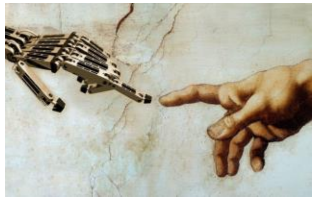
互联网技术演进：回顾和思考

李星

2021-09-20

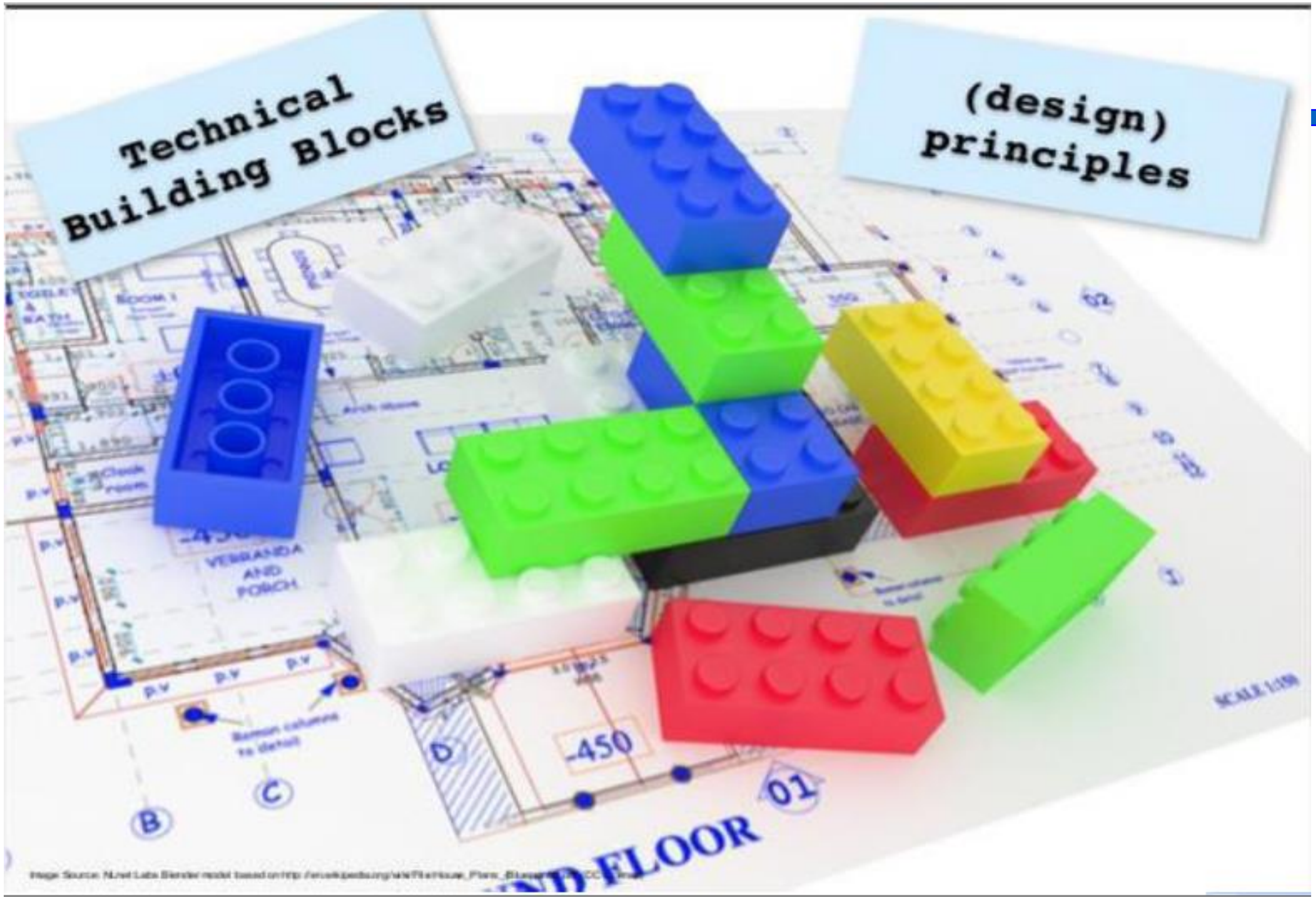
Engineering vs. science



Words

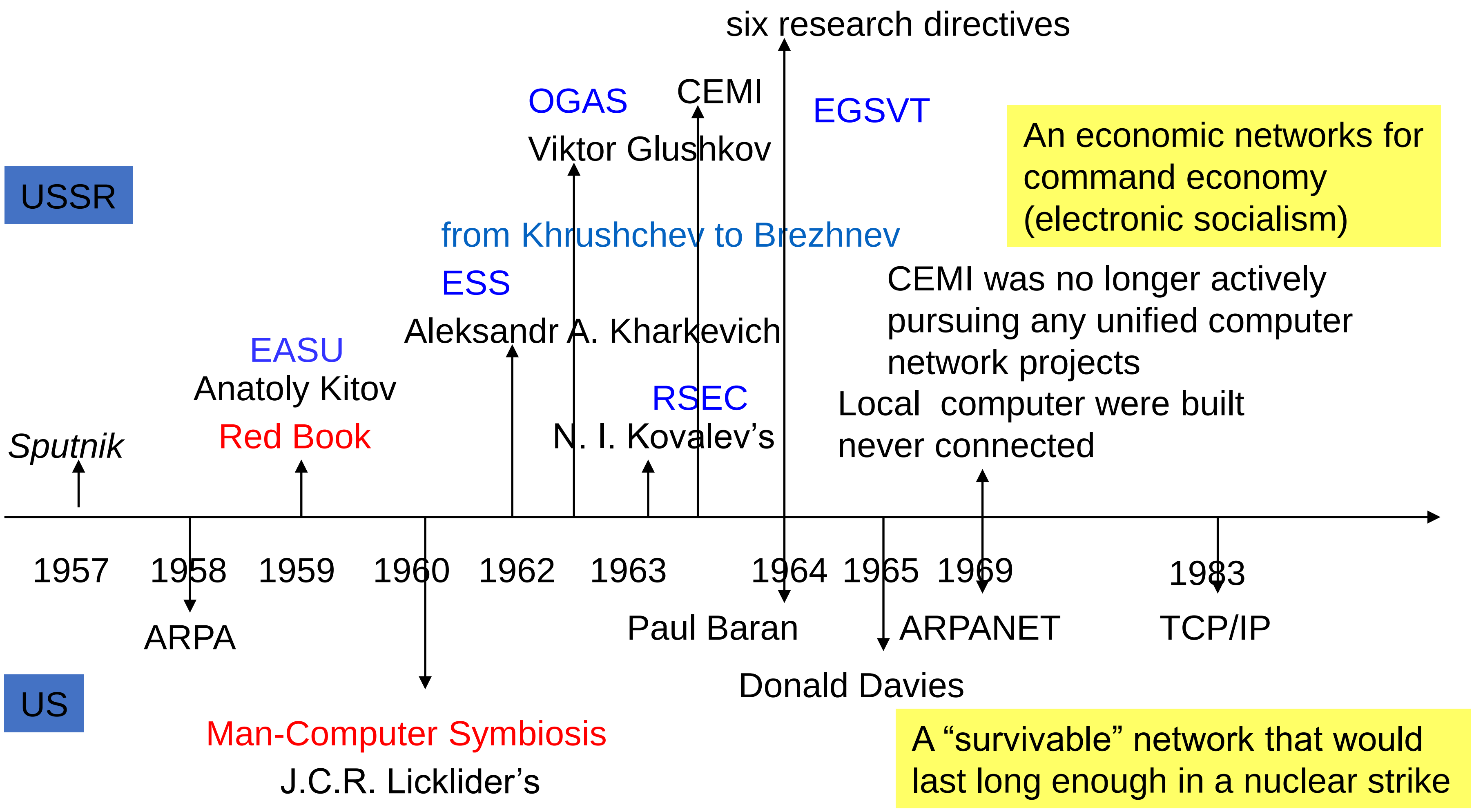
- History repeats itself.
- The future is a door, the past is the key. The farther backward you can look, the farther forward you can see.
- If we could do it again?


GROUND FLOOR 01
Internet Society
5 About the IETF | March 2016


Technical
Building Blocks
(design)
principles

案例1： OGAS vs APANET

USSR vs US


six research directives
OGAS
CEMI
EGSVT
Viktor Glushkov
An economic networks for command economy (electronic socialism)
USSR
from Khrushchev to Brezhnev
ESS
CEMI was no longer actively pursuing any unified computer network projects
Aleksandr A. Kharkevich
EASU
Anatoly Kitov
RSEC
Local computer were built never connected
Red Book
N. I. Kovalev's
Sputnik
1957
1958
1959
1960
1962
1963
1964
1965
1969
1983
ARPA
Paul Baran
ARPANET
TCP/IP
Donald Davies
US
A "survivable" network that would last long enough in a nuclear strike
Man-Computer Symbiosis
J.C.R. Licklider's

Paul Baran’s paper, ARPANET and RFC1

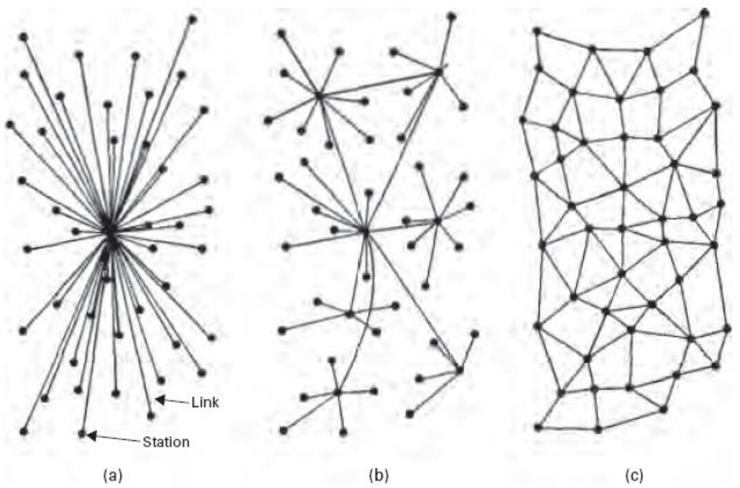
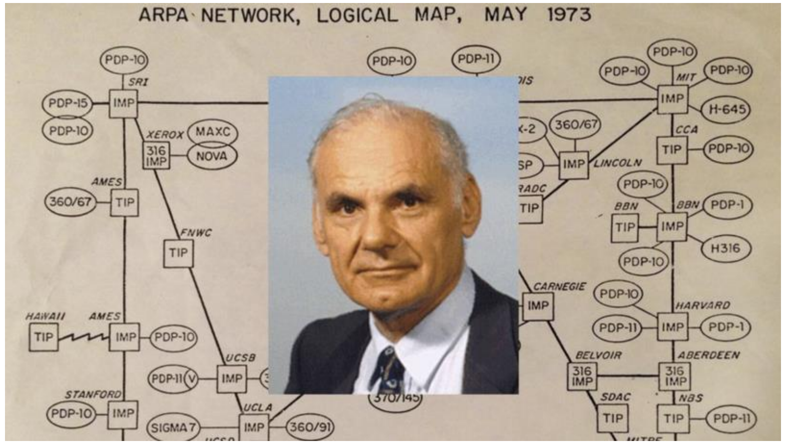


Figure 3.2
Three network types: (a) Centralized, (b) decentralized, and (c) distributed. *Source:* From Paul Baran, “Introduction to Distributed Communication Networks.” *On Distributed Communications,* RAND Corporation Memorandum RM-3420-PR, August 1964, 2. Reproduced with permission of The Rand Corp.




```
Network Working Group                                    Steve Crocker
Request for Comments: 1                                           UCLA
                                                          7 April 1969

                    Title:   Host Software
                   Author:   Steve Crocker
             Installation:   UCLA
                     Date:   7 April 1969
     Network Working Group Request for Comment:   1
```

OGAS

- Six research directives
  1. Develop a theory of optimal planning and management for a unified mathematical model of national economy;
  2. Develop a unified system of economic information;
  3. Standardize and algorithmize the planning and management processes;
  4. Develop mathematical methods for solving economic problems;
  5. Design and create a unified state network of computer centers;
  6. Derive a specialized planning and management system based on mathematical methods and computer technology.

- Resistance from at least five groups
  1. The military wanted nothing to do with civilian affairs, especially when that meant fixing the command economy that already fed its coffers;
  2. The economic ministries (particularly the Central Statistical Administration and the Ministry of Finance) wanted the OGAS Project under their control and fought to the point of mutiny to keep competing ministries from controlling it;
  3. The bureaucrats administering the plan feared that the network would put them out of a job;
  4. Factory managers and factory workers worried that the network would pull them out of informal gray economy; and
  5. Liberal economists fretted that the network would prevent the market reforms that they sought to introduce.

