



US007103594B1

(12) **United States Patent**
Wolfe

(10) **Patent No.:** **US 7,103,594 B1**

(45) **Date of Patent:** ***Sep. 5, 2006**

(54) **SYSTEM AND METHOD FOR
INFORMATION RETRIEVAL EMPLOYING A
PRELOADING PROCEDURE**

(76) Inventor: **Mark A. Wolfe**, 1076 Tamberwood Ct.,
Woodbury, MN (US) 55125

(*) Notice: Subject to any disclaimer, the term of this
patent is extended or adjusted under 35
U.S.C. 154(b) by 468 days.

This patent is subject to a terminal dis-
claimer.

(21) Appl. No.: **10/611,077**

(22) Filed: **Jul. 1, 2003**

Related U.S. Application Data

(63) Continuation of application No. 09/974,242, filed on
Oct. 9, 2001, now Pat. No. 6,604,103, which is a
continuation of application No. 09/620,651, filed on
Jul. 20, 2000, now abandoned, which is a continua-
tion of application No. 09/083,382, filed on May 22,
1998, now abandoned, which is a continuation-in-part
of application No. 08/918,912, filed on Aug. 27, 1997,
now Pat. No. 5,946,682, which is a continuation of
application No. 08/474,921, filed on Jun. 7, 1995,
now Pat. No. 5,715,445, which is a continuation of
application No. 08/300,343, filed on Sep. 2, 1994,
now abandoned, said application No. 09/083,382.

(60) Provisional application No. 60/047,554, filed on May
22, 1997.

(51) **Int. Cl.**
G06F 17/30 (2006.01)

(52) **U.S. Cl.** **707/5; 707/10**

(58) **Field of Classification Search** 707/3,
707/5, 10
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,723,211 A	2/1988	Barker et al.	364/300
4,730,252 A	3/1988	Bradshaw	364/403
4,760,606 A	7/1988	Lesnick et al.	382/48
4,815,029 A	3/1989	Barker et al.	364/900
4,839,853 A	6/1989	Deerwester et al.	364/900
4,855,725 A	8/1989	Fernandez	345/173
4,860,197 A	8/1989	Langendorf	
4,899,292 A	2/1990	Montagna	707/501
4,931,935 A	6/1990	Ohira	364/419
4,943,908 A	7/1990	Emma	
4,945,476 A	7/1990	Bodick et al.	364/413.02
4,954,969 A	9/1990	Tsumura	364/521

(Continued)

OTHER PUBLICATIONS

Roberts, "Controlling Internet Explorer 4.0 with Browser Helper
Objects," Microsoft Interactive Developer, May 1998, p. 66.

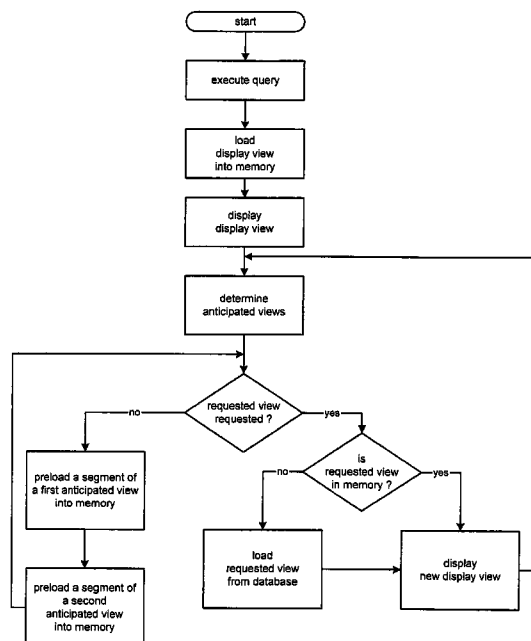
(Continued)

Primary Examiner—Chun Cao

(57) **ABSTRACT**

A document retrieval system having improved response
time. During the time the user spends viewing the displayed
information, other information that the user is likely to read
or study later is preloaded into memory. If the user later
requests the preloaded information, it can be written to the
display very quickly. As a result, the user's request to view
new information can be serviced quickly.

29 Claims, 43 Drawing Sheets



U.S. PATENT DOCUMENTS

4,980,823 A	12/1990	Liu		5,822,537 A	10/1998	Katseff	
4,982,344 A	1/1991	Jordan	364/521	5,822,539 A	10/1998	van Hoff	395/200.66
4,996,642 A	2/1991	Hey	705/27	5,822,731 A	10/1998	Schultz	704/256
5,014,327 A	5/1991	Potter		5,826,025 A	10/1998	Gramlich	395/200.47
5,021,989 A	6/1991	Fujisawa et al.	364/900	5,826,261 A	10/1998	Spencer	707/5
5,062,074 A	10/1991	Kleinberger	364/900	5,826,267 A	10/1998	McMillan	707/9
5,122,951 A	6/1992	Kamiya	364/419	5,826,269 A	10/1998	Hussey	707/10
5,123,088 A	6/1992	Kasahara et al.	395/600	5,835,922 A	11/1998	Shima	707/522
5,157,783 A	10/1992	Anderson et al.	395/600	5,860,074 A	1/1999	Rowe	707/526
5,206,949 A	4/1993	Cochran et al.	395/600	5,864,850 A	1/1999	Nordman	707/10
5,220,648 A	6/1993	Sato	395/146	5,870,552 A	2/1999	Dozier	395/200.49
5,235,680 A	8/1993	Bijnagte	395/161	5,870,770 A	2/1999	Wolfe	707/501
5,241,671 A	8/1993	Reed et al.	395/600	5,873,076 A	2/1999	Barr	707/3
5,243,149 A	9/1993	Comerford et al.	178/18	5,877,760 A	3/1999	Onda	345/341
5,253,337 A	10/1993	Hirose	395/161	5,878,421 A	3/1999	Ferrel	707/100
5,255,386 A	10/1993	Prager	395/600	5,890,172 A	3/1999	Borman	707/501
5,257,367 A	10/1993	Goodlander	395/600	5,893,914 A	4/1999	Clapp	707/507
5,261,071 A	11/1993	Lyon	395/425	5,905,492 A	5/1999	Straub	345/333
5,265,065 A	11/1993	Turtle	395/600	5,913,040 A	6/1999	Rakavy	395/200.62
5,289,569 A	2/1994	Taniguchi	395/145	5,915,256 A	6/1999	Rogers	707/501
5,295,243 A	3/1994	Robertson	395/160	5,933,811 A	8/1999	Angles	705/14
5,301,109 A	4/1994	Landauer et al.	364/419.19	5,946,682 A *	8/1999	Wolfe	707/5
5,305,389 A	4/1994	Palmer		5,948,040 A	9/1999	DeLorme	701/201
5,325,298 A	6/1994	Gallant	364/419.19	5,948,061 A	9/1999	Merriman et al.	709/219
5,341,293 A	8/1994	Vertelney	395/600	5,951,620 A	9/1999	Ahrens	701/200
5,345,551 A	9/1994	Shelley	395/157	5,953,005 A	9/1999	Liu	345/302
5,347,632 A *	9/1994	Filepp et al.	709/202	5,959,621 A	9/1999	Nawaz	345/329
5,390,281 A	2/1995	Luciw et al.	395/12	5,959,623 A	9/1999	van Hoff	345/333
5,392,387 A	2/1995	Fitzpatrick et al.	395/156	5,960,409 A	9/1999	Wexler	705/14
5,404,442 A	4/1995	Foster	395/159	5,963,205 A	10/1999	Sotomayor	345/333
5,408,655 A	4/1995	Oren et al.	395/600	RE36,422 E	11/1999	Pazel	707/104
5,414,811 A	5/1995	Parulski		5,995,976 A	11/1999	Walker	707/104
5,418,948 A	5/1995	Turtle	395/600	6,006,252 A	12/1999	Wolfe	709/203
5,421,008 A	5/1995	Banning et al.	395/600	6,011,537 A	1/2000	Slotznick	345/115
5,444,615 A	8/1995	Bennett		6,012,053 A	1/2000	Pant	707/3
5,471,575 A	11/1995	Giansante	395/144	6,023,698 A	2/2000	Lavey	707/10
5,485,609 A	1/1996	Vitter		6,023,726 A	2/2000	Saksena	709/219
5,495,581 A	2/1996	Tsai	395/154	6,034,680 A	3/2000	Kessenich	345/329
5,511,160 A	4/1996	Robson	395/162	6,038,601 A	3/2000	Lambert et al.	709/226
5,524,193 A	6/1996	Covington	707/512	6,094,675 A	7/2000	Sunaga	709/204
5,526,520 A	6/1996	Krause	707/104	6,138,128 A	10/2000	Perkowitz	707/501
5,535,422 A	7/1996	Chiang	395/155	6,151,021 A	11/2000	Berquist	345/339
5,544,352 A	8/1996	Egger	395/600	6,151,603 A *	11/2000	Wolfe	707/10
5,572,643 A	11/1996	Judson	395/793	6,182,068 B1	1/2001	Culliss	
5,617,526 A	4/1997	Oran et al.	395/326	6,216,141 B1	4/2001	Straub	707/513
5,619,247 A	4/1997	Russo	348/3	6,226,655 B1	5/2001	Borman	707/501
5,623,679 A	4/1997	Rivette et al.	395/773	6,263,351 B1	7/2001	Wolfe	707/501
5,632,022 A	5/1997	Warren	345/350	6,263,505 B1	7/2001	Walker	725/110
5,644,686 A	7/1997	Hekmatpur	706/45	6,282,548 B1	8/2001	Burner	707/104
5,675,788 A	10/1997	Husick	395/615	6,292,813 B1	9/2001	Wolfe	707/513
5,692,107 A	11/1997	Simudis	395/50	6,301,576 B1	10/2001	Wolfe	707/5
5,696,965 A	12/1997	Dedrick	395/610	6,321,991 B1	11/2001	Knowles	235/472.01
5,706,452 A	1/1998	Ivanov	395/329	6,336,131 B1	1/2002	Wolfe	709/203
5,706,507 A	1/1998	Schloss	395/615	6,341,305 B1	1/2002	Wolfe	709/203
5,714,971 A	2/1998	Shalit	345/119	6,366,947 B1	4/2002	Kavner	
5,715,445 A *	2/1998	Wolfe	707/5	6,539,377 B1	3/2003	Culliss	
5,717,914 A	2/1998	Husick	395/605	6,604,103 B1 *	8/2003	Wolfe	707/5
5,721,902 A	2/1998	Schultz	395/604	6,816,850 B1	11/2004	Culliss	
5,729,594 A	3/1998	Klingman	379/93.12				
5,737,734 A	4/1998	Schultz	707/5				
5,742,284 A	4/1998	Duggan	345/335				
5,742,816 A	4/1998	Barr	395/615				
5,754,766 A	5/1998	Shaw					
5,765,138 A	6/1998	Aycock	705/7				
5,768,578 A	6/1998	Kirk	707/100				
5,778,398 A	7/1998	Nagashima	707/501				
5,781,189 A	7/1998	Holleran	345/335				
5,801,702 A	9/1998	Dolan	345/357				
5,802,292 A	9/1998	Mogul	395/200.33				
5,806,077 A	9/1998	Wecker	707/501				
5,808,611 A	9/1998	Johnson	345/349				

OTHER PUBLICATIONS

Hannon, "New Window on Web Ads," PC Week, Feb. 2, 1998.
 Netscape Communicator Plug-in Guide (Netscape Communications Corporation Jan. 1998).
 Finnie, "The Browser Push Is On," PC Magazine, Sep. 9, 1997, pp. 123-128.
 Dragan, "Advice From The Web," PC Magazine, Sep. 9, 1997, pp. 133-144.
 Bremser, "HTML Workshop: Making Frames Fit," Internet World, Sep. 1997, pp. 90-92.
 Rapoza, "Alexa's Theory of Relativity," PC WEEK, Aug. 18, 1997, p. 42.
 Finnie, "Not Just Browsing," PC Magazine, Aug. 1997, pp. 45-54.

- "Internet 101: Web Rings," CNET.COM, Jul. 17, 1997 (www.cnet.com).
- Bannan, "Navigating Made Easy," PC Magazine, Jul. 1997, p. 81.
- "And Now, the Webtop," PC Magazine, Jun. 24, 1997, p. 9.
- Blight, "Annotated Reference List Agents," IEEE 1997 Conference on Communications, Power, and Computing WESCANEX97 Proceedings, May 22, 1997, pp. 7-12.
- Randal, "Web Central," PC Magazine, Apr. 22, 1997, p. 65.
- Clyman, Mendelson, and Nadel, "In Your Face Browsing," PC Magazine, Apr. 22, 1997, pp. 45-48.
- Shankar, "Blaze browsing tool packs punch," Infoworld, Apr. 14, 1997 (available at <http://www.infoworld.com/cgi-bin/displayArchives.pl?97-iw01-15.72c.htm>).
- Sirapyan, "Pushing Content Along," PC Magazine, Apr. 8, 1997, p. 40.
- Munro, "Internet Filtering Utilities," PC Magazine, Apr. 8, 1997, pp. 235-240.
- Mendelson, "Internet Off-line Search Utilities," PC Magazine, Apr. 8, 1997, pp. 227-232.
- Mendelson, "Internet Off-line Browsers," PC Magazine, Apr. 8, 1997, pp. 207-218.
- InterCast Press Releases, Apr. 7-8, 1997 (from www.intercast.com).
- Frelehoux, "An Architecture to Support Personalized Web Applications," Apr. 7, 1997 (available at <http://www.scope.gmd.de/info/www6/posters/726/poster726.html>).
- Robertson, "A Personalized Web," Internet World, Apr. 1997, pp. 32-34.
- Bell, "Make Java Fast: Optimize!," JavaWorld Apr. 1997 (JavaWorld 1997) (available at <http://www.javaworld.com/>).
- Lidsky, "The Web Delivers," PC Magazine, Feb. 18 1997, pp. 128-143.
- Proffit, "InterCast Brings the Web to TV," PC Magazine, Jan. 21, 1997, pp. 203-204.
- Isaacs, Inside Dynamic HTML (Microsoft Press 1997).
- Moeller, "Caching Software Speeds Browsing," PC Week, Dec. 2, 1996, p. 33.
- Lidsky, "Personal Webcasting," PC Magazine, Dec. 17, 1996, p. 40.
- Livingston, "Will Too Many Users Destroy the Internet? I Hardly Think So," Infoworld, Dec. 9, 1996, p. 35.
- Sun Microsystems, Inc., The JavaBeans 1.0 API Specification (Sun Microsystems 1996) (available at <http://java.sun.com/beans>).
- Singh and Lidsky, "All-Out Search," PC Magazine, Dec. 3, 1996, pp. 213-249.
- Vanhelsuwe, "How to Make Java Applets Start Faster," NetscapeWorld, Dec. 1996 (from www.netscapeworld.com).
- Pulver, "A Personal Cast," Internet World, Dec. 1996, pp. 110-112.
- Haskin, "Clearing a Path on the Web," PC Magazine, Nov. 19, 1996, p. 58.
- Mendelson, "The Finder of Lost Sites," PC Magazine, Nov. 5, 1996, p. 48.
- Livingston, "Java-based app speeds Web pages to your browser," Infoworld, Nov. 4, 1996 (available at <http://www.infoworld.com/cgi-bin/displayNew.pl?/livingst/bl110496.htm>).
- "Internet Performance Products: Peak Net.Jet" (available at <http://www.peak-media.com/netjet/netjet.html>), Dec. 29, 1996.
- Levin, "Searching for Meaning," PC Magazine, Sep. 24, 1996, p. 36.
- Perenson, "Your Web Surfing Companion," PC Magazine, Sep. 10, 1996, p. 73.
- Levin, "New Web Tool Reads Minds," PC Magazine, Sep. 10, 1996, p. 36.
- Venditto, "Safe Computing," Internet World, Sep. 1996, pp. 49-58.
- Metz, "PC Broadcast News: PointCast Delivers over the Internet," PC Magazine, Jul. 1996, p. 60.
- Perenson, "Make a Dash for the Web," PC Magazine, Jun. 11, 1996, p. 73.
- Canter, "Internet Search Tools," PC Magazine, May 14, 1996, pp. 207-210.
- Venditto, "Search Engine Showdown," Internet World, May 1996, pp. 79-86.
- Ayre and Mace, "Internet Access: Just Browsing," PC Magazine, Mar. 12, 1996, pp. 100-146.
- Siegel, CORBA Fundamentals and Programming, (John Wiley & Sons 1996).
- Prosisie, Programming Windows 95 with MFC (Microsoft Press 1996).
- Petzold, Programming Windows 95 (Microsoft 1996).
- Lemay, Official Marimba Guide to Castanet, (SamsNet 1997).
- Lemay, Java in 21 Days (SamsNet 1996).
- Graham, HTML Sourcebook (John Wiley & Sons 1996).
- Danesh, JavaScript in a Week (Sams.net 1996).
- Chappell, Understanding ActiveX and OLE, (Microsoft Press 1996).
- Chapman, Building Internet Applications with Delphi 2 (Que 1996).
- Cantu, Mastering Delphi 3, (2d ed. Sybex 1996).
- Campione and Walrath, The Java Tutorial (Addison Wesley 1996).
- "Blaze Web Performance Pack: xSpeed Technology," (from <http://www.xspeed.com/xspeed.htm>), Dec. 30, 1996.
- Kovel, Quirk, & Gabin, The Lotus Notes Idea Book (Addison-Wesley 1996).
- Nick Winfield, "Web-wise apps," Infoworld, Aug. 28, 1995 pp. 1, 20.
- Bederson, "Pad plus plus: A Zoomable Graphical Interface System," Proceedings of the Conference on Human Factors in Computing Systems, May 1995, pp. 23-24.
- Andrews, "Hyper-G and Harmony: Towards the Next Generation of Networked Information Technology," Proceedings of the Conference on Human Factors in Computing Systems, May 1995, pp. 33-34.
- Roscheisen, "Beyond Browsing: Shared Comments, SOAPS, Trails, and Online Communities," Computer Networks and ISDN Systems, Apr. 10, 1995.
- Huser, "Knowledge-based Editing and Visualization for Hypermedia Encyclopedias," Communications of the Association of Computing Machinery, v. 38, Apr. 1995, pp. 49-51.
- "Netscape's DDE Implementation," (from www.netscape.com) Mar. 22, 1995.
- Grunin, "Publish Without Paper," PC Magazine, Feb. 7, 1995, pp. 110-171.
- Ayre & Reichard, "The Web Untangled," PC Magazine, Feb. 7, 1995, pp. 173-196.
- The End of the Beginning (pamphlet) (1995 LEXIS/NEXIS).
- Software Products: Prepare to Practice (pamphlet) (1995 LEXIS/NEXIS).
- Jamsa, Internet Programming (Jamsa Press 1995).
- Discovering Westlaw: The Essential Guide (4th ed.) (1995 West Publishing Company).
- Denning, OLE Controls Inside Out (Microsoft 1995).
- Brown, Using Netscape 2, pp. 773-786 (Que 1995).
- Brockschmidt, Inside OLE (2d ed. Microsoft Press 1995).
- "More" Product Literature (two pages), (© 1995 LEXIS/NEXIS).
- Joel Snyder, "Taming the Internet," MacWorld, Dec. 1994, pp. 115-117.
- Mark Brownstein, "Fast Drives, Tiny Packages," PC/Computing, Nov. 1994, pp. 140-141.
- Fowler, "Experience with the Virtual Notebook System: Abstraction in Hypertext," Transcending Boundaries, Proceedings of the Conference on Computer Supported Cooperative Work, Oct. 1994, pp. 133-143.
- Robertson, "The Hypermedia Authoring Research Toolkit (HART)," Proceedings of European Conference on Hypermedia Technology, Sep. 1994, pp. 177-185.
- Bieber, "Backtracking in a Multiple-Window Hypertext Environment," ACM European Conference on Hypermedia Technology, Sep. 1994, pp. 158-166.
- Using Mosaic (Que Corporation 1994), pp. 31-44 & 80-85.
- Shepard's United States Citations: CD-ROM Edition, © 1994 McGraw-Hill, Inc.
- PatentWorks Workbench User's Guide, Waverly Systems, Inc., pp. 1-6, 67-95, and Quick Reference Card, © 1994 Waverly Systems, Inc.
- Freestyle Quick Reference (pamphlet) (1995 LEXIS/NEXIS).
- Comer, Internetworking with TCP/IP, vols. 1, 2 & 3 (Prentice Hall 1994-1997).

- Instone, "Empirically-based Redesign of a Hypertext Encyclopedia," Conference Proceedings on Human Factors in Computing Systems, Apr. 1993, pp. 500-506.
- Questions & Answers (to be used with "How to Shepardize") (1993 McGraw-Hill, Inc.).
- How to Shepardize, © 1993 McGraw-Hill.
- Foss, "Tools for Reading and Browsing Hypertext," Information Processing & Management, vol. 25, No. 4, pp. 405-418, 1988.
- Tanenbaum, Computer Networks (2d ed. Prentice Hall 1981).
- U.S. Appl. No. 09/620,651, filed Jul. 20, 2000 (inventor: Mark Wolfe).
- U.S. Appl. No. 10/021,761, filed Dec. 12, 2001 (inventor: Mark Wolfe).
- U.S. Appl. No. 09/258,873, filed Feb. 26, 1999 (inventor: Mark Wolfe).
- U.S. Appl. No. 09/903,923, filed Jul. 12, 2001 (inventor: Mark Wolfe).
- U.S. Appl. No. 09/974,242, filed Oct. 9, 2001, now U.S. Patent No. 6,604,103 (inventor: Mark Wolfe).
- U.S. Appl. No. 09/083,382, filed May 22, 1998 (inventor: Mark Wolfe).
- U.S. Appl. No. 60/047,554, filed May 22, 1997 (inventor: Mark Wolfe).
- U.S. Appl. No. 09/784,469, filed Feb. 16, 2001 (inventor: Mark Wolfe).
- U.S. Appl. No. 09/912,932, filed Jul. 26, 2001 (inventor: Mark Wolfe).
- U.S. Appl. No. 08/918,912, filed Aug. 27, 1997, now U.S. Patent No. 5,946,682 (inventor: Mark Wolfe).
- U.S. Appl. No. 08/474,921, filed Jun. 7, 1995, now U.S. Patent No. 5,715,445 (inventor: Mark Wolfe).
- U.S. Appl. No. 08/300,343, filed Sep. 2, 1994 (inventor: Mark Wolfe).
- U.S. Appl. No. 10/634,318, Mark Wolfe, filed Aug. 5, 2003 (continuation-in-part of U.S. Patent No. 6,292,813).
- U.S. Appl. No. 10/652,670, Mark Wolfe, filed Aug. 29, 2003 (continuation of U.S. Patent No. 6,263,351).
- U.S. Appl. No. 10/652,670, Wolfe, filed Aug. 29, 2003.
- U.S. Appl. No. 10/634,318, Wolfe, filed Aug. 5, 2003.
- U.S. Appl. No. 10/021,761, Wolfe, filed Dec. 12, 2001.
- Hawblitzel, "SLK: A Capability System Based on Safe Language Technology," Cornell University, Apr. 15, 1997.
- Joachims, "WebWatcher: A Tour Guide for the World Wide Web," Carnegie Mellon University, Sep. 1996.
- Cao, "Implementation and Performance of Integrated Application-Controlled File Caching, Prefetching, and Disk Scheduling," 1996 no date on document, but believed to be published in 1996.
- Padmanabhan, "Using Predictive Prefetching to Improve World Wide Web Latency," 1996 no date on document, but believed to be published in 1996.
- Wang, "Prefetching in World Wide Web," University College London, 1996 no date on document, but believed to be published in 1996.
- Vitter, "Optimal Prefetching via Data Compression," Duke University, Nov. 13, 1995.
- Davis, "Shared Annotation for Cooperative Learning," Xerox Corporation, Design Research Institute, Sep. 1995.
- Lieberman, "Letizia: An Agent That Assists Web Browsing," Massachusetts Institute of Technology, 1995 no date on document, but believed to be published in 1995.
- Roscheisen, "Beyond Browsing: Shared Comments, SOAPs, Trails, and On-line Communities," Stanford University, 1995 no date on document, but believed to be published in 1995.
- Pazzani, "Learning from Hotlists and Coldlists," University of California—Irvine, 1995 no date on document, but believed to be published in 1995.
- Roscheisen, "Shared Web Annotations as a Platform for Third-Party Value-Added, Information Providers," Stanford University, 1995 no date on document, but believed to be published in 1995.
- Griffioen, "Reducing File System Latency Using a Predictive Approach," University of Kentucky, Jun. 1994.
- Davis, "Believable Advice," Cornell University, Feb. 2, 1994.
- Davis, "A Server for Distributed Digital Technical Report Library," Xerox Corporation, Jan. 15, 1994.
- Cao, "Application-Controlled File Caching Policies," Princeton University, 1994 no date on document, but believed to be published in 1994.
- Glassman, "A Caching Relay for the World Wide Web," Digital Equipment Corporation, 1994 no date on document, but believed to be published in 1994.
- Davis, "Drop-in publishing with the World Wide Web," Xerox Inc. and Cornell University, 1994 no date on document, but believed to be published in 1994.
- Griffioen, "Automatic Prefetching in a WAN," University of Kentucky, Oct. 1993.
- Patterson, "A Status Report on Research in Transparent Informed Prefetching," Carnegie Mellon University, Feb. 1993.
- Wiil, "Issues in the Design of EHTS: A Multiuser Hypertext System for Collaboration," The University of Aalborg, 1992 no date on document, but believed to be published in 1992.
- CheckCite 3.0, The Definitive Citation Checker, User's Manual, by Jurisoft Mead Data Central, Inc., 1992, pp. 11-33, 67-81, 83-112, 127-129, and 131.
- Vitter, "Optimal Prefetching via Data Compression," Brown University, Jul. 1991.
- Tait, "Detection and Exploitation of File Working Sets," Columbia University, 1991 no date on document, but believed to be published in 1991.
- Ousterhout, "Why Aren't Operating Systems Getting Faster as Fast as Hardware?," Digital Western Research Laboratory, Oct. 1989.
- Nelson, "Caching in the Sprite Network File System," University of California—Berkeley, 1988 no date on document, but believed to be published in 1988.

* cited by examiner

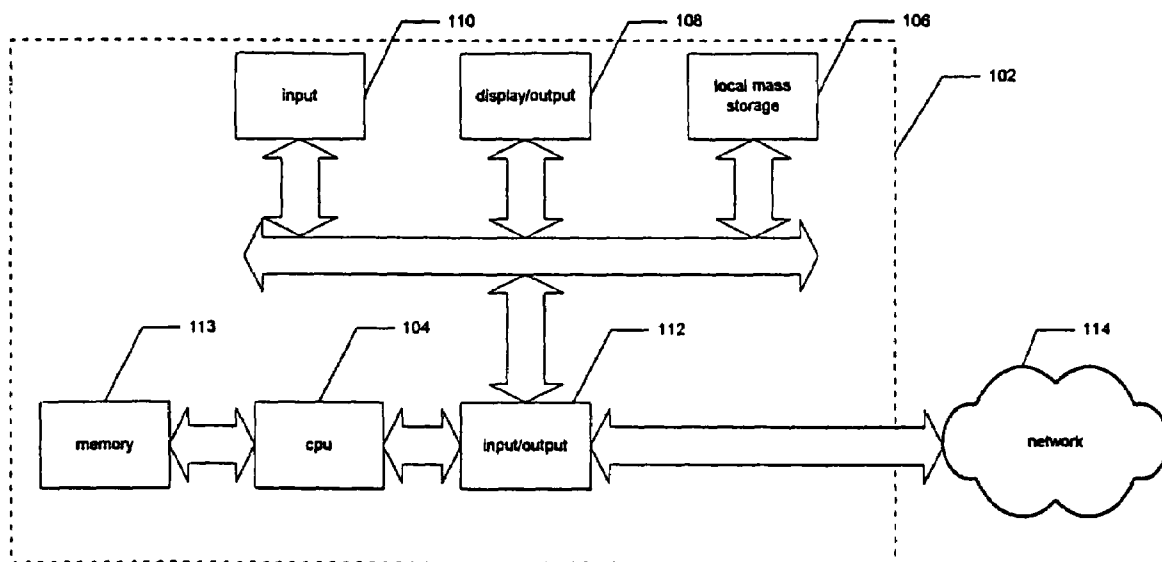
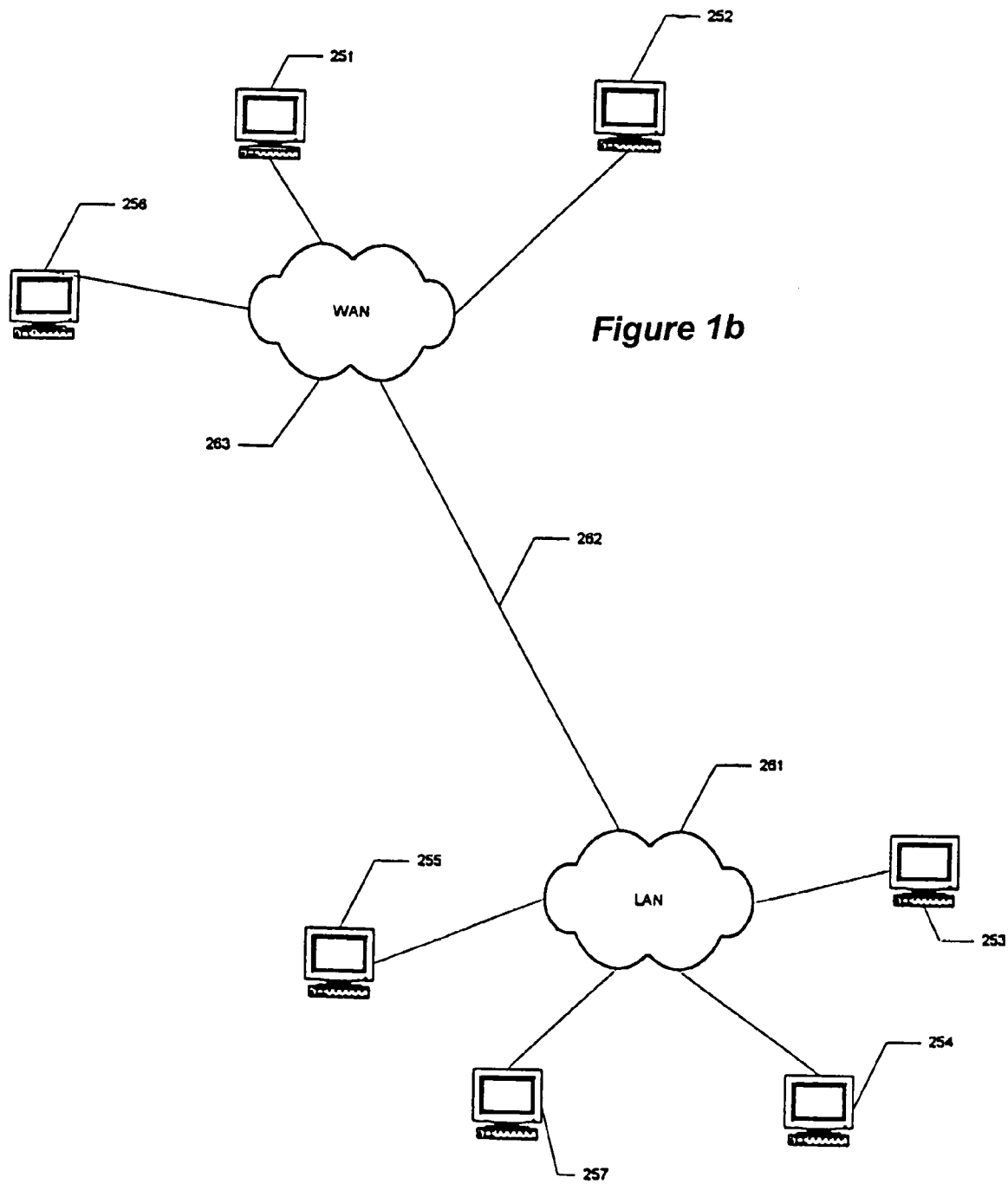


