**软件复用 讨论课2**

**复用产品方案**

1352914 计鹏玥

参考业界架构，讨论程序的扩展

[一． 负载均衡 1](#_Toc28648)

[1.主要应用 1](#_Toc17656)

[2.四层负载均衡VS七层负载均衡 2](#_Toc19581)

[二． 系统监控 4](#_Toc25150)

[1.监控系统软件的框架设计--设计复用 4](#_Toc6844)

[2.监控客户端构件的实现和管理--代码复用 6](#_Toc18416)

[三． ID分配 7](#_Toc12406)

[1.ID分配的几种方式 7](#_Toc2542)

[2.取模块或分段的分布式分配 8](#_Toc23066)

[3.均衡性上的改进 9](#_Toc21197)

[四． 通信可靠高效 9](#_Toc28734)

[1.有效性技术 10](#_Toc30651)

[2.可靠性技术 10](#_Toc7815)

[五． 垃圾消息过滤 11](#_Toc7245)

[1.架构 12](#_Toc3753)

[2.客户端负载均衡 12](#_Toc28151)

# 负载均衡

负载均衡是建立在现有的网络结构之上，提供了一种廉价有效透明的方法来扩展网络设备和服务器的带宽，增加吞吐量，加强网络数据处理能力，提高网络的灵活性与可用性。

## 1.主要应用

（1）DNS负载均衡

最早的负载均衡技术是通过DNS来实现的，它是一种简单有效的方法，但是它不能区分服务器的差异，也不能反应服务器的当前运行状态。

1. 代理服务器负载均衡

使用代理服务器可以将请求转发给内部的服务器，这种加速模式可以提升静态网页的访问速度。

1. 地址转换网关负载均衡

支持负载均衡的地址转换网关可以将一个外部IP地址映射为多个内部IP地址，对每次TCP连接的请求动态使用其中一个内部地址，达到负载均衡的目的。

1. 协议内部支持负载均衡

有的协议内部支持与负载均衡相关的功能，例如HTTP协议中的重定向能力等，HTTP运行于TCP连接的最高层。

1. NAT负载均衡

它将一个IP地址转换为另一个IP地址，一般用于未经注册的内部地址与合法的、已获注册的Internet IP地址之间进行转换。适用于解决Internet IP地址紧张、不想让网络外部知道内部网络结构等的场合下。

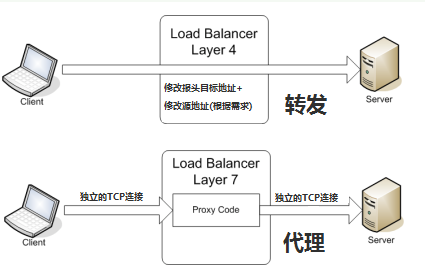
1. 反向代理负载均衡

反向代理（Reverse Proxy）方式是指以代理服务器来接受Internet上的连接请求，然后将请求转发给内部网络上的服务器，并将从服务器上得到的结果返回给Internet上请求连接的客户端，此时代理服务器对外就表现为一个服务器。反向代理负载均衡技术是把将来自Internet上的连接请求以反向代理的方式动态地转发给内部网络上的多台服务器进行处理，从而达到负载均衡的目的。

1. 混合型负载均衡

在有些大型网络，由于多个服务器群内硬件设备、各自的规模、提供的服务等的差异，可以考虑给每个服务器群采用最合适的负载均衡方式，然后又在这多个服务器群间再一次负载均衡或群集起来以一个整体向外界提供服务，从而达到最佳的性能。将这种方式称之为混合型负载均衡。此种方式有时也用于单台均衡设备的性能不能满足大量连接请求的情况下。

## 2.四层负载均衡VS七层负载均衡



（1）四层负载均衡

四层负载均衡是通过报文中的目标地址和端口，再加上负载均衡设备设置的服务器选择方式，决定最终选择的内部服务器与请求客户端建立TCP连接，然后发送Client请求的数据。

在四层负载设备中，把client发送的报文目标地址，根据均衡设备设置的选择web服务器的规则选择对应的web服务器IP地址，这样client就可以直接跟此服务器建立TCP连接并发送数据。。

1. 七层负载均衡设备

七层负载均衡设备也称内容交换，也就是主要通过报文中的真正有意义的应用层内容，再加上负载均衡设备设置的服务器选择方式，决定最终选择的服务器。

其实七层负载均衡服务器起了一个代理服务器的作用，我们知道建立一次TCP连接要三次握手；而client要访问web server要先与七层负载设备进行三次握手后建立TCP连接，把要访问的报文信息发送给七层负载均衡；然后七层负载均衡再根据设置的均衡规则选择特定的web server，然后通过三次握手与此台web server建立TCP连接，然后web server把需要的数据发送给七层负载均衡设备，负载均衡设备再把数据发送给client；所以，七层负载均衡设备起到了代理服务器的作用。