Viktor Glushkov speaking about management information systems. Image: ResearchGate

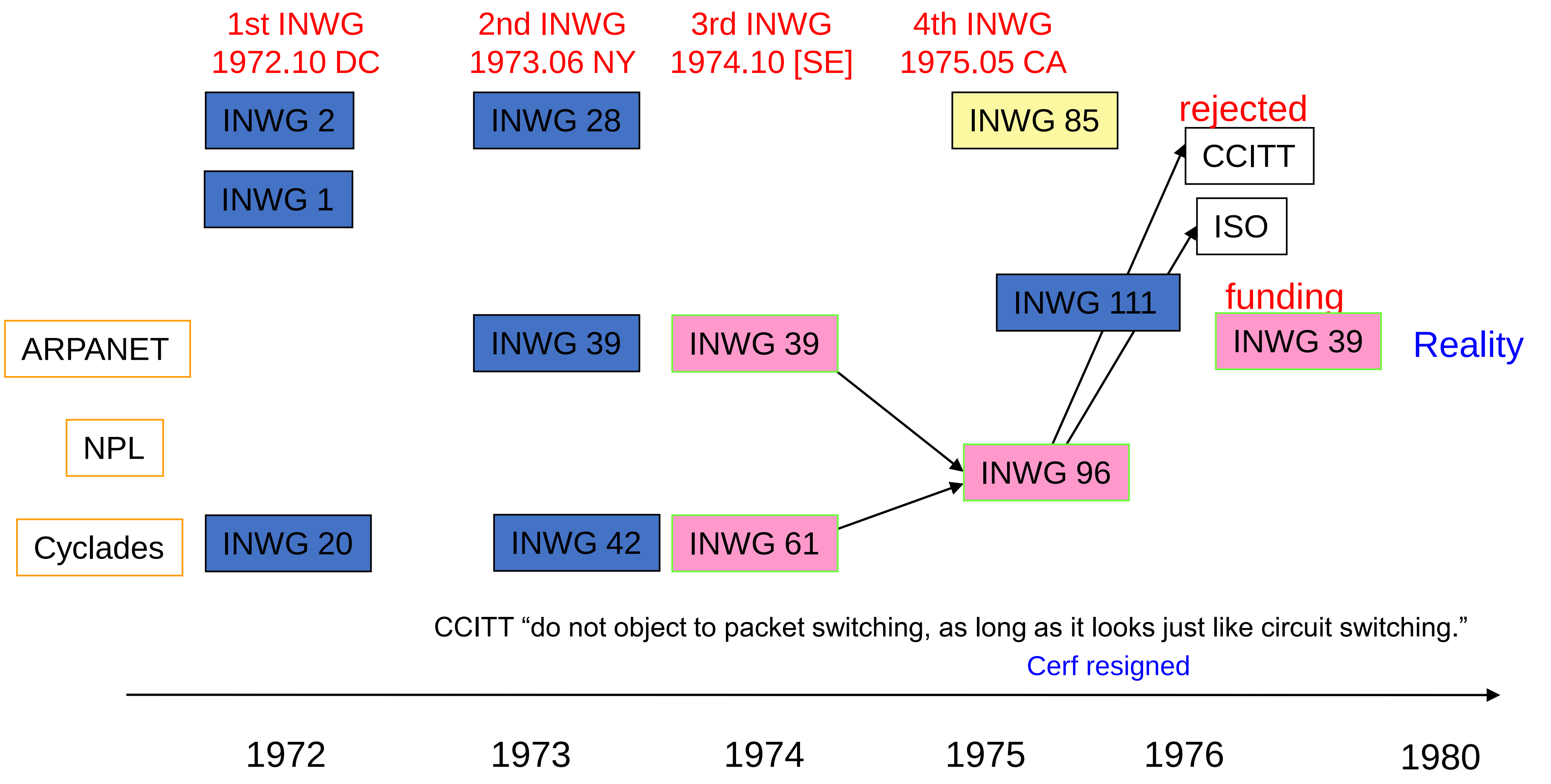
Viktor Glushkov

互联网和苏联计算机网比较

	互联网（Internet）	苏联计算机网（OGAS）
网络性质	开放网络	管制网络
网络拓扑	扁平结构	层次结构
网络应用	平等协作	指令和控制
网络类型	分组交换	电路交换

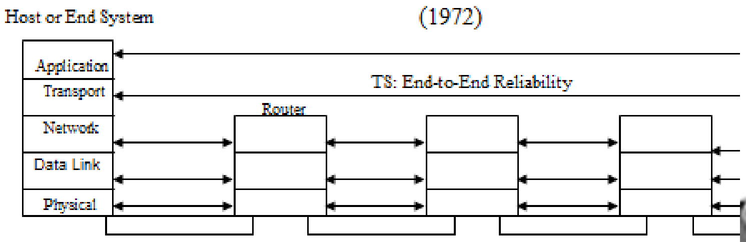
案例2： NCP vs. Cyclades

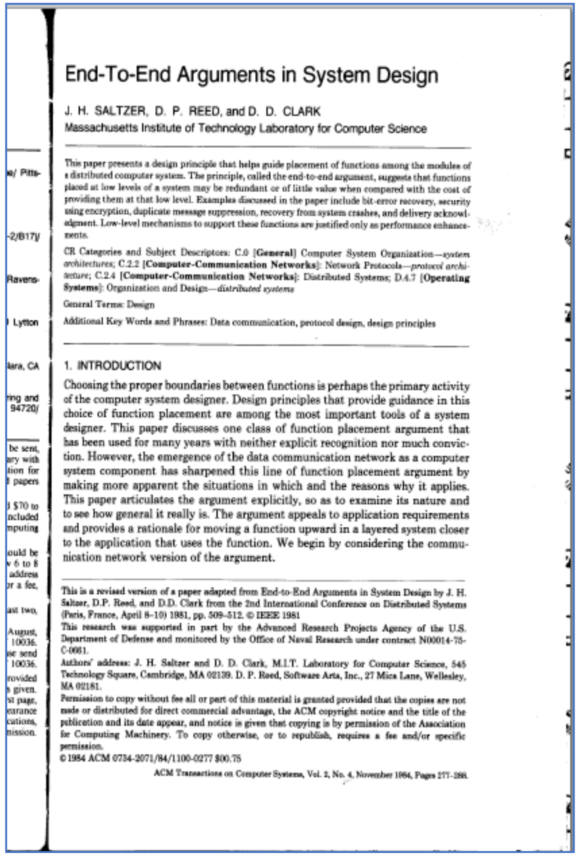
INWG



1st INWG
1972.10 DC
2nd INWG
1973.06 NY
3rd INWG
1974.10 [SE]
4th INWG
1975.05 CA
INWG 2
INWG 28
INWG 85
rejected
CCITT
INWG 1
ISO
INWG 111
funding
ARPANET
INWG 39
INWG 39
INWG 39
Reality
NPL
INWG 96
Cyclades
INWG 20
INWG 42
INWG 61
CCITT “do not object to packet switching, as long as it looks just like circuit switching.”
Cerf resigned
1972
1973
1974
1975
1976
1980

The Cyclades Architecture and David Clark (1988)





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

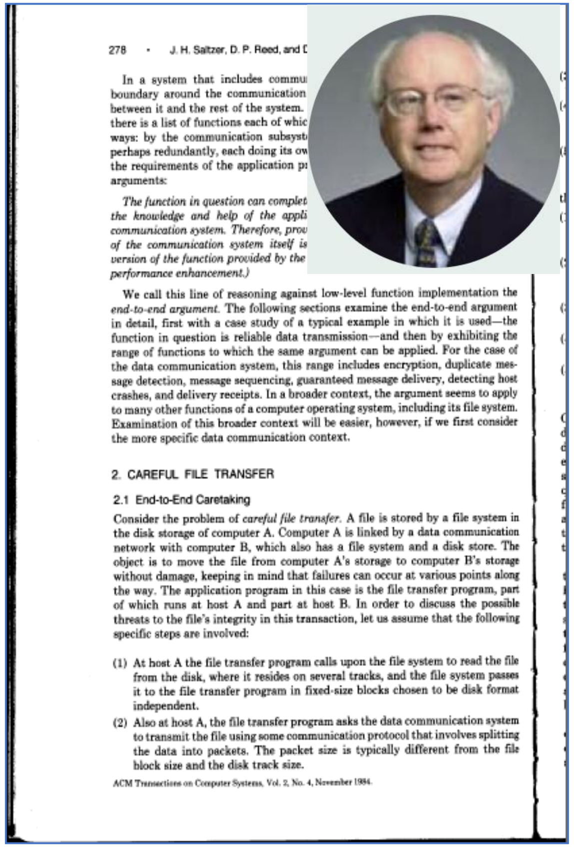
This research was supported in part by the Advanced Research Projects Agency of the U.S. Department of Defense and monitored by the Office of Naval Research under contract N00014-75-C-0661.

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.

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission.

© 1984 ACM 0734-2071/84/1100-0277 $00.75

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



278 • J. H. Saltzer, D. P. Reed, and D

In a system that includes commu
boundary around the communication
between it and the rest of the system.
there is a list of functions each of whic
ways: by the communication subsyst
perhaps redundantly, each doing its ow
the requirements of the application p
arguments:

The function in question can complet
the knowledge and help of the appli
communication system. Therefore, prov
of the communication system itself is
version of the function provided by the
performance enhancement.)

We call this line of reasoning against low-level function implementation the *end-to-end argument*. The following sections examine the end-to-end argument in detail, first with a case study of a typical example in which it is used—the function in question is reliable data transmission—and then by exhibiting the range of functions to which the same argument can be applied. For the case of the data communication system, this range includes encryption, duplicate message detection, message sequencing, guaranteed message delivery, detecting host crashes, and delivery receipts. In a broader context, the argument seems to apply to many other functions of a computer operating system, including its file system. Examination of this broader context will be easier, however, if we first consider the more specific data communication context.

2. CAREFUL FILE TRANSFER

2.1 End-to-End Caretaking

Consider the problem of *careful file transfer*. A file is stored by a file system in the disk storage of computer A. Computer A is linked by a data communication network with computer B, which also has a file system and a disk store. The object is to move the file from computer A's storage to computer B's storage without damage, keeping in mind that failures can occur at various points along the way. The application program in this case is the file transfer program, part of which runs at host A and part at host B. In order to discuss the possible threats to the file's integrity in this transaction, let us assume that the following specific steps are involved:

(1) At host A the file transfer program calls upon the file system to read the file from the disk, where it resides on several tracks, and the file system passes it to the file transfer program in fixed-size blocks chosen to be disk format independent.
(2) Also at host A, the file transfer program asks the data communication system to transmit the file using some communication protocol that involves splitting the data into packets. The packet size is typically different from the file block size and the disk track size.

ACM Transactions on Computer Systems, Vol. 2, No. 4, November 1984.

Host won't trust the network →
the network does not have to be perfect (and can't be) →

- Best-effort →
- End-to-end

A protocol for packet network intercommunication

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 individual network and which strongly influence the way in which internetwork communication can take place.

A typical packet switching network is composed of a set of computer resources called HOSTS, a set 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 reformatted into smaller pieces.

3) The success or failure of a transmission and its performance in each network is governed by different time delays in accepting, delivering, and transporting the data. This requires careful development of internetwork timing procedures to insure that data can be successfully delivered through the various networks.

4) Within each network, communication may be disrupted due to unrecoverable mutation of the data or missing data. End-to-end restoration procedures are desirable to allow complete recovery from these conditions.

Paper approved by the Associate Editor for Data Communications of the IEEE Communications Society for publications without oral presentation. Manuscript received November 5, 1973. The research reported in this paper was supported in part by the Advanced Research Projects Agency of the Department of Defense under Contract DAHC 15-73-C-0370.
V.G. Cerf is with the Department of Computer Science and Electrical Engineering, Standford University, Stanford, Calif.
R.E. Kahn is with the Information Processing Technology Office, Advanced Research Projects Agency, Department of Defense, Arlington, Va.

……

[9] S. Carr, S. Crocker, and V. Cerf, "HOST-HOST Communication Protocol In the ARPA Network," in *Spring Joint Computer Conf., AFIPS Conf. Proc.*, vol. 36. Montvale, N.J.: AFIPS Press, 1970, pp. 589-597.
[10] A. McKenzie, "HOST/HOST protocol for the ARPA network," in *Current Network Protocols*, Network Information Cen., Menlo Park, Calif., NIC 8246, Jan. 1972.
[11] L. Pouzin, "Address format in Mitranet," NIC 14497, INWG 20, Jan. 1973.
[12] D. Walden, "A system for interprocess communication in a resource sharing computer network," *Commun. Ass. Comput. Mach.*, vol. 15, pp. 221-230, Apr. 1972.
[13] B. Lampson, "A scheduling philosophy for multiprocessing system," *Commun. Ass. Comput. Mach.*, vol. 11, pp. 347-360, May 1968.
[14] F. E. Heart, R. E. Kahn, S. Ornstein, W. Crowther, and D. Walden, "The interface message processor for the ARPA computer network," in *Proc. Spring Joint Computer Conf., AFIPS Conf. Proc.*, vol. 36. Montvale, N.J.: AFIPS Press, 1970, pp. 551-567.
[15] N. G. Anslow and J. Hanscoff, "Implementation of international data exchange networks," in *Computer Communications: Impacts and Implications*, S. Winkler, Ed. Washington, D. C., 1972, pp. 181-184.
[16] A. McKenzie, "HOST/HOST protocol design considerations," INWG Note 16, NIC 13879, Jan. 1973.
[17] R. E. Kahn, "Resource-sharing computer communication networks", *Proc. IEEE*, vol. 60, pp. 1397-1407, Nov. 1972.
[18] Bolt, Beranek, and Newman, "Specification for the interconnection of a host and an IMP," Bolt Beranek and Newman, Inc., Cambridge, Mass., BBN Rep. 1822 (revised), Apr. 1973.

Vinton G. Cerf was born in New Haven, Conn., in 1943. He did undergraduate work in mathematics at Stanford University, Stanford, Calif., and received the Ph.D. degree in computer science from the University of California at Los Angeles, Los Angeles, Calif., in 1972.

He was with IBM in Los Angeles from 1965 through 1967 and consulted and/or worked part time at UCLA from 1967 through 1972. Currently he is Assistant Professor of Computer Science and Electrical Engineering at Stanford University, and consultant to Cabledata Associates. Most of his current research is supported by the Defense Advanced Research Projects Agency and by the National Science Foundation on the technology and economics of computer networking. He is Chairman of IFIP TC6.1, an international network working group which is studying the problem of packet network interconnection.

Robert E. Kahn (M'65) was born in Brooklyn, N.Y., on December 23 1938. He received the B.E.E. degree from the City College of New York, New York, in 1960, and the M.A. and Ph.D. degrees from Princeton University, Princeton, N.J., in 1962 and 1964, respectively.

