

**EDI문서에 대한 XML 변환규칙 및 성능분석**  
*(XML Translation Rule and Performance Analysis For EDI Document)*

Aug. 15, 2003  
 Samsung SDS  
 Soo-Cheol Shin

Samsung SDS

**Abstract**

요즈음의 전자상거래 환경하에서 ebXML 및 XML/EDI 보급이 활발하게 되면서 여러가지로 표준과 문서의 XML 문서화와 문서화된 XML 문서를 해독하고 작성하는 부분에 대한 내용이 주요 ISSUE 사항이 될 것이다. 위 사항과 관련하여 XML/EDI 및 ebXML 내에서의 사용 가능한 XML화된 표준문서의 Schema 생성과 생성된 Schema의 성능을 Test 한 결과를 제시하고자 함.

Samsung SDS

**Agenda**

시각화면서..  
 기존 EDIFACT의 구분 정의  
 KEDIFACT Mapping 1(문서의 Header 구조)  
 KEDIFACT Mapping 2(문서의 Body 구조)  
 KEDIFACT Mapping 3(문서의 Tail 구조)  
 Schema 구성 형식(문서의 Header 구조)  
 Schema 구성 형식(문서의 Body 구조)  
 Schema 구조  
 Test 대상 Schema 구조  
 Test 결과

Samsung SDS

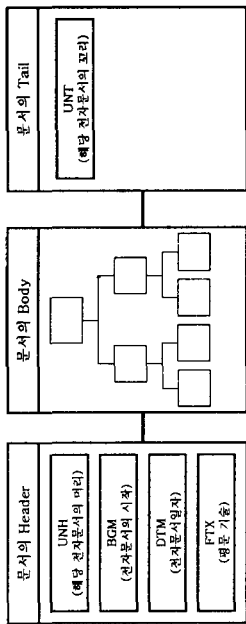
**시각화면서..**

- 기존 EDI 구조가 너무 복잡성으로 인하여 XML로 변환시 변환 규칙을 적용하여 단순화 시킴.
- XML/EDI Format에 Header를 정의함으로써 KEDIFACT MIG의 전자 문서 Header부분을 소화함.
- 기존의 KEDIFACT에서 제시하는 기준을 100% 따름
- KEDIFACT를 효율적으로 수용하기 위하여 고정부분과 변동부분으로 분리하여 Schema 생성
- XML/EDI Group의 Guide를 기준으로 작성
- XML Template 문서구조 단순화로 속도 및 가독성 증가
- 본 문서는 자동차 회사를 기준으로 Test 한 결과를 제시함.

Samsung SDS

### 기본 EDIFACT의 구분 정의

- 기본 EDIFACT의 MIG를 Header, Body, Tail의 3가지로 구분하는 정의이다.
- 문서의 Header와 Tail의 구분은 UNH, BGM, DTM, FTX와 같은 요소가 포함된 경우이다.
- 문서의 Body와 Tail의 구분은 UNT와 같은 요소가 포함된 경우이다.
- 문서의 Tail: 해당 문서의 종료에 사용되는 정보를 담고 있다. (모든 문서 공통)



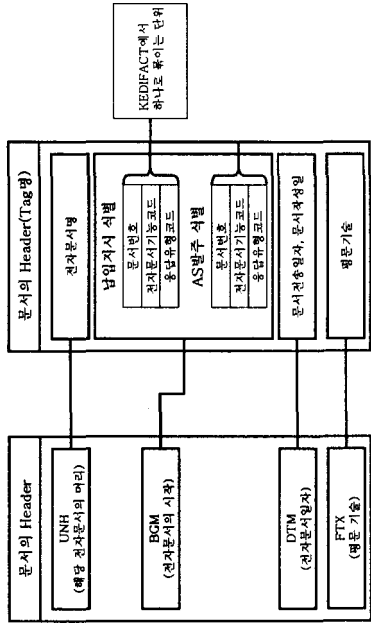
- EDIFACT에 적용되는 건재문서의 구조를 다음의 구조에 Mapping할 수 있다.
- Header부분에 포함된 모든 요소가 Body로 포함될 수 있다.
- 기본 EDIFACT에서 이야기하는 모든 구조를 포함시키는 XML/EDI schema는 EDIFACT의 정보내용을 중심으로 작성하여 각 환경처에서 문서작성시 형식 및 용도를 용기시킴.

