(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2017년06월09일
(11) 등록번호 10-1744892
(24) 등록일자 2017년06월01일

(51) 국제특허분류(Int. Cl.)
G06F 17/30 (2006.01)
(52) CPC특허분류
G06F 17/30194 (2013.01)
G06F 17/30094 (2013.01)
(21) 출원번호 10-2016-0032100
(22) 출원일자 2016년03월17일
심사청구일자 2016년03월17일
(56) 선행기술조사문헌
JP2011070257 A*
KR1020130093202 A*
JP2005284520 A*
JP2010224705 A*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
주식회사 젠시스템즈
서울특별시 영등포구 경인로71길 70, 1601-3호(문래동5가, 벽산디지털밸리)
(72) 발명자
호경중
서울특별시 마포구 독막로28길 10 101동 2102호 (신수동, 밤섬경남아너스빌)
김태근
경기도 의정부시 능곡로 69, 4동 805호 (신곡동, 신성벽산아파트)
(74) 대리인
특허법인메이저

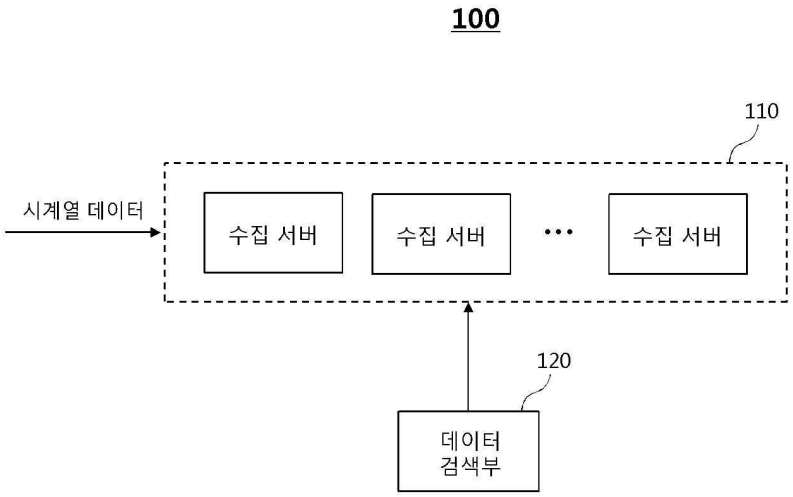
전체 청구항 수 : 총 4 항 심사관 : 김경완

(54) 발명의 명칭 **시계열 계층 인덱싱을 이용한 데이터 검색 시스템 및 데이터 검색 방법**

(57) 요 약

본 발명의 일 기술적 측면에 따른 데이터 검색 시스템은, 수집된 시계열 데이터의 시간 정보를 이용하여 시계열 인덱스를 생성하고, 상기 시계열 인덱스를 포함하는 로그 파일을 생성하여, 저장 기간에 따라 계층적으로 구분되는 복수의 수집 서버에 분산하여 저장하는 데이터 수집부 및 상기 복수의 수집 서버에 대한 계층적인 구조 정보를 유지하고, 사용자의 검색 요청에 대하여 상기 계층적인 구조 정보를 참조하여 검색을 수행할 적어도 하나의 수집 서버를 선택하는 데이터 검색부를 포함할 수 있다.

대 표 도 - 도1



(52) CPC특허분류
G06F 17/30185 (2013.01)
G06F 17/30221 (2013.01)
G06F 17/30386 (2013.01)

명 세 서

청구범위

청구항 1

수집된 시계열 데이터의 시간 정보를 이용하여 시계열 인덱스를 생성하고, 상기 시계열 인덱스를 포함하는 로그 파일을 생성 및 분산하여 저장하며, 서로 다른 성능을 갖는 저장매체를 사용하고, 상기 로그 파일의 저장 기간에 따라 계층적으로 구분되는 복수의 수집 서버로 이루어지는 제1 및 제2 수집 서버 그룹을 포함하는 데이터 수집부; 및

상기 복수의 수집 서버에 대한 계층적인 구조 정보를 유지하고, 사용자의 검색 요청에 대하여 상기 계층적인 구조 정보를 참조하여 검색을 수행할 적어도 하나의 수집 서버를 선택하는 데이터 검색부를 포함하고,

상기 제1 수집 서버 그룹의 수집 서버는,

수집된 시계열 데이터의 시간정보를 추출하는 시간정보 추출기, 상기 시간정보에 따라 시계열 인덱스를 생성하는 인덱스 생성기 및, 상기 시간정보에서 로(row) 데이터를 추출하고, 상기 로(row) 데이터와 상기 시계열 인덱스를 포함하여 로그 파일을 생성하는 로그 파일 생성기를 포함하며, 상기 로그 파일을 제1 기간까지 저장하여 상기 제1 기간이 도과하면 상기 제2 수집 서버 그룹의 수집 서버에 제공하고,

상기 제2 수집 서버 그룹의 수집 서버는 제공된 로그 파일을 제2 기간까지 저장하며,

상기 데이터 검색부는,

상기 검색 요청에 대하여 상기 제1 수집 서버 그룹과 연동하여 우선적으로 검색을 수행하며, 상기 제1 수집 서버 그룹의 모든 수집 서버에 대하여 검색이 완료된 후 상기 제2 수집 서버 그룹과 연동하여 검색을 수행하는 분석 서버를 포함하고,

상기 분석 서버는,

상기 제1 및 제2 수집 서버 그룹의 수집 서버간 계층 구조 정보를 저장하는 구조 정보 저장기, 상기 계층 구조 정보를 생성하여 상기 구조 정보 저장기에 저장하는 구조 정보 관리기 및, 사용자의 요청에 따라 상기 계층 구조 정보를 확인하여 검색을 수행할 수집 서버를 결정하고, 순차적으로 검색을 수행하는 검색기

를 포함하는 데이터 검색 시스템.

