



电信级视频会议系统的并发能力评测方案

丁明玲,余勇昌,王志中

(中国电信股份有限公司广州研究院,广东 广州 510630)

摘要:视频会议系统的评测目前主要停留在功能测试上,而对于系统性能评测,特别是系统支持的最大并发能力评测,目前并没有较好的方案,一般依靠厂商提供的指标,或使用厂商自研的模拟器进行评测。视频系统的并发能力直接影响到大型客户的体验,为了对视频会议系统特别是电信级系统的并发能力有一个客观评价,对其并发能力评测方案进行探讨,并以某电信运营商的视频会议系统为例进行实证,最后给出系统并发能力评测的建议。

关键词:电信级视频会议系统;应用层;最大并发数;最大新建速率;评测方案

中图分类号: TN919

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2016283

Concurrent performance testing method of carrier-class video conference system

DING Mingling, YU Yongchang, WANG Zhizhong

Guangzhou Research Institute of China Telecom Co., Ltd., Guangzhou 510630, China

Abstract: Today, testing method of video conference system is mostly functional, and there is no good solution about performance testing, especially the maximum concurrent ability. The performance indicators are usually provided by manufacturer of system or based on the simulator of system manufacturer. The concurrent capability of video conference system affects the experience of customers. In order to give the objective performance indicators of video conference systems, especially the carrier-class ones, the testing and analysis method of concurrent ability were discussed, then a video conference system of one telecom operator was taken as an example, and the suggestions of testing concurrent ability of video conference system were given.

Key words: carrier-class video conference system, application layer, maximum concurrent connection, maximum connections per second, testing and analysis method

1 引言

随着集约化管理和节能减排的推广以及宽带中国的发展,视频会议系统在企业中的应用越来越多。近年来国内电信运营商积极实施宽带提速工程,这不仅使得数据协作、数据传输、音视频交互更加流畅,而且大大降低了视频

会议系统用户部署和使用视频会议的成本,视频会议市场保持稳定增长, IDC 预计 2018 年中国视频会议市场规模达到 4.9 亿美元。目前除了 QQ 软件视频会议, 华为、Polycom 等厂商的硬件视频会议系统外, 电信运营商以其渠道优势, 也逐步推广自有的视频会议系统。

视频会议系统的客户并不关心系统的原理和技术, 而



是关心使用的体验和效果,这涉及两方面,即系统的功能和性能。功能的测试依据各个功能模块进行,而性能的测试则要基于业务的使用特点。目前的测试基本上以功能为主,性能指标取自厂商标称值,或以厂商自研的模拟器来评测,适用性小,且实际使用情况与模拟器结果差异较大。

在性能的多个指标中,系统的并发能力是大型客户特别敏感的指标,本文主要研究电信级视频会议系统并发能力的客观、适用性强的评测方案。

2 视频会议系统并发能力指标定义

目前主流的视频会议系统一般由终端设备、通信网络和多点控制单元(MCU)组成,而大规模会议系统多采用分级结构,以实现业务的负载均衡,这种结构解决了MCU的单点故障,同时单台MCU的故障不会影响全网业务。典型的多级视频会议的拓扑如图1所示。

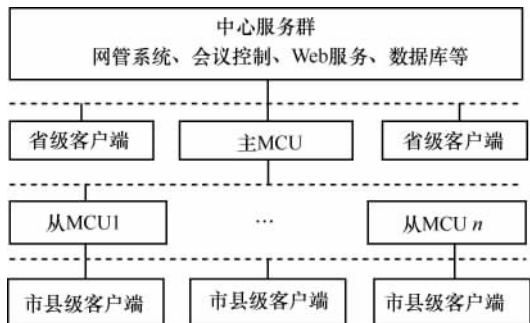


图1 多级视频会议系统典型拓扑

对视频会议系统的并发能力进行评测,首先要明确系统并发能力的真实概念,并结合用户角度研究评测方案,这样评测的结果对实际运营更有意义。从常见错误观点、并发能力指标及业务流量影响等方面进行分析,为评测方案提供依据。

2.1 常见错误观点

关于并发处理能力,有两种常见的错误观点:一种是把并发用户数理解为系统配置的全部用户数,这种观点认为配置用户可能同时使用系统;另一种是把用户在线数量理解为并发用户数,而在线用户不一定同时发起同一个操作,但用户在线数量是统计并发用户数的依据之一。

