第三届国家期刊奖百种重点期刊 | 中国科技核心期刊
工信部优秀科技期刊 | 中国五大文献数据库收录期刊

ISSN 1009-6868
CN 34-1228/TN

中兴通讯技术

ZTE TECHNOLOGY JOURNAL

www.zte.com.cn/magazine

2012年12月 • 第6期

专题：云存储与云灾备

ISSN 1009-6868
12>
9 771009 686007

专题：云存储与云灾备

专|题|导|读

云计算概念自提出后，因其有效而清晰的商业模式受到广泛关注，并逐步得到工业界和学术界的认可，成为近几年来最受关注的IT技术之一。云计算是一种IT资源的交付和使用模式，可以将各种互联的计算能力、存储、数据、应用等资源进行有效整合，并根据需要以可扩展的方式提供用户所需的硬件、平台、软件服务等资源。提供资源的网络即被称为“云”，其计算能力通常是由分布式的大规模集群和服务器虚拟化软件搭建。云计算通过多层次的虚拟化与抽象，将用户从复杂细琐的底层硬件逻辑、网络协议、软件架构中解放出来。

各国政府对云计算给予了高度关注，纷纷出台战略规划和相关政策措施，加快推动云计算的发展和应用，力争抢占云计算发展的产业制高点。美国奥巴马政府于2009年9月正式启动云计算计划，旨在降低基础建设成本、减少政府信息系统对环境的影响；日本通过“加速重要研究开发”、“强化研究开发的基础系统”等预算项目支持云计算研发；韩国将云计算技术研发列为战略重点，投资3.49亿元人民币用于关键技术的研发。目前，各国际标准化组织正在积极开展云计算标准的研究制订工作，如美国国家标准和技术研究院(NIST)、云计算互操作论坛(CCIF)、云安全联盟(CSA)等。在中国，国家发改委、工信部、科技部等多部委共同参与的《关于加快发展高技术服务业的指导意见》指出：云计算产业将作为未来高技术服务业的主角。

虚拟化和云给灾备带来了巨大的挑战。云灾备作为灾备领域的一个新兴概念，它的出现为企业提供了一个行之有效的解决方案。云灾备是指将灾备看做一种服务，由客户付费使用灾备服务提供商提供的灾备服务的模式。采用这种模式，客户可以利用服务提供商的优势技术资源、丰富的灾备项目经验和成熟的运维管理流程，快速实现客户的灾备目标，降低客户的运维成本和工作强度，降低灾备系统的总体拥有成本。

本期专题一共收集8篇文章，分别涉及云灾备数据恢复技术、基于物联网环境的云存储及安全技术、云计算的应用、云存储的安全技术等问题的研究。希望文章有助于读者加深对云存储/云灾备的相关理论、技术、应用和发展趋势的了解。在此，也对各位作者的积极支持和辛勤工作表示最衷心的感谢。

杨义先

本期专题策划人

杨义先

灾备技术国家工程实验室主任，网络与信息攻防教育部重点实验室主任，北京邮电大学信息安全中心主任、博导，首批长江学者奖励计划特聘教授，首届国家杰出青年基金获得者；获得包括国家发明奖和省部级科技进步一等奖等在内的各类科技奖励20余项，授权发明专利4项，主持和参与多项国家“863”、国家自然科学基金、省部级等科研项目；在网络信息安全、现代密码学和纠错编码等领域取得众多学术研究，发表高水平论文500余篇，被SCI、EI、ISTP引用收录数百次，出版专著及教材20多部，其中包括中国密码学方面的第一部专著《编码密码学》。

2012年第1—6期专题计划

1 智能管道及其运营
续合元 电信研究院通信标准所总工

2 物联网与行业信息化
唐雄燕 中国联通国家工程实验室副总工

3 智能终端技术
糜正琨 南京邮电大学教授

4 数据中心网络关键技术
杜军朝 西安电子科技大学副教授
郭得科 国防科技大学副研究员

5 光与无线融合接入技术
陈建平 上海交通大学教授

6 云存储与云灾备
杨义先 北京邮电大学教授

目次

办刊宗旨

以人为本，荟萃通信技术领域精英；迎接挑战，把握世界通信技术动态；
立即行动，求解通信发展疑难课题；励精图治，促进民族信息产业崛起。

中兴通讯技术
ZHONGXING TONGXUN JISHU
双月刊 1995年创刊 总第107期
2012年12月 第18卷第6期 卷终

主管：安徽省科学技术厅
主办：中兴通讯股份有限公司
安徽省科学技术情报研究所
编辑：《中兴通讯技术》编辑部

总编：谢大雄
副总编：赵今明
常务副总编：黄新明
责任编缉：杨勤义
编辑：徐烨，卢丹，朱莉，Paul Sleswick
排版制作：余刚
发行：王萍萍
编务：王坤

《中兴通讯技术》编辑部
地址：合肥市金寨路329号凯旋大厦12楼
邮编：230061
网址：www.zte.com.cn/magazine
投稿平台：www.zte.com.cn/paper
电子信箱：magazine@zte.com.cn
电话：(0551)5533356
传真：(0551)5850139

出版、发行：中兴通讯技术杂志社
发行范围：全球发行
印刷：合肥中建彩色印刷厂
出版日期：2012年12月10日
刊号：ISSN 1009-6868 / CN 34-1228/TN
广告经营许可证：皖合工商广字0058
定价：每册10.00元，全年60.00元

专题：云存储与云灾备

专家视点

运营应用

研究论文

开发园地

系列讲座

综合信息

期刊基本参数：CN 34-1228/TN * 1995 * b * 16 * 64 * zh * P * ¥10.00 * 15000 * 14 * 2012-12

Contents

ZTE TECHNOLOGY JOURNAL Vol.18 No.6 Dec. 2012

Special Topic: Cloud Storage and Cloud Disaster Recovery

Expert View

Operational Application

Research Paper

Development Field

Lecture Series

云灾备的数据修复技术

Data Recovery Technology for Cloud Disaster Recovery

中图分类号：TP393 文献标志码：A 文章编号：1009-6868 (2012) 06-0001-004

李挥/LI Hui
侯韩旭/HOU Hanxu
黄显霞/HUANG Xianxia
（北京大学深圳研究生院，深圳 518055）
（Peking University Shenzhen Graduate School, Shenzhen 518055, China）

摘要： 基于纠错码的云灾备系统的关于降低修复带宽的最新研究成果，文章讨论了云灾备系统中存储节点失效的修复问题。文章指出数据修复问题有3种模型：精确修复、功能修复和系统部分精确修复。在精确修复中，失效的模块需要修复精确的丢失编码包；在功能修复中，新产生的编码包可以包含不同于丢失节点的数据，只要修复的云灾备系统支持最大可分离距离（MDS）码属性。系统部分精确修复是精确修复和部分修复的之间的一个混合的修复模型。

关键词： 云灾备；纠错码；网络编码；数据修复

Abstract: In this paper, we discuss latest research on cloud disaster recovery systems based on an erasure code that reduces repairing bandwidth. We also discuss problems with repairing in cloud disaster recovery systems when there is storage node failure. Three models for addressing the repairing problem are proposed: exact repair, functional repair, and exact repair of systematic parts. In exact repair, the failed blocks are exactly regenerated. In functional repair, the newly generated blocks can contain data different to that of the failed node as long as the repaired system maintains the MDS code property. Exact repair of the systematic part is a hybrid repair model and lies between exact repair and functional repair.

Keywords: cloud disaster recovery; erasure codes; network coding; information recovery

云计算平台通常给用户提供的是一个统一的接口。大量用户通过这个统一的接口对云计算平台进行访问。云计算平台是以服务为中心的，因而云计算平台必须保证提供服务平台的可靠性，云计算应该能够克服时常出现的普通计算机的断电、存储数据的失效等故障。处理这些故障的技术，我们称为云灾备技术。这也是云计算平台的稳定持续可用性要求。

云灾备系统设计中数据分布机制会对许多系统的具体实现产生影响。比如，为了保证数据存储的可靠性和系统的容错性，需要增加存储的冗余；为了提高系统的处理能力，需要将相同数据的副本分布在更多的节点上；为实现负载均衡需要使数据分布得更加合理。最常见的数据分布机制是复制，它通过增加存储空间和副本存储数量来实现系统的可靠性和容错性；为了在保证存储的可靠性和容错性的基础上进一步减少额外的存储空间消耗，纠错编码如Reed Solomon（R-S）编码逐渐在分布式存储环境中得到应用。

近些年来，在编码理论中新出现的网络编码改变了传统路由器中单纯复制转发的模式，证明在中间节点增加编码能力之后，网络可以达到拓扑对应的网络流图的最大流量。一般的网络编码实现时需要中心节点了解网络拓扑选择编码系数，而随机线性网络编码通过随机选择编码系数去除了这一限制，并选择相对较大的Galois域上进行编码，可以以很大的概率保证编解码的性能。针对分布式存储环境，随机线性网络编码可以代替纠错编码对数据进行编码分块。凭借其独特的编码方式，相信随机线性网络编码可以在减轻中心节点负载、负载均衡和节点修复等方面带来好处。

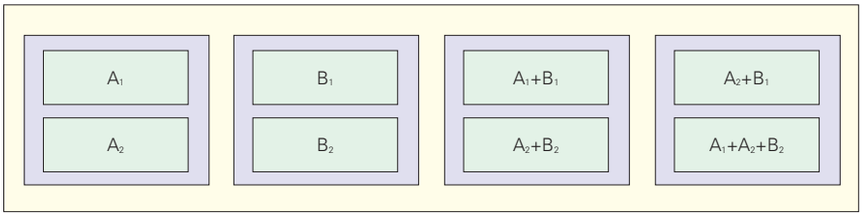
在云灾备系统必须引入冗余来提高节点失效时的可靠性。最简单且常用的冗余方式是直接在多个存储节点上进行数据复制。而对于相同的冗余度，纠错码技术较复制技术可成倍提高可靠性。为了实现编码增强可靠性，编码必须要支持一个纠错码形式。

给定两个正整数n,k（$n>k$），(n,k)最大可分离距离（MDS）编码的可靠性为：首先把数据均分为k个信息包，然后使用MDS码这k个信息包编码成n个数据包（相同大小），这样n个编码包中任k个均可修复原始数据。一个（4,2）MDS二进制纠错码如

基金项目：国家重点基础研究发展（“973”）规划（2012CB315904）；国家自然科学基金（61179028）

图1所示。

在图1中，每个存储节点存储两个信息包，都是原始数据A_1，A_2，B_1，B_2的二进制线性组合。在这个例子中，存储大小为4。在这n=4个存储节点中任意k=2个存储节点均可修复原始数据。



▲图1 一个(4,2)MDS二进制纠错码

MDS码是在冗余和可靠性方面最好的折衷方案，因为k个数据包包含了能够修复原始数据所需要的最少信息。在一个云灾备系统中，把n个编码数据包放置在不同存储节点，这样云灾备这可以容忍($n-k$)个节点失效而不丢失原始信息。纠错码的容错能力显著。R-S码可能是MDS码中最流行的。此外类似的信息扩散算法(IDA)也已经用于云灾备技术的研究中。喷泉码和低密度校验码(LDPC)是近几年的研究称成果，用来估计MDS属性和稳固快速的编解码。本文主要讨论云灾备系统中存储节点失效的修复问题[1-9]。

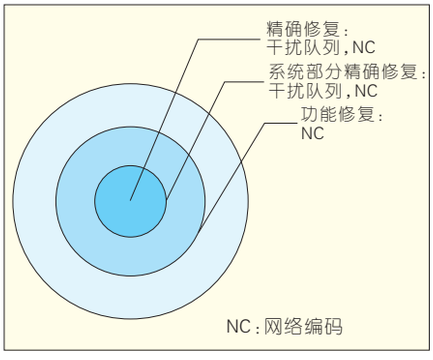
修复问题和相应的再生编码技术在文献[10]中有介绍。文献[11-20]中进一步研究了该问题。与对于纠错码，再生编码的最大优势就是其可以大量地减少修复带宽。

1 数据修复模型

修复模型通常分为3种：精确修复、功能修复和系统部分精确修复。在精确修复中，失效的模块需要修复精确的丢失编码包。在功能修复中，新产生的编码包可以包含不同于丢失节点的数据，只要修复的云灾备系统支持MDS码属性。系统部分精确修复是精确修复和部分修复的之间的一个混合的修复模型。在这个混合模型中，存储节点经常是一个系统节点(系统节点存储未编码的数据包)和非系统节点。系统节点在故障时是精确修复，非系统节点遵照功能修复模型。不同修复模型和主要的构造技术如图2所示。

其中精确修复问题是最令人关注的问题，它也是最有挑战性的和一个大部分保持开放的研究方向。功能修复问题其实是很容易理解的，正如文献[7]中所述，可以用信息流图构造成一个多播问题。Ahlsewede等人通过证明最下割集的界限来实现组播的最大流特征。进一步研究表明线性网络编码能够达到该最大流特征。这里随机线性结合以高概率构造网络编码。由于功能修复可以简化成组播，我们可以通过评估最小割界限来描述最小修复带宽，使网络编码提供有效的、创造性的解决方法。

精确修复问题比功能修复问题要难。在精确修复中，新的节点访问一些现有的存储节点并精确地再生出丢失的编码包。随后我们将给出修复问题伴随的存储开销和修复带宽的基本折衷。系统精确修复模型是功能修复和精确修复的整合，在本文中讨论功能修复和精确修复。



▲图2 不同修复模型和主要的构造技术

1.1 功能修复

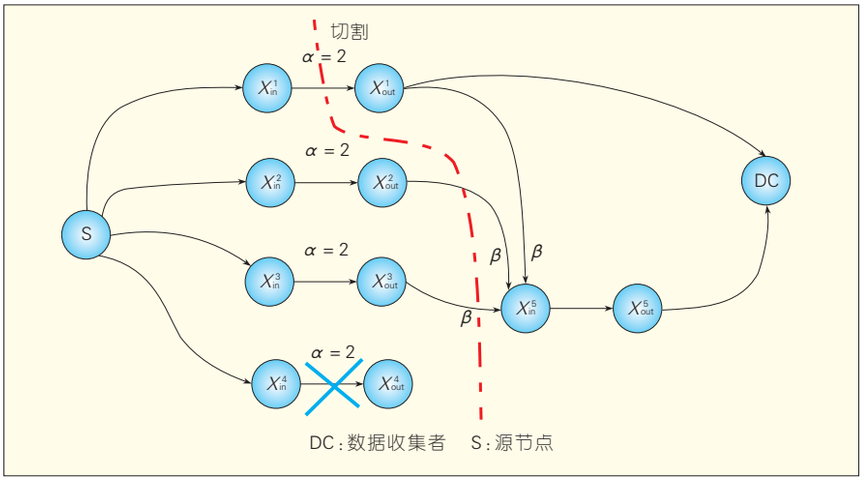
如文献[7]所述，功能修复问题可以说是在一个信息流图中的组播问题。信息流图代表信息流的传输，如节点加入和离开存储网络。图3给出一个信息流图的例子。存储节点通过一对节点X^i_{in}，X^i_{out}来代表，这对节点通过一个容量为该节点的存储容量的边来连接。虚拟的源节点s相当于原始的数据对象。假设在4个节点存储一个大小为M=4的文件，每个节点存储α=2的编码包，这个文件可以通过任意两个存储节点来修复。虚拟的信宿节点叫做数据的收集者，连接任意k节点子集并确保该编码具有MDS属性(n中任意k个节点可以修复原始数据)。假设存储节点4失败，为保持MDS属性创造一个新的节点5，其他的存储节点通过传输最小的信息给节点5并使得节点5存储的编码包为2。

在云灾备系统中使用(4,2)纠错码，其中任意两个节点能够修复原始数据。如果节点X^4失效，一个新节点会加入系统，需要在X^5构造新的编码包。节点X^4_{in}与3个存储节点相连。假设β为与每个存储节点流量的比特数，我们关心的是最小β值。最小割要大于M=4，模块才能修复原始数据。图3中最小割是$\alpha+2\beta$。这意味着流量$\beta\geqslant 1$。总的修复带宽为$\gamma=d\beta=3$。

对于组播，众所周知其最大流为信源和所有接受者的最小割容量，它可以通过线性网络编码实现。由于修复问题可以被看成一个组播问题，网络编码提供了有效的建设性解决方案。接下来，我们提出信息流图切割分析，并给出相应的最小修复带宽和最小存储开销的两个极点。

1.1.1 信息流图的割集分析

假设在云灾备系统中总是有n个活动的存储节点。每个节点可以存



▲图3 功能修复信息流图

储 α 比特。我们称每个失效/修复为一个阶段。在每个阶段，一个存储节点失效，通过从任何 d 个活动节点下载 β 比特来修复失效节点的信息。因此，总共修复带宽为 $\gamma=d\beta$。

以图3为例说明。在初始化阶段，系统由节点1、2、3和4组成；在第二阶段，系统由节点2、3、4和5组成。对于每个参数集合 $(n, d, \alpha, \gamma=d\beta)$，都有相应的有限或无限的信息流图。我们定义有向非循环图为 $g(n, d, \alpha, \gamma)$。要求任何 k 个存储节点均可修复原始信息。新节点从其他活动节点接收相同数量的信息。对于编码存储量 α 和修复带宽 β，(n, d, α, γ) 元组是切实可行的。注意 n, k, d 必须是整数。如果有一个节点失效，新节点最多可以跟所有 n−1 个活动节点连接，因此 $d\leqslant n-1$，α，β，$\gamma=d\beta$ 是修复过程的非负实值的参数。

观察最小修复带宽 $\gamma=d\beta$ 可以发现其是参与修复的节点个数 d 的递减函数。新节点与多个节点交互，每个连接的流量为 β。因此，最小修复带宽可以通过 $d=n-1$ 来实现。

1.1.2 两种特例

我们感兴趣的是最优折衷曲线上的两个极值点，分别相当于最优的存储效应和最小修复带宽。我们称达到这些点的编码分别为最小存储再生编码（MSR）和最小带宽再生编码（MBR）。当 $d=k$，总的修复流量是 M（原始文件的大小）。因此，如果一个新节点只允许和 k 个节点链接，它必然要下载整个数据对象来修复一个新的故障，这是对于执行任何MDS码的原始修复方法。然而，如果允许新节点和多于 k 个节点链接，MSR码可以降低修复带宽。另一个折衷方式是MBR码，也就是最小修复带宽。在最小带宽再产生码中，修复过程中的流量总比特数等于存储大小。

1.2 精确修复

在功能修复下，每当故障发生云灾备就需要不断更新修复，基于随机网络编码的功能修复解决方案需要一个巨大的有限域来支持连续不断的修复。这样计算复杂性会很高，而且功能修复不容易防止窃听。这些缺点促进了节点失效的精确修复的研究。

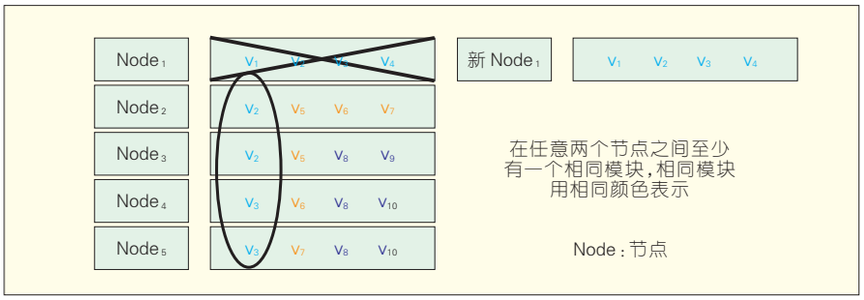
对于MSR点，Wu提出当 $d=n-1$ 时，可以达到 $k=2$ 和 $k=n-1$ 情况。Shah指出对于 $(k/n)>(1/2)+(2/n)$，在标量线性编码下（如 $\beta=1$）精确修复的割集界限不能实现。Suh给出精确MSR码在 $(k/n)>(1/2)$，$d\geqslant 2k-1(2/n)$ 条件下匹配割集界限。在中间的状态 $(k/n)\in[1/2,(1/2)+(2/n)]$，Cullina和Suh提出当 $k=3$ 时可以实现割集界限。Papailiopulos[21-23]进一步探索了连接性。文献[23]的构造是不切实际的，因为它们要求指数分布的领域大小和子包。MSR和MBR以外的中间点，存储空间和修复带宽的限制仍然是一个挑战性开放问题。

1.2.1 精确的MBR码

精确的MBR码：对于 $d=n-1$，最小割集界限可以用一个决定性的方案来实现，它要求有限域字母大小不能找过 $(n-1)n/2$。

一个(5,3)码精确修复节点1的例子如图4所示。通过例子 $(n, k, d, \alpha, \beta)=(5,3,4,4,4)$ 描述了这种思想，其中能够存储的文件大小为 $M=9$（相应的割集界）。每个节点存储4个模块的 $a^t v_t$，v_t 可以被理解为数据文件的一维子空间。我们用子空间向量来代表一个实际的存储模块。注意 d 等于修复的存储模块的大小，可利用的方程数匹配一个节点精确修复所需要的变量。因此，对于精确修复，在节点1和所有 $i\neq 1$ 的节点 i 之间必须至少有一个相同的模块。

这个思想是使其他节点 $i(i\neq 1)$ 分别存储节点1的每个模块：节点2、3、4和5在它们自己位置分别存储 $a^t v_1$、$a^t v_2$、$a^t v_3$、$a^t v_4$。为了确保修复，在任意两个存储节点之间只能够拥有一个相同模块。因此，节点2可以在剩下的其他位置存储其他3个信息块 $a^t v_5$、$a^t v_6$、$a^t v_7$。依据上面的过程，节点3、4和5在它们的空间分别复制3个块中的一个。我们重复这个过程直到总共 $10=(4+3+2+1)$ 块都被存储。这个构造可以保证任何节点失效的精确修复，因为在任意两个节点之间至少有一个模块是相同的，并且各相同模块间是不同的。剩下的问题是设计这10个子空间向量 v_i，$i=1\cdots\cdots 10$。详细的构造来自于MDS码属性，5个节点中任意3个就可以修复原始信息文件。因此，任意(10,9) MDS码可以构造这些 v_i。文献[16]中



▲图4 一个(5,3)码精确修复节点1的例子

的思想可延伸至任意的(n,k)码字。

1.2.2 精确MSR码

这个新的思想是干扰队列。干扰队列思想是将多个干扰信号排成一个信号子空间，它的维数比干扰数量要小。特别当一个译码器需要解码一个需要的信号，这个信号被两个独立的干扰信号干扰。这个译码器需要多少个线性方程来修复它想要的信号？由于跨越需要和不需要信号的总维数不超过3，编译器可以通过在3个未知信号里访问3个线性无关的方程组来修复原始信号。然而，由于译码器只对3个信号的其中一个感兴趣，可以解码出它想要的未知信号。详见文献[11–13]。

2 结束语

本文给出了基于纠错码的云灾备系统的关于降低修复带宽的最新研究。修复问题有3种模型：精确修复、功能修复和系统部分精确修复。在精确修复模型中，丢失的内容要正确的再生成；在功能修复中，在修复前后过程中，只要保持相同MDS码属性；在系统部分精确修复中，系统部分正确的再构造而非系统部分依照功能修复模型。

功能修复问题本质上是在一个无限图中，从一个信源到一个无限的接受者的组播问题。功能修复在存储和修复带宽之间有一个折衷，并且这两个极点通过MBR和MSR码来实现。修复带宽通过最小割集界限定，因此功能修复问题很好理解。

精确修复问题相当于信宿覆盖子集的网络编码问题。对于这些问题，割集界一般不是密封的并且线性编码甚至不能满足[24]。在文献[16]中的讨论表明，对于MBR码修复带宽在$d=n-1$情况下，可以实现给定的割集界限。最小存储点看起来更难理解。现有最好构造中，为了连通性，有效状态为$d\in[2,k-1,n-1]$，对$k/n\leqslant 1/2$提供匹配割集界。然而，符号拓展方式表明割集界限可以通过非常大的子包β来渐近逼近。目前，有一些令人关注的开放问题。

问题1：网络拓扑的影响，如最近文献[20]中描述的树状结构。到目前为止，所有先前对于存储网络的工作，均假设完全的连通性拓扑结构。然而，好多网络有不同的通信容量和稀疏的拓扑结构。对于这种情况，通信将会有不同的开销，需要为其制订一个优化问题。

问题2(修复阵列码)：阵列码在数据存储系统[2–3,25]中广泛使用。对于特需的奇偶码，通过因子0.75来改善以前的数据对象修复方式建立，离最小割集界限仍有一定距离，如果我们实施奇偶码构造，是否可以实现最小修复带宽仍是开放的。

问题3：将精确修复连接到局部译码的深入理论看起来是一个令人关注的研究方向。

问题4：分布式存储的安全和隐私问题是重要的。在修复过程中使用编码可能会混合编码包的传播，因而要有故障控制机制。还有就是在修复过程中信息泄露到敌人的数据隐私性问题。

3 参考文献

[1] BLAUM M, BRADY J, BRUCK J, et al. EVENODD: An efficient scheme for tolerating double disk failures in raid architectures [J]. IEEE Transactions on Computers, 1995, 44(2):192–202.
[2] BLAUM M, BRUCK J, VARDY A. MDS array codes with independent parity symbols [J]. IEEE Transactions on Information Theory, 1996, 42(2):529–542.
[3] BLAUM M, FARRELL P G, VAN TILBORG H. Book chapter on array codes [M]. PLESS V S, HUFFMAN W C. Handbook of Coding Theory. Amsterdam, Netherlands: Elsevier, 1998.
[4] WANG Z, DIMAKIS A G, BRUCK J. Rebuilding for array codes in distributed storage systems [C]//Proceedings of the 2010 IEEE GLOBECOM Workshops(GC Wkshps'10), Dec 6–10, 2010, Miami, FL, USA. Piscataway, NJ, USA: IEEE, 2010: 1905–1909.
[5] LI S Y R, YEUNG R W, CAI N. Linear network coding [J]. IEEE Transactions on Information Theory, 2003, 49(2): 371–381.
[6] DOUGHERTY R, FREILING C, ZEGER K. Insufficiency of linear coding in network information flow [J]. IEEE Transactions on Information Theory, 2005, 51(8): 2745–2759.
[7] MADDAH–ALI M A, MOTAHARI S A, KHANDANI A K. Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis [J]. IEEE Transactions on Information Theory, 2008, 54(8): 3457–3478.
[8] CADAMBE V R, JAFAR S A. Interference alignment and the degrees of freedom for the K user interference channel [J]. IEEE Transactions on Information Theory, 2008, 54(8): 3425–3441.
[9] SUH C, TSE D. Interference alignment for cellular networks [C]//Proceedings of the 46th Annual Allerton Conference on Communications, Control, and Computing (Allerton'08), Sep 23–26, 2008, Urbana–Champaign, IL, USA. Boston, MA, USA: Kluwer Academic Publisher, 2008.
[10] DIMAKIS A G, GODFREY P G, WU Y, et al. Network coding for distributed storage systems [J]. IEEE Transactions on Information Theory, 2010, 56(9): 4539–4551.
[11] WU Y, DIMAKIS A G, RAMCHANDRAN K. Deterministic regenerating codes for distributed storage [C]//Proceedings of the 45th Annual Allerton Conference on Communications, Control, and Computing (Allerton'07), Sep 26–28, 2007, Urbana–Champaign, IL, USA. Boston, MA, USA: Kluwer Academic Publisher, 2007.
[12] WU Y. Existence and construction of capacity–achieving network codes for distributed storage [C]//Proceedings of the IEEE International Symposium on Information Theory(ISIT'09), Jun 28–Jul 3, 2009, Seoul, Republic of Korea. Piscataway, NJ, USA: IEEE, 2009: 1150–1154.

➡下转第11页

云存储与云灾备的原理与短板分析

Mechanism and Challenges Associated with Cloud Storage and Cloud Disaster Recovery

中图分类号：TP393 文献标志码：A 文章编号：1009-6868 (2012) 06-0005-007

邓玉辉/DENG Yuhui
（暨南大学计算机系，广州 510632）
(Department of computer science, Jinan University, Guangzhou 510632, China)

摘要： 通过资源的整合和虚拟化技术的应用，云计算可以对用户提供按需的资源分配和服务，以充分提高资源的利用率。文章认为将该理念拓展到云存储和云灾备，则由于海量数据的高度聚集将会对系统带来一系列的挑战，是云存储与云灾备的短板。实现海量存储系统从传统的纵向扩展向横向扩展转化，实现系统的性能和规模线性可扩展，处理海量存储系统的高度聚集带来的能耗和冷却等问题是在进行云存储和云灾备系统设计时必须要考虑的重要因素。

关键词： 云存储；云灾备；数据；能耗

Abstract: Cloud computing can allocate resources on demand to users through technology that integrates and virtualizes resources. In this way, resource use can be greatly improved. Extending this idea to cloud storage and cloud disaster recovery means highly aggregating mass data, and this will create a number of challenges for system design. Scale-out model, linear scalability, energy consumption, and cooling are very important factors that need to be considered when designing cloud storage system and cloud disaster recovery system.

Key words: cloud storage; cloud disaster recovery; data; energy consumption

来自IDC的报告显示，2011年，共有1 800 EB的数据被创建和拷贝，且数据年增长率达到60%[1]。如果将所有的数据都存储在CD光盘上，堆起来的高度是地球到月球距离的6倍。另外，随着各种家庭数字终端的兴起以及Web 2.0的广泛应用，大众成为信息创造的主体。移动互联网把信息的生产从PC拓展到手机，物联网把信息的生产从人拓展到物，IDC预测2020年全球产生的信息将达到350亿TB。这些数据中的绝大部分将存储在世界各地的大型数据中心。图灵奖获得者Jim Gray曾断言，现在每18个月新增的数据量等于有史以来的数据量之和。信息数字化所产生的呈指数级增长的数据对存储系统的容量提出了非常严峻的挑战。

随着社会信息化程度的不断提高，对数据存储的急剧提升，导致了以“计算”为中心到以“数据存储”为中心的观念革新[2-3]。在过去的十多年中，磁盘的区域密度、轨密度和线密度分别获得了100%、50%和30%的增长[4]。在存储领域有两个重要的技术对存储系统的发展和存储容量的扩展产生了重要的影响。第一个是并行存储技术，比如磁盘阵列技术[5]；第二个就是网络技术对存储系统体系结构的影响。通过将网络引入存储系统，改变主机与外部存储节点间的连接模式，出现了若干新型存储体系结构：附网存储（NAS）和存储区域网（SAN）。网络与存储设备不同的结合方式可以形成不同拓扑结构的网络存储系统，不同的拓扑结构对于系统性能的影响也各不相同[6-7]。但由于性能、价格、可扩展性等各方面的原因，这些仍不足以应对爆炸性的数据增长。另外，许多大型企业的IT基础设施的利用率只有35%。在某些企业中可能会低至15%[8]。Google报告称其服务器的利用率往往在10%到15%之间[9]。这使得工业界不得不重新思考所面临的问题，并努力寻求解决的方法。

2001年，Google首次提出云计算的概念。2007年年底，Google的一名工程师再次提出了云计算。云计算将应用和计算机资源包括硬件和系统软件虚拟化之后包装成服务，通过按需付费的方式，穿越Internet来满足用户各种不同的需求。用户可以不再需要购买昂贵的计算机系统，不再因为需要短时间使用某个软件而不得不购买该软件的使用版权。这种服务模式在过去的十多年中有过充分的探讨，这两年的重新兴起并以一个新的技术名词出现，并不是因为产生了某种技术上的突破，而是由于信息数字化导致数据的爆炸性增长所带来的一系列问题让我们不得不

基金项目：国家自然科学基金（61272073、61073064）；广东省重大科技专项基金（2012A080102002）；广州市科技计划项目基金（2012J4100109）

重新思考计算机系统发展的新走向。另外，由于技术进步所带来的部分老技术的重新复苏也对云计算的发展起到了推波助澜的作用。借助于云计算的理念，将存储资源进行整合，并实现存储资源的按需分配。于是就产生了云存储。

1 云存储面临的挑战

云存储面向个人的应用主要由网盘、在线文档编辑、工作流及日程安排；面向企业的应用主要有企业空间的租赁服务，企业级数据备份和归档、视频监控系统等。无论是哪种应用，海量数据的高度聚集都要导致存储系统从少数的存储引擎向连在网络上的成千上万的商用化存储设备进行转变，从传统的烟囱式的建设模式转变为集约化的建设模式。在过去的十多年中集群网络的重要进展之一是可以将成千上万的节点连起来，同时保证高可扩展性和相对较低的通信开销。因此，我们认为，采用商用化的技术来构造可扩展的集群是云存储的基本组件。因为我们可以以搭积木的形式来聚合存储组件以构造大规模的存储系统。但是现有的存储系统进行规模的扩展之后还存在很多待解决的问题。

1.1 名字空间

存储器空间的组织和分配，数据的存储、保护和检索都依赖于文件系统。文件系统由文件和目录组成。数据按其内容、结构和用途命名成不同的文件，而目录则构建文件系统的层次化结构。现代的文件系统一般都是按树形的层次架构来组织文件和目录。集群文件系统往往也采用树形架构来构造名字空间。然而，当数据的访问从树根走向树叶的时候，访问的延迟会相应地增加。另外，还有两个重要的因素导致树形架构不适合于云存储环境。第一，树根本身就是一个单一失效点，而且很容易形成系统的“瓶颈”；第二，树形架构很难在Internet上扩展到地理上分布的规模。另外，层次化结构使得文件的访问效率不高。每一层目录都隐藏了它所包含的子目录和文件，用户很难知道一个目录下面到底有哪些文件和子目录。因此，用户访问某个文件时，必须通过层次型的目录树结构到达其保存位置，如果不知道文件保存位置，则必须遍历整个目录。因此云存储只有采用非集中式的名字空间来避免潜在的性能“瓶颈”和单点失效。

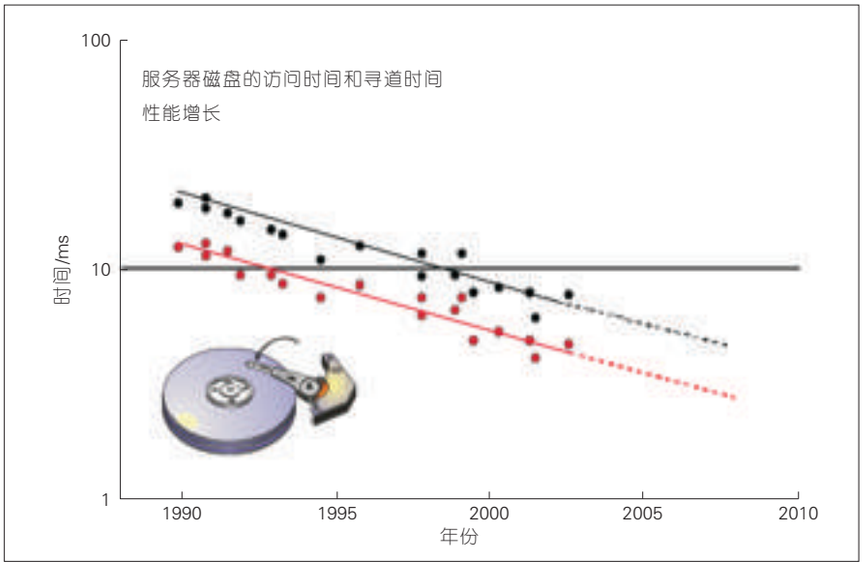
1.2 元数据组织

元数据是描述数据的数据，主要用来反映地址信息和控制信息，通常包括文件名、文件大小、时间戳、文件属性等等。元数据主要是用来管理的操作数据。研究表明，在文件系统的操作中，超过50%的操作是针对元数据的[10]。元数据最重要的特点是其往往是小的随机请求。一般来讲，元数据都是存储在磁盘上的，然而，和磁盘存储容量的增长不同的是，由于机械组件所带来的延迟，磁盘的平均访问时间每年的降低不足8%[11]。Hitachi的磁盘在过去10年里磁盘访问时间和寻道时间的发展趋势[12]如图1所示。对于这种由小的随机请求所组成的数据访问流，磁盘的寻道时间是磁盘访问延迟中最主要的部分[13]。因此，对于大规模系统来讲，元数据的访问往往成为制约整个系统性能的“瓶颈”。

很多分布式的存储系统将数据访问和元数据的访问分离开来。在这样的系统中，客户端首先和元数据服务器通信来获取元数据包括文件名、文件位置等信息。然后，利用该元数据，客户端直接和数据服务器通信去访问相应的数据。一般来讲，元数据服务器的内存可以满足大部分的读请求，但服务器不得不周期性地访问磁盘来读取需要的数据，并且所有元数据的更新也要写回到磁盘。存储系统空间的增长可以通过增加额外的存储服务器来保证。然而，对于一个管理数以亿计的数据文件的云存储系统，保证元数据的访问性能和可扩展性比较困难。对于像云这样的需要高可扩展性的环境，对元数据的依赖给系统设计带来了巨大的挑战。

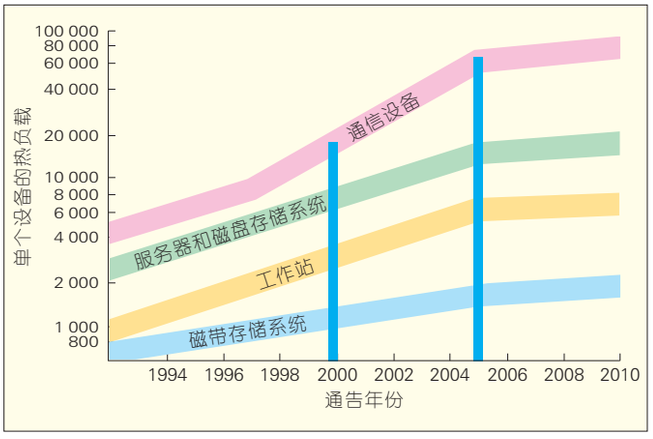
1.3 能耗与地板空间

2005年美国新建立的数据中心需要消耗的能量相当于加利福尼亚州所消耗能量的10%（大约5 GW），



▲图1 磁盘访问时间和寻道时间的发展趋势

需要花费大约40亿美金[14]。英国的1 500个数据中心每年消耗的能量和英国第十大城市莱卡斯特所需要的能量相当。2010年，英国单个数据中心每年在能量上的花费达到大约740万英镑[15]。在这些数据中心中，存储系统所消耗的能量达到了总能耗的27%。另外，消耗的能量除了供各种计算机组件工作外，还会产生大量的热量。由于大部分计算机组件只能在一定的温度环境下才能保证足够的可靠性，因此，还需要额外的能量驱动制冷设备。Netapp的调查表明大型数据中心中制冷系统的能耗仅次于服务器[16]。数据中心主要设备的热密度趋势[17]如图2所示。可以认为，数据中心的能耗问题处于一个恶性循环的状态。



◀图2
数据中心的热密度趋势图

另外，由于数据的增长导致数据中心对新设备需求的不断增加，但是数据中心的可扩展性完全受限于其地板空间。在数据中心的空间未扩展的情况下，随着单位地板面积内计算机设备的不断增加，传统数据中心的设备容量必将达到极限。因此，能耗和地板空间成为当前设计和管理大型数据中心所面临的主要挑战。

2 云灾备

国际上对于IT系统灾难的定义是指由于人为或自然的原因，造成信息系统运行严重故障或瘫痪，使信息系统支持的业务功能停顿或服务水平不可接受，并达到特定的时间的突发性事件。虽然数据是企业的命脉，然而在传统的存储系统下，数据丢失很难避免。数据丢失的原因如图3所示。图3表示人为因素是导致数据丢失的最重要的原因。由于管理员或员工的活动造成数据的损失或变更，使数据的完整性与真实性受到影响，如误删除、误格式化或误分区、误克隆等误操作，系统管理员出错或蓄意破坏、窃取等等。因此，如果在云计算环境下，专业的工程技术人员将能最大限度地避免由于人为因素所导致的数据丢失。然而，设备和硬件故障所带来的数据丢失则很难避免。例如，硬盘损坏是极为常见的导致数据丢失的原因，一般来讲，磁盘阵列（RAID）系统能够一定程度上避免硬盘故障导致的数据丢失，如RAID1、RAID5都能够在一块硬盘失效后对数据进行修复。但在两块硬盘失效的情况下，则仅有RAID6数据保护模式能够保护数据不丢失，而RAID6由于复杂冗余和校验算法导致系统大量的开销，一般企业采用时存在顾虑。另外，大型存储系统中磁盘的失效往往是具有相关性的，一块大容量磁盘失效后要进行长时间的重构（例如，1 TB容量的磁盘重构可能需要数小时），会对系统带来极高的存储I/O率，这可能导致另一块磁盘的失效，从而引发连锁效应。因此，利用蝴蝶效应来描述毫不为过。

