

# OSICC 시스템을 위한 워크플로우 관리자

한영주\*, 윤찬현\*\*, 심은보\*\*\*, 노경태\*\*\*\*, 남기엽\*\*\*\*\*

\*한국과학기술원 정보통신공학과

\*\*한국과학기술원 전기및전자공학과

\*\*\*강원대학교 기계의용공학과

\*\*\*\*연세대학교 생명공학과

\*\*\*\*\*분자설계연구소

e-mail : { y.han, chyoun}@kaist.ac.kr, ebshim@kangwon.ac.kr, ktno@yonsei.ac.kr, kyn@bmdrc.org

## Workflow Manager for OSICC System

Youngjoo Han\*, Chan-Hyun Youn\*\*, Eun Bo Shim\*\*\*, Kyoung Tai No\*\*\*\*, Ky-Youb Nam\*\*\*\*\*

\*Dept. of Information and Communication Engineering, KAIST

\*\*Dept. of Electrical Engineering, KAIST

\*\*\* Dept. of Mechanical & Biomedical Engineering, Kangwon National University

\*\*\*\* Dept. of Biotechnology, Yonsei University

\*\*\*\*\* Bioinformatics & Molecular Design Research Center

### 요 약

IT 기술과 바이오 기술의 융합을 통해 경제적, 기술적 측면에서 효과적인 연구 및 개발을 위한 노력들이 증가하고 있다. 이를 위해서는 계산 집약적인 다양한 응용 프로그램들이 서로 유기적으로 통합되어 결과를 도출하도록 관리할 수 있어야 한다. 본 논문에서는 인체 생리기능 단위들을 모사한 신약개발 시뮬레이션을 지원하는 OSICC (e-Organ Simulator-Integrated Cyber Computing) 시스템을 위한 워크플로우 관리자를 제안한다. 제안하는 시스템은 복잡한 약물독성 시뮬레이션을 위해 워크플로우를 이용해 시뮬레이션의 구성과 조합을 효율화하였으며, 프로토타입으로 구현되어 효율적인 시뮬레이션 환경 제공이 가능함을 확인하였다.

### 1. 서론

IT 기술의 발전을 기반으로 바이오 기술의 결합을 통해 경제적, 기술적 측면에서 효과적인 연구 및 개발을 위한 노력들이 증가하고 있다. 특히, 신약개발 분야에서 대용량 컴퓨팅 기술을 이용한 시뮬레이션의 활용은 신약개발 기간을 단축할 수 있으며 신약개발 비용의 획기적 절감에 이바지할 수 있을 것으로 전망되고 있다[1].

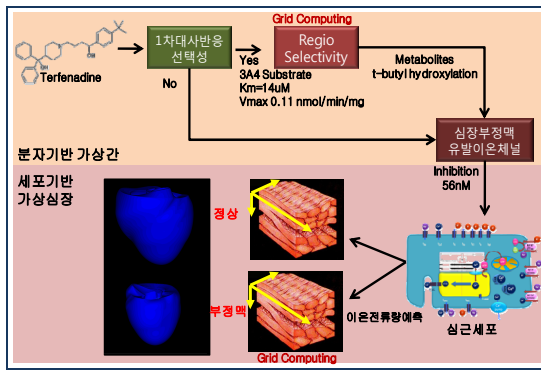
일반적으로 신약개발을 위한 약물독성 예측 시뮬레이션은 세포-조직-장기(e-Organ)에 이르는 인체 생리기능 단위들을 모사한 시뮬레이션을 기반으로 약물 스크리닝 및 가상 임상 시험의 과정을 거친다. 약물독성 예측을 위한 시뮬레이션은 크게 2 가지 특성을 가진다. 첫째, 약물독성 시뮬레이션은 여러 프로그램의 조합으로 구성된다. 따라서, 이러한 시뮬레이션이 효율적으로 수행되어 정확한 결과를 생성하기 위해서는 다양한 응용 프로그램들이 서로 유기적으로 통합되어 결과를 도출해야 한다. 즉, 응용들의 선후관계를 관리하면서 수행을 조절하는 자율적이고 지능적인 워크플로우 관리기술이 필요하다. 둘째, 시뮬레이션을 구성하는 각 개별 컴포넌트들은 계산 집약적인 특성을 가진다. 많은 컴퓨팅 파워를 이용하며 비교적 긴 실행 시간을 요구한다. 현재 널리 사용되고 있는 비즈니스

