PROJECT	IEEE 802 Broadband Wireless Access Study Group
DATE	March 5, 1999
TITLE	**Functional Requirements for Broadband Wireless Access Networks**
SOURCE	James F. Mollenauer Ensemble Communications 37 Silver Birch Road Newton, Massachusetts 02468 Phone and fax: 617-244-0077 E-mail: jmollenauer@TSAnet.com
DIST	IEEE P802.BWA Study Group
ABSTRACT	This contribution addresses the functional requirements for the forthcoming IEEE 802.BWA standard. It suggests what the standardized equipment should be able to do. The most significant issue is that an 802BWA network should be able to support data, voice, and video applications, both in their legacy form and in emerging forms such as voice over IP. The standard must address the MAC and physical layer aspects of systems in conformance with the charter of IEEE 802. Emphasis should be on point-to-multipoint networks operating in the vicinity of 28-31 GHz, but flexibility in upstream/downstream balance as well as in frequency usage and distance assumptions would be very desirable.
NOTICE	

I. Introduction: the IEEE 802 context

Project 802 was set up by the IEEE in 1980 to write an official standard for Ethernet local area networks (LANs), and then go home. Nearly 20 years later, it still finds that it has not exhausted the list of LAN-related standards that it is called to work on. Much of the reason for this is that the success of Ethernet in the marketplace has resulted in Project 802, known as the LAN/MAN Standards Committee, being the normal venue for work on local area networks, and also metropolitan areas networks (MANs) as well.

A standard is in many ways like a normal product. It has a market window; if it misses the window, its value is greatly diminished. If it catches the imagination of the user population, it will do well, but it can fall flat if not adequately publicized. And it is not forever static; in the interest of finishing the design, bells and whistles can be left for next year's version.

As with other products, it helps a great deal to know what the goals are before we start designing. Selecting the right goals and establishing commitment to them are half the battle in a successful project, but they should not take half the total time. Having established the broad functional requirements, we can start designing hypothetical equipment that will be defined by the standard we write. Fine details of the functional requirements need not be finalized before we start working on the standard per se; they will fall naturally out of the discussion of particular aspects of the standard.

It is important to bear in mind that the utility of standards depends on establishing interoperability and on achieving economies of scale. Internal details of the products are not at issue: all vendors are free to innovate, but across the air interface the right bits have to be present.

A. Standardizing the MAC and PHY layers

Early on, the scope of Project 802 was defined to be the physical and medium-access layers, the lower 2 layers of the OSI Reference Model. This kept the work distinct from activity in, for example, in the Internet Engineering Task Force on the TCP/IP family at layers 3 and 4. While the TCP/IP standards were separate from the official international standardization on layers 3 and 4, the IEEE work has always been the definitive work for international standards on LANs in layer 1 and 2. IEEE standards have been jointly published with ISO, the International Standards Organization, although the dominance of TCP/IP has left the continuing benefit of international standards somewhat open to question.

Since we do not control the upper layers, we need to be able to work with existing (and possibly future) upper layers. If we cannot provide the services needed by particular higher layers, then our standard will not be used. Clearly, the wider the variety of higher layers supported, the wider the market for our equipment, within limits of cost-effective implementation.

The way to serve multiple higher layers is through convergence sublayers. In a layered model, it is always possible to split a given layer into two layers; if there are two alternates at the same level, then they can be shown side by side. Also, it is possible to have more than one level of a given type in the stack. For example, a packet-based MAC can carry ATM cells as payload.

These concepts are shown in Figure 1.

<table>
<tr><td>IP</td><td>ATM</td><td>(PBX)</td></tr>
<tr><td>LLC</td><td rowspan="2">ATM convergence</td><td rowspan="2">STM convergence</td></tr>
<tr><td>Packet convergence</td></tr>
<tr><td colspan="3">MAC</td></tr>
<tr><td colspan="3">PHYSICAL</td></tr>
</table>

Figure 1: Convergence layers. The same physical layer is common to all upper layers. The MAC layer is responsible for scheduling the base station's transmissions and for allocating requests for upstream transmissions made by users. The packet convergence layer is a part of the MAC layer; it converts between the format needed by the MAC and any format needed by the higher data layers. LLC provides logical link control; often it is omitted. Likewise, a different convergence layer operates for ATM, interworking between ATM cells and the format used in our MAC layer. Finally, STM, or synchronous transfer mode, is the mode generally used for voice lines utilizing T1 and other fixed-rate technologies. It too requires a convergence layer.

