US005819292A

United States Patent [19]

Hitz et al.

[11] Patent Number: 5,819,292

[45] Date of Patent: Oct. 6, 1998

[54] **METHOD FOR MAINTAINING CONSISTENT STATES OF A FILE SYSTEM AND FOR CREATING USER-ACCESSIBLE READ-ONLY COPIES OF A FILE SYSTEM**

[75] Inventors: **David Hitz**, Sunnyvale; **Michael Malcolm**, Los Altos; **James Lau**, Cupertino; **Byron Rakitzis**, Mountain View, all of Calif.

[73] Assignee: **Network Appliance, Inc.**, Santa Clara, Calif.

[21] Appl. No.: **454,921**

[22] Filed: **May 31, 1995**

Related U.S. Application Data

[63] Continuation of Ser. No. 71,643, Jun. 3, 1993, abandoned.

[51] **Int. Cl.**[6] **G06F 17/30**

[52] **U.S. Cl.** **707/203**; 707/205

[58] **Field of Search** 395/621, 619

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,075,691	2/1978	Davis et al.	395/884
4,156,907	5/1979	Rawlings et al.	395/200.01
4,333,144	6/1982	Whiteside et al.	395/672
4,377,843	3/1983	Garringer et al.	395/890
4,399,503	8/1983	Hawley	395/440
4,456,957	6/1984	Schieltz	395/824
4,459,664	7/1984	Pottier et al.	395/675
4,488,231	12/1984	Yu et al.	395/868
4,527,232	7/1985	Bechtolsheim	395/416
4,550,368	10/1985	Bechtolsheim	395/416
4,685,125	8/1987	Zave	379/96
4,710,868	12/1987	Cocke et al.	395/311
4,719,569	1/1988	Ludemann et al.	395/729
4,742,447	5/1988	Duvall et al.	364/200
4,761,785	8/1988	Clark et al.	371/51
4,766,534	8/1988	DeBenedictis	395/200.14
4,780,821	10/1988	Crossley	395/670

(List continued on next page.)

OTHER PUBLICATIONS

Bach, "The Design of the UNIX Operating System", Prentice–Hall, 1986, pp. 38–90 and 325–329, Jan. 1986.

Beach, Richard J., et al., "The Message is The Medium: Multiprocess Structuring of an Interactive Paint Program", Computer Graphics, vol. 16, No. 3, Jul. 1982, pp. 277–287.

Britton, Dianne E. et al., "A Interprocess Communication Facility for Distributed Applications", RCA Laboratories, David Sarnoff Research Center. Princeton, N.J., Copyright 1980 IEEE, Jan. 1980.

Carr, Robert et al., The Power of PenPoint, Ch 5, the Class Manager, Copyright 1991, Addison–Wesley, Jan. 1991.

Cashin, "Interprocess Communication", Bell–Northern Research, Document No. 8005014, Jan. 1980.

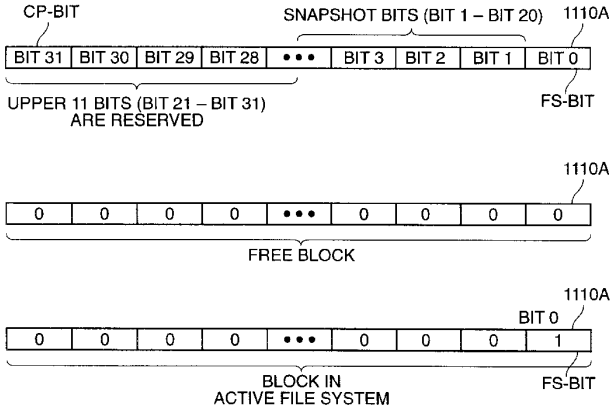
(List continued on next page.)