Samsung SDS

6

### KEDIFACT Mapping 1(문서의 Header)

- 기본 EDIFACT의 문서 Header의 내용을 XML/EDI Schema에 사용한 Tag명들의 Mapping한 그림
- 결과문서의 Header는 아래의 구조로 항상 동일하게 적용된다.

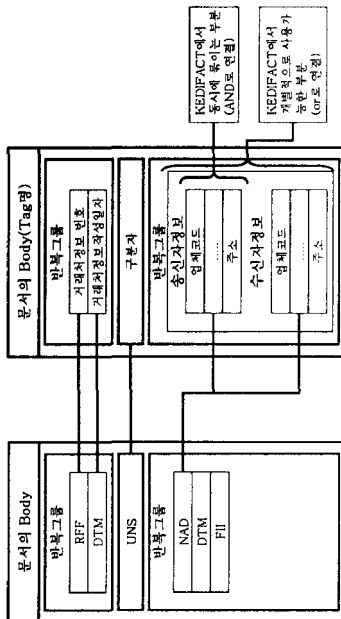


Samsung SDS

6

### KEDIFACT Mapping 2(문서의 Body)

- 기본 EDIFACT의 문서 Body의 내용을 XML/EDI Schema에 사용한 Tag명들의 Mapping한 그림.
- 결과문서의 Body는 EDIFACT의 각 정보(항목)를 그대로 사용.
- 해당 Schema를 정의하는 XML/EDI Schema에 사용.
- MIG를 정의한 문서로 단편시 해당 MIG에 맞는 항목을 해당 정보 이름의 속성으로 정의하여 사용.

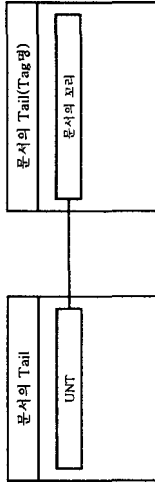


Samsung SDS

7

### KEDIFACT Mapping 3(문서의 Tail)

- 기본 EDIFACT의 문서 Tail의 내용을 XML/EDI Schema에 사용한 Tag명들의 Mapping한 그림
- 결과문서의 Body는 EDIFACT에서 제공하는 MIG의 Mapping되는 모든 정보의 이름을 명시하여 사용



Samsung SDS

8



## Test Plan

- EDI 구조의 모든 MIG Tag를 기본으로 하여 Schema를 작성할 것입니다.
- Node의 속성은 기존 EDI에 모든 MIG Tag를 사용하고 각각에서 사용되는 Qualifier를 모두 XML Tag로 분리해내기 때문에 전체 노드수는 170개가 됩니다.

Samsung SDS

## Test Plan

- Message Message의 Sender-Info가 들어가게 되고 실제 데이터가 포함되지는 MessageBody에 각 MessageBody의 MessageBody를 포함시키는 형식입니다.
- 실제 Data의 양이 적어지고 각 노드와 숫자 및 Depth가 줄어들게 됩니다.
- 그리고 각 Node안에 Attribute, EDIFACT의 MIG 데이터를 포함하여 처음으로 하게 되므로 해당 데이터의 내용은 EDI Format으로 전송하는 것이 가능합니다.
- 또한, 실제 데이터를 전송할 때에서 XML/EDI로 전송하는 것이 부가적으로 한층과 일체에서 현재 EDI 문서를 작성하고 통합체에서 실제 XML/EDI문서를 확인할 수 있는 것이 가능합니다.
- 본문서는 DELING후의 데이터를 반복하여 작성함으로써 전체적으로 데이터를 풍부함을 가져오면 실제 협력업체에서 작업을 할 때를 풍부합니다.

