#### UNITED STATES PATENT AND TRADEMARK OFFICE

#### BEFORE THE PATENT TRIAL AND APPEAL BOARD

UNIFIED PAENTS, INC.
Petitioner

v.

CATONIAN IP MANAGEMENT, LLC Patent Owner

United States Patent No. 8,799,468

\_\_\_\_\_

10 August 2017

DECLARATION OF SCOTT BENNETT, Ph.D.

#### **TABLE OF CONTENTS**

		Page
I.	INTRODUCTION	1
II.	BACKGROUND AND QUALIFICATIONS	1
III.	PRELIMINARIES	3
IV.	OPINIONS REGARDING INDIVIDUAL DOCUMENTS	9
V.	ATTACHMENTS	15
VI.	CONCLUSION	16

I, Scott Bennett, hereby declare under penalty of perjury:

#### I. INTRODUCTION

- 1. I have personal knowledge of the facts and opinions set forth in this declaration, I believe them to be true, and if called upon to do so, I would testify competently to them. I have been warned that willful false statements and the like are punishable by fine or imprisonment, or both.
- 2. I am a retired academic librarian working as a Managing Partner of the firm Prior Art Documentation LLC at 711 South Race Street, Urbana, IL, 61801-4132. Attached as Appendix A is a true and correct copy of my Curriculum Vitae describing my background and experience. Further information about my firm, Prior Art Documentation Services LLC, is available at <a href="https://www.priorartdocumentation.com">www.priorartdocumentation.com</a>.
- 3. I have been retained by Oblon, McClelland, Maier & Neustadt, LLP to authenticate and establish the dates of public accessibility of certain documents in an *inter partes* review proceedings for U.S. Patent No. 8,799,468. For this service, I am being paid my usual hourly fee of \$91/hour. My compensation in no way depends on the substance of my testimony or the outcome of this proceeding.

#### II. BACKGROUND AND QUALIFICATIONS

- 4. I was previously employed as follows:
- University Librarian, Yale University, New Haven, CT, 1994-2001;

- Director, The Milton S. Eisenhower Library, The Johns Hopkins
   University, Baltimore, MD, 1989-1994;
- Assistant University Librarian for Collection Management,
   Northwestern University, Evanston, IL, 1981-1989;
- Instructor, Assistant, and Associate Professor of Library
   Administration, University of Illinois at Urbana-Champaign, Urbana,
   IL, 1974-1981; and
- Assistant Professor of English, University of Illinois at Urbana-Champaign, 1967-1974.
- 5. Over the course of my work as a librarian, professor of English, researcher, and author of nearly fifty scholarly papers and other publications, I have had extensive experience with catalog records and online library management systems built around Machine-Readable Cataloging (MARC) standards. I also have substantial experience in authenticating printed documents and establishing the date when they were accessible to researchers.
- 6. In the course of more than fifty years of academic life, I have myself been an active researcher. I have collaborated with many individual researchers and, as a librarian, worked in the services of thousands of researchers at four prominent research universities. Over the years, I have read some of the voluminous professional literature on the information seeking behaviors of

academic researchers. And as an educator, I have a broad knowledge of the ways in which students in a variety of disciplines learn to master the bibliographic resources used in their disciplines. In all of these ways, I have a general knowledge of how researchers work.

#### III. PRELIMINARIES

- 7. Scope of this declaration. I am not a lawyer and I am not rendering an opinion on the legal question of whether any particular document is, or is not, a "printed publication" under the law.
- 8. I am, however, rendering my expert opinion on the authenticity of the documents referenced herein and on when and how each of these documents was disseminated or otherwise made available to the extent that persons interested and ordinarily skilled in the subject matter or art, exercising reasonable diligence, could have located the documents before 18 November 2003.
- 9. I am informed by counsel that an item is considered authentic if there is sufficient evidence to support a finding that the item is what it is claims to be. I am also informed that authenticity can be established based on the contents of the documents themselves, such as the appearance, contents, substance, internal patterns, or other distinctive characteristics of the item, taken together with all of the circumstances. I am further informed that an item is considered authentic if it

is at least 20 years old, in a condition that creates no suspicion of its authenticity, and in a place where, if authentic, it would likely be.

- 10. I am informed by counsel that a given reference is publicly accessible upon a satisfactory showing that such document has been disseminated or otherwise made available to the extent that persons interested and ordinarily skilled in the subject matter or art exercising reasonable diligence, can locate it. I have also been informed by counsel that materials available in a library constitute printed publications if they are cataloged and indexed (such as by subject) according to general library practices that make the references available to members of the interested public.
- 11. *Materials considered*. In forming the opinions expressed in this declaration, I have reviewed the documents and attachments referenced herein. These materials are records created in the ordinary course of business by publishers, libraries, indexing services, and others. From my years of experience, I am familiar with the process for creating many of these records, and I know these records are created by people with knowledge of the information in the record. Further, these records are created with the expectation that researchers and other members of the public will use them. All materials cited in this declaration and its attachments are of a type that experts in my field would reasonably rely upon and refer to in forming their opinions.

- 12. Persons of ordinary skill in the art. I am told by counsel that the subject matter of this proceeding relates to computer networking and distributed systems.
- 13. I have been informed by counsel that a "person of ordinary skill in the art at the time of the inventions" is a hypothetical person who is presumed to be familiar with the relevant field and its literature at the time of the inventions. This hypothetical person is also a person of ordinary creativity, capable of understanding the scientific principles applicable to the pertinent field.
- 14. I am told by counsel that persons of ordinary skill in this subject matter or art would have had a bachelor's degree in Computer Science, or related discipline, and two years of relevant experience and knowledge of regulating network access and designing such systems, TCP/IP-based networking as practiced in the Internet, routers, web proxies, web caches, and web servers, and distributed systems and their advantages and management.
- 15. It is my opinion that such a person would have been engaged in academic research, learning though study and practice in the field, through formal instruction, and through the bibliographic resources relevant to his or her research. In the 1980s and 1990s such a person would have had access to a vast array of long-established print resources in the areas of computer science and electrical/computer engineering, as well as to a rich and fast changing set of online

resources providing indexing information, abstracts, and full text services in those same areas.

- 16. *Library catalog records*. Some background on MARC formatted records, OCLC, WorldCat, and OCLC's Connexion is needed to understand the library catalog records discussed in this declaration.
- 17. Libraries world-wide use the MARC format for catalog records; this machine readable format was developed at the Library of Congress in the 1960s.
- 18. MARC formatted records provide a variety of subject access points based on the content of the document being cataloged. All may be found in the MARC Fields 6XX. For example, MARC Field 600 identifies personal names used as subjects and the MARC Field 650 identifies topical terms. A researcher might discover material relevant to his or her topic by a search using the terms employed in the MARC Fields 6XX.
- 19. The MARC Field 040, subfield a, identifies the library or other entity that created the original catalog record for a given document and transcribed it into machine readable form. The MARC Field 008 identifies the date when this first catalog record was entered on the file. This date persists in all subsequent uses of the first catalog record, although newly-created records for the same document, separate from the original record, will show a new date. It is not unusual to find multiple catalog records for the same document.

- 20. WorldCat is the world's largest public online catalog, maintained by the Online Computer Library Center, Inc., or OCLC, and built with the records created by the thousands of libraries that are members of OCLC. WorldCat provides a user-friendly interface for the public to use MARC records; it requires no knowledge of MARC tags and codes. WorldCat records appear in many different catalogs, including the Statewide Illinois Library Catalog. The date a given catalog record was created (corresponding to the MARC Field 008) appears in some detailed WorldCat records as the Date of Entry.
- 21. Whereas WorldCat records are very widely available, the availability of MARC formatted records varies from library to library.
- 22. When an OCLC participating institution acquires a document for which it finds no previously created record in OCLC, or when the institution chooses not to use an existing record, it creates a record for the document using OCLC's Connexion, the bibliographic system used by catalogers to create MARC records. Connexion automatically supplies the date of record creation in the MARC Field 008.
- 23. Once the MARC record is created by a cataloger at an OCLC participating member institution, it becomes available to other OCLC participating members in Connexion and also in WorldCat, where persons interested and

ordinarily skilled in the subject matter or art, exercising reasonable diligence, can locate it.

- 24. When a book has been cataloged, it will normally be made available to readers soon thereafter—normally within a few days or (at most) within a few weeks of cataloging.
- 25. *Indexing*. A researcher may discover material relevant to his or her topic in a variety of ways. One common means of discovery is to search for relevant information in an index of periodical and other publications. Having found relevant material, the researcher will then normally obtain it online, look for it in libraries, or purchase it from the publisher, a bookstore, a document delivery service, or other provider. Sometimes, the date of a document's public accessibility will involve both indexing and library date information. Date information for indexing entries is, however, often unavailable. This is especially true for online indices.
- 26. Indexing services use a wide variety of controlled vocabularies to provide subject access and other means of discovering the content of documents.

  The formats in which these access terms are presented vary from service to service.
- 27. Online indexing services commonly provide bibliographic information, abstracts, and full-text copies of the indexed publications, along with a list of the documents cited in the indexed publication. These services also often

provide lists of publications that cite a given document. A citation of a document is evidence that the document was publicly available and in use by researchers no later than the publication date of the citing document.

- 28. Prominent indexing services include:
- 29. ACM Digital Library. This index is produced by the Association for Computing Machinery, the world's largest scientific and educational computing society. AMC Digital Library contains the full text of all AMC publications, hosted full-text publications from selected publishers, and the ACM Guide to Computing Literature—a comprehensive bibliography of computing literature beginning in the 1950s with more than a million entries. All metadata in the database are freely available on the Web, including abstracts, linked references, citing work, and usage statistics. Full-text articles are available with subscription.

#### IV. OPINIONS REGARDING INDIVIDUAL DOCUMENTS

Document 1. Norbert Pohlmann and Tim Crothers, Firewall Architecture for the Enterprise (New York: Wiley, 2002).

#### Authentication

- 30. Document 1 is a book by Norbert Pohlmann and Tim Crothers published by John Wiley & Sons in 2002.
- 31. Attachment 1a is a true and accurate copy of the book's cover, half title, title page, title page verso, table of contents, and pp.114-135, 149-155, 174-181, and 308-315 from the Library of Congress. Attachment 1b is a true and

accurate copy of the Library of Congress catalog record, in MARC format, for Document 1, showing the book's bibliographic information and, in MARC Field 020, the book's International Standard Book Number (ISBN), 076459926X. Since 1970, each published book has had an ISBN unique to it.

- 32. Attachment 1a is in a condition that creates no suspicion about its authenticity. Specifically, the sequences of pages copied in Attachment 1a are not missing any intermediate pages, the text on each page appears to flow seamlessly from one page to the next, and there are no visible alterations to the document. Attachment 1a was found within the custody of a library a place where, if authentic, it would likely be found.
- 33. Attachment 1c is a true and accurate copy of the ACM Digital Library index record for Document 1, showing the International Standard Book Number (ISBN) 076454926X, identical (except for format) to that provided on the verso of the title page in Attachment 1a: 076459926X.
- 34. I conclude, based on finding Document 1 in a library and on finding library catalog records and an online record for Document 1, that Document 1 is an authentic document and that Attachment 1a is an authentic copy of Document 1.

#### Public accessibility

35. Attachment 1d is a true and accurate copy of a Statewide Illinois
Library catalog record for Document 1, showing this book held by 108 libraries

world-wide. Attachment 1d also indicates that Document 1 was cataloged or indexed in a meaningful way—including being cataloged by subject. The date of entry in Attachment 1d is 27 April 2002, somewhat before the publication of Document 1 (discussed below). This is the same date as in the Attachment 1b, the Library of Congress catalog record for Document 1, where the MARC Field 008 date is 27 April 2002. In my opinion, Document 1 was bibliographically identifiable by 27 April 2002.

- 36. Attachment 1e is a true and accurate copy of the United States

  Copyright Office record for Document 1. It shows the book was published on 3

  July 2002 and registered for copyright on 2 August 2002. I conclude from this copyright record that Document 1 was publicly available from its publisher on or about 3 July 2002.
- 37. Attachment 1f is a true and accurate copy of the Northern Michigan University Library catalog record, in MARC format, for Document 1. In Attachment 1f, the MARC Field 008 indicates this catalog record was created on

<sup>&</sup>lt;sup>1</sup> On the half title page of Attachment 1a, a cop of Document 1 from the Library of Congress, there is a 2 August 2002 date stamp of the Library of Congress Copyright Office.

16 August 2002.<sup>2</sup> Allowing for some time between the cataloging of Document 1 and its arrival on library shelves, where it would be publicly available, I conclude that Document 1 was accessible to the public interested in the art, and that an ordinarily skilled researcher, exercising reasonable diligence, would have had no difficulty finding Document 1 in at least one library by September 2002.

#### Conclusion

38. Based on the evidence presented here—book publication, Copyright Office record, online indexing, and library cataloging—it is my opinion that Document 1 is an authentic document that was bibliographically identifiable by 27 April 2002, was publicly available from its publisher on or about 3 July 2002, was publicly available in at least one library by September 2002.

# Document 2. Andrew S. Tanenbaum, Computer Networks, 3<sup>rd</sup> ed. (Upper Saddle River, NJ: Prentice Hall, 1996).

#### Authentication

- 39. Document 2 is a book by Andrew Tanenbaum published by Prentice Hall in 1996.
- 40. Attachment 2a is a true and accurate copy of the book's fly leaves (one with a circulation slip and circulation stamps), half title, title page, title page

<sup>&</sup>lt;sup>2</sup> MARC Field 040, subfield a, indicates this record was created by GSA, a now obsolete OCLC member code that may have identified the Georgia State Library.

verso, table of contents, and pp.4-8, 50-56, and 408-413 from the University of Illinois at Urbana-Champaign Library. Attachment 2b is a true and accurate copy of that library's catalog record, in MARC format, for Document 2, showing the book's bibliographic information and, in MARC Field 020, the book's International Standard Book Number (ISBN), 0133499456. Since 1970, each published book has had an ISBN unique to it.

- 41. Attachment 2a is in a condition that creates no suspicion about its authenticity. Specifically, the sequences of pages copied in Attachment 2a are not missing any intermediate pages, the text on each page appears to flow seamlessly from one page to the next, and there are no visible alterations to the document. Attachment 2a was found within the custody of a library a place where, if authentic, it would likely be found.
- 42. Attachment 2c is a true and accurate copy of the ACM Digital Library index record for Document 2, showing the International Standard Book Number (ISBN) 0-13-349945-6, identical to that provided on the verso of the title page in Attachment 2a and, except for format, to the ISBN in the Attachment 2b catalog record for Document 2.
- 43. I conclude, based on finding Document 2 in a library and on finding library catalog records and an online record for Document 2, that Document 2 is an authentic document and that Attachment 2a is an authentic copy of Document 2.

#### Public accessibility

- 44. Attachment 2d is a true and accurate copy of a Statewide Illinois
  Library catalog record for Document 2, showing this book held by 710 libraries
  world-wide. Attachment 2d also indicates that Document 2 was cataloged or
  indexed in a meaningful way—including being cataloged by subject. The date of
  entry in Attachment 2d is 2 January 1996, identical to the MARC Field 008 date in
  the Attachment 2b catalog record for Document 2, and somewhat before the
  publication of Document 1 (discussed below). That this is a cataloging-inpublication (CIP) record is indicated by the presence of CIP information on the
  verso of the title page in Attachment 2a In my opinion, Document 2 was
  bibliographically identifiable by 2 January 1996.
- 45. Attachment 2e is a true and accurate copy of the United States

  Copyright Office record for Document 2. It shows the book was published on 6

  March 1996 and registered for copyright on 23 May 1996. I conclude from this

  copyright record that Document 2 was publicly available from its publisher on or

  about 6 March 1996.
- 46. Attachment 2f is a true and accurate copy of a second Statewide
  Illinois Library catalog record for Document 2, showing this book held by an
  additional 9 libraries world-wide. Attachment 2f also indicates that Document 2
  was cataloged or indexed in a meaningful way—including being cataloged by

subject. The date of entry in Attachment 2f is 10 June 1996. Allowing for some time between the cataloging of Document 2 and its arrival on library shelves, where it would be publicly available, I conclude that Document 2 was accessible to the public interested in the art, and that an ordinarily skilled researcher, exercising reasonable diligence, would have had no difficulty finding Document 2 in at least one library by July 1996.

#### Conclusion

47. Based on the evidence presented here—book publication, Copyright
Office record, online indexing, and library cataloging—it is my opinion that
Document 2 is an authentic document that was bibliographically identifiable by 2
January 1996, was publicly available from its publisher on or about 6 March 1996,
was publicly available in at least one library by July 1996.

#### V. ATTACHMENTS

- 48. The attachments attached hereto are true and correct copies of the materials identified above. Helen Sullivan is a Managing Partner in Prior Art Documentation Services LLC (see <a href="http://www.priorartdocumentation.com/hellen-sullivan/">http://www.priorartdocumentation.com/hellen-sullivan/</a>). One of her primary responsibilities in our partnership is to secure the bibliographic documentation used in attachments to our declarations.
- 49. Ms. Sullivan and I work in close collaboration on the bibliographic documentation needed in each declaration. I will sometimes request specific

bibliographic documents or, more rarely, secure them myself. In all cases, I have carefully reviewed the bibliographic documentation used in my declaration. My signature on the declaration indicates my full confidence in the authenticity, accuracy, and reliability of the bibliographic documentation used.

50. Each Attachment has been marked with an identifying label on the top of each page. However, no alterations other than these noted labels appear in these attachments, unless otherwise noted. All attachments were created on 4-10 August 2017 and all URLs referenced in this declaration were available 6 August 2017.

#### VI. CONCLUSION

- 51. I reserve the right to supplement my opinions in the future to respond to any arguments that Patent Owner or its expert(s) may raise and to take into account new information as it becomes available to me.
- 52. I declare that all statements made herein of my knowledge are true, and that all statements made on information and belief are believed to be true, and that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Executed this 10<sup>th</sup> day of August, 2017 in Urbana, Illinois.

Swed Burnet

Scott Bennett

#### Appendix A

### SCOTT BENNETT Yale University Librarian Emeritus

711 South Race Urbana, Illinois 61801-4132 2scottbb@gmail.com 217-367-9896

#### **EMPLOYMENT**

Retired, 2001. Retirement activities include:

- Managing Partner in Prior Art Documentation Services, LLC, 2015-. This firm provides documentation services to patent attorneys; more information is available at <a href="http://www.priorartdocumentation.com">http://www.priorartdocumentation.com</a>
- Consultant on library space design, 2004-. This consulting practice is rooted in a research, publication, and public speaking program conducted since I retired from Yale University in 2001. I have served more than 50 colleges and universities in the United States and abroad with projects ranging in likely cost from under \$50,000 to over \$100 million. More information is available at <a href="http://www.libraryspaceplanning.com/">http://www.libraryspaceplanning.com/</a>
- Senior Advisor for the library program of the Council of Independent Colleges, 2001-2009
- Member of the Wartburg College Library Advisory Board, 2004-
- Visiting Professor, Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign, Fall 2003

University Librarian, Yale University, 1994-2001

Director, The Milton S. Eisenhower Library, **The Johns Hopkins University**, Baltimore, Maryland, 1989-1994

Assistant University Librarian for Collection Management, **Northwestern University**, Evanston, Illinois, 1981-1989

Instructor, Assistant and Associate Professor of Library Administration, University of Illinois at Urbana-Champaign, 1974-1981

Assistant Professor of English, University of Illinois at Urbana-Champaign, 1967-1974

Woodrow Wilson Teaching Intern, St. Paul's College, Lawrenceville, Virginia, 1964-1965

#### **EDUCATION**

University of Illinois, M.S., 1976 (Library Science) Indiana University, M.A., 1966; Ph.D., 1967 (English) Oberlin College, A.B. magna cum laude, 1960 (English)

HONORS AND AWARDS

Morningside College (Sioux City, IA) Doctor of Humane Letters, 2010

**American Council of Learned Societies** Fellowship, 1978-1979; Honorary Visiting Research Fellow, Victorian Studies Centre, **University of Leicester**, 1979; **University of Illinois** Summer Faculty Fellowship, 1969

**Indiana University** Dissertation Year Fellowship and an **Oberlin College** Haskell Fellowship, 1966-1967; **Woodrow Wilson** National Fellow, 1960-1961

#### PROFESSIONAL ACTIVITIES

**American Association for the Advancement of Science**: Project on Intellectual Property and Electronic Publishing in Science, 1999-2001

American Association of University Professors: University of Illinois at Urbana-Champaign Chapter Secretary and President, 1975-1978; Illinois Conference Vice President and President, 1978-1984; national Council, 1982-1985, Committee F, 1982-1986, Assembly of State Conferences Executive Committee, 1983-1986, and Committee H, 1997-2001; Northwestern University Chapter Secretary/Treasurer, 1985-1986

**Association of American Universities**: Member of the Research Libraries Task Force on Intellectual Property Rights in an Electronic Environment, 1993-1994, 1995-1996

**Association of Research Libraries**: Member of the Preservation Committee, 1990-1993; member of the Information Policy Committee, 1993-1995; member of the Working Group on Copyright, 1994-2001; member of the Research Library Leadership and Management Committee, 1999-2001; member of the Board of Directors, 1998-2000

Carnegie Mellon University: Member of the University Libraries Advisory Board, 1994

Center for Research Libraries: Program Committee, 1998-2000

**Johns Hopkins University Press**: Ex-officio member of the Editorial Board, 1990-1994; Codirector of Project Muse, 1994

**Library Administration and Management Association**, Public Relations Section, Friends of the Library Committee, 1977-1978

**Oberlin College**: Member of the Library Visiting Committee, 1990, and of the Steering Committee for the library's capital campaign, 1992-1993; President of the Library Friends, 1992-1993, 2004-2005; member, Friends of the Library Council, 2003-

**Research Society for Victorian Periodicals**: Executive Board, 1971-1983; Co-chairperson of the Executive Committee on Serials Bibliography, 1976-1982; President, 1977-1982

**A Selected Edition of W.D. Howells** (one of several editions sponsored by the MLA Center for Editions of American Authors): Associate Textual Editor, 1965-1970; Center for Editions of American Authors panel of textual experts, 1968-1970

Victorian Studies: Editorial Assistant and Managing Editor, 1962-1964

Wartburg College: member, National Advisory Board for the Vogel Library, 2004-

Some other activities: Member of the **Illinois State Library** Statewide Library and Archival Preservation Advisory Panel; member of the **Illinois State Archives** Advisory Board; member of a committee advising the **Illinois Board of Higher Education** on the cooperative management of research collections; chair of a major collaborative research project conducted by the **Research Libraries Group** with support from Conoco, Inc.; active advisor on behalf of the **Illinois Conference AAUP** to faculty and administrators on academic freedom and tenure matters in northern Illinois.

