Some Important Problems in Communications Theory

A View from the Front

Terrence P. McGarty[1]

Abstract

This paper presents several "problems" in communications which, in the authors opinion, if solved, may have significant impact on the world in general. The issues of "more, better, cheaper, quicker" are the driving factors that we all look towards as we envision the deployment of telecommunications in the twenty first century. We always seek more, of a better level, always at a lower price, and always implemented faster. To achieve this in the broadest possible way there are ten problems which must be solved, and their solutions, are a complex blend of technology, economics factors, and regulatory sensibility. This paper attempts to focus on those problems which, if solved, may have the greatest impact on the development and deployment of telecommunications in its broadest sense.

Contents

---

[1] This paper was written while the author was a Research Affiliate at MIT Laboratory for Information and Decision Sciences in the Department of Electrical Engineering and Computer Science at MIT. mcgarty@mit.edu and tmcgarty@telmarc.com The opinions stated herein are solely those of the author and in no way reflect the faculty, staff or students at MIT.

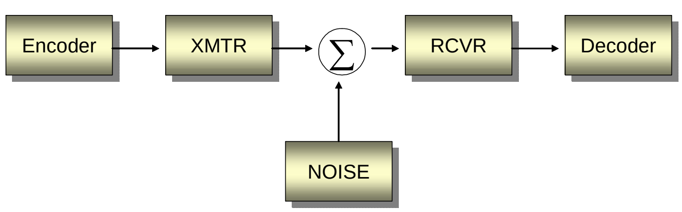
1 INTRODUCTION

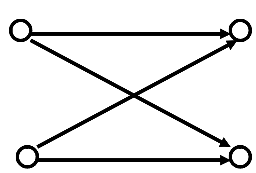
Over the past one hundred years one could pose a set of key research questions relating to telecommunications. This paper takes a large personal risk by attempting to define and characterize key problem for the area of telecommunications for the next generation, say next twenty years. The author is claiming no particular wisdom or insight other than having even an observer and in some cases a participant in the past forty years of the development, namely the last two generations. Does past in any way become prolog to the future, one could only hope. I am always reminded of Kahn's law, "what we expect tomorrow frequently never occurs, and what we anticipate in the distant future frequently occurs tomorrow".[2] Thus anyone who prognosticates the future will inevitably bear the burden of totally missing the target. The prognosticator may be focusing on all of the things which never come to be. The strength of any guesser in the future is one who has keen insight into the obvious, that is also the strength of the securities trader, focus on the next step, chess games are not always the rules of life, Brownian motion is a more likely model!

1.1 Fifty Years Ago

Fifty years ago, for example, the recent discoveries of information theory permitted questions such as what was the maximum capacity of certain links, how could one efficiently use wireless links, what modulation and coding techniques were best, and how could one create more efficient switching systems. Shannon's work was seminal at that time and the top ten topics in many ways revolved around the goals potentially attainable in his analysis.

The following is the Shannon paradigm. It is simple, it is direct, it includes enough that anyone can start here and follow the process, and anyone can generalize from this point. However it took brilliance to simplify it to this well circumscribed view. There were years of looking at communications and looking at various modulations schemes and how best to optimize them. Detailed analysis of loop noise, FM clicking, and the like, provided elegant solutions to specific problems but failed to provide the key paradigm to take the next major steps. The question is do we need to have more of these simple paradigms or is the world too complex at this time to return to this path.





The Shannon paradigm is the five elements, a combiner and the binary symmetric channel. Everything can follow from this simple example.

[2] The author hear Bob Kahn, the key player when at ARPA in creating what is now the Internet and also not at the Center for National Research Initiatives.

Thus the topics of that period were still fundamental, and formal. They asked about modulation and multiple access, they asked about technologies for transmission, they focused on coding techniques. The first set of questions were those relating to ultimate capacities and capabilities. They related to how far could we go and then how did we implement this. Whether it was the development of convolutional codes and ultimately the Viterbi decoder in the 70s, it was generally all focused on fundamentals. Transmission and channel characterization were important since we really did not know what could be done, In addition memory and processing were expensive.

The other constraining issue during this period was the existence of the Bell System. The classic story epitomizing this as told to me by Bob Kahn, was how Kahn had to work around Bell Labs and ultimately led to the development of modems which were not AT&T approved. In a meeting in the early 1970s, Kahn brought several key players to a Bell Labs meeting in Murray Hill, the Bell Labs headquarters and center for research. He wanted to deploy the ARPA net, a packet based network in the US. At the meeting attended by a mass of Bell Labs management as was the case in those days, the Bell Labs hierarchy said that he could not have access to their modems, and that he should give Bell Labs the contract and that they would build all of it and give it to him, wrapped in a cover.

Kahn stated that he needed an open system and that this was critical. They stood firm. He walked. From that he then funded such entities as would become modem manufacturers, integrated chip manufacturers, and the whole plethora of what we now take as common in telecommunications. That one event was the beach head which opened telecommunications to outside ideas, it freed telecommunications from the strangle hold of the Bell System mentality, and led ultimately to the Internet as we see it today.

1.2 Twenty Years Ago

Twenty years ago, when I headed Research for NYNEX I was asked to pose the same set of problems which we were to focus on, along with our research group at Bellcore. The topics that I focused on then were as follows:

1. Multimedia Communications
2. Network Management
3. Software Development Methodologies
4. Transaction Processing System and Networks
5. Broadband Communications Systems
6. Advanced Wireless Systems

Let me discuss each in some detail.

1.2.1 Multimedia Communications:

How could we develop a fully integrated multimedia communications system allowing uses to share multimedia objects in a fully open environment. We understood the impact of the ARPA net at that time but we did not understand how best to integrate multimedia objects. We created a system called MEDOS, the Media Distributed Operating System, which was a layered approach enabling users to fully integrate many types of multimedia objects in a conversational manner. This was before the web and Berners Lee and the enablers of the web. We actually were addressing a much more complex problem.[3] The seminal work in this area was a small book by Winograd and Flores, Understanding Computers and Cognition, which set forth a strong philosophical base for multimedia communications.[4]

---

[3] See the paper on Multimedia Communications. This was an attempt to understand what was being requested in the broadest possible way. It also attempted in a crude way to address the philosophy of multimedia communications. In the fall of 1989 I also taught a course on multimedia communications at MIT. Unfortunately it was too early and the topic needed more time to evolve.

[4] Terry Winograd and Fernando Flores, Understanding Computers and Cognition, Addison Wesley (Reading, MA), 1986. This book is the basis of a philosophical understanding of communications from a human interaction point of view.

To understand multimedia communications I drew upon several key real life experiences. In 1981 I designed the first full motion video on demand system on a cable television network, the Warner Cable System.[5] It was done jointly with participants from Atari and American Express and eventually became a joint venture in 1984 with Bank of America, DEC, GTE and Bell Atlantic. It was the first system which tries to combine transaction processing, information access, and entertainment, in a fully interactive manner. We called it the TIES system, Transaction, Information, Entertainment System. We actually deployed the first such system in Pittsburgh using a combination of what was called video text systems and also systems using both Atari equipment and IBM PCs, the first generation of them. Our video storage was a set of banks of Pioneer video disks, all analog systems, and video was distributed in packets of analog bursts. We allowed migration thru this world by means of Atari joy sticks and visual cues of a shopping mall, thus rather than space invaders we had shopping mall invaders! It worked, we had suppliers or seller and we had customers. The challenge was to understand how people communicated and how did we create an environment supportive of transactions.

However, it was the set of questions we asked in the early 80s which led to the next attempt to develop true multimedia communications. I believe that it is till not where we were looking at taking it. The web as we know it fails to create a conversational environment, it fails to create a promotional environment. Towards the latter let me recount a tale collected during the Warner days. As we were trying to effect electronic home shopping, we visited many retail stores and their owners. In Houston we met the CEO of one of the largest at the front of his largest store.

It was about 9 AM. We had told him all about the new world of electronic shopping. He said we would learn a great deal by just watching. As the store opened we saw many women approaching the store. He said, “See each of them, they are looking for something, but my job is to get them to buy something, not necessarily what they are looking for. Wait a bit and talk with them when the leave.” We did, when the left we spoke with many of the shoppers and indeed the all bought several items, and when asked why the first went to the store we found that without an exception it was for something else, not for what they eventually bought. Upon meeting again with the owner he drove thru the point that the physical experience of shopping is essential, to promote and persuade, a true multimedia environment, perfumes, visual displays, goods to be touched, and an overall environment conducive to selling, using all the senses. Thus multimedia communications is effecting all of these in a true conversational mode, then and only then can we see true results. This problem is not as clear cut as minimal bit error rate, not as simple as the channel coding theorem, but it is the challenge of true communications.

1.2.2 Network Management

Network management was a second area. In this period of time issues of managing disparate network elements were becoming an overwhelming problem. The problem was actually solved via SNMP on IP based networks. It became a simple solution by industry agreements on standards. This was a short lived problem with a simple initial solution.

However, there was still a residual issue, namely anticipating network failures and acting to mediate against them. This problem has still not been solved. In the late 80’s approaches using neural networks were tried but the solution of basic network management was adequate for most practical demands. There is a second residual issue, namely customer perception of quality versus system perception. The system may be fine but the customers think the service is quite poor. This is the classic problem of coverage in a cellular system. Frequently the US based approaches, especially the TDMA approached, have very poor coverage as well as limited capacity. Combining these two factors results in overall low customer satisfaction. This was typically what almost drove ATT Wireless out of business. The ability to combine human perception of service quality

---

[5] There were may members of the team that I led who contributed, some but not all were Glenn Shapiro, one of the most brilliant marketing people I have ever met, Brenda Pomerance, now an attorney, but the conscience for completeness, Richard Veith, the creator in almost real time of the video system and content, and Gus Hauser, our Chairman, who always asked the right question to help us focus. We also had a vice Chairman, Lou Gerstner, who was less than enthusiastic of any form of electronic marketing and distribution in an interactive manner. He went on to become CEO of IBM.

with the actual optimization of the system was an issue we attempted to address but the problem is still quite complex as well as there is a rapidly moving field of service implementations.

1.2.3 Software Development Methodologies

In the late 80s software development was costly and there was clearly a need especially in the telecommunications space to improve productivity. This was another area we investigated but it became clear that many other groups were doing so and that the whole software production process would soon change. Now, of course, we outsource this and the combination of tool and lower manpower costs have driven down the software problem. This was really never a true technology problem, it was resources and technique.

1.2.4 Transaction Processing System and Networks

Transaction processing is a complex set of tasks which allow for the use of a telecommunications network to have two or more entities interact for the effecting of a transfer of value between the parties. Thus it entails a communications between the parties but also has elements of security and validation and verification. Namely in a transaction processing system we want to certain that only the parties involved are aware of the transaction, that we know that the parties are who they are and that the “value” transferred between the parties is what is truly expected by all parties. Transaction processing is a part of the multimedia world but in many ways it transcends this world as well. In the late 80s our focus was trying to create an environment for distributed transaction processing and allowing the telecommunications network to be an economically viable element in this potentially new market.

1.2.5 Broadband Communications Systems

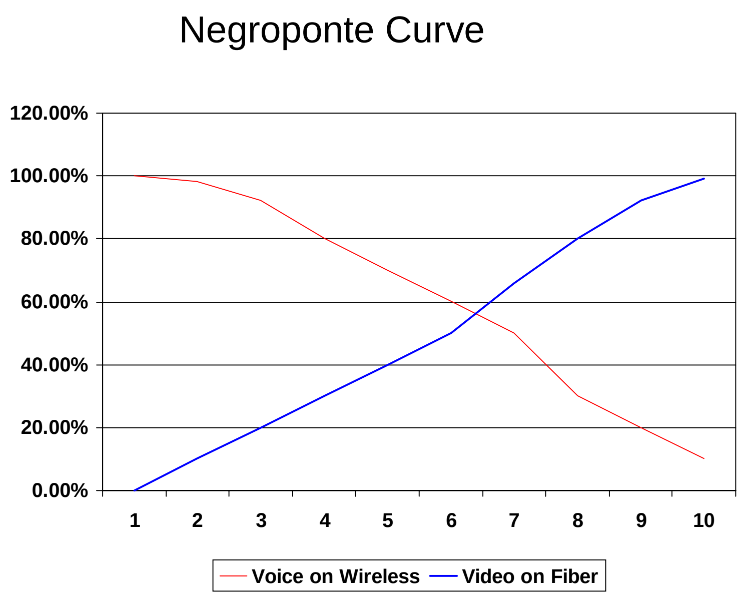
Broadband in the late 80s was a challenge, even its definition was somewhat up in the air. In this period we considered 45 Mbps as the beginning of a broadband application. T1 or 1.544 Mbps was much too slow. It is ironic that in 2005 the FCC defines broadband as anything in excess of 300 Kbps! In 1986 we defined broadband as anything over 45 Mbps.

The challenge of broadband twenty years ago was one of interconnection and standards, thus evolved SONET, of using multiple bandwidths on a single fiber, this DWDM and its derivatives. The growth was in backbone networks. In addition one of our research projects with MIT was the deployment of TCP/IP on a broadband network, at first it was not viewed as practical but soon evolved as the common network fabric.

The main issue was the deployment of technology, not network architecture issues nor even performance issues. Admittedly many theoretical analyses of large scale fiber networks were performed but the problems were not of the real world nor were they of any lasting import, academic exercises only.

The next expansion of fiber networks was to the end user directly. Here architectural issues do abound. In the period from 1988 till 2005 the options ranged from hybrid fiber coax, to ATM Passive Optical systems to active multi giga bit layer 3 systems. The question as to which is best is left unanswered. In addition one also asks how does one integrate wireless into this mix.

In the late 80s Prof Negroponte introduced the concept of a transition from wired to wireless and from wireless to wired. Namely voice would transition from twisted pair to wireless and video from wireless broadcast to fiber.



The question left both unposed and unanswered at the time is what does IP do to all of this, namely if everything is just a packet then what difference does one transmission media make over another. Furthermore, is there any optimal design or even further what design rules should there be?

1.2.6 Advanced Wireless Systems

Wireless was just breaking out into the mobile, cellular, world as we now know it. In 1984 I had developed the Motorola data radio network business which became ARDIS. This was a 1-2 Mbps packet radio network for the use of field service and distributor companies. The challenge was to use limited radio spectrum but to provide a data only service. I had worked closely with customers for the service and attempted to understand what it was that would make this a viable business. It was clear that wireless data was not the driver, customers had to have applications which demanded wireless connectivity. The ARDIS system was the first packetized data radio system deployed nationally. It service commercial users and it was really limited to that user base because Motorola wanted to maximize revenue from high end terminals. We attempted to convince management of the production side to deliver PC interfaces for the lap top PCs which were just beginning to evolve. However, there was not believed to be a demand. Indeed at that time there was no world wide web and no Internet as we know it. There was limited email and some access to private networks or bulletin boards but the driver was truly missing. In the twenty years since, the drivers now are p[resent but still in limited form, for now again the infrastructure is limiting the content side. We shall discuss that in a bit.

In 1988 Irwin Jacobs approached us at NYNEX to consider use of CDMA. This he felt would be a better approach to wireless, aka cellular, than what was being proposed, namely TDMA. We at NYNEX along with Pac Bell and Ameritech joined together and invested $15 million in what was Qualcomm. Jacobs vision was brilliant, because hidden at the upper levels was an all IP based approach, although he did not promulgate that at the time.

1.3 Controlling Issues

The focus on new technologies cannot be accomplished in a vacuum. Technology may create economic and political change but all too often in lives in a world where the economic, political, and social forces are boundary conditions which delimit its progress. The typical case of technology development which goes nowhere is the development of something too early. Too early for a viable market, too early for the regulatory environment to accept it, too early for people to recognize its overwhelming social value. Let me spend a few paragraphs on each of the potentially controlling elements and then we will place each of our suggested research areas in the context of what their separate implications may be.

