



IBM PureData System for Analytics

Powered by Netezza technology

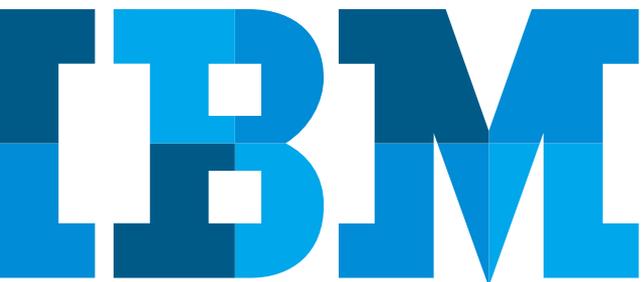
중점 사항

- 관리 및 배치 용이성, 데이터 웨어하우스 및 분석 인프라의 급격한 간소화 가능
- 디자인 단계부터 내장형 전문 지식이 통합되어 있어 설치 후 즉시 실행 가능
- 비즈니스 인텔리전스 및 고급 분석이 통합된 고성능 플랫폼
- 수천 명의 사용자와 복잡한 분석 워크로드 지원
- 동시 분석 및 모듈의 기본 제공으로 분석 개발 간소화, 동시 프로그래밍이 필요하지 않음
- Netezza® 기술 적용

기업은 경쟁 우위를 차지하기 위해 엄청난 양의 이력 데이터에 대한 심층적 고급 분석에 의존해야 합니다. 그럼에도 많은 기업들이 새로운 분석 기능의 출시와 기존 분석에 대한 SLA 유지 문제를 해결해야 하는 과제를 안고 있습니다.

IBM PureData™ System for Analytics는 고객들이 엄청난 양의 데이터를 대상으로 분석을 수행할 수 있게 해주는 고성능 확장형 병렬 시스템입니다. 빅 데이터 용량은 더 간단해지고, 빠르고 액세스가 용이해지고 있습니다. Netezza 기술이 적용된 이 시스템은 엄청난 양의 데이터에 대하여 타사 솔루션보다 더 빠른 속도로 복잡한 분석을 수행하기 위해 특수 설계되었습니다. PureData System for Analytics는 비즈니스 요구를 충족하는 검증된 성능, 확장성, 인텔리전스 및 단순성을 제공합니다. 또한, 관리 또는 튜닝을 최소화하여 TCO(total cost of ownership)를 절약할 수 있다는 비용적 이점도 있습니다.

데이터 분석에 대한 IBM 접근법은 특허 인증 및 검증 받은 방식입니다. 데이터 이동을 최소화하고 매우 빠른 속도로 처리합니다. 이런 업무들을 탁월한 간편성의 데이터 웨어하우스 어플라이언스를 이용하여 한 번에 동시에 처리하여 시간을 단축하고 비용을 절약할 수 있습니다. 과거에는 불가능한데다 비실용적이었던 비즈니스 인텔리전스(BI) 및 고급 분석도 실행합니다.



IBM PureData System for Analytics는 데이터베이스, 서버, 스토리지 및 고급 분석 기능을 관리하기 쉬운 단일 시스템 구조로 통합시킨 목적지향형 표준 기반 데이터 웨어하우스 및 분석 어플라이언스입니다. 페타바이트급 데이터를 빠르고 심층적 분석으로 처리할 수 있게 고안된 이 시스템은 업계에서 가장 낮은 총 소유비용을 실현할 수 있게 해줍니다.

분석 기회

대용량 데이터에 대한 심층적 고급 분석은 똑똑한 경제 환경에서 기업에 없어서는 안될 필수 요소로 기업에 경쟁 우위를 제공합니다. 그럼에도 많은 기업들이 새로운 분석 기능의 출시와 기존 분석에 대한 SLA 유지 문제를 해결해야 하는 과제를 안고 있습니다. 복잡성 관리에 대한 초점을 단순성과 민첩성으로 바꿔야 하는 시대가 도래했습니다.

전략적 분석은 구축 및 관리가 복잡하지 않아야 합니다.

PureData System for Analytics는 고객이 기존 시스템보다 10 – 100배 더 빠른 속도로 엄청난 양의 데이터로부터 통찰력을 얻을 수 있게 해주는 하드웨어 기반의 확장형 대용량 병렬 시스템입니다¹.

