

초대형 워크플로우 엔진의 로깅 메커니즘

안형진, 박민재, 김광훈
경기대학교 전자계산학과
e-mail : ctrl_workflow@kyonggi.ac.kr

Logging Mechanism of Very Large Scale Workflow Engine

Hyung-Jin Ahn, Mean-Jae Park, Kwang-Hoon Kim
*Dept. of Computer Science, Kyonggi University

요 약

워크플로우 시스템은 비즈니스 환경에서 프로세스의 자동화 수행을 통해 업무 처리의 효율성 및 성능을 극대화시켜주는 미들웨어 시스템이며 워크플로우 엔진은 이러한 비즈니스 서비스의 실질적인 수행을 컨트롤 및 관리해주는 역할을 한다. 워크플로우 클라이언트로부터의 서비스 요청에 대한 처리를 위해 워크플로우 엔진은 엔진 내부의 핵심 컴포넌트들의 연동에 의해 생성되는 서비스 인스턴스들의 처리 행위를 통해 서비스를 수행하며, 서비스 처리를 하면서 발생하는 이벤트들에 대해서 로그를 기록한다. 이러한 로그 데이터들은 워크플로우 모니터링 분석에 중요한 근거 자료로서 사용되며, 워크플로우 웨어하우징 및 마이닝등의 분야에서 주요 근거 데이터로서 사용될 수 있다. 본 논문에서는 자체 제작된 e-chautauqua 초대형 워크플로우 시스템을 배경으로 초대형 워크플로우라는 환경에서 대용량의 로그를 어떻게 구성하는지에 대해서도 살펴볼 것이며, 워크플로우 엔진을 구성하는 핵심 컴포넌트들의 연동에 의해 수행되는 서비스 인스턴스들의 이벤트들이 어떠한 모습으로 로그 메시지를 구성하게 되는지에 대한 로그 메시지 포맷에 대한 전반적인 워크플로우 로깅 메커니즘에 대해 기술하고자 한다.

1. 서론

워크플로우 엔진은 비즈니스 프로세스 분야에서 사용되는 미들웨어 시스템이다. 비즈니스 환경을 구성하는 프로세스, 사용자 및 애플리케이션들 간에 발생하는 서비스들을 연계시켜주는 것이 워크플로우 엔진의 역할이다. 워크플로우 시스템은 모델러, 모니터링, 런타임 클라이언트등의 애플리케이션들이 워크플로우 엔진이라는 비즈니스 미들웨어의 연동에 의해 패키지 형태로 구성되어 비즈니스 환경에서 일어나는 서비스들을 처리하게 되는 거대 시스템이라 할 수 있다. 워크플로우 엔진은 컴포넌트 기반의 시스템으로서 개개의 컴포넌트들에서 발생하는 이벤트들에 대해서 그에 상응하는 상태들과 컴포넌트들이 위치한 호스트에서의 장애 발생시 해당 결합 내용이 로그로서 기록된다. 로그의 형태는 파일 또는 데이터베이스, XML 등 여러 모습을 가질 수 있는데, 이러한 로그를 통해 워크플로우 시스템 영역에서 발생하는 모든 이벤트들에 대한

상태들의 히스토리를 알 수 있으며, 이와 같은 정보들은 모니터링 및 워크플로우 마이닝 및 마이닝의 근거가 되는 워크플로우 웨어하우스 구축을 위한 중요한 데이터로서 활용된다.

본 논문에서 기술되는 워크플로우 엔진은 거대량의 프로세스 인스턴스 처리에 초점을 맞추어 자체 개발된 e-chautauqua 워크플로우 엔진을 말한다. 본 논문에서는 e-chautauqua 초대형 워크플로우 엔진이 어떠한 메커니즘을 통해 컴포넌트들의 상태 관련 히스토리 데이터를 기록하는지에 대한 워크플로우 로깅 메커니즘을 설명한다. 먼저 초대형 워크플로우라는 환경에 적합하게 대용량의 로그를 어떻게 구성하는지에 대해서도 살펴볼 것이다. 그리고 워크플로우 엔진을 구성하는 핵심 컴포넌트들의 연동에 의해 수행되는 서비스 인스턴스들의 이벤트들이 어떠한 모습으로 로그 메시지를 구성하게 되는지에 대한 로그 메시지 포맷에 대해 기술하고자 한다.

