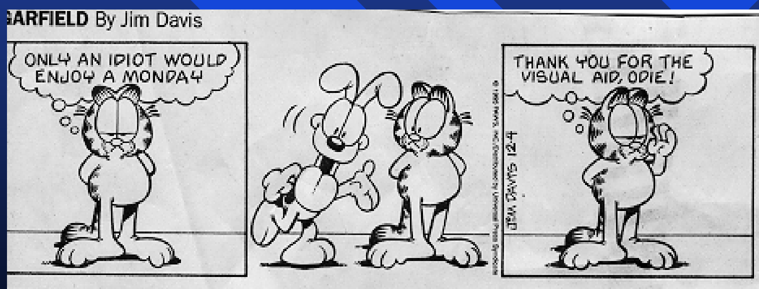
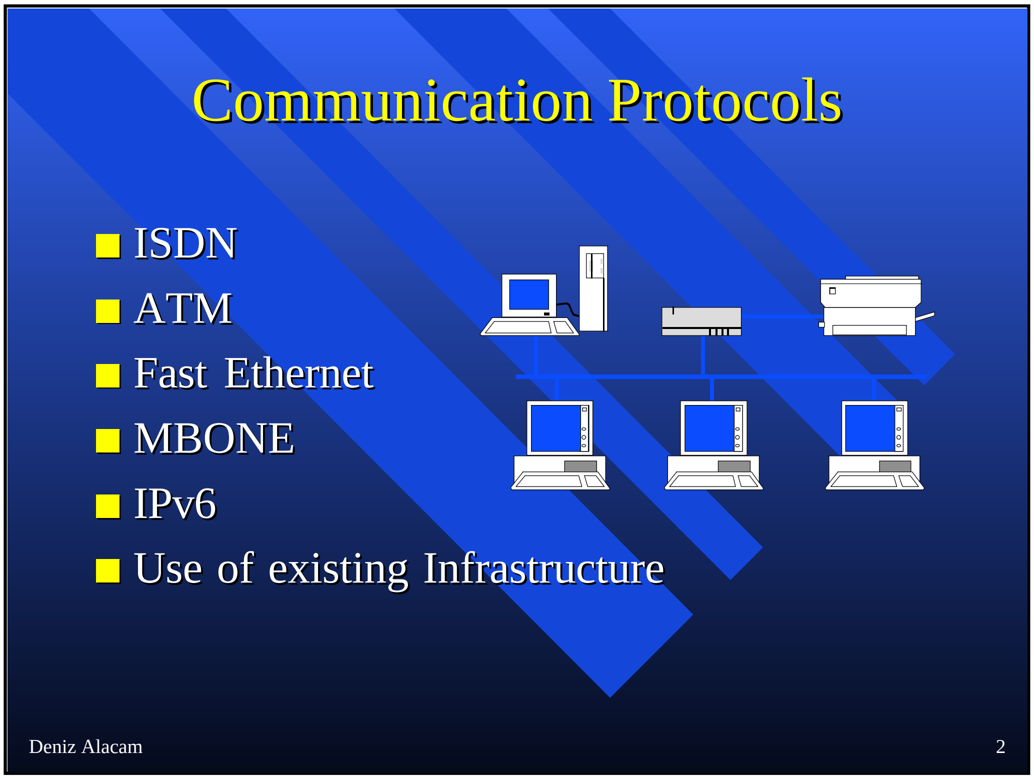
Advanced Distributed Systems and Communication

Communication Protocols





The type of data communications facility used is a function of the nature of the application, the number of computers involved, and their physical separation.

ISDN (integrated services digital networks)

Many public carriers are converting their existing public switched telephone networks to enable data to be transmitted without modems. The resulting networks, which operate in all-digital mode are known as ISDN.

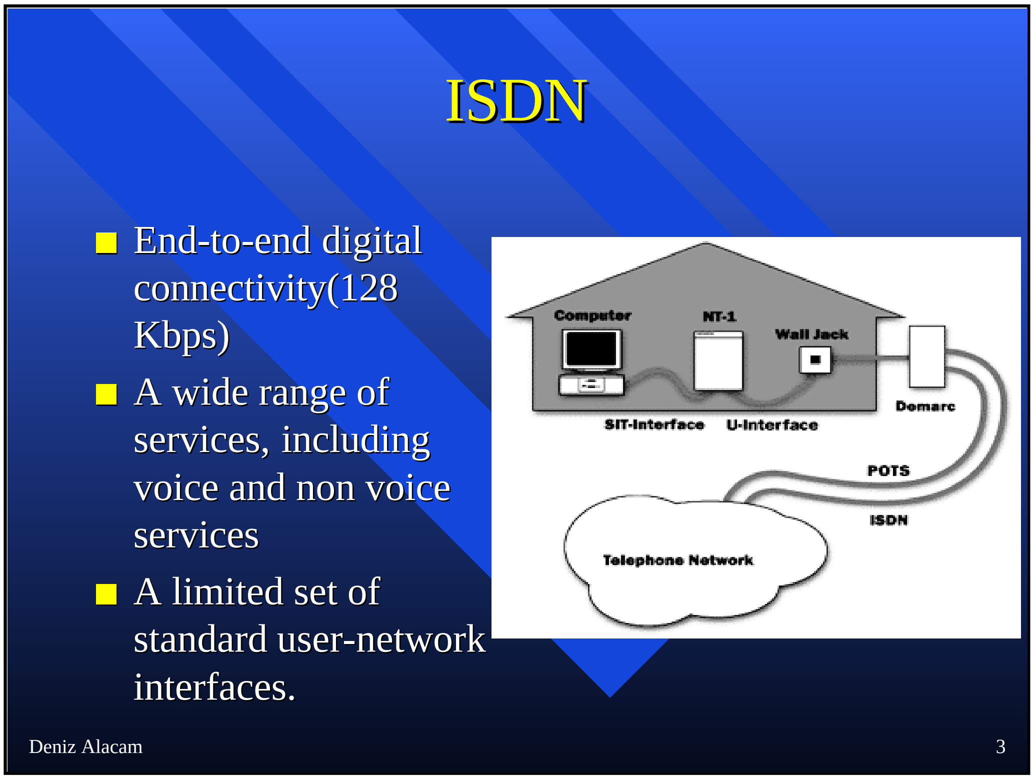
Broadband multi-service networks -- ATM

High-speed and bridged local area networks -- Fast Ethernet

Multicast Backbone -- MBONE

Internetworking -- IPv6

Use of existing infrastructure -- Cable Modems



ISDN uses switched digital connections. Perhaps the most important single feature of ISDN, however, is that it offers inexpensive dialed digital access to the worldwide telecommunications network. It is no longer necessary to lease costly dedicated lines for high-speed digital transmission, or to limit data speed and accuracy by using modems to convert digital signals to analog pulses.

ISDN stands for Integrated Services Digital Network - the name for digital telephone service that works over existing copper telephone wiring. There are several types of ISDN service, but the most appropriate type for individual computer users, and the type that this site focuses on, is the ISDN Basic Rate Interface (BRI).

Basic Rate ISDN divides the telephone line into 3 digital channels: 2 "B" channels and one "D" channel, each of which can be used simultaneously. The B channels are used to transmit data, at rates of 64k or 56k (depending on your telephone company). The D channel does the administrative work, such as setting up and tearing down the call and communicating with the telephone network. With two B channels, you can make two calls simultaneously.

Most of the world's existing telephone network is already digital. The only part that typically isn't digital is the section that runs from the local exchange to your house or office. ISDN makes that final leg of the network digital.

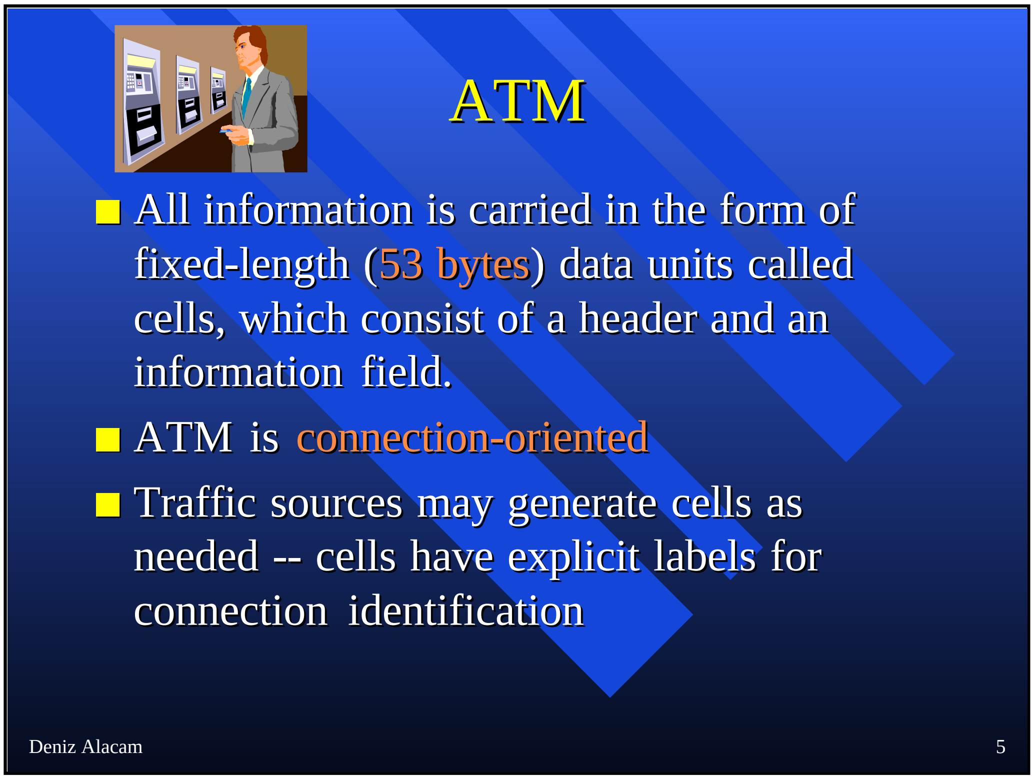
Bearer channels (or B channels) transmit user information at relatively high speeds, while separate Data channels (or D channels) carry call set-up, signaling and other information.

