This project is supported by the European Commission's FP7 Capacities Programme for the period April 2013 - March 2017 under the Grant Agreement number 312495.

www.solarnet-east.eu

# DELIVERABLE D100.1

# Assessment on Access to Databases

WP100 Access to Science Data Centres. Space Missions

1ST Reporting Period

November 2014

This project is supported by the European Commission's FP7 Capacities Programme for the period April 2013 - March 2017 under the Grant Agreement number 312495.

# PROJECT GENERAL INFORMATION

Grant Agreement number: 312495

Project acronym: SOLARNET

Project title: High-Resolution Solar Physics Network

Funded under: FP7-INFRASTRUCTURES: INFRA-2012-1.1.26 - Research Infrastructures for High-Resolution Solar Physics

Funding scheme: Combination of Collaborative Project and Coordination and Support Action for Integrating Activities

From: 2013-04-01 to 2017-03-31

Date of latest version of Annex I against which the assessment will be made: **13/02/2013**

Periodic report: 1st **X** 2nd ☐ 3rd ☐ 4th ☐

Period covered: from **01/04/2013** to **30/09/2014**

Project's coordinator: Dr. Manuel Collados Vera, IAC.

Tel: (34) 922 60 52 00

Fax: (34) 922 60 52 10

E-mail: mcv@iac.es

Project website address: http://solarnet-east.eu/

This project is supported by the European Commission's FP7 Capacities Programme for the period April 2013 - March 2017 under the Grant Agreement number 312495.

# WP 100 Access to Science Data Centres. Space missions.

Leader organization: Universitetet i Oslo (UiO)

Participants: MPG, UiO, ROB

The description of Work specifies the following for WP 100:

*"Access to the most demanded European Science Data Centre, providing data gathered by the solar satellite HINODE, and the Solar Dynamics Observatory (SDO), will be offered. As previously mentioned, this access to space-based data for solar research, offered for the first time under an I3 supported by the EC, will expand significantly the content of the Programmes for the high-resolution solar physics community supported in the past.*

*The Hinode Science Data Centre Europe, the Belgian Web Incessant Screening and the German Science Center for the Solar Dynamics Observatory, are the three facilities offering access and services under this WP101, WP102 and WP103.*

*The main efforts, concerning these web-based databases, will be to widely advertise these resources, to increase the number of EU researchers accessing to them.*

*Modality of Access*

*The previously mentioned project web-page will include also reference and full information about these WP10 activities. Some specific information about the modality of access, support offered and periodical assessment is provided under the facilities' description.*

*Outreach of new users*

*As a complement of the publicity made through the project website, a set of publicity actions will be undertaken to attract users, specially targeted groups (new users):*

*– Participation in conferences, info days and meetings. Brochures, talks, etc.*

*– General advertisements. A special effort will be made to inform about the new facilities and instruments offering access for the first time under en EC programme: HINODE-UiO, BEWISSDOM and GSC-SDO (and the same applies for the access offered (WP9.2, WP9.5 and WP9.6) to GREGOR, IBIS/DST and ROSA/DST)."*

The centers have all streamlined the access to the facilities and have significantly added new resources to the databases through the SOLARNET project.

The three centers have increased awareness of the existence of the space-based data to the solar physics community in Europe. This has been accomplished through participation in conferences and meetings. Several dedicated talks were given at the 1st SOLARNET meeting in Oslo August 5-8 2013 and at the 1st CASSDA-SOLARNET workshop meeting in Freiburg February 18-20 2014. Hands-on exercises for the search and retrieval of IRIS data from the Hinode Science Data Centre at UiO was given at the 1st SOLARNET School in Wroclaw March 24 – April 4 2014.

More details about the individual centers are given in the following reports from the sub-Workpackages 101, 102 and 103.

This project is supported by the European Commission's FP7 Capacities Programme for the period April 2013 - March 2017 under the Grant Agreement number 312495.

# SubWorkpackage 101: HINODE/IRIS. Access to Science Data Centres

Leader Organization: Universitetet i Oslo (UiO)

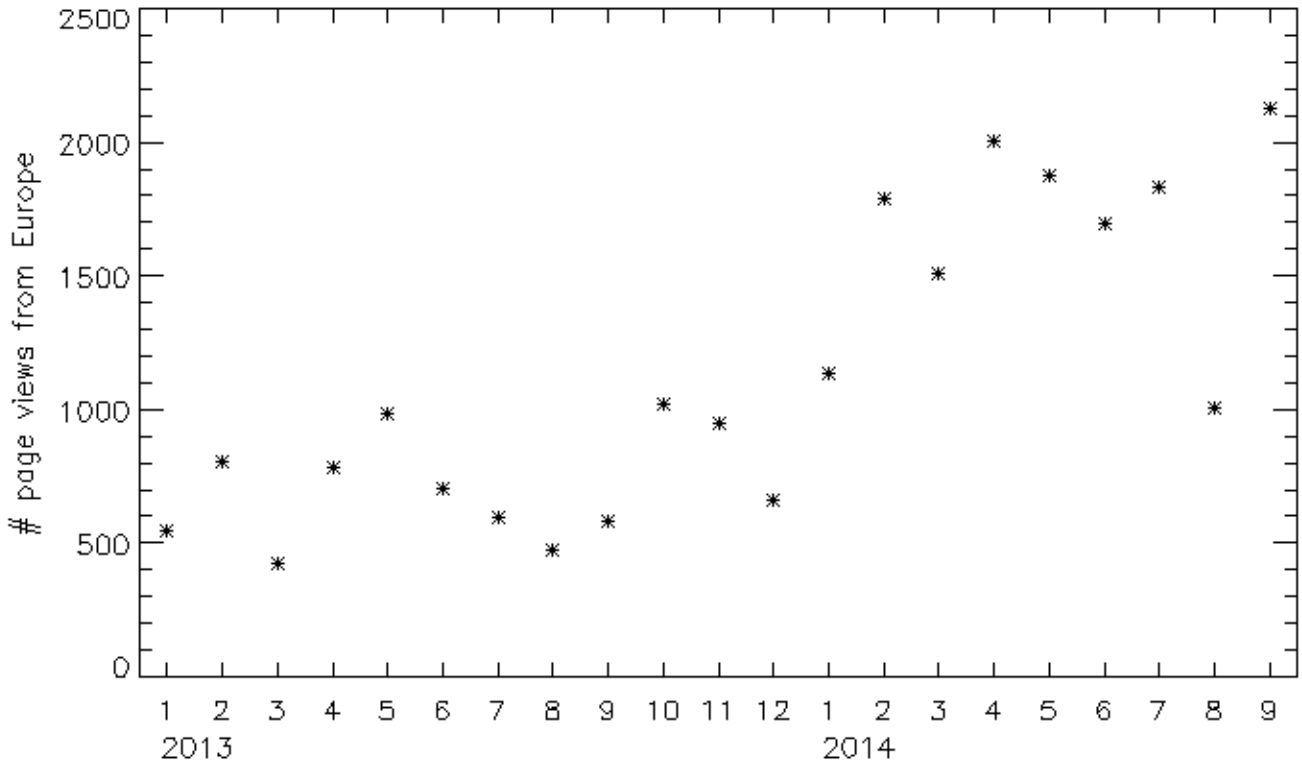
Participants: UiO

We first give a summary, then detailed monthly reports on the activity in the Science Data Centre followed by detailed usage statistics (complete data available from http://sdc.uio.no/sdc/webstats).

**Summary**

Access to the Hinode Science Data Centre was provided throughout the period. The most significant improvement came through the addition of all data from the Interface Region Imaging Spectrograph (IRIS). The data was made available from the first day of open data policy, October 31st 2013. Additional search possibilities were added and the data presentation was improved through improved thumbnail pictures.

At the start of the SOLARNET Project the average number of distinct hosts accessing the data centre was around 600-700 per month. This has increased throughout the Project period and is now 1100-1400. The same increase is evident in the number of page-views from Europe where the average monthly numbers have increased from 590 (January-March 2013) to 1655 (July-September 2014), see figure




.

This project is supported by the European Commission's FP7 Capacities Programme for the period April 2013 - March 2017 under the Grant Agreement number 312495.

## Monthly reports from the Hinode SDC

**April 2013**

Most of the effort in April was focused on IRIS, both in terms of discussing fits file keywords with the IRIS team, and in preparing for the inclusion of IRIS files in the archive.

The generalization of the file-fetching routines has now been completed. Modifying the quick-look image production software is proceeding, and we now also have some dummy fits files to "practice" on.

Simulated observations of numerical simulations (produced in Oslo) are an integral part of the IRIS project, and we have started preparations for inclusion of these "observations" as well.

**May 2013**

In May (as in April), most of the focus was on IRIS. We have now started including IRIS keywords in the database definition file, which is also used for stripping information out of the fits headers. Modifications to the quick-look software continue. These modifications are to a large extent analogous to the database modifications – having two spectrometers highlights the need for a class of "spectrometer-common" features (i.e. keywords/tables & data access/manipulation methods). In the database, this ideal solution is not implemented at this time, but it will be implemented in the quick-look software.

In mid-May we encountered a problem that should have been caught before it happened: The file id numbers reached the ceiling for what can be stored in an unsigned 3-byte integer (16,7 million). We only have 15 million files, but a large number of file ids have been used up by e.g. XRT files that change name from quick-look to level 01. Changing the type used to store file ids is in principle a simple task – even though we have over 1600 tables. But due to our lack of experience with this kind of operation, some MySQL technicalities caused the operation to take about one week to complete. The archive was available to users the entire time, but no new files were ingested during this period.

**June 2013**

Besides "business as usual", the inclusion of IRIS data in the archive is progressing. The experience is proving to be very useful: Not only will it allow Hinode data users to see and compare Hinde and IRIS observations at a glance, but it is also a great learning experience with respect to the inclusion of other "archive-enriching" data sets in the future

**July-August 2013**

We are on track to include IRIS data in the archive at some point during the first half of September, assuming we get the go-ahead from the IRIS team. We will not have quicklook images at that point, however, because our Science Data Specialist is on a 5-week sickleave.

While revising our system to include IRIS data, we have used the opportunity to tidy up some of the middleware programs to increase readability and maintainability. We also have plans to improve the

functionality by changing the method used to pick fields to be displayed or for entering search criteria. This is currently done with an extremely long multiple-selection list for display fields, and on/off buttons for large groups of search criteria. The new method will be using a drop-down auto-complete/auto-suggest text box.

**September 2013**

We are ready to include IRIS data in the archive at relatively short notice as soon as we get the go-ahead from the IRIS team. We estimate that it will take a few days at most to iron out small bugs that will most likely appear when IRIS is included in the normal cycle of fetching, metadata extraction, and ingestion into the regular database (not a test database).

Our Science Data Specialist is now partially back from sick leave, so there is some progress being made on the quicklook image production for IRIS. It is unclear, however, exactly how long it will take before we have images for the archive.

While waiting for the go-ahead to serve IRIS data, we are using the opportunity to develop new functionality for the archive, since we have reviewed a significant part of the middleware.

The developments include a drop-down search/suggest entry field for adding new columns to be shown in the result listing. The entry field appears both on the query form and on the result page. The same mechanism has been implemented for adding search criteria to the query form. The mechanism is a "globbing" (wildcards * and ? permitted), case-insensitive substring search. Entering e.g. "ir_*status" will match strings like "IR__FLT : Flat field status" (not an actual field!). The earlier mechanisms were very cumbersome – an incredibly long multiple selection list for fields to be shown, and on/off buttons for showing or hiding very large groups of criteria in bulk.

The chosen extra criteria are displayed line by line together with its description, which used to be shown only when hovering over the input fields. This is not done with input fields added in bulk with the on/off buttons, because it would take up very much space on the query page.

The next feature to be implemented is an "expand result to include all files in its group" option. I.e. an entire group of files is to be selected if *one or more* files in that group match(es) the search criteria. This can be very useful when searching for groups that contain one or more observation where e.g. a flare flag has been raised.

**October 2013**

IRIS data was made available through the Hinode archive on 31. October. The structure of the source directory at Lockheed Martin had been changed only a few days before: The released data are now stored compressed – with raster files tar'ed together before compression, and slit-jaw files compressed individually.

The source directory is continuously monitored and mirrored with multiple parallel rsync connections. To maintain compatibility with the fetching system, a decompression and unpacking script had to be written to make another copy with a structure that is analogous to the Hinode directories.

For technical reasons, however, this copy cannot yet be handled as if it were "just another Hinode source directory". A normal Hinode fetch cycle looks only 20 days backwards in time for quicklook and level 0 files. The limit is both for efficiency reasons1 and due to the problem of "disappearing quicklook files": some quicklook files change name/identity during level 0 processing due to clock adjustments. A process that runs every two days removes quicklook files older than 20 days if they have not been replaced by level 0 files.

IRIS data released so far stems from a period prior to the 20-day period, and processing continues *backwards* in time. Extending the look-back time during normal fetch cycles to catch "old" IRIS files would re-fetch a large number of outdated/deleted Hinode quicklook files. So until IRIS data processing has caught up, IRIS file fetching has to be handled semi-manually.

The improved front-end/middleware was put on-line at the same time as the IRIS data release. With our science data specialist being busy with the quick look image production software over the last few weeks, the new software had been subject to much less beta testing than we would have liked. Since the release, our data specialist has have discovered one bug in the new middleware, and certain log messages indicates that there may be one additional bug. It seems like no external users have triggered or noticed these bugs thus far.

The option to "expand result to include all files in parent groups" was implemented before the release of the new middleware, and seems to work quite well.

We will soon be able to offer quick look images for the IRIS data.

**November 2013**

With IRIS data and software upgrades on-line, the effort has turned to more extensive testing and bug fixing.

The Lockheed Martin pipeline is still not mature enough to have a thorough test of "normal operations", but it is getting closer and closer.

The web statistics software had to be modified to work with the software upgrade otherwise it would not count the traffic using the new URLs. A notable increase in traffic can be seen in November. There is a small increase in October as well, which reflects traffic on our test server from the LMSAL/UiO IRIS teams. Some of the traffic increase, however, may be due to the fact that people using the old URLs will be registered, then redirected to the new URL.

We now have quick look images for most of the IRIS data, but work continues to improve the coverage. For IRIS/FUV, the current quick look images only show integrated intensity, but we are working on the production of EIS-style intensity/velocity/line width images. For IRIS/NUV we currently show intensity

images of the Mg II k central line depression. In the future we will show both intensity and velocity images for the three most significant line features (the emission peaks and the central absorption) of the two Mg II lines in the NUV band.

**December 2013**

The Lockheed Martin pipeline producing IRIS data is now *mostly* stable, but there are occasional hiccups and issues. However, we expect to see several changes to the level 2 fits files in the coming weeks/months – the developer of the IRIS level 2 formatting software has been on an extended vacation, and will soon start implementing requested/planned changes. Based on our experience this will most likely cause occasional hiccups on our end.

We have not yet implemented a fix for the web statistics over-count caused by the use of old URLs that are redirected to the new one. The web statistics are therefore likely to be somewhat overstated.

Work continues to improve the coverage of quick-look images for IRIS.

The two middleware bugs mentioned in the report for October have been located and fixed. Both of them were due to the introduction of IRIS into the archive – not from the introduction of new features. One was caused by the fact that a single IRIS file could go above the size limit for packaged files, and the other one had an unusually simple fix: three letters, "|IR" had to be inserted into a regular expression in order to recognize the prefix for IRIS-specific data columns.

**January-February 2014**

We are still expecting a reprocessing of IRIS level 2 fits files to occur "soon".

The possibility of overcounted web statistics due to redirection of old URLs - mentioned in the previous report - turns out to be non-existent, since redirections were already being filtered out.

Work continues to improve the coverage of quick-look images for IRIS.

**March 2014**

We are still expecting a reprocessing of IRIS level 2 fits files to occur "soon". Work continues to improve the coverage of quick-look images for IRIS. We have discovered that some of the files in the LMSAL official release directory are "dummy" files/placeholders with no data due to missing telemetry, which of course makes it impossible to create quick-look images. We intend to remove these from the database until they have been reprocessed. We have also encountered some errors in the IRIS "prepping" software that were quickly fixed by the IRIS team.

**April 2014**

The number of IRIS files that fail during QL image production is now down to about 3%. The vast majority of these files fail b/c of inherent flaws (typically, they are missing almost all the data, or they are special calibration studies). Although LMSAL still has not reprocessed the level 2 fits files, we have now reprocessed the current archive with the new QL software.

This project is supported by the European Commission's FP7 Capacities Programme for the period April 2013 - March 2017 under the Grant Agreement number 312495.


**May-June 2014**

The number of IRIS files that fail during QL image production is at about 5% - the vast majority of which fail b/c of inherent flaws (missing data, special calibration files, etc). LMSAL has now reprocessed their files. Due to a bug in our two-step fetching system (we unpack tar'ed files from LMSAL), the IRIS/SPEC files were not refreshed. The bug has now been corrected, and the files are being refreshed.

The data centre is set up with three database servers: a primary one and a secondary one, both running our special storage engines. The third machine runs using only standard MySQL, as a "ground truth" against which the others are compared. Some weeks ago, it was discovered that the secondary machine would get "out of sync" with the two others. It was found that the issue is related to deletion of files (files that have changed names due to clock corrections between QL and level 0 formatting rounds). Since the primary machine stays in sync with the pure MySQL machine, it seems like the issue is not caused by the storage engines. The issue is being investigated.

**July-August 2014**

The issue with the secondary database server getting out of sync compared with the others has now been fixed (see last report for details). In fact, two separate bugs were the cause of the issue. The reason why the symptoms became visible only now was due to a clean-up of surrounding code that exposed the "vulnerability". The bugs were in the system/control layer of the software, not in the custom-made storage engines.

**September 2014**

Since the identification and correction of two bugs in the system/control software, there have been no further cases of the databases getting out of sync (see reports for June, and July/August for details).

In creating statistics plots for the upcoming Hinode extension proposal and a SOLARNET report, we have discovered a "suspicious feature":

In May, Iran seems to have downloaded about 2.7 terabytes of data, and in July, China seemingly downloaded 2.3 terabytes, while the grand total downloaded over the lifetime of the archive is "only" 10 terabytes. In other words, those two months doubled the total downloaded volume! We will investigate this to see if this is "real", a bug in the statistics software, or some new access pattern that gives misleading figures and should be filtered out, as with the "China syndrome" (see report for December 2010).

Also noteworthy (but most likely unrelated) is that that several statistics measures have significantly increased in September (distinct hosts accessing site/search/result pages, number of IDL searches, number of single-file downloads, and the number of downloads).

We will investigate this together with the peculiarities regarding download volumes in May and July.

There seems to be an issue with the creation of quick-look images for EIS files that started around the last week of September. Initial investigations indicate that the process crashes inside the EIS prepping software, however. This will be followed up.

## Access statistics (taken from http://sdc.uio.no/sdc/webstats):

| Jan-2013 | Feb-2013 | Mar-2013 | Apr-2013 | May-2013 | Jun-2013 | Jul-2013 | Aug-2013 | Sep-2013 | Oct-2013 | Nov-2013 | Dec-2013 | SUM-2013 | Jan-2014 | Feb-2014 | Mar-2014 | Apr-2014 | May-2014 | Jun-2014 | Jul-2014 | Aug-2014 | Sep-2014 | Oct-2014 | Nov-2014 | SUM-2014 | *Average*/Sum | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 798 | 690 | 659 | 694 | 701 | 720 | 771 | 649 | 546 | 711 | 940 | 1248 | 4842 | 1021 | 932 | 1107 | 1213 | 1217 | 1148 | 1114 | 1060 | 1415 | 1302 | 245 | 6360 | *488* | Distinct hosts accessing site |
| 120 | 125 | 120 | 160 | 155 | 176 | 154 | 116 | 143 | 178 | 261 | 178 | 1217 | 162 | 147 | 207 | 198 | 197 | 192 | 159 | 159 | 237 | 195 | 26 | 1300 | *148* | Distinct hosts accessing search page |
| 100 | 97 | 100 | 129 | 135 | 152 | 143 | 113 | 125 | 153 | 202 | 165 | 1002 | 173 | 145 | 177 | 169 | 185 | 204 | 142 | 128 | 228 | 172 | 19 | 1155 | *109* | Distinct hosts accessing results pages |
| 18 | 28 | 32 | 35 | 54 | 33 | 46 | 24 | 31 | 39 | 46 | 48 | 325 | 50 | 38 | 46 | 55 | 48 | 47 | 40 | 30 | 40 | 31 | 10 | 310 | *29* | Distinct hosts downloading packaged files |
| 1252 | 1887 | 2248 | 2296 | 2183 | 1217 | 1233 | 1221 | 1147 | 1850 | 1918 | 1408 | 19860 | 2143 | 2359 | 2409 | 3021 | 3076 | 3435 | 5272 | 3204 | 4189 | 2133 | 307 | 31548 | **177757** | Number of result page views |
| 98% | 99% | 98% | 97% | 89% | 94% | 96% | 98% | 97% | 97% | 98% | 98% | 97% | 98% | 99% | 99% | 99% | 99% | 99% | 99% | 99% | 99% | 94% | 99% | 99% | *75%* | Fraction of result pages using thumbnails |
| 75% | 61% | 59% | 72% | 86% | 92% | 95% | 63% | 89% | 71% | 76% | 89% | 76% | 87% | 96% | 81% | 85% | 87% | 96% | 97% | 95% | 89% | 89% | 98% | 91% | *73%* | Fraction of result pages using EPOCH |
| 1477 | 1201 | 1280 | 652 | 3160 | 158 | 146 | 3 | 0 | 60 | 142 | 141 | 3998 | 124 | 88 | 57 | 44 | 113 | 74 | 78 | 89 | 605 | 96 | 12 | 1380 | **119741** | Number of (IDL) client searches |
| 63 | 74 | 31 | 169 | 116 | 67 | 61 | 60 | 52 | 117 | 91 | 80 | 981 | 230 | 53 | 95 | 69 | 174 | 89 | 106 | 132 | 170 | 130 | 22 | 1270 | **86430** | Number of fits header views |
| 10194 | 128 | 115 | 213 | 7283 | 180 | 1268 | 568 | 39 | 1264 | 392 | 5400 | 27044 | 219 | 3108 | 5448 | 209 | 13542 | 361 | 7465 | 4012 | 120290 | 20534 | 42 | 175230 | **1694049** | Number of single-file downloads |
| 6 | 6 | 2 | 26 | 102 | 50 | 9 | 16 | 8 | 13 | 8 | 6 | 252 | 5 | 10 | 13 | 27 | 13 | 9 | 21 | 5 | 14 | 33 | 0 | 150 | **4507** | Number of 'show details' views |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | **177** | Number of 'summarise search' views |

