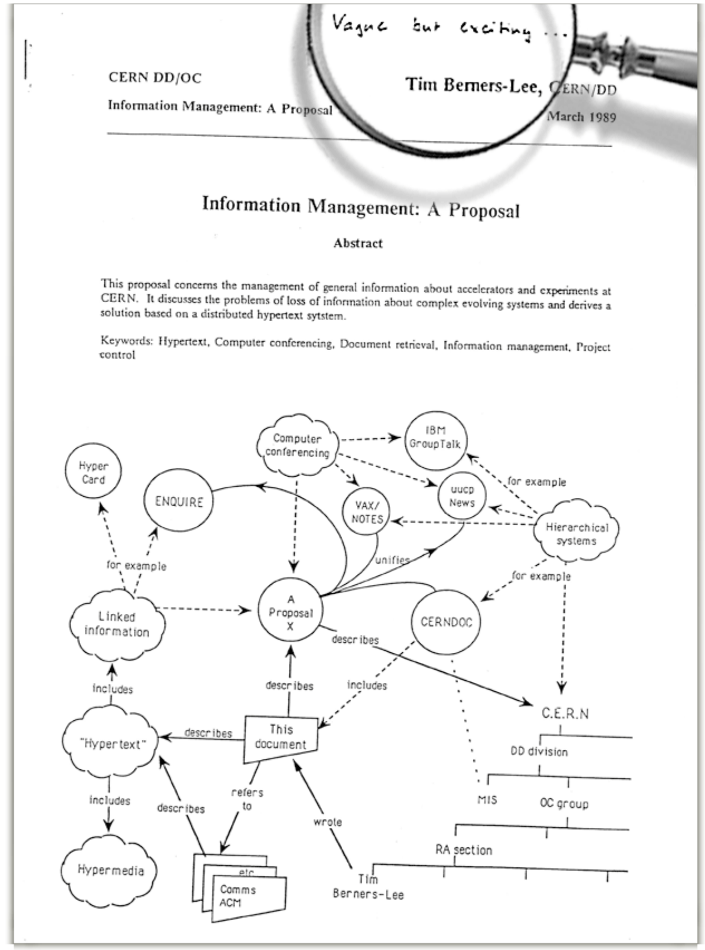
Computer Networks

6.033 Lecture 9

Hari Balakrishnan
hari@mit.edu

Applications on the Internet


Vague but exciting ...