Figure 1a



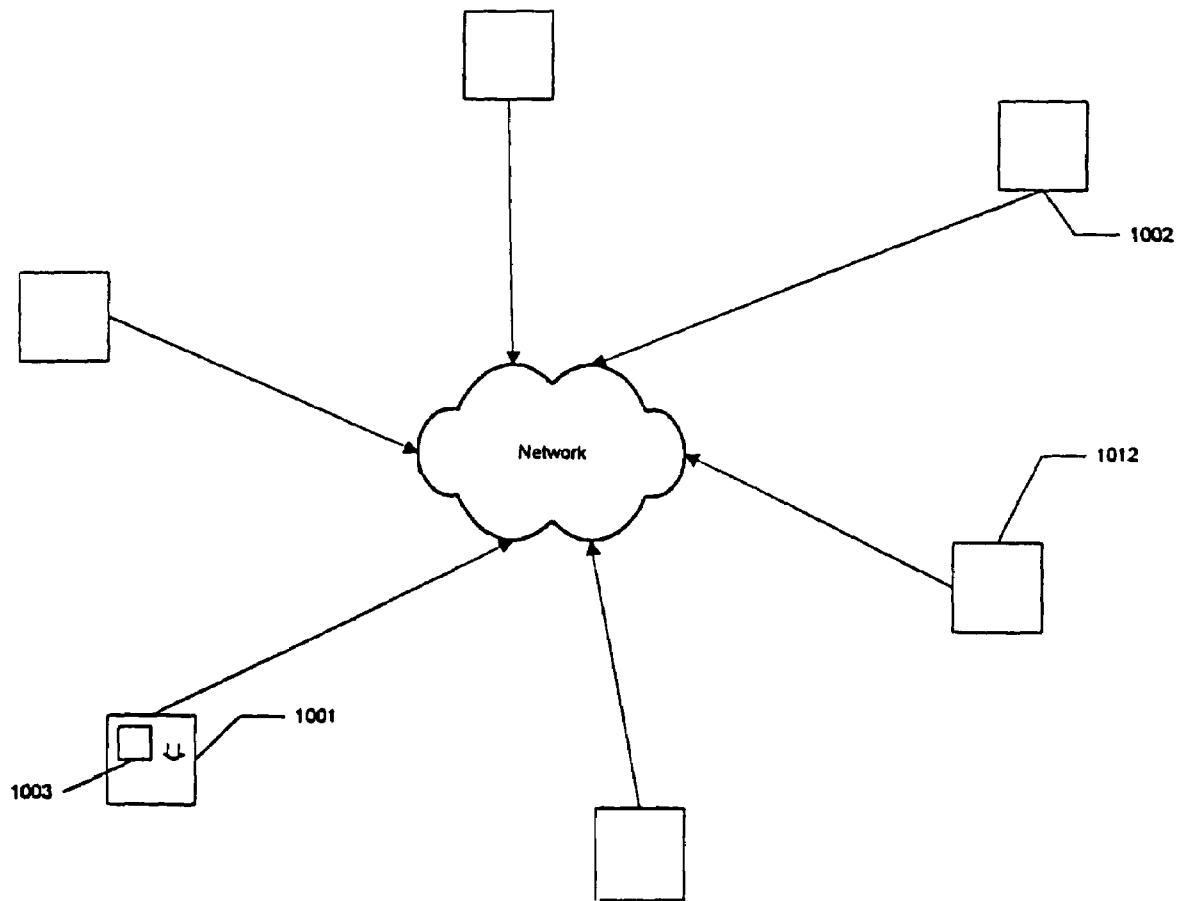


Figure 1c

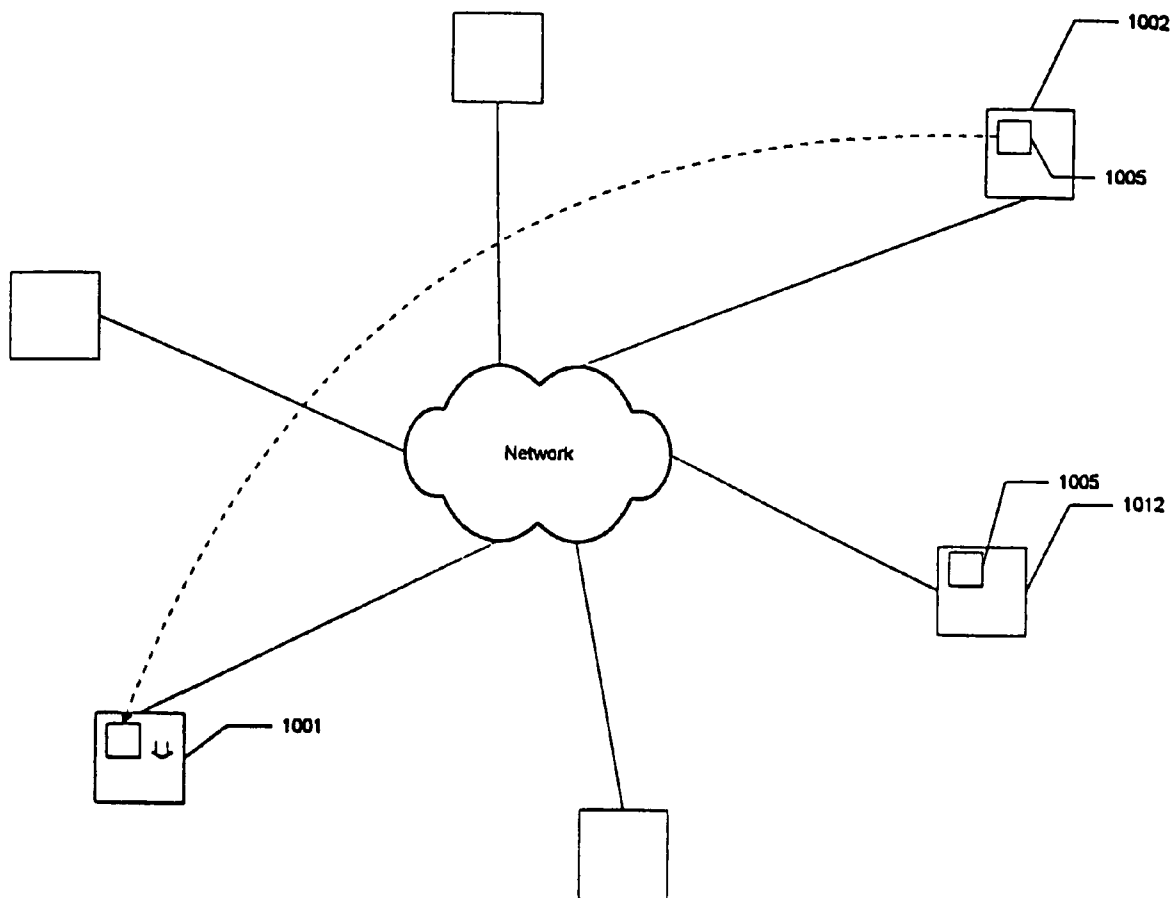
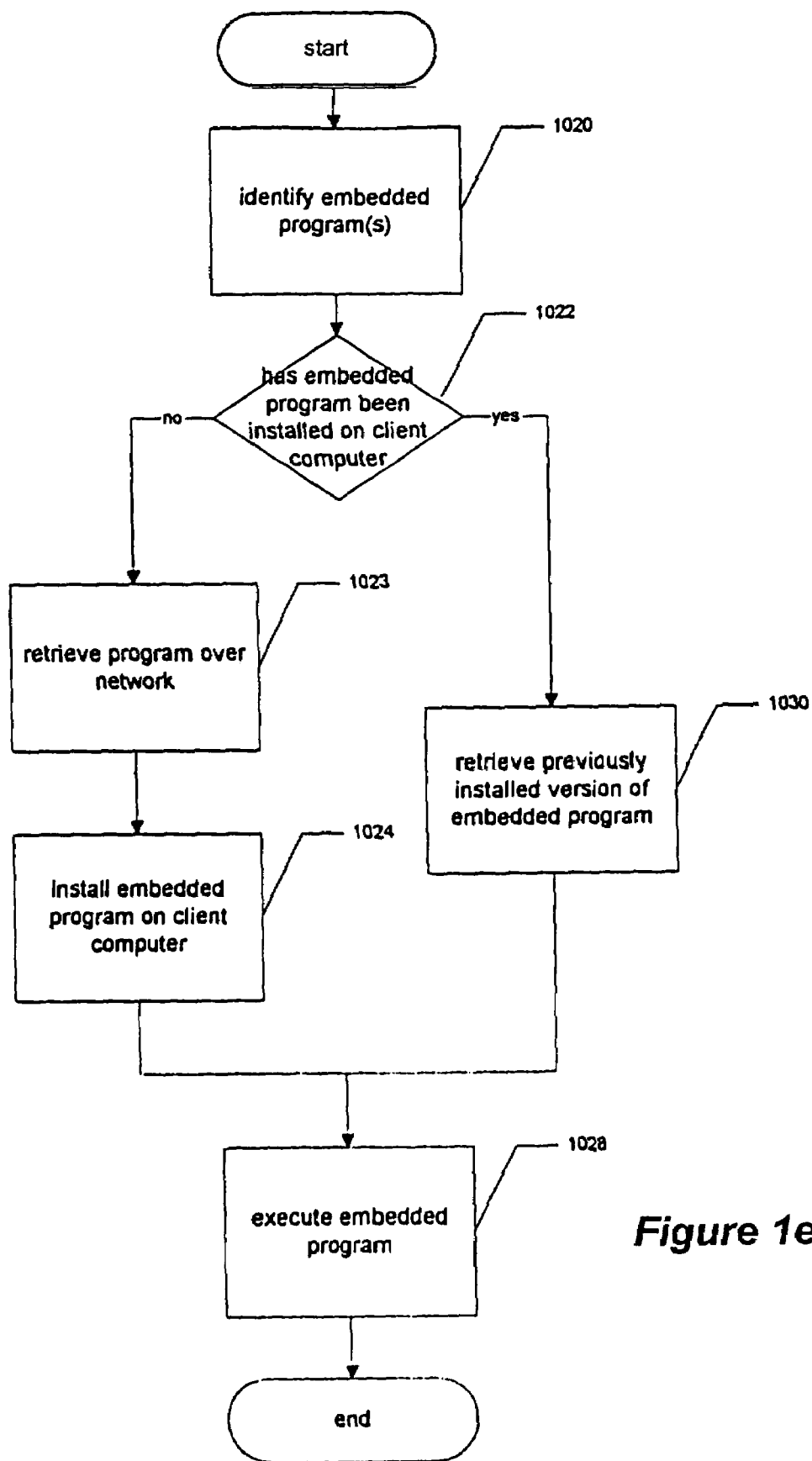
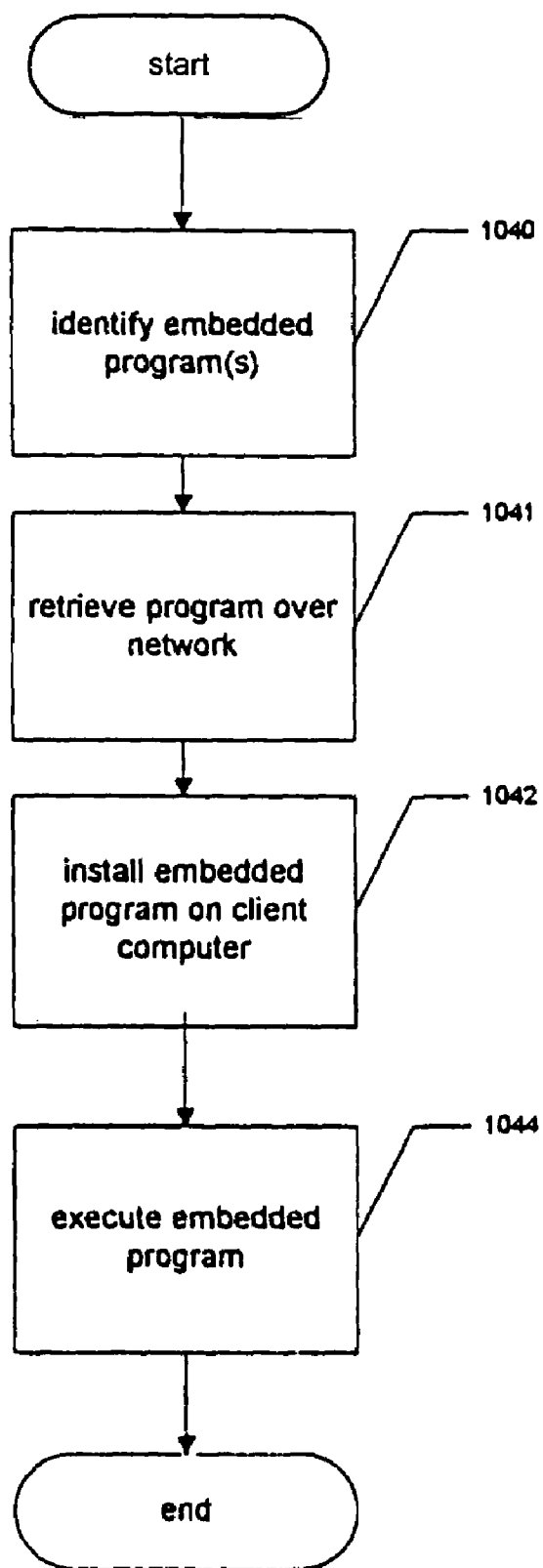


