Advances in Network Architecture

Celina Hardaway

First Edition, 2012

ISBN 978-81-323-4603-6

Published by:
The English Press
4735/22 Prakashdeep Bldg,
Ansari Road, Darya Ganj,
Delhi - 110002
Email: info@wtbooks.com

Table of Contents

Introduction

Network architecture is the design of a communications network. It is a framework for the specification of a network's physical components and their functional organization and configuration, its operational principles and procedures, as well as data formats used in its operation.

In telecommunication, the specification of a network architecture may also include a detailed description of products and services delivered via a communications network, as well as detailed rate and billing structures under which services are compensated.

The network architecture of the Internet is predominantly expressed by its use of the Internet Protocol Suite, rather than a specific model for interconnecting networks or nodes in the network, or the usage of specific types of hardware links.

OSI Network Model

The Open Systems Interconnection model (OSI model) is a product of the Open Systems Interconnection effort at the International Organization for Standardization. It is a way of sub-dividing a communications system into smaller parts called layers. A layer is a collection of similar functions that provide services to the layer above it and receives services from the layer below it. On each layer, an instance provides services to the instances at the layer above and requests service from the layer below.

Physical Layer

The Physical Layer defines the electrical and physical specifications for devices. In particular, it defines the relationship between a device and a transmission medium, such as a copper or optical cable. This includes the layout of pins, voltages, cable specifications, hubs, repeaters, network adapters, host bus adapters (HBA used in storage area networks) and more. Its main task is the transmission of a stream of bits over a communication channel.

Data Linking Layer

The Data Link Layer provides the functional and procedural means to transfer data between network entities and to detect and possibly correct errors that may occur in the Physical Layer. Originally, this layer was intended for point-to-point and point-to-multipoint media, characteristic of wide area media in the telephone system. Local area network architecture, which included broadcast-capable multiaccess media, was developed independently of the ISO work in IEEE Project 802. IEEE work assumed sublayering and management functions not required for WAN use. In modern practice, only error detection, not flow control using sliding window, is present in data link protocols such as Point-to-Point Protocol (PPP), and, on local area networks, the IEEE 802.2 LLC layer is not used for most protocols on the Ethernet, and on other local area networks, its flow control and acknowledgment mechanisms are rarely used. Sliding window flow control and acknowledgment is used at the Transport Layer by protocols such as TCP, but is still used in niches where X.25 offers performance advantages. Simply it's main job is to create and recognize the frame boundary. This can be done by attaching special bit patterns to the beginning and the end of the frame. The input data is broken up into frames.

Network Layer

The Network Layer provides the functional and procedural means of transferring variable length data sequences from a source host on one network to a destination host on a different network, while maintaining the quality of service requested by the Transport Layer (in contrast to the data link layer which connects hosts within the same network). The Network Layer performs network routing functions, and might also perform fragmentation and reassembly, and report delivery errors. Routers operate at this layer—sending data throughout the extended network and making the Internet possible. This is a logical addressing scheme – values are chosen by the network engineer. The addressing scheme is not hierarchical. It controls the operation of the subnet and determine the routing strategies between IMP and insures that all the packs are correctly received at the destination in the proper order.

Transport Layer

The Transport Layer provides transparent transfer of data between end users, providing reliable data transfer services to the upper layers. The Transport Layer controls the reliability of a given link through flow control, segmentation/desegmentation, and error control. Some protocols are state and connection oriented. This means that the Transport Layer can keep track of the segments and retransmit those that fail. The Transport layer also provides the acknowledgement of the successful data transmission and sends the next data if no errors occurred. Some Transport Layer protocols, for example TCP, but not UDP, support virtual circuits provideconnection oriented communication over an underlying packet oriented datagram network. Where it assures the delivery of packets in the order in which they were sent and assure that they are free of errors. The datagram transportation deliver the packets randomly and broadcast it to multiple nodes. Notes: The transport layer multiplexes several streams on to 1 physical channel. The transport headers tells which message belongs to which connnection.

The Session Layer

This Layer provide a user interface to the network where the user negotiate to establish a connection, the user must provide the remote address in with he want to contact. The operation of setting up a session between 2 process is called "Binding" in some protocols it is merged with the transport layer.

Presentation Layer

The Presentation Layer establishes context between Application Layer entities, in which the higher-layer entities may use different syntax and semantics if the presentation service provides a mapping between them. If a mapping is available, presentation service data units are encapsulated into session protocol data units, and passed down the stack.This layer provides independence from data representation (e.g., encryption) by translating between application and network formats. The presentation layer transforms data into the form that the application accepts. This layer formats and encrypts data to be sent across a network. It is sometimes called the syntax layer. The original presentation structure used the basic encoding rules of Abstract Syntax Notation One (ASN.1), with capabilities such as converting an EBCDIC-coded text file to an ASCII-coded file, or serialization of objects and other data structures from and to XML.

Application Layer

The Application Layer is the OSI layer closest to the end user, which means that both the OSI application layer and the user interact directly with the software application. This layer interacts with software applications that implement a communicating component. Such application programs fall outside the scope of the OSI model. Application layer functions typically include identifying communication partners, determining resource availability, and synchronizing communication. When identifying communication partners, the application layer determines the identity and availability of communication partners for an application with data to transmit.

Distributed computing

In distinct usage in distributed computing, the term *network architecture* often describes the structure and classification of a distributed application architecture, as the participating nodes in a distributed application are often referred to as a *network*. For example, the applications architecture of the public switched telephone network (PSTN) has been termed the Advanced Intelligent Network. There are any number of specific classifications but all lie on a continuum between the dumb network (e.g., Internet) and the intelligent computer network (e.g., the telephone network). Other networks contain various elements of these two classical types to make them suitable for various types of applications. Recently the context aware network, which is a synthesis of two, has gained much interest with its ability to combine the best elements of both.

A popular example of such usage of the term in distributed applications, as well as PVCs (permanent virtual circuits), is the organization of nodes in peer-to-peer (P2P) services and networks. P2P networks usually implement overlay networks running over an underlying physical or logical network. These overlay network may implement certain organizational structures of the nodes according to several distinct models, the network architecture of the system.

Network architecture is a broad plan that specifies everything necessary for two application programs on different networks on an Internet to be able to work together effectively.

Chapter- 1

Passive Optical Network

A **passive optical network** (**PON**) is a point-to-multipoint, fiber to the premises network architecture in which unpowered optical splitters utilizing Brewster's angle principles are used to enable a single optical fiber to serve multiple premises, typically 32-128. A PON consists of an optical line terminal (OLT) at the service provider's central office and a number of optical network units (ONUs) near end users. A PON configuration reduces the amount of fiber and central office equipment required compared with point to point architectures. A passive optical network is a form of fiber-optic access network.

Downstream signals are broadcast to each premises sharing a single fiber. Encryption is used to prevent eavesdropping.

Upstream signals are combined using a multiple access protocol, usually time division multiple access (TDMA). The OLTs "range" the ONUs in order to provide time slot assignments for upstream communication.

Standards

- IEEE 802.3
    - **EPON** (Ethernet PON) is part of IEEE standard Ethernet with options for 1/1 Gbit/s 10/1 Gbit/s and 10/10 Gbit/s. There are currently over 40 million installed EPON ports making it the most widely deployed PON technology globally. EPON is also the foundation for cable operators business services as part of the DOCSIS Provisioning of EPON (DPoE) specifications.

- ITU-T
    - G.983
        - **APON** (ATM PON). This was the first Passive optical network standard. It was used primarily for business applications, and was based on ATM.
        - **BPON** (Broadband PON) is a standard based on APON. It adds support for WDM, dynamic and higher upstream bandwidth allocation, and survivability. It also created a standard management interface, called OMCI, between the OLT and ONU/ONT, enabling mixed-vendor networks.

    - G.984
        - **G-PON** (Gigabit PON) is an evolution of the BPON standard. It supports higher rates, enhanced security, and choice of Layer 2 protocol (ATM, GEM, Ethernet). By mid-2008, Verizon had installed over 800 thousand lines. British Telecom, Mobily-SaudiArabia, Etisalat-UAE, and AT&T are in advanced trials. It is the successor to G.983.
    - G.987
        - **10G-PON** has 10 Gbit/s downstream and 2.5 Gbit/s upstream – framing is "G-PON like" and designed to coexist with GPON devices on the same network.
- SCTE IPS910
    - **RFoG** (RFoverGlass) is an SCTE Interface Practices Subcomittee standard in development for carrying HFC RF signals over a passive optical Network (PON).

History

Early work on efficient fiber to the home architectures was done in the 1990s by the Full Service Access Network (FSAN) working group, formed by major telecommunications service providers and system vendors. The International Telecommunications Union (ITU) did further work, and has since standardized on two generations of PON. The older ITU-T G.983 standard is based on Asynchronous Transfer Mode (ATM), and has therefore been referred to as APON (ATM PON). Further improvements to the original APON standard – as well as the gradual falling out of favor of ATM as a protocol – led to the full, final version of ITU-T G.983 being referred to more often as broadband PON, or BPON. A typical APON/BPON provides 622 megabits per second (Mbit/s) (OC-12) of downstream bandwidth and 155 Mbit/s (OC-3) of upstream traffic, although the standard accommodates higher rates.

The ITU-T G.984 (GPON) standard represents a boost, compared to BPON, in both the total bandwidth and bandwidth efficiency through the use of larger, variable-length packets. Again, the standards permit several choices of bit rate, but the industry has converged on 2.488 gigabits per second (Gbit/s) of downstream bandwidth, and 1.244 Gbit/s of upstream bandwidth. GPON Encapsulation Method (GEM) allows very efficient packaging of user traffic with frame segmentation.

The IEEE 802.3 Ethernet PON (EPON or GEPON) standard was completed in 2004, as part of the Ethernet First Mile project. EPON uses standard 802.3 Ethernet frames with symmetric 1 gigabit per second upstream and downstream rates. EPON is applicable for data-centric networks, as well as full-service voice, data and video networks. 10Gbit/s EPON or 10G-EPON was ratified as an amendment IEEE 802.3av to IEEE 802.3. 10G-EPON supports 10/1 Gbit/s. The downstream wavelength plan support simultaneous operation of 10 Gbit/s on one wavelength and 1 Gbit/s on a separate wavelength for operation of IEEE 802.3av and IEEE 802.3ah on the same PON concurrently. The upstream channel can support simultaneous operation of IEEE 802.3av and 1 Gbit/s 802.3ah simultaneously on a single shared (1,310 nm) channel.

Network elements

A PON takes advantage of wavelength division multiplexing (WDM), using one wavelength for downstream traffic and another for upstream traffic on a single nondispersion-shifted fiber (ITU-T G.652). BPON, EPON, GEPON, and GPON have the same basic wavelength plan and use the 1,490 nanometer (nm) wavelength for downstream traffic and 1310 nm wavelength for upstream traffic. 1550 nm is reserved for optional overlay services, typically RF (analog) video.

As with bit rate, the standards describe several optical budgets, most common is 28 dB of loss budget for both BPON and GPON, but products have been announced using less expensive optics as well. 28 dB corresponds to about 20 km with a 32-way split. Forward error correction (FEC) may provide another 2–3 dB of loss budget on GPON systems. As optics improve, the 28 dB budget will likely increase. Although both the GPON and EPON protocols permit large split ratios (up to 128 subscribers for GPON, up to 32,768 for EPON), in practice most PONs are deployed with a split ratio of 1x32 or smaller.

A PON consists of a central office node, called an optical line terminal (OLT), one or more user nodes, called optical network units (ONUs) or optical network terminals (ONTs), and the fibers and splitters between them, called the optical distribution network (ODN). ONT is an ITU-T term to describe a special, single-user case of an ONU. In Multiple Tenant Units, the ONU may be bridged to a customer premise device within the individual dwelling unit using technologies such as Ethernet over twisted pair, G.hn (a high-speed ITU-T standard that can operate over any existing home wiring - power lines, phone lines and coaxial cables) or DSL. An ONU is a device that terminates the PON and presents customer service interfaces to the user. Some ONUs implement a separate subscriber unit to provide services such as telephony, Ethernet data, or video.

The OLT provides the interface between the PON and the service providers network services. These typically include:

- Internet Protocol (IP) traffic over gigabit/s, 10 Gbit/s, or 100 Mbit/s Ethernet
- standard time division multiplexed (TDM) interfaces such as SONET or SDH
- ATM UNI at 155–622 Mbit/s

The ONT or ONU terminates the PON and presents the native service interfaces to the user. These services can include voice (plain old telephone service (POTS) or voice over IP (VoIP)), data (typically Ethernet or V.35), video, and/or telemetry (TTL, ECL, RS530, etc.). Often, the ONU functions are separated into two parts:

- the ONU, which terminates the PON and presents a converged interface – such as xDSL, coax, or multiservice Ethernet – toward the user, and

- network termination equipment (NTE), which provides the separate, native service interfaces directly to the user

A PON is a shared network, in that the OLT sends a single stream of downstream traffic that is seen by all ONUs. Each ONU only reads the content of those packets that are addressed to it. Encryption is used to prevent eavesdropping on downstream traffic.

Upstream bandwidth allocation

The OLT is responsible for allocating upstream bandwidth to the ONUs. Because the optical distribution network (ODN) is shared, ONU upstream transmissions could collide if they were transmitted at random times. ONUs can lie at varying distances from the OLT, meaning that the transmission delay from each ONU is unique. The OLT measures delay and sets a register in each ONU via PLOAM (physical layer operations and maintenance) messages to equalize its delay with respect to all of the other ONUs on the PON.

Once the delay of all ONUs has been set, the OLT transmits so-called grants to the individual ONUs. A grant is permission to use a defined interval of time for upstream transmission. The grant map is dynamically re-calculated every few milliseconds. The map allocates bandwidth to all ONUs, such that each ONU receives timely bandwidth for its service needs.

Some services – POTS, for example – require essentially constant upstream bandwidth, and the OLT may provide a fixed bandwidth allocation to each such service that has been provisioned. DS1 and some classes of data service may also require constant upstream bit rate. But much data traffic – internet surfing, for example – is bursty and highly variable. Through dynamic bandwidth allocation (DBA), a PON can be oversubscribed for upstream traffic, according to the traffic engineering concepts of statistical multiplexing. (Downstream traffic can also be oversubscribed, in the same way that any LAN can be oversubscribed. The only special feature in the PON architecture for downstream oversubscription is the fact that the ONU must be able to accept completely arbitrary downstream time slots, both in time and in size.)

In GPON there are two forms of DBA, status-reporting (SR) and non-status reporting (NSR).

In NSR DBA, the OLT continuously allocates a small amount of extra bandwidth to each ONU. If the ONU has no traffic to send, it transmits idle frames during its excess allocation. If the OLT observes that a given ONU is not sending idle frames, it increases the bandwidth allocation to that ONU. Once the ONU's burst has been transferred, the OLT observes a large number of idle frames from the given ONU, and reduces its allocation accordingly. NSR DBA has the advantage that it imposes no requirements on the ONU, and the disadvantage that there is no way for the OLT to know how best to assign bandwidth across several ONUs that need more.

In SR DBA, the OLT polls ONUs for their backlogs. A given ONU may have several so-called traffic containers (T-CONTs), each with its own priority or traffic class. The ONU reports each T-CONT separately to the OLT. The report message contains a logarithmic measure of the backlog in the T-CONT queue. By knowledge of the service level agreement for each T-CONT across the entire PON, as well as the size of each T-CONT's backlog, the OLT can optimize allocation of the spare bandwidth on the PON.

EPON systems use a DBA mechanism equivalent to GPON's SR DBA solution. The OLT polls ONUs for their queue status and grants bandwidth using the MPCP GATE message, while ONUs report their status using the MPCP REPORT message.

Current status

TDM-PON

Both APON/BPON and EPON/GEPON have been deployed widely, but most networks designed in 2008 use GPON or GEPON. GPON has fewer than 2 million installed ports. GEPON has approximately 30 million deployed ports. For TDM-PON, a passive power splitter is used as the remote terminal. Each ONUs (Optical network units) signals are multiplexed in the time domain. ONUs see their own data through the address labels embedded in the signal.

DOCSIS Provisioning of EPON or DPoE

Data Over Cable Service Interface Specification (DOCSIS) Provisioning of Ethernet Passive Optical Network, or DPoE, is a set of Cable Television Laboratory specifications that implement the DOCSIS service layer interface on existing Ethernet PON (EPON, GEPON or 10G-EPON) Media Access Control (MAC) and Physical layer (PHY) standards. In short it implements the DOCSIS Operations Administration Maintenance and Provisioning (OAMP) functionality on existing EPON equipment. It makes the EPON OLT look and act like a DOCSIS Cable Modem Termination Systems (CMTS) platform (which is called a DPoE System in DPoE terminology). In addition to the offering the same IP service capabilities as a CMTS, DPoE supports Metro Ethernet Forum (MEF) 9 and 14 services for the delivery of Ethernet services for business customers.

RFoG

Radio Frequency over Glass (RFoG) is a type of passive optical networking, that transports RF signals that are now transported over copper (principally over a hybrid fiber and coaxial cable) over PON. In the forward direction RFoGis either a stand alone P2MP system or an optical overlay for existing PON such as GEPON/EPON. The overlay for RFoG is based on Wave Division Multiplexing (WDM) -- the passive combination of wavelengths on a single strand of glass. Reverse RF support is provided by transporting the upstream or return RF into on a separate lambda from the PON return wavelength. The Society of Cable and Telecommunications Engineers (SCTE) Interface

Practices Subcomittee (IPS) Work Group 5, is currently working on IPS 910 RF over Glass. RFoG offers backwards compatibility with existing RF modulation technology, but offers no additional bandwidth for RF based services. Although not yet completed, the RFoG standard is actually a collection of standardized options which are not compatible with each other (they cannot be mixed on the same PON). Some of the standards may interoperate with other PONs, others may not. It offers a means to support RF technologies in locations where only fiber is available or where copper is not permitted or feasible. This technology is targeted towards Cable TV operators and their existing HFC networks. Some describe RFoG as "all of the cost of fiber to the home with none of the benefits."

WDM-PON

Wavelength Division Multiplexing PON, or WDM-PON, is a non-standard type of passive optical networking, being developed by some companies.

The multiple wavelengths of a WDM-PON can be used to separate Optical Network Units (ONUs) into several virtual PONs co-existing on the same physical infrastructure. Alternatively the wavelengths can be used collectively through statistical multiplexing to provide efficient wavelength utilization and lower delays experienced by the ONUs.

There is no common standard for WDM-PON nor any unanimously agreed upon definition of the term. By some definitions WDM-PON is a dedicated wavelength for each ONU. Other more liberal definitions suggest the use of more than one wavelength in any one direction on a PON is WDM-PON. It is difficult to point to an un-biased list of WDM-PON vendors when there is no such unanimous definition. PONs provide higher bandwidth than traditional copper based access networks. WDM-PON has better privacy and better scalability because of each ONU only receives its own wavelength.

Advantages: The MAC layer is simplified because the P2P connections between OLT and ONUs are realized in wavelength domain, so no P2MP media access control is needed. In WDM-PON each wavelength can run at a different speed and protocol so there is a easy pay-as-you-grow upgrade

Challenges: High cost of initial set-up, the cost of the WDM components. Temperature control is another challenge because of how wavelengths tend to drift with environmental temperatures.

Long-Reach Optical Access Networks

The concept of the Long-Reach Optical Access Network (LROAN) is to replace the optical/electrical/optical conversion that takes place at the local exchange with a continuous optical path that extends from the customer to the core of the network. Work by Davey and Payne at BT showed that significant cost savings could be made by reducing the electronic equipment and real-estate required at the local exchange or wire center. A proof of concept demonstrator showed that it was possible to serve 1024 at 10GBit/s with 100km reach.

This technology has sometimes been termed Long-Reach PON, however, many argue that the term PON is no longer applicable as, in most instances, only the distribution remains passive.

Enabling technologies

Due to the topology of PON, the transmission modes for downstream (i.e., from OLT to ONU) and upstream (i.e., from ONU to OLT) are different. For the downstream transmission, the OLT broadcasts optical signal to all the ONUs in continuous mode (CM), i.e., the downstream channel always has optical data signal. However, in the upstream channel, ONUs can not transmit optical data signal in CM. Use of CM would result in all of the signals transmitted from the ONUs converging (with attenuation) into one fiber by the power splitter (serving as power coupler), and overlapping. To solve this problem, burst mode (BM) transmission is adopted for upstream channel. The given ONU only transmits optical packet when it is allocated a time slot and it needs to transmit, and all the ONUs share the upstream channel in the time division multiplexing (TDM) mode. The phases of the BM optical packets received by the OLT are different from packet to packet, since the ONUs are not synchronized to transmit optical packet in the same phase, and the distance between OLT and given ONU are random. Since the distance between the OLT and ONUs are not uniform, the optical packets received by the OLT may have different amplitudes. In order to compensate the phase variation and amplitude variation in a short time (e.g., within 40 ns for GPON), burst mode clock and data recovery (BM-CDR) and burst mode amplifier (e.g., burst mode TIA) need to be employed, respectively. Furthermore, the BM transmission mode requires the transmitter to work in burst mode. Such a burst mode transmitter is able to turn on and off in short time. The above three kinds of circuitries in PON are quite different from their counterparts in the point-to-point continuous mode optical communication link.

Fiber to the premises

Passive optical networks do not use electrically powered components to split the signal. Instead, the signal is distributed using beam splitters. Each splitter typically splits the signal from a single fiber into 16, 32, or 64 fibers, depending on the manufacturer, and several splitters can be aggregated in a single cabinet. A beam splitter cannot provide any switching or buffering capabilities; the resulting connection is called a point-to-multipoint link. For such a connection, the optical network terminals on the customer's end must perform some special functions which would not otherwise be required. For example, due to the absence of switching capabilities, each signal leaving the central office must be broadcast to all users served by that splitter (including to those for whom the signal is not intended). It is therefore up to the optical network terminal to filter out any signals intended for other customers. In addition, since beam splitters cannot perform buffering, each individual optical network terminal must be coordinated in a multiplexing scheme to prevent signals leaving the customer from colliding at the intersection. Two types of multiplexing are possible for achieving this: wavelength-division multiplexing and time-division multiplexing. With wavelength-division multiplexing, each customer transmits their signal using a unique wavelength. With time-division multiplexing (TDM), the customers "take turns" transmitting information. TDM equipment has been on the market longest; WDM-PON equipment became available in 2005.

Passive optical networks have both advantages and disadvantages over active networks. They avoid the complexities involved in keeping electronic equipment operating outdoors. They also allow for analog broadcasts, which can simplify the delivery of analog television. However, because each signal must be pushed out to *everyone* served by the splitter (rather than to just a single switching device), the central office must be equipped with a particularly powerful piece of transmitting equipment called an optical line terminal (OLT). In addition, because each customer's optical network terminal must transmit all the way to the central office (rather than to just the nearest switching device), customers can't be as far from the central office as is possible with active optical networks.

Passive optical components

The drivers behind the modern passive optical network are the optical components that enable Quality of Service (QoS).

