

Tiers and GRID computing



2012. 8. 10

김민석 (성균관대)

The higgs, the history, and the grid

28 years since the idea of the LHC

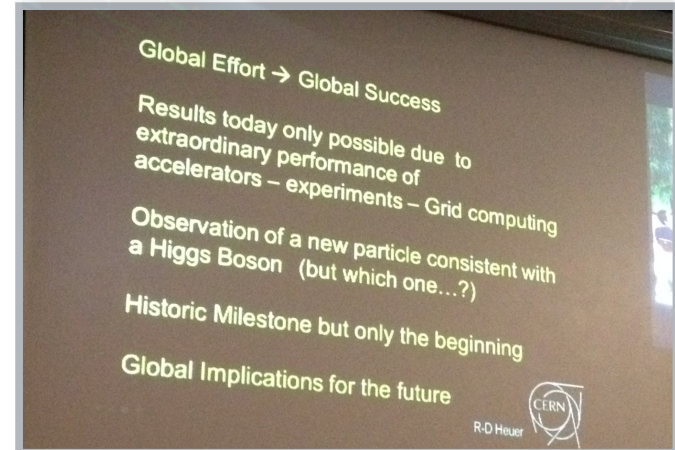
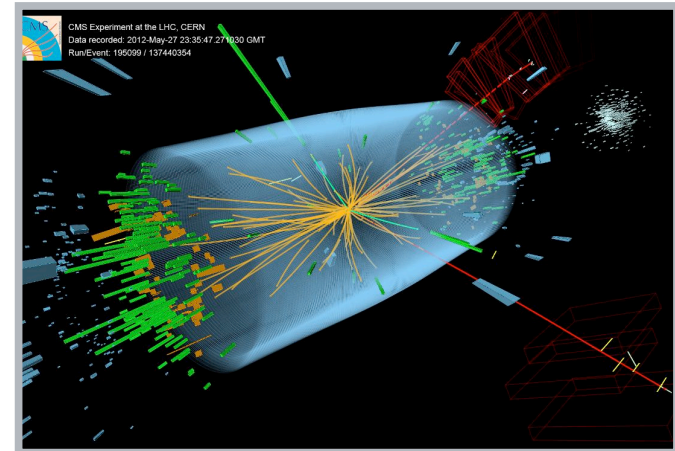
thousands of people worldwide

thousands of computers worldwide

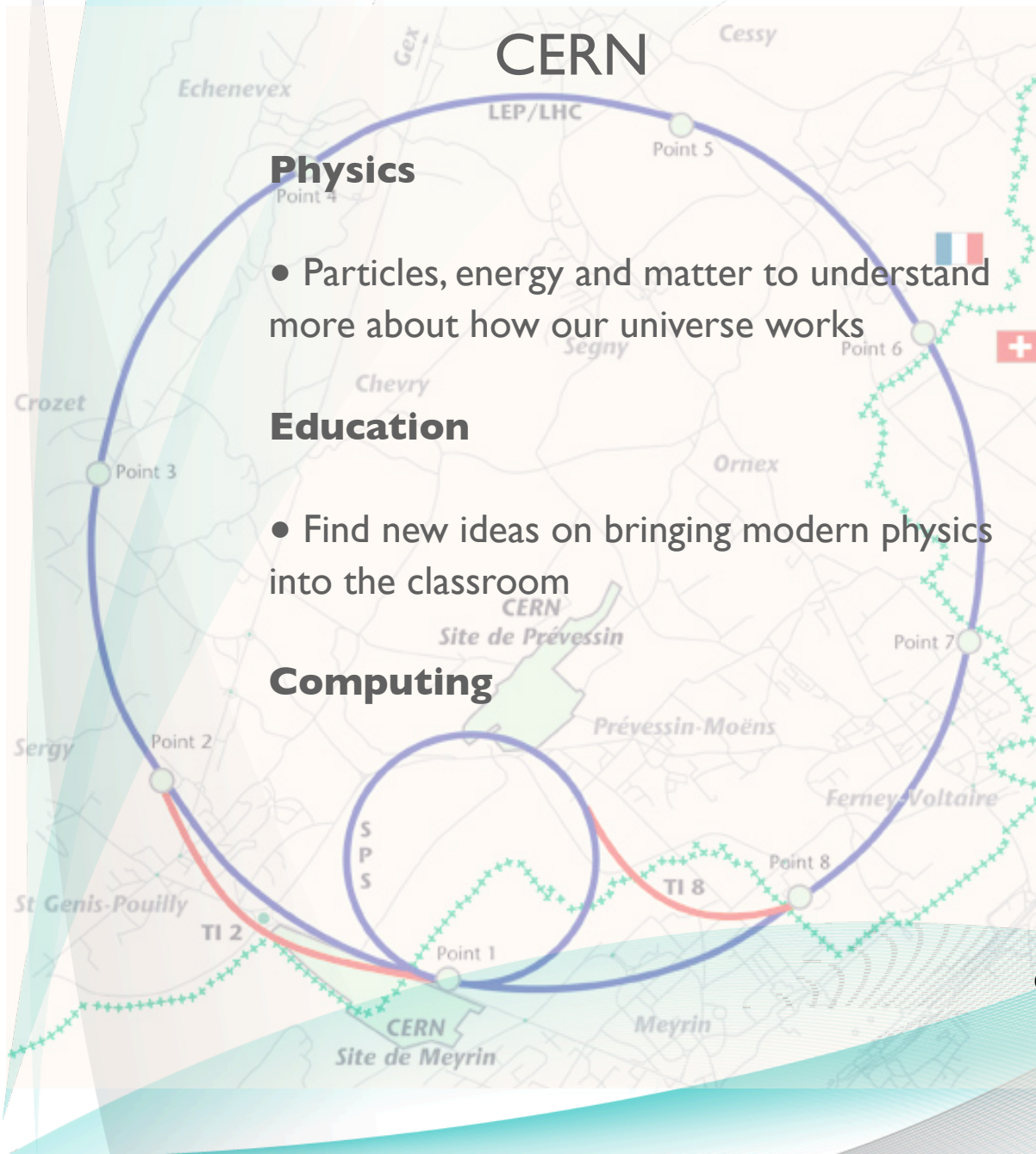
Global Effort → Global Success

Results today only possible due to
extraordinary performance of

**accelerators - experiments - Grid
computing**



*July 4, 2012: a final comment
from Rolf Heuer,
the director general of CERN*



Dec 25, 1990: the first successful communication between a Hypertext Transfer Protocol (HTTP) and server via the Internet by Tim Berners-Lee, the inventor of the **World Wide Web**

