

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(43) 国际公布日:
2005年12月8日(08.12.2005)

PCT

(10) 国际公布号:
WO 2005/117461 A1

- (51) 国际分类号⁷: H04Q 7/00
- (21) 国际申请号: PCT/CN2004/001303
- (22) 国际申请日: 2004年11月15日(15.11.2004)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
200410042583.0 2004年5月25日(25.05.2004) CN
- (71) 申请人(对除美国以外的所有指定国): 中兴通讯股份有限公司(ZTE CORPORATION) [CN/CN]; 中国广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦, Guangdong 518057 (CN)。
- (72) 发明人: 及
- (75) 发明人/申请人(仅对美国): 傅凯(FU, Kai) [CN/CN]; 袁知贵(YUAN, Zhigui) [CN/CN]; 中国广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦, Guangdong 518057 (CN)。

- (81) 指定国(除另有指明, 要求每一种可提供的国家保护):
AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW
- (84) 指定国(除另有指明, 要求每一种可提供的地区保护):
ARIPO(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚专利(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), 欧洲专利(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

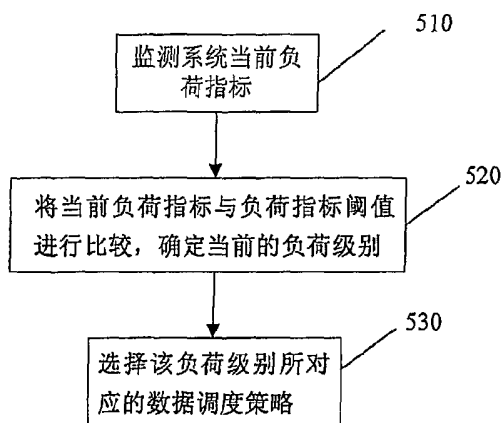
本国际公布:
— 包括国际检索报告。

- (74) 代理人: 北京市中咨律师事务所(BEIJING ZHONGZI LAW OFFICE); 中国北京市西城区阜成门北大街6号-9国际投资大厦C座17层, Beijing 100034 (CN)。

所引用双字母代码和其它缩写符号, 请参考刊登在每期PCT公报期刊起始的“代码及缩写符号简要说明”。

(54) Title: A INTELLIGENT DATA SCHEDULING METHOD OF SHORT MESSAGE SYSTEM

(54) 发明名称: 短消息系统中智能调度数据的方法



- 510 MONITORING THE CURRENT TRAFFIC LOAD INDICATION OF THE SYSTEM
- 520 COMPARING THE CURRENT TRAFFIC LOAD INDICATION WITH THE THRESHOLD OF THE TRAFFIC LOAD INDICATION, DETERMINING THE CURRENT TRAFFIC LOAD LEVEL
- 530 SELECTING THE DATA SCHEDULING STRATEGY WHICH CORRESPONDING TO THE TRAFFIC LOAD LEVEL

(57) Abstract: The present invention discloses an intelligent data scheduling method of short message system. It comprises the steps of dividing the operation mode of short message system into normal mode and strategic mode in the case of heavy traffic service, configuring the switching manner and switching condition between the operation modes, and the data scheduling strategy which corresponding to each mode, in the strategic mode, selectively discarding the data that generating in processing short message traffic; detecting the traffic load indication of the short message system; and comparing the traffic load indication with the switching condition, when it satisfies the said switching condition, carrying out the switching between the operation mode of the short message system according to the configured switching manner; and in the case of generating data in processing short message traffic, scheduling the data according to the data scheduling strategy which corresponding to the current operation mode. The present invention improves the traffic load capability, improves the robustness, stability and adaptability of the system, and ensures the short message successfully sent in the case of the traffic load is heavy.

[见续页]



(57) 摘要

本发明公开一种短消息系统中智能调度数据的方法，包括以下步骤：将短消息系统的运行模式分为正常模式和话务压力过大时的抗压模式，配置模式间的切换方式和切换条件，以及各模式对应的数据调度策略，在抗压模式下对短消息业务处理过程产生的数据有选择性地进行丢弃；检测短消息系统的负荷指标；将所述负荷指标与所述切换条件进行比较，在满足所述切换条件时，按配置的切换方式完成短消息系统运行模式间的切换；以及在短消息业务处理过程产生数据时，按照当前运行模式所对应的数据调度策略对数据进行调度。本发明方法提高了系统抗话务冲击能力，增强了系统健壮性、稳定性和灵活性，并在话务压力大时优先保证了短消息的及时成功发送。

短消息系统中智能调度数据的方法

技术领域

本发明涉及通信领域的各类短消息系统，尤其涉及短消息系统中数据的调度方法。

背景技术

短消息作为电信行业最重要的增值业务，地位日渐重要，现有的 CDMA 网络、GSM 网络和固定电话网中都设置了短消息系统或类似产品。在为广大用户带来便捷服务、为运营商带来丰厚利润的同时，各方对短消息系统的要求也越来越严格，现在短消息系统的安全性和健壮性已向交换机等主设备看齐。在任何情况下都保证短消息系统的正常运作是最基本的需求。