Single-mode, *passive optical components* include branching devices such as *Wavelength-*Division Multiplexer/Demultiplexers– (WDMs)*, isolators, circulators, and filters. These components are used in interoffice, loop feeder, Fiber In The Loop (FITL), Hybrid Fiber-Coaxial Cable (HFC), Synchronous Optical Network (SONET), and Synchronous Digital Hierarchy (SDH) systems; and other* telecommunications networks employing optical communications systems that utilize Optical Fiber Amplifiers (OFAs) and Dense Wavelength Division Multiplexer (DWDM) systems. Industry proposed requirements for these components are detailed in GR-1209, *Generic Requirements for Passive Optical Components*.

The broad variety of passive optical components applications include multichannel transmission, distribution, optical taps for monitoring, pump combiners for fiber amplifiers, bit-rate limiters, optical connects, route diversity, polarization diversity, interferometers, and conherent communication.

WDMs are optical components in which power is split or combined based on the wavelength composition of the optical signal. *Dense Wavelength Division Multiplexers (DWDMs)* are optical components that split power over at least four wavelengths. *Wavelength insensitive couplers* are passive optical components in which power is split or combined independently of the wavelength composition of the optical signal. A given component may combine and divide optical signals simultaneously, as in bidirectional (duplex) transmission over a single fiber. Passive optical components are data format transparent, combining and dividing optical power in some predetermined ratio (*coupling ratio*) regardless of the information content of the signals. WDMs can be thought of as *wavelength* splitters and combiners. Wavelength insensitive couplers can be thought of as *power* splitters and combiners.

An optical *isolator* is a two-port passive component that allows light (in a given wavelength range) to pass through with low attenuation in one direction, while isolating (providing a high attenuation for) light propagating in the reverse direction. Isolators are used as both integral and in-line components in laser diode modules and optical amplifiers, and to reduce noise caused by multi-path reflection in highbit-rate and analog transmission systems.

An optical *circulator* operates in a similar way to an optical isolator, except that the reverse propagating lightwave is directed to a third port for output, instead of being lost. An optical circulator can be used for bidirectional transmission, as a type of branching component that distributes (and isolates) optical power among fibers, based on the direction of the lightwave propagation.

A fiber optic *filter* is a component with two or more ports that provides wavelength sensitive loss, isolation and/or return loss. Fiber optic filters are in-line, wavelength selective, components that allow a specific range of wavelengths to pass through (or reflect) with low attenuation for classification of filter types).

GR-1221-CORE, *Generic Reliability Assurance Requirements for Passive Optical Components*, addresses the long-term reliability of passive optical components.

Chapter- 2

Ambient Network and Bus Network

Ambient network

Ambient Networks is a network integration design that seeks to solve problems relating to switching between networks to maintain contact with the outside world. This project aims to develop a network software-driven infrastructure that will run on top of all current or future network physical infrastructures to provide a way for devices to connect to each other, and through each other to the outside world.

The concept of Ambient Networks comes from the IST Ambient Network project, which is a research project sponsored by the European Commission within the Sixth Framework Programme (FP6).

The Ambient Networks Project

Ambient Networks is a large-scale collaborative project within the European Union's Sixth Framework Programme that investigates future communications systems beyond today's fixed and 3rd generation mobile networks. It is part of the Wireless World Initiative. The project works at a new concept called Ambient Networking, to provide suitable mobile networking technology for the future mobile and wireless communications environment. Ambient Networks aims to provide a unified networking concept that can adapt to the very heterogeneous environment of different radio technologies and service and network environments. Special focus is put on facilitating both competition and cooperation of various market players by defining interfaces, which allow the instant negotiation of agreements. This approach goes clearly beyond interworking of well-defined protocols and is expected to have a long-term effect on the business landscape in the Wireless World. Central to the project is the concept of composition of networks, which is an approach to address the dynamic nature of the target environment. The approach is based on an open framework for network control functionality, which can be extended with new capabilities as well as operating over existing connectivity infrastructure.

- **Phase 1** of the project (2004-2005) has laid the conceptual foundations. The Deliverable D1-5 "Ambient Networks Framework Architecture" summarizes the work from phase 1 and provides links to other relevant material.

- Ambient Networks **Phase 2** (2006-2007) focuses on validation aspects. One key result of phase 2 is an integrated prototype that will be used to study the feasibility of the Ambient Networks concept for a number of typical network scenarios. The ACS prototype will be used to iteratively test the components developed by the project in a real implementation. In parallel, the top-down work is being continued which will lead to a refined System Specification. This document, referred to as the System Description, is available on the Ambient Networks website. Furthermore, standardization of the composition concept is addressed in 3GPP.

Interfaces and their use

The ACS (Ambient Control Space) is the internal of an Ambient Network. It has the functions that can be accessed and it is in full control of the resources of the network. The Ambient Networks infrastructure does not deal with nodes, instead it deals with networks, though at the beginning, all the "networks" might only consist of just one node: these "networks" need to merge to form a network in the original sense of the word. A composition establishment consists of the negotiation and then the realization of a Composition Agreement. This merging can happen be fully automatic. The decision to merge or not is decided using pre-configured policies.

There are three interfaces present to communicate with an ACS. These are:

- ANI: Ambient Network Interface. If a network wants to join in, it has to do so through this interface.
- ASI: Ambient Service Interface. If a function needs to be accessed inside the ACS, this Interface is used.
- ARI: Ambient Resource Interface. If a resource inside a network needs to be accessed (e.g. the volume of the traffic), this interface is used.

Interfaces are used to hide the internal structures of the underlying network.

If two networks meet, and decide to merge, a new ACS will be formed of the two (though the two networks will have their own ACS along with the interfaces inside this global, new ACS). The newly composed ACS will of course have its own ANI, ASI and ARI, and will use these interfaces to merge with other Ambient Networks. Other options for composition are to not merge the two Ambient Networks (Network Interworking) or to establish a new virtual ACS that exercises joint control over a given set of shared resources (Control Sharing).

ACS Functional Entities

Functions are divided into Functional Entities (FEs). The ACS provides a flexible and extensible framework to run these FEs as a distributed system. Examples are

- Composition Functional Entity: Controlling composition of ANs

- Bearer Management FEs
- Overlay Management FEs

More information on FEs is contained in the Ambient Networks Framework Architecture and the latest version of the System Description.

Example Situation

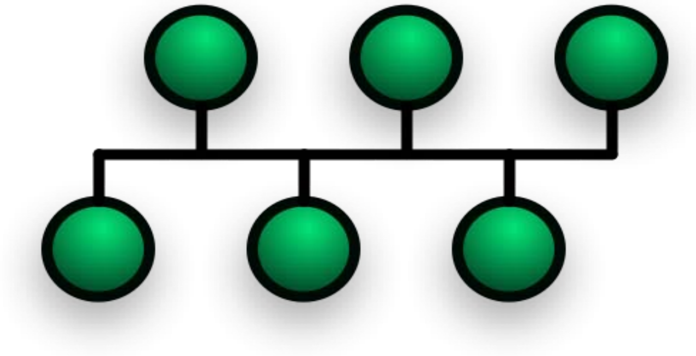
Alice has a PAN, a Personal Area Network on her body: she has a Bluetooth enabled PDA, mobile phone and laptop that she is carrying, and are all currently turned on, and forming a network. Her laptop also has the ability to connect using an available WLAN, and her mobile phone has the ability to connect through GPRS, though GPRS is slower and much more costly for Alice to use. She is now on the move, and her laptop is downloading her emails using the GPRS connection on the mobile:

Laptop -> (Bluetooth) -> Mobile -> (GPRS) -> Mobile phone network

While walking, she passes into an area covered by a free WLAN hotspot: Her PAN now immediately starts to initiate a connection with the hotspot. This is called "merging" of the networks (that of the hotspot and that of her PAN). Once this merging is complete, the downloading of her email continues totally unaffected, but instead of using the expensive and slow GPRS connection, it is now using the newly established WLAN connection. If she now wants to browse the web with her PDA, the PDA will also use the WLAN connection of the laptop:

PDA-> (bluetooth) -> Laptop-> (WLAN) -> Hotspot

Bus network



Bus network layout

A **bus network topology** is a network architecture in which a set of clients are connected via a shared communications line, called a bus. There are several common instances of the bus architecture, including one in the motherboard of most computers, and those in some versions of Ethernet networks.

How it works

Bus networks are the simplest way to connect multiple clients, but may have problems when two clients want to transmit at the same time on the same bus. Thus systems which use bus network architectures normally have some scheme of collision handling or collision avoidance for communication on the bus, quite often using Carrier Sense Multiple Access or the presence of a bus master which controls access to the shared bus resource.

A true bus network is passive – the computers on the bus simply listen for a signal; they are not responsible for moving the signal along. However, many active architectures can also be described as a "bus", as they provide the same logical functions as a passive bus; for example, switched Ethernet can still be regarded as a logical network, if not a physical one. Indeed, the hardware may be abstracted away completely in the case of a software bus.

With the dominance of switched Ethernet over passive Ethernet, passive bus networks are uncommon in wired networks. However, almost all current wireless networks can be viewed as examples of passive bus networks, with radio propagation serving as the shared passive medium.

The bus topology makes the addition of new devices straightforward. The term used to describe clients is station or workstation in this type of network. Bus network topology uses a broadcast channel which means that all attached stations can hear every transmission and all stations have equal priority in using the network to transmit data.

The Ethernet bus topology works like a big telephone party line — before any device can send a packet, devices on the bus must first determine that no other device is sending a packet on the cable. When a device sends its packet out over the bus, every other network card on the bus sees and reads the packet. Ethernet's scheme of having devices communicate like they were in chat room is called Carrier Sense Multiple Access/ Collision Detection (CSMA/CD). Sometimes two cards talk (send packets) at the same time. This creates a collision, and the cards themselves arbitrate to decide which one will resend its packet first. All PCs on a bus network share a common wire, which also means they share the data transfer capacity of that wire – or, in tech terms, they share its bandwidth.

This creates an interesting effect. Ten PCs chatting on a bus each get to use a much higher proportion of its total bandwidth than, for instance, 100 PCs on the same bus (in this case, one – tenth compared to one – hundredth). The more PCs on a bus, the more likely you'll have a communication traffic jam.

Advantages and disadvantages of a bus network

Advantages

- Easy to implement and extend.
- Easy to install.
- Well-suited for temporary or small networks not requiring high speeds (quick setup), resulting in faster networks.
- Cheaper than other topologies (But in recent years has became less important due to devices like a switch)
- Cost effective; only a single cable is used.
- Easy identification of cable faults.
- Reduced weight due to fewer wires.

Disadvantages

- Limited cable length and number of stations.
- If there is a problem with the cable, the entire network breaks down.
- Maintenance costs may be higher in the long run.
- Performance degrades as additional computers are added or on heavy traffic (shared bandwidth).
- Proper termination is required (loop must be in closed path).
- Significant Capacitive Load (each bus transaction must be able to stretch to most distant link).
- It works best with limited number of nodes.
- Commonly has a slower data transfer rate than other topologies.
- Only one packet can remain on the bus during one clock pulse.

Chapter- 3

Delay-tolerant Networking

Delay-tolerant networking (DTN) is an approach to computer network architecture that seeks to address the technical issues in heterogeneous networks that may lack continuous network connectivity. Examples of such networks are those operating in mobile or extreme terrestrial environments, or planned networks in space.

Recently, the term **disruption-tolerant networking** has gained currency in the United States due to support from DARPA, which has funded many DTN projects. Disruption may occur because of the limits of wireless radio range, sparsity of mobile nodes, energy resources, attack, and noise.

History

In the 1970s, spurred by the micronization of computing, researchers began developing technology for routing between non-fixed locations of computers. While the field of ad-hoc routing was inactive throughout the 1980s, the widespread use of wireless protocols reinvigorated the field in the 1990s as mobile ad-hoc networking (MANET) and vehicular ad-hoc networking became areas of increasing interest.

Concurrently with (but separate from) the MANET activities, DARPA had funded NASA, MITRE and others to develop a proposal for the Interplanetary Internet (IPN). Internet pioneer Vint Cerf and others developed the initial IPN architecture, relating to the necessity of networking technologies that can cope with the significant delays and packet corruption of deep-space communications. In 2002, Kevin Fall started to adapt some of the ideas in the IPN design to terrestrial networks and coined the term *delay-tolerant networking* and the DTN acronym. A paper published in 2003 SIGCOMM conference gives the motivation for DTNs. The mid-2000s brought about increased interest in DTNs, including a growing number of academic conferences on delay and disruption-tolerant networking, and growing interest in combining work from sensor networks and MANETs with the work on DTN. This field saw many optimizations on classic ad-hoc and delay-tolerant networking algorithms and began to examine factors such as security, reliability, verifiability, and other areas of research that are well understood in traditional computer networking.

Routing

The ability to transport, or route, data from a source to a destination is a fundamental ability all communication networks must have. Delay and disruption-tolerant networks (DTNs), are characterized by their lack of connectivity, resulting in a lack of instantaneous end-to-end paths. In these challenging environments, popular ad hoc routing protocols such as AODV and DSR fail to establish routes. This is due to these protocols trying to first establish a complete route and then, after the route has been established, forward the actual data. However, when instantaneous end-to-end paths are difficult or impossible to establish, routing protocols must take to a "store and forward" approach, where data is incrementally moved and stored throughout the network in hopes that it will eventually reach its destination. A common technique used to maximize the probability of a message being successfully transferred is to replicate many copies of the message in the hope that one will succeed in reaching its destination. This is feasible only on networks with large amounts of local storage and internode bandwidth relative to the expected traffic. In many common problem spaces, this inefficiency is outweighed by the increased efficiency and shortened delivery times made possible by taking maximum advantage of available unscheduled forwarding opportunities. In others, where available storage and internode throughput opportunities are more tightly constrained, a more discriminate algorithm is required.

Other concerns

Bundle protocols

In efforts to provide a shared framework for algorithm and application development in DTNs, RFC 4838 and RFC 5050 were published in 2007 to define a common abstraction to software running on disrupted networks. Commonly known as the Bundle Protocol, this protocol defines a series of contiguous data blocks as a bundle—where each bundle contains enough semantic information to allow the application to make progress where an individual block may not. Bundles are routed in a store and forward manner between participating nodes over varied network transport technologies (including both IP and non-IP based transports). The transport layers carrying the bundles across their local networks are called *bundle convergence layers.* The bundle architecture therefore operates as an overlay network, providing a new naming architecture based on Endpoint Identifiers (EIDs) and coarse-grained class of service offerings.

Protocols using bundling must leverage application-level preferences for sending bundles across a network. Due to the store and forward nature of delay-tolerant protocols, routing solutions for delay-tolerant networks can benefit from exposure to application-layer information. For example, network scheduling can be influenced if application data must be received in its entirety, quickly, or without variation in packet delay. Bundle protocols collect application data into bundles that can be sent across heterogeneous network configurations with high-level service guarantees. The service guarantees are generally set by the application level, and the RFC 5050 Bundle Protocol specification includes 'bulk', 'normal', and 'expedited' markings.

Security

Addressing security issues has been a major focus of the bundle protocol.

Security concerns for delay-tolerant networks vary depending on the environment and application, though authentication and privacy are often critical. These security guarantees are difficult to establish in a network without persistent connectivity because the network hinders complicated cryptographic protocols, hinders key exchange, and each device must identify other intermittently-visible devices. Solutions have typically been modified from mobile ad hoc network and distributed security research, such as the use of distributed certificate authorities and PKI schemes. Original solutions from the delay-tolerant research community include: 1) the use of identity-based encryption, which allows nodes to receive information encrypted with their public identifier, and 2) the use of tamper-evident tables with a gossiping protocol;

Research efforts

Various research efforts are currently investigating the issues involved with DTN:

- The The Delay-Tolerant Networking Research Group.
- The Technology and Infrastructure for Developing Regions project at UC Berkeley
- The KioskNet research project at the University of Waterloo.
- The DieselNet research project at the University of Massachusetts, Amherst.
- The ResiliNets Research Initiative at the University of Kansas and Lancaster University.
- The Haggle EU research project.
- The N4C EU/FP7 research project.
- The WNaN DARPA project.
- The EMMA project at TU Braunschweig
- The DTN networking at Helsinki University of Technology.
- The SARAH project, funded by the French National Research Agency (ANR).
- The development of the DoDWAN platform at the University of South Brittany.
- The CROWD project, funded by the French National Research Agency (ANR).
- The PodNet project at KTH Stockholm and ETH Zurich.

Some research efforts look at DTN for the Interplanetary Internet by examining use of the Bundle Protocol in space:

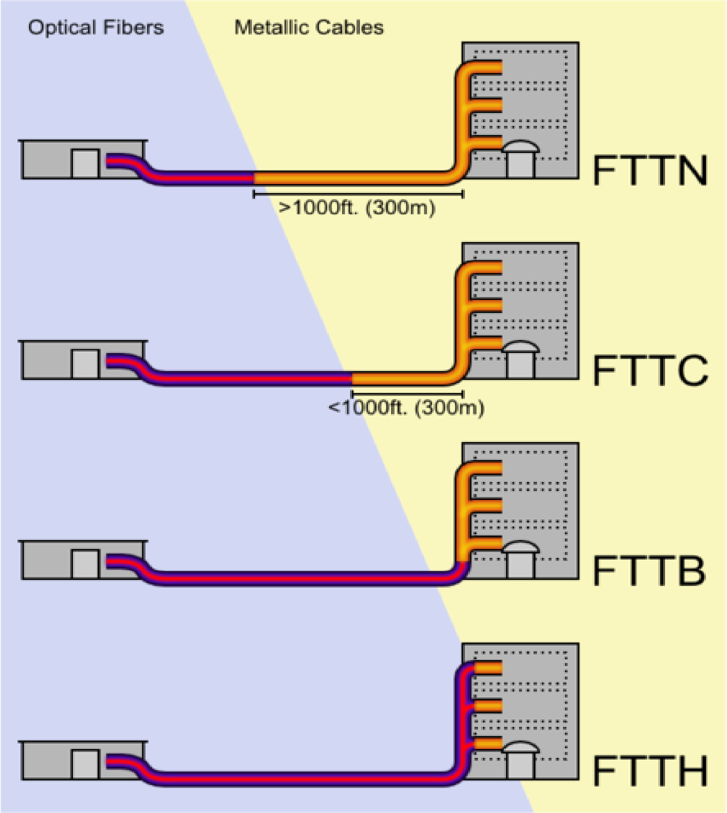
- The Saratoga project at the University of Surrey, which was the first to test the bundle protocol in space on the UK-DMC Disaster Monitoring Constellation satellite in 2008.
- NASA JPL's Deep Impact Networking (DINET) Experiment on board the Deep Impact/EPOXI spacecraft.
- BioServe Space Technologies, one of the first payload developers to adopt the DTN technology, has utilized their CGBA (Commercial Generic Bioprocessing

Apparatus) payloads onboard the ISS, which provide computational/communications platforms, to implement the DTN protocol.

Chapter- 4

Fiber to the x

Fiber to the *x* (**FTT*x***) is a generic term for any broadband network architecture that uses optical fiber to replace all or part of the usual metal local loop used for last mile telecommunications. The generic term originated as a generalization of several configurations of fiber deployment (FTTN, FTTC, FTTB, FTTH...), all starting by FTT but differentiated by the last letter, which is substituted by an *x* in the generalization.



A schematic illustrating how FTTx architectures vary — with regard to the distance between the optical fiber and the end-user. The building on the left is the central office; that on the right is one of the buildings served by the central office. Dotted rectangles represent separate living or office spaces within the same building.

Definition of terms

The telecommunications industry differentiates between several distinct configurations. The terms in most widespread use today are:

- FTTN - Fiber-to-the-node - fiber is terminated in a street cabinet up to several kilometers away from the customer premises, with the final connection being copper.
- FTTC - Fiber-to-the-cabinet or fiber-to-the-curb - this is very similar to FTTN, but the street cabinet is closer to the user's premises; typically within 300 m.
- FTTB - Fiber-to-the-building or Fiber-to-the-basement - fiber reaches the boundary of the building, such as the basement in a multi-dwelling unit, with the final connection to the individual living space being made via alternative means.
- FTTH - Fiber-to-the-home - fiber reaches the boundary of the living space, such as a box on the outside wall of a home.
- FTTP - Fiber-to-the premises - this term is used in several contexts: as a blanket term for both FTTH and FTTB, or where the fiber network includes both homes and small businesses.

To promote consistency, especially when comparing FTTH penetration rates between countries, the three FTTH Councils of Europe, North America and Asia-Pacific have agreed upon definitions for FTTH and FTTB. The FTTH Councils do not have formal definitions for FTTC and FTTN.

It is worth pointing out that fiber to the telecom enclosure (FTTE) is not considered to be part of the FTTx group of technologies, despite the similarity in name. FTTE is a form of structured cabling typically used in the enterprise local area network, where fiber is used to link the main computer equipment room to an enclosure close to the desk or workstation. Similarly, in fiber-to-the-desk a fiber connection is installed from the main computer room to a terminal at the desk.

Benefits of fiber in the access network

The speeds of fiber optic and copper cables are both limited by length, but copper is much more sharply limited in this respect. For example, gigabit Ethernet runs over relatively economical category 5e, category 6, or augmented category 6 unshielded twisted pair copper cabling but only to 100 meters. However, over the right kind of fiber, gigabit ethernet can easily reach distances of tens of kilometers.

Even in the commercial world, most computers have copper communication cables. But these cables are short, typically tens of meters. Most metropolitan network links (e.g., those based on telephone or cable television services) are several kilometers long, in the range where fiber significantly outperforms copper. Replacing at least part of these links with fiber shortens the remaining copper segments and allows them to run much faster.

Fiber configurations that bring fiber right into the building can offer the highest speeds since the remaining segments can use standard Ethernet or coaxial cable. Fiber configurations that transition to copper in a street cabinet are generally too far from the users for standard Ethernet configurations over existing copper cabling. They generally use VDSL at (downstream) speeds of several tens of megabits per second.

Fiber is often said to be 'future proof' because the speed of the broadband connection is usually limited by the terminal equipment rather than the fiber itself, permitting at least some speed improvements by equipment upgrades before the fiber itself must be upgraded. Still, the type and length of employed fibers chosen, e.g. multimode vs single mode, are critical for applicability for future high gigabit connections.

Fiber to the node

Fiber to the node (FTTN), also called fiber to the neighborhood or fiber to the cabinet (FTTCab), is a telecommunication architecture based on fiber-optic cables run to a cabinet serving a neighborhood. Customers typically connect to this cabinet using traditional coaxial cable or twisted pair wiring. The area served by the cabinet is usually less than 1,500 m in radius and can contain several hundred customers. (If the cabinet serves an area of less than 300 m in radius then the architecture is typically called fiber to the curb.)

Fiber to the node allows delivery of broadband services such as high speed Internet. High speed communications protocols such as broadband cable access (typically DOCSIS) or some form of DSL are used between the cabinet and the customers. The data rates vary according to the exact protocol used and according to how close the customer is to the cabinet.

Unlike the competing fiber to the premises technology, fiber to the node often uses the existing coaxial or twisted pair infrastructure to provide last mile service. For this reason, fiber to the node is less costly to deploy. In the long-term, however, its bandwidth potential is limited relative to implementations which bring the fiber still closer to the subscriber.

