

基于 FEC 的可靠组播在分布式交换机上的设计与应用

罗 婷¹ 余少华² 汪学舜¹

(华中科技大学计算机学院 武汉 430074)¹ (武汉邮电科学技术研究院 武汉 430074)²

摘 要 提出了一种分布式交换机的通用架构,重点分析了其内部通信机制和特点,发现了可靠组播在其中具有良好的适应性。因此,提出了一种基于 FEC 恢复策略的自适应可靠组播算法,并将其应用于提出的分布式系统,进行了理论分析和计算,最后在试验环境中验证了它的有效性。

关键词 分布式交换机, ForCES, 可靠组播, FEC

中图分类号 TP393 **文献标识码** A

Designing and Implementation of Reliable Multicast Based on FEC in Distributed Switch

LUO Ting¹ YU Shao-hua² WANG Xue-shun¹

(School of Computer Science, Huazhong University of Sciences and Technologies, Wuhan 430074, China)¹

(Wuhan Research Institute of Posts and Telecommunications, Wuhan 430074, China)²

Abstract We presented the general architecture of a distributed ethernet switch, focusing on analysis of the model of internal communication subsystem. According to its character, a novel reliable multicast communication mechanism based on FEC recovery algorithm was applied and evaluated in our experiment.

Keywords Distributed switch, ForCES, Reliable multicast, FEC

互联网的发展使得对以太网交换机的处理能力需求大幅度增加,目前的核心交换机具有数十 G 甚至数百 Gbps 的交换容量。因此,核心交换机多采用分布式转发结合大容量交换结构的体系架构来实现对分组的高速转发,而转发系统可以通过专用 ASIC 实现。

IETF 虽然提出了 ForCES 框架,但是并没有定义 CE 和 FE 之间的消息通信机制,而由于分布式交换机工作时都是一个 CE 对应多个 FE,具有一个发送者多个接收者即组播通信的特点,如何将可靠组播应用到分布式交换机的内部通信系统,就成为值得研究的课题了。

本文描述的研究主要解决了以下两个问题:(1)目前的可靠组播协议都是根据不同的应用发展而来的,并非先天就适应分布式交换机内部通信;(2)以较低的开销实现丢失数据的检测和恢复,以达到分布式系统的实时性要求。

本文首先介绍了分布式通信系统、可靠组播和 FEC 算法的基本理论;其次,建立了分布式交换机内部通信系统模型,并分析了它的通信特征;然后,提出了一种基于 FEC 的可靠组播算法并进行了理论分析和计算;最后,介绍了我们的试验系统,验证了可靠组播算法的有效性。

1 分布式通信系统

随着越来越多的关键应用依赖于 IP 网络,网络的可靠性问题日益成为人们关注的焦点。最新一代的路由技术创新性地提出了将路由功能和转发功能分离来实现网络设备的高可

靠性。当路由功能和转发功能分离后,由专门处理路由的处理器根据邻居信息计算路由信息并导入转发表,在此,我们简称路由功能为控制平面,而转发功能为数据平面或转发平面。

控制平面和数据平面分离带来的最大好处就是:过载的路由消息引起的路由振荡只会影响控制平面的不稳定甚至是崩溃,但是数据转发平面仍然会正常工作,这种机制通常称为“不停顿转发”,目前这种机制多用于单机的路由器。

IETF 的 ForCES^[1] 框架定义了物理上 CE(控制单元)和 FE(转发单元)分离的模型以及它们之间的通信协议,但是目前这些通信协议都是不可靠的。CE 和 FE 之间的通信涉及不同服务水平的可靠性要求。例如,hello 消息的传输不需要考虑服务的可靠性;但是,配置消息和一些同步消息就要求可靠传输,以避免消息丢失引起的数据转发不一致问题。例如,在控制单元到转发单元之间通信的路由信息丢失时,可能引起数据转发出现黑洞转发和错误转发等异常转发行为,有些异常转发甚至会导致网络出现环路。因此,在分布式通信系统中,CE 和 FE 之间的各种通信消息的一致性问题成为实现分布式系统高可靠性的关键问题。