Delegate to Maryland Governor's Conference on Libraries and Information Service; principal in initiating state-wide preservation planning in Maryland; principal in an effort to widen the use of mass deacidification for the preservation of library materials through cooperative action by the Association of Research Libraries and the Committee on Institutional Cooperation; co-instigator of a campus-wide information service for Johns Hopkins University; initiated efforts with the Enoch Pratt Free Library to provide information services to Baltimore's Empowerment Zones; speaker or panelist on academic publishing, copyright, scholarly communication, national and regional preservation planning, mass deacidification.

Consultant for the University of British Columbia (1995), Princeton University (1996), Modern Language Association, (1995, 1996), Library of Congress (1997), Center for Jewish History (1998, 2000-), National Research Council (1998); Board of Directors for the Digital Library Federation, 1996-2001; accreditation visiting team at Brandeis University (1997); mentor for Northern Exposure to Leadership (1997); instructor and mentor for ARL's Leadership and Career Development Program (1999-2000)

At the **Northwestern University Library**, led in the creation of a preservation department and in the renovation of the renovation, for preservation purposes, of the Deering Library book stacks.

At the **Milton S. Eisenhower Library**, led the refocusing and vitalization of client-centered services; strategic planning and organizational restructuring for the library; building renovation planning. Successfully completed a \$5 million endowment campaign for the humanities collections and launched a \$27 million capital campaign for the library.

At the **Yale University Library**, participated widely in campus-space planning, university budget planning, information technology development, and the promotion of effective teaching and learning; for the library has exercised leadership in space planning and renovation, retrospective conversion of the card catalog, preservation, organizational development, recruitment of minority librarians, intellectual property and copyright issues, scholarly communication, document delivery services among libraries, and instruction in the use of information resources. Oversaw approximately \$70 million of library space renovation and construction. Was co-principal investigator for a grant to plan a digital archive for Elsevier Science.

Numerous to invitations speak at regional, national, and other professional meetings and at alumni meetings. Lectured and presented a series of seminars on library management at the **Yunnan University Library**, 2002. Participated in the 2005 International Roundtable for Library and Information Science sponsored by the **Kanazawa Institute of Technology** Library Center and the Council on Library and Information Resources.

#### **PUBLICATIONS**

- "Putting Learning into Library Planning," portal: Libraries and the Academy, 15, 2 (April 2015), 215-231.
- "How librarians (and others!) love silos: Three stories from the field "available at the Learning Spaces Collaborary Web site, <a href="http://www.pkallsc.org/">http://www.pkallsc.org/</a>
- "Learning Behaviors and Learning Spaces," portal: Libraries and the Academy, 11, 3 (July 2011), 765-789.
- "Libraries and Learning: A History of Paradigm Change," *portal: Libraries and the Academy*, 9, 2 (April 2009), 181-197. Judged as the best article published in the 2009 volume of *portal*.
- "The Information or the Learning Commons: Which Will We Have?" *Journal of Academic Librarianship*, 34 (May 2008), 183-185. One of the ten most-cited articles published in JAL, 2007-2011.
- "Designing for Uncertainty: Three Approaches," *Journal of Academic Librarianship*, 33 (2007), 165–179.
- "Campus Cultures Fostering Information Literacy," *portal: Libraries and the Academy*, 7 (2007), 147-167. Included in Library Instruction Round Table Top Twenty library instruction articles published in 2007
- "Designing for Uncertainty: Three Approaches," *Journal of Academic Librarianship*, 33 (2007), 165–179.
- "First Questions for Designing Higher Education Learning Spaces," *Journal of Academic Librarianship*, 33 (2007), 14-26.
- "The Choice for Learning," Journal of Academic Librarianship, 32 (2006), 3-13.
- With Richard A. O'Connor, "The Power of Place in Learning," *Planning for Higher Education*, 33 (June-August 2005), 28-30
- "Righting the Balance," in *Library as Place: Rethinking Roles, Rethinking Space* (Washington, DC: Council on Library and Information Resources, 2005), pp. 10-24
- Libraries Designed for Learning (Washington, DC: Council on Library and Information Resources, 2003)
- "The Golden Age of Libraries," in *Proceedings of the International Conference on Academic Librarianship in the New Millennium: Roles, Trends, and Global Collaboration*, ed. Haipeng Li (Kunming: Yunnan University Press, 2002), pp. 13-21. This is a slightly different version of the following item.
- "The Golden Age of Libraries," Journal of Academic Librarianship, 24 (2001), 256-258
- "Second Chances. An address . . . at the annual dinner of the Friends of the Oberlin College Library November 13 1999," Friends of the Oberlin College Library, February 2000

- "Authors' Rights," *The Journal of Electronic Publishing* (December 1999), http://www.press.umich.edu/jep/05-02/bennett.html
- "Information-Based Productivity," in *Technology and Scholarly Communication*, ed. Richard Ekman and Richard E. Quandt (Berkeley, 1999), pp. 73-94
- "Just-In-Time Scholarly Monographs: or, Is There a Cavalry Bugle Call for Beleaguered Authors and Publishers?" *The Journal of Electronic Publishing* (September 1998), http://www.press.umich.edu/jep/04-01/bennett.html
- "Re-engineering Scholarly Communication: Thoughts Addressed to Authors," *Scholarly Publishing*, 27 (1996), 185-196
- "The Copyright Challenge: Strengthening the Public Interest in the Digital Age," *Library Journal*, 15 November 1994, pp. 34-37
- "The Management of Intellectual Property," Computers in Libraries, 14 (May 1994), 18-20
- "Repositioning University Presses in Scholarly Communication," *Journal of Scholarly Publishing*, 25 (1994), 243-248. Reprinted in *The Essential JSP. Critical Insights into the World of Scholarly Publishing. Volume 1: University Presses* (Toronto: University of Toronto Press, 2011), pp. 147-153
- "Preservation and the Economic Investment Model," in *Preservation Research and Development.* Round Table Proceedings, September 28-29, 1992, ed. Carrie Beyer (Washington, D.C.: Library of Congress, 1993), pp. 17-18
- "Copyright and Innovation in Electronic Publishing: A Commentary," *Journal of Academic Librarianship*, 19 (1993), 87-91; reprinted in condensed form in *Library Issues: Briefings for Faculty and Administrators*, 14 (September 1993)
- with Nina Matheson, "Scholarly Articles: Valuable Commodities for Universities," *Chronicle of Higher Education*, 27 May 1992, pp. B1-B3
- "Strategies for Increasing [Preservation] Productivity," *Minutes of the [119th] Meeting [of the Association of Research Libraries]* (Washington, D.C., 1992), pp. 39-40
- "Management Issues: The Director's Perspective," and "Cooperative Approaches to Mass Deacidification: Mid-Atlantic Region," in *A Roundtable on Mass Deacidification*, ed. Peter G. Sparks (Washington, D.C.: Association of Research Libraries, 1992), pp. 15-18, 54-55
- "The Boat that Must Stay Afloat: Academic Libraries in Hard Times," *Scholarly Publishing*, 23 (1992), 131-137
- "Buying Time: An Alternative for the Preservation of Library Material," ACLS *Newsletter*, Second Series 3 (Summer, 1991), 10-11
- "The Golden Stain of Time: Preserving Victorian Periodicals" in *Investigating Victorian Journalism*, ed. Laurel Brake, Alex Jones, and Lionel Madden (London: Macmillan, 1990), pp. 166-183

- "Commentary on the Stephens and Haley Papers" in *Coordinating Cooperative Collection Development: A National Perspective*, an issue of *Resource Sharing and Information Networks*, 2 (1985), 199-201
- "The Editorial Character and Readership of *The Penny Magazine*: An Analysis," *Victorian Periodicals Review*, 17 (1984), 127-141
- "Current Initiatives and Issues in Collection Management," *Journal of Academic Librarianship*, 10 (1984), 257-261; reprinted in *Library Lit: The Best of 85*
- "Revolutions in Thought: Serial Publication and the Mass Market for Reading" in *The Victorian Periodical Press: Samplings and Soundings*, ed. Joanne Shattock and Michael Wolff (Leicester: Leicester University Press, 1982), pp. 225-257
- "Victorian Newspaper Advertising: Counting What Counts," Publishing History, 8 (1980), 5-18
- "Library Friends: A Theoretical History" in *Organizing the Library's Support: Donors, Volunteers, Friends*, ed. D.W. Krummel, Allerton Park Institute Number 25 (Urbana: University of Illinois Graduate School of Library Science, 1980), pp. 23-32
- "The Learned Professor: being a brief account of a scholar [Harris Francis Fletcher] who asked for the Moon, and got it," *Non Solus*, 7 (1980), 5-12
- "Prolegomenon to Serials Bibliography: A Report to the [Research] Society [for Victorian Periodicals]," *Victorian Periodicals Review*, 12 (1979), 3-15
- "The Bibliographic Control of Victorian Periodicals" in *Victorian Periodicals: A Guide to Research*, ed. J. Don Vann and Rosemary T. VanArsdel (New York: Modern Language Association, 1978), pp. 21-51
- "John Murray's Family Library and the Cheapening of Books in Early Nineteenth Century Britain," *Studies in Bibliography*, 29 (1976), 139-166. Reprinted in Stephen Colclough and Alexis Weedon, eds., *The History of the Book in the West: 1800-1914*, Vol. 4 (Farnham, Surrey: Ashgate, 2010), pp. 307-334.
- with Robert Carringer, "Dreiser to Sandburg: Three Unpublished Letters," *Library Chronicle*, 40 (1976), 252-256
- "David Douglas and the British Publication of W. D. Howells' Works," *Studies in Bibliography*, 25 (1972), 107-124
- as primary editor, W. D. Howells, *Indian Summer* (Bloomington: Indiana University Press, 1971)
- "The Profession of Authorship: Some Problems for Descriptive Bibliography" in *Research Methods in Librarianship: Historical and Bibliographic Methods in Library Research*, ed. Rolland E. Stevens (Urbana: University of Illinois Graduate School of Library Science, 1971), pp. 74-85
- edited with Ronald Gottesman, *Art and Error: Modern Textual Editing* (Bloomington: Indiana University Press, 1970)--also published in London by Methuen, 1970

"Catholic Emancipation, the *Quarterly Review*, and Britain's Constitutional Revolution," *Victorian Studies*, 12 (1969), 283-304

as textual editor, W. D. Howells, *The Altrurian Romances* (Bloomington: Indiana University Press, 1968); introduction and annotation by Clara and Rudolf Kirk

as associate textual editor, W. D. Howells, *Their Wedding Journey* (Bloomington: Indiana University Press, 1968); introduction by John Reeves

"A Concealed Printing in W. D. Howells," *Papers of the Bibliographic Society of America*, 61 (1967), 56-60

editor, Non Solus, A Publication of the University of Illinois Library Friends, 1974-1981

editor, Robert B. Downs Publication Fund, University of Illinois Library, 1975-1981

Reviews, short articles, etc. in Victorian Studies, Journal of English and German Philology, Victorian Periodicals Newsletter, Collection Management, Nineteenth-Century Literature, College & Research Libraries, Scholarly Publishing Today, ARL Newsletter, Serials Review, Library Issues, S[ociety for] S[cholarly] P[ublishing] Newsletter, and Victorian Britain: An Encyclopedia



WTS Number: 167943

Order #133748

**Standard** 

**Estimated Delivery:** 

08/07/2017 11:21am

Maxcost:

ARTICLE DELIVERY

**Request Date:** 

08/03/2017 11:21am

Requester:

Helen Sullivan - Prior Art **Documentation Services LLC** 

Phone:

2174465370

Reference:

17803

**Delivery email:** 

helen@priorartdocs.com

Instructions:

Citation:

Norbert Pohlmann et al., Firewall Architecture for the Enterprise, Wiley Publishing, Inc. 2002, p.114-135, 149-155, 174-181, 308-315

I am aware this is not in your collection but hope you might be able to obtain via ILL. Please ask that this be copied from print only and that the cover, title page, verso, table of contents and the pages cited above are all copied. Any library date stamps are useful though I thought it unlikely since this is a monograph. Thanks so much for your excellent service!

Library/Supplier:

ART-LOC 11257

ISSN/ISBN/OCLC: 076454926X

Base Service Fee: \$15

Rush fee: \$0.00

Copyright: WTS will pay copyright royalties and bill me:\$\_\_\_\_

Supplier: WTS staff may obtain materials from outside UW - Madison: \$\_\_\_\_\_

Shipping: \$\_\_\_\_ Research: \$\_\_\_\_

THIS IS NOT AN INVOICE

An outreach service of the Wendt Commons Library, University of Wisconsin - Madison http://wts.wisc.edu | wts@engr.wisc.edu | 608.262.5917



# Firewall Architecture for the Enterprise

Norbert Pohlmann Tim Crothers



# Firewall Architecture for the Enterprise

# Firewall Architecture for the Enterprise

Norbert Pohlmann and Tim Crothers



Best-Selling Books • Digital Downloads • e-Books • Answer Networks e-Newsletters • Branded Web Sites • e-Learning Firewall Architecture for the Enterprise

Published by Wiley Publishing, Inc. 909 Third Avenue

New York, NY 10022

www.wiley.com

Copyright © 2002 by Wiley Publishing, Inc., Indianapolis, Indiana

Library of Congress Control Number: 2002102445

ISBN: 0-7645-4926-X

Manufactured in the United States of America

10987654321

10/0T/QX/QS/IN

Published by Wiley Publishing, Inc., Indianapolis, Indiana

Published simultaneously in Canada

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, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4744. Requests to the Publisher for permission should be addressed to the Legal Department, Wiley Publishing, Inc., 10475 Crosspoint Blvd., Indianapolis, IN 46256, (317) 572-3447, fax (317) 572-4447, E-Mail: permcoordinator@wiley.com.

LIMIT OF LIABILITY/DISCLAIMER OF WARRANTY: WHILE THE PUBLISHER AND AUTHOR HAVE USED THEIR BEST EFFORTS IN PREPARING THIS BOOK, THEY MAKE NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE ACCURACY OR COMPLETENESS OF THE CONTENTS OF THIS BOOK AND SPECIFICALLY DISCLAIM ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. NO WARRANTY MAY BE CREATED OR EXTENDED BY SALES REPRESENTATIVES OR WRITTEN SALES MATERIALS. THE ADVICE AND STRATEGIES CONTAINED HEREIN MAY NOT BE SUITABLE FOR YOUR SITUATION. YOU SHOULD CONSULT WITH A PROFESSIONAL WHERE APPROPRIATE. NEITHER THE PUBLISHER NOR AUTHOR SHALL BE LIABLE FOR ANY LOSS OF PROFIT OR ANY OTHER COMMERCIAL DAMAGES, INCLUDING BUT NOT LIMITED TO SPECIAL, INCIDENTAL, CONSEQUENTIAL, OR OTHER DAMAGES.

For general information on our other products and services or to obtain technical support, please contact our Customer Care Department within the U.S. at 800-762-2974, outside the U.S. at 317-572-3993 or fax 317-572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Trademarks: Wiley and the Wiley Publishing logo are trademarks or registered trademarks of Wiley Publishing, Inc., in the United States and other countries, and may not be used without written permission. All other trademarks are the property of their respective owners. Wiley is not associated with any product or vendor mentioned in this book.

Wiley Publishing, Inc. is a trademark of Wiley Publishing, Inc.

2002-106044

## **Contents**

	Preface
Chapter 1	Business Transformation, IT Security, and
	Introduction to the Firewall
	Developments in Information Technology and IT Security 1
	The Internet Revolution
	Characteristics of the Internet as the Global Data Network3
	Electronic Services4
	The Internet as a World-Wide Source of Information 5
	Dangers from the Internet
	General Threats from the Internet
	The Need for IT Security8
	Why is Security So Important in Information Technology? 8
	What Role Does IT Security Play in the Information Society? 9
	Global Expansion and Changing Business Processes 10
	Industrial and Economic Espionage
	Economic Liability
	IT Security in Context
	Opportunities and Risk12
	Analogies to Firewall Systems
	Firewall14
	Security Guard14
	Firewalls are Not Automatic Protection
	Purpose of a Central Firewall System16
	Advantages of the Common Point of Trust Method 17
	General Objectives of Firewall Systems
	Additional Objectives of a Firewall System
Chapter 2	TCP/IP Technology for the Internet and Intranet 21
	Advantages of TCP/IP Technology
	The OSI Reference Model
	TCP/IP Protocol Architecture24
	Internet Addresses
	The Communication Protocols
	Internet Protocol
	Routing Protocols31
	Internet Control Message Protocol
	Port Numbers35
	User Datagram Protocol
	Transmission Control Protocol38
	Domain Name Service

#### xiv Contents

	Telnet
	File Transfer Protocol41
	Simple Mail Transport Protocol
	Hypertext Transfer Protocol
	Network News Transfer Protocol
	Other Common Network Protocols
Chantor 2	Threats in Networks
Chapter 3	Attack Possibilities in Communication Systems
	Passive Attacks
	Special Dangers Related to the Use of Local Area Networks 54
	Log of a Telnet Session
	Active Attacks
	Opportunities for Accidental Harm
	Other Aspects of Potential Threats
	in Communication over the Internet
	Communication Paths of IP Packets in the Internet 61
	Attack Tools from the Internet
	Implementation Errors in Applications and
	Incorrect Configurations
	How High Is the Risk?62
	Damage Categories and the Consequences of Damage 64
	Violation of Laws, Regulations, or Contracts 64
	Physical Injury65
	Impaired Performance of Duties65
	Negative Effects on External Relationships
	Financial Consequences67
	Methods of Attack and Principle Countermeasures
	Based on the TCP/IP Protocols67
	Idea Behind an Attack
	Analysis of the Network Using Scanner Programs
	Password Snooping and IP Masquerade71
	Exploitation of Incorrect Configuration74
	Hopping
	Exploitation of Implementation Errors in Applications
	such as Internet Information Server
	IP Address Spoofing78
	ICMP Attacks80
	Internet Routing Attacks
	Results of the 2001 CSI/FBI Crime and Security Survey 84
Chapter 4	Elements of a Firewall System 87
•	Active Firewall Elements
	Architecture of Active Firewall Elements
	Design Concept for an Active Firewall Element90

Contents Checks in the Transport Layer .......96 Strategies for Setting Up and Evaluating the Filter Rules ..... 100 Example of the Use of a Packet Filter ......101 Dynamic Packet Filters ......104 Security-Relevant Information in a Packet Filter . . . . . . . . . . 109 When to Use Application Gateways .......149 Capabilities and Limitations of an Adaptive Proxy ........... 151 Firewall Elements and the Speed-Versus-Security Security Management for Active Firewall Elements . . . . . 152 Requirements for a Security Management Module ........... 152 Coupling to a Network Management System (NMS) .......... 154 Concepts of Firewall Systems ................................ 157 Application Gateways ......159 Evaluation of Application Gateways . . . . . . . . . . . . . . . . . . 160 Combination of Firewall Elements . . . . . . . . . . . . . . . . . . 160

 Dual-Homed Application Gateway
 160

 Single-Homed Application Gateway
 160

 Packet Filter Plus Single-Homed Application Gateway
 161

 Packet Filter Plus Dual-Homed Application Gateway
 163

 The Screened Subnet
 165

Application Gateway ...... 166

Application Gateway (High-Level Security Firewall System)... 168

Two Packet Filters as Screened Subnet and Single-Homed

Two Packet Filters as Screened Subnet and Dual-Homed

Chapter 5

 $\chi v$ 

#### xvi Contents

	Operational Environment of High-Level Security
	Firewall Systems
	Capabilities of a High-Level Security Firewall System 171
	Internet Server
	Intranet Servers
	Several Application Gateways in Parallel 176
	The Right Firewall Concept for Every Application 179
	Mail Security
•	Configuration in the Protected Network
	DNS Security
	The Application Gateway Queries the Internal DNS Server 184
	Split DNS
01	Dedicated DNS Server or DNS Proxy
Chapter 6	Firewall Systems and Encryption 187
	Security Mechanisms for Encryption and
	Digital Signatures 187
	Private Key Systems
	Public Key Systems
	One-Way Hash Function190
	Hybrid Encryption Technology191
	Certification Systems
	Smart Card195
	E-Mail Security
	Envelope and Signature for E-Mail
	E-Mail Security System Services
	E-Mail Security from the User's Point-of-View 202
	Object Encryption and Firewall Systems
	Electronic Contracts
	Other Object-Oriented Security Concepts203
	Success Is Determined by the Market
	Virtual Private Networks208
	Security System as a Transparent Solution
	Black Box Solution
	Security Sublayer in the Client Device:
	End-to-End Encryption
	Security in LAN Segments
	Linking LAN Segments with a Security Bridge
	Linking LAN Segments over Public Networks 214
	Formation of Cryptographically Protected Logical
	Networks (VPNs)
	Security Client
	Applications
	Transparent Encryption and Firewall Systems 221
	Comparison of E-Mail Security and VPN Security 221

		Contents
Chapter 7	Authentication Procedures	
	Identification and Authentication	
	General Authentication Procedures	
	Password Method224	
	One-Time Passwords	
	Challenge-Response Procedures226	
	Biometric Authentication	
	Extensible Authentication	
	Authentication Procedure for Firewall Systems 227	
	S/Key	
	Authentication Procedure with a Security Token 230	
	Signature Card	
	Authentication Procedure Using Mobile Phone	
	Authentication Procedure Using Kerberos	
Chapter 8	Evaluating Firewall Solutions	
	Public Domain Software or Firewall Product? 239	
	Public Domain (and Open Source) Software for	
	the Firewall Application	
	Firewall Products	
	Software Solution or Turn-Key Solution? 242	
	What Does a Pure Software Solution Offer?242	
	What Is a Turn-Key Solution?	
	Criteria for Assessment of the Actual Security	
	Achieved by a Firewall Product	
	Open and Transparent Security	
	Tested, Demonstrable Security	
	Security without Governmental Restrictions	
Chapter 9	Practical Use of Firewall Systems249	
	Secure Connection of the Organization's Intranet	
	to the Internet	
	Internet Server253	
	Intranet Security255	
	Partitioning Organizational Units257	
	Scalable Security	
	Remote Access (Telecommuting, Mobile Workstations) 258	
	Connection of Special Organizational Units 259	
	External Modem Connections	
	Modem Connections from the Insecure Network	
	to the Protected Network	
	Anti-virus and Anti-malware System	
	General Problem Regarding the Detection of Viruses263	
	Compression of Files	
	Encryption of Files	
	Integration of Virus Scanners at the Common Point of Trust 264	
	Additional Technical, Personnel-related,	
	and Organizational Measures	

xvii

#### xviii Contents

	Intrusion Detection Systems	270
	Personal Firewall	
	Practical Implementation	273
	Central Administration	
	Application of Firewall Components	. 275
Chapter 10	Firewall Security Policy	. 283
•	Firewall Security Policy	
	Security Objectives	
	Description of the Resources to Be Protected	285
	Definition of Communications Requirements	285
	Definition of Services and Applications	286
	Additional Security Measures	
	Infrastructural Measures	289
	Organizational Measures	290
,	Personnel-Related Measures	
/	Contingency Plans	298
	Conceptual Limitations of a Central Firewall System	299
	Back Doors	299
	Internal Attacks	299
	Attacks at the Data Level	299
	Correct Security Policy and Implementation	300
	The Man-in-the-Middle Attack	300
	Security versus Connectivity - Risk versus Opportunity	302
	Trustworthiness of the Communications Partner	
	and the Received Data	303
	Practical Security	303
Chapter 11	Special Issues Related to Firewall Systems	. 307
	Network Address Translation	307
	Firewall Systems and Network Address Translation	309
	Problems for Networks That Work with Illegal IP Addresses	310
	Network Address Translation Problems	310
	Domain Names	311
	Administration of Several Firewall Systems via a	
	Security Management Module	
	Nested Firewall Configurations	312
	Availability	313
	Performance	
Chapter 12	Secondary Issues Related to Firewall Systems	
	Logbooks Burden or Benefit?	317
	Purpose of Logging	
	Logging Events	
	Alarm Function	
	Preserving Evidence	
	Protecting the Logged Data	
	Reactions to a Security Breach	
	Data Protection Issues	224