Fiber to the last amplifier

FTTLA are the initials of *Fiber To The Last Amplifier*. The network cables being able to use several amplifiers, the FTTLA aims at replacing the coaxial cable to the last amplifier (towards the subscriber) by optical fiber. It acts as a new technology aiming at re-using the network cables existing in particular on the final part while installing of optical fiber more closely to the subscriber while using the coaxial cable of the networks cables for the "last mile" or "last meters" connected with the subscriber.

Fiber to the last amplifier (FttLA) node is an efficient tool to deploy fiber deeper into the CATV network architecture and add most desirable aspects of scalability (performance and reliability) which are necessary when new services (i.e. "triple play", video on demand, gaming) are introduced.

FTTLA is a technology which assists hybrid fiber-coaxial CATV networks to provide to their customers more bandwidth. Using a replacement of all coaxial active equipments by nodes (optical receiver) with high power output (up to 117 dBuV). The coaxial is maintained from the node to the customer without any active equipment in between.

From the optical sender to the node, it uses fiber which is split by 4 or by 8 depending on the distance and on the output power of the optical sender (from 6 to 16 dBm).

Also, IM2, IM3 and C/N are modified for a better network and it also has other benefits such as power saving in the network, as the power consumption is lower than a normal HFC network (up to 40%).

Fiber to the curb

Fiber to the curb (FTTC) is a telecommunications system based on fiber-optic cables run to a platform that serves several customers. Each of these customers has a connection to this platform via coaxial cable or twisted pair.

Fiber to the curb allows delivery of broadband services such as high speed internet. High speed communications protocols such as broadband cable access (typically DOCSIS) or some form of DSL are used between the cabinet and the customers. The data rates vary according to the exact protocol used and according to how close the customer is to the cabinet.

FTTC is subtly distinct from FTTN or FTTP (all are versions of Fiber in the Loop). The chief difference is the placement of the cabinet. FTTC will be placed near the "curb" which differs from FTTN which is placed far from the customer and FTTP which is placed right at the serving location.

Unlike the competing fiber to the premises (FTTP) technology, fiber to the curb can use the existing coaxial or twisted pair infrastructure to provide last mile service. For this reason, fiber to the curb costs less to deploy. However, it also has lower bandwidth potential than fiber to the premises.

In the United States of America and Canada, the largest deployment of FTTC was carried out by BellSouth Telecommunications. With the acquisition of BellSouth by AT&T, deployment of FTTC will end. Future deployments will be based on either FTTN or FTTP. Existing FTTC plant may be removed and replaced with FTTP.

Fiber to the premises

Fiber to the premises is a form of fiber-optic communication delivery in which an optical fiber is run from the central office all the way to the premises occupied by the subscriber.

Fiber to the premises is often abbreviated with the acronym FTTP. However, this acronym has become ambiguous and may instead refer to a form of fiber to the curb where the fiber terminates at a utility pole without reaching the premises.

FTTH vs. FTTB

Fiber to the premises can be categorized according to where the optical fiber ends:

- FTTH (fiber to the home) is a form of fiber optic communication delivery in which the fiber extends from the central office to the subscriber's living or working space. Once at the subscriber's living or working space, the signal may be conveyed throughout the space using any means, including twisted pair, coaxial cable, wireless, power line communication, or optical fiber.

- FTTB (fiber to the building, also called fiber to the basement) is a form of fiber optic communication delivery in which the optical fiber terminates before actually reaching the subscribers living or working space itself, but does extend to the property containing that living or working space. The signal is conveyed the final distance using any non-optical means, including twisted pair, coaxial cable, wireless, or power line communication. By definition, FTTB necessarily applies only to those properties which contain multiple living or working spaces.

An apartment building may provide an example of the distinction between FTTH and FTTB. If a fiber is run to a panel at each subscriber's apartment, this is FTTH. If instead the fiber goes only as far as the apartment building's shared electrical room, then this is FTTB.

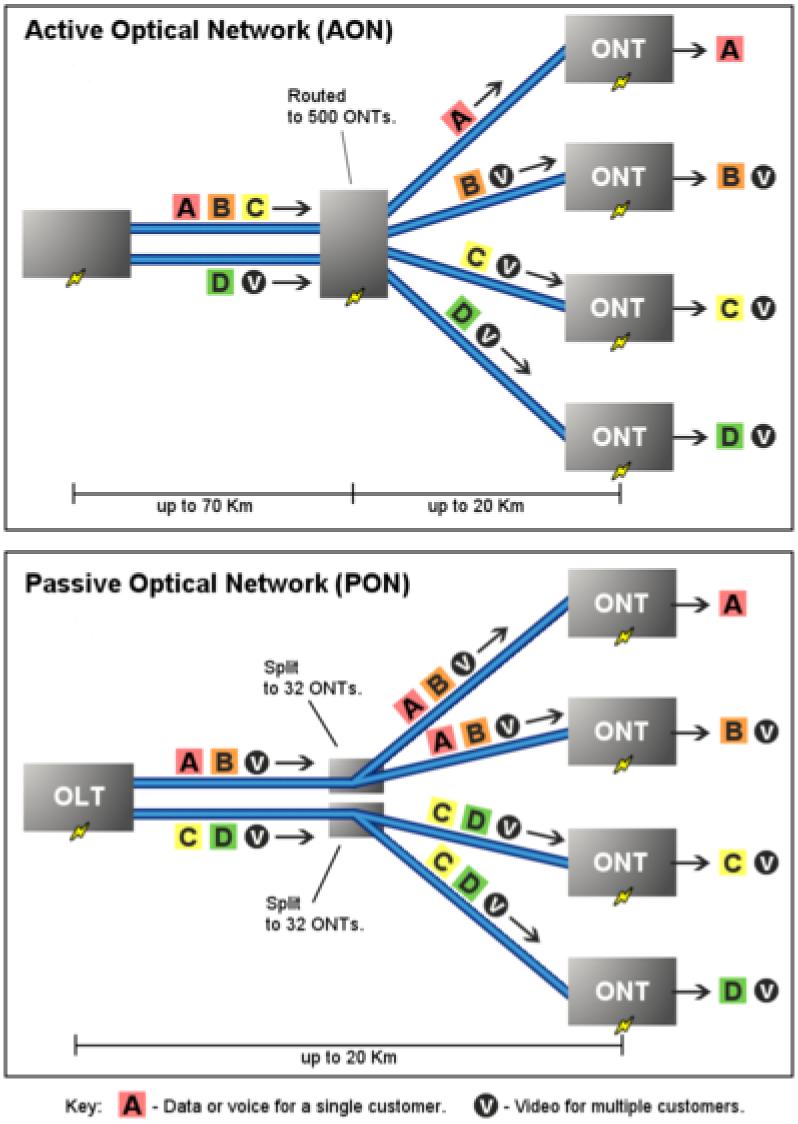
Direct fiber

The simplest optical distribution network can be called direct fiber. In this architecture, each fiber leaving the central office goes to exactly one customer. Such networks can provide excellent bandwidth since each customer gets their own dedicated fiber extending all the way to the central office. However, this approach is about 10% more costly due to the amount of fiber and central office machinery required. The approach is generally favored by new entrants and competitive operators. A benefit of this approach is that it doesn't exclude any layer 2 networking technologies, be they Passive optical network, Active Optical Network, etc. From a regulatory point of view it leads to least implications as any form of regulatory remedy is still possible using this topology.

Shared fiber

More commonly each fiber leaving the central office is actually shared by many customers. It is not until such a fiber gets relatively close to the customers that it is split into individual customer-specific fibers. There are two competing optical distribution network architectures which achieve this split: active optical networks (AONs) and passive optical networks (PONs).

Active optical network



Comparison showing how a typical active optical network handles downstream traffic differently than a typical passive optical network. The type of active optical network shown is a star network capable of multicasting. The type of passive optical network shown is a star network having multiple splitters housed in the same cabinet.

Active optical networks rely on some sort of electrically powered equipment in Optical Distribution Network(ODN) to distribute the signal, such as a switch or router. Normally, optical signals need O-E-O transformation in ODN. Each signal leaving the central office is directed only to the customer for which it is intended. Incoming signals from the customers avoid colliding at the intersection because the powered equipment there provides buffering.

As of 2007, the most common type of active optical networks are called active Ethernet, a type of Ethernet in the first mile (EFM). Active Ethernet uses optical Ethernet switches to distribute the signal, thus incorporating the customers' premises and the central office into one giant switched Ethernet network. Such networks are identical to the Ethernet computer networks used in businesses and academic institutions, except that their purpose is to connect homes and buildings to a central office rather than to connect computers and printers within a campus. Each switching cabinet can handle up to 1,000 customers, although 400-500 is more typical. This neighborhood equipment performs layer 2/layer 3 switching and routing, offloading full layer 3 routing to the carrier's central office. The IEEE 802.3ah standard enables service providers to deliver up to 100 Mbit/s full-duplex over one single-mode optical fiber to the premises depending on the provider. Speeds of 1Gbit/s are becoming commercially available.

Passive optical network

A passive optical network (PON) is a point-to-multipoint, fiber to the premises network architecture in which unpowered optical splitters are used to enable a single optical fiber to serve multiple premises, typically 32-128. A PON configuration reduces the amount of fiber and central office equipment required compared with point to point architectures.

Downstream signal coming from the central office is broadcast to each customer premises sharing a fiber. Encryption is used to prevent eavesdropping.

Upstream signals are combined using a multiple access protocol, invariably time division multiple access (TDMA). The OLTs "range" the ONUs in order to provide time slot assignments for upstream communication.

Electrical portion

Once on private property, the signal typically travels the final distance to the end user's equipment using an electrical format.

A device called an Optical Network Terminal (ONT), also called an Optical Network Unit (ONU), converts the optical signal into an electrical signal. (ONT is an ITU-T term, whereas ONU is an IEEE term, but the two terms mean exactly the same thing.) Optical network terminals require electrical power for their operation, so some providers connect them to back-up batteries in case of power outages. Optical network units use thin film filter technology to convert between optical and electrical signals.

For fiber to the home and for some forms of fiber to the building, it is common for the building's existing phone systems, local area networks, and cable TV systems to connect directly to the ONT.

If all three systems cannot directly reach the ONT, it is possible to combine signals and transport them over a common medium. Once closer to the end-user, equipment such as a router, modem, and/or network interface module can separate the signals and convert them into the appropriate protocol. For example, one solution for apartment buildings uses VDSL to combine data (and / or video) with voice. With this approach, the combined signal travels through the building over the existing telephone wiring until it reaches the end-user's living space. Once there, a VDSL modem copies the data and video signals and converts them into Ethernet protocol. These are then sent over the end user's category 5 cable. A network interface module can then separate out the video signal and convert it into an RF signal that is sent over the end-user's coaxial cable. The voice signal continues to travel over the phone wiring and is sent through DSL filters to remove the video and data signals. An alternative strategy allows data and / or voice to be transmitted over coaxial cable. In yet another strategy, some office buildings dispense with the telephone wiring altogether, instead using voice over Internet Protocol phones that can plug directly into the local area network.

Chapter- 5

GINA: Global Information Network Architecture

The concept for the **Global Information Network Architecture (GINA)** evolved from a realization that the current technologies provided an unprecedented opportunity to create a useful Global Information Grid (GIG) that could transform the possibilities for Net-Centric Operations.

The Global Information Network Architecture (GINA) Team was created in 2004 to address this possibility. Originally developed under Cooperative Research and Development Agreement (CRADA) with The US Naval Postgraduate School (NPS) in Monterey CA, the projects initial title was Network Aware Business Data Management System (NABDMS).

In late 2008, the United States Army Corps of Engineers (USACE) Engineer Research and Development Center (ERDC) began the second phase of GINA's development. Currently focus is on purposing GINA as a High Level Architecture (HLA) for System Fusion Networks (SFN), GINA is being evolved and deployed globally to facilitate interoperability and a new form of computational design.

The Global Information Grid

"The globally interconnected, end-to-end set of information capabilities, associated processes, and personnel for collecting, processing, storing, disseminating and managing information on demand to warfighters, policy makers, and support personnel. The Global Information Grid (GIG) includes all owned and leased communications and computing systems and services, software (including applications), data, security services, and other associated services necessary to achieve Information Superiority. It also includes National Security Systems as defined in section 5142 of the Clinger-Cohen Act of 1996. "

The United States Department of Defense (DoD) recognized back in 1996 the need to have a GIG. We have struggled, largely unsuccessfully to bring a reasonable facsimile of a GIG forward. There have been numerous research projects directed at creating a GIG, but actually turning DoD networks into a functioning GIG has proved elusive. The GIG is a very hard problem that requires a rethinking of current approaches to interoperability.

Vector Relational Data Modeling (VRDM)... Modeling Models.

The dictionary defines modeling as, “The representation, often mathematical, of a process, concept, or operation of a system.”

While software modeling languages such as Unified Modeling Language (UML) or Object Role Modeling (ORM) attempt to provide an iconic representation to articulate the structure or architecture of an application, they are not executable and have limited applicability beyond the software design environment.

The GINA Team created GINA to enable model-based software engineering, but we did so in such a way that the model, once-defined, represented a working application. By properly bounding the problem space to "Information Applications", i.e., non-algorithmically-intense applications with linear relationships representing the vast majority of software applications, the GINA team was able to create a configurable Component Based Object Model (CBOM) for information management where the configuration represented the instructions for assembling a working implementation of the model. Moreover, the GINA team made the decision early in its development to make GINA a GINA model. By doing so the configuration itself could be controlled by configuration. As will be illustrated later, that turned out to be an important decision. Enabling GINA’s deep configurability required the development and implementation of multiple models.

The **Control Model** assembles the components of the component model, according to the assembly instructions in the Application Model into the structures defined in the Implementation Model to create GINA information objects.

The **Application Model** describes actual GINA applications. Both the description of applications and the GINA and Application Models themselves. These applications are described in terms of components in the Component Model. It represents the set of components that are assembled in order to create GINA information objects as specified in the application model.

The Control Model assembles the components of the component model, according to the assembly instructions in the Application Model into the structures defined in the **Implementation Model** to create GINA information objects.

Ultimately, we have to define GINA applications using a development model that is appropriate for developing GINA applications. And again, the GINA **Development Model** is itself described as a GINA application.

GINA, at a high level, is a model for modeling. Designed to facilitate a principle of development where relationships between objects can be objectified and wielded as objects in their own right, GINA itself is an executable model configured using VRDM.

VRDM is a core concept that is embodied by GINA. GINA could be looked at as an environment that turns collected data into a multi-dimensional object environment with each object being connected to other objects through vectors. This environment makes many of the information-centric tasks that a user might want to perform far easier than any other approach.

A key concept of VRDM is that relationships among information objects should themselves be defined as information objects, and be fully configurable.

Taking relationships and implementing them as GINA objects enables GINA to take configurations and assemble models that can perform most, if not all of the work done by typical hard-coded information applications, such as most enterprise systems, integrations, and information sharing systems.

As a result of VRDM, it is possible to specify an application through describing the required components as a series of objects and their relationships, or vectors. VRDM enables disparate data, from disparate sources to be invoked and configured to relate in a “System of Systems” model called a specification. The behavior of a specification can be the much same as a contemporary application. The difference, however, is that with VRDM, there is no programming. GINA is a true CBOM for Object Modeling, where the models are themselves executable.

Critical to this approach is the concept of **Reflexivity**, i.e., the GINA model is described as a GINA model, that permits deep configurability. When one is using the interactive development environment used to create GINA applications, one is using a GINA application. More importantly, GINA assembles GINA applications according to a GINA model for GINA applications. Deeper still, the GINA model is itself an example of a GINA model. This deep configurability is the key to GINA’s power as interoperability and multi-level security ("MLS") engine.

As a combination of hardware and software-based components, which meets the requirements of a true GIG, GINA is configured as a universal virtual network of data of any type from any source in any location on collected physical networks. GINA is a product of several key concepts which collectively enable it to represent a complete, configurable interoperability environment. These key concepts are embodied in a series of layers and components which collectively allow GINA to perform the types of functions and provide the type of services it needs to be a fully functional interoperability environment.

Vector Relational Data Modeling Core Concepts

Descriptive Programming

Just as Assembler made machine language programming faster and more accurate, and descriptive languages for specifying procedures (“4GL”s like SQL) made procedural programming faster and more accurate, GINA's VRDM brings a new level of speed and accuracy to object-oriented programming.

In the long run GINA's encapsulation of the management of network-available data may be more important that even its speed and accuracy for large organizations.

Component Objects

With VRDM, Data Agnostic Objects can be created to represent common relationships called Mechanisms. These Mechanisms can be reused and combined with others, new and existing, to create systems and subsystems. This facilitates rapid deployment and non-programmatic implementation.

Complexity

Just as everything in the world is assembled from remarkably few elements, arbitrarily complex systems can be assembled from relatively few objects. The key in both cases is that objects have to be designed for interaction and assembly. GINA is designed in that way. Fundamentally, at the bottom level are very few objects, e.g., objects, or XTypes in VRDM, and relationships between XTypes, or Vectors in VRDM. These objects are the primitives on which the VRDM object management model is based. In turn, instances of these primitives are assembled into the basic building blocks of VRDM: fully defined objects representing XTypes and Vectors, as well as constraints and simple entities. These can then be assembled to fully describe the GINA environment, and to allow the administrator to create the data objects to support a Task Oriented User Interface (TOUI), or a specific application.

WorldSpace

A central concept in GINA is that objects can be referenced in multiple WorldSpaces, depending on how a user gets to that object. A WorldSpace determines the applicability of an object's vectors, e.g. attributes and relationships, when assembled for a particular event or usage. WorldSpaces are inherently hierarchical: as one more tightly defines the WorldSpace associated with an event or usage, the more tightly one must define, and the more granularly one needs to specify associated behaviors.

HyperPlanes

If we look back at the concepts associated with GINA, we could say that an object exists in a 3 dimensional data object space. Its location in that space is defined by its order of complexity, its usage and related components, and the user and WorldSpace in which it is being accessed. At any given time the behavior of a system is dictated by all of its objects locations in this 3-dimensional object- space. However, this behavior is not the same for every user, and is influenced by the characteristics of that user that effectively define hyperplanes in this object space. Thus, the appropriate object model is only summarized in three dimensions, with multiple user-based dimensions of behavior effectively being summarized on this graphic. In effect, at any given time an object exists as a point in 7+ dimensional object behavior space. Remarkably, GINA not only models this space effectively, but does so in an environment where most specifications are created through configuration, not programming.

Directory Sub System (DSS)

GINA is implemented through a software-based, multi-layer, configurable data object management environment. Just as the entirety of GINA can be viewed as a series of well-structured layers, the data object management environment is also structured and layered, with multiple layers of the object management environment corresponding to each of the top three layers in the overall GINA. GINA's "DSS" layer is actually composed of two separate implementation layers: a content server layer that consists of a collection of configurable objects that know how to navigate the network, acquire data, and present it in a consistent way; and an aggregation layer that homogenizes all incoming data, in both format and name, and presents itself as a universal object repository that insulates the information consumer from the complexities of managing the underlying data stores.

Data Access Layer (DAL)

GINA collects data from aggregated systems using a collection of adaptors called Content Servers which structure the protocols, formats, and syntax of collected data into a common representation that then becomes the base data that can then be managed through the GINA model. Just as the providers of data to GINA operate on multiple protocols, formats, and syntaxes, the prospective consumers of GINA data may require information using their own protocols, formats, and syntax. GINA exposes itself using a standard "Data Access Layer" ("DAL") that can be—and has been—used to provide data to standardized or customized DALs such as SOAP Web Services, ODBC interfaces, etc.

Task Oriented User Interface (TOUI)

Another model that has been built in GINA is called the Task-Oriented User Interface, or "TOUI". The current mainstream approach to user interfaces ("UIs") involves a process where a developer "paints", or in some other way creates a mark-up of the UI, and then defines the binding of components of that the UI to the underlying application using some standardized approach. The TOUI model takes a different approach: a UI is assembled at the time of request from components according to a set of vectors that take into the account model states and the user during the assembly process. As a result, the UI no longer represents the application that uses information, but rather becomes the external expression of the information model that represents the application. Moreover, because the definition of the UI is done as a set of metadata-defined GINA components, the expression of those components can be done in any environment that has sufficiently strong semantics for representing applications, whether that is Java, .NET, Python, or even a 3-D visualization environment.interfaces, etc.

Chapter- 6

IBM Systems Network Architecture

Systems Network Architecture (**SNA**) is IBM's proprietary networking architecture created in 1974. It is a complete protocol stack for interconnecting computers and their resources. SNA describes the protocol and is, in itself, not actually a program. The implementation of SNA takes the form of various communications packages, most notably Virtual telecommunications access method (VTAM) which is the mainframe package for SNA communications. SNA is still used extensively in banks and other financial transaction networks, as well as in many government agencies. While IBM is still providing support for SNA, one of the primary pieces of hardware, the 3745/3746 communications controller has been withdrawn from marketing by the IBM Corporation. However, there are an estimated 20,000 of these controllers installed and IBM continues to provide hardware maintenance service and micro code features to support users. A robust market of smaller companies continues to provide the 3745/3746, features, parts and service. VTAM is also supported by IBM, as is the IBM Network Control Program (NCP) required by the 3745/3746 controllers.

Objectives of SNA

IBM in the mid-1970s saw itself mainly as a hardware vendor and hence all its innovations in that period aimed to increase hardware sales. SNA's objective was to reduce the costs of operating large numbers of terminals and thus induce customers to develop or expand interactive terminal based-systems as opposed to batch systems. An expansion of interactive terminal based-systems would increase sales of terminals and more importantly of mainframe computers and peripherals - partly because of the simple increase in the volume of work done by the systems and partly because interactive processing requires more computing power per transaction than batch processing.

Hence SNA aimed to reduce the main non-computer costs and other difficulties in operating large networks using earlier communications protocols. The difficulties included:

- A communications line could not be shared by terminals whose users wished to use different types of application, for example one which ran under the control of CICS and another which ran under TSO.
- Often a communications line could not be shared by terminals of different types, as they used different "dialects" of the existing communications protocols. Up to

the early 1970s, computer components were so expensive and bulky that it was not feasible to include all-purpose communications interface cards in terminals. Every type of terminal had a hard-wired communications card which supported only the operation of one type of terminal without compatibility with other types of terminals on the same line.
- The protocols which the primitive communications cards could handle were not efficient. Each communications line used more time transmitting data than modern lines do.
- Telecommunications lines at the time were of much lower quality. For example, it was almost impossible to run a dial-up line at more than 300 bits per second because of the overwhelming error rate, as comparing with 56,000 bits per second today on dial-up lines; and in the early 1970s few leased lines were run at more than 2400 bits per second (these low speeds are a consequence of Shannon's Law in a relatively low-technology environment). Telecommunications companies had little incentive to improve line quality or reduce costs, because at the time they were mostly monopolies and sometimes state-owned.

As a result running a large number of terminals required a lot more communications lines than the number required today, especially if different types of terminals needed to be supported, or the users wanted to use different types of applications (.e.g. under CICS or TSO) from the same location. In purely financial terms SNA's objectives were to increase customers' spending on terminal-based systems and at the same time to increase IBM's share of that spending, mainly at the expense of the telecommunications companies.