속도, 확장성, 지능, 단순성의 완전한 통합

PureData System for Analytics는 엄청난 양의 데이터를 가지고 타사 솔루션보다 더 빠른 속도로 복잡한 분석을 수행하기 위해 특수 설계되었습니다. 이 시스템의 검증된 성능, 확장성, 인텔리전스 및 단순성은 조직이 자사 데이터를 깊이 분석할 수 있게 지원합니다.



Fast

타사 분석 옵션보다 더 뛰어난 PureData System의 성능 이점은 개방형 IBM 블레이드 서버 및 디스크 스토리지와 FPGA(Field Programmable Gate Array)를 이용한 IBM 특허 데이터 필터링을 결합한 AMPP™(Asymmetric Massively Parallel Processing) 아키텍처에서 기인합니다. 이러한 결합은 수만여 명에 달하는 BI 및 데이터 웨어하우스 사용자를 지원하는 분석 워크로드에 대한 쿼리 성능 시간을 단축하고, 빠른 분석 능력과 페타바이트수준의 확장성을 제공합니다.

Increased performance and concurrency

IBM PureData System for Analytics는 새로운 Directed Data Processing 알고리즘을 이용해 기존 Netezza 어플라이언스보다 20배 더 빠른 속도로 전술적 쿼리를 수행할 수 있습니다.² 이 시스템은 관련 데이터를 포함하는 처리 노드에만 워크로드를 자동으로 배포하고 혼합 및 순수 분석 워크로드 공히 효율성을 극대화하고 어플라이언스 단순성을 유지시켜 줍니다.

Scalable

기업은 PureData System을 통해 자사 데이터 용량 및 워크로드에 맞는 환경을 배치할 수 있고, 데이터 용량이 증가함에 따라 얼마든지 대용량 IBM PureData Systems를 빠르고 쉽게 배치할 수 있습니다. PureData System for Analytics는 쿼터 랙 시스템의 경우 기본 데이터 용량(8TB)으로 시작하여 300TB까지 확장할 수 있습니다. 내장 하드웨어 압축을 이용하여 대형 시스템의 경우 사용 가능한 용량은 1.2 페타바이트를 상회합니다.

PureData System for Analytics는 어플라이언스 장비에 따라 선형에 가까운 성능 확장성을 제공합니다. 이것은 조직이 적절한 크기의 어플라이언스를 선택하여 SLA가 요구하는 데이터 용량과 성능을 모두 충족할 수 있고, 예측 및 확장 가능한 성능을 통해 막대한 리소스를 추가하지 않고도 데이터 볼륨 증가에 따라 어플라이언스를 관리할 수 있음을 의미합니다.

Smart

PureData System for Analytics는 상세 데이터가 저장된 영역으로 모든 분석 활동을 통합하여 줌으로써 분석활동의 복잡성을 급격하게 간소화시켜 줍니다. IBM PureData System으로 분석을 이동하는 것은 IBM의 임베디드 분석 플랫폼을 통해 간편해졌습니다. PMML 4.0 모델 지원을 통해 데이터 모델러 및 정량 팀은 개별 인프라로 데이터를 오프로드하여 관련 데이터를 처리, 변환 및 이동시키지 않고도 어플라이언스에 저장된 데이터를 조작할 수 있습니다. 데이터 연구원들은 모든 기업 데이터를 이용하여 자체 모델을 구축한 다음, 다른 모델을 통해 더 빨리 반복하여 최적의 모델을 찾을 수 있습니다. 개발된 분석 모델을 어플라이언스에 있는 관련 데이터에 대해 원활하게 적용할 수 있습니다. 예측 및 스코어링은 데이터가 상주하는 바로 그 위치에서 이루어지고 기타 처리 등도 필요에 따라 함께 처리할 수 있습니다. 사용자들은 자사의 예측 스코어를 실시간으로 얻을 수 있어 고급 분석을 조작할 수 있게 지원하고 기업 전체에 걸쳐 사용할 수 있게 배포할 수 있습니다.