CERN in our universe

Largest and biggest

- 세계에서 가장 큰 기계

Fastest

- 가장 빠른 레이스트랙

Hottest spots

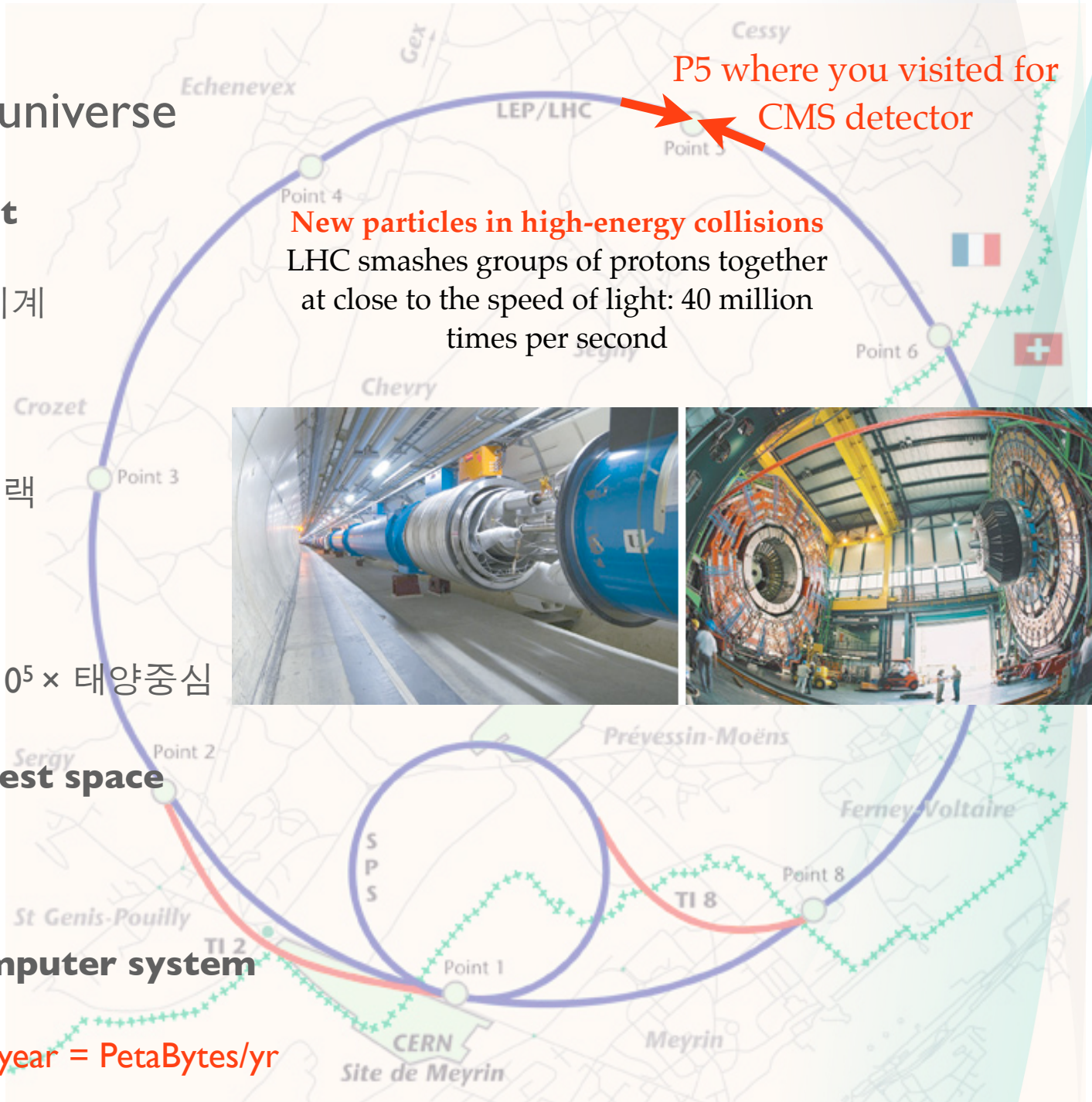
- 양성자충돌온도 = $10^5 \times$ 태양중심

Coldest and emptiest space

- 영하 271.2도 (1.9 K)

Most powerful computer system

- 10^5 dual layer DVDs/year = PetaBytes/yr



Computing for Physics Research

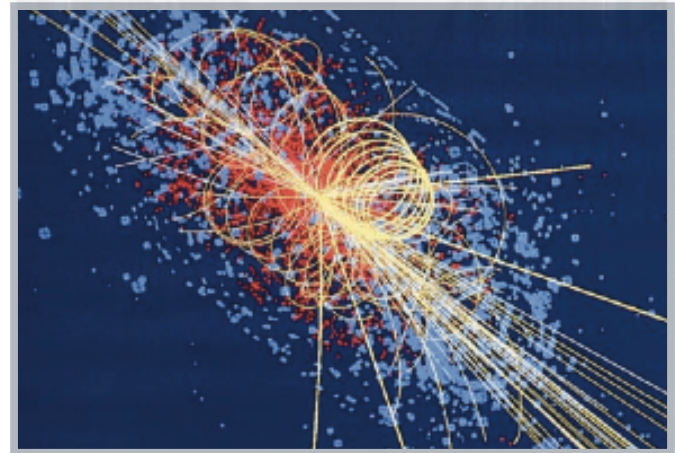
Data Technologies

- Data storing and management
distributed, parallel and cloud computing network

called the Grid

Data Analysis

- Algorithm and tools
simulation, reconstruction and visualization





Computing system

Grids (a super virtual computer)

- Have computers connected to a network by a conventional **network interface**, such as Ethernet
- Geographically distributed computing
- Parallel computing for computation

Worldwide LHC Computing Grid (WLCG)