2 可靠组播

在分布式交换机的内部通信系统中包含一个发送者(路由控制单元)和多个接收者(转发单元),这种一对多的模式正好符合组播的应用环境,因此,我们使用可靠组播来研究 CE 和 FE 之间的通信协议。

到稿日期:2009-10-20 返修日期:2009-12-29 本文受 863 国家重点基金项目(2005AA121410),国家电子信息产业发展基金资助。

罗 婷(1976—),女,博士生,主要研究方向为 IP 城域网等,E-mail:luoting@fhn.com.cn;余少华(1962—),男,博士生导师,主要研究方向为 IP 城域网等;汪学舜(1972—),男,博士生,主要研究方向为 IP 城域网等。

近年来,一些新的 Internet 应用如多媒体会议、远程教学等都需要组播的支持,然而 Internet 尽力而为的本质不能保证应用对可靠性和实时性的需求,必须在上层增加差错恢复机制以提供可靠的组播传输。如何为不同的应用需求提供合适的可靠性或实时性是可靠组播研究的重要领域之一。在单播环境中,TCP 协议取得了巨大的成功,基本上能满足各种应用的可靠性需求。但在组播环境下实现可靠传输要复杂得多,存在诸如反馈风暴、重复恢复、恢复延迟、异质性支持以及拥塞与流量控制等一些难以解决的问题。研究表明,满足各种应用需求或者适应所有网络环境的通用型协议在目前的 Internet 中是不现实的。因此人们根据特定的应用需求和网络环境提出了数十种可靠组播协议。

可靠组播分为以下几类^[2]:基于 ACK 方式、基于 NACK 方式、基于 ALC(Asynchronous Layered Coding)方式和基于路由协议方式的可靠组播。

基于 ACK 的方法也称为基于 Sender-Initiated 方法:接收端为每个包产生惟一正确反馈 ACK,发送端匹配反馈和已发送的数据包,没有收到反馈的数据包被认为是被丢失的报文。该机制将导致发送端负担过重:首先,它需要在发送端保存所有接收者的状态信息来确保收到所有的 ACK;其次 ACK 风暴将导致发送端的资源消耗过多。针对上述问题,提出了一种新的改进机制,即基于聚集树的 ACK 方法,通过在组接收成员中选举组服务器,来代替其它接收端产生的 ACK 消息及发送聚集之后的 ACK 消息,典型的应用是 RMTTP。

基于 NACK 的方法又称为基于 Receiver-Initiated 的方法:接收端检查正常接收的数据包的序号,不连续的序号意味着数据包的丢失,接收端向发送端报告标识数据丢失的 NACK 反馈。基于 NACK 的机制减少了发送端的负担,但是 NACK 的风暴仍然可能导致网络拥塞和接收端繁忙,通常的策略是采用随机延时实现 NACK^[3]抑制,例如 SRM(Scalable Reliable Multicast)协议。

为了减少接收端发送的 ACK 或 NACK 反馈信息,提出了基于 ALC 的可靠组播协议。基于 ALC 的协议是在尽力而为的 IP 组播上实现的,它可以为多个不同类的接收端传输批量数据。在 ALC 协议中,数据被编码而使不同的组接收者以分层模型的方式接收。数据的传输速率取决于组播组的数量,因此一个接收者可以根据自己的处理能力决定是加入一个还是多个组播组。Digital Fountain 是一个典型的 ALC 协议的实现,目前,IETF 正在将基于 ALC 的可靠组播协议进行标准化。

基于路由协议的可靠组播是依靠组播路由协议来实现的,例如 PGM(Pragmatic General Multicast)。它通过组播路由协议构建的组播转发树,实现了大量接收者接收组播数据的能力,而采用 NACK 和 FEC 机制实现了错误检测和恢复。