目前，短消息系统的基本架构和数据处理模式决定了它在抗话务冲击和防范数据库服务器故障等方面有很大局限，处理手段较为单一。普遍采用的用于抗话务冲击的过负荷方法会拒绝或缓发大量的正常消息，导致短消息系统发送率和响应率低下，而且在长时间的大话务冲击下，如果积累的缓发消息达到一定量级，则可能产生溢出并导致消息丢失，直至整个系统由于负荷过高而宕机。

例如公开日为 2004 年 1 月 14 日的中国专利申请 02123903.7 中公开了一种应对大话务冲击短消息系统的方法，该方法在负载增大时的处理策略主要是暂存缓发，但这会导致系统的响应率低下；该方法还通过将短消息的临时错误修改为永久错误，从而丢弃发送失败的短消息。然而，话务压力越高，导致的发送失败就越多，如果直接丢弃失败消息而不重发，会使系统的发送成功率大幅降低。

数据库服务器是短消息系统的核心部件，一旦发生故障会对整个短消息系统产生致命影响，因此也需要增强其安全性。现有技术中一般采用系统冗余容灾技术对整个系统进行备份，有效控制整个系统的故障发生率，但其实现较为复杂并且成本很高。

因此，目前的短消息系统急需一种高效易行的智能调度数据的方法，以增强系统的健壮性和安全性，同时保证消息发送的响应率和成功率。

发明内容

本发明正是鉴于现有技术中的上述问题而提出的，其目的在于提供一种短消息系统中智能调度数据的方法，可以使短消息系统根据业务量的变化采用相应的运行模式以对数据进行智能调度，从而提高短消息系统抗话务冲击的能力，并保证短消息系统的发送率和响应率。

本发明提供了一种短消息系统中智能调度数据的方法，根据当前短消息系统的负荷指标选择系统的运行模式，并按照所选的运行模式对短消息业务数据进行调度。

优选地，本发明所述短消息系统中智能调度数据的方法，具体包括以下步骤：

将短消息系统的运行模式分为正常模式和话务压力过大时的抗压模式，配置模式间的切换方式和切换条件，以及各模式对应的数据调度策略，在抗压模式下对短消息业务处理过程产生的数据有选择性地丢弃；

检测短消息系统的负荷指标；

将所述负荷指标与所述切换条件进行比较，在满足所述切换条件时，按配置的切换方式完成短消息系统运行模式间的切换；以及

在短消息业务处理过程产生数据时，按照当前运行模式所对应的数据调度策略对数据进行调度。

优选地，所述智能调度数据的方法还包括：将短消息的业务处理过程分为短消息的起呼处理和短消息的终呼处理两个独立的处理过程，并设置终呼暂停负荷阈值；其中，

在将所述负荷指标与所述切换条件进行比较的步骤之前，还包括：在正常模式下，如果所述负荷指标超过所述终呼暂停负荷阈值，则暂停短消息的终呼处理；如果所述负荷指标低于所述终呼暂停负荷阈值，则进行终呼处理。

由上可知，本发明方法提供了短消息系统的多种运行模式，各模式对应于不同的数据调度策略，并可以灵活地相互切换，使短消息系统可以根据业务量的变化采用相应的运行模式，从而提高了系统抗话务冲击能力，增强了系统健壮性、稳定性和灵活性，由于没有改变现有的发送策略，因此也保证了系统的响应率。另一方面，本发明还可以将短消息业务处理过程分为起呼处理和终呼处理两个独立的过程，并提供了暂停终呼处理的系统减压方式，可以发挥系统潜能并减小浪涌式话务冲击对短消息系统的影响。进一步地，通过优先保存影响发送成功率的短消息数据，因而在话务压力大时优先保证了短消息的成功发送。另外，本发明由于采用容灾模式，可以防范短消息系统因数据库服务器故障而导致的灾难性影响，并以简易方式实现了容灾，节约了系统成本。

附图说明

图 1 是采用根据本发明的一个实施例的智能调度数据的方法的短消息系统的功能模块示意图；

图 2 是短消息业务处理过程的示意图；

图 3 是根据本发明的一个实施例的智能调度数据的方法中短消息系统各运行模式之间的关系示意图；

图 4 是根据本发明的一个实施例的智能调度数据的方法在抗压模式下的业务处理流程图；

图 5 是根据本发明的一个实施例的智能调度数据的方法在抗压模式下选择数据调度策略的流程图；

图 6 是根据本发明的一个实施例的智能调度数据的方法在容灾模式下的业务处理流程图。

具体实施方式

相信通过下面结合附图对本发明的优选实施例的详细说明，可以更清楚地了解本发明的上述和其它目的、特征和优点。

本发明的基本思想是：对短消息系统设置多种运行模式，根据当前短消息系统的负荷指标选择系统的运行模式，并按照所选的运行模式对短消息业务数据进行调度。

根据本发明的一个实施例，首先将短消息系统的运行模式分为正常模式和抗压模式，其中正常模式为现有技术的基本运行模式，抗压模式定义为短消息系统在话务压力过大时的运行模式。然后配置各运行模式间的切换方式和切换条件，以及各运行模式对应的数据调度策略，其中在抗压模式下，对短消息业务处理过程产生的数据进行有选择性地丢弃。检测系统的负荷指标，将其与上述切换条件进行比较；如果负荷指标满足切换条件，则按照上述切换方式完成运行模式间的切换。对于短消息业务处理过程中产生的数据，按照短消息系统的当前运行模式所对应的数据调度策略对其进行调度。