A supercomputer

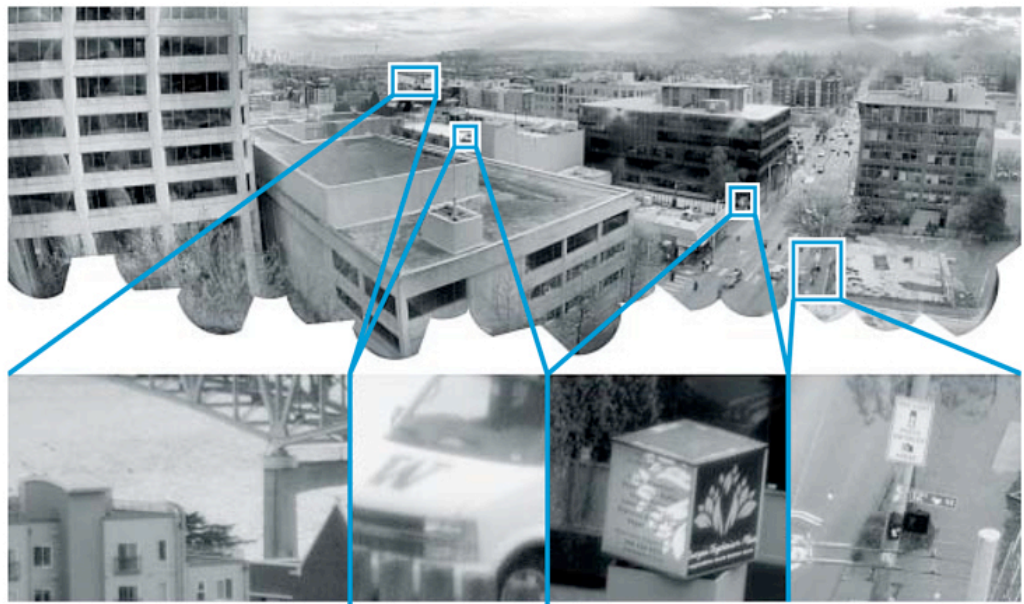
- Has many processors connected by a local-speed **computer bus**, which is a subsystem that transfers data between components

Europe



10억화소... '천리안' 카메라 나왔다

크기는 최소·촬영범위는 최대... 듀크대팀 "5년 뒤 상용화"

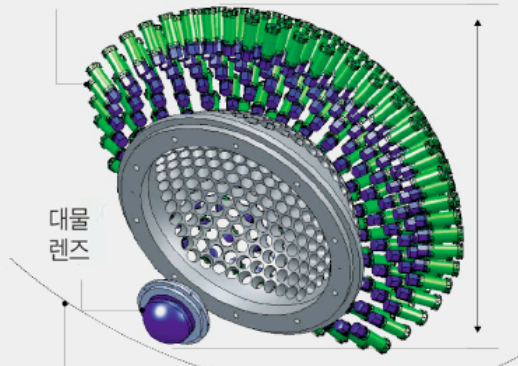


수백m 밖 건물의 세부적인 모습 포착
 시력이 1.0인 사람보다 5배 뛰어난 풍경 인식 기능
 서로 다른 거리에 있는 입간판이나 표지판도 크게 확대하면 선명하게 나옴
 듀크대 제공

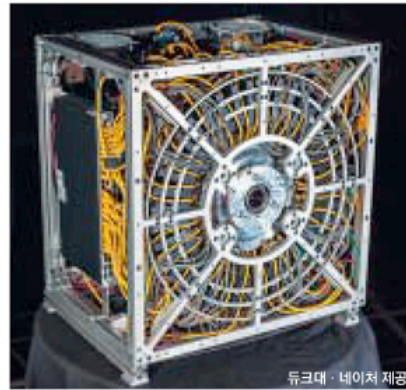
10⁹ 기가픽셀 카메라

1400만 화소짜리 마이크로카메라
 98개가 반구(半球)형으로 배치

프레임(받침대)
 을 제외한
 카메라 높이 66cm



좌우 120도, 상하 50도 범위 와이드
 (광각·廣角) 촬영 기능
 전체 무게는 45kg
 한 장면 촬영과 저장에 걸리는 시간은 18초
 그래픽=조선일보 디자인편집팀



1400만 화소 카메라, 98개 결합
 가로·세로 75cm '기존 절반 크기'
 거리 사진 찍어 확대하면
 사람·간판 등 또렷하게 보여

천리안카메라

- 무게 45kg, 가로/세로 75 cm, 높이 50 cm 로 일반 디지털카메라보다 100배 이상 선명한 10억화소급 소형 카메라: 크기는 최소-촬영범위는 최대

거리 사진 찍어 확대하면 사람/
간판 등 또렷하게 보여

- 가격: 10만달러
- 도심촬영시 800m밖 우표구별, 광활한 지역 생태 연구시 장소 옮길 필요없음
- 데이터 처리 한계 때문에 흑백사진만 찍을 수 있음

천리안카메라 vs. CMS 검출기

- 무게 45kg, 가로/세로 75 cm, 높이 50 cm 로 일반 디지털카메라보다 100배 이상 선명한 10억화소급 소형 카메라: 크기는 최소-촬영범위는 최대

거리 사진 찍어 확대하면 사람/간판 등 또렷하게 보여

- 가격: 10만달러
- 도심촬영시 800m밖 우포구별, 광활한 지역 생태 연구시 장소 옮길 필요없음
- 데이터 처리 한계 때문에 흑백사진만 찍을 수 있음

- 무게 14,000톤, 길이 22 m, 직경 15 m로 지구 자기장의 100,000배 이상 큰 자기장을 이용한 초대형 입자 검출기: 크기는 최대-검출기시스템은 콤팩트 (초대형 초전도 솔레노이드 전자석안에 설치됨)

강력한 자기장을 이용하면 입자의 궤적(운동량)을 정확하게 측정

- 가격: 모름 (LHC~100억달러)
- 우주생성의 비밀(새물리)을 밝히고 세상을 바꿀 새로운 기술도 개발
- CERN연구소내 데이터 처리 한계

Tier sites

- Tier-0: CERN Computer Centre

- Tier-1 (11 sites)

