# EXHIBIT A

US007415565B2

(12) **United States Patent**
**Bullen et al.**

(10) **Patent No.: US 7,415,565 B2**
(45) **Date of Patent: *Aug. 19, 2008**

(54) **METHODS AND SYSTEMS FOR A STORAGE SYSTEM WITH A PROGRAM-CONTROLLED SWITCH FOR ROUTING DATA**

(75) Inventors: **Melvin James Bullen**, Reston, VA (US); **Steven Louis Dodd**, Reston, VA (US); **William Thomas Lynch**, Apex, NC (US); **David James Herbison**, Arvada, CO (US)

(73) Assignee: **Ring Technology Enterprises, LLC**, Reston, VA (US)

( * ) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 179 days.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **10/284,278**

(22) Filed: **Oct. 31, 2002**

(65) **Prior Publication Data**

US 2004/0088514 A1 May 6, 2004

(51) **Int. Cl.**
***G06F 13/16*** (2006.01)

***G06F 12/02*** (2006.01)

(52) **U.S. Cl.** ..... **710/316**; 710/38; 710/31; 711/5

(58) **Field of Classification Search** ..... 711/109, 711/167, 202, 118, 162, 173; 365/189.12, 365/201, 240, 221, 73; 714/718; 370/242, 370/250, 359, 363, 379, 392, 395.7; 725/145, 725/147, 92, 93, 114; 709/217; 710/316
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

| | | | | |
|---|---|---|---|---|
| 3,713,096 | A | 1/1973 | Comfort et al. | 709/251 |
| 3,735,362 | A | 5/1973 | Ashany et al. | 710/316 |
| 3,748,647 | A | 7/1973 | Ashany et al. | 710/316 |
| 3,812,476 | A | 5/1974 | Cragon | 365/73 |
| 4,064,556 | A | 12/1977 | Edelberg et al. | 711/110 |
| 4,065,756 | A | 12/1977 | Panigrahi | 365/49 |
| 4,193,121 | A | 3/1980 | Fedida et al. | 711/110 |
| 4,302,632 | A | 11/1981 | Vicari et al. | 365/214.01 |
| 4,334,305 | A | 6/1982 | Girardi | 370/359 |
| 4,363,125 | A | 12/1982 | Brewer et al. | 714/824 |
| 4,506,387 | A | 3/1985 | Walter | 398/66 |
| 4,510,599 | A * | 4/1985 | Ulug | 370/463 |

(Continued)

FOREIGN PATENT DOCUMENTS

WO WO 97/07458 2/1997

(Continued)

OTHER PUBLICATIONS

Farley, Marc, "Building Storage Networks, Second Edition", Osbourne/McGraw-Hill, 2001, entire book.

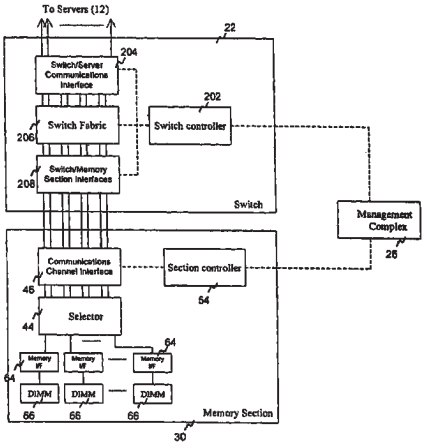
(Continued)

*Primary Examiner*—B. James Peikari
(74) *Attorney, Agent, or Firm*—Finnegan, Henderson, Farabow, Garrett & Dunner, L.L.P.

(57) **ABSTRACT**

A storage system that may include one or more memory devices, a memory interface device corresponding to one or more of the memory devices, which are organized in sections, a section controller, and a switch. The switch is capable of reading a data request including a data block identifier and routing the data request and any associated data through the switch on the basis of this data block identifier, such that a data request may be routed to a memory section. The section controller, in response, determines the addresses in the memory devices storing the requested data, and it transfers these addresses to those memory devices storing the requested data.

**27 Claims, 15 Drawing Sheets**



U.S. PATENT DOCUMENTS

| | | | | |
|---|---|---|---|---|
| 4,538,174 | A | | 8/1985 | Gargini et al. ..... 725/120 |
| 4,646,270 | A | | 2/1987 | Voss |
| 4,672,602 | A | | 6/1987 | Hargrave et al. ..... 370/360 |
| 4,709,418 | A | | 11/1987 | Fox et al. ..... 398/67 |
| 4,763,317 | A | | 8/1988 | Lehman et al. ..... 370/358 |
| 4,790,418 | A | | 12/1988 | Brown et al. ..... 701/51 |
| 4,796,231 | A | * | 1/1989 | Pinkham ..... 365/189.05 |
| 4,980,857 | A | | 12/1990 | Walter et al. |
| 4,984,240 | A | | 1/1991 | Keren-Zvi et al. |
| 4,995,078 | A | | 2/1991 | Monslow et al. ..... 380/240 |
| 5,003,591 | A | | 3/1991 | Kauffman et al. ..... 380/232 |
| 5,014,125 | A | | 5/1991 | Pocock et al. ..... 725/93 |
| 5,027,400 | A | | 6/1991 | Baji et al. ..... 725/116 |
| 5,060,068 | A | | 10/1991 | Lindstrom ..... 725/32 |
| 5,062,059 | A | | 10/1991 | Youngblood et al. ..... 709/217 |
| 5,084,839 | A | | 1/1992 | Young ..... 365/73 |
| 5,119,481 | A | | 6/1992 | Frank et al. ..... 710/100 |
| 5,130,792 | A | | 7/1992 | Tindell et al. ..... 725/146 |
| 5,132,992 | A | | 7/1992 | Yurt et al. ..... 375/240 |
| 5,133,079 | A | | 7/1992 | Ballantyne et al. ..... 725/146 |
| 5,153,884 | A | | 10/1992 | Lucak et al. ..... 714/748 |
| 5,163,024 | A | * | 11/1992 | Heilveil et al. ..... 365/219 |
| 5,191,410 | A | | 3/1993 | McCalley et al. ..... 725/114 |
| 5,200,925 | A | | 4/1993 | Morouka et al. ..... 365/219 |
| 5,247,347 | A | | 9/1993 | Litteral et al. ..... 725/114 |
| 5,253,341 | A | | 10/1993 | Rozmanith et al. ..... 709/219 |
| 5,261,114 | A | | 11/1993 | Raasch et al. ..... 709/221 |
| 5,285,451 | A | | 2/1994 | Henson et al. |
| 5,369,784 | A | | 11/1994 | Nelson et al. ..... 455/503 |
| 5,371,532 | A | | 12/1994 | Gelman et al. ..... 725/88 |
| 5,374,952 | A | | 12/1994 | Flohr ..... 348/14.08 |
| 5,400,331 | A | | 3/1995 | Lucak et al. ..... 370/401 |
| 5,553,311 | A | | 9/1996 | McLaughlin et al. ..... 710/64 |
| 5,581,479 | A | | 12/1996 | McLaughlin et al. ..... 725/145 |
| 5,604,682 | A | | 2/1997 | McLaughlin et al. ..... 709/219 |
| 5,636,139 | A | | 6/1997 | McLaughlin et al. ..... 709/219 |
| 5,729,763 | A | | 3/1998 | Leshem |
| 5,768,623 | A | | 6/1998 | Judd et al. |
| 5,771,367 | A | | 6/1998 | Beardlsey et al. |
| 5,883,831 | A | | 3/1999 | Lopez et al. ..... 365/185.04 |
| 5,908,333 | A | | 6/1999 | Perino et al. |
| 5,909,564 | A | * | 6/1999 | Alexander et al. ..... 710/316 |
| 5,953,263 | A | | 9/1999 | Farmwald et al. |
| 5,954,804 | A | | 9/1999 | Farmwald et al. |
| 5,968,114 | A | * | 10/1999 | Wentka et al. ..... 718/100 |
| 5,978,295 | A | | 11/1999 | Pomet et al. ..... 365/221 |
| 6,032,214 | A | | 2/2000 | Farmwald et al. |
| 6,034,918 | A | | 3/2000 | Farmwald et al. |
| 6,185,644 | B1 | | 2/2001 | Farmwald et al. |
| 6,317,377 | B1 | | 11/2001 | Kobayashi |
| 6,356,973 | B1 | | 3/2002 | McLaughlin et al. ..... 711/100 |
| 6,356,975 | B1 | | 3/2002 | Barth et al. |
| 6,498,741 | B2 | | 12/2002 | Matsudera et al. ..... 365/63 |
| 6,587,909 | B1 | | 7/2003 | Olarig et al. |
| 6,684,292 | B2 | | 1/2004 | Piccirillo et al. |
| 6,697,368 | B2 | * | 2/2004 | Chang et al. ..... 370/395.1 |
| 6,728,799 | B1 | * | 4/2004 | Perner et al. ..... 710/52 |
| 6,879,526 | B2 | | 4/2005 | Lynch et al. |
| 6,981,173 | B2 | | 12/2005 | Ferguson et al. |
| 7,069,468 | B1 | | 6/2006 | Olson et al. |
| 7,197,662 | B2 | | 3/2007 | Bullen et al. |
| 7,266,706 | B2 | | 9/2007 | Brown et al. |
| 2003/0018930 | A1 | | 1/2003 | Mora et al. |
| 2003/0135782 | A1 | | 7/2003 | Matsunami et al. |
| 2003/0187945 | A1 | | 10/2003 | Lubbers et al. |
| 2004/0044744 | A1 | | 3/2004 | Grosner et al. |
| 2004/0068561 | A1 | | 4/2004 | Yamamoto et al. |
| 2004/0168101 | A1 | | 8/2004 | Kubo |
| 2005/0025321 | A1 | * | 2/2005 | Ajamian ..... 381/119 |

FOREIGN PATENT DOCUMENTS

| | | |
|---|---|---|
| WO | WO 2004/025476 A1 | 3/2004 |

OTHER PUBLICATIONS

Clark, Tom, "Designing Storage Area Networks—A Practical Reference for Implementing Fibre Channel SANs", Addison-Wesley, 1999.

U.S. Appl. No. 10/284,198 including Specification, Claims and Figures.

U.S. Appl. No. 10/284,199 including Specification, Claims and Figures.

U.S. Appl. No. 10/284,268 including Specification, Claims and Figures.

Preston, W. Curtis, "Using SANs and NAS," O'Reilly & Associates, Inc., Feb. 2002.

Mauro, Doouglas R., et al., "Essential SNMP," O'Reilly & Associates, Inc., Jul. 2001.

"Data-In Tristate Buffer" (retrieved Mar. 17, 2002) <http://www.mit.edu:8001/people/tairan/6371/node64.html>, MIT, Jun. 26, 1996, 1 page.

"Data-out Precharging Circuits and Control Circuits" (retrieved Mar. 17, 2002) <http://www.mit.edu:8001/people/tairan/6371/node65.html>, MIT, Jun. 26, 1996, 1 page.

"Output Multiplexer" (retrieved Mar. 17, 2002), <http://www.mit.edu:8001/people/tairan/6371/node66.html>, MIT Jun. 26, 1996, 1 page.

"Other Circuits" (retrieved Mar. 17, 2002) <http://www.mit.edu:8001/people/tairan/6371/node67.html>, MIT, Jun. 26, 1996, 1 page.

"Simulations" (retrieved Mar. 17, 2002) <http://www.mit.edu:8001/people/tairan/6371/node68.html>, MIT, Jun. 26, 1996, 1 page.

"Comments" (retrieved Mar. 17, 2002) <http://www.mit/edu:8001/people/tairan/6371/node70.html>, MIT, Jun. 26, 1995, 1 page.

"HHSPICE Verification of Latch" (retrieved Mar. 17, 2002), <http://www.mit.edu:8001/people/tairan/6371/node73.html>, MIT, Jun. 26, 1996, 1 page.

"Introduction" (retrieved Mar. 17, 2002) http://www4.tomshardware.com/mainboard/98q4/981024/index.html, 2 pages.

"Basic DRAM Operation" (retrieved Mar. 17, 2002) <http://www4.tomshardware.com/mainboard/98q4/981024/ram-01.html>, 2 pages.

"Asynchronous Operation" (retrieved Mar. 17, 2002), <http://www4.tomshardware.com/mainboard/98q4/981024.ram-02.html>, 2 pages.

"Synchronous Operation" (retrieved Mar. 17, 2002) <http://www4.tomshardware.com/mainboard/98q4/981024/ram-07.html>, 2 pages.

"DDR SDRAM" (retrieved Mar. 17, 2002), <http://www4.tomshardware.com/mainboard/98q4/981024/ram-10.html>, 2 pages.

"170 MHz FIFOs Using the Virtex Block SelectRAM+ Features" Xilinx Application Note XAPP131, Jun. 5, 2001, 6 pages.

"Using the Virtex Block SelectRAM+ Features" Xilinx Application Note XAPP130, Dec. 18, 2000, 11 pages.

"API Networks, enabling technology for next generation product . . . HyperTransport technology licensed by HP" (retrieved Mar. 18, 2002) <http://www.api-networks.com/pressreleases/pr121001.shtml>, API Networks, Dec. 10, 2001, 2 pages.

Richmond, Robert, "AMD 64-Bit K8 Platform Preview" (retrieved Mar. 11, 2002) Sep. 14, 2000, 4 pages.