1.3.1 Technical

The technical issue is what is the technical problem. What has to be solved, what are the elements of the solution, and which of the elements are the most critical. Technical elements also address questions of when is enough, namely bounds on performance, capacity, etc.

1.3.2 Economic

The technical questions are always in the forefront for the discussion but the economic questions are just as important. What will it cost and how much will it save. The classic example was the introduction of CDMA. Technically the work of Jacobs and the others at Qualcomm was brilliant, it skirted around the regulatory issues and the end users wanted more bandwidth. Yet the question was did it save a great deal. The answer was quite simple, it was the classic back of the envelope discussion of effective capital deployment. Jacobs and others could readily demonstrate a significant costs savings as well as a significant service level improvement using the new CDMA system.

1.3.3 Regulatory

Regulatory control is one of the greatest obstacles to technology development. Until 1984 in the United States the FCC mandated that only AT&T was able to connect something to the telephone network. This made it impossible for any amount of creativity. All development was in the control of Bell Labs and much of this work was consistent with maintaining monopoly control. This is one of the reasons why digital switching took so long. In this pre-84 period the profit to AT&T was based upon a return on assets, thus the more equipment the greater the product, the more costly the equipment the greater the profit. What we now see in the computer world of PCs, reducing costs and prices, would never have occurred in this old regime.

1.3.4 Impact

Finally, one must always try to assess the impact of the new technology on the end user, will the dog eat the dog food. The question is one of use and acceptance. The difference between engineering and science it the difference between peer acceptance, as a scientist looks for, and user acceptance, which is the engineers goal. Will someone deploy this technological innovation, and if so why. It is the better faster cheaper argument which one sees again and again.

2 THE PROBLEMS

As we have discussed, the issues were now less fundamental issues of ultimate capacity and how to implement elements to achieve that capacity, but ones of interconnectivity and utilization. Namely we now are asking how do we make all of this work together. In many ways there are two seminal steps along this path; first was the Kahn and Cerf 1974 paper on TCP/IP[6], and the second was the 1988 Berners Lee construct of the web. Both were simple, as seen after the fact, both were implemented quickly, and both addressed problems whose solutions were immediately useful to the community at large. Both also are on the path of development which we further outline in the problems posed. Both also focused on deploying a workable solution in a short period of time to demonstrate how the problem could be solved. Both also led to the

[6] What should be understood about the Kahn and Cerf contribution was that at the same time IBM had developed SNA, the architecture for communicating in a computer environment. This was a multi layer approach. Simultaneously ISO, the international standards organization did the same with the now classic seven layer approach. Kahn needed a simple and real time solution to a set of problems he was addressing at ARPA. Thus TCP/IP. Simple, without detailed architectures and implementable. Kahn had the network to try this on. He did and the rest is history.

creation of a community of followers in a short period leading to adoption, the IETF for example in the Internet world.

2.1 How can many users efficiently access wireless spectrum in a real time manner without having a centralized spectrum regulatory body?

The use of spectrum is controlled by the FCC or other regulatory bodies. The policies for the allocation of spectrum follow concepts of 12th century property rights. That is if you are given rights to use spectrum then you have those rights for say period of time and the spectrum is your property, and no one else can have access to it. If there is an intruder the enforcer is the FCC. Owning spectrum was a barrier to entry to any new competitor.

However, most spectrum lays fallow. Even cellular spectrum, in many parts of the United States is unused most of the time. For years this was not as problem since there was neither the demand nor the technology to utilize this unused spectrum. Now, however, technology can allow us to use it, in a non-interfering manner, and the demand for spectrum is growing in a near unbounded manner.

In 1993 there was an article by a popular writer, George Gilder, who before the dot.com bubble burst, was viewed as the seer of all technology. Gilder looked at spectrum and said that it should be shared.[7] He used the paradigm of the Qualcomm CDMA technology to argue for the existence of technology to do it. In response to Gilder, I and Prof Medard wrote a policy/technology paper in 1994 which addressed the Gilder conjectures.

Let me return to what we said in 1994 about Gilder. Gilder's conjectures were as follows ad our comments at that time followed as well:

"(1) Many Users can occupy the same spectrum at one time.[8] ***There exists a well defined set of protocols that allow this and prevent collisions.***[9] ***There further exists a set of workable multiple access/interface technologies that can be interchangeably used.***[10]

Gilder assumes that there is a well developed technology base that can be operationally available and that permits multiple systems to operate simultaneously and that the industry as a whole has agreed to how best to handle the interference problem."

Gilder was at that time looking to CDMA as the enabler. He failed to understand that CDMA had limitations and that he was stretching a bit too much. However, looking back he had a point, one which said that if we could develop a set of protocols then many users in a relatively free and open environment can share spectrum. Thus Gilder may eventually be correct if we were to interpret his prognostications on future inventions.

"(2) Frequency and modulation/multiple access schemes are utterly unnecessary.[11]

Gilder assumes that worrying about the technical details such as modulation and multiple access is a secondary factor, at best."

---

[7] Gilder, G., "Auctioning the Airwaves", Forbes ASAP, April, 1994, pp. 99-112.

[8] Gilder, p. 100.

[9] Gilder, p. 112.

[10] Gilder p. 112.

[11] Gilder, p. 104.

This is an interesting conjecture. Again we have over the past fifty years been looking at the physical and what we now call the MAC layers as the place to do battle with access to spectrum. Perhaps we are looking at the wrong place. Perhaps the battle should be at a higher layer or at some new place.

"(3) Networks can be made open and all of the processing done in software.[12]

Gilder assumes that hardware is de minimis in terms of its interaction with the operations and that all changes and operational issues are handled in software."

(4) Broadband Front Ends replace cell sites in functionality at lower costs.[13]

This conjecture is based upon the Steinbrecher hypothesis, namely that some simple device can replace all of the features and functions of a cell site, such as network management, billing, provisioning, and many other such functions.

(5) It is possible to manufacture spectrum at will.[14] ***Spectrum is abundant.***[15]

This conjecture assumes or posits that spectrum can be "created" de novo from a combination of what is available and the technological "productivity" gains.

(6) Spectrum can be used any way one wants as long as one does not interfere.[16] ***New technology makes hash of the need to auction off exclusive spectrum, spectrum assignment is a technological absurdity.***[17]

The last conjecture is the one that says that given the above five conjectures, spectrum can be used in an almost arbitrary and capricious fashion, allowing the assumed technology to handle the conflicts, and not having to have the FCC handle the conflicts via a spectrum allocation process. The last Gilder conjecture states that technology obviates the needs for spectrum allocation of any form.

We can now view the problem in the following simple graphic.

---

[12]Gilder, p. 104.

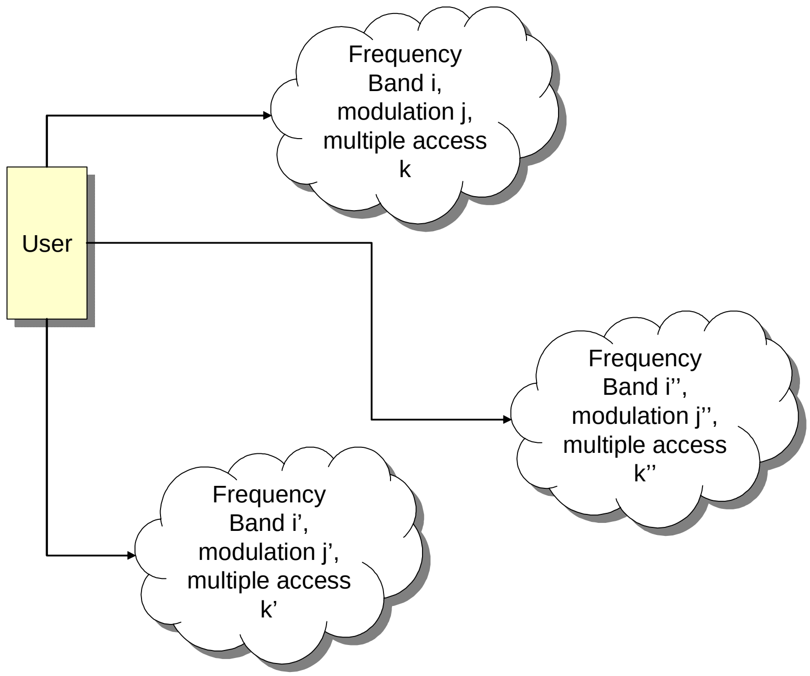
[13]Gilder, p. 110.

[14]Forbes, p. 27.

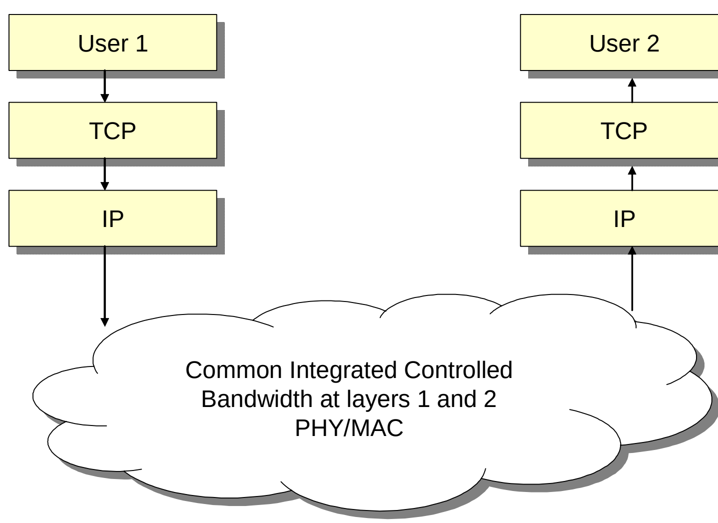
[15]Gilder, p. 100.

[16]Gilder, p. 111.

[17]Forbes, p. 27.



Specifically there are several available bandwidth spectra available but using different modulation and multiple access schemes. The details on all of these systems is known or can be readily known to the user. The user now wants to create an amalgam of these disparate elements and provide to itself what would appear as a single broadband channel. This if the user employs TCP/IP, what is below IP looks as if it were a single integrated broadband channel.



How then do we make this work? This is the technical challenge. Several technical questions arise:

How do we look into the multiple spectra and determine what the PHY/MAC or similar protocols are? How do we determine how to talk with the cloud? This is a physical and logical question. We need some form of broadband radio which can then interpret with a minimal data set what it needs to talk and then to do so.

How do we know who has capacity and whether it is useful for our application?

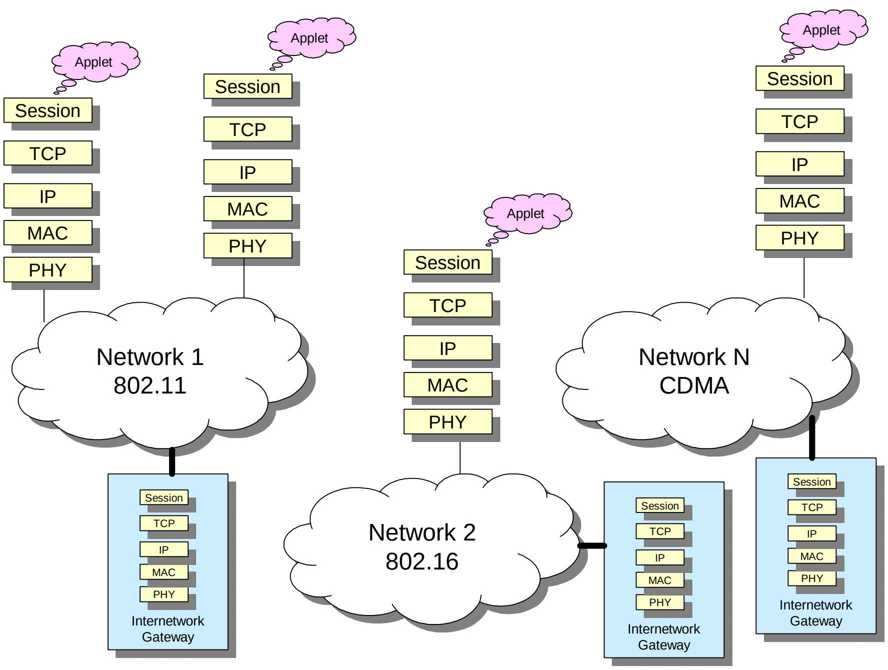
How do we negotiate with the other users and owners of spectrum in a real time fashion creating what might best be called a micro transaction? What compensation do we provide and who and how is all of this kept track of?

What is the performance gain achieved by this process. What are the roadblocks in the delivery of the ultimate performance levels. What are the optimal processes for the execution of this approach?

What functions are performed by the elements of an architecture and how well do we have to perform them. A suggestion is the following:

1. Assess: Determine who is out there and what capacity they have
2. Demand: Determine what capacity is needed now
3. Authenticate: Check entities in domain and determine who is trusted
4. Bid: Enter into auction with entities in domain for capacity
5. Assign: Assign capacity to source entity to meet demand and resulting from auction
6. Manage: Coordinate multiple simultaneous demands between entities and self entity
7. Pay: Conclude payment for capacity
8. Clear: Clear resources when no longer needed

The next question is where do we perform these functions? Clearly they are not PHY/MAC functions since this would require massive changes to a plethora of PHY/MAC implementations and frankly they would not function well there. They are not IP nor are they TCP related functions. Then where. We suggest layer 5, the session layer. This can be shown below. Specifically they become shareable applets which can be placed at this layer and sent back and forth amongst users employing the TCP/IP fabric.



The above architecture contains the following:

1. Variable PHY/MAC layers common over multiple subnetworks.

2. Common TCP/IP across all networks.

3. Applets on all users desiring interconnectivity. The applets can be transferred between any set of users who want to join the shared internetworking facility.

4. Gateways which permit disparate internetwork connectivity. The gateways me be a distributed set of shared facilities allowing PHY/MAC intercommunications and thus permit full interconnectivity. These can be broadband interfaces or may be effected as we discuss in the appendices.

There are now several key questions which can be posed:

1. What is the capacity in terms of say data carrying which a single network can effect with n users distributed in various locations?[18]

2. What is the capacity of an interconnection of disparate networks?

3. What schemes can be developed to maximize throughput?

4. What changes can be made to the architecture to optimize internetwork flow?

5. What optimizations can be made on each of the functional elements, what are their limitations?

2.2 How can one create a bandwidth on demand market so that excess bandwidth, wired and/or wireless, can be interconnected by any user(s)?

Bandwidth is typically provided on a determined amount basis. The "pipe" providing the backbone or local access is engineered to allow a certain amount of capacity. Total capacity is backbone limited and end user capacity is last mile, foot, or inch limited. If we can provide a Gbps or Tbps backbone, then how do we extend that flexibility to the end user? If some users have multiple connections, can we integrate them, can they be used in a manner which allows a perception of integrated bandwidth, or must we allocate on path by path.

2.3 How can one have access to a reasonable amount bandwidth anywhere in the world at a reasonable price?

Rural and third world access to broadband is something that is quite limited. Fiber to users is not reasonable and wireless using terrestrial coverage is also potentially quite expensive. This has always been the opportunity for satellites. The low earth orbit satellites ("LEO"s) were often thought of being the solution. However they never truly materialized for a variety of reasons. Again the question is one of using shared wireless as we have discussed above. Possible in conjunction with some form of a satellite solution. Are there others?

2.4 How does a fully multimedia communications system operate so that any collection of users can communicate in a fully integrated and interactive manner (e.g. using as many senses as possible and in a conversational manner)?

Multimedia communications is a bit messy. It is not sending bits back and forth, there are no zeros and ones, it involves humans communicating with each other in groups. It is not just the Internet with a video link and

[18] See the paper by Gupta and Kumar, The Capacity of Wireless Networks, Univ Illinois.

audio. It is a great deal more, it is conversationality and community and using whatever senses are available. Multimedia communications is creating an experience in displaced communications.

