Enabling Gigabit Network Access to End Users

SHERALI ZEADALLY AND LIQIANG ZHANG, STUDENT MEMBER, IEEE

Invited Paper

The increasing demand for high-performance voice, video, image, and data communication networks has challenged researchers to design innovative network architectures capable of delivering high-speed data transfers to end users. To achieve this objective, considerable efforts have been invested in minimizing or eliminating bottlenecks that exist in high-speed network environments. These bottlenecks exist primarily at two levels, namely, network data transmission to the end system and data delivery within the end system to the user. For wired networks, improvements in fiber optic technologies have shifted the bottleneck from the underlying physical network to the end system. However, for wireless networks, we still face obstacles at both levels to achieve high, end-to-end performance data delivery, particularly at gigabit per second rates. In this paper, we first present current wireless communication technologies aimed at delivering gigabit per second transmission rates to end systems. We then investigate the bottleneck at the end system by exploring experimentally the performance benefits of a network interface architecture designed for enabling high-performance, low-latency applications using minimal host resources. We compare the performance of the network interface architecture with the traditional network interface architecture, using commodity PCs connected by gigabit per second local area networks running protocols such as TCP/IP and UDP/IP. We argue that such a network interface architecture will eliminate the bottlenecks prevalent in current end systems and, consequently, enables users to reap the full benefits of high-speed networks available today.

Keywords*—*Broadband communications, communication system performance, computer network performance, networks, network interfaces, protocols, radio communications, transport protocols.

I. INTRODUCTION

In recent years, we have witnessed increasing demands for high bandwidth, low-delays, and predictable performance on an end-to-end basis. These demands include multimedia, metacomputing, computer clustering, storage networking, and emergence of Internet-based applications. Coupled with these requirements, the cost to efficiently deliver and support these services have also become a significant factor in determining the success of both new technologies and applications deployment. In this paper, we explore technologies currently being deployed to deliver gigabit per second rates to end users for both wired and wireless networking environments. To achieve end-to-end gigabit rates, we require support at two levels. First, the underlying physical transmission technology should be capable of delivering high data transmission rates. Second, once the network data reaches the end system, we need to maintain the high-speed data path transmission through the end system (network interface, system bus, protocol stacks, operating system) all the way to the user application. Failure to maintain the performance at either level (i.e., the physical layer or within the end system) will dramatically degrade the end-to-end performance. It is worth noting that wired networks can now deliver data at the physical layer at multigigabit per second speed [21], [39], [57], [67], [70], [99]. However, these performance improvements have not translated into corresponding improvements at the application level because of bottlenecks at the end system. In the case of wireless networks, end-to-end high-speed data delivery is still a challenge at *both* levels, namely, the physical transmission layer and within the end system.

The rest of the paper is organized as follows. In Section II, we discuss wireless access technologies currently being deployed to enable gigabit networks. Section III investigates the implementation of a new network interface architecture designed to eliminate existing bottlenecks in commodity end systems. In Section IV, we describe our experimental procedures and testbed configurations, and discuss the results. In Section V, we make some concluding remarks and summarize the main contributions of this work.

II. WIRELESS ACCESS COMMUNICATION TECHNOLOGIES

Several technological changes are enabling the deployment of high-speed wired and wireless networking technologies including the availability of components (laser driver, multiplexor/demultiplexor, amplifier), application-specific integrated circuit (ASIC) technologies

---

Manuscript received October 18, 2002; revised October 24, 2003. This work was supported by grants from Sun Microsystems Inc., Palo Alto, CA, and Microsoft Corporation, Seattle,WA.

The authors are with the Department of Computer Science, Wayne State University, Detroit, MI 48202 USA (e-mail: zeadally@cs.wayne.edu; liqiang@cs.wayne.edu).

Digital Object Identifier 10.1109/JPROC.2003.821905

(SiGe, Hybrid Bipolar), cost reduction of technologies due to high volume of production, advances in fiber optics, radio transmission technologies, infrared (IR), laser, and processor performance, which have led to the emergence of several high-speed networks (100 Mb/s and above).

Over the last few years, we have witnessed the emergence of high-speed wired networks such as industry-standard Gigabit Ethernet, and other less popular networks such as Myrinet [26], Giganet [41], and POLO [81]. In addition, the recent 10-Gb Ethernet (IEEE 802.3ae) [52], [95] standard technology is anticipated to be used as a switch-to-switch interconnection for statistically multiplexing packet traffic from lower rate (10/100/1000 Mb/s) Ethernets. It is expected that the 10-Gb Ethernet will be used primarily as a backbone technology targeting the enterprise local area network (LAN) or telecommunication wide area network (WAN). However, no end station has a 10-Gb/s connection yet.

In recent years, a variety of wireless communication technologies and products have emerged in many areas including manufacturing, education, travel, and business sectors [85]. Common physical layer options used by these wireless technologies include microwave, IR, and laser. Wireless communications provide the benefits of network connectivity without the restrictions of being tied to a location or tethered wires. For wireless technologies, many factors need to be considered during design, implementation, and deployment, for instance, transmitters and receivers, antennas, link margin, modulation, propagation considerations (atmospheric attenuation, rain, snow, fog, gas effects), frequency reuse, and others. Today, we have wireless solutions available for wireless LANs (WLANs) and broadband wireless access (BWA). It is worth pointing out that while both provide access to emerging wireless applications, they do have some key differences. First, WLANs are designed for relatively short distances. In such cases, it takes a few tens of nanoseconds for the echoes to relax, whereas with BWA systems which are designed for outdoor environments, echoes spread apart a few hundreds of nanoseconds and up to few microseconds, which makes the "delay spread" problem more pronounced. Another significant issue between BWA systems and WLANs is the media access control layer used. WLANs are mostly used for data services where packets of which tend to be rather long. BWA systems, on the other hand, support a broader variety of services including voice (well known for using small packets). In addition, the carrier sense multiple access /collision avoidance (CSMA/CA) protocol, commonly used in WLANs, incurs significant overhead due to contention for the medium. But for BWA systems, stations seldom hear each other, either because of frequency duplexing or due to directional antennas in the stations, which is why BWA systems typically rely on centralized scheduling of transmissions by the base station [63].

A. Wireless LANs

The IEEE 802.11 project has set up several universal standards [46]–[48], [106] for WLANs operating in the 2.4- and 5-GHz bands. (Table 1 gives a brief overview of the evolution of IEEE 802.11 standards in terms of frequency range, air access scheme, data rate, and compatibility with other technologies.) In addition to these standards, we also include a European standard called Higher Performance Radio LAN2 (HiperLAN2) [40]. HiperLAN2 was developed under the European Telecommunications Standardization Institute (ETSI) Broadband Radio Access Networks (BRAN) project [64]. HiperLAN2 is similar to IEEE 802.11a in that both use the 5-GHz band and orthogonal frequency division multiplexing (OFDM). The 802.11 Study Group (the 5GSG) addresses the interoperability between 802.11a and HiperLAN. (See Fig. 1.)

Several modulation techniques have been adopted in 802.11 standards (as shown in Table 1). These include complimentary code keying (CCK), OFDM, and packet binary convolutional coding (PBCC) [106]. CCK is a "single carrier" system where all of the data (including the preamble/header and the payload) is transmitted by modulating a single radio frequency (RF), or carrier. OFDM is a "multicarrier" modulation scheme where the data is split up among several closely spaced "subcarriers." By doing so, OFDM systems are able to provide reliable operation even in environments that result in a high degree of signal distortion due to multipath. OFDM systems also support higher data rates than single-carrier systems without increasing system complexity. In addition, by exploiting OFDM modulation, IEEE 802.11a [47] reduces the preamble overhead using a shorter preamble length of 16 μs compared to 72 μs (802.11b [48] with short preamble) and 144 μs (802.11b with long preamble) of CCK. PBCC is a single-carrier system, but different from CCK in that it employs a more complex signal constellation [eight-phase shift keying (PSK) for PBCC versus BPSK/QPSK for CCK] and a convolutional code structure (block code structure used in CCK).

WLANs are built from two basic topologies often termed as infrastructure and *ad hoc*. The infrastructure topology extends the wired LAN to wireless devices by providing a base station (called an access point). The access point bridges wireless and wired LAN and coordinates transmission and reception from multiple wireless devices within a specific range. In contrast, the *ad hoc* topology is created solely by the wireless devices themselves, with no access point.

There are several issues in the design of high-speed WLANs including technical, economic, and regulatory. For indoor environments, a major technical factor is the channel response behavior (multipath) as a function of frequency band, building type, and the radio system architecture. To achieve high speed, design approaches generally focus on either alleviating multipath with multitone or equalization techniques (by emphasizing on communication/modulation techniques) or using line-of-sight (LOS) links with narrow beam antennas to eliminate virtually all multipath, and use simple unequalized modulation techniques such as frequency shift keying or PSK. For instance, in [25], Driessen demonstrated a variety of reliable 622-Mb/s binary PSK (BPSK) indoor radio links at 19 GHz on both LOS

	FrequencyRange	Air Access Scheme	Data Rate	Compatibility
802.11	2.4 GHz	FHSS, DSSS, Ir	1 and 2 Mbits/s	N/A
802.11b	2.4 to 2.4835 GHz	DSSS using CCK	Up to 11 Mbits/s	Compatible with 802.11 DSSS 1 and 2 Mbits/s
802.11a	5.15 to 5.25 GHz (50 mW), 5.25 to 5.35 GHz (250 mW), 5.725 to 5.825 GHz (1 W)	OFDM	Up to 54 Mbits/s	Not compatible with 802.11 and 802.11b
802.11g	2.4 to 2.4835 GHz	Mandatory CCK and OFDM, Optional PBCC and CCK/OFDM	Up to 54 Mbits/s	Compatible with 802.11b; not compatible with 802.11 FHSS, Ir
HiperLAN2	5.15 to 5.35 GHz, 5.470 to 5.725 GHz	OFDM	Up to 54 Mbits/s	Not compatible with 802.11, 802.11b, 802.11g

FHSS: Frequency Hoping Spread Spectrum
DSSS: Direct Sequence Spread Spectrum
Ir: Infrared

OFDM: Orthogonal Frequency Division Multiplexing
CCK: Complementary Code Keying
PBCC: Packet Binary Convolution Coding

Fig. 1. IEEE 802.11 standards and HiperLAN2.

and non-LOS (NLOS) links. The latter approach explores antenna/beamsteering techniques.

