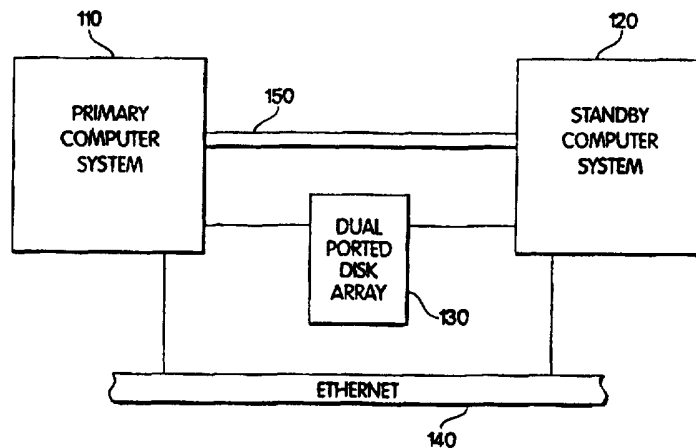




INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : G06F 11/20, 11/14	A3	(11) International Publication Number: WO 97/22046 (43) International Publication Date: 19 June 1997 (19.06.97)
<p>(21) International Application Number: PCT/US96/18980</p> <p>(22) International Filing Date: 27 November 1996 (27.11.96)</p> <p>(30) Priority Data: 08/564,023 29 November 1995 (29.11.95) US</p> <p>(71) Applicant: SEQUOIA SYSTEMS, INC. [US/US]; 400 Nicker- son Road, Marlborough, MA 01752 (US).</p> <p>(72) Inventor: STIFFLER, Jack, J.; 34 Lake Shore Drive, Hopkin- ton, MA 01748 (US).</p> <p>(74) Agent: GORDON, Peter, J.; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210 (US).</p>	<p>(81) Designated States: DE, GB, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).</p> <p>Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i></p> <p>(88) Date of publication of the international search report: 25 September 1997 (25.09.97)</p>	

(54) Title: REMOTE CHECKPOINT MEMORY SYSTEM AND PROTOCOL FOR FAULT-TOLERANT COMPUTER SYSTEM



(57) Abstract

A mechanism for maintaining a consistent, periodically updated state in main memory without constraining normal computer operation is provided, thereby enabling a computer system to recover from faults without loss of data or processing continuity. In this invention, a first computer includes a processor and input/output elements connected to a main memory subsystem including a primary element. A second computer has a remote checkpoint memory element, which may include one or more buffer memories and a shadowy memory, which is connected to the main memory subsystem of the first computer. During normal processing, an image of data written to the primary memory element is captured by the remote checkpoint memory element. When a new checkpoint is desired (thereby establishing a consistent state in main memory to which all executing applications can safely return following a fault), the data previously captured is used to establish a new checkpointed state in the second computer. In case of failure of the first computer, the second computer can be restarted to operate from the last checkpoint established for the first computer. This structure and protocol can guarantee a consistent state in main memory, thus enabling fault-tolerant operation.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 96/18980

A. CLASSIFICATION OF SUBJECT MATTER
 IPC 6 G06F11/20 G06F11/14

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
 IPC 6 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	EP 0 441 087 A (IBM) 14 August 1991 see abstract see page 4, line 52 - line 57 see page 5, line 49 - page 6, line 30 see page 7, line 8 - line 17 see page 9, line 32 - line 40 see figures 1-3 ---	1-4,7 11,15
Y	DE 41 36 729 A (MITSUBISHI ELECTRIC CORP) 21 May 1992 see abstract see column 10, line 45 - line 59 see column 15, line 9 - column 16, line 19 --- -/--	11,15

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *&* document member of the same patent family

Date of the actual completion of the international search <p style="text-align: center; font-size: 1.2em;">18 July 1997</p>	Date of mailing of the international search report <p style="text-align: center; font-size: 1.2em;">14.08.97</p>
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax (+ 31-70) 340-3016	Authorized officer <p style="text-align: center; font-size: 1.2em;">Masche, C</p>

INTERNATIONAL SEARCH REPORT

Inter. Application No
PCT/US 96/18980

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 4 905 196 A (KIRRMANN HUBERT) 27 February 1990 cited in the application see column 7, line 60 - column 8, line 61 see column 9, line 32 - line 64 see column 10, line 28 - line 45 see figures 4,5,7,8 ---	1,3,11, 15
X A	EP 0 299 511 A (FUJITSU LTD) 18 January 1989 see abstract see column 5, line 36 - column 7, line 9 ---	1 2-4,7, 11,12
X Y	FR 2 606 184 A (THOMSON CSF) 6 May 1988 see page 3, line 5 - line 19 see page 7, line 4 - line 24 see page 8, line 20 - line 22 see page 16, line 30 - page 17, line 2 see figure 1 ---	8,16 9
Y	US 4 228 496 A (KATZMAN JAMES A ET AL) 14 October 1980 see column 13, line 38 - line 54 see column 14, line 7 - line 11 see column 33, line 1 - line 24 see column 34, line 6 - line 50 see figure 1 ---	9
X A	FR 2 028 517 A (PLESSEY BTR LIMITED) 9 October 1970 see page 4, line 36 - page 5, line 2 see page 5, line 36 - page 6, line 6 see page 15, line 22 - line 34 see page 19, line 16 - line 24 see figure 1 ---	8,16 9,10
A	PROCEEDINGS OF THE REAL TIME SYSTEMS SYMPOSIUM, SAN ANTONIO, DEC. 4 - 6, 1991, no. SYMP. 12, 4 December 1991, INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, pages 2-11, XP000337134 MORIKAZU TAKEGAKI ET AL: "THE DIFFUSION MODEL BASED TASK REMAPPING FOR DISTRIBUTED REAL-TIME SYSTEMS" see page 5, left-hand column, line 17 - line 21 see page 6, left-hand column, line 33 - right-hand column, line 5 see figure 4 ---	8,10,16
	-/--	

