

Internet from its inception to the early 1990s⁶—and the design principles that were used to create that architecture: modularity, layering, and the end-to-end arguments.

While network engineers agree that the end-to-end arguments are among the few architectural principles underlying the architecture of the Internet,⁷ other scholars have offered widely differing and often contradicting views on what the end-to-end arguments are, what they say, and how they relate to the architecture of the Internet. For example, some suggest that the end-to-end principle is “an important architectural principle that has governed the Internet since its inception”⁸; others say that “the end-to-end principle simply does not dictate a robustly specified functional design for the network”⁹ but “follows from (and is an articulation of) the implicit design principle inherent to the layers model of the TCP/IP protocol [sic]”¹⁰; still others argue that “the Internet was never wholly end-to-end”¹¹ and that the end-to-end argument “is not an organizing principle; . . . if it is a principle, it is probably not true, and . . . even if it is true, it is probably not useful.”¹² In policy debates concerning the architecture of the Internet, opponents of regulation often argue that proponents of certain regulatory interventions (for example, of open-access rules or network-neutrality rules) have stretched the end-to-end principle beyond its original meaning.¹³ Even networking engineers often disagree about whether a certain technical solution violates the end-to-end arguments or not. In view of the high level of confusion, I discuss the end-to-end arguments and their relationship to the architecture of the Internet in detail.

My analysis yields an important insight: there is no single version of the end-to-end arguments, but two different ones that embody different rules for architectural design. The first version, which I call *the narrow version*, was presented by Jerome Saltzer, David Reed, and David Clark in the 1981 paper in which the end-to-end arguments were first named and identified as a design principle¹⁴; the second version, which I call *the broad version*, is the focus of later papers by these authors.¹⁵ The difference between the two versions is not immediately apparent, and Saltzer, Reed, and Clark never explicitly drew attention to the change in definition. There are, however, real differences in scope, content, and validity that make it necessary to distinguish between the two versions. At the same time, the silent coexistence of two different design principles under the same name explains some of the confusion that surrounds the end-to-end arguments.

Chapter 2 describes the design principles that were used to create the original architecture of the Internet and highlights the trade-offs involved in each of them. Chapter 3 describes how these design principles shaped

Internet Architecture and Innovation

Barbara van Schewick

The MIT Press
Cambridge, Massachusetts
London, England

© 2010 Barbara van Schewick

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

For information about special quantity discounts, email special_sales@mitpress.mit.edu.

Set in Stone Sans and Stone Serif by Toppan Best-set Premedia Limited. Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Van Schewick, Barbara.

Internet architecture and innovation / Barbara van Schewick.

p. cm.

Includes bibliographical references and index.

ISBN 978-0-262-01397-0 (hardcover : alk. paper) 1. Internet. 2. Computer network architectures. 3. Technological innovations. 4. Business—Data processing. I. Title.

TK5105.875.I57V378 2010

004.6'5—dc22

2009037130

10 9 8 7 6 5 4 3 2 1