"Block SelectRAM Overview" (retrieved Mar. 18, 2002). <http://www.xilinx.com/xil_prodcat_systemsolution.jsp?title=xam_memory_embedded_blockram_pag>, 2 pages.

McComas, Bert "PCI-X or InfiniBand Complementary New Technologies Go Head to Head" (retrieved Mar. 18, 2002) <http://www.inqst.com/articles/pcixvib/pciarticle.htm>, Inquest Market Research, Jan. 19, 2001, 10 pages.

"API Networks Accelerates Use of HyperTransport™ Technology With Launch of Industry's First HyperTransport-to-PCI Bridge Chip" API Networks Press Release, Apr. 2, 2001, 2 pages.

"HyperTransport-to-PCI Bridge Chip from API Networks" Cahners Business Information 2002, 1 page.
"API Networks Unveils Industry's First HyperTransport™ Switch to Bring Products to Market Quickly and Cost-Effectively" API Networks Press Release, Nov. 5, 2001, 4 pages.
"74F1763 Intelligent DRAM Controller Product Information" Philips Semiconductors 2002, 2 pages.
"Basic DRAM Cell" (retrieved Mar. 17, 2002) <http://www.mit.edu:8001/people/tairan/6371/node59.html>, MIT, Jun. 26, 1996, 1 page.
"Row Address Decoder and Row Driver" (retrieved Mar. 17, 2002) <http://www.mit.edu:8001/people/tairan/6371/node60.html>, MIT, Jun. 26, 1996, 1 page.
"Column Decode and Refresh Control Logic" (retrieved Mar. 17, 2002) <http://www.mit.edu:8001/people/tairan/6371/node61.html>, MIT, Jun. 26, 1996, 1 page.
"Refresh Circuit" (retrieved Mar. 17, 2002) <http://www.mit.edu:8001/people/tairan/6371/node62.html>, MIT, Jun. 26, 1996, 1 page.
"Refresh Address Control" (retrieved Mar. 17, 2002) <http://www.mit.edu:8001/people/tairan/6371/node63.html>, Jun. 26, 1996, 1 page.
Malavalli, Kumar, "Fibre Channel Classes of Service for Data Transport," Brocade Communications Services, Inc. 1997, 15 pages.
"HyperTransport™ I/O Link Specification," HyperTransport Technology Consortium, Rev. 1.03, Oct. 10, 2001, 217 pages.
Tran, Jennifer, "Synthesizable 1.6 Gbytes/s DDR SRAM Controller" Xilinx Application Note XAPP200, Mar. 21, 2000, 12 pages.
"Using Block SelectRAM+ Memory in Spartan-II FPGAs," Xilinx Application Note XAPP173, Dec. 11, 2000, 15 pages.
"200-MHz SDRAM Controller Core Product Specification" Alliance Core, Jan. 10, 2000, 4 pages.
Bapat, Sheker, "Synthesizable 200 MHz ZBT SRAM Interface" Xilinx Application Note XAPP136, Jan. 10, 2000, 6 pages.
"Synthesizable High Performance SDRAM Controller" Xilinx Application Note XAPP134, Feb. 1, 2001, 16 pages.
Ma, Alex, "Synchronous DRAM Controller" Powerpoint slides, EE527 Spring 1998, 21 pages.
PCT Search Report dated Jan. 14, 2005 for International Application No. PCT/US03/33665.
Prince, Betty, "High Performance Memories New architecture DRAMs and SRAMs—evolution and function Revised Edition," John Wiley & Sons, Ltd., 1996, entire book.
RDRAM® Memory: Leading Performance and Value Over SDRAM and DDR, Document WP0001-R, Version 1.2, 2001.
RDRAM®: Maximizing the Value of PCs and Workstations, Document WP0003-R, Version 1.0, 2001.
The Economist, Jun. 22nd, 2002, 28th edition, A Match for Flash?, pp. 22-23.
Office Action for Application No. 10/284,268 dated Mar. 18, 2005, (14 pages).
Office Action for Application No. 10/284,268 dated Sep. 21, 2005, (13 pages).
Office Action for Application No. 11/030,881 dated Jan. 11, 2006, (5 pages).
Office Action for Application No. 11/030,881 dated Jul. 27, 2005, (8 pages).
Office Action for Application No. 11/030,881 dated Jul. 10, 2006, (5 pages).
Office Action for Application No. 10/284,199 dated Jun. 15, 2006, (12 pages).
Office Action for Application No. 10/284,268 dated Apr. 24, 2006, (16 pages).
European Patent Office Action dated Mar. 27, 2007, for European Application No. 38 810 796.7 (4 pages).
Office Action for U.S. Appl. No. 11/030,881 dated Feb. 28, 2007, (7 pages).
Notice of Allowance for U.S. Appl. No. 11/030,881 dated Jun. 28, 2007, with Reply filed May 8, 2007, listing allowed claims (14 pages).
Office Action for U.S. Appl. No. 10/284,268, dated Apr. 11, 2007 (5 pages).
Supplemental European Search Report for European Appl. No. EP 03 77 7844 dated Oct 12, 2007 (2 pages).
Communication Pursuant to Article 94(3) EPC for EPO Appl. No. 03 777 844.6 dated Feb. 11, 2008 (10 pages).
Communication Pursuant to Article 94(3) EPC for EPO Appl. No.03 810 796.7 dated Apr. 25, 2008 (4 pages).
Office Action for U.S. Appl. No. 11/710,407 mailed Mar. 27, 2008 (10 pages).
Office Action for U.S. Appl. No. 10/284,268 mailed Dec. 31, 2007 (15 pages).

* cited by examiner

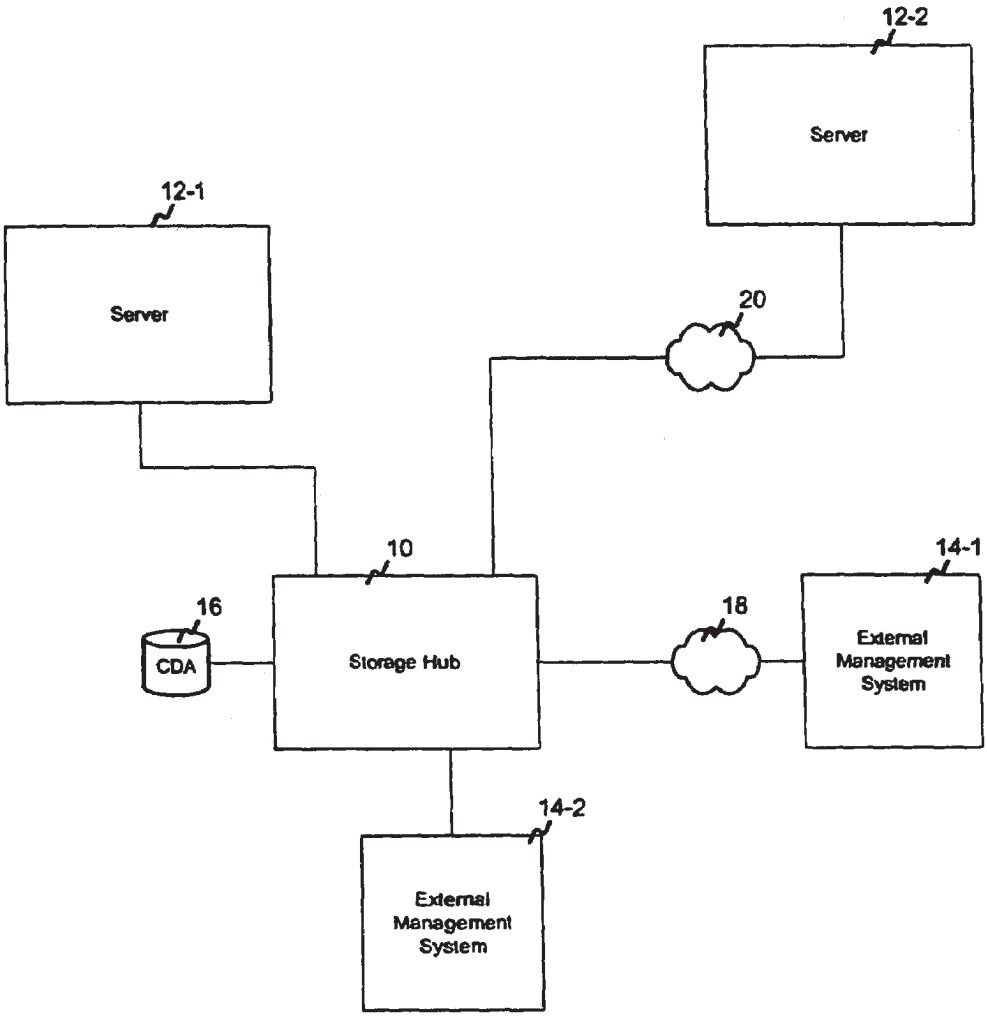

12-2
Server
12-1
Server
20
10
14-1
16
CDA
Storage Hub
18
External Management System
14-2
External Management System

FIG. 1

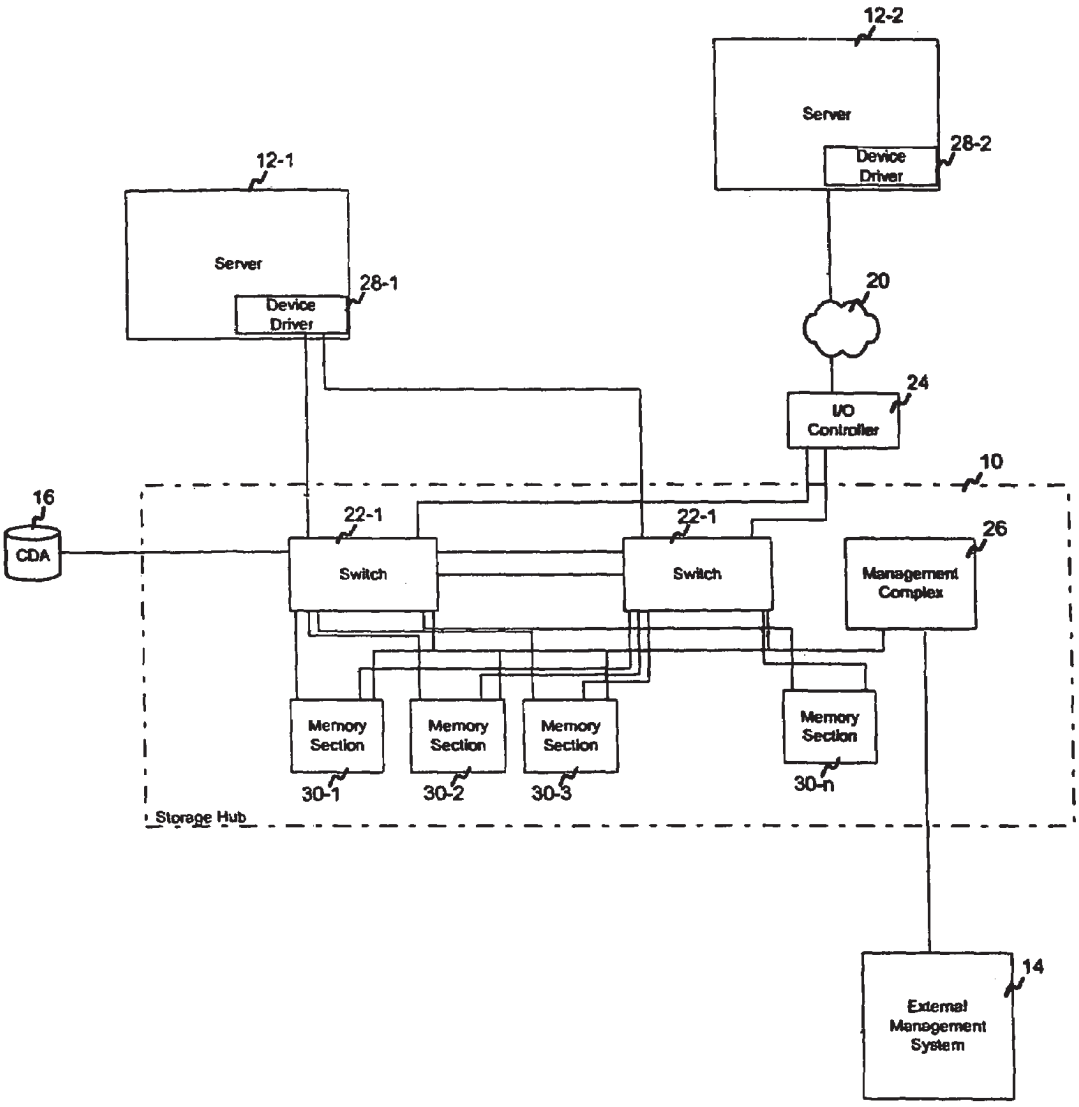

12-2
Server
Device Driver
28-2
12-1
Server
Device Driver
28-1
20
24
I/O Controller
10
16
CDA
22-1
Switch
22-1
Switch
26
Management Complex
Memory Section
30-1
Memory Section
30-2
Memory Section
30-3
Memory Section
30-n
Storage Hub
14
External Management System