1) Performance Enhancement Techniques for Wireless LANs: Current WLANs still do not provide high link speeds compared to wired networking technologies. Many factors limit capacity access to end users. These factors include high, variable, and unpredictable error rate and delay (due to fading, interference, noise), and mobility. In recent years, a lot of research effort has been invested investigating new methods, schemes, and models in an attempt to maximize utilization of wireless links. Among these efforts, link layer adaptation has received considerable attention. Link layer adaptation mechanisms proposed include frame size and rate adaptation. Frame size significantly influences the network throughput. Generally, the larger the frames, the higher throughput. But for wireless transmissions, larger frames sometimes also mean higher retransmission overheads. To maximize wireless link throughput, adaptive sizing of the MAC layer frame has been proposed [16], [17], [68], [98]. The support of multiple physical transmission rates in IEEE 802.11 WLANs makes rate adaptation possible. Prior research efforts [44], [77], [80] propose that different data rates should be used under different channel conditions to achieve maximum link efficiency. Based on received signal strength measurements, a station can adaptively select the best rate suitable for transmissions [77]. A hybrid approach combining frame size and rate adaptations was also proposed in [80] for IEEE 802.11a. Other approaches such as packet frame grouping and PiggyData to improve throughput performance have also been discussed in [91], [92].

Although link throughput performance constitutes an important component in delivering high performance to wireless end users, the widespread use of multimedia networked applications involving continuous media such as audio/video have brought further requirements including delay and jitter. Current wireless network architectures are inadequate to efficiently support multimedia applications. In addition, technical challenges of these networks cannot be met with a layered design [90]. A promising approach to achieve this, which has received a lot of attention in recent years is *cross-layer design*. Most cross-layer solutions proposed exploit the cooperation of different layers to achieve optimal performance goals including reliability, power efficiency, network utilization, quality of service (QoS), and others. However, different approaches may integrate different layers depending on different goals. For instance, in [15], the cross-layer design approach for *ad hoc* networks uses two major layers of the mobile end system, namely, the routing layer and the middleware layer. These two layers are likely to use the same information to make specific decisions such as location of nodes or the topology of network. Another major benefit of this design is that in fast changing mobile computing environments, different layers need to cooperate closely to meet the changing requirements of mobile applications. In [104], a cross-layer adaptation framework and its prototype implementation using adaptations of CPU frequency, CPU allocation, and application quality in mobile systems are discussed. The goal was to deliver high system utility in the presence of time and energy constraints. In [105], a cross-layer architecture for multimedia delivery over wireless Internet combining the application, transport, and link layer is introduced. In particular, dynamic estimation of channel variations, adaptive error control in application and link layers, efficient congestion control, header compression, and others are discussed in the architecture. A cross-layer model for efficient resource management which includes application, network, and MAC layers is presented in [73]. Other approaches exploiting cross-layer optimizations in code division multiple-access (CDMA) cellular networks are discussed in [14] and [45].

In the context of multimedia support over wireless networks, it is worth pointing out that priority schemes at the MAC layer are needed to provide service differentiation. A full discussion of this topic is beyond the scope of this paper, but interesting schemes have been discussed in [1], [24], [88], and [96]. The IEEE 802.11 task group E is currently working toward the IEEE 802.11e [49] standard, which addresses QoS enhancement of the 802.11 MAC layer. Several performance evaluations of the draft standard are given in [43], [71], and [93].

2) Indoor IR: An alternative approach to radio-based WLANs is IR communications. IR networking uses electromagnetic radiation with wavelengths of 820 to 890 nm and its main benefits include no need for licenses, high potential capacity, and good control of interference. However, IR does not penetrate walls; therefore, IR-based WLANs should be contained within the room. IR transmitters and receivers can be designed either for diffused or directed uses. With diffused transmissions, the IR light transmitted by the sender unit fills the area and the receiver can be located anywhere it can receive the signal. For directed transmissions, the IR light is focused before transmitting the signal.

Indoor IR communications use photo detectors as receivers and light-emitting diodes (LEDs) and lasers as transmitters. Intensity mod/direct detection (IM/DD) [23], [103] is used in transceivers. Major sources of performance degradation with IR include multipath dispersion, shadowing, and background noise (sources include sunlight, incandescent light, and fluorescent light). Several research efforts [4], [59], [65] have investigated different techniques and key components to make IR communications possible at bit rates ranging from 10 to 100 Mb/s and above. In [59], Kahn *et al.* studied the use of multibeam transmitters and angle diversity receivers to improve performance by minimizing power consumption, multipath distortion, cochannel interference as well as power-efficient modulation and equalization techniques for IR transmission. A prototype demonstrating 70 Mb/s IR link was demonstrated. In [10], Bynoe and Carruthers focus at the system level which includes the network (*ad hoc* routing protocol), data link (WDMA protocol), and physical layer (IR) in the design and implementation of a multichannel broadband IR Wireless LAN architecture which assumes the existence of small, low-cost, tunable IR transceivers to realize the bandwidth potential of the IR medium in WLANs. The main goal of their work is to investigate techniques that can applied to design a wireless LAN capable of delivering greater than 100 Mb/s.

Although IR communication generally requires LOS, recent research at Pennsylvania State University, University Park [86], has demonstrated a wireless, IR LAN with non-LOS transmission. Their IR LAN uses a signaling scheme where each computer is equipped with a low power IR source and a holographic beam splitter. The splitter separates the original low-power beam into several narrow beams, which strike the ceiling and walls at points that form an invisible grid throughout the entire volume of the room to support non-LOS transmission (which diffuse beams also provide). In addition, their technique provides the benefits of low power (well below 1 W) and low error rate (one error per billion bits in 99% of the coverage area at bit rates up to a few hundred megabits per second) which cannot be achieved with diffuse beams.

3) Ultrawideband (UWB) Technology: A new wireless communication technology that could lead to short-range, high-speed data services that are faster than leading-edge wireless networks is the UWB transmission technology [34], [35]. UWB is an RF technology that transmits binary data using low-energy and extremely short duration (on the order of picoseconds) impulses or bursts of RF energy. Data can be transmitted over a wide spectrum of frequencies for short to medium distances around 15 to 100 m. Unlike conventional wireless systems that upconvert baseband signals to RF carriers, UWB can be used at baseband and can be thought of as a baseband transmission scheme that propagates at RF frequencies. Current UWB devices can transmit data up to 100 Mb/s [100], compared to the 1 Mb/s of Bluetooth and 54 Mb/s of 802.11. It is anticipated that UWB can be used to send data over short-distance personal area networks that link wireless devices such as notebooks, PCs, and PDAs in businesses and homes. UWB is also fast enough to accommodate multimedia traffic, including video, voice, and data. Developers claim that the technology could hit speeds up to 480 Mb/s [36]. Moreover, UWB operates on microwatts of power, less than 1/1000 the power required by conventional cellular phones. UWB is emerging as a promising technology, but several challenges still remain to be addressed, including that antennas must be nonresonant, phase-linear, and have a fixed phase center; transmitters must control pulse shaping; and silicon implementation of nonresonant circuits can be challenging.

B. BWA Technologies

LAN and WAN capacities are keeping pace with the tremendous growth in data traffic, but the “local loop bottleneck” between them is getting worse. Wireline technologies such as copper, cable, and fiber are not viable alternatives for several reasons.

1) Attempts by phone companies to reuse aging voice-grade copper lines by exploiting digital subscriber line (DSL) technology are hindered by limitations on distance, copper quality, and bandwidth.
2) Cable modems may satisfy the needs of residential users (asymmetrical network), but business users prefer more symmetrical networks [78], [54].
3) Today, fiber access is the primary transmission medium for broadband delivery, but its high costs have limited its wide deployment.

The solution to the local loop bottleneck is BWA extensions from existing high-speed fiber optics. Generally, BWA has several advantages over other common access technologies: bypassing the local loop (there is no need to "dig up the streets" to lay new infrastructure; compared to fiber optic, which costs on average of $1 million–2.5 million per metro mile and months to deploy, BWA can be deployed in hours costing up to $80 000 per metro mile [56], [69]); systems can be scaled up incrementally; high bit rates can be achieved (up to OC-48 or higher, depending on the frequency bands used); fast provisioning of services; and flexible partitioning of uplink/downlink bandwidth.

Current broadband wireless technologies based on spectrum allocation for services include local multipoint distribution services (LMDS) [42], [55], [75], multichannel multipoint distribution services (MMDS) [5], [20], [84], license-free fixed wireless services that comprise the industrial scientific and medical (ISM) bands, and the unlicensed national information infrastructure (U-NII) [33] bands. In addition, several efforts [60], [61], [79], [101] are also focusing on exploring the use of upper millimeter-wave bands (i.e., 60 GHz–300 GHz) and laser wireless communication, also known as free space optics (FSO), to provide gigabit per second or higher data rates [87], [101]. In the following sections, we briefly discuss these technologies based on the frequency range used by each of them.

