

# IGN 2002 Global Data Center Report

*Édouard Gaulué*

Ecole Nationale des Sciences Géographiques  
6-8 Avenue Blaise Pascal  
Cité Descartes / Champs-sur-Marne  
77455 Marne-la-Vallée CEDEX 2  
France

## **Introduction**

The *Institut Géographique National* (IGN) has been involved in IGS since its beginning through the *Laboratoire de Recherche en Géodésie* (LAREG). For a long time, a single person has administrated IGN Global Data Center (GDC): Loïc DANIEL. Since May 1<sup>st</sup> 2001, Edouard GAULUE has been appointed to assist him in his mission as IGS services around data kept growing (new stations, products and data transfer strategies, web developments). Here is a summary of IGN GDC activities from September 2001 to December 2002.

## **Service Overview**

As a global data center, the IGN server gets observation files from regional and operational data centers and product files from analysis centers. All collected files (since 1997) are accessible to the user community through anonymous FTP. Moreover IGN also mirrors the central bureau directory for administrative reasons.

## **History**

### *Hardware*

IGN GDC has been working on a VAX/VMS machine until 1997. Many of the files obtained during this period had been archived on magneto-optical disks (75 Gb) that we expect to restore soon. In 1997, this server has been changed to an HP D230 machine running HP-UX. At the same time, the service moved to the *Ecole Nationale des Sciences Géographiques* (ENSG) equipped with a 2 Mb/s bandwidth Internet link. Five years later, due to the addition of external disks, a CD jukebox and a saving-robot, this machine has grown from half a cubic meter to 3. It has also become a little obsolete. Two machines have been ordered at the end of 2001 and should be delivered in 2003. In February 2002, the bandwidth of the ENSG has raised to 10 Mb/s.

### *Online service*

In the middle of 2001, the FTP server has been refurbished taking into account a new stations (and name changes) and new products. In 2002, a new web site has been developed but it had to remain on a test server because of the delay in the delivery of the new servers.

Due to everyday service, our first 150 CDs jukebox accessible online is today more than half-full (about 90 Gb). In addition to the jukebox, 25 Gb are reserved for data storage on the server hard disks. Files online on hard disks represent about 200 days. After this time a CD is burned and stored in the jukebox from which data are still accessible online.

### *Service availability in 2002*

In 2002, the IGN GDC FTP service has not experienced an interruption exceeding 3 days. Most of the interruptions were due to (i) our Internet provider during the bandwidth change, or (ii) the lack of space on the server. In fact, the principal difficulty we had to deal with in 2002 came from the continuous increasing delay concerning our new machines arrival.

### **How Does IGN GDC Work?**

A diagram presenting the operating mechanism of the IGN IGS global data center server is given in Figure 1.

IGN GDC stores information about partner FTP sites in a MySQL database. In the meantime, using an automatic procedure on sitelog it also gets information concerning station sites. A third table links stations and centers according to the different information sources IGS provide and IGN GDC policy. This table gathers the station observations files (as well as navigation, meteorological and quality checks) for each center. For products, filenames are directly store with partner FTP sites information. Every night, based on these information, a “mirror configuration file” is generated for each partner data center.

This file is then processed by the so-called “mirror” program. This program questions the distant site and compares it to the local one before starting transfers. Files arrive directly in the public area. This procedure is the same for data centers which “put” data to IGN GDC: the distant site in this case is their local deposal directory at IGN. More information concerning the mirror program is displayed on Figure 2.

All mirroring tasks are scheduled using the CRON system. Moreover tasks are placed in specified queues depending of their aim to avoid download conflicts and control server overload.

Once in public area, automatic files scanning is done either for IGS community through check import or to feed our database. The “file” table is then used for statistical analysis, building data and products holding reports and web requests.

### **IGN Policy**

#### *Near Real Time (NRT) data*

As NRT data doesn't need to be archived, IGN GDC tries to get data from a maximum number of stations (about 100 in 2002).

### *Observation files*

Due to the server limitation, we prefer not to deal with more than 150 stations. A new policy will be defined when new machines will be in place. In 2002, priority is first given to regional stations (coming mainly from IFAG, CNES and ORSTOM) and then to global station. A few other stations have been dealt on request.

From May 2001 to the end of 2002, the number of processed stations rose from 70 to 150.

### *Product files*

In 2002, IGN GDC tried to deal with more product files, taking into account new ultra rapid products. Moreover ionospheric and tropospheric data are always accessible in special tropo and iono directory.