FIG. 2

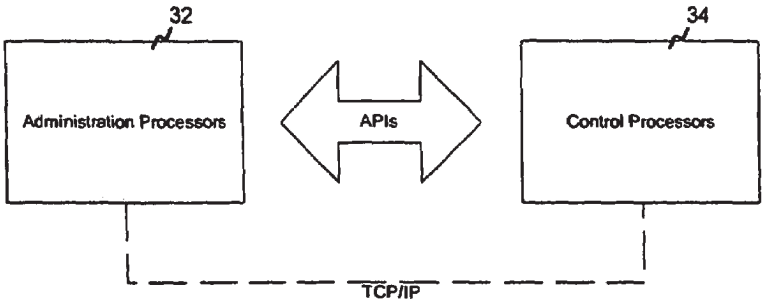

32
34
Administration Processors
APIs
Control Processors
TCP/IP

FIG. 3

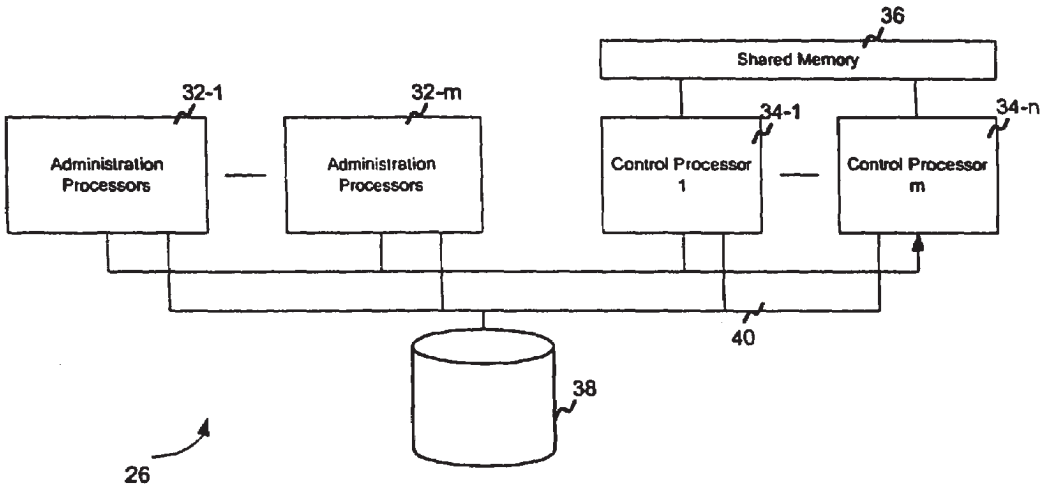

36
Shared Memory
32-1
32-m
34-1
34-n
Administration Processors
Administration Processors
Control Processor 1
Control Processor m
40
38
26

FIG. 4

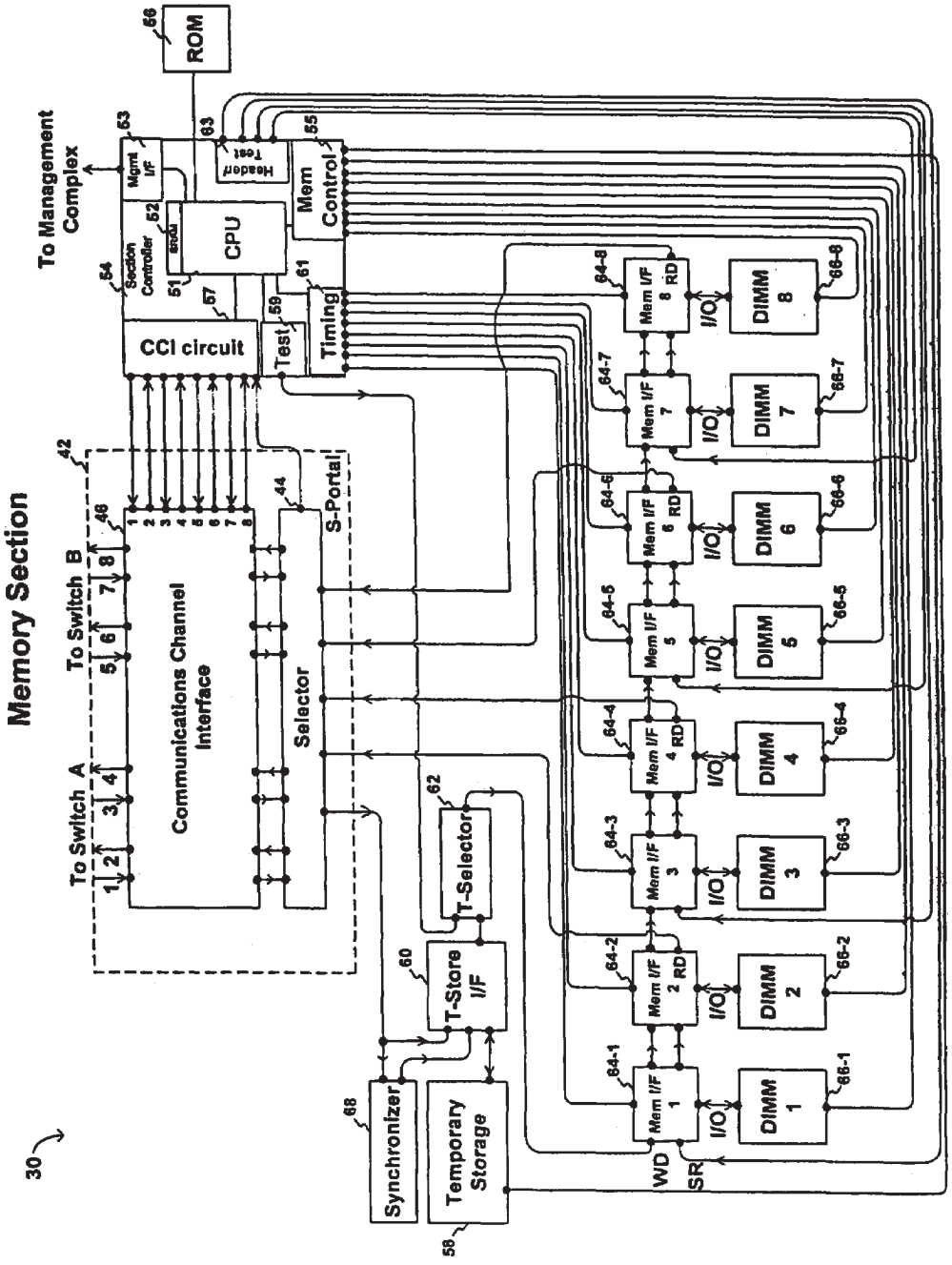

Memory Section
30
To Management Complex
ROM
CPU
Mem Control
Header Test
Mgmt I/F
Section Controller
CCI circuit
Test
Timing
To Switch A
To Switch B
Communications Channel Interface
Selector
S-Portal
T-Selector
T-Store I/F
Synchronizer
Temporary Storage
Mem I/F
RD
I/O
DIMM
WD
SR

FIG. 5

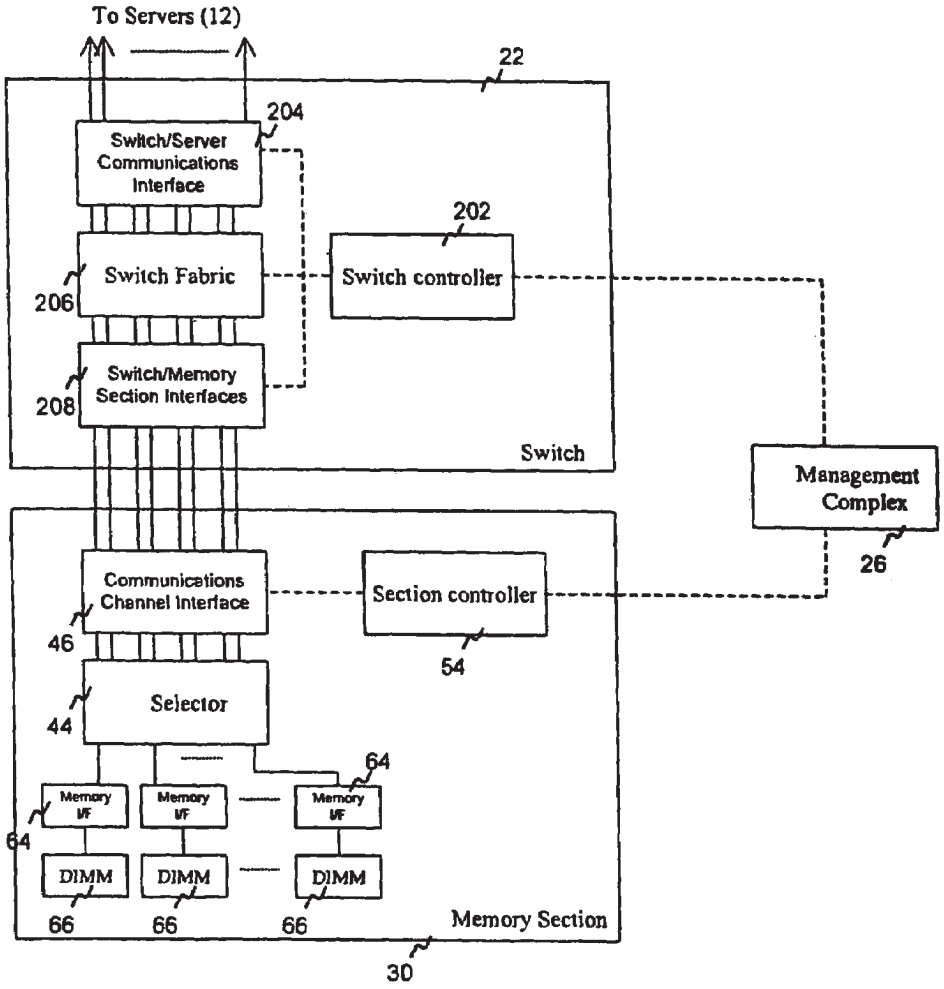

To Servers (12)
22
204
Switch/Server Communications Interface
202
Switch Fabric
206
Switch controller
Switch/Memory Section Interfaces
208
Switch
Management Complex
26
Communications Channel Interface
46
Section controller
54
Selector
44
64
Memory I/F
64
DIMM
66
66
66
Memory Section
30

FIG. 6

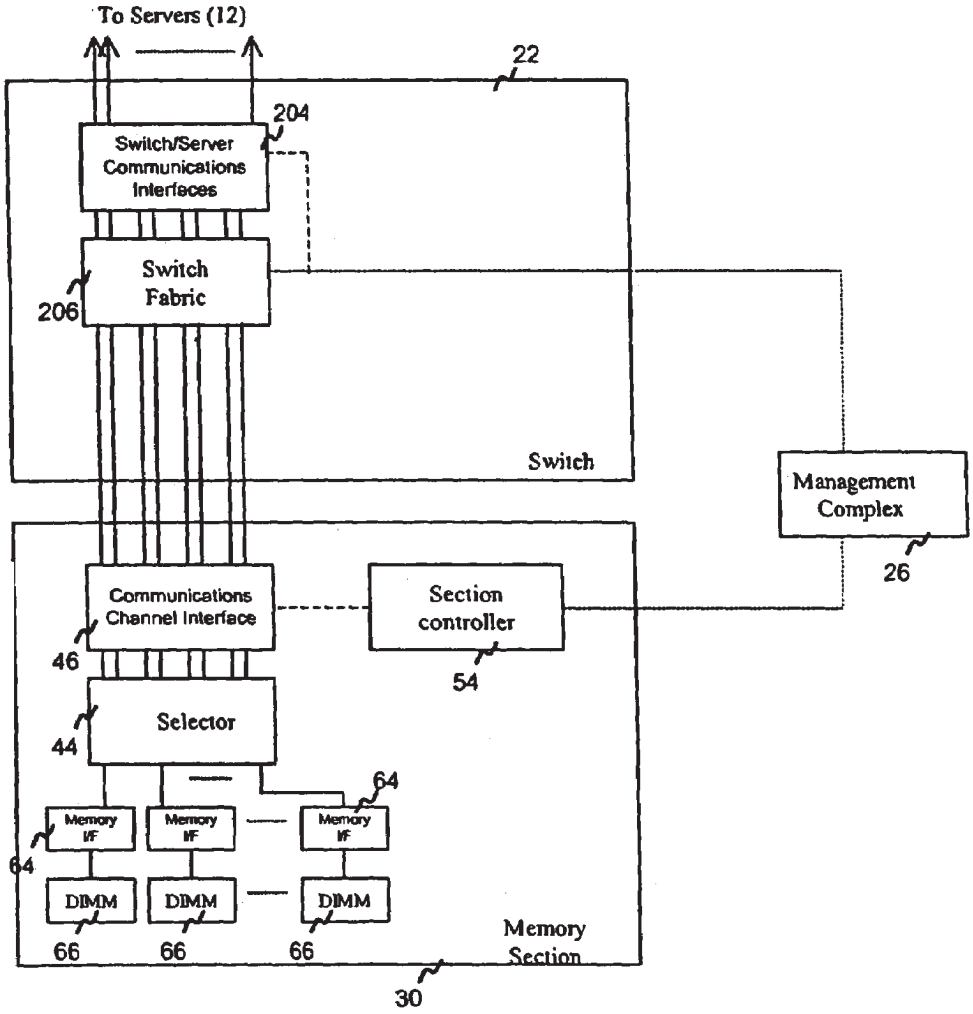

To Servers (12)
22
204
Switch/Server Communications Interfaces
Switch Fabric
206
Switch
Management Complex
26
Communications Channel Interface
46
Section controller
54
Selector
44
64
Memory I/F
64
DIMM
66
66
66
Memory Section
30

