

국가 R&D 사이버인프라로서의 클라우드 서비스에 관한 연구

함재균⁰, 정기문*, 정현미*

⁰한국과학기술정보연구원

e-mail: jaehahm@kisti.re.kr⁰, {kmjeong, hmjung}@kisti.re.kr*

A Study on Cloud Computing Services as a National R&D Cyberinfrastructure

Jaegyoon Hahm⁰, Kimoon Jeong*, Hyun Mi Jung*

⁰Korea Institute of Science and Technology Information

● 요약 ●

본 논문에서는 다양한 과학기술 분야의 연구자들이 R&D 수행에 있어서 필요한 인프라로서 클라우드 서비스를 활발히 이용하도록 하기 위해서 요구되는 정책적·기술적 개선 방안을 도출하였다. 이를 위해서 R&D를 위한 클라우드 서비스의 해외 사례를 조사하고, 국내 출연(연)·대학 등 연구현장의 다양한 분야의 연구자들로부터 클라우드 서비스에 대한 인식 조사 및 수요 분석을 실시하였다. 이러한 조사·분석 결과 과학기술 및 공학·ICT 등 다양한 분야의 연구에서 요구되는 수요를 충족할 수 있는 클라우드 서비스의 필요성을 도출하였으며, 보안 및 행정제도 등에 있어서도 개선점을 도출할 수 있었다. 연구를 통하여 우리는 국가 R&D를 수행하는 연구자들이 기존 클라우드 서비스에 대한 제약을 해소하고 더 나은 연구환경을 지원받도록 하기 위한 R&D 사이버인프라스트럭처로서의 클라우드 서비스를 제안하였다.

키워드: 클라우드컴퓨팅(Cloud Computing), R&D 클라우드 서비스(R&D Cloud Service), 사이버인프라스트럭처(Cyberinfrastructure)

I. 개요

국내에서는 2013년 제정된 『클라우드컴퓨팅 발전 및 이용자 보호에 관한 법률』을 계기로 다양한 클라우드 활성화 정책이 추진되고 있으며 특히 공공 부문의 민간 클라우드 이용 활성화를 위해 국가 R&D에서의 클라우드 활용을 촉진하기 위한 노력을 기울이고 있다. 클라우드 산업 활성화 측면에서는 국가 R&D에서 발생하는 다양한 ICT 자원 및 서비스 수요를 민간 클라우드로 연계하여 활용하도록 함으로써 지속적인 클라우드 컴퓨팅에 대한 수요 창출이 가능하다. 그러나 기존 R&D 사업의 경우 자체적으로 ICT 장비 및 SW를 구축 활용하고 있으며 새로운 서비스에 대한 생소함 등의 이유로 연구자들이 클라우드로 전환하는데 어려움이 있는 것이 현실이다. 본 연구에서는 이러한 어려움을 해소하기 위해 연구자들의 클라우드에 대한 인식 조사 및 제약 사항을 파악하여 국가 R&D에서의 클라우드 컴퓨팅 활성화를 위한 개선 방안을 제안하고자 한다.

First Policy’를 통해 정부와 공공기관에 클라우드를 선제 도입하고 클라우드 컴퓨팅 확산에 노력하고 있다.

미국의 경우 2011년부터 국가 정보화 예산의 25%를 클라우드 도입에 활용하도록 권고하고 있으며, R&D 분야에서도 NSF(National Science Foundation)의 BigData 프로그램과 같은 사업을 통해 연구자들이 AWS, Google, Microsoft와 같은 퍼블릭 클라우드를 활용하도록 적극적으로 지원하고 있다.

이 밖에도 영국, 일본, 호주, 중국 등 많은 국가에서 공공부문의 클라우드 활성화를 위해 보안 및 인증제도 개선, 클라우드 스토어 등 활용저변 확대, 관련 법 제도 개선 등의 노력을 기울이고 있다.