1. 七层负载均衡设备的优点

公司用的负载均衡设备是七层的，其优点有：

1. 使整个网络更“智能化”，能把对图片类的请求转发到图片服务器，对文字的请求转发到文字服务器。
2. 可以有效防止SYN Flood的攻击，使网站更安全。
3. 七层负载均衡设备的缺点

因为七层负载均衡设备其实是一个代理服务器，则对此设备的要求也很高。

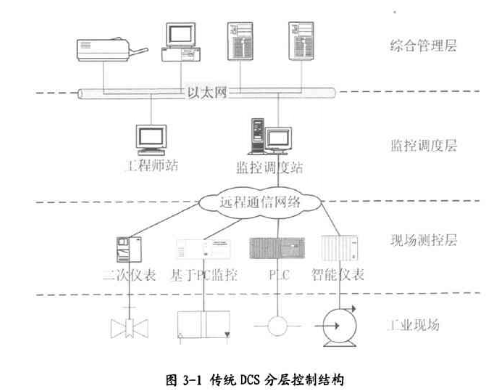
# 系统监控

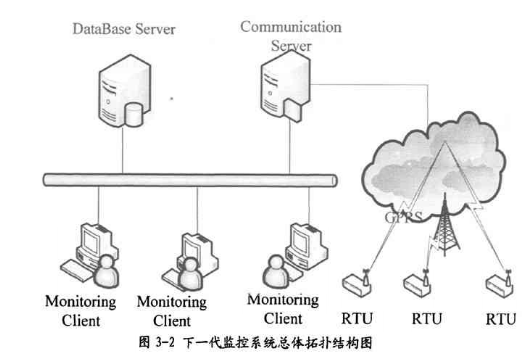
案例：监控系统软件中软件复用技术的应用研究

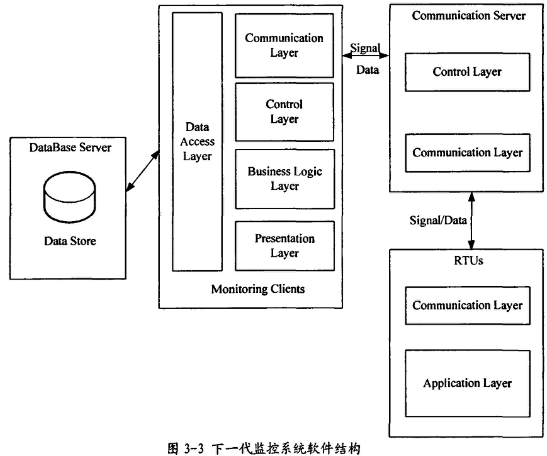
作者：麦通

出版源：北京邮电大学，2009

## 1.监控系统软件的框架设计--设计复用

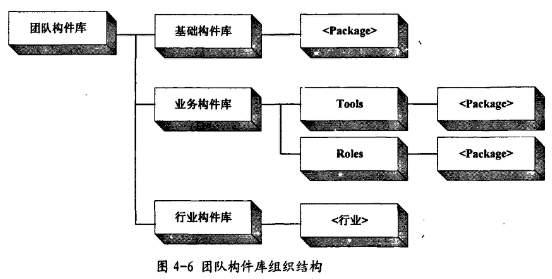




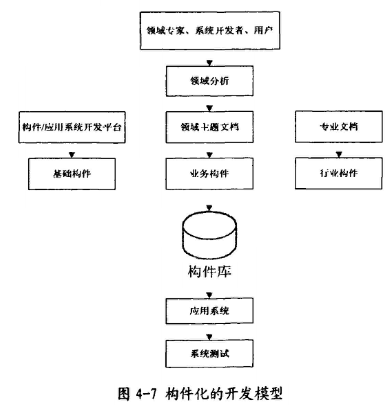


作者对监控系统的可信逻辑，例如监控功能、系统管理、配置管理等bong没有进行本质化的改变，只针对用户的特殊业务逻辑进行了扩展和修改，交媾文档和概要设计文档已达到50%到70%的重用，重用率很高，提高了架构设计和概要设计的效率，同时，一个紊动的框架也有助于开发人员更好地了解整个系统，便于沟通和交流。