워크플로우는 복잡한 과학적 문제를 풀기 위한 계산 집약적이고 데이터 분석에 초점을 맞춘 문제를 해결하기에 적합하지 않다[2].

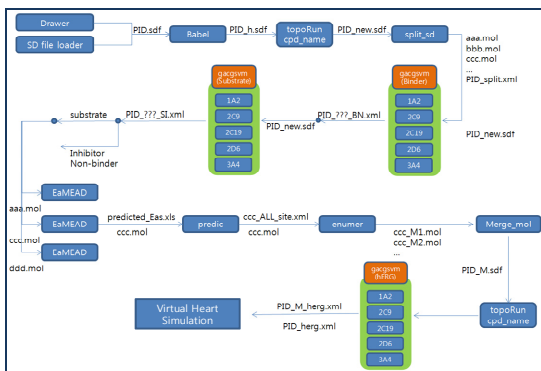
이에 본 논문에서는 세포-조직-장기에 이르는 인체 생리기능 단위들을 모사한 신약개발 시뮬레이션을 지원하는 OSICC 시스템을 위해 상기에 제시한 두 가지 관점에서 효율적인 관리를 제공하는 워크플로우 관리 시스템을 제안한다. 본 제안하는 시스템은 복잡한 약물독성 예측 시뮬레이션을 위해 워크플로우 관리자를 이용해 시뮬레이션의 구성과 조합을 효율화하였으며, 작업의 처리를 위해 분산된 자원을 효율적으로 이용하도록 하였다. 제안한 시스템은 프로토타입으로 개발되어 실제 테스트베드에서 신약개발 시뮬레이션을 위한 OSICC 시스템에 사용되고 있으며 효율적인 시뮬레이션 환경 제공이 가능함을 확인하였다.

### 2. OSICC 시스템 모델

OSICC 시스템은 세포-조직-장기에 이르는 인체 생리기능 단위들을 모사한 시뮬레이션을 기반으로 약물 스크리닝 및 가상 임상 시험이 가능하도록 하는 시스템이다. 본 시스템은 연구기관에서 신약개발에 활용될 수 있다. 새로운 약물의 개발 시에 컴퓨터 시뮬레이션을 통하여 약물독성을 예측하여 부작용 유무를



(a) 가상간 및 가상심장 기반 e-Organ 시뮬레이션



(b) e-Organ 시뮬레이션의 워크플로우

(그림 1) 약물독성 예측을 위한 e-Organ 시스템

검증할 수 있으며 약물의 부작용이 발생하는 원인에 대한 분석이 가능하므로 약물의 성질 개선 방향에 대한 정보를 제공할 수 있다. 또한, 의료 기관에서 환자별 맞춤형 처방에 활용될 수 있다. 예를 들어, 선천적으로 또는 후천적으로 간에서의 대사 효소, 심장의 이온 채널 등에 이상이 있는 환자에 대하여 처방 약품이 야기할 수 있는 부작용을 예측하여 약물로 인한 부작용을 줄일 수 있다.

