



面向新型消息业务的垃圾消息管控系统

丁志刚,肖子玉,李雪芳,吉翔川

(中国移动通信集团设计院有限公司,北京 100080)

摘要:从研究新业务背景下的消息业务需求开始,分析了针对新消息监控业务的监控点选择、监控系统架构、系统需要具备的功能以及系统部署需注意的要点。提出了一套满足现阶段监控需求的二级系统架构以及监控系统部署指导原则。

关键词:垃圾消息监控;融合通信;新业务

中图分类号:TN918.91

文献标识码:A

doi: 10.11959/j.issn.1000-0801.2016026

Anti-spam system of new message business

DING Zhigang, XIAO Ziyu, LI Xuefang, JI Xiangchuan

China Mobile Group Design Institute Co., Ltd., Beijing 100080, China

Abstract: In the background of new business requirements, the monitory point selection for new monitoring business, anti-spam system architecture, the function of the system need and the key points that should pay attention to system deployment were analyzed. Finally a set of two level monitoring system architecture and deployment guidelines for monitoring system were put forward.

Key words: anti-spam, RCS, new business

1 引言

随着 LTE 网络的逐渐部署完善,运营商推出了新型的消息类业务以升级替代传统的短/彩信业务,并应对来自 OTT 业务提供商的竞争。随后,对新型消息业务的内容安全进行的监控需求被提出,本文将针对融合通信业务属性、系统建设情况对融合通信垃圾消息监控系统建设思路进行分析。

2 新型消息业务的监控点分析

2.1 现有传统短信业务监控点情况

传统短信业务路由情况如图 1 所示:UE→电路域核心网→SMSC→电路域核心网→UE。

由图 1 可见,传统短信业务中,任何短信均需通过短

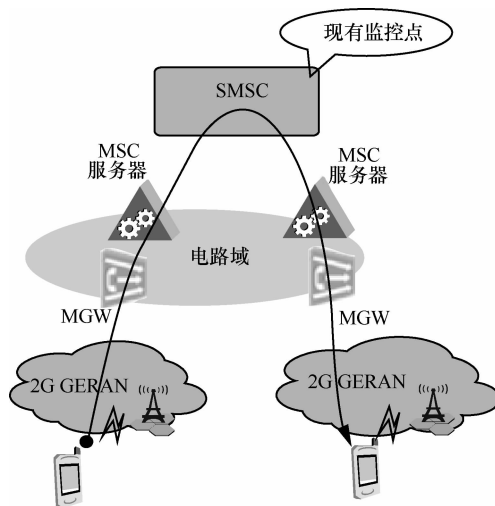


图 1 传统短信业务路由情况

信中心存储转发,因此,监控系统的监控点选择短信中心能监控所有业务;原则上只需要在发送端进行监控,无需接收端监控即可实现监控目的;为了保障监控效果,目前监控系统采用了实时监控的方式,即每条消息均需监控系统鉴权后方可下发。

2.2 新型消息类业务的监控点分析

新型的消息类业务主要包括两种:融合通信业务提供的新消息业务以及 VoLTE 引入后支持的传统短消息业务。以下具体分析两种场景的监控点需求。

融合通信新消息业务路由示意如图 2 所示:UE→分组域核心网(或其他 IP 接入网)→IMS 核心网→RCS→IMS 核

心网→分组域核心网(或其他 IP 接入网)→UE。

VoLTE 支持的传统短/彩信路由示意如图 3 所示:UE→分组域核心网→IMS 核心网→IP-SM-GW→SMSC→IP-SM-GW→IMS 核心网→分组域核心网→UE。

由以上路由分析可见,融合通信新消息业务的大部分消息都不经过现有短信中心,故现有监控系统无法对这部分消息进行监控,需要新增监控点进行监控;而 VoLTE 支持的传统短信仍会通过现网短信中心转发,已在现有监控系统的监控范围内,且业务属性基本一致,故可沿用现有监控系统在现有监控点进行监控。