IBM Netezza Analytics의 내장 분석 인프라 및 이와 관련된 통계/수학적 함수 라이브러리는 다양한 분석 툴 및 프로그래밍 언어를 지원합니다. 함께 제공된 200개 이상의 내장된 확장형 내부 데이터베이스 분석 함수 라이브러리는 동시 분석 실행은 물론 개발자, 사용자 및 DBA의 동시 프로그래밍 작업으로 인한 복잡성을 해소해 줍니다.

Netezza Analytics 기능에 포함된 내부 데이터베이스 지리 정보 분석은 업계 표준 ESRI GIS 형식과 호환됩니다. 이로써 기존 지리 정보 분석 환경에 쉽게 통합할 수 있으며, 기업 전역에 걸쳐 빠른 속도로 데이터를 동시에 분석할 수 있습니다.

Simple and completely integrated

이 모든 새로운 기능들은 모두 IBM PureSystems™ 제품군 오퍼링을 특징 짓는 동일한 단순성과 사용 용이성과 함께 제공됩니다. 검증된 Netezza 기술을 토대로 한 이 단순성과 개발 및 배치 용이성은 IBM PureData System for Analytics를 특징 짓는 독창적인 요소입니다. PureData System은 성능을 자동으로 극대화 시켜 주므로 인덱싱 또는 튜닝이 필요하지 않습니다.

하드웨어, 소프트웨어 및 스토리지의 통합체인 어플라이언스가 자동으로 처리해 주므로 배치 시간을 단축하고 단기간에 BI 및 분석 이니셔티브에 대한 가치 실현을 극대화할 수 있습니다. 이 어플라이언스는 즉각적 데이터 로딩 및 쿼리 실행이 가능하고, ODBC, JDBC 및 OLE DB 인터페이스를 통해 선도적 ETL, BI 및 분석 애플리케이션과 통합합니다.

PureData System for Analytics는 처음부터 고가용성을 염두에 두고 설계한 제품입니다. 모든 구성요소는 내부적으로 이중화되어 있기 때문에 처리 노드(S-Blade)에 문제가 생겨도 데이터 센터에 어플라이언스를 연결하면 고성능 운영 환경에 장애가 올 정도로 심각한 영향을 미치지 않습니다.

IBM은 모든 단계에서 복잡성 문제를 해소하여 중요한 리소스를 목표 달성에 실질적 영향을 미치는 이니셔티브에 재할당 할 수 있습니다.

주요 기능

데이터베이스:

IBM Netezza Platform Software 7 이상

지원 API:

SQL, OLE DB, ODBC 3.5, JDBC 3.0 Type 4

SQL 표준:

SQL-92 호환(SQL-99 확장 포함)

프로그래밍 언어:

Java, Python, R*, Fortran, C/C++, Perl, Lua

Netezza Analytics foundation

In-Database Analytics, R*, Matrix, MapReduce, 지리 정보 기술(ESRI 지원)

고속 로딩/엔로딩:

ETL 및 EAI 톨로 상호 운영 가능(시간당 비율: 최대 5TB)

백업 및 복구:

IBM® Tivoli®, EMC Legato 및 Symantec Netbackup으로 상호 운영 가능(시간당 비율: 4TB 이상)

데이터베이스 이동성:

IBM DB2®, Informix®, Microsoft SQL Server, MySQL, Oracle, Red Brick, Sybase IQ, Teradata

운영 체제:

Red Hat Linux Advanced Server 5

추가 들:

Windows 및 웹기반 DB Admin GUI, CLI 및 고속 로딩/엔로딩(AIX®, HP-UX, Linux, Solaris 및 Windows)

IBM InfoSphere® BigInsights™ 1.4 – Starter Edition – 5인용

IBM InfoSphere Streams 2.0 – 개발자 라이선스 – 2인용

IBM PureData System for Analytics는 여러 업체에 걸쳐 보조 기술 파트너, 리셀러, 시스템 통합업체 및 서비스 공급업체 등 선도적인 비즈니스 파트너사들이 지원하고 있습니다. 전체 목록이 필요하거나 당사 프로그램에 특정 기업 또는 솔루션이 해당하는 경우 가까운 IBM 담당자에게 문의하시기 바랍니다.

