Guest Editorial
Recent Advances in Service Overlay Networks

I. Service Overlay Networks

THE best effort Internet was designed when host connectivity was the primary concern. To achieve this purpose, the Internet followed a simple and robust design philosophy and has become a phenomenal success in the last decade. Today's Internet users are very different from those of a decade ago and their expectations with regard to the Internet have changed considerably. Users are now more concerned about timely delivery of desired services than host connectivity. The role of the Internet has also evolved; the Internet has become a commercial infrastructure for service delivery. However, due to the Internet protocol (IP)'s addressing scheme, routing paradigm, and other historical reasons, it is currently not well suited for the purpose of service delivery. In response to these challenges, various forms of overlay networks have been proposed and some deployed over the Internet. For example, commercial content delivery service providers have used proprietary server networks over the existing Internet to bring content closer to users. In the meantime, grassroots users have actively contributed their computers to form various kinds of peer-to-peer networks for file swapping and content sharing.

The purpose of this issue is to disseminate state-of-the-art research results that address service delivery over the Internet with overlay networks.

II. Overview of the Issue

The 15 papers in this issue fall into the following areas: network construction, routing and discovery, measurement, overlay multicasting, performance evaluations and design tradeoffs, security and resilience.

A. *Network Construction*

The first three papers in this special issue investigate architectural issues in overlay networks. Internet flash crowds, also known as "hot spots," are a phenomenon that result from a sudden, unpredicted increase in an online object's popularity. When content reaches its apex in popularity, it becomes unavailable to the majority of users who seek it. One approach to address this problem is to have clients form a peer-to-peer (P2P) overlay network that allows those clients who have received copies of popular content to forward it to those clients who also desire but have not yet received it. The first paper in this issue, "A Lightweight, Robust P2P System to Handle Flash Crowds," describes one implementation of such an approach. This paper proposes to use randomized overlay construction and randomized, scoped searches to efficiently locate and deliver objects that are under heavy demand to all users who desire them. The authors show via a mixture of theoretical results, simulation, and experimentation that by relying on randomness, their proposed P2P implementation can achieve low latency delivery using modest traffic levels, even when membership to the overlay changes dynamically with time and when there exist members that limit their participation in the system.

In the second paper, "A Construction of Locality-Aware Overlay Network: mOverlay and its Performance," the authors propose to construct an overlay network by exploiting the locality in the underlying network using group concept. This approach is based on the observation that a complete decoupling of overlay topology and underlying data network topology may not be completely desirable. For instance, if the P2P overlay is constructed randomly, hosts that are nearby in the overlay network could, in fact, be far away from the underlying data network, which translates into inefficiency in network resource use and degradation in user-perceived performance. The authors propose a new protocol to construct an unstructured overlay network to achieve local characteristics.

The third paper, "QRON: QoS-Aware Routing in Overlay Networks," envisions an overlay service network constructed by deploying so-called "overlay brokers" (OBs) in each autonomous systems. By having these OBs cooperate with each other, an overlay service network can perform various network services to support other applications, such as resource allocation and routing. The authors focus on quality-of-service (QoS)-aware routing protocols within such overlay networks; they adopt a hierarchical approach in their design of routing algorithms to achieve scalability. Through simulations, the authors show that their routing algorithms can effectively find and provide QoS-satisfied overlay paths and can balance the overlay traffic burden among OBs and overlay links.

B. *Routing and Discovery*

The next two papers focus on routing and discovery for service overlay networks. The fist paper, "Tapestry: A Resilient Global-Scale Overlay for Service Deployment," presents a P2P overlay routing infrastructure offering location-independent routing of messages directly to nearby copies of an object or service using only localized resources. Tapestry supports a generic decentralized object location and routing (DOLR) applications programming interface (API) that uses a self-repairing, soft-state-based routing layer. This paper presents the Tapestry architecture, algorithms, and implementation. It also explores the behavior of Tapestry deployment on a global testbed under a variety of conditions.

