

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2012년 6월 28일 (28.06.2012)



(10) 국제공개번호
WO 2012/086914 A1

- (51) 국제특허분류:
G06F 11/10 (2006.01)
- (21) 국제출원번호: PCT/KR2011/008216
- (22) 국제출원일: 2011년 10월 31일 (31.10.2011)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보:
10-2010-0132416 2010년 12월 22일 (22.12.2010) KR
- (71) 출원인 (US 을(를) 제외한 모든 지정국에 대하여): (주) 케이티 (KT CORPORATION) [KR/KR]; 경기도 성남시 분당구 정자동 206, 463-711 Gyeonggi-do (KR).
- (72) 발명자: 겸
- (75) 발명자/출원인 (US 에 한하여): 박창식 (PARK, Chang-Sik) [KR/KR]; 서울시 강남구 일원본동 가람아파트 108동 407호, 135-947 Seoul (KR). 김미점 (KIM, Mi-Jeom) [KR/KR]; 경기도 안양시 동안구 관양동 1608-2, 현대 아이스페이스 606호, 431-060 Gyeonggi-do (KR). 김효민 (KIM, Hyo-Min) [KR/KR]; 서울시 서초구 우면동 40-1 103호, 137-900 Seoul (KR). 이어형 (LEE, Eo-

Hyung) [KR/KR]; 경기도 용인시 수지구 죽전 1동 현대홈타운 3차 1단지아파트 723동 102호, 448-160 Gyeonggi-do (KR). 황진경 (HWANG, Jin-Kyung) [KR/KR]; 경기도 과천시 중앙동 삼성레미안 APT 1108동 705호, 427-010 Gyeonggi-do (KR).

(74) 대리인: 리엔목특허법인 (Y.P.LEE, MOCK & PARTNERS); 서울시 서초구 서초동 1575-1, 137-875 Seoul (KR).

(81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

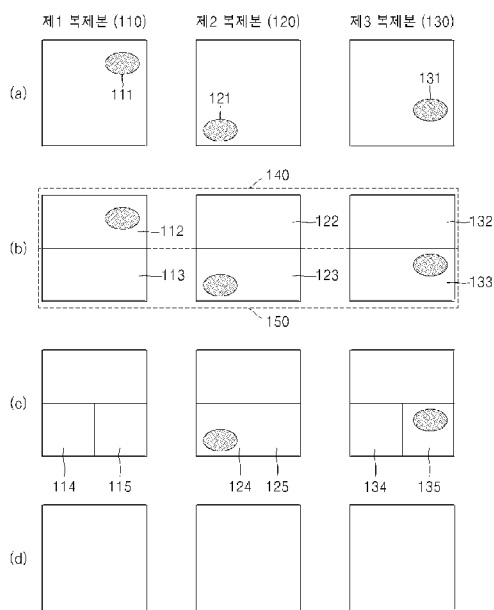
(84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, GH, GM,

[다음 쪽 계속]

(54) Title: METHOD FOR RECOVERING ERRORS FROM A PLURALITY OF ERROR REPLICAS AND STORAGE SYSTEM USING SAME

(54) 발명의 명칭 : 복수의 오류 복제본으로부터 오류를 복구하는 방법 및 상기 방법을 이용하는 스토리지 시스템

[Fig. 1]



110 ... First replica
120 ... Second replica
130 ... Third replica

(57) Abstract: The present invention provides a method for recovering errors when errors occur in every replica including the originals by using information about each error location in a plurality of replicas and information about errors which do not occur.

(57) 요약서: 본원 발명에서는 복수의 복제본 각각의 오류의 위치 및 오류가 발생하지 않은 정보를 이용하여 원본을 포함한 모든 복제본에 오류가 발생하는 경우에도 오류를 복구하는 방법을 제공한다.



KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

공개:

- 국제조사보고서와 함께 (조약 제 21 조(3))
- 청구범위 보정 기한 만료 전의 공개이며, 보정서를 접수하는 경우 그에 관하여 별도 공개함 (규칙 48.2(h))

명세서

발명의 명칭: 복수의 오류 복제본으로부터 오류를 복구하는 방법 및 상기 방법을 이용하는 스토리지 시스템

기술분야

- [1] 본 발명은 다수의 복제본(replica)를 가진 스토리지 시스템에서 원본을 포함한 모든 복제본에 오류가 있어 데이터의 일치성(consistency)이 없을 때 오류 복제본으로부터 오류를 복구하여 데이터의 일치성을 얻기 위한 방법에 관한 것이다.

배경기술

- [2] 대부분의 스토리지 시스템에서는 파일을 저장할 때 파일의 오류에 대비하여 다수의 복제본을 단일 서버 또는 분산된 서버들에 저장하며, 또한 수시로 이들 간의 데이터 일치성을 검사한다.
- [3] 데이터 일치성을 검사한 후 복제본 중 일부에 오류가 있는 경우, 오류가 발생한 파일을 오류가 발생하지 않은 파일로 대체함으로써 오류를 복구하였다. 그러나, 기존에 사용된 스토리지 시스템에서 사용되는 데이터 일치성 검사방법의 경우, 원본을 포함한 모든 복제본에 오류가 발생할 경우 오류를 복구할 수 없는 문제점이 있다.

발명의 상세한 설명

기술적 과제

- [4] 본원 발명에서는 원본을 포함한 모든 복제본에 오류가 발생한 경우에도 오류를 복구하는 방법을 제시한다. 구체적으로, 오류가 난 원본 파일 및 복제본으로부터 정상 상태의 파일로 오류 복구가 불가능했던 종래의 문제점에 대한 해결책을 제시한다.

과제 해결 수단

- [5] 본원 발명에서는 복수의 복제본 각각의 오류의 위치 및 오류가 발생하지 않은 정보를 이용하여 원본을 포함한 모든 복제본에 오류가 발생하는 경우에도 오류를 복구하는 방법을 제공한다.

발명의 효과

- [6] 스토리지 시스템이 각 파일의 마스터 패리티를 가지고 있는 경우 저장된 파일의 신뢰성 및 안전성 보장을 위해 다수의 복제본을 단일 서버 또는 분산된 서버들에 저장하는 스토리지 시스템에서 데이터 일치성 검사에서 모든 복제본의 패리티 값이 마스터 패리티 값과 서로 일치하지 않을 때 오류를 복구할 수 없지만, 본 발명에 따르면 모든 복제본들의 오류가 너무 많더라도 각 복제본에서의 오류의 비트 위치가 동일하지 않으면 오류의 복구가 가능하여 사용자의 파일을 안전하고 신뢰성 있게 관리할 수 있는 장점이 있다.

- [7] 또한, 스토리지 시스템이 각 파일의 마스터 패리티를 가지고 있지 않는 경우, 비록 1개의 복제본이 오류가 없더라도 오류가 없다는 사실을 증명할 수 없으며, 따라서 오류 복구가 불가능하지만, 본 발명에 따르면 각 복제본에서의 오류의 비트 위치가 동일하지 않으면 오류의 복구가 가능하여 마스터 패리티를 안전하게 보관하고 있어야 한다는 조건 없이도 사용자의 파일을 안전하고 신뢰성 있게 관리할 수 있는 장점이 있다.