### *GLONASS*

No GLONASS data has been processed since the end of 2001. It has been decided at the Ottawa workshop to improve IGEX services by gathering observations and products in the same directories than IGS. All the IGN GLONASS services will be refurbished and implemented in 2003.

### *DORIS*

Since December 2002, IGN has become a global DORIS data center. This event implied a transfer of the service to another server. Concerning observation, IGN DORIS GDC is just a mirror of the CDDIS archive. Concerning products, everything as been done to propose IGS-like services.

## **IGN GDC Usage in 2002**

### *Incoming files*

Figure 3 displays the average delay for incoming daily observation files in 2002. We can point out that 2/3 of the files arrive in less than 3 hours and 3/4 in less than half a day.

Figure 4 displays the same information for NRT data. One can notice that data arrival at IGN is in fact largely driven by the scheduled time in the CRON. The first scan of remote data centers begins approximately 10 minutes after the beginning of each hour and lasts 10 minutes. The next scan (for “*retardataires*”) starts 45 min past the hour. The last files of the day (x files) follow a particular schedule for administrative reason. That explains why download begins 30 minutes after the end of the day.

Figure 5 and 6 show mean delays for data or products download at IGN GDC according to their origin. The delays for incoming observation are rather short. The mean delay for IGS final products arrival at IGN was about 17 to 18 days in 2002.

## *Outgoing files*

Figure 7 and 8 show general statistics on observation and data downloads from IGN server. On the observation graph, the three first months show an unusual download: someone was downloading identical observation files 64 times a day. Nevertheless we can remark a systematic trend during spring and summer months certainly due to the global activity.

Figure 9 shows downloads and associated average transfer rates grouped by Internet domain. We can notice a particularly good connection to Luxembourg and Switzerland certainly due to our provider connection.

## **Future Development**

In 2003, the IGN GDC activity will critically need its new computers to function properly. This would be the opportunity to completely rethink its procedures and use newer and better tools. Servers should be running under debian Linux distribution and should provide a MySQL 4.x database and a PHP-apache web server. Those machines will be shared with the permanent French geodetic GPS array. They should in term work in a mirroring scheme, one in Paris and the other one in Marne-La-Vallée with a heart-bit connection. Load balancing will have to be considered.

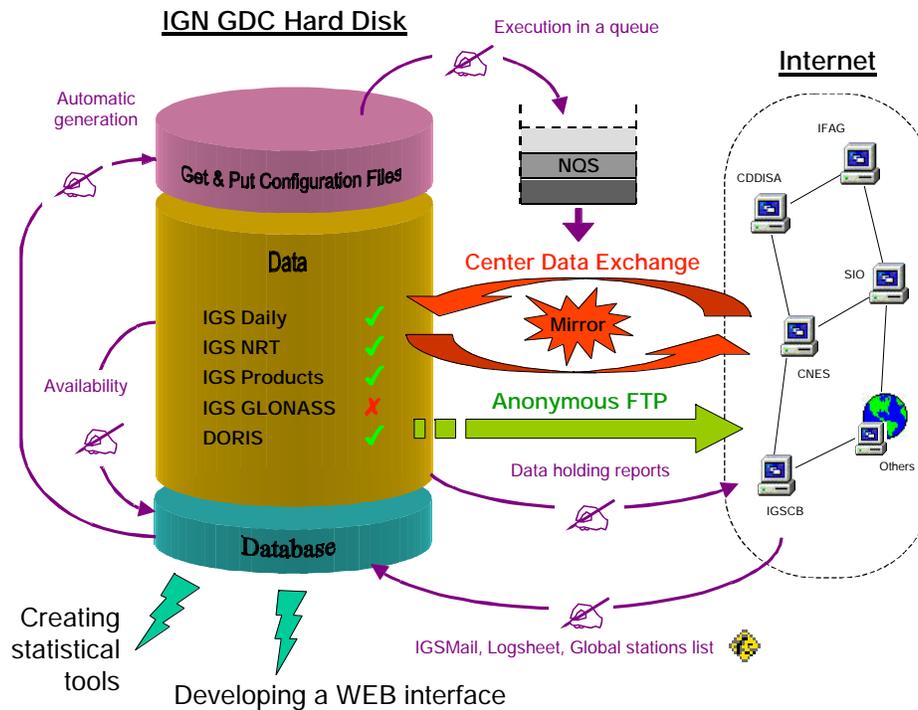
IGN GDC would also like to participate in real time transfer experiments. Contacts have been taken in Ottawa and at AGU to move towards this objective.

