- 网络技术巨擘Jeff Doyle经典力作
- 《TCP/IP路由技术》（第一卷、第二卷）完全进阶
- 钜细靡遗地剖析、对比OSPF与IS-IS协议

PEARSON

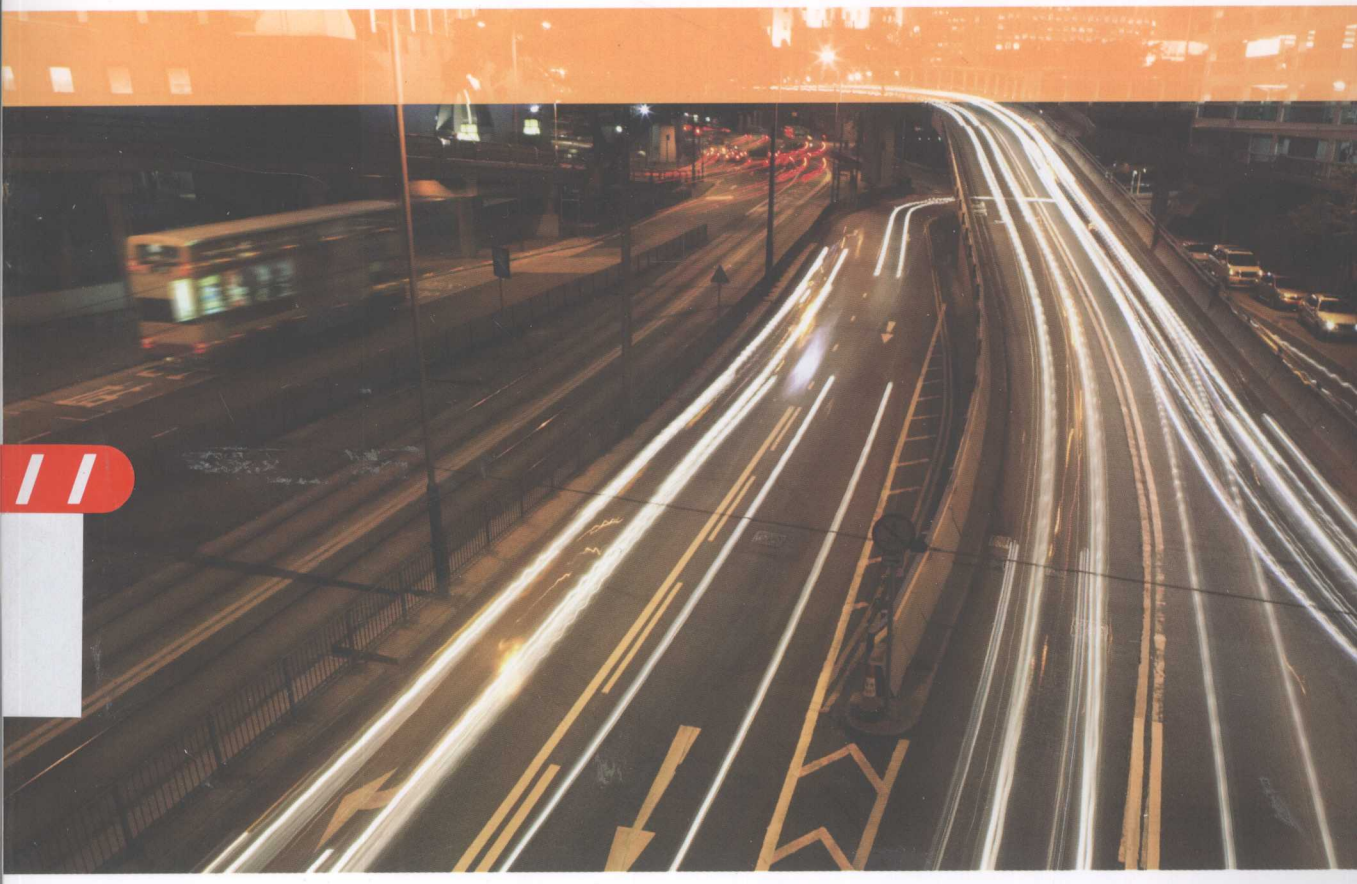
OSPF和IS-IS详解

OSPF and IS-IS

Choosing an IGP for Large-Scale Networks

[美] Jeff Doyle 著

孙余强 译



人民邮电出版社
POSTS & TELECOM PRESS

014041411

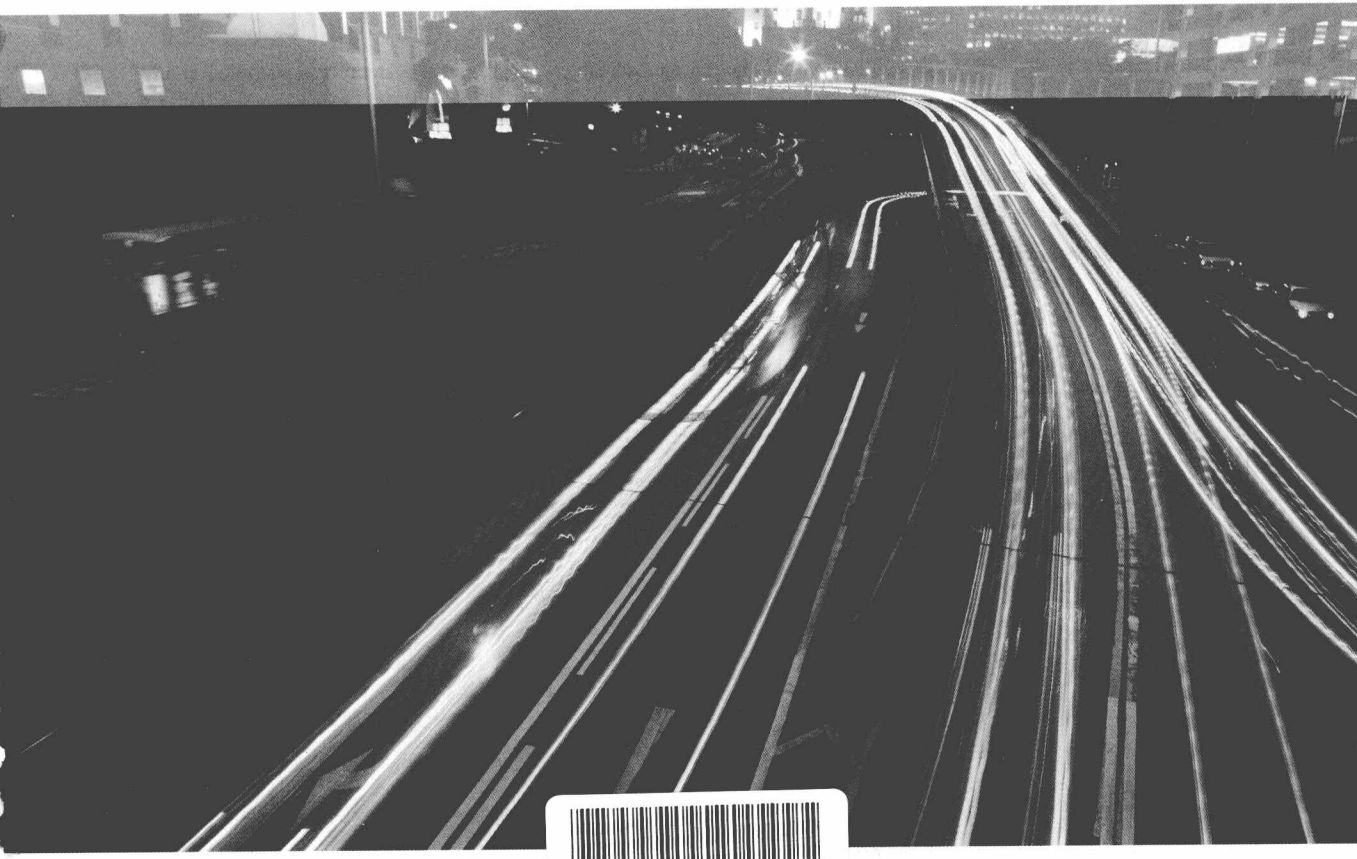
TN915.05
139

OSPF和IS-IS详解

OSPF and IS-IS

Choosing an IGP for Large-Scale Networks

[美] Jeff Doyle 著
孙余强 译



北航 C1724588

TN915.05
139

人民邮电出版社
北京

图书在版编目（CIP）数据

OSPF和IS-IS详解 /（美）多伊尔（Doyle, J.）著；
孙余强译. -- 北京：人民邮电出版社，2014.5
ISBN 978-7-115-34788-6

Ⅰ. ①O… Ⅱ. ①多… ②孙… Ⅲ. ①互联网络一路由协议一指南 Ⅳ. ①TN915.05-62

中国版本图书馆CIP数据核字(2014)第039719号

版 权 声 明

Jeff Doyle: OSPF and IS-IS: Choosing an IGP for Large-Scale Networks

