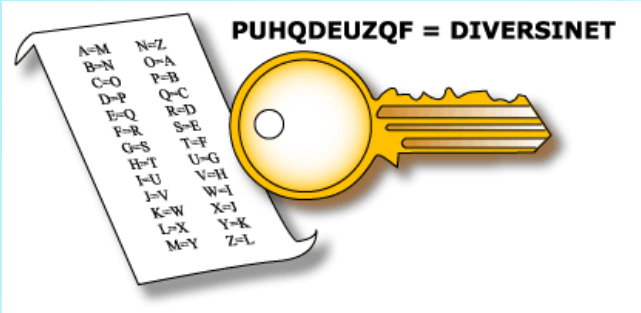
Network Security

Conventional Encryption

Caesar Cipher

```
plain:   abcdefghijklmnopqrstuvwxyz
key:     defghijklmnopqrstuvwxyzabc
```



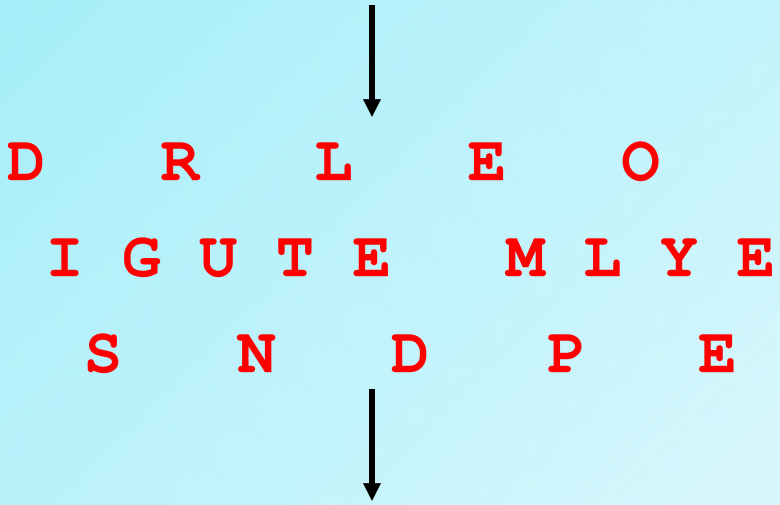

```
cipher:  PHHW PH DIWHU WKH WRJD SDUWB
plain:   MEET ME AFTER THE TOGA PARTY
```

Basic Types of Ciphers

- **Transposition ciphers** – rearrange bits or characters in the data
- **Substitution ciphers** – replace bits, characters, or blocks of characters with substitutes

“Rail-Fence” Cipher

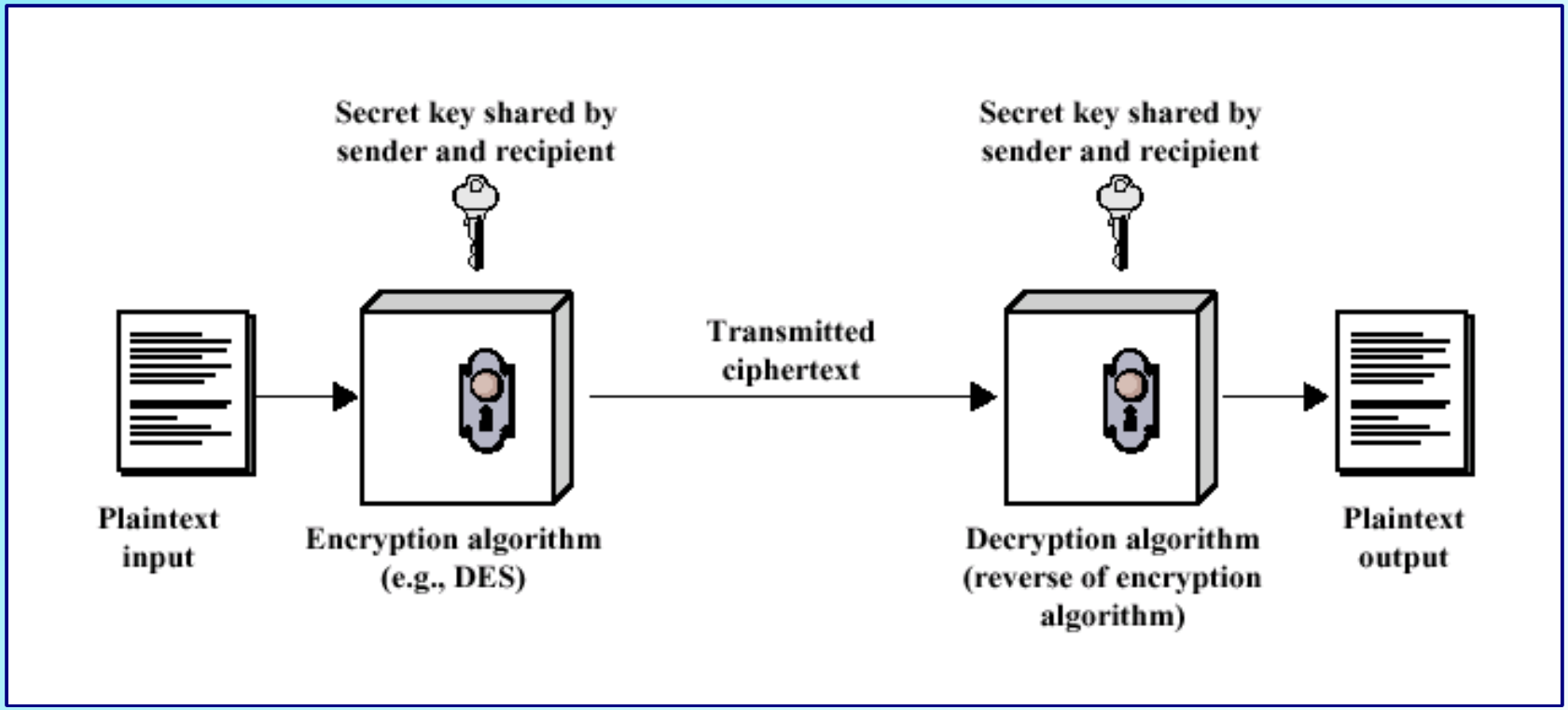
DISGRUNTLED EMPLOYEE



Encryption Methods

- The essential technology underlying virtually all automated network and computer security applications is **cryptography**
- Two fundamental approaches are in use:
  - **Conventional Encryption**, also known as symmetric encryption
  - **Public-key Encryption**, also known as asymmetric encryption

Conventional Encryption Model



Conventional Encryption

- The only form of encryption prior to late 1970s
- Long history
- Most widely used

Conventional Encryption

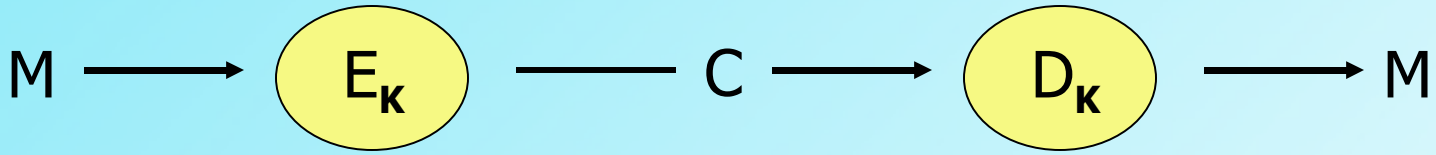
Five components to the algorithm

- **Plaintext**: The original message or data
- **Encryption algorithm**: Performs various substitutions and transformations on the plaintext
- **Secret key**: Input to the encryption algorithm. Substitutions and transformations performed depend on this key
- **Ciphertext**: Scrambled message produced as output. depends on the plaintext and the secret key
- **Decryption algorithm**: Encryption algorithm run in reverse. Uses ciphertext and the secret key to produce the original plaintext

