Enabling Large-scale Wireless Broadband: The Case for TAPs

R. Karrer, A. Sabharwal, E. Knightly
Rice University, Houston, TX
{karrer,ashu,knightly}@ece.rice.edu

Abstract— **The vision is tantalizing: a high-performance, scalable, and widely deployed wireless Internet that facilitates services ranging from radically new and unforeseen applications to true wireless "broadband" to residences and public spaces at rates of 10s of Mb/sec. However, while high-speed wireless access is easy to achieve in an enterprise network via low-cost IEEE 802.11 (WiFi) access points, wireless technology in public spaces is in its infancy. "Hot spots" provide high-speed wireless access, but do so in very few isolated "islands" at immense costs. Likewise, while fixed wireless (e.g. LMDS) and 3G can provide ubiquitous coverage and 3G can support mobility, throughputs can often be two orders of magnitude slower than WiFi.**

In this paper, we formulate the challenges of building a high-performance, scalable and widely deployed wireless Internet along 10 premises. We make the case for the requirement of a fundamental new architecture based on beamforming antennas deployed on fixed, wire-powered *Transit Access Points (TAPs)* that form a multi-hopping wireless backbone with a limited number of *wired* ingress/egress points. To address scalability, deployability, and performance challenges we present distributed, opportunistic and coordinated resource management problems and a novel "network is the channel" framework that searches for fundamental information-theoretic tradeoffs between protocol overhead and capacity.

I. INTRODUCTION

Over the last decade, we have witnessed an explosion in wireless access to the Internet. In 2002, revenue from IEEE 802.11 (WiFi) network cards and access points totaled an estimated \$2.1 billion on 23.9 million devices with 73% growth predicted for 2003 alone.[1] Moreover, advances in the physical layer and media access protocols have enabled transmission rates of 54 Mb/sec in IEEE 802.11a, and even higher rates are projected in future revisions.

However, in spite of these advances, we remain in the infancy of of the long-standing vision of a high-speed ubiquitous wireless web. To date, the overwhelming majority of deployed WiFi networks are in the enterprise or home, restricting high-speed wireless data communication to small wireless "islands."

There are two simultaneous efforts to providing wireless Internet beyond these islands. The first is deployment of "WiFi hot spots", typically consisting of IEEE 802.11b access points connected to a (T1) wired backbone. However, the great fanfare with which each hot spot is announced [1] is immediately tuned down by sheer numbers: at the end of year 2002, the U.S. had approximately 3,000 hot spots, attracting an estimated 20,000 users, resulting in a net revenue of \$2 million – and yielding a large net loss, given fixed costs as high as \$10k and recurring costs of approximately \$400/month per hot spot.[2] Ironically, the overwhelming costs of providing wireless hot spots is due to fixed and recurring costs of the *wired* infrastructure.

As a consequence, deployment is low and "coverage" is dismal. Even adding Cometa's plans for an additional 20,000 hot spots to the existing 3000, and optimistically estimating that each hot spot covers 100x100 m^2, coverage will be as low as approximately 4 km^2 per metro area, or 0.4% of the area of a moderate sized city such as Indianapolis. Thus, today's hot spot architecture is slow to deploy, costly, and unscalable, and is not on any path to provide large-scale coverage.

A second major effort is 3G and fixed wireless services such as LMDS. However, in both cases, speeds are typically 2 to 3 orders of magnitude slower than WiFi, with maximum per-user speeds in the 100s of kb/sec range. Moreover, because of multi-billion dollar spectral license costs and high infrastructure costs, such systems have proven costly to deploy and hence lead to expensive, yet moderate speed, wireless Internet services. Thus, while having the promise of near-ubiquitous coverage and allowing high mobility speeds, such technologies have significant performance and cost limitations. Moreover, given their small Internet subscriber base, scalability to many data users remains unproven. Therefore, despite a decade of strong progress in wireless data communication, it is clear that with the current evolutionary path, a large-scale high-speed wireless web is not on the horizon.

This paper describes the challenges of building a wireless Internet that simultaneously achieves deployability, scalability, high-performance, and a cost-effective economic model along 10 fundamental premises. We believe that these premises - some of which are inherent to any wireless network, some of which are specific to the outlined challenges - provide the basic framework to realize the above vision.

