

계산 및 데이터 그리드 환경을 위한 Information Providers 에 관한 연구

홍원택
한국과학기술정보연구원 슈퍼컴퓨팅센터
e-mail : wthong@kisti.re.kr

A study on the Information Providers for computational and data Grid environments

Won Taek Hong
Supercomputing Center, Korea Institute of Science and Technology Information

요 약

본 논문은 계산 및 데이터 그리드 환경에서 필요한 Information Provider 들의 제공 방법에 관한 것이다. 그리드 정보 서비스의 한 요소로 필요한 Information Provider 를 구현하기 위해서는 일반적으로 요구 사항을 분석한 후, 스키마를 정의하고 그것에 기반하여 Information Provider 를 구현하는 과정을 거친다. 본 논문에서는 Globus Toolkit 과 SRB 를 기반으로 구축된 계산 및 데이터 그리드 환경에서 MCE 스키마와 MSE 스키마를 정의하고, 그에 따른 각각의 Information Provider 들을 구현한다. 각각의 Information Provider 는 계산 자원에 관한 정보 및 MCAT 기반의 SRB 서버에서 유지되는 리플리카에 대한 정보를 제공한다. 이러한 Information Provider 들은 유사하게 구축된 그리드 환경에서 그리드 정보 서비스의 한 요소로 사용될 수 있을 것이다.

1. 서론

그리드 환경에서는 기본적으로 인터넷 환경에서 제공되는 단순한 정보 공유 외에 다양한 종류의 컴퓨팅 자원을 공유할 수 있게 된다. 따라서 그리드 환경에서 사용자는 자신이 원하는 자원 및 서비스를 제공 받기 위해 요구 조건을 만족시키는 자원 및 서비스를 선택해야 한다. 그리드 정보 서비스는 이러한 사용자의 요구 조건을 반영하는 자원 및 서비스의 발견, 선택 및 최적화에 대한 정보를 제공한다. 즉, 사용자는 그리드 정보 서비스를 통하여 그리드 환경의 운영 및 어플리케이션의 실행에 필요한 정보를 얻게 된다. 특히, 계산 및 데이터 그리드 환경에서 요구되는 많은 정보를 사용자에게 제공하는 것은 매우 중요한 요소라고 할 수 있고, 이러한 실질적인 정보는 그리드 정보 서비스의 한 요소인 Information Provider 를 통해 이루어진다.

계산 및 데이터 그리드 환경에서 제기되는 다양한 요구사항을 만족시키기 위해 범용적으로 Globus toolkit (GT)[4]와 Storage Resource Broker (SRB)[3]가 널리 사용

되고 있다. SRB 는 네트워크상의 이질적인 스토리지 자원들을 연결하고, 리플리카 데이터를 접근하는 일관된 인터페이스를 제공하는 클라이언트 서버 기반의 미들웨어이고, GT3 는 Open Grid Service Architecture(OGSA)[2] 프레임워크 하에서 일반적으로 서비스가 구현해야 하는 인터페이스들의 상세한 스펙을 묘사한 Open Grid Service Infrastructure(OGSI)[1]의 참조 구현물이다.

현재, GT3 는 그리드 정보 서비스의 기본적인 기능을 제공한다. 이러한 그리드 정보 서비스에는 다양한 자원 정보들이 유지되어 있어야 하고, 그렇게 하기 위해 정보의 다양성과 정확성을 고려한 Information Provider 가 제공되어야 한다. 본 논문에서는 계산 및 데이터 그리드 환경에서 사용자들로부터 요구되는 정보들을 분석해 보고, 그에 따라 스키마를 설계하여 필요한 Information Provider 를 구현한다.

2. 요구사항 분석

그리드 환경을 구축하기 위해서는 응용의 특성에 따라 다양한 미들웨어들과 자원들이 요구된다. 일반적인 계산 및 데이터 그리드 환경으로 그림 1 과 같이 GT3 와 SRB 가 미들웨어로 구축된 그리드 환경을 가정한다. 이러한 환경에는 계산 자원들, 스토리지 자원들, 과학 장비들이 있을 수 있다. 그리드 사용자는 일반적으로 자신의 작업을 제출하고 할 때, 다음과 같은 절차를 따른다.

