

e-Science: IT 기반의 또 다른 혁신

국민대학교 황선태 · 허대영

건국대학교 정갑주

한국과학기술정보연구원 조금원 · 이상동 · 이지수

1. 서론

e-Business(전자상거래), e-Government(전자정부), e-Learning(사이버교육) 등과 같은 IT 기반의 경제·사회적 혁신은 인류 사회 삶의 방식과 기업 및 국가 경쟁력을 바꾸어 놓았다. 이러한 IT 기반의 혁신들의 공통적 특징으로 첫째, 오프라인 정보(서류상의 정보)의 인터넷에서 검색·접근 가능한 온라인 콘텐츠화, 둘째, 이러한 온라인 콘텐츠를 이용해서 복잡하고, 비효율적이고, 고비용인 오프라인 비즈니스와 사무를 단순하고, 효율적이고, 저렴한 인터넷상에서 처리가능한 온라인 서비스화, 마지막으로 이러한 온라인 콘텐츠와 서비스를 바탕으로 현실에서 불가능한 새로운 창의적 온라인 복합 서비스 창출 등을 들 수 있다. 이를 통해서 단순히 편리성 증대뿐만 아니라, 창의적 아이디어만 있으면, 대규모 경제적 투자 없이도 엄청난 새로운 부가가치 창출이 가능하다. 이러한 점에서 국토가 작고, 천연 자원이 부족한 우리나라는 21세기 이후 세계화된 경제·사회에서 주도적 역할을 하기 위해서 이러한 다양한 분야의 IT 기반의 혁신에 충력을 다해야 한다.

미국 NSF Atkins 보고서[1]에서 전문가들은 "e-Science는 한 세대를 대표할 수 있는 획기적인 사업이다."라고 언급하였다. e-Science는 국내외에 위치한 연

구자, 연구장비, 연구정보 등의 연구자원을 사이버 공간에서 공동 활용하며 연구를 수행하는 차세대 연구활동을 말한다. e-Science의 목표가 실현되는 경우, 과학기술분야에서도 e-Business, e-Government 분야와 유사한 거대한 변화가 발생할 것으로 기대된다. 우선, 다양한 과학실험 결과의 온라인 콘텐츠화(과학지식에 대한 인터넷을 통한 검색 및 접근 가능), 실험·계측장비의 온라인화(첨단 고가장비의 공동활용 가능), 과학기술 온라인 커뮤니티를 통해 과학기술전문가의 온라인화(인터넷을 통한 공동연구 가능)가 구현되어 전 세계 과학계가 하나의 거대한 IT 기반 과학기술 연구환경을 구성하게 될 것이다.

이러한 IT 기반 연구환경이 국내 또는 국가간 대용량의 협업연구가 요구되는 지구환경, 고에너지물리학, 생명공학, 천문우주, 나노물질 해석 등의 연구분야에 적용할 경우, 연구생산성의 획기적인 개선이 기대된다. 즉 과학자들의 전통적 실험실 중심 연구방식에서 인터넷 기반·컴퓨터 기반 연구기법의 역할이 점차 증대될 것이고, 더 나아가 이러한 IT 기반의 과학기술 연구환경상에서만 가능한 새로운 방식의 첨단 연구방법들이 개발되어질 것이다.

최근에는 e-Science의 적용범위가 순수과학의 범주를 벗어나 자동차, 항공, 선박 등의 산업기술 뿐만 아

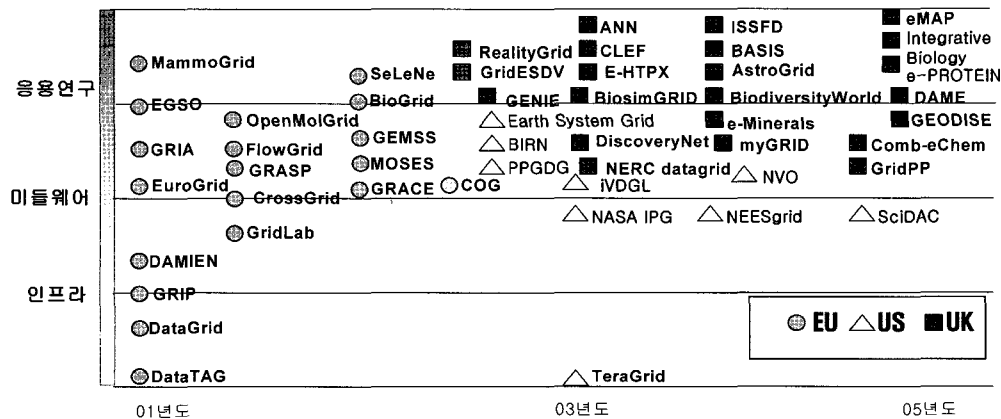


그림 1. 미국, 영국 및 EU의 e-Science 사업의 추진 동향

나라 교육, 국방, 경제 등의 분야까지 확대되고 있다. 또한 e-Science 환경에서는 일반인도 첨단 과학지식(온라인 콘텐츠)을 쉽게 접근할 수 있기 때문에, 고가의 실험시설이나 전문 과학자가 없더라도, 일반 교육기관에서도 우수한 과학기술 전문인력 양성을 위한 첨단 과학기술 교육프로그램을 제공할 수 있게 된다.

이러한 이유에서 성공적 e-Science 환경구축은 향후 세계 과학기술분야를 우리나라가 선도하는데 매우 중요하다. 그림 1은 2000년도부터 최근까지 미국, 영국, 그리고 EU가 국가주도로 추진하고 있는 e-Science 사업의 변화추이다. 그림 1에서 보는 것과 같이 2000년도에는 e-Science 프로그램이 이의 기반이 되는 인프라 구축에 중점을 두었으나, 2003년 이후로는 이를 기반으로 하여 실제 활용을 위한 연구개발로 전환되어 추진되고 있음을 알 수 있다.

미국, EU, 영국, 일본, 그리고 후발국인 중국 등에서도 e-Science를 차세대 연구환경으로 인식하고 있으며, 국가별로 연간 1,000억원 이상 투자하고 있다. 기술의 발전 속도에 의하면 e-Science가 수년 내에 보편적인 연구환경으로 활용될 수 있을 것으로 예상되고 있다[2,3].