图 1 是采用根据本发明的一个实施例的智能调度数据的方法的短消息系统的功能模块示意图，该短消息系统主要包括业务处理机 101、与该业务处理机 101 相连的操作维护中心 102 和主用数据库 103。其中业务处理机 101 是进行短消息业务处理的核心部件，各种运行模式都在业务处理机 101 上实现。操作维护中心 102 是人机接口，用于提供配置运行模式的人机界面。主用数据库 103 是主用的数据存储设备，用于存放短信息系统中的主要数据，如业务处理所需的用户信息、业务处理生成的短消息数据等等。

以下对短消息系统的运行模式进行说明，如上所述，这些运行模式都在业务处理机 101 上实现。

正常模式：是短消息系统无异常时的运行模式。在该模式下，业务处理机 101 连接主用数据库 103，其业务和数据的处理均沿袭现有标准的处理流程；系统正常读取业务处理所需的用户信息，保存业务处理产生的数据（如短消息等）；此外短消息系统对数据进行完备的保存。

抗压模式：是短消息系统话务压力过大时的运行模式。在该模式下，业务处理机 101 连接到主用数据库 103，对短消息业务处理过程产生的数

据有选择性地丢弃；此外在该模式下，短消息系统仅对数据进行部分保存或全部不保存。优选地，为了优先保证短消息发送的成功率，在丢弃数据时，优先丢弃对短消息发送成功率无影响的数据。

对短消息系统的运行模式进行配置主要通过操作维护中心 102 完成，相关的配置项主要包括：

1) 配置切换方式：各运行模式间的切换方式可以是人工切换或者自适应切换或者两者结合。如果是人工切换，则根据短消息系统的现场环境指标（包括实时的性能统计、话务流量、系统告警类型等）进行手工切换；如果是自适应切换，则由业务处理机 101 根据系统负荷指标（如话务压力和数据库服务可用状况等）按照配置的切换策略自行完成运行模式的切换，各种运行模式切换即时生效，无需复位系统。优选地，短消息系统的负荷指标采用的是 CPU 使用率、系统读写等待比以及缓存使用率中的一种或任意组合；而系统负荷指标的实时获取可以使用业界普遍使用的检测 CPU 使用率、系统读写等待比（IO wait）的应用程序接口（API）函数获得。例如，在自适应切换方式下，如果短消息系统处于正常模式，则检测系统的负荷指标，如果发现负荷指标满足切换条件，则自动切换到抗压模式运行，直到系统的负荷指标满足从抗压模式到正常模式的切换条件后，再从抗压模式切换回正常模式。

2) 配置切换条件：例如对于自适应切换方式，采用切换检测时长和切换检测次数等。优选地，自适应切换方式下的切换条件是：在正常模式下，如果所述负荷指标超过抗压模式负荷阈值且持续一设定时间，则自动切换到抗压模式；在抗压模式下，如果负荷指标小于抗压模式负荷阈值且持续一设定时间，则自动切换到正常模式。例如，短消息系统运行在正常模式下，由于话务风暴等原因造成业务处理机 101 的负荷指标持续 10 分钟超过正常水平，则系统自动由正常模式切换到抗压模式；等到话务风暴过后，短消息系统检测到负荷指标回归到正常水平并持续 10 分钟后，从抗压模式切换回正常模式。

通过上述描述可以知道，采用本实施例所述方法可以使短消息系统根

据当前的话务压力和数据库服务器可用状况等指标，选择合适的运行模式和对应的数据调度策略对短消息业务产生的数据进行处理，从而达到增强短消息系统抗话务冲击能力的目的。

为了充分利用系统的硬件资源，提升应用程序的处理性能，在业务处理实施上可以将短消息业务处理分为两个彼此独立的处理过程：短消息的起呼处理和短消息的终呼处理，其中，起呼处理是指主叫用户发送短消息到短消息系统；终呼处理是指短消息系统将短消息投送到目的用户。图 2 给出了短消息业务处理过程的示意图。起呼处理（线程）210 接收起呼请求并处理后，一方面将完整的短消息保存到主用数据库 103 中，另一方面将短消息的简明信息（包括短消息标识号 ID、目的用户号码等）插入短消息缓冲队列 220 中，然后向主叫用户发送起呼成功的响应，上述操作完成后，起呼处理 210 接收并处理下一个起呼请求。起呼处理 210 无论什么情况均无条件地接收起呼请求并处理。终呼处理（线程）230 按先进先出的原则从短消息缓冲队列 220 中取出短消息的简明信息，并由此从主用数据库 103 中查找对应得短消息进行终呼，终呼处理完成后，将改动的信息再次保存到主用数据库 103 中。每个终呼完成后继续进行下一条消息的终呼。

