



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2017년04월18일
 (11) 등록번호 10-1727508
 (24) 등록일자 2017년04월11일

(51) 국제특허분류(Int. Cl.)
 G06F 17/30 (2006.01) H03M 7/30 (2006.01)
 (52) CPC특허분류
 G06F 17/30318 (2013.01)
 G06F 17/30194 (2013.01)
 (21) 출원번호 10-2015-0106457
 (22) 출원일자 2015년07월28일
 심사청구일자 2015년07월28일
 (65) 공개번호 10-2017-0014042
 (43) 공개일자 2017년02월08일
 (56) 선행기술조사문헌
 US20150172209 A1
 US20140258650 A1
 KR1020070092720 A
 WO2014105323 A1

(73) 특허권자
 서울과학기술대학교 산학협력단
 서울특별시 노원구 공릉로 232 (공릉동, 서울과학기술대학교)
 (72) 발명자
 장지훈
 경기도 남양주시 순화궁로 18 별내신안인스빌아파트 4115동1902호
 이승은
 서울특별시 강서구 허준로 23 106동 601호 (가양동, 한강타운아파트)
 (뒷면에 계속)
 (74) 대리인
 원대규

전체 청구항 수 : 총 3 항

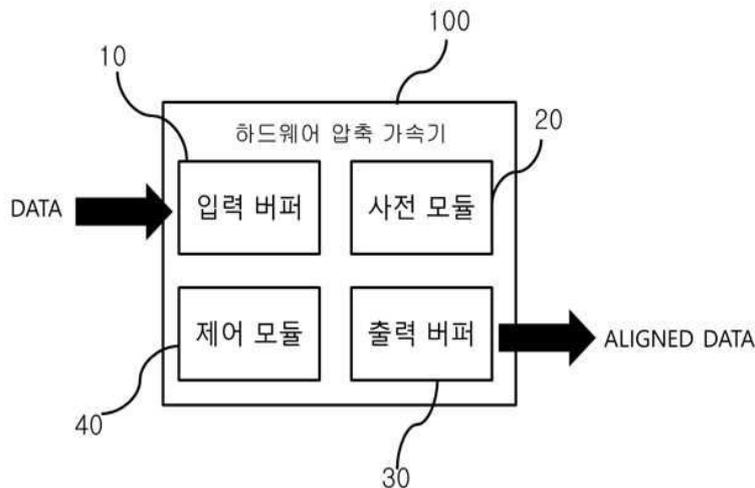
심사관 : 윤혜숙

(54) 발명의 명칭 하둡 기반 하드웨어 압축 고속화 장치 및 방법

(57) 요약

본 발명은 하둡 기반의 하드웨어 압축 고속화 장치에 관한 것이다. 본 발명은 저전력 하둡 스토리지 어플라이언스에서 하둡 미들웨어가 수행하는 압축 및 압축해제 과정을 하드웨어를 통해 수행하여 저전력 CPU의 성능을 보완한다. 이를 위해, 본 발명에 따른 하둡 기반의 하드웨어 압축 고속화 장치는 압축 또는 압축해제가 되어야 하는 데이터 블록을 수신하는 입력 버퍼와 사전 등록 및 검색을 수행하고, 윈도우를 통해 상기 데이터 블록에 대한 압축을 수행하는 사전 모듈과 상기 압축이 수행된 결과를 출력하는 출력버퍼 및 하둡 스토리지 어플라이언스 정보를 기초로 상기 입력 버퍼, 상기 사전 모듈 및 상기 출력 버퍼를 제어하는 제어모듈을 포함할 수 있다.