SNA also aimed to overcome a limitation of the architecture which IBM's System/370 mainframes inherited from System/360. Each CPU could connect to at most 16 "channels" (devices which acted as controllers for peripherals such as tape and disk drives, printers, card-readers) and each channel could handle up to 16 peripherals - i.e. there was maximum of 256 peripherals per CPU. At the time when SNA was designed, each communications line counted as a peripheral. Thus the number of terminals with which powerful mainframe could otherwise communicate is severely limited.

Principal components and technologies

Improvements in computer component technology made it feasible to build terminals that included more powerful communications cards which could operate a single standard communications protocol rather than a very stripped-down protocol which suited only a specific type of terminal. As a result several multi-layer communications protocols were proposed in the 1970s, of which IBM's SNA and ITU-T's X.25 became dominant later.

The most important elements of SNA include:

- IBM Network Control Program (NCP) is a primitive switching protocol, implemented in 3705 communications processors. The protocol performed two main functions:

    - It is a packet forwarding protocol, acting like modern switch - forwarding data packages to the next node, which might be a mainframe, a terminal or another 3705. The communications processors supported only hierarchical networks with a mainframe at the center, unlike modern routers which support peer-to-peer networks in which a machine at the end of the line can be both a client and a server at the same time.
    - It is a multiplexer that connected multiple terminals into one communication line to the CPU, thus relieved the constraints on the maximum number of communication lines per CPU. A 3705 could support a larger number of lines (352 initially) but only counted as one peripheral by the CPUs and channels. Since the launch of SNA IBM has introduced improved communications processors, of which the latest is the 3745.
- Synchronous Data Link Control (SDLC), a protocol which greatly improved the efficiency of data transfer over a single link:
    - SDLC included much more powerful error detection and correction codes than earlier protocols. These codes often enabled the communications cards to correct minor transmission errors without requesting re-transmission, and therefore made it possible to pump data down a line much faster.
    - It enabled terminals and 3705 communications processors to send "frames" of data one after the other without waiting for an acknowledgement of the previous frame - the communications cards had sufficient memory and processing capacity to "remember" the last 7 frames sent or received, request re-transmission of only those frames which contained errors that the error detection and correction codes could not repair, and slot the re-transmitted frames into the right place in the sequence before forwarding them to the next stage.
    - These frames all had the same type of "envelope" (frame header and trailer) which contained enough information for data packages from different types of terminal to be send along the same communications line, leaving the mainframe to deal with any differences in the formatting of the content or in the rules governing dialogs with different types of terminal.

  Remote terminals (i.e. those connected to the mainframe by telephone lines) and 3705 communications processors would have SDLC-capable communications cards.
  This is the precursor of the so called "packet communication" that eventually evolved into today's IP technology, and SDLC itself evolved into HDLC that is one of the base technology for dedicated telecommunication circuit.

- VTAM, a software package to provide log-in, session keeping and routing services within the mainframe. A terminal user would log-in via VTAM to a specific application or application environment (e.g. CICS or TSO). A VTAM device would then route data from that terminal to the appropriate application or application environment until the user logged out and possibly logged in to another application. The original versions of IBM hardware could only keep one session per terminal. In the 1980s further software (mainly from third-party vendors) made it possible for a terminal to have simultaneous sessions with different applications or application environments.

Advantages and disadvantages

SNA removed link control from the application program and placed it in the NCP. This had the following advantages and disadvantages:

Advantages

- Localization of problems in the telecommunications network was easier because a relatively small amount of software actually dealt with communication links. There was a single error reporting system.
- Adding communication capability to an application program was much easier because the formidable area of link control software that typically requires interrupt processors and software timers was relegated to system software and NCP.
- With the advent of APPN, routing functionality was the responsibility of the computer as opposed to the router (as with TCP/IP networks). Each computer maintained a list of Nodes that defined the forwarding mechanisms. A centralized node type known as a Network Node maintained Global tables of all other node types. APPN stopped the need to maintain APPC routing tables that explicitly defined endpoint to endpoint connectivity. APPN sessions would route to endpoints through other allowed node types until it found the destination. This was similar to the way that TCP/IP routers function today.

Disadvantages

- Connection to non-SNA networks was difficult. An application which needed access to some communication scheme, which was not supported in the current version of SNA, faced obstacles. Before IBM included X.25 support (NPSI) in SNA, connecting to an X.25 network would have been awkward. Conversion between X.25 and SNA protocols could have been provided either by NCP software modifications or by an external protocol converter.

- A sheaf of alternate pathways between every pair of nodes in a network had to be predesigned and stored centrally. Choice among these pathways by SNA was rigid and did not take advantage of current link loads for optimum speed.

- SNA network installation and maintenance are complicated and SNA network products are (or were) expensive. Attempts to reduce SNA network complexity by adding IBM Advanced Peer-to-Peer Networking functionality were not really successful, if only because the migration from traditional SNA to SNA/APPN was very complex, without providing much additional value, at least initially. SNA software licences (VTAM) cost as much as $10000 a month for high-end systems.

And SNA IBM 3745 Communications Controllers typically cost over $100K. TCP/IP was still seen as unfit for commercial applications e.g. in the finance industry until the late 1980s, but rapidly took over in the 1990s due to its peer-to-peer networking and packet communication technology it deployed.

- The design of SNA was in the era when the concept of layered communication was not fully adopted by the computer industry. Applications, databases and communication functions were mingled into the same protocol or product, to make it difficult to maintain or manage. That was very common for the products created in that time. Even after TCP/IP was fully developed, X window system was designed with the same model where communication protocols were embedded into graphic display application.

- SNA's connection based architecture invoked huge state machine logic to "keep track" of everything. APPN added a new dimension to state logic with its concept of differing node types. While it was solid when everything was running correctly, there was still a need for manual intervention. Simple things like watching the Control Point sessions had to be done manually. APPN wasn't without issues; in the early days many shops abandoned it due to issues found in APPN support. Over time, however, many of the issues were worked out but not before the advent of the Web Browser which was the beginning of the end for SNA.

Logical unit types

Network Addressable Units in an SNA network are any components that can be assigned an address and can send and receive information. They are distinguished further as follows:

- *System Service Control Points*, provide services to manage a network or subnetwork (typically in the mainframe),
- *Physical Units*, a physical device or communications link (relating to boxes),
- *Logical Units*, an access point to the network (relating to applications or subsystems such as CICS and TSO) or terminals.

SNA essentially offers transparent communication: equipment specifics don't impose any constraints onto LU-LU communication. But eventually it serves a purpose to make a distinction between LU types, as the application must take the functionality of the terminal equipment into account (e.g. screen sizes and layout).

SNA defines several kinds of devices, called Logical Unit types:

- LU0 provides for undefined devices, or build your own protocol.
- LU1 devices are printers.
- LU2 devices are dumb IBM 3270 display terminals.
- LU3 devices are printers using 3270 protocols.

- LU4 devices are batch terminals.
- LU5 has never been defined.
- LU6 provides for protocols between two applications.
- LU7 provides for sessions with IBM 5250 terminals.

The primary ones in use are LU1, LU2, and LU6.2 (an advanced protocol for application to application conversations).

Within SNA there are two types of data stream to connect local terminals and printers; there is the 3270 data stream mainly used by mainframes (zSeries family) and the 5250 data stream mainly used by minicomputers/servers such as the S/36, S/38, and AS/400 (now System i).

Starting from version 5.2 of OS/400, SNA for client-access is no longer supported.

The term 37xx refers to IBM's family of SNA communications controllers. The 3745 supports up to eight high-speed T1 circuits, the 3725 is a large-scale node and front-end processor for a host, and the 3720 is a remote node that functions as a concentrator and router.

Implementation and publication

SNA was made public as part of IBM's "Advanced Function for Communications" announcement in September, 1974, which included the implementation of the SNA/SDLC (Synchronous Data Link Control) protocols on new communications products:

- IBM 3767 communication terminal (printer)
- IBM 3770 data communication system

They were supported by IBM 3704/3705 communication controllers and their Network Control Program, and by System/360 and System/370 and their VTAM and other software such as CICS and IMS. This announcement was followed by another announcement in July, 1975, which introduced the IBM 3760 data entry station, the IBM 3790 communication system, and the new models of the IBM 3270 display system.

SNA was mainly designed by the IBM Systems Development Division laboratory in Research Triangle Park, North Carolina, USA, helped by other laboratories that implemented SNA/SDLC. The details were later made public by IBM's System Reference Library manuals and IBM Systems Journal.

Competitors

The proprietary networking architecture for Honeywell Bull mainframes is Distributed Systems Architecture (DSA). Communications package for DSA is VIP. Like SNA, DSA is also no longer supported for client access. Bull mainframes are fitted with Mainway for translating DSA to TCP/IP and VIP devices are replaced by TNVIP Terminal Emulations (GLink, Winsurf). GCOS 8 supports TNVIP SE over TCP/IP.

Chapter- 7

Next Generation Network and Open Access Network

Next generation network

A **Next generation network** (**NGN**) is a broad term to describe key architectural evolutions in telecommunication core and access networks. The general idea behind the NGN is that one network transports all information and services (voice, data, and all sorts of media such as video) by encapsulating these into packets, like it is on the Internet. NGNs are commonly built around the Internet Protocol, and therefore the term "all-IP" is also sometimes used to describe the transformation toward NGN.

Description

Challenges
for New Generation Network
Seoul, Korea
Hanyang University, December 9th, 2009
Hiroaki Harai (harai@nict.go.jp)
NICT (National Institute of Information and Communications Technology)
http://akari-project.nict.go.jp/
NiCT

NGN Seminar in Fusion Technology Center by NICT(Japan) researcher

According to ITU-T, the definition is:

> A Next generation network (NGN) is a packet-based network which can provide services including Telecommunication Services and able to make use of multiple broadband, Quality of Service-enabled transport technologies and in which service-related functions are independent from underlying transport-related technologies. It offers unrestricted access by users to different service providers. It supports generalized mobility which will allow consistent and ubiquitous provision of services to users..

From a practical perspective, NGN involves three main architectural changes that need to be looked at separately:

- In the core network, NGN implies a consolidation of several (dedicated or overlay) transport networks each historically built for a different service into one core transport network (often based on IP and Ethernet). It implies amongst others the migration of voice from a circuit-switched architecture (PSTN) to VoIP, and also migration of legacy services such as X.25, Frame Relay (either commercial migration of the customer to a new service like IP VPN, or technical emigration by emulation of the "legacy service" on the NGN).
- In the wired access network, NGN implies the migration from the dual system of legacy voice next to xDSL setup in local exchanges to a converged setup in which the DSLAMs integrate *voice ports* or VoIP, making it possible to remove the voice switching infrastructure from the exchange.
- In the cable access network, NGN convergence implies migration of constant bit rate voice to CableLabs PacketCable standards that provide VoIP and SIP services. Both services ride over DOCSIS as the cable data layer standard.

In an NGN, there is a more defined separation between the transport (connectivity) portion of the network and the services that run on top of that transport. This means that whenever a provider wants to enable a new service, they can do so by defining it directly at the service layer without considering the transport layer - i.e. services are independent of transport details. Increasingly applications, including voice, tend to be independent of the access network (de-layering of network and applications) and will reside more on end-user devices (phone, PC, set-top box).

Underlying technology components

Next Generation Networks are based on Internet technologies including Internet Protocol (IP) and Multiprotocol Label Switching (MPLS). At the application level, Session Initiation Protocol (SIP) seems to be taking over from ITU-T H.323.

Initially H.323 was the most popular protocol, though its popularity decreased in the "local loop" due to its original poor traversal of Network address translation (NAT) and firewalls. For this reason as domestic VoIP services have been developed, SIP has been more widely adopted. However in voice networks where everything is under the control of the network operator or telco, many of the largest carriers use H.323 as the protocol of choice in their core backbones. So really SIP is a useful tool for the "local loop" and H.323 is like the "fiber backbone". With the most recent changes introduced for H.323, it is now possible for H.323 devices to easily and consistently traverse NAT and firewall devices, opening up the possibility that H.323 may again be looked upon more favorably in cases where such devices encumbered its use previously. Nonetheless, most of the telcos are extensively researching and supporting IP Multimedia Subsystem (IMS), which gives SIP a major chance of being the most widely adopted protocol.

For voice applications one of the most important devices in NGN is a Softswitch - a programmable device that controls Voice over IP (VoIP) calls. It enables correct integration of different protocols within NGN. The most important function of the Softswitch is creating the interface to the existing telephone network, PSTN, through Signalling Gateways and Media Gateways. However, the Softswitch as a term may be defined differently by the different equipment manufacturers and have somewhat different functions.

One may quite often find the term Gatekeeper in NGN literature. This was originally a VoIP device, which converted (using gateways) voice and data from their analog or digital switched-circuit form (PSTN, SS7) to the packet-based one (IP). It controlled one or more gateways. As soon as this kind of device started using the Media Gateway Control Protocol, the name was changed to Media Gateway Controller (MGC).

A Call Agent is a general name for devices/systems controlling calls.

The IP Multimedia Subsystem (IMS) is a standardised NGN architecture for an Internet media-services capability defined by the European Telecommunications Standards Institute (ETSI) and the 3rd Generation Partnership Project (3GPP).

Implementations

In the UK another popular acronym was introduced by BT (British Telecom) as 21CN (21st Century Networks, sometimes mistakenly quoted as C21N) — this is another loose term for NGN and denotes BT's initiative to deploy and operate NGN switches and networks in the period 2006-2008 (the aim being by 2008 BT to have only all-IP switches in their network)

The first company in the UK to roll out a NGN was THUS plc which started deployment back in 1999. THUS' NGN contains 10,600 km of fibre optic cable with more than 190 points of presence throughout the UK. The core optical network uses Dense Wave Division Multiplexing (DWDM) technology to provide scalability to many hundreds of gigabits per second of bandwidth, in line with growth demand. On top of this, the THUS backbone network uses MPLS technology to deliver the highest possible performance. IP/MPLS-based services carry voice, video and data traffic across a converged infrastructure, potentially allowing organisations to enjoy lower infrastructure costs, as well as added flexibility and functionality. Traffic can be prioritised with Classes of

Service, coupled with Service Level Agreements (SLAs) that underpin quality of service performance guarantees. The THUS NGN accommodates seven Classes of Service, four of which are currently offered on MPLS IP VPN.

In the Netherlands, KPN is developing a NGN network in a network transformation program called all-IP — this is another loose term for NGN that is increasingly used. Next Generation Networks also extends into the messaging domain and in Ireland, Openmind Networks has designed, built and deployed Traffic Control to handle the demands and requirements of all IP networks.

In Bulgaria, BTC (Bulgarian Telecommunications Company) has implemented the NGN as underlying network of its telco services on a large scale project in 2004. The inherent flexibility and scalability of the new core network approach resulted in an unprecedented rise of classical services deployment as POTS/ISDN, Centrex, ADSL, VPN, as well as implementation of higher bandwidths for the Metro and Long-distance Ethernet / VPN services, cross-national transits and WebTV/IPTV application.

In Israel, Bezeq announced in a June 2009 press release the move to NGN in selected areas. The service will allow enhanced services to phone subscribers as well as upgraded speed capabilities for ADSL users (up to 50Mbps DL, 1000Kbps UL).

In Canada, upstart Wind Mobile owned by Globalive is deploying an all-ip wireless backbone for its mobile phone service.

Open Access Network

In telecommunications, **Open Access Network** (**OAN**) refers to horizontally layered network architecture and business model that separates physical access to the network from service provisioning. The same OAN will be used by a number of different providers that share the investments and maintenance cost.

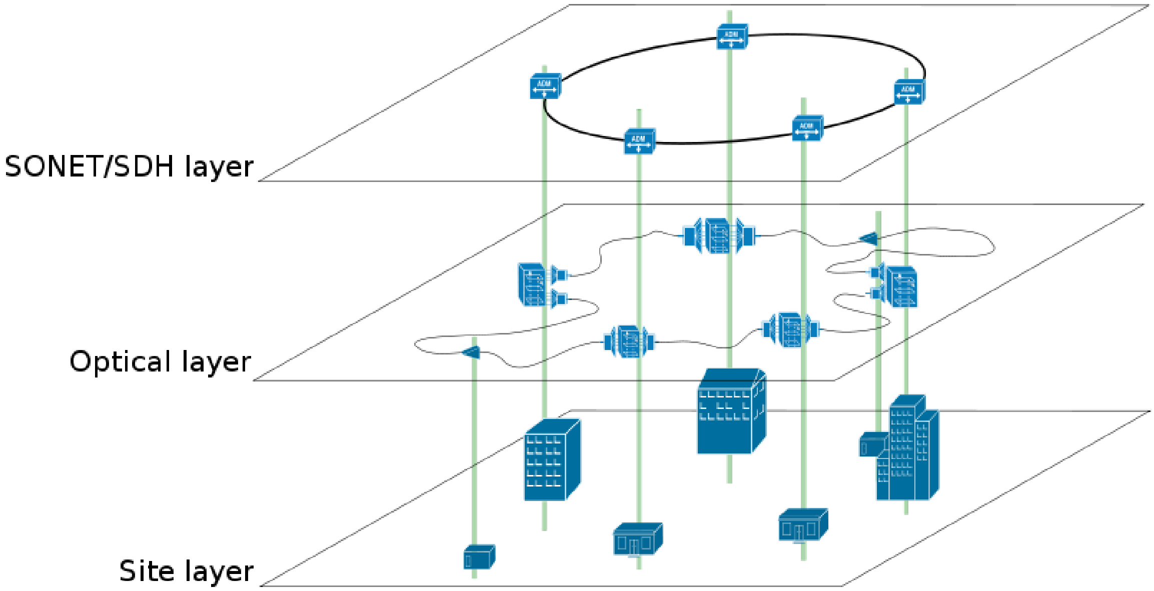
The OAN concept is appropriate for both fiber and WiFi Access Networks, especially where exclusivity can not be allowed. The shared maintenance costs make it very appropriate for distant rural areas, where traditional ISP are reluctant to offer their services. An open access network uses a different business model than traditional telecom networks. In the twentieth century, analog telephone and cable TV networks were designed around the limitations of the technology; copper-based twisted pair networks were not able to carry TV programming, and copper-based coaxial networks were not able to carry voice telephony. Near the end of the twentieth century, with the rise of packet-based switching (i.e. the Internet) and IP-based fiber and wireless technologies, it became possible to design, build, and operate a single high performance network capable of delivering dozens or even hundreds of services from multiple, competing providers.

Open access networks are also viewed as a feasible way of deploying next-generation broadband networks in low population density areas where service providers cannot obtain a sufficient return on investment to cover the high costs associated with trenching, right-of-way encroachment permits, and the requisite network infrastructure. In contrast to traditional municipal networks where the municipality owns the network and there is only one service provider, the open access model allows multiple service providers to compete over the same network at wholesale prices. This allows service providers to make money in the short-term and the municipality or cooperative to recoup its costs over the long-term. The build-out and infrastructure is typically financed through low-cost bonds. Open access networks have been very successful in the U.S., Europe, and Asia. One of the best known and most mature open access networks is Vasteras, city of about 40,000 homes in Sweden. The Vasteras open access network has dozens of providers and more than a hundred services available on the network.

In the United States, open access networks like The Wired Road have been able to attract both local and regional service providers quickly, and the cost of Internet access and telephone service for business users in The Wired Road service area have declined by 50% to 70% because of the increased competition between providers. The Wired Road is a municipal network owned by the counties of Carroll and Grayson and the City of Galax, and is operated as a regional authority. The Wired Road Broadband Authority sells no services, so it does not compete with private sector providers. The Wired Road provides open access transport to any service provider that meets minimum technical and financial qualifications, and incumbent providers are also able to use the system to deliver enhanced services to existing and new customers.

Chapter- 8

Optical Mesh Network



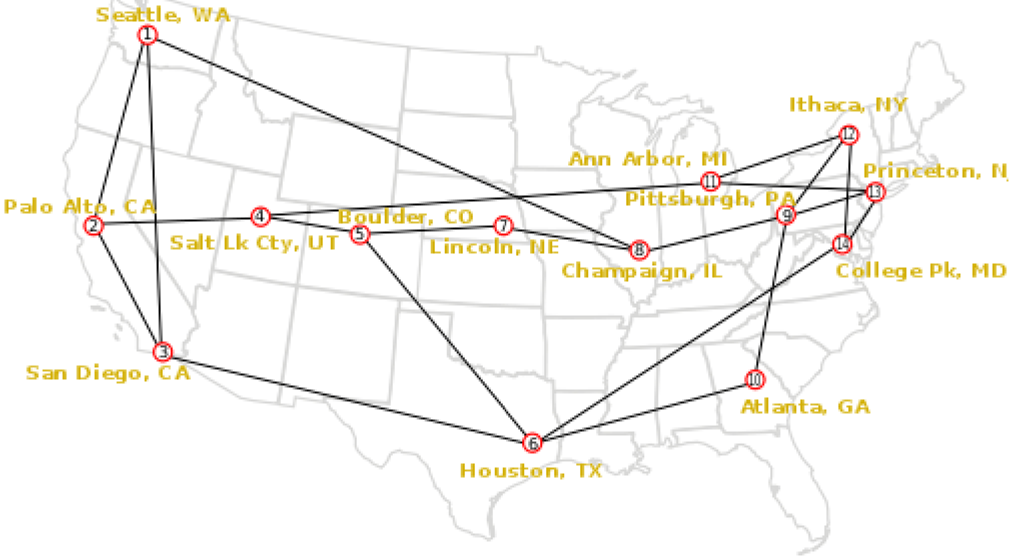
Transport network based on SONET/SDH ring architecture

Optical mesh networks are a type of telecommunications network.

Transport networks, the underlying optical fiber-based layer of telecommunications networks, have evolved from DCS (Digital Cross-connect Systems)-based mesh architectures in the 1980s, to SONET/SDH (Synchronous Optical Networking/Synchronous Digital Hierarchy) ring architectures in the 1990s. Technological advancements in optical transport equipment in the first decade of the 21st century, along with continuous deployment of DWDM systems, have led telecommunications service providers to replace their SONET ring architectures by mesh-based architectures. The new optical mesh networks support the same fast recovery previously available in ring networks while achieving better capacity efficiency and resulting in lower capital cost.

Optical mesh networks today not only provide trunking capacity to higher-layer networks, such as inter-router or inter-switch connectivity in an IP, MPLS, or Ethernet-centric infrastructure, but also support efficient routing and fast failure recovery of high-bandwidth services. This was made possible by the emergence of optical network elements that have the intelligence required to automatically control certain network functions, such as fault recovery.

Optical mesh networks enable a variety of dynamic services such as bandwidth-on-demand, Just-In-Time bandwidth, bandwidth scheduling, bandwidth brokering, and optical virtual private networks that open up new opportunities for service providers and their customers alike.



