TECHNOLOGY INNOVATION

AN INTRODUCTION TO HOSPITAL NETWORKS

P.S. Ganney [1]

[1] University College London Hospitals NHS Trust, London, England

Abstract— **Computer networks in healthcare settings increasingly contain systems that fall under the classification of Medical Devices. Such items are under the jurisdiction of Clinical Engineering and Medical Physics staff who have not been formally trained in networking. This article provides an introduction to many of the concepts.**

Keywords— **Computer, Network, Connectivity, LAN, Medical Device**

X. Introduction

The emergence of the computer-based medical device has led naturally to a desire to interconnect such devices. This paper gives an overview of networking concepts, together with the applicability of them to a healthcare environment.

XI. Networking concepts

Whilst a computer on its own is a powerful device, the possibilities and the power increase greatly when such devices are linked together in a network.

The minimum number of devices in a network is 2 (otherwise you're talking to yourself). The maximum number depends on the addressing method: standard IP addresses allow for 4,294,967,296 (256^4) but there are lots of ways to extend this, as we shall see.

In a hospital environment, devices are usually connected physically – i.e. with a cable. This improves reliability as well as giving larger bandwidth and higher speed. There are multiple ways of connecting devices, but the simplest is via a hub, which is essentially a connection box where all incoming signals are sent to all connected devices.

A. The Network Packet

All networking is described in terms of packets so it is useful to describe this first. A network packet is a formatted unit of data carried by a packet-switched network. Computer communications links that do not support packets, such as traditional point-to-point telecommunications links, simply transmit data as a bit stream.

A packet consists of control information and user data, which is also known as the payload. Control information provides data for delivering the payload, for example: source and destination network addresses, error detection codes, and sequencing information. Typically, control information is found in packet headers and trailers.

Different communications protocols use different conventions for distinguishing between the elements and for formatting the data. For example, in Point-to-Point Protocol, the packet is formatted in 8-bit bytes, and special characters are used to delimit the different elements. Other protocols like Ethernet, establish the start of the header and data elements by their location relative to the start of the packet. Some protocols format the information at a bit level instead of a byte level.

A good analogy is to consider a packet to be like a letter: the header is like the envelope, and the data area is whatever the person puts inside the envelope.

In the seven-layer OSI model of computer networking (see later), packet strictly refers to a data unit at layer 3, the Network Layer. The correct term for a data unit at Layer 2, the Data Link Layer, is a frame, and at Layer 4, the Transport Layer, the correct term is a segment or datagram. For the case of TCP/IP communication over Ethernet, a TCP segment is carried in one or more IP packets, which are each carried in one or more Ethernet frames.

B. Hardware

We shall first examine the three main hardware components – hub, switch and router.

Fig. 1. A network hub (source: Wikimedia Commons)

A hub is a simple connection box, operating at the Physical Layer (layer 1) of the OSI model (see later). Like a transport hub, it's where everything comes together. Unlike a transport hub, though, whatever comes in on one connection goes out on all other connections. It's up to the receiver to decide whether or not the message is for them. Hubs therefore work well for small networks, but get messy and slow down for larger ones. It is therefore often common to find them in small networks (e.g. at home) or in sub-networks (e.g. in an office).

Fig. 2. A network switch (source: Wikimedia Commons)

A switch is a device that filters and forwards packets between LAN segments. Switches operate at the data link layer (layer 2) and sometimes the network layer (layer 3) of the OSI Reference Model (see later) and therefore support any packet protocol. LANs that use switches to join segments are called switched LANs or, in the case of Ethernet networks, switched Ethernet LANs.

Fig. 3. A network router (source: Wikimedia Commons)

A router forwards data packets along networks. A router is connected to at least two networks and is located at a gateway, the place where two or more networks connect. Routers use headers and forwarding tables to determine the best path for forwarding the packets, and they use protocols such as ICMP (Internet Control Message Protocol) to communicate with each other and configure the best route between any two hosts.

A large network will therefore contain all 3 of these types of devices.