2. 초대형 워크플로우 엔진

초대형 워크플로우는 컴퓨팅 환경과 기업 환경의 발전으로 인하여 발생하는 대량의 워크플로우 인스턴스를 관리할 수 있는 능력을 가진 워크플로우를 말한다. 초대형 워크플로우는 워크플로우 시스템을 사용하는 조직이 점차 거대화되어 조직에서 사용되는 워크플로우 프로세스의 수가 많아지고, 그에 따른 작업의 수가 대량으로 일어남에 따라 발생하는 대량의 워크플로우 인스턴스를 처리하는 것에 목적을 둔 워크플로우이다. 본 연구에서 사용하고 있는 초대형 워크플로우 시스템은 e-chautauqua 라는 이름의 워크플로우 시스템으로서 대량의 워크케이스에 대한 수행을 처리할 수 있도록 자체 제작된 워크플로우 시스템이다. 아래의 그림 1 은 e-chautauqua 워크플로우 시스템의 아키텍처 구성도를 나타내고 있다.

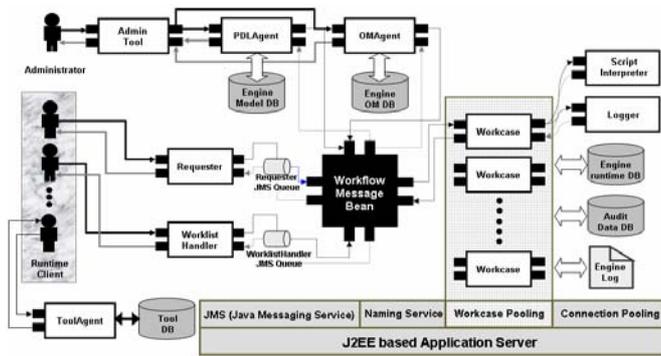


그림 1. e-chautauqua 워크플로우 시스템 아키텍처

e-chautauqua 워크플로우 엔진은 J2EE 의 EJB 컴포넌트 구조를 기반으로 개발되었다. 엔진을 구성하는 핵심 컴포넌트들은 Requester, Worklist Handler, Workcase 이다. Requester 는 모니터링 또는 런타임 클라이언트로부터 작업 요청을 받아들여 워크케이스를 생성하고 시작하며 워크케이스 및 액티비티 인스턴스등과 관련된 서비스에 대한 결과를 제공해주는 역할을 하는 컴포넌트이다. Worklist Handler 는 비즈니스 업무를 수행할 사용자에게 할당되는 단위 업무인 워크아이템과 관련된 서비스를 처리해주는 컴포넌트이며, Workcase 는 프로세스 인스턴스인 워크케이스와 대응되는 컴포넌트로서 실질적인 프로세스 업무 처리의 주체가 되는 컴포넌트이다. e-chautauqua 워크플로우 엔진은 외부로부터의 서비스 요청에 대해서 핵심 구성 컴포넌트들의 상호작용을 통해 그에 대한 서비스 처리를 하게 된다. 엔진 컴포넌트들이 작업 수행을 하면서 발생하는 이벤트들에 대해서는 로깅을 통해 해당 기록을 남기게 되는데, 이러한 로그 자료는 워크플로우 모니터링과 워크플로우 마이닝을 위한 워크플로우 웨어하우스의 구축을 하는데에 중요한 근간 자료로서 활용된다. 또한 엔진의 로그를 통해 워크플로우 시스템에 장애가 발생할 시 복구를 위한 근거 자료로서 활용되어진다. 이와 같이 워크플로우의 로깅은 모니터링, 분석, 관리 측면 뿐만 아니라 시스템 디버그등의 다양한 작업 수행을 위한 참조 데이터로서 사용될 수 있다.