本文提出了一种可靠的分布式组播通信协议,即 Reliable Distributed Communication Protocol(RDCP)。RDCP 是基于 NACK 的可靠组播协议,主要为传输块数据而设计,例如在分布式交换机中控制单元传输大容量的转发表至多个转发单元,并且 RDCP 是不依赖于 IP 层及应用层的协议,它是基于链路层的协议,适用于多种网络拓扑环境。

3 FEC

可靠组播主要解决网络中一个发送者发送数据包到多个

接收者时发生的包丢失问题。目前有两种技术用于解决该类问题:ARQ(Automatic Repeat Request),该算法通过直接重传丢失的数据包来实现可靠传输;FEC(Forward Error Correction),它通过传输冗余数据,即原始数据加上校验数据来实现可靠传输。

本文不对 ARQ 作出研究,而是重点研究 FEC 算法。FEC^[4-6]技术是可靠组播中进行数据恢复的一种非常健壮的解决方法,它可以很好地缓解拥塞发生时带来的不利影响,并能充分利用网络可用带宽资源。FEC 是一种单轮(无需反馈)的协议,发送者对源内容消息按照一定算法进行编码后发送给接收者,而接收者接收到足够的编码消息后,可以重构出源内容消息。

假设有一个大数据文件 F ,划分成 n 个大小相同的消息块(Message Blocks) B_1, B_2, \dots, B_n 。FEC 编码机制使用对消息块进行位异或(\oplus)运算的方法,生成相应的校验块(Check Blocks)^[5]。例如,一个校验块 C_1 可定义为 $C_1 = B_2 \oplus B_1$ (校验块的度数是由组成它的消息块的数量来决定的,此例中, C_1 的度数就是 2。因此,消息块也可以被认为是度数为 1 的校验块)。那么编码后的文件 F 就变成由一连串校验块组成。

接收者通过进行反复迭代运算得到消息块的方式来解码编码文件。如接收者已接收到上面定义的校验块 C_1 ,并且接收到 $C_2 = B_2$,那么就可以恢复出 $B_1(B_1 = C_1 \oplus B_2)$ 。

纠错编码的特性可以确保接收者能够在接收到 $(1+\epsilon)n$ 个校验块后重构源文件 F , ϵ 的大小取决于具体纠错编码。 $1/(1+\epsilon)$ 则代表了 FEC 编码的有效性。

采用 FEC 算法,通过校验数据恢复原始数据而不是直接重传原始数据有以下优点:

增强传输有效性:一个单独的校验包可以修复 n 个数据包中的任意一个数据包,这意味着不同的数据接收者可以使用该包修复不同的丢失数据。

增强扩展性:ARQ 传输数据包时,需要发送者了解接收者丢失的数据包序号。而使用 FEC 校验,只需要发送者知道在一个数据块中接收者丢失数据包的最大数量。

减少冗余的接收:通过多播重传来修复丢失数据将导致接收者收到过多的冗余数据,将浪费链路带宽,而通过 FEC 可以避免该情况,从而极大减少链路带宽的占用。

4 分布式模型分析

图 1 是对分布式交换机内部通信抽象之后的模型。为了对其内部的分布式通信机制进行分析,首先作以下定义:

用 $S = (N, C, f, M)$ 来表示一个分布式通信系统,其中:
 N 表示分布式通信系统中转发单元(FE)的集合;
 C 表示分布式通信系统中内部通信子系统的集合;
 $f: N \rightarrow P(C)$ 表示将各个 FE 连接到通信子系统的映射函数;
 $M = \{(s, r, c, id)\}$ 表示分布式交换机中内部通信的消息类型,其中 s 表示发送者, r 表示接收者, c 表示承载消息的通信实体, id 表示通信的消息类型。