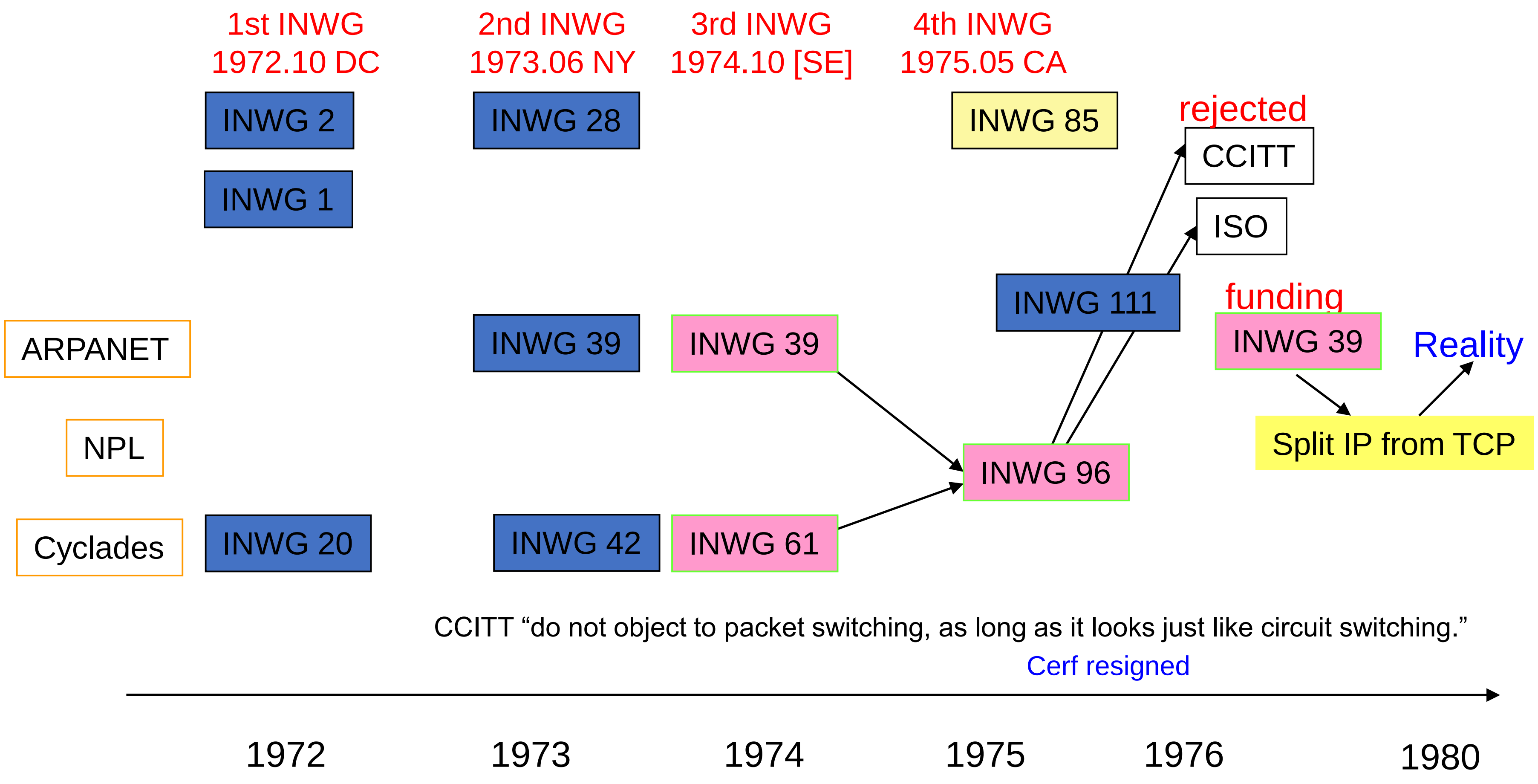
From 1960 to 1962 he was a Member of the Technical Staff of Bell Telephone Laboratories, Murray Hill, N.J., engaged in traffic and communication studies. From 1964 to 1966 he was a Ford Postdoctoral Fellow and an Assistant Professor of Electrical Engineering at the Massachusetts Institute of Technology, Cambridge, where he worked on communications and information theory. From 1966 to 1972 he was a Senior Scientist at Bolt Beranek and Newman, Inc., Cambridge, Mass., where he worked on computer communications network design and techniques for distributed computation. Since 1972 he has been with the Advanced Research Projects Agency, Department of Defense, Arlington, Va.

Dr. Kahn is a member of Tau Beta Pi, Sigma Xi, Eta Kappa Nu, the Institute of Mathematical Statistics, and the Mathematical Association of America. He was selected to serve as a National Lecturer for the Assocation for Computing Machinery in 1972.

© 1974 IEEE. Reprinted, with permission, from IEEE Trans on Comms, Vol Com-22, No 5 May 1974

案例3： Split IP from TCP

INWG 39 and TCP/IP



Split TCP and IP (1)

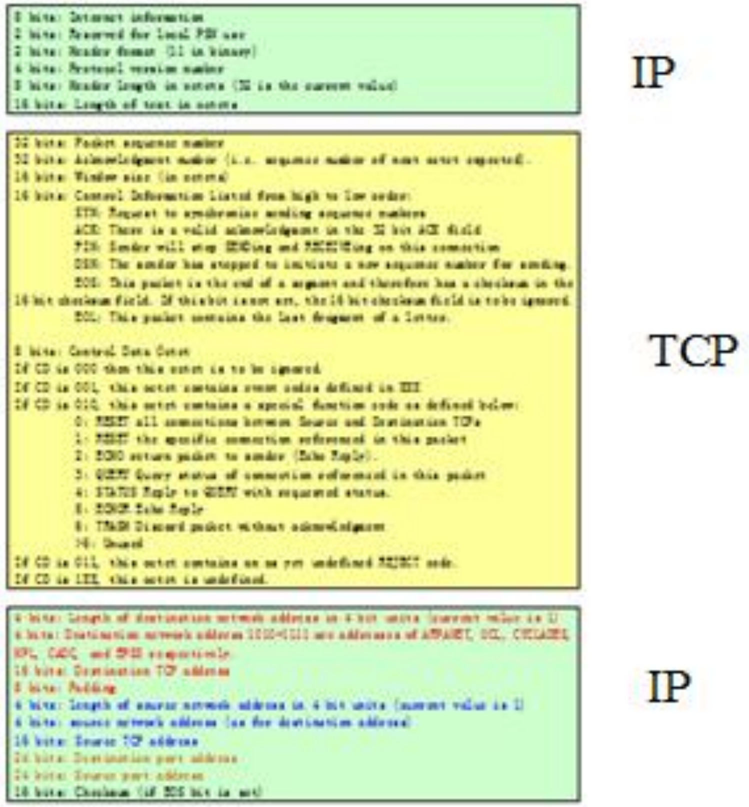
Network Working Group   Vinton Cerf
Request for Comments: 675   Yogen Dalal
NIC: 2   Carl Sunshine
INWG: 72   December 1974

SPECIFICATION OF INTERNET TRANSMISSION CONTROL PROGRAM

December 1974 Version

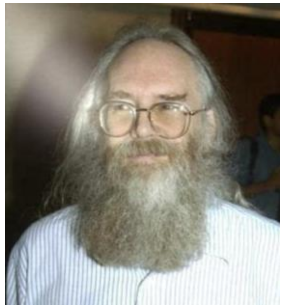
1. INTRODUCTION

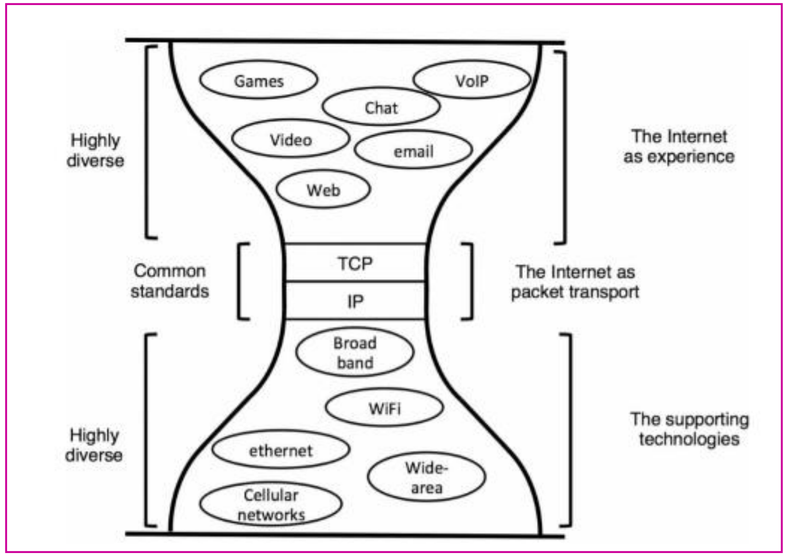
This document describes the functions to be performed by the internetwork Transmission Control Program [TCP] and its interface to programs or users that require its services. Several basic assumptions are made about process to process communication and these are listed here without further justification. The interested reader is referred to [CEKA74, TOML74, BELS74, DALA74, SUNS74] for further discussion.



Split TCP and IP (2)

- The initial effort to implement TCP resulted in a version that only allowed for virtual circuits.
- Different applications
  - File transfer
  - Voice
- In 1978, Cerf, Danny Cohen and Jon Postel, split the functions of TCP into two protocols, TCP and the Internet Protocol (IP).

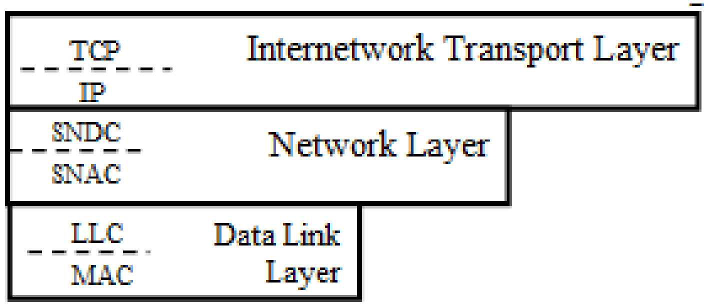




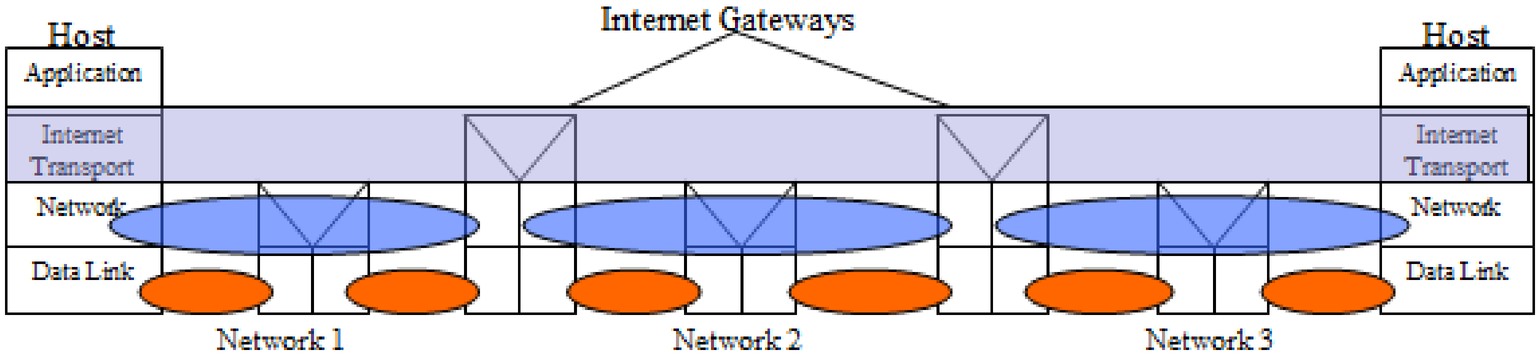
Vint Cerf

- Not designed for any specific applications: just move packets
- Designed to run over any communication technology
- Permission innovation at the edges
- Design to scale
- Open to new protocols, new technologies, new applications