Figure 1d

**Figure 1e**

**Figure 1f**

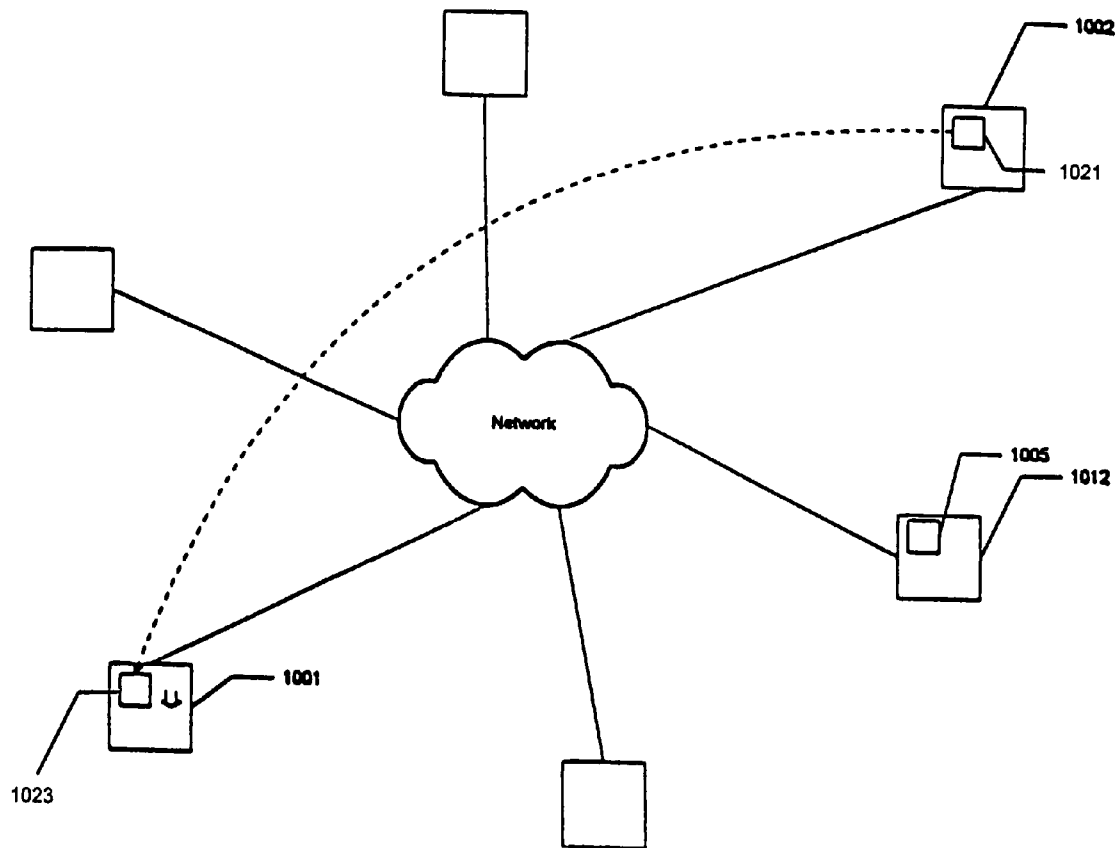


Figure 1g

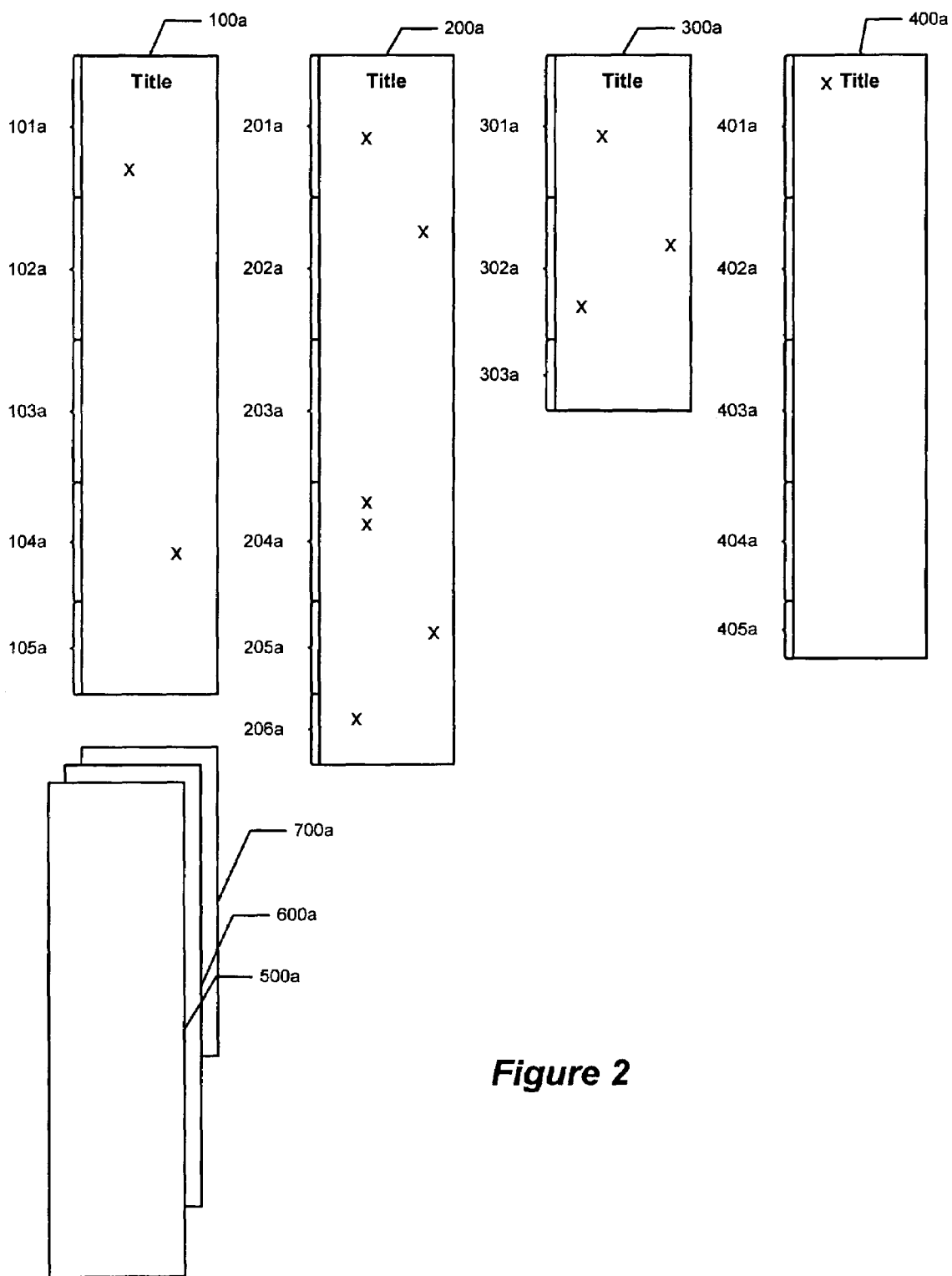


Figure 2

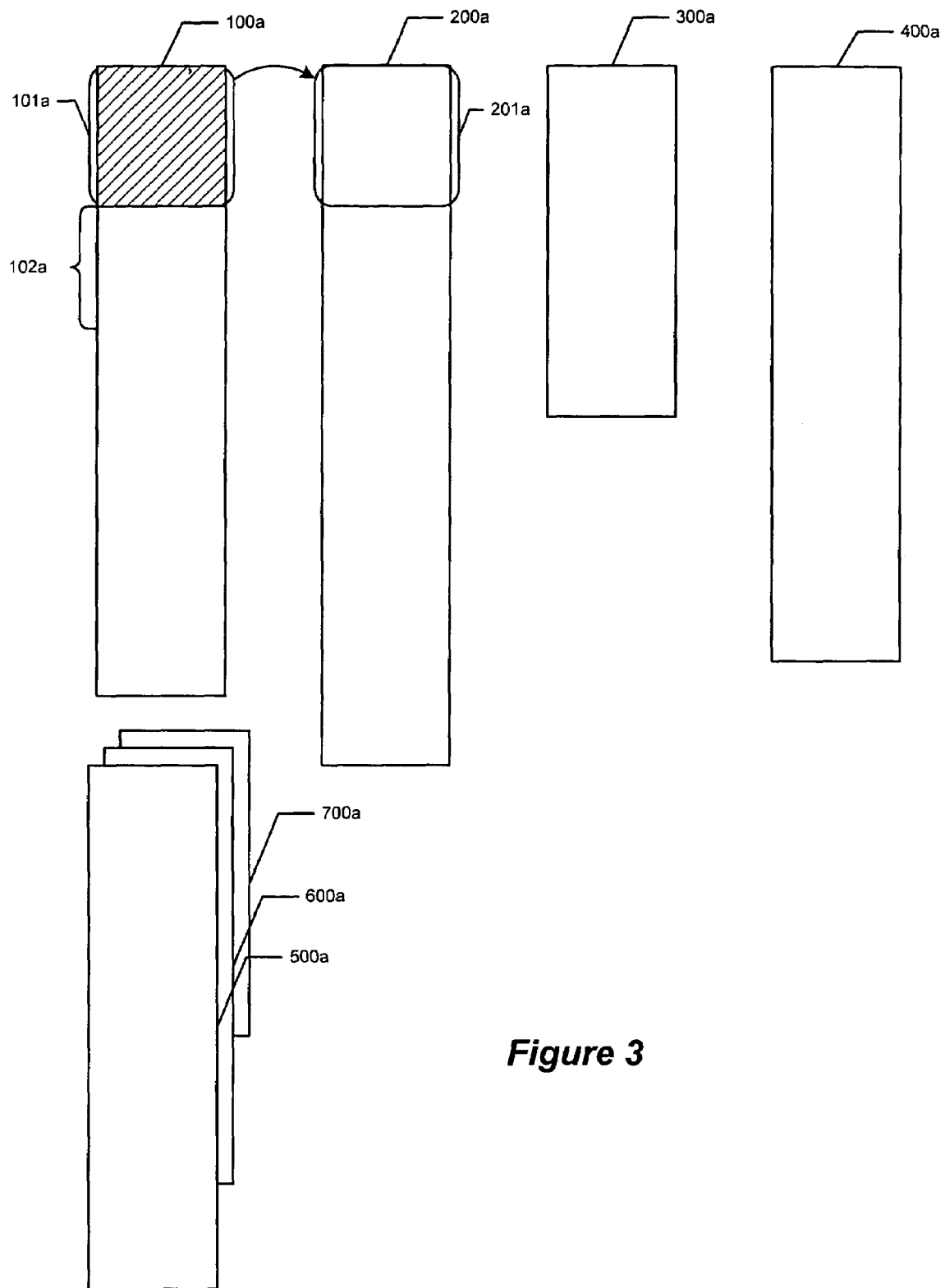


Figure 3

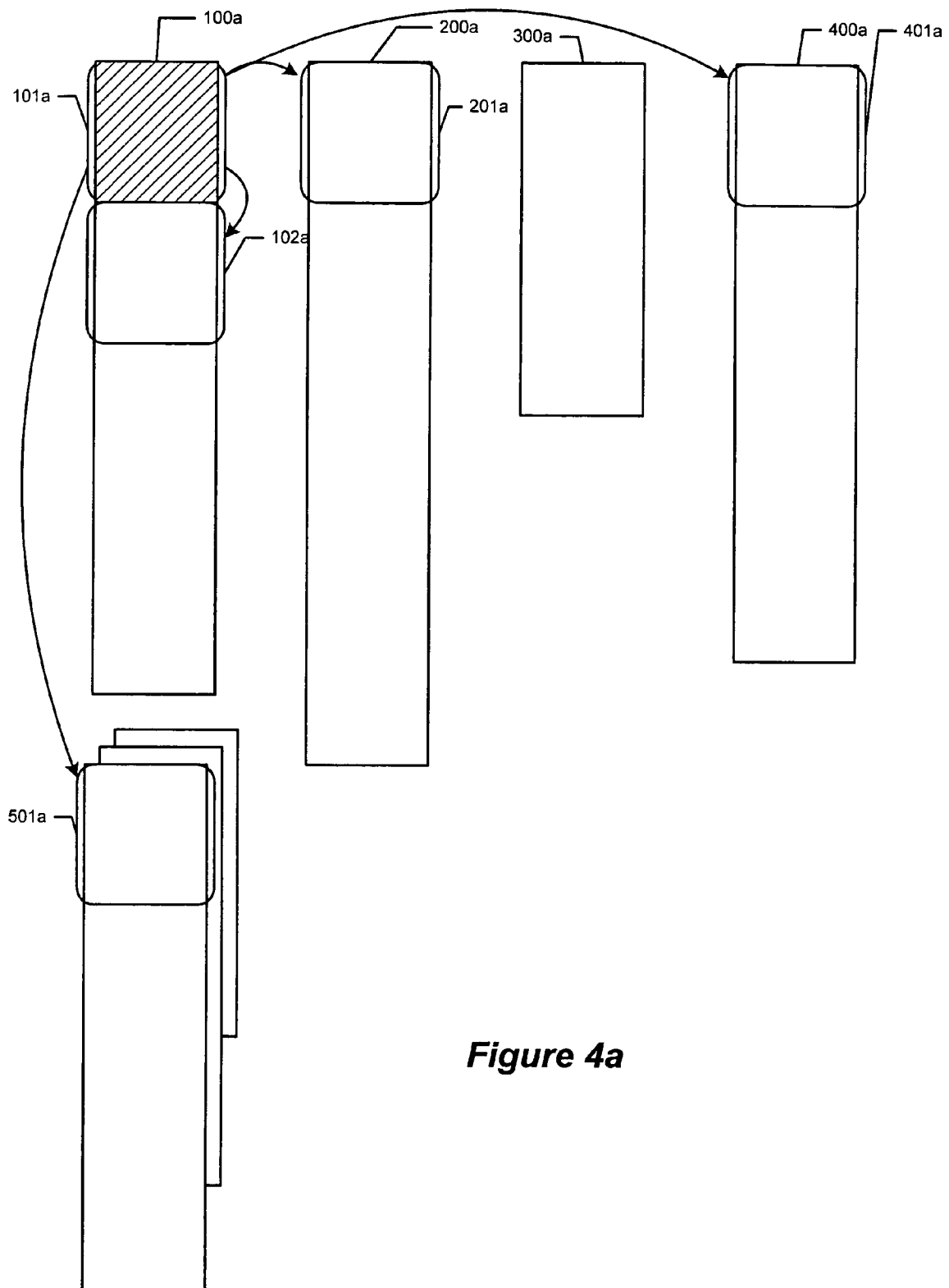


Figure 4a

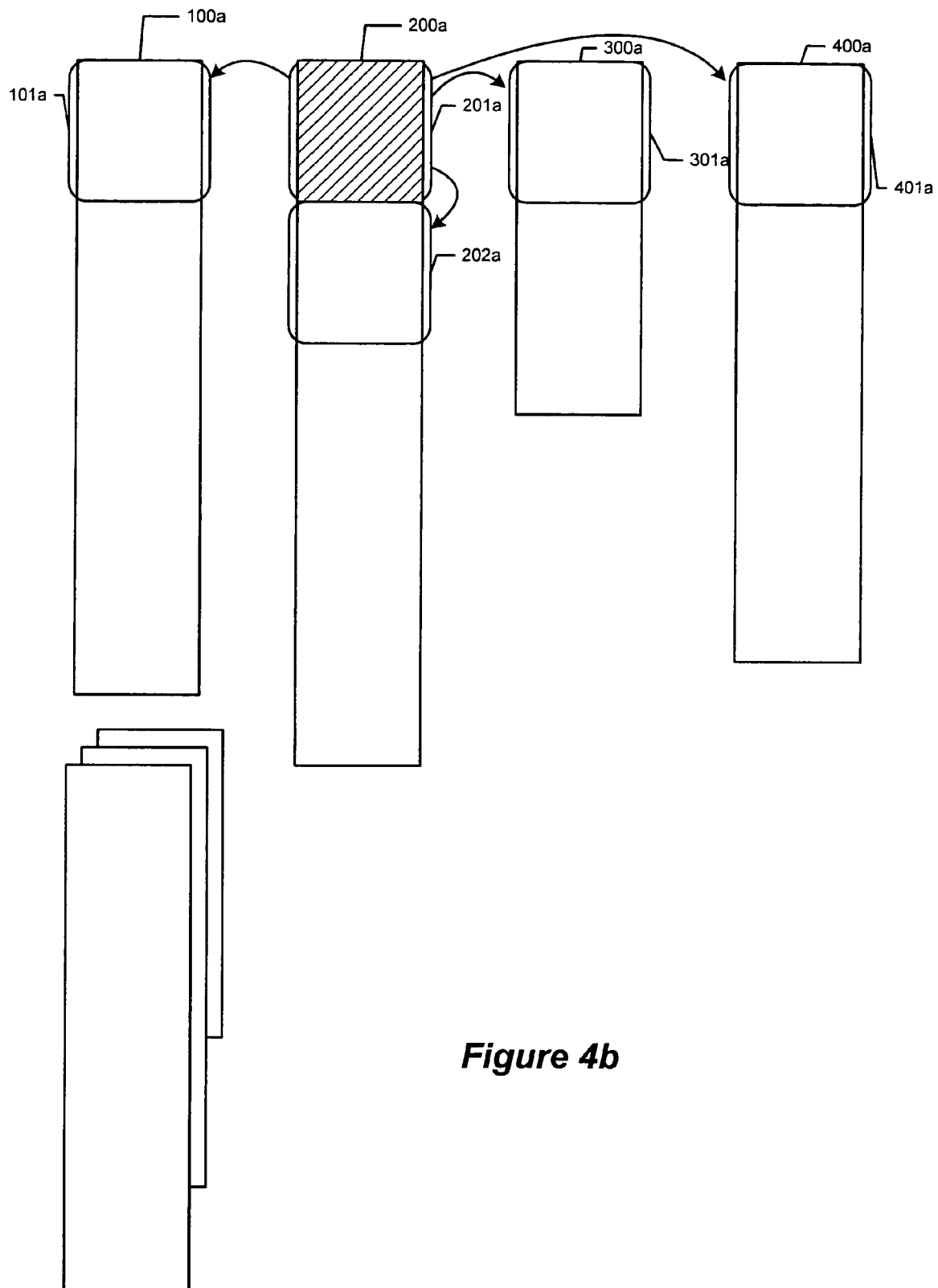


Figure 4b

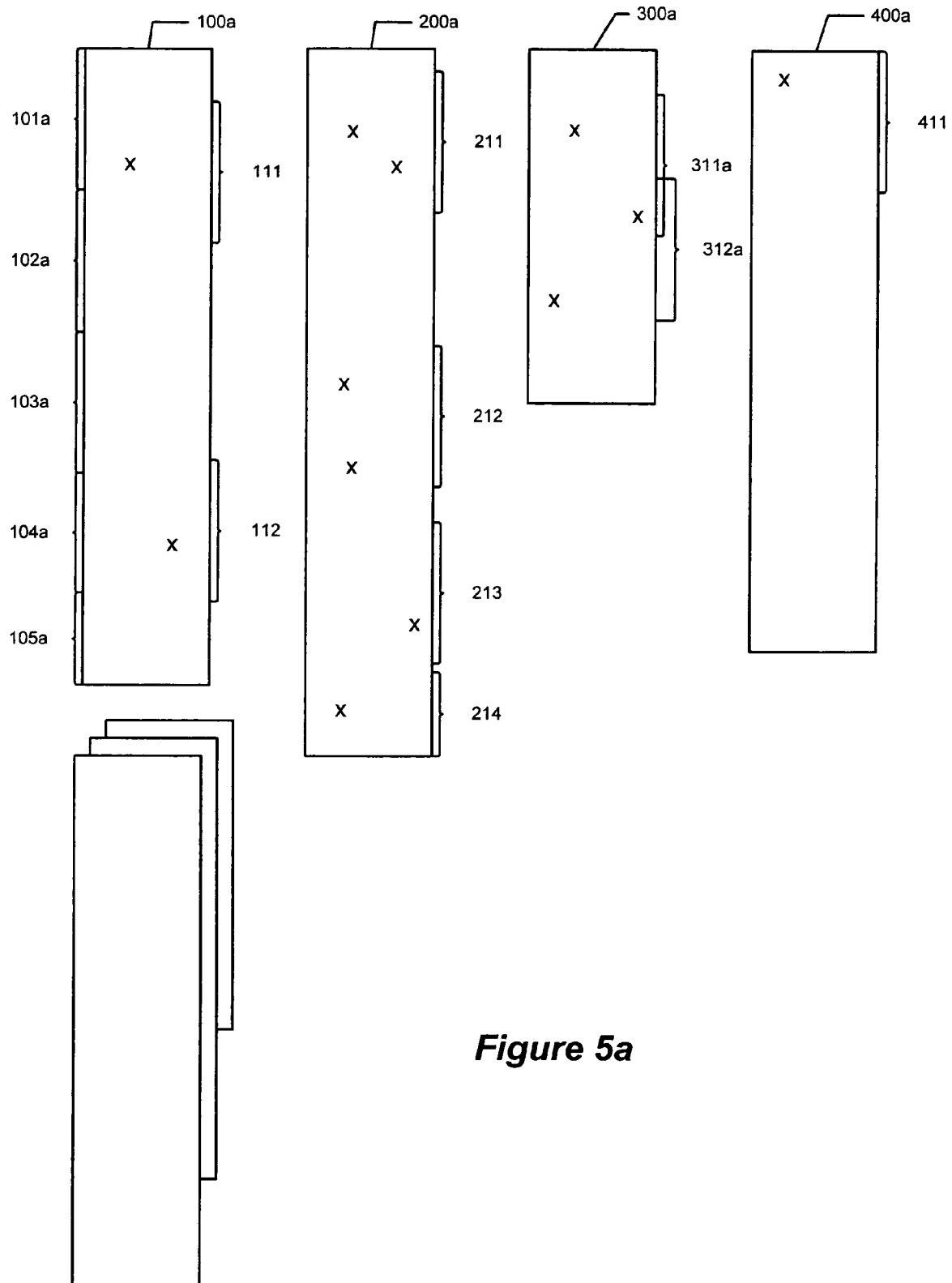


Figure 5a

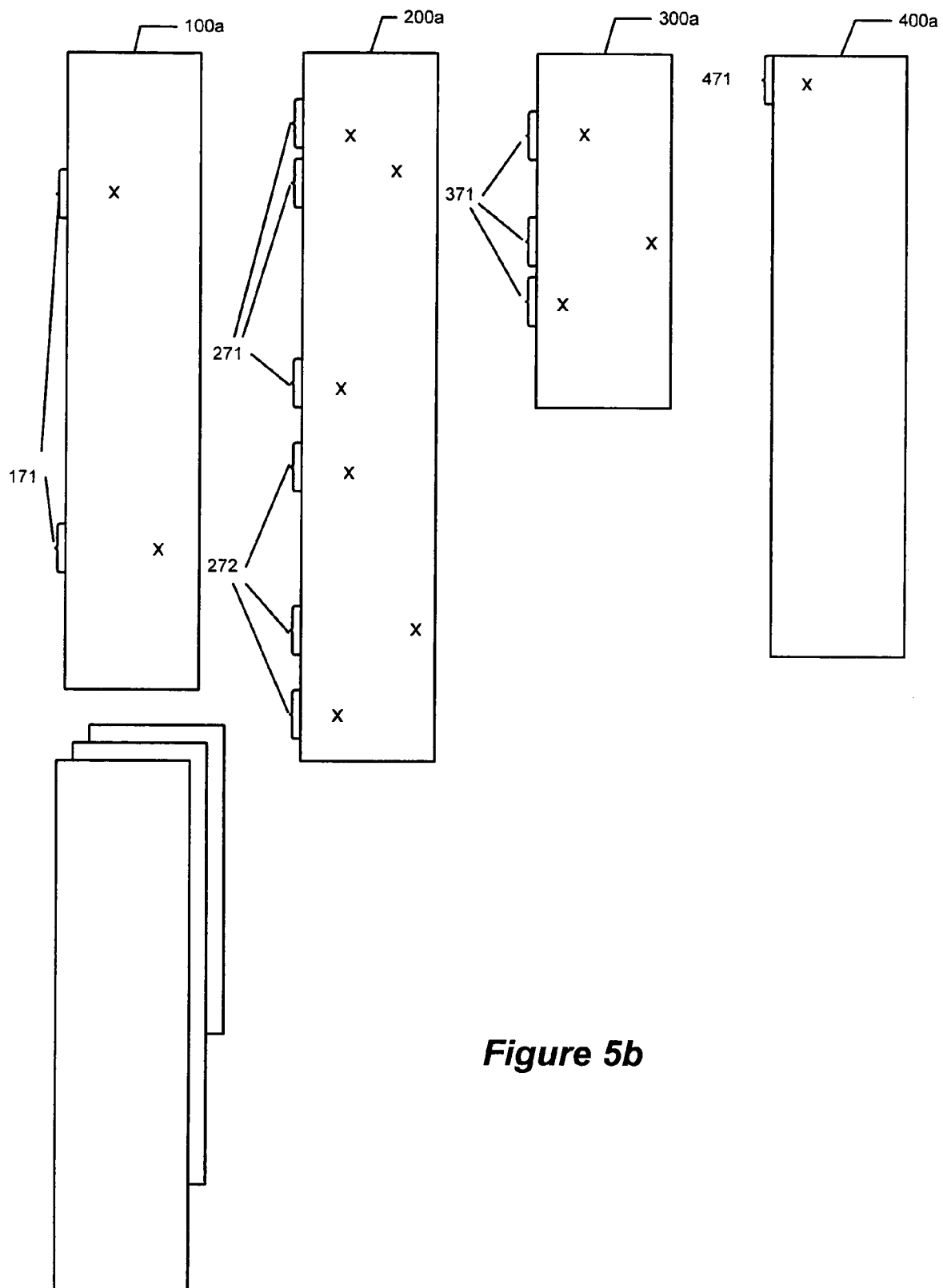


Figure 5b

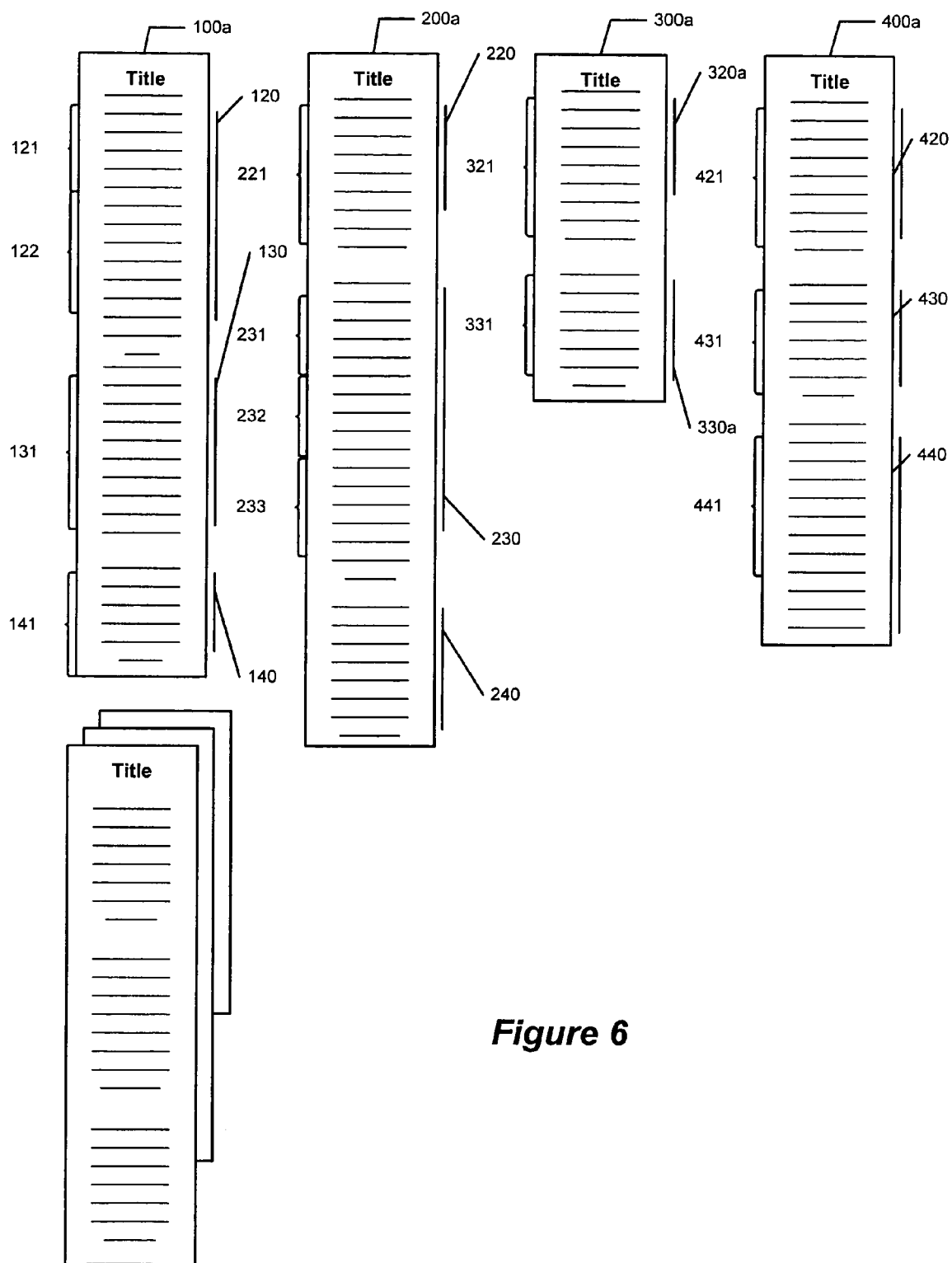
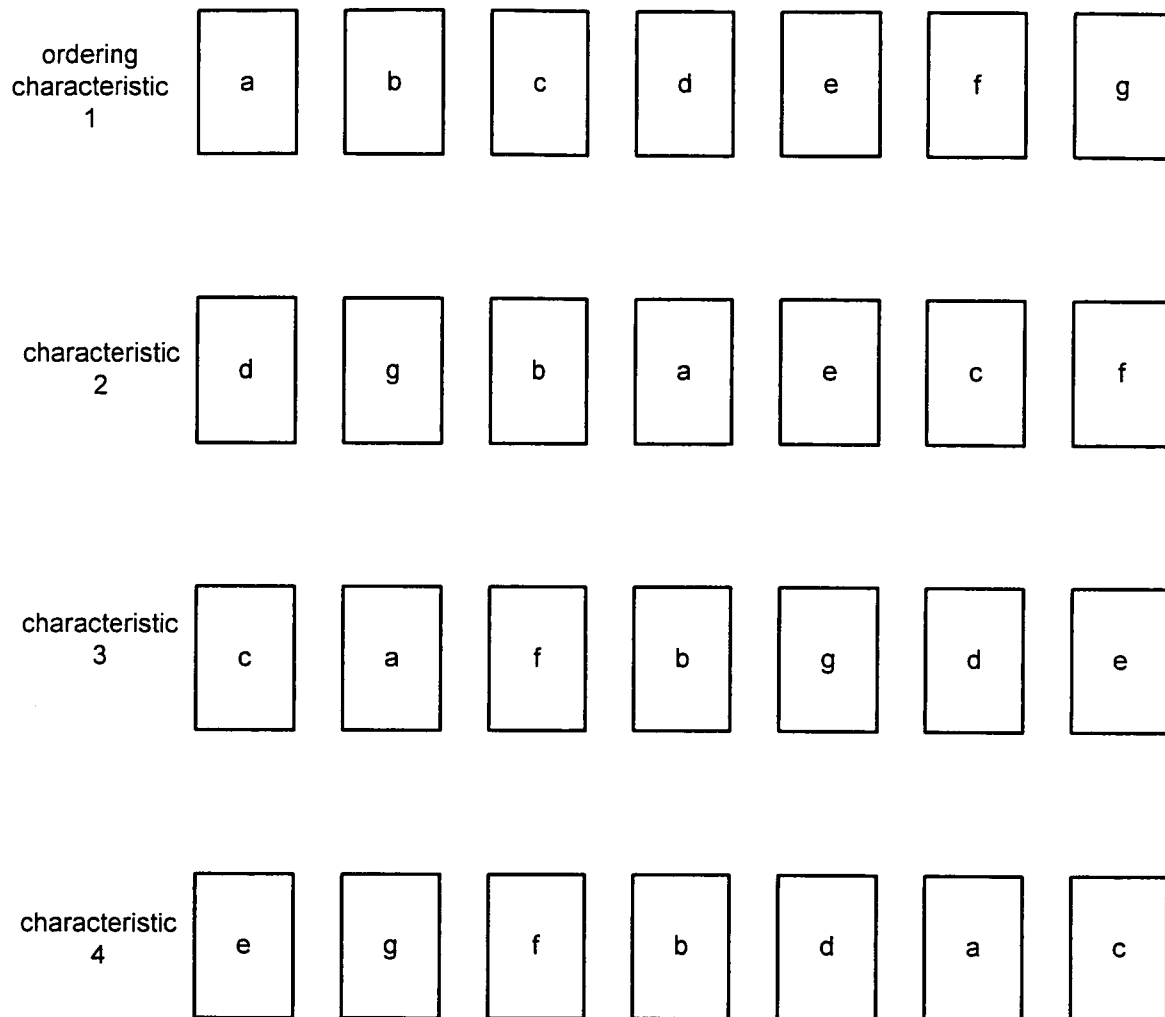
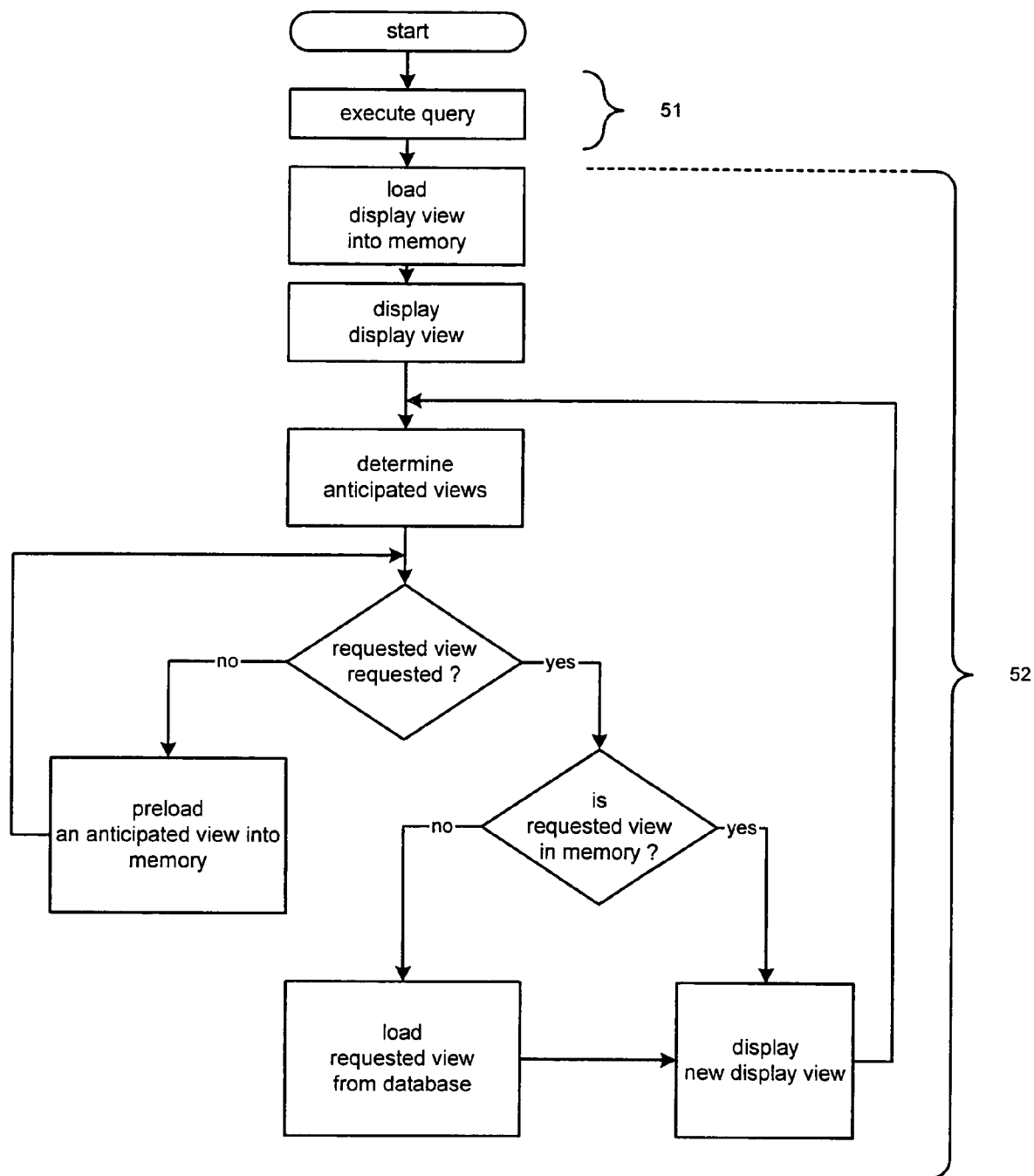


Figure 6

**Figure 7**

**Figure 8**

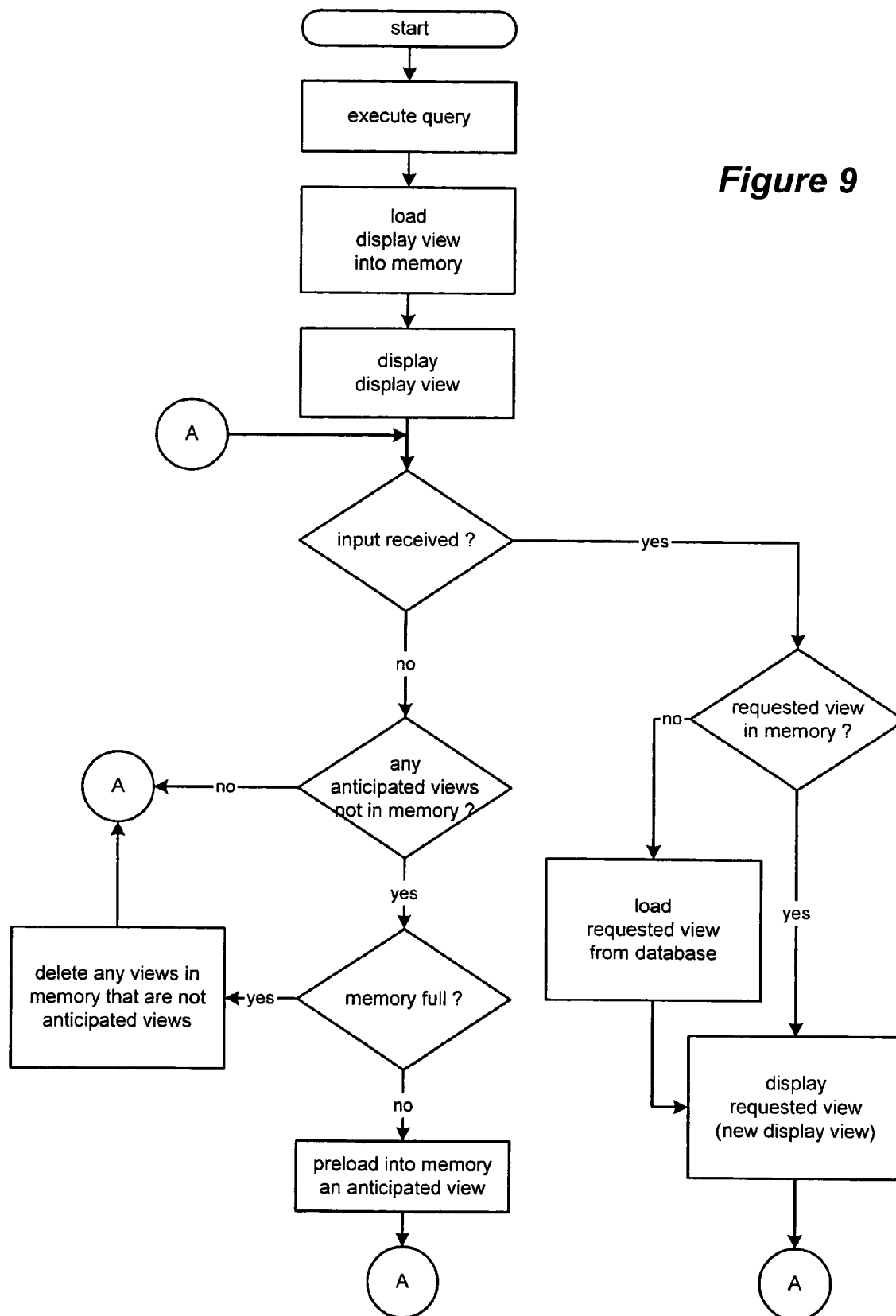
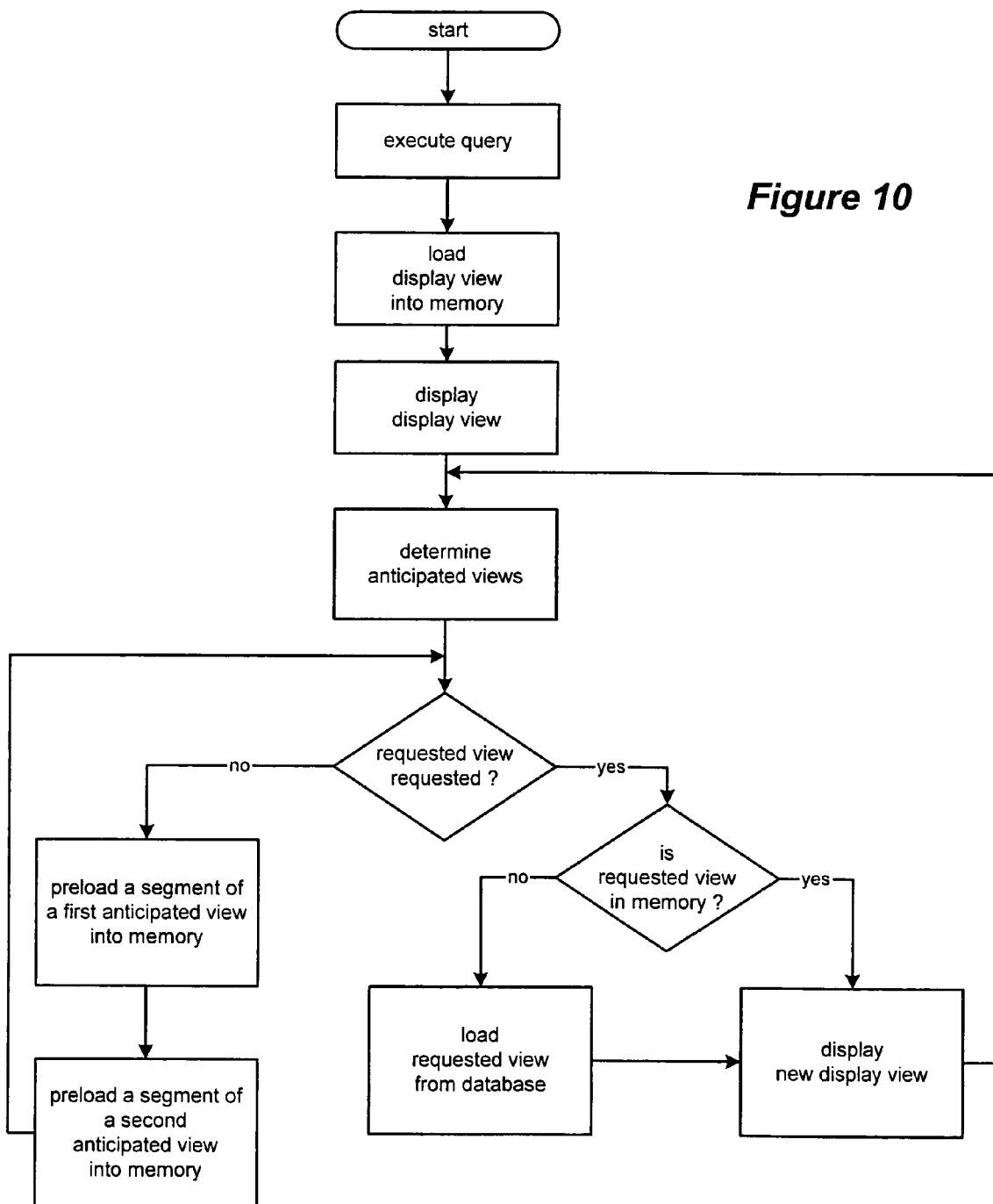