综上,对于新型消息类业务的监控系统建设仅需考虑

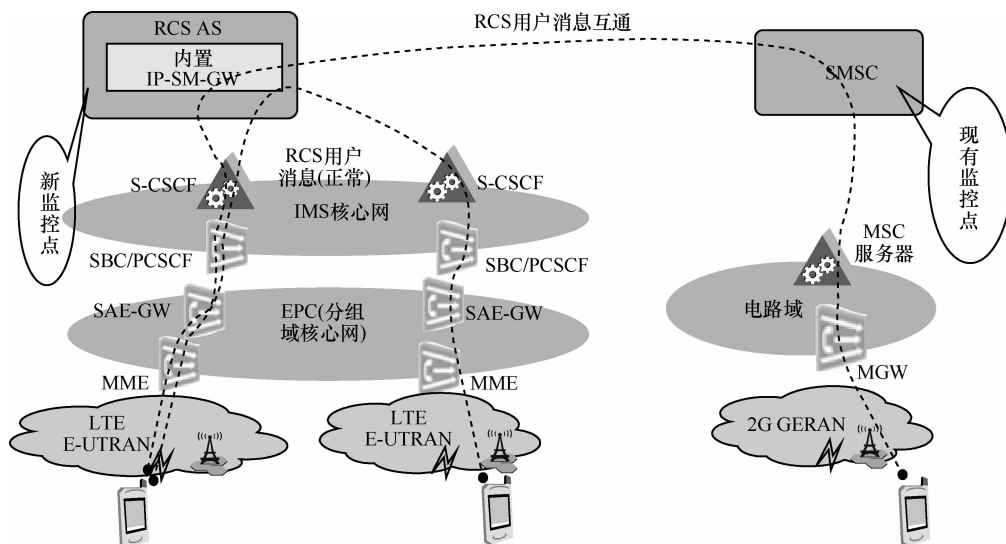


图 2 融合通信新消息业务路由示意

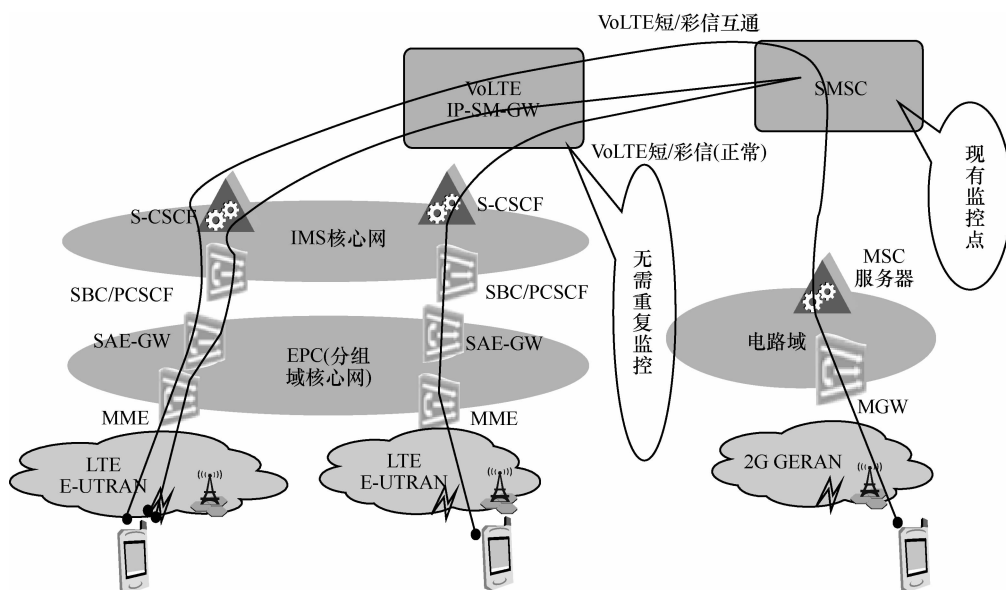


图 3 VoLTE 支持的传统短/彩信路由示意



融合通信新消息业务部分。

3 融合通信新消息业务监控采集方案

本节讨论监控系统对融合通信待监测消息的采集方案。在现有融合通信业务架构下,消息(文本消息、图片、音视频)通过 IMS 标准 SIP/MSRP 经过融合通信业务平台 (rich communication suit application sever, RCS AS) 进行存储转发。因此,对于监控内容采集,采集点选择靠近业务平台效果最好,在此原则下,采集点选择有如下几种方案,如图 4 所示。