CERN DD/OC

Information Management: A Proposal

Tim Berners-Lee, CERN/DD

March 1989

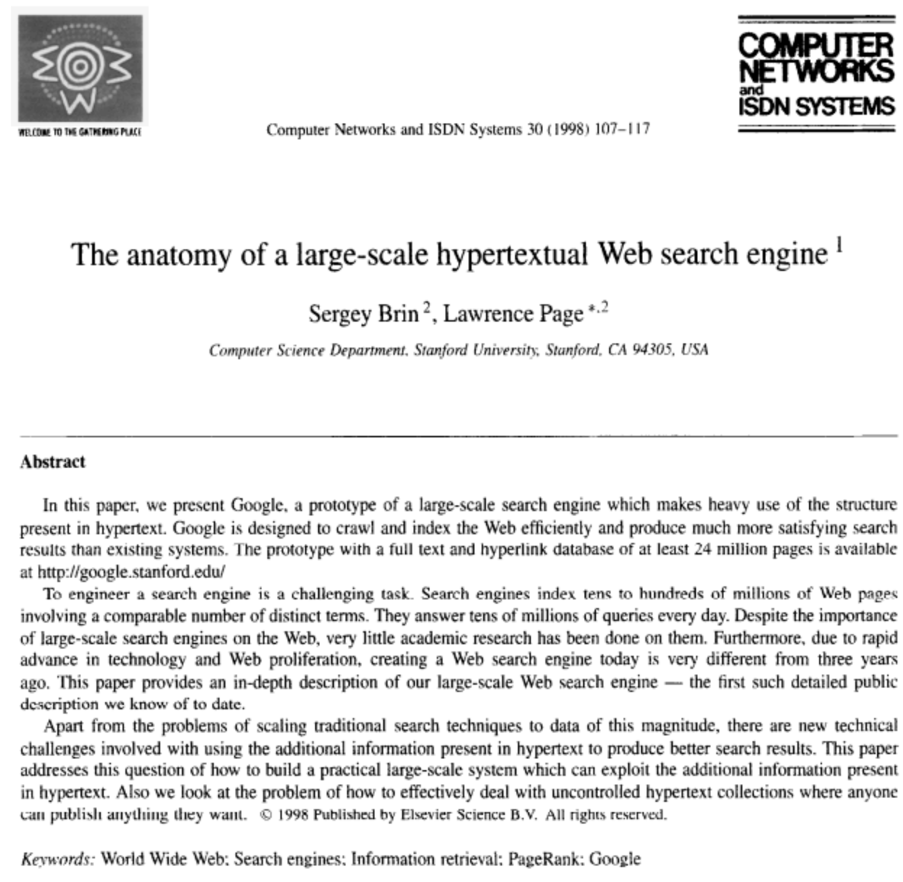
Information Management: A Proposal

Abstract

This proposal concerns the management of general information about accelerators and experiments at CERN. It discusses the problems of loss of information about complex evolving systems and derives a solution based on a distributed hypertext sytstem.

Keywords: Hypertext, Computer conferencing, Document retrieval, Information management, Project control

Applications on the Internet


Computer Networks and ISDN Systems 30 (1998) 107–117

COMPUTER NETWORKS and ISDN SYSTEMS

The anatomy of a large-scale hypertextual Web search engine [1]

Sergey Brin [2], Lawrence Page [*,2]

Computer Science Department, Stanford University, Stanford, CA 94305, USA

Abstract

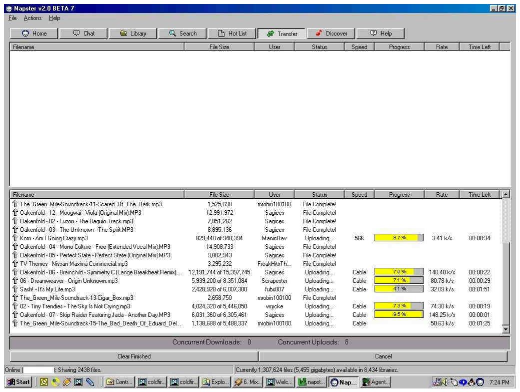
In this paper, we present Google, a prototype of a large-scale search engine which makes heavy use of the structure present in hypertext. Google is designed to crawl and index the Web efficiently and produce much more satisfying search results than existing systems. The prototype with a full text and hyperlink database of at least 24 million pages is available at http://google.stanford.edu/

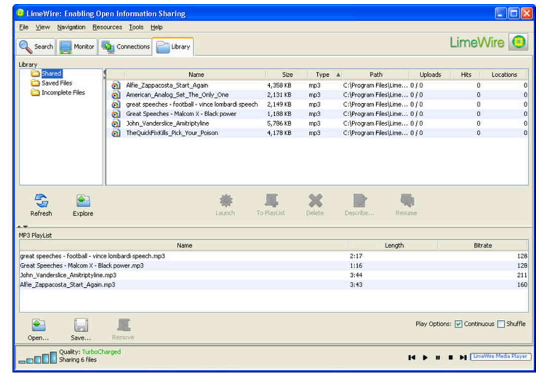
To engineer a search engine is a challenging task. Search engines index tens to hundreds of millions of Web pages involving a comparable number of distinct terms. They answer tens of millions of queries every day. Despite the importance of large-scale search engines on the Web, very little academic research has been done on them. Furthermore, due to rapid advance in technology and Web proliferation, creating a Web search engine today is very different from three years ago. This paper provides an in-depth description of our large-scale Web search engine — the first such detailed public description we know of to date.