II. THE CASE FOR TAPS

Premise 1: Designing a wireless Internet that simultaneously achieves deployability, scalability, high-performance, and a cost-effective economic model requires a new architecture. This architecture is based on fully-wireless beamforming Transit Access Points (TAPs) that form a multihop backbone mesh which interconnects TAPs, mobile units (MUs), and the wired Internet.

The dominant infrastructural costs of traditional hot spots lead to the (logical) conclusion that many of the AP's wires

---

This work was partially supported by National Science Foundation under Grant ANI-0325971.The authors will like to thank their collaborators, Behnaam Aazhang, Patrick Frantz and Dave Johnson for invaluable discussions and comments.

[1] Source: Gartner Dataquest.

[2] Sources: Jupiter Media Metrix analyst Dylan Brooks and Insight on Wireless analyst Andrew Luan.

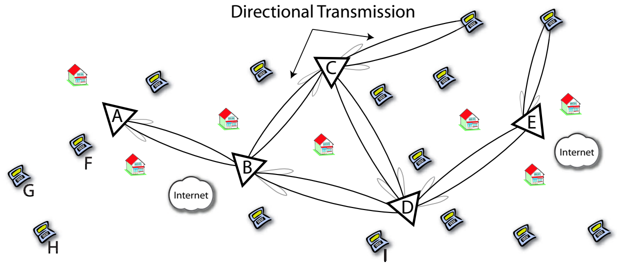

Fig. 1. TAP network

must be removed. However, simply removing the wires creates an ad hoc-like network where packet forwarding of wireless APs and user access compete for the scarce spectrum, pushing the system capacity dramatically to the well-known scalability limits of ad hoc networks [9].

We envision therefore an architecture as depicted in Figure 1, where TAPs equipped with sector antennas allow for geographically focused as well as omni-directional transmission of data. Being equipped with multiple antenna arrays, TAPs can form multiple orthogonal beams and hence communicate with different destinations simultaneously. Likewise, simultaneous transmissions can be separated in frequency via use of multiple orthogonal subbands (e.g., IEEE 802.11b has 3 non-overlapping channels and 11 channels within the WiFi unlicensed spectrum). In practice, the actual number of possible simultaneous transmissions will be limited by the number of actual air interfaces that can be mounted on a TAP.

Since TAPs are not mobile, their relative spatial location does not change. This stability allows the use of directional transmission, known as *beamforming*. Beamforming improves the system throughput in two ways. First, there is an increased received energy at the destination as well as a higher per-link capacity because beamforming does not spread its energy in all directions. Second, directional transmission creates little or no interference to ongoing transmissions to and from mobile units, which increases spatial reuse.

Building a TAP architecture introduces new research challenges at the physical layer. First, state-of-the-art beamforming techniques [4] assume that only either sender or receiver are equipped with multiple antenna elements, but the TAP architecture assumes both. Second, MIMO space-time encoding (e.g., [26]) assumes that antenna elements are spaced sufficiently far apart to create independent fading at each element so that the antenna beam patterns are *not* focused, whereas TAPs require focusing.

Thus, using directional antennas, an interconnected TAP wireless "backbone" can be formed with high speed and a high degree of spatial reuse. This backbone efficiently forwards traffic from and to multiple *wired TAPs*, which additionally have a connection to the wired Internet with possible capacities up to 100s of Mb/sec (e.g., Ethernet, Gigabit Ethernet, and OC-X access links). Since the TAPs may not necessarily provide complete coverage for economic or environmental reasons (obstructions), mobile users, such as G and H in Figure 1, can have their packets forwarded by other mobile users over multiple hops before reaching a TAP.

This combination requires fundamental research in deriving the transmit and receiver array coefficients to maximize signal to interference plus noise ratio (SINR) at the receiver while ensuring that the ongoing transmissions do not suffer any degradation in SINR.

III. COORDINATED AND OPPORTUNISTIC RESOURCE MANAGEMENT

To achieve system-wide high performance, the TAP network must address fundamental new challenges to coordinate and opportunistically exploit available resources system-wide.