1) Wave Frequency Less Than 10 GHz:

Centimeter-Wave Bands (MMDS): MMDS is a broadcasting and communications service that operates in the ultrahigh frequency (UHF) portion of the radio spectrum between 2.1- and 2.7-GHz RFs and can achieve data rates up to 10 Mb/s. MMDS service was originally licensed as a one-way service providing wireless video programming sometimes referred to as "wireless cable." The wireless cable industry largely failed in its effort to compete with wired and satellite-based video programming providers. As a result, recently in many countries MMDS has been permitted to be used for bidirectional services, enabling it as a transport mechanism for high-speed Internet access. The MMDS wireless system generally consists of the head-end equipment (satellite signal reception equipment, radio transmitter, other broadcast equipment, and transmission antenna) and the reception equipment at each subscriber location (antenna, frequency conversion device, and set-top device).

U-NII Bands: The U-NII [33] frequencies cover a bandwidth of 300 MHz and span three bands defined as follows: lower band, which covers 5.15 to 5.25 GHz reserved for indoor use; midband, which covers 5.25 to 5.35 GHz for outdoor use; and upper band, which covers 5.725 to 5.825 GHz, also for outdoor use. The transmitter output power is limited to 200 mW for lower band, 1 W for midband, and 4 W for upper band. These bands are designated for wideband, high data rate digital communications. The U-NII bands are unlicensed bands with no defined modulation or multiple access scheme. If devices sharing these bands follow regulations such as limiting the power of transmitters, these devices do not interfere with each other. U-NII bands are generally not susceptible to rain, fog, and snow. It is worth noting that it is possible to achieve a dedicated symmetric connection ranging from 128 Kb/s to 45 Mb/s depending on the equipment type with U-NII.

2) Wave Frequency Greater Than 10 GHz:

a) Millimeter-Wave Bands: The traditional microwave frequency allocations (less than 10 GHz) have become inadequate in supporting the exploding bandwidth demand because of the scarcity of unallocated spectrum and the need for interference-free channel separation. This led the wireless industry to investigate higher, previously unallocated portions of the spectrum such as the millimeter-wave frequency ranging from 10 to 300 GHz and laser.

Lower Millimeter-Wave Bands: Lower millimeter-wave bands are typically used by LMDS systems. Originally designed for wireless digital television transmission (one-way), LMDS was expected to serve the wireless subscription television needs, but nowadays LMDS is more used as a broadband wireless point to multipoint communication system that provides reliable digital two-way voice, data and Internet services. An LMDS system consists of a series of cells defined by individual base stations interconnected to a network operations center (NOC) using a fiber-based infrastructure. In every cell, the customer premise equipment is connected to the base stations via wireless links which require LOS. LMDS operates at 24, 28, 31, 38, and 40-GHz RFs in different countries and can achieve data rates up to 155 Mb/s. Various network architectures are possible with an LMDS system including asynchronous transfer mode (ATM) and Internet protocol (IP) transport technologies. A comparison between LMDS and MMDS is given in [20].

Upper Millimeter-Wave Bands: The maturity and success of lower millimeter-wave radios have driven researchers to investigate higher frequency bands, namely, the upper millimeter-wave bands,[1] especially the bands above 60 GHz. Several features make upper millimeter-wave communication an attractive and deployable technology. They include the capability to deliver multigigabit communication services (up to 10 Gb/s with 4 or 5 nines for distances ranging up to 3.5 km [69]); their short signal wavelength makes it possible to use smaller antennas; and their focused signal and low energy make it safer than lower frequency RF transmissions, since the upper millimeter-wave system energy fully dissipates in the upper levels of the skin. In addition, recent developments in microwave device technology such as the integration of solid-state and optical components with monolithic microwave integrated circuit, and upgraded microwave power devices are making upper millimeter-wave systems possible. However, the deployment of upper millimeter-wave systems is still limited by the effect of atmospheric factors [2], [38] and lack of support from technology (e.g., the need to develop supporting signal processing technology for digital wideband communications), standards, and frequency regulations.

b) Laser: Laser can also be used to provide broadband wireless communications and Laser wireless communication

---

[1] It is worth pointing out that the first demonstration of 60-GHz communications was shown in 1895 by J. C. Bose and first used in astronomy and military systems in late 1960s [79].

is often called FSO [12], [62], [101]. There are several benefits associated with FSO technologies: high bandwidth data rates; narrow laser beamwidth precludes interference with other laser-based systems; secure transmissions, since laser systems have narrow optical beam paths which are not accessible unless directly into the transmitter path (potential eavesdropping interrupts data transmission); and since the equipments operate around the IR spectrum, they are not subjected to government licensing and no spectrum fees have to be paid [37]. As with all the technologies described above, FSO also has some drawbacks: it has strict limitation on the directions, it can be interrupted (objects, heavy fog, snow, smoke), and although an emitted laser beam is invisible to the unaided eye, it can cause eye damage if viewed directly at close range for a long time. Another disadvantage of FSO is that the laser power attenuation through the atmosphere is variable and difficult to predict, since it is weather dependent, which limits the high availability range of FSO. In [60], a practical solution was proposed to extend the high availability range by using a backup low data rate RF link. Several companies (e.g., LightPointe Communications, Infrared Communication Systems) currently offer FSO products which can support data rates up ranging from 155 Mb/s to 2.5 Gb/s with distance ranging from 1 to 5 km, respectively.

3) IEEE 802.16 Standard: Based on the proliferation on BWA technology, IEEE 802.16 study group issued the Wireless Metropolitan Area Network air interface specification for broadband wireless access. The IEEE 802.16 Standard 802.16 [50], [72] was designed to evolve as a set of air interfaces based on a common MAC protocol but with physical layer specifications dependent on the spectrum of use and the associated regulations. The current draft of the standard, 802.16–2001, as approved in 2001, addresses frequencies from 10 to 66 GHz. A new project, IEEE 802.16a [51], approved in January 2003, extends the air interface support to lower frequencies in the 2–11-GHz band, including both licensed and license-exempt spectra. Compared to the higher frequencies, such spectra offer the opportunity to reach more customers at a lower cost although at generally lower data rates. It is anticipated that such services will be targeting primarily individual homes or small to medium-sized enterprises.

Significant research efforts are currently focusing on various physical layer technologies to enable gigabit per second data rates over wireless transmission links. These technologies include OFDM [82], [83], space-timing processing [3], [76](e.g., smart antennas [6], space-time codes (STC) [3]), adaptive modulation [13], and advanced signal processing techniques. We do not cover these technologies further, since they are already addressed in depth by other papers in this issue. Instead, the rest of this paper addresses the end-sytem bottleneck issue, which is another important factor impeding the delivery of gigabit per second data rates to end users.

III. ELIMINATING THE END-SYSTEM BOTTLENECK

Processor and network hardware performance continues to improve. Despite these technological improvements, applications continue to have difficulty to reap the full benefits (e.g., link speed) of these high-speed networks. The main reason that prevents application users from reaping the benefits of high-speed networks is primarily because of the bottlenecks that exist in the end system. The software overheads introduced by the network interface, operating system, and protocols constitute the dominant factor that slows down the path between the network and the user application. The following factors contribute mainly to the end system bottleneck of systems connected to high-speed networks [18], [27], [28], [58].

- Protocols have remained almost unchanged despite the fact that there have been tremendous improvements in the underlying physical network bandwidth.
- Operating systems also have remained essentially the same. For instance, most network interface interactions are interrupt driven, and under heavy loads successive interrupts make interrupt processing a significant overhead.
- The frequency and overhead of context switches (such as the need to switch between communication stacks) also add significant load on the end system.
- The number of data copies incurred as data is transferred between the user application and the network. Network data is copied at least twice as it passes between the user and the network. One copy occurs between the user and a kernel buffer and the second copy is between the kernel buffer and a network buffer. Data copying operations are expensive given the increasing mistmatch of processor speeds and memory speeds in recent years.

The virtual interface (VI) architecture was developed in response to the need for a new communication paradigm to achieve low latency and high bandwidth between computers using minimal host processor cycles. The VI architecture specification [19], [29] was jointly developed by Intel Corporation, Compaq Computer Corporation, and Microsoft Corporation.

In this section, we investigate the potential of a new network interface architecture known as the VI architecture, designed to address the bottlenecks that currently exist in systems connected to wired networks using the conventional network interface design and protocol stack architectures. We also report on practical experiences gained with two commercial gigabit per second networks. We explore experimentally the performance improvements obtained with the VI architecture when used on gigabit LANs.

As depicted in Fig. 2, VI provides for one-copy data transfers and defines a way to bypass the protocol stack layers and the underlying operating system. With the traditional network interface architecture, all controls and data pass through the kernel. With the VI architecture, only control and set up go through the kernel but data is transferred *directly* between the application and the network interface. In contrast to the traditional network interface architecture, VI performs one data copy only, since the user application writes directly to the network interface buffers on board the adapter. Only control and setup go through the kernel. Eliminating data copies not only boosts communication

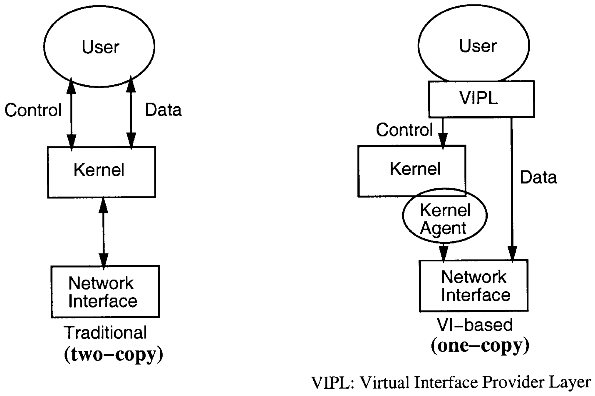

Fig. 2. Two-copy approach versus the one-copy approach using the VI architecture.

performance but also increases the host processor cycles availability.