Figure 10

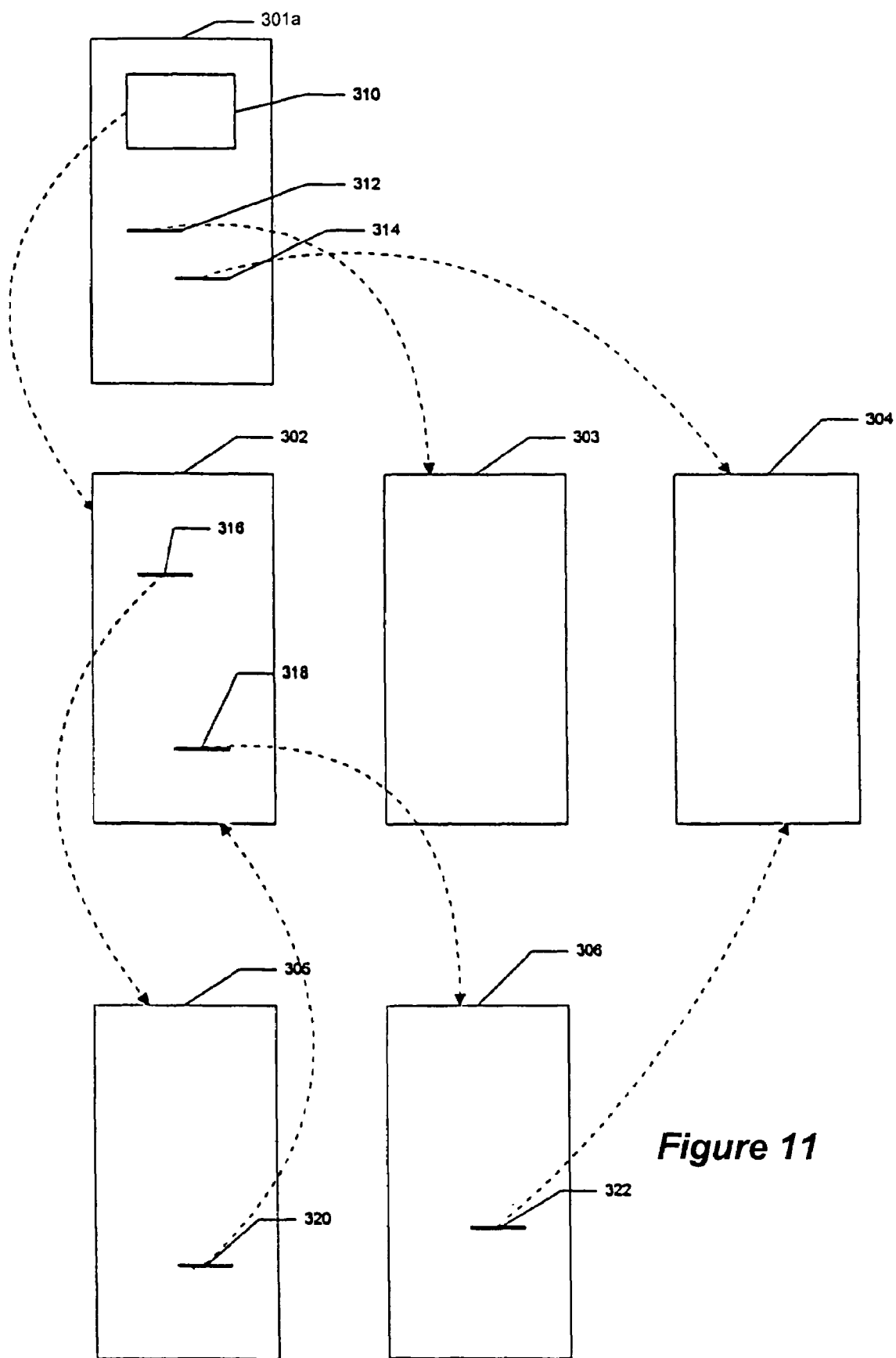
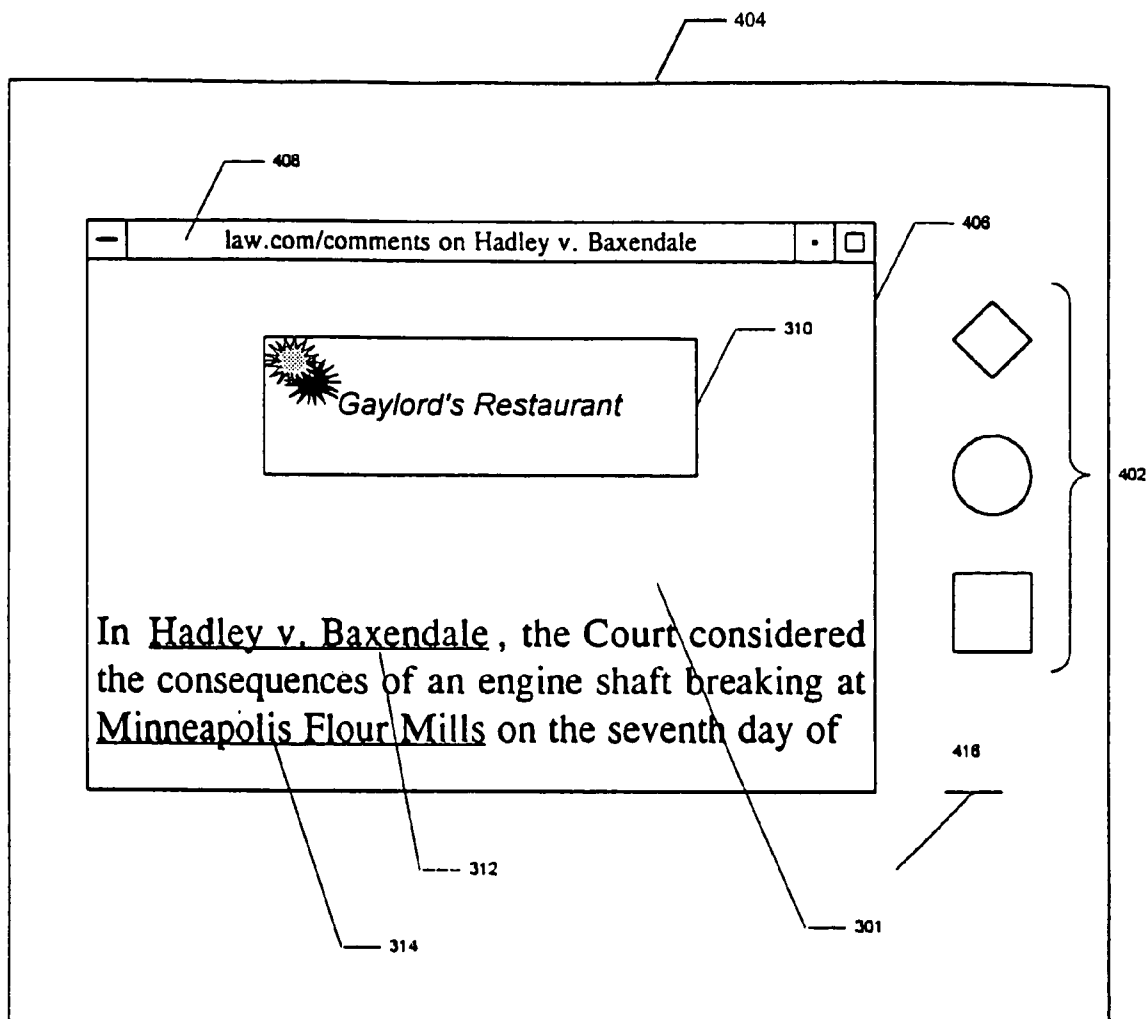
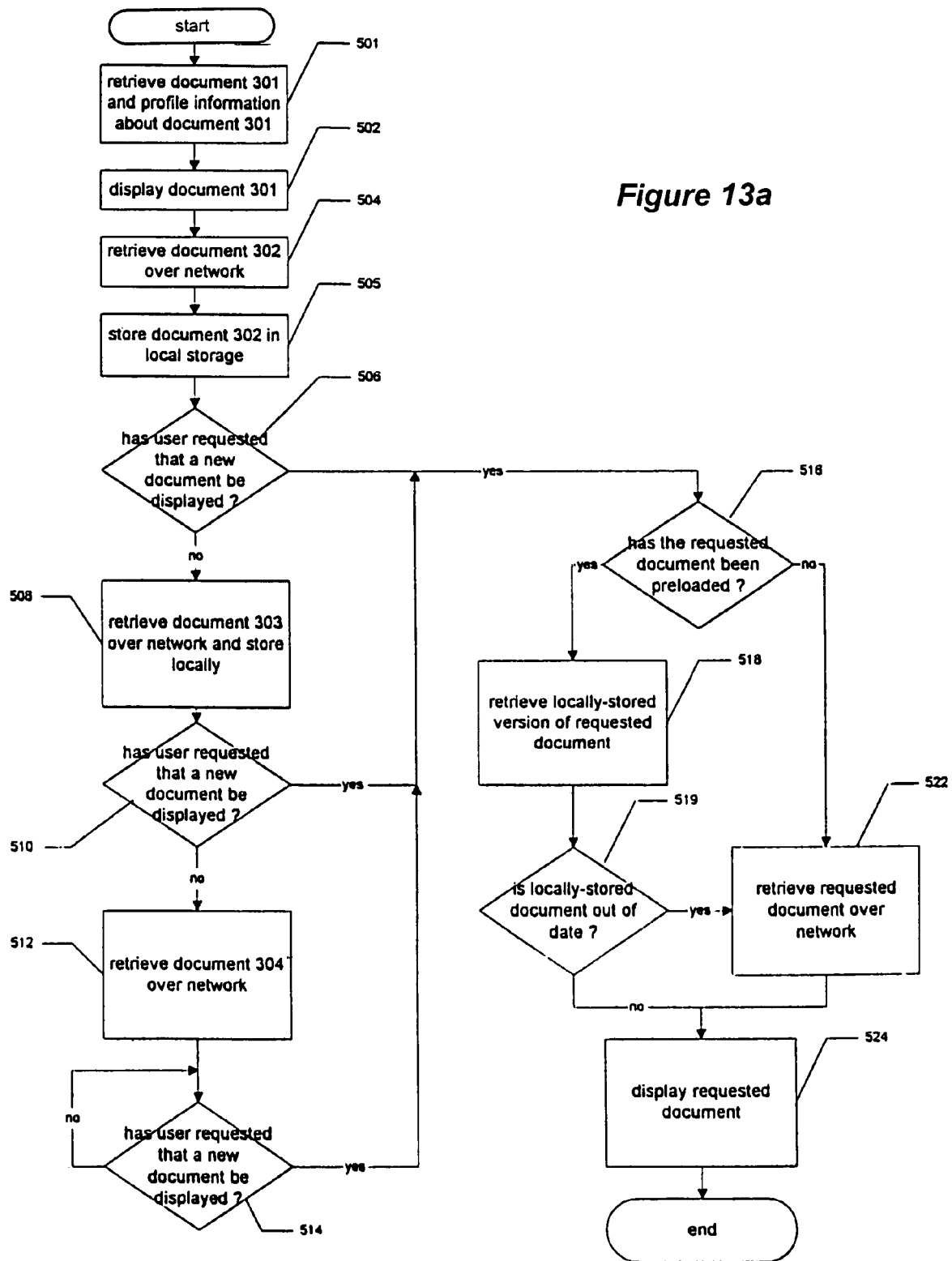
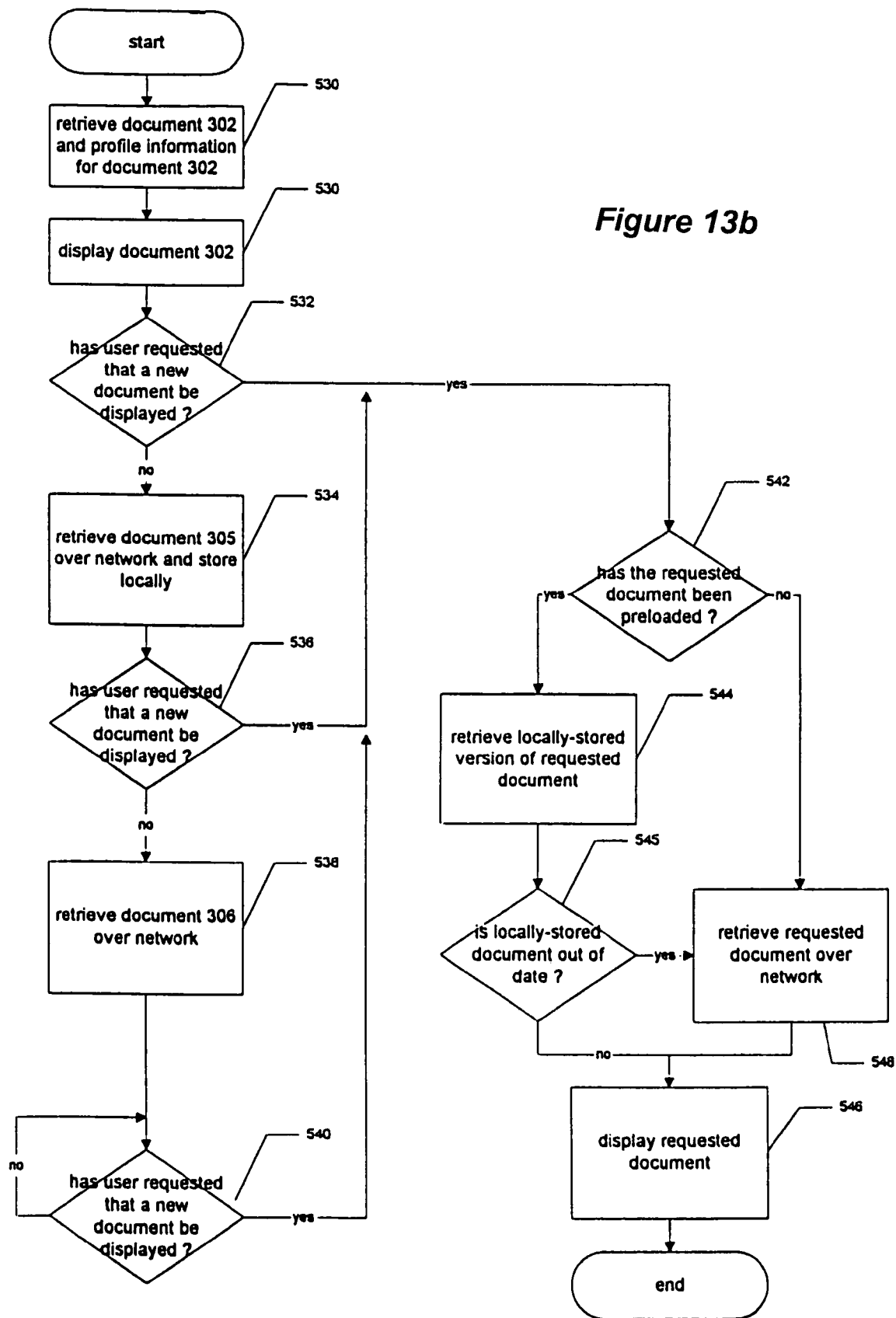
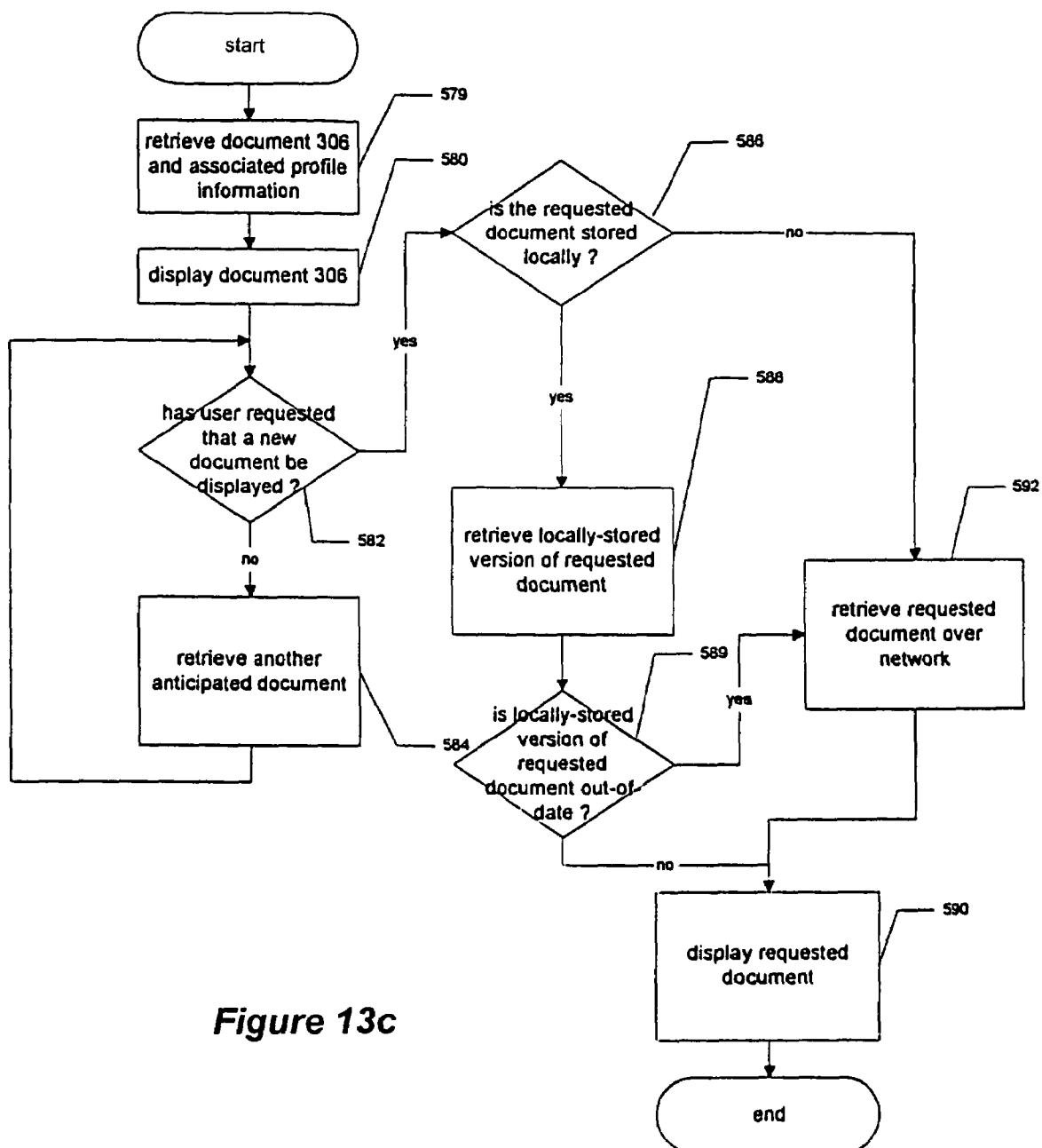


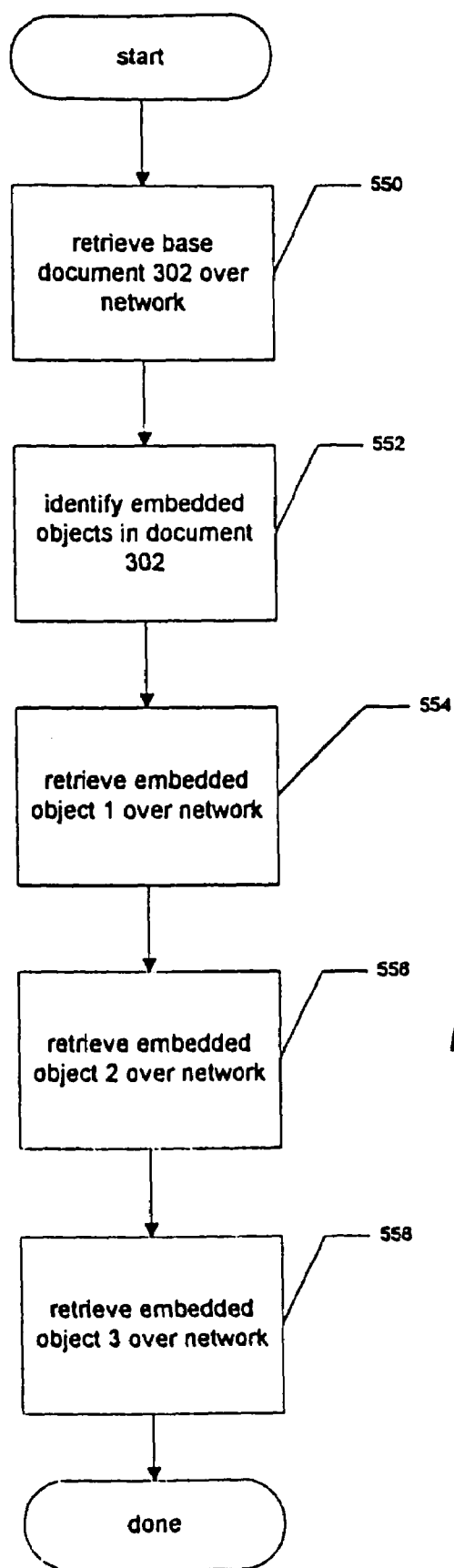
Figure 11

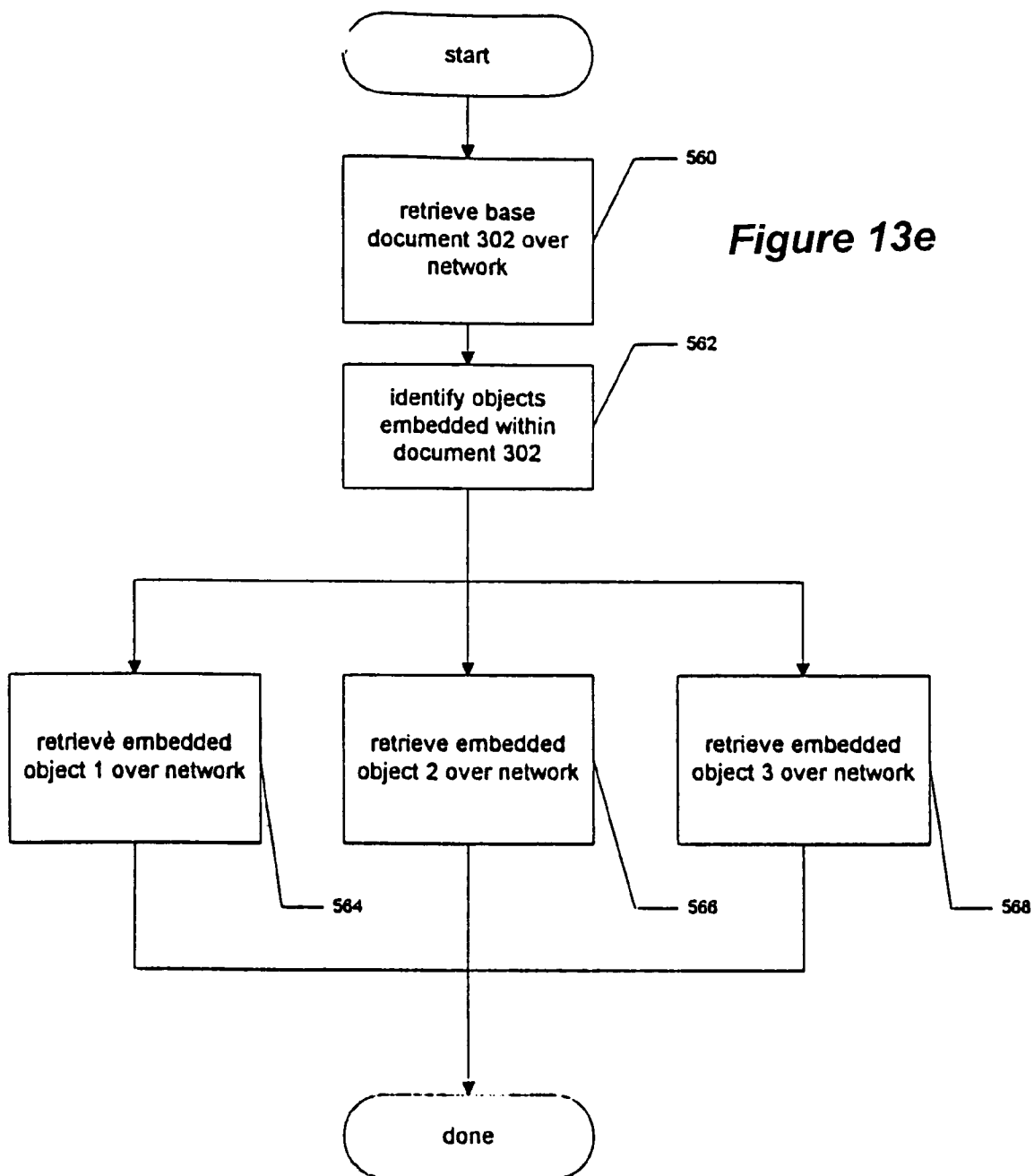
**Figure 12**





**Figure 13c**

**Figure 13d**



Profile information for Document 301

	doc. 302	doc. 303	doc. 304	doc. 319
server on which the document is stored	gaylords.com	same	mfc.com	delta.com
% of users who select	23%	21%	4%	2%

Figure 14

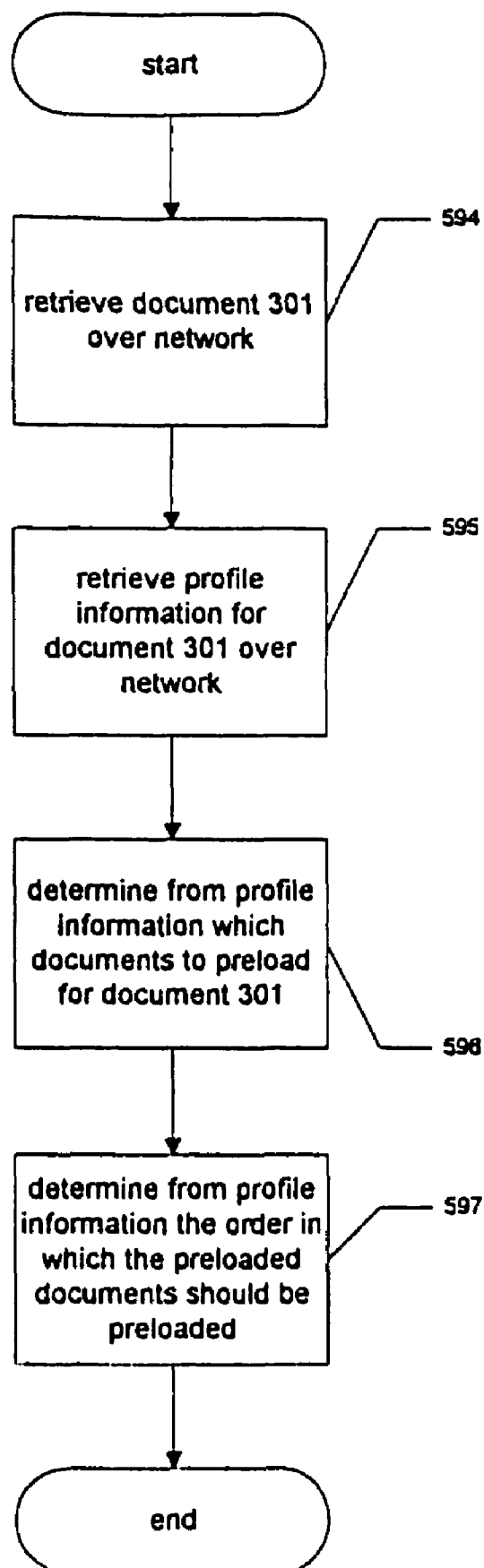
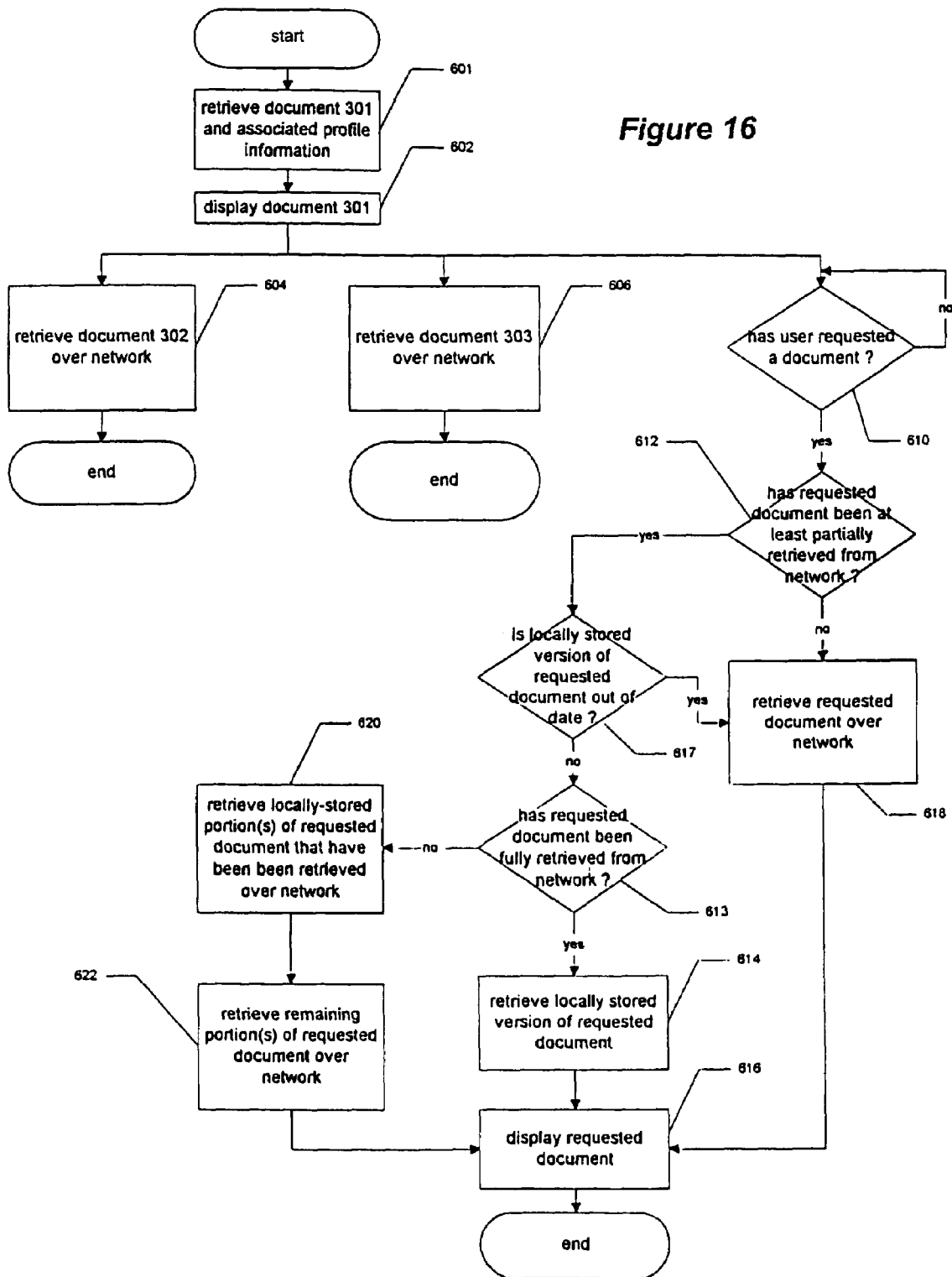
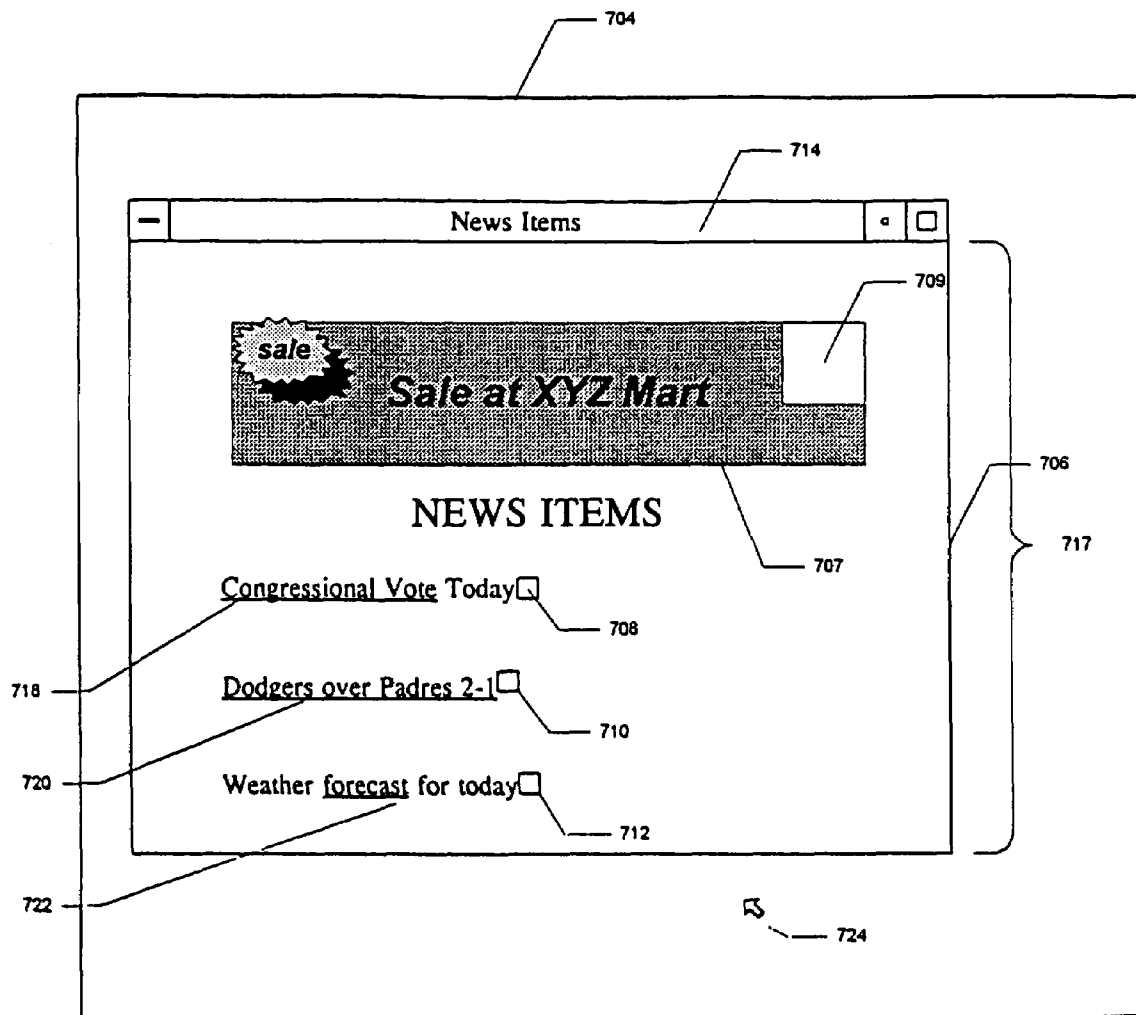
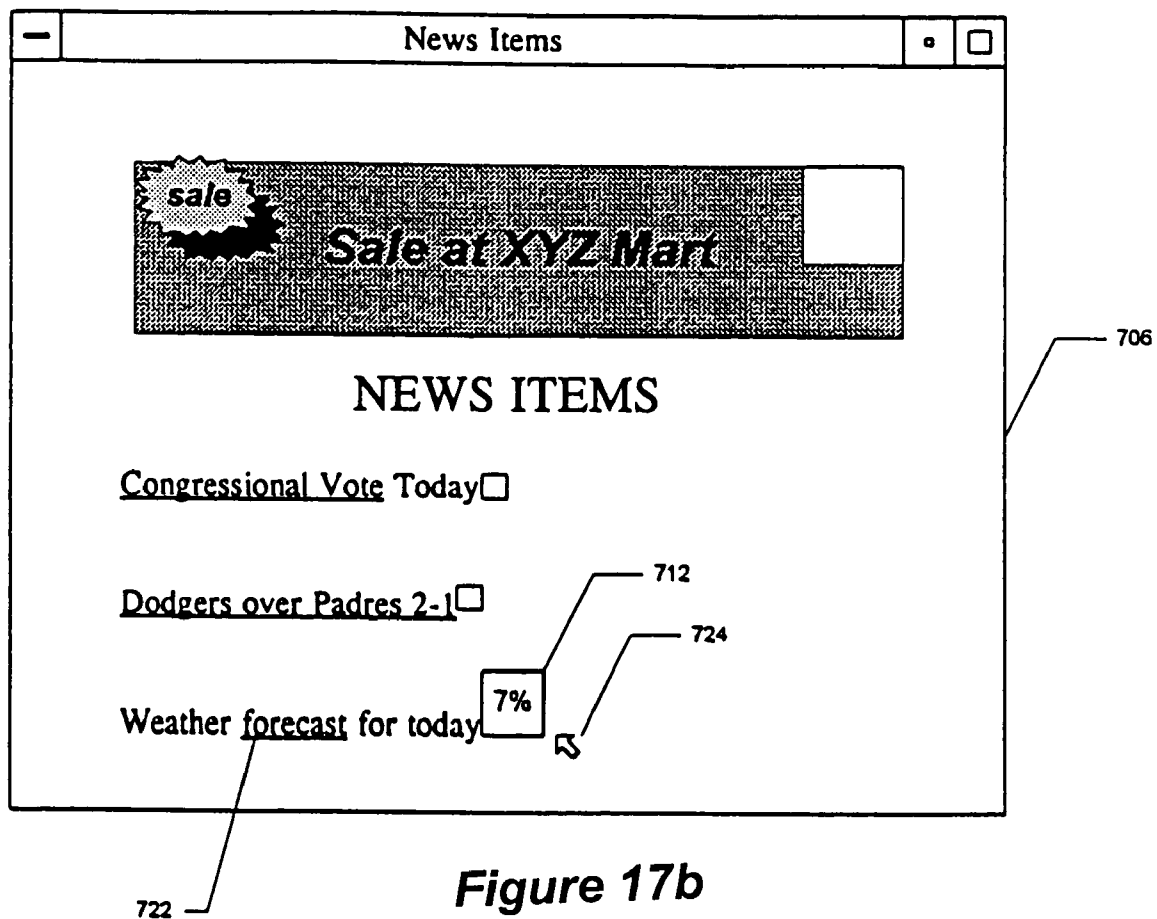
**Figure 15**