2011年4月，亚马逊的网络服务经历了长时间断电，造成停机等一系列问题，并且影响到了云计算的服务。在长达4天的时间里，一些客户无法使用亚马逊的存储服务，并且会出现网络配置错误。2011年4月25日，Vmware的Cloud Foundry在发布13天后连续两天发生服务中断事件。第一次是由于某供电柜发生故障，在停机持续了10小时后，故障得到修复。但在第二天，当Vmware的官方工作人员在尝试实施先期检测方案以避免前一天的事故再一次发生时，导致了新一轮的停机。2011年8月，都柏林的亚马逊和微软的数据中心因遭遇雷击而停电，两家企业都经历了数天才完成修复。国际最知名的IT企业也无法保证其IT基础设施的24×7×365业务连续性。再者，不可预测的自然灾害也会导致数据丢失，如日本的广岛地震，中国的汶川地震等。因此，对数据进行有效的灾备，并经常性的进行恢复演练确保备份的有效性能够最大程度的降低因为

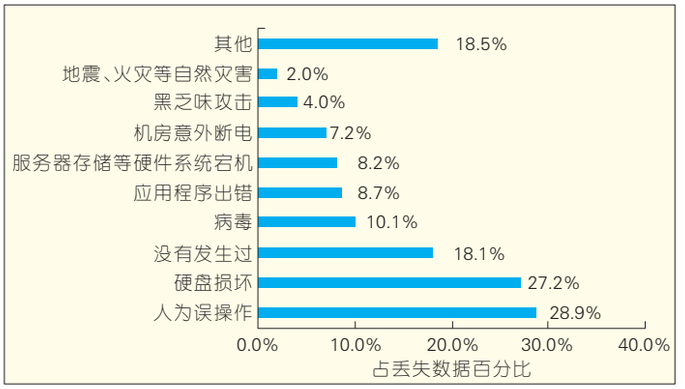


图3▶
数据丢失的原因

硬件故障导致数据丢失的可能性，充分得到云存储用户的信任。

2.1 灾备的技术指标

在容灾体系中，人们往往采用恢复点目标（RPO）和恢复时间目标（RTO）这两个指标来衡量容灾体系的应急能力和系统保护能力。RPO体现为灾难发生后，恢复运转时数据丢失的可容忍程度。RTO表示需要恢复的紧迫性也即多久能够得到恢复的问题。然而，在设计一个容灾系统时，并不意味着RPO和RTO越小越好。因为系统投资会随着RPO和RTO的降低而增加。因此，最佳的容灾方案不一定是性价比最好的方案。

2.2 数据备份

数据备份是指为防止系统出现操作失误或系统故障导致数据丢失，而将数据集合从应用系统中以备份格式存储到处于离线的存储介质的过程。在数据备份过程中，一般采用备份软件配合磁带库的物理介质保存系统来进行。数据备份分为完全备份、差异备份和增量备份。完全备份是指对某一个时间点上的所有数据或应用进行的一个完全拷贝。差异备份则备份自上一次完全备份之后有变化的数据。增量备份则备份自上一次备份（包含完全备份、差异备份、增量备份）之后有变化的数据。无论哪种模式都完全服从备份计划的规定，即在固定的时间点开始备份。

传统的备份系统并不保证数据的实时性或近实时性。而且，备份后的数据格式是专用的备份格式，并非应用系统中的数据原有格局。因此，当发生灾难时，备份数据是不能立即使用的，必须先恢复。恢复时要通过格式转换进行导回操作，这导致无法保证恢复的快捷。例如，如果按T h的时间间隔来进行增量备份。如果在A时间点发生了系统故障，那只能回复到上一个备份点$A-T$，而且还要进行数据格式的转换。随着T的增加和数据量的增涨，需要恢复的时间也随之线性增涨。因此，指标RPO和RTO都会较高，也很难保证IT基础设施的$24\times7\times365$业务连续性。另外，为了提高RPO，必须提高数据备份的频度。但大多数情况下，仅仅增加备份的频度会带来一系列的问题。例如：应用的高峰时段无法进行备份操作；备份数据所花时间太长。因此，需要有一个契机和一个新的技术的诞生，来达到以用户为中心的数据安全和系统安全的要求。

2.3 数据复制

为了保证较低的RPO和RTO目标，数据复制技术常应用于各种灾备系统。数据复制是将原卷或原文件直接复制到目标卷或目标文件系统中，分别称为卷复制和文件复制。由于数据复制的目标卷（目标文件）和源卷（源文件）的数据格式一致，可以消除备份系统中数据格式的转换时间。数据复制又分为同步复制和异步复制。

2.3.1 同步复制

同步复制表示，在数据复制系统的源端，主机发出的I/O请求在写入本地磁盘的同时，通过专用的数据网络或通道将数据从本地磁盘系统同步地复制到异地磁盘系统。当异地系统完成该I/O操作后，通知本地系统I/O完成，本地的主机系统才能发出第二个I/O请求。利用同步复制方式建立异地数据灾备，可以保证异地系统和本地系统数据的完全一致性。但同步复制方式对性能的要求非常高。由于每一次本地I/O必须要等到数据成功地写到异地系统，才能进行下一个I/O操作，因此同步复制的性能受网络带宽、网络的距离、中间设备及协议转换等多方面的影响。

2.3.2 异步复制

异步复制是指在数据复制系统的源端，主机发出的/O请求在写入本地磁盘的同时，向本地磁盘系统上预留的空间发出相同的写请求（决定于不同的策略），然后通知本地系统I/O完成。此时，本地的主机系统可以发出第下一个I/O请求。在设定的复制规则满足后（基于时间、基于变化量等），系统的复制功能模块再将数据通过专用的数据网络或通道复制到异地的存储系统中。

2.4 灾备分析

与同步复制相比，异步复制对网络带宽和距离的要求低很多，只要在某个时间段内能将数据全部复制到异地即可，同时异步复制对应用系统的性能影响也很小。但是，当本地系统发生灾难时，异地系统上的数据可能会短暂缺失（在复制的时间间隔内数据未完整地从源端发送到目的端）。因此，当源端灾难发生时，同步复制的RPO接近于0，异步复制的RPO则取决于复制时间间隔。同时，在业务恢复时间上，相对于传统的备份系统而言，由于不存在数据格式的转换，可以在较短的时间内恢复业务系统，从而具有较好的RTO。对于规模1 000亿元人民币以上的银行，银监会要求建立200 km以上的备份系统。因此只能使用远程复制模式。同城复制可以使用光纤，但是远程复制由于成本方面的因素，全光纤传输还很遥远。因此，不可能采用同步复制。所以，远程异步复制模式会越来越多。

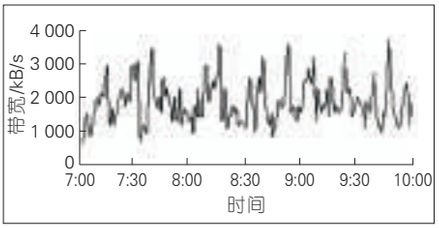
3 云存储与云灾备的短板

当用户向云存储系统中进行数据备份时，网络对系统性能的影响起到了至关重要的作用。当云存储服务提供商在进行后台的云灾备时，远程的云备份和云复制也依赖于网络的性能。

3.1 网络短板

按照Nielsen法则[18]，终端用户的

网络带宽以每年50%的速度增长。然而，和局域网形成鲜明对照的是，广域网的性能不尽人意。例如，一条T1线路的带宽只相当于千兆网的千分之一，许多帧中继线路的带宽只有256 kb/s。Garfinkel[19]通过测量发现从美国伯克利大学到西雅图的平均网络写带宽大约是5~18 Mb/s。通过使用网络测试工具iperf，采用256个数据流测量，数据表明在格林尼治标准时间下午7点到10点，从英国剑桥大学到中国北京的平均网络带宽大约是14 Mb/s，如图4所示[20]。



▲图4 英国剑桥大学到中国北京的网络带宽

基于以上的测试数据，如果假设网络带宽为20 Mb/s，Armbrust[21]等人作了简单的计算，计算结果表明从美国伯克利大学传输10 TB数据到西雅图需要45 d的时间（10 × 1 012 B/（20 × 106 b/s）=4 000 000 s=45 d）。如果通过亚马逊来进行该数据传输，需要另外向亚马逊支付1 000美元的网络传输费用。另外，由于广域网物理距离的原因，不可避免的时延也会对带宽造成影响。例如，一个T3链路（44.736 Mb/s），当时延超过40 ms时，其带宽很快就下降到与T1链路（1.544 Mb/s）相当。

如果是进行云备份，时间上的开销相对还可以忍受，因为用户在本地还有一个数据拷贝可供使用。但如果是从云存储系统中恢复数据，这是无法让人接受的，特别是对于那些需要提供24 × 7 × 365业务连续性的企业级用户。为了缓解这个问题，对于云存储系统中大数据量的恢复，云存储提供商Mozy[22]和CrashPlan[23]提供了一个不得已的选择，在用户许可的情况下，将数据转存在DVD或者硬盘上，然后通过特快专递的形式交付给用户。

3.2 网络优化

针对广域网数据传输的协议优化如图5所示。为了优化广域网环境下大规模数据传输的性能，我们曾将数据在套接字层在发送端进行分割，然后利用多个套接字流进行并行传输，最后在接收端进行数据重组（如图5（c）所示）。理论上讲，对传输控制协议（TCP）管道而言，其最大的吞吐量为带宽延迟乘积，即容量=带宽 × 环回时间。在传输窗口一定的情况下（图5中红色的方形区表示传输窗口，缺省为64 kB），按通常100 Mb/s的网络带宽来计算，传统的单套接字流显然无法填满TCP管道（如图5（a）所示），使得其效率极低。通过加大传输窗口可以在一定程度上提高TCP管道的利用率（如图5（b）所示），但在丢包的情况下，会导致每次重传的数据增加。因此，通过多个套接字流来并行传输的效果较好。另外，由于采用了多流，不同的数据流在必要的情况下可以走不同的路由，也能够进一步优化广域网的性能。

正如前面提到的，云基础设施必须是地理上分布的，因为云的成功在很大程度上决定于其规模效应。虽然计算和存储相对便宜，然而，由于广域网环境下的低带宽、高延迟和较高的丢包率，使得广域网成为云环境下那块最短的木板。因此，在地理上分布的云环境下进行大规模的数据传输是非常昂贵的。图灵奖获得者Jim Gray[24]在2006年就指出在广域网上处理大数据集时，应该将程序传给数据，而不是将数据传给程序。另外，也可以通过数据压缩、数据去重等方法来减少网域网上的数据传输流量，降低对网络带宽的需求。还可以采用动态缓存、IP流量管理以及服务质量（QoS）控制等方法来降低广域网的延迟。但是，这些方法只能在一定程度上来缓解网络“瓶颈”问题，不能从根本上解决问题。因此，在设计云存储和云灾备系统时，必须要考虑广域网的带宽、延迟和包丢失率所带来的影响。

4 云存储实例分析

对于企业用户而言，现有的云存储更多的是一种在线远程备份的系统[22-23,25-26]。Hu等人[27]针对Mozy、Carbonite、Dropbox、Crashplan 4种云存储系统进行了测试、比较和分析。当将8 GB的文件备份到云存储系统中时，有的系统备份时间超过了30 h，还有的系统经过4 d的时间还未备份完成。当他们将数据集减小到2 GB左右时，云备份系统才回复到基本正常的工作状态。

图6表示Hu等人在Mozy、Carbonite、Dropbox、Crashplan 4个不同

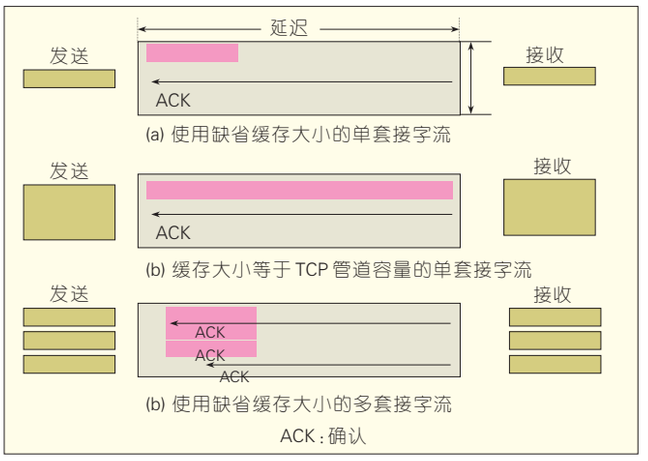
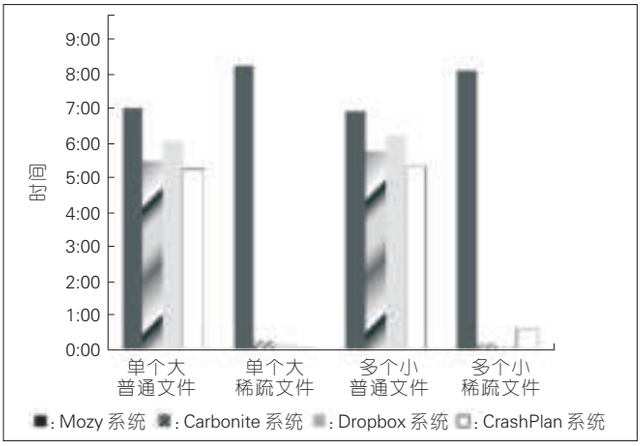


图5▶ 针对广域网数据传输的协议优化



◀图6 2.12 GB 数据的备份时间

的云存储系统下备份 2.12 GB 数据时的远程备份时间。其中横坐标从左到右的 4 种常见情况分别表示单个 2.12 GB 的大普通文件、单个 2.12 GB 的大稀疏文件、很多小的普通文件组成 2.12 GB 的数据集、很多小的稀疏文件组成 2.12 GB 的数据集。稀疏文件表示该文件不包含用户数据，也没有分配用来存储用户数据的磁盘空间。当数据被写入稀疏文件时，文件系统（例如微软的 NTFS）才逐渐地为其分配磁盘空间。可以看到对于正常 2.12 GB 的文件数据 4 个系统的备份时间都超过了 5 h。

图 7 表示相应的恢复时间。恢复比备份要相对快很多，这主要是由于网络的上行链路和下行链路带宽的不对称造成的。通过大量的测试分析，Hu 等人得出了一下结论：

（1）云存储系统必须对于网络失效具有回弹性，同时能够实现大文件的增量备份。

（2）云存储提供商在进行大数据的网络传输时还要进行加密、压缩等预处理以避免网络延迟。

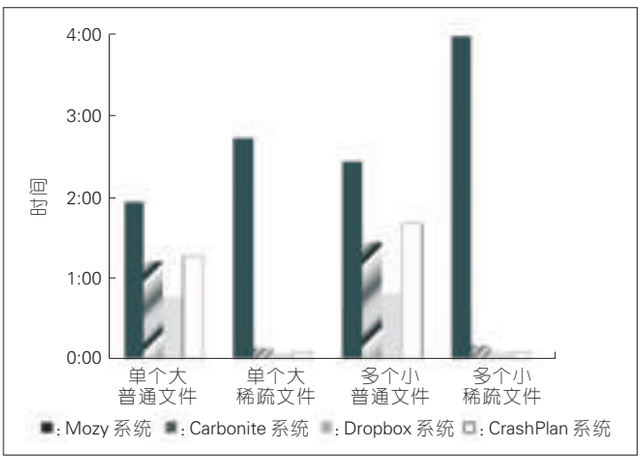
（3）云存储用户需要手动检测重要的文件是否都已经进行了备份。

（4）云存储用户应该将云存储系统作为本地备份系统的一种补充，而不能将其当成主要的备份策略。

本文认为，现有的云存储应对普通用户小数据的备份与恢复应该问题不大，但是企业级用户大数据量的存储与恢复则要慎重考虑。

5 结束语

云存储面向个人的应用主要有网盘、在线文档编辑、工作流及日程安排。面向企业的应用主要有企业空间的租赁服务，企业级数据备份和归档、视频监控系统等。云灾备则主要用于保证云存储服务商后台系统的可靠性和可用性。对两者而言，海量数据的高度聚集会对系统带来一系列的挑战。例如，如何实现海量存储系统从传统的纵向扩展向横向扩展转化？如何实现系统的性能和规模线性可扩展？如何处理海量存储系统的高度聚集带来的能耗和冷却？等问题[13,28–29]都是我们在进行云存储和云灾备系统设计时必须要考虑的重要因素。当然，云存储最终能否成功，还受到其他很多因素的影响，如大量的数据存储在云端如何保证数据的安全和用户隐私等。



◀图7 2.12 GB 数据的恢复时间

6 参考文献

[1] GANTZ J F, Chute C, Manfrediz A, et al. The diverse and exploding digital universe: An updated forecast of worldwide information growth through 2011 [R]. IDC.2008.
[2] KATZ R H. High performance network and channel based storage [J]. Proceedings of the IEEE, 1992, 80(8): 1238–1261.
[3] CHRISTENSEN G S, .FRANTA W R, PETERSEN W A. Future directions of high speed networks for distributed storage environments [C]//Proceedings of the 11th IEEE Symposium on Mass Storage Systems, Oct 7–10, 1991, Monterey, CA, USA. Piscataway, NJ, USA: IEEE, 1991:145–148.
[4] Hitachi global storage technologies – HDD technology overview charts [R]. Hitachi. 2011.
[5] PATTERSON D A, GIBSON G, KATZ R H. A case for redundant arrays of inexpensive disks (RAID)[J]. SIGMOD RECORD, 1988, 17(3):109–116.
[6] DENG Y H. RISC: A resilient interconnection network for scalable cluster storage systems [J]. Journal of Systems Architecture, 2008, 54(1/2):70–80.
[7] DENG Y H. Deconstructing network attached storage systems [J]. Journal of Network and Computer Applications, 2009, 32(5): 1064–1072.
[8] BARROSO L, HÖLZLE U. The case for energy–proportional computing [J]. Computer, 2007, 40(12):33–37.
[9] DENG Y H, WANG F, HELIAN N, et al. Dynamic and scalable storage management architecture for grid oriented storage devices [J]. Parallel Computing, 2008, 34(1):17–31.
[10] WEIL S, POLLACK K, BRANDT S, et al. Dynamic metadata management for petabyte–scale file systems [C]// Proceedings of the ACM/IEEE Conference on Supercomputing (SC'04), Nov 6–12, 200, Pittsburgh, PA, USA. Piscataway, NJ, USA: IEEE, 2004:4p.
[11] HSU W W, SMITH A J. The performance impact of I/O optimizations and disk improvements [J]. IBM Journal of Research and Development, 2004, 48(2): 255–289.

[12] Hitachi global storage technologies–HDD technology overview charts [R]. Hitachi. 2011.
[13] DENG Y H. What is the future of disk drives, death or rebirth? [J]. ACM Computing Surveys, 2011, 43(3):Article 23.
[14] CHASE J S, DOYLE R P. Balance of power: Energy management for server clusters [C]// Proceedings of the 8th Workshop on Hot Topics in Operating Systems (HotOS'01), May 20–23, 2001, Elmau/Oberbayern, Germany. Los Alamitos, CA, USA: IEEE Computer Society, 2001.
[15] FILDES N. UK data–centres use more power in a year than city of Leicester [M]. London, UK: The Independent , 2006.
[16] BATTLES B, BELLEVILLE C, GRABAU S, et al. Reducing data center power consumption through efficient storage [R]. Netapp. 2007.
[17] HUGHES R. Design guidelines for a high density data center [EB/OL]. [2011–09–17]. http://www.availability.com. 2011.
[18] Nielsen's law of Internet bandwidth [EB/OL]. [2011–04–05]. http://www.useit.com/alertbox/980405.html. 2011.
[19] GARFINKEL S. An evaluation of amazon's grid computing services: ec2, s3 and sqs [R]. TR–08–07. Harvard University, 2007.
[20] WANG F Z, WU S, HELIAN N, et al. Khare. Grid–oriented storage: A single–image, cross–domain, high–bandwidth architecture [J]. IEEE Transactions on Computers, 2007, 56(4): 474–487.
[21] ARMBRUST M, FOX A, GRIFFITH R, et al. Above the clouds: A Berkeley view of cloud computing [R]. UCB/EECS–2009–28. University of California at Berkeley. 2009.
[22] Mozy [EB/OL]. [2011–04–9]. http://mozy.com/. 2011.
[23] Crashplan [EB/OL]. [2011–06–05]. http://www.crashplan.com/. 2011.
[24] SZALAY A, GRAY J. Science in an exponential world [J]. Nature, 2006, 440: 413–414.
[25] Carbonite [EB/OL]. [2011–08–06]. http://www.carbonite.com. 2011.
[26] Dropbox [EB/OL]. [2011–08–06]. http://www.dropbox.com. 2011.
[27] HU W, YANG T, MATTHEWS J N. The good, the bad and the ugly of consumer cloud storage [J]. Operating Systems Review, 2010, 44(3): 110–115.
[28] DENG Y H, PUNG B. Conserving disk energy in virtual machine based environments by amplifying bursts [J]. Computing, 2011, 91 (1): 3–21.
[29] DENG Y H, WANG F, HELIAN N. EED: Energy efficient disk drive architecture [J]. Information Sciences, 2008, 178(22): 4403–4417.

收稿日期：2012–10–05

作者简介

邓玉辉，暨南大学计算机系教授，ACM 会员，中国计算机协会存储专委会委员，Thapar 大学兼职教授；研究范围包括云计算、网络存储、集群计算、计算机系统性能评估等；已主持基金项目 7 项；已发表学术论文 50 余篇。

⬅上接第 4 页

[13] WU Y. Existence and construction of capacity–achieving network codes for distributed storage [J]. IEEE Journal on Selected Areas in Communications, 2010, 28 (2): 277–288.
[14] WU Y, DIMAKIS A G. Reducing repair traffic for erasure coding–based storage via interference alignment [C]//Proceedings of the IEEE International Symposium on Information Theory(ISIT'09), Jun 28–Jul 3, 2009, Seoul, Republic of Korea. Piscataway, NJ, USA: IEEE, 2009: 2276–2280.
[15] CULLINA D, DIMAKIS A G, HO T. Searching for minimum storage regenerating codes [C]//Proceedings of the 47th Annual Allerton Conference on Communications, Control, and Computing(Allerton'09), Sep 30–Oct 2, 2009, Urbana–Champaign, IL, USA. Boston, MA, USA: Kluwer Academic Publisher, 2009.
[16] RASHMI K V, SHAH N B, KUMAR P V, et al. Exact regenerating codes for distributed storage [C]//Proceedings of the 47th Annual Allerton Conference on Communications, Control, and Computing(Allerton'09), Sep 30–Oct 2, 2009, Urbana–Champaign, IL, USA. Boston, MA, USA: Kluwer Academic Publisher, 2009.
[17] SHAH N B, RASHMI K V, KUMAR P V, et al. Explicit codes minimizing repair bandwidth for distributed storage [C]//Proceedings of the 2010 IEEE Information Theory Workshop on Networking and Information Theory(ITW'10), Aug 30–Sep 3, 2010, Dublin, Ireland. Piscataway, NJ, USA: IEEE, 2010: 5p.
[18] WU Y. A construction of systematic MDS codes with minimum repair bandwidth [J]. IEEE Transactions on Information Theory, 2011, 57(6):3738–3741.
[19] DUMINUCO A, BIERSACK E. A practical study of regenerating codes for peer–to–peer backup systems [C]// Proceedings of the 29th International Conference on Distributed Computing Systems (ICDCS'09), Jun 22–26, 2009, Montreal, Canada. Piscataway, NJ, USA: IEEE, 2009:376–384.
[20] LI J, YANG S, WANG X, et al. Tree–structured data regeneration in distributed storage systems with regenerating codes [C]//Proceedings of the 29th Annual Joint Conference of the IEEE Computer and Communications(INFOCOM '10), Mar 14–19, 2010, San Diego, CA, USA. Piscataway, NJ, USA: IEEE: 9p.
[21] PAPAILIOPOULOS D, DIMAKIS A G. Interference alignment as a rank constrained rank minimization [C]// Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM'10):Vol 2, Dec 6–10, 2010, Miami, FL, USA. Piscataway, NJ, USA: IEEE, 2010: 895–900.
[22] PAPAILIOPOULOS D, DIMAKIS A G. Connecting distributed storage codes and secure degrees–of–freedom of multiple access channels [C]//Proceedings of the 48th Annual Allerton Conference on Communications, Control, and Computing (Allerton'10), Sep 29–Oct 1, 2010, Urbana–Champaign, IL, USA. Boston, MA, USA: Kluwer Academic Publisher, 2010.
[23] CADAMBE V, JAFAR S, MALEKI H. Distributed data storage with minimum storage regenerating codes –– VExact and functional repair are asymptotically equally efficient [C]//Proceedings of the 3rd IEEE International Workhsop on Wireless Network Coding(WiNC'10), Jun 21, 2010, Boston, MA, USA. Piscataway, NJ, USA: IEEE, 2010.
[24] DIKALIOTIS T, DIMAKIS A G, HO T. Security in distributed storage systems by communicating a logarithmic number of bits [C]//Proceedings of the IEEE International Symposium on Information Theory(ISIT' 10), Jun 13–18, 2010, Pasadena, CA, USA. Piscataway, NJ, USA: IEEE, 2010: 1948–1952.
[25] HUANG C, XU L. STAR: An efficient coding scheme for correcting triple storage node failures [C]//Proceedings of the 4th USENIX Conference on File and Storage Technologies(FAST'05): Vol 4, Dec 14–16, 2005, San Francisco, CA, USA. Berkeley, CA, USA: USENIX Association, 2005:15p.

收稿日期：2012–10–05

作者简介

李挥，北京大学教授、博导；北京大学深圳研究生院信息学院集成电路与系统系主任、先进网络技术实验室主任、深圳市云计算重点实验室副主任；研究方向为三网合一流媒体云计算网络视频关键技术、下一代网络体系结构、网络路由和宽带交换结构、网络编码理论及其应用、嵌入式系统。

侯韩旭，北京大学在读博士研究生；主要研究多信源网络编码理论及应用、三网合一流媒体云计算网络视频关键技术。

黄显霞，北京大学在读硕士研究生；主要研究分布式云存储的关键技术。

基于物联网环境的云存储及安全技术研究

Cloud Storage and Security Technology Based on the Internet of Things

中图分类号：TP393 文献标志码：A 文章编号：1009-6868 (2012) 06-0012-005

杨继慧/YANG Jihui
周奇年/ZHOU Qinian
张振浩/ZHANG Zhenhao
（浙江理工大学，杭州 310018）
（Zhejiang Sci-Tech University, Hangzhou 310018, China）

摘要： 文章基于云存储的架构模型和云存储的关键技术，从可用性、可靠性和数据共享3个方面分析了云存储与物联网的结合的可行性，探讨了云存储在物联网环境下所面临的安全问题及相应的解决办法和物联网环境下的云存储平台结构模型，并对云存储在物联网中的未来发展趋势进行了展望。

关键词： 云存储；海量数据存储；物联网；存储安全

Abstract: In this paper, we discuss the feasibility of combining cloud storage with the Internet of things (IoT). We look at three aspects—availability, reliability, and data resource sharing—and explore security problems IoT cloud storage. Solutions to these problems are proposed. We also present a cloud storage structure model and forecast trends in IoT cloud storage.

Keywords: cloud storage; mass data storage; the internet of things; storage security

物联网是将具有计算、通信和信息感知能力的设备嵌入到物品中，然后按照约定的协议来把物品与互联网连接起来，进行信息交换和通信，以实现智能化识别、定位、跟踪、监控和管理的一种网络[1]。通常在物联网中有数以亿计的传感设备，这些传感设备时刻都在收集、传输和交换数据，因此，物联网是一个数据的海洋，需要一个强有力的存储平台来满足其应用需求。同时，物联网最主要的目的不在于物物相连，而在于通过物与物之间的互联交换来为用户提供智能化服务。对于物联网应用服务提供商来说，传统存储模式要求企业一次性投入大量资金购置软硬件存储设备，搭建平台。而且随着新业务新应用的出现，企业还要对这些设备不断的维护和升级。在这种模式下，存储设施将占用企业很大的资金投入。

近年来随着云计算技术的兴起，云存储受到了人们的广泛关注。云计算为用户提供两种服务，一种是计算资源服务，把计算能力作为一种服务提供给用户；另一种是存储服务，将存储作为服务提供给用户，即本文所讲的云存储。云存储通过一系列软件集合将各种异构存储设备集合在一起构成海量存储空间供用户使用，需要存储服务的用户不再需要建立自己的数据中心，只需向云存储服务商申请存储服务，将自己的数据存放在云存储服务商提供的存储空间中。云存储模式使企业避免了存储平台的重复建设，节约了昂贵的软硬件基础设施投资。当前，云存储模式得到了众多厂商的支持和关注，众多知名厂商纷纷推出自己的云存储服务如Amazon公司推出的简单存储服务S3、谷歌推出的在线存储服务GDrive、微软公司推出的Windows Azure存储服务等。

1 云存储的基本概念和关键技术

1.1 云存储的概念和通用结构模型

云存储是在云计算的概念上延伸和发展出的一个新的概念，它是指通过集群应用、网格技术或分布式文件系统等功能，将网络中大量各种不同类型的存储设备通过应用软件集合起来协同工作，共同对外提供数据存储和业务访问功能的一个系统[2]。对比传统的存储设备，云存储是一个由网络设备、存储设备、服务器、应用软件、公用访问接口、接入网和客户端程序等多个部分组成的系统[3]。在云存储系统中，各个部分以存储设备为核心，通过应用软件集合来对外提供数据存储和业务访问服务。云存

基金项目：国家自然科学基金项目(60702069、50802089)

储的通用结构模型如图1所示[2-4]。

（1）存储层

存储层是云存储最基础的部分，包括存储设备层和存储设备管理层。存储层由各种各样的存储设备和网络设备组成，为了实现低成本，这些存储设备以及网络设备通常都是普通的商业产品，而不是可靠性更高的高端设备，系统的可靠性由一系列软件集合来保证。存储设备可以是网络连接式存储（NAS）和由Internet小型计算机系统接口（ISCSI）所建立和管理的存储区域网等IP存储设备，可以是服务器连接存储（SAS）和小型计算机系统接口（SCSI）磁盘阵列等直连式存储（DAS）存储设备。

存储设备层之上是存储设备管理层，用来实现对存储设备的逻辑虚拟化管理、多链路冗余管理、硬件设备状态监控及故障维护等功能。

（2）基础管理层

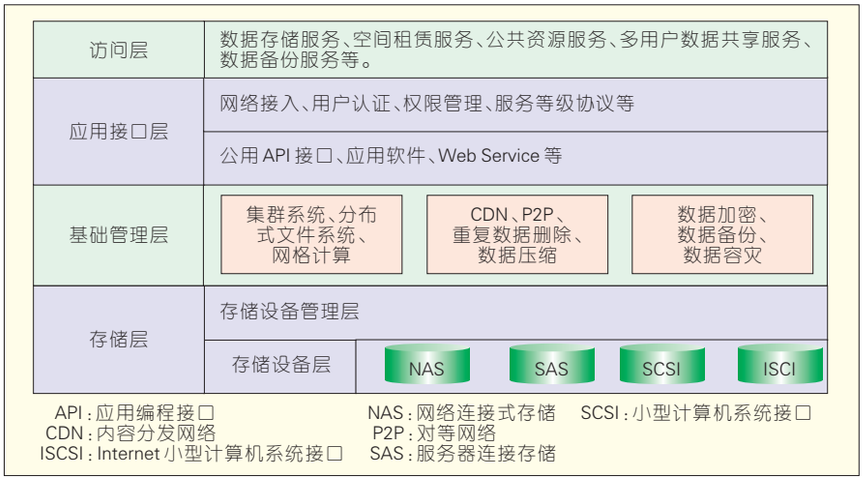
基础管理层是云存储系统中最为核心的部分。云存储并不是一个单点存储系统，而是一个有成千上万个存储设备和服务器构成的一个集合体。基础管理层通过集群、分布式文件系统和网格计算等技术，实现了云存储中多个存储设备之间的协同工作，使多个存储设备可以对外提供同一种服务，并在实现了良好的扩展性的同时，也满足了高可用性以及性能的需求。基础管理层还需要负责数据的加密、备份以及容灾。

（3）应用接口层

应用接口层是云存储中最灵活多变的部分，是用户利用云存储资源进行应用开发的关键部分。云存储供应商通过应用接口层，对客户提供统一的协议和编程接口，通常这些协议都是与平台无关的。

（4）访问层

访问层是基于云存储开发的应用程序的入口，云存储系统通过提供标准的公用应用接口来使授权用户享受云存储服务。云存储服务商可以根据服务类型和用户的不同来提供不同访问手段，从而保证数据的安全性和服务质量。



▲图1 通用云存储的结构模型

1.2 云存储的关键技术

与传统的存储技术相比，云存储最大的特点是可扩展性好。当对存储空间的需求增加时，只需要简单增加存储设备就可以达到目的，并不需要对存储结构进行重新设计。而且随着存储空间的增加，系统的整体性能并不会下降。可以说云存储是对虚拟化的存储资源管理和使用。云存储的关键技术包括存储虚拟化、弹性存储空间扩展、分布式存储、数据隔离与保护等。

（1）存储虚拟化技术

存储虚拟化技术是将物理资源存储进行了替换，通过将多个存储设备整合成一个与物理存储资源有相同功能和接口的虚拟存储资源，这样系统可以提供海量存储空间给用户，这些存储空间可能是建立在一个实际的物理资源上，也可能是跨多个物理资源，用户不需要了解底层的物理细节。

（2）弹性存储空间扩展技术

云存储提供了一个巨大的存储资源池，但是应用对存储资源的使用具有时空性，即不同的时间段和不同地点对存储资源的需求是不同的，这就要求系统具有良好的弹性存储空间扩展功能。

（3）分布式存储技术

分布式存储，就是将数据分散存储在多台独立的设备上，利用多个存储设备的存储资源来满足单个存储设备所不能满足的存储需求，并且实现对数据的并行访问。分布式存储的特征是存储资源能够被抽象表示和统一管理，并且能够保证数据读写与操作的安全性、可靠性等各方面要求。最典型分布式存储技术有Google的分布式文件系统GFS和Hadoop的开源HDFS。

（4）数据隔离与保护

数据隔离与保护技术保证大量用户共享云存储中的统一存储资源时，每个用户只能访问自身存储空间中的数据，对存储空间的分配管理不影响其他用户的使用。

2 云存储作为物联网存储平台的可行性分析

物联网是一个物理世界与信息世界相连接的网络，通过将信息的载体扩展到“物”，在对这些信息分析处理的基础上实现对物理世界的智能化控制。因此物联网是一个规模庞大的信息计算系统，这个系统需要一个强有力的平台提供计算和存储服务来支撑其应用需求。当前云计算

模式兴起，并迅速从概念走向应用。云计算通过互联网将计算能力和存储空间有限的一系列IT设备整合成一个具有强大计算能力和海量存储空间的系统，其超大规模、高扩展性、高可靠性正好满足物联网对计算资源和存储资源的需求，成为支撑物联网应用的一个强有力的平台[5]。云计算面向物联网提供计算资源和存储资源两种服务，其中，存储资源服务就是云存储。云存储服务是整个云计算平台最底层的服务，是与云计算模式相匹配的存储模式，满足整个系统对数据的存取访问。下面从可用性、可靠性、数据资源高度共享3个方面分析云存储在物联网环境中的应用。

（1）可用性

从处理的数据对象上来讲，物联网传感层所采集的数据大都是非结构化的数据，例如图像、视频、文本、病例数据等，而云存储最擅长的数据处理对象便是这些数据。

物联网中的传感信息具有大数据量、实效性、高度并发等特征[6]。面对海量对象和海量汇聚的信息，快速存取成为物联网评价一个存储系统高可用性的一个关键指标。同时，高度并发的应用需求也要求系统必须以很小的响应时间来完成信息的快速处理和访问。云存储中的存储资源采用集中式的存放管理，而对这些资源的分配调度采用分布式。当用户提交数据访问请求时，云存储系统中便会有多个存储设备和服务器提供服务，大大提高了数据存取速度，同时，采用分布式存储架构可以实现对数据的并行读写，满足物联网中并发业务的数据存取需求。

物联网对存储资源的需求具有时空性，也就是不同的时间段和不同的地区对存储资源的需求是不同的。这就要求系统具有动态扩展存储空间和负载均衡功能。云存储系统采用弹性存储空间扩展技术和虚拟化技术，可以根据用户的需求或负载对存储空间大小动态伸缩，而这些过程对用户来说是透明的。

（2）可靠性

可靠性主要是从对数据存储的安全性方面来说的。物联网通过对感知数据的采集和分析处理来提供智能化服务，因此信息的可靠存储就更加关键。云存储通过以下两种方法来保证系统的可靠性：一是加强系统的容错性，增加备份数据；二是通过全网全资源监控管理来保障系统各环节的健壮性。

云存储提供多种级别的容错技术，如硬盘级、节点级和Domain/Site级的数据可靠性技术，可以被运用到物联网中以满足不同数据的不同存储需求。同时云存储中的每一份数据都是冗余存储的，数据可以根据用户的需求而创建不同数目的副本，并且这些副本是存储在不同地方的，这样可以提高系统的健壮性，当某些存储节点失效时不会影响整个系统的稳定性。

云存储的全网全资源管理的特性可以对全网资源的性能进行监控，使得系统可以据此快速定位故障并修复并根据资源使用状况来优化存储节点的性能从而保障整个系统各个环节的健壮性。

（3）数据资源共享

单一的物联网应用是物联网发展的必要过程，一个个单一的应用构成了未来物联网建设的基本单元[7]。通过对众多单一物联网应用的互联和集成能够提高对物理世界的管理水平，可以形成覆盖范围更广的未来物联网。这就要求海量数据的共享，通过数据共享，众多单一的物联网应用才能互联相互协作，从而为用户提供更好的智能化服务。云存储通过将收集到的海量感知信息按照应用需求统一存放在不同的数据中心中，这种集中存放的模式通过高速传输的互联网使得不同应用服务提供商之间的数据共享更为方便，可提高共享数据的访问速度。

从以上3个方面的分析可以看出，云存储解决了物联网所面临的海量数据存储这个难题，是物联网环境下一种比较好的存储方案。但是，物联网与云存储结合有一个前提条件是规模化，也就是说，只有当物联网的应用达到相当大的规模后才有必要采用云存储，对于一些小型的物联网应用例如家庭物联网应用则没有必要结合云存储。

3 云存储在物联网应用中所面临的安全挑战及安全技术

物联网中的应用都是数据密集型的，传感设备与存储平台之间、用户与存储平台之间和用户与传感设备之间时刻都在进行数据交互，一旦数据的丢失和损坏都将早成难以预料的后果。同时，由于云存储是通过虚拟化技术来按照用户需求来分配存储空间，实际上所有用户的数据都存放在一个相同的物理存储系统中，不再像传统存储系统一样有物理的隔离和防护边界。这种集中存放模式容易丧失不同的企业数据和用户数据在存储和传输过程中的保密性，造成商业信息和隐私信息的泄露。因此，云存储平台必须采取适当的安全策略来保证物联网中数据的完整性、保密性和不可抵赖性。

3.1 云存储基础设施安全

云存储基础设施安全的主要目的是保证数据存储的完整性和保密性[8]。下面主要从数据备份、数据检错和纠错、文件系统安全性、访问控制和身份鉴别等几个方面来阐述。

（1）备份和数据检错纠错

云存储中的物理存储设备都是一些比较廉价的商用设备，存储设备故障是一种正常现象而不是异常。通常采用的做法是冗余备份数据，并将数据存放在不同的数据中心中，以保证个别存储设备的故障不影响整个存储系统的可用性。系统能够迅

速发现错误并找寻备份数据来完成数据存取访问。同时，廉价的商用存储设备也要求系统具有良好的数据检错和纠错技术来保证数据的正确读写。

（2）文件系统安全性

文件系统是云存储系统中的一个重要组成部分。文件系统加密是实现存储系统安全最简单最直接的方法。文件系统的安全性一方面通过数据加密的方式来保证，同是也可采用当前常见的安全文件系统，主要有以下几种：Blaze 等人提出的加密文件系统（CFS）、Howaro 等人提出的一个附加安全措施的分布式文件（AFS）、Kaashoek 等人提出的 SFS-RO 系统、Wylie 等人提出的 PASIS 系统等保证文件安全性的技术方法[9]。