Today most routers have become something of a Swiss Army knife, combining the features and functionality of a router and switch/hub into a single unit (and may contain a basic Domain Name Service (DNS)).

The functions of a router, hub and a switch are all quite different from one another, even if at times they are all integrated into a single device. We will start with the hub and the switch since these two devices have similar roles on the network.

Each serves as a central connection for all of the network equipment and handles a data type known as frames. Frames carry data. When a frame is received, it is amplified and then transmitted on to the port of the destination device. The big difference between these two devices is in the method in which frames are being delivered.

In a hub, a frame is passed along or "broadcast" to every one of its ports. It doesn't matter that the frame is only destined for one port. The hub has no way of distinguishing which port a frame should be sent to. Passing it along to every port ensures that it will reach its intended destination. This places a lot of traffic on the network and can lead to poor network response times.

Additionally, a 10/100Mbps hub must share its bandwidth with each and every one of its ports. So when only one device is broadcasting, it will have access to the maximum available bandwidth. If, however, multiple devices are broadcasting, then that bandwidth will need to be divided among all of those systems, which will degrade performance.

A switch, however, keeps a record of the MAC addresses of all the devices connected to it. With this information, a switch can identify which system is sitting on which port. So when a frame is received, it knows exactly which port to send it to, without significantly increasing network response times. And, unlike a hub, a 10/100Mbps switch will allocate a full 10/100Mbps to each of its ports. So regardless of the number of devices transmitting, users will always have access to the maximum amount of bandwidth. For these reasons a switch is considered to be a much better choice than a hub.

Routers are completely different devices. Where a hub or switch is concerned with transmitting frames, a router's job, as its name implies, is to route packets to other networks until that packet ultimately reaches its destination. One of the key features of a packet is that it not only contains data, but the destination address of where it's going.

A router is typically connected to at least two networks, commonly two Local Area Networks (LANs) or Wide Area Networks (WANs) or a LAN and its ISP's network. For example, a PC or workgroup and Broadband.

Routers might have a single WAN port and a single LAN port and are designed to connect an existing LAN hub or switch to a WAN. Ethernet switches and hubs can be connected to a router with multiple PC ports to expand a LAN. Depending on the capabilities (kinds of available ports) of the router and the switches or hubs, the connection between the router and switches/hubs may require either straight-through or crossover (null-modem) cables. Some routers even have USB ports, and more commonly, wireless access points built into them.

Besides the inherent protection features provided by the NAT, many routers will also have a built-in, configurable, hardware-based firewall. Firewall capabilities can range from the very basic to quite sophisticated devices. Among the capabilities found on leading routers are those that permit configuring TCP/UDP ports for games, chat services, and the like, on the LAN behind the firewall.

So, in summary, a hub glues together an Ethernet network segment, a switch can connect multiple Ethernet segments more efficiently and a router can do those functions plus route TCP/IP packets between multiple LANs and/or WANs; and much more.

C. Network Topologies

The topology of the network can be thought of as its shape. Not its physical shape, but its logical one: much like the tube map shows how stations connect, not where they are. The five basic topologies are bus, ring, star, tree and mesh, which we now examine.