VI avoids the need to switch between communication stacks, thereby eliminating context switch overheads. In addition, the VI architecture enables a process to avoid interrupts under heavy loads, since the architecture allows interrupts only on wait-for-completion (compared to the traditional case where the process.

All the features above enable VI implementations to deliver low latency and high bandwidth to end-user applications using minimum host CPU cycles.

It is worth pointing out that recently, VI functions have been incorporated into a new industry-standard architecture known as the InfiniBand Architecture (IBA) [53], [102]. It is a switched-fabric architecture designed for next-generation I/O systems and data centers. The architecture promises to replace bus-based I/O architectures, such as PCI, with a switch-based fabric whose benefits include higher performance, higher reliability, availability, and scalability (RAS), and the ability to create modular networks of servers and shared I/O devices.

A. The VI Architecture

In the traditional network architecture model, all communication operations trap into the operating system kernel making them expensive to execute. The VI architecture [7], [97] (shown in Fig. 3) provides a user-level networking architecture and eliminates the system processing overhead of the traditional model by providing each process with a directly accessible, protected interface to the network hardware known as a *virtual interface*. A VI is owned and maintained by a single process. The efficiency of the VI architecture relies on the removal of the TCP/IP layers, which gives users direct access to the network hardware for data transfer and reception.

The VI user agent is a software library provided by the operating system vendor that implements a user-level API for the VI architecture. The application and the VI user agent constitute a VI consumer. The VI consumer is the user of a VI. An application accesses communication services through standard operating system programming interfaces such as sockets implemented as a library linked with the application. To create a VI, the operating system makes system calls to the kernel agent. The newly created VI then connects to a VI on a remote system. Once the connection is established, the operating system posts the application's send and receive requests directly to the local VI.

The kernel agent is usually a driver supplied by the VI network adapter vendor that performs setup and resource management functions (such as the creation/destruction of VIs, interrupt management, and management of system memory used by the VI adapter) needed to maintain a VI between VI consumers and VI network adapters. The kernel agent and the network interface card together comprise the VI provider. VI consumers access the kernel agent using system calls. Kernel agents interact with the VI network adapter via standard device management functions.

The work queues shown in Fig. 3 are allocated when a kernel agent makes a remote connection. The queues hold descriptors for messages that are to be sent or received (one send and one receive queue per VI). A descriptor is a memory structure that contains all the information (e.g., pointers to data buffers) the provider needs to process the request. After a request is processed, the VI provider marks the descriptors with a status value. VI consumers remove completed descriptors from the work queues and use them for subsequent requests. Moreover, each queue has an associated doorbell to inform the VI network adapter that a descriptor has been posted. The doorbell is directly implemented by the adapter and requires no operating system involvement. The completion queue (CQ) notifies the application that operations are complete.

Typical application to network data transfers require that the memory pages holding the data to be transmitted to be locked and their virtual addresses translated to physical addresses. The pages are unlocked after the data transfer is complete. Traditional network subsystems perform these operations for every data transfer request. This adds to the overheads of the data transfers. With the VI architecture, the VI consumer identifies the memory used for a data transfer before requesting the data transfer. The memory registration step avoids the locking, unlocking, and translation operations normally incurred during data transfers by typical network subsystems. Moreover, the registration process also allows reuse of registered memory buffers. As a result, the processing overheads are removed from the network-application data path. In this case, data is directly transferred between the VI consumer buffers and the network without any intermediate data copies such as user to kernel and kernel to network buffers.

The improved performance offered by the VI architecture model results from the elimination of the intermediate data copies in the application-network data path. In addition, reducing the need for buffer management with the VI architecture also contributes significantly to minimizing the communication overheads found in traditional network communication subsystems.

B. Related Work

Previous research efforts have influenced many of the architectural concepts used in the VI architecture.

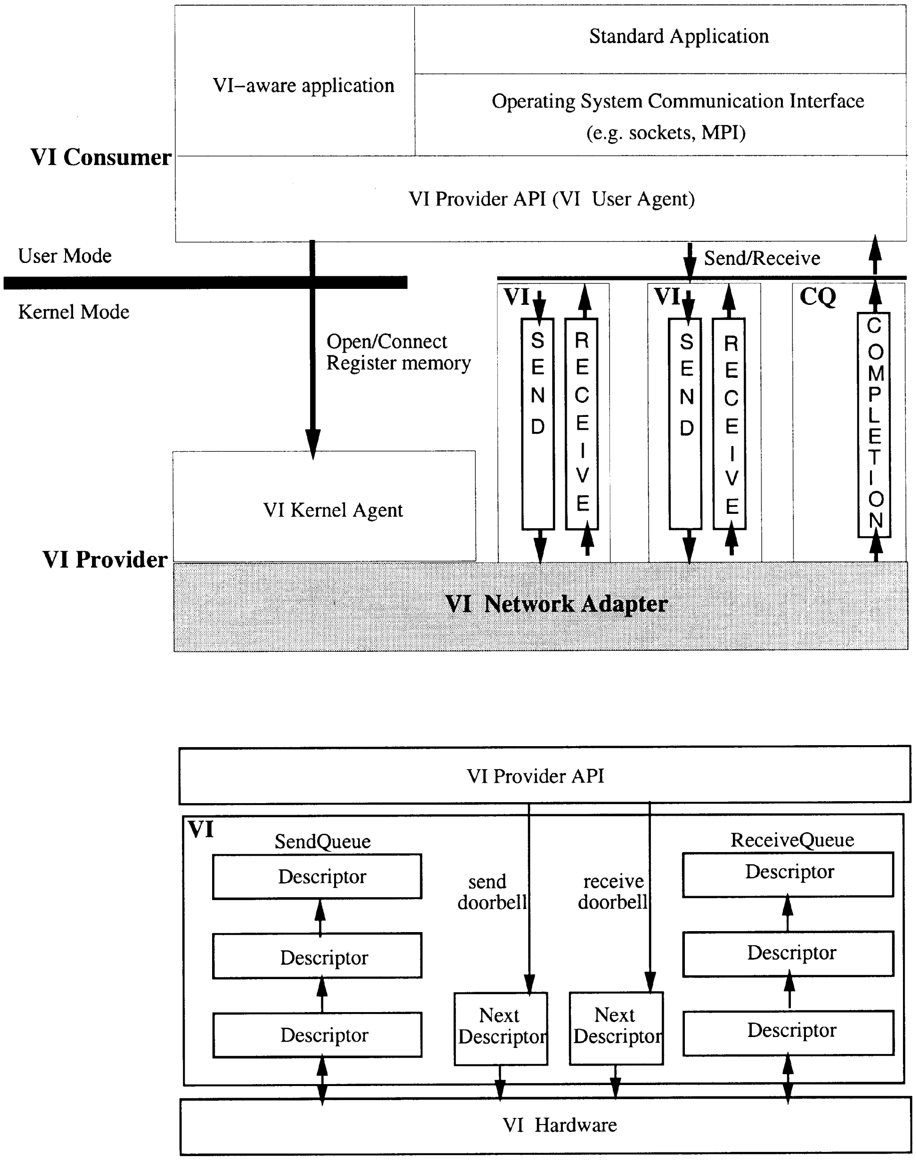

Fig. 3. The VI architectural model and the VI work queues used by the VI architecture.

The Message Driven Processor (MDP) project [22] that fast context switches and offloading the host CPU from handling messages (copying, buffering) improves message passing performance. Although processor speed has increased dramatically in the last few years, we have not seen a corresponding decrease in context switching time. Thus, VI architecture avoids context switches while off-loading the processor.

The Active Messages [31] and Fast Messages projects [74] at the University of California, Berkeley, and the University of Illinois Urbana–Champaign, both demonstrated the benefit of the asynchronous communication mechanism. In both architectures, a handler is executed on message arrival. Similarly, in the VI architecture, all message passing operations are asynchronous although VI does not however use the concept of a handler.

Several user-level networking architectures have been proposed, designed, and implemented [8], [9], [11], [30]. The U-net project at Cornell University, Ithaca, NY, demonstrated user-level access to a network. In the U-net architecture, each process is provided with the illusion of its own protected interface to the network. A set of send/receive/free queues are allocated for each endpoint (an endpoint enables U-net to provide protection boundaries among multiple processes). The Hamlyn architecture from Hewlett Packard laboratories also provides applications direct access to the network interface. In the Shrimp project at Princeton University, Princeton, NJ, a memory mapped network interface maps the send buffer to main memory. The network interface maps the source physical pages to the destination physical pages. All the page mapping information is stored in the page tables of network interfaces of the two communicating nodes.

Providing users direct access to a network adapter is not a new concept. However in the past, application code was written based on one system with a specific network adapter, operating system, and CPU. With the VI architecture, the concepts, data structures, and semantics are specified and defined by the computing industry in a *standard* way, independent of the operating system, CPU, and network adapter. The VI architecture standards remove that barrier if applications adhere to VI standards. Application code will benefit from future technological improvements, which lead to higher performance without the need of rewriting code. This reuse of application code is the key to the VI architecture acceptance.

IV. MEASUREMENT PROCEDURES, TESTBED CONFIGURATION, AND EXPERIMENTAL RESULTS

All experimental tests were performed using the following host, networks, and API configurations.