FIG. 7

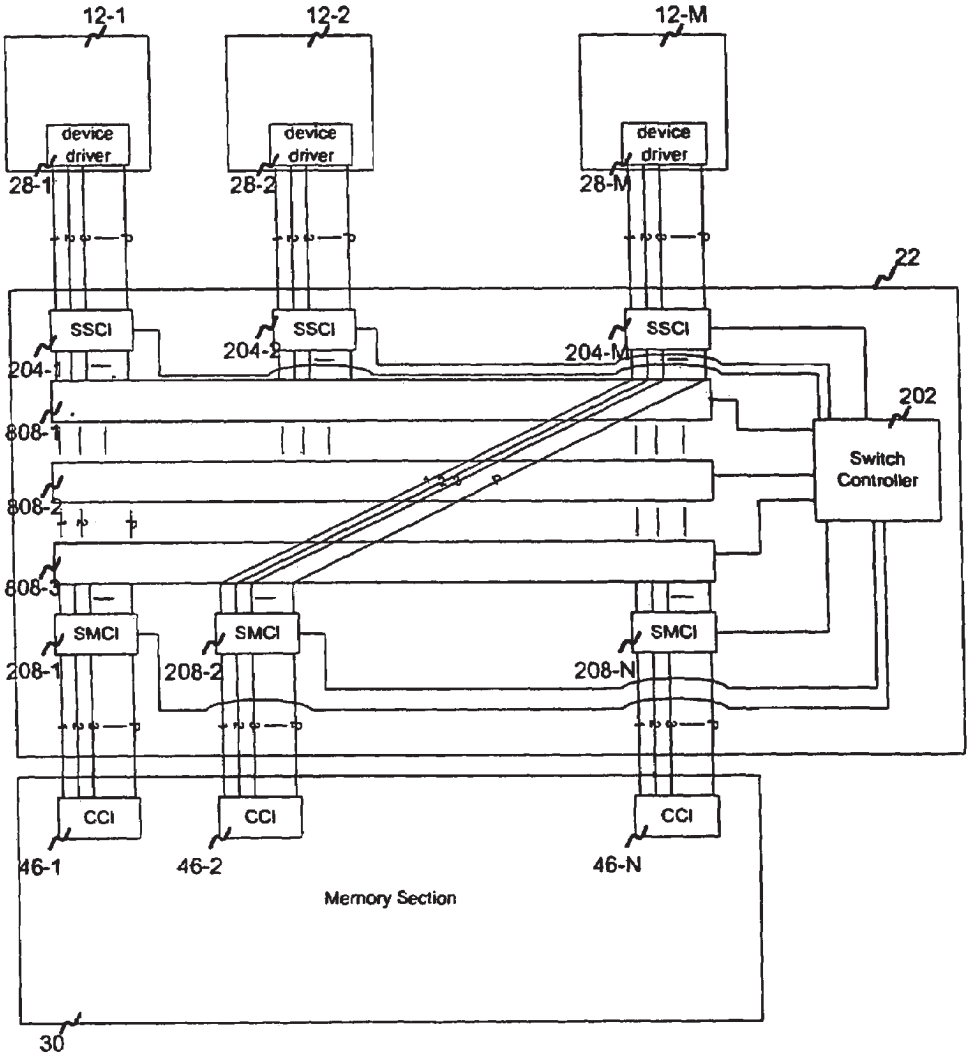

12-1
12-2
12-M
device driver
28-1
28-2
28-M
22
SSCI
204-1
204-2
204-M
202
Switch Controller
808-1
808-2
808-3
SMCI
208-1
208-2
208-N
CCI
46-1
46-2
46-N
Memory Section
30

FIG. 8

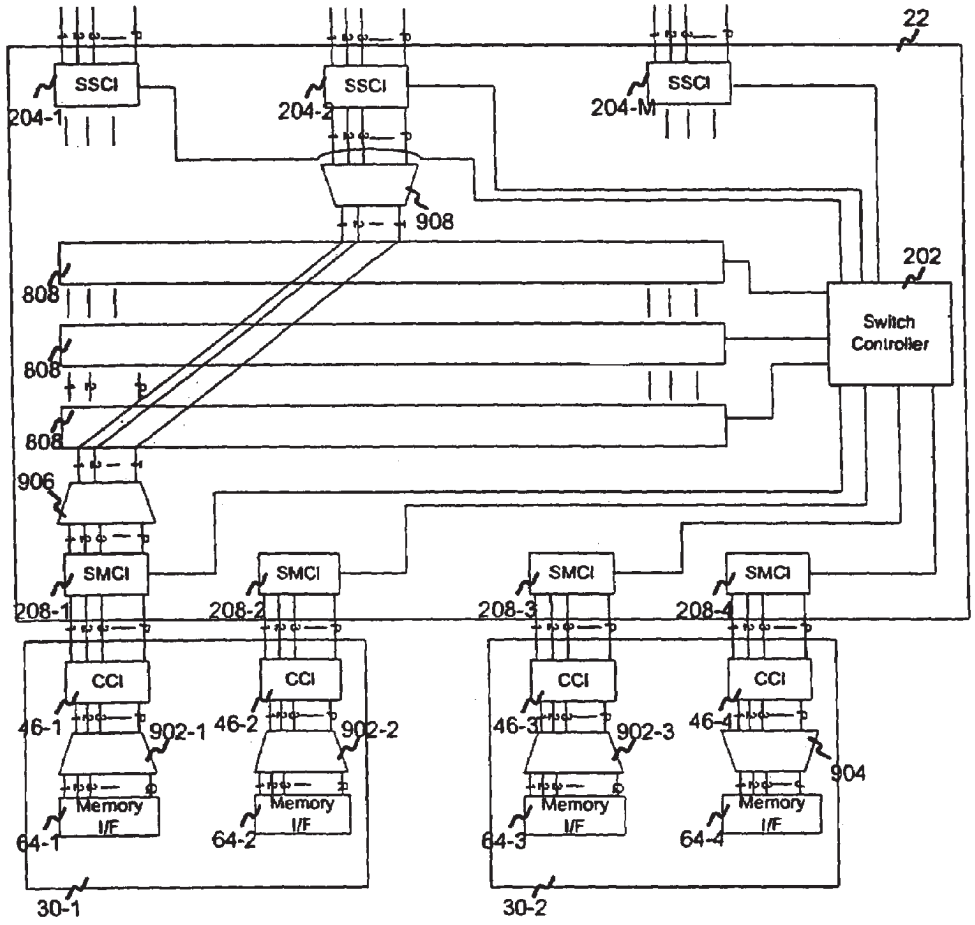

22
SSCI
204-1
SSCI
204-2
SSCI
204-M
908
202
Switch Controller
808
808
808
906
SMCI
208-1
SMCI
208-2
SMCI
208-3
SMCI
208-4
CCI
46-1
902-1
CCI
46-2
902-2
CCI
46-3
902-3
CCI
46-4
904
Memory I/F
64-1
Memory I/F
64-2
Memory I/F
64-3
Memory I/F
64-4
30-1
30-2

FIG. 9

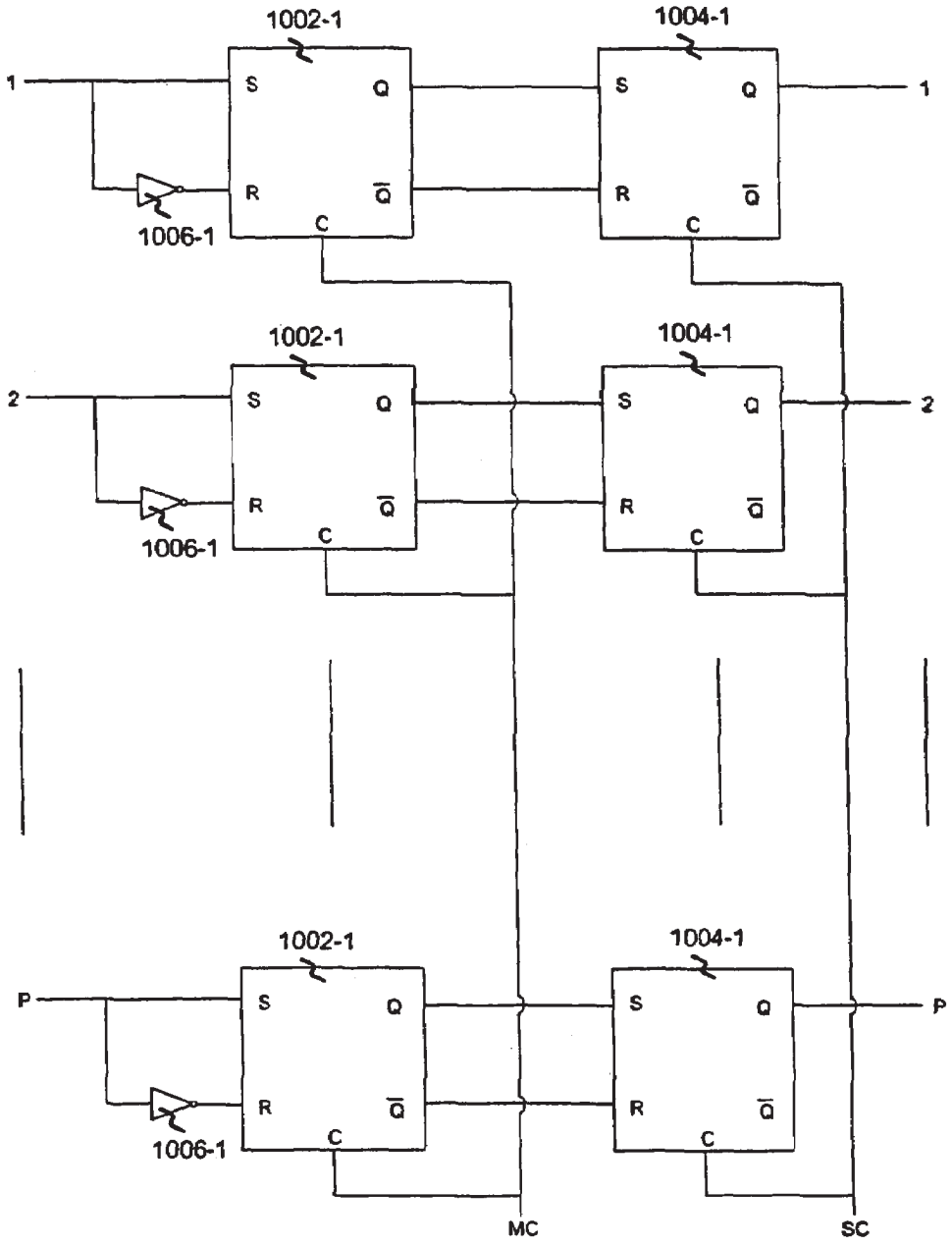



FIG. 10

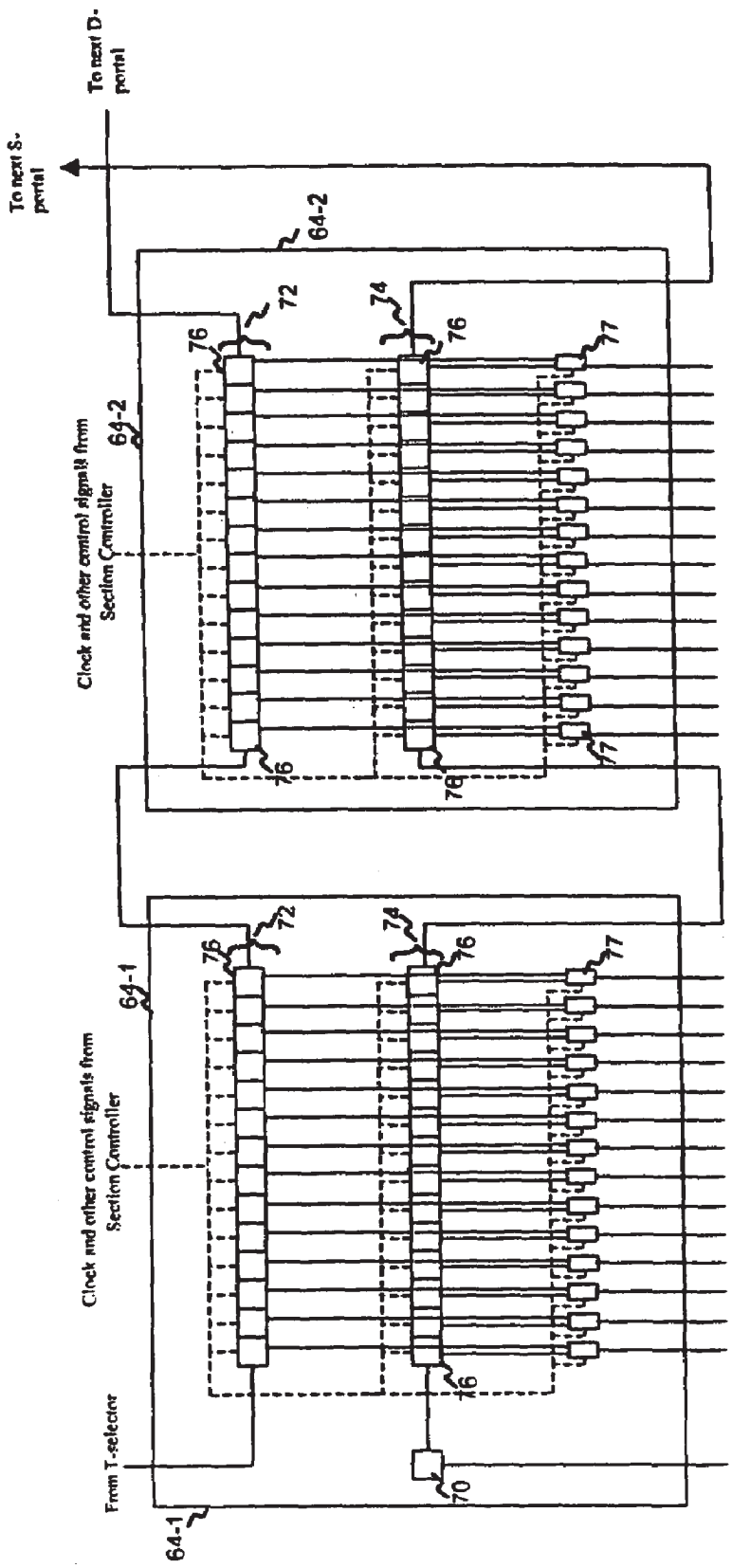

To next S-portal
To next D-portal
64-2
72
74
76
77
78
64-2
Clock and other control signals from Section Controller
To Pins of Memory Device (66-2)
64-1
72
74
76
77
78
64-1
Clock and other control signals from Section Controller
From T-selector
70
To Pins of Memory Device (66-1)