Example of mesh network: NSFNET 14nodes

History of transport networks

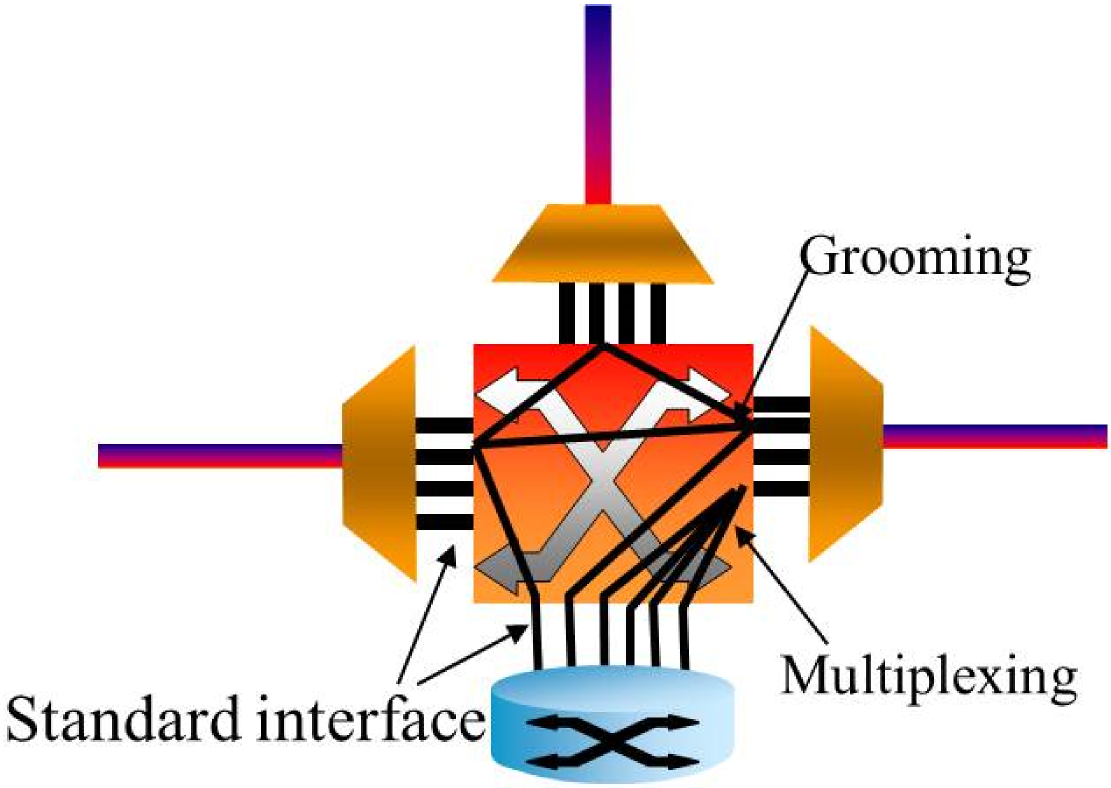
Transport networks, the underlying optical fiber-based layer of telecommunications networks, have evolved from Digital cross connect system (DCS)-based mesh architectures in the 1980s, to SONET/SDH (Synchronous Optical Networking/Synchronous Digital Hierarchy) ring architectures in the 1990s. In DCS-based mesh architectures, telecommunications carriers deployed restoration systems for DS3 circuits such as at&t FASTAR (FAST Automatic Restoration) and MCI Real Time Restoration (RTR), restoring circuits in minutes after a network failure. In SONET/SDH rings, carriers implemented ring protection such as SONET Universal Path Switched Ring (UPSR) (also called Sub-Network Connection Protection (SCNP) in SDH networks) or SONET Bidirectional Line Switched Ring (BLSR) (also called Multiplex Section - Shared Protection Ring (MS-SPRing) in SDH networks), protecting against and recovering from a network failure in 50 msecs or less, a significant improvement over the recovery time supported in DCS-based mesh restoration, and a key driver for the deployment of SONET/SDH ring-based protection.

There have been attempts at improving and/or evolving traditional ring architectures to overcome some of its limitations, with trans-oceanic ring architecture (ITU-T Rec.

G.841), “P-cycles” protection, next-generation SONET/SDH equipment that can handle multiple rings, or have the ability to not close the working or protection ring side, or to share protection capacity among rings (e.g., with Virtual Line Switched Ring (VLSR).

Technological advancements in optical transport switches in the first decade of the 21st century, along with continuous deployment of dense wavelength-division multiplexing (DWDM) systems, have led telecommunications service providers to replace their SONET ring architectures by mesh-based architectures for new traffic. The new optical mesh networks support the same fast recovery previously available in ring networks while achieving better capacity efficiency and resulting in lower capital cost. Such fast recovery (in the 10's to 100's of msecs) in case of failures (e.g., network link or node failure) is achieved through the intelligence embedded in these new optical transport equipment, which allows recovery to be automatic and handled within the network itself as part of the network control plane, without relying on an external network management system.

Optical mesh networks

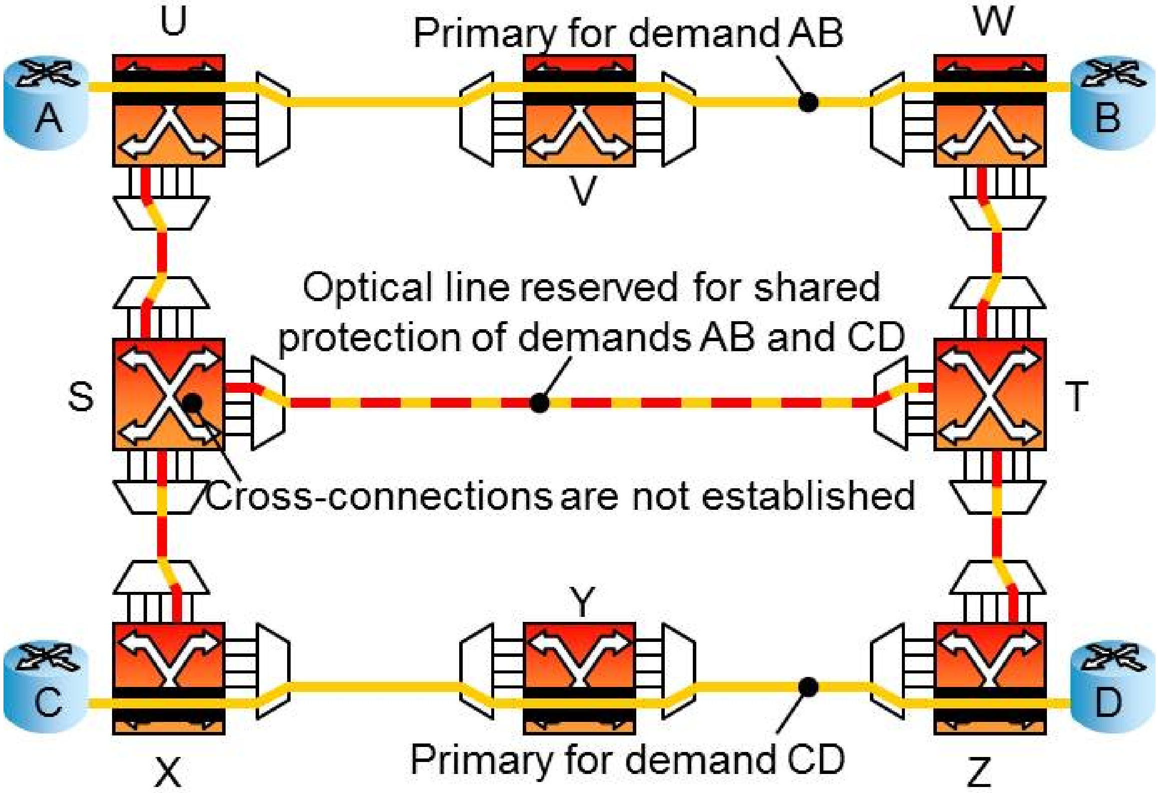


Switching, multiplexing, and grooming of traffic in an OEO device

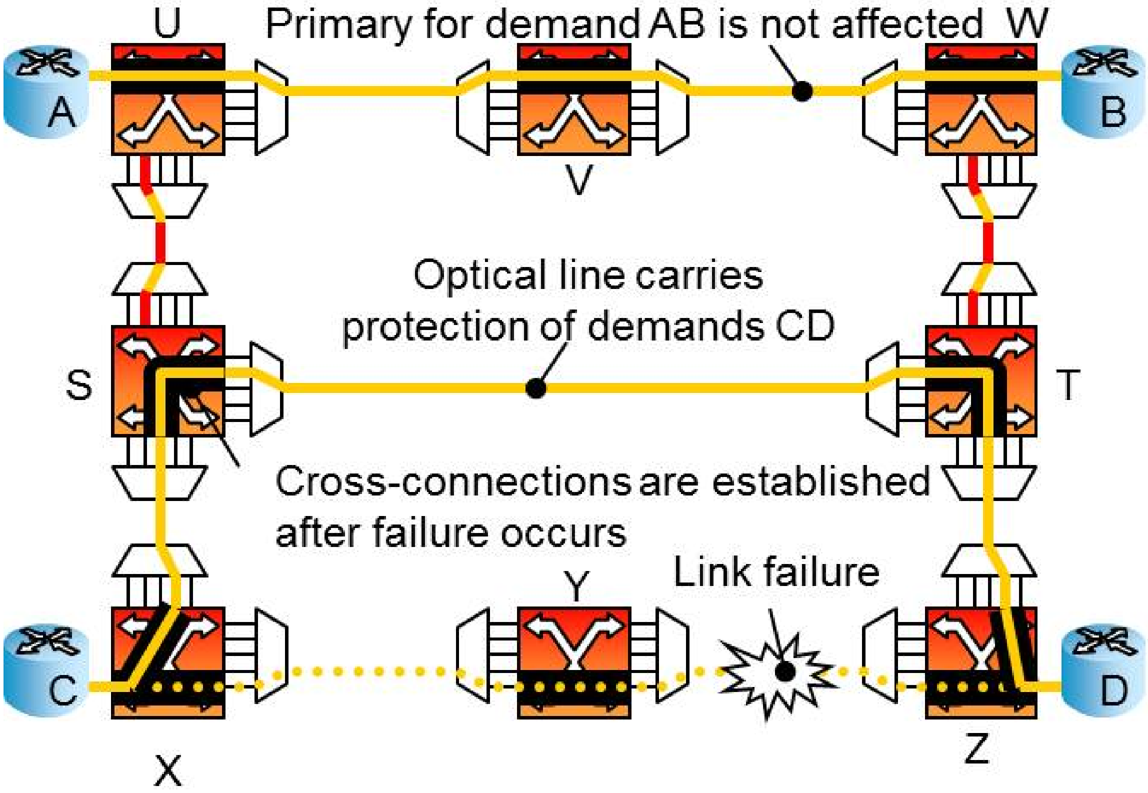
Optical mesh networks refer to transport networks that are built directly off the mesh-like fiber infrastructure deployed in metropolitan, regional, national, or international (e.g., trans-oceanic) areas by deploying optical transport equipment that are capable of switching traffic (at the wavelength or sub-wavelength level) from an incoming fiber to an outgoing fiber. In addition to switching wavelengths, the equipment is typically also able to multiplex lower speed traffic into wavelengths for transport, and to groom traffic. Finally, these equipment also provide for the recovery of traffic in case of a network failure. As most of the transport networks evolve toward mesh topologies utilizing intelligent network elements (optical cross-connects or optical switches) for provisioning and recovery of services, new approaches have been developed for the design, deployment, operations and management of mesh optical networks.

Optical mesh networks today not only provide trunking capacity to higher-layer networks, such as inter-router or inter-switch connectivity in an IP, MPLS, or Ethernet-centric packet infrastructure, but also support efficient routing and fast failure recovery of high-bandwidth point-to-point Ethernet and SONET/SDH services.

Recovery in optical mesh networks


U
Primary for demand AB
W
A
B
V
Optical line reserved for shared protection of demands AB and CD
S
T
Cross-connections are not established
Y
C
D
X
Primary for demand CD
Z

Shared backup path protection - before failure

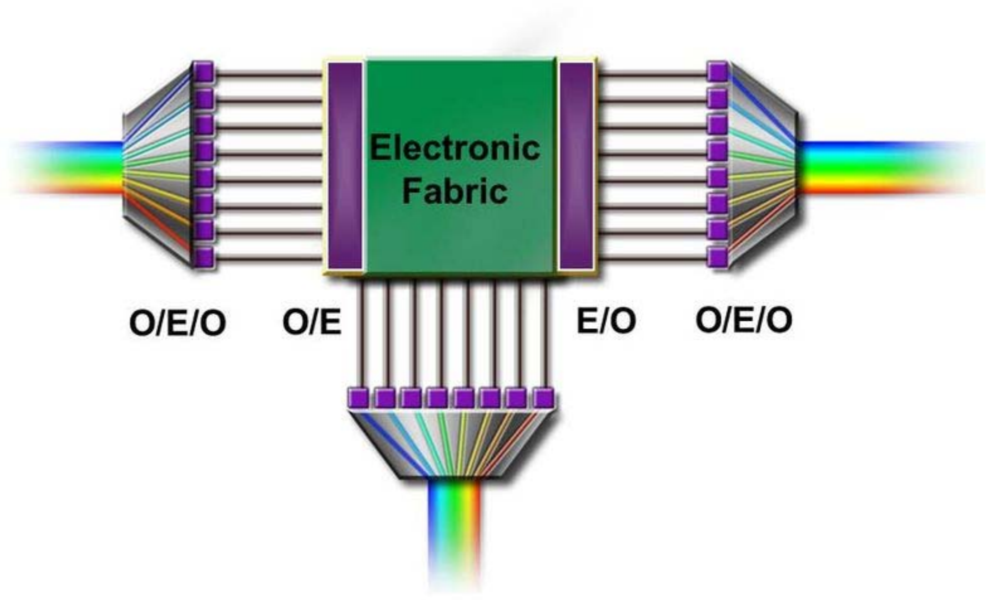

U
Primary for demand AB is not affected
W
A
B
V
Optical line carries protection of demands CD
S
T
Cross-connections are established after failure occurs
Y
Link failure
C
D
X
Z

Shared backup path protection - after failure and recovery

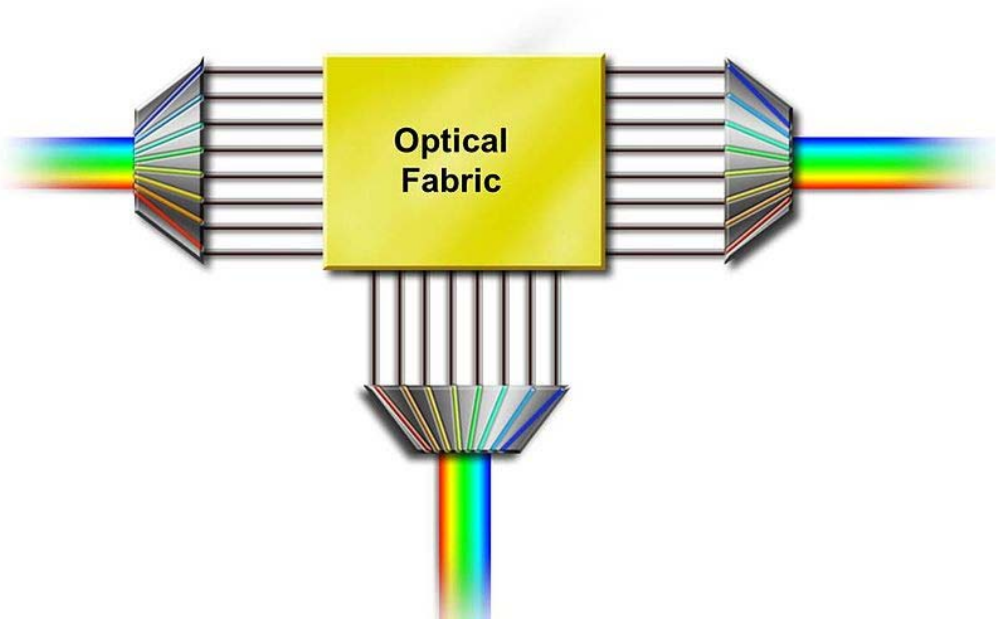
Optical mesh networks support the establishment of circuit-mode connection-oriented services. Multiple recovery mechanisms that provide different levels of protection or restoration against different failure modes are available in mesh networks. Channel, link, segment and path protection are the most common protection schemes. P-cycles is another type of protection that leverages and extends ring-based protection. Restoration is another recovery method that can work on its own or complement faster protection schemes in case of multiple failures.

In path-protected mesh networks, some connections can be unprotected; others can be protected against single or multiple failures in various ways. A connection can be protected against a single failure by defining a backup path, diverse from the primary path taken by the connection over the mesh network. The backup path and associated resources can be dedicated to the connection (aka Dedicated Backup Path Protection), or shared among multiple connections (aka Shared Backup Path Protection), typically ones whose primary paths are not likely to fail at the same time, thereby avoiding contention for the shared resources in case of a single link or node failure. A number of other protection schemes such as the use of pre-emptible paths, or only partially diverse backup paths, can be implemented. Finally, multiple diverse routes can be designed so that a connection has multiple recovery routes and can recover even after multiple failures (examples of mesh networks across the Atlantic and Pacific oceans).

Transparency



Opaque switching of traffic between fiber links



Transparent switching of traffic between fiber links

Traditional transport networks are made of optical fiber-based links between telecommunications offices, where multiple wavelengths are multiplexed to increase the capacity of the fiber. The wavelengths are terminated on electronic devices called transponders, undergoing an optical-to-electrical conversion for signal Reamplification, Reshaping, and Retiming (3R). Inside a telecommunications office, the signals are then handled to and switched by a transport switch (aka optical cross-connect or optical switch) and either are dropped at that office, or directed to an outgoing fiber link where they are again carried as wavelengths multiplexed into that fiber link towards the next telecommunications office. The act of going through Optical-Electrical-Optical (O-E-O) conversion through a telecommunications office causes the network to be considered opaque. When the incoming wavelengths do not undergo an optical-to-electrical conversion and are switched through a telecommunications office in the optical domain using all-optical switches (also called photonic cross-connect, optical add-drop multiplexer, or Reconfigurable Optical Add-Drop Multiplexer (ROADM) systems), the network is considered to be transparent. Hybrid schemes can provide limited O-E-O conversions at key locations across the network.

Transparent optical mesh networks have been deployed in metropolitan and regional networks. In 2010, operational long distance networks still tend to remain opaque.

Routing in optical mesh networks

Routing is a key control and operational aspect of optical mesh networks. In transparent or all-optical networks, routing of connections is tightly linked to the wavelength selection and assignment process (so-called routing and wavelength assignment, or "RWA"). This is due to the fact that the connection remains on the same wavelength from end-to-end throughout the network (sometimes referred to as wavelength continuity constraint, in the absence of devices that can translate between wavelengths in the optical domain). In an opaque network, the routing problem is one of finding a primary path for a connection and if protection is needed, a backup path diverse from the primary path. Wavelengths are used on each link independently of each other's. Several algorithms can be used to determine a primary path and a diverse backup path (with or without sharing of resource along the backup path) for a connection or service, such as shortest path, including Dijkstra's algorithm, k-shortest path, edge and node-diverse or disjoint routing, including Suurballe's algorithm, and numerous heuristics.

Applications

The deployment of optical mesh networks is enabling new services and applications for service providers to offer their customers, such as

- Dynamic services such as Bandwidth-on-Demand (BoD), Just-In-Time (JIT) bandwidth, bandwidth scheduling, and bandwidth brokering
- Optical virtual private networks

It also supports new network paradigms such as

- IP-over-optical network architectures

Chapter- 9

Internet Protocol Suite

The **Internet Protocol Suite** is the set of communications protocols used for the Internet and other similar networks. It is commonly also known as **TCP/IP**, named from two of the most important protocols in it: the Transmission Control Protocol (TCP) and the Internet Protocol (IP), which were the first two networking protocols defined in this standard. Modern IP networking represents a synthesis of several developments that began to evolve in the 1960s and 1970s, namely the Internet and local area networks, which emerged during the 1980s, together with the advent of the World Wide Web in the early 1990s.

The Internet Protocol Suite, like many protocol suites, is constructed as a set of layers. Each layer solves a set of problems involving the transmission of data.

Every layer provides a well-defined service to the upper layer protocols and relies on using services from the lower layers. The upper layers provide a higher level of abstraction, being closer to the application and the end user. The lower layer protocols are more concerned providing specific solutions to deal with the actual physical transmission of the data.

The TCP/IP model consists of 4 layers (RFC 1122). From lowest to highest, these are the Link Layer, the Internet Layer, the Transport Layer, and the Application Layer.

History

The Internet Protocol Suite resulted from research and development conducted by the Defense Advanced Research Projects Agency (DARPA) in the early 1970s. After initiating the pioneering ARPANET in 1969, DARPA started work on a number of other data transmission technologies. In 1972, Robert E. Kahn joined the DARPA Information Processing Technology Office, where he worked on both satellite packet networks and ground-based radio packet networks, and recognized the value of being able to communicate across both. In the spring of 1973, Vinton Cerf, the developer of the existing ARPANET Network Control Program (NCP) protocol, joined Kahn to work on open-architecture interconnection models with the goal of designing the next protocol generation for the ARPANET.

By the summer of 1973, Kahn and Cerf had worked out a fundamental reformulation, where the differences between network protocols were hidden by using a common internetwork protocol, and, instead of the network being responsible for reliability, as in the ARPANET, the hosts became responsible. Cerf credits Hubert Zimmerman and Louis Pouzin, designer of the CYCLADES network, with important influences on this design.

The design of the network included the recognition that it should provide only the functions of efficiently transmitting and routing traffic between end nodes and that all other intelligence should be located at the edge of the network, in the end nodes. Using a simple design, it became possible to connect almost any network to the ARPANET, irrespective of their local characteristics, thereby solving Kahn's initial problem. One popular expression is that TCP/IP, the eventual product of Cerf and Kahn's work, will run over "*two tin cans and a string.*"

A computer called a *router* (a name changed from *gateway* to avoid confusion with other types of *gateway*s) is provided with an interface to each network, and forwards packets back and forth between them. Requirements for routers are defined in (Request for Comments 1812).

The idea was worked out in more detailed form by Cerf's networking research group at Stanford in the 1973–74 period, resulting in the first TCP specification.(Request for Comments 675) (The early networking work at Xerox PARC, which produced the PARC Universal Packet protocol suite, much of which existed around the same period of time, was also a significant technical influence; people moved between the two.)

DARPA then contracted with BBN Technologies, Stanford University, and the University College London to develop operational versions of the protocol on different hardware platforms. Four versions were developed: TCP v1, TCP v2, a split into TCP v3 and IP v3 in the spring of 1978, and then stability with TCP/IP v4 — the standard protocol still in use on the Internet today.

In 1975, a two-network TCP/IP communications test was performed between Stanford and University College London (UCL). In November, 1977, a three-network TCP/IP test was conducted between sites in the US, UK, and Norway. Several other TCP/IP prototypes were developed at multiple research centres between 1978 and 1983. The migration of the ARPANET to TCP/IP was officially completed on flag day January 1, 1983, when the new protocols were permanently activated.