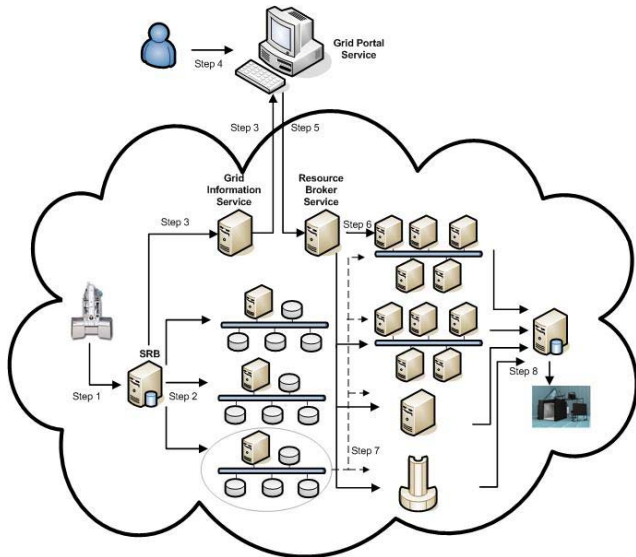


그림 1. 작업 제출을 위한 일반적인 흐름

첫째, 과학 장비에서 응용을 통해 대용량의 데이터들이 생성된다.

둘째, 이러한 데이터는 로컬에 있는 스토리지 뿐만 아니라 SRB 에 의해 관리되는 다른 외부 사이트에도 복제된다.

셋째, 이렇게 복제된 리플리카에 대한 정보는 그리드 정보 서비스로 전송된다.

넷째, 사용자는 그리드 포털 서비스나 인터페이스를 통해 그리드 정보 서비스에서 제공되는 리플리카들 중에서 최선의 것을 선택하게 된다.

다섯째, 위에서 선택된 리플리카를 이용하는 작업 스펙을 자원 브로커 서비스에게 제출하게 된다.

여섯째, 자원 브로커 서비스는 작업 스펙을 계산 자원들에게 분배하게 된다.

일곱째, globus-url-copy 와 같은 도구를 이용하여 위에서 선택된 리플리카는 각각의 계산 자원들로 이동하게 되고, 그것을 바탕으로 작업이 수행된다.

마지막으로 작업이 끝나면 자원 브로커 서비스는 해당 결과를 그리드 사용자에게 돌려준다.

위의 흐름의 3 번 단계에서 리플리카에 대한 정보를 그리드 정보 서비스로 제공하기 위해서는 리플리카에 대한 정보를 제공해 주는 Information Provider 가 필요하다. 그리고, 이렇게 제공된 정보는 4 번 단계에서 최선의 리플리카를 선택하는 기준이 된다.

6 번 단계에서 자원 브로커 서비스가 계산 자원들에게 작업을 분배하는 과정에서는 계산 자원들 중에

서 최선의 것을 선택하는 과정이 필요하다. 이러한 판단 기준은 계산 자원들에 대한 상태 정보를 제공해주는 Information Provider 의해 이루어진다.

7 번 단계에서 GT3 의 도구 중 하나인 globus-url-copy 는 근원지 URL 에서 명시된 파일을 목적지 URL 에 명시된 위치로 복사한다. 이러한 과정 중에 GASS transfer API 가 사용된다. 그것은 작업을 실행하는 과정에서 저장 장치간에 파일들을 stage in / stage out 을 하기 위해 사용된다. GASS 가 지원하는 모든 프로토콜들(예를들어, local file, http, https, ftp, gsiftp)이 지원된다. 그러나 GASS 에 의해 지원되지 프로토콜을 사용하지 않는 스토리지로의 데이터 접근은 불가능하다. SRB 의 경우도 이 경우에 해당한다. SRB 는 HRM, HPSS, DB2, Oracle, Illustra, ObjectStore, ADSM, UniTree 등을 포함하는 많은 스토리지 장치로의 인터페이스를 지원한다. 그림 1 에서의 흐름을 지원하기 위해 단순히 GASS protocol 들 뿐만 아니라, 다양한 저장 장치들과 리플리카 데이터들의 접근을 가능하게 하기 위한 SRB 프로토콜을 지원할 필요가 있고, 그러한 목적으로 globus-url-copy 의 수정이 이뤄졌다. [7]

또한, 7 번 단계에서 선택된 리플리카가 각각의 계산 자원들로 이동될 때, 작업을 실행할 수 있는 권한을 갖는 사용자가 사용할 수 있는 로컬 스토리지에 대한 용량 정보가 필요할 것이다. 이러한 정보 또한 Information Provider 에서 제공되어야 할 하나의 항목이 된다.