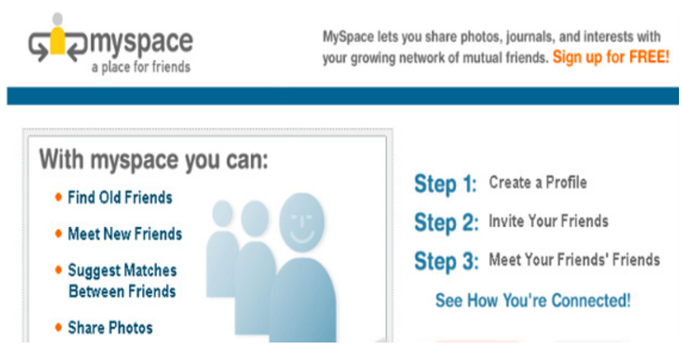
Apart from the problems of scaling traditional search techniques to data of this magnitude, there are new technical challenges involved with using the additional information present in hypertext to produce better search results. This paper addresses this question of how to build a practical large-scale system which can exploit the additional information present in hypertext. Also we look at the problem of how to effectively deal with uncontrolled hypertext collections where anyone can publish anything they want.

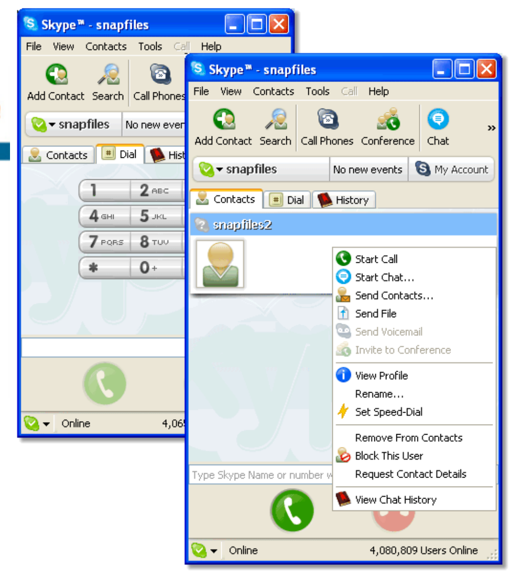
Keywords: World Wide Web; Search engines; Information retrieval; PageRank; Google