Figure 16



**Figure 17a**



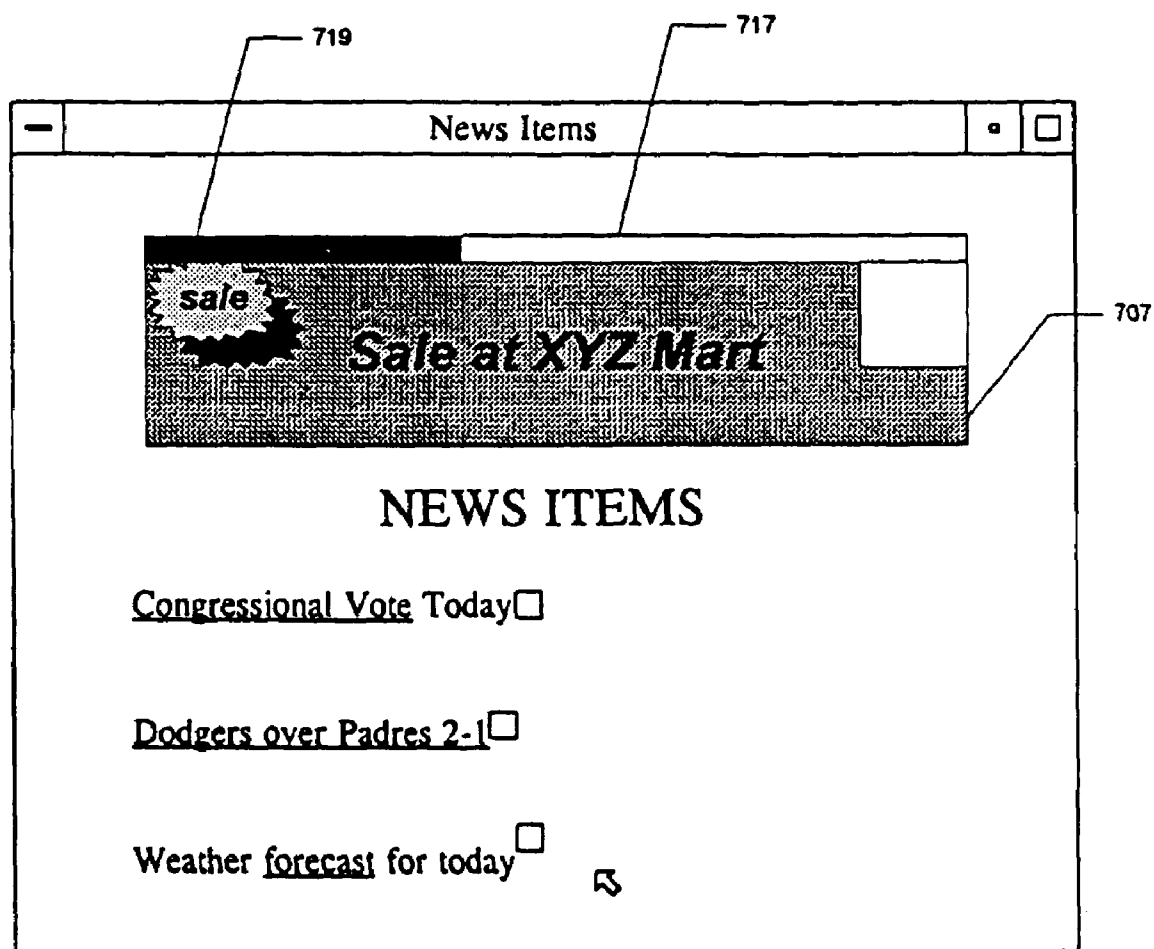


Figure 17c

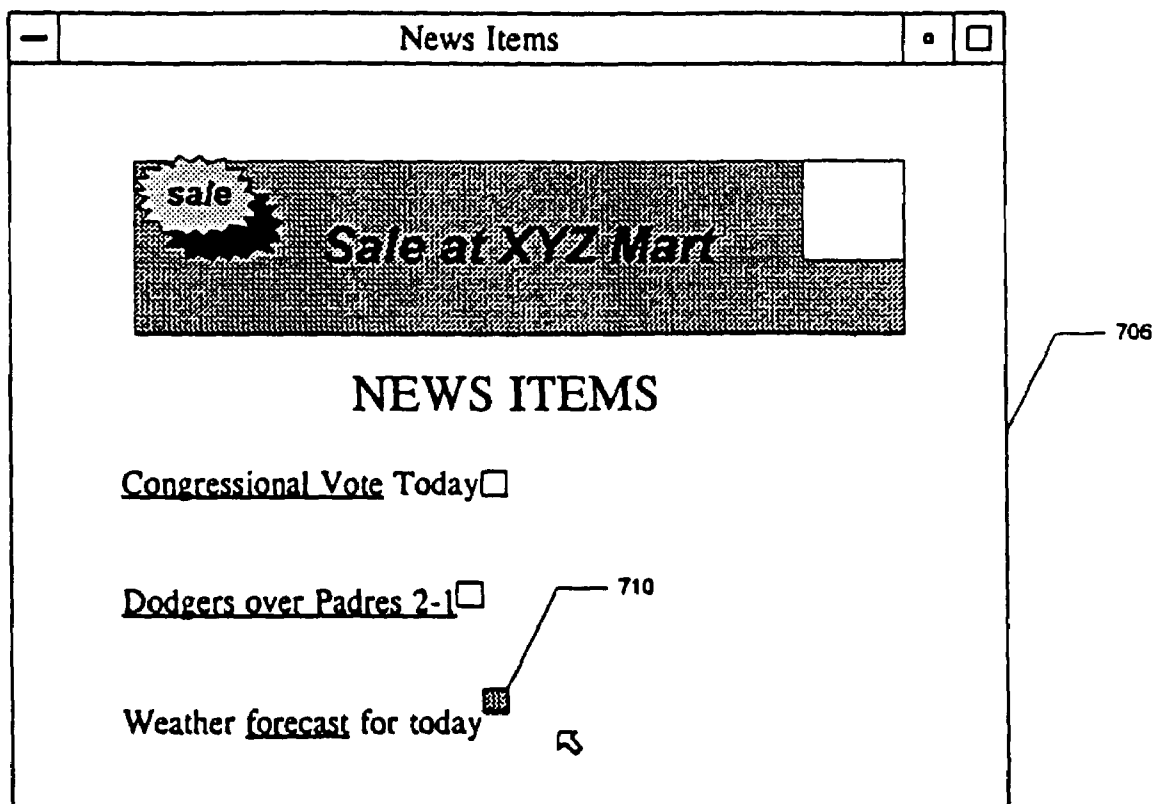
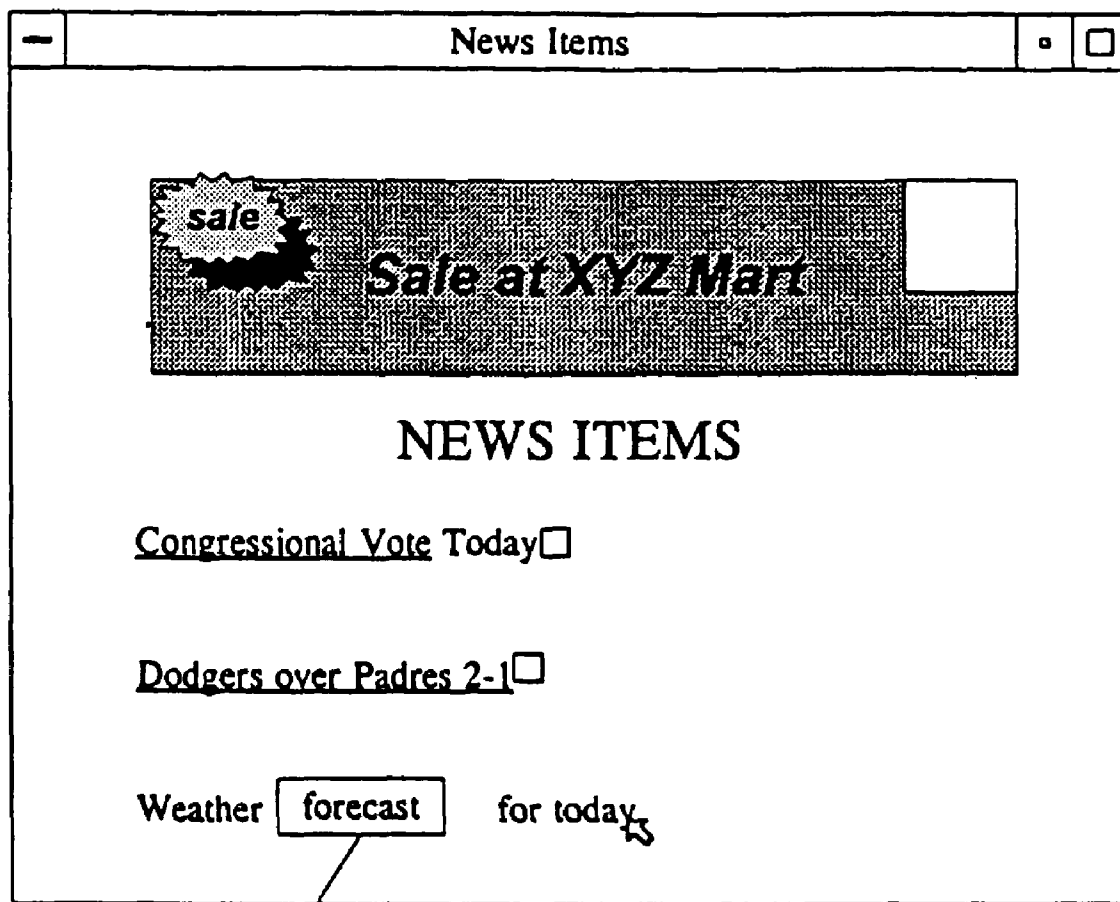


Figure 17d

**Figure 17e**

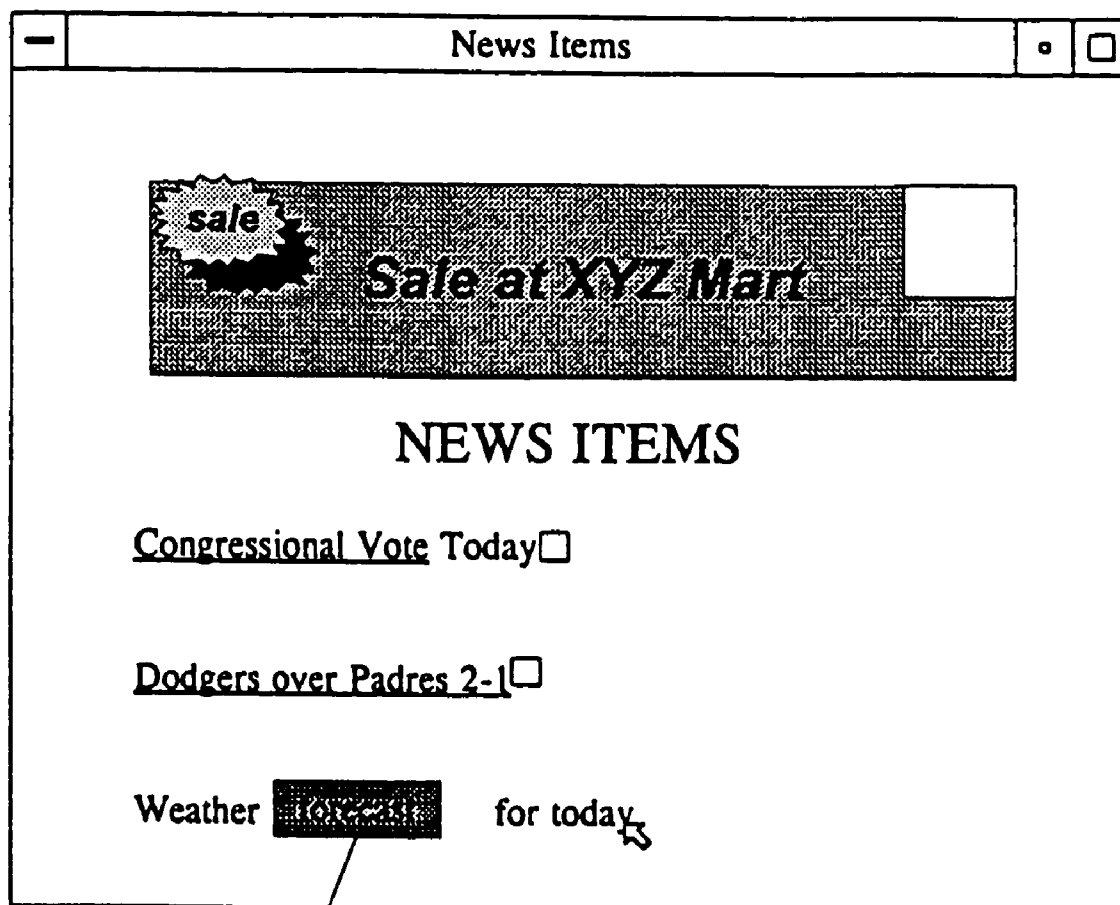
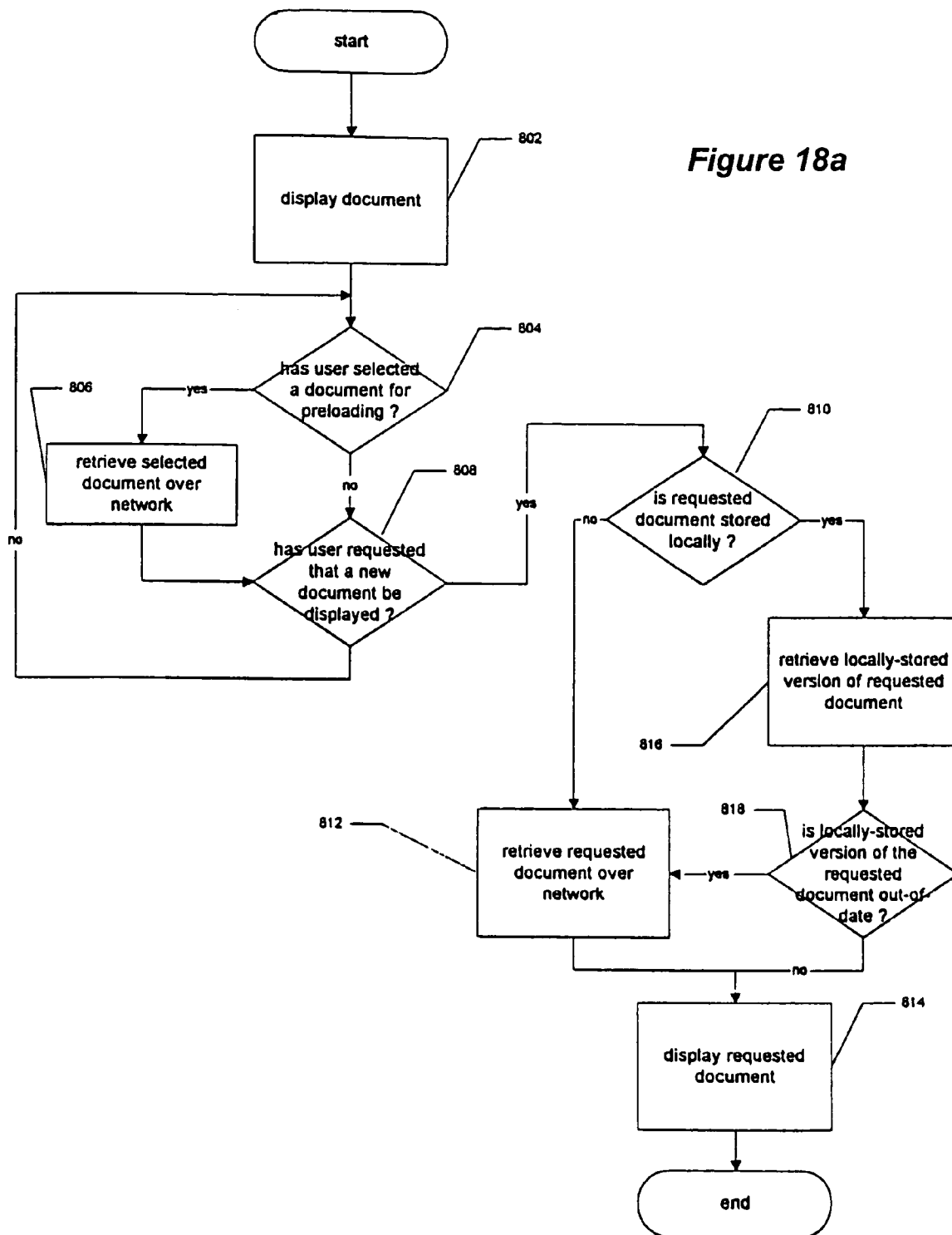
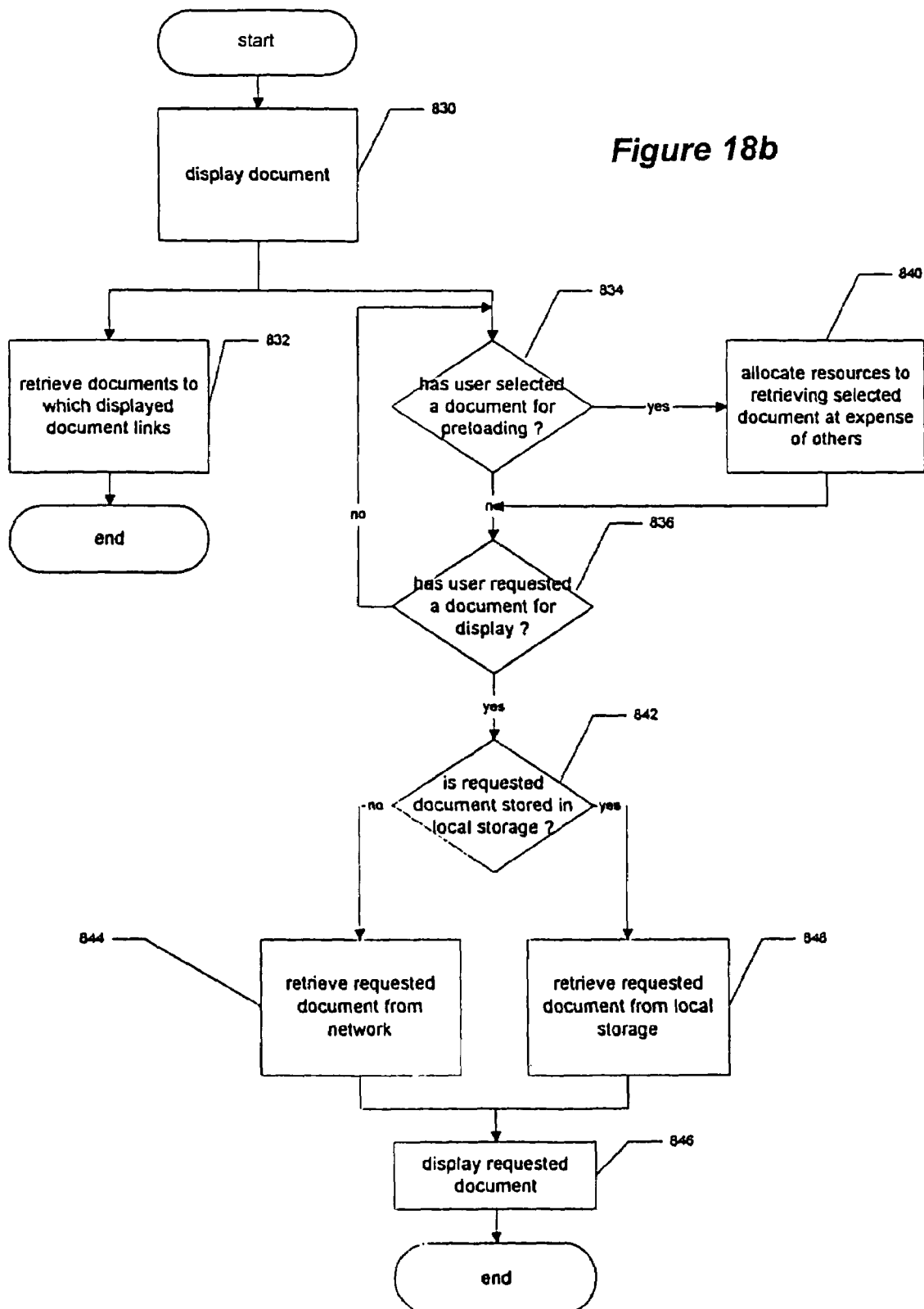
*Figure 17f*

Figure 18a





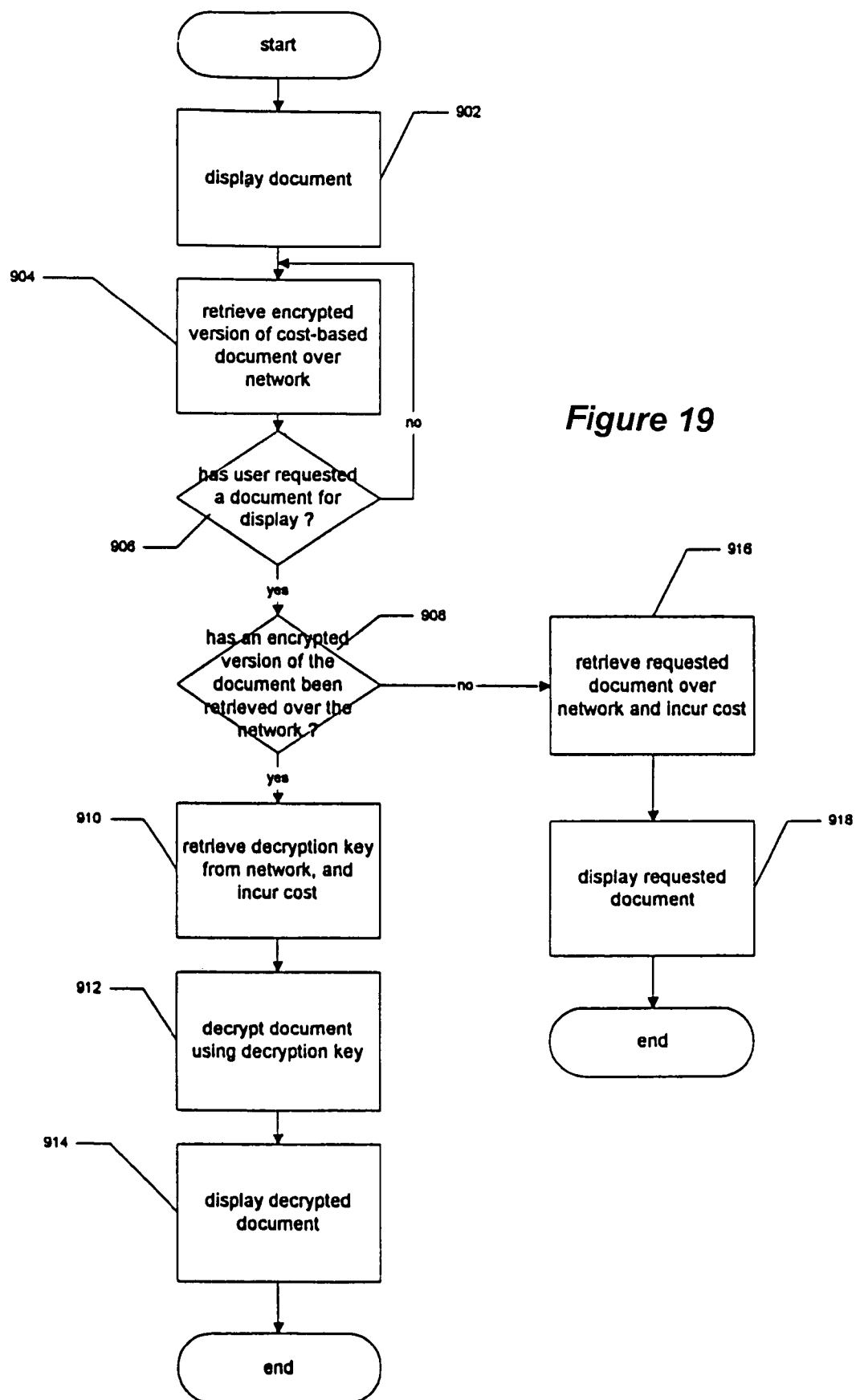
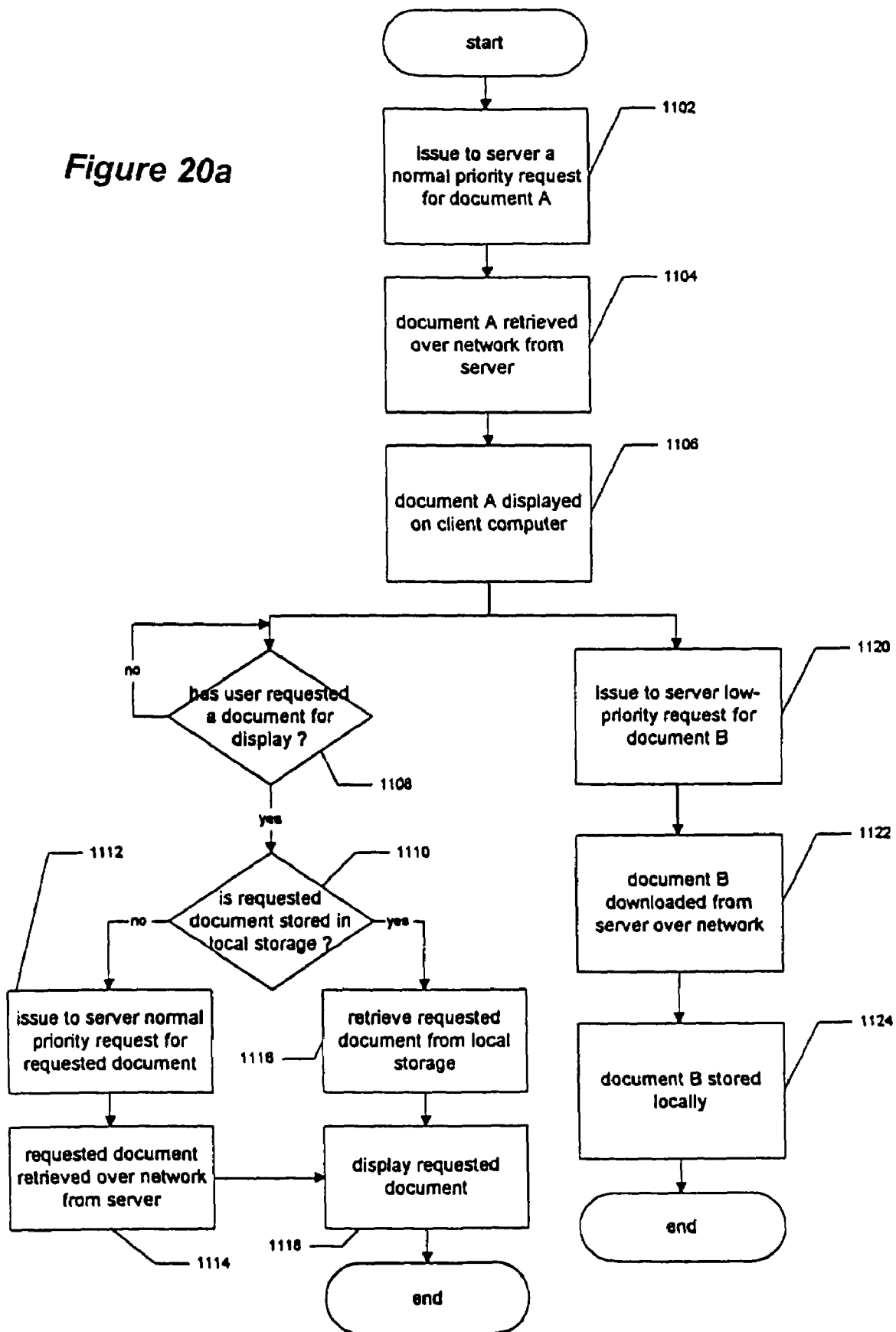
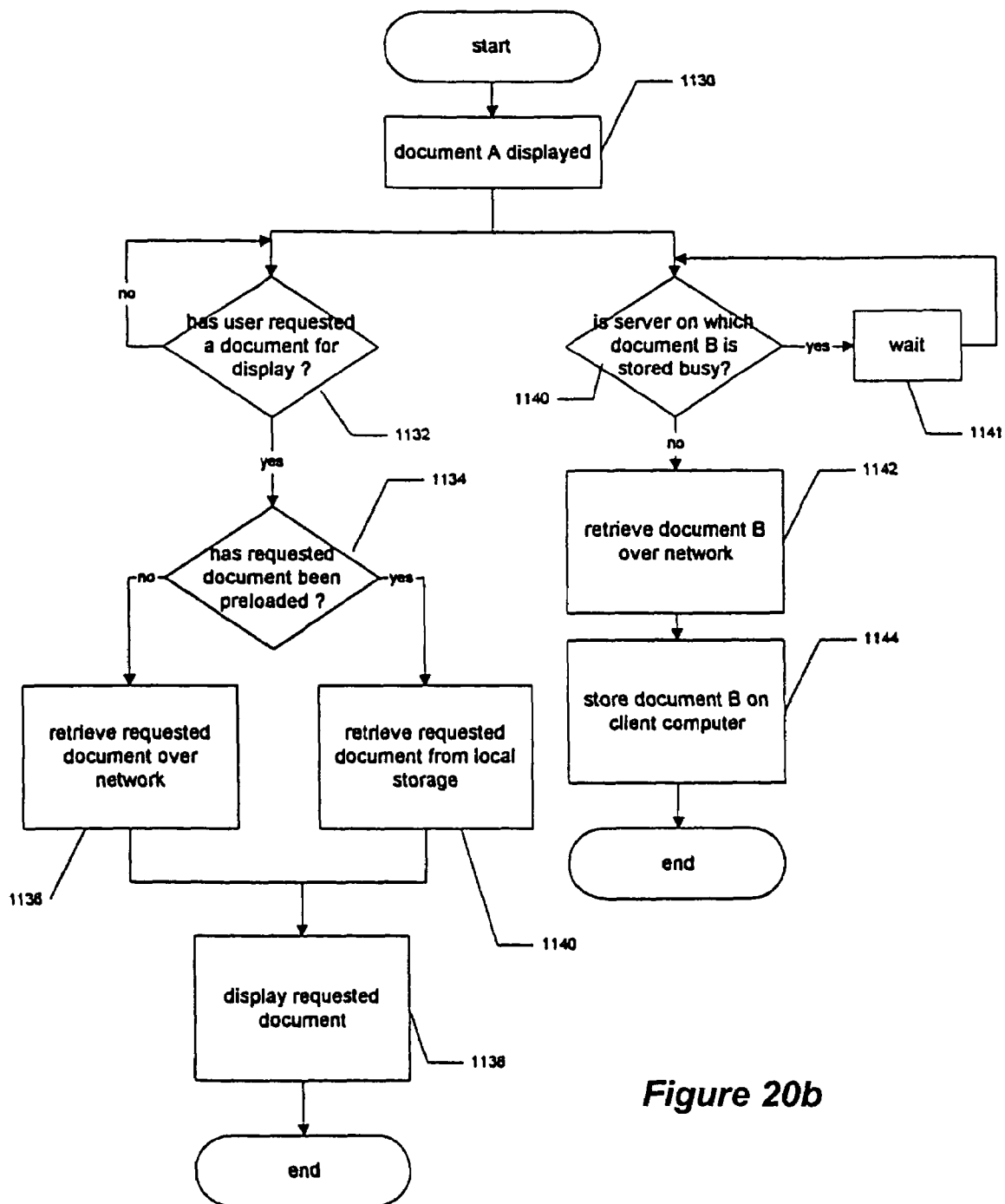