3. 스키마 설계

위의 요구사항 분석 단계에서 도출된 요구 사항들을 반영할 수 있는 스키마들의 정의가 필요하다. 그리드 환경에서 정보 모델을 정의하는 활동들이 세계적으로 진행되고 있다. 그 예로 EU-DataTAG[5]와 US-iVDGL[6] 프로젝트에 의한 결과물로 Glue 스키마가 있다. Glue 스키마는 그리드 자원들을 표현하기 위해 정보 모델과 스키마들간의 관계를 명료하게 정의한다. 그러나, 이렇게 진행되는 스키마 결과물이 모든 환경에 맞는 것은 아니다. 그림 1 에서 분석된 요구 사항들을 반영하기 위해 기존의 Glue 스키마의 내용을 변경 또는 확장하고, 새로운 스키마를 정의할 필요가 있다. 본 논문에서는 계산 그리드 환경에서 필요한 스키마를 More Computing Element (MCE) 스키마라고 정의하고, 데이터 그리드 환경에서 필요한 스키마를 More Storage Element (MSE) 스키마라 명명하였다.

MCE 스키마는 Glue 스키마의 주요 구조를 따르고 그림 1 의 요구 사항에서 6,7 번 단계에서 분석된 다음과 같은 항목을 추가적으로 반영한다. 실제로 ProcessorLoad 와 MainMemory 는 기존의 Glue 스키마에도 존재하는 요소이다. ProcessorLoad 는 Host 밑에 존재하고, MainMemory 는 Subcluster 의 하위 요소로 존재한다. 하지만, 실질적으로 각각의 계산 자원으로 사용되는 시스템의 경우 모든 자원들이 이상적으로 Homogenous 한 형태로 존재하는 경우보다는 각각의 특성에 따라 Processor Load 와 Main Memory 의 상태가 가변적인 경우가 발생한다. 따라서 Main Memory 와

같은 요소는 Host 밑에 존재하여 각각의 시스템 특성을 반영할 필요가 있다고 판단된다. 한편, UserStorage는 로컬 스토리지에서 작업을 수행할 수 있는 사용자가 이용할 수 있는 용량 정보를 제공해 준다. 사용자의 식별을 위한 Local ID 정보와 Distinguished Name(DN) 정보를 기본적으로 제공하고, 사용자에게 할당된 Quota 정보 및 Quota가 정해지지 않을 경우 기본적으로 사용할 수 있는 Filesystem의 용량 정보를 제공한다.

표 1. MCE 스키마 주요 항목

Element	Attribute	Description	Unit
UserStorage	LocalID	Local User ID	
	DN	User DN	
	Quota	Local User Quota	MB
	DefaultCapacity	Local User default Quota	MB
MainMemory	RAMSize	RAM size	MB
	RAMAvailable	Available RAM size	MB
	VirtualSize	Virtual RAM size	MB
	VirtualAvailable	Available virtual RAM size	MB
ProcessorLoad	Last1Min	1-minute average processor availability	%
	Last5Min	5-minute average processor availability	%
	Last15Min	15-minute average processor availability	%

resource는 SRB URI 정책에 포함되지 않는다[8]. replica는 도처에 존재하는 리플리카를 구별하는 식별자로 사용되고, resource는 SRB에 의해 관리되는 스토리지에 새로운 데이터를 생성하여 스토리지에 저장할 때 해당 스토리지 이름을 명시하는데 사용된다.

MCAT-enabled SRB 서버에서 생성되는 replica와 resource를 포함한 다양한 정보를 제공하기 위해 MSE 스키마는 정의된다. 기본 구조는 그림 3과 같고, 주요 항목은 표 2와 같다. 각각의 요소들은 다음과 같은 의미를 갖는다.