| Jan-2013 | Feb-2013 | Mar-2013 | Apr-2013 | May-2013 | Jun-2013 | Jul-2013 | Aug-2013 | Sep-2013 | Oct-2013 | Nov-2013 | Dec-2013 | SUM-2013 | Jan-2014 | Feb-2014 | Mar-2014 | Apr-2014 | May-2014 | Jun-2014 | Jul-2014 | Aug-2014 | Sep-2014 | Oct-2014 | Nov-2014 | SUM-2014 | | Number of result page views, by affiliation |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1252 | 1887 | 2248 | 2296 | 2183 | 1217 | 1233 | 1221 | 1147 | 1850 | 1918 | 1408 | 19860 | 2143 | 2359 | 2409 | 3021 | 3076 | 3435 | 5272 | 3204 | 4189 | 2133 | 307 | 31548 | **177753** | Grand total (SUM) |
| 543 | 808 | 422 | 783 | 987 | 701 | 597 | 472 | 583 | 1023 | 947 | 658 | 8524 | 1138 | 1789 | 1510 | 2004 | 1874 | 1697 | 1834 | 1005 | 2125 | 1148 | 52 | 16176 | **97366** | European countries(*) (EUROPE) |
| 168 | 38 | 82 | 346 | 139 | 144 | 100 | 154 | 105 | 158 | 236 | 192 | 1862 | 295 | 227 | 139 | 222 | 366 | 669 | 202 | 121 | 130 | 182 | 37 | 2590 | **27020** | United States (US) |
| 120 | 4 | 105 | 460 | 641 | 167 | 222 | 224 | 131 | 88 | 176 | 241 | 2579 | 239 | 26 | 172 | 272 | 220 | 324 | 1005 | 371 | 1660 | 371 | 100 | 4760 | **22306** | China (CN) |
| 89 | 90 | 139 | 65 | 482 | 43 | 90 | 23 | 33 | 289 | 293 | 111 | 1747 | 401 | 141 | 377 | 401 | 681 | 241 | 393 | 65 | 79 | 562 | 2 | 3343 | **21789** | Great Britain* (GB) |
| 62 | 161 | 37 | 35 | 51 | 218 | 272 | 88 | 59 | 201 | 156 | 123 | 1463 | 350 | 191 | 434 | 233 | 271 | 161 | 194 | 168 | 258 | 74 | 0 | 2334 | **19900** | Germany* (DE) |
| 12 | 31 | 45 | 29 | 63 | 39 | 126 | 167 | 28 | 24 | 54 | 101 | 719 | 37 | 17 | 144 | 244 | 119 | 58 | 105 | 97 | 1018 | 189 | 0 | 2028 | **15095** | Norway* (NC |
| 160 | 244 | 31 | 34 | 85 | 102 | 82 | 78 | 181 | 129 | 47 | 102 | 1275 | 322 | 107 | 165 | 127 | 196 | 58 | 1484 | 1379 | 147 | 80 | 93 | 4158 | **9121** | India (IN) |
| 29 | 2 | 510 | 67 | 40 | 53 | 195 | 77 | 24 | 242 | 42 | 55 | 1336 | 66 | 65 | 102 | 282 | 201 | 398 | 291 | 252 | 50 | 227 | 0 | 1934 | **7814** | Iran (IR) |
| 12 | 3 | 3 | 8 | 14 | 22 | 12 | 10 | 28 | 100 | 79 | 13 | 304 | 54 | 0 | 2 | 1 | 137 | 67 | 25 | 8 | 112 | 24 | 0 | 430 | **6517** | Spain* (ES) |
| 117 | 169 | 77 | 152 | 192 | 77 | 21 | 7 | 16 | 40 | 30 | 160 | 1058 | 36 | 39 | 5 | 0 | 76 | 70 | 19 | 70 | 16 | 0 | 0 | 331 | **4913** | Italy* (IT) |
| 7 | 109 | 24 | 333 | 39 | 58 | 52 | 4 | 0 | 93 | 56 | 58 | 833 | 54 | 905 | 81 | 23 | 0 | 10 | 42 | 12 | 0 | 0 | 0 | 1127 | **4355** | Poland* (PL) |
| 0 | 109 | 15 | 46 | 62 | 155 | 8 | 4 | 67 | 57 | 43 | 18 | 584 | 1 | 19 | 9 | 4 | 145 | 51 | 0 | 7 | 70 | 2 | 7 | 315 | **4002** | Greece* (GR) |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 242 | 78 | 5 | 0 | 326 | 49 | 152 | 322 | 744 | 275 | 132 | 913 | 444 | 323 | 65 | 9 | 3428 | **3857** | Bulgaria* (BC |
| 40 | 22 | 6 | 11 | 8 | 50 | 0 | 13 | 8 | 28 | 91 | 10 | 287 | 64 | 56 | 16 | 58 | 41 | 538 | 22 | 18 | 26 | 13 | 1 | 853 | **3580** | France* (FR) |
| 217 | 579 | 766 | 558 | 185 | 2 | 2 | 206 | 65 | 133 | 291 | 46 | 3050 | 15 | 27 | 26 | 12 | 11 | 23 | 1 | 25 | 2 | 16 | 1 | 159 | **3383** | Commercial (com) |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 2 | 3 | 0 | 0 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | **2807** | Netherlands* (NL) |
| 15 | 93 | 34 | 61 | 0 | 5 | 4 | 0 | 16 | 38 | 126 | 10 | 402 | 4 | 1 | 7 | 0 | 3 | 15 | 0 | 18 | 18 | 94 | 15 | 175 | **2783** | Slovak Republic* (SI |
| 0 | 16 | 4 | 5 | 0 | 20 | 1 | 1 | 14 | 38 | 114 | 57 | 270 | 21 | 34 | 53 | 10 | 102 | 27 | 43 | 0 | 1 | 15 | 0 | 306 | **2679** | Japan (JP) |
| 3 | 193 | 31 | 2 | 3 | 7 | 1 | 2 | 5 | 26 | 24 | 23 | 320 | 9 | 0 | 62 | 15 | 23 | 85 | 210 | 5 | 5 | 23 | 13 | 450 | **2581** | ? (unresolved) |
| 8 | 1 | 260 | 17 | 98 | 18 | 25 | 3 | 16 | 0 | 9 | 3 | 458 | 14 | 13 | 31 | 23 | 12 | 5 | 48 | 1 | 21 | 22 | 11 | 201 | **2532** | Network (net) |
| 3 | 0 | 20 | 21 | 3 | 2 | 0 | 2 | 23 | 13 | 26 | 30 | 143 | 21 | 70 | 27 | 45 | 50 | 142 | 51 | 41 | 16 | 12 | 0 | 475 | **1885** | Russian Federation (RU) |
| 168 | 18 | 8 | 24 | 29 | 18 | 2 | 120 | 0 | 13 | 2 | 0 | 402 | 8 | 10 | 3 | 0 | 0 | 4 | 17 | 0 | 3 | 19 | 16 | 80 | **1512** | Czech Republic* (CZ) |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 83 | 37 | 3 | 18 | 142 | 61 | 197 | 30 | 98 | 47 | 206 | 94 | 83 | 18 | 10 | 2 | 846 | **1453** | Austria* (AT) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 1 | 0 | 0 | 4 | 0 | 0 | 7 | 3 | 0 | 2 | 0 | 0 | 47 | 0 | 0 | 59 | **1308** | Ireland* (IE) |
| 0 | 2 | 0 | 16 | 42 | 4 | 0 | 5 | 0 | 21 | 0 | 17 | 107 | 7 | 0 | 1 | 96 | 67 | 63 | 0 | 0 | 0 | 0 | 0 | 234 | **1165** | Belgium* (BE) |
| 20 | 1 | 34 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 16 | 72 | 12 | 17 | 62 | 7 | 0 | 12 | 4 | 8 | 15 | 84 | 0 | 221 | **1102** | Switzerland* (CH) |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 0 | 8 | 0 | 6 | 0 | 73 | 12 | 6 | 0 | 0 | 10 | 6 | 0 | 113 | **393** | Sweden* (SE) |
| 0 | 1 | 1 | 0 | 2 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 6 | 0 | 0 | 117 | 2 | 21 | 0 | 79 | 4 | 17 | 21 | 0 | 261 | **374** | Brazil (BR) |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 19 | 3 | 0 | 0 | 0 | 23 | 0 | 37 | 1 | 8 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 47 | **252** | Ukraine* (UA) |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 5 | 0 | 0 | 2 | 0 | 8 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | **185** | Romania* (RO) |
| 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **166** | Argentina (AR) |
| 0 | 0 | 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **150** | Saudi Arabia (SA) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 112 | 0 | 0 | 119 | **133** | Serbia* (RS) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 58 | 0 | 0 | 0 | 0 | 0 | 58 | **108** | Denmark (DK) |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 3 | 0 | 3 | 2 | 0 | 0 | 0 | 0 | 12 | 11 | 0 | 31 | **81** | Australia (AU) |
| 0 | 0 | 0 | 0 | 5 | 6 | 10 | 4 | 0 | 0 | 0 | 1 | 26 | 0 | 0 | 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 9 | **50** | Finland* (FI) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 0 | 0 | 0 | 6 | **44** | Mexico (MX) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **37** | Taiwan (TW) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 7 | 18 | 0 | 0 | 0 | 0 | 27 | **32** | South Africa (ZA) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **31** | United Arab Emirates (AE) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 6 | 0 | 0 | 2 | 0 | 9 | **31** | European Union* (EU) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 3 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 8 | **21** | Canada (CA) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **21** | Hungary* (HU) |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **19** | Hong Kong (HK) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **15** | Portugal* (PT) |
| 0 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **12** | Egypt (EG) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **11** | Kazakhstan (KZ) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **11** | Peru (PE) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **10** | Non-Profit Making Organisations (ORG) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 4 | 0 | 6 | **9** | Latvia* (LV) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **9** | Israel (IL) |

| | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 8 | Georgia* (GE) |
| 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 | United Kingdom (UK) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 | Algeria (DZ) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | Colombia (CO) |
| 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 1 | 1 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | Indonesia (ID) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | Luxembourg* (LU) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | (BIZ) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 5 | 5 | (AP) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | Lithuania* (LT) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | Polynesia (French) (PF) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Estonia (EE) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Slovenia* (SI) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Kenya (KE) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Former Czechoslovakia (CS) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | Croatia (HR) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | Uzbekistan (UZ) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | Thailand (TH) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | Costa Rica (CR) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | Iraq (IQ) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | Venezuela (VE) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | New Zealand (NZ) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | Chile (CL) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | Philippines (PH) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | Pakistan (PK) |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | Bangladesh (BD) |

## SubWorkpackage 102: BE-WISSDOM. Access to Science Data Centres

Leader Organization: KONINKLIJKE STERRENWACHT VAN BELGIE (ROB)

Participants: ROB

The Solar Dynamics Observatory (SDO) was launched in January 2010 carrying three primary instruments the Atmospheric Imaging Assembly (AIA), the Helioseismic and Magnetic Imager (HMI) and the Extreme ultraviolet Variability Experiment (EVE). These instruments produce between 1 and 2 TB of data per day.

The goal of the European SDO data centre is to be a powerful cashe for the data produced by SDO. By doing this we diminish the load on the USA SDO data centre at the Joint Science Operations Center (JSOC) at Stanford university and provide the Solar community with a quick and reliable access to SDO data.

Currently we are in the process of purchasing a new server setup to replace our current one and with this we add to the services we are currently providing.

In this report we'll explain:

- the current setup for the SDO datacentre and the services it provides.
- the planned setup and services the new server setup will provide
- and provide web data on the access to the SDO datacentre
- the outreach we've done.
- deviations from the statement of work in the WP 100 and WP 102 and corrective actions.

# The SDO datacentre

Here we'll describe the hardware and the services the current servers of the SDO data centre provide.

The servers of the SDO data centre are made out of database servers, a storage array and a webserver.

The webserver is a DELL PowerEdge R308 server running Apache software.

The storage array is made out of 1 and 2 TB SAS and SATA hard disks in total providing 293 TB of storage. These are controlled by redundant pairs of MD3000i and MD1000i controllers.

The database servers are two redundant DELL PowerEdge 1950 with a 500GB SAS hard disk and a 100 GB SSD each.

Foto 1: current SDO data centre

Currently we provide the following data:

- Near real time (15 min) 1k by 1k AIA quickook data. A movie of the last 24 hours can be found on our website. http://sdodata.oma.be/latest/
- synoptic 4k by 4k AIA/HMI (HMI continium and magnetogram) Level 1 fits files at a cadence of 1h for all wavelengths.
- 9 Months of 4k by 4k AIA/HMI data with a cadence higher than one hour.

The data is available from ROB through the following means:

- Directly from ROB through the ftp and http protocol
- In IDL trough the Virtual solar observatory.
- In Python trough the Virtual solar observatory.
- Through a web aplication named the SDO wizard. http://sdodata.oma.be/sdo_wizard/

Here we describe deviation from the statement of work in the DOW Annex I and actions taken to correct these deviations.

Access to data for complete archive transfer is offered to collaborating institutes in the SDO data distribution project in particular the University of central Lancashire (uclan) and Institute d'Astrophysique Spatiale (IAS). While the access is still offered IAS and uclan are now downloading their data directly from from the Joint Science Operations Center (JSOC), headquartered at

Stanford University. This is because JSOC has significantly upgraded their infrastructure and can now handle direct download of multiple data centres.

The SDO data centre stopped providing the 4k by 4k data series in the period from September 2013 until December 2013. This was due to an attempt of a large scale software update.

## On going work and future changes

We decided to replace the current server setup with a new one. The reason for this is in part that Dell has stopped making the hard disk controllers that are used in the old system and stopped providing updates for the firmware. Also the system is starting to show signs of wear. So we decided to let the system go out of warranty. A second reason is that the new system will be faster and more powerful allowing us to improve on our services to meet the needs of the Solar community.

We have already purchased a new storage system that is currently being set up and tested and we are planning on buying a new database server. The former was bought by ROB on Solar-Terrestrial Centre of Excellence (STCE) budget and the later will be bought with the funds allocated to us from the SOLARNET project.

The storage array is smaller than the one in the current system because we are changing our services to be more focused on on-demand download and faster storage. The direction we are taking with the new system will also make it more suited for testing and hosting the Solar Virtual Observatory (SVO) prototype in development at ROB as part of WP50.

The servers of the future SDO data centre are made up of database servers, a storage array and a webserver. The data storage is shown in the photo on the right.

The storage array is made up of a SuperMicro storage server and a JBOD (just a bunch of disks) storage extension.

These are equipped with 45 4TB harddisks and two 80GB solid state drives, providing about 120 TB of usable storage after formatting.

The storage array is configured to allow any two harddisks to fail simultaneously without any data loss. It can easily be extended with additional hard disks, or JBOD extensions.

The database server is to be ordered before the end of 2014.

Foto 2: new storrage array

When the new SDO data centre is fully operational we will provide the following services.

- Near real time (15 min) 1k by 1k AIA quicklook data. A movie of the last 24 hours can be found on our website http://sdodata.oma.be/latest/
- Synoptic 4k by 4k AIA/HMI (HMI continuum and magnetogram) Level 1 fits files at a cadence of 1h for all wavelengths.
- 1k by 1k AIA data at a 2 minute cadence.
- We will add Spaceweather HMI Active Region Patch (SHARP) data.
- Instead of providing 9 months of high cadence data we'll move to a buffer system were data is kept by user request. With an alert service for when your data is available.

Access will remain the same. The access tools are being updated where needed.

## Outreach

Here we mention some outreach highlights for the data centre and a summary of the results of our recent survey.

In the past years the European SDO data centre has been advertised at conferences and meetings by the ROB team members attending.

Earlier this year we sent around a survey to the SOLARNET partners and to the solar science community in general through the "SolarNews" newsletter. Asking for improvements that the solar science community would like to see from the data centre.

An e-poster has been accepted at the European space weather week which will show a demo of the new SVO tool in development which is part of WP50.

Here we'll give a summery of the results of our survey on the SDO data centre. http://sdoatsidc.oma.be/web/sdoatsidc/Survey We had a total of 25 participants who filled in the survey.

On the question of what kind of synoptic cadence people preferred most said one hour. Most of the participants also preferred a higher cadence for specific events and gave some examples.

On the question if people would like have access to the AIA level 1.5 synoptic dataset at ROB? The answers of yes and no were divided about fifty fifty.

## Interpretation of the web statistics

In order to collect and analyse the web data we use the server log file analysis method. This works by reading the logfiles in which the web server records file requests by browsers. We use two web log analysis software namely awstats and webalyzer. Only webalyzer was installed in April and May of 2013 so we only have data from that software for those months.

There are two caveats that need to be kept in mind when dealing with web data from log file analysis. Log file analysis software doesn't always distinguish between a browser operated by a human and a web crawler collecting data for a search engine this can lead to an overestimation of the number of visitors. Secondly due to web cache a person revisiting a page may not be counted towards the number of visits causing an underestimation of the number of visitors.

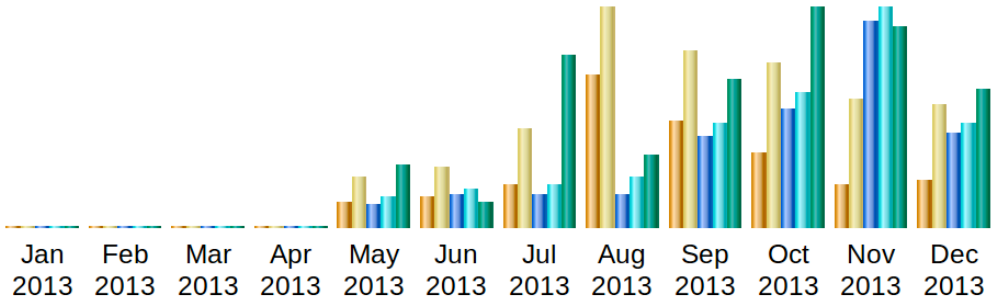
Both these problems can be solved by excluding known web crawlers and stopping browser caching but this hampers visibility in search engines and result in degraded performance for the visitor and bigger load on the servers. An option for the future to deal with this is to start using the page tagging method.

## Web statistics

Here we present the statistics from Awstats and webalyser from Arpil 1, 2013 to September 30, 2014

Below we present the table and graphs from the Awstats tool. It was installed mid May and the data was taken in October. The definitions from the manual are below.

## Monthly history




| Month | Unique visitors | Number of visits | Pages | Hits | Bandwidth |
|---|---|---|---|---|---|
| Jan 2013 | 0 | 0 | 0 | 0 | 0 |
| Feb 2013 | 0 | 0 | 0 | 0 | 0 |
| Mar 2013 | 0 | 0 | 0 | 0 | 0 |
| Apr 2013 | 0 | 0 | 0 | 0 | 0 |
| May 2013 | 339 | 667 | 62529 | 84454 | 11.84 GB |
| Jun 2013 | 395 | 803 | 85329 | 104327 | 5.10 GB |
| Jul 2013 | 562 | 1328 | 88984 | 111752 | 33.28 GB |
| Aug 2013 | 2029 | 2944 | 88278 | 131899 | 14.07 GB |
| Sep 2013 | 1418 | 2358 | 244524 | 275317 | 28.60 GB |
| Oct 2013 | 986 | 2201 | 318872 | 361628 | 42.23 GB |
| Nov 2013 | 559 | 1702 | 554289 | 588905 | 38.65 GB |
| Dec 2013 | 626 | 1650 | 252277 | 280003 | 26.28 GB |
| Total | 6914 | 13653 | 1695082 | 1938285 | 200.05 GB |

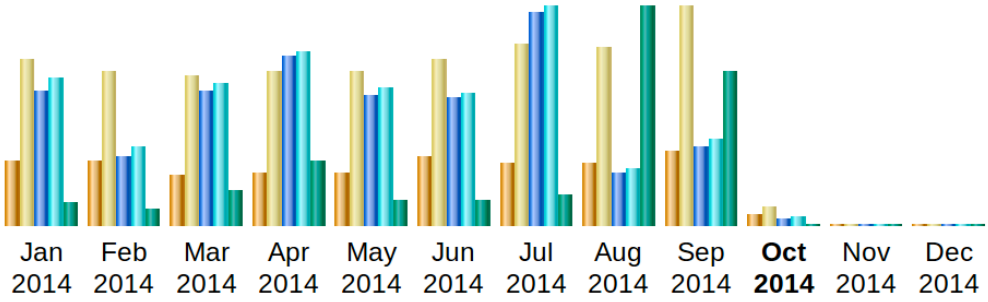
Graph and table 1: Awstats data 2013

**Unique Visitor**:
A unique visitor is a person or computer (host) that has made at least 1 hit on 1 page of our web site during the period of one month. If this user makes several visits during this period, it is counted only once. Visitors are tracked by IP address, so if multiple users are accessing our site from the same IP (such as a office network), they will be counted as a single unique visitor.

**Visits:**
Number of visits made by all visitors. Think "session" here, say a unique IP accesses a page, and then requests three other pages within an hour. All of the "pages" are included in the visit, therefore you should expect multiple pages per visit and multiple visits per unique visitor (assuming that some of the unique IPs are logged with more than an hour between requests)

**Pages:**
The number of "pages" viewed by visitors. Pages are usually HTML, PHP or ASP files, not images or other files requested as a result of loading a "Page" (like js,css... files).

## Monthly history




| Month | Unique visitors | Number of visits | Pages | Hits | Bandwidth |
|---|---|---|---|---|---|
| Jan 2014 | 587 | 1529 | 372115 | 406403 | 30.28 GB |
| Feb 2014 | 598 | 1421 | 193835 | 219512 | 23.42 GB |
| Mar 2014 | 460 | 1380 | 369634 | 388757 | 47.61 GB |
| Apr 2014 | 480 | 1413 | 461907 | 479997 | 87.94 GB |
| May 2014 | 476 | 1422 | 358644 | 376833 | 34.10 GB |
| Jun 2014 | 634 | 1541 | 348251 | 367006 | 33.78 GB |
| Jul 2014 | 569 | 1677 | 582918 | 602377 | 42.38 GB |
| Aug 2014 | 579 | 1648 | 143812 | 160164 | 301.79 GB |
| Sep 2014 | 690 | 2017 | 217248 | 239129 | 213.19 GB |
| **Oct 2014** | 96 | 171 | 20018 | 21874 | 747.01 MB |
| Nov 2014 | 0 | 0 | 0 | 0 | 0 |
| Dec 2014 | 0 | 0 | 0 | 0 | 0 |
| Total | 5169 | 14219 | 3068382 | 3262052 | 815.21 GB |

Graph and table 2: Awstats data 2014

**Hits:**
Any files requested from the server (including files that are "Pages").

**Bandwidth:**
Total number of bytes for pages, images and files downloaded by web browsing.
Note 1: Of course, this number includes only traffic for web only (or mail only, or ftp).
Note 2: This number does not include technical header data size used inside the HTTP or HTTPS protocol or by protocols at a lower level (TCP, IP...).

So the bandwidth here gives an estimate of the files downloaded and played by our AIA quicklook data.