（3）访问控制和身份鉴别技术

访问控制和身份鉴别技术可以有效地控制用户对存储资源的访问，根据用户身份的不同，系统可以授予用户不同的访问权限，并设置相应的策略保证合法用户获得资源的访问权。这种策略可以将用户对存储系统的访问限制在一定的范围内，从而保证其他用户数据的安全性，防止越界访问。例如登录访问控制可以使有权用户能够登录到网络存储系统并获取存储资源，目录访问控制可以控制用户对目录、文件、存储设备的访问等。

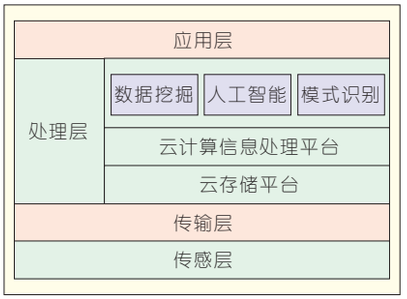
3.2 云应用安全

云存储应用安全主要是面向用户提供一些安全手段来保证用户数据在传输、交换和存取访问过程中的安全性，防止用户数据被非法访问和泄露。应用安全常采用的安全机制有以下几种：存储加密、交换加密、身份认证与访问控制和接口安全等。

存储加密在访问云入口对数据进行加密，保障传输的安全性，只有授权用户才能访问数据。在云存储平台中，各个用户的数据不是相互孤立的，各个用户之间需要时刻进行数据交换来满足物联网中的各种应用。交换加密保证了用户数据在交互过程中的安全性，交换加密常采用的技术手段是数字信封。身份认证与访问控制机制确保授权用户在自己的权限范围内进行数据操作，从而防止非法用户对数据的访问，也可以防止授权用户的越界访问。云存储根据物联网的应用需求不同而提供不同的应用接口，因此，接口安全可以有效的保证应用程序对存储资源的安全访问，多接口模式和加密技术可以有效的保证接口安全。

4 物联网环境下云存储系统的模型

基于云存储的物联网体系结构如图 2 所示。物联网环境下云存储平台的结构模型如图 3 所示。图 2 给出了云存储平台在整个物联网的体系结构中的位置，物联网整个体系结构从下到上依次可分为感知层、传输层、处理层、应用层 4 层[10]。感知层主要用来收集周围可被感知物品的信息，并将这些感知信息简单处理后通过各种接入网传递到传输层；传输层将融合后的感知信息传输到处理层，再将处理层的反馈信息传递到感知层的各个设备；处理层提供存储和处理功能，提供数据分析、局势判断和控制决策等处理功能，云存储便设立



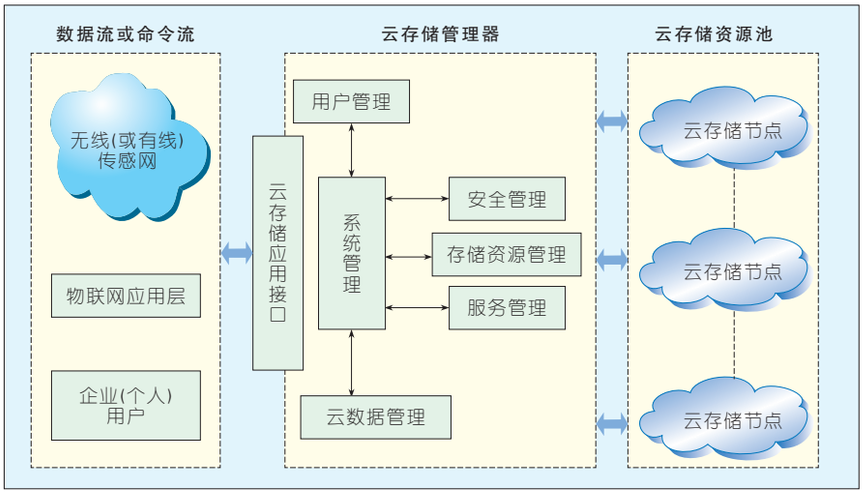
▲图 2 基于云存储的物联网体系结构

在这一层来提供存储服务来满足整个系统对数据的访问。最上层的应用层建立不同领域中的各种应用。

云存储接受传输层提交的感知数据，并采用合适的策略（如按照地理位置就近原则）存储数据，提供访问接口供云计算信息处理平台对数据进行分析处理，并将处理的结果存放起来供应用层访问。

图 3 中整个云存储平台由云存储管理器[11-12]和云存储资源池[13-14]两个部分构成：

云存储管理器是整个云存储平台的“大脑”，主要任务是负责存储资源的管理和调度。云存储管理器以系统管理为核心，以用户管理、安全管理、存储资源管理、服务管理和云数据管理等为子管理模块。这些子管理模块分别分布在各个不同的服务器节点上，接受系统管理模块的统一管理。



▲图 3 云存储平台结构模型

系统管理模块是云存储管理器中的核心部件，掌握着整个云存储系统的拓扑结构和底层操作系统以及所采用的文件系统的基本信息，协调各个子管理模块的工作。同时，系统管理模块制订各种各样的策略并将其分发到各个子管理模块中去，例如文件分块大小、存储分区大小、存储空间分配策略、冗余副本个数、节点故障处理、安全等级等。子管理模块根据这些策略实现自身功能。

存储资源管理模块实现存储资源的管理和分配，主要功能如下：监控整个系统中云存储节点的运行状态，及时发现故障节点并采取相应的策略选择新的存储节点来代替原有节点提供存储服务；发现新添加的存储节点，实现对其透明访问；采取合适的负载均衡功能保证系统的可靠性；将大的数据块划分为较小的数据并将这些小型数据块分配到不同的存储节点上，实现分布式存储。

云数据管理模块的主要功能实现对元数据的管理，元数据是存储文件基本信息的一种数据。当一个文件被创建的时候，文件的一些基本信息如文件名、文件大小、文件存储位置、文件访问控制等将被作为一个元数据项存放在元数据文件中。当有文件访问请求时，系统首先查找该文件的元数据信息，根据元数据信息找出该文件的存储位置，然后向相应的存储节点发出数据访问请求。

服务管理模块定义了不同等级的云存储服务和用户为此支付的费用，用户可根据自己的实际需求选择合适的存储服务；用户管理模块负责管理整个云存储系统中用户的基本信息，实现对用户的访问认证及访问权限的授予；安全管理模块实现对整个云存储系统的安全管理，采用各种安全技术如防火墙技术、入侵检测技术、加密技术等来保证整个系统的安全性。

云存储资源池是由多个云存储节点构成，这些云存储节点可能分布在不同地点。通过存储虚拟化技术将其整合为一大容量虚拟存储设备，用户可以像使用物理存储设备那样来使用。

5 结束语

云存储以其成本低廉、可扩展性高、易于管理等优点已成为未来存储发展的一种趋势。对于物联网应用服务提供商来说，云存储模式使其节省了构建存储平台所带来的昂贵的成本投入，应用服务商可以以较低的代价享受到先进的存储技术。数据的存储和管理工作完全由云存储服务商按照自己的要求来做，这样可以将更多的资金投入到本身的业务开发中去。物联网是一种数据密集型的信息系统，从传感层到应用层，时刻都有海量数据需要存取，同时对数据的敏感度比较高，一旦数据丢失或损坏，将严重影响系统的可用性。因此，云存储必须在安全性、可用性及可靠性等方面不断加以改进以适应物联网的应用需求。此外，物联网中云存储的大规模应用也不是一朝一夕能够实现的，随着物联网应用规模的不断扩大，云存储将经历着从为小范围物联网应用服务的私有云存储发展到为某个行业应用服务的行业云存储，最后发展到的各种云存储互联泛在阶段。

6 参考文献

[1] International Telecommunication Union. The Internet of things [R]. ITU Report, 2005.
[2] 拓守恒. 云计算与云数据存储技术研究 [J]. 电脑开发与应用, 2010, 23(9):1–3.
[3] 周可, 王桦, 李春花. 云存储技术及其应用 [J]. 中兴通讯技术, 2010, 16(4):24–27.
[4] 王德政, 申山宏, 周宁宁. 云计算环境下的数据存储 [J]. 计算机技术与发展, 2011, 21(4): 82–83.
[5] 赵均. 构建基于云计算的物联网运营平台 [J]. 电信科学, 2010, 26(6):48–52.
[6] 李玲娟. IoT的数据管理与智能处理 [J]. 中兴通讯技术, 2011, 17(1):38–41.
[7] 韩燕波, 赵卓峰, 王桂玲, 刘晨. 物联网与云计算 [J]. 中国计算机学会通讯, 2010, 6(2): 58–62.
[8] Security guidance for critical areas of focus cloud computing [R]. Cloud Security Alliance (CSA). 2011.
[9] 赵俊杰, 詹永照, 蔡涛.网络存储安全系统研究综述 [J]. 计算机应用与软件, 2008, 25(2): 272– 273.
[10] 孙利民, 沈杰, 朱红松. 从云计算到海计算:论物联网的体系结构 [J]. 中兴通讯技术, 2011, 17(1): 3–7.
[11] WU Jiyi, PING Lingdi, GE Xiaoping, et al. Cloud storage as the infrastructure of cloud computing [C]//Proceedings of the 2010 International Conference on Intelligent Computing and Cognitive Informatics(ICICCI '10), Jun 22–23, 2010, Kuala Lumpur, Malaysia. Los Alamitos, CA, USA:IEEE Computer Society, 2010:380–383.
[12] HUO Yanmei, WANG Hongyuan, HU Liang, et al. A cloud storage architecture model for data–Intensive applications [C]//Proceedings of the 2011 International Conference on Computer and Management(CAMAN'11), May 19–21, 2011, Wuhan, China. Los Alamitos, CA, USA: IEEE Computer Society, 2011: 4p.
[13] HE Qinlu, LI Zhanhuai, ZHANG Xiao. Analysis of the key technology on cloud storage [C]//Proceedings of the 2010 International Conference on Future Information Technology and Management Engineering(FITME'10), Oct 9–10, 2010, Beijing, China. Los Alamitos, CA, USA: IEEE Computer Society, 2010:426–429.
[14] HE Qinlu, LI Zhanhuai, ZHANG Xiao. Analysis of the key technology on cloud storage [C]//Proceedings of the 2010 International Conference on Future Information Technology and Management Engineering(FITME'10), Oct 9–10, 2010, Beijing, China. Los Alamitos, CA, USA: IEEE Computer Society, 2010:426–429.

收稿日期：2012–10–05

作者简介

杨继慧，甘肃教育学院毕业；工作于浙江理工大学；主要从事数字图书馆与信息存储方面的研究；已主持和参与完成科研项目7项，发表研究论文12篇。

周奇年，浙江理工大学信息学院副院长、教授，中国计算机学会会员；主要研究方向为云计算及物联网；已完成与合作完成国家基金项目3项；表发研究论文60余篇，其中大多被SCI、EI及CA等收录。

张振浩，浙江理工大学在读硕士研究生；主要研究方向为云计算及数据挖掘。

云计算在智能电网中的应用及其安全问题研究

Smart Grid Cloud Computing Applications and Security

中图分类号：TP393 文献标志码：A 文章编号：1009-6868 (2012) 06-0017-005

陈杰/CHEN Jie
张跃宇/ZHANG Yueyu
（西安电子科技大学通信工程学院，西安 710071）
（School of Telecommunications Engineering, Xidian University, Xi'an 710071, China）

摘要： 文章认为在智能电网中引入云计算，构建智能电网云可以为智能电网的发展提供有效支持。基于智能电网的特征，依托作为一种崭新的存储和计算模式的云计算技术，文章阐述了云计算技术如何为智能电网的数据存储和分析提供技术支持，设计了一种智能电网云计算的结构，分析了云计算技术给智能电网带来的诸多益处；并就智能电网云可能存在的安全威胁和相应的防范措施进行讨论。

关键词： 云计算；智能电网；智能电网云；安全威胁

Abstract: A smart grid cloud supports smart grid development. In this paper, we discuss how cloud computing—a new storage and computation model—provides technical support for data storage and analysis in a smart grid. We discuss the structure and benefits of a smart grid cloud. Security threats and prevention measures in a smart grid cloud are discussed.

Key words: cloud computing; smart grid; smart grid cloud; security threats

1 智能电网的特征

电网的智能化体现为：能够全面、及时地掌握电网运行的信息，综合各自动化功能系统对信息分析结果做出最优的反应。智能电网在大量交互式数据的基础上实现精细化、智能化的管理，贯穿了发电、输电、配电、用电的全过程，它通过智能终端将用户之间、用户和电力部门之间形成即时连接的网络，从而实现数据采集的实时、高速和双向，从整体上提高电网的综合运行效率，实现节能减排[1-3]。智能电网的特征概括起来主要有如下几点[4-5]：

（1）自愈

智能电网发生故障时，设备应根据故障类型及时发出信息并采取正确的措施，实现问题部件的隔离或恢复正常运行，最小化或避免用户的供电中断。智能电网通过实时地评估和自测，可以检测、分析、响应、恢复电力元件或局部网络的异常运行。

（2）安全

当物理系统或计算机遭到外部攻击时，智能电网能有效抵御由此造成的对电力系统本身和对其他领域形成的伤害。一旦发生中断，也能很快恢复运行。

（3）兼容

智能电网中不仅包括常规的远端集中式大型电厂，还有大量的分布式电源、电力电子元件和储能设备，这些分布式电源通常是新能源发电。通过在电源互联领域引入类似于计算机中“即插即用”技术（尤其是分布式发电资源），智能电网可以兼容包括集中式发电在内的多种不同的发电类型，甚至是储能装置。

（4）交互

传统电网中用户不参与电力系统管理，而智能电网中用户可根据电价高低选择性地用电。电网运行中与之交互的用户设备和行为被视为电力系统的完整组成部分之一，可以促使电力用户发挥积极作用，实现电力运行和环境保护等多方面收益。

（5）协调

智能电网应拥有成熟、完整的电力市场运行模式，以减小输电阻塞和其他限制，从而与电力批发和零售市场实现无缝衔接。

（6）高效

引入最先进的IT和监控技术优化设备和资源的使用效益，提高单个资产的利用效率，从整体上实现网络运行和扩容的优化，降低运行维护成本和投资。

由智能电网的特征可以看出：精确、快速、开放、共享的信息系统是智

基金项目：国家自然科学基金（61201132，61102056）

能电网的基础，也是智能电网与传统电网的最大区别[6]。由于云计算具有分布式的存储和计算特性，易于扩展和管理，特别适合解决智能电网信息系统所面临的一系列新问题。因此，在智能电网信息系统中引入云计算，可以为智能电网技术的发展提供有效的支持。

2 云计算技术

云计算是通过把大量高度虚拟化的资源管理起来，组成一个庞大的资源池，用来统一为云租户提供快速部署、易于管理、按需分配的云服务。目前，Amazon、Google、IBM、Microsoft、Sun等公司已纷纷建立并对外提供各种云服务[7-8]。根据美国国家标准与技术研究院（NIST）的定义，当前云服务可分为3个层次，分别是：基础设施即服务（IaaS）、平台即服务（PaaS）、软件即服务（SaaS）。

典型的云计算架构由3个层次构成：物理层、虚拟层和云服务层，如图1所示。其中物理层提供最基本的存储介质和计算能力，是云计算中心的根基；虚拟层通过对物理层设备的虚拟化，将物理层的计算和存储能力进行了分割，为按需分配的云服务提供了基础；云服务层通过对虚拟层设备的统一管理和调度，为用户提供IaaS、PaaS、SaaS的服务。

在云计算模式下，客户通过互联网对数据资产进行访问和使用，也就是说，客户数据资产的所有权和管理权已经分离，数据资产统一由云服务器进行管理。云计算数据存储系统的原理如图2所示。某个客户对具体数据资产的访问受到访问控制服务器的约束，数据搜索服务器提供搜索服务，使得客户可以得到相应的数据资产。基于第三方审计服务器，云存储系统可以向客户提供数据资产的完整性验证。

可见，云计算是一种基于网络的超级计算模式。本质上，云计算就是将数据、应用和服务存储在云端，充分利用一组内部互连的虚拟机组成具有强大计算能力的并行和分布式计算系统，以实现用户业务系统的自适应性[9-10]。基于云计算的这些特点，将其引入智能电网，建立电力系统的云计算体系，在电力系统广域网络硬件不变的情况下，最大限度地整合当前系统的数据资源和处理器资源，可极大地提高电网数据的存储、处理和交互能力。

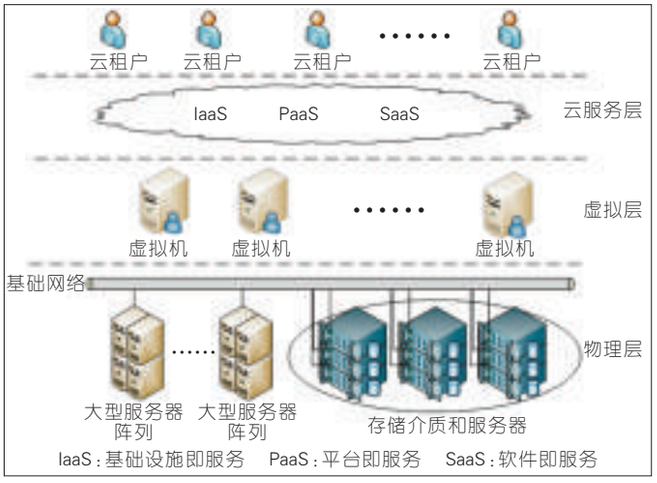
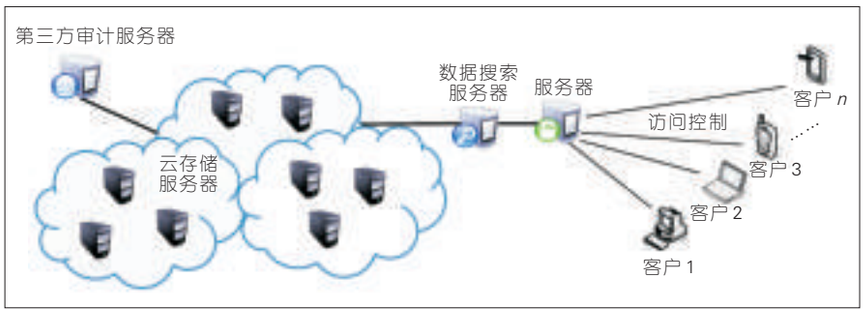


图1▶ 云计算体系结构

3 智能电网中的云计算技术

智能电网的信息种类繁多，用户的服务请求不断增多，数据呈现海量化，电力系统现有的数据库系统随着智能电网的发展已难以满足海量数据的存储要求。智能电网数据分布广泛，各类数据查询和处理的频度以及性能要求也不尽相同，此外，智能电网需要在海量数据的基础上，进行大规模的电力系统计算、分析、仿真、优化、设计和决策，电力系统现有的计算能力随着智能电网的发展将难以满足数据实时处理的需求。云计算采用分布式存储方式存储海量数据，并采用冗余存储和高可靠性软件的方式来保证数据的可靠性，它的数据管理技术能够高效地管理智能电网信息平台中类型和性能要求不同的各类多元数据，可以为电力系统计算提供高性能的并行处理能力，并提供并行编程模式，使并行计算的开发变得简单[11]。由此可见，在智能电网中引入云计算技术可以更好地解决电网海量数据的存储、管理和分析中所面临的问题。在下文中，我们将智能电网中引入的云计算称为智能电网云。

智能电网云以透明的方式向用户和电力系统应用提供各种服务，通过虚拟化的计算和存储资源池进行动态部署、动态分配/重分配、实时监控，向用户或电力系统应用提供满足服务质量（QoS）要求的计算服务、数



▲图2 云计算数据存储系统的原理

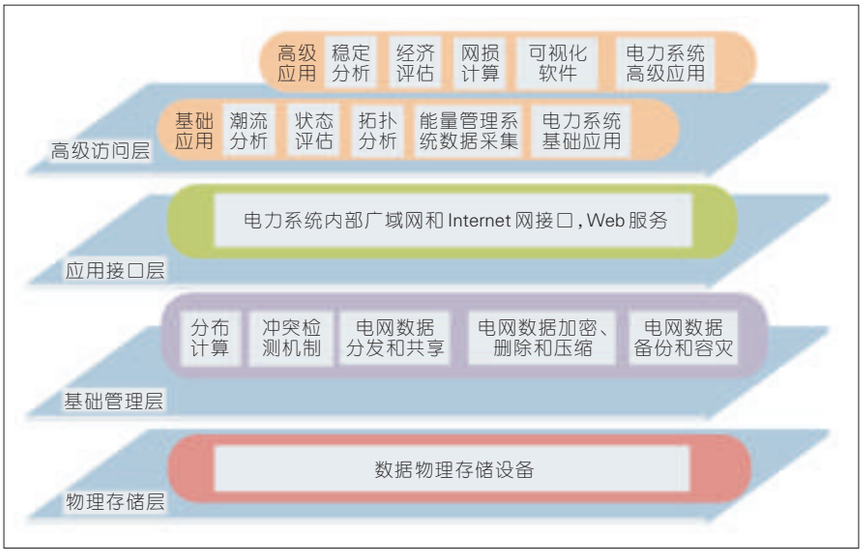
据存储服务以及平台服务。智能电网云体系结构可以分为4层，如图3所示。自下而上分别为：物理储存层、基础管理层、应用接口层、高级访问层。

电力系统内部通过集群应用、分布式计算等系统功能，将几乎所有网络和计算应用软件资源整合起来协同工作，共同对各级电网和终端提供数据储存和计算服务，形成一个电力系统内部的私有云[12]。因为电力系统分级管理的层次清晰，可以按级别建立主云和子云来提供资源分配，从而减少系统内部不必要的资源调配，使电力系统私有云的资源使用更为合理。

因为智能电网中新的设备越来越多，用户的信息数据也急剧增加，仅依靠电力系统私有云将难以从根本上解决海量的数据存储、管理、计算和分析所面临的问题。随着智能电网的发展，电力用户对供电服务的要求越来越高，希望享受更加个性的、多样的、便捷的、互动的服务，然而电力系统私有云并不能满足用户日益增长的服务需求。Sebnem等人提出了智能电网公有云的模型，完全依靠访问控制和分级管理来实现电网公有云的安全性，可以达到更好的可扩展性和经济性，相应的一些云服务甚至可以由第三方服务提供商来提供。

我们考虑将安全性要求较高的数据交由电力系统私有云存储和管理，对于安全性要求相对较低的数据可以考虑在公有云中存储和管理，并设置相应的访问权限。在电网计算量较大或部分电力系统遭到破坏出现紧急情况时，可以充分利用公有云来实现电网的稳定运行。

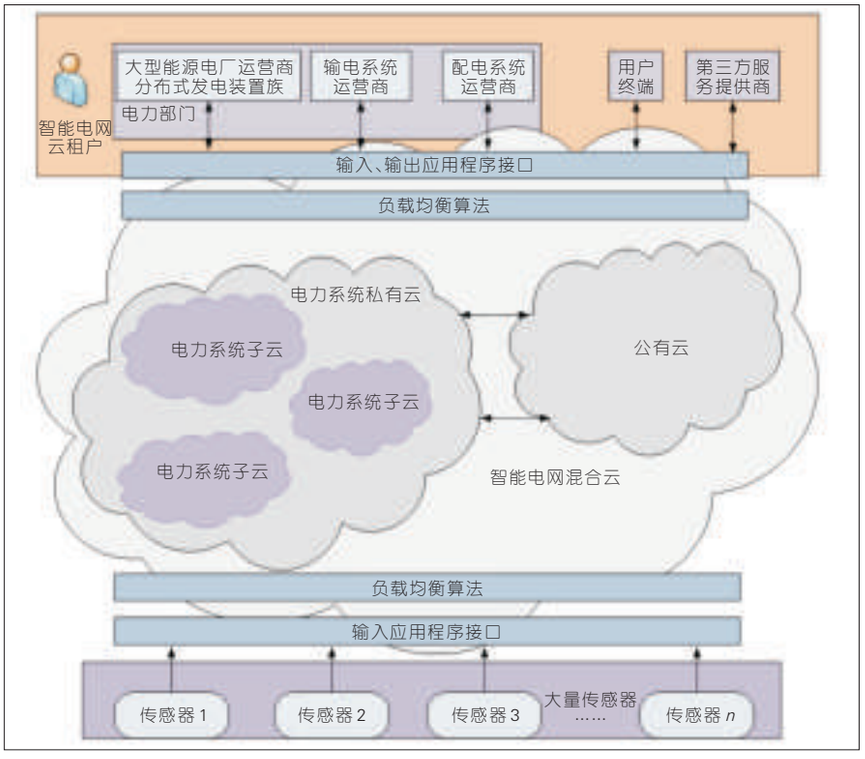
智能电网混合云的结构如图4所示。大量的传感器通过输入应用程序接口将采集到的不同类型数据传送至智能电网混合云中进行存储和管理。电力部门、用户终端和第三方服务提供商相当于智能电网的云租户，通过输出应用程序接口可以接收来自电力控制系统的控制信息，并根据具体的权限可访问相应的数据信息；智能电网云租户通过输入应用程序接口传送各类数据信息至智能电网云中进行存储、管理和分析，并发送相应的服务请求信息。根据数据类型、数据来源以及数据检索的存储模式，智能电网混合云设置负载均衡算法，以尽可能降低网络负载。



▲图3 智能电网云结构层次模型

智能电网的云租户主要有3类：用户、电力部门和第三方服务提供商[13]。

（1）用户可以分为：住宅用户、商



▲图4 智能电网混合云结构

业用户和工业用户。一般在每个用户端安装智能电表或在一片区域的末端也会安装智能电表，形成相应的家域网和区域网。智能电表用于收集电能使用数据和给相应的智能电器设备传送控制信号。智能电网信息网络有相应的软件算法，可以根据用户选择的参数和电力部门给出的需求响应信息来进行分析，最优化电能的使用。用户之间通过智能电网云可以共享由电力部门提供的一些附加信息，并可以根据已有的使用数据得到电网运行的预测模型，从而更好地规划用电方案。

住宅用户的信息包括：用户的身份信息、用电数、已安装的智能电器设施和电动汽车、社交网络活动以及用户的用电特点等。这些信息的泄露将对用户身份识别信息和行为特征的安全性构成威胁。例如：在电动汽车入网中需要保护车辆的位置信息、用户身份、使用充电电池类型、用户支付信息、充电站处理信息等。如果用户个人身份和电动汽车停放位置等信息泄露，攻击者还可以根据大量的用户数据，例如充电时间、充电地点以及充电量等信息推断出用户的活动范围、行驶路径和行驶距离等生活习惯信息。

工业和商业用户的附加数据包括机器使用数据、生产计划、销售活动、电动汽车车队运营等信息。而这些信息都是十分敏感的数据，涉及到工业和商业的一些机密信息。例如：通过分析某工厂计划的能源使用信息可以预测工厂的产量。

因此，智能电网云需要对用户数据进行隐私保护。根据数据信息的安全级别以及用户类型，设置相应的访问控制列表。用户甚至可以设置个人访问控制列表，在允许范围内对自己的个人数据进访问。访问控制列表可以动态变化，这就需要经常更新访问控制列表，由智能电网云提供资源的可扩展性。智能电网云在多个虚拟机上并行对数据进行加、解密，需要保证分布式资源接入控制的权限保持一致。

（2）电力部门是智能电网的核心部分，它包含很多职能，例如：智能电网的稳定运行，发电、输电和配电，维护客户满意度，遵守各种管理规范等。智能电网中电力部门利用云基础架构存取和处理大量数据，它们包括来自智能电表和设备以及智能电网中部署的传感器传输的大量直接数据，也包括由直接数据分析预测得到的间接数据，而多样性的数据源需要安全管理。根据用户角色的不同，数据还可以分为保密数据和公开数据，如果管理不当可能导致保密数据公开化，从而泄露用户的隐私信息，所以智能电网云数据平台需要支持对不同的信息数据进行安全采集和存储。

（3）随着智能电网的发展，需要开发各种类型的应用服务，这些都由第三方服务提供商来提供，因而需要通过电力部门云架构在不可信环境中提供服务，这必将增加第三方服务提供商的处理安全和隐私保护问题的困难程度。

由上面的分析可以看出：云计算为智能电网海量数据的存储、管理、分析和决策提供了解决方案，智能电网的安全稳定发展离不开云计算的安全。

4 智能电网中云计算的安全问题和防范措施

云计算能够给智能电网持续稳定发展带来诸多益处，但也相应地引入了一些新的安全威胁。下面分别从物理层、虚拟层和云服务层3个层面讨论所面临的威胁。

（1）物理层安全威胁

在智能电网云中，云租户数据都要在云计算物理层进行存储和计算，这将导致用户不信任云服务商对其数据私密性的保护；同时，某些云租户希望通过服务的漏洞窃取其他用户的数据信息。由此，物理层引入了如下新的风险：智能电网云服务商窃取云租户数据，侵犯云租户的隐私或者破坏云租户数据的保密性；恶意云租户非法访问其他云租户的数据，破坏其他云租户数据的保密性及侵犯他人的隐私；云租户可以访问其他云租户的历史遗留信息，导致云租户数据私密性遭到破坏；恶意云租户渗透电网云计算中心的虚拟机或宿主机，进而越过虚拟机的束缚破坏其他虚拟机或宿主机；物理环境不安全，可能对智能电网云基础设施造成破坏，进而导致电网云基础设施的不安全；传统互联网的攻击手段对智能电网云的物理网络同样构成安全威胁。

（2）虚拟层安全威胁

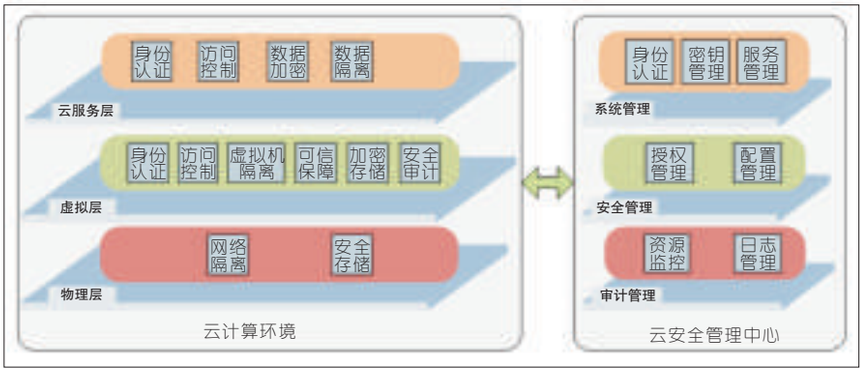
虚拟控制器负责对虚拟机的管理，拥有对虚拟机的所有控制权，虚拟机主要为云租户提供按需分配的计算和存储资源。云计算共享资源的特点导致虚拟层也面临如下安全威胁：虚拟机之间的非授权通信，造成云租户数据的机密性或完整性遭到破坏；攻击者利用虚拟机逃逸技术，可以获取整个智能电网云计算中心的数据；宿主机通过控制虚拟机，可以任意窃取虚拟机的网络数据；虚拟机通过控制其他虚拟机，可使数据的机密性和完整性遭到破坏；虚拟机强制占用一些资源，从而使得其他虚拟机的服务遭到拒绝；智能电网云服务商维护人员通过虚拟机窃取云租户隐私。

（3）云服务层安全威胁

由于智能电网云为不同云租户提供服务，因此云服务层需要满足云租户和云租户资源的映射关系，并且需要处理不同云租户的所有数据，这就给智能电网云引入了一些新的安全威胁：云租户勾结电网云服务商修改云服务的相关配置，使得非法云租户可以访问其他云租户的数据；恶意云租户利用电网云服务可能存在的安全漏洞，窃取电网云服务的权限，从而获得穿越该层的所有云租户数据；电网云服务商可以通过提供的服

务，窃取云租户的数据信息，破坏用户数据的私密性；传统应用服务的安全攻击，如结构化查询语言（SQL）注入、网页挂马等，同样威胁电网云服务的安全。

根据以上的安全风险分析，为保证智能电网云计算各参与方的安全需求，需要采用相应的安全机制，构建智能电网云计算环境下的安全体系。针对不同的层次，应采取合理的安全防范措施。智能电网云安全保障体系结构如图5所示。



▲图5 智能电网云安全保障体系结构

（1）物理层

物理层采取隔离措施，限制不同虚拟机对智能电网云基础网络访问的权限；通过数据加密等手段，确保系统内的数据资源所在的存储空间被释放或重新分配给其他用户前得到完全清除。

（2）虚拟层

虚拟层对登录虚拟机的用户身份进行真实性鉴别；实现电网云服务商对资源的访问控制，保证虚拟机中数据免遭非授权访问；对不同虚拟机进行隔离；保障宿主机中程序等在可信环境下运行，防止遭到恶意代码等破坏；实现用户数据的加密存储；通过安全审计，记录智能电网云服务商的非法访问行为。

（3）云服务层

云服务层对用户身份进行真实性鉴别；实现用户对电网云服务资源的访问控制；通过加密技术保证用户数据的私密性；实现对电网云应用中不同用户数据信息的有效隔离；实现对用户行为的安全审计。

（4）安全管理层

安全管理层实现对智能电网云计算环境中不同用户的身份管理、密钥管理，以及电网云服务商的云服务管理；完成云环境下的授权、安全配置等管理；实现整个智能电网云环境中监控及审计管理。

5 结束语

智能电网是当前全球电力工业关注的热点，引领了电网未来的发展方向。随着智能电网技术的不断发展，电网数据和信息量急剧增加，现有的电力系统已难以满足海量数据的存储、管理、分析和决策。在智能电网中引入云计算，构建智能电网云可以为智能电网技术的发展提供有效的支持。本文通过介绍智能电网的特点，分析了智能电网所面临的挑战，云计算技术作为一种崭新的存储和计算模式，为智能电网中的数据存储和分析问题的解决提供了机遇。文章给出智能电网云计算的结构以及云计算给智能电网带来的诸多益处，并就智能电网云可能存在的安全威胁和相应的防范措施进行讨论。

6 参考文献

[1] RUSITSCHKA S, EGER K, GERDES C. Smart grid data cloud: A model for utilizing cloud computing in the smart grid domain [C]// Proceedings of the 1st IEEE International Conference on Smart Grid Communications (Smart Grid Comm'10), Oct 4–6, 2010, Gaithersburg, MD, USA. Piscataway, NJ, USA: IEEE, 2010:483–488.
[2] SHARGAL M, HOUSEMAN D. The big picture of your coming smart grid [EB/OL]. [2009–03–05]. http://www.smartgridnews.com/artman/publish/commentary/The_Big_Picture_of_Your_Coming_Smar_Grid–529.html. 2009.
[3] 潘睿, 刘俊勇, 郭晓鸣. 电力系统云计算初探 [J]. 四川电力技术, 2010, 33(3):71–76.
[4] COLLIER S E. Ten steps to a smart grid [J]. IEEE Industry Applications Magazine, 2010, 16(2):62–28.
[5] 孙晶. 智能电网及其通信技术 [J]. 电力系统通信, 2010, 32(12):1–7.
[6] 江道灼, 申屠刚, 李海翔, 等. 基础信息的标准化和规范化在智能电网建设中的作用与意义 [J]. 电力系统自动化, 2009, 33(20):1–6.
[7] IBM blue cloud solution(in Chinese)[EB/OL]. [2011–05–05]. http://www–900.ibm.com/ibm/ideasfromibm/cn/cloud/solutions/index.shtml. 2011.
[8] BARROSO L A, DEAN J, HOLZLE U. Web search for a planet: The Google cluster architecture [J]. IEEE Micro, 2003, 23(2): 22–28.
[9] DIKAIAKOS M D, KATSAROS D, MEHRA P, et al. Cloud computing: Distributed Internet computing for IT and scientific research [J]. IEEE Internet Computing, 2009, 13(5):10–13.
[10] 陈全, 邓倩妮. 云计算及其关键技术 [J]. 计算机应用, 2009, 29(9): 2562–2567.
[11] 王德文, 宋亚奇, 朱永利. 基于云计算的智能电网信息平台 [J]. 电力系统自动化, 2010, 34(22):7–12.
[12] 余勇, 林为民, 邓松, 等. 智能电网中云计算应用及安全研究 [J]. 信息网络安全, 2011(6): 41–43.
[13] SIMMHAN Y, KUMBHARE A G, CAO B, et al. An analysis of security and privacy issues in smart grid software architectures on clouds [C]//Proceedings of the IEEE 4th International Conference on Cloud Computing(Cloud'11), Jul 4–9, 2011, Washington, DC, USA. Washington, DC, USA: IEEE Computer Society, 2011:582–589.

收稿日期：2012–10–05

作者简介

陈杰，西安电子科技大学通信工程学院副教授、硕士生导师；主要研究领域为密码算法分析、智能电网安全、信息及网络安全。

张跃宇，西安电子科技大学通信工程学院副教授、硕士生导师；主要研究领域为物联网技术、无线网络安全。

云灾备中系统级管理技术的关键问题

System-Level Management Problems in Cloud Disaster Backup and Recovery

中图分类号：TP393 文献标志码：A 文章编号：1009-6868 (2012) 06-0022-004

姚文斌/YAO Wenbin
叶鹏迪/YE Pengdi
（北京邮电大学 灾备技术国家工程实验室，北京 100876）
(National Engineering Laboratory for Disaster Backup and Recovery, Beijing University of Posts and Telecommunications, Beijing 100876, China)

摘要： 文章指出云计算技术的核心在于对计算资源、存储资源和传输资源的系统级管理；云灾备的核心则在于利用云中海量资源，为用户提供系统级、数据级和业务级的灾难恢复服务和持续性服务。为了解决云灾备遇到的管理方面的挑战，文章有针对性地对云灾备管理中的最大化收益管理、高可用服务管理、业务连续服务管理、安全保障管理4个方面进行了分析和探讨。

关键词： 云灾备；云计算；服务；系统级管理

Abstract: The core of cloud computing is system-level management of computing, storage, and transmission resources. Cloud disaster backup and recovery requires a massive number of resources in the cloud to provide system-level, data-level, and business-level disaster backup, recovery, and durative services. In this paper, we discuss problems related to maximizing revenue and managing highlyavailable services, managing business continuity, and managing security protection.

