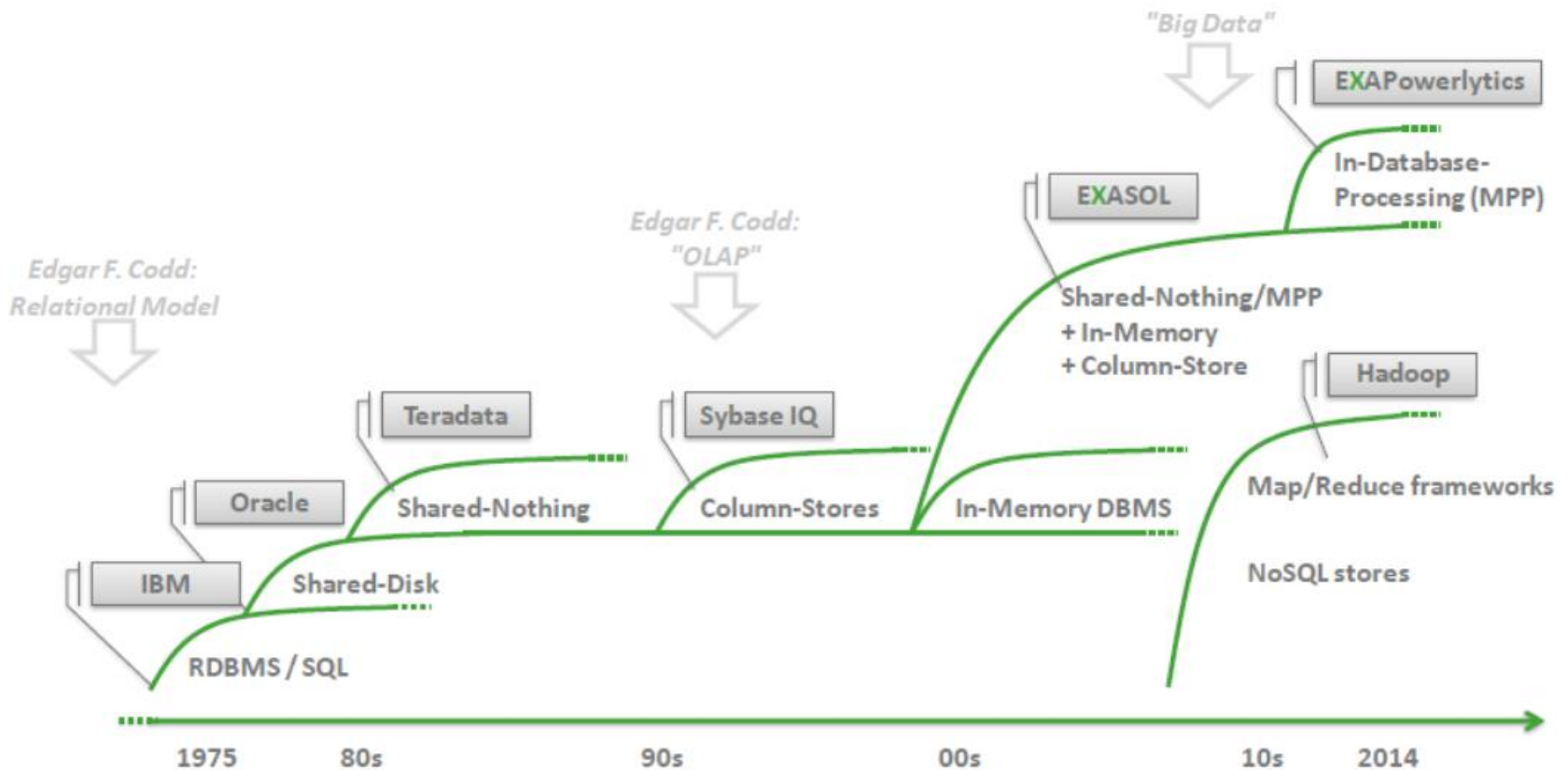


Big Data를 위한 In-Memory DB

# DBMS의 발전





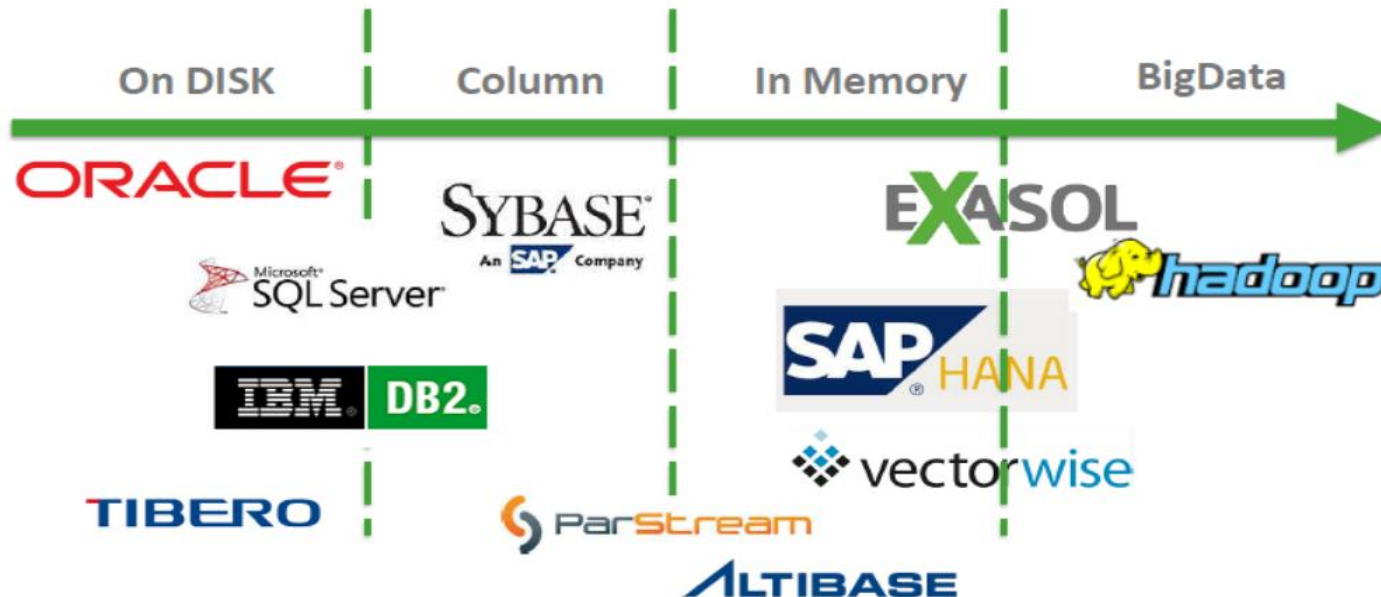
# EXASOL 개요

- **Traditional Database**

- 장점: 트랜잭션 처리에 집중되어 있다.
- 단점: 데이터 사이즈가 증가하면 처리에 한계를 보인다.

- **Hadoop**

- 장점: 비정형 데이터를 처리하는데 최적화 되어 있다.
- 단점: 사용이 어렵고, 분석쿼리를 빠르게 처리하기 위해서는 수많은 노드가 필요하다.  
(c.f: RDBMS BEP= 10 Node 이상 필요)





# EXASOL 개요

## EXASOL

- 사용이 쉬운 Traditional Database의 장점(SQL 사용)을 살리면서도, 빅 사이즈의 데이터를 실시간으로 손쉽게 빠르게 처리 한다.