표 2. MSE 스키마 주요 항목

Element	Attribute	Description	Unit
SRB Resource	ResourceName	Resource name	
	ResourceLocation	Resource location	IP addr
	ResourceType	Resource description	
	ResourceClassName	Resource class name	
	AdminName	Admin name	
	DomainDesc	Domain description	
	ZoneID	MCAT Zone ID	
SRB Replica	CollectionName	Collection name	
	UserName	User name	
	ServerLocation	SRB server location	IP addr
	FileName	File name	
	FileSize	File size	Byte
	Filetype	File description	
Replica Detail	FileReplicationID	Replica ID	
	ResourceLocation	Resource location	IP addr
	ResourceName	Resource name	

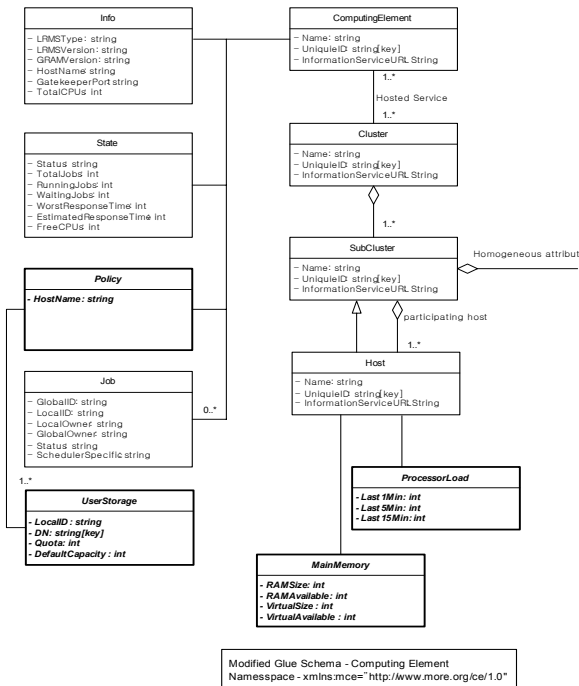


그림 2. MCE 스키마

MSE 스키마는 요구사항 분석의 3,7 번 단계에서 분석된 내용을 반영하기 위해 새로운 정보 구조를 구성한다. 수정된 globus-url-copy를 사용하기 위해 데이터 파일의 위치는 아래와 같이 URL 형식으로 표현될 수 있다.

```
srb:// [username.mdasdomain [zone] [:password] @] host [:port]
[?replica=replica_id][?resource=resources_name]/[path]
```

SRB에서 유지되는 데이터의 위치는 위와 같이 URI 형식으로 표현될 수 있다. 실제로 replica와

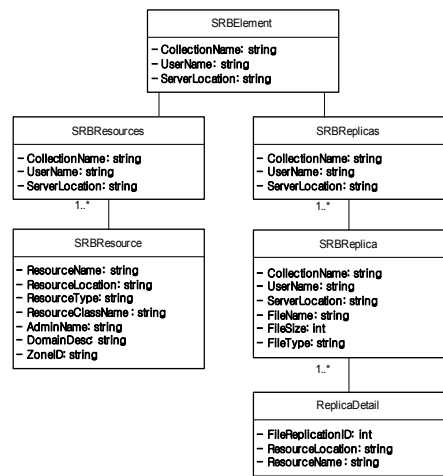


그림 3. MSE 스키마

- SRB Element: SRB Element는 MCAT-enabled SRB 시스템을 접근할 수 있는 시작점을 표현한다. SRB 계층당 SRB Element가 존재할 수 있다. 속성으로 CollectionName, UserName, ServerLocation 정보를 포함한다. replicas와 physical resources에 대한 모든 정보는 SRB Replicas와 SRB Resources 요소에 의해 구성된다.

- SRB Resources: SRB Resources는 SRB Resource를 그룹화 시키는 컨테이너 역할을 한다. 복수개의 SRB Resource 요소들을 유지할 수 있다.

- SRB Resource: SRB Resource는 ResourceName, ResourceLocation, ResourceType, AdminName 등과 같은

물리적인 자원들에 대한 속성 정보를 포함한다.

- SRB Replicas: SRB Replicas 는 SRB Replica 를 그룹화 시키는 컨테이너 역할을 한다. 복수개의 SRB Replica 요소들을 유지한다.

- SRB Replica: SRB Replica 는 FileName, FileSize, FileType 과 같은 파일에 대한 기본적인 정보를 유지한다.

- Replica Detail: Replica Detail 은 FileReplicationID, ResourceLocation, ResourceName 과 같은 각각의 리플리카에 대한 자세한 정보를 제공한다.