청구항 2

삭제

청구항 3

제1항에 있어서, 상기 데이터 수집부는

상기 제2 기간이 도과된 후, 상기 제2 수집 서버 그룹으로부터 상기 로그 파일을 제공받아 저장하는 제3 수집 서버 그룹;

을 더 포함하고,

상기 제3 수집 서버 그룹에 포함되는 적어도 하나의 수집 서버는

상기 제2 수집 서버 그룹에 포함되는 복수의 수집 서버와 1:N(N은 2 이상의 자연수)의 관계로 대응되는 데이터 검색 시스템.

청구항 4

삭제

청구항 5

삭제

청구항 6

삭제

청구항 7

청구항 1에 기재된 데이터 검색 시스템에 의한 데이터 검색방법으로서,

성능차이를 갖는 저장매체를 사용하며, 계층적으로 구분되는 복수의 수집 서버로 이루어지는 제1 및 제2 데이터 수집 서버 그룹이 시계열 데이터를 저장 기간에 따라 분산하여 저장하는 단계;

데이터 검색부가 상기 복수의 수집 서버에 대한 계층적인 구조 정보를 생성 또는 유지하는 단계; 및

사용자의 검색 요청에 대하여, 상기 데이터 검색부가 상기 계층적인 구조 정보를 참조하여 상기 복수의 수집 서버 중, 검색을 수행할 적어도 하나의 수집 서버를 선택하여 검색을 수행하는 단계를 포함하고,

상기 데이터 검색부가 하나의 수집 서버를 선택하여 검색을 수행하는 단계는,

분석 서버의 구조 정보 관리기가 상기 제1 및 제2 수집 서버 그룹의 수집 서버간 계층 구조를 생성하여 상기 분석 서버의 구조 정보 저장기에 저장하는 단계; 및

사용자의 요청에 따라, 상기 분석 서버의 검색기가 상기 구조 정보 저장기에 저장된 계층 구조를 확인하여 검색을 수행할 수집 서버를 결정하고, 순차적으로 검색을 수행하는 단계

를 포함하는 데이터 검색 방법.

청구항 8

삭제

청구항 9

제7항에 있어서,

상기 복수의 수집 서버가 시계열 데이터를 분산하여 저장하는 단계는

상기 제1 수집 서버 그룹에 포함된 수집 서버가 제1 기간까지 로그 파일을 저장하는 단계;

상기 제1 기간이 도과되면, 상기 제1 수집 서버 그룹에 포함된 수집 서버가 저장된 로그 파일을 상기 제2 수집 서버 그룹에 포함된 수집서버에 제공하는 단계; 및

상기 제2 수집 서버 그룹에 포함된 수집서버가 제공된 로그파일을 저장하는 단계

를 포함하는 데이터 검색 방법.

청구항 10

삭제

발명의 설명

기 술 분 야

[0001] 본 발명은 시계열 계층 인덱싱을 이용한 데이터 검색 시스템 및 데이터 검색 방법에 관한 것이다.

배 경 기 술

[0002] 데이터의 꾸준한 증가에 따라, 대용량 데이터들에 대한 검색이나 처리가 중요한 이슈가 되고 있다. 또한, 데이터의 효과적인 처리를 위하여, 데이터의 시간 정보가 반영된 시계열 데이터에 대한 검색이 중요한 이슈가 되고 있다.

[0003] 종래에는, 이러한 대용량의 시계열 데이터 처리를 위하여 분산 처리 기술 등이 적용되고 있었다. 그러나, 이러한 종래의 시계열 데이터 처리 방식은 분산 처리 등을 기반으로 동작되므로, 검색 시스템 자체의 규모를 크게 요구하거나, 또는 분산 처리 대상의 일차적인 관리에 의하여 검색 효율의 비효율이 존재하는 문제점이 있다.

[0004] 이는, 특히 데이터의 시계열적인 특성이 부여될 때, 단순 분산 처리 방식에 의한 경우 검색 비효율이 증가되는 문제가 있다.

[0005] 이러한 종래기술에 대해서는, 한국 등록특허공보 제10-1426673호, 한국 등록특허공보 제10-0898465호 또는 한국 공개특허공보 제2002-0066836호 등을 통하여 보다 쉽게 이해할 수 있다.

선행기술문헌

특허문헌

[0006] (특허문헌 0001) 한국 등록특허공보 제10-1426673호

(특허문헌 0002) 한국 등록특허공보 제10-0898465호

(특허문헌 0003) 한국 공개특허공보 제2002-0066836호

발명의 내용

해결하려는 과제

[0007] 본 발명은 상기한 종래 기술의 문제점을 해결하기 위한 것으로써, 시계열 데이터의 검색을 최소의 자원으로도 효율적이고 빠르게 수행할 수 있는 시계열 계층 인덱싱을 이용한 데이터 검색 시스템 및 데이터 검색 방법을 제공하는 것이다.

과제의 해결 수단

[0008] 본 발명의 일 기술적 측면은 데이터 검색 시스템을 제안한다. 상기 데이터 검색 시스템은, 수집된 시계열 데이터의 시간 정보를 이용하여 시계열 인덱스를 생성하고, 상기 시계열 인덱스를 포함하는 로그 파일을 생성하여, 저장 기간에 따라 계층적으로 구분되는 복수의 수집 서버에 분산하여 저장하는 데이터 수집부 및 상기 복수의 수집 서버에 대한 계층적인 구조 정보를 유지하고, 사용자의 검색 요청에 대하여 상기 계층적인 구조 정보를 참조하여 검색을 수행할 적어도 하나의 수집 서버를 선택하는 데이터 검색부를 포함할 수 있다.