## 2.监控客户端构件的实现和管理--代码复用



随着团队构件库的逐步完善，开发团队逐步转变为构件化的开发模式，需求人员对特殊的监控领域进行分析，找到领域中的应用的共同点和不同点，根据用户提交的文档资料形成主题文档和特殊的专业文档，从而将共用的可复用部分和可变部分分开。



这种开发模式有利于团队的长期积累，随着团队的发展和项目的开展，基础构件将逐渐成熟，开发量也将越来越少；随着开发人员的主要精力会更多的集中于业务构件的组织和行业构件的开发，从而能够开发出符合用户需求、让用户满意的软件，提高开发团队的可信竞争力。

# **ID分配**

案例：多IDC环境下的分布式ID分配方案

实际情况：

1. 一些老模块依赖于从0开始自增的ID，数据在内存或文件中以ID为偏移来进行存储。
2. 一些系统依赖于ID的增长做数据分片，例如按取除后分表，因此要求ID在整体上是比较均衡的增长。
3. 在多IDC环境，高延迟加不稳定的网络环境，要求各个分配器彼此之间无需协作，或者可以容忍短期内不可协作。
4. 对一些古董级的老系统来说，还在使用32位的ID，63位ID还是太大了。

因此，我们需要一种分布式高可用、从0开始自增、基本均衡、能够兼容老系统的ID分配方案。

## 1.ID分配的几种方式

（1）单点自增分配

全局由一个模块来负责生成ID，可保证ID从0开始连续递增，数据一般放在本地文件。简洁，但致命的问题是单点故障会导致服务整体不可用。

改进：

为该模块提供主从复制的能力，或者干脆将数据放在mysql里，利用mysql的主从复制，都一定程度上增强了可用性，减轻了单点故障的影响。

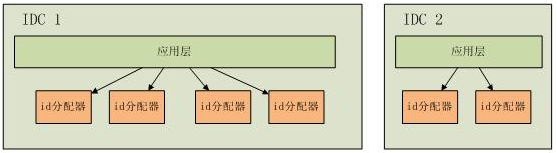
1. 随机/散列分配

通过一些hash算法，比如以时间+随机串为key的md5生成一个唯一的id，关键点在于算法和key的选择要避免冲突。

改进：

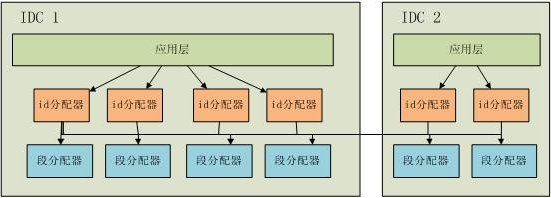
有条件地进行压缩。Twitter的snowflake使用time-41 bits+configured machine id-10 bits+sequence number-12 bits的形式分配ID，但对系统时间的依赖性非常强，需要关闭NTP的时间同步功能，或者当检测到NTP时间调整后，拒绝分配ID。

## 2.取模块或分段的分布式分配



基于方案一再改进一步，将整个id空间按取模或分段等分为若干个独立的id子空间，每个id子空间由一个独立的分配器负责。

## 3.均衡性上的改进



将ID分配分为两层：

1. 上层的“ID分配器”对应用暴露，提供一次申请一个ID的接口，一般本IDC的应用只连本IDC的ID分配器。
2. 下层的“段分配器”对“ID分配器”提供服务。ID分配器“知晓”所有IDC的所有段分配器的存在，使用均衡策略向段分配器申请一个ID段，当所持有的ID段快耗尽时，再请求下一个段。基于方案一再改进一步，将整个ID空间按取模或分段等分为若干个独立的ID子空间，每个ID子空间由一个独立的分配器负责。

# 通信可靠高效

影响通信质量的因素可分为两个方面，一是前面已经介绍过的广义信道的特征及种种限制因素，二是表示信息本身的信号或编码方式和传输（调制解调）方式。有效性与可靠性是相辅相承的两个质量指标体系，模拟与数字通信又有所不同。

## 1.有效性技术

（1）模拟通信技术

模拟通信系统中，每一路模拟信号需占用一定信道带宽，如何在信道具有一定带宽时充分利用它的传输能力，可有几个方面的措施：

1. 多路信号通过频率分割复用/频分复用（FDM）

以复用路数多少来体现其有效性，如同轴电缆最高可容纳10800路4kHz模拟话音信号。目前使用的无线频段从105~1012Hz范围的自由空间，更是利用多种频分复用方式实现各种无线通信。