Below we give the data from webalizer.

| | April, 2013 | May, 2013 | June, 2013 | July, 2013 | August, 2013 | September, 2013 | October, 2013 | November, 2013 | December,2013 |
|---|---|---|---|---|---|---|---|---|---|
| Hits | 30928 | 30929 | 36578 | 382452 | 863716 | 1333782 | 2034810 | 1996520 | 1605356 |
| Files | 11049 | 20195 | 24645 | 211160 | 234041 | 177803 | 224235 | 471303 | 211464 |
| Pages | 10356 | 25319 | 33420 | 190248 | 184452 | 104331 | 125811 | 389916 | 158383 |
| Visits | 4781 | 9207 | 13327 | 5327 | 5114 | 4054 | 4146 | 5525 | 4606 |
| Total KB | 1781007 | 1784149 | 1479037 | 90073525 | 104428705 | 352731689 | 563257522 | 459872936 | 334373708 |
| | January, 2014 | February, 2014 | March, 2014 | April, 2014 | May, 2014 | June, 2014 | July, 2014 | August, 2014 | September, 2014 |
| Hits | 1355707 | 1001421 | 1091996 | 1322210 | 1157286 | 1270803 | 1534221 | 629264 | 697597 |
| Files | 353503 | 152843 | 172729 | 164404 | 178176 | 226234 | 268209 | 142797 | 188159 |
| Pages | 282083 | 97635 | 116094 | 102774 | 118582 | 125768 | 192095 | 124051 | 156441 |
| Visits | 5311 | 3878 | 4509 | 3490 | 3876 | 4102 | 4876 | 6385 | 5633 |
| Total KB | 322576070 | 240522351 | 171728866 | 250545859 | 184653792 | 305468383 | 246055385 | 424981885 | 307969727 |

Table 1: webalizer data

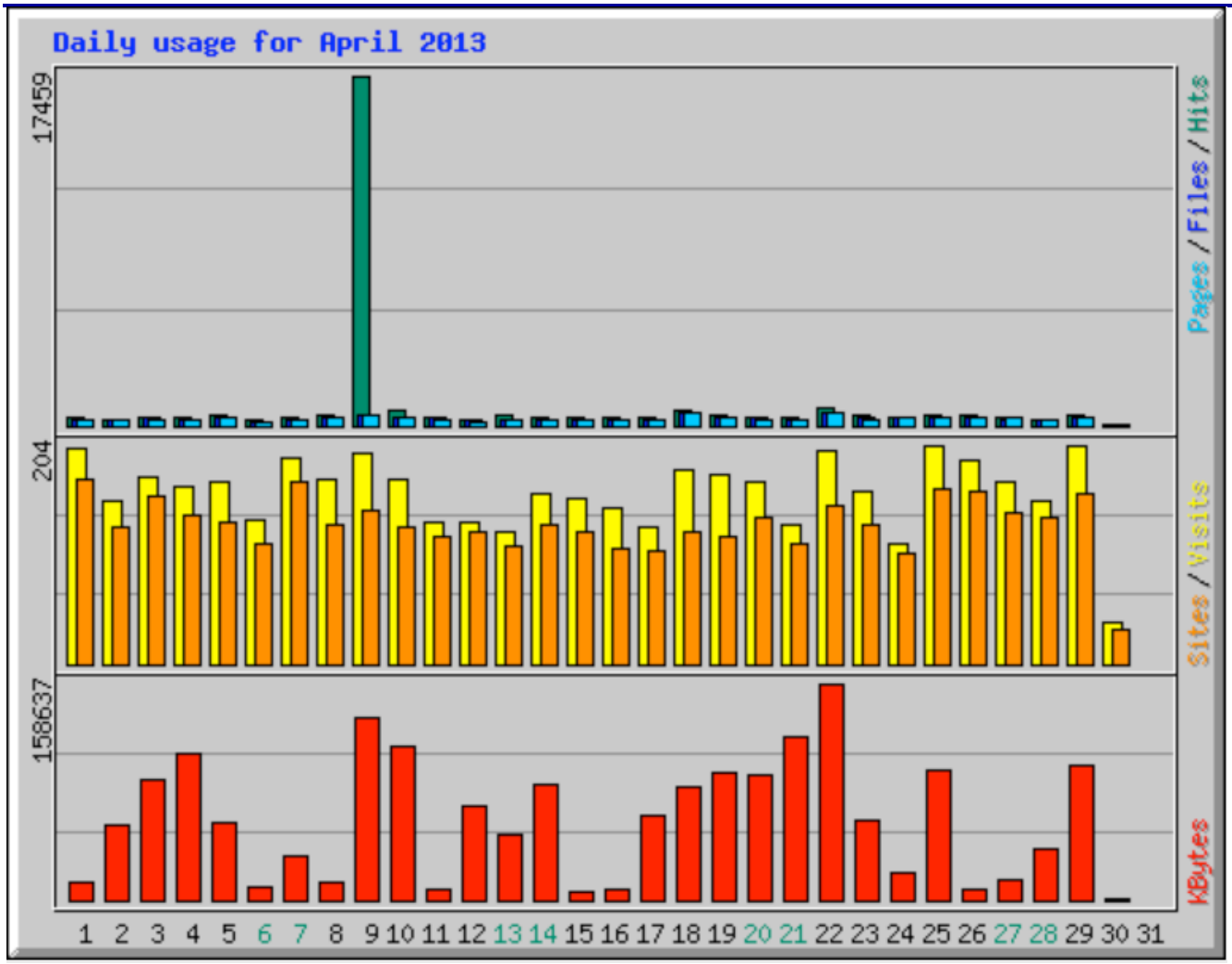
**Hits** represent the total number of requests made to the server during the given month.

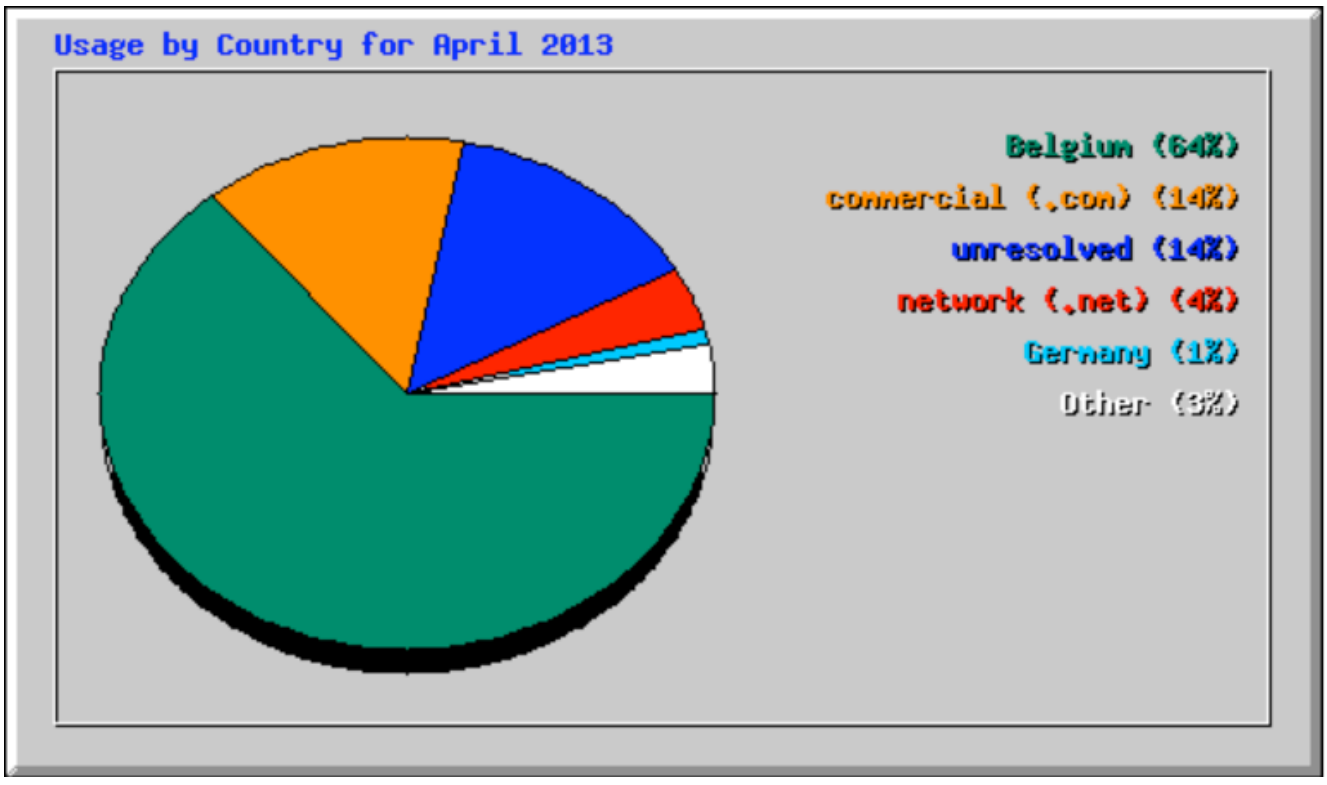
**Files** represent the total number of hits (requests) that actually resulted in something being sent back to the user. Not all hits will send data, such as 404-Not Found requests and requests for pages that are already in the browsers cache.

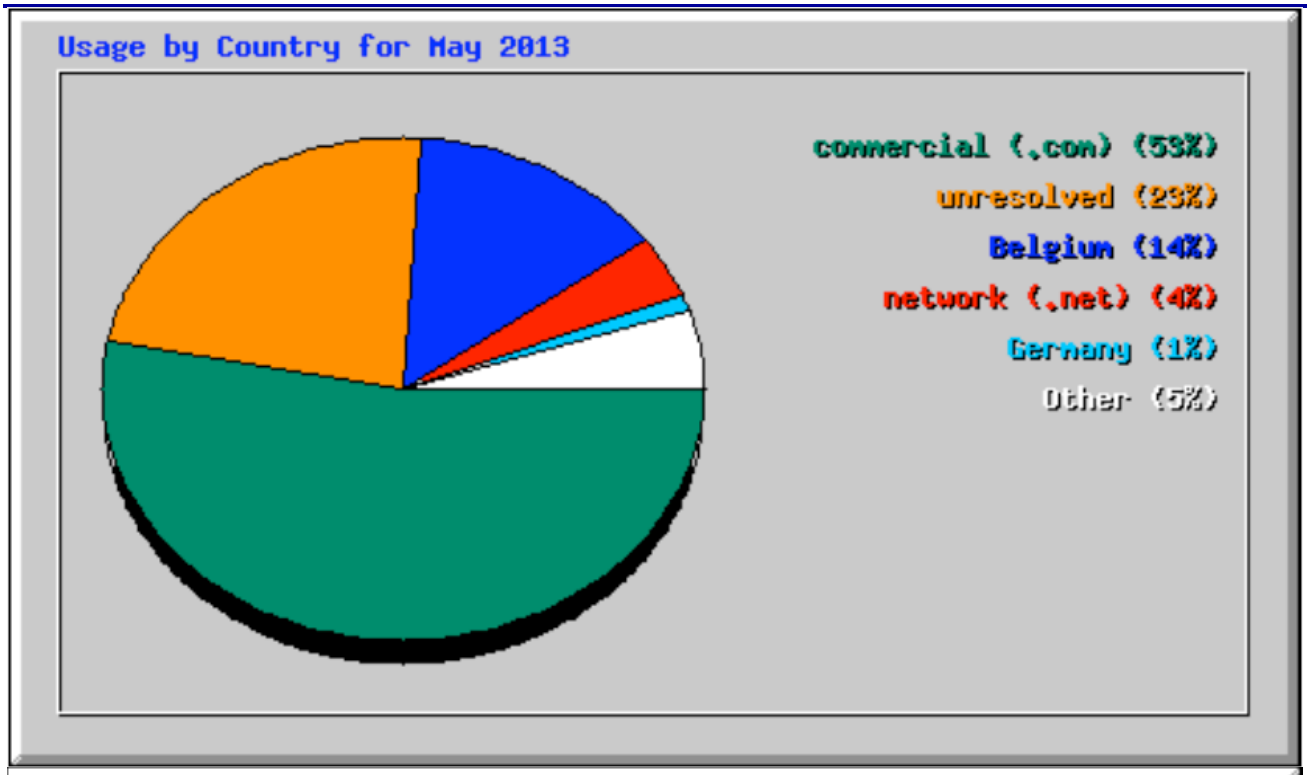
**Pages** are those URLs that would be considered the actual page being requested, and not all of the individual items that make it up (such as graphics and audio clips).

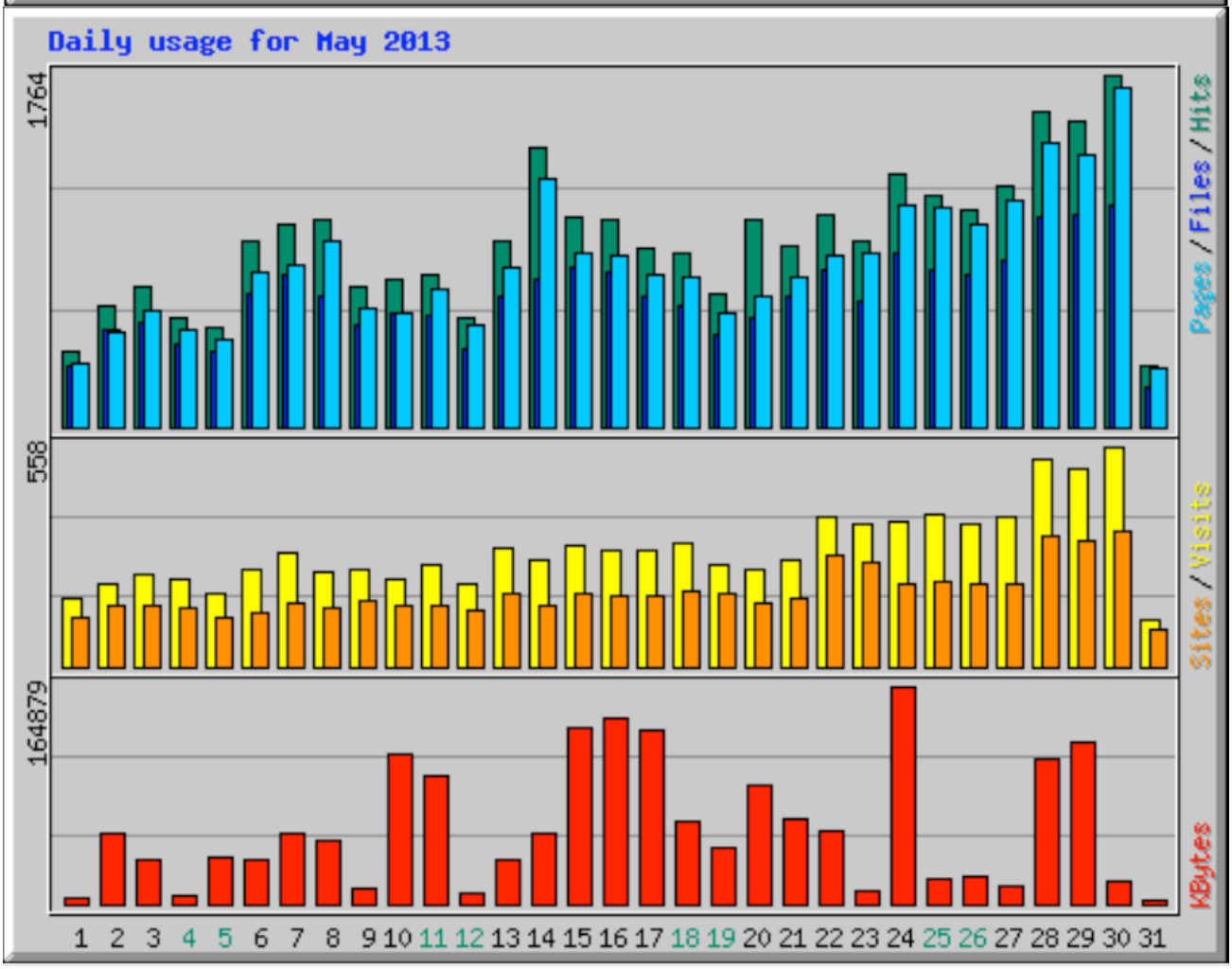
**Visits** occur when some remote site makes a request for a page on your server for the first time. As long as the same site keeps making requests within 30 min, they will all be considered part of the same visit. If the site makes a request to our server, and the length of time since the last request is greater than 30 min, a new visit is started and counted, and the sequence repeats. Since only pages will trigger a visit, remotes sites that link to graphic and other non- page URLs are not counted in the visit totals, reducing the number of false visits.

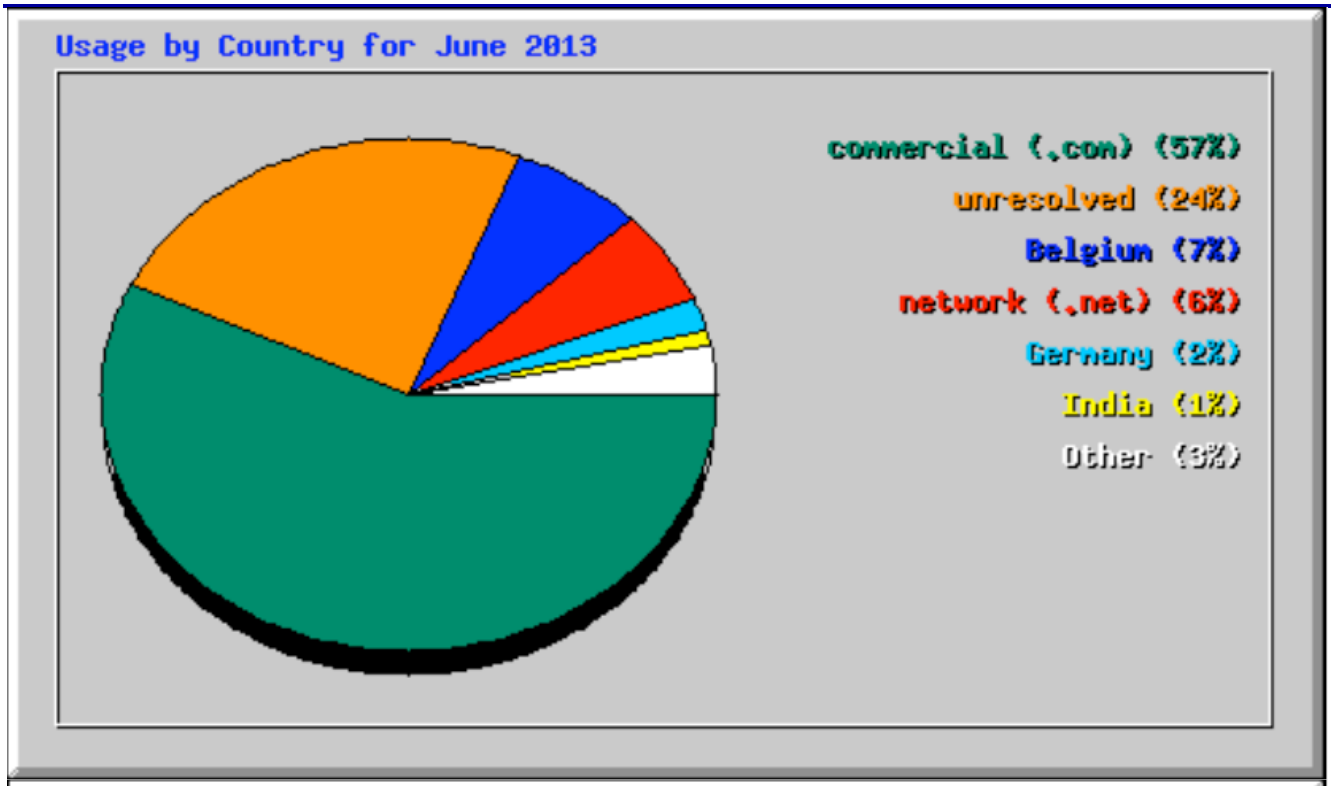
**Total KByte** (KB) is used to show the amount of data that was transfered between the server and the remote machine, based on the data found in the server log. These numbers are an estimate of the amount of fits files that are downloaded

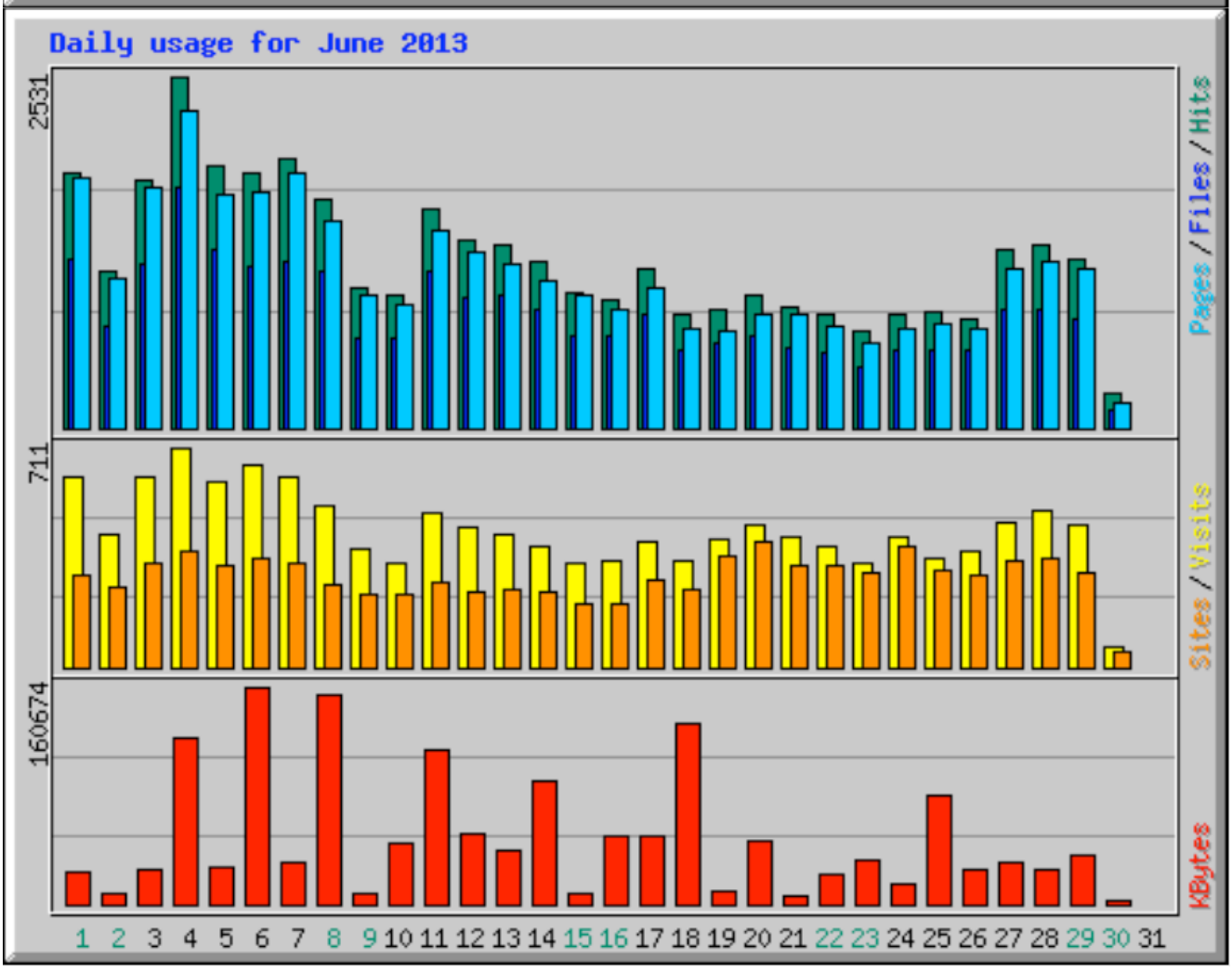

Daily usage for April 2013
17459
204
158637
Pages / Files / Hits
Sites / Visits
KBytes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