도면의 간단한 설명

- [8] 도 1 (a) 내지 (d)는 본 발명의 바람직한 일 실시예로서, 오류 복구를 위해 이용하는 기본 개념을 도시한다.
- [9] 도 2는 본 발명의 바람직한 일 실시예로서, 스토리지 시스템이 각 파일의 마스터 패리티를 지닌 경우 오류 복구를 위한 서비스 흐름도를 도시한다.
- [10] 도 3은 본 발명의 바람직한 일 실시예로서, 스토리지 시스템이 각 파일의 마스터 패리티를 지닌 경우 오류 복구를 위한 서비스 흐름도를 구체화한 예를 도시한다.
- [11] 도 4는 본 발명의 바람직한 일 실시예로서, 스토리지 시스템이 각 파일의 마스터 패리티를 지니지 못한 경우 오류 복구를 위한 서비스 흐름도를 구체화한 예를 도시한다.
- [12] 도 5는 본 발명의 바람직한 일 실시예로서, 모든 복제본에 오류가 있는 경우 오류복구가 가능한 스토리지 시스템의 내부 구성도를 도시한다.

발명의 실시를 위한 최선의 형태

- [13] 바람직한 일 실시예로서, 스토리지 시스템에서 모든 복제본에 오류가 있는 경우 오류를 복구하는 방법은 (a)각 복제본을 이등분하는 단계; (b)이등분된 각 복제본 조각들 중 복제본들 간에 상호 대응되는 조각들의 패리티를 비교하는 단계; (c)상기 복제본들 간에 상호 대응되는 조각들 중 최소 2개 이상의 조각의 패리티가 일치하는 경우 해당 조각을 정상조각으로 판단하여, 상기 정상조각에 대응되는 상기 각 복제본 조각들의 오류를 복구하는 단계; (d)상기 복제본들 간에 상호 대응되는 조각들 간의 패리티가 모두 불일치하는 경우, 해당조각을 오류포함조각으로 판단하는 단계; 및 (e) 각 복제본에 대해 상기 오류포함조각에 대해서만 (a) 내지 (d) 단계를 반복 수행하는 단계;를 포함한다.
- [14] 또 다른 바람직한 일 실시예로서, 모든 복제본에 오류가 있는 경우 오류복구가 가능한 스토리지 시스템은 각 복제본의 전부 또는 일부를 이등분하는 분할부; 이등분된 각 복제본 조각들 중 복제본들 간에 상호 대응되는 조각들의 패리티를 비교하는 데이터일치검증부; 상기 복제본들 간에 상호 대응되는 조각들 중 최소 2개 이상의 조각의 패리티가 일치하는 경우 해당 조각을 정상조각으로 판단하여, 상기 정상조각에 대응되는 상기 각 복제본 조각들의 오류를 복구하는 정상조각참조부; 및 상기 복제본들 간에 상호 대응되는 조각들 간의 패리티가 모두 불일치하는 경우 해당조각을 오류포함조각으로 판단하고, 오류가 복구될

때까지 반복적으로 상기 오류포함조각을 상기 분할부, 상기 데이터일치검증부 및 상기 정상조각참조부를 통해 오류를 복구하는 오류복구순환부;를 포함한다.

발명의 실시를 위한 형태

- [15] 이하 첨부된 도면들을 참조하여 본 발명의 실시예들을 설명한다. 하기의 설명 및 첨부된 도면은 본 발명에 따른 동작을 이해하기 위한 것이며, 본 기술 분야의 통상의 기술자가 용이하게 구현할 수 있는 부분은 생략될 수 있다.
- [16] 또한 본 명세서 및 도면은 본 발명을 제한하기 위한 목적으로 제공된 것은 아니고, 본 발명의 범위는 청구의 범위에 의하여 정해져야 한다. 본 명세서에서 사용된 용어들은 본 발명을 가장 적절하게 표현할 수 있도록 본 발명의 기술적 사상에 부합하는 의미와 개념으로 해석되어야 한다.
- [17] 이하 첨부된 도면을 참조하여 본 발명의 실시예들을 설명한다.
- [18] 도 1 (a) 내지 (d)는 본 발명의 바람직한 일 실시예로서, 오류 복구에 이용되는 기본 개념을 도시한다.
- [19] 본원 발명에서는 원본을 포함한 모든 복제본이 일치하지 못하는 경우, 오류가 없는 부분의 유효한 정보를 통합하여 오류를 복구하고, 복제본들 간의 데이터 일치성을 제공한다.
- [20] 도 1 (a)는 세 개의 복제본이 모두 오류가 난 경우의 예를 도시한다.
- [21] 도 1(a)에 도시된 제 1 복제본(110), 제 2 복제본(120) 및 제 3 복제본(130)에서 검은색으로 표시된 부분(111, 121, 131)은 모두 오류를 포함하고 있는 부분이다.
- [22] 복제본들(110, 120, 130)은 모두 오류를 포함하고 있으므로, 제 1 복제본(110), 제 2 복제본(120) 및 제 3 복제본(130) 각각의 패리티와 마스터 패리티를 비교하면 일치하는 패리티는 하나도 없다.
- [23] 원본을 포함한 모든 복제본이 일치하지 못하는 경우, 오류가 없는 부분의 유효한 정보를 통합하기 위해 도 1(b)에 도시된 바와 같이 각 복제본을 이등분한다.
- [24] 이 후 복제본들(110, 120, 130)의 이등분된 영역 각각의 패리티를 상호 비교한다. 먼저, 도 1(b)의 제1영역(140)에 해당하는 각 복제본들의 1st_제1조각들(112, 122, 132)의 패리티를 상호 비교하여 오류를 복구하는 과정은 다음과 같다.
- [25] 각 복제본들의 1st_제1조각(112, 122, 132)의 패리티들을 비교한다. 비교 결과, 제 2 복제본(120)과 제 3 복제본(130)의 1st_제1조각들의(112,132) 패리티가 일치하며, 제 1 복제본의 1st_제1조각(112) 패리티는 제 2 복제본 및 제 3 복제본의 1st_제1조각(122, 132) 패리티 값과 일치하지 않는다.
- [26] 이상의 비교를 통해, 제 2 복제본(120)의 1st_제1조각(122)의 패리티와 제 3 복제본(130)의 1st_제1조각(132)의 패리티에 오류가 없는 것으로 판단하고, 제 1 복제본의 1st_제1조각(112)을 오류 없는 것으로 대체한다. 이로써 모든 복제본(110,120, 130)의 제 1영역(140)에 해당하는 각 복제본들의