Applications on the Internet

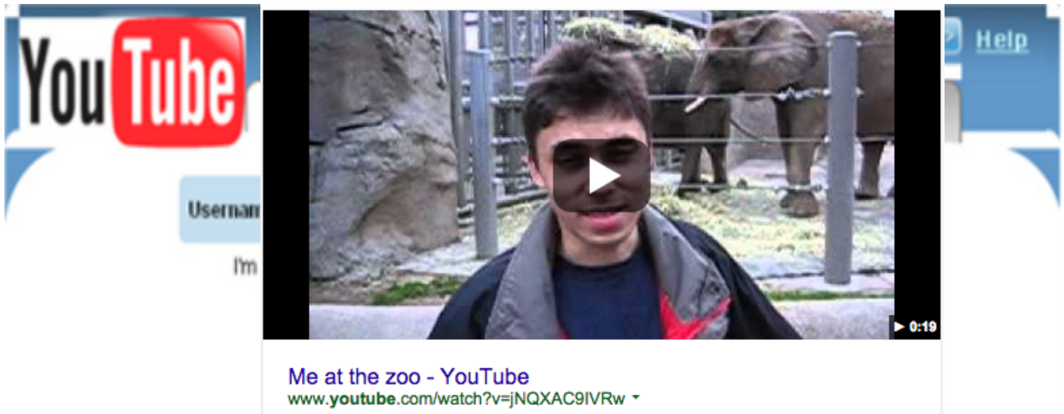


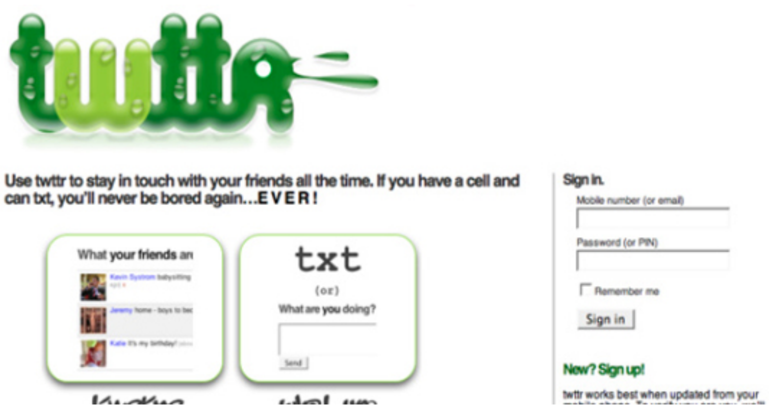






Applications on the Internet





Greatest Engineering Achievements OF THE 20TH CENTURY

National Academy of Engineering (http://www.greatachievements.org/)

Welcome!

How many of the 20th century's greatest engineering achievements will you use today? A car? Computer? Telephone? Explore our list of the top 20 achievements and learn how engineering shaped a century and changed the world.

1. Electrification
2. Automobile
3. Airplane
4. Water Supply and Distribution
5. Electronics
6. Radio and Television
7. Agricultural Mechanization