It handles all types of information. Unlike some other digital communications technologies, ISDN handles all types of information voice, data, studio-quality sound, still and moving images. They are all digitized, and transmitted at high speeds in the same flow of data.

It handles many devices and many telephone numbers, on the same line. Up to eight separate telephones, fax machines or computers can be linked to a single ISDN connection, and have up to 64 "call appearances" of the same or different telephone numbers.

It supports up to three calls at the same time. Two voice, fax or PC "conversations," and one data "conversation" can take place at the same time, through the same ISDN connection.

It offers variable, responsive transmission speeds. Two or more channels can be combined into a single larger transmission "pipe." Channels can be assembled as needed for a specific application (a large video conference, for example), and then broken down and reassembled into different groups for different applications (normal voice or data transmissions). Combining B channels in this manner is called inverse multiplexing , or bonding.



ATM Technology is based on powerful, yet flexible concepts.

1.When information needs to be communicated, the sender NEGOTIATES a "requested path" with the network for a connection to the destination. When setting up this connection, the sender specifies the type, speed and other attributes of the call, which determine the end-to-end quality of service.

2.Another key concept is that ATM is a switched based technology. By providing connectivity through a switch (instead of a shared bus) several benefits are provided:

Dedicated bandwidth per connection

Higher aggregate bandwidth

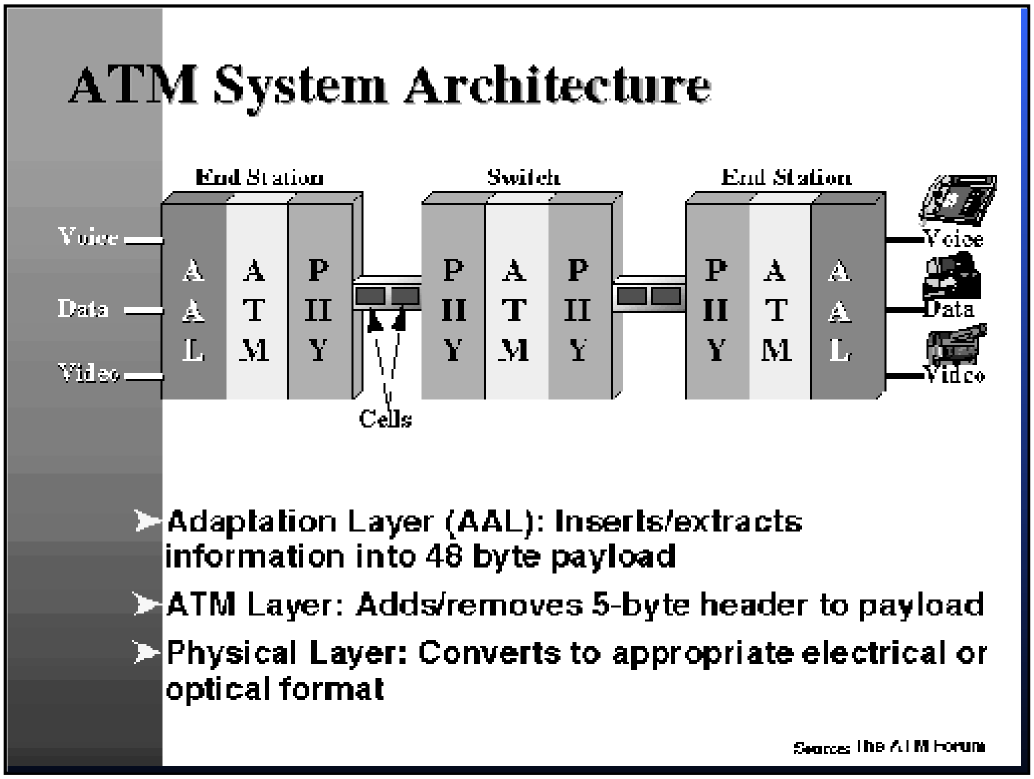
Well defined connection procedures

Flexible access speeds

Using ATM, information to be sent is segmented into fixed length cell, transported to and re-assembled at the destination. The ATM cell has a fixed length of 53 bytes. Being fixed length allows the information to be transported in a predictable manner.

This predictability accommodates different traffic types on the same network.

The cell is broken into two main sections, the header and the payload. The payload (48 bytes) is the portion which carries the actual information-either voice, data, or video. The Header (5 bytes) is the addressing mechanism.



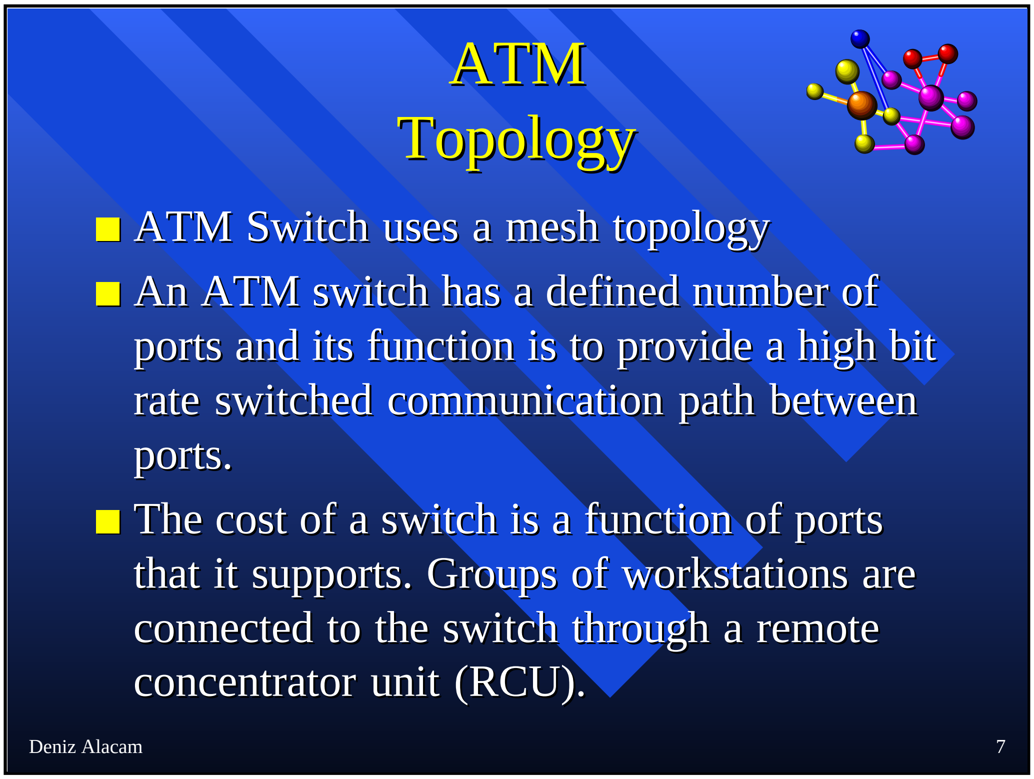
ATM System Architecture

ATM is a layered architecture allowing multiple services like voice, data and video, to be mixed over the network. Three lower level layers have been defined to implement the features of ATM.

The Adaptation layer assures the appropriate service characteristics and divides all types of data into the 48 byte payload that will make up the ATM cell.

The ATM layer takes the data to be sent and adds the 5 byte header information that assures the cell is sent on the right connection.

The Physical layer defines the electrical characteristics and network interfaces. This layer "puts the bits on the wire." ATM is not tied to a specific type of physical transport.



ATM uses a mesh topology. An ATM switch has a defined number of ports and its function is to provide a high bit rate switched communication path between ports.

The cost of a switch is a function of ports that it supports. To minimize the number of ports, groups of workstations are connected to the switch through a remote concentrator unit (RCU).

The role of RCU is simply to multiplex/demultiplex the cell streams from/to those workstations that are involved in a network transaction onto/from the link connecting the concentrator to a switch port.

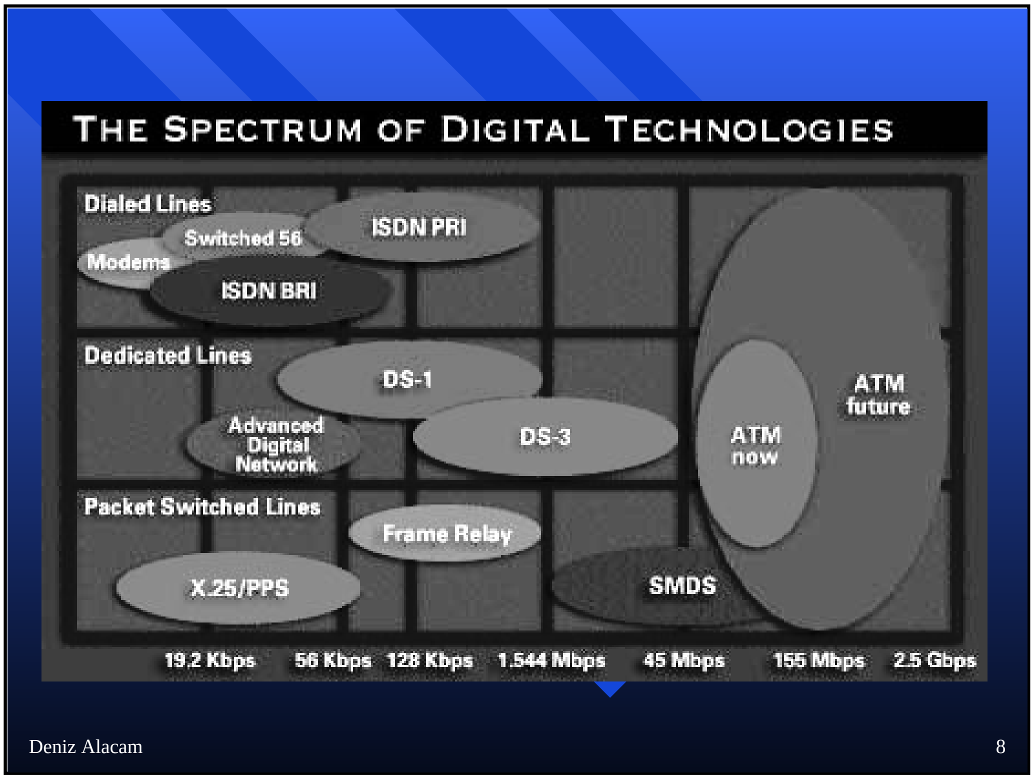
ATM has several key benefits:

1.One Network-ATM will provide a single network for all traffic types-voice, data, video. ATM allows for the integration of networks improving efficiency and manageability.