- 1st_제1조각(112, 122, 132)의 오류가 복구된다.
- [27] 다음으로, 도 1(b)의 제 2 영역(150)에 해당하는 각 복제본들의 1st_제2조각들(113, 123, 133)의 패리티를 상호 비교하여 오류를 복구하는 과정은 다음과 같다.
- [28] 각 복제본들의 1st_제2조각들(113, 123, 133)의 패리티들을 비교한다. 비교 결과, 세 개의 복제본 모두 1st_제 2조각들(113,123,133)의 패리티가 모두 다르다. 따라서, 아직은 어느 이등분된 복제본에 오류가 있고 어느 것이 오류가 없는 지가 불분명하므로 제2 영역(150)을 도 1의 (c)와 같이 다시 이등분한다.
- [29] 도 1 (b)와 마찬가지로, 도 1(c)에서도 도 1(b)의 제 2영역(150)을 이등분한 두 개의 영역 중 하나의 영역에 해당하는 2nd_제1조각들(114, 124, 134)의 패리티를 비교한다.
- [30] 각 복제본들(110, 120, 130)의 2nd_제1조각들(114, 124, 134)의 패리티들을 비교한다. 비교 결과, 제 1 복제본(110)과 제3복제본(130)의 2nd_제1조각(114,134)들의 패리티가 일치하며, 제 2 복제본의 2nd_제1조각(124) 패리티는 제 2 복제본 및 제 3 복제본의 2nd_제1조각(114, 134) 패리티 값과 일치하지 않는다.
- [31] 이상의 비교를 통해, 제 1 복제본(110)과 제3복제본(130)의 2nd_제1조각(114,134)들의 패리티에 오류가 없는 것으로 판단하고, 제 2 복제본의 2nd_제1조각(124)을 오류 없는 것으로 대체한다. 이로써 모든 복제본(110,120, 130)의 2nd_제1조각(114, 124, 134)의 오류가 복구된다.
- [32] 도 1 (c)에서 2nd_제2조각들(115, 125, 135)의 오류도 위와 유사하게 복구한다. 또한, 도 1(c)의 경우와 달리 여전히 세 개의 2nd_제2조각들(115, 125, 135)의 패리티가 모두 다른 경우, 2nd_제2조각들(115, 125, 135)을 이등분하여 3rd_제1조각들(미 도시) 및 3rd_제2조각들(미 도시)로 나눌 수 있다.
- [33] 도 1(d)는 도 1(a) 내지 (c)의 과정을 통해 원본을 포함한 모든 복제본이 일치하지 못하는 경우, 오류가 없는 부분의 유효한 정보를 통합하여 오류를 복구한 예를 나타낸다. 도 1(d)에서는 모든 오류가 복구되었으므로 1st_제 1조각들(112, 122, 132), 1st_제 2조각들(113,123,133), 2nd_제1조각들(114, 124, 134) 및 2nd_제2조각들(115, 125, 135)을 각각 합하여 패리티를 계산하고, 그 결과를 스토리지 시스템이 지니고 있는 각 파일의 마스터(master) 패리티(parity)와 비교한다. 마스터 패리티와 일치하는 패리티를 가진 복제본은 오류가 없는 것이며, 만일 마스터 패리티와 다른 패리티를가진 복제본이 있으면 오류 없는 복제본으로 대체함으로써 모든 복제본의 오류를 복구한다.
- [34] 본 발명에서 데이터 일치성을 이용하여 오류를 복구하는 방법으로는 크게 2가지 방법이 있다.
- [35] 첫 째는 스토리지 시스템이 각 파일의 마스터 패리티를 갖고 있고, 각 복제본의 패리티와 마스터 패리티를 비교함으로써 마스터 패리티와의 일치성을 검사할 때 오류를 복구하는 방법이다. 두 번째는 스토리지 시스템이 마스터 패리티를

찾고 있지 않고 각 복제본의 패리티를 상호 비교함으로써 데이터 일치성을 검사하여 오류를 복구하는 방법이다. 첫 번째 오류 복구 방법에 대해서는 도 2 및 3에서, 두 번째 오류 복구 방법에서는 도 4에서 서술하겠다.

- [36] 도 2는 본 발명의 바람직한 일 실시예로서, 스토리지 시스템이 각 파일의 마스터 패리티를 지닌 경우 오류 복구를 위한 서비스 흐름도를 도시한다.
- [37] 원본을 포함한 복제본들 각각의 패리티를 계산한다. 복제본들의 패리티 중 최소한 1개 이상의 패리티가 마스터 패리티와 일치하면 오류 있는 복제본은 오류 없는 복제본으로 대체함으로써 오류를 복구할 수 있다. 따라서, 도 2에서는 모든 복제본의 패리티가 마스터 패리티와 일치하지 경우 오류를 복구하는 흐름도를 도시한다.
- [38] 먼저, 각 복제본을 제1조각, 제2조각으로 이등분한다(S210). 각 복제본의 제1조각 및 제 2 조각은 각각 대응되는 영역을 나타낸다. 예를 들면, 제 1 조각은 각 복제본을 이등분한 영역 중 Upper half 부분을 제 2 조각은 각 복제본을 이등분한 영역 중 Lower half 부분을 나타낼 수 있다. 도 1(b)의 경우를 참고하면, 각 복제본의 제1조각들은 112, 122, 132를 각각 나타내고, 각 복제본의 제 2 조각들은 113, 114, 115 를 각각 나타낸다.
- [39] 이후 이등분된 복제본 조각들의 패리티를 계산한다(S220). 제1조각들 간에 패리티를 비교하여, 최소한 2개 이상의 패리티가 서로 일치하는 경우, 일치하는 복제본 제1조각들은 오류가 없는 것으로 판단하고(S230), 이를 기초로 제 1 조각들 중 패리티가 일치하지 않는 복제본의 오류를 복구한다(S240). 즉, 최소한 2개 이상의 제 1 조각들의 패리티가 서로 일치하면 이들 데이터에는 오류가 없거나 오류가 이미 복구된 것으로 판단할 수 있으며, 이들로 오류가 있는 복제본을 복구할 수 있다. 이상의 과정을 제 2 조각에 대해서도 동일 또는 유사하게 수행한다.
- [40] 일치하지 않는 경우, 제1조각들을 다시 이등분하여(S240), 이등분된 각 조각들의 패리티를 계산한 후(S220), 대응되는 복제본 조각들의 패리티간 최소한 2개 이상의 패리티가 서로 일치하는지를 판단한다(S230).
- [41] 일치하는 경우, 일치하는 복제본 조각들은 오류가 없는 것으로 판단하여, 이를 기초로 대응되는 복제본 조각들 중 패리티가 일치하지 않는 복제본의 오류를 복구한다(S240). 그리고 이전에 오류가 복구된 복제본 조각에 통합한다(S241).
- [42] 오류가 복구된 복제본 조각이 있는지 여부는 패리티가 일치하지 않는 복제본 조각이 없을 때까지 지속된다(S250). 패리티가 일치하지 않는 복제본 조각이 없는지를 구현하는 일 예는 도 3의 s351 단계를 참고한다.
- [43] 패리티가 일치하지 않는 복제본 조각이 여전히 존재하는 경우에는, 패리티가 일치하지 않는 복제본 조각의 크기가 1비트가 될 때까지 반복한다(S231). 다만, 주의할 것은 본 발명에서 기술한 1비트라는 숫자는 예시적인 것으로, 이는 변형, 변경이 가능함을 유의하여야 한다. 이상의 과정을 제2조각에 대해서도 동일 또는 유사하게 수행한다.