Primary Examiner—Paul R. Lintz
Attorney, Agent, or Firm—D'Alessandro & Ritchie

[57] **ABSTRACT**

A method is disclosed for maintaining consistent states of a file system. The file system progresses from one self-consistent state to another self-consistent state. The set of self-consistent blocks on disk that is rooted by a root inode is referred to as a consistency point. The root inode is stored in a file system information structure. To implement consistency points, new data is written to unallocated blocks on disk. A new consistency point occurs when the file system information structure is updated by writing a new root inode into it. Thus, as long as the root inode is not updated, the state of the file system represented on disk does not change. The method also creates snapshots that are user-accessible read-only copies of the file system. A snapshot uses no disk space when it is initially created. It is designed so that many different snapshots can be created for the same file system. Unlike prior art file systems that create a done by duplicating an entire inode file and all indirect blocks, the method of the present invention duplicates only the inode that describes the inode file. A multi-bit free-block map file is used to prevent data referenced by snapshots from being overwritten on disk.

20 Claims, 39 Drawing Sheets



U.S. PATENT DOCUMENTS

4,783,730	11/1988	Fischer	395/825
4,803,621	2/1989	Kelly	395/405
4,819,159	4/1989	Shipley et al.	395/182.17
4,825,354	4/1989	Agrawal et al.	395/610
4,875,159	10/1989	Cary et al.	364/200
4,887,204	12/1989	Johnson et al.	395/610
4,897,781	1/1990	Chang et al.	395/617
4,914,583	4/1990	Weisshaar et al.	395/680
5,043,871	8/1991	Nishigaki et al.	395/600
5,043,876	8/1991	Terry	395/600
5,134,619	7/1992	Henson et al.	371/40.1
5,155,835	10/1992	Belsan	395/425
5,163,131	11/1992	Row et al.	395/200
5,163,148	11/1992	Walls	395/600
5,195,100	3/1993	Katz et al.	371/66
5,218,695	6/1993	Noveck et al.	395/600
5,218,696	6/1993	Baird et al.	395/600
5,235,601	8/1993	Stallmo et al.	371/40.1
5,255,270	10/1993	Yanai et al.	371/10.2
5,274,799	12/1993	Brant et al.	395/575
5,274,807	12/1993	Hoshen et al.	395/650
5,276,840	1/1994	Yu	395/425
5,276,867	1/1994	Kenley et al.	395/600
5,305,326	4/1994	Solomon et al.	371/11.1
5,313,626	5/1994	Jones et al.	395/575
5,315,602	5/1994	Noya et al.	371/40.4
5,355,453	10/1994	Row et al. .	
5,357,509	10/1994	Ohizumi	371/10.1
5,379,417	1/1995	Lui et al.	395/575

OTHER PUBLICATIONS

Chao, Chia et al., "MIME: A High Performance Storage Device With Strong Recovery Guarantees", Hewlett–Packard Co, Mar. 18, 1992.

Charlton, David, The Thoth System: Multiprocess Structuring and Portability, Ch2 Thoth Program Environment, and Ch 3: THoth Kernel Design, (1982, North–Holland), Jan. 1982.

Cohen, Jacques, Garbage Collection of Linked Data Structures < Department of Physics, Brandeis University, Waltham, Ma, Computing Surveys, vol. 13, No. 3, Sep. 1981.

Deitel, An Introduction to Operating Structures, Addison–Wesley, Jan. 1984, pp. 302–334.

de Jonge, Wiebrand et al., The Logical Disk, A New Approach to Inmproving File Systems, Jan. 1984, p. 1–14, unpublished.

English et al., Loge: A Self–Organizing Disk Controller, Software and Systems Laboratory, Jan. 1991 (Hewlett–Packard.

Anonymous, File Systems, Chapter 5, pp. 251–273, Jan. 1995.

Mulqueen, John T., Product Analysis Review, Communications Week, vol. 452, pp. 25, May 3, 1993.

Hitz et al., File System for an NFS File Server Appliance, Abstract only from INSPEC Abstravt # C9502–6120–020, Proceedings of the 1994 UNIX Winter Conference, 17–21 Jan. 1994.