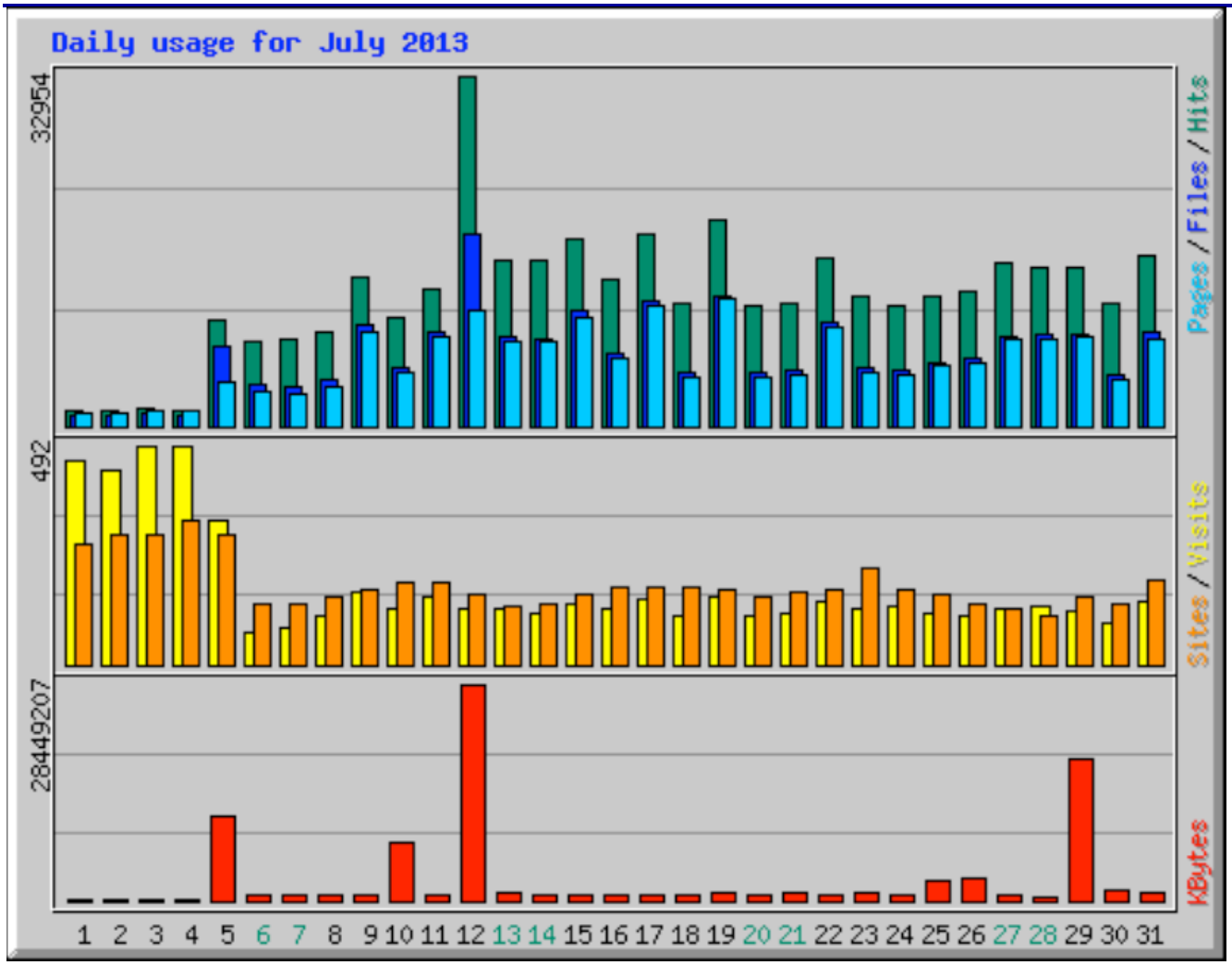

Usage by Country for April 2013
Belgium (64%)
commercial (.com) (14%)
unresolved (14%)
network (.net) (4%)
Germany (1%)
Other (3%)

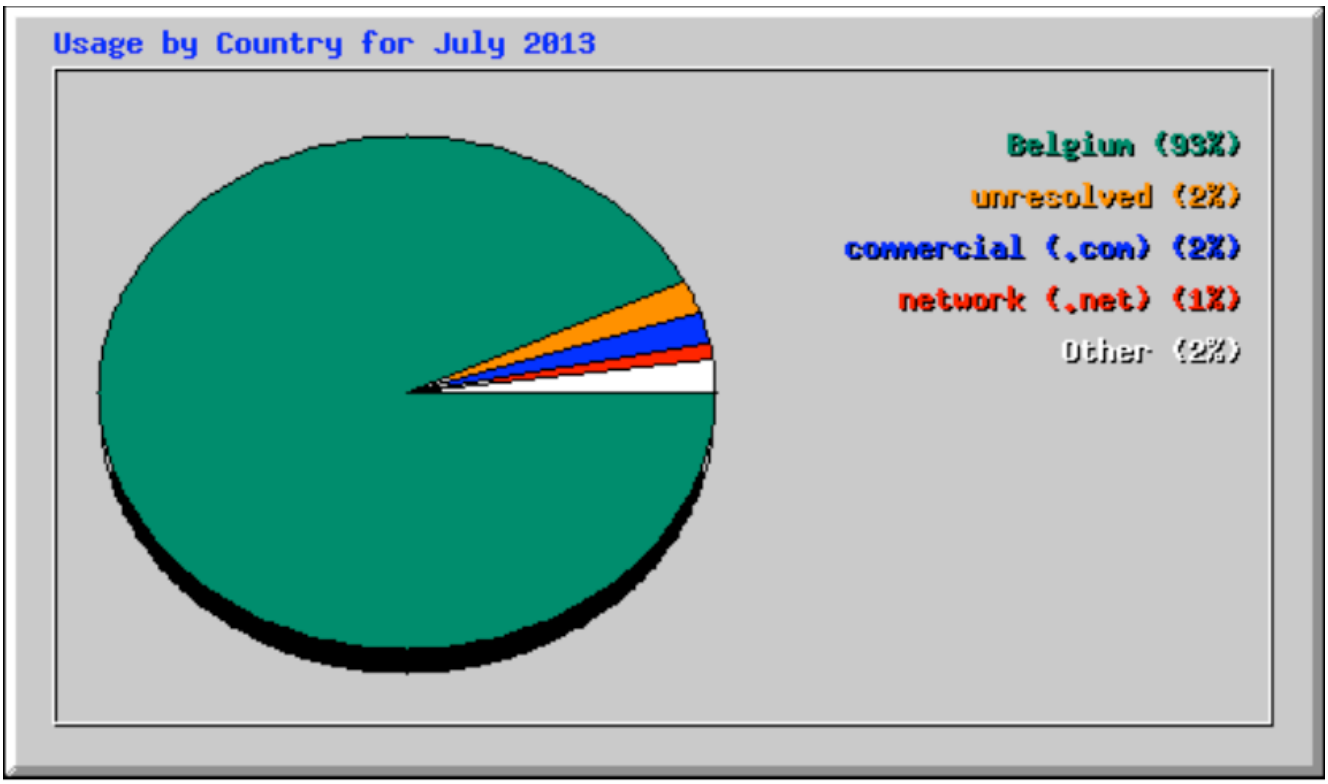

Usage by Country for May 2013
connercial (.con) (53%)
unresolved (23%)
Belgium (14%)
network (.net) (4%)
Germany (1%)
Other (5%)

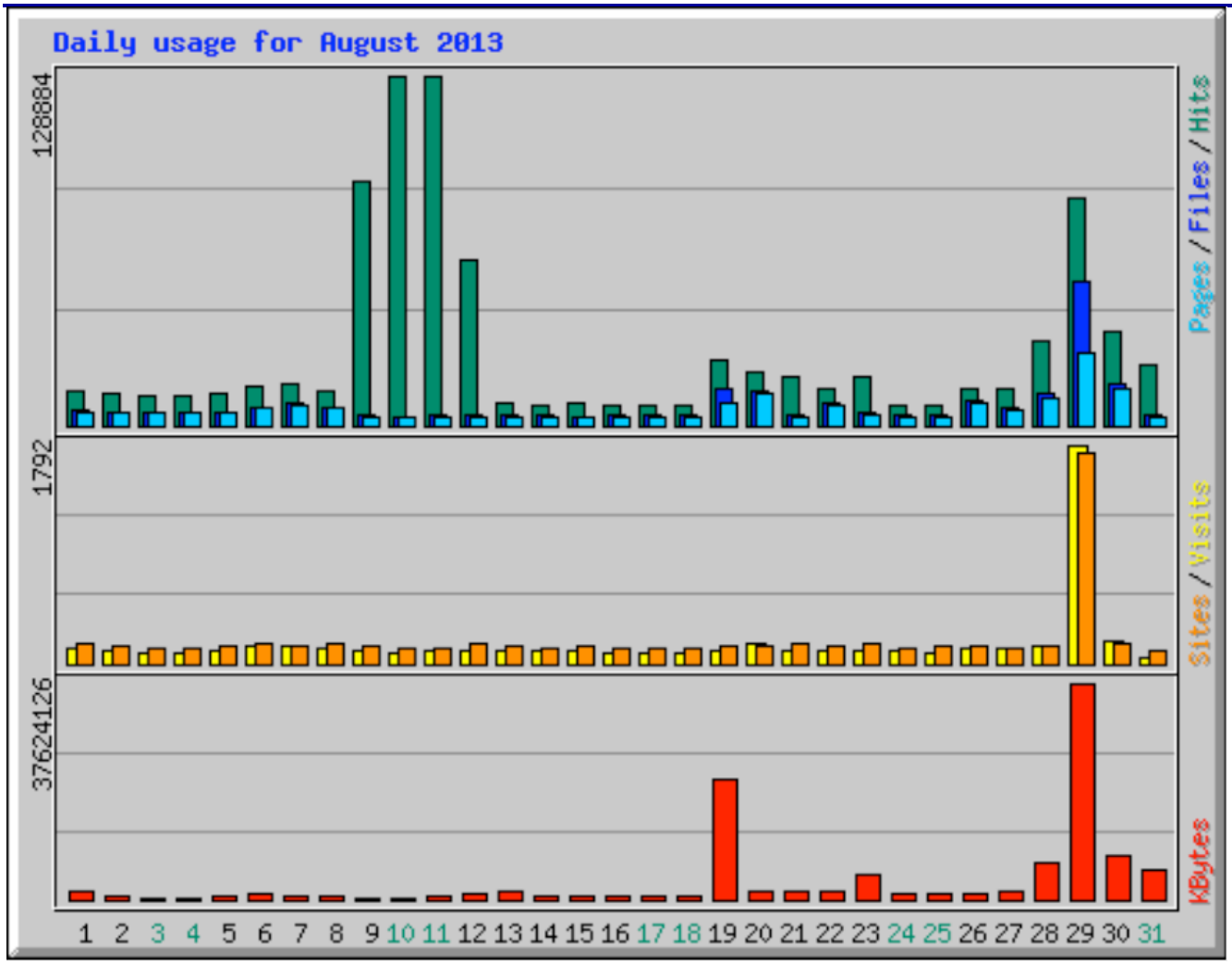

Daily usage for May 2013
1764
558
164879
Pages / Files / Hits
Sites / Visits
KBytes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

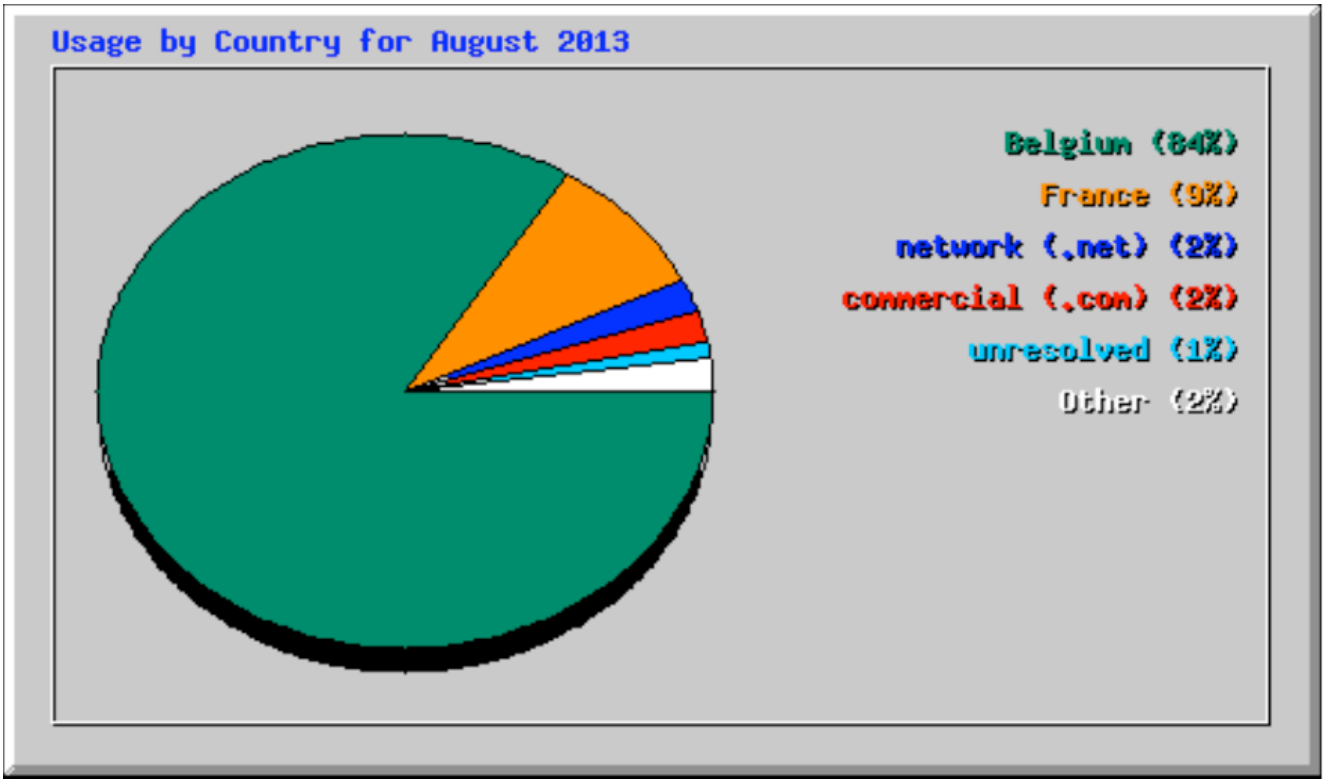

Usage by Country for June 2013
connercial (.com) (57%)
unresolved (24%)
Belgium (7%)
network (.net) (6%)
Germany (2%)
India (1%)
Other (3%)

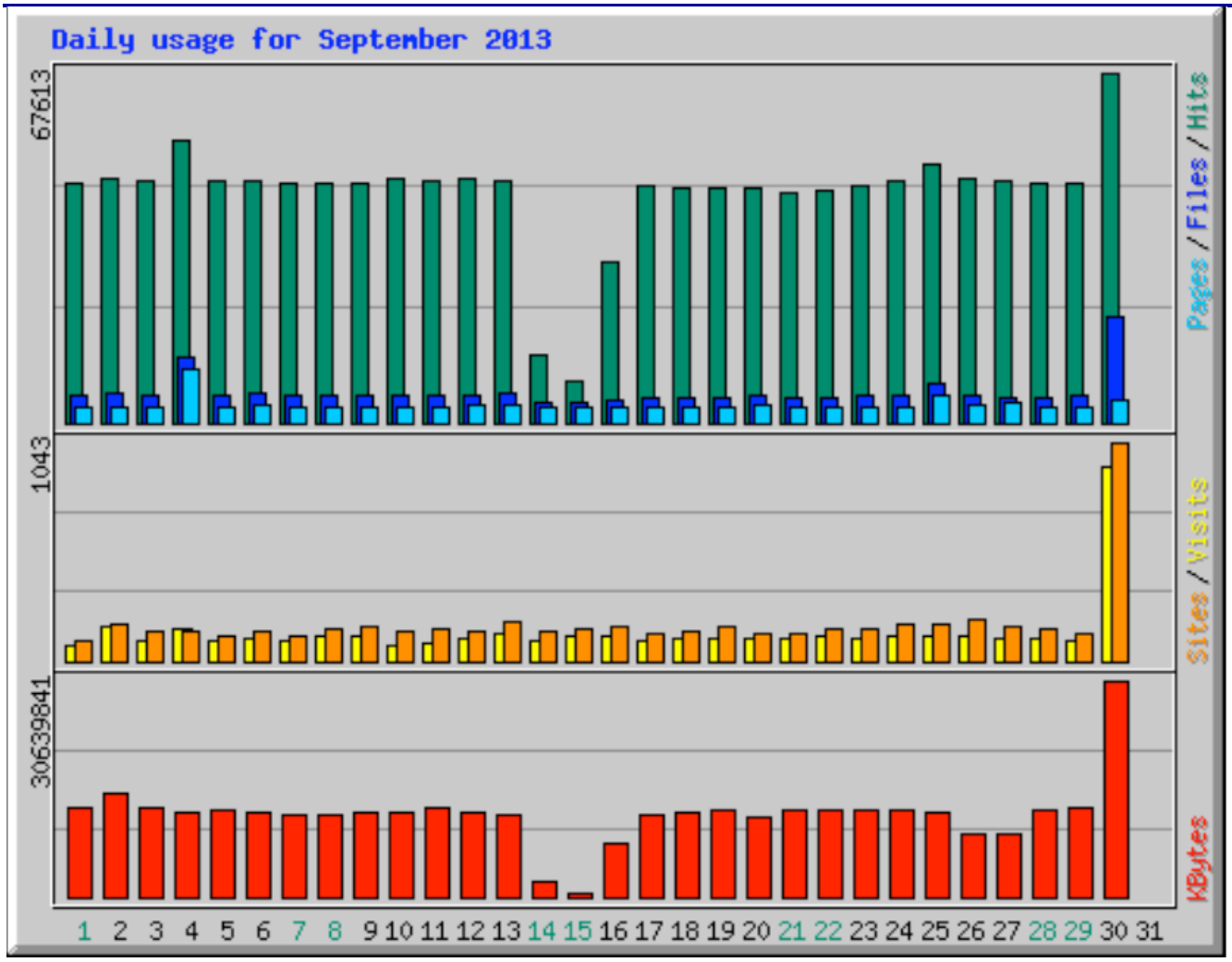

Daily usage for June 2013
2531
711
160674
Pages / Files / Hits
Sites / Visits
KBytes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

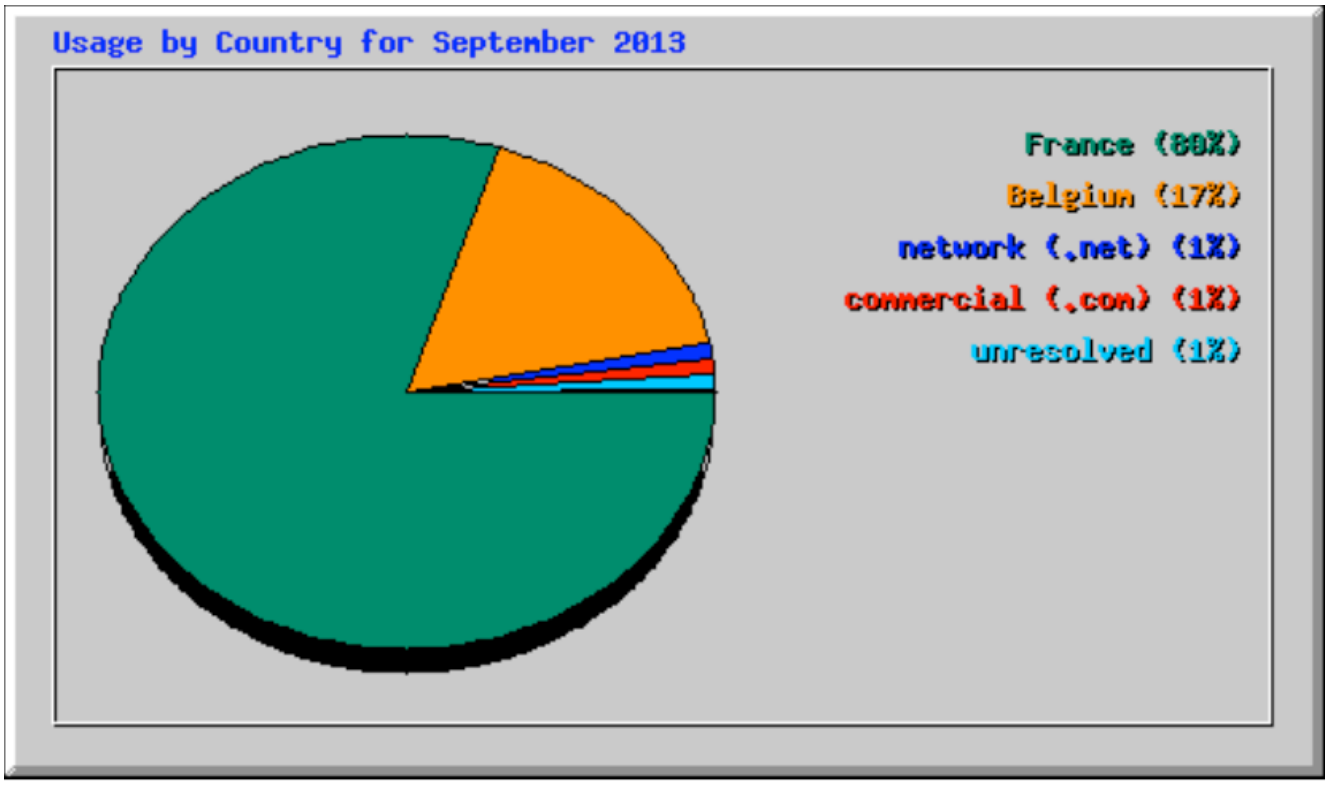

Daily usage for July 2013
32954
492
28449207
Pages / Files / Hits
Sites / Visits
KBytes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

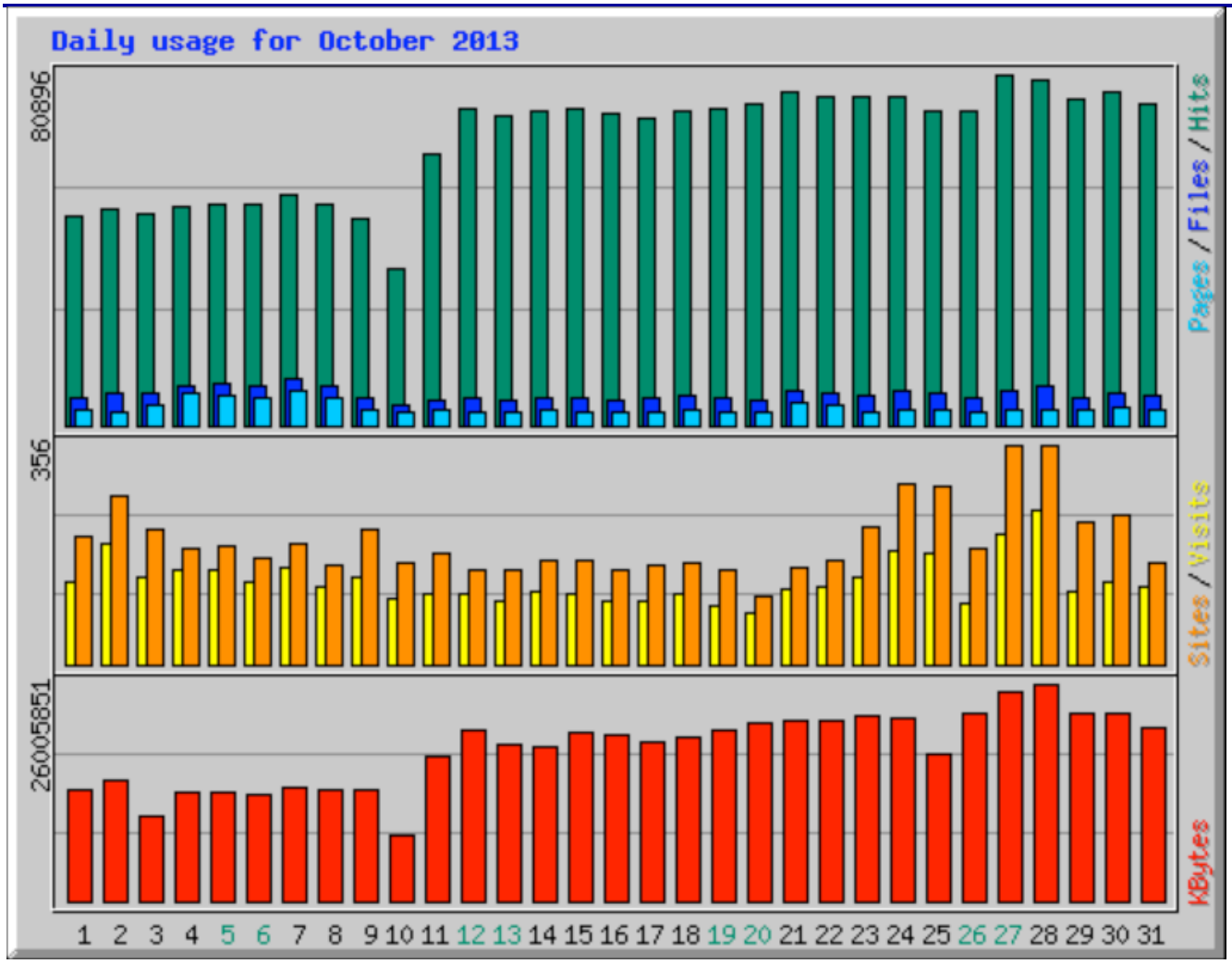

Usage by Country for July 2013
Belgium (93%)
unresolved (2%)
commercial (.com) (2%)
network (.net) (1%)
Other (2%)