2. e-Science 관련 해외 동향

미국 등 선진국들은 디지털경제 추진기인 1990년대부터 국가 경쟁력 차원에서 정보통신 기반 확충에 국가적 역량을 집중하고 있다. 특히, 경제 및 사회제도, 시장구조, 산업구조, 과학기술구조에 있어서 총체적인 전환노력을 기울이고 있으며, 더 나아가 국익을 위해 국제사회에서 정보통신기술(IT) 산업의 주도권을 확보하려 하고 있다. 이에 따라 선진국에서는 최신 정보통신기술을 이용하여 과학연구 환경에서 대규모 공공연구, 그리고 시간, 공간 및 물리적 장애를 넘어 서로 협력 작업이 가능하도록 하는 등 국제경쟁력 제고를 위해 다양한 형태의 e-Science 프로젝트를 정부차원에서 수립하여 추진하고 있다.

2.1 영국 e-Science 프로그램

영국은 e-Science의 용어를 처음으로 사용한 국가이며, e-Science 프로그램(기간: '01~'06, 5천억원)을 범부처 차원에서 추진하고 있다[4-6,7,8]. 2000년도 영국정부의 과학혁신백서[9]에 따르면 e-Science를 최우선 사업으로 지정하였으며, 이를 기반으로 순수 과학에서 산업기술까지의 대형 및 융합적 연구개발 프로그램을 추진하고 있다. 영국의 e-Science 추진 배경은 역동적인 지식기반 경제의 토대구축 및 혁신을 위

한 기회증대 및 과학에 대한 대중의 신뢰회복에 있다. 특히, 미국 및 타 유럽국가에 비해 뒤떨어진 과학기술 투자 및 고급인력의 해외유출 방지, 과학기술의 아이디어 상품화, IT와 응용연구를 결합한 융합형 연구 등을 진행하고 있으며, 이러한 모든 과정이 웹을 통해 해결될 수 있는 방법을 연구하고 있다.

영국 e-Science 프로그램은 두 가지 형태로 운영되고 있다. 첫째는 참여하는 각각의 연구회의 전문분야에서 처한 문제들을 해결하기 위한 각 연구회별 연구사업이며, 둘째는 모든 연구회가 직면하고 있는 공통적인 문제를 해결하기 위한 코어(Core) 프로그램이다. 이러한 사업들의 조정체계는 각 연구회가 OST(Office of Science and Technology)와 협력을 체결하고, 목표와 진행상황을 보고하는 것이다. 또한, 매년 All Hands Meeting 등을 개최하여, 성과발표, 데모 및 산업체 전시를 통해 일반 국민들까지도 e-Science를 이해할 수 있도록 기회를 제공하고 있다.

표 1 영국 e-Science 사업 추진 조직도

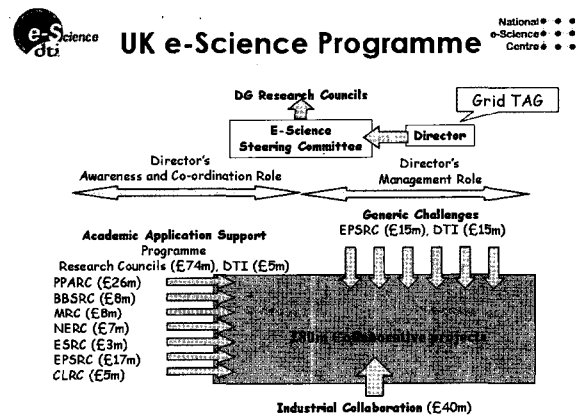
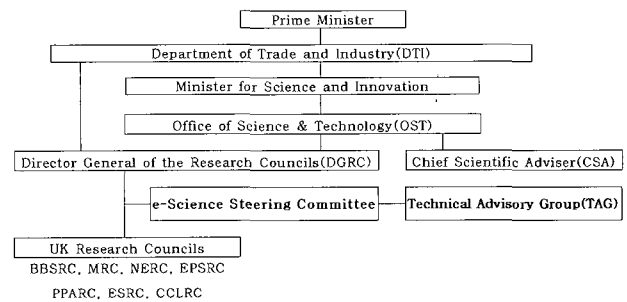


그림 2 참여기관별 구성도

영국의 e-Science 프로그램의 수행체계는 OST 산하에 과제책임자가 있으며(표 1 참조), EPSRC(Engineering and Physical Science Research Council)[10]가 연구회를 대표하여 운영위원회를 구성한다. 운영위원회는 사업의 진도 및 결과를 점검하는 역할을 한다. 더불어 e-Science 사업에 참여하는 과제는 그림 2와 같이 산업체를 반드시 포함하도록 하여, 연구결과의 활용 및

산업화를 적극 권장하고 있다.

영국 e-Science 프로그램의 핵심이 되는 코어(Core) 프로그램은 그림 3과 같이 하나의 National e-Science Centre와 8개의 지역 센터로 구성되는 네트워크 조직이 핵심주체가 되어 진행되고 있다. 코어 프로그램의 역할은 지역별로 선정된 e-Science 센터간의 네트워크 구축, 각 산업별로 특성에 맞는 미들웨어 개발, 대국민 활용을 위한 e-Science 응용과제 발굴 등이다. 그리고, 각 지역의 e-Science 센터는 산업체와 함께 산업적 및 상업적인 응용연구를 개발함으로써, 실제로 e-Science가 활용되는 환경을 구축하고 있다.

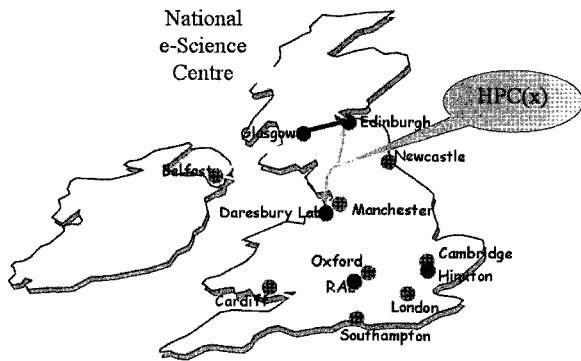


그림 3 영국의 e-Science 센터

• e-Science 응용별 환경구축 현황

영국의 e-Science 응용환경 구축과제는 코어 프로그램에서 지원하고 각 연구회 별로 추진하고 있다. 응용영역은 크게 7개 분야 즉, 생태과학, 공학 및 물리학, 의학, 미들웨어(전산학), 분자물리학 및 천문학, 사회과학으로 나누고 있다. 본 원고에서는 EPSRC와 PPARC(Particle Physics and Astronomy Research Council)[11] 및 NESC(National e-Science Center)[12]에서 추진하는 응용에 대해서 간단히 소개하기로 한다.