Conventional Encryption

- More rigorous definition
- Five components to the algorithm
  - A Plaintext message space, $\mathcal{M}$
  - A family of enciphering transformations, $E_K: \mathcal{M} \rightarrow \mathcal{C}$, where $K \in \mathcal{K}$
  - A key space, $\mathcal{K}$
  - A ciphertext message space, $\mathcal{C}$
  - A family of deciphering transformations, $D_K: \mathcal{C} \rightarrow \mathcal{M}$, where $K \in \mathcal{K}$

Conventional Encryption



E_K defined by an encrypting algorithm E
D_K defined by an decrypting algorithm D

For given K, D_K is the **inverse** of E_K, i.e.,

$$D_K(E_K(M))=M$$

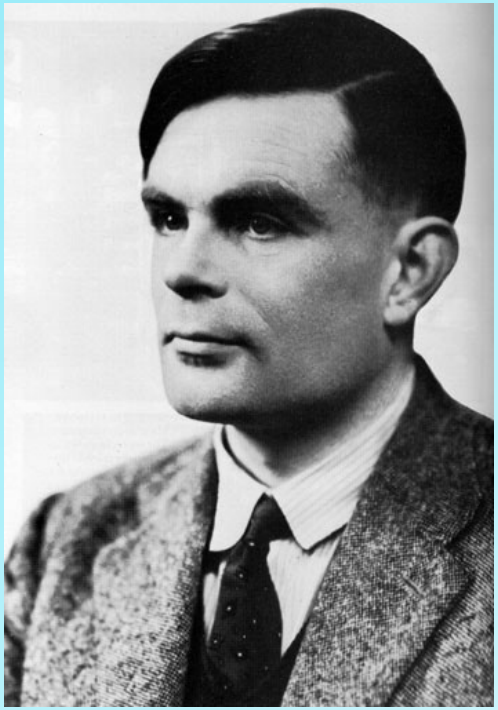
for every plain text message M

Requirements & Weaknesses

- Requirements
  - A **strong** encryption algorithm
  - Secure **process** for sender & receiver to **obtain secret keys**
- Methods of Attack
  - **Cryptanalysis**
  - **Brute force**

Cryptanalysis

- The process of attempting to discover the plaintext or key



Alan Turing broke the Enigma Code in WWII

Cryptanalysis

- Security depends on the key...
- ...NOT the secrecy of the algorithm
- Low cost chips are possible
- Principal security problem is maintaining the secrecy of the key!

Cryptographic Systems

- **Type of Transformation** – substitution and/or transposition; no information must be lost, i.e., reversible
- **Number of Keys Used** – symmetric, single key, conventional; asymmetric, two-key, public-key encryption
- **Plaintext Processing** – block or stream cipher

Attacks On Encrypted Msgs

Type of Attack	Known to Cryptanalyst
Ciphertext only	•Encryption algorithm •Ciphertext to be decoded
Known plaintext	•Encryption algorithm •Ciphertext to be decoded •One or more plaintext-ciphertext pairs formed with the secret key
Chosen plaintext	•Encryption algorithm •Ciphertext to be decoded •Plaintext message chosen by cryptanalyst, together with its corresponding ciphertext generated with the secret key
Chosen ciphertext	•Encryption algorithm •Ciphertext to be decoded •Purported ciphertext chosen by cryptanalyst, together with its corresponding decrypted plaintext generated with the secret key
Chosen text	•Encryption algorithm •Ciphertext to be decoded •Plaintext message chosen by cryptanalyst, together with its corresponding ciphertext generated with the secret key •Purported ciphertext chosen by cryptanalyst, together with its corresponding decrypted plaintext generated with the secret key

Computationally Secure

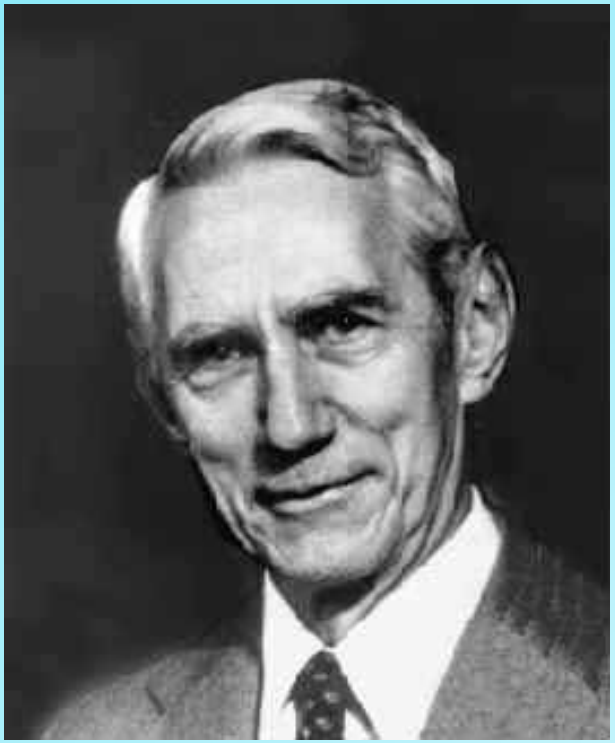
- Cost of breaking cipher exceeds value of encrypted information
- Time to break cipher exceeds useful lifetime of the information

Exhaustive Key Search

Key Size (bits)	Number of Alternative Keys	Time required at 1 encryption/μs	Time required at 10^6 encryptions/μs
32	$2^{32} = 4.3 \times 10^9$	2^{31} μs = 35.8 minutes	2.15 milliseconds
56	$2^{56} = 7.2 \times 10^{16}$	2^{55} μs = 1142 years	10.01 hours
128	$2^{128} = 3.4 \times 10^{38}$	2^{127} μs = 5.4×10^{24} years	5.4×10^{18} years
168	$2^{168} = 3.7 \times 10^{50}$	2^{167} μs = 5.9×10^{36} years	5.9×10^{30} years
26 characters (permutation)	$26! = 4 \times 10^{26}$	2×10^{26} μs = 6.4×10^{12} years	6.4×10^6 years

Brute Force with massively parallel processors

Claude Shannon



- *A Mathematical Theory of Communication* (1948), outlining what we now know as Information Theory
- Described ways to measure data using the quantity of disorder in any given system, together with the concept of entropy
- *The Magna Carta of the information age*
- Retired at age 50

Claude Shannon

- Concept of entropy - equivalent to a shortage in the information content in a message
- Second law of thermodynamics – **entropy** is the degree of randomness in any system
- Many sentences can be significantly shortened without losing their meaning
- Shannon proved that in a noisy conversation, signal could always be send without distortion

Claude Shannon

- If the message is encoded in such a way that it is **self-checking**, signals will be received with the same accuracy as if there were no interference on the line
- A language has a built in **error-correcting code**
- http://cm.bell-labs.com/cm/ms/what/shannonday
- http://cm.bell-labs.com/cm/ms/what/shannonday/paper.html

Information Theory

- Information theory measures the **amount of information** in a message by the average number of bits needed to encode all possible messages in an optimal encoding
- SEX field in a database: only one bit of information (Male:0; Female:1)
- Encoded in ASCII – more space, but *no more information*

Information Theory

- **Amount of information** in a message is formally measured by the **entropy** of the message
- **Entropy** is a function of the probability distribution over the set of all possible messages

Information Theory

- **Entropy** of a given message is defined by the weighted average over all possible messages X:

$$H(X) = \sum_{X} p(X) \log_2\left(\frac{1}{p(X)}\right)$$

Information Theory Example

p(male) = p(female) =1/2, then

$$H(X) = \frac{1}{2}\left(\log_2 2\right) + \frac{1}{2}\left(\log_2 2\right)$$

$$= \frac{1}{2} + \frac{1}{2} = 1$$

- There is 1 bit of information in the SEX field of a database

Information Theory

- Text files can be reduced by about 40% without losing information
- Because $1/p(x)$ decreases as $p(x)$ increases, an **optimal encoding** uses *short codes for frequently occurring messages; longer codes for infrequent*
- *Morse code*
  E •, T –, J • – – –, Z – – • •

