

GAIS : OGSA 기반 그리드 정보 서비스 구조

임민열, 김은성, 홍원택, 박형우
한국과학기술정보연구원 슈퍼컴퓨팅센터
e-mail : {philipus, eskim74, wthong, hwpark}@kisti.re.kr

An architecture of Grid Information Service based on OGSI

Minyeol Lim, Eunsung Kim, Wontaek Hong, Hyungwoo Park
Supercomputing center, Korea Institute of Science and Technology
Information(KISTI)

요 약

최근 그리드 미들웨어 개발 연구는 과학적 어플리케이션 뿐만 아니라 상업적인 비즈니스 어플리케이션에서 활용가능하도록 웹서비스(Web service) 기술을 활용하고 있다. 이를 통해 정의된 그리드 서비스 구조를 Open Grid Service Architecture(OGSA)라고 한다. OGSA 를 통해 정의된 그리드 서비스 개념은 GGF 를 통해 OGSI 라는 표준 기술 명세서로 채택되었다. OGSI 를 지원하는 그리드 서비스는 확장성있고 상호운용성을 높여 다른 서비스와의 연동 및 통합을 쉽게 이루게 한다.

본 논문에서는 기존의 그리드 정보 서비스를 OGSI 기반의 서비스 구조인 Grid Advanced Information Service(GAIS)로 재구성한다. GAIS 는 그리드 미들웨어의 기본 서비스로써 다음의 특징을 가진다. 우선, 기존의 GT3 기반의 그리드 정보 서비스와의 서비스 인터페이스 측면에서 호환성을 가지면서, 내부적으로 동적 Virtual Organization(VO)관리 기능을 추가로 가진다. 그리고, 여러 GAIS 들간의 VO 정보를 관리할 수 있도록 Peer-to-peer 기반의 정보 분산기능을 가진다. 또한, GAIS 내 등록된 서비스들에 대해서는 Service data 에 따른 그룹화를 통해 서비스 discovery 기능을 강화하게 된다.

그리드 사용자는 GAIS 를 통해 다양한 VO 에 대한 정보를 검색하고, VO 의 생성 및 참여할 수 있고, 자신이 속한 VO 에 대하여 효율적으로 자원정보 및 서비스 데이터(Service Data)를 접근할 수 있다.

1. 서론

우리가 항상 접하는 인터넷 기반의 컴퓨팅 환경은 정보의 공유(Information sharing)을 가능하게 한다. 이를 이용하여 전자상거래와 같은 다양한 서비스가 생성되었다. 하지만, 단순히 정보의 공유를 통해서만 실제로 우리가 할 수 있는 다양한 일들을 실현할 수 없다. 그리드 컴퓨팅[1]은 이 한계를 뛰어 넘는 것을 가능하게 한다. 즉, 앞에서 말했던 정보의 공유뿐만 아니라, 다양한 종류의 자원의 공유(Resource sharing)를 가능하게 하기 때문이다. 예를 들어, 어떤 사용자가 엄청난 컴퓨팅 자원을 필요로 한다면, 그리드 컴퓨팅환경은 전세계에 존재하는 컴퓨팅 자원을 그리드 컴퓨팅 환경을 통해 마치 자신이 슈퍼컴퓨터

를 보유한 것처럼 이용할 수 있다.

위와 같은 그리드 컴퓨팅 환경을 구축하기 위해서는 기본적으로 분산 컴퓨팅 환경에 존재하는 자원들에 대한 정보를 효율적으로 통합하여 접근할 수 있도록 지원하는 것이 필수적이다. 이것을 일반적으로 그리드 정보서비스라고 말한다[2]. 그리드 정보 서비스는 이기종의 다양하고 분산되어진 자원들에 대한 정보(Metadata)를 효율적으로 통합 관리하여, 그리드 사용자가 원하는 자원에 대한 정보를 효율적으로 서비스해야 한다. 그리드 사용자는 이를 통해 실질적이고 분산된 자원들을 자신이 직접 보유한 로컬 자원이고 또 하나의 자원인 것처럼 볼 수 있게 된다.