在本发明的另一个实施例中，在将短消息系统的运行模式分为正常模式和抗压模式的同时，将短消息的业务处理过程分为短消息的起呼处理和短消息的终呼处理两个独立的处理过程，并且配置了终呼暂停负荷阈值，这样，终呼处理线程 230 可以根据短消息系统实时检测的系统负荷指标与配置的终呼暂停负荷阈值的比较结果，决定是否进行终呼处理，从而为系统减压。如果当前系统负荷指标超过了终呼暂停负荷阈值，则暂停终呼处理；一旦系统负荷指标低于终呼暂停负荷阈值，则继续进行终呼处理。当然终呼线程也可以无条件进行终呼处理。在执行完上述操作后，再将系统负荷指标与切换条件进行比较，然后执行后续步骤。

在该实施例中，需要通过操作维护中心 102 配置下述配置项：是否无条件终呼；非无条件终呼时的终呼暂停负荷阈值。优选地，终呼暂停负荷阈值小于或等于抗压模式负荷阈值。

通过上述描述可知，采用本实施例，可以对系统的暂时过负荷（即系统指标超过了正常水平，但又没有达到切换条件）进行减压，一方面发挥了系统的潜能，削平了浪涌式话务冲击对系统的冲击影响，另一方面通过与抗压模式相结合，为短消息系统提供应对不同话务量的数据调度策略，对于系统的业务处理能力和抗冲击能力有本质的提升。

在本发明的另一个实施例中，为了防止短消息系统因为数据库的故障而导致灾难性影响，短消息系统的运行模式还包括容灾模式。容灾模式是在短消息系统的主用数据库无法提供服务（如系统宕机、软件升级等）时的运行模式。容灾模式与上述两个运行模式之间进行切换的切换条件是：短消息系统的主用数据库是否发生故障：如果在正常模式或抗压模式下，主用数据库发生故障，则切换到容灾模式；如果在容灾模式下，主用数据库恢复正常，则切换回主用数据库发生故障前短消息系统的运行模式。在容灾模式下，短消息系统的业务处理机连接到备用数据库，从备用数据库获取业务处理所需的信息。如图 1 所示，备用数据库 104 是备用的数据存储设备，作为主用数据库 103 的备份，在容灾模式下，业务处理机 101 连接到备用数据库 104 上进行短消息业务处理，待主用数据库 103 恢复正常后，业务处理机 101 再断开与备用数据库 104 的连接，恢复与主用数据库 103 的连接，并切换回主用数据库 103 故障前短消息系统所处的运行模式。

在该实施例中，需要通过操作维护中心 102 配置下述配置项：容灾模式下的备用数据库类型、节点位置等，以及容灾模式下的数据调度策略。在本实施例中，容灾模式下的数据调度策略是仅保存待重发的短消息，当然也可以设定数据调度策略为所有短消息都不保存。

优选地，在备用数据库中只保存业务处理所必须的数据，对于业务处理过程产生的数据，只保存待重发的短消息或者全部不保存；备用数据库通过定时导入的方式与主用数据库的关键数据进行数据同步。

通过以上描述可知，采用本实施例，可以保证短消息系统在主用数据库发生故障时仍可正常运行，增强了系统的健壮性，并且大大节约了备用数据库的设备成本。

图 3 给出了根据本发明的一个实施例的智能调度数据的方法中短消息系统各运行模式之间的关系示意图, 其中正常模式、抗压模式和容灾模式三种模式之间都可以采用人工切换或者自适应切换的切换方式。

图 4 是根据本发明的一个实施例的智能调度数据的方法在抗压模式下的业务处理流程图。当短消息系统切换到抗压模式下后, 如果短消息系统是从容灾模式切换过来的, 则执行步骤 410, 短消息系统断开备用数据库, 连接到主用数据库上。在步骤 420, 短消息系统进行正常的业务处理。当在业务处理过程中发起数据操作请求时, 在步骤 430, 判断数据操作的类型。如果是数据读取请求, 则在步骤 440 中, 进行正常的的数据读取操作, 然后返回步骤 420, 继续下一轮的业务处理。如果是数据写入请求, 则在步骤 450 中, 按照当前选择的数据调度策略保存或丢弃该数据, 然后返回步骤 420, 继续下一轮的业务处理。

下面描述如何选择步骤 450 中使用的的数据调度策略。在抗压模式中, 还可以进一步地划分有若干负荷级别, 每一负荷级别配置对应的负荷指标阈值以及相应的数据调度策略。当短消息系统进入抗压模式后, 业务处理机需实时监测短消息系统的如 CPU、输入输出接口 I/O、缓存等的使用情况, 从而判断话务量对短消息系统的负荷级别, 根据不同的负荷级别选择数据调度策略。如图 5 所示的根据本发明的一个实施例的智能调度数据的方法在抗压模式下选择数据调度策略的流程图, 在本实施例中, 假设抗压模式中划分了三个负荷级别: 较高、高和超高, 三个负荷级别的负荷指标阈值为: 较高级别为 CPU 使用率=60%, 读写等待比 IOwait=30%; 高级别为 CPU 使用率=70%, 读写等待比 IOwait=40%; 超高级别为 CPU 使用率=80%, 读写等待比 IOwait=50%。在步骤 510, 检测短消息系统当前的负荷指标, 假设测得的当前负荷指标为 CPU 使用率=70%, 读写等待比 IOwait=45%。在步骤 520 中, 将测得的当前负荷指标与各负荷级别的负荷指标阈值进行比较, 确定短消息系统当前的负荷级别, 在本实施例中, 经过比较可以确定当前的负荷级别为高级别。确定了负荷级别后, 在步骤 530 中选择该负荷级别所对应的数据调度策略, 而各负荷级别对应的数据调度策略是: 在较高级