Samsung SDS

## Test Plan

- 내림차순 데이터의 모든 내용 MessageBody에 포함되게 되도록 실제 반복되는 부분이 될 수 없이 들어가게 됩니다.
- 각각의 한층과 일체에서 필요한 데이터의 상위에 별도의 MessageBody의 Tag를 붙임으로써 노드가 증가됩니다.
- Depth가 6이라는 것은 최상위의 데이터를 처리하는 부분에서는 볼 수리가 없어도 많은 양의 데이터를 처리할 경우에는 무리가 없습니다.
- Schema의 구조에서 Depth는 같은 노드도 되나 실제 노드의 숫자가 줄어들게 됩니다.

Samsung SDS

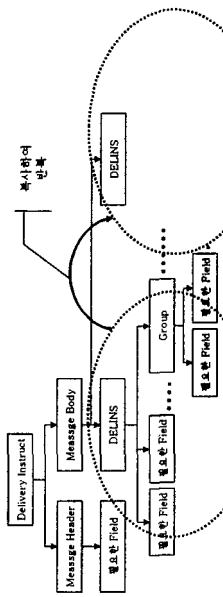
## Test 환경

- Server : HP-UX K460 (OS: HP-UX 11.11)
- Web Server : Netscape Web Server 4.2i (Servlet Engine Include)
- Web Server Tuning : Thread 200으로 Up, Heap Size 130M로 설정.
- DOM API를 이용하여 template XML공시를 Load하여 해당 Node에 값을 Append한 후 반복되는 Node를 집합을 삭제하여 Insert하는 작업을 수행

Samsung SDS

## Test 결과

### Test 결과(전체문서를 반복할 경우)



- DELINS 아래의 모든 부분을 반복하는 형태로 전체적으로 중복필드를 포함하고 있음.
- DELINS아래의 정보의 수가 많으면 많을수록 속도는 느려질수 있음.
- 전체의 구조는 각 안정화 단계별로 필요한 필드만을 가져가므로 DELINS 아래의 필드는 최소화되어 있을 수 있음.

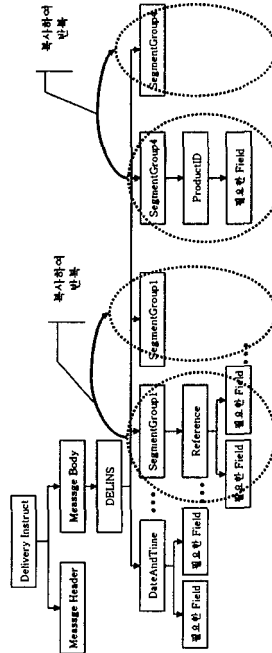
## Test 결과

### Test 결과(전체문서를 반복할 경우)

기본 Schema의 Template 문서	Schema1의 Template 문서	Schema2의 Template 문서
Template 문서의 Node 수(MsgBody)	170개	38개
Web상에서의 DOM API 이용속도 Test(DB Data 280건 기준)	30초	12초
Web상에서의 DOM API 이용화일 크기 Test(DB Data 290건 기준)	1.3M	641K
Web상에서의 DOM API 이용속도 Test(DB Data 1140건 기준)	128초	40초
Web상에서의 DOM API 이용화일 크기 Test(DB Data 1140건 기준)	4M	2M
		3M

## Test 결과

### Test 결과(SegmentGroup별로 반복)



- SegmentGroup별로 반복되는 데이터의 횟수만큼 반복이 되어지므로 반복되는 정보의 양은 적어짐.
- SegmentGroup별로 반복되는 부분이 많아지므로 Program에 부하를 주게됨.
- 전체적인 문서의 양은 Schema1보다는 적어짐.