최근에 전세계의 그리드 관련 표준화 기구인 글로

	GT2 기반 MDS	GT3 기반 정보 서비스
Data source	Resource information	Resource information + Service data
Data format	LDIF	XML
Data query	LDAP query	XPATH, XQuery
Service architecture	GIIS + GRIS	Index service(Aggregator, Provider)

표 1. GT2.x 와 GT3.x 에서의 정보 서비스 비교

벌 그리드 포럼(Global Grid Forum, GGF)는 그리드 컴퓨팅을 실현하기 위한 기반 기술로써 웹 서비스 구조를 확장한 Open Grid Services Infrastructure(OGSI)를 발표하였다. 이는 그리드 컴퓨팅을 효율적으로 구축하기 위해서 개발자는 그리드 서비스를 기본 단위로 쉽게 개발할 수 있는 프레임워크를 제공한다. 개발된 그리드 서비스들은 단독 또는 다른 서비스와의 통합을 통해 배치되어 서비스되게 된다. 이는 웹 서비스의 확장성 및 상호운용성을 그리드 컴퓨팅 기술에 접목한 구조라 할 수 있다.

본 논문에서 우리는 기존의 그리드 정보 서비스 특징들을 OGSA 기반의 그리드 서비스 구조로 재구성한다. 이를 통해, 기존의 정보 서비스의 특징들을 발전시키면서 그리드 정보서비스의 확장성 및 활용도를 높일 수 있게 된다. 여기서 우리는 OGSA 기반의 그리드 정보 서비스를 Grid Advanced Information Service(GAIS)라고 표현한다. 기본적으로 GAIS 는 그리드 사용자에게 직접 서비스하거나 다른 그리드 서비스를 통해 활용될 수 있는 기본 서비스(Base service)로써 VO 관리, VO 내 서비스의 등록 및 관리, 그리고 자원 및 서비스 데이터를 제공할 수 있는 기능을 가지게 된다.

본 논문의 구성은 우선 2 장에서 글로벌스 툴킷 3.0 기반의 그리드 정보 서비스 구조를 살펴보고, 3 장에서는 이를 개선한 GAIS 구조를 살펴본다. 4 장에서는 GAIS 서비스 구현 및 서비스 활용 시나리오 대해 기술한다. 그리고 마지막 5 장에서 GAIS 의 향후 발전 방향을 제공하고 본 논문을 마무리한다.

2. GT3 기반의 그리드 정보 서비스

앞서 언급했듯이, OGSI 는 OGSA 를 통해 정의된 그리드 서비스 개념을 공식적이고 표준화된 형태로 기술된 테크니컬 스펙을 일컫는다. 현재, 글로벌스의 3.0 버전은 OGSI 의 명세된 내용을 따라 Java 기반으로 구현하였고, 기존의 2.x 버전에서 존재하는 서비스 개념을 그대로 재구성하였다.

특히, 글로벌스 툴킷내 여러 서비스 컴포넌트중에서 그리드 정보 서비스부문에 있어서 변화가 가장 두

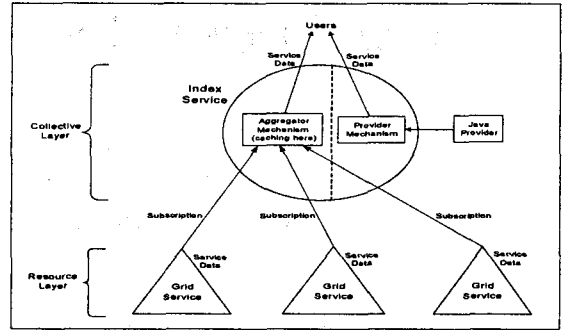


그림 1. GT3 의 인덱스 서비스(Index service) 구조

드러진다. 기존 글로벌스 툴킷 2.x 버전의 그리드 정보 서비스 컴포넌트인 Monitoring and Discovery Service(MDS)에 비해 서비스 구조가 대부분 바뀌었다고 할 수 있다.