Figure 20a

*Figure 20b*

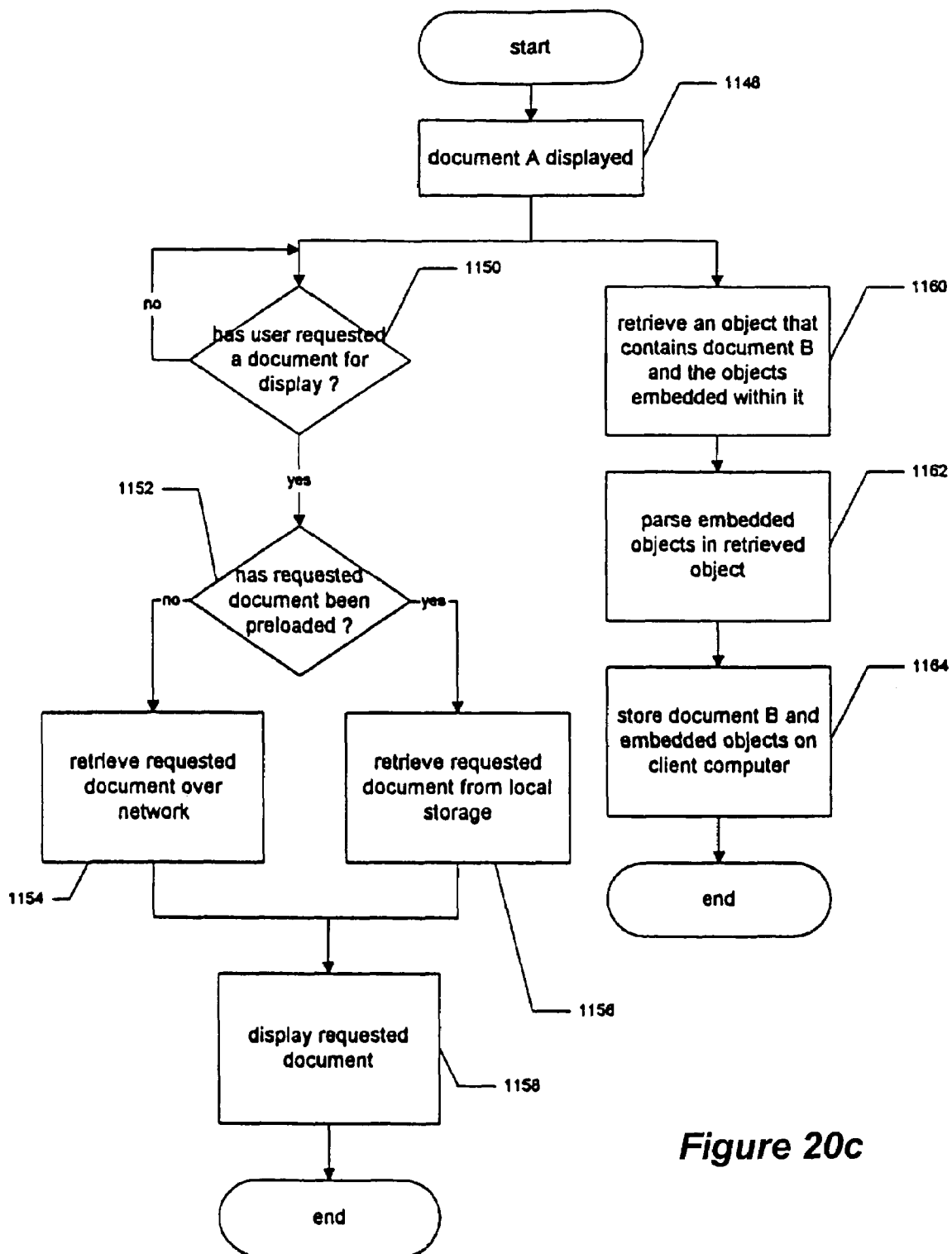
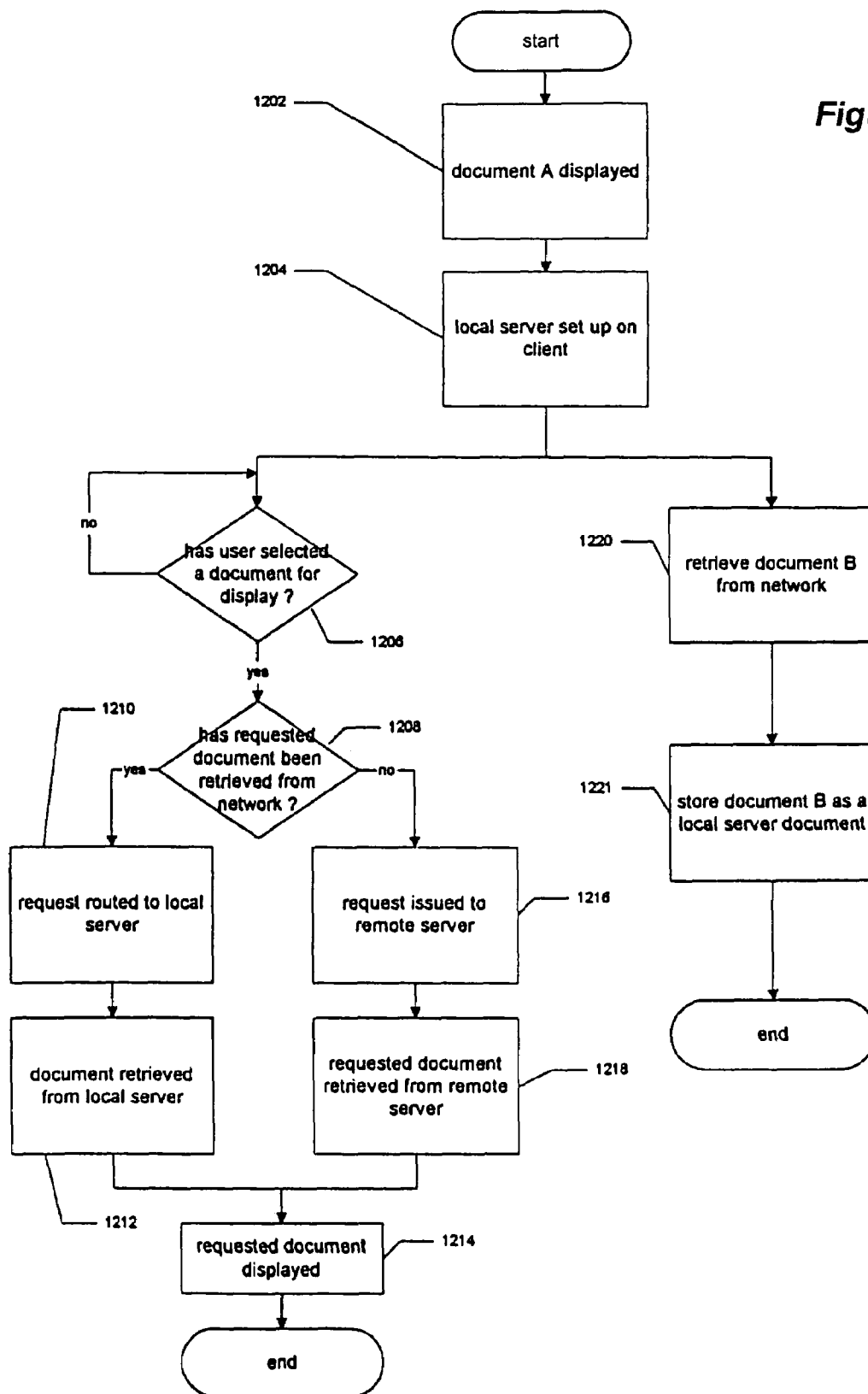
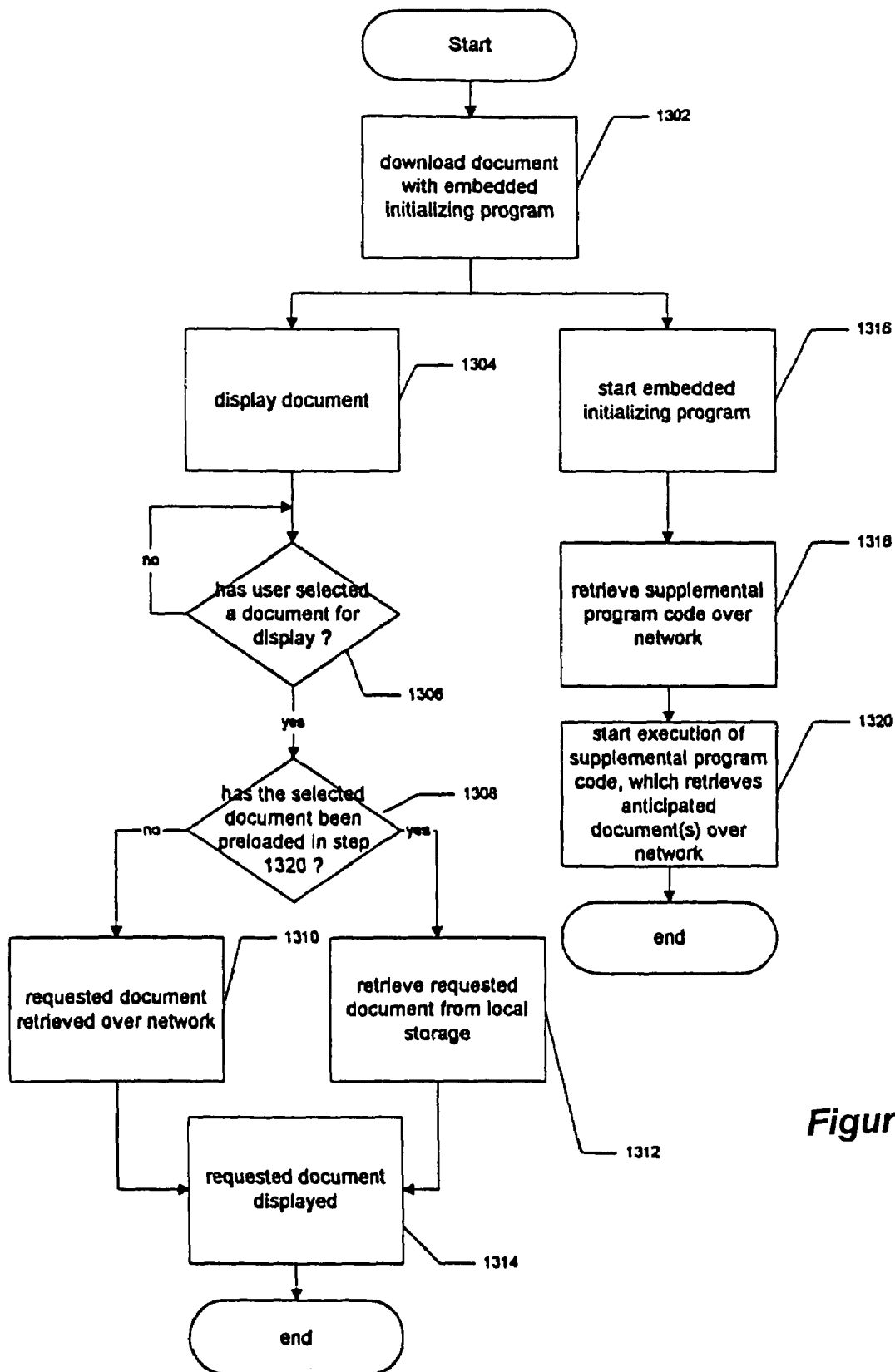
**Figure 20c**

Figure 21



**Figure 22**

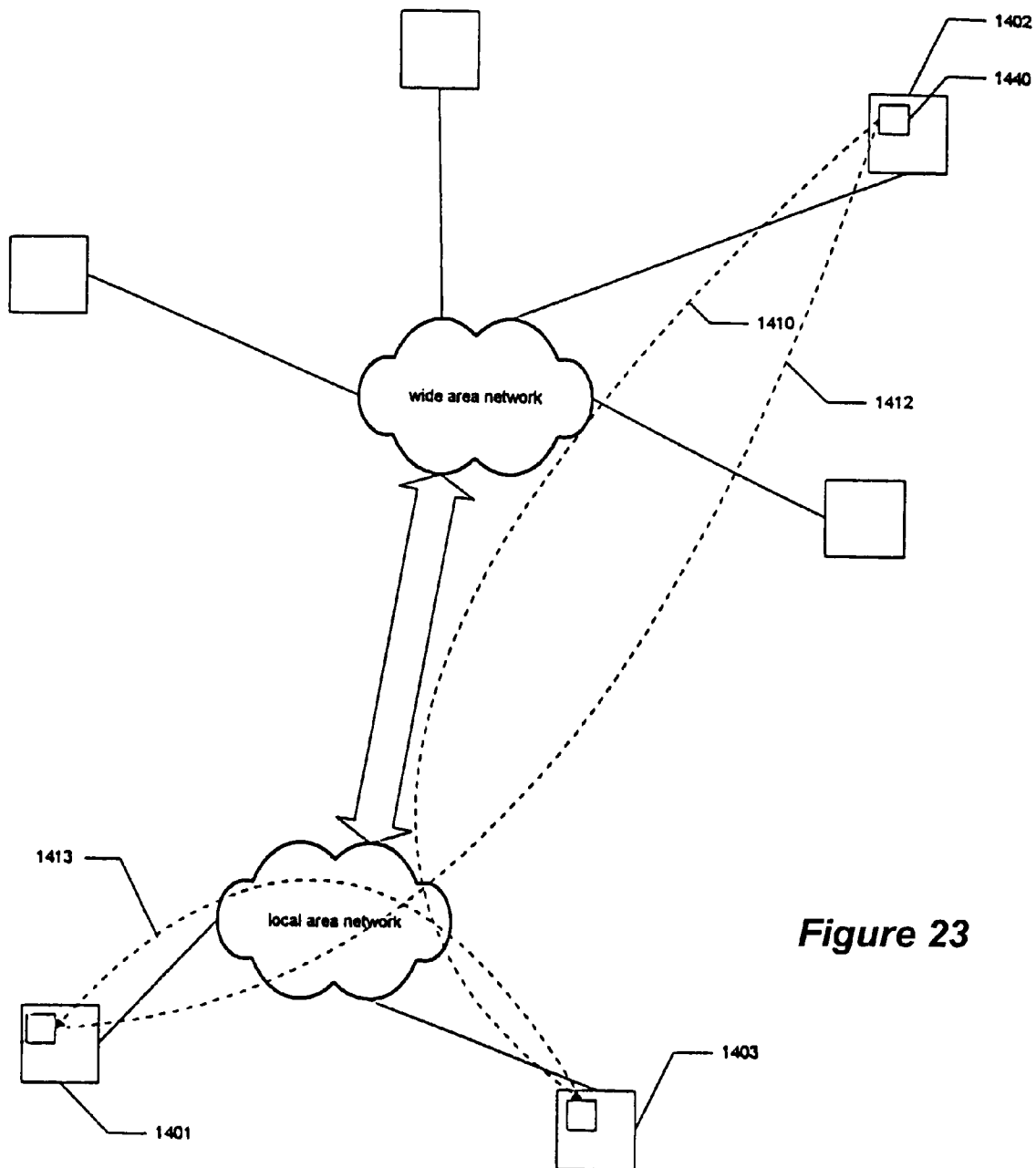


Figure 23

1

SYSTEM AND METHOD FOR INFORMATION RETRIEVAL EMPLOYING A PRELOADING PROCEDURE

CROSS REFERENCE TO RELATED APPLICATIONS

This application is a continuation of application Ser. No. 09/974,242, filed Oct. 9, 2001, now U.S. Pat. No. 6,604,103, issued Aug. 5, 2003, which is a continuation of Ser. No. 09/620,651, filed Jul. 20, 2000 now abandoned, which is a continuation of Ser. No. 09/083,382, filed May 22, 1998 now abandoned, which is a continuation-in-part of Ser. No. 08/918,912, filed Aug. 27, 1997, now U.S. Pat. No. 5,946,682, issued Aug. 31, 1999, which is a continuation of Ser. No. 08/474,921, filed Jun. 7, 1995, now U.S. Pat. No. 5,715,445, issued Feb. 3, 1998, which is a continuation of Ser. No. 08/300,343, filed Sep. 2, 1994, now abandoned. Application Ser. No. 09/083,382 filed May 22, 1998. This application also claims the benefit of provisional application Ser. No. 60/047,554, filed May 22, 1997. All of these applications are hereby fully incorporated by reference.

FIELD OF THE INVENTION

The present invention relates to a system for retrieving information from a database. More specifically, the present invention improves a system's response time so that a user's request to view new information is serviced quickly.

BACKGROUND AND SUMMARY

The recent proliferation of electronic text and multimedia databases has placed at society's fingertips a wealth of information and knowledge. Typically, a computer is employed that locates and retrieves information from the database in response to a user's input. The requested information is then displayed on the computer's monitor. Modern database systems permit efficient, comprehensive, and convenient access to an infinite variety of documents, publications, periodicals, and newspapers. Yet retrieving information from databases is often slow. Sometimes, this is caused by bandwidth limitations, such as when information is retrieved from remotely-located databases over an ordinary telephone line, a very narrow bottleneck. In other cases, slow retrieval is caused by a relatively slow local mass storage device (e.g., a CD-ROM drive).

There exists a compelling need for a database system that has a quicker response time so that information is displayed very soon after the user requests it. This need can be satisfied by effectively utilizing the time the user spends studying information on the display screen. In a database system or document retrieval system in one embodiment of the present invention, information that the user is likely to eventually request is preloaded into memory while the user is viewing other information. In some embodiments, the present invention takes advantage of the fact that it is possible to accurately predict the information that the user will eventually request be shown on the display. Some embodiments of the present invention also take advantage of the fact that the time that the user spends viewing displayed information is often sufficient to advantageously preload a substantial amount of information.

With these and other objects, advantages, and features of the invention that may become hereinafter apparent, the nature of the invention may be more clearly understood by

2

reference to the following detailed description of the invention, the appended claims, and to the several drawings herein.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1a is a block diagram of a general purpose computer.

FIG. 1b is a diagram of multiple computers connected together to form a network of computers and/or networks.

FIGS. 1c, 1d, and 1g are diagrams illustrating various procedures for installing and executing software.

FIGS. 1e and 1f are flow charts illustrating procedures for installing and executing software.

FIG. 2 is a representation of four search documents and three related documents.

FIG. 3 is a representation of four search documents and three related documents with a display view and one anticipated view designated.

FIGS. 4(a) and 4(b) are each a representation of four search documents and three related documents showing a display view and four anticipated views.

FIGS. 5(a) and 5(b) are each a representation of four search documents and three related documents showing various term views.

FIG. 6 is a representation of four search documents and three related documents showing various subdocument views.

FIG. 7 shows seven documents ordered according to four different ordering characteristics.

FIGS. 8, 9, and 10 are flow charts illustrating alternate embodiments of the present invention.

FIG. 11 is a diagram of the relationships between six statically-related documents.

FIG. 12 is a representation of a video display screen for a computer such as that of FIG. 1a.

FIGS. 13a, 13b, 13c, 13d, and 13e are flow charts illustrating the operation of embodiments of the present invention.

FIG. 14 is a chart illustrating one example of the type of profile information that may be provided in connection with a given document.

FIG. 15 is a flow chart illustrating how profile information can be used in an embodiment of the present invention.

FIG. 16 is a flow chart of the operation of a system in one embodiment of the present invention where a plurality of documents are preloaded in separate threads of execution.

FIGS. 17a to 17f are representations of a video display screen illustrating various features and embodiments of the present invention.

FIGS. 18a and 18b are flow charts illustrating the operation of various embodiments of the invention.

FIG. 19 is a flow chart illustrating how an embodiment of the present invention can effectively operate in a fee-based environment.

FIGS. 20a, 20b, and 20c are flow charts of the operation of embodiments of the present invention illustrating how server demands can be reduced in some circumstances.

FIG. 21 is a flow chart that illustrates the operation of a computer program illustrating some aspects of the present invention.

FIG. 22 is a flow chart that illustrates the use of an embedded program in connection with the present invention.

FIG. 23 is a network diagram illustrating an alternate method of preloading information on a network.

FIG. 1a is a block diagram of a general purpose computer 102 that can be used to implement the present invention. The computer 102 has a central processing unit (CPU) 104, memory 113, and input/output (i/o) circuitry 112. The CPU 104 is connected to the memory 113 and the i/o circuitry 112. The i/o circuitry permits the CPU 104 to access various peripheral devices, such as the display or monitor 108, local storage 106, and input device(s) 110. The input device(s) 110 may include a keyboard, mouse, pen, voice-recognition circuitry and/or software, or any other input device. Some type of secondary or mass storage 106 is generally used, and could be, for example, a hard disk or optical drive. The storage 106 can also be eliminated by providing a sufficient amount of memory 113. Either the storage 106 or the memory 113 could act as a program storage medium that holds instructions or source code. The i/o circuitry 112 is also connected to a network 114, thereby connecting the computer 102 to other computers or devices.

FIG. 1b is a representation of multiple computers (251, 252, 253, 254, 255, 256, and 257) connected together to form a network of computers and/or networks. Computers 251, 252, and 256 are shown connected to wide area network (WAN) 263, whereas computers 253, 254, 255, and 257 are shown interconnected by local area network (LAN) 261. The LAN 261 is connected to the WAN 263 by connection 262. Various network resources, such as databases of documents or other objects, are stored on one or more the computers shown in FIG. 1b.

In a networked environment, such as that of FIG. 1b, there are numerous ways in which software can be installed, distributed, and/or executed on the various computers on the network. FIG. 1c illustrates a conventional way in which desktop software is installed and executed. In FIG. 1c, a computer program 1003 is installed at the computer 1001 through some type of installation program typically started by the user of the computer 1001, and executed on the computer 1001. During installation, the program 1003 may need to be configured at the computer 1001 for use with the network in order to enable access to other computers on the network (e.g., 1002 and 1012). After installation, the computer program 1003 resides and executes at the computer 1001, and is persistent. When the computer 1001 is shut down or restarted, the program continues to be stored at the client on non-volatile storage media. Upon restarting the computer 1001, the program 1003 is available for use without reinstallation.

FIG. 1d shows a different embodiment. When the network-connected computer 1001 connects to or downloads an object stored on the remote computer 1002 over the network, a program 1005 embedded within the downloaded document or object is installed on the computer 1001 and is executed on the computer 1001. FIG. 1e is a flow chart that illustrates one possible installation procedure that is carried out when the computer 1001 accesses the program 1005.

The computer 1001 identifies at 1020 one or more programs embedded within the accessed object. The client computer then determines whether each embedded program has been installed previously on the computer 1001. This can be done by searching the computer's storage or system registry for the program or for the program's identifying characteristics. In Microsoft's ActiveX/COM architecture, for example, this is done by searching the registry for an instance of the program's globally unique identifier (GUID) in the system registry.

If the embedded program has been installed on the client computer, the previously installed program is retrieved from local storage at 1030, and executed at 1028. However, if the program has not been already installed on the client computer, it is retrieved over the network (1023), and installed on the client computer. The installation process will typically involve updating a system registry or other persistent storage with identifying information on the computer 1001.

Preferably, the program is installed at 1024 such that it need not be downloaded again over the network when it is encountered embedded within another object. For example, if the computer 1001 were to access an object on computer 1012 that had program 1005 embedded within it, the program 1005 would not need to be installed again because it has already been installed when computer 1002 was accessed.

FIG. 1f is flow chart illustrating a different embodiment of the present invention. In this system, when the computer 1001 encounters an object on computer 1002, it identifies at 1040 each program embedded within the object. It then retrieves one or more programs over the network, and then installs them at the client computer 1001, but without the use of a persistent storage mechanism. Thus, although the program is executed on the client computer 1001, the embedded program must be downloaded each time it is encountered because no persistent storage mechanism is used. This type of installation procedure may be more secure, and has been used in some of the early Java implementations.

A system in which software is downloaded over the network, perhaps from an untrusted server, has significant security risks associated with it, and for this reason, security restrictions may be placed on computer programs downloaded from the network. Thus, a downloaded computer program may be unable access some of the resources of a client computer or of the network generally. In some embodiments, however, a downloaded program may be tested for authenticity and safety through a code signing procedure, or through a code verifying procedure. If such a program passes such authenticity tests, it may be given more complete access to system or network resources.

In yet another embodiment, shown in FIG. 1g, the network-connected computer 1001 connects to the remote computer 1002 over the network, but the program 1021 executes on the remote computer 1002. Display information and/or instructions are sent from the remote computer 1002 to the computer 1001, and this information and/or instructions are interpreted by a terminal program or a thin client program 1023, which updates the display. Input device events (e.g., mouse clicks and keyboard events) caused by the user at the computer 1001 are sent to the remote computer 1002 so as to appropriately alter the execution of the program 1021 executing at the remote computer. In some implementations, it appears to the user of computer 1001 that the program 1021 is executing on computer 1001, even though the program 1021 is actually executing on the computer 1002. This scenario, or one similar to it, is employed by some of the thin-client computing environments, such as Citrix Corporation's WinFrame solution, or Microsoft's forthcoming Hydra initiative.

Each of the described techniques for installation and/or use of software can be implemented in connection with the present invention. For example, software used for carrying out one or more embodiments of the present invention may be installed and executed in accordance with the techniques described above.

In FIG. 2, four documents that might correspond to search documents found as a result of a query are shown. The query

5

may be formulated to find all the documents in a given database that include the phrase "Hadley v. Baxendale." Each X in the search documents **100A**, **200A**, **300A**, and **400A** represents an occurrence of the phrase "Hadley v. Baxendale." As can be seen, the phrase "Hadley v. Baxendale" can be found in search document **100A** at two separate locations. Document **200A** has six occurrences, and search document **300A** has three. Search document **400A** has one occurrence—the title of search document **400A** is "Hadley v. Baxendale."

There are also "related documents" (**500A**, **600A**, and **700A**) shown in FIG. 2. A related document is a document that is somehow explicitly associated, linked, or otherwise connected to one of the search documents. For example, if search document **100A** is a judicial opinion, a related document might be a subsequent opinion in the same case (e.g., an decision on appeal). Other related documents might be an opinion or scholarly article that cites or discusses search document **100A**, or a list of judicial opinions that cite the search document. Any document that is usefully associated with the search document can be considered a related document. Often the related document does not satisfy the query, so it is usually not one of the search documents. In some circumstances, however, the related document might satisfy the query, so it can be a search document.

Related documents may also be related only to a particular view within a search document. For example, a search document that is a judicial opinion may have numerous other judicial opinions cited in the text of the opinion. These cited opinions may be "related documents," but often they relate only to a particular view within the document. Depending on the implementation of the database system, they might not be considered to be "related" to the search document as a whole. Thus, they are available as related documents only when the corresponding cite is within the currently displayed view. In such an implementation, the related documents are dependent on the view shown on the monitor at any given time.

FIG. 3 shows the representation of the four search documents that satisfy the user's query. The search documents are ordered by an ordering characteristic, such as the date of publication. Other ordering characteristics can be used as appropriate for a given situation (e.g., number of query terms in a document, statistical relevance of the documents, type of document, etc.). Any ordering characteristic that permits the search documents to be distinguished from one another can be appropriate. In the example of FIG. 3, search document **100A** is the first search document according to the ordering characteristic, and view **101A** (shaded) in search document **100A** is the display view shown on the monitor. (The view shown on the monitor at any given time is the "display view.") Once view **101A** is displayed on the monitor, the user reads, studies or otherwise observes the displayed information. When the user wishes to change the display view, he or she uses an input device to cause the system to display either (a) a different view in the search document **100A**, or (b) a view from one of the other documents **200A**, **300A**, **400A**, **500A**, **600A**, or **700A**.

The user uses one or more input devices to request particular views. For example, an input device might be a keyboard that includes a "next page" key and a "next document" key. The "next page" key requests the next successive view (view **102A**) within the document currently being viewed (document **100A**). The "next document" view requests the first view (view **201A**) of the next successive search document according to the ordering characteristic (document **200A**). Many database systems have "next page"

6

and "next document" commands or keys (e.g., Westlaw, LEXIS/NEXIS, and West Publishing Company's CD-ROM products), as well as others (e.g., "previous document," "previous page"). Westlaw also permits a user to request a particular search document or "page" by typing a command. For example, to view search document three (**300A**), the user types "r3"; to request page 2 (i.e., view 2) within the currently displayed document, the user types "p2." And in some systems, multiple commands can be executed together by separating them with a semicolon, so page two from document three (view **302A**) can be requested with a single command: "r3;p2."

In the systems of the prior art, when the database system receives the command to display a different view, the requested view must be loaded from the database before it can be displayed on the monitor or display. Since retrieving information from the database is time-consuming, this loading process is undesirably slow. But in a system employing the present invention, the time required to respond to the user's request for a different view (the "requested view") is reduced by taking advantage of the fact that it is often possible to predict the requested view before the user actually requests it. In the present invention, the view(s) that the user is likely to next request are preloaded while the user is reading the displayed view.

Thus, in one embodiment of the present invention, the view or views (i.e., anticipated view(s)) that are likely to be next requested by the user are "preloaded" (e.g., in the background) to the extent permitted by the time the user spends reading or studying the display view. When the user does request that a different view be displayed (i.e., the user requests a "requested view"), the requested view can be very quickly displayed on the monitor if it has already been preloaded into memory. Thus, if the requested view is one of the anticipated views, the database system is able to quickly respond to the user's request for the requested view.

As shown in FIG. 3, while the user is reading or studying the display view **101A**, view **201A** is identified as an anticipated view (signified by the arrow from view **101A** to view **201A**). View **201A** is likely to be requested by the user because it is the first view of the "next" search document (as defined by the ordering characteristic) following search document **100A**. And while the display view **101A** is being viewed by the user, the database system will preload view **201A** from the database into memory, before it is actually requested by the user. After view **201A** is preloaded into memory, the input device is checked to see if the user has requested that another view be displayed. If the user has requested that a requested view be displayed, the database system checks to see if the requested view has been loaded into memory (e.g., as the preloaded anticipated view). If the requested view is view **201A**, it will have been loaded into memory as the anticipated view, so view **201A** is retrieved from memory and displayed on the monitor. Since loading the requested view from memory is much faster than loading the requested view from the database, the time required to respond to the user's request for the requested view is shortened dramatically. If the requested view is not in memory, however, it must be retrieved from the database.

Instead of loading the entire anticipated view before checking the input device, in other embodiments of the present invention the input device is monitored during the time the anticipated view is being preloaded into the database. If the user requests a requested view, the preloading of the anticipated view stops and the user's request is serviced. This ensures that the system is very responsive to the user's input. Such an embodiment can be implemented by check-

ing the input device each time a segment (i.e., a portion) of the anticipated view is preloaded. If the computer retrieving information from the database is running a multitasking and/or multithreading operating system, such an embodiment can alternatively be carried out using the various techniques appropriate for such an operating system.

FIG. 4(a) shows a situation where view **101A** (shaded) is the display view, and the retrieval system has identified four views **102A**, **501A**, **201A**, and **401A** as anticipated views. View **102A** is likely to be requested by the user when the displayed view is view **101A** because it is the next view in the document that the user is currently viewing. View **501A** is a candidate for the requested view because it is the first view from a document (**500A**) that relates to the search document (**100A**) that the user is currently viewing. View **401A** is also an anticipated view because the user might wish to view the document that represents the opposite extreme of the ordering characteristic (e.g., the oldest document). And as described above, view **201A** is also likely to be requested by the user.

In the embodiment of FIG. 4(a), the retrieval system will attempt to load as many of these anticipated views as possible while the user is studying the display view **101A**. If enough time passes before the user requests a requested view, the retrieval system may preload all four of the anticipated views, thereby enhancing the likelihood that the next requested view will be in memory.

Once the user issues a request for a requested view, the requested view is loaded from memory (or from the database, if necessary) and displayed on the monitor. The process of determining and preloading anticipated views then starts over. For example, if the requested view is view **201A**, the display view will then become view **201A** (shaded) as shown in FIG. 4(b). The anticipated views would also change, and might be identified as indicated by the arrows.

FIG. 5(a) shows another representation of four search documents showing term views **111**, **112**, **211**, **212**, **213**, **214**, **311A**, **312A**, and **411**. In FIG. 5(a), a term view is a view that has at least one search term from the query. And as can be seen from document **100A** in FIG. 5(a), the boundaries of these term views may or may not correspond to the boundaries of views **101A**, **102A**, **103A**, and **104A**. Term views may also be anticipated views because the user might request as a requested view the next view having one or more of the terms in the query. Some systems provide a command for this purpose (e.g., in Westlaw, the command is "t").

FIG. 5(b) shows the representation of the four search documents showing other term views **171**, **271**, **272**, **371**, and **471**. These term views are made up of a small number of words surrounding each occurrence of a search term in the search documents. Since the number of words surrounding the search terms is small, more than one set of words can fit on the screen at a given time. Thus, the term view in this embodiment includes information from different parts of the document. The "KWIC" display format in the LEXIS/NEXIS system operates similarly.

FIG. 6 shows another representation of the four search documents showing subdocument views **121**, **122**, **131**, **141**, **221**, **231**, **232**, **233**, **321**, **331**, **421**, **431**, and **441**. The subdocuments are shown in FIG. 6 as **120**, **130**, **140**, **220**, **230**, **240**, **320A**, **330A**, **420**, **430**, and **440**. A subdocument is any logically separable or identifiable portion of a document. For example, if a document is a judicial opinion, there might be subdocuments for the title and citation for the case,

for each of the headnotes, for the opinion itself, and for any dissenting opinions. A subdocument view is a view within a subdocument.

Subdocument views may be anticipated views because the user is often particularly interested in a particular portion of the search documents. If the search documents consist of a series of judicial opinions, for example, a user may only wish to view, for each of the search documents, the subdocument for the majority opinion (and not the headnotes, dissenting opinions, etc.). Thus, it may be appropriate for the anticipated views to be drawn primarily from a particular type of subdocument.

In other situations, however, the user may only wish to see the first subdocument view for each subdocument. It would be appropriate in these situations for the anticipated views to be primarily the first views from the various subdocuments within each document.

The retrieval system of the present invention identifies anticipated documents by focussing on the current display view. The current display view gives clues as to which view might be requested by the user because the display view identifies the user's progress in browsing the search documents. In other words, the current display view identifies which search document in the sequence of search documents is currently being viewed. This information is useful because the search document immediately following and preceding the current search document (as defined by the ordering characteristic) is often the search document next requested by the user.