Premise 2: Opportunistic selection of high-quality paths, sub-bands, and channels is required due to fast timescale of variations in channel conditions and the availability of multiple paths to and from wires.

The traffic behavior of a TAP network is unique in two ways. First, unlike cellular and ad hoc networks, traffic does not have a unique, fixed destination, but rather can be delivered to "any wire." In Figure 1, data from MU I can reach the wire via TAP B or E. Based on prior information about the end-to-end available bandwidth on each route (Section IV) and fast timescale channel measurements (Section II) traffic can opportunistically be scheduled to the best current path. Second, TAPs are equipped with more than one air interface, thereby enabling more than one simultaneous channel via beamforming or orthogonal frequency bands.

The scheduling challenge is to design a distributed *opportunistic multi-channel, multi-destination* scheduler. In an ideal case, a scheduler could utilize information regarding channels and queue backlogs of *all* flows to maximize throughput by exploiting high-quality channels, best-quality paths to different wires, and multiple air interfaces. In practice, this decision must be made while incorporating the distributed nature of TAP resource coordination (TAPs do not have perfect knowledge of other TAPs and MUs), subject to constraints on limits on the number of simultaneous transmissions (imposed by the number of air interfaces), and subject to balancing transmission of ingressing and transit traffic to provide a fair allocation of time shares in the system.

Premise 3: Avoiding contention by adaptively selecting backoff times allows the system to scale.

Scalability is seldom associated with scheduling and medium access protocols. Yet, as the number of users increases, the amount of side information (about channels and queues) and overhead in contention resolution increases without bound for current protocols, making them unfit to scale for our envisioned system.

Common random access MACs are limited in scalability because contention incurs long backoff periods and high collision rates, thereby severely throttling system goodput and increasing delay. If a contention-free system is assumed, information theoretic bounds predict that as the number of users in the system increases, the net throughput should increase unboundedly [13]. Thus, it is not evident from existing results whether non-scalability of contention resolution MACs is fundamentally unavoidable.

To facilitate design of load-scalable medium access protocols, the average time spent in contention per packet must be

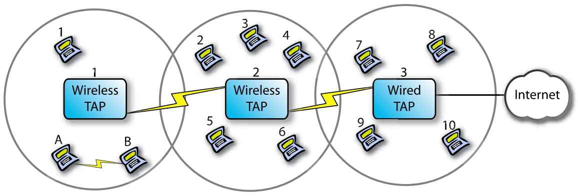

Fig. 2. Coordinated resource management

traded with queuing delay. For example, if a node acquires the channel and retains it for a duration of multiple packet transmission times, then the time used in contention per packet is effectively reduced by a corresponding factor. To achieve scalability, the time that a node retains the channel should increase with the number of contending flows. Thus, while the delay in this system necessarily increases with the number of nodes, the system goodput can scale. By further scaling the number of consecutive packet transmissions in direct proportion to the current channel conditions, MAC scaling can be integrated with opportunistic scheduling while still maintaining compliance with IEEE 802.11.

Premise 4: Any medium access and scheduling decision requires distributed resource management rather than a purely local decision.

Unlike schedulers designed for cellular and wired-AP networks, (e.g., [3], [5], [11], [17]), scheduling in TAP networks is inherently a distributed operation. Nodes in the network are not aware of the channel conditions or queue backlog of other nodes. It is evident that naive exchange about everyone's local information will lead to protocol overhead explosion and in turn, a scheduling discipline that will *not* scale. Thus, new techniques are needed to enable a scalable opportunistic scheduler for networks with distributed control.

In a first step, a centralized solution for opportunistic scheduling may provide an upper bound to performance. In a second step, the broadcast nature of the wireless medium to share information can be used to ensure scalability. Namely, by piggybacking information on data and control packets such that other nodes can overhear, nodes can obtain a partial, but necessarily incomplete view of the "distributed queue." Thus our thesis is that with perfect information and centralization, the net system throughput will grow with an increasing number of users even under fairness constraints (as in [13]), whereas with partial information sharing, scalability can be maintained albeit with stochastically bounded deviation from the centralized solution. Note that in contrast to centralized systems where scheduling and medium access are typically addressed independently, the distributed nature here implies that their effects are tightly coupled, adding to the challenge of this problem.