别时，只丢弃（即不保存）成功发送的短消息；在高级别时，进一步丢弃过期短消息、永久失败的短消息（重发失败 3 次以上）和短消息回执；在超高级别时，再进一步丢弃所有处理完成的短消息，如等待重发的短消息等。因此本实施例中，数据调度策略应采用“丢弃成功发送的短消息、过期短消息、永久失败的短消息和短消息回执”。

当然，抗压模式下负荷级别的级数和对应的数据调度策略可以灵活地设置。需要说明的是，如果当前短消息系统的负荷指标小于较高级别的负荷指标阈值，但短消息系统还处于抗压模式，则采用较高级别所对应的数据调度策略。

上述抗压模式下的负荷级别、各负荷级别下的数据调度策略以及各负荷级别的负荷指标阈值都需要通过短消息系统中的操作维护中心 102 进行配置。

此外，抗压模式还可进一步分为若干对应于不同话务压力的模式，并对这些模式进行切换条件和数据调度策略的配置。

通过以上的描述可以看出，本实施例在抗压模式中进一步划分了不同的负荷级别，并针对不同的负荷级别采用了不同的数据调度策略对短消息系统进行减压，其策略是随着系统负荷级别的逐步升高依次丢弃成功短消息、过期短消息、永久失败短消息、短消息回执直至所有处理完成的短消息。本实施例通过不对短消息进行暂存缓发，保证了消息发送的响应率，而在选择丢弃的数据类型时，任何时候都优先保证消息发送的成功率。

图 6 是根据本发明的一个实施例的智能调度数据的方法在容灾模式下的业务处理流程图。当短消息系统切换到容灾模式后，在步骤 610，业务处理机断开与主用数据库的连接，连接到备用数据库上。然后在步骤 620，进行正常的业务处理。在业务处理过程中，当发起数据操作请求时，在步骤 630 中，判断数据操作的类型，如果是数据读取请求（如读取用户信息进行鉴权），则执行步骤 640，从备用数据库中读取业务处理必要的信息，然后返回步骤 620，继续下一轮的业务处理；如果是数据写入请求（如保存处理过的短消息），则在步骤 650 中，根据容灾模式下的数据调度策略

保存或丢弃该数据，在本实施例中仅保存待重发的短消息，然后返回步骤620，继续下一轮的业务处理。

通过以上描述可以看出，本实施例采用容灾模式后，可以防范短消息系统因数据库服务故障而导致的灾难性影响，增强了系统的健壮性、稳定性和灵活性，并且实现容灾的方式简单易行，节约了成本。

应当理解的是，本发明作为实施例提出的各种运行模式可以视应用需要进行自由组合和变换，例如，可以只采用正常模式和抗压模式进行组合；又如还可以将抗压模式进一步细分为普通话务压力模式和高危话务压力模式，对这些模式也配置切换条件和数据调度策略；此外短消息系统的负荷指标还可以采用缓存使用率等其它指标，对于多个指标可采用加权算法等；在数据调度策略方面，如果负荷持续增大，还可以进一步丢弃所有非必要的短消息，如定时短消息等等。

权利要求书

1. 一种短消息系统中智能调度数据的方法，根据当前短消息系统的负荷指标选择系统的运行模式，并按照所选的运行模式对短消息业务数据进行调度。

2. 根据权利要求 1 所述的短消息系统中智能调度数据的方法，其特征在于，包括以下步骤：

将短消息系统的运行模式分为正常模式和话务压力过大时的抗压模式，配置模式间的切换方式和切换条件，以及各模式对应的数据调度策略，在抗压模式下对短消息业务处理过程产生的数据有选择性地进行丢弃；

检测短消息系统的负荷指标；

将所述负荷指标与所述切换条件进行比较，在满足所述切换条件时，按配置的切换方式完成短消息系统运行模式间的切换；以及在短消息业务处理过程产生数据时，按照当前运行模式所对应的数据调度策略对数据进行调度。

3. 根据权利要求 2 所述的短消息系统中智能调度数据的方法，其特征在于，所述智能调度数据的方法还包括：将短消息的业务处理过程分为短消息的起呼处理和短消息的终呼处理两个独立的处理过程，并设置终呼暂停负荷阈值；其中，

在将所述负荷指标与所述切换条件进行比较的步骤之前，还包括：在正常模式下，如果所述负荷指标超过所述终呼暂停负荷阈值，则暂停短消息的终呼处理；如果负荷指标低于终呼暂停负荷阈值，则进行终呼处理。