2.Enables new applications-Due to its high speed and the integration of traffic types, ATM will enable the creation and expansion of new applications such as multimedia to the desktop.

3.Compatibility-Because ATM is not based on a specific type of physical transport, it is compatible with currently deployed physical networks. ATM can be transported over twisted pair, coax and fiber optics

4.Incremental Migration-Efforts within the standards organizations and the ATM Forum continue to assure that embedded networks will be able to gain the benefits of ATM incrementally-upgrading portions of the network based on new application requirements and business needs..



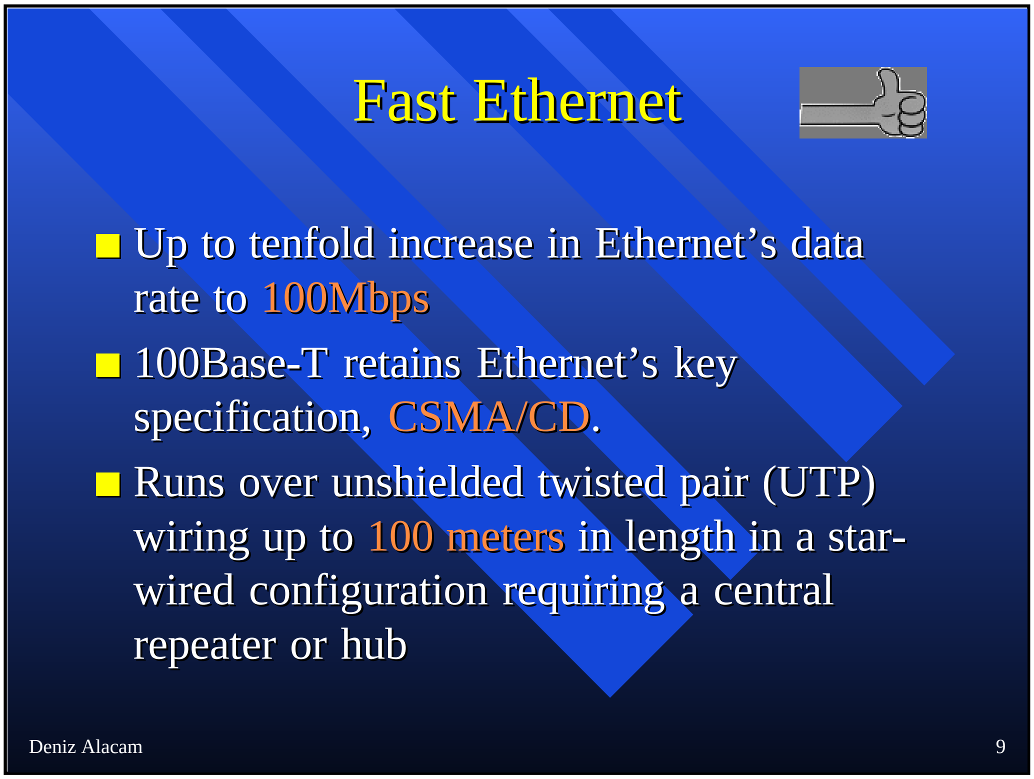
5.Simplified Network Management-ATM is evolving into a standard technology for local, campus/backbone and public and private wide area services. This uniformity is intended to simplify network management by using the same technology for all levels of the network.

6.Long Architectural Lifetime-The information systems and telecommunications industries are focusing and standardizing on ATM. ATM has been designed from the onset to be scaleable and flexible in:

Geographic distance, number of users, access and trunk bandwidths (As of today, the speeds range from Megabits to Gigabits)

The IEEE 802.3 specification allows for a total cabel length -with repeaters - of 2.5 km. The worst case signal propagation delay is the time for a signal to propagate twice this length. The standard allows a worst case signal propagation delay - including repeater delay - of 50ms which is equivalent to 500 bits at 10Mbps. A safety margin is added to give a minimum frame size of 512 bits. If this maximum length is reduced, then CSMA/Cd access method can be used with higher bit rates. This is the basis of the Fast Ethernet standard.

In practice, the vast majority of 10 Base T installations use less than 100 m of cable to connect each node to the hub. This means the maximum distance between any two nodes is 200 m and hence the worst-case path length for collision detection purposes is 400 m. Cllearly, a higher bit rate can be used while still retaining the same CSMA/CD MAC method and minimum frame size of 512 bits. In the standard the bit rate is set at 100 Mbps.



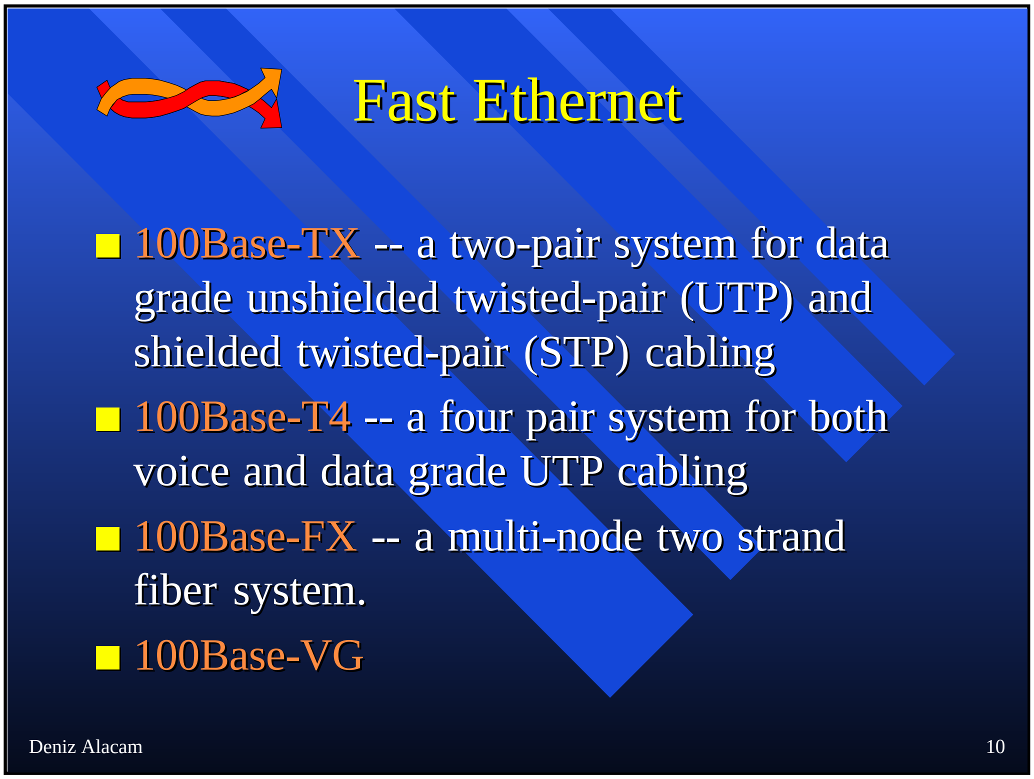
When the IEEE standardization committee met to begin work on a faster Ethernet system, two approaches were presented.

One approach was to speed up the original Ethernet system to 100-Mbps, keeping the original CSMA/CD medium access control mechanism. This approach is called 100BASE-T Fast Ethernet.

Another approach presented to the committee was to create an entirely new medium access control mechanism, one based on hubs that controlled access to the medium using a "demand priority" mechanism. This new access control system transports standard Ethernet frames, but it does it with a new medium access control mechanism. This system was further extended to allow it to transport token ring frames as well. As a result, this approach is now called 100VG-AnyLAN.

The IEEE decided to create standards for both approaches. The 100BASE-T Fast Ethernet standard described here is part of the original 802.3 standard. The 100VG-AnyLAN system is standardized under a new number: IEEE 802.12.

Compared to the 10-Mbps specifications, the 100-Mbps system results in a factor of ten reduction in the bit-time, which is the amount of time it takes to transmit a bit on the Ethernet channel. This produces a tenfold increase in the speed of the packets over the media system. However, the other important aspects of the Ethernet system including the frame format, the amount of data a frame may carry, and the media access control mechanism, are all unchanged.

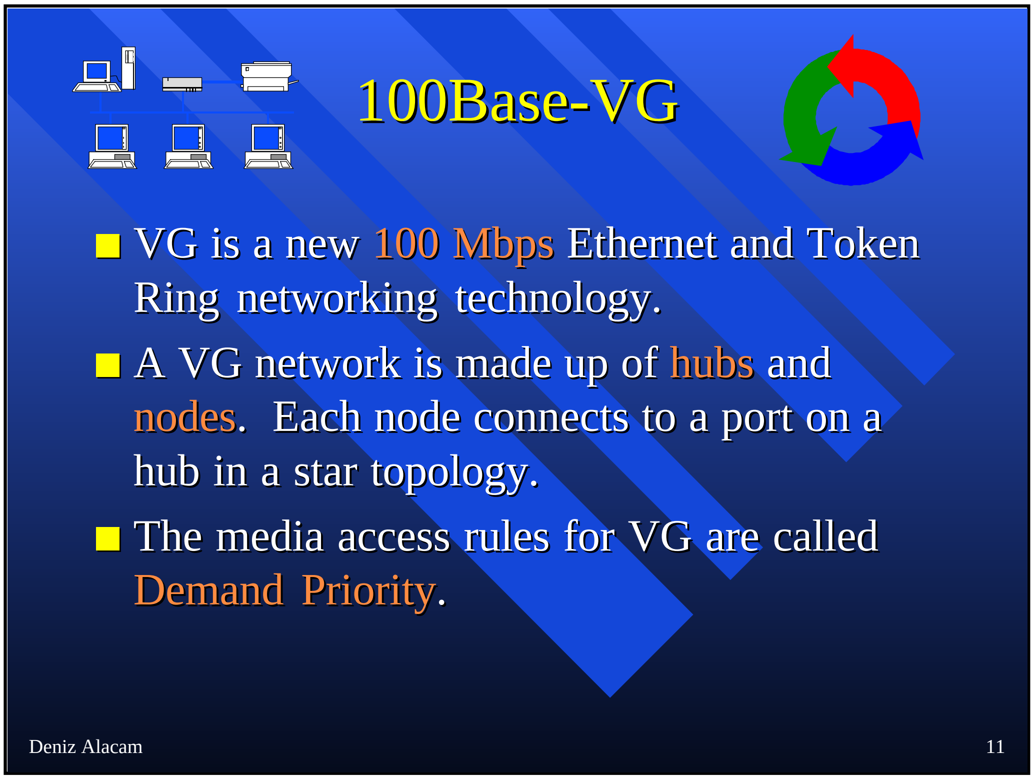


The Fast Ethernet specifications include mechanisms for Auto-Negotiation of the media speed. This makes it possible for vendors to provide dual-speed Ethernet interfaces that can be installed and run at either 10-Mbps or 100-Mbps automatically.

