下一代互联网体系结构研究

林 闯 雷 蕾
(清华大学计算机科学与技术系 北京 100084)

摘 要 互联网发展到今天在很多方面已经无法满足用户的需求,面临着一系列问题,如安全性差、难以管理以及不可预测等.当前互联网的体系结构已经不能适应甚至阻碍互联网应用的进一步发展.文章对当前互联网体系结构存在的问题及面临的挑战进行了分析,并对国内外研究动态、在该领域的若干研究方向及提出的解决方案进行了总结评述.在此基础上重新审视互联网最初的基本设计原则是否符合当前和新兴网络应用环境的需求,并对互联网未来的发展进行了展望.

关键词 互联网体系结构;安全;服务质量;网络管理
中图法分类号 TP393

Research on Next Generation Internet Architecture

LIN Chuang LEI Lei
(*Department of Computer Science and Technology, Tsinghua University, Beijing* 100084)

Abstract Internet is in many ways not meeting the needs of its users. Today, it is not secure, hard to use, and unpredictable. Its architecture has created barriers rather than stimulants to its further development. This paper analyzes the problems and challenges of the current Internet architecture, summarizes several research areas of this field and introduces some proposed solutions. We rethink the design principles of Internet architecture, examining whether they still meet the current and new network environment demand, and analyze the development prospect of future Internet.

Keywords Internet architecture; security; QoS; network management

1 引 言

在 30 年前互联网刚刚问世之时,个人计算机和局域网还没有出现,光纤正处于实验室研究阶段,而商用的国际骨干线路最高速率大约只有 50Kb/s,最普遍的用户界面是基于文字的.在这 30 年,互联网无疑取得了巨大的成功,显示出其强大的生命力.然而发展到今天,互联网在很多方面已经无法满足用户的需求,面临着一系列问题和挑战,如安全性差、难以管理以及不可预测等.这主要是由于当前的网络环境与互联网最初设计之时已经出现了很大的不同,不仅用户的数目和网络的带宽都已经产生了巨大的增长,更重要的是新的网络和计算技术以及新的业务应用和商业模式的出现对互联网提出了新的要求.在这种情况下,人们开始认识到当前互联网的体系结构已经不能适应甚至正在阻碍互联网应用的进一步发展,并积极寻求一种方法能够解决目前互联网体系结构所处的尴尬境地.

与其它网络技术不同,互联网的体系结构很难进行准确统一的定义.相比之下,对互联网本身的定义则由其具体的组成协议(如 TCP/IP 和 SNMP

收稿日期:2007-02-02.本课题得到国家自然科学基金(60673187,90412012)和国家"九七三"重点基础研究发展规划前期研究专项基金(2006CB708301)资助.林 闯,男,1948 年生,教授,博士生导师,主要研究领域为计算机网络和系统性能模型及评价.雷 蕾,女,1980 年生,博士后,主要研究方向为网络性能评价、无线网络、互联网体系结构.E-mail: leilei@csnet1.cs.tsinghua.edu.cn.

等)构成,相对简单得多.然而,互联网的设计的确遵循若干基本原则,主要包括:(1)利用节点位置对其进行标识;(2)分层的协议栈;(3)在网络层部署唯一的网络互联协议(细腰);(4)基于数据报的无连接服务;(5)由端系统负责复杂的网络功能,如可靠传输和拥塞控制等,而连接这些端节点的网络只负责简单的传送功能(边缘论).这些体系结构的基本原则对于互联网今天的成功具有深刻的意义,然而现在却面临着越来越大的压力,需要我们在新的网络环境下重新对其进行审视.

目前业界对未来互联网发展的方式普遍持两种观点,可以归纳为革命与演进.前一种观点认为,当前阶段是互联网社会效用的变革时期,目前互联网技术创新的数量和速度都在减少,而30年前设计的体系结构限制了互联网应对不断出现的挑战和机遇的能力.因此,这种观点认为目前需要对互联网的体系结构进行一种"从零开始"(clean slate)的全新设计.另一种观点则认为,未来对互联网的创新应该以在现有互联网之上运行新业务和应用的形式进行.随着时间的推移,这些创新将逐渐对互联网进行改变.而在当前阶段希望对现有互联网的体系结构进行全面的更新是不现实的.最近出现的一种观点提出这种争论是不必要的①.首先,所谓"未来互联网"应该广义地包含体系结构中任何层次的创新,而不局限于某一个协议或层次.对未来互联网的研究既可能导致网络内部运行的协议和体系结构的改变,也同样可能导致现有网络之上运行新的业务和应用.所有这些协议和层次共同构成下一代互联网.其次,对未来互联网的研究思路和想法应该不受当前互联网的束缚,即"从零开始".但这并不意味着研究结果一定导致一个全新互联网的出现.换言之,"从零开始"是一个过程,而不是结果.最后,这两种观点可以进行融合,即探索现有互联网如何逐渐演进以支持新兴的重叠业务.

目前业界对未来互联网存在的形式也普遍持两种观点,可以归纳为纯化论和多元论[1].这两种不同的观点在对未来互联网体系结构的定义和评价方面都具有根本性的区别.前一种观点认为互联网体系结构应该是单一的,即在所有网元中都部署同一个具有核心地位的、唯一的协议(如目前的IP协议).互联网体系结构应该很少发生变化,新旧两种体系结构存在着较长的变革时间,而重叠(over-lay)网络只是互联网体系结构发生改变的一种手段,而不是结果.后一种观点则认为IP只不过是互联网整体系统中的一个组成部分,重叠网络则是另一种传送用户所需数据的方式,两者处于并列的地位.互联网体系结构应该是动态的和演进的,在任何时刻都被定义为各种重叠网络和协议的集合,而互联网体系结构的核心是支持多种重叠网络共存.在前一种观点下,互联网体系结构的灵活性是追求的目标,而后一种观点则更看重短期的性能改善.最近出现的一种观点认为,互联网体系结构当前最核心的作用是将"分布式系统与应用"和"网络"划分为两个领域,而正是这一点阻碍了互联网的发展[2].未来互联网将抛弃这种"端系统"与"路由器"的区分,而在这种情况下,纯化论与多元论之争可以解释为是通过对多种体系结构进行实验最终选出一种,还是允许多种体系结构同时存在并运行.

由上述分析可见,目前国际上对下一代互联网体系结构的研究还处于最初的摸索阶段,在未来互联网的发展方式和存在形式等方面都存在分歧.在这种情况下,对下一代互联网体系结构的研究可以让我们紧跟国际上的最新研究动态,重新审视互联网最初的基本设计原则是否符合当前和新兴网络环境的需求,从而明确未来互联网的发展目标,为互联网的下一步发展指明方向.

2 研究动态

目前对下一代互联网体系结构的研究主要有美国的GENI计划①,FIND项目②和IETF的IPng项目组等.

2.1 GENI,FIND和PlanetLab

GENI,即网络创新的全球环境(Global Environment for Network Innovation),是美国自然基金委(Natural Science Foundation,NSF)为促进对网络体系结构、业务和应用的研究而设计的一个共享全球设施.GENI旨在增加网络和分布式系统试验研究成果的质量和数量,加速这些成果的产业化进程,最终实现现有互联网的变革.

GENI的设施将与其它试验平台有所不同,它将同时支持研究和实用,有效填补实验室小规模试验和成熟技术的商用部署之间的裂缝.研究者可以使用GENI在大规模(全球)系统上评估新的网络系统,随着用户开始使用他们的系统而了解真实负载下的情况,并随着用户使用数量的增加而增强那些

① GENI: Global Environment for Network Innovations. http://www.geni.net/, August 2006

② FIND: Future Internet Network Design. http://find.isi.edu/

成功的系统.

GENI 与一些美国其它的网络计划有着重要的联系，如 PlanetLab[3] 和 FIND. 未来互联网计划 PlanetLab 是 GENI 的原型，它是用于开发下一代互联网技术的开放式全球性测试平台，目前在全球拥有 300 多个站点、600 多个节点. 在 GENI 计划的初始阶段，将整合 PlanetLab 的 30 个国家的 300 多个站点的资源. 相比较 PlanetLab，GENI 具有更好的发展前景. 首先，GENI 支持更为丰富的节点技术并且暴露低层次的网络连接，使更多的互联网实验成为可能. 同时 GENI 支持多种类型的服务，降低了研究者使用 GENI 的门槛.

GENI 与未来互联网设计项目 FIND（Future Internet Network Design）也有着重要联系. FIND 是 NSF 网络系统与技术计划（Networking Technology and Systems，NeTS）的一个长期项目，研究网络系统及相关技术. FIND 将可能成为 GENI 计划的重要组成部分，但是它不会涵盖 GENI 的所有研究领域，FIND 更侧重于未来互联网体系框架的研究.

2.2　IETF IPng 工作部

1990 年后期，IETF 开始了一项长期的工作，选择接替现行 IPv4 的协议. 此后，人们开展了许多工作，以解决 IPv4 地址的局限性，同时提供额外的功能. 1991 年 11 月，IETF 组织了路由选择和地址工作组（ROAD），以指导解决以上问题. 1992 年 9 月，ROAD 工作组提出了关于过渡性的和长期的解决方案建议，包括采用无类别域间路由（Classless InterDomain Routing，CIDR）方案以降低路由表增长的速度以及建议成立专门工作组以探索采用较大 Internet 地址的不同方案.

1993 年末，IETF 成立了 IPng 工作部以研究各种方案，并建议如何开展工作. 该工作部制订了 IPng 技术准则，并根据此准则来评价已经提出的各种方案. 在经过深入讨论之后，SIPP（Simple Internet Protocol Plus）工作组提供了一个经过修改的方案，IPng 工作部建议 IETF 将这个方案作为 IPng 的基础，称为 IPv6，并集中精力制定有关的文档. 自 1995 年末起，陆续发表了 IPv6 规范等一批技术文档.

IPv6 是 Internet 协议的一个新版本，其设计思想是对 IPv4 加以改进，而不是对其进行革命性的改造. 在 IPv4 中运行良好的功能在 IPv6 中都给予保留，而在 IPv4 中不能工作或很少使用的功能则被去掉或作为选项. 为适应实际应用的要求，在 IPv6 中增加了一些必要的新功能. IPv6 的主要特点包括：(1)经过扩展的地址和路由选择功能；(2)定义了任一成员（anycast）地址，用来标识一组接口；(3)简化的首部格式；(4)支持扩展首部和选项；(5)支持验证和隐私权；(6)支持自动配置；(7)服务质量能力等.

3　存在的问题及面临的挑战

现有互联网体系结构正在承受着来自内部和外部的双重压力，如图 1 所示. 首先分析来自内部的压力. 当前互联网所处的网络环境和社会效用与 30 年前相比都发生了巨大变化，某些在设计之初并不重要甚至并不需要的功能对于当前互联网来说却至关重要，最典型的包括网络安全，尤其是对拒绝服务攻击的抵御，以及端到端的服务质量保证等. 然而，现有互联网体系结构很难有效地支持这些功能. 在互联网的发展过程中，人们逐渐向互联网中加入一些并不符合互联网体系结构基本原则的模块，如多协议标记交换（Multiprotocol Label Switching，MPLS）、防火墙、网络地址转换器（Network Address Translator，NAT）以及其它各种中间件等. 这主要是为支持各种额外功能而采取的实用化解决方案，但却破坏了互联网体系结构的完整性和一致性. 例如，防火墙通过进行访问控制增强了网络的安全性，却导致网络无法系统地确定两个主机之间端到端的连接性. 其次分析来自外部的压力. 无线通信等物理层新技术的发展以及新的分布式应用和系统的出现分别从底层和上层为互联网体系结构带来挑战. 另一方面，互联网体系结构最主要的贡献是通过网络层使得处于底层的各种异构网络可以进行互联互通. 然而，当前存在的各种网络与互联网的关系不再如此单纯. 例如，各种话音网络（如移动电话、SMS 信令等）实际上只是通过网关接入互联网，而并没有在网络层运行 IP 协议，而无线传感器网络也很可能采取这种

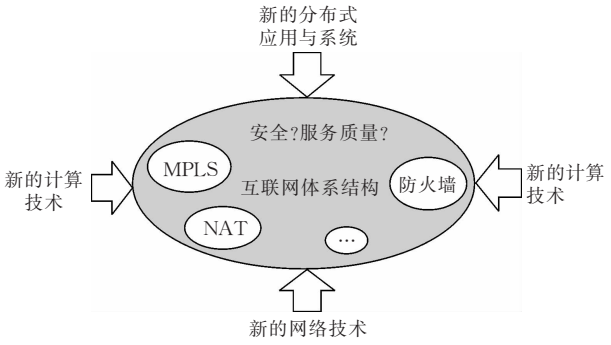

图 1　互联网体系结构面临的挑战

方式与互联网连接.因而,传统互联网很可能只是未来互联网的一个组成部分.在这种情况下,现有互联网体系结构也必然无法适应未来互联网的需求.