Premise 5: Coordinated resource management is required to eliminate spatial bias of throughput and to exploit spatial reuse.

The TAP network must ensure that all nodes in the network receive a proportionately fair share of the network capacity. Our view of fairness is that a node should get the same bandwidth share independent of whether the node is just 1 hop away from a wired TAP or whether reaches the wire via multihopping. Consider the scenario of Figure 2. Suppose that the link capacity to the wired Internet is the current bottleneck. If the TAPs provide only local fairness, each of the depicted nodes communicating with wired TAP 3 (MUs 7 to 10 and TAP 2) would receive an equal bandwidth share to the wire. However, wireless TAP 2 requires a far greater bandwidth share than MUs 7-10, as it is forwarding aggregated traffic both from its own serviced mobiles (MUs 2-6) as well as aggregated traffic from farther upstream (from TAP 1).

Consequently, flows must be throttled to ensure fairness. This throttling must be done at the first TAP to achieve efficient spatial reuse, and therefore scalability. Returning to the scenario of Figure 2, suppose that the flow from MU 1 is bottlenecked at the wire of "Wired TAP 3" to a "fair" rate of 1 Mb/sec and that MUs A communicate only locally. Then, only by throttling the flow of MU 1 by TAP 1 can MUs A and B use the full remaining capacity for their local communication.

This flow throttling is explicitly necessary for a TAP network, although, at a high level of abstraction, TCP addresses fairness and spatial reuse via additive-increase multiplicative-decrease congestion control. However, relying on TCP alone is not enough. First, TCP's congestion control has well-documented performance limitations over both multi-hop and single-hop wireless networks (e.g., [2], [6], [10]). Second, TCP's congestion control necessarily operates at *end-to-end* timescales of 100s of milliseconds – too coarse to address the fast timescale dynamics of contention and realistic channels. Finally, TCP naturally biases flow throughput to favor flows traversing fewer hops. In contrast, the objective of a TAP network is to provide fair or minimum bandwidth targets independent of spatial location.

Likewise, significant progress has been made in distributed media access and scheduling algorithms designed to balance fairness and spatial reuse objectives in ad hoc networks (e.g., [18], [22], [27]). There are two critical aspects of the TAP network that require a fundamentally new look at distributed resource allocation. First, the network has a distinct structure as compared to general ad hoc networks because the TAPs act as points of centralization through which most traffic passes. Namely, combined with the use of directional antennas, the TAP network has a unique concept of transit traffic traversing a backbone. Second, as described below, the performance objective (fairness reference model) is different for TAPs as compared to general ad hoc networks.

Premise 6. 'TAP-aggregates", and not MU flows, should be the basic fairness element.

Returning to the example of Figure 2, our notion of fairness is that all TAPs should get the same fair bandwidth share. As a consequence, however, not all MUs are given the same share: since TAP 2 is serving more MUs than TAP 1, MUs 2-6 are given a smaller share than MU 1.

Since it is impossible to achieve TAP-aggregated and per-MU fairness, we advocate for TAP-aggregated fairness for three reasons. First, TAP-aggregated fairness provides exactly the same service level that would be achieved if each of the TAPs were a traditional "wired" hot spot, namely, the MUs equally share the capacity of the local wireless channel and the wired link. Second, a fairness reference model of *TAP aggregates* enables us to design scalable coordinated resource management algorithms that would not be possible with per-MU approaches. Finally, TAP-aggregated fairness removes spatial bias of throughput that would occur with only local fairness mechanisms.

Spatial bias of throughput must be addressed by designing a formal reference model for achieving fairness and spatial reuse in TAP networks. This TAP-aggregated fairness model differs fundamentally from both proportional fairness as approximately achieved by TCP [12], [19], [20] and max-min fairness as targeted by some ATM congestion control algorithms [14]. While the solution to achieve this desired reference model for the scenario of Figure 2 is immediate, the general case provides significant challenges due to variable rate channels, MU mobility, dense TAP meshes, bi-directional traffic, etc.

Premise 7. New coordinated resource management algorithms are required to achieve the *TAP-aggregated* fairness reference model.

For the above reasons, the reference model must provide a coordinated and distributed resource management algorithm and protocols that have both a proactive and reactive component.