Copyright © 2006 Pearson Education, Inc.

ISBN: 978-0321168795

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise without the prior consent of Addison Wesley.

版权所有。未经出版者书面许可，对本书任何部分不得以任何方式或任何手段复制和传播。

本书中文简体字版由人民邮电出版社经 Pearson Education, Inc.授权出版。版权所有，侵权必究。

本书封面贴有 Pearson Education（培生教育出版集团）激光防伪标签。无标签者不得销售。

◆ 著 [美] Jeff Doyle
译 孙余强
责任编辑 傅道坤
责任印制 彭志环 杨林杰

◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路 11 号
邮编 100164 电子邮件 315@ptpress.com.cn
网址 http://www.ptpress.com.cn
北京艺辉印刷有限公司印刷

◆ 开本：800×1000 1/16
印张：30.5
字数：634 千字 2014 年 5 月第 1 版
印数：1－3 000 册 2014 年 5 月北京第 1 次印刷
著作权合同登记号 图字：01-2012-3156 号

定价：99.00 元

读者服务热线：(010)81055410 印装质量热线：(010)81055316
反盗版热线：(010)81055315

内容提要

本书是在大型网络中部署 OSPF 和 IS-IS 协议的权威指南，作者以对比的方式讲解了如何在部署大型网络时分别实施 OSPF 和 IS-IS 协议，并从这两种协议的可扩展性、可靠性，以及安全性等方面给出了契合实际的建议和答案。

关于作者

Jeff Doyle，IP 路由协议、MPLS 及 IPv6 技术专家，主持或参与设计过的大型 IP 服务提供商网络遍及北美、欧洲、日本、韩国及中国大陆。Jeff 是《TCP/IP 路由技术》第 1 卷和第 2 卷的作者，同时还是 *Juniper Networks Routers: The Complete Reference* 一书的编辑及特约作者。他代表 Juniper 公司出席过无数场企业研讨会，并同时在 NANOG、JANOG、APRICOT 以及 IPv6 论坛会议上发表过多次演讲。

加盟 Juniper 公司之前，Jeff 在 International Network Services 公司任资深网络系统咨询顾问一职，自那时起，他就开始专攻 IP 路由协议的设计。Jeff 持有孟菲斯州立大学文学学士学位，还在新墨西哥大学学习过电气工程专业。目前，Jeff 与妻子和 4 个孩子居住在科罗拉多州丹佛市。

试读结束，需要全本PDF请购买 www.ertongbook.com

献辞

谨将本书献给我的父母 L.H.和 Louise Doyle，我爱你们。

致谢

首先要感谢 Catherine Nolan 以及 Addison-Wesley 出版社全体编辑、制作以及营销人员，感谢你们对我本人以及我迟迟不能交稿的极度容忍。还要感谢本书的项目编辑 Laurie McGuire。这并不是我与 Laurie 的第一次合作，和以前一样，这次她又使我的文字看起来远高于我的真实水平。