Baseband signaling simply means that Ethernet signals are the only signals carried over the media system.

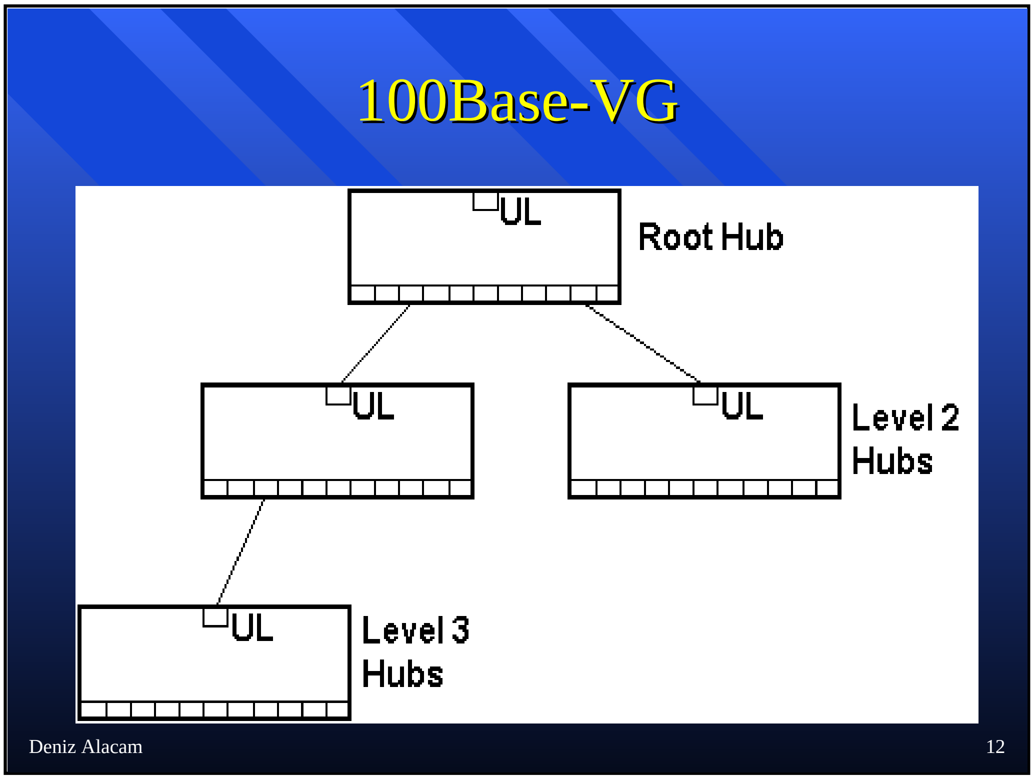
The third part of the identifier provides an indication of the segment type. The "T4" segment type is a twisted-pair segment that uses four pairs of telephone-grade twisted-pair wire. The "TX" segment type is a twisted-pair segment that uses two pairs of wires and is based on the data grade twisted-pair physical medium standard developed by ANSI. The "FX" segment type is a fiber optic link segment based on the fiber optic physical medium standard developed by ANSI and that uses two strands of fiber cable. The TX and FX medium standards are collectively known as 100BASE-X.

The 100BASE-TX and 100BASE-FX media standards used in Fast Ethernet are both adopted from physical media standards first developed by ANSI, the American National Standards Institute. The ANSI physical media standards were originally developed for the Fiber Distributed Data Interface (FDDI) LAN standard (ANSI standard X3T9.5), and are widely used in FDDI LANs.



The VG specification (802.12) states that the design goals are to:

1. Provide a minimum data rate of 100 Mb/s,
2. Provide smooth migration from ISO/IEC 8802-3 (Ethernet) and ISO/IEC 8802-5 (Token Ring) LANs,
3. Support either ISO/IEC 8802-3 (Ethernet) or ISO/IEC 8802-5 (Token Ring) frame formats and MAC service interface to the LLC,
4. Support a cascaded star topology over twisted pair Category 3, 4, and 5 and fiber optic wiring,
5. Allow topologies of 2.5 km and greater with 3 levels of cascading,
6. Provide a physical layer bit error rate of less than 10-8,
7. Provide fair access and bounded latency,
8. Provide two priority levels, normal and high priority,
9. Provide a low latency service through high priority for support of multimedia applications over extended networks,
10. Support an option for filtering individually addressed packets at the repeater to enhance privacy,
11. Support network management to monitor network performance, isolate faults, and control network configuration,



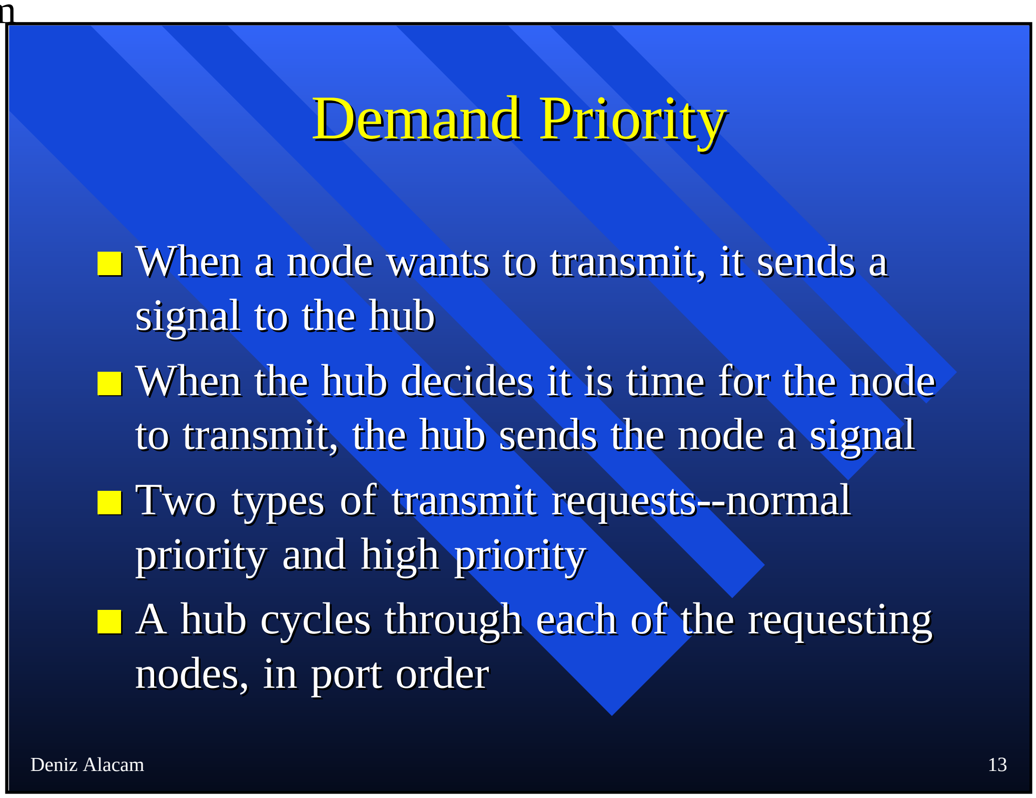
12.Enable low cost implementation and high levels of integration, and

13.Provide for robust operation by testing the Physical Layer connection before allowing an end node to enter the network and by removing disruptive nodes.

In simple English, the goal is to provide a 100 mega-bit networking that provides fair access to the network and to support both Ethernet and Token Ring frame types.

A VG network is made up of hubs and nodes. Each node connects to a port on a hub in a star topology.

Hubs connect to each other in a tree topology where each hub can have zero, one or more children hubs and zero or one parent hub. The top most hub (the one without a parent) is usually called the root hub.



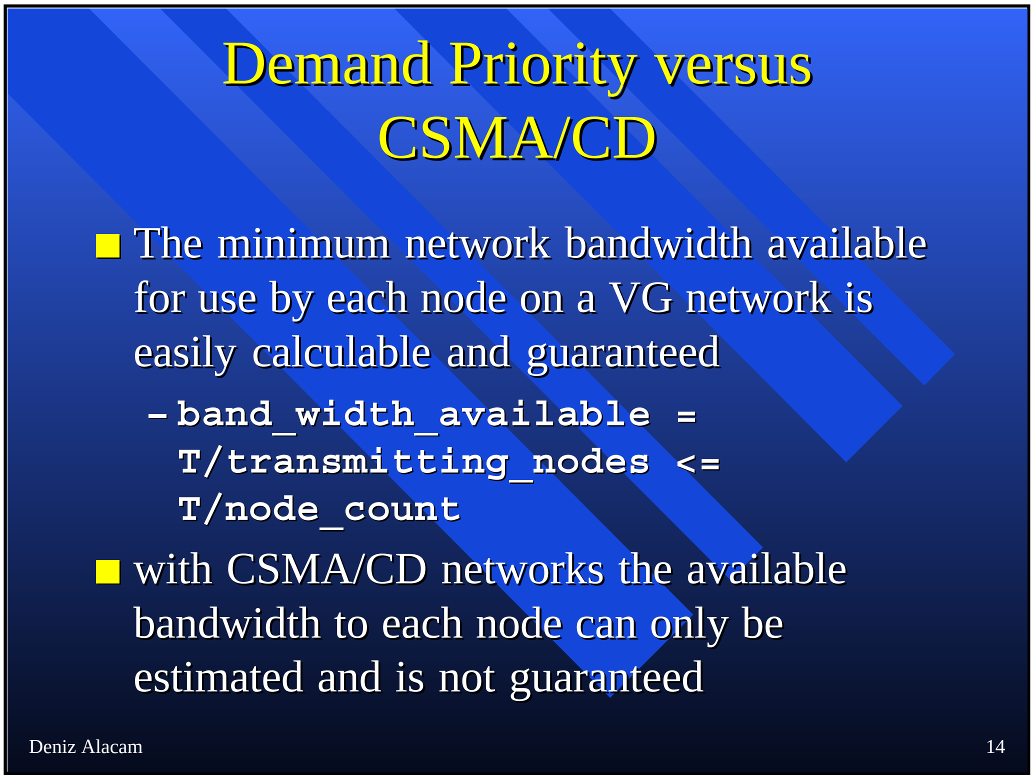
The media access rules for VG are called Demand Priority. Other networking technologies, including Ethernet, Fast Ethernet, Token Ring, Arcnet, and FDDI, implement their media access rules in the nodes themselves (there is no central access control). In contrast, all the media access rules for VG are implemented in the hub. The hub, or hubs, decide when each node can transmit. Nodes do not cooperate with each other to decide when they are able to transmit. VG nodes don't pass tokens (even when using Token Ring frame types), or detect and resolve collisions (except when plugged into a 10Base-T hub).