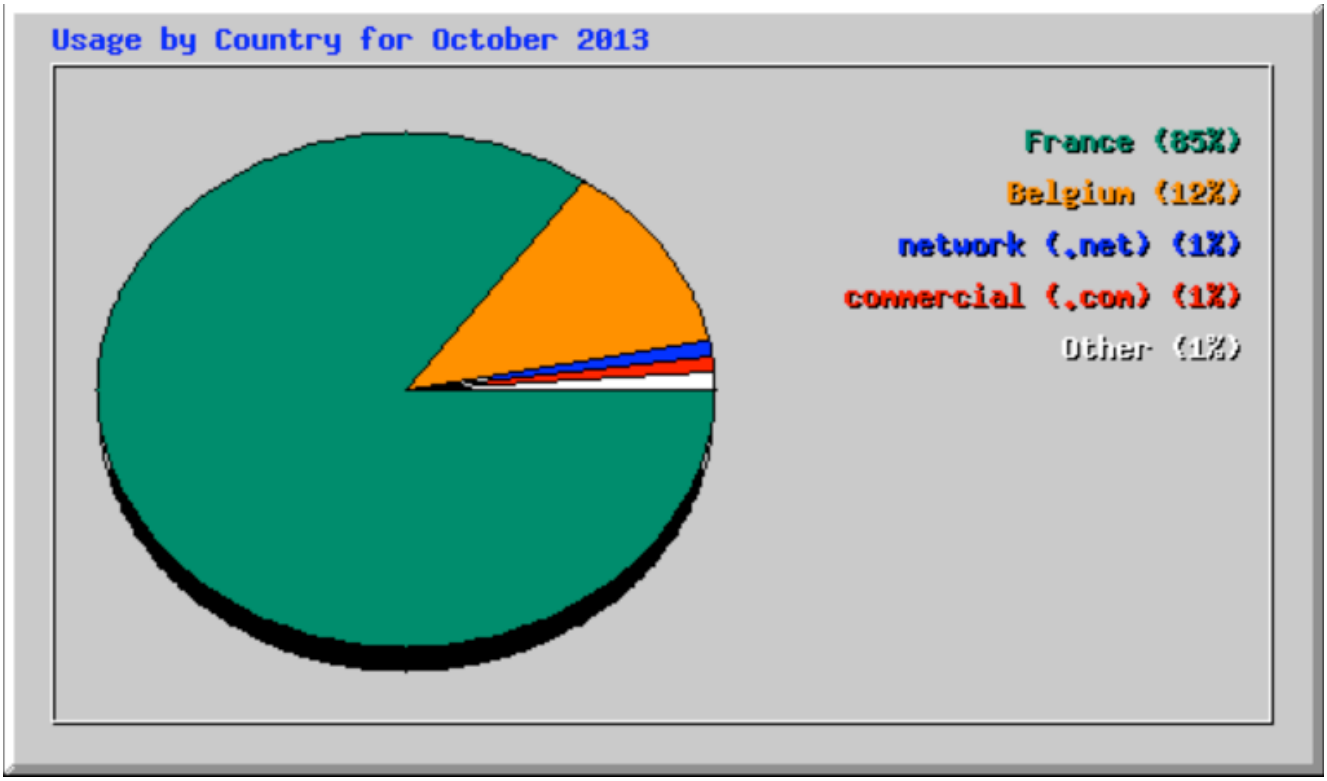

Daily usage for August 2013
128884
1792
37624126
Pages / Files / Hits
Sites / Visits
KBytes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

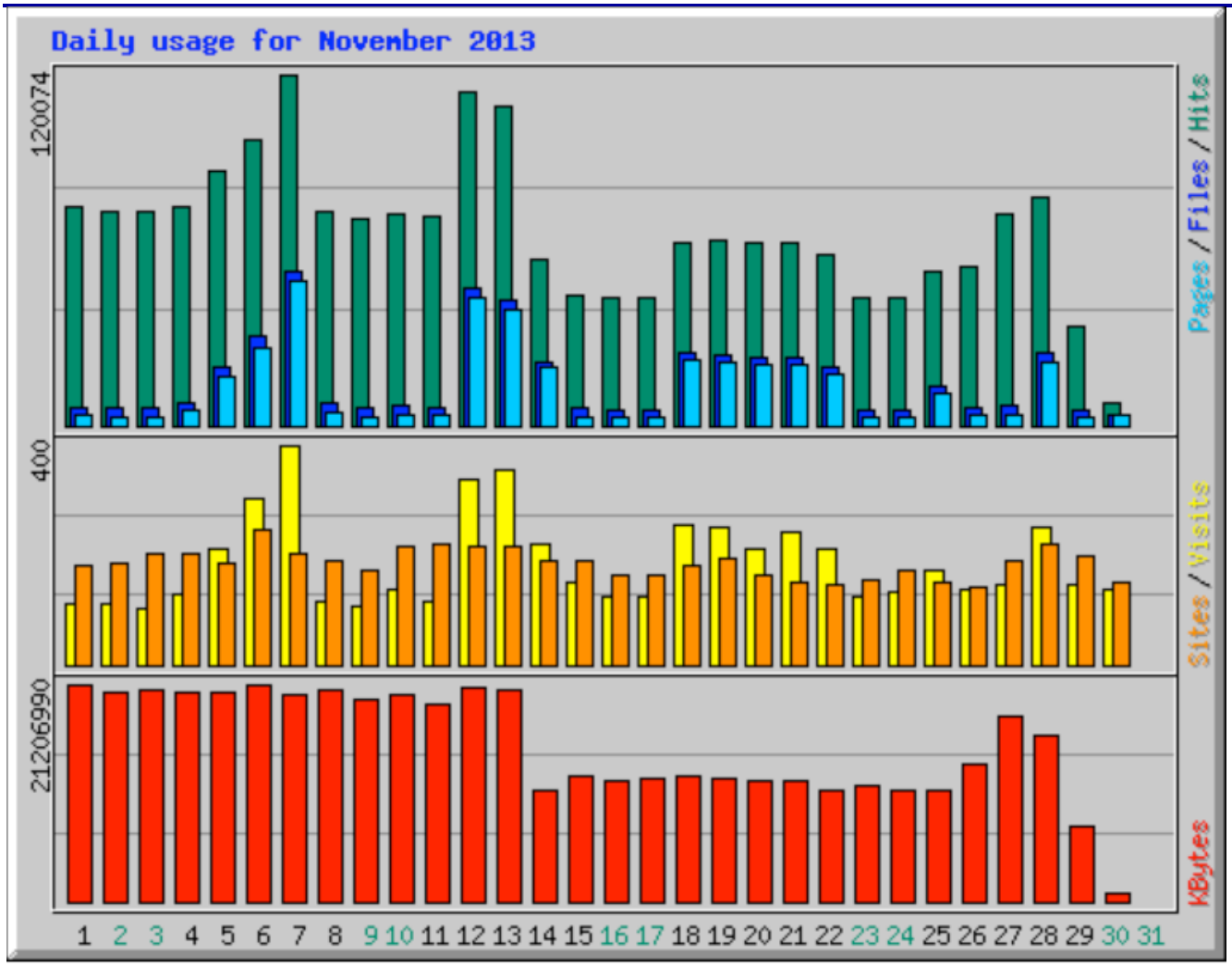

Usage by Country for August 2013
Belgium (84%)
France (9%)
network (.net) (2%)
connercial (.com) (2%)
unresolved (1%)
Other (2%)

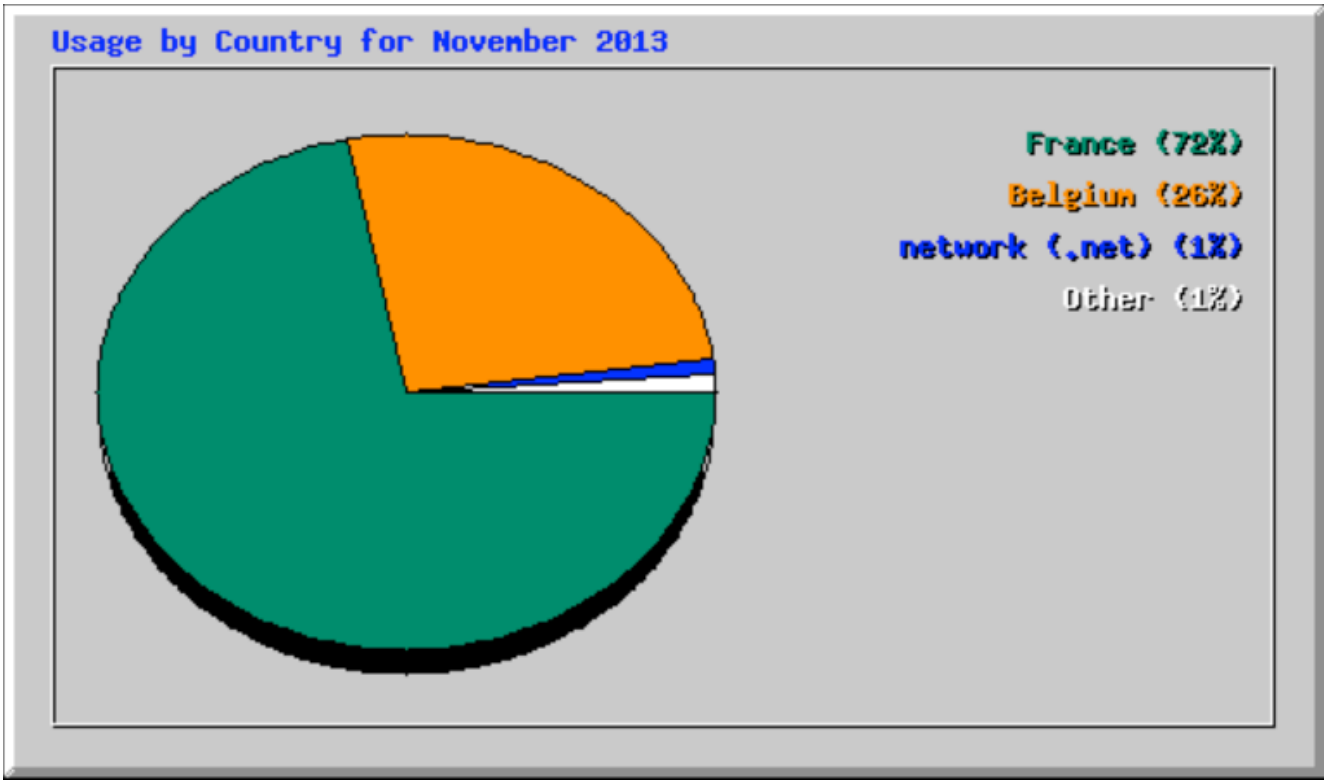

Daily usage for September 2013
67613
1043
30639841
Pages / Files / Hits
Sites / Visits
KBytes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

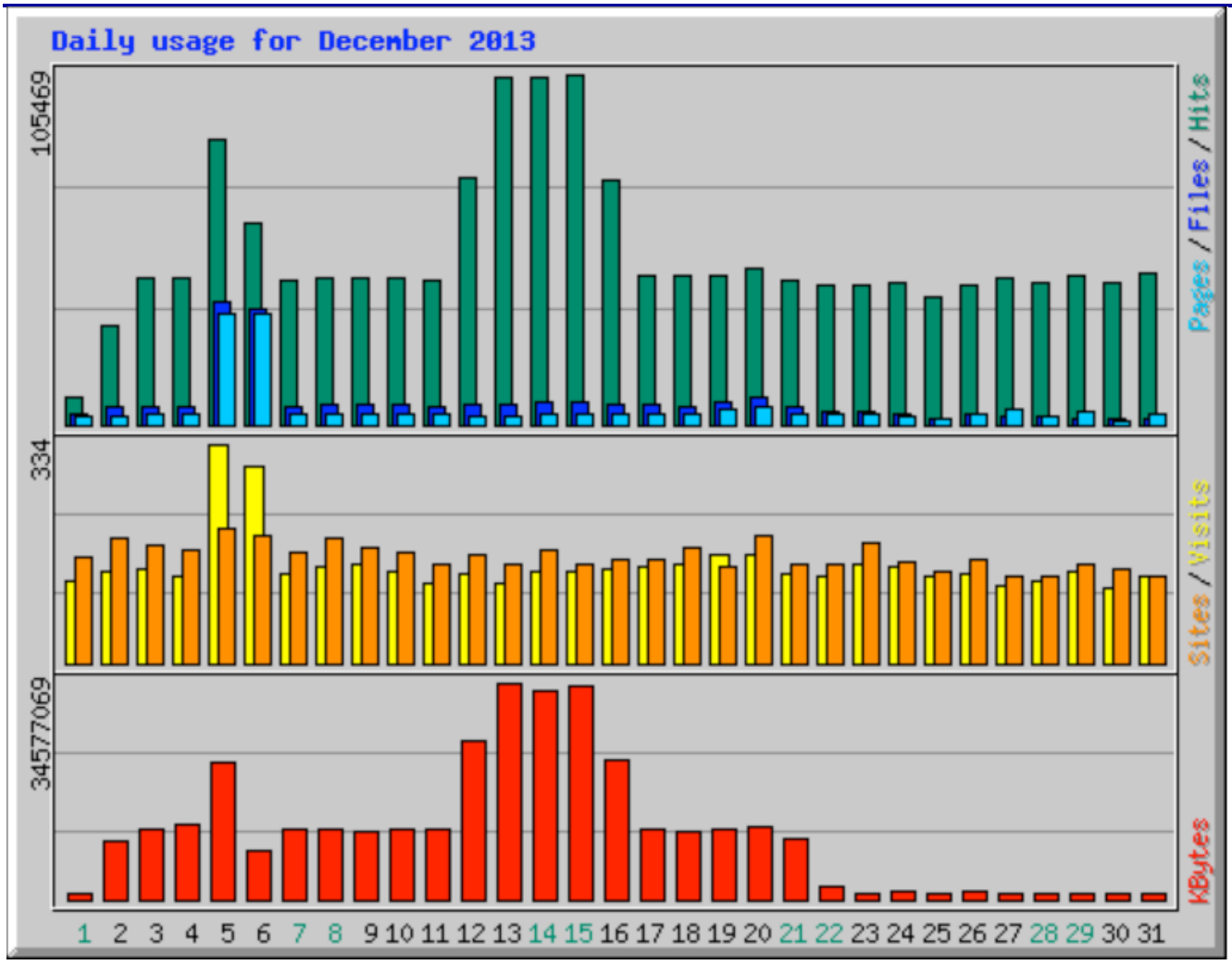

Usage by Country for September 2013
France (80%)
Belgium (17%)
network (.net) (1%)
commercial (.com) (1%)
unresolved (1%)

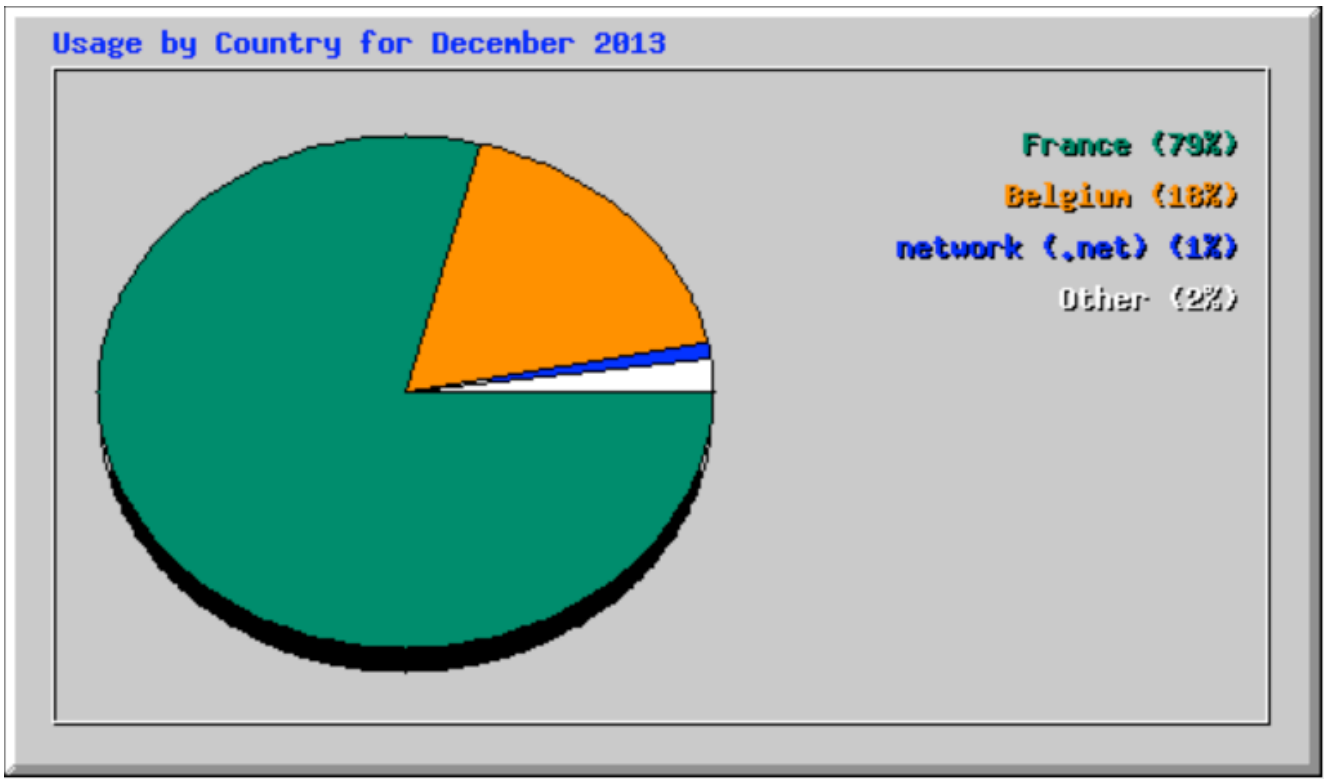

Daily usage for October 2013
80896
356
26005851
Pages / Files / Hits
Sites / Visits
KBytes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

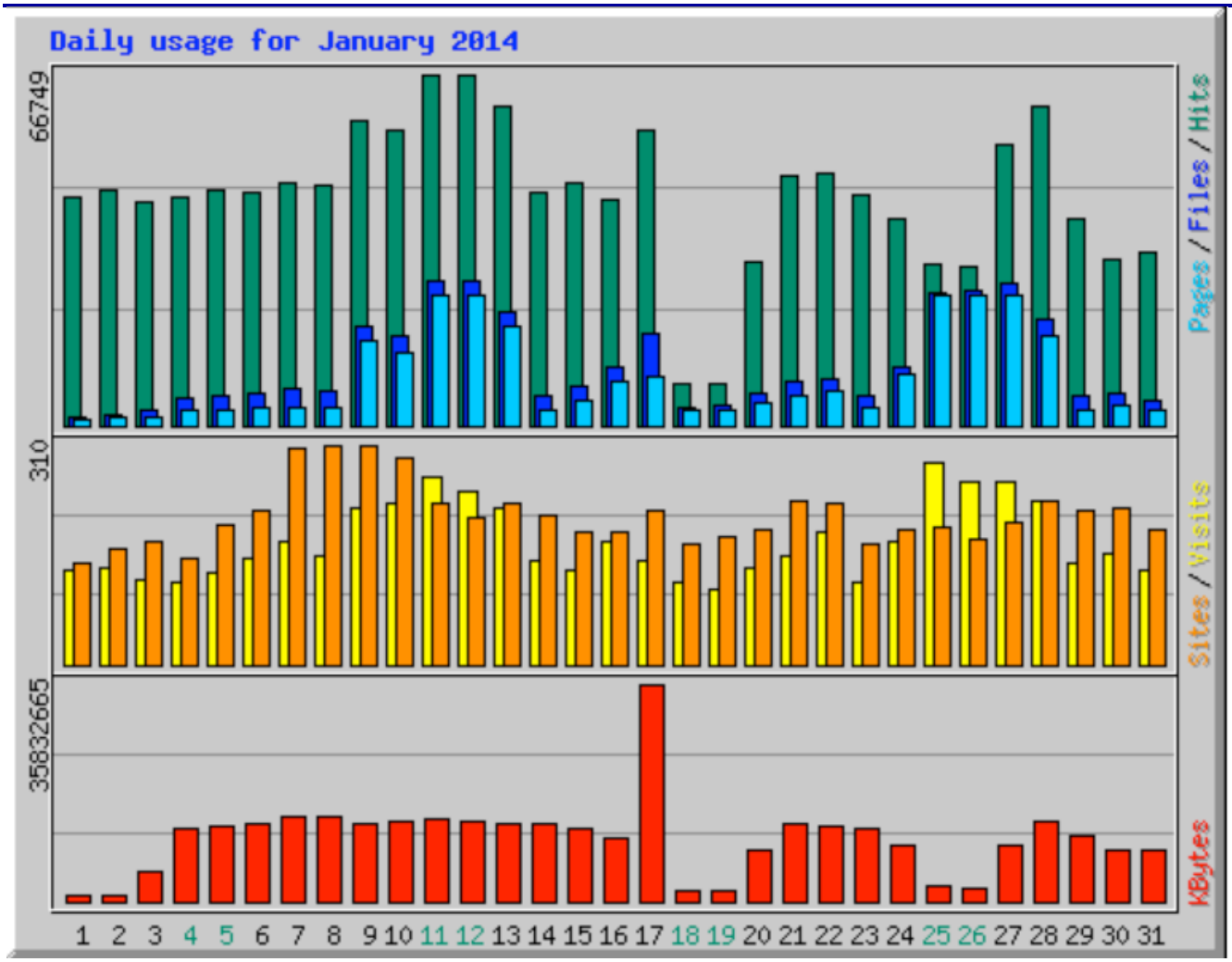

Usage by Country for October 2013
France (85%)
Belgium (12%)
network (.net) (1%)
commercial (.com) (1%)
Other (1%)