EPSRC에 의해 운영되는 프로젝트는 농축물질과 농축재료의 분석 수단을 제공하는 RealityGrid, 물질의 구조와 특성을 연구하는 Comb-e-Chem, 분산된 항공기 유지보수 환경을 위한 DAME, 항공기 설계를 위한 GEODISE, 바이오 연구를 위한 MyGrid 등 6개 과제로 이루어져 있다.

PPARC에서는 스위스 CERN 연구소의 초대형 입자가속기에서 발생하는 데이터를 분석하기 위한 과제로 GridPP[13] 프로젝트를 수행하고 있으며, 다양한 파장대의 방대한 천문 관측데이터를 공동활용 할 수 있는 가상천문대(Virtual Observatory)를 만드는 AstroGrid[14] 프로젝트를 진행하고 있다. 그리고,

국민 의료복지 분야인 유방암 진단 시스템(eDIAmond)[15]이 그림 4와 같이 EPSRC의 후원으로 Oxford 대학에서 수행되고 있으며, 이미 상업화 단계에 있다.

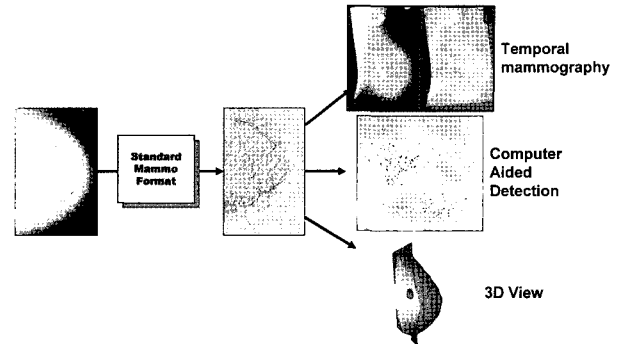


그림 4 eDIAmond 프로젝트 활용 예시

현재의 유방암 등 진단을 위한 이미지는 매우 비효율적인 방법으로 체계화가 되어있거나 통합되지 않은 방식으로 관리되고 있다. 많은 정보들이 기존에 있음에도 불구하고 이를 십분 활용하지 못하고 계속 새로운 이미지를 찍어내기에 바쁘다. 예를 들어, 한 이미지를 다른 병원에서 사용하기 위해 전송하는 것은 매우 불편한 일이거나 불가능하기까지 하다. eDIAmond 프로젝트는 우선 이러한 이미지를 효율적이고 체계적이며 통합적인 이용을 가능하게 하여 언제 어느 곳에서든지 필요한 이미지를 사용할 수 있게 할 뿐 아니라 질병의 진단 등 여러 파급효과를 노리고 있다. 이러한 통합관리 클리닉 시스템을 완성하는 것을 목표로 eDIAmond 프로젝트에서는 우선 표준안을 만들고 있으며, 이미지의 촬영에서부터 저장, 관리, 열람, 심지어 이를 이용한 컴퓨터의 자동 질병진단까지 진행하고 있다.

2.2 미국의 e-Science 프로그램

미국에서 추진하고 있는 e-Science를 설명하기 위해 우선 사이버인프라스트럭처(CI: CyberInfrastructure)를 소개한다. CI는 미국의 차세대 슈퍼컴퓨팅환경으로 미국내의 대표적인 슈퍼컴퓨팅센터를 고속의 네트워크로 연동하여 단일자원처럼 활용하고자 추진하는 인프라구축 사업이다. CI의 대표적인 프로젝트가 테라그리드(TeraGrid)[16]이며, 여기에는 샌디에고 슈퍼컴퓨팅센터, 일리노이 슈퍼컴퓨팅센터, 피츠버그 슈퍼컴퓨팅센터, 아르곤 국가 연구소 등이 참여하고 있다. 이러한 인프라 구축사업의 기반으로 ITR(Information Technology Research), DDDAS(Dynamic Data Driven Application Simulation) 등의 e-Science 프로그램이 추진되고 있으며, 각 사업별로 기초과학, 재난방지, 산업기술 등의 분야에 다수의 응용과제가 수

표 2 새로운 지식환경의 통합을 위한 사이버인프라스트럭처 구성도

연구와 교육을 위한 공동체 중심의 지식환경 (collaboratory, co-laboratory, e-Science 연합체, 가상 연합체) Customization for discipline-and project-specific applications				
고성능 컴퓨팅 서비스	데이터, 정보 지식관리 서비스	실험장치 의 관찰, 측정, 보정 서비스	인터페이스, 가시화 서비스	협력환경 서비스
네트워킹, 운영체제, 미들웨어				
기본 기술: 계산, 저장장치, 통신				

행되고 있다. 표 2는 CI의 구성도를 나타낸다.

지난 10년 동안 미국은 NSF의 테라그리드, NASA의 IPG(Information Power Grid), DoE의 ASCII (Accelerated Strategic Computing Initiative) 등 다수의 그리드 프로젝트, e-Science 구축, 그리고 응용개발 관련 프로젝트 수행을 통해 많은 e-Science 응용이 개발되고 있으며, 구현 가능하다고 예상하고 있다. 현재 대두되고 있는 분야는 뇌지도, 의료 영상, 원격 의료, 분자 정보학 등의 건강과 의료, 공학, 입자 물리학, 천문학, 화학과 재료공학, 환경과학, 바이오학과 유전학, 교육 및 디지털 도서관 등이다.

e-Science에서 가장 빠르게 성장하는 분야 중의 하나는 생명과학이다. e-Science는 계산생물학, 생물정보학, 유전학, 계산신경과학 등에서 데이터 액세스, 수집 및 마이닝을 위한 방법으로(예를 들면, 단백질 은행 등) 대규모 모의시험 및 분석을 위해, 그리고 전자현미경 같은 원격실험장비연동을 위해 활용되고 있다. 대표적인 프로젝트가 그림 5의 BIRN(Biomedical Information Research Network)[17] 프로젝트이다.