Key words: cloud backup and disaster recovery; cloud computing; service; system-level management

云计算技术作为一种可以有偿提供IT服务的模式得到了产业界的大力宣传和推广。目前，对于云计算还没有一个确定性的定义，通常云计算被认为是一种基于IT基础设施的服务模式，即运营服务商通过将大量的异构设备资源统一组织、集中管理，将计算、存储和传输作为销售资源对用户提供的一种有偿服务。

随着云计算技术的发展，云灾备技术应运而生，正逐步成为云服务的一项基础性服务。

云灾备，即利用云环境中的资源冗余特性，为用户提供容灾抗毁能力的一种保护用户信息的服务模式。云灾备为灾备技术的发展提供了更广阔的发展空间。

基金项目：基金项目：国家发改委09信息安全专项项目中央高校基本科研业务费专项（BUPT2011RCZJ16）

1 云灾备概述

1.1 云计算技术

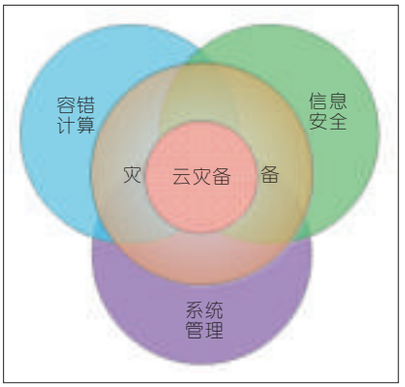
云计算技术是涵盖了网格计算、分布式计算、并行计算、效用计算、网络存储、虚拟化、负载均衡等诸多技术并发展融合的产物，是诸多技术的综合。云计算环境具有资源众多（包括设备、软件、网络等）、信息高度冗余、集中式管理、用户透明使用等一系列特点。

1.2 灾备技术

灾备，即灾难备份与恢复，信息系统灾备涵盖了灾难前的冗余与备份、灾难发生时刻的应急响应和灾难发生后的系统恢复。历史上的灾备主要是容错计算的一个研究方向，现在的灾备则是容错计算、信息安全和系统管理3个研究领域的综合，其总体架构如图1所示。

具体说，灾备信息系统可以理解为是以存储系统作为基本支撑系统、以网络作为基本传输手段、以容错软硬件技术为直接技术手段、以管理技术为重要辅助手段的综合系统。其技术本身至少包含了容错技术、信息安全技术、应急管理技术、存储技术、传输技术等多方面技术的综合，因此，灾备技术难以简单地归纳到计算机领域、通信领域、应急管理领域等任何一个单一领域范畴。

信息系统灾备是一门综合工程，其最大挑战是从上述研究方向中选



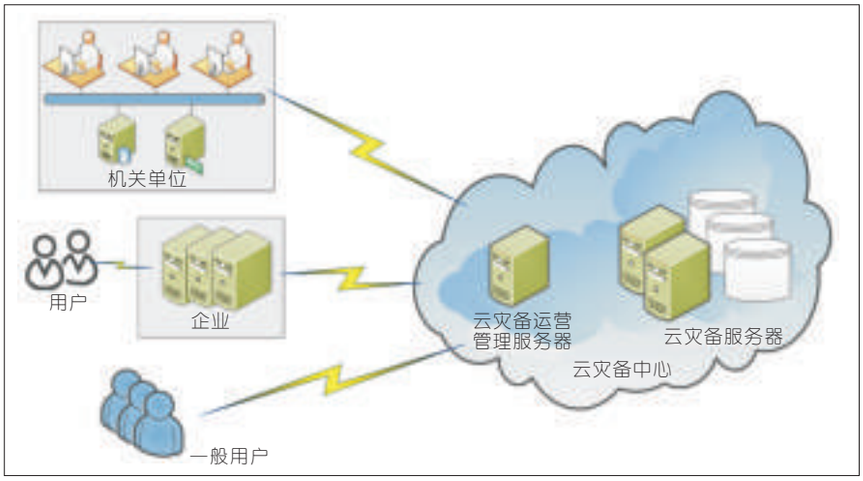
▲图1 灾备技术总体架构

择和优化合理组合来应对多样的用户需求，也就是说仅仅发展某一方面技术并不能对整体发生根本性的作用，行之有效的灾备管理需要企事业单位重视每一个环节，需要很高的人力、物力投入。

1.3 云灾备架构

云灾备，顾名思义，就是基于云环境下的灾备服务，即由用户付费使用服务提供商提供的灾备服务。

云灾备涵盖了云计算、容错计算、可信计算、可生存技术等技术的综合。从用户的角度来看，云灾备结构示意图如图2所示。



▲图2 云灾备示意图

作为云灾备服务的使用者，可以不用关心数据备份和系统备份、业务连续规划、人员架构、通信保障、危机公关、灾难恢复规划、灾难恢复预案、业务恢复预案、紧急事件响应、第三方合作机构和供应链危机管理等一系列应对灾难的复杂技术和管理问题。使用者所看到的就是随时随地可以接入、有无限的存储可供使用、有无限的计算能力、提供安全可靠的灾备服务，所做的就是按实际使用情况计量付费。

云灾备系统整体架构可划分为4个层次，自底向上依次是：设备层、统一管理层、服务管理层以及业务层。云灾备系统架构如图3所示。

设备层是云灾备系统中基础部分，它由海量服务器、存储设备、网络设备互连而成。这些设备不但数量庞大，而且可能分布在不同地域，彼此之间通过广域网、互联网或者光纤通道网络连接在一起。

统一管理层通过虚拟化技术、集群技术等方式将底层的海量设备整合成一个资源池，实现对硬件设备的集中管理、逻辑虚拟化管理、硬件设备的状态监控和故障维护以及资源的动态扩展。

服务管理层为上层不同服务提供统一的公共管理接口，主要负责对用户信息、数据副本、容灾策略、数据安全等进行管理，将底层资源和上层灾备业务无缝衔接起来，对外提供多种服务。

业务层是云灾备系统中直接面向用户、可以灵活扩展的部分。根据用户需求，业务层提供相应的服务，如数据容灾、应用容灾。其中，数据容灾指通过定期备份、持续数据保护等技术手段将用户本地关键应用数据复制到云中；应用容灾是在备份中心建立一套完整的与本地生产系统相当的备份应用系统（可以是互为备份），在灾难情况下，云灾备中心的系统可以迅速接管业务运行。

1.4 云灾备与云计算的关系

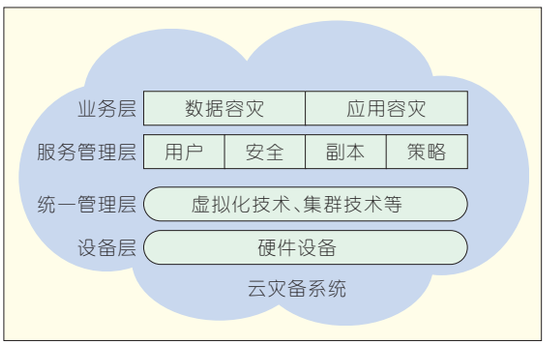
在云计算模式下，我们认为：云灾备将成为云计算环境的基础性服务。理由如下：

- 云计算环境中存在着海量的服务器、存储设备和多传输路径，这种资源的冗余性为容灾的设计与实施提供了广阔的物理容灾基础。
- 在云计算环境中，为了能够为用户提供随时随地的快速服务，需要在云系统内根据不同的调度策略，提供多数据版本，这些用于快速服务的多数据版本为灾备提供了天然的软件容灾基础。
- 为了便于管理，云环境中普遍使用集中式管理和虚拟化技术，这为云灾备提供了良好的逻辑实现基础，使得容灾的实施变得更加简单。
- 云灾备事实上是一种第三方灾备的安全保护模式，用户可以利用服务提供商的优势技术资源、丰富的灾备项目经验和成熟的运维管理流程，快速实现客户的灾备目标，降低客户的运维成本和工作强度，降低灾备系统的总体拥有成本。

云灾备的上述特性使得灾备整体门槛大大降低，使得灾备技术从传统的主要应用于“高精尖”的“贵族”应用逐步成本下移，形成了“贵族技术平民化”。

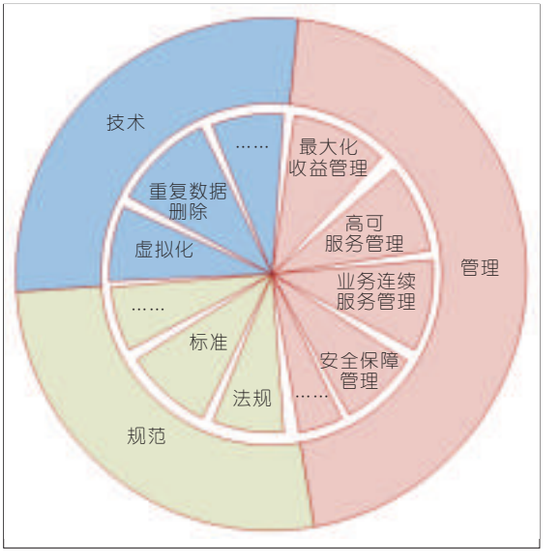
2 云灾备管理存在问题

虽然云灾备服务能够快速实现



▲图3 云灾备架构图

用户的灾备目标，降低用户的运维成本和工作强度，降低灾备系统的总体拥有成本，而且当前云计算平台也已经包含了很多可以支撑灾难恢复的有用特性，但对云灾备服务的提供商来说，依然存在诸多技术和管理方面的挑战。云灾备管理存在的主要问题如图4所示。这里，分别从最大化收益管理、高可用服务管理、业务连续服务管理、安全保障管理4个分别阐述云灾备管理中遇到的问题。



▲图4 云灾备管理存在的主要问题

2.1 最大化收益管理

所谓的最大化收益服务[1]是指在日常应用过程中，由于灾难是一种突发性的小概率事件，而信息化系统的高可用性则是一种常态的大概率要求，这就意味着灾备服务提供上需要做好专属容灾设备、软件和管理开销在日常中的高效率问题。简单地说，灾备服务就如同用户的保险，只有当灾难发生了，其价值才能真正体现。那么就需要在没有灾难的时候，合理利用冗余资源实现服务提供商效益最大化。最大化收益服务是服务提供商关注的最重要的核心问题之一。

云灾备服务提供商需要提供服务的客户包括非容灾需求客户、部分容灾需求客户和容灾需求客户。客户的服务等级还可以更进一步划分为一般客户、VIP客户、重点客户等多个等级。从管理的角度来说，云灾备服务管理的目标就是将IT服务资源在规定的时间以规定的质量等级合理地分配给用户。云灾备服务管理的核心在于资源的分配和调度。

通常，云灾备服务提供商同时也提供传统的云计算服务，并将资源租给客户用作非容灾目的使用。云计算系统在有新的虚拟机或者网络资源请求时，通常会尽力提供服务，这在单纯的云计算服务体系中来说没有问题。但在灾难备份的应用环境中却有可能行不通：由于保证灾备资源在指定的恢复时间目标（RTO）内能够提供服务是更加重要的事情，那么在这种情况下，云灾备服务提供商需要向非灾备用户提供资源服务，并且在高优先级的灾备用户需要时，还应按需、动态地调整系统各级资源来提供灾备服务；同时，在没有容灾服务要求时，系统也能够从空闲的灾备服务器上获得收益。

Amazon使用“现买现卖”的价格机制对这种按需访问资源的服务模式进行收费[2]。其云服务弹性计算云（EC2）针对那些已知其虚拟机需要运行很长时间的用户，支持保留其虚拟机实例，所以其价格机制是以适中的年费和更便宜的按小时计费的方式来反映此特点的。然而对灾难备份来说，用户系统需要多久才能完全修复是难以估计的，因此也很难估计灾备站点的虚拟机实例需要运行多久。同时，价格机制还必须设计用来保证允许“优先资源”在规定的时间可用，所以，其价格机制可以比基本云计算服务的按小时计费要高，这能够向用户提供更好的容灾服务，同时也能增加云服务提供商的收益。

2.2 高可用服务管理

云灾备服务提供商提供服务的前提是能够为用户提供持续不间断的系统功能服务，因此，其IT系统必须具备高可用的特性，高可用性管理自然也成为其关注的核心。

一般说来，云灾备服务提供商选择让多个同等级的用户共享其海量资源池。海量资源池可以在地理上广泛分布，也可以是在某地固定集中。对于地域集中的资源池的做法的逻辑核心是大量系统设备同时发生故障的概率远远低于单台设备发生故障的概率。因此，假设不是所有用户都会同时发生故障，可用的空闲设备往往会少于所有用户的最高需求，用户可以感觉不到发生故障。另外，这样做的好处是可以提高系统响应的效率、降低企业成本。

然而，对于灾难来说，上述观点并不适用。通常的自然灾难、设备大规模损毁灾难、战争威胁等灾难的发生，往往导致区域性的灾难发生。例如，供电网络故障、或者一场自然灾害比如洪水都会导致一个地区大范围用户对其灾备站点同时进行故障切换。一旦上述事件发生，即使对于拥有海量资源的云灾备服务，从管理上来说，依然是一个巨大的挑战。

具体说来，为防止任何一个数据中心出现这种因并发故障而导致的压力陡增，云服务提供商需要实现一种方法将其用户分流到不同的云数据中心来减小潜在的冲突，比如，很多同一地区的用户应该备份到不同的云数据中心。然而用户与云数据中心之间的延迟等限制又会使得选址问题更加复杂化。为了更好的解决这些问题，云服务提供商需要采用风险模型来解决以下两个问题：估计如果一个数据中心某一组用户提供服务，应该准备多少服务器以供使用；怎样将一个地区的用户分配到不同的数据中心中去来实现风险均摊。

同时，当任何一个数据中心因为

并发错误而产生负载压力时，还需动态地将其中一组用户迁移到其他可以使用的站点上去。为了无缝地完成这些任务，云服务提供商需要将其所有数据中心中的资源整合成一个资源池协同对用户提供服务[3-4]。

实际上，现有的数据中心往往是一个个孤立的实体，然而数据中心之间的存储和计算资源的移动或复制却是件十分重要的事情。在未来的云架构将可以依靠网络虚拟化来为数据中心提供无缝连接，而且广域网虚拟机和存储迁移将令跨数据中心的资源管理成为可能。

2.3 业务连续服务管理

在实际应用中除了数据容灾，还需要业务灾备，即将数据复制到云端的同时，也将受保护的应用系统的状态复制到云端，当灾难发生时可以立即切换到云端的应用系统运行，保证业务连续性。

通过虚拟机的迁移或复制可以实现用户应用系统的迁移。迁移方式既可以是用户应用系统到云灾备中心的迁移，也可以是云灾备中心内部的虚拟机池之间的迁移，以实现根据前端用户的需求自动地调节云灾备服务提供商有限的硬件与软件资源，动态地、弹性地反应前端业务对灾备的需求。另外，在故障处理完毕后，将应用移回其原来的站点，来简化故障恢复过程，而且这也是一个实现计划停机维护的实用方式。

因此虚拟机的迁移或复制是一个极其有价值的特性。然而，当前的云计算平台基本都不支持虚拟机实例迁入或者迁出云环境。为支持这种特性，云灾备服务提供者需为其用户开发附加的虚拟层功能，以支持虚拟机迁移或复制的特性，同时针对广域网环境下带宽资源有限的问题，还需要优化的数据迁移技术。

2.4 安全保障管理

虽然现在云计算平台已经包含了很多可以支撑灾难恢复的有用特性，但如果想要将灾难恢复作为一个云服务来提供，还需要满足一些额外的需求。

云计算平台的公有特性是一些用户担心的地方。实施云灾备以后，用户的数据存放在由服务提供商管理和维护的服务器上，不再受用户的直接控制。鉴于云计算环境与传统灾备系统之间的内在差异，使得云计算环境下的攻击模式和破坏行为更为多样和严重，传统的网络安全与隐私技术已不能完全适应具有多样、随机、动态等特点的云计算环境。云安全和隐私面临如安全域的模糊性、用户/数据的动态性、数据安全性、服务的隐私性、服务的可审计性等新挑战。这些都增加了用户数据的潜在风险，如云服务提供商的系统故障、服务器被攻击、云服务提供商内部人员的泄密或蓄意破坏等，有可能造成用户数据的泄密、损坏或丢失。用户难以衡量服务提供者所提供灾备服务的安全等级、内容如何分布、存储地点如何选择等，因此对云灾备以及灾难恢复存在疑虑。由于用户不知道需要做什么，只能被动信任云存储提供者。

因此，为了让一个企业愿意在发生灾难时，将其信息化系统从其私有的数据中心迁移到云中，需要对存储、网络及虚拟机的安全提供强有力的保障。同时，云灾备服务提供商还需要保证正在云中运行的应用系统性能不会被其他用户应用系统的灾难恢复活动影响。

3 结束语

本文并从最大化收益管理、高可用服务管理、业务连续服务管理、安全保障管理这4个方面讨论了云灾备服务管理在实现当中所面临的相关问题。信息技术未来发展趋势必然是以信息服务模式满足多样用户应用需求，因此，云计算模式作为典型应用模式得到了大力发展。正如同计算机系统的可靠性是保证其广泛应用的前提一样，未来云灾备技术必然成为云环境中的基本功能和结构。但如何实现云灾备，妥善解决服务过程中所面临的问题，尤其是管理问题，并使其广泛应用，还需要更多的研究和努力。

4 参考文献

[1] WOOD T, CECCHET E, RAMAKRISHNAN K, et al. Disaster recovery as a cloud service: Economic benefits & deployment challenges [C]//Proceedings of the 2nd USENIX Workshop on Hot Topicsin Cloud Computing (HotCloud'10), Jun 22–25, 2010, Boston, MA, USA. Berkeley, CA, USA: USENIX Association, 2010.
[2] AMAZON elastic compute cloud(Amazon EC2)[EB/OL]. [2011–10–05]. http://aws.amazon.com/ec2/. 2011.
[3] WOOD T, GERBER, RAMAKRISHNAN K K, et al. The case for enterprise ready virtual private clouds [C]//Proceedings of the 1st USENIX Workshop on Hot Topicsin Cloud Computing(HotCloud'09), Jun 14–19, 2009, San Diego, CA, USA. Berkeley, CA, USA: USENIX Association, 2009.
[4] BUYYA R, RANJAN R, CALHEIROS R N. InterCloud: Utility–oriented federation of cloud computing environments for scaling of application services [C]//Proceedings of the 10th International Conference on Algorithms and Architectures for Parallel Processing (ICA3PP'10): Part 1, May 21–23, 2010, Busan, Republic of Korea.LNCS 6081. Berlin, Germany: Springer–Verlag, 2010:13–31.

收稿日期：2012–10–05

作者简介

姚文斌，北京邮电大学教授、博士导师，北京邮电大学灾备技术国家工程实验室总工程师，教育部高等学校计算机科学与技术专业教学指导分委员会专家组成员，中国通信学会云计算专家委员会委员，中国计算机学会容错专委会成员；主要研究领域为灾备技术、信息安全、可信计算等；主持和参加了基金项目20余项；已发表学术论文60余篇(SCI/EI检索40余篇)，申请国家发明专利20项(6项已授权)。

叶鹏迪，北京邮电大学灾备技术国家工程实验室在读博士研究生；主要研究领域为云存储和灾备技术。

一种基础设施云系统——YUN

YUN: An Infrastructure-as-a-Service Cloud System

中图分类号：TP393 文献标志码：A 文章编号：1009-6868 (2012) 06-0026-004

刘川意/LIU Chuanyi
林杰/LIN Jie
（北京邮电大学 可信分布式计算与服务教育部重点实验室，北京 100876）
(Key Laboratory of Trustworthy Distributed Computing and Service Ministry of Education, Beijing University of Posts and Telecommunications, Beijing 100876, China)

摘要： 文章设计了一个虚拟化平台层——YUN系统，介绍了YUN系统的设计目标、系统架构及主要子系统的功能。YUN系统是一种基础设施云系统。YUN系统管理和虚拟化异构、分布的物理资源（包括服务器设备、网络设备、存储设备），并以即付即用的方式将这些资源按需提供给用户使用；通过虚拟化和网络技术，将操作系统功能边界扩展到网络侧，使得终端侧和网络侧能够在一个统一的架构内进行计算和资源的配置、调度、管理，实现网络化操作系统充分、合理地使用和共享网络系统内的资源。

关键词： 云计算；基础设施即服务；虚拟机

Abstract: In this paper, we design a virtualization platform and introduce the design goal, system architecture, and function of the main subsystems in YUN—an infrastructure-in-a-service cloud system. YUN virtualizes and manages heterogeneous and distributed physical resources such as server, network, and storage devices. Using a pay-as-you-go model, these are exported as on-demand services to customers. Virtualization and network technology extend the functionality of an OS to the network side, enabling both the terminal and network sides to allocate and schedule computing and other resources in a unified manner. Networking operating systems can use and share resources reasonably within the network system.

Key words: cloud computing; infrastructure as a service; virtual machine

云计算是一种基于互联网的新计算方式，通过互联网上异构、自治的服务为个人和企业提供按需即取的计算[1]。其核心思想就是网络上的资源和能力如何更有效地被分享，以达到高效率、低成本计算的目标。

云计算是一个具有更广泛含义的计算平台，支持广泛企业计算、Web2.0模式的网络应用程序，商业模型比较清晰，普适性更强。云计算强调对IT资源进行合理化分配重组，轻化客户端，在大型数据中心的服务器上对数据信息进行存储与处理。在云计算环境下，所有的计算资源，比如服务器，能够动态地从基础架构上产生出来，并加以修改以适应工作任务的需求。可以这样定义：云计算是一种应用模式，在这种模式中，应用、数据和IT资源以服务的方式通过网络提供给用户使用。

云计算具有5个特点[2]：

- 按需自助服务
- 普适网络接入
- 地点独立的资源池
- 快速伸缩
- 按需付费

云计算的交付模型，按照其在体系结构中的层次可分3种方式[3]：

- 软件即服务（SaaS）
- 平台即服务（PaaS）
- 基础设施即服务（IaaS）

基础设施云是一种商业模式的创新，它把IT资源、数据、应用作为服务通过互联网提供给用户，同时把IT基础资源（计算、网络、存储）形成资源池按需提供。终端用户（企业）可以通过网络获得自己所需要的计算资源，运行自己的业务系统。对于用户来说，云计算为他们提供了无限的计算能力和存储空间，使得他们方便地访问这些资源：只需要通过网络来获得相应的服务，按需使用。用户不必具有专业的技术知识，了解具体的技术细节。

从体系结构层次上看，基础设施云位于云平台体系结构的最底层，是其他各种形式的云的基础，其他的云都需要建立在基础设施云上，因此，基础设施云是整个云计算生态链的根本所在。

1 YUN系统设计目标

为了实现管理和虚拟化异构、分布的物理资源（包括服务器设备、网络设备、存储设备），并实现这些物理资源的按需使用，我们设计并搭建了一个虚拟化平台层，即YUN系统，目的是给用户按需提供虚拟运行环境、

基金项目：国家自然科学基金（61202081）

虚拟磁盘、基于Web的云存储，以及超市货架模式的软件设备等服务，使得YUN系统在服务器资源、网络资源和存储资源被有效利用的同时，大大降低用户的投资成本。YUN系统提供了统一的平台管理接口，且与Amazon EC2接口和S3接口兼容，同时提供了完整的支撑工具，包括部署工具、监控和管理工具等。YUN实现了支持“软件服务化、计算虚拟化、位置透明化、交互普适化”等特征的基础设施云系统。具体来说，YUN系统主要管理以下资源：

（1）虚拟运行环境资源

YUN系统通过虚拟机及多种虚拟化技术，在一台物理计算机上模拟出一台或多台虚拟的计算机，每个虚拟的计算机运行在一个完全隔离的环境中，并具有完整的软、硬件系统功能。实现异构、分布的软硬件资源（包括计算资源、存储资源、网络资源、服务资源及数据资源）的按需构建与重组，形成满足应用需求的资源映像，以服务的方式提供给使用者。

（2）计算资源

YUN系统整合服务器资源，形成多个虚拟运行环境，并通过虚拟运行环境之间的协同工作，从而大大提高了计算处理速度，使得用户在逻辑上拥有大量的可计算资源。

（3）网络资源

YUN系统对网络资源的管理，是通过桥接、虚拟局域网、隧道等网络技术，实现YUN系统的多种网络模式配置，根据用户需求将虚拟运行环境组成一个或多个用户组，在加快了用户组内虚拟运行环境之间通信速度的同时，也实现了YUN中虚拟运行环境内部及虚拟运行环境之间网络通信的可靠性和安全性。

（4）存储资源

YUN系统所管理的存储资源分为两种。一种为所有用户提供在线存储资源，保证用户数据完整性、保密性，同时方便用户随时随地上传、下载数据；另一种为用户的虚拟运行环境绑定虚拟磁盘存储资源，用户在使用虚拟运行环境时，将数据存储到该存储资源中，保证用户数据不丢失，方便数据移植。

（5）服务与数据资源

YUN系统中拥有众多的虚拟运行环境镜像资源，用户可根据自身系统及软件需求，向YUN申请相应镜像资源，用户无需自己安装操作系统、应用软件等繁琐操作。

2 YUN系统架构

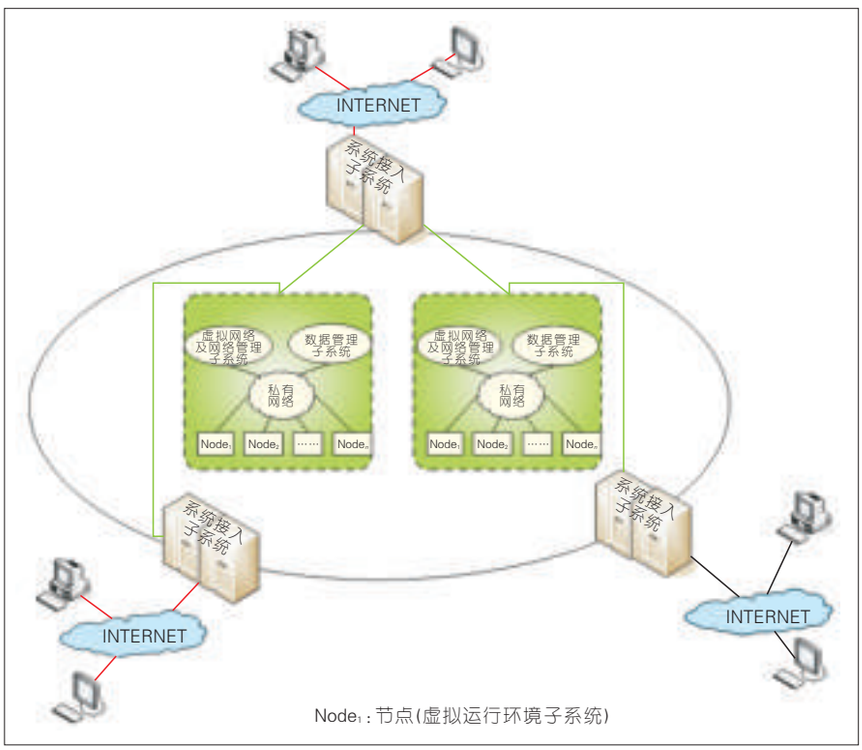
YUN系统可以聚合和管理来自一个或多个集群的资源。一个集群是指连接到相同网段的一组机器。同理，多个集群也就是指来自于多个网段的多组机器[4–10]。

多集群操作系统总体框架如图1所示。YUN系统包括以下子系统：系统接入子系统、数据管理子系统、虚拟网络及网络管理子系统、虚拟运行环境子系统、用户管理和安全管理子系统，这些子系统独立部署于单独的服务器上。

为了提高YUN系统整体性能，可以部署相应组件服务器的增强型硬件配置。例如，我们可以为系统接入子系统和虚拟网络及网络管理子系统的服务器安装超快CPU、超大内存；为数据管理子系统的服务器增加硬盘存储空间；为数据管理子系统和虚拟运行环境子系统的服务器装上具有较快I/O速度的存储设备等。使用多集群体系结构的目的是为了提高可用性的同时，保证负载和资源的跨集群分布，使得资源可以跨多个可用性区域分配，从而确保一个区域内的故障不会影响到整个系统。

2.1 系统接入子系统

系统接入子系统主要的功能是向用户提供软件应用服务和数据信息服务。主要负责处理由用户发起的请求或系统管理员发出的管理请求，做出高层的虚拟运行环境调度决策等操作，并且处理服务等级协议和维护系统及用户相关的元数据。该子系统由一组服务组成，这些服务用



▲图1 多集群操作系统总体框架

于处理用户请求、验证和维护系统及用户元数据，并管理和监视虚拟运行环境的运行。

系统接入子系统强调透明度和简单性，以便促进YUN系统的实验和扩展。为了达到这一要求，该系统的组件包括虚拟运行环境调度器、服务等级协议（SLA）引擎、用户接口和管理接口等。它们是模块化的且彼此独立的组件，对外提供定义良好的接口，企业服务总线（ESB）负责控制和管理它们之间的交互及有机配合，通过使用Web服务和Amazon的EC2查询接口与EC2的客户端工具交互操作。

系统接入子系统主要包括以下接口的设计和管理：

（1）B/S架构的用户访问接口

在B/S架构的用户访问接口方面采用了一种JAVA开发框架GWT，其优势在于简化AJAX应用程序的开发。在Web Server方面使用了易用和易嵌入的开源Servlet容器Jetty。

（2）YUN系统管理及使用接口

在YUN系统管理及使用接口方面使用了Netty，提供异步的、事件驱动的网络应用程序框架和工具，用以快速开发高性能、高可靠性的网络服务器和客户端程序。

（3）YUN系统元数据管理接口

在YUN系统元数据管理方面，使用关系型数据库对数据进行存储，并通过对象持久化技术，提供数据库访问接口。在存储方面，选择了Hsqldb，它是一个开放源代码的Java数据库，具有标准的SQL语法和Java接口，可以自由使用和分发，简洁并且快速。在访问方面，选择了Hibernate，一个开放源代码的对象关系映射框架，它对JDBC进行了非常轻量级的对象封装，使得Java程序员可以随心所欲地使用对象编程思想来操纵数据库。

（4）子系统之间通信接口及互联机制

在子系统之间通信及互联方面，使用Web Service进行各个子系统之间的互联，使用Apache+Axis2C完成发布服务的功能。Apache源于NSCAhttpd服务器，经过多次修改，目前已经成为世界上最流行的Web服务器软件之一。Axis2C采用名为AXIOM的新核心XML处理模型，灵活按需构造对象模型，支持不同的消息交换模式，提供阻塞和非阻塞客户端API，支持内置的Web服务寻址，支持HTTP、SMTP、JMS、TCP协议。

2.2 数据管理子系统

数据管理子系统的功能是为用户提供安全、可靠的数据存储服务，该子系统主要提供以下两个方面的服务：

（1）基于文件的存储服务

为用户提供数据流上传和下载的Web service，通过Web service用户可以申请存储空间，上传及下载数据，数据在上传、下载及存储的整个过程中都是加密的，因此，除数据拥有者外，包括管理员在内的所有合法或非法用户都不能读取数据的明文。

（2）基于块的存储服务

通过逻辑卷管理，以逻辑卷的形式为用户提供存储空间，并将逻辑卷与用户的虚拟运行环境绑定，实现动态扩展虚拟运行环境的存储空间。另外，由于该逻辑卷属于且仅属于申请用户，因此用户存储在该逻辑卷中的数据可以长期保存，其隐私性也能得到保证，同时这一服务也保证了用户在切换虚拟运行环境（将该逻辑卷与其他的虚拟运行环境绑定）时存储数据不丢失，从而实现存储资源的稳定、安全、便捷和可靠。

此外，为了保证数据管理子系统中海量数据的存储和一致性，数据管理子系统还应包括下述技术：重复数据删除；归档存储管理；连续数据保护容灾等。

2.3 虚拟网络及网络管理子系统

虚拟网络及网络管理子系统的主要功能是通过配置YUN系统中的虚拟网络，实现网络通信的方便、可靠和安全。虚拟运行环境的互联问题是构建YUN系统的重要工作之一。不同于物理机器组成的物理网络，虚拟运行环境组成的虚拟网络不具有严格、复杂的拓扑逻辑。通过虚拟化处理，虚拟网络具有简单和易配置等特点。在该子系统中，虚拟网络实现以下功能：

（1）通过配置虚拟网络，虚拟运行环境之间能够正常通信。

（2）定义安全组，运行在同一组内的虚拟运行环境具有相同的访问规则（如Ping、Ssh）。

（3）用户能够动态设置虚拟运行环境的IP。YUN系统对IP资源池进行动态管理，弹性地分配和释放IP，保证IP资源合理分配。

2.4 虚拟运行环境子系统

虚拟运行环境子系统的功能是根据用户的个性化需求，弹性地为用户分配虚拟运行环境。与此同时，通过虚拟运行环境之间的协同工作，大大提高了计算处理能力，在逻辑上使用户拥有大量的计算资源。该子系统除了为用户分配虚拟运行环境之外，还要管理虚拟运行环境，包括对虚拟运行环境的创建、修改、监控、关闭、清除、热迁移以及将逻辑卷与虚拟运行环境绑定等等。

该子系统通过集群控制器和节点控制器完成其主要功能。一个集群控制器可以管理多个节点控制器。集群控制器负责从其所属的节点控制器收集节点的状态信息，根据这些节点的资源状态信息，调度进入的虚拟运行环境执行请求到各个节点控制器上，并负责管理公共网络和私有网络的配置。集群控制器接口是通过WSDL文档来描述的，这些操作包括运行、描述和终止虚拟运行环境。描述和终止虚拟运行环境的操作会直接传给相关节点控制器。当集群控制器接收到一个运行实例的

请求后，它执行一个简单的调度任务，该任务通过调用描述资源命令来查询每一个节点控制器，选择第一个具有足够空闲资源的节点控制器来执行运行实例请求。集群控制器还实现了查询资源操作，该操作将一个实例需要占据的资源作为输入，并返回可以同时在其所属节点控制器上执行的实例个数。

为了准确地描述该子系统中虚拟运行环境的个数，在系统接入子系统中，要提供虚拟运行环境控制服务来管理虚拟运行环境元数据（硬盘、CPU及内存资源的状态及容量）的创建。虚拟运行环境控制器不间断的维护一个基本资源状态的简单本地描述，如一个集群控制器能够创建的虚拟运行环境个数。当发起一个虚拟运行环境创建请求时，虚拟运行环境控制器将和系统接入子系统进行协作，将用户的请求分解成镜像、网络和安全组等，并根据相应的元数据和资源应用配置策略预先生成一个解决方案，然后将消息发送至其涉及到的集群控制器，集群控制器将调度这些请求到其所辖节点控制器，最后由节点控制器创建虚拟运行环境来运行用户作业和应用程序。

2.5 用户管理和安全管理子系统

用户管理和安全管理子系统主要包括用户类型设置、用户身份认证两个方面。YUN系统对不同用户开放不同的权限，用户管理和安全管理子系统将用户分为管理员用户和普通用户。管理员用户负责通过YUN系统对系统所使用服务器资源、存储资源、网络资源等各种资源进行管理，而普通用户只能对属于自己的存储资源、虚拟运行环境进行管理。

在用户身份认证方面，YUN系统通过3个层次对用户进行认证：第1层，用户名、口令认证；第2层，证书认证；第3层，公-私钥密钥对认证。通过这3层认证，可以防止恶意用户窃取用户数据、攻击YUN系统、植入恶意代码等破坏性操作。

3 结束语

YUN系统是“核心电子器件、高端通用芯片及基础软件产品”国家科技重大专项“面向新型网络应用模式的网络化操作系统”的研发成果。YUN系统通过虚拟化技术和现代网络技术，将操作系统功能边界扩展到网络侧，将终端侧和网络侧在一个统一的架构内进行计算和资源的配置、调度、管理，使得网络化操作系统可以最大化地使用和共享网络系统内的资源。与其同时，YUN系统还支持用户终端的位置无关性，计算和资源配置的透明性，为用户提供一致的业务体验。

4 参考文献

[1] CLOUD COMPUTING [EB/OL]. [2011-10-05]. HTTP://EN.WIKIPEDIA.ORG/WIKI/CLOUD_COMPUTING. 2011.
[2] OLSEN K B, MINSTER J B, CUI Y, et al. SCEC TeraShake simulations: High resolution simulations of large southern San Andreas earthquakes using the TeraGrid [C]// Proceedings of the TeraGrid 2006 Conference, Jun 13-15, 2006, Indianapolis, IA, USA.2006.
[3] US Government Printing Office. Sarbanes-oxley act of 2002 [M]. USA Public Law 107-204.
[4] RUSSELL T. New regulations & compliance issues: How to stay one step ahead [C]// Proceedings of the Storage Networking World Spring 2008 Conference, Apr 7-10, 2008, Orlando, FL, USA.2008.
[5] EMC. Centera: World's first content-addressed storage(CAS)solution [R]. Hopkinton, MA, USA: Business Wire, 2003.
[6] GHEMAWAT S, GOBIOFF H, LEUNG S T. The Google file system [C]//Proceedings of the 19th ACM SIGOPS Symposium on Operating Systems Principles(SOSP'03), Oct 19 – 22, 2003, Bolton Landing, NY, USA. New York, NY, USA: ACM, 2003:29-43.
[7] COHEN R E. High-performance computing requirements for the computational solid earth sciences [M]. Washington, DC, USA: Carnegie Institution of Washington, 2005.
[8] CHIN R. Content archiving for rapid eDiscovery, compliance and disaster recovery [C]//Proceedings of the Storage Networking World Spring 2008 Conference, Apr 7-10, 2008, Orlando, FL, USA.2008.
[9] Object-based storage: The next wave of storage technology and devices [R]. Intel. 2003.
[10] MESNIER M, GANGER G R, RIEDEL E. Object-based storage [J]. IEEE Communications Magazine, 2003, 41(8): 84-90.

收稿日期：2012-10-05

作者简介

刘川意，清华大学博士毕业；北京邮电大学软件学院讲师、博士后；主要研究方向为云计算与云安全、可信计算与服务、数据保护与数据安全；已参与基金项目3项；已发表论文20余篇，其中SCI收录4篇（EI收录10余篇），申请国家发明专利9项（获得专利号3项，获得申请号6项）。

林杰，云南财经大学硕士毕业；北京邮电大学计算机学院在读博士研究生；主要从事云计算、可信计算、计算机网络、信息安全方面的研究；已参与基金项目3项；已发表论文4篇（EI收录2篇）。

综合信息

中兴通讯荣获BBWF 2012最佳宽带合作伙伴Infovision大奖

【本刊讯】2012年10月18日，在荷兰阿姆斯特丹举行的欧洲宽带论坛（BBWF）上，中兴通讯联合瑞典运营商Wexnet以创新的ICP Store解决方案荣获2012年最佳宽带合作伙伴大奖。该奖项是本次欧洲宽带论坛颁发的唯一一个联合奖项，彰显了中兴通讯以创新的技术与产品和客户实现共赢的经营理念。

云存储安全分析
Cloud Storage Security

中图分类号：TP393 文献标志码：A 文章编号：1009-6868 (2012) 06-0030-004

刘建毅/LIU Jianyi[1]
王枞/WANG Cong[1]
薛向东/XUE Xiangdong[2]
（1. 北京邮电大学，北京 100876；
2. 东华软件股份有限公司，北京 100190）
(1.Beijing University of Posts and Telecommunications, Beijing 100876, China;
2.Donghua software company, Beijing 100190, China)

摘要： 文章认为云存储是社会发展、技术发展的必然趋势，但其安全问题是云存储普及最大的障碍。基于云存储系统架构、安全分析及安全保护机制，文章给出了一种云存储服务应用支撑系统架构。该架构应用于云存储服务系统，是基于互联网应用的新型服务化存储模式，分为业务应用层、应用接口层、平台软件层和基础设备层，提供信息服务管理、运营统计分析和多种安全措施。

关键词： 云存储安全；加密存储；访问控制

Abstract: Cloud storage is an inevitable trend, but security is the biggest obstacle. In this paper, we discuss the safety and protection mechanism in the architecture of a cloud storage system. This architecture is divided into business application layer, application interface layer, platform software layer, and infrastructure layer. These provide information services management, operational statistical analysis, and a variety of safety measures.