方案 1(应用级):定义业务平台与监控平台的鉴权接口,RCS 业务平台完成将监控所需的内容、主被叫、时间等信息解析还原,并依安全监控要求提供接口。

方案 2(信令级):业务平台使用现有消息使用的 SIP/MSRP 信令消息协议栈向监控平台提供鉴权接口,实现内容监控。

方案 3(网络级):在业务平台接入 IP 承载网的 CE 节点镜像数据,通过 DPI 设备解析原始流量获取所需监控内容。

几种方案的优劣比较见表 1。

综合考虑各方案的实现代价和对工程项目进度的影响,建议采用方案 1,即定义应用级接口实现实时的鉴权

功能。具体的接口协议可以选择现有垃圾彩信接口协议 MMS-S 或 SOAP 等。

4 融合通信垃圾消息管控系统的系统架构和功能

4.1 融合通信垃圾消息管控系统架构

现行消息类业务管控系统主要包括垃圾短信和垃圾彩信管控系统,信息集中治理采取“两级共管”的治理模式:垃圾短信,31 省监控系统对发送端疑似垃圾短信进行监控识别,并送中央进行二次过滤及审核处置;垃圾彩信,31 省彩信中心将彩信送集中建设的中央监控平台进行实时/非实时鉴权后,再由中央管理平台进行二次过滤及审核处置,如图 5 所示。

这种架构逻辑上分明,管理界面结构清晰,因此建议对新消息的监控沿用现有消息类业务监控系统的二级架构,由集中管控系统负责统一策略管理下发、人工仲裁,由监控系统负责具体监控操作执行。逻辑功能架构建议如图 6 所示。