3. e-chautauqua 워크플로우 엔진 로깅 메커니즘

3.1. 로깅 환경 설정

워크플로우 로그는 앞서 설명하였듯이 분석 및 장애 복구 데이터로서 사용되는 중요한 자료이기 때문에 체계적으로 기록 및 관리가 되어야만 한다. 무엇보다도 로그의 가장 중요한 역할은 히스토리의 기능이라고 할 수 있다. 엔진을 구성하는 컴포넌트 중 어떤 컴포넌트를 통해 수행된 서비스에서 특정 상태 변화를 일으켰는지, 또는 컴포넌트들간의 상호작용에 의해 생성되어 실행되는 인스턴스들 중 특정 인스턴스에서 어떠한 상태 변화를 보였는지등의 서비스에 대한 변화의 기록들이 자세히 관리되어야 한다. 그 이유는 워크플로우 모니터링 및 마이닝은 워크플로우 환경의 영역에서 비즈니스 부가 가치 창출의 결과를 돌려주기 때문이다. 예를 들면, 워크플로우 로그를 이용하여 워크플로우 모니터링을 통해 현재 수행되고 있는 워크케이스들에 대한 정보들을 파악하여 어떤 프로세스가 업무 처리를 위해 자주 사용되는지를 알 수 있고, 어떤 사용자가 작업 처리를 많이 수행하는지에 대한 성과도를 분석해낼 수 있는 등의 비즈니스를 운영해나감에 있어 이윤 획득을 높이기 위해 어떠한 점을 개선해야 하는지에 대한 내용을 빠른 시간에 정확히 분석해낼 수 있다.

e-chautauqua 워크플로우 시스템은 초대형 워크플로우라는 환경의 특성 상 대량의 작업을 처리하기 때문에 그에 따른 빈번한 상태 변화가 발생하게 되며, 엔진의 각 구성 컴포넌트들 간에 서비스 처리에 의해 발생하는 이벤트들에 대한 로그의 양은 대량으로 기록되게 된다. 따라서, 대량의 로그 기록을 효율적으로 관리하기 위한 로깅 환경 설정을 로그 파일이 일정 용량 이상이 저장되어질 경우 다른 로그 파일을 생성하여 그곳에 연이은 로깅이 가능하도록 하는 것이 효과적이다. 그리고 로깅 우선 순위를 적용하여 시스템 레벨에 해당하는 이벤트일 경우엔 시스템 로그로서 저장하고, 엔진 컴포넌트의 서비스 레벨에 해당할 경우는 서비스 로그로서 저장하여 분석 자료로서 사용하도록 한다.

3.2. e-chautauqua 엔진 로깅 메커니즘

e-chautauqua 워크플로우 시스템의 로그는 워크플로우를 구성하는 인터페이스들 간의 서비스 수행에 의해 발생된 이벤트에 대한 기록들이다. 워크플로우 클라이언트 측으로부터 서비스 요청이 들어오게 되면 e-chautauqua 워크플로우 엔진은 그에 해당하는 서비스의 결과를 돌려주기 위해 구성 컴포넌트들 간의 연동 작업을 엔진 내부적으로 실행하게 된다. 워크플로우 클라이언트의 요청에 의해 발생하는 인스턴스들의 서비스 특성에 의해 로깅 내용이 분류되어지게 되는데, 해당 서비스 인스턴스들은 워크케이스, 액티비티 인스턴스, 워크아이템이다. 워크케이스는 워크플로우 모델러에 의해서 정의되어진 비즈니스 프로세스 정의가 엔진의 가용 데이터로서 변환되어 사용되는 실질적인 프로세스 인스턴스 데이터이며, 액티비티 인스턴스는 워크케이스가 생성되어 엔진에 의해 실질적인 업무가