- TPC-H is the industry standard benchmark for analytical databases

TPC Transaction Processing Performance Council

- 분석 데이터베이스 부문 1위
- 분석 속도 부문 1위
- 데이터 확장성 부문 1위
- 가격 대비 성능 부문 1위



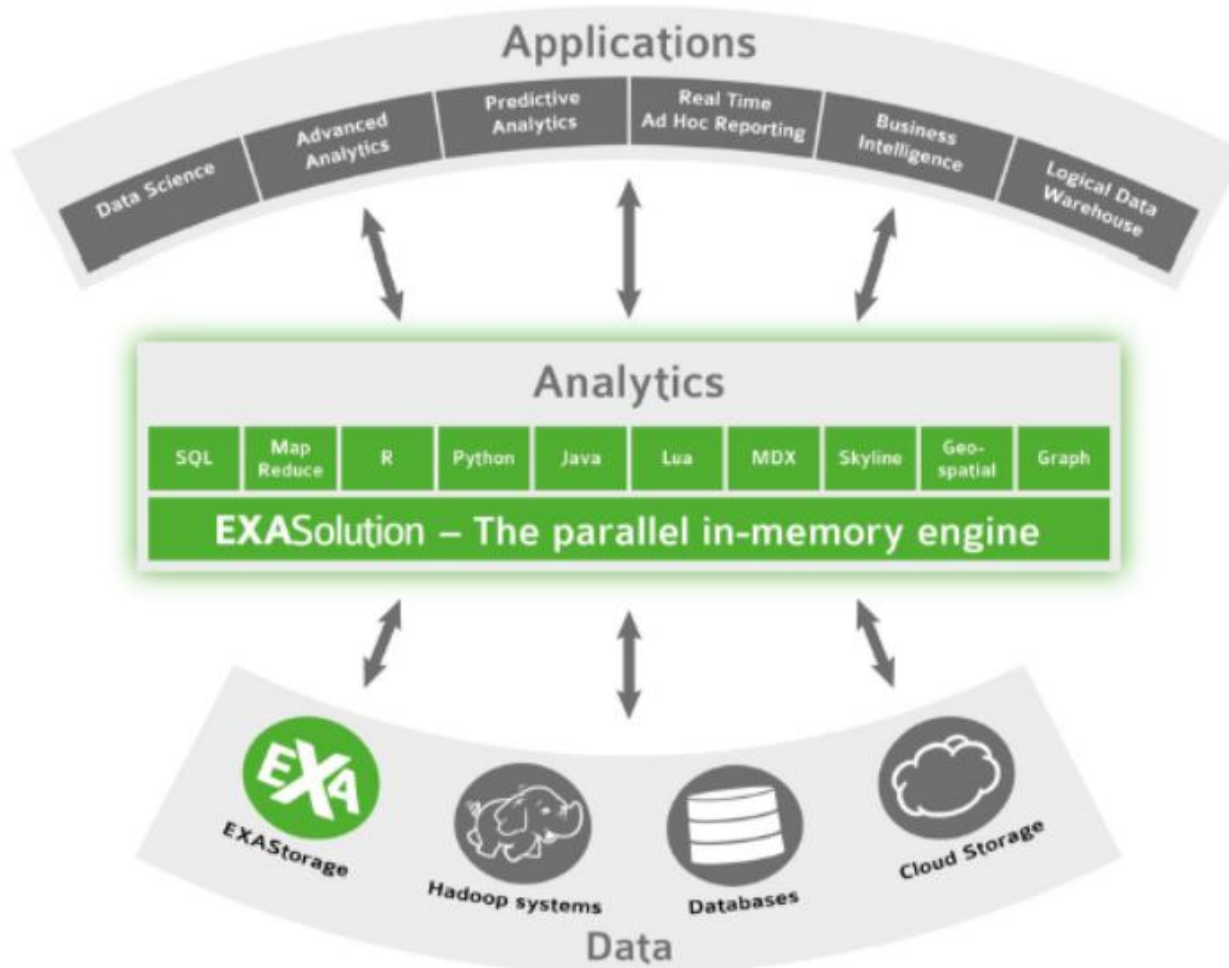
- DB 모델링 없이 다양한 형태의 데이터를 기존DB와 같은 방식으로 사용한다.

- 관련 분야 15년 이상의 경험과 노하우
- 핵심역량 : 대용량 분산 처리 (MPP)
- 독일 회사의 엄격한 품질관리
- 수 백 개의 고객사를 통해 검증된 솔루션
- USA / UK / Israel / Brazil 지사 보유





# EXASOL Architecture



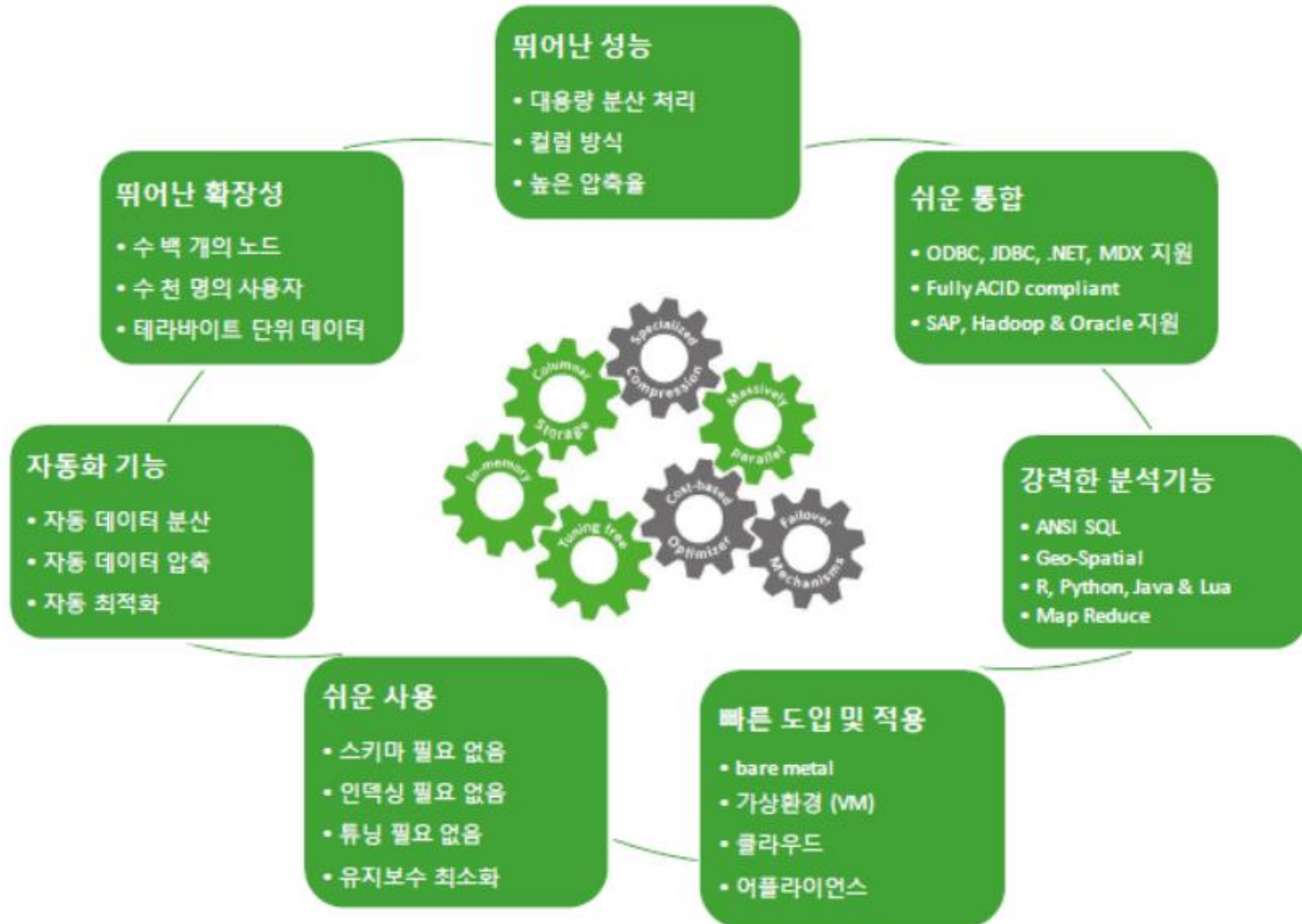
**Leading edge technology**

- World's fastest in-memory engine
- Extendable analytics platform
- Unbeatable scalability
- Minimal TCO