4.2 融合通信垃圾消息管控中央管理平台功能

根据以上系统架构分析,中央管理平台应实现的主要监测功能如下。

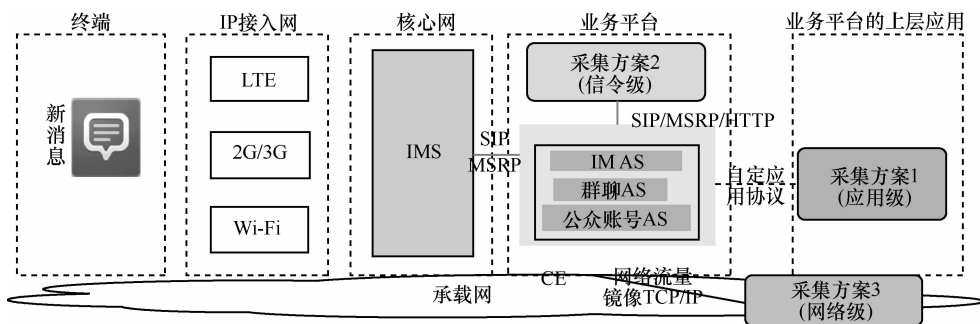


图 4 几种不同的采集点选择方案

表 1 几种采集点选择方案对比

比较点	方案优劣	方案 1(应用级)	方案 2(信令级)	方案 3(网络级)
对 RCS 业务平台影响	3>2>1	对 RCS 业务平台要求较高,业务平台需从信令中解析出原始内容、主被叫及监控所需其他信息并按接口封装后传送给监控系统	需在 RCS IM 业务逻辑中设计有监控平台参与的信令流程,对 RCS 业务平台本身处理能力消耗不大	此方案对于 RCS 业务平台没有任何影响
监控平台技术要求	1>2>3	接口可采用类似现有垃圾短/彩信监控系统与短/彩中心之间的鉴权接口,对监控平台厂商的开发要求较低,亦不需要监控平台具备从信令解析内容的能力	监控平台作为融合通信业务的一个逻辑实体存在,监控平台需要具备处理 SIP、MSRP、RTP 等协议的能力	对监控平台部署的 DPI 设备能力要求较高,需能够基于 IP、TCP、UDP 解析重组 SIP、MSRP、RTP 等数据分组,还原原始样本内容
获取数据、监控效果	1=2>3	监控效果较好,能实现实时监控	监控效果较好,能实现实时监控	无法进行实时监控,依赖于网络流量还原,不一定能保证所有数据完整
未来监控策略灵活性	2>3>1	灵活性较低,面对监控需求的变更时,需要频繁对业务平台接口进行改造	灵活性较高,所有数据均被监控平台掌握,不依赖于业务平台接口	灵活性适中,所有数据均被监控平台掌握,但对 DPI 解析能力要求高

