下一代互联网体系结构

苏金树 吴纯青 胡晓峰 彭 伟 等
国防科学技术大学

关键词：计算机网络 体系结构 下一代互联网

引言

为解决目前互联网存在的安全性、可扩展性等诸多问题，满足不断增长的应用需求，美国国家科学基金会（National Science Foundation，NSF）分别于2005年和2006年启动了GENI（Global Environment for Network Innovation，全球网络创新环境）计划[1]和FIND（Future Internet Network Design，未来互联网网络设计）[2]计划，希望从互联网的基本设计原则和体系结构上寻求解决方案。GENI计划试图建立一个全球性的网络试验平台，用来进行各种新的网络技术试验，并以此为基础逐渐演进到下一代互联网。GENI计划认为，未来互联网应该具有很强的生存性，能在国家出现危机时提供服务，应当增强在遭受攻击或者局部故障情况下网络的可用性和恢复技术等方面的研究。FIND计划试图从底层开始对互联网体系结构进行重新设计，并把安全性和鲁棒性作为设计的基本要求。除了美国以外，欧洲发达国家和日本也对未来网络安全性的基础研究给予了高度重视，如欧盟的FIRE项目[3]、日本的AKARI项目[4]等。

美国国防部认为，当前的互联网技术不足以作为可确保全球网络（Assurable Global Networks，AGN）[5]的基础。2006年12月中旬，美国国防部高级研究计划署（Defense Advanced Research Projects Agency，DARPA）战略科技办公室（Strategic Technology Office，STO）发布征求意见书，征求能给AGN奠定基础的研究思想和方法，并于2007年2月召开了战略研讨会。与会者普遍认为，当前网络脆弱性的根源在于互联网初始设计原则的优先顺序以及互联网体系结构；需要着眼于军事需求，从基础理论和体系结构上对AGN进行研究。下一代互联体系结构研究中一个重要的分歧是：一个学派希望通过建立新型的网络体系结构全面提升网络能力，而另一个学派希望提出基于层叠方法的新型体系结构，在兼容目前结构的前提下，扩展当前网络的适应性。同时，他们希望深化网络威胁问题描述能力，为网络生存性奠定基础；希望增强安全防御能力，提高网络生存性；希望增强路由技术，提高网络的生存性和适应性等。下面介绍这些研究进展。

新型的网络体系结构

由于诸多网络问题源于互联网初始设计原则和由此产生的体系结构，研究人员开始着手对新型网络体系结构进行探索。2000年，麻省理工学院的克拉克（Clark）等人承担了美国国防部高级研究计划署的下一代互联网体系结构NewArch项目，提出了知识平面（Knowledge Plane）[6]的概念，拓展了互联网体系结构研究的方法和思路。

体系结构研究主要集中在地址和身份分离、控制和数据分离等方面。目前TCP/IP[1]体系中的IP地址用于位置标识和端点的身份标识，这种双重功能限制了网络的移动性，也加大了访问控制的复杂性和

[1] Transmission Control Protocol/Internet Protocol，传输控制协议/英特网互联协议

难度。因此，在新型网络体系结构设计中严格区分位置和身份标识是许多研究机构的共识。

2004年，互联网研究工作组（The Internet Engineering Task Force，IETF）提出了主机标识协议（Host Identity Protocol，HIP）[7]，以实现位置与身份标识的分离。主机标识协议将传输层与网络层解耦合，在两层之间加入主机标识层。这种分离提供了一种安全的处理移动性和多归属（Multi-Homing）的方式。

2005年，卡耐基梅隆大学阿尔伯特·格林伯格（Albert Greenberg）等人提出了4D（Decision，Dissemination，Discovery，Data）结构[8]，即把网络控制平面分解为4个概念性组件：决策、分发、发现和数据平面。自治域的决策平面创建满足自治域级目标的网络配置要求，且具有全网拓扑和业务信息，直接控制数据平面的操作。决策平面的输出结果通过分发平面发送给路由器和交换机。网络节点中的分布式协议不再具备任何决策功能，如路由选择等。4D结构直接控制网络，决策逻辑把安全策略规范作为输入，来控制可以进行交换报文的源和目的子网对。4D结构对后来的工作有较大影响。