The *proactive* aspect of such protocols must consist of messages exchanged among TAPs to convey information about a TAP's aggregated traffic demand and channel conditions. With this information, TAPs can make a coordinated decision as to the relative service rate of ingressing and transiting traffic. The objective is to balance throttling flows to their bottleneck fair rate with more aggressive forwarding that ensures that a sufficient number of packets are backlogged at TAPs to exploit opportunistic medium access when high quality channels permit, or when contention and congestion is temporarily reduced. Addressing this issue requires the development of a performance analysis framework to gain fundamental understanding of the relationship between local channel-dependent medium access decisions and system-wide performance.

The *reactive* aspect must operate on a per-packet basis (versus per-TAP and per-MU throttling). Here, the critical issue is to ensure that each packet meets its targeted performance objective, despite multi-hopping across highly variable channel conditions. To solve this problem, coordinating packets' priority indexes among nodes is essential. In *wired* networks, a class of coordinated schedulers has been developed that allows packets that are "late" or under-serviced upstream to catch up at downstream nodes by coordinating a packet's priority index across multiple nodes [15]. In TAP networks, multi-hop coordination to best achieve system-wide performance objectives must take variable channel conditions into account and must interact with the random access MAC protocol.

IV. THE NETWORK IS THE CHANNEL – ESTIMATION, PROTOCOLS AND CAPACITY SCALING

In addition to the protocol design, which is driven by capacity and scaling issues, the TAP architecture also provides a unique possibility for protocol-driven capacity analysis.

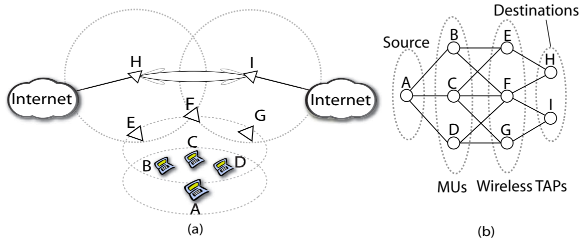

Fig. 3. *Network channel* depiction for the wireless TAP architecture for MU to wired TAP communication.

Premise 8: Information-theoretic channel capacity analyses are overly optimistic because they ignore the performance impact of protocols (e.g., MAC, scheduling, and routing). A new view of the "whole network is the channel" is needed to understand fundamental tradeoffs between protocol performance and system capacity.

The key to a high-performance scalable system is to ensure that packets consume minimal system resources to reach their destination. In particular, the scalability limitations of purely ad hoc networks [9] arise because each forwarding hop consumes additional resources. Moreover, equally crucial scaling impediments can be observed in measurement studies [8], [16] which show that actual implementations perform significantly worse than the predicted information theoretic bounds [7], [9] because they assume perfect "zero-overhead" protocols. Thus, while representing an important step in understanding the behavior of large-scale wireless ad hoc networks, existing theoretical capacity results provide limited insights on system design issues. But to understand the real-world scaling behavior of the TAP architecture as well as ad hoc networks in general, a capacity analysis that incorporates the critical impact of protocols is essential.

The design of a routing protocol for the TAP network, e.g., must contend with two unique issues not previously addressed. To achieve high performance, it is essential that the routing protocol consistently discover high-quality routes. However, this discovery must be balanced with the resulting routing protocol overhead. Second, the TAP network is inherently heterogeneous in terms of power limitations, transmission ranges, channel qualities and (wired and wireless) bandwidth. Thus, routing protocols must contend with a dynamic and highly non-homogeneous TAP backbone in addition to mobility and dynamics encountered in ad hoc networks. The challenge is to develop an analysis and protocol design methodology based on treating the *whole network as a channel*, which clearly identifies the role and the impact of protocols.

Premise 9: The "network is the channel" framework allows for an integral solution that addresses the heterogeneity in timescales and transmission modes in a TAP.

The spatial distribution of MUs and TAPs, as depicted in Figure 3(a) is noted by an individual MU as a composite channel between itself and its destination (in most cases the wired Internet). This notion of a composite channel, labeled *network channel*, is depicted in Figure 3(b), with MUs, wireless TAPs and wired TAPs represented in different sets to emphasize their difference in power limitations and capacity. Analogous to any other channel studied in information theory, the network channel has a capacity.[3] To understand the fundamental limits on protocol overhead and the network channel capacity, it is necessary to study the different timescale variations and transmission modes.

