Jak znaleźć igłę w stogu siana

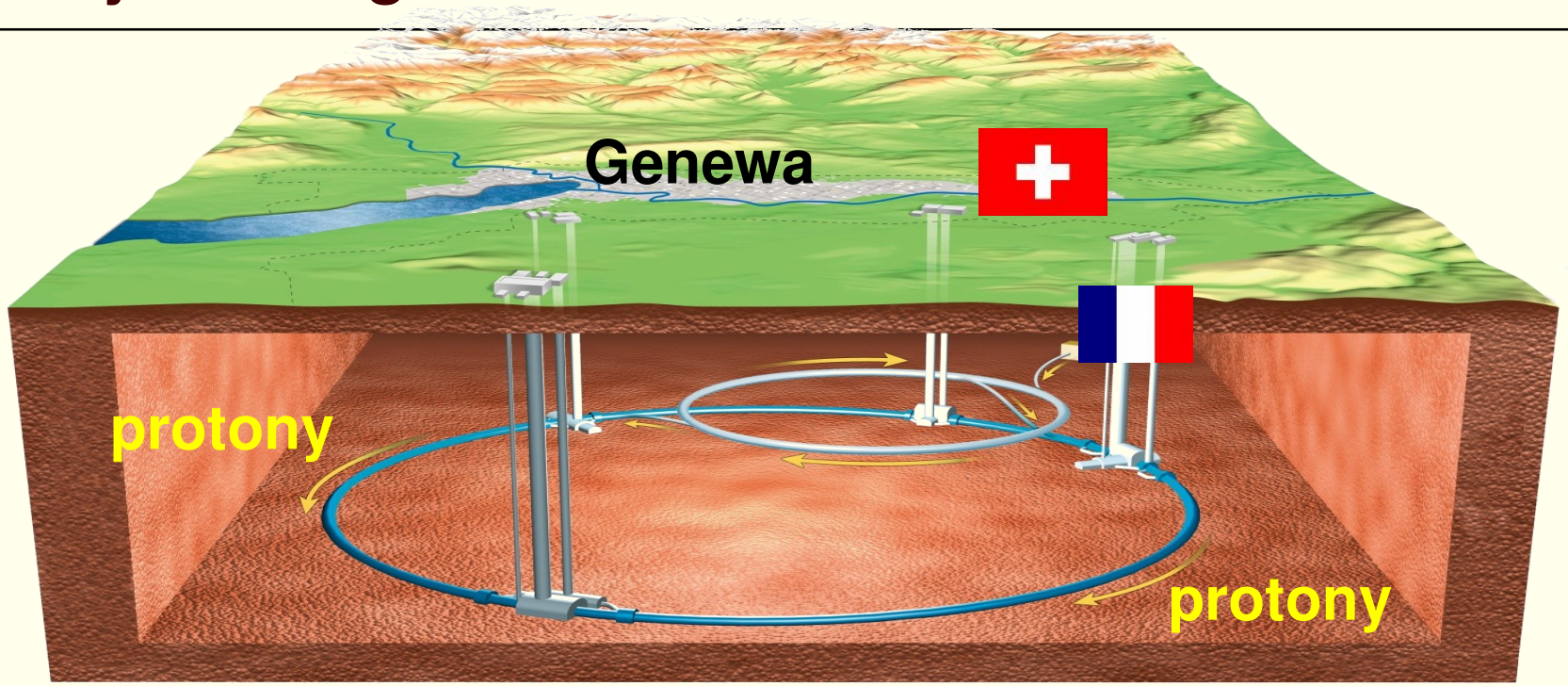
Rola obliczeń komputerowych w eksperymentach fizyki wysokich energii

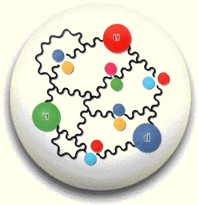
Piotr Golonka
CERN EN/ICE-SCD

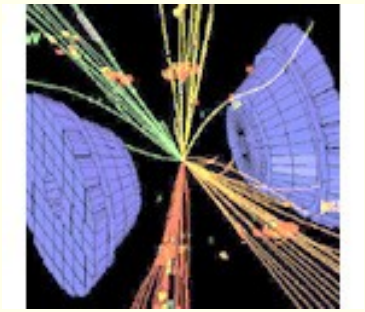
Plan

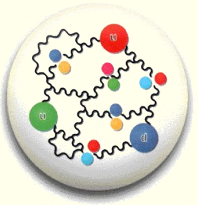
- Co jest igłą a co stogiem siana
  - ... między teorią a doświadczeniem
- Seria złożonych problemów
  - Zbieranie danych
  - Opracowywanie danych
  - Analiza danych
- Technologie
  - Klastry, GRID

Co jest stogiem siana? - LHC



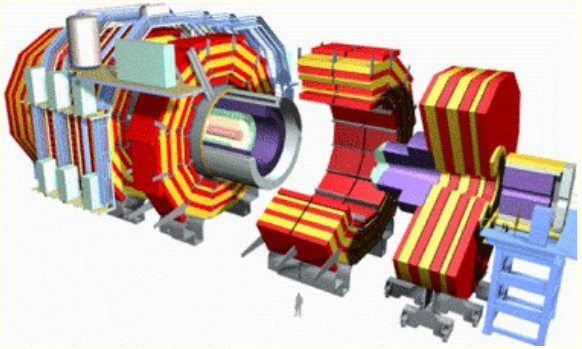


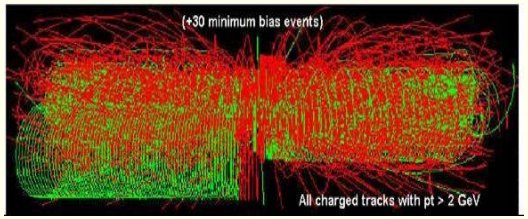




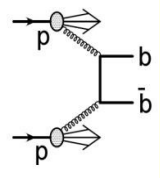
14 TeV = 14·10^{12} eV

Co jest stogiem siana? ... i detektory





LHC → Eksperyment → Interesujące oddziaływania proton-proton



Co jest igłą ?

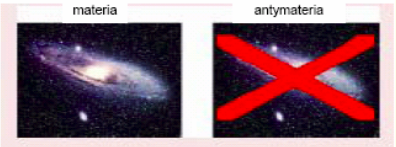
- Większość zderzeń to zwykłe, dobrze znane oddziaływania.
- Te interesujące które niosą informacje o nieznanych zjawiskach są bardzo rzadkie.
  - Typowo 1 przypadek na 100 mln!
- Jak „zauważyć” taki przypadek w aparaturze i zapisać go na taśmę gdy mamy tak duży strumień danych ≡ tak mało czasu na analizę „on-line”?

Czego szukamy w eksperymentach?

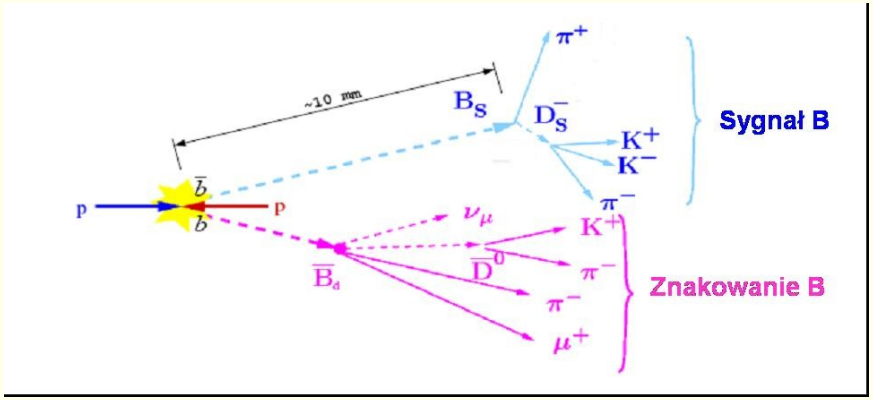
Przykład: eksperyment LHCb

Łamanie symetrii CP

- *Parzystość CP*
  - jednoczesne wykonanie operacji sprzężenia ładunkowego C i inwersji przestrzennej P, tzn. operacji $q \to -q$ oraz $\vec{r} \to -\vec{r}$. CP przeprowadza cząstkę w jej antycząstkę o przeciwnym pędzie i skrętności
  - niezachowanie parzystości P (1956) oraz parzystości C (1957) w oddziaływaniach słabych
- *Łamanie parzystości kombinowanej CP*
  - łamanie symetrii CP w rozpadach długożyciowych mezonów K^0 (1964)
- *Asymetria materia-antymateria*
  - wszystkie obserwacje wskazują, że obecny wszechświat składa się głównie z materii
  - hipoteza Sacharowa (1967)
    - jednym z 3 warunków powstania asymetrii materia-antymateria jest **łamanie CP**



Badanie łamania CP od 40 lat (zjawiska subtelne)
LHCb: duże próbki danych, rozszerzenie obszaru badań

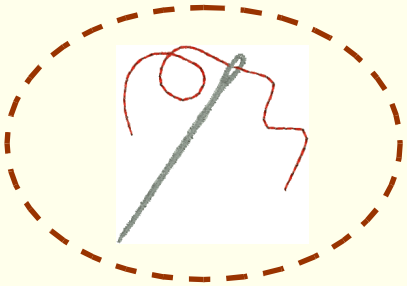


Łamanie CP w układzie mezonów pięknych (B)