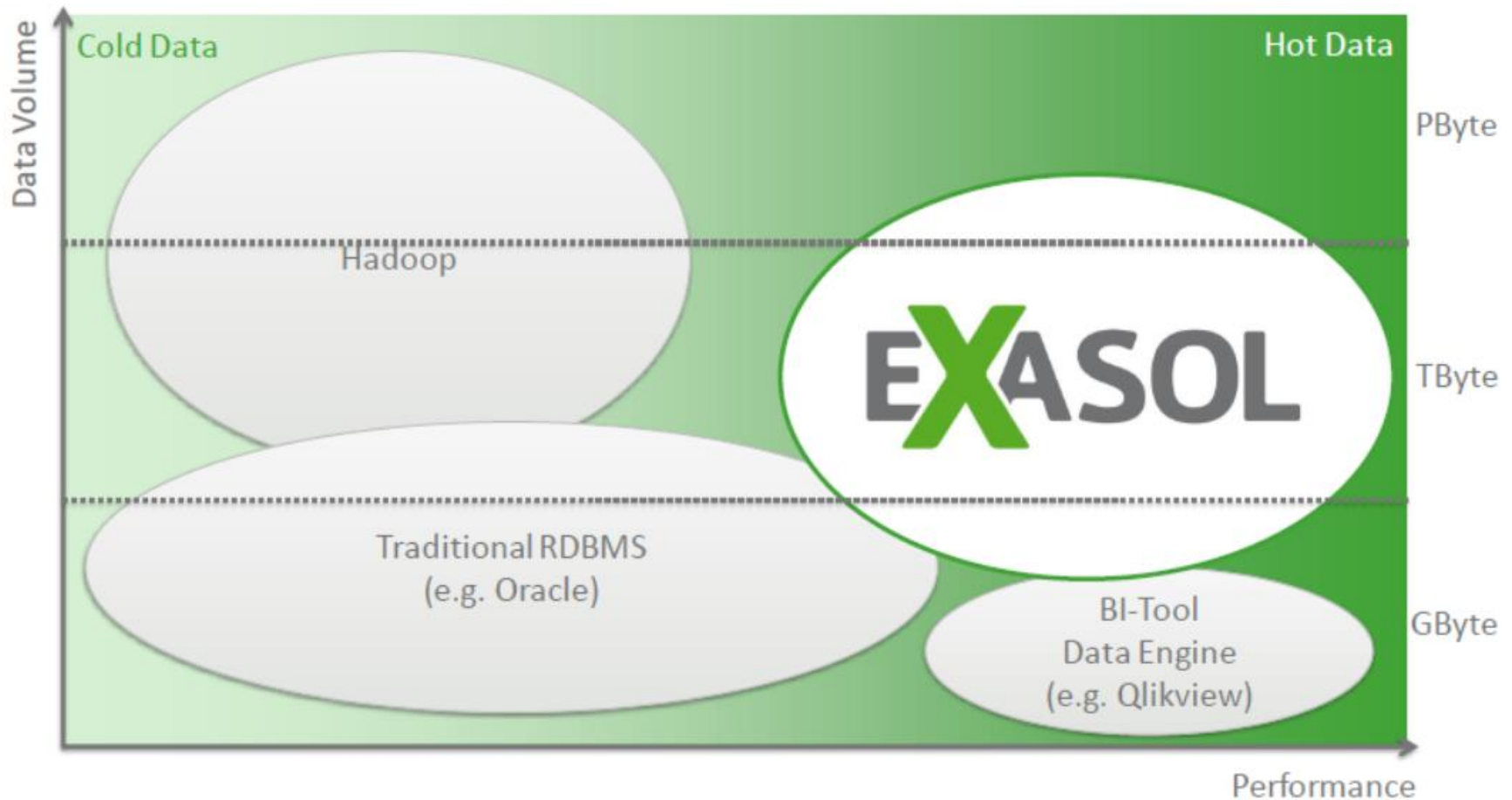


# EXASOL 특징

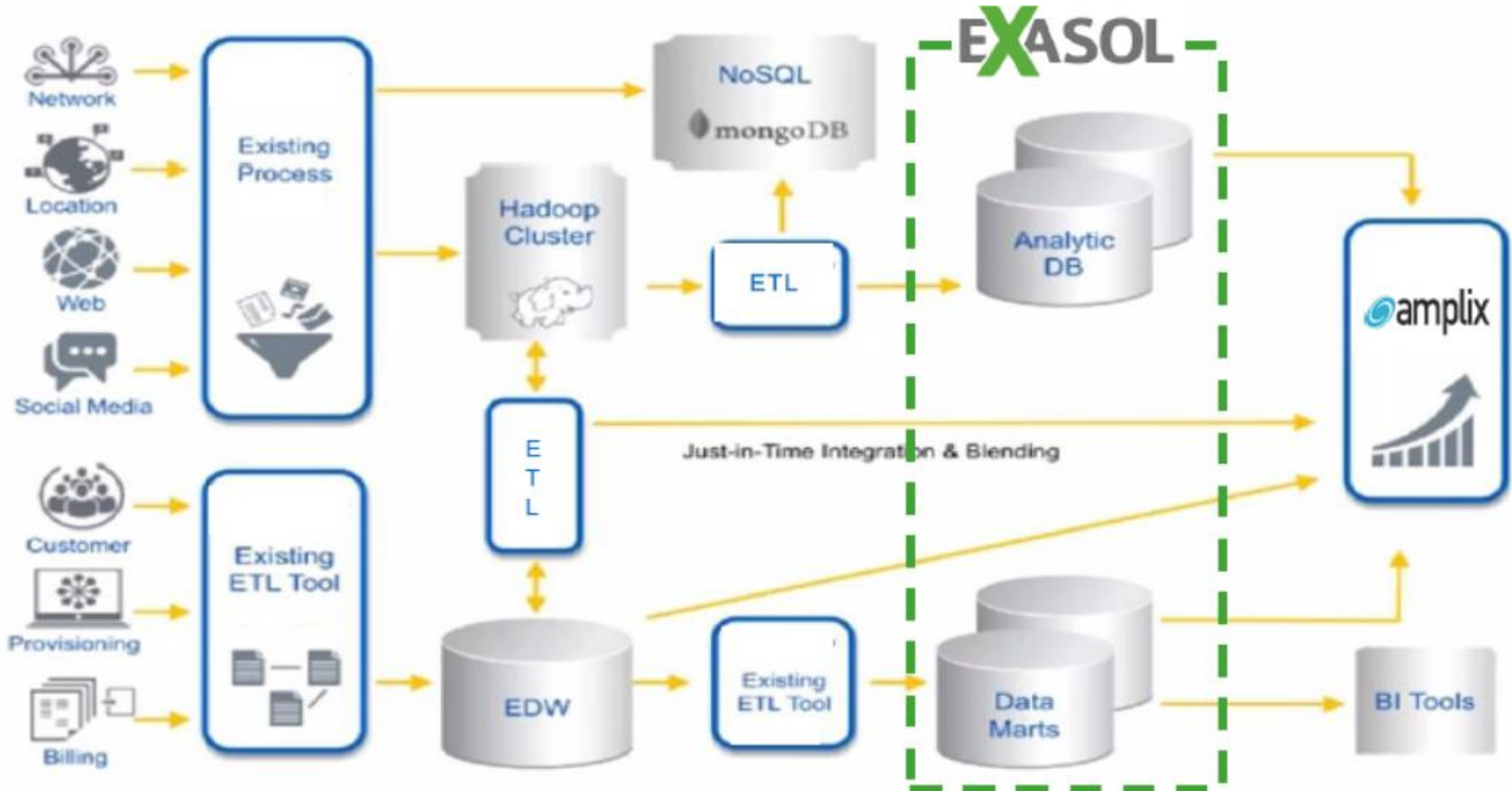




# EXASOL Positioning