Last, IGN GDC will have its new web site online. Figure 10 displays a screenshot of what it could look like. Some modules are already operational but lots of work still has to be done until official opening. In relation to this topic, the IGN GDC team is interested to exchange on topics related to XML and particularly by the way to introduce this technology in IGS documentation.

## **Contacts**

*IGN GDC Team :*

Édouard GAULUÉ & Loïc DANIEL  
Ecole Nationale des Sciences Géographiques  
6-8 Avenue Blaise Pascal  
Cité Descartes / Champs-sur-Marne  
77455 Marne-la-Vallée CEDEX 2  
France  
Phone: (33) 1.64.15.32.43  
FAX: (33) 1.64.15.31.07  
Email: [edouard.gaulue@ensg.ign.fr](mailto:edouard.gaulue@ensg.ign.fr)



Global IGS GDC functioning - Edouard GAULUE - 2003/4/10

**Figure 1.**

## Transfers and monitoring: The mirror program

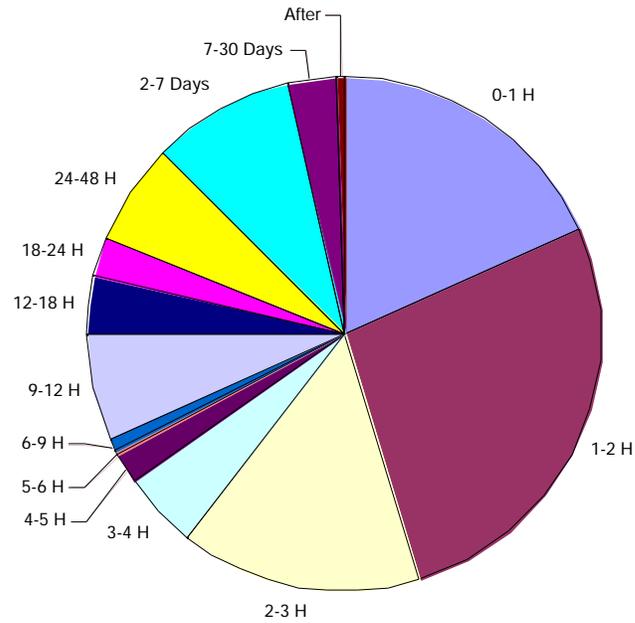
- **Actions:**
  - Make a FTP connection to the remote site
  - List remote files (or use ls-lR)
  - Compare this list to the local one according to numerous criteria (size, date, type, compression, ...)
  - Only get/put/delete/modify necessary files
  - Close the connection and writing log
- **Pros and cons:**

<ul style="list-style-type: none"> <li>+ Got a proper FTP implementation,</li> <li>+ Manage simultaneous transfers to remote site</li> <li>+ can be driven through configuration files</li> <li>+ detailed logging: help issues detection and allow statistic making</li> <li>+ open source: largely used and tested, free, adjustable</li> </ul>	<ul style="list-style-type: none"> <li>- hard to handle</li> <li>- few bugs haven't been solved</li> <li>- need much memory</li> </ul>
---	--

