，云计算行业确实存在激烈的竞争，金山云保留在必要的情况下对离职员工采取法律手段保护公司知识产权的权利。

金山云设有产品经理职位，对部分产品实行产品经理负责制。朱桦也坦诚地认为自身对产品经理的定位存在困惑。在过去两年的实践中，具备一线研发与运维经验的产品经理的表现比较出色；缺乏技术背景的传统产品经理的表现相对较差，但是依然有个别成功的转型案例。

在设有产品经理的情况下，技术驱动的思路又是如何体现的呢？研发团队决定的是产品的路线图，还是限于产品的具体实现方案？朱桦认为**技术创新是为了解决业务问题，因此研发团队需要满足业务部门的要求**。这是一个大的原则，但是在处理具体的问题时允许偏离这个原则。这些具体的问题包括功能与成本的平衡，稳定性和功能多样性的平衡，自动化和可靠度的平衡。譬如说在设计数据中心布局时，全球范围内部署几个区域，每个区域部署几个可用区，可用区之间的延迟参数，都有严格的规定。但是在起步阶段，即使是国内顶尖的团队，也达不到这个严格的标准，一定会有偏离的情况出现。**金山云内部设立了一个技术评议委员会，当某个具体问题可能发生偏离时，由技术委员会来决定是否允许偏离以及偏离的容忍度。**

除此之外，一个创业团队必须是目标驱动的，不能以部门利益为驱动。在公司里哪个部门有话语权并不重要，重要的是谁的认识更加清晰，谁能够给出最终的解决方案。

金山云在市场、营销、商务拓展等领域的投入相对较小，相关功能团队大概百余人。在行业内和社会上，现阶段金山云的声音相对较小。朱桦认为这种表现符合雷军一贯的低调态度。

美团云

从管理体系上看，美团云是美团的基础设施团队。**美团云对内向美团各业务单位提供服务，称为私有云；美团云将类似的功能对外部企业级用户开放，称为公有云**。2015年美团与大众点评战略合作成立新美大之后，美团云同时也给大众点评提供基础设施服务。

2013年的员工全部都是研发人员；2014年团队开始出现少数的非研发人员；2015年团队规模扩大了数倍，其中研发人员占70%，产品、市场、销售等等占

30%。美团云的负责人李爽也是在 2015 年加入美团云的。

李爽认为美团云这个团队相对年轻。在员工招聘上，级别较高的员工寻找业界资深人才，侧重考察候选人的专业技术以及在云计算领域的声望；普通员工则偏向于招聘应届毕业生，侧重考察候选人的学习能力。在过去三年中，未出现关键员工流失的情况，存在极少数普通员工的正常流动，并且有流出再度流入的情况发生。总体来说，美团云的人员流失率低于市场平均水平。

美团云的组织架构包括产品、研发、运维、市场、销售几个部分。在研发方面，又根据产品线分为主机、网络、存储、数据库几个不同的方向。工程师被分配到特定的产品线，工作内容通常聚焦在该领域，以打造业内领先的云产品为目标。李爽表示美团云最近也在考虑优化更新的研发组织架构。

美团云的产品设计，重点在于持续收集客户反馈，根据客户反馈不断对产品进行迭代。这些客户既包括内部客户也包括外部客户，内部客户的实际应用场景更多，对于需求的优先级会更高。美团云的外部客户，很多是与美团合作的 O2O 行业相关厂商，主要集中在传统服务行业。李爽认为传统服务行业的信息化程度是不高的。国内其它公有云服务提供商对传统服务行业的兴趣不大，这恰好给美团云带来一些机会。李爽认为云计算是一个 TO B 的生意，在执行层面和 TO C 类生意的运作存在很大差异。对于当时的美团云来说，**重要的不仅是产品如何设计，更是如何提升团队的执行力。**

七牛云

七牛云成立于 2011 年，同年获得经纬创投的百万美元级别 A 轮投资。2012 年和 2014 年又分别获得启明创投和宽带资本的 B 轮和 C 轮投资，但是具体的投资金额均未做披露。

2011 年员工总数 10 人左右，主要为技术人员；2012 年员工总数 40 人，技术人员比例为 3/4；2013 年员工总数 60 人，技术人员比例为 2/3；2014 年员工总数 120 人；技术人员比例略高于 1/2；2015 年员工总数 250 人，技术人员比例略高于 1/2。公司从 2012 年开始设立市场部门，从 2013 年开始设立销售部门。

在工程师招聘方面，七牛在早期面试中看重考察候选人的自制力以及对相关

领域的激情。在技术面试环节，不太关心候选人所掌握的具体技能，更侧重于考察候选人对其所使用的技术了解到什么层次。七牛通过一套严格的招聘流程来保证新员工的质量，特定级别的职位需要经过的面试次数以及面试者的级别要求都有具体的规定。对于面试者的面试方法不做具体的要求，因此不同的面试者可以有不同的面试风格。

七牛 CEO 许式伟认为任何企业——尤其是创业企业——都会面临招聘困难的问题。七牛早期的招聘策略是内推，也就是团队成员推荐自己的熟人前来面试。这样的方法招到的新员工一般都比较靠谱，但是团队扩张的速度比较慢。七牛曾经尝试过招聘应届毕业生，但是成功率接近于零。许式伟认为他们所接触到的大部分应届毕业生更希望找一份稳定的工作，很难找到有创业梦想的人。为了解决这个问题，七牛大量地采用了实习生。在面试实习生的时候，不关心候选人是否有云计算相关领域的经验，只关心候选人是否能够很快地接受新技术。实习生入职之后，再对其进行相关领域的技术培训，大部分实习生都能够迅速适应相应岗位的技术要求。因为实习的时候对公司有了充分了解也培养了感情，在七牛实习的实习生大概有 2/3 的概率会选择留下。除此之外，七牛也对 HR 团队进行扩容，每天有 3 到 4 个 HR 人员在大量简历中进行筛选，加大社招的力度。

在过去 5 年中，很少出现员工流失的情况。公司采取过一些主动淘汰的措施，但是这种情况也极少发生。用许式伟的话来说：“除非是特别不靠谱的人，基本上还是能留下的。”这种特别的情况，在信息技术行业相关公司相对缺乏的地区比较常见，但是在北京、上海、广州是比较少见的。流失员工的流向通常是规模更大的公司，2B 类和 2C 类的都有。公司正在建立主动流失人员的数据库，但是目前还处于雏形阶段。

公司在员工入职时与员工签订保密协议，没有对离职员工执行过同业竞争禁止协议。许式伟认为员工不能简单地看成是员工，希望人才在自然流动的过程当中成长并促进整个行业的进步。

在七牛内部，尽可能淡化产品经理这个概念。许式伟认为 CEO 本人必须是产品经理，具体子业务的负责人必须担负产品经理的职能，但是他要做的事情比对

产品经理的要求要高很多。虽然公司的研发人员占大多数，在公司发展早期的作用非常重要，但是许式伟也在淡化以研发为中心的观念，尤其是技术人员是整个公司最重要的角色这样的观念。许式伟认为任何公司都是做服务的，七牛虽然以技术立业但不以技术为中心。七牛要达到的目的不是技术导向，而是产品与服务导向。技术只是构建产品的手段，要通过产品来支撑公司给客户所提供的服务。

盛大云

根据公开的资料，盛大云成立于 2011 年，同年 7 月对外提供服务。实际上盛大云起源于盛大创新院，是盛大集团内部的技术支撑平台。2015 年，盛大云正式独立出来。目前盛大云研发团队的大部分人员在盛大集团的服务时间超过 5 年，算是一个非常稳定的团队。

盛大云的研发团队分成三个部分。核心团队负责 Xen、KVM、OpenStack 等虚拟化底层技术；业务中间件团队负责业务逻辑，侧重于前端与底层之间的通讯；前端团队面向客户，侧重于产品的人机交互和用户体验。除此之外，运维人员负责服务器、机房网络和操作系统层面的优化。

盛大云认为要在国内云计算市场里存活下来，需要从三个方面满足用户需求。**第一是用户对计算资源的基本需求；第二是用户对可靠性、稳定性、安全性的需求；第三是用户对易用性、高性能的需求。**研发团队之前侧重于第一和第二个需求，最近通过设立产品经理岗位来关注第三个需求。产品经理会对相关企业进行走访，哪怕是他们没有使用盛大云，也会咨询他们对正在使用的云计算服务的观点。在了解用户痛点的基础上，产品经理会对研发团队提出建议，设法对产品的易用性进行改进。

目前，研发团队在盛大云中的比重较大，产品路线图基本上由研发团队来决定。但是盛大云认为**用户需求会越来越重要，运营和市场部门的话语权也会越来越多。**对于盛大云来说，存在内部与外部两种不同类型的客户。内部客户泛指盛大集团内其它子公司；外部客户泛指盛大集团以外的其它客户，外部客户通过盛大云的运营团队给研发团队提出需求。外部客户的获取，主要是通过运营团队的线上和线下推广，包括与科技园区、校友会、创业机构等等进行合作。盛大云正

在从技术导向往市场、产品、技术共同导向的方向去发展。

腾讯云

根据公开的资料，腾讯云在 2013 年下半年开始公测，在 2014 年成立于腾讯云计算公司，是一家相对年轻的云计算公司。腾讯云副总裁王慧星则认为腾讯集团是包括多个不同公司的一个整体，2014 年在北京注册的云计算公司只是众多公司其中的一个。除了北京的云计算公司之外，腾讯集团内部多个公司里都有从事云计算的人员。所以，从时间上来讲不能说腾讯是从 2014 年才开始做云计算的。在云计算这个行业，像网宿这些都算是早期的公司，腾讯比阿里要稍晚一些。腾讯云的历史，最早要追溯到腾讯开放平台。后来腾讯集团意识到云计算这个业务需要在财务上变成一个实体，所以单独成立了北京的云计算公司。

腾讯的云计算团队成长非常快，从业人员并不局限于北京的云计算公司。其中，研发人员的比例远远超过 50%。在过去几年里，整个云计算团队的年成长速度均超过 100%，处于高速成长期。腾讯云设有独立的团队负责市场宣传、公共关系、品牌建设、对外交流等等事物。王慧星认为腾讯云对外提供的服务是基于腾讯集团从成立至今的技术积累之上的。以前腾讯集团的基础设施团队只对腾讯集团内部提供服务，腾讯云是将这些现成的技术能力加上产品化的服务等级协议（SLA）后对外提供服务。腾讯云以及腾讯集团内其它公司的的基础设施团队是一个整体，不能简单地称为人员复用。

在招聘过程中，腾讯云希望看到候选人对技术有深入理解和，具备行业视野，认同腾讯价值观。王慧星认为他并没有觉得腾讯云存在招聘上的困难，不存在人才缺乏对腾讯云的发展构成障碍的情况。**腾讯集团在发展过程当中积累的很多人才，也对外输出了很多人才。**人才的流动对中国的发展是有用的。在过去几年中，员工流出的主要原因是薪酬竞争力和管理问题。王慧星认为，和国内友商相比较腾讯云的员工流失率是相当低的。HR 部门会与离职员工签订保密协议和同业竞争禁止协议。但是王慧星也认为整个互联网行业里各个公司对相关问题都持开放的态度，并且保密协议和同业竞争禁止协议实际上起不到阻止离职员工泄密或者加入竞争对手的作用。

腾讯云设有产品经理职位。王慧星认为云计算产品也是产品，其使用对象是研发人员。只要是产品，就一定有产品经理，和其它产品没有什么差别。王慧星进一步指出产品团队和研发团队之间必须有很好的合作。以其在腾讯的经验，很难理解产品团队和研发团队之间会存在沟通问题。

UCloud

UCloud 成立于 2012 年，2013 年 11 月获得 A 轮融资 1000 万美元，2014 年 6 月获得 B 轮投资 5000 万美元，2015 年 4 月获得 C 轮融资近 1 亿美元。

2012 年员工总数 20 人，2013 年员工总数 80 人，2014 年员工总数 280 人，2015 年员工总数近 500 人，计划 2016 年增至 700 人以上。研发团队占公司员工总数的 50% 左右，在过去几年中仅有小幅度的浮动。

UCloud 在招聘工程师过程中侧重考察三个方面的内容：一是对云计算有认同感，二是具备发展潜质，三是具有研发经验。换句话说，就是先看兴趣爱好，再看资质，最后看经验。UCloud 认为整个云计算行业都一直存在人才缺乏的情况，而 UCloud 的主要招聘途径是行业内的人脉推荐和猎头挖掘。当然，UCloud 也使用正常的招聘渠道，筛选简历以及校招等，并尝试通过与企业合作为自己培养潜在的候选人。在过去三年里存在一些人员流失的情况，一方面是竞争对手挖角，另一方面是自己主动淘汰。UCloud 不考虑和离职员工签订保密协议或者同业竞争禁止协议，因为 CEO 季昕华认为：“我们自己也做过员工。如果公司好，员工自然会留下来，如果不好，员工自然会走。我们的心态非常开放，有离开的员工后来又再回到我们团队。”

UCloud 设有产品经理职位。产品经理的职责主要包括三个方面：**一是搜集用户需求，二是了解整个行业的发展趋势和变化，三是把用户需求和行业变化汇总成产品特性，和研发人员一起进行研发并把产品推向市场。**UCloud 的产品经理均具有研发或者运维背景，这个要求是因为 UCloud 的客户大部分是研发人员（70%）或者运维人员（30%）。季昕华认为 UCloud 并不存在研发团队掌握话语权的情况。作为一家公司，UCloud 的使命是为用户创造价值，研发团队在话语权多寡在于其是否为用户创造价值。

又拍云

又拍云源自 2005 年成立的又拍网。又拍网是一个类 Flickr 的图片分享网站。为了满足网站本身的功能需求，研发团队开发了一个分布式文件系统允许用户之间分享文件，此后又整合了内容分发网络（CDN）的功能。公司注意到当时国内国外均缺少类似的产品和服务，决定将其作为一个发展方向。到了 2010 年，又拍网的图片分享功能已经非常稳定，也达到了很好的性能。很多做电子商务的公司都直接使用又拍网的图片服务。公司看到了这个机会，顺势推出了“存储+分享+分发”的一体化产品，后来又逐渐加入编码转码、屏幕适配等等多种特性。从产品发展的角度来看，这个产品是根据又拍网的业务需求和客户需求逐步演进出来的。在公司层面，并没有专门规划要去做一个云产品，甚至没有意识到这个产品可以被归类为云产品。因此，又拍网认为自己“在不知不觉当中成了国内第一家云计算公司”。

又拍网是原本是一个 2C 类的公司，其业务是按照 2C 的模式进行运作的。新推出的“存储+分享+分发”产品是一款 2B 的产品，其业务是按照 2B 的模式进行运作的。一家公司同时运营 2C 和 2B 两类不同的产品和业务，在管理上遇到了一些问题。2014 年，2B 部分业务被剥离出来，成立了又拍云，进行独立运营。同年，又拍云引入鼎辉创投和松树资本，获得数千万元人民币的 A 轮融资，并完成管理层收购（MBO）。

在 2010 年到 2014 年之间，由于重视程度的问题，“存储+分享+分发”的业务发展得不够快。相应团队需要自己去拓展业务，有了收入再反哺研发，因此发展相对缓慢。**2014 年成立又拍云并引进 A 轮融资后，业务发展相对较快。**2014 年初，整个团队不到 50 人；2015 年底，整个团队共有 130 人。又拍云认为自己是一家“技术驱动的产品型公司”，技术团队占整个团队的 60% 以上。技术团队包括底层研发（包括核心研发、操作系统和网络）、前端研发、产品运维、技术支持四个部分。公司设有独立的产品部门，主要功能是衔接前端、客户、销售、市场，把来自客户的需求转化成产品的需求。产品部门负责采集客户需求并提出产品功能需求，在产品设计环节占主导地位；研发部门基于产品部门所提出

的产品需求决定产品的架构和技术选型，在产品实现环节占主导地位。市场部门占整个团队的 10%，主要负责公共关系、社区运营、线下或等等内容。在人员配置方面，公司认为在营销方面还需要加强。

又拍云的总部和研发团队设在杭州。公司认为，与技术人员相对密集的北京和上海相比较，杭州的技术人员相对较少，相关领域的人才缺乏在一定程度上妨碍了公司的发展。在工程师招聘方面，公司注重招聘渠道的开拓，尤其注重在技术社区和市场活动中穿插的招聘缓解，基本上不依赖于专业的招聘网站。公司鼓励技术团队在技术社区进行分享，树立技术团队在社区中的良好声誉，也让潜在的应聘者了解公司的环境氛围和成长机会。技术团队的主要负责人看到有潜力的工程师时，也会主动去邀请加盟。公司有很多优秀的工程师，都是通过技术社区这个渠道招聘进来的。在对候选人进行考察时，重点考核候选人的学习能力和接受能力。候选人现有的技术水平固然重要，但是公司更注重候选人的求知欲望和自主学习的能力。

又拍云的员工流失率相对较低。2014 年前 8 个月没有出现员工流失的情况，2014 年 9 月到 2015 年 12 月之间有个位数的员工离职。研发团队员工离职的主要理由是回乡（江苏、福建）发展，并且回乡后基本上都还从事技术相关工作。公司会与离职员工签订保密协议和同业竞争禁止协议，但是没有真正执行过。公司举办年会活动时，也会邀请离职的员工回来参加。

2015年中国公有云服务发展报告

产品研发篇

在这个部分，我们以产品研发为主线，探讨不同公有云服务提供商的发展理念、研发重点、决策过程。我们关心的问题包括云计算产品与特性，重点发展的业务方向，产品与特性的优先级，过去几年中取得的关键性进展，等等。

金山云

金山云的朱桦以 AWS 为例阐述了他对云计算产品的理解。**朱桦认为 EC2 代表计算的虚拟化，是 IaaS 层面的核心；EBS 和 S3 代表存储虚拟化；VPC 代表网络的虚拟化。**云计算产品多种多样，有偏向计算的，有偏向存储的，有偏向网络的。每一个领域都有很专门的技术，需要很大的投入。但是要解决客户的问题，则需要将计算虚拟化、存储虚拟化和网络虚拟化有机地整合起来。AWS 成功地将专门的技术包装成普通技术人员能够理解的概念，又进一步包装出多种多样面向场景和面向用户的云计算产品，走出了一条将技术研发和产品包装相结合的道路。在这一点上，国内云计算公司做得还不够，阿里云做得稍好一点。和 AWS



相比较，国内很多云计算公司在技术上的投入不够，往往通过创造新概念来标榜自己的产品与众不同。金山云追求的是在技术上的差异化，而不是产品或者概念上的差异化。金山云认为 AWS 是云计算领域的先行者，在技术和产品方面需要尊重和认同 AWS 的理念。与 AWS 相比较，国内云计算服务提供商的用户和场景都不够大。在方向不够明确的时候，需要虚心地接受 AWS 的经验和学习 AWS 的教训，而不是盲目地追求所谓的微创新。金山云坦率地将 AWS 和阿里云视为行业标杆，并将 AWS 和阿里云作为赶超的目标。

金山云的研发重点包括类似于 EC2 的云主机服务、类似于 S3 的对象存储服务、类似于 VPC 的虚拟网络服务。金山云的云主机服务满足了所有游戏类客户的苛刻要求，在行业内是水平一流的云服务商。在对象存储领域，金山云的规模国内最大，比其他竞争对手要大一个数量级。金山云认为 VPC 是一个技术难点，其 VPC 产品已经内测了很长时间，争取在 2016 年内正式推出（作者注：VPC 产品已经于 2016 年推出）。国内的云计算公司当中，目前能够同时提供云主机、云存储、VPC 这三大类别产品的可能只有阿里云。云主机产品和云存储产品多家云服务提供商都有，金山云认为自己的产品性能更好些。云计算行业要求巨大而艰辛的技术资源投入。**过去金山云看到市场对云计算产品有多种多样的需求，也做了很多产品，由于缺乏聚焦走了很多弯路。**目前金山云有六百余人，这样的团