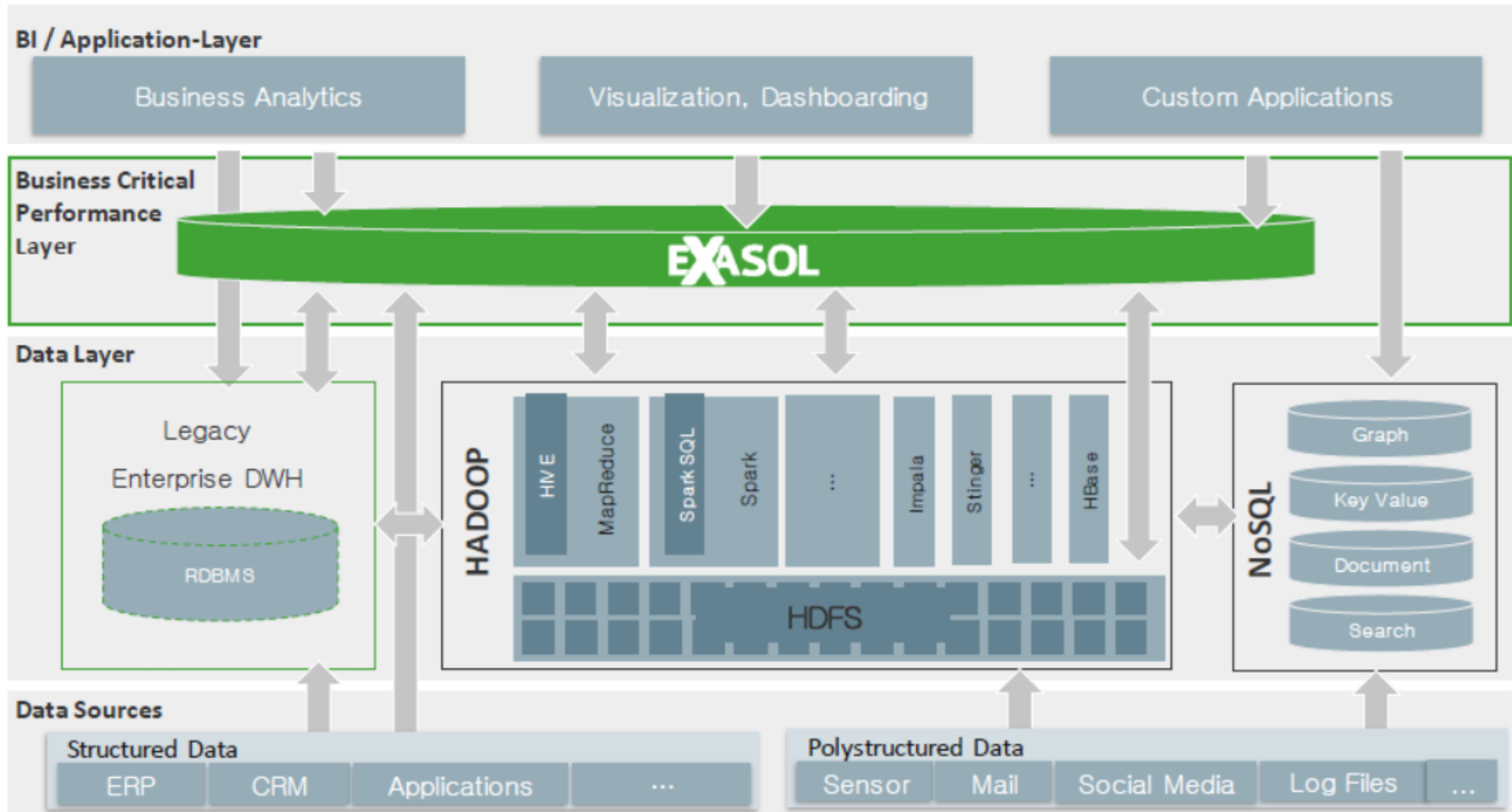
# EXASOL & DW 조합(1/2)







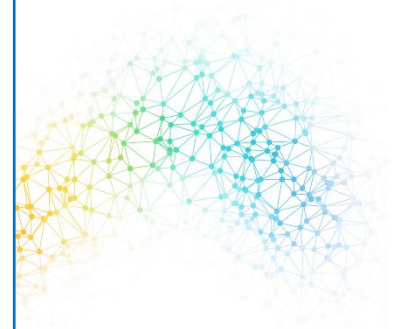
# EXASOL & DW 조합(2/2)