Multimedia communications is characterized by the following factors:

1. Multi Sensory: It uses several of the human senses in transferring, processing, and creating information.

2. Multi-User: It interconnects several users of the information into a conversational mode and allows a dialog based on a fully interconnected set of media.

3. Displaceable: It allows for the establishment of communications and information transfer that is displaced in both space and time from the source.

4. Interactive: It permits a real time interaction between any of the users of the medium, whether the users be human or databases or applications software.

The multimedia environment is one that is user centered and is designed to meet the users needs in interfacing with complex images and in conveying information from one location to another. Multimedia is not just a description of how the data is stored, it is, more importantly, the description of a philosophy of human interaction with complex data elements in a multi sensory fashion.

We must ask the question of what does the human want to do with multimedia that the human does not either want of is able to do with the classic single media information sources. There are several processes that are necessary;

1. Define: The user desires first to define a multimedia object. This object is in sharp contrast to a normal data object that is typically a structured and bounded alphanumeric data element, convertible into a digital representation. A multimedia object is the concatenation of video, voice, text, and other sensory representations of the event at hand. A multimedia event is the analog of a data object. The data object is the representation of a definable and measurable term used in common communication, such as the word, NAME. In contrast the multimedia event is a collection of multimedia presentations that have a temporal and spatial extent to them. That is a multimedia even is a set of voice segments, a set of video elements, and a set of text frames. To define a multimedia event means to concatenate in a rational form the set of disparate multimedia elements into a connected event. The process of connection is complex but it goes to the heart of human communications and understanding.

2. Query: The query in a data object case is a way to do one of two simple tasks. A data object may be either selected or enumerated. The selection process is based upon the ability to take a data object and recognize that it has a unique representation in some stable denumerable set. For example we can use the alphabet or a binary representation for any data element. Once we select a data object we can then enumerate all of the objects that meet a certain criteria, by again matching and now counting. Thus we can answer the question of how many patients over forty have high blood pressure. We first use the select process on high blood pressure and then the enumeration on patients. In a multimedia environment, we are now posing much more complex queries.

3. Store: This means that we must store complex multimedia objects, composed of video, voice, image, pointer movement, text etc. that may reside on different storage devices at different locations. We must be able to retrieve them in the same order and timing that they were stored and do so in a minimum time.

4. Process: We must be able to process multimedia objects, to alter, enhance, combine them. We must perform the processing in a fully distributed fashion, using the resources from multiple processors.

5. Display: The display of the multimedia objects includes not only the display of the image or video but the "display" of the voice and other sensory elements of the multimedia object.

6. Communicate: Communications means the development of a conversational mode. Conversation is key to communicating in a multimedia environment. Thus, we must not reproduce a communications environment that is attuned for the computer but one that is matched to the human user. The essence is the ability to effectively share the multimedia objects in a dialog fashion, interactive and interpretive.

These multimedia processes are to be done in a fashion that is transparent to the user. They must also be done in a fashion that is resonant with the way the users currently performs the tasks.

There are several dimensions that can be used to characterize the extent of the multimedia environment. These dimensions are;

1. Time and Duration: This dimension shows the amount of simultaneity that the medium allows both for a single user as well as for a collection of users. Further a dimension of durability to the environment is essential as the complexity of a multimedia object requires that time pass until it has its full representation. Thus unlike a mono-media object that can be represented to a single user in a fixed period, the interlining of media and users requires a sustainability of the environment.

2. Communication and Conversationality: This characteristic is one of allowing for a multimedia multi-user environment that permits a full sharing of the environment a an equal basis amongst all of the users. It further allows the users to interact with any other user while at the same time allowing this interaction along any one of the multimedia dimensions.

3. Interactivity and Responsiveness: This dimension relates to the ability of the environment to allow one or several users to utilize all elements of the medium and at the same time to pose questions that are robust in a multimedia sense and to obtain adequate answers.

4. Presentation and Interaction: The interactiveness of the environment is a key element of understanding the

5. Non-Linearity and Hyper-Dimensionality: This dimension of characterization allows for the movement amongst the object in an unbounded fashion. It allows for movement in space, all dimensions, and time, as well as in point of reference. The spatial movement entails the ability to view at different magnifications that is common amongst hypermedia environments. The ability to view at different points of reference allows one users to accept the reference frame of another to view the object.

6. Sense-Complexity and Representation: This dimension allows for the combining of multiple sensory elements into the multimedia objects as well as the presentation of those elements either as direct manifestations or as appropriate analogs.

In the context of this paper there will be three elements that define multimedia and multimedia communications; the message, the medium and the messenger. The ultimate result is the impact of the information created on the environment. Without the result, there has been no transfer of information. The objective of this paper is to show the relationships between all three of these elements and the blend with them a set of philosophical underpinnings that will allow them to be used in analyzing the development of multimedia communications.

We first define the three elements of message, messenger and medium.

Medium: The medium is the collections of all physical elements outside of the mind of the creator of the message that facilitates the externalization of the message. Paper, a video screen, a hypermedia environment, a set of signaling flags, a stone tablet are all the elements of the medium.

Message: The message is the "idea" to be transferred from one individual to another. It is the information content to be transferred and thus to acted upon. It is an actionable element of internalized conceptualization.

Messenger: The messenger generally is thought of as an individual. In our context, the messenger is the collection of any and all entities that move the message from one point to another. Recall that the movement may be in space and time. Recall also that the channel used in the movement may be a "noisy" channel that can introduce errors

In the past twenty years we have looked at many dimensions of multimedia communications. The web is a simple example of this approach, elegant but simplistic. There have been many who have studied gimmicks and games which purport to be attempts to address what multimedia communications is. We know it when we see it, it works, its is the conversationality of daily human intercourse. It is not a web camera, video teleconferencing, or other attempts to place several sense to work at the same time.

Clearly true broadband should be the facilitator for this effort of multimedia communications. But a true broadband is just a part of the solution. We need to develop that concept of displaced conversationality.

2.5 How does one create the most effective broadband local infrastructure and what are the measures of the effectiveness?

Are local networks nothing more than mini Internet backbones or are they mini-Internets or are they nothing more that last mile connectors to the real Internet. What is the local network and is it one thing or many, is it one technology or many, is it open or is it proprietary?

There are a great deal of fundamental questions regarding local broadband networks. The first is generally the question of how fast. The development of the Internet was a gradual process with the initial driving factor being connecting university and government computer systems. The first set of computers were well defined and the evolution was controlled by a small group of collegial designers. The local broadband infrastructure we now envision is being developed in a highly contentious environment; ILECs fighting with cable companies and again destroying any new entrant. Cable companies have no interest in asking what the best architecture is, it is a CATV system slowly changed to maintain monopoly control. The ILECs no longer have any R&D groups and thus rely on whatever the vendors pitch at them. Thus there is no independent thinking about what the “best” network architecture should be. There is no Bob Kahn and Vint Cerf, there is no group at MIT, USC, Illinois, talking through the issues based upon intelligent give and take. Add to this is the fact that the FCC, the main regulatory body is generally clueless about what this means, it is managed by lawyers and there is no venue fr open technical or even business discussions.

Before continuing with the technology determination, we should first stop and look at where broadband and broadband-like networks are progressing. The following is a brief summary which sets the framework for evaluating municipal networks:

Phase 1: The Internet started with local communities needing a backbone network to provide interconnection between the communities. The IP protocol allowed the routing of the packets in the most efficient fashion over the backbone network.

Phase 2: DARPA spins off the commercial Internet backbone and multiple large scale players take positions as what we now call Tier 1 Internet carriers; ATT, MCI, now Level 3 and others. They need to interconnect and the evolution of the MAEs, East and West (Virginia and California), occurred as well as the introduction of the regional NAPs with the support of Sprint. The Internet is becoming broadband BUT from the top down. This interconnection is called peering, since the Tier 1 ISPs agree to interconnect at no cost to each other, but not to anyone else.

Phase 3: ISPs, local and national, evolved allowing dial up connections. The ISPs need to interconnect to the backbone and they do that via Transit agreements agreeing to pay the Tier 1 carriers. In this phase the local end users has to connect via phone, local ISP, transit connections, then peering, and eventually as it goes back down the chain to the information they seek. Typically this may entail 50-100 packet hops. This further entails an individual seeking information for a collection of larger purveyors.

Phase 4: Ersatz broadband begins; enter cable and telcos. This is the phase of DSL and cable modems. The only material change from Phase 3 is the slightly higher speed of local access. The same backbone bottlenecks exist. The local carriers such as the cable and telco companies can increase the speed to multiple megabits per second, but one must always remember that there is that bottleneck to the Internet backbone. Thus, for example, if one has a transit agreement with say a Tier 2 or 3 carrier in central New Hampshire, one may as a carrier pay that transit carrier a fee of $200 to $400 per Mbps per month. That is we average the traffic and come up with an average rate and this is the basis for billing. Now take this example a step further. Assume we have 20 Mbps HDTV and we watch this 6 hour a day or 25% of the time. The transit costs alone for that one TV set would be; $400 X 20 X 25% or $2,000 per month per channel! It is this economic factor which drives architecture change as well as true technical improvements.

Phase 5: This is an interesting Phase an generally not noted by many. It is the development of the NIXs. The NIX is the national Internet exchange. Take for example the Czech Republic. There are many small ISPs and a few larger ones, such as that of Czech Telecom. In 1995, almost all Czech Internet traffic went from a user to some web site in the US, or possible Western Europe. By 2002, it was 50% to the US and Western Europe but the other was all within the Czech Republic. This super “municipal” network saw that they need not pay for transit fees in Frankfurt to connect all Czech users. They instead agreed to create a NIX, National Internet Exchange, to interconnect all Czech networks who talk with Czech Networks. Two observations come from this; first, when one allows localism and openness of the network, people have a tendency to communicate within their regional group, second, costs can be avoided and performance increased by local interconnection. This step is the first step towards localism. All members of the NIX agree to peer at no cost. Thus the high and almost extortionary fees of the Tier 1 carriers were eliminated. This did not happen in the US!

Phase 6: Municipal networks in the US and other countries have deployed what are best described as wide area LANs, local area networks, and these networks allow any user to communicate with any other user of the network and the interconnections are performed in the network using IP. This is NOT the approach of the proprietary networks such as cable and DSL. In the proprietary cases they still use Tier 1 interconnection architecture. They go to a Tier 1 node and pay the fee and go back again. Municipal networks connect “within” the network! This means lower costs, zero marginal costs in fact, and improved performance, 2-5 hops, not 50-100.

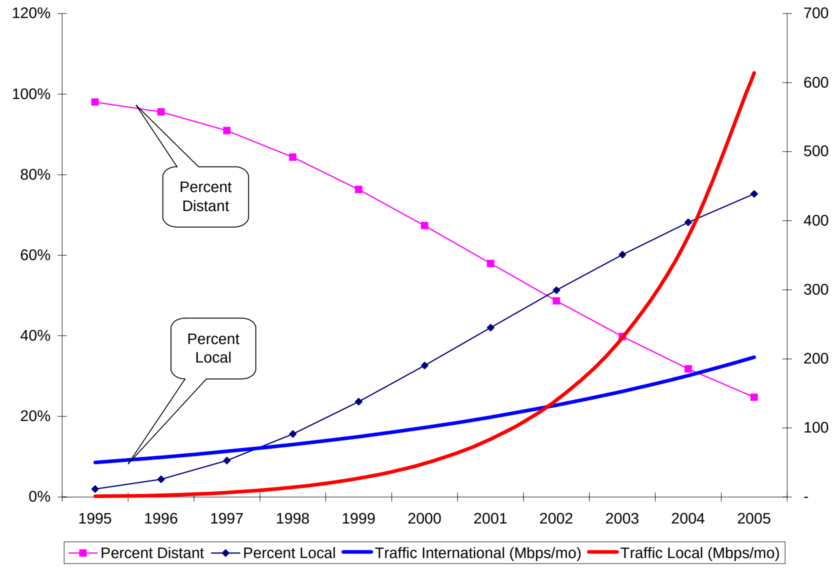
Phase 7: Regional Municipal Clusters: This we believe is the next step. There are some initial efforts in this stage already. One may consider this being what a smaller country may do, such as the Netherlands. In the Netherlands, the cities may build their own Municipal Networks, then agree to interconnect, and then the entire country is one Clustered set of Municipal Networks. The economic power is now within the network not external to it. This is the contrast of externalities to internalities.

The following Figure is characteristic of NIX countries and we believe is characteristics of Phase 7 Municipal Clusters. It shows several characteristics:

First, total distant traffic continues to grow at a steady rate. This is in many ways consistent with the theory of externalities, namely one has more people to connect with.

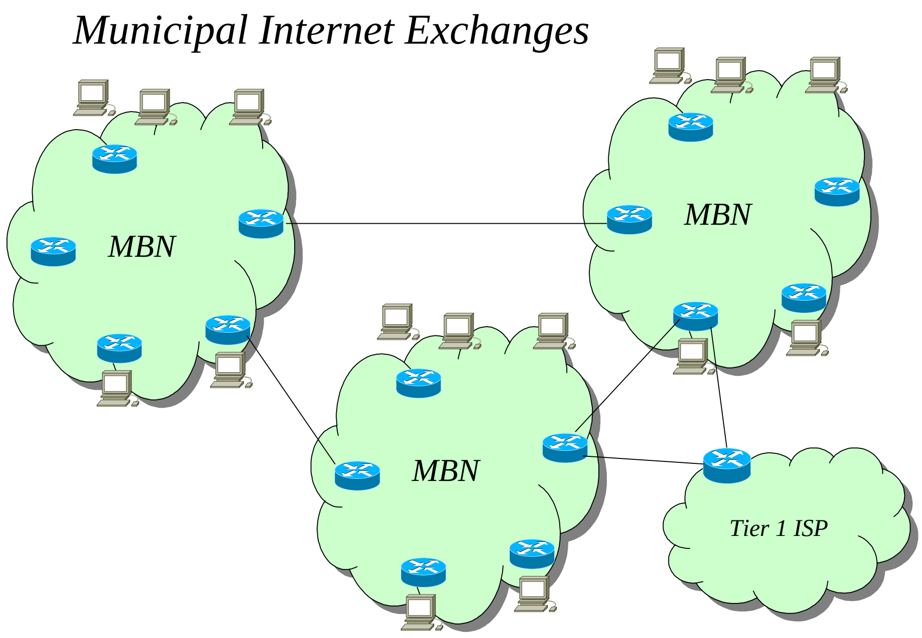
Second, local traffic grows at a much more rapid rate, this is the most significant factor in total growth. This is consistent with the corollary of externalities, namely internalities.

Third, there reaches a point where percent of internal traffic exceeds percent external traffic. This stage is the stage of true economic development. It is this point of inflection that we anticipate occurring with the deployment of truly open local networks. The existence proof is the projection of the deployment of NIXs.



Thus given this evolving structure, can we go back to the drawing boards and ask the same questions that Kahn et al were asking almost forty years ago, how should we engineer this network and what are the overall performance and design factors. We also drop into this mix the combined elements of multimedia as well as integrating wireless.

The first step in a new architecture is the development of the local Internet Exchange or Municipal Broadband Network ("MBN") Internet Exchange, MIX. The MIX is the next evolutionary step in the evolution of this network. It is the peering point for local participants in the network. In many ways it looks like a NIX except the countries become municipalities and in each municipality there is a connection point. Thus the NIX approach is the forerunner of a MIX. The following Figure depicts a collection of MIXs and their interconnection to the Internet backbone via a Tier 1 ISP. Multiple other connections are also possible.



2.6 How does one interconnect disparate networks (wire and wireless) in a real time manner and with de minimis costs of interconnection?

