NETWORK AND TRANSPORT PROTOCOLS

Ghalib A. Shah and **Ozgur B. Akan**
Next-generation and Wireless Communication Laboratory (NWCL)
Department of Electrical & Electronics Engineering, Koc University, Turkey, 34450

Keywords: Internet, wireless networks, Quality of Service (QoS), IPv4, IPv6, Mobile IP, TCP, TCP for wireless networks, Multimedia communication, Internet of Things.

Contents

Summary

Surveys on Internet usage suggest that the dramatic growth of wireless infrastructure and the devices will eventually replace the wired network access in future. Therefore, this chapter discusses the fundamental network and transport protocols of Internet and then investigates their support for wireless networks and identifies the issues that need to be addressed for successful transition from wired to wireless Internet. The fundamental network, i.e., IP, and transport protocols, i.e., TCP and UDP, which have become the de facto standard for Internet are described in more detail to understand the functioning of current Internet. Moreover, the quality of service requirements of the ever increasing growth of diverse applications is also investigated and its support in current Internet protocols suit is explored. It is analyzed that the existing Quality of Service (QoS) architectures are not suitable for wireless network and new architectures need to be explored by taking into account the quality of links exhibited in the wireless network to support consistent QoS. Moreover, a new paradigm Internet of Things (IoT) has recently been introduced in which it is envisioned to provide all of the everyday devices, with possibly wireless interface facing some resource constraints, connectivity with each other or to the Internet. These end devices are heterogeneous and usually mobile without having any continuous source of power. This raises new challenges to incorporate the energy efficiency and mobility in the design of network and transport protocols. Therefore, the chapter also highlights the current efforts on the design and development of new networking protocols to support this emerging paradigm of IoT.

1. Introduction

The growth of modern communication infrastructures, such as, the Internet, various broadband wireless networks and the realization of *4G* over the last decade has surpassed many expectations. This growth is mainly fostered due to the rapid deployment of heterogeneous broadband networks that mainly include IEEE 802.11 wireless local area networks (WLANs), IEEE 802.16 metropolitan area wireless networks and cellular mobile networks (LTE). Cisco forecast suggests that Wi-Fi devices will for the first time use more bandwidth of global IP traffic than wired devices and will consume 37.2 Exabytes of data worldwide per month in 2015, taking up 46.2 percent of all IP traffic. On the other hand, the share of wired IP traffic will sink from 63 percent in 2010 to 46.1 percent in 2015. Similarly, mobile IP traffic will grow 26 times over the period, compared to three-fold growth for wired and five-fold growth for Wi-Fi traffic and will make up 8 percent of global IP traffic in 2015. These trends suggest the dramatic growth of wireless infrastructure as shown in Figure 1 and the devices that will eventually replace the wired network access in future. Therefore, this chapter first discusses the fundamental network and transport protocols of Internet and then investigates their support for wireless networks with the proposed enhancements.

The wide availability of wireless network access is leveraged by a variety of services and applications, which are rapidly evolving to facilitate users in enormous ways, such as, information or knowledge sharing, entertainment, healthcare, online auctioning, shopping, banking, to mention few. Apart from the traditional infrastructure based network architecture, the technological advancements in networking have also permitted the users to spontaneously form ad hoc networks for various reasons. Wireless networks operating in open-spectrum bands will be used in environments characterized by higher levels of interference, capture and disruption phenomena. This deteriorates the performance of widely recognized networking and transport protocol (TCP/IP), a de facto transport and network layer protocols in the Internet protocols suit of wired networks. Moreover, the applications and services running over the modern heterogeneous networks are posing many different kinds of requirements on underlying protocols. Hence, the networking protocol faces different challenges that mainly include the heterogeneity of end devices, mobility, scalability and diverse quality of service (QoS) requirements. Communication under highly variable network conditions will require the transport layer protocol to handle not only packet losses due to network congestion but also transmission errors in addition to supporting QoS requirements imposed by the user applications.