Key words: cloud storage security; encrypted storage; access control

随着信息技术的高速发展和社会经济的发展进步，人们对计算能力的需求不断提高，数据的访问形式也发生了巨大的变化：从单个节点的独享访问，到集群、多机系统的共享访问；从数据的分散存储，到集中存放、统一管理；从单个数据存放节点，向数据中心发展，到建立跨城市、跨洲际的数据存储和备份体系。这些变化，对传统的存储系统的体系架构、管理模式提出了挑战。云存储是一个有效地解决这些挑战的途径，并且已成为信息存储领域的一个研究热点。

云存储是在云计算基础上延伸和发展出来的。它遵循了云计算共享基础设施的服务理念，以传统的大规模、可扩展的海量数据存储技术为基础，集成存储、网络、虚拟化和文件系统等多种技术，以超大规模、高性能、高效率、低能耗、高度可扩展、可靠性、可定制、动态组合和面向规模庞大的群体服务为系统目标，研究一种新的存储服务理念，为用户提供高效廉价、安全可靠、可扩展、可定制和按需使用的强大存储服务。云存储以其独特的特点和优势，集成并突破多种传统存储技术，避免了用户进行昂贵的设备采购、高额的管理和维护费用，提高了资源利用率，屏蔽了海量异构的数据存储管理的复杂性，增强了存储系统可扩展性、可伸缩性、可靠性和健壮性。作为一种新型服务化存储模式，云存储可广泛服务于经济建设、科学研究和国家安全等领域，具有重要而广阔的应用前景[1]。

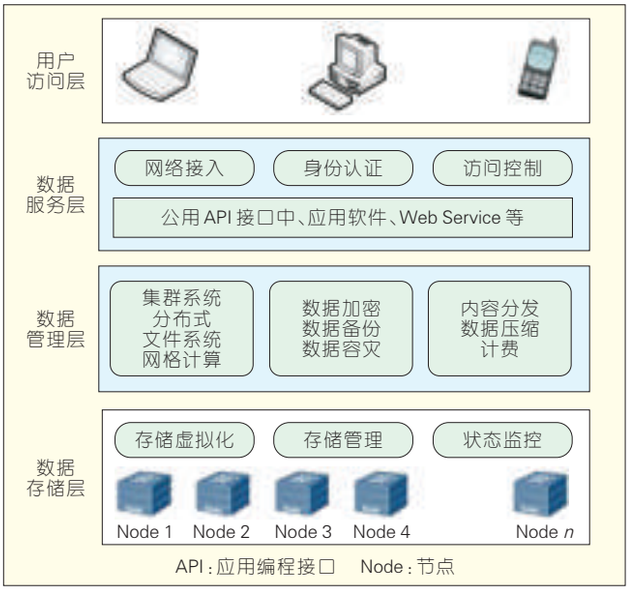
然而，云存储服务在带来便利的同时，也引起了用户对于安全性的广泛担忧，调查显示，出于安全方面的考虑，多达70%的用户仍然不愿意将关键数据置于自身控制域之外。事实上，Google Docs、The Linkup等多家著名云服务商都曾出现过各种安全问题，并导致了严重的后果。安全技术的缺失已经成为云存储普及的最重要的障碍。

1 云存储系统架构

2009年4月，全球网络存储工业协会（SNIA）主持组建了云存储技术工作组TWG[2]。该组织的主要任务是引领云存储的发展方向，制订相关的行业规范。目前已发布了关于云存储规范的第一个版本云数据管理接口（CDMI）[3]，在数据对象、容器、计算、计费、性能、队列、元数据6个方面提出了初步规范。与传统存储系统相比，云存储系统面向多种类型的网络在线存储服务，是一个由网络设备、存储设备、服务器、应用软件、公用访问接口、接入网和客户端程序等多个部分组成的复杂系统[4]，对外提供安全、可靠、高效的数据存储和业务访问服务。云存储系统的体系结构可划分为4个层次，如图1所示。

数据存储层是云存储最基础的

基金项目：国家高技术研究发展（“863”）计划（2012AA012606）



◀图1
云存储系统架构

部分，由不同类型的存储设备和网络设备组成。数据存储层实现海量数据的统一管理、存储设备管理、状态监控等。数据管理层是云存储最为核心的部分，也是最复杂的部分。数据管理层采用集群技术、分布式存储技术，实现多存储设备之间的协同工作，对外提供高可用性、可扩展性的服务，同时还负责数据加密、备份、容灾以及必要的计费等任务。数据服务层是利用云存储资源进行应用开发的关键部分，云存储提供商通过数据服务层为用户提供统一的协议和编程接口，进行应用程序的开发。用户访问层是基于云存储开发的应用程序的入口，授权用户可以通过标准的公用应用接口来登录云存储系统，享受云存储服务。

2 云存储系统安全分析

数据的安全性及可用性是云存储系统最为突出的问题，受到产业界的普遍关注。云安全联盟（CSA）发布的《云安全指南》是业界备受关注的安全参考。该指南着重总结了云计算的技术架构模型、安全控制模型以及相关的合规性模型之间的映射关系，在13个领域提出了具体的安全建议，包括事故响应、加密和密钥管理，身份和访问管理、以及法规和电子化搜寻等，并期望得到企业、机构和个人的关注、研究和采纳。

在学术界，全球对云存储安全方面的研究还比较少。Bowers等提出了分布式加密系统[5]；Cachin等通过使用加密工具来解决数据完整性和一致性问题，研究数据可恢复机制[6]；Tamleek Ali等提出了基于云计算的使用控制模型，对UCON模型进行了改进，定义了文件管理模块、访问控制模块、认证模块等共同保护文件的整个生命周期[7]；Xiaosong Lou等提出了一种基于数据毒化的版权保护方法，通过时间戳和签名判断用户是否合法，若为非法用户，针对他关于受保护文件的请求回应以不可用链接[8]，消耗非法用户的计算能力来抵御攻击；Kaiwang等提出了一种基于数字水印技术的可信云平台，通过对数据片嵌入标识用户信息的水印来保证数据的可审计性及完整性，此外该文中还建立了云计算的信任模型建立云计算资源和数据中心间的信任关系[9]；Jeremie Tharaud等则针对电子医疗信息云存储系统，提出用嵌入水印的方法来追踪电子医疗信息的分发者和使用者，在使用时提取水印进行验证，防止医疗信息被篡改[10]。中国清华大学、华中科技大学、国防科技大学、北京邮电大学等科研院校也开始在云存储技术相关领域进行基础性研究工作。

云存储系统的安全威胁主要表现如下[11]：

（1）云存储提供可伸缩的数据服务，无法清晰定义安全边界及保护设备，给云存储的安全保护措施增加了难度。

（2）云存储通过IP网络传输数据，因此传统网络上的安全威胁也存在于云存储系统上，如数据破坏、数据窃取、数据篡改、拒绝服务等，影响了数据的安全存储。

（3）数据存储的安全性包括静态存储安全和动态存储安全，静态存储安全是确保云存储系统上最终存储数据的存放安全，动态存储安全是确保在数据传输时的完整性和保密性。

（4）云存储需要保证数据的容错性、可恢复性和完整性，在灾难发生时如何避免数据服务中断及数据丢失等问题。

（5）云存储系统作为一个公共数据中心，具有多客户连接、高交互性、数据安全保障要求高等特点，对入侵、攻击、病毒和恶意软件十分敏感，有必要对云存储中的数据流进行实时主动地检测和防御。

可以说，不能保证安全的云存储，是无法运营的。但目前的研究只是对云存储的安全性做出提醒与建议，而没有提出具体的防护措施。因此，需要进一步研究云存储安全技术，制订云存储业务安全标准，让云存储业务安全、健康的发展。

3 云存储系统安全机制

在基于云安全联盟（CSA）的云安全模型[12]中，云存储安全机制可以归纳为3方面：平台安全机制、管控安全机制和应用安全机制。

云存储平台安全机制保护整个云存储平台系统自身的安全，主要实现技术是密码技术和系统加固技

术。密码技术保证系统和应用程序的完整性、提供强身份鉴别以及存储节点的透明加密。系统加固技术保障服务器和主机的安全性，如采用操作系统内核加固实现对计算/存储节点、虚拟主机的保护，免遭病毒木马攻击。

云存储管控安全机制主要解决云存储平台安全管理的问题，包括对云存储节点服务器密钥的统一管理、密钥生命周期的可控性、云数据接口/云客户端密钥的自主性等。

云存储应用安全机制主要解决云存储平台应用和服务的安全问题，主要技术有加密、身份认证、访问控制等。数据加密保证云存储静态数据和动态数据的机密性和完整性（静态数据，即磁盘数据、生产数据库中的数据及备份媒介中的数据等；动态数据，即网络中传输的数据，如信用卡号、密码和私钥等）。此外，加密信息检索是云存储数据加密中必须解决的问题之一。身份认证与访问控制必须包括身份供应、认证和访问控制三个功能，身份认证在云存储系统的各个层次都会涉及，如何实现单点登陆，使得用户的数据访问控制的身份能够具有应用的一致性，是一个需要解决的问题

4 一种云存储系统安全架构

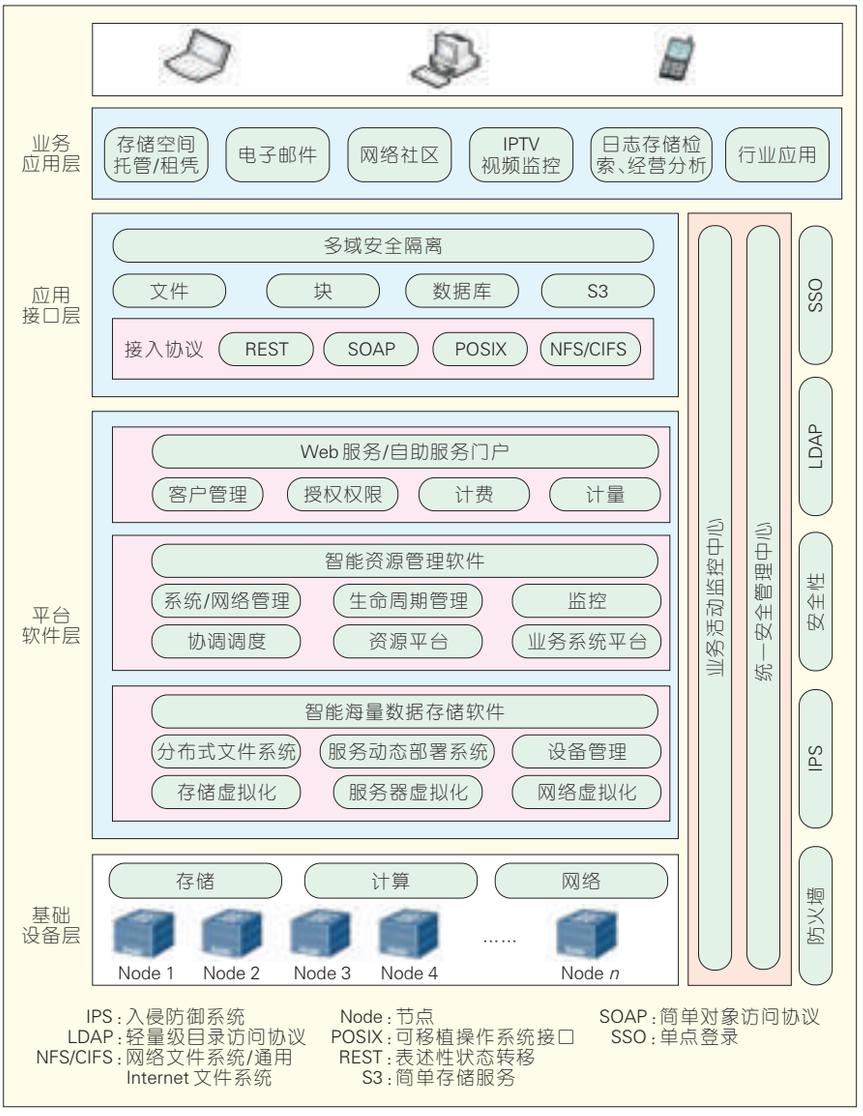
中国政府十分重视云存储系统及其安全技术的发展和产业化进程。2011年电子信息产业发展基金设立《云计算关键支撑技术及产业化（云存储服务）》项目，旨在研发具有自主知识产权的高性能、高可靠性、高安全性、高可扩展性、开放的云存储服务应用支撑系统。通过突破云存储的关键技术和先进的云存储服务应用支撑系统的创新研发和应用推广，促进中国云计算产业健康、快速的发展。

本文介绍一种云存储服务应用支撑系统架构。它应用于云存储服务系统，是基于互联网应用的新型服务化存储模式，分为业务应用层、应用接口层、平台软件层和基础设备层，提供信息服务管理、运营统计分析和多种安全措施。云存储服务应用支撑系统总体应用结构如图2所示。其功能覆盖了平台软件层、业务活动监控中心、统一安全管理中心。

业务应用层支持存储空间的托管和租赁、云邮件、网络社区、IPTV音视频监控等基本云存储服务，支持区域医疗云计算服务、村镇银行云计算服务等行业应用。任何一个授权用户都可以通过标准的公用应用接口来登录云存储系统，享受各种云存储服务。

应用接口层提供多域安全隔离策略，支持文件、块、数据库和简单存储服务第三方（S3）云接口等。提供遵循POSIX标准的文件级接口，提供开放的SOAP/REST和NFS/CIFS标准协议，通过应用编程接口（API）支持各种应用软件的开发。

平台软件层支持Web服务和自助服务门户，提供客户权限管理、访问控制策略、统计与计费管理。提供以智能资源管理为核心的系统/网络管理、生命周期管理、协调调度与监控。通过集群、分布式文件系统或网



▲图2 一种云存储服务系统的安全架构

格计算等技术，实现云存储中多个存储设备之间的协同工作，使多个的存储设备可以对外提供同一种服务，并提供更大更强更好的数据访问性能。提供服务动态部署系统、设备管理等运营软件。支持存储虚拟化、服务器虚拟化和网络虚拟化。通过各种数据备份和容灾技术和措施可以保证云存储中的数据不会丢失，保证云存储自身的安全和稳定。

基础设备层提供高性能以及高可靠性的后端存储。存储设备可以是光纤通道（FC）存储设备，可以是网络连接式存储（NAS）和 Internet 小型计算机系统接口（iSCSI）等 IP 存储设备，也可以是小型计算机系统接口（SCSI）或串行连接 SCSI（SAS）等直连式存储（DAS）存储设备。存储设备之上是一个统一存储设备管理系统，可以实现存储设备的逻辑虚拟化管理、多链路冗余管理，以及硬件设备的状态监控和故障维护。

统一安全管理中心的根本目标是保证存储数据的安全，即数据的保密性、完整性和可用性；主要实现身份认证、访问授权、数据加密、数据隔离、入侵检测与恶意攻击处理与安全审计功能，解决信息存储安全，实现威胁来源的收集、分析与处理，即时进行安全升级和威胁防护，并完成安全信息的高效分发等功能。

安全管理涉及的过程有数据生成、传输、保存、访问。在云存储服务应用支撑系统的各层次上都有相应的安全技术：如应用接口层的单点登陆、平台软件层涉及数据流转的数据加密技术、平台软件层的入侵检测处理技术、基础设备层的防火墙技术等等。

业务活动监控中心主要负责云存储服务应用支撑系统中各类运营数据的采集与分析，快速发现和恢复系统故障，实现云存储服务应用支撑系统的高效可控运行。

5 结束语

云存储是社会发展和技术发展的必然趋势，但其安全问题是云存储普及最重要的障碍。可以说云存储安全不单单是技术问题，还涉及到标准化、管理模式、法律法规等诸方面的问题。需要学术界、产业界以及政府相关部门共同努力才能实现。

6 参考文献

[1] 金海, 吴松, 廖小飞, 等. 云计算的发展与挑战 [M]. 中国计算机学会学术工作委员会. 2009 中国计算机科学技术发展报告. 北京: 清华大学出版社, 2009.
[2] SNIA Cloud [EB/OL]. [2011–10–15]. http://www.snia.org/cloud. 2011.
[3] SNIA. Cloud Data Management Interface (Version 1.0)[EB/OL]. [2011–10–15]. http://cdmi.sniacloud.com. 2011.
[4] 王庆波, 等. 虚拟化与云计算 [M]. 北京: 电子工业出版社, 2009.
[5] BOWERS K D, JUELS A, OPREA A. HAIL: A high availability and integrity layer for cloud storage [C]//Proceedings of the 16th ACM Conference on Computer and Communications Security (CCS'09), Nov 9–13, 2009, Chicago, IL, USA. New York, NY, USA: ACM, 2009: 489–501.
[6] CACHIN C, KEIDAR I, SHRAER A. Trusting the cloud [J]. ACM SIGACT News, 2009, 40 (2):455–461.
[7] ALI T, NAUMAN M , HADI F E, et al. On usage control of multimedia content in andthrough cloud computing paradigm [C]// Proceedings of the 5th International Conference on Future Information Technology (FutureTech'10), May 21–23, 2010, Busan, Republic of Korea. Los Alamitos, CA, USA: IEEE Computer Society, 2010:5p.
[8] LOU X, HWANG K. Collusive piracy prevention in P2P content delivery networks [J]. IEEE Transactions on Computers, 2009, 58(7): 970–983.
[9] HWANG K, LI D Y. Trusted cloud computing with secure resources and data coloring [J]. IEEE Internet Computing, 2010, 14(5): 14–22.
[10] THARAUD J, WOHLGEMUTH S, ECHIZEN I, et al. Privacy by data provenance with digital watermarking [C]//Proceedings of the 6th International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH–MSP'10), Oct 15–17, 2010, Darmstadt, Germany. Piscataway, NJ, USA: IEEE, 2010:510–513.
[11] STEVE M. Danger in the clouds [J]. Network security, 2008(12):9–11.
[12] 云安全联盟. 云计算关键领域安全指南 V2.1 [EB/OL]. [2011–07–05]. http://www.cloudsecurityalliance.org/guidance/csaguide.v2.1.pdf. 2011.

收稿日期：2012–10–05

作者简介

刘建毅，北京邮电大学博士毕业；北京邮电大学信息安全中心副教授、博士；主要研究领域为灾难备份、信息内容安全；已主持和参加国家级基金项目 10 余项，获 1 项科研成果奖；已发表论文 40 多篇，其中被 SCI/EI 检索 20 余篇。

王枞，北京科技大学控制理论与控制工程专业博士毕业；北京邮电大学教授；长期从事信息安全、智能信息处理、人机交互、电子商务等方面的研究；已发表论文 40 余篇（其中 SCI、EI 收录 30 余篇），出版著作 5 部。

薛向东，湖南大学计算机科学系毕业；东华软件股份公司董事长。

综合信息

中兴通讯推出业内首款基于个人 PC 的 LTE 容量设计规划工具

【本刊讯】2012 年 10 月 22 日消息，中兴通讯对外宣布推出业内第一款基于个人 PC 的 LTE 无线网络容量规划工具（CPT）。

该规划工具突破了传统的容量规划理念及方法的局限，包含 4 项技术专利，是业界第一个系统性、专业性的 LTE 容量规划工具，填补了业内空白。与传统方案相比，该方案在规划精度及效率上分别实现 20% 及 80% 以上的提升。

基于IaaS云计算平台的弹性计费模型

An Elastic Billing Model for IaaS Cloud Computing

中图分类号：TP393 文献标志码：A 文章编号：1009-6868 (2012) 06-0034-004

袁玉宇/YUAN Yuyu
胡文博/HU Wenbo
（北京邮电大学，北京 100876）
(Beijing University of Posts and Telecommunications, Beijing 100876, China)

摘要： 文章认为计费系统是云计算平台不可缺少的重要组成部分。在分析中国外著名企业的云计费模式的基础上，文章针对基础设施即服务（IaaS）云计算平台，提出了一种基于租赁实体的计费模型与基于事件触发机制的计费模型相结合的计费模型，利用租赁实体的计费模型对空闲的主机进行调度，通过事件触发机制对计费信息进行统计。该计费模型使云计费方式更加灵活。

关键词： 云计算；租赁模型；事件触发机制；弹性计费模型

Abstract: A billing system is one of the most important parts of a cloud computing platform. In this paper, we analyze the cloud billing models of some well-known enterprises and propose an elastic billing model. This billing model combines a leasing instance model with an event-triggered mechanism. The leasing-instance model can schedule the idle host, and the event-triggered mechanism can create statistics on billing information. This billing model makes cloud computing billing more flexible.

Key words: cloud computing; leasing instance model; event-trigged mechanism; elastic billing model

云计算作为一种新的资源共享方式，可以按需提供软硬件资源和数据信息等。云计算的兴起将使很多企业摒弃原有购买服务器、存储设备的模式，转而选择更灵活的服务模式。计量、定价和收费等新商业模式已成为云计算供应商需要解决的核心问题，因此，针对云计算的计费模式即云计费模型的研究变得十分重要。

目前Amazon和Saleforce.com两大云计算提供商公司都已经进入了盈利阶段[1]。根据相关预测，2014年亚马逊的AWS（Amazon Web Service）部门的收入将达到25亿美元左右。

云计算的计费模式成为云计算用户去留的重要因素，因此，云计算提供商必须提供一种灵活、快速、方便的云计算计费模式。

1 云计算服务类型

云计算服务类型可以分为基础设施即服务（IaaS）、平台即服务（PaaS）、软件即服务（SaaS）三大类。

IaaS将设备作为服务出租给客户。IaaS无法计量部署在设备上的应用程序的使用量，因此，IaaS的计费模型以使用时间和存储空间作为计费标准。

PaaS通过云提供商的计算平台和解决方案框架，将用户的应用程序部署到云端。虽然用户不能控制云下层的基础设施，但是可以控制部署的应用程序和主机环境设置。因此，PaaS平台的计费不仅能在操作系统层面计量CPU和内存的使用量等；而且还能在应用层面计量运行时间和所消耗的事务多少等。在PaaS中，一般采用“随用随付”的计费方式[2]，不仅可以将硬件资源作为计量标准，也可以将应用软件的动作行为作为计量标准。可以通过系统日志记录硬件资源消耗、网络流量等，实现基于用户行为分析的计费方案[3]。

SaaS将特定软件运行于云的基础设施上，用户可以通过浏览器连接该服务并根据特定的需求配置选项，SaaS计费方案大多采用按需服务模式和按用户人数计费。按需服务模式可以通过用户所租用的软件模块计费，这种模式延伸了传统的管理软件。按用户人数模式则根据统计每月每用户的方式进行收费，有利于小企业的应用。

2 云计费模式分析

云计费的先行者Zuora，提出了系统的定价和组装、管理整个用户的使用生命周期、在管理过程中分析和把控各个计算点3个云计费问题[4]。

定价和组装需要很大自由度，不同的公司有不同的云计算需求，计费可能是时长，也可能是带宽。就像电信计费的灵活性一样，每个时间段的费用可能不同，将这些多样的需求放

基金项目：国家高技术研究发展（“863”）计划（2011AA01A204）；国家自然科学基金（91118002）

在一个计费系统当中就会非常复杂。而个性化用户可能还会有配合自身产品的新的计费要求，因此一定要考虑到各种新的需求，将它们放到成型的计费系统中。系统的定价和组装是云计费的一个重要环节，本身的自由度又很大，所以在设计时就会出现很多问题，所以需要通过各种方案的比较来确定一套能够符合需求的方案。

管理用户应用的生命周期租用的方式非常灵活。比如用户第1个月用100 min，第2个月使用900 min，但是第3个月可能暂停计费，因此这种计费方式更像是一种订阅方式，是一个长期的商业关系。这同样是在云计费中的重要因素，应该通过灵活的租用方式，达到“用多少，付多少”的订用型商业模式。

为了将系统做得更出色，需要通过在管理过程中进行分析，控制好每个计算点，不仅可以实现精确计费，还可以帮助用户对自己的产品进行数据分析。基于用户行为分析的云计费系统就是通过统计分析各个关键计算点进行计费和用户信息分析。

2.1 亚马逊的云计费

亚马逊的云计算服务AWS[5]，包含了众多的细分产品服务，满足不同用户的需求，“计算能力”服务是其中最主要的细分服务。它向用户提供高性能的计算能力，满足用户的数据计算和数据吞吐需求。“计算能力”服务包括弹性计算云（EC2）、Amazon MapReduce、Auto Scaling等。

以EC2为例，亚马逊在全球共有5个EC2数据中心。EC2的服务价格会因云计算数据中心所处地区不同而有所差异，其中以东京的最贵，新加坡的最便宜。EC2服务的报价分为4个方面：计算实例、数据传输、存储、其他增值服务。

（1）计算实例

计算实例即“逻辑上的计算机”。客户租用这样一台“逻辑上的计算机”，来配置不同的计算资源，包括CPU、内存、硬盘、I/O总线等。亚马逊提供多种不同配置的计算实例，另外根据租用方式不同，租金也有所不同。亚马逊总共有3种租用方式：按需租用、预留租用和现场租用。

（2）数据传输

客户在租用计算实例的同时也需要租用EC2的网络连接服务，数据传输服务按照流量收取费用。使用不同连接线路传输数据，该费用价格会有差异。连接线路分为3种：互联网连接、数据中心之间连接以及可用性区域之间连接。

（3）存储

如果客户的存储需求非常高则可能需要租用亚马逊提供的存储服务EBS。EBS的计费方式比较简单，按照客户所占用的存储空间、时间以及读写次数来计算费用。

2.2 微软的云计费

微软的Azure平台2010年开始投入商用[6]。目前Azure推出了5项托管服务，包括.NET应用服务、SQL服务、SharePoint服务、Dynamics CRM服务，以及Live服务等，以帮助客户建立云计算的应用，或将现有的业务扩展到云端[7]。

Azure平台的计费模型分为以下几个方面：计算费用、存储费用、请求事务费用、连接费用和带宽费用。

（1）计算费用

计算费用=使用小时数×实例数（即虚拟机（VM）的数目）×计费基准，其中计费基准根据VM的具体规格确定。

（2）存储费用

存储费用分为Windows Azure存储费用和Sql Azure数据库存数费用。Windows Azure提供的存储服务包括Table、Queue和Blob存储；而Sql Azure根据创建的数据库大小不同，收费也不同。两者均以月作为一个周期，统计一个月内上传的容量，分摊到每一天得到日平均值，然后根据日平均值收费。

（3）请求事务费用

向Azure平台服务发起的HTTP请求称为事务（Transaction），不同的请求事务类型有不同的收费标准。

（4）连接费用

Azure按照一个月内平均的日最大并发连接数目进行计费。可以选择按需计费，也可以选择购买固定最大连接数。

（5）带宽费用

一旦数据中心节点内外之间发生任何数据传输，均需要计入带宽费用，但同一个数据中心内的数据传输不收费。微软的数据中心可以分为主节点和内容分发网络（CDN）节点，这两种节点的带宽收费标准也是不同的。

2.3 Google的云计费

Google的云计算服务GAE平台[8]共有3类收费内容：可购买的资源，数据操作和API使用费用。

可以购买的资源可分为5类：CPU时间，传入带宽，传出带宽，存储和电子邮件。用户可以根据应用选择合适的配置。

数据操作分为3种类型：读操作、写操作和小型操作。每种操作根据应用情况计费。

API使用情况可以看作是一种高级的数据操作方式，可以根据API的功能换算成相应的低级数据操作集合，借以作为计费的依据。

2.4 中国云计费

盛大云推出流量计费收费方式，包括云主机费用（云主机套餐费用+带宽套餐费用）、云硬盘容量费用、云存储费、云分发费等[9]。

腾讯开放平台对云服务的计费包括3个部分：虚拟机计费，带宽计费和云存储计费[10]。腾讯开放平台通过信用账户对每项服务采取预扣费形式。

新浪的SAE用云豆作为虚拟货

币进行支付，按开发者使用的各服务对应消耗的资源扣除云豆，云豆消耗完后将不能继续使用服务，即SAE采用预付费方式，资源消耗的同时，扣除对应数量的云豆[11]。

3 基于IaaS模式的计费模型

目前，IaaS模式下主要采用租赁实体的计费模型，这种计费模型已经相当成熟，并已被广大云服务商接受，但是这种方式的灵活性差。赵旭等提出了基于事件触发机制和用户行为的计费模型，这种方式计算繁琐。本文提出了一种基于租赁实体的计费模型与基于事件触发机制的计费模型相结合的计费模型，利用租赁实体的计费模型对空闲的主机进行调度，通过事件触发机制对计费信息进行统计。

3.1 基于租赁实体的计费模型

目前，IaaS模式下主要采用租赁实体的计费模型[12]，主要内容包括主机配置、网络流量和存储。其中主机配置包括CPU、内存和硬盘大小；网络流量则包括流入和流出的网络数据；存储是指除了主机配置中的硬盘之外扩展的硬盘空间。

主机配置将主机作为服务租借给用户，主机的不同配置直接影响云应用的运行速度与时间。若主机配置过低，则会造成云计算所需时间过长，导致费用的增加；若主机配置过高，则会造成浪费，同样会增加费用。

网络流量占据了大部分费用，主要源于电信运营商的费用。由于大部分云运营商不是电信运营商，他们需要向电信运营商租用带宽以保证云端的应用能够通过网络正常使用。

云应用都需要大量的空间存储各种数据及备份。针对不同的存储安全及冗余策略，有不同的存储解决方案和计费标准。

在租赁实体的计费模型下，云服务商将主机租用给初级用户使用，但是当该主机空闲时，初级用户可以将其再次租赁给二级用户，从而降低租用主机的费用，甚至可以从中获利，租赁过程如图1所示。

租赁具体步骤如下：

（1）当初级用户所租用的主机空闲时，主机会被提交到云计算中心。云计算中心会将主机按照系统类型和配置类型安排到不同的租赁实体队列。

（2）二级用户根据他们的需求提交实习应用。应用包括系统类型，实体配置。然后应用会被放入到实体请求队列。

（3）云计算中心按照一定策略将租赁实体队列与请求队列匹配。只要租赁队列中的实体与请求队列中的任务需求的实体配置相符合，则认为匹配成功。若租赁队列中的实体没有匹配成功，则将其放入等待队列，并会被优先匹配。

（4）如果租赁实体匹配成功，则将其租赁给二级用户。

（5）初级用户恢复租赁实体的请求已经被提交到云计算中心。

（6）若实体没有被分配，则直接恢复。若实体已经被分配给二级用户，此时的云计算中心需要保存二级用户的运行环境，然后再恢复实体。

当第6步完成时，实体退回给初级用户。

在该模型下，供求关系通过一个指数函数计算：

$$D(X_t)=\begin{cases}\frac{X_t-1}{e^{c}} & X_t\in N, X_t\geqslant 2,\\ 1, & X_t=0.1\end{cases}$$

其中C表示云资源池中二级用户的最大数目，X_t则是时间t时资源池中二级用户的数目。租赁实体的价格会随着竞争用户数目的增加而增加。

3.2 基于事件触发机制的计费模型

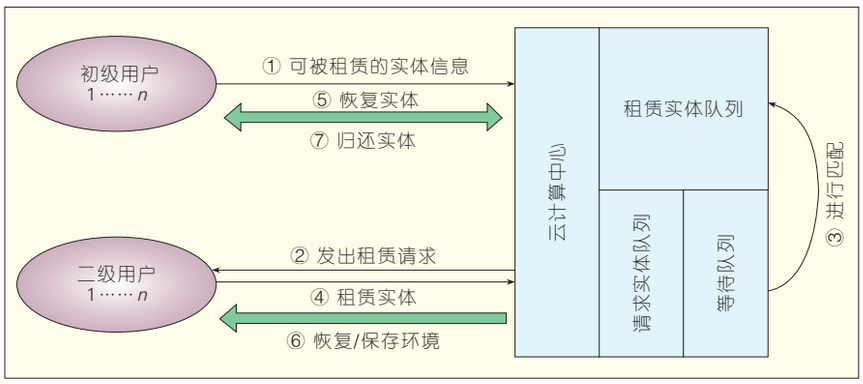
基于事件触发机制的计费模型实时监控用户行为，事件状态流程转换如图2所示。作业提交后，作业若调度及分派成功，则进入已分配状态，否则进入全局等待状态，处于分派状态的作业若启动成功，并且执行完毕，则进入执行的状态，否则进入本地等待的状态，当作业处于完成执行状态时，触发信息收集进程，并成功计费后交作业结束。

基于事件触发机制的计费模型采用了计费信息即时收集算法[13]，算法流程如图3所示。在任务完成时，将通过配适器执行计费信息提取命令，成功后，对返回的文本进行解析，最后存储解析结果，完成计费信息的收集。

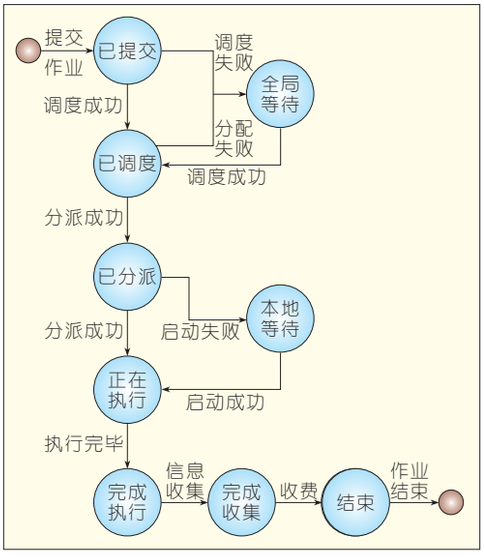
3.3 基于租赁模型和事件触发的弹性云计费模型

基于租赁模型和事件触发的弹性云计费模型，通过基于租赁实体的计费模型对空闲的主机进行调度，并根据该模型中的供求关系作为计费权重因子。再通过基于事件触发机制的计费模型对计费信息进行统计。

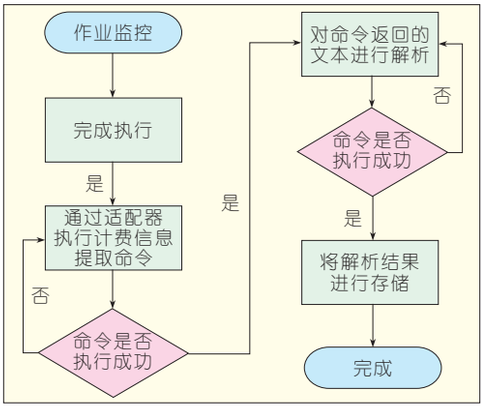
当初级用户在不急切需要主机



▲图1 基于实体租赁的过程



▲图2 事件状态流程



▲图3 计费信息即时收集算法

时，可以向云运营商提出外租请求，运营商将其所占用的主机放入资源池中。并会根据资源池中主机的数目和申请主机的二级用户的数目计算出供需关系，并按照申请主机的队列顺序分配资源池中的主机，同时根据当前的供需关系动态的对租用主机的价格进行调整。当初级用户申请主机时，二级用户则无条件退还主机，并将作业转移到其他符合要求的主机上继续执行。

另外，在统计二级用户租用主机的费用时，可以根据系统的计费信息采集进程所采集的数据和供需关系进行弹性计算。当二级用户较多而资源池中的主机较少时，二级用户租用主机租用的费用将升高；相反地，二级用户较少而资源池中的主机较多时，租用的费用将降低，计费函数为：

$$C_t = \sum_{i=1}^{n} P_i(D(X_t)) \times U_i$$

其中，C_t表示在时间段t内产生的费用，$D(X_t)$表示时间段t内的供需因子，P_i表示根据供需因子所计算出的计费项目i的单价，U_i则表示时间段t内所使用的项目i的量。

这样，既可以根据空闲主机和二级用户的数量动态的分配主机，提高了主机的利用率；同时又根据供需关系作为计费的权重因子，保证了初级用户及运营商的利益。

对于一些长期不紧急的任务，可以尽量在硬件资源充足的情况下获取硬件资源；对于一些需要时间较短且紧急的任务，则在相应的时间段内尽可能的多分配硬件资源，以达到资源利用最大化。所以面向IaaS的弹性计费模型将不同硬件的使用时间和流入流出流量作为计量要素，并在此基础上最大限度的减少硬件资源的空闲时间，弹性的将硬件资源分配给最需要的用户。

4 结束语

合理的计费模式，是云计算从技术研发转向服务管理的关键环节。针对云服务的计费模式，各云服务提供商在实践中进行了不同的尝试与创新，与用户多样化的需求相比，云服务的计费模式依旧显得较为单一，弹性计费模型是实现按需服务的关键，还需要不断探寻新的计费模式，促进云计算服务的健康发展。

5 参考文献

[1] 陈红, 任怡, 刘晓建. 云计算平台下计费机制研究 [C]//第二届中国计算机学会服务计算学术会议论文集(CCF NCSC 2011), 2011年8月18–19日, 济南. 2011:48–52, 68.
[2] PaaS计量与计费 [EB/OL]. [2011–09–07]. http://www.oracle.com/technetwork/cn/topics/cloud/ paas–metering–090434–zhs.html. 2011.
[3] 唐箭. 基于用户行为分析的云计算计费系统的分析与设计 [J]. 辽宁经济管理干部学院(辽宁经济职业技术学院)学报, 2009, 45(5): 46–47.
[4] Zuora:云计费计费产业的领军者 [N]. 中国计算机报, 2010–12–13(084).
[5] Amazon Web Service, LLC. Amazon elastic compute cloud(Amazon S3)[EB/OL]. [2011–09–07]. http://aws.amazon.com/ec2/. 2011.
[6] Azure定价概述 [EB/OL]. [2011–10–05]. http://www.windowsazure.com/zh–cn/pricing/details/. 2011.
[7] Azure_百度百科 [EB/OL]. [2011–10–05] http://baike.baidu.com/view/978088.htm. 2011.
[8] Billing and budgeting resources – Google app engine [EB/OL]. [2011–10–05]. http://code.google.com/intl/zh–CN/appengine/ docs/billing.html. 2011.
[9] 价格——盛大云计算 [EB/OL]. [2011–10–05]. http://www.grandcloud.cn/index/price. 2011.
[10] 腾讯开放平台云服务计费标准 [EB/OL]. [2011–10–05]. http://wiki.open.qq.com/wiki/. 2011.
[11] SAE资费业务说明 [EB/OL]. [2011–10–05]. http://sae.sina.com.cn/?m=devcenter&catId=155. 2011.
[12] Yuan Qin, Liu Zhixiang. A leasing instances based billing model for cloud computing[M], Advances in Grid and Pervasive Computing, LNCS 6646. Berlin, Germany: Springer–Verlag, 2011:33–41
[13] 赵旭, 吕太强, 梅一多, 等. 基于事件触发机制和用户行为的网格计费模型 [J]. 华中科技大学学报: 自然科学版, 2011, 39(z1): 28–32.

收稿日期：2012–10–05

作者简介

袁玉宇，北京邮电大学软件学院教授，ISO/IEC JTC1/SC7/WG6工作组中国代表团团长，全国信息技术标准化技术委员会软件工程分技术委员会软件质量测试工作组（SAC/TC28/SC7/WG1）组长，中国国家软件标准化推广中心副主任，中国电子学会数据库专家委员会委员，第5届IEEE认知信息学国际会议委员会主席；从事可信软件(服务)度量、评估、认证的理论、方法与技术研究；已发表论文5篇，出版专著(译著)2部。

胡文博，北京邮电大学软件学院在读硕士研究生；主要从事可信分布式计算与服务研究。

新一代移动承载网：IP RAN 网络

Next-generation Mobile Backhaul Network: IP RAN

中图分类号：TN929.5　文献标志码：A　文章编号：1009-6868 (2012) 06-0038-04

摘要：文章认为采用IP无线接入网(IP RAN)网络逐步代替多业务传送平台(MSTP)网络已是运营商网络发展的重要趋势。部署新型IP RAN网络也面临着很多挑战，需要考虑组网模型、多业务承载、QoS保障和端到端管理等诸多问题。文章指出如何实现MSTP与IP RAN的互联互通成为了IP RAN网络演进中的重要课题。

关键词：IP无线接入网；多协议标签交换；分组传送网；多业务传送平台；互联互通

Abstract: Using IP radio access network (IP RAN) instead of multiservice transfer platform (MSTP) is a trend in network development. There are many challenges associated with networking model, multiservice, QoS, and end-to-end management for IP RAN. In this paper, we discuss these issues and describe how to interconnect and interwork MSTP and IP RAN. This is an important aspect of IP RAN network evolution.