- [44] 각 복제본의 모든 복제본 조각이 오류를 복구하여 통합된 경우, 마스터 패리티와 패리티가 일치하는 복제본 패리티가 최소 1개 이상인지 판단한다(S260). 복제본 패리티가 최소 1개 이상인 경우, 오류 복구가 성공적으로 이루어진 경우이다.
- [45] 마스터 패리티와 패리티가 일치하는 복제본 패리티가 1개도 없는 경우에는 오류복구가 실패한 경우이다. 이 경우는 모든 복제본의 동일 비트 위치에 오류가 있어 오류 발견이 불가능하여 더 이상 오류를 복구할 수 없고 따라서, 모든 과정을 종료한다.
- [46] 도 3은 본 발명의 바람직한 일 실시예로서, 스토리지 시스템이 각 파일의 마스터 패리티를 지닌 경우 오류 복구를 위한 서비스 흐름도를 구체화한 예를 도시한다.
- [47] 각 복제본들의 패리티를 계산한다(S300). 계산된 각 복제본들의 패리티들을 스토리지 시스템이 미리 갖고 있는 마스터 패리티와 비교한다(S301). 계산된 패리티가 마스터 패리티와 일치하는 복제본은 오류가 없는 복제본이고 일치하지 않으면 오류가 있는 복제본이다.
- [48] 다수의 복제본 중에 최소한 1개 이상의 계산된 패리티가 마스터 패리티와 일치하면 오류 있는 복제본은 오류 없는 복제본으로 대체함으로써 오류를 복구하고 모든 과정을 종료한다. 만일 모든 복제본의 패리티가 마스터 패리티와 일치하지 않으면 모든 복제본에 오류가 있는 것으로서 판단한다.
- [49] 오류가 있는 것으로 판단되면, $n=1$ 및 $i=0$ 으로 지정한다(S302). 여기서 n 은 오류 복구를 위해 복제본을 이등분하는 레벨을 나타내는 것으로서 n 번 이등분 했다는 의미이다. i 는 이등분한 복제본을 나타낸다. i 가 0이면 이등분한 복제본의 제1조각을 i 가 1이면 이등분한 복제본의 제2조각을 나타낸다.
- [50] 이후, 각 복제본을 이등분한다(S310). 이등분된 복제본은 $R_{n,i}(x)$ 로 표기한다. 여기서, x 는 몇 번째 복제본인지를 나타낸다. 따라서, 예를 들어 $R_{1,0}(2)$ 는 스토리지 시스템이 갖고 있는 2번째 복제본을 처음 이등분했을 때의 제1조각을 나타낸다.
- [51] 이등분된 각 복제본 조각의 패리티를 계산하고(S320), 계산된 패리티들을 비교한다(S330).
- [52] S330 단계에서 최소한 2개 이상의 이등분된 복제본의 패리티가 서로 일치하면 패리티가 같은 이등분된 복제본에는 오류가 없는 것으로 판단한다.
- [53] 오류가 없는 것으로 판단된 경우, i 값이 0인지 검사한다(S340). i 값이 0이면 S310 및 S320에서 이등분된 복제본의 제1조각을 검사한 것으로, 이제 나머지 반쪽인 제2조각을 검사하기 위해 $i=1$ 로 지정한 후(S341), 제2조각들에 대해 패리티를 계산하고(S320), 제2조각들 중 최소한 2개 이상의 이등분된 복제본의 패리티가 서로 일치하면 패리티가 같은 이등분된 복제본에는 오류가 없는 것으로 판단한다(S330).
- [54] $i=1$ 이면 이등분한 복제본의 제2조각까지 검사를 마친 경우로서 제1조각 및

- 제2조각 모두가 오류가 없거나 오류가 복구된 것으로 판단되므로 모든 복제본의 제1조각 및 제2조각을 합하여 패리티를 계산하고 계산된 패리티들을 비교한다(S342).
- [55] 이후, 만일 최소한 2개 이상의 패리티가 서로 일치하면(S350) 제1조각 및 제2조각을 합한 이들 데이터에는 오류가 없거나 오류가 이미 복구된 것으로 판단할 수 있으며, 이들로 오류가 있는 복제본을 복구한다.
- [56] 제1조각 및 제2조각을 1개로 합했다는 의미는 오류 복구를 위한 레벨을 1단계 상위로 올렸다는 의미이므로 $n = n-1$ 로 n 값을 1 감소한다(S351).
- [57] 이때, 만일 $n = 0$ 이 아니고(S352), $i = 0$ 이면(S353), $i = 1$ 로 바꾼다(S341). 이후 다시 이등분된 각 복제본 조각의 패리티를 계산하고 계산된 패리티들을 비교한다(S320). 만일 최소한 2개 이상의 이등분된 복제본의 패리티가 서로 일치하면 패리티가 같은 이등분된 복제본에는 오류가 없는 것으로 판단하고(S330), S340 단계 이후를 진행한다.
- [58] 만일 $n = 0$ 이 아니고(S352), $i = 1$ 이면(S353), 이등분한 복제본의 제2조각까지 검사를 마친 경우로서 다시 모든 복제본의 제1조각 및 제2조각을 합하여 패리티를 계산하고 계산된 패리티들을 비교한다(S342).
- [59] 만일 $n = 0$ 이면, 모든 복제본 조각의 오류가 복구된 것이므로 마스터 패리티와 최소 1개 이상의 복제본 패리티가 일치하는지 판단한다(S380). 일치하면, 오류복구를 성공한 것으로 보고, 그렇지 않은 경우 오류 복구가 실패된 경우로 본다.
- [60] 오류 복구를 실패하는 경우는 모든 복제본의 동일한 비트 위치에서 오류가 있을 때이다. 그러나, 실제로 파일의 크기가 크고 파일의 모든 비트 위치에서의 오류 확률이 동일하다고 가정할 때 모든 복제본의 동일한 비트 위치에서 오류가 있을 가능성은 무시할 수 있을 정도로 아주 작다.
- [61] S330 단계에서 모든 복제본의 패리티가 일치하지 않는 경우 n 값을 1 증가시킨다(S360). 이후 이등분된 복제본의 크기($|R_{n,i}(x)|$)가 1보다 큰지를 비교한다(S370). 만일 $|R_{n,i}(x)| > 1$ 인 경우, 즉 복제본의 크기가 1비트 이상인 경우 $i = 0$ 으로 하고, S310 단계로 이동한다.
- [62] S360 단계에서 만일 $|R_{n,i}(x)| = 1$ 이면 이 경우는 이등분된 복제본의 크기가 1비트인 경우로서 이 비트 위치에서의 값을 다수의 원리에 따라 정하고 S340 단계로 이동한다.
- [63] 스토리지 시스템에 저장된 파일의 복제본들간 데이터 일치성 검사에서 발견된 오류는 도 3에 도시된 흐름도를 통해 복구된다. 다만, 오류 복구가 불가능한 경우는 모든 복제본이 모두 동일한 비트 위치에 오류가 있을 경우이다. 그러나 앞에서 언급한 바와 같이 실제로 파일의 크기가 크고 파일의 모든 비트 위치에서의 오류 확률이 동일하다고 가정할 때 모든 복제본의 동일한 비트 위치에서 오류가 있을 가능성은 무시할 수 있을 정도로 아주 작다.
- [64] 즉, 복수 개의 복제본들의 특정 부분에 오류가 있더라도, 모두 동일한 위치에