The fastest timescale variations (on the order of several packets) impact the performance of beamforming and opportunistic scheduling which utilize channel measurements made at that timescale. A fundamental bound on the capacity of beamforming for a system with M transmit antenna elements and a single receive antenna using B bits of channel information was presented in [21]. First of its kind, this bound uses *no* asymptotic approximations and is thus valid for all practical cases of interest. These results can form the basis to study the relationship between channel coherence time and the channel measurement rate for TAP to TAP and TAP to MU communication, where the receiver can have more than one receive antenna element.

At longer timescales, variations in traffic patterns, channel conditions, and contention impact network capacity such that coordinated resource management using message passing is essential for fairly throttling flows and maximizing throughput and spatial reuse. While increased protocol information on network channels at this time scale can provide increasingly precise control, the overhead of message exchange will eventually overwhelm performance. A delay-limited capacity theorem characterizes the fundamental relationship between queuing delay and average transmit power for single link communication [24]. Based on this work, new capacity results can be derived that consider the case that only limited information is available to MUs from fast timescale channel estimation and coordinated resource management thus providing a realistic characterization of capacity and scaling.

Finally, at the longest timescales, node mobility leads to unpredictable changes in the probability distribution function (pdf) governing channel variations. This leads to a fundamentally different situation compared to traditional information theoretic analysis where transmitters and receivers are assumed to know the channel pdf. However, in multi-hop networks like TAP networks, nodes are unaware of the network channel pdfs and must estimate them as a precursor to actual communication. With the above conceptual organization, we observe that *routing protocols are network channel estimators*.[4] Similar to the establishment of the relationship between the number of channel measurements and long-term route throughput for simple linear topologies [25], the routing protocol overhead is related to the level of network mobility and the resulting system capacity.

These analytical tools provide critical foundations for a complete scaling analysis that incorporates protocol overhead in measuring fast and slow timescale channel variations, the impact of traffic patterns on spatial reuse, and the relationship between routing overhead, mobility and quality of discovered paths. Such an analysis is particularly crucial for a TAP network: because it includes protocol overhead at various layers, it can already be employed in protocol design to study scalability and throughput limitations.

Premise 10: The "network is a channel" view allows for designs of hybrid and scalable routing protocols.

Since the bandwidth and stability of MU and TAP links differ significantly, a two-tiered hybrid routing protocol is required to exploit node heterogeneity. In particular, because the TAP to TAP links have relatively high reliability and bandwidth, a proactive (periodic) routing protocol is needed for TAP to TAP routing. In contrast, routing to and from MUs can be reactive to address mobility and the variable channel dynamics of MUs.

To address the scalability challenge, the following two key innovations are required. First, by decoupling the TAP to TAP routing from routing involving MUs, DSR-like routing only to the first TAP ensures that requests originating from MUs never traverse the TAP network. Furthermore, the route request from MUs can target the nearest TAP(s) and can be restricted to traverse the MU's local neighborhood. This restriction bounds the average path length traversed by route requests, resulting in improved traffic scalability [16], and providing a foundation for scalable routing. Moreover, by exploiting the *network is the channel* framework, the overhead in discovering new and better routes can be balanced with the quality of the resulting paths.

Second, a scalable location management protocol is required. We envision a distributed system of "home agents", similar to Mobile IP [23], located at TAPs. In particular, each mobile unit can register with the closest home agent. This agent is also used to discover the intended MU for traffic originating from a wire. The registration can be *reactive* and be performed in the process of uplink route discovery initiated by MUs. Note that this contrasts to the current cellular approach, which is completely proactive in nature. During route discovery over the TAP backbone (which is not the same as MU to TAP route discovery), either one or more TAPs can be associated with the intended MU.