队规模尚不足以支撑金山云在多个产品线上全面出击。因此，金山云要在有限的几个重点方向上投入研发力量，抵制某些特别有诱惑力的需求。

金山云认为其主要产品具备如下特性：

1. 在云主机领域，在其他友商尚在钻研虚拟化技术本身时，金山云较早地注意到了云主机产品对 I/O 性能的要求。当时市面上所有的云主机产品在 I/O 性能上都满足不了游戏行业的需求。金山云果断地将游戏行业作为云主机产品的重点应用场景，云主机的 I/O 性能遥遥领先。后来其他友商也开始注意 I/O 性能了，部分友商的云主机 I/O 性能已经赶上了金山云的水平。在网络性能方面，金山云的云主机延迟很低，可以支持很高的并发连接，适用于缓存等特别频繁的小包通讯类业务。金山云还采取了大量细颗粒度的策略来实现云主机之间的资源隔离。金山云认为网络和磁盘 I/O 的隔离是比较容易实现的，比较难实现的是运行在同一物理机上的云主机之间的 CPU 隔离。很多友商对 CPU 进行超售，并且将 CPU 超售能力视为公司的核心竞争力。这样的做法无法保证云主机的服务等级协议（SLA），最终会损害客户的利益。金山云认为在这个领域做得最好的是 AWS。如果对云主机的性能进行批量测试的话，AWS 的云主机性能曲线很平滑，其他友商的云主机性能曲线存在严重的跳跃。这就说明在资源隔离方面大家做得还不够好。在云主机领域，金山云要不断地提升自身在计算虚拟化、存储虚拟化、网络虚拟化方面的核心能力，通过不断迭代提升性能、降低成本、增强隔离。

2. 在云存储（类似于 S3 的对象存储）领域，金山云将可用性和成本作为核心指标。云存储的规模越大，对可用性的挑战也越大。金山云现有的云存储规模接近 200PB（作者注：2016 年 8 月的规模为 300 PB）。金山云认为，云存储每产生 10 倍的规模增长都会对系统的架构提出新的要求。

3.VPC 是一个难点，跨数据中心的 VPC 实现甚至涉及到数据中心的规划，实现起来非常复杂。金山云内测中的 VPC 在功能上可以对标 AWS 的 VPC，在性能上稍差。

在决定产品与特性的优先级时，金山云有两个大的原则。第一个原则是坚持聚焦，避免把资源分散到多个项目；第二个原则是把握方向，资源投入要对重点

研发方向有促进作用。所有的需求都有其存在的合理性，但是并不意味着所有的需求都要由技术团队来支撑，要做减法。在做减法是，要反复思考金山云最核心的能力与最核心的业务是什么。只有与核心业务相符合的需求才会被处理，和核心业务离得比较远的需求通常会被放弃。

金山云认为**云计算卖的不是产品而是服务能力**，包括商务能力和技术支持能力，**其核心在于服务提供商要贴近客户的业务**，要明白厂商对客户所承担的责任。对于厂商来说，好像是在卖一个产品；对于客户来说，他们最重要的资产都放在厂商这里了。一次事故对于厂商来说不过是一次事故，对于客户来说可能就是业务停顿甚至是数据丢失。如果在关键时间发生业务停顿或者是关键数据丢失，对于客户来说可能就是灭顶之灾。金山云的团队在不断变大变强，但是依然保持对客户的敬畏，这是金山云最核心的能力。金山云有信心在拥有几十万到上百万客户的时候依然能够保持这样的能力。

人们常说技术是第一生产力，对于云计算来说要体现在服务等级协议（SLA）能不能被严格的执行，厂商所承诺的 N 个 9 能不能在理论和实践上都得到严格的保障。在服务出现问题的时候，能不能在 SLA 规定的时间内被解决掉。这些问题，在理想的情况下应该是依靠技术手段来解决的。金山云相信技术产品的创新能力能够带来行业的巨变，更理解这种巨变不是每时每刻都在发生的，需要一个从量变到质变的过程。**将来哪个厂商能够走得更远，取决于哪个厂商能够坚持，能够聚焦。**

美团云

美团云的发展始终建立在对稳定性的极致要求上。在很长的一段时间内，美团云都只有云主机和云存储两个产品，主要的考量就是要把最基础、用户使用最多的产品先做扎实。其它方面的需求——包括网络和 API——都作为特性看待。2015 年，美团云打磨了更丰富的产品线，能够满足企业级用户大部分的需求。

美团云认为国内和国外用户使用云计算的场景不太相同。**目前大部分国内用户只是把云计算当做资源来使用，把云计算当做服务来使用的情况还是比较少的。**譬如说云主机，很多用户只是把云主机当作虚拟主机——也就是传统物理服务器

的替代品——来使用，没有考虑到弹性扩展等等云服务特性。这些思维定式在国内用户使用云主机、云数据库、云缓存等产品时都有明显的体现。

美团云的云主机提供了丰富的 API，但是使用 API 的用户比较少。随着时间的推移，云计算的理念得到普及了，可能会有更多的用户使用。**李爽认为美团云的云主机的特性还包括高性能与高稳定性，但是这两个特性比较容易被用户所忽略。**在云主机的基础上，美团云也为用户提供了预先配置好的 Hadoop 数据处理平台，这样用户就不需要自己搭建 Hadoop 集群了。在网络方面，美团云将外网 IP 设计为浮动 IP，可以随时切换外网 IP 与云主机之间的映射。在网络带宽方面，网关可以任意调整带宽，峰值的时候可以调高，也可以根据需求随时下调。目前美团云的虚拟网络是基于 VxLAN 的，比在主机上做 Overlay 网络的性能更好。VPC 产品正在研发中，很快将对外正式发布。

美团云正在研发基于 Docker 的产品。从 2015 年开始，美团内部希望用 Docker 来部署一些服务。在内部使用一段时间之后，美团云就会把内部的 Docker 服务提供给公有云的用户。这也是美团云一贯的政策 **内部业务环境先试，再推出给外部用户。**

美团云根据基础产品的完整度来决定产品与特性的优先级。一方面，参考业内的情况制定一个标准特性列表；另一方面，根据大客户的反馈对标准特性列表上的项目进行排期。譬如说对象存储这个产品，开始只做了 PUT 和 GET 功能。后来根据客户的反馈增加了转码处理等功能。

作为美团的基础设施部门，美团云的研发动力来自内部需求。美团云的产品，都是先在美团内部得到广泛应用，然后才对外部客户提供服务。因此，美团云自认为其产品在稳定性上具有明显的优势。在另一方面，美团内部没有需求但是外部客户需要的产品美团云也不会做。因此，从产品丰富度上来看较欠缺。

和国内友商进行比较，美团云认为各个云服务提供商所提供的产品和服务类型都很接近。某些云的产品或者特性大小无数百个，数量上可能比较多，但是只有少部分是常用的。美团云的短期目标，就是把用户常用的产品或者特性做好。从技术角度来看，国内云计算厂商处于同一水平，差异不大。

七牛云

七牛云在很长一段时间里只有一个产品。从 2015 年开始，才有了子产品的概念。从单一产品到多个子产品的决策是很自然的。一个企业在刚刚起步的时候不能一心多用，得单点切入。七牛花了 3 年多的时间来打磨一个产品，并且取得了很好的成绩。但是云计算很复杂，单一产品不能满足用户的闭环需求。基于单一产品，七牛只能是用户产品的一部分，而不是全部。七牛希望满足某类用户的全部需求，就必须从单一产品往多个子产品转变。

七牛的子产品有四个，对象存储、融合 CDN、数据处理、直播云服务。前三个子产品，可以认为是一个一体化的 PaaS 服务，是相互依存的关系。从外部看来，这是一个整合在一起的产品；从内部看来，这三个子产品由不同的团队负责，各自有不同的 KPI。直播云服务则是一个相对独立的产品。在许式伟看来，这四个子产品是一种互补的关系。和大部分人对存储的定义有所不同，许式伟认为存储需要满足用户远程沟通的要求。譬如说一个用户通过云存储向另一个用户展示了一张图片，实际上是完成了一次沟通的过程。因此，**云存储是用户和用户之间交互的媒介。传统的云存储服务提供的仅仅是单向交互的功能。**某个用户浏览了一张图片，图片的拥有者并不知道是哪个用户浏览了这张图片。这样的单向服务，没有办法满足用户的“互动性”交互要求。譬如说在线教育这个应用场景，教师给学生授课，学生向教师提问，就是一个需要“互动性”交互的场景。这种“互动性”交互和电话沟通类似，是双向的。七牛的直播产品解决了用户的“互动性”交互问题。七牛的直播产品和大部分人所理解的直播是不一样的。七牛的直播产品不是一个非技术人员可以使用的独立终端产品，而是提供给直播网站构建直播服务的后台支撑产品。

在七牛，由子业务负责人（兼具产品经理的职能）决定子产品中各个特性的优先级。简单地说，就是**子业务负责人说了算，是一个由上往下的过程。**市场部门的员工从客户那里采集需求信息，入库管理，通过七牛内部对相关产品的理解来判断某项需求的普适程度，最终由子产品负责人做出决策。

很多人都搞不清楚七牛云和 CDN 服务提供商的区别。CDN 这个市场从互联

网兴起之后就在发展，现在已经非常成熟。相较于市面上的 CDN 服务，七牛云提供融合 CDN 服务。融合 CDN 是在传统 CDN 基础上实现的对数据网络加速进一步优化的融合管理服务。除了服务于音视频点播，文件、应用与 Web 加速，以及各类增值场景外，七牛融合 CDN 还通过全方位的 CDN 质量监控，以及智能易用的加速节点调度等功能，保障用户服务的连续性，提供稳定快速的网络访问质量。

七牛云与 AWS、阿里云属于同类产品。在许式伟看来，AWS 的 S3 和阿里云的 OSS 的功能局限于 PUT 请求和 GET 请求。**七牛把存储的概念给扩展开来，变成一个存储的 PaaS。这样的扩展能够帮助用户更好地使用所存储的数据，为用户提供围绕存储数据的增值服务。**譬如说七牛平台上已经有了 20 到 30 个用于数据处理的 API。这些数据处理 API 是对外开放的，任何用户都可以在七牛平台上提供数据处理的产品，去满足其客户的数据处理需求。七牛是唯一一家把存储做成 PaaS 的厂商。七牛云、CDN 服务提供商、以及其他厂商，可以说是分别位于同一个产业链上的不同位置。七牛希望在这个链条上进行资源整合，把整个链条穿起来，降低用户构建产品的成本。

盛大云

盛大云认为云计算技术有几个发展阶段。早期大家对市场还不够了解，中国客户和国外客户也存在差别。国内客户大都是从中小市场转化过来的，这些早期客户对 AWS 所提出的 EC2 或者 EBS 等新概念并不了解，只是认定了云计算就是虚拟机，云主机就是传统物理主机的替代品。因此，早期的云计算产品是以 VPS 的形式出现的。随着市场逐渐成熟，客户的需求也在慢慢扩展，自然而然地围绕云主机发展出一系列产品。这些产品可以是独立的，但是又不一定必须独立存在。譬如说云硬盘是盛大云 2011 年发明的概念。当时盛大云发现客户对虚拟机数据盘的性能要求比较高，自然而然地就想到了云硬盘这个产品。但是云硬盘是不能独立于云主机使用的，所以不能算是一个独立的产品。为了解决存储与虚拟机分离的问题，又想到了云存储这个产品。客户把云存储当成一个存储池，通过 PUT 和 GET 这样的 API 接口来读写对象。等到客户的业务增长了，单台云主机无法满

足客户业务上的需求了，负载均衡产品就自然而然地出现了。客户所服务的用户在地理范围上的分布变广了，自然而然地产生了对内容分发的需求，盛大云的 CDN 产品也就随之出现了。

所以说，**盛大云的发展与用户需求的演进基本上是同步的**。在国内，盛大云是最早向 AWS 学习的云计算服务提供商。在盛大发展的早期，在 AWS 的 EC2、EBS、S3 等已有产品的基础上，盛大云又开发了 CDN、视频云、云安全等等新产品。当国内其它云计算服务提供商还没有起步的时候，盛大云就已经推出了这么多产品。但是盛大云随之发现了两个问题。第一个问题是国内客户对云主机的接受程度并不高，相比之下现在的情况好多了。第二个问题是运营商之间的互联互通。在这种情况下，客户用量最多的还是云主机，对其他产品的需求并没有那么大。因此，盛大云在很长一段时间里都专注于云主机业务。除了云主机，客户用量最大的产品是云存储、云硬盘、CDN。因此，盛大云也把云主机、云存储、云硬盘、CDN 作为我们的基础性产品看待。

在产品特性上，盛大云注意到市场上对云主机的要求也是分阶段发展的。早期的时候，各个厂商在技术上并没有多少距离，因为客户对磁盘容量和读写能力的要求并不高。随着市场变得成熟，用户对云主机稳定性和性能的要求慢慢提高了。客户对磁盘 IOPS 的要求提高了，就有厂商推出基于 SSD 硬盘的云主机。客户对云主机覆盖的网络范围提高了，就有厂商推出基于 BGP 带宽的云主机。客户对云主机的安全要求提高了，就有厂商推出具备高防能力的云主机。国内的云计算圈子有一个特点，就是厂商之间的沙龙和分享很多。厂商和厂商之间都在有意无意的产生很多沟通。虽然友商不会说底层的实现方法，但是他会讲他们是怎样满足用户需求的。在微信和主流媒体上，时不时地会发布一些新技术。一个新技术被发布出来后，其它厂商都会很快跟上。在技术水平上拉不开，造成厂商之间产品同质化的局面。在这种情况下，盛大云选择在某些特性上突出我们和其他厂商的区别。例如在云主机的稳定性和可靠性方面，盛大云会多下一些功夫。

盛大云在 2015 年在产品大方向上做了一些优化，在技术实现上做了一些改进。这些改进体现在产品线更加丰富，机房做了异地灾备，服务可靠性得到提升。

这些改进从用户界面上可能反映不出来，但是在用户的使用过程当中会慢慢地感受到。譬如做镜像的速度和做备份的速度，这些只能在使用过程中感受到。2016年盛大云会对 Web 控制台做一些大的改动，包括对资源进行操作的流程的改动。在对产品进行改进的时候，我们既会做横向的比较，也会做纵向的比较。我们现在与用户进行沟通的过程当中，用户很难表达出他真正的需求。因此，我们需要对用户所表达的需求进行梳理和挖掘，得出正确的用户需求再和用户进行确认。为此我们设立了一个产品经理的职位，负责与用户沟通和改进我们的评估。

盛大云将缺陷（bug）分为 ABC 三个等级，A 级是影响主要功能的缺陷，B 级是影响用户直接使用的缺陷，C 级是影响用户体验的缺陷。盛大云的运维团队 7x24 小时保持手机开机，在有需要的情况下可以随时随地做出响应。在产品与特性的优先级排序方面，盛大云优先考虑能够带来性能提高的特性，或者是来自大客户的迫切需求。除此之外，研发团队也会主动提出新的特性。研发团队看到友商的新功能和新思路也会主动跟用户沟通。用户未来可能会使用的功能，也会排上开发日程。最终由研发经理和产品经理共同对产品和特性进行优先级排序。

盛大集团内部也有 IDC 业务，提供服务器和带宽租赁。云计算和 IDC 业务这两块业务在某些领域是充分竞争的，但是在某些领域又是相互补充的。传统 IDC 属于资源粗放型业务，采购周期上架周期比较长。传统 IDC 业务需要多为用户考虑，譬如用户需要某个量级的资源，但是我们会增加资源配备来满足用户的突发性需求。就是因为传统 IDC 有这个弊端，云计算的按需使用使用按量横向扩展等等特点就有了明显的优势。用户不再需要关心采购周期和上架周期。需要使用的时候申请所需的计算资源就可以了。**盛大云认为 2016 年的一个大的方向是：用户将敏感数据放在自己这里，需要横向扩展的部分或者承担业务压力的部分放到云上。**从这个角度来看，IDC 业务和云计算业务是互补的。用一面扇子来打个比方，扇子打开的时候，扇子的上沿是公有云，横向扩展会很快；下面是私有云，横向扩展相对慢，譬如业务前端和数据库，前端扩充五倍，后面可能才扩充一倍。这样的架构能够保证数据的安全性，可控，备份方便，对成本增加不大。盛大云注意到很多用户愿意付少量的成本来保证数据的私有性。

在公司层面，盛大云的业务分成三块：公有云平台，定制云（私有云 + 混合云业务，可以根据客户需求进行定制），综合的云方案（针对政府部门，园区，从园区的层面提供综合的解决方案）。前面提到的私有云属于定制云方面的业务。我们的定制云和公有云在同一个机房，解决了定制云和公有云之间的网络互联互通问题。定制云和公有云两个业务所使用的物理服务器，是由同一个运维团队进行管理的。定制云在底层也有横向扩展的需求，涉及到物理服务器的变动（上线、下线、维修）。盛大云准备有备机提供给用户。另外，盛大云还可以在用户当地机房提供定制云服务，可以通过专线打通用户当地机房和盛大云的机房。

盛大云认为国内本土的云计算服务提供商比 AWS 和微软有竞争优势。这是因为国内厂商对国内互联网用户、国家政策、国内网络状况都有特别的理解。和国内友商相比，盛大云的优势有两点。第一是盛大云起步较早，通过多年的运营积累了大量的运维经验。盛大云多年为盛大集团服务，处理了很多突发状况。第二是盛大云对用户需求有很深的理解，打造的产品更加贴合国内用户的需要。

腾讯云

在腾讯云内部，并没有特意去对产品和特性进行界定。王慧星认为，**客观上人的认知存在多样性，由于评价维度不同，便会出现不同的理解。**从公有云的发展来看，公有云已经远远超越了 IaaS 这个层面。腾讯云认为 AWS 最基础的产品并不一定就是 EC2、S3 和 VPC。对于腾讯云来说，比较重要的包括设施层的 IaaS、平台层的 PaaS，还有生态层的合作伙伴。设施层的 IaaS 是一大类产品，整合了计算、网络 and 存储；平台层的 PaaS 也是一类产品，包括腾讯的深度学习能力和音频视频处理能力；生态层的合作伙伴则是另外一个大的类别。但是腾讯云内部从来都没有讨论过哪几个是最基础的产品。譬如说我们对安全很重视，在研发方面有很大的投入，那么安全算不算最基础呢？网络攻击是所有用户的痛点。我们在这方面做了很多事情。腾讯云的目标是将底层的基础设施以一种更上层的方式提供给客户。计算、存储、网络会透明化，而不是以裸机的方式去提供给客户。

腾讯云认为设施层面的 IaaS 在国内和国外基本上大同小异。譬如说 CDN 服务，

各家都说自己的性能领先，谁都互相说服不了。2015 年 9 月，腾讯云发布了基于 MySQL 的数据库服务，QPI 达到了每秒 3.7 万，在国内其竞争对手的 QPI 只有每秒 2.5 万左右。等到 2016 年新产品发布的时候，性能还会再增长一倍。在云主机的性能优化方面，腾讯云的块存储的 IOPS 达到了非常高的水平。但是，腾讯云更多地考虑在 PaaS 和 SaaS 层面体现出与友商的差异。譬如说腾讯的音频视频服务，又譬如说与 Oracle 合作联合发布 Oracle 数据库服务。腾讯的音频视频服务与苹果的 Facetime 类似，但是腾讯在跨国的互联互通方面做得更好。在中国的三大运营商之间，还存在互联互通的问题，因此在基础设施层面还有很大的竞争空间。腾讯的网络互联互通能力是非常强大的。腾讯在中国有几个大的节点，彼此之间互联互通能够很好地覆盖全国甚至是全球。又譬如说人脸识别，腾讯在这个领域有很深的积累。腾讯云的客户可以在腾讯的人脸识别功能之上构建上层应用。譬如说银行的免见面开户业户，譬如说某些特殊行业的活体检测业务。腾讯在 QQ、QZone 等产品上积累了很多独有的能力，它们都是腾讯云独有的特性，是其它 IaaS 服务提供商所无法提供的。