INWG transport protocols



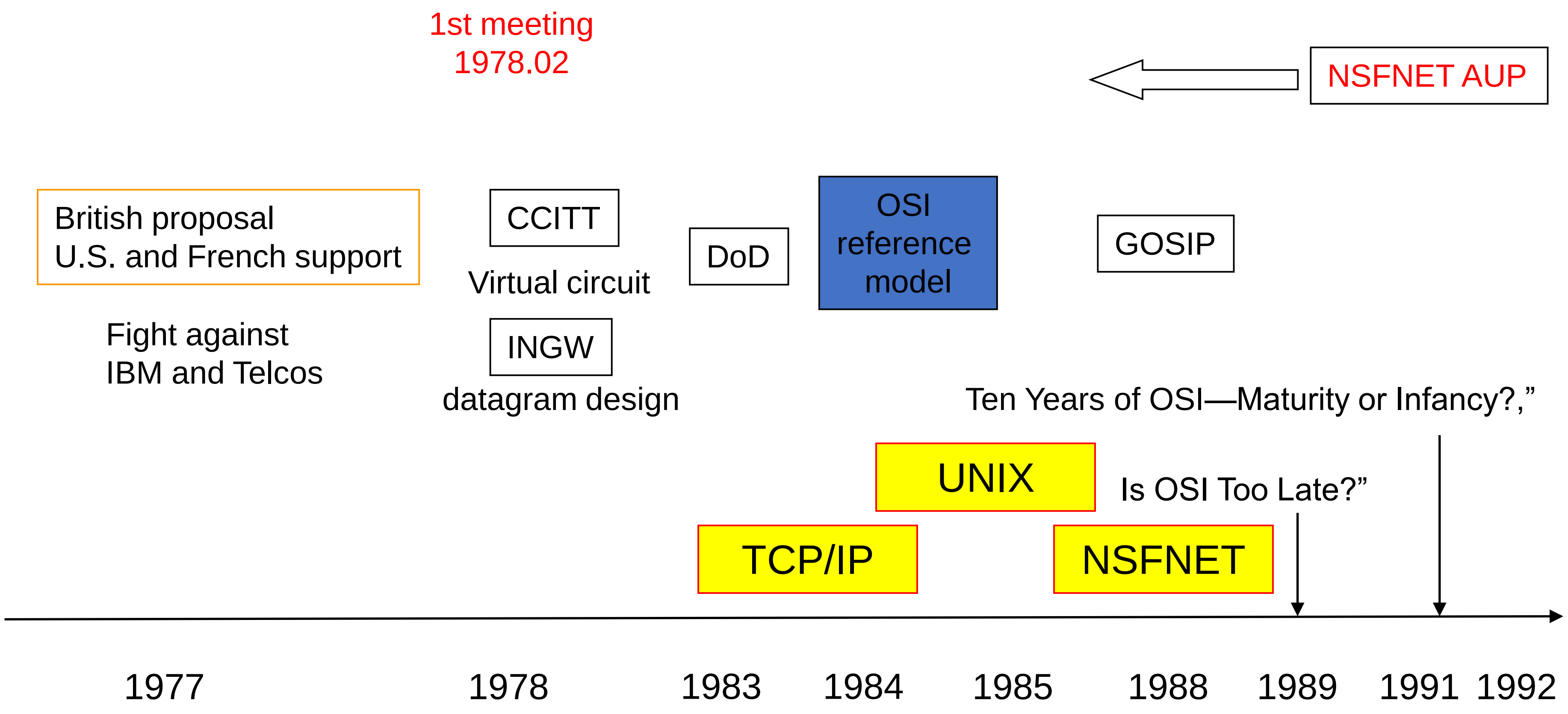
Three layers of different scope each with addresses



- Inter-domain routing: Internet layer
- Intra-domain routing: network layer
- Subnets: link layer

案例4： TCP/IP vs. OSI

OSI



Can you imagine trying to get the representatives from ten major and competing computer corporations, and ten telephone companies and PTTs], and the technical experts from ten different nations to come to any agreement within the foreseeable future?”

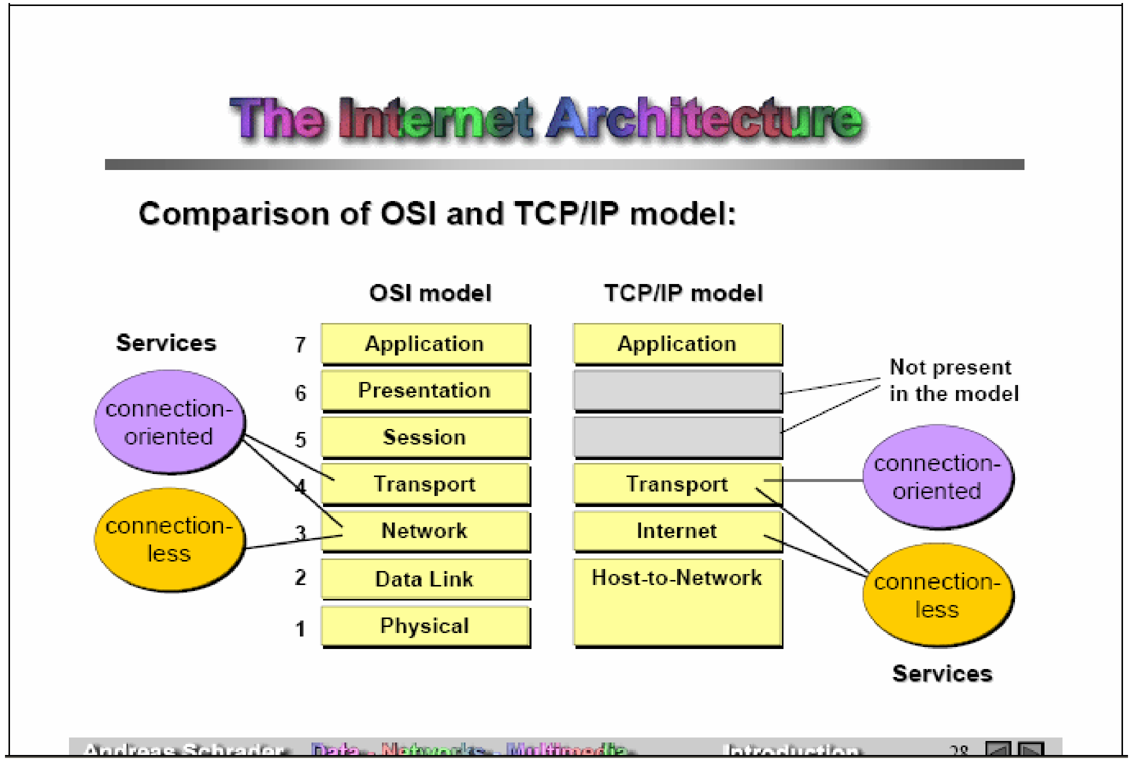
为什么TCP/IP战胜了OSI

At the time the second edition of this book was published (1989), it appeared to many experts in the field that the OSI model and its protocols were going to take over the world and push everything else out of their way. This did not happen. Why? A look back at some of the lessons may be useful. These lessons can be summarized as:

1. Bad timing.
2. Bad technology.
3. Bad implementations.
4. Bad politics.

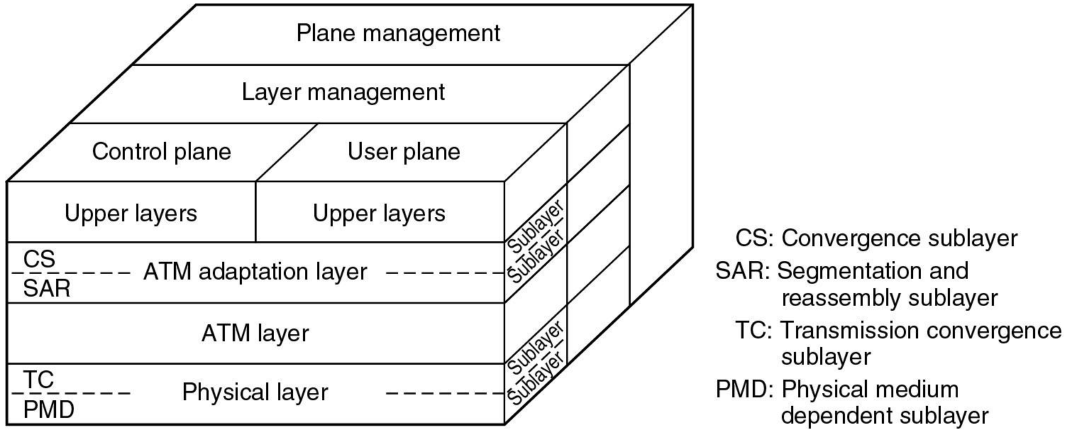
The OSI model, along with the associated service definitions and protocols, is extraordinarily complex.

from 《Computer Networks》 Fourth Edition by Andrew S. Tanenbaum, 2003

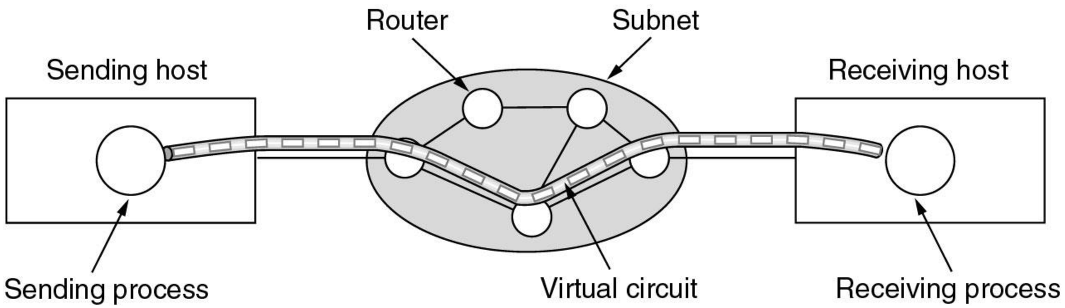


为什么TCP/IP战胜了ATM

三维模型



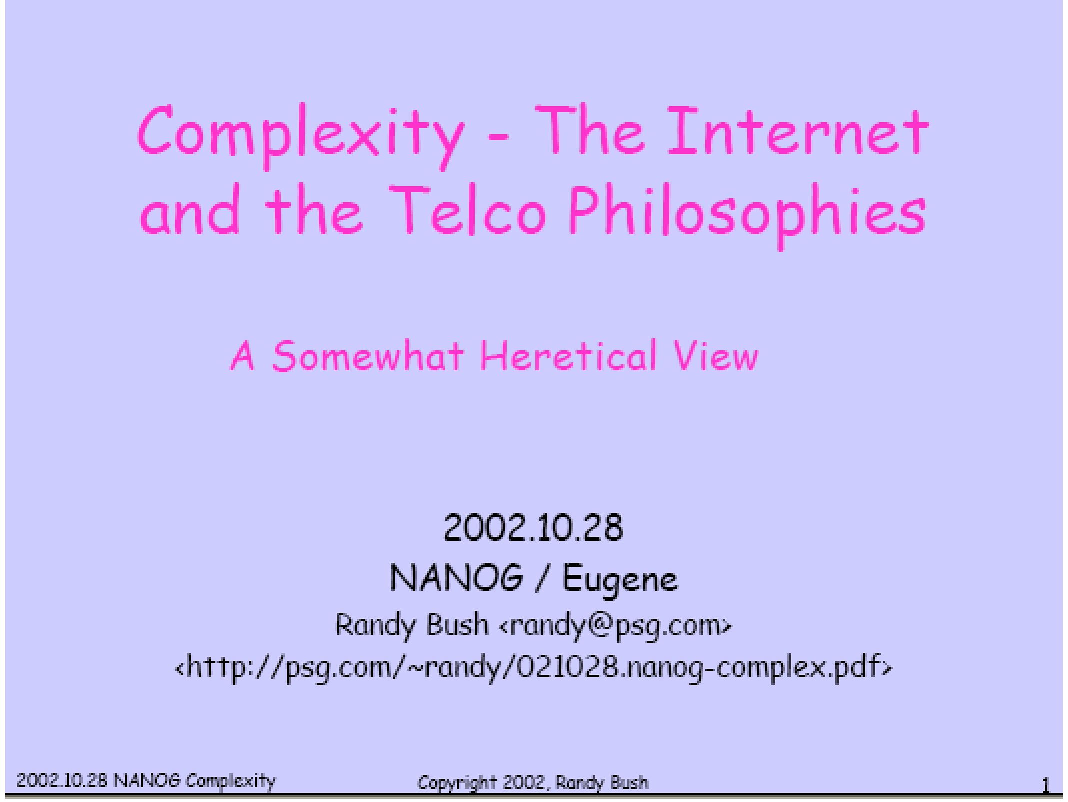
ATM was going to solve all the world's networking and telecommunications problems by merging voice, data, cable television, telex, telegraph, carrier pigeon, tin cans connected by strings, tom-toms, smoke signals, and everything else into a single integrated system that could do everything for everyone.



面向连接 虚电路

信元定长53字节

原因


Complexity - The Internet and the Telco Philosophies
A Somewhat Heretical View
2002.10.28
NANOG / Eugene
Randy Bush <randy@psg.com>
<http://psg.com/~randy/021028.nanog-complex.pdf>
2002.10.28 NANOG Complexity
Copyright 2002, Randy Bush
1


The Only Real Problem is
Scaling
All the others inherit from that one
If you can scale, everything else must be working.
-- Mike O'Dell - Chief Technologist UUNET
2002.10.28 NANOG Complexity
Copyright 2002, Randy Bush
4

互联网和X.25及ATM比较

	互联网（Internet）	X.25，ATM等
网络类型	无连接分组交换	虚电路分组交换
复杂度	端系统	网络
网络服务	尽力而为	保证服务质量
互联互通	容易	困难

案例5：　SIPP vs. CLNP

IPng

INET92

Kobe meeting

IAB proposal
for CLNP

IETF - 24

Cambridge, Massachusetts meeting

IETF participants protested to the Internet Society

iab-minutes-1992-06-18

In discussing TUBA, which proposes using CLNP in place of IP, several important limitations of CLNP were brought up: the way it embeds routing hierarchy in the addresses (Huitema), and the even- tual need to support realtime service (Braden). Cerf proposed a principle: if CLNP were to be adopted for use in the Internet, the IAB/IETF must retain flexibility for its evolution in the Internet context. It was suggested that we call the new Internet layer "IP v7", to make clear that the intent is to use only the CLNP format and that the IETF "owns" the spec. Kent suggested that change control be explicitly cited as one selection criterion.

draft-iab-ipversion7-00.txt

Internet growth has created serious problems of address space consumption and routing information explosion. A solution to these problems requires a new version of the Internet Protocol, which we call IP version 7 ("IPv7"). This memo presents architectural guidelines that any IPv7 should meet. It then discusses how an IPv7 based upon the OSI CLNP protocol would meet these requirements, and presents the reasons for the IAB's preference for this solution. Finally, it makes a three-part recommendation: (1) proceed at full speed on CIDR; (2) do the design work on IPv7 based on CLNP; and (3) continue to pursue research in advanced routing and other future extensions of the Internet architecture.

IETF24

7:00-10:00 pm — Tuesday, July 14, 1992 - Evening Sessions

BOF — Internet Society Q&A (isoc) (Vint Cerf/CNRI)

4:00-6:00 pm — Technical Presentations

- "Trusted NFS: Protocol Extensions for MultiLevel Security" (Fred Glover/DEC)
- "IP Encapsulation" (Bob Hinden/Sun and Dave Crocker/TBO)
- "A Cloudy Crystal Ball - Visions of the Future" (Dave Clark/MIT)

1992.06 — 1992.07

Why IPng?

