Performance in Broadband Wireless Access Systems

White Paper
Performance in Broadband Wireless Access (BWA) Systems

As Broadband Wireless systems begin to proliferate in the market, the array of choices between systems can become mind-boggling. As with any industry, and especially in the broadband access world, system specifications and performance play a critical role in the customer's decision process. In most cases performance has a definitive meaning, and well established measurements are applied to enable meaningful comparisons between products.

In the case of Broadband Wireless Access however, this is not a simple matter. Partly because it is an emerging industry, and partly due to the complexity of the technology, there are no industry acknowledged or sanctioned test suites available which can be used to compare systems. This leaves customers on their own when it comes to comparing competing products, creating and conducting their own performance tests, often times without understanding the impact seemingly less innocuous test plan choices can have on the results.

What to Measure?

The first question that has to be answered is, "What are the relevant performance metrics that should be looked at?" When it comes to BWA, performance is typically characterized in terms of how many Subscriber Modules (SMs) can be supported by a single Access Point (AP), how far can it go, how fast in terms of Mbits/sec, and what is the latency introduced in the wireless portion of the network.

When it comes down to how many SMs per AP or cell site, this is a number that needs little clarification. A given system supports X number of SMs per AP while System B supports X+50; the comparison is simple. While these numbers can be obtained from a simple data sheet, the real life numbers may be something quite different. It is important to also consider how much bandwidth or data throughput is available to be allocated and the demands of each SM in the network. Many times the available throughput will be the limiting factor in terms of how many SMs an AP can really support.

Likewise when one wishes to determine the range of a BWA system, while there are numerous examples of exaggerations in the market, in most cases a simple examination of the receiver sensitivity specification for the radio allows an apples to apples comparison between different products.
However when it comes to how many Mbits/sec and with what latency, the problem of meaningful comparisons becomes much more challenging. Again this is in large part due to the underlying technology and the relative immaturity of the market. For example, some BWA products have vastly different results in terms of total throughput based on the farthest SM in the network, or the system load, or even how many SMs are on a given AP. Systems based on 802.11 technology for instance, have a significant drop off in data throughput based on the distance between an SM and an AP. The 11 Mbps raw data rate advertised is typically only available within the first dozen meters in a cluttered environment, and tens of meters in a clear line of site scenario. This is due to less complex modulation they employ for "longer" links and the construct of the 802.11 MAC (Media Access Control) layer.

In addition how these systems are tested can have a big impact. Ethernet/IP based wire-line networks are interacting with a network element (the wireless portion) that has a radically different MAC layer. This can have a huge impact on the results. For example when standard wire-line IP throughput tests such as FTP are used to measure throughput in a wireless system, unless care is taken in executing the tests, the results can be skewed.

**How to measure throughput and latency**

Many times the testing that is carried out on a BWA system by a customer is by necessity done with only a few SMs installed. This does not give an accurate indication of how the network will perform under load. There are numerous examples of products on the market that do well when lightly loaded, but throughput declines and latency climbs when placed under heavy traffic loads. These results are affected largely by the design of the underlying point to multi-point wireless MAC protocol.

Today, every system on the market is either an IEEE 802.11 based MAC or a proprietary one, with wide variations in performance. The net effect: how many Mbits/sec a system can deliver has to be tempered by how well the MAC layer works, and all of this must be done under real world conditions. 20Mbits/sec from product A can be significantly different from the 20Mbps available from Product B.

When it comes to measuring throughput, the FTP protocol has been
around for years, is well understood (at least in terms of how to use it) and is a simple means for measuring throughput in the overall broadband access environment. For these and other reasons, when end users install a BWA network and they wish to measure the throughput, FTP is the first and many times the only method used.

However many times the test results are skewed by factors beyond the BWA system. For example, the following parameters can significantly influence the test results, and not one of them has anything to do with the wireless portion of the network.

- File Size
- Processing power of the hardware (both client and server)
- Operating system used (both client and server)
- TCP settings such as the window size and congestion control settings
- FTP program (dedicated software or within a browser)

Given the ubiquitous nature of FTP, it will continue to be used as a throughput measurement tool, and if the above mentioned factors are understood and properly adjusted for, a service provider can have confidence in the results.