在产品特性的优先级方面，腾讯云聚焦在 IaaS 和 PaaS 层面。有一些重点投资的产品，其优先级是最高的。其它的产品，则要根据具体的场景进行具体的分析。优先级排序是一个 trade off 的结果。腾讯云首先关注的是性能优化，其次是腾讯内部特有的能力。譬如说人工智能与深度学习能力，如果能够对外开放提供服务，也会加进来。这里的关键是产品和特性要对用户有价值，也要符合腾讯的发展规划。在腾讯云内部，每个产品都有一个对应的团队，腾讯云把决策权下放给执行的团队。在一个具体的问题上，其实并没有一个明确的决策流程。不管是 2C 类产品还是 2B 类产品，产品经理试错都是需要的。用户提出一个需求时，他真正的需求也未必就是他提出的那个需求。要考虑在现有的产品里面是不是有现成的东西可以达成他的目标。如果一个团队总是在纠结需求的优先级，这个团队可能存在一些结构性的问题。

与 AWS 和 Windows Azure 等外来的公有云服务提供商相比较，**腾讯云认为大家在基础的 IaaS 层面基本上大同小异，在计算、存储、网络方面的能力并没有**

根本性的差别。腾讯过去十多年的积累是非常了不得的，在产品磨练过程中有一些独到之处。在腾讯内部，以前是不同的产品团队单独支撑自己的产品。现在腾讯内部有一些上万台机器的业务，是全部跑在腾讯云上面的。腾讯云的产品可以把微信和 QQ 支撑好，是非常了不得的一件事情。腾讯云希望把支撑微信和 QQ 的这种能力提供给外部客户，给他们帮助和支持。

在 API 接口方面，腾讯云给腾讯内部提供的 API 接口比给外部客户提供的 API 接口要更加丰富。对外提供服务的产品和 API 接口，都是到已经非常成熟了才发布出来。王慧星认为 **AWS 声称“亚马逊的电子商务平台是构建在 AWS 之上的”不过是市场宣传上的一种说法**。具体实现上必定有多个数据中心，有的数据中心是对内的，有的数据中心是对外的，涉及到的 API 接口不可能是完全一致的。AWS 给 Netflix 和 Apple 所提供的支持必然是经过深度定制化的。譬如说 S3 服务的 SLA，苹果的 iCloud 可能会要求 SLA 要达到 12 个 9，但是对外的 SLA 只有 11 个 9。同一套系统给不同的用户提供服务，这个从逻辑上来说并不成立。大系统给大公司用，小系统给小公司用，才是符合逻辑的做法。

腾讯公司高层希望腾讯应该是同一套系统同时对内对外提供服务。这个愿望在逻辑上没有问题，但是在具体实现上是有问题的。

UCloud

对于 UCloud 来说，产品与特性之间的界定是比较简单的。比如说云主机 UHost 是一个产品，负载均衡 ULB 是另一个产品，但是 ULB 的某个功能点是一个特性。产品是一个能够单独完成某个任务的东西，而特性仅仅是某一个小功能而已，围绕产品的功能点来做的。举个例子来说，云数据库 UDB 是我们的产品，但是 UDB 的双机热备是一个特性。

UCloud 的产品也分为计算、存储、网络三个大类。在计算方面，UCloud 既有云主机也有物理机，还提供 GPU 云主机。UCloud 的云主机具有很好的稳定性与性能，可以在物理机之间进行动态迁移，可以实现弹性扩展。底层物理机可以在不重启操作系统的情况下给操作系统内核打补丁，称之为内核热补丁技术，这项技术在整个云计算领域都是领先的。UCloud 的云存储具有很高的可靠性，使

用非常灵活。UCloud 的数据方舟特性，允许用户将数据恢复到 72 小时以内的任意零点，24 小时以内的任意整点，12 小时以内的任意一秒。UCloud 于 2015 年推出了基于 PCI-E 的存储，可以达到非常高的性能。UCloud 的普通存储也使用了一些有助于提高性能的专利技术。在网络方面，UCloud 在 2012 年实现了软件定义网络，2013 年实现了同城多机房布局，2015 年实现了跨城容灾备份。

UCloud 在 2015 年集中精力做了三件事情。第一件是推出为金融业务服务的金融云解决方案。金融云围绕金融行业的特殊需求，从 IDC 的物理环境、到网络的同城容灾、到防 DDOS 攻击，都达到了金融行业的要求。第二件是基于 Hadoop 的数据分析产品 UHadoop 集群，为用户提供大数据分析的平台。第三件是直播云服务。UCloud 的很多用户都有实现视频直播、点播的需求，直播云服务能够让更多小团队和个人快速创业。

UCloud 内部有一个文化，叫做“用户的需求就是我们的下一个产品”。

UCloud 会经常性地跟客户进行沟通，了解什么产品用户有需求。比如说混合云产品，UCloud 自研的 SDN 网络服务允许用户将业务在一个 VPC（虚拟私有网络）中无缝地迁移、整合、连接，也允许用户将公有云上的资源和私有云中的资源组成一个互通的资源池，这些互联方式以物理网络的不同属性可以分为内网互联、专线互联、以及公网互联。从用户的角度看，他们的计算节点（不论是容器还是虚拟机还是物理机）都是在一个统一的虚拟网络内是可达的。而在底层，UCloud 则是通过各类 NFV 化的网关（比如物理云主机有物理云主机的网关，托管云主机有托管云主机的网关等等）为这些不同的节点集群之间提供了高性能互联的机制。

和传统的数据中心服务器租赁业务相比，UCloud 最大的特点在于 API 化。

有了 API 后，任何操作都可以在线地通过 API 实现，而不是人工实现。这样才能够真正实现云计算的弹性。小型的客户，基本上不需要使用 API。规模大一些的客户（譬如说几十台）才有 API 的需求（需要写程序来做管理）。UCloud 的客户，有 3 台以上服务器或者云主机的基本上都会使用 API。

UCloud 的定位，是做国内最大的中立云服务提供商。UCloud 和用户之间没有竞争，这一点是其他公司所不具备的。他们既要做平台，又要做内容，和

客户产生了竞争关系。UCloud 关注的客户很聚焦，只做行业客户，不做个人客户。在行业用户中，UCloud 争取的是创业型互联网公司以及向互联网转型的传统企业。UCloud 对这些客户有很深的研究，才使得他们愿意把身家性命都放在 UCloud 这里。平台好、稳定性高、性能高、服务周到是他们的需求。但是 UCloud 的品牌还不如一些云计算巨头那么知名，在产品的种类和数量也还有上升的空间。比如，阿里云进入行业的时间更长，研发人员相对更多。在获取资源的能力和价格方面，阿里云的采购能力也比 UCloud 更占优势。

AWS 是 UCloud 的标杆。和 AWS 类似，UCloud 也是先从行业客户做起，UCloud 觉得能长大的客户才是好客户。UCloud 要先做企业客户，再做政府客户。行业客户是真正需要使用云计算的，他们对云计算的需求是强烈的。UCloud 的发展目标，是让用户像搭积木一样方便地创建应用。目前中国的云计算用户还不够成熟，很多时候需要架构师的支持。AWS 的架构咨询支持是收费的，属于增值服务。而 UCloud 的架构及解决方案咨询、售前及售后的技术支持全部是免费的。除此之外，AWS 在提供 IaaS 服务以外，也做 SaaS 应用，难免和用户业务产生冲突。

又拍云

又拍云将公司所提供的全部服务视为一个大的产品，所有的功能都是为了满足不同用户需求而开发的特性。又拍云的主要功能有三个：第一是 CDN 类功能，又称为网络加速，是又拍云的核心功能；第二是云存储功能，用户可以方便地从手机等不同客户端上传数据；第三是云处理功能，允许用户在分享图片的时候针对客户端进行适配。

每年年底，公司层面会决定第二年的产品方向，确定在某个时间点上要实现哪些功能或者特性。产品部门和研发部门根据这些大的方向对具体的功能设定优先级和安排资源。在每个季度中，都会对研发计划进行梳理和细化。

和传统的 CDN 业务相比较，又拍云认为其最大的优势在于行业内第一张云 CDN 网络。传统的 CDN 有很多不同的平台，不同的产品由不同的平台来提供服务。但是计算资源的总量是有限的，传统 CDN 的资源分配方法降低了每个平台的性能和服务能力。又拍云充分了解传统 CDN 的弊端，根据当前的网络条件和用户需求

设计出全新的架构。这个新的架构只有一个开放的 CDN 平台，所有的计算资源都云化了，能够按需为所有功能和特性提供服务，从而提高了平台的整体性能。计算资源云化之后，新增加的机房，所有用户都可以使用。这样就提高了所有用户的访问速度，也提升了服务的稳定性。

跟国内友商相比，又拍云将把自己定位成一个服务化的平台。又拍云更侧重于云服务，把云存储、云分发、计算处理等不同服务集成在一起，做成一个整体的服务。又拍云希望通过这样的功能集成组合，能够和用户的需求更好地衔接。传统上，用户需要用不同的接口来实现云存储、CDN、数据处理，但是又拍云只需要一个 API 接口。因此，又拍云是一个融合了不同云服务的平台，是一个云化的服务网络。又拍云认为，云计算说到底还是以服务为主，和卖产品不一样，注重的是一个服务的过程。

2015年中国公有云服务发展报告

服务运营篇

在这个部分，我们以服务运营为主线，探讨不同公有云服务提供商在运维运营和市场拓展方面的思路。我们关心的问题包括数据中心选址、服务等级协议（SLA）、运行维护、应急预案、安全防护、扩容策略、IPv4 储备、服务规模、客户筛选、定价策略、市场竞争，等等。

金山云

金山云在国内有华东和华北两个大的区域（Region）。华北主要是北京、河北、天津；华东在上海、浙江。国外有美国（北美）和香港（东南亚）两个区域，未来会有扩张的考虑。**数据中心都是远程管理的，并不是每个数据中心都需要现场部署运维人员。**系统层面的扩容通过软件实现，硬件层面的扩容需要外包人员进场。

金山云认为国内公司和国外公司对于服务等级协议（SLA）的定义是一致的，但是对待 SLA 的态度是不同的。此外，用户和服务提供商对于 SLA 的理解也



是不同的。用户往往从自身的角度去看待 SLA。譬如说，对于服务提供商来说，网络不可达的现象是不可感知的；对于用户来说，网络不可达的现象是可以感知的。因此，**用户和服务提供商之间的理解和感知是不一致的**。金山云承认这种认知差异是客观存在的，云服务提供商应该设法获得客户端的真实认知。

在过去几年中，金山云处理了各种各样的突发事件，包括人为事故和不可抗力。供电失效导致机房不可用，光纤挖断导致网络不可用，都遇到过。金山云认为任何可能发生的事情一定会发生，绝对不能有任何侥幸心理。至于什么时候发生只是时间问题和概率问题。而是**在方案论证时不能放过任何可能发生的事情，所有可以预知的问题都要有应急预案**。譬如 AZ 对称设计，要理解为什么一个 AZ 要跨数据中心部署。在不可抗力的处理上，取决于人力的实力、技术的实力、公司的实力。譬如天津港发生爆炸的时候，某大型互联网服务提供商的天津机房全部出现了问题。云计算在这个层面上，真的特别需要资源。小型服务提供商基本上没有资源去应对这个事情，他们可能就是在赌一个小概率事件。金山云已经做到了对任何突发事件都有应急预案。

安全是云计算服务的重中之重。AWS 最近发布的产品，都和安全性有关。来自内网和外网的攻击行为，譬如端口扫描，金山云每天都遇到。对付外部攻击没有什么好的办法，基本上就是流量分析，把攻击流量引导到黑洞。核心问题在于

流量识别的效率，流量黑洞的容量，需要从资源不断投入和技术不断迭代两个方面来解决。用户关心的是多长时间自己的业务能够恢复，厂商关心的是怎样处理才不对其他用户造成影响。除了网络攻击行为，还有一些人为的、机制的、管理的漏洞会导致安全盲点，有一些意识不到的没看见的问题。在安全问题上，厂商需要尊重客户隐私。云计算服务提供商往往会把自己当上帝，忽视用户的自主意志。所有的技术都是为了解决业务场景而产生的。**金山云倾向于用技术来解决业务场景，但是不建议用技术来替代客户做决定。**

如果客户需要使用金山云的服务，金山云的商务团队和技术支持团队会同时和客户沟通，了解客户需求及其需求和金山云产品的匹配程度。金山云的业务增长速度很快。云存储领域，用了两年时间接近 200 PB，规模和增速都是全国第一。云主机领域，吃下了游戏行业的大部分客户。今年新推出的手机游戏，大部分运行在金山云上。2015 年金山云在海南开了一次游戏生态大会，到场的有 400 家游戏厂商。金山云目前的服务器总量还没有超过 10 万台。金山云也不认为现在有任何一家国内公有云服务提供商能够达到这个规模。依照目前的增长速度 and 市场需求进行估算，金山云可能在两、三年内达到这个规模。

金山云的计费模式是按天结算。金山云认为计费模式不是一项核心能力，修改计费模式是一件很简单的事情。未来市场竞争极其激烈的时候，金山云可能会修改这个计费模式。金山云有自己的 IPv4 储备，具体数量不方便透露，但是能够满足金山云未来几年的增长需要。

谈到国内云计算市场的竞争，朱桦认为关键在如下几点。

资金。无论是资源还是技术都不足以成为吸引客户的第一要素，第一要素其实是钱。云服务提供商给客户提供的是一种租赁服务，其实是一种财务服务。大家对互联网公司的认知就是互联网公司是不挣钱的。假如你用虚拟机，云计算公司不挣钱，那么云计算公司需要不断的投入，累计下来是一个很恐怖的数字。所有不投钱搞公有云的，基本上都是不可靠的。我觉得对中国的云计算，有本质深刻的理解。钱是第一位的。

必须是技术驱动，绝对不能是商务导向和市场导向，必须要在经营上坚定不

移的持续投入。做商务和市场推广总是能够吸引很多眼球和机会，但是本质上来说，产品和技术的迭代才是最重要的。战略必须聚焦，如果投入大量的资源在市场和商务拓展，必然要牺牲技术产品迭代的资源。因此，你要看他在技术产品上投入的决心。

规模。云计算说白了是靠规模来驱动整个商业模式的产业。如果没有规模，做到3年或者更长时间还没有达到一定规模，就只能掉队。规模本身也是实力的表现。

人才。需要一个没有权威，不迷信，不自负，不冒进，坚定务实朴素的团队。

金山云只知道要全力以赴去做事情，并不关心其他厂商做什么和怎么做。云计算这个行业，从来都没有说应该怎么做。中国作为世界第一第二大的IT产业集群，中国的云计算公司应该有信心超越其他外国公司。中国的市场没有天花板。中国的云计算市场已经衍生出很多和AWS不一样的东西。

金山云将IDC厂商视为金山云的合作伙伴。**现在的客户，越来越倾向于一站式服务。**以前创业的时候，需要自己买服务器，自己跟戴尔打交道，自己把服务器拉到机房，自己上架自己管理，需要协调各种单位和人，流程非常复杂，整个行业的门槛比较高。云计算行业出现所带来的契机，就是创新创业更容易了。金山云注入能力和资源，把计算、存储、网络方面的门槛降低下来。未来很有可能是所有的开发工作都跟创业者没关系了。创业者如果要做视频类的APP，只要用金山云提供的方案就可以了。**IDC和云服务提供商是个共生体。**云计算向上延伸能力更强，能够把客户的需求都扛下来。这时云服务提供商就处于产业链的中枢地位，向下驱动对IDC的需求。目前云计算这个行业还不够成熟，才会出现IDC和云服务提供商之间的矛盾。再过几年，可能就没有这个矛盾了。

金山云没有过多地考虑过OpenStack和OpenStack生态圈这个事情。金山云采用OpenStack，主要是因为团队在起步阶段使用了OpenStack。开始的时候，金山云迎合上游或者意见领袖提出的观点和路线图。慢慢地做下来，金山云发现**开源意见领袖的架构和规划其实无法和金山云的一线需求相比较。**这个时候，金山云和开源意见领袖之间就出现了分歧。金山云是一个具备超强研发能力

的团队，对自己的技术实力有信心。因此，金山云的观点是金山云要自己决定研发方向，不冒进，不轻信。金山云关注社区的变化，同时保留自己的看法。目前 OpenStack 在金山云里仅仅是一个管理工具，并不是金山云的核心。金山云认为**核心能力上的进步才是硬进步**。公关宣传出来的进步不是真正的进步。如果说 OpenStack 代表云计算的话，这完全是偷换概念。金山云是真正把云计算当作事业来做的，切身经历了中国云计算领域的迭代。OpenStack 的目的并没有那么纯粹，只是把目前的很多概念装进去。金山云内部的 OpenStack 版本，在功能和性能上都远超社区版本。金山云对开源的态度是开放的，也热心地希望将金山云的能力反馈给社区和其他开发者。但是作为一个小创业公司，金山云还没有能力影响 OpenStack 社区。等到金山云的能力强大了，金山云还是希望回馈社区的。

中国云计算的格局，2015 年在资金、规模、人力上差别已经很明显了。**目前整个行业正处在一个转折点上**，未来还是取决于产品技术能力的突破上。在这一点上，金山云通过量变已经开始出现质变。纵观整个云计算行业，2016 年相对来说应该会更加蓬勃向上。社会对云计算的认识越来越深刻，这个趋势会自然而然地加强。2014 年不知道什么是云，没有人用；2015 年开始拥抱云，有些人用；2016 年云真的来了，越来越多的人用。

金山云的观点是：**只要务实，不断加强产品技术能力，任何一个云厂商都有巨大的机会**。如果不着力加强产品技术能力，可能就没有机会了。金山云希望有更多靠谱的云厂商出现。这是一个巨大的市场，目前所有的厂商加一起，也只不过占很小的一块。这个行业参与的人越多，在互相竞争中一起往前走，这种状态就越是健康，这个行业就越是会往前走得更快。

美团云

美团云通过位于北京的三个数据中心向客户提供服务，这是因为美团的机房就在北京。**机房值守的工作是外包出去的，不需要美团自己的运维人员**。对于美团来说，机房内网是互通的。对于用户来说，机房内网是不通的。目前，美团云正在把北京 1 区和北京 2 区的用户往北京 3 区迁移。北京 1 区和北京 2 区用的都是第三方机房，服务质量不够好。北京 3 区是美团和运营商合建的机房，服务质

量更好。目前美团正在和运营商合建北京 4 区。未来美团会在上海开设一个区域，这样华东和华北也都有资源了。美团目前尚不考虑建设独立机房。独立机房如果规模不大的话，得到的网络资源就很有限，抗攻击的能力就不够了。**美团给用户使用的是自己的 IPv4 储备，目前投入的是一个 B 段。**美团云手头还有几个 B 段，但是尚未投入使用。

在 SLA 方面，美团云认为国内用户对 SLA 的数字并不看重，但对于实际的服务质量要求很高。服务出现故障时，告诉客户 SLA 是几个 9，以及现在没有超过 SLA 承诺的范畴，用户是不会接受的。只有用户打算诉诸法律时才会拿 SLA 作为一个赔偿的基准。

对于突发事件（包括不可抗力和人为事故）的处理，美团云有一套应急预案并且定期进行演练。突发事件发生时，美团云会在第一时间通知用户。譬如说某个机房出口交换机宕机了，运维人员会手工地把客户的流量从一个区域切换到另外一个区域。美团云会给大客户提供定制化的咨询，譬如建议大客户不要把服务都放在同一个机房，因为这样比较危险。