Enron is a know economic and business fiasco from the dot.com period. Hidden in the Enron story was an attempt to create a real time market for bandwidth. Energy, electric power, is fungible but is a commodity that can be traded back and forth. It is like corn, but unlike corn cannot be stored, it has to be used when generated. Like corn it is the same everywhere, power companies can connect to the "grid" and down load their desired amounts. In any time sensitive commodity market there is a secondary, call it derivative, market which deals with futures and future contracts. I can enter into an agreement for a price to be assured that I will have so much power at such and such a time at such and such a price. It costs me but I know it is there. Thus the futures help to stabilize the market by smoothing out fluctuations by compensating people for standby power, albeit it may never be used.

Bandwidth is the same. If I have 150 Mbps bandwidth I have used it or it goes away.

2.7 How does one provide a communications environment which establishes a level of privacy and even anonymity while allowing protection from abuse and threat?

Privacy, unfortunately, is not a Constitutional right delineated in the US Constitution or even the Bill of Rights. The penumbra, shadow, of privacy was constructed by the Supreme Court for reproductive rights but there the coverage seems to end. Furthermore the right to be left alone, as stated by Justice Brandeis in the famous Weaver and Brandeis paper, is just a wish, and anonymity does not exist. In today's world, created out of the slaughter of innocents, decimates any sense of privacy and destroys any attempt at anonymity.

In the mid 70s there was a burst of work on encryption with the evolution of the public key systems. The response to this was the Governments attempt to close the lid on this research and stop it totally. In 1978 a conference at Cornell was held to discuss the new encryption systems and approaches. I attended and for reasons I will leave unsaid, I was personally aware of the attendance of a Russian agent, the head of the Czechoslovakian intelligence service, the woman in charge of Cuban counter intelligence and the Chinese head of intelligence were also present. They were not there to enjoy Ithaca, New York. It was like a Mad comic book spy versus spy meeting and in addition there were US agents from every branch including I believe the Department of Agriculture, after all Cornell was a land grant college formed for the purpose of advancing agriculture at the end of the Civil War. Now matter what attempts were made the genie was let out of the bag.

3 CONCLUSIONS

The thoughts in this paper are both backward looking and forward looking. They in some sense reflect the changes we are seeing in the world of telecommunications, changes that relate to networking and services, that reflect challenges of regulation and opportunities of markets. Fifty years ago there was hardly an question of who would buy what. The regulated environment of the PTTs and AT&T made any change by an independent third party but impractical and illegal. Universities could perform research and hopefully that may influence places like Bell Labs. However, if Bell Labs did not respond then it remained an academic exercise. The change in my view was not the breakup of AT&T in 1984, but was ARPA and people like Bob Kahn. People who were to build networking as we see it today. Their challenge was technical not entrepreneurial. Today the challenge is broad and the entrepreneur has a significant role to play.

The following table summarizes all the questions, technical issues, regulatory and market issues. The Table presents only some questions of a technical nature and is hardly inclusive. The regulatory and market issues are also only approached in a preliminary manner.

Question	*Technology* *Some Questions of Interest*	*Regulatory* *Barriers to Implementation*	*Market* *Incentives to Implementation*
How can many users efficiently access wireless spectrum in a real time manner without having a centralized spectrum regulatory body?	What is capacity of a multi-element wireless network? What are the optimization criteria for a multi-element network? How can the session layer approach be implemented/optimized? If nano-transactions are effective, what are the optimal nano-transaction mechanisms? How do you implement a multi-element gateway for optimal performance? How does one use a software defined radio to effect this approach?	FCC issues exclusive licenses, how can the regulatory body provide incentive for sharing?	Incumbents want to protect their assets, can this be monetized to be in their best interest?
How can one create a bandwidth on demand market so that excess bandwidth, wired and/or wireless, can be interconnected by any user(s)?	How does one create a universal networking gateway between all networks and users? Is the layered architecture a basis for integrability or is some other paradigm more appropriate? If such a network is implemented, how does one manage such a network? How does one implement the last mile?	International regulatory bodies have control of carriers. Trans-border issues are also a concern.	Incumbents want to control end to end but this is not totally effected in current markets. Customers want this very much and at a premium.
How can one have access to a reasonable amount bandwidth anywhere in the world at a reasonable price?	What is the maximum coverage and capacity for a wireless network? How can wireless grids be combines using satellites to provide large area coverage. How does one optimize the TCP/IP capabilities in a fully integrated grid using satellite elements?	Regulatory bodies specify who can be a carrier and demand licenses. There is a limit to creating a market. Last mile extensions are typically the most costly and of greatest delay because of local ordinances.	Low earth orbit satellites were shown to be uneconomical compared to their expectations; the time constants of satellites exceed that of the network technologies. How can satellites be part of an integrated global network?

Question	*Technology* *Some Questions of Interest*	*Regulatory* *Barriers to Implementation*	*Market* *Incentives to Implementation*
How does a fully multimedia communications system operate so that any collection of users can communicate in a fully integrated and interactive manner (e.g. using as many senses as possible and in a conversational manner)?	How does one model a multimedia communications network (e.g. Petri nets?) Can a multimedia network be built on top of the web or does it require a full new fabric? Is the session layer the correct approach to establishing multimedia sessioning, if so what are its performance factors and how can it be optimized? If not, what is?	This may be a standards issue rather than a regulatory issue.	This may have a sweeping change on distribution channels in many product and services areas and as such may be viewed as a threat to entrenched players.
How does one create the most effective broadband local infrastructure and what are the measures of the effectiveness?	Is there a fundamental difference between a local broadband IP network and an international broadband backbone IP network, if so why and what are the optimality criteria?	Franchise control delimits anything done locally. There are 30K or more franchising entities and if the product is in any way video enabled then it requires local contractual approval.	This could be threat to incumbents and local towns themselves may want to play in this area. It also competes with Internet backbone players by creating a new Internet outside of the existing Internet.

Question	*Technology* *Some Questions of Interest*	*Regulatory* *Barriers to Implementation*	*Market* *Incentives to Implementation*
How does one interconnect disparate networks (wire and wireless) in a real time manner and with de minimis costs of interconnection?	What are the optimal uses of optical, broadband and conventional signalling processors in a multi element wireless fabric? What are the optimal uses in a multi element wireless network for optical signal processing elements? What are the tradeoffs between broadband and optical processing, how can they be optimally integrated? How does one create a real time software configurable gateway between disparate fiber networks? Can this be done on an on demand basis, on an auction basis, and what are the optimality criteria?	Interconnection and access is controlled by the FCC and all regulatory bodies. The incumbent defines the entry costs. The FCC is generally clueless about opening this interface.	Incumbents have strong vested interests.
How does one provide a communications environment which establishes a level of privacy and even anonymity while allowing protection from abuse and threat?	How does one design an anonymous communications network? How does one design an anonymous transaction network (e.g. cash purchases without video cameras to blind salespersons)? As with encryption, how can authentication and anonymity be balanced, are there analogs?	Federal law is a major block to entry since September 11th.	There are demands to protect identity from theft but not clear that this is the way.

4 REFERENCES

The following are a list of some of the authors papers which relate to the topics discussed herein.

1. Processing Voice and Data in Mobile Satellite Communications Systems, International Communications Conference, Chicago, IL, 1977.

2. Multiple Beam Satellite Optimization, IEEE AES-13, September, 1977.

3. Multiple Access Techniques for Low Data Rate Satellite Communications Systems, National Telecommunications Conference, Los Angeles, CA, 1977.

4. Multiple Beam Satellite System Optimization, National Telecommunications Conference, Los Angeles, CA, 1977.

5. Communications Satellites; Looking to the 1980s, IEEE Spectrum, December, 1977.

6. Unattended Earth Stations for Global Data Collection, International Communications Conference, Boston, MA, 1979.

7. Financial Data Networking and Technological Impacts, INTELECOM, 80', Los Angeles, CA, 1980.

8. EFT Networks and Systems, CASHFLOW Magazine, November, 1981.

9. Local Area Wideband Data Communications Networks, IEEE EASCON, Washington, DC, 1981.

10. The Confluence of Policy and Technology in Cable Communications, Communications Symposium, Harvard University, New York, 1982.

11. CATV for Computer Communications Networks, IEEE Computer Conference, Washington, DC, 1982.

12. QUBE: The Medium of Interactive Direct Response, Direct Marketers Compendium, Direct Marketing Association (New York), pp 162-165, 1982.

13. Impacts of Consumer Demands on CATV Local Loop Communications, International Communications Conference, Boston, MA, 1983.

14. Hybrid Cable and Telephone Computer Communications, Computer Conference, Washington, DC, 1983.

15. Cable Based Metro Area Networks, IEEE, JSAC-1, November, 1983.

16. Network Management and Control Systems, IEEE NOMS Conference, 1988.

17. Alternative Networking Architectures; Pricing, Policy, and Competition, Information Infrastructures for the 1990s, John F. Kennedy School of Government, Harvard University, November, 1990.

18. Image Processing in Full Multimedia Communications, Advanced Imaging, pp 28-33, November, 1990.

19. Applications of Multimedia Communications Systems for Health Care Transaction Management, HIMMS Conference, San Francisco, CA, January, 1991.

20. Multimedia Communications Technology in Diagnostic Imaging, *Investigative Radiology*, Vol. 26, No 4, pp 377-381, April, 1991.

21. Multimedia Communications: Architectural Alternatives, SPIE Conference, Boston, MA, September, 1991.

22. Information Architectures and Infrastructures; Value Creation and Transfer, Nineteenth Annual Telecommunications Research Conference, Plenary Address and Paper, Solomon's Island, September, 1991.

23. Communications Networks; A Morphological and Taxonomical Approach, Private Networks and Public Policy Conference, Columbia University, New York, October, 1991.

24. Alternative Networking Architectures, B. Kahin Editor, McGraw-Hill (New York), October, 1991.

25. Multimedia Session Management, IEEE Proceedings on Communications, 1990.

26. Wireless Communications Economics, Advanced Telecommunications Institute Policy Paper, Carnegie Mellon University, February, 1992.

27. Broadband Telecom Rate Projections, AMIA Conference, Spring, 1992 (Portland, OR).

28. Communications Network Morphological and Taxonomical Policy Implications, Telecommunications Policy Research Conference, Solomon's Island, MD, September, 1992.

29. Multimedia Communications in Medicine, IEEE JSAC, November, 1992.

30. Architectures et Structures de L'Information, Reseaux, No 56, pp. 119-156, December, 1992, Paris.

31. Economic Structural Analysis of Wireless Communications Systems, Advanced Telecommunications Institute Policy Paper, Carnegie Mellon University, February, 1993.

32. Access to the Local Loop; Options, Evolution and Policy Implications, Kennedy School of Government, Harvard University, Infrastructures in Massachusetts, March, 1993.

33. Wireless Access to the Local Loop, MIT Universal Personal Communications Symposium, March, 1993.

34. Spectrum Allocation Alternatives; Industrial; Policy versus Fiscal Policy, MIT Universal Personal Communications Symposium, March, 1993.

35. Access Policy and the Changing Telecommunications Infrastructures, Telecommunications Policy Research Conference, Solomon's Island, MD, September, 1993.

36. Internet Architectural and Policy Implications, Kennedy School of Government, Harvard University, Public Access to the Internet, May 26, 1993.

37. A Précis on PCS Economics and Access Fees, presented at the NPC SC Seminar on "Wireless Technology and Policy Implications" at MIT Lincoln Laboratory, in Lexington, MA, May 18, 1994

38. Wireless: The Challenge of Using Bandwidth Intelligently, presented at the Symposium on Communications , Optics and Related Topics, held in honor of the 60th birthday of Professor Robert S. Kennedy, at Endicott House of the Massachusetts Institute of Technology, June 4, 1994

39. Wireless Architectural Alternatives: Current Economic Valuations versus Broadband Options, The Gilder Conjectures; Solomon's Island, MD, September, 1994.

40. From High End User to New User: A New Internet Paradigm, McGraw Hill (New York), 1995.

41. Disaggregation of Telecommunications, Presented at Columbia University CITI Conference on The Impact of Cybercommunications on Telecommunications, March 8, 1996.

42. The Economic Viability of Wireless Local Loop, and its Impact on Universal Service, Columbia University CITI seminar on "The Role of Wireless Communications in Delivering Universal Service", October 30, 1996.

43. Internet Voice: Regulatory and Legal Implications, Presented at the VocalTec Seminar on September 9, 1996, New York, NY.

44. Communications Networks; A Morphological and Taxonomical Approach, Private Networks and Public Objectives (Noam, Editor),Elsevier (London), 1996.

45. Competition in the Local Exchange Market: An Economic and Antitrust Perspective, *MIT ITC Working Paper, September, 1977.*

46. The Economic Viability of Wireless Local Loop, and its Impact on Universal Service, *Telecommunications Policy*, Elsevier (London), 1997.

47. Economic Factors on International Internet/Intranet Telecommunications, MIT Research Program on Communications Policy Conference Internet Telephony Interoperability Forum, Bristol, England, June 11, 1997

48. Comparative Deregulation of Far Eastern Telecommunications Markets, Telecommunications Policy Research Conference, Washington, DC, September 28-30, 1997.

49. Telecommunications Infrastructure, Technology, and Policy in Russia, A Plan for the Development of an Information Based Economy, Russian Freedom Channel Report, September, 1997.

50. International IP Telephony, MIT ITC Working Paper, September, 1999.

51. The Internet Protocol (IP) and Global Telecommunications Transformation, Tufts University, Fletcher School, March, 1999.

52. The Application of IP Telephony to Local Exchange Carriers, MIT, Internet Telephony Consortium, March, 1999.

53. IP Telecommunications QoS (Quality of Service), Is Service Quality a Sustainable Metric?, MIT Internet Consortium, Aquila, Italy, June 2000.

54. The Evolution of International Internet Telephony, TPRC, Arlington VA, September 2000.

55. Virtual Global Telcos: International Internet Telephony Architectures, in Internet Telephony, MIT Press (Cambridge), 2001.

56. Internet Telephony Markets and Services, in Internet Telephony, MIT Press (Cambridge), 2001.

57. The Internet Protocol and the Creative Destruction of Revenue, in Creative Destruction, MIT Press (Cambridge), 2001.

58. Peering, Transit, Interconnection: Internet Access In Central Europe, MIT Internet Consortium, January 2002.

59. Privacy in the Internet Environment, MIT ITC Conference, December, 2002.

60. Municipal Broadband Networks: A Revised Paradigm of Ownership, MIT ITC Conference, December, 2002.

61. The Imminent Collapse of the Telecommunications Industry, MIT ITC Working Paper, August, 2002.

62. Privacy in the Internet Environment, MIT ITC Working Paper, November 2002.

63. Current Telecommunications Legal Issues, Litigation v. Legislation: Is the 1996 Act a Beginning or an End?, MIT ITC Working Paper, December 2002.

64. Municipal Broadband Networks, A Local Paradigm, Working Paper for Dutch Government, July, 2004.

65. The Hidden Cost of Broadband, “The Franchise”, Working Paper Telmarc, November, 2004.

66. The Impact of Broadband Options on the Disaggregation of the Media Industry, Telmarc Working Paper, September, 2004.

67. New Wireless Spectrum Policies, Alternatives to Outdate Spectrum Management, to be published, February 2005.

5 APPENDIX 1: SUMMARY OF WIRELESS TECHNOLOGY OPTIONS

In this section we details a broad spectrum of technological challenges which address several of the areas in the body of the paper, specifically; disparate network connectivity, spectrum utilization, broadband connectivity. In each of the areas we have considered each in significant analytical detail and have in some cases pursued actual implementations and execution.

5.1 Key Technologies for Wireless

The following technology areas will be surveyed for new and innovative challenges.

