



주요 특징

- **뛰어난 성능 및 유연성** – x86 기반 경쟁 솔루션보다 최대 1.7배 우수한 성능을 발휘합니다.¹ 다양한 Linux® 서버 옵션으로 워크로드에 적합한 규모로 구축하십시오.
- **효율 극대화** – IBM® Elastic Storage™ Server(IBM Spectrum Scale™ erasure coding)와 통합하면 스토리지 인프라 요구 사항을 HDFS의 1/3 수준으로 줄일 수 있습니다.²
- **코그니티브 지원** – IBM PowerAI 딥러닝 플랫폼에서 머신 러닝과 딥 러닝을 효과적으로 통합할 수 있습니다.
- **고객의 성공 및 개방적인 혁신에 주력** – 철저한 테스트를 거쳐 업계 최고의 자원과 전문성으로 제공되는 개방형 하드웨어/소프트웨어 기술을 기반으로 하는 완전한 엔터프라이즈급 솔루션입니다.

IBM Power Systems 기반 Hortonworks Data Platform

빅데이터 분석 및 인공지능을 지원하는 최고의 개방형 서버를 위한
안전한 엔터프라이즈급 오픈소스 Apache Hadoop 배포판

IBM Power Systems™ 기반 Hortonworks Data Platform(HDP)은 커넥티드 엔터프라이즈 데이터 플랫폼을 위한 최강의 솔루션입니다. 최상의 개방형 기술 혁신과의 조합으로 업계 최고의 성능과 IT 효율을 제공하여 빅데이터 분석과 인공지능의 발전을 가속화함으로써, 기업은 이전과 다른 방식으로 비즈니스를 위한 데이터 중심의 인사이트를 발굴하고 확장할 수 있습니다.

Hortonworks Data Platform

업계 최고의 안전한 엔터프라이즈급 오픈소스 Apache Hadoop 배포판인 HDP는 저장 데이터(data-at-rest)와 관련된 각종 활용 사례에서 실시간 고객 애플리케이션을 지원하고 빠른 의사결정과 혁신을 위한 강력한 분석 기능을 제공합니다.

HDP는 서비스 중단없이 확장 가능한 빅데이터 스토리지로 HDFS(Hadoop Distributed File System)를 리소스 및 워크로드 관리에 Hadoop의 중앙 집중식 YARN(Yet Another Resource Negotiator)을 사용합니다. YARN은 SQL, 실시간 스트리밍, 배치 프로세스 등 다양한 데이터 처리 엔진이 공유 데이터 세트와의 상호 작용을 동시에 수행할 수 있게 함으로써 불필요하고 비용 부담이 큰 데이터 사일로를 방지하고 전혀 새로운 방식의 분석을 가능하게 합니다.

개방적이고 유연한 데이터 플랫폼인 HDP는 데이터 액세스, 거버넌스, 통합뿐 아니라 보안 및 운영 관리 등 종합적인 기능을 갖추고 있습니다. HDP의 오픈소스 커뮤니티 개발 모델을 통해 전사적 범위에서 빠른 혁신과 긴밀한 통합의 이점을 누릴 수 있습니다.