The best value

IBM PureData System for Analytics는 투자 대비 비용 효율이 우수한 분석 시스템입니다. 총 소유비용(TCO)을 줄이기 위해서는 진행 중인 관리 또는 튜닝과, 내부 리소스, 구현 비용을 최소화시켜야 합니다. IBM 어플라이언스의 성능 및 확장성은 튜닝, 인덱싱, 수집 등의 과정을 사전에 준비하여 두지 않고도 즉시 쿼리를 수행할 수 있습니다.

기업이 변화하는 시장 상황과 증가하는 분석 요구에 민첩하게 대응하려면 당연히 놀라울 정도의 빠른 성능으로 페타급 크기의 데이터를 분석할 수 있는 설치가 용이하고 간편한 시스템이 필요합니다.

IBM은 빠른 가치 실현을 통해 기업 목표 달성에 막대한 영향을 미치는 주요 BI 및 분석 이니셔티브를 제공합니다. IBM과 협력하는 기업은 더욱 정확한 인텔리전스로 무장하여 시장에 존재할 수 있는 기회 또는 위험을 재빨리 인식하고 정확하게 대응합니다.

IBM PureData System for Analytics 사양	단일 랙 시스템			다중 랙 시스템	
IBM PureData System for Analytics	IBM PureData System for Analytics N1001-002	IBM PureData System for Analytics N1001-005	IBM PureData System for Analytics N1001-010	랙 2개	랙 3개 이상
랙	1	1	1	2	3-10
액티브 S-Blade	4	7	14	28	랙 당 14개
CPU 코어	32	56	112	224	랙 당 112개
FPGA 코어	32	56	112	224	랙 당 112개
사용자 데이터(TB) [†]	32	64	128	256	랙 당 128개
전원/랙(최대 W/랙)	2,820	3,960	7,635	7,400	7,000
냉각/랙 - BTU/시간	9,600	13,500	26,100	25,500	24,000
무게/랙 kg	453.6	589.7	907.2	907.2	907.2
높이/랙 cm	202	202	202	202	202
깊이/랙 cm	101.6	101.6	101.6	101.6	101.6
너비/랙 cm	64.8	64.8	64.8	64.8	64.8
전원	200-240VAC 50/60Hz, 단상 16A	200-240VAC 50/60Hz, 단상 24A	200-240VAC 50/60Hz, 단상 48A 200-240VAC 50Hz, 3상 WYE 24A 200-240VAC 60Hz, 3상 Delta 32A	200-240VAC 50/60Hz, 단상 48A 200-240VAC 50Hz, 3상 WYE 24A 200-240VAC 60Hz, 3상 Delta 32A	200-240VAC 50/60Hz, 단상 48A 200-240VAC 50Hz, 3상 WYE 24A 200-240VAC 60Hz, 3상 Delta 32A
드롭/랙	2	2	2	2	2
안전	US/CSA/EN60950				
배출량	FCC 15부, ICES-003, AUS/NZ C-Tick, VCCI 및 EN55022 Class A; 유럽 조약: EN55024				

How is this possible?

특히 인증된 하드웨어 가속기술

IBM PureData System for Analytics는 데이터를 processing unit로 이동하는 것이 아니라 processing 기능이 데이터 저장소로 이동해야 한다는 IBM의 기본 원칙을 고수합니다. 각각의 PureData System에 포함된 다중 Snippet Blade 또는 S-Blade에서는 SQL 쿼리 코드 세그먼트(또는 "snippet") 및 복합 분석 처리가 수

행 됩니다. S-Blade는 어플라이언스의 대용량 동시 처리 엔진을 구성하는 지능형 처리 노드입니다. 각각의 S-Blade는 독립형 서버로, 강력한 멀티코어 Intel CPU, IBM 고유의 멀티엔진 FPGA 및 기가바이트급 RAM 그리고 전용 스토리지 장치를 가지고 있습니다. 이들 모두 상호 간 균형을 통해 동시에 작동하여 성능을 극대화합니다.