[0009] 일 실시예에서, 상기 데이터 수집부는, 상기 로그 파일을 제1 기간까지 저장하는 제1 수집 서버 그룹 및 상기 제1 기간이 도과된 후, 상기 제1 수집 서버 그룹으로부터 상기 로그 파일을 제공받아 제2 기간까지 저장하는 제2 수집 서버 그룹을 포함할 수 있다.

[0010] 일 실시예에서, 상기 데이터 수집부는, 상기 제2 기간이 도과된 후, 상기 제2 수집 서버 그룹으로부터 상기 로그 파일을 제공받아 저장하는 제3 수집 서버 그룹을 더 포함하고, 상기 제3 수집 서버 그룹에 포함되는 적어도 하나의 수집 서버는 상기 제2 수집 서버 그룹에 포함되는 복수의 수집 서버와 1:N(N은 2 이상의 자연수)의 관계로 대응될 수 있다.

[0011] 일 실시예에서, 상기 데이터 검색부는, 상기 검색 요청에 대하여 상기 제1 수집 서버 그룹과 연동하여 우선적으로 검색을 수행하고, 상기 제1 수집 서버 그룹의 모든 수집 서버에 대하여 검색이 완료된 후 상기 제2 수집 서버 그룹과 연동하여 검색을 수행하는 분석 서버를 포함할 수 있다.

[0012] 일 실시예에서, 상기 수집 서버는, 상기 시계열 데이터의 시간 정보를 추출하는 시간 정보 추출기, 상기 시간 정보에 따라 시계열 인덱스를 생성하는 인덱스 생성기 및 상기 시계열 데이터에서 로 데이터를 추출하고, 상기 로 데이터와 상기 시계열 인덱스를 포함하여 상기 로그 파일을 생성하는 로그 파일 생성기를 포함할 수 있다.

[0013] 일 실시예에서, 상기 분석 서버는, 상기 복수의 수집 서버에 대한 계층적인 구조 정보를 생성 또는 갱신하는 구조 정보 관리기 및 상기 구조 정보를 확인하여 검색을 수행할 적어도 하나의 수집 서버를 결정하고, 상기 적어도 하나의 수집 서버와 연동하여 상기 사용자의 검색 요청에 따른 검색을 수행하는 검색기를 포함할 수 있다.

[0014] 본 발명의 다른 일 기술적 측면은 데이터 검색 방법을 제안한다. 상기 데이터 검색 방법은, 수집된 시계열 데이터를, 저장 기간에 따라 계층적으로 구분되는 복수의 수집 서버에 분산하여 저장하는 단계, 상기 복수의 수집 서버에 대한 계층적인 구조 정보를 생성 또는 유지하는 단계 및 사용자의 검색 요청에 대하여 상기 계층적인 구조 정보를 참조하여 검색을 수행할 적어도 하나의 수집 서버를 선택하여 검색을 수행하는 단계를 포함할 수 있다.

[0015] 일 실시예에서, 상기 복수의 수집 서버에 분산하여 저장하는 단계는, 상기 시계열 데이터의 시간 정보를 확인하여 시계열 인덱스를 생성하는 단계, 상기 시계열 인덱스 및 상기 시계열 데이터의 로 데이터를 포함하여 로그 파일을 생성하는 단계 및 상기 로그 파일을 복수의 수집 서버에 저장 기간에 따라 구분하여 저장하는 단계를 포함할 수 있다.

[0016] 일 실시예에서, 상기 복수의 수집 서버는, 제1 수집 서버 그룹 및 제2 수집 서버 그룹을 포함하고, 상기 복수의 수집 서버에 분산하여 저장하는 단계는, 제1 기간까지 상기 로그 파일을 상기 제1 수집 서버 그룹에 저장하는 단계 및 상기 제1 기간이 도과되면, 상기 로그 파일을 상기 제1 수집 서버 그룹에서 삭제하고, 상기 제2 수집 서버에 저장하는 단계를 포함할 수 있다.

[0017] 일 실시예에서, 상기 검색을 수행하는 단계는, 상기 검색 요청에 대하여 상기 제1 수집 서버 그룹과 연동하여 우선적으로 검색을 수행하는 단계 및 상기 제1 수집 서버 그룹의 모든 수집 서버에 대하여 검색이 완료된 후 상기 제2 수집 서버 그룹과 연동하여 검색을 수행하는 단계를 포함할 수 있다.

[0018] 상기한 과제의 해결 수단은, 본 발명의 특징을 모두 열거한 것은 아니다. 본 발명의 과제 해결을 위한 다양한 수단들은 이하의 상세한 설명의 구체적인 실시형태를 참조하여 보다 상세하게 이해될 수 있을 것이다.

발명의 효과

[0019] 본 발명의 일 실시형태에 따르면, 시계열 데이터의 검색을 최소의 자원으로도 효율적이고 빠르게 수행할 수 있는 효과가 있다.

[0020] 또한, 본 발명의 일 실시형태에 따르면, 시계열 데이터를 저장 기간에 따라 계층적으로 구분하여 분산 저장하므로, 최신 정보가 우선되는 시계열 데이터의 특성에 따라 보다 효율적인 검색 시스템을 구축할 수 있는 효과가 있다.

[0021] 또한, 본 발명의 일 실시형태에 따르면, 시계열 데이터에 대한 계층 구분에 있어서 계층의 단계에 따라 차등적인 서버 개수를 가지므로, 시스템 구축의 효율성을 증대할 수 있는 효과가 있다.

도면의 간단한 설명

[0022] 도 1은 본 발명의 일 실시예에 따른 데이터 검색 시스템을 설명하기 위한 블록 구성도이다.

도 2는 도 1에 도시된 데이터 수집부의 일 예를 설명하기 위한 블록 구성도이다.