Java and Its Relations	324
Basics	32
The Common Gateway Interface	32
Active Server Pages	32
PHP	32
Java	32
JavaScript	33
ActiveX	33
Microsoft .NET and Web Services	33
Firewall Systems from a Business Perspective	. 332
The Procurement Phase	33
The Installation Phase	33
The Maintenance Phase	33
Summary of Costs	33
Cost-Benefit Analysis of Firewall Systems	339
Evaluation and Certification of Firewall Systems	34
ITSEC Certification	34
ICSA Certification	35
Further Development of Comprehensive	
Firewall Systems	. 356
Increasing Innovations	350
Integrative, Central Security Management of All	
Security Mechanisms	350
Ever Greater Speed Combined with an Ever Higher	
Protection Requirement	356
New Proxies for New Services	
Universal Identification and Authentication Procedures	357
Uniform Representation of Attacks and Security Services and	
Mechanisms	
Log Analysis	
Security Audits	. 357
Security Audits	
Components of a Security Audit	
Data Collection Tools	
Additional Sources of Information	
Intrusion Detection/Response Systems	
Analogy to Video Surveillance	367
Difference between Intrusion Detection	
and Intrusion Response	
Integration into Firewall Systems	
Primary Task of the Intrusion Detection System	
Design and Operation of Intrusion Detection Systems $\dots$	
Analysis Concepts	
Limitations of IDS	375

#### xx Contents

	Distributed Denial-of-Service Attacks —
	Description and Evaluation
	General Mode of Operation of DDoS Attacks 377
	Personal Firewall379
	Central Firewall System380
	Aim of a Personal Firewall
	Personal Firewall Components
	Personal Firewall Security Components
	Secure Environment for Digital Signatures 385
	Display, Logging, and Statistics about
	Security-Relevant Events
Chapter 13	Theoretical Foundations of Firewall Systems 387
_	The Communications Model387
	Layers in the TCP/IP Protocol Architecture
	Simplified Logical Communications Model
	Transmitters
	Receivers
	Protocol Elements
	Actions
	Communication Sequences
	Sequence of Actions at the Receiver's End
	Attacks from the Network
	Firewall Elements
	Integration and Enforcement Module
	Analysis Module $\rightarrow$ analysis(x <sub>i</sub> )401
	Decision Module
	Ruleset→Security-Management (rules) 401
	The Communications Model with Integrated
	Firewall System
	Functions for Action Selection at the Receiver's End for $\boldsymbol{r}_n$ 402
	Attacks on the Firewall System
	Basic Factors That Affect the Selection and Execution of
	Actions at the Receiver's End
	Defects That Arise from Network Attacks 406
	Sources of Defects in the Communications Solution
	in Use at the Receiver's End
	Sources of Defects in the Firewall System
	A Firewall System's Security Services
	Factors Influencing the Security and Trustworthiness
	of Firewall Systems409
	Effectiveness of Security Services
	The Effect of the Operational Environment
	The Effect of a Trustworthy Implementation
	of a Firewall System
	The Effect of Influencing Factors within the Organization 414

xxi

	The Effectiveness of Different Firewall Design Concepts 415
	The Effect of Additional Security Mechanisms
	Summary of the Effect of a Comprehensive Firewall System 416
	Security Mechanisms and Their Effectiveness 418
	•
	High-Level Security Firewall Systems
	Encryption
	Anti-Malware Systems
	Intrusion Detection Systems
	Personal Firewalls
	Non-Technical Security Measures419
	Trustworthiness, Audits, Security Policy,
	and Secure Operation
	Attacks and the Effectiveness of Different
	Security Mechanisms
Appendix A	Security Standards425
	SSL426
	Assessment
	SSL Key Management
	Secure Shell
	Assessment
	IPv6 430
	Assessment
	Internet Protocol Security Architecture
	The Authentication Header431
	Encapsulated Security Payload433
	Security Association
	IPSec Databases434
	Key Management under IPSec
Appendix B	References
Appendix C	Glossary
Appendix D	Sample Security Policy451
• •	Security Policy
	Computer Usage Guidelines
	Escalation Procedures for Security Incidents
	Indo-

extended state-oriented packet filters as *stateful inspection*, *smart filtering*, or *adaptive screening*. With this extended functionality, they are often offered as user-oriented packet filters. Figure 4-15 illustrates state-oriented packet filters.

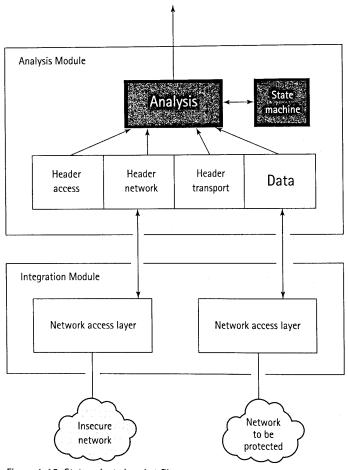


Figure 4-15: State-oriented packet filters

State-oriented packet filters have the same advantages as packet filters, but they can also check the applications. Some risks remain because the services are not directly isolated from each other.

Because it is a complex matter to simultaneously hold and interpret the communications data in the different communications layers, state-oriented packet filters generally have a shallower depth of analysis, or they are particularly prone to

This material may be protected by Copyright law (Title 17 U.S. Code)

errors connected with their very powerful software. Basically, you can't test the complex software of state-oriented packet filters sufficiently or comprehensively to prove that errors cannot occur in any operating state. For this reason, one must continue to assume that the complex programs contain potential security risks that could be used to perpetrate an attack.

A better and more secure way to analyze the application data is to use application gateways with proxies. This approach is described in the next section.

## **Network Address Translation**

As its name implies, NAT (Network Address Translation) works by using one set of addresses for communications on the Internet and a separate set of addresses for communications on the internal network. To fully support this translation, the IANA set aside three ranges of IP addresses in RFC 1918:

- ◆ 10.0.0.0 through 10.255.255.255 (10.0.0.0/8)
- ◆ 172.16.0.0 through 172.31.255.255 (172.16.0.0/12)
- ◆ 192.168.0.0 through 192.168.255.255 (192.168.0.0/16)

These addresses are reserved for internal use only and, as a consequence, are nonroutable on the Internet. Attempts to communicate with any of these ranges through the Internet result in ICMP "network unreachable" errors.

An organization implementing NAT uses one of the preceding ranges for their internal network addressing. The external interface of the firewall is assigned a normal routable IP address. When the firewall transmits a packet from the internal network to the Internet, it actually creates a new packet destined for the same address but originating from the external address(es) of the firewall. This packet is then transmitted to the destination. The firewall keeps a table of current communications so that when the return communications reach the firewall, they are taken and placed into a new packet destined for the internal computer and transmitted internally.

The NAT process affords a substantial degree of security. Since all direct communications are prevented (as long as the systems are properly configured), an external attacker is forced to compromise the firewall or find a means of passing his communications through it successfully rather than attacking the internal host directly. Given the firewall protection module and hardening, this is a relatively significant challenge.

Not inconsequentially, NAT prolongs the life expectancy of IPv4 on the Internet. Were it not for address translation, the supply of Internet addresses would have been exhausted long ago. Using NAT, a company with hundreds of internal computers can communicate fully with the Internet using only a handful of routable addresses.

In addition to being used dynamically, NAT is also used in a static translation mode. *Static translation NAT* uses a one-to-one correspondence of an external routable address to an internal nonroutable address. It is used for application servers such as Web, e-mail, DNS, and FTP where externally originated communications are necessary. By using NAT in these circumstances, the firewall is still required to be a bridge between the internal server and the Internet, thus requiring the attacker to deal with the firewall in addition to the host security of the protected server.



Some of the challenges posed by NAT are covered in Chapter 11.

# **Application Gateways and Proxies**

This section describes how the application gateway firewall element works. The distinguishing feature of the application gateway is that it separates the two networks logically and physically, as shown in Figure 4-16.

In some firewall configurations, the application gateway is the only computer system that can be accessed from the insecure network (for example, the Internet); thus, the application gateway requires particular protection. For this reason, the computer system on which the application gateway is implemented is also referred to as a *bastion host*.

The application gateway—implemented as a dual-homed gateway—has two network interfaces, one in the network to be protected and the other in the insecure network. The term "dual-homed" refers to the fact that the application gateway has complete control over the packets that are to be passed between the insecure network and the network to be protected.

The application gateway can also be operated as single-homed — in other words, operated with only one network interface. In this case, however, an attacker could bypass the application gateway.

#### ANALOGY TO THE SECURITY GUARD

The "application gateway security guard" does not just inspect the addresses of inbound deliveries. He also opens every packet, examines its contents, and checks the shipping documents prepared by the originator against a clearly defined set of evaluation criteria. After he has completed his detailed security check, the security guard signs the delivery note and sends the truck on its way again. This time, however, he arranges for a trustworthy driver from his own company to take the packets to the actual recipient. The security check at this point is significantly more reliable than just packet filtering, and the driver of the outside company does not see the company premises. It is true that the check takes longer, but as a result, any activities that threaten security can be ruled out.

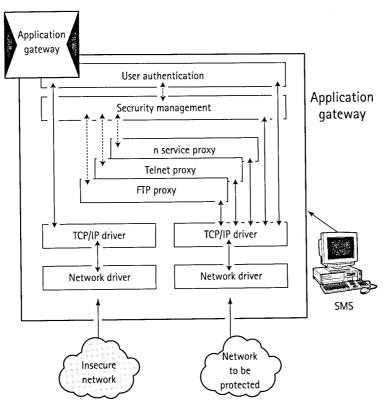


Figure 4-16: Application gateway

## How Application Gateways Work

A user who wants to communicate through the application gateway must first identify himself and undergo authentication. Application gateways generally offer different authentication procedures.

To authenticate, the user first establishes a connection with the application gateway. The direct communication partner is not the destination computer system but the application gateway. However, once identification and authentication are complete, the application gateway is transparent so that the user has the impression of working directly on the destination computer system.

#### **BASIC APPROACH**

The application gateway receives the packets via the network access and TCP/IP drivers on the appropriate ports. If only one service is to be allowed on a given port, software must be available on the application gateway that will transfer the packet

that corresponds to that service from the network on one side of the application gateway to the network on the other side, and vice versa. Such software, which only allows packet transmissions for one particular service (FTP, HTTP, Telnet, and so on) through the application gateway, is known as a *proxy*. The term *proxy* is used because, as far as the user accessing the facilities is concerned, he is communicating with the actual server process of the service on the destination computer system.

Each proxy on the application gateway can offer additional security services tailored to the service for which it is responsible. Because each proxy specializes in one service, the scope of the security and logging functions that are possible on the application gateway is greater. A particularly thorough analysis is possible in this communication layer, as the context of the application data is clearly defined for the relevant service. The proxies concentrate on what is essential. The advantage is that small, straightforward modules are used, so that the susceptibility to implementation errors is reduced (see Figure 4-17).



Re-encryption or re-coding can be performed in the proxy.

#### SECURITY CONCEPT OF AN APPLICATION GATEWAY

For every service that is to be used over the application gateway, a special proxy must be provided. If certain services are to be barred completely, you should have no proxies for those services on the application gateway, nor should any other software be present that would enable them to run.

Thus, as little software as possible should be installed on the application gateway to avoid the possibility that, either by mistake or deliberate introduction by an attacker, some other software can adopt the role of proxy (packet transmission in the application gateway) for a service that ought not to be allowed.

Security Management is intended to make work as easy as possible for the user and is therefore supplied with powerful software (X-Terminal, database, and so on). However, for purposes of security, it should not be run on the same computer system (or at least not at the same time) as the application gateway.

To prevent the possibility that the proxies will be bypassed, application gateways should, for security reasons, have no routing functionality.

Since the application gateway is linked during communication to both the computer system of the insecure network and to the computer system of the protected network, the application gateway provides Network Address Translation. The application gateway has an IP address in the insecure network (for example, an official Internet IP address such as 194.173.3.1) and an IP address in the network to be protected (for example, a private IP address such as 192.168.1.60 that is reserved for this purpose). During communication with computer systems in the insecure network, the application gateway uses the IP addresses of the insecure network; during

communication with computer systems on the network to be protected, it uses the IP addresses of that particular network.

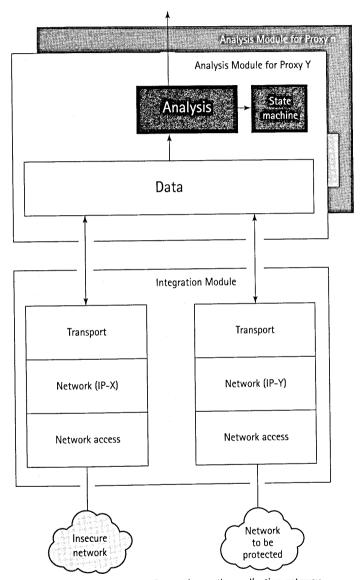


Figure 4-17: Analysis Modules for proxies on the application gateway

#### The Proxies

A distinction can be made between application level proxies and circuit level proxies in their implementation.

#### APPLICATION LEVEL PROXIES

Application level proxies are implemented for particular services and/or applications. In other words, they know the commands of the particular application protocol involved and can analyze and monitor them. Application level proxies work with the standard client software for FTP or Telnet (no modifications are necessary) or with browsers. However, the procedure followed for user-oriented services may differ from the one that is usually followed. For example, identification and authentication with the application level proxy is initially required before transparent communication is available to the user.

In the next section, some application level proxies are described in detail and illustrated with particular implementation types, and the basic ideas behind proxy technology are presented. Some proxies function according to the store-and-forward principle (SMTP), while others are interactive and user-oriented (Telnet, FTP, and HTTP).

#### SMTP PROXY

An SMTP proxy works according to the *store-and-forward principle*. Under this principle, the SMTP proxy accepts the mail in its entirety, stores it temporarily, and then forwards it. No end-to-end link is required between the actual transmitter and the receiver.

#### ANALOGY WITH A COMPANY'S MAILBOX (MAIL PROXY)

A mail proxy can be compared to a company's mailbox. If an employee wants to send a letter to another person in the company, he can put it directly or indirectly into the company's mailbox. The letters go to the internal mail room and are distributed by a messenger who works for the organization. In this way, the external mailman does not need to enter the building, and hence presents no risk. The opening to the outside is a potential area exposed to attack.

With SMTP proxies, solutions work either with or without a Message Transfer Agent (MTA) on the same system. Figure 4-18 shows an SMTP proxy with an MTA available.

#### DESCRIPTION

The SMTP proxy does not work in a user-oriented fashion, so no user authentication is required.

An SMTP proxy on port 25 receives inbound mail and, after the originator (IP address and computer name of the mail server) has been checked, it is stored on the application gateway in a special directory. The SMTP daemon checks periodically to see whether any new mail has arrived. The Mail Transfer Agent (MTA) delivers the mail to the addressee either directly or through one or more other MTAs. The SMTP proxy thus prevents direct access to the internal MTA from the insecure network.

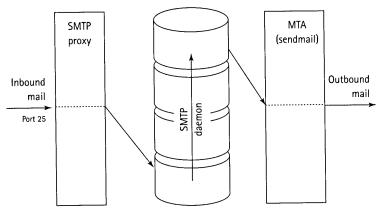


Figure 4-18: SMTP proxy

One example of a commonly used MTA is Sendmail, which is known to exhibit a number of security weaknesses and implementation errors.

An SMTP proxy only processes the following commands, which are security-neutral:

- ◆ It processes HELO, MAIL, RCPT, DATA, QUIT, RSET, NOOP.
- Some additional commands are supported with standard replies in order to make communication possible: HELP, VRFY, and EXPN.

Security-relevant commands, such as DEBUG, can send a spontaneous message to Security Management. If the DEBUG command is detected in an SMTP proxy, no resulting error can occur because the SMTP proxy does not react to it. However, if an outsider attempts to execute a DEBUG command, the fact can be interpreted to mean that it is concealing an attempted attack. This information on an attempted attack can be important.

With the store-and-forward principle, it is possible to neutralize the complex and error-prone Sendmail program (MTA). In this way, attacks known to exploit the shortcomings of Sendmail are prevented, as it is possible to prevent Sendmail from being addressed directly and ensure that only the substitute software of the SMTP proxies can be accessed. The SMTP proxy is straightforward and thus the software is easy to test.

#### LOGBOOK

With an SMTP proxy or MTA, the following log data can be recorded in the application gateway's logbook:

- ◆ IP address and name of the source computer system
- ◆ Time and date of connection setup

- Originator of the mail (as specified in the mail header)
- Addressee of the mail (as specified in the mail header)
- Number of bytes transmitted
- ◆ Time and date of disconnection

If a problem arises, the extensive log data covering events in the SMTP proxy can be used to resolve it.

## USER-ORIENTED APPLICATION-LEVEL PROXIES

The proxies for Telnet, FTP, and HTTP are user-oriented proxies that enforce authentication of the user concerned, analogous to a security guard. Assuming the user is successfully identified and authenticated, this authentication is only good for that particular proxy. If the user wants to use another service (that is, a different proxy), he must undergo identification and authentication again. The advantage of user-oriented proxies is that the user and IP address are, without exception, correctly assigned to the desired service.

# PROCESS OF COMMUNICATIONS THROUGH AN APPLICATION PROXY

The following example presents a connection setup over the application gateway with the aid of a simple password procedure for user-oriented services (refer to Figure 4-19).

- ◆ Phase 1: Establishing a Connection to the Application Gateway. The user attempts to establish a connection from his source computer system to a desired destination computer system over the application gateway. The application gateway accepts the connection setup and asks the person seeking access to undergo identification and authentication.
- Phase 2: User Authentication. The user enters his user identification and the destination computer system (name or address). A check is performed on the application gateway as to whether the user is allowed to access the desired destination computer system from his source computer system and what restrictions apply to such access. In this case, the user is then asked to enter his password. A check is performed on the application gateway as to whether the user has entered the correct password (as with the security guard).

Authentication in firewall systems can be implemented in a number of ways. For example, one can use a standard password procedure, a one-time password, or Challenge/Response. Authentication procedures that make use of cryptographic algorithms require the user to have a security token or smart card. The particular authentication procedure that is used generally depends on the protection requirement and the direction of communication through the firewall system. Communication through the

firewall system can be implemented from a network to be protected to an insecure network with a simple authentication procedure, or even without one. Where communication is directed from an insecure network to a network to be protected, encryption (involving a security token or smart card) should always be used.

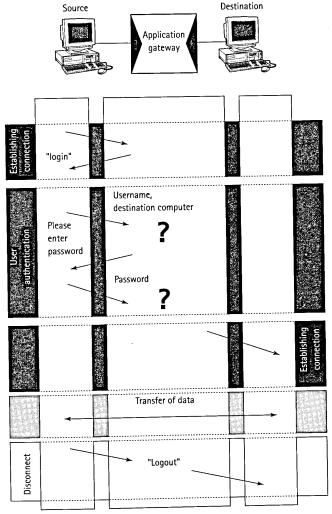


Figure 4-19: Communication over an application-level proxy

- ◆ Phase 3: Establishing a Connection to the Destination Computer. After the user seeking access has been successfully identified and authenticated, a second connection is established from the application gateway through the proxy to the desired and permitted destination computer system.
- ◆ Phase 4: Data Transfer. Data transfer occurs. Depending on the proxy concerned, the transfer of data through the proxy is monitored, controlled, and logged on the application gateway. This phase is transparent to the user.
- Phase 5: Disconnection. In the final phase, the connection through the application gateway is terminated.

## **Telnet Proxy**

The Telnet proxy is responsible for controlled communications using Telnet. It provides appropriate special security-enforcing functions for this service. A connection is established from the source computer system (client) to port 23 of the application gateway (the port for the Telnet service).

The Telnet proxy takes over the connection on port 23. The user on the source computer system identifies and authenticates himself, informing the Telnet proxy of the connection destination. Once identification and authentication have been successfully completed, a user profile containing entries that correspond to the following information is activated:

- IP address of the source computer system that wants to establish the connection
- User name that was used during identification and authentication
- ◆ IP address of the destination computer system

The Telnet proxy establishes a second connection from the application gateway to port 23 of the destination computer system. The user can now use the Telnet service of the destination computer system from the source computer system via the Telnet proxy (see Figure 4-20).

#### CONTROL MONITOR

During the Telnet session, a *control monitor* can check whether or not the user is accessing the permitted destination computer system from the source computer system, or a different computer system without permission. The monitor must check the data stream for byte sequences that could possibly be used to hop to a different computer. It can also look for other information, such as control characters, that are not supposed to be used (for example, Ctrl+C).

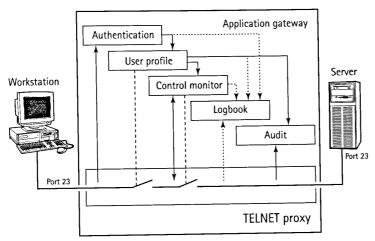


Figure 4-20: Telnet proxy

#### LOGBOOK

The Telnet proxy can write the following protocol entries to the logbook of the application gateway:

- ◆ IP address and name of the source computer system
- ◆ IP address and name of the destination computer system
- Time and date of connection setup
- Name of the user
- ◆ Number of bytes transmitted
- ◆ Time and date of disconnection

With a Telnet connection, it is often appropriate to make a recording of the complete communication for audit purposes. This security-enforcing function not only permits subsequent analysis of the recording, but also exercises a warning effect that should not be underestimated.