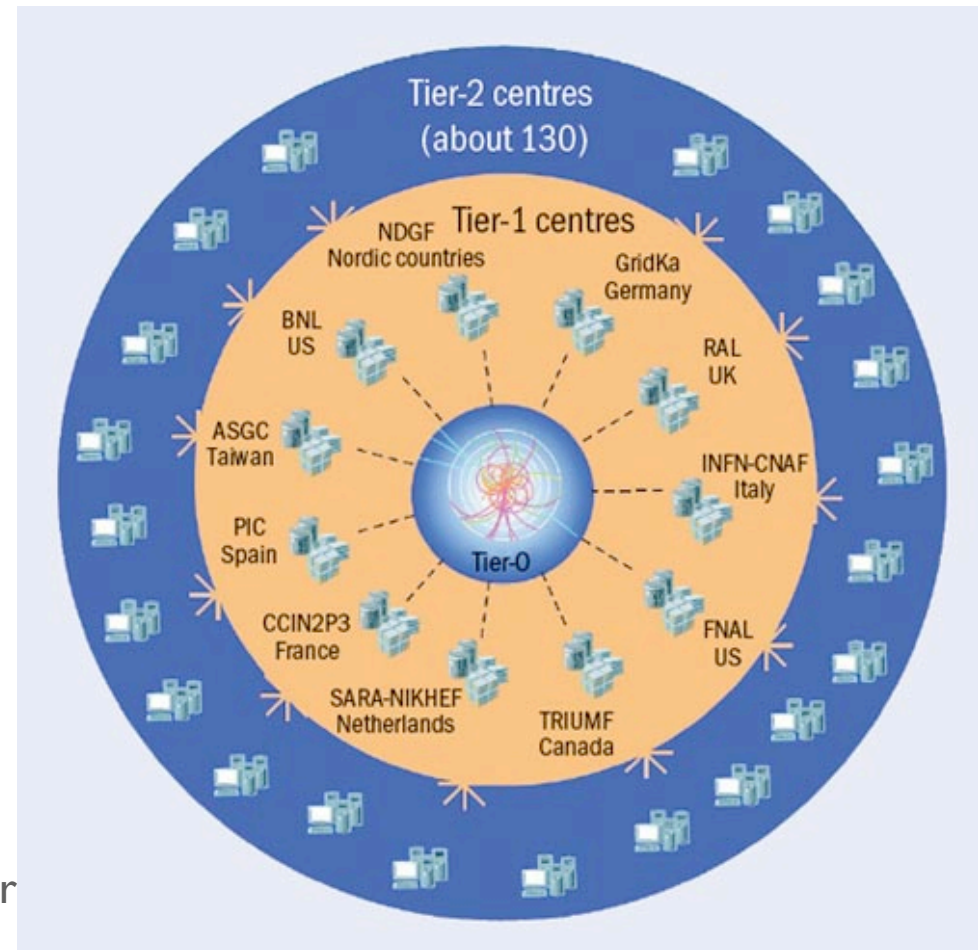
| | |
|------------------|------------------------------------|
| Canada | TRIUMF |
| Germany | KIT |
| Spain | Port d'Informació Científica (PIC) |
| France | IN2P3 |
| Italy | INFN |
| Nordic countries | Nordic Datagrid Facility |
| Netherlands | NIKHEF / SARA |
| Taipei | ASGC |
| United Kingdom | GridPP |
| USA | Fermilab-CMS |
| USA | BNL ATLAS |

- Tier-2 (140 sites)

universities and other scientific institutes

- Tier-3 (lots)

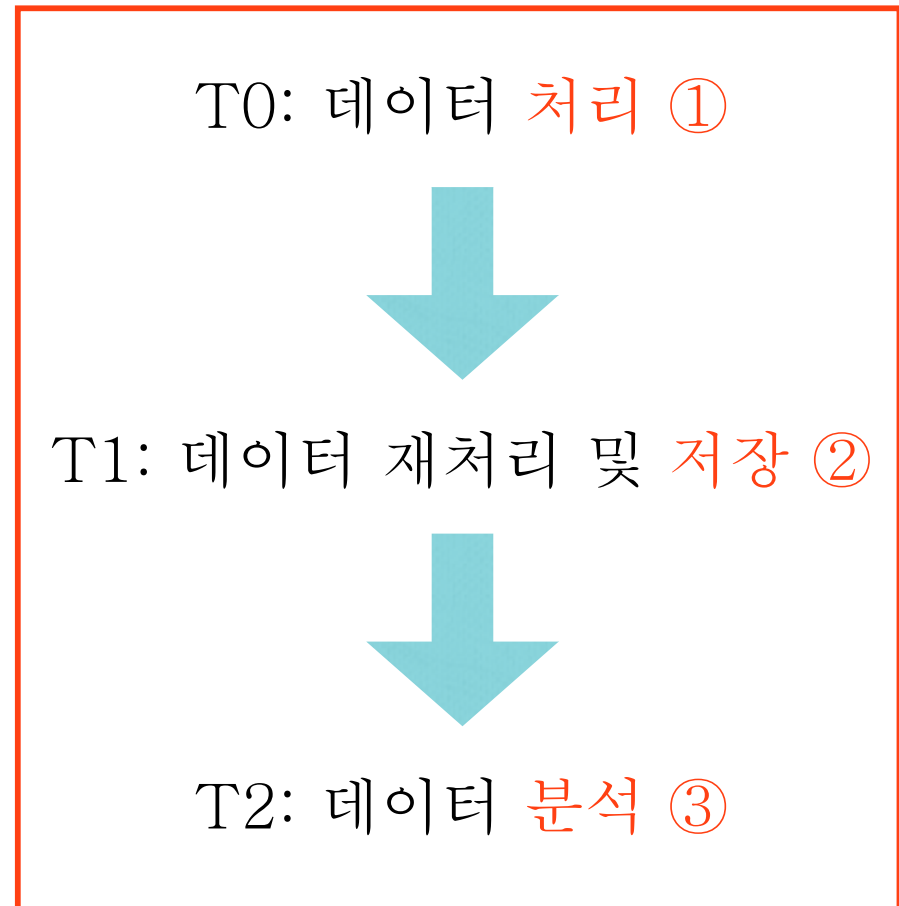
local clusters in a university department or an individual PC



Why tiered?

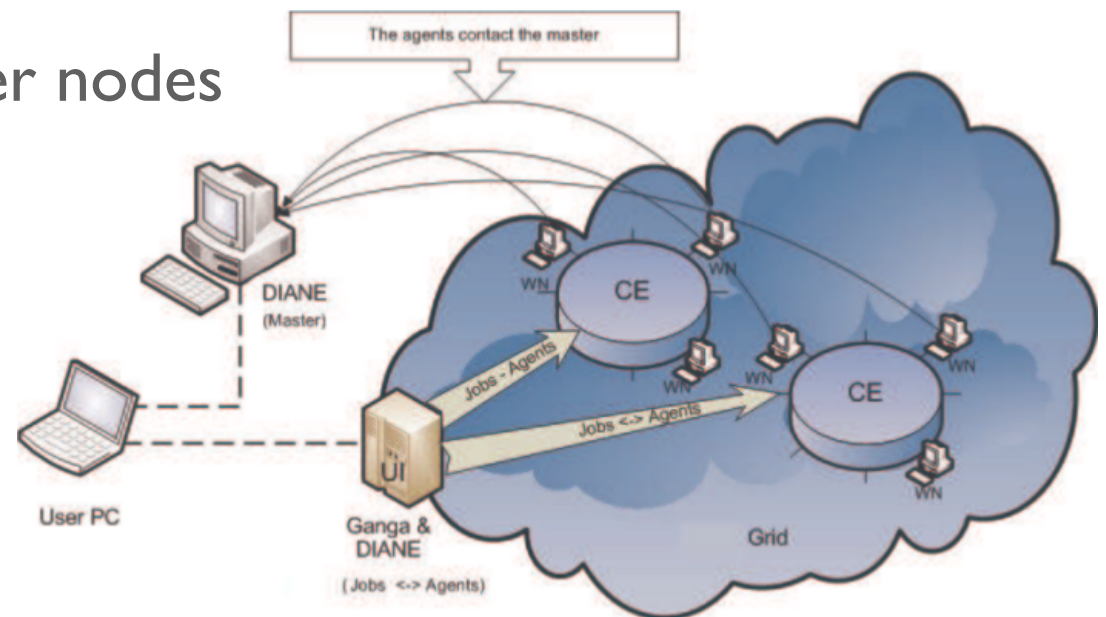
Tier functions

- **Tier-0: CERN Computer Centre**
 - first accepts RAW data
 - repacks the RAW data into primary datasets
 - archives the RAW data to tape
 - distributes into T1 (i.e, two copies)
 - prompt calibration to get calibration constants
 - prompt first pass reconstruction (RECO data)
 - distributes into T1
- **T1**
 - re-reconstruction, skimming, calibration
 - distributes into other T1, CERN, T2
 - stores data
- **T2**
 - grid-based analysis
 - Monte Carlo (MC) simulation