When a node has one or more packets to transmit, it sends a signal (a demand) to the hub. When the hub decides it is time for the node to transmit, the hub sends the node a signal. A node can send a hub two types of transmit requests--normal priority and high priority.

A hub cycles through each of the requesting nodes, in port order, allowing each to transmit one packet. A hub will service each of the nodes asserting a high priority request before servicing any nodes with normal priority requests. It is important to note that the hub pays no attention to nodes that don't need to transmit; they are skipped and do not take time in the hub's round robin algorithm. It is important to note that the round robin scanning is extremely fast and is implemented in hardware by the RMAC chips in the hub.

There two types of transmit demands--normal priority and high priority. This feature is designed to support multi-media video streams or other time-critical applications that require low transmit latency.

To prevent starving normal priority nodes by allowing high priority transmits to hog the network, the hub maintains a priority promotion timer for each node. This timer is started when a node first asserts a normal priority transmit request. If the timer expires before the node gets a chance to transmit, then its request will be upgraded to a high priority request.



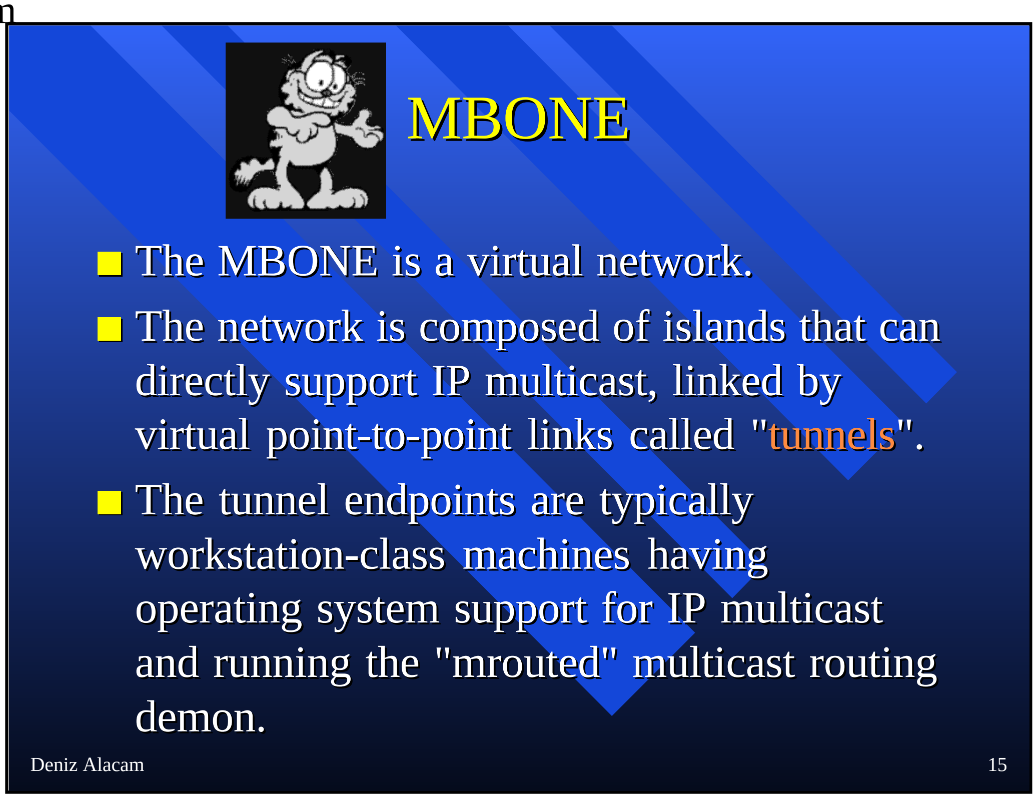
The demand priority scheme is designed to improve on the CSMA/CD access method used by 10mpbs Ethernet and 100mpbs fast Ethernet. VG is different from CSMA/CD networks in that for any period of time the minimum network bandwidth available for use by each node on a VG network is easily calculable and guaranteed, where with CSMA/CD networks the available bandwidth to each node can only be estimated and is not guaranteed. This is the difference between a deterministic network (VG) and probabilistic network (CSMA/CD Ethernet). In other words, other nodes can never cause the available bandwidth to a node to fall bellow T / node_count.. VG is designed to work well, even under high (90%+) network utilization (often called network load).

Note that this does not mean that the CSMA/CD access method is 'bad' or inferior to Demand Priority, they are just different. Ethernet (10 and 100) networks usually perform quite well. However, there are some instances where Ethernet performance can suffer greatly. VG is designed to work well in these situations.

The big benefit of DP is that it allocates network bandwidth fairly between all nodes on the network that need to transmit.

There are two key ideas here--the notion of fairness and the fact that network bandwidth is divided between only those nodes that need to transmit.

Fairness means that for any given cycle all transmitting nodes get to send exactly one packet. One node (or a small set of nodes) cannot hog a majority of the available network bandwidth.



The MBONE is called a virtual network because it shares the same physical media -- wires, routers and other equipment --as the Internet.

The MBONE allows multicast packets to travel through routers that are set up to handle only unicast traffic. Software that utilizes the MBONE hides the multicast packets in traditional unicast packets so that unicast routers can handle the information.

The scheme of moving multicast packets by putting them in regular unicast packets is called tunneling. In the future, most commercial routers will support multicasting, eliminating the headaches of tunneling information through unicast routers.

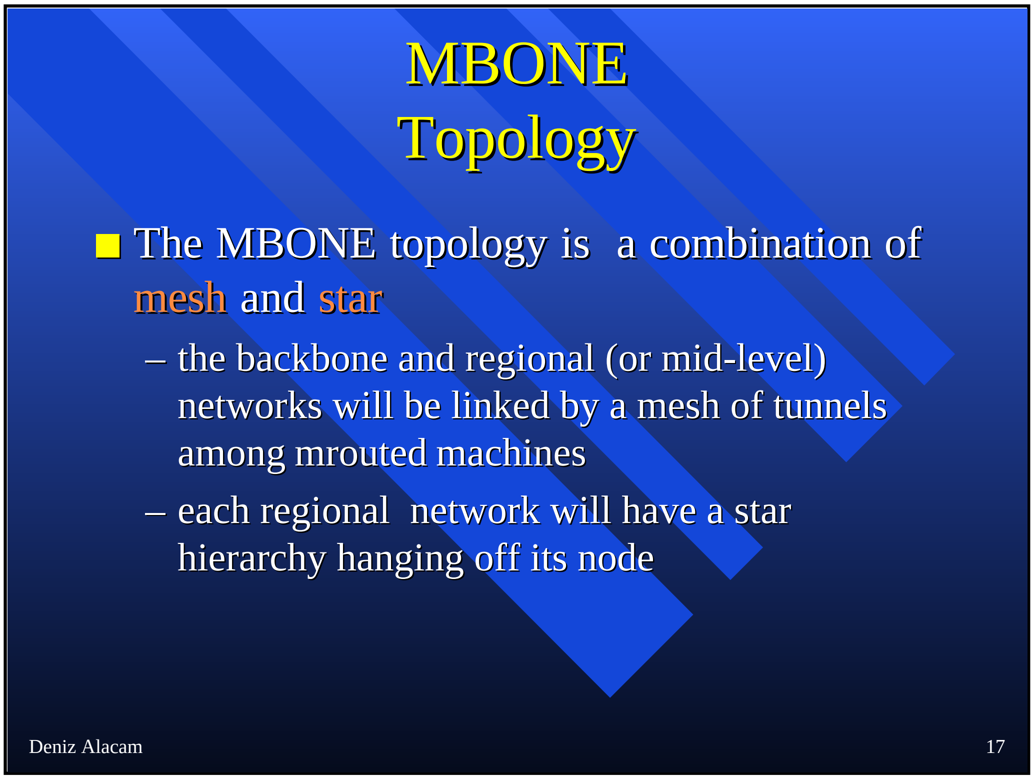
When the multicast packets that are hidden in unicast packets reach a router that understands multicast packets, or a workstation that's running the right software, the packets are recognized and processed as the multicast packets they really are.

Machines (workstations or routers) that are equipped to support multicast IP are called mrouters (multicast routers). Mrouters are either commercial routers that can handle multicasting or (more commonly) dedicated workstations running special software that works in conjunction with standard routers.

The MBONE is a loose confederation of sites that currently implement IP multicasting. It is at best a temporary utility that will eventually become obsolete when multicasting is a standard feature in Internet routers. By then there will be an established base of MBONE users (which should make the router manufacturers happy). The utilities and programs that work on today's MBONE will undoubtedly work on the multicast backbone of tomorrow.