2. 해외 R&D 클라우드 서비스 사례

2.1 NeCTAR Research Cloud(호주)

II. 관련 동향 및 사례

1. 해외 정책동향

세계 각국은 정부 주도형 클라우드 서비스 활성화 정책으로 “Cloud

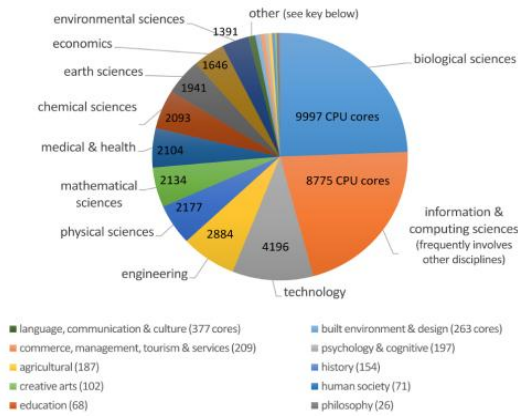


Fig. 2. Research Areas in NeCTAR Cloud

호주에서는 국가 과학기술 연구분야 연구자들을 대상으로 2012년부터 R&D 클라우드 서비스인 NeCTAR Research Cloud를 구축하여 연구자들이 필요로 하는 컴퓨팅·데이터·SW 인프라를 신속하게 이용하도록 온디맨드 셀프 서비스를 제공하고 있다[1]. 또 NeCTAR에서는 일반적인 클라우드 서비스 이외에도 공동연구를 수행하는 연구자들이 함께 협업할 수 있는 가상 연구실(Virtual Lab) 환경을 서비스하는 것이 특징이다. NeCTAR 클라우드 서비스는 현재 NCI, University of Melbourne 등 8개 기관이 보유한 34,000여 CPU 코어 규모의 자원을 연동이 연동되어 있고 잘 알려진 오픈소스 클라우드 플랫폼인 오픈스택(OpenStack)을 활용하여 서비스를 구축하였다 [Fig.1].

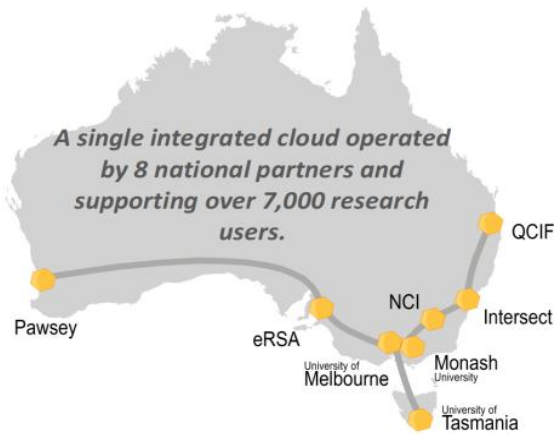


Fig. 1. NeCTAR Cloud Collaborating Institutes

지난해까지의 통계에 따르면 호주 전역에 걸쳐 7,000명 이상의 사용자를 지원하고 있으며 연구 분야로 보면 호주 과학기술 연구분야 중 95%를 지원하고 있다[Fig.2].

2.2 JetStream(미국)

JetStream은 미국의 대표적 사이버인프라스트럭처인 XSEDE 서

비스의 일부로서, 자체적인 인프라를 가지고 있지 않은 과학기술, 공학 분야 연구자들이 손쉽게 컴퓨팅 및 데이터 분석 서비스를 이용할 수 있도록 R&D 클라우드를 제공한다[2]. 2014년부터 NSF 사업으로 진행하고 있으며, Indiana University, TACC(Texas Advanced Computing Center), Chicago University, University of Arizona 등이 자원을 연동하여 참여하고 있다. JetStream에서는 계산 및 데이터 분석에 관련된 공용 라이브러리를 포함하는 미리 정의된 가상머신 이미지를 서비스할 뿐만 아니라 개인화된 가상머신 이미지를 활용할 수 있게 하여 다양한 응용분야 연구자, 기관, 대학 등에서 활용하고 있다. 클라우드 인프라 구축·운영 관리는 오픈스택을 활용하였으며 안전한 데이터 전송을 위해 Globus 기술을 활용하고 있다. 최근에는 오픈스택 Magnum 프로젝트를 활용하여 컨테이너 오케스트레이션 서비스도 제공한다.