2

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 96/18980

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	COMPUTER, vol. 21, no. 2, February 1988, pages 37-45, XP000048793 BERNSTEIN P A: "SEQUOIA: A FAULT-TOLERANT TIGHTLY COUPLED MULTIPROCESSOR FOR TRANSACTION PROCESSING" see page 43, column 1, line 48 - line 58 ---	17
A	IBM TECHNICAL DISCLOSURE BULLETIN, vol. 36, no. 8, 1 August 1993, pages 161-165, XP000390179 "EFFICIENT CACHE ACCESS THROUGH STORE COMPRESSION" see page 162, line 1 - line 24 -----	17

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 96/ 18980

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see annexed sheet

1. As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- The additional search fees were accompanied by the applicant's protest.
- No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/210

1. Claims 1-7,11-15: Terminating the task of the second computer if the first computer fails, solving the objectively determined problem of performance, if the second computer has to take over the workload of the failed first.
2. Claims 8-10,16: A plurality of computers, one of which is a spare not processing a task, connected in a logical ring; redistribution of the tasks among the remaining computers including a spare, if one computer fails, solving the objective problem of optimal redistribution of tasks among computers after a processing fault without breaking off a task.
3. Claim 17: Means for queuing input/output events between checkpoints and flushing this queued events to the first memory at a checkpoint, solving the objective problem of recoverability of input/output operations after a processing fault.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 96/18980

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0441087 A	14-08-91	DE 69021712 D	21-09-95
		DE 69021712 T	18-04-96
		JP 2505928 B	12-06-96
		JP 4213736 A	04-08-92
		US 5235700 A	10-08-93
DE 4136729 A	21-05-92	JP 5006308 A	14-01-93
		DE 4143452 A	02-09-93
		GB 2251103 A,B	24-06-92
		GB 2277821 A,B	09-11-94
US 4905196 A	27-02-90	DE 3566314 A	22-12-88
		EP 0163096 A	04-12-85
		EP 0254247 A	27-01-88
		JP 6019722 B	16-03-94
		JP 60235246 A	21-11-85
EP 0299511 A	18-01-89	JP 1019438 A	23-01-89
		JP 2514208 B	10-07-96
		CA 1310430 A	17-11-92
		CN 1030653 A,B	25-01-89
		DE 3855031 D	04-04-96
		DE 3855031 T	05-09-96
		US 5123099 A	16-06-92
FR 2606184 A	06-05-88	DE 3781873 A	29-10-92
		EP 0272165 A	22-06-88
		JP 63141139 A	13-06-88
		US 4891810 A	02-01-90
US 4228496 A	14-10-80	BE 892627 A	16-07-82
		CA 1121481 A	06-04-82
		CA 1176338 A	16-10-84
		CA 1135809 A	16-11-82
		CA 1147474 A	31-05-83
		CA 1137582 A	14-12-82
		CA 1147824 A	07-06-83
		CA 1142619 A	08-03-83
		CA 1136728 A	30-11-82
		CA 1147417 A	31-05-83

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 96/18980

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 4228496 A		CA 1185670 C	16-04-85
		DE 2740056 A	16-03-78
		FR 2473197 A	10-07-81
		FR 2485227 A	24-12-81
		FR 2485228 A	24-12-81
		FR 2547082 A	07-12-84
		GB 1588804 A	29-04-81
		GB 1588805 A	29-04-81
		GB 1588806 A	29-04-81
		GB 1588807 A	29-04-81
		GB 1588803 A	29-04-81
		HK 62281 A	24-12-81
		HK 62381 A	24-12-81
		HK 62481 A	17-12-81
		HK 62581 A	24-12-81
		HK 62681 A	24-12-81
		JP 1353183 C	11-12-86
		JP 58050062 A	24-03-83
		JP 61020017 B	20-05-86
		JP 1257068 C	29-03-85
		JP 53033027 A	28-03-78
		JP 59025257 B	15-06-84
		JP 61286962 A	17-12-86
		JP 1408240 C	27-10-87
		JP 60100252 A	04-06-85
		JP 62018951 B	25-04-87
		JP 1353197 C	11-12-86
		JP 60100253 A	04-06-85
		JP 61020018 B	20-05-86
		JP 1353198 C	11-12-86
		JP 60100254 A	04-06-85
		JP 61020016 B	20-05-86
		JP 1356211 C	24-12-86
		JP 60100255 A	04-06-85
		JP 61024740 B	12-06-86
	JP 1359881 C	30-01-87	
	JP 60100256 A	04-06-85	
	JP 61029028 B	03-07-86	
	JP 1353199 C	11-12-86	
	JP 60100257 A	04-06-85	

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 96/18980

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
FR 2028517 A	09-10-70	DE 2001832 A	30-07-70
		GB 1243464 A	18-08-71
		NL 7000738 A,B,	21-07-70
