The World Wide Web

An Introduction

Mark Handley
Department of Computer Science
University College London

Outline

- A little about the internet
  - what it is
  - how it functions
- A little bit of history
  - ftp
  - gopher
  - other information services
- An Introduction to the Web
  - hypertext and hyperlinks
  - why the history was relevant
- More details about the Web

The Internet

- **What the Internet isn't:**
  - **it's not dial up**
  - **it's not a bullitin board**
  - **it's not run by one organisation**
  - **it's not one thing at all**
- **What it is:**
  - **a network of networks**
  - **all the bits of it are run locally**
  - **most computers on it are *permanently* connected, although you can access it by dial up**
  - **the one thing in common in the *Internet Protocol***

Internet Protocol (IP)

- take the information you want to send
- put it in a packet
- put the address of the computer you want to send it to on the packet
- let it loose
- the network tries to get your packet to it's destination

IP and TCP

- the internet protocol is a standard way of putting information into packets and putting addresses on those packets
- the internet (IP) does not guarantee to get your packets there!
- TCP (Transmission Control Protocol) attempts to make the internet reliable
- the end systems communicate with each other using TCP in such a way that lost packets get replaced and all your data gets there eventually and the network gets shared out between all the people who want to use it.

IP addresses

- these are the addresses that get added to a packet to get it somewhere
- an IP address is written as 128.16.64.5 or 193.64.58.2 or whatever
- each router on the way looks at the address and decides where to send the packet next
- machines use addresses
- people should normally use names
- machines convert names into addresses using DNS (Domain Name Service)

Host Names

- a hostname looks something like
  - speedy.cs.ucl.ac.uk
  - www.sics.se
  - ftp.ee.lbl.gov
  - www.isi.edu
- the first bit is the local host name
- the rest is the domain (or site) name
- together they form the full (globally unique) host name

in the beginning...

...there was **ftp**...

- File Transfer Protocol
- to transfer files

...and **telnet**...

- for logging in remotely

...and **SMPT**...

- Simple Mail Transfer Protocol
- for electronic mail exchange

...and **NNPT**...

- Net News Transfer Protocol
- for discussion and arguements

...all built on top of TCP

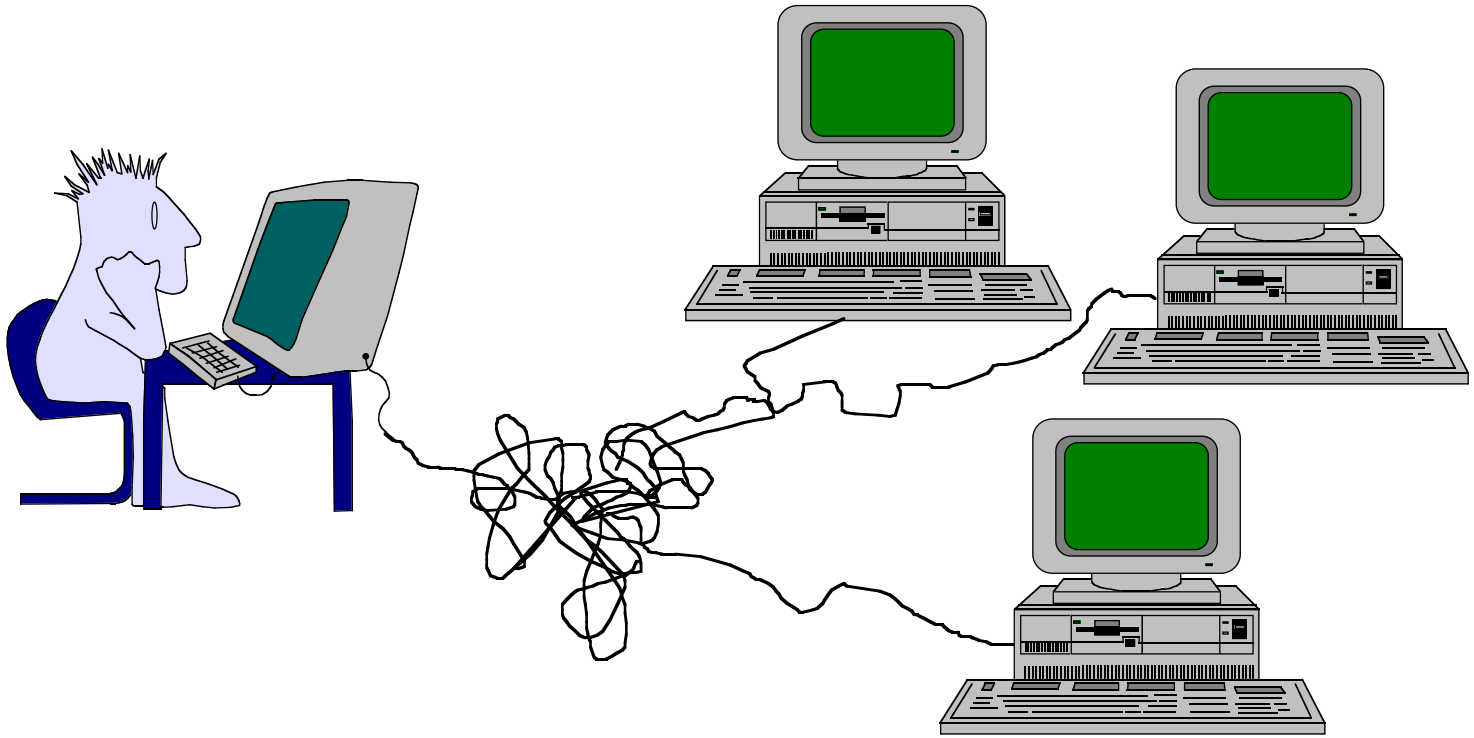
in the beginning...