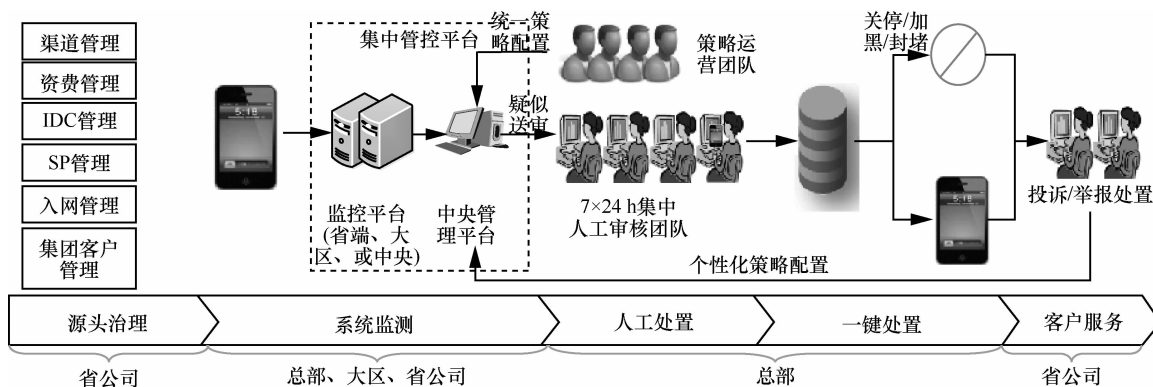


图5 “两级共管”的治理模式

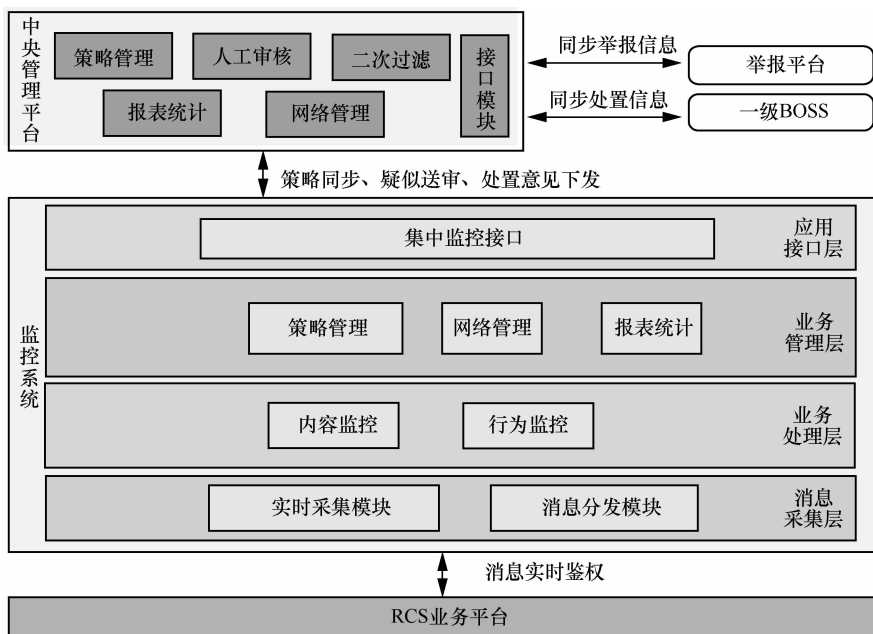


图6 “两级共管”逻辑功能架构

- 人工审核：是监控所有不良信息都应具备的基本手段，主要完成疑似不良信息的审核处理，原则上所有对用户的加黑处理都应该通过人工审核。
- 二次过滤：人工审核之前可对疑似信息执行二次过滤策略以减轻人工审核工作量，二次过滤功能在现有消息监控体系下主要通过一些优质客户名单类筛选实现过滤。
- 策略同步和下发：对监控系统要执行的监控策略进行统一管理，下发给监控系统执行；监控策略可同现网已有各类不良信息集管平台、策略运营分析平台进行联动。

由于被监控的新消息业务初期业务量较小、其业务功能本身也在日趋完善中，而部分管控功能实现代价较高、效果未知，因此在系统建设时建议可以分阶段实施这些功

能，初期实现经验证明最有实施效果、实现代价较低的功能，其他功能根据业务发展考虑实施。具体针对中央管理平台，建议初期实现人工审核、策略同步下发功能，而策略联动功能可暂不通过与现网平台自动接口实现，采用通过人工局数据配置的方式实现。二次过滤功能主要用于辅助人工审核，在初期审核业务量较小时，建议可先不实施。

4.3 融合通信垃圾消息管控监控系统监测功能

根据现网消息类监控实际情况及与主要系统集成厂商调研情况，目前对于融合通信新消息业务中存在的各种内容可采用的监测手段多种多样，分为基于内容的监测手段和基于行为的监测手段两类。同样的，部分功能技术上并不成熟且在业务初期效果未知，表2对各种监测功能原理进行了说明，并分析了各监测功能在初期阶段实现的必要性。