Ottawa IGS Workshop - April 2002

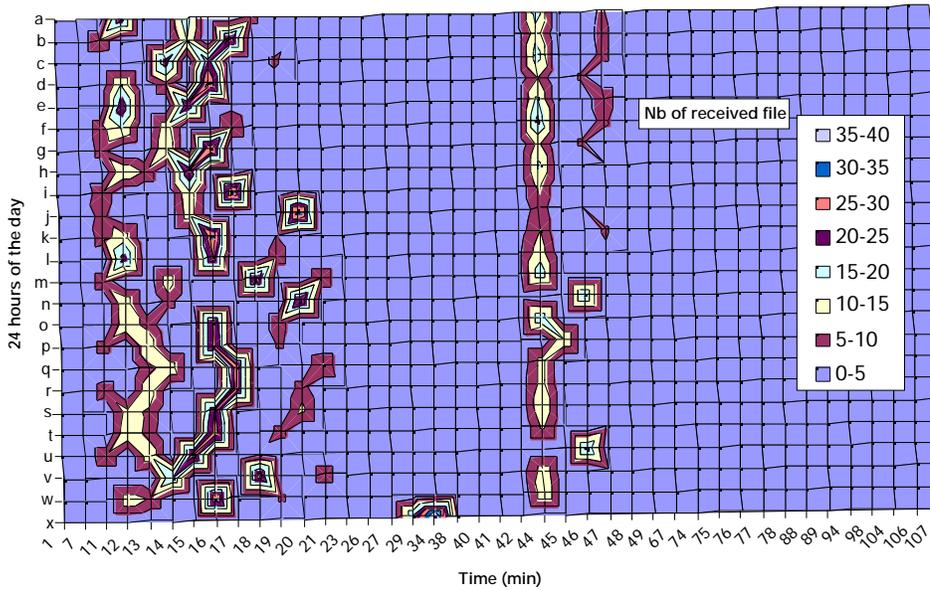
**Figure 2. Mirror Program**

Delay from station to IGN GDC for 2002 observation files  
(53 987 files)