상기에 제시된 목적으로 사용될 수 있는 e-Organ의 독성예측 시뮬레이션은 기본적으로 3 단계로 구성되어 있다. 1 단계는 인간이 약물을 복용하였을 때, 간에서의 대사 반응을 중심으로 약물이 어떻게 변하는지에 대하여 가상간에서의 대사반응에 대한 시뮬레이션이다. 2 단계는 약물과 대사반응에 의해 생성된 대사체가 심장에 어떤 영향을 줄 수 있는지를 시뮬레이션 한다. 마지막 3 단계에서는 최종적으로 심장에 부작용이 어떻게 발생하는지에 대하여 가상심장 시뮬레이션[3]을 수행하여 간에서의 대사 현상을 고려한 심장 부작용 예측이 가능하도록 하는 것이다. 그림 1(a)는 약물독성 예측을 위한 OSICC 시스템의 가상간과 가상심장에 대해 나타낸 것이다. 또한 1(b)는 3 단계로 구성된 OSICC 시스템의 수행을 위한 워크플로우 구성을 나타낸 것이다.

약물독성 시뮬레이션을 수행하기 위해 사용자는 먼저 작업실행을 위한 워크플로우를 디자인 및 생성하여야 한다. 워크플로우의 디자인은 정의된 워크플로우 기술언어(Workflow Description Language)를 이용하

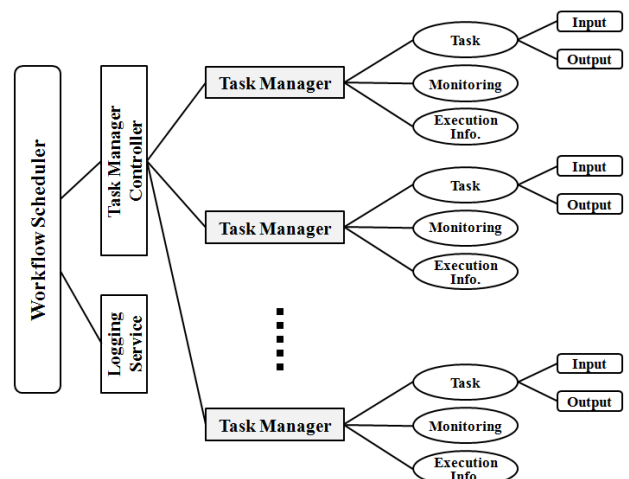
여 작업의 순서 및 작업들의 연결관계를 정하고 작업들 사이에 흐르는 데이터를 제대로 주고 받을 수 있도록 데이터 타입을 정의하는 것들을 의미한다. 본 OSICC 시스템에 수행 요청이 이루어진 워크플로우는 일련의 과정을 거쳐 워크플로우 관리자에게 전달된다. 복잡한 작업으로 구성된 응용을 효율적으로 관리 및 통제하기 위한 워크플로우관리자는 상위의 응용계층에서 요청한 작업을 인지하고 이를 효율적으로 처리한다. 워크플로우 관리자에 대해서는 다음 장에 자세히 논의하도록 한다.

### 3. 워크플로우 관리자

상기에 제시된 약물독성 예측 시뮬레이션을 효율적으로 지원할 수 있는 워크플로우 관리자는 가상간과 가상심장으로 구성되는 복잡한 약물독성 시뮬레이션을 위해 워크플로우를 이용해 시뮬레이션의 구성과 조합을 효율화하였으며, 분산된 자원을 효율적으로 이용하여 빠른 시간 안에 시뮬레이션을 처리하기 위해 기 연구 개발된 자원 관리 시스템을 사용하였다[3]. 워크플로우 관리자는 크게 워크플로우 스케줄러와 태스크 관리자로 구성된다. 각각의 기능에 대해서는 3.1과 3.2에서 자세히 논의한다.

