Astrofizyka: Środowisko gridowe dla projektu LOFAR (wycofana)

Usługa Astrofizyka: Środowisko gridowe dla projektu LOFAR została wycofana

Krótki opis usługi

Usługa przeznaczona jest dla osób redukujących dane interferometryczne uzyskane za pomocą obserwacji wykonywanych interferometrem radiowym LOFAR (http://www.lofar.org/). Usługa pomaga w przygotowaniu danych do redukcji (np. flagowanie), ich kalibracji, oraz wykonania map radiowych. Na usługę składają się dedykowane skrypty do optymalizacji poszczególnych kroków obliczeń oraz obsługa obliczeń wielowątkowych dla oprogramowania LOFAR.

Aktywowanie usługi

Aby móc skorzystać z usługi należy mieć aktywne konto w portalu PL-GRID. Aktywowanie usługi wymaga dostępu do klastra obliczeniowego ACK Cyfronet Zeus. Należy zatem w pierwszej kolejności zaaplikować o usługę dostępową "Dostęp do klastra ZEUS", a następnie aktywować usługę "Środowisko gridowe dla projektu LOFAR".

Pierwsze kroki

W środowisku obliczeniowym należy załadować moduł LOFAR:

```
module load plgrid/apps/lofar/29388
```

oraz wykonać inicjalizację zmiennych:

```
. lofarinit.sh
```

Moduł ten ma w zależnościach między innymi moduły CASA core i Python.

Jeżeli moduł lofar/29388 nie jest już dostępny (przykładowo: z powodu zmiany wersji na nowszą) należy użyć polecenia:

```
module spider lofar
```

w celu wyświetlenia aktualnie dostępnych wersji oprogramowania, a następnie stosownie zmodyfikować podaną wyżej ścieżkę polecenia module load.

Operacje na modułach

Operacje na modułach możliwe są tylko w środowisku obliczeniowym (nie na maszynach dostępowych - UI).

Uruchamianie (ładowanie) modułu. Pozwala na korzystanie z programów i bibliotek zainstalowanych w ramach modułu. W przeciwnym wypadku są one niedostępne.

```
module load nazwa_moduu
```

Wyświetlanie listy załadowanych modułów:

```
module list
```

Wyświetlanie listy wszystkich dostępnych modułów (uwaga! długa lista, może potrwać; przerwanie Ctrl-C):

```
module avail
```

Informacje o module:

```
module show nazwa_moduu
```

Środowisko obliczeniowe (kolejki):

Wszelkie obliczenia oraz używanie programów interaktywnych wykonywane jest w klastrze w ramach kolejek zadań określających priorytet i czas wykonania zadania. Bezpośrednio na maszynach dostępowych (UI, np. zeus) można wykonywać operacje plikowe: tworzenie, modyfikowanie, usuwanie plików i kartotek. Poprzez środowisko UI możliwe jest też kopiowanie danych do i z klastra. UI pozwala też na wykonywanie niektórych bardziej złożonych operacji stricte plikowych, takich jak tworzenie i rozpakowywanie archiwów tar.

Dostępne jest kilka kolejek:
* *l_interactive* do zadań wymagających interakcji, także grafika przez X,
* *l_test* do zadań do 15 minut czasu wykonania (rzeczywistego).
* *l_prio* do zadań szybkich do 1 godziny,
* *l_short* do zadań do 3 godzin
* *l_long* do zadań do 3 tygodni

W przypadku kolejek stricte obliczeniowych (l_prio, l_short) wskazane jest uruchamianie zadań poprzez skrypty PBS:
`qsub -q nazwa_kolejki nazwa_skryptu`

Informacje o zadanich w danej kolejce, między innymi można odczytać numer zadania PBS_JOBID.
`qstat -q` nazwa kolejki

Lista zadań podanego użytkownika:

```
qstat -u username
```

Przerwanie i usunięcie zadania:

```
qdel PBS_JOBID
```

Uruchamianie kolejek

Uruchomiając zadanie w kolejce podaje się liczbę wymaganych nodów (węzłów obliczeniowych) i rdzeni na przykład 2 węzły po 12 rdzeni (wskazane jest używanie pełnych węzłów z maksymalną liczbą rdzeni, czyli 12) :

```
nodes=2:ppn=12
```

Można użyć bezpośrednio w linii poleceń, albo w skrypcie. Im więcej zażąda się nodów, tym dłuższy jest przeciętny czas uruchomienia zadania (nie wykonania). Do testowania i oglądania wyników wskazane jest użycie kolejki trybu interaktywnego. Uruchomienie na jednym rdzeniu na trzy godziny:
`qsub -IX -q l_interactive -l nodes=1:ppn=1 -l walltime=03:00:00`
Uruchomienie na 2 węzłach po 12 rdzeni z domyślnym czasem maksymalnym (3 godziny):
`qsub -IX -q l_interactive -l nodes=2:ppn=12`

Zaawansowane użycie

1. Łączenie poszczególnych subbandów

Przed wykonaniem kalibracji konieczne jest uzyskanie dobrego stosunku sygnału do szumu w danych obserwacyjnych. W tym celu pojedyncze subbandy łączone są w paczki. Przy założeniu, że np. subbandów z danych obserwacji jest 324, dobre wyniki uzyskuje się przy podziale na 12 paczek po 27 subbandów. Służy do tego polecenie uruchamiające program NDPPP:

```
concat.NDPPP.script <numer pierwszej obserwacji> <numer ostatniej obserwacji> 000
```

