

Data interchange

- Unix-style files
- serialization (marshalling): convert structured data into linear stream of bytes
- for files and across networks – used to be separate
- part of RPC:
 - Sun RPC
 - Corba
 - Java
- ASN.1
- XML

Data interchange

- Work across different platforms
 - byte order, floating point, character sets, ...
 - convert to destination platform
 - common intermediate representation
- Efficient

Structured files

- Older OS had records and punch card columns
- Unix model: lines separated into columns (tabs, spaces, commas)
- Also: csv in Excel and kin
- Sometimes # for comments

Structured files

- Examples:

- /etc/passwd

hgs:8D6uxb.jefyxz:5815:92:Henning G.
Schulzrinne:/home/hgs:/bin/tcsh

- files for sort

- /etc/services

time	37/tcp	timeserver	
time	37/udp	timeserver	
rlp	39/tcp	resource	# resource location
rlp	39/udp	resource	# resource location
nameserver	42/tcp	name	# IEN 116
nameserver	42/udp	name	# IEN 116

XML

- IBM: SGML (structured general markup language) → HTML (hypertext mark-up) → XML → XHTML
- idea: label content instead of presentation
- subset of SGML
- “documents”, but really structured (tree) data objects
- human readable, but not necessarily terse

XML

- entities = storage units containing
 - parsed data: characters + markup
 - unparsed data
- markup = start tags, end tags, entity references, ...
- starts with document type declaration:
`<?xml version="1.0"?>`
`<!DOCTYPE greeting SYSTEM "hello.dtd">`
- comments: `<!-- comment -->`
- verbatim: `<![CDATA[<greeting>Hello, world!</greeting>]]>`

XML

- Document type definition (DTD) defines structure of XML

```
<?xml version="1.0" encoding="UTF-8" ?>
  <!DOCTYPE greeting [
    <!ELEMENT greeting (#PCDATA)>
    <!ATTLIST poem xml:space (default|preserve)
      'preserve'>
  ]>
  <greeting>Hello, world!</greeting>
```

- Other mechanisms: XSD (later)

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/315042221000011034>