Simpson, David, "'Appliances' Take Over File Server Role", Digital News and Review, vol. 11, No. 6, pp. 1–2, Mar. 21, 1994.

Chutani et al, "The Episode File System", USENIX—Winter 1992, pp. 43–60.

Tan, et al., "SOS—Stan's Own Server: A NFS File Server for the IBM PC," Computer Science Research Department, Information and Computing Sciences Division, Lawrence Berkeley Laboratory, Aug. 17, 1988.

Sandberg, "The Sun Network File System: Design, Implementation and Experience," Sun Microsystems, Inc. Technical Report, 1986.

Hitz, "A System Administrator's Performance Monitor for Tuning NFS Network Servers," Auspex Systems, Inc. Technical Report 7, May 1991.

Nelson, Benchmark Methodology and Preliminary Performance Specifications for the Auspect NS5000 Network Server, Auspex Systems, Inc. Technical Report 2, Oct. 1989.

Nelson, et al., "How and Why SCSI Is Better than IPI for NFS," Auspex Systems, Inc. Technical Report 6, 2nd edition, Jul. 1992.

Schröder, "PEACE: The Distributed SUPRENUM Operating System," Parallel Computing, vol. 7 (1988) pp. 325–333.

Malcolm, "A Process Abstraction and its Application," Proc. Eighth Manitoba Conference on Numerical Math and Computing, University of Manitoba at Winnipeg, 1978, pp. 33–50.

Nelson, et al., "The Myth of MIPS for I/O: An Overview of Functional Multiprocessing for NFS Network Servers," Auspex Systems, Inc., Technical Report 1, 6th edition, 2nd printing, Aug. 1992.

Cheriton, et al., "Thoth, a Portable Real–Time Operating System," Communications of the ACM, Feb. 1979, vol. 22, No. 2, pp. 102–115.

Cheriton, "Multi–Process Structuring and the Thoth Operating System," Ph.D. Thesis submitted to Department of Computer Science, University of British Columbia at Vancouver, 1979.

Chutani, Sailesh, et al., "The Episode File System", USENIX Winter Conference, Jan. 30–24, 1992.

Sun Microsystems, "Network Programming Guide", Chapter 4, Remote Procedure Call Programming Guide, Revision A of 27 Mar. 1990, pp. 65–128.

Osadzinski, Alex, "The Network File System (NFS)", 8202 Computer Standards & Interfaces, 8 (1988/89) No. 1, pp. 45–48, Amsterdam, The Netherlands.

Tannebaum, Andrew S., "Computer Networks" (1988), 2nd Edition, Prentice–Hall, pp. 35, 36, Chap. 9.

IEEE Computer, "I/O subsystem", Sep. 1988, pp. 23–25 and 106.

Carlson, et al., "HP AdvanceNet: A Growth–Oriented Computer Networking Architectural Strategy", Hewlett–Packard Journal (Oct. 1986), p. 2, pp. 6–10.

Tribby, David M., "Network Services for HP Real–Time Computers", Hewlett–Packard Journal (Oct. 1986), pp. 22–27.

Motorola, Inc., Microsystems Products Technical Data Sheet (1986), MicoMAP1–7, "MicroMAP MAnufacturing Automation Protocol Software".

Hammond, Richard A. *Expression with the Series/1 Distributed System*, Department of Electrical Engineering, U of Delaware, Newark, Delaware, © 1980 IEEE.

Hanson, Per Brinch (editor), brochure: *RC 400 Software Multiprogramming System*, Ch 4: Process Communication, A/s Regnecentralen, Copenhagen—Apr. 1969.

Hartman, John H., et al., *The Zebra Striped Network File System*.

Hitz, David, *Technical Report TR01: An NFS File Server Appliance*, Network Appliance Corporation.