(1) Adaptive Network Management: Adaptive Network Management, ANM, is a system that uses in-situ sensors to monitor the power and signal quality throughout the network. The number of sensors will greatly exceed the number of cell locations. This set of dynamic measurements will then be used in a feedback schemes to adaptive change the characteristics of the cell transmit power and other characteristics to maximize the service quality. Specifically, we have designed a proprietary network management system that uses the in-situ sensors that monitor all key signal elements. These elements are power, frequency, interference, noise, and other significant signal parameters. The system then transmits these signals back to a central processor which then generates an optimal signal to control the cell site transmission characteristics, such as power, frequency and other factors. The overall objective is to optimize the system performance from the users perspective.

(2) Gateway RF Digital Front Ends: The use of wideband RF Front Ends,[19] where we are using highly linear RF front ends, the RF signal may be sampled at the Nyquist rate and at 16 bits per sample or even higher. This then allows a large dynamic range reconstitution of the signal that allows for the processing of the RF in a fully digital mode. A broadband, digital front end will be used to act as a gateway to interface the air interfaces of CDMA, TDMA and other access methods through the same cell and in the same frequency band. This system will permit multiple air interfaces to be gateway into the same network access point thus reducing the need for a single standard, and increasing the ability to provide a national network. The system element allows, through its use of large gain bandwidth product front end and fully digital RF processing, the ability to handles many different and simultaneous multiple access methods, such as TDMA and CDMA. This ability goes to the heart of interoperability and standards.

(3) Co-Located Distributed Switch Access: Unlike other proposed schemes which use redundant MTSO accesses, this trial will focus on Central Office Co-Location methods that reduce capital and operating cost redundancies. The co-location approach, will minimize access line costs and eliminate the need for a MTSO. The adjunct processors at the Central Offices will be interconnected by a high speed bus to allow for adequate control and call hand-off. Co-Location is achieved via the intelligence that is contained in the CDMA cell sites and the adjunct processor distribute communications and processing capabilities. The fundamental existence of this capability was demonstrated by QUALCOMM in their CDMA trial, albeit not in the Co-Location context. The QUALCOMM QTSO was in effect a no Co-Located adjunct.

(4) Adaptive Beam Forming Phased Array Technology: One of the current problems with a wireless systems is the use of broad beam antennas and the inability to provide additional antenna gain on both transmit and receive to the individual portables. With the use of adaptive beam forming antennas, the service to lower power portables may be improved. The approach would include such capabilities. Time dynamic control of these multiple bean antennas will permit higher localized gain on portables, which will in turn allow for lower transmit power and thus longer portable battery life.

(5) Multi Layered Processing: This approach assumes that the sensors and processing equipment for a wireless network in the field are de minimis in form and structure. Namely we have a collection of

---

[19]Gilder, G., Telecosm, The New Rule of Wireless, Forbes, ASAP, March, 1993, pp. 96-111. Gilder details the work between Steinbrecher and QUALCOMM.

transmit/receive array elements, possibly placed in collections in some definite format but not even that is to be necessary. In the example we use herein the sensors are placed on what we would normally see as a cylindrical phased array. Each array element has its own broadband connection to a from a central site, and the signal is nothing more than the capturing of what is received at the carrier band by the array element. It is placed upon the fiber strand and then on a wavelength by wavelength the array sensors are packet on one or several fiber strands. Now with this form of system we can first process the many array element on transmit or receive by using optical processing techniques. We receive all the data at one point and thus all the data can be used. Further we can achieve significant "beam forming" capabilities by having an optical signal, we can correlate, mix, down convert and a variety of other techniques that would be more difficult if we were to do this in a digital format for an RF signal. Then we can do wideband processing, allowing us to facilitate the extraction and segmentation of various PHY layers. Finally separate baseband allows MAC and other layers to be extracted and processed.

The choice of the above areas was based upon the demonstration that there are several keys for success for wireless services and these are dominated by the need to reduce the capital per subscriber and the cost per access minute. These technologies and the related unique efforts will clearly demonstrate that these are achievable goals.

5.2 Goals of Technology Development

This section presents a detailed review of the general work efforts envisioned during the course of this effort. The major goals of this efforts are to demonstrate that there are key technological advances that can be made that will allow wireless to be a viable service and business. The focus of this is on providing total customer satisfaction through a transparent set of technologies. The intention is to avoid ambiguity of expectations on the part of the customer. Namely, a wireless telephone should provide toll grade quality of service or better in any wireless environment, fixed or mobile.

The above *five technology areas* are focused on *nine critical success criteria* identified for the development of wireless. These are:

(1) **Assuring Real Time Interference Elimination**: The issue of ensuring that interference with other carriers that may be located or using the assigned bands is critical. This can be accomplished in a dynamic fashion through the use of the Wideband dynamic front end and the use of digital processing of the signals and the insertion of nulls in the spread spectrum transmit band.

(2) **Extending the Life of Batteries and Reducing Portable Costs**: Estimates have been made on the extension of the size and lifetime requirements of the portable batteries. It can be shown that a 10 dB improvement on transmit and receive link dBs can be achieved with adaptive antenna formation. This directly results in power reduction in the portable. Further reductions of power drain are achieved through use of single ASIC in the portable.

(3) **Improving Customer Satisfaction through improved System Performance**: Balancing the network management and customer service functions to maximize both Customer Satisfaction and System Performance. This will result in lower churn amongst the customer base.

(4) **Minimizing the Capital per Subscriber:** This success of wireless is a quality and cost combination. Cost is driven by three factors; capital per subscriber, cost per access minute and cost to acquire a net new customer. The capital per sub is the driving factor in multiple access design. In this first report, the Experiment has developed an extension that appears to allow for multiple re-rads per cell site, thus permitting capital per subscriber of approximately $100.

(5) **Minimize the Overall Operating Costs**: The overall operating costs are minimized by effective network management.

(6) **Minimizing the Cell Site Life Cycle Costs**: The overall operations costs are dominated by cell site life cycle costs. The overall costs are minimized by combining low capital per cell along with total network management.

(7) **Minimizing the Access Costs to the Local Exchange**: The access costs must be brought under control. Currently Wireless companies pay two to three times what a IEC pays per minute for comparable access and the IEC is charged four to five times what the internal LEC transfer price is. If true competition is sought, disaggregation of LEC infrastructure on an equitable marginally based pricing system must be achieved.

(8) **Assuring Open Network Interfaces, namely Standards**: The recognizes that open interfaces, standards, are key to quality, cost and competition. Thus the will continue to share this information with others as necessary.

(9) **Establishing a Seamless Interoperable National Network**: The issue is to use common resources where appropriate and to empower entrepreneurial skills where necessary.

5.3 System Architecture

The system architecture depicts the overall concept that has been presented by in earlier presentations. The key elements of the system are:

(1) **Sets:** The sets are the key factor in low consumer access costs.

(2) **Propagation Analysis and Optimization:** From prior studies, the results for the Boston tests in and around Logan demonstrated multipath with delays of 1 to 5 nanoseconds and typical amplitude reductions of 10 to 35 dB. Using the 1.25 Mbps transmission rates, the delays of less than 1 microsecond will cause coherent delays, so that the typical delay pattern shown by the tests clearly show a single or in some cases multiple coherent signal in this band in the Boston market. In particular, results from the tests in the L Band region in suburban Boston for range measurements show the results of three range tests in L Band demonstrating the extensive power variation, also showing the R^4 law.

(3) **Signaling:** The signaling used will be that of CDMA.

(4) **Multiple Access:** CDMA has been chosen as the selected access scheme since it provides the least scale effect and lowest capital per cell or per customer.

(5) **Physical/Electronic Interface:** This represents the cell site electronics. The first is the wideband RF digital processing front end and the second is the use of adaptive beamforming for unit optimization.

(6) **Network Management:** It has been has argued that it is necessary to view the provision of such a new service in a holistic sense, recognizing that it must meet the needs of the public by providing the highest quality at the lowest price. This can be achieved if and only if an overall interactive, dynamic, and adaptive network management system that integrates system performance with customer perception is deployed.

5.4 Adaptive Network Management

The Network Management element monitors and controls all elements of the network including all physical elements, all propagation paths and integrates the customer perception of the service to maintain overall quality.

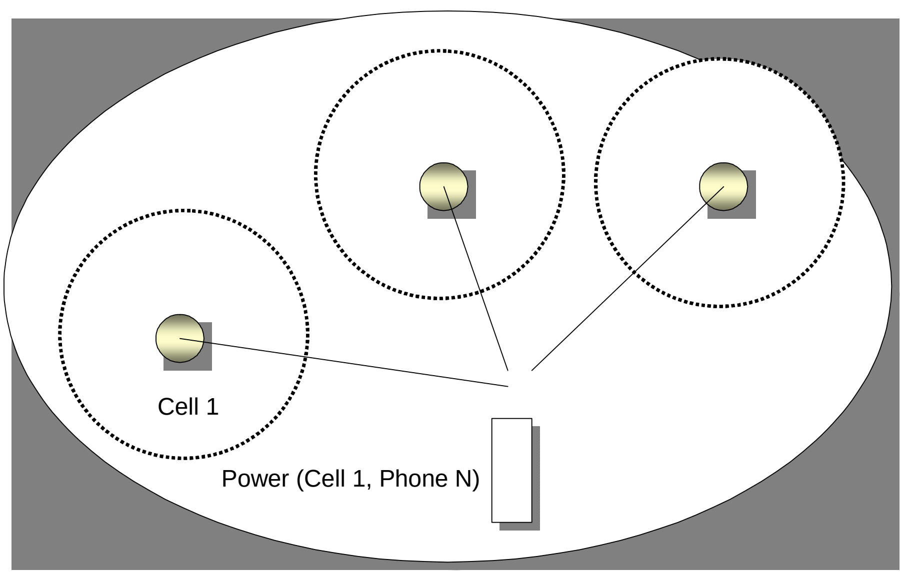
System Performance, SP, can be generally viewed as any single measure of the system operations that is controllable, directly or otherwise. A typical example of SP is the power received at a specific location from the transmitter, the bit error rate at a specific location at a specific time or any other such physical system measure. In contrast, Customer Perception, CP, is a quantitative measure of a more qualitative phenomenon.

CP is a measure of how well the customer perceives the service. It may be inferred via a set of test customers that respond to service perception as part of their usage of the system, it may be inferred as a result of the calls to customer service, or through any other controllable psychometric test procedures. The objective of an ANMS is to maximize CP by controlling SP factors.

The current approach to cell layout is to provide service through maximizing the power per cell and is capacity limited, is controlled in a passive form, and is designed in a classic "cookbook" fashion. The new paradigm that will be required for a PCN network is coverage driven, allows for flexible installation, -demonstrates real-time adaptive control, is driven by the desire to minimize the costs, while maximizing customer perception.

The approach herein to an ANMS is such that within each cell site a sensor is placed that measures the signal generated by that cell and all other cells. The sensor measures each cell transmit characteristics at a large number of locations throughout the system. Moreover, the sensors can be equipped with a GPS (Global -Positioning System) transceiver and be used in a dynamic fashion to monitor the signals very wide area. The system then transmits these sensor data elements back to the ANMS processor. The system then takes these signals and passes them into the optimizer for estimating the field over the entire coverage area and allows correlation of the performance with perception.

The design considers two factors: Portable Power Control and Cell Power Control. In the QUALCOMM portable design, the portables each provide power control to ensure low levels of CDMA interference. This is the Portable Power Control function, PPC. The Cell Power Control Function, CPC, is designed to provide a constant level of receive power at all cell sites and to minimize the interference from adjacent cell sites. The combination of these two elements ensures global Macro Power Control, MPC. MPC establishes an overall power field density such that all power levels are equal. Consider the field descriptions shown below;



Consider the two cases discussed; Portable Power Optimization and Cell Power Optimization.

5.4.1 Case 1: Portable Power Optimization

In this case we assume that we will use a simple modification of that used in CDMA with the QUALCOMM approach. Namely, let a portable transmits power, $P_k(t)$. Let us further assume that the propagation path to the receiver is given by;

$G_{ik}(\mathbf{x}_i,\mathbf{x}_k,t)$ = Signal propagation gain from point $\mathbf{x}_k$ to point $\mathbf{x}_i$, at time t.

Then the received power at the cell site from portable k is:

$$r_k(t) = G_{ik}(\mathbf{x}_i,\mathbf{x}_k,t)P_k(t)$$

Then the system determines what the power $P_k(t)$ should be to maximize the reception of all other signals. It then sends this adjustment back to the portable, which readjusts the power to $P^*_k(t)$, the optimum power. This is the essence of dynamic power control. The power control is with the portables and not with the cell sites. It works under the CDMA self noise control assumption.

The problem with this scheme is that there are propagation and processing delays. Let us assume that t_{Prop} is the propagation delay and that t_{Proc} is the processing delay. The we have the fact that P* is optimum at time t, but it may not be optimum at time:

$$t + t_{Prop} + t_{Proc}$$

Namely, if the channel is changing quickly enough such that;

$G_{ik}(\mathbf{x}_i,\mathbf{x}_k,t)$ is not equal to $G_{ik}(\mathbf{x}_i,\mathbf{x}_k,t + t_{Prop} + t_{Proc})$

Namely, with a one point reading it is not possible to assure effective one way power control. For example, if we define S(f) as the averaged spectrum of the process, $G_{ik}(\mathbf{x}_i,\mathbf{x}_k,t)$, where S(f) is equal to;

$$S(f) = \int G_{ik}(\mathbf{x}_i,\mathbf{x}_k,t)G_{ik}(\mathbf{x}_i,\mathbf{x}_k,t+t^*) \exp(-j2\pi ft^*)\, dt^*$$

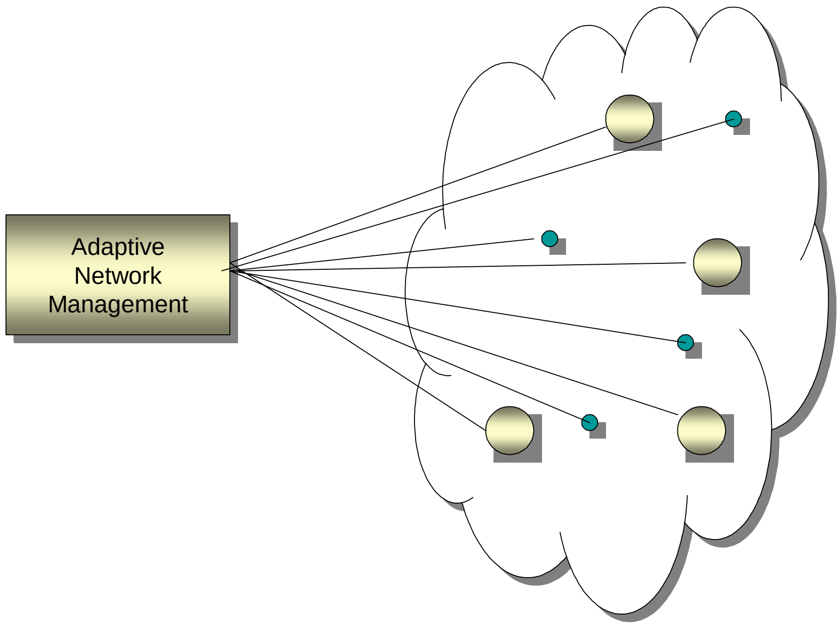
Now if we define BW as the bandwidth of the above spectrum, and if ;

$$t_{Prop} + t_{Proc} < 1/BW$$

then we have an adequate spread. Otherwise the channel is changing too fast.

5.4.2 Case 2: Cell Controller Site Power Control

Now consider the case of cell power control. This is the essence of the ANM system developed as part of this effort. Consider the system as shown in below.



