

TOP500: la llista dels ordinadors més potents del món

Pep Turró Mauri

<2018-04-16 Mon>

Descripció

Un registre dels ordinadors comercials més potents del món al llarg dels anys que permet estudiar l'evolució de la tecnologia dels computadors d'altres prestacions.



Figura 1: El superordinador MareNostrum del Barcelona Supercomputing Centre. 16è a la llista TOP500 de Novembre de 2017.

1 Context

Les dades obtingudes mitjançant web scraping són un històric del llistat dels superordinadors més potents del món, en funció del seu rendiment a les proves [LINPACK](#). Aquestes llistes les manté el [projecte TOP500](#), que publica dos cops l'any (cada Juny i Novembre) la llista actualitzada dels 500 ordinadors amb millor rendiment Linpack del moment.

2 Contingut

El conjunt de dades conté les entrades de les diferents llistes TOP500, amb un registre per cada entrada a una llista. Cal dir que un mateix sistema sol aparèixer en més d'una llista: des de la seva entrada al TOP500, en edicions posteriors pot anar baixant, és a dir, pot aparèixer de nou en una posició diferent en un nou registre.

2.1 Procés de descàrrega

El **dataset adjunt al repositori** inclou les dades completes de les 50 edicions de la llista TOP500 publicades fins ara, és a dir, $50 \times 500 = 25.000$ registres. El **codi de l'scraper** permet triar quines llistes i quants registres volem. La captura del dataset adjunt es va fer especificant la data de la primera llista (Juny de 1993) com a punt de partida (per defecte, el programa captura des del punt de partida fins a la darrera publicada, que es calcula en temps d'execució en funció de la data actual):

```
$ time python3 -u scrape.py --year 1993 --month 6 > scraping.log
```

```
real 57m7,175s
user 11m38,674s
sys 0m4,049s
```

Aquest procés va comportar la descàrrega de 13.524 pàgines. L'scraper no descarrega la mateixa pàgina dues vegades, es desa els valors ja recuperats internament per quan cal reutilitzar-los, per exemple quan es torna a trobar el mateix ordinador en una llista posterior:

```
$ grep Download scraping.log | wc -l
13524
```

2.2 Estructura de les dades capturades

Per a cada registre desem 28 camps, tot i que uns quants d'ells tenen una proporció elevada de valors buits degut a que la informació no està disponible. Els camps són els següents (agrupats per temàtica/tipus):

- Identificadors: 6 camps amb informació per identificar el superordinador llistat:
 - **system_id**: l'identificador numèric que la llista TOP500 fa servir per referir-se a un ordinador en concret. Ens serveix per identificar un ordinador, ja que cap altre camp és prou fiable per fer-ho.
 - **site_id**: l'identificador numèric que TOP500 fa servir per identificar el centre de supercomputació on es troba un ordinador.
 - **system_url** és l'enllaç a la web de l'ordinador (alguns superordinadors tenen lloc web propi, com per exemple el [MareNostrum](#) de la figura 1).
 - **site_name**: el nom del centre de supercomputació.

- **site_url**: l'enllaç a la web del centre de supercomputació.
- **name**: el nom de l'ordinador. Més detalls a [origen dels camps](#).
- Descriptius quantitativs: 8 camps amb dades numèriques:
 - **rmax**: velocitat màxima aconseguida (TFlops) a les proves Linpack. Aquest valor és el que utilitza la llista TOP500 per elaborar el rànquing.
 - **rpeak**: velocitat màxima teòrica (TFlops).
 - **cores**: número de nuclis de computació.
 - **memory**: quantitat de memòria (GB).
 - **nmax**: mida del problema per assolir *rmax*.
 - **nhalf**: mida del problema per assolir la meitat de *rmax*.
 - **hpcg**: resultat de proves [HPCG](#) (TFlops).
 - **power**: potència consumida (kW).
- Descriptius qualitativs: 11 camps textuais categòrics que permeten classificar els sistemes:
 - **segment**: àmbit d'ús/aplicació del superordinador (universitari, fabricant, govern, recerca, indústria, classificat, altres).
 - **country, city**: país i ciutat on està ubicat el superordinador.
 - **manufacturer**: fabricant.
 - **processor**: processador principal.
 - **gpu**: processador secundari, co-processador.
 - **interconnect**: tipus de xarxa interna.
 - **os**: sistema operatiu.
 - **compiler**: compilador.
 - **math**: llibreria matemàtica.
 - **mpi**: llibreria de pas de missatges.
- Posició en les llistes TOP500: 3 camps que identifiquen l'edició de la llista i la posició que ocupa: **year, month, rank**.

2.3 Origen dels camps

Els diferents camps provenen de diferents parts del lloc web.

De les pàgines dedicades als centres de computació (p.e. <https://www.top500.org/site/49748>) n'obtenim **site_name** (a la capçalera), **site_url**, **country**, **city** i **segment** (a la taula de detalls).