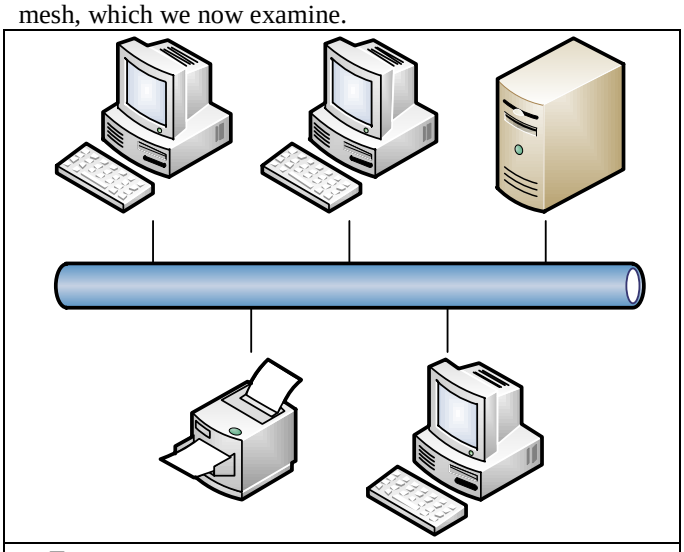

Fig. 4. Bus topology

Bus networks use a common backbone to connect all devices. A single cable, the backbone functions as a shared communication medium that devices attach or tap into with an interface connector. A device wanting to communicate with another device on the network sends a broadcast message onto the wire that all other devices see, but only the intended recipient actually accepts and processes the message.

Ethernet bus topologies are relatively easy to install and don't require much cabling compared to the alternatives. However, bus networks work best with a limited number of devices. If more than a few dozen computers are added to a network bus, performance problems will likely result. In addition, if the backbone cable fails, the entire network effectively becomes unusable.

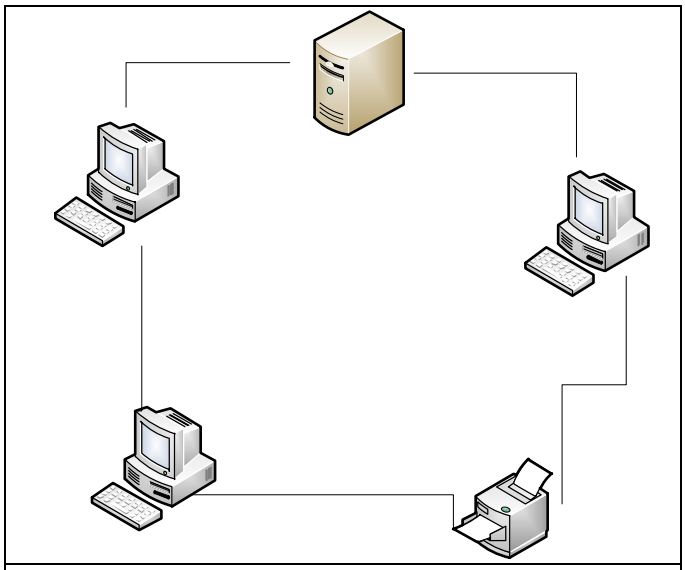

Fig. 5. Ring topology

In a ring network, every device has exactly two neighbours for communication purposes. All messages travel through a ring in the same direction (either "clockwise" or "counterclockwise"). A failure in any cable or device breaks the loop and can take down the entire network.

To implement a ring network, one typically uses FDDI[1], SONET[2], or Token Ring technology[3].

1. [1] Fiber Distributed Data Interface – a set of ANSI and ISO standards for data transmission on fibre optic lines in a LAN that can extend in range up to 200 km (124 miles).
2. [2] Synchronous Optical Network – the American National Standards Institute standard for synchronous data transmission on optical media.
3. [3] In a token ring, a “token” is passed around the network. The device holding the “token” is permitted to transmit – nothing else is. If a device has nothing to transmit, it passes the token on.