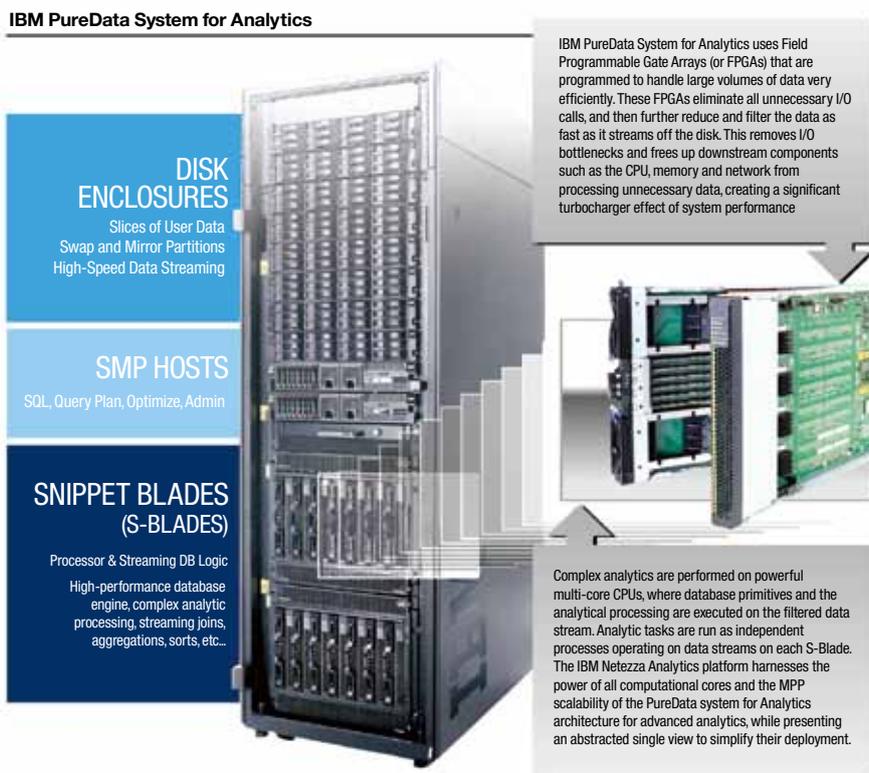


Figure 1: IBM PureData System for Analytics

See for yourself-무료 테스트

IBM PureExperience Program의 일환으로 PureData System for Analytics 무료 평가판을 사용할 수 있습니다. 이 프로그램은 무료로 제공되며 고객님의 데이터로 시스템을 테스트 평가할 수 있게 해줍니다. 이 프로그램은 비즈니스 가치 극대화를 위한 현장 설치 및 데모, 교육/데이터 마이그레이션 서비스, 특정 기간 동안 시스템 사용 및 단일 지원 라인을 제공합니다. 프로그램에 대한 자세한 사항과 가까운 지역에서 제공되는 제품은 다음 사이트를 방문하시기 바랍니다. ibm.com/PureExperience 또는 IBM 담당자에게 문의하시기 바랍니다.

IBM PureData System for Analytics 정보

Netezza 기술이 적용된 IBM PureData System for Analytics는 데이터베이스, 서버 및 스토리지를 쉽게 관리할 수 있도록 단일화 하여 최소한의 설치와 관리 부담만으로도 보다 빠르고 더욱 안정적인 분석성능을 제공 할 수 있게 되었습니다. IBM PureData System for Analytics는 모든 분석 활동을 데이터가 저장된 어플라이언스에 통합하여 비즈니스 분석을 극적으로 간소화하고 업계 최고의 성능을 발휘합니다. 자세한 내용은 ibm.com/PureData 를 통해 IBM 전문가 통합 시스템 제품군으로 어떻게 모든 단계의 복잡성 문제를 한 번에 해소하고 조직을 위해 진정한 비즈니스 가치를 실현할 수 있게 지원하는지 확인해 보시기 바랍니다.

IBM 데이터 웨어하우징 및 분석 솔루션에 관하여

IBM은 가장 폭넓고 포괄적인 데이터 웨어하우징, 정보 관리 및 비즈니스 분석 소프트웨어, 하드웨어 및 솔루션으로 구성된 포트폴리오를 통해 고객이 자사의 정보 자산 가치를 극대화하고, 새로운 통찰력을 발견하여 빠르고 정확한 의사결정을 내리고, 비즈니스 성과를 최적화할 수 있게 도움을 드립니다.