感谢本书的技术审阅团队：Eural Authement、Hannes Gredler、Dave Humphrey、Pete Moyer Mike Shand 以及 Rena Yang。你们都身经百战，对技术无比精通，试问还有那位作者能够请得动这样一支技术审阅“梦之队”。我还要感谢 Ross Callon、Vint Cerf、Steve Crocker、Paul Goyette、Matt Kolon、Chelian Pandian、Russ White 以及我在 Juniper 公司专业服务团队的全体同仁，感谢你们对本书某些特定章节的评审与建议。

我的妻子 Sara，我的孩子 Anna、Carol、James 和 Katherine，都是我生命中最重要的人。你们的支持与鼓励是本书成功付梓的关键；你们所带给我的爱和欢笑，对我而言，都是无价之宝。

最后，我要感谢我前几本书的所有读者，来自世界各地的宝贵意见和慷慨赞美使我万分荣幸，我希望你们也能从本书中获益。

前言

本书起源于一份 PPT 文档。多年来，作者就用这套 PPT 文档在 IP 运营商和服务提供商的网络设计座谈会上，来比较并对比 OSPF 和 IS-IS。而这份 PPT 文档则是作者日积月累，从无数次非正式的网络技术培训中积聚而成，其主要用途就是要让某些网络工程师能从容面对 IS-IS，这些网络工程师都经常接触 OSPF，但对 IS-IS 却知之不深，或一无所知。

随着大型网络数量的不断增多，迫切需要更多能同时掌握 OSPF 和 IS-IS，且能分清以上两种协议各自特征的网络工程师。为了让网络工程师更方便地理解这方面的知识，作者把那些多年来在现场演讲以及技术研讨会中所使用的素材集结成书。本书的内容也确实包含了某些读者一贯想要弄清的问题、疑虑以及难点。作者坚信本书能够对读者的学习和工作有所帮助。

本书的读者

本书主要面向那些精通 OSPF，并希望继续钻研 IS-IS 的网络工程师和架构师。因此，在本书的每一章节里，几乎都是 OSPF 的内容在前，IS-IS 的内容在后。之所以把对相关主题的 OSPF 实现的介绍置于 IS-IS 实现之前，是为了让读者用已然掌握的知识先行铺垫。

本书同样适用于那些不甚精通 OSPF 的读者。此外，有些读者对 OSPF 和 IS-IS 这两种路由协议中的一种或两种有着一般性的了解，但都迫切希望深入研究这两种协议；还有些读者只精通 IS-IS，对 OSPF 知之甚少（这样的人比较少见）。由于作者会在书中对 OSPF 和 IS-IS 做详尽论述，不偏不倚，因此本书亦适用于以上两类读者。

若读者对网络技术涉猎不深，但却想开拓自己在链路状态路由协议上的知识面，则可把本书作为教科书来读。本书的第 2 章专为这些读者而著，这一章内容为本书的后续章节打下了基础。在第 2 章中，作者全面介绍了路由协议的基本概念，尤其是链路状态路由协议。作者对其后各章内容也做了精心编排，以方便读者在阅读过程中由简入深。

对于那些立志参加网络技术认证（如 Cisco 公司 CCIE 认证或 Juniper 公司 JNCIE 认证）考试的读者来说，本书同样适合他们阅读。此类读者能从本书中获得为通过认证考

试所必须掌握的与 OSPF 和 IS-IS 有关的基本概念。然而，本书并不包含相关配置及排障方面的内容，请读者通过其他渠道获取这方面的信息，以便更为全面地准备认证考试。本书每一章的末尾都附有若干习题，其目的是让读者在阅读下一章之前巩固并测试自己对本章内容的理解。因习题答案均可见于正文，故不再单独给出。

什么是大型数通网络

OSPF 和 IS-IS 是仅有的两种适合在大型数通网络中运行的路由协议。但何谓大型数通网络，又该如何精确定义呢？回顾一下自计算机诞生之前就已经存在的数通网络及其相关理念，将会对此有所帮助。