#### 3.1. 워크플로우 스케줄러

워크플로우 스케줄러는 워크플로우의 실행을 관리하는 역할을 한다. 그림 2에 도시된 것처럼 워크플로우 스케줄러는 워크플로우의 각 태스크를 관리하기 위한 태스크 관리자를 생성하고 로깅서비스(logging service)를 통해 각 태스크의 실행 상태를 모니터링 한다. 워크플로우 스케줄러는 전체 워크플로우의 실행을 관리하기 위해 다음과 같은 일련의 과정을 거치게 된다. 먼저, 워크플로우 스케줄러는 각각의 워크플로우를 위해 로깅서비스에 워크플로우를 등록하고 워크플로우의 첫 번째 레벨 태스크를 위한 태스크 관리자를 생성한다. 태스크 관리자로부터 작업 실행 상태 정보를 받는 동안 워크플로우 스케줄러는 자식 태스크들(children tasks)을 위한 태스크 관리자를 생성하여



(그림 2) 워크플로우 스케줄러

태스크 수행을 준비한다. 그리고 이것은 전체 워크플로우에 속한 모든 태스크가 종료될 때까지 반복적으로 수행된다. 워크플로우 스케줄러가 “Done” 메시지를 받았을 경우, 다른 태스크 관리자들이 수행 중인지를 체크한다. 만약 수행 중인 태스크 관리자가 있을 경우에는 진행 중인 태스크들이 종료될 때까지 기다린 후 모두 종료되었을 때 워크플로우의 수행을 완료한다. “Fail” 메시지를 받았을 경우, 워크플로우 스케줄러는 실패한 태스크를 다시 실행하여 워크플로우 전체 작업을 종료하도록 한다.

### 3.2. 태스크 관리자 (Task Manager)

태스크 관리자는 워크플로우에서 태스크를 관리하기 위해 개발되었다. 각 태스크 관리자는 각각 자신의 모니터를 가지고 있으며 분산된 노드에서 수행 중인 태스크 실행의 상태를 수집하는 역할을 한다. 태스크 관리자가 제공하는 기능들은 아래와 같다.

- **Self-healing:** 서비스나 자원에 의해서 부적당한 수행을 탐지하고 적절하지 않은 상태가 되었을 때, 정책 기반의 수정된 액션을 수행한다.
- **Self-optimizing:** 사용자의 요구에 최적화되도록 조절한다.
- **Self-configuring:** 관리 시스템의 의해 제공되는 정책을 기반으로 분산된 시스템의 동적인 상태변화에 적응하도록 한다.

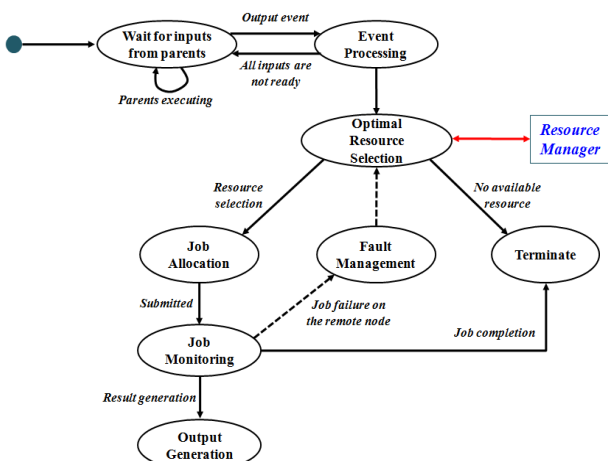
그림 3 은 태스크 관리자의 상태전이도를 나타낸 것이다. 태스크 관리자는 상위 부모 태스크에 의해서 생성된 output 이벤트들을 등록하고 output 이벤트를 기다린다. 이벤트가 발생했을 때, 태스크 관리자는 이벤트 처리 상태로 이동하고 이때, input 데이터의 정보를 기록한다. 만약 다른 어떤 input 데이터라도 여전히 가용하지 않을 때는 태스크 관리자는 대기 상태로 간다. 그렇지 않으면 자원 선정 상태로 이동한다. 자원 선정에서는 자원관리자로부터 생성된 자원 그룹 중 하나의 자원을 선정한다. 만약 선정된 자원이 사용 가능하면, 태스크 관리자는 선정된 자원에 태스크

를 할당하고 원격 자원에서 작업의 상태를 모니터링한다. 만약 태스크 실행이 실패하였을 경우, 태스크 관리자는 자원 선정 작업을 다시 수행하고 새롭게 선정된 자원에 태스크를 할당하여 자원을 재실행한다. 태스크 실행을 위한 자원이 충분하지 않을 경우에는 태스크 관리자는 종료된다.

