

IoT 기반 가축 관리 빅데이터 시스템

이진실* · 곽호** · 김기영** · 김동균** · 이상정*

순천향대학교*, (주)버추얼스톰**

Livestock Management Bigdata System using Internet of Things

Jin-Sil Lee · Ho Gwak · Gi-Young Kim · Dong-Kyun Kim · Sang-Jeong Lee
SoonChunHyang University, Virtualstorm

E-mail : thdwjdo@naver.com · guohao1206@gmail.com · k71077@nate.com ·
kdk70@sch.ac.kr · sjlee@sch.ac.kr

요 약

최근 차세대 IT 핵심 분야로 떠오르는 IoT와 빅데이터를 농업에 적용해 생산성을 높이는 시도가 확대되고 있다. 수많은 센서 노드에서 수집한 데이터를 효율적으로 저장하고 분석하기 위해 전통적인 관계형 데이터베이스 시스템을 이용하는 데에는 한계가 있다. 따라서 본 논문에서는 센서기반 IoT 기술을 이용하여 수집한 가축 활동 데이터를 빅데이터 클러스터를 구축하여 관리하는 가축 관리 시스템을 설계 구현하고 검증한다. 제안된 가축 관리 시스템의 효율성 검증을 위해 전통적인 관계형 데이터베이스 시스템 환경과 빅데이터 시스템 환경 상에서 가축의 활동량의 분석하는 데에 소요되는 시간을 측정 비교하여 제안된 빅데이터 시스템 환경이 더 효율적임을 입증한다.

키워드

빅데이터, 가축 관리 시스템, 모니터링, 활동량 분석

1. 서 론

최근에 가축을 사육하는 규모의 확대에 따라 가축 관리의 난이도가 높아지고 질병을 예방하기 힘든 여러 가지 문제에 직면하여 가축을 수시로 관찰 또는 모니터링 하는 모델이 필요하다. 최근에 모니터링을 하기 위한 기술적인 가축 관리 시스템에 대한 요구가 급증하면서 3축 가속도 센서를 많이

이용하고 있다 [1].

가축의 발정 감지는 농가의 생산성 및 소득과 직결되는 부분이기 때문에 번식우를 사육하는 생산 농가에게 가장 중요하다 [2]. 가축의 발정을 감지하기 위해 3축 가속도 센서로 초당 7번의 활동량 데이터를 추출한다. 한 농장에는 여러 마리의 번식우가 존재하고 이를 국가 단위로 확대하면 다수의 농장에 수십 마리 이상의 번식우가 존

재하게 된다. 이러한 환경에서 각 농장 단위로 시스템을 사용한다면 데이터 양이 기하학적으로 증가함에 따라서 관계형 데이터베이스로 저장하고 분석하는 데에는 한계가 있다.

본 논문에서는 대규모 가축 관리 데이터의 저장과 분석을 위해 빅데이터를 이용한 가축 관리 시스템을 설계 구현 한다. 제안된 시스템은 하둡 빅데이터 클러스터 상에서 HBase로 설계 구현한다. 또한 설계 구현된 시스템의 효율성 검증을 위해 기존의 관계형 데이터베이스 시스템과 제안된 시스템의 성능을 비교하여 제안된 빅데이터 시스템 환경이 더 효율적임을 입증한다.

II. 빅데이터와 하둡 프레임워크

빅데이터란 기존 데이터베이스 관리도구로 데이터를 수집, 저장, 관리, 분석할 수 있는 역량을 넘어서는 대량의 정형 또는 비정형 데이터 집합 및 이러한 데이터로부터 가치를 추출하고 결과를 분석하는 기술을 의미한다. 데이터를 넓은 영역에서 활용하려는 시도들로 막대한 양의 데이터가 축적되었고, 이를 자산으로 활용해 가치를 창출하려는 시도들이 현실화되고 있다. 현재의 빅데이터 기술은 공공 데이터, 민간 SW 컨설팅 등 다양한 분야에 적용되어 그 가능성을 인정받고 있다 [3].

하둡(Hadoop)은 대용량 데이터를 처리하는 분산 응용 프로그램을 작성하고 실행시키기 위한 오픈 소스 프레임워크이다. 하둡 플랫폼에서는 서비스 목적에 의해 HBase, Pig, Zookeeper 등의 하둡 하부 프레임워크를 선택하여 구성한다 [4]. HBase는 대용량 파일의 처리량을 높이기 위해 설계된 하둡의 분산 파일 시스템인 HDFS 기반의 컬럼 기반 데이터베이스이다.

III. 가축 관리 빅데이터 시스템

3.1 관계형 데이터베이스 vs HBase

표 1. 관계형 데이터베이스와 HBase 비교

	관계형 데이터베이스	HBase
Date layout	로우 또는 컬럼 지향형	컬럼 지향형
Query language	SQL	API 명령어
Index	임의의 컬럼	Rowkey

표 1은 관계형 데이터베이스와 HBase를 비교해 정리해 놓은 것이다. 컬럼 지향 데이터베이스는 데이터를 컬럼 단위로 묶어 저장한다. 특정 쿼리에 대해서는 로우의 모든 데이터가 필요하지 않은 점을 이용하여 분석적 데이터베이스에서는 컬럼 지향의 저장스키마를 주로 이용한다. 컬럼식 데이터베이스는 데이터에 대한 실시간 분석적 접근 기능을 제공하는 데 강점을 보이지만, HBase는 특정 데이터 셀에 키 기반으로 접근하거나 셀에 순차적으로 접근하는 데 탁월한 성능을 보인다.

