

P-HILS 실험데이터 특성을 고려한 데이터 관리 클러스터 구조 설계

허성민, 이지현, 임정택, 함경선, 김대형

한국전자기술연구원

heosm0402@keti.re.kr, jhlee00@keti.re.kr, jtlim@keti.re.kr, ksham@keti.re.kr, thkim@keti.re.kr

Design of Data Management Cluster Structures Considering P-HILS Data Characteristics

Seongmin Heo, Jihyeon Lee, Jeongtaek Lim, Kyung Sun Ham, Taehyoung Kim

Korea Electronics Technology Institute

요약

전 세계적으로 기후 위기에 대응하기 위해 많은 국가가 재생에너지와 관련된 연구를 진행하고 있다. 재생에너지와 관련된 연구는 크게 설비의 스펙 등의 하드웨어와 관련된 연구와 제어 로직 등의 소프트웨어와 관련된 연구로 구분된다[1]. 이 두 연구의 교차 영역에서 두 분야 모두를 실험하는 연구도 있는데, 대표적인 사례가 P-HILS 실험이다. P-HILS 실험은 실험 대상 하드웨어에 전력을 연결하고, 개발된 제어 로직 소프트웨어를 적용하는 방식으로 이루어진다. 개방형 통합 플랫폼에서도 재생에너지 관련 연구 검증을 위해 P-HILS 관련 서비스를 제공하며 실험을 구성하는 하드웨어와 제어 로직으로부터 다양한 형태의 데이터가 생산된다. 실험의 초기 조건과 같은 정적 데이터부터 주파수 변동과 같이 50ms 단위로 측정되는 시계열 데이터가 대표적인 사례이다. 형태와 생성 조건이 상이한 이러한 데이터를 연구에 활용하기 위해서는 각 데이터의 개별 특성을 고려한 데이터 관리 클러스터가 필요하다. 초당 100건 이상 데이터를 생성하는 시계열데이터는 시계열DB에 적재하고 응답속도를 최대한 끌어올려 데이터 요청에 빠른 응답을 가능하게 했다. 비교적 정적인 데이터는 관계형DB에 적재해 JOIN 등의 기능으로 다양한 분석을 할 수 있도록 설계했다. 더불어 CacheDB를 구축하여 요청될 가능성이 높은 데이터를 미리 준비하고 응답 속도를 극대화하는 구조를 설계했다. 본 논문에서는 P-HILS 사용자의 원활한 연구 진행을 위해 시계열 DB, 관계형DB, CacheDB를 활용한 데이터 적재, 관리 클러스터 구조를 제안한다.

I. 서론

최근 기후 위기에 대응하기 위해 대부분의 국가에서 분산에너지를 정책 목표 중 하나로 삼으며 탄소중립 달성을 목표로 하고 있다. 분산에너지 자원(Distributed Energy Resources, DER)이란 중앙에서 전력을 공급하는 전통적인 시스템과 달리 에너지 수요와 인접한 곳에서 생산되고 소비되는 에너지 자원을 의미한다. 많은 에너지 관련 연구에서 분산에너지자원 시스템 개발과 검증에 P-HILS를 활용한다.[2]

는 현상을 포착하고 분석하기 때문에 ms 단위 시간해상도의 데이터가 발생한다. 사용자가 P-HILS 시스템 내에서 발생한 데이터를 분석할 때 짧은 쿼리 응답속도를 보장하고 데이터 가용성을 보장하는 등의 편의를 제공하기 위해서는 데이터 특성에 따른 관리 클러스터 구조가 필요하다. 이에 본 논문에서는 P-HILS 데이터의 개별적 특성을 고려하여 설계한 데이터 관리 클러스터 구조를 제안하고자 한다.

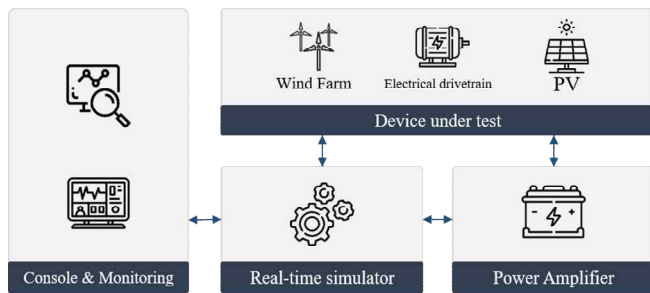


그림 1 P-HILS 구성도

P-HILS는 HILS를 기반으로 한 시뮬레이션 시스템으로 HILS에 전력을 발생하는 파워앰프를 결합한 시스템이다. 결합된 파워앰프가 전력계통, 배터리 등 다양한 형태의 전력을 시뮬레이션하여 시험환경을 구성하고 현실을 반영한 실험이 가능해진다. 이러한 장점으로 NREL(National Renewable Energy Laboratory)과 SNL(Sandia National Laboratories), AIT SmartEST Laboratory[3] 같은 기관들도 P-HILS를 통해 연구를 검증한다. P-HILS를 활용하는 대부분의 연구에서 찰나의 시간 동안 변화하