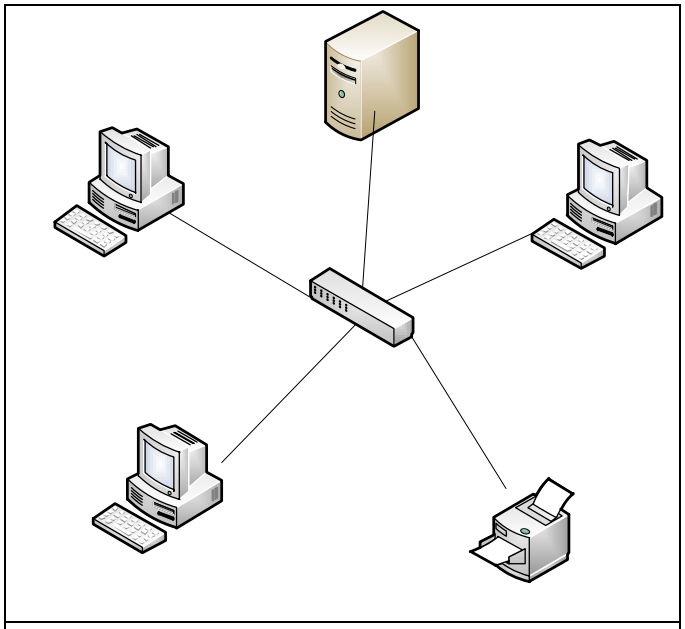

Fig. 6. Star topology

Most small (e.g. home) networks use the star topology. A star network features a central connection point called a "hub node" that may be a network hub, switch or (more likely) a router. Devices typically connect to the hub with Unshielded Twisted Pair (UTP) Ethernet.

Compared to the bus topology, a star network generally requires more cable, but a failure in any star network cable will only take down one computer's network access and not the entire LAN. (If the hub fails, however, the entire network also fails.)

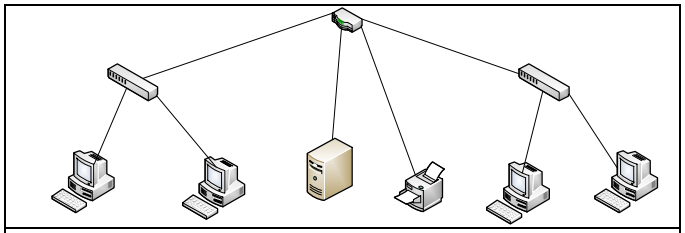

Fig. 7. Tree topology

Tree topologies integrate multiple star topologies together onto a bus. In its simplest form, only hub devices connect directly to the tree bus, and each hub functions as the root of a tree of devices. This bus/star hybrid approach supports future expandability of the network much better than a bus (limited in the number of devices due to the broadcast traffic it generates) or a star (limited by the number of hub connection points) alone.

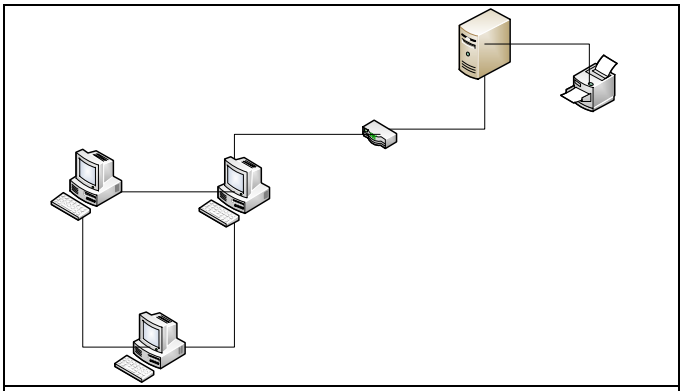

Fig. 8. Mesh topology

Mesh topologies involve the concept of routes. Unlike each of the previous topologies, messages sent on a mesh network can take any of several possible paths from source to destination. (Recall that even in a ring, although two cable paths exist, messages can only travel in one direction.) Some WANs, most notably the Internet, employ mesh routing, specifically for the resilience that it brings.

A mesh network in which every device connects to every other is called a full mesh. As shown in the illustration above, partial mesh networks also exist in which some devices connect only indirectly to others.

D. IP addressing and DNS

People like to communicate using names. Therefore we call our devices “Linac B PC”, “Endoscopy control” and so on rather than “1”, “2” or “3.1412”.

Computers prefer numbers. When they’re communicating, they require unique numbers. Therefore they use IP addresses. An IP (Internet Protocol) version 4 address is formed of 4 groups of digits, separated by dots[4]. Each group of digits can range in value from 0 to 255 – 256[5] unique numbers. The combination of these 4 groups should uniquely identify the device on the network. If it’s not unique then chaos ensues.