并发用户数指同一时刻与系统进行相同交互行为的在线用户数。这些用户的最大特征是和服务器产生了交互,这种交互既可以是单向的,也可以是双向的。

并发用户从应用角度可分为注册用户、在线用户和并发用户。并发用户是同时执行某个操作的用户,这个并发在不同场景的情况不尽相同,在实际评测中需要根据具体需求制定相应的方案。

笔者为某电信运营商视频会议系统提供技术支撑的过程中,了解到政企大客户(最大的客户有2 000以上分点)抱怨主要分为两大类:接入过程频繁掉线、清晰度问题。其中,接入过程频繁掉线主要由系统的处理能力引起,本文将深入探讨相应的评测方案。

政企类大客户分部众多、覆盖面广、接入会议时间集中、个体操作的差异性大,评估系统的并发处理能力,要考虑极端情况,即在会议开始时所有分部同时接入及可能出现分部断线、重新上线的情况。这些都是本文案例实证中考虑的问题。

2.2 系统并发能力的主要指标

系统并发能力在评测行业内属于4~7层的性能指标,主要有以下3项。

(1)最大并发数

最大并发数,在评测行业内也称为maximum concurrent connection,代表系统同一时间可以处理的最大请求数,若一个用户有多个连接,比如多个TCP连接,则计算为多个连接数。

(2)最大新建速率

最大新建速率,在评测行业内也称为maximum connection per second,代表系统每秒可以处理最大新发起的请求数,与最大并发数的区别在于,最大并发数着重于最终处理成功的连接总数,每秒处理的速率可以比较小,而最大新建速率着重于系统处理请求的速率,而在这种速率前提下一定时间后也可以达到一个最大并发连接数。笔者认为两项在评估时同时考虑,因为大量用户能够快速成功接入,是保证用户良好体验的门槛。

(3)业务流量对并发能力的影响

系统在处理请求时,无业务流量前提下的最大并发数比带业务流量的并发数高很多,而实际应用中,设备/系统收到大部分请求都带有业务流量,有些多、有些少,所以在考虑评测模型时,根据系统不同操作类型,应设计加载不同的流量,以验证系统真实处理能力。

对于不同的系统,并发能力可能存在不同的过程中,如用户接入过程、使用应用等,所以实际评测时,根据主要过程进行相应的并发能力指标的评测。

3 案例实证

本文以某电信运营商的视频会议系统为例进行评测方案实证。

3.1 案例概述

该视频会议系统定位于政企大客户,最大客户有 2 000 多个地市节点。根据前文分析结果,将其应用于该系统并发能力的评测中,并对比系统开发商的内部测试结果及网上实际运营结果,讨论该评测方案的可行性和参考性。

测试仪表决定了评测方案的一致性、客观性及可回归性,同时只有当测试仪表性能远高于被测系统时,才能在模拟复杂协议、协议流程、加载流量前提下准确获取评估结果。因此,本文选择 4~7 应用层测试领域的主流仪表进行评测,该仪表可发起百万级并发连接数。

3.2 评测模型及评估

3.2.1 评测模型

对评测对象进行研究,是制定评测方案的前提。根据被测系统的功能模型以及收集的用户抱怨,系统的并发处理主要存在两个阶段:会议接入阶段及系统提供的网页服务部分。

按照系统在现网部署的网络结构,同时考虑现网环境复杂性及对指标的影响不可控,并发能力指标应为裸性指标,相应的方案才是可回归、真实、客观的,故应排除网络因素的影响。

验证测试床搭建在隔离的网络环境中,排除网络不可控因素影响;采用具备线速转发能力的 12 口吉比特交换机,每个终端连接一个吉比特端口,排除辅助设备交换能力影响;所有连接线路经过线速测试,排除线路质量因素影响;终端设备采用视频会议系统全新原装设备,包括服务器、终端、多媒体资源等。

测试床如图 2 所示。

其中,tester 为 4~7 层评测领域主流测试仪表,AS tester 为视频会议系统厂商的性能测试工具,终端为视频会议系统原装客户端及多媒体设备。AS tester 是厂商给出并发能力指标的依据,但在实际使用中远未达到此指标就引起掉线等故障,引入该工具是为了与评测方案结果做比较。