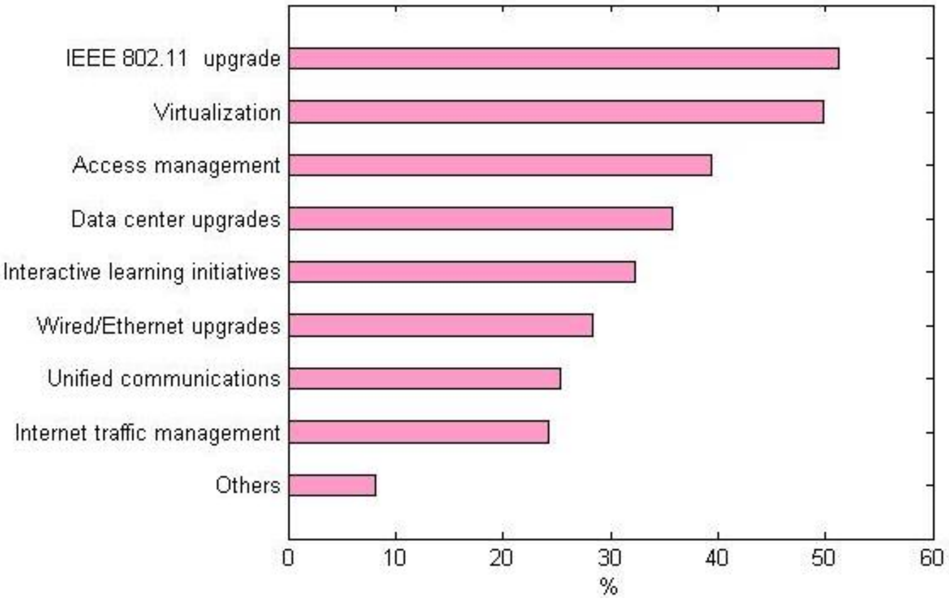

Figure 1. Cisco forecasts the wired and wireless share of Global IP traffic in 2015.

This chapter investigates the design and performance of existing networking protocols (IP) as well as TCP and TCP-friendly transport protocols in the modern heterogeneous networks and also their prospects for multimedia communication. Furthermore, the applications domain of Internet has dramatically changed the type of end devices that introduces a new paradigm, beyond the traditional Internet paradigm, called Internet of Things (IoT). Hence, this new paradigm will also be discussed with the current progress towards its design, applications and standardization.

1.1. Network Architecture

Wireless networks are generally classified as infrastructure based networks and ad hoc networks depending on the type of users and services.

1.1.1. Infrastructure Mode Network

In Infrastructure mode wireless networks, all the end devices or mobile stations are controlled and centrally coordinated with the help of Access Point (AP), also called the base station. Wireless APs are usually routers or switches which are connected to Internet by a broadband modem or Ethernet.