8. Computers
9. Telephone
10. Air Conditioning and Refrigeration
11. Highways
12. Spacecraft
13. Internet
14. Imaging
15. Household Appliances
16. Health Technologies
17. Petroleum and Petrochemical Technologies
18. Laser and Fiber Optics
19. Nuclear Technologies
20. High-performance Materials

The Internet? Bah!

Hype alert: Why cyberspace isn't, and will never be, nirvana

By Clifford Stoll | NEWSWEEK
From the magazine issue dated Feb 27, 1995

Why the Web Won't Be Nirvana

BY CLIFFORD STOLL 2/26/95 AT 7:00 PM

Why the Web Won't Be Nirvana

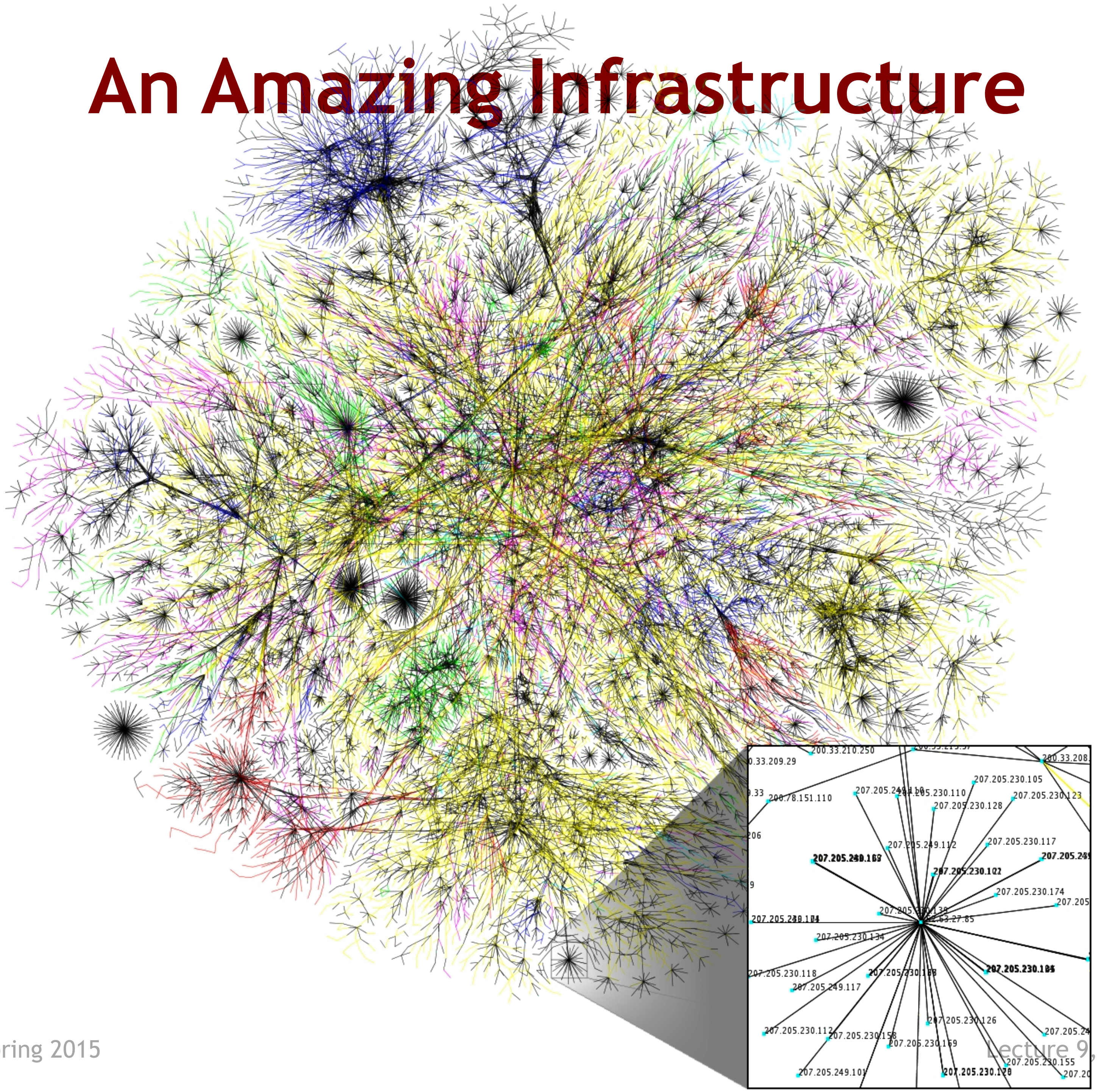
BY **CLIFFORD STOLL** 2/26/95 AT 7:00 PM