Information Theory

- The entropy of a message measures its **uncertainty**. The number of bits that must be learned when the message is hidden in ciphertext
- **English** is a highly **redundant**
- occurring frequently => ocrng frq

English Redundancy

- Delete vowels and double letters

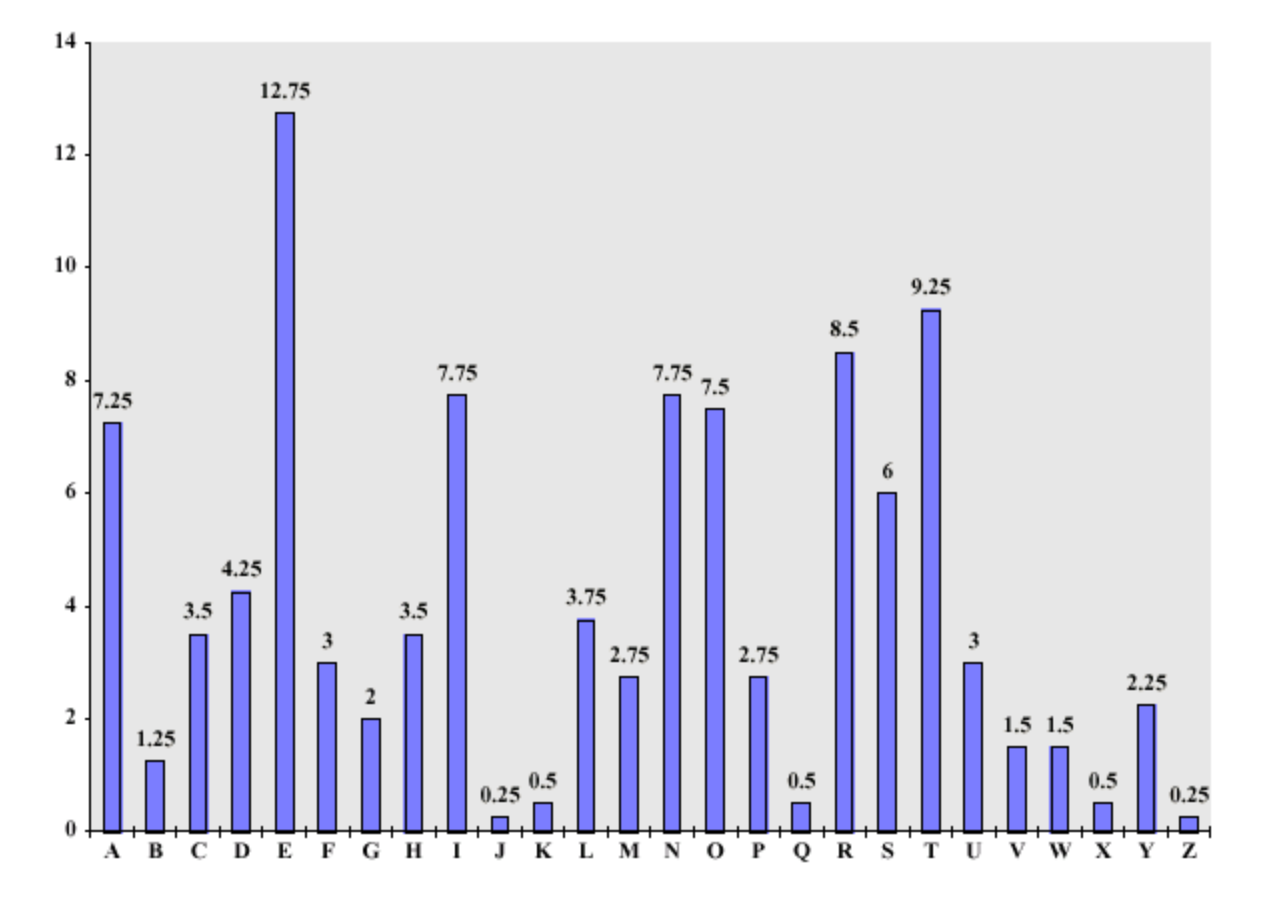
mst ids cn b xprsd n fwr ltrs,
bt th xprnc s mst nplsnt

Simple Cryptanalysis

CIPHERTEXT:

UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ

Letter Frequency In the English Language



Simple Cryptanalysis

PLAINTEXT:

IT WAS DISCLOSED YESTERDAY THAT SEVERAL INFORMAL BUT DIRECT CONTACTS HAVE BEEN MADE WITH POLITICAL REPRESENTATIVES OF THE VIET CONG IN MOSCOW

20th Century Encryption

- 20's & 30's bootleggers made heavy use of cryptography
- FBI create an office for code-breaking
- Japanese Purple Machine
- German Enigma Machine
- Navajo Code Talkers - Windtalkers

Hedy Lamarr

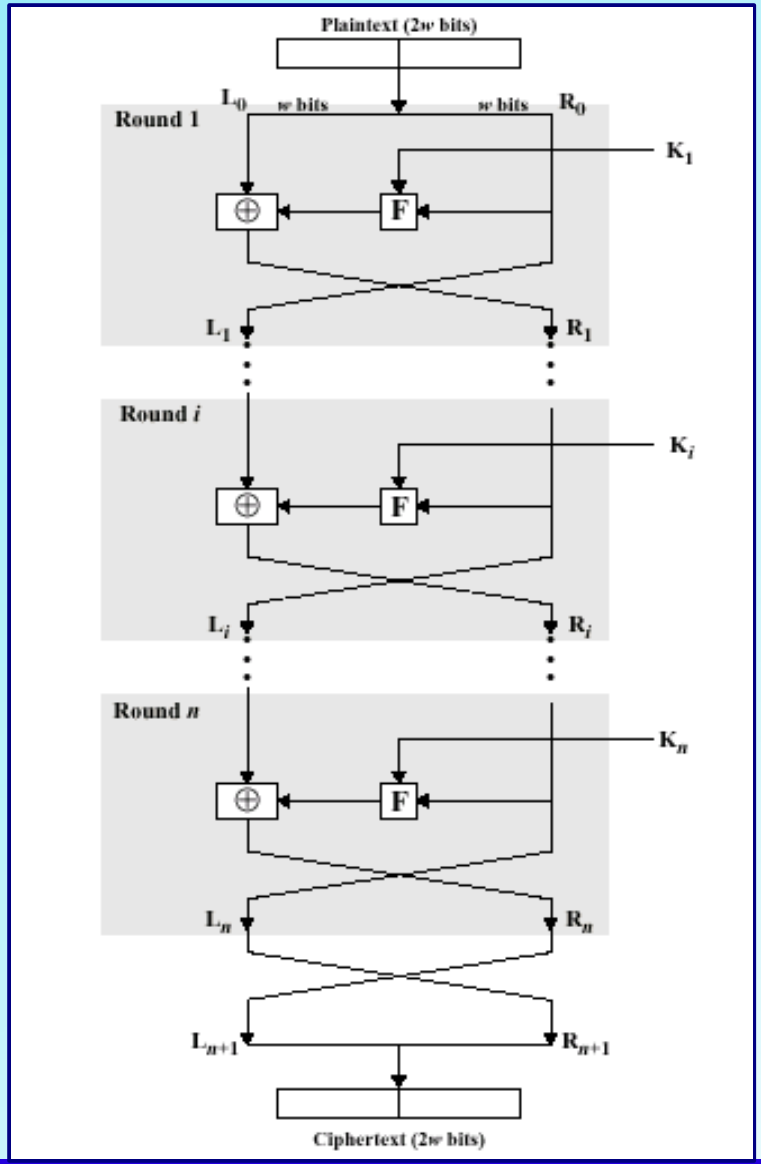
- 1941, Lamarr and composer George Antheil received a patent for their invention of a classified communication system that was especially useful for submarines
- It was based on radio frequencies changed at irregular periods that were synchronized between the transmitter and receiver
- **Spread Spectrum** – wireless devices