대표도 - 도1



(52) CPC특허분류
H03M 7/30 (2013.01)

임동일

서울특별시 노원구 동일로 1077 (공릉동) 스테이션 빌딩 802호

(72) 발명자

이현화

서울특별시 양천구 신정로 293 신트리1단지아파트 106동 1405호

한재용

인천광역시 계양구 봉오대로 517번길 12-1(효성동) 풍림아파트 102동 2003호

이 발명을 지원한 국가연구개발사업

과제고유번호 10047088

부처명 미래창조과학부

연구관리전문기관 정보통신 기술 진흥 센터 (IITP)

연구사업명 SW컴퓨팅산업원천기술개발

연구과제명 단일 노드 48TB 이상을 지원하는 개방형 하둡 스토리지 어플라이언스(Hadoop Storage Appliance) 개발 / (영문 : Development of open type hadoop storage appliance to support more than 48TB per single data node)

기여율 1/1

주관기관 디포커스(주)

연구기간 2013.09.01 ~ 2016.08.31

명세서

청구범위

청구항 1

사전 모듈의 윈도우와 같은 너비의 엔디안으로 데이터를 정렬하는 입력 버퍼가 맵리듀스 (MapReduce) 태스크로부터 압축 또는 압축해제가 되어야 하고 데이터의 너비(Data width)가 하둡 스토리지 어플라이언스가 사용하는 버스의 너비와 동일한 데이터 블록을 수신하는 제1 단계;

제어 모듈이 하둡 클러스터에 포함된 하둡 스토리지 어플라이언스 정보를 기초로, 상기 데이터 블록의 압축을 위한 LZ 4 압축 알고리즘을 포함한 복수의 압축 알고리즘을 미리 저장하고 상기 미리 저장된 압축 알고리즘 중 어느 하나를 결정하는 제2 단계;

오프셋의 사전 값을 저장하는 메모리와 메모리 주소를 제어하는 해쉬 함수를 포함하는 사전 모듈이 윈도우의 사이즈만큼 병렬로 사전 등록 및 검색을 동시에 수행하고, 윈도우를 통해 상기 입력 버퍼로부터 수신한 상기 데이터 블록을 상기 제어 모듈에 의해 결정된 압축 알고리즘에 따라 압축을 수행하는 제3 단계;

상기 제어 모듈이 상기 사전 모듈을 통해 수행되는 압축을 위한 연산량을 측정하는 제4 단계;

상기 제어 모듈이 측정되는 연산량이 기 설정된 연산량을 만족하는지 판단하는 제5 단계; 및

출력 버퍼가 상기 압축이 수행된 결과를 FIFO(First In First Out)에 따라 하둡 스토리지 어플라이언스의 버스 너비와 엔디안에 맞게 데이터를 정렬하여 출력하는 제6 단계;를 포함하고,

상기 제5 단계에서 상기 측정되는 연산량이 상기 기 설정된 연산량을 만족하지 않는 경우 상기 제어 모듈은 상기 하둡 스토리지 어플라이언스 정보 및 상기 측정된 연산량을 기초로 압축 알고리즘을 변경하고 상기 데이터 블록의 크기 및 상기 윈도우의 크기 중 적어도 하나를 변경하고, 변경된 압축 알고리즘과 변경된 크기로 상기 사전 모듈이 상기 데이터 블록을 압축하여 상기 제어 모듈이 상기 제5 단계를 다시 수행하고,

상기 제6 단계는 상기 제5 단계에서 상기 측정되는 연산량이 상기 기 설정된 연산량을 만족하는 경우, 상기 제3 단계에서의 상기 사전 모듈이 상기 데이터 블록에 대한 압축의 수행을 완료 한 다음 수행되고,

상기 하둡 스토리지 어플라이언스 정보는 하둡 클러스터에 포함된 하둡 스토리지 어플라이언스의 개수, 상기 하둡 스토리지 어플라이언스 각각의 가용 저장 공간 및 상기 하둡 스토리지 어플라이언스의 저장 공간을 합산한 전체 저장 공간 정보를 포함하고

상기 제어 모듈은 상기 입력 버퍼가 상기 맵리듀스 태스크로부터 전달받는 압축 또는 압축해제가 되어야 하는 데이터 블록의 크기를 결정하고,

상기 압축 알고리즘 변경에 따라 압축을 위한 연산량을 조절하는

하둡 기반의 하드웨어 압축 가속화 방법.