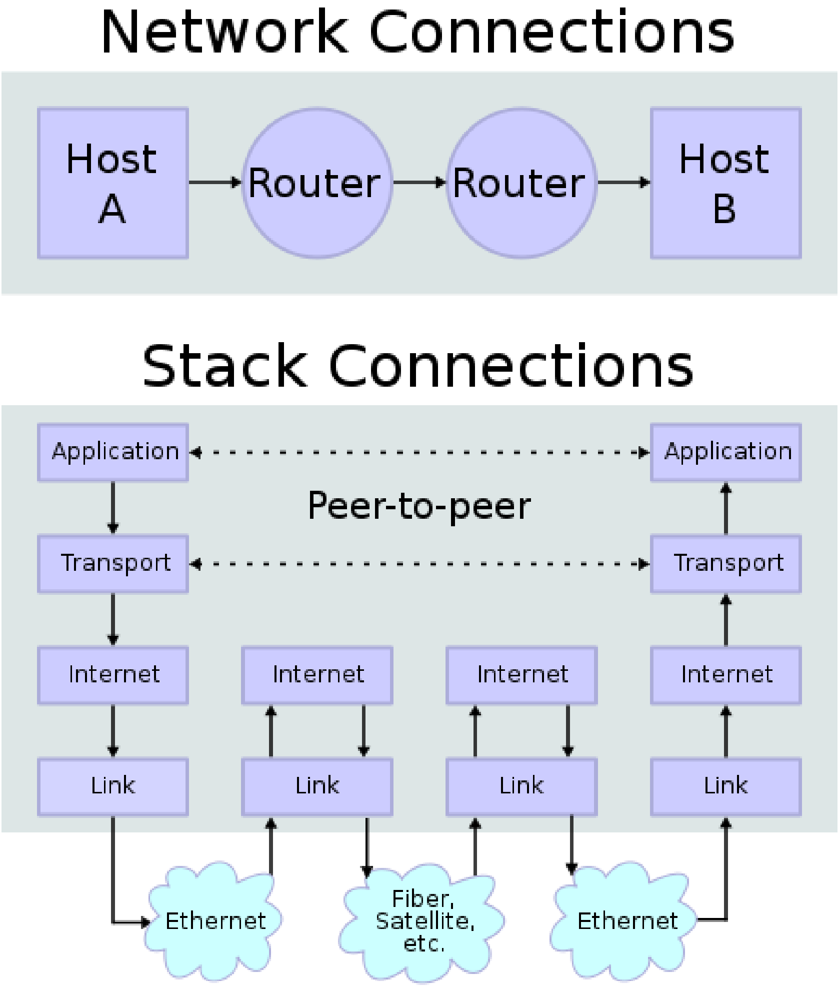
In March 1982, the US Department of Defense declared TCP/IP as the standard for all military computer networking. In 1985, the Internet Architecture Board held a three day workshop on TCP/IP for the computer industry, attended by 250 vendor representatives, promoting the protocol and leading to its increasing commercial use.

Layers in the Internet Protocol Suite

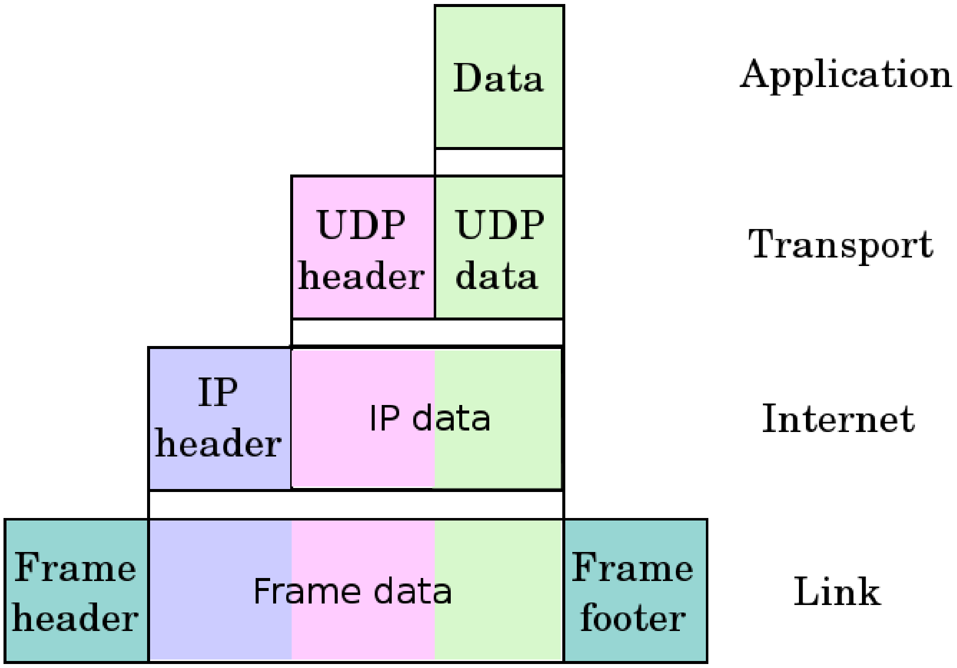
The concept of layers

The TCP/IP suite uses encapsulation to provide abstraction of protocols and services. Such encapsulation usually is aligned with the division of the protocol suite into layers of general functionality. In general, an application (the highest level of the model) uses a set of protocols to send its data down the layers, being further encapsulated at each level.

This may be illustrated by an example network scenario, in which two Internet host computers communicate across local network boundaries constituted by their internetworking gateways (routers).



TCP/IP stack operating on two hosts connected via two routers and the corresponding layers used at each hop



Encapsulation of application data descending through the protocol stack

The functional groups of protocols and methods are the Application Layer, the Transport Layer, the Internet Layer, and the Link Layer (RFC 1122). This model was not intended to be a rigid reference model into which new protocols have to fit in order to be accepted as a standard.

The following table provides some examples of the protocols grouped in their respective layers.

Application	DNS, TFTP, TLS/SSL, FTP, Gopher, HTTP, IMAP, IRC, NNTP, POP3, SIP, SMTP, SMPP, SNMP, SSH, Telnet, Echo, RTP, PNRP, rlogin, ENRP Routing protocols like BGP and RIP which run over TCP/UDP, may also be considered part of the Internet Layer.
Transport	TCP, UDP, DCCP, SCTP, IL, RUDP, RSVP
Internet	IP (IPv4, IPv6), ICMP, IGMP, and ICMPv6 OSPF for IPv4 was initially considered IP layer protocol since it runs per IP-subnet, but has been placed on the Link since RFC 2740.
Link	ARP, RARP, OSPF (IPv4/IPv6), IS-IS, NDP

Layer names and number of layers in the literature

The following table shows the layer names and the number of layers of networking models presented in RFCs and textbooks in widespread use in today's university computer networking courses.

RFC 1122	Tanenbaum	Cisco Academy	Kurose Forouzan	Comer Kozierok	Stallings	Arpanet Reference Model 1982 (RFC 871)
Four layers	*Four layers*	*Four layers*	*Five layers*	*Four+one layers*	*Five layers*	*Three layers*
"Internet model"	"TCP/IP reference model"	"Internet model"	"Five-layer Internet model" or "TCP/IP protocol suite"	"TCP/IP 5-layer reference model"	"TCP/IP model"	"Arpanet reference model"
Application	Application	Application	Application	Application	Application	Application/Process
Transport	Transport	Transport	Transport	Transport	Host-to-host or transport	Host-to-host
Internet	Internet	Internetwork	Network	Internet	Internet	
Link	Host-to-network	Network interface	Data link	Data link (Network interface)	Network access	Network interface
			Physical	(Hardware)	Physical	

These textbooks are secondary sources that may contravene the intent of RFC 1122 and other IETF primary sources.

Different authors have interpreted the RFCs differently regarding the question whether the Link Layer (and the TCP/IP model) covers Physical Layer issues, or if a hardware layer is assumed below the Link Layer. Some authors have tried to use other names for the Link Layer, such as *network interface layer*, in view to avoid confusion with the Data Link Layer of the seven layer OSI model. Others have attempted to map the Internet Protocol model onto the OSI Model. The mapping often results in a model with five layers where the Link Layer is split into a Data Link Layer on top of a Physical Layer. In literature with a bottom-up approach to Internet communication, in which hardware issues are emphasized, those are often discussed in terms of Physical Layer and Data Link Layer.

The Internet Layer is usually directly mapped into the OSI Model's Network Layer, a more general concept of network functionality. The Transport Layer of the TCP/IP model, sometimes also described as the host-to-host layer, is mapped to OSI Layer 4 (Transport Layer), sometimes also including aspects of OSI Layer 5 (Session Layer) functionality. OSI's Application Layer, Presentation Layer, and the remaining functionality of the Session Layer are collapsed into TCP/IP's Application Layer. The argument is that these OSI layers do usually not exist as separate processes and protocols in Internet applications.

However, the Internet protocol stack has never been altered by the Internet Engineering Task Force from the four layers defined in RFC 1122. The IETF makes no effort to follow the OSI model although RFCs sometimes refer to it. The IETF has repeatedly stated that Internet protocol and architecture development is not intended to be OSI-compliant.

RFC 3439, addressing Internet architecture, contains a section entitled: "Layering Considered Harmful".

Implementations

Most computer operating systems in use today, including all consumer-targeted systems, include a TCP/IP implementation.

Minimally acceptable implementation includes implementation for (from most essential to the less essential) IP, ARP, ICMP, UDP, TCP and sometime IGMP. It is in principle possible to support only one of transport protocols (i.e. simple UDP), but it is rarely done, as it limits usage of the whole implementation. IPv6, beyond own version of ARP (NBP), and ICMP (ICMPv6), and IGMP (IGMPv6) have some additional required functionalities, and often is accompanied with integrated IPSec security layer. Other protocols could be easily added later (often they can be implemented entirely in the userspace), for example DNS for resolving domain names to IP addresses or DHCP client for automatic configuration of network interfaces.

Most of the IP implementations are accessible to the programmers using socket abstraction (usable also with other protocols) and proper API for most of the operations. This interface is known as BSD sockets and was used initially in C.

Unique implementations include Lightweight TCP/IP, an open source stack designed for embedded systems and KA9Q NOS, a stack and associated protocols for amateur packet radio systems and personal computers connected via serial lines.

Chapter- 10

Network Planning & Design and TCP/IP Model

Network planning and design

Network planning and design is an iterative process, encompassing topological design, network-synthesis, and network-realization, and is aimed at ensuring that a new network or service meets the needs of the subscriber and operator. The process can be tailored according to each new network or service.

This is an extremely important process which must be performed before the establishment of a new telecommunications network or service.

A network planning methodology

A traditional network planning methodology involves five layers of planning, namely:

- business planning
- long-term and medium-term network planning
- short-term network planning
- IT asset sourcing
- operations and maintenance.

Each of these layers incorporates plans for different time horizons, i.e. the business planning layer determines the planning that the operator must perform to ensure that the network will perform as required for its intended life-span. The Operations and Maintenance layer, however, examines how the network will run on a day-to-day basis.

The network planning process begins with the acquisition of external information. This includes:

- forecasts of how the new network/service will operate;
- the economic information concerning costs; and
- the technical details of the network's capabilities.

It should be borne in mind that planning a new network/service involves implementing the new system across the first four layers of the OSI Reference Model. This means that even before the network planning process begins, choices must be made, involving protocols and transmission technologies.

Once the initial decisions have been made, the network planning process involves three main steps:

- **Topological design**: This stage involves determining where to place the components and how to connect them. The (topological) optimisation methods that can be used in this stage come from an area of mathematics called Graph Theory. These methods involve determining the costs of transmission and the cost of switching, and thereby determining the optimum connection matrix and location of switches and concentrators.

- **Network-synthesis**: This stage involves determining the size of the components used, subject to performance criteria such as the Grade of Service (GoS). The method used is known as "Nonlinear Optimisation", and involves determining the topology, required GoS, cost of transmission, etc., and using this information to calculate a routing plan, and the size of the components.

- **Network realization**: This stage involves determining how to meet capacity requirements, and ensure reliability within the network. The method used is known as "Multicommodity Flow Optimisation", and involves determining all information relating to demand, costs and reliability, and then using this information to calculate an actual physical circuit plan.

These steps are interrelated and are therefore performed iteratively, and in parallel with one another. The planning process is highly complex, meaning that at each iteration, an analyst must increase his planning horizons, and in so doing, he must generate plans for the various layers outlined above.

The role of forecasting

During the process of Network Planning and Design, it is necessary to estimate the expected traffic intensity and thus the traffic load that the network must support. If a network of a similar nature already exists, then it may be possible to take traffic measurements of such a network and use that data to calculate the exact traffic load. However, as is more likely in most instances, if there are no similar networks to be found, then the network planner must use telecommunications forecasting methods to estimate the expected traffic intensity.

The forecasting process involves several steps as follows:

- Definition of problem;
- Data acquisition;

- Choice of forecasting method;
- Analysis/Forecasting;
- Documentation and analysis of results.

Dimensioning

The purpose of dimensioning a new network/service is to determine the minimum capacity requirements that will still allow the Teletraffic Grade of Service (GoS) requirements to be met. To do this, dimensioning involves planning for peak-hour traffic, i.e. that hour during the day during which traffic intensity is at its peak.

The dimensioning process involves determining the network's topology, routing plan, traffic matrix, and GoS requirements, and using this information to determine the maximum call handling capacity of the switches, and the maximum number of channels required between the switches.. This process requires a complex model that simulates the behavior of the network equipment and routing protocols.

A dimensioning rule is that the planner must ensure that the traffic load should never approach a load of 100 percent. To calculate the correct dimensioning to comply with the above rule, the planner must take on-going measurements of the network's traffic, and continuously maintain and upgrade resources to meet the changing requirements.. Another reason for "overprovisioning" is to make sure that traffic can be rerouted in case a failure occurs in the network.

Because of the complexity of network dimensioning, this is typically done using specialized software tools. Whereas researchers typically develop custom software to study a particular problem, network operators typically make use of commercial network planning software (e.g. OPNET Technologies, SevOne, WANDL, VPISystems, Cariden, Aria Networks). However, there is one notable open source network planning software available by the name of TOTEM named after TOolbox for Traffic Engineering Methods.

Traffic engineering

Comparing to network engineering, which adds resources such as links, routers and switches into the network, traffic engineering targets to change traffic paths on the existing network to alleviate traffic congestion or accommodate more traffic demand.

This technology is critical when the cost of network expansion is prohibitively high and network load is not optimally balanced. The first part provides financial motivation for traffic engineering while the second part grants the possibility of deploying this technology.

The available technologies for traffic engineering include MPLS and ATM for current Internet backbone. For example, MPLS allows carriers to provision LSPs with dynamic or explicit routes. The dynamic routes is controlled by CSPF while the explicit routes are optimized in an offline tool or through a path computation element which is under study by IETF. Fast reroute has been implemented by major vendors, such as Cisco and Juniper Networks, to provide localized resilient capability for MPLS networks. End-to-end protection is an alternative resilient approach. It provisions a backup route for each primary route. Pre-planning enough bandwidth for these backup routes is one of the active topic for survivable network design.

Provisioning a large number of LSPs also brought up a scalability problem. Various solutions have been proposed and it is still an active topic under study.

Survivability

Network survivability enables the network to maintain maximum network connectivity and quality of service under failure conditions. It has been one of the critical requirements in network planning and design. It involves design requirements on topology, protocol, bandwidth allocation, etc.. Topology requirement can be maintaining a minimum two-connected network against any failure of a single link or node. Protocol requirements include using dynamic routing protocol to reroute traffic against network dynamics during the transition of network dimensioning or equipment failures. Bandwidth allocation requirements pro-actively allocate extra bandwidth to avoid traffic loss under failure conditions. This topic has been actively studied in conferences, such as the International Workshop on Design of Reliable Communication Networks DRCN.

TCP/IP model

The **TCP/IP model** is a description framework for computer network protocols created in the 1970s by DARPA, an agency of the United States Department of Defense. It evolved from ARPANET, which was the world's first wide area network and a predecessor of the Internet. The TCP/IP Model is sometimes called the *Internet Model* or the *DoD Model*.

The TCP/IP model, or Internet Protocol Suite, describes a set of general design guidelines and implementations of specific networking protocols to enable computers to communicate over a network. TCP/IP provides end-to-end connectivity specifying how data should be formatted, addressed, transmitted, routed and received at the destination. Protocols exist for a variety of different types of communication services between computers.

TCP/IP, sometimes referred to as the *Internet model*, has four abstraction layers as defined in RFC 1122. This layer architecture is often compared with the seven-layer OSI Reference Model; using terms such as *Internet reference model*, incorrectly, however, because it is descriptive while the OSI Reference Model was intended to be prescriptive, hence being a reference model.

The TCP/IP model and related protocols are maintained by the Internet Engineering Task Force (IETF).

Key architectural principles

An early architectural document, RFC 1122, emphasizes architectural principles over layering.

- End-to-End Principle: This principle has evolved over time. Its original expression put the maintenance of state and overall intelligence at the edges, and assumed the Internet that connected the edges retained no state and concentrated on speed and simplicity. Real-world needs for firewalls, network address translators, web content caches and the like have forced changes in this principle.
- Robustness Principle: "In general, an implementation must be conservative in its sending behavior, and liberal in its receiving behavior. That is, it must be careful to send well-formed datagrams, but must accept any datagram that it can interpret (e.g., not object to technical errors where the meaning is still clear)." "The second part of the principle is almost as important: software on other hosts may contain deficiencies that make it unwise to exploit legal but obscure protocol features."

Even when the layers are examined, the assorted architectural documents—there is no single architectural model such as ISO 7498, the OSI reference model—have fewer and less rigidly-defined layers than the OSI model, and thus provide an easier fit for real-world protocols. In point of fact, one frequently referenced document, RFC 1958, does not contain a stack of layers. The lack of emphasis on layering is a strong difference between the IETF and OSI approaches. It only refers to the existence of the "internetworking layer" and generally to "upper layers"; this document was intended as a 1996 "snapshot" of the architecture: "The Internet and its architecture have grown in evolutionary fashion from modest beginnings, rather than from a Grand Plan. While this process of evolution is one of the main reasons for the technology's success, it nevertheless seems useful to record a snapshot of the current principles of the Internet architecture."

RFC 1122, entitled *Host Requirements*, is structured in paragraphs referring to layers, but the document refers to many other architectural principles not emphasizing layering. It loosely defines a four-layer model, with the layers having names, not numbers, as follows:

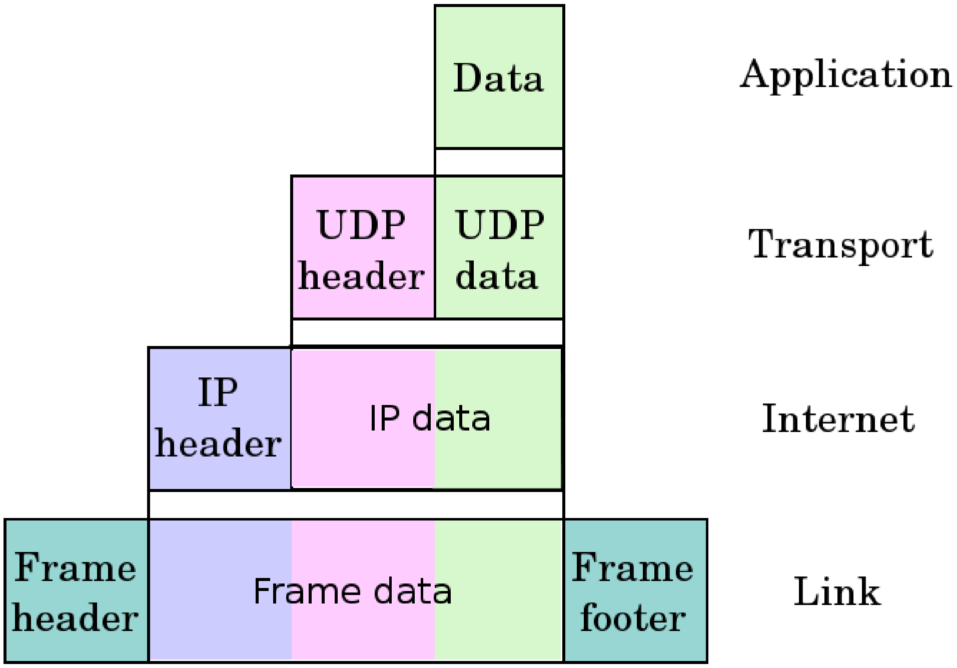
- Application Layer (process-to-process): This is the scope within which applications create user data and communicate this data to other processes or applications on another or the same host. The communications partners are often called *peers*. This is where the "higher level" protocols such as SMTP, FTP, SSH, HTTP, etc. operate.
- Transport Layer (host-to-host): The Transport Layer constitutes the networking regime between two network hosts, either on the local network or on remote networks separated by routers. The Transport Layer provides a uniform

networking interface that hides the actual topology (layout) of the underlying network connections. This is where flow-control, error-correction, and connection protocols exist, such as TCP. This layer deals with opening and maintaining connections between Internet hosts.

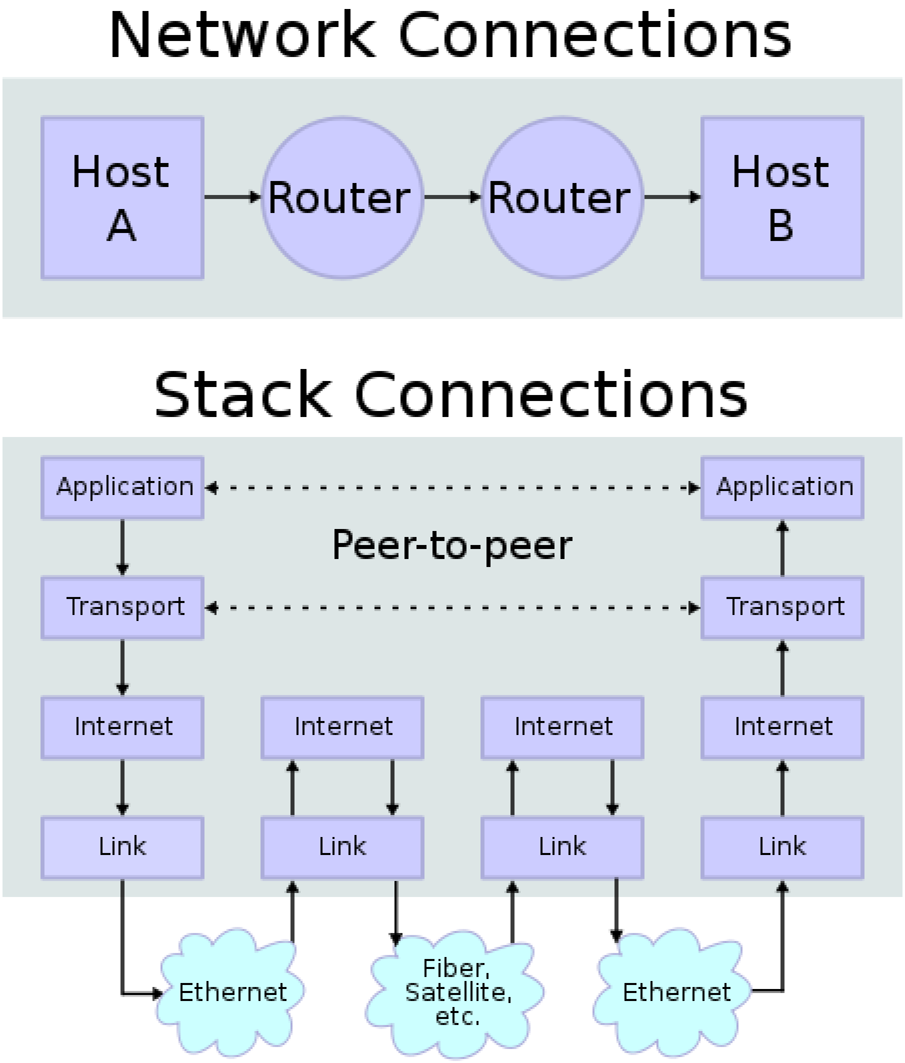
- Internet Layer (internetworking): The Internet Layer has the task of exchanging datagrams across network boundaries. It is therefore also referred to as the layer that establishes internetworking, indeed, it defines and establishes the Internet. This layer defines the addressing and routing structures used for the TCP/IP protocol suite. The primary protocol in this scope is the Internet Protocol, which defines IP addresses. Its function in routing is to transport datagrams to the next IP router that has the connectivity to a network closer to the final data destination.
- Link Layer: This layer defines the networking methods with the scope of the local network link on which hosts communicate without intervening routers. This layer describes the protocols used to describe the local network topology and the interfaces needed to affect transmission of Internet Layer datagrams to next-neighbor hosts. (cf. the OSI Data Link Layer).