Zjawiska rzadkie ~ 10^{-4} – 10^{-8}

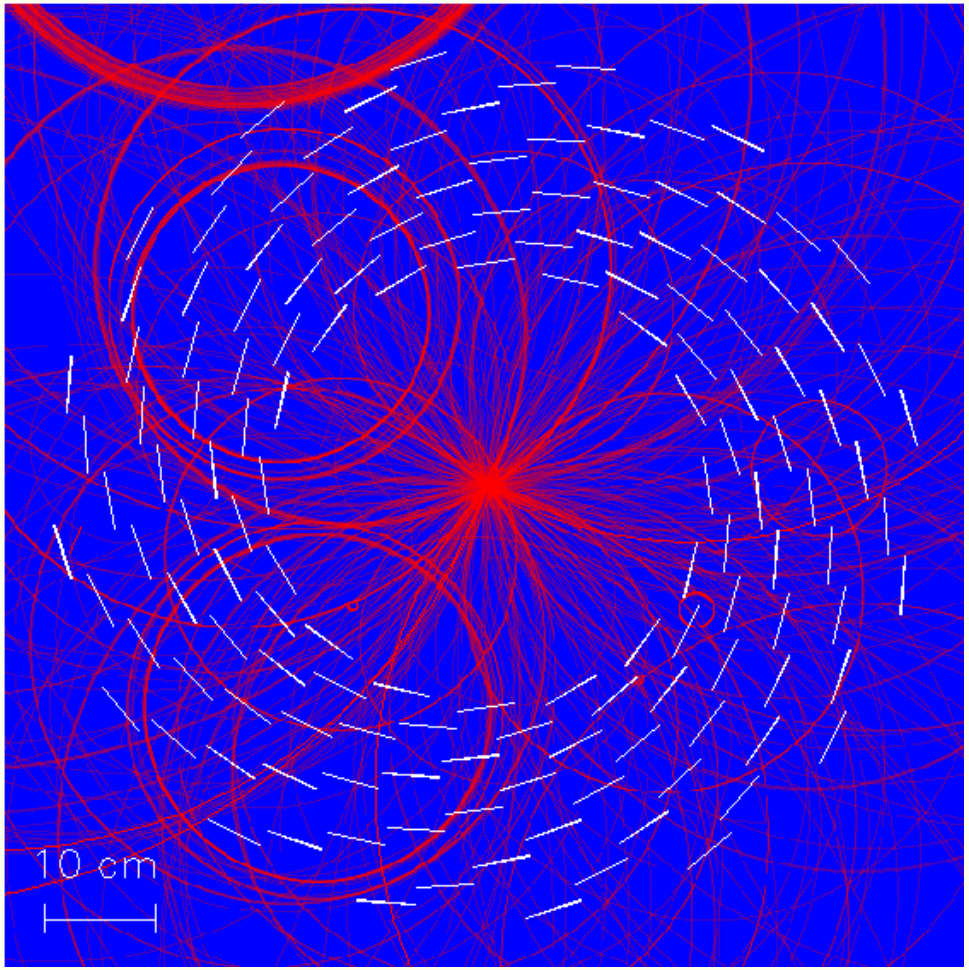
Pozostałe eksperymenty na LHC mają swoje zestawy „igieł”

- Cząstka Higgsa (ATLAS, CMS)
- Supersymetria (ATLAS, CMS)
- Plazma kwarkowo-gluonowa (ALICE)
- **...**



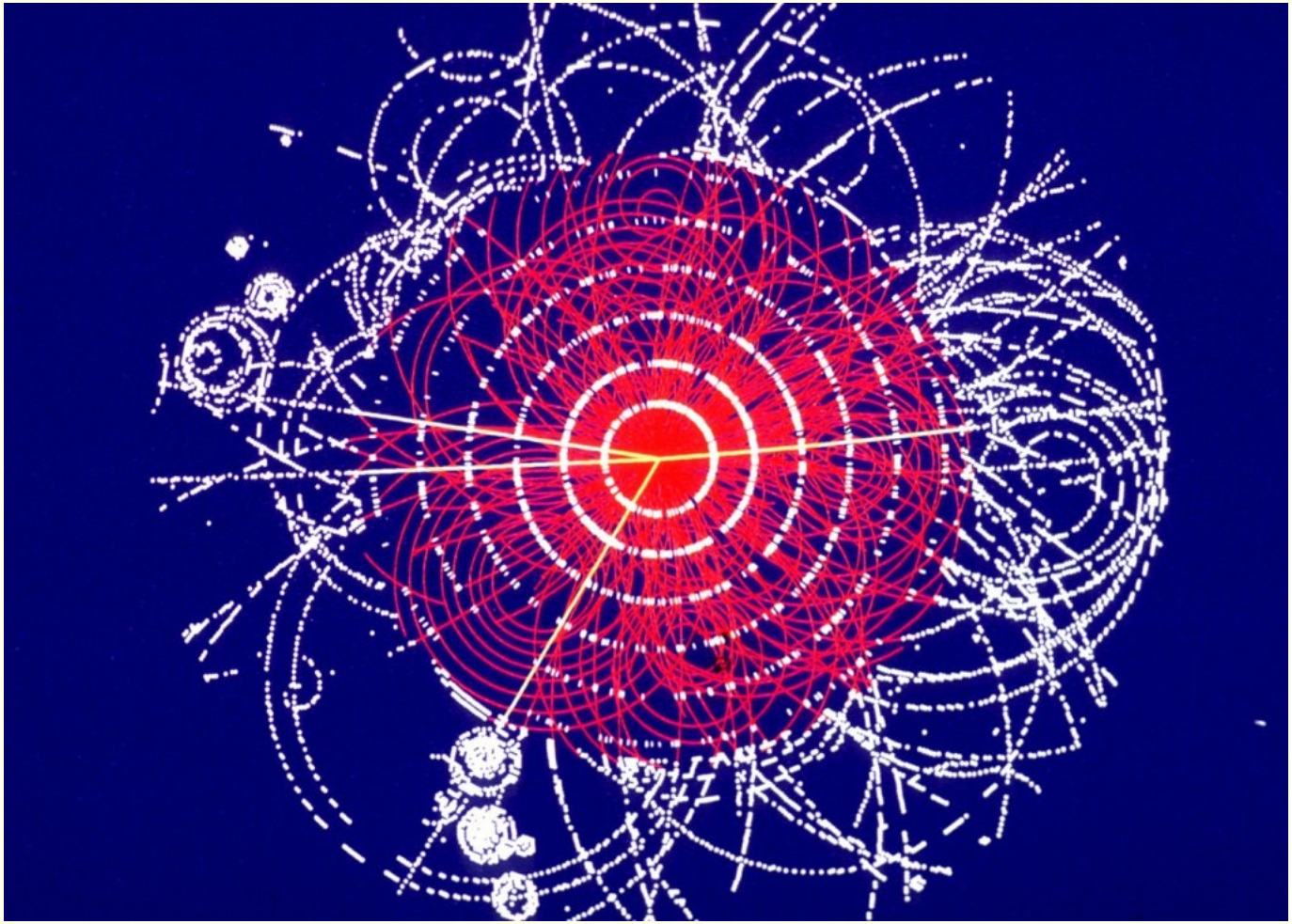
Przykład: rekonstrukcja śladów cząstek

Symulacja komputerowa danych z 1 przypadku w detektorze wewnętrznym eksperymentu ATLAS



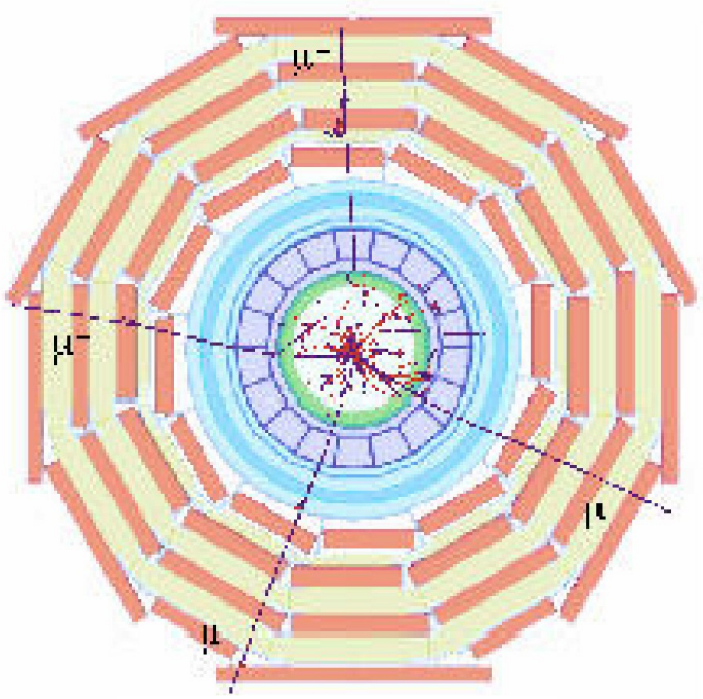
Zagadka:
znajdź 4
proste ślady

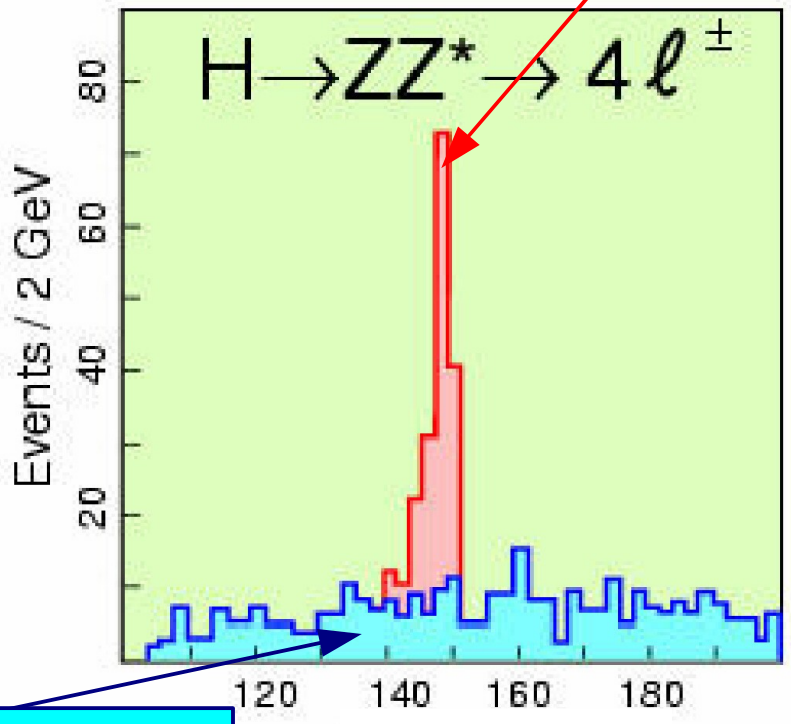
Przykład: rekonstrukcja śladów cząstek



Rozwiązanie:
Rozpad cząstki Higgsa na 4 miony: H → 4μ

Analiza danych: co zobaczy Fizyk:

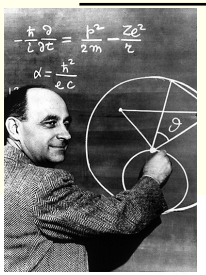


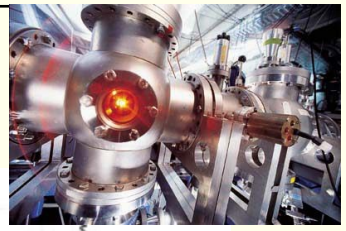


Cząstki Higgsa
o masie ~150 GeV

Tło od innych rozpadów
dających 4 miony

Teoria, doświadczenie i ... komputery

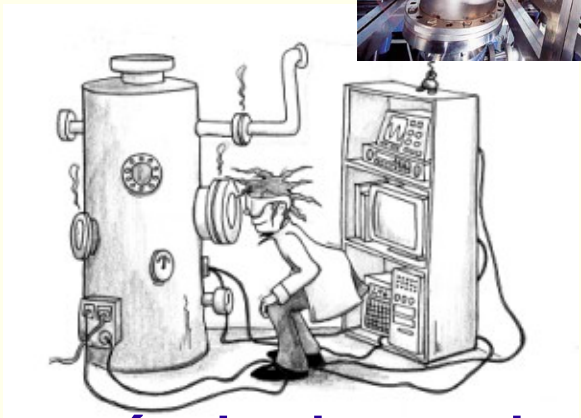




Teoria

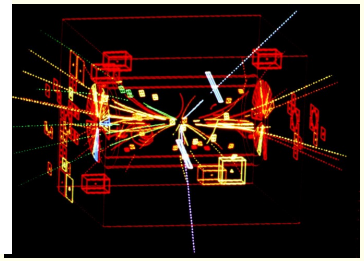
(?)