Feistel Cipher Structure

- **Horst Feistel** of IBM, 1973
- Input is plaintext block of length $2w$ bits (usually 64) and a key K
- Block is divided into two halves, L_0 and R_0
- Each round i has inputs L_{i-1} and R_{i-1}, derived from the previous round, along with subkey K_i
- Substitution is performed on the left half of the data
- **Round function** F applied to right half and then XOR'd with left

Feistel Cipher Structure

Things to consider:
-Block size (64)
-Key Size (128)
-# of rounds (16)
-SubKey Generation
-Round function

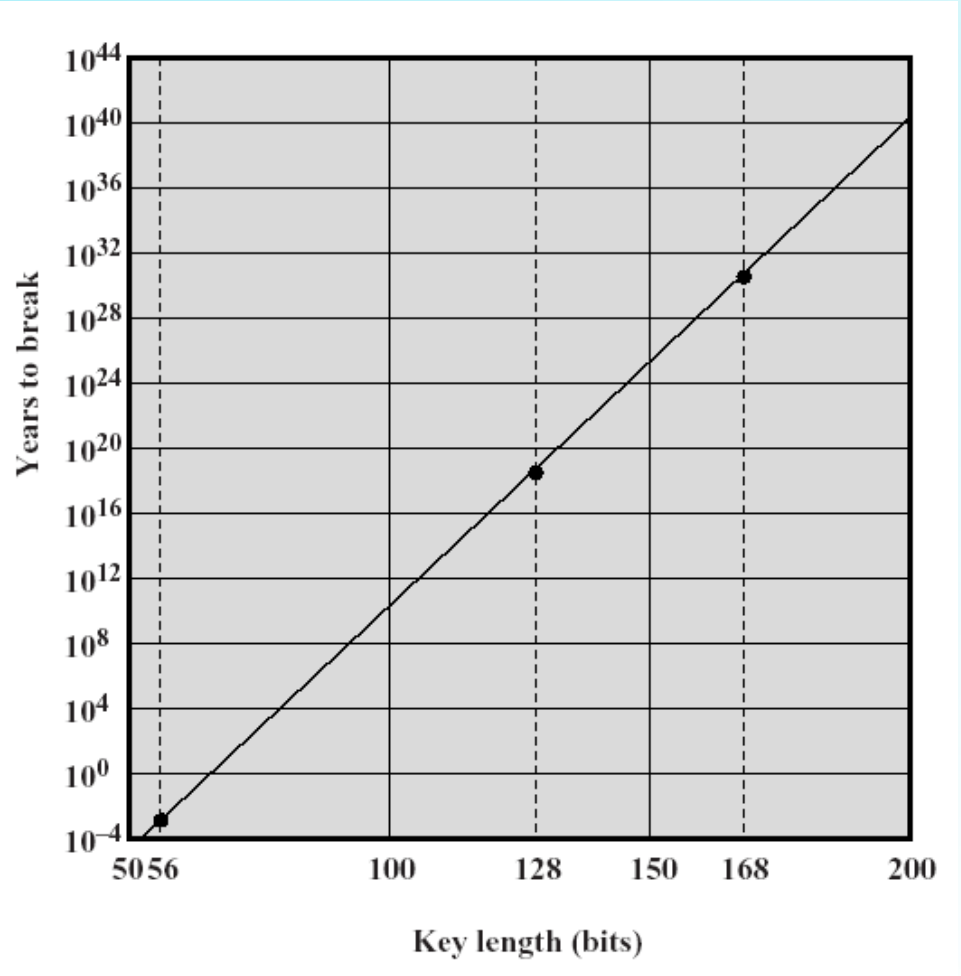


Data Encryption Standard (DES)

- Adopted in 1977, reaffirmed for 5 years in 1994, by NBS(NIST)
- Plaintext is 64 bits (or blocks of 64 bits), key is 56 bits
- Plaintext goes through 16 iterations, each producing an intermediate value that is used in the next iteration
- DES is now too easy to crack to be a useful encryption method

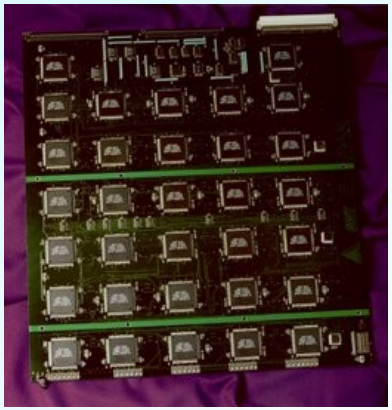
Strength of DES

- Concerns about the algorithm itself
- Concerns about 56-bit key – this is the biggest worry



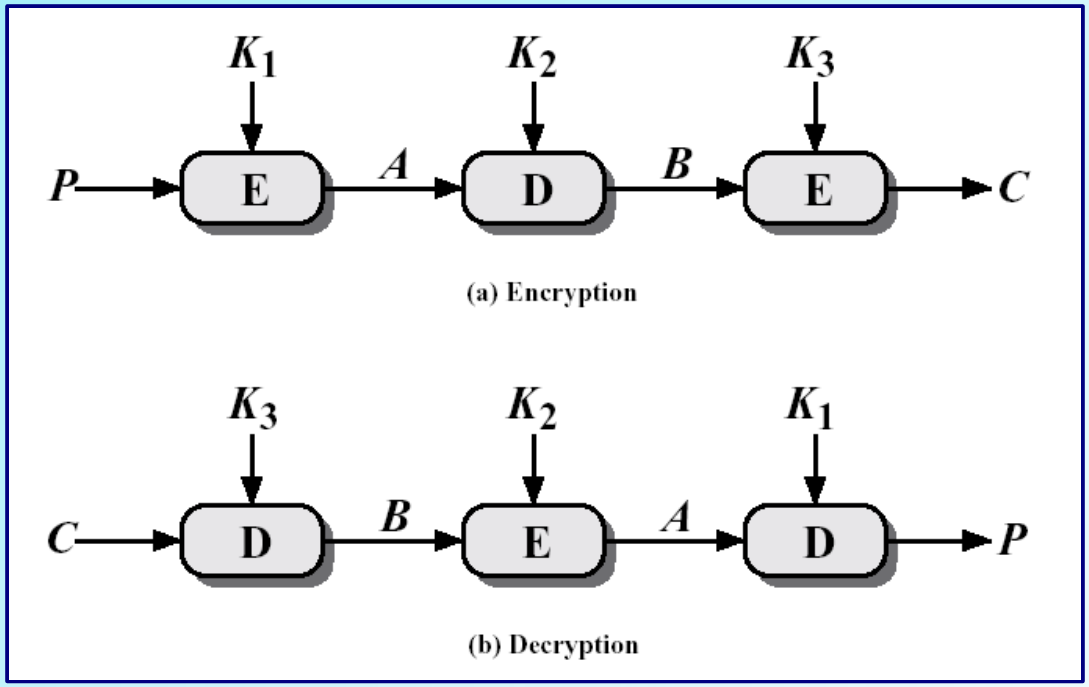
Strength of DES

- DES is the most studied encryption algorithm in existence
- No one has succeeded in discovering a fatal weakness
- 1998, DES Cracker from Electronic Frontier Foundation, built for $250,000
- Solution: Use a bigger key



Triple DES

$$C = E_{K_3}\,[D_{K_2}\,[E_{K_1}\,[P\,]]]$$



Triple DES

- Alternative to DES, uses multiple encryption with DES and multiple keys
- With three distinct keys, 3DES has an effective key length of 168 bits, so it is essentially immune to brute force attacks
- Backward compatible with DES
- Principal drawback of DES is that the algorithm is relatively sluggish in software