选用的 4~7 层评测领域主流仪表,可以根据协议类别设置并发能力的连接参数,也可以根据系统私有协议自定



图 2 并发能力测试床

义协议流程并加载并发能力的参数及流量,所以直观性、客观性、回归性强。

搭建完测试床后,研究该视频会议系统的协议流程是关键一步。各个厂商系统研发的关键协议流程不尽相同,且多使用私有协议,必须深入被测对象的协议流程,将其“复制”到测试仪表中,并与被测系统交互验证,保证协议流程准确、可实施。

根据对系统的协议研究,被测视频会议系统并发能力主要体现在两个阶段:会议接入阶段和网页服务。会议接入阶段由客户端发起,属于纯协议请求,故评测模型中只配置相关协议的流量。网页服务页面有大量的图片,故评测模型中,除了协议流量外,还在每个请求连接中设置了业务流量,根据采集实际操作产生的平均流量配置业务流量大小。

并发处理能力指标包括最大并发数和最大新建速率,两项指标应用在被测系统的两个阶段,并结合分析。根据系统工作模式,会议接入只设置协议流量、网页服务增加业务流量。前文提到用户在加入会议过程中,有多次接入、退出动作,评测将这种行为极端化,所有模型全程模拟部分用户连接、断连的行为。

确定了测试模型后,编辑被测系统并发的两个阶段客户端的协议流程,导入测试仪表,测试仪表模拟视频会议系统客户端与系统对接联调通过后,在测试仪表上修改相应的参数,主要包括并发能力类别的选择,这里主要选择最大并发数或最大新建连接速率;性能加载的阶段设置,并发能力指标值可在这里设置爬升时间和稳定持续时间。协议、参数设置完成,开始运行,同时观察真实客户端上的效果,保证语音、视频运行流畅的前提下,记录测试结果。



3.2.2 会议接入评测

会议接入过程为协议流量,无业务流量,故在测试仪表设置相关协议流程、并发连接数、完成并发连接数时间,时间越长,代表最大新建连接速率越小。设置并发压测时长时考虑开大会的特点,最少设置1 h,考虑实际使用过程中常见的上下线情况,在并发压测时间内,有几十个用户不断上下线,以评估系统在该常见条件下的处理能力。

评估前,系统开发商给出最大并发数为2 000,通过其自研的AS tester工具测试,确实可以达到。

使用本文模型及第三方仪表进行评测,第一次尝试600个并发请求,失败请求非常多,仪表测试结果如图3所示。

图3(a)代表测试仪表设置的最大并发用户数,且按设置正常向系统发起;图3(b)表示仪表发起的请求、成功

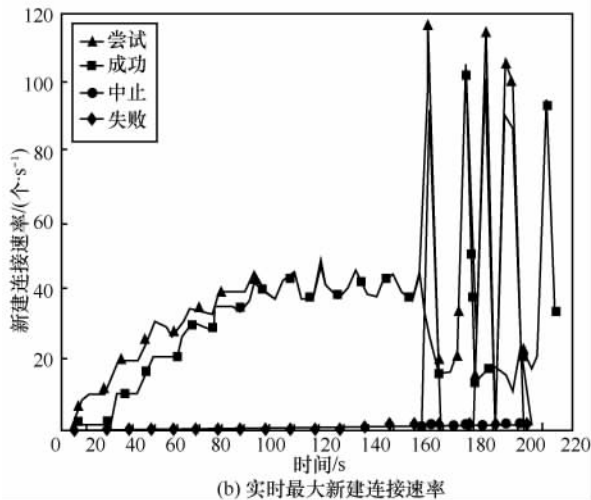
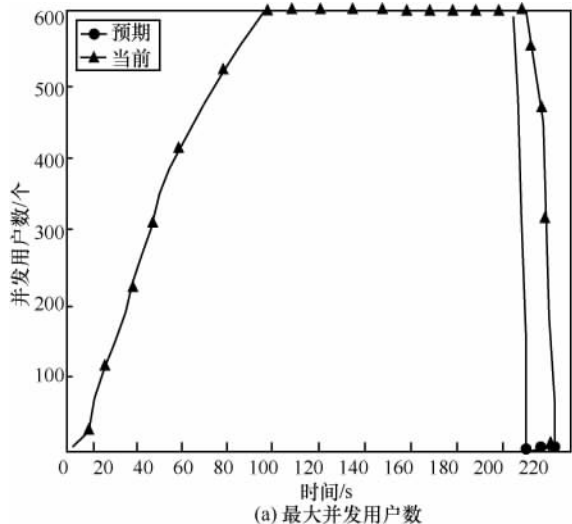