- Problems with IPv4
  - "Address is running out!"
  - Routing table explosion
- Short term solutions
  - NAT
  - CIDR
- Long term solution
  - IPng

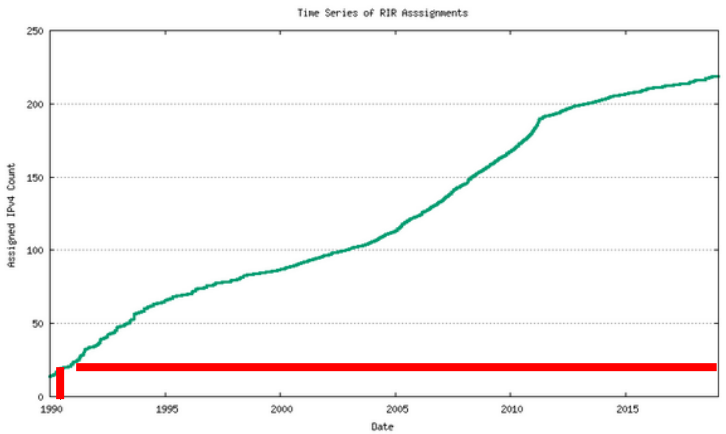
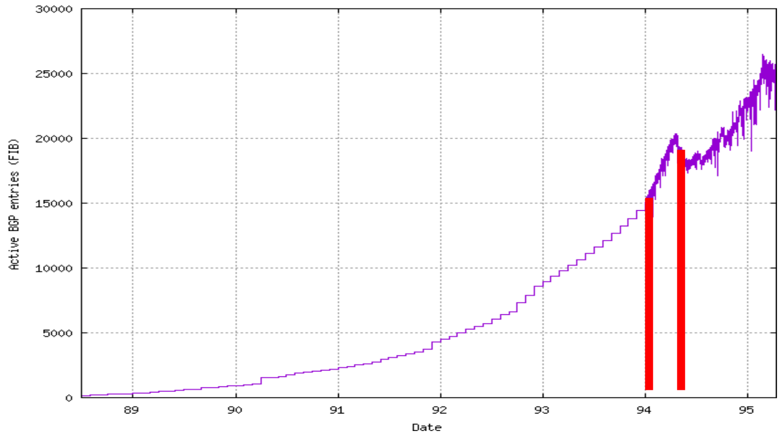
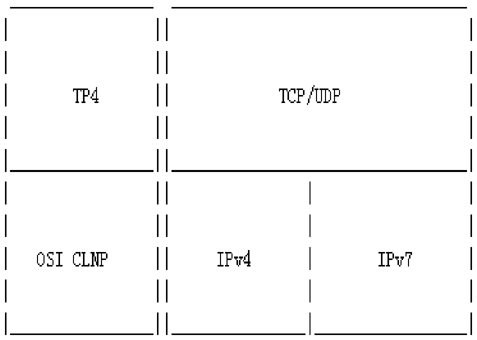


Figure 8 - Cumulative RIR Address assignments



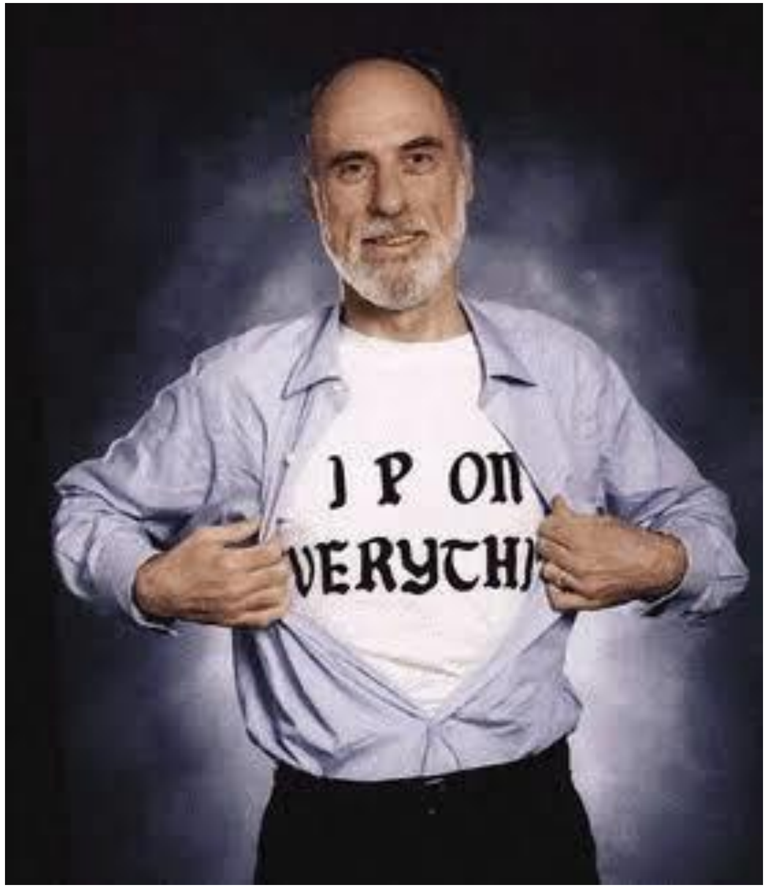
IAB’s proposal

Internet growth has created serious problems of address space consumption and routing information explosion. A solution to these problems requires a new version of the Internet Protocol, which we call IP version 7 ("IPv7"). This memo presents architectural guidelines that any IPv7 should meet. It then discusses how an IPv7 based upon the OSI CLNP protocol would meet these requirements, and presents the reasons for the IAB's preference for this solution. Finally, it makes a three-part recommendation: (1) proceed at full speed on CIDR; (2) do the design work on IPv7 based on CLNP; and (3) continue to pursue research in advanced routing and other future extensions of the Internet architecture.



- 1992年6月IAB宣布一项OSI技术-无连接网络协议(CLNP)—将成为互联网路由和编址的未来标准时，遭到了IETF参与者的直接反抗。
  - IAB之所以选择CLNP,是因为它能够迅速解决编址和路由问题。IAB拒绝了IETF指导组(IESG)关于允许对此问题进一步研究和实验六个月的建议。IAB主席断言：。IP地址太少、路由条目太多，这一问题切切实实而且迫在眉睫，对未来全球互联网的发展威胁显而易见。传统的IETF决策流程在这一问题面前不合适，我们不能指望在百花齐放的过程中，让'好的选择'在部署和实验中自动浮现出来，这样整个互联网都太危险了，在整个选择流程走完之前，互联网可能已经在自己爆炸式的成功中淹死了"(Lyman Chapin,JUl 1,1992,cited in Hofmann 1998, 15).
  - IAB的这一决定引发了一次抗议风暴，在接下来的一次IETF会议中该决定被全部撤销。这个争议迫使IETF不得不面对互联网社区中谁做决策以及决策者如何选择这样的基本问题。
  - Steve Crocker领导成立了新工作组——互联网标准组织流程，后来被称为POISED工作组(RFC 1396)然而，这个正式选举领导人的建议很快便终止了。IETF没有开始选举的最重要原因是，Jon Postel和其他几位资深人物郑重声明他们拒绝参与任何选举制度。这些技术精英对民主方法和公共问责(public accountablity)非常厌恶。
  - 1992年那次致命的失策破坏了对IAB好几年的信任，同样也玷污了互联网协会。直到1996年，ISOC和IAB才在更广泛的互联网社区中重新获得足够的权威来维持领导位置。

Vint Cerf



- The Internet could not have been so successful in the past years if IPv4 had contained any major flaw.
- Ten years of experience brought lessons.

David Clark

VIEWS OF THE FUTURE

A Cloudy Crystal Ball
--
Visions of the Future

David D. Clark

M.I.T. Laboratory for Computer Science

IETF, July 1992

Alternate title: Apocalypse Now

DDC 7/16/92 19:39 COPYRIGHT © David Clark 1992

VIEWS OF THE FUTURE

ATM -- A really big elephant

Myths from New Jersey:
- “They” will supply the scalable address space.
- “They” will solve the routing problem.
- ATM will solve the problem of real-time and QOS.
- “They” will be here real soon.

What are the real issues here?
- The network designers with telephony background do not understand multi-application networks.
- The phone companies have no history or approach to rapid deployment.
- They do not know how to do QOS either.

An example: why ATM LANs.
- My personal research: Everyone -> Sun -> standard.
- WHEN will the standard come? Mismatch possible.

VIEWS OF THE FUTURE

The last force on us -- us

The standards elephant of yesterday -- OSI.

The standards elephant of today -- it’s right here.

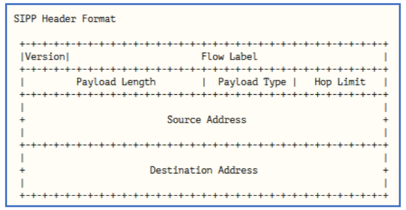
As the Internet and its community grows, how do we manage the process of change and growth?
- Open process -- let all voices be heard.
- Closed process -- make progress.
- Quick process -- keep up with reality.
- Slow process -- leave time to think.
- Market driven process -- the future is commercial.
- Scaling driven process -- the future is the Internet.

We reject: kings, presidents and voting.
We believe in: rough consensus and running code.

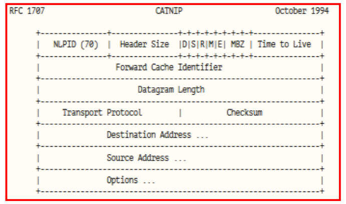
SLIDE 19

IP Versions

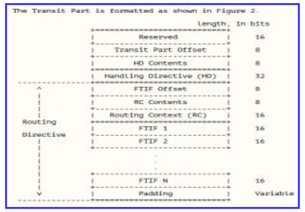
0	IP	March 1977 version	(deprecated)	
1	IP	January 1978 version	(deprecated)	
2	IP	February 1978 version A	(deprecated)	
3	IP	February 1978 version B	(deprecated)	
4	IPv4	September 1981 version	(current widespread)	RFC791
5	ST	Stream Transport (not a new IP, little use)		
6	IPv6	December 1998 version	(formerly SIP, SIPP)	RFC8200, RFC4291
7	CATNIP	IPng evaluation	(formerly TP/IX; deprecated)	RFC1707
8	Pip	IPng evaluation	(deprecated)	RFC1621
9	TUBA	IPng evaluation	(deprecated)	RFC1347, RFC1606
10-15		unassigned		



IPv6 (SIPP: RFC1710)

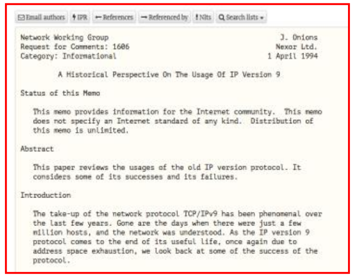


IPv7 (CATNIP: RFC1707)



IPv8 (Pip: RFC1621)

IPv9 (TUBA: RFC1347)



IPv9 (April 1: RFC1606)

案例6： World Wide Web

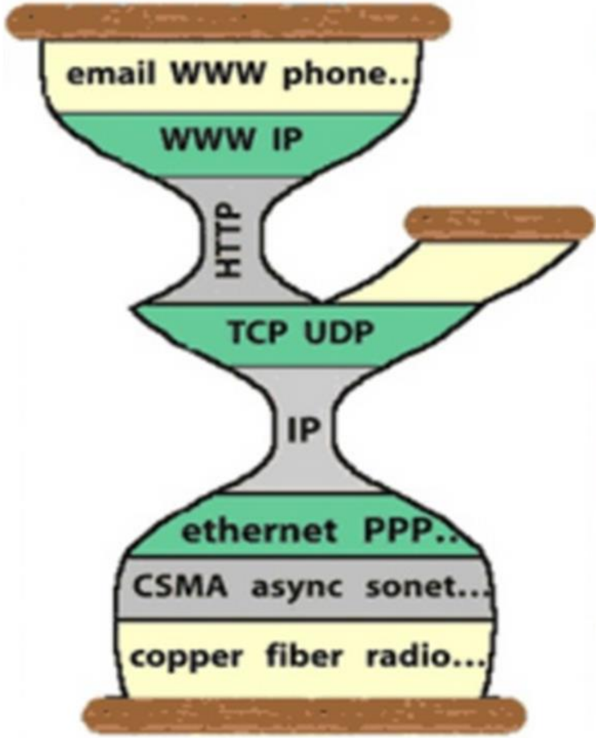
WWW

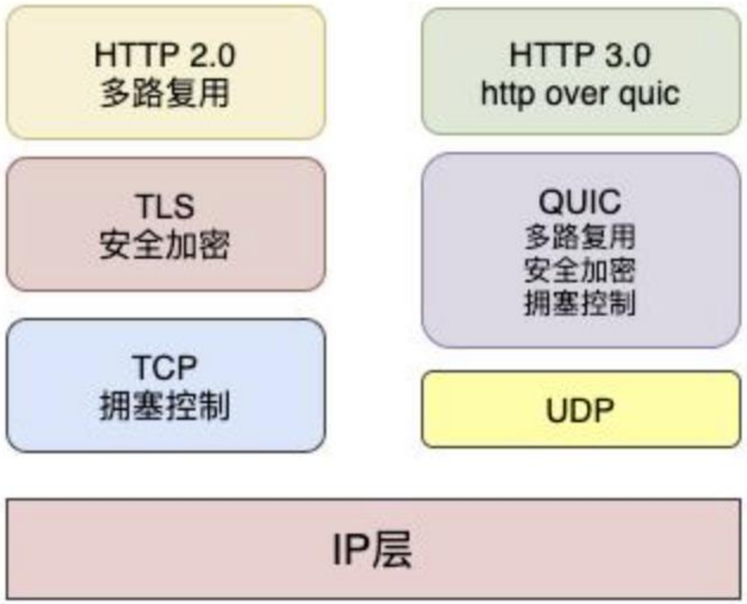
- 1980年，伯纳 · 李（Tim Berners-Lee）到CERN工作半年，需要对大量的内部资料（声图文）进行关联和查询，他编了一段 “内部问询（Enquire Within）” 程序
- 经过10年努力，1989年再次来到CERN，提出开发World Wide Web 的建议，但未获支持，只能业余进行研究
  - 能否不按照物理地址，而是根据内容进行查询
  - 能否用统一的格式对声图文等各类资源进行标注
- 1990年12月，编写了世界上第一个万维网服务器和浏览器的程序，1991年在CERN内部使用，并在网上免费提供全部技术资料
- 1993年，第一个图形界面浏览器研制成功，取名Mosaic
- 1995年，图形界面的浏览器Netscape Navigator上市，最普及的是Netscape公司的Navigator和微软的Internet Explorer