After two decades online, I'm perplexed. It's not that I haven't had a gas of a good time on the Internet. I've met great people and even caught a hacker or two. But today, I'm uneasy about this most trendy and oversold community. Visionaries see a future of telecommuting workers, interactive libraries and multimedia classrooms. They speak of electronic town meetings and virtual communities. Commerce and business will shift from offices and malls to

Baloney. Do our computer pundits lack all common sense? The truth in no online database will replace your daily newspaper, no CD-ROM can take the place of a competent teacher and no computer network will change the way government works.

Then there's cyberbusiness. We're promised instant catalog shopping—just point and click for great deals. We'll order airline tickets over the network, make restaurant reservations and negotiate sales contracts. Stores will become obselete. So how come my local mall does more business in an afternoon than the entire Internet handles in a month? Even if there were a trustworthy way to send money over the Internet—which there isn't—the network is missing a most essential ingredient of capitalism: salespeople.

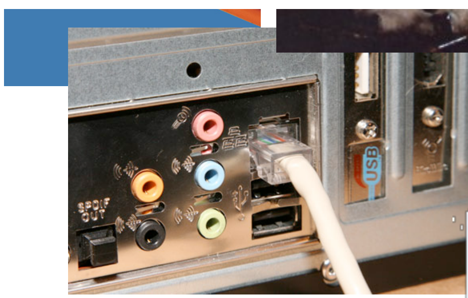
An Amazing Infrastructure

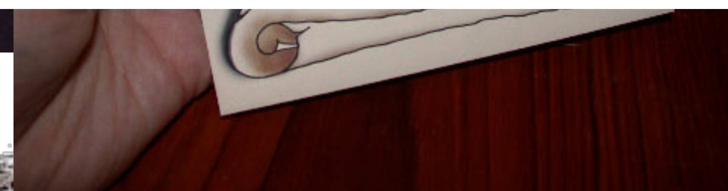


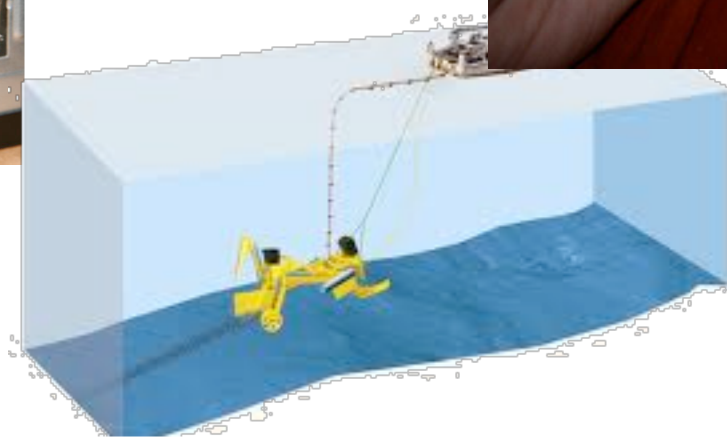
Physical Communication Channels

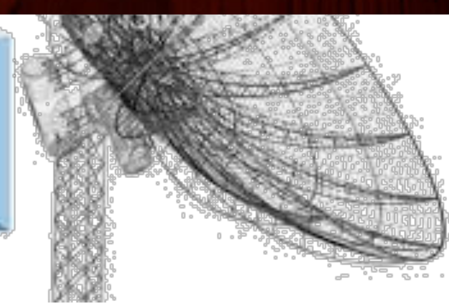
Network Working Group   D. Waitzman
Request for Comments: 1149   BBN STC
                                                                             1 April 1990

