



专题:云计算与语音业务

媒体资源云化关键技术研究

李 青,唐哲红,宋阿芳,王发光

(中国电信股份有限公司上海研究院 上海 200122)

摘 要:提出一种媒体资源池云化中的多租户资源调度、负载均衡实现技术及流程,介绍了云媒体在 IMS 业务平台中的应用方案、云媒体资源技术的实验情况,并对云媒体商业模式等进行了阐述,为推动本土软硬件云媒体服务器生产商及相关产业链发展提供了新思路。

关键词:云计算;媒体服务器;业务平台云化;负载均衡

doi: 10.3969/j.issn.1000-0801.2013.06.007

System Design for Media Server Based on Cloud Computing Technology and the Application Analysis

Li Qing, Tang Zhehong, Song Afang, Wang Faguang

(Shanghai Research Institute of China Telecom Co., Ltd., Shanghai 200122, China)

Abstract: A media resource pool of multi-tenant cloud-based resource scheduling, load balancing implementation techniques and processes were proposed. Cloud media in the IMS service platform application solutions, cloud technology, media resources test situation were introduced, and cloud media business model was presented. In order to promote cloud media server hardware and software for the local manufacturers and related industry chain development provides a new idea.

Key words: cloud computing, media server, cloud based service platform, load balancing

1 背景

我国信息技术资源分布不均匀,市场需求复杂多样,发展云计算对提升我国信息化水平、培育信息消费等新经济增长点有着积极的促进作用,下一步需要加强核心技术攻关,健全完善云计算技术标准体系,并以此在未来云计算产业发展中谋取话语权。

中国电信股份有限公司(以下简称“中国电信”)作为我国互联网基础接入服务的主导者,在 IP 网络关键资源、IT 技术架构,甚至上下游供应链的关系等方面拥有更多优势,为云计算技术创新提供了难得的机遇和巨大的发展空间。同时,中国电信作为国内最大的云计算服务提供商,在“去电信化”改革和云计算服务创新方面也进行了大胆尝

试。其中,关于媒体资源云技术及标准的研究与试验是众多云计算应用创新中的一个案例,它为中国电信实时语音类业务平台云化迁移工作、媒体资源类服务平台的云化改造与转型以及全网媒体资源类服务能力的集约化运营开创了新局面,也为本土软硬件媒体服务器设备生产商以及系统集成商和云计算平台提供商等在新型产业发展中创造了新的商机。

2 媒体资源云化关键技术研究

媒体资源云化的一项关键技术是要求支持多租户服务能力,即按需的资源分配、负载均衡及承诺的个性化等级服务能力。在云计算架构中的多租户云媒体资源池中,每个租户都有权获得指定的媒体资源处理能力、网络及存



储资源的 SLA(service-level agreement, 服务等级协议)。

根据业务模式或组织层次结构的不同,一个租户的 SLA 要求可能比另一个租户高。例如,租户 A 的语音资源计算处理和网络带宽要求可能比租户 B 高,而租户 B 的会议、视频资源计算处理和存储容量要求可能比租户 A 高。所以,多租户模式下虚拟云媒体资源池的主要目标是确保此环境中的各租户正确获得订购的 SLA,同时将其数据、通信和应用程序环境安全地与其他租户分离和隔离,并免受其他租户的干扰。按照网络能力、媒体资源计算处理能力、存储能力等服务等级的不同,进行排列组合,从 level 0~level N,可定义出若干种不同的等级,且要求在同一租户内,允许有不同的服务等级存在。

根据上述设计思路,本文提出了一种媒体资源池云化系统(以下简称云媒体系统)的体系架构,如图 1 所示。该系统包括六大核心逻辑单元:云媒体业务访问门户、性能管理单元、资源调度单元、健康检查单元、云媒体处理单元池以及用户业务属性数据库。

