


7.2 빅데이터 분석에 활용되는 기술

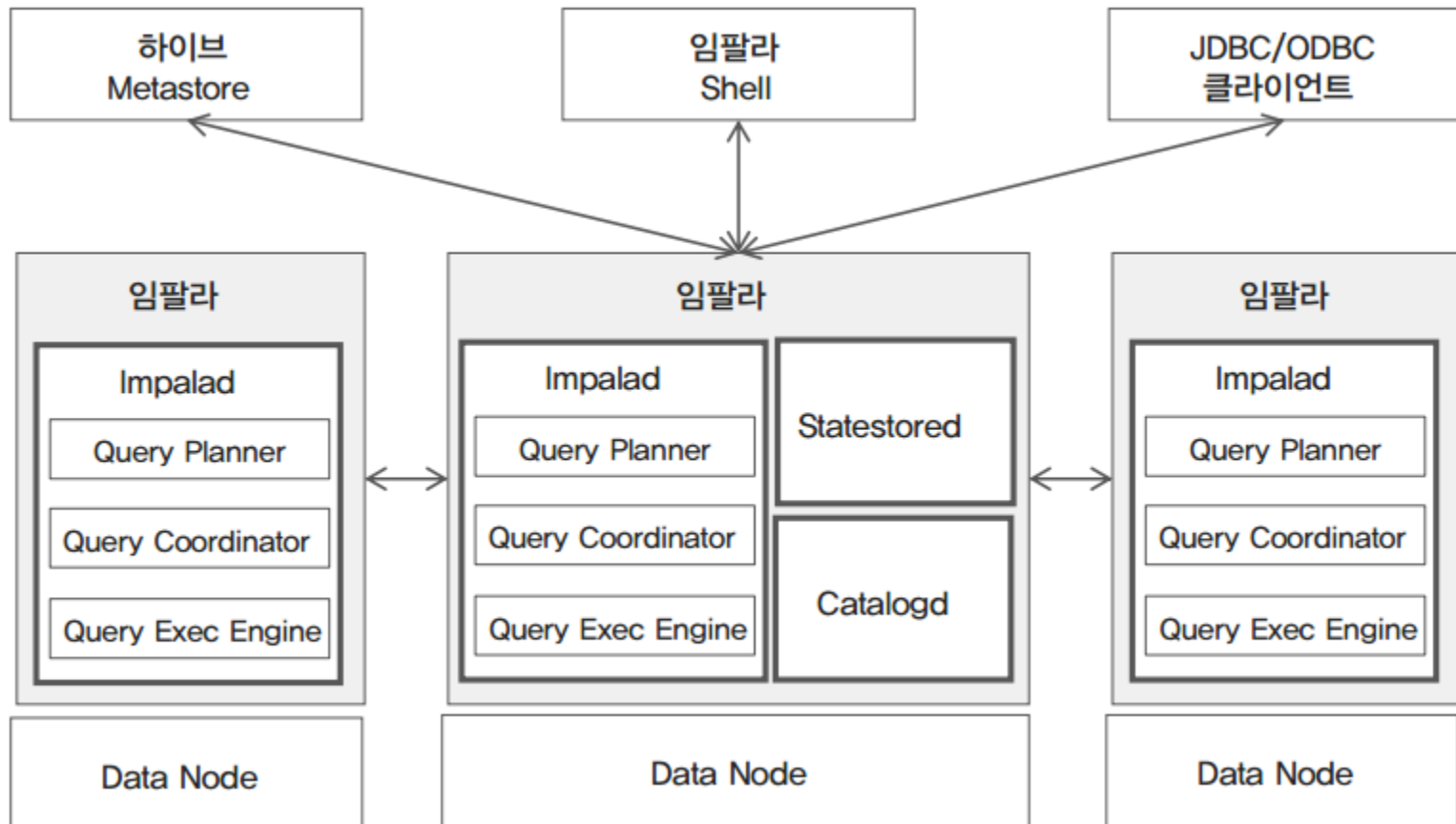
임팔라 소개

표 7.1 임팔라의 기본 요소

| | | |
|----------|---|---|
| 공식 홈페이지 |  | http://impala.apache.org/ |
| 주요 구성 요소 | Impalad | 하둡의 데이터노드에 설치되어 임팔라의 실행 쿼리에 대한 계획, 스케줄링, 엔진을 관리하는 코어 영역 |
| | Query Planner | 임팔라 쿼리에 대한 실행 계획을 수립 |
| | Query Coordinator | 임팔라 잡리스트 및 스케줄링을 관리 |
| | Query Exec Engine | 임팔라 쿼리를 최적화해서 실행하고, 쿼리 결과를 제공 |