FIG. 11

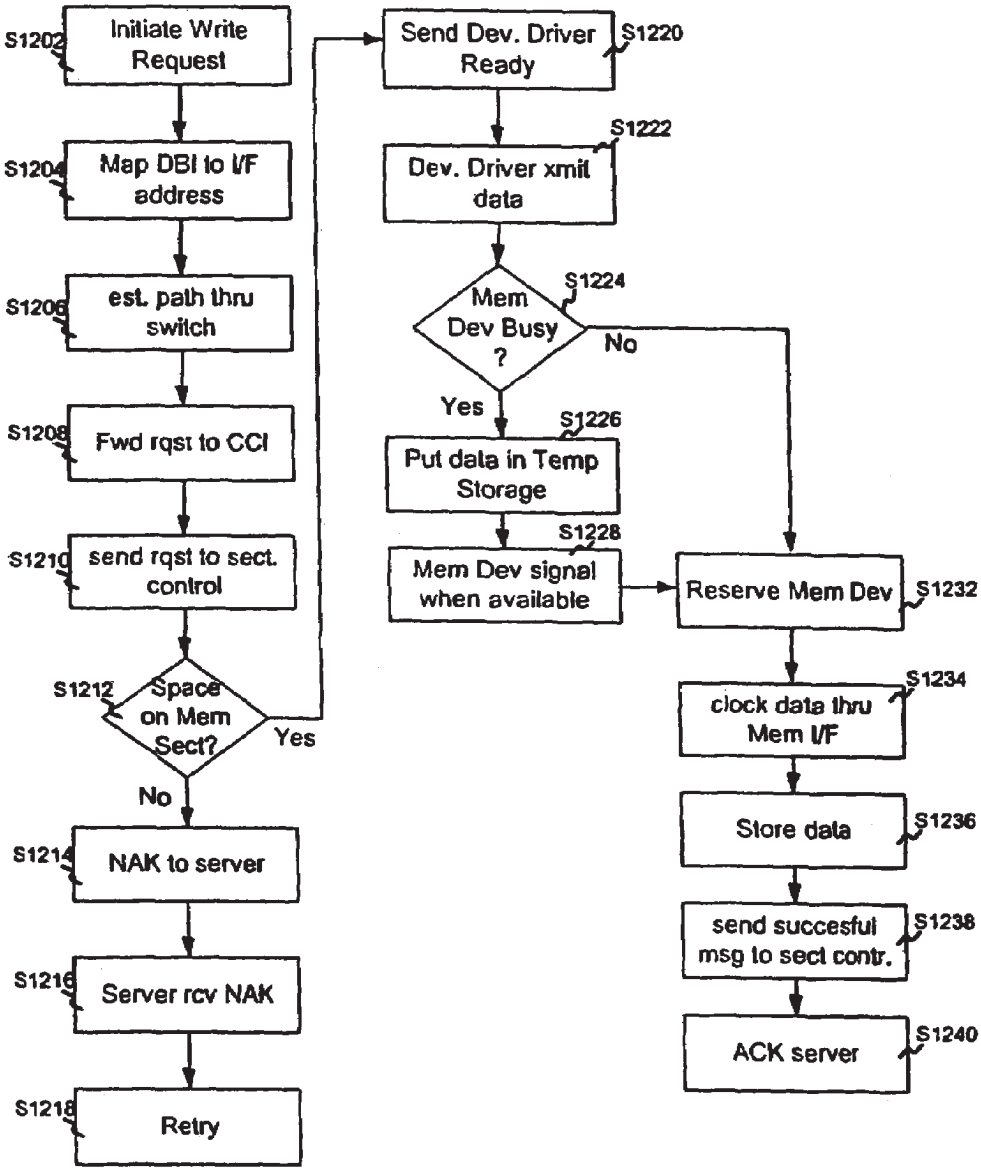

S1202
Initiate Write Request
S1204
Map DBI to I/F address
S1206
est. path thru switch
S1208
Fwd rqst to CCI
S1210
send rqst to sect. control
S1212
Space on Mem Sect?
Yes
No
S1214
NAK to server
S1216
Server rcv NAK
S1218
Retry
S1220
Send Dev. Driver Ready
S1222
Dev. Driver xmit data
S1224
Mem Dev Busy ?
No
Yes
S1226
Put data in Temp Storage
S1228
Mem Dev signal when available
S1232
Reserve Mem Dev
S1234
clock data thru Mem I/F
S1236
Store data
S1238
send succesful msg to sect contr.
S1240
ACK server

FIG. 12

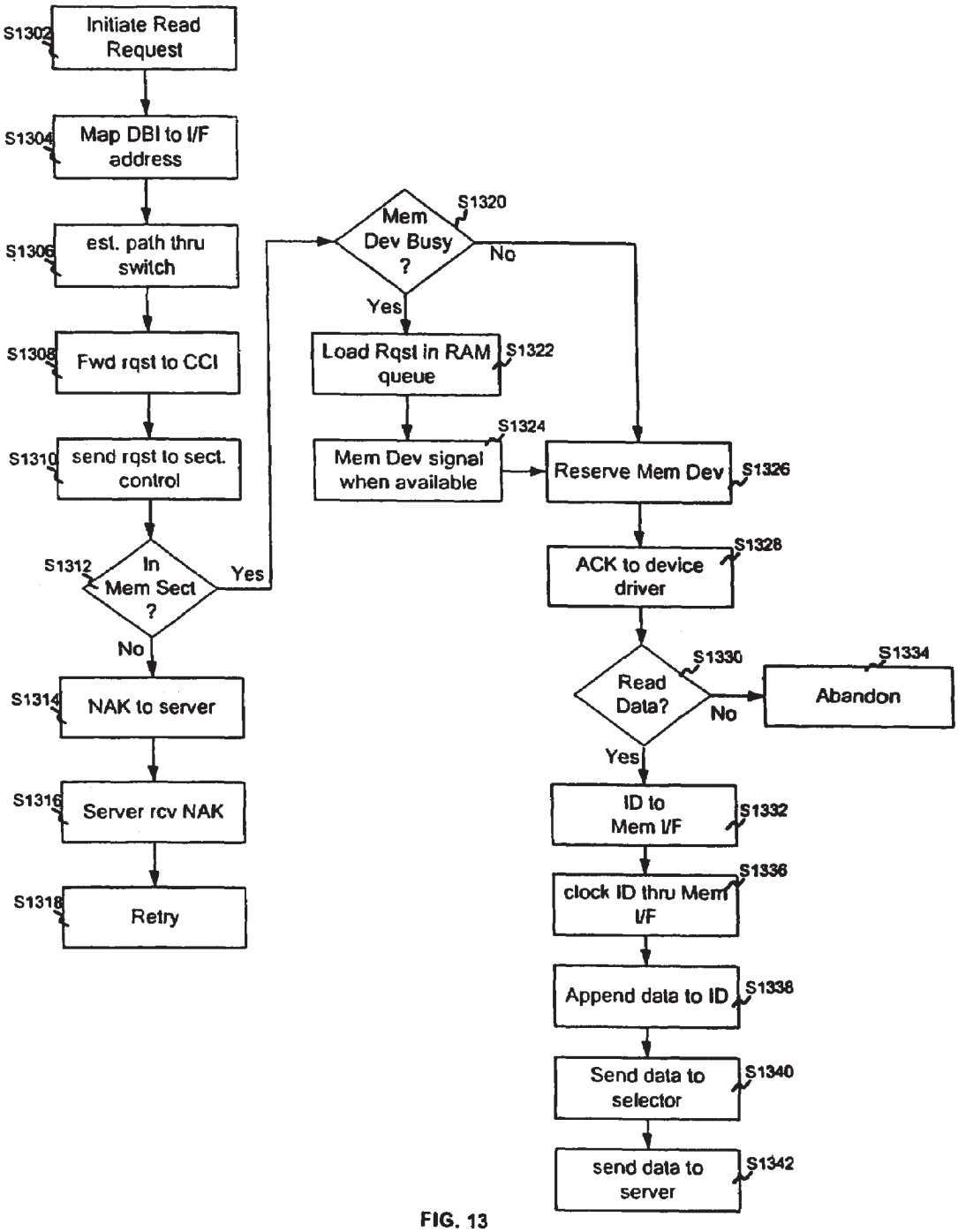
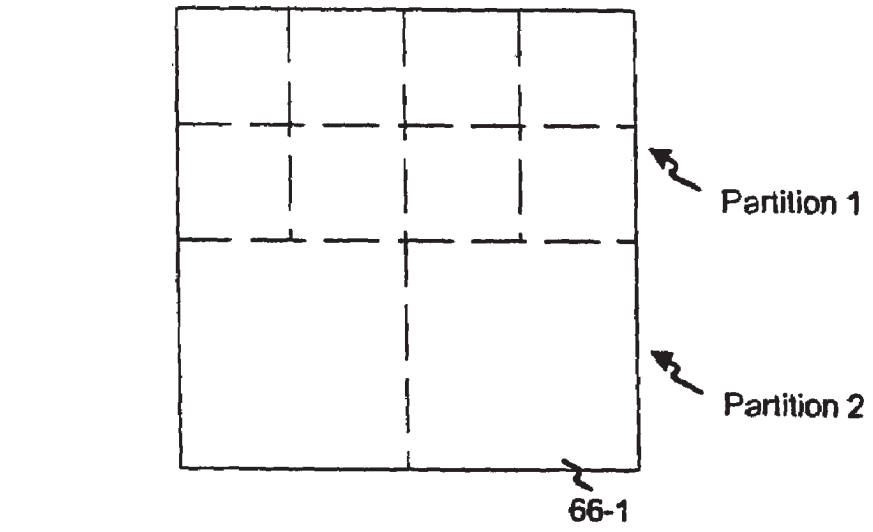

S1302
Initiate Read Request
S1304
Map DBI to I/F address
S1306
est. path thru switch
S1308
Fwd rqst to CCI
S1310
send rqst to sect. control
S1312
In Mem Sect ?
Yes
No
S1314
NAK to server
S1316
Server rcv NAK
S1318
Retry
S1320
Mem Dev Busy ?
Yes
No
S1322
Load Rqst in RAM queue
S1324
Mem Dev signal when available
S1326
Reserve Mem Dev
S1328
ACK to device driver
S1330
Read Data?
Yes
No
S1334
Abandon
S1332
ID to Mem I/F
S1336
clock ID thru Mem I/F
S1338
Append data to ID
S1340
Send data to selector
S1342
send data to server