Latency in a BWA system is another metric that is difficult to compare. This measure of performance is heavily dependent on the number of SMs, the average packet size being transmitted, and the distribution of the traffic over time and amongst the SM devices. Again to truly gain an understanding of how a given system will handle latencies, testing must be done under load. Almost every BWA product on the market will give excellent, low latency results when one or two SMs are in the network. What happens when 20 SMs attempt to transmit a series of small VoIP packets? This is a more accurate test scenario and will give an operator a better idea of how the system actually handles latency in a real world situation.

**What to Expect from a Well Designed BWA system - Throughput**

As discussed above, the real measure of a BWA system throughput performance should be done under load, with varying numbers of SMs. In addition to testing under load, they should also be tested
with a small number of users. The reason for this is simple: a common but less than desirable characteristic of many BWA systems is the underutilization of the available bandwidth when only a few SMs or applications are accessing the network. Ideally the system should ramp up and provide access to the maximum available bandwidth as quickly as possible, and maintain fairness as load increases.

The Canopy system has been designed to provide maximum through-

![Figure 1. Percentage of bandwidth used vs. active SMs.](image)

put performance the network load. Figure 1 gives an illustration of how the Canopy system performs when compared to both 802.11 based systems and a typical proprietary MAC protocol. The Canopy system reaches full capacity quickly. When one SM is active on a Canopy network in point to multipoint mode, the aggregate useable throughput reaches slightly more than 4Mbps. When the second SM is added, the total aggregate throughput climbs to over 6Mbps thus fully utilizing the link bandwidth.

The second major factor affecting throughput in some systems is the distance between the SM and the AP. As Table 1 illustrates, 802.11 based systems have decidedly poorer performance as the distance grows, while Canopy delivers full bandwidth over the entire range.
Both numbers, raw (total data rate) and effective (useable data rate) are compared.

**What to Expect from a Well Designed BWA system - Latency**

There are many applications that operate over a BWA network where latency control is important, and two of the more prominent are VoIP and database queries.

As noted above, many systems provide excellent latencies when a few number of SMs are configured. This has to do with the design of the over the air frame. There are two key factors in the design of this frame that have a large impact on latency control.

The first has to do with fragmentation of the incoming Ethernet frame. Some systems have large over the air payloads as one packet, such as 802.11 based systems as well as many proprietary solutions. In this case incoming Ethernet frames are either not fragmented at all or very little. The problem with this approach occurs when many small packets arrive for transmission. In this case, a system that does not fragment Ethernet frames can either send out the air frame instantly to avoid introducing latency (at the cost of wasting air frame payload and degrading throughput) or, the system can wait to collect several small packets and send them together. This latter approach improves bandwidth efficiency but introduces significant amounts of latency.

The Canopy system actually approaches this issue by fragmenting every incoming Ethernet frame into many smaller data packets and keeping the overall airframe duration low, 2.5ms. The net effect of this design is to deliver both efficiency and low latencies when transporting small packet traffic, e.g. VoIP.

The second critical factor in a BWA system that has an impact on latency control is important, and two of the more prominent are VoIP and database queries.

As noted above, many systems provide excellent latencies when a few number of SMs are configured. This has to do with the design of the over the air frame. There are two key factors in the design of this frame that have a large impact on latency control.

The first has to do with fragmentation of the incoming Ethernet frame. Some systems have large over the air payloads as one packet, such as 802.11 based systems as well as many proprietary solutions. In this case incoming Ethernet frames are either not fragmented at all or very little. The problem with this approach occurs when many small packets arrive for transmission. In this case, a system that does not fragment Ethernet frames can either send out the air frame instantly to avoid introducing latency (at the cost of wasting air frame payload and degrading throughput) or, the system can wait to collect several small packets and send them together. This latter approach improves bandwidth efficiency but introduces significant amounts of latency.

The Canopy system actually approaches this issue by fragmenting every incoming Ethernet frame into many smaller data packets and keeping the overall airframe duration low, 2.5ms. The net effect of this design is to deliver both efficiency and low latencies when transporting small packet traffic, e.g. VoIP.

The second critical factor in a BWA system that has an impact on latency control is important, and two of the more prominent are VoIP and database queries.

As noted above, many systems provide excellent latencies when a few number of SMs are configured. This has to do with the design of the over the air frame. There are two key factors in the design of this frame that have a large impact on latency control.