表 2 几种监控功能原理

监控内容	可采用的监测手段	功能说明	实施建议
文本	关键字匹配	指基于关键词库,通过设置各种关键词组合监测不良内容,可配合基于行为监测的手段提高监测效率。目前消息类业务的监控系统均支持,是目前消息类监控中的有效手段,配置容易,对系统消耗一般	建议此功能优先支持,可作为系统的实时策略配置
	正则表达式	指基于正则表达式的高级文本匹配功能。此功能目前消息类业务的监控系统未支持,因此具体效果未知,功能对策略配置人员技术素质要求高,且对系统性能消耗高,厂商均反馈实现难度较大	建议初期不实现
	相似文本匹配	指基于已有不良文本特征模糊匹配类似的文本内容进行监控,目前消息类业务的监控系统尚未支持,不法分子实际发送消息经常变种,此种监控效果应该是不错的,但相关配置模型复杂,且对系统消耗较高	考虑到现网实施经验不足,系统建设初期可主要依靠关键字匹配功能,建议未来再考虑建设
图片	特殊字符处理、拆分子、拼音转换等	指将文本中特殊字符转换成正常字符,以便进行内容匹配。目前消息类业务的监控系统均支持此功能,效果较好,配置难度较高,对系统消耗一般	建议优先支持,可作为实时策略配置
	基于已知库的图片精确匹配	指基于已有不良图片库,采用散列值比对等方法识别不良图片。此功能目前消息类业务的监控系统已支持,是图片类监控中的有效手段,配置容易,对系统消耗较低	建议优先支持,可作为实时策略配置
	基于已知库的图片模糊匹配	指基于已有不良图片库,通过特征提取等方法识别相似图片。此功能目前消息类业务的监控系统未支持,但对图片监控有重要意义,此功能的监测准确性比较依赖于算法,算法需要提供方有长时间科研经验才能提供,一般系统集成厂商不一定具备能力,另外考虑功能对系统消耗较高	建议初期可暂不实现,未来可通过引入第三方识别引擎的方式纳入系统功能中
	OCR 处理	针对采用图片形式发送的文本内容,可通过 OCR 技术转为文本后执行文本识别算法监控。此功能目前消息类业务的监控系统已有支持案例,但效果一般,对系统消耗高	建议未来支持
音频、视频	图片涉黄特征识别	采用智能科学技术建模的图片识别算法,可分析图片是否涉黄,目前消息类业务的监控系统已支持,效果依赖于采用的核心算法,对配置者技术要求高,对系统消耗较高,涉黄图片算法属于专有类技术,需要提供方科研经验丰富才能提供	建议此功能推后提供,未来可通过引入第三方识别引擎的方式纳入系统功能中
	基于已知库的音频视频精确匹配	基于已有音视频文件库,基于散列值比对等方法识别不良音视频,此功能目前消息类业务的监控系统已支持,由于传统短/彩信语音类内容较少,因此效果未知,此功能配置容易,对系统消耗一般	考虑融合通信消息语音内容较多,建议优先支持,针对短音视频作为实时策略配置,长音视频离线检测
	基于已知库的音频视频模糊匹配	基于已有音视频文件库,通过特征提取算法识别出类似不良音视频,此功能目前消息类业务的监控系统未支持,根据调研,业界没有成熟的算法,实现功能对系统消耗影响大	建议初期暂不实施
	声音转文字	把自然语音智能地转化为文字,之后执行文本识别算法。目前消息类业务的监控系统未支持,效果未知,配置难度一般,对系统消耗高	建议未来支持
	对视频抽部分帧后进行图片识别	此功能目前消息类业务的监控系统未支持,效果未知,配置一般,对系统消耗高	建议未来支持
	URL 库精确匹配	基于已有的恶意 URL 库监测恶意 URL 传播,目前消息类业务的监控系统未支持,在不良信息、恶意软件集管系统中有应用,是控制 URL 传播的有效手段,配置容易,对系统消耗低	建议优先支持,作为实时策略配置
文件	通过正则表达式匹配 URL	基于正则表达式的高级匹配功能,目前消息类业务的监控系统未支持,效果未知,对策略配置人员要求高,对系统消耗较高	建议未来支持
	恶意文件散列比对	基于已有恶意文件散列库,根据散列值比对识别传播的恶意文件,目前消息类业务的监控系统未支持,在恶意软件集管系统采用,是控制恶意软件传播的有效手段,配置简单,对系统消耗较低	建议通过融合通信现有平台能力支持,暂时可不纳入监控系统中
	对可执行文件恶意软件引擎监测	基于已有恶意软件监测引擎对文件进行监测,目前消息类业务的监控系统未支持,在恶意软件集管系统采用,控制恶意软件传播的有效手段,配置复杂,对系统消耗高	
	主叫行为分析	基于主叫用户流量特征进行监测,可结合内容监控应用,目前消息类业务的监控系统支持此功能,是对消息类监控最主要监控手段,配置难度一般,对系统消耗一般	建议优先支持,作为实时策略配置
	被叫行为分析	基于被叫用户离散度、主被叫比例、被叫号码规律分析进行监测,可结合内容监控应用,目前消息类业务的监控系统不支持,各厂商均反馈能够实现,但现网无技术经验,效果未知,配置难度一般,对系统消耗较高	建议未来支持
行为监测功能	基于名单的监测功能	对于基于传统消息业务的黑名单、白名单、高风险名单进行监测的功能,目前消息类业务的监控系统均支持,也是消息类监控的必备监控手段,配置难度低,对系统消耗低 配合新消息的业务属性,也可以基于好友名单、通讯录名单、历史记录等来进行信息安全监测,这种监控方式属于新需求,目前消息类业务的监控系统不支持,效果方面可预见是比较有效的监测手段,配置难度一般,但此功能相比一般名单类监测消耗大,需要开通多个系统接口同步名单,对其他系统改造要求高	建议优先支持作为实时策略配置 建议初期阶段暂不实施