FIG. 13


Partition 1
Partition 2
66-1

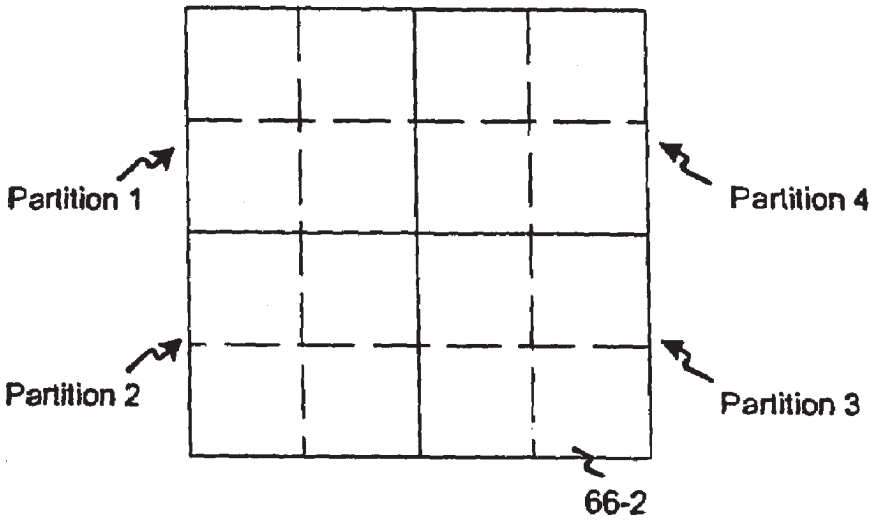

Partition 1
Partition 4
Partition 2
Partition 3
66-2

FIG. 14

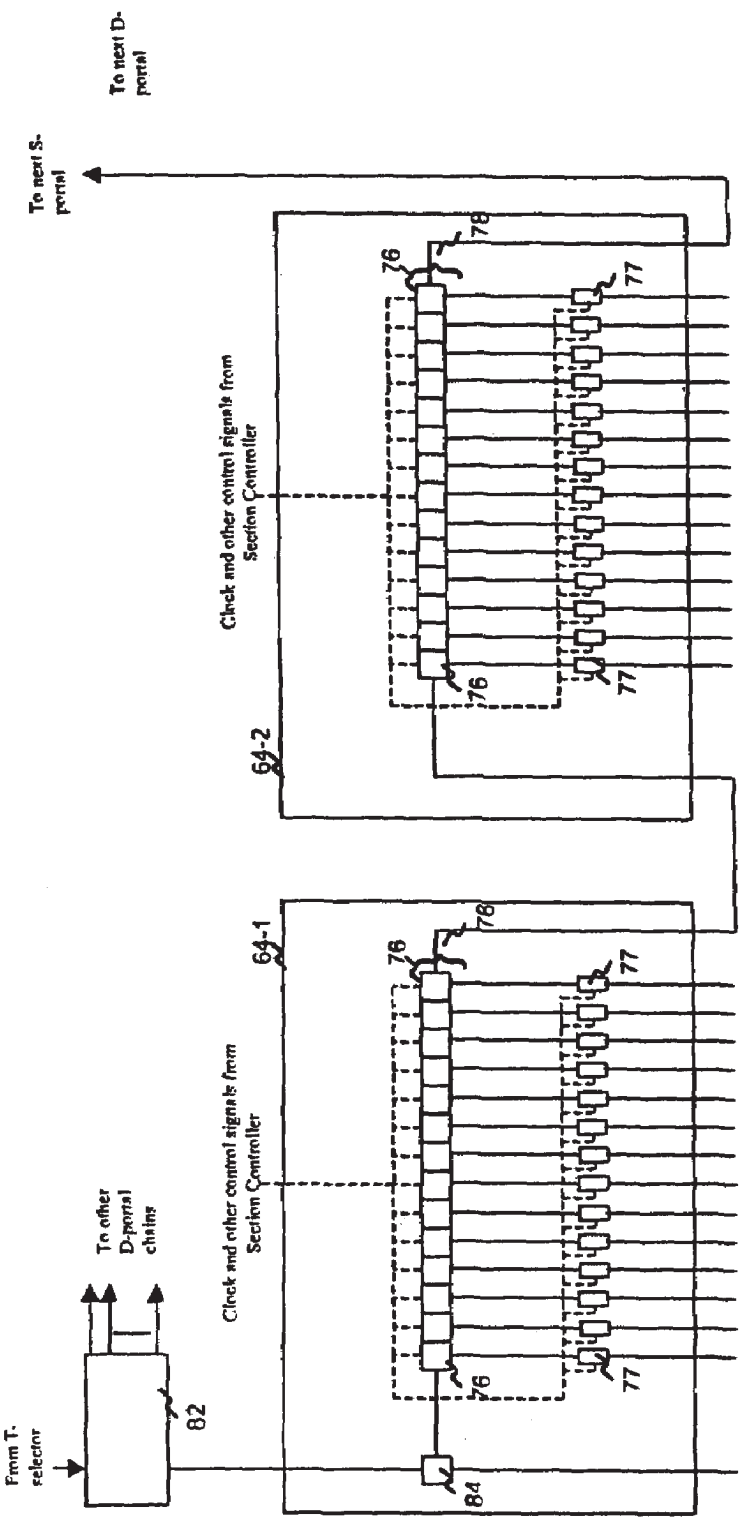

To next D-port
To next S-port
Clock and other control signals from Section Controller
64-2
64-1
76
77
78
82
84
From T-selector
To other D-port chains
To Pins of Memory Device
To Pins of Memory Device

FIG. 15

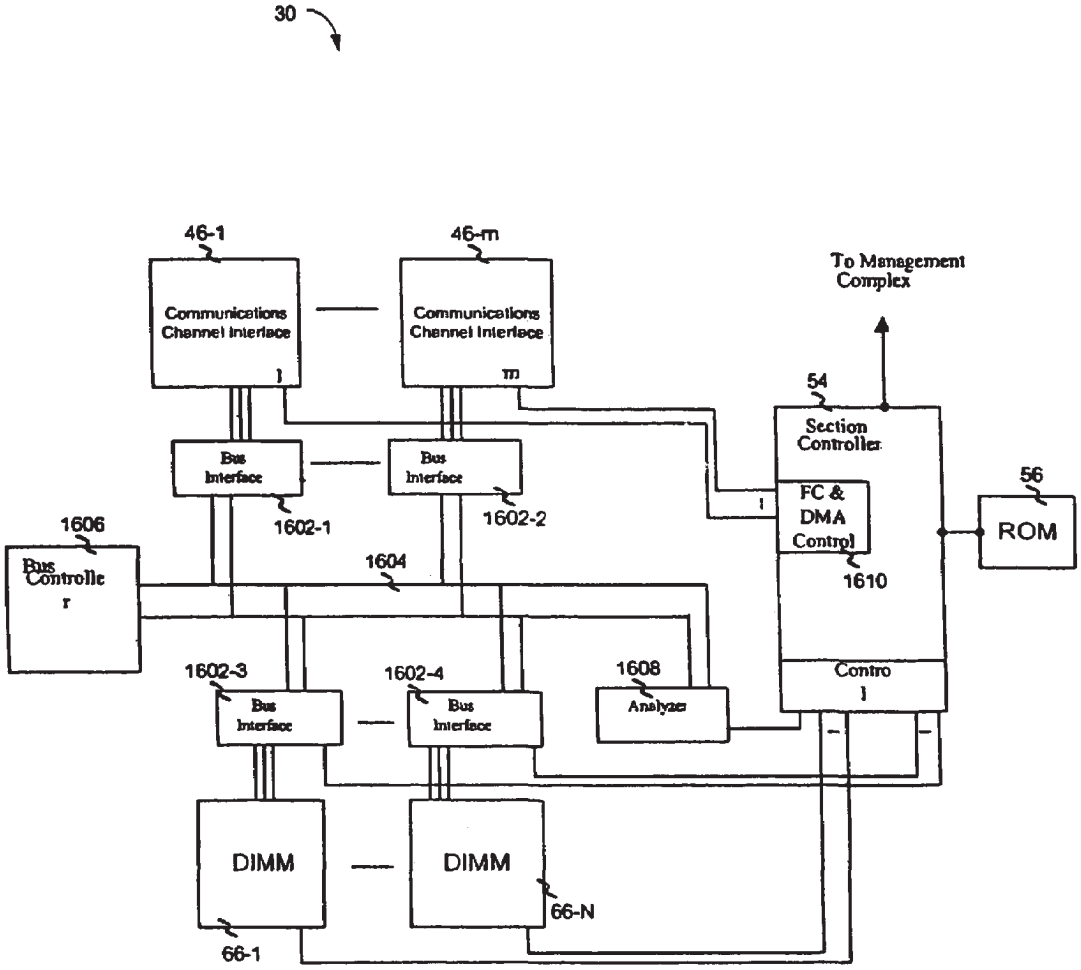

30
46-1
46-m
Communications Channel Interface
1
Communications Channel Interface
m
To Management Complex
54
Section Controller
Bus Interface
Bus Interface
1602-1
1602-2
FC & DMA Control
1610
56
ROM
1606
Bus Controller
1604
1602-3
1602-4
1608
Bus Interface
Bus Interface
Analyzer
Control
DIMM
DIMM
66-N
66-1

FIG. 16

# METHODS AND SYSTEMS FOR A STORAGE SYSTEM WITH A PROGRAM-CONTROLLED SWITCH FOR ROUTING DATA

## RELATED APPLICATIONS

The present application relates to the U.S. patent application Ser. No. 10/284,199 by M. James Bullen, Steven L. Dodd, David J. Herbison, and William T. Lynch, entitled "Methods and Systems for a Storage System," and the U.S. patent application Ser. No. 10/284,268 by M. James Bullen, Steven L. Dodd, David J. Herbison, and William T. Lynch, entitled "Methods and Systems for a Memory Section," both of which are incorporated by reference herein in their entireties.

## BACKGROUND

The present invention relates to data storage, and more particularly, to methods and systems for a high throughput storage device.

A form of on-line transaction processing (OLTP) applications requiring a high number of data block reads or writes are called H-OLTP applications. A large server or mainframe or several servers typically host an H-OLTP application. Typically, these applications involve the use of a real time operating system, a relational database, optical fiber based networking, distributed communications facilities to a user community, and the application itself. Storage solutions for these applications use a combination of mechanical disk drives and cached memory under stored program control. The techniques for the storage management of H-OLTP applications can use redundant file storage algorithms on multiple disk drives, memory cache replications, data coherency algorithms, and/or load balancing.

A brief overview of the storage management technologies of cached disk arrays (CDAs) and solid-state disk storage systems (SSDs) follows.

Cached disk arrays (CDAs) combine disk drives and solid-state memory systems under common program control. The disk drives in CDAs are servo-mechanical devices. Advances in motor technology currently allow the platters of the disk drives to spin at 15,000 revolutions per minute; advanced systems may spin their platters at 18,000 revolutions per minute.

CDAs combine several racks of rotating disks with a common memory cache in an architecture where capacity may be added through the addition of more racks of devices, more cache, or both. CDAs often are used by companies to provide storage services in their mission critical applications, including H-OLTP applications.

The on-board cache of a CDA stores frequently used data because access times for data in cache memory can be short relative to access times for data on the drives. Such high-end storage system devices with rotating media, such as CDAs, include less than ideally desirable characteristics in terms of total throughput and memory cache size.

A solid-state disk (SSD) is a storage device corresponding to the solid-state memory attached to a computer's central processing unit through its internal bus structure. To an external computer (server or mainframe) the SSD appears as a very fast disk drive when it is directly attached to the computer over a fast communications link or network. Operating under stored program control, SSDs store frequently used information like transaction logs, database indices, and specialized data structures integral to the efficient execution of a company's mission critical applications.



It would be desirable for large capacity storage to provide sufficient throughput for high-volume, real-time applications, especially, for example in emerging applications in financial, defense, research, customer management, and homeland security areas.

## SUMMARY

Accordingly, the present invention is directed to methods and systems that address the problems of prior art.

In accordance with the purposes of the invention, as embodied and broadly described herein, methods and systems for an apparatus are provided, including one or more memory sections, and one or more switches. The one or more memory sections include one or more memory devices including storage locations for storing data, and a memory section controller for providing addresses to the memory devices, the addresses identifying storage locations for a memory device, wherein the memory devices use the provided addresses to perform a function selected from the set of reading and writing data to/from the memory devices. The one or more switches for receiving a data request include a data block identifier and switching the data request based on the data block identifier to one or more of the memory sections, the data block identifier identifying a set of storage locations. Additionally, the memory sections to which the data request was switched forward the received data block identifier to its memory section controller which maps the data block identifier to a set of addresses for the storage locations identified by the data block identifier, and provides the set of addresses to one or more of the memory section's memory devices.

The summary and the following detailed description should not restrict the scope of the claimed invention. Both provide examples and explanations to enable others to practice the invention. The accompanying drawings, which form part of the description for carrying out the best mode of the invention, show several embodiments of the invention, and together with the description, explain the principles of the invention.

## BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** is a block diagram of a storage hub environment, in accordance with methods and systems provided;

FIG. **2** is a more detailed block diagram of a storage hub, in accordance with methods and systems provided;

FIG. **3** illustrates a logical architecture for a management complex, in accordance with methods and systems provided;

FIG. **4** is a block diagram of a physical architecture for a management complex, in accordance with methods and systems provided;

FIG. **5** is a block diagram of a exemplary memory section, in accordance with methods and systems provided;

FIG. **6** illustrates a functional diagram of a switch and memory section, in accordance with methods and system consistent with the invention;

FIG. **7** illustrates an alternative functional diagram of a switch and memory section, in accordance with methods and systems provided;

FIG. **8** illustrates a diagram of an alternative exemplary switch, in accordance with methods and systems provided.

FIG. **9** illustrates a diagram of an alternative switch, in accordance with methods and systems provided;

FIG. **10** illustrates an exemplary pipeline shift register, in accordance with methods and systems provided;

FIG. **11** includes a more detailed block diagram of an exemplary embodiment of a memory interface device, in accordance with methods and systems provided;

FIG. **12** illustrates a flow chart for an exemplary writing operation, in accordance with methods and systems provided;

FIG. **13** illustrates a flow chart for an exemplary reading operation, in accordance with methods and systems provided;

FIG. **14** illustrates a logical diagram of partitioned memory devices, in accordance with methods and systems provided;

FIG. **15** illustrates an alternative embodiment of a memory interface devices, in accordance with methods and systems provided; and

FIG. **16** illustrates an alternative memory section, in accordance with methods and systems provided.