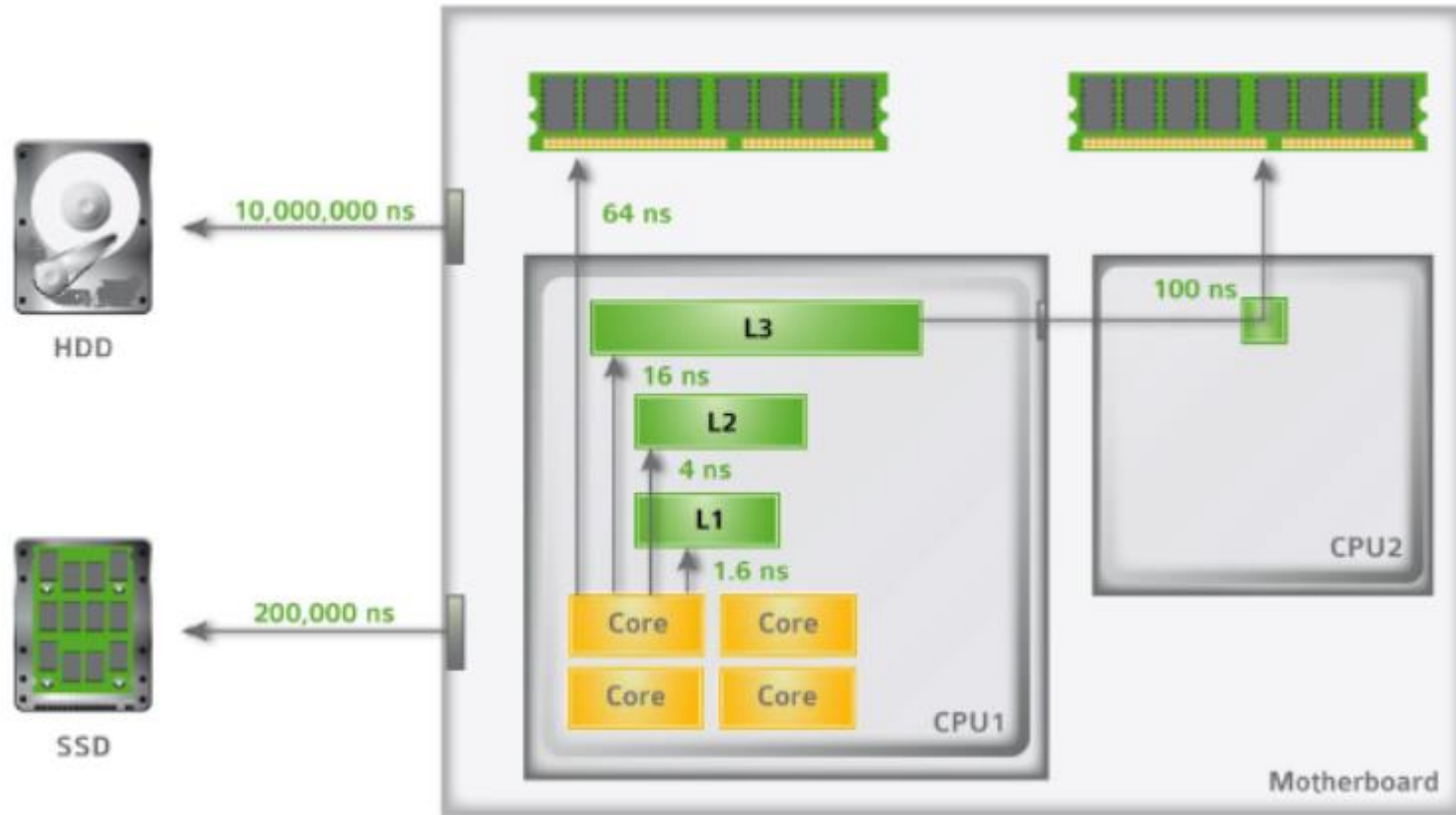


# EXASOL 주요 기능

구분	세부 기능
In-Memory Database	<ul style="list-style-type: none"> <li>• 모든 데이터를 메모리에서 처리</li> <li>• 메모리는 디스크와 같이 I/O 병목 현상이 발생하지 않으므로 대용량의 데이터를 빠르고 쉽게 처리</li> </ul>
Compression & Columnar storage	<ul style="list-style-type: none"> <li>• 압축된 컬럼형 데이터 구조는 Sybase IQ에서 처음 도입된 기술로 지금은 대부분의 분석용DBMS 에 적용되어 있고 OLAP 분석 환경에서 Row 형 데이터 구조에 비해 10~100 배 빠름</li> </ul>
Massively Parallel Processing (MPP)	<ul style="list-style-type: none"> <li>• 오래전 부터 Teradata는 MPP방식으로 시스템을 구성하여 대용량 데이터를 빠르게 처리</li> <li>• 하나의 하드웨어가 아닌 여러 하드웨어가 데이터를 분산하여 처리 하는 방식으로 하드웨어간의 로드 밸런싱 및 최적화가 매우 중요하여 하드웨어와 OS, DBMS을 하나로 묶어 제공</li> <li>• EXASOL은 Linux를 변형하여 자체OS 개발 및 OS 와 DBMS 를 같이 제공하여 안정적인 MPP 보장</li> </ul>



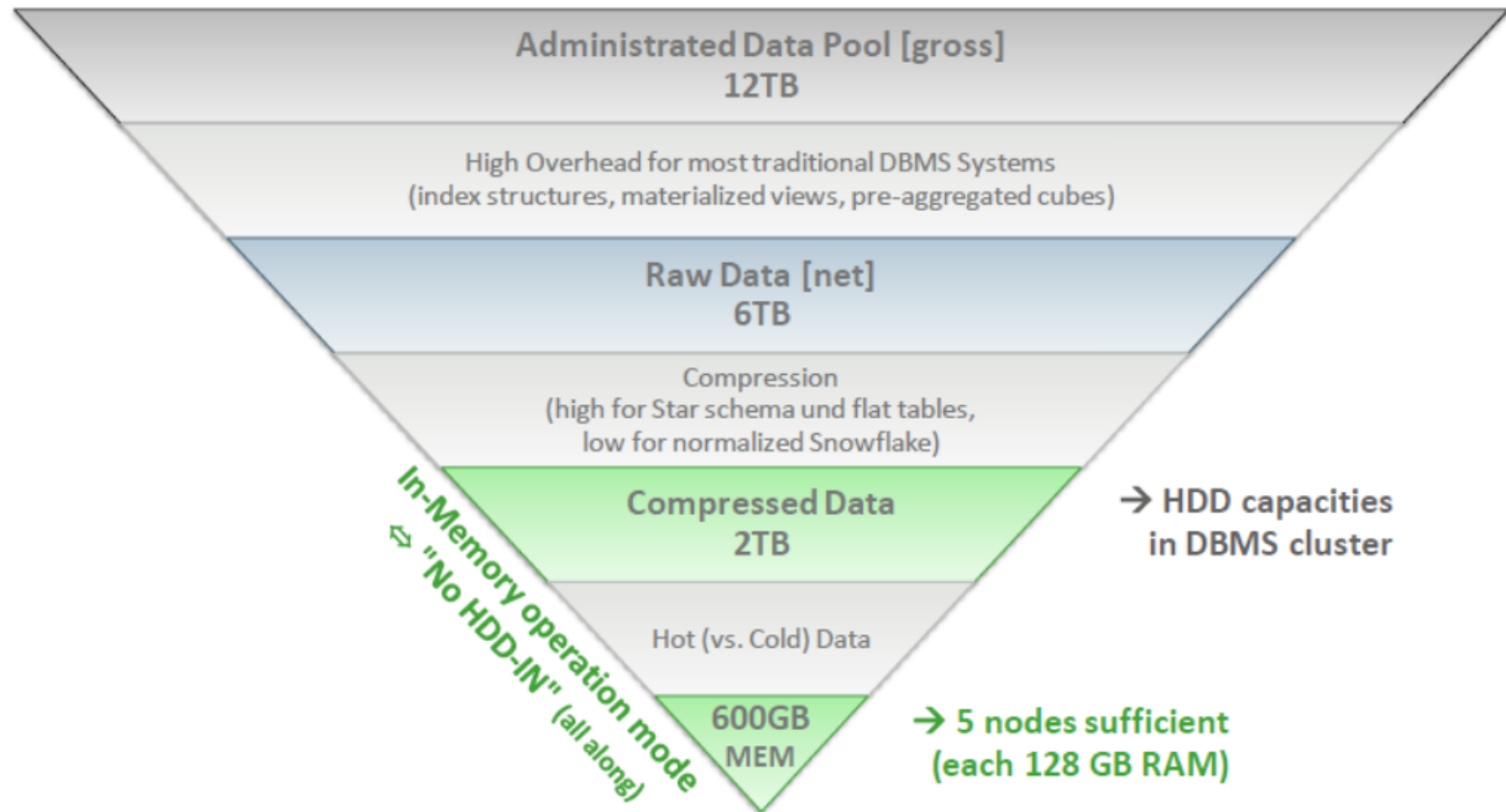
# 기능 : In-Memory DB (Disk 중심의 한계 극복)



- 실제 "In-Memory" 성능은 CPU나 알고리즘 보다는 메모리의 크기에 가장 큰 영향을 받음.