2006年，斯坦福大学马丁·卡萨多（Martin Casado）等提出了SANE（Scanner Access Now Easy）结构[9]，与互联网“默认打开+特殊过滤”的机制相对应，SANE是“默认关闭+特殊允许”的机制。SANE集中管理全网实体的安全认证、接入控制和路由控制等，在核心上保障网络的安全性。端系统必须通过域控制器认证来获得网络接入权限，用户在访问资源之前需要进行准入检测；在发布服务之前，用户向域控制器注册服务的访问控制策略，以规范访问用户的身份和权限等。域控制器代理通信双方进行协商，并根据协商结果指定路由。

2007年，斯坦福大学与加州大学伯克利分校合作，吸收了4D结构的思想，进一步提出了Ethane[10]，总体思路是：由策略（Policy）控制整个网络包含的实体；策略指定每个报文流经过的路径；将报文与其产生源强制绑定。该结构包含控制器和Ethane交换机。中央控制器包含全局知识，对网络实体进行安全绑定和认证，维护全局安全策略，检测每个流是否违反了安全策略，并为每个流确定路由。Ethane交换机在中央控制器的测量指导下转发报文。Ethane将整个网络纳入控制范围，各种设备之间相互协作，提高了网络的安全性，但目前该结构对广播、应用层路由等支持不足。

基于层叠方法的新型体系结构

保证TCP有效工作的基本前提是端到端存在持

学会将重组中文信息技术专委
授权CCF理事肖建国负责重组事宜

2010年1月30日，CCF九届四次常务理事会会议决定保留并重组中文信息技术专委，同时通过决议授权CCF理事、北京大学教授肖建国负责重组事宜。

中文信息技术专委已连续两年评估不合格，根据《CCF专业委员会条例》的规定，该专委可以撤销也可以重组。CCF九届三次常务理事会议委托CCF学术工委，就是否保留该专委进行研究和调查。CCF学术工委为此向相关专业领域的专家征集意见、召开研讨会，调研结果认为，该专委应予以保留并重组。此次常务理事会通过了该项决议。

续的连接、网络传输延迟小、双向通信对称性强。但是这些条件在卫星通信、偏远山区等应用场景无法保证。为此，2003年英特尔研究院的研究人员提出了延迟容忍网络（Delay Tolerant Network，DTN）[11]。这是一个建立在网络传输层上的层叠网，在网络协议中增加一个称为Bundling（绑定）的协议层，提供了编址、报文封装、路由、数据重传和流量控制等机制。延迟容忍网络采用“存储—转发”工作模式，将暂时无法端到端连接的报文存储在转发结点，待重建通信路径后，再将报文发送到下一跳，实现在网络拓扑断续频繁、传输延迟长、错误率高等条件下的可靠传输。此后，研究人员又针对断续拓扑的路由技术开展了相关研究，提出多种路由算法。随后人们又将这些算法引入到传感器网络和无线网络中。延迟容忍网络技术应用的广阔前景，引起了美国国防部高级研究计划署及广大研究人员的高度重视。目前，国际互联网研究工作组成立了延迟容忍网络研究小组，不仅对延迟容忍技术，还对网络中断条件下的传输技术开展研究。

研究人员认为通过在应用层构建层叠网，结合应用需求对网络实施控制，能够有效提高网络在故障条件下的持续服务能力。例如，2001年，麻省理工学院实验室的大卫·安德森（David Andersen）等人提出了弹性层叠网（Resilient Overlay Networks）[12]，它是在现有互联网结构基础上建立的一种应用层层叠网络。弹性层叠网结点可以监控结点之间路径的质量，并根据这些信息决定是否直接利用互联网转发分组或者通过其它弹性层叠网结点转发分组。利用该网络的选路机制，可以为加入弹性层叠网的用户提供更有弹性和容错能力更强的互联网服务。实验测试表明弹性层叠网路由机制能够快速检测失效，发现路径并恢复路由。

网络威胁问题的描述能力

随着网络技术的发展，人们认识到仅从应用层考虑无法有效提高网络安全生存性，需要在网络协议体系中增加安全机制。相关研究逐步过渡到安全网络协议设计阶段。研究人员提出了多个安全协议设计，例如提出IPsec协议[2]以增强网络层的安全性，提出安全传输层协议（Transport Layer Security Protocol，TSLP）以增强传输层的安全性，提出安全超文本传输协议（Secure Hypertext Transfer Protocol，HTTPS）以增强HTTP协议的安全性。