4.4 对融合通信垃圾消息的控制功能

识别出不良消息后,必须对消息进行控制,诸如封堵同样内容的消息、封停传播消息的非法账户等。融合通信是富媒体通信业务,实现丰富的多种媒体的通信形式,对监测出的不良内容的控制手段在现有消息类业务控制手段的基础上也应该更加丰富。以下罗列了可采用的一些控制手段及其实现方式。

- 消息封堵:监控系统给融合通信业务平台发鉴权失败消息,业务平台不下发消息,这是整治不良信息的必要手段,近期建议实现。
- 用户处置:在针对运营商系统架构,主要有3种方式实现用户处置,分别是:监控系统处置、BOSS侧处置、业务系统处置。其中在监控系统通过加黑名单方式对用户进行处置对于信息安全管理人员是最灵活和最直接的方式,需要实现。但仅实现监控系统处置只能够封停监控系统所辖业务,而无法进一步对不法用户在其他业务的行为进行控制,因此,BOSS侧处置或业务系统侧处置也是必要的。根据传统的垃圾短/彩信的处置流程,建议采用BOSS侧处置的方式,一方面,这种方式在符合正规的管理流程,即由业务支撑人员负责用户的业务开通和关停;另一方面,在BOSS侧进行处置可最大范围地对不法用户业务功能进行管控。
- URL封堵:对于在管控系统中发现的恶意URL,运营商可在其管理的网络内进行控制,对于在传播服务源在网内的URL域名采用管理手段处置、对于服务源在网外的URL可通过CMNet流控和网间流控系统进行封堵。在初期垃圾消息管控系统只具备基于已知URL的监测手段,因此不会发现新的恶意URL,因此针对URL的控制建议暂不支持。

5 融合通信垃圾消息管控系统部署要点

针对以上系统需求分析,垃圾消息管控系统在建设部署时需注意的要点如图7所示。

(1)监控系统应尽可能与融合通信业务平台同局址部署:按照1亿个用户规模估算监控系统与融合通信业务平台之间的承载带宽,约为46 Gbit/s,若分局址部署,网络带宽压力过大,对承载网能力要求很高,难以承受。

(2)中央管理平台建议与监控系统物理分设:初期阶段,融合通信业务平台仅设置一个节点,随着业务增长,业务平台可能分裂为多个节点,形成多节点的架构,监控系

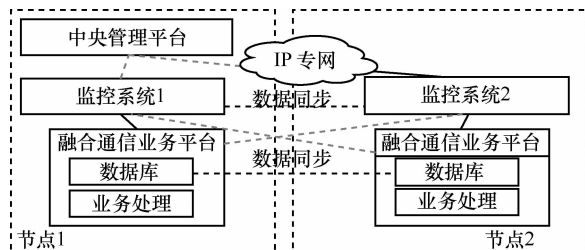


图7 垃圾消息管控系统建设部署要点

统与中央管理平台物理上分散部署便于管理和扩展。

6 结束语

本文从研究新业务背景下的消息业务需求开始,分析了针对新消息监控业务的监控点选择、监控系统架构、系统需要具备的功能以及系统部署需注意的要点。

[作者简介]



丁志刚(1983-),男,中国移动通信集团设计院有限公司工程师,主要研究方向为信息安全、核心网、业务网、自然语言处理。



肖子玉(1969-),女,中国移动通信集团设计院有限公司网络所总工程师、教授级高级工程师,主要研究方向为5G移动核心网、IMS、信息安全监控等。



李雪芳(1984-),女,中国移动通信集团设计院有限公司工程师,主要研究方向为信息安全、核心网数据交换等。



吉翔川(1987-),男,中国移动通信集团设计院有限公司助理工程师,主要研究方向为信息安全、核心网数据交换等。