In the above Figure, we show a set of transmit paths and receive paths. The objective is to set the transmit paths so as to optimize the power received by the portables. We further assume that the portables are measuring the transmitted power from each of the different cell sites. Let us assume that there are N cell sites and M sensors. Let us assume that each sensor measures power $P_{n,m}(t)$, where n is the cell site and m is the sensor location. This we have a set of measurements that are:

$$P_{n,m}(t) = G_{nm}(\mathbf{x}_n,\mathbf{x}_m,t)\, P_n(t)$$

and here $P_n(t)$ is the power transmitted to sensor m from cell n. The total received signal is the collection:

$$P_{n,m}(t)\ ;\ n=1,N,\ m=1,M$$

Where we have the following vector for the received signal, $\mathbf{r}(t)$:

$$G(x_1, x_{N+1}, t) P_1(t)$$
$$G(x_1, x_{N+2}, t) P_1(t)$$
$$.$$
$$.$$
$$.$$
$$G(x_N, x_{N+M-1}, t) P_N(t)$$
$$G(x_N, x_{N+M}, t) P_N(t)$$

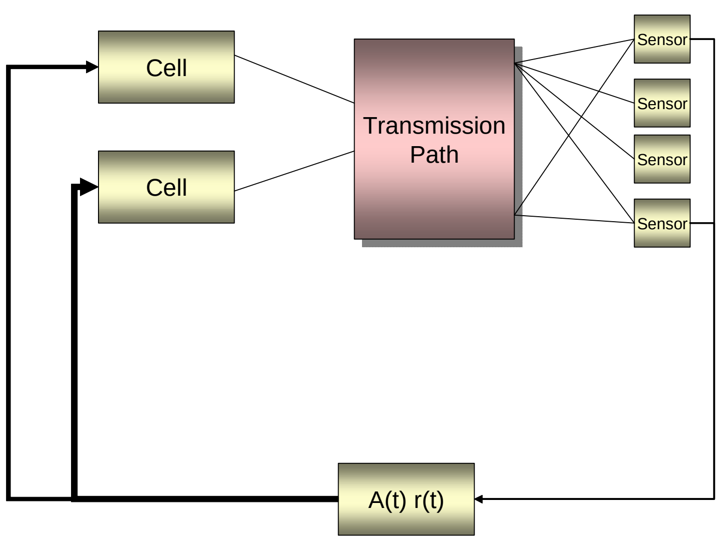
We now use this NM vector to be the control driver for the adaptive power control. Specifically we want to have a system that performs the following. Let set the objective function to be:

$\min(\| G_{nm}(\mathbf{x}_n,\mathbf{x}_m,t)\ P_n(t)-P_0 \|)$; for all n,m.

This is obtained through a feedback scheme where we have:

$\mathbf{P}(t) = \mathbf{A}(t)\ \mathbf{r}(t)$

where $\mathbf{P}(t)$ is the N vector of cell powers, $\mathbf{r}(t)$ the set of NM measured field elements, and $\mathbf{A}(t)$ a N X NM matrix based on the minimization principle stated above. The system is demonstrated in below.



The algorithm, as described in below shows that the system measurements are made and then passed through a control matrix A(t) that is time varying. The selection of the control matrix is chosen so as to reach the optimization criteria stated. This is a standard least squares approach in a distributed environment.[20]

[20]McGarty, Stochastic Systems and State Estimation, Wiley, 1974. The author develops the exact algorithm for array systems.

This system further allows for the determination of the random field, $G(\mathbf{x}_1,\mathbf{x}_2,t_1,t_2)$. Specifically, $G(\mathbf{x}_1,\mathbf{x}_2,t_1,t_2)$, represents the effects of a field at x_2 and time t_2 at point x_1 at time t_1. This algorithm allows for the calculation of this total field. Recall that $\mathbf{r}(t)$ is the sample vector of the field at M points at a single instance of time. We can use the sampling theorem as applied to random fields to determine the field at all points in the domain. This is in essence what the ANM is ultimately doing.

The unique sensor then is the element to do this. The sensor is a stand alone RF device that measures each signal sent on each frequency channel from each cell or re-radiator. It accomplishes this by means of a coded signal being imbedded in each transmitted signal. It decodes this message, measures the power, co-channel interference, the noise level, the interference level due to other carriers, and transmits all of these on an event and time driven basis to the cell and in turn to the network manager. The sensor uses standard chip Architectures but is uniquely designed via a programmable microprocessor. The sensor has been augmented to include a GPS transceiver to include sensor location if the sensor is on a mobile platform, such as a bus or taxi.

5.5 RF Broadband Digital Front End

The broadband RF front end processor is developed for military uses and recently has become an element in the digital radio effort. The Broadband Air Interface (BAI) accepts large bandwidth RF signals and converts them without distortion of any kind to a digital representation which can be digitally processed using existing computer hardware and highly developed digital signal processing (DSP) algorithms. The digital representation produced by the BAI provides more than 90 dB spurious free dynamic range, SFDR, within the representation. The SFDR of the modern DSP algorithms is better than 130 dB. Thus post processing SFDR of the combination is greater than 90 dB.

In practical terms, signals differing by as much as 90 dB can be simultaneously processed without distortion even though they may coexist in the same segment of the spectrum. In more practical terms, large RF bandwidths anywhere within the currently used portions of the RF spectrum can be converted to digital representations with virtually no distortion, and without the need for pre selectors of any kind, even in the most crowded section of spectrum.

The principal advantages that accrue as a result of adopting the BAI concept fall into two principal categories: first, those which are to the advantage of the provider, and, second, those which are advantage to the service user. The following list provides a summary of both categories:

(i) **The system level hardware is generic.** Application specific portions of the system reside in software. The system is software reprogrammable to accommodate new users. This factor is key to the approach of leveraging from software and allowing fixed infrastructure to be an enabler of flexible and mutable software.

(ii) **Protocol standards are minimized if not eliminated**. A user-equipment provider can be given the software architecture constraints in order to develop software necessary to utilize the PCN for their -specific Application. Then, the developer can immediately begin using the system to provide their -unique service. The open architecture capability will foster many new business applications which can be provided with minimal in-band or between-band interference. It is clear that new users will have to ensure orthogonality of access schemes in the digital domain, yet this allows co-location of such access protocols as TDMA and CDMA.

(iii) **Interoperability is assured**. Because an extremely large number of software processors can be made available with only the addition of memory, an almost unlimited number of protocol possibilities can be accommodated with only a small additional cost. This further ensures the low scale economy factor already presented for this business. Thus simultaneous operations of TDMA, FDMA and CDMA are possible.

(iv) **Co-location, the ability to directly locate and connect to the CO in the LEC public network is further enhanced and simplified.** Once the RF environment is converted to digital format by the BAI, the replica can be transmitted in toto and error free to LEC and IEC connections for post processing. The only penalty is more bandwidth, but with co-location, access to large bandwidths, especially in the form of dark fiber, is readily available from the LEC. This site extender concept minimizes the hardware requirements ate cells and microcells and provides maximum versatility in the selection of sites, further enhancing the ANMS approach and reducing the overall operating costs.

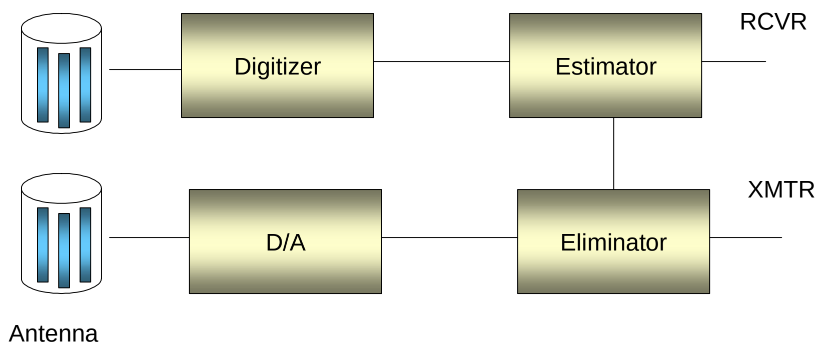
(v) **Adaptive network management (ANM) can be implemented**. Since the entire network is software based with the BAI, the implementation of detailed diagnostics and control schema is readily deployed. The quality of service indicators to ensure maximized CP are readily optimized as the user perception is better understood.

(vi) **Adaptive networks can be provided with the use of such technology as GPS in a low cost design.** Thus time tagging of data, time difference of arrival (TDOA), can be obtained for each signal. TDOA accuracies in the picosecond range are readily achievable and are an integral part of the BAI output processor, and this information will allow for the adaptive optimization of the overall signal processing and electromagnetic field control.

The use of the wide dynamic range has been an enabling point for the development of a interference proof front end design. Specifically, the Wideband dynamic front end takes the incoming CDMA bands and digitizes them directly. Assuming a 15 MHz front end and a sample rate of 30 Msps, and a sample density of 16 bits per sample, this is a 480 Mbps data stream. This can now be used for interference sensing and interference protection. below depicts the scheme that is currently under design.

In this scheme, called a Wideband interference sense and reject system (WISR), the RF is automatically digitized. The digital signal is then detection processed to determine if there are any transmitters in the band. If there are transmitters, the transmit, which is also digitized, is then notch filtered digitally, to ensure that the transmit spectrum is 60 dB below the interferer level. This approach utilizes all of the processing capabilities of high speed digital electronics and can be achieved in a very low cost fashion. It also gets tied into the cell management system so that the environment of each cell can be viewed and optimized.

The developments to date are shown in below.



In this design, the received signal is given as follows:

$$\mathbf{r}(t) = \sum_{i=1}^{M} \mathbf{s}_i(t) + \sum_{i=1}^{N} \mathbf{w}_j(t) + \mathbf{n}(t)$$

Here we have the spread signals, s_i, and the interferers, w_j. We assume M spread signals and M interferers. The approach is to digitize the front end received signal and then to apply interference sensing estimators to the signal. Specifically, we have developed an estimator that allows for the estimation of interference in noise. [21] Thus the estimator, using a Kalman Filtering scheme on the filtered digital received signals, provides and estimate:

$$\mathbf{w}_{i*} = \boldsymbol{F}\,(\mathbf{r}i;\ i=1,..,M)$$

where ***F*** is the estimator operator on the received data.

5.6 Adaptive Beamforming

The adaptive beamforming efforts are currently under R&D developments at MIT Lincoln Laboratory and discussions are in progress towards the transfer of this technology.

The design objectives of the adaptive beam forming antenna system are as follows:

(i) Reduce Power Consumption of portables by 50% or more.

(ii) Extend the range of a cell by a factor of 50% or more. The target range is 4.5 miles.

---

[21] McGarty, The Effect of Interfering Signals, IEEE Vol AES 10, No 1, pp. 70-77, Jan. 1974.

(iii) Extend the capacity of CDMA to 90% of the theoretical limit in major urban and suburban markets.

(iv) Integrate into the ANMS and Wideband front end systems.

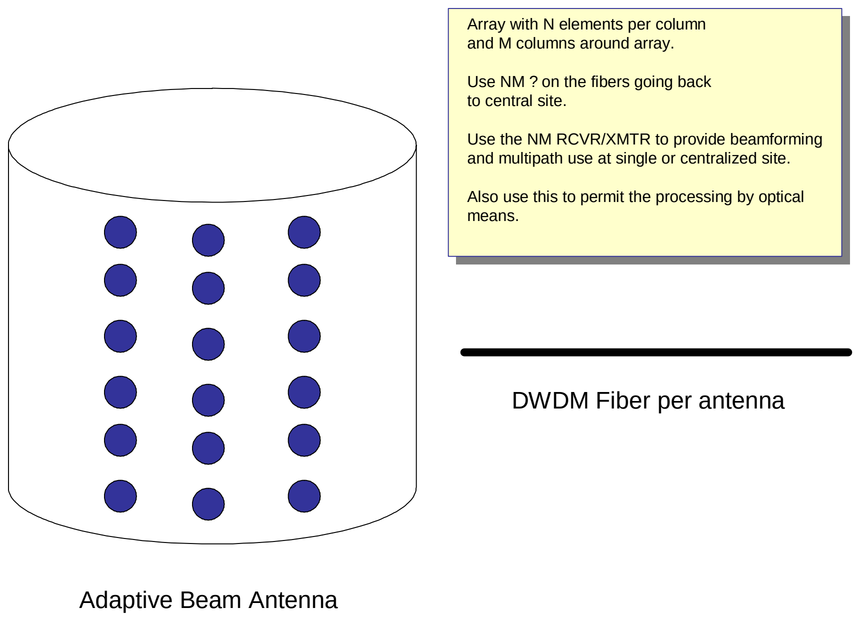
The ABF system works as follows:

(i) A portable requests service by either requesting a cell be set up or by having a call placed to it.

(ii) The Portable acknowledges its presence and the ABF antenna locks onto its signal.

(iii) The ABF, using a real time adaptive beam forming algorithm, uses the portable signal as a tone, and forms a beam on the portable to raise its relative transmit and receive gain by 10 dB.

The system is as shown in below.



Let us assume that there are N signals coming from portables. Let the jth signal be given by:

$$s_j(t) = \sqrt{E_j/T}\; s_0(t-t_j)\,\exp(-j\,\mathbf{k}_j \cdot \mathbf{r})$$

where; E is the energy of this pulse, T the time duration, K the wave vector of direction, and r the position vector from the antenna to the portable. Specifically:

$$\mathbf{k} = k_0 \begin{bmatrix} \cos\alpha \\ \cos\beta \\ \cos\gamma \end{bmatrix}$$

and the cosines are the directional cosines.

The total received signal is:

$$r(t) = \sum_{i=1}^{N} s_j(t) + n(t)$$

We then use an estimator, based upon maximum likelihood estimation to determine the set of $\mathbf{k}_j$. Specifically; $\mathbf{k}^*_j$ is determined by the Likelihood Function.[22]

Determining the angle of arrival allows for the development of a highly accurate adaptive beam forming antenna. Let w_{ij} be the antenna weights of the adaptive array. Let us further assume that we adapt the weights according to a least squares fit for the forming of beams on the targets.[23] The system must provide:

(i) a carrier locking system that optimizes the beam power on a specific signal that in turn can provide a 10-25 dB advantage over a fixed antenna pattern. This system uses the techniques developed by the Principal in adaptive target tracking in other applications areas.

(ii) a diffuse multipath integration system that allows the use of diffuse reflections, those from incoherent sources, to add a further 3 to 6 dB increase in adaptive gain adjustment.

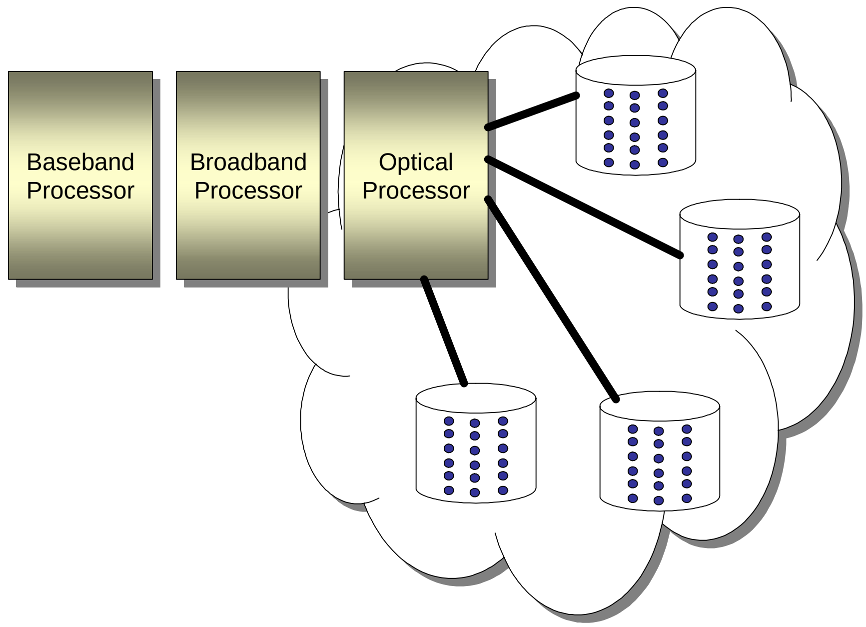
(iii) an interferer rejection system that reduces the C/N from interferers that may add an additional 7 to 10 dB of protection.