概念和机理

- 两个概念
  - 超文本概念-Hypertext：不是顺序关系，而是链接关系
  - 通用资源定位概念-Universal Resource Locator：对各种资源统一定位
- 两项实现技术
  - 超文本传输协议-HTTP
  - 超文本表识语言-HTML：如何链接、是哪种资源（声图文等）

演进





HTTP2.0和HTTP3.0能力矩阵对比

理念

- *Principles such as simplicity and modularity are the stuff of software engineering*
- *Decentralisation and tolerance are the life and breath of Internet*

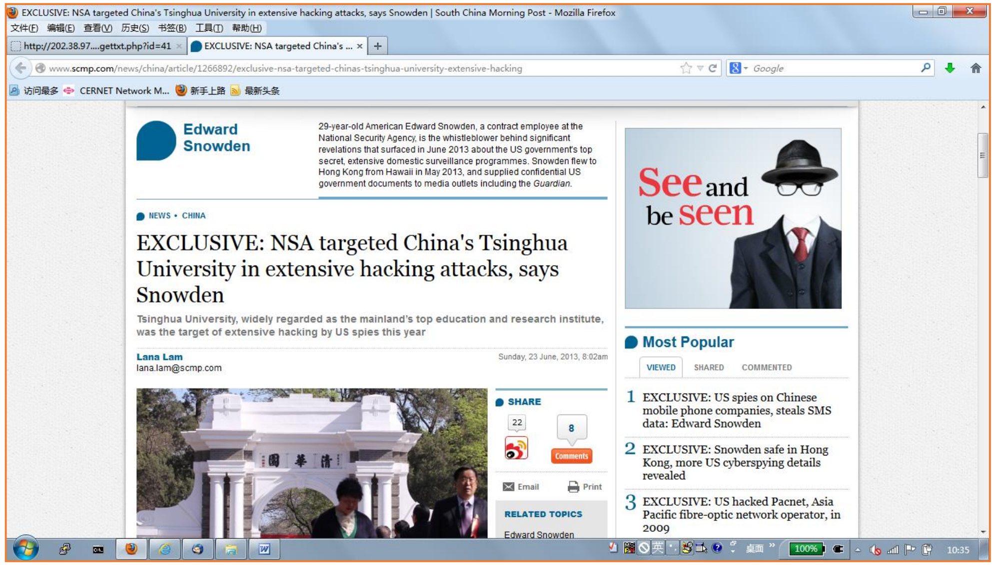
- 你可以让围墙花园变得非常动人，但从长远来看，外面的丛林永远都是更有吸引力的那一个。

  – Tim Berners-Lee

案例7：IAB

Open Internet Keys

Snowden



IETF87

IETF88

Encryption without authentication

Sweden event

- The IETF is willing to respond to the pervasive surveillance attack?
  - Overwhelming YES. Silence for NO.
- Pervasive surveillance is an attack, and the IETF needs to adjust our threat model to consider it when developing standards track specifications.
  - Very strong YES. Silence for NO
- The IETF should include encryption, even outside authentication, where practical.
  - Strong YES. Silence for NO
- The IETF should strive for end-to-end encryption, even when there are middleboxes in the path.
  - Mixed response, but more YES than NO.
- Many insecure protocols are used in the Internet today, and the IETF should create a secure alternative for the popular ones.
  - Mostly YES, but some NO.

[Hardening The Internet]

分层模型

Application（HTTP, HTTPS, [1.1, 2.0, 3.0]）

Names: DNS,DNSSEC, DOT, DOH

Certificate: DV 、OV 、 EV

Transport（TCP/UDP）：TLS、dTLS（传输层两元组）

IPv6 （网络层三元组）

link: Ethernet（fiber, UTP, WIFI）

Link: 2G、3G、4G、5G、6G

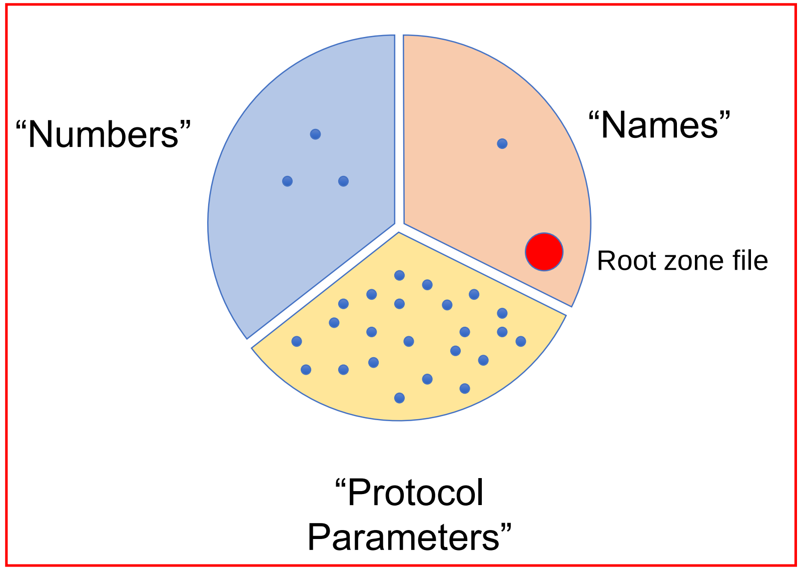
基础设施安全

- 路由系统安全
  - rPKI
- 域名空间安全
  - Root
  - DNSSEC
- 公钥证书安全
  - HTTPS
- 供应链安全
  - 商业软件
  - 开放源码
  - 硬件
  - 芯片

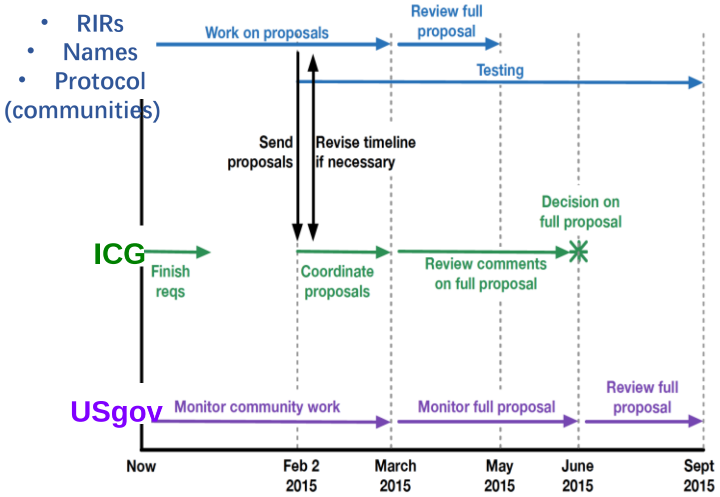
USG statement

- *The US government's role in IANA is purely clerical*
- *There are four key principles – and that's it*
  - Support and enhance the multistakeholder model
  - Maintain the security, stability, and resiliency of the Internet DNS
  - Meet the needs and expectation of the global customers and partners of the IANA services, and
  - Maintain the openness of the Internet
- *Governments are only one stakeholder and cannot be in charge*
- *The answer to the transition lies in IANA's 'customers'*
- *US domestic politics are a factor*
- *The bigger picture is developing countries and the multistakeholder process*
- *ICANN accountability is something for the community to figure out*

IANA transition

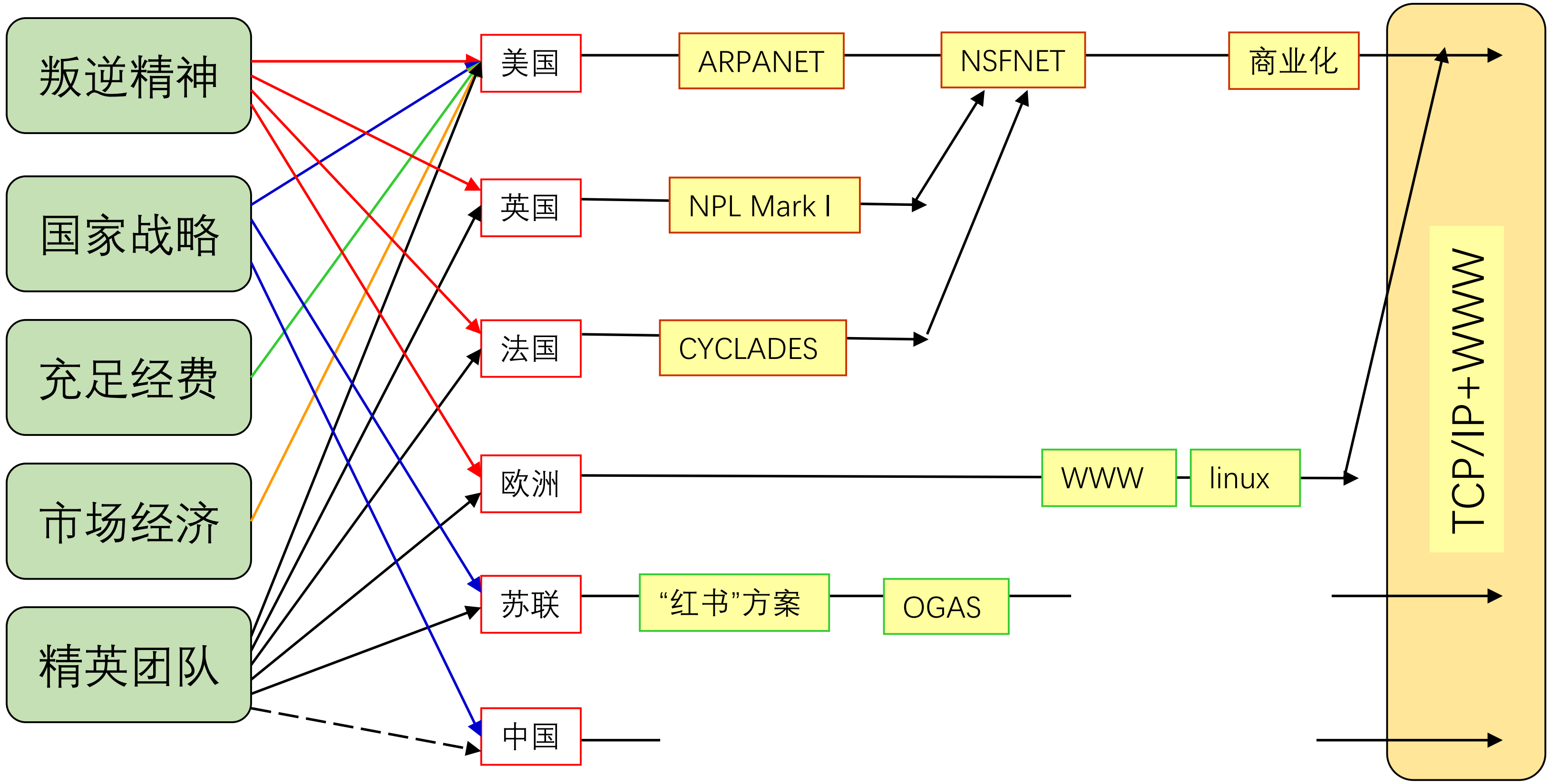


- Support and enhance the multistakeholder model
- Maintain the security, stability, and resiliency of the Internet DNS
- Meet the needs and expectation of the global customers and partners of the IANA services, and
- Maintain the openness of the Internet