## Test 결과

### Test 결과(SegmentGroup별로 반복)

기본 Schema의 Template 문서	Schema1의 Template 문서	Schema2의 Template 문서
Web상에서의 DOM API 이용속도 Test(DB Data 280건 기준)	30초	15초
Web상에서의 DOM API 이용화일 크기 Test(DB Data 280건 기준)	985KByte	492KByte
Web상에서의 DOM API 이용속도 Test(DB Data 1140건 기준)	120초	52초
Web상에서의 DOM API 이용화일 크기 Test(DB Data 1140건 기준)	4,022KByte	2,009KByte
		657KByte
		62초
		2,681KByte

( 각 Segment 그룹별로 데이터를 구성하여 각 Segment Group별로 노드를 반복 )

## Test 결과

### Test용 XML 문서 생성 속도에 영향을 미치는 요인(총 5가지)

- Servlet Engine에 Java의 Option
  - 최대 Thread크기를 200으로 증가시켜서 성능을 향상
  - Heap Memory의 size를 130Mbyte로 증가시켜서 성능 향상
- DOM API를 이용한 Program의 방식
  - DOM API를 이용하는 방법에 따라 성능이 틀림.
- 노드의 개수에 따라서 XML문서를 생성하고 해석하는 속도가 차이가 발생한다.
- Segment Group별도의 반복을 통한 것은 실제 XML문서를 생성하는 Program의 복잡도가 증가하고 Module의 수행시간이 증가함에 따라 속도향상에 영향을 주지 못함.
- 실제 Test는 Web상으로 이루어졌기 때문에 Java Application으로 실행할 경우 20%정도의 성능 향상을 보일수 있음.

Samsung SDS

21

## Test 결과

### 결론(기술적인 관점)

- 본 Test의 시간수치는 프로그램 실행후 가시적으로 측정할 수치로서 절대적인 수치가 될수 없으며 단지 각 Template 문서간의 비교로서 활용되어 갈수 있음.
- 문서 전체 반복과 Segment Group별 반복일 경우에 성능 비교
  - 문서 전체 반복 : 한번에 하나의 NodeList를 Copy
  - 문서 Segment Group별 반복 : 한개의 문서 생성시 SegmentGroup별로 Program을 수행
  - 두가지 방식이 비슷한 성능을 보임.
- Node의 개수가 많은 것부터 Node의 수치가 적은 것을 비교하여 보면,
  - 소량의 데이터를 XML문서로 생성시 생성자이는 거의 미미함.
  - 대량의 데이터, 즉 데이터가 많아질수록 성능차이가 발생함.
- 문서의 Size만큼의 양이 Memory에 존재하는 것임, 그러므로 문서의 크기가 클경우 성능이 저하됨.

Samsung SDS

22

## Test 결과

### 결론(업무적인 관점)

- Schema 2에서는 문서를 전체 반복하여 데이터를 전송하게 되어 컴터업체 및 원자 업체에서 문서의 생성과 해석이 귀찮음.
- Schema 2에서는 문서를 SegmentGroup별로 반복하여야 하므로 합리업체 및 원자업체 업무의 문서를 생성과 해석이 어려워짐.
- Schema 3은 기존의 EDIFACT의 구조를 그대로 따르고 있으며 그 문서의 복잡도가 증가함.
- Schema 2은 기존 EDIFACT에서 필요로 하는 모든 필드를 가져가며 그 필드를 이용하여 EDI문서를 생성할수 있음.
- XML/EDI System에서 가장 중요한 것은 원자업체와 합리업체간의 신속하고 정확하게 정보를 공유하기 위한 것으로 EDI문서의 구조를 따라갈 필요가 있음 것으로 판단되어 지며, 다만, 실제 기존 EDI문서를 사용하려면 EDI문서로 변환할수 있는 방법만 존재하고 기존 EDI문서를 받는 측에서 이상없이 문서를 받거나 우리측에서 해당할 수있으면 되더라도 생각이 됨.
- Schema2와 Schema3의 속도를 비교하여 보면 거의 차이가 미미하나 데이터량의 증가시 약간의 차이를 보임. 이 차이는 데이터량이 증가시 더욱 큰 차이로 벌어질수 있음.

Samsung SDS

23