청구항 2

삭제

청구항 3

제1 항에 있어서,

상기 사전 모듈은,

상기 LZ 4 압축 알고리즘에 따라, 상기 데이터 블록에 대한 압축을 수행하는,

하둡 기반의 하드웨어 압축 가속화 방법.

청구항 4

제1 항에 있어서,
 상기 입력 버퍼가 수신하는 데이터 블록의 크기는 256 KB 인,
 하둡 기반의 하드웨어 압축 고속화 방법.

청구항 5

삭제

청구항 6

삭제

청구항 7

삭제

청구항 8

삭제

청구항 9

삭제

발명의 설명

기술 분야

[0001] 본 발명은 저전력 하둡 스토리지 어플라이언스의 고속 처리를 위한 하드웨어 압축 고속화 장치 및 방법에 관한 것이다.

배경 기술

[0002] 최근 빅데이터(Big Data)를 효과적으로 분산 처리하는 방법으로 하둡(Hadoop) 클러스터가 이용되고 있다.

[0003] 하둡 클러스터를 이용하더라도, 처리해야 할 데이터의 양이 증가함에 따라 데이터의 저장 및 분석을 위하여 더 많은 수의 서버가 요구된다. 이러한 서버의 증설은 클러스터 운용에 있어 많은 전력 소모를 초래하고, 클러스터의 관리에 있어 고비용이 발생한다.

[0004] 따라서, 저전력 하둡 스토리지 어플라이언스의 필요성이 대두되고 있다. 이러한 저전력 하둡 스토리지 어플라이언스를 이용하기 위해서는 저전력 CPU의 사용이 요구된다. 그러나, 이와 같은 저전력 CPU의 부족한 연산 능력을 보완하기 위한 장치 및 방법은 제공되고 있지 않다.

발명의 내용

해결하려는 과제

[0005] 본 발명은 상기와 같은 문제점을 해결하기 위해 안출된 것으로, 본 발명의 목적은 어플라이언스에 하둡 기반의 하드웨어 압축 고속화 장치를 제공하기 위한 것이다.

[0006] 보다 구체적으로, 본 발명은 상기 하드웨어 압축 고속화 장치를 통해 하둡 시스템의 데이터 압축을 수행함으로써, 빅데이터를 분산 처리하고 분석하는데 있어 소모되는 시간을 최소화하는 장치 및 방법을 제공하기 위한 것이다.

과제의 해결 수단

[0007] 본 발명의 일 실시예에 따른 하둡 기반 하드웨어 압축 고속화 장치는 압축 또는 압축해제가 되어야 하는 데이터 블록을 수신하는 입력 버퍼와 사전 등록 및 검색을 수행하고, 윈도우를 통해 상기 데이터 블록에 대한 압축을 수행하는 사전 모듈과 상기 압축이 수행된 결과를 출력하는 출력버퍼 및 하둡 스토리지 어플라이언스 정보를 기

초로 상기 입력 버퍼, 상기 사전 모듈 및 상기 출력 버퍼를 제어하는 제어모듈을 포함할 수 있다.

발명의 효과

- [0008] 본 발명에 따른 하둡 기반 하드웨어 압축 가속화 장치 및 방법은 하드웨어 병렬성을 활용하여 압축 알고리즘의 사전 검색 및 등록 과정을 가속화할 수 있다.
- [0009] 보다 구체적으로, 하둡 기반 하드웨어 압축 가속화 장치 및 방법은, 사전 검색 및 등록 과정의 가속화를 통해 기존의 소프트웨어 압축 보다 처리량을 향상 시킨다.
- [0010] 또한, 본 발명에 따른 하둡 기반 하드웨어 압축 가속화 장치 및 방법에 따르면, 저전력 CPU의 부족한 연산 능력을 보완함으로써, 저전력 하둡 스토리지 어플라이언스가 적용된 하둡 클러스터를 저비용으로 관리할 수 있는 장점이 있다.