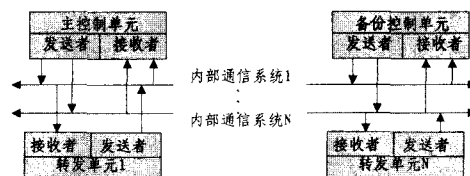


图 1 分布式通信系统的内部通信模型

一般而言,分布式交换机中的内部通信子系统 C 有 3 种方式:总线方式、以太网介质和交换结构。其中,总线方式(如 PCI、CPCI 和私有总线)速度较慢,且无法扩展到跨网络设备单元(NE)使用,因此不在本文的研究范围以内。而以太网和交换结构能够对单播和多播等多种数据传送方式提供硬件支持,能够提供服务质量(QoS)保障,因此是内部通信系统的主要方式。由于交换结构主要用于提供高速的数据转发通道,本系统采用以太网作为分布式通信的物理介质。考虑到系统的可用性,内部通信子系统 C 采用 1+1 的保护方式,即通信的数据正常运行时在主用(Primary)通道上传送,当主用通道出现故障时,切换到备用(Backup)通道。

N 为系统的 FE 数目,可能多达 1000 个左右。但由于分布式交换机多用于网络的边缘汇聚节点,其集合 N 一般小于 64 左右,因此,本文中考虑 N 的范围在 100 以内。

映射函数 f_c 是各个卡到内部通信子系统 C 的映射方式, f_c 可以是基于 IP、UDP 或 TCP 的,也可以是直接基于链路层的。 M 即内部通信协议,它包括发送方、接收方和消息的定义。 M 的具体定义由具体的通信内容和通信环境决定。

在分布式交换机内部,通信的方式一般为主从模式,即流量主要分布在 CE 和 FE 之间;而 FE 之间和 CE 之间的流量相对较少。从流量特征来看,分布式交换机通信机制的流量有以下 3 个特点:

实时性:由于交换机需要即时地对外部事件进行响应,因此,需要实时地将相关信息发送到转发单元。

同步性:在分布式架构的交换机中,每个转发单元中需要维护全局的转发信息,以保证转发行为的一致性。因此,一个转发表项需要同时转发到多个转发单元上。

突发性:正常情况下,控制单元和转发单元之间只需要很少的通信内容,如链路健康检测报文以及其它协议的控制报文。但当交换机重新启动时或网络拓扑发生变化时,控制单元在短时间内需要将数十甚至数百 k 的路由表项发送到转发单元。因此,其流量特点具有突发性。

此外,由上文对 C 的分析可知,分布式交换机的内部通信子系统 C 有以下特点:

平坦性:在分布式交换机内部通信中,各个发送者和接收者的地域位置非常接近,系统结构和性能比较接近,传播延迟和往返时间非常小而且比较平均,这种情况属于“平坦型”的网络拓扑。因此,在广域网中比较突出的“异构性”问题在 C 中并不明显。

高带宽和低传输误码率:在内部通信子系统中,传输带宽较高且误码率低。数据丢失大多数是由于网络拥塞或者是通信终端处理能力不足,传输过程中出现错误的比例极小,在这种环境中考虑可靠传输的性能,更多的是考虑如何降低错误恢复过程占用的终端处理能力。

静态性:在分布式交换机内部通信中,发送者 s 和接收者 r 的关系相对比较稳定。除了启动和主备切换时,接收者 r 的动态添加和删除并不频繁。

优先级策略:内部通信系统中的控制单元和转发单元均采用智能化的专用 ASIC 实现,可以利用自身的带宽分配和等优先级控制策略,为可靠组播协议中的不同数据提供不同的传输等级。例如,原始数据包采用最高优先级的组播组,反馈等控制报文采用低一级优先级的组播组,重传报文采用最

低优先级的组播组,同时可以对控制报文和重传报文实行带宽限制。采用这种分级传输和流量控制的 QoS 方法,可以避免错误恢复过程影响原有的组播发送过程,避免在网络负担较重的情况下引起整个传输质量的下降。