=

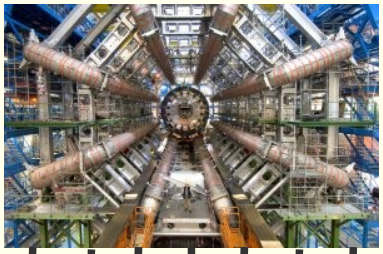


Doświadczenie

Metody komputerowe

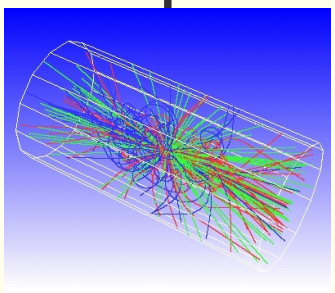


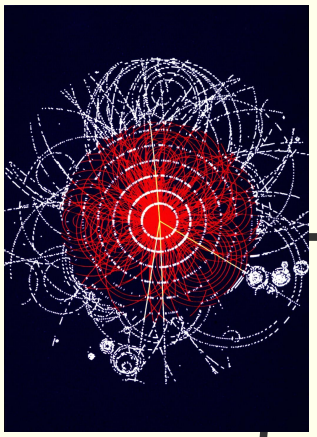
“doświadczenie = teoria” (?)



Filtr przypadków

Generator(y) przypadków

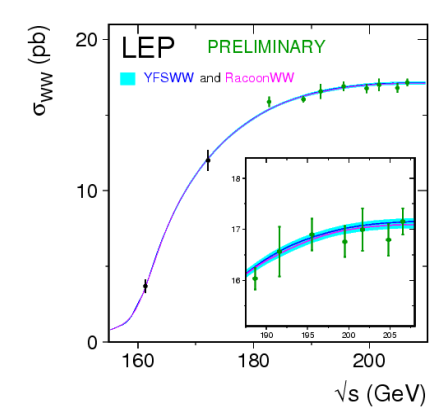




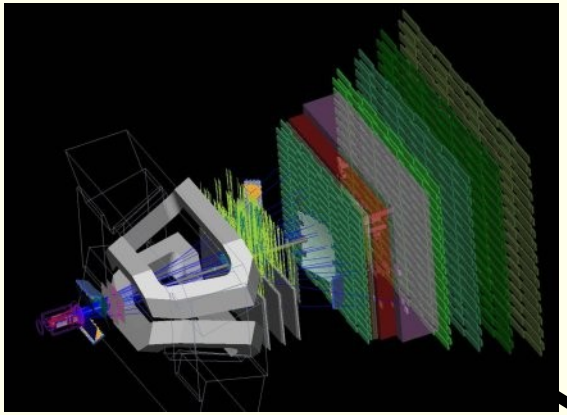
Symulacja detektora

Dane

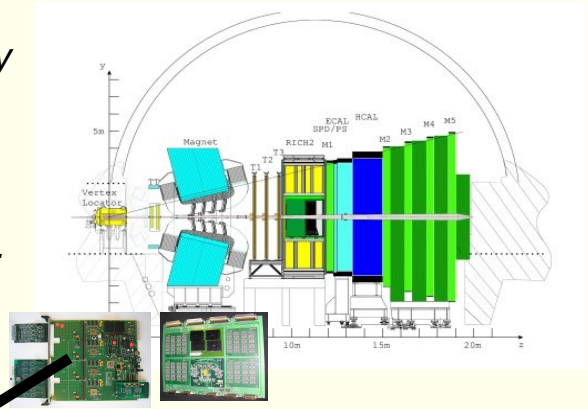
$$\mathcal{L}_H = -\frac{1}{2}[(\partial_\mu - iW^a_\mu t^a - iB_\mu)\phi]^2 - \frac{\mu^2}{2}\phi^*\phi - \frac{\lambda}{4}(\phi^*\phi)^2,$$



Symulacja i prawdziwy eksperyment



W efekcie modelowania otrzymujemy dane symulowane w formacie takim jakie będą napływać z elektroniki odczytu aparatury eksperymentalnej.

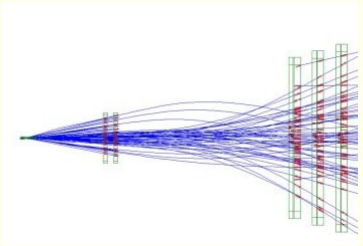


Symulacja

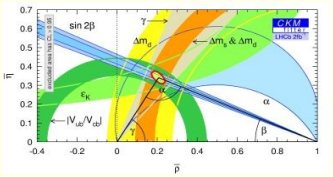
1. Generator przypadków (teoria)
2. Modelowanie odpowiedzi aparatury (Geant4)