- 오류가 있는 경우가 아닌 한 본원 발명의 방법을 통해 오류 복구가 가능하다.
- [65] 도 4는 본 발명의 바람직한 일 실시예로서, 스토리지 시스템이 각 파일의 마스터 패리티를 지니지 못한 경우 오류 복구를 위한 서비스 흐름도를 구체화한 예를 도시한다.
- [66] 도 4의 오류복구 흐름도는 도 3의 s301 과 s380 단계만 상이하고, 그 외의 단계는 유사하므로, 이하에서는 도 3과 상이한 s401 및 s480 단계만 기술하겠다. 그 외의 단계는 도 3의 설명을 참고한다.
- [67] s401 단계에서는 다수의 복제본 중에 최소한 2개 이상의 패리티가 서로 일치하면 오류 있는 복제본은 오류 없는 복제본으로 대체함으로써 오류를 복구한다. 즉, 패리티가 서로 일치하는 복제본은 오류가 없는 복제본으로 판단하고 일치하지 않으면 오류가 있는 복제본으로 판단한다. 만일 모든 복제본의 패리티가 서로 일치하지 않으면, 비록 실제로는 1개의 복제본이 오류가 없는 경우일 지라도 모든 복제본에 오류가 있는 것으로서 판단한다.
- [68] s480 단계에서는 오류 복구 작업이 모두 수행된 것으로 판단되는 복제본의 패리티를 계산한 후, 계산된 패리티들을 비교하여 최소한 두 개의 복제본의 패리티가 상호일치하면, 오류 복구가 성공적으로 수행된 것으로 판단한 후, 이 복제본을 기초로 패리티가 다른 복제본들의 오류 복구를 수행한다.
- [69] 도 5 는 본 발명의 바람직한 일 실시예로서, 모든 복제본에 오류가 있는 경우 오류복구가 가능한 스토리지 시스템의 내부 구성도를 도시한다.
- [70] 본 발명의 바람직한 일 실시예로서 스토리지 시스템(500)은 분할부(510), 데이터일치검증부(520), 정상조각참조부(530), 오류복구순환부(540) 및 복구부(550)를 포함한다.
- [71] 분할부(510)는 각 복제본의 전부 또는 일부를 이등분한다. 최초의 복제본을 이등분할 때는 복제본 전부를 이등분하나, 이후 복제본 중 일부가 오류가 없는 데이터로 검증되면, 오류가 있는 부분만을 이등분한다. 이 경우 상이한 기준으로 이등분 할 경우 대응되는 복제본 조각이 상이할 수 있으므로, 각 복제본의 전부 또는 일부를 모두 동일한 기준으로 이등분한다.
- [72] 데이터일치검증부(520)는 이등분된 각 복제본 조각들 중 복제본들 간에 상호 대응되는 조각들의 패리티를 비교한다. 이에 대해서는 도 1(b) 내지 (c)를 참고한다.
- [73] 정상조각참조부(530)는 복제본들 간에 상호 대응되는 조각들 중 최소 2개 이상의 조각의 패리티가 일치하는 경우 해당 조각을 정상조각으로 판단하여, 상기 정상조각에 대응되는 각 복제본 조각들의 오류를 복구한다.
- [74] 오류복구순환부(540)는 복제본들 간에 상호 대응되는 조각들 간의 패리티가 모두 불일치하는 경우 해당조각을 오류포함조각으로 판단한다. 이후 복제본 전체가 아니라 오류포함조각만 반복해서 이등분하여 부분적인 패리티 검사를 통해 오류가 없는 부분을 찾아내어 정상조각으로 판단하는 방식으로 오류를 복구한다. 이를 위해, 오류복구순환부(540)는 오류가 복구될 때까지 반복적으로

오류포함조각을 분할부(500), 데이터일치검증부(510) 및 정상조각참조부(520)를 통해 오류를 복구한다. 오류복구순환부(540)는 오류포함조각이 없거나 또는 오류포함조각의 크기가 1비트일때까지 동작하도록 구현된다.

[75] 복구부(550)는 더이상 오류포함조각이 없는 경우, 정상조각들을 합하여 각 복제본을 복구한다. 이후 복구된 복제본의 오류 복구가 성공적인지 확인하기 위하여, 스토리지 시스템이 각 파일의 마스터 패리티를 가지고 있는 경우 오류가 복구된 복제본들의 패리티들 중 적어도 하나의 패리티와 마스터 패리티가 일치하는지 판단한다. 스토리지 시스템이 각 파일의 마스터 패리티를 가지고 있지 않는 경우에는 오류가 복구된 복제본들의 패리티 값 중 적어도 두 개의 패리티 값이 일치하는지 판단하여 오류복구가 성공적으로 이루어졌는지 확인할 수 있다.

[76] 한편, 본 발명은 컴퓨터 판독가능 저장매체에 컴퓨터가 판독 가능한 코드를 저장하여 구현하는 것이 가능하다. 상기 컴퓨터 판독가능 저장매체는 컴퓨터 시스템에 의하여 판독될 수 있는 데이터가 저장되는 모든 종류의 저장 장치를 포함한다.

[77] 컴퓨터 판독 가능 저장 매체의 예로는 ROM, RAM, CD-ROM, 자기 테이프, 플로피디스크, 광 데이터 저장장치 등이 있으며, 또한 방송파(예를 들어 인터넷을 통한 전송)의 형태로 구현하는 것을 포함한다. 또한, 컴퓨터 판독 가능 저장 매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산 방식으로 컴퓨터가 판독 가능한 코드가 저장되고 실행될 수 있다.

[78] 이제까지 본 발명에 대하여 바람직한 실시예를 중심으로 살펴보았다. 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 본 발명을 구현할 수 있음을 이해할 것이다. 그러므로 상기 개시된 실시예들은 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 한다. 본 발명의 범위는 전술한 설명이 아니라 특허청구범위에 나타나 있으며, 특허청구범위에 의해 청구된 발명 및 청구된 발명과 균등한 발명들은 본 발명에 포함된 것으로 해석되어야 한다.

[79]

[80]

청구범위

- [청구항 1] 스토리지 시스템에서 모든 복제본에 오류가 있는 경우, 복수의 오류 복제본으로부터 오류를 복구하는 방법으로서,
 (a) 각 복제본을 이등분하는 단계;
 (b) 이등분된 각 복제본 조각들 중 복제본들 간에 상호 대응되는 조각들의 패리티를 비교하는 단계;
 (c) 상기 복제본들 간에 상호 대응되는 조각들 중 최소 2개 이상의 조각의 패리티가 일치하는 경우 해당 조각을 정상조각으로 판단하여, 상기 정상조각에 대응되는 상기 각 복제본 조각들의 오류를 복구하는 단계;
 (d) 상기 복제본들 간에 상호 대응되는 조각들 간의 패리티가 모두 불일치하는 경우, 해당조각을 오류포함조각으로 판단하는 단계; 및
 (e) 각 복제본 내의 상기 오류포함조각에 대해서만 (a) 내지 (d) 단계를 반복 수행하는 단계;를 포함하는 것을 특징으로 하는 방법.
- [청구항 2] 제 1 항에 있어서,
 상기 오류포함조각이 없는 경우, 상기 정상조각들을 합하여 각 복제본을 복구하는 단계를 더 포함하는 것을 특징으로 하는 방법.
- [청구항 3] 제 1 항에 있어서, 상기 (e) 단계는
 상기 오류포함조각의 크기가 1비트일 때까지 수행되는 것을 특징으로 하는 방법.
- [청구항 4] 제 2 항에 있어서,
 스토리지 시스템이 각 파일의 마스터 패리티를 가지고 있는 경우
 상기 오류가 복구된 복제본들의 패리티들 중 적어도 하나의 패리티와 상기 마스터 패리티가 일치하는지 판단하는 단계를 더 포함하는 것을 특징으로 하는 방법.
- [청구항 5] 제 2 항에 있어서,
 스토리지 시스템이 각 파일의 마스터 패리티를 가지고 있지 않는 경우, 상기 오류가 복구된 복제본들의 패리티 값 중 적어도 두 개의 패리티 값이 일치하는지 판단하는 단계를 더 포함하는 것을 특징으로 하는 방법.
- [청구항 6] 제 1 항에 있어서, 상기 (a) 단계는
 각 복제본에 대해 모두 동일한 기준으로 이등분하는 것을 특징으로 하는 방법.
- [청구항 7] 제 1 항에 있어서, 상기 스토리지 시스템은 최소 3개의 복제본을 포함하는 것을 특징으로 하는 방법.
- [청구항 8] 모든 복제본에 오류가 있는 경우에도 오류복구가 가능한 스토리지

시스템으로서,

각 복제본의 전부 또는 일부를 이등분하는 분할부;

이등분된 각 복제본 조각들 중 복제본들 간에 상호 대응되는 조각들의 패리티를 비교하는 데이터일치검증부;

상기 복제본들 간에 상호 대응되는 조각들 중 최소 2개 이상의 조각의 패리티가 일치하는 경우 해당 조각을 정상조각으로 판단하여, 상기 정상조각에 대응되는 상기 각 복제본 조각들의 오류를 복구하는 정상조각참조부; 및

상기 복제본들 간에 상호 대응되는 조각들 간의 패리티가 모두 불일치하는 경우 해당조각을 오류포함조각으로 판단하고, 오류가 복구될 때까지 반복적으로 상기 오류포함조각을 상기 분할부, 상기 데이터일치검증부 및 상기 정상조각참조부를 통해 오류를 복구하는 오류복구순환부;를 포함하는 것을 특징으로 하는 스토리지시스템.