A. *Testbed*

- *Hosts:* We used Pentium III workstations each with a 500-MHz processor, and equipped with 512-MB RAM, 18-GB hard disk, loaded with three operating systems, namely, Windows 2000, Windows NT 4.0, and Linux. The hardware configuration we used for tests over cLAN gigabit network consisted of two uniprocessor 500-MHz Pentium III PCs, connected back-to-back via cLAN1000 gigabit network adapters capable of 1 Gb/s (line rate).
- *Networks:* Each workstation used was equipped with two 32-b, 33-MHz PCI network adapters: one gigabit Ethernet adapter (from Altheon Networks Inc.)), and one cLAN1000 network adapter from Giganet Networks Inc., which can support link speeds up to 1 and 1.25 Gb/s, respectively. It is worth noting that we chose the cLAN1000 adapter because it is the industry's first hardware-based (ASIC) implementation of the VI standard. We were interested in hardware support, since software implementations of VI have already been studied by other researchers [29], [89].
- *Application Programming Interfaces (APIs):* We implemented a native VI-based application using the VI API for all VI tests and a standard socket-based application (using Winsock 2 on Windows platform and BSD on Linux) for the tests conducted using TCP/IP-based stacks.

B. *Performance Metrics*

We used throughput, round-trip latency, and processor utilization as our performance metrics and they were measured as follows.

- Throughput: Average application-to-application throughput was measured by timing bulk data transfers over a sufficiently long period using test programs we have developed for a range of message sizes.
- Round-Trip Time: The test application on one host (one for each API used) echoes a message of a specified size to the peer application running on the remote host. Basically, the client machine sends an M-byte message to the server (timing starts) and waits to receive the M-byte message back; the interaction was repeated N times between client and server after which timing stops. From the N readings obtained, an average round-trip time for exchanging an M-byte message between the two workstations was calculated.

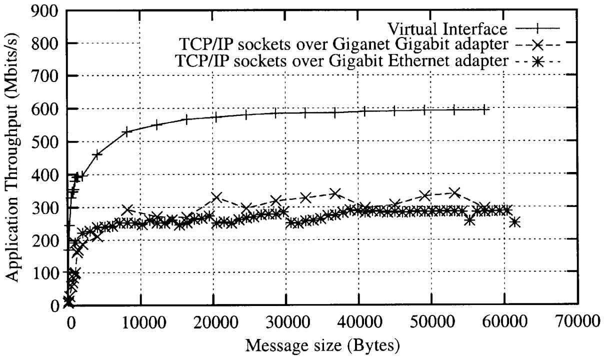

Fig. 4. Throughput of VI API, TCP/IP on cLAN1000 Giganet adapter, and TCP/IP on Altheon Gigabit Ethernet adapter.

- Processor utilization: CPU usage on the Windows and Linux platforms were measured using performance monitor [94] and Top [66] respectively. The CPU utilization reported is the average observed over the duration of a test.

C. *Experimental Results*

This section presents and discusses the results of our tests.

1) Performance of TCP/IP Versus VI-Based Protocol Stacks: The first test we performed was to investigate the current performance that can be achieved with current commercial gigabit per second networks using conventional TCP/IP protocol stacks. We conducted the test on Windows 2000 operating system. As Fig. 4 illustrates, both the cLAN Gigabit from Giganet and the gigabit Ethernet adapter from Altheon delivers almost the same application throughput performance at around 300 Mb/s. We also measured the corresponding CPU usage corresponding to that performance level for both networks. We obtained a host CPU usage of around 90%–100%, which becomes the bottleneck. Only 33% of the underlying network bandwidth can be delivered to end-user applications. This simple test clearly demonstrates that we continue to have difficulty pushing data at gigabit rates through end systems connected to gigabit networks. Using the host processor to move data between the network and the application remains of the major bottlenecks for end systems connected to gigabit networks.

To understand the performance improvement that can be achieved with the VI architecture for a VI-based application (that is bypassing the conventional TCP/IP stack), we repeated the above test using the VI API over the cLAN gigabit network. Using VI, we observed a doubling of the application throughput performance to 600 Mb/s. We also observed that the CPU usage in this case for the messages sizes tested was around 4%. The low CPU consumption was expected, since the VI implementation uses hardware-based

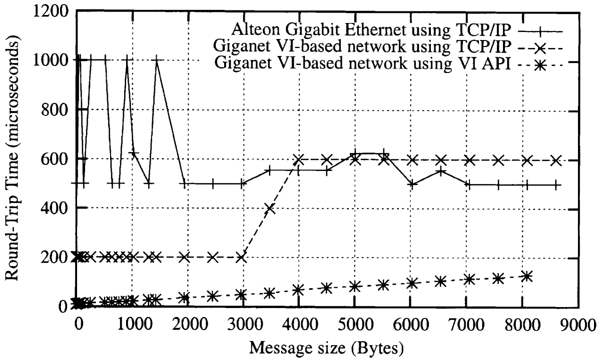

Fig. 5. Round-trip time using TCP/IP over Altheon Gigabit Ethernet adapter, TCP/IP and VI over cLAN on Windows 2000.

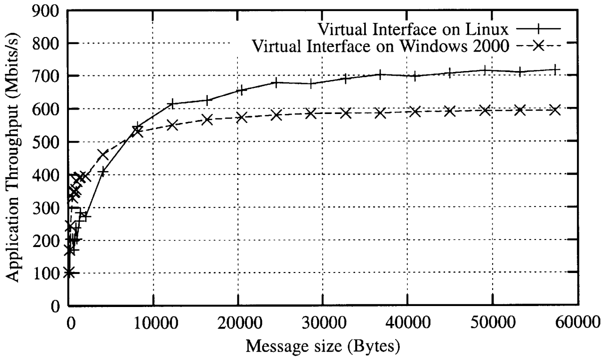

Fig. 6. VI throughput on Linux and Windows 2000.

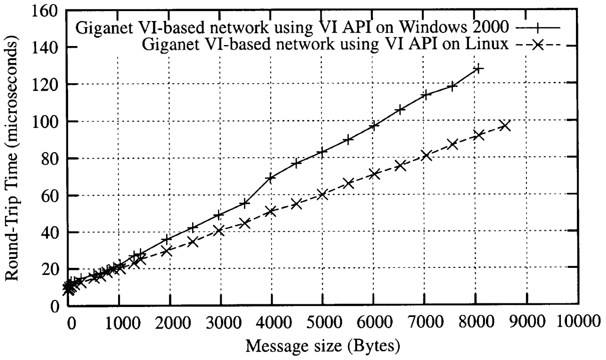

Fig. 7. VI round-trip time on Windows 2000 and Linux using cLAN gigabit adapters.

DMA to move data to application directly between the network adapter memory to application's memory without host CPU involvement.

We also measured the round-trip latencies using TCP/IP stacks running over the gigabit networks. As Fig. 5 illustrates, for messages up to around 4 KB, cLAN gigabit network adapter yields far lower latencies (almost by a factor of three to four) than gigabit Ethernet. Above 4 KB messages, latencies over TCP/IP for both networks are fairly close to each other and are around 500 to 600 μs. In contrast, in the case of the VI architecture using the VI API, the round-trip latencies obtained are well below 100 μs as observed in Fig. 5. We obtained a *fivefold to sixfold* improvement in latency with the VI implementation compared to the TCP/IP-based stacks running over the same gigabit networks.

2) Performance of VI Using Giganet Gigabit per Second Network on Windows 2000 and Linux: This section explores the performance of VI on UNIX and Windows platforms. We used Windows 2000 and Linux operating systems, both of which were installed on the same pair of hosts the test was carried out. The throughput obtained is shown in Fig. 6. We note that for messages up to 8 KB, Windows 2000 yields higher throughput, but beyond 8 KB message sizes, Linux yields higher performance approaching almost 700 Mb/s with a 20% increase over Windows 2000. Since we do not have the source code of the VI implementations for both Linux and Windows 2000, it is hard to explain the better performance for Windows 2000 over Linux for messages up to 8 KB, and the higher performance for Linux over Windows 2000 for messages bigger than 8 KB. The lack of device driver source code for the VI implementations on Windows 2000 and Linux makes it hard to explain this result. Further tests are required to explain this behavior and we plan to investigate this in the future.

We also performed round-trip latency tests, and the graphical results are shown in Fig. 7. For small message sizes up to 1 KB, the round-trip latencies obtained using VI are almost the same on both Windows 2000 and Linux. However, for larger messages, Linux gives better latencies than Windows 2000. Furthermore, the latency increase with message size on Linux is almost linear compared to the trend for Windows 2000. In fact, using the result in Fig. 7, the round-trip time on Linux can be calculated using the following equation:

$$\begin{aligned}&\text{Round--Trip Time (microseconds)}\\&\qquad = 10 + (0.01 * \text{Message Size})\ \text{(bytes)}\end{aligned}$$

A fixed cost of 10 μs is incurred by any message regardless of its size.

3) Winsock 2 and VI API System Call Cost on Windows 2000 and Windows NT 4.0 With Gigabit per Second Networks cLAN and Gigabit Ethernet: One of the goals of the application developer, operating systems designer, and network designer is to deliver optimal performance to end-user applications. Among some of the enhancements made to Windows 2000 compared to Windows NT 4.0 include features such as IPV6 and QoS support. To the best of the authors' knowledge, no comparison has ever been made on the communication performance between Windows 2000 and Windows NT 4.0. To reap the benefits of these new features in Windows 2000, we investigated any performance differences between the two Windows operating systems. In this work, we were only interested in the communication networking API performance of the two operating systems. Other operating system comparisons are beyond the scope of this work. We conducted simple tests that measure the time to perform basic socket system calls (for a STREAM socket) on each operating system using the same hardware platform (in this case a Pentium III using a 500-MHz processor). For the `send()/recv()` calls, the test measured the time to send or receive 1 B of data.