II. 본론

시뮬레이션 과정은 연구의 목적에 따라 연결되는 PV, 풍력발전단지, 마이크로그리드 등의 DUT(Device Under Test)에 따라서 달라지며 발생하는 데이터 역시 달라진다. 각각의 DUT는 고유한 특징에 따라 천차만별의 데이터를 생성한다. 예를 들면, 마이크로그리드는 ms 단위의 주파수 데이터를 생성하고, 풍력발전단지의 경우에는 ms 단위의 유효전력출력 데이터와 초단위의 Met mast 풍향 데이터를 동시에 생성할 수 있다. 또한 단지 정보 데이터는 시뮬레이션 동안 변경이 없을 가능성도 존재한다. 이렇듯 발생하는 데이터의 특성은 DUT의 종류와 역할에 따라 다양하다. 데이터의 활용성을 높이기 위해서는 특성을 고려한 적재 시스템이 필요하다. 따라서 본 논문에서는 시계열DB와 관계형DB를 결합한 데이터 관리 클러스터 구조를 제안한다.

시계열DB는 관계형DB에 비해 조회와 집계에서 탁월한 성능을 보인다. 삽입된 100만 개의 데이터에 대해 무작위 추출 및 조회를 1,000회 반복하는 테스트에서 시계열DB는 평균 약 40ms가 소요되었고, 같은 테스트에서 관계형DB는 평균 약 800ms 소요되었다.[4] 삽입된 100만 개 데이터에 대해 함께, 평균, 개수세기와 같은 집계함수를 250회씩 실행한 실험에서도

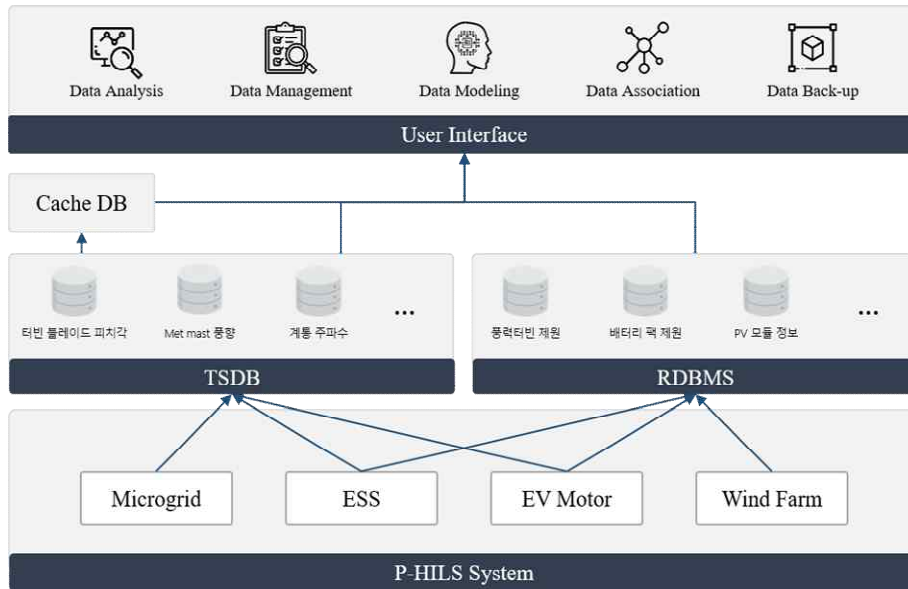


그림 2 P-HILS 실험 데이터 관리를 위한 클러스터 구조

시계열DB는 요청 수행에 평균 약 55ms가 소요되었고, 관계형DB는 평균 약 650ms가 소요되었다.