도 3은 도 1에 도시된 데이터 수집부의 다른 일 예를 설명하기 위한 블록 구성도이다.

도 4는 본 발명의 일 실시예에 따른 수집 서버의 일 예를 설명하기 위한 블록 구성도이다.

도 5는 로그 파일의 일 예를 설명하기 위한 도면이다.

도 6은 본 발명의 일 실시예에 따른 분석 서버 의 일 예를 설명하기 위한 블록 구성도이다.

도 7은 본 발명의 일 실시예에 따른 데이터 검색 방법을 설명하기 위한 블록 구성도이다.

발명을 실시하기 위한 구체적인 내용

[0023] 이하, 첨부된 도면을 참조하여 본 발명의 바람직한 실시 형태들을 설명한다.

[0024] 그러나, 본 발명의 실시형태는 여러 가지 다른 형태로 변형될 수 있으며, 본 발명의 범위가 이하 설명하는 실시형태로 한정되는 것은 아니다. 또한, 본 발명의 실시형태는 당해 기술분야에서 평균적인 지식을 가진 자에게 본 발명을 더욱 완전하게 설명하기 위해서 제공되는 것이다.

[0025] 도 1은 본 발명의 일 실시예에 따른 데이터 검색 시스템을 설명하기 위한 블록 구성도이다.

[0026] 도 1을 참조하면, 데이터 검색 시스템(100)은 데이터 수집부(110)와 데이터 검색부(120)를 포함할 수 있다.

[0027] 데이터 수집부(110)는 수집된 시계열 데이터의 시간 정보를 이용하여 시계열 인덱스를 생성하고, 시계열 인덱스를 포함하는 로그 파일을 생성할 수 있다. 데이터 수집부(110)는 생성된 로그 파일을 복수의 수집 서버에 분산하여 저장할 수 있다.

[0028] 데이터 수집부(110)는 저장 기간에 따라 계층적으로 구분되는 복수의 수집 서버를 포함할 수 있다. 이에 대해서는 이하의 도 2 내지 도 5을 참조하여 설명한다.

[0029] 데이터 검색부(120)는 복수의 수집 서버에 대한 계층적인 구조 정보를 유지할 수 있다. 데이터 검색부(120)는 사용자의 검색 요청에 대하여 계층적인 구조 정보를 참조하여 검색을 수행할 적어도 하나의 수집 서버를 선택할 수 있으며, 선택된 수집 서버와 연동하여 사용자의 검색 요청에 대한 검색 결과를 제공할 수 있다.

[0030] 데이터 검색부(120)는 수집 서버와 연동하여 검색을 수행하는 분석 서버를 포함할 수 있다. 이러한 데이터 검색부(120)에 대해서는 이하의 도 6을 참조하여 설명한다.

[0031] 도 2는 도 1에 도시된 데이터 수집부의 일 예를 설명하기 위한 블록 구성도이다.

[0032] 도 2를 참조하면, 데이터 수집부(110)는 시계열 데이터의 종류(데이터 A 내지 데이터 D)에 대하여, 각각 계층적인 수집 서버가 구비된 예를 도시하고 있다.

[0033] 데이터 수집부(110)는 제1 수집 서버 그룹(210)과 제2 수집 서버 그룹(220)을 포함할 수 있다. 제1 수집 서버 그룹(210)과 제2 수집 서버 그룹(220)은 데이터의 저장 기간의 차이에 따라서 구별될 수 있다.

[0034] 제1 수집 서버 그룹(210)은 복수의 수집 서버(A1 내지 D1)을 포함하며, 복수의 수집 서버(A1 내지 D1)는 시계열 데이터로부터 로그 파일을 생성하고, 생성한 로그 파일을 제1 기간까지 저장할 수 있다. 제1 수집 서버 그룹(210)에 포함된 복수의 수집 서버(A1 내지 D1)는 제1 기간이 도과하면 로그 파일을 자신의 하위 계층에 있는, 제2 수집 서버 그룹(220)에 포함된 복수의 수집 서버(A2 내지 D2)에 제공할 수 있다.

[0035] 제2 수집 서버 그룹(220) 에 포함된 복수의 수집 서버(A2 내지 D2)는 제1 기간이 도과된 후, 상기 제1 수집 서버 그룹으로부터 상기 로그 파일을 제공받아 제2 기간까지 저장할 수 있다.

[0036] 예를 들어, 제1 수집 서버 그룹(210)의 각 수집 서버(A1 내지 D1)들은 3개월까지 로그 파일을 저장하고, 3개월이 경과하면 이를 제2 수집 서버 그룹(220)에 포함된 복수의 수집 서버(A2 내지 D2)에 각각 제공할 수 있다.

[0037] 제1 수집 서버 그룹(210)의 수집 서버들(A1 내지 D1)은, 제2 수집 서버 그룹(220)의 수집 서버들(A2 내지 D2)보다 높은 성능을 가질 수 있다. 예컨대, 제1 수집 서버 그룹(210)의 수집 서버들(A1 내지 D1)는 SSD(Solid State Disk) 등의 고속 저장 매체를 사용할 수 있고, 제2 수집 서버 그룹(220)의 수집 서버들(A2 내지 D2)은 HDD(Hard Disk Drive) 등 일반적인 저장 매체를 사용할 수 있다.

[0038] 이와 같이, 시계열적 데이터의 시간 정보에 따라 계층적으로 구성된 복수의 서버 구조를 이용하게 되므로, 비교적 최근의 데이터에 대한 검색 요구가 많은 시계열 데이터의 검색 환경을 개선할 수 있다.