또 하나의 대표적 e-Science 프로젝트가 그림 6의 GEON(지구과학 네트워크)[18] 프로젝트이다. GEON은 여러 기관의 정보기술분야 연구자와 지구과학연구

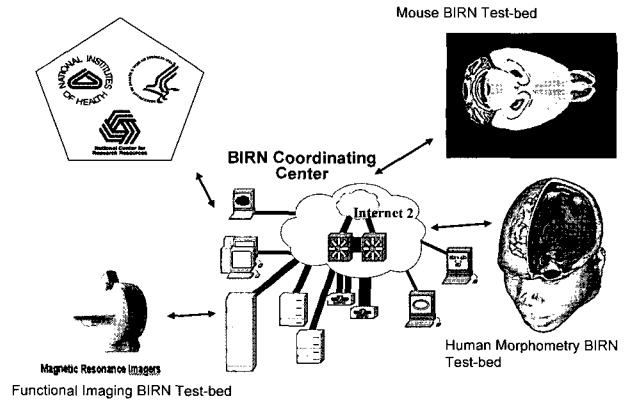
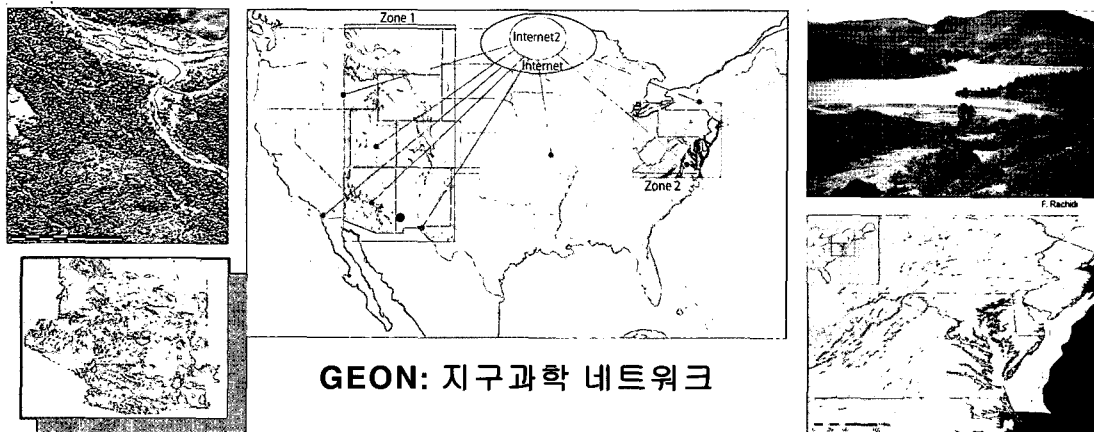


그림 5 BIRN의 시스템 구성도

자들을 대규모 협력구조로 참가시킨 NSF의 정보기술 연구 프로젝트이다. 사이버 연구환경 구조를 제공하기 위해 GEON은 데이터 공유프레임을 구축하고, 최고의 관행을 식별하고 지구과학이 이루어지는 방식에 극적인 진보를 이룰 수 있게 해주는 유용한 기능과 도구를 개발하고 있다. 이러한 노력은 과학적인 도구와 데이터에 대한 액세스를 위해 사용자 친화적인 과학환경을 제공하고 답변할 수 있는 과학적 질문의 범위를 상당히 넓힐 것으로 기대하고 있다. 또한 연구자들은 광범위한 자원에 대한 제어된 액세스를 활용하기 위해 GEONGrid 포털을 통해 데이터, 도구들 그리고 모델들을 발견할 수 있다.

재난대비를 위한 대표적 e-Science 프로젝트가 EarthScope[19]이다. EarthScope 프로젝트는 슈퍼컴퓨터, 초고속 연구네트워크, 센서 네트워크 등을 이용하여 북아메리카 대륙의 지질구조 분석, 지진발생, 화산폭발 등의 원인을 규명하는 연구를 수행한다. 최근 e-Science의 활용이 지진, 화산폭발과 같은 위험에 대한 연구를 진행하거나 천연자원을 발굴할 때 수반되는 위험요소를 줄이는 데에 일조하고 있다. 더불어, 인터넷 홈페이지에서 실시간으로 미국 전역에 설치돼 있는



GEON: 지구과학 네트워크

그림 6 GEON 프로젝트의 구성도

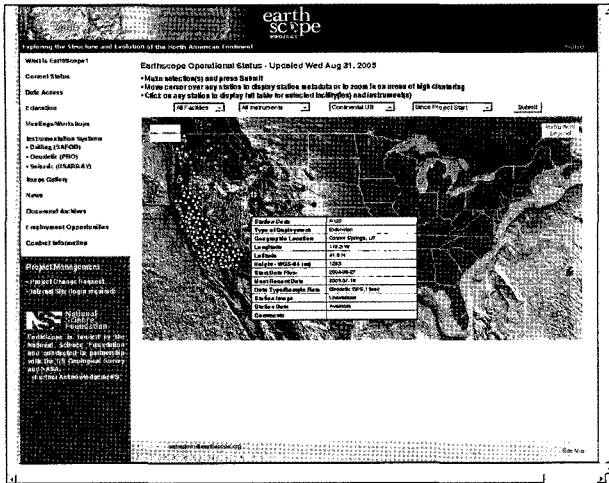


그림 7 EarthScope 프로젝트의 구성도

센서 정보를 조회하거나 3차원 영상의 재구성, 시각적 데이터해석 등을 지원하고 있다.

미국에서 진행되고 있는 e-Science는 특정 분야에 대해서 대규모로 추진되고 있으며, 대표적인 응용과제는 NEON(National Ecological Observatory Network) [20], LOOKING(Laboratory for the Ocean Observatory Knowledge Integration Grid)[21], NEES(Network for Earthquake Engineering Simulation)[22], OptIPuter(Optical networking, Internet Protocol, computer storage, processing and visualization technologies)[23], LEAD (Linked Environments for Atmospheric Discovery) [24] 등이 있다.

2.3 EU의 e-Science 사업

EU 정부는 차세대 지식기반 사회에서 세계최고의 경쟁력을 지속하기 위한 전략을 2000년에 리스본(Lisbon)에서 수립하였으며, 이후 2001년 3월에 스톡홀름에서 EU 정상들간의 협의를 통해 eEurope, European Research Area에 대한 연구개발 인프라 구축을 결정하였다[25]. 이를 통해 각 국가에서 독립적인 e-Science 프로젝트를 추진하지만 범 EU 차원에서도 체계적인 e-Science 프로젝트를 추진하고 있다. EU 프로젝트들은 그리드 기술을 기반으로 과학기술 및 산업분야까지 확장을 추진하고 있으며, 예산도 FP6에서 2배정도 증액하여 추진하고 있다.