Key words: IP RAN; multi-protocol label switching (MPLS); packet transport network; MSTP; interconnection and interworking

唐雄燕/TANG Xiongyan
简伟/JIAN Wei
张沛/ZHANG Pei

（中国联合网络通信有限公司研究院，北京 100048）
(China Unicom Research Institute, Beijing 100048, China)

- 基于分组技术的多业务传送网络是运营商城域网演进的主流趋势
- IP RAN的部署需要兼顾3G与LTE网络的回传需求
- 城域网络的演进是基于分组化的IP RAN和基于TDM的MSTP融合协作并逐渐替代的过程

IP RAN是无线接入网(RAN)IP化的产物，通常指基于IP技术的移动回传网络。随着移动互联网的迅猛发展，无线接入网正朝着全IP化的方向演进，包括基站的IP化和移动回传的IP化。随着第三代合作伙伴计划(3GPP)对无线接入网IP化标准的大力推进，从2006年开始，主流设备厂家所提供的3G基站产品就具备较完善的IP协议支持能力。然而由于移动回传一直依赖基于同步数字体系(SDH)的多业务传送平台(MSTP)网络，因此难以满足无线接入网IP化和宽带化的发展要求。IP RAN作为基于路由器/交换机等数据设备所构建的动态IP承载网，成为实现移动回传IP化的重要技术选择。

1 业务需求与应用定位

随着移动通信从3G向LTE的演进，以及智能手机等终端对网络覆盖和传输带宽的不断增加，中国主流运营商均加大了3G网络的部署，进一步推动业务从传统的语音式向数据式演进。目前移动网络业务主要是通过MSTP网络进行回传，其典型的接入环容量为155 M或者625 M。从现有网络的实际流量监控来看，在相当长时期内，MSTP技术仍可作为以2G、3G移动回传为主的、带宽需求不大的站点的综合接入手段。但是由于应用环境的变化和自身特点的限制，现有的MSTP技术在高带宽供给、三层网络功能提供等方面已经不能适应面对业务IP化和宽带化的发展趋势。因此，IP RAN的部署还需要兼顾3G与LTE网络的回传需求。

对于全业务运营商，他们则希望实现移动与固定宽带业务的融合承载。其期望承载的业务主要包括：移动回传业务、家庭客户接入业务（固定宽带、语音及IPTV业务等），以及集团客户业务（时分复用模式(TDM)专线、异步传送方式(ATM)专线、以太网专线，及多协议标签交换(MPLS)业务等）。从业务的服务质量要求、对承载传送网络的要求、流量特点等方面综合考虑，承载的业务划分为两大类：公众平面业务和电信平面业务。前者主要指互联网宽带业务，也包括集团客户的互联网专线接入，其业务特点为流量大、突发性强、控制难度大、无明确质量要求，并且均为尽力而为业务；后者主要包括移动回传、固定语音、交互式网络电视(IPTV)、集团客户专线等业务。其业务特点为流量模型相对稳定、便于控制，主要为运营商网内业务或集团客户业务，对于安全性和可控性的要求比较高，并且有严格的质量要求，承载网络的

封闭性要求强。运营商对于不同类型的业务往往采用不同的承载技术和网络：公众平面业务主要由传统的IP城域网进行承载，业务包括互联网业务、固定宽带业务以及Wi-Fi热点业务等；电信平面业务则采用相对封闭的专用IP承载网，业务包括下一代网络(NGN)、移动分组域业务等。从网络运营的集约化角度考虑，在城域层面实现多业务统一承载是必然的趋势[1]。

总之，为满足3G/LTE回传及全业务统一承载的需求，IP RAN需要具备如下能力：

• 多业务承载。目前运营商网络承载的业务包括互联网宽带业务、大客户专线业务、固话NGN业务和移动2G/3G业务等，既有二层业务，又有三层业务。尤其是当移动网演进到LTE后，S1和X2接口的引入对于底层承载提出了三层交换的需求。由于业务类型丰富多样，目前各业务的承载网独立发展，造成承载方式多样、组网复杂低效、优化难度大等问题。新兴的承载网需要朝着多业务承载的方向发展。

• 超高带宽。随着业务日趋宽带化，固网宽带提速后家庭接入可达20 M，并在向100 M迈进；移动宽带高速分组接入(HSPA+)已规模商用，带宽达21 M甚至42 M；未来LTE部署后用户带宽可达300 M。因此移动回传与城域承载网必须有足够强的带宽扩展能力。

• 服务质量(QoS)保障。带宽的提升和业务类型的多样化对网络QoS保障能力提出了更高的要求。移动回传网同时承载移动PS域和CS域的业务，CS域业务通常需要更高的QoS保证。此外，承载网还承载大客户专线等高价值业务，网络必须具备完备的QoS能力。

• 高可靠性。为保证网络质量，承载网需要具备端到端的操作、管理和维护(OAM)故障检测机制，可以从业务层面和隧道层面对业务质量和网络质量进行管控。此外，网络还需要电信级的保护倒换能力，以更好地确保语音、视频等高实时性业务的服务质量。

近年来，国际上许多运营商纷纷采用IP RAN建设移动回传网或多业务承载网。如北美运营商AT&T、Verizon、Sprint等都已确立IP RAN的建设思路，他们认为IP RAN便于扩展网络规模，统一的承载网能有效简化网络结构，同时在产业链和运维方面的优势也会大幅降低运营成本。西班牙电信Telefonica已确定向IP RAN方向发展，在马德里、巴塞罗那等7个城市部署了IP综合承载网。英国电信BT的IP RAN建设使BT的移动承载网能满足综合业务需求，适应快速的业务变化与演进。中国运营商中国联通与中国电信也积极开展了IP RAN的全面测试和现场试验，探索IP RAN规模应用的可行性，中国联通已开始IP RAN规模部署的实践[2-3]。

2 技术特点与主要优势

IP RAN是指IP化的无线接入网，虽然狭义上的IP RAN专注于陆地无线接入网(UTRAN)到核心网(CN)之间的IP化。但是目前业界常说的IP RAN是指使用IP/MPLS技术来承载业务的组网方式。由于使用IP/MPLS组网可以从根本上实现移动回传的IP化，因此业内就直接使用IP RAN来命名这种组网方式。IP RAN具有如下技术特点和优势：

(1)端到端的IP化。端到端的IP化使得网络复杂度大大降低，简化了网络配置，能极大缩短基站开通、割接和调整的工作量。另外，端到端IP减少了网络中协议转换的次数，简化了封装解封装的过程，使得链路更加透明可控，实现了网元到网元的对等协作、全程全网的OAM管理以及层次化的端到端QoS。IP化的网络还有助于提高网络的智能化，便于部署各类策略，发展智能管道。

(2)更高效的网络资源利用率。面向连接的SDH或MSTP提供的是刚性管道，容易导致网络利用率低下。而基于IP/MPLS的IP RAN不再面向连接，而是采取动态寻址方式，实现承载网络内自动的路由优化，大大简化了后期网络维护和网络优化的工作量。同时与刚性管道相比，分组交换和统计复用能够大大地提高网络利用率。

(3)多业务融合承载。IP RAN采用动态三层组网方式，可以更充分满足综合业务的承载需求，实现多业务承载时的资源统一协调和控制层面统一管理，还可以提升运营商的综合运营能力。

(4)成熟的标准和良好的互通性。IP RAN技术标准主要基于Internet工程任务组(IETF)的MPLS工作组发布的RFC文档，已经形成成熟的标准文档百余篇。IP RAN设备形态基于成熟的路由交换网络技术，大多是在传统路由器或交换机基础上改进而成，因此有着良好的互通性。

3 关键问题与发展方向

基于IP/MPLS组网的IP RAN应用于移动回传以及多业务承载依然面临多方面的挑战，技术和产品也需要不断地改进和发展。

(1)规模组网问题。IP RAN的组网规模是业内一直有争议的问题之一。从目前由路由器组网的现网部署情况看，还没有上千个节点的单个IP承载网存在。但是从全网角度看，整个互联网就是由自治的多个IP网通过边界网关协议(BGP)注入形成的完整网络。因此，基于IP/MPLS的IP RAN网络也可以采用分域管理，不同的域使用不同的内部网关协议(IGP)协议，并互相使用静态路由注入的方式解决规模组网的问题。静态路由配合动态路由，也利于网络路由收敛、故障恢复和自愈。

(2)OAM管理问题。传统路由器组网的网络配置管理是采用命令行

方式(CLI)。命令行方式的特点是可以使用各种平台和网络，配置速度快、命令丰富。而传统传输设备如MSTP的配置方式是图形界面(GUI)，特点是配置直观，适合批量管理，使用简单。为了减少管理方式变化给运维带来的影响，部分路由器厂家已经开发了基于图形界面的管理方式，并遵循了传统的MSTP网络管理习惯。由于IP RAN的承载方式打破了运营商传统传输专业的运维和管理思路，如何平稳过渡还需要在实践中进一步探讨。

(3)保护恢复问题。IP/MPLS采用了快速重路由(FRR)机制可以提供50 ms级别的故障恢复，但属于局部网络保护方式，当链路或节点故障发生在TE域之外，系统的故障恢复需要IGP收敛实现，整网保护倒换可能在几百ms左右。

(4)端到端QoS保障问题。在传统IP承载网中，高品质的QoS保障往往要靠大带宽轻载来实现，网络带宽的利用效率较低。为此需要考虑部署端到端的QoS解决方案，以提高网络利用率。

(5)与现网的互联互通问题。由于运营商原有MSTP部署的规模较大，虽然MSTP提供了以太网接口，以满足IP化和多业务的承载，但内核仍为TDM。另一方面MSTP承载了现网中大部分业务，新部署的IP RAN还不能完全取代MSTP，业务的割接有一个渐进过程，因此在MSTP与IP RAN共存的情形下，必须解决MSTP与IP RAN的互联互通问题，包括业务的互联互通、OAM的互联互通以及网络保护的互联互通。

4 IP RAN与MSTP的互联互通

现阶段，IP RAN将主要承载移动基站回传、集团客户业务、L2和L3专线等电信级业务；而传统的TDM业务目前还是由MSTP网络进行承载。但是随着3G基站量的不断增长和HSPA+的规模部署，MSTP网络难以满足移动回传新增的带宽需求，采用IP RAN网络逐步代替MSTP网络已是运营商网络发展的重要趋势。对于MSTP网络应严格控制建设规模、充分挖掘其潜力。但由于在较长时间内存在MSTP与IP RAN共存的局面，因此MSTP与IP RAN的互联互通成为了网络演进中的重要课题。

4.1 基于UNI对接的IP RAN与MSTP互通组网技术

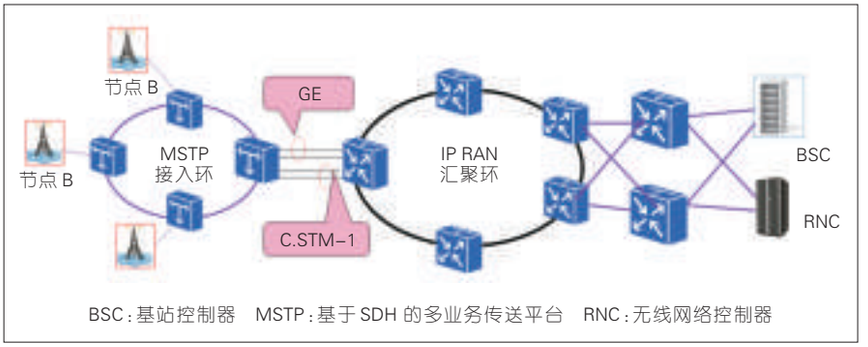
如图1所示，在基于用户节点接口(UNI)对接的场景中，MSTP设备和IP RAN设备互通组网所承载的TDM业务和数据业务分别通过SDH接口和GE接口进行互通传输，因此要求互通的MSTP设备和IP RAN汇聚设备需要同时具备GE接口和SDH接口。但是由于TDM业务和数据业务的配置是相对隔离的，即是在IP RAN网络中和MSTP网络中各自配置其端到端的业务，因此如何实现网络的时钟同步不可避免。而在这一点上，我们建议可采用线路抽取时钟和外定时等方式实现，从而解决整个MSTP网络和IPRAN网络互连组网的时钟同步问题。

在运维方面，MSTP网络和IP RAN网络采用不同的网管操作界面，IP RAN网络应该延续MSTP的网管理念，维护界面清晰，以快速告警和定位网络的故障。对于在实际组网中遇到的异厂家组网的情况，不同厂家MSTP和IP RAN网络混合组网的问题也可以依据此方法实现对接。

在OAM方面，MSTP和IP RAN设备各自采用其OAM标准，MSTP设备采用基于SDH开销的OAM管理机制，而IP RAN设备依据设备形态的不同可以采用基于G.ACH+Y.1731或者双向转发检测(BFD)的OAM管理机制。SDH接口和GE接口互通部分的OAM是采用基于802.3ah或者802.1ag的接入链路OAM管理机制。

在保护倒换方面，MSTP部分基于SDH的复用段保护实现小于50 ms的保护倒换，而IP RAN设备采用基于1:1/1+1线性或者环网保护实现小于50 ms的保护倒换。在互通部分，对于GE接口的数据采用以太网的链路聚合组(LAG)保护，保护倒换时间小于200 ms；而对于SDH接口的业务，采用1:1/1+1 MSP保护，实现小于50 ms的保护倒换。

基于UNI接口实现MSTP和IP RAN网络的互连互通也有如下两个方面的局限性：一方面，基于UNI互通会在IP RAN网络与MSTP网络边界存在单节点失效的隐患。由于在本互通场景中，不论是IP RAN网络和MSTP网络，均只有一个节点进行互通，链路的保护可以通过SDH复用段和MSP线性保护实现，但是对于节点的保护并不完备，因此存在MSTP网络或者IP RAN网络互通节点失效时，网络中断的隐患。另一方面，基于UNI互通会使现有部分IP RAN设备缺乏支持STM-4业务承载能力。目前来看，各厂家的IP RAN设备对



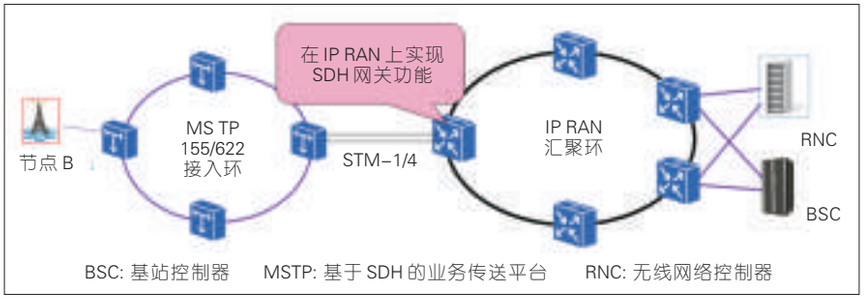
BSC：基站控制器　MSTP：基于SDH的多业务传送平台　RNC：无线网络控制器

▲图1 采用UNI对接的IPRAN技术与MSTP的互通组网方案

TDM 业务的支持程度不一,并不是所有分组均支持 STM-4 的业务承载能力,当 MSTP 通过 STM-4 进行 TDM 业务的 UNI 口互通时,会存在 GE 接口的数据互通,而 SDH 接口的业务无法互通的状况。

4.2 基于 NNI 对接的 IP RAN 与 MSTP 互通组网技术

如图 2 所示,在本场景中,MSTP 网络的接入环和 IP RAN 网络的汇聚环所承载的业务是利用 NNI 接口实现对接,即通过 SDH 接口进行互通传输。在 IP RAN 网络中,IP RAN 汇聚节点设备需要提供网关板卡支实现 SDH 信号到 PW/LSP 信号的转换功能。SDH 设备中的以太网业务和 TDM 业务适配到 IP RAN 设备中的以太网业务和 TDM 业务进行传递,SDH 中的 STM 和 VC 层的承载与 IP RAN 设备中的 PW 和 LSP 层适配,从而实现业务的互通。



▲图 2 采用 NNI 对接的 IP RAN 与 MSTP 互通组网

另一方面,由于 MSTP 网络采用的是 SDH 的告警机制,而 IP RAN 网络采用的 Pseudowire (PW)、标签交换路径(LSP)、段层的多层次告警机制或者基于 BFD、LSP ping 和 Traceroute 等 MPLS 的告警方式。因此分组设备的网关板卡需要支持 SDH 信号、VC 12 告警与 IPRAN 网络 PW 告警的无缝转换,从而实现互通组网中告警 OAM 的端到端管理。

对于基于 NNI 接口实现 MSTP 网络和 IP RAN 网络的互联互通方式,其保护倒换功能也是需要重点关注的方面,因为这直接关系到该方式的互联互通能否满足电信级网络的生存要求。在前面所讨论的基于网关板卡实现 MSTP 网络与 IP RAN 网络互通组网的端到端 OAM 对接与转换的基础上,可以通过 OAM 迅速定位故障的程度和类型,及时触发网络保护倒换,从而实现基于电信级的 50 ms 端到端保护。

由于在互连互通时能实现端到端的 OAM 互通,又能满足端到端 50 ms 保护倒换要求,基于 NNI 接口的互通方式比基于 UNI 接口的互通方式在网络的 OAM 和保护倒换方面有优势,但是通过 NNI 接口对接存在如下的问题:一方面,由于通过 NNI 接口实现对接需要 IP RAN 设备具备 SDH 的网关板卡,但是目前业界仅有一部分厂家可提供 SDH 网关板卡;另一方面,在进行 MSTP 网络和 IP RAN 网络之间的业务适配时,SDH 和以太网开销转换比较复杂,目前业内还没有针对这个问题提出解决方案,缺乏相关标准规范的支持,因而无法实现异厂家设备的互通,这为 NNI 接口对接方式的广泛应用带来了挑战。

5 结束语

随着移动通信日趋宽带化和 IP 化,基于 TDM 的 MSTP 无论从容量还是技术上都无法满足移动回传的需求,建设新型的分组化移动回传网势在必行。在此背景下,基于 IP/MPLS 组网的 IP RAN 成为了重要的技术选择。IP RAN 采用成熟的 IP 组网技术,同时吸取了传统传输网的管理理念,是实现移动与固定宽带业务统一承载的重要手段。但 IP RAN 设备和网络依然处于发展过程中,面临组网模式、QoS、保护、运维等多方面的挑战,有待进一步改进。尤其是在 MSTP 已大量部署的情形下,如何实现 IP RAN 与现有 MSTP 的互通和融合,确保传送网由 TDM 向分组化平滑演进是需要关注的关键问题。总之,IP RAN 作为新一代移动承载网必将随着移动宽带的发展得到更广泛的部署,其技术也需要在实践中不断发展和完善。

6 参考文献

[1] NIVEN-JENKINS B, BRUNGARD D, BETTS M,et al. Requirements of an MPLS Transport Profile [S].IETF RFC 5654. 2009.
[2] 张沛,赵正一,简伟等.城域网的业务需求及其与 MSTP 网络互联互通场景分析[J]. 信息通信技术,2012 27(2): 6-12.
[3] 唐雄燕,张沛. IP RAN:移动回传向全 IP 化演进[N]. 人民邮电报, 2012-05-10.

收稿日期:2012-08-22

作者简介

唐雄燕,中国联通研究院副总工程师、博士、教授级高工,兼任北京邮电大学兼职教授、博士生导师,中国通信标准化协会泛在网技术工作委员会副主席;长期在电信运营企业从事宽带通信和信息应用方面的研发和技术管理工作;已出版专著 5 部,发表技术论文 100 余篇。

简伟,北京邮电大学和美国佐治亚理工大学联合培养博士;现工作于中国联通研究院网络技术中心;研究方向包括(超)100G 光传输技术、光/分组传送网、超高速毫米波无线传感器网络以及毫米波无线通信系统等方面的研究。

张沛,北京邮电大学光通信中心博士毕业;现工作于中国联通研究院网络技术中心;研究方向是高速光传输系统、承载网关键技术研究;向 ITU-T 提交国际文稿二十余篇,发表文章十余篇,专利 3 项,著作 3 本。

对宽带的再认识
Re-Understanding Broadband

中图分类号：TN929.5 文献标志码：A 文章编号：1009-6868 (2012) 06-0042-05

摘要：文章从宽带发展的现状、内涵、作用及前景4个方面，对当今宽带进行了再认识。文章认为宽带不仅是推进信息化、持续发展经济社会的重要基础设施，而且是在网络空间中维护国家主权和安全的坚实后盾；宽带还能够提高整个经济领域的生产率和效率，从而拉动GDP的发展；宽带是实现未来可持续发展目标（SDG）的基本技术。文章还指出中国的宽带的建设，需要注入更多的投资，并在政府的带领下落实各种政策，才有可能更好地发挥宽带的巨大作用。

关键词：宽带；GDP；可持续发展

Abstract: In this paper, we look again at the status, connotations, effects, and prospect of current broadband. We suggest that a) broadband can be the critical infrastructure for economic and social development and can provide a solid backing for national sovereignty and security in network cyberspace; b) it improves economic productivity and efficiency and can boost GDP; and c) it is the basic technology for future sustainable development goals. Broadband construction in China needs more investment, and government policies.

Key words: broadband; GDP; sustainable development

雷震洲/LEI Zhenzhou

（工业和信息化部电信研究院，北京 100191）
(China Academy of Telecommunication Research of MIIT, Beijing 100083, China)

- 低收入经济体的宽带拉动作用大于高收入经济体
- 宽带具有应对可持续性挑战和气候变化挑战的独特潜力
- 政府最重要的角色是把宽带放在优先发展的位置，领导、制订和实施发展战略、行动计划和配套政策

1 全球宽带发展迅猛

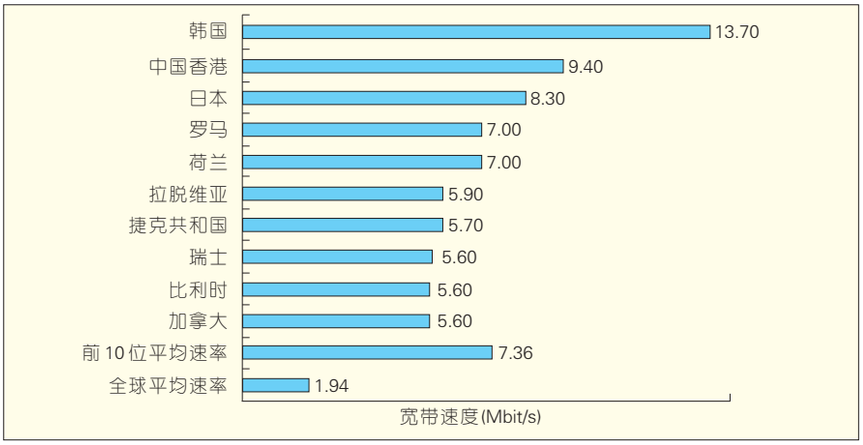
目前，宽带连接正在全球扩展。2005—2010年，有线宽带普及率平均增长59%，从每百人3.3个用户增长到8个。全球移动宽带用户在2011年已经突破10亿，占总移动用户数的五分之一。表1按地区列出了2011年6月每百人的有线和无线宽带用户数[1]。按此推算，截止到2011年6月，有线宽带用户数和无线宽带用户数大约分别达到5.67亿和9.5亿。

另外，宽带平均速度也在逐年提高，2010年互联网全球平均连接速度已达到1.9 Mbit/s。排在最前面的10个国家，其连接速度已经达到或超过5 Mbit/s的“高宽带门限”（见图1）[2]。宽带终端设备则在能力、移动性和便携性方面得到了明显提高。芯片主频已经从600 MHz全面进入1 GHz时代，终端功能不断丰富，产品形态日趋多样化，已从普及时代进入了细分化服务时代。最典型的例子是平板电脑在2010年第3季度的出货量比第2季度高45%，2011年的出货量超过2010年160%以上，预计2012年将超过2011年60%。

宽带连接水平的提升，为开发新的、更健全、更有价值的业务、应用与内容创造了良好条件。以下由宽带促成的业务、应用和内容越来越成为

▼表1 每百人有线和无线宽带用户数(2011年)

地区	每百人有线宽带用户数/个	每百人无线宽带用户数/个
撒哈拉以南非洲	0.3	2.9
东亚与太平洋	10.5	16.6
东欧与中亚	9.2	14.5
欧盟与西欧	27.6	45.9
拉美与加勒比海	7.1	12.2
中东与北非	2.5	13.1
北美	24.5	34.0
南亚	0.8	1.6
全球	8.1	13.6



▲图1 平均宽带速度排在前10位的国家与地区(2010年)

个人和企业日常不可缺少的一部分，并将进一步激发对宽带业务的需求。它们包括：

• 视频。内容越来越丰富、质量越来越高的视频将继续推动对大容量宽带业务的需求。

• 各种应用和云计算。由云提供的各种应用正在日益推动宽带的使用和发展，尤其在无线宽带环境下。诸如办公套件之类的生产性应用已经开始面市(Google与Microsoft等公司都有提供)。

• Web 2.0应用。利用计算和连接技术的进步，创造了以用户为中心的可互操作的协作环境，在此环境中，用户可以实时制作、分发和分享内容。

• 社交网。应用越来越细分的社交网现已深受人们欢迎。Facebook的5亿多活跃用户每月分享300亿件内容(例如，照片、视频、更新材料、web链接、新故事和博客帖子)。YouTube已成为全世界最受欢迎的在线视频共享网站，而其内容的70%是在美国境外创作的。宽带将进一步推动这些社交网应用的使用，继而成为宽带需求的主要驱动力。

然而，在我们步入宽带时代时，还必须看到发达国家与发展中国家之间的数字鸿沟依然存在。在发展中国家每百人只有4.4个宽带用户，而发达国家为24.6个。平均宽带速度排在最前面的10个经济体中，没有一个是在南半球的。对数字鸿沟的度量已从过去的接入差异变为如今的接入质量差异，下载速度慢意味着可能丢失经济机遇。

2 宽带内涵更为深刻

一提到宽带，人们马上就会联想到速度，因为以往都是用速度来定义宽带的，但是这样的定义并不科学。首先，由于每个国家的独特需要和不同情况，基于速度的宽带定义在各个国家和机构之间是不同的。例如，在印度、南非等国家中，ITU和OECD规定下载速度在256 kbit/s以上即为宽带；而在加拿大等国家ITU和OECD则规定高于1.5 Mbit/s才为宽带。其次，基于速度的定义跟不上技术的进步，或者跟不上特定业务与应用所要求的速度。当初定义的256 kbit/s或1.5 Mbit/s还算高速，若干年后就沦为低速了。因此，任何基于速度的宽带定义都需要不时更新。再有，这样的定义并不能反映最终用户得到的实际速度，宽带提供商宣传的速度可能比政府定义的宽带速度高很多。

实际上，现在对宽带的定义并非只是纯技术。它涉及网络和业务的多个方面，通常包括：用来向用户交付业务的基础设施或管道；能高速接入的互联网；通过宽带网向用户提供的业务与应用，如IPTV、“三重播放”套餐等。并且进入21世纪以后，时代赋予宽带的内涵远比简单的定义深刻。

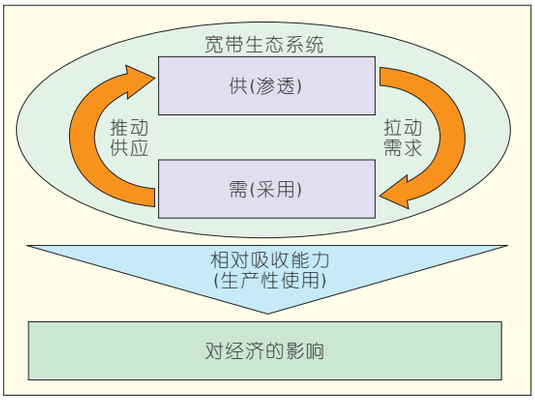
宽带是21世纪的新一代信息高速公路。宽带标志着从20世纪资源密集型物理基础设施向21世纪高效率信息基础设施的转变。宽带既促成了“大数据”时代的到来，又成为承载大数据的大管道。宽带是完成信息化、构建信息社会必不可少的国家基础设施。

宽带是一个信息通信技术(ICT)使能平台。在过去的10年中，基于ICT的互联网发生了翻天覆地的变化，它从一个传递信息的工具变为一个使能平台，不仅为个人、企业、社会和政府提供了无数不断延伸的新用途，而且还为它们提供了一个激发智慧、鼓励创新的环境。而宽带可以把ICT的益处进一步延伸到整个经济领域，成为各行各业的重要输入，包括教育、医疗卫生、交通运输、能源和金融等。宽带为每个国家提供了一个推动经济发展、延伸公共服务、加强企业实力和惠及人民的使能平台。

宽带是一种通用技术(GPT)。GPT的概念诞生于20世纪90年代，它包括3个关键特点：可以在许多行业中普遍使用；技术上具有动态性(技术演进的固有潜力)；随着GPT的演进、提高和在整个经济领域的推广使用，将提高整个国家的劳动生产率。GPT具有影响整个经济的能力，能在根本上改变经济活动如何组织和在何处组织。最典型的例子是电，200年来电作为GPT进入各行各业、家家户户，极大地影响到每个经济体的生产率、经济增长及创新，甚至重新定义了各经济体的运转方式。如今，宽带同样能够成为其他行业的重要输入，影响一个国家和地区的整个经济。宽带是21世纪助力经济转型的GPT。

宽带是一个生态系统。为了形成一整套条理清晰、相互协调的政策，并使各行各业与社会各界能最大

程度地获得宽带全部的潜在益处，必须把宽带看作一个由供给和需求构成的生态系统，如图2所示。供给主要指宽带网络平台，包括国际连接、国内骨干网、城域连接和本地连接4个基本要素。需求主要靠业务、应用和内容的不断开发、创新与推广使用来拉动。但是形成良性互动的供和需只是促进宽带发展的必要条件。在此生态系统中，还需要有一种体现吸收能力的机制，才能够最终保证宽带作为一种通用技术而释放出它作用于经济社会领域中的全部潜力。其吸收能力主要表现为：企业、政府和民众创造和开发能够提高效率和劳动生产率的宽带业务与应用的能力；企业、政府和民众以高效、生产性方式使用宽带业务与应用的能力。一个国家即便宽带覆盖全国并广泛采用，如果其吸收能力有限的话，它也只能获得有限的整体经济以及社会效益。



▲图2 宽带生态系统及其对经济的影响

3 宽带影响日趋重要

3.1 保障国家安全

宽带不仅是推进信息化、持续发展经济社会的重要基础设施，而且是在网络空间中维护国家主权和安全的坚实后盾。

网络空间是英语Cyberspace的译名，该词源自于美国科幻作家William Gibson1984年的科幻小说《神经漫游者》(Neuromancer)。不过现今我们所说的网络空间，指的是由互相依存的信息基础设施、通信网和电脑系统构成的全球网络。在这个广袤的空间里，看不到物理世界，只有许多庞大的信息库和高速流动的各种信息，但人们照样可以在其中交换思想、分享信息、经营事业、指导行动、创办媒体、畅玩游戏、提供社会支持、开展政治讨论等等。实际上，互联网现在已经成为最大的网络空间。

美国为了在网络空间保持世界霸主的地位，把布什政府在2004年制订的“确保网络空间安全的国家战略”于2011年进一步提升为奥巴马政府的“网络空间国际战略”，把网络安全提升为与经济和军事安全同等重要的位置，把网络空间视作新的战略制高点。该报告宣称，针对重大网络攻击，美国将保留使用所有包括外交、情报、经济和军事一切必要手段的权利。

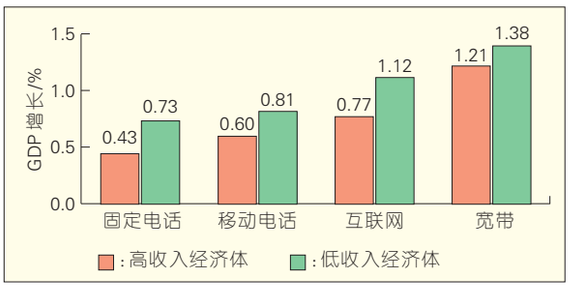
如今，网络攻击的水平越来越高、种类越来越多，来自网络的威胁明显增加(包括频次、影响、规模和代价)，网络犯罪正席卷全球，并有可能成为世界上最危险的犯罪威胁之一。针对网络安全的“攻”与“防”也在不断升温，不少国家已经成立“网络司令部”并相继进行网络战演习，网络战绝非子虚乌有。将来，如果防备不当，来自恐怖组织或黑客的网络攻击很有可能造成一次“数字911事件”，使核电厂和供电设施等基础设施瘫痪，或者无须任何潜伏凶犯登机就让客机坠毁。

网络安全的“攻”与“防”、恐怖事件或其他重大突发事件的处理都必须要有宽带网络作为坚实后盾。在网络空间中，同样甚至更加需要“兵贵神速”。美国电视连续剧《反恐24小时》讲的是如何在24小时内搞定恐怖事件的故事，其中充分表现出高速网络在处理恐怖事件中的威力，连指挥外勤人员执行任务都是通过网络实现的。可以说，没有宽带，在网络战中必败无疑；没有宽带，在应对恐怖、动乱事件或重大灾难时就难以速战速决。在美国911恐怖事件中，曾经有过这样的教训：当时由于没有足够的宽带接入，使一些本可以用会议电视来处理的事不得不通过面对面会议来处理，既增加了交通负担，又延误了救援行动。难怪美国政府把网络空间的战略地位提得如此之高。

3.2 拉动GDP发展

宽带也可以拉动GDP的发展，因为除了发展宽带自身的拉动作用外，宽带在各行各业中得到使用后，还能够提高整个经济领域的生产率和效率，必然对GDP产生正面的影响。

据世界银行研究发现：在2000—2006年期间，低收入经济体的宽带普及率每提高10%，人均GDP就可以增加1.38个百分点；而高收入经济体的宽带普及率每提高10%，人均GDP仅增加1.21个百分点(见图3)。低收入经济体的宽带拉动作用大于高收入经济体。这表明低收入经济体的生产率和效率提升空间更大。另一个发现是：宽带潜在的经济拉动作用大于其他ICT领域，包括固定电话、移动电话与互联网。在低收入经济体，固定电话、移动电话、互联网普及率每提高10%对人均GDP的拉动作用



▲图3 各种ICT在高收入和低收入经济体中对GDP增长的影响(2000–2006年)

分别为0.73%和0.43%、0.81%；而在高收入经济体，此组比例为0.60、1.12%和0.77%。

关于宽带对GDP拉动作用的研究中，宽带的拉动作用大小的估算不尽相同。例如，据麦肯锡估算，一个国家的宽带家庭普及率每提高10%，将使GDP增加0.1～1.4%；Booz & Company发现，在经济合作与发展组织(OECD)国家中，GDP平均年增长与宽带普及率之间有很强的相关性，宽带普及率排在靠前的国家，其GDP增长比宽带普及率排在靠后的那些国家高2%。综合多个研究可得宽带普及率提高10%，能使GDP增长率达到约0.24~1.50%。

宽带可以通过3个渠道来创造就业机会：为了部署宽带基础设施直接创造的就业机会；部署宽带基础设施过程中间接产生的就业机会；由于宽带网络的外部经济效应和意外效应创造的就业机会。全世界各个国家通过多项研究得到的结论是：每增加1 000个宽带用户，将创造大约80个新的就业机会。2006年由于宽带的部署，整个欧洲创造了105 000个就业机会；2009年，12个拉美国家宽带普及率提高大约8个百分点，结果使它们的就业率平均提高了将近12%。

宽带对企业的运行、研发与创新具有潜在的正面作用，因为宽带能够促进更大范围的联网；促成知识与构想的及时共享，加速创业；降低产品和工序创新的壁垒；改善企业协作，尤其是促成小企业与大型研发机构的协作，扩大它们的研发能力；缩短从构想到最终产品的时间；推动用户主导的创新。例如，Apple的iPhone应用商店拥有10万个注册的应用开发商，其中大多数是小公司。2008年，这些小公司开发了在应用商店中销售最好的10款应用中的5个。

宽带对发展中国家而言，在制造业、零售业、服务业、教育、卫生等方面尤其能起到很好的促进作用。例如，像印度、毛里求斯和中国现在都是承接外包服务的主要离岸地，通过宽带可以与客户快速沟通，改善客户关系，使外包服务成为它们持续高增长的一个新途径。又如，通过移动教育实验室(或移动学校)的部署可以把宽带接入和ICT扩展到农村和偏远地区，从而改善教育状况。蒙古在21个省里部署了100个移动“帐篷”学校，效果十分明显。移动宽带还是医疗工作者联络患者的一个极有用的工具。目前在全世界60多亿人口中，医生和护士不到2 700万，在最低收入国家中只有120万医生和护士，有了移动宽带就可以在那里实现远程医疗。一般的远程医疗需要10到100 Mbit/s的速度，高清远程医疗则需要100 Mbit/s以上的速度。卢旺达的远程医疗计划正是因为缺乏高速宽带连接而被推迟。

以下几个案例是值得大家进一步关注。早在2005年，加拿大工业部在不列颠哥伦比亚省对用户作了一次关于宽带接入重要性的调查，80%以上的受访企业认为如果没有宽带接入，它们的企业将受损；18%以上的受访企业表示如果没有宽带，它们的企业就无法运作；62%的企业称，宽带提高了企业的劳动生产率，其中多数企业的劳动生产率因宽带提高了10%以上。此案例充分说明了宽带对提高企业效率和生产率的重要性。如果各行各业都认识到这一点，对宽带给予互补性的投资或出台相应的宽带支持政策，整个国家的GDP就可能得到富有成效的增长。仅仅凭借ICT业一家的力量是远远达不到这一目的的。

印度也对移动宽带的GDP拉动作用进行了研究，其结论是在2015年以前，移动宽带普及率每提高一个百分点将使GDP增加0.11%。其拉动作用主要来自直接贡献(宽带业务与移动终端设备收入)、间接贡献(劳动生产率提高带来的收入或成本节省)、生态系统贡献(移动宽带促成的附加值及其他服务收入)。南非也作了类似研究，预计到2015年，无线宽带能使南非GDP增加1.8%(超过94亿美元)，还能创造28 000个直接就业机会。这两个案例说明，移动/无线宽带对农村地区广阔的新兴国家而言，其经济拉动作用是不可小视的。

3.3 促进可持续发展

2012年4月23日联合国数字发展宽带委员会向全世界，尤其是向各国领导人、政策制订者和产业界领袖发出“行动呼吁”(Call to Action)，殷切希望在2012年6月即将召开的联合国可持续发展大会上，整个国际社会能够充分认识到宽带推动社会进步和可持续发展的潜在力量，把“宽带惠及一切”列为联合国可持续发展目标(SDG)之一。

当今人类面临的最大挑战之一是通过可持续发展来改善全世界人民的生活。要可持续发展就必须管好地球上的宝贵资源，既要让有限的资源满足可持续发展的长远需要，又要实现消除贫困的目标。这是一个极富挑战性，但在全世界共同努力下有望完成的任务。

宽带是实现联合国未来SDG的基本技术。宽带具有应对可持续性挑战和气候变化挑战的独特潜力，在迎接挑战的同时将加快经济社会发展并提高人们生活质量。在宽带的支持下，不断涌现的应用与业务创新能把可持续发展的三大支柱——经济增长、社会发展和环境保护集成在一起，使我们更加有效地使用自然资源，更好地保护环境，形成可持续的发展模式，创造可持续发展的绿色和低碳经济。

宽带还能促进社会平等、信息流通、提供通信、实时获取与交换在线信息和知识。并且宽带还能使人们对其所面临的各种挑战与问题有更好的了解，帮助他们作出更明智的选择和决定。因此，联合国数字发展宽带委员会认为，对所有国家而言，“把宽带接入普及到每一个人”应该被列

为具有最高优先权的政策。

宽带基础设施建设和宽带应用与业务的开发肯定需要可观的投资，但投资将换来长期的可持续经济发展、劳动生产率的提高、就业机会的增加和可观的经济回报。对发展中国家而言，如果想消除贫困和参与21世纪的数字经济，就必须对宽带给予足够的投资。千万不要因为投资不足而享受不到宽带带来的经济社会效益（如远程医疗和远程教育等），把自己置于更加落伍、难以脱贫的困境之中。

4 宽带发展前途无量

表2是从500 kbit/s到10 Gbit/s所支持的各种业务与应用，从中可见宽带速度越高，所支持的业务与应用的价值越高、对经济社会发展的推动作用越大。世界正在大踏步走向宽带，为的是让人类在网络空间上开创新的工作方式、生产方式、管理方式、商贸方式、金融方式、思想交流方式、文化教育方式、医疗保健方式以及消费与生活方式。但是，宽带业务与应用比具有100多年历史的电话业务要复杂得多，市场弹性很大，可涉及的领域甚至是无止境的，短期内难以搞清人们真正的宽带需求是什么。实际上，传统的宽带已经走到极限。现在几乎所有新的技术和业务，诸如高清IPTV、移动电视、游戏和互动性娱乐以及新兴的云服务等都需要比传统宽带高得多的速度。

实际上宽带能覆盖的应用范畴远远不止表2所列的这些。随着宽带的繁荣发展，相信所有对“发展宽带”的疑虑都将烟消云散。

联合国数字发展宽带委员会最近呼吁：把宽带构想变成实际行动。全球领导人要让宽带政策遍及世界，制订政策首先把宽带接入扩展到医疗卫生和教育领域；产业界要通过积极投资宽带建设、开发创新型商业模式等行动作出贡献；媒体和网络要广泛宣传宽带对可持续发展的一些深远作用。

显然，政府在发展宽带过程中，最重要的角色是把宽带放在优先的位置上，领导制订和实施发展战略、行动计划和一系列支持宽带网络建设和宽带业务提供的配套政策，往往涉及市场监管、普遍接入与普遍服务、灵活许可、基础设施直接投资、消除瓶颈以及市场税收等方面的政策。

5 结束语

世界正在从窄带走向宽带。迄今为止，全球已有110多个经济体发布了宽带战略或计划，还有22个经济体正计划近期发布。国际电联（ITU）也提出了新的全球宽带发展目标，并呼吁所有经济体2015年都要出台宽带计划。这充分说明发展宽带高度凝聚了全球共识。

中国的“宽带中国”战略已经启动，“宽带普及提速工程”也将全面实施。然而，中国是一个地域辽阔、发展不平衡的新兴经济体。只有在政府的带领下，让各项政策真正落地，各行各业和社会各界都行动起来，“宽带中国”战略才能顺利实施，并可能与国际社会一起为实现全球可持续发展目标做出贡献。

6 参考文献

[1] 李德毅.云计算支撑信息服务社会化、集约化和专业化[J].重庆邮电大学学报（自然科学版），2010,22(6):698–702.
[2] 夏侯允德.三网融合背景下IPTV何去何从[J].数字通信,2010,(4):41–44.

