

# 빅데이터 기반 대용량 시계열 에너지 데이터 처리 시스템

강정훈\*, 유준재\*

\*전자부품연구원

e-mail: [budge@keti.re.kr](mailto:budge@keti.re.kr)

## Time-series big data analytics software on IoT streaming data

Jeonghoon Kang\*, June-Jae Yoo\*

\*Korea Electronics Technology Institute

### 요 약

본 논문은 에너지 빅데이터를 분석하기 위해 대용량의 시계열 데이터를 처리하는 시스템의 설계, 구축 방법을 제시한다. 이미 사용 중인 건물이나 공장의 에너지 효율화를 위해서 정부는 효율자원 시장 지원 사업을 수행하고 있다, 에너지 소비 설비에 따라 고효율 자원으로 변경 설치하는 데 필요한 자금의 일부를 지원하고 있다. 정부지원으로 고효율 설비로 변경함에 따라 실증 사이트에서는 측정 데이터를 수집하여, 효율화 정도를 파악하기 위한 에너지 데이터 분석 시스템을 구축하여 운영하였다. 해당 측정 정보는 IoT 전력량계를 통해 수집되며, 수집된 데이터는 클라우드 시스템에서 다양한 머신러닝 알고리즘에 적용되어, 에너지 소비 효율 평가에 필요한 성능 지표를 연산한다. 구현된 진단 시스템은 기축 건물의 에너지 효율향상 상황을 분석하는데 기여할 수 있다. 빅데이터 기반의 에너지 분석 기능을 사용하여 에너지 고효율 장비의 운영시간, 부하율 등의 효율성과 성능통계를 연산할 수 있다

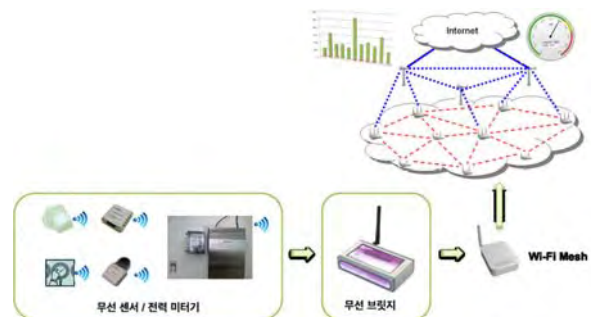
### 1. 서론

본 논문에서는 건물이나 공장에 설치된 IoT 전력 미터기 데이터를 기반으로 에너지 소비패턴을 분석할 수 있는 고속 검색 시스템과 데이터 수집, 처리 시스템을 제안한다. 지금까지는 건물의 에너지 분석을 위해서 Measurement & Verification (M&V) 프로토콜 절차를 수행하여 수동으로 건물 에너지 요소 및 변수를 파악하고, 에너지 효율 성능을 도출 하였다. 건물에 데이터 수집 시스템이 적용된 경우, 12개월 데이터 저장 용량은 기가바이트 이상의 대용량이며, 도시에서 다수 건물의 데이터를 동시에 분석할 경우, 데이터 용량이 테라바이트 규모이다. 이런 빅데이터를 관리자 또는 전문가들이 수동으로 분석, 정리하는 것은 한계가 있다. 이런 어려움을 해결하고, 자동화된 건물 에너지 분석을 위해 에너지 고속 검색 시스템은 신속한 통계 결과를 연산할 수 있다.

제안한 시스템은 무선 메쉬 네트워크 기반의 IoT 센서를 이용하여 데이터를 수집하고 데이터베이스 구축과 기반 시스템을 구축한다. 이를 기반으로 상위 고속 시계열 데이터 접근과 유사 데이터의 분류, 분석을 담당하는 데이터 마이닝 분석 부분으로 구성되어 있다. 또한 이런 빅데이터 규모를 관리자 또는 전문가들이 수동으로 분석, 정리하는 한계를 줄이기 위해 비쿼리방식(Query-Free Approach)의 분석 방법을 적용 하였다.

### 2. 빅데이터 수집 시스템

무선 메쉬기술은 에너지 측정 센서인 전력량계, 온도, 습도, 최대수요 기기, 냉난방 공조기 설비 등의 감지 정보를 수집할 수 있도록 구성되며, 와이파이 메쉬와 TinyOS가 탑재된 무선 센서 네트워크 기술을 기반으로 구성된다. 기존의 유선 측정 시스템에 비해 설치와 유지관리 업무를 최소화할 수 있는 기능이 특징이며, 센서와 서버 구간은 저전력 무선 네트워크와 와이파이 메쉬 네트워크의 두 가지 네트워크 기술이 통합 적용된다.



(그림 1) 무선 메쉬 시스템

수집된 에너지 측정 정보를 저장하는 시스템은 분산 빅데이터 소프트웨어를 적용하여 구현하였다. 시스템의 데이터베이스 기능에 사용되는 HBases는 Hadoop기반의 데이

터베이스로 Hadoop의 HDFS 위에 설치되며, Zookeeper를 Node관리에 사용한다. RowKey를 기반으로 조회에 빠른 액세스 속도를 내지만, 쿼리문은 지원하지 않는다[1]. 대용량 접근 기능과 고속 스캔을 지원하고, 빅데이터는 1,000개 이상의 클러스터 노드를 지원하며 분산 컬럼 기반(distributed column oriented) NoSQL 데이터베이스 기능을 제공한다.