## EXAMPLE ILLUSTRATING THE USE OF THE AUDIT TRAIL

The audit security mechanism can be agreed upon in the service contract with a company, perhaps in a case where remote maintenance is being provided. With this mechanism in place, the employee who performs the maintenance knows that everything he does will be recorded and will thus be motivated to perform only those actions that are required to complete the job. Should anything untoward occur, the log can be used to identify any impermissible or unnecessary actions that were carried out via remote access. In other words, the employee's actions can be reliably determined after the event.

# EXAMPLE OF THE USE OF AN APPLICATION GATEWAY WITH TELNET PROXY

Security can be implemented between two networks that have different protection requirements, using an application gateway with a Telnet proxy. Figure 4-21 shows two networks: Network X with the IP address 192.168.3.X, and Network Y with the IP address 192.168.5.Y. These two networks are independent of each other; there is no direct connection between them. The administrator of Server 1 in Network X is working at Workstation A in Network Y and wants to have remote access to Server 1 in Network X. The two networks will be connected using an application gateway with Telnet proxy in such a manner that only a Telnet session from Workstation A in Network Y to Server 1 in Network X is permitted.

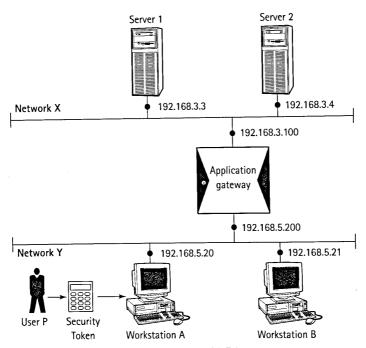


Figure 4-21: Use of an application gateway with Telnet proxy

# CHARACTERISTICS OF THE TELNET COMMUNICATION SERVICE Telnet communications has the following attributes:

- ◆ Telnet is based on TCP.
- ◆ The standard port number used by Telnet servers is 23 (TCP destination port number).

◆ The TCP source port number used by Telnet clients is any port number greater than 1023.

#### FILTER RULES SPECIFIED FOR USER P

Here are some typical rules that might apply to a user.

- ◆ P can only use the Telnet service on server 1 on working days between 7 a.m. and 6 p.m.
- P can establish this connection under the following conditions:
  - Workstation A must use IP address 192.168.5.20.
  - Server 1 must use IP address 192.168.3.3.
  - The transport protocol used is TCP.
  - The Telnet port on server 1 must be 23.
  - The source port on workstation A must be greater than 1023.
- ◆ P must authenticate himself using a security token.
- ◆ The complete connection should be logged.
- ◆ The actions that take place during connection should be monitored (control monitor).
- ◆ User P can only use the Telnet service at the specified times, and cannot use any other services.
- ◆ User P can't access any other computer systems in network X (for example, server 2).

#### OTHER USERS

Different users in an organization have different rights to access systems and data. Continuing the example rules, users other than P might be subject to the following limitations:

- ◆ Cannot access servers 1 or 2.
- Other users cannot obtain any information as to what computer systems exist in the other network.

Table 4-4 shows the filter rules for the Telnet proxy that are necessary to achieve the sample rules just outlined.

Table 4-4 FILTER RULES									
User	Source Address	Destination Address	Authenti- cation Procedure	Audit	Monitor	Week- days	Time Window		
Р	192.168.5.20	192.168.3.3	Security token	Yes	Yes	Mon-Fri	7 a.m. to 6 p.m.		

These filter rules specify precisely with which source address workstation A may access which destination address of server 1. In addition, precise rules specify the time frame within which access is permitted. Finally, the authentication procedure and complete logging (audit) and monitoring of actions (control monitor) are also specified.

#### RESULT

Using an application gateway, you can precisely specify that workstation A can have a Telnet session with server 1 at particular times; no other communications links through the application gateway are possible.

The Telnet proxy detects whether any other service has been activated on server 1 on port 23, so that this other service cannot be passed through the application gateway with the Telnet proxy. The IP addresses of networks X and Y remain concealed.

Furthermore, user P can only communicate through the application gateway with Telnet after he has been authenticated. Since all actions are logged on the application gateway, the actions of user P can be traced from the logbooks. The procedure is different for user P than when he is communicating without an application gateway, as he must authenticate himself to the proxy before he is granted access to server 1. Once the authentication phase is complete, however, operation of the Telnet proxy is transparent.

#### FTP PROXY

The FTP proxy is responsible for controlled communications using FTP and provides appropriate special security-enforcing functions for this service.

The connection for the command channel is established from the source computer system (client) to port 21 (the FTP command port) of the application gateway. The user on the source computer system identifies and authenticates himself, informing the FTP service of the connection destination. Once identification and authentication have been successfully completed, a user profile containing entries that correspond to the following information is activated:

 IP address of the source computer system that wants to establish the connection

- ◆ User name used during authentication
- ◆ IP address of the destination computer system

The FTP proxy now establishes a second command channel from the application gateway to port 21 of the destination computer system (see Figure 4-22).

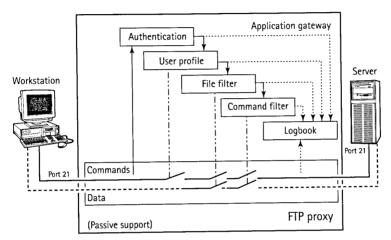


Figure 4-22: FTP proxy

#### COMMAND FILTER

The command filter analyzes and checks all the FTP commands entered by the user to ensure that they are all entered in the access rights file (user profile). For example, for the FTP proxy, you can define which commands (cd, put, get, del, and so on) can and cannot be used.

When the user enters an allowed command that entails a data transfer, connection setup of the data channel is affected, depending on whether an active or passive FTP connection has been requested on the source computer system (client side).

If a user attempts to use a command for which he lacks usage rights, an error message is displayed, the unauthorized attempt is entered in the logbook of the application gateway and, if specified, sent to Security Management in the form of a spontaneous message.

#### FILE FILTER

With FTP proxies, restricting the names of files that can be transmitted is usually possible with a file filter. Examples of such file filtering rules are:

- ◆ Only files with the name Input.new and Output.new may be transferred.
- ◆ No files with the suffix .exe can be transferred.

#### LOGBOOK

The FTP proxy can routinely write the following protocol entries to the application gateway's logbook:

- IP address and name of the source computer system
- IP address and name of the destination computer system
- Time and date of connection setup
- Name of the user
- Number of bytes transmitted
- ◆ Name of the files transmitted
- ♦ Commands used
- Time and date of disconnection

#### USE OF AN FTP PROXY

Using the FTP proxy, the commands used in an FTP session can be precisely specified. If, for example, a software house wants to send an update to a particular server, an employee of the software house is permitted to use the commands cd and put. These commands are sufficient, enabling him to perform his work.

Reducing the number of permitted commands prevents any unintentional or deliberate damage from occurring as a consequence of this action. For example, an attempt to execute the del (delete) command will be detected in the FTP proxy of the application gateway and indicated to the user. The event is entered in the logbook and, if the ruleset so specifies, a spontaneous message is sent to Security Management with the corresponding log data.

#### HTTP PROXY

The HTTP proxy is responsible for controlled communications using HTTP. It provides appropriate special security-enforcing functions for this service.

A connection is established from the source computer system (client) to port 80 of the application gateway (the port for the HTTP service). The user on the source computer system (client side) now identifies and authenticates himself, informing the HTTP service of the connection destination. Once identification and authentication have been successfully completed, a user profile containing entries that correspond to the following information is activated:

- ◆ IP address of the source computer system that wants to establish the connection
- User name used during authentication
- ◆ IP address of the destination computer system

The HTTP proxy establishes a second connection from the application gateway to port 80 of the destination computer system. The user can now use the HTTP service of the destination computer system from the source computer system via the application gateway (HTTP proxy) (see Figure 4-23).

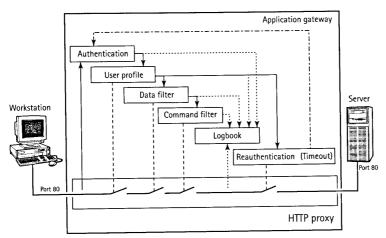


Figure 4-23: HTTP proxy

#### REAUTHENTICATION

The HTTP protocol does not work in a session-oriented manner. In other words, the HTTP proxy cannot tell when a session is ended on its own. Every time a WWW page is requested, a connection is established through the firewall system, the WWW page is transmitted, and then cleared. On the first occasion, authentication is performed prior to transmission. For this reason, a timer that records the beginning of the session is set. Following expiry of the timer, the HTTP proxy automatically shuts down the associated HTTP session. As soon as any user activity is detected in this session, the timer is restarted. The proxy can be configured so that, if the timer has previously been switched off, identification and authentication must be repeated if communication is resumed.

#### COMMAND FILTER

The command filter analyzes and checks the methods (FTP, HTTP, NNTP, SMTP) and commands (put, get, post, for example) that are used.

Any attempt to use an invalid method or an illegal command is indicated as such to the user, and a corresponding entry is made in the application gateway logbook. If the ruleset has been so defined, a spontaneous message is also sent to Security Management in such cases, together with the log data.

#### DATA FILTER

A data filter can be incorporated in the HTTP proxy so that only predefined URLs are allowed, effectively acting as a URL blocker. For example, a data filter can specify that users may only use HTTP servers with the domain \*.de. The data filter can also be used within the proxy to filter out known, undesirable files or HTTP pages. This filtering can be used, for example, to block out files known to contain viruses or HTTP pages displaying pornographic material.

#### **CONTENT SECURITY**

Content security refers to the security mechanisms that are used to protect against the threats associated with active content in HTML pages.

APPLET FILTER With the aid of an applet filter, you can prevent the use of Java, Java scripts, and ActiveX. This enables implementation of that part of an organization's security policy that relates to the use of dynamic program parts. For example, you could allow Java in the network to be protected for the Intranet applications but prohibit communication through the firewall system to computer systems in the insecure network.

MALWARE FILTER With a malware filter, you can locate viruses, worms, and Trojan horses and prevent them from causing any damage.

#### LOGBOOK

The HTTP proxy can enter the following log data in the application gateway's logbook:

- IP address and name of the source computer system
- IP address and name of the destination computer system
- ◆ Time and date of connection setup
- Name of the user
- Number of bytes transmitted
- ◆ Name of the file or HTML page (name of the page and IP address of the server/destination computer system) transmitted
- ◆ Time and date of disconnection

#### JAVA PROXY

Java proxies enable Java applets to be used with confidence. Given the continuing discovery of security vulnerabilities in the Java sandbox model establishing protection from Java applets is a necessity in most cases. There is also a bit of a "war" waging between Microsoft and Sun (the original inventor of Java). Microsoft refuses to license the Java run-time from Sun and has instead created its own Java interpreter. Microsoft's Java interpreter is not fully compatible with Sun's implementation of Java. The incompatibilities between Java implementations have lead to security issues due to the necessity to maintain and patch two different Java environments.

#### BASIC OPERATION

The basic operation of Java is as follows:

- The Java runtime environment used in the enterprise to execute Java applets is standard.
- ◆ The technology of the Java runtime environment used has been evaluated and is thus viewed as highly trustworthy.
- ◆ Using check mechanisms (special Java applets that are sent by the Java proxy), it is possible to check whether any changes have taken place in the Java runtime environment on the client.

If a user accesses a WWW server from his computer in a network to be protected and an applet is downloaded, the Java relay detects it. Using the Java Policy Manager, the Java relay determines which policy has been specified for which user. If necessary, a Java applet is sent down with check mechanisms to check the trustworthiness of the Java runtime environment. Only when a reply is received to the effect that the Java runtime environment is still trustworthy and has not been manipulated is the real Java applet downloaded from the proxy and securely executed in the trustworthy Java runtime environment. Figure 4-24 illustrates a Java proxy.

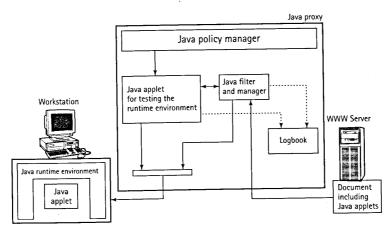


Figure 4-24: Java proxy

### JAVA PROXY MANAGER

The Java Proxy Manager is responsible for the security policy for Java applications in the entire organization. Accordingly, definitions are entered here to define which permissions the individual user can implement. The Java Proxy Manager also specifies what Java class operating system access a given user is allowed in the relevant Java runtime class library.

#### LOGBOOK

The Java proxy can enter the following log data in the application gateway's logbook:

- ◆ IP address and name of the source computer system
- IP address and name of the destination computer system
- ◆ Time and date of connection setup
- Name of the user
- ◆ Time and date of disconnection

#### E-COMMERCE PROXY

An application gateway is not limited to just proxies for protocols like HTTP or Telnet. Increasingly, such protocols are being integrated into tunnel protocols in order to retrospectively equip these protocols with security functions or to make them proxy-capable.

Examples of such tunnel protocols are SSL and SOCKS. The most well known case is HTTPS (HTTP in an SSL tunnel). Such tunnels can be integrated as functionalities into an existing tunnel, or they can be implemented as wrappers. A wrapper accepts the tunneled protocol, unpacks the protocol, and conveys the data to the actual protocol proxy. The advantage of this concept is that many tunnels can be combined with many protocols relatively easily.

## TLS WRAPPER (SSL GATEWAY)

TLS (previously known under the name of SSL) is a tunnel protocol that can be used for various protocols. The most frequent application is HTTP. There are two implementation methods available:

◆ A firewall system is on the client side (that is, a browser is to access many different servers on the Internet with HTTPS through a firewall system) is only possible if the firewall system tunnels an encrypted SSL connection. This ensures end-to-end encryption between the client and the server. This mode of operation can be implemented through the normal HTTP proxy. The task of the firewall system is to protect the client. Tunneling of an encrypted connection is not ideal from the point of view of the firewall system, as the client can only be protected to a limited extent. If the client demands an end-to-end encrypted connection to the server, the firewall system can only accept or reject it.

- If the firewall system is on the server side, it is expected to protect the server. In this case, the firewall system can take over the entire security functionality for the Web site. This entails:
  - Secure and logged authentication of the client to the server using a client certificate. This function is optional. A bank might perhaps only grant access to the online banking server to existing customers. An Internet shop, on the other hand, will want to allow access to its product range to all new customers. In this case, client authentication can be omitted.
  - Secure and logged authentication of the server to the client using a server certificate.
  - Secure storage of the server certificate. If a hacker steals the server certificate, the hacker can assume the identity of the Web site on the Internet and cause extensive damage, as clients will rely on the server certificate to securely authenticate the server. The secure storage can either be on a separate system behind the SSL gateway a separate authentication server or it can be realized in a suitable hardware module. The certificate can never leave this module, which operates like that of a smart card.
  - Efficient processing of encryption operations, including authentication via asymmetric encryption algorithms and data encryption via symmetric encryption algorithms. This offers an advantage, since many Web sites operate at the limits of their capabilities, and it is often impossible to implement SSL on the Web server itself without experiencing a significant loss of performance.
  - Independent logging of attacks. This logging is reasonable, since Web sites are a popular target for attacks on the Internet. The firewall system offers a uniform logging and alarm mechanism, continuously monitored by the administrator. In this case, it makes sense to supplement the firewall system with an intrusion detection system. However, the firewall system can also log information about the authentication process that is not accessible to the intrusion detection system, which operates as an outsider.



Pure SSL tunnels through a firewall system are a disadvantage from a security standpoint, since they place the responsibility for security on the server, which is usually unable to guarantee it.

## When to Use Application Gateways

An application gateway is an ideal active firewall element whenever protective measures must be made available to applications. The possibility of logging in the Application layer is also a good reason to consider the application gateway in a firewall concept.

Wherever a connection to the Internet is planned and the computer systems in the network to be protected have a high protection requirement (see also Chapter 3), inclusion of an application gateway in the firewall configuration should be considered.

Organizational units that want to be isolated from other parts of the enterprise network can achieve special protection using an application gateway.

# CAPABILITIES, ADVANTAGES, AND SPECIAL ASPECTS OF AN APPLICATION GATEWAY

The use of an application gateway yields several desirable security conditions.

- Secure design concept, since modules are small and straightforward (proxies)
- ◆ Concentration on what is essential
- ◆ Higher security, since without exception all packets are passed through the proxy
- ◆ The communication partner of the computer systems that communicate through the application gateway is the proxy. This means that services can be kept truly separate.
- Connection data and application data can be logged, enabling the actions
  of users who communicate through/over the application gateway to be
  recorded.
- ◆ Concealment of internal network structure
- Security-enforcing functions are made available to the applications (command filters, file filters, data filters, and so on).
- ◆ Network Address Translation takes place

# DISADVANTAGES AND LIMITATIONS OF AN APPLICATION GATEWAY

Application gateways have some disadvantages as well as advantages.

- ◆ Low flexibility, since a new proxy must be provided for every new service
- Application gateways are generally expensive

- ◆ Not transparent, since a different procedure is used during communication over the application gateway
- Some application gateways cannot detect IP spoofing, although this is not a common problem

# **Adaptive Proxy**

Some vendors of security applications attempt to combine the advantages of packet filter and application gateway into an *adaptive proxy*. The adaptive proxy functions like an application proxy during the connection setup phase and acts like a packet filter during the data transfer phase (see Figure 4-33). The advantages of this approach are obvious: a very high level of security is achieved in the first phase, and thereafter, only the rapid packet filter functions are performed. Assuming that all attacks involve the first phase – that is, the establishment of a communications link – a high level of security could be achieved.

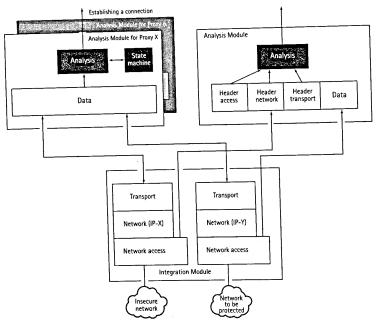


Figure 4-33: Adaptive proxy

## Analogy to the Security Guard

In the connection setup phase, the adaptive proxy works like the application proxy: it not only inspects the address of inbound packets, it also opens them and examines their contents. If the "adaptive proxy security guard" has known the supplier for a long time, then he sends the supplier's truck through the gate, allowing the driver to deliver the goods directly. However, if he does not know the supplier, he sends the truck driver away after the delivery is unloaded and arranges for a company driver to take the packet to the receiver in a company vehicle.

## Capabilities and Limitations of an Adaptive Proxy

Since an electronic security guard cannot establish and rely on personal, human bonds, the adaptive proxy appears to be more interesting in theory than in practice, as it can scarcely attain the quality of an application proxy. Alternatively, the equivalent of the personal, human bond could be implemented using trustworthy networks and/or encryption systems. However, an accurate analysis of the threats and the operational environment must be performed in that case.

## Virtual Private Networks

Virtual Private Networks (VPNs) allow for secure external communications between remote entities and the internal network. As such, they form an important mechanism for firewalling. VPNs function by using encryption to encapsulate packets inside other packets, a process commonly referred to as *tunneling*. This encapsulation allows entities at either end of the "tunnel" to transparently hold private communications across the Internet. VPNs, along with encryption, are covered at length in Chapter 7.

# Firewall Elements and the Speed-Versus-Security Tradeoff

Figure 4-34 presents a classification of firewall elements. The relative merits of the various firewall elements with regard to speed and security are presented qualitatively.



If several parallel application gateways (with application proxies) are used, then together they can produce a higher capability (throughput)), thereby surmounting the speed penalty shown. (See also Chapter 6.)

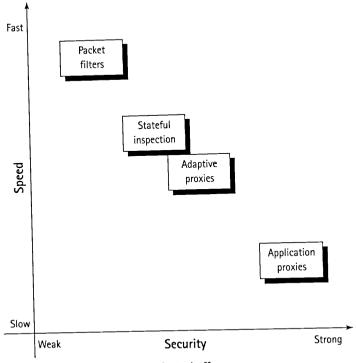


Figure 4-34: The speed-versus-security tradeoff

# Security Management for Active Firewall Elements

With a Security Management module, simple, reliable, and verifiable administration of the active firewall elements should be possible.

Because of the special design requirement that active firewall elements be implemented with minimal software, it is essential that Security Management be separated from the active firewall element. A separate Security Management module can, for example, be implemented on a separate computer system, or with the aid of removable hard disks that are only introduced when the ruleset is to be modified.

## Requirements for a Security Management Module

The Security Management module must itself be resistant to attack; otherwise, attackers could disable the security functions of the active firewall elements by penetrating the Security Management software. To this end, the Security

Management module itself must offer security mechanisms, such as identification and authentication, role distribution, logging with audit capabilities, and encryption of security-relevant information.

Security Management for active firewall elements should at least fulfill the security services *Ease of use* and *Consistency of rules*.

- ◆ Ease of use: The Security Management menus should have a reliable design and permit easy navigation around the system. Furthermore, you should not have to enter the same data twice.
- Consistency of the filter rules: Input errors in the data entry fields (MAC address, IP address, protocol number, port numbers) should be detected by the validation process. Syntactic validation should be used here.
- ◆ Filter rules should only be configurable for the appropriate protocol layer. It should not be possible for users to be entered as both active and barred from a service. Multiple entries should be eliminated. The filter rules should also be subjected to semantic validation.

To ensure a high overall security of the entire firewall system, the additional security functions of access control, access rights administration, encryption, and logging must be available as well.

- ◆ Access control in Security Management: Identification and authentication of users should be performed in order to prevent unauthorized persons from using the Security Management functionality.
- ◆ Access rights management (roles) in Security Management: To ensure secure operation of the Security Management functions, Security Management should offer the roles of Security Administrator, Operator, Data Entry Operator, Observer, and Auditor.
  - The Security Administrator is responsible, for example, for the personalization of the Security Management, the granting of access rights to the Security Management, and the creation and restoration of backups.
  - The Operator is responsible for inputting the rights of users who are allowed to communicate through the active firewall elements under the security policy of his organization.
  - A Data Entry Operator is responsible for capturing non-security-critical data, such as user names, computer systems, profiles, and so on. He cannot grant or withdraw any rights.
  - The Observer observes the operation of the active firewall element and, if necessary, analyzes any problems. He cannot grant or withdraw any rights.

■ The Auditor assumes the task of examining the Security Management logbook data for security-critical actions. He cannot grant or withdraw any rights.

For quite special, security-critical actions within Security Management, a twoperson rule can be required, in which two or more persons must simultaneously enter a separate password to initiate a given action.

The security-relevant information, such as passwords or keys for authentication with the active firewall elements, should be stored within the Security Management in an encrypted form to eliminate the possibility of misuse.

The various functions within the Security Management should be logged in separate logbooks. For this purpose, the Security Management should provide the following:

- ◆ A function logbook that contains a record of all actions performed using the Security Management (for example, the granting of user rights, the deletion of logbooks, and so on). In this logbook, the actions of the various users of the Security Management module (Security Administrator, Operator, and so on) can be recorded.
- ◆ All the logins that are sent to the Security Management must be recorded in a login logbook.
- ◆ The error logbook records must detail all errors that are detected in the Security Management.
- ◆ The backup logbook is used to log all the backup actions the Security Administrator performs within the Security Management.