...no-one had heard of the internet

...too much like Rocket Science

A note on clients and servers

A user uses a client program...



...to access information on an information server

First there was ftp...

- ftp is the DOS of Information Retrieval systems
- you connect to a remote machine
- you then log on (possibly anonymously)
- then you issue commands to move around the file system, list file directories, “get” files, etc
- you have to know where you’re going, what the file is called, which directory it’s in, plus a load of commands

...but it works well

Then came gopher...

- **Gopher is to Information Retrieval what Windows is to DOS**
  - **You can't actually do anything more with gopher that you can with ftp**
  - **But it looks nicer...**

    ...and you don't need to remember the commands

Gopher is a menu driven information retrieval system

Menus help you find the data

Menus don't help you find the server!

..and Archie

- Archie is a set of indexes to ftp servers
- It is searchable, to let you find something if you know what it's called but not where it is stored.

and still *almost* no-one had heard of the internet...

...and now you need to understand lots of different programs!

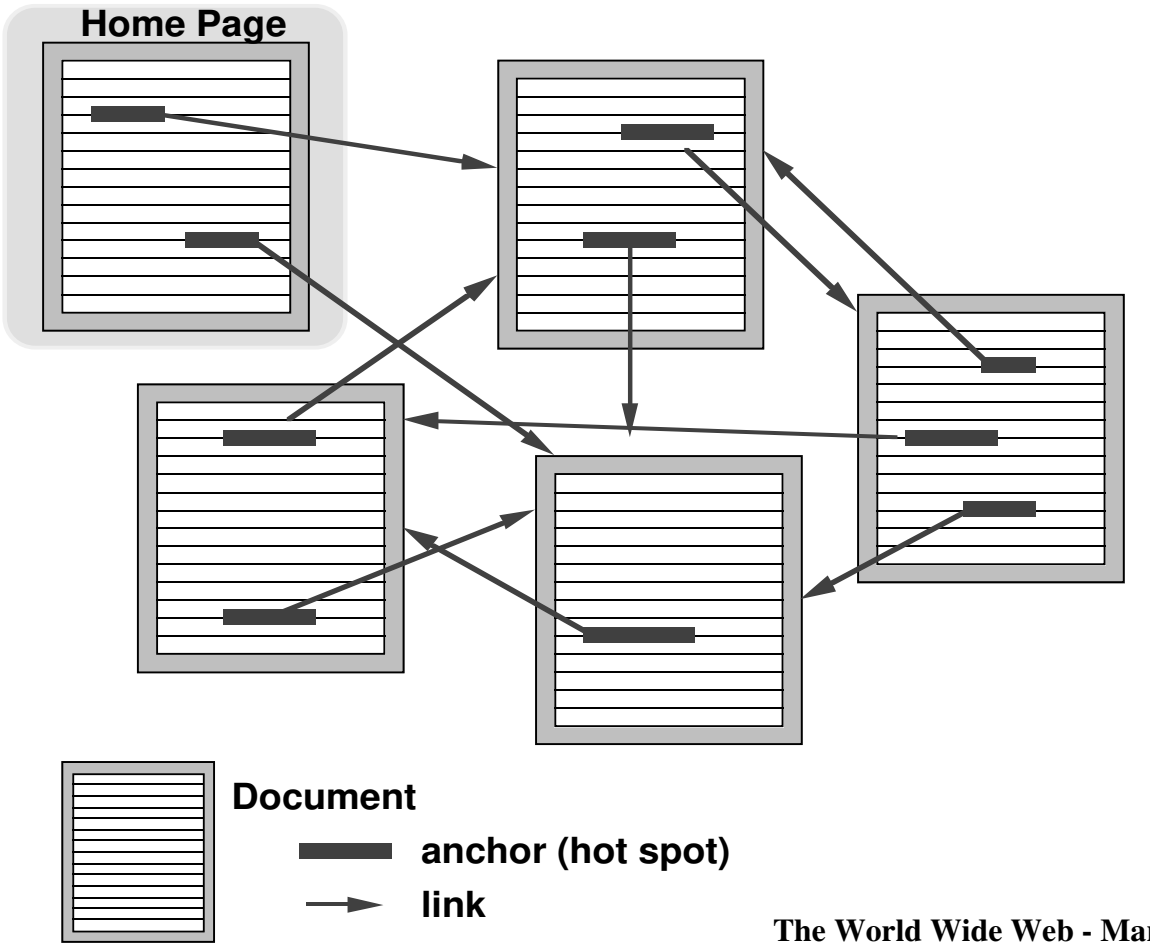
...but what are those people at CERN doing?