도면의 간단한 설명

- [0011] 도 1은 본 발명과 관련된 하둡 기반 하드웨어 압축 가속화 장치 구조 예시도.
- 도 2는 본 발명과 관련된 하둡 기반 하드웨어 압축 가속화 장치의 동작 순서도.

발명을 실시하기 위한 구체적인 내용

- [0012] 이하, 첨부된 도면을 참조하여 본 발명의 바람직한 실시예를 상세히 설명한다. 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 명세서 전체에 걸쳐 동일 참조 부호는 동일 구성 요소를 지칭한다.
- [0013] 다른 정의가 없다면, 본 명세서에서 사용되는 모든 용어(기술 및 과학적 용어를 포함)는 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 공통적으로 이해될 수 있는 의미로 사용될 수 있을 것이다. 또 일반적으로 사용되는 사전에 정의되어 있는 용어들은 명백하게 특별히 정의되어 있지 않는 한 이상적으로 또는 과도하게 해석되지 않는다.
- [0014] 본 명세서에서 사용된 용어는 실시예들을 설명하기 위한 것이며 본 발명을 제한하고자 하는 것은 아니다. 본 명세서에서, 단수형은 문구에서 특별히 언급하지 않는 한 복수형도 포함한다. 명세서에서 사용되는 "포함한다(comprises)" 및/또는 "포함하는(comprising)"은 언급된 구성요소 외에 하나 이상의 다른 구성요소의 존재 또는 추가를 배제하지 않는다.
- [0015] 본 명세서에서 어플라이언스란 서버나 스토리지 같은 하드웨어를 의미한다. 상기 어플라이언스는 판매 전에 미리 소프트웨어를 장착해 특정 업무에 최적화된 상태에서 판매되는 정보기기일 수 있다. 사용자는 통합 장비 운영 체계나 응용 소프트웨어의 설치, 설정 등의 별도 프로그램의 설치를 행하지 않고도, 구입 시 전원을 접속하면 상기 어플라이언스를 사용 할 수 있다.
- [0016] 특히, 상기 하둡 스토리지 어플라이언스는 하둡 기반의 데이터 분산 저장을 수행하는 어플라이언스를 의미한다.
- [0017] 도 1은 본 발명과 관련된 하둡 기반의 하드웨어 압축 가속화 장치의 예시도이다. 도 1을 참조하면, 하둡 기반의 하드웨어 압축 가속화 장치(100)는 입력 버퍼(10), 사전 모듈(20), 출력 버퍼(30) 및 제어 모듈(40)을 포함할 수 있다.
- [0018] 하드웨어 압축 가속화 장치(100)는 하둡 스토리지 어플라이언스와 충분한 대역폭을 보장하는 인터페이스를 통하여 연결되어야 한다. 예를 들어, 상기 인터페이스는 PCIe 2.0 x4 Lane일 수 있다. 이때, 하둡 기반 하드웨어 압축 가속화 장치(100)는 예를 들어, FPGA (Field Programmable Gate Array) 또는 SoC (System On Chip) 상에서 구현될 수 있다.
- [0019] 또한, 상기 하둡 기반 하드웨어 압축 가속화 장치(100)는 압축 알고리즘 연산 회로를 구비할 수도 있다. 이때, 압축 알고리즘은 사전 기반의 무손실 압축 알고리즘이 사용될 수 있다. 예를 들어, 압축 알고리즘은 LZ4 압축 알고리즘일 수 있다. 또한, 상기 LZ4 압축 알고리즘 연산 회로가 처리하는 블록의 크기는 256 KB인 것일 수 있다.