3.1 安全性与健壮性

当前互联网不存在网络安全的整体解决方案,这是迫切需要对下一代互联网体系结构进行研究的主要原因之一.当前互联网虽然存在许多安全机制,然而却没有一组规则能够将这些机制有效结合以实现网络的高度安全和可信.换言之,当前互联网虽然存在越来越多的"安全补丁",却不存在一个安全的体系结构.

对下一代互联网网络安全的研究不仅要针对数据的非法泄漏和破坏,还要解决攻击和故障等情况下网络的可用性和生存性等问题.未来的互联网必须能够提供类似电话系统的高可用性以支持"任务关键型"业务.另外,目前许多安全问题不是由互联网本身,而是由连接到互联网上的个人计算机造成的.因此,除了传统的网络安全问题,还要对操作系统和分布式系统的设计进行研究.最后,对安全问题的研究不仅只涉及到技术层面,而且还牵涉到人类行为与技术的相互作用.例如,对互联网用户身份更详细的确认可以便于对攻击的追踪,然而互联网匿名性的丢失以及对用户持续的监视却可能对互联网的使用方式产生巨大的影响.因此,互联网安全的重新设计需要从社会学的角度对特定设计产生的影响进行评估.

构建安全健壮的下一代互联网面临下列挑战:

(1)任何一组"行为良好"的主机之间应该能够任意进行具有高可靠性和可预测性的通信,而使恶意或遭到破坏的节点无法中断这种通信.用户应该可以达到或超过目前电话系统的可用性.

(2)应该实现跨层(cross-layer)的安全设计.这是由于用户的安全性和可靠性同时依赖于网络层和分布式应用的健壮性.

(3)要求用户提交身份信息一方面有助于阻止和威慑网络违法行为,另一方面却使得用户可能遭到不合法的监视和跟踪,因而需要在两者之间实现合理的平衡.

3.2 服务质量与网络可管性

传统互联网为面向非实时的数据通信而设计,多媒体应用的飞速发展对网络提出了不同于数据应用的服务质量(Quality of Service,QoS)要求,需要提供端到端的服务质量控制和保障.虽然从20世纪80年代末期开始,服务质量的控制和管理机制已经得到了广泛的研究,并提出了两种不同的互联网QoS体系结构:综合服务(Integrated Services,IntServ)[4]和区分服务(Different Services,DiffServ)[5],然而这些体系结构和机制都没有在当前的互联网络中成功应用.服务质量的有效提供从根本上依赖于对网络资源的有效管理和控制,而互联网服务质量研究的困境源于当前互联网较差的可管可控性.

互联网难于管理在体系结构上存在深层次的原因,仅仅依靠更加先进的网络管理工具不能从根本上解决问题.一方面,互联网"边缘论"和面向非连接的设计思想保障了高效的互通,但控制手段很薄弱,难以满足解决服务质量保障的需要.另一方面,互联网络发展至今,已成为一个庞大的非线性复杂系统,各种中间件的加入以及网络协议体系的庞杂导致互联网的"腰部"不断变粗,分布式控制平面中由于各种协议的决策逻辑交织在一起而产生的非线性作用使网络行为呈现出相当的复杂性并且难以预测.如何为用户提供满意的服务质量,设计有效的机制与算法来管理和控制这个已经具备相当规模并不断发展的大尺度复杂系统,在理论上和技术上仍是当前学术界的一个难题.

构建可管理和能够提供服务质量保障的下一代互联网面临下列挑战:

(1)应该在网络的关键部分增加服务质量控制机制使网络更可管;应该在网络中维护一定的状态信息,施加必要的控制,使网络具有某种程度面向连接的特性;应该在不同层次上可以对它进行监管.

(2)允许为每个网络配置对其性能、可靠性和策略的高层需求和目标;网络管理控制系统应该具有数据平面当前全局状态的准确信息;网络管理控制系统应该能够直接对数据平面进行控制,而将决策逻辑与分布式系统分离.

3.3 对网络新技术的支持

当前互联网的设计使其可以利用许多底层的网络技术.值得一提的是,互联网的出现早于局域网和光纤技术,却可以很好地与两者结合.然而,当前互联网也面临着一些挑战.

毫无疑问,从WiFi到无线传感器网络的各种无线技术是当前最主要的新技术.例如,笔记本电脑的销量在2003年已经超过了台式机的销量;2005年全球有超过2亿的手机用户,并且大约20%的手机具有数据功能.在未来的五年中,所有的手机都将成为互联网设备,这意味着网络应用与体系结构都需要支持移动性、定位以及移动终端有限的处理能力和无线链路有限的带宽等.很明显,未来互联网必须能够与无线通信很好地结合在一起.

支持移动技术的下一代互联网面临下列挑战：

(1) 未来互联网必须将支持节点移动性作为首要目标. 节点必须能够改变它们在互联网上的接入点.

(2) 未来互联网必须为应用提供有效的方式以发现无线链路变化的特性并进行适应.

(3) 未来互联网(或者在互联网上运行的业务)必须使邻近的节点能够互相发现.

(4) 未来互联网必须开发能够在互联网环境下很好工作的无线技术，具有安全健壮性，可以进行资源控制并能够与有线网络互通.

另一种正在经历巨大变革的网络技术是底层的光传送技术，这与 1960 年早期电子技术的变革类似. 光子集成电路(Photonic Integrated Circuit，PIC)允许不断增加复杂度的光电路与功能被置于一个单独的芯片之上，从而实现一种只靠电子技术无法达到的网络和通信模式. PIC 技术的成熟，将使得可重构的更加灵活的具有更高容量和更低造价的网络成为可能. 这将使网络产生一系列变革，包括从环状网到网状网的变革，从固定带宽分配到可调谐发送和接收的变革，从不具备光缓存的网络到具有智能控制平面和足够光缓存网络的变革，从将光纤带宽作为固定电路到允许光纤带宽被动态访问和使用的变革等.

利用新兴光传送技术的下一代互联网面临下列挑战：

(1) 未来互联网必须使用户能够利用底层光传送网所提供的新能力，包括通过跨层分析实现更好的可靠性，通过跨层流量工程实现更好的可预测性以及为用户提供更好的性能等.

(2) 未来互联网必须允许动态重构的光节点实现电层对光纤带宽的动态使用.

(3) 未来互联网必须包括控制和管理软件，使得由动态可重构节点组成的网络能够作为一个稳定的网络层运行.

3.4　对新计算技术的支持

互联网的发展主要在个人计算机(PC)时代，因而主要支持这种计算方式. 然而，10 年之后的计算方式会与今天产生很大的不同. 在计算机非常昂贵的时代，人们采用“多人一机”的方式共享计算资源. 随着计算机价格的降低，在 PC 机时代人们采用“一人一机”的方式. 虽然这种方式已经持续了将近 20 年，然而随着计算机价格的持续下降，人们将进入一个新的时代，即“一人多机”. 这个时代已经开始来临，并将加快步伐.

10 年之后，大部分计算机将与目前的 PC 机有很大不同，它们将是置于建筑，汽车和环境中的小的传感器，用来监视健康、交通、天气、污染、科学试验和军事行动等. 现在这些计算机的原形组成了专用的传感器网络，而并未直接连接到互联网中. 一个传感器网络可以连接到互联网中以实现远端访问，但目前互联网并不能解决这些计算机的特殊需求. 10 年之后这些新型计算机将成为主流，而互联网必须考虑如何支持这种计算模式的需求.

支持新计算技术的下一代互联网面临下列挑战：

(1) 未来互联网必须考虑支持未来计算设备的专用设备网络，这意味着对其体系结构提出一系列要求，如不间断的连接、数据驱动的通信、对位置相关应用的支持以及区分应用的性能等.

(2) 未来互联网应该使得一个特定的传感器应用可以跨过互联网的核心，将属于一个传感器网络的两个部分连接起来.

3.5　新的分布式应用和系统

新的网络和计算技术为向用户发布新一代分布式数据业务提供了前所未有的机遇. 多媒体用户到用户通信已经成为普遍，其中每个用户既是内容发布者又是内容消费者，信息很容易获取却又受到数字产权的保护；端系统联合的能力实现了全新的并行计算和通信模式等.

面向新业务和应用的下一代互联网面临如下挑战：

(1) 未来互联网必须定义一组新的抽象，以安全灵活地管理复杂的分布式业务.

(2) 未来互联网必须指明应该向应用设计者透露的监视和控制信息，并包含相应的规范和接口. 例如，未来互联网可能在特定点之间透露一些吞吐量和时延的信息.

(3) 未来互联网应该包含一个不同名字空间的统一设计. 这个设计应该在对不同设计方案进行分析的基础上得到. 它必须对从分组到应用的不同层需要何种标识给出答案.

4　研究方向及若干解决方案

目前对下一代互联网体系结构的研究方向主要集中在地址与标识、网络安全、服务质量和网络管理等方面. 下面将对一些近年来各个研究方向提出的较为典型的下一代互联网体系结构解决方案进行介绍. 由于篇幅有限，本文不可能对每个领域的研究成果进行十分全面和深入的分析，有兴趣的读者可以在某一领域展开进一步的研究.

4.1 地址与标识

4.1.1 地址耗尽问题及解决方案——IPv6 和 NAT 扩展

IP 地址的耗尽是互联网所面临的一大紧迫问题,为此 IETF 分别提出了长期和短期解决方案. 短期解决方案主要是 CIDR,而长期解决方案则是具有更多地址的新互联网协议 IPv6. 然而,由于担心在长期方案部署到位之前 CIDR 无法满足 IP 地址的紧迫需求,IETF 又提出了一种新的方案——网络地址转换器(NAT)通过地址重利用来更好地解决 IP 地址耗尽问题.

NAT 在 1994 年作为一种解决 IP 地址耗尽问题的短期方案被提出. 同年,IETF 选择 IPv6 作为下一代互联网协议. 由于 NAT 不需要对主机和路由器做任何修改,非常便于部署,并且能够很好地解决 IP 地址缺乏问题,因而得到了广泛的应用. 相比之下,IPv6 的部署进程则非常慢,一个主要原因在于从 IPv4 到 IPv6 的转变非常复杂. 由于无法将所有的 IPv4 主机和路由器都同时由 IPv6 替代,因而必须解决 IPv4 和 IPv6 的互联和共存问题. 由于 IPv6 所遇到的困难以及 NAT 的广泛应用,一些文章开始提出将 NAT 作为一个长期方案与 IPv4 共存,并在此基础上对主机和 NAT 设备进行扩展,解决 NAT 目前存在的问题,从而形成下一代互联网体系结构. 根据这种思路提出的体系结构被称为 NAT 扩展的体系结构,典型的有 IPNL[6],IPv4+4[7]等.