4. 根据权利要求 2 或 3 所述的短消息系统中智能调度数据的方法，其特征在于，所述切换方式配置为人工切换或自适应切换或两者的结合。

5. 根据权利要求 2 或 3 所述的短消息系统中智能调度数据的方法，其特征在于，所述自适应切换方式下的切换条件是：在正常模式下，如果所述负荷指标超过抗压模式负荷阈值且持续一设定时间，则自动切换到抗压模式；在抗压模式下，如果负荷指标小于抗压模式负荷阈值且持续一设定时间，则自动切换到正常模式。

6. 根据权利要求 2 或 3 所述的短消息系统中智能调度数据的方法，其特征在于，在所述抗压模式下对业务产生的数据进行选择性丢弃时，优先丢弃对短消息发送成功率无影响的数据。

7. 根据权利要求 6 所述的短消息系统中智能调度数据的方法，其特征在于，所述抗压模式进一步分为若干负荷级别，划分有若干负荷级别，每一负荷级别配置对应的负荷指标阈值以及相应的数据调度策略；所述方法还包括：将测得的当前负荷指标与各负荷级别的负荷指标阈值进行比较，确定短消息系统当前的负荷级别，再选择对应的数据调度策略。

8. 根据权利要求 7 所述的短消息系统中智能调度数据的方法，其特征在于，所述各负荷级别对应的数据调度策略应满足：随着负荷级别的提高，按下述顺序调度数据：丢弃发送成功的短消息；丢弃过期短消息；丢弃永久失败短消息；丢弃短消息回执；丢弃所有处理完成的短消息。

9. 根据权利要求 2 或 3 所述的短消息系统中智能调度数据的方法，其特征在于，所述负荷指标采用 CPU 使用率、系统读写等待比和缓存使用率中的一种或任意组合。

10. 根据权利要求 2 或 3 所述的短消息系统中智能调度数据的方法，其特征在于，所述系统的运行模式还包括在主用数据库故障时启用的容灾模式；在该模式下，短消息系统连接到备用数据库进行业务处理，等主用数据库恢复正常后，再切换回主用数据库发生故障前短消息系统的运行模式，断开与备用数据库的连接，连接到主用数据库。

11. 根据权利要求 10 所述的短消息系统中智能调度数据的方法，其特征在于，所述备用数据库只保存业务处理所必须的数据，对于业务处理过程产生的数据只保存待重发的短消息或全部不保存；所述备用数据库通过定时导入的方式与所述主用数据库的关键数据进行数据同步。

12. 根据权利要求 2 或 3 所述的短消息系统中智能调度数据的方法，其特征在于，所述抗压模式进一步分为若干对应于不同话务压力的模式，并对这些模式进行切换条件和数据调度策略的配置。

