REPORT #1

과 목: Data Communications

담 당 교 수 : 이 태 진 교수님

학 과: 철학 / 컴퓨터공학

학 번: 2014314130

이 름:이영민

제 출일: 2019.03.29

1. What is the coding rule and the advantage/disadvantage of the Korean character encoding system? Where are the Korean letters in the Unicode?

인코딩(encoding)한다는 것은 정보의 형태나 형식을 처리하는 방식을 의미한다. 사용자가 컴퓨터와 문자를 가지고 소통하기 위하여 이 문자를 컴퓨터가 해석할 수 있는 방식으로 표현해야 한다. 대표적인 표준 방식으로 ASCII 문자 인코딩이 있다. 예를 들어, A는 십진수 65로 변환되며 이를 다시 이진수로 변환하여 컴퓨터가 해석하게 되는 것이다. ASCII는 7비트 인코딩으로서, 0x00~0x7F까지의 총 127개의 문자로 이루어진다. 이는 미국에서 제정된 표준이었기에 유럽권의 문자를 표현하기 위하여 1비트를 추가하였고 확장 아스키(Extended ASCII)를 통하여 유럽 표준 통일안이 만들어졌다. 하지만 256개의 코드로는 한글을 표현하지 못하기에 별도의 방법이 필요하다.

컴퓨터에서 한글을 표현하는 방법은 크게 조합형과 완성형으로 나눌 수 있다. 먼저 조합형의 경우 초성, 중성, 종성을 나누어 각각에 코드를 할당하는 방식이다. 반면 완성 형이란 '가', '나', '힣'과 같은 완성된 문자에 코드를 할당하는 방식이며 완성형이 한글 표준안이다.

조합형의 경우 간단한 계산으로 초,중,종성을 분리할 수 있다는 장점이 있지만 여러 바이트 단위로 비트를 쪼개어 해석하기 때문에 오버헤드가 있으며 다른 문자 체계와 호환이 안된다는 단점이 있다.

이로 인하여 완성형이 표준으로 채택되었다. 하지만 EUC-KR의 경우 사용빈도가 높은 2350자만을 채택하여 지원하였기에 표현할 수 없는 한글이 있다. 이를 보완한 것이 CP949이다. 또 다른 유니코드로는 UTF-8이 있으며 전세계 모든 문자를 동일하게 표현할 수 있는 인코딩 방식으로 사용도가 높다.

한글 인코딩은 장점보다는 단점이 더 눈에 들어온다. 예를 들어, 한글 한 자를 2byte 혹은 3byte로 할당하기에 문서의 용량이 커지게 된다. 뿐만 아니라 몇몇 IDE나 프레임워크는 한글을 지원하지 않아 추가적인 작업을 해줘야 한다.

유니코드는 전 세계적으로 사용하는 모든 문자 집합을 하나로 모은 것으로서, 유니코드 값을 나타내기 위하여 코드 포인터인 U+를 붙여 표시된다. 유니코드에서 한글의 표현하기 위한 코드 영역은 다음과 같다.

한글 자모: U+1100 ~ U+11FF

호환용 한글 자모 : U+3130 ~ U+318F 한글 자모 확장 A : U+A960 ~ U+A97F 한글 소리 마디 : U+AC00 ~ D&AF 한글 자모 확장 B : U+D7B0 ~ D&FF

2. What is the difference between port numbers, IP addresses, and MAC addresses?

네트워크는 수많은 end-point로 구성된다. 네트워크 상에서 이들을 구별해주기 위한 고유한 식별자가 Port Num, IP 주소 그리고 MAC 주소이다.