# Coupling to a Network Management System (NMS)

In general, firewall systems' particularly high availability requirements means that certain spontaneous messages that provide information about the availability of the system must be sent from the firewall elements to the Network Management System (NMS). In larger organizations, the NMS is manned 24 hours a day and can respond quickly in the event of failures. For this purpose, the Security Management module should be able to exchange SNMP traps and simple GET commands with the NMS using an SNMP proxy.

## COMMUNICATION PROTECTION FOR SECURITY MANAGEMENT

In many firewall system configurations, it makes sense to isolate the Security Management functionality itself with the aid of a packet filter. If a packet filter with

encryption (VPN box) is used, it is also possible to access the Security Management remotely. If there were a number of local Security Managers in the various organizational units, they would be able to access a central Security Management function remotely.

# Summary

This chapter covers the functioning of several firewall elements in detail. Understanding how each element works is necessary for proper use of each element. Comprehending what each element can and can't do is necessary to insure proper choice of security elements for fulfilling your organizational security requirements.

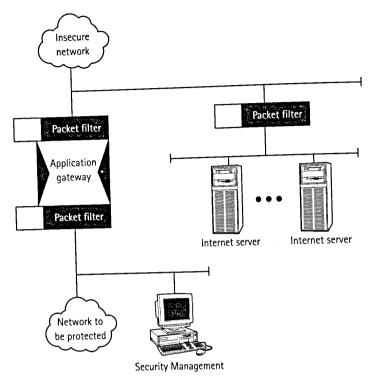


Figure 5-10: Internet servers 2

## **Intranet Servers**

Information for users from the protected network can be made available to the intranet on one or more intranet servers. The intranet server is located between the inner packet filter and the application gateway in the internal screened subnet. This positioning means that users do not need to access the insecure network from inside the protected network.

Isolation of the networks means that no direct link exists between the insecure network and the protected network. Data from the protected network is only transported as far as the intranet server (or servers) located behind the application gateway, as shown in Figure 5-12.

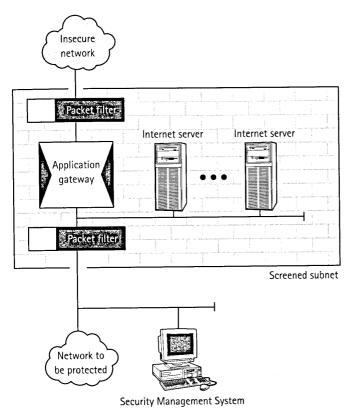


Figure 5-11: Internet servers 3

Data intended for users from the protected network is stored on the intranet server in the internal screened subnet for retrieval. All users with access to the intranet server are controlled through the inner packet filter. Communications between computer systems in the protected network and the intranet servers can also pass through the application gateway. The application gateway operates as a single-homed application gateway, providing security services such as user authentication and logging of user actions.



Chapter 9 describes other ways intranet servers can be used.

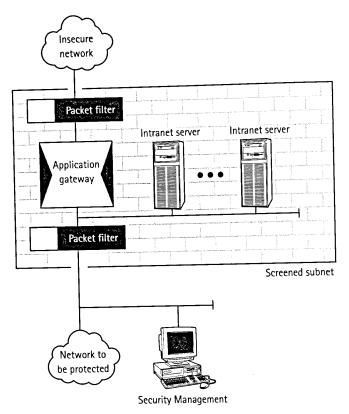


Figure 5-12: Intranet servers

## Several Application Gateways in Parallel

In practice, there are applications for which it is appropriate to operate several application gateways in parallel in the screened subnet. This is the case, for example, when performance is to be increased, redundancy is to be achieved, or separation is to be maintained between certain services. These performance or redundancy requirements are used in large organizations to cope with future requirements flexibly and securely.

The implementation of parallel firewalls results in avenues for both redundancy and fail-over. Actually, implementing parallel firewalls presents two challenges when it comes to synchronizing rules and consolidating log information.

Normally, the Security Management function is used to implement parallel rules on the multiple gateways. Proper execution requires routing support as well as firewall configuration. In addition, the use of parallel firewalls necessitates the consolidation of log information from the gateways. Failure to consolidate this

information makes detecting malicious activity more difficult in most configurations because an attacker's activity is spread across the parallel systems, effectively reducing the level on each system (note Figure 5-13). This reduced level may cause the attacks to fall below malicious activity thresholds. By consolidating the activity from the parallel gateways, you regain the advantage of analysis.

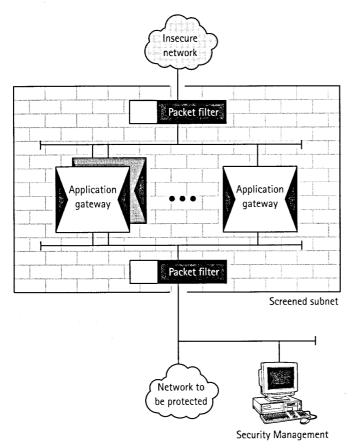


Figure 5-13: Several application gateways operated in parallel

#### SEPARATION BETWEEN PARTICULAR SERVICES

One application gateway can be responsible for the communications that pass from an insecure network to the protected network, and a second application gateway can be responsible for communications in the opposite direction, from the protected network to the insecure network. With this configuration, you can more clearly

#### Firewall Architecture for the Enterprise 178

define different operating times, security policies, and individuals who are responsible for operations.

Experience shows that large firewall systems benefit from having separate application gateways for applications, such as the e-mail service, in order to ensure an even level of performance.

#### INCREASED PERFORMANCE

The use of several application gateways can achieve a higher overall level of performance (throughput). This increased throughput can be achieved, for example, by dividing the services among different application gateways. Proxies for services that operate according to the store-and-forward principle and do not require any user authentication, such as an SMTP or NNTP proxy, as well as proxies for services that operate in a user-oriented fashion, are configured separately on one or more application gateways.

Application gateways can also be used for different groups or organizational units. In this case, the firewall systems can respond flexibly to additional requirements.

For a greater throughput, you can offer a given service on several application gateways. In this case, it is particularly important that Security Management implement the same security policy on both application gateways to prevent one application gateway from allowing something that is forbidden by the other.

Different user groups can be offered different throughput assurances. One group comprising only a limited number of users might use a high-performance application gateway that guarantees a particular throughput, while everyone else uses the other application gateway, whose total load is such that it cannot offer any throughput guarantees.

True load balancing is often required in cases where multiple parallel gateways are required on a single service. To achieve load balancing, parallel systems must communicate their overall utilization and spread new incoming activity amongst themselves appropriately. This additional communication between gateways and the protocols necessary for achieving the load balancing run counter to the desired design concept of minimal software. Given the complexity of the TCP/IP suite of protocols and the lack of a load-balancing mechanism within TCP/IP, innovative techniques are required to achieve load balancing. These techniques vary in terms of methodology but are all rather complex. This is not to say that load balancing precludes security by any means. Just keep in mind that a trade-off of security for performance occurs when load-balancing mechanisms are utilized.

#### CREATING REDUNDANCY

Another reason for operating several application gateways in parallel is high availability. If, for example, the application gateways operate in parallel in hot standby mode, a higher level of availability can be assured.

# The Right Firewall Concept for Every Application

Different decision criteria are used when choosing a particular firewall system design. The primary goal is always to achieve a level of security that matches the protection requirement of the secure network.

The comparison shown in Table 5-1 was developed to determine which firewall system best meets the security requirements of different organizations. Table 5-1 rates several factors—trustworthiness of the insecure network, trustworthiness of the communications partner, and attack potential—based on whether the insecure network is inside or outside the organization.

A firewall system is primarily used to reduce the vulnerability of an organization's sensitive data. If the data on a network does not need to be protected, a firewall system isn't necessary. However, if protection is required, the relevant circumstances must be considered and an appropriate firewall system must be selected.

TABLE 5-1 SECURITY REQUIREMENTS MATRIX		
Criteria	The Insecure Network is within the Organization	The Insecure Network is Outside of the Organization
Trustworthiness of the Insecure Network	Very high Company is responsible for the insecure network	Dependent on special factors that are difficult to measure
	It is checked regularly	Company itself is not responsible for the insecure network
		All the risks must be considered
Trustworthiness of the Communications Partner	Very high The communications partners belong to the same organization and the same security policy applies to them One still needs to beware of insiders	Probably very low In principle, no assumptions can be made regarding the trustworthiness of communications partners from outside
		Continued

#### 180 Firewall Architecture for the Enterprise

TABLE 5-1 SECURITY REQUIREMENTS MATRIX (Continued)		
Criteria	The Insecure Network is within the Organization	The Insecure Network is Outside of the Organization
Attack Potential	Very low	Very high
	The communications partners belong to the same organization and the same security policy applies to them	The network partners have very different protection requirements (for example, hackers and professional attackers from the Internet)
	One still needs to beware of insiders	

#### **EVALUATION OF PROTECTION REQUIREMENT**

Even when IT systems are used within organizations, *blind trust* in employees would be a mistake and would mean that security was only deceptive, as suggested by case histories involving insiders. The 2002 FBI/Computer Security Institute crime survey for 2001 found that 33 percent of respondents suffered an attack from internal employees. Of even greater concern is the fact that insider attacks are almost always more severe (and monetarily more costly) in nature because of the inside information possessed by the attacker.

Table 5-2 presents a decision matrix for, depending on the protection requirement and the scenario, what active firewall element or combination of active firewall elements should be used.

TABLE 5-2 DEC	ISION MATRIX FOR T	HE FIREWALL SYSTEM	
Protection Requirement	Risks	Scenario	Firewall System
Low	Minor infringement of laws	Within the organization	Packet filter
	Limited negative external effect	Outside the organization	Dual-homed application gateway
	Financial loss <\$10,000		

Protection Requirement	Risks	Scenario	Firewall System
High	Major infringement of laws Wide negative external effect	Within the organization	Packet filter plus single-homed application gateway or stateful inspection or adaptive proxy
	Financial loss >\$10,000 but <\$2.5 million	Outside the organization	Packet filter plus dual-homed application gateway
Very High	Fundamental violation of laws  Extremely serious negative external	Within the organization	Screened subnet with packet filter plus single-homed application gateway
	effect Financial loss >\$2.5 million	Outside the organization	Screened subnet with packet filter plus dual-homed application gateway (high-level firewall system)

If a network requires any type of protection, the firewall system must include a dual-homed application gateway.

Depending on the protection requirement, a dual-homed application gateway should be used either on its own or in combination with a packet filter. A screened subnet should probably be used as well.

If the insecure network is under the control of the organization (for example, an intranet), it is sufficient to use only packet filters or a combination of packet filters with single-homed application gateways, depending on the protection requirement. Alternatively, stateful inspection or adaptive proxy concepts, solutions that also provide security functions at the application level, can be used.

A very high level of security is possible if a dual-homed application gateway is combined with a packet filter or screened subnet.

In addition to the basic strength of the firewall system, which determines how it is used, the following security-relevant aspects also play a role.

- The actual security services offered
- ◆ The depth of analysis at the different communication levels
- ◆ The firewall solution design concept



The reserved IP address ranges are always contiguous.

As previously stated, these IP addresses are reserved for private purposes and are never granted on the Internet. Furthermore, because they are only used in the protected network, all networks can use the same reserved address ranges internally.

The address range chosen depends on the organization's particular needs. In large protected networks, in which subnet masks do not constitute a problem, the Class A address 10.0.0.0 would probably be used. In a small, protected network, a Class C address (for example, 192.168.1.0) could be utilized.

Additional information can be found in RFC 1918. RFC 1918 is the document that reserves the addresses listed in Table 11-1 for internal use only.

The Internet Assigned Numbers Authority (IANA) assigns external addresses. Major providers, such as Cable and Wireless, obtain addresses from the IANA and re-assign them to subscribers (including ISPs). You can obtain addresses directly from the IANA, but this is normally only done for very large organizations. Due to the lack of available addresses, all external address assignments must be justified closely.

In a small organization with three departments (in which no department has more than 250 computer systems), reserved addresses can be shared in the following way:

- ◆ Department 1: IP addresses 192.168.1.0-192.168.1.255
- ◆ Department 2: IP addresses 192.168.2.0-192.168.2.255
- ◆ Department 3: IP addresses 192.168.3.0-192.168.3.255

This particular address allocation has the following advantages:

- Standard network size (not subnetted)
- ◆ A clear separation between individual departments
- The standard network mask can be used in the individual subnets

Big organizations with an extremely large number of computer systems (in which many departments are broken down into sections or sub-departments), address allocation can follow these guidelines:

- ◆ The address range 10.0.0.0 should be chosen to allow plenty of room for expansion.
- ◆ Each department is given its own Class B network, which can be further subdivided as required for example, into a Class C subnet or smaller.

Don't use subnets with the addresses 0 and 255. Doing so may cause ambiguities with broadcasts. The subnets 10.0.xxx.yyy (problem with network addresses 10.0.0.0), 10.255.xxx.yyy (problem with broadcast 10.255.255.255), or 10.1.255.xxx (as in last example) should not be assigned.

Medium-sized organizations can fill their requirements using Class B addresses. In general, however, take the small solution with 192.168.xxx.yyy, or the large solution with the network 10.0.0.0.

#### Firewall Systems and Network Address Translation

Firewall systems and firewall elements such as application gateways may not perform network address translation in the strict sense of the word, but they achieve the same end by using official Internet IP addresses for the insecure network and reserved IP addresses for the intranet.

When communications occur through a firewall system, an application gateway has a link with the computer system (or workstation) in the protected network and a link with the computer system in the insecure network (or server). In other words, the firewall system is the communications partner for both ends of the connection. The firewall system communicates in the insecure network with the official Internet IP address and in the protected network (intranet) with the reserved IP address. This behavior is similar to network address translation, except that translation tables at the application layer, rather than the network layer achieve it.

The firewall system's IP address, which belongs to the reserved IP address range of the intranet, is always stated as the communications partner within the protected network. The firewall system's official Internet IP address is always given as the communications partner in the insecure network (see Figure 11-1). An attacker can't tell which computer system is hidden behind the firewall system in the protected network.

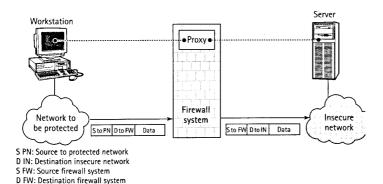


Figure 11-1: Firewall system and network address translation

# Problems for Networks That Work with Illegal IP Addresses

Occasionally, a network doesn't use reserved IP addresses, which creates problems when connecting to the Internet. If any IP addresses officially assigned to other organizations are used illegally in the protected network, problems result at the interface because the IP addresses exist both in the insecure and in the protected networks.

There are two possible solutions to resolving the problem of using external addresses internally:

- Assignment of reserved IP addresses in the intranet. Replace the illegally used IP addresses with unproblematic reserved IP addresses. Large organizations may experience problems with this approach; the procedure can be time-consuming if a large number of computer systems are involved.
- ◆ IP address mapping. Map the illegal IP addresses onto reserved addresses. Depending on which illegal addresses are used, this procedure can be implemented using a substitution table, for example, or a simple mapping rule. Some firewall products offer IP address mapping as an additional functionality.

#### **Network Address Translation Problems**

The use of network address translation introduces some problems not experienced when using non-translated IP addresses. Ironically, security protocols often don't operate correctly with NAT. The majority of virtual private network (VPN) protocols do not function correctly because of the address translation. Protocols that embed the source client address in the packet generally do not work. When the remote application compares the source address of the packet (the external routable translation address) with the address in the packet (the client's internal IP address) and discovers a mismatch, it disallows communications.

Using NAT for these situations requires either the use of an external non-translated address or an alternative protocol that works successfully with NAT. In the case of a VPN, you could put the VPN system at the network perimeter and assign it an external address—an example of the first solution. This allows network-to-network VPN access. You could also use a protocol, such as SSH, that works with NAT—an example of the second solution. Some vendors have released modified versions of standard protocols that also work with NAT. Cisco has developed a VPN black box solution called a *VPN concentrator*. The VPN concentrator supports a modified IPSEC implementation that works through NAT by tunneling IPSEC using UDP packets.

#### **Domain Names**

In general, there is a primary DNS server on the Internet and a primary DNS server in the intranet. In mid-size organizations and larger, these are usually separate physical machines. In smaller companies the two different host groups may be referenced in different zones on the same physical DNS server. The intranet names and IP addresses should only be known to the DNS server in the intranet to keep the assignment of addresses in the Internet unique.

Due to the uniqueness of names throughout the Internet (including an organization's own intranet), the domain names for the intranet should also be adapted to the Internet's name schema. For this purpose, the domain name in the Internet can be appended to the name in the intranet, as follows:

Internet domain: company.com
Intranet domain: intra.company.com

In this way you can implement a simplified procedure for the DNS server and firewall configuration in large intranets.

# Administration of Several Firewall Systems via a Security Management Module

Large organizations may require implementing several interfaces for insecure networks. In this case, administering and controlling a number of firewall systems centrally via a Security Management module makes sense.

You can administer several firewall systems from the central Security Management module. This enables an organization with several interfaces to insecure networks to easily implement its security policy. Information is exchanged between Security Management and the firewall systems in a secure way. Other firewall elements in the intranet, such as packet filters and application gateways, can also be administered with the central Security Management).



For more information on other firewall modules, refer to Chapter 9.

## **Nested Firewall Configurations**

A nested firewall configuration can occur in several applications in an organization. Figure 11-2 depicts a simplified example of such a configuration.

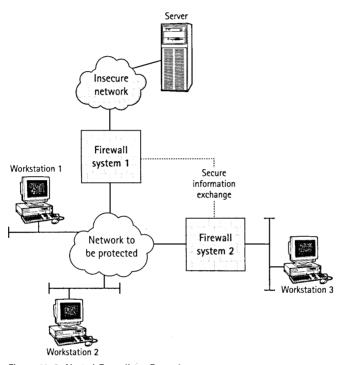


Figure 11-2: Nested firewall configuration

A firewall system is set up between the insecure and protected networks. Another firewall system is installed within the protected network in order to segregate a particularly sensitive organizational unit, such as a board of directors. If a user from Workstation 1 or 2 wants to use the services of a server in the insecure network, Firewall System 1 alone (shown in Figure 11-2) can handle the connection. If, however, a user from Workstation 3 wants to use this server, both firewall systems must process the request.

The configuration illustrated in Figure 11-2 can be implemented either with two independent Security Management modules or with one Security Management module. If one Security Management module is used, special requirements apply. Information about the authentication and destination of the data must be exchanged in a secure form between the two firewall systems.

### **Availability**

Internet users often need continuous and reliable access to the services they use to perform their tasks. This availability requirement affects not only the firewall system components, but also the availability of the network, especially its availability over the Internet. The extent to which organizations can guarantee Quality of Service (QoS) on the Internet depends largely on whether Internet Service Providers (ISPs) can guarantee availability and performance globally in the future. Availability issues must be thoroughly considered before any availability problem occurs. By anticipating where issues can arise, you can implement fail-over and redundant systems for critical infrastructures.



Chapters 4 and 5 discuss the different components and concepts behind firewalls. These components are the failure points potentially requiring redundancy.

From the user's point of view, minimum requirements in the area of QoS must be guaranteed so that the risk of non-availability can be properly estimated. Redundancy solutions that guarantee a high level of availability must be accessible for firewall systems.

An availability concept for firewall system components must be developed. Accordingly, the most important components of the system must be designed to be redundant. Moreover, load balancing is necessary to achieve the degree of availability that will in turn achieve the QoS rating required for the application. Fortunately, several mechanisms, including the following, can be used to address availability:

- ♦ Redundancy
- ♦ Hot fail-over
- ◆ Clustering

Some firewalls include Quality of Service modules. QoS modules allow you to specify traffic quality of service profiles. For example, you might specify that traffic to and from the corporate Web servers takes precedence over internal employee Web browsing. In the event that traffic has to be discarded, the employee traffic is discarded, preserving maximum availability for the more critical traffic. The firewall then makes decisions about allowing or disallowing traffic based on these policies. QoS policies work by giving you the ability to classify traffic based upon type. Traffic parameters include the following packet characteristics:

#### 314 Firewall Architecture for the Enterprise

- Source or destination
- Time of day
- ◆ Specific URLs
- Users
- Applications

Critical business functionality can be classified as priority, and other traffic can be held or disallowed as necessary.

A second option involves hot fail-over or clustering systems. Hot fail-over can be as simple as maintaining a redundant system with duplicate rulesets that can be activated quickly in the event of a failure on the primary system. At the other end of the spectrum is full-system clustering. Clustering usually involves two or more systems sharing a central drive subsystem. In the event of a system failure, the other system or systems continue transparently. Because of the shared storage system, no data is lost during the switch.

Availability problems may result from more than simple hardware or software failure. Connectivity failure, for example, makes using systems as desired rather difficult. Failures in other services, such as connectivity, require firewall configuration support as well. Having a redundant T1 in place won't do any good if the firewall isn't also configured to use the alternate T1 in the event of a failure of the primary T1.

Denial-of-Service attacks can pose a particular challenge for availability planning. While the problem of DoS attacks is not solved yet by any means, several of the techniques being developed to combat Denial-of-Service attacks look promising. In the interim, planning for redundancies and response plans is your best defense.

#### Performance

Performance is, in many ways, the other side of the availability coin. Performance considerations often take precedence over security considerations in an organization. Given the importance of performance, a firewall system should be designed and implemented from the beginning with *maximum* performance in mind.

Performance is directly impacted by the firewall policies you implement. Packet filtering has the least negative impact on performance, while application gateways produce the most performance degradation. The order in which rules are implemented also affects performance. Putting the most used rules first often results in a measurable, positive performance impact.

The way rules impact both security and performance should also be considered. Traffic rules that have very little impact on security may significantly affect performance. The appropriateness of rules that yield only small security improvements should be carefully evaluated.

Virtually all commercial, firewall software or black box vendors release performance specifications as well as hardware guidelines for optimizing traffic handling in your environment.

Ultimately, performance impact can only be measured by specific environmental testing. Software packages are readily available for stress-testing servers in various capacities. Using a Web server stress suite from outside your firewall can yield useful performance information. You can evaluate specific rule impact by running the same test sequence with different rules enabled or disabled. To perform specific rule testing, run a tool, such as a Web server performance tester, against the Web server while all of the firewall rules are enabled. Then re-run the test while disabling rules in the firewall. Run the test for each rule. Testing in this manner will clearly indicate the performance impact of each rule.

Firewall performance enhancements are also available. These enhancements range in scope from hardware add-ins designed to offload processor-intensive activity such as encryption/decryption to load-balancing tools for balancing system utilization between multiple firewalls and/or computer systems.