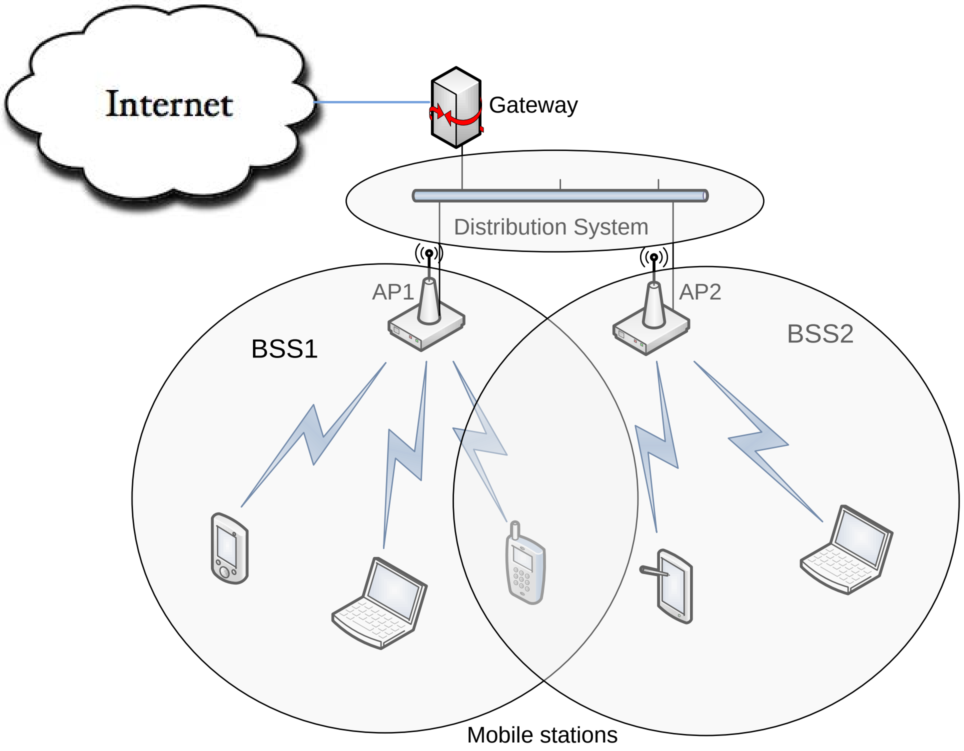

Figure 2. Infrastructure mode wireless network scenario.

Infrastructure mode deployments are more suitable for larger organizations or facility to provide continuous service or Internet access to the users with ease of configuration and mobility. This kind of deployment is made to substitute the conventional local area network with the wireless network to simplify the network management operations, and allows the facility to address operational concerns.

The basic networking unit of the infrastructure based architecture is the basic service set (BSS), which is a set of all devices that can communicate with each other or access Internet as shown in the example scenario in Figure 2. Every BSS has an identification (ID) called the BSSID, which is the MAC address of the access point servicing the BSS. The coverage in wireless network is extended by deploying as many access points as required that forms the extended service set (ESS). Access points in an ESS are connected by a backbone distribution system. A distribution system (DS) connects access points in an extended service set. Thus, the distribution system allows the mobile users to move from the coverage area of one AP to the other without losing connectivity. Each ESS also has an ID called the SSID, which is a 32-byte (maximum) character string. This configuration is similar to the architecture of cellular networks except that the deployment of APs do not strictly adhere the hexagonal architecture of cellular networks.

1.1.2. Ad Hoc Networks

Ad hoc network is created spontaneously to enable a group of users to share resources without the support of infrastructure and therefore, forms an independent basic service set (IBSS) of wireless network. Ad-hoc mode allows the wireless devices within the range of each other to discover and directly communicate in a peer-to-peer fashion without involving central access points. For setting up ad hoc mode, all the participating devices are required to be manually configured at ad hoc mode instead of infrastructure mode, and all adapters must use the same channel and same SSID for making the connection active. However, the performance of ad hoc network suffers with the growth of number of devices. Disconnections of random device may occur frequently and also, it is a tedious task for the administrator to manage the network in ad hoc mode. Another limitation of ad hoc network is that they cannot bridge to wired local area network and also cannot access Internet without the support of any specialized gateway installed at any of the participating device. However, ad hoc networks are very useful in a small environment when there is no open access to Internet and a group of users wants to share data, for example, airports, public parks, shopping malls, classrooms, meeting halls, etc. Thus, ad hoc mode does not need any extra access point and therefore, it reduces the cost and can be set up anywhere any time. The life of ad hoc networks is usually limited to the battery time of the participating devices.

1.2. What is carried on Networks?

Initially, the Internet applications were developed for Web, FTP, Email, Telnet, etc, which generates bursty data with lesser or no timing constraint but some reliability requirements for FTP, Email. However, a tremendous growth has been observed in last decade to many different kinds of applications, such as, online gaming, shopping, banking, stock trading, data storage, real-time streaming, video on demand (VoD), voice over IP (VoIP), peer-to-peer instant messaging, file sharing etc, that has resulted in a diverse nature of traffic on these networks. Thus, the current applications traffic is classified into bursty, interactive, real-time and non-real-time. The bursty traffic is characterized as uneven pattern of data transmission in a cycle of burst of packets and silent periods, for example, FTP, VoD. The interactive traffic is produced by sessions that consist of comparatively short request/response pairs with small number of packets, such as, web browsing, telnet sessions, messaging etc. The real-time traffic is more demanding and puts latency constraint on the delivery of data that the operator has to meet. This includes VoIP, online gaming, video conferencing, live streaming. On the other hand, non real-time traffic does not require timely delivery but secure and reliable data delivery is highly important that mainly includes banking, online shopping.