5.7 Multiple Layered Signal Processing

One of the approaches that we have taken is a multilayered signal processing approach. In the architecture discussed above there is a full broadband optical link back to a single or multiple but small number of processing points. Now processing can be accomplished on several levels; optical, broadband, baseband. This is shown graphically below.

[22]McGarty, Azimuth-Elevation Estimation Performance of Spatially Dispersive Channel, IEEE Vol AES-10, No 1, pp. 58-69, January, 1974. The author details the use of adaptive arrays for the estimation of arrival angles and determines the performance of these estimators. The work was done while the author was on the Faculty of MIT and in the Research Staff at MIT Lincoln Laboratory.

[23]Monzingo, R.A., T. W. Miller, Adaptive Arrays, Wiley, 1980. This reference details most of the major algorithms that are used in adaptive beam forming.



5.7.1 Optical Processing

By having the full RF spectrum at an optical carrier, there is a significant amount of processing achievable by the optical components themselves. In addition this processing can be achieved by means of optical elements which makes for less expensive processing especially if this is performed in bulk. Finally one now has access in one place to all signals and thus one anticipates that full and positive use of multipath can be achieved. Finally all of this can be placed upon an integrated fiber backbone.

5.7.2 Broadband Processing

Broadband processing assumes the signal is fully in tact in a RF manner and that we can then process the signal in RF segments. This has seen significant use in military and intelligence areas and now can be applied to the commercial uses of wireless.

5.7.3 Baseband Processing

The baseband processing must now take into account the normal elements which we are so accustomed but also any and all elements which arise by have so much more information.

The use of these three levels makes the Shannon paradigm look simplistic, namely we throw a great deal away to achieve that simplicity. The question is what is the value of what we keep or have thrown away?

5.8 Cell Allocation Optimization

In the development of the several technologies during this quarter, there has been considerable progress also on the optimization of cell site allocation. Specifically, many of the current s have contended that small cell radii are optimum. We have, as a result of our detailed design including but not limited to the ANM efforts, concluded that an optimum design can be obtained.

In this section we summarize these efforts. From a policy perspective the larger cell site reduces cost, but requires more power per cell. There are petitioners who argue against larger cells with higher power. We have argued elsewhere that these arguments are based on anti competitive attempts to drive cost up and protect their entrenched positions.

Let us consider the cell life cycle costs. We make the following assumptions:

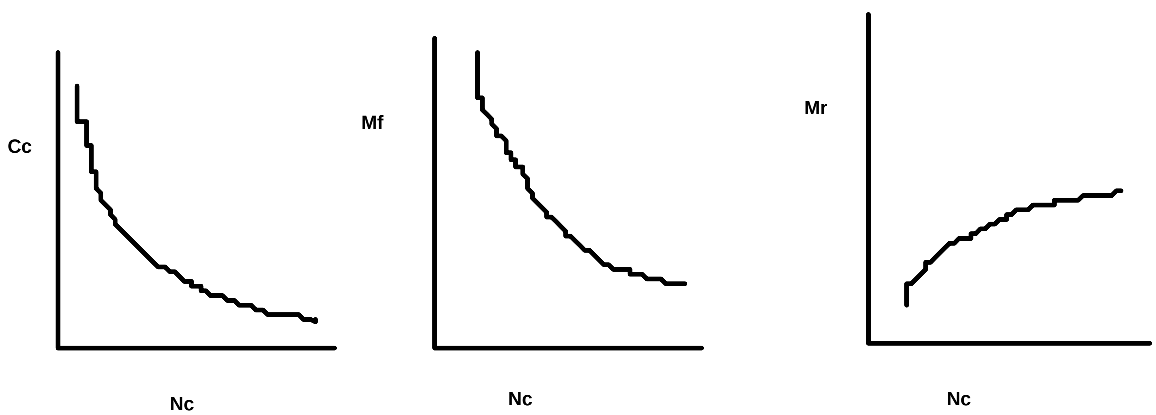
> C_C *is the capital per cell*
> T_C *is the life of a cell*
> M_F *is the MTBF, mean time between failure per cell.*
> M_R *is the mean time to repair a cell*
> N_C *is the number of cells needed to cover an area.*
> R_C *is the cell Radius*
> A_C *is the cell area*
> P_C *is the power per cell*
> S_{OPS} *is the salary per operational person per hour*
> C_T *is the total life cell cost*

Now we can state:

$$C_T\ (N_C) = N_C\ C_C + N_C\ \ T_C(\ M_F / M_R)\ S_{OPS}$$

Note that this has been parameterized on N_C. We shall now show that, given a fixed coverage area and a fixed radius of coverage, or that is setting N_C as the variable, we can demonstrate that there is an optimum, namely a N_C *, that minimizes the total life cycle cost. To do this we examine the impact of the number of cells on the capital per cell, the MTBF, and the MTTR. We show these factors in below.

below



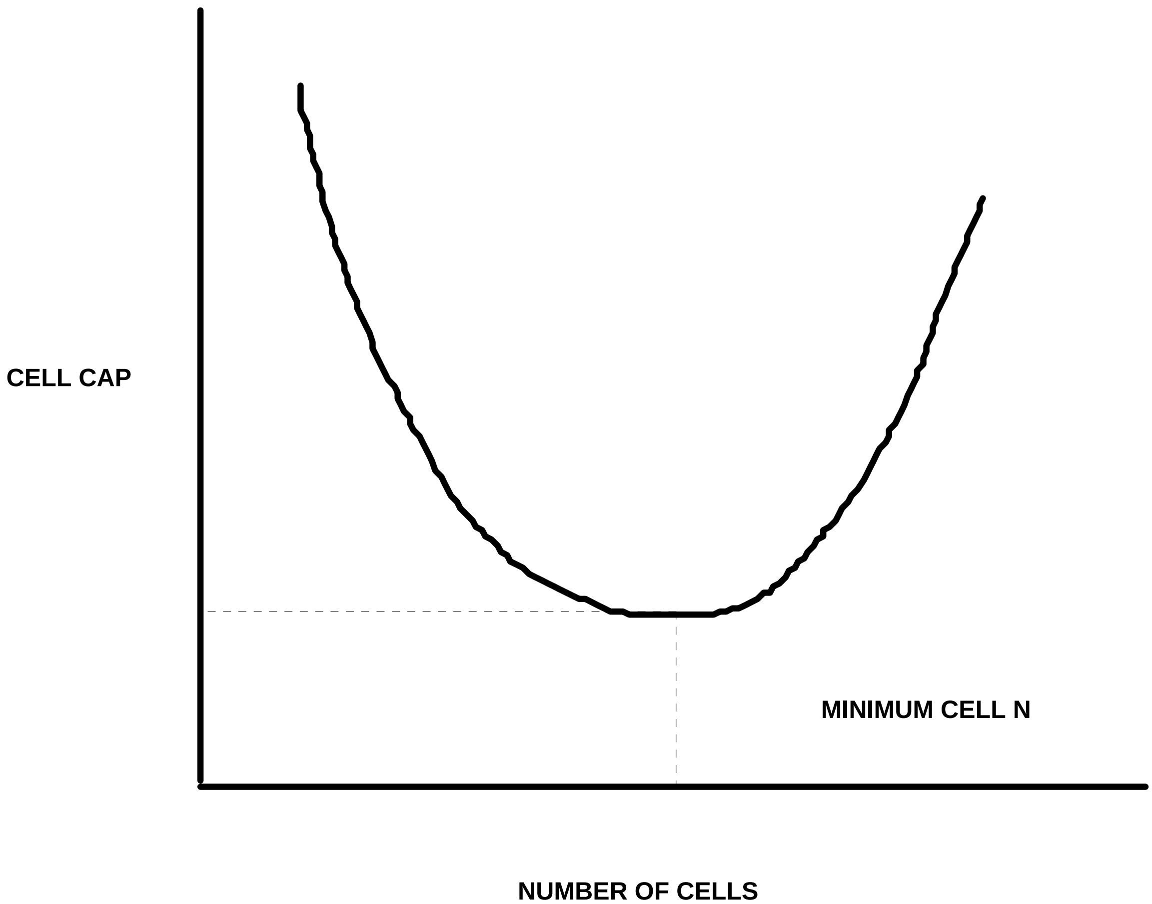
Let us note the result of the above:

(1) **Capital Per Cell**: As the number of cells decrease, or equivalently as the cell size decreases, the capital per cell decreases. However, there is a point of diminishing decreases.

(2) **MTBF Per Cell:** As the cell size shrinks, the complexity of the cell decreases. This may drive the cell MTBF lower. However, as the cell decreases the cost per cell overpowers this factor and makes the cell MTBF get larger.

(3) **MTTR Per cell:** The MTTR per cell increases as the number of cells increase since the cells are more inconveniently located and less centrally located.

This model then implies the following as shown in below.



We can now determine the cell minimum by differentiating the cost equation with respect to N_C. Specifically, this results in;

$$C_C(N_C) + N_C\ \partial C_C(N_C) / \partial N_C + C_L\ A_C(N_C) + N_C\ C_L\ \partial A_C(N_C) / \partial N_C = 0$$

The solution of this equation, where we have defined A_C as M_R/M_F, is the minimum cost cell size. In our analysis to date we have found that this is about two to three miles. This dramatically differs from the lower cell radii as proposed by others who have not performed this minimization.

6 APPENDIX 2: SUMMARY OF MULTIMEDIA ARCHITECTURAL AND TECHNOLOGY ISSUES

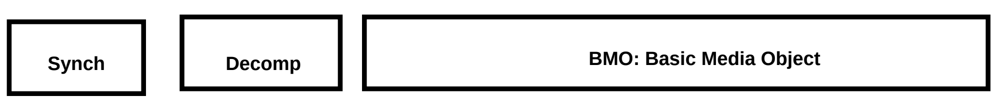
This appendix presents the various details on the elements which we have developed in the area of multimedia communications.

6.1 Multimedia Data Objects

In a more standard computer communications environment, the data objects have significant structure and they are frequently integrated into a system wide data base management system that ensures the overall integrity of the data structures. In a multimedia environment, the data elements are more complex, taking the form of video, voice, text, images and may be real time in nature or can be gathered from a stored environment. More importantly, the separate data objects may combined into more complex forms so that the users may want to create new objects by concatenating several simpler objects into a complex whole. Thus we can conceive of a set of three objects composed of an image, a voice annotation and a pointer motion annotating the voice annotation. The combination of all three of these can also be viewed as a single identifiable multimedia object.

Before commencing on the issues of communications, it is necessary to understand the data objects that are to be communicated. We can consider a multimedia data object to be composed of several related multimedia data objects which are a voice segment, an image and a pointer movement (e.g. mouse movement). As we have just described, these can be combined into a more complex object. We call the initial objects Simple Multimedia Objects (SMOs) and the combination of several a Compound Multimedia Object (CMO). In general a multimedia communications process involves one or multiple SMOs and possibly several CMOs.

The SMO contains two headers that are to be defined and a long data sting. The data string we call a Basic Multimedia Object (BMO). There may be two types of BMOs. The first type we call a segmented BMO or SG:BMO. It has a definite length in data bits and may result from either a stored data record or from a generated record that has a natural data length such as a single image screen or text record. We show the SMO in Figure 21..



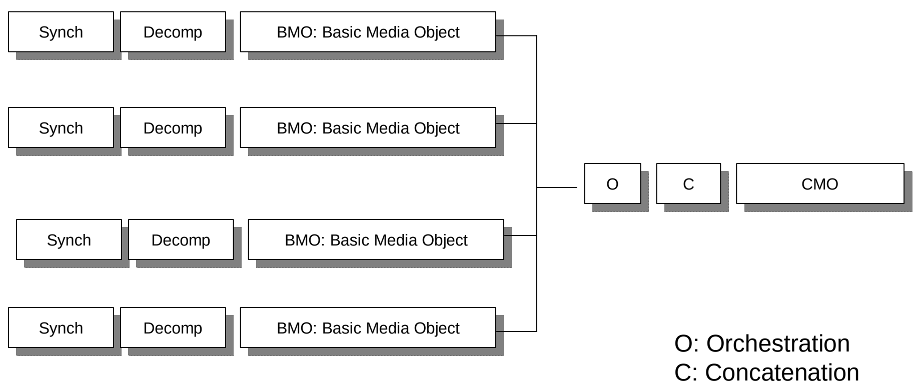
The second type of BMO is a streamed BMO, ST:BMO. This BMO has an a priori undetermined duration. Thus it may be a real time voice or video segment.

A simple multimedia object, SMO, is a BMO with two additional fields; a Synchronization field (Synch) and a Decomposition field (Decomp). The above depicts the SMO structure in detail. The Synch field details the inherent internal timing information relative to the BMO. For example it may contain the information on the sample rate, the sample density and the other internal temporal structure of the object. It will be a useful field in the overall end to end timing in the network.

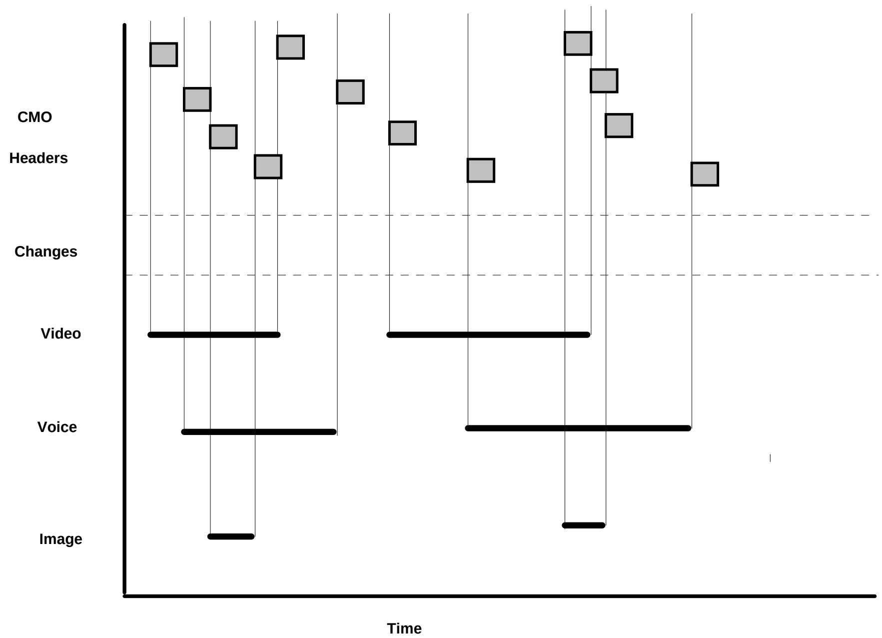
The second field is called the Decomp field and it is used to characterize the logical and spatial structure of the data object. Thus it may contain the information on a text object as to where the paragraphs, sentences, or words are, or in an image object, where the parts of the image are located in the data field.

These fields are part of an overall architecture requirement finds it necessary to provide an "out-of-band" signaling scheme for the identification of object structure. The object structure is abstracted from the object itself and becomes an input element to the overall communications environment. Other schemes use in-band signaling which imbeds the signal information with the object in the data stream. This is generally an unacceptable approach for this type of environment.

When we combine these objects together we can create a compound multimedia object. This is shown below. A CMO has two headers, the Orchestration header and the Concatenation header. The Orchestration header describes the temporal relationship between the SMOs and ensures that they are not only individually synchronized but also they are jointly orchestrated. The orchestration concept has also been introduced by Nicolaou. In this paper we further extend the orchestration function beyond that of Nicolaou. The concatenation function provides a description of the logical and spatial relationships amongst the SMOs.