**Figure 3.**

Delay of NRT files per hour for day 2002-090



**Figure 4.**

Median delay and received files by center at IGN GDC in 2002

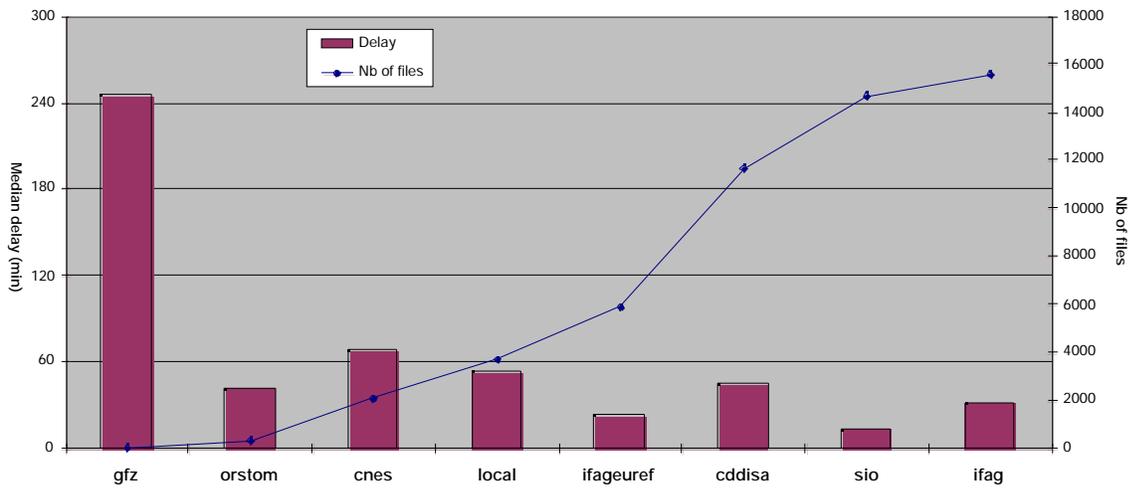


Figure 5

Mean delay for various products from various sources received at ign in 2002

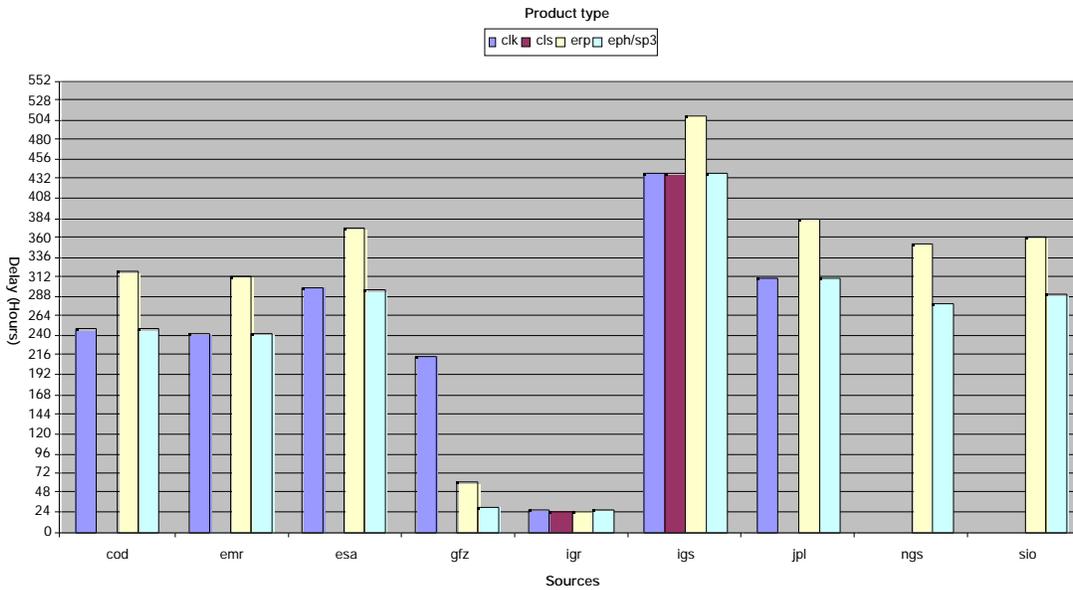
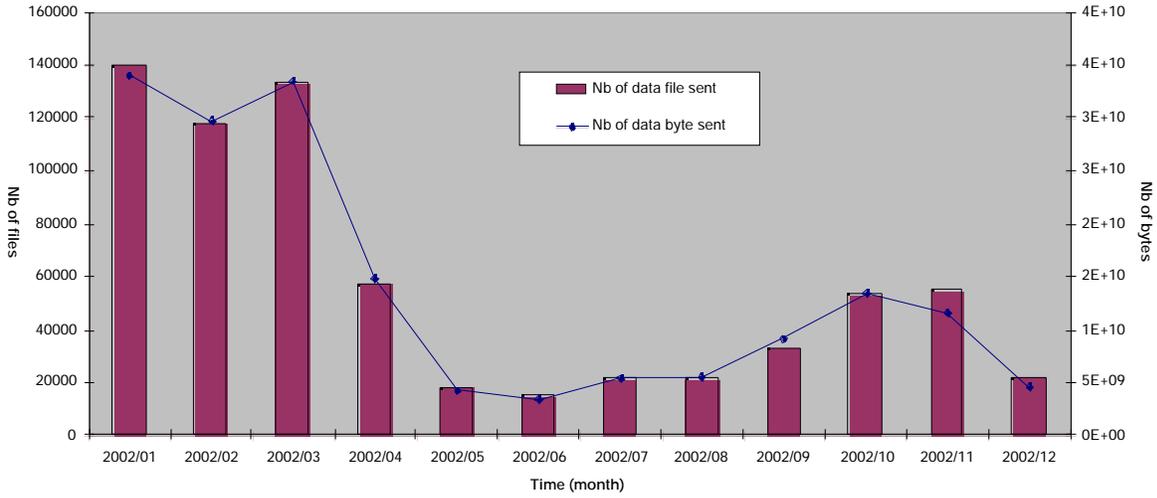


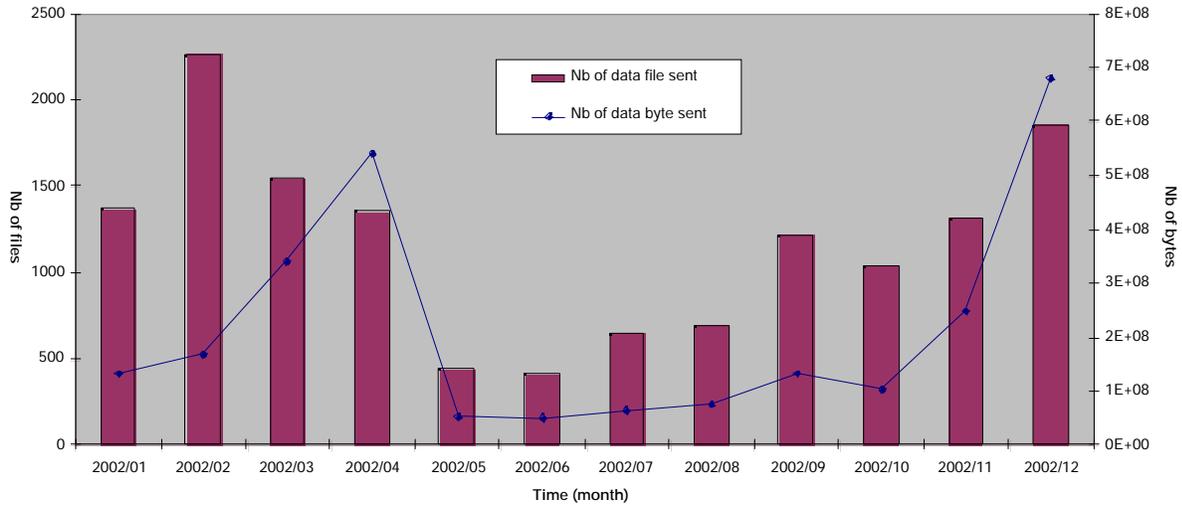
Figure 6.