3.2 빅데이터 클러스터 구성

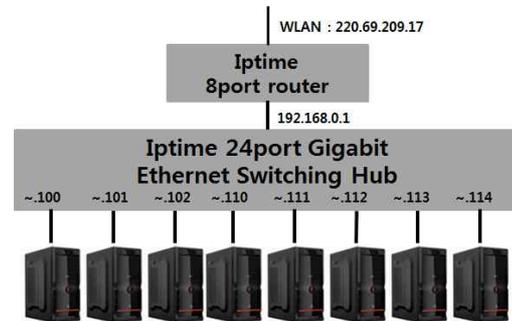


그림 1. 실제 하둡 클러스터 구성도

HDFS는 하나의 서버로 동작하는 것이 아니라 여러 서버에 설치되어 운영된다. 그림 1과 같이 총 8대의 서버를 하나의 고속 이더넷 허브에 연결하고 최상단에 라우터를 위치시켜 실제 하둡 클러스터를 구축하였다.

hostname	IP	hadoop	Hbase
Master	192.168.0.100	Namenode, Datanode	Hmaster, HRegionServer
Slave0	192.168.0.101	Secondary Namenode, Datanode	HRegionServer
Slave1	192.168.0.102	Datanode	HRegionServer
Slave2	192.168.0.110	Datanode	HRegionServer
Slave3	192.168.0.111	Datanode	HRegionServer
Slave4	192.168.0.112	Datanode	HRegionServer
Slave5	192.168.0.113	Datanode	HRegionServer
Slave6	192.168.0.114	Datanode	HRegionServer

그림 2. 각 노드들의 고정 IP와 역할

그림 2는 각 노드의 고정적 IP와 소프트웨어에서 하는 역할을 정리해 놓은 것이다. 네임노드는 데이터 노드의 동작상태와 분산 저장되어 있는 데이터에 대한 메타데이터를 관리하는 일을 한다. 각 노드들은 관리, 처리, 저장을 위하여 큰 디스크 공간이 필요하다. 다음 그림 3은 하드웨어 스펙을 정리한 내용이다.

hostname	CPU	RAM	HDD
Master	Intel Core i3-4150 Processor	8GB	2TB
Slave0	Intel Core i3-4150 Processor	8GB	2TB
Slave1	Intel Core i3-4150 Processor	8GB	2TB
Slave2	Intel Core i3-4150 Processor	8GB	2TB
Slave3	Intel Core i3-4150 Processor	8GB	2TB
Slave4	Intel Core i3-4150 Processor	8GB	2TB
Slave5	Intel Core i3-4150 Processor	8GB	2TB
Slave6	Intel Core i3-4150 Processor	8GB	2TB

그림 3. 하드웨어 스펙

3.3 HBase 테이블 설계

HBase에서 데이터는 테이블에 저장된다. 테이블은 로우와 컬럼으로 구성되고, 로우는 테이블의 주 키인 rowkey에 의해 정렬

되며 모든 테이블의 접근은 테이블의 rowkey를 통해 이루어진다.

국가코드							
Rowkey	DATA						
	시퀀스	x	y	z	지자기	위도	경도

그림 4. HBase 테이블 구조

본 연구에서는 센서를 이용하여 수집된 데이터들을 그림 4와 같은 구조로 HBase 테이블에 저장한다. 데이터에 접근하는 시간을 줄이기 위해 지역코드, 농장코드, mote_id, datetime, micro_second 까지 여러 데이터를 통합하여 rowkey로 지정하고 나머지의 데이터를 컬럼으로 추가하여 효율적으로 저장하였다.

```
public static void addRecord(String tableName, String rowKey, String family, String qualifier, String value) throws IOException{
    try{
        HTable table = new HTable(conf, tableName);
        Put put = new Put(Bytes.toBytes(rowKey));
        put.add(Bytes.toBytes(family), Bytes.toBytes(qualifier), Bytes.toBytes(value));
        table.put(put);
    } catch (IOException e) {
        e.printStackTrace();
    }
}
```

그림 5. 데이터 저장 소스코드

그림 5의 소스코드는 tableName, rowKey 등의 정보를 인수로 받아 데이터를 HBase 테이블에 저장하는 함수이다. HBase에서 제공하는 Put 메소드의 add()를 이용하여 컬럼을 하나씩 추가해서 데이터를 저장한다.

3.4 빅데이터 시스템 모니터링

그림 6는 전체 시스템의 구성도를 그린 것이다. 소의 목에 달린 3축 가속도센서를 이용하여 활동량 데이터를 수집하고 임베디드 보드를 통하여 데이터를 저장을 한다. 또한 외부의 사용자의 직접 접근을 방지하

기 위하여 프록시 서버를 통해 웹 서비스를 제공한다. 외부 사용자는 웹 서비스를 이용하여 모니터링을 할 수 있다.