Of course, we still like to call our devices by names, so a network service called a DNS (Domain Name Server) is usually available to translate “LinacA” into 123.45.67.89 so that the command “ping LinacA[6]” can be issued and a reply can come from 123.45.67.89 without having to know the IP address of LinacA.

The four-byte IP address allows us to perform grouping. A set of devices may be given addresses in the same range – i.e. they have the same first 2 or 3 bytes, differing only in the final one or two. Given that we have a DNS to do the translation and can therefore give our devices sensible

4. [4] IPv6 also exists, but is not widespread (yet), but does allow 296 addresses.
5. [5] Due to binary: 256 is 2^8 so each group of digits is composed of 8 bits.
6. [6] A command that sends an “are you there” message.

names, this may seem unnecessary, but it does allow us to segregate our network using masking.

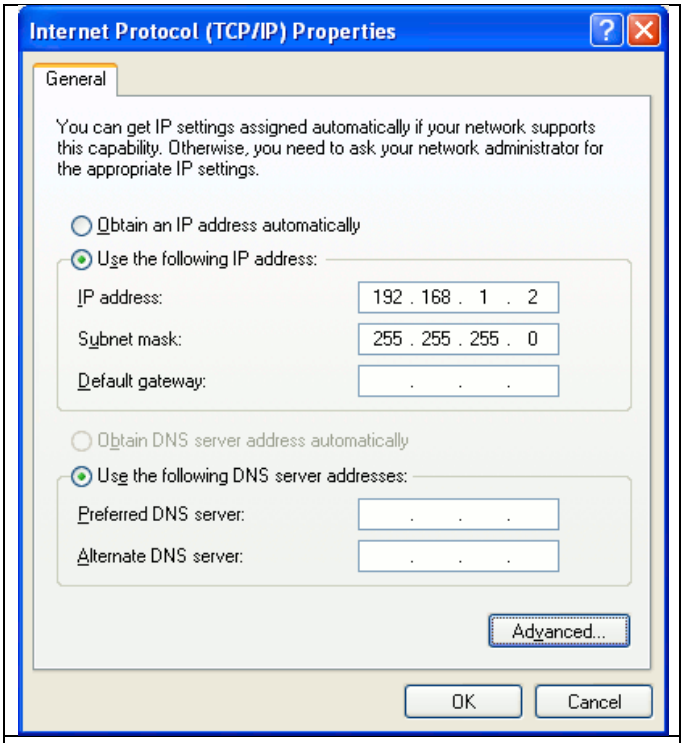

Fig. 9. An IP setup dialog from Windows XP

An IP mask of 255.255.255.0 allows us to separate out the network prefix and the host number using bitwise AND[7].

Thus, in the example in Fig. 9, the network address is 192.168.1 and the host number is 2. Networks can be further divided to produce smaller subnets by putting more bits into the mask. For example a mask of 255.255.255.192 leaves this example with the same host number, but a host number previously of 130 would now be rendered as 2.

The modern way of specifying a network mask is to specify the number of bits in it, as masks are always 1s to the left and 0s to the right and no intermingling. Thus 192.168.1.2, netmask 255.255.255.0 is written as 192.168.1.2/24. This is known as CIDR (Classless Inter-Domain Routing) notation.

It is this subnetting that allows us to write simple router rules as they can be produced for groups of IPs instead of for each one individually.

There are 3 main ways of giving a device an IP address. Firstly, and most simply, is static. This is where the 4-byte number is given to that device and that device only (see Fig. 9 where this is the case).

The second is dynamic. This is where a device requests an IP address from the DHCP[8] controller and is provided with the next spare one. This system works well with hot-spots and networks where there are more devices than IP addresses, but never has all the devices switched on at once. It's also very useful in networks where the number of devices changes frequently (e.g. a hospital network): it is not necessary to keep note of the addresses issued so far and no-one has to check whether a device has been decommissioned. Administratively it's the simplest method, but it too has drawbacks.