随着网络应用的日益深入，人们发现认识网络生存环境对提高安全生存性具有重要意义。网络威胁建模成为一个重要的研究方向。早期的威胁建模研究主要采用基于规则的脆弱性建模方法，分析病毒、黑客等利用网络漏洞产生的威胁。后来，人们使用攻击树、攻击图、渗透图和Petri网等形式化工具对网络威胁进行研究，并从宏观上对网络威胁进行建模。

2001年，杜克大学的龚（F. Gong）等人提出了一种基于状态的随机网络威胁模型，建立了攻击情况下系统的运行状态集及状态转移关系。假定状态停留时间服从指数分布，相应的基于状态的随机模型具有马尔可夫性，如果给定状态之间的转移速率，就可以利用马尔可夫过程求得稳态概率分布。

2004年，意大利卡塔尼亚大学（Università degli Studi di CATANIA）的克鲁斯地（P. Crucitti）和麻省理工学院的密斯罗尼（M. Marchiori）提出了在ER随机网络和BA网络中同时考虑节点和边的随机攻击威胁模型[13]。当某个节点被攻击并从网络中去除后，将会改变网络节点之间效率最优的路径，导致负载重新分布，引起其它节点过载。结果表明，ER随机网络对随机攻击的抵抗力比BA网络强。

2006年，英国兰开斯特大学（Lancaster University）的史密斯（P. Smith）等人提出了基于行为的网络威胁模型。由于入侵检测系统主要是对

[2] 给IP和上层协议提供安全的IP协议扩展。最初是为IPv6标准而开发的，后来返过来又支持IPv4。

协议功能进行检测，因此高级攻击者会适应非正常协议功能阈值并使报警低于正常水平，导致系统难以检测攻击。通过建立正常协议行为模型，并以此作为入侵检测系统的输入，可实现对隐蔽攻击的有效检测。

增强安全防御能力

针对网络安全问题的研究经历了三个阶段：安全构件孤立防御阶段，开发出了如入侵检测系统（Intrusion Detection System，IDS）和防火墙等独立承担防御任务的安全构件；基于网络边缘的端安全技术、安全构件联动阶段，攻击的不断演化使得孤立防御难以满足安全需求，出现了入侵阻止系统之类的工具，实现了不同安全构件如入侵检测系统和防火墙之间的联动和信息反馈，同时还出现了协同入侵检测系统等研究；基于网络核心的协作安全阶段，主要在域内实现不同安全构件、不同协议层次、安全构件与网络设备之间的协同，能有效检测和阻止联合攻击行为，该阶段尚处于理论研究中。

安全构件孤立防御阶段

为了从多个层面保证网络安全，研究人员逐步引入了防病毒程序、防火墙、入侵检测系统、虚拟专用网络（Virtual Private Network，VPN）和AAA（Authentication Authorization Accounting，鉴别、授权、计费）等大量异构的孤立的单点安全防御技术。这一阶段主要研究集中在增强单个安全防御构件的性能上。1988年DEC公司最早实现的实用防火墙技术是静态防御的典型代表，其发展经历了包过滤、状态/动态检测、应用级代理和NAT（Network Address Translation，网络地址转换）防火墙技术。1987年斯坦福研究所丹宁（Denning）最早提出了入侵检测模型，从1998年至今，入侵检测方面的新理论不多，研究人员更偏重于对检测算法进行改进。

基于网络边缘的端安全技术

这个阶段主要是安全构件联动。由于原有的孤立安全防御体系中防护、检测和响应没有形成高效的闭环，各安全构件响应手段单一，缺乏联动，导致网络对联合攻击响应能力低，难以及时、准确地进行整体防御。2000年前后出现了安全构件间的联动概念。2004年思科公司提出“自防御网络”的“网络准入控制（Network Access Control，NAC）”。网络中的每个组件都作为防御点，各种组件联动工作，根据终端是否符合安全策略，通过路由器和交换机等来决定它们是否接入网络。2003年众多著名IT企业发起组建了可信计算组织（Trusted Computing Group，TCG），通过可信网络连接解决接入终端引起的网络安全问题。

2005年，哥伦比亚大学洛卡斯托（M.E.Locasto）等人提出了P2P[3]结构的协同入侵检测系统（Collaborative Intrusion Detection System，CIDS）[14]，使用布隆过滤器（Bloom Filter）保留隐私，并采用动态层叠网（Overlay）进行分布式数据关联。协同入侵检测系统中的每个实体都采用网络化的入侵检测系统（Network Instrusion Detection System）产生“观察表（Watchlist）”，使用分布式关联调度算法实现信息交换，以实现入侵监测的协同。