# 기능 : In-Memory DB (Data Size)



## 기능 : Column 방식 (빠른 속도)

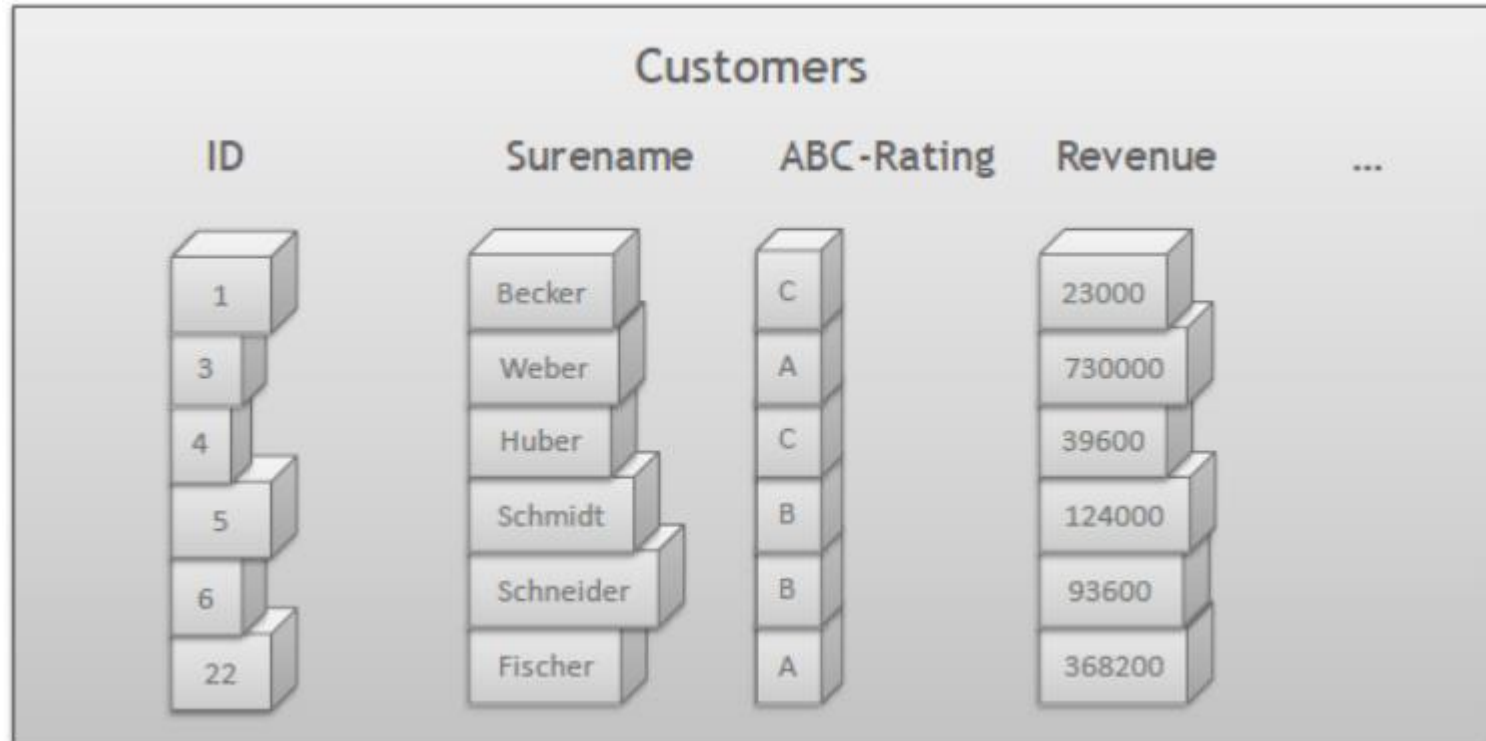
- Row 방식의 DB에서는 불필요한 Column 까지 처리 하는 낭비가 일어남
- Column 방식의 DB에서는 계산에 필요한 Column만을 사용해서 RAM과 IO를 절약
- 그 결과 RDBMS (Row 방식)과는 비교할 수 없을 정도의 빠른 속도로 계산을 수행





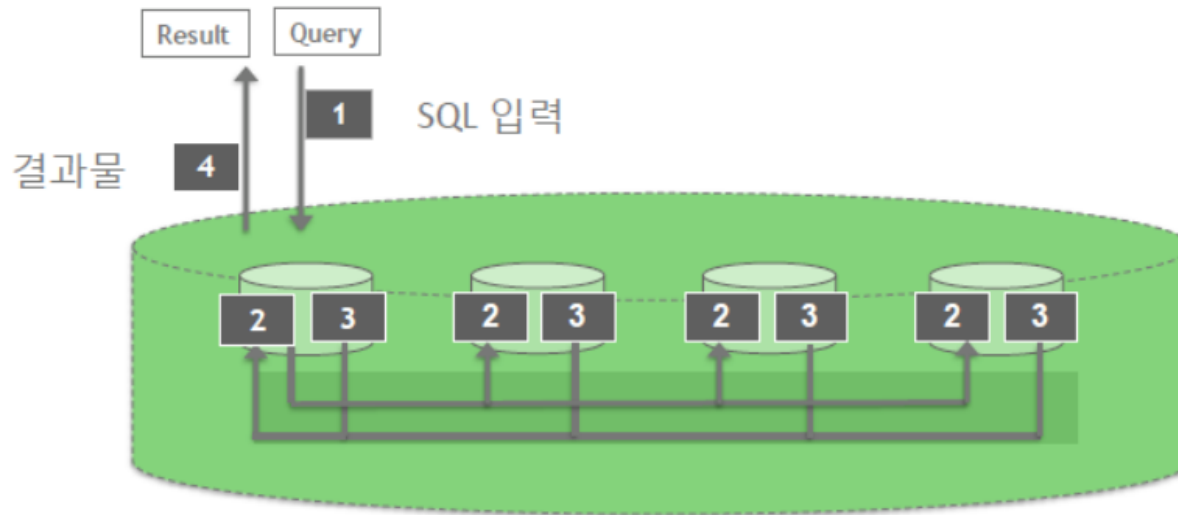
## 기능 : Column 방식 (높은 압축률)

- 같은 Column에 속한 데이터는 단일한 형태이며 비교적 비슷한 길이를 갖고 있음
- 이러한 이유에서 Column단위의 압축은 최고의 성능을 낼 수 있음(약1/5 ~ 1/10)
- 높은 압축률을 바탕으로 메모리 사용을 줄이고 경제성 확보