美团云的云主机和公网 IP 之间不存在一一对应关系，部分云主机没有公网 IP，所以通过端口扫描得到的云主机数量结果会小一些。这些云主机中 90% 以上运行 Linux 操作系统，运行 Windows 操作系统的非常少。这是美团云的用户特点决定的。美团云的用户中传统企业不是很多，传统的公司用 Windows 操作系统多一些。美团云的用户大部分是互联网企业，或者是给传统企业提供软件和服务的独立软件提供商。

美团云认为用户忠诚度的问题在国内的公有云服务中普遍存在。在选型的时候，口碑很重要；用户开始使用后，品质就很重要。美团云有一些用户是从某国际知名云计算服务提供商那里迁移过来的。2015 年里这家知名的云计算服务提供商出现了几个大的问题，但是都没有被报道出来。这些问题包括整个机房宕机，还有机房发生故障的情况。

OpenStack 为云计算的普及做了很大贡献。从技术上来讲，OpenStack 是一个非常重要的项目。但是把 OpenStack 作为一个产品去提供服务，和市场需求

之间有很大差距。把 OpenStack 用在公有云上，需要大量后续的改进和维护工作。**OpenStack 就像是乐高模型，你可以用乐高模型快速搭建原型。但是乐高做的房子是不能住的，你还是需要自己搭建一套。**

在公有云和私有云之间进行比较，**公有云的份额会越来越大，私有云不过是将基础设施迁移到公有云的一个过度阶段。**那些认为私有云比公有云更安全的论调，其实是不成立的。举个例子来说，郊区的别墅和市区的小区，哪个更加安全？当然是市区的安保措施比郊区的安保措施更好。最近几年被报道出来的黑客攻击和网站拖库事件，很多发生在私有云上。黑客在公有云上做这些动作，往往是可以被及时发现及时制止的。**美团云有自己的流量清洗系统，暴力攻击行为累积到一定的阈值时，美团云会自动地进行流量清洗，用户根本就感知不到。**针对账号破解等行为，美团云和做安全的公司合作，一起去进行防范。美团云把安全厂商的产品集成到美团云的产品里面，给用户提供更安全的服务。在使用 agent 和不使用 agent 这个问题上，其实没有谁对谁错。暴力攻击行为可以在物理机和防火墙的层面进行检测和拦截的，但是文件级别的替换和渗透还是需要在主机层面去做。这里的关键是要让用户去选择，要告知用户具体的技术实现。如果用户觉得 agent 可信，就选择用；用户觉得 agent 不可信，就选择不用。

在定价策略方面，早期整体云计算行业利润较高，**美团云希望做一个低毛利的业务通过规模来盈利，因此将“业界最低”作为定价策略。**实际执行下来，美团云发现这个策略并不是一个很好的选择。美团云新的定价策略还没有实施，目前这个定价策略还会维持一段时间。目前国内云计算行业只有降价的没有涨价的，美团云单方面对产品涨价并不现实。**目前美团云的规模还比较小，暂时不考虑赢利的问题。**未来美团云希望能够通过用户满意度来影响定价策略。

中国的公有云市场的发展速度很快。公有云在中国互联网基础设施中的比例，2015 年应该比 2014 年翻一番。公有云领域的竞争，一半是云计算服务提供商之间的竞争，另一半是所有云计算服务提供商跟传统 IDC 之间的竞争。从 2015 年的数据来看，云计算的份额增长很快，可能 2016 年会更快。再往后做下去，会有一些比较大的挑战。现在市场对 IaaS 型产品的接受度已经很高，但是

对 PaaS 型产品和 SaaS 型产品的接受度还不高。传统软件 SaaS 化谁来做？这对于中国的云计算公司是个挑战，中国的 Sales Force 还没有出现。

七牛云

七牛做事有一个优先级的考虑，如果这个领域已经有人做了，七牛就将其整合进来不要自己做，杜绝重复制造轮子。CDN 大趋势是已经从资源为中心转为效果为中心，七牛成熟的多 CDN 融合管控方案，精选主流 CDN 厂商优质节点，与七牛自有高质量节点相结合，全面覆盖各地区各运营商网络，真正做到无盲区。

突发事件（包括不可抗力和认为事故）的处理，在运维的层面是一个需要不断完善的过程。**七牛侧重于通过历史数据识别出常见故障，争取通过自动化手段发现这些故障，然后自动化触发故障处理预案。**这些措施都需要随着时间不断迭代优化。来自外部的攻击行为主要是 CDN 厂商去处理。从 CDN 的角度来看，七牛自建 BGP 网络中间源，消除跨运营商网络回源慢甚至不可达等问题，保护源站带宽不受边缘节点请求波动影响，节省回源成本。

在七牛发展的早期，年增长率曾经达到十倍甚至是百倍，现在年增长率依然保持在五到六倍左右。在七牛所管理的数据当中，冷数据占绝大多数，热数据不到 10%。随着云服务商服务年份的增加，热数据的比例会越来越少。从云存储领域的市场影响力来说，七牛认为我们是国内第一。**以前大部分云存储的厂商都在学习 AWS，现在国内做云存储的厂商都在学习七牛。**在定价策略上，七牛的基本原则是和 AWS 持平。不会刻意把价格压低。

七牛非常关心用户忠诚度的问题。**七牛之所以从存储领域切入云计算市场，就是出于用户忠诚度的考虑。**从产品属性来讲，存储的用户黏性是很高的，因为他的迁移难度最高，因此用户的稳定性也最高。

盛大云

盛大云设置华东和华北两个节点，是从用户分布出发的。北方是北京，南方是无锡，可以同时服务盛大云内部和外部的用户。盛大云的节点没有阿里云这么多，但是在选址方面有很多严格的考量。中国互联网用户的分布有一个明显的特

点，就是集中分布在北京、上海、广州、深圳几个区域。无锡在江苏骨干网上，盛大云在无锡选择了一个 T4 级别的机房，基本上可以覆盖所有南方的用户。在此基础上，以北京的 BGP 线路来覆盖北方的用户。从使用状况来看，这两个节点的用户比例是比较均衡的。2016 年可能会增加一些新的节点。盛大云在各个数据中心都有自己的运维人员，7x24 小时驻场。我们互联网圈子里有一个玩笑，就是**技术再高也经不住挖掘机一铲子**。盛大云选择在当地投入自己的团队，就是要大力保障服务的稳定性。

在服务等级方面，盛大云分为两个部分来设置。**第一个部分是数据中心，为云计算平台提供服务**。这个部分没有设置具体的 SLA，但是可以认为是达到 99.99%。这是因为盛大云在数据中心有自己的运维团队，响应速度非常快。**第二个部分是云计算平台，为云计算客户提供服务**。盛大云有 7x24 小时值班的客户团队，能够很快地响应来自客户的问题，这个部分的 SLA 是 99.9%。有的用户提出了新的需求，或者是发现了软件缺陷。这种情况会有一个流转过程，盛大云会第一时间响应客户，然后转入内部处理流程，最后把处理过程告知给客户。盛大云做了这么长时间的运营，和一些新的云服务提供商相比，在服务方面更深的理解。

对于突发事件（包括不可抗力力和人为事故）的处理，盛大云通过异地灾备来对用户数据进行保护。当发生不可抗力事件时，盛大云内部会首先做一个预判，通过公告向用户说明事件状况。如果短时间内问题无法得到解决，盛大云会把线上的资源切到灾备机房，减少对用户业务的影响。盛大云的云主机主目前还做不到跨机房迁移，但是用户可以登录到控制台去申请其他的资源。跨机房迁移的功能在盛大云的研发列表里面。目前没有任何一家云服务提供商能够解决这个问题。这也是为什么盛大云在选址的时候非常看重机房的管理，包括硬件配置、人的因素、供电是否双路、机柜是否双电源、机房隶属于谁、上级单位的性质等等。

目前运行在盛大云上的云主机数量超过一万台，但是尚未达到十万台，使用运营商提供的 IP 地址。其中，95% 以上的云主机是带公网 IP 的，不带公网 IP 的云主机数量极少。运行 Windows 操作系统的云主机占 60%，运行 Linux 操作

系统的云主机占 40%。盛大云在国内云计算市场起步比较早，早期的用户使用 Windows 操作系统更多一些。操作系统比例也是用户自由选择的结果，体现了用户的性质。盛大云的用户，大都喜欢易用性比较好的 Windows 操作系统。盛大云只想低调务实地在云计算领域做事，不太关心在市场份额的问题。盛大云的用户中，有 40% 到 50% 都是在 2012 年以前注册的，这说明盛大云的用户忠诚度非常高。这些用户一直在持续使用，持续付费。从 2015 年下半年开始，陆续有一些用户从其他友商那里转过来，这是盛大云不断提高用户体验的结果。

盛大云也提供基于 OpenStack 的私有云解决方案。一开始的时候，OpenStack 想要进入公有云市场，希望打破 AWS 的封闭式的公有云格局。盛大云在实际运营过程中发现，客户对于使用 OpenStack 来构建公有云还是处于观望状态，导致 OpenStack 在公有云领域的成就并不大。OpenStack 要打入云公有云，还是需要一点时间。它对公有云的冲击，现在和未来都不会很大。但是 OpenStack 在私有云市场的接受程度还是比较高的，必然会在私有领域占据一席之地。

腾讯云

腾讯云从北京、上海、广州、香港、多伦多提供服务。**腾讯云的选址决策是基于 QQ 和微信产品的热力图做出的。**哪里人多，腾讯云就选址在哪里。中国的经济版图，基本上就是华南、华中、华北，集中了中国的人才和财力。选址香港是因为腾讯总部在深圳，选址北美是因为腾讯当时使用过 AWS 等相关服务，多伦多这个点是和 SoftLayer 合作的，北美的网络连通性比较简单，具体的地理位置不是很重要。2016 年腾讯云将东南亚数据中心提上日程。**腾讯云在各个数据中心都有经理，有一个本地化的团队进行运维。**

不同地理位置之间的网络是打通的。譬如说一个用户在北京和上海两个区域分别开一台云服务器，这两台云服务器是可以放在在同一个子网里的。有这种需求的用户并不多，通常是超大体量的公司才会有这样的需求。**开发这个功能，主要是考虑到容灾和负载分布。**譬如说，北方的用户连接到北方的 IDC，用户体验会好一些。

对于突发事件（包括不可抗力和人为事故）的处理，腾讯云有一套应急预案。譬如说天津港爆炸事件，腾讯云启动了应急措施，非常好地保障了服务，完整的服务没有受到任何影响。

如何应对来自内网和外网的攻击行为，譬如端口扫描，腾讯云给用户提供了安全服务，用户可以选择是否开通。如果用户选择开通，腾讯云会在用户的操作系统里面安装 agent。**用户可以选择用户信得过的 agent。**腾讯云认为需要给公众的客户传播这样的理念：选择大的公司如腾讯云、阿里云、AWS，使用他们提供的 agent 其实是没有风险的。这些东西是放在阳光下的，agent 做了什么事情，都是可以知道的，没有必要太担心。大的厂商，在安全方面都有巨大的投入，用 agent 比不用 agent 要好。

腾讯云的帐号体系是与 QQ 号码绑定的。QQ 号码背后是一整套的账号管理系统，可以是 QQ、可以是邮箱、可以是电话号码。如果用户不打算使用 QQ，简单地把它当作一个邮箱使用也是 OK 的。腾讯云慢慢地也要支持更多的实名认证方式。譬如说 2C 类型的用户，已经可以使用微信支付的方式来购买云主机了。在用户体验这一块，腾讯云是愿意不断优化的。

在公网 IP 方面，腾讯云使用自己的 IPv4 储备。腾讯云也有弹性 IP(Elastic IP, EIP) 的功能，用户可以按需使用，不需要的时候就释放掉。腾讯云的公网 IP 是够用的，不会对腾讯云未来的发展造成障碍。

在腾讯云所提供的虚拟机中，Windows 操作系统的比例要远高于 15%，但是也没有 50%。在腾讯云里还有一部分物理机托管业务，这部分的业务从外界看不出来。在游戏领域，腾讯云是最大的服务提供商。**游戏类客户使用 Windows 的比例非常高，因此腾讯云的 Windows 操作系统比例会比阿里云高一些。**在 IDC 里面，如果游戏类客户比较多的话，可能会发现 Windows 操作系统占 40% 到 50%。

腾讯云的主要竞争对手更多还是比较传统的 IDC，比如说世纪互联。**迁移到腾讯云的客户，很多都是从传统 IDC 迁移过来的。**腾讯云的估算是传统 IDC 在数据中心市场的占比还在 9 成以上。

国内公有云服务的价格是比较低的。如果和 AWS 的价格相比，AWS 的定价比国内要高。两、三年前，国内友商都在通过下调价格来占领市场。但是最近大家有一些默契，认为应该提高产品的能力，提高对用户的价值，不要把价格压下去。**企业客户并不是真的非常对价格敏感，而是希望价格合理。**腾讯云不希望在某个产品或者功能点上做价格的 PK 去迎合用户，这样的模式很难持续。

在用户忠诚度这个问题上，腾讯云觉得首要在于练好内功，**不能要求客户对你忠诚。就像任何人之间的交往，你要给对方带去价值，对方才会给你带来回报。**

关于 OpenStack 和 OpenStack 生态圈，腾讯云认为这样一套体系非常完整。腾讯云也在和国内一些 OpenStack 厂商合作推出混合云解决方案。在金融领域，私有云的需求很强烈。**用户对资源的要求弹性比较大的，建议优先在公有云上做。**如果用户愿意支付相应的成本，则可以在 OpenStack 上实现。OpenStack 和公有云的发展，应该是互相促进的关系。私有的 IDC 必定是单点的，用户需要为所有的资源买单。使用公有云，则只需要为你使用的时段和服务买单。**公有云很重要的一点，在于它的网络是否真的四通八达。**

UCloud

UCloud 分别从北京、广东、香港和北美四个不同的地理区域提供服务。这个**选址决定是因为 UCloud 最初的客户分布，他们的用户基本上都在北京、上海、广深。**选择香港和北美，是因为 UCloud 的客户有往香港和北美拓展业务的需求。UCloud 在北京和广东部署了自己的运维人员，但是在香港和北美没有。UCloud 在海外节点的部署和扩容基本上是自动化的，只要厂商把设备运到机房，机房帮我们把设备放上去就 OK 了。

在服务等级协议（SLA）的设计方面，UCloud 分为三个阶段去进行。**第一阶段是了解用户需求**，也就是他们能够接受什么样的 SLA。根据 UCloud 研发人员从在腾讯和盛大学习到的经验，对用户的需求和期望值有一个初步了解。**第二阶段是设计**，从理论上计算我们能达到什么水平。**第三阶段是改进**，从实际运行数据计算、验证、改进。

在突发事件（包括不可抗力和人为事故）的处理方面，有两个问题不是

UCloud 能够控制的。一个是机房问题，一个是全国性网络问题。其他友商遇到过光纤被挖断、电力被雷击的问题。**UCloud 的侧重点在于机房掉电之后如何迅速地把用户的业务从一个机房转移到另外一个机房。**

应对来自内网和外网的攻击行为，譬如端口扫描，是 UCloud 的优势和特长。在安全方面，UCloud 采取应对措施有三个。第一是区分职责，UCloud 处理网络底层，保障网络的安全可靠，防 DDoS 等等虚拟机以外的攻击行为，用户则负责虚拟机以内的攻击行为。第二是引进第三方安全解决方案。第三是建立内部操作审计系统，进行镜像分析。**UCloud 严格限制内部工作人员接触或者读取用户的镜像。对于在用户的虚拟机内部预装 agent 的行为，UCloud 是严格禁止的。**虽然这样在管理上非常不方便，但是用户会认为 UCloud 是可信的。在网络隔离方面，UCloud 从 2012 年开始就给所有用户默认创建一个私有网络，内网不能扫描，但是公网还是能扫描。内部隔离是必然要做的事情。用户还可以通过公网的防火墙来限制访问。

UCloud 通常提前半年对机房资源进行整体准备，根据用户业务需求对资源进行动态扩容，保证 25% 到 30% 之间的空闲资源。**当空闲资源小于 25% 时就会对机房进行扩容。**每个月都会对下个月的资源需求情况进行预测，制定每周的资源增长计划表。对于大的客户，有客户经理事先沟通，提前了解他们对计算资源的需求情况。通过如上三个措施，确保我们的资源是足够的充分的。

UCloud 没有超售的情况，因为我们的客户和其他客户很不一样。个人客户一般就放一个网站，一个月可能只有一两次百度爬虫的访问。对于这种类型的客户，是有很大超售空间的。但是 **UCloud 的客户都是企业客户，他们对 CPU 资源的使用要求是很高的，所以很难去做超售。**CPU 和内存是最难复用的。IP 资源是绝对复用不了的。

在公网 IP 方面，UCloud 有两个 B 段的 IPv4 储备，同时也使用运营商提供的公网 IP。此外，UCloud 还在准备自己的 BGP 和 AS 编号。在 UCloud 所提供的云主机中，Linux 操作系统大概占 90% 左右，Windows 操作系统只有 10% 左右。**UCloud 没有个人客户。**就好像一句话这样说：“如果你什么都信仰的话，

等于你没有信仰。”服务客户也是如此，如果你谁都想服务好，那只能谁都服务不好。

在虚拟化产品和非虚拟化产品的比例方面，92%到95%是虚拟服务器，5%到8%是物理服务器。使用物理服务器的客户，基本上都是2013年之前的客户，我们也称为有历史包袱的客户。2013年之后的客户，大部分都用云计算了，没有历史包袱。在2013年之前，云计算还没有那么流行，只能用物理机。这些用户的体量可能不大，但是对UCloud的战略意义很强，所以要把它们拉进来。

UCloud的目标是先把他们拉进来，等到他们的服务器生命周期到了再逐步转到云上，这就是这些客户的价值。UCloud云认为0和1之间是有灰度的。UCloud提供的产品，在帮助客户彻底转换到云计算之前，能够有一个混合的模式来保护客户的投资。这也是一种对客户负责任的态度。

UCloud的主要竞争对手是阿里云。市场份额则不好确定，因为大部分公司都没有公布业务数据，分母和分子都是未知数。在定价策略上，UCloud的价格略高于行业平均水平。UCloud认为企业客户关心稳定性和可靠性，UCloud要让优质的产品有更好的价格，这样才能够真正解决中国大部分行业要靠低价取胜的思路。这也就是供给侧改革，好的商品是有价值的，好的服务是有价值的。所以UCloud的价格虽然会贵一些，但是企业是会接受的。这也就是为什么UCloud不做个人客户。这个定价策略和UCloud的目标客户群是一致的。UCloud的客户都是企业级客户。**只要你服务好，稳定性好，他是不会因为价格问题离开的。所以UCloud的客户流失率非常低，并且越是大客户越是低。**

在规模上，UCloud最大的一个客户拥有的主机数量超过1万台。UCloud的整体规模大于10万台云主机。运行在UCloud上的云主机，大部分都需要公网IP，除非是数据库等内部业务。但是**UCloud并没有一个详细的统计数据**。从数量来说，因为UCloud的中小客户比较多，他们的主机大多数是有公网IP的。对于大客户来说，内网IP会比外网IP多。针对“云主机是一个赔钱的生意，云计算赚钱的业务是存储和网络”这样的观点，UCloud没有这样分开计算过。

未来CDN是一个非常重要的方向。娱乐产业起来了，基于手机的视频点播

和直播需求很大。CDN 厂商有明显的优势，譬如网宿。地区性的 IDC 还需要时间来改进和完善他们的全国布点。

OpenStack 最近的发展趋势是非常快的，UCloud 也在学习和了解 OpenStack。**UCloud 和全球第一大 OpenStack 服务提供商 Mirantis 合资在国内成立一个公司，通过这个合资公司去满足客户在私有云方面的要求。**目前 OpenStack 已经成为开源社区最有活力、贡献最大的开源项目了。但是 UCloud 没有看到用 OpenStack 做公有云做的好的公司。OpenStack 在私有云领域有相当大的影响力，能够满足相当多用户的需求，但是它做公有云还是有压力的。因此，UCloud 的公有云是自己研发的，私有云则和 Mirantis 联合提供基于 OpenStack 的解决方案，能够满足用户的混合云的需求。