B. Speed range

In network speed, the scope of Project 802 was taken originally to be from 1 to 20 megabits per second. Telephone-based networks fell below the limit, and existing work on FDDI (done in a non-IEEE committee) was above 50 megabits per second. This appeared to leave considerable room above the 10 megabit Ethernet rate. In fact, Ethernet eventually went to 100 Mbps and later to a gigabit, but the speed limit was first breached with the advent of metropolitan area networks; the 802.6 standard could run over Sonet up to 155 Mbps. Along the way, the speed maximum for Project 802 was removed, but the group is still focused on the MAC and PHY layers.

C. Use of shared media

For the first 10 years of its life, Project 802 concentrated on standards for shared-medium networks. In effect, the wire was the switch. The burstiness of data traffic made it desirable to have the full bandwidth of the LAN on an occasional basis rather than a small part of the bandwidth all of the time, which switches generally provide. The latter case fits voice quite well, but not data. However, economies of installation dictated moving to twisted pair, which did not support the shared-medium mechanisms developed for coaxial cable. Soon central hubs evolved into switches that could support many users transmitting at the same time, greatly raising the total capacity of the LAN.

While Ethernet is now normally star-wired, the shared-medium approach is still necessary in branching-coax environments like cable TV and in wireless. Hence the problem of achieving good efficiency in the face of many users wishing to transmit is one that must still be solved in the broadband wireless context.

D. Supporting multiple services

Another area of evolution has been a result of the great convergence between data, voice, and video. As voice and video are digitized, they behave more like data in some respects, but not all. The mechanisms adopted early in the life of Project 802 to ensure clean protocol layering for data no longer suffice. However, we are not the first to break the ice on these issues. In the mid-1980's, the emerging work on metropolitan area networks in 802.6 and the 802.9 isochronous LAN established the validity of serving applications with requirements differing from data.

The normal protocol layering model for LANs shows the MAC as a sublayer sitting under the LLC (Logical Link Control) sublayer. LLC has functions which are useful for data in some contexts, but it is not necessary to support TCP/IP and is omitted as often as not. The IP network layer sits above LLC or directly on top of the MAC convergence layer. The convergence layer might, for example, add header information that would identify the data as part of a packet on arrival across the wireless link.

However, this layering is not appropriate in other cases. ATM, for example, may be supported by the MAC. In that case it is layered on top of the MAC, with a convergence layer in between to perform such functions as segmentation of the packets into ATM cells. Legacy voice lines may also be supported; in that case a convergence layer to interface with the telephony application (e.g., PBX) would also sit directly above the MAC. This is shown in Figure 1.

E. Private versus public networks

Likewise, IEEE 802 networks have also broadened their applicability from private networks to include public metropolitan networks. A LAN may administrated informally because the same enterprise pays all the bills and has a variety of methods to enforce responsible behavior on the part of users. A public network, however, requires more formality in the way it provides service and assures its customers that their needs will not be adversely impacted by the actions of others. This area has already been addressed by the 802.6 standards, which form the basis of the

SMDS service offered by many carriers here in the US and the similar CBDS service in Europe, and also by the 802.14 standards.

II. Applications supported

A. The role of broadband access

We expect that wireless broadband will serve as an access network: it will be the first network encountered by the customer's traffic outside the building. That traffic will be delivered to a transit network (i.e. WAN) to take it to its destination. Communication between two stations served by the same base station will be the exception, not the rule.

Meeting external networks

Serving as an access network means that the external interfaces that must be met by the 802.BWA network equipment are largely already defined. Ideally, the user's application should see no difference whether an access network is present or not. A router should operate the same way as if it were connected directly to a Sonet link; a PBX should not know that its T1 interface is delivered by wireless rather than fiber or copper.

This is not to say that we should define the mechanisms used to for interworking between, say, Ethernet and our air link. On the contrary: we must concentrate on interoperability, so that one vendor's base station works seamlessly with another vendor's user station. To do this, we must concentrate on the air interface and define the bits that flow between the user station and the base station. How the user station handles received traffic in order to get it on an Ethernet in the building is not an issue that affects interoperability. Each vendor can exercise maximum ingenuity to make this happen efficiently and economically.

Exposed inter-layer interfaces

The OSI Reference Model defines 7 layers, but interfaces between them need not be standardized except with respect to their logical function. Exactly how the data moves between the layers is up to the implementation. An exposed interface between the physical layers at the bottom of the stack on each side of the communication link is all that is required to achieve interoperability: the bits transmitted contain the contributions of all the layers above.