## DETAILED DESCRIPTION

Reference will now be made in detail to exemplary embodiments, examples of which are illustrated in the accompanying drawings. Wherever possible, the same reference numbers will be used throughout the drawings to refer to the same or like parts.

FIG. **1** is a block diagram of one embodiment storage hub environment, in accordance with methods and systems provided. As illustrated, the storage hub environment includes a storage hub **10**, servers **12-1** and **12-2**, external management systems **14-1** and **14-2**, a non-volatile storage device **16**, an IP network **18** and a connection to another network **20**. The storage hub **10** may include a large amount of storage (not shown) and stores the data in data blocks. Although the data may be stored in data blocks, any other mechanism for storing the data may be used without departing from the scope of the invention. The non-volatile storage device **16** may be a magnetic storage device, such as a CDA as described above. The non-volatile storage device **16** may be used to store back-up versions of the data stored by the storage hub **10**.

The description below is organized in the following manner. First, a brief overview of the storage hub **10** environment illustrated in FIG. **1** is presented. Then, more detailed descriptions of the components of the storage hub **10** are presented, after which a more detailed description of exemplary methods for writing data to the storage hub, reading data from the storage hub **10**, and a testing operation for the storage hub **10** are presented. Then, exemplary alternatives to these components are presented. It should, however, be understood that these are all exemplary descriptions regarding example methods and systems for implementing the invention. As such, one of skill will recognize that there are other methods and systems that may be used for practicing the invention that is defined by the claims of this application.

The servers **12-1** and **12-2** are, for example, standard commercially available servers or farms of servers that can be connected to internal or external networks (not shown). For example, the servers **12-1** and/or **12-2** may be connected to an internal network such as an Ethernet for receiving requests for the retrieval or storage of information from end users connected to the network. Alternatively, the servers **12-1** and/or **12-2** could be connected to external networks, such as the Internet, for receiving requests for retrieval or storage of information from end users connected to the external network. Further, although two servers **12-1** and **12-2** are illustrated, the storage hub **10** may be connected to any number of servers **12**.

When an application being executed by the server **12** requires data, the server **12** determines if the storage hub **10** stores the data. The servers **12** may store a record showing whether the data their applications require is on the storage hub **10**. The server **12** then sends a data request to the storage hub **10** requesting the data. The storage hub **10** reads the data from the location in which it is stored and sends it to the server requesting the data **12-1** or **12-2**. The server may run different types of applications and database management systems that may require data from the storage hub **10**. Examples of typical applications include, by way of example only, billing systems, customer relationship management systems, reservations systems, ordering systems, security systems, etc. Examples of database management systems include ORACLE, DB2, Sybase, Informix, etc.

Additionally, the storage hub **10** may receive a request from a server **12-1** or **12-2** to store data. Thereafter, the storage hub **10** preferably provides the server **12** with either an acknowledgement that the write occurred (i.e., the storage of the data) or a failure message. Such messages could include, for example, an acknowledgement that the data block was safely stored on both the storage (not shown) in the storage hub **10** and on the CDA **16** when a CDA **16** is used as backup for the storage hub **10**, an acknowledgement that the data block is safely stored in the storage hub's **10** storage (not shown), no acknowledgement of any sort, or a failure message.

The external management system **14** may be directly connected to the storage hub **10**, such as external management system **14-2**. Or, the external management system **14** may be connected to the storage hub **10** via a network, such as external management systems **14-1** that is connected to the storage hub **10** via network **18**. Network **18** may be any type of network, such as an internal Ethernet network, an IP network, or the Internet. Although FIG. **1** illustrates both external management systems **14-1** and **14-2** connected to the storage hub **10**, in other implementations there may be only one or any number of external management systems, or an external management system **14** need not be included. For example, in other implementations it may be desirable to have 3 or more external management systems. Additionally, the external management system may be a computer running proprietary or commercially available software, such as, for example, HP Openview. The storage hub **10** may provide surveillance and administration information to the external management system **14**, such as the status and location of stored data blocks.

FIG. **2** illustrates a more detailed block diagram of the storage hub **10**, in accordance with methods and systems provided. As illustrated, the storage hub **10** includes a switch or switches **22-1** and **22-2**, a management complex **26**, and memory sections **30-1** thru **30**-n. In this embodiment, both switches **22-1** and **22-2** may be active or one of the switches may be active while the other is a redundant switch for use in the event the active switch suffers a problem. Although FIG. **2** illustrates two switches, the storage hub **12** may include only one switch or any number of switches.

In FIG. **2**, server **12-2** connects to the storage hub **10** via a network **20** thru an input/output (I/O) controller **24**. The network may be any type of internal or external network, such as an Ethernet network or the Internet. The I/O controller **24** preferably is an appropriate I/O controller for connecting to the particular network **20**. Preferably, the I/O controller **24** converts signals between a native protocol of the network **20** and a local protocol used by the storage hub **10**. Potential protocols include, but are not limited to, Telecommunications Control Protocol/Internet Protocol (TCP/IP), System Network Architecture (SNA)-based protocols, Serial Communications Control Interface (SCCI), Intelligent Serial Communications Interface (ISCI), Fibre Channel, Infiniband, and other third generation input/output (3GIO) protocols

The memory sections **30** preferably include the storage for the storage hub **10** along with other hardware for accessing the storage. As used herein, the term "memory section" refers to any subsystem including one or more memory devices that may be used for storing information. This architecture is applicable to any device that can store data. Thus, when the storage hub **10** receives a request to store data, the data is forwarded to a memory section **30**, which stores the data. Likewise, when a request for data is received by the storage hub **10**, the request is directed to the memory section **30** storing the requested information. The memory section **30** then reads the requested data, after which it is sent to the server **12** requesting the data. More detailed descriptions of exemplary memory sections **30** and their operations are presented below.

The management complex **26** of the storage hub **10** performs management-type functions for the storage hub **10** and connects the storage hub **10** with the external management system **14**. As used herein the term "management complex" refers to any software and/or hardware for performing management of the storage hub **10**. A more detailed description of the management complex **26** is presented below.

The I/O Controller **24** and switches **22-1** and **22-2** are preferably under common management control by the management complex **26** to allow data blocks to be sent to and received from the storage hub in the native protocol of the network **20**.

Each server **12-1** and **12-2** preferably includes a device driver **28-1** and **28-2**, respectively. The device driver **28** is a program running in software on a server that permits applications on the server to cause data to be read from or written to (i.e., stored in) the storage hub **10**. When a server **12** receives a request to read or write data, the device driver **28** of the server **12** forwards the request to the switch in the storage hub **10**. The device driver **28** may be, for example, a standard device driver supplied as part of server-resident software, or it may be, for example, proprietary software supplied by a vendor of storage devices. Additionally, in some applications, the device driver **28** may be independent of any application resident on the server.

The switches **22-1** and **22-2** are connected to the server **12-1**, the I/O controller **24**, the CDA **16**, the memory sections **30-1** thru **30**-n, and each other via an industry standard communications interface protocol. These communications interface protocols may be, for example, Fibre Channel, Asynchronous Transfer Mode (ATM), Ethernet, Fiber Distributed Data Interface (FDDI) a Systems Network Architecture (SNA) interface, or X.25. Any type of physical connection, e.g., copper or fiber optic cables, may be used for connecting these various components. The management complex **26** is preferably connected to the switches **22**, memory sections **30-1** thru **30**-n, the I/O controller **26**, and the external management system **14** via gigabit Ethernet connections. Although these are preferable connections, persons skilled in the art will recognize there are numerous other protocols and physical media that may be used to connect these devices. Further, the memory sections **30** may simultaneously support multiple protocols and physical media for connecting these devices.

The switches **22** may be any type of switch using any type of switch fabric, such as, for example, a time division multiplexed fabric or a space division multiplexed fabric. As used herein, the term "switch fabric" the physical interconnection architecture that directs data from an incoming interface to an outgoing interface. For example, the switches **22** may be a Fibre Channel switch, an ATM switch, a switched fast Ethernet switch, a switched FDDI switch, or any other type of switch. The switches **22** may also include a controller (not shown) for controlling the switch.

For write operations, the data block, in addition to being written to the memory sections **30** of the storage hub **10**, may also be written to the cached disk array **16** or another storage hub (not shown). After the data is written, the storage hub **10** may send an acknowledgement to the device driver **28** of the server **12** depending upon the configuration management parameters in the management complex **26**. Examples of configuration management parameters are status parameters, write-acknowledgement parameters, routing parameters, reporting interval parameters, and the current date and time.

For a read data block request and at the request of the device driver **28** requesting the data block, the switches **22** direct the request to the appropriate memory section **30**, which retrieves the data block and transmits it through a switch **22** to the device driver **28** of the server **12** from which the request originated.

During read and write data block operations and depending on the configuration management parameters in the management complex **26**, the memory section **30** gathers administrative data that it sends to the management complex **26**. The management complex **26** then makes this data available to the external management system **14**.

Additionally, the management complex **26** may gather and provide the external management system **14** with surveillance and administrative information. Surveillance information may include, for example, memory section heartbeats (i.e., a signal that shows that the memory section can still communicate), alarms, and acknowledgement of alarms. Administration information may include, for example, statistics about data read and written, statistics about the number of active memory sections, statistics about memory section availability, and reports that present the preceding information to the external management system.

The external management system **14** may also provide the management complex **26** with configuration management data. This configuration management information may include, for example, valid communications network addresses, a period for heartbeat intervals, data block sizes, and command sets.

The storage hub **10** may also perform bit-level error recovery using standard means available in the industry. For example, error correction codes (ECC), also referred to as error detection and correction (EDAC) codes, using circuitry and/or software may be used to test data for its accuracy. These codes and techniques include parity bit or cyclic redundancy checks, using multiple parity bits in order to detect and correct errors, or more advanced techniques (e.g., Reed-Solomon codes) to detect multiple errors. Further, each memory section **30** of the storage hub **10** may include its own error correction scheme.

The following provides a more detailed description of the components of the storage hub **10** illustrated in FIG. **2**: the management complex **26**, the switches **22**, and the memory sections **30**. After which, more detailed descriptions of exemplary reading, writing, and testing operations are presented. Then, alternative exemplary embodiments of the memory sections **30** are provided along with exemplary characteristics of the storage hub **10** and its components.

## Management Complex

FIG. **3** illustrates a logical architecture for a management complex **26**, in accordance with methods and systems provided. As illustrated, the management complex **26** may include functions that manage administrative processor **32**

and functions that manage control processor **34**. These management functions can include one or more central processing units (CPUs) for executing their respective processes. Additionally, the management complex **26** may use one or more application program interfaces (APIs) for communications between these functions.

FIG. **4** is a block diagram of a physical architecture for a management complex **26**, in accordance with methods and systems provided. As illustrated, the management complex includes one or more control processors **34-1** thru **34**-n, a shared memory **36**, one or more administration processors **32-1** thru **32**-m, a storage device **38**, and a communications network **40**. As discussed above, the control processors **34** may include one or more central processing units (CPUs). These control CPUs **34-1** thru **34**-n interface with the shared memory **36**. The communications network **40** may be an internal network and may use any type of communications protocol, such as Gigabit Ethernet.

One or more of the control processor (e.g., **34-1** thru **34**-m) may function as the master(s), while remaining control processors (e.g., **34**-(m+1) thru **34**-n) may be kept in a hot standby mode, so that they can be quickly switched to in the event one of the master control processor (e.g., **34-1**) fail.

The control CPU's **34** may be attached to a communications network, such as a Gigabit Ethernet network, and be directly attached to the magnetic storage device **38**.

The administrative processors **32** each may include a memory (not shown) and also be attached to the communications network **40**. These administration processors may also connect to the magnetic storage device **38**. The magnetic storage device **38** stores various control and administrative information from the control processors **34** and administration processors **32**. The magnetic storage device **38** may be any type of magnetic storage device, such as, for example, servo-mechanical disc drives. In other embodiments, the storage device **38** need not be included.

The control processors **34** perform configuration management functions for the memory sections **30**, I/O controllers **24**, switches **22**, and the device drivers **28** of the servers **12**. As used herein, the term "configuration" is a broad term that encompasses the various possible operating states of each component of the storage hub. As used herein, an "operating state" refers to a possible way in which the storage hub or one of its components operates as defined by parameter values. These parameter values, for example, may be set by a user of the storage hub, such as, for example, a system administrator, through, for example, an external management system **14**. Operating states may include, for example, how often a component (e.g., a memory section **30**) sends performance statistics to the management complex **26**, the list of events that causes a component (e.g., a memory section, etc.) to report an alarm, and/or the type of alarm reported (e.g., catastrophic failure of component, minor fault with component, etc.). Further, as used herein, the term "configuration management" means the understanding of the current operating states of the storage hub's components and the capability to react to changes in the states of those components as defined by software running in the control processors **34**. For example, the control processors **34** may control in real time the number of active memory sections **30** in the storage hub **10**, the switches **22**, and the device drivers **28** of the servers **12**, if any, and any external servers **22** connected to the storage hub.

The software in the control processors **34** may also be capable of bringing new memory sections into service and taking memory sections out of service independently of other functions that the management complex performs and without materially affecting the operation of other memory sections **30** or adversely affecting the overall performance of the storage hub. The instructions to perform this function are carried from the control process **34** to the switches **22** and may be carried to the device drivers **28** in the servers **12**. In the case that new capacity is added to the storage hub **10**, then it is possible to bring new memory sections **30** into service with the software capability in the control processors **32**. In the case that a memory section **30** has failed, then the faulty memory section **30** may be replaced and a new one brought into service. A further description of fault management follows.