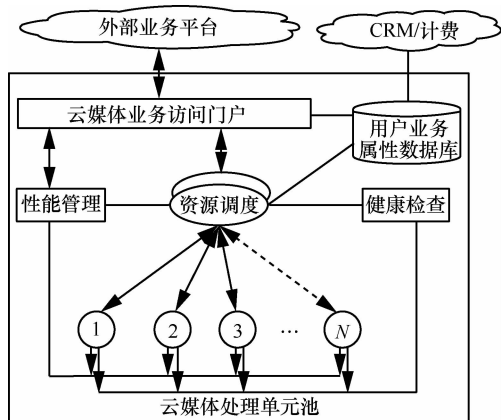


图 1 云媒体系统逻辑架构示意

(1)云媒体业务访问门户对外部业务平台发起的各种媒体类资源请求提供请求转发服务,完成用户 SLA 服务等级和访问权限相关鉴权操作,并屏蔽后端云媒体处理单元的个体差异。同时,云媒体业务访问门户能根据用户业务访问需求、云媒体资源池能力现状等,将云媒体资源池虚拟化为若干个能力集单元,分不同单元独立对外提供服务,例如语音播放能力集群、电话会议能力集群、传真业务能力集群、视频播放能力集群等。

(2)性能管理单元实时采集各云媒体处理单元的业务负荷、并发连接数以及虚拟机资源占用情况等。当云媒体处理单元当前性能参数超过或低于设定的告警门限时,主动启动或关闭云媒体处理逻辑单元,实现云媒体处理能力

的弹性扩展。即当检测到资源池超负荷时,先启动新的媒体处理单元,加入服务队列,然后通知资源调度单元新资源加入服务;当资源池轻载时,选择负载最轻的媒体处理单元,通知资源调度单元,不再分配新的访问至该节点,且当检测到该待关闭单元当前业务链接数归零时关闭该指定单元。

(3)健康检查单元实时采集云媒体处理单元池各实体运行状态。当检测到某媒体资源处理单元无法访问(含虚拟机状态、媒体处理状态)时,通知资源调度单元某云媒体处理单元退出服务队列。

(4)资源调度单元负责接应门户转发的平台访问请求,根据事先配置的负载均衡及业务转发策略,选择对应的媒体处理单元进行资源调配。负载均衡策略包括全局请求轮询、轻载优先、指定处理单元转发、指定区域处理单元转发等。特别是针对电话会议的服务,由于跨媒体处理单元承载一个电话会场时存在混音困难,因此要求资源调度单元在选择媒体处理单元过程中,从用户业务数据中提取用户业务属性,并根据用户发起的会场规模、权限等情况,选择并分配拥有足够会议资源的媒体处理单元,并确保同一用户发起的电话会议的后继加入者请求能转发至同一个会议媒体处理单元负责接应。

(5)云媒体处理单元是构成庞大的云媒体资源池的独立单元实体,具备接应业务平台发起各种媒体类资源处理请求。

(6)用户业务属性数据库记录了用户订购的媒体资源池服务属性及用户标识等信息,为门户访问鉴权及资源调度模块的分配策略执行等提供支持。同时,用户业务属性数据库对用户媒体资源占用情况进行统计与记录。

2.1 云媒体资源池工作原理

基于本架构的云媒体系统中,用户业务属性数据库首先从 CRM 系统获取用户或业务平台订购的服务能力、SLA 等业务属性信息,当用户或业务平台发起请求时,首先对多租户访问权限、业务属性等进行鉴权认证,以确保每个租户的个性化服务需求得以满足。另外,本系统架构中,要求媒体资源处理单元事先需注册在媒体资源调度单元上,由媒体资源调度单元根据当前已在线的各媒体资源处理单元的负荷状态等情况动态控制其业务状态(在线或下线)。

当已在线的媒体资源处理单元均超负荷运行或存在工作状态异常时,性能管理单元和健康检查单元将根据其

主动检查或被动接收到的媒体资源处理单元异常状态报告给媒体资源调度单元并发出告警消息,媒体资源调度单元可综合分析当前在线的各媒体资源处理单元的负荷及工作状态,决定是否唤醒其他下线状态的媒体资源处理单元上线工作或屏蔽当前出现异常情况的媒体资源处理单元服务能力。媒体资源调度单元也可根据各在线媒体资源处理单元上报的低负荷运行信息,逐步激活轻载的媒体资源处理单元的工作状态,从而实现性能弹性扩展。

媒体资源调度单元同样接收性能管理与健康检查单元的监督,当主媒体调度单元负荷超载时,性能管理单元将主动创建新的媒体调度单元实例,并复制当前主用媒体资源调度单元上保持的当前注册及在线的各媒体资源处理单元的工作状态等信息,并通知媒体资源访问门户,将新的媒体处理请求优先转发至新创建的媒体资源调度单元处理。当多个媒体调度单元中出现轻载的实例时,性能管理单元可再次通知媒体资源访问门户,聚焦新到媒体呼叫请求优先转发至某个(或某几个)指定的媒体资源调度实例处理,并逐渐关闭多余的媒体调度单元实例。