The second paper in this area, "Design and Evaluation of a Distributed Scalable Content Discovery System," focuses on

Digital Object Identifier 10.1109/JSAC.2003.818774

a content discovery system (CDS), which is used by nodes in the system to discover contents published by some other nodes in the system. The main challenge in the design of a CDS is to achieve both rich functionality and scalability. This paper presents a distributed and scalable CDS that uses rendezvous points (RPs) that avoid network-wide message flooding at both registration and query time. This paper also presents a mechanism that uses load balancing matrices (LBMs) to dynamically balance both registration and query load in a system in order to improve the system's throughput under skewed load.

C. Measurement

Measurement is an important component for operation and management of overlay networks. The next two papers address the techniques and infrastructure for conducting measurements. The first paper in this area, "Computing the Unmeasured: An Algebraic Approach to Internet Mapping," investigates algorithms for distance estimation. The authors present an algorithm that increases the effectiveness of end-to-end distance measurements at no additional overhead. Given end-to-end distance measurements and routes along which the measurements are conducted, the proposed algorithm computes additional distances, both to and between intermediate nodes. This information can be used to build a more detailed map, which provides better distance estimation.

The second paper, "Overlay Networks of *In Situ* Instruments for Probabilistic Guarantees on Message Delays in Wide-Area Networks," proposes an overlay network of in-situ instruments over a wide-area network to collect delay measurements, compute paths, and route messages. The overlay network is implemented using user-level daemons that realize paths among themselves without explicit support from underlying network routers. The authors propose regression methods to compute a path whose message delay is close to the optimal expected delay with a high probability, and which is based entirely on measurements.

D. Overlay Multicasting

An important application for overlay networks is to provide multicast services. There are three papers in this issue that address this subject. The first paper, "oStream: Asynchronous Streaming Multicast in Application-Layer Overlay Networks," proposes that asynchronous multicast be used to take advantage of the strong buffering capabilities of end hosts in application-layer overlay networks. Based on this concept, the authors propose an overlay multicast strategy, oStream, to address the on-demand distribution problem. Through a mixture of analytical and experimental results, the authors show that the required server bandwidth of oStream defeats the theoretical lower bound of traditional multicast-based solutions. Furthermore, with respect to bandwidth consumption on the backbone network, the benefit introduced by oStream overshadows the topological inefficiency of application overlay.

The second paper in this area, "Multicast With Network Coding in Application-Layer Overlay Networks," seeks to improve end-to-end throughput in application-layer multicast by exploiting the unique processing capability and topological characteristics of overlay networks. The authors propose to apply a network coding mechanism on intermediate overlay nodes and suggest construction of a two-redundant multicast graph as the multicast topology, on which network coding will be applied. The authors design a set of distributed algorithms intended to construct such multicast graphs and to assign linear codes and apply network coding. By using both analytical and simulation results, the authors show that their algorithm is able to increase end-to-end session throughput.

The third paper, "A Peer-to-Peer Architecture for Media Streaming," investigates the applicability of P2P to the problem of streaming live media. There are several problems that need to be addressed in this area, including end-to-end delay, unpredictable receiver behavior, and control overhead. The authors propose a solution that organizes receivers into a hierarchy of clusters, and builds a multicast tree atop this hierarchy according to a set of rules. By using both theoretical and simulation results, the authors show that the proposed solution is scalable, robust, and efficient for a large-scale P2P streaming systems.

E. Performance Evaluations and Design Tradeoffs

The next three papers in this issue investigate some fundamental performance issues and design tradeoffs in overlay networks. The first paper, "On Expiration-Based Hierarchical Caching Systems," builds a simple model for a time-to-live (TTL)-based hierarchical caching system. By analyzing the intrinsic timing behavior of the basic model, the authors develop important performance metrics from the perspective of both caching systems and users. Then, the authors propose several directions that might be taken to enhance the performance of the basic model, concluding with insights on various system design tradeoffs.