图3 最大并发数为600及实时最大新建连接速率

的请求、失败请求数量,可以看出3 min左右开始失败请求爆发式增长。

通过多次尝试,无失败请求的最大并发数为5。需要说明该结果并不代表系统只能处理5个用户,而是代表在1 h会议过程中,最大新建速率为5个/s,且在有用户不断上、下线的情况下,系统能百分百成功处理请求的结果,对于电信级产品,并发能力指标应在这种条件下评估。当时试用客户反映的情况也是每次都有用户掉线,非常不稳定,符合评测结果。最大并发数分别为10个、5个时的结果参考图4、图5。

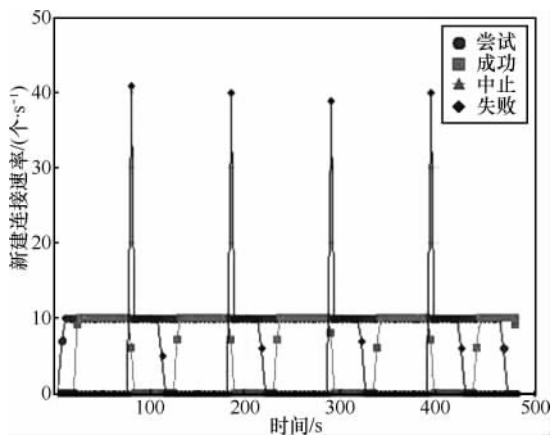


图4 实时最大新建连接速率(最大并发数10个)

在最大并发数为10个、最大新建速率为10的条件下,系统处理得很有规律,约100 s就失败一段时间。

最后尝试成功最大并发数为5个,最大新建速率5个/s,运行1 h,无失败记录,并发3 h,流量处理包括上、下及合并流量,系统处理及流量较平稳。

系统开发人员经过定位、修改、回归测试,能力逐步提升。系统性能完善最后版本,最大并发数达到400个,最大新建速率25个/s,运行3 h,无失败记录,实时处理速率、流量有一定波动,但能正常处理,代表已达系统极限,尝试更大的参数失败。测试的结果如图6所示。

评测结束,参考网络部署的经验,现网部署峰值不超过系统能力的80%,在现网部署多台虚拟机进行负载均衡,保证每台虚拟机上不超过300个用户。完善后的版本上线运行半年,客户表示使用体验良好,无异常掉线、卡死等问题,与评测结果相符。

3.2.3 网页服务评测

网页服务评测的过程与会议接入类似,但在协议流程中增加业务流量。本案例同时做了有/无业务流量的对比

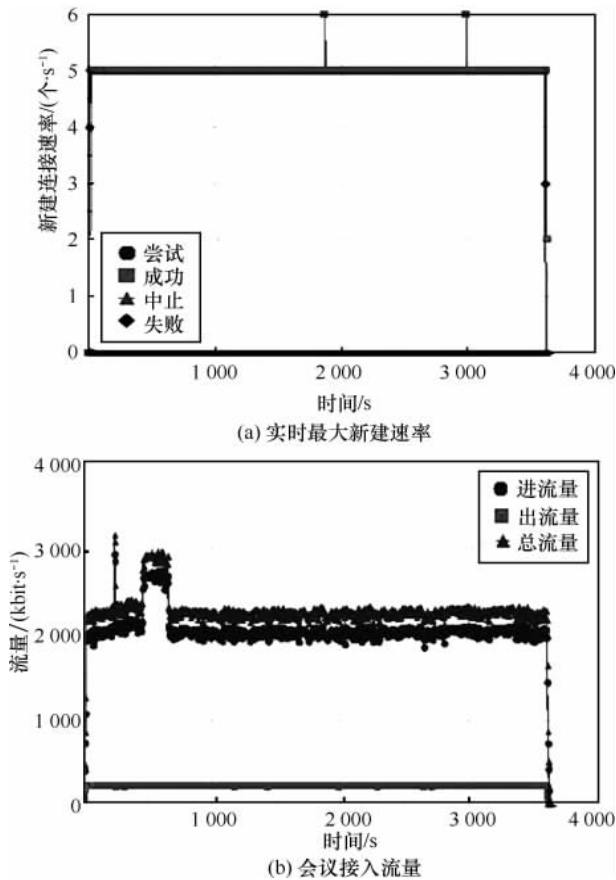


图 5 实时最大新建速率及会议接入流量(最大并发数 5 个)