1) Port Numbers

멀티프로그래밍 환경에서는 하나의 컴퓨터에서도 여러 프로그램이 돌아갈 수 있다. 웹 클라이언트로 요청을 보냈다면 웹 서버를 거쳐 해당 응답은 웹 클라이언트로 돌아와야 할 것이다. 이처럼 한 컴퓨터 위에서 특정 프로세스를 식별해주는 번호가 Port Number인 것이다. 포트 번호는 16bit로 이루어지며 well-known port, registered port, dynamic port로 범위가 나눠진다. 예를 들어, HTTP는 port 80을 예약하고 있다. 간단한 소켓 프로그래밍을 해본다면 dynamic port 범위 내에서 포트를 선택해야 할 것이다. 네트워크 계층에서 Transport Layer에서 포트번호를 처리하는데 이 계층에서 다루는 헤더에 송신, 수신 포트번호가 추가되며 이를 바탕으로 프로세스를 식별하여 Application Layer로 보내게 된다.

2) IP addresses

Port Number가 프로세스 간의 식별자였다면 IP 주소는 인터넷에 연결된 기기를 식별하는 주소이다. Communication Device가 인터넷에 참가하면 IP주소를 할당 받게 되며 이를 기반으로 네트워크에 참여할 수 있게 된다. IP주소는 유동적인데 어떤 인터넷망에 접속하냐에 따라 그 값이 달라지기 때문이다. (이를 할당해주는 것이 DHCP의 역할이다.) IP 주소는 32bit로 구성되어 있으며 4개의 옥텟 (octet)으로 구별된다. 예를 들면, 127.0.0.1과 같은 방식이다. 나아가 IP주소 자체는 네트워크 주소와 호스트 주소로 구분되는데 전자는 해당 기기가 속한 네트워크를 식별하며, 후자는 해당 네트워크 내에서의 기기를 식별한다. 두 주소를 나누는 방법은 원래 Classful addressing이었으나, 주소 고갈로 인하여 Classless addressing을 사용하여 네트워크 주소와 호스트 주소를 나누고 있다. 현재 IPv4 자체의 주소 고갈로 인하여 32비트에서 128비트로 IP주소를 확장하는 IPv6주소

체계로 나아가고 있다.

3) MAC addresses

컴퓨터는 IP 주소 외에도 MAC 주소를 가지고 있다. 이는 IP와 달리, 기기가만들어질 때부터 부여되는 고유한 식별번호이다. IP주소는 인터넷에 참가하기 위하여 임시적으로 할당되는 주소라면 MAC주소는 기기의 하드웨어 자체에 부여되는 고유한 번호인 것이다. MAC주소가 사용되는 한 예로, 스마트폰이 무선 랜에 접속 시 인증에 사용된다. 단말기 마다 부여된 MAC주소에 대하여 통신사는이 주소가 가입되어 있는지를 확인하여 무선 랜을 사용할 수 있게 해주는 것이다.

3. What is the difference between the network layer and the transport layer?

. 네트워크 계층의 경우 네트워크 상에서 패킷을 전달하기 위하여 라우팅을 담당하며 전송 계층은 주로 응용 계층에 서비스를 제공하는 역할을 수행한다.

우선 네트워크 계층은 host-to-host 간에 최적의 경로를 선택하는 라우팅 기능을 수행한다. 이 계층에서는 네트워크 간 통신(inter-network)이 주된 목적이기에 IP주소를 활용하며 이 IP주소를 활용하여 어떤 경로로 보낼지 결정하는 라우팅을 수행하는 것이다. (만일 inner-network 통신이라면 MAC 주소로도 충분할 것이다.) 이 계층에서도 데이터에 대한 캡슐화가 발생하며 여기서 다뤄지는 헤더를 IP헤더라고 한다. 여기에는 Source IP addr과 Dest IP addr 등이 들어있고 이렇게 캡슐화 된 패킷을 datagram이라고 한다.

그러나 많은 라우터를 경유하는 도중 데이터 유실 등과 같은 문제가 발생할 수 있는데 목적지까지 신뢰할 수 있는 데이터를 전달하기 위해 필요한 계층이 전송 계층이다.이 계층에서는 보낼 데이터의 용량과 속도, 목적지 등을 처리한다. 통신 방식에 따라두 가지로 나눌 수 있는데 신뢰성과 정확성을 추구하는 연결형 통신(TCP)과 효율성을 추구하는 비연결형 통신(UDP)이다. 패킷의 경우 TCP를 사용했다면 Segment, UDP를 사용했다면 User Datagram이라 불리게 된다. 뿐만 아니라 패킷이 어떤 어플리케이션에서 사용되는지도 구분해준다.