Table 1
Winsock 2 Socket System Call Cost (Microseconds) on Windows 2000 and Windows NT 4.0 Using TCP/IP Protocol and Running Over Gigabit per Second cLAN and Ethernet Networks.

Network Type	**Gigabit cLAN**	**Gigabit cLAN**	**Gigabit Ethernet**	**Gigabit Ethernet**
Socket System Call	Windows 2000 (microseconds)	Windows NT 4.0 (microseconds)	Windows 2000 (microseconds)	Windows NT 4.0 (microseconds)
socket()	210	139	111	89
bind()	198	135	198	138
connect()	436	318	2004	1149
send()/recv()	414	331	1178	539
closesocket()	182	134	84	70

Table 2
Cost of VI API Functions in Microseconds.

VI API	**Windows 2000** (microseconds)	**Windows NT 4.0** (microseconds)
VipOpenNic	366	302
VipNSGetHostByName	495	600
VipCreateVi	230	179
VipRegisterMem	22	17
VipConnectRequest	850	618
VipPostSend	24	18
VipSendWait	6	5
VipPostRecv	3	2
VipRecvWait	34	32
Total cost of above 4 items	67	57
VipDisconnect	57	68
VipDestroyVi	93	60
VipCloseNic	474	311

The results are given in Table 1. Measurements were made for the socket system calls over both the cLAN and Gigabit Ethernet LANs. From Table 1, it is interesting to note that Windows NT 4.0 outperforms Windows 2000. In fact over the gigabit cLAN network the improvement in performance of NT 4.0 over Windows 2000 is around 26% to 34%. The results are also consistent when the tests were repeated over the gigabit ethernet network. NT 4.0 outperforms Windows 2000 (in this case from around 17% to 54%). We note that although Windows 2000 has incorporated some enhanced networking features as mentioned above, the network communication API performance has *worsened* in Windows 2000 compared to Windows NT 4.0. This affects the application performance particularly for those specific socket calls which are called many times during execution of a network-based application [such as `send()` and `recv()`, which has an improvement of around 20% for cLAN and 54% for gigabit Ethernet].

We also compared the system call cost associated with the VI API on NT 4.0 and Windows 2000. Similar to the results above, performance of the VI API on NT 4.0 is better than on Windows 2000. It is worth noting from Table 2 that with the VI architecture and VI networking API, the cost of setting up a connection with VI (`VipOpenNic`, `VipNSGetHostByName`, `VipCreateVi`, `VipRegisterMem`, `VipConnectRequest`) and the cost of closing a connection (`VipDisconnect`, `VipDestroyVi`, `VipCloseNic`), are relatively high compared to their Winsock 2 counterparts. The high costs are due to the overheads involved in setting up all the data structures (memory allocation, pinning, doorbell setup, and so on) needed by the descriptors for sending and receiving data. In contrast, the Winsock 2 socket calls to make a connection are fairly low compared to the VI equivalent calls. The main point to note here is the cost to send and receive data from the network. Once the network connection is established, we incur minimal overheads in the data transfer with the VI API, around 67 μs, which is the aggregate cost of `VipPostSend`, `VipSendWait`, `VipPostRecv`, `VipRecvWait` (from Table 2). This is significantly lower than the `send()/recv()` call cost obtained with Winsock 2 (as shown in Table 2). The API cost improvement for sending and receiving data is almost *seven* times better with VI than the equivalent Winsock 2 `send()/recv()`.

V. CONCLUSION

In this paper, we argue that high, end-to-end performance delivery over high-speed networks requires support at two levels, namely, network data transmission to the end system and data delivery *within* (i.e., through the network interface, system bus, operating system, protocol stack) the end system to the end user. For optimal performance over high-speed networks, achieving and maintaining high data rates at both levels are needed. In the case of wired networks, it is now possible to deliver data at very high speed over the networks shifting the bottleneck to the end system. In contrast, for wireless networks, the challenges still remain to deliver data at high speeds at *both* levels (network transmission speed, and data path through the end system). It is imperative that we investigate solutions at both of these levels since wireless access is being widely adopted as the "last mile" solution.

In this paper, we reviewed the wireless communication technologies that have been deployed or currently being investigated to enable gigabit speed access to end users. We investigated the VI network architecture designed to eliminate the end system bottleneck which degrades the performance of the data path between the network and the application. We compared the benefits of the VI architecture with the traditional communication architecture using conventional TCP/IP stacks and APIs in gigabit per second LAN environments. We also compared TCP/IP-based stacks versus VI-based protocol stacks. In addition, we evaluated the performance of the VI architecture implementation on several operating systems, namely, Windows NT 4.0, Windows 2000, and Linux, all running on the same hardware platform. We obtained the following major empirical results.

- We deliver only 33% of the underlying network bandwidth to end-user applications using TCP/IP-based protocol stacks (using a 500-MHz Pentium III processor) over current commercial gigabit per second networks.
- Using the VI architecture and the VI API (bypassing TCP/IP), we can deliver 66% of the underlying bandwidth to user applications—a twofold performance improvement over conventional TCP/IP-based stacks.
- Round-trip latencies using VI are well within 100 μs. This is a fivefold to sixfold improvement over standard TCP/IP stacks running over the same underlying gigabit network.
- The Linux operating system yields better throughput and latency performances than Windows 2000 with VI for the cLAN gigabit network used in our experiments.
- Windows NT 4.0 gives better (i.e., lower) system call cost than Windows 2000 for both the VI and Winsock communication APIs.

ACKNOWLEDGMENT

The authors would like to thank H. Jiang and D. Ford for their help in executing the experimental tests performed. The authors would also like to thank J. Calder for his encouragements and rapid feedbacks. The authors would also like to thank N. Jayant who helped us throughout the preparation of this paper. His valuable feedback and comments on several drafts of this paper significantly helped us improve the originality, quality, organization, and presentation of this work. We would like to thank him for his dedication, time, and efforts throughout the revision process despite his busy schedule. We express our gratitude to the anonymous reviewers for their comments and suggestions. Finally, the authors would also like to thank K. Faggett for her kind help and support during the manuscript preparation.

REFERENCES

[1] I. Aad and C. Castelluccia, "Differentiation mechanisms for IEEE 802.11," presented at the INFOCOM 2001, Boston, MA, 2001.

[2] G. Ajayi and R. Olsen, "Modeling of a tropical raindrop size distribution for microwave and millimeter wave applications," *Radio Sci.*, vol. 20, pp. 193–202, Mar./Apr. 1985.

[3] N. Al-Dhahir *et al.*, "Space-time processing for broadband wireless access," *IEEE Commun. Mag.*, vol. 40, pp. 136–142, Sept. 2002.

[4] J. Barry, *Wireless Infrared Communications*. Boston, MA: Kluwer Academic, 1994.

[5] E. Bartsch, I. Wassell, and M. Sellars, "Equalization requirement study for broadband MMDS wireless access systems," presented at the Int. Symp. Communications (ISCOM01), Tainan, Taiwan, R.O.C., 2001.

[6] R. Becher *et al.*, "Broadband wireless access and future communication networks," *Proc. IEEE*, vol. 89, pp. 58–75, Jan. 2001.

[7] F. Berry, E. Deleganes, and A. Merritt, "The virtual interface architecture proof-of-concept performance results," Intel Corp., 1997.

[8] M. Blumrich *et al.*, "Virtual memory mapped network interface for the SHRIMP multicomputer," in *Proc. 21st Annu. Int. Symp. Computer Architecture*, 1994, pp. 142–153.

[9] A. Blumrich *et al.*, "Two virtual memory mapped network interface designs," in *Proc. Hot Interconnects II Symp.*, 1994, pp. 134–142.

[10] W. Bynoe and J. Carruthers, "A multichannel broadband infrared wireless LAN," presented at the 3rd IEEE Workshop on Wireless LANs, Boston, MA, 2001.

[11] G. Buzzard *et al.*, "An implementation of the Hamlyn sender managed interface architecture," *Oper. Syst. Rev.*, vol. 29, no. 5, pp. 40–53.

[12] T. Carbonneau and D. Wisely, "Opportunities and challenges for optical wireless; The competitive advantage of free space telecommunication links in today's crowded market place," in *Proc. SPIE Wireless Technologies and Systems: Millimeter Wave and Optical*, vol. 3232, 1997, pp. 119–128.

[13] S. Catreux *et al.*, "Adaptive modulation and MIMO coding for broadband wireless data networks," *IEEE Commun. Mag.*, vol. 40, pp. 108–115, June 2002.

[14] Y. Chan and J. Modestino, "An integrated/cross-layer approach for video delivery over CDMA cellular networks," presented at the 13th Int. Packet Video Workshop (PV 2003), Nantes, France, 2003.

[15] K. Chen, S. Shah, and K. Nahrstedt, "Cross-layer design for data accesibility in mobile ad hoc networks," in *Wireless Personal Communications*. Boston, MA: Kluwer Academic, 2002, vol. 21, pp. 49–76.

[16] S. Ci and H. Sharif, "Adaptive approaches to enhance throughput of IEEE 802.11 wireless LAN with bursty channel," presented at the 25th IEEE Conf. Local Area Network (LCN2000), Tampa, FL.

[17] ——, "Adaptive optimal frame length predictor for IEEE 802.11 wireless LAN," in *6th IEE Int. Symp. Digital Signal Processing for Communication Systems (IEE DSPCS2002)*, Sydney, Australia.