We can now add dynamics to this process and we show this in the following. In this Figure we show first the real time display of video, voice, image, pointer and text. In the Figure we depict the time that these object are involved in the system dynamics. We then also plot the times that the CMO, the concatenation of all simultaneous objects, change in this system. We depict the change element below. Then we also show the CMO headers that are flowing in the network at each change interval. It is this dynamic process of data elements that must be controlled by the session layer to be discussed in the next session.



We can also expand the concept of a CMO as a data construct that is created and managed by multiple users at multiple locations. In this construct we have demonstrated that N users can create a CMO by entering multiple SMOs into the overall CMO structure.

The objectives of the communications system are thus focused on meeting the interaction between users who are communicating with CMOs. Specifically we must be able to perform the following tasks:

1. Allow any user to create an SMO and a CMO.

2. Allow any user or set of users to share, store, or modify a CMO.

3. Ensure that the user to user communications preserves the temporal, logical and spatial relationships between all CMOs at all users at all times.

4. Provide an environment to define, manage and monitor the overall activity.

5. Provide for an environment to monitor, manage and restore all services in the event of system failures or degradation.

We shall see in the next section that the session layer service address all of these requirements.

6.2 Session Layer Functions

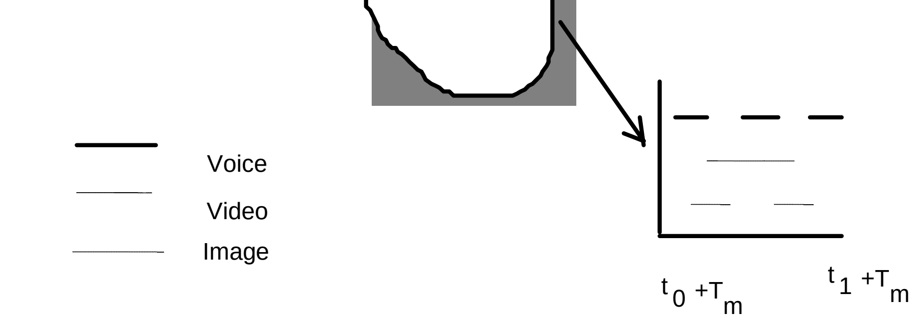
The OSI layered communications architecture has evolved to manage and support the distributed communications environment across error prone communications channels. It is presented in detail in either Tannenbaum or Stallings. A great deal of effort has been spent on developing and implementing protocols to support these channel requirements. Layer 7 provides for the applications interface and generally support such applications as file, mail and directory. The requirements of a multimedia environment are best met by focusing on layer 5, the session layer whose overall function is to ensure the end to end integrity of the applications that are being supported. Some authors (See Couloris and Dollimore or Mullender) indicate that the session function is merely to support virtual connections between pairs of processes. Mullender specifically deals with the session function in the context of the inter-process communications (IPC). In the context of the multimedia object requirements of the previous section, we can further extend the concept of the session service to provide for IPC functionality at the applications layer and specifically with regards to multimedia applications and their imbedded objects.

Here we have shown the session entity which is effectively a session service server. The entity is accesses from above by a Session_Service Access Point (S_SAP). The session entities communicate through a Protocol Data Unit (PDU) that is passed along from location to location. Logically the session server sits atop the transport server at each location.

The servers are conceptually at a level above the transport level. We typically view the transport servers as communicating distributed processes that are locally resident in each of the transmitting entities. This then begs the question as to where does one place the session servers. Are they local and fully distributed, can they be centralized, and if so what is their relationship to the Transport servers. Before answering these questions, let us first review how the session services are accessed and how they are communicated.

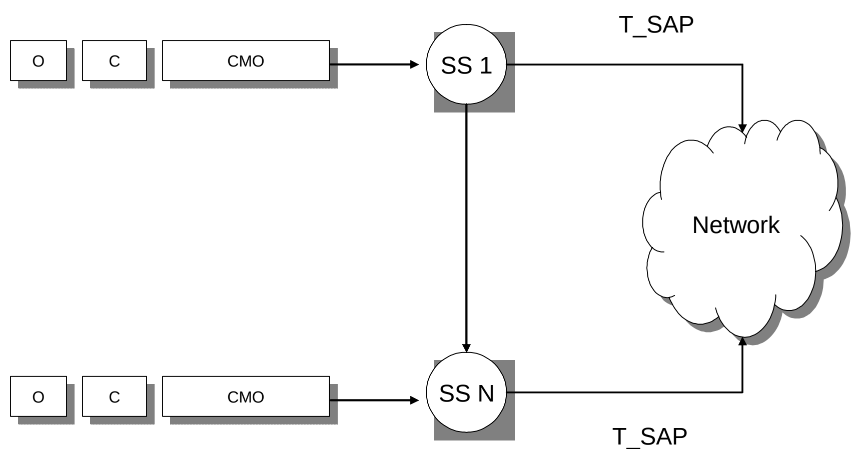
Session services are accessed by the higher layer protocols by invoking session service primitives. These primitives can invoke a dialog function such as Token_Give. The application may make the call to the S_SAP and this request may be answered. There are typically four steps in such a request, and these are listed in Stallings who shows that the requests are made of the session server by entity one and are responded to by entity two. The model does no however say where the session server is nor even if it is a single centralized server, a shared distributed server, or a fully distributed server per entity design. We shall discuss some of the advantages of these architectural advantages as we develop the synchronization service.

Synchronization is a session service that ensures that the overall temporal, spatial and logical structure of multimedia objects are retained. Consider the example shown in Figure 6.1. In this case we have a source generating a set of Voice (VO), video (VI), and Image (IM) data objects that are part of a session. These objects are simple objects that combined together form a compound multimedia object. The object is part of an overall application process that is communicating with other processes at other locations. These locations are now to receive this compound object as show with the internal timing retained in tact and the absolute offset timing as shown for each of the other two users.



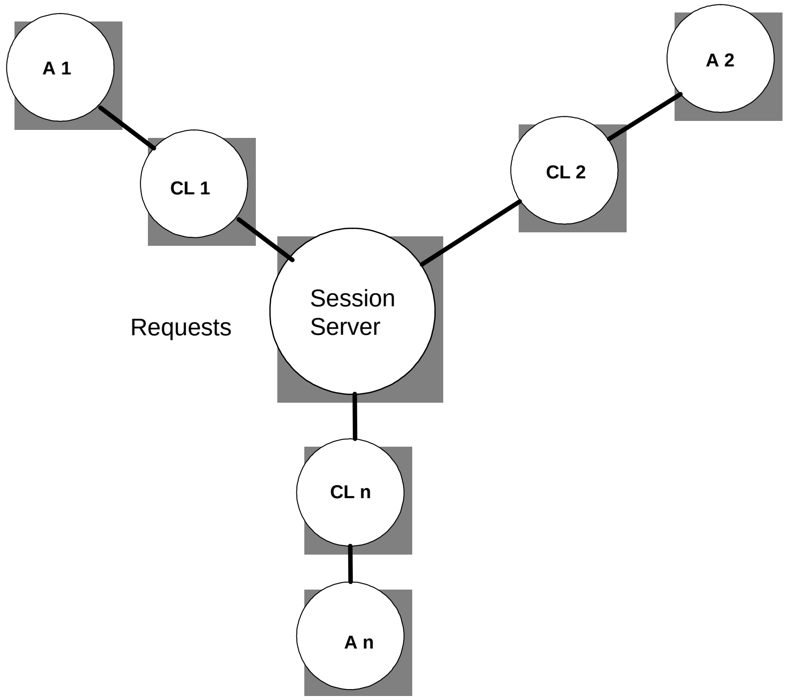
In this example, the synchronization function provided by the session server to the applications processes at the separate locations is to ensure both the relative and absolute timing of the objects. The location of the functionality can be centralized or distributed. Let us first see what the overall timing problem is. Consider a simple SMO synchronization problem. The network than transmits the packets and they arrive either in order or out of order at the second point. The session server must then ensure that there is a mechanism for the proper reordering of the packets at the receiving end of the transmission.

The architecture for the session synchronization problem is shown in the following Figure. Here we have a CMO entering the network, knowing that the session server at Server 1 must not only do the appropriate interleaving but it must also communicate with the other servers (in this case K and N) to ensure that de-interleaving is accomplished. We show the session servers communicating with the network through the T_SAP and that in turn takes care of the packetizing. However, we also show that the session server, 1 and N, communicate in a out of band fashion, using some inter process communications (IPC) scheme, to ensure that the relative actions are all synchronized amongst each other.

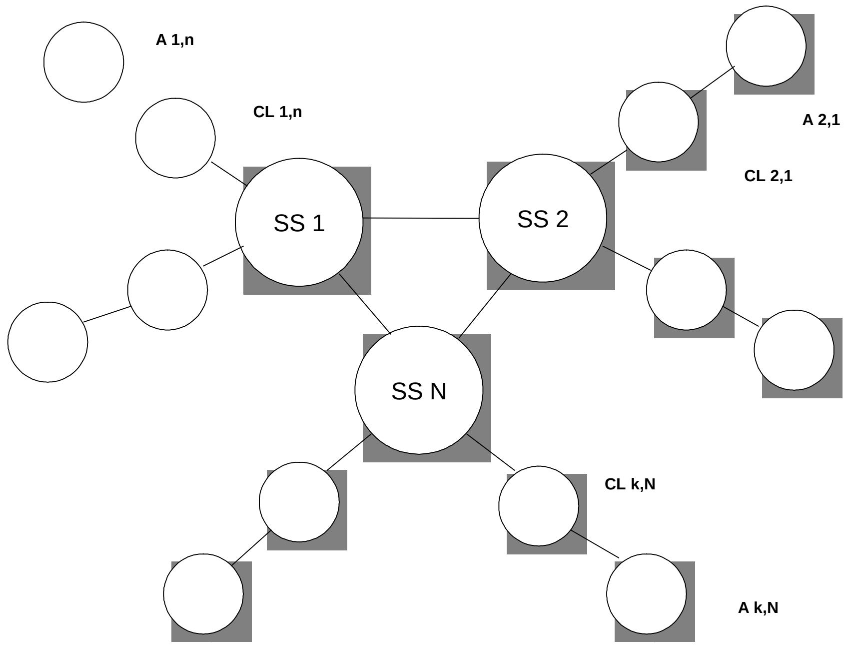


We can now envision how the architecture for this can be accomplished. There are two schemes:

Centralized: The following Figure depicts the centralized synch scheme for the session service. It assumes that each application (A) has a local client (CL). The application communicates with the local client (CL) to request the session service. The session server is centrally located and communicates with the application locally by means of a client at each location. This is a fully configured client server architecture and can employ many existing techniques for distributed processing (See Mullender or Coulouris et al).



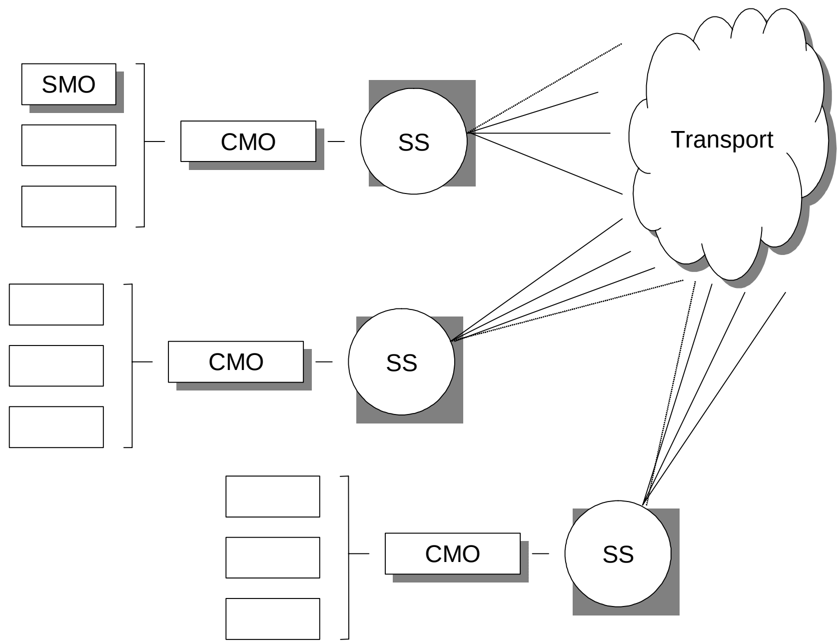
Distributed: In contrast to the centralized scheme, we can envision a fully distributed session server architecture as shown in the following Figure. In this case we have a set of applications, and cluster several applications per session server. We again user local clients to communicate between the session server and the applications. The clients then provide local clusters of communications and the session servers allow for faster response and better cost efficiency. However, we have introduced a demand for a fully distributed environment for the session managers to work in a distributed operating system environment. As a further extreme, we could eliminate the clients altogether by attaching a session server per applications and allow for the distributed processing on a full scale.



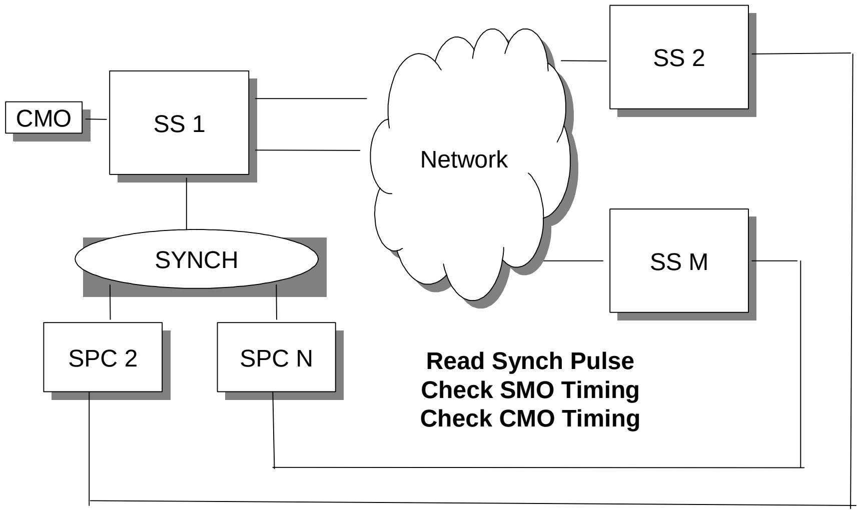
The major functions of the session server in its synch mode are:

1. bind together simple objects into compound objects as requested by the application.
2. provide intra object synchronization to ensure that all timing within each object is met.
3. orchestrate amongst objects to provide inter object timing.
4. minimize delay, slippage, between simple objects.
5. minimize delay, latency, between different users.

To effect these requirements, we have developed and implemented a scheme that is based on a paradigm of the phased locked loop found in communications. We show this configuration below. Here we have a distributed session server architecture receiving a CMO from an application. The session server passes the message over several paths to multiple users. On a reverse path, each server passes information on the relative and absolute timing of the CMO as it is received using the session services primitives found in the OSI model. Generally for segmented BMOs this is a simple problem but with streamed BMOs this becomes a real time synchronization problem.



The specific implementation is shown below. Here we show M session servers and at the sending server we do the pacing of the packets to the T_SAP and allow for the interleaving of the SMOs. Based on the commands from the feedback system we provide delay adjustment, through caching and resetting priorities to the T_SAP for quality of service adjustments for the lower layer protocols.



True multimedia communications thus is not just more web front ends and enhanced games. It is fundamentally different. For the moment we must assume that there will be an evolution of human interfaces to balance the contribution of all the senses and add to the ability to have displaced conversationality. The question on the table is how do we design networks to support these in the most effective as well as most inclusive manner.