Tier services by middleware

- Storage Element (SE)
Data access protocols & interfaces
- Computing Element (CE)
Job-manager and worker nodes
- Security
- User interface
- Information service
- Workload management
- Software components



데이터처리

데이터저장



데이터분석

데이터용량

It's time to talk about...

Powers of ten

The powers of ten are commonly used in physics and information technology. They are practical shorthand for very large or very small numbers.

| Power of ten | Number | Symbol |
|--------------|-----------------------|---------------|
| 10^{-12} | 0.000000000001 | p (pico) |
| 10^{-9} | 0.000000001 | n (nano) |
| 10^{-6} | 0.000001 | μ (micro) |
| 10^{-3} | 0.001 | m (milli) |
| 10^{-2} | 0.01 | |
| 10^{-1} | 0.1 | |
| 10^0 | 1 | |
| 10^1 | 10 | |
| 10^2 | 100 | |
| 10^3 | 1000 | k (kilo) |
| 10^6 | 1 000 000 | M (mega) |
| 10^9 | 1 000 000 000 | G (giga) |
| 10^{12} | 1 000 000 000 000 | T (tera) |
| 10^{15} | 1 000 000 000 000 000 | P (peta) |

Volume of data
produced at LHC

천
백만
십억
조
천조

Powers of ten

The powers of ten are commonly used in physics and information technology. They are practical shorthand for very large or very small numbers.

| Power of ten | Number | Symbol |
|--------------|-----------------------|---------------|
| 10^{-12} | 0.000000000001 | p (pico) |
| 10^{-9} | 0.000000001 | n (nano) |
| 10^{-6} | 0.000001 | μ (micro) |
| 10^{-3} | 0.001 | m (milli) |
| 10^{-2} | 0.01 | |
| 10^{-1} | 0.1 | |
| 10^0 | 1 | |
| 10^1 | 10 | |
| 10^2 | 100 | |
| 10^3 | 1000 | k (kilo) |
| 10^6 | 1 000 000 | M (mega) |
| 10^9 | 1 000 000 000 | G (giga) |
| 10^{12} | 1 000 000 000 000 | T (tera) |
| 10^{15} | 1 000 000 000 000 000 | P (peta) |

Volume of data produced at LHC

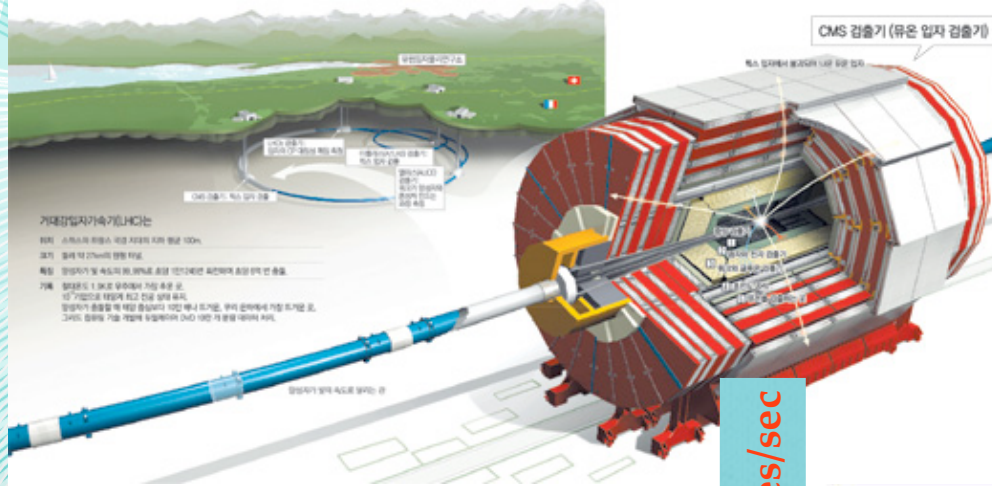
Size of interest

10^{-15}

천조분의일

f (femto)