The Internet Protocol Suite and the layered protocol stack design were in use before the OSI model was established. Since then, the TCP/IP model has been compared with the OSI model in books and classrooms, which often results in confusion because the two models use different assumptions, including about the relative importance of strict layering.

Layers in the TCP/IP model



Encapsulation of application data descending through the TCP/IP layers



Two Internet hosts connected via two routers and the corresponding layers used at each hop

The layers near the top are logically closer to the user application, while those near the bottom are logically closer to the physical transmission of the data. Viewing layers as providing or consuming a service is a method of abstraction to isolate upper layer protocols from the nitty-gritty detail of transmitting bits over, for example, Ethernet and collision detection, while the lower layers avoid having to know the details of each and every application and its protocol.

This abstraction also allows upper layers to provide services that the lower layers cannot, or choose not to, provide. Again, the original OSI Reference Model was extended to include connectionless services (OSIRM CL). For example, IP is not designed to be reliable and is a best effort delivery protocol. This means that all transport layer implementations must choose whether or not to provide reliability and to what degree. UDP provides data integrity (via a checksum) but does not guarantee delivery; TCP provides both data integrity and delivery guarantee (by retransmitting until the receiver acknowledges the reception of the packet).

This model lacks the formalism of the OSI reference model and associated documents, but the IETF does not use a formal model and does not consider this a limitation, as in the comment by David D. Clark, "We reject: kings, presidents and voting. We believe in: rough consensus and running code." Criticisms of this model, which have been made with respect to the OSI Reference Model, often do not consider ISO's later extensions to that model.

1. For multiaccess links with their own addressing systems (e.g. Ethernet) an address mapping protocol is needed. Such protocols can be considered to be below IP but above the existing link system. While the IETF does not use the terminology, this is a subnetwork dependent convergence facility according to an extension to the OSI model, the Internal Organization of the Network Layer (IONL).
2. ICMP & IGMP operate on top of IP but do not transport data like UDP or TCP. Again, this functionality exists as layer management extensions to the OSI model, in its *Management Framework* (OSIRM MF)
3. The SSL/TLS library operates above the transport layer (uses TCP) but below application protocols. Again, there was no intention, on the part of the designers of these protocols, to comply with OSI architecture.
4. The link is treated like a black box here. This is fine for discussing IP (since the whole point of IP is it will run over virtually anything). The IETF explicitly does not intend to discuss transmission systems, which is a less academic but practical alternative to the OSI Reference Model.

The following is a description of each layer in the TCP/IP networking model starting from the lowest level.

Link Layer

The Link Layer is the networking scope of the local network connection to which a host is attached. This regime is called the *link* in Internet literature. This is the lowest component layer of the Internet protocols, as TCP/IP is designed to be hardware independent. As a result TCP/IP has been implemented on top of virtually any hardware networking technology in existence.

The Link Layer is used to move packets between the Internet Layer interfaces of two different hosts on the same link. The processes of transmitting and receiving packets on a given link can be controlled both in the software device driver for the network card, as well as on firmware or specialized chipsets. These will perform data link functions such as adding a packet header to prepare it for transmission, then actually transmit the frame over a physical medium. The TCP/IP model includes specifications of translating the network addressing methods used in the Internet Protocol to data link addressing, such as

Media Access Control (MAC), however all other aspects below that level are implicitly assumed to exist in the Link Layer, but are not explicitly defined.

The Link Layer is also the layer where packets may be selected to be sent over a virtual private network or other networking tunnel. In this scenario, the Link Layer data may be considered application data which traverses another instantiation of the IP stack for transmission or reception over another IP connection. Such a connection, or virtual link, may be established with a transport protocol or even an application scope protocol that serves as a tunnel in the Link Layer of the protocol stack. Thus, the TCP/IP model does not dictate a strict hierarchical encapsulation sequence.

Internet Layer

The Internet Layer solves the problem of sending packets across one or more networks. Internetworking requires sending data from the source network to the destination network. This process is called routing.

In the Internet Protocol Suite, the Internet Protocol performs two basic functions:

- *Host addressing and identification*: This is accomplished with a hierarchical addressing system.
- *Packet routing*: This is the basic task of getting packets of data (datagrams) from source to destination by sending them to the next network node (router) closer to the final destination.

IP can carry data for a number of different upper layer protocols. These protocols are each identified by a unique protocol number: for example, Internet Control Message Protocol (ICMP) and Internet Group Management Protocol (IGMP) are protocols 1 and 2, respectively.

Some of the protocols carried by IP, such as ICMP (used to transmit diagnostic information about IP transmission) and IGMP (used to manage IP Multicast data) are layered on top of IP but perform internetworking functions. This illustrates the differences in the architecture of the TCP/IP stack of the Internet and the OSI model.

Transport Layer

The Transport Layer's responsibilities include end-to-end message transfer capabilities independent of the underlying network, along with error control, segmentation, flow control, congestion control, and application addressing (port numbers). End to end message transmission or connecting applications at the transport layer can be categorized as either connection-oriented, implemented in Transmission Control Protocol (TCP), or connectionless, implemented in User Datagram Protocol (UDP).

The Transport Layer can be thought of as a transport mechanism, e.g., a vehicle with the responsibility to make sure that its contents (passengers/goods) reach their destination safely and soundly, unless another protocol layer is responsible for safe delivery.

The Transport Layer provides this service of connecting applications through the use of service ports. Since IP provides only a best effort delivery, the Transport Layer is the first layer of the TCP/IP stack to offer reliability. IP can run over a reliable data link protocol such as the High-Level Data Link Control (HDLC). Protocols above transport, such as RPC, also can provide reliability.

For example, the Transmission Control Protocol (TCP) is a connection-oriented protocol that addresses numerous reliability issues to provide a reliable byte stream:

- data arrives in-order
- data has minimal error (i.e. correctness)
- duplicate data is discarded
- lost/discarded packets are resent
- includes traffic congestion control

The newer Stream Control Transmission Protocol (SCTP) is also a reliable, connection-oriented transport mechanism. It is Message-stream-oriented — not byte-stream-oriented like TCP — and provides multiple streams multiplexed over a single connection. It also provides multi-homing support, in which a connection end can be represented by multiple IP addresses (representing multiple physical interfaces), such that if one fails, the connection is not interrupted. It was developed initially for telephony applications (to transport SS7 over IP), but can also be used for other applications.

User Datagram Protocol is a connectionless datagram protocol. Like IP, it is a best effort, "unreliable" protocol. Reliability is addressed through error detection using a weak checksum algorithm. UDP is typically used for applications such as streaming media (audio, video, Voice over IP etc) where on-time arrival is more important than reliability, or for simple query/response applications like DNS lookups, where the overhead of setting up a reliable connection is disproportionately large. Real-time Transport Protocol (RTP) is a datagram protocol that is designed for real-time data such as streaming audio and video.

TCP and UDP are used to carry an assortment of higher-level applications. The appropriate transport protocol is chosen based on the higher-layer protocol application. For example, the File Transfer Protocol expects a reliable connection, but the Network File System (NFS) assumes that the subordinate Remote Procedure Call protocol, not transport, will guarantee reliable transfer. Other applications, such as VoIP, can tolerate some loss of packets, but not the reordering or delay that could be caused by retransmission.

The applications at any given network address are distinguished by their TCP or UDP port. By convention certain *well known ports* are associated with specific applications.

Application Layer

The Application Layer refers to the higher-level protocols used by most applications for network communication. Examples of application layer protocols include the File Transfer Protocol (FTP) and the Simple Mail Transfer Protocol (SMTP). Data coded according to application layer protocols are then encapsulated into one or (occasionally) more transport layer protocols (such as the Transmission Control Protocol (TCP) or User Datagram Protocol (UDP)), which in turn use lower layer protocols to effect actual data transfer.

Since the IP stack defines no layers between the application and transport layers, the application layer must include any protocols that act like the OSI's presentation and session layer protocols. This is usually done through libraries.

Application Layer protocols generally treat the transport layer (and lower) protocols as "black boxes" which provide a stable network connection across which to communicate, although the applications are usually aware of key qualities of the transport layer connection such as the end point IP addresses and port numbers. As noted above, layers are not necessarily clearly defined in the Internet protocol suite. Application layer protocols are most often associated with client–server applications, and the commoner servers have specific ports assigned to them by the IANA: HTTP has port 80; Telnet has port 23; etc. Clients, on the other hand, tend to use ephemeral ports, i.e. port numbers assigned at random from a range set aside for the purpose.

Transport and lower level layers are largely unconcerned with the specifics of application layer protocols. Routers and switches do not typically "look inside" the encapsulated traffic to see what kind of application protocol it represents, rather they just provide a conduit for it. However, some firewall and bandwidth throttling applications do try to determine what's inside, as with the Resource Reservation Protocol (RSVP). It's also sometimes necessary for Network Address Translation (NAT) facilities to take account of the needs of particular application layer protocols. (NAT allows hosts on private networks to communicate with the outside world via a single visible IP address using port forwarding, and is an almost ubiquitous feature of modern domestic broadband routers).

Hardware and software implementation

Normally, application programmers are concerned only with interfaces in the Application Layer and often also in the Transport Layer, while the layers below are services provided by the TCP/IP stack in the operating system. Microcontroller firmware in the network adapter typically handles link issues, supported by driver software in the operational system. Non-programmable analog and digital electronics are normally in charge of the physical components in the Link Layer, typically using an application-specific integrated circuit (ASIC) chipset for each network interface or other physical standard.

However, hardware or software implementation is not stated in the protocols or the layered reference model. High-performance routers are to a large extent based on fast non-programmable digital electronics, carrying out link level switching.

OSI and TCP/IP layering differences

The three top layers in the OSI model—the Application Layer, the Presentation Layer and the Session Layer—are not distinguished separately in the TCP/IP model where it is just the Application Layer. While some pure OSI protocol applications, such as X.400, also combined them, there is no *requirement* that a TCP/IP protocol stack needs to impose monolithic architecture above the Transport Layer. For example, the Network File System (NFS) application protocol runs over the eXternal Data Representation (XDR) presentation protocol, which, in turn, runs over a protocol with Session Layer functionality, Remote Procedure Call (RPC). RPC provides reliable record transmission, so it can run safely over the best-effort User Datagram Protocol (UDP) transport.

The Session Layer roughly corresponds to the Telnet virtual terminal functionality, which is part of text based protocols such as the HTTP and SMTP TCP/IP model Application Layer protocols. It also corresponds to TCP and UDP port numbering, which is considered as part of the transport layer in the TCP/IP model. Some functions that would have been performed by an OSI presentation layer are realized at the Internet application layer using the MIME standard, which is used in application layer protocols such as HTTP and SMTP.

Since the IETF protocol development effort is not concerned with strict layering, some of its protocols may not appear to fit cleanly into the OSI model. These conflicts, however, are more frequent when one only looks at the original OSI model, ISO 7498, without looking at the annexes to this model (e.g., ISO 7498/4 Management Framework), or the ISO 8648 Internal Organization of the Network Layer (IONL). When the IONL and Management Framework documents are considered, the ICMP and IGMP are neatly defined as layer management protocols for the network layer. In like manner, the IONL provides a structure for "subnetwork dependent convergence facilities" such as ARP and RARP.

IETF protocols can be encapsulated recursively, as demonstrated by tunneling protocols such as Generic Routing Encapsulation (GRE). While basic OSI documents do not consider tunneling, there is some concept of tunneling in yet another extension to the OSI architecture, specifically the transport layer gateways within the International Standardized Profile framework. The associated OSI development effort, however, has been abandoned given the overwhelming adoption of TCP/IP protocols.

Layer names and number of layers in the literature

The following table shows the layer names and the number of layers of networking models presented in RFCs and textbooks in widespread use in today's university computer networking courses.

Kurose, Forouzan	**Comer, Kozierok**	**Stallings**	**Tanenbaum**	**RFC 1122, Internet STD 3 (1989)**	**Cisco Academy**	**Mike Padlipsky's 1982 "Arpanet Reference Model" (RFC 871)**
Five layers	*Four+one layers*	*Five layers*	*Four layers*	*Four layers*	*Four layers*	*Three layers*
"Five-layer Internet model" or "TCP/IP protocol suite"	"TCP/IP 5-layer reference model"	"TCP/IP model"	"TCP/IP reference model"	"Internet model"	"Internet model"	"Arpanet reference model"
Application	Application	Application	Application	Application	Application	Application/Process
Transport	Transport	Host-to-host or transport	Transport	Transport	Transport	Host-to-host
Network	Internet	Internet	Internet	Internet	Internetwork	
Data link	Data link (Network interface)	Network access	Host-to-network	Link	Network interface	Network interface
Physical	(Hardware)	Physical				

These textbooks are secondary sources that may contravene the intent of RFC 1122 and other IETF primary sources such as RFC 3439.

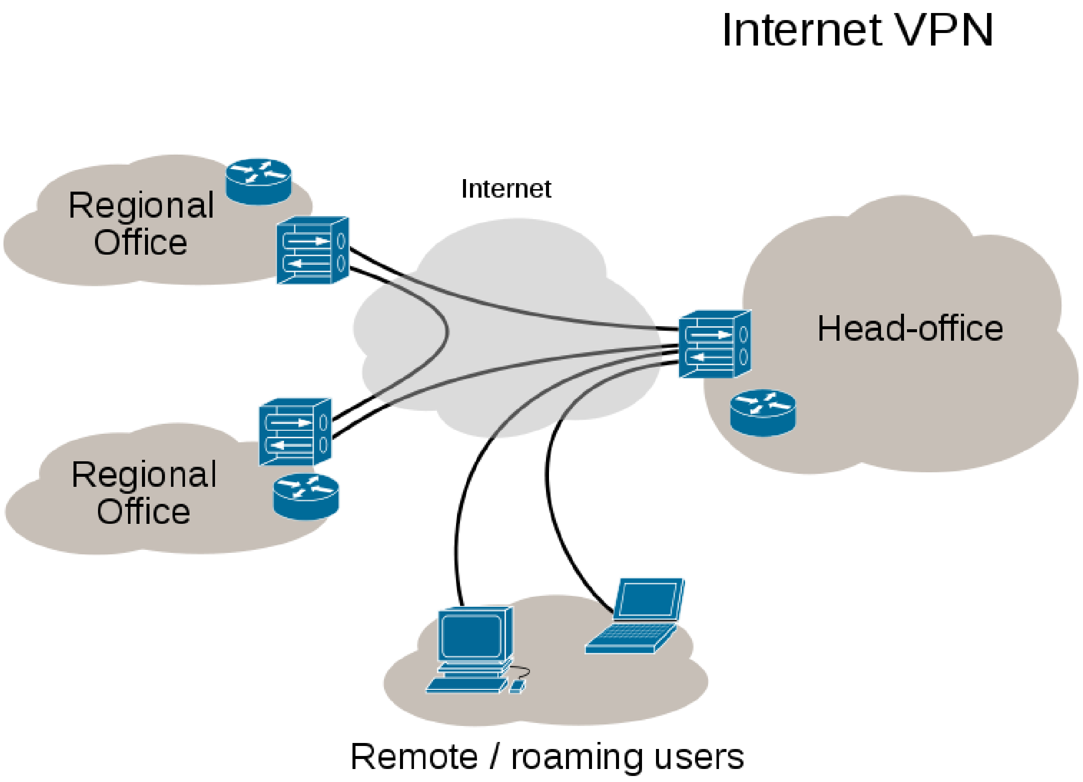
Different authors have interpreted the RFCs differently regarding the question whether the Link Layer (and the TCP/IP model) covers Physical Layer issues, or if a hardware layer is assumed below the Link Layer. Some authors have tried to use other names for the Link Layer, such as *network interface layer*, in view to avoid confusion with the Data Link Layer of the seven layer OSI model. Others have attempted to map the Internet Protocol model onto the OSI Model. The mapping often results in a model with five layers where the Link Layer is split into a Data Link Layer on top of a Physical Layer. In literature with a bottom-up approach to Internet communication, in which hardware issues are emphasized, those are often discussed in terms of physical layer and data link layer.

The Internet Layer is usually directly mapped into the OSI Model's Network Layer, a more general concept of network functionality. The Transport Layer of the TCP/IP model, sometimes also described as the host-to-host layer, is mapped to OSI Layer 4 (Transport Layer), sometimes also including aspects of OSI Layer 5 (Session Layer) functionality. OSI's Application Layer, Presentation Layer, and the remaining functionality of the Session Layer are collapsed into TCP/IP's Application Layer. The argument is that these OSI layers do usually not exist as separate processes and protocols in Internet applications.

However, the Internet protocol stack has never been altered by the Internet Engineering Task Force from the four layers defined in RFC 1122. The IETF makes no effort to follow the OSI model although RFCs sometimes refer to it and often use the old OSI layer numbers. The IETF has repeatedly stated that Internet protocol and architecture development is not intended to be OSI-compliant. RFC 3439, addressing Internet architecture, contains a section entitled: "Layering Considered Harmful".

Chapter- 11

Virtual Private Network



VPN Connectivity overview

A **virtual private network** (**VPN**) is a computer network that uses a public telecommunication infrastructure such as the Internet to provide remote offices or individual users with secure access to their organization's network. It aims to avoid an expensive system of owned or leased lines that can be used by only one organization.

It encapsulates data transfers between two or more networked devices which are not on the same private network so as to keep the transferred data private from other devices on one or more intervening local or wide area networks. There are many different classifications, implementations, and uses for VPNs.

History

Until the end of the 1990s networked computers were connected through expensive leased lines and/or dial-up phone lines.

Virtual Private Networks reduce network costs because they avoid a need for many leased lines that individually connect to the Internet. Users can exchange private data securely, making the expensive leased lines unnecessary.

VPN technologies have myriad protocols, terminologies and marketing influences that define them. For example, VPN technologies can differ in:

- The protocols they use to tunnel the traffic
- The tunnel's termination point, i.e., customer edge or network provider edge
- Whether they offer site-to-site or remote access connectivity
- The levels of security provided
- The OSI layer they present to the connecting network, such as Layer 2 circuits or Layer 3 network connectivity

Some classification schemes are discussed in the following sections.

Security Mechanisms

Secure VPNs use cryptographic tunneling protocols to provide confidentiality by blocking intercepts and packet sniffing, allowing sender authentication to block identity spoofing, and provide message integrity by preventing message alteration.

Secure VPN protocols include the following:

- IPsec (Internet Protocol Security) was originally developed for IPv6, which requires it. This standards-based security protocol is also widely used with IPv4. L2TP frequently runs over IPsec.
- Transport Layer Security (SSL/TLS) can tunnel an entire network's traffic, as it does in the OpenVPN project, or secure an individual connection. A number of vendors provide remote access VPN capabilities through SSL. An SSL VPN can connect from locations where IPsec runs into trouble with Network Address Translation and firewall rules.
- Datagram Transport Layer Security (DTLS), is used in Cisco's next-generation VPN product, Cisco AnyConnect VPN, to solve the issues SSL/TLS has with tunneling over TCP.
- Microsoft Point-to-Point Encryption (MPPE) works with their PPTP and in several compatible implementations on other platforms.
- Microsoft introduced Secure Socket Tunneling Protocol (SSTP) in Windows Server 2008 and Windows Vista Service Pack 1. SSTP tunnels Point-to-Point Protocol (PPP) or L2TP traffic through an SSL 3.0 channel.

- MPVPN (Multi Path Virtual Private Network). Ragula Systems Development Company owns the registered trademark "MPVPN".
- Secure Shell (SSH) VPN -- OpenSSH offers VPN tunneling to secure remote connections to a network or inter-network links. This should not be confused with port forwarding. OpenSSH server provides limited number of concurrent tunnels and the VPN feature itself does not support personal authentication.

Authentication

Tunnel endpoints must authenticate before secure VPN tunnels can establish.

User-created remote access VPNs may use passwords, biometrics, two-factor authentication or other cryptographic methods.

Network-to-network tunnels often use passwords or digital certificates, as they permanently store the key to allow the tunnel to establish automatically and without intervention.

Routing

Tunneling protocols can be used in a point-to-point topology that would theoretically not be considered a VPN, because a VPN by definition is expected to support arbitrary and changing sets of network nodes. But since most router implementations support a software-defined tunnel interface, customer-provisioned VPNs often are simply defined tunnels running conventional routing protocols.

On the other hand provider-provided VPNs (PPVPNs) need to support coexisting multiple VPNs, hidden from one another, but operated by the same service provider.

PPVPN Building blocks

Depending on whether the PPVPN runs in layer 2 or layer 3, the building blocks described below may be L2 only, L3 only, or combine them both. Multiprotocol Label Switching (MPLS) functionality blurs the L2-L3 identity.

RFC 4026 generalized the following terms to cover L2 and L3 VPNs, but they were introduced in RFC 2547.

Customer edge device. (CE)

a device at the customer premises, that provides access to the PPVPN. Sometimes it's just a demarcation point between provider and customer responsibility. Other providers allow customers to configure it.

Provider edge device (PE)

A PE is a device, or set of devices, at the edge of the provider network, that presents the provider's view of the customer site. PEs are aware of the VPNs that connect through them, and maintain VPN state.

Provider device (P)

A P device operates inside the provider's core network, and does not directly interface to any customer endpoint. It might, for example, provide routing for many provider-operated tunnels that belong to different customers' PPVPNs. While the P device is a key part of implementing PPVPNs, it is not itself VPN-aware and does not maintain VPN state. Its principal role is allowing the service provider to scale its PPVPN offerings, as, for example, by acting as an aggregation point for multiple PEs. P-to-P connections, in such a role, often are high-capacity optical links between major locations of provider.

User-visible PPVPN services

This section deals with the types of VPN considered in the IETF; some historical names were replaced by these terms.

OSI Layer 1 services

Virtual private wire and private line services (VPWS and VPLS)

In both of these services, the service provider does not offer a full routed or bridged network, but provides components to build customer-administered networks. VPWS are point-to-point while VPLS can be point-to-multipoint. They can be Layer 1 emulated circuits with no data link structure.