The view displayed just prior to the displayed view might also be a consideration in determining the anticipated views if it tends to show a pattern that can identify the user's next requested view. For example, referring to FIG. 6, if the user requests view **131** of search document **100A**, and then requests view **231** of search document **200A**, the retrieval system can consider these two consecutive display views and determine that an appropriate anticipated view is view **331** of search document **300A**. View **331** is the first view of subdocument **330A**, which could be the same type as subdocuments **130** and **230**, the two subdocuments previously viewed by the user. Since the goal is to accurately predict the next view, considering the views that the user requested in the past may be helpful if it tends to identify the views that the user will request in the future.

In general, any appropriate adaptive prediction scheme can be used that uses the user's history of requested views (and display views) to accurately determine which views are likely to be next requested by the user. It might be appropriate in some cases to consider many display views in determining appropriate anticipated views. Longer histories may tend to identify patterns that would not show up if only a small number of recent display views are considered.

Tendencies can even be monitored over more than one research session in situations where a particular user or group of users tend to request views in a particular pattern each time research is done. In addition, the user could be prompted to indicate the type of research being undertaken, which may give clues as to what type of anticipated views are appropriate for efficient operation. Finally, the particular databases used or type of research being done can be monitored by the database system and advantageously taken into account in determining anticipated views.

In the preferred embodiments of the present invention, the anticipated views are drawn from both related documents and search documents. A fundamental distinction between related documents and search documents is that related documents are statically-related to the search documents,

whereas search documents are dynamically-related to one another. This difference is significant because unlike statically-related documents, no predefined link needs to be set up for search documents. A statically-related document is always associated with a particular document, regardless of the query (the related document is therefore statically-related). The search documents, on the other hand, are related to each other by the query. Since the query changes with each search, the search documents are considered dynamically-related to one another:

Some of the recent CD-ROM products have implemented features such as hyperlinked text, and timeline-linked text (clicking on a timeline item will take the user to a relevant article). See *The Top 100 CD-ROMs*, *PC Magazine*, Sep. 14, 1994, p. 115. Links of this nature are static because they always apply and do not depend on any particular query run by the user.

The search documents are ordered by an ordering characteristic as described previously. Thus, when a "next document" is requested, it is assumed that the search document requested by a "next document" command is the search document that is "next" according to the ordering characteristic. If the search documents are ordered by publication date, for example, the "next document" will be interpreted as a request for the search document with the next oldest publication date.

In one embodiment of the present invention, it is possible to make a number of different ordering characteristics available for use by the user in browsing the search documents. For example, FIG. 7 shows seven documents labeled "a" through "g" ordered according to four different ordering characteristics. When the display view is in document "a," the "next document" command can be a request for four different documents (i.e., "b," "e," "f," or "c"), depending on the particular ordering characteristic used. More than one ordering characteristic must therefore be considered when determining anticipated views if the user is capable of moving to a "next document" in the context of more than one ordering characteristic. This feature can be enabled by an input device command that allows the user to select the desired ordering characteristic.

The present invention is applicable to single-user, multiple-user, and many-user databases, but the present invention is most effective when used in connection with single-user databases. The efficient operation of the invention depends on being able to retrieve data from the database very frequently, perhaps continually. The present invention is quite effective with single-user databases such as those on CD-ROM or other mass storage devices (this might also include a hard drive implementation). In a single-user database, no other demands are being made on the database by the other users, so the database is often idle.

But since a many-user or multiple-user database must be shared among more than one user, such a database will often be receiving simultaneous and continual requests for data. Databases in such a system are rarely idle, so there is little time to preload anticipated views into memory. In such a situation, the present invention will not be as effective in improving the response time to users' requests for requested views. But in many-user or multiple-user database systems where the database is not as busy, the present invention can be effective in reducing response times to users' requests for information.

FIG. 8 is a flow chart of the operation of the database system in one embodiment of the present invention. A system in one embodiment of the present invention begins by executing a query to identify the search documents. This

step is carried out by search logic 51. The remaining steps shown in FIG. 8 (described below) are carried out by retrieval logic 52. Both the search logic 51 and the retrieval logic 52 are often software, but need not be. As one skilled in the art will recognize, in a software implementation the search logic 51 and the retrieval logic 52 may or may not be integral or intertwined parts of the same computer program.

As dictated by the retrieval logic 52, the database system then loads into memory a view from one of the search documents. See FIG. 8. This first display view is then displayed on the monitor. Normally the user will take a few moments to read or study the display view. During this time, one or more anticipated views are identified. The anticipated views are views that the user is likely to request be displayed on the monitor after the display view.

The database system then begins to preload these anticipated views into memory from the database, while also continually monitoring the input device to determine if the user has issued a request to display a different view (i.e., a "requested view") on the monitor. Anticipated views are loaded into memory until the user requests a requested view.

When the user does makes such a request, the database system then determines whether the requested view is in memory. The requested view may be in memory because it could have been preloaded into memory as an anticipated view. If the requested view is in memory, the requested view becomes the new display view, and it is displayed on the monitor. But if the requested view is not in memory, the requested view must first be loaded from the database before it can be displayed on the monitor as the display view.

The anticipated views are a function of the display view because the views that the user is likely to request depend to some degree on the view the user is currently reading. In other words, those views that are anticipated views when view 101A is the display view are not likely to be the same as the anticipated views when view 202A is the display view. Therefore, as shown in FIG. 8, the anticipated views are determined each time the display view changes.

When the display view is changed, the anticipated views for the prior display view can remain in memory so that they are available if they are ever requested by the user. But if memory is limited, the anticipated views for the prior display view can be deleted from memory, preferably in an efficient matter (e.g., anticipated views common to both the new display view and the prior display view are not deleted from memory). It is best to delete those views that are not likely to be requested by the user. It may also be appropriate to consider whether a view is likely to become an anticipated view in the future.

FIG. 9 shows a flow chart representing another embodiment of the present invention where anticipated views from prior display views are deleted if memory is full. The views deleted are those that are not anticipated views for the new display view. This will presumably make room for new anticipated views to be preloaded into memory (if not all of the anticipated views are already in memory).

The number of anticipated views for a given display view does not have to be a predetermined or constant number, but rather can vary depending on memory available. Typically, the number of anticipated views for a display view is a trade-off between the amount of memory available and the desired speed of retrieval. In instances where memory is plentiful, where the number of search documents is few, and/or where the search documents are small, it may be possible for all of the search documents to be completely loaded into memory. In such a situation, the number of anticipated views for a given display view could be as high

11

as the total number of views in the search documents. At the other end of the spectrum, there might be only one or two anticipated views for each display view if memory is limited.

Embodiments of the present invention can vary as to how anticipated views are preloaded into memory. In the embodiments of FIGS. 8 and 9, one anticipated view at a time is preloaded into memory, and the retrieval system does not begin preloading a second anticipated view into memory until the prior anticipated view is completely preloaded into memory. In other embodiments, anticipated views are simultaneously preloaded.

Simultaneous preloading of multiple anticipated views can be done in a number of ways. In a multitasking operating system, for example, an appropriate time-slicing procedure can be used to preload the anticipated views so that they are preloaded simultaneously. In another embodiment, one segment from each anticipated view is preloaded in turn, and the cycle is repeated until all the anticipated views are fully preloaded into memory (or until the user's request for a requested view interrupts the preloading process). A segment is any portion of an anticipated view, such as one or two lines or even a single byte of the anticipated view.

FIG. 10 shows a simple implementation of the simultaneous preload concept, where the database system preloads a segment of a first anticipated view into memory, and then preloads a segment of a second anticipated view into memory. These steps continue until either the user requests a requested view, or both anticipated views are fully preloaded into memory. When the user requests a requested view, the database system checks to see if that requested view is in memory. If the requested view is only partially preloaded into memory, that portion in memory can be written to the monitor and the remaining portion loaded from the database. The response time in this situation will still be better than if the entire requested view has to be loaded from the database.

In another embodiment of the invention, the use of profile information is employed to assist in the selection of views or documents to preload, as illustrated in FIGS. 11–12, 13a–13e, and 14–16. For example, FIG. 11 is a diagram of the relationships between six objects or documents 301–306. The six documents are linked to each other in the manner shown and hereinafter described. Document 301 contains three links (310, 312, and 314); one to each of the documents 302, 303, and 304. Document 302 contains two links, one link 316 to document 305, and another link 318 to document 306. Document 305 contains a link 320 back to document 302, and document 306 contains a link 322 to document 304. Each of these documents is stored on a server within a network, and may incorporate or have embedded within it objects stored on other servers. The documents 301–306 may be stored on the same server, or may be stored on various computers distributed throughout the network.

FIG. 12 shows a representation of a video display screen 404 for a computer such as that of FIG. 1. The area 404 represents the area on a screen within which images, text, video, and other types of data or multimedia objects can be displayed and manipulated. On the display 404 shown in FIG. 12, a number of icons or objects 402 are arranged, along with another type of object, window 406. The window 406 is a representation of a document retrieval, browsing, and/or viewing program that is used to view information either stored locally on the computer or retrieved over a network. The window 406 has a title area 408 that displays the title of the document being displayed. The title area 408 could also display the location or address of the document

12

being displayed, or also the universal resource locator of the document being displayed. Alternatively, an additional area within the window could be used for displaying the universal resource locator. The contents of the document are shown in displayed within the window 406 in FIG. 12, but it should be understood that the contents could be displayed in other ways. For example, the contents could be displayed on the entire desktop, or a portion of the desktop. In another embodiment, the contents might be scrolled on the screen, perhaps under other windows. In addition, it should be understood that where a document has only a single view, or is treated as having only a single view, the “document” essentially becomes the same as the “view.” In such a situation, a scroll bar (not shown) may be used to allow a user to effectively expand the size of the display, thereby allowing the user to see the entire document/view.

Shown within the document viewing area of the window 406 in FIG. 12 is the contents of the document 301 from FIG. 11. The document 301 has been displayed in the window 406 in response to a user query, which might involve a key word search or might involve the user specifying the identity, address, or resource locator of document 301. The document 301 could be also be displayed within the window 406 in response to the selection of a link in another document (not shown) that points to the document 301.

FIG. 13a is flow chart representing the operation of one embodiment of the present invention in which profile information is used to assist in the selection of documents to preload. In FIG. 13a, the documents 302 and 303 are assumed to be stored on the network on a (preferably) remote database server. At 501 in FIG. 13a, document 301 is retrieved from the server by a document retrieval program that executes on a client computer, such as, for example, the computer 257 in FIG. 1b. The document 301 may be stored on any of the other computers shown in FIG. 1b, including computer 257. Along with the document 301, the document retrieval program retrieves profile information that is preferably (but not necessarily) embedded within or is stored with document 301. The profile information can be used to provide information about document 301, including information about documents that are related to document 301, as described below.

At 502, the document viewing program renders document 301 in the window 406, as shown in FIG. 12a. Once document 301 is retrieved, the viewing program begins at 504 retrieving from the network the document 302. During this time, document 301 continues to be displayed in window 406, and the user is free to read, scroll through, or otherwise interact with the document 301 shown in the window 406 in FIG. 12a. Thus, document 302 is retrieved over the network (504) and stored into the memory or local storage (505) while the user is viewing document 301.

After document 302 has been retrieved, if the user has not requested (e.g., through the input device) at 506 that another document be displayed, the document viewing program at 508 retrieves document 303 over the network, and this document is then stored in memory or local storage. The document 301 is still displayed in the window 406 at this point. If the user still has not requested that another document be displayed at 510, the document 304 is retrieved from the network and stored in memory or local storage by the document viewing program in 512.

At 514, the document viewing program in the embodiment of FIG. 13a stops preloading documents, and waits until the user requests that a new document be displayed. When the user does request that a new document be dis-

13

played in the viewing program at **514**, the viewing program determines at **516** whether the requested document is one of the documents that has already been retrieved and stored in local storage. If so, then the locally-stored version of the requested document is retrieved from memory or local storage at **518**. The locally-stored version of the requested document is then checked at **519** to see if it is out of date. If the requested document has content that may change often, it may be that the version of the requested document that is stored in local storage is not sufficiently new and is out of date or "stale." This condition can be determined by monitoring the amount of time since the document was originally preloaded, or by inspecting the time stamp on the requested document stored on the network and comparing it to the time stamp for the locally-stored version to determine whether the locally-stored version has changed.

If the locally stored version of the requested document at **519** is not out of date, then it is displayed at **524**. If the preloaded version at **519** is out of date, or if the requested document has not been preloaded at all (**516**), then the requested document is retrieved over the network and is displayed at **524**.

In another embodiment, the document viewing program may continue preloading additional documents at **512** in FIG. **13a**. For example, the document viewing program could begin to preload the documents that are linked to by the documents that are already preloaded (i.e., documents **302**, **303**, and **304**). In the set of documents shown in FIG. **11**, this would mean that the document viewing program would download over the network additional documents **305** and **306**. Thus, the present invention need not be limited to the preloading of only a single level of linked documents, but rather, could extend to the preloading of two or more levels.

As described in connection with FIG. **13a**, the document viewing program retrieves over the network documents **302**, **303**, and **304** while the user is viewing document **301**. And as shown in FIG. **13a**, document **302** is preloaded first, followed by document **303**, and then by document **304**. In some embodiments, the document viewing program executing on the client computer in FIG. **13a** uses the profile information to determine the order in which the documents **302**, **303**, and **304** are to be retrieved. In other words, in some embodiments, the database server determines the order in which the document viewing program executing on the client preloads the links within document **301**. This procedure can be quite effective because the database server may have useful information that can help to predict the documents that the user will request be displayed at the client computer. For example, a server that keeps track of the frequency that users select the links within document **301** may find that one or two links are selected very often, whereas other links are selected rarely. The server can use this information to instruct the client as to the most efficient order in which to preload documents.

FIG. **14** is a chart illustrating one example of the type of profile information that could be provided with document **301** at **501** of FIG. **13a**. As shown in FIG. **14**, for each of the documents identified in the profile information, the historical percentage of users that have selected the document are identified. The first three documents (**302**, **303**, and **304**) are documents that are linked to by the document **301** as shown in FIGS. **3** and **4**. The fourth document (document **319**) is not linked to or otherwise related to document **301**, but the profile information nevertheless tells us that 2% of the people request document **319** when document **301** is displayed. Thus, the profile information tells us that document

14

302 is, historically, the most likely document to be selected. The document viewing program can use this information to ensure that document **302** is preloaded when document **301** has been rendered in the window **406** because at least according to this statistical information, document **302** is likely to be requested by the user who is viewing (or who has at least retrieved) document **301**.

Other data that might be included in the document profile might be the server or database in which each document is stored. This information is shown in FIG. **14**, and identifies the document **302** as being from the server "gaylords.com," and document **303** as being from the "same" database server, which means the database server from which document **301** has been retrieved. Normally, the profile information of FIG. **14** would be stored in a particular format or data structure, or even as source code, and either embedded within the document **301**, or downloaded by the document viewing program along with the document **301**.

In operation, the server sends to the document viewing program the information of FIG. **14**, and the document viewing program can choose to ignore it, or could choose to act upon it in some way. Thus, the document viewing program in one embodiment engages in some form of interpretation of the profile information and it exercises some control over how the information is used. In another embodiment, however, the profile information could simply consist of a list of documents that the document viewing program uses to select what documents to preload. The list of documents might be ordered so the document viewing program could determine relative priorities among the documents, but the document viewing program may not engage in interpretation of any statistical data or other data sent from the server. Such an embodiment may be implemented by actually programming a program embedded within a document to retrieve certain documents, thus effectively hard-coding the documents that are to be preloaded. Another embodiment may use a program embedded within an object or a document that reads a parameter list and uses the parameter list to select the documents to be preloaded. Where the profile information is not used, some type of predefined procedure could be followed for selecting documents to preload, and this procedure may involve preloading the documents that are linked to by the document **301** (the displayed document).

FIG. **15** is a flow chart which illustrates how the profile information might be used in an embodiment of the present invention. At **594**, the document **301** is retrieved over the network, and at **595**, the profile information for document **301** is retrieved over the network. The profile information is then analyzed at **596** to determine which documents to preload when document **301** is being displayed. At **597**, the profile information is further analyzed to determine the order in which the documents identified at **596** should be preloaded. Thus, in this embodiment, the profile information not only identifies the order in which to preload documents, but also identifies at **596** the documents that are to be preloaded.

FIG. **13b** is a continuation flow chart of FIG. **13a**, where the user has requested that document **302** be displayed. In other words, the requested document at **516** of FIG. **12a** is document **302**. Thus, initially in FIG. **13b** the document **302** and associated profile information associated with document **302** is retrieved at **529** and then displayed at **530** within window **406**. At **532**, the viewing program checks to see if the user has requested that another document be displayed. If not, document **305**, which is linked to by document **302** (see FIGS. **3** and **4b**), is retrieved from the network and

15

stored in local storage. At 536, the viewing program checks again to determine whether the user has requested that another document be displayed, and then proceeds to preload document 306, which is the other document linked to by document 302. In this particular embodiment, the order in which documents 305 and 306 are preloaded is dictated by the profile information.

When the user does request that a different document be displayed in FIG. 13b, the viewing program determines at 542 whether the requested document has been already loaded into local storage. If it has been loaded into local storage, the requested document is retrieved from the higher-speed local storage (544) and the contents of the requested document are analyzed to determine if the information in the preloaded version is out of date (545). If not, the preloaded version of the document is rendered in the window 406 (546). Otherwise, the document is retrieved over the network (548), and then rendered in the window 406 (546).

FIG. 13c is a continuation of the flow chart of FIG. 13b, where the user has requested in FIG. 13b that document 306 be displayed, and at 579 the document 306 and its profile information is retrieved, and then at 580 of FIG. 13c, document 306 is displayed. At 582, the viewing program checks to see if the user has requested that another document be displayed. If not, another document is preloaded into the memory unit or into local storage at 584. As can be seen in FIG. 11, document 306 contains only link 322, which is a link back to document 304. Thus, at 584, the document 306 may be preloaded into memory or local storage if it is still in memory from a preloading operation at 538 in FIG. 13b. If it is in memory, it is not necessary to preload it, so the document viewing program can preload from the network some other document that the user may be likely to request. Such a document could be identified in the profile information for document 306 (as described above), or such a document could be a document that was linked to a previously-viewed document, but wasn't fully preloaded into local memory. For example, if in FIG. 13a document 303 was not preloaded, this document could be preloaded at 584 because it may, at some point, be requested by the user. Alternatively, the document preloaded at 584 may be one of the bookmarked documents maintained by the document viewing program (e.g., at the client), or a document from some other popular site.

When the user requests a new document, the document viewing program checks at 586 to determine whether the requested document has been preloaded (586). If it has, the preloaded version is retrieved at 588 and analyzed to determine whether it is out of date at 589. If the preloaded document is not out of date, it is displayed at 590. Otherwise, the document is retrieved over the network at 592, and displayed at 590.

FIG. 13d is a partial flow chart in an alternate embodiment of the present invention that can be used to replace 504 in FIG. 13a. FIG. 13d illustrates that each document normally contains a number of objects, and in order to retrieve from the network the entire document, each of these embedded objects must be also be retrieved. In the embodiment of FIG. 13d, document 302 comprises a base document which is retrieved over the network at 550. The base document 302 contains references to the embedded objects within document 302. When the base document is retrieved from the network, it is analyzed at 552 to determine the additional embedded objects (if any) that must be retrieved to complete the document. If document 302 has three embedded objects, each is retrieved from the network as shown in FIG. 13b in succession (steps 554, 556, 558).

16

In an alternate embodiment, shown in FIG. 13e, a separate thread of execution is started for retrieving each of the embedded objects. This embodiment recognizes that it may be more efficient to download the embedded objects simultaneously, rather than one at a time as shown in FIG. 13d. The embodiment of FIG. 13d, however, has the advantage that it may be able to completely download at least one of the embedded objects before it is interrupted by a request to display another document. Depending on the implementation of the viewing program, a fully preloaded object may be more useful than a partially-preloaded object. Preloading documents simultaneously may increase the chance of having a large number of partially-preloaded objects, and fewer fully-preloaded objects. Thus, where partially-preloaded objects are less useful than fully-preloaded objects, the embodiment of FIG. 13d may be more efficient than FIG. 13e.

FIG. 16 is flow chart of the operation of a system in which document 301 and its associated profile information is retrieved at 601, and the document 301 is displayed at 602. The documents 302 and 303 are then preloaded in two separate threads of execution (steps 604 and 606) so that they are retrieved from the network substantially simultaneously. Unlike FIG. 15, in the embodiment of FIG. 16 the document viewing program does not preload document 304 while document 301 is displayed. The decision not to preload document 304 may be based on the profile information for document 301 retrieved over the network at 601 in FIG. 16. For example, the profile information could instruct the document retrieval program to not preload document 304, or the profile information could indicate that the document 304 has been (historically) so rarely selected by other users that the document retrieval program decides not to retrieve document 304.

A third thread of execution in FIG. 16 (610) monitors the user's actions (e.g., manipulation of the input device) to determine if the user has requested that a different document be displayed. When the user does request a document at 610, the system (e.g., the document viewing program) determines (612) whether the requested document has been at least partially preloaded. Where it has not, the requested document is retrieved over the network at 618 and displayed at 616.

However, if the requested document has been at least partially preloaded, the preloaded version is checked for staleness at 617. At 613 the document viewing program determines whether the requested document has been partially or fully preloaded. If the requested document has been fully preloaded, it is retrieved from local storage (614) and rendered on the display (616). If it has been only partially preloaded, the partially preloaded version is retrieved from local storage (620), and any portion not in local storage is retrieved from the network (622), and then rendered on the display (616).

In some situations, it may be useful to have the user exercise some direct control over the preloading process. For example, FIG. 17a shows a screen 704 having a window 706, which includes a title bar 714 and an area 717 in which to visually render the contents of a document. A cursor 724 corresponding to a pointing-type input device is also shown in the embodiment of FIG. 17a. The document shown in the window 706 includes hypertext links 718, 720, and 722, and the document also comprises graphical object 707, which includes area 709. The graphical object 707 also acts as a link to another document. Also shown in the window 706 are buttons 708, 710, and 712, which each correspond to one of

17

the hypertext links. Button **708** corresponds to link **718**, button **710** corresponds to link **720**, and button **712** corresponds to link **722**.

In FIG. **17b**, the cursor **724** has been moved to the button **712** so that the button **712** is selected. Upon selection, the document viewing program begins preloading the document linked by the "forecast" hyperlink **722**, which corresponds to the button **712**. The "forecast" document is not yet displayed within the window **706**, however, and the "News Items" document shown in FIGS. **17a** and **17b** continues to be displayed. While the document corresponding to link **722** is retrieved over the network, the progress of the preloading operation is displayed on the button **712**. At the point shown in FIG. **17b**, the document corresponding to the "forecast" link is 7% retrieved.

FIG. **18a** is a flow chart illustrating the operation of an embodiment of the present invention that is similar to that described in connection with FIGS. **17a** and **17b**. Initially at **802**, a document is displayed by the document viewing program. The document viewing program then alternatively monitors the user's input to determine whether the user has selected a document to preload (**804**) or a document to be displayed (**808**). A document is selected to be preloaded by the user when one of the buttons **708**, **710**, or **712** is selected, or when area **709** within the graphical object **707** is selected. Upon such a selection, the document corresponding to the selected button or to the selected graphical object is retrieved over the network and stored in local storage (**806**).

In the embodiment of FIG. **17a**, the user requests a document to be displayed by selecting a hypertext link or by selecting the graphical object **707**. When the user has requested a document to be displayed, the viewing program determines at **810** whether the requested document has already been retrieved into local memory or storage. If it has been preloaded, it is retrieved from local storage (**816**), and displayed in the window **706** (**814**).

In some embodiments, the document may also be checked at **818** to determine whether it is sufficiently new or current. If its download date, modification date, or other information indicates that the contents of the document are not sufficiently new, or are out of date, the requested document is again retrieved over the network at **812**.

FIG. **18b** illustrates an embodiment of the present invention in which a document is displayed at **830** by the viewing program, and then one or more threads of execution begin preloading the documents that are linked to by the displayed document (**832**). Another thread of execution monitors the user's input to determine whether the user has selected a link to preload (**834**) or whether the user has selected a link to display (**836**).

When the user selects a link to preload at **834**, such as by selecting one or more of the buttons **708**, **710**, or **712** in FIG. **17a**, the viewing program begins preloading the selected link, and does so at a higher priority at **840** than any of the other links that are being preloaded at **832**. In other words, when the user selects a link to be preloaded, the viewing program allocates more resources to preloading the selected document at **840** than to any other preloading operations it is carrying out on any other documents at **832**.

Where more than one link has been selected by the user, each could be preloaded at a priority higher than that of the documents being preloaded at **832**. In another embodiment, the document most recently selected for preloading could be given a priority higher than any other, so that the resources of the document viewing program are being applied to the preloading of the most recently selected-document.

18

Once the user selects a document to be displayed, the viewing program determines at **842** whether the document has been preloaded into local storage. If so, the preloaded version is retrieved from local storage (**848**), and displayed (**846**). Otherwise, the requested document is retrieved over the network (**844**) and displayed (**846**).

In the embodiments described in FIGS. **18a** and **18b**, the user selects the document that he or she wishes to preload, and in the embodiments of FIGS. **17a** and **17b**, this is done by selecting button that corresponds to the desired link. In other embodiments, selecting the link that the user wishes to preload can be done in a number of other ways. For example, selection of a link to preload could be carried out by simply passing the mouse or pointing device cursor over the desired link or over an object that corresponds to the link. Such an action could communicate to the document viewing program the link that the user wishes to preload. In another embodiment, the user could select the desired link with a right mouse click (or some other keyboard or pointing device action), or by directing the document viewing program to preload a given link by selecting an appropriate option from a menu that is displayed when the link is selected with the pointing device.

In FIG. **17a**, selection of the document linked to by the graphical object **707** is carried out by the selection of the space **709** within the object **707**, or by passing the cursor over the area **709** in FIG. **17a**. Selection of any other portion of the object **707** constitutes a request that the document corresponding to the document to the object **707** be displayed, rather than preloaded. FIG. **17b** shows one way in which the progress of the preloading can be communicated to the user. FIGS. **17c**, **17d**, **17e**, and **17f** show other embodiments in which the progress being made in the preloading operation is communicated to the user. In FIG. **17c**, the graphical object **707** from FIGS. **17a** and **17b** has been selected for preloading, and a progress gauge **717** has a shaded area **719** which is used to show what portion of the document has been preloaded. When the shaded area **719** fills the gauge **717** entirely, the document linked to by the graphical object **707** has been preloaded.

The button **710** in FIG. **17d** operates in a manner similar to the button **712** of FIG. **17b**, but it changes colors to indicate the progress of the preloading operation. For example, the button **710** could get progressively darker (or lighter) while the linked document is being preloaded. FIG. **17e** shows a text button **760** that is used as a hyperlink. Selection of the button for preloading (e.g., by passing the cursor over the button) causes the button to change color or shade (see FIG. **17f**) as the preloading proceeds. Any type of visual or audio progress indicator could be used to indicate progress of the preloading, and is useful to the user because he or she will know when a desired document has been preloaded. The user can continue to read or interact with the currently displayed document until the visual or audio indicator signifies that the document has been preloaded. Thus, the user can be assured that when the preloaded document is selected for display, it will be quickly displayed.

In some document retrieval systems, the user may incur a cost for each document or set of information that he or she retrieves from a database or over a network. In such a system, preloading documents before they are requested by the user could incur fees for documents that the user has never intended to see, use, or retrieve from the network. In other words, some documents may be retrieved in such a system simply because they are linked or otherwise related to one of the documents that the user did retrieve. This can undesirably increase costs for the user.

19

FIG. 19 is a flow chart illustrating how an embodiment of the present invention can effectively operate in an environment where the user incurs fees for each document or set of documents retrieved over the network. In this example, when a document is displayed at 902, the document viewing program proceeds to preload one or more documents that are linked to the document that the user is viewing. However, the viewing program does not preload the version of the linked documents that incur a fee. Instead, the viewing program preloads a free (or reduced cost) encrypted version of the linked document(s). This encrypted version is distributed free or at a lower charge because it is unreadable (or at least difficult to read) to anyone that attempts to view the encrypted version. However, the encrypted version can be easily converted into the normal, readable version of the content or the document by processing the encrypted version of the document with a password or a key.

When the user selects a document to be displayed at 906, the viewing program determines at 908 whether an encrypted version of the requested document has been preloaded. If so, the viewing program retrieves the password or key required to decrypt the encrypted version of the document, and at that time, the cost of retrieving the document is incurred at 910. The encrypted version of the document stored locally is decrypted at 912 and then displayed at 914.

By retrieving only the password or key over the network and then decrypting the locally-stored encrypted version of the requested document, the document will typically be displayed much more quickly than if the entire document would have to be retrieved from the network. Normally, the size of the key will be much smaller than the size of the document. Retrieving only the key, and processing the encrypted document at the client will therefore typically be much faster than retrieving the unencrypted version of the document over the network upon receiving a request for it.

The use of the procedures described herein may, in some environments, substantially increase the number of requests that are issued to network servers, and may also increase the amount of bandwidth required for a given network. This can be exacerbated where each document has embedded within it additional objects that must be separately requested from the server. Thus, it may be desirable to implement techniques to alleviate, eliminate, or avoid these effects. In one embodiment of the present invention, each time a request is issued to a network server, additional information is included within the request so that the database server (or any other network hardware or resources) is notified of the type of the request. This will allow requests to be prioritized so that server and/or other network resources are not allocated to tasks that may have less priority (e.g., a request to preload a document) than other tasks (e.g., a normal document request).

FIG. 20a is a flow chart of a system in which the document viewing program communicates to the database server (or to the network itself) a priority for each request. At 1102, the document viewing program issues a normal priority request to the database server for document A. The database server responds to this request, and at 1104, document A is retrieved over the network by the document viewing program. When it is received, it is displayed by the document viewing program at 1106.

The document viewing program then starts a thread that monitors the user input at 1108 to determine whether the user has requested a document for display. Another thread is also started, and this thread at 1120 issues a low priority request to the server for document B (one of the documents

20

it seeks to preload). The user at this point has not requested that document B be displayed, so the retrieval of document B is done based on the expectation that the user may wish to view document B at some point. For this reason, the request for document B is issued on a low-priority basis. (Document B may be a document that is linked to document A, that is identified in profile information, or that is otherwise related to document A.) When the server responds to the request, document B is downloaded over the network at 1122, and stored locally at 1124.

The low-priority request allows the network server to respond to other normal or high priority requests in advance of responding to the low-priority request for document B. This can be used to ensure that when the user actually requests a document from the server, the server will service that request before other low-priority requests by either that user or by other users. This information can also be used by the network hardware (e.g., network routers) itself to prioritize the routing of the requests or the routing of packets of data.

When a request that a document be displayed is made by the user at 1108, the document viewing program determines whether the requested document is in local storage at 1110. If it is, it is retrieved from local storage at 1116, and displayed at 1118. However, if the requested document is not stored in local storage, a normal-priority request is issued to the server at 1112. The request is a "normal" priority request because the user has actually requested a document, in contrast to the request made at 1120 of FIG. 20a. The document is then retrieved over the network at 1114, and then displayed at 1118.

FIG. 20b is another embodiment of the present invention that deals generally with the types of problems addressed in FIG. 20a. After document A is displayed at 1130, a thread that monitors whether the user has requested a document for display is started at 1132. Another thread is started at 1140 to determine whether the server on which document B is stored is busy. If it is, a wait state is entered at 1141 so that requests are not issued to the server over the network. This procedure thus preserves network and/or server resources. After a period of time, the server is then checked again. When the server is not busy, document B is retrieved over the network at 1142, and stored on the client computer at 1144.

When the user requests a document for display, the document viewing program determines whether the requested document has already been preloaded. If necessary, the requested document is retrieved over the network at 1136; otherwise, it is retrieved from local storage at 1140. After it is retrieved, it is displayed at 1138.

FIG. 20c is a flow chart of an embodiment of the present invention in which an anticipated document, document B, is stored in a file along with the objects that are embedded within the document B, are referenced by the document B, or are linked to by document B. By downloading such a file, the number of requests that must be issued to the server can be reduced. And if data compression is used to reduce the size of the file at the server and decompress the file at the client, the number of bits that must be downloaded to the client computer can be reduced.