분산형 빅데이터 데이터베이스는 지속적인 데이터의 수집, 추가 확장이 가능하지만, 효율적인 에너지 서비스를 위해서는 고속 에너지 데이터 처리 기능이 필요하다. 접근 속도를 높이기 위해 Time-Series 기능을 제공하는 어댑테이션 레이어 OpenTSDB를 적용하여 구축하였다.

OpenTSDB는 Time Series Daemon(TSD)로 구성되고, 하나 이상의 TSD, Tcollector가 동작하고 각각 독립적으로 데이터를 주고받는다[2][3]. Hbase 스키마는 최소한의 공간에 유사 시계열의 데이터를 빠르게 저장하며 TSD 유저는 직접 HBase에 액세스할 필요가 없고 간단한 텔넷 프로토콜, HTTP API 또는 간단한 GUI를 통해 TSD와 통신할 수 있다.

IoT를 통해 1초 주기로 수신되는 전력미터기의 원본 데이터가 수집된 이후에는, 다음 단계인 데이터 전처리 과정에서는 데이터 요약, 압축에 사용되는 센서 데이터를 데이터베이스로부터 전송 받아, 분석 포맷으로 변환하여 다음 단계 입력 형식으로 준비한다. 준비된 에너지 센서 포인트 데이터들의 시간동기를 맞추고, 수집되지 않은 데이터를 보간하여 데이터 충실도를 높인다[4].

데이터 요약, 압축과정은 입력변환을 거쳐 정규 이벤트(Regular Event Classification)와 비정규 이벤트(Irregular Event Classification)를 탐색한다. 이후, 자동화된 affinity propagation을 사용하여 센서 데이터를 유사도에 따라 클러스터링 한다.

실시간 모니터링을 통해 에너지 피크를 실시간 관제할 수 있으나, 측정값의 비용 및 시스템 의미를 분석하기 위해서는 6개월에서 2년 정도까지의 과거데이터를 동시에 처리하여, 현재 측정되는 데이터의 의미를 분석, 정리, 표출해야 한다. 개발 소프트웨어에는 지난 1년간의 전력 소비량 데이터에서 현재 데이터와 작년 피크 전력의 95%, 90%, 85%를 소비한 시점에 대한 고속 분석을 수행 하였다. 이런 고속 전력 분석 데이터를 기반으로 현재의 에너지 소비에 대한 전반적 판단이 가능하도록 결과를 제공한다. 사용자는 다양한 통계 결과에 따라 현재 에너지 소비가 낭비 중인지, 현재 상태로 인해 전력 지출 비용이 크게 상승할 것인지에 대한 예측 결과를 실시간으로 확인할 수 있다.

이런 기능은 사용자의 에너지 사용 패턴을 분석하고 개선하는데 많은 기여를 할 수 있는데, 절감 목표로 하는 전력 사용량이 지난 1년간 1회 발생하였다면, 이 경우는 투입 시간에 따른 이익을 추정하면 경제성이 있는 것으로 판단할 수 있으며, 수십 ~ 수백 개의 경우가 결과로 나오

면 예상되는 경제성에 대해 세밀히 분석해볼 필요가 있을 을 의미한다.

피크 분석은 15분 사용량을 기준으로 하므로 현재 사용량과 예상 사용량을 주기적으로 모니터링 해야 할 필요가 있다. 아래 그래프는 실시간 전력 사용량을 게이지 차트로 보여주고 피크 이력을 확인할 수 있으며 최대 피크기반 기본요금과 최대피크의 90%인 경우의 기본요금을 비교하였다.

시스템 표출은 전력 사용량의 히스토리 그래프로서 현재까지의 최대 피크를 빨간색 점선으로 기준값을 표시하였다. 어제와 오늘 그리고 10일 동안의 전력 사용량 패턴을 확인할 수 있고 기상정보 중 외기 온도를 통해 피크 전력 예측에 도움이 되도록 하였다.



(그림 2) 피크 검색

### 3. 결론

제안한 시스템은 수동 M&V 방법에 비해, 데이터 분석 시간을 대폭 줄일 수 있기 때문에, 인지하고 있지 못했던 데이터 사이의 연관관계를 검색할 수 있다. IoT 전력량계의 에너지 소비 형태를 분석하고 절감 방법을 찾아낼 수 있는 핵심 정보를 도출할 수 있다. 빅데이터 소프트웨어를 적용하여 절감결과 분석을 제공하여 이후 장기적으로 에너지 운영효율을 극대화할 수 있을 것으로 기대된다.

### ACKNOWLEDGMENT

This work was supported by the Energy Efficiency & Resources of the Korea Institute of Energy Technology Evaluation and Planning (KETEP) grant funded by the Korea government Ministry of Trade, Industry and Energy. (No.20152010103110)

### 참고문헌

[1] openTSDB <http://opentsdb.net/>  
 [2] HBase <http://hbase.apache.org/>  
 [3] Hadoop <http://hadoop.apache.org/>  
 [4] Deokwoo Jung, Varun Badrinath Krishna, William G. Temple, David K. Y. Yau, "Data-Driven Evaluation of Building Demand Response Capacity" in IEEE SmartGridComm 2014.