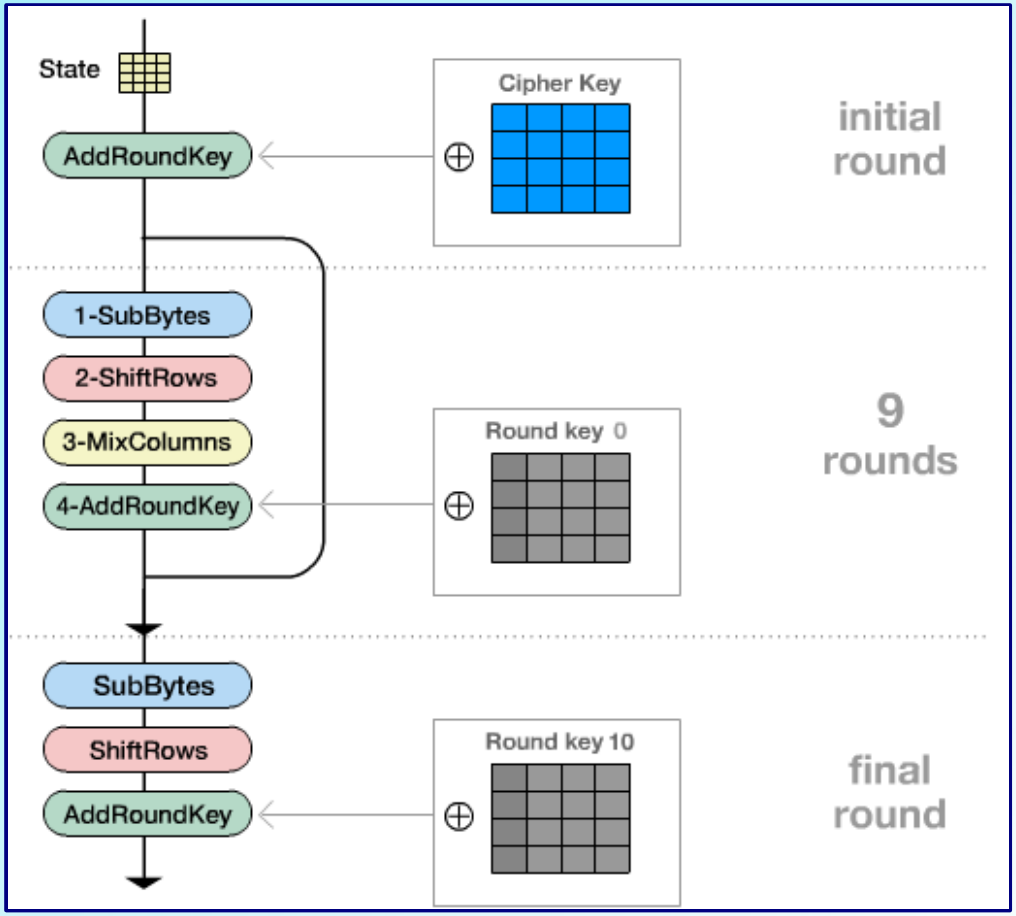
Advanced Encryption Standard

- NIST call for proposals in 1997
- Nov, 2001 – **Rijndael** [rain´ dow]
- Symmetric block cipher (128 bits) and key lengths 128, 192, 256
- Two Flemish cryptographers: Joan Daeman and Vincent Rijmen

Overview of AES

4Transformations:

- Substitute Bytes
- Shift Rows
- Mix Columns
- Add Round Key



AES URLS

- http://csrc.nist.gov/CryptoToolkit/aes/rijnd
  - NIST AES
- http://www.esat.kuleuven.ac.be/~rijmen/r
  - Rijndael Home Page
- http://www.esat.kuleuven.ac.be/~rijmen/r
  - Great Animation

IDEA International Data Encryption Algorithm

- 1991 by Swiss Federal Institute of Technology
- Uses 128-bit key
- Complex functions replace S-boxes
- Highly resistant to cryptanalysis
- Used in PGP

Blowfish

- 1993 by Bruce Schneier
- Easy to implement; high execution speed
- Variable key length up to 448 bits
- Used in a number of commercial applications

RC5

- 1994 by Ron Rivest, one of the inventors of RSA algorithm
- Defined in RFC2040
- Suitable for hardware and software
- Simple, fast, variable length key, low memory requirements
- High security

CAST-128

- 1997, Entrust Technologies
- RFC 2144
- Extensively reviewed
- Variable key length, 40-128 bits
- Used in PGP

Conventional Encryption Algorithms

Algorithm	Key Size (bits)	Block Size (bits)	Number of Rounds	Applications
DES	56	64	16	SET, Kerberos
Triple DES	112 or 168	64	48	Financial key management, PGP, S/MIME
AES	128, 192, or 256	128	10, 12, or 14	Intended to replace DES and 3DES
IDEA	128	64	8	PGP
Blowfish	variable to 448	64	16	Various software packages
RC5	variable to 2048	64	variable to 255	Various software packages