Once document A is displayed at 1148 in FIG. 20c, the document viewing program monitors the user at 1150 for a request to display a new document. At the same time (i.e., in another thread of execution), an object that contains document B and the objects embedded within it is retrieved over the network at 1160. This object may also include one or more documents that are linked to by document B. When the

21

object is downloaded, it is parsed at the client computer at **1162** so as to extract document B and the embedded objects, which are then stored at **1164** on the client computer. When the user requests a document for display, the document viewing program determines at **1152** whether the requested document has already been preloaded. If necessary, requested document is retrieved over the network (**1154**), but if possible, it is retrieved from local storage (**1156**). After it is retrieved, it is displayed at **1158**.

The present invention is suitable for implementation as an ActiveX or Java control, which could be downloaded as part of a web page into a browser or an operating system for execution on a client computer. In such an embodiment, there may be security restrictions placed on the downloaded control. Appendix A is an outline of a Java program or pseudocode for applet written in Java that can be inserted into a web page, and appears on the web page as a button that it is associated with an HTML link. When the button is selected, the document corresponding to the associated link is preloaded onto the client, and the base HTML document and at least some of the embedded objects are stored on the client's local file storage system. The client's file system is typically much faster than the network.

In some Java environments, the client's local file system is not accessible because of security restrictions if the applet is downloaded from a remote host. These security restrictions can be avoided by using an insecure environment, or by using a code signing technique that allows the user to verify the author of the applet. Once the code is identified as being written by a trusted author, the security restrictions can be safely eased or eliminated.

In another embodiment, a secure means of accessing the client's file system can be used to securely and safely store data on the client's file system. In such a system, the applet may only be allowed to write files to certain directories. The applet may also be limited to reading only files that it had created. One such secure file system for Java has been referred to as "protected domains," and can be useful in implementing some embodiments of the present invention.

Appendix B is another listing of Java code/pseudocode in an implementation of the computer program or applet that does not use the local file system for storing preloaded documents. Instead, the Java program in Appendix B stores preloaded documents in memory, and implements a web server on the client to serve the documents back to the document viewing program running on the client. Thus, when the user selects a link, the link is redirected to the server running on the client, and that local server responds with the preloaded document if it is available. The document viewing program would, in effect, be retrieving preloaded documents from local memory, thereby making access to preloaded documents quite fast. Such an implementation may also avoid the security restrictions placed on accesses to the local file system in some embodiments.

FIG. 21 is a flow chart that illustrates the high-level operation of the pseudocode of Appendix B. At **1202**, a document is displayed, and a local document or object server is then set up at **1204**. Two threads of execution are started, one to retrieve an anticipated document (document B) from the network and store it as a document capable of being served by the local server (steps **1220** and **1221**), and another to monitor the user's actions to determine when the user has selected a document for display (**1204**).

When the user selects a document for display at **1206**, the request is routed to the local server (**1210**) if the requested document had been stored locally. Otherwise, the request for the document requested by the user is issued to the server

22

(usually a remote server) on the network (**1216**). When the document is retrieved (or as it is retrieved), the document is displayed at **1214**.

Because a web page control may have to be downloaded with each page, it may be desirable to implement techniques to speed the amount of time that a user has to wait for a document to be retrieved from the network. One such procedure is to embed a small applet into the web page that is downloaded by the user, where the small applet then retrieves a larger program that carries out the remaining steps. Such a procedure will allow the user to begin interacting with the web page after the small applet is downloaded, and will not require that the user wait for a larger program to be downloaded before interacting with the web page. Once the small applet is downloaded, the larger applet is downloaded in the background while the user is viewing or interacting with the web page.

FIG. 22 is a flow chart that illustrates the use of a small embedded program that retrieves a larger embedded program, and causes the latter to execute. At **1302**, a document or object with an embedded initializing program stored within it is downloaded from the network. The document (or object) is displayed on the screen at **1304**, and in a separate thread of execution, the initializing program is started at **1316**. The initializing program retrieves supplemental program code over the network, and execution of this code is started at **1320**. As indicated in FIG. 22, this new code retrieves anticipated documents over the network, also at **1320**.

Because the initializing program is small, it takes relatively little time to download, and the document viewing program is able to promptly start the execution of the initializing program. This may allow the display of the document at **1304** to take place more quickly. The effect is a more responsive program that does not cause the user significant delay while an applet implementing the present invention is being downloaded.

Many embodiments of the present invention have been described as storing preloaded documents into local storage at the client computer. However, the present invention need not be limited to contexts in which information is stored at the client computer or in local storage at the client computer. The present invention is useful in any environment where it is possible to store preloaded information in an area where access to the preloaded information is faster than that of the original location for the information. For example, FIG. 23 shows a network where computer **1401** is preloading (**1412**) a document **1440** on server **1402** while viewing another document on the network. In the embodiments described previously, the document **1440** is retrieved over the network and stored in local storage at the computer **1401**. However, other embodiments of the present invention can be performed by storing the preloaded document elsewhere, but still in a location that can be accessed quickly.

An example is shown in FIG. 23 where the computer **1401** retrieves document **1440** from the server **1402** as part of a preloading procedure. At the direction (**1412**) of computer **1401**, the preloaded document is retrieved (**1410**) and stored in the computer **1403**, which is accessible by the computer **1401** over the LAN. Information on computer **1403** can be accessed by the computer **1401** quickly because these two computers are connected over a relatively fast (local) network. This is unlike the connection between the computers **1401** and **1402**, which are connected over the lower speed WAN.

When the preloaded document **1440** is stored on the computer **1403**, it can be more quickly retrieved from

computer 1403 than from computer 1410. Thus, significant enhancements to the responsiveness of the document viewing program can be made in the present invention, even if the preloaded documents or objects are not stored directly in local storage, but instead, are stored elsewhere where they can be retrieved quickly.

Some embodiments of the present invention have been described in the context of accessing the database and identifying search documents through a search term query. The present invention can be applicable in other research-related contexts where search documents are identified using another type of entry path. For example, a time-line can be used for locating information or documents that are associated with a given time or time-frame. Another information access method uses a topic tree that permits a user to choose from successively narrowing topics until the desired topic is located. It is possible for the present invention to be applicable even in other non-research contexts where similar preloading techniques may permit efficient navigation of information and/or short response times. The present invention can also be used in combination with caching systems where previously-displayed documents or views are stored for repeated use.

The present invention has been primarily described in the context of a general purpose computer implementation. As one skilled in the art will recognize, however, it is possible to construct a specialized machine that can carry out the present invention.

The additional references listed below are hereby fully incorporated by reference to the extent that they enable, provide support for, provide a background for, or teach methodology, techniques, and/or procedures employed herein.

- Reference 1: Yellin, The Java Application Programming Interface: Volumes 1 & 2 (Addison Wesley 1996)
- Reference 2: Campione, The Java Tutorial (Addison Wesley 1996)
- Reference 3: Chappell, Understanding ActiveX and OLE (Microsoft Press 1996)
- Reference 4: Denning, OLE Controls Inside Out (Microsoft 1995)
- Reference 5: Brockschmidt, Inside OLE (2d ed. Microsoft 1995)
- Reference 6: Graham, HTML Sourcebook (2d ed. John Wiley & Sons 1996)
- Reference 7: Tanenbaum, Computer Networks (2d ed. Prentice Hall 1989)
- Reference 8: Jamsa, Internet Programming (Jamsa Press 1995)
- Reference 9: Corner, Internetworking with TCP/IP, Volumes 1, 2, & 3 (2d ed. Prentice Hall 1995)
- Reference 10: Petzold, Programming Windows 95 (Microsoft 1996)
- Reference 11: Prosise, Programming Windows 95 with MFC (Microsoft Press 1996)
- Reference 12: Chapman, Building Internet Applications with Delphi 2 (Que 1996)
- Reference 13: Schneier, Applied Cryptography (2nd edition John Wiley & Sons 1995)
- Reference 14: Chan, The Java Class Libraries (Addison Wesley 1997)
- Reference 15: Siegel, CORBA Fundamentals and Programming (John Wiley & Sons 1996)
- Reference 16: Lemay, Official Marimba Guide to Castanet (Sams.Net 1997)
- Reference 17: Adkins, Internet Security Professional Reference (New Riders 1996)

Reference 18: Microsoft Corporation, Windows NT Server Resource Kit (Microsoft Press 1996)

Reference 19: Russel, Running Windows NT Server (Microsoft Press 1997)

Reference 20: Lemay et al., Java in 21 Days (Sams.Net 1996)

Reference 21: Danesh, JavaScript in a Week (Sams.Net 1996)

Reference 22: Kovel et al., The Lotus Notes Idea Book (Addison Wesley 1996)

Reference 23: Sun Microsystems, Inc., The JavaBeans 1.0 API Specification (Sun Microsystems 1996) (available at <http://java.sun.com/beans>)

Reference 24: Sun Microsystems, Inc., The Java 1.1 API Specification (Sun Microsystems 1997) (available at <http://java.sun.com/>)

Reference 25: Bell, "Make Java fast: Optimize!," JavaWorld April 1997 (JavaWorld 1997) (available at <http://www.javaworld.com/>)

Reference 26: Vanhelsuwe, "How to make Java applets start faster," JavaWorld December 1996 (JavaWorld 1996) (available at <http://www.javaworld.com/>)

Although the present invention has been shown and described with respect to preferred embodiments, various changes and modifications, even if not shown or specifically described herein, are deemed to lie within the spirit and scope of the invention and the following claims. Importantly, it should be understood that any specific features or aspects of the embodiments described or illustrated herein are not intended to limit the scope and interpretation of the claims in a manner not explicitly required by the claims.

What is claimed is:

1. A system comprising:

- (a) a large number of documents that are stored on a network;
- (b) a search engine that is capable of finding documents that are relevant to a search query from among the large number of documents that are stored on the network; and
- (c) a computer that is operated by a user and that is in communication with the search engine and the large number of documents, wherein the computer is programmed to carry out the operations of:
 - monitoring input from the user,
 - receiving a text-based search query as input from the user,
 - submitting the text-based search query to the search engine,
 - allowing the search engine to identify a plurality of search documents that may be of interest to the user based on the text-based search query,
 - displaying a first document for the user on a display screen associated with the computer,
 - identifying a second document as a document that the user may wish to display on the display screen, wherein the second document is one of the plurality of search documents that may be of interest to the user based on the text-based search query,
 - retrieving over the network information from the second document, wherein the information from the second document is retrieved from a remote location, and wherein the information from the second document is retrieved while information from the first document is displayed on the display screen, and before the user inputs a request to display the second document,
 - storing the information from the second document in local storage associated with the computer,
 - continuing to monitor input from the user, and

25

displaying the second document on the display screen for the user when the user inputs a request to display the second document, wherein the second document is displayed on the display screen by retrieving at least a portion of the second document from local storage so that the second document is displayed quickly. 5

2. The system of claim 1, wherein the operation of retrieving information over the network from the second document includes the operation of:

retrieving the entire second document over the network before the user inputs a request to display the second document. 10

3. The system of claim 1, wherein the operation of identifying a second document includes the operation of:

identifying the second document as a document the user may wish to display on the display screen because the second document is one of the search documents identified by the search engine. 15

4. The system of claim 1, wherein the operation of identifying a second document includes the operation of: identifying the second document as a document the user may wish to display on the display screen based on the content of the displayed first document. 20

5. The system of claim 4, wherein the first document contains a link to the second document. 25

6. The system of claim 1, wherein the operation of displaying a first document for the user on the display screen includes the operation of:

displaying a portion of one of the plurality of search documents by displaying an excerpt that includes a term from the text-based search query. 30

7. The system of claim 1, wherein the first document is one of the search documents.

8. The system of claim 1, wherein the search engine ranks at least some of the search documents it identifies as being of interest to the user based on the text-based search query, wherein the search documents are ranked based on the degree to which each of the search documents is determined to be relevant to the text-based search query. 35

9. The system of claim 8, wherein the operation of identifying a second document includes the operation of: identifying the second document as a document the user may wish to display on the display screen because the second document is one of the search documents identified by the search engine and because it has a higher degree of relevance to the text-based search query than another search document. 40

10. A system for transferring information over a network, wherein the system is configured to carry out the operations of: 50

monitoring input from a user operating a client computer that is in communication with the network;

receiving a text-based search query as input from the user operating the client computer;

identifying a plurality of search documents that are relevant to the text-based search query, wherein the search documents are identified from among a large number of documents that are stored on the network;

ranking at least some of the search documents based on the degree to which each search document is determined to be relevant to the text-based search query; 60

displaying a document for the user on a display screen associated with the client computer;

determining that a first search document is an appropriate document to anticipate that the user will request be displayed on the display screen associated with the client computer, wherein the first search document is 65

26

one of the search documents identified as relevant to the text-based search query, and wherein the first search document is determined to be an appropriate document to anticipate that the user will request be displayed at least because the first search document is a search document;

transferring information from the first search document over the network to the client computer, wherein the information from the first search document is transferred over the network from a remote location relative to the client computer, and wherein the information from the first search document is transferred over the network in the background while a different document is being displayed on the display screen associated with the client computer, and before the user inputs a request to display the first search document on the display screen associated with the client computer;

storing the information from the first search document in local storage associated with the client computer;

continuing to monitor input from the user operating the client computer; and

displaying the first search document for the user on the display screen associated with the client computer when the user inputs a request to display the first search document, wherein the first search document is displayed on the display screen by retrieving at least a portion of the first search document from local storage associated with the client computer so that the first search document is displayed more quickly for the user than if the entire first search document was transferred over the network following the user's request to display the first search document.

11. The system of claim 10, wherein the operation of determining that a first search document is an appropriate document to anticipate that the user will request be displayed on the display screen includes the operation of:

determining that the first search document is an appropriate document to anticipate that the user will request be displayed on the display screen because the first search document has a higher degree of relevance to the text-based search query than a second search document.

12. A computer operated by a user for retrieving information over a network, wherein the computer is programmed to carry out the operations of:

monitoring input from the user;

receiving a search query as input from the user;

submitting the search query to a search engine, wherein the search engine uses the search query to find a plurality of search documents that may be of interest to the user based on the search query;

displaying a first document on a display screen associated with the computer;

identifying a second document, a third document, and a fourth document as documents that the user may wish to display on the display screen, wherein the second document, the third document, and the fourth document are each search documents found by the search engine, and wherein the second document, the third document, and the fourth document are documents that the user may wish to display on the display screen because they are search documents found by the search engine;

retrieving information from the second document over the network, wherein the information from the second document is retrieved from a remote location, and wherein the information from the second document is retrieved before the user inputs a request to display the

27

second document, and wherein the information from the second document is retrieved in anticipation of the user requesting that the second document be displayed; storing the information from the second document in local storage associated with the computer; retrieving information from the third document over the network before the user inputs a request to display the third document, wherein the information from the third document is retrieved in anticipation of the user requesting that the third document be displayed; 10 storing the information from the third document in local storage associated with the computer; retrieving information from the fourth document over the network before the user inputs a request to display the fourth document, wherein the information from the 15 fourth document is retrieved in anticipation of the user requesting that the fourth document be displayed; storing the information from the fourth document in local storage associated with the computer; continuing to monitor input from the user; and 20 displaying the second document for the user when the user inputs a request to display the second document, wherein the second document is displayed on the display screen associated with the computer by retrieving at least a portion of the second document from local storage associated with the computer so that the second document is displayed more quickly than if the entire second document was retrieved over the network following the user's request to display the second document. 30

13. The computer of claim 12, wherein the operation of retrieving information from the second document over the network includes the operation of:

retrieving the entire second document over the network before the user inputs a request to display the second document. 35

14. The computer of claim 13, further programmed to carry out the operations of:

monitoring input from the user while the second document is being displayed on the display screen associated with the computer; 40

displaying the third document for the user when the user inputs a request to display the third document while viewing the second document, wherein the third document is displayed on the display screen associated with the computer by retrieving the third document from local storage so that the third document is displayed more quickly than if the entire third document was retrieved over the network following the user's request to display the third document; 50

monitoring input from the user while the third document is being displayed on the display screen associated with the computer; and

displaying the fourth document for the user when the user inputs a request to display the fourth document while viewing the third document, wherein the fourth document is displayed on the display screen associated with the computer by retrieving the fourth document from local storage so that the fourth document is displayed more quickly than if the entire fourth document was retrieved over the network following the user's request to display the fourth document. 55

15. The computer of claim 12, wherein the operation of submitting the search query to a search engine includes the operation of: 65

submitting the search query to a search engine that ranks at least some of the plurality of search documents based

28

on the degree to which each of the search documents is determined to be relevant to the search query.

16. The computer of claim 15, wherein the operation of retrieving information over the network from the second document includes the operation of: 5

starting to retrieve information over the network from the second document before starting to retrieve information over the network from the third document, wherein the degree to which the second document is determined to be relevant to the search query is higher than that of the third document. 10

17. A computer readable storage medium having computer program instructions stored on it, where the computer program instructions are executable on a client computer operated by a user, and where the computer program instructions configure the client computer to be capable of carrying out the operations of: 15

monitoring input from the user operating the client computer by monitoring input to an input device associated with the client computer; 20

receiving a search query as input to the client computer, where the search query is entered by the user using the input device;

submitting the search query to a search query processor, where the search query processor uses the search query to find a plurality of search documents that may be of interest to the user based on the search query, where the search query processor finds the plurality of search documents from among a large number of documents that may be available to the client computer over a connection providing access to the large number of documents; 25

identifying a first document as a document that the user may wish to display on a display screen associated with the client computer, where the first document is one of the search documents found by the search query processor, and where the first document is a document that the user may wish to display on the display screen at least because it is a search document found by the search query processor; 30

retrieving information that can be used to display the first document on the display screen, where the information that can be used to display the first document is retrieved over the connection by the client computer from a remote location relative to the client computer, where the information that can be used to display the first document is retrieved after the operation of submitting the search query to the search query processor and before the user inputs a request to display the first document on the display screen, and where the information that can be used to display the first document is retrieved in anticipation of quickly displaying the first document upon the user requesting that the first document be displayed on the display screen; 35

storing the information that can be used to display the first document in local storage associated with the client computer; 40

continuing to monitor input from the user operating the client computer by monitoring input to the input device;

detecting as input from the user a request to display the first document on the display screen associated with the client computer; and 45

upon detecting as input from the user a request to display the first document on the display screen associated with the client computer, displaying the first document on the display screen by accessing the information that can

29

be used to display the first document stored in local storage associated with the client computer so that the first document is quickly displayed for the user on the display screen upon detecting as input from the user a request to display the first document on the display screen associated with the client computer.

18. The computer readable storage medium of claim 17, where the computer program instructions further configure the client computer to be capable of carrying out the operations of:

identifying a second document as a document that the user may wish to display on the display screen associated with the client computer, where the second document is one of the search documents found by the search query processor, and where the second document is a document that the user may wish to display on the display screen at least because it is a search document found by the search query processor;

retrieving information that can be used to display the second document on the display screen, where the information that can be used to display the second document is retrieved over the connection by the client computer after the operation of submitting the search query to the search query processor and before the user inputs a request to display the second document on the display screen, and where the information that can be used to display the second document is retrieved in anticipation of quickly displaying the second document upon the user requesting that the second document be displayed on the display screen;

storing the information that can be used to display the second document in local storage associated with the client computer;

monitoring input from the user while the first document is being displayed on the display screen;

detecting as input from the user a request to display the second document while the first document is being displayed on the display screen; and

upon detecting as input from the user a request to display the second document while the first document is being displayed on the display screen, displaying the second document on the display screen by accessing the information that can be used to display the second document stored in local storage associated with the client computer so that the second document is quickly displayed for the user on the display screen upon detecting as input from the user the request to display the second document.

19. The computer readable storage medium of claim 17, where the operation of identifying a first document includes the operation of identifying a search document that has been assessed as having the highest relevance to the search query.

20. A computer readable storage medium having computer program instructions stored on it, where the computer program instructions are executable on a computer operated by a user and connected to a network, and where the computer program instructions configure the computer to be capable of carrying out the operations of:

monitoring input from the user operating the computer; receiving a search query as input to the computer through an input device associated with the computer, where the search query is entered by the user using the input device associated with the computer;

submitting the search query to a search engine, where the search engine uses the search query to find a plurality of search documents that may be of interest to the user based on the search query, where the search engine

30

finds the plurality of search documents from among a large number of documents that may be accessible to the computer over the network;

identifying a first anticipated document, a second anticipated document, and a third anticipated document as documents that the user may wish to display on a display screen associated with the computer, where the first anticipated document, the second anticipated document, and the third anticipated document are each search documents found as a result of the operation of submitting the search query to the search engine, and where the first anticipated document has been assessed as having a higher degree of relevance to the search query than another search document;

starting to retrieve information from the first anticipated document over the network from a remote location relative to the computer, where the operation of starting to retrieve information from the first anticipated document occurs after the operation of submitting the search query and before the user inputs a request to display the first anticipated document;

starting to retrieve information from the second anticipated document, where the operation of starting to retrieve information from the second anticipated document occurs after the operation of submitting the search query and before the user inputs a request to display the second anticipated document, and where the operation of starting to retrieve information from the second anticipated document occurs after the operation of starting to retrieve information from the first anticipated document;

continuing to retrieve information from the first anticipated document over the network, where the operation of continuing to retrieve information from the first anticipated document over the network occurs after the operation of submitting the search query and before the user inputs a request to display the first anticipated document, and where the information from the first anticipated document is retrieved over the network in the background by the computer in anticipation of responding quickly upon the user requesting that the first anticipated document be displayed;

storing the information from the first anticipated document in local storage associated with the computer;

continuing to retrieve information from the second anticipated document over the network, where the operation of continuing to retrieve information from the second anticipated document over the network occurs after the operation of submitting the search query and before the user inputs a request to display the second anticipated document, and where the information from the second anticipated document is retrieved over the network in the background by the computer in anticipation of responding quickly upon the user requesting that the second anticipated document be displayed;

storing the information from the second anticipated document in local storage associated with the computer;

retrieving information from the third anticipated document, where the operation of retrieving information from the third anticipated document occurs after the operation of submitting the search query and before the user inputs a request to display the third anticipated document, and where the information from the third anticipated document is retrieved over the network by the computer in anticipation of responding quickly upon the user requesting that the third anticipated document be displayed;

31

storing the information from the third anticipated document in local storage associated with the computer;
 continuing to monitor input from the user operating the computer;
 detecting as input from the user a request to display the first anticipated document;
 upon detecting as input from the user a request to display the first anticipated document, displaying the first anticipated document for the user on the display screen, where the first anticipated document is displayed on the display screen by retrieving the information from the first anticipated document from local storage associated with the computer so that the first anticipated document is quickly displayed for the user upon detecting as input from the user the request to display the first anticipated document;
 monitoring input from the user while the first anticipated document is being displayed on the display screen;
 detecting as input from the user a request to display the second anticipated document, where the detection occurs while the first anticipated document is being displayed;
 upon detecting as input from the user a request to display the second anticipated document, displaying the second anticipated document on the display screen, where the second anticipated document is displayed on the display screen by retrieving the information from the second anticipated document from local storage associated with the computer so that the second anticipated document is quickly displayed for the user upon detecting as input from the user the request to display the second anticipated document while the first anticipated document is being displayed;
 monitoring input from the user while the second anticipated document is being displayed on the display screen;
 detecting as input from the user a request to display the third anticipated document, where the detection occurs while the second anticipated document is being displayed; and
 upon detecting as input from the user a request to display the third anticipated document, displaying the third anticipated document on the display screen, where the third anticipated document is displayed on the display screen by retrieving the information from the third anticipated document from local storage associated with the computer so that the third anticipated document is quickly displayed for the user upon detecting as input from the user the request to display the third anticipated document.

21. A system for transferring information over a network, where the system is configured to be capable of carrying out the operations of:

- monitoring input from a user operating a computer that is connected to the network;
- receiving a search query from the user as input to the computer;
- processing the search query to identify a plurality of search documents that are relevant to the search query from among a large number of documents that may be available to the computer over the network;
- determining that a first anticipated document, a second anticipated document, and a third anticipated document are documents that the user may wish to display on a display screen associated with the computer, where the first anticipated document, the second anticipated document, and the third anticipated document are docu-

32

- ments that the user may wish to display on the display screen associated with the computer at least because they are each search documents identified as a result of the operation of processing the search query, and where the first anticipated document has a higher degree of relevance to the search query than another search document;
- starting to transfer over the network to the computer information that can be used to display the first anticipated document, where the information that can be used to display the first anticipated document is transferred from a remote location relative to the computer, where the operation of starting to transfer information that can be used to display the first anticipated document occurs after the operation of processing the search query and before the user inputs a request to display the first anticipated document on the display screen;
- starting to transfer over the network to the computer information that can be used to display the second anticipated document, where the operation of starting to transfer information that can be used to display the second anticipated document occurs after the operation of processing the search query and before the user inputs a request to display the second anticipated document on the display screen, and where the operation of starting to transfer information that can be used to display the second anticipated document occurs after the operation of starting to transfer information that can be used to display the first anticipated document;
- continuing to transfer over the network information that can be used to display the first anticipated document, where the operation of continuing to transfer over the network information that can be used to display the first anticipated document occurs after the operation of processing the search query and before the user inputs a request to display the first anticipated document on the display screen;
- storing the information that can be used to display the first anticipated document in local storage associated with the computer;
- continuing to transfer over the network information that can be used to display the second anticipated document, where the information that can be used to display the second anticipated document is transferred over the network while the computer remains responsive to user input, and where the operation of continuing to transfer over the network information that can be used to display the second anticipated document occurs after the operation of processing the search query and before the user inputs a request to display the second anticipated document on the display screen;
- storing the information that can be used to display the second anticipated document in local storage associated with the computer;
- transferring over the network to the computer information that can be used to display the third anticipated document, where the operation of transferring information that can be used to display the third anticipated document occurs after the operation of processing the search query and before the user inputs a request to display the third anticipated document on the display screen;
- storing the information that can be used to display the third anticipated document in local storage associated with the computer;
- continuing to monitor input from the user operating the computer;

33

detecting as input from the user a request to display the first anticipated document on the display screen associated with the computer;

upon detecting as input from the user a request to display the first anticipated document on the display screen associated with the computer, displaying the first anticipated document on the display screen associated with the computer, where the first anticipated document is displayed on the display screen by accessing in local storage associated with the computer the information that can be used to display the first anticipated document so that the first anticipated document can be displayed quickly for the user upon detecting as input from the user the request to display the first anticipated document;

monitoring input from the user while information is being displayed on the display screen associated with the computer;

detecting as input from the user a request to display the second anticipated document on the display screen associated with the computer;

upon detecting as input from the user a request to display the second anticipated document on the display screen associated with the computer, displaying the second anticipated document on the display screen associated with the computer, where the second anticipated document is displayed on the display screen by accessing in local storage associated with the computer the information that can be used to display the second anticipated document so that the second anticipated document can be displayed quickly for the user upon detecting as input from the user the request to display the second anticipated document;

continuing to monitor input from the user while information is being displayed on the display screen associated with the computer;

detecting as input from the user a request to display the third anticipated document on the display screen associated with the computer; and

upon detecting as input from the user a request to display the third anticipated document on the display screen associated with the computer, displaying the third anticipated document on the display screen associated with the computer, where the third anticipated document is displayed on the display screen by accessing in local storage associated with the computer the information that can be used to display the third anticipated document so that the third anticipated document can be displayed quickly for the user upon detecting as input from the user the request to display the third anticipated document.

22. The system of claim 21, where the operation of starting to transfer over the network to the computer information that can be used to display the first anticipated document is carried out by a computer program that is downloaded over the network to the computer, checked for authenticity using a code signing procedure, and stored on the computer using a persistent storage mechanism.

23. The system of claim 21, where the operation of starting to transfer over the network to the computer information that can be used to display the first anticipated document is carried out by a computer program that is downloaded over the network to the computer and executed on the computer without the use of a persistent storage mechanism.

34

24. The system of claim 23, where the computer program is hard-coded to determine that the first anticipated document is a document that the user may wish to display on the display screen.

25. The system of claim 21, where the operation of determining that a first anticipated document, a second anticipated document, and a third anticipated document are documents that the user may wish to display on a display screen associated with the computer includes the operation of:

determining that a first anticipated document, a second anticipated document, and a third anticipated document are documents that the user may wish to display on the display screen based at least in part on profile information transferred over the network.

26. The system of claim 21, where the system is further configured to carry out the operations of:

upon completing the operation of transferring over the network information that can be used to display the first anticipated document, providing an indication to the user that the operation is complete.

27. The system of claim 26, where the indication to the user includes changing the appearance of a display element shown on the display screen associated with the computer.

28. In a system comprising a client computer and a search query processor connected by a network, a method comprising the acts of:

monitoring input from a user operating the client computer;

enabling the user to enter a search query as input to the client computer;

enabling the processing of the search query to identify a plurality of search documents that are relevant to the search query from among a large number of documents stored on the network and that may be accessible to the client computer over the network;

determining that a first anticipated document, a second anticipated document, and a third anticipated document are documents that the user may wish to display on a display screen associated with the client computer, where the first anticipated document, the second anticipated document, and the third anticipated document are each determined to be documents that the user may wish to display on the display screen based on the content of a document being viewed by the user on the display screen;

starting to transfer over the network to the client computer information that can be used to display the first anticipated document, where the information that can be used to display the first anticipated document is transferred from a remote location relative to the client computer, and where the act of starting to transfer information that can be used to display the first anticipated document occurs after the processing of the search query and before the user inputs a request to display the first anticipated document on the display screen;

starting to transfer over the network to the client computer information that can be used to display the second anticipated document, where the act of starting to transfer information that can be used to display the second anticipated document occurs after the processing of the search query and before the user inputs a request to display the second anticipated document on the display screen, and where the act of starting to transfer information that can be used to display the second anticipated document occurs after the act of

35

starting to transfer information that can be used to display the first anticipated document;

continuing to transfer over the network information that can be used to display the first anticipated document, where the act of continuing to transfer over the network information that can be used to display the first anticipated document occurs after the processing of the search query and before the user inputs a request to display the first anticipated document on the display screen, and where the information that can be used to display the first anticipated document is transferred over the network to the client computer in anticipation of quickly responding upon the user requesting that the first anticipated document be displayed on the display screen;

storing the information that can be used to display the first anticipated document in local storage associated with the client computer;

continuing to transfer over the network information that can be used to display the second anticipated document, where the act of continuing to transfer over the network information that can be used to display the second anticipated document occurs after the processing of the search query and before the user inputs a request to display the second anticipated document on the display screen, and where the information that can be used to display the second anticipated document is transferred over the network to the client computer in anticipation of quickly responding upon the user requesting that the second anticipated document be displayed on the display screen;

storing the information that can be used to display the second anticipated document in local storage associated with the client computer;

transferring over the network to the client computer information that can be used to display the third anticipated document, where the act of transferring information that can be used to display the third anticipated document occurs after the processing of the search query and before the user inputs a request to display the third anticipated document on the display screen, and where the information that can be used to display the third anticipated document is transferred over the network to the client computer while the client computer is responsive to user input and in anticipation of quickly responding upon the user requesting that the third anticipated document be displayed on the display screen;

storing the information that can be used to display the third anticipated document in local storage associated with the client computer;

continuing to monitor input from the user operating the client computer;

detecting as input from the user a request to display the first anticipated document on the display screen associated with the client computer;

36

upon detecting as input from the user a request to display the first anticipated document on the display screen associated with the client computer, displaying the first anticipated document on the display screen associated with the client computer, where the first anticipated document is displayed on the display screen by retrieving from local storage associated with the client computer the information that can be used to display the first anticipated document so that the first anticipated document is displayed quickly for the user upon detecting as input from the user the request to display the first anticipated document;

monitoring input from the user while a document is being displayed on the display screen associated with the client computer;

detecting as input from the user a request to display the second anticipated document on the display screen associated with the client computer;

upon detecting as input from the user a request to display the second anticipated document on the display screen associated with the client computer, displaying the second anticipated document on the display screen associated with the client computer, where the second anticipated document is displayed on the display screen by retrieving from local storage associated with the client computer the information that can be used to display the second anticipated document so that the second anticipated document is displayed quickly for the user upon detecting as input from the user the request to display the second anticipated document;

monitoring input from the user while information is being displayed on the display screen associated with the client computer;

detecting as input from the user a request to display the third anticipated document on the display screen associated with the client computer; and

upon detecting as input from the user a request to display the third anticipated document on the display screen associated with the client computer, displaying the third anticipated document on the display screen associated with the client computer, where the third anticipated document is displayed on the display screen by retrieving from local storage associated with the client computer the information that can be used to display the third anticipated document so that the third anticipated document is displayed quickly for the user upon detecting as input from the user the request to display the third anticipated document.

29. A system comprising a client computer and a search query processor connected by a network, where the system is configured to carry out the method of claim 28.

* * * * *