5 分布式交换机的实现和性能分析

图 2 显示了一个分布式交换机原型系统的设计,根据 IETF ForCES 中的定义,将一个网络设备划分为控制单元(CE)和转发单元(FE)两部分。其中,CE 负责各种网络配置管理、信令协议、路由协议的控制报文的处理;而 FE 则仅根据转发信息表(FIB)对数据报文进行分类、流量监管、整形和转发。在我们的系统实现中,主要包含两个重要的模块:即 RDCP(Reliable Distributed Communication Protocol)和 IDP(Information Delivery Protocol)。

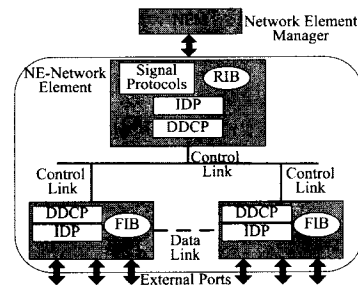


图 2 分布式交换机原型系统的体系架构

可靠组播通信协议(RDCP)实现以下功能:(1)为 CE 和 FE 提供自动发现机制,实现 CE 和 FE 的内部绑定;(2)CE 接收 FE 发送的控制消息,并进行鉴权后再传输到 IDP 处理;(3)支持不同可靠级别的消息传输。

信息分发协议(IDP)主要实现各种信令消息的管理,例如转发信息表、过滤表、地址转换表,该协议要确保 CE 和 FE 之间的各种控制管理信息的同步,以避免不一致性引起的数据丢失或错误转发等问题。

IETF 的 ForCES 工作组定义了 CE 和 FE 之间的通信机制,但是 IETF 的工作仅局限于通信接口的标准化,而我们的研究更加关注它的实现细节,使其更可靠、灵活和安全。在此,本文将详细描述 RDCP 协议,在实现中,假设所有的 CE 和 FE 位于一个网络单元,对于多网络单元的讨论不在本文的讨论范围内。

RDCP 提供两种可靠传输的方式:在 C 上的基于 NACK 的可靠多播方式和基于 TCP 套接口(Socket)的单播传输方式。之所以在 RDCP 中保留了基于 TCP Socket 的单播传输方式,是因为考虑到在核心交换机中的 FE 较少的情况下,基于 TCP 单播的方式仍能具备较高的性能,我们将在实验部分比较两种方式下的性能。而在可靠多播的实现中,由于 C 具有高可靠性的特点,为了避免 ACK 泛滥,我们采用了收端检测的 NACK 方式,而不是发送端检测的 ACK 方式。

RDCP 的协议模型如图 3 所示,主要组成为 RM(reliable multicast)层和 FEC 层。FEC 层负责对分组进行 FEC 编码以及从 FEC 冗余数据中重构原始数据,而 RM 层和应用层接口,将数据传递给 FEC 层或接收来自 FEC 层的数据,采用 NACK 机制实现分组的可靠发送和接收。

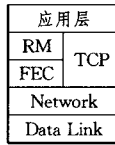


图3 RDCP的分层模型

RDCP的可靠多播的基本过程包括4个阶段:发送端初次发送;接收端错误检测,包括NACK产生和NACK抑制;接收端进行FEC校验和发送端FEC恢复策略;接收端成功接收并根据消息内容执行相应操作。RDCP采用数据包级的FEC算法实现可靠多播的数据包修复。

下面对RDCP算法进行理论分析和计算。为方便分析,假设各节点丢失报文的概率相互独立。分析中用到的符号作如下定义:

p :表示每个接收端丢失报文的概率;

M_r :表示节点 r 成功接收一个报文所需的重发次数;

M :表示一个组播组成功组播一个报文所需的重发次数。

首先来推导几个表达式:

在没有采用FEC编码的RDCP算法中,接收端 r 经过 m 次传输,成功接收报文的概率是:

$$P(M_r \leq m) = 1 - p^m, m = 1, 2, \dots$$

$$E(M_r) = 1 / (1 - p^m)$$

所有接收端正正确接收任一分组需要的传送次数 M 的概率分布为:

$$P(M \leq m) = \prod_{r=1}^R P(M_r \leq m) = (1 - p^m)^R$$

$$E[M] = \sum_{m=0}^{\infty} [1 - (1 - p^m)^R]$$

在采用FEC编码的RDCP算法中,令 $q(k, n, p)$ 表示接收端没有接收到来自发送端发送的任意数据块的概率。对于发送端发送来的组播报文,如果RM层丢失,则只要FEC接收的数据包大于 $n-1$ 个报文中的任意 $h-1 = n-k-1$ 个即可恢复出原始数据, RM层的包丢失概率可以表示为:

$$q(k, n, p) = p \left(1 - \sum_{j=0}^{n-k-1} \binom{n-1}{j} p^j (1-p)^{n-j-1} \right)$$

M' 表示为使所有RM接收端都接收到组播数据而需要的报文重传次数;单个RM接收端的报文重传次数小于 $i+1$ 的概率为 $1 - q(k, n, p)^i, i = 0, 1, \dots$,则 M' 的概率分布为:

$$P(M' \leq i) = (1 - q(k, n, p)^i)^R, i = 0, 1, \dots$$

M_{FEC} 等价于 M' 加上FEC层的奇偶校验,其平均值可由以下公式表示:

$$E[M_{FEC}] = n/k \cdot E[M'] = n/k \cdot \sum_{i=0}^{\infty} (1 - p(M' \leq i))$$

FEC编码可提高分组传输的可靠性,但引入了冗余,降低了传输的有效性,特别是固定的FEC编码,在网络传输质量很好、分组丢失率很低(如 10^{-6})的情况下,FEC提高可靠性的作用减小,但是仍然会产生相同的冗余。

在分布式交换机中,许多信息需要CE转发到FE以保证数据转发的一致性。例如,二三层转发表、访问控制列表、MPLS标记转发路径表等。

如果能使编码的纠错能力和发送的数据块大小以及分组的丢失率相匹配,那么有效性和可靠性的矛盾就能得到很好的解决。针对本系统的特点,本文提出的RDCP协议中使用了一种新的自适应FEC编码机制:

(1)发送者在发送数据前,对 k 个原始数据分组进行FEC编码,产生 h 个冗余数据。发送者以Multicast方式发送 k 个原始数据及 α 个FEC编码数据, $\alpha < h$ 。

(2)接收者接收的数据大于 k 时,恢复出原始数据。如果接收分组数 $< k$,等待一随机时间后,以Multicast方式发送NACK,其中含丢失分组数据信息 f ,并启动定时器 $T1$ 。

(3)如果某一接收者在发送NACK前的随机等待时间中收到一个NACK,判断NACK中的 lm ,如果大于等于自己丢失分组数 ln ,那么启动定时器 $T1$,当作自己发送了一个NACK分组;否则,发送一个含 ln 的NACK分组,启动定时器 $T1$ 。

(4)发送者收到NACK之后,将FEC聚合周期内收到的NACK进行聚合,得到冗余的FEC编码分组 $lout$,以满足大多数丢失者的需要,发送者以Multicast方式发送 $lout$,启动定时器 $T2$,如果在 $T2$ 未超时前收到NACK,且 $lin > lout$,那么发送 $lin - lout$ 个恢复分组。

(5)如果接收者在定时器 $T1$ 超时前没有收到足够的恢复分组,那么协议处理同步骤(2)。

6 试验评估

根据上述理论推导的公式,我们知道 M_{FEC} 表示任意数据成功发送到所有的组播接收者时需要发送的次数,它直接衡量了组播性能的优劣。影响 M 性能的关键参数如下:组播接收端的数量、CE和FE之间的丢失概率、原始数据块大小和FEC校验数据块大小。

我们的试验环境为2个CE加上多个FE构建的分布式交换机。试验时,假定线路丢失概率为0.01,我们通过不同的接收者数量、不同的FEC参数、传输大容量的3层转发表来验证RDCP的性能。