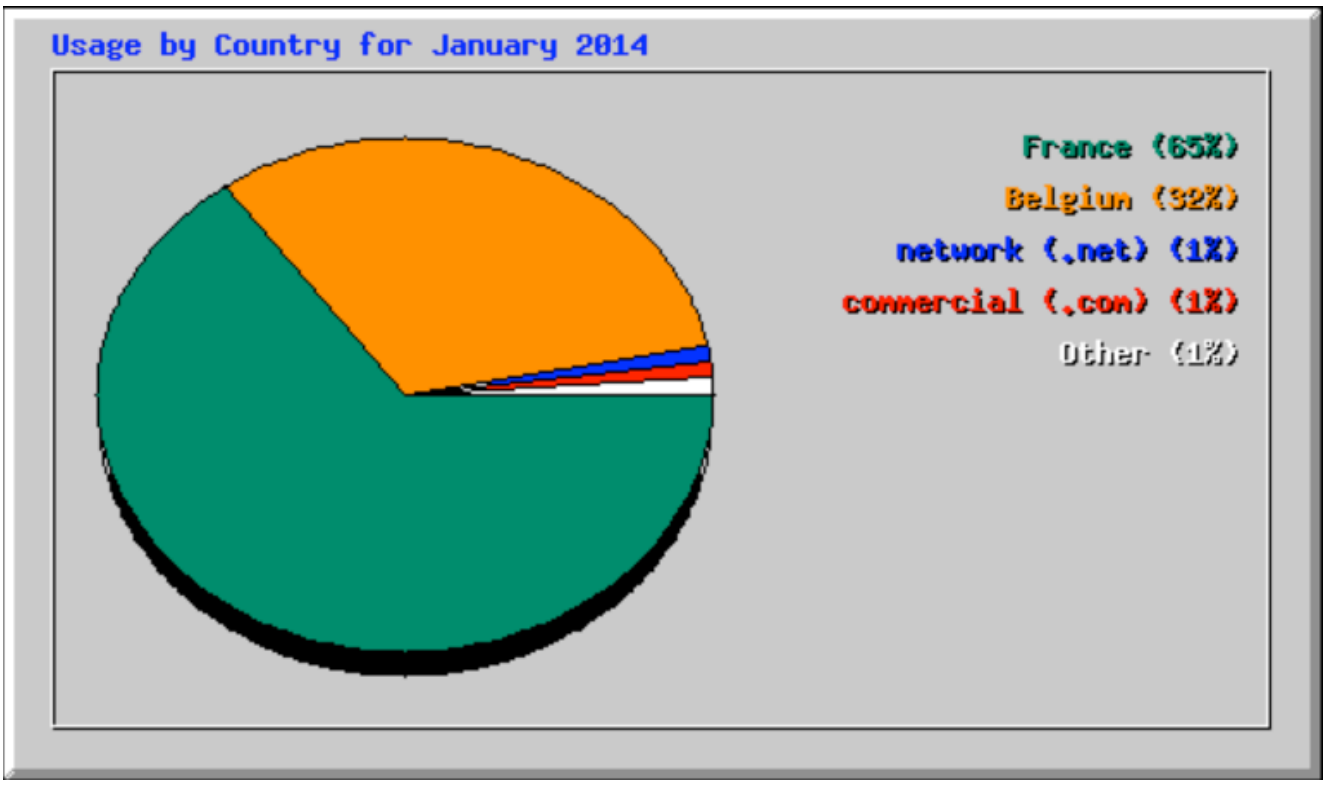

Daily usage for November 2013
120074
400
21206990
Pages / Files / Hits
Sites / Visits
KBytes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

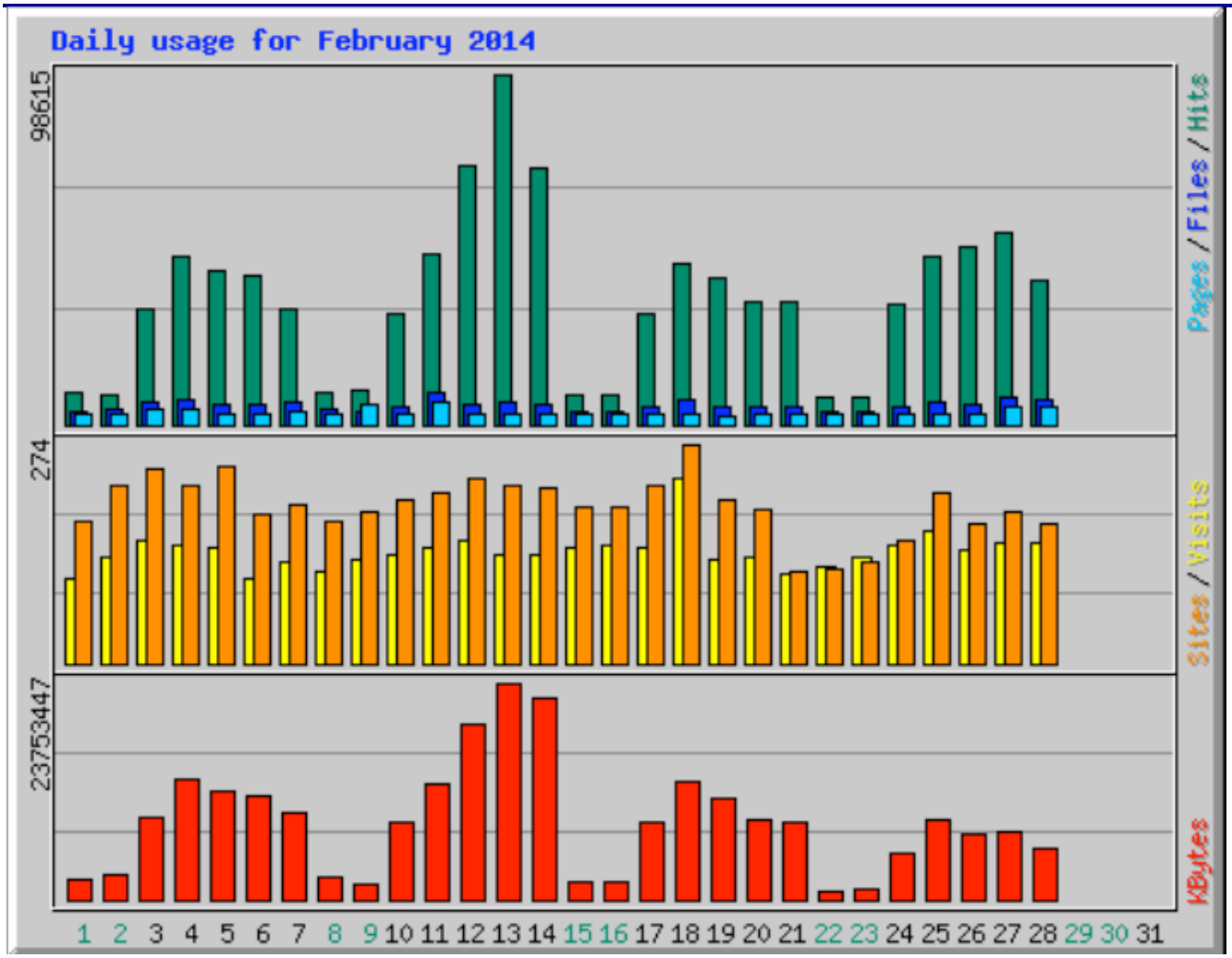

Usage by Country for November 2013
France (72%)
Belgium (26%)
network (.net) (1%)
Other (1%)

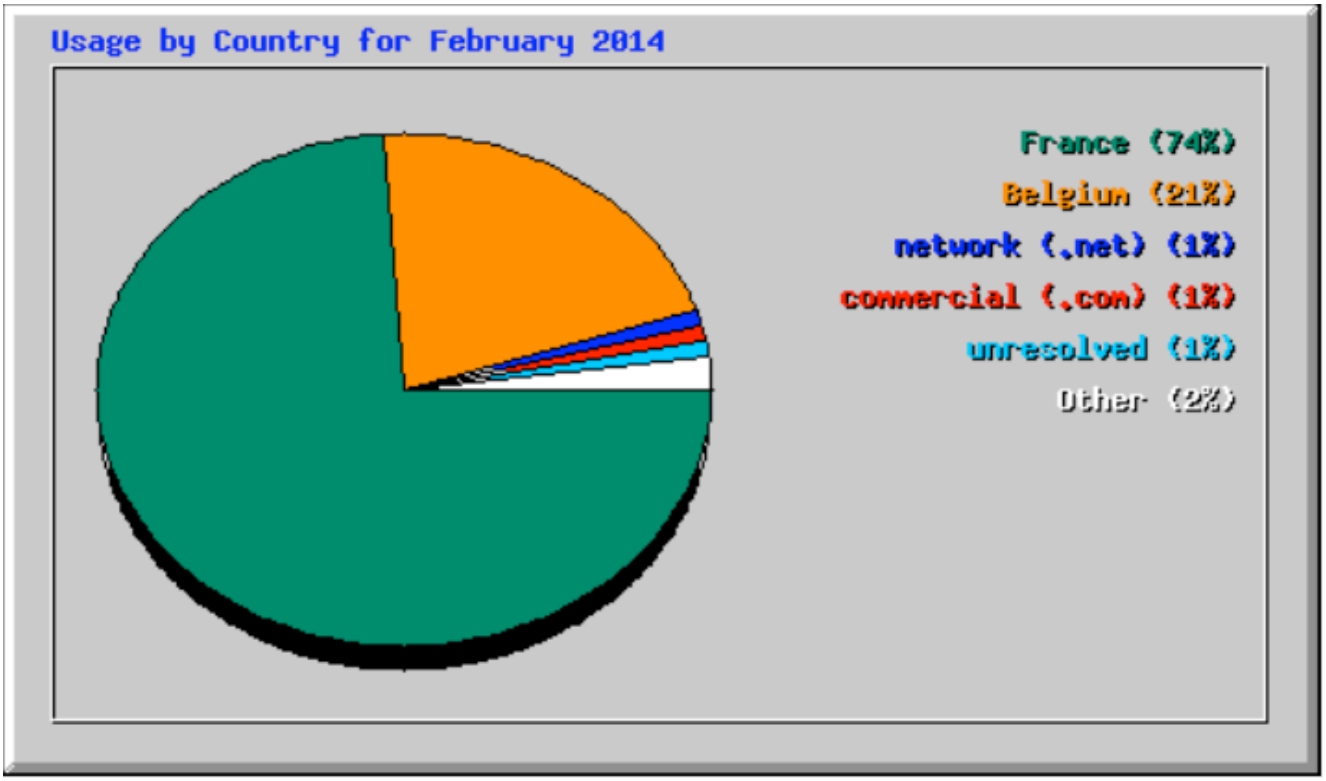

Daily usage for December 2013
105469
334
34577069
Pages / Files / Hits
Sites / Visits
KBytes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

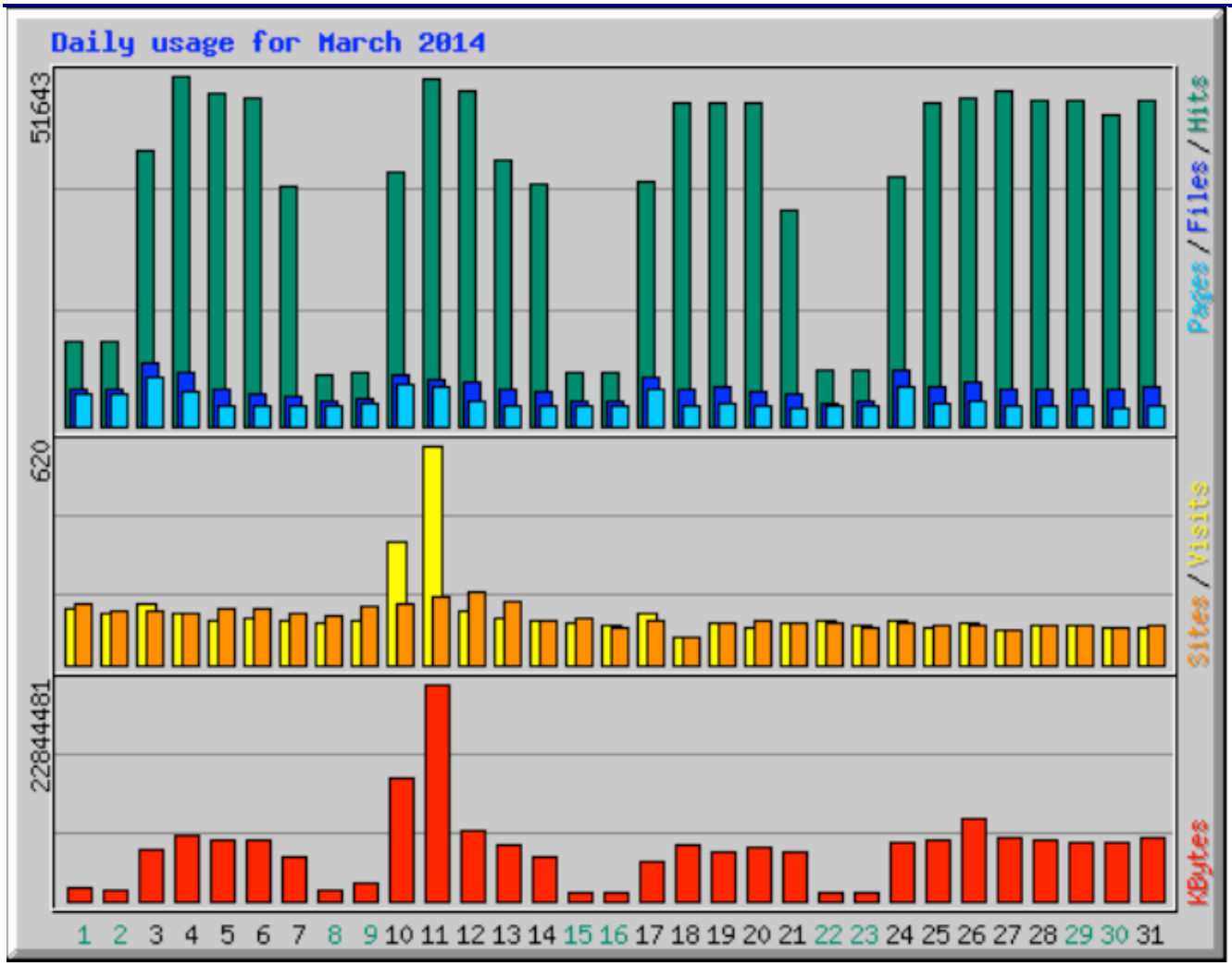

Usage by Country for December 2013
France (79%)
Belgium (18%)
network (.net) (1%)
Other (2%)

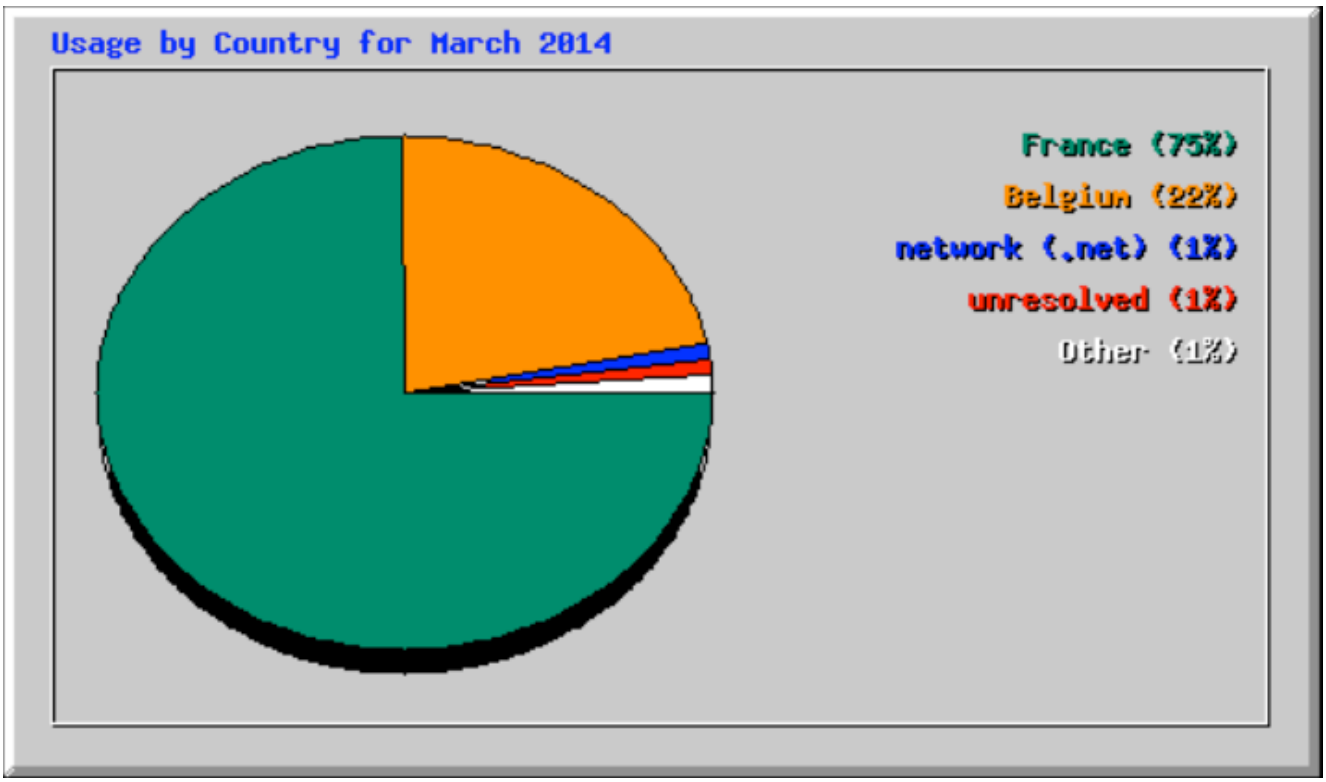

Daily usage for January 2014
66749
310
35832665
Pages / Files / Hits
Sites / Visits
KBytes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

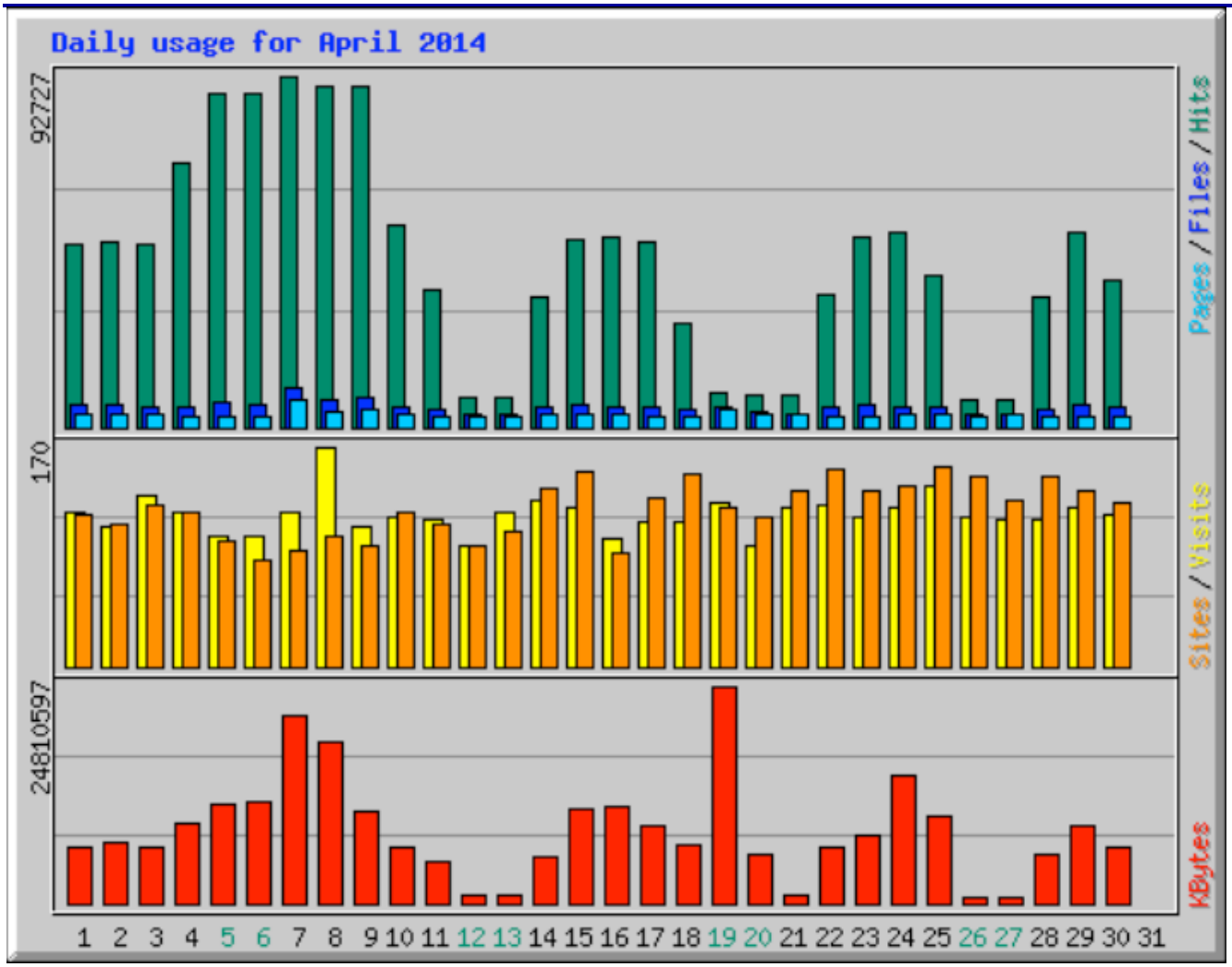

Usage by Country for January 2014
France (65%)
Belgium (32%)
network (.net) (1%)
commercial (.com) (1%)
Other (1%)

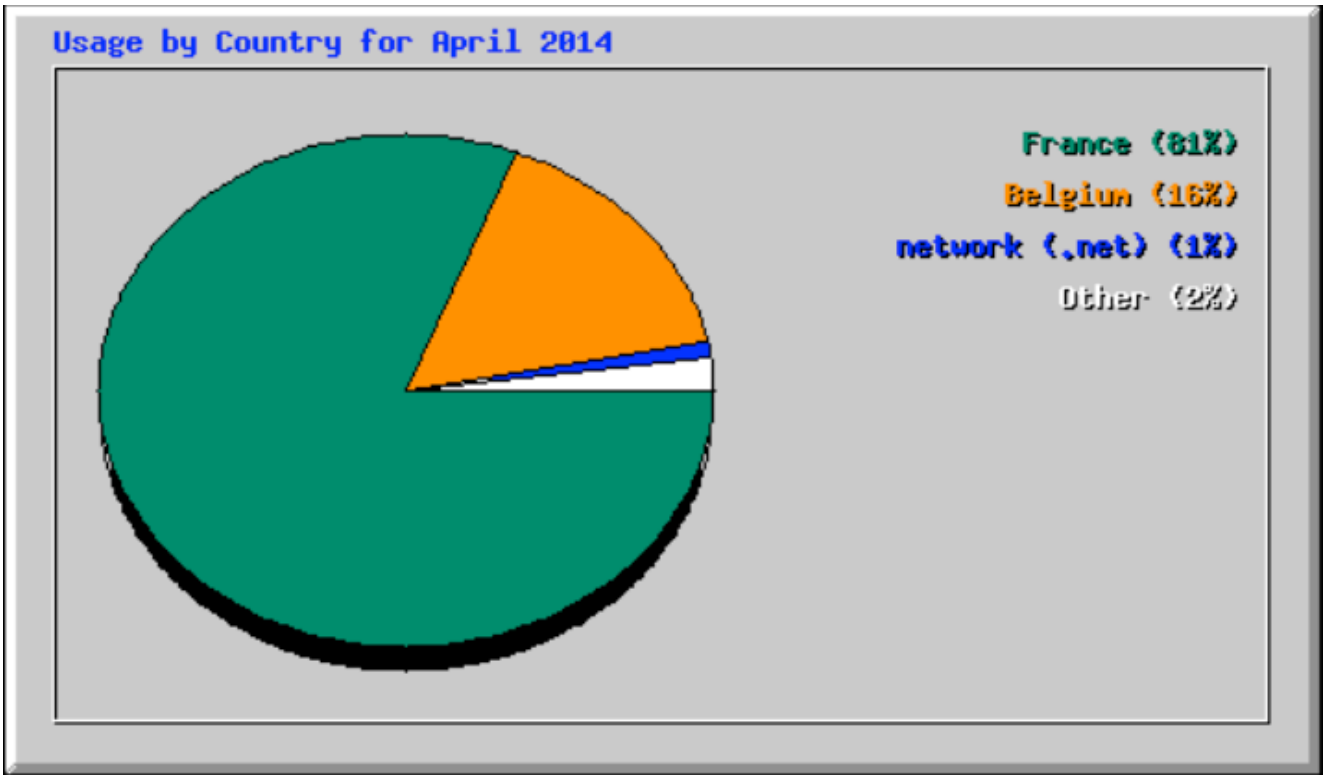

Daily usage for February 2014
98615
274
23753447
Pages / Files / Hits
Sites / Visits
KBytes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