추가 정보

IT가 귀사 전략의 중심으로 이동하여 검증된 전문지식을 통해 업계 선두를 차지할 수 있도록 도움이 되기를 바랍니다. PureSystems 제품군 및 PureData System for Analytics에 대한 자세한 사항은 IBM 담당자 또는 IBM 비즈니스 파트너에게 문의하거나 다음 웹사이트를 방문하시기 바랍니다. ibm.com/PureSystems/PureData/

추가적으로, IBM Global Financing은 가장 비용 효율적 방법과 전략적 방법으로 비즈니스에서 필요로 하는 소프트웨어 기능을 취득할 수 있도록 지원합니다. IBM은 신용 있는 고객과 협력하여 귀사의 비즈니스 및 개발 목표에 적합하고 효과적인 현금 관리를 가능하게 하며 귀사의 총소유 비용을 개선하는 맞춤형 재무 솔루션을 제공합니다. IBM Global Financing으로 중대한 IT 투자에 자본을 투입하고 귀사의 비즈니스를 발전시키십시오. 자세한 정보는 다음 웹사이트를 참조하시기 바랍니다. ibm.com/kr/financing



© Copyright IBM Corporation 2012

(135-270) 서울시 강남구 도곡동 467-12
군인공제회관빌딩

한국아이비엠주식회사
고객만족센터

TEL: (02)3781-7114
www.ibm.com/kr

2012년 11월

Printed in Korea
All Rights Reserved

IBM, IBM 로고, ibm.com, PureData, Tivoli, DB2, Informix, AIX, InfoSphere 및 PureSystems는 전 세계에 등록된 International Business Machines Corp.의 상표입니다. 기타 회사, 제품 및 서비스 이름은 타사의 상표 또는 서비스표입니다. 현재 IBM 상표 목록은 웹 "저작권 및 상표 정보"(ibm.com/legal/copytrade.shtm)에 있습니다.

Netezza는 IBM Company인 IBM International Group B.V.의 등록상표입니다.

Intel은 미국 및 기타 국가에서 Intel Corporation 또는 그 자회사의 등록 상표입니다.

Linux는 미국 또는 기타 국가에서 사용되는 Linus Torvalds의 등록 상표입니다.

Java 및 모든 Java 기반 상표 및 로고는 Oracle 및/또는 해당 자회사의 상표 또는 등록상표입니다.

Microsoft, Windows 및 Windows NT는 미국 또는 기타 국가에서 사용되는 Microsoft Corporation의 상표입니다.

이 문서는 처음 발행될 당시의 날짜를 기준으로 업데이트되었으며 IBM은 언제 든지 문서 내용을 변경할 수 있습니다. 일부 오퍼링은 IBM 매장이 있는 국가에서 제공되지 않습니다.

여기에 언급된 성능 데이터는 특정 운영 환경에서 파생된 결과를 나타낸 것입니다. 실제 결과와 다를 수 있습니다.

IBM 제품 및 프로그램과 함께 사용되는 기타 제품 또는 프로그램을 평가 및 검증하는 것은 사용자의 책임입니다.

이 문서의 정보는 상품성에 대한 보증, 특정 목적의 적합성 여부 및 저작권을 침해하지 않는다는 보증 또는 조건을 포함해 명시적 또는 암묵적 보증 없이 "있는 그대로" 제공됩니다. IBM 제품은 제공된 약정에 명시된 조항 및 조건에 따라 보증됩니다.

실제 사용할 수 있는 스토리지 용량은 비압축 및 압축된 데이터 모두에 대해 보고된 내용일 수 있으며 보고된 수치보다 더 낮을 수 있습니다.

* Revolution Analytics Enterprise R for IBM PureData System for Analytics 사용

† 압축된 사용자 데이터 크기는 4:1 압축으로 가정된 것입니다.



재활용 하십시오.

¹ IBM 고객이 답변한 결과를 토대로 작성한 것입니다. "기존 맞춤형 시스템은 전문적 출고, 테스트 및 최적화 공정"을 거치지 않은 시스템을 말합니다. 실제 결과와 다를 수 있습니다.

² IBM 내부 성능 벤치마킹을 토대로 만들었습니다.