The second paper, "On the Fundamental Tradeoffs Between Routing Table Size and Network Diameter in Peer-to-Peer Networks," studies the tradeoff between routing table size and network diameter in designing distributed hash table (DHT) in P2P networks. There are several findings in this paper. First, the authors show that there are straightforward routing algorithms that achieve better asymptotic tradeoffs than existing algorithms. Next, the authors prove that such tradeoffs are asymptotic optimal for uniform algorithms. The authors then study the exact (instead of asymptotic) optimal tradeoffs for uniform algorithms, going on to propose a new routing algorithm that reduces the routing table size and network diameter of Chord. Finally, the authors present a congestion-free, nonuniform algorithm that achieves a better asymptotic "state-efficiency" tradeoff than existing schemes in the probabilistic sense.

The third paper, "Evaluation of Architectures for Reliable Server Pooling in Wired and Wireless Environments," investigates reliable server pooling (RSP), which is used to allow a pool of redundant information sources to be viewed as a single transport endpoint so as to provide persistent connections and balanced traffic for different applications. The Internet Engineering Task Force (IETF) RSerPool working group has proposed an architecture to implement RSP that defines an overlay network providing an upper layer protocol or an application with a range of reliability services. Via simulations conducted in both wired and wireless environments, this paper shows that the current version of the RSerPool works well only in fixed and relatively reliable environments, and that its performance worsens rapidly as the networks become more unreliable or mobile. The authors identify potential issues that may cause such performance degradation, and then propose alternative design options to improve RSPs performance in wireless and mobile environments.

F. Security and Resilience

The last two papers investigate security and resilience issues. The first paper, "SOS: An Architecture for Mitigating DDoS Attacks," proposes an architecture that proactively prevents DoS attacks, in which communication is between a predetermined location and a set of well-known users, located in the wide-area network, who have authorization to communicate with that location. The proposed architecture uses a combination of secure overlay tunneling, routing via consistent hashing, and filtering. The probability of successful attacks is reduced by performing intensive filtering near protected network edges and by introducing randomness and anonymity in the forwarding architecture. Through simple analytical models, the authors show that DoS attacks directed against the proposed architecture have a negligible probability of disrupting communication between two parties.

The second paper, "Resilient Self-Organizing Overlay Networks for Security Update Delivery," investigates how to rapidly and widely disseminate security updates throughout the Internet. A successful system must out-pace the propagation of threats, handle complexities in a large-scale environment, deal with interruption attacks on dissemination, and also secure itself. This paper addresses these problems by proposing a large-scale, self-organizing, and resilient overlay network that resides on top of the Internet. The authors discuss how to secure the dissemination procedure and the overlay network, and consider possible attacks and countermeasures.

ACKNOWLEDGMENT

The guest editors would like to thank L. Peterson, JSAC Board Representative, and S. McDonald, from the JSAC editorial office, for their advice and assistance in the preparation of this special issue.

BO LI, *Leading Guest Editor*
Hong Kong University of Science and Technology
Department of Computer Science
Kowloon, Hong Kong

Y. THOMAS HOU, *Guest Editor*
Virginia Polytechnic Institute and State University
The Bradley Department of Electrical and Computer Engineering
Blacksburg, VA 24061 USA

KAZEM SOHRABY, *Guest Editor*
University of Arkansas
Department of Computer Science and Computer Engineering
Fayetteville, AR 72701 USA

MEHMET ULEMA, *Guest Editor*
Manhattan College
Department of Computer Information Systems
Riverdale, NY 10471 USA

ZHENSHENG ZHANG, *Guest Editor*
San Diego Research Center
San Diego, CA 92108 USA

LARRY L. PETERSON, *J-SAC Board Representative*

REFERENCES

[1] "Web caching and content delivery," in *2001 Int. Web Content Caching and Distribution Workshop (WCW)*, A. Bestavros and M. Rabinovich, Eds., Boston, MA, June 2001.

[2] E. Biersack, P. Danzig, and E.W. Zegura (Guest Editors), "Special issue on Internet proxy services," *IEEE J. Select. Areas Commun.*, vol. 20, no. 7, Sept. 2002.

[3] "Peer-to-peer systems," presented at the 1st Int. Workshop on peer-to-peer systems (IPTPS 2002) Lecture Notes in Computer Science 2429, P. Druschel, A. Rowstron, and F. Kasshoek, Eds., Cambridge, MA, Mar. 7–8, 2002.