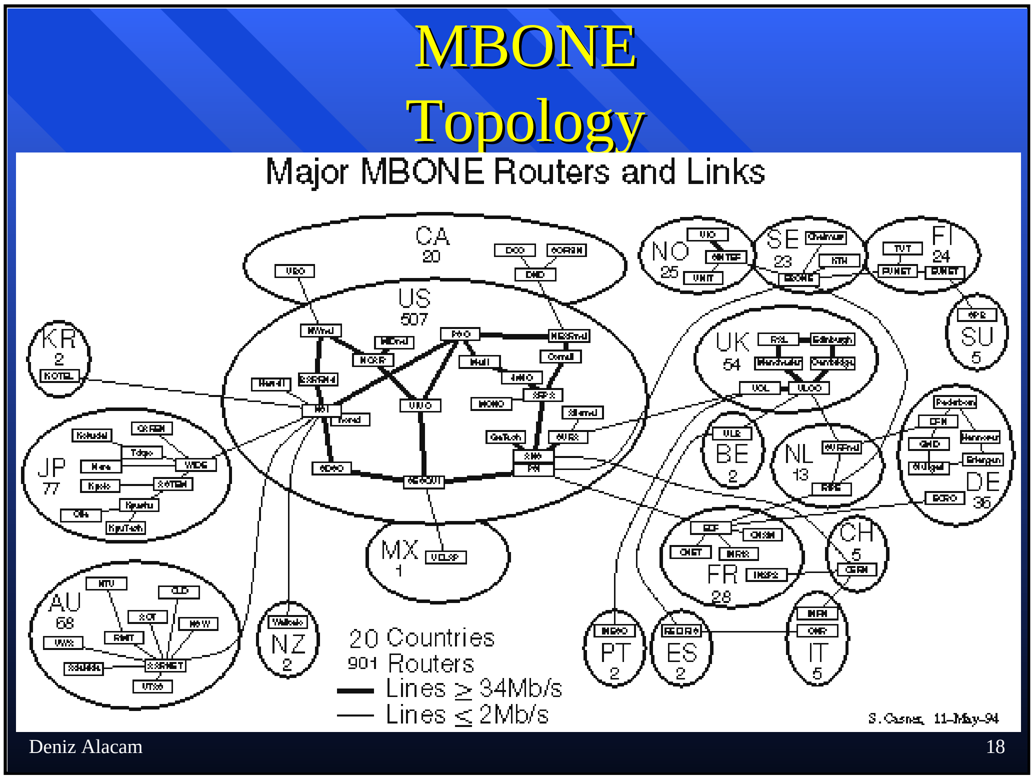
MBONE

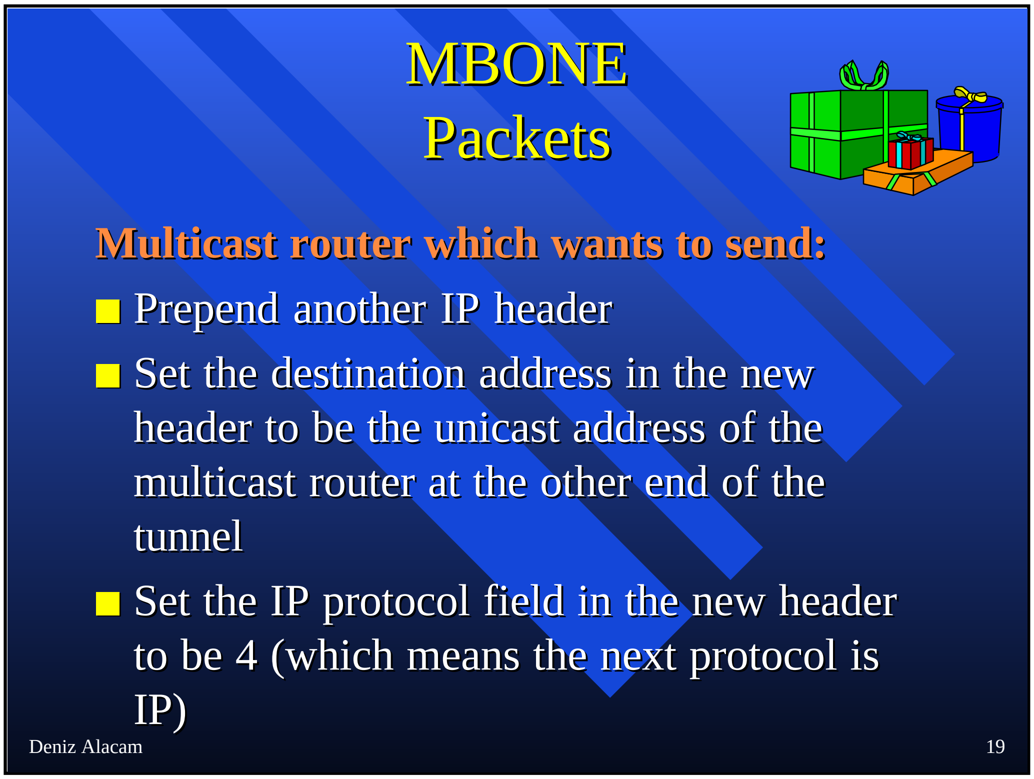
- The MBONE is an outgrowth of the first two IETF "audiocast" experiments in which live audio and video were multicast from the IETF meeting site to destinations around the world.
- Videoconferencing, audioconferencing, shared collaborative workspaces
- About 1,700 networks (in about 20 countries) are on the MBONE



Within a continent, the MBONE topology will be a combination of mesh and star: the backbone and regional (or mid-level) networks will be linked by a mesh of tunnels among mrouted machines located primarily at interconnection points of the backbones and regionals. Some redundant tunnels may be configured with higher metrics for robustness. Then each regional network will have a star hierarchy hanging off its node of the mesh to fan out and connect to all the customer networks that want to participate.

Between continents there will probably be only one or two tunnels, preferably terminating at the closest point on the MBONE mesh.


MBONE
Topology
Major MBONE Routers and Links
CA 20
US 507
KR 2
JP 77
AU 58
NZ 2
MX 1
NO 25
SE 23
FI 24
SU 5
UK 54
BE 2
NL 13
DE 35
CH 5
FR 28
PT 2
ES 2
IT 5
20 Countries
901 Routers
Lines ≥ 34Mb/s
Lines ≤ 2Mb/s
S. Casner, 11-May-94
Deniz Alacam
18

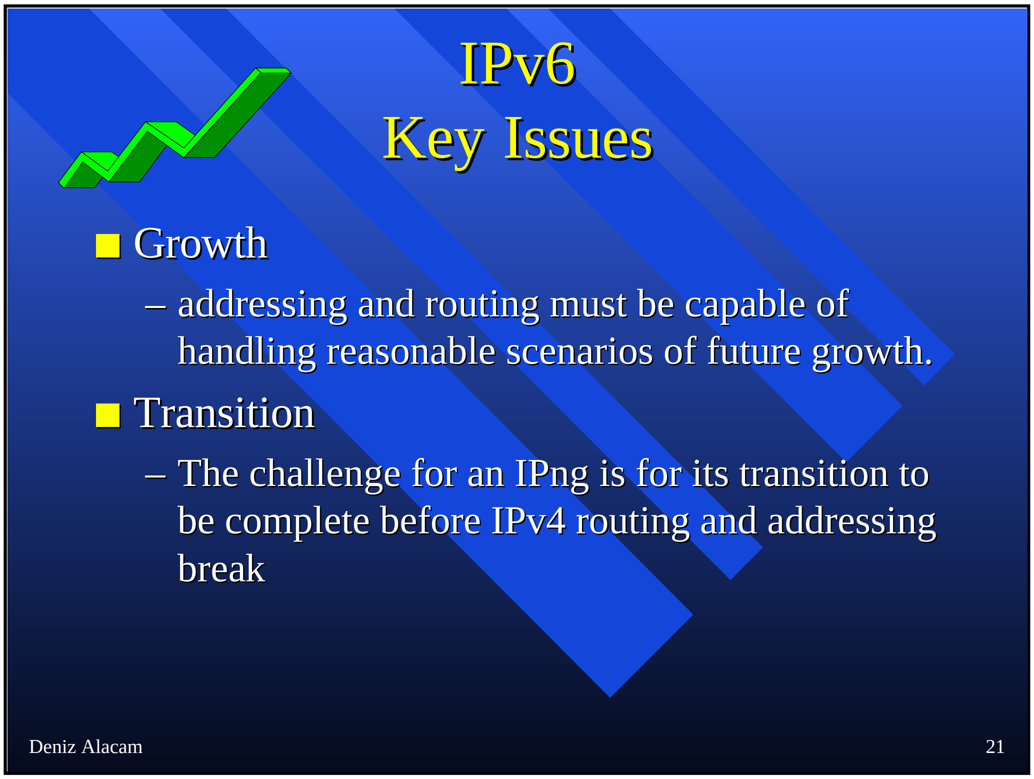


IP multicast packets are encapsulated for transmission through tunnels, so that they look like normal unicast datagrams to intervening routers and subnets. A multicast router that wants to send a multicast packet across a tunnel will prepend another IP header, set the destination address in the new header to be the unicast address of the multicast router at the other end of the tunnel, and set the IP protocol field in the new header to be 4 (which means the next protocol is IP). The multicast router at the other end of the tunnel receives the packet, strips off the encapsulating IP header, and forwards the packet as appropriate.

