Part III.C: Internet protocols

- Overview
- HTTP/1.1
- HTTP-NG
- WebDAV
- Push
- Protocols and security
- XML-based protocols
What’s a protocol?

- Two or more computers connected on a network
- Rules for interaction:
  - what can a system send?
  - What can it expect to get back?
Design criteria for protocols

- performance (make it faster)
- bandwidth (amount of data sent in a particular time)
- reliability (entire system is stable even if some things go wrong)
- extensibility (can new features be added and still work with old implementations)
- security (doesn’t let others mess with you)
Internet protocols for different purposes

- Electronic mail *(SMTP, POP, IMAP)*
- Web *(HTTP)*
- Network news *(NNTP)*
- directory access *(LDAP)*
- interactive sessions *(TELNET)*
- ... and many many more...
HyperText Transfer Protocol (HTTP)

- A simple protocol, designed for the 1990 vision of the World Wide Web
  - open connection to widget.com
  - send “GET /product.html”
  - read headers
  - read body
  - close connection
HTTP/1.0 added features

- More kinds of content
  - Accept, language, charset, content-type
- More information about client & server
  - User-Agent, From, error codes
- Simple caching
  - last-modified, if-modified-since
- Basic Authorization
Proxy cache

- Between client and server
- Remembers what was retrieved before
- Don’t retrieve again unnecessarily
HTTP/1.1 Improvements

- **Performance**
  - pipelining: send several requests together
  - persistent connections:
    - don’t open and close connections all the time
  - better caching

- **Reliability**
  - clearer semantics for many headers

- **New features**
HTTP/1.1 Draft Standard

- Resolved over 100 “issues” with RFC 2068
  - problems discovered during implementation and deployment
  - each a serious design problem
- Additional features
- Improved security
Content Negotiation

- Different recipients have different capabilities
  - Cellphone
  - Reading machine
  - Print vs. display
- How to tune content for recipient?
- How to describe recipients
HTTP Content Negotiation

- Language (Accept-Language)
- Character set (Accept-Charset)
- Capabilities to handle media (Accept)
- Brand of software (User-Agent)

need more: active working group
HTTP is \emph{not} a good protocol

- HTTP/1.0 didn’t work well as web evolved
- HTTP/1.1 fixed some problems
  - backward compatibility was more important
- It still has lots of problems!
  - Don’t copy it for new protocols
  - Session Initiation Protocol, Real Time Streaming Protocol do
  - See RFC 2324: HTCP CP
HTTP-NG: “Next Generation”

- New design
- Not required to interwork with HTTP/1.1

Design goals:
- simple
- performance
- asynchronous operation

- uses distributed object technology
  - Compatibility with CORBA, RMI, DCOM
WebDAV:
Distributed Authoring and Versioning

- Locking
- Compound objects
- Version management
- Directory management

*WebDA* nearly finished, versioning, search language in progress
XML-based application protocols

- Define data exchange in terms of XML
- controlled extension with core interoperability
- E-commerce applications
  - Internet Content & Exchange (ICE)
  - Open Buying on the Internet (OBI)
  - Trading Protocol (OTP)
Internet Security

- Everyone connected
  - including people you should not trust
  - security is easy: disconnect the net
  - “firewall”: selective disconnection
- “Security” provides assurance against threats
  - must analyze the threats!
- Encryption: scrambling the data
  - sometimes only “locking the front door”
Internet security: threat analysis

- Someone will access my private files...
- Someone will modify my files...
- Someone will discover my password and later pretend to be me...
- Someone will watch what I am reading...
- Someone will pretend to be someone that I trust...

Hard to predict, to protect against threats
Security measures

- Authentication: prove you are you
  - passwords
  - scrambled passwords (digest authentication)
- Protected channel: keep connection secret
  - “SSL” (Secure Socket Layer)
  - TLS (Transaction Level Security)
  - VPN (Virtual Private Network)
Protected objects

- use “insecure” network to transmit encrypted objects
- S/MIME (Secure MIME)
- PGP (Pretty Good Privacy)
- S-HTTP (Secure HTTP)

Standards are difficult because of patents
Large amount of ongoing work

- Internet Payment
- Content Rating (PICS)
- “Push”: broadcasting
- Messaging
- Real-time multimedia
Tutorial Review

- Part I: Internet and Digital Libraries
- Part II: Standards and Organizations
- Part III: Technology for
  - Content
  - Naming
  - Protocols
Key points

- The Internet is evolving
- Many different elements are used to weave each application
- Digital library applications will use these … and more