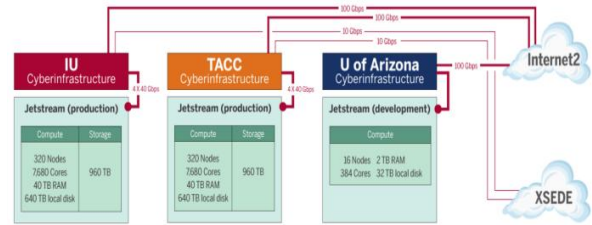


Fig. 3. JetStream Cloud Architecture

2.3 CERN Cloud (EU)

CERN은 스위스 제네바에 위치한 입자물리 연구소로서 LHC(Large Hadron Collider)이라는 세계에서 가장 큰 입자가속기를 통해 ALICE, ATLAS, CMS 등의 거대 실험을 수행하고 있다. 이러한 입자가속기 충돌실험에서 얻어지는 거대 데이터를 분석하기 위해서는 그리드 컴퓨팅 등 다양한 인프라를 활용하고 있는데 그중에는 클라우드 컴퓨팅 서비스도 포함된다.

CERN은 데이터량이 증가함에 따라 데이터 분석 효율성 및 자원 활용의 효율성을 높이기 위해 2012년부터 컴퓨팅 자원의 클라우드 전환 도입을 시작하였고, 2016년 기준 4개의 사이트에서 20만개 이상의 코어를 오픈스택 클라우드로 제공하고 있다[3]. CERN 클라우드 서비스의 목적은 크게 두가지로, 연구진 도입을 위한 테스트베드, 장기 실험 프로그램 지원, 멀티 OS 요구사항 등의 수요를 지원하여 서비스 효율성 및 대응능력을 향상하고, 유휴자원 활용, 가변성 높은 부하 대응 등의 자원 활용성 향상에 있다. 이러한 서비스를 오픈스택 기술에 기반을 두고 구축하였으며, 최근에는 자체 클라우드 서비스와 상용 클라우드와의 하이브리드 형태의 클라우드 서비스에 대한 다양한 시도를 하고 있다.

III. 국내 연구자 인식 및 수요조사

우리는 국내 연구자들의 클라우드 서비스에 대한 인식과 수요를 파악하기 위해 선별된 연구자 그룹과의 심층인터뷰와 그보다 넓은 범위의 연구자들을 대상으로 설문조사를 실시하였다.

1. 분야별 연구자 심층인터뷰

실제 연구자들의 클라우드에 대한 활용현황, 인식수준, 수요 등을 조사하기 위해 과학기술 출연(연) 연구현장을 방문하여 인터뷰를 시행하였으며, 이에 대한 개요는 [Table 1]과 같다

Table 4. Focus Group Interview Overview

구분	내용
대상	R&D 사업을 수행중인 출연(연) 연구원 16명
방법	현장 방문을 통한 개별 인터뷰
분야	전기, 재료, 천문, 생명공학, 환경, 핵융합(플라즈마 기술) 분야 등
목적	과학기술 분야 연구자들의 클라우드 활용현황, 인식 수준, 수요 파악

분야별로 연구방법, 기술수준 등에 대한 차이가 있어 다양한 의견이 도출되었으나 공통적으로 파악한 요구사항은 [Table 2] 와 같이 요약할 수 있다.

Table 5. Focus Group Interview Result Summary

대상 항목	내용
연구 데이터	<ul style="list-style-type: none"> 천문학, 생명공학 등의 분야에서는 특히 데이터 기반 연구가 활발하며 대규모 데이터에 대한 관리, 공유 및 분석 환경에 필요한 클라우드를 필요로 함 대부분의 분야에서 실험을 통해 발생하는 데이터의 관리 및 분석을 위한 플랫폼을 필요로 하며 소프트웨어 환경에 대한 요구사항이 큼
연구 자원	<ul style="list-style-type: none"> 연구에 필요한 인프라로서 계산자원 및 스토리지 등에 대한 요구가 기본적으로 존재 보안정책 등의 이유로 연구실 내 자체 장비를 구축하는 경우가 많으며, 이경우 자원관리를 위해 연구자 중 추가 업무를 떠맡음 IT 기술에 대한 전문성이 부족하여 필요한 소프트웨어, 플랫폼 설치 및 유지보수에 애로사항이 있음
정책 및 제도	<ul style="list-style-type: none"> 공공기관 보안정책에 따라 클라우드 등 외부 서비스 이용에 제약이 큼 클라우드 이용료에 대한 연구비 집행 등에 있어서 기존 연구비 관리체계와 괴리가 있어 연구현장에서 혼선과 어려움 상존