Sometimes it may be useful to define an additional exposed interface. For example, the ATM Forum defined the Utopia interface to the physical-layer chip; the IEEE 802.4 committee defined an exposed modem interface. Such interfaces should be optional; they are not required in order to achieve interoperability across the communication link.

The link to the outdoor equipment

It is important that such an exposed interface, the indoor-outdoor link in our case, not be specified until after the MAC and PHY are stable. The IOL must be able to carry in a well-specified way all the information needed to operate the actual rooftop unit, and until we define the MAC, and especially the physical layer, we will not know what information is required.

B. Applications

Data
Voice
Video

In this era of convergence, the choice of applications to support is easy. Data, voice, and video are the normal choices, and for good reason. In wide-area networks, voice has been the largest part of the traffic; it is just now being equaled by data propelled by Internet usage, and both may well be overtaken by digital video in the future.

Legacy services: fixed rate

The easiest way for a user to justify the adoption of a new technology like broadband wireless, committing equipment money and risking existing applications, is to save substantial money. Saving money now being spent is a more powerful argument than reducing potential future expenditure. To provide immediate savings, the wireless network must be capable of taking over at least a part of the current communications load of the enterprise. This means serving legacy applications without impacting them adversely; in most cases it will not be practical to modify them for a different communications medium.

These legacy applications are typically based around fixed-rate, fixed-delay links. These range from 56-kilobit DDS lines through T1, E1 and on up to Sonet and SDH. They have evolved, of course, from the needs of voice traffic, which is well served by symmetric bandwidth at a constant rate.

One aspect of these line types is worth mentioning: they are widely spaced in rate. A T1 line is 24 times faster than a single voice line, and T3 is 28 times faster than T1. These speeds were not devised to match user needs; rather they were designed for internal multiplexing within the carrier networks. Only later were they made available end users as the users' requirements increased.

Therefore it would be desirable for the BWA network to offer fixed-rate access with a finer granularity of steps than is otherwise available from the basic telecommunications line rates. This flexibility could be utilized by the system operator in providing a variety of service plans for the user.

Supporting data: bursty rates

Even though data is well known to exhibit highly bursty behavior, the wide-area lines used to interconnect user sites are still fixed-rate. This rate defines the maximum amount of data that can be transported; of course lower amounts are sent most of the time. For an individual user sitting at a computer, the communication load may well exhibit a peak/average ration of 100-1 or even higher; when aggregated with other users over an external link from the building, the bursty properties of the traffic are still very evident. Backbone traffic at Bellcore was shown about 5 years ago to have fractal characteristics, quantitatively as well as qualitatively. The randomness of the traffic was equally apparent when displayed over time lines varying by 5 or 6 orders of magnitude.

One reason that IP networks are being widely adopted for all data applications is that they have the ability to do statistical multiplexing between users. The Internet backbone lines are dedicated to no single user, and the costs borne by users are dictated much more by their average usage than by their peak rate. The point-to-multipoint wireless access network can provide similar statistical economies to the access links, substantially reducing costs. While one user on a base station is transmitting heavily, another is idle; in effect the network capacity is much greater than a collection of fixed-speed lines with the same nominal total rate. Such economies are not possible on any dedicated medium running to the user premises.

On the user premises, we can expect that most of the traffic will flow over Ethernet, generally the 100 Mbps Fast Ethernet version. Average rates on the Ethernet lines of course are very much lower than 100 Mbps. The access network should provide variable bandwidth that can respond to peaks in usage that represent a fair fraction of 100 Mbps: perhaps 40 or 50 Mbps. If history is a guide, this number is sure to rise over time. This traffic may be delivered to the public Internet, to virtual public networks (VPNs) built on top of the Internet, to frame relay networks, and to LAN bridges; we should be able to work smoothly with all of them.

Compressed video: somewhere in between

Video is a great bandwidth consumer. Unlike the desktop computer scenario, video links (most probably video conferencing and in the future video phone calls) are not bursty at the level of 100-1, but more like 3-1. Video compression algorithms generally produce variable-rate output; it requires an additional step to force this output into a fixed-rate output channel, typically by discarding precision during peak times. Variable-rate video can be combined with data, with the data packets filling the gaps in the video. For the commercial market, rate requirements are not as severe as for the home: video conferencing is now run at about 384 Kbps, compared with several megabits per second for digital standard-definition video.

The advent of priorities in IP networks will make it practical to run video conferencing over private IP networks and even the public internet. Such video will run in normal IP packets, but they will have delay and discard priorities that will be distinct from normal packets.