[0039] 한편, 데이터 검색부(120, 도 1에 도시됨)는 사용자로부터 검색 요청을 입력받고, 데이터 수집부(110)의 계층적 구조에 대응하여 검색을 수행할 수 있다. 즉, 데이터 검색부(120, 도 1에 도시됨)는 사용자의 검색 요청을 수신하면, 먼저 제1 수집 서버 그룹(210)과 연동하여 우선적으로 검색을 수행할 수 있다. 제1 수집 서버 그룹(210)의 모든 수집 서버에 대하여 검색이 완료된 후, 데이터 검색부(120, 도 1에 도시됨)는 제2 수집 서버 그룹(220)과 연동하여 검색을 수행할 수 있다.

[0040] 도 3은 도 1에 도시된 데이터 수집부의 다른 일 예를 설명하기 위한 블록 구성도이다.

[0041] 도 3을 참조하면, 데이터 수집부(111)는 제1 수집 서버 그룹(310), 제2 수집 서버 그룹(320) 및 제3 수집 서버 그룹(330)을 포함할 수 있다.

[0042] 제1 수집 서버 그룹(310)은 복수의 수집 서버(A1 내지 D1)을 포함하며, 복수의 수집 서버(A1 내지 D1)는 시계열 데이터(데이터 A 내지 데이터 D)로부터 각각 로그 파일을 생성하고, 생성한 로그 파일을 제1 기간까지 저장할 수 있다. 제1 수집 서버 그룹(310)에 포함된 복수의 수집 서버(A1 내지 D1)는 제1 기간이 도과하면, 로그 파일을 자신의 하위 계층에 있는, 제2 수집 서버 그룹(320)에 포함된 복수의 수집 서버(A2 내지 D2)에 제공할 수 있다.

[0043] 제2 수집 서버 그룹(320) 에 포함된 복수의 수집 서버(A2 내지 D2)는 제1 기간이 도과된 후, 상기 제1 수집 서버 그룹으로부터 상기 로그 파일을 제공받아 제2 기간까지 저장할 수 있다. 제2 수집 서버 그룹(320)은 제2 기간이 도과하면, 로그 파일을 자신의 하위 계층에 있는, 제3 수집 서버 그룹(330)에 제공할 수 있다.

[0044] 제3 수집 서버 그룹(330)은 제2 기간이 도과된 후, 제2 수집 서버 그룹(320)으로부터 로그 파일을 제공받아 저장할 수 있다.

[0045] 예를 들어, 예를 들어, 제1 수집 서버 그룹(310)의 각 수집 서버(A1 내지 D1)들은 3개월까지 로그 파일을 저장하고, 제2 수집 서버 그룹(320)에 포함된 복수의 수집 서버(A2 내지 D2)는 3개월이 경과된 로그 파일을 6개월까지 저장할 수 있다. 6개월이 경과된 로그 파일은 제3 수집 서버 그룹(330)에 저장될 수 있다.

[0046] 제3 수집 서버 그룹(330)에 포함된 적어도 하나의 수집 서버(A3 내지 C3)는, 제2 서버 그룹(320)에 포함된 복수의 수집 서버(A2 내지 D2)와 1:1로 대응되지 않을 수 있다. 즉, 제3 수집 서버 그룹(330)에 포함되는 적어도 하나의 수집 서버(A3 내지 C3)는, 제2 수집 서버 그룹(320)에 포함되는 복수의 수집 서버(A2 내지 D2)와 1:N(N은 2 이상의 자연수)의 관계로 대응될 수 있다. 도 3에서는 1:2의 비율로 대응되는 예를 도시하고 있다.

[0047] 즉, 데이터 수집부(111)는 저장 기간에 따라 계층적으로 구분되는 복수의 수집 서버를 구비하며, 그러한 저장 기간에 따른 계층이 낮을 수록, 즉, 제1 수집 서버 그룹(310)에서 제3 수집 서버 그룹(330)으로 내려감에 따라 차등적인 서버 설정을 가질 수 있다. 예컨대, 도시된 바와 같이 제3 수집 서버 그룹(330)은 타 서버 그룹(310, 320)에 비해 더 적은 수의 서버로 이루어질 수 있다.

[0048] 이는, 최신 데이터에 대한 활용도가 높게 요구되는 시계열 데이터의 특성에 대응한 구조로서, 이러한 구조에 의하여 검색 속도 및 검색 시스템의 구성 효율을 증대시킬 수 있다.

[0049] 이상에서는, 도 2 내지 도 3을 참조하여, 시계열 특성에 대응하여 구성된, 복수의 수집 서버의 계층적인 구조에 대하여 설명하였다. 이하, 도 4 내지 도 5를 참조하여, 이러한 수집 서버의 일 구성예에 대하여 설명한다.

[0050] 도 4는 본 발명의 일 실시예에 따른 수집 서버의 일 예를 설명하기 위한 블록 구성도이고, 도 5는 로그 파일의 일 예를 설명하기 위한 도면이다.

[0051] 도 4를 참조하면, 수집 서버는 시간 정보 추출기(410), 인덱스 생성기(420) 및 로그 파일 생성기(430)를 포함할 수 있다.

[0052] 시간 정보 추출기(410)는 입력받은 시계열 데이터의 시간 정보를 추출할 수 있다. 시간 정보는, 예를 들어, 이벤트 시간이나, 수집 시간일 수 있다.

[0053] 인덱스 생성기(420)는 상기 시간 정보에 따라 시계열 인덱스를 생성할 수 있다.

[0054] 로그 파일 생성기(430)는 상기 시계열 데이터에서 로(raw) 데이터를 추출하고, 상기 로 데이터와 상기 시계열 인덱스를 포함하여 로그 파일을 생성할 수 있다.

[0055] 도 5를 더 참조하면, 로그 파일은 데이터 식별 정보, 시계열 인덱스 및 로 데이터를 포함할 수 있다. 데이터 식별 정보는 데이터의 종류 등을 구분하는 정보이며, 이는 도 2 내지 도 3의 예에서 데이터 A 내지 데이터 D로 표현된다. 로그 파일 생성기(430)는 시계열 데이터의 종류 등을 구분하여 데이터 식별 정보를 생성할 수 있다.