시작되면서 발생하는 단위 업무들의 인스턴스이고, 워크아이템은 액티비티 인스턴스 중 클라이언트 측의 사용자에게 할당되는 실제 비즈니스 업무 인스턴스이다. 워크플로우 로그 메시지 포맷은 워크플로우 표준 단체인 WfMC(Workflow Management Coalition)의 워킹 그룹인 인터페이스 5 에서 표준으로 제시하고 있는 워크플로우 감사 데이터(Workflow Audit Data)의 데이터 포맷을 기반으로 제작되었으며, 워크플로우 서비스 인스턴스들이 자신의 특성화된 서비스와 관련된 이벤트 내용들을 체계화된 로그 포맷을 통해 기록하게 된다. 워크케이스와 관련된 이벤트에 대한 로그 메시지 포맷은 아래의 표 1 과 같다.

Log elements	Description
WorkcaseLog	워크케이스 로그 최상위 Element
WorkcaseID	워크케이스 ID
ParentWorkcaseID	현재 워크케이스가 서브-프로세스 인스턴스일 경우 부모 워크케이스의 ID 가 기록됨
WorkcaseName	워크케이스 이름
State	워크케이스의 현재 상태값
PackageID	워크케이스의 모체인 워크플로우 프로세스 정의가 속한 패키지 ID
WorkflowID	워크플로우 프로세스 정의 ID
EventCode	워크케이스에서 발생한 이벤트에 대한 코드 메시지를 기록
EventTimestamp	이벤트가 발생한 시간을 기록
CreatedTimestamp	워크케이스가 생성된 시간을 기록
StartTimestamp	워크케이스가 시작 구동된 시각을 기록

표 1. 워크케이스 로그 메시지 포맷

위에 나타난 바와 같이 워크케이스와 관련된 이벤트의 내용들은 표 1 의 Element 들에 value 를 할당하여 로그를 기록하게 된다. 이 때, 워크케이스 관련 이벤트들은 특정 이벤트에 대해서 다음의 표 2 에서 기술하는 코드 메시지들 중 하나를 나타내게 된다.

Event codes	Description
WMCreatedWorkcase	워크케이스 생성 시 기록되는 이벤트 코드
WMStartedWorkcase	워크케이스 시작 시 기록되는 이벤트 코드
WMChangedWorkcaseState	워크케이스 상태 변경 시 기록 되는 이벤트 코드
WMCompletedWorkcase	워크케이스 정상 종료 시 기록 되는 이벤트 코드
WMTerminatedWorkcase	워크케이스 강제 종료 시 기록 되는 이벤트 코드
WMAbortedWorkcase	워크케이스 예외상황 발생 시 기록되는 이벤트 코드

표 2. 워크케이스 관련 이벤트 코드 메시지

다음의 표 3 은 액티비티 인스턴스가 실행될 시 발생하는 이벤트에 대한 로그 메시지 포맷을 나타낸다.

Log elements	Description
RTActivityLog	액티비티 인스턴스 로그 최상위 Element
WorkcaseID	액티비티 인스턴스가 속해 있는 워크케이스의 ID
State	액티비티 인스턴스의 현재 상태값
PackageID	액티비티 정의가 속한 패키지 ID
WorkflowID	액티비티 정의가 속한 워크플로우 프로세스 정의 ID
ActivityID	액티비티 정의 ID
ActivityName	액티비티의 이름
ActivityType	액티비티 타입
EventCode	액티비티 인스턴스에서 발생한 이벤트에 대한 코드 메시지를 기록
EventTimestamp	액티비티 인스턴스의 이벤트 발생 시간을 기록

표 3. 액티비티 인스턴스 로그 메시지 포맷

워크케이스의 경우와 마찬가지로 액티비티 인스턴스 수행과 관련된 이벤트의 내용들은 표 3 의 Element 들에 대한 value 할당으로 기록된다. 아래의 표 4 는 액티비티 인스턴스의 이벤트들에 대한 코드 메시지 목록을 나타낸다.