[청구항 9]

제 8 항에 있어서,

상기 오류복구순환부에서 상기 복제본들 간에 상호 대응되는 조각들 간의 패리티가 모두 불일치하는 경우가 없어 오류포함조각이 없는 경우, 상기 정상조각들을 합하여 각 복제본을 복구하는 복구부;를 더 포함하는 것을 특징으로 하는 스토리지시스템.

[청구항 10]

제 8 항에 있어서, 상기 오류복구순환부는

상기 오류포함조각의 크기가 1비트일 때까지 오류복구를 수행하는 것을 특징으로 하는 스토리지시스템.

[청구항 11]

제 9 항에 있어서, 상기 스토리지 시스템은

상기 스토리지 시스템이 각 파일의 마스터 패리티를 가지고 있는 경우 상기 오류가 복구된 복제본들의 패리티들 중 적어도 하나의 패리티와 상기 마스터 패리티가 일치하는지 판단하는 것을 특징으로 하는 스토리지시스템.

[청구항 12]

제 9 항에 있어서, 상기 스토리지 시스템은

상기 스토리지 시스템이 각 파일의 마스터 패리티를 가지고 있지 않는 경우, 상기 오류가 복구된 복제본들의 패리티 값 중 적어도 두 개의 패리티 값이 일치하는지 판단하는 것을 특징으로 하는 스토리지시스템.

[청구항 13]

제 8 항에 있어서, 상기 분할부는

각 복제본의 전부 또는 일부를 모두 동일한 기준으로 이등분하는 것을 특징으로 하는 스토리지시스템.

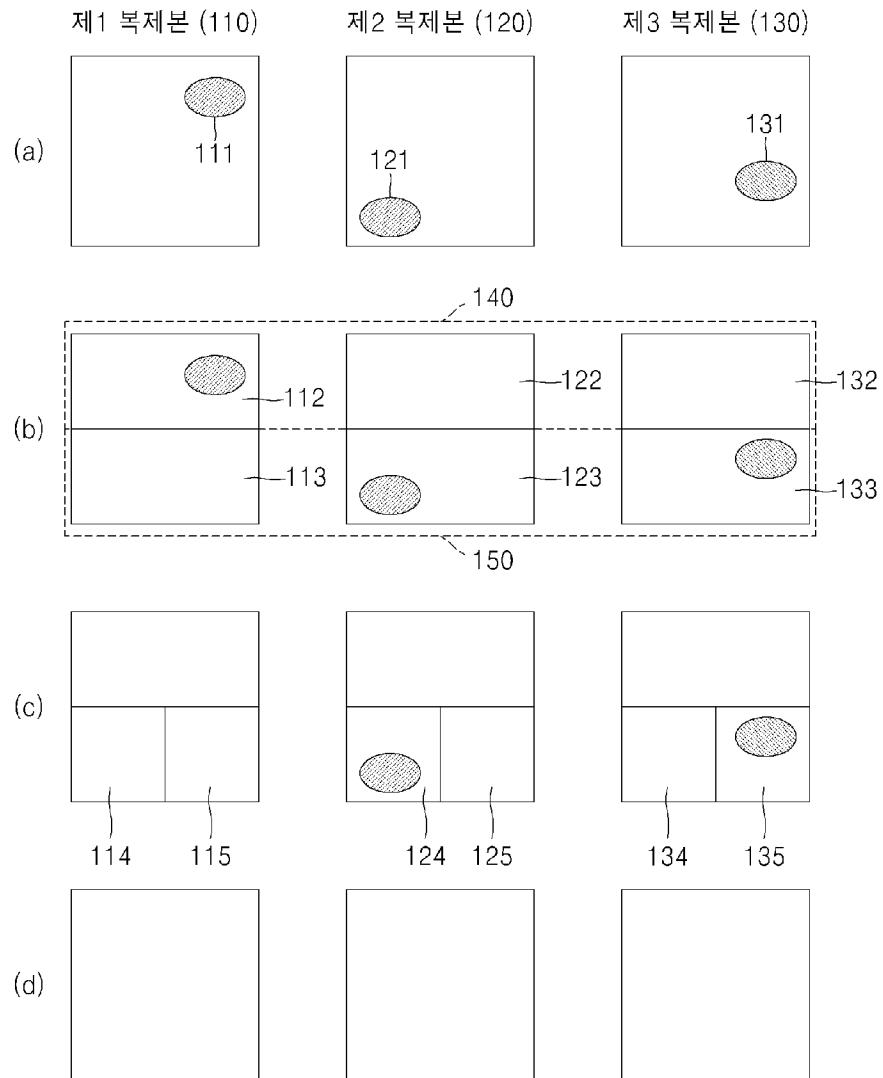
[청구항 14]