4. For n devices in a network, what is the number of cable links required for a mesh, ring, bus and star topology?

1) Mesh: n(n-1)/2 개

2) Ring: n 개 (+ Ring 자체를 구성하는 cable)

3) Bus: 하나의 backbone line 과 n 개의 drop line

4) Star: n 개

5. What are the RFC numbers of TCP, IP, HTTP/1.0?

1) TCP: RFC 793

2) IP: RFC 791

3) HTTP/1.0 : RFC 1945

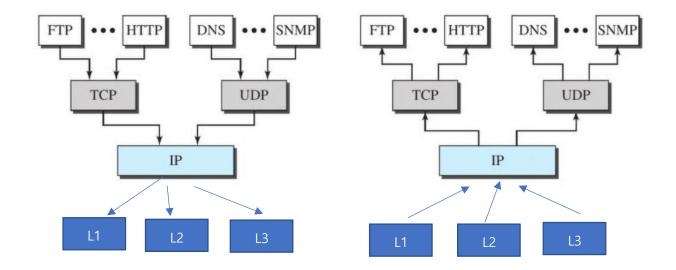
6. A color image uses 16 bits to represent a pixel. What is the maximum number of different colors that can be represented?

• 2^16

- 7. We have two computers connected by an Ethernet hub at home. Is this a LAN or a WAN? Explain the reason.
 - LAN
 - Ethernet 은 LAN Protocol 이며 집에서 두 대의 컴퓨터를 연결한 망 자체의 규모를 보았을 때도 Local Area Network에 속한다. LAN은 빌딩, 학교와 같은 비교적 가까운 거리(30m~3000m)의 네트워크 망을 의미하며, WAN은 나라에서부터 세계까지 그 규모가 확장된 네트워크 망이다.

- 8. In the bus topology in Figure 1.6, what happens if one of the stations is unplugged?
 - 스테이션들 중 하나가 unplugged 되면, 전체 시스템이 다운된다.
- 9. Match the following to one or more layers of the TCP/IP protocol suite:
 - 1) route determination : Network Layer (라우팅의 의미를 확장시켜 스위치까지 포함시킨다면 data-link Layer 도 포함된다.)
 - 2) connection to transmission media: Physical Layer
 - 3) providing services for the end user: Application Layer
- **10.** In Figure 2.10, when the IP protocol decapsulates the transport-layer packet, how does it know to which upper-layer protocol(UDP or TCP) the packet should be delivered?
 - 각 계층에서 관리되는 헤더에는 상위 계층의 어디로 가야 하는지 (Demultiplexing)에 대한 정보를 가지고 있다. 이를 통해 상위 계층의 어떤 프로토콜로 포워딩할지를 결정한다.
- **11.** Assume a private internet uses three different protocols at the data-link layer(L1, L2, and L3). Redraw Figure 2.10 with this assumption. Can we say that in the data-link layer, we have demultiplexing at the source node and multi-plexing at the destination node?
 - Multiplexing : Many to One
 - Demultiplexing : One to Many

Source node에서 멀티플렉싱이, Destination node에서 디멀티플렉싱이 가능하다.



- **12.** In an internet, we change the LAN technology to a new one,. Which layers in the TCP/IP protocol suite need to be changed?
 - Data-link 계층과 physical 계층
 - LAN 통신을 담당하는 부분은 위의 두 계층이며 네트워크 계층은 네트워크 간의 통신을 담당한다. 그렇기에 새로운 LAN 기술이 도입되더라도 모든 Communication Device 의 하위 두 계층을 바꾸면 된다. 그럼에도 data-link 계층과 네트워크 계층 간 datagram을 주고받을 수 있다면 상위 계층들은 바꾸지 않아도 되는 것이다. 이것은 Protocol Layering 의 주요 이점 중 하나이며 각 계층별로 모듈화를 함으로써, 서로에 대한 최소한의 인터페이스를 제공함으로써 결합도를 낮출 수 있는 것이다.