Dla danych testowych numery pierwszej i ostatniej obserwacji to 168883 i 168952. Program wymaga umieszczenia pliku z parametrami NDPPP-dummy.parset-proto w katalogu, w którym jest uruchamiany. Rezultatem pracy jest powstanie plików postaci *LXXXXXX_ALL_uv.dppp.MS.ndppp.ndppp*

2. Usuwanie oryginalnych tablic flagowych

Tablice flagowe zawierają informacje o odrzuconych fragmentach z surowych obserwacji, które nie będą brane pod uwagę przy kalibracji i tworzeniu mapy należy je „wyrzucić”. Na wielu danych LOFAR przeprowadzono błędną, testową metodę usuwania wpływ silnych źródeł promieniowania, przez co ich tablice mogą zawierać te niepotrzebne informacje. Aby je usunąć, można skorzystać ze skryptu:

```
casapy --nologger -c clearflag.py
```

Program tworzy pliki z dodaną nazwą flagversions, zawierające wcześniejszą, niewłaściwą tablicę flagową. Program wymaga umieszczenia pliku flagdata.last w katalogu, w którym jest uruchamiany.

3. Tworzenie nowych tablic flagowych

W usuniętych tablicach znajdują się zwykle użyteczne informacje, które przed dalszą redukcją danych należy przywrócić. Służy do tego polecenie:

```
python ndppp_generator.py
```

Program wymaga umieszczenia pliku `ndpppbody.file` w katalogu, w którym jest uruchamiany. Efektem jego poprawnego działania jest utworzenie się plików *NDPPP_LXXXXXX_ALL_uv.dppp.MS.ndppp.ndppp.parset*.

Następnie należy wywołać polecenie uruchamiające program NDPPP dla każdego ze stworzonych w poprzednim kroku plików i w rezultacie przeflagowanie wszystkich plików z obserwacjami:

```
python ndppp_trigger.py
```

4. Informacje o elementach antenowych

Dane mogą zawierać błędne informacje o elementach antenowych, co pogarsza jakość kalibracji. Aby je poprawić, należy użyć polecenia:

```
python fixinfo_trigger.py
```

5. Demixing

Jest to metoda mająca na celu usuwanie wpływu odległych i silnych radioźródeł, realizowana przy pomocy programu NDPPP. W pierwszym kroku należy utworzyć pliki typu parset za pomocą polecenia:

```
python demixing_generator.py
```

Wymaga to obecności plików demixing.body.file i demixing.interpart.file w katalogu, w którym jest uruchamiane. Po jego poprawnym wykonaniu powinny pojawić się pliki postaci:

```
demixing_LXXXXXX_ALL_uv.dppp.MS.ndppp.ndppp.parset
```

Następnie wywołuje się program

```
python ~demixing_trigger.py
```

Wymaga on wcześniejszego umieszczenia pliku Ateamhighresdemix.sourcedb (model źródeł, które należy usunąć z danych) w katalogu, w którym jest uruchamiany. Po jego zakończeniu powinny powstać pliki postaci:

LXXXXXX_ALL_uv.dppp.MS.ndppp.ndppp.demix.

6. Kalibracja

Kolejnym krokiem redukcji jest kalibracja, czyli wyliczenie poprawek pozwalających na dopasowanie danych do określonego modelu nieba. Uruchamia ją polecenie:

```
../pckg_loop.script <nazwa iteracji> <numer pierwszej obserwacji> <numer ostatniej obserwacji> <parset> <model>
```

Plik "model" musi znajdować się w katalogu, w którym uruchamiany jest skrypt. Dla danych testowych należy wybrać obserwacje o numerach 168883 – 168952, model SQ_gsm_7.5_0.2.model oraz parset bjorn.bbs_SQ.gsm.parset . Program w czasie pracy tworzy logi zawierające informacje o tym, jak przebiegała kalibracja (za pomocą tee). Poprawnie zakończony wyświetla komunikat

```
BBS-reducer terminated successfully
```

Po jego zakończeniu (obliczenia mogą trwać do kilku godzin) należy sprawdzić jakoś kalibracji, np. korzystając z polecenia

```
cat <element wspólny dla nazw logów>* | grep conve .
```

Polecenie to wyświetla listę iteracji wraz z informacją, ile z nich zakończyło się sukcesem. Duża ilość zatrzymanych (kolumna stopped) iteracji w stosunku do poprawnie zakończonych (converged) oznacza zwykle, że źle dobrano parametry kalibracji.

Gdzie szukać dalszych informacji?

Dalszej pomocy, jak również informacji na temat praktycznego zastosowania oprogramowania do redukcji danych LOFAR, można szukać na stronach LOFAR:

- http://www.lofar.org/wiki/doku.php?id=start (konieczna rejestracja)
- http://www.lofar.org/
- http://www.astron.nl/radio-observatory/lofar/lofar-imaging-cookbook (podręczniki - przewodniki po redukcji danych interferometrycznych, aktualizowane co kilka miesięcy).