유럽 전체를 대상으로 추진하는 대표적인 e-Science 프로젝트가 EGEE(Enabling Grids for e-Science in Europe)[26]이다. EGEE 프로젝트는 EC(European Commission)에서 주관하는 사업이며, HEP(High Energy Physics)[27], 생명정보학 등 과학기술의 여러 분야를 지원한다. HEP의 예로 그림 8과

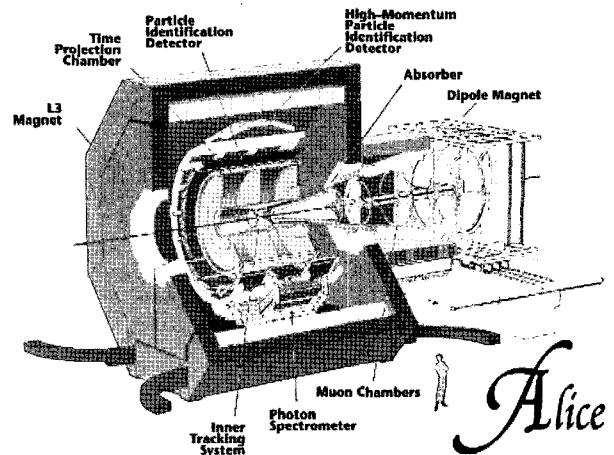


그림 8 EGEE에서 활용될 Alice 장비

같이 과학적 실험(ALICE: A Large Ion Collider Experiment)[28]에 의해 발생하는 대용량의 자료 분석을 위해 필요한 컴퓨팅 자원을 공유하고, 이를 상호 활용할 때 필요로 하는 기술을 개발하기 위한 프로젝트이다. 현재 여러 연구소에서 수 천명의 연구자가 수 만대의 컴퓨터에 동시에 접속하여 테라급(10^{12})의 막대한 자료를 공유하여 해석할 수 있도록 확장성 있는 소프트웨어와 테스트베드를 구축하고 있다. EGEE 프로젝트는 스위스에 위치한 CERN과 유럽 연합 소속의 원자력 연구 조직체를 포함한 75개의 연구소들이 주도적으로 참여하고 있다.

EGEE와는 별개로 EU FP6에서 수행하는 여러 가지의 프로젝트는 타 국가의 사업과 다른 특징을 가지고 있다. 그것은 대형의 극한 문제에 대한 e-Science 응용분야를 선정하고 이를 지원할 수 있는 요소기술(워크플로우, 모바일, 지식관리 등) 개발과 결과의 상업화에 중점을 두고 있다. 이러한 다양한 기술은 포괄적으로 SIMDAT(<http://www.simdat.org/>)[29] 프로젝트에 적용되고 있으며, SIMDAT 프로젝트는 분산된 데이터를 지식기반으로 구축하고 이를 이용하여 제품의 해석 및 설계 등을 수행하는 사업이다.

2.4 일본의 e-Science 현황

일본은 2000년 IT 기본법을 제정하고, 5년 이내 세계 최고의 IT 강국을 건설하기 위해 2001년 e-Japan 전략을 수립하였다. 일본의 e-Science 정책은 e-Japan 전략을 기반으로 추진된다. 일본의 e-Science는 문부과학성(MEXT)과 종합과학기술회의(CSTP: Council for Science and Technology Policy) 소속 정보기술위원회에 의해 주도되고 있다. 일본은 2001년부터 가상 연구환경인 ITBL(Information Technology Based Laboratory)[30], Biogrid[31] 프로젝트 등

의 e-Science 프로젝트를 수행하고 있다.

Biogrid는 대용량의 생명과학 분야 데이터를 처리하기 위해 필요한 환경을 구축하는 프로젝트로서 e-Science 환경구축을 위한 핵심 그리드 기술 그룹, 대용량 데이터의 분석이나 시뮬레이션을 위한 계산 그리드 기술 그룹, 대용량 데이터를 검색하기 위한 데이터 그리드 기술 그룹, 원격에서도 데이터를 수집할 수 있는 원격 데이터 수집 기술그룹으로 나누어 추진하고 있다. 일본의 e-Science 동향의 특징은 대규모의 사업을 추진하고 있으나, 대부분이 인프라 구축에 중점을 두고 있다는 점이다

3. e-Science 성공 전략

앞에서 살펴본 해외 사례를 볼 때에 e-Science의 성공을 위해서 필요한 핵심 사항은 두 가지로 압축될 수 있다. 즉 IT 인프라에 대한 투자와 그 인프라 위에 올라갈 콘텐츠의 확보이다. 두 가지 다 소수의 특정 개인, 기업 또는 기관에 의해서 해결되기 어려운 많은 장애요인이 존재한다. 따라서 범 국가적 차원의 인프라에 대한 투자와 새로운 인프라의 개념을 살릴 수 있는 정책에 의한 동기 부여가 절실히 요구된다.

3.1 e-Science를 위한 IT 인프라에 대한 지속적인 대규모 투자

e-Science를 구축하는데 사용되는 핵심기술 중에 하나가 Grid 인데 Grid라는 것이 유휴 자원을 서로 공유한다는 취지를 담고 있어서 정책적인 오판을 하게 하기 쉽다. 그러나 해외의 사례를 볼 때 “유휴” 자원을 활용한다는 취지보다는 공유할 자원을 “확보”한다는 취지에서 매우 과감한 투자를 해오고 있다. 미국의 TeraGrid 프로젝트의 경우 컴퓨팅 장비, 저장장치, 초고속망 등의 기본 IT 컴퓨팅·네트워크 인프라 구축에만 2001년부터 2004년까지 약 천억대의 예산을 투자한 바 있다. 그 결과 9개의 사이트가 30-40Gb/s의 초고속망으로 연결되었으며 총 40Tflop의 컴퓨팅 파워, 2PB의 저장 공간을 형성하고 있다. TeraGrid는 개방된 과학 연구를 위한 세계에서 가장 크고, 포괄적이면서 분산된 인프라를 개발하고 배포하기 위해서 장기간의 노력의 결과이다 현재 이러한 인프라를 기반으로 해서 다양한 응용연구 환경 등이 구축되고, 대형 응용연구 프로젝트들이 진행되고 있다.