2.2 云媒体资源池工作流程

(1)外部应用系统通过负载均衡器向媒体资源访问门户发出请求。媒体资源访问门户访问用户业务属性数据库,获取用户 SLA 服务等级及标准,检查当前可用的媒体资源调度单元列表及转发策略,将请求转发至某选定的媒体资源调度单元处理。

(2)媒体资源调度单元进行访问鉴权,确定合法性后根据外部系统请求访问的媒体属性(语音、传真、视频、E-mail等)检查当前在线的对应媒体资源处理单元列表、分配策略及当前各媒体资源处理单元负荷/健康状态等分配承担本次任务的目标对象。若需启动新的媒体资源处理单元,则即时唤醒媒体资源处理单元并转发呼叫至该服务对象。媒体资源调度单元同时提供面向用户的流量、请求等日志记录。

(3)媒体资源处理单元接收媒体资源分配单元的状态控制及任务分配,启动基于多线程机制的媒体信息编解码处理过程,执行媒体处理任务。当本媒体处理单元负荷超过设定阈值时,则向系统性能管理单元发送告警信息。

(4)性能管理单元定期检查各媒体处理单元、媒体分配单元实例的负荷状态,并接收相关单元主动式性能告

警。当察觉出媒体处理单元性能告警后,性能管理单元则向媒体资源分配单元发出告警消息,由媒体分配单元决策是否需要激活新的媒体处理单元实例。

(5)健康检查单元定期检查媒体处理单元,媒体分配单元工作状态。当发现媒体处理单元状态异常时,则向媒体分配单元发出状态告警消息,由媒体分配单元屏蔽该媒体处理单元的工作能力。

(6)配置管理单元提供面向用户的并发访问能力需求订制、用户鉴权,以及媒体资源分配单元的资源调度策略和系统管理等控制能力。

3 云媒体在 IMS 业务平台中的应用方案

媒体资源云系统在 IMS 业务网络中的位置处于业务平台或应用平台一侧,主要配合 IMS 业务/应用平台提供各种媒体放音类服务,在 IMS 业务网络中的位置如图 2 所示。

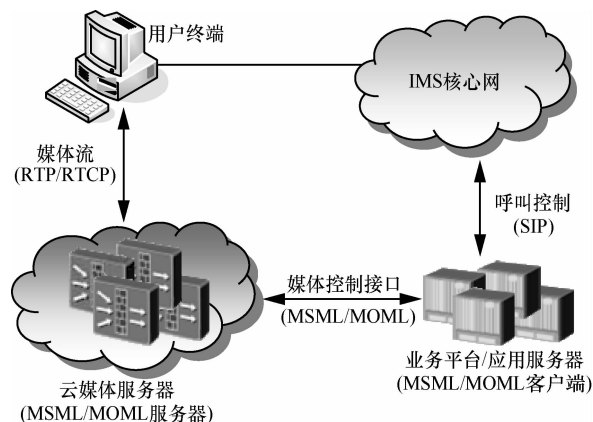


图 2 云媒体在 IMS 业务平台中的位置

云媒体系统要求支持 SIP+MSML/MOML 协议,在业务/应用平台 AS(例如改号通知、彩铃、电话会议等)的控制下,与用户间进行多种媒体类服务的交互,包括向用户播放语音提示、检测接收 DTMF 拨号、播放视频短片、发送/接收传真、创建电话/视频会议等。

云媒体系统可作为传统媒体服务器的备份节点或负载分担节点与传统媒体服务器同期并存,也可独立承担 IMS 业务平台对媒体资源的能力需求。当媒体服务器作为 IMS 业务平台的独立节点存在时,建议基于该云媒体系统构建的媒体资源能力池作将其为“私有云”单独建设,与面向互联网用户提供多租户媒体能力出租的“公有云”在地理位置及网络连接上相互隔离,以满足内部云与外部云在安全等级上的不同要求。云媒体系统平台组网方式如图 3 所示。