1.3. What are the User Demands?

Based on the type of applications, users lay down different requirements that the service providers are obliged to support according to the service agreement. These requirements are usually specified as quality of service (QoS) using various performance metrics that include response time, bandwidth, latency, jitter, reliability, security and cost. These are defined as follows:

(a) *Response time:* The user's expected response time is the time elapsed between sending a request and the reception of the first response by the user.

(b) *Latency (delay):* Latency refers to the network transmit time that elapses between the emission of the data by the application to the reception at the receiving end-system.

(c) *Jitter:* Jitter is the measure of the variation in latency that the receiver observes in receiving data by the sender.

(d) *Data rate:* Data rate refers to the raw data rate generated by an application, for example, encoded multimedia data before transmission.

(e) *Bandwidth:* Bandwidth is defined as the required data transfer rate, measured in bits per second that also account for the overhead in delivering the application data.

(f) *Reliability:* It is defined as the proportion of the transmitted data packets that has been successfully and error-free received by the receiver.

(g) *Security:* It refers to the protection of transmitted data from an unauthorized access and alteration except the intended authorized receiver.

(h) *Cost:* Though it is not related to the application QoS but it is an important metric that user considers in selecting service provider. It has different models, e.g., bits/dollars, bandwidth leased with fixed period.

In addition to the QoS specifications for a service, quality of experience (QoE) is a new subjective measure that represents the user's experiences with a service. QoE cannot be taken simply the effective quality of service but must also take into consideration every factor that contributes to overall user experience such as suitableness, flexibility, mobility, security, cost, personalization and choice. Apart from its dependence to user, it is also influenced by the user's device (for example the wireless device support to 3.5G for higher rate), his/her environment (urban/rural, mobility), his expectations (cellular or corded telephone), the nature of the content and its importance (a simple text message or live streaming). Mean Opinion Score (MOS) is the assessment scale of QoE commonly used for video streaming and VoIP services. MOS has been expanded to assess the quality of television over ADSL (IPTV), by observing jittering, blockness and artifacts of a viewed video signal.

1.4. Internet Protocol Suite

The Internet protocol suite organizes the set of protocols and methods into four layers to facilitate the growth and enhancements, according to [RFC 1122]. The layered design of protocol suit consists of four layers; the link layer, the Internet layer, the transport layer, and the application layer from bottom to top as illustrated in Figure 3. This model was not intended to be a rigid reference model into which new protocols have to fit in order to be accepted as a standard. Application layer defines a set of functions and specifications in order to help user developing different applications to run over Internet. There are many applications available in the TCP/IP suite of protocols. The most commonly used are hyper text transfer protocol (HTTP) for web, file transfer protocol (FTP) for files exchange, simple mail transfer protocol (SMTP) for email, displaying web pages (HTTP), domain name server (DNS) for resolving URL addresses, dynamic host control protocol (DHCP) for dynamic address assignment, PING for querying host/server and secure HTTP (SHTTP) for secure transfer of http contents.