2. 클라우드 서비스 활용의견 설문조사

보다 다양한 사용자를 대상으로 과학기술 및 ICT 분야의 국가 R&D 사업에서의 클라우드 이용 현황을 파악하고 R&D 수행을 위해 필요로 하는 ICT 기반 연구 환경 서비스에 대한 수요를 파악하기 위해 대학 및 출연(연) 연구자를 대상으로 설문조사를 [Table 3]과 같이 시행하였다.

Table 6. Online Survey Overview

구분	내용
설문대상	R&D 사업을 수행하는 연구자(기업, 대학, 출연(연) 등)
설문내용	R&D에 이용하는 ICT 활용 현황, 클라우드 서비스 인식, 클라우드 서비스 수요
설문문항	15문항
응답수	180명(기업 20명, 대학 54명, 출연(연) 106명)

2.1 ICT 자원 활용 현황

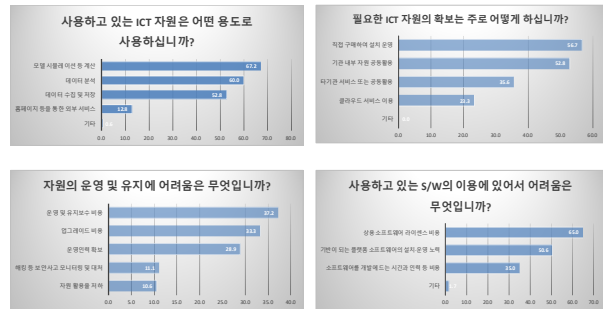


Fig. 4. ICT Resource Use for R&D

연구자들은 R&D를 수행하면서 모델 시뮬레이션 등 계산, 데이터 분석, 수집 및 저장 등의 목적으로 ICT 자원을 사용하고 있으며, 이에 필요한 자원은 직접 구매하여 설치·운영하는 경우가 가장 많고 클라우드를 이용하는 경우는 적은 편이었다. 이러한 자원의 운영에 있어서 어려움은 유지보수·업그레이드 비용이 가장 많았으며 운영인력 확보도 어려움으로 꼽았다. 또 소프트웨어 이용에 있어서 어려움은 상용 소프트웨어 라이선스 비용이 가장 많았고, 플랫폼 소프트웨어와 응용 소프트웨어의 관리·개발에 소요되는 노력에 어려움 겪고 있었다.

2.2 클라우드 이용 관련 의견

클라우드 서비스의 경우 응답자 중 60%가 클라우드 이용 경험이 있으며, 35%는 이용경험은 없으나 향후 이용할 용의가 있다고 응답하였다. 클라우드를 이용한 목적은 데이터의 저장 및 관리가 가장 많았으며, 클라우드를 이용한 이유는 ‘유저 인터페이스의 편리해서’, ‘비용이 저렴해서’, ‘필요한 소프트웨어를 설치없이 이용할 수 있어서’ 등의 순서로 응답하였다. 클라우드의 장점으로는 ‘데이터 백업 및 보안 등 안정성’, ‘협업 R&D 용이성’, ‘유연한 자원의 확장성’, ‘자원의 빠른 획득’으로 꼽았고, 반대로 클라우드 이용의 장애물은 ‘보안에 대한 우려’, ‘데이터 및 보안 정책 등 규제’, ‘데이터 전송시간’ 등을 꼽았다.