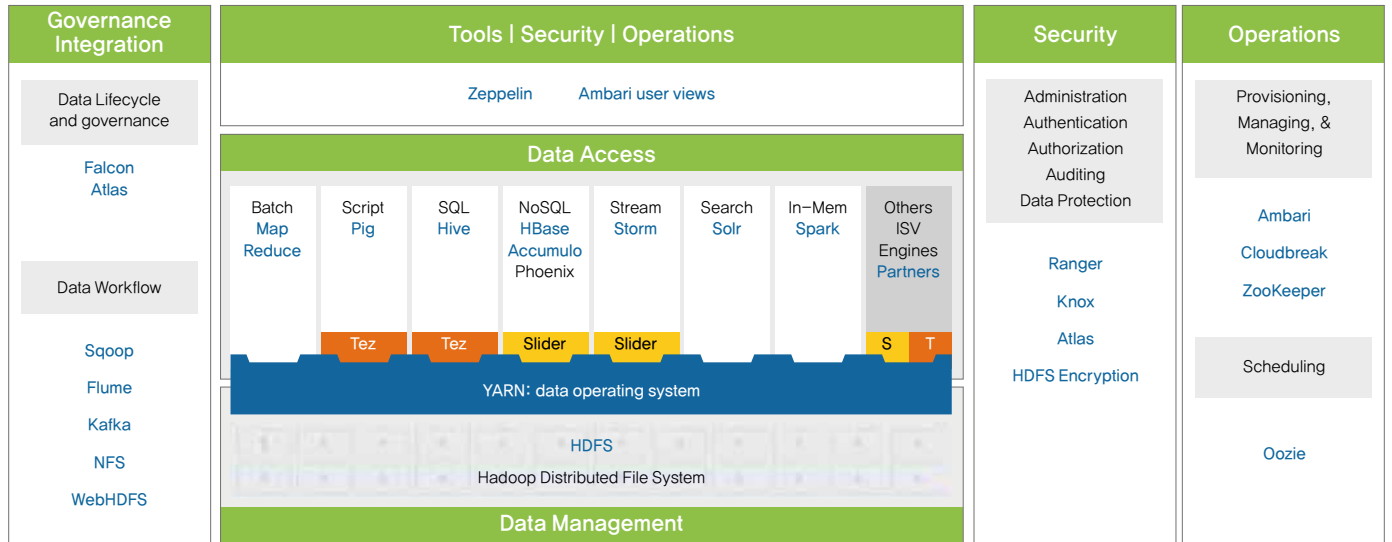


그림 1: Hortonworks Data Platform

Power Systems 기반 Hortonworks Data Platform – 효율 극대화

IBM POWER8® 프로세서와 차별화된 하드웨어 가속 기술을 사용하는 Power Systems는 빅데이터를 분석하는 데 획기적인 성능을 제공합니다. POWER8 프로세서는 빠른 분석을 위한 멀티스레딩(코어당 스레드 8개), 지속적인 데이터 로드와 빠른 응답을 위한 멀티레벨 캐시(L4 캐시 포함), 데이터 중심 애플리케이션의 처리량을 극대화하는 대용량 고대역폭 메모리 워크페이스를 탑재하여 HDP에서 실행 중인 빅데이터 분석 애플리케이션에 업계 최고 수준의 성능을 제공합니다.

IBM Power Systems OpenPOWER LC 서버 제품군

새로운 IBM OpenPOWER LC 서버 제품군은 유연성은 물론이고, 기존의 클러스터나 클라우드와의 원활한 통합을 염두에 두고 개발되었으며, 데이터 집약형(data-crushing) POWER8 프로세서를 갖추고 있어 고밀도 컴퓨팅부터 대용량 스토리지까지 다양한 목적에 맞는 전문 시스템 구성이 가능합니다. OpenPOWER Foundation과의 제휴를 통해 탄생한 LC 제품군의 혁신적인 구성은 하드웨어 액셀러 레이터가 컴퓨팅, 스토리지, 네트워크 워크로드에 대한 부담을 줄임으로써 데이터 분석의 속도를 크게 높이고 데이터 이동의 효율을 극대화합니다.

OpenPOWER는 고성능 컴퓨팅 환경, 엔터프라이즈 IT 및 초대형 데이터 센터, 세계적인 시스템 디자이너 등이 포함된 에코시스템의 파트너와 일반 사용자가 보유한 최상의 기술과 IBM의 첨단 프로세서를 접목시킵니다.

여기서 설명한 IBM Power Systems 기반 HDP의 구성은 뛰어난 성능과 넉넉한 스토리지를 갖추고 빅데이터를 지원하는 Systems S822LC와 가볍지만 강력한 컴퓨팅 성능을 갖춘 Power Systems S821LC의 조합을 토대로 고객의 구체적인 필요에 부합하는 적정 규모의 솔루션을 구성할 수 있는 옵션을 제시합니다. 자세한 내용은 그림 2를 참조하십시오.