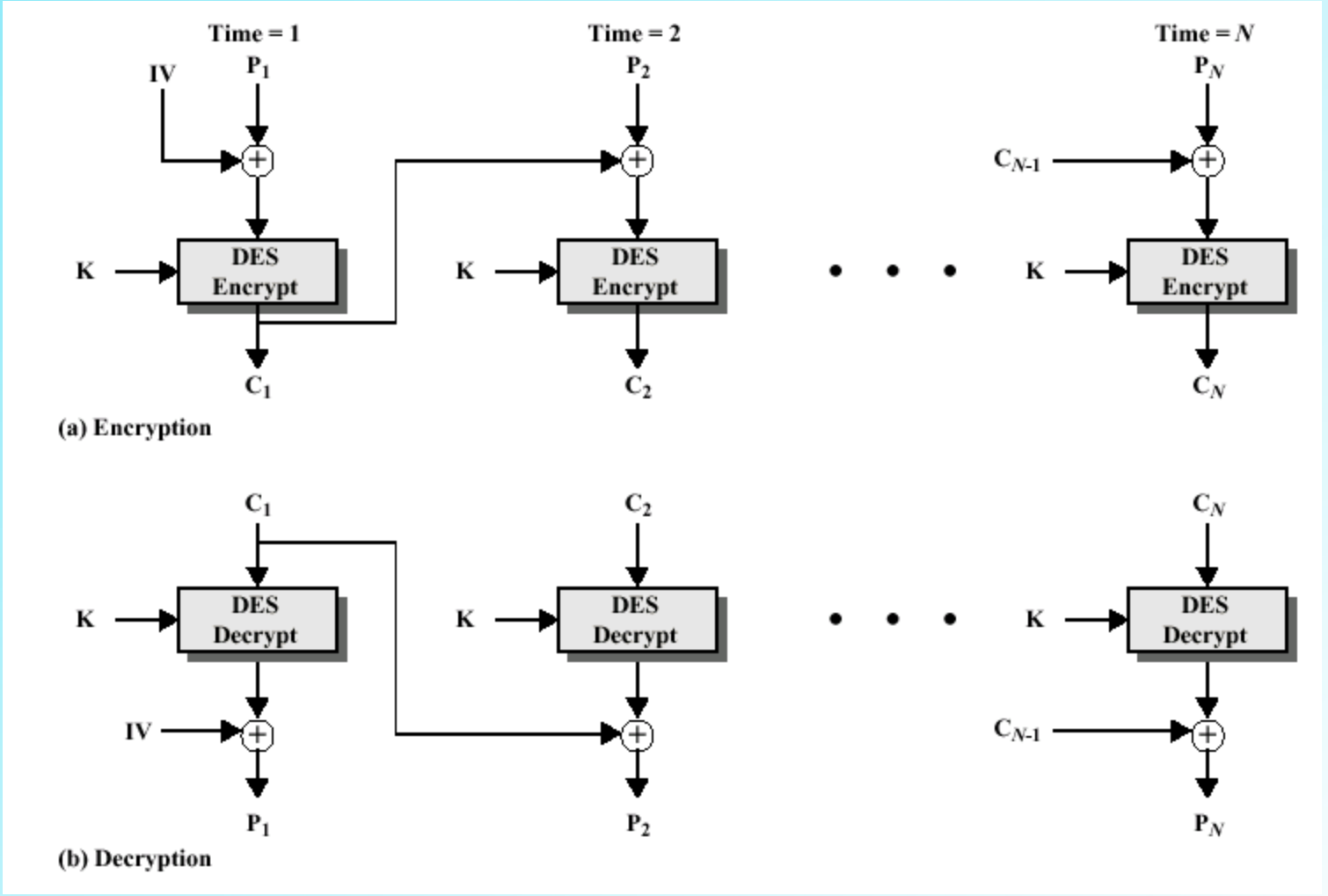
Cipher Block Modes of Operation

- Block ciphers process **one n-bit block** of data at a time
- Break long amounts of plaintext into **64-bit blocks**
- Use **Electronic Code Book** (ECB)
  - Each block of plaintext is encrypted using the same key
  - Entry for every possible 64-bit plaintext pattern
  - Block appears more than once, produce same ciphertext
  - Repeating patterns become a problem

Cipher Block Chaining Mode

- Input to algorithm is the **XOR** of current plaintext block and preceding ciphertext block
- **Repeating patterns** are **not** exposed

Cipher Block Chaining Mode



Cipher Feedback Mode

- Convert DES into a **stream cipher**
- **Eliminates** need to **pad** a message
- Operates in **real time**
- Each character can be **encrypted** and **transmitted immediately**

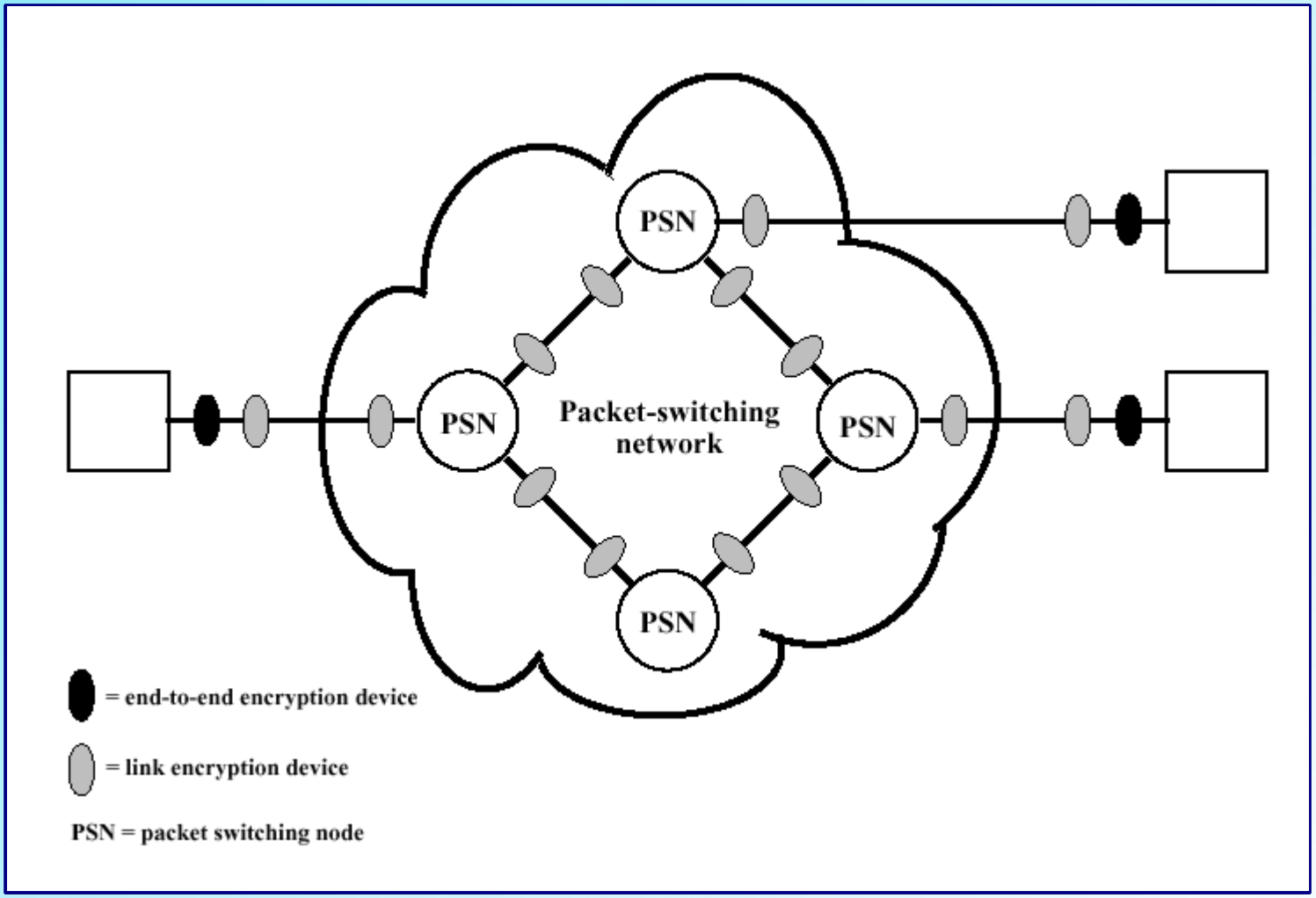
Location of Encryption Devices

- **Link Encryption**
  - Each vulnerable communications link is equipped on both ends with an encryption device
  - All traffic over all communications links is secured
  - Vulnerable at each switch

Location of Encryption Devices

- **End-to-end Encryption**
  - The encryption process is carried out at the two end systems
  - Encrypted data are transmitted unaltered across the network to the destination, which shares a key with the source to decrypt the data
  - Packet headers cannot be secured