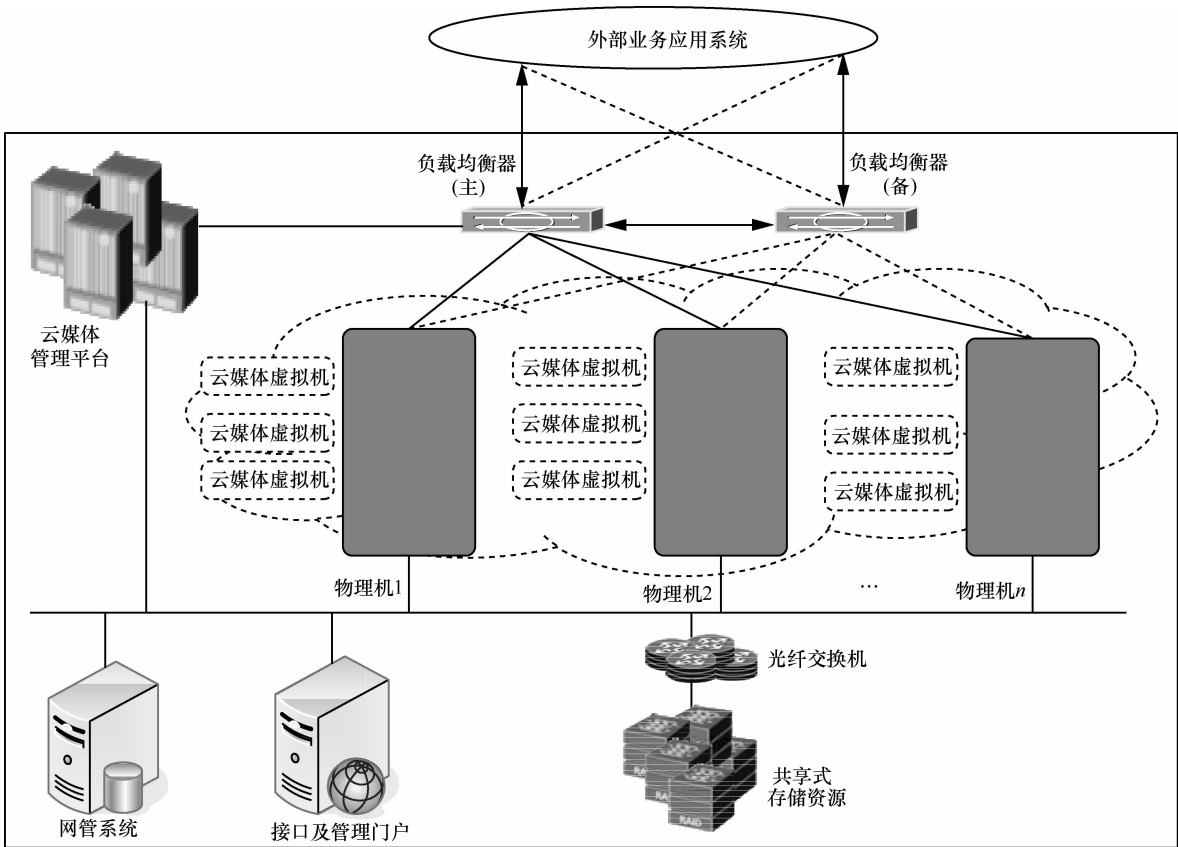


图 3 云媒体系统组网方式

云媒体系统平台主要包括以下组网设备：云媒体管理平台、负载均衡器、云媒体物理主机集群、共享式存储资源、光纤交换机、网管系统、接口机与管理门户等。考虑云媒体服务器的电信级可靠性、安全性要求，系统平台组网上应坚持以下原则。

- 不应出现影响业务的单点设备。
- 物理机集群支持异地组网方式以提高平台整体灾备能力。
- 内外网应完全隔离。
- 具备功能完善的网管监控系统，提升平台运营维护管理水平。

云媒体系统的组网要求如表 1 所示。

4 基于虚拟化环境的媒体资源云技术实验情况

基于本文提出的云媒体服务器系统架构，项目组进行了超大规模组网的可行性、系统性能等的初步技术验证。测试以播放 500 KB 左右 G711 编码的 60 s 语音片文件为例进行。在 1 000 路并发访问情况下，需要 $500\text{ KB}\times 1\,000\times 8=4\text{ Gbit/s}$ 的带宽。在如此高的带宽下保持低延迟显然对网络要求很高。因此，在本次测试过程中，要求云媒体资源控制器必须具备网络缓存能力。