The World Wide Web

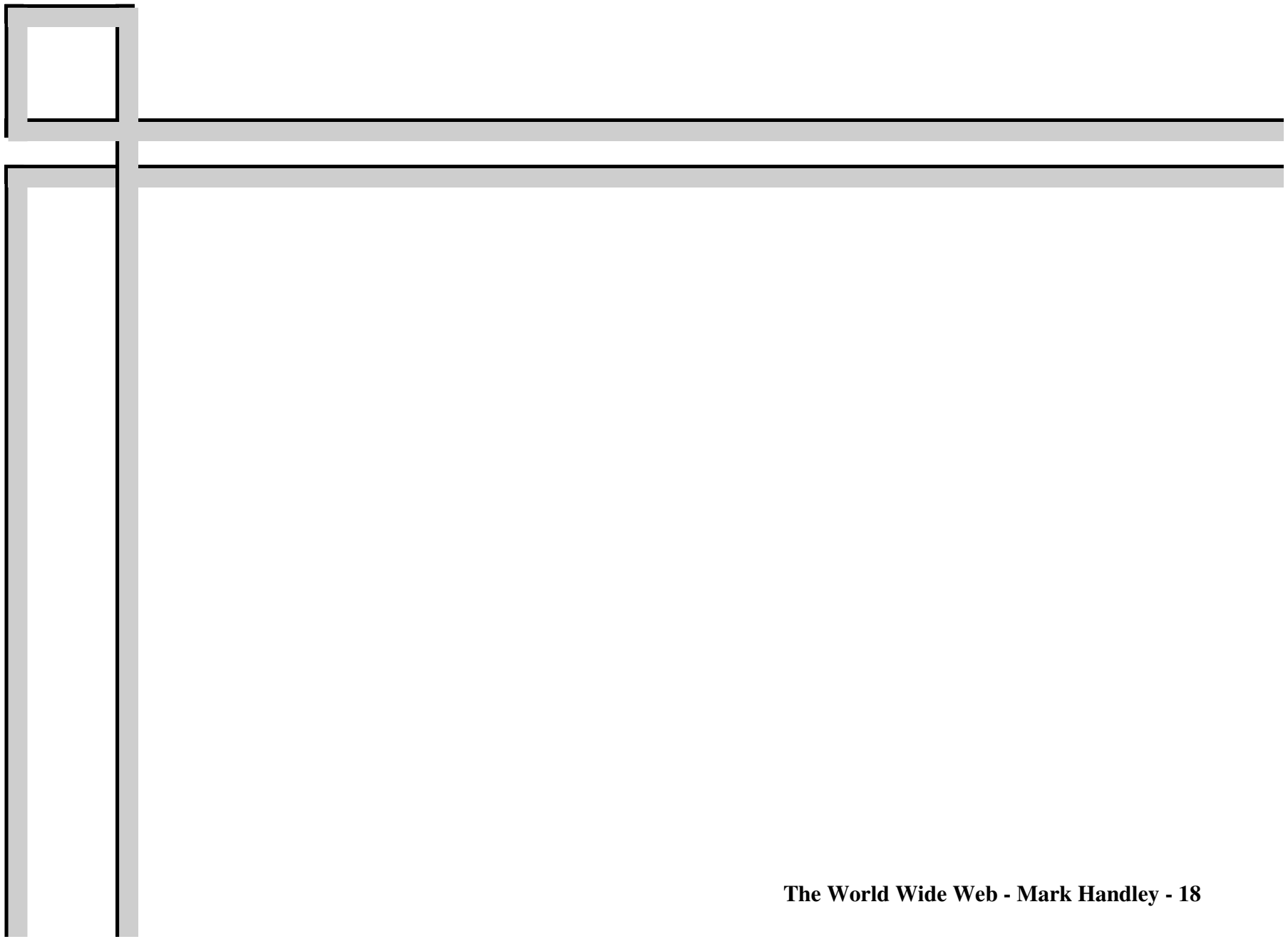
- **A different way of organising information**
- **Pages of text with pictures**
- **Links to other pages of text**
- **Information Content is Important**
- **Information Location is unimportant**

- ***Hypertext*** **and** ***Hypermedia***

Basic Hypertext



Multiprotocol Clients

- There's lots of existing information out there on gopher and ftp and wais and other servers
- It just needs to be easier to find
- So why not point a link from a WWW document to information on an FTP server?

- Needs a *multiprotocol client*
- Mosaic and other WWW clients are multiprotocol clients