[0056] 시계열 인덱스는 인덱스 생성기(420)에 의하여 생성되며, 시계열 인덱스는 저장 기간을 판단하는데 사용된다.

[0057] 로 데이터는 시계열 데이터의 내용에 해당하며, 로그 파일 생성기(430)에 의하여 생성된다.

[0058] 이상에서는 도 2 내지 도 5를 참조하여, 데이터 수집부의 다양한 예들에 대하여 설명하였다. 이하에서는 도 6을 참조하여 데이터 검색부에 포함되는 분석 서버에 대하여 보다 상세히 설명한다.

[0059] 도 6은 본 발명의 일 실시예에 따른 분석 서버 의 일 예를 설명하기 위한 블록 구성도이다.

[0060] 도 6을 참조하면, 분석 서버는 구조 정보 관리기(610), 구조 정보 저장기(620) 및 검색기(630)를 포함할 수 있다.

[0061] 구조 정보 관리기(610)는 데이터 수집부에 포함되는 복수의 수집 서버에 대한 계층적인 구조 정보를 생성 또는 갱신할 수 있다. 구조 정보 관리기(610)에 의해 생성 또는 갱신되는 구조 정보는 구조 정보 저장기(620)에 저장될 수 있다.

[0062] 즉, 구조 정보 관리기(610)는 도 2 또는 도 3에서 예시한 바와 같은 저장 기간에 따라 계층적으로 구분되는 복수의 수집 서버의 계층적인 구조에 대한 정보를 유지할 수 있다.

[0063] 검색기(630)는 구조 정보를 확인하여, 검색을 수행할 적어도 하나의 수집 서버를 결정할 수 있다. 검색기(630)는 결정된 적어도 하나의 수집 서버와 연동하여 사용자의 검색 요청에 따른 검색을 수행할 수 있다.

[0064] 예를 들어, 검색기(630)는 사용자의 검색 요청에 대해 검색 기간에 대한 제한이 없는 경우, 데이터 수집부의 계층을 순차적으로 검색할 수 있다.

[0065] 다른 예를 들어, 검색기(630)는 사용자의 검색 요청에 대해 검색 기간에 대한 제한이 있는 경우, 해당 검색 기간에 대응되는 수집 서버를 선택하여 검색을 수행할 수 있다.

[0066] 이와 같이, 분석 서버는 복수의 수집 서버에 대한 계층적인 구조 정보를 기반으로, 시계열 데이터의 검색 효율을 증대시킬 수 있다.

[0067] 이상에서는 도 1 내지 도 6을 참조하여, 본 발명의 일 실시예에 따른 데이터 검색 시스템에 대하여 설명하였다. 이하에서는, 도 7을 참조하여 본 발명의 일 실시예에 따른 데이터 검색 방법에 대하여 설명한다. 다만, 이하에서 설명할 데이터 검색 방법은 상술한 데이터 검색 시스템에서 수행되므로, 도 1 내지 도 6에서 상술한 설명을 참조하여 보다 쉽게 이해할 수 있다.

[0068] 도 7은 본 발명의 일 실시예에 따른 데이터 검색 방법을 설명하기 위한 블록 구성도이다.

[0069] 도 7을 참조하면, 데이터 검색 시스템은, 수집된 시계열 데이터를, 저장 기간에 따라 계층적으로 구분되는 복수의 수집 서버에 분산하여 저장할 수 있다(S710).

[0070] 데이터 검색 시스템은 상기 복수의 수집 서버에 대한 계층적인 구조 정보를 생성 또는 유지할 수 있다(S720).

[0071] 데이터 검색 시스템은 사용자의 검색 요청에 대하여 상기 계층적인 구조 정보를 참조하여 검색을 수행할 적어도 하나의 수집 서버를 선택하고(S730), 선택된 수집 서버와 연동하여 검색을 수행할 수 있다(S740).

[0072] 단계 S710에 대한 일 실시예에서, 데이터 검색 시스템은 시계열 데이터의 시간 정보를 확인하여 시계열 인덱스를 생성할 수 있다. 이후, 데이터 검색 시스템은 시계열 인덱스 및 시계열 데이터의 로 데이터를 포함하여 로그 파일을 생성할 수 있다. 데이터 검색 시스템은 로그 파일을 복수의 수집 서버에 저장 기간에 따라 구분하여 저장할 수 있다.

[0073] 일 실시예에서, 상기 복수의 수집 서버는 제1 수집 서버 그룹 및 제2 수집 서버 그룹을 포함할 수 있다. 데이터 검색 시스템은 제1 기간까지 상기 로그 파일을 상기 제1 수집 서버 그룹에 저장할 수 있다. 상기 제1 기간이 도과되면, 데이터 검색 시스템은 로그 파일을 제1 수집 서버 그룹에서 삭제하고, 제2 수집 서버에 저장할 수 있다.

[0074] 단계 S740에 대한 일 실시예에서, 데이터 검색 시스템은 사용자의 검색 요청에 대하여 상기 제1 수집 서버 그룹과 연동하여 우선적으로 검색을 수행할 수 있다. 데이터 검색 시스템은 상기 제1 수집 서버 그룹의 모든 수집 서버에 대하여 검색이 완료된 후, 상기 제2 수집 서버 그룹과 연동하여 검색을 수행할 수 있다.

[0075] 이상에서 설명한 본 발명은 전술한 실시예 및 첨부된 도면에 의해 한정되는 것이 아니고 후술하는 특허청구범위에 의해 한정되며, 본 발명의 구성은 본 발명의 기술적 사상을 벗어나지 않는 범위 내에서 그 구성을 다양하게 변경 및 개조할 수 있다는 것을 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 쉽게 알 수 있다.