Hitz, Dave et al., *Technical Report: 3002: File System Design for an NFS File Server Appliance*Network Appliance, Jan. 19, 1994.
Hitz, David, et al., *Using Unix as One Component of a Lightweight Distributed Kernel for Multiprocessor File Servers*, Auspex Technical Report 5, ©1990 Auspex Systems Inc.
Jones, Anita K., et al., *StarOS, a Multiprocessor Operating System for the Support of Task Forces*, Association for Computing Machinery, 1979, pp. 117–127.
Lantz, Keith A., et al., *Rochester's Intelligent Gateway*, Computer © 1982 IEEE.
Leffler, et al., *The Design and Implementation of the 4.3BSD UNIX Operating System*, Addison–Wesley, 1989, pp. 187–245.
Ousterhout, John, et al., *Beating the I/O Bottleneck: A Case for Log–Structured File Systems*, Oct. 30, 1988.
Patterson, David A., et al., A Case for Redundant Arrays of Inexpensive Disks (RAID), UC Berkeley.
Plum, Thomas, *Reliable Data Structures in C*, section 4.9: Pointers to Functions, © 1985 Plum Hall, Inc., pp. 4–23 to 4–27.
Rashid, Richard F. et al., *Accent: A communication oriented network operating system kernel*, Department of Computer Science, Carnegie–Mellon University, Pittsburgh, PA © 981.
Ritchie, D.M. *The UNIX System: A Stream Input–Output System*, AT&T Bell Laboratories Technical Journal, Oct. 1984, vol. 63, No. 8 Part 2, © 1984 AT&T.
Robertazzi, Thomas G., *Computer Networks and Systems: Queuing Theory and Performance Evaluation*, Ch 1: The Queuing Paradigm, © 1990 Springer–Verlag.
Row, John, *Lan Software Links Diverse Machines, OS's*, Mini–Micro Systems, Sep. 1965, pp. 141–146.
Row, John, et al., *Operating System Extensions Link Disparate Systems*, Computer Design, Jul. 1984, pp. 171–183.
Schwartz, Allan M., et al., *LFS—A Local File System for Multiprocessor NFS Network Servers*, Auspex Technical Report 4, © 1989 Auspex Systems, Inc.
Seltzer, Margo, et al., *An Implementation of a Log–Structured File System for UNIX, 1993 Winter UNIX*, San Diego, CA.
Seltzer, Margo, *File System Performance and Transaction Support*, 1992.
Sessions, Roger, *Class Construction in C and C++ Object Oriented Programming Fundamentals*, International Business Machines Corporation, section 4.4: Object Method Inflexibility, © 1992 Prentice–Hall, NJ.
Silberschatz, Abraham, et al., *Operating Concepts*, Ch 3.7: Interprocess Communication, pp. 127–147, © 1988 Addison–Wesley Publishing Company, Inc.
Sincoskie, W., David, et al., *The Series/1 Distributed Operating System: Description and Comments*, IEEE, 1980, pp. 579–584.
Stern, Hal, *Managing NFS and NIS*, NFS Daemons, © 1991 O'Reilly & Associates, Inc.
Tanenbaum, Andrew S., *Operating Systems Design and Implementation*, pp. 251–273.
Tweten, David, *Hiding Mass Storage Under Unix: NASA's MSS II Architecture*, IEEE, 1990, pp. 140–145.
Unix System Laboratories, Inc., *Unix System V Release 4 Programmer's Guide: STREAMS*, section 2: What is STREAMS?, AT & T Unix System Laboratories, Inc., Unix System V Release 4 Programmers's Guide: STREAMS. 1990 Prentice Hall.
UNIX International, *UI–Atlas Distributed Computing Architecture: A Technical Overview*, Oct. 1991.
VRTX Versatile Real–Time Executive for the Microprocessors C: User's Guide, Software Release 1, Document No. 541331001, Apr. 1987.
VRTX32/68020 Versatile Real–Time Executive for the MC68020 Microprocessor: User's Guide, Software Release 3, Document No. 592103003, Jan. 1987.
Walton, Robert L., *Rational for a Queueable Object Distributed Interprocess Communication System*, IEEE Transactions on Communications, vol. Com–30, No. 6, Jun. 1982, pp. 1417–1425.
Wood, B.J., et al., *A Local–Area Network Architecture Based on Message–Passing Operating System Concepts*, 7th Conference on Local Computer Networks, Miniapolis, MN, Oct. 1932.
Sun Microsystems, Inc., Sun OS5.0 Network Interfaces Programmer's Guide, p. 108, Jun. 1992.