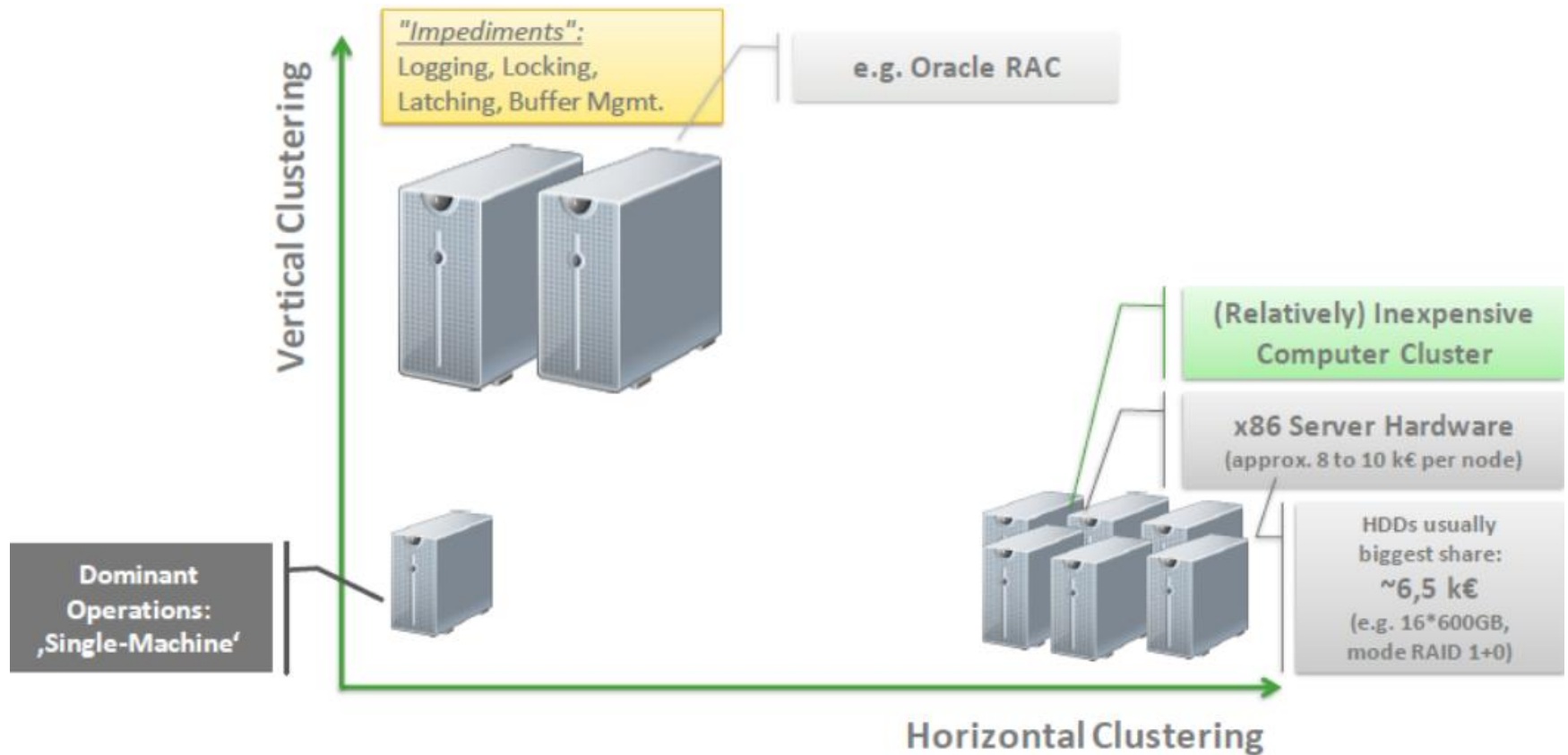


# 기능 : MPP (분산된 SQL 쿼리 프로세스 지원)

- 대용량 병렬 처리를 위한 Shared-nothing 아키텍처



# 기능 : MPP (저렴한 서버의 병렬 연결을 통한 대용량 확장)



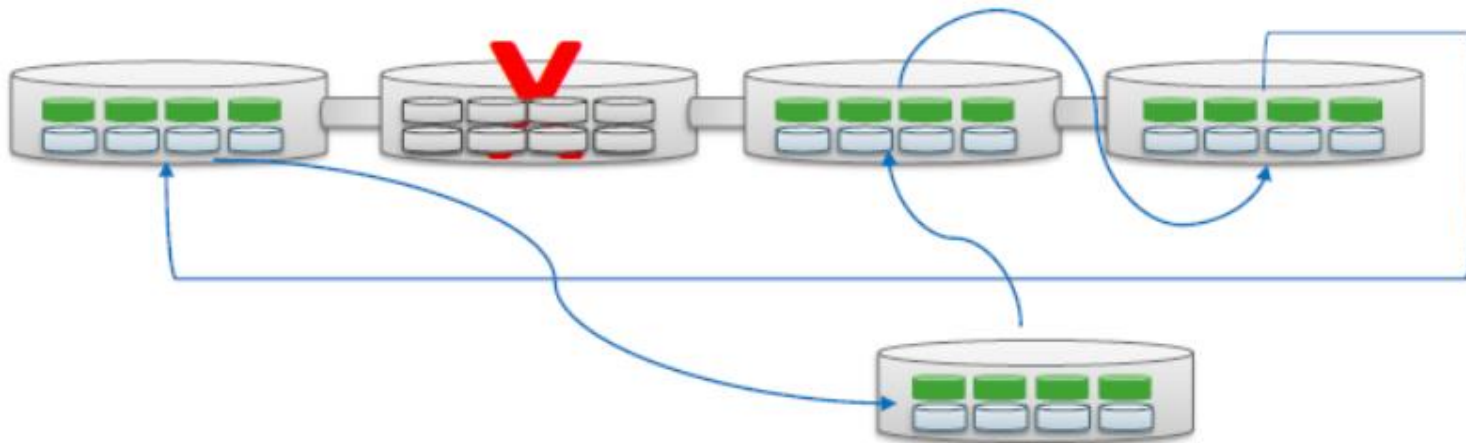
# 기능 : MPP (SPMD 아키텍처와 뛰어난 확장 성능)

- PB단위까지 자연스럽게 데이터 확장 가능 (저렴한 서버의 쉬운 병렬 추가)
- SPMD-Architecture -> IO 프로그램 단의 병목현상 없음
- 자동화된 병력 관리 및 데이터 재분배 지원



# 기능 : MPP (Node Failure 자동 관리)

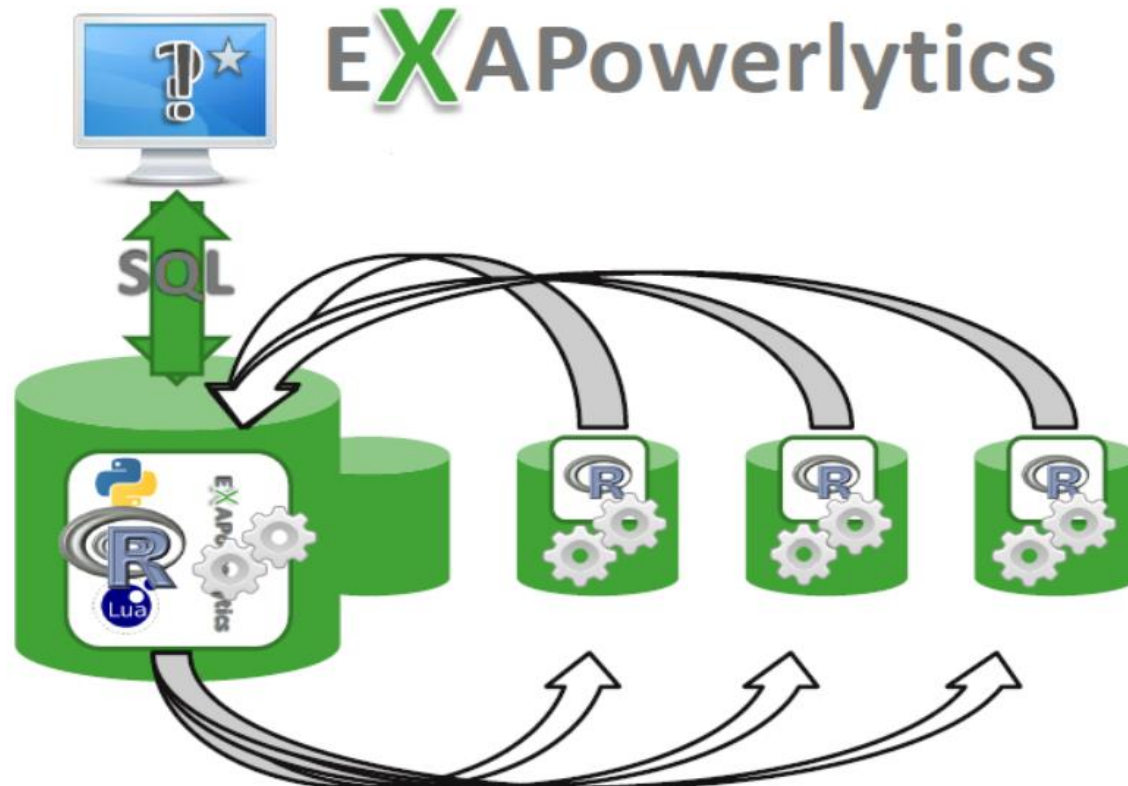
- 더 높은 Redundancy Level 도 대처 가능 : 2개 이상의 노드가 동시에 Fail 할 경우