## Summary

Issues such as NAT, nested firewall configurations, availability, and performance should not be the primary criteria for designing the architecture and configuration of your firewall. That being said, these issues are critical to the success of the firewall. Always bear in mind that firewalls exist to eliminate some of the disruptions to conducting business. If the firewall itself becomes a significant disruption to business, the firewall's primary purpose has failed. Spending the extra planning and time to evaluate criteria, such as performance and availability, enable you to take your firewall to the highest possible level of security and functionality.

#### Attachment 1b: Library of Congress catalog record for Document 1

Records 1 through 1 of 1 returned.

```
LDR 01379cam 22003254a 4500
001 12758590
005 20160907154050.0
008 020427s2002
                                     001 0 eng
                                b
                     nyua
906 $a7$bcbc$corignew$d2$eepcn$f20$gy-gencatlg
925 0 $aacquire$b1 shelf copy$xpolicy default
925 1 $aacquire$b2 shelf copies$xpolicy default
      $apc24 2002-04-27$apv06 2003-03-27 to ASCD$ijg12 2003-03-28$ejg08
   2003-04-17 to Dewey
010 $a 2002102445$z 2002106044
020 $a076454926X
       $a076454926X
035
       $a(DLC) 2002106044
      $aDLC$cDLC$dDLC
040
       $apcc
050 00$aTK5105.59$b.P64 2002
100 1 $aPohlmann, Norbert.
245 10$aFirewall architecture for the enterprise /$cNorbert Pohlmann and Tim
   Crothers.
260
       $aNew York, NY :$bWiley Pub.,$cc2002.
       $axxi, 481 p. :$bill. ;$c24 cm.
504
       aIncludes bibliographical references (p. 437-438) and index.
650 0$aFirewalls (Computer security)
     0$aComputer networks$xSecurity measures.
650 0$aComputer network architectures.
700 1 $aCrothers, Tim, $cMCSE.
856 42$3Contributor biographical information
$uhttp://www.loc.gov/catdir/bios/wiley045/2002102445.html
856 42$3Publisher description
   $uhttp://www.loc.gov/catdir/description/wiley038/2002102445.html
856 41$3Table of contents
   $uhttp://www.loc.gov/catdir/toc/wiley023/2002102445.html
<u>Labeled display</u> | <u>Brief Record Display</u> | <u>New Search</u>
```

This display was generated by the CNIDR Web-Z39.50 gateway, version 1.08, with Library of Congress Modifications.

#### Attachment 1c: ACM Digital Library index record for Document 1

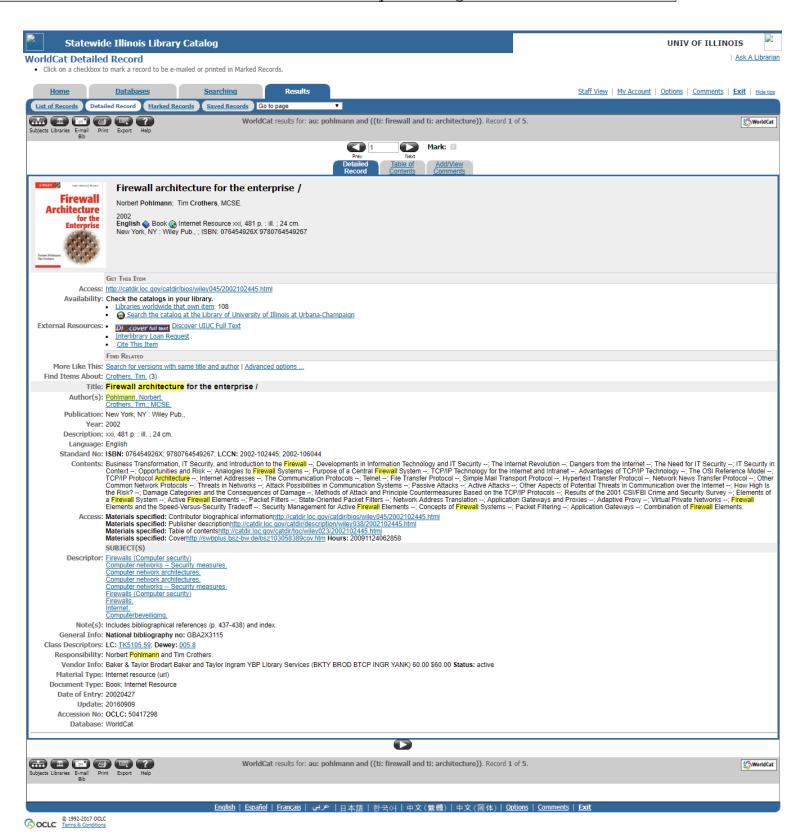


Powered by THE ACM GUIDE TO COMPUTING LITERATURE

The ACM Digital Library is published by the Association for Computing Machinery. Copyright © 2017 ACM, Inc.

<u>Terms of Usage Privacy Policy Code of Ethics Contact Us</u>

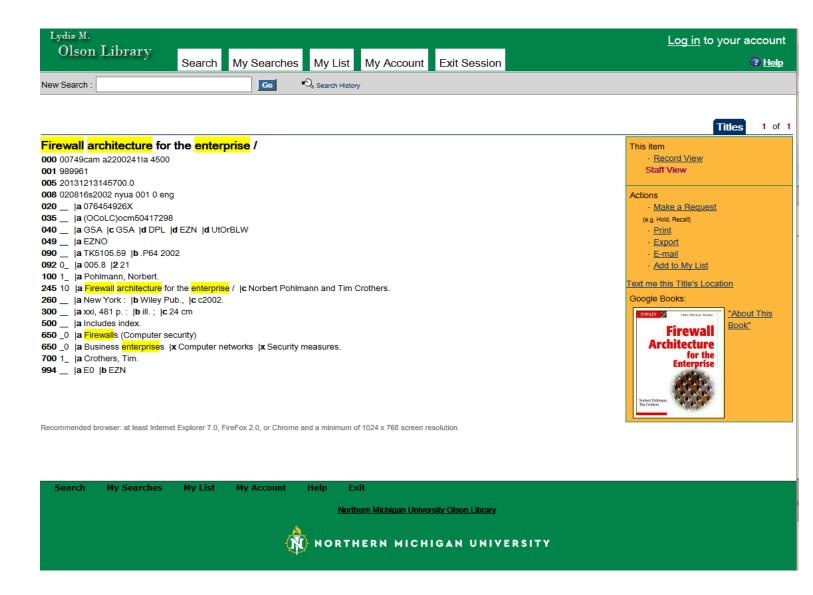
#### Attachment 1d: Statewide Illinois Library Catalog record for Document 1

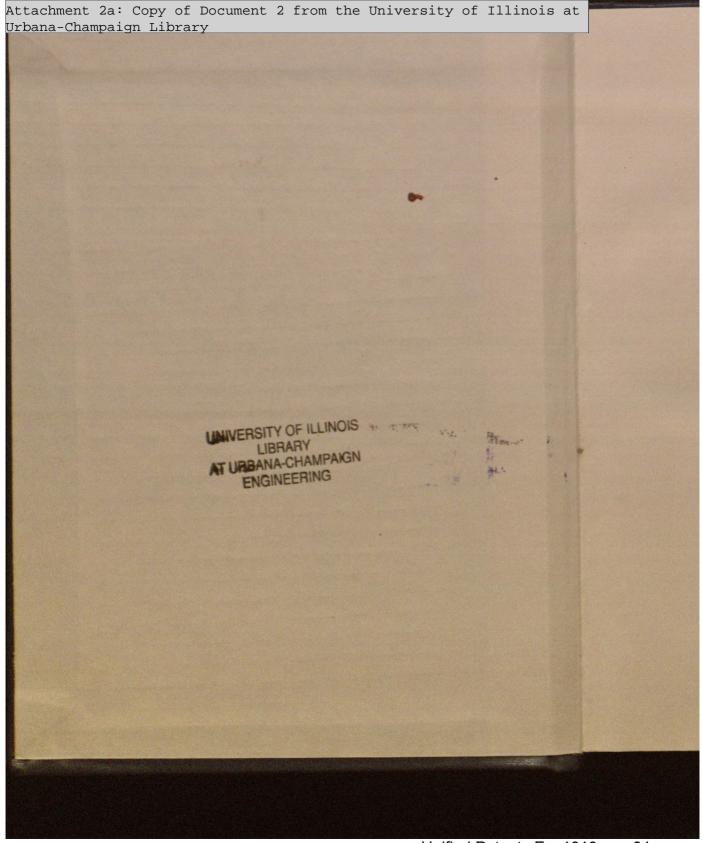


#### Attachment le: United States Copyright Office record for Document 1

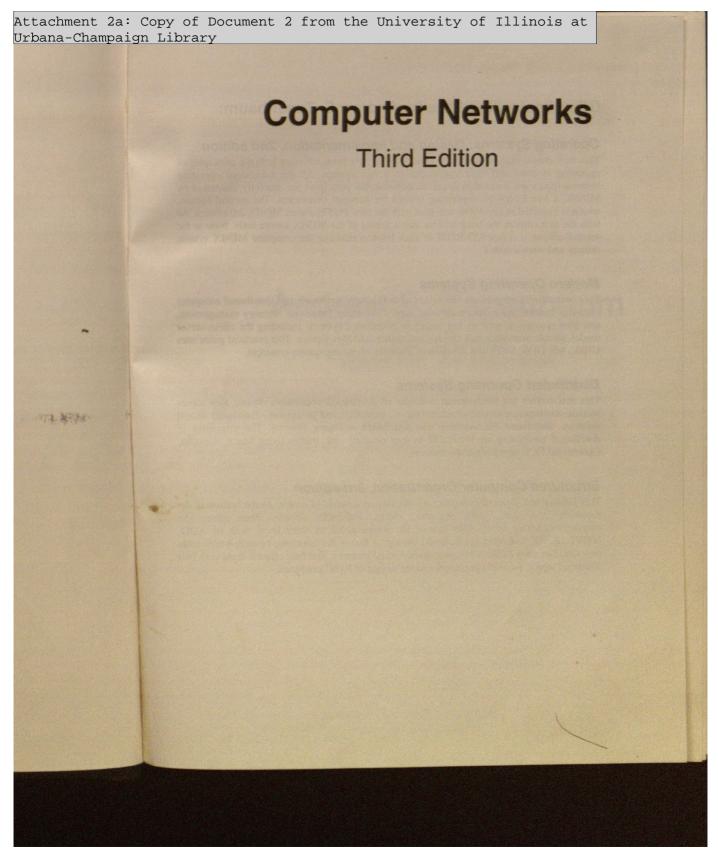


#### Attachment 1f: Northern Michigan University Library catalog record for Document 1





Attachment 2a: Copy of Document 2 from the University of Illinois at Urbana-Champaign Library UNIVERSITY LIBRARY UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN The person charging this material is responsible for its renewal or return to the library on or before the due date. The minimum fee for a lost item is \$125.00, \$300.00 for bound journals. Theft, mutilation, and underlining of books are reasons for disciplinary action and may result in dismissal from the University. Please note: self-stick notes may result in torn pages and lift some inks. Renew via the Telephone Center at 217-333-8400, 846-262-1510 (toll-free) or circlib@uiuc.edu. Renew online by choosing the My Account option at http://www.library.uiuc.edu/catalog/ 6/ JUL 1 1 2007 OCT 0 3 2008 'AUG 11 2011 DEC 07 2012 MNY 29 7016 UIC-REC'D SEP 0 8 2017

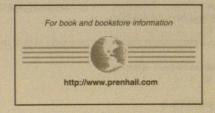


## **Computer Networks**

Third Edition

## Andrew S. Tanenbaum

Vrije Universiteit Amsterdam, The Netherlands



ne D,



Prentice Hall PTR
Upper Saddle River, New Jersey 07458

CONTENTS	
PREFACE	xv
1 INTRODUCTION	1
1.1 USES OF COMPUTER NETWORKS 3 1.1.1 Networks for Companies 3 1.1.2 Networks for People 4 1.1.3 Social Issues 6	
1.2 NETWORK HARDWARE 7 1.2.1 Local Area Networks 9 1.2.2 Metropolitan Area Networks 10 1.2.3 Wide Area Networks 11 1.2.4 Wireless Networks 13 1.2.5 Internetworks 16	
1.3 NETWORK SOFTWARE 16 1.3.1 Protocol Hierarchies 17 1.3.2 Design Issues for the Layers 21 1.3.3 Interfaces and Services 22 1.3.4 Connection-Oriented and Connectionless 1.3.5 Service Primitives 25 1.3.6 The Relationship of Services to Protocols	Control States States and Additional Control of the
1.4 REFERENCE MODELS 28 1.4.1 The OSI Reference Model 28 1.4.2 The TCP/IP Reference Model 35 1.4.3 A Comparison of the OSI and TCP Reference 1.4.4 A Critique of the OSI Model and Protoco 1.4.5 A Critique of the TCP/IP Reference Model 1.4.5 A Critique 0.4.5 A Critiq	ls 40
1.5 EXAMPLE NETWORKS 44 1.5.1 Novell NetWare 45 1.5.2 The ARPANET 47 1.5.3 NSFNET 50 1.5.4 The Internet 52 1.5.5 Gigabit Testbeds 54	
vi	

Attachment	2a:	Сору	of	Document	2	from	the	University	of	Illinois	at
Urbana-Cham	npaig	yn Lik	orai	ſУ							

CONTENTS		V

- 1.6 EXAMPLE DATA COMMUNICATION SERVICES 56
  - 1.6.1 SMDS—Switched Multimegabit Data Service 57
  - 1.6.2 X.25 Networks 59
  - 1.6.3 Frame Relay 60
  - 1.6.4 Broadband ISDN and ATM 61
    - 1.6.5 Comparison of Services 66
- 1.7 NETWORK STANDARDIZATION 66
  - 1.7.1 Who's Who in the Telecommunications World 67
  - 1.7.2 Who's Who in the International Standards World 69
  - 1.7.3 Who's Who in the Internet Standards World 70
- 1.8 OUTLINE OF THE REST OF THE BOOK 72
- 1.9. SUMMARY 73

#### 2 THE PHYSICAL LAYER

77

- 2.1 THE THEORETICAL BASIS FOR DATA COMMUNICATION 77
  - 2.1.1 Fourier Analysis 78
  - 2.1.2 Bandwidth-Limited Signals 78
  - 2.1.3 The Maximum Data Rate of a Channel 81
- 2.2 TRANSMISSION MEDIA 82
  - 2.2.1 Magnetic Media 82
  - 2.2.2 Twisted Pair 83
  - 2.2.3 Baseband Coaxial Cable 84
  - 2.2.4 Broadband Coaxial Cable 85
  - 2.2.5 Fiber Optics 87
- 2.3 WIRELESS TRANSMISSION 94
  - 2.3.1 The Electromagnetic Spectrum 94
  - 2.3.2 Radio Transmission 97
  - 2.3.3 Microwave Transmission 98
  - 2.3.4 Infrared and Millimeter Waves 100
  - 2.3.5 Lightwave Transmission 100
- 2.4 THE TELEPHONE SYSTEM 102
  - 2.4.1 Structure of the Telephone System 103
  - 2.4.2 The Politics of Telephones 106
  - 2.4.3 The Local Loop 108
  - 2.4.4 Trunks and Multiplexing 118
  - 2.4.5 Switching 130

viii	CONTENTS	
2.5	NARROWBAND ISDN 139 2.5.1 ISDN Services 140 2.5.2 ISDN System Architecture 140 2.5.3 The ISDN Interface 142 2.5.4 Perspective on N-ISDN 143	
√2.6	BROADBAND ISDN AND ATM 144 2.6.1 Virtual Circuits versus Circuit Switching 145 2.6.2 Transmission in ATM Networks 146 2.6.3 ATM Switches 147	
2.7	CELLULAR RADIO 155 2.7.1 Paging Systems 155 2.7.2 Cordless Telephones 157 2.7.3 Analog Cellular Telephones 157 2.7.4 Digital Cellular Telephones 162 2.7.5 Personal Communications Services 162	
2.8	COMMUNICATION SATELLITES 163 2.8.1 Geosynchronous Satellites 164 2.8.2 Low-Orbit Satellites 167 2.8.3 Satellites versus Fiber 168	
2.9	SUMMARY 170	
3 тн	HE DATA LINK LAYER 17	5
3.1	DATA LINK LAYER DESIGN ISSUES 176 3.1.1 Services Provided to the Network Layer 176 3.1.2 Framing 179 3.1.3 Error Control 182 3.1.4 Flow Control 183	
3.2	ERROR DETECTION AND CORRECTION 183 3.2.1 Error-Correcting Codes 184 3.2.2 Error-Detecting Codes 186	
3.3	ELEMENTARY DATA LINK PROTOCOLS 190 3.3.1 An Unrestricted Simplex Protocol 195 3.3.2 A Simplex Stop-and-Wait Protocol 195 3.3.3 A Simplex Protocol for a Noisy Channel 197	

4.4 BRIDGES 304 4.4.1 Bridges from 802.x to 802.y 307 4.4.2 Transparent Bridges 310 4.4.3 Source Routing Bridges 314 4.4.4 Comparison of 802 Bridges 316 4.4.5 Remote Bridges 317  4.5 HIGH-SPEED LANS 318 4.5.1 FDDI 319 4.5.2 Fast Ethernet 322 4.5.3 HIPPI—High-Performance Parallel Interface 325 4.5.4 Fibre Channel 326  4.6 SATELLITE NETWORKS 327 4.6.1 Polling 328 4.6.2 ALOHA 329	
4.5.1 FDDI 319 4.5.2 Fast Ethernet 322 4.5.3 HIPPI—High-Performance Parallel Interface 325 4.5.4 Fibre Channel 326  4.6 SATELLITE NETWORKS 327 4.6.1 Polling 328	
4.6.1 Polling 328	
4.6.3 FDM 330 4.6.4 TDM 330 4.6.5 CDMA 333	
4.7 SUMMARY 333	
5 THE NETWORK LAYER  5.1 NETWORK LAYER DESIGN ISSUES 339 5.1.1 Services Provided to the Transport Layer 340 5.1.2 Internal Organization of the Network Layer 342	
5.1.3 Comparison of Virtual Circuit and Datagram Subnets 344 5.2 ROUTING ALGORITHMS 345 5.2.1 The Optimality Principle 347 5.2.2 Shortest Path Routing 349 5.2.3 Flooding 351	
5.2.4 Flow-Based Routing 353  5.2.5 Distance Vector Routing 355  5.2.6 Link State Routing 359  5.2.7 Hierarchical Routing 365  5.2.8 Routing for Mobile Hosts 367  5.2.9 Broadcast Routing 370  5.2.10 Multicast Routing 372	

CONTENTS

xi

52	CONGESTION	CONTROL AL	GORITHMS	374
----	------------	------------	----------	-----

- 5.3.1 General Principles of Congestion Control 376
- 5.3.2 Congestion Prevention Policies 378
- 5.3.3 Traffic Shaping 379
- 5.3.4 Flow Specifications 384
- 5.3.5 Congestion Control in Virtual Circuit Subnets 386
- 5.3.6 Choke Packets 387
- 5.3.7 Load Shedding 390
- 5.3.8 Jitter Control 392
- 5.3.9 Congestion Control for Multicasting 393

#### 5.4 INTERNETWORKING 396

- 5.4.1 How Networks Differ 399
- 5.4.2 Concatenated Virtual Circuits 401
- 5.4.3 Connectionless Internetworking 402
- 5.4.4 Tunneling 404
- 5.4.5 Internetwork Routing 405
- 5.4.6 Fragmentation 406
- 5.4.7 Firewalls 410

#### 5.5 THE NETWORK LAYER IN THE INTERNET 412

- 5.5.1 The IP Protocol 413
- 5.5.2 IP Addresses 416
- 5.5.3 Subnets 417
- 5.5.4 Internet Control Protocols 419
- 5.5.5 The Interior Gateway Routing Protocol: OSPF 424
- 5.5.6 The Exterior Gateway Routing Protocol: BGP 429
- 5.5.7 Internet Multicasting 431
- 5.5.8 Mobile IP 432
- 5.5.9 CIDR—Classless InterDomain Routing 434
- 5.5.10 IPv6 437

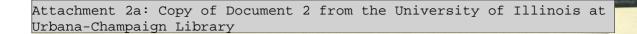
#### 5.6 THE NETWORK LAYER IN ATM NETWORKS 449

- 5.6.1 Cell Formats 450
- 5.6.2 Connection Setup 452
- 5.6.3 Routing and Switching 455
- 5.6.4 Service Categories 458
- 5.6.5 Quality of Service 460
- 5.6.6 Traffic Shaping and Policing 463
- 5.6.7 Congestion Control 467
- 5.6.8 ATM LANs 471

#### 5.7 SUMMARY 473

	CONTENTS xiii
	6.6.3 System Design for Better Performance 561
	6.6.4 Fast TPDU Processing 565
	6.6.5 Protocols for Gigabit Networks 568
	6.7 SUMMARY 572
	(if gaW att no accumulate property continued
	7 THE APPLICATION LAYER 577
	7.1 NETWORK SECURITY 577
	7.1.1 Traditional Cryptography 580
	7.1.2 Two Fundamental Cryptographic Principles 585
1 1 2 2 2	7.1.3 Secret-Key Algorithms 587
	7.1.4 Public-Key Algorithms 597
	7.1.5 Authentication Protocols 601
	7.1.6 Digital Signatures 613
	7.1.7 Social Issues 620
	7.2 DNS—DOMAIN NAME SYSTEM 622
The second	7.2.1 The DNS Name Space 622
	7.2.2 Resource Records 624 7.2.3 Name Servers 628
	1.2.3 Name Servers 020
	7.3 SNMP—SIMPLE NETWORK MANAGEMENT PROTOCOL 630
	7.3.1 The SNMP Model 631
	7.3.2 ASN.1—Abstract Syntax Notation 1 633
	7.3.3 SMI—Structure of Management Information 639
	7.3.4 The MIB—Management Information Base 641
	7.3.5 The SNMP Protocol 642
	7.4 ELECTRONIC MAIL 643
	7.4.1 Architecture and Services 645
	7.4.2 The User Agent 646
	7.4.3 Message Formats 650
	7.4.4 Message Transfer 657
	7.4.5 Email Privacy 663
	7.5 USENET NEWS 669 7.5.1 The User View of USENET 670
	7.5.1 The User View of USENET 676 7.5.2 How USENET is Implemented 675
	7.5.2 How USENET is implemented 075

xiv	CONTENTS
7.6	THE WORLD WIDE WEB 681 7.6.1 The Client Side 682 7.6.2 The Server Side 685 7.6.3 Writing a Web Page in HTML 691 7.6.4 Java 706 7.6.5 Locating Information on the Web 720
	MULTIMEDIA 723 7.7.1 Audio 724 7.7.2 Video 727 7.7.3 Data Compression 730 7.7.4 Video on Demand 744 7.7.5 MBone—Multicast Backbone 756 SUMMARY 760
7.0	SUNIVIART 700
8 RE	EADING LIST AND BIBLIOGRAPHY 767
8.1	SUGGESTIONS FOR FURTHER READING 767 8.1.1 Introduction and General Works 768 8.1.2 The Physical Layer 769 8.1.3 The Data Link Layer 770
	<ul> <li>8.1.4 The Medium Access Control Sublayer 770</li> <li>8.1.5 The Network Layer 771</li> <li>8.1.6 The Transport Layer 772</li> <li>8.1.7 The Application Layer 772</li> </ul>
8.2	ALPHABETICAL BIBLIOGRAPHY 775
INI	DEX 795
199	



INTRODUCTION

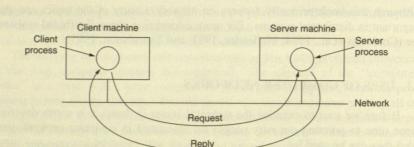


Fig. 1-1. The client-server model.