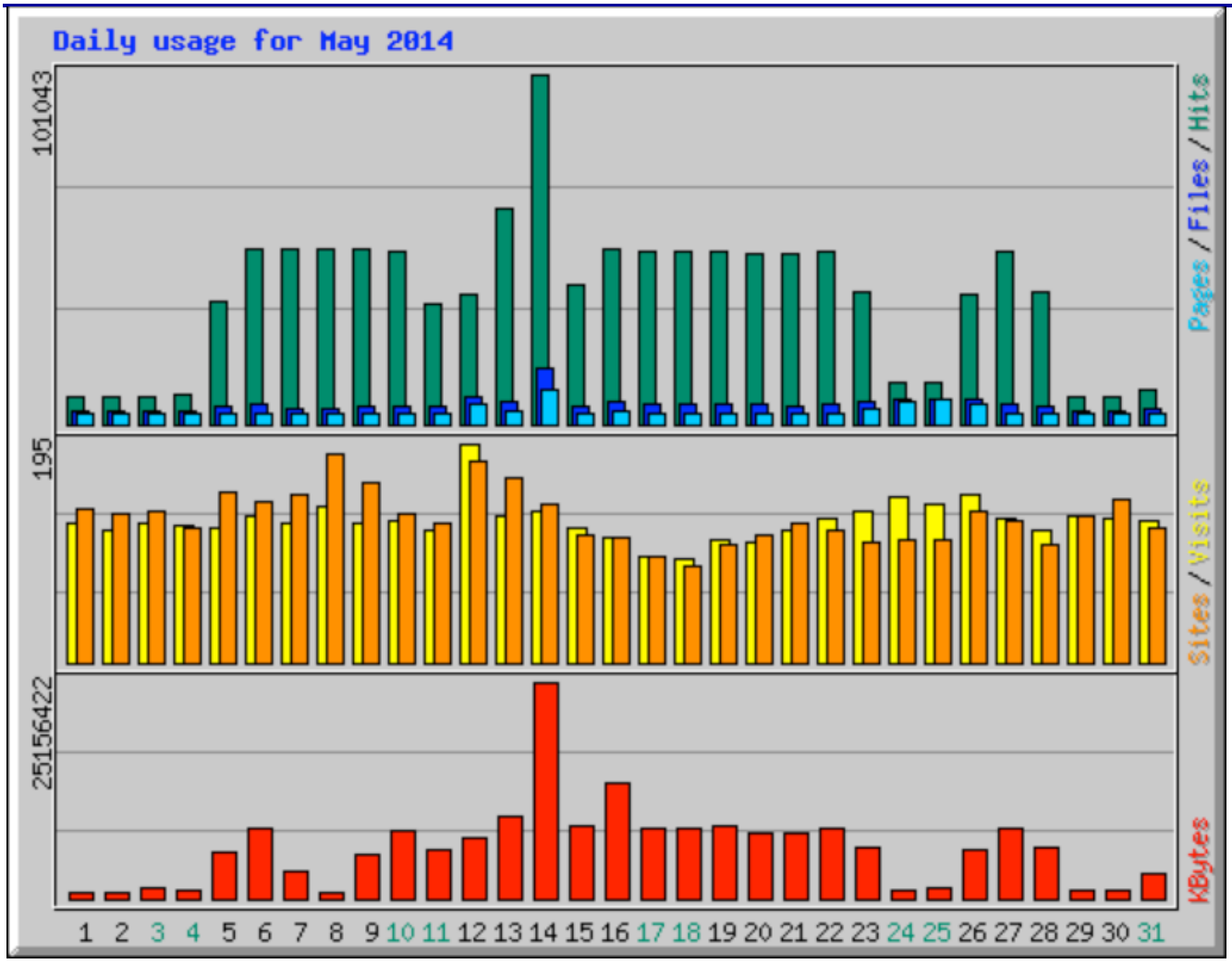

Usage by Country for February 2014
France (74%)
Belgium (21%)
network (.net) (1%)
connercial (.com) (1%)
unresolved (1%)
Other (2%)

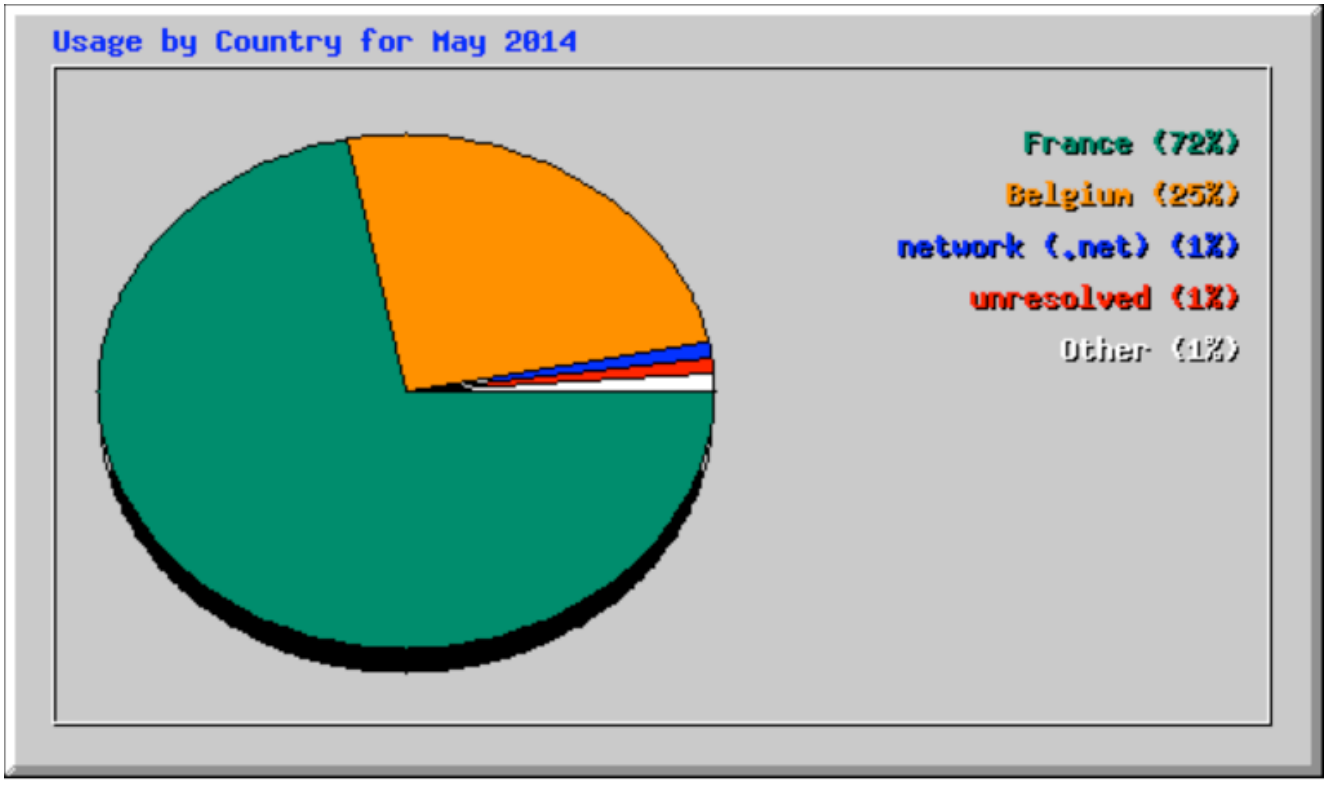

Daily usage for March 2014
51643
620
22844481
Pages / Files / Hits
Sites / Visits
KBytes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

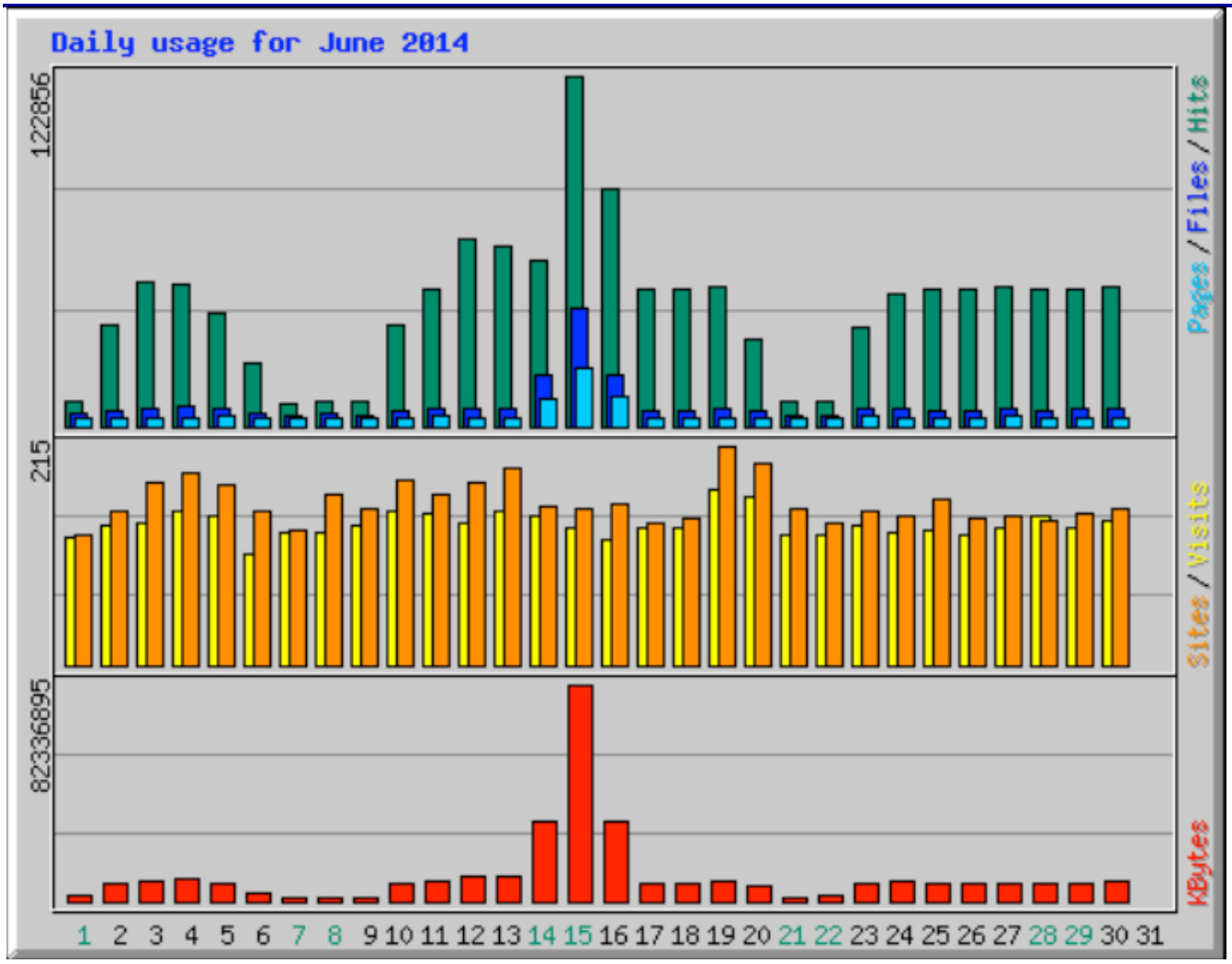

Usage by Country for March 2014
France (75%)
Belgium (22%)
network (.net) (1%)
unresolved (1%)
Other (1%)

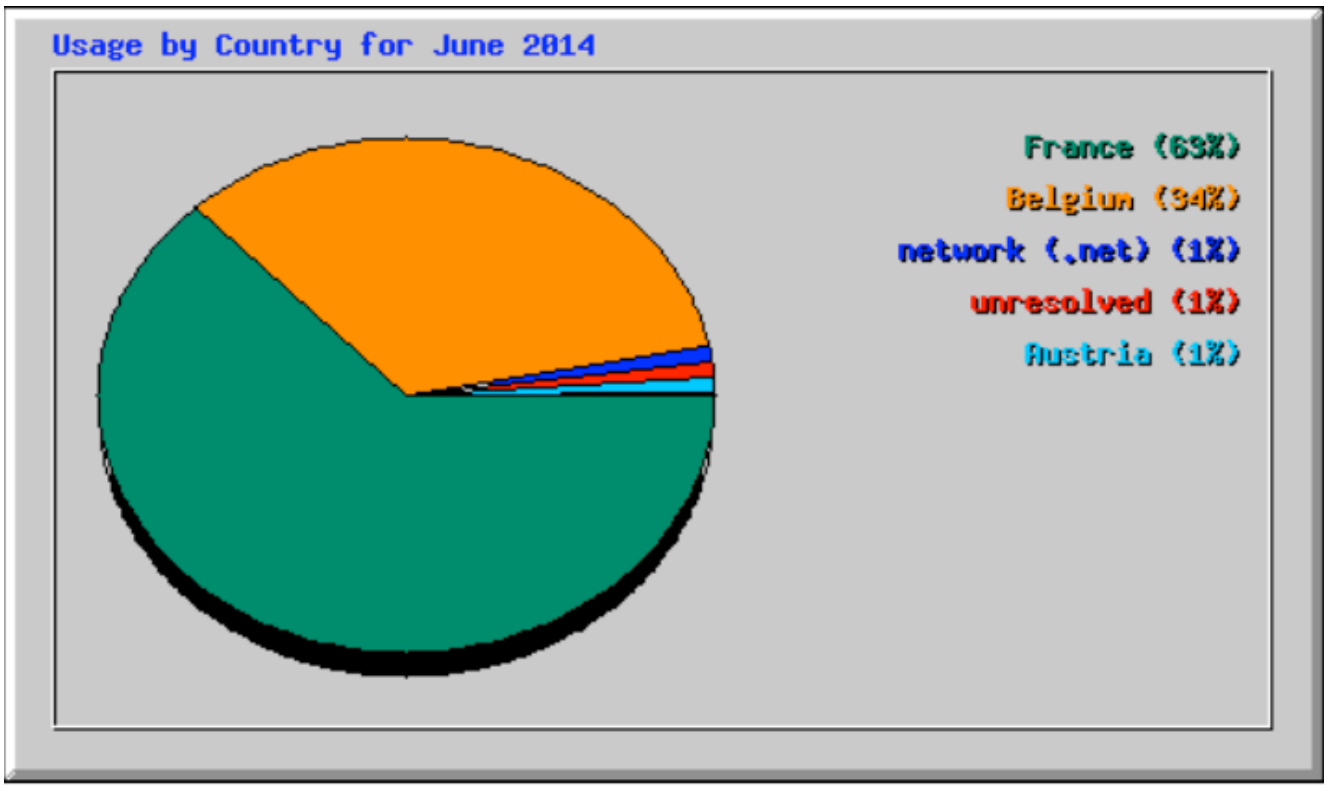

Daily usage for April 2014
92727
170
24810597
Pages / Files / Hits
Sites / Visits
KBytes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

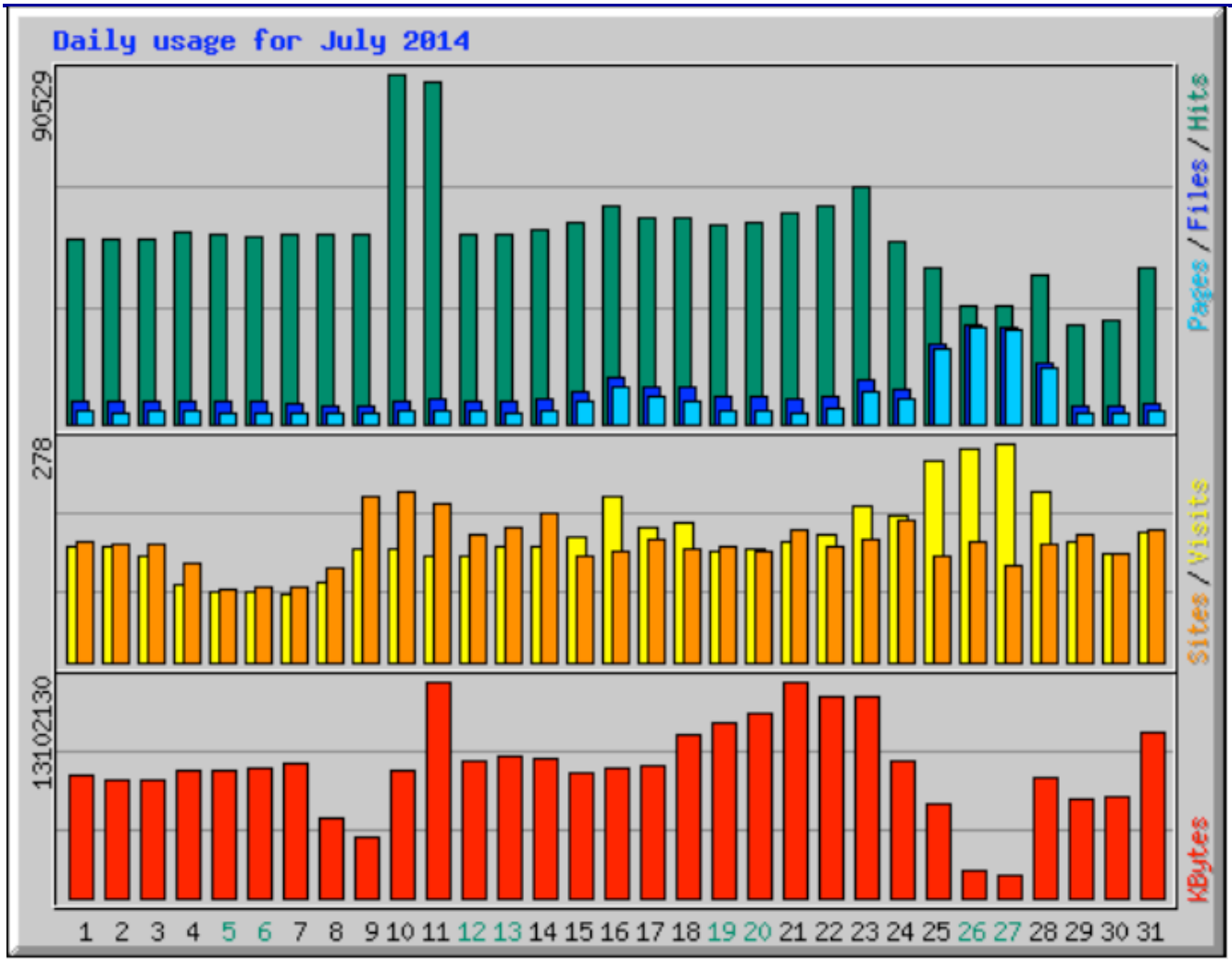

Usage by Country for April 2014
France (81%)
Belgium (16%)
network (.net) (1%)
Other (2%)

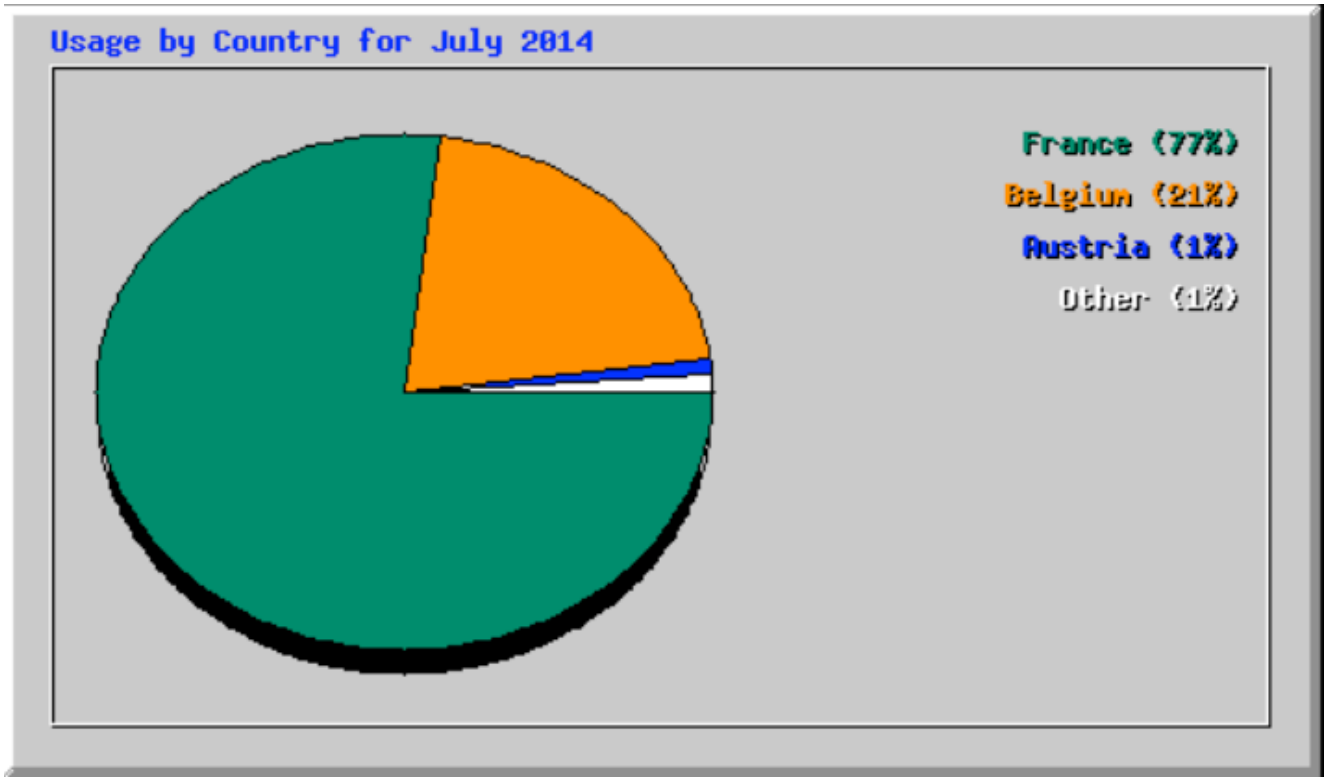

Daily usage for May 2014
101043
195
25156422
Pages / Files / Hits
Sites / Visits
KBytes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