De les pàgines que mostren detalls de cada superordinador (com per exemple <https://www.top500.org/system/179067>) n'obtenim de la taula de detalls: **system_url**, **manufacturer**, **memory**, **processor**, **interconnect**, **nmax**, **nhalf**, **hpcg**, **os**, **compiler**, **math** i **mpi**.

Les dades de rendiment i potència no les agafem d'aquesta taula, ja que hi ha sistemes que evolucionen al llarg del temps (s'amplia la seva capacitat, i

en conseqüència el seu consum) i això no ho reflecteix la pàgina de detalls del sistema. Aquesta és una petita inconsistència a la web de TOP500.

Així doncs, els camps **rmax**, **rpeak**, **cores** i **power** els agafem de les pàgines amb la llista (p.e. <https://www.top500.org/list/2016/06/?page=2>), juntament amb la posició que ocupa cada superordinador: la variable **rank**.

Alguns camps els obtenim de les adreces/enllaços (*href*) de les diferents pàgines comentades anteriorment: **system_id**, **site_id**, **year** i **month**.

Hi ha dos camps que són calculats a partir d'analitzar el contingut del valor de la columna *System* a les pàgines del llistat: el nom de l'ordinador (**name**) i els co-processadors (**gpu**). Aquesta informació no es troba detallada de forma individual enlloc més, i malauradament és també impossible extreure el valor d'aquests dos camps de manera completament fiable a partir només de la informació que proporciona el lloc. Això és perquè el valor de la columna *System* té un contingut força inconsistent al llarg de les múltiples llistes, i pot contenir una combinació de processador principal, interconnector, nom del sistema i múltiples co-processadors, en un ordre que no sempre és el mateix.

Per tant, el contingut d'aquests dos camps al conjunt de dades conté algunes imprecisions. Per arreglar-les caldria combinar la informació disponible amb altres fonts de dades, com per exemple una base de dades de co-processadors coneguts. Aquest processament va més enllà de l'abast d'aquest projecte.

Finalment, comentar que a les pàgines de detalls de sistemes hi trobem, en alguns pocs casos, dos detalls més que he decidit no capturar: *Power Measurement Level* i *Measured Cores*. D'una banda, no s'explica el significat d'aquests camps i per tant en desconec la utilitat. D'altra banda, el valor de *Measured Cores* sol correspondre al valor de *Cores* que ja es captura. I, a més, són presents només en una minoria de sistemes: el programa d' scraping informa dels detalls ignorats, i això permet saber aquesta proporció:

```
$ grep 'Measured Cores' scraping.log | sort -u | wc -l
90
$ grep 'Power Measurement' scraping.log | sort -u | wc -l
42
```

Aquestes xifres representen un 0.86% i un 0.4% respectivament del total de sistemes.

3 Agraïments

Les dades recollides provenen del [projecte TOP500](#), a qui cal agrair la feina de recopilació i ordenació, així com el fet de fer-les accessibles al seu lloc web.

Existeix un projecte amb objectius similars (processar les dades de les llistes TOP500): <https://github.com/dlenski/top500>. Aquest altre projecte fa scraping de les pàgines HTML, sino que es basa en els arxius en format Excel que el projecte TOP500 proporciona a usuaris registrats.

La fotografia de la figura 1 prové de [Wikimedia Commons](#), sota llicència Creative Commons BY-SA.

4 Inspiració

Les dades recopilades són útils per fer estudis de l'evolució de la computació d'altres prestacions al llarg dels darrers 25 anys. En veiem un exemple al projecte citat als agraïments, que va estar motivat per realitzar les gràfiques d'evolució de les arquitectures de processadors de supercomputadors que actualment són visibles a la [Wikipedia](#).

Més enllà d'aquest exemple, es poden fer anàlisi d'altres factors a partir de les diferents variables capturades: països que amb recursos invertits en supercomputació, sistemes operatius utilitzats, evolució de la potència, etc.

5 Llicència

El **codi** de l'scraper està llicenciat amb la llicència [GPLv3](#).

El **conjunt de dades** adjunt es proporciona sota llicència [CC0 1.0 Public Domain](#).

Els motius per aquesta elecció són:

- La voluntat de contribuir el codi com a programari lliure.
- La llicència del codi que processa les dades no aplica a les dades que processa, i aquestes tenen una llicència pròpia.
- La voluntat de maximitzar l'ús lliure de les dades recopilades, que són de naturalesa factual.

Hi ha una explicació detallada sobre els motius a la [wiki de Creative Commons](#) sobre tractament de dades.

6 Codi

El codi que duu a terme el web scraping està desenvolupat en Python i disponible en un [repositori a github](#). El programa principal és `scrape.py` i fa servir el mòdul ubicat a la carpeta `top500` per trobar el contingut i fer l'scraping pròpiament dit.

Executant `scrape.py` sense opcions addicionals generarà un fitxer anomenat `top500.csv` amb el contingut sencer (500 entrades) de la darrera edició de la llista TOP500. El programa té opcions per controlar quines edicions descarregar, el nombre d'entrades per edició, i el fitxer de sortida. Més detalls d'ús i algunes notes d'implementació es troben al fitxer `README` del repositori.