Event codes	Description
WMChanged-ActivityInstanceState	액티비티 인스턴스의 상태 변경 시 기록되는 이벤트 코드
WMCompleted-ActivityInstance	액티비티 인스턴스의 정상 종료 시 기록되는 이벤트 코드
WMTerminated-ActivityInstance	액티비티 인스턴스가 특정 사용자에 의해 강제 종료될 시 기록되는 이벤트 코드
WMAborted-ActivityInstance	액티비티 인스턴스가 예외 상황에 의해 강제 종료될 시 기록되는 이벤트 코드

표 4. 액티비티 인스턴스 관련 이벤트 코드 메시지

다음의 표 5 는 워크아이템과 관련된 이벤트에 대한 로그 메시지 포맷을 나타낸다.

Log elements	Description
WorkitemLog	워크아이템 로그 최상위 Element
WorkitemID	워크아이템 ID
PackageID	워크아이템의 모체가 되는 액티비티 정의가 속한 패키지 ID
WorkflowID	워크아이템의 모체가 되는 액티비티 정의가 속한 워크플로우 프로세스 정의 ID
ActivityID	워크아이템의 모체인 액티비티 정의 ID
ActivityName	액티비티 이름
EventCode	워크아이템에서 발생한 이벤트에 대한

	코드 메시지를 기록
EventTimestamp	워크아이템에서 이벤트가 발생한 시간을 기록
Performer	현재의 워크아이템을 수행하는 수행자의 ID를 기록
State	워크아이템의 현재 상태값 기록

표 5. 워크아이템 로그 메시지 포맷

위에서 설명된 인스턴스들과 유사한 방식으로 워크아이템에 대한 이벤트가 발생할 시 기록되는 코드 메시지는 다음의 표 6 과 같다.

Event codes	Description
WMAssignedWorkitem	워크아이템을 적정 수행자에게 할당 시 기록되는 이벤트 코드
WMGetWorkitem	사용자가 자신에게 할당된 워크아이템을 얻고자할 때 발생하는 이벤트 코드
WMChangedWorkitemState	워크아이템에 대한 수행 상태가 변경될 시 기록되는 이벤트 코드
WMCompletedWorkitem	워크아이템 작업이 정상적으로 종료될 시 기록되는 이벤트 코드

표 6. 워크아이템 관련 이벤트 코드 메시지

위에서 설명된 세 가지의 워크플로우 서비스 인스턴스들은 상태 변화에 의한 이벤트가 발생할 때마다 다음의 목록에 나오는 상태값들 중 하나의 값을 로그의 State 라는 Element 에 기록하게 된다.

States	Description
INACTIVE	인스턴스가 생성되었으나 구동은 되지 않은 상태임을 나타냄
ACTIVE	인스턴스가 현재 진행중인 상태임을 나타냄
SUSPENDED	인스턴스가 현재 특정 사용자에게 의해 일시 중지됨을 나타냄
COMPLETED	인스턴스가 정상적으로 완료되었음을 나타냄
TERMINATED	인스턴스가 특정 사용자에게 의해 강제 종료되었음을 나타냄
ABORTED	인스턴스가 시스템 또는 예외 상황에 의해 강제 종료되었음을 나타냄

표 7. 워크플로우 서비스 인스턴스들의 상태값 리스트

e-chautauqua 엔진은 클라이언트 측으로부터 임의의 서비스를 요청받게 되면 그에 해당하는 작업의 수행 주체가 되는 인스턴스를 생성하여 처리하게 되며, 이로부터 발생하는 이벤트에 대한 내용들을 위에서 설명한 로그 메시지 포맷 룰을 이용하여 기록하게 된다. 이 때, 로그 메시지가 기록되는 파일의 형태는 XML 을 이용하는데, 그 이유는 로그 데이터가 워크플로우