评估。

无业务流量情况下, 网页服务的最大并发数为 2 000 个, 最大新建速率为 65 个/s, 流量只有协议流量, 因无业务流量的模式与现实不符, 只测试 5 min, 如图 7 所示。

预期并发连接速率与实际能处理的重合, 图线平缓、重合度高, 说明处理性能正常, 反之说明服务器处理压力大。

无业务流量时的网页服务流量如图 8 所示。

根据采集网页数据, 增加每个用户 2.5 Mbit/s 的业务流量后, 最大并发连接数为 1 000, 最大新建速率为 33 个/s, 均为无业务流量的一半, 预验证成功后, 再设置 3 h, 结果也通过。具体如图 9 所示。

仪表发起的最大新建连接速率是 33 个用户/s, 因业务流量是每个用户发起多个连接请求, 所以实时连接速率为 33 的多倍。从用户角度上看, 系统声称的并发能力应该是处理用户数量, 连接数量可以作为另一项指标。有业务流量条件下并发连接的处理波动大很多, 也可以看出服务器处理压力很大, 再增大压力, 便引起请求失败。业务流量处

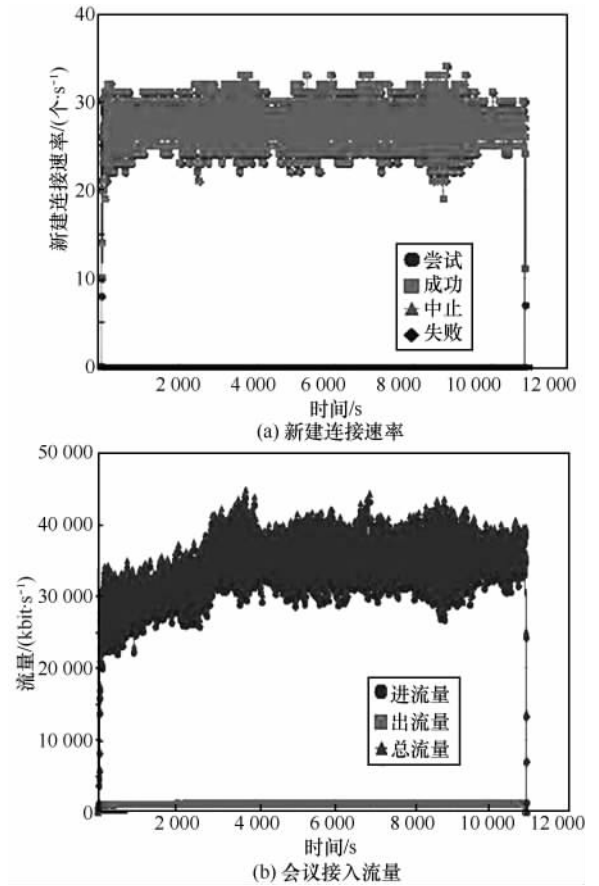


图 6 最大并发数 400 个下实时最大新建速率及会议接入流量

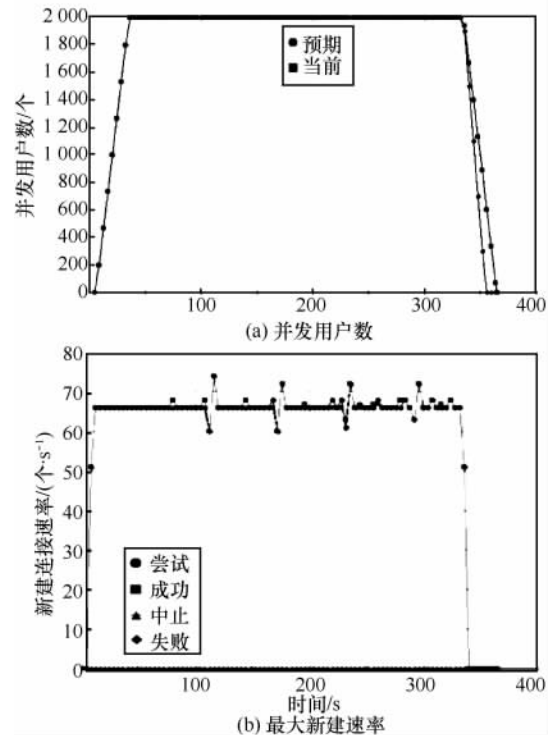


图 7 最大并发用户数及实时最大新建连接速率(无业务流量)