Prawdziwy eksperyment

Rekonstrukcja

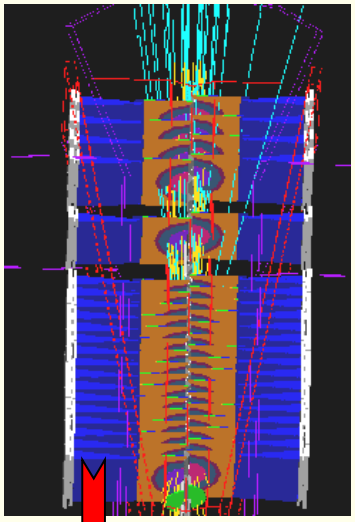


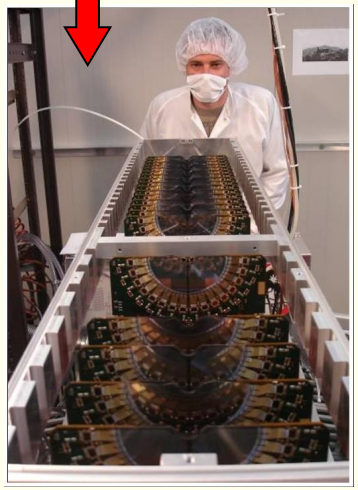
Analiza fizyczna



Modelowanie aparatury

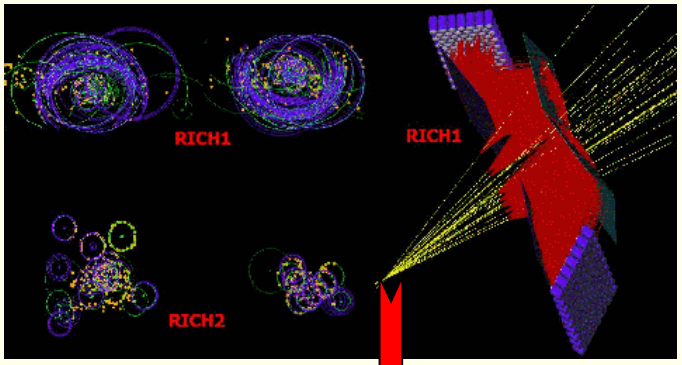
Detektor wierzchołka



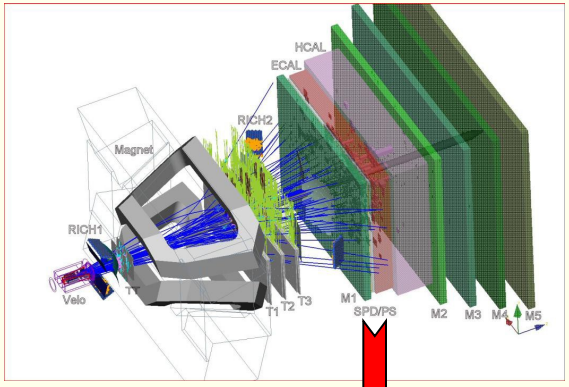


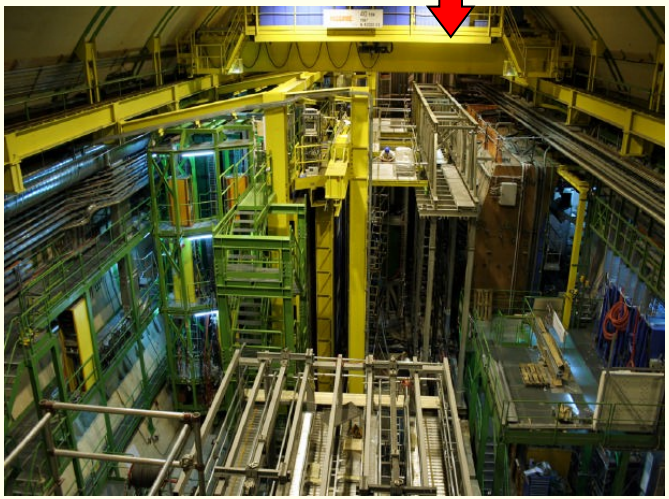
Detektor RICH

Identyfikacja cząstek

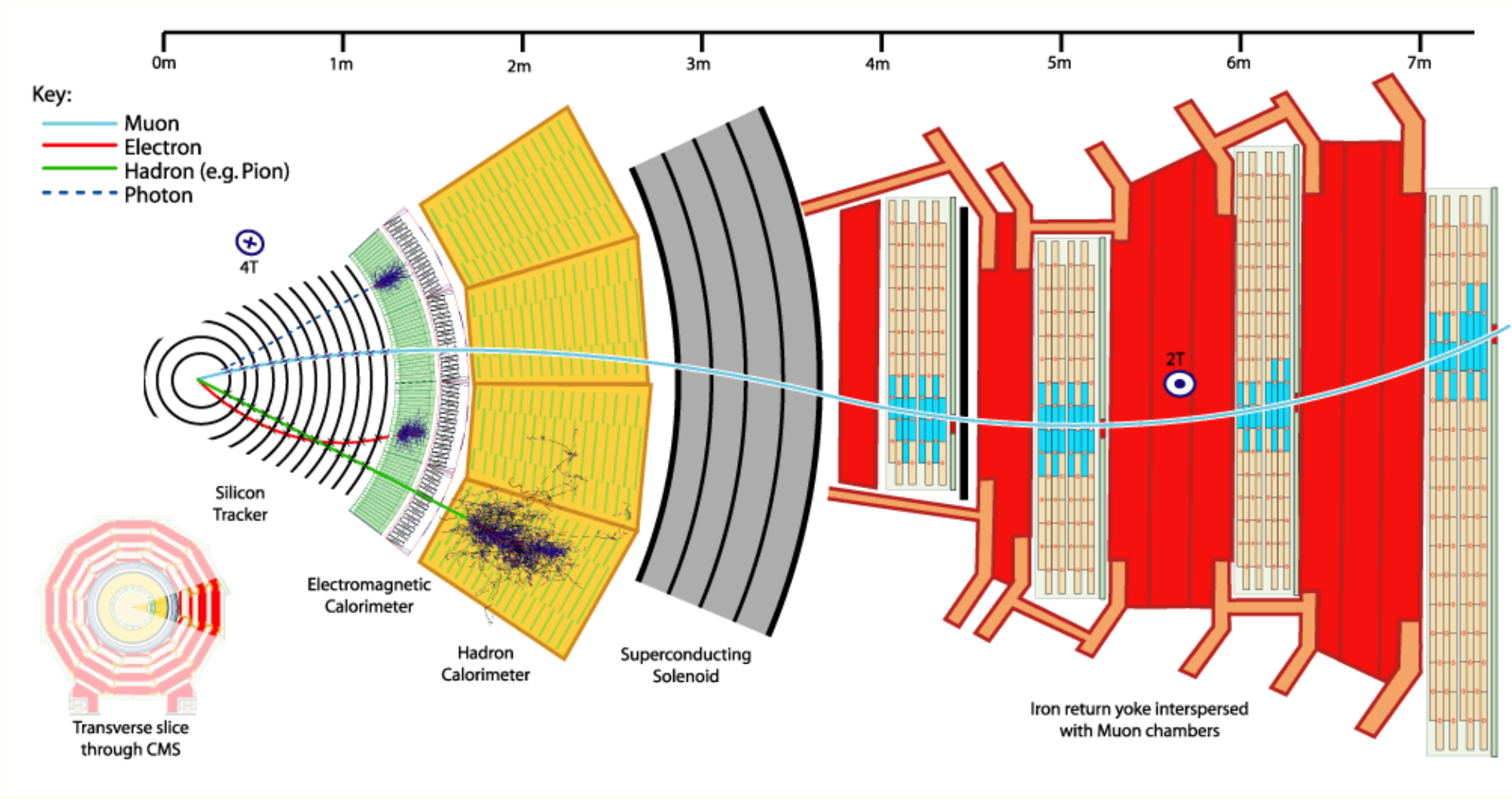


Pełny spektrometr





Symulacja zjawisk w detektorach



Slajd ze strony www eksperymentu CMS

Co jest stogiem siana

- Przecięcia pęków wiązek: co 25 nanosekund (częstotliwość LHC: 40 MHz)
- Nie wszystkie “pęczki” są pełne – zderzenia: 31 MHz
- 10 * 10^6 zderzeń w ciągu 1 sekundy widocznych w eksperymencie LHC-b!

Dane przepływające przez eksperyment

	Sekunda	Dzień	Rok
l. przypadków	**10^7**	**10^{12}**	**10^{14}**
Ilość danych	**300 GB**	**30 PB**	**3 EB ($3*10^{18}$)**
l. dysków (300 GB)	**1**	**100 tys**	**10 mln**

- Dla eksperymentów ATLAS i CMS jest znacznie gorzej
  - ATLAS 2PB/sek ≡ 3 mln CD/s

Dygresja... ile to bajtów?

1 Megabajt (1MB)

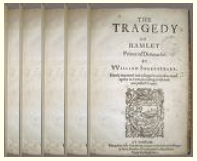
zdjęcie z aparatu cyfrowego: 2 MB

1 Gigabajt (1GB) = 1000MB

film na nośniku DVD: 6 GB

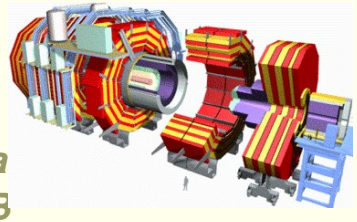
1 Terabajt (1TB) = 1000GB

Światowa roczna produkcja książek



1 Petabajt (1PB) = 1000TB

Roczne składowanie danych w LHC: 14 PB



1 Eksabajt (1EB) = 1000 PB

Roczna produkcja informacji w 2002 (zapisa

Wszystkie dotąd wypowiedziane słowa: 5EB

Przepływ cyfrowej informacji w 2002: 18 EB (z czego 98% to telefony)

Ilość danych "widzianych" przez eksperyment na LHC: 3 EB

Źródło: HMI Report 2003, http://hmi.ucsd.edu/howmuchinfo.php ; najnowszy raport (HMI 2008) spodziewany pod koniec roku)

http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/execsum.htm

Co jest stogiem siana

- Przecięcia pęków wiązek: co 25 nanosekund (częstotliwość LHC: 40 MHz)
- Nie wszystkie “pęczki” są pełne – zderzenia: 31 MHz
- 10 * 10^6 zderzeń w ciągu 1 sekundy widocznych w eksperymencie LHC-b!

Dane przepływające przez eksperyment

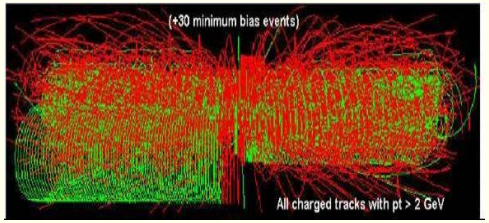
	Sekunda	Dzień	Rok
l. przypadków	**10^7**	**10^{12}**	**10^{14}**
Ilość danych	**300 GB**	**30 PB**	**3 EB ($3*10^{18}$)**
l. dysków (300 GB)	**1**	**100 tys**	**10 mln**

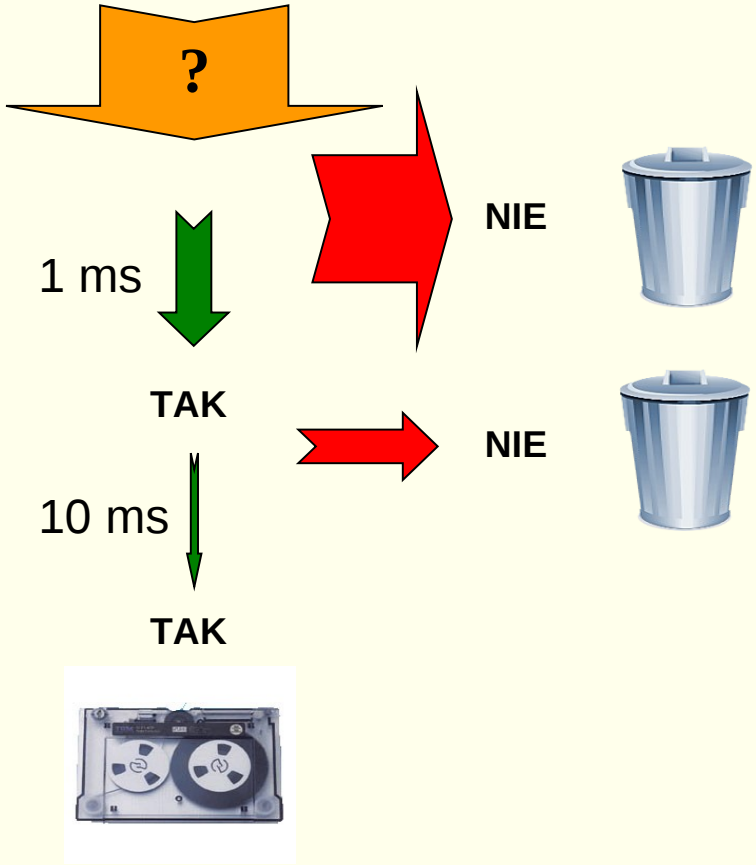
- Dla eksperymentów ATLAS i CMS jest znacznie gorzej
  - ATLAS 2PB/sek ≡ 3 mln CD/s

- Zapis wszystkich przypadków jest niemożliwy
- Na każdy przypadek należy jednak „zerknąć” i zadecydować czy jest interesujący.
  - Do „przeglądania” przypadków służy **system wyzwalania** czyli filtracji przypadków (tzw. **tryger**).

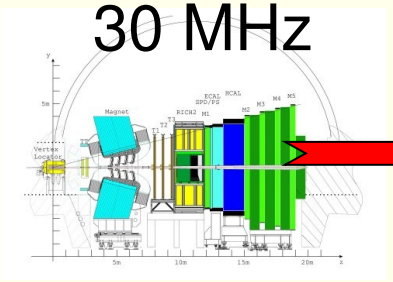
Zbieranie danych - system wyzwalania

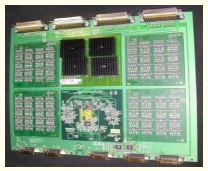
- Zadaniem systemu wyzwalania jest przede wszystkim **jak najszybsze odrzucanie zbędnych przypadków** (a nie wybieranie tych interesujących).
- W kolejnych stopniach filtrowania mamy do dyspozycji coraz więcej czasu na dokładniejszą analizę przypadków zaakceptowanych i dokładniejsze odrzucanie.