[18] D. Clark *et al.*, "An analysis of TCP processing overhead," *IEEE Commun. Mag.*, vol. 27, pp. 23–29, June 1989.

[19] "Virtual interface architecture specification," Compaq Corp., Intel Corp., and Microsoft Corp., 1.0 ed., 1997.

[20] F. Creede, "Datacasting with LMDS and MMDS systems," *Applied Microwave and Wireless*, vol. 13, no. 8, pp. 36–41, 2001.

[21] O. Crochat, J. Le Boudec, and O. Gerstel, "Protection interoperability for WDM optical networks," *IEEE/ACM Trans. Networking*, vol. 8, pp. 384–395, June 2000.

[22] W. J. Dally *et al.*, "Architecture of a message-driven processor," in *Proc. 14th Int. Symp. Computer Architecture*, 1987, pp. 189–196.

[23] T. Darcie, "Subcarrier multiplexing for lightwave networks and video distribution systems," *IEEE J. Select. Areas Commun.*, vol. 8, pp. 1240–1248, Sept. 1990.

[24] D. Deng and R. Chang, "A priority scheme for IEEE 802.11 DCF access method," *IEICE Trans. Commun.*, vol. E82-B, no. 1, pp. 96–102, Jan. 1999.

[25] P. Driessen, "Gigabits indoor wireless systems with directional antennas," *IEEE Trans. Commun.*, vol. 44, pp. 1034–1043, Aug. 1996.

[26] C. Dubnicki, A. Bilas, K. Li, and J. Philbin, "Design and implementation of virtual memory-mapped communication on Myrinet," in *Proc. International Parallel Processing Symp.*, 1997, pp. 388–396.

[27] P. Druschel and L. Peterson, "Fbufs: A high-bandwidth cross-domain transfer facility," in *Proc. 14th Symp. Operating System Principles*, 1993, pp. 189–202.

[28] P. Druschel, L. Peterson, and B. Davie, "Experiences with a high-speed network adaptor: A software perspective," in *Proc. Conf. Communication Architectures, Protocols, and Applications*, 1994, pp. 2–13.

[29] D. Dunning *et al.*, "The virtual interface architecture: A protected, zero copy, user-level interface to networks," *IEEE Micro*, vol. 18, pp. 66–76, Mar. 1998.

[30] T. Eicken, A. Basu, V. Buch, and W. Vogels, "U-Net: A user-level network interface for parallel and distributed computing," in *Proc. 15th ACM Symp. Operating Systems Principles*, 1995, pp. 40–53.

[31] T. von Eicken *et al.*, "Active messages: A mechanism for integrated communications and computation," *Comput. Arch. News*, vol. 20, no. 2, pp. 256–266, May 1992.

[32] C. Eklund, E. Marks, and K. Stanwood. IEEE Standard 802.16: A Technical Overview of the Wireless MAN Interface for Broadband Wireless Access. [Online]. Available: http://grouper.ieee.org/groups/802/16/docs/02/C80216-02_05.pdf

[33] "FCC, First Report and Order, Unlicensed NII Devices in the 5 GHz Band,", ET Docket no. 96-102, RM-8648 & RM-8653, Document FCC97–005, 1997.

[34] R. Fontana, "Recent trends in ultra wideband communication systems, multispectral solutions," presented at the 3rd IEEE Workshop Wireless LANs, Newton, MA, 2001.

[35] ——, "A novel ultra wideband communications system," presented at the MILCOM 97, Monterey, CA.

[36] ——, "Recent applications of ultra wideband radar and communications systems," in *Ultra-Wideband, Short Pulse Electromagnetics*. Boston, MA: Kluwer Academic/Plenum, 2000.

[37] Free Space Laser Communications Systems [Online]. Available: http://www.bakom.ch/imperia/md/content/english/funk/forschungundentwicklung/studien/FreeSpaceLaser.pdf

[38] Y. Furuhanma *et al.*, "Propagation experiment program at millimeter waves," *Radio Res. Lab. Tech. J.*, vol. 26, no. 138, pp. 507–553, June 1980.

[39] O. Gerstel, R. Ramaswami, and G. Sasaki, "Cost-effective traffic grooming in WDM rings," *IEEE/ACM Trans. Networking (TON)*, vol. 8, pp. 618–630, Oct. 2000.

[40] V. Ghazi-Moghadam, "Performance results for SPW implementations of IEEE 802.11a and HIPERLAN/2 WLAN standards," in *Proc.3rd IEEE Workshop Wireless LANs*, Boston, MA, 2001.

[41] GigaNet cLAN Product Family. Giganet Inc. [Online]. Available: http://www.giga-net.com/products

[42] D. Gray, "A broadband wireless access system at 28 GHz," in *Proc. NIST-Cosponsored 1997 Wireless Communications Conf.*, 1997, pp. 1–7.

[43] A. Grilo and M. Nunes, "Performance evaluation of IEEE 802.11e," presented at the 13th IEEE Int. Symp. Personal, Indoor and Mobile Radio Communications, Lisboa, Portugal, 2002.

[44] G. Holland, N. Vaidya, and P. Bahl, "A rate-adaptive MAC protocol for multi-hop wireless networks," presented at the ACM/IEEE Int. Conf. Mobile Computing and Networking (MOBICOM'01), Rome, Italy, 2001.

[45] Y. Huh *et al.*, ""MAI-JSQ:A Cross-Layer Design for Real-Time Video Streaming in Wireless Networks", Internal Technical Report," Dept. Electrical Eng, Arizona State Univ. , 2002.

[46] *Wireless LAN Medium Access Control (MAC) Sublayer and Physical Layer Specifications*, IEEE Standard 802.11–1997, 1997.

[47] *Part11: Wireless LAN Medium Access Control (MAC) Sublayer and Physical Layer Specifications; High-Speed Physical Layer in the 5 GHz Band*, IEEE Standard 802.11a-1999, 1999.

[48] *Part11: Wireless LAN Medium Access Control (MAC) Sublayer and Physical Layer Specifications: Higher Speed Physical Layer Extension in the 2.4 GHz band*, IEEE Standard 802.11b-1999, 1999.

[49] IEEE P802.11 – Task Group E. IEEE. [Online]. Available: http://grouper.ieee.org/groups/802/11/Reports/tge_update.htm

[50] *IEEE Standard for Local and Metropolitan Area Networks – Part 16: Air Interface for Fixed Broadband Wireless Access Systems*, IEEE Standard 802.16–2001, Apr. 2002.

[51] *IEEE Standard for Local and Metropolitan Area Networks – Part 16: Air Interface for Fixed Broadband Wireless Access Systems – Amendment 2: Medium Access Control Modifications and Additional Physical Layer Specifications for 2–11 GHz*, IEEE Standard 802.16a-2003, Jan. 2003.

[52] 10 G/s Status and Technology. IEEE. [Online]. Available: http://grouper.ieee.org/groups/802/3/10G_study/public/march99/van_doorn_1_03 991.pdf

[53] (2001) IBTA 1.0.a Specifications. Infiniband (SM) Trade Association. [Online]. Available: http://www.infinibandta.org/specs

[54] Broadband Wireless. Infy. [Online]. Available: http://www.infy.com/knowledge_capital/thought-papers/broadband2.pdf

[55] H. Izadpanah, "LMDS: A broadband wireless access technology: An overview," presented at the 3rd IAA Annu. Computer and Commmunications Conf., New York, 1998.

[56] I. Izadpanah *et al.*, "MM-wave wireless access technology for the wideband wireless local loop applications," presented at the IEEE RAWCON'98, Colorado Springs, CO, 1998.

[57] S. Kartapoulos, *Introduction to DWDM Technology: Data in a Rainbow*. Washington, DC: SPIE Optical Engineering Press, 2000.

[58] J. Kay and J. Pasquale, "The importance of nondata touching processing overheads in TCP/IP," *Comput. Commun. Rev.*, vol. 23, no. 4, pp. 259–268, Oct. 1993.

[59] J. Kahn and J. Barry, "Wireless infrared communications," *Proc. IEEE*, vol. 85, pp. 265–298, Feb. 1995.

[60] I. Kim and E. Korevaar, "Availability of free space optics (FSO) and hybrid FSO/RF systems," presented at the SPIE ITCOM 2001, Wireless Communications IV, Denver, CO, 2001.

[61] I. Kim *et al.*, "Wireless optical transmission of fast ethernet, FDDI, ATM, and ESCON protocol data using the TerraLink laser communication system," *Opt. Eng.*, vol. 37, no. 12, pp. 3143–3155, Dec. 1998.

[62] I. Kim, B. McArthur, and E. Korevaar, "Comparison of laser beam propagation at 785 nm and 850 nm in fog and haze for optical wireless communications," in *Proc. SPIE Optical Wireless Communications III*, vol. 4214, 2000, pp. 26–37.

[63] J. Klein, "TDD v/s FDD: The drive for effective bandwidth management," *RF Design*, pp. 35–55, Aug. 1999.

[64] J. Kruys. The ETSI Project Broadband Radio Access Networks. Technologies, Nieuwegein, The Netherlands. [Online]. Available: http://www.watmag.com/tech/Broadband/Broadband-002/Broadband-002.html

[65] D. Lee, J. Kahn, and M. Audeh, "Trellis-coded pulse-position modulation for indoor wireless infrared communications," *IEEE Trans. Commun.*, vol. 45, pp. 1080–1087, Sept. 1997.

[66] W. LeFabvre. (2000) Unix Top. [Online]. Available: http://www.groupsys.com/top/

[67] A. Levi, "Optical interconnects in system," *Proc. IEEE*, vol. 88, pp. 750–757, June 2000.