很多人认为云计算就是做 IDC 生意的，UCloud 对此并不赞同。我们认为云计算有五个阶段。第一是替换 IDC 和服务器设备。这个领域大概 1100 亿到 1300 亿的年产值，目前云计算所占的比例还非常小，需要把基础的计算、存储、网络做好。第二是在计算、存储、网络的基础上，用户会需要托管服务（Hosted Services）。托管服务的好处是资源利用率更高，这是一种软件服务，不是卖主机。软件服务的利润值更高，黏性更强。第三是 UCloud 在 2015 年主推的软件市场——U 市场。软件厂商把他们的解决方法放到 UCloud 上来，UCloud 帮助他们推广，形成企业级的软件应用市场（企业级的 App Store）。第四是打造服务的应用市场，引进为中小企业和传统企业做服务的公司，譬如运维服务、开发服务、安全服务。很多公司一个月付十到二十块钱，就能够享受这些服务。第五是大数据阶段，UCloud 目前还没有完善的产品模式。

云计算的价值不仅仅是赚钱。**云计算的核心价值是“使用比拥有更有价值”。**用户不需要拥有计算资源，他只需要简单的使用就好了。这种转变会改变社会的组织运转模式。人们常说，一个组织最大的资产就是人。现在一个公司不再需要招聘运维人才，这些可以通过云计算自动实现。这就像现在的分享经济，由不属于你公司的员工来帮助你。**UCloud 的目标就是通过 API 的模式，帮助客户获得他所需要的资源。**淘宝非常伟大，可以帮助客户获得他所需要的商品，但是这

还不够。UCloud 的目的，是让有一技之长的人能够在 UCloud 的平台上为用户提供服务，发挥自己的价值。

又拍云

又拍云将云计算定义成一个服务，对服务等级协议（SLA）的要求是第一的。

SLA 是一个比较重要的指标。跟客户签约的时候，又拍云会把 SLA 放到合同里面。在研发层面，又拍云通过架构优化设计了自我修复的功能。在运维层面，又拍云通过 7x24 小时的监控，确保出现问题能够及时修复。

众所周知国内网络非常复杂，单纯依靠一个较好的机房或者一个较好的地区去提供服务，不可能达到客户要求。**又拍云依靠全局的网络去解决局部性的问题。**比如，从南到北，电信和联通之间的长距离传输经常有拥塞，又拍云架构了一个内部网络，不会由于主干网的拥塞而导致网络不可用。公司下了很大的力气来克服国内复杂网络的单点故障问题。又拍云在全国有 150 多个机房，如果某个机房发生了网络中断，公司可以直接把这些机房从基础设施中移除，流量就不会被导向这些机房。对于用户而言，某个机房的单点失效是不可感知的。

云服务更新迭代的速度非常快。**又拍云在发布周期和发布进度两个方面都有严格控制。**又拍云有一个专门的房间做发布，叫做发布室。新产品上线时需要进行两次发布，一次是预发布（灰度发布），一次是全局发布。灰度发布是发布到 150 多个机房中的 3 个，这时产品和运维人员要集中到发布室，发现问题要及时处理。灰度发布稳定运行 24 小时之后，再全局发布。**又拍云不允许某一个人员单独把某一个功能单独发布到产品平台上。**

又拍云有三个比较大的防攻击的机房。在遇到 DDoS 攻击时，公司根据 DDoS 的量级来做不同的处理。又拍云的 150 个机房都有一定的上行容量，几百个 GB 的流量攻击属于小量级的攻击，基本上不需要理会。**遇到大规模攻击时，又拍云会把相应的流量调度到这 3 个防攻击机房，再把流量分散到其它的机房。**在这个方面，公司已积累了一定的算法，筛选攻击来源和攻击对象，在此基础上采取一些更细致的应对措施。同时，又拍云还跟运营商合作，在省级网关那里就把攻击流量给过滤掉，不会下发给又拍云的服务器。

在成长速度方面，又拍云现在全国有 150 多个机房，4000 多台服务器，带宽到了 1.5 TB。日均 PV 数量超过 500 亿，新增存储容量超过 5 PB。在又拍云所管理的数据当中，热数据比例在 10% 以内，90% 以上是冷数据，带宽主要服务于 10% 以下的热数据。

又拍云已经实现了自动化运维，运维体系非常完善，无需在机房部署又拍云的运维人员。服务器厂商出货，配上又拍云的 U 盘发到数据中心。机房人员帮又拍云上架、加电、开机，从 U 盘启动操作系统，启动完成后自动加入又拍云的集成平台。

又拍云坚持自主定价，不会参考其它友商报价。得益于整体规模的扩大，又拍云的优势得以展现，在此前提下，又拍云做过几次幅度很大的调价行为。又拍云的产品定价非常纯粹，基于服务能力和成本，制定一个比较合理的价格。因此，又拍云的调价并不是为了获得竞争优势或者占领市场。

又拍云本身是一家创业型公司，非常注重处于创业期的客户。**又拍云是目的培养两三年以后的客户，会对创业型的客户做很大的支持工作。**公司把自己和客户绑定在一起，和客户一起发展。

目前，**又拍云客户流失率非常低，客户忠诚度非常高。**客户把又拍云当成亲密的合作伙伴，又拍云也把客户作为公司的合作伙伴。这跟又拍云的基因有很大的关系。云计算本身就是一种服务，服务质量好，产品够稳定、够好，用户把东西放在这里会非常放心。对于服务型的公司来说，客户的忠诚度体现在你能够给客户做什么东西。

又拍云的目标是打造一个综合的云服务，或者说是一个可编程的 CDN。用户可以在又拍云的 CDN 平台上做开发，不需要单独购买云主机或者是物理机，也不需要区分云主机和云存储。在又拍云这个综合的云服务上，计算可以发生在不同的层面，譬如说计算可以发生在最前端。从大的方向来说，又拍云类似于 AWS 推出的 Lambda 产品，但是要比 Lambda 更贴近用户。

2015年中国公有云服务发展报告

用户体验篇

在这个部分，我们以用户体验为主线，对不同公有云服务提供商的产品进行一些小规模测试。这些测试旨在探测客户关心的几个关键参数：

- (1) 服务规模；
- (2) 网络与存储吞吐能力；
- (3) 资源隔离状况；
- (4) 客服能力。

需要说明的是，这些测试仅仅是试探性的探测（probe），并非严谨的基准测试（benchmark），测试结果反映的只是测试当时的用户体验。作者本人的技术背景与云主机类产品比较接近，对云存储领域的了解相对有限。因此，这些测试仅针对设施层面的云主机类产品，并且没有完整覆盖所有国内的 IaaS 服务提供商。（在此作者谨向七牛云和又拍云这两家公有云服务提供商致以诚挚的歉意。）

服务规模

针对服务规模的测试，是通过端口扫描进行的。针对一个特定的 IaaS 服务提供商，这个测试分为两个步骤进行：

1. 在所有区域分不同时段（时间跨度长达一个月）大量创建云主机，通过枚举得出云主机所用公网IP所在的B段列表，并通过公开的信息进行矫正；
2. 对所有的 B 段针对22、80、443、3389端口进行扫描，将扫描结果记录到数据库。

有些 B 段 IP 地址，可能超出了 IaaS 服务提供商所拥有 IP 资源范围。譬如某些服务提供商使用了运营商提供的 IP 地址，在同一个 B 段里面还有用于其它用途的 IP 地址。有些 B 段 IP 地址，虽然由某个服务提供商拥有，但是并非用于 IaaS 服务。因此，端口扫描得到的结果，反映的是从外界可以探测到的服务规模上限。考虑到防火墙、安全组、部分云主机未配置公网等等多种因素，端口扫描的结果是小于实际服务规模的。

网络与存储吞吐能力

针对网络吞吐能力的测试，是在同一个区域内启动 N 对云主机。在所有的云主机内安装 Apache 服务，提供一个 100 MB 大小的文件下载。在每一对云主机之间，在每台云主机上启动多个线程从对方下载如上所述 100 MB 大小的文件，单次测试持续时间 15 分钟。由于供下载的文件是同一个，该文件在第一次被读取之后便驻留在内存当中，不再产生新的磁盘 I/O。因此，这个测试探测的是两台云主机之间的内网带宽。N 的取值范围，从 1 逐渐增加到 10，目的在于探测单个用户可以使用的网络带宽边界。测试中使用的第一对云主机，一台在用户账号 A 中，一台在用户账号 B 中，目的在于测试网络资源隔离状况。针对一个特定的 IaaS 服务提供商，这个测试在不同时段进行多次，以了解不同时段对网络性能的影响。

针对存储吞吐能力的测试，是在同一个区域内启动 N 台云主机。在每台云主机上挂载 M 块云硬盘创建一个 RAID0 磁盘阵列。在云主机上启动多个线程，分别往磁盘阵列上写入多个远大于云主机物理内存的大文件。单次测试持续 15 分钟，记录测试过程中的磁盘写入带宽。这个测试分为三个步骤进行：

1. M 的取值为1，探测单台云主机上单块云硬盘的存储带宽上限；

2. M 的取值在2到4之间，探测单台云主机上一个磁盘阵列的存储带宽上限；
3. N 的取值在1到10之间，探测单个用户可以使用的存储带宽上限。测试中使用的前两台云主机，一台在用户账号 A 中，一台在用户账号B中，目的在于测试存储资源隔离状况。针对一个特定的 IaaS 服务提供商，这个测试在不同时段进行多次，以了解不同时段对存储性能的影响。

作者也注意到一些公有云服务提供商采取了“地理区域——可用区——集群”这样的结构设计。在同一个可用区中，尽可能将同一用户所使用的计算资源分配到同一个集群。因此，针对网络吞吐能力的测试结果和针对存储吞吐能力的测试结果反映的可能是一个可用区中某一个集群的网络吞吐能力和存储吞吐能力。

客服能力

针对客服能力的测试，是在云服务提供商的 Web 控制台里提交工单。工单的内容包括要求提高配额、询问基础性的使用问题、报告缺陷等等。这部分的测试，一方面在于了解客服的响应速度，另一方面在于了解客服处理能力。

阿里云

阿里巴巴集团在自治域 AS37963、AS45102 中共声明了 120 个 B 类 IP 地址段以及多个 C 类 IP 地址段。

2016 年 3 月，从公网对全部 120 个 B 类 IP 地址段针对 22 (SSH) 和 3389 (RDP) 端口进行扫描，有 26.5 万个 IP 地址可以通过 22 端口创建网络连接，有 21.5 万个 IP 地址可以通过 3389 端口创建网络连接。

2016 年 9 月，从公网对如上所述 IP 地址段针对 22 (SSH)、80 (HTTP)、443 (HTTPS) 和 3389 (RDP) 端口进行扫描。有 35 万个 IP 地址可以通过 22 端口创建网络连接，有 92 万个 IP 地址可以通过 80 端口创建网络连接，有 9 万个 IP 地址可以通过 443 端口创建网络连接，有 25 万个 IP 地址可以通过 3389 端口创建网络连接。

需要说明的是，如上所述 120 个 B 类 IP 地址段并非全部用于阿里云的公有云服务。阿里巴巴集团下的其他业务譬如淘宝网和支付宝所使用的 IP 地址也

都在这 120 个 B 类 IP 地址段中。根据章文嵩 2011 年 5 月在第三届中国云计算大会上的演讲，淘宝网的生产服务器大约为 20,000 台^[11]。根据高山渊 2012 年 6 月在 QClub 深圳站上的演讲，阿里巴巴集团的服务器规模接近 10 万^[12]。根据工信部电信研究院发布的《云计算白皮书（2014 年）》^[3]，截止到 2013 年 9 月运行在阿里云上的 Web 服务器数量达到 18,000 个，比 2012 年增长了 500%。根据 NetCraft 在 2015 年 6 月发布的数据^[13]，阿里云所管理的 Web 服务器达到 45,000 个。**考虑到阿里巴巴集团过去五年中的业务增长对计算资源的需求，阿里云公有云部分所使用的 IP 地址（包括物理机和虚拟机）可能只占如上所述活跃 IP 地址中的一小部分。**

2016 年 3 月，在阿里云各个区域内创建云主机，并对云主机所在的 A 类内网 IP 地址段针对 22 和 3389 端口进行扫描，有 39 万个内网地址可以通过 22 端口创建网络连接，有 8 万个内网地址可以通过 3389 端口创建网络连接。在可以通过 22 端口连接的 IP 地址中，又发现了大量活跃的 3306（MySQL）端口和 11211（Memcached）端口。运行在 11211 端口的服务，大部分可以通过 SET 和 GET 命令直接进行操作。运行在 3306 端口的服务，有一定数量可以基于社会工程数据库使用 root 帐号通过自动化测试程序登录。在可以通过 3389 端口连接的 IP 地址中，发现了部分活跃的 1433（SQL Server）端口。运行在 1433 端口的服务，也有一定数量可以基于社会工程数据库使用 Administrator 帐号通过自动化测试程序登录。由于 SQL Server 服务可以使用 Windows 身份验证，有理由认为一定数量运行 Windows 操作系统的云主机已经沦为肉鸡。

作者还注意到，**在阿里云各个区域进行内网扫描获得的端口数量是高度一致的**。深圳、杭州、青岛、北京、上海、香港、美西这七个区域的活跃端口数量，精确到千位数都是完全相同的。唯一的一个例外是新加坡区域，原因不明。

在如上所述端口扫描和自动化登录测试中，无论测试流量来自公网还是阿里云内网，测试程序均未检测到连接被拒绝或者重置等主动防御行为。在针对 1433、3306 和 11211 端口的测试中，测试程序仅进行计数而不记录任何可以识别对方主机的数据。

阿里云内网带宽测试（MB/s）

测试 编号	节点 1	节点 2	节点 3	节点 4	节点 5	节点 6	节点 7	节点 8	节点 9	节点 10	节点 11	节点 12	节点 13	节点 14
1	60	60												
2	60	60	60	60										
3	60	60	60	60	60	60								
4	60	60	60	60	60	60	60	60						
5	60	60	60	60	60	60	60	60	60	60				
6	60	60	60	60	60	60	60	60	60	60	60	60		
7	60	60	60	60	60	60	60	60	60	60	60	60	60	60

上表所示是在阿里云杭州区域进行网络带宽探测的结果。测试中使用了 7 对云主机，所有云主机都部署在同一个可用区内。**我们首先使用同样的测试程序对不同的云主机实例类型进行测试，发现不同的云主机实例类型所能够达到的内网带宽是一样的。**

考虑到批量测试的费用问题，如上测试使用的云主机实例类型为“系列二：通用型 n1”，配置 1 颗 vCPU 和 1GB 内存。在参与测试的 14 台云主机中，1 号云主机在一个用户帐号中，2 ~ 14 号云主机在另外一个用户帐号中。1 号云主机和 2 号云主机配为一对，3 号云主机和 4 号云主机配为一对，以此类推。在编号为 1 的测试中，只有第一对云主机产生网络流量，其他云主机处于空闲状态；在编号为 2 的测试中，第一对和第二对云主机产生网络流量，其他云主机处于空闲状态，以此类推。如上测试在一个月中的不同时段进行了多次，不同批次的测试结果之间高度一致。作者将云主机的总量增加到 10 对（共 20 台），可以得到同样的测试结果。基于如上测试，可以认为阿里云的网络质量达到了较高的水平，具体表现在：

- 以云主机为单位进行精确限流，吞吐量指标基本没有发生抖动；
- 在小规模测试中，未能探测到单个用户能够使用的网络带宽上限；
- 在小规模测试中，未能探测到单个用户大量占用网络带宽对其他用户使用网络产生影响。

阿里云存储带宽测试 (MB/s)

测试编号	节点1	节点2	节点3	节点4	节点5	节点6	节点7	节点8	节点9	节点10
1	400									
2	400	400								
3	400	400	400							
4	400	400	400	400						
5	400	400	400	400	400					
6	400	400	400	400	400	400				
7	400	400	400	400	400	400	400			
8	400	400	400	400	400	400	400	400		
9	400	400	400	400	400	400	400	400	400	
10	400	400	400	400	400	400	400	400	400	400

上表所示是在阿里云杭州区域进行存储带宽探测的结果。测试中使用了 10 台云主机，所有云主机都部署在同一个可用区内。

首先，我们使用不同的云主机实例类型挂载单块云硬盘进行测试，可以达到阿里云文档所标注的 256MB/s 带宽上限。此外，我们发现存储带宽上限与云硬盘的容量有关，但是与云主机实例类型无关。

其次，我们在同一台云主机上挂载多块云硬盘创建 RAID0 磁盘阵列进行同样测试。与单块云硬盘相比，用两块云硬盘创建的 RAID0 磁盘阵列可以达到 400MB/s 的存储带宽。用三块或者四块云硬盘创建的 RAID0 磁盘阵列，其存储带宽和用两块云硬盘创建的 RAID0 磁盘阵列是同样的。

考虑到批量测试的费用问题，如上测试使用的云主机实例类型为“系列二：通用型 n1”，配置 1 颗 vCPU 和 1GB 内存，挂载两块 500GB 的云硬盘配置成 RAID0 磁盘阵列。在参与测试的 10 台云主机中，1 号云主机在一个用户帐号中，2 ~ 10 号云主机在另外一个用户帐号中。在编号为 1 的测试中，只有 1 号云主机产生存储流量，其他云主机处于空闲状态；在编号为 2 的测试中，1 号和 2 号对云主机产生存储流量，其他云主机处于空闲状态，以此类推。如上测试在一个月中的不同时段进行了多次，不同批次的测试结果之间高度一致。将云主机的总量增加到 20 台，可以得到同样的测试结果。基于如上测试，可以认为阿里云的

存储质量达到了较高的水平，具体表现在：

1. 以云主机和云硬盘为单位进行精确限流，吞吐量指标基本没有发生抖动；
2. 在小规模测试中，未能探测到单个用户能够使用的存储带宽上限；
3. 在小规模测试中，未能探测到单个用户大量占用存储带宽对其他用户使用存储产生影响。

在针对客服能力的测试中，作者通过阿里云 Web 控制台里提交了两个工单。第一个工单的响应时间为 40 分钟，第二个工单的响应时间为 70 分钟。两个工单询问的是同一个问题：一台云主机挂载多块云硬盘创建 RAID0 磁盘阵列可以达到的存储性能。**在两个工单的答复中，作者均未获得正确的解答。**通过阿里云 Web 控制台对云主机进行销毁操作时需要进行短信验证，作者在测试过程中遇到了短信功能失效的情况。

为了观察阿里云的故障发现与处理效率，作者未通过工单系统报告此故障。等待了四个小时之后，故障依然存在。于是作者通过微博与一个包括多位阿里云员工的微信群公布了此故障。在微博和微信上，均有阿里云的员工主动联系作者了解情况，45 分钟之后故障得到解决。这个事件似乎表明在接近五个小时的时间里没有其他阿里云用户发现同一故障。换句话说，在接近五个小时的时间里，没有其他阿里云用户通过阿里云 Web 控制台进行销毁云主机的操作。如果这个推断成立，则**意味着阿里云的用户基本上是把云主机当成是长期运行的 VPS 服务器来使用的。**

金山云

金山云对客户的挑选比较苛刻。作者自主在**金山云的网站上注册帐号，可以完成注册但是无法激活帐号**。未激活帐号依然可以对帐号进行充值，但是充值完成之后无法创建云主机，也无法使用金山云提供的任何其他资源。作者通过在线客服功能联系到金山云的客服人员，客服人员提供了一个激活帐号的连接，但是依然无法成功激活帐号。（注：2016 年 5 月 31 日前，金山云客户网上注册，需要通过线下人员审核后可激活账户。6 月 1 日后，金山云客户可实现网上自助注册。）