Zbieranie danych





35 GB/s

1 MHz

2 kHz

L0 (hardware)

Wstępny poziom wyzwalania

HLT (algorytmy)

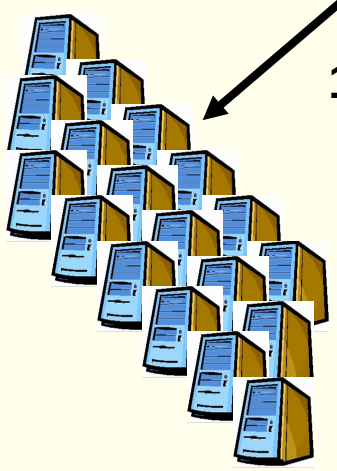
Wyższy stopień wyzwalania

$2 \cdot 10^{10}$ przyp./rok

1 MHz → 1μs/przypadek.

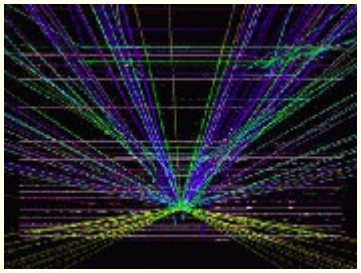
W ciagu 1 μs nie można nawet zdekodowac danych

Co zrobic ???



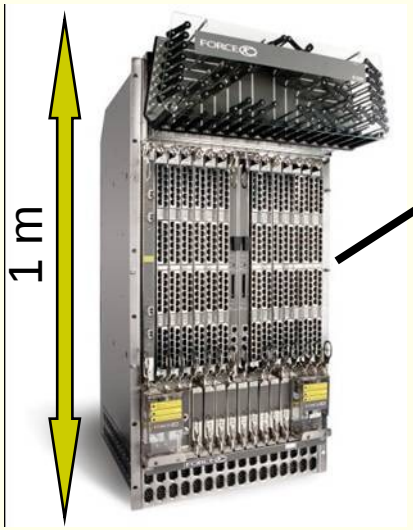
1000 procesorów → 1 ms/przypadek

Możliwa staje się rekonstrukcja sladow w detektorze wierzcholka w rzucie rφ

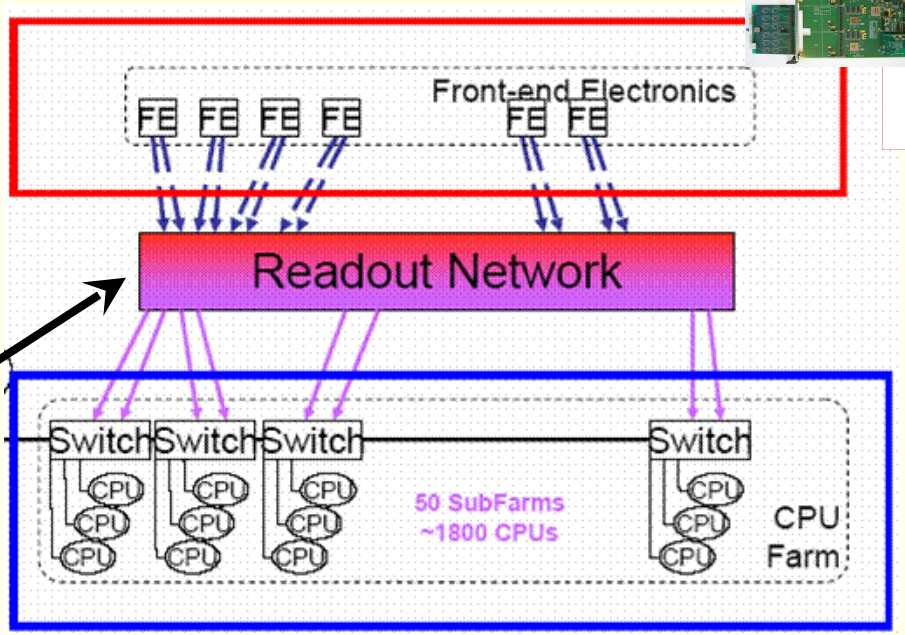


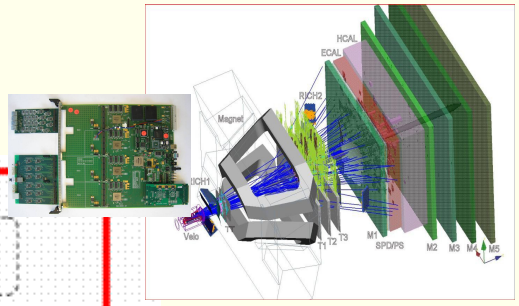
Farma procesorów on-line

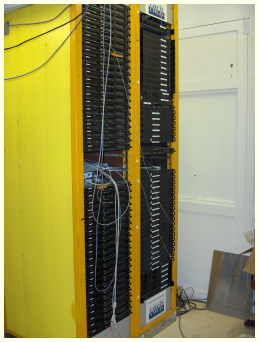
1 MHz na wejściu
→ 30 GB/s

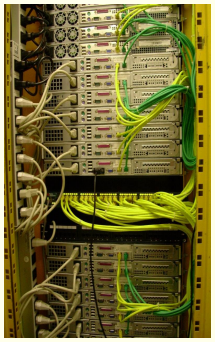


Force10 E1200:
1260 portów GbE
Przepustowość 50 GB/s



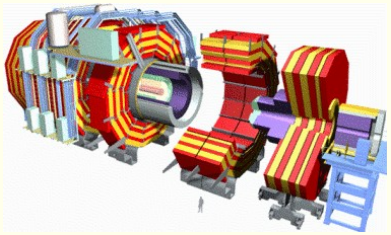


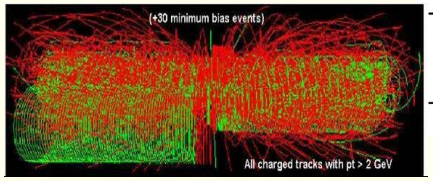


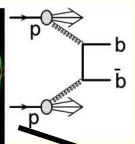


1000 CPU
2 kHz na wyjściu
Zapis 250 MB/s

Igła czy stóg siana?







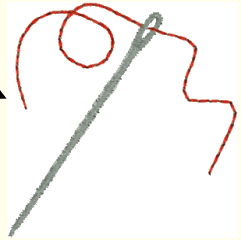
LHC → Eksperymenty → Interesujące oddziaływania p-p

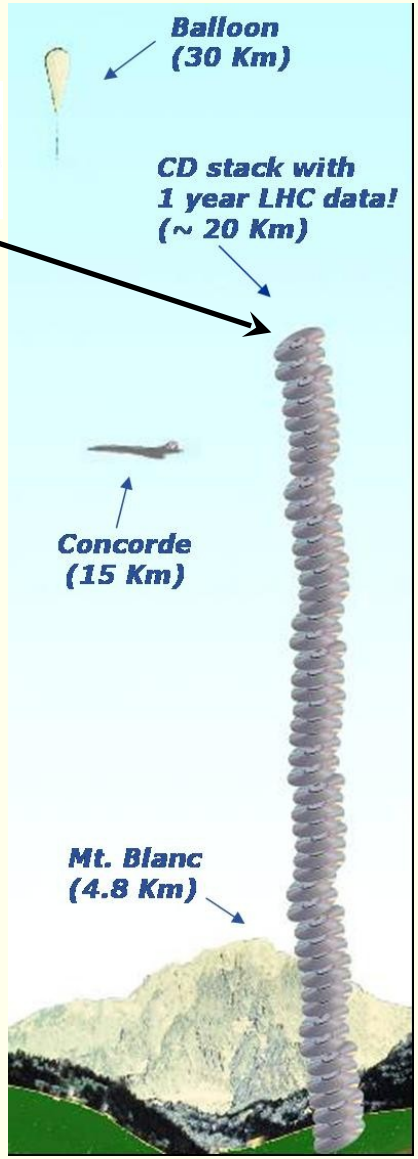
Filtrowanie

Redukcja > 10^5

3 Eksabajty/rok → 15 Petabajtów/rok

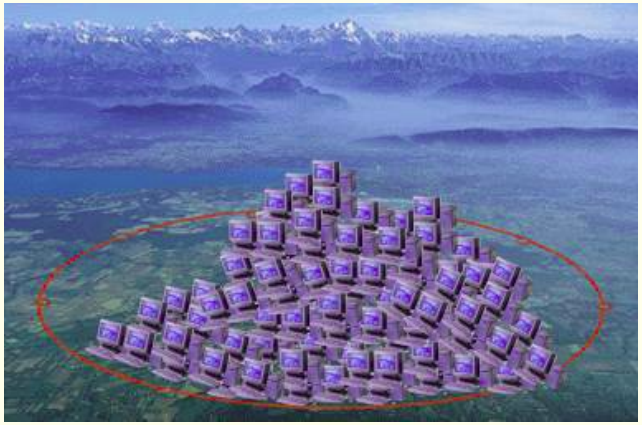
15 Petabajtów / rok przefiltrowanych danych z czterech eksperymentow na LHC





Analiza danych

- Analiza danych zapisanych na taśmy
  - Przypadki przefiltrowane
- Niespotykane dotąd wyzwania
  - ~kilkanaście PB danych na rok
  - 200 tys procesorów
  - 10 tys dysków
- Skoncentrowanie infrastruktury w jednym miejscu jest niesłychanie trudne (i niecelowe bo blokuje rozwój innych ośrodków).
  - Instalacja, zasilanie, chłodzenie, obsługa systemów operacyjnych i oprogramowania
- Konieczny model rozproszonych obliczeń czyli zespól współpracujących ze sobą farm komputerowych.
  - **Czy można jednak stworzyć coś bardziej uniwersalnego co byłoby przydatne także dla innych dziedzin?**