无论何种 NAT 扩展的体系结构,其主要目的在于解决目前 NAT 技术存在的重要缺陷. NAT 方案在 stub 域的边界加入 NAT 设备,该设备存储一个本地地址和全局唯一地址的转换表. Stub 域中的 IP 地址不是全局唯一的,它们可以与其它 stub 域的 IP 地址相同,从而解决 IP 地址缺乏问题. 当一个 stub 域中的主机与域外的主机通信时,NAT 设备根据地址转换表将分组的源地址由本地地址替换为相应的全局唯一地址. 传统 NAT 只能完成 stub 域内主机到域外主机的单向通信,一些改进的 NAT 版本能够完成由 stub 域内主机发起的双向通信. 但对于一些在 IP 分组头之外携带 IP 地址的业务来说,必须依靠应用层网关(Application Level Gateway,ALG)对分组进行额外的处理才能进行. 尤其是当 stub 域外主机向域内主机发起会话时,必须要通过 DNS-ALG 根据域内主机的主机名得到它相应的全局唯一地址. 因此,要通过 NAT 实现目前新兴的 P2P 应用较为困难. 总之,NAT 使网络失去了端到端的透明性,使新业务的实现复杂化,必须进行改进才能够作为下一代互联网体系结构的可选方案. 下面介绍一种典型的 NAT 扩展体系结构——IPNL(IP Next Layer).

IPNL 只对主机和 NAT 设备进行扩展而保持 IPv4 路由器和支持的协议不变. IPNL 在 IP 层和传送层之间插入一个新的协议层,如图 2 所示. 扩展的 NAT 设备被称作 nl-router,它们将分组逐域(realm)路由到目的地. 连接私有域和公共域的 nl-router 被称作 frontdoor,一个 frontdoor 可以与多个私有域相连,每个私有域具有一个局部唯一的域编号. 这样,每个私有域都可以通过与其相连的 frontdoor 的全局 IP 地址和其域编号的组合进行标识. 私有域上的每个主机可以通过被称作 IPNL 地址的(frontdoor 地址,域编号,主机地址)进行标识,这里的"主机地址"是本地 IP 地址,不必全局唯一,因而节省了大量的地址空间. IPNL 地址不是主机的长期标识符,它只作为主机的位置标识而可以经常变化. IPNL 使用 FQDN(Fully Qualified Domain Name)作为主机长期标识符. Nl-router 既可以通过 FQDN 也可以通过 IPNL 地址对分组进行路由. 在最初的分组交互中双方在分组头中使用 FQDN 标识对方. 此后,双方得知自己以及对方的 IPNL 地址并使用该地址进行后续的通信.

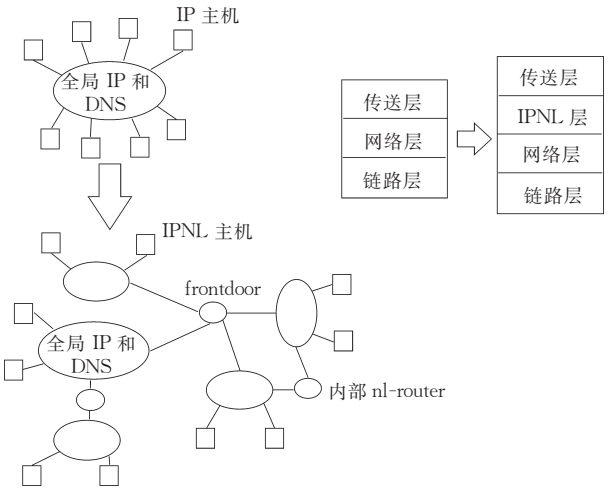

图 2 IPNL 拓扑

4.1.2 主机移动性问题及解决方案——地址与标识分开

在当前的互联网中,主机通过 IP 地址进行标识,而 IP 地址依赖于该主机的拓扑位置. 因此,IP 地址在语义上被进行了重载——一个 IP 地址描述了连接到网络的一个网络接口的拓扑位置,同时该 IP 地址也用来标识该接口所属的主机. 简而言之,IP 地址既标识了主机又标识了拓扑位置. 当主机具有移动性时,这种方式出现了内在矛盾. 一方面,IP

地址作为主机标识不应被改变；而另一方面，由于 IP 地址描述拓扑位置，因而当主机在网络中的位置出现改变时它必须随之改变.

移动 IP(Mobile IP)对该问题的解决方案是为每个节点提供一个对应固定的归属位置的“归属地址”. 当该节点在归属地时，归属地址同时为节点提供标识和固定位置. 当节点移动到其它位置时，节点当前的位置信息以暂时地址(care-of address)的形式提供以供路由使用.

另一种解决该问题的方法是将标识和位置功能分离. 其实，在 4.1.1 小节介绍的 IPNL 方案中就采用了将节点标识和位置信息分开的思路，即用 FQDN 作为主机的固定标识，而用 IPNL 地址作为主机当前的位置信息. 主机标识协议(Host Identity Protocol，HIP)[8-10]是一种采用该思路的典型方案. 它不仅介绍了一种分离位置与主机标识信息的方法，而且与 IPNL 不同的是，它为主机标识引入了一个新的加密命名空间，而 IP 地址则继续用来作为位置信息进行分组路由.

HIP 在网络层和传输层之间加入一个主机标识层，如图 3 所示. 它引入一个新的名字空间——主机标识. 主机标识本质上是一个公共-私有密钥对中的公共密钥. 公共密钥标识唯一私有密钥的拥有者. 拥有私有密钥的主机可以直接证明他拥有用于在网络中标识它的公共密钥. 这种分离也提供了一种安全的处理移动性和多归属(multi-homing)的方式. 与 IPNL 相比，HIP 具有更高的安全性.

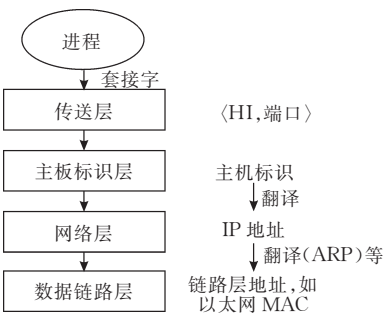

图 3　HIP 体系结构

4.1.3　以资源为中心问题及解决方案

随着 WWW(World Wide Web)的出现，互联网的主要用途成为内容的发布，包括网页的发布以及向用户应用发布声音和图像等. 然而，互联网并不是为可扩展的内容发布而设计的. 为了增强内容传递的可扩展性以支撑这些需求，越来越多的内容提供给网站在世界各地进行复制.

主机移动性和内容复制改变了分布式业务的域名解析模式. 计算机系统中的域名解析在传统意义上是指将人可识别的字符串映射为机器可识别的标识. 这种映射通常是静态的和一对一的. 例如，当访问一个远端互联网主机时，主机的域名必须首先被翻译成一个 IP 地址，然后再使用此 IP 地址向远端主机请求服务. 然而现在当用户使用域名访问资源和服务提供商时，这些资源的域名不再只对应一个单独的主机. 一个命名的资源可以从许多被复制的服务提供商或者从一组竞争的服务提供商中得到. 另外，移动性和不同的网络性能意味着客户可以通过当前状况隐含对一个命名资源不同版本的要求. 例如，一个使用无线 PDA 的客户可以与一个使用 T3 链路工作站的客户使用相同域名的资源(如 cnn.com). 然而，这两个客户隐含着对同一命名资源不同版本的要求.

综上所述，当前互联网的主要应用和业务感兴趣的不是一个资源的地址，而是该资源本身. 由于这些资源可以被复制和移动，因而传统 DNS 将资源域名解析为地址的方法存在着弊端. 为此，人们提出了一系列新的体系结构解决该问题，如 TRIAD[11-12]，Active Name[13]和 INS[14]等.

TRIAD 主要解决内容路由(content routing)问题，即如何为客户找到一个距离他最近的内容复制站点. 在基于 DNS 的内容路由中，客户首先向 DNS 服务器发起请求从而得到一个与所需内容相关的特殊域名服务器的地址. 然后，客户再向该域名服务器查询所需内容的最佳内容服务器的地址. 最后，客户向最佳内容服务器请求所需的内容. 这个过程的延时很大.

TRIAD 将域名解析与路由过程合为一体. 它将防火墙、网关路由器和 BGP 级路由器升级为内容路由器(content router)，使它们同时作为传统 IP 路由器和域名服务器，不仅进行传统的 IP 路由，还进行基于名称的路由(name-based routing). 这些内容路由器形成了一层重叠在网络层之上的内容层(content Layer)，而大量的普通 IP 路由器并不知道内容层的存在.

TRIAD 通过 DRP(Directory Relay Protocol)进行域名查找. 客户首先将包含待查找内容 URL 的请求发送给最近的内容路由器(与发送给预配置的 DNS 服务器类似). 内容路由器包含“域名到下一跳”的映射表，从而将请求转发给下一跳内容路由器. 这样，就可以逐跳将查找请求向最佳的内容服务器进行转发. 当与最佳内容服务器相邻的内容路由器接收到请求时，它产生一个包含最佳内容服务器地址的回复，并将其沿着与请求相反的路径返回给

客户. TRIAD 内容路由器中“域名到下一跳”映射表是通过 NBRP(Named-based Routing Protocol)路由协议形成的. NBRP 与 BGP 类似,但它在内容路由器之间发布的是名称前缀可达性信息.

与 IPNL 和 HIP 不同,TRIAD 将内容与数据而不是它们所在的主机作为命名的对象,并将其解析为主机地址. 从这种意义上来说,TRIAD 也支持标识与地址的分离,从而能够支持移动业务. 另外,TRIAD 也是一种 NAT 扩展的体系结构.

4.1.4 间接定向问题及解决方案

当前的互联网是基于点对点通信的抽象模型建立的. 这个简化的抽象模型是当前互联网具有可扩展性和有效性的一个重要原因. 然而,随着互联网逐渐演化为一个全球通信平台,对多种通信模型的需求越来越大,如组播、任意播和端节点移动性等. 然而,尽管经过多年的研究,这些业务还未在互联网中得到部署. 其主要原因在于当前互联网体系结构所假设的一个发送端和一个在固定地点的接收端的点到点通信模型并不适合于这些业务. 为此,目前的解决方法通常采用一种简单有效的技术——间接定向(Indirection).

所谓间接定向,是指在源节点和目的节点之间通过一个物理或逻辑的间接定向点对分组进行中继,源节点与中继节点而不是目的节点进行直接通信,从而利用间接定向点对目的节点的位置和数目进行屏蔽. 例如,移动 IP 利用归属代理屏蔽目的节点的移动性,而 IP 组播利用逻辑间接定向点(地址)隐藏目的节点的数目和位置. 换言之,间接定向是指由用户对网络中分组的路径进行控制. 互联网在最初的设计中并未考虑需要具备间接定向能力,但随着互联网的发展,越来越多的应用需要间接定向功能的支持,如移动 IP、NAT、代理、路由选择以及各种重叠网络.

虽然利用间接定向可以实现上述业务,但在 IP 层对间接定向进行有效的部署却非常困难. 松散源路由是解决间接定向问题的一种方法,但由于其难以控制和管理,因而大多数 ISP 都不采用该方法. 为此,一些文章提出了一系列新的体系结构以解决间接定向问题,如 i3(Internet Indirection Infrastructure)[15] 和 SelNet[16-17] 等.

Selnet 在数据链路层和网络层之间增加一个虚拟链路层. 它采用标记交换的方式在本层转发分组. Selnet 包含两部分:XRP(eXtensible Resolution Protocol)和 SAPF(Simple Active Packet Format). 在 Selnet 中,XRP 替代 ARP 协议作为接入和配置 Selnet 资源的统一接口. SAPF 是分组头的格式,它包含数据链路层地址和选择器(selector)地址. Selnet 中的选择器类似于路径或流 ID,用来作为分组转发的依据. Selnet 与前面提到所有的体系结构最大的不同之处在于,它是在 IP 层之下而不是之上增加一个新的协议层.