B.另一方面提高模拟通信有效性是根据业务性质减少信号带宽，如话音信号的调幅单边带（SSB）为4kHz，就比调频信号带宽小数倍，但可靠性较差。

（2）数字通信系统

数字通信的有效性主要体现在一个信道通过的信息速率。对于基带数字信号可以采用时分复用（TDM）以充分利用信道带宽。其它复用方式还有前面提到的空分复用（SDM）、码分复用（CDM）、极化复用（PDM）和波分复用（WDM）以及相应的“多址”方式。数字信号频带传输，可以采用多元调制提高有效性。 另外，为了利用有限的信道带宽支持信源信息量大的通信业务传输，根据信息理论可以采用信源压缩编码，即消除源信息中冗余部分。更进一步，根据不同应用要求的精度，由仙农率失真理论，还可以去掉一些次要信息，这种有损压缩编码，往往可以压缩上百倍以上。

## 2.可靠性技术

（1）模拟系统

对于模拟通信系统，可靠性通常以整个系统的输出信噪比来衡量。一般通信系统特别是卫星通信，发送信号功率总是有一定限量，而信道噪声（主要是热噪声）则随传输距离而增长，其功率不断累积，并以相加的形式来干扰信号，这种干扰称为加性干扰。信号加噪声的混合波形与原信号相比则具有一定程度失真。模拟通信的输出信噪比越高，通信质量就越好。

在实际中，常用折衷办法来改善可靠性，即以带宽（有效性）为代价换取可靠性，可提高输出信噪比涉及信号的调制方式。

1. 数字系统

数字通信可靠性因素就主要地、本质地说，主要还是信噪比问题，另一因素是设计的信号本身抗扰能力。但数字信号传输最终反映在判决输出的码元符号是否正确，因此其可靠性指标均为码元或码字的差错概率如何保证通信的有效性和可靠性，即一定时间内的平均差错率。一般通信系统，差错率主要决定于信噪比输出大小。

# 垃圾消息过滤

案例：垃圾信息过滤系统Mollom

作者：Todd Hoff

来源：<http://highscalability.com/blog/2011/2/8/mollom-architecture-killing-over-373-million-spams-at-100-re.html>

Mollom是一项Web服务，用于将各种类型的垃圾信息从用户生成的内容中过滤掉，这些内容包括：评论留言、论坛帖子、博客帖子、民意调查、联系表单、登记表单和密码请求表单。确定垃圾信息的依据不仅仅有所发的内容，还有发帖者过去的行为和信誉。

## 1.架构

Mollom非常注重技术设计。Mollom过去很重视在代码和服务器资源使用方面尽可能高效。实际上，每台服务器都能处理所有请求，但它们都有完整的故障切换机制。工作负载在多台机器之间分配。如果一台机器出了故障，那么工作负载会转移到另一台机器上。

Mollom过去有三台服务器，但由于大大提升了性能，所以可以将第三台服务器用作登台服务器（staging server）。每台服务器每秒能够处理整整100个连接，每个连接执行整个流程：全文本分析、计算作者的信誉以及提供验证码。

Mollom真正为低延迟进行了优化。由于垃圾信息检测是内容提交到网站的整个过程的一部分，要是检测所花时间很长，用户会觉得很烦人。

## 2.客户端负载均衡

Mollom使用的客户端负载均衡基于延迟等标准来分配负载。作为一家新兴公司，Mollom没有钱来买大型负载均衡工具。他们还有一个目标：能够在多个数据中心之间进行全局负载均衡，这就需要安装一套昂贵又复杂的系统。

客户列表通过API来进行管理。每个客户都有一份服务器列表，列出了可以使用的服务器。每个客户可以使用不同的服务器。为付费客户提供了位置更近的服务器，以缩短延迟。当一台服务器发生故障时，客户会尝试连接列表中的下一台服务器。

客户端负载均衡有助于从旧系统迁移到基于Glassfish的新系统。新用户获得了迁移后的机器的地址，老用户仍在旧机器上工作，只要更新服务器列表，就可以有条不紊地迁移过去。这就允许进行测试，那样用户就可以测试功能，然后测试扩展性和性能。他们观察响应时间、连接队列中有多少个连接，以及连接在队列中停留多长时间。他们可以测试：如果增加线程池中的线程、改变JDBC连接的数量以及其他配置，会出现什么样的情况。一旦每个人都迁移完毕，旧服务器就被关闭。迁移过程中遇到的停机时间极短，这对于高可用性的系统来说很关键。系统停运期间，垃圾信息还是在进来。

客户端方法的一个缺点是，要是第三方客户软件编得很差，就会有问题。比如说，客户获得服务器列表后，可能以相反的顺序对列表进行迭代处理，这其实是不对的。现在Mollom与客户的开发人员紧密合作，提供优质的参考代码示例，那样开发人员可以学习最佳实践。