The first has to do with fragmentation of the incoming Ethernet frame. Some systems have large over the air payloads as one packet, such as 802.11 based systems as well as many proprietary solutions. In this case incoming Ethernet frames are either not fragmented at all or very little. The problem with this approach occurs when many small packets arrive for transmission. In this case, a system that does not fragment Ethernet frames can either send out the air frame instantly to avoid introducing latency (at the cost of wasting air frame payload and degrading throughput) or, the system can wait to collect several small packets and send them together. This latter approach improves bandwidth efficiency but introduces significant amounts of latency.

The Canopy system actually approaches this issue by fragmenting every incoming Ethernet frame into many smaller data packets and keeping the overall airframe duration low, 2.5ms. The net effect of this design is to deliver both efficiency and low latencies when transporting small packet traffic, e.g. VoIP.

The second critical factor in a BWA system that has an impact on latency control is important, and two of the more prominent are VoIP and database queries.

As noted above, many systems provide excellent latencies when a few number of SMs are configured. This has to do with the design of the over the air frame. There are two key factors in the design of this frame that have a large impact on latency control.

The first has to do with fragmentation of the incoming Ethernet frame. Some systems have large over the air payloads as one packet, such as 802.11 based systems as well as many proprietary solutions. In this case incoming Ethernet frames are either not fragmented at all or very little. The problem with this approach occurs when many small packets arrive for transmission. In this case, a system that does not fragment Ethernet frames can either send out the air frame instantly to avoid introducing latency (at the cost of wasting air frame payload and degrading throughput) or, the system can wait to collect several small packets and send them together. This latter approach improves bandwidth efficiency but introduces significant amounts of latency.

The Canopy system actually approaches this issue by fragmenting every incoming Ethernet frame into many smaller data packets and keeping the overall airframe duration low, 2.5ms. The net effect of this design is to deliver both efficiency and low latencies when transporting small packet traffic, e.g. VoIP.
latency performance has to do with the number of what are referred to as reservation slots. When an SM has data to send, it notifies the AP via these reservation slots. A given system has so many slots per air frame. It is easy to envision a system with, let’s say, 1 reservation slot per air frame, and the frame is 10ms in duration. In this case, if there are 20 SMs with packets to send simultaneously, it will take 20 frames or 200ms before the last SM can even register its request, with several more frames before it is allowed to send the data. Conversely, too many reservation slots can end up consuming valuable bandwidth.

The Canopy system has taken this into account in its MAC layer design. With the Canopy system, with standard telephony traffic models put into play, up to 60 VoIP end users can be supported on a given AP without introducing more than 20ms of latency. As the graph below illustrates, the Canopy system delivers consistent, predictable low latency over a range of the number of Subscriber Modules accessing the network.

While this performance is very good, the Canopy system takes this one step further by allowing the service provider to dedicate a portion

![Graph showing latency vs. active SMs.](image)
of the upstream bandwidth to VoIP traffic, and then prioritizing this traffic above low priority, latency insensitive data traffic.

Other approaches to controlling latency invoke a method referred to as dedicated polling. Dedicated polling refers to the technique where all SMs are polled, or given a chance to send data every time, whether there is data to send or not. This approach, even when the polling is "smart," works well only with a few SMs in the network. For these types of systems, initial latencies will be low as a poll for a given SM occurs every 5ms or so. Thus for a network with 4 SMs, latency is limited to 5ms * 4 SMs or 20ms. If this same network has 50 SMs, the latency for all traffic climbs dramatically to 50*5ms or 250ms.

Summary

As the BWA industry moves forward and gains maturity, performance comparisons will become more and more standardized. Until that time service providers and end users will be tasked with evaluating systems themselves. The basic performance parameters that must be considered are:

- The number of SMs supported per sector or AP, not just from a spec sheet perspective but in practical terms relating to how much throughput is available.
- The actual throughput of the system
  - With SMs near and far
  - With few SMs and with many
  - Lightly loaded and heavily loaded
  - With varying types of traffic, data, VoIP, etc.
- The latency introduced by the system, under varying conditions as listed above.

In the end how a BWA system utilizes the bandwidth available via the MAC layer rather than the raw bandwidth available determines the ultimate performance in the field. The Canopy BWA system with its hundreds of engineering years of design, more than 60 patents, and hundreds of commercially deployed networks has the proven design to truly deliver Broadband Wireless Access for all applications.