4.2 网络安全

由于计算机网络具有开放性和互联性等特点,致使网络容易受到黑客、病毒和恶意软件等不轨行为的攻击,所以网络上的信息安全和保密是一个至关重要的问题. 互联网最初的体系结构是面向善意和可信的网络环境而设计的,这种假设显然不适用于当前连接全球数以百万计网民、计算机和企业的复杂网络环境. 目前互联网在安全问题上的研究困境究其根源在于僵化和脆弱的体系结构. 因此,对下一代互联网体系结构的设计必须把网络安全与健壮性作为重要的要求进行考虑.

下一代互联网安全体系结构的设计存在着巨大的挑战. 无论是在网络协议栈中增加一个“安全层”,还是目前在各个协议层中分别设计各自独立的安全机制的方式都存在问题. 前者使得网络中只有某一层具有安全机制,无法阻止攻击者对网络中其它层的攻击;后者需要对当前存在的各个协议进行安全性设计,使得协议复杂化以及多个协议的交互存在不可预测性,而系统的复杂化又进一步增加了网络的不安全因素.

保证网络安全的手段大体可以分为四种:防御、检测恢复、可生存性以及威慑. 目前广泛采用的网络安全技术主要有加密和访问控制两种防御机制. 加密是指利用数学方法来重新组织数据,使得除合法接收者之外的其他人恢复原始数据(明文)或读懂变化后的数据(密文)都非常困难. 访问控制用来防止对资源的未授权使用,分为自主访问控制和强制访问控制等. 通过防御手段虽然能够阻止多数已知类型的黑客攻击和病毒等,却无法完全消除网络中现在和将来潜在的所有威胁,缺乏灵活性.

目前业界和学术界对下一代互联网的安全体系结构没有统一意见,甚至对保证网络安全应采用的主要手段也不确定. IETF 提出 IPv6 中的安全体系结构 IPSecurity(IPSec)①,增强了 IP 层的安全性,但却仍然无法确保整个网络的安全. 目前对拒绝服务(Denial-of-Service,DoS)[18] 攻击的防御是网络安

① Friedl. An Illustrated Guide to IPsec. http://www.unixwiz.net/techtips/iguide-ipsec.html

全体系结构研究的一个热点，这些研究大多提出在网络中的某些关键点上增加对数据包的控制能力. 然而，虽然防御 DoS 攻击是目前互联网面临的一个重要挑战，却不是互联网安全问题的全部. 我们认为，对互联网安全体系结构的研究应该跳出传统的研究思路，借鉴系统可信性的概念，将传统孤立的研究内容融合到网络可信这一目标下，构建可信网络体系结构，面向用户提供系统的安全服务[19-20]. 下面将分别对这三类安全体系结构进行介绍.

4.2.1　IPv6 安全体系结构——IPSec

IP 协议在设计之初没有充分考虑其安全性. IETF 在 IPv6 中提出了全新的网络安全体系结构 IPSec. IPSec 定义了在 IP 层使用的安全性服务，其功能包括对网络单元的访问控制、数据源身份验证、提供无连接完整性服务、重放数据报的监测和拒绝、使用加密来提供保密性. IPSec 能够为上层的数据提供透明保护，解决了 TCP/IP 协议的各种安全漏洞，对 IPv4 和 IPv6 都可用. 目前 IPSec 在网络安全领域得到了广泛应用，尤其被广泛用于构建 VPN.

IPSec 体系结构采用模块化设计，主要由三个协议：AH（验证头）、ESP（封装安全载荷）和 IKE（Internet 密钥交换）来提供认证、数据完整、机密性三种保护形式，如图 4 所示.

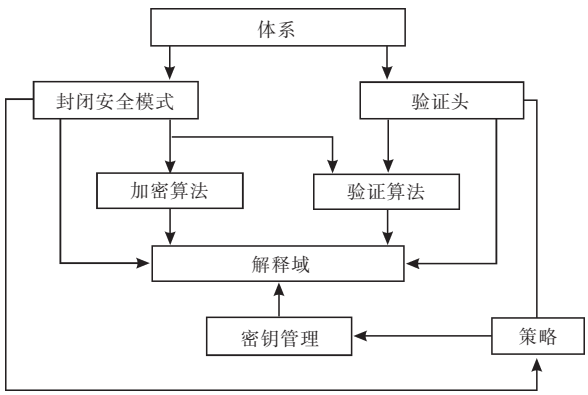

图 4　IPSec 体系结构

IPSec 从协议上保证了数据传输的安全性. 它提供网络层端到端的安全服务，通信双方为实现通信保护必须维护所需的安全策略和安全联盟，任何第三方如果没有关于此次通信的安全参数，不可能伪造、偷窥通信数据. 尽管 IPSec 具有这样的优点，但并不是说它可以确保整个系统的安全了. IPSec 只能保证网络层的安全，而安全包括各个层次和多方面的问题，网络系统还可能遭到以下方面的攻击：

（1）来自网络层以上应用层发起的攻击，如利用系统缓冲区溢出或木马的方法；

（2）由于加密解密都需要花费大量的 CPU 时间，黑客通过发送大量的貌似正确的随意填充的加密数据包给目的主机，目的主机花费大量的 CPU 时间来进行解密运算而不能响应其它请求，从而造成拒绝服务攻击；

（3）黑客窃取了通信双方协商的密钥，冒充合法用户向目标主机发动攻击，这种攻击将具有极大的隐蔽性，也是致命的.

4.2.2　防御 DoS 攻击的互联网安全体系结构

目前网络上存在的攻击类型根据产生的后果大体可以分为两类：破坏和窃取型攻击以及资源消耗型攻击. 前者指攻击者非法入侵受害者（主机或路由器）系统，直接对受害者的软硬件进行破坏使其不能正常工作，或者修改和盗取受害者的重要信息（帐号或密码等）. 后者是指攻击者不直接对受害者进行破坏，却通过大量消耗受害者的资源使其无法正常工作. 一种常见的资源消耗型攻击是拒绝服务（DoS）攻击. DoS 攻击的目标包括客户、服务器、防火墙、路由器和通信链路等；它消耗的资源包括处理能力、内存、带宽等. 除了以上两种攻击类型外，还有一种攻击方式将两者结合在一起，即首先通过非法入侵大量其它主机以达到对这些主机进行控制的目的，再利用这些受控主机组成“僵尸网络”对受害者进行资源消耗型攻击. 这种攻击被称为分布式拒绝服务（Distributed Denial-of-Service，DDoS）攻击.

目前没有十分有效的安全技术以应对资源消耗型攻击，即 DoS 和 DDoS 等. 这在很大程度上是由于互联网体系结构内在的开放性和普遍连通性造成的. 因此，在不对互联网体系结构进行重大改变的情况下，对 DoS 攻击进行彻底防御是非常困难的[26]. 目前，已经提出一系列解决方案，对互联网体系结构进行不同程度的改变，从而在不同程度上减轻 DoS 攻击带来的危害.

（1）基于业务过滤的防御方案

前面提到的访问控制是目前互联网中一种主要的安全技术，可以用来应对破坏和窃取型攻击. 但将访问控制技术用于对 DoS 攻击的防御中却存在着问题. 在应对破坏和窃取型攻击时，访问控制技术一般在目的端进行业务过滤，但这种方式却不适于 DoS，原因主要有以下 3 点：① 对于那些消耗带宽资源的 DoS 攻击来说，在目的端进行业务过滤无法解决问题；② 由于攻击源可以伪造攻击业务的源地址，因而受害者无法确定攻击源从而建立有效的过滤规则；③ DDoS 攻击可能包含几百万攻击源，而路由器受到造价和体积的约束，每个网络接口最多有几万个过滤器，因而无法对所有的攻击源进行过滤.

基于以上三点原因,应对 DoS 攻击最好的方法是在离攻击源最近的路由器设置过滤器,从而可以利用足够多的过滤器阻止攻击业务.然而,由于攻击者可能伪造源地址,以及在当前互联网中不存在确定离攻击源最近路由器的机制,因而受害者无法确定过滤器的最佳设置位置.针对这一问题,文献[21]提出一种主动互联网业务过滤(AITF)机制,通过使用一种变化的 IP 路由记录(Route Record,RR)机制确定每条攻击路径上的最后一个可信点,并在此点设置过滤器阻止攻击业务.文献[22]则提出一种为服务器子网设置保护边界的思路,在服务器子网边界设置控制点对去往服务器的数据进行 IP 封装并发往一个解封点.当一个服务器受到攻击时,解封点可以在攻击业务进行封装的封装点设置过滤器.

(2)基于资源共享的防御方案

与业务过滤的思路不同,资源共享方案不试图区分合法和攻击业务,而是通过为所有用户更优化地分配资源以减轻 DoS 攻击的影响.例如,通过公平队列等可以减轻对带宽进行消耗的攻击,而通过操作系统的 CPU 调度和内存分配可以保护主机的资源.

DoS 攻击与 flash crowds 类似,拥塞不是由单个网络流或业务的普遍增长造成的,而是由一组广泛分布在很多流中的分组,或者说聚合造成的.因而,传统的基于单个流的拥塞控制机制不适用于 DoS 攻击.文献[23]提出一种基于聚合的拥塞控制(ACC)机制对以聚合为颗粒度的业务进行流量控制.ACC 机制介于传统单个流拥塞控制和主动队列管理之间,进行聚合的检测和控制.

(3)基于网络能力的防御方案

基于能力的通信分两步进行:能力建立和数据传输[24].这两步都涉及到发送端、接收端以及两者之间路径上的一组认证节点.能力建立阶段与 TCP 连接建立类似:① 发送端向接收端发送一个能力请求;② 路径上每个认证节点为该请求分组打上一个特定标记并将其转发,所有标记一起构成"能力";③ 接收端将能力返回给发送端.在数据传输节点,发送端在给接收端的所有数据包中包含能力.每个认证节点对能力中自己创建的标记进行认证,如果正确则将该数据包转发.

与传统数据过滤不同的是,能力由于使用智能标记技术从而无须在网络中加入每个连接的状态.即认证节点无须知道一个分组的发送端和接收端就可以对其能力进行认证.文献[25]提出的 TVA (Traffic Validation Architecture)就是一种基于网络能力的互联网体系结构.

4.2.3 可信网络体系结构

通过上面对互联网安全体系结构的讨论可以看出,互联网在设计之初对安全问题考虑不足,是产生当前网络众多脆弱性的一个重要因素.一段时间以来网络体系结构的研究过度集中于如何提高数据传输的效率,形成了如今一个核心简单、边缘复杂的互联网体系模型.这种核心网络的简单性方便了新业务的部署,但同时造成核心网络对业务过于透明,基本不存在特定于应用的运行模式,导致难以检测到应用业务层面出现的问题,更难将攻击行为和新业务区分开来①.

此外,网络安全已经超出传统信息安全的可用性、完整性和私密性的内涵,服务的安全作为一个整体属性为用户所感知的趋势日益凸现.然而目前的许多网络安全设计很少触及体系结构的核心内容,大多是单一的防御、单一的信息安全和打补丁附加的机制,遵从"堵漏洞、作高墙、防外攻"的建设样式,以共享信息资源为中心在外围对非法用户和越权访问进行封堵,以达到防止外部攻击的目的.在攻击方式出现复合交织的趋势下,当前的安全系统将变得越来越臃肿,严重地降低了网络性能,甚至破坏了系统设计开放性、简单性的原则.并且,安全系统自身在设计、实施和管理各个环节上也不可避免地存在着脆弱性,严重影响了其功效的发挥.因此基于这些附加的、被动防御的安全机制上的网络安全是不可信的.另一方面,网络安全研究的理念已经从被动防御转向了积极防御,需要从访问源端就开始进行安全分析,尽可能地将不信任的访问操作控制在源端[27].