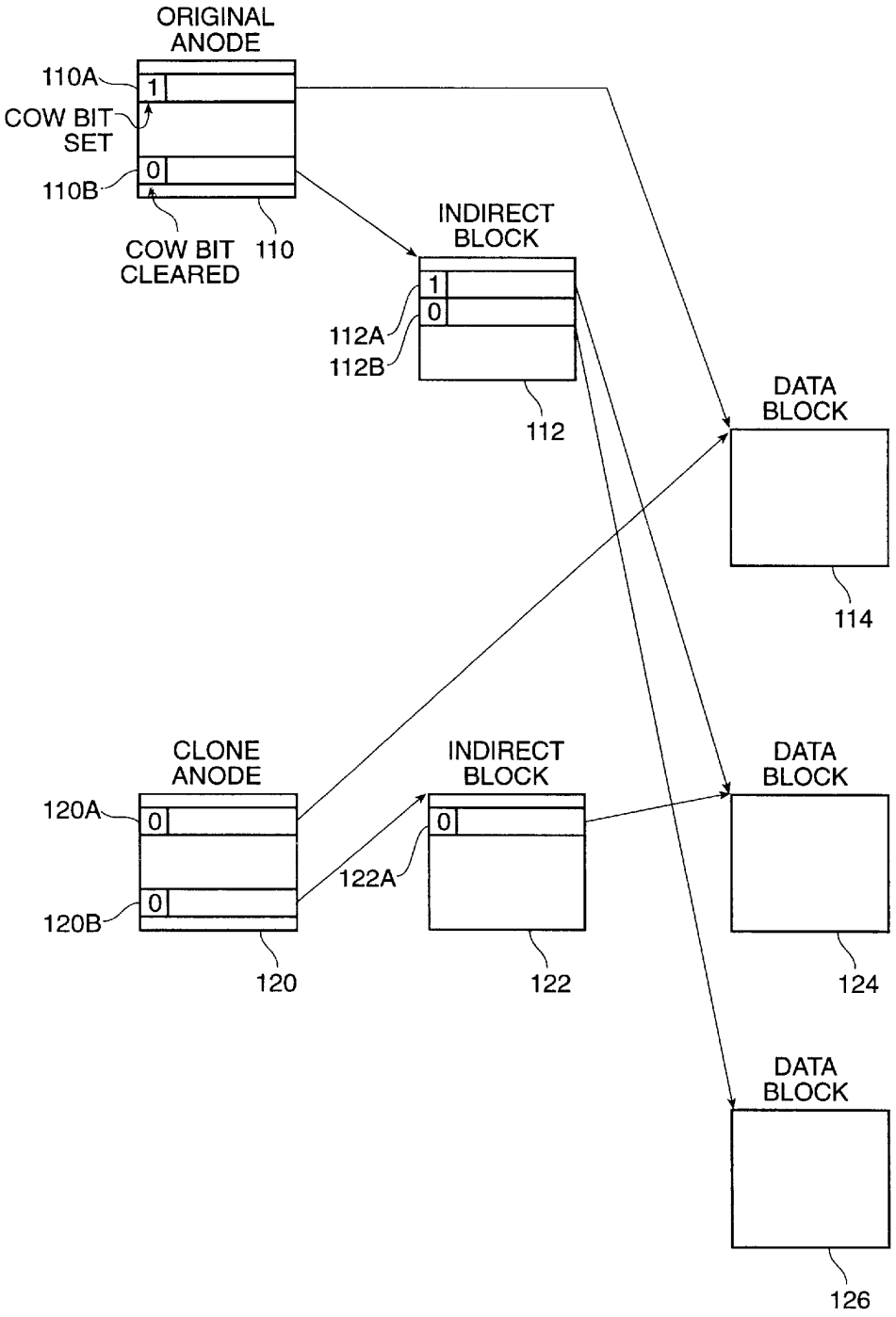


FIG. 1
PRIOR ART

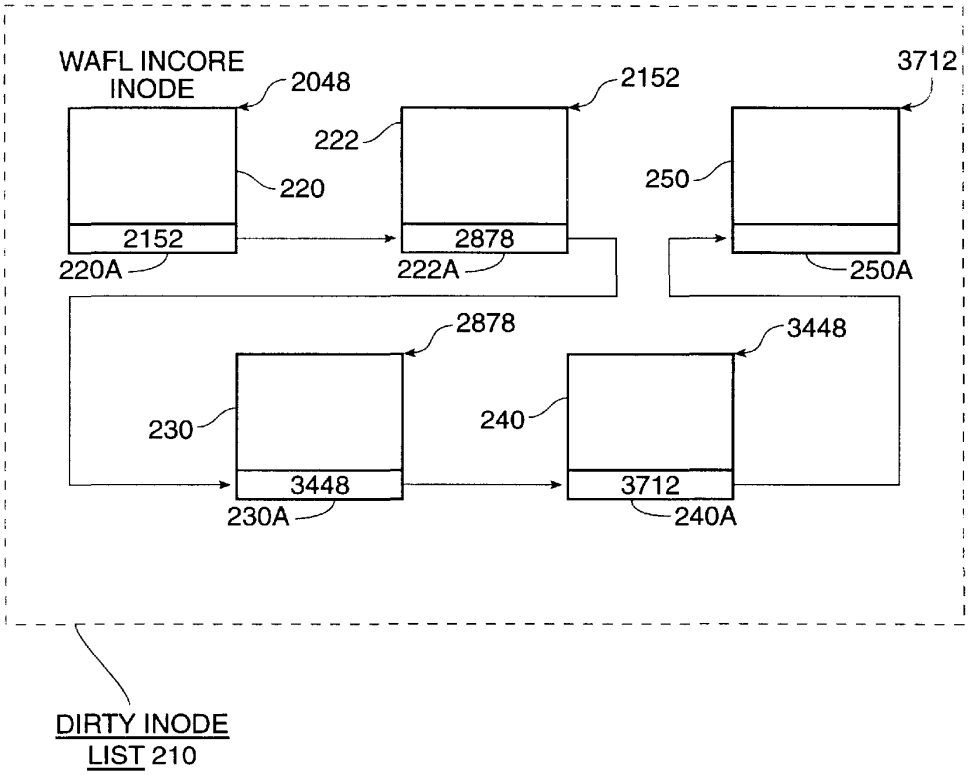


FIG. 2

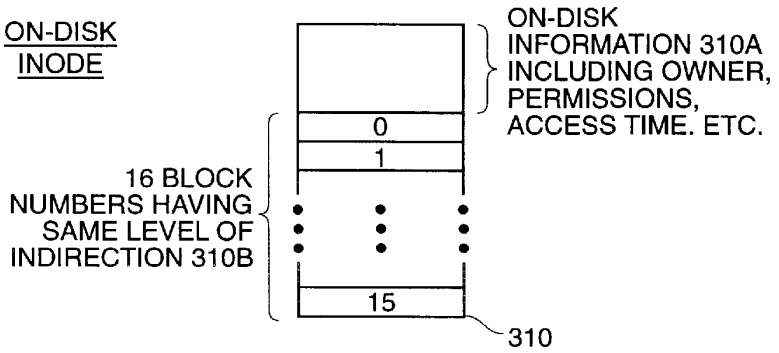


FIG. 3

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.