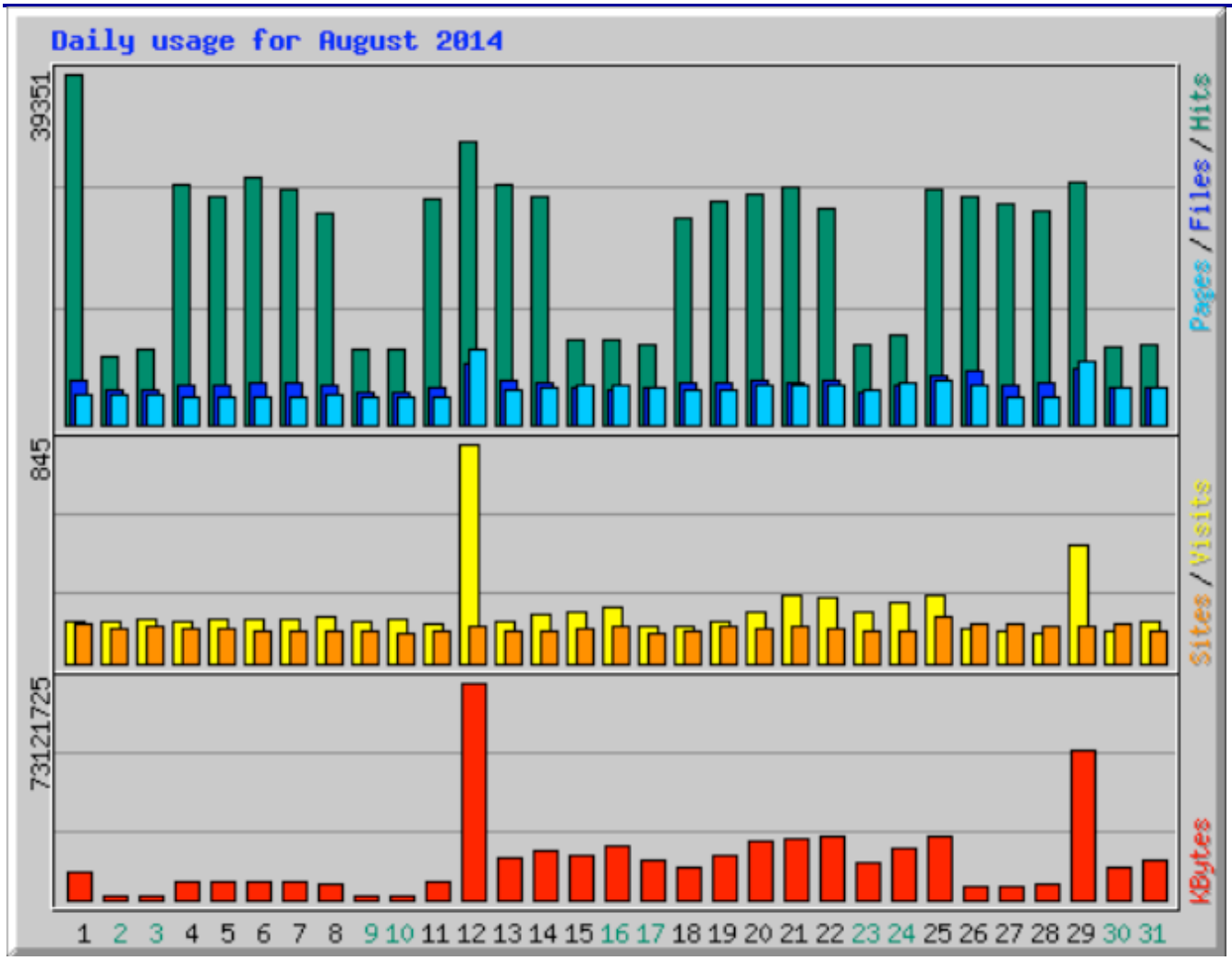

Usage by Country for May 2014
France (72%)
Belgium (25%)
network (.net) (1%)
unresolved (1%)
Other (1%)

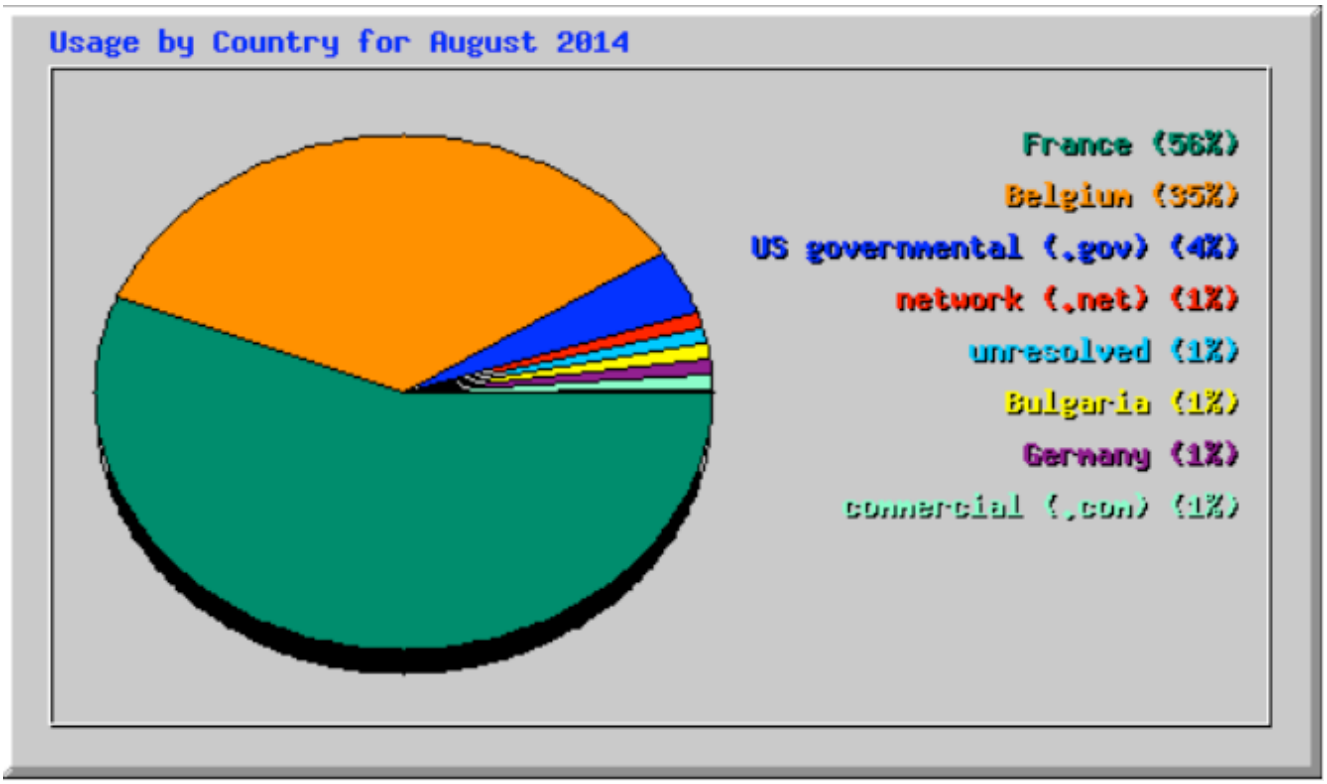

Daily usage for June 2014
122856
215
82336895
Pages / Files / Hits
Sites / Visits
KBytes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

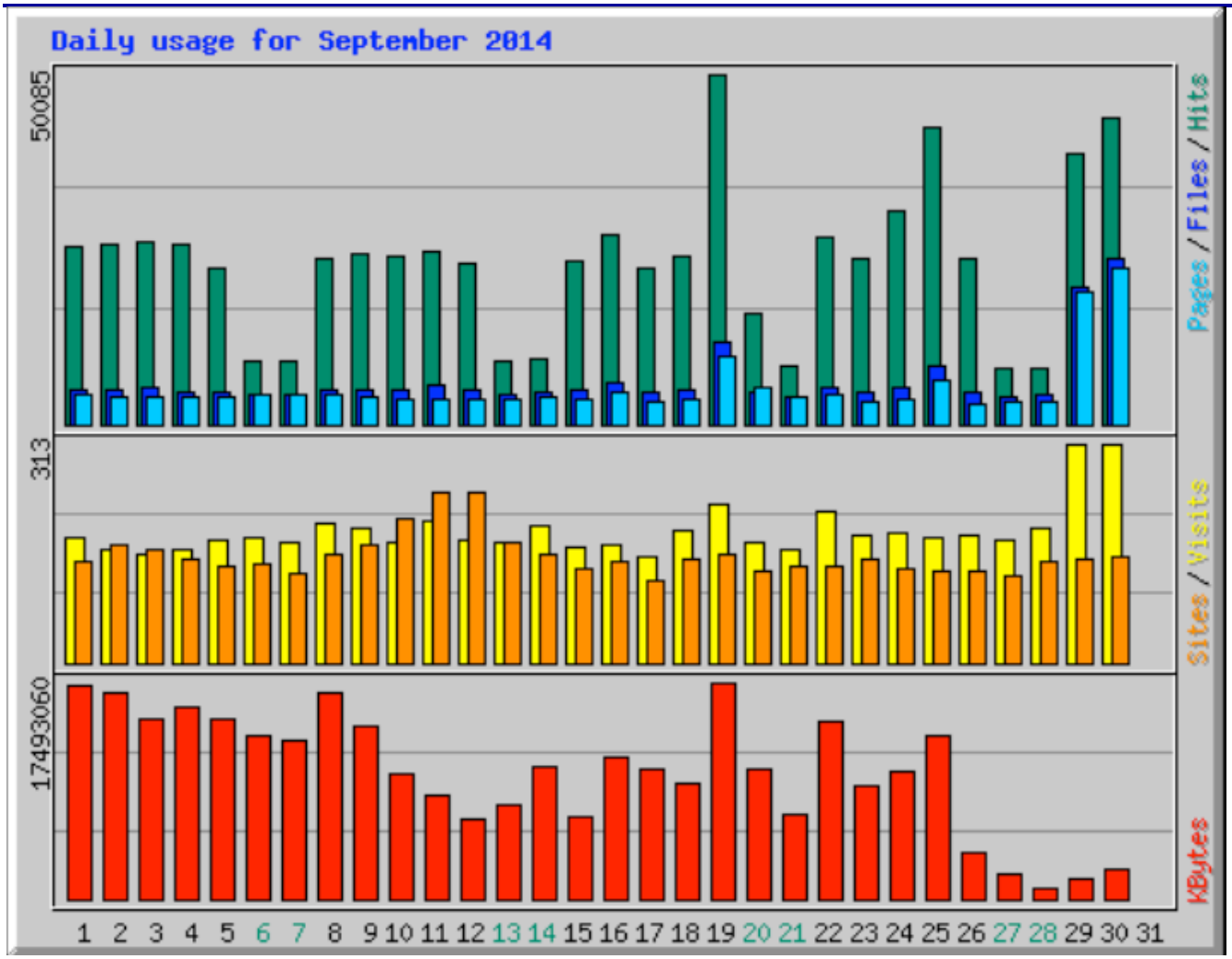

Usage by Country for June 2014
France (63%)
Belgium (34%)
network (.net) (1%)
unresolved (1%)
Austria (1%)

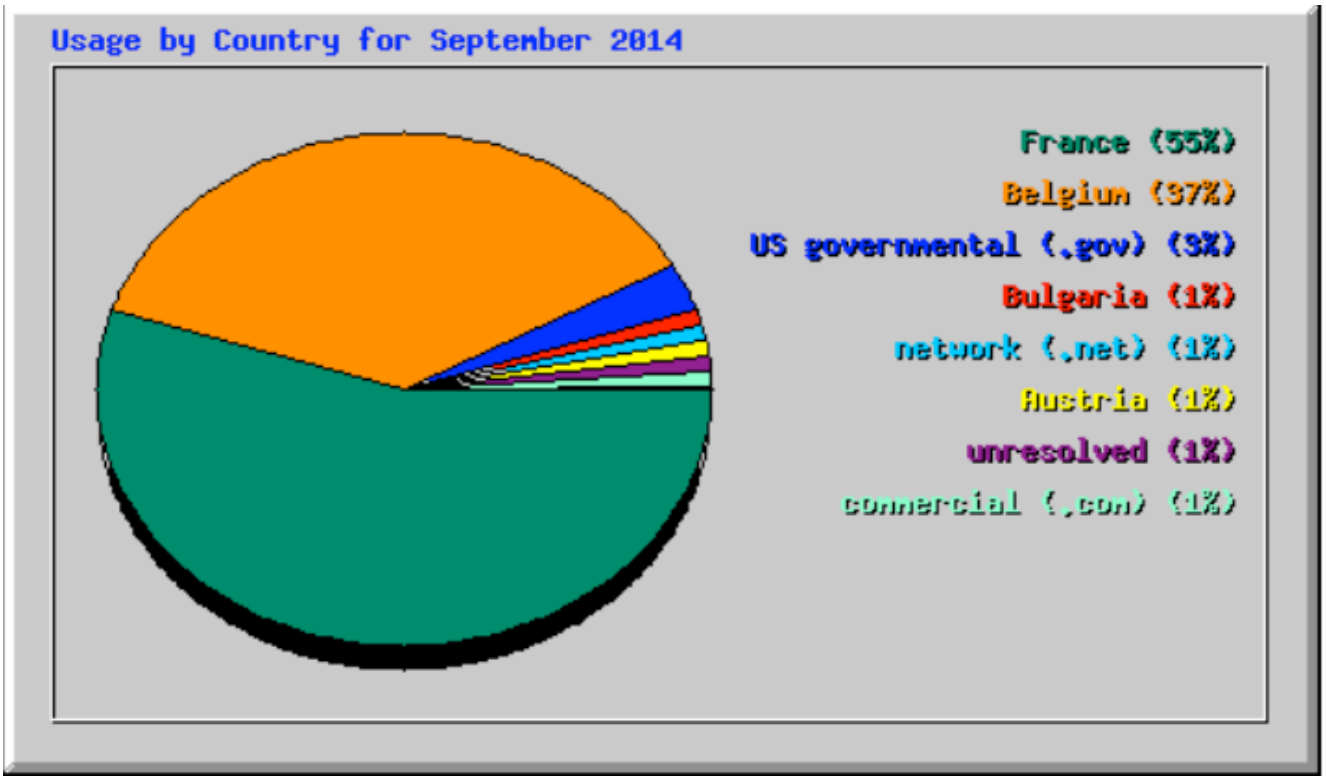

Daily usage for July 2014
90529
278
13102130
Pages / Files / Hits
Sites / Visits
KBytes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31


Usage by Country for July 2014
France (77%)
Belgium (21%)
Austria (1%)
Other (1%)


Daily usage for August 2014
39351
845
73121725
Pages / Files / Hits
Sites / Visits
KBytes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31


Usage by Country for August 2014
France (56%)
Belgium (35%)
US governmental (.gov) (4%)
network (.net) (1%)
unresolved (1%)
Bulgaria (1%)
Germany (1%)
commercial (.com) (1%)


Daily usage for September 2014
50085
313
17493060
Pages / Files / Hits
Sites / Visits
KBytes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31


Usage by Country for September 2014
France (55%)
Belgium (37%)
US governmental (.gov) (3%)
Bulgaria (1%)
network (.net) (1%)
Austria (1%)
unresolved (1%)
commercial (.com) (1%)

This project is supported by the European Commission's FP7 Capacities Programme for the period April 2013 - March 2017 under the Grant Agreement number 312495.
www.solarnet-east.eu

# SubWorkpackage 103: GSC-SDO. Access to Science Data Centres

Leader Organization: Max-Planck-Gesellschaft zur förderung der Wissenschaften e.V.

Participants: MPS (Laurant Gizon, Raymond Burston)

The German Science Center for the Solar Dynamics Observatory (GSC-SDO) is dedicated to the acquisition and distribution of data from the Helioseismic and Magnetic Imager (HMI) instrument that is onboard the Solar Dynamics Observatory (SD0). The GSC-SDO is hosted by the Max Planck Institute for Solar System Research (MPS) in Goettingen, Germany.

The ultimate goal of this work package (WP103) is to make HMI data (and other relevant data) available to the European solar physics community through the Data Record Management System, NetDRMS. NetDRMS is a software suite developed by Stanford University specifically for managing SDO observations and sharing the data worldwide:

http://jsoc.stanford.edu/netdrms

NetDRMS is fully operational at the GSC-SDO and data are automatically transferred directly from Stanford University as soon as they are available. The GSC-SDO has a storage capacity of 800 TB (online and near-line) dedicated to SDO/HMI data and other related data products.

The NetDRMS system at the Kiepenheuer-Institut für Sonnenphysik (KIS) in Freiburg, Germany was recently reconstructed. This was due to new personnel at KIS plus new and expanded hardware infrastructure at KIS. The MPS provided significant support and training to help get the renewed NetDRMS system operational at KIS.

KIS have now "subscribed" to the several data series from the HMI instrument. They are also subscribed to data series from HMI's predecessor, the Michelson Doppler Imager (MDI) onboard the Solar & Heliospheric Observatory (SOHO).

When a NetDRMS site subscribes to a particular data series, that site by default receives all the meta-data, i.e. keyword, value pairs. Thus, KIS has all the meta-data for the following data series:

- HMI Dopplergrams at a 45s cadence
- HMI Magnetograms at a 720s cadence
- HMI Continuum at a 720s cadence
- MDI Dopplergrams at a 60s cadence
- MDI Magnetograms at a 96m cadence

There is a separate mechanism that must be triggered in order to also receive the images. Furthermore, this can be done selectively so you can choose which images you want in your local NetDRMS system. This feature is particularly useful if local storage resources are not sufficient to store all the images, allowing the site to only transfer the images of interest.

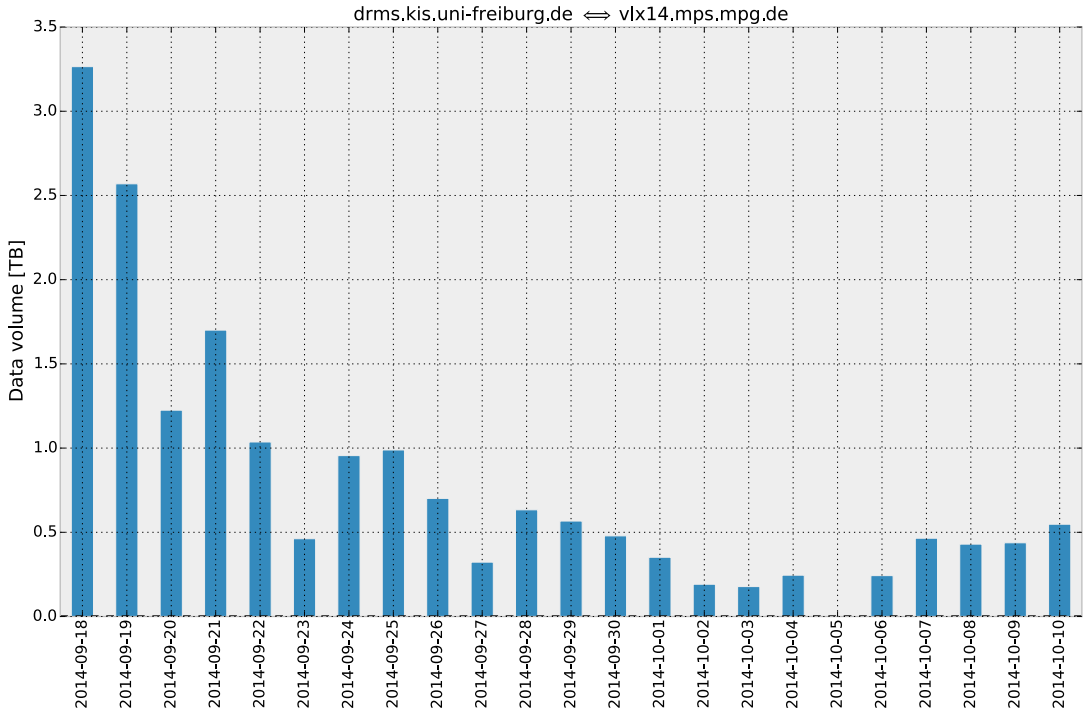
At the time of writing this report, KIS had already transferred the following images from the GSC-SDO into their local NetDRMS system in Freiburg:

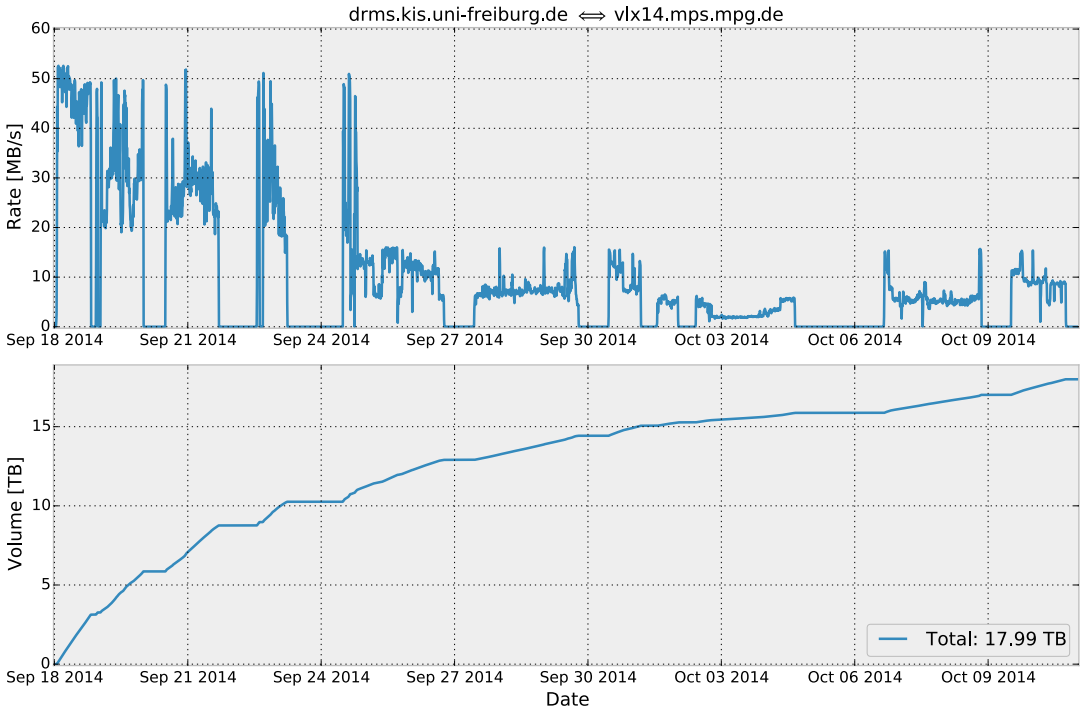
- 12 months of HMI Dopplergram images at a 45s cadence
  - 350400 images (4096x4096 pixels)
  - 12 TB
- 6 months of HMI Magnetogram images at a 720s cadence
  - 21900 images (4096x4096 pixels)
  - 352 GB
- 6 months of HMI Continuum images at a 720s cadence
  - 21900 images (4096x4096 pixels)
  - 352 GB
- all the Doppler grams from MDI
  - 5 TB

The current data transfer rate from MPS to KIS is around 200 Mbps, which means it would take several days to transfer 1 year of HMI Dopplergrams from the GSC-SDO.

In the future, KIS intends to expand on their current HMI archive and transfer all the images for the HMI Dopplergrams (45s cadence), HMI Magnetograms (720s cadence), and HMI Continuum (720s cadence). Furthermore, this will be fully automated using NetDRMS utilities. The MPS has already started supporting the implementation of a fully automated system at KIS.

Figures 1 and 2 shows how the data volume has grown plus data transfer rates since the new implementation of NetDRMS was functional at KIS






In addition, the IAS in France also wants to receive data from the GSC-SDO via NetDRMS. Initial discussions have commenced and support from the MPS is expected to begin in the coming months.