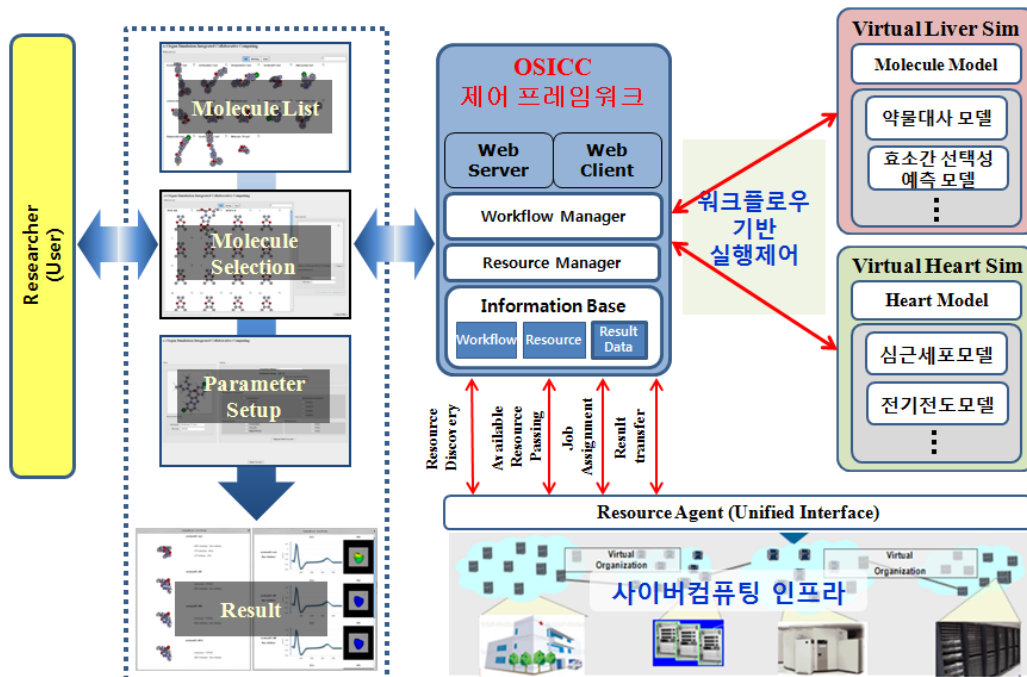
### 4. 워크플로우 관리자 기반 OSICC 시스템 구현

OSICC 시스템은 IT 비전문가인 사용자가 약물독성 예측 시뮬레이션의 과정을 디자인하고 분산된 사이버컴퓨팅 인프라를 이용하여 약물독성 시뮬레이션을 수행하고 결과를 가시화할 수 있는 통합서비스를 제공한다. 따라서, 사용자는 복잡한 시뮬레이션의 일련의 과정에 대한 컨트롤 없이 작업수행 및 분산된 자원 사용에 대한 복잡한 제어 없이 사이버컴퓨팅 자원을 이용하여 쉽게 시뮬레이션을 실행하고 결과를 유지 관리할 수 있다. 그림 4 에 도시된 것처럼 OSICC 시스템은 3 부분으로 구성된다: 사이버컴퓨팅 인프라, OSICC 제어프레임워크, 사용자 서비스 포털.