V. INDUSTRIAL EFFORTS

The long-standing vision of a high-speed ubiquitous wireless web has also attracted several companies. Ricochet Networks,[5] a daughter company of Metricom, was the first to deploy a commercial architecture with multi-hop wireless transmission consisting of a grid of proprietary "radio receivers" spaced within a half mile of each other, and covering 17 metropolitan areas. Unfortunately, Metricom's approach led to economic failure and eventually bankruptcy in July 2001. While quite innovative compared to alternative solutions at the time, Metricom failed technically at many levels in both its architecture and protocols. It did not achieve scalability (the 50,000 subscribers were spread over half as many radio receivers), nor high performance (peak rates were limited to 128 kb/sec), nor cost-effective deployability (high deployment

---

[3] The capacity of the network channel is the maximum rate at which the source node can transmit such that it can be reliably (with vanishingly small probability of error) received at the destination.

[4] Note that the objective is not necessarily to form a highly accurate estimate of the network channel, but rather to obtain an estimate that satisfies the routing objective such as finding a minimum hop path subject to performance constraints.

[5] `http://www.ricochet.com`

and operating costs without exploitation of economies of scale for many users resulted in high subscription costs and a small subscriber base). Such past failures highlight the need to rethink the fundamentals of algorithms and architectures for large-scale wireless systems and illustrates the requirement to leverage the attractive economics and installed base of existing IEEE 802.11 hardware.

There are also many ambitious industry efforts that provide a small piece of the solution for a wireless Internet, such as directional antennas (e.g., AirNet, SkyPilot, Vivato), multi-hopping (e.g., MeshNetworks, RoofTop Communications), IP-centric base stations (e.g., Flarion), and Hot Spot operators (e.g., Boingo, Cometa, T-Mobile). While their success or failure is not yet established, the missing link for achieving scalability, deployability, and high performance is not simply integration of these components, but rather requires holistic and fundamental research into the foundations of TAP-like architectures, an objective that is not being addressed by any current industrial effort.

VI. CONCLUDING REMARKS

The development of the described TAP architecture impacts a set of critical application scenarios. First, by removing the dominant costs of hot spots associated with wired infrastructure, a wireless TAP network will provide an economically viable and deployable architecture to provide large-scale high speed wireless access to large user populations. In particular, TAPs will exploit the cost-effectiveness of mass market wireless devices that have driven markets to the $50 access point, a cost-performance curve that cannot be achieved by a fully wired AP infrastructure due to the physical necessities of wires (such as the expenditures of digging trenches and laying cables) and their device requirements (such as router line cards). These advantages will enable large-scale WiFi-based deployments with broad coverage versus today's small-scale hot spot islands.

Second, TAPs provide a key technology for true broadband to the home. Today's broadband-to-the-home efforts require that each person independently purchase a relatively low speed (100's of kb/sec) "broadband" connection from an ISP. The resulting high costs, moderate data rates, and requirement to use existing infrastructure (phone or CATV lines) has resulted in disappointing service and a dismal penetration rate of less than 10% of households. With TAPs, communities (through local governments) or new access providers can purchase a neighborhood T3 connection (for example) and deploy pole-top TAPs to provide low-cost, high-performance broadband to the home. Since the TAP network aims at maintaining IEEE 802.11 compliance, users are not required to buy expensive cards. They rather can use the same WiFi cards at the office and wherever a TAP network is available.

Thus, by addressing the outlined challenges, the TAP network provides a critical foundation for the wireless Internet, and has the potential to transform from today's frustratingly slow, overpriced, unreliable wireless data services into a new wireless Internet at an unprecedented scale, economy, deployment, and performance.

REFERENCES

[1] A. Baer. The Wi-Fi boom; out and about, and online. The New York Times, December 12, 2002.

[2] H. Balakrishnan, V. Padmanabhan, S. Seshan, and R. Katz. A comparison of mechanisms for improving TCP performance over wireless links. In *Proceedings of ACM SIGCOMM'96*, Stanford, CA, August 1996.

[3] V. Bharghavan, S. Lu, and T. Nandagopal. Fair scheduling in wireless networks: Issues and approaches. *IEEE Personal Communications*, 6(1):44–53, February 1999.

[4] J. S. Blogh and L. Hanzo. *Third Generation Systems and Intelligent Wireless Networking: Smart Antenna and Adaptive Modulation.* John Wiley & Sons, Ltd., 2002.