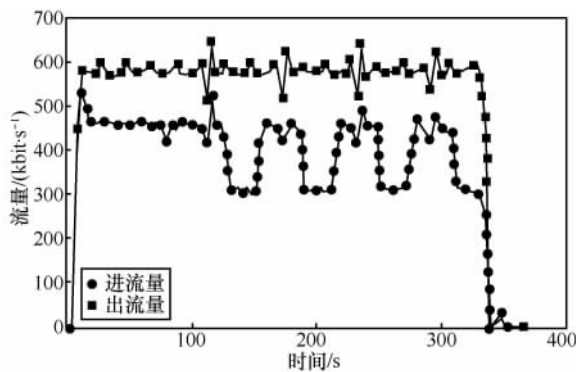


图8 网页服务流量(无业务流量)

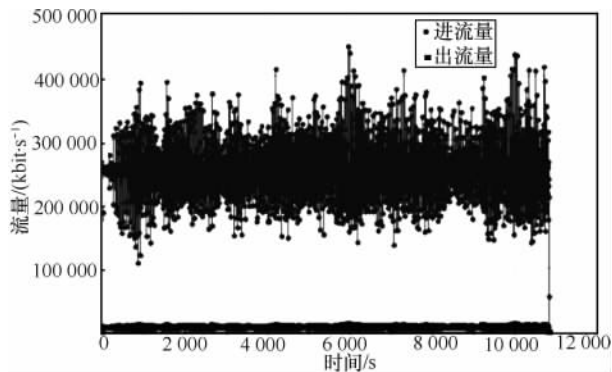


图10 网页服务流量(叠加业务流量)

3.2.4 小结

对于软件系统,并发能力的指标是衡量系统性能的关键指标,并发能力指标主要有最大并发数、最大新建连接数和其在业务流量下表现。并发能力指标是目前最接近系统处理业务能力的指标。

完成系统并发能力的评测,一般的步骤包括:

- 研究被测系统的功能模块、确定并发处理的主要流程,编辑该流程的协议过程;
- 选用业界主流的测试仪表,保证测试的公平、公正、可回归;
- 根据系统在现网应用的网络结构,搭建相同结构的测试床,测试床的辅测设备及线路应保证性能良好,不成为系统并发能力的瓶颈;
- 协议流程“复制”到测试仪表中,在测试床上验证准确性;
- 调整测试仪表中并发能力的参数,叠加相应的协议、业务流量,并考虑可能的异常操作,验证出系统的最大指标;
- 实际部署时,根据最大值的80%或实际要求进行部署,不建议以实验室的最大值部署。

4 评测方案的应用建议

电信级的视频会议系统,除了具备完备的功能外,高性能的处理能力也是必备条件。在处理性能的指标中,本文认为系统并发处理能力非常重要。除了评测出系统的真实能力外,还应达到公正性、一致性、回归性等目标,建议将最大并发连接数、最大新建连接速率、叠加业务流量等指标作为必测项。同时,评测并发能力时建议考虑以下3方面。

(1)采用第三方仪表,提高评测可信度

既然是电信级系统的评测,那么使用的工具必须是业

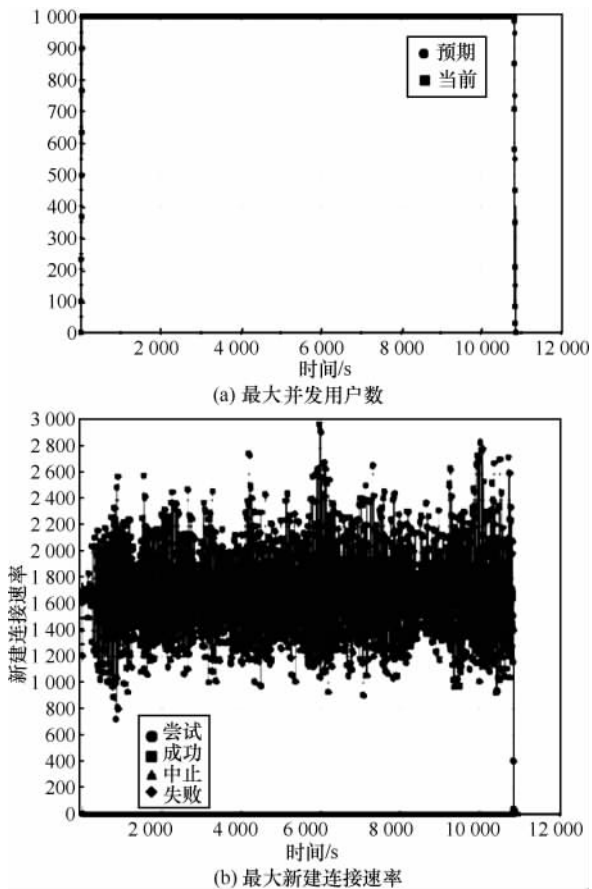


图9 最大并发用户数及实时最大新建连接速率(有业务流量)