### User observation download statistics from IGN GDC FTP server in 2002



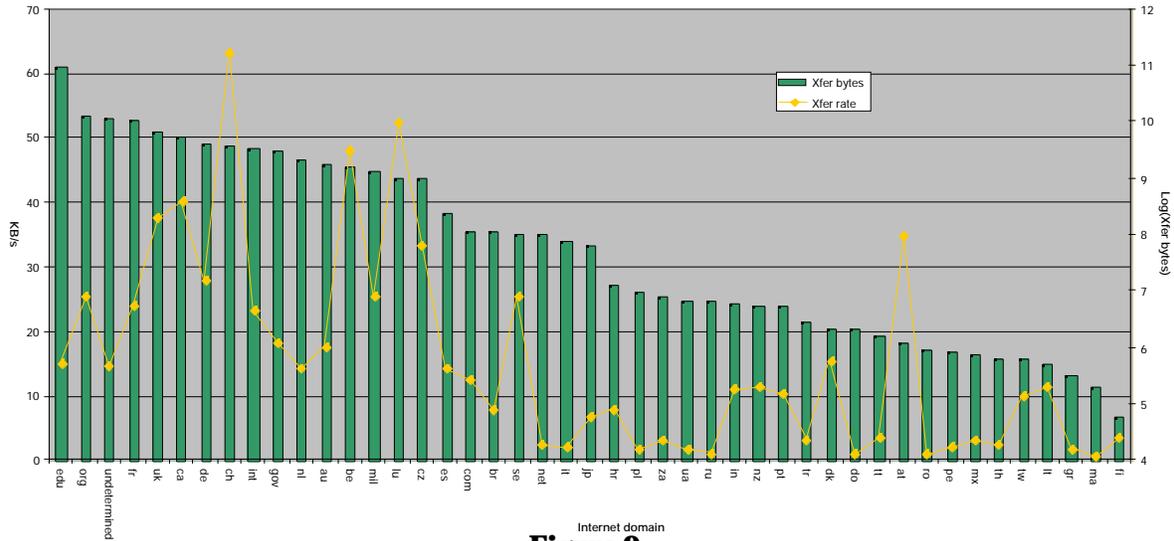
**Figure 7.**

### User products download statistics from IGN GDC FTP server in 2002



**Figure 8.**

## Outgoing bytes and average transfer rate from IGN GDC FTP server in 2002



Internet domain  
**Figure 9.**

**Today, 14:50 GUT**

**GLOBAL INFORMATION ON IGS NETWORK**  
 The IGS network comprises 320 stations  
 124 of them are global stations and 142 give NRT data  
 We are day of year 274 and GPS week 1238 (dow 3)

**IGS GLOBAL DATA SERVER STATISTICS**  
 The IGN archive holds 678534 files:  
 > Daily files :  
 175 station data files are expected each day  
 Last 4 days results :  
 D270 : 146      D271 : 146      D272 : 145      D273 : 131  
 > Near Real Time (NRT) files :  
 106 station data files are expected each hour  
 Last 4 hours results :  
 8-9H : 87      9-10H : 87      10-11H : 86      11-12H : 75

**Editorial**  
 By Edouard GAULUE on 2002-08-09.

**Welcome to new IGN's IGS web site**

Dear colleagues,

For a long time, you may have seen that IGN web site upon IGS hasn't changed a lot. That's surely why it was great time to give it a face-lift after the reinforcement of the FTP site as you may have already seen.

As well, today, I'm pleased to welcome you on the new IGN Global Data Center web site even if everything is not finished yet. In fact, quite more job is in preparation for the year to come.

Starting today, you can :

- make basic or advanced IGSMail research using the "IGSMail tool" link in the top banner, or the "Station Related" and "Others" links in the "IGSMail 24h/24" frame. You can also use the "Keyword search" tool that only works for IGSMail for the moment
- use the date tool to make any date translation or find a GPS calendar
- ask/answer to a question concerning data management in the Forum. It will help us to fill a FAQ list.
- find any basic information on IGS, in the 'IGS for beginners' section

**Figure 10.**