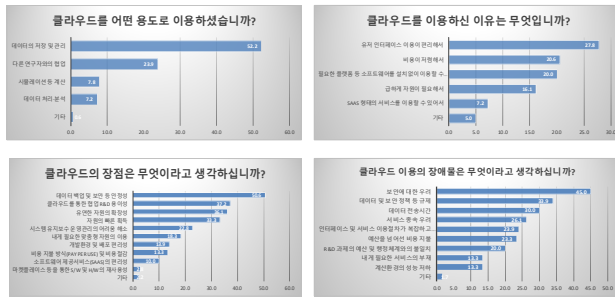


Fig. 5. Understandings on Cloud Service

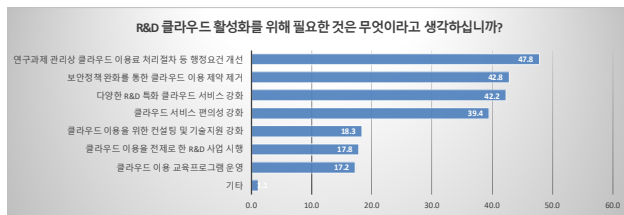


Fig. 6. Challenges on Stimulation of Cloud Services for R&D Use

특히 R&D 클라우드의 활성화를 위해서는 R&D과제관리의 행정 절차 개선, 보안정책 완화, 다양한 R&D 특화 클라우드 서비스 강화 등이 필요하다는 응답을 하였다.

3. 수요조사 결과 시사점

수요분석 결과 R&D 연구자들의 클라우드 서비스의 활성화를 위해서 고려할 사항으로 기능적 측면과 정책·제도적 측면으로 나누어 볼 수 있을 것이다.

기능적 측면에서는 가장 기본적인 인프라 서비스로부터 안정적 서비스 제공이 우선되어야 하고, 분야별로 오픈소스 소프트웨어를 편하게 이용할 수 있는 플랫폼 형태의 서비스의 필요성을 도출할 수 있다. 특히 연구데이터의 활용을 효율적으로 하기 위한 데이터 중심의 서비스가 필요하다고 여겨진다.

정책·제도적 측면에서는 연구데이터의 외부 저장 및 클라우드 서비스 이용을 위한 공공기관 보안 정책을 보다 세밀하게 제시하고 공공기관 특성을 고려한 상세 클라우드 서비스 가이드라인의 필요성이 있다. 또한 연구과제 관리 체계상 클라우드 서비스 이용을 위한 행정절차의 개선이 필요하고 연구자들의 거부감을 최소화 할 수 있는 이용료 정책 또한 필요하다고 생각할 수 있다.

IV. 결론 및 제안사항

클라우드 서비스는 연구자들에게는 R&D에 필요한 ICT 자원 및 SW 등을 서비스를 통해 이용함으로써 R&D 생산성을 향상시킬 수 있는 좋은 도구이다. 우리는 수요조사 및 연구를 통해 R&D에 특화된 클라우드 서비스가 필요하다는 결론을 도출하였다. 이는 R&D에서 요구하는 자원 및 서비스는 그 분야에 따라 수요가 매우 다양하여 고도의 유연성 있는 서비스가 필요하며, 단순한 인프라 서비스에 머무르지 않고 다양한 계층(인프라-플랫폼-SW)에서 맞춤형 서비스가

필요하기 때문이다. 특히 R&D에 요구되는 보안의 수준을 고려한 맞춤형 보안 정책 및 서비스가 제공되어 한다. 이러한 수요를 고려하여 [Fig.7]와 같은 서비스를 제공할 수 있는 R&D 특화 클라우드의 구축이 필요하다.

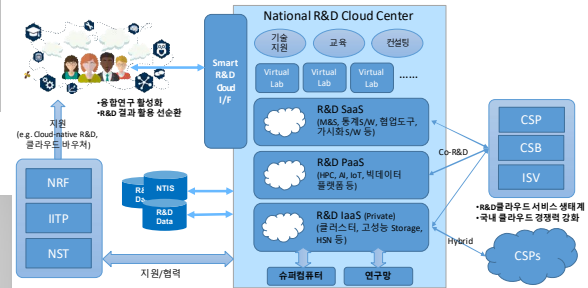


Fig. 7. R&D Cloud Service Architecture

REFERENCES

- [1] National eResearch Collaboration Tools and Resources, <https://nectar.org.au>
- [2] Jetstream Cloud, <https://jetstream-cloud.org>
- [3] CERN Cloud, <https://home.cern/about/experiments/cloud>
- [4] White House(2011), "Federal Cloud Computing Strategy".
- [6] Australian Government Department of Finance(2014), "Australian Government Cloud Computing Policy Version3.0".
- [7] EUROPEAN COMMISSION(2012), "Unleashing the Potential of Cloud Computing in Europe"