3.2 연구자원의 온라인 콘텐츠화에 대한 동기부여

앞에서 언급한 것처럼 e-Science는 연구자원을 사 이버 공간에서 공동 활용하며 연구를 수행하는 차세대

연구활동으로 보고 있으며 여기서 연구자원이라 함은 국내외에 위치한 연구자, 연구장비, 연구정보 등으로 구성된다. 따라서 연구자의 공유라는 것은 새로운 IT 인프라 위에서 협동 연구가 이루어지는 것으로 볼 수 있고 이미 기존의 인프라(전화, 인터넷 등) 위에서도 이루어지고 있는 일이기 때문에 더 좋은 인프라가 제공된다면 협동연구를 하는 자체에 대해서는 논란이 없는 것 같다. 연구장비의 경우 매우 고가인 것 들이 있는데 이를 공동 사용하는 것 또한 기존의 인프라(도로에 의한 현장 접근이나 전화에 의한 지시 등)에서도 제한적이지만 해오던 일이기도 하다. 하지만 연구정보의 공유에 대해서는 정보의 종류 및 공유의 범위에 따라서 논란이 많다. 이에 대한 해외의 성공 사례가 BIRN (Biomedical Informatics Research Network) 프로젝트이다. 다음은 국내 KTV(한국정책방송)가 BIRN 프로젝트 관계자와 인터뷰 한 내용들이다.

- 제프리 그레테, NCMIR - “연구자로서 제게 BIRN은 많은 국내외 협력을 가능케 해 줍니다. 초고속 네트워크와 방대한 양의 정보 리소스가 있어서 서로 공유하기 용이하며 접속 속도가 빨라서 다른 협력 연구자들이 수집한 연구 자료를 쉽게 접속할 수 있다. 분석 면에서도 BIRN은 계산 자원을 확보해서, 가령 테라그리드를 통해 대용량의 정보를 분석하게 도와 준다. 과거엔 불가능했다.”

- 밀러 박사, 존스 홉킨스 대학 CIS 센터 - “미국에서 연구자 대부분은 개별적으로 연구를 한다. BIRN은 협력 연구의 사례이다. 원래 대부분 연구 지원도 개별적이다. 다른 대학과 잘 협력을 안했던 편이었는데 개별 대학이 연구를 진행했다. 그러나 이제는 변하고 있다. 미국 내 연구가 특정 영역을 넘어서 협력적으로 이루어지기 시작했다.”

실제로 BIRN 프로젝트에서는 연구정보, 특히 가공되지 않은 상태의 데이터 공유가 매우 활발히 일어나고 있다. 하지만 이런 식의 정보공유 및 협력 연구는 매우 강한 동기 부여가 요구된다. 즉 과학데이터나 실험 장비를 실제 IT 인프라 상에서 공유할 수 있도록 온라인 콘텐츠화 하는 것은 많은 시간과 비용이 요구되고, 전통적 과학기술 연구관점에서 연구 성과로 인정받지 못하고 있다. 실제 과학기술 경쟁력 차원에서도 공개하지 않는 것이 더 유리한 측면이 있어서, 개별 연구자나 그룹이 자발적으로 하기는 어렵다. 따라서 국가적 차원에서 정책적 활성화 전략과 지원이 요구된다.

• BIRN 프로젝트의 데이터 공유를 통한 협력연구 사례

BIRN 프로젝트는 미국 NCMIR에서 총괄하는 뇌 과학 연구의 효율성 제고를 위해 바이오-의료과학과 IT 기술을 접목한 연구자간 협력 프로젝트이다. 이 프

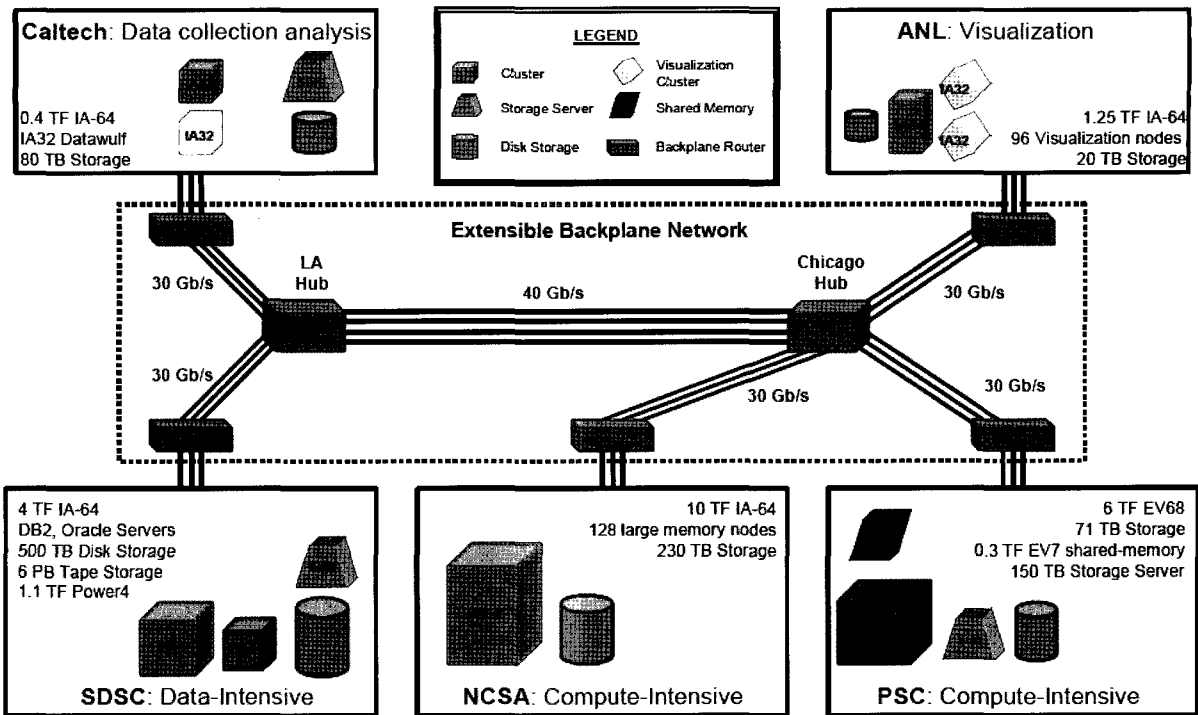


그림 9 TeraGrid의 초고속 네트워크

Enabling collaborative research at 28 research institutions comprised of 37 research groups.