천
백만
십억
조
천조



LHC / CMS Tiered Data System

PBytes/sec

MBytes/sec

PBytes/sec

CERN/Outside Resource Ratio ~1:2
 Tier0/(Σ Tier1)/(Σ Tier2) ~1:1:1

Online System

~200-1500 MBytes/sec

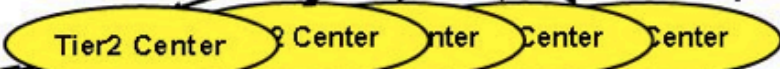
Tier 0 + 1

CERN Center
 PBs of Disk;
 Tape Robot

Tier 1 10 - 40 Gbps



Tier 2



Tier 3 1-10 Gbps



Physics data cache

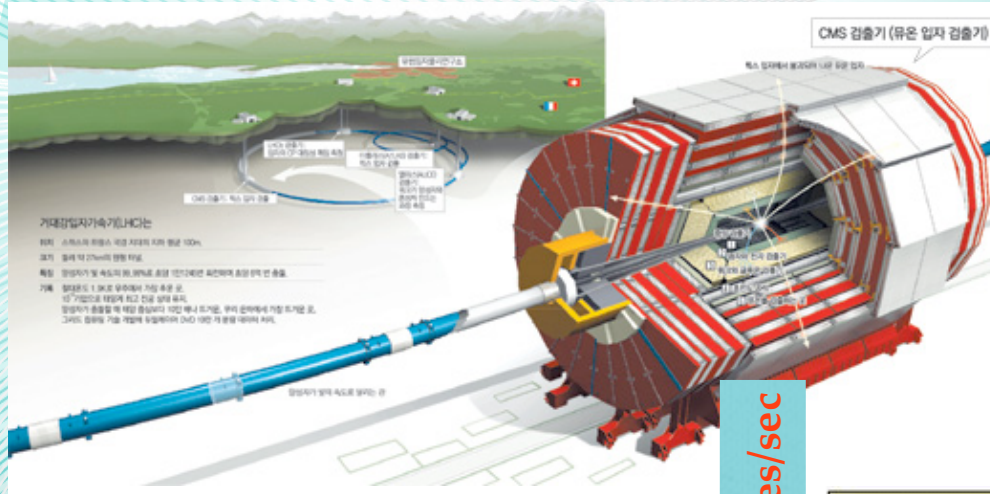
1 to 10 Gbps

Workstations

Tier 4

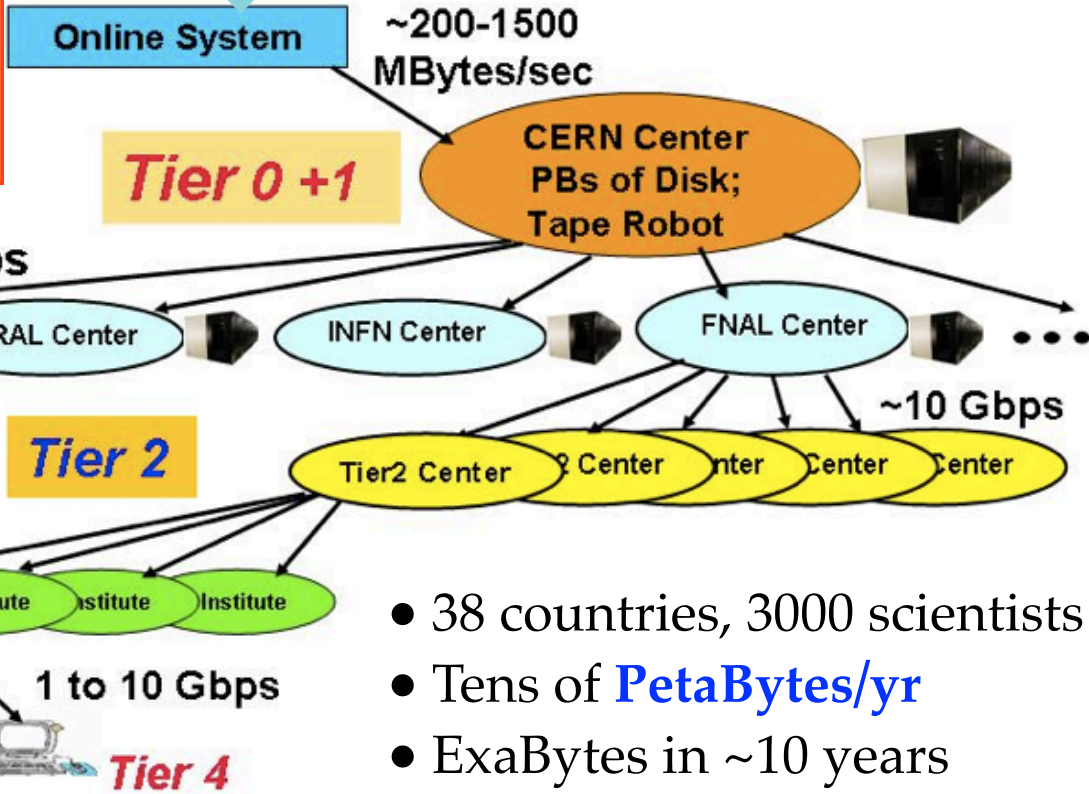
- 38 countries, 3000 scientists
- Tens of PetaBytes/yr
- ExaBytes in ~10 years

LHC / CMS Tiered Data System



PBytes/sec
 ↓ trigger
 MBytes/sec

CERN/Outside Resource Ratio ~1:2
 Tier0/(Σ Tier1)/(Σ Tier2) ~1:1:1



- 38 countries, 3000 scientists
- Tens of **PetaBytes/yr**
- ExaBytes in ~10 years

GBytes per second

● A record of data on backup tape with a transfer rate of 1.1 GB/s for several hours (백업속도) =
Recording a movie on DVD every 4 s

● A record of data transfer over 10,000 km between CERN and California, with a throughput of 2.38 GB/s for over an hour (전송속도) =
Sending 200 DVD films in an hour



Storage: CASTOR → EOS

- CERN stores 12 million new files per month in the CASTOR. Expected to increase 3 times until 2015 reaching 0.125 EB. ~2000 storage servers are used to store experiment/user data
- EOS project started in Apr 2010 to provide fast and reliable disk only storage technology

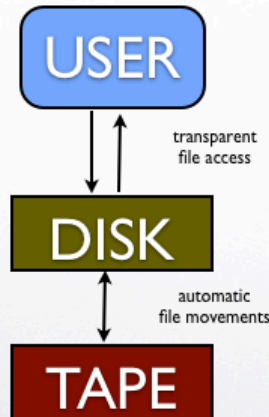
Exabyte Scale

Pessimistic calculation assuming 1 MB file size

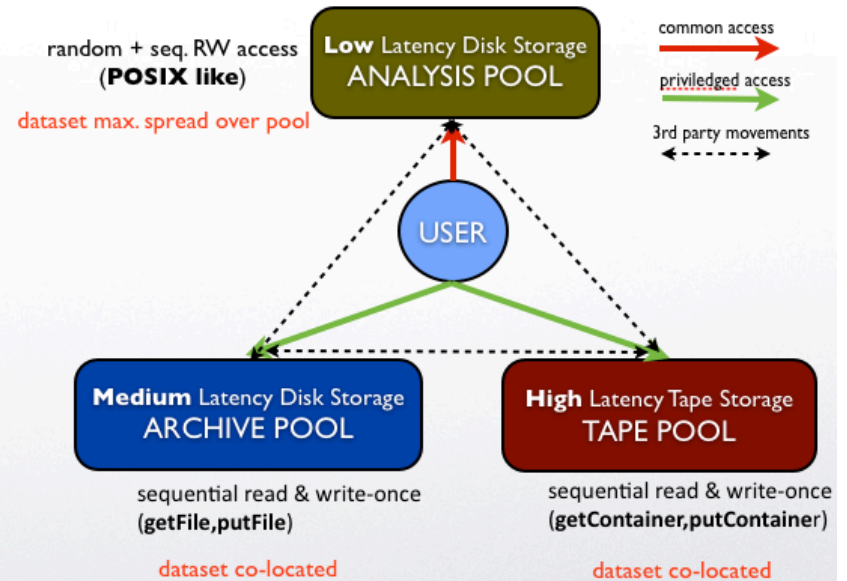
| | Access Latency [s] | Files | File Container | Volume [bytes] |
|---------------|---------------------|------------------------------|-------------------------------|----------------------------------|
| Analysis Pool | $10^{-3} - 10^{-2}$ | $10^9 - 10^{10}$ Billions | $10^6 - 10^7$ Millions | $10^{15} - 10^{16}$ Petabytes |
| Archive Pool | $10^{-2} - 10^1$ | $10^{11} - 10^{12}$ | $10^8 - 10^9$ 100 Million+ | $10^{17} - 10^{18}$ Exabytes |
| Tape Pool | $10^1 - 10^3$ | $10^{11} - 10^{12}$ | $10^8 - 10^9$ 100 Million+ | $10^{17} - 10^{18}$ Exabytes |