표 1 은 기존 MDS 와의 차이점은 정리하고 있다. MDS 와 가장 큰 차이점은 제공되는 정보가 그리드 서비스들에 대한 서비스 데이터를 포함한다는 것이다. 이는 기존의 컴퓨팅 자원에 대한 정보가 전부였던 것에 비해 서비스 구조에 있어서 변화를 가져오게 된다. 이 때문에, GT3 에서 정보서비스 구조는 GT2 에서 처럼 GIIS 와 GRIS 로 떼어내기 어렵다.

GT3 의 그리드 정보 서비스에서 자원 정보는 기존의 MDS 와 유사하게 정보 제공자(Information providers)를 통해 수집되어 진다. 이는 Provider Manager 를 통해 동기식(Synchronous) 또는 비동기식(Asynchronous)으로 호출되어 진다. 한편, 그리드 서비스에 정의된 서비스 데이터는 Aggregator 를 통해 Notification mechanism 으로 수집되어 진다. 이는 특정 서비스 데이터에 대한 변화가 발생하면 이벤트가 발생하여 Aggregator 에 의해 그 값이 반영되게 된다.

그림 1 은 GT3 의 수집 계층(Collection layer) 서비스인 인덱스 서비스(Index service) 구조를 보여준다. 인덱스 서비스는 Aggregator 와 Provider 를 통해 수집된 정보들을 서비스 데이터로 유지하면서 다양한 검색 방법을 통해 자원 및 서비스 데이터에 대한 발견(Discovery) 및 접근(Lookup)기능을 제공한다.

그러나, GT3 인덱스 서비스의 한계는 Virtual Organization(VO)에 대한 생성/제거 및 참여/탈퇴와 같은 기능을 제공하지 않습니다. 그리고, 자원 정보가 여러 정보 서비스를 통해 분산되어 있을 경우 전역적으로 탐색할 수 있는 기능이 없다는 것이다. 즉, 그리드 컴퓨팅에서 가장 중요한 개념인 VO 에 대한 동적인 관리기능을 지원하는 것이 그리드 정보 서비스가 가져야 될 필수 기능이다.

3. Grid Advanced Information Service(GAIS)

결론적으로 GAIS 는 기본적으로 GT3 인덱스 서비스와 같은 기본 서비스(Base service) 기능을 제공하고,

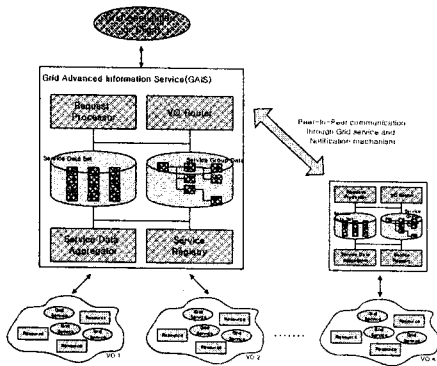


그림 2. GAIS 서비스 구조

GT3 인덱스 서비스와 인터페이스의 호환성을 유지한다. 추가로, 인덱스 서비스에서 제공되지 않는 동적 VO 관리 기능 및 전역적 정보 서비스 기능을 포함한다. 이에 추가로, 인덱스 서비스를 통해 등록된 그리드 서비스를 다양한 형태로 그룹화할 수 있도록 지원한다. 이는 그리드 사용자가 등록된 서비스를 효율적으로 검색할 수 있는 인덱싱 방법을 포함한다.

그림 2 는 GAIS 서비스 구조를 보여준다. GAIS 는 크게 요청처리기(Request processor), VO 라우터(VO router), 서비스데이터 수집기(Service Data Aggregator), 서비스 레지스트리(Service Registry) 4 가지 모듈로 나뉘어진다. 각각에 대한 설명은 다음과 같다.