收稿日期：2012–10–08

作者简介

雷震洲，教授级高工，工业与信息化部电信研究院科技委副主任委员，现为中国人民政治协商会议全国委员会委员、中国通信学会会士；已出版译著4部、专著4部，发表论文350余篇。

▼表2 各种业务与应用所需的上行和下行速度

500 kbit/s ~ 1 Mbit/s	5 ~ 10 Mbit/s	100 Mbit/s ~ 1 Gbit/s	1 ~ 5 Mbit/s	10 ~ 100 Mbit/s	1 ~ 10 Gbit/s
• VoIP • 短信 • 电子邮件（基本的） • Web浏览（简单网站） • 流音乐（高速缓存） • 低质量视频（高度压缩）	• 居家办公 • 文件共享（大文件） • 标清IPTV（多频道） • 可交换数字视频 • 标清视频点播 • 标清广播视频 • 视频流（2–3个频道） • 高清视频下载 • 低清遥现 • 游戏 • 医疗文件共享（基本的） • 远程诊断（基本的） • 远程教育 • 建筑物控制与管理	• 高清远程医疗 • 多种教育服务 • 全高清广播视频 • IPTV全频道支持 • 高清视频点播 • 沉浸游戏 • 居家办公远程服务器服务	• Web浏览（综合网站） • 电子邮件（大附件） • 遥信 • 标清IPTV（1–3个频道） • 文件共享（小、中文件） • 居家办公（普通） • 数字广播视频（1个频道） • 流音乐	• 远程医疗 • 教育服务 • 标清与某些高清广播视频 • 高清IPTV • 复杂游戏 • 居家办公（高质量视频） • 高质量遥现 • 高清监视 • 智能大楼控制	• 研究应用 • 高清无压缩视频流遥现 • 现场直播数字电影 • 科学或医疗设备远程遥控 • 互动式遥视或虚拟现实 • 兆兆字节数据集搬移 • 远程超级计算

大数据场景下的云存储技术与应用

Cloud Storage Technology and Applications for Big Data

陈杰/CHEN Jie
（中兴通讯股份有限公司，广东 深圳 518057）
(ZTE Corporation, Shenzhen 518057, China)

摘要：文章认为随着大数据应用规模的扩大，新业务环境和场景对海量云存储需求的迫切，云存储平台需要打破原有的框架，改变组网和管理方式，以满足新的业务需求。文章分析了各种场景，提出了云存储的需求及关键技术等。文章指出大数据需求促进了云存储的发展，而云存储的发展则带动了新的业务应用。

关键词：大数据；云存储；安全

Abstract: With the expansion of big data applications, mass cloud storage has become a more important requirement. To meet service demands, cloud storage needs a new framework and new networking and management methods. In this paper, we discuss the various scenarios of cloud storage and discuss the demands and key technology of cloud storage. Big data requirements promote the development of cloud storage, and cloud storage development creates new service applications.

Key words: big data; cloud storage; safety

中图分类号：TP393　文献标志码：A　文章编号：1009-6868 (2012) 06-0047-05

1 大数据应用场景与需求

1.1 大数据的发展

随着互联网、移动互联网、物联网的发展，“大数据”逐渐成为发展的趋势。数据不仅仅正变得更加可用，同时也变得更易被计算机所理解。大数据发展趋势中所增加的大部分数据都来源于物联网世界中商品、物流信息，企业内部经营交易信息，互联网世界中人与人交互信息、位置信息等。

2011年分析调研机构IDC发布的研究报告——《从混沌中提取价值》显示：2011年全球被创建和被复制的数据总量为1.8 ZB，和2010年同期相比，这一数据上涨了超过1 ZB，而且这些数据大部分是非结构化的数据。预测到2020年，全球数据量暴增44倍（相比2009年），总量会达到35 ZB。根据图灵奖获得者Jim Gray提出的数据增长的经验定律——网络环境下每18个月产生的数据量等于有史以来数据量之和，因此海量数据处理的需求会变得非常普遍。

2012年，《纽约时报》称“大数据时代”已经降临，决策行为将基于数据和分析，而并非基于经验和直觉。这不是简单的数据增多的问题，而是全新的问题。旨在从互联网时代非结构化数据的庞大“宝藏”中获取知识和洞察力的计算机工具也正在迅速发展中。大数据技术领域的竞争，事关国家安全和未来，国家层面的竞争力将部分体现为一国拥有数据的规模、活性以及解释、运用的能力。2012年3月29日，奥巴马政府投资2亿美元启动“大数据研究与开发计划”，是大数据技术从商业行为上升到国家意志的分水岭。预计欧盟、中国等大型经济体很快也会出台相应倾斜性政策，大数据相关产业链公司也将迎来快速发展的机会[1]。

1.2 大数据的类型

在经历了20世纪的计算浪潮和网络浪潮之后，信息存储技术已经发展成为信息领域的三大支撑技术之一。随着云计算、物联网等信息技术的飞速发展，异构数据源越来越多，数据信息量在飞速增长，数据的类型也复杂多样，不仅使得信息系统规模日益庞大，也导致海量非结构化数据管理复杂、异构数据存储利用率低下、资源不易扩展等问题。

海量非确定性异构数据产生的原因复杂多样，在应用中也具有新的特点：随着各种应用规模及领域的扩大，数据量会呈现爆炸性增长及海量数据存储的趋势和特点；在非确定数据的典型应用中数据源很多，数据种类也繁多，数据资源具有异构性特点；数据还呈现数据块大小、数据类型和数据访问方式等不确定的特点；云计算、物联网等应用的不断丰富，数据产生、应用、访问方式十分复杂，还使得数据具有时效性和空间性，高频度访问和高并发的特点[2]。

1.3 大数据对存储的需求

非确定数据应用中的海量数据对数据的存储体系结构带来了很大

的挑战。首先，海量数据的组织必然采用分布式数据组织与管理策略，这需要实现适合于非确定数据应用的（元）数据和数据组织方式；其次，由于海量数据是通过持续增长积累而成，而积累的过程需要很长的时间，因此需要存储支持可保证规模与性能同时扩展的存储组织模式以及相应的索引机制。

针对海量不确定性数据，使用基于传统的信息存储结构和对象查询方法的实际运行效率呈现下降趋势，因此必须采用新的元数据组织结构和查询方法来提高效率，为用户提供高性能的多并发数据查询服务。

由于在分布式环境中，数据源分布在不同的网络结点，这就存在网络传输性能低的问题。而各个数据源有很强的自治性，它们可以自治地改变自身的结构和更新数据，这就会给数据集成系统的一致性带来了困难。由于数据存在非确定性，针对海量非确定性异构数据的集成工作将变得更为复杂，可以采用分布式并行处理技术实现计算资源和存储资源的全局最优化的管理。

在信息化时代和全球经济竞争的新环境下，要做出一项决策，往往需要查询多个业务系统和外部系统，并进行大量的数据分析，工作量大且易出现人为差错，影响决策的质量。这就需要对海量非确定性异构进行整合集成，整合、集成后的数据必须保证一定的集成性、完整性、一致性和访问安全性。

数据的海量性、非确定性以及异构性为传统的数据挖掘算法提出了挑战。由于数据的异构、海量、分布性和决策控制的实时性，需要调整数据挖掘引擎的布局及多引擎的调度策略。结构化或者非结构化数据都涉及数据的存储、管理（索引、并发、一致性、查询等）等，这是因为用户对大数据使用方面的要求（对海量非结构化数据查询仍然要准确和快速），导致对数据逻辑结构和物理存储方式的新要求。

随着大数据各种应用规模的扩大，数据量会呈现爆炸性增长及海量存储的趋势和特点。传统的数据存储系统达到了瓶颈，无法及时地完成各项运作任务。大数据的新特点对存储提出了新的挑战，为了适应大数据的发展，存储需要支持纵向无限扩容（存储扩容）和横向无限扩容（能力扩容），并以对象作为基本的存储形式，以提高系统的扩展性，降低系统维护复杂度。

2 云存储的关键技术

针对数据的飞速发展和数据安全要求的不断提高，如何建立安全、性价比高的存储成为业界的普遍需求。云存储成为首要选择，因为它能够根据所需容量大小对用户进行定制，用户不需要进行硬件的管理维护，缩减了用户成本和人力投入。而且云存储具有易扩容、易管理、价格低、数据安全、服务不中断等优点。

云存储是在云计算概念上延伸和发展出来的新概念，通过集群应用、网络技术活分布式文件系统等功能，将网络中大量各种不同类型的存储设备通过应用软件集合起来协调工作，共同对外提供数据存储和业务访问功能的一个系统。云存储是一个以数据存储和管理为核心的云计算系统。

大数据时代的来临对云存储提出了很多关键需求：

(1) 大规模级别存储系统的构建

随着数据的爆炸性增长，存储的规模越来越大。2012 年云存储的建设规模是几十 PB 级，存储的文件数或者对象数是几十亿。到 2013 或 2014 年，就会有百 PB 级和 EB 级的需求，过几年将会增长到 ZB 级，文件数或对对象数也会超过百亿、千亿。

传统存储通常是在一个设备、一个机架或一个数据中心内完成资源组织管理，而当存储容量上升到 EB 级或 ZB 级后存储则很难在一个数据中心内完成。大规模的存储需要跨数据中心，跨城市、省、甚至国家进行存储设备、存储数据、存储服务的组织和管理，并支持跨域的访问、备份、容灾等功能。同时大规模的存储要求存储提供不同等级的管理和服务权限，并按照区域、级别分配不同的权限。系统对资源的访问必须经过严格的权限控制。只有用户确认共享的资源才能被其他用户或业务进行访问，即使是被授权的访问也会根据不同的权限控制方式受到访问权限控制。

云存储就是将不同种类的存储设备协调起来进行工作。这些存储设备使用的存储介质也是多种多样的，而且随着技术的发展，设备种类和存储介质种类会越来越多，如何调度这些设备和存储介质协调工作，需要在云存储管理软件上考虑和优化，以保证组织好的资源被高效利用。

(2) 存储设备在线扩展和收缩

在存储设备的使用过程中，会遇到调整存储资源池的需求，这则要求存储资源池根据业务的需求增加或者减少存储设备。在调整的过程中，业务不能被中断，也不能使上层业务感受存储资源池的变化，同时被裁剪设备的数据要在较短的时间内在其他设备上恢复、备份，并在较短的时间内完成增加存储设备和原有存储你设备的数据均衡。

云存储系统要优化和调整数据组织和管理方法，即使存储规模增加后，性能要随之线性增加。数据变得庞大后，元数据管理要考虑中心化或多节点方式，以降低元数据管理对整个系统读写性能的影响。对于热点数据支持自动的多副本复制，则会在多个存储节点提供读能力，以降低硬盘、网口、处理器对性能能力的限制。采用多级缓存技术，热点数据则会先读入智能加速卡，并由智能加速卡对外提供读服务，在写数据时，也是先写入到智能加速卡，由加速卡组织分发到存储设备上。

(3) 实现面向应用的专业化管理策略

实现面向应用的专业化管理策略呈现出一些特点：传统存储系统存储资源与应用独立，存储资源利用效率低；海量存储系统把资源进行了整合，但是针对所有应用都采用统一存储策略；在云存储系统中如何做到资源整合并且针对应用进行专业化的策略管理，根据应用的变化进行弹性配额管理是一个较大的挑战。

云存储必须提供基于容器的多层次租户/应用隔离技术：系统必须提供数据隔离功能，保证数据不被非法访问，并保证用户数据的隐私。云存储可以通过物理隔离与权限控制相结合，实现对数据的隔离。

- 提供以用户为单位的数据隔离：业务系统为每个用户创建独立的存储空间，业务系统根据用户标识和对应权限对用户存储空间的数据进行访问控制，这样可以避免未授权用户访问到其他用户的数据以及用户信息。
- 提供以业务为单位的数据隔离：在进行数据的存储和读取时，每个业务都必须拥有自己独立访问的权限，系统根据不同的业务将数据隔离，避免数据被未授权的业务访问。
- 提供以存储容器为单位的数据隔离：可以设定数据存储在指定的存储容器中，不同的存储容器有不同的访问授权。访问授权可以是基于用户的，也可以是基于业务的。

云存储提供基于系统或者应用的多种服务管理策略：提供压缩策略，用户可以根据文件类型与活跃度设置压缩条件；提供系统级、业务级和用户级的流量控制策略设置；提供系统级、业务级和用户级的数据分片设置，业务可以设置对象存取的分片大小、分片存储区域（跨盘、跨节点、跨区域），同一对象的各分片可并发存取；提供系统级、业务级和用户级的热点对象设置，业务可以依据对象的访问活跃度设置热点对象和热点对象的存取方式；提供系统级、业务级和用户级的文件归档设置，业务可以设定归档区域和归档条件，包括对象活跃度、容器活跃度、文件类型等；提供系统级、业务级和用户级的隐私性保护，对象及其元数据必须归属用户，非用户授权任何用户不允许访问；提供系统级、业务级和用户级的重复数据删除策略设置，可以按命名空间、存储容器、存储区域、文件类型、执行时间等设定重复数据删除的策略，并可以查询重复数据删除的操作记录和效果分析。

云存储要提供弹性资源伸缩和共享：系统要支持根据业务使用情况自动的增加和缩减存储空间，同时利用重复数据删除技术，提升存储资源的利用率。

(4) 系统全局自动负载均衡

在云存储的系统中，物理存储主机节点规模从几万到几十万，多为数据密集型应用，比如每天亿次级别的网络搜索访问。面对超大规模的数据请求和节点数量，如何高效进行节点负载均衡，如何发挥空闲节点的作用是保障高水平服务质量，提高系统运行效能一个较大的挑战。

云存储系统要求云存储具备基于服务质量(QoS)的多层次自动负载均衡与调度功能。

- 实现基于请求类别及前端节点负载进行的均衡和差异化调度：系统参照业务诉求和区域的QoS信息（存储总容量、总的IO吞吐能力、当前系统繁忙程度等）为业务选择最合适的区域归属，如对于IO要求较高的应用可以放到IO吞吐能力较强的区域里。
- 实现基于请求类别及数据分布进行的均衡和差异化调度：系统必须提供对多组云存储系统之间的动态调度能力，并根据区域内每组系统的IO繁忙程度，将业务的访问请求尽可能发送给那些IO负荷不重的组，以实现组间IO的负载均衡。
- 实现数据中心之间的负载调度，均衡各中心利用率：通过调度服务和资源策略实现资源的跨域整合，还可以通过访问重定向和数据迁移等多种技术手段对外提供统一的存储资源服务，并对用户屏蔽资源的具体位置信息并自动实现就近访问。

此外，云存储设备还采用数据压缩技术构建分布式缓存系统，提高给定缓存加载的数据量，提升系统性能 。同时，在广域网数据传输前进行重复数据检测，相同数据只传一份，就可以实现基于删冗的广域网数据传输加速。

(5) 云存储数据安全和数据保护

云存储系统需要支持的用户数量巨大，且存储了用户生活、工作、学习等各种类型的数据，具有私密性，另外对于数据的可靠性和完整要求也非常高。因此如何解决用户数据的共享和隐私保护之间的矛盾、用户数据的可靠保护和存储高效之间的矛盾是一个很大的挑战。

系统对资源的访问必须经过严格的权限控制，只有用户确认共享的资源才能被其他用户或业务进行访问，即使是被授权的访问也会根据不同的权限控制方式受到访问权限的相关控制。

云存储还需要具备以下的一些基本功能：

系统必须提供数据隔离功能，以保证数据不被非法访问并保证用户数据的隐私。通过物理隔离以及与权限控制相结合，可以实现对数据的隔离。

系统必须提供信息加密的功能，防止用户信息被盗取。用户的关键信息，如登录密码和系统访问等其他鉴权信息，无论是传输时还是在存储时必须加密。

系统必须提供数据传输加密功能。数据的传输加密可以通过客户端软件的传输设置实现。用户设置采取加密通道传输，系统应当在重新登录后进行数据传输时使用Https通道进行数据的传输。

系统必须提供有效的硬盘保护形式，保证即使硬盘被窃取，非法用户也无法从硬盘中获取有效的用户数据。

系统必须支持数据加密存储，用户在使用客户端软件时可以选择对存储数据加密。采用数据加密存储的客户端软件在上传数据时对数据进行自动加密，在线备份获取加密数据后能够在客户端自动解密，而在线存储获取加密数据后必须手动解密。

系统必须将数据切片存储在不同的云存储节点和硬盘上，数据无法通过单个硬盘恢复。故障硬盘无需进行数据清除即可直接废弃，用户数据不会通过硬盘泄露。

(6) 云存储节能降耗

云存储系统规模巨大，且需要提供高质量的对外服务，传统的构建方法提高性能和能耗增长近似线性关系，需要用新系统的架构打破这种关系，解决系统性能和能耗的矛盾。

- 建立不同级别的存储池：高写高读、高写低读、低写高读、低写低读等，并按照业务模型进行资源分配。对于不同的存储资源池采用不用的策略：如重删、压缩、按需分配等，提高存储的利用率，从而降低系统能耗以及二氧化碳的排放量。
- 监控业务访问量，控制系统内部部分设备可以轮流进行休眠及CPU降频。同时考虑使用低功耗的处理器。
- 控制业务流向能源供应充分或者消耗能源少的存储资源池。

3 云存储的应用

从2010年开始，云存储的运用越来越广泛，在互联网、平安城市、视频制作、数字传媒、家庭娱乐、个人网盘等方面都有很多应用。

3.1 视频监控应用

随着城市的现代化建设和经济的快速发展，构建和谐社会的必要性与日俱增，每个城市都在努力打造“平安城市”。平安城市等大规模高清视频监控系统中的主要问题就是如何处理庞大的高清视频数据，因此就必须从视频的采集、编解码、传输、实时监控、录像回放等环节全面支持大规模高清。这样就给高清监控系统带来了一系列问题：网络带宽紧张、存储空间庞大、对性能的要求成倍增长、系统扩容升级压力等。如果采用1 080 P的高清视频监控，即使使用具有高压缩比的H.264编解码技术压缩高清视频，输出码率也将可达到6 Mbit/s，那么每台摄像机每天将大约产生50 GB左右的数据量，一个月就是1.5 TB左右的存储量。一个城市有上千个摄像头，一个月的数据量就达到PB级以上。面对PB级的海量网络存储需求，传统的开放系统的直连式存储(DAS)和网络存储技术(NAS)在容量和性能的扩展上存在瓶颈，已经不能满足对高清视频的存储。

远程云端用户可以通过以太网访问云存储中的视频，如图1所示，云存储提供多种应用接口给视频监控系统中的应用、管理和终使用。采用基于浏览器的方法，云端用户不需要安装任何播放以及管理软件，就可以远程对监控录像进行回放、视频的再分析等。

云存储为实现大规模高清视频的存储和处理提供了一种新的解决方案。云存储突破了传统存储方式的性能和容量瓶颈，实现了性能和容量的线性扩展，让海量数据存储成为可能。同时云存储可以实现存储完全虚拟化，所有设备对云端用户完全透明，任何云端、任何被授权用户都可以通过一根网线与云存储连接，从而让用户拥有相当于整片云的存储能力。由于各个监控区域地理范围分布广阔，监控点的数据巨大，采用云存储系统，便于分布式管理和随时扩容。云存储由元数据管理节点和大量存储节点组成。元数据节点负责存储虚拟化、资源管理以及存储数据的命名空间、存取控制、存放位置等信息，是云存储的核心部分。云端不直接通过元数据节点读取数据，而是从元数据节点获取视频存储的位置信息后，直接和存储节点进行读写操作。将控制部分和业务数据部分分离，有助于提高系统的可扩展性和数据处理的读写带宽。

3.2 互联网应用

一个用户拥有多个终端的现象

DVS: 网络视频服务器 DVR: 硬盘录像机

▲图1 视频监控中云存储的应用

越来越普遍，用户的同步和分享需求则会变得越来越强烈，同时移动终端以及移动互联网的快速发展均为个人云存储的发展提供了良好的发展环境。

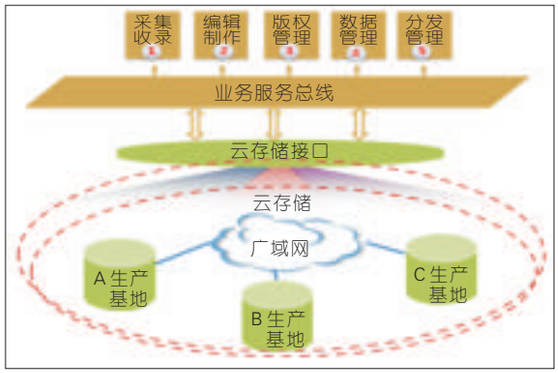
互联网公司的纷纷加入云存储的研发和应用队伍中，并将互联网产品中需要存储的个人信息与云存储应用绑定，加快了用户的接收速度。基于互联网提供的云存储平台，各类消费电子产品实现了前所未有的互联互通，包括文字、图片、音乐、视频、应用等在内的数字内容开始实现跨越时间和空间的自由流通，为社交网络中的用户提供了良好的交流元素。这些数字内容目前越来越向诸如 Dropbox、Box.net、115 网盘等一些专业云存储服务商集中。

要求云存储支持 10 PB 以上容量空间，提供网络文件系统(NFS)等文件或对象访问接口。如图 2 所示，云存储提供满足多种应用需求的接口。由于互联网对云存储的访问速度要求相对较低，关注建设成本，因此需要云存储提供纠删码方式的数据保护、提供压缩、重删等多种策略。并可以基于容器或者用户为单位对外提供租赁服务，按照空间大小、访问流量进行收费。

▲图 2 互联网中云存储提供的接口方式

3.3 视频编排应用

云存储系统通过对物理存储设备的虚拟化管理，实现视频数据与物理存储位置无关性，降低了视频数据管理的复杂性，增强了系统的灵活性和可扩展性，满足了海量视频数据的存储需求。图 3 所示说明了视频数据可以跨地域存储和调用。



▲图 3 云存储跨地域组网

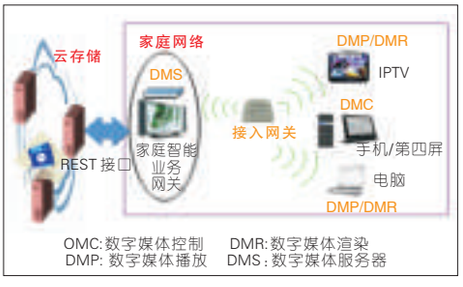
云存储系统提供高速的读写接口，满足采编等应用的视频加工需求，实现对原始素材、成品节目、再加工节目等不同类型节目的分层存储和分类管理。

3.4 家庭娱乐应用

实现家庭网络内各终端通过云存储访问代理与云存储平台交互，则可以完成媒体文件的上传和在线播放，并实现家庭网络内各终端信息在云存储上共享和备份。

基于家庭的云存储可以实现家庭内多终端间的多媒体资源共享及多屏互动、多终端交互控制，优化用户感知。

图 4 所示为家庭媒体云。通过家庭媒体共享和交互控制功能，可以实现第四屏、机顶盒等的互操作。



▲图 4 家庭媒体云

4 结束语

随着互联网、移动互联网、物联网的发展，“大数据”逐渐成为发展的趋势，数据产生的原因复杂多样，在应用中也具有新的特点。随着各种应用规模的扩大，数据量会呈现爆炸性增长的趋势及海量数据存储的特点。新业务环境和应用场景对海量云存储需求越来越迫切，这需要海量存储平台打破原有的框架，改变组网和管理方式，满足业务需求。文章分析了各种场景，提出了对云存储的需求、关键技术和指标。文章认为云存储的应用越来越广泛，云存储技术的发展促进了业务融合，并衍生出新的业务应用。总之，大数据的需求促进了云存储的发展，云存储的发展带动了新的一系列业务应用。

5 参考文献

[1] 王伟，柯尊友. 云存储的进化：云存储解决方案[J].中兴通讯技术(简讯):2012(8):18-19.
[2] 李群. 国内个人云存储应用风生水起发展迅猛[EB/OL]. http://net.chinabyte.com/115/12393115.shtml

收稿日期：2012-10-12

作者简介

陈杰，南京邮电大学通信专业、纽约大学计算机专业双硕士学位毕业；曾任中兴半导体有限公司开发部主任、AT&T 公司贝尔实验室高级研究员与研究部主任、中兴通讯美国分公司总裁、中兴通讯网络事业部总经理，自 2003 年以来，任中兴软创董事长、中兴通讯高级副总裁；1992 年获得国家科学技术进步奖一等奖，2007、2008 年分别获得国家科学技术进步奖二等奖，2012 年获得深圳市高层次专业领军人才称号。

SDPaaS云平台架构及其关键技术研究

Architecture and Key Technology of SDPaaS Cloud Computing Platform

屠要峰/TU Yaofeng，黄震江/HUANG Zhenjiang，陈心哲/CHEN Xinzhe

(中兴通讯股份有限公司 业务研究院，江苏 南京 210012)
(Communication Service R&D Institute, ZTE Corporation, Nanjing 210012, China)

中图分类号：TP393.03　文献标志码：A　文章编号：1009-6868 (2012) 06-0052-04

摘要：文章提出了SDPaaS云平台的解决方案，该平台是电信SDP和云计算PaaS的融合。SDPaaS云平台针对ICT网络融合的大环境，在支持第三方开发人员快速开发和交付业务的基础上，进一步实现了业务自动部署和托管运行、专业化集中式维护等功能。该方案是对传统电信增值业务交付方案的一种拓展，为创建良好的社会化增值服务创新的生态环境提供了必需的技术支撑。

关键词：业务创新；云计算；SDPaaS

Abstract: In this paper, we present SDPaaS cloud platform solutions, which combine telecom SDP with PaaS cloud computing. This platform allows third-party developers to rapidly develop and deliver services. It also allows for automatic deployment and hosting run of business service and has a special centralized maintenance function. It is an extension of the delivery program of traditional telecommunications value-added services, and it provides the necessary technical support for the ecological environment of socialized value-added service innovation.

Key words: service innovation; cloud computing; SDPaaS

SDPaaS云平台是电信SDP和云计算PaaS的融合，是在传统SDP的技术基础上，结合云计算的技术特性，以服务为理念而实现的新的解决方案。它针对信息通信技术(ICT)网络融合的大环境，在支持第三方开发人员快速开发和交付业务的基础上，进一步实现了业务自动部署和托管运行、专业化集中式维护等功能，是对传统电信增值业务交付方案的一种拓展。

1 SDPaaS云平台系统架构

SDPaaS云平台构建的主旨是提供一种社会化的平台服务，使得专业的增值业务开发团队和普通开发者都能基于平台进行ICT融合服务的创新，并降低服务创新的难度及服务创新的成本。图1为SDPaaS云平台的总体架构，该方案采用松耦合、组件化的理念设计[1-5]。

SDPaaS云平台主要由开发者社区、应用开发测试环境、应用执行环境、资源汇聚网关、能力引擎、托管环境和管理平台七大部分组成，该平台可以为应用开发者(AP)提供应用开发、测试、发布、运维相关的资源和服务，同时支持资源提供者(RP)接入新的能力资源到SDPaaS平台中。SDPaaS云平台可以灵活部署，和基础设施即服务(IaaS)[6-10]没有必然的依赖关系。

(1) 开发者社区

开发者社区一方面可以让应用开发者学习如何开发应用、获取相关资源等；另一方面，应用开发者可以将自己开发出来的应用通过运营商的发布渠道进行销售，而开发者社区提供相应的自助服务功能。

开发者社区主要包括开发者注册、能力超市、资源下载、应用发布管理和论坛等功能。

(2) 应用开发测试环境

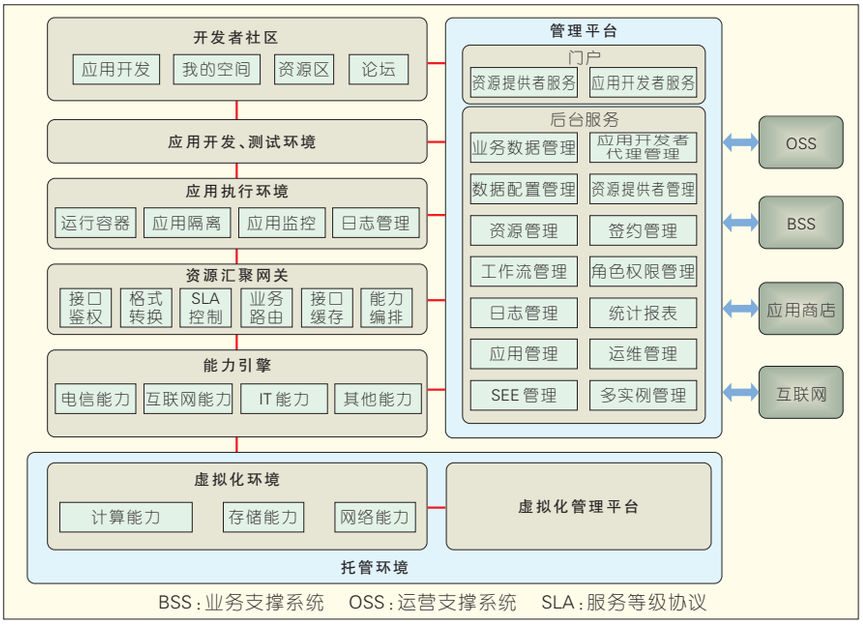
SDPaaS应用开发环境满足应用开发者简易快速完成应用逻辑生成、编译的功能要求，业务开发环境有离线业务开发环境和在线业务开发环境两种。从应用种类上讲，SDPaaS应用开发环境可以支持系统类的应用，也可以支持移动终端上的应用，例如Android应用。

- 在线业务开发环境

应用开发者可通过Web页面在线进行应用开发，无需安装客户端环境。在线业务开发环境主要用于以创意为主并且应用逻辑简单的应用开发，比如功能相对简单的融合应用或者基于内容聚合的移动应用。它主要面向对程序语言不太熟悉的开发者，并且对开发者的技能没有太高要求。

- 离线业务开发环境

应用开发者需要下载并安装客户端环境，对开发者有一定的硬件配



▲图1 SDPaaS 云平台系统架构图

置要求。对于应用逻辑复杂、功能要求复杂的应用，特别是需要开发者自行扩展较多组件的应用，可以使用离线的业务开发环境来开发。离线业务开发环境主要面向熟悉编程语言的专业开发者。

离线业务开发环境提供图形化的开发界面和软件开发工具包(SDK)，开发者可以根据自己的使用习惯选择使用图形化的开发界面，或者基于SDK进行应用开发。

应用测试环境(STE)提供业务模拟测试环境，应用开发者可以进行端到端的模拟测试，并可以查看应用模拟运行结果。

(3) 应用执行环境

应用执行环境负责提供应用运行所必需的容器或沙箱，它是应用运行时的环境，也是应用隔离的实现保证。第三方应用的执行或者二次编排的组合能力的执行，都需要应用执行环境进行控制和检测，并将应用的运行状态和资源使用情况等参数上报给管理平台。

(4) 资源汇聚网关

资源汇聚网关负责封装能力引擎所提供的能力服务接口，并以通用的服务接口方式暴露出来，它为应用开发者提供统一的开放接口。资源汇聚网关可汇聚电信能力、互联网能力，并提供开发工具供开发者编排出新的能力。

资源汇聚网关上汇聚的能力可以动态扩展。资源提供者可通过管理平台提交新的资源接口，并通过审核的资源将自动加载到资源汇聚网关上，这样便可实现动态加载，也不会影响现有能力资源和应用的运行。

资源汇聚网关是能力调用的入口，它可以在管理平台的配合下，由资源汇聚网关负责对能力调用的请求进行认证、鉴权和计费。

(5) 能力引擎

能力引擎负责平台所提供的能力的实现。对于传统的电信业务能力，能力引擎负责电信业务能力的适配，并通过与多种电信网元按照标准的电信协议进行交互，从而实现电信业务能力。对于互联网业务能力、IT能力和其他的业务能力，能力引擎的主要作用是接口适配。

(6) 管理平台

管理平台提供SDPaaS云的管理门户和后台管理服务，包括AP管理、RP管理、工作流管理、资源管理和签约管理等功能，实现了应用从提交、审核、部署到发布全生命周期的一系列管理。

管理平台还负责向开发者社区、应用开发测试环境、应用执行环境和资源汇聚网关同步数据，并且其开发出的应用可以通过管理平台发布到应用商店。

此外，管理平台需要和业务支撑系统(BSS)/运营支撑系统(OSS)、应用商店等外部网元进行交互，以便实现网管、计费、应用交易等功能。

(7) 托管环境

应用托管环境是应用部署和运行的物理环境。在应用测试完毕后，开发者可以根据应用运行的实际需要申请应用部署所需的资源和服务，包括需要部署的应用实例数量、应用实例的动态调整策略和应用的生命周期等。审批通过后，系统根据应用的类型和所需要的执行环境，自动打包并生成应用的虚拟机模板(即镜像文件)，并加载到虚拟机中并完成启动。通过这些简单方便的操作，就可以用最快的速度完成应用的部署并即刻开始提供服务。

2 SDPaaS 云平台关键技术及其解决方案

SDPaaS云平台是传统电信域能力开放与云计算技术的结合，主要涵盖了以下关键技术。

2.1 能力集成与开放

服务能力集成与开放是云平台服务需要解决的关键问题之一。SDPaaS平台需要对外提供包括电信能力、互联网能力和IT能力在内的多种能力服务接口，并实现参数封装与映射、接口鉴权、服务等级协议(SLA)控制、消息路由和资源有效性监控等功能，从技术实现上讲，需要重点考虑以下内容：

(1) 接口封装与适配

接口封装基于面向服务的体系

结构(SOA)架构，并采用互联网通用的接口服务形式，提供基于简单对象访问协议(SOAP)的Webservice和RESTful Webservice等多种形式，从实现语言的角度，支持java、C和Android等多种编程语言。

(2) 接口鉴权

应用在调用资源汇聚网关能力接口时，平台需要根据应用所携带的信息，如资源的应用标志(APPID)和开发者标志(APID)等对接口调用进行鉴权和认证，判断应用是否有权限来使用相应的能力资源。

(3) SLA控制

SLA控制是平台即服务的一个重要特征，它允许不同的应用申请不同水平的服务策略(能力许可、流量控制)。平台则会根据这些策略对应用的接入请求进行控制，并确保业务能力调用的服务质量。对于违反SLA中规定的请求予以拒绝，并返回对应的异常错误代码，并记录日志。

(4) 消息路由

注册到平台的能力接口，有从网络侧发起的上行消息，也有从应用侧发起的下行请求，平台需要根据该消息类型和接口规范来实现消息的分发和路由。同时，在存在多个能力引擎的情况下，还需要根据一定的路由策略实现下行消息的分发。

(5) 资源有效性监控

平台需要对能力资源的状态进行监控，当对能力引擎的调用出现连续异常，且次数超过系统配置值时，资源汇聚网关会向管理平台上报该资源异常。

(6) 能力编排

平台需要提供可视化的能力编排工具，该工具能够将多种下层网元提供的能力接口按照一定的逻辑进行编排，形成一种新的能力，进行接口封装后，可将新能力对外开放。

2.2 应用隔离

应用隔离是应用执行环境安全运行的关键，对于非信任域内的第三方应用，需确保某一个应用运行时故障不会引起整个系统出现故障。应用隔离的实现方案主要有3种：虚拟机隔离、容器隔离和应用实例隔离。

(1) 虚拟机隔离

虚拟机隔离功能，即一个应用独占一个虚拟机，每个虚拟机上安装独立的应用执行环境。应用部署时，系统根据应用类型从虚拟机资源池中获取对应的虚拟机及应用执行环境，并将应用安装在虚拟机上。

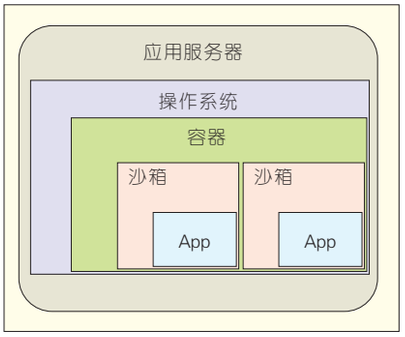
虚拟机部署的优点是应用隔离比较彻底，应用之间相互不影响，比较安全。但是对于移动互联网应用来说，大部分是小众应用，如果一个应用独占一个虚拟机，则会消耗大量的虚拟机资源，成本较高。因此虚拟机隔离适合于有大量并发用户访问的应用。

(2) 容器隔离

容器隔离是以应用运行的容器为单位来进行应用隔离的，一个应用独立部署在一个容器中，一个虚拟机或者一台物理机上可以部署多个容器。容器隔离方案的缺点是无法给某一容器分配或者限定资源，只能对容器所消耗的资源进行监控，一旦发现某一容器占用的资源达到指定阈值，就执行既定的策略(比如杀掉该容器进程)，但这样会中断业务，影响用户体验。

(3) 应用实例隔离

应用实例隔离是指在一个应用运行的容器中部署多个应用，每个应用的实例运行在一个独立的沙箱中，由沙箱对应用运行时的安全进行控制，如图2所示。



▲图2 应用实例隔离

应用实例隔离方案具有以下一些特点：

- 进程内隔离

虚拟机上安装一个或多个容器，其中一个容器就是一个独立的进程。应用部署在容器内的沙箱中，应用实例的运行安全则由沙箱去保证。这种隔离方式是进程内的隔离，占用系统资源最少。同时，采用沙箱技术也可以控制单个应用实例对资源的占用。

开放服务网关协议(OSGI)[8]是解决应用实例隔离的方案之一，其缺点是目前只能针对java应用。Google为了解决跨平台的浏览器插件隔离问题，还提出了NACL的解决方案，但是这些技术都有它们的局限性。

- 资源访问控制

在容器内部署多个相同或不同的应用实例，通过类加载白名单机制对应用可访问资源进行控制，应用不能访问系统内部的资源文件，只能访问其自身应用的资源文件和类文件。该方案还禁止应用对某些关键系统资源(如本地存储、线程、网络等)的访问，在屏蔽这些系统资源调用的同时，也必须为应用提供必须的替代功能，这也是GAE[10]、SAE等提供分布式存储、任务队列等计算和存储能力暴露接口的原因之一。

2.3 弹性伸缩与调度

SDPaaS云平台可以在两个层面实现弹性伸缩调度。

(1) 硬件基础设施层面。SDPaaS云平台通过与IaaS的整合，使系统具有IaaS所拥有的虚拟化计算与存储服务能力。SDPaaS与IaaS的整合需要虚拟化平台暴露相关的接口，主要包括创建虚拟机、启动虚拟机、关闭虚拟机、删除虚拟机和状体上报等5类接口。

(2) SDPaas层面。SDPaaS云平台可以在一个虚拟机上部署多个应用

实例，通过进程级别的监控来识别每个应用的CPU/内存使用情况，并根据每个应用的CPU/内存使用情况进行应用级别的动态伸缩。

2.4 负载均衡与域名转换

SDPaaS需要实现动态的应用请求二级域名转换和负载均衡，以便支持应用的动态加载、即时服务。SDPaaS在功能实现上可以基于Apache或者NGINX等开源软件。同时，考虑到应用请求接入协议多样，负载均衡设备需要支持四层交换、七层交换、GSLB等功能。

在应用部署时，通过接口动态更新负载均衡设备的配置策略，可以实现应用的二级域名地址到虚拟机IP地址的映射转换，并可以动态分发请求。当一个应用部署到多个虚拟机时，还可以起到负载均衡器的作用，如图3所示。

3 SDPaaS云平台的应用场景

SDPaaS云平台目前已经有3种商用的应用场景，分别是能力开放平台、企业应用数据中心和应用工厂，每种应用场景都有自己的侧重点。

(1) 能力开放平台

SDPaaS云平台接入各种电信能力、互联网能力及其它能力，并提供能力开放服务。电信能力包括短信能力、彩信能力、wappush能力、数据类能力、IP多媒体子系统(IMS)能力以及定位能力等。互联网能力主要有股票行情查询能力、航班查询能力等互联网上开放出来的能力。开发者也可以通过能力编排，创新出新的能力，并作为服务开放出去。

(2) 企业应用数据中心

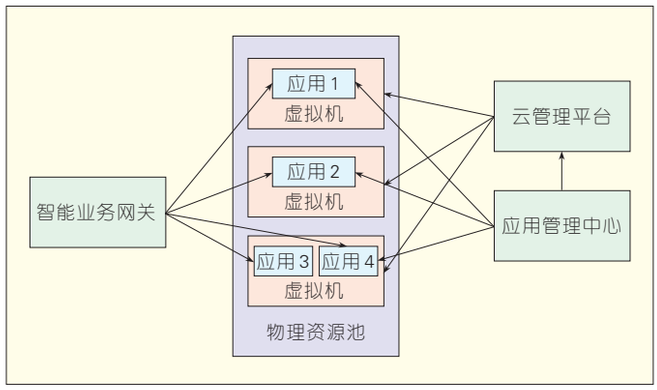
对于中小企业或专业开发者，开发出的服务端应用需要托管到互联网数据中心(IDC)机房，不仅费用高昂，而且技术要求高、调试运维难。SDPaaS云平台底层可以基于IaaS平台，而计算能力、存储能力和网络能力等可以由SDPaaS云平台统一监控、管理，自动具有弹性伸缩、按需分配和故障迁移等云计算特性。中小企业开或者专业开发者可以直接将自己的应用托管在SDPaaS云上，构建企业私有的应用数据中心(ADC)。

(3) 应用工厂

SDPaaS云平台提供了应用开发环境、应用执行环境、应用仿真测试环境，并为应用的开发、测试、部署提供完整的生命周期管理，开发者基于平台可以方便地开发、测试、部署自己的应用。开发出来的应用既可以托管在SDPaaS云平台，又可以快捷地放到运营商的应用商店里销售，从而形成“前店后厂”的模式。

4 结束语

作为促进电信和互联网“长尾”业务模式快速发展的重要手段之一，能力开放平台的发展得到了全球产业界的广泛关注。我们在长期增值业务平台技术实践的基础上，通过吸收、消化互联网和云计算技术，创造性地提出了SDPaaS这一具有云计算特性的融合业务开放平台并付诸研发实践。以SDPaaS平台提供了完整的端到端产品解决方案，是电信运营商向移动互联网转型的尝试，也是电信设备商从传统的卖产品向卖服务的战略转型的尝试。



◀图3
负载均衡的实现

5 参考文献

[1] 董振江,陆平,杨勇.具备云计算特性的业务交付平台及其关键技术研究[J].中兴通讯技术,2011,17 (5):55-57.
[2] 倪洪章.IBM SDP 实现电信运营商业务水平整合[N]. 计算机世界, 2008-02-04.
[3] 马苏安.云计算在电信领域的应用[J].中兴通讯技术,2010,16(4):44-47.
[4] 罗黎霞.基于云计算的服务平台--Google APP Engine[J].信息与电脑:理论版,2009(8):93-94.
[5] 刘鹏,云计算[M]. 2版. 北京:电子工业出版社,2011.
[6] 吴朱华. 云计算核心技术剖析[M]. 北京:人民邮电出版社,2011.
[7] 雷葆华,饶少阳,江峰等.计算解码:技术架构和产业运营[M]. 北京:电子工业出版社,2011.
[8] 林昊,曾宪杰.OSGI原理与最佳实践[M]. 北京:电子工业出版社,2009.
[9] MCGRATH M R.Understanding PaaS[M].New York,NY,USA:O'Reilly Media,2012.
[10] 唐学韬,何继业等. SANDERSON D. GAE编程指南[M].北京:机械工业出版社,2011.