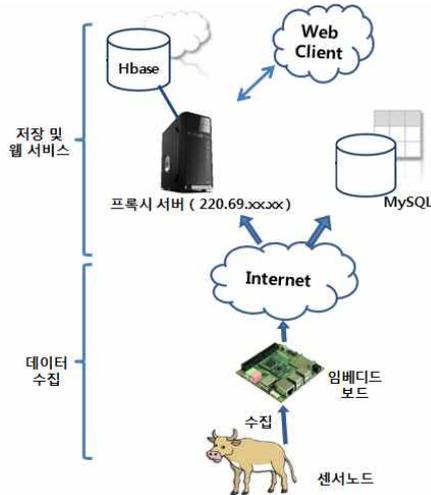


그림 6. 전체 시스템 구성도

데이터의 값과 변화를 쉽게 알아보기 위해 Javascript 기반의 HighChart를 이용하여 모니터링 시스템을 구축하였다. 날짜로 특정 구간을 지정하여 조회가 가능하며 해당 날짜에 포함되는 데이터만을 HBase 테이블에서 가져와 그래프로 표현한다.



그림 7. 그래프로 표현한 데이터

그림 7은 활동량을 분석하기 위해 필요한 x, y, z 값의 데이터를 보기 쉽게 그래프로 나타낸 것이다.

3.5 시스템 환경 성능 평가

본 연구에서는 같은 양의 센서 데이터를 두 시스템에 저장하였다. 성능을 비교하기 위해 데이터 수집 기간과 데이터의 양을 동일하게 하였고, 데이터를 검색하고 활동량을 계산하는 연산을 처리하는 데에 소요되는 시간을 측정하였다.

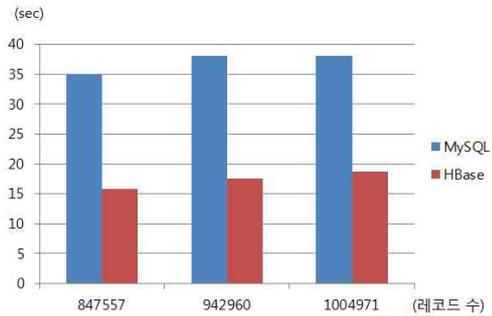


그림 8. 시스템 환경 성능 비교

그림 8이 실제로 소요되는 시간을 측정된 결과이다. x축은 정해진 시간동안의 데이터 양을 가리키고 y축은 소요된 시간을 가리킨다. 양이 늘어남에 따라 소요되는 시간이 증가하고 관계형 데이터베이스인 MySQL 보다 HBase가 소요된 시간이 적은 것을 볼 수 있다. 이 결과로 관계형 데이터베이스 시스템보다 빅데이터 시스템이 더 효율적임을 입증하였다.

IV. 결 론

본 논문에서는 최근 주목받고 있는 빅데

이터 기술을 활용하여 가축의 활동량 데이터를 수집하고 그 결과로 가축을 효율적으로 관리 할 수 있는 시스템을 구축하였다. 가축을 사육하는 규모의 확대에 따라 데이터 양이 많아짐으로써 빅데이터 시스템인 하둡 클러스터를 구성하였다. HBase의 테이블을 이용하여 데이터를 저장하였고, 그 데이터를 이용하여 그래프로 그려 쉽게 모니터링 할 수 있도록 하였다. 또한, MySQL와 HBase의 성능을 비교하는 실험을 통해 빅데이터 시스템이 더 효율적임을 입증하였다.

감사의 글

본 연구는 농림축산식품부 첨단생산기술 개발사업에 의해 이루어진 것임

참고문헌

- [1] 곽호, USN 기반 가축 관리 시스템, 순천향대학교 컴퓨터학과 석사학위논문, 2015.6
- [2] 김현기, ICT 기반의 가축 발정 감지를 위한 시스템 연구, 순천대학교 정보통신공학과 석사학위논문, 2013.12
- [3] 신동희, 이재길, 빅데이터 동향 및 시사점, 인터넷정보학회지 제14권 제2호, 2013.06,
- [4] 한성근, 공간 빅 데이터를 위한 Spatial HBase의 설계 및 구현, 건국대학교 정보통신공학과 석사학위논문, 2013.02
- [5] 조영탁, 이원진, 이인규, 온병원, 최중인, 하둡기반 빅데이터 시스템을 이용한 스마트그리드 전력 데이터 분석, 전기학회논문지 제64P권 제2호, 2015.06
- [6] D. U. Jeong, S. J. Kim, W. Y. Chung, Classification of posture and movement using a 3-axis accelerometer. IEEE International

Conference on Convergence Information Technology, pp. 837-844. November, 2007

- [7] Markandeya, M. N., Mhetre, M. Automatic Estrus Detection System for Dairy Animals. International Journal on Recent Trends in Engineering & Technology, September, 2013