The control processors **34** may also, for example, be able to perform fault management for the storage hub **10**. The term "fault management" as used herein means attempting to detect faults and take corrective action in response to the detection of a fault. For example, the control processors may recognize an operational failure of a memory section **30** or part of a memory section **30** and re-map data to working memory sections **30**. Then, the control processors **34** may communicate this re-mapping to the external management system **14** and the device drivers **28** running on servers **12** attached to the storage hub **10**.

The control processors **34** may also manage "bad-block" re-mapping functions when a memory section fails **30** and the writing of data to the magnetic storage device **38** in the event of power failures. Bad block remapping is a process wherein data blocks discovered by the section controller **54** or management complex **26** to be in a damaged memory device are, if possible, recovered.

For example, if the control processors **34** discover that block **65,000** in memory section **30-2** does not read correctly, the control processor **34** may decide to remap block **65,000** in memory section **30-2** to block location **1,9999,998** in memory section **30-2**. The control processor **34** may then direct the CDA **16** to read the data block and cause it to be written in location **1,9999,998** in memory section **30-2**. Once completed, the control processor **34** may inform the switches **22** and memory section **30-2** that block **65,000** may now be read from location **1,9999,998**.

As another example of bad block remapping, if for example only one memory device on a memory section is faulty, a control processor **34** in the management complex **26** may inform the section controller **54** about the bad device, determine where the data on the faulty memory device is backed-up (e.g., CDA **16**), and direct the backed-up data to be loaded into a replacement memory device on the same memory section or on a different memory section. In the latter case, the management complex also informs the switch about the data being relocated to a new memory section.

As yet another example, in the event the control processors **34** determine that a memory section **30** is faulty, the control processors **34** may direct that the entire memory section **30** is taken out of service and that a replacement memory section takes its place. To accomplish this, the control processors **34** may, for example, direct the CDA **16** to transfer a back-up version of the data for the faulty memory section **30** to another memory section **30**-N that may be, for example, a spare memory section **30** for use in the event a memory section **30** goes bad. The new memory section **30**-N then may operate as though it were the now faulty memory section **30**. The control processors **34** may then communicate this information to the various device drivers **28** and the external management system **14**.

The control processors **34** may also provide the memory sections **30**, the switch controller(s) **202**, and the I/O Controllers **24** with updated and new software. For example, if software used by the memory sections **30** or the switches **22**

become corrupted and/or fails, the control processors **34** can load backup copies of current or previous versions of a software image from its storage **38**. A software image is a binary object code that may be run directly by a computer. The software image for the control processor **34** in one embodiment is stored on the magnetic storage **38**. Further, the control processors **34** may also control the loading of a data block from the CDA **16** into the memory sections **30** and visa versa.

In addition, the control processors **34** may receive information such as, for example, the time a component sent an alarm or the total elapsed time a component was in alarm from the components of the storage hub **10** over a communications interface.

The control processors **34** also may allow the administration processors **32** to gather data on parameters like the number of active memory sections **30**, the total throughput of the storage hub **10** over time, the size of memory section queues, etc., that comprise the operating state of the storage hub. (Note that memory section queues are those queues in the section controller that comprise the list of yet-to-be completed read operations and write operations). In addition, the control processors **34** are responsible for monitoring their own operational status, such as, for example determining which control processor is active as Master, which are on standby, and which, if any, are not operational. Additionally, the control processors **34** may monitor the Storage Hub's environment for extreme temperatures or humidity, etc.

The control processors **34** may also store a copy of the software (i.e., a software image) run by the switches **22**. A more thorough description of the switches **22** is present below. If the need arises, it can reload the switch software to one or more of the switches. As discussed below, the switch **22** may include one or more switch controllers (not shown) for executing this software to control the switch **22**. In the event the switch **22** uses multiple controllers configured in a master-slave architecture, the control processor **34** may determine which of the controllers in the switch is(are) the master(s) and which is(are) the slave(s).

Additionally, the control processors **34** may determine the status (active, idle, out-of-service) of ports (not shown) on the switch **22**, whether the ports are used to connect to servers **12** or to memory sections **30**. The control processors **34** may also provide configuration management data to the switches **22**. Examples of configuration management data include the date, the time, a routing algorithm to use, an interval for a status check, the identity of active server ports, etc. Further, the control processors **34** may instruct the switch to use different "hunt" algorithms to find idle ports that may be used in establishing connections. These algorithms may be included in the software executed by the switch controller, examples of which include rotary hunt, skip route, and least-used.

The administration processors **32** preferably collect information and statistics from the I/O controllers **24**, memory sections **30**, switches **22**, and the control processors **34**. The information and statistics collected may include information for generating statistical reports, telemetry data, and other alarms and administrative data. The administration processors **32** provide this information to the external management system **14** using a protocol, such as, for example, TCP/IP or any other suitable protocol. The administration processors **32** may collect data on such parameters from the device drivers **28**, the switches **22**, and the memory sections **30**.

Users of the external management system, such as for example, a system administrator, may request a change in the configuration management parameters of the storage system **10**. This change may, for example represent the addition of new memory sections **30**. Users of the external management system **14**, such as for example, a system administrator, may also request the administration processors **36** to collect statistical data from a storage area network environment (a set of storage devices connected by a network dedicated solely to the storage devices) including one or more storage hubs **10**, a network area storage environment (a set of storage devices connected by a network shared with other traffic) including one or more storage hubs **10**, and other external systems. For example, this statistical data may include the total incoming requests from each storage environment or from a particular server.

The administration processors **32** may execute a database program such that the administration data is stored in a standard database, which can then be used to provide the information to system administrators of the storage hub **10** in reports and graphs on a computer screen or on paper. For example, the system administrators of the storage hub may use an external management system **14** to gain access to this information. Alternatively, the system administrators of the storage hub **10** may access this information directly through an interface to the administration processors. Like the control processors **34**, the administration processors **36** can monitor themselves and communicate their own operational state to the control processor **34**, which determines which administration processors **34** are active or inactive for any reason.

The management complex **26** may instruct a non-volatile storage device to load data into one or more of the memory sections **30**. For example, as illustrated in FIG. **2**, the storage hub **10** may be connected to a non-volatile storage device such as a CDA **16**. The management complex **26** may then be able to send instructions to the CDA **16**, switches **22**, and memory sections **30** to perform various activities. These activities may include the loading of the memory sections **30** from the non-volatile storage device **16** when the storage hub **10** is powered, when the storage hub **10** has been restarted after, for example, having lost power in an outage, as a result of administrative changes to the configuration of the storage hub **10**, as a result of the failure of a memory section **30**, or as a result of a user-initiated command.

Although the above presents numerous management and control functions capable of being performed by the management complex **26**, it should be understood that the management complex **26** may perform all, a subset, or even entirely different functions. Additionally, although FIGS. **3** and **4** illustrate an exemplary management complex being implemented using separate administration processors **32** and control processors **34**, a management complex may be implemented using only one, none, or any number of processors.

## Memory Section

FIG. **5** is a block diagram of an exemplary memory section **30**, in accordance with methods and systems provided. As illustrated, the memory section **30** may include a switch portal ("S-portal") **42**, a section controller **54**, a read only memory (ROM) **56**, a temporary storage **58**, a temporary storage interface device **60**, a temporary store selector ("T-selector") **62**, a synchronizer **68**, one or more memory interface devices **64-1** thru **64-8**, and one or more memory devices **66-1** to **66**-n.

The memory devices **66** may be any type of memory devices, such as, for example, dynamic random access memory (DRAMs), synchronous dynamic random access memory (SDRAMs), Rambus DRAMs (RDRAMs), magnetic random access memory, resistance random access memory, ferroelectric random access memory, polymer random access memory, chalcogenide random access memory,

single in-line memory module (SIMMs), dual in-line memory module (DIMMs), rambus in-line memory modules (RIMMs), rotating media, etc. Although, the term memory interface device is used herein, it should be understood that this term should be interpreted broadly to include any type of access device capable of accessing information stored in a memory device. A more detailed description of exemplary memory interface devices is presented below.

The section controller **54** may, for example, include a microprocessor **51**, internal memory **52**, a management complex interface(s) **53**, memory device control circuitry **55**, communications channel interface (CCI) control circuitry **57**, test circuitry **59**, timing circuitry **61**, and a Header/test interface **63**. The microprocessor **51** may be, for example, a chip such as the Motorola G2 executing appropriate software. The internal memory **52** may be, for example, **32** megabytes of useable SRAM for program and data storage. This internal memory **52** may be included in the microprocessor **51**, such as for example in a Motorola G2. The management complex interface **53** may, for example, be a TCP/IP running over gigabit Ethernet interface that the section controller **54** may use in communicating with the management complex **26**. The header/test interface **63** may be an appropriate interface for providing information from the section controller **54** to the memory interface devices **64**.

The section controller **54** further may access bootstrap read only memory **56** that may be used by it when power is first applied. This bootstrap read only memory **56** may, for example, contain a small software image that allows the section controller **54** to communicate with the control processors **34** to obtain the current software image via the management interface **53**. The section controller **54** may further include CCI control circuitry **57** that may, for example contain a direct memory address circuitry for use in the management of the communications channel interface **46**.

The section controller **54** may also include memory device control circuitry **55** for controlling the memory devices **66**. This memory device control circuitry **55** may, for example include a memory latching circuit for controlling the state of the memory devices **66** through the binary states of the memory latch. A further description of memory latching is presented below. The section controller **54** may further include test circuitry **59** for testing the memory section **30**. A more detailed description of an exemplary test procedure is presented below.

Additionally, the section controller may include a header/test interface **63** for providing header type information (e.g., a data block identifier, destination address, etc.) and testing the memory section **30**. Also, the section controller **54** may include timing circuitry **61** that may provide master and slave clock signal and other timing signals, such as start and stop read or write signals, etc. for use by the memory section.

The S-portal **42** may include a selector **44** and a communications channel interface **46**. The communications channel interface **46** provides the interface for connecting the memory section **30** with the one or more servers **12** via the switches **22**. This connection may be, for example, via one or more fiber optic or copper cables. The selector **44** may include circuitry for connecting the communications channel interface **46** with the one or more memory interface devices **64**, such that the selector **44** may connect any memory interface device **64** with any I/O port of the communications channel interface **46**. The section controller **54** via the CCI circuitry **57** may provide control signals to the selector **44** regarding how the selector should connect the memory interface devices **64** and communication channel interface **46**. Additionally, the selector **44** may be directed to send data, such as, for example, test data, from a memory interface device **64** to the section controller **54** via the CCI circuitry **57**.

The communications channel interface **46** can use any type of protocol, such as, for example, any standard channel interface protocol and the selector **44** may or may not be included. Exemplary standard channel interface protocols include Fibre Channel, System Network Architecture-based protocols, Intelligent Serial communications Control Interface, and other third generation input/output (3GIO) protocols.

The temporary storage interface device **60** is any type of device capable of accessing the temporary storage device **58**. For example, the temporary storage interface device **60** may include one or more shift register arrays (not shown), including a plurality of shift registers interconnected in series, such that the data may be serially clocked through the shift register arrays. For a further description of shift register arrays and their use in accessing storage media such as memory devices, see the patent application by William T. Lynch and David J. Herbison, entitled "Methods and Systems for Improved Memory Access," filed on the same day as this application, which is incorporated by reference herein in its entirety.

The temporary storage **58** may be any type of memory device, such as a DRAM, SDRAM, SIMM, DIMM, a disk drive etc. The T-selector **62** may be any type of selector for selecting between a plurality of inputs.

The storage hub **10** may use a synchronizer **68** in embodiments where the temporary storage interface device **60** includes shift register arrays. In such an embodiment, the synchronizer **68** may, for example, accept data to be stored in the memory section **30** and use phase lock loop circuitry to extract a clock frequency from the incoming data stream. A temporary storage interface device **60** including shift register arrays may then use this clock signal to shift the data in writing data to the temporary storage device **58**. This clock signal may be used, for example, to compensate for possible differences in either the phase or frequency of the incoming data from the memory section's system clock. When data is shifted out of the temporary storage interface device **60** for storage in the memory devices **66**, the system clock for the memory section is preferably used to shift the data.

The section controller **54** may be capable of detecting faults in the memory section **30**. For example, the section controller **54** may detect errors in the hardware or protocol used by the communications channel interface **42** through the communications channel interface circuit **57**. Additionally, the section controller **54** may, for example, detect errors in the memory interface device **64** through the use of the Header/Test interface **63**. Further, if the memory devices **66** include circuitry for detecting and/or correcting faults, such as, for example, electronic error correction circuitry (e.g, DIMMs), the memory devices **66** may communicate detected faults to the section controller **54** through the memory control **55**. In the event the section controller **54** detects a fault, the section controller **54** may transmit information regarding the fault (e.g., time, component, type of fault) through the management interface **53** to the management complex **26**.

The section controller **54** may also include an interface available for an external system (not shown) that permits the external system to obtain information about the section controller **54** through interaction with the microprocessor **51**. This interface may, for example support a keyboard and display for direct diagnostic observations. The external system interface (not shown) may also, for example support an interface to a personal computer or similar system for direct diagnostic observations. The external system, not shown, may