当 Alexander Graham Bell 和 Theodore Vail 于 1885 年创立美国电话电报（American Telephone and Telegraph Company，AT&T）公司时，并没有打算提供电报服务。但二位都知道自己所构建的网络能够传输的东西不单是电话信号。除了知道像电话信号那样的信息表示方式之外，他们对其他的信息表示方式并无先见之明。电报反映出了 Bell 和 Vail 当时所理解的数通网络的雏形。

人们若要把信息传递到声音所不能企及的范围以外，就必须借助形形色色的广域数据通信手段。许多文明古国都使用火光信号来实现长途快速通信。在封建时期的日本，邻近的几个村庄都会在夜晚释放纸灯笼。灯笼会借“腹”中之火所产生的热空气冉冉上升，乃至明火高悬，这些村庄就以这样的方式来互报平安。在整个美国西南部，可以发现许多岩石雕刻——在岩石的侧面刻有数字和符号。这些雕刻出自几百年前生活在美国的土著人之手，其中包括猎人、士兵以及游客。尽管有一些雕刻只是人们自娱自乐，但也有许多雕刻被认为用来传递信号和消息，这都是过客们为路过此地的后来者而留。那些雕刻给数据通信带来了某些意想不到的启迪：信号本身总是固定不变，而信源和信宿（传递和接收信号的人）却总是在不停地移动。

电报网络是世界上首个利用电子数字信号（传递信息）的数通网络。毋庸置疑，虽然该网络属于广域网络——连接了多个国家，甚至通过海底电缆连接了几大洲——但就复杂性而言，它并不能算是我们现在所探讨的大型网络。在该网络内，通过人工操作，可以轻而易举地生成、传递以及接收信号，而且其维护手段，也是通过人工监控并干预。

那么，到底应该怎样给大型网络下定义呢？虽然网络所包括的节点数和链路数对此有所影响，但只根据那些数字来下定义，则未免以偏概全。相反的是，网络的规模应由其复杂程度来决定。通过人工直接干预就能管理得井井有条的网络，只能算是小型网络；也就是讲，早期的电报网络虽然在地理上横跨多个区域，但也只能视之为小型网络。在 IP 领域内，可以人为方式指定 IP 包的发送路线（routed statically），无需通过自动化管理系统来管理的网络，被称为小型网络。请读者注意上一句话中的“可”字，在一个特定的小型 IP 网络中，也同时有可能会运行某种动态路由协议或自动化管理软件，这也正是作者未说“‘必须’以人为方式指定 IP 包的发送路线”的原因所在。

随着网络的复杂程度不断增加，让其具备“自动运行”的能力，将会变得愈发重要。在一个中等规模的网络中，要想通过人工监控和干预的方式，来执行 IP 包的静态路由及相关网络管理工作，则未免显得不太实际。对于中等规模的 IP 网络，需要运行如 RIP 这样的动态路由协议，来维护 IP 包在多条路径之间的转发。

然而，当网络的规模达到一定程度时，自动化系统自身的健壮性将会成为重中之重。如第 2 章所述，在充满“变数”的复杂网络中运行简单动态路由协议（如 RIP），很有可会故障百出。那么，现在可以总结出大型网络的一个重要特征了。所谓大型网络，其内所运行的自动化管理系统需将网络本身作为一个实体来统一管理，而非只能单独管理节点间的链路。以下所列为管理大型网络所要考虑的因素：

- 各个节点之间复杂的交互方式；
- 复杂的与流量转发路径有关的设计考量（其中牵涉流量负载均衡）、流量的监控及分布方式、健壮的环路避免功能；
- 复杂的评判路由优劣的手段（complex link metrics）；
- 数据传输需求的多样性；
- 对安全性及可靠性的严格要求。

无论网络的规模如何，OSPF 和 IS-IS 都能轻松满足需求。这两种路由协议的核心价值就是，当运行两者的网络的规模不断增大时，两者本身在性能方面不会有丝毫折扣。尚无第三种 IGP 能满足世界上最大的 IP 网络——Internet——的路由选择需求。