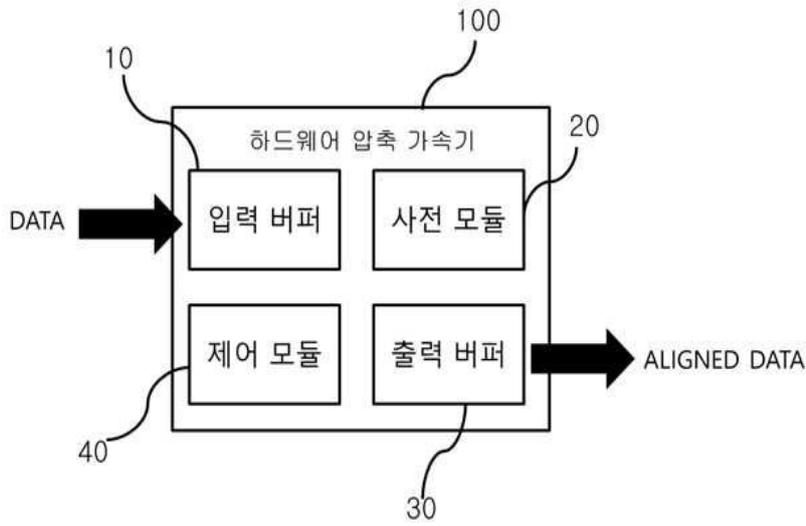
- [0020] 이와 같이 하둡 기반 하드웨어 압축 고속화 장치(100)가 압축 알고리즘 연산 회로를 구비하는 경우, 상기 입력 버퍼(10), 사전 모듈(20), 출력 버퍼(30) 및 제어 모듈(40)은 상기 압축 알고리즘 연산 회로를 구성하는 구성요소일 수 있다.
- [0021] 입력 버퍼(10)는 맵리듀스 (MapReduce) 태스크로부터 압축 또는 압축해제가 되어야 하는 데이터 블록을 전달 받는다. 데이터 너비 (Data width)는 하둡 스토리지 어플라이언스가 사용하는 버스의 너비와 같으며, 입력 버퍼(10)는 하드웨어 압축 고속화 장치에서 데이터를 처리하기 위하여 사전 모듈의 윈도우와 같은 너비의 엔디안으로 데이터를 정렬한다. 예를 들어, 하둡 스토리지 어플라이언스로부터 입력 버퍼로 32-bit 데이터가 들어오면 128-bit 빅엔디안으로 데이터를 정렬할 수 있다.
- [0022] 상기 엔디안은 하드웨어 압축 고속화 장치(100)의 사용 환경에 따라 정해질 수 있다. 즉, 하둡 클러스터에 적용된 CPU의 종류에 따라 결정될 수 있다. 또는, 상기 엔디안은 사용자 또는 제조자의 설정에 따라 결정될 수 있다.
- [0023] 입력 버퍼(10)는 정렬된 데이터를 사전 모듈(20)에 제공할 수 있다.
- [0024] 사전 모듈(20)은 오프셋 등의 사전 값을 저장할 메모리와 메모리 주소를 제어하는 해쉬 함수를 포함할 수 있다. 사전 모듈(20)은 해쉬 함수를 이용하여 메모리 및 메모리 주소를 제어하는 적어도 하나의 프로그램을 저장할 수 있다.
- [0025] 또한, 사전 모듈(20)은 사전 등록 및 검색을 수행하는 로직과 윈도우를 포함할 수 있다. 윈도우는 압축이 수행되는 크기와 같으며 병렬처리가 수행되는 단위이다. 이때, 사용되는 메모리의 크기와 수는 사용하는 해쉬 함수에 따라 달라질 수 있다. 또한, 사전 모듈(20)은 사전 등록 및 검색 과정을 윈도우 사이즈만큼 병렬로 동시에 수행할 수 있다. 사전 모듈(20)은 입력 버퍼(10)로부터 제공 받은 상기 정렬된 데이터에 대한 압축 동작을 상기 윈도우를 이용하여 수행할 수 있다.
- [0026] 출력 버퍼(30)는 압축이 수행된 결과를 하둡 스토리지 어플라이언스의 버스 너비와 엔디안에 맞게 데이터를 정렬하여 출력한다. 압축된 데이터를 출력하는 과정의 오버헤드로 인하여 유발되는 압축 과정의 지연을 막기 위해 FIFO(First In First Out)가 사용될 수 있다.