Priorities

The key to supporting all these applications is careful application of priorities. With a small number of priority groups, it is possible to support almost all prospective applications. This is the direction of wide-area facilities like IP networks as well. The benefits of economy and widespread connectivity have made TCP/IP the protocols of choice for many applications with more stringent requirements than data. As a result, priority mechanisms are now being added to the TCP/IP protocols under the flag of Differentiated Services. Voice over IP seems to be the biggest thing to hit the telephone industry since divestiture; to be viable in the marketplace we must strive to be consistent with these efforts.

Basically, the various needs can be served by prioritizing the traffic as follows:

1. Constant bit rate

2. Real-time variable bit rate

3. Variable bit rate, non-real-time

4. Available bit rate

Constant-rate traffic would be scheduled first in each scheduling cycle; it generally has stringent delay limitations as well as a need for a constant bandwidth.

Real-time services like video conferencing would come next; it includes mainly statistically-multiplexed video, though IP voice would also fall in this category since it is typically compressed and the bit rate depends (among other things) on who is talking.

High-priority data would be the next category.

Available-bit-rate follows last, serving normal data, which can tolerate reasonable delays waiting for the higher-priority classes to go through. (Naturally, a continuing overload causes problems in any network.) ABR traffic sources are able to accommodate their transmission to the amount of capacity that is available at the moment. This is true of all LAN equipment, which operates under the control of the MAC protocol.

The MAC is the key to serving all of these varied data types effectively. It must be able to handle all of the above services, in any proportion. Wireless access networks are new, and it is very hard to tell what will be the most important uses of the technology in the immediate future, let alone 5 years from now.

III. Operating environment

A. Who are the customers?

Not likely consumers in their homes. They already have two ways of getting high-speed access without installing new wiring: cable modems and ADSL. While these technologies are not yet widely installed, the telephone industry and especially the cable industry are pushing fast to offer service wherever they can. And since broadband wireless has a limited range, regions that are too rural to justify cable TV are likely not to justify wireless either.

Also, homes, especially those of people with sufficient disposable income to be early adopters of a wireless service, are well surrounded by trees; moreover they lack high rooftops where antennas could be located. Tests in Sweden showed only about 30% of homes within an adequate line of sight.

Commercial customers are almost certain to be the best market. Cable TV facilities rarely come near them, and the types of telephone wiring used (such as T1) cannot support ADSL. For commercial customers to get high bandwidth, installation of fiber is necessary. Wireless can compete very effectively against fiber, if the cost of the wireless equipment is balanced against the cost of digging up the street to bury the fiber.

As a result, we should optimize the standard to serve commercial users rather than homes. If wireless service providers wish to target homeowners, that is fine, but it is not likely to be the largest part of the market.

B. Business considerations

Access networks are public facilities, like wide-area networks but unlike LANs. Third-party service providers transport bits for a fee. Who the service providers are may vary from country to country, but the model of a third party is not changed if it is the same third party who provides WAN services as well as access.

The important thing is that the relationship is a formalized one with a service contract. If the network does not perform as intended, or fails frequently, business may evaporate or penalties may be invoked. Use of a standardized network is a way to minimize such misfortunes, if the standard is well drafted.

Aside from the capability of offering a well-characterized service, perhaps the most important thing for the service provider is to isolate one user from another. There should be ways for the service provider to allocate transmission opportunity such that users are isolated from each other, and if the user does not abide by service constraints, to cut off the service. It is be desirable for multiple users within the same organization to compete against each other for a share of the organizations's contracted service, without impacting service rendered to other organizations.

Wireless networks will find their place in the spectrum of communications facilities because they require less infrastructure than wireline networks. The last mile is the most expensive part of any network because there are far more end users than backbone nodes. If the carrier can avoid digging up the street or leasing rights of way from power companies and the like, the network can be much more competitive.

C. Encryption

Privacy is another critical issue. Especially with data going through the air, physical isolation must be enhanced by encryption. This is a very active development area currently, because of considerations of privacy over the Internet and also because cable modems are another public network with low physical isolation.

Officially, LAN security is the province of the 802.10 group. We should review their work to see how well it applies to our case. In general the steps that the wireless access network needs to take include authentication, key exchange via public-key methods, and then data encryption perhaps with DES or triple-DES. More modern and stronger cryptographic algorithms, such as IDEA from Europe should also be considered. We will be sending proprietary corporate information across the countryside with no physical protection. The appearance as well as the fact of strong security is vital in retaining the customer's confidence.