A Standard for the Transmission of IP Datagrams on Avian Carriers

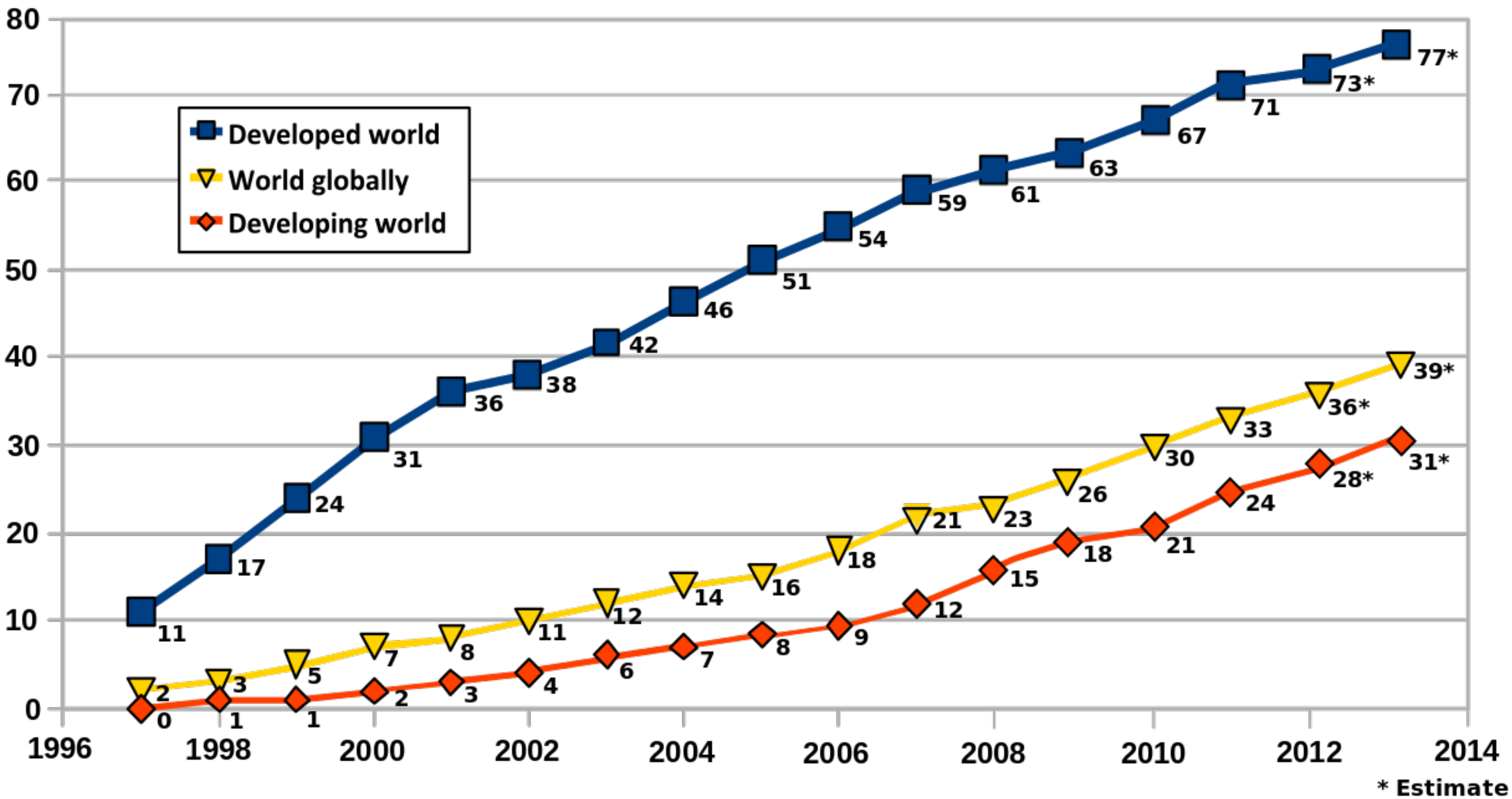








Internet users per 100 inhabitants
Source: International Telecommunications Union



Universal Network Layer: IP

IP packet format (v4 and v6)

32 bits

Version	IHL	Type-of-Service	Total Length
Identification		Flags	Fragmentation offset
Time-to-live	Protocol	Header Checksum	
Source Address			
Destination Address			
Options (+ padding)			
Data (variable)			

Version	Traffic class	Flow label
Payload length	Next header	Hop limit
Source address		
Destination address		

Layering

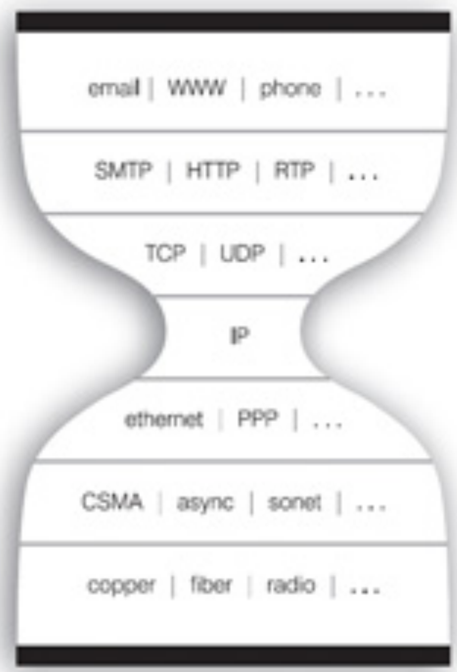


Figure 4.1 Hourglass architecture of the Internet

Evolution of design

A Protocol for Packet Network Intercommunication

VINTON G. CERF AND ROBERT E. KAHN,
MEMBER, IEEE

Abstract — A protocol that supports the sharing of resources that exist in different packet switching networks is presented. The protocol provides for variation in individual network packet sizes, transmission failures, sequencing, flow control, end-to-end error checking, and the creation and destruction of logical process-to-process connections. Some implementation issues are considered, and problems such as internetwork routing, accounting, and timeouts are exposed.

INTRODUCTION

IN THE LAST few years considerable effort has been expended on the design and implementation of packet switching networks [1]-[7],[14],[17]. A principle reason for developing such networks has been to facilitate the sharing of computer resources. A packet communication network includes a transportation mechanism for delivering data between computers or between computers and terminals. To make the data meaningful, computer and terminals share a common protocol (i.e, a set of agreed upon conventions). Several protocols have already been developed for this purpose [8]-[12],[16]. However, these protocols have addressed only the problem of communication on the same network. In this paper we present a protocol design and philosophy that supports the sharing of resources that exist in different packet switching networks.

After a brief introduction to internetwork protocol issues, we describe the function of a GATEWAY as an interface between networks and discuss its role in the protocol. We then consider the various details of the protocol, including addressing, formatting, buffering, sequencing, flow control, error control, and so forth. We close with a description of an interprocess communication mechanism and show how it can be supported by the internetwork protocol.