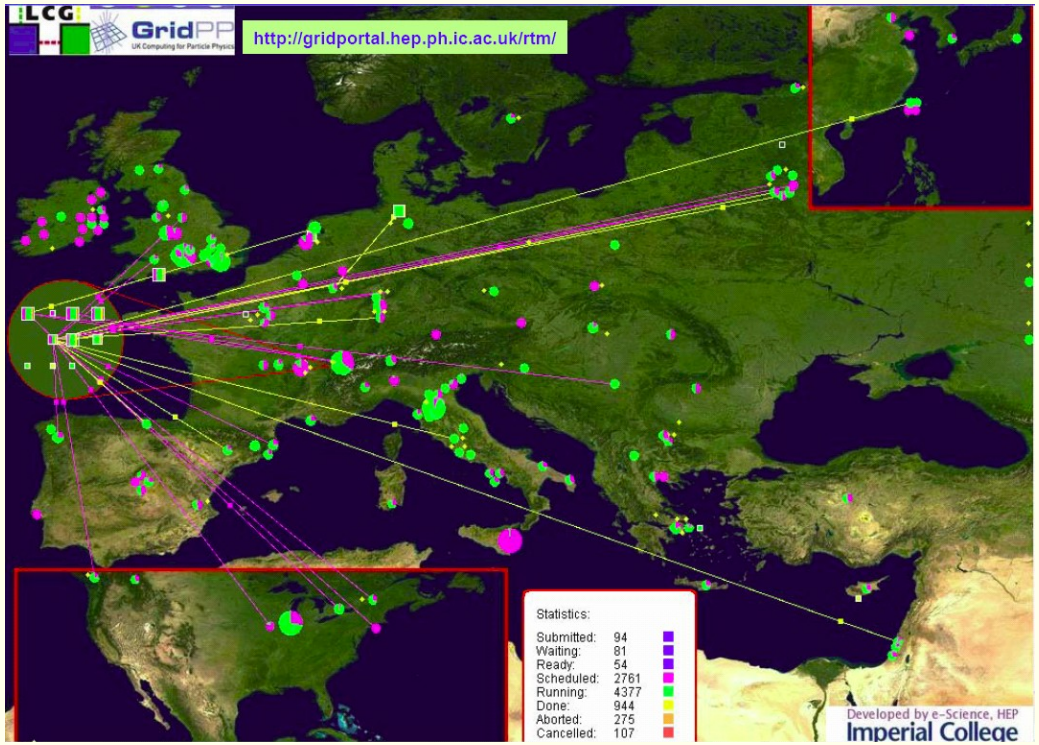
GRID – globalny komputer

- WWW – przewrót w dostępie do informacji (wynalezione w CERN).
  - Jednolity dostęp do informacji bez względu na miejsce na Ziemi.
- GRID – jednolity dostęp do zasobów obliczeniowych.
  - GRID widziany przez każdego użytkownika jako **jeden wielki komputer**.
  - Dobra platforma współpracy dla dużych projektów.
    - Wspólne narzędzia i dane, dostęp do mocy obliczeniowej.
  - Analogia z **siecią elektryczna** (także w przypadku opłat za pobraną moc obliczeniowa)

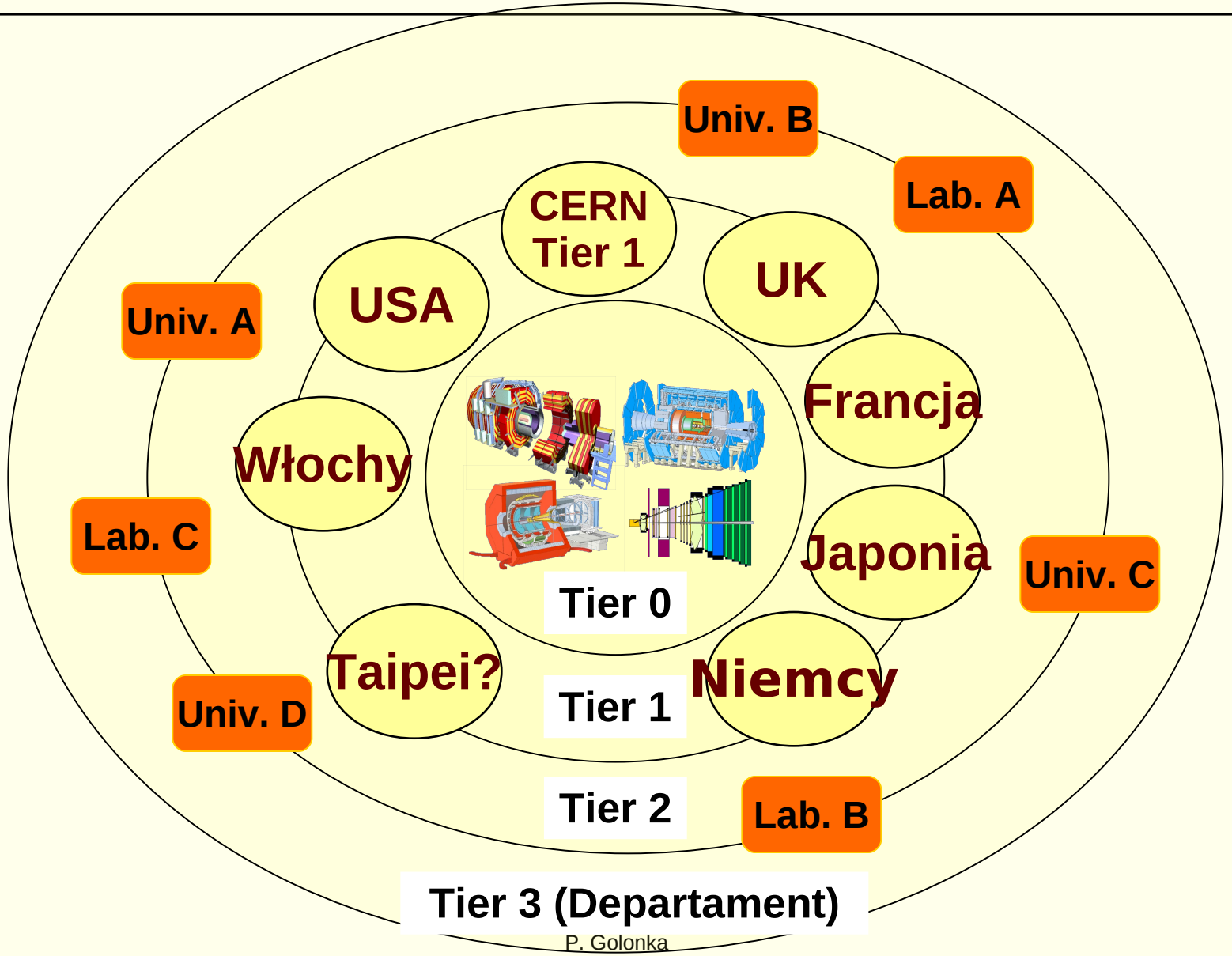
- Pierwszą udaną inicjatywą na dużą skalę były obliczenia w ramach projektu SETI (*Search for Extra-Terrestrial Intelligence*)
  - Każdy posiadacz PC mógł udostępnić moc obliczeniowa np. w nocy i uruchomić program analizujący sygnały przychodzące z kosmosu.
  - Projekt zaoszczędził wiele milionów $

Ogólnoświatowa sieć GRID

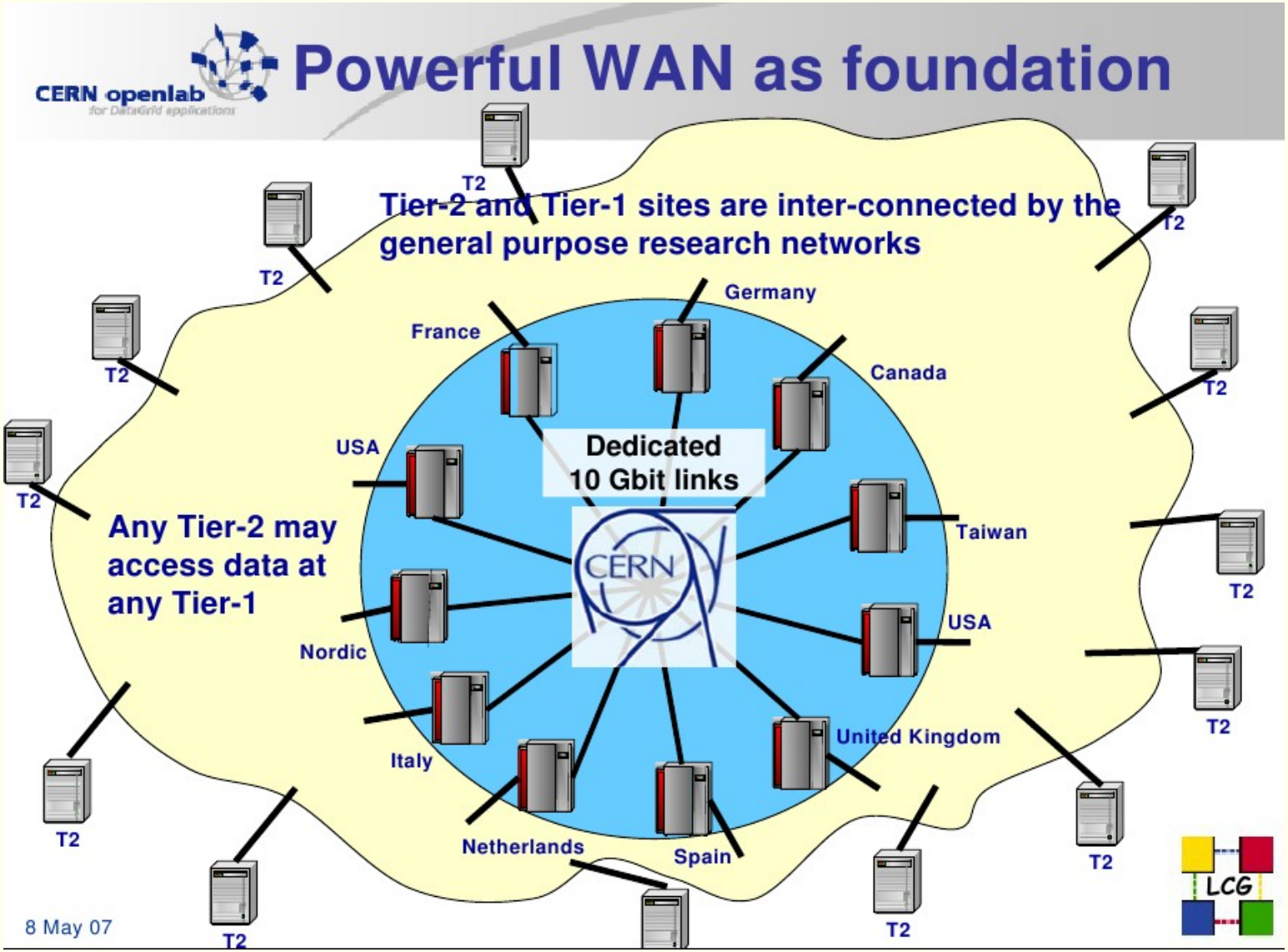
- EGEE jest największym ogólnoświatowym projektem GRID
  - Obecnie ponad 80 tys procesorów
  - 300 ośrodków naukowych w 50 krajach na całym świecie
  - 10 000 użytkowników
  - 100 tys zadań wykonywanych jednocześnie
  - 15 PB przestrzeni dyskowej
- Koordynatorem projektu jest CERN
- Fizyka cząstek jest głównym użytkownikiem (największe potrzeby)