Apache Hadoop 및 Spark 워크로드를 위한 탁월한 성능

Power Systems 기반 HDP는 많은 데이터를 더 빨리 전달하기 때문에 보다 유익한 분석으로 합리적인 의사결정을 신속하게 내릴 수 있습니다. 일반 Apache Hadoop 워크로드를 대상으로 한 IBM 성능 테스트에서 Power Systems 기반 HDP는 x86 기반 솔루션과 비교하여 시간당 70% 더 많은 쿼리를 처리하고(평균 응답 시간 기준), 쿼리 응답 시간을 평균 40% 단축했습니다.¹

	System Management Node	Worker Node				
		Master Node	Edge Node	Balanced	Performance	Storage Dense
Server Type	1U S821LC	1U S821LC	1U S821LC	2U S822LC	2U S822LC	2U S822LC
Count(Min/Max)	1 / 1	3 / Any	1 / Any	8 / Any	8 / Any	8 / Any
Core	8	20	20	22	22	11
Memory	32GB	256GB	256GB	256GB	512GB	128GB
Storage—HDD	2x 4TB HDD	4x 4TB HDD	4x 4TB HDD	12x 4TB HDD	8x 6TB HDD	12x 8TB HDD
Storage—SSD					+ 4x 3.8TB SSD	
Storage Controller	Marvell (internal)	LSI MegaRAID 9361-8i (2GB Cache)	LSI MegaRAID 9361-8i (2GB Cache)	LSI MegaRAID 9361-8i (2GB Cache)	LSI MegaRAID 9361-8i (2GB Cache)	LSI MegaRAID 9361-8i (2GB Cache)
Network—1GbE	4 ports (internal)	4 ports (internal)	4 ports (internal)	4 ports (internal)	4 ports (internal)	4 ports (internal)
Network—10GbE	2 ports	2 ports	2 ports	2 ports	2 ports	2 ports

그림 2: IBM Power Systems 기반 HDP 구성

POWER8 프로세서의 뛰어난 쓰레드 밀도, 더 커진 캐시 및 메모리 대역폭, 탁월한 I/O 기능은 Hadoop 워크로드를 빠르게 처리할 뿐 아니라 SQL, 스트리밍, 그래프, 머신 러닝 분석 등 인메모리 Apache Spark 워크로드에도 탁월합니다.

IBM Spectrum Scale과 Elastic Storage Server

IBM Power Systems에서 IBM Spectrum Scale 소프트웨어를 실행하는 통합 스토리지 시스템인 IBM Elastic Storage Server는 HDP를 위한 기본 스토리지의 역할을 할 수 있습니다. IBM Spectrum Scale은 병렬 파일 시스템 아키텍처 기반의 소프트웨어 정의 스토리지 시스템이며, 파일(NFS, SMB, POSIX) 및 오브젝트(S3, Swift) 액세스를 제공하고 HDFS API를 지원합니다. HDFS API를 지원하므로써 엔터프라이즈 스토리지의 데이터를 분석 사일로로 복사할 필요 없이 엔터프라이즈 스토리지에서 곧바로 분석을 수행할 수 있습니다. 이러한 인플레이스(in-place) 분석은 데이터를 복제할

필요도 없고, 오래된 데이터를 대상으로 분석을 실행하게 되는 문제도 없습니다. 뿐만 아니라 Spectrum Scale은 HDP에 대한 공유 스토리지를 제공하므로 컴퓨팅과 스토리지를 분리하여 구성을 최적화할 수 있습니다.

Elastic Storage Server는 광범위한 데이터 유형과 애플리케이션을 저장하고 표준 액세스 방식을 지원하면서 필요에 따라 용량과 성능을 확장할 수 있는 탄력적인 공유 데이터 오션을 구현하는 통합 스토리지 솔루션입니다. 공유 데이터 레이크(방대한 원시 데이터를 원래의 형식으로, 필요한 시점까지 보관하는 스토리지 저장소)를 사용하면 서로 다른 애플리케이션 도메인과 위치에서 동일한 데이터를 공유할 수 있어서 데이터 이동과 복사의 필요성이 줄어듭니다. 따라서 스토리지, 저장 공간, 관리에 드는 비용도 크게 절약됩니다.