2005年，剑桥大学和微软实验室提出了基于主机的协同式蠕虫抑制系统Vigilante。系统中的每台主机不需要相互信任，主机运行指定的软件检测蠕虫，并通过Pastry结构的层叠网传播自证明报警（Self-Certifying Alert），接收报警主机对将要到来的蠕虫消息进行过滤以抑制蠕虫。

基于网络核心的协作安全阶段

2003年，美国电话电报公司实验室马哈扬（Mahajan）和弗洛伊德（Floyd）等人，在核心路由器增加集中拥塞控制ACC（Advanced Clock

[3] Peer to Peer，对等计算。

Calibration，高级时钟校准）功能，检测由分布式拒绝服务攻击（Distribution Denial of Service，DDoS）引起的拥塞，并把从速率上限制（Rate Limiting）高带宽的集中作为响应的手段。通过研究识别合法流量并使其得到优先处理。不能处理流量集中的拥塞路由器通过Pushback机制向邻近的ACC路由器发布速率限制的请求。这种协同在进行检测时体现出较大的优越性，形成了一个分布式的防范分布式拒绝服务攻击的体系。

2004年美国南加州大学陈璇（Xuan Chen，音译）和海德曼（Heidemann）提出了一个基于网络路由器端的蠕虫检测系统（Detector for Early Worm Propagation，DEWP）。其方法基于以下观察：由于大多数蠕虫都是针对某一种网络服务的漏洞，因此当这种蠕虫爆发时，在路由器的两个方向上都会出现大量与目的端口相同的网络流量，并且不同目的地址的数量急剧增加，当增长量超过一定阈值时就认为发生蠕虫。蠕虫检测系统主要包含蠕虫检测和包过滤功能。这种从网络核心进行检测、防御的方法对于大规模的蠕虫防御具有很好的效果，其不足之处在于用目前的算法可能会产生误报。

事实证明，安全措施叠加并不能从根本上解决网络的安全问题，反而会带来一些问题：缺乏灵活性，难以支持移动性和新协议；自身的脆弱性，改变某一构件上的安全规则往往会破坏整体安全策略；容易造成混乱，需要维护太多的状态，安全技术本身可能引发新的安全问题。

目前，缺乏一个完整的安全体系结构，来满足多种不同防护技术的要求。上述孤立的安全要素还无法组合起来为用户提供良好的全面安全保障。不同安全设备具有不同的安全防护功能，管理与协作面临巨大的复杂性。需要从网络核心采用协同的方法，将众多安全要素有机联合起来，达到网络安全最大效果。

因此，笔者认为从网络核心开始，结合已有安全系统方面的研究成果，进一步深化设备、协议层次、域间的协同技术，是未来安全技术的发展方向。

增强型路由技术

提高网络生存能力的重要途径是设计灵活高效的路由机制，从而提高网络对故障的动态适应与屏蔽能力，减小故障对正常通信的影响。例如，普林斯顿大学的费姆斯特（N. Feamster）在2005年提出了RCP（Routing Control Platform）路由体系结构[15]。与传统路由机制不同，RCP将域间路由从IP路由器中分离出来，路由器主要进行报文转发。在每个自治系统内，由一个分离的路由控制平台代替自治系统内的路由器选择路由，并和其它域交换可达信息。RCP体系结构能克服许多与内部边界网关协议（Border Gateway Protocol，BGP）相关的问题，例如转发环路和信令分割。RCP还简化了底层的路由机制和配置语言，更易于实施流量工程，但这一结构引入了影响健壮性、可扩展性、收敛速度及路由一致性等的潜在危险。

近年来，研究人员考虑在没有全局收敛的情况下实现快速路由，提出了FIR（Failure Insensitive Routing，故障非敏感路由）、FIFR（Failure Inferencing Based Fast Rerouting，基于故障推理的重路由）、FCP（Failure-Carrying Packets，故障携带报文）等新型算法。其中，FCP[4]算法[16]能够避免网络变化在全网范围内的广播式扩散，有效防止路由收敛过程导致协议处理开销增加、产生瞬时环路等问题。但是，该算法存在路由计算开销大，故障条件下路由选择非最优化等问题。