金山云在自治域 AS59019 中声明了多个 C 类 IP 地址段，IP 地址总数接近一个 B 段。

2016 年 9 月，从公网对如上所述 IP 地址段针对 22（SSH）、80（HTTP）、443（HTTPS）和 3389（RDP）端口进行扫描。有 1500 个 IP 地址可以通过 22 端口创建网络连接，有 1900 个 IP 地址可以通过 80 端口创建网络连接，有 1300 个 IP 地址可以通过 443 端口创建网络连接，有 300 个 IP 地址可以通过 3389 端口创建网络连接。

除此之外，作者未能对金山云进行其他用户体验层面的测试。

美团云

美团云启用的公网 IP 地址只有一个 B 段。通过 ip-tracker.org 进行查询，未能确认这些 IP 地址属于美团云。

2016 年 3 月，从公网对该地址段中针对 22（SSH）和 3389（RDP）端口进行扫描。有 3,700 个 IP 地址可以通过 22 端口创建网络连接，有 1600 个 IP 地址可以通过 3389 端口创建网络连接。

2016 年 9 月，从公网对该地址段针对 22（SSH）、80（HTTP）、443（HTTPS）和 3389（RDP）端口进行扫描。有 5,500 个 IP 地址可以通过 22 端口创建网络连接，有 6,400 个 IP 地址可以通过 80 端口创建网络连接，有 3,000 个 IP 地址可以通过 443 端口创建网络连接，有 2,000 个 IP 地址可以通过 3389 端口创建网络连接。

由于美团云的规模相对较小，作者未对 1433、3306 和 11211 等端口进行扫描和自动化登录测试。**基于同样的原因，作者也未对美团云进行网络、存储、客服等方面的测试。**

青云

青云启用的公网 IP 地址有 4 个 B 段。通过 ip-tracker.org 进行查询，未能确认这些 IP 地址属于青云。

2016 年 3 月，从公网对全部 4 个 B 类 IP 地址段针对 22（SSH）和 3389（RDP）端口进行扫描，有 7,000 个 IP 地址可以通过 22 端口创建网络连接，有 2,000

个 IP 地址可以通过 3389 端口创建网络连接。在青云的基础网络上创建云主机，可以在云主机所在的 A 类内网 IP 地址段扫描到大量活跃的云主机。在青云的私有网络上创建云主机，则无法扫描到不属于用户自己的云主机。

2016 年 9 月，从公网对全部 4 个 B 类 IP 地址段针对 22 (SSH)、80 (HTTP)、443 (HTTPS) 和 3389 (RDP) 端口进行扫描。有 5,500 个 IP 地址可以通过 22 端口创建网络连接，有 14,100 个 IP 地址可以通过 80 端口创建网络连接，有 4,400 个 IP 地址可以通过 443 端口创建网络连接，有 1,700 个 IP 地址可以通过 3389 端口创建网络连接。

基于如上端口扫描结果，青云的总体规模还是比较小的。即使是规模最大的一个可用区，云主机的数量级也不过是千位数而已。这样的规模，对于一个内部自用的私有云来说可能不小，但是对于面向公众提供服务的公有云的确不大。作者注意到青云于 2016 年 1 月高调发布了一个“超大规模网络 SDN/NFV 2.0 网络”。与青云的实际规模相比，这样的宣传未免名不副实。

由于青云的规模相对较小，作者未对 1433、3306 和 11211 等端口进行扫描和自动化登录测试。同时，考虑到青云在国内云计算行业具有很高的知名度，作者也对青云进行了网络、存储、客服等方面的测试。

青云内网带宽测试 (MB/s)

测试 编号	节 点1	节 点2	节 点3	节 点4	节 点5	节 点6	节 点7	节 点8	节 点9	节 点10	节 点11	节 点12	节 点13	节 点14
1	115	115												
2	115	115	115	115										
3	105	95	90	105	95	90								
4	70	65	65	75	65	75	65	65						
5	60	55	55	60	55	65	55	55	65	60				
6	50	50	50	55	50	55	50	50	55	55	55	55		
7	40	40	40	45	40	45	40	40	45	45	40	50	45	45

上表所示是在青云北京 2 区 (PEK-2) 进行网络带宽探测的结果。测试中使

用了 7 对云主机，所有云主机都部署在同一个区域内。我们首先使用同样的测试程序对不同的云主机实例类型进行测试，发现不同的云主机实例类型所能够达到的内网带宽是一样的。考虑到批量测试的费用问题，如上测试使用的云主机实例类型为“超高性能主机”，配置 1 颗 vCPU 和 1GB 内存。在参与测试的 14 台云主机中，1 号云主机在一个用户帐号中，2～14 号云主机在另外一个用户帐号中。1 号云主机和 2 号云主机配为一对，3 号云主机和 4 号云主机配为一对，以此类推。在编号为 1 的测试中，只有第一对云主机产生网络流量，其他云主机处于空闲状态；在编号为 2 的测试中，第一对和第二对云主机产生网络流量，其他云主机处于空闲状态，以此类推。如上测试在一个月中的不同时段进行了多次，不同批次的测试结果之间基本一致。**基于如上测试，可以认为青云的网络质量相对较低，具体表现在：**

- 1. 没有对云主机采取限流措施，吞吐量指标存在大规模抖动；
- 2. 在小规模测试中，仅用6台云主机即可探测到网络性能恶化的迹象；
- 3. 随着参与测试的云主机数量的增加，网络性能恶化极快；
- 4. 单个用户可以使用的网络带宽上限低于 700MB/s；
- 5. 在小规模测试中，可以观察到单个用户大量占用网络带宽对其他用户使用网络产生影响。

青云存储带宽测试（MB/s）

测试编号	节点1	节点2	节点3	节点4	节点5	节点6	节点7	节点8	节点9	节点10
1	800									
2	800	800								
3	800	800	800							
4	330	800	800	350						
5	320	770	730	360	800					
6	330	800	800	360	800	800				
7	320	800	800	350	330	780	300			
8	380	800	400	340	320	800	300	300		
9	340	800	360	380	330	280	320	270	620	
10	340	350	350	380	330	290	330	270	670	350

上表所示是在青云北京 2 区（PEK-2）进行存储带宽探测的结果。测试中使用了 10 台云主机，所有云主机都部署在同一个区域内。

首先，我们使用不同的云主机实例类型挂载单块云硬盘进行测试，发现存储带宽上限为 200MB/s。这个存储带宽上限既与云硬盘的容量无关，也与云主机实例类型无关。

其次，我们在同一台云主机上挂载多块云硬盘创建 RAID0 磁盘阵列进行同样测试。与单块云硬盘相比，用两块云硬盘创建的 RAID0 磁盘阵列可以获得 400MB/s 的存储带宽。用三块或者四块云硬盘创建的 RAID0 磁盘阵列，则可以获得 600MB/s 和 800MB/s 的存储带宽。

考虑到批量测试的费用问题，如上测试使用的云主机实例类型为“超高性能主机”，配置 1 颗 vCPU 和 1GB 内存，挂载四块 50GB 的云硬盘配置成 RAID0 磁盘阵列。在参与测试的 10 台云主机中，1 号云主机在一个用户帐号中，2～10 号云主机在另外一个用户帐号中。在编号为 1 的测试中，只有 1 号云主机产生存储流量，其他云主机处于空闲状态；在编号为 2 的测试中，1 号和 2 号对云主机产生存储流量，其他云主机处于空闲状态，以此类推。如上测试在一个月中的不同时段进行了多次，不同批次的测试结果之间基本一致。基于如上测试，可以认为青云的存储质量相对较低，具体表现在：

1. 没有对云主机或者云硬盘采取限流措施，吞吐量指标存在大规模抖动；
2. 在小规模测试中，仅用 4 台云主机即可探测到存储性能恶化的迹象；
3. 随着参与测试的云主机数量的增加，存储性能恶化极快；
4. 单个用户可以使用的存储带宽上限为 4000MB/s；
5. 在小规模测试中，可以观察到单个用户大量占用存储带宽对其他用户使用存储产生影响。

在针对客服能力的测试中，作者通过青云 Web 控制台里提交了多个工单。所有工单的响应时间均在 30 分钟以内。**不同工单分别涉及配额上调、使用方法、缺陷报告等等内容，处理所需要的时间也有不同。所有工单咨询的问题最终都得到很好的解决。**

盛大云

盛大云启用的公网 IP 地址有 3 个 B 段。通过 ip-tracker.org 进行查询，未能确认这些 IP 地址属于盛大云。

2016 年 3 月，从公网对全部 3 个 B 类 IP 地址段针对 22 (SSH) 和 3389 (RDP) 端口进行扫描，有 6,000 个 IP 地址可以通过 22 端口创建网络连接，有 4,000 个 IP 地址可以通过 3389 端口创建网络连接。

2016 年 9 月，从公网对全部 4 个 B 类 IP 地址段针对 22 (SSH)、80 (HTTP)、443 (HTTPS) 和 3389 (RDP) 端口进行扫描。有 5500 个 IP 地址可以通过 22 端口创建网络连接，有 36000 个 IP 地址可以通过 80 端口创建网络连接，有 3600 个 IP 地址可以通过 443 端口创建网络连接，有 3,200 个 IP 地址可以通过 3389 端口创建网络连接。

由于盛大云的规模相对较小，作者未对 1433、3306 和 11211 等端口进行扫描和自动化登录测试。基于同样的原因，作者也未对盛大云进行网络、存储、客服等方面的测试。

UCloud

UCloud 启用的公网 IP 地址有 8 个 B 段。通过 ip-tracker.org 进行查询，仅有一个 B 段可以确认属于 UCloud。

2016 年 3 月，从公网对全部 8 个 B 类 IP 地址段针对 22 (SSH) 和 3389 (RDP) 端口进行扫描，有 24,000 个 IP 地址可以通过 22 端口创建网络连接，有 9,000 个 IP 地址可以通过 3389 端口创建网络连接。**UCloud 给每个用户均缺省地提供了一个私有网络，用户在自己的私有网络上创建云主机，无法扫描到不属于用户自己的云主机。**

2016 年 9 月，从公网对全部 8 个 B 类 IP 地址段针对 22 (SSH)、80 (HTTP)、443 (HTTPS) 和 3389 (RDP) 端口进行扫描。有 43,100 个 IP 地址可以通过 22 端口创建网络连接，有 54,200 个 IP 地址可以通过 80 端口创建网络连接，有 27,500 个 IP 地址可以通过 443 端口创建网络连接，有 22,800 个 IP 地址可以通过 3389 端口创建网络连接。

UCloud 给每个用户均缺省地提供了一个私有网络，用户在自己的私有网络上创建云主机，无法扫描到不属于用户自己的云主机。由于 UCloud 的规模相对较小（相对于阿里云而言），作者未对 1433、3306 和 11211 等端口进行扫描和自动化登录测试。

UCloud 内网带宽测试（MB/s）

测试编号	节点1	节点2	节点3	节点4	节点5	节点6	节点7	节点8	节点9	节点10	节点11	节点12	节点13	节点14
1	175	175												
2	175	175	175	170										
3	175	175	170	175	175	165								
4	170	165	165	175	165	175	165	165						
5	175	175	175	175	175	165	165	165						
6	175	175	175	130	175	180	170	160	165	170	180			
7	175	175	165	180	190	175	185	155	150	190	170	190	190	150

上表所示是在 UCloud 北京 D 区（PEK-D）进行网络带宽探测的结果。测试中使用了 7 对云主机，所有云主机都部署在同一个区域内。我们首先使用同样的测试程序对不同的云主机实例类型进行测试，发现不同的云主机实例类型所能够达到的内网带宽是一样的。考虑到批量测试的费用问题，如上测试使用的云主机实例类型为“SSD 高性能主机”，配置 1 颗 vCPU 和 2GB 内存。在参与测试的 14 台云主机中，1 号云主机在一个用户帐号中，2～14 号云主机在另外一个用户帐号中。1 号云主机和 2 号云主机配为一对，3 号云主机和 4 号云主机配为一对，以此类推。在编号为 1 的测试中，只有第一对云主机产生网络流量，其他云主机处于空闲状态；在编号为 2 的测试中，第一对和第二对云主机产生网络流量，其他云主机处于空闲状态，以此类推。如上测试在一个月中的不同时段进行了多次，不同批次的测试结果之间基本一致。基于如上测试，可以认为 UCloud 云的网络质量相对较好，具体表现在：

- 1、似乎对云主机采取了限流措施，但是吞吐量指标存在一定范围的抖动；

- 2、在小规模测试中，未探测到网络性能明显恶化的迹象；
- 3、在小规模测试中，未观察到单个用户大量占用网络带宽对其他用户使用网络产生影响。

基于现象（1），作者倾向于认为 UCloud 尚未实现以云主机为单位进行精准限流。在小规模测试中观察到现象（2）和（3）的根本原因是因为 UCloud 的规模较大（相对于青云而言），其网络资源总量足以消化小规模测试所产生的网络流量。

UCloud 存储带宽测试（MB/s）

测试编号	节点1	节点2	节点3	节点4	节点5	节点6	节点7	节点8	节点9	节点10
1	280									
2	280	260								
3	200	260	240							
4	200	250	210	220						
5	200	230	250	200	200					
6	200	200	210	230	160	180				
7	140	220	140	140	140	170	140			
8	140	190	140	140	140	140	140	180		
9	140	180	140	140	140	140	140	140	140	
10	120	160	120	120	120	140	140	140	140	120

上表所示是在 UCloud 北京 C 区（PEK-C）进行存储带宽探测的结果。测试中使用了 10 台云主机，所有云主机都部署在同一个区域内。

首先，我们使用不同的云主机实例类型挂载单块云硬盘进行测试，发现存储带宽上限在 100MB/s 上下波动，但是并不稳定。这个存储带宽上限既与云硬盘的容量无关，也与云主机实例类型无关。

其次，我们在同一台云主机上挂载多块云硬盘创建 RAID0 磁盘阵列进行同样测试。与单块云硬盘相比，用两块云硬盘创建的 RAID0 磁盘阵列可以获得 200MB/s 的存储带宽。用三块或者四块云硬盘创建的 RAID0 磁盘阵列，则可以获得 300MB/s 和 400MB/s 的存储带宽。用六块云硬盘创建的 RAID0 磁盘阵列，最高可以获得 580MB/s 的存储带宽，但是均值只有 400MB/s。

考虑到批量测试的费用问题，如上测试使用的云主机实例类型为“SSD 高性能主机”，配置 1 颗 vCPU 和 2GB 内存，挂载四块 50GB 的云硬盘配置成 RAID0 磁盘阵列。在参与测试的 10 台云主机中，1 号云主机在一个用户帐号中，2 ~ 10 号云主机在另外一个用户帐号中。在编号为 1 的测试中，只有 1 号云主机产生存储流量，其他云主机处于空闲状态；在编号为 2 的测试中，1 号和 2 号对云主机产生存储流量，其他云主机处于空闲状态，以此类推。如上测试在一个月中的不同时段进行了多次，不同批次的测试数据存在较大差别，但是所观察到的现象基本一致。基于如上测试，可以认为 UCloud 的存储质量相对较低，具体表现在：

1. 没有对云主机或者云硬盘采取限流措施，吞吐量指标存在大规模抖动；
2. 在小规模测试中，仅用 4 台云主机即可探测到存储性能恶化的迹象；
3. 随着参与测试的云主机数量的增加，存储性能恶化极快；
4. 在小规模测试中，可以观察到单个用户大量占用存储带宽对其他用户使用存储产生影响；
5. 在被测试区域中可能存在其他用户运行的磁盘 I/O 密集型应用，并且其磁盘 I/O 资源使用模式随时间发生变化。

注：在针对青云的测试中，并未观察到同类现象。青云的可观测规模不足 UCloud 的 1/3，如果青云上存在其他用户运行的磁盘 I/O 密集型应用，应该比 UCloud 更容易观察到。因此，作者倾向于认为在青云的被测试区域中存在其他用户运行的磁盘 I/O 密集型应用的可能性较小。

针对存储带宽的测试结果，也从侧面验证了作者在网络带宽测试中所做的判断。UCloud 尚未实现以云主机为单位对网络和存储流量进行精准限流。在网络带宽测试中，UCloud 的网络资源总量足以消化小规模测试所产生的网络流量，从测试结果中仅能观察到网络带宽的波动。在存储带宽测试中，UCloud 的存储资源总量不足以消化小规模测试所产生的存储流量，从测试结果中可以观察到显著的性能恶化。

在针对客服能力的测试中，作者通过 UCloud 的 Web 控制台里提交了多个

工单。所有工单的响应时间均在 30 分钟以内。不同工单分别涉及配额上调、使用方法、缺陷报告等等内容，处理所需要的时间也有不同。工单所咨询的问题，大部分得到很好的解决，小部分工单没有得到解决。

腾讯云

腾讯集团在自治域 AS45090、AS132203、AS132591、AS134103 中共声明了 12 个 B 类 IP 地址段以及多个 C 类 IP 地址段。

2016 年 3 月，从公网对全部 12 个 B 类 IP 地址段针对 22（SSH）和 3389（RDP）端口进行扫描，有 40,000 个 IP 地址可以通过 22 端口创建网络连接，有 40,000 万个 IP 地址可以通过 3389 端口创建网络连接。

2016 年 9 月，从公网对全部 12 个 B 类 IP 地址段针对 22（SSH）、80（HTTP）、443（HTTPS）和 3389（RDP）端口进行扫描。有 65000 个 IP 地址可以通过 22 端口创建网络连接，有 90,000 个 IP 地址可以通过 80 端口创建网络连接，有 20,000 个 IP 地址可以通过 443 端口创建网络连接，有 50,000 个 IP 地址可以通过 3389 端口创建网络连接。

需要说明的是，如上所述 12 个 B 类 IP 地址段并非全部用于腾讯云的公有云服务。腾讯集团下的其他业务譬如 QQ 和微信所使用的 IP 地址也都在这 12 个 B 类 IP 地址段中。根据华为集团 2014 年发布的成功案例《华为服务器助力腾讯构建十万级高效部署》^[21]的成功案例，腾讯集团现网服务器超过 30 万台，其中华为服务器超过 10 万台。假设腾讯集团所使用的服务器当中只有 10% 配置公网 IP，需要占用的 IP 地址数量就超过 3 万个。考虑到腾讯集团本身对计算资源的需求还在增长，同时也会占用更多的 IP 地址。因此，腾讯云的公有云服务所使用的 IP 地址（包括物理机和虚拟机）只占如上所述活跃 IP 地址中的一小部分。

由于腾讯云的规模相对较小，作者未对 1433、3306 和 11211 等端口进行扫描和自动化登录测试。基于同样的原因，作者也未对腾讯云进行网络、存储、客服等方面的测试。

腾讯云使用 QQ 的帐号管理体系，可能是腾讯云用户最大的风险之一。众所

周知，QQ 用户在密码丢失、手机停机、切换地理位置的时候，均有 QQ 号码被腾讯集团回收的可能。作者于 2000 年成为 QQ 早期用户，十多年如一日地使用同一个 QQ 号码。在此期间，作者从美国伊利诺州移居加州，又从加州移居北京，再从北京移居海南，从未放弃过使用该 QQ 号码与亲朋好友进行联系。2014 年 2 月，作者从海南移居悉尼，QQ 以登录地理位置可疑为由拒绝作者登录。由于作者居住在北京时向腾讯登记的密码保护手机号码已经停用，作者选择通过早期好友确认的方式找回 QQ 号码。尽管所有三位早期好友均向腾讯作出了确认，腾讯方面依然拒绝作者继续使用该 QQ 号码。作者先前设置了 QQ 邮件转发，尽管作者已经不再拥有该 QQ 号码的使用权，但是发向该 QQ 号码的电子邮件依然被转发到作者的常用邮箱里。假设一家创业公司选择在腾讯云部署服务，而其所使用的 QQ 号码由于某种已知或者未知的原因被腾讯回收，必定会对其业务产生不可知的重大影响。**创业者最不愿意看到的情形之一，可能是你所提供的服务还在正常运行，但是你已经不再拥有运行这些服务的计算资源的使用权和管理权了。**