[5] S. Borst and P. Whiting. Dynamic rate control algorithms for HDR throughput optimization. In *Proceedings of IEEE INFOCOM'01*, Anchorage, Alaska, April 2001.

[6] Z. Fu, P. Zerfos, H. Luo, S. Lu, L. Zhang, and M. Gerla. The impact of multihop wireless channel on TCP throughput and loss. In *Proceedings of IEEE INFOCOM'03*, San Francisco, CA, March 2003.

[7] M. Grossglauser and D. N. C. Tse. Mobility increases the capacity of ad hoc wireless networks. *IEEE/ACM Transactions on Networking*, 10(4):477–486, August 2002.

[8] P. Gupta, R. Gray, and P. R. Kumar. An experimental scaling law for ad hoc networks. Technical report, University of Illinois at Urbana-Champaign, http://decision.csl.uiuc.edu/∼prkumar, May 2001.

[9] P. Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46(2):388–404, March 2000.

[10] G. Holland and N. Vaidya. Analysis of TCP performance over mobile ad hoc networks. In *Proceedings of ACM MobiCom'99*, Seattle, WA, August 1999.

[11] G. Holland, N. Vaidya, and P. Bahl. A rate-adaptive MAC protocol for multi-hop wireless networks. In *Proceedings of ACM MOBICOM '01*, Rome, Italy, August 2001.

[12] F. Kelly, A. Maulloo, and D. Tan. Rate control in communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research Society*, 49:237–252, 1998.

[13] R. Knopp. Achieving multiuser diversity under hard fairness constraints. In *Proceedings of IEEE International Symposium on Information Theory (ISIT'02)*, Lausanne, Switzerland, June 2002.

[14] H. T. Kung and R. Morris. Credit based flow control for ATM networks. *IEEE Network*, 9(2):40–48, March 1995.

[15] C. Li and E. Knightly. Coordinated multihop scheduling: A framework for end-to-end services. *IEEE/ACM Transactions on Networking*, 10(6):776–789, December 2002.

[16] J. Li, C. Blake, D. DeCouto, H. I. Lee, and R. Morris. Capacity of ad hoc wireless networks. In *Proceedings of ACM MobiCom'01*, Rome, June 2001.

[17] X. Liu, E. Chong, and N. Shroff. Transmission scheduling for efficient wireless network utilization. In *Proceedings of IEEE INFOCOM'01*, Anchorage, Alaska, April 2001.

[18] H. Luo, S. Lu, and V. Bharghavan. A new model for packet scheduling in multihop wireless networks. In *Proceedings of ACM MobiCom'00*, Boston, MA, August 2000.

[19] L. Massoulie and J. Roberts. Bandwidth sharing: objectives and algorithms. In *Proc. of IEEE INFOCOM'99*, New York, March 1999.

[20] J. Mo and J. Walrand. Fair end-to-end window-based congestion control. *IEEE/ACM Transactions on Networking*, 8(5):556–567, October 2000.

[21] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang. On beamforming with finite rate feedback in multiple antenna systems. *Submitted to IEEE Transactions on Information Theory*, 2002.

[22] T. Nandagopal, T. Kim, X. Gao, and V. Bharghavan. Achieving MAC layer fairness in wireless packet networks. In *Proceedings of ACM MobiCom'00*, Boston, MA, August 2000.

[23] C. Perkins, editor. IP mobility support. RFC 2002, October 1996.

[24] D. Rajan, A. Sabharwal, and B. Aazhang. Delay bounded scheduling of bursty traffic over wireless channels. *IEEE Transactions on Information Theory*, 2003.

[25] A. Sabharwal. On the capacity of relay-assisted communications. In *Proceedings of IEEE GLOBECOM '02*, Taipei, Taiwan R.O.C., November 2002.

[26] V. Tarokh, N. Seshadri, and A. Calderbank. Space-time codes for high data rate wireless communication:performance criterion and code construction. *IEEE Transactions on Information Theory*, 44(2):744–765, March 1998.

[27] N. Vaidya, P. Bahl, and S. Gupta. Distributed fair scheduling in a wireless LAN. In *Proceedings of ACM MobiCom'00*, Boston, MA, August 2000.