

실시간 관세행정 경영지원이 가능한 CDW 환경 고도화

박근호, 부장, 펜타시스템테크놀러지(주)



Agenda

1. 프로젝트 개요
2. 실시간 DW 구축 배경 및 기대효과
3. CDC 솔루션 선정 기준
4. 관세청 CDW 구성도
5. CDC를 활용한 CDW Architecture
6. 실시간 DW 구축 시 고려 사항
7. 향후 발전 방향

프로젝트 개요

1. 프로젝트명 : 실시간 관세행정 경영지원이 가능한 CDW 환경 고도화
2. 프로젝트 기간 : 2008.06 ~ 2008.12
3. 적용 솔루션 :
 - IBM InfoSphere DataStage 7.x
 - IBM InfoSphere CDC 6.0 For Oracle
 - IBM InfoSphere CDC 6.0 For DataStage
4. 적용 범위 :
 - 전체 CDW 관련 테이블 3,000여개
 - 이중 Real-Time 요건이 있는 중요 400여개 테이블 CDC 연계

실시간 DW 구축 배경 및 기대효과

배치 형태의 CDW 운영 환경으로 인해 CDW 시스템 운영과 활용에 많은 제약 사항을 CDC 솔루션을 도입 활용함으로써 이를 개선 Speed한 관세 행정 시스템 구축을 할 수 있었습니다.

구축배경

- 기업정보, 입출항정보, 적하목록 정보 및 여행자정보 등 **실시간 분석에 필요한 정보가 야간배치작업(Batch) 으로 실시간 정보조회 곤란**
- CDW ETL프로그램이 원시데이터의 변경 사항을 Time Stamp를 이용하는 방식으로 되어 있어 **운영계 Data와 CDW의 Data 불일치 발생**
- **매년 데이터가 증가**되므로 데이터를 추출하는데 소요되는 시간이 그만큼 증가하여 결과적으로 **익일 업무 오픈 시간까지 배치가 수행**되는 경우가 발생함