The server then does the work and sends back the reply. Usually, there are many clients using a small number of servers.

Another networking goal is scalability, the ability to increase system performance gradually as the workload grows just by adding more processors. With centralized mainframes, when the system is full, it must be replaced by a larger one, usually at great expense and even greater disruption to the users. With the client-server model, new clients and new servers can be added as needed.

Yet another goal of setting up a computer network has little to do with technology at all. A computer network can provide a powerful **communication** medium among widely separated employees. Using a network, it is easy for two or more people who live far apart to write a report together. When one worker makes a change to an on-line document, the others can see the change immediately, instead of waiting several days for a letter. Such a speedup makes cooperation among far-flung groups of people easy where it previously had been impossible. In the long run, the use of networks to enhance human-to-human communication will probably prove more important than technical goals such as improved reliability.

#### 1.1.2. Networks for People

The motivations given above for building computer networks are all essentially economic and technological in nature. If sufficiently large and powerful mainframes were available at acceptable prices, most companies would simply choose to keep all their data on them and give employees terminals connected to them. In the 1970s and early 1980s, most companies operated this way. Computer networks only became popular when networks of personal computers offered a huge price/performance advantage over mainframes.

Starting in the 1990s, computer networks began to start delivering services to private individuals at home. These services and the motivations for using them

CHAP. 1

are quite different than the "corporate efficiency" model described in the previous section. Below we will sketch three of the more exciting ones that are starting to happen:

- 1. Access to remote information.
- 2. Person-to-person communication.
- 3. Interactive entertainment.

Access to remote information will come in many forms. One area in which it is already happening is access to financial institutions. Many people pay their bills, manage their bank accounts, and handle their investments electronically. Home shopping is also becoming popular, with the ability to inspect the on-line catalogs of thousands of companies. Some of these catalogs will soon provide the ability to get an instant video on any product by just clicking on the product's name.

Newspapers will go on-line and be personalized. It will be possible to tell the newspaper that you want everything about corrupt politicians, big fires, scandals involving celebrities, and epidemics, but no football, thank you. At night while you sleep, the newspaper will be downloaded to your computer's disk or printed on your laser printer. On a small scale, this service already exists. The next step beyond newspapers (plus magazines and scientific journals) is the on-line digital library. Depending on the cost, size, and weight of book-sized notebook computers, printed books may become obsolete. Skeptics should take note of the effect the printing press had on the medieval illuminated manuscript.

Another application that falls in this category is access to information systems like the current World Wide Web, which contains information about the arts, business, cooking, government, health, history, hobbies, recreation, science, sports, travel, and too many other topics to even mention.

All of the above applications involve interactions between a person and a remote database. The second broad category of network use will be person-to-person interactions, basically the 21st Century's answer to the 19th Century's telephone. Electronic mail or **email** is already widely used by millions of people and will soon routinely contain audio and video as well as text. Smell in messages will take a bit longer to perfect.

Real-time email will allow remote users to communicate with no delay, possibly seeing and hearing each other as well. This technology makes it possible to have virtual meetings, called **videoconference**, among far-flung people. It is sometimes said that transportation and communication are having a race, and whichever wins will make the other obsolete. Virtual meetings could be used for remote school, getting medical opinions from distant specialists, and numerous other applications.

Worldwide newsgroups, with discussions on every conceivable topic are already commonplace among a select group of people, and this will grow to

5

INTRODUCTION CHAP. 1

include the population at large. These discussions, in which one person posts a message and all the other subscribers to the newsgroup can read it, run the gamut from humorous to impassioned.

Our third category is entertainment, which is a huge and growing industry. The killer application here (the one that may drive all the rest) is video on demand. A decade or so hence, it may be possible to select any movie or television program ever made, in any country, and have it displayed on your screen instantly. New films may become interactive, where the user is occasionally prompted for the story direction (should MacBeth murder Duncan or just bide his time?) with alternative scenarios provided for all cases. Live television may also become interactive, with the audience participating in quiz shows, choosing among contestants, and so on.

On the other hand, maybe the killer application will not be video on demand. Maybe it will be game playing. Already we have multiperson real-time simulation games, like hide-and-seek in a virtual dungeon, and flight simulators with the players on one team trying to shoot down the players on the opposing team. If done with goggles and 3-dimensional real-time, photographic-quality moving images, we have a kind of worldwide shared virtual reality.

In short, the ability to merge information, communication, and entertainment will surely give rise to a massive new industry based on computer networking.

#### 1.1.3. Social Issues

The widespread introduction of networking will introduce new social, ethical, political problems (Laudon, 1995). Let us just briefly mention a few of them; a thorough study would require a full book, at least. A popular feature of many networks are newsgroups or bulletin boards where people can exchange messages with like-minded individuals. As long as the subjects are restricted to technical topics or hobbies like gardening, not too many problems will arise.

The trouble comes when newsgroups are set up on topics that people actually care about, like politics, religion, or sex. Views posted to such groups may be deeply offensive to some people. Furthermore, messages need not be limited to text. High-resolution color photographs and even short video clips can now easily be transmitted over computer networks. Some people take a live-and-let-live view, but others feel that posting certain material (e.g., child pornography) is simply unacceptable. Thus the debate rages.

People have sued network operators, claiming that they are responsible for the contents of what they carry, just as newspapers and magazines are. The inevitable response is that a network is like a telephone company or the post office and cannot be expected to police what its users say. Stronger yet, having network operators censor messages would probably cause them to delete everything with even the slightest possibility of their being sued, and thus violate their users' rights to free speech. It is probably safe to say that this debate will go on for a while.

Another fun area is employee rights versus employer rights. Many people read and write email at work. Some employers have claimed the right to read and possibly censor employee messages, including messages sent from a home termi-

nal after work. Not all employees agree with this (Sipior and Ward, 1995).

Even if employers have power over employees, does this relationship also govern universities and students? How about high schools and students? In 1994, Carnegie-Mellon University decided to turn off the incoming message stream for several newsgroups dealing with sex because the university felt the material was inappropriate for minors (i.e., those few students under 18). The fallout from this event will take years to settle.

Computer networks offer the potential for sending anonymous messages. In some situations, this capability may be desirable. For example, it provides a way for students, soldiers, employees, and citizens to blow the whistle on illegal behavior on the part of professors, officers, superiors, and politicians without fear of reprisals. On the other hand, in the United States and most other democracies, the law specifically permits an accused person the right to confront and challenge his accuser in court. Anonymous accusations cannot be used as evidence.

In short, computer networks, like the printing press 500 years ago, allow ordinary citizens to distribute their views in different ways and to different audiences than were previously possible. This new-found freedom brings with it many unsolved social, political, and moral issues. The solution to these problems is left as an exercise for the reader.

#### 1.2. NETWORK HARDWARE

It is now time to turn our attention from the applications and social aspects of networking to the technical issues involved in network design. There is no generally accepted taxonomy into which all computer networks fit, but two dimensions stand out as important: transmission technology and scale. We will now examine each of these in turn.

Broadly speaking, there are two types of transmission technology:

- 1. Broadcast networks.
- 2. Point-to-point networks.

Broadcast networks have a single communication channel that is shared by all the machines on the network. Short messages, called packets in certain contexts, sent by any machine are received by all the others. An address field within the packet specifies for whom it is intended. Upon receiving a packet, a machine checks the address field. If the packet is intended for itself, it processes the packet; if the packet is intended for some other machine, it is just ignored.

As an analogy, consider someone standing at the end of a corridor with many rooms off it and shouting "Watson, come here. I want you." Although the packet

7

may actually be received (heard) by many people, only Watson responds. The others just ignore it. Another example is an airport announcement asking all flight 644 passengers to report to gate 12.

Broadcast systems generally also allow the possibility of addressing a packet to *all* destinations by using a special code in the address field. When a packet with this code is transmitted, it is received and processed by every machine on the network. This mode of operation is called **broadcasting**. Some broadcast systems also support transmission to a subset of the machines, something known as **multicasting**. One possible scheme is to reserve one bit to indicate multicasting. The remaining n-1 address bits can hold a group number. Each machine can "subscribe" to any or all of the groups. When a packet is sent to a certain group, it is delivered to all machines subscribing to that group.

In contrast, **point-to-point** networks consist of many connections between individual pairs of machines. To go from the source to the destination, a packet on this type of network may have to first visit one or more intermediate machines. Often multiple routes, of different lengths are possible, so routing algorithms play an important role in point-to-point networks. As a general rule (although there are many exceptions), smaller, geographically localized networks tend to use broadcasting, whereas larger networks usually are point-to-point.

Interprocessor distance	Processors located in same	Example
0.1 m	Circuit board	Data flow machine
1 m	System	Multicomputer
10 m	Room	
100 m	Building	Local area network
1 km	Campus	
10 km	City	Metropolitan area network
100 km	Country	The state of the sail
1,000 km	Continent	Wide area network
10,000 km	Planet	The internet

Fig. 1-2. Classification of interconnected processors by scale.

An alternative criterion for classifying networks is their scale. In Fig. 1-2 we give a classification of multiple processor systems arranged by their physical size. At the top are **data flow machines**, highly parallel computers with many functional units all working on the same program. Next come the **multicomputers**, systems that communicate by sending messages over very short, very fast buses. Beyond the multicomputers are the true networks, computers that communicate

CHAP. 1

interface to the network (sockets) and wrote many application, utility, and management programs to make networking easier.

The timing was perfect. Many universities had just acquired a second or third VAX computer and a LAN to connect them, but they had no networking software. When 4.2BSD came along, with TCP/IP, sockets, and many network utilities, the complete package was adopted immediately. Furthermore, with TCP/IP, it was easy for the LANs to connect to the ARPANET, and many did.

By 1983, the ARPANET was stable and successful, with over 200 IMPs and hundreds of hosts. At this point, ARPA turned the management of the network over to the Defense Communications Agency (DCA), to run it as an operational network. The first thing DCA did was to separate the military portion (about 160 IMPs, of which 110 in the United States and 50 abroad) into a separate subnet, MILNET, with stringent gateways between MILNET and the remaining research subnet

During the 1980s, additional networks, especially LANs, were connected to the ARPANET. As the scale increased, finding hosts became increasingly expensive, so **DNS** (**Domain Naming System**) was created to organize machines into domains and map host names onto IP addresses. Since then, DNS has become a generalized, distributed database system for storing a variety of information related to naming. We will study it in detail in Chap. 7.

By 1990, the ARPANET had been overtaken by newer networks that it itself had spawned, so it was shut down and dismantled, but it lives on in the hearts and minds of network researchers everywhere. MILNET continues to operate, however.

#### **1.5.3. NSFNET**

50

By the late 1970s, NSF (the U.S. National Science Foundation) saw the enormous impact the ARPANET was having on university research, allowing scientists across the country to share data and collaborate on research projects. However, to get on the ARPANET, a university had to have a research contract with the DoD, which many did not have. This lack of universal access prompted NSF to set up a virtual network, CSNET, centered around a single machine at BBN that supported dial-up lines and had connections to the ARPANET and other networks. Using CSNET, academic researchers could call up and leave email for other people to pick up later. It was simple, but it worked.

By 1984 NSF began designing a high-speed successor to the ARPANET that would be open to all university research groups. To have something concrete to start with, NSF decided to build a backbone network to connect its six supercomputer centers, in San Diego, Boulder, Champaign, Pittsburgh, Ithaca, and Princeton. Each supercomputer was given a little brother, consisting of an LSI-11 microcomputer called a **fuzzball**. The fuzzballs were connected with 56 kbps leased lines and formed the subnet, the same hardware technology as the

ARPANET used. The software technology was different however: the fuzzballs spoke TCP/IP right from the start, making it the first TCP/IP WAN.

NSF also funded some (eventually about 20) regional networks that connected to the backbone to allow users at thousands of universities, research labs, libraries, and museums to access any of the supercomputers and to communicate with one another. The complete network, including the backbone and the regional networks, was called **NSFNET**. It connected to the ARPANET through a link between an IMP and a fuzzball in the Carnegie-Mellon machine room. The first NSFNET backbone is illustrated in Fig. 1-26.

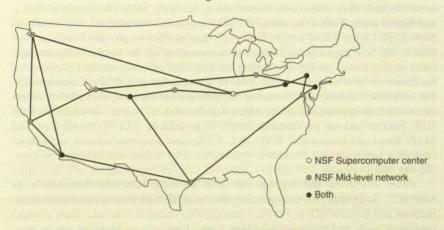


Fig. 1-26. The NSFNET backbone in 1988.

NSFNET was an instantaneous success and was overloaded from the word go. NSF immediately began planning its successor and awarded a contract to the Michigan-based MERIT consortium to run it. Fiber optic channels at 448 kbps were leased from MCI to provide the version 2 backbone. IBM RS6000s were used as routers. This, too, was soon overwhelmed, and by 1990, the second backbone was upgraded to 1.5 Mbps.

As growth continued, NSF realized that the government could not continue financing networking forever. Furthermore, commercial organizations wanted to join but were forbidden by NSF's charter from using networks NSF paid for. Consequently, NSF encouraged MERIT, MCI, and IBM to form a nonprofit corporation, ANS (Advanced Networks and Services) as a step along the road to commercialization. In 1990, ANS took over NSFNET and upgraded the 1.5-Mbps links to 45 Mbps to form ANSNET.

In December 1991, the U.S. Congress passed a bill authorizing NREN, the National Research and Educational Network, the research successor to NSFNET, only running at gigabits speeds. The goal was a national network

51

running at 3 Gbps before the millenium. This network is to act as a prototype for the much-discussed information superhighway.

CHAP. 1

By 1995, the NSFNET backbone was no longer needed to interconnect the NSF regional networks because numerous companies were running commercial IP networks. When ANSNET was sold to America Online in 1995, the NSF regional networks had to go out and buy commercial IP service to interconnect.

To ease the transition and make sure every regional network could communicate with every other regional network, NSF awarded contracts to four different network operators to establish a NAP (Network Access Point). These operators were PacBell (San Francisco), Ameritech (Chicago), MFS (Washington, D.C.), and Sprint (New York City, where for NAP purposes, Pennsauken, N.J. counts as New York City). Every network operator that wanted to provide backbone service to the NSF regional networks had to connect to all the NAPs. This arrangement meant that a packet originating on any regional network had a choice of backbone carriers to get from its NAP to the destination's NAP. Consequently, the backbone carriers were forced to compete for the regional networks' business on the basis of service and price, which was the idea, of course. In addition to the NSF NAPs, various government NAPs (e.g., FIX-E, FIX-W, MAE-East and MAE-West) and commercial NAPs (e.g., CIX) have also been created, so the concept of a single default backbone was replaced by a commercially-driven competitive infrastructure.

Other countries and regions are also building networks comparable to NSFNET. In Europe, for example, EBONE is an IP backbone for research organizations and EuropaNET is a more commercially oriented network. Both connect numerous cities in Europe with 2-Mbps lines. Upgrades to 34 Mbps are in progress. Each country in Europe has one or more national networks, which are roughly comparable to the NSF regional networks.

#### 1.5.4. The Internet

The number of networks, machines, and users connected to the ARPANET grew rapidly after TCP/IP became the only official protocol on Jan. 1, 1983. When NSFNET and the ARPANET were interconnected, the growth became exponential. Many regional networks joined up, and connections were made to networks in Canada, Europe, and the Pacific.

Sometime in the mid-1980s, people began viewing the collection of networks as an internet, and later as the Internet, although there was no official dedication with some politician breaking a bottle of champagne over a fuzzball.

Growth continued exponentially, and by 1990 the Internet had grown to 3000 networks and 200,000 computers. In 1992, the one millionth host was attached. By 1995, there were multiple backbones, hundreds of mid-level (i.e., regional) networks, tens of thousands of LANs, millions of hosts, and tens of millions of users. The size doubles approximately every year (Paxson, 1994).

Much of the growth comes from connecting existing networks to the Internet. In the past these have included SPAN, NASA's space physics network, HEPNET, a high energy physics network, BITNET, IBM's mainframe network, EARN, a European academic network now widely used in Eastern Europe, and many others. Numerous transatlantic links are in use, running from 64 kbps to 2 Mbps.

The glue that holds the Internet together is the TCP/IP reference model and TCP/IP protocol stack. TCP/IP makes universal service possible and can be compared to the telephone system or the adoption of standard gauge by the railroads in the 19th Century.

What does it actually mean to be on the Internet? Our definition is that a machine is on the Internet if it runs the TCP/IP protocol stack, has an IP address, and has the ability to send IP packets to all the other machines on the Internet. The mere ability to send and receive electronic mail is not enough, since email is gatewayed to many networks outside the Internet. However, the issue is clouded somewhat by the fact that many personal computers have the ability to call up an Internet service provider using a modem, be assigned a temporary IP address, and send IP packets to other Internet hosts. It make sense to regard such machines as being on the Internet for as long as they are connected to the service provider's router.

With exponential growth, the old informal way of running the Internet no longer works. In January 1992, the **Internet Society** was set up, to promote the use of the Internet and perhaps eventually take over managing it.

Traditionally, the Internet had four main applications, as follows:

- Email. The ability to compose, send, and receive electronic mail has been around since the early days of the ARPANET and is enormously popular. Many people get dozens of messages a day and consider it their primary way of interacting with the outside world, far outdistancing the telephone and snail mail. Email programs are available on virtually every kind of computer these days.
- 2. News. Newsgroups are specialized forums in which users with a common interest can exchange messages. Thousands of newsgroups exist, on technical and nontechnical topics, including computers, science, recreation, and politics. Each newsgroup has its own etiquette, style, and customs, and woe be to anyone violating them.
- 3. **Remote login**. Using the Telnet, Rlogin, or other programs, users anywhere on the Internet can log into any other machine on which they have an account.
- 4. File transfer. Using the FTP program, it is possible to copy files from one machine on the Internet to another. Vast numbers of articles, databases, and other information are available this way.

54

CHAP. 1

Up until the early 1990s, the Internet was largely populated by academic, government, and industrial researchers. One new application, the WWW (World Wide Web) changed all that and brought millions of new, nonacademic users to the net. This application, invented by CERN physicist Tim Berners-Lee, did not change any of the underlying facilities but made them easier to use. Together with the Mosaic viewer, written at the National Center for Supercomputer Applications, the WWW made it possible for a site to set up a number of pages of information containing text, pictures, sound, and even video, with embedded links to other pages. By clicking on a link, the user is suddenly transported to the page pointed to by that link. For example, many companies have a home page with entries pointing to other pages for product information, price lists, sales, technical support, communication with employees, stockholder information, and much more.

Numerous other kinds of pages have come into existence in a very short time, including maps, stock market tables, library card catalogs, recorded radio programs, and even a page pointing to the complete text of many books whose copyrights have expired (Mark Twain, Charles Dickens, etc.). Many people also have personal pages (home pages).

In the first year after Mosaic was released, the number of WWW servers grew from 100 to 7000. Enormous growth will undoubtedly continue for years to come, and will probably be the force driving the technology and use of the Internet into the next millenium.

Many books have been written about the Internet and its protocols. For more information, see (Black, 1995; Carl-Mitchell and Quarterman, 1993; Comer, 1995; and Santifaller, 1994).

#### 1.5.5. Gigabit Testbeds

The Internet backbones operate at megabit speeds, so for people who want to push the technological envelope, the next step is gigabit networking. With each increase in network bandwidth, new applications become possible, and gigabit networks are no exception. In this section we will first say a few words about gigabit applications, mention two of them, and then list some example gigabit testbeds that have been built.

Gigabit networks provide better bandwidth than megabit networks, but not always much better delay. For example, sending a 1-kbit packet from New York to San Francisco at 1 Mbps takes 1 msec to pump the bits out and 20 msec for the transcontinental delay, for a total of 21 msec. A 1-Gbps network can reduce this to 20.001 msec. While the bits go out faster, the transcontinental delay remains the same, since the speed of light in optical fiber (or copper wire) is about 200,000 km/sec, independent of the data rate. Thus for wide area applications in which low delay is critical, going to higher speeds may not help much. Fortunately, for

Unified Patents Ex. 1010, pg. 114

some applications, bandwidth is what counts, and these are the applications for which gigabit networks will make a big difference.

One application is telemedicine. Many people think that a way to reduce medical costs is to reintroduce family doctors and family clinics on a large scale, so everyone has convenient access to first line medical care. When a serious medical problem occurs, the family doctor can order lab tests and medical imaging, such as X-rays, CAT scans, and MRI scans. The test results and images can then be sent electronically to a specialist who then makes the diagnosis.

Doctors are generally unwilling to make diagnoses from computer images unless the quality of the transmitted image is as good as the original image. This requirement means images will probably need 4K × 4K pixels, with 8 bits per pixel (black and white images) or 24 bits per pixel (color images). Since many tests require up to 100 images (e.g., different cross sections of the organ in question), a single series for one patient can generate 40 gigabits. Moving images (e.g., a beating heart) generate even more data. Compression can help some but doctors are leary of it because the most efficient algorithms reduce image quality. Furthermore, all the images must be stored for years but may need to be retrieved at a moment's notice in the event of a medical emergency. Hospitals do not want to become computer centers, so off-site storage combined with high-bandwidth electronic retrieval is essential.

Another gigabit application is the virtual meeting. Each meeting room contains a spherical camera and one or more people. The bit streams from each of the cameras are combined electronically to give the illusion that everyone is in the same room. Each person sees this image using virtual reality goggles. In this way meetings can happen without travel, but again, the data rates required are stupendous.

Starting in 1989, ARPA and NSF jointly agreed to finance a number of university-industry gigabit testbeds, later as part of the NREN project. In some of these, the data rate in each direction was 622 Mbps, so only by counting the data going in both directions do you get a gigabit. This kind of gigabit is sometimes called a "government gigabit." (Some cynics call it a gigabit after taxes.) Below we will briefly mention the first five projects. They have done their job and been shut down, but deserve some credit as pioneers, in the same way the ARPANET does.