- 2×10^9 events/yr
- 1 event = 1.6 MB
- 3.2 PB/yr from CMS
- 8.5 GB DVD
- DVD thickness = 1.2 mm
- How many DVDs?
- $3.2\text{PB} / 8.5\text{GB} \approx 40\text{k DVDs}$
- How tall?
- $40\text{k DVDs} \times 1.2\text{mm} \approx 500\text{m}$

HSM Model CASTOR2



Tier Model



Tier 2 in Korea

● CMS
경북대 (KNU)

● ALICE
한국과학기술정보연구원
(KISTI)



세계 최고속 110Gbps 연구망 구현

KISTI·경북대
대전 = 박희범기자 hbpark@etnews.co.kr

국제 연구망이 개통된 이래 최고인 110Gbps 속도가 미국 주도로 국내에서 달성됐다. 이 속도는 세계 최고 수준이다. 기초기술연구의 산하 한국과학기술정보연구원(KISTI·원장 박영서)과 경북대학교(총장 노동일)는

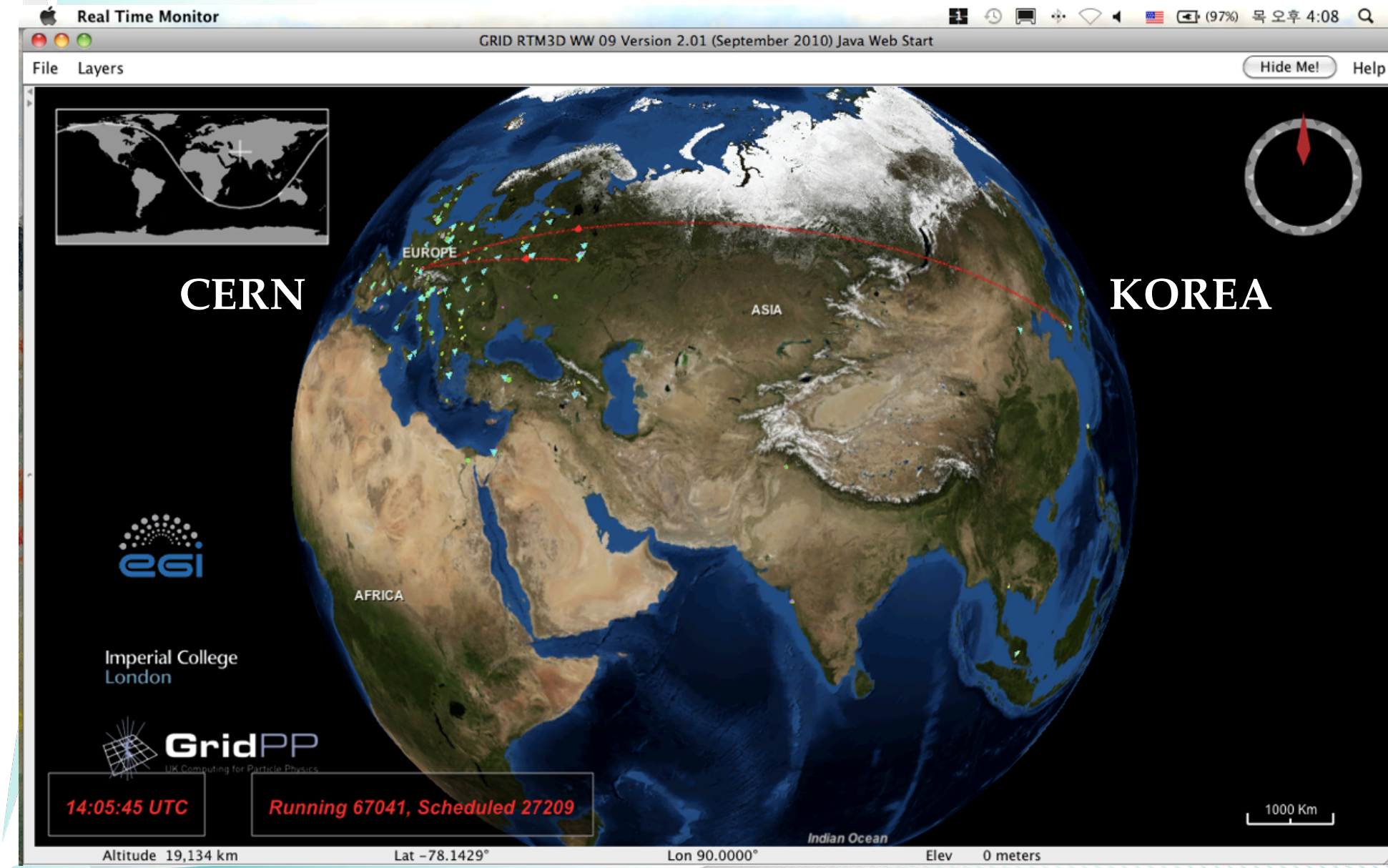
지난 달 열린 '슈퍼컴퓨팅 콘퍼런스 2009'에서 고에너지물리분야의 벤드위스 챌린지(고성능 데이터 전송 시험)를 개최한 결과 이 같은 세계 신기록이 나왔다고 8일 밝혔다. 고에너지 물리 밴드위스 챌린지에는 미국의 고에너지물리 연구기관 칼텍(Caltech)의 주도 아래 KISTI와 경북대 등이 공동 참여했다. 시험에 활용한 글로벌과학기술 협업연구망(GLORIAD)은 전 세계

11개국이 지구를 10기가용 네트워크로 연결한 세계 첫형 국제 연구망이다. 경북대 손동철 고에너지연구소장은 "유럽원자핵공학(CERN)의 대형강입자충돌기(C)가 재가동되기 하루 전됐다"며 "향후 글로벌과업연구망과 슈퍼컴퓨팅 센터를 지속적으로 연계해 나가고자 한다"고 말했다.