- ✓ 요청 처리기 : VO 생성/제거 및 참여/탈퇴와 관련된 사용자 요청, 자원 및 서비스 데이터에 대한 검색 요청, 특정 서비스 검색 요청등을 받아 각 모듈로 분기하는 기능을 가진다.
- ✓ VO 라우터 : 자신이 관리하지 않는 다른 VO 에 대한 정보를 피어-투-피어 기반으로 검색하여 GAIS 의 서비스 핸들(Handle)을 찾는다.
- ✓ 서비스 데이터 수집기 : 서비스 데이터 수집기는 다양한 정보 제공자 또는 그리드 서비스로부터 제공되는 서비스 데이터를 수집하는 역할을 수행한다.
- ✓ 서비스 레지스트리 : GAIS 에 등록된 서비스들을 관리하는 역할을 수행한다. 그리드 사용자 또는 어플리케이션은 GAIS 를 통해 원하는 서비스를 찾을 수 있게 된다.

하나의 GAIS 는 기존의 인덱스 서비스와 달리 다중의 VO 를 동적으로 관리할 수 있는 기능을 가진다. 이는 특정 VO 에 관련된 정보를 다른 VO 정보와 분리해서 유지관리하며 그리드 사용자가 GAIS 를 통해 VO 를 동적으로 생성/소멸 및 가입/탈퇴할 수 있도록 한다. GAIS 의 또 다른 특징은 여러 GAIS 들 간에 각각 유지되는 VO 정보를 피어-투-피어 구조로 공유한다는 것이다. 이는 그리드 사용자가 특정 자원 또는 서비스를 검색할 때, VO 도메인을 넘어서 검색할 수 있는 메커니즘을 제공한다. 이를 통해, 그리드 컴퓨팅 내 존재하는 자원 및 서비스들을 전역적으로 접근할

수 있게 된다. 또한, GAIS 는 기존의 인덱스 서비스가

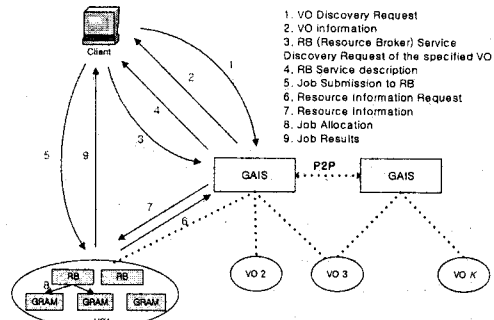


그림 3. GAIS 활용 시나리오

제공하지 않는 서비스 레지스트리 기능을 가진다. 이는 OGSi 스펙에 기술된 *ServiceGroup*, *ServiceGroupEntry*, *ServiceGroupRegistration portType* 을 지원하는 것을 의미한다. 이들을 통해, GAIS 는 등록된 서비스들을 그룹화하고 각 그룹별로 인덱싱할 수 있다. 즉, 사용자가 관심있는 서비스에 대해서 효율적으로 검색할 수 있다. 예를 들어, 자신이 속한 VO 내 에 자원 브로커링을 제공하는 그리드 서비스들이 여러 개 존재할 경우, 이들을 GAIS 내에서 서비스 그룹화하여 유지할 경우, 사용자 요청에 대해 빠르게 서비스할 수 있게 된다.

4. 구현 및 활용 시나리오

GAIS 의 구현은 우선 GT3 인덱스 서비스와의 호환성을 위해 인덱스 서비스에 정의된 *portType* 을 기반으로 확장하게 된다. 여기에 VO 관리에 필요한 *portType* 과 서비스 레지스트리 기능을 제공하기 위한 *ServiceGroup portType* 을 기본 *portType* 으로 추가하게 된다.

GAIS 들 간에 VO 정보를 공유할 수 있도록 피어-투-피어 메커니즘을 활용하기 위해서 GAIS 내부에 Java 피어-투-피어 라이브러리인 JXTA 를 활용한다. GAIS 는 다른 GAIS 로부터 자원 및 서비스에 대한 검색 요청을 받으면 자신이 관리하는 VO 중에서 검색한 뒤 그 결과를 원래 GAIS 에게 전달하게 된다.