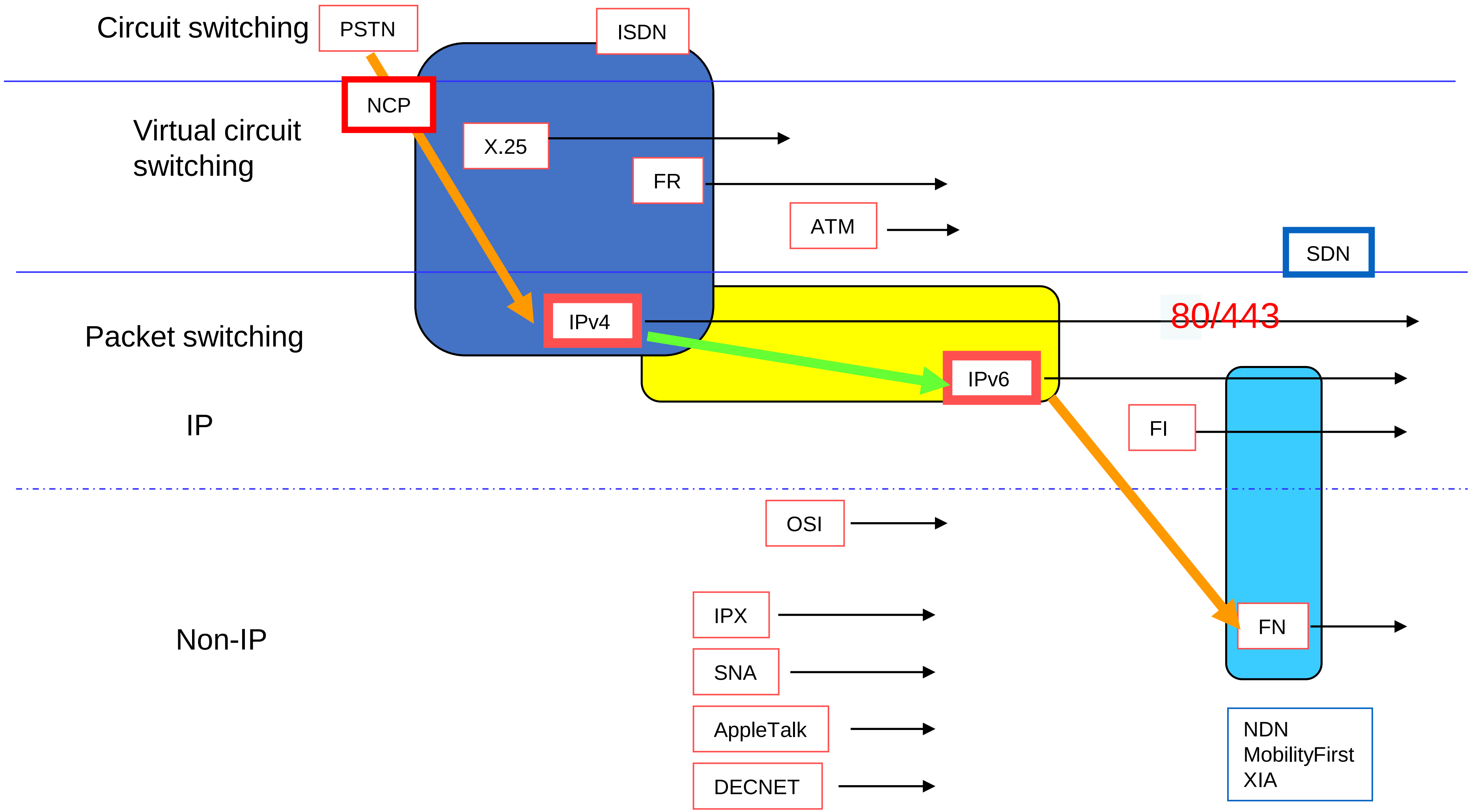


小结

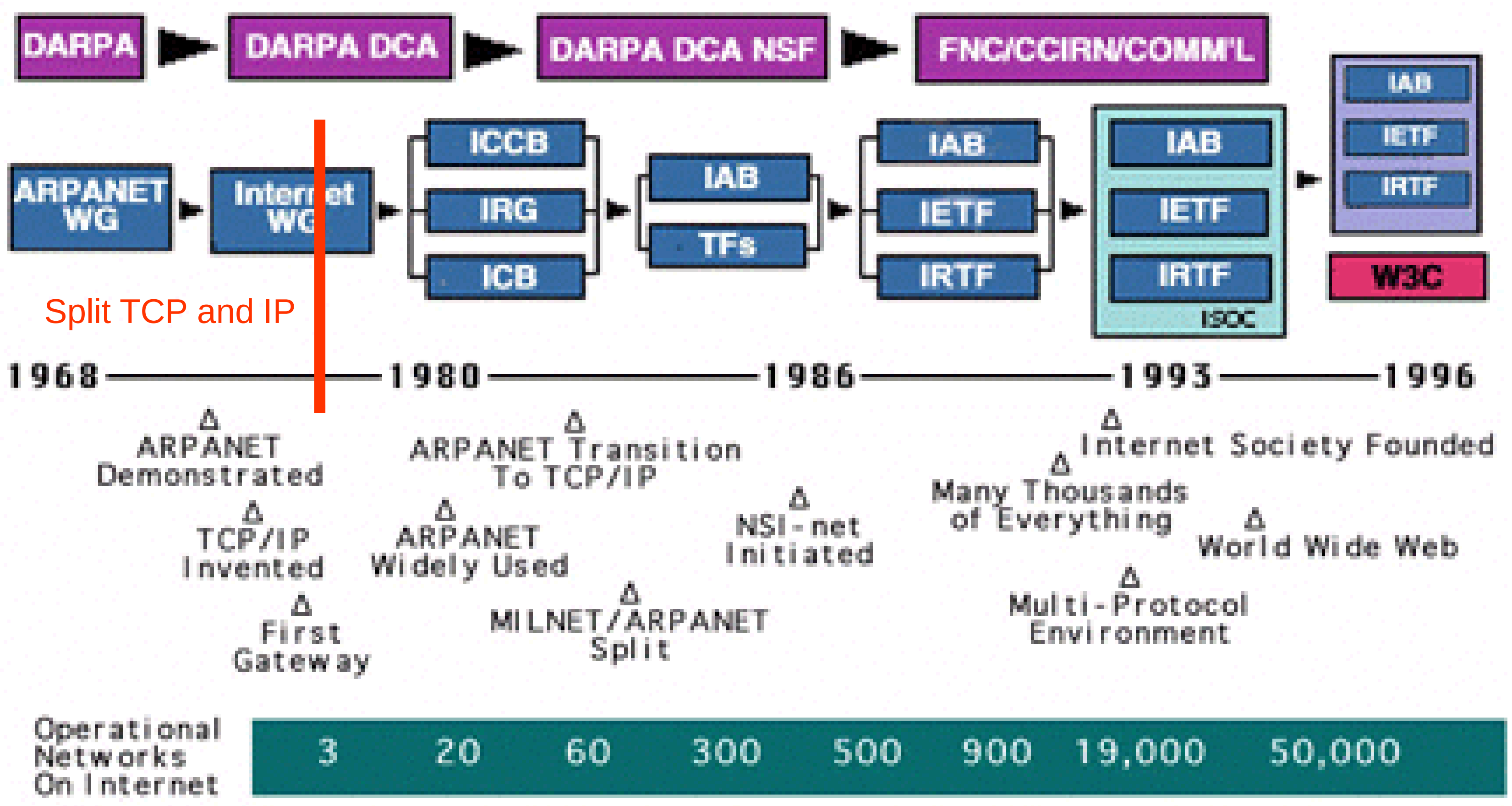
为什么是ARPANET?



网络体系结构演进


Circuit switching
PSTN
ISDN
NCP
Virtual circuit switching
X.25
FR
ATM
SDN
Packet switching
IPv4
80/443
IPv6
IP
FI
OSI
Non-IP
IPX
SNA
AppleTalk
DECNET
FN
NDN
MobilityFirst
XIA

TCP/IP标准化组织演进



Internet的设计原则

- 网络协议必须适应异种机之间的互联。
- 选择某一个方法（靠标准）。
- 具有很好的扩展性。
- 性能、成本和所能实现的功能的平衡点。
- 保持简单性。
- 模块化。
- 不要等待找到完美的解决方案。
- 尽量避免选项和参数。
- 在发送时应严格，在接收时应宽容。
- 小心处理自己没有请求而收到的分组。
- 避免循环依赖性。
- 对象应该能够自我描述。必须使用由IANA授权所使用的编码。
- 任何协议都应使用统一术语、注释、比特和字节顺序。
- 只有当实现了几个能够运行的程序实现后，Internet的协议才能成为标准。

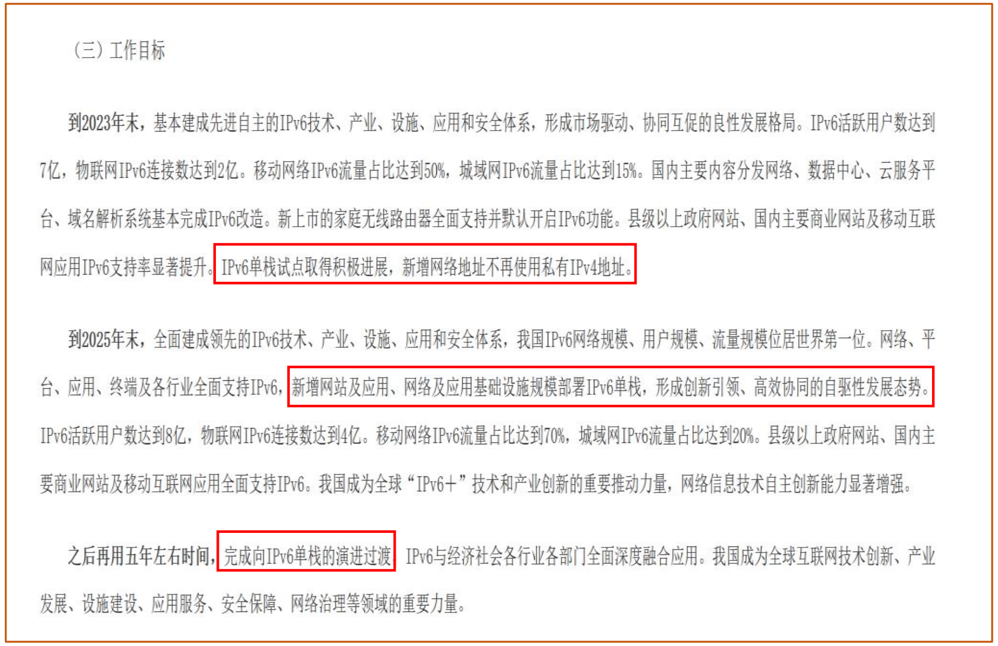
Project to 2051 (?!)

Characteristic	1981	2016	2051
Backbone Channel Capacity	5×10^4 bps	10^{11} bps	2×10^{17} bps
Personal Computer Storage	$< 10^6$ bytes	$\sim 10^{12}$ bytes	$\sim 10^{18}$
Telecomm Service Providers	$\sim 10^2$	$\sim 10^4$ (?)	10^6
Computers/User	~0.1	~10	10^3
Computers Connected	$\sim 10^4$	10^9+	10^{14}

Conclusion: We will not get there with the IP paradigm
(naming host interfaces, datagram service)
as the interoperability layer.

11th CFI Nanjing, June 2016

2021年网信办加快推进IPv6规模部署和应用

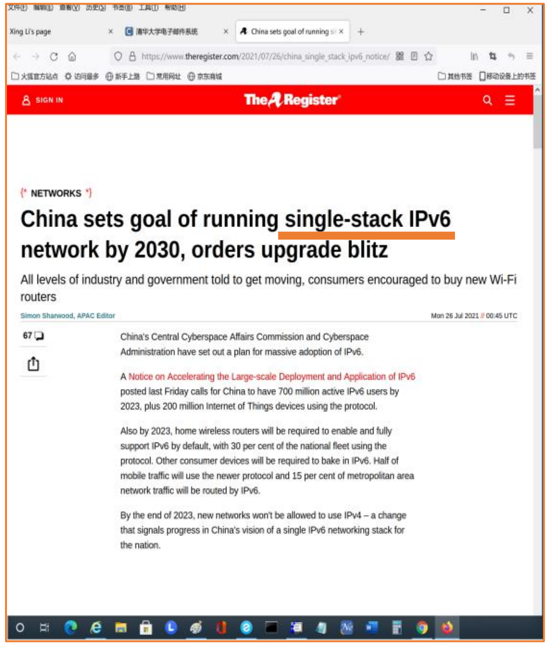


（三）工作目标

到2023年末，基本建成先进自主的IPv6技术、产业、设施、应用和安全体系，形成市场驱动、协同互促的良性发展格局。IPv6活跃用户数达到7亿，物联网IPv6连接数达到2亿。移动网络IPv6流量占比达到50%，城域网IPv6流量占比达到15%。国内主要内容分发网络、数据中心、云服务平台、域名解析系统基本完成IPv6改造。新上市的家庭无线路由器全面支持并默认开启IPv6功能。县级以上政府网站、国内主要商业网站及移动互联网应用IPv6支持率显著提升。IPv6单栈试点取得积极进展，新增网络地址不再使用私有IPv4地址。

到2025年末，全面建成领先的IPv6技术、产业、设施、应用和安全体系，我国IPv6网络规模、用户规模、流量规模位居世界第一位。网络、平台、应用、终端及各行业全面支持IPv6，新增网站及应用、网络及应用基础设施规模部署IPv6单栈，形成创新引领、高效协同的自驱性发展态势。IPv6活跃用户数达到8亿，物联网IPv6连接数达到4亿。移动网络IPv6流量占比达到70%，城域网IPv6流量占比达到20%。县级以上政府网站、国内主要商业网站及移动互联网应用全面支持IPv6。我国成为全球“IPv6+”技术和产业创新的重要推动力量，网络信息技术自主创新能力显著增强。

之后再用五年左右时间，完成向IPv6单栈的演进过渡，IPv6与经济社会各行业各部门全面深度融合应用。我国成为全球互联网技术创新、产业发展、设施建设、应用服务、安全保障、网络治理等领域的重要力量。



The Register

{* NETWORKS *}

China sets goal of running single-stack IPv6 network by 2030, orders upgrade blitz

All levels of industry and government told to get moving, consumers encouraged to buy new Wi-Fi routers

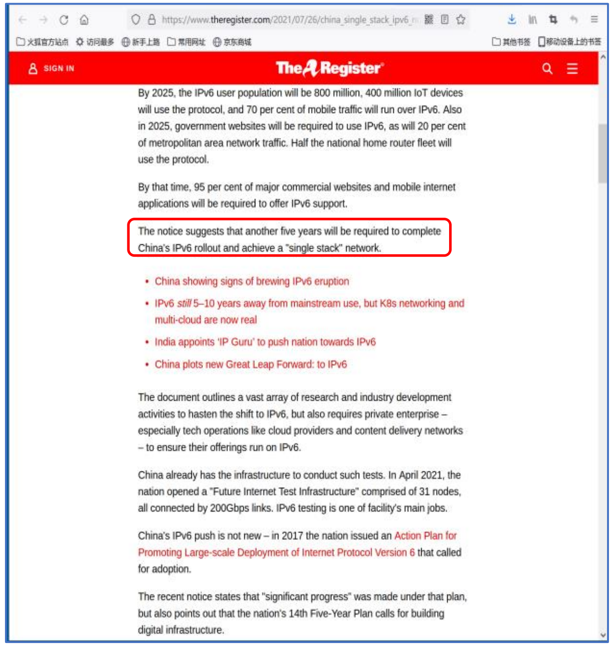
Simon Sharwood, APAC Editor — Mon 26 Jul 2021 // 00:45 UTC

China's Central Cyberspace Affairs Commission and Cyberspace Administration have set out a plan for massive adoption of IPv6.

