Chapter 4

by

David G. Messerschmitt

Technical properties of information

by

David G. Messerschmitt
Two fundamental concepts

• The computer is the first machine whose functionality is not determined at the time of manufacture
  – added by software later

• Any form of information can be represented or approximated within the information technology infrastructure

Key concept

• The key commodity manipulated by information technology is information

• To be manipulated in a computing/networking environment, information must be represented by data

What is information?
Information

• From a user (human) perspective…
  ....recognizable patterns that influence you in some way
  (perspective, understanding, behavior…)
• In the computing infrastructure, information
  has a somewhat different connotation as structure and interpretation added to data

Data

• A bit is “0” or “1” — the atom of the information economy
• Data is a collection of bits, like
  – “0101110111010110”
  – “0000011”
  – “11101110111010101101111011011010”
• Note: the terms data and information are not always used consistently!
Representation

- Take the place of the original
- Equivalent to, in the sense that the original can be reconstructed from its representation
- Often the original can only be approximately reconstructed, although it may be indistinguishable to the user
  - e.g. audio or video

ASCII

<table>
<thead>
<tr>
<th>Alphabet</th>
<th>Hex</th>
<th>Binary</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;7&gt;</td>
<td>/x37</td>
<td>00110111</td>
</tr>
<tr>
<td>&lt;8&gt;</td>
<td>/x38</td>
<td>00111000</td>
</tr>
<tr>
<td>&lt;9&gt;</td>
<td>/x39</td>
<td>00111001</td>
</tr>
<tr>
<td>&lt;:&gt;</td>
<td>/x3A</td>
<td>00111010</td>
</tr>
<tr>
<td>;&gt;</td>
<td>/x3B</td>
<td>00111011</td>
</tr>
<tr>
<td>&lt;&lt;&gt;</td>
<td>/x3C</td>
<td>00111100</td>
</tr>
<tr>
<td>&lt;=&gt;</td>
<td>/x3D</td>
<td>00111101</td>
</tr>
<tr>
<td>=&gt;</td>
<td>/x3E</td>
<td>00111110</td>
</tr>
<tr>
<td>&gt;?</td>
<td>/x3F</td>
<td>00111111</td>
</tr>
<tr>
<td>&lt;At&gt;</td>
<td>/x40</td>
<td>01000000</td>
</tr>
<tr>
<td>&lt;A&gt;</td>
<td>/x41</td>
<td>01000001</td>
</tr>
<tr>
<td>&lt;B&gt;</td>
<td>/x42</td>
<td>01000010</td>
</tr>
<tr>
<td>&lt;=</td>
<td>/x43</td>
<td>01000011</td>
</tr>
<tr>
<td>&lt;D&gt;</td>
<td>/x44</td>
<td>01000100</td>
</tr>
</tbody>
</table>

Note that this representation is not unique…

…this one happens to be a standard (ANSI X3.110-1983)
A picture

This picture conveys information

This information is represented in this computer, but how?

Representation of picture: image

Expanding a small portion of the picture, we see that it is represented by square pixels…

….300 tall by 200 wide…..

….with a range of 256 intensities per pixel

An approximation!

300 • 200 • 8 bits = 480,000 bits (but it can be compressed)
A color picture can be represented by three monochrome images…

At the expense of three times as many bits

Representation needs to be standardized

If the representation is not standardized, the information is garbled!

Communicate data to another user or organization
**Regeneration**

- Make a precise copy of the data (copy bit by bit)
- If you know the representation, this is equivalent to making a precise copy of the information
- Each such precise copy is called a generation, process is called regeneration

**Replication of information**

Anything that can be regenerated can be replicated any number of times

This is a blessing and a curse
Analog information cannot be regenerated

Analog information can be copied, but not regenerated

We will never know exactly what the original of this Rembrandt looked like

Discrete information can be regenerated

Regeneration can preserve data (but not its original physical form)

Regeneration is possible for information represented digitally (which is tolerant of physical deterioration)

\[ 0 + \text{noise} \rightarrow 0 \]
\[ 1 + \text{noise} \rightarrow 1 \]
Example

Analog information

Digital information

Replication of information requires knowledge of representation

- Replication of information also presumes knowledge of its representation

- Replication preserves the integrity of the data, but that is not sufficient

Every .xxx DOS file is a representation
Implications

• Digitally represented information can be preserved over time or distance in its precise original form by occasional regeneration
  – digital library
  – digital telephony

• Replication of data is easy and cheap

Implications (con’t)

• Replication of information requires knowledge of the structure and interpretation
  – Standardization or some other means

• Extreme supply economies of scale

• You can give away or sell and still retain

• Unauthorized replication or piracy relatively easy
Understanding Networked Applications: A First Course

Architecture

by

David G. Messerschmitt

Outline

• Architecture
  – Decomposition
  – Modularity
  – Interfaces
• Hardware
• Software
A system is decomposed into interacting subsystems.

Each subsystem may have a similar internal decomposition.

A company is organized into interacting divisions.

Each division may be organized into departments.
Three elements of architecture

Decomposition
Organization
Functionality
Responsibility
Interaction
Cooperation

Some building blocks

User
Client
Communications
Software
Server

Which of these can be subsystems?
System examples

- Let’s quickly look at some system decomposition examples
- Quick tour of information technology systems

Time sharing

Point-to-point wire
(no network)

Mainframe
(database and application server)

ASCII terminal
(no graphics)
Two-tier client/server

- Micro/server
- Local-area network
- Mainframe

Three-tier client/server

- Client
- Application server
- Enterprise data server
Inter-organizational computing

Consumer access
Emergence

• Subsystems are more specialized and simpler functionality
• Higher-level system functionality arises from the interaction of subsystems
• Emergence includes capabilities that arise purely from that interaction (desired or not)
  – e.g. airplane flies, but subsystems can’t
System integration

Architecture ➔ subsystem implementation ➔ system integration

• Bring together subsystems and make them cooperate properly to achieve desired system functionality
  – Always requires testing
  – May require modifications to architecture and/or subsystem implementation

Why system decomposition?