제 8 항에 있어서, 상기 스토리지시스템은

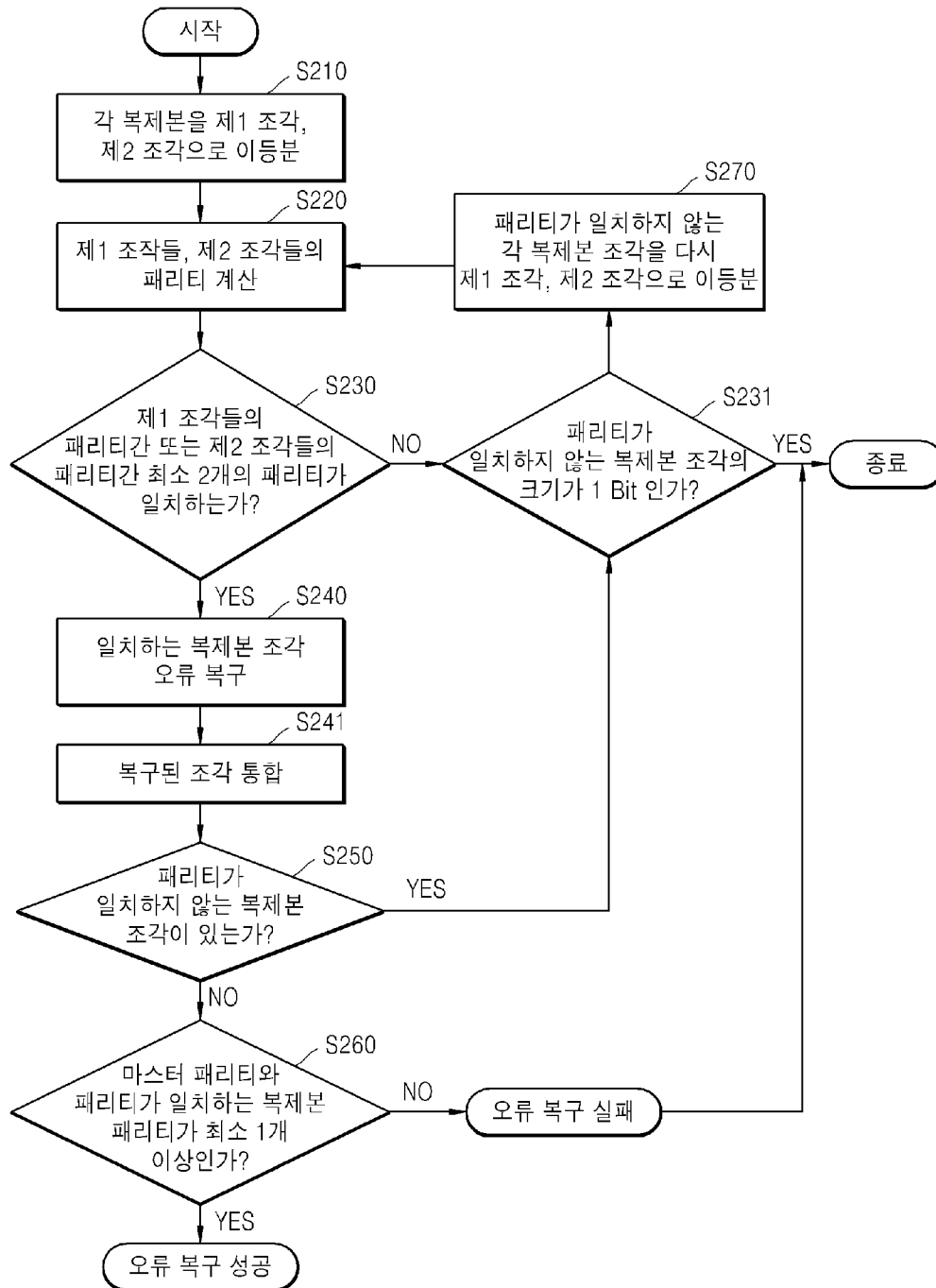
최소 3개의 복제본을 포함하는 것을 특징으로 하는

스토리지시스템.

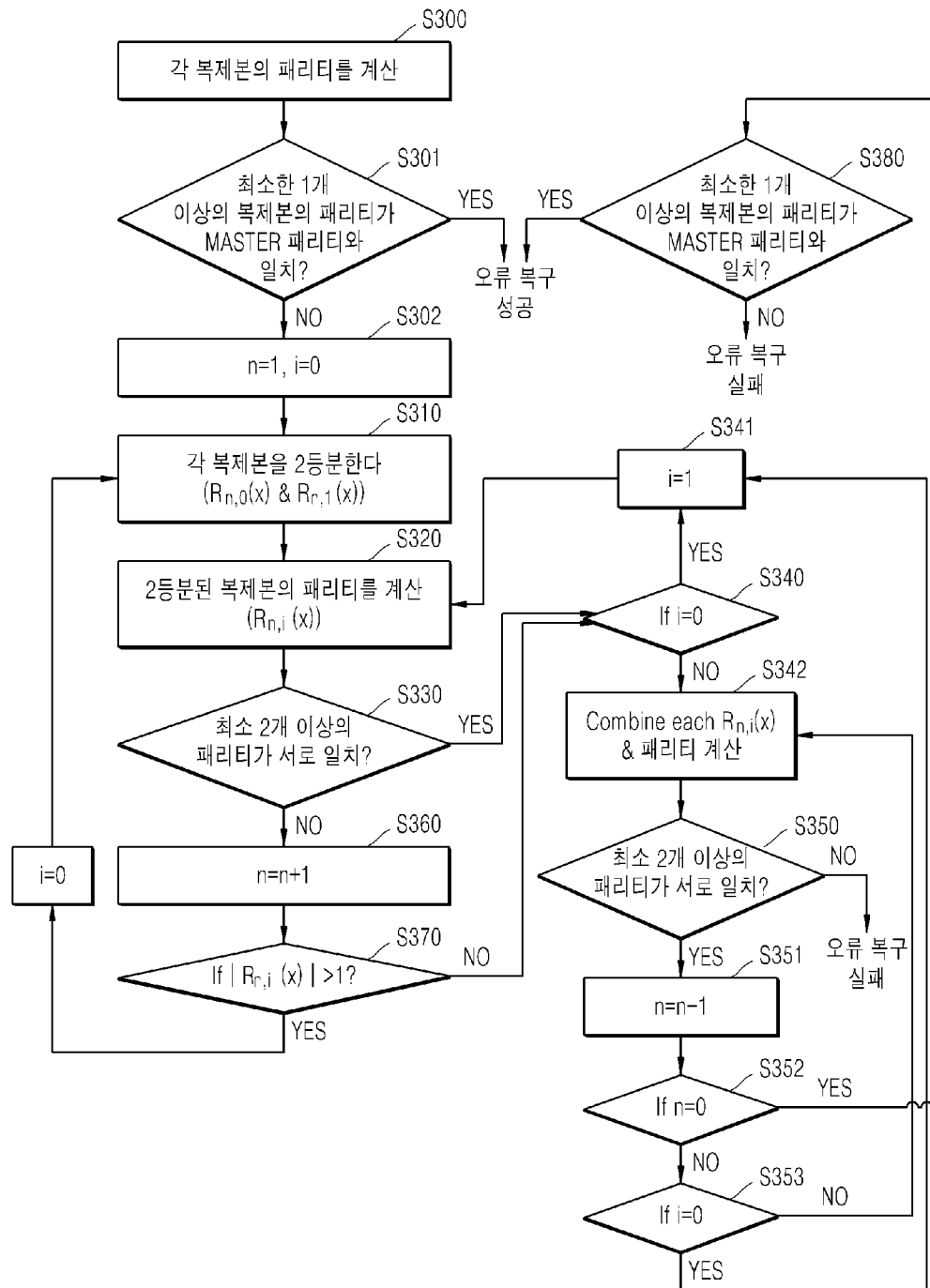
[Fig. 1]



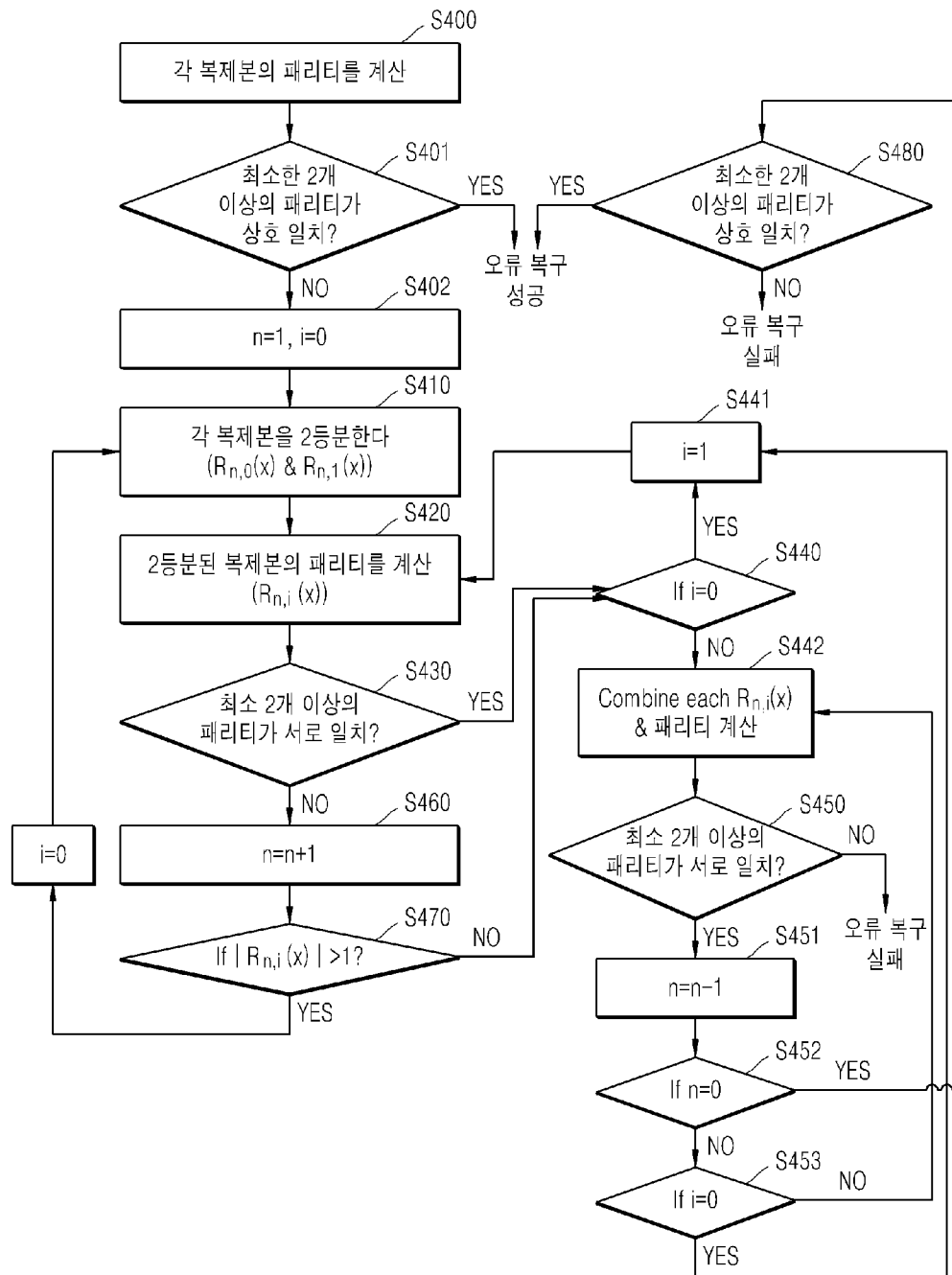
[Fig. 2]