云媒体处理单元测试物理宿主机采用 IBM 3650 (Intel^(R) Xeon^(R) CPU X5550 @ 2.67 GHz 16 核，CentOS

表 1 云媒体系统的组网要求

类型	组网要求	备注
物理服务器与存储设备间的通信网络	光纤链接 双链路冗余设计	为了避免物理服务器与存储间的通信线路出故障而导致所有虚拟媒体服务器应用停机，方案采用双链路冗余机制，即每台物理服务器与存储间都有两条独立的链路
物理服务器与物理服务器间的通信网络	1 000 MB 以太网连接	部署独立的物理服务器间的交换网络，且该网络只用于各虚拟软件服务器之间的数据传输。采用双链路冗余机制

Release 5.8 (Final) on VMware ESXi 5.0.0),其上分配的虚拟机资源为 8 核,4 GB 内存。采用大呼模拟器进行拨打测试,分为拨入接入号、应答、放音 30 s、挂断 4 个步骤,测试结果如表 2 所示。

由表 1 可见,在接近或超过 1 000 线并发时,虚拟云媒体处理单元 CPU 负载会急剧上升,极端情况下甚至会超过 50%。以上测试为最简单的 G711 编码的纯放音(30 s)场景,在采用 G729 等压缩编码或放音收码等相同操作后,CPU 负载进一步上升。同时,由于虚拟机本身会占用一定的宿主 CPU 负荷,因此在高负载情况下,云媒体资源处理软件会导致虚拟机 CPU 负载接近 100%。由此可见,云媒体资源处理软件属于 CPU 高负荷应用,为保证媒体播放质量、降低播放延迟、避免因媒体处理缓慢导致呼损的情况发生,云媒体系统应采用硬件配置尽可能高的物理宿主机进行承载,同时为保证云媒体应用能够在高负载情况下充分利用 CPU 能力,应根据实际应用,尽量降低在同一台物理宿主机上的虚拟机数量。

另外,若为固定语音放音类业务,则可以利用缓存技术以大幅度减少文件传输带宽。目前,软件媒体服务器支持 HTTP 缓存技术,在此场景下,文件若不发生变化则仅

需传输一次,软件媒体服务器会定期查询 HTTP 缓存状态来自动判断是否需要重新传输语音文件,文件传输带宽因此可以忽略不计,平时网络占用仅为放音带宽。

5 关于媒体资源云的商业模式分析

5.1 建设模式

云媒体资源池可采用集中建设模式,依托中国电信云数据中心或传统 IDC 机房的基础设施,从云数据中心或 IDC 机房划分云媒体资源池专区,基于云主机(传统物理机需先部署虚拟化软件)部署软件媒体资源服务器及云媒体管理平台,提供放音、视频播放、音/视频会议、实时转码、传真、收号等各类媒体资源服务。图 4 为云媒体数据中心建设模式示意。

5.2 运营模式

云媒体可采用集约化运营模式,参考目前云数据中心及 IDC 的集约化运营模式,部署云数据中心的云媒体资源池,并将云媒体纳入云数据中心资源管理平台或 IDC 资源管理系统进行集中管理。

云媒体即可作为独立产品面向 ICT 行业客户提供各种媒体资源类服务,也可作为基础能力对外提供整体的行

表 2 云媒体性能测试结果

路数	云媒体虚拟机 CPU 占用	云媒体物理宿主机 CPU 占用	云媒体虚拟机内存(MB)	云媒体虚拟机内存(MB)	呼叫成功率
100 线	5.3%	0.0%	169	646	100%
300 线	16.6%	1.0%	241	979	100%
500 线	28.8%	1.0%	315	1 408	100%
1 000 线	51.2%	8.3%	500	2 511	100%