IOS 与 JUNOS

本书通篇都以 Cisco 公司的 IOS 或 Juniper 公司的 JUNOS 来举例。但是在本书的前几章，作者会偶尔同时用以上两种 OS 来举例。作者的本意是要让读者理解 OSPF 和 IS-IS 协议本身，无意传授某种特定 OS 的配置方法。书中用 IOS 或 JUNOS 来举例，一是因为这两种 OS 都是路由器操作系统，二来则是作者恰好会用，而且能够接触到这两种 OS。读完本书之后，读者应该能够具备这样一种能力：只要捧起任一厂商的设备配置手册，就能轻松自如地在该厂商的设备上配置 OSPF 和 IS-IS，并排除与这两种路由协议有关的故障。

目录

试读结束，需要全本PDF请购买 www.ertongbook.com

第 1 章

链路状态路由协议之由来

本书的开篇方式极为特别。只要读者愿意，第 1 章可略过不读。若读者只准备了解 OSPF 和 IS-IS 技术方面的内容，请直接阅读第 2 章。本章不涉及技术内容，为非必读章节。作者之所以非要在这里说一说与链路状态路由协议有关的历史故事，理由很简单，那就是作者对某些事物的关注程度甚至还要超过网络技术，而历史正是其中之一。研究历史不但能帮助我们以正视听，而且还能使我们免遭满嘴谎言的奸商、政客以及其他奸诈小人的蒙蔽。专注于技术，通晓某些网络协议的运作方式固然是好事，但了解路由协议的历史，则既可以加深对它的理解，又能够拓展自己的知识面。此外，还能对为自己的网络甄选合适的路由协议有所帮助。退一万步来讲，研究历史也比去看那些胡编乱造的小说有趣多了。

要说链路状态路由协议的历史，就不能不提 Internet 及其前身 ARPANET 的历史，它们之间有很深的渊源。像 Internet 这样的大型网络的出现，带动了链路状态协议的需求和发展。因此，本章的内容实际上是 Internet 以及链路状态路由协议的发展简史。

1.1 星际网络

为当今 Internet 的诞生及发展做出过重要贡献的前辈数不胜数。但有一位名叫 J. C. R. Licklider 的智者做出了开创性的贡献，他为人谦逊，总让人喊他“Lick”，而不是“Licklider 博士”。他也不介意别人用他的创意来争名夺利。Licklider对许多事物都有好奇心，但（据他自己而言）均属浅尝辄止。以上特质加之他在解决问题方面的天赋，促使他成为对多个领域都有深入研究的通才。其实，他只是一位心理学家，并非工程师，这也解释了一

开始他在提出与计算和网络有关的创意时，为什么更偏重于两者在文化方面的作用而非技术本身。

20 世纪 50 年代中期，Licklider 在 MIT（麻省理工学院）研究心理声学（psychoacoustics）时，对计算机产生了浓厚的兴趣。他以计算机为工具，对人类的认知建模。在此期间以及担任 BBN（Bolt Baranek and Newman）公司（当时为一家专门从事声学工程咨询业务的公司）副总裁的任期内，他和他的弟子们一直宣扬一个理念，即要把计算机作为人类认知及通信的工具。1955 年之前，Licklider 还没怎么接触过计算机，但到了 1960 年初，他便成为此中高手，同时也是受到普遍认可的计算机科学领导者[1]。由他提出的多项开创性理论呈现在各种备忘录，以及后来的两篇重要论文（“Man-Computer Symbiosis”和“The Computer as a Communication Device”）中[2]。

实时性、交互式计算处理，是 Licklider 提出的重要理论之一，发表在其论文“Man-Computer Symbiosis”中。20 世纪 50 年代，计算由批处理来完成：人们将问题确定下来后，再编制程序，由计算机来计算答案。批处理所面临的问题恰恰是解决复杂问题的过程自身可能会改变原始问题。不可预见的变故（unforeseen alternatives）意味着重回起点，开始批处理的新轮回。Licklider 援引庞加莱的话，并概况为：“对问题来说，答案是什么并不重要，重要的是问题的本身。”（The question is not，“What is the answer?” The question is，“What is the question?”）人机实时交互，就可以随时引入问题解决过程中发现的新信息，对问题加以修正。