부호의 설명

[0076] 100 : 데이터 검색 시스템

110 : 데이터 수집부 120 : 데이터 검색부

410 : 시간 정보 추출기 420 : 인덱스 생성기

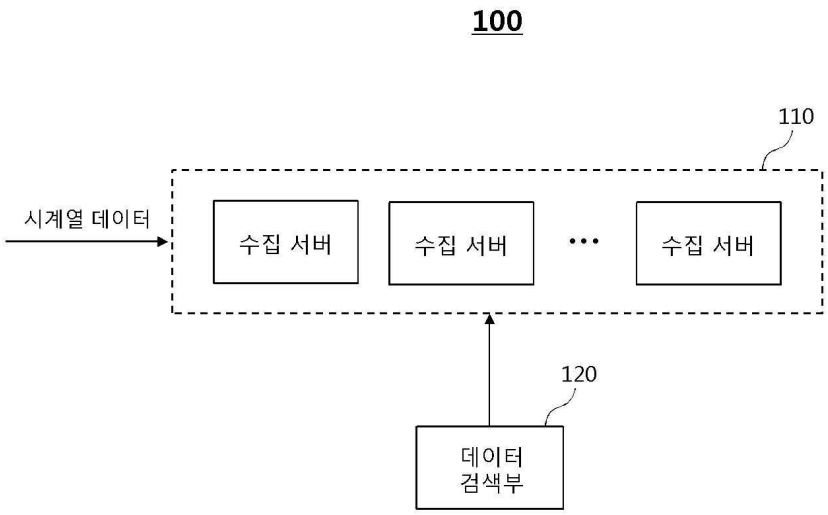
430 : 로그 파일 생성기

610 : 구조 정보 관리기 620 : 구조 정보

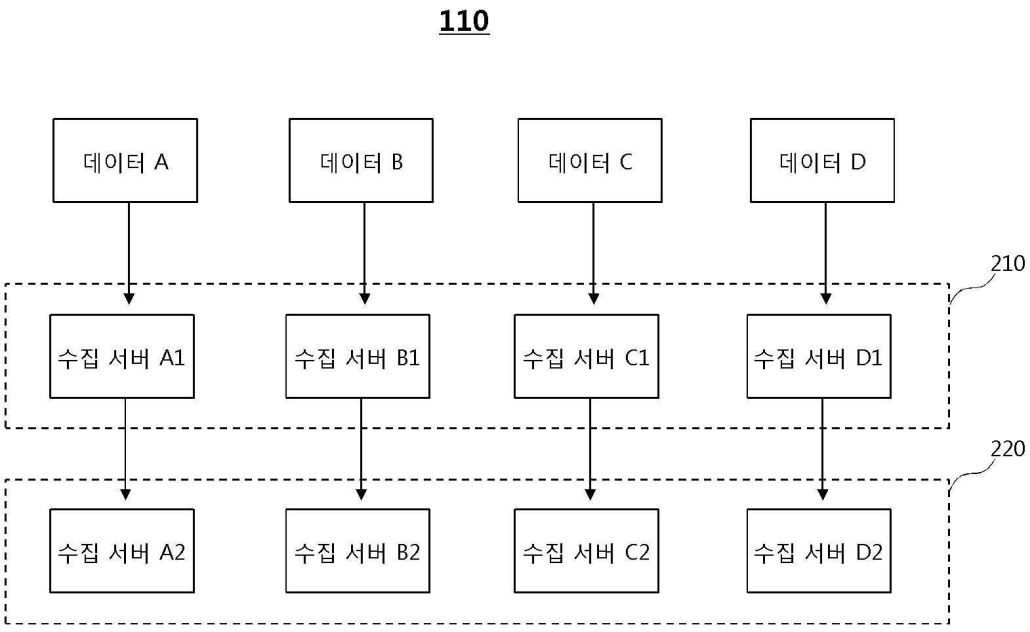
630 : 검색기

도면

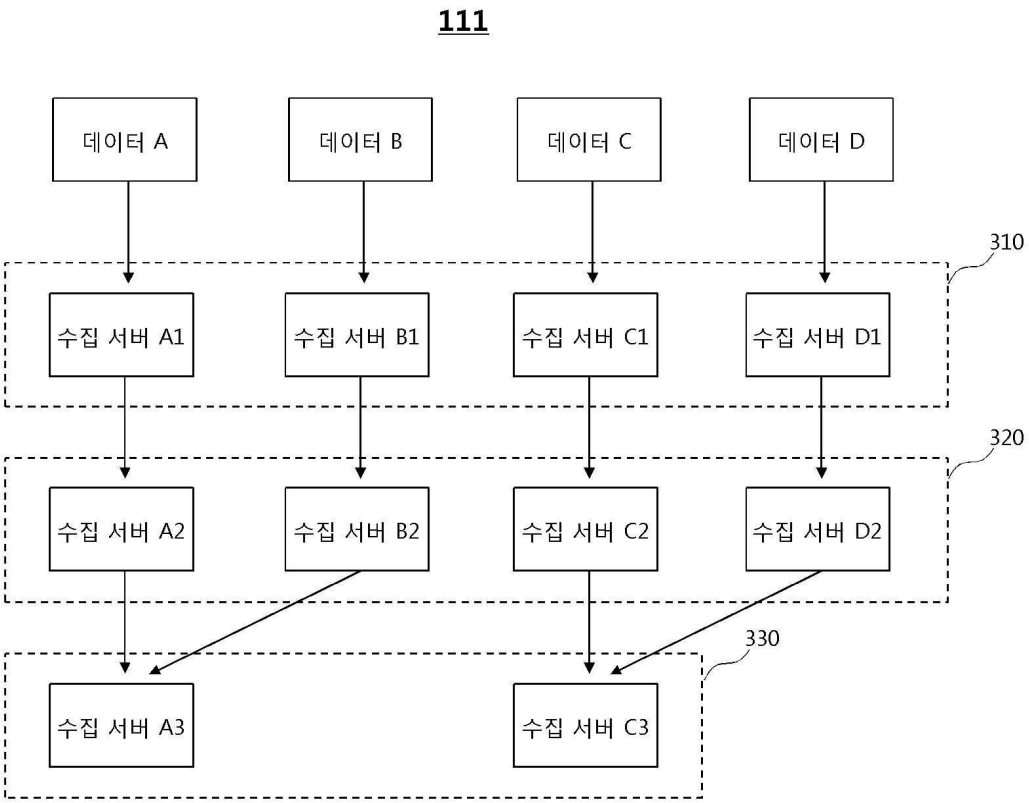
도면1


100
110
시계열 데이터
수집 서버
수집 서버
수집 서버
120
데이터
검색부

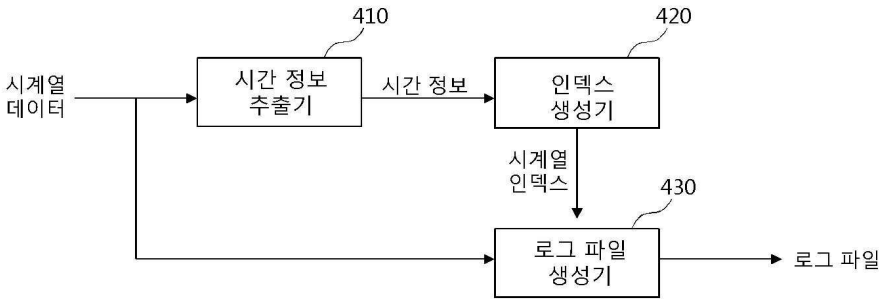
도면2


110
데이터 A
데이터 B
데이터 C
데이터 D
210
수집 서버 A1
수집 서버 B1
수집 서버 C1
수집 서버 D1
220
수집 서버 A2
수집 서버 B2
수집 서버 C2
수집 서버 D2

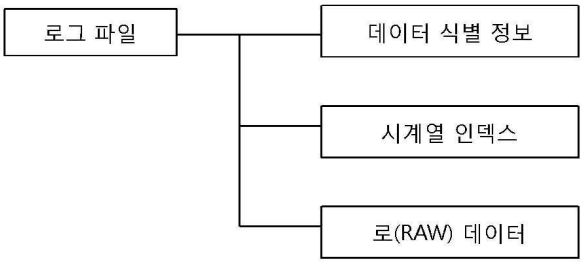
도면3