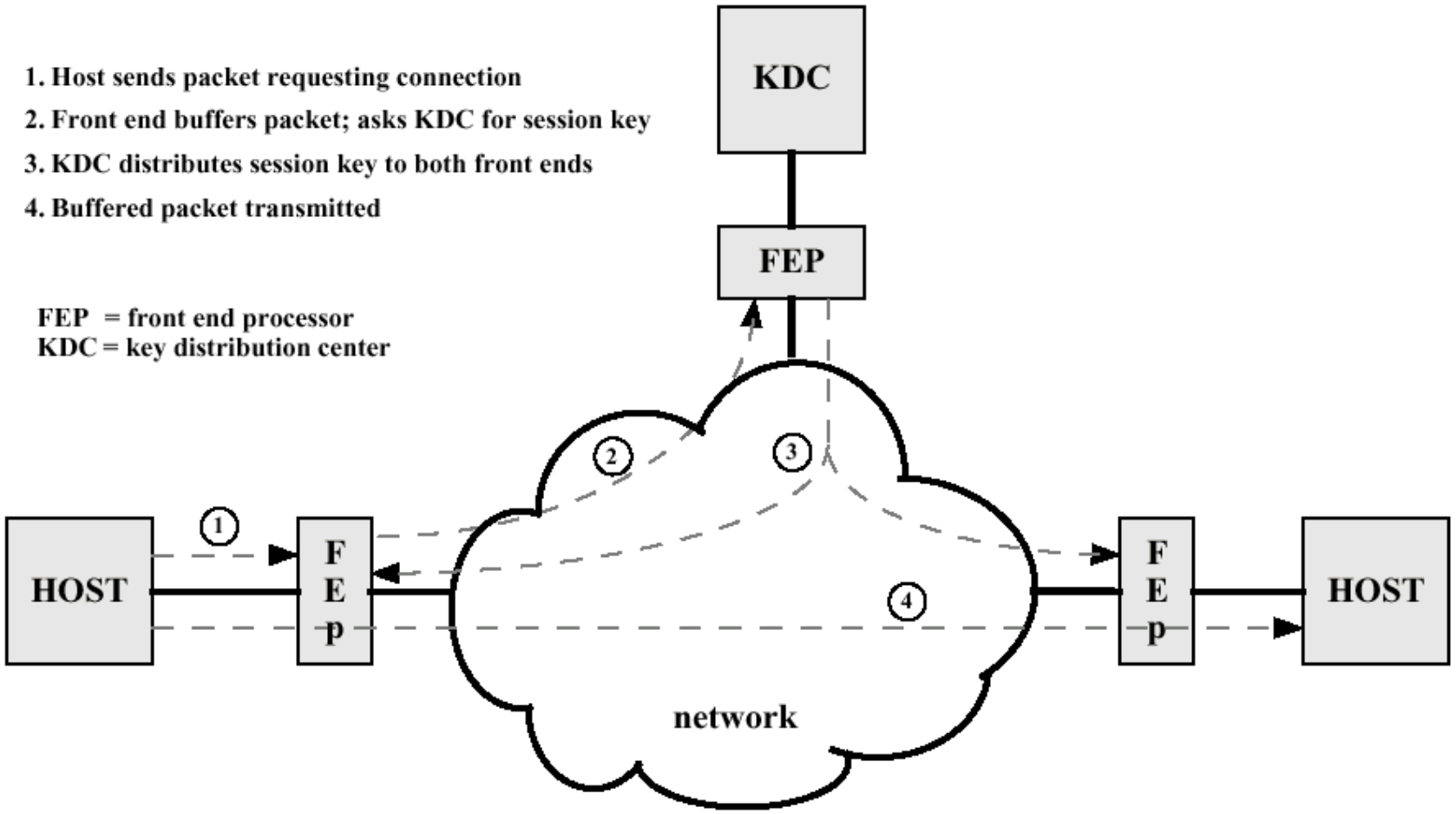
Location of Encryption Devices



Key Distribution

- **Both parties** must have the **secret key**
- Key is **changed frequently**
- Requires either manual **delivery** of keys, or a third-party encrypted channel
- Most effective method is a **Key Distribution Center** (e.g. Kerberos)

Key Distribution



Network Security

DNS & Addressing

Internet History

- Evolved from ARPANet (Defense Department's Advanced Research Projects Agency Network)
- ARPANet was developed in 1969, and was the first packet-switching network
- Initially, included only four nodes: UCLA, UCSB, Utah, and SRI

NSF and the Internet

- In the 1980s, NSFNet extended packet-switched networking to non-ARPA organization; eventually replaced ARPANet
- Instituted Acceptable Use Policies to control use
- CIX (Commercial Internet eXchange) was developed to provide commercial internetworking

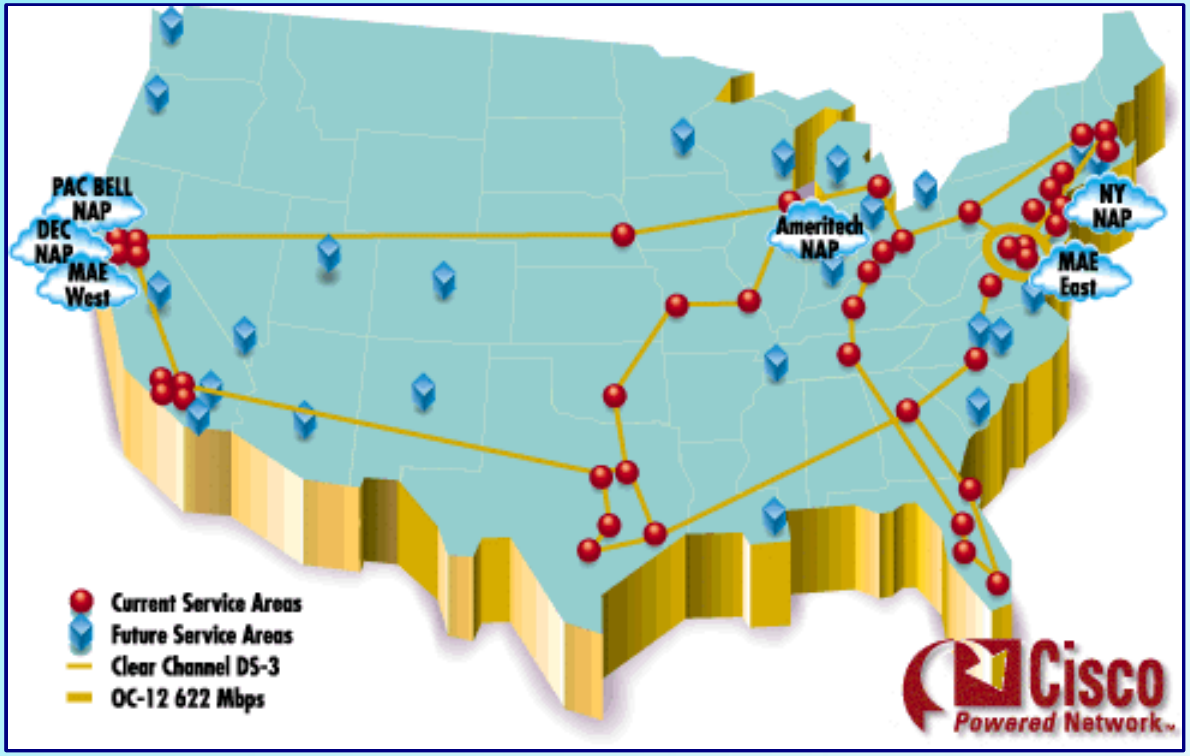
The World Wide Web

- Concept proposed by **Tim Berners-Lee** in 1989, prototype WWW developed at CERN in 1991
- First graphical browser (**Mosaic**) developed by Mark Andreessen at NCSA
- Client-server system with browsers as clients, and a variety of media types stored on servers
- Uses **HTTP** (**H**yper **T**ext **T**ransfer **P**rotocol) for retrieving files

Connecting to the Internet

- End users get connectivity from an ISP (Internet Service Provider)
  - Home users use dial-up, ADSL, cable modems, satellite, wireless
  - Businesses use dedicated circuits connected to LANs
- ISPs use “wholesalers” called network service providers and high speed (T-3 or higher) connections