1. Aurora was a testbed linking four sites in the Northeast: M.I.T., the University of Pennsylvania, IBM's T.J. Watson Lab, and Bellcore (Morristown, N.J.) at 622 Mbps using fiber optics provided by MCI, Bell Atlantic, and NYNEX. Aurora was largely designed to help debug Bellcore's Sunshine switch and IBM's (proprietary) plaNET switch using parallel networks. Research issues included switching technology, gigabit protocols, routing, network control, distributed virtual memory, and collaboration using videoconferencing. For more information, see (Clark et al., 1993).

- 2. Blanca was originally a research project called XUNET involving AT&T Bell Labs, Berkeley, and the University of Wisconsin. In 1990 it added some new sites (LBL, Cray Research, and the University of Illinois) and acquired NSF/ARPA funding. Some of it ran at 622 Mbps, but other parts ran at lower speeds. Blanca was the only nationwide testbed; the rest were regional. Consequently, much of the research was concerned with the effects of speed-of-light delay. The interest here was in protocols, especially network control protocols, host interfaces, and gigabit applications such as medical imaging, meteorological modeling, and radio astronomy. For more information, see (Catlett, 1992; and Fraser, 1993).
- 3. CASA was aimed at doing research on supercomputer applications, especially those in which part of the problem ran best on one kind of supercomputer (e.g., a Cray vector supercomputer) and part ran best on a different kind of supercomputer (e.g., a parallel supercomputer). The applications investigated included geology (analyzing Landsat images), climate modeling, and understanding chemical reactions. It operated in California and New Mexico and connected Los Alamos, Cal Tech, JPL, and the San Diego Supercomputer Center.
- 4. Nectar differed from the three testbeds given above in that it was an experimental gigabit MAN running from CMU to the Pittsburgh Supercomputer Center. The designers were interested in applications involving chemical process flowsheeting and operations research, as well as the tools for debugging them.
- 5. VISTAnet was a small gigabit testbed operated in Research Triangle Park, North Carolina, and connecting the University of North Carolina, North Carolina State University, and MCNC. The interest here was in a prototype for a public switched gigabit network with switches having hundreds of gigabit lines, meaning that the switches had to be capable of processing terabits/sec. The scientific research focused on using 3D images to plan radiation therapy for cancer patients, with the oncologist being able to vary the beam parameters and instantaneously see the radiation dosages being delivered to the tumor and surrounding tissue (Ransom, 1992).

# 1.6. EXAMPLE DATA COMMUNICATION SERVICES

Telephone companies and others have begun to offer networking services to any organization that wishes to subscribe. The subnet is owned by the network operator, providing communication service for the customers' hosts and terminals.

56

CHAP. 1

CHAP. 5

provide transparent fragmentation of packets into cells and then reassembly of cells into packets. In the ATM world, fragmentation is called segmentation; the concept is the same, but some of the details are different.

Transparent fragmentation is simple but has some problems. For one thing, the exit gateway must know when it has received all the pieces, so that either a count field or an "end of packet" bit must be included in each packet. For another thing, all packets must exit via the same gateway. By not allowing some fragments to follow one route to the ultimate destination, and other fragments a disjoint route, some performance may be lost. A last problem is the overhead required to repeatedly reassemble and then refragment a large packet passing through a series of small-packet networks.

The other fragmentation strategy is to refrain from recombining fragments at any intermediate gateways. Once a packet has been fragmented, each fragment is treated as though it were an original packet. All fragments are passed through the exit gateway (or gateways), as shown in Fig. 5-41(b). Recombination occurs only at the destination host.

Nontransparent fragmentation also has some problems. For example, it requires *every* host to be able to do reassembly. Yet another problem is that when a large packet is fragmented the total overhead increases, because each fragment must have a header. Whereas in the first method this overhead disappears as soon as the small-packet network is exited, in this method the overhead remains for the rest of the journey. An advantage of this method, however, is that multiple exit gateways can now be used and higher performance can be achieved. Of course, if the concatenated virtual circuit model is being used, this advantage is of no use.

When a packet is fragmented, the fragments must be numbered in such a way that the original data stream can be reconstructed. One way of numbering the fragments is to use a tree. If packet 0 must be split up, the pieces are called 0.0, 0.1, 0.2, etc. If these fragments themselves must be fragmented later on, the pieces are numbered 0.0.0, 0.0.1, 0.0.2, ..., 0.1.0, 0.1.1, 0.1.2, etc. If enough fields have been reserved in the header for the worst case and no duplicates are generated anywhere, this scheme is sufficient to ensure that all the pieces can be correctly reassembled at the destination, no matter what order they arrive in.

However, if even one network loses or discards packets, there is a need for end-to-end retransmissions, with unfortunate effects for the numbering system. Suppose that a 1024-bit packet is initially fragmented into four equal-sized fragments, 0.0, 0.1, 0.2, and 0.3. Fragment 0.1 is lost, but the other parts arrive at the destination. Eventually, the source times out and retransmits the original packet again. Only this time the route taken passes through a network with a 512-bit limit, so two fragments are generated. When the new fragment 0.1 arrives at the destination, the receiver will think that all four pieces are now accounted for and reconstruct the packet incorrectly.

A completely different (and better) numbering system is for the internetwork protocol to define an elementary fragment size small enough that the elementary

I IRRARY II TIFI I IRRANA-CHAMPAIGN

fragment can pass through every network. When a packet is fragmented, all the pieces are equal to the elementary fragment size except the last one, which may be shorter. An internet packet may contain several fragments, for efficiency reasons. The internet header must provide the original packet number, and the number of the (first) elementary fragment contained in the packet. As usual, there must also be a bit indicating that the last elementary fragment contained within the internet packet is the last one of the original packet.

This approach requires two sequence fields in the internet header: the original packet number, and the fragment number. There is clearly a trade-off between the size of the elementary fragment and the number of bits in the fragment number. Because the elementary fragment size is presumed to be acceptable to every network, subsequent fragmentation of an internet packet containing several fragments causes no problem. The ultimate limit here is to have the elementary fragment be a single bit or byte, with the fragment number then being the bit or byte offset within the original packet, as shown in Fig. 5-42.

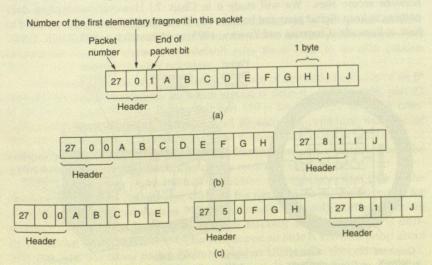


Fig. 5-42. Fragmentation when the elementary data size is 1 byte. (a) Original packet, containing 10 data bytes. (b) Fragments after passing through a network with maximum packet size of 8 bytes. (c) Fragments after passing through a size 5 gateway.

Some internet protocols take this method even further and consider the entire transmission on a virtual circuit to be one giant packet, so that each fragment contains the absolute byte number of the first byte within the fragment. Some other issues relating to fragmentation are discussed in (Kent and Mogul, 1987).

CHAP. 5

#### 547 Firewalls

The ability to connect any computer, anywhere, to any other computer, anywhere, is a mixed blessing. For individuals at home, wandering around the Internet is lots of fun. For corporate security managers, it is a nightmare. Most companies have large amounts of confidential information on-line—trade secrets, product development plans, marketing strategies, financial analyses, etc. Disclosure of this information to a competitor could have dire consequences.

In addition to the danger of information leaking out, there is also a danger of information leaking in. In particular, viruses, worms, and other digital pests (Kaufman et al., 1995) can breach security, destroy valuable data, and waste large amounts of administrators' time trying to clean up the mess they leave. Often they are imported by careless employees who want to play some nifty new game.

Consequently, mechanisms are needed to keep "good" bits in and "bad" bits out. One method is to use encryption. This approach protects data in transit between secure sites. We will study it in Chap. 7. However, encryption does nothing to keep digital pests and hackers out. To accomplish this goal, we need to look at firewalls (Chapman and Zwicky, 1995; and Cheswick and Bellovin, 1994).

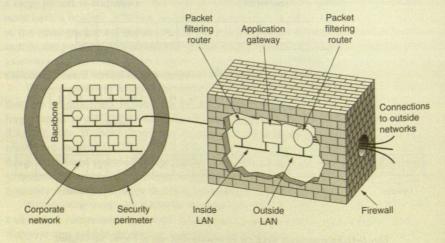


Fig. 5-43. A firewall consisting of two packet filters and an application gateway.

**Firewalls** are just a modern adaptation of that old medieval security standby: digging a deep moat around your castle. This design forced everyone entering or leaving the castle to pass over a single drawbridge, where they could be inspected by the I/O police. With networks, the same trick is possible: a company can have many LANs connected in arbitrary ways, but all traffic to or from the company is forced through an electronic drawbridge (firewall), as shown in Fig. 5-43.

IRRARY II THE IIRRANA-CHAMPAIGN

The firewall in this configuration has two components: two routers that do packet filtering and an application gateway. Simpler configurations also exist, but the advantage of this design is that every packet must transit two filters and an application gateway to go in or out. No other route exists. Readers who think that one security checkpoint is enough clearly have not made an international flight on a scheduled airline recently.

Each **packet filter** is a standard router equipped with some extra functionality. The extra functionality allows every incoming or outgoing packet to be inspected. Packets meeting some criterion are forwarded normally. Those that fail the test are dropped.

In Fig. 5-43, most likely the packet filter on the inside LAN checks outgoing packets and the one on the outside LAN checks incoming packets. Packets crossing the first hurdle go to the application gateway for further examination. The point of putting the two packet filters on different LANs is to ensure that no packet gets in or out without having to pass through the application gateway: there is no path around it.

Packet filters are typically driven by tables configured by the system administrator. These tables list sources and destinations that are acceptable, sources and destinations that are blocked, and default rules about what to do with packets coming from or going to other machines.

In the common case of a UNIX setting, a source or destination consists of an IP address and a port. Ports indicate which service is desired. For example, port 23 is for Telnet, port 79 is for Finger, and port 119 is for USENET news. A company could block incoming packets for all IP addresses combined with one of these ports. In this way, no one outside the company could log in via Telnet, or look up people using the Finger daemon. Furthermore, the company would be spared from having employees spend all day reading USENET news.

Blocking outgoing packets is trickier because although most sites stick to the standard port naming conventions, they are not forced to do so. Furthermore, for some important services, such as FTP (File Transfer Protocol), port numbers are assigned dynamically. In addition, although blocking TCP connections is difficult, blocking UDP packets is even harder because so little is known a priori about what they will do. Many packet filters simply ban UDP traffic altogether.

The second half of the firewall mechanism is the application gateway. Rather than just looking at raw packets, the gateway operates at the application level. A mail gateway, for example, can be set up to examine each message going in or coming out. For each one it makes a decision to transmit or discard it based on header fields, message size, or even the content (e.g., at a military installation, the presence of words like "nuclear" or "bomb" might cause some special action to be taken).

Installations are free to set up one or more application gateways for specific applications, but it is not uncommon for suspicious organizations to permit email in and out, and perhaps use of the World Wide Web, but ban everything else as

CHAP. 5

too dicey. Combined with encryption and packet filtering, this arrangement offers a limited amount of security at the cost of some inconvenience.

One final note concerns wireless communication and firewalls. It is easy to design a system that is logically completely secure, but which, in practice, leaks like a sieve. This situation can occur if some of the machines are wireless and use radio communication, which passes right over the firewall in both directions.

# 5.5. THE NETWORK LAYER IN THE INTERNET

At the network layer, the Internet can be viewed as a collection of subnetworks or **Autonomous Systems** (**ASes**) that are connected together. There is no real structure, but several major backbones exist. These are constructed from high-bandwidth lines and fast routers. Attached to the backbones are regional (midlevel) networks, and attached to these regional networks are the LANs at many universities, companies, and Internet service providers. A sketch of this quasihierarchical organization is given in Fig. 5-44.

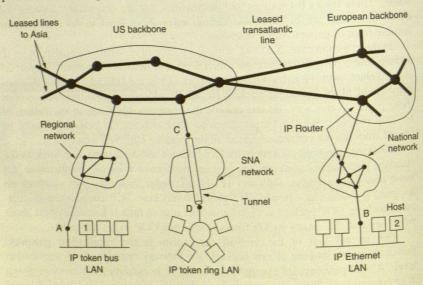


Fig. 5-44. The Internet is an interconnected collection of many networks.

The glue that holds the Internet together is the network layer protocol, IP (Internet Protocol). Unlike most older network layer protocols, it was designed from the beginning with internetworking in mind. A good way to think of the network layer is this. Its job is to provide a best-efforts way to transport datagrams

IRRANA-CHAMPAIGN

from source to destination, without regard to whether or not these machines are on the same network, or whether or not there are other networks in between them.

Communication in the Internet works as follows. The transport layer takes data streams and breaks them up into datagrams. In theory, datagrams can be up to 64 Kbytes each, but in practice they are usually around 1500 bytes. Each datagram is transmitted through the Internet, possibly being fragmented into smaller units as it goes. When all the pieces finally get to the destination machine, they are reassembled by the network layer into the original datagram. This datagram is then handed to the transport layer, which inserts it into the receiving process' input stream.

#### 5.5.1. The IP Protocol

IP

ned

netms An appropriate place to start our study of the network layer in the Internet is the format of the IP datagrams themselves. An IP datagram consists of a header part and a text part. The header has a 20-byte fixed part and a variable length optional part. The header format is shown in Fig. 5-45. It is transmitted in big endian order: from left to right, with the high-order bit of the *Version* field going first. (The SPARC is big endian; the Pentium is little endian.) On little endian machines, software conversion is required on both transmission and reception.

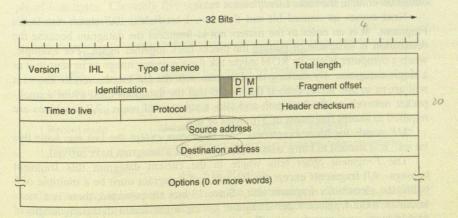


Fig. 5-45. The IP (Internet Protocol) header.

The *Version* field keeps track of which version of the protocol the datagram belongs to. By including the version in each datagram, it becomes possible to have the transition between versions take months, or even years, with some machines running the old version and others running the new one.

Since the header length is not constant, a field in the header, *IHL*, is provided to tell how long the header is, in 32-bit words. The minimum value is 5, which

Unified Patents Ex. 1010, pg. 122



# UNIVERSITY LIBRARY

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

504 | a Includes bibliographical references (p. 767-794) and index.

# Library Catalog

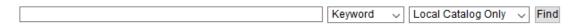
? ASK A LIBRARIAN FOR LIVE ASSISTANCE

What happened to the Library Catalog? Tell us what you think of the Library Catalog Your Accoun Keyword Logged in as Advanced Search | Classic Search | Course Reserves | E-Reserves | Search History Cite this Email this Add to favorites Staff view Computer networks / Andrew S. Tanenbaum. Main Tanenbaum, Andrew S. THIRD EDITION Author: COMPUTER NETWORKS Published: Upper Saddle River, N.J.: Prentice Hall PTR, c1996. Edition: 3rd ed. Topics: Computer networks. | Computacao (metodologia e tecnicas) | Processamento de dados | Redes de computadores e tecnicas | Computernetwerken. | Réseaux d'ordinateurs. | Protocoles de réseaux d'ordinateurs. | Réseaux d'ordinateurs -Architectures. | Rechnernetz. | Bilgisayar ağları. Add Tags: No Tags, Be the first to tag this record! More Location & Table of User Published Request Availability Reviews Details Contents Reviews Item 00002253cam a2200529 a 4500 0013836295 003UIUdb 00520120209200858.0 008960102s1996 njuab b 001 0 eng 010 a 96004121 015 aGB97-28628 019|a36916242|a50759000|a638929021 020 a0133499456 020 a9780133499452 020 a0133942481 (pbk.) 020 a9780133942484 (pbk.) 035 | a(OCoLC)ocm34079009 037 a 57642 b TVG 040|aDLC|cDLC|dUKM|dUBA|dNLGGC|dBTCTA|dYDXCP|dLVB|dOCLCG|dAU@|dEYR|dSINLB|dHEBIS|dDEBBG|dTVG|dL 049 aUIUU 05000 aTK5105.5 b.T36 1996 08200 a004.6 221 084 a54.32 2bcl 084 aDAT 250f 2stub 1001 | aTanenbaum, Andrew S., | d1944-24510 a Computer networks / c Andrew S. Tanenbaum. 250 a3rd ed. 260 aUpper Saddle River, N.J.: bPrentice Hall PTR, cc1996. 300 axvii, 813 p. : |bill., maps ; c24 cm.

50500|g1.|tIntroduction --|g2.|tPhysical Layer --|g3.|tData Link Layer --|g4.|tMedium Access Sublayer

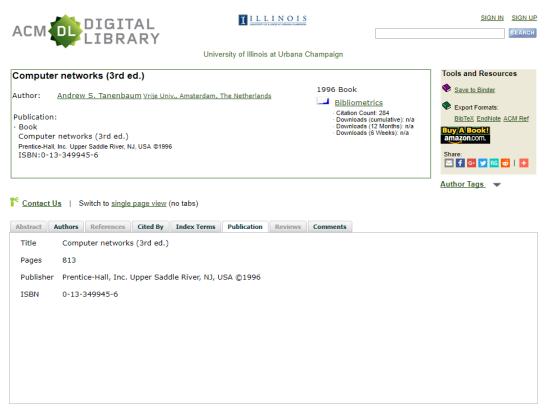
# Attachment 2b: University of Illinois at Urbana-Champaign Library catalog record for Document 2

520 | a I his classic reference for students, and anyone who wants to know more about connectivity, has been totally rewritten to reflect the networks of the 1990s and beyond. 6500 a Computer networks. 6507 | a Computação (metodologia e tecnicas) | 2 larpcal 6507 | a Processamento de dados | 2 larpcal 6507 | a Redes de computadores e tecnicas | 2 larpcal 65017 a Computernetwerken. 2gtt 6506 | a Réseaux d'ordinateurs. 6506 | a Protocoles de réseaux d'ordinateurs. 6506 | a Réseaux d'ordinateurs | x Architectures. 65007 a Rechnernetz. 2 swd 65004 aBilgisayar ağları. 77608 iOnline version: aTanenbaum, Andrew S., 1944- tComputer networks. b3rd ed. dUpper Saddle River, N.J.: Prent Hall PTR, c1996 | w(OCoLC)604516214 938 aBaker and Taylor bBTCP n96004121 938 a YBP Library Services bYANK n69989 938 aBaker & Taylor bBKTY c52.19 d52.19 i0133942481 n0003515003 sactive 994 a02 bUIU



Advanced Search | Classic Search | Course Reserves | E-Reserves | Search History

# Attachment 2c: ACM Digital Library index record for Document 2

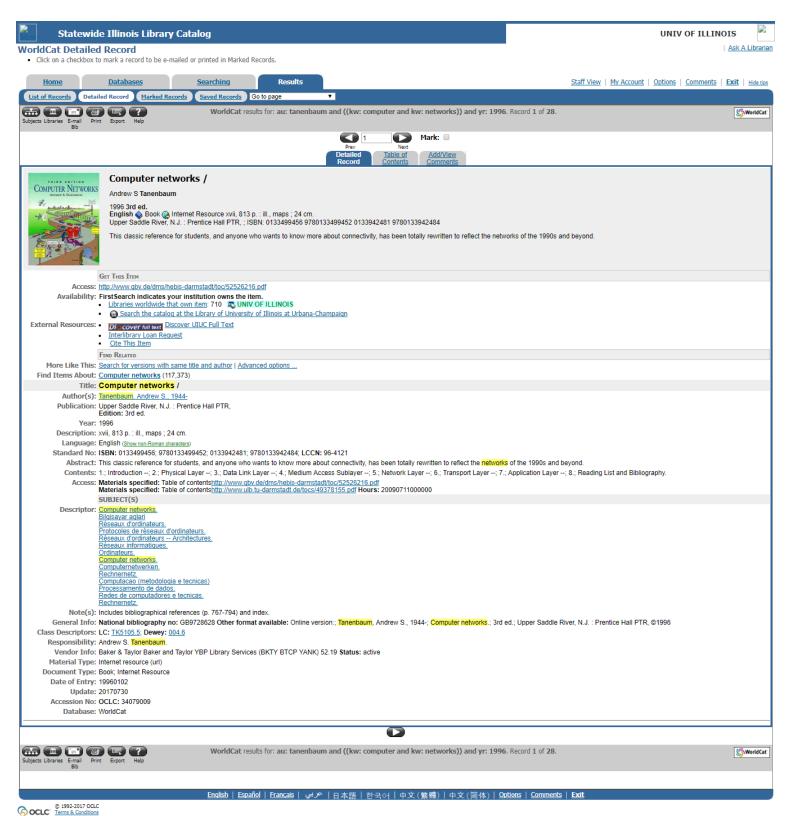


Powered by THE ACM GUIDE TO COMPUTING LITERATURE

The ACM Digital Library is published by the Association for Computing Machinery. Copyright © 2017 ACM, Inc.

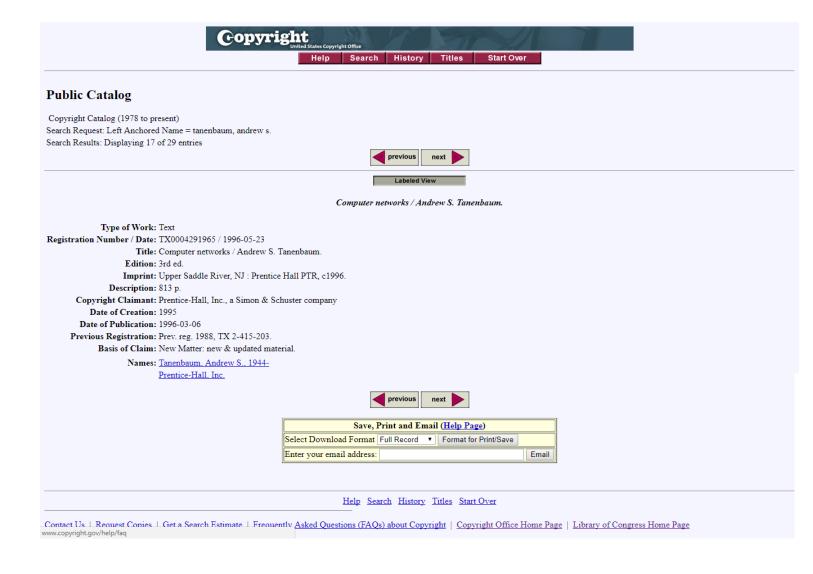
<u>Terms of Usage Privacy Policy Code of Ethics Contact Us</u>

## Attachment 2d: Statewide Illinois Library Catalog record (1) for Document 2



0 0020

# Attachment 2e: United States Copyright Office record for Document 2



## Attachment 2f: Statewide Illinois Library Catalog record (2) for Document 2