먼저, 사이버컴퓨팅 인프라는 지리적으로 분산되어 있는 이기종 자원으로 구성되어 있으며 실제 작업의 처리가 일어날 컴퓨팅 자원의 상태는 각 자원에 설치되어 있는 자원모니터링 에이전트를 통해 자원상태 정보 수집시스템에 수집 및 관리된다. 또한, OSICC 제어 프레임워크는 크게 워크플로우 관리자와 자원관리자로 구성된다. 워크플로우 관리자는 3 장에서 자세히 논의하였다. 자원관리자는 실시간 모니터링된 자원 상태를 통해 실제 작업의 처리가 일어날 물리계층의 자원 상태를 인지하고 이를 기반으로 최적의 자원을 선정하여 제공한다. 복잡한 작업으로 구성된 워크플로우를 효율적으로 관리 및 통제하기 위해 OSICC 제어 프레임워크는 상위의 응용계층에서 요청한 작업과 자원 상태 정보를 기반으로 최적의 작업 할당 정책을 생성한다. 마지막으로 사용자 서비스 포털은 사용자가 시뮬레이션을 수행하고 관리하기 위한 웹 기반 관리 틀이다. 그림 4 의 왼쪽은 사이버컴퓨팅 기반 OSICC 시스템의 사용자를 위한 서비스 포털에서의 수행단계를 나타낸 것이다. 사용자는 서비스 포털에 접속하여 자신의 계정에 접속하여 기 수행된 시뮬레이션의 약물과 결과 등을 확인할 수 있는 “Molecule List” 로 이동할 수 있다. 또한, 새로운 molecule 의 시뮬레이션을 위한 “Molecule Selection” 기능을 통해 molecule 파일을 업로드하여 시뮬레이션을 수행할 수 있다. Molecule 파일의 업로드 이후 시뮬레이션을 위한 파라미터 설정 과정을 수행한다. 시뮬레이션을 위한 예비 단계가 완료된 후에 워크플로우 관리 시스템은 하부의 컴퓨팅 자원을 이용하여 시뮬레이션을 수행한다. 시뮬레이션 수행 시 워크플로우 진행 정보는 실행상태 표시자를 통해 사용자에게 전달되며 수행이 종료된 시뮬레이션 최종결과는 웹 상에서 가시화된다.



(그림 3) 태스크 관리자의 상태전이도



(그림 4) OSICC 시스템 플랫폼 구조

## 5. 결론

본 논문에서는 약물독성 예측 시뮬레이션 서비스를 제공하는 OSICC 모델이 계산 집약적인 다양한 응용 프로그램의 조합으로 구성되어 있는 것을 확인하였다. 또한, 이를 효율적으로 지원할 수 있는 워크플로우 관리자에 대해 논의하였다. 워크플로우 스케줄러와 태스크 관리자로 구성된 워크플로우 관리자는 계산 집약적인 컴포넌트로 구성된 복잡한 시뮬레이션의 효율적인 실행과 관리 기능을 제공한다. 제안하는 워크플로우 관리자의 유효성을 평가하기 위해 약물독성 예측 시뮬레이션을 지원하는 시스템 프로토타입을 구현하여 OSICC 시스템에서 효율성을 확인하였다.

## Acknowledgement

본 연구는 지식경제부 및 한국산업기술평가관리원의 산업원천기술개발사업의 일환으로 수행하였음. [KI001826, 사이버컴퓨팅 기반 e-Organ 시스템 개발]

## 참고문헌

- [1] PricewaterhouseCoopers report ([www.pwc.com](http://www.pwc.com)), 1999.
- [2] Workflow Management Coalition. <http://www.wfmc.org/>
- [3] U. Im, S. Kwon, K. Kim, Y. Lee, Y. Park, C.-H. Youn, E. Shim, "Theoretical analysis of the magnetocardiographic pattern for reentry wave propagation in a three-dimensional human heart model," Progress in Biophysics & Molecular Biology, vol. 96, no. 1, pp. 339-56.
- [4] Youngjoo Han, Hyewon Song, Byungsang Kim, and Chan-Hyun Youn, "SLA-Constrained Policy-based Scheduling Mechanism in Grid," IEICE Trans. on Comm., Vol. E91-B, No. 12, pp. 4009-4012.