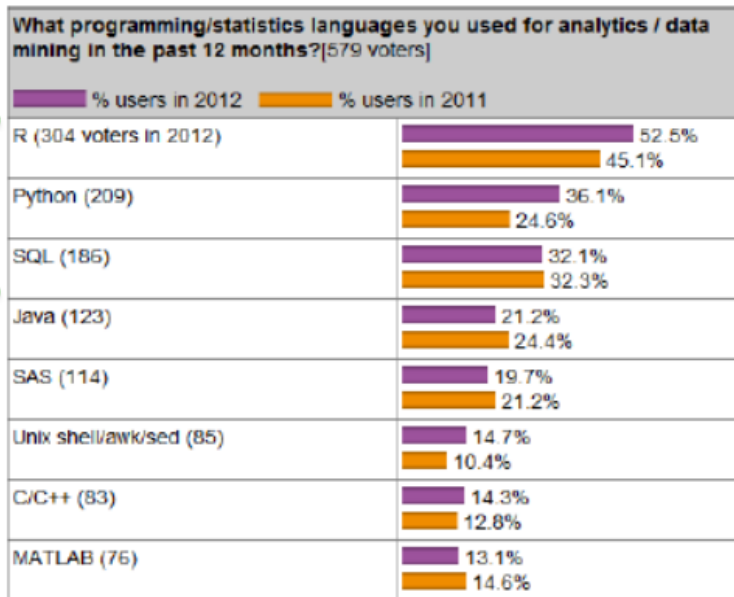


# 기능 : EXAPowerlytics (병렬처리 분석)

- 복잡한 분석을 대용량 병렬 처리를 통해 해결



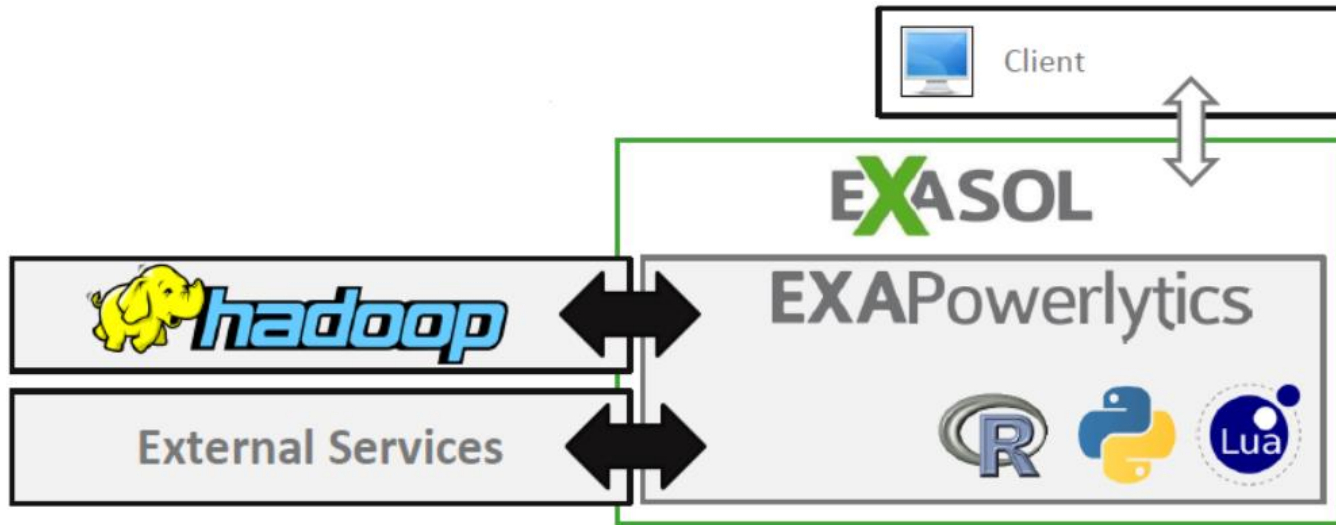
# 기능 : EXAPowerytics (R, Python, SQL 지원)



Perl (52)	9.0%	7.9%
Pig, Hive, or other Hadoop-based languages (39)	6.7%	6.1%
GNU Octave (34)	5.9%	N/A for 2011
Lisp/Clojure (25)	4.4%	0.7% (Lisp only)
Ruby (22)	3.8%	N/A for 2011
Scala (14)	2.4%	N/A for 2011
Julia (2)	0.3%	N/A for 2011
Other (66)	11.6%	12.3%
None (4)	0.7%	1.2%

## 기능 : EXAPowerlytics (Big Data 연계)

- Hadoop을 이용한 대용량 데이터 전처리
- R을 이용한 대용량 통계 분산 분석



XING

OTTO

xplosion  
interactive

myThings  
making impressions personal

BAUR

united  
internet  
dialog

# 기능 : 자동화된 Database 관리

## 기존 방식의 DBMS

- Partition Pruning
- Materialized Views
- Bitmap Indexes
- Bitmap Join Indexes
- Index Organized Tables
- DIMENSION Objects
- Current Statistics
- Histograms
- Table Compression
- CACHE
- PCTFREE
- ...

## EXASOL

- Optional: Distribution Keys
- Optional: Replication Boarder
- ✓ 자동 테이블 분석
- ✓ 자동 인덱스 생성
- ✓ 강력한 쿼리 최적화
- ✓ 자동 데이터 분산 관리
- ✓ 자동 데이터 중복 관리
- ✓ 자동 Failover 관리

최대 95%  
노력 절감

# 기능 : 자동화된 Database 관리 (Backup Schedule)

- Services
  - EXASolution
  - EXASStorage
  - EXAoperation
  - Monitoring
  - Support
- Legacy Services
  - Scheduler
  - Backup
- Configuration
  - Software
  - Nodes
  - Network
  - Access Management

### EXASolution Backup Scheduling

Active	Minute	Hour	Day	Month	Weekday	Archive	Level	Expire	Comment	Add	Delete
<input checked="" type="checkbox"/>	0	0	*	*	0	v0004	0	2w	Weekly Full-Backu		<input type="checkbox"/>
<input checked="" type="checkbox"/>	0	0	*	*	1,2,3	v0004	1	3d	Daily Increment		<input type="checkbox"/>
<input checked="" type="checkbox"/>	0	0	*	*	*	v0004	0			<input type="checkbox"/>	<input type="checkbox"/>

### Backup Schedules

	Creation Date	Expire Date	2/2 - 2/8/14							2/9 - 2/15/14							2/16 - 2/22/14							2/23 - 3/1/14							3/2 - 3/8/14						
			2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	1	2	3	4	5	6	7	
+ Current backup schedule	2/9/14	2/23/14	14 Days																																		
Level 0 backup	2/9/14	2/23/14	14 Days																																		
Level 1 backup	2/10/14	2/13/14	3 Days																																		
Level 1 backup	2/11/14	2/14/14	3 Days																																		
Level 1 backup	2/12/14	2/15/14	3 Days																																		
Level 1 backup	2/13/14	2/16/14	3 Days																																		
Level 1 backup	2/14/14	2/17/14	3 Days																																		
Level 1 backup	2/15/14	2/18/14	3 Days																																		
+ Next backup schedule	2/16/14	3/2/14	14 Days																																		
Level 0 backup	2/16/14	3/2/14	14 Days																																		
Level 1 backup	2/17/14	2/20/14	3 Days																																		
Level 1 backup	2/18/14	2/21/14	3 Days																																		
Level 1 backup	2/19/14	2/22/14	3 Days																																		
Level 1 backup	2/20/14	2/23/14	3 Days																																		
Level 1 backup	2/21/14	2/24/14	3 Days																																		
Level 1 backup	2/22/14	2/25/14	3 Days																																		

Format:  Hour  Day  Week





# 감사합니다



서울시 구로구 디지털로 30길 28 마리오타워7층 701호  
TEL: 02-558-9130 / FAX: 02-558-7868 / [www.in-con.biz](http://www.in-con.biz)