[4] B. Krishnamurthy and J. Rexford, *Web Protocols and Practice: HTTP/1.1, Networking Protocols, Caching, and Traffic Measurement*. Reading, MA: Addison-Wesley, 2001.

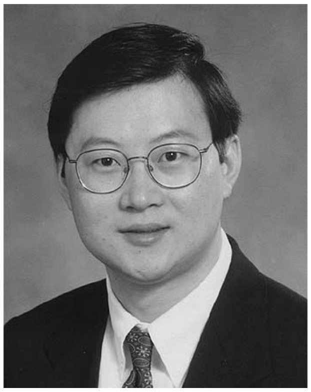
[5] A. Oram, Ed., *Peer-to-Peer: Harnessing the Power of Disruptive Technologies*. Sebastopol, CA: O'Reilly & Associates, 2001.

[6] M. Rabinovich and O. Spatscheck, *Web Caching and Replication*. Reading, MA: Addison-Wesley, 2001.

Bo Li (S'89–M'92–SM'99) received the B.S. and M.S. degrees in computer science from Tsinghua University, Beijing, China, in 1987 and 1989, respectively, and the Ph.D. degree in computer engineering from the University of Massachusetts, Amherst, in 1993.

From 1994 and 1996, he worked on high-performance routers and ATM switches in the IBM Networking System Division, Research Triangle Park, NC. Since January 1996, he has been with the Computer Science Department, Hong Kong University of Science and Technology, Kowloon, where he is now an Associate Professor. He is also an Adjunct Researcher at Microsoft Research Asia (MSRA), Beijing, China. His current research interests include wireless and mobile networks supporting multimedia, video multicast, and all-optical networks with WDM. He has published over 140 technical papers in refereed journals and conference proceedings. He has been an Editor or a Guest Editor for 14 journals and involved in the organization of over 30 conferences.

Dr. Li currently serves as the Chair of the Technical Program Committee of the IEEE INFOCOM 2004. He is a Member of the Association for Computing Machinery (ACM).



Y. Thomas Hou (S'91–M'98) received the B.E. degree (*summa cum laude*) from the City College of New York, New York, NY, in 1991, the M.S. degree from Columbia University, New York, NY, in 1993, and the Ph.D. degree from Polytechnic University, Brooklyn, NY, in 1998, all in electrical engineering.

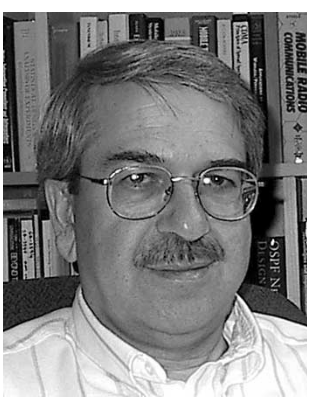
From 1997 to 2002, he was a Research Scientist and Project Leader in the IP Networking Research Department, Fujitsu Laboratories of America, Sunnyvale, CA (Silicon Valley). He is currently an Assistant Professor in the Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA. His research interests include wireless sensor networks, multimedia delivery over wireless networks, scalable architectures, protocols, and implementations for differentiated services Internet, and service overlay networking.

Dr. Hou has published extensively in the above areas and is a corecipient of the 2002 IEEE International Conference on Network Protocols Best Paper Award and the 2001 IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY Best Paper Award. He is a Member of the Association for Computing Machinery (ACM).

Kazem Sohraby (S'74–M'78–SM'89) received the B.S., M.S., and Ph.D. degrees in electrical engineering, and the M.B.A. degree from the Wharton School, University of Pennsylvania, Philadelphia.

He is a Professor and Head of the Department of Computer Science and Computer Engineering, College of Engineering, University of Arkansas, Fayetteville. Before joining academia, he was with Bell Laboratories, Holmdel, NJ. His areas of interest include computer networking, signaling, switching, performance analysis, and traffic theory. He has over 20 patents on computer protocols, wireless and optical systems, circuit and packet switching, and optical Internet. He has numerous publications, including a book titled *Control and Performance in Packet, Circuit, and ATM Networks* (New York, NY: Kluwer, 1995).