The main one is that some protocols have to communicate via fixed IP addresses. DICOM is one such (although later devices can often take the device name, many can't). One solution to this is to run a mixed addressing network, with some static IPs (usually in the same range so they can be easily administered) and the rest dynamic. One other solution is to use reserved IPs – in this case it is again all controlled by the DHCP controller (often part of the DNS), but when a device requests an IP it is always served the same one, which has been reserved for it. That way the network runs only in dynamic mode but the needs of static addressing are met.

Which brings us to the question of 2 Devices with the same IP address. Locally this is impossible, but globally not so. If a device's IP address is only visible up to the router, then it is not visible beyond it (e.g. on the Internet) and therefore another device on the other side of the router can have exactly the same IP address and they will never conflict.

E. Security

The first question to address is: where is your data? It may be on a local machine, on a local server or in "the cloud".

Cloud Computing is where the software and data do not reside on local servers, in the local organisation or possibly even in the same country. Whilst this frees up a lot of infrastructure and makes mobile computing more possible, it does have 2 main drawbacks:

The first is that a reliable network connection is essential to use the cloud.

The second is that the laws of data protection that apply to data are those that exist in the country in which the data resides: in the USA, for example, companies are allowed to sell the data they have on their network. Not to anyone, of course, but they can sell health records to insurance agencies, email addresses to marketers etc. So it is important, for health records, to know where the data is being held (for further information see the Data Protection Act, which is beyond the scope of this article). In the UK there is a G-Cloud, a cloud solution hosted there for public service use, thus making it subject to the UK's data protection laws. There are also more and more assured clouds being marketed.

There are great advantages in connecting together ICT equipment to enable data sharing, together with the enhanced safety from a reduction in transcription errors and

7. [7] Where two binary numbers are compared, bit by bit and an AND operation is performed on them to produce the result.
8. [8] Dynamic Host Configuration Protocol

the increased availability and speed of access to information. However, this connectedness brings with it additional system security issues: a failure in one part may be swiftly replicated across the IT estate. There are many ways to tackle these issues and this section details some of these. It should be noted that best practice will utilise a range of security methods.

The first method is one of segregation, using a firewall. A firewall is, in the simplest sense, a pair of network cards (or a router) and a set of rules. A network packet arrives at one card, is tested against the rules and (if it passes) is passed to the other card for transmission. In this way, a part (or the whole of) a network can be protected from activity on the rest of the network by restricting the messages that can pass through it to a predefined and pre-approved set. The rules controlling this may be as simple as only allowing a predefined set of IP addresses through. Refinements include port numbers, the direction the message is travelling in, whether the incoming message is a response to an outgoing one (e.g. a web page) and specific exceptions to general rules. This is all achieved via packet filtering, where the header of the packet is examined in order to extract the information required for the rules.

The above description is of a hardware firewall. Software firewalls run on the device after the network traffic has been received. They can therefore be more sophisticated in their rules in that they can have additional information such as the program that made the request. Software firewalls can also include privacy controls and content filtering. As the software firewall runs on the device, if the device becomes compromised, then the firewall may also be compromised. The Windows 7 firewall only blocks incoming traffic, so will not prevent a compromised device from sending malicious network packets.

F. Bandwidth

Bandwidth in a computer network sense is its transmission capacity, which (as it is a function of the speed of transmission) is usually expressed in bps (bits per second). The most common wired bandwidths are 1 Gbps (often called "Gigabit Ethernet"), 10 Mbps (standard Ethernet) and 100 Mbps (fast Ethernet). Wireless is generally slower – 802.11g supports up to 54 Mbps, for example. Note that these are maximums and a wired network stands a better chance of providing the full bandwidth due to less interference. As bandwidth is actually the capacity, binding together several cables can increase the total bandwidth whilst not increasing the speed – although this would not normally be done in a departmental network, the point at which a hospital meets the national N3 network may be implemented this way (provided both sides of the connection can handle it – which is usually by routing pre-defined packets to specified lines, e.g. by IP address range).