批处理还意味着计算机只能同时运行一个程序，每个用户需要轮流使用计算机。倘若能达成人与计算机间的实时交互，多个用户就应该可以同时使用计算机了。于是，时间共享（分时）理论成为了实时性、交互式计算理论的衍生。

Licklider 在同一篇论文里，还提出了另一个理论，即把电脑作为辅助人类思考的工具。他以自身的日常工作为例，据 Licklider 测算，他的大多数行为都属于机械性劳动或文案工作：“我的思考时间有 85%花在了寻找解决问题的出发点、做出决策以及学习一些有必

[1] Licklider 对家用电脑、图形用户界面、点击式输入设备，以及现代化计算机的许多方面都早有预言，其研究成果也使他成为人工智能之父。

[2] J. C. R. Licklider，“Man-Computer Symbiosis”，IRE Transactions of Human Factors in Electronics，Volume HFE-1，1960 年 3 月；“The Computer as a Communication Device”，Science and Technology，1968 年 4 月。
上述两篇论文的重印版本，会同 Bob Taylor 所撰写的一份对 J.C.R. Licklider 的简介，由 DEC 公司系统研究中心于 1990 年结集出版，命名为“怀念 J.C.R. Licklider（1915–1990）”，该书的电子版可由如下网址获得：gatekeeper.dec.com/pub/DEC/SRC/research-reports/SRC-061.pdf。

要掌握的知识上了，寻找和获得信息的时间要远远多于消化知识。”与人类相比，计算机发现和组织信息的速度要快很多，因此，合理的人机关系应当是：由计算机来担负信息获取和数据处理的枯燥工作，让人类腾出精力，专注于引入新信息，解决新问题。

基于上述理论，他又提出了另一个名为“思维中心”的理论。要想利用计算机来获取信息，需能访问到单台计算机所存储不下的海量信息。Licklider 提出的思维中心理论的主旨是，分处异地的多台计算机可以互相连网，并构成某种新型信息库。在“思维中心”内，随便哪一台计算机上的用户都能访问到另一台计算机里的信息。于是，一个重要理念——由分处异地的计算机构成的互连网络——应运而生。

其实，刊载于论文“The Computer as a Communication Device”里的上述理念的主旨，还不是要通过连网的计算机把不同地域的人与人“连接”在一起，而是要设法实现人与人之间最基本的心灵沟通。富于创造性的交互式通信要求作为载体的介质必须具有可塑性，只有如此，推理的前提方能借此而转化为结论。更为重要的是，那种介质必须是一种共有介质，如若不然，人们便无法集思广益、实验验证。Licklider 把计算机视为介质，利用其来创建人脑所想象不出来的模型。“迄今为止，数量最多、最精密、也最为重要的模型当属人的思维。人类思维之丰富、可塑、经济和便利是难以匹敌的，不过在其他方面，也并非没有短板。它既无法在静止状态钻研问题，也无法重复运转。运作方式则无从知晓。重感性，轻理性。所能访问者，唯一人所知。一举一动，唯一人洞悉，唯一人操弄。”上述理念立足于人机交互的早期理论，但是引出了利用广域网来执行分布式数据处理的基本思路。

1962 年 10 月，Licklider 受雇于美国国防部高级研究计划署（Advanced Research Projects Agency，ARPA）[3]，去领导其下辖的命令与控制研究室，以及行为科学处。命令与控制研究室又很快被更名为信息处理技术办（Information Processing Techniques Office，IPTO）。Licklider 不仅带来了他的分时理论和人机交互理论，还为 ARPA 吸引来了同时代的众多顶尖计算机科学家。

尽管 Licklider 只在 IPTO 工作到 1964 年，但他的理念却为星际计算机网络的诞生起到了播种的作用。追随他的那些科学家则成为了 ARPANET 发展中的关键性人物。

[3] 1972 年 ARPA 被莫名其妙地更名为国防高级研究项目计划署（Defense Advanced Research Projects Agency，DARPA）。

试读结束，需要全本PDF请购买 www.ertongbook.com