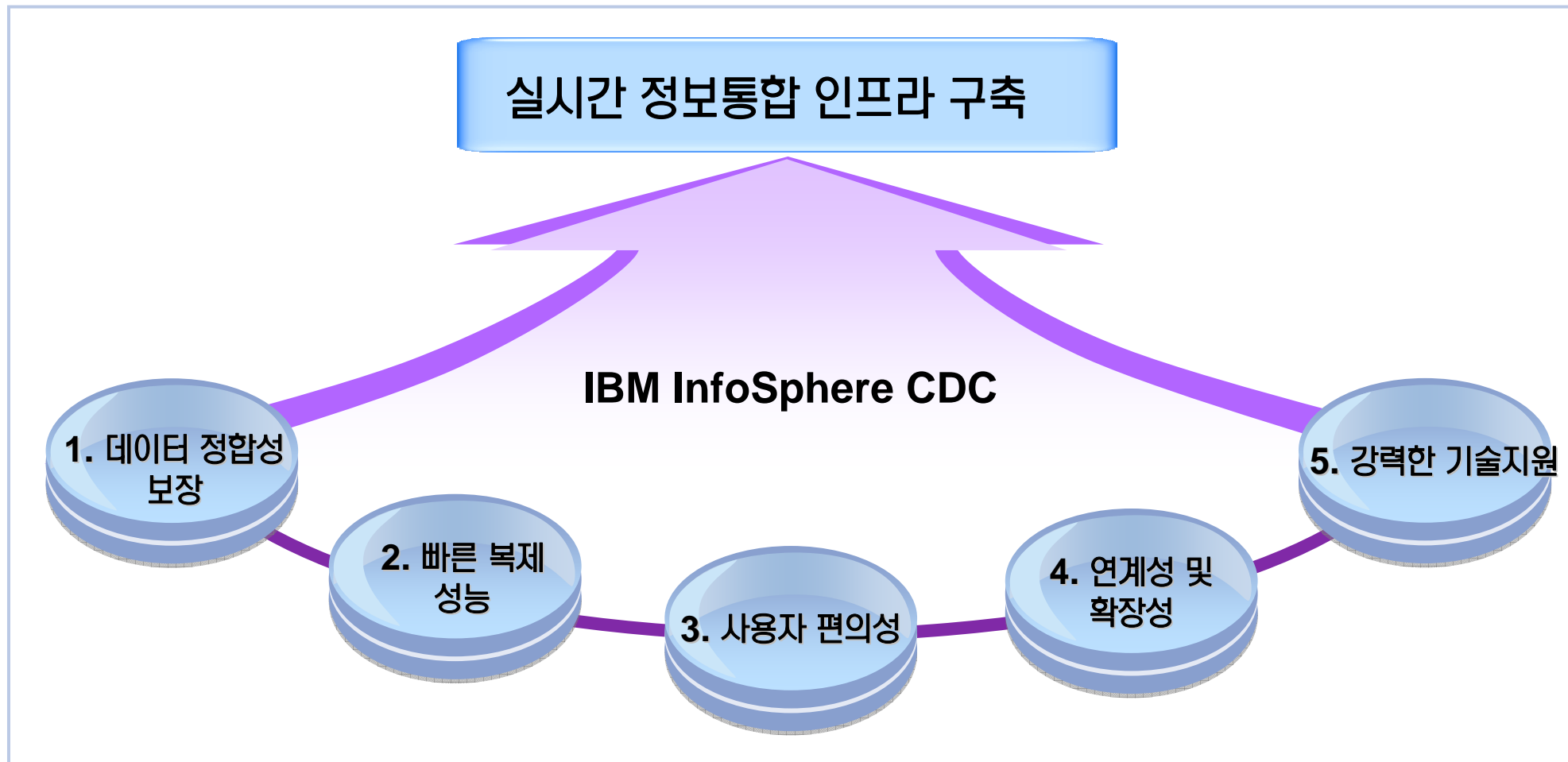
CDC
솔루션
도입

기대 효과

- 기존 CDW 축적 정보와 **실시간 생성 정보간의 연계분석** 및 정보 정합성 확대로 차세대 분석환경 구축
- 정보제공 창구를 CDW로 일원화하여 통계정보 등의 일관성 확보
- CDC추출이 수행되고 있는 상황에서도 Application구동이 가능하여 가용성 향상
- 변경데이터 추출 과 동시에 변환이 가능하므로 많은 배치 ETL작업 축소와 배치 실생 시간이 감소됨
- 실시간 추출하여 전달되는 구조이므로 데이터 양이 증가하더라도 시스템 및 운영에 미치는 영향이 적음

CDC 솔루션 선정 기준

CDC 솔루션 도입에 따른 선정 기준은 정합성, 복제 성능, 사용자 편의성, 연계성 및 확장성과 마지막으로 솔루션 공급 업체의 기술 지원 능력을 고려하였습니다.



관세청 CDW 구성도

관세청 CDW 시스템은 운영 시스템에서 발생한 Data를 일 배치 형식으로 DW시스템 및 Dart에 적재하는 전통적인 방식의 DW 운영 시스템 이었습니다.