It is never a good idea to reach 100% bandwidth utilisation and the average in order to avoid this may be as low as 30%, although 50% would be more common. The amount of "spare" capacity is often termed "headroom". At 75% the throughput verses offered traffic curve starts to depart from a linear proportional increase of throughput for increase of offered traffic. At 80% the channel could be approaching overload. Much is dependent upon the traffic type - data traffic can cope with higher utilisation levels than voice as delay and jitter have more effect on the user experience for voice traffic than data traffic. Optimisation techniques such as QOS[9] can be used to prioritise voice traffic (or any other traffic that is time-critical).

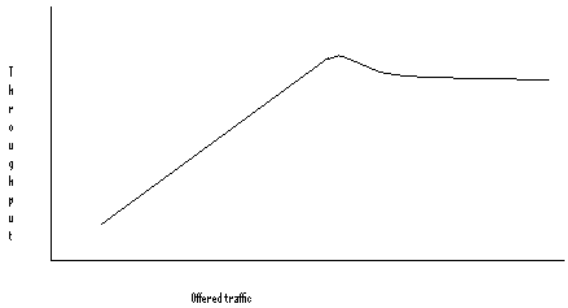

Fig. 10. A generalised throughput verses offered traffic curve showing a deterioration in performance when optimal levels are breached. – from TakTak et al [1]

The above utilisation levels are generally for non collision based channels. In the case of Ethernet which uses CSMA[10] with collision detection as the access mechanism, utilisation should be much lower. An overdriven CSMA channel can result in throughput reduction rather than an increase with increasing offered traffic. Retrys as a result of a collision lead to more retrys and more collisions and so on. Collision detection with a limitation on the number of retrys and back off between the retrys is aimed at keeping the channel stable but throughput will tail off. Kleinrock [2] provides good further reading.

All the resilience methods outlined here require a level of redundancy: be it a copy, checksums or headroom. Thus a resilient system will always be over-engineered – in the case of bandwidth, over-engineering can remove the need for optimisation systems such as QOS, thus making the design (and therefore the support) simpler.

XII. STANDARDS

There are several international standards covering IT networks, but the one of particular interest here is IEC 80001-1 (2010). This standard, titled "Application of risk management for IT-networks incorporating medical devices -- Part 1: Roles, responsibilities and activities" contains

9. [9] Quality-of-service, a Cisco product.
10. [10] Carrier Sense Multiple Access, a protocol in which a node verifies the absence of other traffic before transmitting on a shared transmission medium.

many definitions. The main one for consideration here is that of the "*Medical IT Network*", which is defined as "*an IT-NETWORK that incorporates at least one MEDICAL DEVICE*". An IT-NETWORK is defined as "*a system or systems composed of communicating nodes and transmission links to provide physically linked or wireless transmission between two or more specified communication nodes*" and is adapted from IEC 61907:2009, definition 3.1.1. The MEDICAL DEVICE definition is from the Medical Device Directive. Thus a hospital that connects even one medical device into its standard network (or, indeed, loads medical device software onto a non-medical device so connected) has thereby created a medical IT-Network. The bounds of this network are that of the responsible organisation[11] but do bring different responsibilities into play, as detailed in the standard. In particular, the role of the medical IT-network risk manager, the person accountable for risk management of the medical IT-network is specified.

This family of standards could stimulate cross disciplinary teams in hospitals, involving IT departments, Informatics Departments and Clinical Departments in establishing quality systems for Clinical Computing. This would include assurance of finance, planning of procurement and upgrades, and the monitoring of adequate support arrangements. Medical Physicists and Clinical Engineers would have an important role to play in these groups, particularly with regard to day to day running and relationships with device suppliers.