通过上面的分析,我们提出通过可信网络的研究重新审视互联网的体系结构设计,减少系统脆弱性并提供系统的安全服务.我们认为未来的网络安全供给模式应该是提供系统的安全服务,一方面安全应成为嵌入到网络内部的一种服务,同时要从体系结构的设计上保障网络服务的安全持续.这也是我们研究可信网络的重要目标.

尽管人们提出可信系统的概念已经有一段历史,增强计算机终端可信性的可信计算也是近年来的一个研究热点,但是国际上对可信网络的探索刚刚开始,基本概念和科学问题的认识还不深入.目前国际上对可信性比较有代表性的阐述主要有:ISO/IEC15408 标准中指出,一个可信的组件、操作或过

① Clark David. A new vision for network architecture. http://www.isi.edu/know-plane/DOCS/DDC_knowledge-Plane_3.pdf

程的行为在任意操作条件下是可预测的，并能很好地抵抗应用程序软件、病毒以及一定物理干扰所造成的破坏；微软公司的比尔盖茨认为可信计算是一种可以随时获得的可靠安全的计算，并包括人类信任计算机的程度，就像使用电力系统、电话那样自由、安全①；Algridas 和 Laprie 等人则将可信性表述为系统提供的服务可以被论证为可信任的，系统能够避免出现不能接受的频繁或严重的服务失效[28].

我们认为一个可信的网络应该是网络系统的行为及其结果是可以预期的，能够做到行为状态可监测、行为结果可评估、异常行为可控制. 具体而言，网络的可信性应该包括一组属性，从用户的角度需要保障服务的安全性和可生存性，从设计的角度则需要提供网络的可控性. 不同于安全性、可生存性和可控性在传统意义上分散、孤立的概念内涵，可信网络将在网络可信的目标下融合这三个基本属性，围绕网络组件间信任的维护和行为控制形成一个有机整体.

如图 5 所示，信任信息的维护过程可以分为信任信息输入、信任信息处理和信任等级或策略输出三个部分. 信任信息采集提供具体的输入方式，主要包括：集中式安全检测，即通过在网络中设置专门的服务器，对某个范围内的网络节点进行脆弱性检测等信任信息的采集，其特点是网络结构简单，但是可扩展性较分布式节点自检方式差些；分布式节点自检，即将部分监测功能交由网络节点中的代理完成，网络只负责接收检测结果，其特点是工作效率高，但是控制机制较为复杂；第三方通告，即由于不能直接对被测节点进行检测等原因，而间接地获得有关信息.

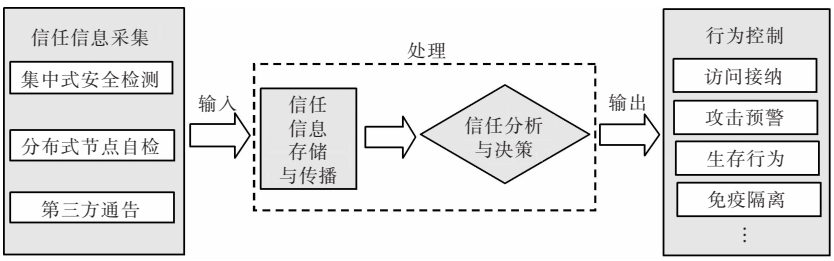

图 5　可信网络的信任维护与行为控制

信任信息经过存储、传播和分析后，通过信任等级和策略输出用于驱动和协调需要采取的行为控制. 典型的行为控制方式有：访问控制，即开放或禁止网络节点对被防护网络资源的全部或部分访问权限，从而能够对抗那些具有传播性的网络攻击；攻击预警，即向被监控对象通知其潜在的易于被攻击和破坏的脆弱性，并在网络上发布可信性评估结果，报告正在遭受破坏的节点或服务；生存行为，即在网络设施上调度服务资源，根据系统工作状态进行服务能力的自适应调整以及故障的恢复等；免疫隔离，即根据被保护对象可信性的分析结果，提供到网络不同级别的接纳服务. 不同于访问控制主要针对的是防护区域外具有攻击和破坏性的节点及行为，免疫隔离更多地是在攻击和破坏行为出现前主动对防护区域内的设施进行处理.

可信网络三个基本属性的紧密联系体现在：通过可信网络安全体系结构设计，改变传统打补丁、附加的安全供给模式，降低整个信任信息维护链体系结构设计上的脆弱性，支持多样的信任信息采集方式，保障信任信息的可靠有效传播，并能有效协同各种行为控制方式，使其能在可信的目标下得到融合；通过可生存性设计，在系统脆弱性不可避免以及攻击和破坏行为客观存在的状况下，提供资源调度等提高服务生存性的行为控制，提高包括安全服务在内等关键服务的持续能力；通过可控性设计，完成对网络节点的监测以及信任信息的采集，根据信任分析决策的结果实施具体的访问接纳和攻击预警等行为控制手段，从而建立起内在关联的异常行为控制体系，结束当前安全系统分散孤立的局面，全面提升对恶意攻击和非恶意破坏行为的对抗能力. 当然，还需要建立网络与用户行为的可信模型，为信任信息的分析决策、行为控制方式的选择及实施效果的评估提供判据.

可信网络体系结构研究必须充分认识到网络的复杂异构性，从系统的角度保障安全服务的一致性. 现实的互联网涵盖了不同类型的传输技术，如有线和无线，存在着不同属性的业务，如数据、图像、语音和视频. 这些差异可能会形成对网络可信性威胁因素的不同关注，然而来自用户的安全服务要求却是明确的，并不会因某个业务需要跨越几个无线和有线的传输路径而改变，当然也不会关心提供安全服务的具体技术细节. 作为网络研究最有价值的经验，

① Bill Gates. Trustworthy computing. http://www.wired.com/news/business/0,1367,49826,00.html

开放系统互联应该是可信网络体系结构研究需要遵从的一个原则.

图 6 给出了可信网络的一种可能的体系结构模型. 数据传输平面负责承载业务,并保障协议的可信性. 可信控制平面则包括一组可信协议,提供完备一致的控制信令,实现对用户和网络运行信息的分布式采集、传播和处理,支持信任信息在可信用户间的共享,并驱动和协调具体的行为控制方式. 数据平面接受可信控制平面的监管,可信控制平面则向数据平面开放某些访问接口,从而使得业务能够获知网络运行是否可信,网络也可以根据用户要求为业务定制某种模式的运行方式,授予更高的信任级别,体现更高的可信保障水平.

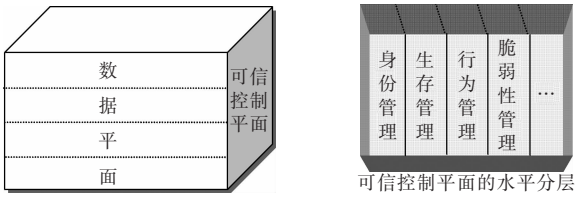

图 6 可信网络的一种体系结构模型

4.3 网络服务质量

随着高速网络技术和多媒体技术的飞速发展,人们越来越多地提出了包括多媒体通信在内的综合服务要求. 传统的分组交换网络,如 Internet,是面向非实时的数据通信(如 FTP 和 E-mail 的传输)而设计的,采用的 TCP/IP 协议主要是为了优化整个网络的数据吞吐量并保证数据通信的可靠性. 而当今分布式多媒体应用(如视频会议、视频点播、IP 可视电话、远程教育)不仅包括文本数据信息,还包括语音、图形、图像、视频、动画这些类型的多媒体信息. 分布式多媒体应用不仅对网络有很高的带宽要求,而且要求信息传输的低延迟和低抖动等,同时,这些应用大都能容忍一定程度的信息丢失和错误. 由此可见,当今高速网络中的多媒体应用对网络提出了不同于数据应用的服务质量要求,需要提供端到端的服务质量控制和保证.

所谓服务质量,是指网络传输数据流时要求满足的一系列服务请求,具体可以量化为带宽、延迟、延迟抖动、丢失率、吞吐量等性能指标. 对计算机网络 QoS 的研究可以追溯到 20 世纪 80 年代初期. 那时,尽管网络的性能还比较低,提供的服务种类也比较少,但一些有远见的研究者已经认识到服务质量的重要性. Seitz 和 Wortendyke 等人在研究 ARPANET 中的 X.25 通信时已提出基于用户的性能评价问题,这也许是关于 QoS 研究的最早文献. 在早期的 OSI 协议制定中,也为服务质量的一些参数留有相应的表示手段,但一直空缺未用. 很长的一段时间,由于计算机网络的性能所限,人们对 QoS 的关注只停留在数据流传输中的正确率、吞吐量和延迟等单一服务质量的评价与控制上. 直到 20 世纪 80 年代末期,随着 B-ISDN 以及 ATM 交换网的出现和分布式多媒体应用的急剧增加,人们才开始系统地对 QoS 管理和控制进行较为深入的研究.

随着 Internet 商业化的巨大成功,网上传输的多媒体信息迅速增多,网络拥塞现象日益严重,Internet 的 QoS 问题研究也随之开始深入. IETF 于 1997 年 9 月开始制定了有关 QoS 定义与服务的一系列 RFC 标准,典型的工作是提出了两种不同的 Internet QoS 体系结构: IntServ 和 DiffServ. 然而,这两种 QoS 体系结构都没有在实际系统中成功应用. 具有面向连接特性的 IntServ 虽然能够很好地提供 QoS 保障,但复杂的管理开销严重影响了它的可扩展性,使其无法在大规模的网络中部署实施. 另一方面,DiffServ 建立在遵守"边缘论"原则和面向无连接的网络体系结构之上,由于缺乏有效的管理控制手段而无法提供有力的 QoS 保障. 因此,如何建立可扩展的端到端 QoS 保障模式,从而达到既能提供类似 IntServ 的强有力的服务,又可继承 DiffServ 具有较小管理开销和扩展性强的优势,是一个非常重要的问题. 我们认为,实现在面向非连接的网络中提供面向连接的特性以及数据传输与服务管理相互独立的功能平面设计是解决这一问题的有效途径.

4.3.1 IP QoS 与过量配置(overprovision)

IntServ 与 DiffServ 作为 IP 网络的两种 QoS 标准目前并没有成功商用. 这除了两者自身存在的局限性之外,一个重要的原因在于目前核心网由于高速传输和交换技术的发展而产生"带宽过剩"的现状. 在这种情况下,许多骨干网运营商都通过为"尽力而为"的互联网配置大量过剩容量以保证用户的 QoS,而不愿对已经大量部署的路由器等设备进行升级从而采用复杂的 QoS 控制机制[29]. 据统计,目前大部分骨干网的负载大约只占总容量的 5%～15%,而核心互联网运营商通常提前配置容量以满足 3～6 个月后增长的用户需求.

这种过量配置的方式虽然能够暂时满足目前部分用户的 QoS 需求,然而从历史发展的角度看,却存在着潜在的威胁. 在网络的发展历史上,由于新接入网和新核心网技术的交错出现,核心网的容量周期性地被人们认为从"带宽无穷大"到"几乎拥塞崩溃"[30],如图 7 所示. QoS 控制机制似乎只有当接入网容量与核心网容量的比值较高时才具有研究价

值. 然而，QoS 控制机制必须作为网络体系结构的一个有机组成部分才能最有效地实现目标. 如果只有当面临迫切需求时才向一个已经大量部署的网络中加入 QoS 机制只能达到事倍功半的效果. 因此，下一代互联网体系结构的设计应该充分考虑如何为用户有效地提供服务质量保证.