• Divide and conquer approach to containing complexity
• Reuse
• Consonant with industry structure (unless system is to be supplied by one company)
• Others?
Networked computing infrastructure

by

David G. Messerschmitt

Major subsystems

Presentation software

Application software

Logic Data

Infrastrcuture software

Infrastructure equipment

Client host

Network

Server host
Layering builds capability incrementally by adding to what exists.
Understanding Networked Applications A First Course

Simplified infrastructure layering

<table>
<thead>
<tr>
<th>Application</th>
<th>Distributed object management</th>
<th>Database management</th>
<th>Middleware management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network software</td>
<td>File system</td>
<td>Operating system</td>
<td></td>
</tr>
<tr>
<td>Network equipment</td>
<td>Storage peripherals</td>
<td>Equipment</td>
<td></td>
</tr>
<tr>
<td>Communications</td>
<td>Storage</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Operating system functions

- Graphical user interface (client only)
- Hide details of equipment from the application
- Multitasking
- Resource management
  - Processing, memory, storage, etc
- etc

Understanding Networked Applications 41 A First Course

Understanding Networked Applications 42 A First Course
File system

• Hides details of storage equipment from applications

• File is:
  – Unit of data managed for the benefit of the application
    • Size known, but unspecified structure and interpretation
  – Name
  – Location in naming hierarchy

Network equipment

- Switches
- Hosts
- Backbone links
- Access links
Messages and packets

- Simplest network communication service is the message
  - Smallest unit of communicated data meaningful to application
  - Size, but unknown structure and interpretation
  - Analogous to file in storage
- Internally, the network may fragment a message into packets, and reassemble those packets back into a message

Communication middleware

- New application-specific communication services
- Location independence
  - makes distributed application look similar to centralized
- Many possible other functions
Storage middleware

- **Database**
  - File with specified structure
  - Example: relational table
  - Oriented toward business applications

- **Database management system (DBMS)**
  - Manage multiple databases
  - Basis of online transaction processing (OLTP)

<table>
<thead>
<tr>
<th>Year</th>
<th>City</th>
<th>Accommodation</th>
<th>Tourists</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>Oakley</td>
<td>Bed &amp; Breakfast</td>
<td>14</td>
</tr>
<tr>
<td>2002</td>
<td>Oakley</td>
<td>Resort</td>
<td>190</td>
</tr>
<tr>
<td>2002</td>
<td>Oakland</td>
<td>Bed &amp; Breakfast</td>
<td>340</td>
</tr>
<tr>
<td>2002</td>
<td>Oakland</td>
<td>Resort</td>
<td>230</td>
</tr>
<tr>
<td>2002</td>
<td>Berkeley</td>
<td>Camping</td>
<td>120000</td>
</tr>
<tr>
<td>2002</td>
<td>Berkeley</td>
<td>Bed &amp; Breakfast</td>
<td>3450</td>
</tr>
<tr>
<td>2002</td>
<td>Berkeley</td>
<td>Resort</td>
<td>390800</td>
</tr>
<tr>
<td>2002</td>
<td>Albany</td>
<td>Camping</td>
<td>8790</td>
</tr>
<tr>
<td>2002</td>
<td>Albany</td>
<td>Bed &amp; Breakfast</td>
<td>3240</td>
</tr>
<tr>
<td>2003</td>
<td>Oakley</td>
<td>Bed &amp; Breakfast</td>
<td>55</td>
</tr>
<tr>
<td>2003</td>
<td>Oakley</td>
<td>Resort</td>
<td>320</td>
</tr>
<tr>
<td>2003</td>
<td>Oakland</td>
<td>Bed &amp; Breakfast</td>
<td>280</td>
</tr>
<tr>
<td>2003</td>
<td>Oakland</td>
<td>Resort</td>
<td>210</td>
</tr>
<tr>
<td>2003</td>
<td>Berkeley</td>
<td>Camping</td>
<td>115800</td>
</tr>
<tr>
<td>2003</td>
<td>Berkeley</td>
<td>Bed &amp; Breakfast</td>
<td>4560</td>
</tr>
<tr>
<td>2003</td>
<td>Berkeley</td>
<td>Resort</td>
<td>419000</td>
</tr>
<tr>
<td>2003</td>
<td>Albany</td>
<td>Camping</td>
<td>7650</td>
</tr>
<tr>
<td>2003</td>
<td>Albany</td>
<td>Bed &amp; Breakfast</td>
<td>6750</td>
</tr>
</tbody>
</table>
Some DBMS functions

- Logical structure separated from physical structure
- Platform independence
- Implement standard queries
- Access from multiple users/applications
- Manage data as asset separate from applications

The Internet

by

David G. Messerschmitt
What is the Internet

- **Internet** = the major global internet
- An *internet* is a “network of networks”
  - Interconnect standard for LAN’s, MAN’s, and WAN’s
- A private internet is called an **intranet**
- An **extranet** is an interconnection of intranets through the Internet

---

**Intranet**

- Private internet
- May be connected to Internet
  - **Firewall** creates a protected enclave
Extranet

• Intranets connected through an unprotected domain (typically the Internet)
• Encryption and other security technologies used to
  – protect proprietary information
  – prevent imposters, vandals, etc
Extranet

Consumers, field workers, etc.

Intranet

Internet

Lock icon indicates this is an extranet
Certificate is the server’s credential

Questions

- What business purposes do nomadic workers serve?
- Mobile?
- What advantage does direct Internet access have over long distance telephony?