- [0027] 제어 모듈(40)은 하둡 기반 하드웨어 압축 고속화 장치(100)의 각 구성의 동작을 제어할 수 있다. 즉, 제어 모듈(30)은 상기 하둡 기반 하드웨어 압축 고속화 장치(100)를 구성하는 각 구성 요소의 신호처리 및 각 구성 요소간 데이터의 송수신을 제어할 수 있다.
- [0028] 상기 압축 과정은 각 단계가 순차적이므로 제어 모듈(40)은 FSM(Finite State Machine)을 사용하여 상태에 따른 순차적으로 제어할 수 있다.
- [0029] 제어 모듈(40)은 데이터 압축을 위한 적어도 하나의 압축 알고리즘을 미리 저장할 수 있다. 제어 모듈(40)은 미리 저장된 압축 알고리즘 중 어느 하나를 결정할 수 있다. 이때, 제어 모듈(40)은 하둡 스토리지 어플라이언스 정보에 기초하여 압축 알고리즘을 결정할 수 있다. 하둡 스토리지 어플라이언스 정보는 하둡 클러스터에 포함된 하둡 스토리지 어플라이언스의 개수, 하둡 스토리지 어플라이언스 각각의 가용 저장 공간 및 모든 하둡 스토리지 어플라이언스의 저장 공간을 합산한 전체 저장 공간 등의 정보를 포함할 수 있다.
- [0030] 또한, 제어 모듈(40)은 하둡 기반 하드웨어 압축 고속화 장치(100)의 데이터 압축 동작 수행에 따른 연산량을 측정할 수 있다. 여기에서 연산량은 연산 속도, 즉 압축 또는 압축 해제 동작 수행 속도에 대한 정보를 포함할 수도 있다. 제어 모듈(40)은 측정된 연산량 및 하둡 스토리지 어플라이언스 정보를 바탕으로 압축 알고리즘을 결정할 수도 있다.
- [0031] 즉, 어느 하나의 압축 알고리즘을 결정하여, 이에 따라 압축 알고리즘을 수행하는 중 측정된 연산량이 기 설정된 연산 속도 및 처리량에 미치지 못하는 경우, 제어 모듈(40)은 압축 알고리즘을 변경할 수 있다. 또한, 하둡 스토리지 어플라이언스 정보에 변경 사항이 있는 경우에, 제어 모듈(40)은 압축 알고리즘을 변경할 수도 있다.
- [0032] 또한, 제어 모듈(40)은 입력 버퍼(10)가 맵리듀스 (MapReduce) 태스크로부터 전달받는 압축 또는 압축해제가 되어야 하는 데이터 블록의 크기를 결정할 수도 있다. 제어 모듈(40)은 데이터 상기 압축 동작 수행에 따른 연산량을 측정하고, 이를 바탕으로 상기 블록의 크기를 결정할 수 있다. 제어 모듈(40)은 상기 블록 크기 결정을 위해 하둡 스토리지 어플라이언스 정보를 이용할 수도 있다.
- [0033] 제어 모듈(40)은 사전 모듈(20)으로부터 출력 버퍼(30)로 제공되는 압축 결과 또는 출력 버퍼(30)가 출력하는