表1和表2表示了RDCP算法的试验结果,试验中采用了不同大小的数据块,其中 $k = 7, 20, 100$,而校验包参数 $h = 2, 7$ 。通过试验数据,主要得到两个结论:采用FEC算法的RDCP机制明显降低了数据丢失时的重传次数;当接收者数量很少时, $E[M] \approx n/k$,这种情况RDCP中将采用TCP协议保证可靠性传输,因为FEC校验将影响性能,而随着接收者数量的逐渐增加,算法的性能将越来越好。

表1 RDCP算法试验结果($p = 0.01, h = 2$)

接收者数量	传输次数 (未使用FEC)	传输次数 (FEC with $k=7$)	传输次数 (FEC with $k=20$)	传输次数 (FEC with $k=100$)
R=1	1	1.2	1.1	1
R=20	1.1	1.3	1.1	1
R=40	1.7	1.3	1.2	1.3
R=80	2.1	1.4	1.4	2
R=100	2.6	1.6	2	2.1

表2 RDCP算法试验结果($p = 0.01, h = 7$)

接收者数量	传输次数 (未使用FEC)	传输次数 (FEC with $k=7$)	传输次数 (FEC with $k=20$)	传输次数 (FEC with $k=100$)
R=1	1	2	1.4	1.1
R=20	1.1	2	1.4	1.1
R=40	1.7	2	1.4	1.1
R=80	2.1	2	1.4	1.1
R=100	2.6	2	1.4	1.7

由〈Target〉、〈Rule〉和〈Obligation〉元素构成,在策略决定点上做出评估,以生成和访问决策^[8]。由于可能发现多项策略对于访问决策来说都是可用的(并且由于单个策略可以包含多个 Rules),因此可以使用组合算法(Combining Algorithms)来将多个结果协调成单个决策。标准组合算法被定义为 Deny-Overrides, Permit-Overrides, First Applicable 和 Only-One Applicable 等。〈Target〉元素用于将被请求的资源关联到一个可用的〈Policy〉上。它包含的请求〈Subject〉、〈Resource〉或〈Action〉必须符合〈Policy〉或〈Rule〉的条件,以便资源更适用。〈Target〉包括一个内置的方案,用于有效地索引/查找策略。〈Rules〉提供了在一个〈Policy〉中测试相关属性的条件。可以使用任何数量的〈Rule〉元素,每一个〈Rule〉元素都可以产生一个真或者假的结果。将这些结果进行综合便产生了对该〈Policy〉的单一决策,可以是 Permit, Deny, Indeterminate 或者 NotApplicable 决策。

结束语 由于 SOA 环境分布性、异构性、动态性的特点,基于角色的访问控制已经不能适应 SOA 环境下访问控制的需求。基于属性的访问控制以主体、资源、环境的属性为基本访问控制粒度,可以更加灵活地适应开放的 SOA 环境。结合 XACML 和 SAML 实现的 ABAC 框架,由于安全基于 XML 标准,更符合 SOA 异构环境的特点。但是大量的 XML 文档会增加网络资源的开销,从而增加访问的延迟时间。而且随着访问决策粒度的细化,SAML 断言中将会包含大量的属性信息,使这种延迟更加明显。所以如何控制访问决策粒度和