신문 2009년 2월 9일(수)

GRID Real Time Monitoring

<http://rtm.hep.ph.ic.ac.uk/webstart.php>





Bonus: Distributed analysis

Car Driving

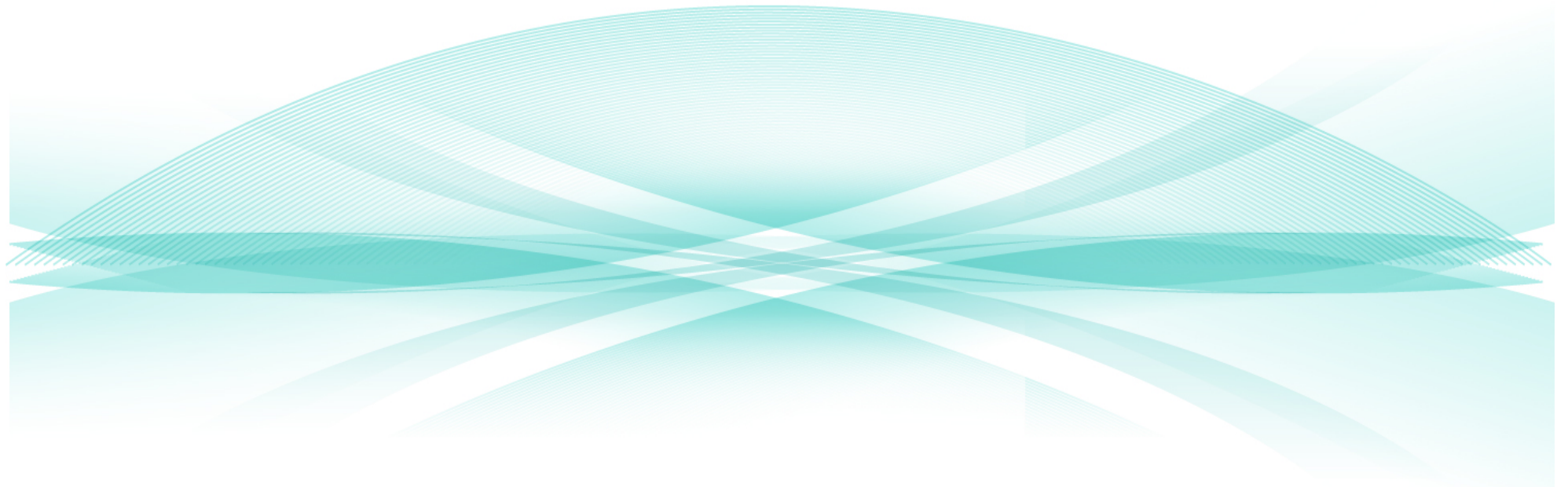
시동(출발)

driving

감속(정지)

GPS

Data analysis



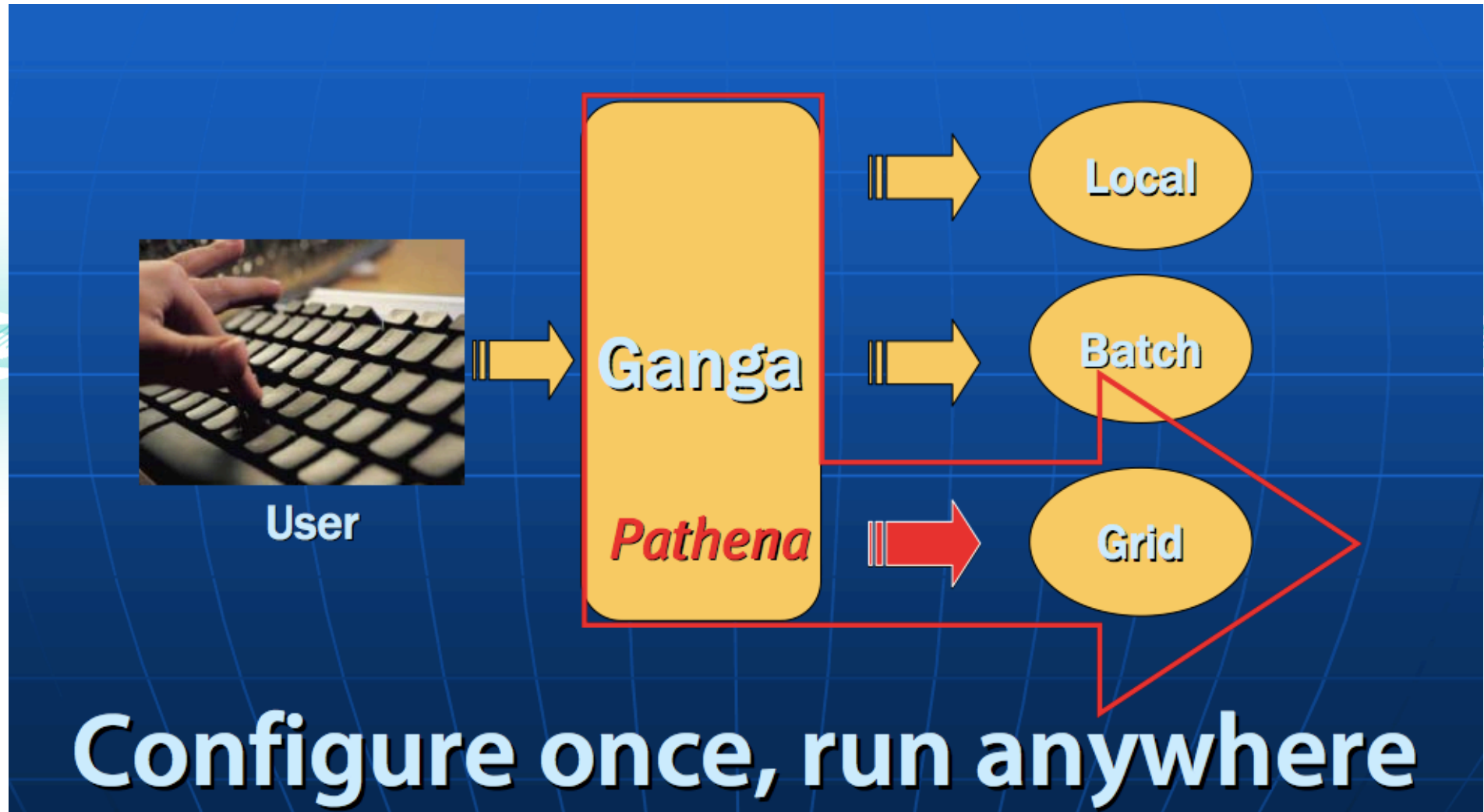
분석(시작)

analysing

분석(끝)

Configuration

Distributed analysis





<http://www.gridcafe.org/>



Interesting! 새로운 물리현상 발견 vs. 미래 인터넷 환경 변화
과학자들간의 정보교환을 위해 개발된 기술 → 생활의 혁명

Thanks 2012