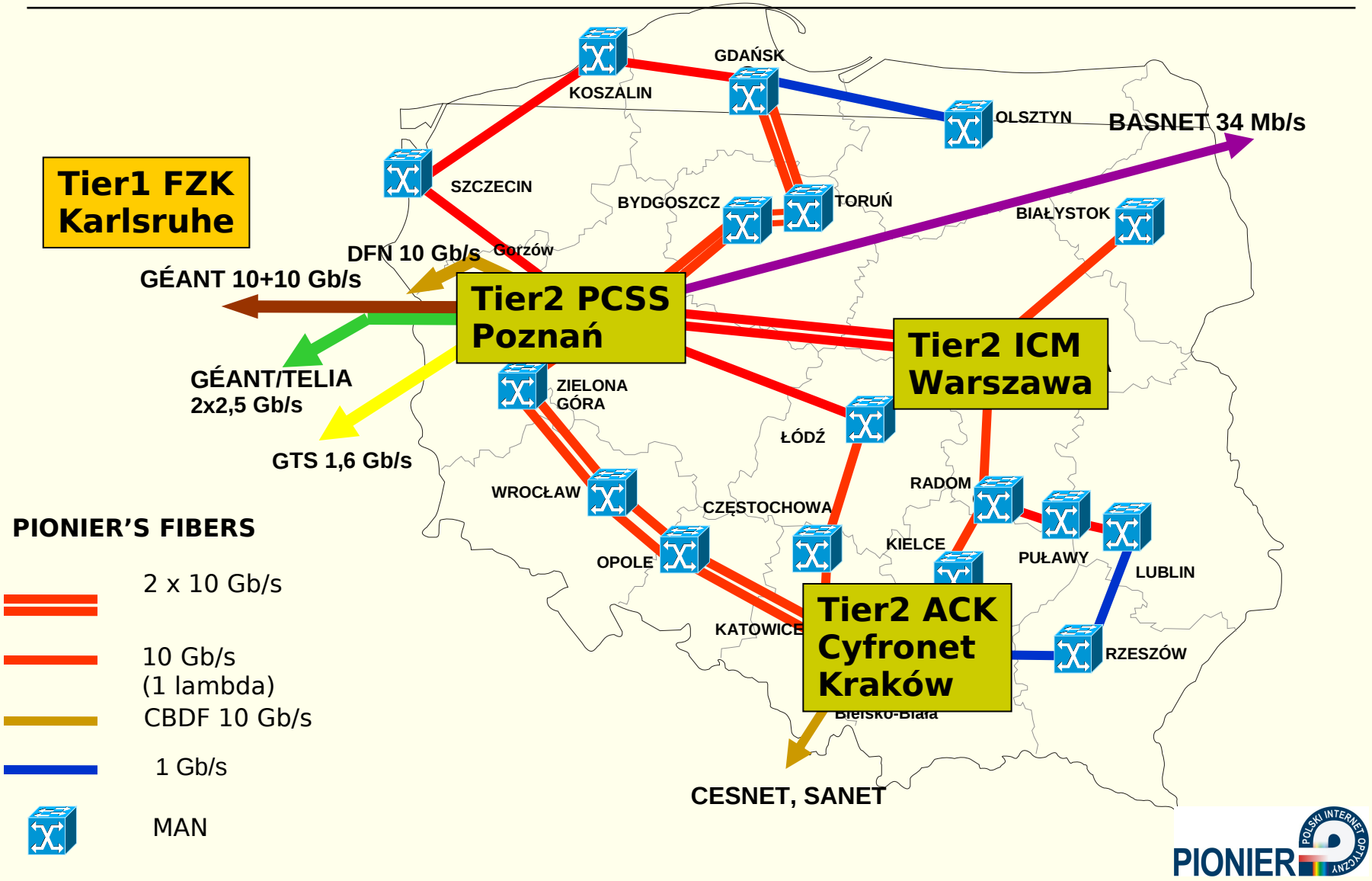
Struktura hierarchiczna (Multi Tier)



Struktura hierarchiczna (Multi Tier)