Operation	TSDB	RDBMS
SELECT	40ms	800ms
SCAN	50ms	750ms
AVG	50ms	550ms
SUM	60ms	550ms
CNT	50ms	550ms

표 1 TSDB, RDBMS 주요 연산 지연속도 비교

반면 관계형DB는 시계열DB에 비해 정규화된 데이터 모델링이 가능하다. 정규화된 데이터 구조를 이용해 장비들의 관계를 나타낼 수 있으며 강력한 JOIN 기능으로 다양한 분석이 가능하다는 점에서 발전단지 정보, 발전설비 스펙 정보와 같은 정적인 정보 관리에 유용하다.

따라서 초당 100건 이상의 고해상도 시계열 데이터는 시계열DB에 적재한다. 응답속도에 초점을 맞추어 저장, 관리하고 사용자의 요구에 빠르게 데이터를 제공하는 데에 목적을 둔다. 비교적 정적인 발전설비 정보, 단지 운영 데이터는 관계형DB에 적재하여 데이터끼리 연계가 가능하게 한다. 더불어 CacheDB에는 시간 지역성, 공간 지역성을 고려하여 요청 가능성이 높은 데이터와 특이사항이 발생했던 시간대 데이터를 미리 적재해 응답 속도 극대화할 수 있다.

사용자가 개방형 통합 플랫폼의 P-HILS를 활용하여 연구 시뮬레이션을 할 때 발생하는 데이터에 대해서 시계열DB, 관계형DB에 나누어 적재, 관리할 수 있다. 양이 많은 데이터의 경우 시계열DB에, 비교적 정적인 데이터의 경우 관계형DB에 적재하여 사용자가 데이터 분석, 활용하고자 할 때 속도와 연계용이성 부분에서 편의성을 제공할 수 있다.

III. 결론 및 향후연구

기후 변화에 대응하기 위해 다양한 국가에서 연구를 진행 중이며, 연구 수행과 검증의 방법 중 하나로 P-HILS를 도입했다. P-HILS는 전력 시스템을 모델링하고 실제 환경을 반영하는 시뮬레이션이 가능하다. 이로 인해 발생하는 데이터는 ms 단위의 고해상도 데이터부터 매우 정적인 데이터까지 특성이 천차만별이다. 이를 효과적으로 관리하기 위해 데이터 특

성에 따른 클러스터 구조가 필요하다. 본 논문에서는 시계열DB와 관계형DB를 결합한 데이터 관리 클러스터 구조를 제안한다. 시계열DB는 고해상도 시계열 데이터를 적재해 빠른 응답속도를 끌어내고, 관계형DB는 정적인 정보를 적재해 설비나 단지, 장비 간의 관계를 이용한 다양한 분석을 가능하게 한다. 더불어 CacheDB를 구조에 추가하여 응답속도를 극대화하는 설계를 제안한다. 이러한 클러스터 구조로 데이터를 관리하며 연구자들이 원활한 연구 검증을 가능하게 돕는다. 하지만 시계열DB의 입력 처리량보다 더 많은 고해상도 데이터가 발생할 경우 입력 처리가 지연되고, 그만큼 데이터 요청 응답도 느려질 수 있다는 점에서 한계가 존재한다. 향후 연구에서는 고해상도 데이터 적재 과정에서 부하가 발생할 경우를 대비한 실시간 데이터 처리 기술을 포함한 클러스터 구조 설계 연구를 진행하여 이러한 한계를 극복하고자 한다.

ACKNOWLEDGMENT

이 논문은 2023년도 정부(산업통상자원부)의 재원으로 한국에너지기술연구원 지원에 의해 수행된 연구임(RS-2023-00231702, 분산형 재생에너지 시스템 개방형 통합 플랫폼 개발)

참고 문헌

- [1] Bernstein, Andrey. 2021. Final Report for ARPA-E NODES "Real-Time Optimization and Control of Next-Generation Distribution Infrastructure" Project. Golden, CO: National Renewable Energy Laboratory. NREL/TP-5D00-78531. <https://www.nrel.gov/docs/fy21osti/78531.pdf>.
- [2] F. Colas, X. Guillaud, P. Rault, J. Desclous, and S. A. Amamra, "Use of real time simulation for power hardware in the loop simulation of multi-terminal DC grid," OPAL-RT Real-Time User Conference, 2013.
- [3] 오형록, 장해근, 한대수. (2018). Power Hardware-In-the-Loop-Simulation(P-HILS) 개념과 활용사례. 전력전자학회지, 23(6), 51-57.
- [4] A. Bader, "Comparison of time series databases." Universität Stuttgart, 11-May-2018.