找到最优化的访问控制策略,从而达到安全和性能的平衡,是未来需要研究的重点。

参 考 文 献

- [1] 邢少敏,周伯生. SOA 研究进展[J]. 计算机科学,2008,35(9): 13-20
- [2] Sandhu R S, Coyne E J, Feinstein H L, et al. Role-based access control models[J]. Computer, 1996, 29(2), 38-47
- [3] 尹刚,王怀民,滕猛. 基于角色的访问控制[J]. 计算机科学, 2002, 29(3): 69-71
- [4] Yuan E, Tong J. Attributed based access control (ABAC) for Web services[C]//IEEE International Conference on Web Services(ICWS'05). 2005
- [5] OASIS. eXtensible Access Control Markup Language(XACML) v2. 0[S]. OASIS Standard. <http://docs.oasis-open.org/xacml/2.0/>, 2005-5
- [6] OASIS. Security Assertion Markup Language (SAML) v2. 0 [S]. OASIS Standard. <http://docs.oasis-open.org/security/saml/v2.0/>, 2005-5
- [7] Steel C, Nagappan R, Lai R. 安全模式[M]. 陈秋萍,罗邓,袁国忠,等译. 北京:机械工业出版社,2006:204-233
- [8] 李晓峰,冯登国,徐震. 基于扩展 XACML 的策略管理[J]. 通信学报,2007(01):103-110
- [9] 罗海,安世全. 网络访问控制及对 RBAC 模型扩展的研究[J]. 重庆邮电大学学报:自然科学版,2008,20(6):714-718

(上接第 116 页)

结束语 本文描述了分布式交换机内部通信系统的设计和实现,并重点论述了基于 FEC 的可靠组播算法的实现。提出了一种自适应的 FEC 动态算法并计算了 RDCP 协议的性能。通过试验,验证了 RDCP 协议的有效性和可扩展性。

参 考 文 献

- [1] Yang L, et al. Forwarding and Control Element Separation (ForCES) Framework[S]. RFC3746. 2004
- [2] Floyd S, et al. A Reliable Multicast Framework For Lightweight Sessions and Application Level Framing[J]. IEEE/ACM Transactions on Networking, 1997, 5(6): 784-803

(上接第 130 页)

- [3] 王映辉,张世现,刘瑜. 基于可达矩阵的软件体系结构演化波及效应分析[J]. 软件学报,2004,15(8):1107-1115
- [4] Magee J, Kramer J. Dynamic structure in software architectures [C]//Proceedings of the 4th ACM SIGSOFT Symposium on Foundations of Software Engineering, 1996:3-14
- [5] Luckham D. Rapide: A language and toolset for simulation of distributed systems by partial orderings of events[C]//Proceedings of the DIMACS Workshop on Partial Order Methods in Verification, 1997:329-357
- [6] Allen R. A formal approach to software architectures[D]. Pittsburgh: Carnegie Mellon University, 1997
- [7] Kacem M H, Miladi M N, Jmaiel M, et al. Towards an UML profile for the description of software architecture[C]//The Conference on Component-Oriented Enterprise Applications,

- [3] Adamson B, et al. Negative-acknowledgment (NACK)-Oriented Reliable Multicast(NORM) Protocol[S]. RFC 3940. 2004
- [4] Rizzo L. Effective erasure codes for reliable computer communication protocols[J]. Computer communication Review, 1997, 27(2):24-36
- [5] McAuley A J. Reliable Broadband Communication Using a Burst Erasure Correcting Code[C]//Proc. ACM SIGCOMM 90. 1990: 287-306
- [6] Huitema C. The case for packet level FEC[C]//Proceedings of IFIP 5th International Workshop on Protocols for High Speed Networks. 1996:110-120
- [8] 马晓星,曹春,余萍. 基于图文法的动态软件体系结构支撑环境[J]. 软件学报,2008,19(8):1881-1892
- [9] Metayer D L. Describing software architecture styles using graph grammars[J]. IEEE Trans. on Software Engineering, 1998, 24(7):521-533
- [10] Kong J, Zhang K, Dong J, et al. A generative style-driven framework for software architecture design[C]// Proc. of the 29th Annual IEEE/NASA Software Engineering Workshop. 2005: 173-182
- [11] Loyall J, Kaplan S, Goering S. Abstraction and composition in D-specifications of concurrent systems[C]// Proc. 6th Int. Workshop Software Specification and Design. 1991:52-59
- [12] Bucchiarone A, Galeotti J P. Dynamic software architectures verification using DynAlloy[C]//Proceedings of GT-VMT. 2008