4. Information Providers 구현

GT3 는 Plug-in 이 가능한 Information Provider 메커니즘을 제공한다. 그것은 실행 가능한 외부 프로그램을 통해 서비스 데이터를 동적으로 생성하는 표준 메커니즘이다. 이러한 외부 Provider 프로그램들은 GT3 에서 기본적으로 제공하는 것이 있을 수 있고, 사용자가 필요에 의해 생성된 프로그램일 수 있다. MCE 스키마와 MSE 스키마에서 정의된 항목들을 GT3 에 적용하기 위해서는 표준 메커니즘을 따르는 새로운 Information Providers 들을 제작해야 한다. 각각의 스키마에 따라 MceScriptProvider 와 MseScriptProvider 를 제작한다. 이러한 Information Provider 들은 GT3 의 주요 서비스 중에 하나인 Resource Information Provider Service(RIPS)에 등록되어 주기적으로 수행되면서 각각의 스키마에서 정의된 정보를 XML 문서 형태로 출력하게 된다. 그림 4 는 제작된 Information Provider 들과 RIPS 와 관계를 보여준다.

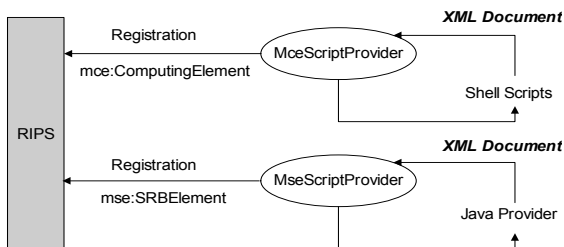


그림 4. Information Providers 구현

5. 결론

본 논문에서는 GT3 및 SRB 로 구축된 일반적인 계산 및 데이터 그리드 환경에서 사용자로부터 요구되는 정보를 분석하고, MCE 및 MSE 스키마를 정의하였다. MCE 스키마는 기존의 Glue 스키마의 구조를 따르며 ProcessorLoad, Mainmemory, UserStorage 등의 요소를 표현한다. 또한 MSE 스키마는 MCAT-enabled SRB 서버에 의해 관리되는 스토리지들과 리플리카를 포함하는 다양한 정보를 제공한다. 그것은 주로 SRB Element, SRB Resources, SRB Resource, SRB Replicas, SRB Replica, Replica Detail 등으로 구성된다. 각각의 스키마는 요구 분석된 정보를 표현하고, 그에 따라 Information Provider 를 제작하여 실제 환경에 적용해 보았다.

향후, 응용 과학 실험에서는 대규모 데이터가 생성될 것이고 이러한 데이터들을 공유하기 위해서는 계

산 및 데이터 그리드 환경이 안정적으로 구축되고 운영되어야 할 것이다. 본 논문에서 설계한 MCE 및 MSE 스키마들과 Information Provider 들이 유사한 그리드 환경에서 응용될 수 있을 것이다.

참고문헌

- [1] S. Tuecke, K. Czajkowski, I. Foster, J. Frey, S. Graham, C. Kesselman, T. Maguire, T. Sandholm and D. Snelling, P. Vanderbilt, "Open Grid Services Infrastructure(OGSI) Version 1.0", Global Grid Forum, 2003
- [2] I. Foster, C. Kesselman, J. Nick, and 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] SRB Project, <http://www.sdsc.edu/srb/>
- [4] Globus Project, <http://www.globus.org/>
- [5] DataTAG Project, <http://datatag.web.cern.ch/datatag/>
- [6] US-iVDGL Project, <http://www.ppdg.net/mtgs/ivdgl/Interoperability.htm>
- [7] Moredream Project, <http://www.moredream.org>
- [8] SRB URI, http://www1.ietf.org/proceedings_new/04nov/IDs/draft-gilbert-srb-uri-00.txt
- [9] Glue schema activity, <http://www.cnaf.infn.it/~sergio/datatag/glue/install/index.htm>
- [10] K. Czajkowski, S. Fitzgerald, I. Foster, and C. Kesselman, "Grid Information Services for Distributed Resource Sharing", in Proceedings of the Tenth IEEE International Symposium on High-Performance Distributed Computing (HPDC-10), IEEE Press, August 2001.