A Notice on Accelerating the Large-scale Deployment and Application of IPv6 posted last Friday calls for China to have 700 million active IPv6 users by 2023, plus 200 million Internet of Things devices using the protocol.

Also by 2023, home wireless routers will be required to enable and fully support IPv6 by default, with 30 per cent of the national fleet using the protocol. Other consumer devices will be required to bake in IPv6. Half of mobile traffic will use the newer protocol and 15 per cent of metropolitan area network traffic will be routed by IPv6.

By the end of 2023, new networks won't be allowed to use IPv4 – a change that signals progress in China's vision of a single IPv6 networking stack for the nation.



https://www.theregister.com/2021/07/26/china_single_stack_ipv6_n

The Register

By 2025, the IPv6 user population will be 800 million, 400 million IoT devices will use the protocol, and 70 per cent of mobile traffic will run over IPv6. Also in 2025, government websites will be required to use IPv6, as will 20 per cent of metropolitan area network traffic. Half the national home router fleet will use the protocol.

By that time, 95 per cent of major commercial websites and mobile internet applications will be required to offer IPv6 support.

The notice suggests that another five years will be required to complete China's IPv6 rollout and achieve a "single stack" network.

- China showing signs of brewing IPv6 eruption
- IPv6 *still* 5–10 years away from mainstream use, but K8s networking and multi-cloud are now real
- India appoints 'IP Guru' to push nation towards IPv6
- China plots new Great Leap Forward: to IPv6

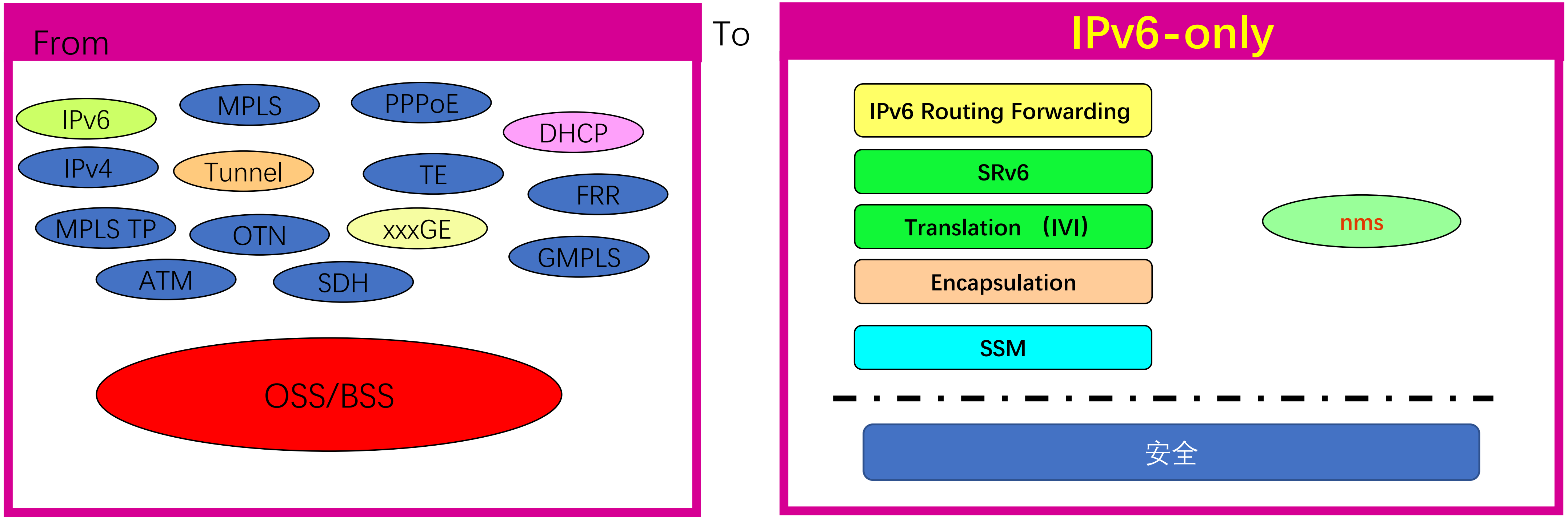
The document outlines a vast array of research and industry development activities to hasten the shift to IPv6, but also requires private enterprise – especially tech operations like cloud providers and content delivery networks – to ensure their offerings run on IPv6.

China already has the infrastructure to conduct such tests. In April 2021, the nation opened a "Future Internet Test Infrastructure" comprised of 31 nodes, all connected by 200Gbps links. IPv6 testing is one of facility's main jobs.

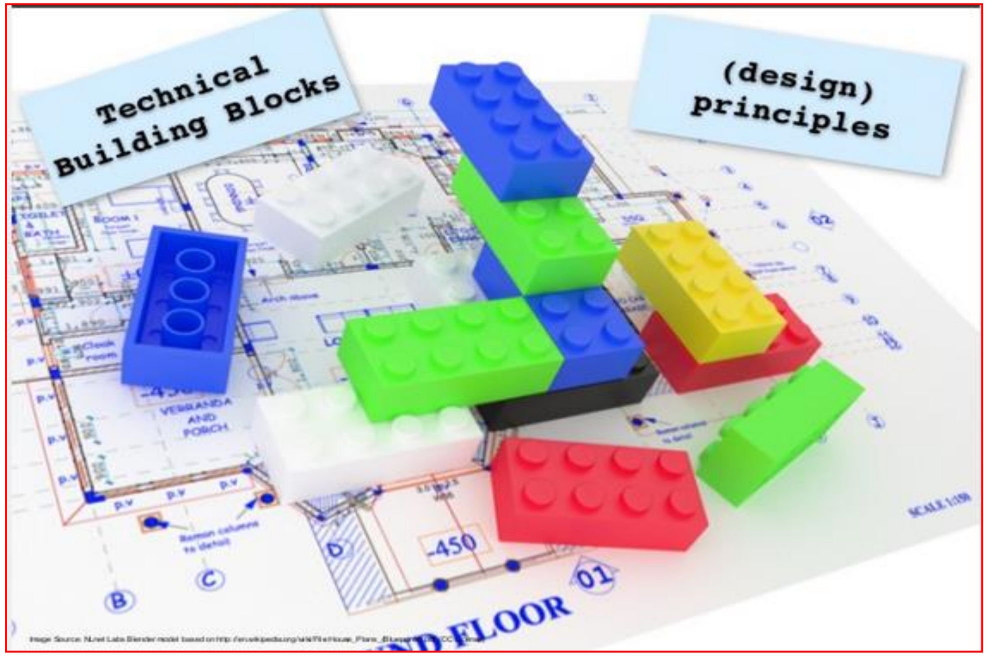
China's IPv6 push is not new – in 2017 the nation issued an Action Plan for Promoting Large-scale Deployment of Internet Protocol Version 6 that called for adoption.

The recent notice states that "significant progress" was made under that plan, but also points out that the nation's 14th Five-Year Plan calls for building digital infrastructure.

大道至简


From
IPv6
MPLS
PPPoE
DHCP
IPv4
Tunnel
TE
FRR
MPLS TP
OTN
xxxGE
GMPLS
ATM
SDH
OSS/BSS
To
IPv6-only
IPv6 Routing Forwarding
SRv6
Translation （IVI）
Encapsulation
SSM
nms
安全

互联网核心技术演进



设计原则

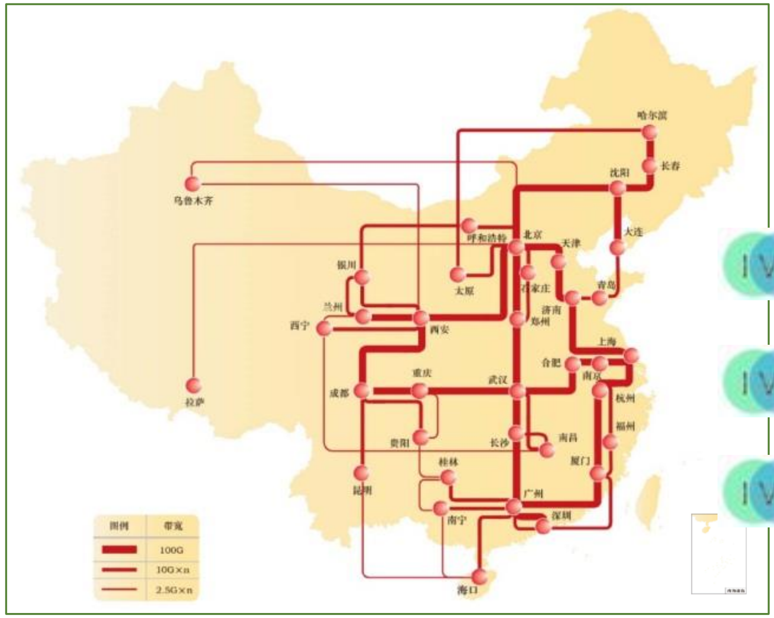
技术模块

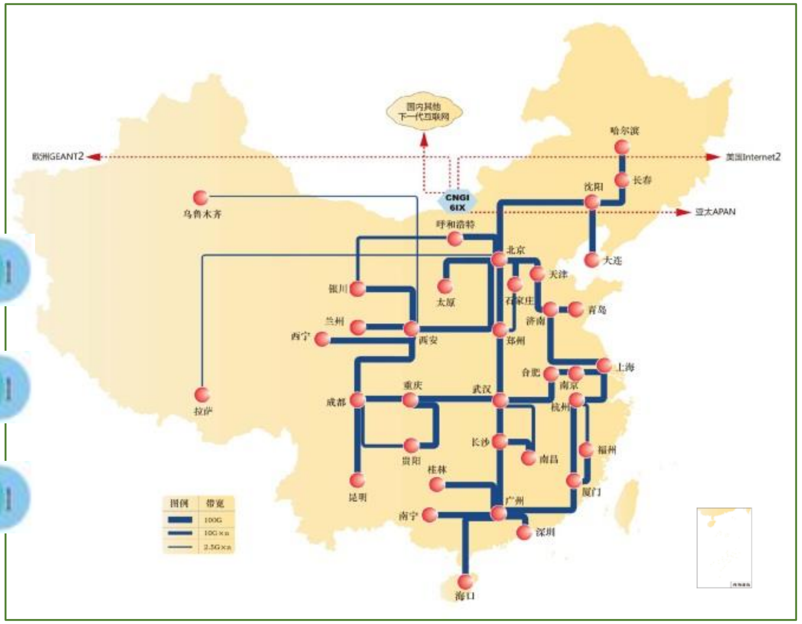
NCP | TCP/IP | DNS, BGP | HTTP | HTTPS | ？？？

1970s | 1980s | 1990s | 2000s | 2010s | 2020s

CERNET

学术互联网CERNET | **下一代互联网示范网络CERNET2** | **未来互联网试验设施FITI**





CERNET lambda (100G)

AS4538
IPv4:CIDR blocks (/8)
IPv6: 2001:250::/32

AS23910+(AS24348-24372)+AS23911
IPv4: none
IPv6： 2001:da8::/32 + 240c:c000::/20 +

AS38272： AS38272+(AS142650 -146745)+AS38255
IPv4: none
IPv6: 240a:a000::/20

RFC1925

```
Network Working Group                                  R. Callon, Editor
Request for Comments: 1925                                          IOOF
Category: Informational                                     1 April 1996


                   The Twelve Networking Truths

Status of this Memo

   This memo provides information for the Internet community.  This memo
   does not specify an Internet standard of any kind.  Distribution of
   this memo is unlimited.

Abstract

   This memo documents the fundamental truths of networking for the
   Internet community. This memo does not specify a standard, except in
   the sense that all standards must implicitly follow the fundamental
   truths.
```

网络的十二条军规（1）

- 真理 #1: 必须有效
- 真理#2: 不管你如何增加推力，你也不可能超越光速。
  - 2a (推论). 不论你怎么努力，也不可能让胎儿在9个月之前正常出生。拔苗助长只会适得其反。
- 真理#3: 只要推力足够，猪也能在天上飞。不过这并不是个好主意，因为着陆可能会是个问题。还有，当猪在天上飞的时候，你在下面会很危险。
- 真理#4: 有些事情如果不是亲自动手，这辈子都不会理解或体会到。如果某人没有真正搭建过商业网络设备或没有运营过网络，是无法完全理解网络的。
- 真理#5: 很多人都希望通过一个解决方案同时解决多个分立的问题。在大多数情况下，这都是个坏主意。

网络的十二条军规（2）

- 真理#7: 总是有原因的
  - 7a (推论): 好用，快速，廉价：三者只能选其二(不可能三者都占着)
- 真理#8:事情总是比你想的复杂.
- 真理#9: 不管什么资源，多多益善
  - 9a (推论) 所有网络问题的解决时间总要比看上去长
- 真理#10: 不同情况分别对待
- 真理#11: 所有的老方案都会换个名字或者换个陈述方法重新被提出来，不论以前这种方案是否成功过。
  - 11a (推论). 参考6a。
- 真理#12: 在协议设计上，完美并不是指该添加的都添加了，而是指该去掉的都去掉了。

小结

- Solve the real problem：解决真实问题
- Focus on the key issues：聚焦关键问题
- Think globally：全局考虑
- Actively participate in the mailing list：主要通过email进行交流
- Speak to the friends：口口相传
- Encourage Young people：鼓励年轻人参与
- Have fun：乐在其中

问题???

- *Net Neutrality*
- *Protocol ossification*
- *Internet fragmentation*

谢谢！