MBONE
Packets

Multicast router at the other end of the Tunnel:

- strip off the encapsulating IP header
- forward the packet as appropriate

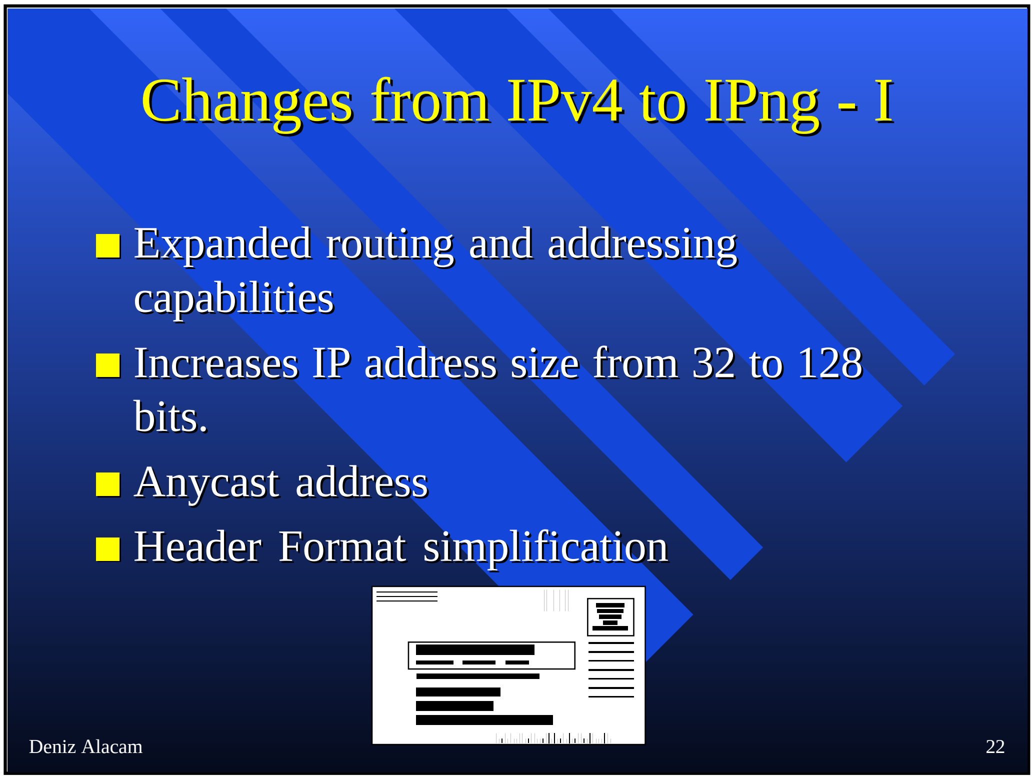


IPng is a new version of IP which is designed to be an evolutionary step from IPv4. It is a natural increment to IPv4. It can be installed as a normal software upgrade in Internet devices and is interoperable with the current IPv4. Its deployment strategy is designed to not have any flag days or other dependencies. IPng is designed to run well on high performance networks (e.g. ATM) and at the same time is still efficient for low bandwidth networks (e.g. wireless). In addition, it provides a platform for new Internet functionality that will be required in the near future.

Growth is the basic issue which caused there to be a need for a next generation IP. If anything is to be learned from our experience with IPv4 it is that the addressing and routing must be capable of handling reasonable scenarios of future growth.

The challenge the IETF faced in the selection of an IPng is to pick a protocol which meets today's requirements and also matches the requirements of emerging markets. There is need for a common protocol which can work over a variety of physical networks. (Some emerging markets: networked entertainment, device control) IPv4 address do not provide enough flexibility to construct efficient hierarchies which can be aggregated.

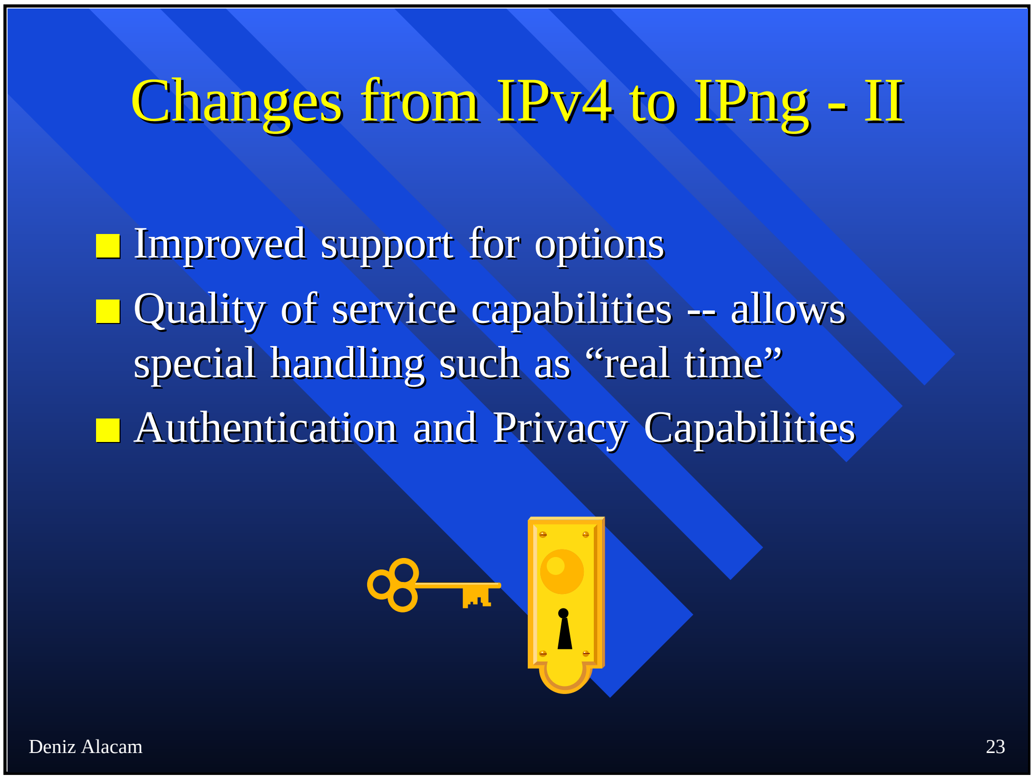
The challenge for an IPng is for its transition to be complete before IPv4 routing and addressing break. The transition will be much easier if IPv4 address are still globally unique. The two transition requirements which are the most important are flexibility of deployment and the ability for IPv4 hosts to communicate with IPng hosts. The capability must exist for IPng-only hosts to communicate with IPv4-only hosts globally while IPv4 addresses are globally unique.



•Expanded Routing and Addressing Capabilities IPng increases the IP address size from 32 bits to 128 bits, to support more levels of addressing hierarchy and a much greater number of addressable nodes, and simpler auto-configuration of addresses. The scalability of multicast routing is improved by adding a "scope" field to multicast addresses.

•A new type of address called a "anycast address" is defined, to identify sets of nodes where a packet sent to an anycast address is delivered to one of the nodes. The use of anycast addresses in the IPng source route allows nodes to control the path which their traffic flows.

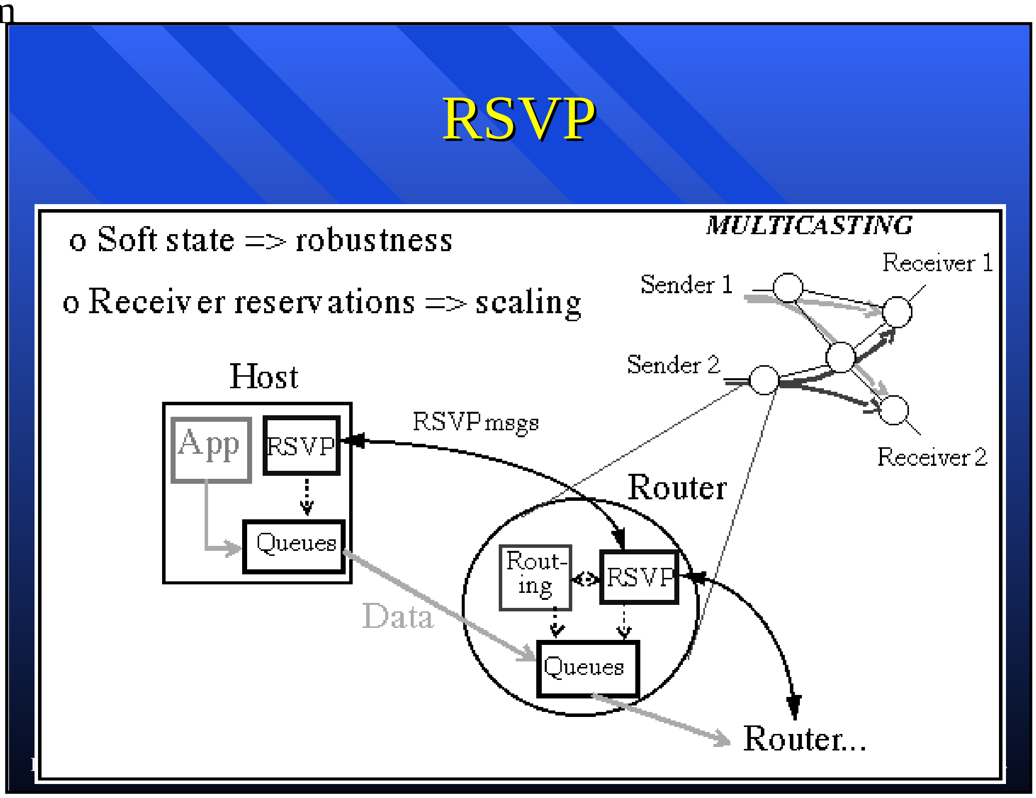
•Header Format Simplification Some IPv4 header fields have been dropped or made optional, to reduce the common-case processing cost of packet handling and to keep the bandwidth cost of the IPng header as low as possible despite the increased size of the addresses. Even though the IPng addresses are four time longer than the IPv4 addresses, the IPng header is only twice the size of the IPv4 header.



•Improved Support for Options Changes in the way IP header options are encoded allows for more efficient forwarding, less stringent limits on the length of options, and greater flexibility for introducing new options in the future.

•Quality-of-Service Capabilities A new capability is added to enable the labeling of packets belonging to particular traffic "flows" for which the sender requests special handling, such as non-default quality of service or "real- time" service.

•Authentication and Privacy Capabilities IPng includes the definition of extensions which provide support for authentication, data integrity, and confidentiality. This is included as a basic element of IPng and will be included in all implementations.