7 Dataset

El fitxer CSV amb totes les edicions de la llista TOP500 fins a dia d'avui és a [data/top500.csv](#).

Podem veure un resum de l'estructura de l'arxiu carregant-lo amb R:

```
col.classes <- c(
  'year'      = 'factor',
```

```

    'month'      = 'factor',
    'site_id'    = 'factor',
    'system_id'  = 'factor',
    'name'       = 'character',
    'site_name'  = 'character',
    'system_url' = 'character',
    'site_url'   = 'character'
  )
top500 <- read.csv('data/top500.csv',
                  na.strings = '',
                  colClasses = col.classes)

str(top500)

'data.frame': 25000 obs. of  28 variables:
 $ site_name      : chr  "Los Alamos National Laboratory" "Minnesota Supercomputer Center" "N
 $ system_url     : chr  "http://www.acl.lanl.gov/" NA NA NA ...
 $ manufacturer: Factor w/ 124 levels "Acer Group","ACTION",...: 118 118 118 118 81 81 118
 $ cores          : int   1024 544 512 512 4 4 256 512 16 16 ...
 $ memory         : int   NA NA NA NA NA NA NA NA NA NA ...
 $ processor      : Factor w/ 463 levels "Alpha 1.25GHz",...: 312 312 312 312 130 129 312 74
 $ interconnect: Factor w/ 71 levels "100G Ethernet",...: 28 28 28 28 50 50 28 NA NA NA ...
 $ rmax           : num   59.7 30.4 30.4 30.4 23.2 20 15.1 13.9 13.7 13.7 ...
 $ rpeak          : num   131 69.6 65.5 65.5 25.6 22 32.8 20.5 15.2 15.2 ...
 $ nmax           : int   52224 36864 36864 36864 6400 6144 26112 25000 10000 10000 ...
 $ nhalf          : int   24064 16384 16384 16384 830 832 12032 7500 650 650 ...
 $ hpcg           : num   NA NA NA NA NA NA NA NA NA NA ...
 $ power          : num   NA NA NA NA NA NA NA NA NA NA ...
 $ os             : Factor w/ 72 levels "AIX","Bullx Linux",...: 7 7 7 7 50 50 7 26 62 62 ...
 $ compiler       : Factor w/ 100 levels "CCE","cce 8.0",...: NA NA NA NA NA NA NA NA NA NA ...
 $ math           : Factor w/ 70 levels "-", "2017.0.2",...: NA NA NA NA NA NA NA NA NA NA ...
 $ mpi            : Factor w/ 103 levels "bullx MPI","bullxmpi 1.1.11.1",...: NA NA NA NA NA NA NA NA NA NA ...
 $ gpu            : Factor w/ 508 levels "1 GHz","1.25 GHz",...: NA NA NA NA NA NA NA NA NA NA ...
 $ country        : Factor w/ 59 levels "Australia","Austria",...: 57 57 57 57 26 7 57 57 57 5
 $ site_id        : Factor w/ 2853 levels "47223","47224",...: 809 881 980 925 927 107 989 233
 $ system_id      : Factor w/ 10421 levels "166659","166660",...: 188 248 247 246 3217 3216 21
 $ name           : chr   "CM-5/1024" "CM-5/544" "CM-5/512" "CM-5/512" ...
 $ site_url       : chr   "http://www.lanl.gov/" "http://www.msi.umn.edu/" "http://www.nsa.gov
 $ city           : Factor w/ 679 levels "Aachen","Aarheiligen",...: 349 387 NA 630 608 167 65
 $ segment        : Factor w/ 7 levels "Academic","Classified",...: 6 4 2 1 7 6 6 1 7 6 ...
 $ year           : Factor w/ 25 levels "1993","1994",...: 1 1 1 1 1 1 1 1 1 1 ...
 $ month          : Factor w/ 2 levels "11","6": 2 2 2 2 2 2 2 2 2 2 ...
 $ rank           : int    1 2 3 4 5 6 7 8 9 10 ...

```

Del resum en podem destacar:

- Un total de 25.000 observacions (500 entrades a cada llista, amb 50 llistes descarregades) amb 28 variables.
- Tenim dades de 10.421 ordinadors (*system_id* únics) que pertanyen a 2.853 centres (*site_id* únics) ubicats a 679 ciutats de 59 països.

8 Referències

- Meuer, H., Strohmaier, E., Dongarra, J., Simon, H., Meuer, M. TOP500 project. <https://www.top500.org/>
- Python Software Foundation. Python Language Reference, version 3.6. <https://www.python.org/>
- Richardson, Leonard (2015). Beautiful Soup Documentation. <https://www.crummy.com/software/BeautifulSoup/bs4/doc/>
- Reitz, Kenneth (2018). Requests: HTTP for Humans. <http://docs.python-requests.org/en/master/>
- Lawson, Richard (2015). Web Scraping with Python. Packt Publishing Ltd.
- R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.