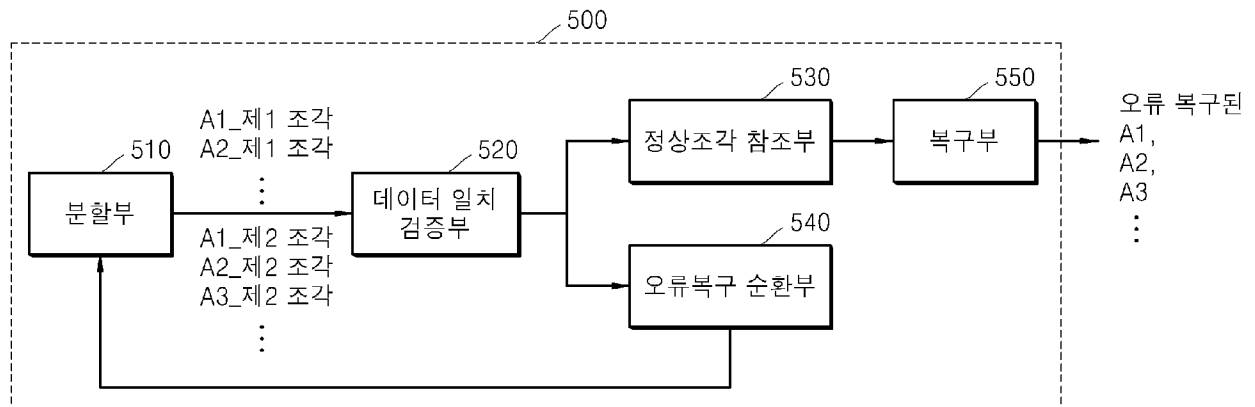
[Fig. 3]



[Fig. 4]



[Fig. 5]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2011/008216**A. CLASSIFICATION OF SUBJECT MATTER****G06F 11/10(2006.01)i**

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 : G06F; H04L; G11B

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

Korean Utility models and applications for Utility models: IPC as above

Japanese Utility models and applications for Utility models: IPC as above

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

partition, error correction, parity, halving

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KR 10-2009-0036276 A (DAEWOO ELECTRONICS CORPORATION) 14 April 2009 See abstract, paragraphs <22-31>, figure 3	1-14
A	KR 10-2001-0091593 A (LG ELECTRONICS INC.) 23 October 2001 See abstract, paragraphs <18-35>, claims 1, 3.	1-14
A	KR 10-1993-0023826 A (SAMSUNG ELECTRONICS CO., LTD.) 21 December 1993 See abstract, claim 1.	1-14
A	KR 10-1993-0007928 A (SAMSUNG ELECTRONICS CO., LTD.) 21 August 1993 See claim 1, figure 3	1-14



Further documents are listed in the continuation of Box C.



See patent family 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 application or patent 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

20 APRIL 2012 (20.04.2012)

Date of mailing of the international search report

23 APRIL 2012 (23.04.2012)

Name and mailing address of the ISA/KR



Korean Intellectual Property Office
Government Complex-Daejeon, 139 Seonsa-ro, Daejeon 302-701,
Republic of Korea

Facsimile No. 82-42-472-7140

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/KR2011/008216

Patent document cited in search report	Publication date	Patent family member	Publication date
KR 10-2009-0036276 A	14.04.2009	NONE	
KR 10-2001-0091593 A	23.10.2001	NONE	
KR 10-1993-0023826 A	21.12.1993	NONE	
KR 10-1993-0007928 B1	21.08.1993	JP 02-644639 B2 JP 04-248630 A US 05315601 A	02.05.1997 04.09.1992 24.05.1994

A. 발명이 속하는 기술분류(국제특허분류(IPC)) G06F 11/10(2006.01)i		
B. 조사된 분야 조사된 최소문헌(국제특허분류를 기재) IPC : G06F; H04L; G11B 조사된 기술분야에 속하는 최소문헌 이외의 문헌 한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우)) 분할, 오류 수정, 패리티, 이등분		
C. 관련 문헌		
카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
A	KR 10-2009-0036276 A (주식회사 대우일렉트로닉스) 2009.04.14 요약, 문단번호 <22-31>, 도면 3 참조	1-14
A	KR 10-2001-0091593 A (엘지전자주식회사) 2001.10.23 요약, 문단번호 <18-35>, 청구항 1, 3 참조.	1-14
A	KR 10-1993-0023826 A (삼성전자주식회사) 1993.12.21 요약, 청구항 1 참조.	1-14
A	KR 10-1993-0007928 B1 (삼성전자주식회사) 1993.08.21 청구항 1, 도면 3 참조	1-14
☐ 추가 문헌이 C(계속)에 기재되어 있습니다. ☒ 대응특허에 관한 별지를 참조하십시오.		
* 인용된 문헌의 특별 카테고리: “A” 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌 “E” 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌 “L” 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌 “O” 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌 “P” 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌 “T” 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌 “X” 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신구성 또는 진보성이 없는 것으로 본다. “Y” 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다. “&” 동일한 대응특허문헌에 속하는 문헌		
국제조사의 실제 완료일 2012년 04월 20일 (20.04.2012)		국제조사보고서 발송일 2012년 04월 23일 (23.04.2012)
ISA/KR의 명칭 및 우편주소  대한민국 특허청 (302-701) 대전광역시 서구 청사로 189, 정부대전청사 팩스 번호 82-42-472-7140		심사관 이정은 전화번호 82-42-481-5391 

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2009-0036276 A	2009.04.14	없음	
KR 10-2001-0091593 A	2001.10.23	없음	
KR 10-1993-0023826 A	1993.12.21	없음	
KR 10-1993-0007928 B1	1993.08.21	JP 02-644639 B2 JP 04-248630 A US 05315601 A	1997.05.02 1992.09.04 1994.05.24