理情况如图10所示。

流量平均约为250 Mbit/s,无丢失分组,流量波动较大,说明系统在该业务流量下的处理能力已接近上限。

部署时,同样采用峰值不超过系统能力80%的原则,系统的网页服务模块在现网部署多台虚拟机进行负载均衡,保证每台虚拟机上不超过800个用户,实施后运行稳定,用户反馈使用满意。

界第三方主流仪表,优点包括:

- 硬件类仪表,性能强,不存在测试瓶颈;
- 协议流程标准,规范化,符合一致性原则;
- 第三方,非某个产品自研,符合公正性原则;
- 配置的参数稳定性高,回归性强。

(2)根据用户使用行为,考虑极端情况

目前很多性能评测方案,因采用开发商自研的小工具,测试流程基本是从系统的角度出发,但使用系统的是用户,用户体验是产品的唯一标准,需要从用户行为的角度出发制定评测模型,考虑极端情况,当系统能高效处理极端情况时,系统因压力造成的故障率相对就低,投诉率也会降低。

(3)参考评测结果,制定部署方案

实验室评测并发能力的结果是现网部署的重要参考依据,但考虑到现网的复杂度,建议按不高于实验室评测结果的 80%进行部署,当并发用户数接近 80%临界时,就要考虑优化扩容等工作,这样可以减少因为系统的能力问题引发的故障和投诉。

5 结束语

在完善用户的个性化需求的功能外,并发处理能力是电信级视频会议系统的重要性能指标,这直接关系到大客户的使用体验,满意的体验是开拓市场、留住大客户的基础,但这项指标在目前的视频会议系统评测中很少使用。随着视频会议系统的持续发展,电信级视频会议系统的占比也会逐步增大,做好电信级系统的评测,特别是系统并发处理能力的评测,辅助系统改进优化,将为其在领域内展现实力、赢得口碑、拓展市场奠定坚实基础。

参考文献:

- [1] 陶劲松. 视频会议系统的发展和研究[J]. 电视技术, 2008, 32(4): 94-96.
TAO J S. Research on development and detection of video conference system[J]. Video Engineering, 2008, 32(4): 94-96.
- [2] 杜从宾. IP 视频会议系统中 MCU 容量测试方案的分析[J]. 电子技术与软件工程, 2015(14): 64-65.

DU C B. Analysis of MCU capacity test method on IP video conference system [J]. Electronic Technology & Software Engineering, 2015(14): 64-65.

- [3] 李伟. 视频会议测试工作研究 [J]. 数字通信世界, 2014(8): 34-36.
LI W. The research of video conferencing testing [J]. Digital Communication World, 2014(8): 34-36.
- [4] 徐惠萍, 张其斌, 王娜, 等. 浅谈影响高清视频会议的几个关键问题[J]. 计算机光盘软件与应用, 2014(19): 229-230.
XU H P, ZHANG Q B, WANG N, et al. Several key problems of HD video conference [J]. Computer CD Software and Applications, 2014(19): 229-230.
- [5] 周城, 冯伟东, 周正, 等. 视频会议系统音视频质量评价体系研究[J]. 电视技术, 2013, 37(8): 80-85.
ZHOU C, FENG W D, ZHOU Z, et al. Research on audio-video quality evaluation system of video conferencing system[J]. Video Engineering, 2013, 37(8): 80-85.

[作者简介]



丁明玲(1978-),女,中国电信股份有限公司广州研究院工程师,主要研究方向为无线接入网、电信城域网、骨干网的网络架构。



余勇昌(1976-),男,中国电信股份有限公司广州研究院项目经理,主要研究方向为宽带接入网、无线传感器网络。



王志中(1971-),男,中国电信股份有限公司广州研究院高级经理,主要研究方向为无线接入网、数据交换网络架构。