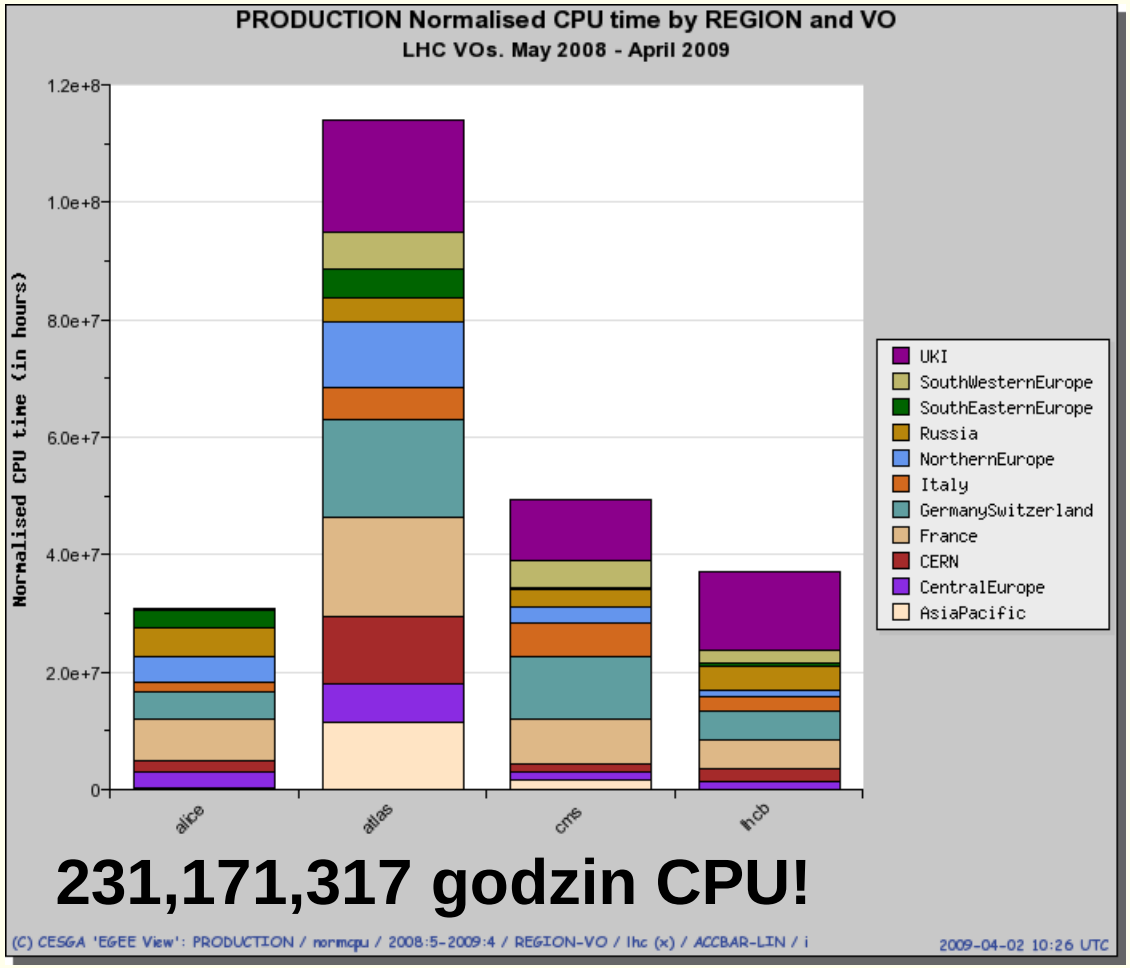


Polska infrastruktura WLCG

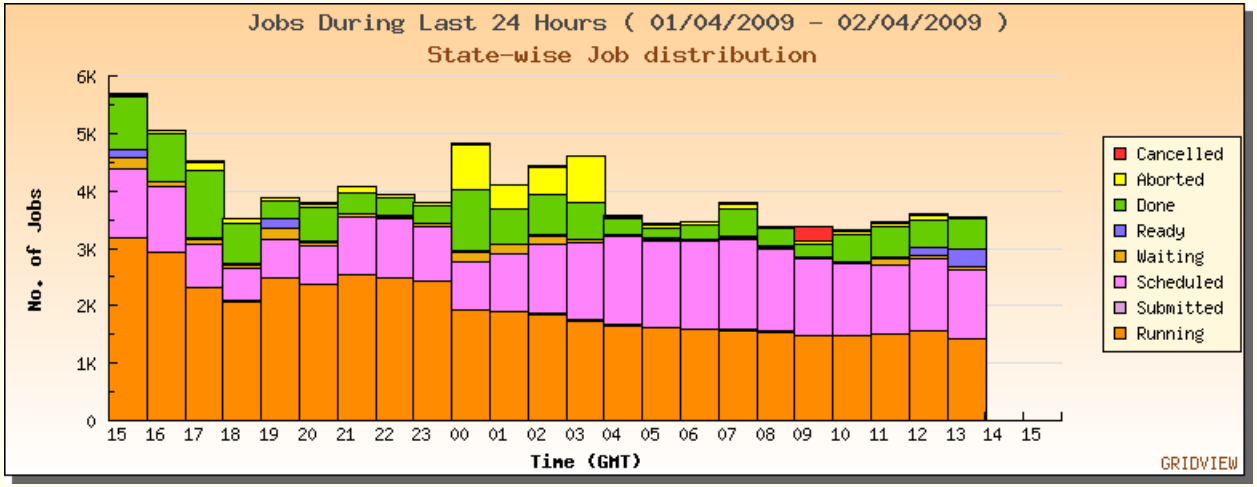


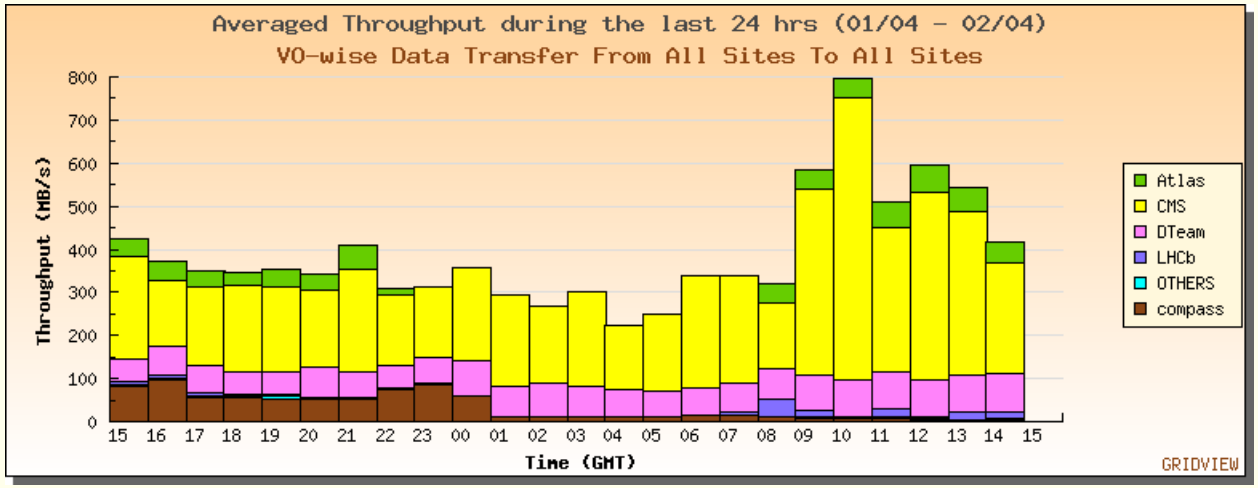
EGEE

Normalised CPU time [units 1K.SI2K.Hours]

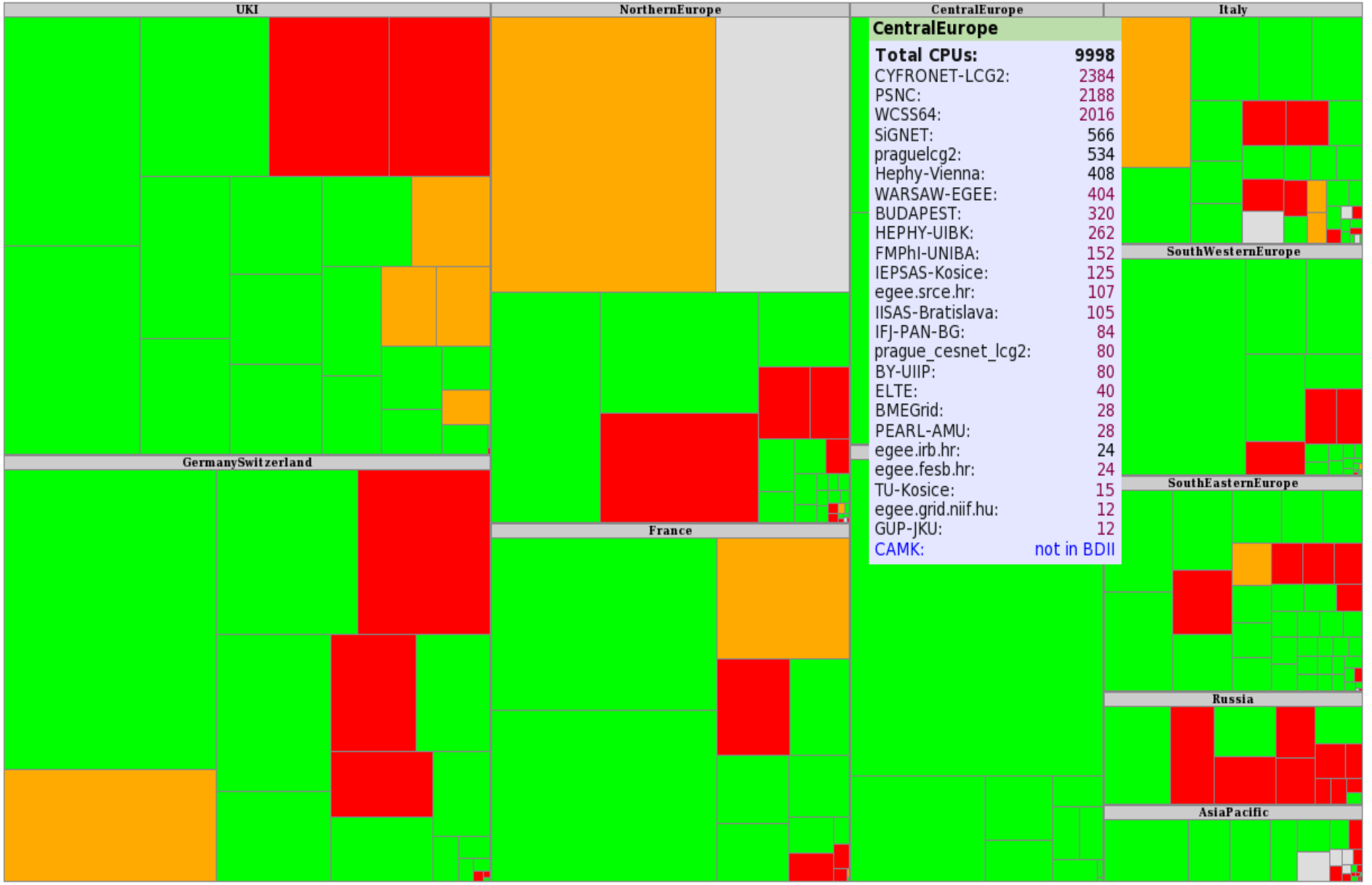


LCG: LHC Computing Grid





GridMap – Visualizing the "State" of the Grid



Latest SAM results, Site Status, for 'OPS' VO, 02 Apr 2009 15:27 GMT.
Size of site rectangles is number of CPUs from BDII.
Certified Production sites, grouped by regions.

Maint Down Degraded Ok

Więcej informacji...

- http://www.gridcafe.org

- http://lcg.web.cern.ch/lcg/monitor.htm

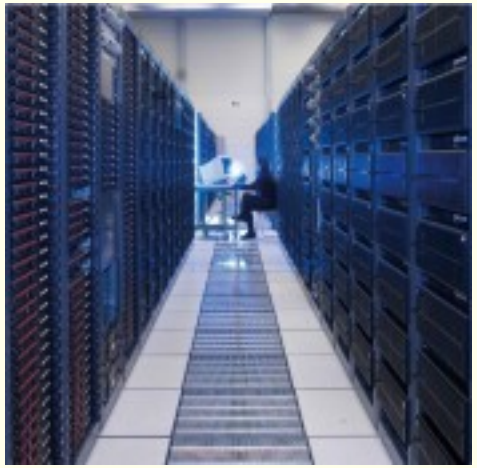
- Real-Time grid monitor: http://gridportal.hep.ph.ic.ac.uk/rtm/

Technologie: farmy PC

- Początkowo CERN zdecydował zakupić komputery w formie zwykłych PC dostępnych w sklepach.
- W międzyczasie technologie RACK i BLADE osiągnęły konkurencyjne ceny.
- Obecnie w dużych ośrodkach wygrywają technologie specjalistyczne RACK i BLADE

Pierwsze klastry w CERN
w standardowej technologii.

Hala w CERN przygotowana do montażu PC w szafach RACK



Serwery typu “rack”w centrum komputerowym CERN

PC typu RACK i BLADE

Technologie - c.d.

- Około 7500 procesorów (rdzeni)
  - Kilka procesorów w komputerze, kilka rdzeni w procesorze
- Ultra-szybkie połączenia sieciowe
  - 1, 10Gb Ethernet
  - Fiber Channel
- Składowanie danych: taśmy
  - np. robot: *Sun StorageTek SL8500*
    - Do 300 tys taśm, 500 GB każda
    - Do 2048 napędów o prędkości 100MB/s

- System operacyjny:
  - CERN Scientific **Linux**

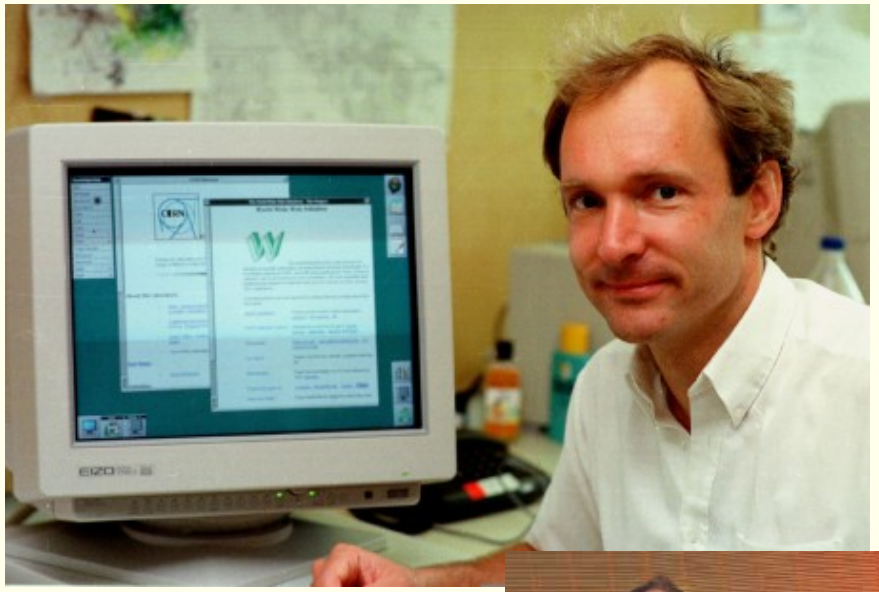
CERN openlab

Partnerstwo z czołowymi firmami informatycznymi:

- Najnowsze procesory i technologie sieciowe
- Bazy danych
- Automatyka przemysłowa
- Bezpieczeństwo

World Wide Web, europejski Internet

WWW:Tim Berners-Lee, 1989-2009!



CERNET: 1984
(first European TCP/IP network)

Podsumowanie

- Eksperymenty na LHC dostarczają ogromnej ilości danych.
- Przypadki których szukamy zdarzają się bardzo rzadko, ~1 na 100 mln.
- Jedynie dzięki komputerom udaje się zebrać, przetworzyć i analizować dane.
- CERN stymuluje rozwój nowych technologii informatycznych (WWW, GRID)