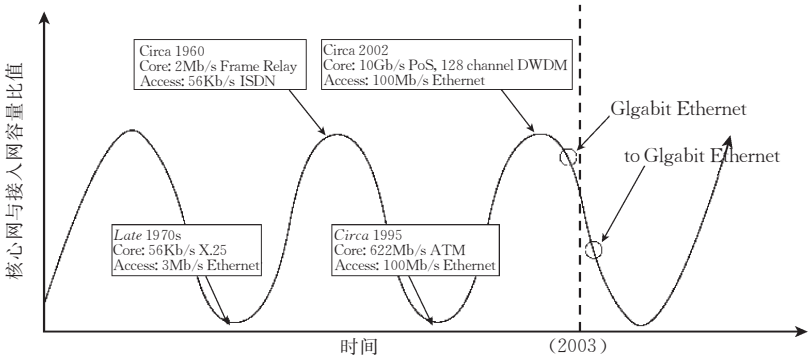

图 7　核心网与接入网容量的比值

4.3.2　基于 IP 的核心网：新型 IP QoS 体系结构

由于 IntServ 与 DiffServ 并没有成功商用，因而人们又开始对两者的局限性进行分析，并在此基础之上提出一些新的 IP QoS 体系结构.

(1) DiffServ 与 IntServ 相结合的 QoS 体系结构

IntServ 和 DiffServ 各有自己的长处和局限，但都不能彻底实现整个网络的端到端 QoS. 为此，人们考虑将 IntServ 和 DiffServ 看作互相补充的技术，将其结合，互相协同，共同实现端到端的 QoS 提供机制. 最终达到既能提供类似状态相关网络的强有力的服务，又能实现与状态无关网络近似的可扩展性和鲁棒性.

IntServ 体系结构提供了一种在异类网络元素之上为应用提供端到端 QoS 的方法. 一般来讲，网络元素可以是单独的节点(如路由器)或链路，更复杂的实体(如 ATM 云或 802.3 网络)也可以从功能上视为网络元素. 在这种意义下，DiffServ 网络也可以视为更大的 IntServ 网络中的一种网络元素. 文献[31]描述了一种在 DiffServ 网络之上实施 IntServ 的框架，给出了在 IntServ 体系结构下 DiffServ 网络支持端到端 QoS 的方法.

在该框架中，端到端的、定量的 QoS 是通过在含有一个或多个 DiffServ 区的端到端网络中应用 IntServ 模型来提供的. 为了优化资源的分配和支持接纳控制，DiffServ 区可以(但并不绝对要求)参加端到端的 RSVP 信令过程. 从 IntServ 的角度来看，网络中的 DiffServ 区被视为连接 IntServ 路由器和主机的虚链路. 该框架的目标是实现 IntServ 区与 DiffServ 区无缝的相互操作，网络管理员可以自由选择网络中哪个区作为 DiffServ 区. 这种 DiffServ 与 IntServ 相结合的 QoS 体系结构虽然在一定程度上改进了两者各自的缺点，但仍没有对运营商产生太大的吸引力.

(2) 基于重叠网络的 QoS 体系结构

鉴于 IntServ 与 DiffServ 没有成功商用的主要原因之一在于需要对大量已经部署的 IP 路由器进行修改，人们提出通过在网络层之上构建重叠网络增强整个网络的 QoS 能力，一个典型的例子是 OverQoS 体系结构[32]. 在一个 OverQoS 网络中，每个路由域中都部署若干个重叠节点，每个重叠节点之间通过 IP 路径互相连通，这些 IP 路径称为重叠链路或虚链路. OverQoS 假设每个路由域的重叠节点部署方式以及源到目的重叠节点之间的重叠路径已经确定，则每个重叠节点可以通过一系列 QoS 控制机制在“尽力而为”的网络层之上为每条重叠链路提供一定的丢失率和带宽保证.

基于重叠网络的 QoS 体系结构虽然能够在一定程度上增加网络 QoS 能力，却很难彻底使网络为用户提供严格的 QoS 保障. 这种方式相当于在已有网络上增加一个控制系统，同时也增加了系统开销和不稳定因素.

4.3.3　无线接入网：跨层的体系结构

随着无线通信与网络快速占据通信网络领域研发活动的重要位置，作为网络设计最基本原则之一的分层协议体系结构开始受到质疑. 分层体系结构将整个网络功能分为若干层，每一层通过特定协议实现本层的功能并为上一层提供服务. 不相邻的层次之间禁止直接通信，且相邻层次之间的通信只能通过有限的程序调用和响应而实现，如 OSI 七层模型和 TCP/IP 五层模型等. 许多研究者认为，无线网络不适于采用有线网络中的层次结构，而需要跨层的体系结构设计. 这主要是由于三方面的原因造成的. 首先，无线链路对协议设计提出了若干新的问题，在层次结构的框架下无法很好地解决. 例如，

TCP 协议发送者会误将无线链路由于信道质量劣化造成分组丢失作为网络拥塞的标志. 其次,在无线网络中可以通过机会通信提高网络的性能,而严格的分层设计使得无法充分发挥机会通信的价值. 例如,时变的无线链路质量使得基站可以通过机会调度算法选择在某一时刻信道质量最佳的用户进行通信,从而得到所谓的多用户分集增益,增大系统吞吐量. 最后,无线媒质提供了层次结构无法支持的丰富通信形态. 例如,无线网络物理层可以同时接收多个分组,因此无线节点可以充分利用这种广播特性彼此协作以提高网络性能,而在协议设计中利用这种"新"通信模式也需要违反层次结构的原则.

跨层设计的本质就是违背层次体系结构的原则,对于目前无线网络的跨层设计来说主要体现为以下几种形式[33]:

(1) 在协议层之间创建新的接口,以便于网络运行时不同协议层之间共享信息. 此类跨层设计可以根据新接口的信息流向分为 3 个子类型:

① 信息流从低层到高层,以便于低层通知高层其网络状态,如图 8(a)所示.

② 信息流从高层到低层,以便于高层通知低层应用数据的处理方式,如图 8(b)所示.

③ 高层和低层之间双向信息流,以便于完成不同功能的两个协议层可以在运行时进行合作,如图 8(c)所示.

(2) 合并相邻的协议层,即共同设计两个或多个相邻的协议层,以使新的"合并层"提供的服务为其组成各层提供服务的合集,如图 8(d)所示. 由于"合并层"可以使用原体系结构中已有的接口与其它层进行通信,因而这种跨层设计方式不需要创建新的接口.

(3) 不通过新接口直接耦合协议层,即在设计时不通过创建新接口的方式耦合两个或更多的协议层,以便于这些协议层可以在运行时共享信息,如图 8(e)所示. 这种跨层设计的后果是如果某一层发生了变化,则相应的耦合层也可能需要进行修改.

(4) 垂直调整,即通过多个协议层共同对某项参数进行调整,如图 8(f)所示. 这是由于应用层的性能参数是其下面各层参数的函数,因而由各层联合对参数进行调整可以获得比各层独立设置参数更好的性能.

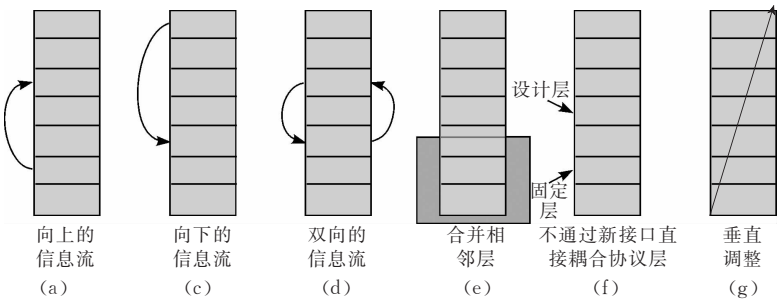

图 8 跨层设计的形式

目前提出的跨层设计的实现方式大致可以分为以下 3 类:

(1) 层间直接通信,即允许某一层的变量在运行时对另一层可见,例如可以通过协议头或内部分组的方式在层间传递信息,如图 9(a)所示. 在严格的分层体系结构中,每一层的变量都由本层进行独立管理而与其它层无关. 这类实现方式比较适用于与分层体系结构相比只需传递少量额外信息的跨层设计,虽然在分层协议栈上打穿了若干小孔,但仍大体保持了原有结构.

(2) 层间共享数据库,即创建一个各层都可以访问的数据库,如图 9(b)所示. 在某种意义上,共享数据库类似于一个新的协议层,向所有其它层提供信息存取的服务. 此类实现方式尤其适用于基于垂直调整的跨层设计,其主要需要解决不同层与共享数据库之间交互的设计问题.

(3) 全新的抽象模型,例如文献[34]提出将协议以堆(heap)而不是堆栈(stack)的方式进行组织. 这类实现方式可以实现协议组成模块之间更丰富的交互,从而增强了网络设计和运行时的灵活性. 但由于其彻底改变了协议原来的组织方式,因而也意味着全新的系统级实现.

虽然目前在无线网络中对跨层设计的研究如火如荼,但跨层体系结构的实现还面临着一系列挑战. 首先,虽然目前已经提出了大量的跨层设计方案,但并不清楚哪些才是最重要的. 例如,目前研究较多的包括 Ad Hoc 网络 MAC 层和网络层的跨层设计;传送层与底层的跨层设计;MAC 层与物理层的跨层设计等. 然而,很少有文献对各种跨层设计的实现复杂度与性能改善进行定量的研究. 其次,各种跨层

设计方案如何共存也是必须解决的一个问题，或者说某一种跨层设计方案是否会使得体系结构变得僵化而阻碍其进一步的发展. 例如，一种 MAC 层与物理层的跨层设计方案是否可以与另一种 MAC 与传送层的跨层设计方案很好地共存?最后，分层体系结构的优点之一在于其很好地规定了各层之间的边界和接口. 跨层设计以各种形式违背了层次结构，然而必须找到一种方式定义这种新的体系结构. 例如，必须明确新体系结构模块之间的边界是如图 9(a)和(b)一样在原有层次结构的基础上进行新的接口定义，抑或是如图 9(c)一样寻找全新边界. 这也需要对跨层设计实现和性能之间的折中进行定量分析，使得新的跨层体系结构既可以有效共享层间信息又不需要以过大的开销和时延作为代价.

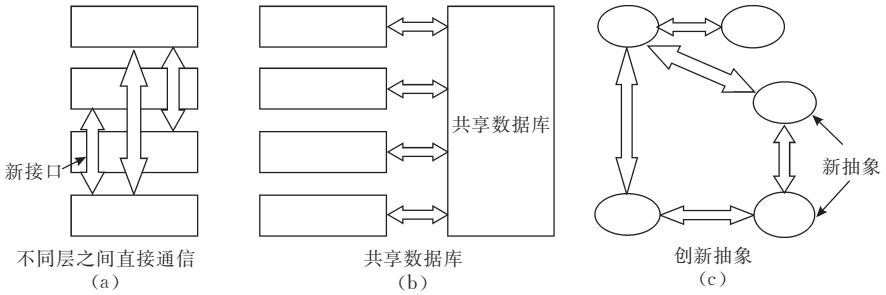

图 9　跨层设计的实现方式

4.4　网络管理

下一代互联网应该是一个可管理的网络. 所谓“可管理的网络”，是指在网络环境受到内外干扰的情况下，不但对网络状态而且对用户行为进行持续的监测、分析和决策，进而对设备、协议和机制的控制参数进行自适应优化配置，使得网络的数据传输、资源分配和用户服务的过程及结果是可以预期的. 可管理的网络是一个应具有充分可管理性的网络，网络的可管理性对于网络的其它本质属性，如安全性、鲁棒性、普适性等也具有重要的支撑作用. 当前 IP 网络的体系结构和管理协议不支持可管理性设计，仅在现有网络的体系结构添加修改网络管理功能无法实现可管理网络的目标. 因此，对可管理的 IP 网络的基础研究应该通过改进 IP 网络体系结构中导致可管理性不足的设计原则，实现 IP 网络的充分可管理性，从而促进 IP 网络管理；并进一步为解决一系列现实的网络本质属性问题，如安全性、鲁棒性、普适性，服务质量(QoS)保障等提供支撑；以及为网络的进一步发展提供自适应的能力. 目前，国内外相关的研究不多，下面主要介绍两种相关的新型互联网体系结构：4D 和 Ambient 网络. 前者主要提出将互联网的决策逻辑和分布式系统分离，加强集中式的管理；后者则从移动网络流媒体发布的角度出发，强调在网络的某些节点中加入业务处理和安全相关的功能，使得网络除了尽力而为传送数据之外具有更多的业务支持能力. 这种思路与 4.2.2 节中介绍的防御 DoS 攻击的安全体系结构思路具有某些一致性，都提出在某种程度上打破传统互联网“边缘论”的思想，加强网络的可控性. 我们认为这种思路是未来互联网体系结构发展的一个重要方向.