收稿日期:2012-10-18

作者简介

屠要峰，南京航空航天大学硕士毕业；中兴通讯业务研究院副院长；长期从事电信增值业务产品的设计以及研发工作；研究方向为移动互联网、云计算、SDP及能力引擎等。

黄震江，郑州大学毕业；中兴通讯主任工程师；有通讯行业十余年从业经验，先后从事过增值业务产品、SDP产品系统设计、产品经理等相关工作。

陈心哲，武汉理工大学毕业；中兴通讯主任工程师；长期从事移动互联网技术和产品研发工作，主要研究方向为移动互联网平台、云计算平台等。

ByPass流量旁路技术组网实现探析

Networking with ByPass Flow Technology

中图分类号：TP393 文献标志码：A 文章编号：1009-6868 (2012) 06-0056-04

摘要：文章根据现网中的核心路由器“过境”流量负担重的问题，提出了流量旁路技术（ByPass）的IP层与光层联合组网策略。文章指出通过引入光层和IP层协同的机制，网络可以得到优化，并且光层OTN设备能够代替核心路由器来转发部分业务流量，这也是解决IP承载网所面临的扩展性问题的一个重要途径。

关键词：ByPass；流量旁路；光传送网(OTN)；IP层；光层

Abstract: To solve the problem of “crossing” traffic in the core network router, we suggest integrating IP layer and optical layer traffic using ByPass technology. By using this technology, the network can be optimized, and the optical transport network (OTN) equipment in the optical layer can forward some traffic instead of core router. This is an important way of solving scalability issues in the IP bearer network.

Key words: Key words: ByPass; traffic bypass; OTN; IP layer; optical layer

钟秀芳/ZHONG Xiufang
张沛/ZHANG Pei
（中国联合网络通信有限公司研究院，北京 100048）
(China Unicom Research Institute, Beijing 100048, China)

通过对运营商骨干网络流量的分析，人们发现在经过核心路由器的业务流量中，大约有50%以上属于“过境”的转发流量，而这些“过境”流量大大加重了核心路由器的负担。如果使用昂贵的路由器线卡处理这类流量，则会造成网络成本和功耗的快速增长。而利用光层和IP层的协同组网调度机制，可以在光层旁路IP层的“过境”流量。

文章针对IP层与光层网络资源协调技术，分析流量旁路技术(ByPass)的联合组网策略，并具体阐述了ByPass流量旁路技术的几种实现方式。文章指出不同的应用场景可选用不同的ByPass流量旁路技术实现方式，从而为不同的网络提供灵活的优化方案。

1 流量旁路技术概述

ByPass技术即指在运营商网络中采取的ByPass[1]组网技术，它可以将“过境”流量有效地通过光传送管道进行旁路，并利用光层大颗粒的调度、疏导和链路保护能力，降低核心路由器的处理压力，降低对骨干路由器的容量与复杂度要求，减少核心路由器的功耗，从而降低投资成本(CAPEX)和运营成本(OPEX)。这种IP层与光层之间的融合与统一调度将成网络演进的方向之一。

为了实现由光层分流IP层的ByPass流量，就要引入两层的协同机制。长期以来在对等模式下，IP层和光层相互隔离，因此两层协同机制大致分为3种实现方式：

- 以光层为主的实现方式
- 以IP层为主的实现方式
- 其他实现方式

以光层为主的实现方式是在光传送层进行分组化传送的方式[2]，该实现方式由分组化OTN来识别IP层的标签（例如：标签交换路由器的多协议标签交换标签(LSR MPLS Label)、虚拟局域网(VLAN)等），从而旁路部分IP层核心路由器的转发流量。这种方式对现有IP层影响较小，在一定程度上依赖于IP层的协议。

以IP层为主的实现方式是将部分光层的特性转移到路由器端口上，这样以来由路由器端口发出的报文会带有“光层的传输标记”（例如：波长、集光纤配线单元(ODU)等）并在光层直接进行交叉。这种方式实际上是由IP层替代了部分光传送层的一些功能，IP层需要预先获取光层的网络拓扑。

另外，一些两层协同的组网方式依赖于控制平面，或者是引进了一些新的设备形态。例如：协同的流量管理、路径集中计算(PCE)、多通道负载均衡(MC-LB)、通用多协议标志交换协议用户节点接口(GMPLS-UNI)方式、基于子接口的多层网络优化、ByPass服务器等等。

在IP层与光层联合组网中引入Bypass机制后，整网可得到如下的一些优化：

- 减轻了核心路由器的转发流量和转发压力。
- 减少了核心路由器的端口，或是降低了核心路由器的端口速率，节约了端口扩容成本。

• 减少了业务流量的转发跳数，有助于提升业务指标、保证业务承载质量。

2 ByPass联合组网策略分析

目前各大运营商均采用两层组网联合组网方案，既上层为IP数据网，下层为光传输网，两层网络之间通过特定光接口进行连接。光接口根据接口类型可以分为彩光接口和白光接口；根据速率可以分为2.5G、10G和40G，未来可能会出现100G速率的需求；根据业务可以分为以太网接口、光传送网(OTN)接口和SONET/SDH上的包传输(POS)接口。

根据运营商的实际需求，现阶段IP层数据网和光层传输网主要可以分为两种组网模型：一种为传统白光组网，另外一种则为彩光组网模型。其中白光组网是目前运营商普遍采用的方案，路由器厂商和波分传输厂商所推出的大部分设备也均是支持白光接口；而彩光组网则代表着未来网络扁平化发展的一种趋势。路由器设备集成光转化单元(OTU)并提供彩光接口，该接口直接与波分设备的彩光口相连，从而节省了组网过程中OTU的配置。目前业内Juniper和思科已推出了包含彩光接口的路由器设备，而德天翔则推出了包含彩光接口的OTN光交叉设备。

2.1 白光接口组网模型

白光接口组网模型，实质上是一种客户层和服务层关系。当OTN作为IP网络的服务层，IP网络作为OTN的客户信号，两者则构成客户—服务关系，并通过路由器提供白光接口支持10GE/40GE/100GE over OTN进行组网。OTN主要进行大颗粒的业务调度、业务在物理层的快速开通以及线路侧故障的保护，以提高整个传送网的链路资源的利用率和增强传送网的生存性，是目前IP网络和OTN网络互存的主要形态。

如图1所示，在白光组网中，路由器直接提供白光信号。该白光信号通过光纤直接连接到OTN设备的OTU客户侧接口中，然后通过OTN设备的OTU单板，将输入的白光信号调制到C波段某个波长上，并最终送入OTN设备的合波器中，和其他波长信号捆绑，在线路侧接口上传送。

在白光组网中，OTN网络的运维和管理主要在光层OTN设备上通过段监控(SM)开销、通道监控(PM)开销和串行连接监控(TCM)开销而实现。路由器网管对路由器所提供的白光信号除了有功率检测等物理信号监测功能之外，不具备其他任何开销、运维、管理功能。

在白光组网模式中，也分为3种接口类型，分别为：POS接口、OTN接口和以太网接口。其中POS接口是目前中国联通现网上普遍采用的接口类型，在路由器和OTN设备对接中可以实现基于同步数字体系(SDH)开销的监控功能，但其成本是最高的。

以太网接口和OTN接口是近几年来新推出的接口类型，相对于POS接口而言，成本较低。以太网接口标准IEEE 802.3工作组所制订，从速率角度分为GE、10GE、40GE，未来会出现100GE需求。

2.2 彩光接口组网模型

彩光接口组网模型是将路由器和OTN设备通过彩光接口对接，并将路由器内置OTU功能模块；它所提供的彩光信号已经被调制到C波段某一个波长上，并直接连接到OTN设备的波分侧，从而实现网络扁平化发展的要求。目前，已有思科、Juniper等公司已经开发了彩光接口路由器。但对于传输厂商，仅有个别厂商开发了具有彩光接口的OTN设备。

如图2所示，在彩光接口组网中，将OTU的转发器内置到路由器中，使得路由器直接提供彩光接口；输出的彩光信号已经被调制到某一特定波长上，直接与OTN设备的线路侧接口连接，然后与其他路由器输出的彩光信号一起输入到合波器中，合为一路信号在线路侧进行传输。

彩光接口遵从ITU-T G.709标准。通过彩光接口互联，网络可以节省路由器与密集波分复用(WDM)/OTN传输系统之间的白光接口。WDM传输系统不需要OTU单元进行波长转换，从而简化了网络架构，方便故障定位，降低管理难度并降低CAPEX和OPEX。

对于彩光接口组网模式，目前仍然存在一些问题：

• 网管界面分工不明确，传输厂家网管系统需要根据不同彩光路由器厂家设备进行定制化开发。

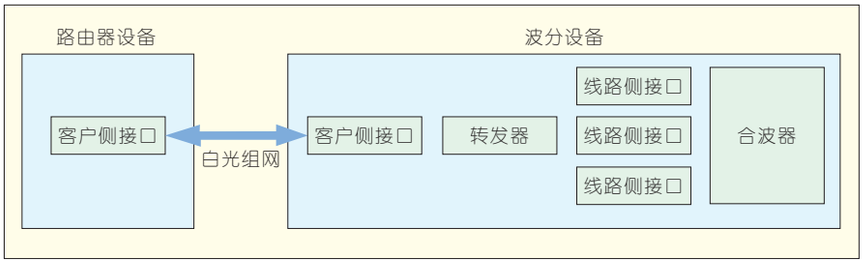
• OTN开销、维护及管理由路由器设备负责，路由器厂家对于SM/PM/TCM段开销管理的理解与传输厂家存在差异。

• 彩光接口组网与现有网络结构存在较大差异。

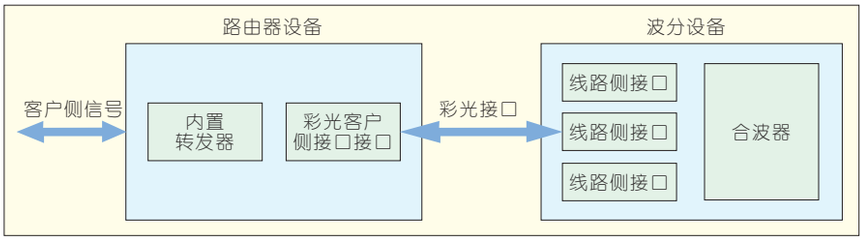
3 ByPass流量旁路组网实现方式

3.1 光层分组化OTN实现方式

在光层分组化OTN实现方式中，



▲图1 白光组网模型



▲图2 彩光接口组网模型

光传送层引入分组化OTN设备，其核心交换矩阵除了支持ODU、VC以外，还支持MPLS-TP协议[3]，接口也支持MPLS-TP功能。

如图3所示，业务流在PE1中被标上不同的标签；在P-OTN4设备上根据不同的标签，业务1被下载到客户侧UNI至核心路由器4，业务2和业务3被分别交叉到OTN2和OTN3；然后业务流量1在P4继续转发，业务2和业务3分别到PE2和PE3之后，根据需求剥去标签，然后下路。

由分组化OTN来识别IP层的标签，有的方案是针对LSR MPLS标签进行识别，还有些是针对VLAN进行识别。利用光层分组化OTN这种方式，可以旁路部分IP层核心路由器的转发流量，对现有IP层影响较小，在一定程度上依赖于IP层的协议。

光层分组化OTN的实现方式适用于在光层有条件将传统OTN设备转换为分组化OTN设备或新引入分组化OTN设备的场景，无需改动IP层的路由器设备。这样既疏导了过境流量，也对原设备做了利旧。

3.2 路由器端口实现方式

路由器端口实现方式，是将光层的波分特性转移到路由器接口上。如图4中所示，与分组化OTN方式比较，业务流在彩光口路由器P1被分配了不同的波长，也可以配置在不同的ODU中。在IP层感知光层拓扑的前提下，IP层把光层作为刚性管道，由不同的波长来决定转发的途经。OTN4设备可以看作是一个基于光交叉的可重构型分插复用器(ROADM)设备。根据不同的波长，业务1被交叉到OTN1设备中，业务2和业务3则被交叉到OTN2和OTN3设备中，然后业务流量1在P4继续转发，业务2和业务3到P2和P中转发。

在路由器端口实现方式中，路由器端口支持彩光特性[4]（或是支持ODU的封装），发送的报文带有某种“光层的传输标记”（例如：波长、ODU等），并将在光层直接进行交叉。这种方式实际上是由IP层替代了部分光传送层的功能，因此IP层要预先获取光层的网络拓扑。

路由器端口实现方式适用于光层充分利旧，在IP层有条件将传统路由器设备转换为彩光接口路由器设备或新引入彩光接口路由器设备的场景，无需改动光层设备的组网。由于彩光接口的路由器仅有少数几个厂商实现，故改造成本将增加。

3.3 其他实现方式

3.3.1 协同的流量管理

IP/OTN协同流量管理（多层流量

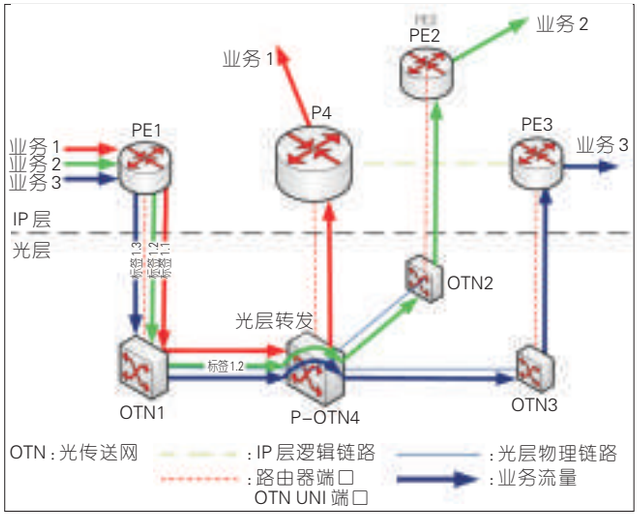


图3▶ 光层分组化OTN组网实现方式

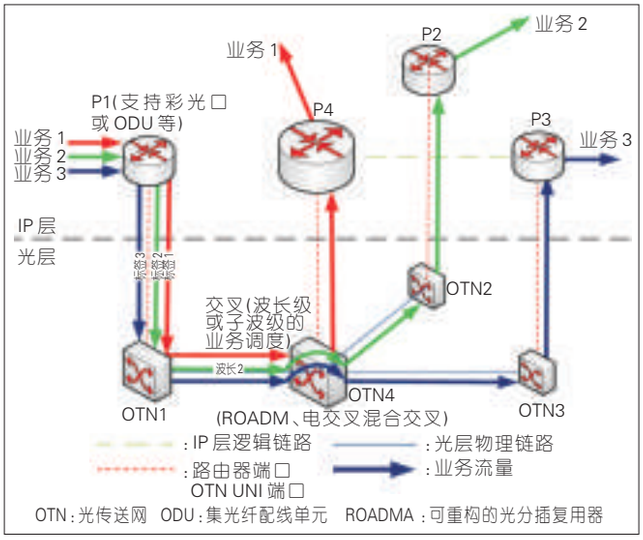


图4▶ 路由器端口实现方式

工程）在提升网络性能的同时，还可降低网络的扩容压力。任意两台路由器之间的流量如果超过事先预设的阈值，路由器就可以通过UNI接口向OTN网提出带宽请求，传送网络在接到路由器的带宽请求之后，通过波长路由算法，在两台路由器之间快速搭建一条光层直达路由。这时，路由器的容量不需要增加，因为达到阈值的流量通过OTN层直达了。路由器IP端口的成本一般是OTN端口的4～5倍。由于光层智能分流了路由器业务，减少了路由跳数，从而减轻了路由器转发压力，减少了骨干IP网络中昂贵的IP端口（路由器高速线卡）的投资，从而可显著降低网络的CAPEX。例如，欧洲某主流运营商正是通过这种IP/OTN双层协同的方式，通过光层自动旁路路由器的流量，使网络的CAPEX节省了40%以上。

协同流量管理的方式适用于路由器与OTN组网的不改变原设备架构的协同管理，能适当减轻过境流量的压力，但在旁路效率上不如上两种方式。

3.3.2 MC-LB方案

MC-LB方案[5]有两个技术要点：一个方向的数据流量可以从多个物理端口转发；一个物理端口可以转发多个方向的数据流量，如图5所示。

尽管MC-LB方案充分发挥了IP设备与光设备的协同优势，但却并不依赖于GMPLS等实现较为复杂的协议，同时，在GMPLS成熟应用后，也将会利用其优势，增强灵活性和易操作性。

3.3.3 路径集中计算单元

Internet工程任务组(IETF) PCE工作组定义的基于PCE的MPLS/GMPLS网络结构[6]使得路径计算功能从网络设备中独立出来成为可能。通过在网络中部署独立的路径计算单元可以解决MPLS/GMPLS网络中大量链路基于约束的路径计算所需的特别资源问题。

在多层的网络结构中，可能上层是IP/MPLS网络，底层是GMPLS控制的光网络，底层网络的流量工程LSP为上层网络构成了一个虚拟的网络拓扑结构(VNT)。这种情况下的流量工程路径计算可以主要分为下面两种模式：

单PCE多层路径计算。这种模式下的PCE称为多层PCE，这种PCE收集各层网络的拓扑信息和流量工程信息，因而可以单独计算跨网络层的流量工程路径。

多PCE的多层路径计算中，每层网络都有至少一个PCE，并且每层网络的PCE相互协调，来计算光层的流量工程路径。

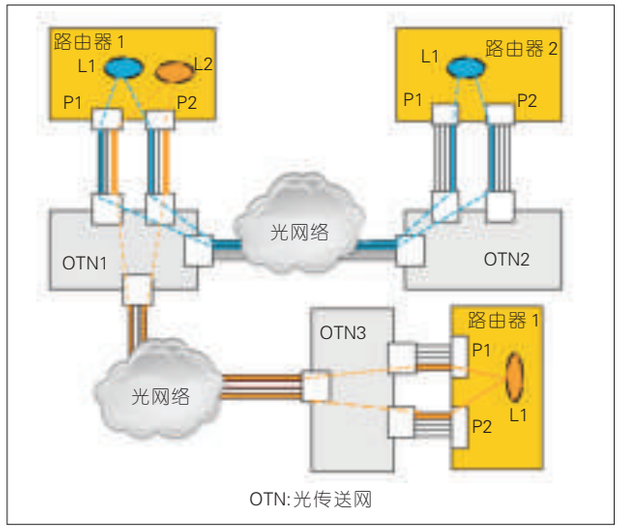
4 结束语

综上所述，在核心路由器上，对于一部分流量的转发，是可以通过光层来直接完成的。通过引入光层和IP层协同的机制，在IP层和OTN的双层协同的基础上进行网络优化，由光层OTN设备代替核心路由器来转发部分业务流量，是解决IP承载网所面临的扩展性问题的一个途径。这些实现方式能够有效地配置网路资源，有效分流核心路由器上的流量，扩大网络的容量，缓解带宽扩容的压力，在不同程度上对应于IP承载网络从重叠模型向对等模型演进的趋势。

5 参考文献

[1] 郑寿涛.运营商互联网流控系统部署分析[J].电信网技术,2011(9):59-63.
[2] 支持多业务的光传送网（OTN）设备技术要求[S].
[3] 传送多协议标记交换MPLS-TP（MPLS Transport Profile）[J]. 通信技术与标准,2011(3).
[4] 资深通信人.彩光口技术的理解[EB/OL].(2010-07-15). http://lsh5768.blog.163.com/blog/static/706630622010615104345869
[5] 颜清华,柏璐.IP路由与光传送的协同组网[J].网络电信,2011(4).
[6] 易小波,孙秀清,唐元春等.基于路径计算单元的MPLS/GMPLS网络结构[J]. 电信网技术,2008(2): 41-47.

收稿日期：2012-08-07



◀图5 MC-LB方案的组网模型

作者简介

钟秀芳，北京邮电大学硕士毕业；现任联通研究院网络技术研究中心中级工程师，硕士；从事电信网络技术接入方面的研究工作；发表论文近10篇。

张沛，北京邮电大学光通信中心博士毕业；现工作于中国联通研究院网络技术中心；研究方向是高速光传输系统、承载网关键技术研究；向ITU-T提交国际文稿二十余篇，发表文章十余篇，专利3项，著作3本。

可信计算 3

姚文斌/YAO Wenbin
(北京邮电大学计算机学院 可信分布式计算与服务教育部重点实验室，北京 100876)

[编者按]随着信息技术的发展和网络应用拓展，可信计算已经成为当前学术界和工业界的研究热点。本讲座将分3期对可信计算的内涵、关键技术和未来进行讨论：第1讲从可信计算的演进过程简单叙述可信计算技术发展历程、可信计算的定义和范畴、相关研究领域；第2讲从技术的角度探讨可信计算研究的系统架构、主要关键技术及其当前研究的热点问题；第3讲介绍可信计算的应用模式和未来。

中图分类号：TN91 文献标志码：A 文章编号：1009-6868 (2012) 06-0060-04

7 可信计算未来研究方向

尽管可信计算的研究已经有很长的时间，但是其未来还有很大的发展空间。在可预见的未来，可信计算将主要围绕着可信性基础理论研究、应用实现平台、监控与评价体系和管理模型等方向展开研究。

7.1 高可信理论研究

在学术领域，国际上普遍承认可信计算研究目前处于技术实践超前于理论研究阶段[6]。这并不代表可信计算研究学者不重视可信计算的理论，而是由于可信计算研究领域涉及的范围宽、交叉学科本身缺乏理论突破而导致的。例如，从社会学的角度出发，“什么样的人可以被信任”、“信任别人的程度如何用量化指标加以度量”等问题，在人类社会同样需要采用多角度加以度量而难以用“是”或“否”的简单断言加以回答。那么，对于使用计算机为工具进行计算的系统管理员来说，如何通过计算行为来度量用户的可信性呢？尤其困难的是如果系统管理员也不一定可信而又需要成为信任评价对象时，这一判断将更为困难。这也触及到人工智能的基础性问题。如果人类认知学方面没有探究清楚，那么计算机作为一种被动工具又如何才能具有类似于人类的思维突变特性，并以此来解决这些信任问题呢？

因此，受上述条件束缚，在可预见的未来，高可信计算理论研究仍将围绕着以计算为核心的相关研究领域展开，具体内容包括：

- 可信故障/攻击模型研究

从可信发展历史看，可信计算针对的故障/攻击目标来自于计算机系统的硬件故障(包括硬件器件老化、硬件设计错误等)、软件故障(包括软件老化故障、设计故障等)、人为恶意故障(也称为Byzantine故障)、系统环境故障(包括人为操作故障、天灾等不可抗力给信息系统造成的打击故障等)等多种形式和形态，其研究包括了当前已知的全部信息系统故障。这也决定了对于信息系统的可信性故障/攻击模型的复杂性以及困难性。

在实践过程中，当前的研究往往局限于故障的某一具体模式。例如，容错领域往往关注的是系统硬件故障、软件故障、环境故障等的综合，重点研究系统的可靠性、可用性和鲁棒性等属性，对于信息系统可信性要求的重点在于保证系统的连续可用和任务的连续可完成性，乃至于系统的容灾抗毁能力和可生存能力；安全领域则往往关注于系统的人为恶意故障，重点研究系统的安全性、信息的防泄漏性等属性，对于信息系统可信性要求的重点在于保证系统的信息机密性、防止非授权用户的恶意欺骗，以及信息的安全可控。显然，当前往往缺乏对信息系统故障/攻击模型的总体性故障模型、系统性故障机理与传播模式等进行的深入研究。

- 可信计算模型研究

当前，对于可信计算模型的研究还很缺乏。学者们对于可信计算涵盖的主要领域(容错计算领域或者安全领域或者管理领域)进行了大量的工作。无论是在针对信息系统容错能力方面的可靠性设计、可用性设计、可测试性设计等设计技术，还是在安全领域的基于密码学的安全保障技术，乃至管理领域的人员、过程等的全方位管理和控制技术，人们都积累了大量的成熟经验和技术。但

是如何将相关关联而交叉的技术综合并加以利用，到目前为止还缺乏系统性研究和富有指导意义的成果。

目前，学术界还没有国际上一致认可的可信计算模型。可信计算组织提出的可信计算模型采用的是信任链模型，即由可信测量根(RTM)→BIOS→OSLoader→OS 构成了一个串行链，并通过信任在信任链上的传递以实现系统的可信。然而，该模型是否通用于所有情况以及模型本身的可扩展性还存在很多探讨空间。

● 可信性度量理论

当前，对于可信性度量理论的研究还处于起步阶段，还有大量的问题亟待解决。

从理论上说，可信性度量应当度量的是信任链中的可信性。但是由于可信性的定义、范畴和属性还存在一些争议，另外可信性的指标体系并未建立，因此，量化的可信性指标难以直接度量。在实践上，可信计算组织则通过在信任链中度量数据完整性来间接评价可信性(在对数据完整性度量时通过校验数据哈希值的方法来实现)。正如前文所述，完整性指标作为可信性中一个侧面，是不足以量化可信性指标的。显然，目前可信计算组织采用的信任链测量具有较大的局限性。另外，可信计算组织为简化模型，将信任值二值化(只考虑可信和不可信两种极端状况)，而且认为在传递过程中没有信任损失，这显然是一种理想化的处理方法。

目前对于可信度量理论的研究主要包括软件可信属性研究与度量、软件可信性的动态演化与软件可信性等级和信任传递模型等。

● 可信软件开发方法

随着计算机硬件系统的不断完善和发展，软件系统的可信性问题却表现得越来越突出。软件功能不完善、软件漏洞频出、软件鲁棒性缺失、软件生存能力不强、软件无法抵抗攻击、软件系统老化等一直是软件系统开发者所面对的棘手问题。为此，改进软件开发过程以控制软件质量，利用相异性等以抵抗设计故障，提高自适应能力以应对运行过程中的异常，增加自检测机制以诊断和屏蔽故障影响等关键技术的研究不断引入到软件开发之中，其研究内容有进一步扩展的趋势。

● 网络行为检测和诊断

可信计算平台不是孤立存在的，计算机网络的飞速发展使得可信网络成为可信计算的关键问题之一。如何对请求接入网络的计算终端进行完整性验证和可信性评估，以杜绝不安全的计算终端接入网络；如何对接入网络的计算终端实施访问控制，以评估计算终端网络行为的可信性等，都是有待解决的关键问题。同时，当前可信计算模型更多是探讨信息系统的静态结构、动态行为和表现，对于软件内容/软件语义方面的可信性研究较少，而在未来网络内容可信、网络舆情可信性都将是极富特色的研究方向。

7.2 高可信计算平台研究

高可信计算平台的研究主要集中在可信计算平台的体系结构、安全芯片、可信密码算法与可信链实现等方面，具体包括：

(1) 可信计算的体系结构。该结构主要包括可信计算平台的硬件结构和软件结构，诸如个人计算机、服务器、移动终端的硬、软件体系结构的可信性增强相关技术实现。

(2) 可信芯片及其相关软件的设计与实现。这主要包括可信计算平台模块(TPM)硬件结构、可信密码模块(TCM)芯片设计、芯片可测试技术、嵌入式软件等。

(3) 可信密码应用技术。该技术主要改进适用于安全芯片的非对称密码算法、对称密码算法、安全散列算法以及随机数生成算法等的芯片级应用。

(4) 可信链应用技术。该技术主要包括以安全芯片为基础的信任传递、可信计算平台的完整性度量、安全存储、可信计算平台状态报告等相关技术。

(5) 可信基础软件的设计。该设计主要包括可信操作系统、可信编译、可信数据库和可信应用软件等。

(6) 可信网络。该网络主要包括可信网络体系结构、可信网络协议、可信网络设备和网络环境下远程证明等。

(7) 可信移动代码。移动代码是指能在发起平台上生成，并在其他一台或多台主机上运行的可执行代码，主要研究包括移动代码可信性模型、移动代码的控制与协商机制、移动代码的认证机制等。

(8) 特殊/典型环境中的可信计算。这主要指移动环境中的可信计算、嵌入式系统中的可信计算、云计算环境中的可信计算和物联网环境中的可信计算。这 4 种典型环境中的可信计算将代表着未来的应用主流方向。

● 移动环境中的可信计算

这里描述的移动环境是泛指一切可以随时接入和退出的动态网络系统架构，这种动态网络可以是 Ad hoc 网络、传感器网络、移动网络、手机、掌上电脑等任何具有动态网络结构的网络系统。由于移动互联网往往具有连接移动性、应用局限性、应用关联性和业务私密性等特点，因此，对于这些网络来说，如何基于可信性方法来感知、评判和使用节点(可以计算节点、存储节点、网络节点等)以确保计算的真实性、准确性、可用性和可检验性，是当前研究的重要问题。

● 嵌入式系统中的可信计算

嵌入式计算机系统是指以应用为中心、以计算机技术为基础、软硬件可裁剪、适应应用环境对功能、实时性、可靠性、成本、体积、功耗等严格约束的专用计算机系统。其基本特征是以应用为中心，系统的功能、性能、组成和外观都按照应用需求而

制订，具有运行环境千差万别、硬件平台丰富多样、软硬件高度融合、运行稳定可靠、系统资源有限等特点，这些特点决定了嵌入式计算机系统的可信性更加侧重于“可信”的基本属性方面，其研究内容呈现出复杂化、应用相关性、内容分散而专有化的特点。

• 云计算环境中的可信计算

云计算是网格计算、分布式计算、并行计算、效用计算、网络存储、虚拟化、负载均衡等传统计算机和网络技术发展融合的产物。从商业角度看，云计算意义在于将信息的系统从传统的设备销售模式为主转变为以提供信息系统服务为主；从技术角度看，云计算对大型复杂信息系统的运营和维护提出了更高的要求，这将极大地促进复杂信息系统管理和信息安全防护相关技术的发展；从实现角度看，云计算要求整合多样、异构的资源(包括计算资源、存储资源、传输资源和信息资源等)以提供高效、便捷的按需服务，这将促进相关服务整合、进一步提高服务质量。

当前云计算发展最主要的问题在于复杂性管理和系统可信性，前者是指对于海量异构资源(包括计算资源、存储资源和传输资源三大类)的合理共享和分配；后者则要求对于云计算系统和云提供的服务的可信赖性。显然，云计算中的可信服务涉及到访问和管理权限、隐私的共享和保护、数据的调度和分发、数据的容灾与抗毁，以及第三方审计、法律法规等各个方面的系统化和规范化问题。

• 物联网环境下的可信计算

物联网(IoT)就是物物相连的互联网，是通过射频识别、红外感应器、全球定位系统、激光扫描器等信息传感设备，按约定协议把任何物品与互联网相连接，并进行信息交换和通信，以实现对物品的智能化识别、定位、跟踪、监控和管理的一种网络。物联网概念的推出，可以说将世界上所有的信息设备互连以实现全球意义上的信息共享和计算。显然，从海量复杂信息中挖掘出有用、可信的信息成为物联网环境中的一个基础性且关键问题。目前，对于该方向研究尚待展开。

7.3 高可信系统监控与评价研究

正如同判断一个人是否可信需要观察其一段时间的行为积累一样，一个系统是否可信同样需要一段时间的观察和判断。这样的操作过程可以表述为对于高可信信息系统的监控和评价。

由于大规模分布式系统是一个动态演化的系统，具有分布范围广、节点状态不稳定和网络强动态等特点，动态演化过程中可能会出现异常，这将导致系统出错，并会降低系统的可信性。信息系统的可信性主要表现在其行为可信上，即运行行为可监测、行为结果可评估、异常行为可控制。行为监测是可信评估的基础，为可信评估提供原始数据。系统行为的可信来源于对系统行为全面、准确、实时的监测，对可信性相关数据的收集及加综合判断等。由于持续可用性一直是容错计算领域关注的重点问题，因此采用持续可用性行为为核心，结合其他行为属性来判断信息系统的可信性常常被当作一种简化的模型方法。

随着以 Internet 为基础平台的、各种大规模的分布式应用(如 P2P、网格计算和普适计算等)的深入研究，系统表现为由多个软件服务组成的动态协作模型。在这种动态和不确定的环境下，传统的容错机制中的故障检测已经不能够满足这些新型应用环境的可信需求。针对这些新型应用环境的动态信任管理技术已成为一个研究热点。大规模动态可信监控是在原有网络安全、容错技术的基础上增加行为可信的分布式系统监控的新方法。该方法强化了对网络状态的动态处理，为实施智能自适应的网络安全和服务质量控制提供策略基础。显然，针对复杂开放环境下网络节点行为的多样性、动态性和不确定性，尤其是系统中各种资源监控对象的行为数据呈现高维度的特征，大规模分布式系统的可信监控研究在未来具有广阔的应用前景。

7.4 高可信管理模型研究

有学者曾经说过“安全问题中有80%是人的问题”。对于信息系统的可信性来说，管理问题同样至关重要。然而，人们当前对于信息系统可信性管理的研究还很薄弱。通常人们认为：只需将现有的可靠性与安全的条例与制度、相应的管理模式与规章照样拷贝过来即可。然而，事实上，高可信管理模型远非如此简单。

虽然从范畴上讲，可信≈可靠+安全，但是从管理学角度出发，可信计算的管理模式远非可靠性管理与安全管理技术的简单叠加，两者之间还存在包容和互斥的内容。例如，从可靠性设计的角度出发，研制高可信的容错软件最有效的方法是采用人员的相异性设计以避免产生共因故障，然而这种技术必然导致相异性人员(人员数量、权限等级、任务等级等)、工具、管理手段等复杂性激增，会给安全性管理带来很大困难。再比如，安全管理介入到软件开发过程中，同样会给软件开发进度管理和质量管理带来很大的负担。

在目前，全球对于可信性管理方面的研究很少，尚处于起步阶段。

8 可信计算未来应用模式

在国际上，虽然可信计算研究领域处于技术实践超前于理论研究，但是这并不意味着可信计算已经被普遍接受并应用于计算领域。恰恰相反，即便在学术界中，对于可信计算的研究也还存在广泛的争议。同时，在当今随着计算模式正从传统的集中式、面向科学计算为主转变为以分布式、面向应用计算为主，可信计算的应用模式问题也随着计算模式的

发展不断进行演化。在可预见的未来，我们认为可信计算将会在基础运行平台和基础服务平台两种主要应用模式方面有进一步的快速发展。

8.1 高可信基础运行平台

目前，信息化系统已经表现出海量化(资源规模庞大)、异构化(软件/硬件/数据异构)、复杂化(不同形态的计算节点构成的复杂巨系统)、大众化(不同背景和文化层次以及不同动机的人们共同参与)等鲜明特征，形成了一个人人参与的集团化、社会化的网络系统。这就如同人类社会需要法律机制作为维护社会稳定的准绳一样，网络系统中需要有可信性来划分可靠和安全的群体，以便将危险和威胁群体加以排除。

高可信基础运行平台，即为用户提供一个可以信赖的基础计算平台，使得用户可以随时、随地、随意地使用计算平台上的各种资源以实现所希望的业务。基于云计算环境下的可信基础平台将是目前最为现实可行的解决方案。

8.2 高可信基础服务平台

进入21世纪以后，随着基于互联网的现代服务业和网络新媒体的兴起，改变了社会政治经济生活方式。同时未来信息系统的核心不仅仅是实现网络通信和资源共享，更是将转变成服务网络。以泛在网络为依托的服务计算跨越多个领域，具有高动态性和不确定性，网上用户行为的合法性以及网上服务系统用户身份的真实性成为所有网上服务参与者关注的焦点。如何为服务双方建立一个可信赖的服务计算环境，提供人与人、人与物、物与物之间按需进行的信息获取、传递、存储、认知、决策、使用等各项服务，已成为信息科学技术领域最重要的课题之一。

服务计算是跨越计算机与信息技术、商业管理、商业质询等领域的一个新的学科，主要用来解决技术平台架构的问题、服务交付的问题、服务本身的整合和管理问题等。可信服务可以有多种形式，例如可信计算资源服务、可信存储资源服务、可信网络资源服务、可信数据资源服务、可信软件资源服务、可信舆情分析服务、可信决策分析服务等。

服务的可信性取决于网络环境的可信性和授权用户的可信性，这要求网络和用户的行为及其结果总是可预期与可管理的，并做到行为状态可监测、行为结果可评估、异常行为可管理。其研究难点在于网络终端行为预期判断评估的模式和信息内容的可信监测。

随着信息化技术的逐步扩展，可信服务很有可能成为具有最为广泛的应用模式、并成为最有希望的发展模式。

9 结束语

从业界第一次提出可信计算的概念至今已经有30多年的时间，可信计算的概念和范畴也随着时间的推移不断进行着演进和发展。尤其是近十年间，学术界和产业界对于可信计算都投入了极大的关注。中国对于高可信软件系统也高度重视：2006年发布的《国家中长期科学和技术发展规划纲要(2006－2020年)》中要求：核心电子器件、高端通用芯片及基础软件产品国家科技重大专项(简称“核高基”)是电子信息产品和国防电子装备的核心，是信息产业核心竞争力最重要的体现，强调必须自主研制高可信的网络化操作系统；自2007年起，国家自然科学基金委员会设立了“可信软件基础研究”重大研究计划，专项支持对于高可信软件基础研究工作，到目前已经取得了若干研究成果。

然而，正如前文所述，更多的研究都还集中于零散的、分散的研究之中，甚至是很多现成研究成果的简单组合和拼接，缺乏整体性、系统性的突破。这主要是由于可信计算研究不仅仅涉及到计算机技术、网络技术、安全技术等计算机相关技术，它还直接取材于管理学、社会学、哲学、心理学等一系列交叉学科的研究成果的综合运用，并且随着计算系统更加深入我们的社会生活。然而，正如“信任体系”必将作为和谐社会的基石一样，“以人为本”的“社会计算”正以前所未有的速度发展和推进，可信计算也必将得到更为广泛的关注和研究。 **(续完)**

10 参考文献

[6] 张焕国，何炎祥，赵波，彭国军等. 我国可信计算研究与发展. 中国计算机学会通讯[J]. 2010, 6(2): 10–19.

收稿日期：2012–08–17

作者简介

姚文斌，北京邮电大学计算机学院教授、博士生导师，主要研究方向为可信计算、信息安全、灾备技术等；已发表论文60余篇、申请国家发明专利20项、制订国家标准1项、通信行业标准3项。

综合信息

中兴通讯勇夺中国电信模块化UPS集采第一

【本刊讯】2012年10月19日消息，中国电信2012年中小容量模块化UPS集采项目结果公布，中兴通讯模块化UPS产品获得综合排名第一，总份额第一的成绩。

《中兴通讯技术》第18卷总目次