注:本表 CPU 占用均指多 CPU 加权平均值。由于虚拟机为多核并发处理,单个 CPU 占用超过 100%为正常现象,全局 CPU 占用进行加权平均计算

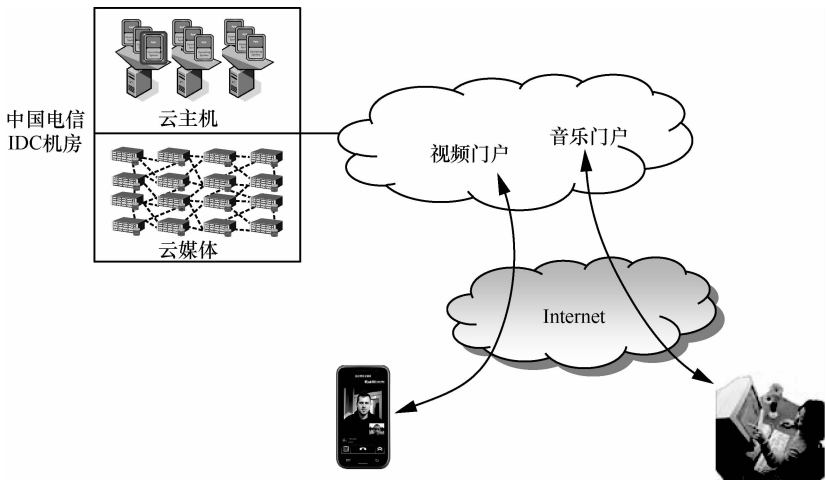


图 4 云媒体数据中心建设模式示意



业解决方案。中国电信可作为云媒体资源需求接收单位对外提供一站式业务受理服务,云媒体资源需求申请流程与云主机产品资源申请流程一致,需纳入中国电信 CRM 系统进行统一管控,包括资源审核、资源审批、维护单位施工、验收、竣工及计费等相关环节。

5.3 推广模式

- 产品应用推广。云媒体资源池可面向视频门户、音乐门户等主流 IDC 行业客户提供按需出租、精确计费的多种音、视频媒体类产品服务,提升云主机产品附加值及客户黏度。云媒体可与云主机产品进行捆绑销售,丰富云主机服务内涵,提升云主机产品功能的竞争力,培养“云主机+云媒体”联合推广的商业模式。
- 内部应用推广。云媒体对内可与中国电信增值业务平台迁移、视讯中心、爱音乐、全球眼等相结合,通过构建弹性媒体资源池,提高中国电信“业务网络”对媒体资源的利用效率。

参考文献

- 1 中国电信股份有限公司. 中国电信云计算技术白皮书
- 2 中国电信股份有限公司. 云计算技术手册
- 3 中国电信股份有限公司. 中国电信云计算 IaaS 服务相关技术

规范

- 4 中国电信股份有限公司. 中国电信服务器虚拟化技术规范

[作者简介]



李青,男,硕士,中国电信股份有限公司上海研究院交换网络部研发工程师,主要从事电信应用新技术研究及增值业务开发等。

唐哲红,女,硕士,中国电信股份有限公司上海研究院部门书记/副主任、高级工程师,主要研究方向为核心网、云计算、增值业务平台技术管理。

宋阿芳,女,硕士,中国电信股份有限公司上海研究院交换网络部研发工程师,主要研究方向为核心网络管理优化、运维新技术研究等。

王发光,男,中国电信股份有限公司上海研究院交换网络部研发工程师,主要研究方向为云计算、电信增值新业务开发等。

(收稿日期:2013-06-03)

· 简讯 ·

2013 年 4 月通信业主要指标完成情况(一)

指标名称	单位	本年本月止累计到达	比上年同期累计(±%)	本月
电信营业收入	亿元	4 232.0	11.5	1 083.4
其中:电信主营业务收入	亿元	3 692.4	8.8	961.8
电信固定资产投资完成额	亿元	652.5	-1.1	164.5
固定本地电话通话时长合计	万分钟	10 068 686.8	-16.9	2 603 093.2
区间电话通话时长	万分钟	1 053 051.7	-21.6	274 255.5
区内电话及拨号上网通话时长	万分钟	9 015 635.1	-16.3	2 328 837.7
固定长途电话通话时长合计	万分钟	1 917 313.4	-17.8	512 593.2
国内长途电话通话时长	万分钟	1 870 436.7	-17.9	500 592.5
国际电话通话时长	万分钟	28 628.3	-12.3	7 295.8
港澳台电话通话时长	万分钟	18 248.4	-14.6	4 704.9
移动电话通话时长合计(含本地)	万分钟	93 637 089.1	6.0	24 261 523.6
移动电话国内长途通话时长	万分钟	21 254 956.9	6.2	5 699 186.2
移动电话国际电话通话时长	万分钟	39 859.0	3.7	10 538.9
移动电话港澳台电话通话时长	万分钟	31 889.1	-5.1	8 098.3
移动短信业务量	万条	30 883 342.8	2.3	7 669 426.0
移动互联网接入流量	万 GB	36 918.3	58.7	9 602.3

注:(1)收入增长率按可比口径计算。(2)固定长途电话通话时长和移动电话通话时长均包含相应的 IP 电话通话时长。(3)通话时长各项指标均为去话通话时长。

(来源:工业和信息化部运行局)