4.4.1　4D 体系结构

当前的互联网非常脆弱和难于管理，一个重要的原因在于其决策逻辑和分布式系统交织在一起，导致控制和管理平面非常复杂. 例如，如图 10 所示，假设某公司有两个工作地点 A 和 B，每个工作地点有若干台供销售人员使用的“前台”计算机($AF1$-2 和 $BF1$-2)以及一个包含若干台服务器($AD1$-2，$BD1$-2)的数据中心. 最初两个工作地点通过前台路由器 $R2$ 和 $R4$ 之间的一条链路相连. 公司的安全策略是每个前台计算机可以与另一地点的前台计算机以及本地数据中心的服务器通信，但不能与另一地点的服务器通信. 这种安全策略在保险等行业非常普遍，即使同一公司不同地点的销售人员彼此之间也存在竞争. 该安全策略通过在控制数据中心入口的路由器上使用分组过滤器将违反规定的分组丢弃来实现. 网络管理员配置接口 i1.1 的分组过滤器使其丢弃所有来自子网 BF 的分组；配置接口 i3.1 的分组过滤器使其丢弃所有来自子网 AF 的分组. 然而，如果由于某种原因需要在两个地点的数据中心之间增加一条新的高速链路. 此时，从 AF 到 BD 的路由变为 $R2$-$R1$-$R3$，因此原来的分组过滤器 i3.1

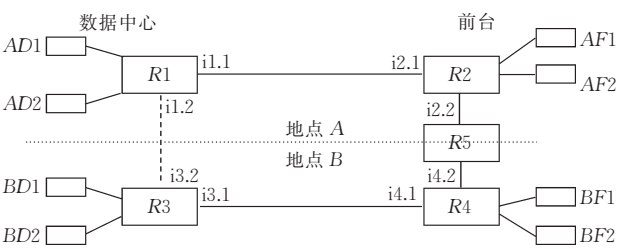

图 10　决策逻辑和分布式系统交织导致网络脆弱性

不再起作用,需要在接口 i3.2 上增加一个分组过滤器将来自子网 AF 的分组丢弃.然而,如果前台路由器 $R2$ 到 $R4$ 之间的链路发生故障,则虽然在两者之间还存在另一条路由 $R2$-$R1$-$R3$-$R4$,但由于分组过滤器的设置使得前台计算机之间仍无法通信.

出现上述问题的主要原因在于网络传递分组的能力依靠路由协议和分组过滤器的合作.虽然路由协议能够根据网络拓扑自动变化,但分组过滤器却不具备自动配置的能力.随着网络规模的扩大,网管系统的功能会越来越复杂.

针对这一问题,文献[35]提出一个全新的 4D 网络体系结构进行网络控制和管理.4D 将互联网从逻辑上分为 4 个平面:数据(data)、发现(discovery)、发布(dissemination)和决策(decision).一个自治域的决策平面创建一个满足自治域级目标的网络配置.决策平面具有全网的拓扑和业务信息,并直接控制数据平面的操作.网络节点中实现的分布式协议不再具备任何决策功能,如路由选择等.决策平面的输出结果通过发布平面发送给路由器和交换机.这种方法的优点在于简化了网络中的协议,同时又可以使网络运行复杂的算法以驱动数据平面的操作.例如,在上面的例子中,4D 体系结构可以指定一个"可达矩阵",并通过决策平面根据当前网络状态自动的建立转发表并配置分组过滤器.然而,4D 体系结构采用的集中式控制方式在实时性和可扩展性方面存在固有的缺陷,而且很容易受到单点故障的影响.

4.4.2 Ambient 网络体系结构

目前移动网络和业务已经逐渐由只支持语音通信向支持基于 IP 的多业务通信发展,如彩信(MMS)、视频流、博客、音乐下载和移动电视等.这些业务的开发和部署大多采用只由端节点(服务器和客户)进行控制和管理的端到端模式,连接这些端节点的网络只负责简单的传送功能,而不对业务产生附加的价值或作用.这种端到端的原则延续了传统互联网"边缘论"和面向非连接的设计思想,保障了高效的互通和良好的可扩展性.然而,各种移动终端的处理能力不尽相同,而且可以通过多种接入技术访问核心互联网.当这些异构的终端设备试图同时访问同一服务时,端到端的模式就显现出其固有的局限性.在此情况下,服务器需要区别对待不同的终端设备,因而只能依靠高性能服务器集群支持大量用户,并且需要大量带宽同时传送不同的数据格式给不同用户.

解决此问题的一种可行方案是使得网络自身对流媒体发布作出贡献.例如,使通信路径上的节点具有流媒体处理能力,包括转码和缓存功能等.当前的移动核心网络也存在某些流媒体处理节点,如流媒体网关或者 MMS 中继服务器等.然而,这些节点是为支持某种特定的服务而设计,无法处理其它类型的业务,也无法被端节点所控制.

为了允许任意业务触发,配置和部署网络的处理能力,就需要能够方便地将这些节点加入到已经部署的网络中,并使得每种业务和每个用户可以动态地触发网络处理能力.由欧盟部分出资支持的 Ambient 网络项目对这些问题进行了研究,提出一种新的网络控制体系结构,尤其适用于具有丰富特性的移动互联网络[36].

移动网络用户在不同接入网络之间移动时体验不同的网络能力和连接质量.因此,一个重要的要求是能够令业务动态适应当前网络和端节点的状态.Ambient 网络通过一个新型的重叠控制层来解决上述问题,即通过 Ambient 控制空间(ACS)注册和组织网络中所有的控制功能.图 11 所示为 ACS 的逻辑组织和主要特点.ACS 主要包括 Ambient 网络接口(ANI)、Ambient 业务接口(ASI)和 Ambient 资源接口(ARI).

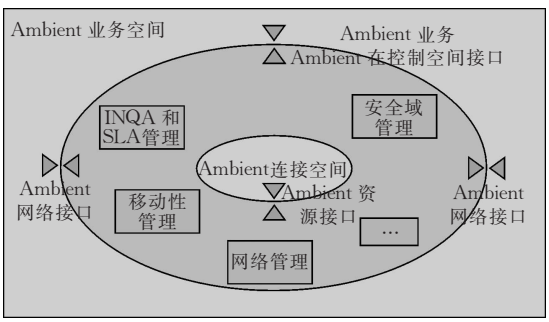

图 11 控制空间和接口

ANI 用于不同网络的控制空间之间的通信.它允许将控制空间的功能发布给其它网络以供他们使用,并将一个 ACS 的功能与另一个 ACS 的功能相连接.因此,它允许不同 ACS 的动态合作,称作组合(composition).简而言之,ANI 是不同 Ambient 网络之间的接口.

ASI 将连接和控制功能以一种统一的方式提供给上层的应用和业务使用.它允许端用户和应用向 ACS 发起请求以建立、维护和终止终端系统之间的端到端连接和传送业务.ASI 也包含管理能力并提供使应用能够得到网络上下文信息的方法.简而言之,ASI 是业务提供者、业务接收者和 Ambient 网络之间的接口.

最后,ARI 在 ACS 和连接层之间,它向 ACS 提供控制机制以管理连接平面的资源.这些资源可以

通过一个抽象层访问，使得 ACS 功能与传送技术无关. 简而言之，ARI 是控制功能和底层传送网络（如 IP 网络）之间的接口.

在传统的端到端流媒体发布概念中，不存在接口使得业务提供者和用户可以向底层传送网络请求特定的功能：网络只提供最简单的业务传送能力. Ambient 网络支持的业务模式则不同，如图 12 所示，ASI 提供网络可能具有的流媒体发布能力以支持尽力而为传送以外的其它业务. 它是任何希望使用 Ambient 网络能力的业务和应用的联系点. Ambient 网络中提供流媒体处理能力的节点称为流媒体端口（MediaPort）. 流媒体端口可以包含多种功能和能力，大体可以分为媒体处理能力、客户化能力、安全相关能力或其它能够监视、处理或重路由属于一个业务的数据分组的能力. ASI 允许业务询问网络所能提供功能的信息，从而为一个传送载体选择特定的支持资源，并最终在该载体上发送流媒体以使用这些功能. ASI 允许用户应用或业务协商和控制 Ambient 网络的传送业务.

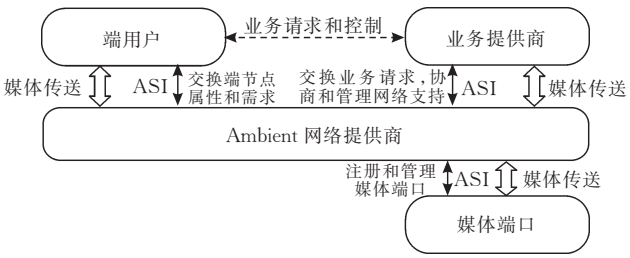

图 12　Ambient 网络具有网络处理能力的媒体发布

下面的例子使用 Ambient 网络提供基于 SIP 的语音业务. 假设移动用户使用其 SIP 客户端建立一个语音会话，并希望对这个会话进行录音，如图 13 所示. 在连接建立时，一个端节点发起话音业务，相关的具有 ASI 的 SIP 服务器或代理检测到对网络录音能力的需求. 它使用 ASI 向底层网络请求并协商一个包含录音节点的传送服务. 当接收到这一请求后，Ambient 网络的控制空间通过查找流媒体端口数据库或使用流媒体端口发现机制查找一个可用的录音节点. 如果存在此节点，则 ACS 提供为此业务建立一个适当的重叠网络. 当控制空间与 SIP 服务器或代理之间就重叠网络的建立达成最终协议时，重叠网络被实际建立，使得 SIP 对等实体之间的端到端路径上包含录音流媒体端口. 作为重叠网络一部分的录音节点通过 ARI 接口进行配置. 当重叠网络成功建立后，语音业务开始使用传统的流协议，录音流媒体端口则记录下业务流. 在通话结束后，它将记录的数据发送到端用户管理的存储空间，重叠网络被拆除.

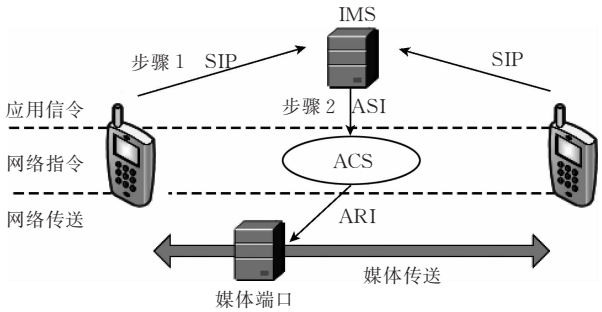

图 13　Ambient 网络基于 SIP 的部署示例

5　总结与展望

通过对下一代互联网领域的调研可以发现一个有趣的现象，目前对下一代互联网体系结构进行研究最活跃的是一些美国的著名大学，如加利福尼亚大学伯克利分校和普林斯顿大学等. 这些大学正是当初互联网萌芽的地方，而目前一些主要研究者则是当年互联网的主要发起者. 这些研究大多集中在技术层面上，主要目的在于根据当前的网络环境，重新思考当初互联网设计的一些基本原则. 然而，目前这些研究提出的体系结构往往只能解决当前互联网存在的一个或几个问题，而且从实现层面上讲很多都难以实现从现有互联网的平滑过渡.