[68] P. Lettieri and M. Srivastava, "Adaptive frame length control for improving wireless link throughput, range, and energy efficiency," in *Proc. IEEE INFOCOM '98*, vol. 2, 1998, pp. 564–571.

[69] D. Lockie. Upper millimeter-wave history, technology and applications for gigabit class communications services IEEE 802 LMSC. Endwave Corp.. [Online]. Available: http://grouper.ieee.org/groups/802/802_tutorials/march02/T80216–02_02b.pdf

[70] B. Madhavan and A. Levi, "A 55 Gb/s/cm data bandwidth data interface in 0.5 μm CMOS for advanced parallel optical interconnects," *Electron. Lett.*, vol. 34, pp. 1846–1847, 1998.

[71] S. Mangold *et al.*, "IEEE 802.11e wireless LAN for quality of service," presented at the Eur. Wireless Conf., Florence, Italy, 2002.

[72] R. Marks. IEEE MTT-S Distinguished Microwave Lecture: Emerging IEEE 802.16 Wireless WAN Standards for Broadband wireless Access. National Institute of Standards and Technology, Chair, IEEE 802.16 Working Group. [Online]. Available: http://nwest.nist.gov/dml_marks.pdf

[73] N. Nikaein and C. Bonnet, "A glance at quality of service models in mobile ad hoc networks," presented at the 16th Conf. New Architectures for Communications, Paris, France, 2002.

[74] S. Pakin, V. Karamcheti, and A. Chien, "Fast messages (FM): Efficient, portable communication for workstations and massively-parallel processors," *IEEE Concurrency*, vol. 5, pp. 60–72, June 1997.

[75] P. Papzian *et al.*, "Study of the LMDS radio channel," *IEEE Trans. Broadcast.*, vol. 43, pp. 175–184, June 1997.

[76] A. Paulyaj and C. Papadias, "Space-time processing for wireless communication," *IEEE Signal Processing Mag.*, vol. 14, pp. 49–83, Nov. 1997.

[77] J. Pavon and S. Choi, "Link adaptation strategy for IEEE 802.11 WLAN via received signal strength measurement," presented at the IEEE ICC, Anchorage, AK, 2003.

[78] P. Perini. Broadband wireless technologies and applications. [Online]. Available: http://www.its.bldrdoc.gov/meetings/art/art00/slides00/per/per_s.pdf

[79] Gigabit MM wave communication. Pleasant Telaxis Commun.. [Online]. Available: http://grouper.ieee.org/groups/802/802_tutorials/march02/T80216–02_02a.pdf

[80] D. Qiao and S. Choi, "Goodput enhancement of IEEE 802.11a wireless LAN via link adaptation," presented at the IEEE Int. Conf. Communications (ICC 2001), Helsinki, Finland, 2001.

[81] B. Raghavan *et al.*, "A Gbyte/s parallel fiber-optic network interface for multimedia applications," *IEEE Network*, vol. 13, pp. 20–28, 1999.

[82] T. Rappaport *et al.*, "Wireless communication: Past events and a future perspective," *IEEE Commun. Mag.*, vol. 40, pp. 148–161, May 2002.

[83] H. Sampath *et al.*, "A fourth-generation MIMO-OFDM broadband wireless system: Design, performance, and field trial results," *IEEE Commun. Mag.*, vol. 40, pp. 143–149, Sept. 2002.

[84] J. Sanford and J. Raymond, "Enhancing the MMDS transport layer," *Appl. Microwave Wireless*, pp. 32–39, Jan. 2001.

[85] H. Sari, "Broadband radio access to homes and businesses," *Comput. Netw.*, vol. 31, pp. 379–393, 1999.

[86] (2001) Low power, highly reliable, wireless, infrared local area networks demonstrated. Science Daily. [Online]. Available: www.sciencedaily.com/releases/2001/07/010 725 081 349.htm

[87] I. Smolyaninov *et al.*, "Long distance 1.2 Gb/s optical wireless communication link at 1550 nm," in *SPIE Conf. Free-Space Laser Communication and Laser Imaging*, San Diego, CA, 2001.

[88] J. Sobrinho and A. Krishnakumar, "Real-time traffic over the IEEE 802.11 medium access control layer," *Bell Labs Tech. J.*, pp. 172–187, 1996.

[89] E. Speight, A. Abdel-Shafi, and J. Bennett, "Realizing the performance potential of the virtual interface architecture," presented at the Int. Conf. Supercomputing, Rhodes, Greece, 1999.

[90] S. Toumpis and A. Goldsmith, "Performance, optimization, and cross-layer design of media access protocols for wireless ad-hoc networks," presented at the ICC, Anchorage, AK, May 2003.

[91] J. Tourrilhes, "Packet frame grouping: Improving IP multimedia performance over CSMA/CA," presented at the IEEE Int. Conf. Universal Personal Communications, Florence, Italy, 1998.

[92] ——, "PiggyData : Reducing CSMA/CA collisions for multimedia and TCP connections," presented at the VTC '99, Amsterdam, The Netherlands.

[93] H. Truong and G. Vannuccini, "The IEEE 802.11e MAC for quality of service in wireless LANs," presented at the SSGRR, L'Aquila, Italy, 2003.

[94] M. Tullocj, "Windows 2000 Administration in a Nutshell," O'Reilly, Sebastopol, CA, 2001.

[95] B. Tolley. 10 Gigabit Ethernet: Customer applications. [Online]. Available: http://grouper.ieee.org/groups/802/3/ae/public/may00-/tolley_1_0500.pdf

[96] N. Vaidya, P. Bahl, and S. Gupta, "Distributed fair scheduling in a wireless LAN," presented at the 6th Annu. Int. Conf. Mobile Computing and Networking, Boston, MA, 2000.

[97] (1997) Virtual Interface Architecture Specification. [Online]. Available: http://www.vidf.org

[98] V. Vishnevsky and A. Lyakhov, "Adaptive features of IEEE 802.11 protocol: Utilization, tuning and modifications," in *Proc. 8th HP-OVUA Conf*, Berlin, Germany, 2001.

[99] R. Walker, "A 10 Gb/s Si-bipolar TX/RX Chipset for computer data transmission," in *Proc. IEEE, Int. Solid State Circuits Conf.*, 1998, pp. 302–302.

[100] J. Walko. (2002) Xtreme spectrum demos UWB chip set. [Online]. Available: http://www.commsdesign.com/story/OEG20 020 716S0033

[101] M. Wright *et al.*, "2.5 Gbps optical communication links depicting LEO-to-GEO communication," presented at the SPIE Conf: LASE (High Power Lasers and Applications) 2002, San Jose, CA, 2002.

[102] J. Wu, A. Gulati, B. Abali, and D. Panda, ""Design of an infiniband emulator over Myrinet: Challenges, implementation, and performance evaluation: Internal report," Dept. Comput. Info. Sci., Ohio State Univ., , Columbus, OSU-CISRC-02/01-TR03.

[103] R. You and J. Kahn, "Average power reduction techniques for multiple sub-carrier intensity modulated optical signals," *IEEE Trans. Commun.*, vol. 49, pp. 2164–2171, Dec. 2001.

[104] W. Yuan *et al.*, "Design and evaluation of a cross-layer adaptation framework for mobile multimedia systems," presented at the SPIE/ACM Multimedia Computing and Networking Conf. (MMCN'03), Santa Clara, CA, 2003.

[105] Q. Zhang, W. Zhu, and Y. Zhang, "A cross-layer QoS-supporting framework for multimedia delivery over wireless internet," presented at the Packet Video Conf., Pittsburgh, PA, 2002.

[106] J. Zyren. 802.11g explained. Intersil Corp.. [Online]. Available: http://www.intersil.com/design/prism/papers/WP_IEEE802gExpla-_12_06.pdf

Sherali Zeadally received the B.A. and M.A. degrees in computer science from University of Cambridge, Cambridge, U.K., and the Ph.D. degree in computer science from University of Buckingham, Buckingham, U.K., in 1996.

In March 1996, he joined the Department of Electrical Engineering at the University of Southern California (USC), Los Angeles, where he was an Assistant Professor until August 1999. While at USC, he was also one of the Principal Investigators of the Integrated Media Systems Center (IMSC), an NSF Engineering Research Center, where he led the networking research group from 1996 to 1999. He joined the Department of Computer Science at Wayne State University, Detroit, MI, as an Assistant Professor in August 1999. He is the Founder and Director of the High-Speed Networking Laboratory and the computer networking research group at Wayne State University. His research in computer networking has been funded by NSF and several companies, including HP, Microsoft, Compaq, Altheon, IBM, Sun Microsystems, and others. He currently serves on the editorial boards of several international journals. His research interests include high-speed networks, quality of service, operating systems, wireless and personal area networks, middleware, and performance analysis of network systems.

Dr. Zeadally is the recipient of two Outstanding Teaching Awards (in 2001 and 2003) from the College of Science at Wayne State University, the highest award given by the college for excellence in teaching.

Liqiang Zhang (Student Member, IEEE) received the B.S. degree in computer science from China University of Mining and Technology in 1993 and the M.S. degree in electrical engineering from China Academy of Telecommunication Technology (CATT) in 1996. He is currently working toward the Ph.D. degree in the Department of Computer Science at Wayne State University, Detroit, MI.

From 1996 to 2000, he participated in several research and development projects on data communications at CATT. He has published several refereed papers on computer networks. His research interests include quality of service over wired and wireless networks, and end systems.

Mr. Zhang was the winner of the 2003 Thomas C. Rumble University Graduate Fellowship.