D. Physical considerations

Radio propagation in the vicinity of 30 GHz is not an easy issue. Absorption due to moisture in the air varies greatly and is much higher in heavy rainstorms. Robustness in the face of such problems must still be the goal, if the wireless network is to be taken seriously for critical business applications.

As a starting goal for availability, I recommend 4-nines: 99.99% of the time rain absorption will not prevent operation of the network, based on published rain figures. This may dictate rate backoff when received signal power falls too far; this is a subject for study in the committee. As a last resort if the data rate falls too drastically, we can back off to 99.9% availability. In short there is a multi-way tradeoff between range, frequency re-use, availability, error rate, spectral efficiency, and transmitter power. Proposals will need to be made on these items.

As a rule, antenna sectorization (4-way, 6-way, etc.), polarization, and frequency reuse among cells are not issues that affect interoperability. However, we may find it desirable to publish a set of guidelines on these issues in order to assist network operators in getting the most out of the standard systems.

IV. Media access control (MAC) Layer

The MAC layer is crucial in enabling the network to support a wide variety of applications with good efficiency. Point-to-multipoint wireless is like cellular telephones in its cellular structure; it is also equivalent topologically to cable TV: users at varying distances transmit to a central station. They do not hear each other, but receive only from the base station. The base station is the only downstream transmitter, but upstream traffic from the user stations must be controlled in a way that minimizes their mutual interference.

This is best accomplished with a request/grant mechanism. Users make requests to the central scheduler, which issues grants specifying when each user can transmit. There may be polling or a contention period during which the requests can be made; contention-based requests are subject to collisions, but most of the time the system operates free of collisions. In many cases, the protocol allows the user station to make additional requests during scheduled transmission times (piggy-backing). This further reduces collisions. Also, constant-rate allocations can be free of collisions since they come at predictable times in each scheduling cycle.

The scheduling process at the base station is relatively straightforward. In each cycle constant-rate traffic is allocated first, then real-time, then priority data, and finally normal data traffic. If there are excess requests for data traffic, the requests can be held for the next cycle. Commitments for higher-priority services, of course, should not exceed the available bandwidth. Separate scheduling categories for different sets of users stations (virtual private networks) add complexity but the process is still tractable.

No one size fits all

It is important to recognize that in the access network field, there is no one size of container that fits all users. At one time it was thought that ATM would fill this role, but it is clear that ATM will never prevail at the desktop. Most carriers are still committed to ATM in their backbone networks and it would be desirable to accommodate ATM as one data type, but IP is on the rise. The adoption of priorities will extend its popularity beyond data, while backbone router speeds will soon equal the fastest ATM switches. In addition, we must find an efficient way to carry the legacy traffic that is not packetized at all: the T1 and E1 lines that are the workhorses for corporate voice.

V. Physical layer considerations

A. Frequencies

Frequency bands assigned in different countries are different. In general, this has little impact on the overall system design: the exact frequency transmitted is controlled at the last stage by the upconversion process, usually in the rooftop unit.

Some issues do exist: range is one. Rain attenuation is higher at higher frequencies hence a 38-GHz system has a shorter range than one at 24 GHz. The worst-case propagation delay assumed in the design of the MAC layer will need to be sized for the lowest frequency that the standard is designed for.

Frequency channelization is a topic that will need to be considered. Unlike the cable TV case, there is no existing practice in LMDS that dictates the channel size. With frequency blocks as large as 850 MHz, it is not possible without very expensive components to process that much spectrum all at once, generating data streams in excess of a gigabit per second. Division of the spectrum into channels will permit use of reasonably-priced components and may well provide much greater flexibility. A possible choice of channel size is 25 MHz; this would permit data rates between 30 and 100 Mbps depending on the type of modulation used. It would also allow efficient coverage of the various-size pieces of the LMDS band, where the size ranges from 150 MHz to 850 MHz.

B. Upstream/downstream bandwidth ratio

Despite the peer-to-peer communication model which has been extant for at three decades, most communication does not occur between peer entities. Telephone calls and video conferences are the only symmetric-bandwidth peer applications that come to mind. Most data communication is actually client-server, and the traffic load is asymmetric. Web traffic is a good example of this, as is the use of file servers in most organizations.

It is clearly advantageous to retain flexibility in this area. Even customers who now generate asymmetric traffic might see differences if video telephony becomes commonplace. And in any case the balance may shift over short intervals as video conferences come and go.

Hence we should retain the capability of addressing this balance dynamically, rather than leaving it rigidly fixed in either a symmetric (as in Sonet) or asymmetric state (as in cable modems).