PowerAI로 코그너티브 지원

IBM Power Systems 기반 HDP 고객은 완벽하게 최적화된 상태로 지원되는 IBM PowerAI 플랫폼과의 원활한 통합으로 딥 러닝을 수행할 수 있습니다. NVIDIA NVLink Technology 및 Tesla P100 GPU를 갖춘 HPC를 위한 Power Systems S822LC에서 탁월한 성능을 발휘하도록 최적화된 PowerAI는 가장 널리 사용되는 딥 러닝 프레임워크를 포함하여 사전 컴파일과 쉬운 구축으로 최고의 처리량과 케넥티드 데이터에 대한 확장 가능한 딥 러닝을 지원합니다.

고객의 성공을 최우선으로

IBM Power Systems 및 Elastic Storage Server 기반 HDP는 뛰어난 산업 전문성을 바탕으로 고객의 성공을 위해 최선을 다하는 Hortonworks와 IBM에서 철저한 테스트와 검증을 거쳤습니다. IBM과 Hortonworks는 오픈소스 Apache Hadoop 및 Spark 커뮤니티의 대표적인 후원 기업으로서 빅데이터 에코시스템의 혁신과 발전을 이끌고 있습니다. 또한 IBM과 Hortonworks는 개방형 데이터 플랫폼 이니셔티브인 ODPi의 창립 멤버이기도 합니다. 이 비영리 기관의 목적은 공통 참조 사양 및 테스트 솔루션으로 빅데이터 에코시스템을 간소화하고 표준화하는 것입니다.

IBM은 Hortonworks와 손잡고 HDP 및 Power Systems 기반 데이터 레이크를 지원하는 다양한 보안 솔루션을 내놓고 있습니다. 예를 들어 [Attunity Replicate](#)는 다양한 엔터프라이즈 데이터 소스의 변경 데이터를 효과적으로 수집하여 데이터 레이크를 채우고, [Zettaset BDEncrypt](#)는 검증된 엔터프라이즈급 데이터 프라이버시 및 보호 기능을 추가하며, [Synerscope](#)는 첨단 고속 데이터 시각화를 지원합니다.

추가 정보

자세한 내용은 IBM 담당자에게 문의하거나

ibm.biz/hortonworksOnPower를 참조하십시오.



© Copyright IBM Corporation 2017.

한국아이비엠주식회사

(07326) 서울시 영등포구 국제금융로10
서울국제금융센터 (Three IFC)

TEL : (02)3781-7800

www.ibm.com/kr

2017년 08월

Printed in Korea

All Rights Reserved

IBM, IBM 로고, ibm.com, IBM Elastic Storage Server, IBM Spectrum Scale, POWER8 및 Power Systems는 미국 또는 기타 국가에서 사용되는 International Business Machines Corporation의 상표 또는 등록상표입니다. 이와 함께 기타 IBM 상표가 기재된 용어가 상표 기호(® 또는 TM)와 함께 이 정보에 처음 표시된 경우, 이와 같은 기호는 이 정보를 발행할 때 미국에서 IBM이 소유한 등록상표 또는 일반 법적 상표입니다. 또한 이러한 상표는 기타 국가에서 등록상표 또는 일반 법적 상표입니다. 현재 IBM 상표 목록은 웹 "저작권 및 상표 정보" (www.ibm.com/legal/copytrade.shtml)에 있습니다.

Linux는 미국 또는 기타 국가에서 사용되는 Linus Torvalds의 등록상표입니다.

기타 회사, 제품 및 서비스 이름은 타사의 상표 또는 서비스표입니다.

이 문서의 정보는 상품성, 특정 목적에의 적합성에 대한 보증 및 타인의 권리 침해에 대한 보증이나 조건을 포함하여(단, 이에 한하지 않음) 명시적이든 묵시적이든 일체의 보증 없이 "현상태대로" 제공됩니다. IBM 제품에 대한 보증은 제품의 준거 계약 조항에 의거하여 제공됩니다.

1. 성능 결과는 예비 IBM 내부 테스트를 기반으로 합니다.
2. HDFS는 데이터 보호 및 가용성을 위해 2개의 복사본을 추가로 생성하지만 IBM ESS는 30%의 오버헤드만 필요하므로 스토리지에서 시간을 1/3까지 단축할 수 있습니다. ESS에서는 NAS 및 Object 프로토콜을 통해 다른 엔터프라이즈 애플리케이션과 동일한 데이터를 공유하므로 복사본을 추가로 생성하지 않아도 됩니다.



Please Recycle