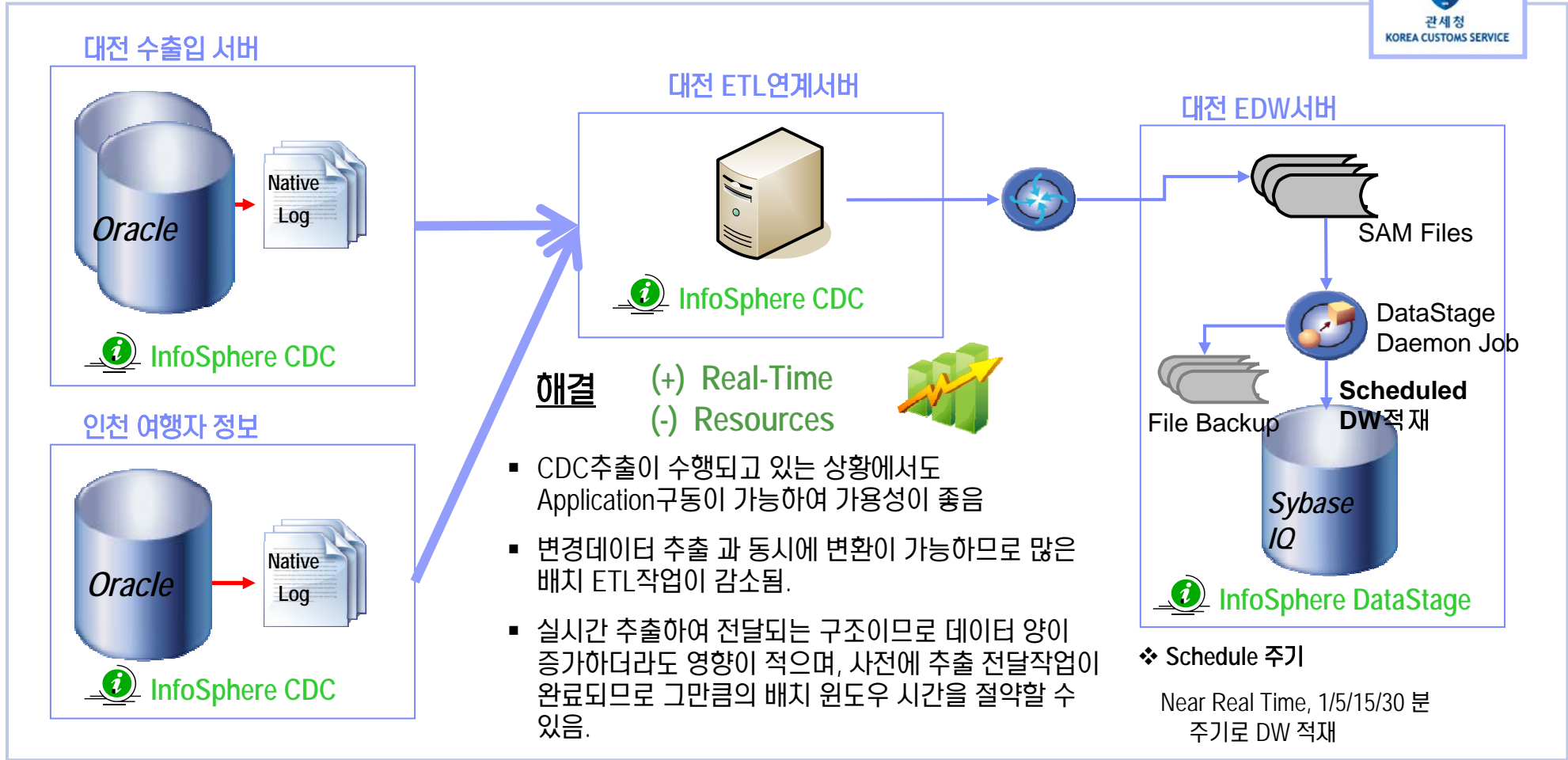
문제점

- 다년간 축적된 Data로 인해 배치 작업 시간의 증가로 익일 업무 오픈 시간까지 배치가 수행되는 경우가 발생함.
- 또한, 야간 배치 작업 도중 에러 발생 시 업무시간에 재 작업이 불가하여 업무에 주는 영향이 매우 큼.
- 실시간 분석 업무 요건 충족 불가



CDC를 활용한 CDW Architecture

Daily 배치 방식의 전통적인 DW 운영 환경에 주요 업무에 Real-Time 솔루션을 적용 현업 사용자의 실시간 업무 분석 요구 충족과 DW 운영 방식을 개선하였습니다.



실시간 DW 구축 시 고려 사항

- 실시간 Job과 Batch Job의 균형
- 시스템 성능을 고려한 Near Real-Time Job 주기 조정
- CDC 및 ETL 솔루션의 연계성
- CDC 솔루션 운영 주체

향후 발전 방향

- 단순 복제 시 DB To DB 방식에서의 복제 방식 전환
- DW -> Mart로의 실시간 업무 확대 적용

Q & A