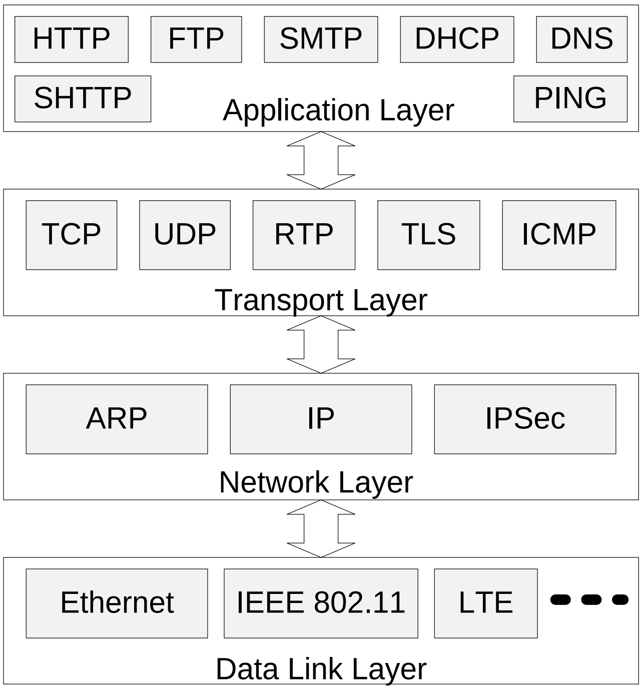

Figure 3. Internet protocol suite with commonly used protocols at each layer.

The next underlying layer is transport that is mainly responsible to transfer data from source application process to destination application process (between two processes) either by establishing an end-to-end connection (TCP) or a connectionless (UDP). The application data is divided into segments of maximum defined segment size to deliver over the network and reassembled at the destination to recover the original application data. To support real-time multimedia applications, real-time transport protocol (RTP) is also developed that basically runs over UDP to provide real-time transport service. Moreover, Internet Control Message Protocol (ICMP) is developed to report error messages to the source about the network congestion and other communication errors that require attention. For secure data transfer, a transport layer secure (TLS) protocol is developed to establish secure connection between the hosts to meet the user security demands. Network layer deals with the communication between two hosts no matter they are directly connected or multiple hops away to each other. Internet Protocol (IP) is the fundamental protocol specified in [RFC 2460] (IPv6) to support communication between hosts on different kinds of networks, i.e., different data-link layers such as Ethernet, IEEE 802.11, Token Ring. Data link layer specifies the organization of data into frames, error recovery over the link and how to send frames over a network. The Address Resolution Protocol (ARP) is used to translate the software-maintained IP addresses to physical addresses that match the addressing scheme of the underlying hardware (for Ethernet, the 48-bit MAC address). IP security (IPSec) is a security protocol that runs with the IP to support secure exchange of packets at the IP layer, which is developed by the IETF.

2. Network Layer

The network-layer moves transport-layer segments from one host to another. At the sending host, the transport layer segment is passed to the network layer. The network layer then gets the segment to the destination host and passes the segment up the protocol stack to the transport layer. The transport layer segments are packetized into maximum packet data unit (PDU) size packets. At the destination host, these packets are then reassembled to recover them into transport layer segments. In this section, we identify the network layer design issues, network services supported by current Internet and its two important functions; addressing and routing.

2.1. Fundamental Design Challenges

The main challenges that the network layer needs to address in the protocol suit of Internet are as follows:

- The services provided should be independent of the underlying technology. Users of the service need not be aware of the physical implementation of the network.
- The transport layer (that is the host computer) should be shielded from the number, type and different topologies of the subnets it uses.
- There is a need for some uniform addressing scheme for network addresses.
- Network layer should efficiently deal with the data delivery over the Internet to isolate not only the topologies of the subnet but also the distance between source and destination host, represented as hops.
- The network layer should be able to notify the senders causing network congestion.
- Mobility should also be supported to deliver data to the growing number of mobile hosts.

-

-

Bibliography

Afanasyev A., Tilley N., Reiher P., and Kleinrock L. (2010), *Host-to-Host Congestion Control for TCP*, IEEE Communication Surveys & Tutorails, pp. 1-39 [TCP performance analysis].

Agrawal P. and Chen J. Y. H. C. (2001), *A Survey of Energy Efficient Network Protocols for Wireless Networks*, Ad Hoc Networks, pp. 343-358 [Routing protocols survey for wireless networks].

Atzori L., Iera A., and Morabito G. (2010), *The Internet of Things: A survey*, Computer Networks, pp. 2787-2805 [The new concept of Internet of Things and its design issues].

Cisco (2011), *Cisco Visual Networking Index: Forecast and Methodology, 2010-2015",* white paper, Cisco [Statistics about the growth of Internet].

Floyd S., Henderson T., and Gurtov A. (2004), *New Reno modification to TCP's fast recovery algorithm*, IETF, RFC 3782 [An enhancement to TCP protocol to better cope with congestion].