然而，就像当年人们无法预想 TCP/IP 技术可以从一个实验室的“玩具”发展成为当今全球最重要的商业网络之一，目前那些在实验室中孕育和发展的网络雏形也许会在将来发展成为下一代互联网的基础. 相比之下，我国对下一代互联网体系结构的研究还远远不够. IPv6 网络从提出到目前终于开始准备进行大规模部署花费了 10～15 年的时间，而 IPv6 绝不是互联网发展的终点，对 10～15 年后互联网体系结构的研究必须从现在开始. 因此，本文希望能够唤起我国科技工作者对下一代互联网体系结构研究的足够重视，从各个层面充分审视现有体系结构的缺陷，在理论研究上保持开放的头脑，突破传统研究思路，并构建实验平台开展相关实验. 我国信息网络发展迅速，网络依赖性日益增强，下一代互联网络体系结构的研究为我国取得此领域实质性突破提供了一次重要机遇.

我们认为，下一代互联网体系结构从技术和实现层面应该遵守若干原则. 从技术层面来说，通过上面的介绍可以看到，传统互联网设计所遵守的基本原则很多已经无法适应当前的网络环境. 下一代互联网应该将主机的名称与主机的位置信息相分离，

将主机上资源的名称与主机名称相分离;将网络安全与服务质量作为与数据传送相并列的网络内在要求进行考虑;在网络管理方面考虑集中与分布式控制相结合的方式,增强网络的可控性.互联网络发展至今,已成为一个庞大的非线性复杂系统,具体表现为:系统的规模和用户数量巨大且仍在不断增长;异构异质的网络融合发展;网络协议体系庞杂;在业务性质上表现为多种业务的集成与综合,突发性日渐明显,且不同业务要求不同的服务质量保证;网络节点间、节点与数据分组间由于协议而产生的非线性作用以及用户之间的合作与竞争,使网络行为呈现出相当的复杂性并且难以预测.如何为用户提供满意的安全性和服务质量,设计有效的机制与算法来管理和控制这个已经具备相当规模,并不断发展的大尺度复杂系统,在理论上和技术上仍是当前学术界的一个难题.我们认为,下一代互联网应该在网络的关键部分增加认证、授权等控制机制使网络更可信;在网络中维护一定的状态信息,施加必要的控制,使网络具有某种程度面向连接的特性;在不同层次上可以对它进行监视和控制.而且,下一代互联网应该包含多种类型的网络安全手段,除了传统"堵漏洞、作高墙、防外攻"等被动防御手段外,还应该通过类似"蜜罐"等的主动防御手段对攻击者进行深入的了解和分析,从而尽可能地将不信任的访问操作控制在源端,并找到攻击者从而产生威慑的作用;通过可生存性设计,在系统脆弱性不可避免以及攻击和破坏客观存在的现实状况下,提高包括安全服务在内等关键服务的持续能力等.

从实现层面来说,当前互联网作为一个已经在全球范围内商用的网络不可能在一朝一夕之间被替换,因而任何新的互联网体系结构只能利用虚拟化等技术逐渐对现有互联网进行改变.另外,只有市场和商业利益才是互联网发生变革的根本推动力.这就要求下一代互联网体系结构必须能够与当前互联网进行互通,只有这样才能让部分 ISP 首先采用下一代互联网技术部署网络,并能够使新互联网用户在访问所有互联网资源的同时感受到新技术带来的优势,从而逐步推动新技术的扩展.

参 考 文 献

[1] Peterson L, Shenker S, Turner J. Overcoming the Internet impasse through virtualization. Computer, 2005, 38(4): 34-41

[2] Roscoe T. The end of Internet architecture//Proceedings of the SIGCOMM Workshop on Hot Topics in Networking. Maryland, USA, 2005

[3] Peterson L, Andersen T, Culler D, Roscoe T. A blueprint for introducing disruptive technology into the Internet//Proceedings of the SIGCOMM Workshop on Hot Topics in Networking. New Jersey, USA, 2002

[4] Braden R, Clark D, Shenker S. Integrated services in the Internet architecture: An overview. IETF RFC 1633, June 1994

[5] Carlson M, Weiss W, Blake S et al. An architecture for differentiated services. IETF Internet RFC 2475, Dec., 1998

[6] Francis P, Gummadi R. IPNL: A NAT-extended Internet architecture//Proceedings of the ACM SIGCOMM. San Diego, CA, 2001: 69-80

[7] Turanyi, Valko. Campbell, Design, Implementation and Evaluation of IPv4+4. ACM SIGCOMM Computer Communication Review, 2003, 33(5): 43-54

[8] Moskowitz R. Host identity protocol. Internet draft, draft-moskowitz-hip-09, 2004

[9] Koponen, Gurtov, Nikander. Application mobility with host identity protocol — Extended abstract//Proceedings of the NDSS'05 Workshop. San Diego, CA, USA, 2005

[10] Nikander, Ylitalo, Wall. Integrating security, mobility, and multi-homing in a HIP way//Proceedings of the Network and Distributed Systems Security Symposium (NDSS'03). San Diego, CA, 2003: 87-99

[11] Cheriton D R, Gritter M. TRIAD: A scalable deployable NAT-based Internet architecture. Stanford Computer Science Technical Report, 2000

[12] Gritter M, Cheriton D R. An architecture for content routing support in the Internet//Proceedings of the USENIX Symposium on Internet Technologies and Systems. San Francisco, California, USA, 2001

[13] Vahdat, Dahlin, Anderson, Aggarwal. Active names: Flexible location and transport of wide-area resources//Proceedings of the USENIX Symposium on Internet Technologies and Systems. Boulder, Colorado, USA, 1999

[14] Adjie-Winoto, Schwartz, Balakrishnan, Lilley. The design and implementation of an intentional naming system. Operating Systems Review, 1999, 34(5): 186-201

[15] Stoica, Adkins, Zhuang, Shenker, Surana. Internet indirection infrastructure//Proceedings of the ACM SIGCOMM Pittsburgh. PA, 2002: 73-88

[16] Gold R, Gunningberg P, Tschudin C. A virtualized link layer with support for indirection//Proceedings of the ACM SIGCOMM Workshop on Future Directions in Network Architecture (FDNA'04). Portland, Oregon, USA, 2004: 28-34

[17] Uddenfeldt, Gold. A content adaptation network using SelNet//Proceedings of the 2nd Swedish National Computer Networking Workshop (SNCNW 2004). Karlstad University, Sweden, 2004

[18] IAB, Handley M. ed. Internet denial of service considerations. draft-iab-dos-00.txt, work-in-progress, IETF, 2004

[19] Lin Chuang, Peng Xue-Hai. Research on trustworthy networks. Chinese Journal of Computers, 2005, 28(5): 751-758(in Chinese)

(林　闯，彭雪海. 可信网络研究. 计算机学报，2005，28(5)：751-758)

[20] Lin C，Peng X-H. Research on network architecture with trustworthiness and controllability. Journal of Computer Science and Technology，2006，21(5)：732-739

[21] Argyraki K，Cheriton D R. Active Internet traffic filtering：Real-time response to Denial- of-service attacks//Proceedings of the USENIX Annual Technical Conference. Anaheim，CA，USA，2005：135-148

[22] Greenhalgh A，Handley M，Huici F. Using routing and tunneling to combat DoS attacks//Proceeding of Usenix Workshop Steps to Reducing Unwanted Traffic on the Internet (SRUTI). Cambridge，MA，USA，2005：1-7

[23] Mahajan R，Bellovin S，Floyd S，Ioannidis J，Paxson V，Schenker S. Controlling high bandwidth aggregates in the network. ACM CCR，2002，32(3)：62-73

[24] Tom Anderson D W，Roscoe Timothy. Preventing Internet denial-of-service with capabilities. ACM SIGCOMM Computer Communication Review，2004，34(1)：39-44

[25] Yang X，Wetherall David，Anderson Thomas. A DoS limiting network architecture. Computer Communication Review，2005，35(4)：241-252

[26] Handley M，Greenhalgh A. Steps towards a DoS resistant Internet architecture//Proceedings of the ACM SIGCOMM Workshop on Future Directions in Network Architecture (FDNA). Portland，USA，2004：49-56

[27] Paulson L D. Stopping intruders outside the gates. IEEE Computer，2002，35(11)：20-22

[28] Algridas A，Laprie J C，Brian R，Carl L. Basic concepts and taxonomy of dependable and secure computing. IEEE Transactions on Dependable and Secure Computing，2004，1(1)：11-33

[29] Bell. Failure to thrive：QoS and the culture of operational networking//Proceedings of the ACM SIGCOMM Workshop on Revisiting IP QoS (RIPQoS′03). Karlsruhe，Germany，2003：115-120

[30] Crowcroft et al. QoS′s downfall：At the bottom or not at all//Proceedings of the ACM SIGCOMM Workshop on Revisiting IP QoS(RIPQoS′03). Karlsruhe，Germany，2003：109-114

[31] Bernet Y，Yavatkar R，Ford P et al. A framework for integrated services operation over diffServ networks. IETF Internet Draft，1999

[32] Subramanian L et al. OverQoS：Offering Internet QoS using overlays//Proceedings of the 1st Workshop on Hot Topics in Networks (HotNets-I). New Jersey，USA，2002

[33] Srivastava V. Cross-layer design：A survey and the road ahead. IEEE Communications Magazine，2005，43(12)：112-119

[34] Braden R，Faber T，Handley M. From protocol stack to protocol heap—Role-based architecture. ACM SIGCOMM Computer Communication Review，2003，33(1)：17-22

[35] Greenberg et al. A clean slate 4D approach to network control and management. ACM SIGCOMM Computer Communication Review，2005，35(5)：41-54

[36] Hartung F et al. Advances in network-supported media delivery in next-generation mobile systems. IEEE Communications Magazine，2006，44(8)：82-89

LIN Chuang，born in 1948，Ph. D.，professor，Ph. D. supervisor. His current research interests include computer networks，performance evaluation，logic reasoning，and Petrinet theory together with its applications.

LEI Lei，born in 1980，Ph. D.. She is especially interested in future Internet architecture and quality of service in wireless network，etc.

Background

This work is supported in part by a grant from the National Grand Fundamental Research 973 Program of China (2006CB708301)，and the National Natural Science Foundation of China (60673187，90412012).

The Internet has been a fantastic success，but in many ways it is not meeting the needs of its users. The Internet architecture has had a profound effect on the Internet′s ability in the past to evolve into possibly dominant networking technology today. Recently，however，the architectural view has come under increasing strain both from within and without.

While there is universal agreement that creating a future Internet that meets the demands of the 21st century is of critical important，there are many schools of thought as to how to pursue this goal. For example，some advocate for a "clean slate" re-conceptualization of the Internet architecture，while others think it is simply not practical to think in terms of replacing all of today′s Internet infrastructure；some deem there should be only one Internet architecture，while others think several different architectures will co-exist in the future.

There are many challenges for the design of future Internet，including addressing issues，security，QoS and network manageability etc. Current researches usually focus on dealing with one of these challenges. Recent years have seen considerable publishing activity in these areas. This paper tries to give a survey of the research status of all the above issues，rethink the design principles of Internet architecture，examining whether they still meet the current and new network environment demand，and analyze the development prospect of future Internet.