모니터링 분석 및 워크플로우 마이닝의 자료로서 사용될 경우, 해당 툴을 통하여 정보를 분석할 시에 XML Parser 를 이용한 파싱을 통해 정보 추출이 쉽다는 장점이 있기 때문이다. 다음의 그림 5 는 e-chautauqua 엔진이 서비스를 처리하면서 기록한 로그 파일의 모습을 보여주고 있다.

```

<WorkcaseLog>
  <WorkcaseID>6561531028633688638</WorkcaseID>
  <ParentWorkcaseID>null</ParentWorkcaseID>
  <WorkcaseName>Performance Test Workcase</WorkcaseName>
  <State>INITIATED</State>
  <PackageID>20041219494658234001</PackageID>
  <WorkflowID>20041219068543911001</WorkflowID>
  <EventCode>WMCreatedWorkcase</EventCode>
  <EventTimestamp>2005-01-27 21:28:30,609</EventTimestamp>
  <CreatedTimestamp>2005-01-27 21:28:30,609</CreatedTimestamp>
  <StartTimestamp>null</StartTimestamp>
</WorkcaseLog>
<RTActivityLog>
  <WorkcaseID>8517011885661850361</WorkcaseID>
  <State>COMPLETED</State>
  <PackageID>20041219494658234001</PackageID>
  <WorkflowID>20041219068543911001</WorkflowID>
  <ActivityID>2004121901034125001</ActivityID>
  <ActivityName>Start Activity</ActivityName>
  <ActivityType>START</ActivityType>
  <EventCode>WMCompletedActivityInstance</EventCode>
  <EventTimestamp>2005-01-27 21:29:10,953</EventTimestamp>
</RTActivityLog>
<WorkitemLog>
  <WorkitemID>381846916557838007</WorkitemID>
  <WorkcaseID>1147826680670526430</WorkcaseID>
  <PackageID>20041219494658234001</PackageID>
  <WorkflowID>20041219068543911001</WorkflowID>
  <ActivityID>2004121901034125002</ActivityID>
  <ActivityName>Draft Making</ActivityName>
  <EventCode>WMAssignedWorkitem</EventCode>
  <EventTimestamp>2005-01-27 21:43:22,14</EventTimestamp>
  <Performer>ahnjh</Performer>
  <State>INACTIVE</State>
</WorkitemLog>
    
```

그림 2. e-chautauqua 엔진 로그 데이터 예제

4. 결론

지금까지 본 논문에서는 워크플로우 환경에서의 서비스 요청에 의한 처리 시 발생하는 이벤트들을 어떠한 메커니즘을 통해 로그 파일을 구성하는지 살펴보았다. 워크플로우 엔진이 클라이언트로부터 요청된 서비스를 처리하기 위해 발생하는 핵심 서비스 인스턴스들인 워크케이스, 액티비티 인스턴스, 워크아이템은 각각 인스턴스들에 관련된 이벤트들에 대해서 적정 로그 메시지 포맷을 이용해 기록되며, 이러한 로그 메시지 포맷은 XML 스키마의 룰에 맞게 엔진에 의해 기록된다. 이렇게 기록된 로그 데이터는 워크플로우 모니터링 및 마이닝 등의 근간 자료로서 중요하게 활용되어 워크플로우 영역의 비즈니스 부가 가치 창출에 기여할 수 있게 될 것이다.

참고문헌

- [1] WfMC(Workflow Management Coalition), "Interface 5 - Audit Data Specification"
- [2] Kim, K, "Architecture for very large scale workflow management systems"
- [3] 심성수, 김광훈, 백수기. "초대형 워크플로우 관리 시스템 구현 및 성능평가",
- [4] 안형진, 김광훈. "초대형 워크플로우 시스템에서의 워크케이스 생성 기법에 대한 구현 분석"

Acknowledgement

본 연구는 정보통신연구진흥원 정보통신 기초기술 연구지원사업(04-기초-0005)의 지원으로 수행되었음.