데이터 정렬 정도 중 적어도 하나를 측정함으로써 상기 연산량을 측정할 수 있다.

- [0034] 또한, 제어 모듈(40)은 상기 하둡 스토리지 어플라이언스 정보 및 상기 측정된 연산량을 토대로 상기 윈도우의 사이즈를 결정할 수도 있다.
- [0035] 도 2는 본 발명과 관련된 하둡 기반 하드웨어 압축 가속화 장치의 동작 순서도이다. 도 2를 참조하여 하둡 기반 하드웨어 압축 가속화 장치(100)의 압축 수행 동작을 구체적으로 설명하도록 한다.
- [0036] 도 2를 참조하면, 하둡 기반 하드웨어 압축 가속화 장치(100)는 입력 버퍼(10)를 통해 압축 또는 압축 해제를 위한 데이터 블록을 수신한다(S210). 다음으로, 하둡 기반 하드웨어 압축 가속화 장치(100)는 사전 모듈(20) 상의 윈도우를 통해 데이터 블록에 대한 압축 또는 압축 해제를 수행할 수 있다(S220).
- [0037] 하둡 기반 하드웨어 압축 가속화 장치(100)는 출력 버퍼(30)를 통해, 압축 또는 압축 해제가 수행된 결과를 출력할 수 있다(S230). 이때, 제어 모듈(40)은 수행된 데이터 압축 또는 압축 해제에 대한 연산량을 측정할 수 있다(S240). 제어 모듈(40)은 측정된 연산량 및 하둡 스토리지 어플라이언스 정보를 기초로 상기 데이터 블록의 크기 및 상기 윈도우의 크기 중 적어도 하나를 변경할 수 있다(S250).
- [0038] 다음으로, 하둡 기반 하드웨어 압축 가속화 장치(100)는 상기 데이터 블록의 크기 및 상기 윈도우의 크기 중 적어도 하나가 변경된 값으로 압축 또는 압축 해제를 수행할 수 있다. 이때, 제어 모듈(40)은 상기 적어도 하나가 변경된 값으로 압축 진행 시, 측정되는 연산량이 기 설정된 연산량을 만족하는지 판단할 수 있다(S260). 상기 판단 결과, 기 설정된 연산량을 만족하면, 하둡 기반 하드웨어 압축 가속화 장치(100)는 압축을 계속 수행할 수 있다(S270).
- [0039] 반면, 측정되는 연산량이 기 설정된 연산량을 만족하지 않는 경우, 단계(S250)에서 다시 제어 모듈(40)은 측정된 연산량 및 하둡 스토리지 어플라이언스 정보를 기초로 상기 데이터 블록의 크기 및 상기 윈도우의 크기 중 적어도 하나를 변경할 수 있다. 특히, 제어 모듈(40)은 기 설정된 연산량이 만족되도록 상기 적어도 하나를 변경할 수 있다. 이때, 기 설정된 연산량은 사용자 또는 제조자의 설정에 따라 결정될 수 있다.
- [0040] 이러한 과정을 통해, 하둡 기반 하드웨어 압축 가속화 장치(100)는 최적의 데이터 압축 또는 압축 해제 상태를 유지할 수 있다.
- [0041] 한편, 도 1의 구성요소들은 기능 또는 동작에 따라 구분된 것에 불과하므로, 다른 기준에 따라 구분될 수도 있다. 또한, 도시된 구성요소들이 필수적 구성요소는 아니므로, 각각 일부 구성요소를 포함하지 않거나 추가적인 구성요소를 더 포함할 수 있다.
- [0042] 한편, 본 명세서와 도면을 통해 본 발명의 바람직한 실시 예들에 대하여 설명하였으며, 비록 특정 용어들이 사용되었으나, 이는 단지 본 발명의 기술 내용을 쉽게 설명하고 발명의 이해를 돕기 위한 일반적인 의미에서 사용된 것일 뿐, 본 발명의 범위를 한정하고자 하는 것은 아니다. 여기에 개시된 실시 예외에도 본 발명의 기술적 사상에 바탕을 둔 다른 변형 예들이 실시 가능하다는 것은 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 자명한 것이다.

도면

도면1



도면2

