Tiers and GRID computing

2011. 8. 12

김 민 석 (성균관대)
CERN

Physics

- Particles, energy and matter to understand more about how our universe works

Education

- Find new ideas on bringing modern physics into the classroom

Computing
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

cern
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

---

해석:

- **Tier-0**: CERN 컴퓨터 센터
  - 최초로 RAW 데이터를 수락
  - RAW 데이터를 기본 데이타 세트로 패키징
  - RAW 데이터를 밸리에 보관
  - T1 (즉, 두 복제본)로 분배
  - 즉시 캘리브레이션을 통해 캘리브레이션 상수를 얻는다
  - 즉시 첫 번째 통과 재구성 (RECO 데이터)
  - T1로 분배

- **T1**
  - 재구성, 스크밍, 캘리브레이션
  - 다른 T1, CERN, T2로 분배
  - 데이터 저장

- **T2**
  - 그리드 기반 분석
  - 몽케로 (MC) 시뮬레이션
It's time to talk about...
<table>
<thead>
<tr>
<th>질문</th>
<th>답</th>
</tr>
</thead>
<tbody>
<tr>
<td>블랙박스 색깔?</td>
<td>블랙</td>
</tr>
<tr>
<td>재래기 침속도?</td>
<td>시속80km이상</td>
</tr>
<tr>
<td>LHC 데이터용량?</td>
<td>많다</td>
</tr>
</tbody>
</table>
LHC 데이터용량?

① 억수로 많다
② 아따 거시기 겁나게 많다
③ 오만 군데 천지빼까리다
CERN in our universe

**Largest and biggest**
- 세계에서 가장 큰 기계

**Fastest**
- 가장 빠른 레이스트랙

**Hottest spots**
- 양성자충돌돌변도 = $10^5 \times$ 태양중심

**Coldest and emptiest space**
- 영하 271.2도 (1.9 K)

**Most powerful computer system**
- $10^5$ dual layer DVDs/year = PetaByte/yr
# 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.

<table>
<thead>
<tr>
<th>Power of ten</th>
<th>Number</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^{-12}$</td>
<td>0.0000000000001</td>
<td>p (pico)</td>
</tr>
<tr>
<td>$10^{-9}$</td>
<td>0.000000001</td>
<td>n (nano)</td>
</tr>
<tr>
<td>$10^{-6}$</td>
<td>0.000001</td>
<td>μ (micro)</td>
</tr>
<tr>
<td>$10^{-3}$</td>
<td>0.001</td>
<td>m (milli)</td>
</tr>
<tr>
<td>$10^{-2}$</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>$10^{-1}$</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>$10^{0}$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$10^{1}$</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>$10^{2}$</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>$10^{3}$</td>
<td>1000</td>
<td></td>
</tr>
<tr>
<td>$10^{6}$</td>
<td>1 000 000</td>
<td>k (kilo)</td>
</tr>
<tr>
<td>$10^{9}$</td>
<td>1 000 000 000</td>
<td>M (mega)</td>
</tr>
<tr>
<td>$10^{12}$</td>
<td>1 000 000 000 000</td>
<td>G (giga)</td>
</tr>
<tr>
<td>$10^{15}$</td>
<td>1 000 000 000 000 000</td>
<td>T (tera)</td>
</tr>
</tbody>
</table>

$10^{15}$
LHC/CMS Tiered Data System

- ~3000 Physicists, ~40 countries
- Tens of PetaBytes/yr
- ExaBytes in ~10 years

Tier 0 +1

Tier 1
10 - 40 Gbps
IN2P3 Center
RAL Center
INFN Center
FNAL Center

Tier 2
1-10 Gbps
Tier 2 Center

Tier 3
1-10 Gbps
Physics data cache
Workstations

Tier 4
1 to 10 Gbps
Institute
Institute
Institute

CERN Center
PBs of Disk; Tape Robot

Online System
PBytes/sec
MBytes/sec

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

~200-1500 MBytes/sec
LHC/CMS Tiered Data System

- ~3000 Physicists, ~40 countries
- Tens of PetaBytes/yr
- ExaBytes in ~10 years

Tier 1
- 10 - 40 Gbps
- IN2P3 Center
- RAL Center
- INFN Center
- FNAL Center

Tier 2
- 1-10 Gbps
- Tier2 Center
- Center
- Center

Tier 3
- 1-10 Gbps
- Institute
- Institute
- Institute

Tier 4
- 1 to 10 Gbps
- Workstations
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
Towers

- $2 \times 10^9$ events/yr
- 1 event = 1.6 MB
- 3.2 PB/yr from CMS
- 8.5 GB DVD (1-side, 2-layers)
- DVD thickness = 1.2 mm

How many DVDs?

$3.2 \text{ PB} \div 8.5 \text{ GB} \approx 4$만 DVDs

How tall?

$4$만 DVDs $\times 1.2 \text{ mm} \approx 500$ m
Tier 2 in Korea

- CMS
  경북대 (KNU)

- ALICE
  한국과학기술정보연구원 (KISTI)
GRID Real Time Monitoring
http://rtm.hep.ph.ic.ac.uk/webstart.php
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
Bonus: Distributed analysis
Car Driving

시동(출발) 가속(진행) 감속(정지)

Map
Data analysis

분석시작  분석중  분석끝

Configuration
Distributed analysis

Configure once, run anywhere
Your interest: 새로운 물리현상 발견 Vs. 미래 인터넷 환경 변화
과학자들간의 정보교환을 위해 개발된 기술 ➔ 생활의 혁명
2020 그들이 온다