XIII. THE OSI 7-LAYER MODEL [3]

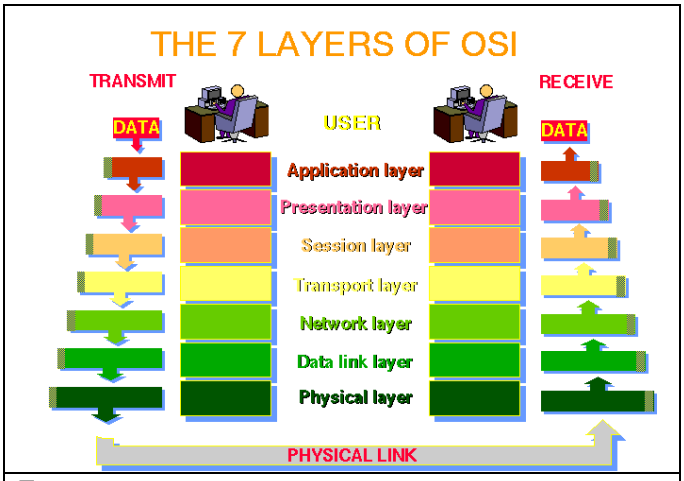

Fig. 11. The 7-layer OSI model. Source: http://www.webopedia.com/img/OSI_Model.jpg

The OSI 7-layer model describes the transmission of messages. In sending a message, each layer (from the highest – the Application – to the lowest – the Physical) adds a wrapper to the message, which is removed after physical transmission as the message makes its way back up the layers into the receiving application.

The physical layer, the lowest layer of the OSI model, is concerned with the transmission and reception of the unstructured raw bit stream over a physical medium. It describes the electrical/optical, mechanical, and functional interfaces to the physical medium, and carries the signals for all of the higher layers. It deal with such matters as what signal state represents a binary 1, how many pins are on a connector (and what they do) and how many volts/db should be used to represent a given signal state.

The data link layer provides error-free transfer of data frames from one node to another over the physical layer, allowing layers above it to assume virtually error-free transmission over the link. The data link layer handles matters such as establishing and terminating the logical link between two nodes, transmitting/receiving frames sequentially and determining when the node "has the right" to use the physical medium (see the description of "token ring" earlier).

The network layer controls the operation of the subnet, deciding which physical path the data should take based on network conditions, priority of service, and other factors. It is concerned with routing, subnet traffic control and logical-physical address mapping.

The transport layer ensures that messages are delivered error-free, in sequence, and with no losses or duplications. It relieves the higher layer protocols from any concern with the transfer of data between them and their peers. It handles such matters as message segmentation, message acknowledgement and session multiplexing (multiplexing several message streams, or sessions onto one logical link and keeping track of which messages belong to which sessions).

The session layer allows session establishment between processes running on different stations. It handles session establishment, maintenance and termination, allowing two application processes on different machines to establish, use and terminate a connection, called a session.

The presentation layer formats the data to be presented to the application layer. It can be viewed as the translator for the network. This layer may translate data from a format used by the application layer into a common format at the sending station, then translate the common format to a format known to the application layer at the receiving station. It handles code conversions such as ASCII to EBCDIC, bit-order and CR/LF. Data compression and encryption take place at this layer.

The application layer serves as the window for users and application processes to access network services. This layer contains a variety of commonly needed functions such as resource sharing, remote printing and file access and directory services.

11. [11] Thus a connection to the Internet does not render a network a medical IT-network.

XIV. CONCLUSIONS

Hospital networks are useful conduits for distributing and sharing information, which thereby enhance patient care. They do, however, require correct implementation in order to do so safely.

ACKNOWLEDGMENT

The author wishes to acknowledge colleagues past and present, especially Bill Webster, who assisted the development of this material over several years.

REFERENCES

8. Taktak AFG, Ganney PS, Long D & White P (2013) Clinical Engineering. Academic Press, Oxford. especially chapters 8-10 by Claridge, Ganney, McDonagh & Pisharody
9. Kleinrock L (1975-2011) Queuing Systems Volumes 1-3. John Wiley & Sons, London
10. https://support.microsoft.com/en-us/kb/103884

Contacts of the corresponding author:

Author:	Paul Ganney
Institute:	University College London Hospitals NHS Trust
Street:	Euston Road
City:	London
Country:	England
Email:	paul.ganney@uclh.nhs.uk