2008年，乔治亚理工学院的莫蒂瓦拉（M. Motiwala）等人提出了路径拼接技术[17]，当出现网络故障时，利用底层网络拓扑的多样性，动态选择结点间的实际网络路径，从而达到动态形成路由路径的目的，使流量可以在中间节点改变路径，达到

[4] 加州大学伯克利分校的拉克斯铭亚南（K. Lakshminarayanan）在2007年提出。

高可靠性和快速恢复性能。

结语

下一代网络体系结构的研究仍在进行，GENI也进入了具体实施阶段。虽然目前的进展没有设想的顺利，但学者们都在努力研究，希望克服当前互联网技术存在的诸多问题，提高互联网的安全性、生存性和扩展性。■

参与本文工作的还有国防科学技术大学计算机学院助理研究员赵锋。

苏金树

CCF高级会员。国防科学技术大学计算机学院教授。主要研究方向为计算机网络、数据链和信息安全。

sjs@nudt.edu.cn

吴纯青

国防科学技术大学计算机学院教授。主要研究方向为计算机网络、卫星网络和信息安全。

xixiwu2001@yahoo.com.cn

胡晓峰

CCF高级会员。国防科学技术大学计算机学院副研究员。主要研究方向为高性能路由交换。xfhu@nudt.edu.cn

彭　伟

国防科学技术大学计算机学院副研究员。主要研究方向为移动自组织网络和传感器网络技术。

wp@ nudt.edu.cn

参考文献

[1] GENI: global environment for network innovations. http://www.geni.net
[2] FIND: future Internet network design. http://find.isi.edu
[3] FIRE: Future Internet Research and Experimentation. http://cordis.europa.eu/fp7/ict/fire/
[4] AKARI Architecture Conceptual Design. http://akari-project.nict.go.jp
[5] AGN REQUEST FOR INFORMATION – Assurable Global Networking. Defense Advanced Research Projects Agency's (DARPA) Strategic Technology Office (STO),2007
[6] D. Clark, C. Partridge , J. C. Ramming, et al , A Knowledge Plane for the Internet. Proc. of ACM SIGCOMM'03 , 2003
[7] P. Jokela, P. Nikander, J. Melen, J. Ylitalo, J. Wall, Host identity protocol - extended abstract. Wireless World Research Forum, 2004
[8] A. Greenberg, G. Hjalmtysson, D. A.Maltz, A. Myers, J. Rexford, G. Xie, H. Yan, J. Zhan, and H. Zhang. A clean slate 4D approach to network control and management. SIGCOMM Comput. Commun. Rev., 2005, 35(5):41~54
[9] M. Casado, T. Garfinkel, A. Akella, M. J. Freedman, D. Boneh, N. McKeown, and S. Shenker, SANE: A protection architecture for enterprise networks. USENIX Security Symposium, 2006
[10] M. Ethane. Casado, M. J. Freedman, J. Pettit, J. Luo, N. McKeown and S. Shenker, Ethane: Taking Control of the Enterprise. Proc. of ACM SIGCOMM '07, 2007
[11] K. Fall, A Delay-Tolerant Network Architecture for Challenged Internets. Proceedings of ACM SIGCOMM'03, 2003
[12] D. Andersen, H. Balakrishnan, M. Kaashoek, R. Morris, Resilient overlay networks. Proc. of SOSP, 2001
[13] P. Crucitti,V.Latora, and M. Marchiori, Model for cascading failures in complex networks. PHYSICAL REVIEW E 69, 045104(R),2004
[14] M.E. Locasto, J. J. Parekh, A.D. Keromytis, S.J. Stolfo, Towards Collaborative Security and P2P Intrusion Detection. Proceedings of the 2005 IEEE Workshop on Information Assurance and Security, 2005
[15] M. Caesar, D. Caldwell, N. Feamster, J. Rexford, A. Shaikh, and J. V.Merwe, Design and implementation of RCP. Proc. of NSDI '05, 2005
[16] K. Lakshminarayanan, M. Caesar, M.Rangan, et al , Achieving Convergence-Free Routing using Failure-Carrying Packets. Proc. of ACM SIGCOMM '07, 2007
[17] M. Motiwala, M. Elmore , N.Feamster, and S. Vempala. Path splicing. Proc. of ACM SIGCOMM '08, 2008