



High Performance Computers in Cyfronet

Andrzej Oziębło
Zakopane, marzec 2009



Plan

- Podział komputerów dużej mocy
- Podstawowe informacje użytkowe
- Opis poszczególnych komputerów
- Systemy składowania danych
- Co się zmieniło od zeszłego roku
- Plany na przyszłość



ACK CYFRONET AGH

Akademickie Centrum Komputerowe CYFRONET

Akademii Górniczo-Hutniczej im. Stanisława Staszica w Krakowie

**ul. Nawojki 11
30-950 Kraków 61
tel. centrali: (12) 632 33 55
tel. sekretariatu: (12) 633 34 26
fax: (12) 634 10 84, 633 80 54
e-mail: cyfronet@cyfronet.pl
<http://www.cyfronet.pl>**





Dwie rodziny komputerów

1. Ogólnego przeznaczenia: **SMP** i klastry
2. Klastry **gridowe** dedykowane do obliczeń w projektach EGEE, WLCG, etc.

SMP (Symmetric multiprocessing, wspólna pamięć): SGI
Altix 3700 - *baribal*, SGI 4700 - *panda*

Klastrer: IBM BladeCenter HS21 - *mars*

Klastry gridowe: HP Cluster Platform BL 2x220 - *zeus*







Podstawowe informacje użytkowe

1. Linux
2. System kolejkowy PBS; stosunkowo mała liczba kolejek
3. Plik READ.ME w katalogu domowym i poradnik użytkownika: http://www.cyf-kr.edu.pl/uslugi_obliczeniowe
4. Ograniczony *home* ale duży *scratch* - niewielkie dyski lokalne i „nieograniczone” zasoby na zewnętrznych macierzach; możliwość korzystania z bibliotek taśmowych (FSE)
5. Informacje o zmianach i awariach: motd i mail; będą zmiany
6. Ograniczenia na pamięć i liczbę procesorów

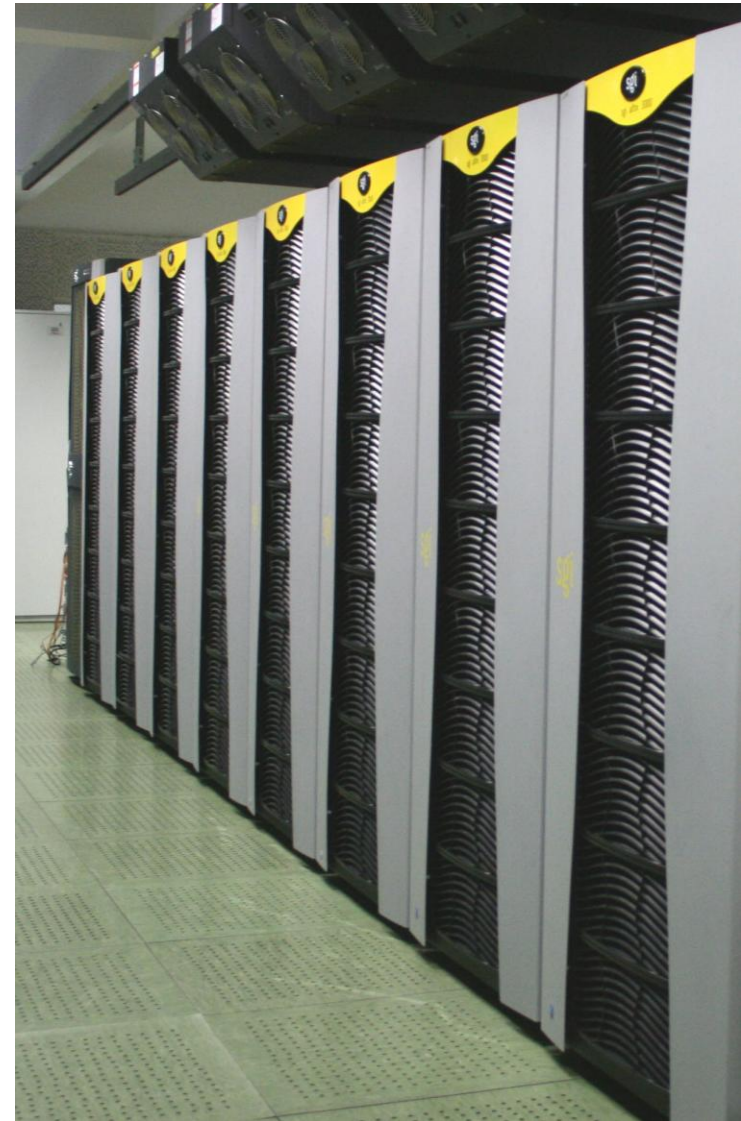


SGI Altix 3700 - *baribal*

system operacyjny:
Linux Suse 10

konfiguracja:

- 256 procesorów Itanium 2 z zegarem 1.5 GHz
- pamięć operacyjna 512 GB
- 256 pamięć dyskowa ~8 TB
- teoretyczna moc - ok. ~1.5 Tflops





SGI Altix 4700 - *panda*

system operacyjny:
Linux Suse 10

konfiguracja:

- 32 procesory Itanium 2 z zegarem 1.66 GHz
- pamięć operacyjna 64 GB
- Pamięć dyskowa ~2 TB
- teoretyczna moc - ok. 190 Gflops
- typu blade
- Wydzielona partycja RASC dla badań w zakresie akceleracji obliczeń

KU KDM 2009



Klaster IBM blade

- *mars*

system operacyjny:
Linux Redhat

konfiguracja:

- 224 rdzeni procesorów Intel Woodcrest z zegarem 2.66 GHz
- pamięć operacyjna 448 GB
- pamięć dyskowa ~5 TB
- typu blade
- wkrótce powiększenie o 392 rdzenie 2.33 GHz i 784 GB RAM



Sun Fire 6800 - *saturn*

system operacyjny:

Solaris 8

konfiguracja:

- 20 procesorów UltraSparc III z zegarem 900 MHz
- pamięć operacyjna 24 GB
- pamięć dyskowa 138 GB





Klaster CYFRONET- LCG2 - zeus

system operacyjny:

- Scientific Linux (SL) 4.x

HP Cluster Platform 3000 BL 2x220 - 256 węzłów typu podwójny "blade", 2 procesorowych, 4-rdzeniowych, Intel Xeon Quadcore 2.5 GHz, 2 GB RAM na rdzeń, 120 GB dysk na węzeł w dwóch rackach (310 miejsce na TOP500!)

konfiguracja:

- 2048 rdzeni Intel Xeon z zegarem 2.5 GHz
- 4096 GB pamięci operacyjnej
- 300 TB pamięci dyskowej
- teoretyczna moc ok. 25 Tflops







Systemy archiwizacji danych

- Moc obliczeniowej musi iść w parze z systemami składowania danych; zamiast rozbudowywać dyski lokalne poszczególnych komputerów postawiliśmy na scentralizowany system zasobów dyskowych oparty o macierze dyskowe i serwery dyskowe
- Dwie nowoczesne macierze HP EVA8000 i HP XP12000 - ponad 200 TB dysków poprzez dedykowaną sieć SAN (*Storage Area Network*)
- Zalety: szybkie połączenie FC, odporność na awarie poszczególnych dysków (RAID) i pełna redundancja połączeń i zasilania
- Działa też hierarchiczny System Składowania Danych FSE (*File System Extender*); dane zapisywane są automatycznie na kilku rodzajach nośników magnetycznych: dyskach i taśmach



Systemy archiwizacji danych cd.

- Od 2008 r. **serwery** dyskowe: 7 x SUN X4540 (*Thor*) - 48 TB w każdym, w sumie ponad 300 TB!
- Tanie i wydajne, połączenia po 10 GbEthernet
- Bardzo przydatne w DPM - systemie używanym w projekcie EGEE/WLCG, świetna skalowalność
- Dobrze fituje się z Lustre





Co się zmieniło od zeszłego roku

- + 2 piętra!
- Duże zmiany w bazie: 1 MW prądu, generator, klimatyzacja (160 + 120 kW), upgrade UPS, porządki na hali, kablologia, zmiany organizacyjne...
- Podwojenie mocy SGI 3700 (baribal)
- Zakup nowych klastrów HP (+ 20 Tflopsow)
- Duże zmiany w storage'u, nowe serwery dyskowe
- Wycofane z eksploatacji: SGI Onyx 300, HP Superdome, część klastrów



Plany na przyszłość

- Najbliższe:
 - Duży switch 10 Gigabit Ethernet
 - Dodatkowe klimatory
 - Lustre dla KDM i klastrów gridowych
 - Twiki i Request Tracker?
- Średnio dystansowe:
 - Nowe UPS-y!
 - Powiększenie klastra IBM HS21 (mars)
 - 1-2 serwery dyskowe
- Długofalowe:
 - Nowe serwery, raczej klastry niż SMP, raczej „blade” niż klasyczne; finansowanie?, Pl-Grid
 - Przeprowadzki?