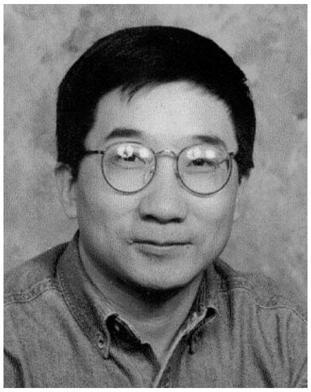
Dr. Sohraby is a Distinguished Lecturer of the IEEE Communications Society and serves as its President's representative on the Committee on Communications and Information Policy (CCIP). He served on the Education Committee of the IEEE Communications Society, is on the Editorial Boards of several publications, and Referee and Reviewer with the National Science Foundation. He served as Grant Referee and Reviewer with the U.S. Army and Natural Sciences and Engineering Research Council of Canada.



Mehmet Ulema (S'77–M'80–SM'93) received the B.S. and M.S. degrees at Technical University of Istanbul, Istanbul, Turkey, and the M.S. and Ph.D. degrees in computer science at Polytechnic University, Brooklyn, NY.

He has more than 25 years experience in the telecommunications field as a Professor, Director, Project Manager, Researcher, Systems Engineer, Network Architect, and Software Developer. Currently, he is a Professor at the Computer Information Systems Department, Manhattan College, New York, NY. Previously, he held Management and Technical Positions in Daewoo Telecom, Middletown, NJ, Bellcore, Red Bank, NJ, AT&T Bell Laboratories, Holmdel, NJ, and Hazeltine Corporations, Long Island, NY. He was also an Adjunct Professor at City College of New York and Stevens Institute of Technology, Hoboken, NJ. He has numerous publications in various international conferences and journals and holds two patents. He gave a number of talks and tutorials on Network Management and Wireless Networks and organized several special issues.

Dr. Ulema is the current Chair and Cofounder of the IEEE Communications Society's Information Infrastructure Technical Committee. Previously, he served as the Chair of the IEEE Communications Society's Radio Communications Technical Committee. He is involved in numerous IEEE conferences. Currently, he is the Technical Program Chair for a new IEEE ComSoc initiative called 2004 IEEE Consumer Communications and Networking Conference, and Technical Program Co-Chair of ICC 2006. More recently, he served as the Technical Program Co-Chair of the IEEE Symposium in Computer and Communications (ISCC 2000) and the IEEE NOMS 2002. He is on the Editorial Boards of the *IEEE Communications Magazine*, the *ACM Wireless Network Journal*, and the *Journal of Communications and Networks*.



Zhensheng Zhang (M'90–SM'01) received the Ph.D. degree in electrical engineering from the University of California, Los Angeles.

He is currently with San Diego Research Center, San Diego, CA. He visited Microsoft Research Asia, Beijing, China, in the summer of 2002, and worked at Sorrento Networks, Department of System Architecture, San Diego, CA, for two years, responsible for designing the next-generation optical metro networks using the GMPLS control framework. Prior to joining Sorrento Networks, he was with Bell Laboratories, Lucent Technologies, Murray Hill, NJ, focusing on research and development in ATM/SONET infrastructure and IP over WDM. He has published more than 100 papers in the IEEE TRANSACTIONS ON NETWORKING, the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, the IEEE TRANSACTIONS ON COMMUNICATIONS, and key IEEE conferences. His research interests include wireless networks, IP over WDM networks, GMPLS, and peer-to-peer networks.

Dr. Zhang was a Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS for the Special Issue on IP Over DWDM Networks, October 2000, and the *Journal of Wireless Networks* issue on Multimedia Wireless Networks, August 1996. Currently, he is Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and on the Editorial Board of *Optical Networks Magazine*. He also served as the Chair of Technical Program Committee for OptiComm 2003 Conference, Dallas, TX, 2003. He is a recipient of the 1988 Phi Beta Kappa Scholarship Award.