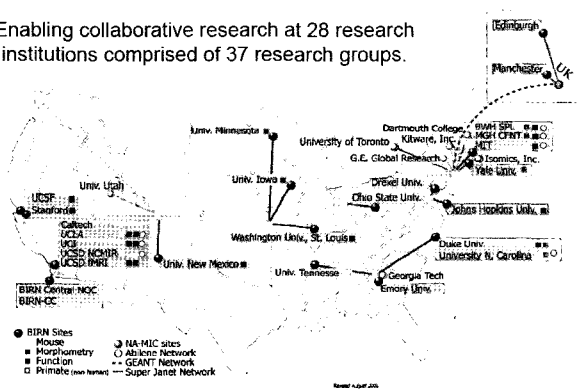


그림 10 BIRN(Biomedical Informatics Research Network) 프로젝트는 연구자간 협력을 위해 테라그리드를 사용한다. 테라그리드는 초고속 네트워크와 방대한 양의 정보를 저장할 수 있다. 테라그리드는 접속 속도가 빨라서 다른 협력 연구자들이 수집한 연구 자료를 쉽게 접근할 수 있어 서로 공유하기 용이하다. BIRN은 이 테라그리드를 위에서 계산 자원을 확보해서 연구자들이 수집한 대용량의 정보를 분석하게 도와준다.

BIRN 프로젝트는 인간의 뇌의 관련된 질병과 형태와의 관계에 대한 연구(Human Structure BIRN), 정신분열증에 대한 이미징 분석(FIRST BIRN), 질병에 걸린 쥐에 대한 다양한 방식의 분석 자료(Mouse BIRN)에 대한 연구를 진행하고 있다. 이 중에서 인간의 뇌에 대한 연구 중에서 알츠하이머에 대한 연구는 인간의 뇌에서 기억을 저장하는 부분인 히포캠포스에

대한 연구가 중심이다. 미국의 존스 홉킨스 대학 병원, 매사추세츠 병원 등 5개 병원에서 환자의 MRI를 촬영하여 수집한다. 수집한 자료중에서 히포캠포스만을 필요로 하기 때문에 수집한 자료는 BIRN 시스템에 저장을 한다. BIRN 시스템에서 다음 연구자가 새로 수집된 MRI 영상으로부터 히포캠포스가 있는 뇌구조를 분리하고, 고화질의 영상을 만든다. 이렇게 만들어진 영상은 다음 연구자를 위해 다시 BIRN 시스템에 저장을 한다. 다음 연구자는 분리된 뇌구조에서 히포캠포스 부분을 찾아 정확히 마킹을 한다. 예전에는 한 대상자당 3-5시간이 걸렸으나, 현재는 15-30분 정도를 BIRN 매커니즘을 통해 준비를 하고 그 자료를 슈퍼컴퓨터로 처리한다. BIRN 시스템과 테라그리드가 연구자간의 협력에 얼마나 도움을 주는지 알 수가 있다. BIRN 시스템에 새로 처리된 정보를 이용하여 매사추세츠 병원 연구진에 의해 히포캠포스를 분리하고, 이를 통해 가공된 정보는 존스 홉킨스 CIS 센터로 전달되어 분석한다. 분석된 정보는 다시 BIRN 시스템에 저장된다. 다음 연구자는 새로 분석된 데이터의 알츠하이머의 여부를 판단한다. 통계치를 통해 정상 그룹으로 표시되는 1 그룹과 알츠하이머로 표시되는 2 그룹을 비교하여 새로운 데이터가 어느 쪽에 가까운지 판단한다. 이렇게 수집되어진 거대한 정보는 다시 병원에서 활용되며, 질병을 조기에 진단하고 질환이 더 악화되기 전에 막을 수 있도록 활용된다.