2015年中国公有云服务发展报告

其他讨论

弹性计算

弹性计算的核心，是负载均衡与自动伸缩的有机结合。负载均衡这个概念出现得比较早，在整个 IT 行业都已经被广泛接受和广泛应用。本文中所讨论的几家公有云服务提供商，基本上都提供了负载均衡的功能或者特性。自动伸缩则是云计算“按需获取、按量计费”理念的具体实现，最早的实现是 AWS 针对其 EC2 服务所提供的 AutoScaling Group (ASG) 功能。本文中讨论的几家公有云服务提供商，只有阿里云（2014 年 9 月）、青云（2015 年 3 月）和 UCloud（2016 年 6 月）提供了类似于 ASG 的自动伸缩功能。

对于一个正常的 Web 应用，其负载通常可以划分成三个档次：长期平均负载，长期高峰负载，短期爆发负载。在每秒只有数百个请求的情况下，云主机集群具备每秒处理一万个请求的能力是没有必要的。在每秒达到数万个请求的情况下，云主机集群只有每秒处理一万个请求的能力是远远不够的。自动伸缩的目的，就是在应用负载降低时自动将多余的云主机从负载均衡上移除并销毁以节省成本，在应用负载升高时自动启动更多的云主机并加入负载均衡以应对压力。通过自动



伸缩，用户自动地按照实际负载购买计算资源，既不存在处理能力不足的问题，也不存在浪费计算资源的问题。

显而易见，自动伸缩要求云主机集群中的每一台云主机都能够稳定地提供一定的处理能力。当云主机数量增加时，集群处理能力随之增加；当云主机数量减少时，集群处理能力随之减少。集群处理能力与云主机数量之间的关系不一定是线性的，但必须是正相关的。在理想的情况下，这种关系应该是可预测的。假设我们有一个网络 I/O 密集型应用，每处理 1 万个请求会产生 100MB 的内网流量，但是对 CPU、内存、存储的要求不高。当应用的负载为每秒 1 万个请求时，要求内网带宽大于 100MB/s。在阿里云上应对这样的负载，要求在云主机集群中部署 2 台云主机。在青云上应对这样的负载，要求云主机集群中部署 1 台云主机。当应用的负载为每秒 10 万个请求时，要求内网吞吐量大于 1,000MB/s。在阿里云应对这样的负载，需要在云主机集群中部署 17 台云主机。在青云上应对这样的负载，不管在云主机集群中部署多少台云主机都无能为力，因为需要的内网带宽超出了单个用户所能够使用的带宽上限。

因此，在弹性计算这个场景中，用户需要了解的并不是某个产品最高可以达到什么性能，而是最低可以达到什么性能。在云主机网络带宽测试中，阿里云两台云主机之间的网络带宽只有 60MB/s，青云两台云主机之间的网络带宽达到 115MB/s。看起来似乎青云的网络性能要好得多，但是阿里云的网络性能是不随

着用户使用量的增加而发生恶化的，青云的网络性能则是随着用户使用量的增加而发生恶化的。在云主机存储带宽测试中，阿里云单台云主机的存储带宽只有 400MB/s，青云单台云主机的存储带宽达到 800MB/s。看起来似乎青云的存储性能要好得多，但是阿里云的存储性能是不随着用户使用量的增加而发生恶化的，青云的存储性能则是随着用户使用量的增加而发生恶化的。

在云计算中，我们常常用坏邻居（noisy neighbours）效应来形容这种那个情形。一台云主机大量使用某种资源造成另一台云主机无法正常使用同种资源，属于坏邻居效应；一个用户大量使用某种资源造成另一个用户无法正常使用某种资源，也属于坏邻居效应。在针对阿里云的测试中，我们没有观察到的坏邻居效应。在针对青云的测试中，同一用户的多台云主机之间存在坏邻居效应，不同用户的云主机之间也存在坏邻居效应。在青云上，通过小规模测试即可在网络和存储两个方面观察到坏邻居效应。基于如上分析，青云虽然提供了自动伸缩的功能，但是其弹性计算能力并不能满足运营一个小规模（10 台云主机以下）高网络 I/O 或者高存储 I/O 型应用的要求。

基于同样的分析，我们在 UCloud 也观察到了如上所述的坏邻居效应。在云主机网络带宽测试中，可以观察到内网带宽存在一定程度的抖动，但是并未观察到明显的性能恶化。在云主机存储带宽测试中，则可以观察到存储带宽随着用户使用量的增加而不断发生恶化。根据作者的判断，UCloud 所提供的自动伸缩功能，可以满足运营一个中等规模（10 到 20 台云主机）高网络 I/O 型应用的要求，但是无法满足运营一个小规模（10 台云主机以下）高存储 I/O 型应用的要求。

不难看出，实现实用意义上的弹性计算要求公有云服务提供商具备两个技术条件。第一，通过细颗粒度的精准限流，保障任意计算资源（CPU、内存、网络、存储）的性能。阿里云的块存储已经具备了细颗粒度的精准限流，其 I/O 能力直接与块存储的容量挂钩，用户可以通过一个简单的公式计算出特定块存储的性能期望值。需要指出的是，阿里云仅对块存储进行了精准限流，但是未对网络 I/O 进行分级限流。在作者的测试中，最高配置的云主机和最低配置的云主机可以使用的网络带宽是相同的。因此，具备一定财务能力的用户依然可以在“合理使用”

的原则下以很低的代价来探测阿里云的内网带宽上限。第二，通过储备大量的冗余资源，应对用户爆发性的资源需求。在这个方面，规模较大的公有云服务提供商比规模较小的公有云服务提供商具有明显的优势。在国内的公有云服务提供商中，只有阿里云具备了这个条件。

阿里云给所有配置的云主机提供相同的网络带宽，并不符合“按需获取、按量计费”的理念。目前阿里云所提供的内网带宽可能可以满足中小型 Web 应用的需求，但是尚远远不能满足大型 Web 应用和科学计算应用的需求。作者曾经试图在阿里云上运行一些对网络 I/O 和磁盘 I/O 同时有较高需求的科学计算应用，但是由于网络 I/O 方面的限制未能取得预期的效果。作者注意到阿里云团队在 2015 年 10 月高调发布了新的 100TB 数据排序世界记录^[13]，但是这个世界纪录不是在阿里云上获得的，而是阿里云团队在物理机上获得的。出于好奇，作者将阿里云团队于 2015 年取得的进展与 Spark 团队于 2014 年报告的成绩^[14]（基于 AWS EC2 服务，使用 i2.8xlarge 实例）以及加州大学圣地亚哥分校（UC San Diego, UCSD）和 Google 团队于 2014 年报告的成绩^[15]（基于 AWS EC2 服务，使用 i2.8xlarge 实例）进行了对比。

100TB 排序	UCSD (2014)	Spark (2014)	Aliyun (2015)
网络类型	10 Gbps, 虚拟网络	10 Gbps, 虚拟网络	10 Gbps, 物理网络
网络配置	EC2 placement group	EC2 placement group	3:1 subscription ratio
主机类型	EC2 i2.8xlarge	EC2 i2.8xlarge	物理机
单机配置	Intel Xeon E5-2670 v2 @2.50GHz, 单机 32 个虚拟核心, 244GB 内存, 8x800GB SSD 硬盘	Intel Xeon E5-2670 v2 @2.50GHz, 单机 32 个虚拟核心, 244GB 内存, 8x800GB SSD 硬盘	3134 台: 双路 Intel Xeon E5-2630 @2.30 GHz (单机 12 个物理核心), 96GB 内存, 12x2TB SATA 硬盘; 243 台: 双路 Intel Xeon E5-2650 v2 @2.60GHz (单机 16 个物理核心), 128GB 内存, 12x2TB SATA 硬盘
操作系统	Amazon Linux 2014.03	Amazon Linux 2014.03	RHEL Server 5.7
主机总数	186	206	3,377
CPU 核心总数	5,952 个虚拟核心	6,592 个虚拟核心	41,496 个物理核心
内存总数 (GB)	45,384	50,264	331,968
排序时间 (秒)	1,378	1,406	377
单颗 CPU 核心每秒所处理的数据量 (MB)	12.19	10.99	6.39
单节点每秒所处理的数据量 (MB)	390	351.77	78.55

作者发现，同样是 100TB 数据排序，Spark 团队使用了 50TB 内存，UCSD 团队使用了 45TB 内存，阿里云使用了 332TB 内存。打个极度简化的比喻，Spark 团队和 UCSD 团队需要 5 个手指头来做 0 到 9 这十个数字的排序，而阿里云团队需要 33 个手指头来做 0 到 9 这十个数字的排序。

学习过算法的从业人员都知道，在内存富余的情况下进行数据排序，其算法复杂度远低于在内存不足的情况下进行数据排序。在内存不足的情况下，需要先将中间数据写入磁盘，再分批从磁盘读出进行排序操作。在 100TB 数据排序中，内存不足的情况产生 100TB 的网络 I/O 以及 200TB 的磁盘 I/O（读和写操作都是 200TB），内存富余的情况则产生 100TB 的网络 I/O 和 100TB 的磁盘 I/O（读和写操作都是 100TB）。在这种情况下，Spark 团队单颗 CPU 核心平均每秒钟处理的数据量（以 MB 计算）是阿里云团队的 1.72 倍，UCSD 团队单颗 CPU 核心平均每秒钟处理的数据量（以 MB 计算）是阿里云团队的 1.91 倍。我们也注意到阿里云团队所使用的处理器相对较老，其处理能力相对较低。就裸机单线程处理能力而言，Intel Xeon E5-2670 v2 @2.50GHz^[18] 是 Intel Xeon E5-2630 @2.30GHz^[16] 的 2.5 倍，是 Intel Xeon E5-2650 v2 @2.60GHz^[17] 的 1.7 倍。此外，我们也注意到 EC2 实例使用的是虚拟处理器核心，存在一定的虚拟化损失。

基于如上两点，可以推测如果阿里云同样采用 Intel Xeon E5-2670 v2 @2.50GHz 处理器的话，则其单颗 CPU 核心平均每秒钟处理的数据量（以 MB 计算）表面上看来可以与 Spark 团队和 UCSD 团队所报告的成绩持平。问题在于，Spark 团队和 UCSD 团队所处理的是一个比较复杂的场景（数据量是内存的 2 倍），其算法复杂度较高；而阿里云所处理的是一个比较简单的场景（内存是数据量的 3.3 倍），其算法复杂度较低。因此，阿里云所报告的成绩只是集群规模增长的自然结果，其系统和算法远远不如 Spark 团队和 UCSD 团队的系统和算法。（阿里云所处理的场景与 Yahoo 团队于 2013 年所处理的场景^[19] 是类似的，并且其系统和算法优于 Yahoo 团队的系统和算法。）

阿里云所报告的 100TB 数据排序成绩，证明的是阿里云作为一家公司已经具备了管理和使用超大型集群的能力。这个能力与阿里云于 2014 年在 VLDB 上发布

的关于伏羲的论文^[20]所报告的能力是类似的。但是，这个能力与阿里云所提供的产品和服务的能力并不直接相关。譬如说，阿里云的 ECS（云主机）服务所提供的计算、网络 and 存储能力，尚远远不足以满足运行这个 100TB 数据排序的要求。作者认为，阿里云尚需加强其 ECS 服务所提供的计算、网络 and 存储能力。什么时候阿里云的用户能够自地主在阿里云所提供的 ECS 服务上运行这个 100TB 数据排序并且达到 Spark 团队于 2014 年所取得的成绩，那么阿里云的 ECS 服务就真的是达到 AWS 的 EC2 服务在 2014 年的水平了。

区域间网络状况

为了了解各个公有云服务提供商的不同区域间的网络状况，我们在阿里云、青云和 UCloud 的华南和华北区域各启动云主机一台，并在两台云主机之间进行 MTR 测试。这部分测试数据，是在 2016 年 8 月底获得的。

下图所示为阿里云华南 1 区到华北 2 区的测试结果，数据包经过 11 跳抵达目标云主机。在这个路由当中，有 5 跳（4 ~ 8）经过的设备使用公网 IP 地址，其余 6 跳经过的设备使用阿里云内网 IP 地址。值得说明的是，下图所示的 4 个公网 IP 全部属于阿里云（属于自治域 AS37963），说明从阿里云华南 1 区到华北 2 区的整个数据链路层都在阿里云的掌控之下。

Start: Mon Aug 29 19:27:23 2016							
HOST:	Loss%	Snt	Last	Avg	Best	Wrst	StDev
1. -- ???	100.0	10	0.0	0.0	0.0	0.0	0.0
2. -- ???	100.0	10	0.0	0.0	0.0	0.0	0.0
3. -- 10.108.216.210	0.0%	10	1.5	1.1	0.8	2.8	0.6
4. -- 42.120.253.9	0.0%	10	3.2	2.8	2.0	3.5	0.0
5. -- 42.120.253.9	0.0%	10	24.3	25.4	24.3	26.0	0.3
6. -- 42.120.239.145	0.0%	10	27.2	26.7	25.8	27.6	0.3
7. -- 101.200.28.126	0.0%	10	50.6	51.2	50.3	52.0	0.0
8. -- 123.56.34.89	0.0%	10	51.6	49.3	49.0	51.6	0.7
9. -- 10.87.88.185	0.0%	10	48.5	49.4	48.4	53.0	1.6
10. -- 10.87.88.158	0.0%	10	50.2	52.2	49.5	67.7	5.7
11. --	0.0%	10	48.8	48.8	48.8	48.9	0.0

下图所示为青云北京 2 区到广东 1 区的测试结果，数据包经过 23 跳抵达目标云主机。在这个路由当中，有 16 跳（3 ~ 18）经过公网，其余 7 跳在青云自己的内网。在整个数据链路上，没有任何一个公网 IP 属于青云。因此，青云不同区域间的互联互通，完全依赖于数据中心服务提供商之间的互联互通。从这

个测试结果来看，青云“1跳进入全球任何网络运营商的主干网”这个目标可能还需要很长时间方可达成。

Start: Mon Aug 29 18:20:40 2016							
HOST:	Loss%	Snt	Last	Avg	Best	Wrst	StDev
1. --	0.0%	10	0.1	0.1	0.1	0.2	0.0
2. -- 10.16.40.241	0.0%	10	0.3	0.3	0.2	0.3	0.0
3. -- 119.254.97.1	0.0%	10	1.7	1.5	1.2	3.1	0.5
4. -- 119.254.106.217	0.0%	10	1.3	1.2	1.0	1.4	0.0
5. -- 119.253.6.93	0.0%	10	2.1	1.9	1.6	2.1	0.0
6. -- 119.253.6.57	0.0%	10	0.7	1.3	0.5	7.2	1.9
7. -- 119.253.5.21	0.0%	10	0.9	0.9	0.8	1.0	0.0
8. -- 119.253.3.237	0.0%	10	0.9	1.0	0.9	1.1	0.0
9. -- 119.57.24.58	0.0%	10	1.0	1.6	1.0	6.9	1.8
10. -- ???	100.0	10	0.0	0.0	0.0	0.0	0.0
11. -- 106.39.117.33	20.0%	10	3.5	3.8	2.8	6.5	1.1
12. -- 106.39.117.33	0.0%	10	3.3	3.2	2.8	3.8	0.0
13. -- 153.234.120.106.static.bj	20.0%	10	3.2	3.2	2.9	3.9	0.0
14. -- 202.97.57.225	0.0%	10	3.6	3.8	3.5	4.4	0.0
15. -- 202.97.65.210	0.0%	10	37.7	37.5	35.8	38.6	0.7
16. -- 183.61.223.113	0.0%	10	35.8	35.8	35.6	36.0	0.0
17. -- 183.61.223.162	0.0%	10	39.0	38.2	37.2	44.3	2.2
18. -- 183.2.241.206	30.0%	10	39.1	61.4	38.8	185.0	54.5
19. -- ???	100.0	10	0.0	0.0	0.0	0.0	0.0
20. -- ???	100.0	10	0.0	0.0	0.0	0.0	0.0
21. -- ???	100.0	10	0.0	0.0	0.0	0.0	0.0
22. -- ???	100.0	10	0.0	0.0	0.0	0.0	0.0
23. --	0.0%	10	40.6	40.7	40.6	41.1	0.0

下图为UCloud广东B区到北京C区的测试结果，数据包经过19跳抵达目标云主机。在这个路由当中，有8跳（6～13）经过公网，其余11跳在UCloud自己的内网。在整个数据链路上，没有任何一个公网IP属于UCloud。因此，UCloud不同区域间的互联互通，也完全依赖于数据中心服务提供商之间的互联互通。

Start: Mon Aug 29 18:55:31 2016							
HOST:	Loss%	Snt	Last	Avg	Best	Wrst	StDev
1. --	0.0%	10	0.1	0.1	0.1	0.2	0.0
2. -- 106.75.136.1	0.0%	10	0.5	0.5	0.5	0.6	0.0
3. -- 172.31.128.13	0.0%	10	1.1	7.0	1.0	60.0	18.6
4. -- 172.31.128.226	0.0%	10	1.1	1.1	1.0	1.2	0.0
5. -- 172.31.128.190	0.0%	10	4.6	7.2	4.2	18.3	4.0
6. -- 163.177.147.1	0.0%	10	1.9	1.9	1.8	2.0	0.0
7. -- 58.254.129.29	0.0%	10	5.3	4.8	3.1	7.0	1.1
8. -- 221.4.0.93	0.0%	10	3.0	3.1	3.0	3.4	0.0
9. -- 219.158.19.225	0.0%	10	38.8	39.1	36.5	42.7	2.4
10. -- 124.65.194.90	0.0%	10	33.7	34.5	33.4	35.6	0.8
11. -- 61.148.153.49	0.0%	10	33.5	35.7	33.5	37.2	1.1
12. -- 61.148.7.134	0.0%	10	34.0	33.5	33.2	34.5	0.3
13. -- 123.125.128.94	0.0%	10	35.5	38.5	35.3	52.3	5.2
14. -- ???	100.0	10	0.0	0.0	0.0	0.0	0.0
15. -- ???	100.0	10	0.0	0.0	0.0	0.0	0.0
16. -- ???	100.0	10	0.0	0.0	0.0	0.0	0.0
17. -- ???	100.0	10	0.0	0.0	0.0	0.0	0.0
18. -- ???	100.0	10	0.0	0.0	0.0	0.0	0.0
19. --	0.0%	10	33.9	33.9	33.8	34.1	0.0

从网络拓扑来看，国内互联网可以分为主干网（公网）、地区网（广域网）、主节点（城域网）几个层次。在计算系统可靠度的时候，又可以进一步将其简化成一个串行系统来处理。我们知道，串行系统的可靠度等于系统中各个组件的可靠度的乘积。串行系统中包含的组件越多，则整个系统的可靠性越低。假设一个串行系统中包含 100 个组件，每个组件的可靠度为 0.999，则整个系统的可靠度为 $0.999^{100}=0.905$ 。我们知道，**中国 ChinaNet 骨干网的拓扑结构在逻辑上分为两层，即核心层和大区层**。核心层由北京、上海、广州、沈阳、南京、武汉、成都、西安等 8 个城市的核心节点组成，其功能主要是提供与国际互联的接口以及大区之间互联的通路。全国 31 个省会城市按照行政区划，以上述 8 个核心节点为中心划分为 8 个大区网络。这 8 个大区网共同构成了大区层，大区之间通信必须经过核心层。

基于这样一个极度简化的模型，云服务提供商某个区域所在的数据中心与核心层之间的串行系统的组件数量，约等于如上所述 MTR 测试结果中数据包经过公网的跳数除以二。以此估算，阿里云的串行系统组件数量为 2，青云的串行系统组件数量为 8，UCloud 的串行系统组件数量为 4。假设这个串行系统中每个组件的可靠度是相同的（实际上并不相同），则阿里云数据链路的可靠度大于 UCloud，而 UCloud 数据链路的可靠度大于青云。