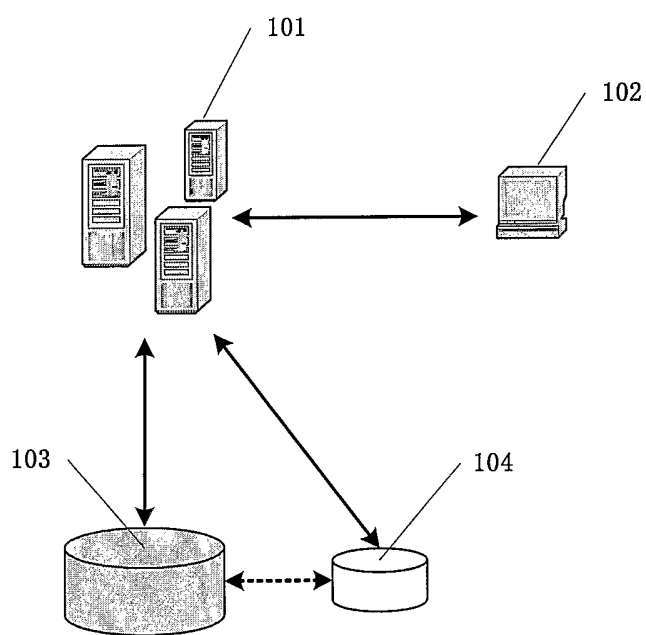


图1

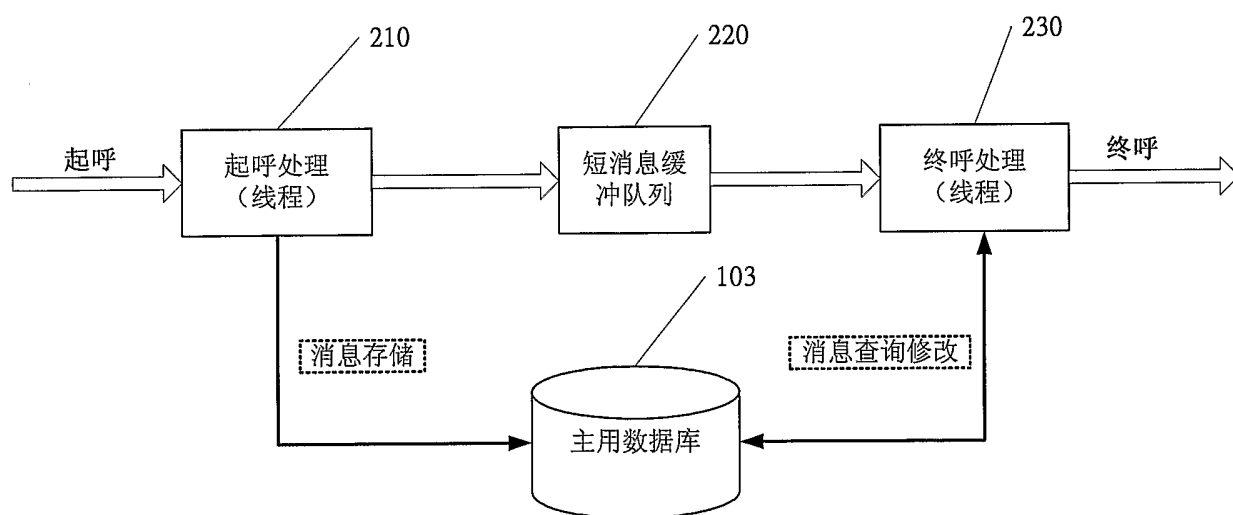


图2

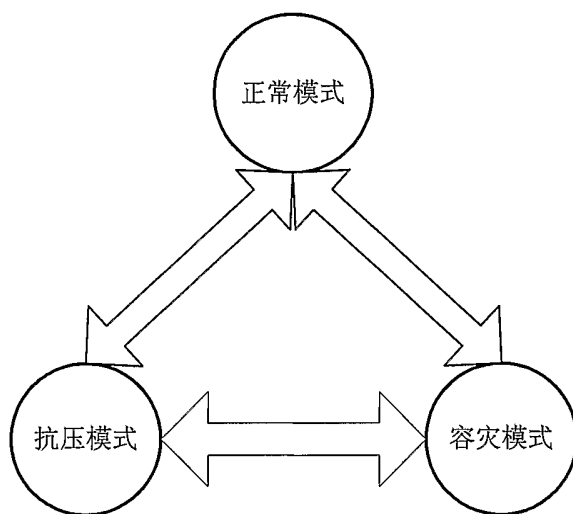


图3

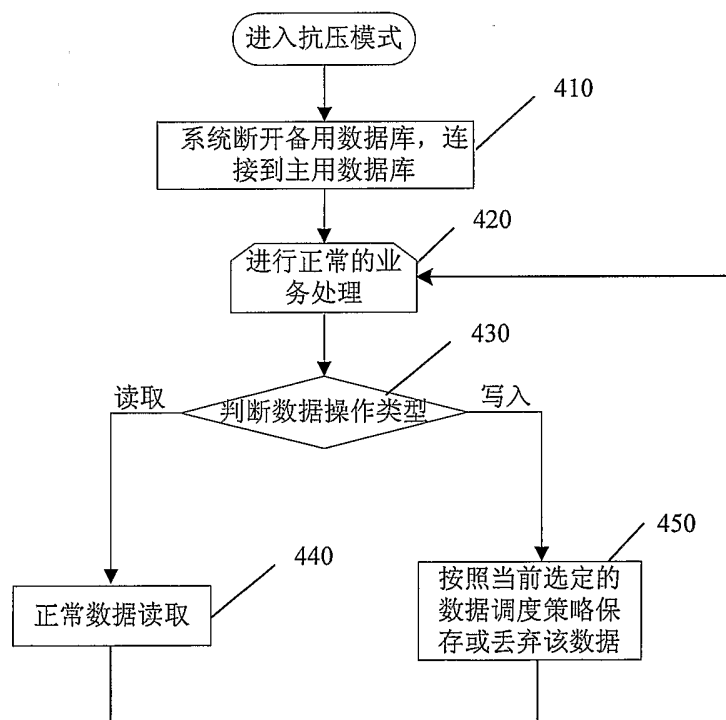


图4

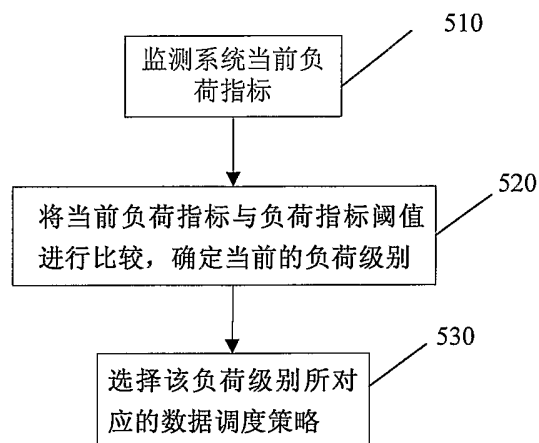


图 5

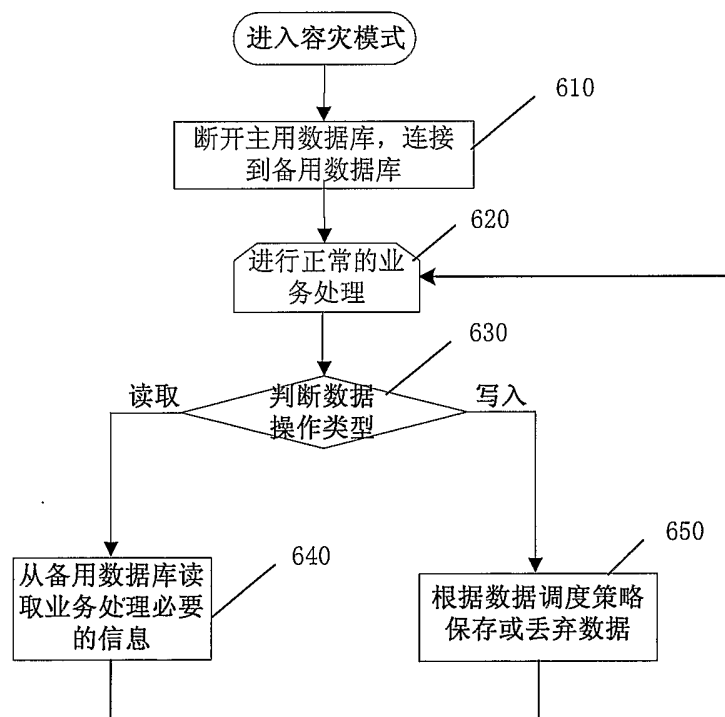


图 6

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2004/001303

A. CLASSIFICATION OF SUBJECT MATTER

H04Q7/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04Q7/00 H04Q7/20 H04L12/24 H04L12/28 H04M3 / 42 H04B7 / 005 (IPC7)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

CNPAT WPI PAJ EPODOC:short message schedul+/al method traffic load+/al indicatoin operation mode

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN, A, 1467952 (HUAWEI TECHNOLOGIES CO.,LTD.) (14.Jan.2004) abstract, description,pag5,line17-page6,line24,pag7,line16-page9,line7	1
A	figure1	2-12
X	CN , A, 1468010 (HUAWEI TECHNOLOGIES CO.,LTD.) (14.Jan.2004) abstract, description,pag2,line16-page3,line6,pag6line19-page7line4,-line2	1
A	pag7line20-28 figure1-6	2-12

☒ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim (S) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 2005.01.15	Date of mailing of the international search report 17 · FEB 2005 (1 7 · 0 2 · 2 0 0 5)
---	---

Name and mailing address of the ISA/
6 Xitucheng Rd., Jimen Bridge, Haidian District,
100088 Beijing, China
Facsimile No. 86-10-62019451

Authorized officer

Telephone No. 86-10-62084567



INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2004/001303

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN , A, 1440596 (TELEFONAKTIEBOLAGET ERICSSON) (03.Sep.2003) Whole document	1-12