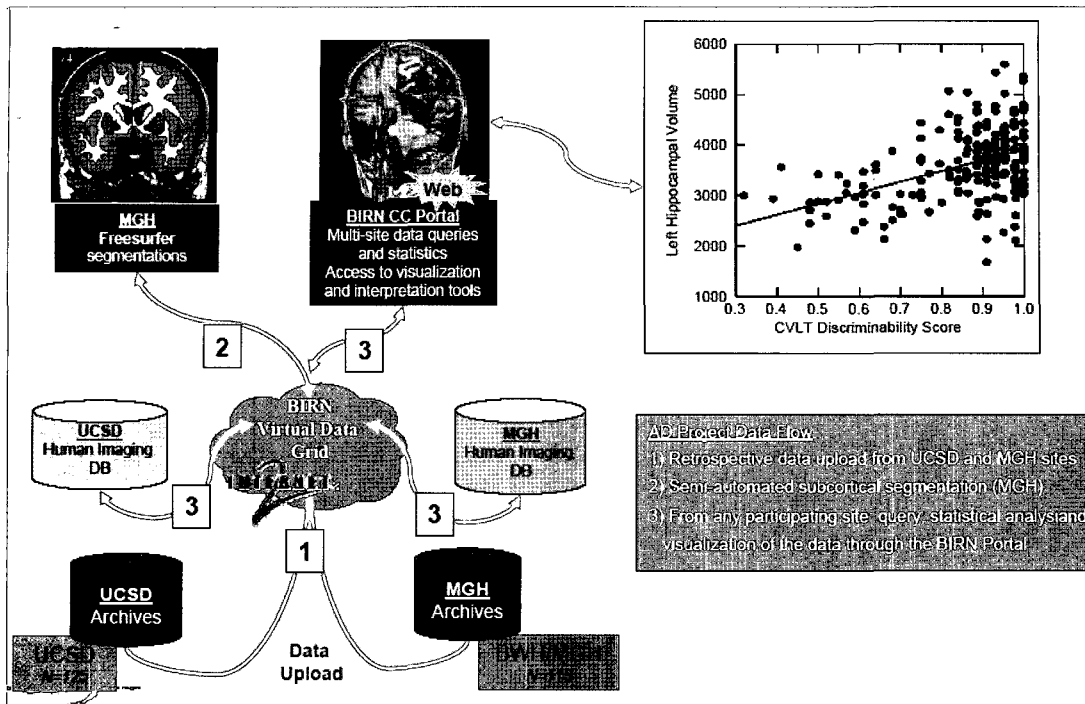


그림 13 BIRN에서 알츠하이머 연구 과정

4. 결 론

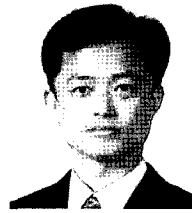
e-Science 성공의 핵심은 응용연구자와 IT 연구자가 긴밀한 결합을 통해 재사용이 가능한 환경을 구축하는 것과 이런 환경 위에서 광범위한 연구자원의 공유를 통해서 새로운 패러다임의 연구가 진행되는 것이다. 해외 사례를 살펴본 것처럼 각국에서는 이미 e-Science를 차세대 연구 환경으로 인식하고 적극적인 투자를 통해 자국의 연구 환경을 개선시키고자 노력하고 있다. 이를 통해서 지역 간, 연구 분야 간, 자원간의 제약 없는 협업이 가능하도록 하여, 연구생산성을 획기적으로 개선하고자 하는 것이다. 국내의 경우 2002년부터 e-Science를 위한 기반기술에 대한 연구 개발은 이루어져왔다. 하지만 새로운 협력연구 패러다임을 통한 IT기반의 혁신이라는 관점에서 볼 때는 아직 해결해야 할 문제가 많고 무엇보다도 e-Science 개념에 대한 인식의 전환이 필요하다. 즉 e-Science 구축 과정을 고속도로 건설과 같은 사회간접자본에 대한 투자의 개념으로 보아야한다는 것과 e-Science 구축이라는 것은 그 자체가 결과물이 아니라 새로운 패러다임의 서비스의 시작이라는 것, 그리고 이 혁신적인 서비스를 성공적으로 제공하기 위해서는 연구 자원 특히 연구 데이터에 대한 공유 개념 도입이 절대적으로 필요하다는 것이다. 앞에서 설명한 TeraGrid와 BIRN 프로젝트는 이에 대한 확실한 교훈을 주는 해외 사례로 볼 수 있다.

참고문헌

- [1] Daniel E. Atkins 등, Revolutionizing Science and Engineering Through Cyberinfrastructure(2003).
- [2] 과학기술부, e-Science 구축방안 연구(2003).
- [3] 국가과학기술자문회의, e-Science 국내 연구환경 영향평가 및 추진타당성 조사연구(2004).
- [4] UK e-Science, <http://www.rcuk.ac.uk/escience>
- [5] e-Science Gap Analysis, UK e-Science (2003).
- [6] The Data Deluge: An e-Science Perspective, UK e-Science(2003).
- [7] The UK e-Science Core Programme and the Grid(2002).
- [8] Multi-Site Videoconferencing for the UK e-Science Programme(2002).
- [9] White Paper: Excellence and Opportunity? - a science and innovation policy for the 21st century, <http://www.ost.gov.uk/enterprise/dtiwhite/index.html>
- [10] Engineering and Physical Sciences Research Council (EPSRC), <http://www.epsrc.ac.uk/>
- [11] Particle Physics and Astronomy Research Council (PPARC), <http://www.pparc.ac.uk/>

- [12] National e-Science Centre (NESC), <http://www.nesc.ac.uk/>
- [13] GridPP, UK Computing for Particle Physics, <http://www.gridpp.ac.uk/>
- [14] AstroGrid Project, <http://www.astrogrid.org/>
- [15] eDiaMoND grid computing project, <http://www.ediamond.ox.ac.uk/>
- [16] TeraGrid, <http://www.teragrid.org/>
- [17] Biomedical Informatics Research Network (BIRN), <http://www.nbirn.net/>
- [18] The Geosciences Network (GEON), <http://www.geongrid.org/>
- [19] Earth Scope Project, <http://www.earthscope.org/>
- [20] National Ecological Observatory Network (NEON), <http://www.neoninc.org/>
- [21] Laboratory for the Ocean Observatory Knowledge Integration Grid (LOOKING), <http://lookingtosea.ucsd.edu/>
- [22] Network for Earthquake Engineering Simulation (NEES), <http://www.nees.org/>
- [23] OPTical networking, Internet Protocol, compUTER storage processing and visualization technologies (OptIPuter), <http://www.optiputer.net/>
- [24] Linked Environments for Atmospheric Discovery (LEAD), <http://lead.ou.edu/>
- [25] K. Baxevanidis, "The European Union policy on the Grid - prospects for the future," (2001)
- [26] Enabling Grids for e-Science in Europe (EGEE), <http://www.eu-egee.org/>
- [27] High Energy Physics Information Center (HEP), <http://www.hep.net/>
- [28] A Large Ion Collider Experiment (ALICE), <http://aliceinfo.cern.ch/>
- [29] SIMDAT, <http://www.simdat.org/>
- [30] Information Technology Based Laboratory (ITBL), <http://www-riken.riken.jp/>
- [31] Biogrid Project, <http://www.biogrid.jp/>
- [32] 과학기술부, 국가 e-Science 구축사업 기본계획 (2005)

황 선 태



1985 서울대학교 컴퓨터공학과(학사)
 1987 서울대학교 컴퓨터공학과(석사)
 1996 Manchester University(PhD)
 1997~현재 국민대학교 컴퓨터학부 부교수
 관심분야: e-Science, 그리드시스템, PSE,
 공개소프트웨어
 E-mail : sthwang@kookmin.ac.kr

허 대 영



2004 국민대학교 컴퓨터학부(학사)
 2005 국민대학교 전산과학(석사)
 2006~현재 국민대학교 전산과학 박사과정
 관심분야: 그리드 시스템, 시스템 아키텍처,
 디자인 패턴
 E-mail : dyheo@cs.kookmin.ac.kr

정 갑 주



1984. 2 서울대학교 컴퓨터공학과(학사)
 1986. 2 서울대학교 컴퓨터공학과 인공
 지능(석사)
 1996. 2 New York University, Com-
 puter Science(박사)
 1995. 12~1997. 8 University of
 Florida, Post Doc.
 1997. 8~2001 건국대학교 컴퓨터공학과
 조교수
 2001~현재 건국대학교 인터넷미디어공학부 부교수
 관심분야: Grid Computing & e-Science, Data Integration
 E-mail : jeongk@konkuk.ac.kr

조 금 원



1993 인하대학교 항공우주공학과(학사)
 1995 한국과학기술원 항공우주공학과
 (석사)
 2000 한국과학기술원 항공우주공학과
 (박사)
 현 재 KISTI 슈퍼컴퓨팅센터 슈퍼컴퓨팅
 응용지원팀장, KISTI e-Science
 사업단응용연구팀장(겸임)
 관심분야: 슈퍼컴퓨팅, e-Science, 그리드
 E-mail : ckw@kisti.re.kr

이 상 동

2000 부산대학교 물리학(박사)
 현 재 한국과학기술정보연구원 슈퍼컴퓨팅사업실
 E-mail : sdlee@kisti.re.kr

이 지 수



1985 서울대학교 물리학과(학사)
1986 피츠버그버그대학 물리학과(석사)
1992 보스턴대학교 물리학과(박사)
현 재 KISTI 슈퍼컴퓨팅센터장,
e-Science 사업단장(겸임)
관심분야: 고성능 컴퓨팅, 고성능 네트워킹,
그리드, 고학기술 연구인프라
E-mail : jysoo@kisti.re.kr