市场规模

根据 Gartner 的统计数据，在 2006 到 2014 年间，全球服务器硬件市场每年的出货量稳定地在 10,000,000 台上下波动。其中，亚太地区占比在 1/4 左右，也就是 2,500,000 台。中国境内服务器出货数量在亚太地区的占比不详，保守地按 1/5 计算也有 500,000 台。按照 5 年折旧周期进行估算，全国范围内现役的计算资源至少有 2,500,000 台物理服务器。（**根据可靠的数据来源，仅阿里巴巴、百度、腾讯三家大型互联网公司所拥有的物理服务器总量就已经超过了 1,000,000 台。全国范围内现役的计算资源总量可能远大于 2,500,000 这个估值。**）作为一家服务于“中国”的产业级别的公有云服务提供商，假设其业务成熟之后拥有全国计算资源的 1%，就是 25,000 台物理服务器。再按 1:4 的虚拟化比例进

行估算（这个虚拟化比例也是一个偏低的估值），则虚拟机的数量为 100,000 台左右。

公有云作为一种新型服务，其市场规模尚有相当程度的自然增长空间。因此，五年之后的公有云可能达到的规模只会比这个数字更大。基于这些估算，我们可以根据其规模将一家公有云创业企业的成长分为五个阶段。

- 概念阶段，小于5,000台虚拟机。公司的终极目标相对模糊，在私有云解决方案提供商和公有云服务提供商之间摇摆不定。在战术层面，缺乏明确的技术路线图，产品形态相对原始并且没有明确的技术指标。
- 原型阶段，小于10,000台虚拟机。公司基本上将其终极目标定位为公有云服务提供商。由于公有云和私有云之间的巨大差异，必然要放弃私有云解决方案服务提供商的身份。在战术层面，基本形成相对清晰的技术路线图，基础产品（云主机）基本定型，在宕机时间和产品性能方面均有明确的技术指标。在云主机的基础上，提供能够承担中低负载的负载均衡、数据库、缓存等等周边产品。
- 成长阶段，小于50,000台虚拟机。基础产品（云主机）能够满足高性能计算的要求，同时发展出一系列模块化的周边产品。普通用户完全依靠云服务提供商所提供的不同模块即可自主创建大规模可伸缩型应用（无需云服务提供商进行干预）。
- 成熟阶段，小于100,000台虚拟机。在技术方面，资源利用率开始提高，规模效应开始出现。在市场方面，客户忠诚度开始提高，马太效应开始出现。这标志着公司在公有云领域已经获得了较有份量的市场份额，其产品和技术获得了一个或者多个细分市场的广泛认可。
- 产业阶段，大于100,000台虚拟机。只有进入这一阶段，才能够认为一个服务提供商已经站稳了脚跟，可以把公有云当作一个产业来做了。至于最后能够做多大，一个好看国内的大环境，另外一个还得看公司自身的发展策略。

根据本文的测试数据，作者认为国内公有云服务提供商中（仅考虑以设施服

务 IaaS 部分) 只有美团云尚处于原型阶段, 盛大云处于成长阶段早期, 青云和腾讯云处于成长阶段中期, UCloud 处于成熟中晚期, 只有阿里云已经进入产业阶段。由于作者缺乏金山云相关数据, 在此不对金山云进行判断。从 2016 年 9 月和 2016 年 3 月的端口扫描结果差异来看, 马太效应在国内公有云服务行业内是非常显著的。

基于本文所述的网络规模探测数据, 作者倾向于认为目前国内公有云市场所提供的云主机总量小于 500,000 台。考虑到公有云的早期用户大都是将云主机当作 VPS 来使用, 并且偏向于使用配置较低的小型云主机实例, 平均意义上的虚拟化比例可能远远高于本文前面提到的 1:4。因此, 作者倾向于认为**目前国内公有云市场所提供的计算能力尚低于国内现有计算能力总量的 2%**。随着公有云服务的逐步成熟, 企业采购计算资源时的偏好将逐渐从采购物理服务器或者使用数据中心托管服务过渡到使用公有云服务。未来五到十年里, 各种运行在中小型数据中心内的负载还会稳步向大型公有云服务提供商迁移。在这种大趋势的影响下, 公有云市场所提供的计算能力在全国计算能力总量中所占的比例也会稳步提高。

公有云服务的未来还远远不止于此。作者于 2012 年在《虚拟化、云计算、开放源代码及其他》^[6] 这篇博客里面提到了英国经济学家威廉杰文斯 (William Stanley Jevons, 1835-1882) 在《煤矿问题》 (The Coal Question) ^[5] 一书中指出的一个似乎自相矛盾的现象: 蒸汽机效率方面的进步提高了煤的能源转换率, 能源转换率的提高导致了能源价格降低, 能源价格的降低又进一步导致了煤消费量的增加。这种现象称为杰文斯悖论, 其核心思想是资源利用率的提高导致价格降低, 最终会增加资源的使用量。在过去 150 年当中, 杰文斯悖论在主要的工业原料、交通、能源、食品工业等多个领域都得到了实证。基于同样的原理, 作者在这篇博客里面又进一步断言: “公共云计算服务的核心价值, 是将服务器、存储、网络等等硬件设备从自行采购的固定资产变成了按量计费的公共资源。虚拟化技术提高了计算资源的利用率, 导致了计算资源价格的降低, 最终会增加计算资源的使用量。”。

因此, 作者在多篇文章 ^[6, 7, 8, 9] 中反复指出, 公有云市场不是一个短期市场,

而是一个未来五到十年尚有充分增长空间的市场。目前国内的公有云市场依然处于早期阶段，所谓“随着国内云计算市场逐步成熟，公有云市场份额基本上大势已定”的论调未免过于目光短浅。在国内现有的几家公有云服务提供商中，除了阿里云可以算是已经站稳了脚跟之外，其他几个公有云服务提供商当中的任何一个都有可能突然倒塌，还可能新的公有云服务提供商脱颖而出。作者倾向于认为，至少要等到2018年才有可能发生中国公有云市场份额“大势已定”的情况。这也是为何作者一直强调公有云服务是一片刚刚显现的蓝海。现在国内各个公有云服务提供商杀得你死我活，看起来似乎已经是一片血海。在作者看来，这些不过都是假象。如果一家公有云领域的创业企业没有这样的大局观，那么作者只有一个建议：“**认怂服输，割肉止损，是为美德。**”

用户习惯

根据作者的观察，目前国内大部分公有云用户还是把云主机当作传统物理服务器的替代品来使用。这个观察在作者与各个公有云服务提供商负责人的访谈中也得到了验证。

在传统的IT架构中，操作系统是安装在物理服务器上的。由于重新安装操作系统需要造成很长的宕机时间，出现软件层面故障时运维或者开发人员往往倾向于寻找问题来源并予以排除。很多时候，运维或者开发人员需要花费很长时间来寻找一些不易发觉的输入或者拼写错误（例如四个空格和一个tab）。在弹性计算这个场景中，操作系统是通过映像模板创建的，获得一台全新的包含正确配置的云主机只需要数分钟甚至更短的时间。在这个时间优势的基础上，云计算服务提供商终于可以直面长期以来被传统IT服务提供商所刻意回避的两个事实。第一个事实是组件的失效是必然的，是不可避免的；第二个事实是组件的失效是随机的，是不可预测的。用AWS首席技术长官Werner Vogels的原话来说，就是“任何组件可在任何时刻失效”（Everything fails, all the time.）。在负载均衡与自动伸缩的帮助下，一个集群中任意云主机均可以在任意时刻由于任意原因（底层硬件、网络环境、操作系统、应用软件）发生失效。在一台云主机发生失效时，自动伸缩功能自动地将其从负载均衡上移除并进行销毁，同时自动地

启动一台新的云主机并加入负载均衡。因此，用户可以将云主机视为“即用即抛型”一次性资源。忽略云主机的失效，不仅不会牺牲应用服务质量，还可以将宝贵的资源集中投入到公司的关键业务。

可惜的是，**由于缺乏对弹性计算的理解，大量系统管理员延续了在使用物理服务器时期培养的习惯。**他们在云主机等计算资源失效时惊慌失措，并且热衷于寻找所谓的“根本原因分析”（root cause analysis）。他们在潜意识里还是将基础设施视为公司的资产，试图去了解和掌握云主机之下每一个层面的信息。他们没有意识到在弹性计算这个场景里这些努力不仅没有必要而且会阻碍整个公司技术进步。解决这个问题，需要所有的公有云服务提供商持之以恒地对用户进行教育。用户的认知水平提高了，也会进一步促进公有云市场的发展。

除了对用户进行教育之外，公有云服务提供商也需要加强对员工的教育。有些一线研发人员虽然是为公有云服务提供商工作，但是对弹性计算也缺乏理解。2016 年春节联欢晚会期间，新浪微博成功地使用阿里云承担了部分高峰负载。这本来是一个值得大书特书的成功案例，但是多位阿里云团队成员通过微博公开指责新浪微博“直到最后一刻才开通所需的计算资源”之举过于小气。虽然这样的指责近乎调侃性质，也不能不说是一个在 KPI 重压之下令人沮丧的失败案例。

安全问题

在“用户体验”这个小节的网络测试部分，作者仅报告了针对阿里云的 1433、3306 和 11211 端口扫描情况。主要的考虑在于小型公有云服务提供商更加需要保护，因此不宜对小型公有云服务提供商进行同类测试；次要的考虑在于公有云服务提供商无法独立承担安全重任，因此需要向从业人员披露公有云服务中存在的安全隐患。需要说明的是，作者在阿里云内网和外网获得的端口扫描结果，并非阿里云独有的现象。任何一个刚刚学会运行 bash 脚本的从业人员，都可以通过类似方法在 AWS 所在的 IP 段扫描到类似的结果。阿里云和 AWS 的不同之处，在于阿里云教育客户“选择云盾，让您的业务安全性如同阿里巴巴一般”，AWS 则教育客户“安全共担”（shared responsibility）模型。云盾的产品介绍页面宣称“每天防御超过 958 万次暴力破解攻击”，但是作者基于社会工程数

数据库的自动化登录测试也获得了部分成功。这样的结果，可以有三个解释：

- 大部分被成功登录的云主机没有使用云盾服务；
- 少部分被成功登录的云主机虽然使用了云盾服务但是其密码设置过于薄弱，因此在云盾未被激活之前就已经被成功登录；
- 类似于Memcached这样免登录的服务，云盾目前是完全无能为力的。在阿里云内网，作者还探测到大量其他免登录或者仅使用弱口令保护的网络安全服务，例如RabbitMQ。

因此，**不管用户是否使用服务提供商所提供的云安全服务，均应对客户进行“安全共担”的教育，引导客户采取必要的措施保护其所使用的计算资源。**遗憾的是，阿里云作为国内规模最大的公有云服务提供商，向用户传达了完全错误的信息。作者建议阿里云安全团队针对阿里云内网进行常规性的分布式端口扫描，充分了解安全隐患的严重程度，并在此基础上向阿里云的用户提出针对性的改进建议。

作者也注意到阿里云于 2014 年 7 月获得了首批数据中心联盟授予的“可信云服务认证”。公开的资料表明，可信云服务认证由数据中心联盟组织、中国信息通信研究院测试评估，是我国唯一针对云服务可信性的权威认证体系。（据其官方网站介绍，数据中心联盟是工业和信息化部通信发展司指导下成立、由中国通信标准化协会管理的非营利性第三方组织。）作者对阿里云进行的测试表明，这个“可信云服务认证”如同虚设，无法保护公有云用户免于弱口令暴力破解等入门级网络攻击威胁。更为严重的是，通过数据中心联盟这样一个半官方机构所谓的“权威认证体系”给不安全的公有云服务背书，引导安全观念原本就极为淡薄的用户更加忽视公有云服务中的安全问题，给互联网信息安全带来了极大的安全隐患。

对公有云服务在安全方面的顾虑，是阻止金融行业等极度保守型客户使用公有云服务的重要原因。例如，招商银行信息技术部副总经理、数据中心总经理高旭磊^[23]明确地认为“**金融云在可靠和可用性方面远高于普通商业云**”。高旭磊在文章中指出：“目前公有云的服务对象仍然是以互联网企业为主，但是运行关

键业务（Mission Critical）的企业非常少，服务质量和数据安全是这些企业踌躇不前的主要原因。仅 2014 年，大型的公有云就出现了多起严重的宕机影响在线服务的事件，其中亚马逊 AWS CloudFront DNS 服务器宕机超过 2 小时，导致一些网站和云服务的内容传输网络全部下线；Microsoft Azure 公有云出现了 4 次严重宕机事件，累计时长超过 20 小时，导致大量用户无法使用。” **尽管高旭磊的论据和论点之间缺乏必要的逻辑联系，又将可用性、可靠性、安全性等多个概念混为一谈，他的文章清楚地反映出金融行业客户对公有云服务在安全方面的不信任心态。**作者注意到阿里巴巴旗下蚂蚁金服集团于 2015 年 10 月推出了针对金融行业客户的蚂蚁金融云，并将安全作为其卖点之一。蚂蚁金融云宣称“基于阿里云计算强大的基础资源”，但是在阿里云内部称为专有云服务（Dedicated Cloud Service）^[24]，并不与阿里云的公有云服务使用同一套基础设施。换句话说，蚂蚁金服云直接否定了阿里云的公有云服务在安全方面的能力。

其他

众所周知，国内的互联网创业环境并不单纯。创业公司夸大创始人资历、融资额度、经营规模、性能参数等等数据，在整个大环境中几乎是家常便饭。在国内的公有云服务领域，类似的现象同样存在并且非常严重，主要集中在服务规模和产品性能两个方面。在撰写本报告的过程中，作者系统地阅读了过去一年中不同公有云服务提供商通过不同渠道发布的各种官方和非官方的规模和性能数据。在服务规模方面，大部分伪造数据者对国内信息行业的总体规模缺乏基本了解。在产品性能方面，大部分性能报告都巧妙地利用了 AWS 在 10 年前发布的“标准型”EBS 卷，通过田忌赛马的技巧将产品层面的对比降级为 SSD 硬盘与机械硬盘的对比，进而得出超越 AWS 的结论。值得注意的是，这样的现象同样不仅仅出现在中小型公有云服务提供商，阿里云这样的大型公有云服务提供商也不能免俗。阿里云级别很高的资深员工也在微博上公开传播第三方伪造的性能测试数据，但是在作者提醒后很快删除类似言论。

除此以外，作者认为国内大部分公有云服务提供商在做事上是还是踏实靠谱的。在与作者的电话访谈中，各位公有云服务提供商的负责人不仅坦诚地阐述创

业当中的决策过程，也坦诚地提出对行业发展的观点和疑问。尤其是七牛云的许式伟和 UCloud 的季昕华，与他们的访谈给作者留下了非常深刻的印象。他们所提供的产品和服务或许尚不完美，他们所报告的数据或许略有夸大，但是作者深深地感觉到他们在做事上是认真的，他们对客户是诚恳的。唯一的例外，却是青云。青云在国内云计算行业的名气很大，是一家有理想有情怀的创业公司。但是与青云 CEO 黄允松的访谈和对青云进行测试的结果都令作者对青云大失所望。作为一家创业企业，有理想有情怀不是坏事。但是，如果一味沉溺在理想和情怀当中不能自拔，甚至是将理想和情怀当成产品和服务来卖，恐怕对创业企业的发展并无益处。

信息披露和参考文献

信息披露

作者蒋清野是悉尼大学信息技术学院的博士研究生，同时也是 AWS 悉尼技术支持中心的员工。他于 1999 年获得清华大学学士学位（土木工程），2000 年获得伊利诺伊大学香槟分校硕士学位（土木工程），2015 年获得悉尼大学硕士学位（计算机科学）。他的研究兴趣包括分布式与高性能计算、开源社区的社会学行为、信息技术领域的微观经济学分析。他是美国电子电气工程师学会（IEEE）的高级会员。

在接受 InfoQ 方面的邀请准备规划这篇报告的时候，作者的内心是兴奋的。在获得所有测试数据准备撰写这篇报告的时候，作者的内心是矛盾的。一方面，作为并行与分布式计算领域的学生，作者希望为业界提供一些有用的信息和观点；另一方面，作为公有云服务领域的从业人员，作者深知发表一份涉及多家友商的报告会带来诸多争议。在 InfoQ 方面的鼓励下，作者选择以真实的身份发布这些的数据和观点，希望能够对国内云计算从业人员有所帮助。

参考文献

1. [工业和信息化部电信研究院](#)（中国信通院，CAICT），《[云计算白皮书（2014年）](#)》，2014年5月
2. 中国电子信息产业发展研究院工业和信息化部赛迪智库，《[云计算发展白皮书（2015版）](#)》，2015年4月
3. [工业和信息化部电信研究院](#)（中国信通院，CAICT），《[云计算白皮书（2016年）](#)》，2016年9月
4. Peter Mell and Timothy Grance, “[The NIST Definition of Cloud Computing](#)”, NIST Special Publication 800-145, September 2011
5. William Stanley Jevons, “[The Coal Question](#)”, 1866
6. 蒋清野，《[虚拟化、云计算、开放源代码及其他](#)》，2012年10月
7. 蒋清野，《[从王博士说起](#)》，2013年12月
8. 蒋清野，《[浅谈“中国”语境下的公有云发展](#)》，2015年5月
9. Qingye Jiang, Young Choon Lee, Albert Y. Zomaya, “[Price Elasticity in the Enterprise Computing Resource Market](#)”, IEEE Cloud Computing, vol.3, no. 1, pp. 24-31, Jan.-Feb. 2016, doi:10.1109/MCC.2016.14
10. 章文嵩，《[淘宝软件基础设施构建实践](#)》，[第三届中国云计算大会](#)，2011年5月
11. 高山渊，《阿里巴巴系统运维实践分享》，QClub深圳站，2012年6月
12. Netcraft, “[Aliyun cloud growth makes Alibaba largest hosting company in China](#)”, May 2015
13. Jiamang Wang, Yongjun Wu, Hua Cai, Zhipeng Tang, Zhiqiang Lv, Bin Lu, Yangyu Tao, Chao Li, Jingren Zhou, Hong Tang, “[FuxiSort](#)”, Alibaba Group Inc, October 2015
14. Reynold Xin, Parviz Deyhim, Ali Ghodsi, Xiangrui Meng, Matei Zaharia, “[GraySort on Apache Spark by Databricks](#)”,

Databricks Inc., October 2014

15. Michael Conley, Amin Vahdat, George Porter, “[TritonSort 2014](#)”, University of California at San Diego, 2014
16. Intel Xeon [E5-2630基准测试数据](#)
17. Intel Xeon [E5-2650基准测试数据](#)
18. Intel Xeon [E5-2670基准测试数据](#)
19. Thomas Graves, “[GraySort and MinuteSort at Yahoo on Hadoop 0.23](#)”, Yahoo!, May 2013
20. Zhuo Zhang, Chao Li, Yangyu Tao, Renyu Yang, Hong Tang, and Jie Xu. “[Fuxi: a fault-tolerant resource management and job scheduling system at internet scale.](#)” Proceedings of the VLDB Endowment 7, no. 13 (2014): 1393–1404.
21. 华为集团, 《[华为服务器助力腾讯构建十万级高效部署](#)》, 2014年7月
22. 陈海泉, 《[下一代超大规模软件定义网络技术实践](#)》, 2016年1月
23. 高旭磊, 《[招商银行关于金融云的思考](#)》, [中国金融电脑](#), 2016年8月
24. 阿里云, 《[阿里云生态路线图](#)》, 2015年7月

Geekbang>

极客邦科技

整合全球优质学习资源，帮助技术人和企业成长

InfoQ

技术媒体

EGO EXTRA GEEKS' ORGANIZATION
NETWORKS

职业社交

StuQ

职业教育