Even though many different and complex problems must be solved in the design of an individual packet switching network, these problems are manifestly compounded when dissimilar networks are interconnected. Issues arise which may have no direct counterpart in an of one or more *packet switches*, and a collection of communication media that interconnect the packet switches. Within each HOST, we assume that there exist *processes* which must communicate with processes in their own or other HOSTS. Any current definition of a process will be adequate for our purposes [13]. These processes are generally the ultimate source and destination of data in the network. Typically, within an individual network, there exists a protocol for communication between any source and destination process. Only the source and destination processes require knowledge of this convention for communication to take place. Processes in two distinct networks would ordinarily use different protocols for this purpose. The ensemble of packet switches and communication media is called the *packet switching subnet*. Fig. 1 illustrates these ideas.

In a typical packet switching subnet, data of a fixed maximum size are accepted from a source HOST, together with a formatted destination address which is used to route the data in a store and forward fashion. The transmit time for this data is usually dependent upon internal network parameters such as communication media data rates, buffering and signalling strategies, routeing, propagation delays, etc. In addition, some mechanism is generally present for error handling and determination of status of the networks components.

Individual packet switching networks may differ in their implementations as follows.

1) Each network may have distinct ways of addressing the receiver, thus requiring that a uniform addressing scheme be created which can be understood by each individual network.

2) Each network may accept data of different maximum size, thus requiring networks to deal in units of the smallest maximum size (which may be impractically small) or requiring procedures which allow data crossing a network boundary to be

Fig. 4. TCP address.

"The choice for network identification (8 bits) allows up to 256 distinct networks. This size seems sufficient for the foreseeable future."

End-To-End Arguments in System Design

J. H. SALTZER, D. P. REED, and D. D. CLARK
Massachusetts Institute of Technology Laboratory for Computer Science

This paper presents a design principle that helps guide placement of functions among the modules of a distributed computer system. The principle, called the end-to-end argument, suggests that functions placed at low levels of a system may be redundant or of little value when compared with the cost of providing them at that low level. Examples discussed in the paper include bit-error recovery, security using encryption, duplicate message suppression, recovery from system crashes, and delivery acknowledgment. Low-level mechanisms to support these functions are justified only as performance enhancements.

CR Categories and Subject Descriptors: C.0 **[General]** Computer System Organization—*system architectures*; C.2.2 **[Computer-Communication Networks]**: Network Protocols—*protocol architecture*; C.2.4 **[Computer-Communication Networks]**: Distributed Systems; D.4.7 **[Operating Systems]**: Organization and Design—*distributed systems*

General Terms: Design

Additional Key Words and Phrases: Data communication, protocol design, design principles

1. INTRODUCTION

Choosing the proper boundaries between functions is perhaps the primary activity of the computer system designer. Design principles that provide guidance in this choice of function placement are among the most important tools of a system designer. This paper discusses one class of function placement argument that has been used for many years with neither explicit recognition nor much conviction. However, the emergence of the data communication network as a computer system component has sharpened this line of function placement argument by making more apparent the situations in which and the reasons why it applies. This paper articulates the argument explicitly, so as to examine its nature and to see how general it really is. The argument appeals to application requirements and provides a rationale for moving a function upward in a layered system closer to the application that uses the function. We begin by considering the communication network version of the argument.

This is a revised version of a paper adapted from End-to-End Arguments in System Design by J. H. Saltzer, D.P. Reed, and D.D. Clark from the 2nd International Conference on Distributed Systems (Paris, France, April 8–10) 1981, pp. 509–512. © IEEE 1981

This research was supported in part by the Advanced Research Projects Agency of the U.S. Department of Defense and monitored by the Office of Naval Research under contract N00014-75-C-0661.

Authors' address: J. H. Saltzer and D. D. Clark, M.I.T. Laboratory for Computer Science, 545 Technology Square, Cambridge, MA 02139. D. P. Reed, Software Arts, Inc., 27 Mica Lane, Wellesley, MA 02181.

ACM Transactions on Computer Systems, Vol. 2, No. 4, November 1984, Pages 277–288.

Growth of the BGP Table - 1994 to Present