The customer determines the overall customer VPN service, which also can involve routing, bridging, or host network elements.

An unfortunate acronym confusion can occur between Virtual Private Line Service and Virtual Private LAN Service; the context should make it clear whether "VPLS" means the layer 1 virtual private line or the layer 2 virtual private LAN.

OSI Layer 2 services

Virtual LAN

A Layer 2 technique that allows for the coexistence of multiple LAN broadcast domains, interconnected via trunks using the IEEE 802.1Q trunking protocol. Other trunking protocols have been used but have become obsolete, including Inter-Switch Link (ISL), IEEE 802.10 (originally a security protocol but a subset was introduced for trunking), and ATM LAN Emulation (LANE).

Virtual private LAN service (VPLS)

Developed by IEEE, VLANs allow multiple tagged LANs to share common trunking. VLANs frequently comprise only customer-owned facilities. The former is a layer 1 technology that supports emulation of both point-to-point and point-to-multipoint topologies. The method discussed here extends Layer 2 technologies such as 802.1d and 802.1q LAN trunking to run over transports such as Metro Ethernet.

As used in this context, a VPLS is a Layer 2 PPVPN, rather than a private line, emulating the full functionality of a traditional local area network (LAN). From a user standpoint, a VPLS makes it possible to interconnect several LAN segments over a packet-switched, or optical, provider core; a core transparent to the user, making the remote LAN segments behave as one single LAN.

In a VPLS, the provider network emulates a learning bridge, which optionally may include VLAN service.

Pseudo wire (PW)

PW is similar to VPWS, but it can provide different L2 protocols at both ends. Typically, its interface is a WAN protocol such as Asynchronous Transfer Mode or Frame Relay. In contrast, when aiming to provide the appearance of a LAN contiguous between two or more locations, the Virtual Private LAN service or IPLS would be appropriate.

IP-only LAN-like service (IPLS)

A subset of VPLS, the CE devices must have L3 capabilities; the IPLS presents packets rather than frames. It may support IPv4 or IPv6.

OSI Layer 3 PPVPN architectures

Here we, discusses the main architectures for PPVPNs, one where the PE disambiguates duplicate addresses in a single routing instance, and the other, virtual router, in which the PE contains a virtual router instance per VPN. The former approach, and its variants, have gained the most attention.

One of the challenges of PPVPNs involves different customers using the same address space, especially the IPv4 private address space. The provider must be able to disambiguate overlapping addresses in the multiple customers' PPVPNs.

BGP/MPLS PPVPN

In the method defined by RFC 2547, BGP extensions advertise routes in the IPv4 VPN address family, which are of the form of 12-byte strings, beginning with an 8-byte Route Distinguisher (RD) and ending with a 4-byte IPv4 address. RDs disambiguate otherwise duplicate addresses in the same PE.

PEs understand the topology of each VPN, which are interconnected with MPLS tunnels, either directly or via P routers. In MPLS terminology, the P routers are Label Switch Routers without awareness of VPNs.

Virtual router PPVPN

The Virtual Router architecture, as opposed to BGP/MPLS techniques, requires no modification to existing routing protocols such as BGP. By the provisioning of logically independent routing domains, the customer operating a VPN is completely responsible for the address space. In the various MPLS tunnels, the different PPVPNs are disambiguated by their label, but do not need routing distinguishers.

Virtual router architectures do not need to disambiguate addresses, because rather than a PE router having awareness of all the PPVPNs, the PE contains multiple virtual router instances, which belong to one and only one VPN.

Plaintext Tunnels

Some virtual networks may not use encryption to protect the data contents. While VPNs often provide security, an unencrypted overlay network does not neatly fit within the secure or trusted categorization. For example a tunnel set up between two hosts that used Generic Routing Encapsulation (GRE) would in fact be a virtual private network, but neither secure nor trusted.

Besides the GRE example above, native plaintext tunneling protocols include Layer 2 Tunneling Protocol (L2TP) when it is set up without IPsec and Point-to-Point Tunneling Protocol (PPTP) or Microsoft Point-to-Point Encryption (MPPE).

Trusted delivery networks

Trusted VPNs do not use cryptographic tunneling, and instead rely on the security of a single provider's network to protect the traffic.

- Multi-Protocol Label Switching (MPLS) is often used to overlay VPNs, often with quality-of-service control over a trusted delivery network.

- Layer 2 Tunneling Protocol (L2TP) which is a standards-based replacement, and a compromise taking the good features from each, for two proprietary VPN protocols: Cisco's Layer 2 Forwarding (L2F) (obsolete as of 2009) and Microsoft's Point-to-Point Tunneling Protocol (PPTP).

From the security standpoint, VPNs either trust the underlying delivery network, or must enforce security with mechanisms in the VPN itself. Unless the trusted delivery network runs among physically secure sites only, both trusted and secure models need an authentication mechanism for users to gain access to the VPN.

VPNs in mobile environments

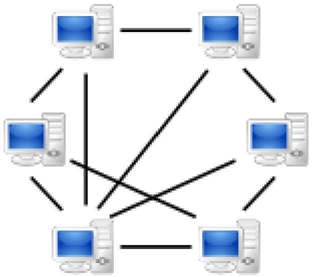
Mobile VPNs are used in a setting where an endpoint of the VPN is not fixed to a single IP address, but instead roams across various networks such as data networks from cellular carriers or between multiple Wi-Fi access points. Mobile VPNs have been widely used in public safety, where they give law enforcement officers access to mission-critical applications, such as computer-assisted dispatch and criminal databases, as they travel between different subnets of a mobile network. They are also used in field service management and by healthcare organizations, among other industries.

Increasingly, mobile VPNs are being adopted by mobile professionals and white-collar workers who need reliable connections. They allow users to roam seamlessly across networks and in and out of wireless-coverage areas without losing application sessions or dropping the secure VPN session. A conventional VPN cannot survive such events because the network tunnel is disrupted, causing applications to disconnect, time out, or fail, or even cause the computing device itself to crash.

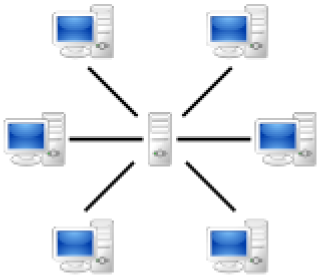
Instead of logically tying the endpoint of the network tunnel to the physical IP address, each tunnel is bound to a permanently associated IP address at the device. The mobile VPN software handles the necessary network authentication and maintains the network sessions in a manner transparent to the application and the user. The Host Identity Protocol (HIP), under study by the Internet Engineering Task Force, is designed to support mobility of hosts by separating the role of IP addresses for host identification from their locator functionality in an IP network. With HIP a mobile host maintains its logical connections established via the host identity identifier while associating with different IP addresses when roaming between access networks.

Chapter- 12

Peer-to-peer



A peer-to-peer system of nodes without central infrastructure



Centralized server-based service model

Peer-to-peer (**P2P**) computing or networking is a distributed application architecture that partitions tasks or work loads between peers. Peers are equally privileged, equipotent participants in the application. They are said to form a peer-to-peer network of nodes.

Peers make a portion of their resources, such as processing power, disk storage or network bandwidth, directly available to other network participants, without the need for central coordination by servers or stable hosts. Peers are both suppliers and consumers of resources, in contrast to the traditional client–server model where only servers supply, and clients consume.

The peer-to-peer application structure was popularized by file sharing systems like Napster. The peer-to-peer computing paradigm has inspired new structures and philosophies in other areas of human interaction. In such social contexts, peer-to-peer as a meme refers to the egalitarian social networking that is currently emerging throughout society, enabled by Internet technologies in general.

Architecture of P2P systems

Peer-to-peer systems often implement an abstract overlay network, built at Application Layer, on top of the native or physical network topology. Such overlays are used for indexing and peer discovery and make the P2P system independent from the physical network topology. Content is typically exchanged directly over the underlying Internet Protocol (IP) network. Anonymous peer-to-peer systems are an exception, and implement extra routing layers to obscure the identity of the source or destination of queries.

In *structured* peer-to-peer networks, peers (and, sometimes, resources) are organized following specific criteria and algorithms, which lead to overlays with specific topologies and properties. They typically use distributed hash table-based (DHT) indexing, such as in the Chord system (MIT).

Unstructured peer-to-peer networks do not provide any algorithm for organization or optimization of network connections.. In particular, three models of unstructured architecture are defined. In *pure peer-to-peer* systems the entire network consists solely of equipotent peers. There is only one routing layer, as there are no preferred nodes with any special infrastructure function. *Hybrid peer-to-peer* systems allow such infrastructure nodes to exist, often called *supernodes*. In *centralized peer-to-peer* systems, a central server is used for indexing functions and to bootstrap the entire system.. Although this has similarities with a structured architecture, the connections between peers are not determined by any algorithm. The first prominent and popular peer-to-peer file sharing system, Napster, was an example of the centralized model. Gnutella and Freenet, on the other hand, are examples of the decentralized model. Kazaa is an example of the hybrid model.

P2P networks are typically used for connecting nodes via largely *ad hoc* connections. Data, including digital formats such as audio files, and real time data such as telephony traffic, is passed using P2P technology.

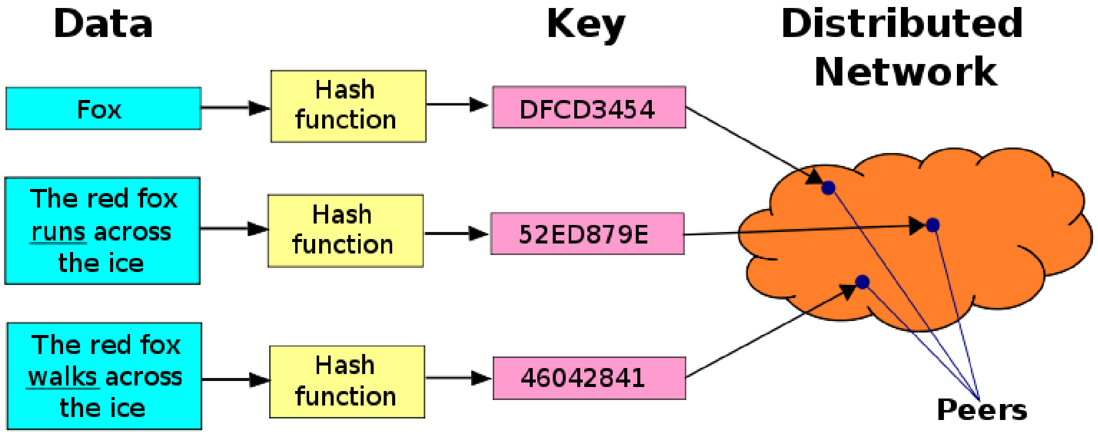
A pure P2P network does not have the notion of clients or servers but only equal *peer* nodes that simultaneously function as both "clients" and "servers" to the other nodes on the network. This model of network arrangement differs from the client–server model where communication is usually to and from a central server. A typical example of a file transfer that does not use the P2P model is the File Transfer Protocol (FTP) service in which the client and server programs are distinct: the clients initiate the transfer, and the servers satisfy these requests.

The P2P overlay network consists of all the participating peers as network nodes. There are links between any two nodes that know each other: i.e. if a participating peer knows the location of another peer in the P2P network, then there is a directed edge from the former node to the latter in the overlay network. Based on how the nodes in the overlay network are linked to each other, we can classify the P2P networks as unstructured or structured.

Structured systems

Structured P2P networks employ a globally consistent protocol to ensure that any node can efficiently route a search to some peer that has the desired file, even if the file is extremely rare. Such a guarantee necessitates a more structured pattern of overlay links. By far the most common type of structured P2P network is the distributed hash table (DHT), in which a variant of consistent hashing is used to assign ownership of each file to a particular peer, in a way analogous to a traditional hash table's assignment of each key to a particular array slot.

Distributed hash tables



Distributed hash tables

Distributed hash tables (DHTs) are a class of decentralized distributed systems that provide a lookup service similar to a hash table: (*key*, *value*) pairs are stored in the DHT, and any participating node can efficiently retrieve the value associated with a given key. Responsibility for maintaining the mapping from keys to values is distributed among the nodes, in such a way that a change in the set of participants causes a minimal amount of disruption. This allows DHTs to scale to extremely large numbers of nodes and to handle continual node arrivals, departures, and failures.

DHTs form an infrastructure that can be used to build peer-to-peer networks. Notable distributed networks that use DHTs include BitTorrent's distributed tracker, the Kad network, the Storm botnet, YaCy, and the Coral Content Distribution Network.

Some prominent research projects include the Chord project, the PAST storage utility, the P-Grid, a self-organized and emerging overlay network and the CoopNet content distribution system.

DHT-based networks have been widely utilized for accomplishing efficient resource discovery for grid computing systems, as it aids in resource management and scheduling of applications. Resource discovery activity involve searching for the appropriate resource types that match the user's application requirements. Recent advances in the domain of decentralized resource discovery have been based on extending the existing DHTs with the capability of multi-dimensional data organization and query routing. Majority of the efforts have looked at embedding spatial database indices such as the Space Filling Curves (SFCs) including the Hilbert curves, Z-curves, k-d tree, MX-CIF Quad tree and R*-tree for managing, routing, and indexing of complex Grid resource query objects over DHT networks. Spatial indices are well suited for handling the complexity of Grid resource queries. Although some spatial indices can have issues as regards to routing load-balance in case of a skewed data set, all the spatial indices are more scalable in terms of the number of hops traversed and messages generated while searching and routing Grid resource queries.

Unstructured systems

An unstructured P2P network is formed when the overlay links are established arbitrarily. Such networks can be easily constructed as a new peer that wants to join the network can copy existing links of another node and then form its own links over time. In an unstructured P2P network, if a peer wants to find a desired piece of data in the network, the query has to be flooded through the network to find as many peers as possible that share the data. The main disadvantage with such networks is that the queries may not always be resolved. Popular content is likely to be available at several peers and any peer searching for it is likely to find the same thing. But if a peer is looking for rare data shared by only a few other peers, then it is highly unlikely that search will be successful. Since there is no correlation between a peer and the content managed by it, there is no guarantee that flooding will find a peer that has the desired data. Flooding also causes a high amount of signaling traffic in the network and hence such networks typically have very poor search efficiency. Many of the popular P2P networks are unstructured.

In *pure* P2P networks: Peers act as equals, merging the roles of clients and server. In such networks, there is no central server managing the network, neither is there a central router. Some examples of pure P2P Application Layer networks designed for peer-to-peer file sharing are Gnutella (pre v0.4) and Freenet.

There also exist *hybrid* P2P systems, which distribute their clients into two groups: client nodes and overlay nodes. Typically, each client is able to act according to the momentary need of the network and can become part of the respective overlay network used to coordinate the P2P structure. This division between normal and 'better' nodes is done in order to address the scaling problems on early pure P2P networks. Examples for such networks are for example Gnutella (after v0.4) or G2.

Another type of hybrid P2P network are networks using on the one hand central server(s) or bootstrapping mechanisms, on the other hand P2P for their data transfers. These networks are in general called 'centralized networks' because of their lack of ability to work without their central server(s). An example for such a network is the eDonkey network (eD2k).

Indexing and resource discovery

Older peer-to-peer networks duplicate resources across each node in the network configured to carry that type of information. This allows local searching, but requires much traffic.

Modern networks use central coordinating servers and directed search requests. Central servers are typically used for listing potential peers (Tor), coordinating their activities (Folding@home), and searching (Napster, eMule). Decentralized searching was first done by flooding search requests out across peers. More efficient directed search strategies, including supernodes and distributed hash tables, are now used.

Many P2P systems use stronger peers (super-peers, super-nodes) as servers and client-peers are connected in a star-like fashion to a single super-peer.

Peer-to-peer-like systems

In modern definitions of peer-to-peer technology, the term implies the general architectural concepts outlined here. However, the basic concept of peer-to-peer computing was envisioned in earlier software systems and networking discussions, reaching back to principles stated in the first Request for Comments, RFC 1.

A distributed messaging system that is often likened as an early peer-to-peer architecture is the USENET network news system that is in principle a client–server model from the user or client perspective, when they read or post news articles. However, news servers communicate with one another as peers to propagate Usenet news articles over the entire group of network servers. The same consideration applies to SMTP email in the sense that the core email relaying network of Mail transfer agents has a peer-to-peer character, while the periphery of e-mail clients and their direct connections is strictly a client–server relationship. Tim Berners-Lee's vision for the World Wide Web, as evidenced by his WorldWideWeb editor/browser, was close to a peer-to-peer design in that it assumed each user of the web would be an active editor and contributor creating and linking content to form an interlinked *web* of links. This contrasts to the broadcasting-like structure of the web as it has developed over the years.

Advantages and weaknesses

In P2P networks, clients provide resources, which may include bandwidth, storage space, and computing power. As nodes arrive and demand on the system increases, the total capacity of the system also increases. In contrast, in a typical client–server architecture, clients share only their demands with the system, but not their resources. In this case, as more clients join the system, less resources are available to serve each client.

The distributed nature of P2P networks also increases robustness, and—in pure P2P systems—by enabling peers to find the data without relying on a centralized index server. In the latter case, there is no single point of failure in the system.

As with most network systems, unsecure and unsigned codes may allow remote access to files on a victim's computer or even compromise the entire network. In the past this has happened for example to the FastTrack network when anti P2P companies managed to introduce faked chunks into downloads and downloaded files (mostly MP3 files) were unusable afterwards or even contained malicious code. Consequently, the P2P networks of today have seen an enormous increase of their security and file verification mechanisms. Modern hashing, chunk verification and different encryption methods have made most networks resistant to almost any type of attack, even when major parts of the respective network have been replaced by faked or nonfunctional hosts.

Internet service providers (ISPs) have been known to throttle P2P file-sharing traffic due to the high-bandwidth usage. Compared to Web browsing, e-mail or many other uses of the internet, where data is only transferred in short intervals and relative small quantities, P2P file-sharing often consists of relatively heavy bandwidth usage due to ongoing file transfers and swarm/network coordination packets. As a reaction to this bandwidth throttling several P2P applications started implementing protocol obfuscation, such as the BitTorrent protocol encryption. Techniques for achieving "protocol obfuscation" involves removing otherwise easily identifiable properties of protocols, such as deterministic byte sequences and packet sizes, by making the data look as if it was random.

A possible solution to this is called P2P caching, where a ISP stores the part of files most accessed by P2P clients in order to save access to the Internet.

Social and economic impact

The concept of P2P is increasingly evolving to an expanded usage as the relational dynamic active in distributed networks, *i.e.*, not just computer to computer, but human to human. Yochai Benkler has coined the term commons-based peer production to denote collaborative projects such as free and open source software. Associated with peer production are the concepts of:

- peer governance (referring to the manner in which peer production projects are managed)

- peer property (referring to the new type of licenses which recognize individual authorship but not exclusive property rights, such as the GNU General Public License and the Creative Commons licenses)
- peer distribution (or the manner in which products, particularly peer-produced products, are distributed)

Some researchers have explored the benefits of enabling virtual communities to self-organize and introduce incentives for resource sharing and cooperation, arguing that the social aspect missing from today's peer-to-peer systems should be seen both as a goal and a means for self-organized virtual communities to be built and fostered. Ongoing research efforts for designing effective incentive mechanisms in P2P systems, based on principles from game theory are beginning to take on a more psychological and information-processing direction.

Applications

There are numerous applications of peer-to-peer networks. The most commonly known is for content distribution

Content delivery

- Many file sharing networks, such as Gnutella, G2 and FastTrack popularized peer-to-peer technologies. From 2004, it is the largest contributor of network traffic on the Internet.
- Peer-to-peer content delivery networks (P2P-CDN) (Giraffic, Kontiki, Ignite, RedSwoosh).
- Software publication and distribution (Linux, several games); via file sharing networks.
- Streaming media. P2PTV and PDTP. Applications include TVUPlayer, Joost, CoolStreaming, Cybersky-TV, PPLive, LiveStation
- Spotify uses a peer-to-peer network along with streaming servers to stream music to its desktop music player.
- Peercasting for multicasting streams.
- Pennsylvania State University, MIT and Simon Fraser University are carrying on a project called LionShare designed for facilitating file sharing among educational institutions globally.
- Osiris (Serverless Portal System) allows its users to create anonymous and autonomous web portals distributed via P2P network.

Networking

- Domain Name System, for Internet information retrieval. ee Comparison of DNS server software
- cloud computing
- Dalesa a peer-to-peer web cache for LANs (based on IP multicasting).

Science

- In bioinformatics, drug candidate identification. The first such program was begun in 2001 the Centre for Computational Drug Discovery at the University of Oxford in cooperation with the National Foundation for Cancer Research. There are now several similar programs running under the United Devices Cancer Research Project.
- The sciencenet P2P search engine.
- BOINC

Search

- YaCy, a free distributed search engine, built on principles of peer-to-peer networks.

Communications networks

- Skype, one of the most widely used internet phone applications is using P2P technology.
- VoIP (using application layer protocols such as SIP)
- Instant messaging and online chat
- Completely decentralized networks of peers: Usenet (1979) and WWIVnet (1987).

General

- Research like the Chord project, the PAST storage utility, the P-Grid, and the CoopNet content distribution system.
- JXTA, for Peer applications.

Miscellaneous

- The U.S. Department of Defense has started research on P2P networks as part of its modern network warfare strategy. In May, 2003 Dr. Tether. Director of Defense Advanced Research Project Agency testified that U.S. Military is using P2P networks.
- Kato et al.'s studies indicate over 200 companies with approximately $400 million USD are investing in P2P network. Besides File Sharing, companies are also interested in Distributing Computing, Content Distribution.
- Wireless community network, Netsukuku
- An earlier generation of peer-to-peer systems were called "metacomputing" or were classed as "middleware". These include: Legion, Globus

Historical perspective

Tim Berners-Lee's vision for the World Wide Web was close to a P2P network in that it assumed each user of the web would be an active editor and contributor, creating and linking content to form an interlinked "web" of links. This contrasts to the current broadcasting-like structure of the web.

Some networks and channels such as Napster, OpenNAP and IRC serving channels use a client–server structure for some tasks (e.g., searching) and a P2P structure for others. Networks such as Gnutella or Freenet use a P2P structure for nearly all tasks, with the exception of finding peers to connect to when first setting up.

P2P architecture embodies one of the key technical concepts of the Internet, described in the first Internet Request for Comments, RFC 1, "Host Software" dated April 7, 1969. More recently, the concept has achieved recognition in the general public in the context of the absence of central indexing servers in architectures used for exchanging multimedia files.

Network neutrality controversy

Peer-to-peer applications present one of the core issues in the network neutrality controversy. In October 2007, Comcast, one of the largest broadband Internet providers in the USA, started blocking P2P applications such as BitTorrent. Their rationale was that P2P is mostly used to share illegal content, and their infrastructure is not designed for continuous, high-bandwidth traffic. Critics point out that P2P networking has legitimate uses, and that this is another way that large providers are trying to control use and content on the Internet, and direct people towards a client-server-based application architecture. The client-server model provides financial barriers-to-entry to small publishers and individuals, and is quite inefficient for sharing large files.