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2004/001303

CN1467952A	13.Sep.2001 (13.09.2001)	None	
CN1468010A	14.Jan.2004 (14.01.2004)	None	
CN1440596A	03.Sep.2003 (03.09.2003)	JP2004503126T T	20040129
		US2002004407 A1	20020110
		AU200168012 A	20020114
		WO0203566 A2	20020110
		EP1297636 A2	20030402

国际检索报告

国际申请号

PCT/CN2004/001303

A. 主题的分类

H04Q7/00

按照国际专利分类表(IPC)或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

H04Q7/00 H04Q7/20 H04L12/24 H04L12/28 H04M3 / 42 H04B7 / 005 (IPC7)

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNPAT WPI PAJ EPODOC : short message schedul+/al method traffic load+/al indicatoin operation mode

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
X	CN, A, 1467952 (华为技术有限公司) (14.01 月.2004) 摘要, 说明书第 5 页第 17—第 6 页第 24 行, 第 7 页第 16 行—第 9 页第 7 行及附图 1	1
A		2-12
X	CN, A, 1468010 (华为技术有限公司) (14.01 月.2004) 摘要, 说明书第 2 页第 16 行—第 3 页第 6 行, 第 6 页第 19 行—7 页第 4	1
A	行, 第 7 页第 20 行—28 行及附图 1—6	2-12

☒ 其余文件在 C 栏的续页中列出。

☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2005. 1. 15

国际检索报告邮寄日期

17. 2月 2005 (17. 02. 2005)

中华人民共和国国家知识产权局(ISA/CN)

中国北京市海淀区蓟门桥西土城路 6 号 100088

传真号: (86-10)62019451

受权官员



电话号码: (86-10)62084567

国际检索报告

国际申请号
PCT/CN2004/001303

C(续). 相关文件

	引用文件，必要时，指明相关段落	相关的权利要求
A	CN , A, 1440596 (艾利森电话股份有限公司) (3.09 月.2003) 全文	1—12

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2004/001303

检索报告中引用的 专利文件	公布日期	同族专利	公布日期
CN1467952A	13.09 月.2001	无	
CN1468010A	14.01 月.2004	无	
CN1440596A	03.09 月.2003	JP2004503126T T	20040129
		US2002004407 A1	20020110
		AU200168012 A	20020114
		WO0203566 A2	20020110
		EP1297636 A2	20030402