111
데이터 A
데이터 B
데이터 C
데이터 D
310
수집 서버 A1
수집 서버 B1
수집 서버 C1
수집 서버 D1
320
수집 서버 A2
수집 서버 B2
수집 서버 C2
수집 서버 D2
330
수집 서버 A3
수집 서버 C3

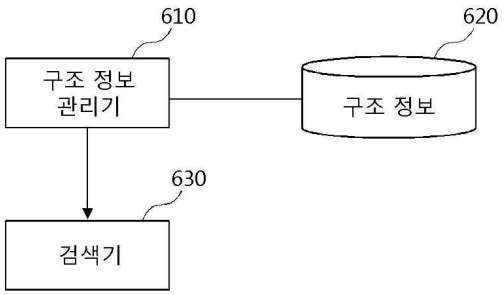
도면4


410
420
시계열
데이터
시간 정보
추출기
시간 정보
인덱스
생성기
시계열
인덱스
430
로그 파일
생성기
로그 파일

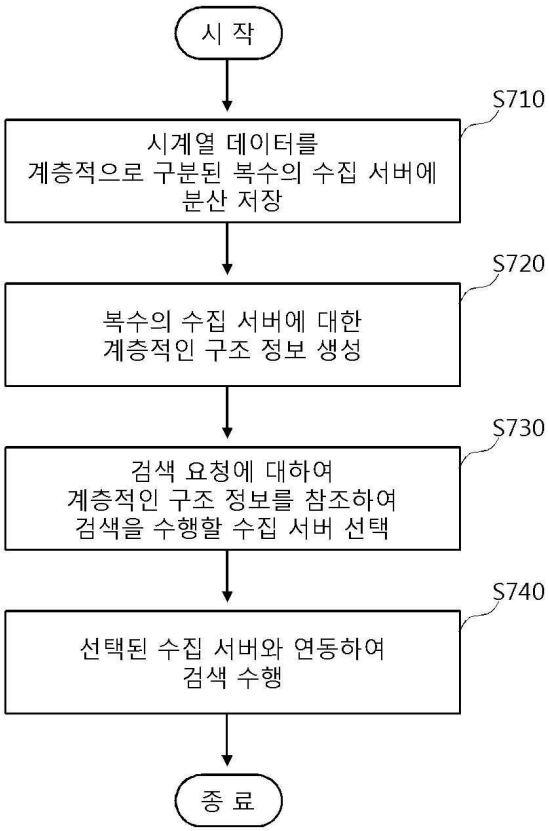
도면5


로그 파일
데이터 식별 정보
시계열 인덱스
로(RAW) 데이터

도면6


610
620
구조 정보
관리기
구조 정보
630
검색기

도면7


시 작
S710
시계열 데이터를
계층적으로 구분된 복수의 수집 서버에
분산 저장
S720
복수의 수집 서버에 대한
계층적인 구조 정보 생성
S730
검색 요청에 대하여
계층적인 구조 정보를 참조하여
검색을 수행할 수집 서버 선택
S740
선택된 수집 서버와 연동하여
검색 수행
종 료