| | Statestored | 분산 환경에 설치돼 있는 Impalad의 설정 정보 및 서비스를 관리 |
| | Catalogd | 임팔라에서 실행된 작업 이력들을 관리하며, 필요 시 작업 이력을 제공 |
| 라이선스 | Apache | |
| 유사 프로젝트 | Tez, Spark SQL, Drill, Tajo | |

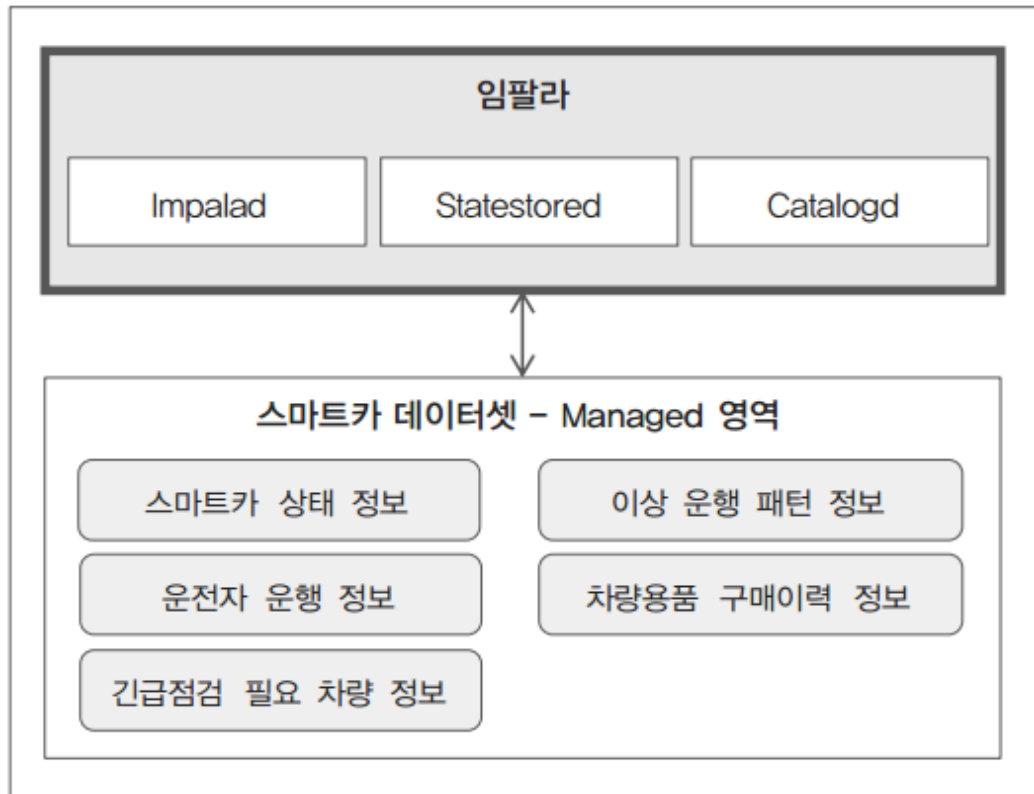
7.2 빅데이터 분석에 활용되는 기술

임팔라 아키텍처



7.2 빅데이터 분석에 활용되는 기술

임팔라 활용방안



- Impala Engine을 이용, 대용량 데이터를 실시간으로 Ad-Hoc 분석
- Managed 영역의 스마트카 데이터셋을 빠르게 탐색