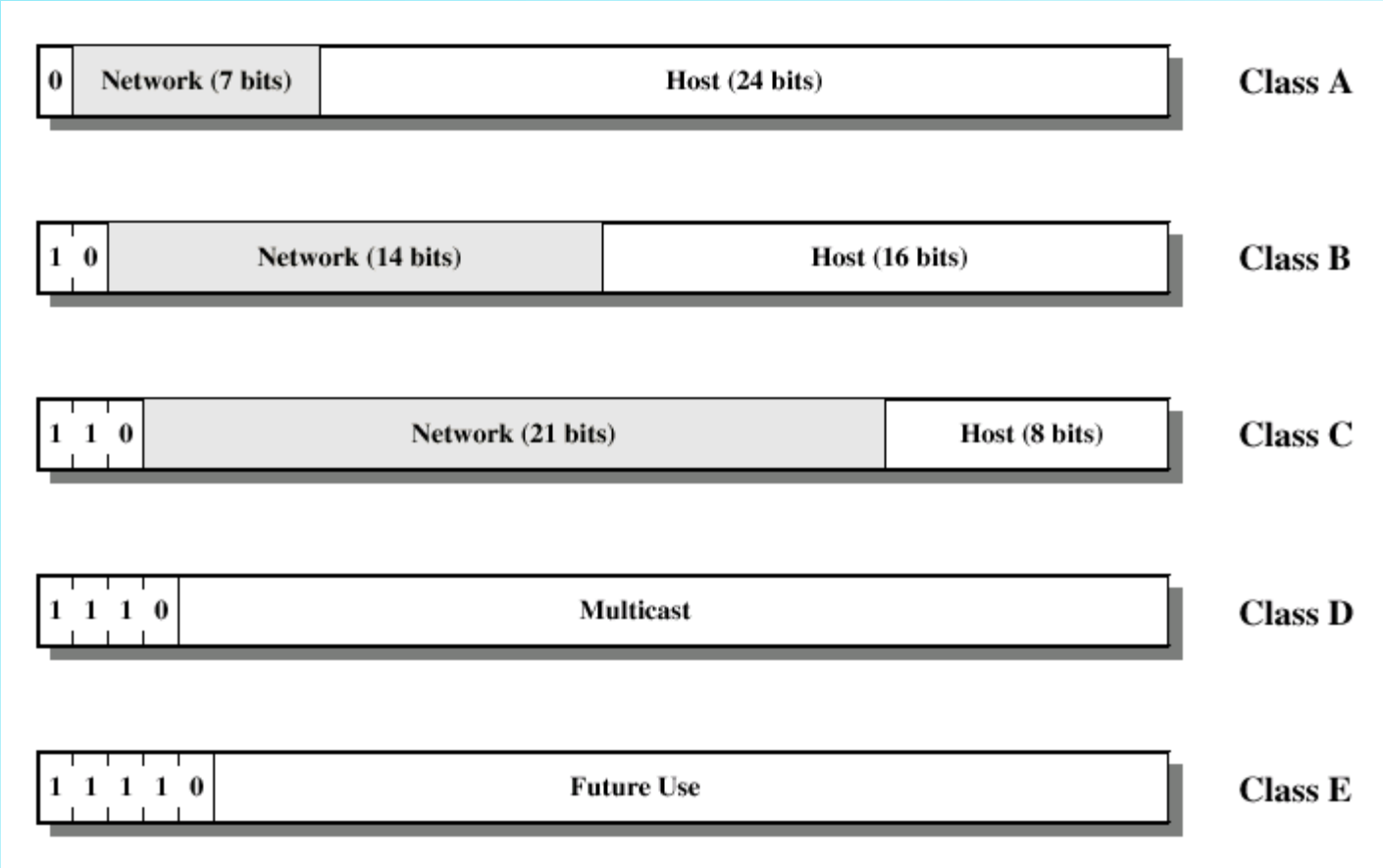
US Internet Access Points



Internet Addressing

- 32-bit global Internet address
- Includes network and host identifiers
- Dotted decimal notation
  - 11000000 11100100 00010001 00111001 (binary)
  - 192.228.17.57 (decimal)

Internet Addressing



Network Classes

- Class A: Few networks, each with many hosts
  All addresses begin with binary 0
  Range: 1-126
- Class B: Medium networks, medium hosts
  All addresses begin with binary 10
  Range: 128-191
- Class C: Many networks, each with few hosts
  All addresses begin with binary 11
  Range: 192-223

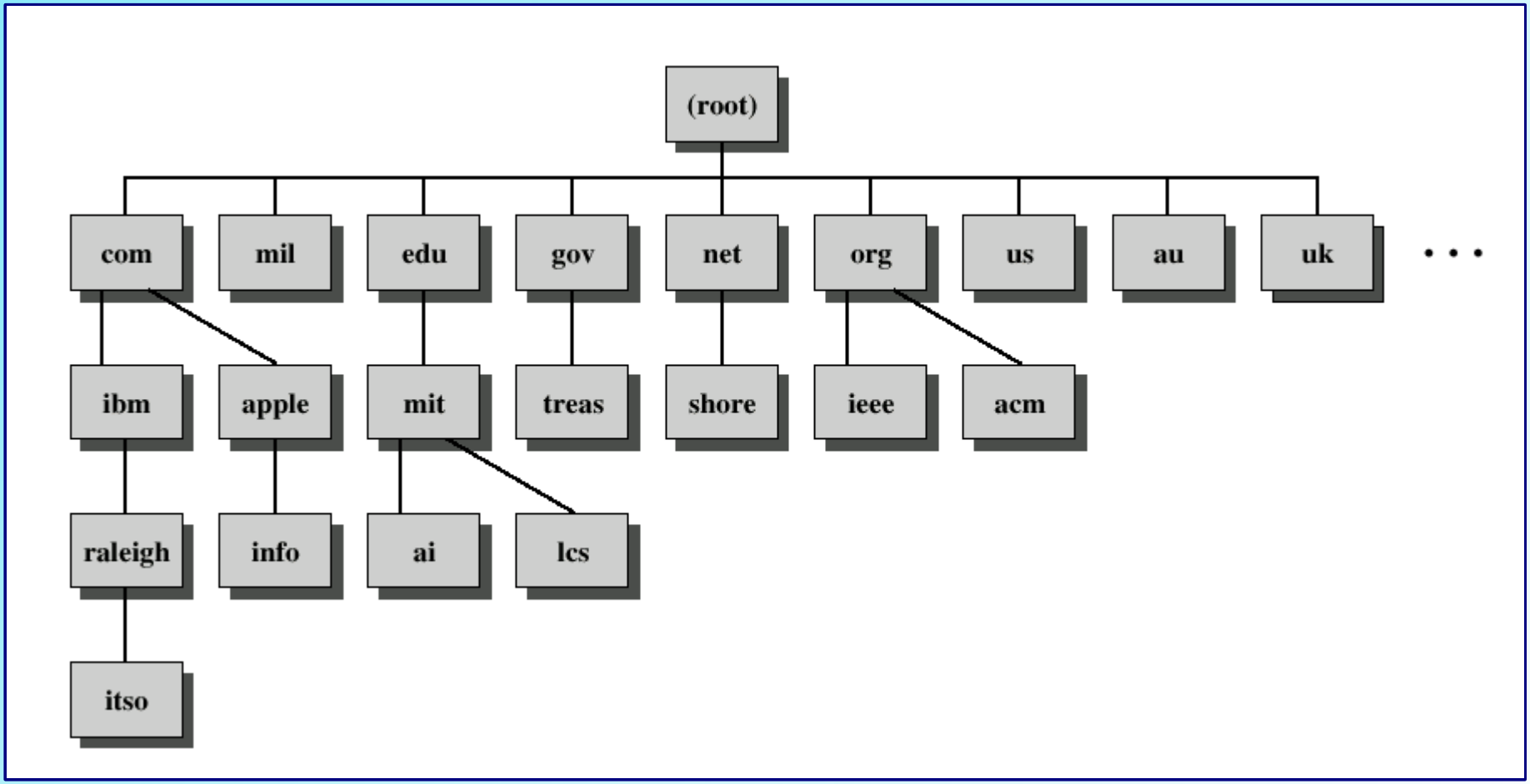
Domain Name System

- 32-bit IP addresses have two drawbacks
  - Routers can’t keep track of every network path
  - Users can’t remember dotted decimals easily
- **Domain names** address these problems by providing a name for each network domain (hosts under the control of a given entity)

DNS Database

- **Hierarchical database** containing name, IP address, and related information for hosts
- Provides **name-to-address** directory services

Domain Tree


(root)
com
mil
edu
gov
net
org
us
au
uk
• • •
ibm
apple
mit
treas
shore
ieee
acm
raleigh
info
ai
lcs
itso

Important URLs

- http://www.networksolutions.com/whois/index.jhtml
  The original InterNIC. This site has the “whois” database
- http://www.arin.net
  American registry for Internet numbers. This site has a “whois” database for IP numbers
- http://www.net.princeton.edu/traceroute.htmlhttp://www.tracert.com/Handy tools: traceroute, ping, nslookup, whois, dig

Homework

- Read **Chapter Two**
- Examine some sites using **whois** and **traceroute** for the domain name and the IP address. See how much you can find out about a site