그림 3 은 그리드 컴퓨팅내에서 GAIS 를 활용한 시나리오를 보여 준다. 그리드 사용자가 자신이 알고 있는 GAIS 를 통해 관심있는 VO 를 관리하는 GAIS 의 주소를 찾게 되고, 그 GAIS 를 통해 관심있는 자원 브로커링 서비스를 찾게 된다. 이를 통해, 자원 브로커링 서비스에 대한 접근 방법을 알게 되고, 그 자원 브로커에게 일(Job)을 던지게 된다. 자원 브로커 서비스 는 실제로 일을 수행할 자원을 선택하기 위해 GAIS 로부터 자원 정보를 활용하게 된다. 그리고, 선택된 자원에 일을 나누어 수행하고 그 결과를 사용자에게 던지게 된다.

위의 시나리오에서 보여주듯이 GAIS 는 사용자가 일을 수행하는 과정에서 필요한 모든 자원 및 서비스에 대한 정보를 원활히 활용할 수 있는 서비스 구조

를 가지게 된다.

5. 결론

현재 GAIS 는 기본적으로 GT3 인덱스 서비스와 같이 자원 및 서비스 데이터를 관리하고 서비스하는 기본 서비스(Base service)이다. 이는 GAIS 를 활용하여 다양한 하이 레벨(High-level) 서비스를 제공할 수 있다는 것을 의미한다. 예를 들어, 사용자가 접근하는 자원 또는 서비스 정보의 시간적(Temporal) 혹은 공간적(Spatial) 지역성(Locality)를 활용하여 자주 또는 가까이 존재하는 자원에 대한 자원 정보에 대해 순위(Ranking)를 매기고 이에 따라 우선적으로 정보 서비스를 제공하는 형태의 서비스가 가능하다. 또한, 정보 서비스의 접근 빈도에 따라 제공되는 정보에 대한 정확성의 레벨을 조절하여 서비스의 성능을 향상시킬 수도 있다. 이와 같은 서비스는 GAIS 가 가져야 상위 서비스로써 구현될 예정이다. 또한, GAIS 는 Grid Security Infrastructure(GSI)기반의 보안 기능을 포함할 예정이다.

GAIS 는 OGSi 스펙에 명시된 표준화된 그리드 서비스로써 그리드 컴퓨팅에 필수적인 정보 서비스를 제공한다. 이를 위해, 기존의 GT3 인덱스 서비스에서 제공하는 기능을 확장하여 동적인 VO 관리 기능을 제공하면서 전역적으로 VO 을 검색할 수 있는 기능을 가진다. 또한, 등록된 서비스들에 대한 효율적인 검색 기능을 제공하기 위해 서비스들의 그룹화 기능을 제공한다. 이와 같은 그리드 정보 서비스는 기존의 정보 서비스 기능을 한층 업그레이드한다고 볼 수 있다.

참고문헌

- [1] I. Foster, C. Kesselman, S. Tuecke, "The Anatomy of the Grid: Enabling Scalable Virtual Organizations", International J. Supercomputer Applications, 15(3), 2001.
- [2] I. Foster, C. Kesselman, J. Nick, S. Tuecke, " The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration", Open Grid Service Infrastructure WG, Global Grid Forum, June 22, 2002.
- [3] S. Tuecke, K. Czajkowski, I. Foster, J. Frey, S. Graham, C. Kesselman, T. Maguire, T. Sandholm, P. Vanderbilt, D. Snelling; "Open Grid Services Infrastructure (OGSI) Version 1.0.", Global Grid Forum Draft Recommendation, 6/27/2003
- [4] K. Czajkowski, S. Fitzgerald, I. Foster, C. Kesselman, "Grid Information Services for Distributed Resource Sharing.", Proceedings of the Tenth IEEE International Symposium on High-Performance Distributed Computing (HPDC-10), IEEE Press, August 2001.
- [5] Globus toolkit, <http://www.globus.org>
- [6] Andy Oram, "Peer-To-Peer", O'REILLY, 2001
- [7] Brendon J. Wilson, "JXTA", New Riders, 2002