RSVP ReServation Protocolis part of a new Internet architecture that is designed to provide multipoint communication with both real-time and best-effort service. In addition to RSVP, the architecture includes:

• a multicast routing protocol that establishes the paths that data will take,

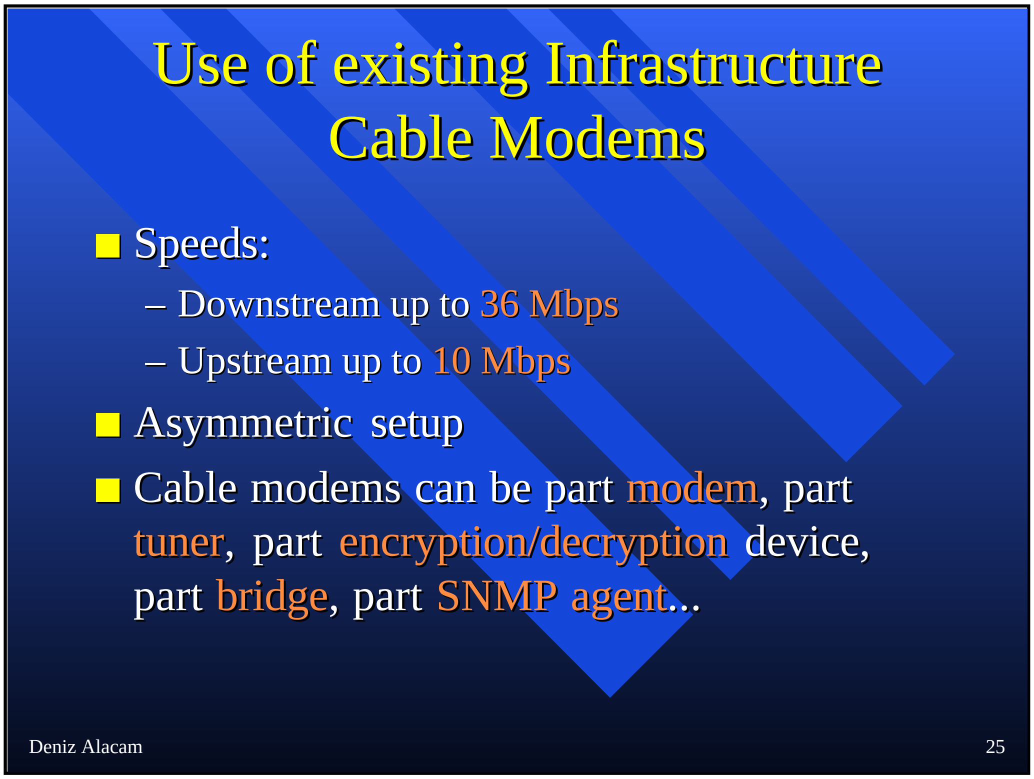
• an admission control module to manage network resources, and

• a packet scheduler that orders packet transmission to meet the quality of service established by admission control

RSVP carries reservation requests to the admission control module for each link of a path provided by the multicast routing protocol.

RSVP is a network control protocol that will allow Internet applications to obtain special qualities-of-service (QoS's) for their data flows. This will generally (but not necessarily) require reserving resources along the data path(s). RSVP is a component of the future "integrated services" Internet, which will provide both best-effort and realtime qualities of service

Each router in the path capable of resource reservation will pass incoming data packets to a Packet Classifier and then queue them in a Packet Scheduler. The Packet Classifier determines the route and the QoS class for each packet. The Scheduler allocates a particular outgoing link for packet transmission (it may also allocate other system resources (e.g., CPU time and/or buffers). A QoS request originating in an application is passed to the RSVP implementation (shown as a user daemon in the figure above) and RSVP passes the request to all the nodes along the reverse data path(s) to the destination(s). At each node, RSVP applies a local decision procedure called admission control to the QoS request, and if admission control succeeds, it sets the parameters to the Classifier and Packet Scheduler to obtain the desired QoS. If admission control fails at any node, RSVP returns an error indication to the application.



A cable modem is a device that allows high speed data access (for example to Internet) via a cable TV network.

Speeds vary widely. In the downstream direction (from the network to the computer), speeds can be anywhere up to ~36 Mbps. (Modems this fast are not currently on the market.) Few computers will be able of connecting at such high speeds, so a more realistic number is 3-10 Mbps. In the upstream direction (from computer to network), speeds can be up to 10 Mbps.

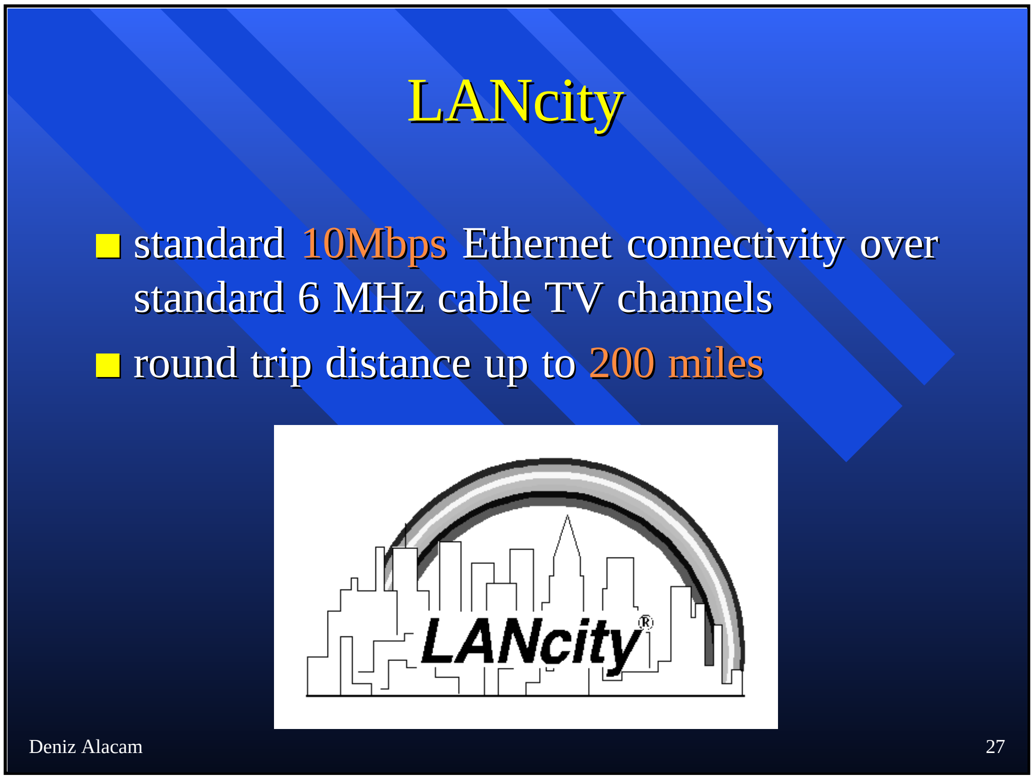
In the first few years an asymmetric setup will probably be more common than a symmetric setup. One reason for that the current Internet applications tend to be asymmetric in nature.

Cable modems are more complicated then their telephone counterparts. Cable modems can be part modem, part tuner, part encryption/decryption device, part bridge, part router, part SNMP agent, part Ethernet hub.

Use of existing Infrastructure
Cable Modems

- LANcity
- Digital -- ChannelWorks
- Hybrid Access System

Deniz Alacam 26



Lancity products provide standard 10 Mbps Ethernet data service connection over standard 6 MHz cable TV channels at round trip distances of up to 200 miles through the cable TV headend.

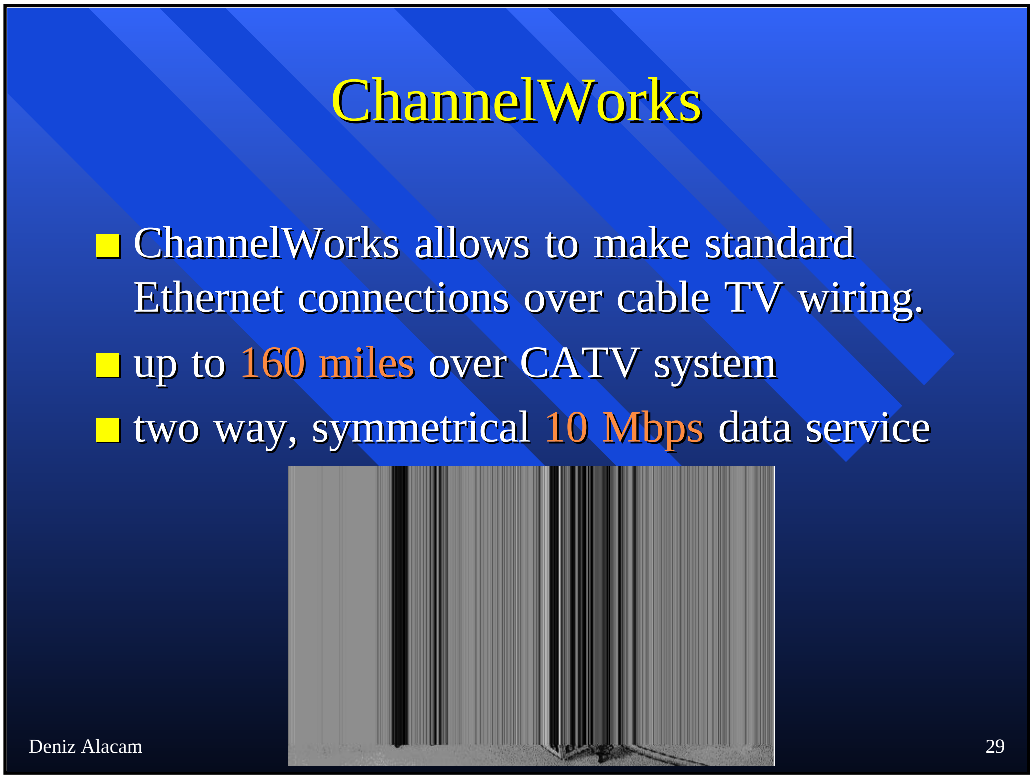
LANcity products are frequency agile, SNMP-network-managed, and coexist with a variety of other services that may be running on the cable TV network.

LANcity’s data over cable TV product families offer a total solution for city-wide connectivity. LANcity’s Personal product line provides high-speed, long-distance and low-cost connectivity to home computer users desiring access to other personal computers, to corporate offices, on-line services and the Internet. LCP, a personal cable TV modem unveiled in 1995, represents the company’s flagship Personal product in the series. LANcity’s Ethernet-to-cable-TV Business solutions employed worldwide consist of an Ethernet-to-cable-TV bridge (LCb); an Internet router (LCr); and SNMP managed workgroup (LCw). Three products comprise LANcity’s cable TV Headend family: an RF frequency translator (LCT); a headend reference node (LCh); and a network server (LCn).

LANcity Products

- Personal
  - – Personal Cable TV Modem (LCP)
- Business
  - – SNMP-Managed Ethernet Bridge to Cable TV (LCB)
  - – SNMP-Managed IP Router to Cable TV (LCR)

Deniz Alacam 28



Provides cost effective, high speed Internet access

Complies with computer industry standards for Ethernet, TCP/IP, and SNMP, and with cable TV industry standards

A ChannelWorks network, which can include ChannelWorks Bridges, Internet Brouters, and future ChannelWorks products; uses only one forward and one reverse 6 MHz channel