Hui J. and Thubert P. (2010), *Compression Format for IPv6 Datagrams in 6LoWPAN Networks*, Internet draft [Header compression technique for low power devices].

Iwata A., Chiang C.-C., Pei G., Gerla M., and Chen T.-W. (1999), *Scalable Routing Strategies for Ad Hoc Wireless Networks*, IEEE Journal on Selected Areas in Communications, Special Issue on Ad-Hoc Networks, pp.1369-1379 [Challenges and design of routing protocols for mobile ad hoc networks] .

Ko J. G., et. al. (2011), *Connecting Low-Power and Lossy Networks to the Internet*, IETF Standards Update, IEEE Communication magzine, pp. 96-101 [Current state of the art on design of protocols for low power devices].

Kurose J. F. and Ross K. W. (2012), *Computer Networking: A Top-Down Approach*, Pearson Education, 862 pages [A comprehensive book on the fundamental concepts and design principles of the Internet protocols at each layer of the abstract layered model].

Liu J. and Singh S. (2001), *ATCP: TCP for mobile ad hoc networks*, EEE JSAC, vol. 19, no. 7, pp. 1300-1315 [A transport layer protocol for ad hoc networks considering frequent link failures].

Montenegro G. et al. (2007), *Transmission of IPv6 Packets over IEEE 802.15.4 Networks,* RFC 4944, IETF [Support of low power devices in IPv6].

Sendra S., Fernandez P. A., Quilez M. A. and Lloret J. (2011), *Study and Performance of Interior Gateway IP Routing Protocols*, Network Protocols and Algorithms, Macrothink, vol. 2, no. 4, pp. 88-117 [analysis of existing interior gateway routing protocols].

Shelby Z., Chakrabarti S., and Nordmark E. (2010), *Neighbor Discovery Optimization for Low-Power and Lossy Networks*, Internet draft [An optimized address resolution protocol for low power devices].

Winter T., Thubert P., and RPL Author Team (2010), *RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks*, Internet draft, 2010 [The routing protocol specification for low power devices].

Wu T.-Y., Huang C.-Y., Chao H.-C. (2005), *A survey of Mobile IP in cellular and Mobile Ad-Hoc Network environments*, Ad Hoc Networks, vol. 3, pp. 351-370 [Challenges and the support of mobile IP protocol for mobile networks].

Biographical Sketches

Ghalib A. Shah [M-09] (ghalib@kics.edu.pk) received his PhD degree in computer engineering from Middle East Technical University, Turkey in 2007. In 2007, he joined National University of Science & Technology, College of E & ME, as an Assistant Professor. He has been at School of Computer Science, Australian National University as a visiting fellow during 2009-2010. He was also awarded a COMSTECH-TWAS joint research grant for young researchers. In 2011, he worked at the Next Generation Wireless Communications Laboratory, Koc University, as a PostDoc fellow. Since July 2012, he is working as an associate professor at Al-Khawarizmi Institute of Computer Science, University of Engineering and Technology. His research interests include the design and analysis of communication protocols from MAC to Transport layer for diverse domain of wireless networks including ad hoc networks, wireless sensor networks, cognitive radio networks. He has authored many papers in peer-reviewed well-reputed journals and conference proceedings and is serving as a reviewer in many journals and TPC member of conferences.

Ozgur B. Akan [M'00, SM'07] (akan@ku.edu.tr) received his Ph.D. degree in electrical and computer engineering from the Broadband and Wireless Networking Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology in 2004. He is currently a full professor with the Department of Electrical and Electronics Engineering, Koc University and the director of the Next-generation and Wireless Communications Laboratory. His current research interests are in wireless communications, nano-scale and molecular communications, and information theory. He is an Associate Editor of IEEE Transactions on Vehicular Technology, International Journal of Communication Systems (Wiley), and Nano Communication Networks Journal (Elsevier). He has served as a General Co-Chair of ACM MOBICOM 2012, IEEE MoNaCom 2012, and TPC Co-Chair of IEEE ISCC 2012.