Report of the TIGGE Working Group on Archiving on Phase 1 implementation

1.0

Draft 1.1	Baudouin Raoult, Manuel Fuentes	17 October 2005
Draft 1.2	Steve Worley	29 October 2005
Draft 1.3	Don Middleton	2 November 2005
Draft 1.4	Baudouin Raoult, Manuel Fuentes	4 November 2005
Draft 1.5	Don Middleton and Steve Worley	7 November 2005
Version 1.0	Baudouin Raoult	8 November 2005

1 Background

The first workshop on TIGGE was held from 1 to 3 March 2005, at ECMWF. The workshop has been initiated by the WMO-THORPEX project. The purpose of this workshop was to collect the views of the community on what the TIGGE science aims should be, what the requirements are for use of the TIGGE data and hence what are the infrastructure requirements. The report document from the first TIGGE workshop is available on the WMO web site[1].

The report concludes that to achieve the TIGGE's key objectives, it is necessary to:

- Determine the user requirements for TIGGE data, including the types, volumes, format of data, access methods and timeliness
- Design the TIGGE infrastructure to meet these requirements
- Determine resource requirements and secure necessary funding if required
- Establish commitments from data contributors, TIGGE Archive Centres and prospective users
- Implement proposed infrastructure, collecting, archiving, and providing access to TIGGE data

[1] http://www.wmo.int/thorpex/pdf/tigge_first_workshop_report.pdf

- Develop and maintain close links with TIGGE users, including other THORPEX sub-programmes, field campaigns and Demonstration Projects, as well as other partners
- Have the flexibility to respond to evolving user needs as scientific understanding increases during the project

Furthermore, the report suggests that the TIGGE infrastructure is developed in two phases:

- Phase-1, during which data are collected in near-real time (via internet ftp) at a small number of central TIGGE data archives. This is to be implemented utilizing the existing infrastructure with little additional cost.
- Phase-2, during which data archives are distributed over a number of repositories, instead of all being held centrally, while efficient and transparent access for the users is maintained. This is a more flexible solution with the potential to eliminate routine transfers of large data volumes. But this will require substantial software development over a number of years, in coordination with the WMO Information System, and will require additional funding

A TIGGE Working Group on Archiving was established and was given the task to propose a technical implementation of Phase 1.

2 Introduction

The TIGGE Working Group on Archiving met at ECMWF on the 19th, 20th and 21st of September 2005. The attendees were TIAN Hao and LANG Hongliang from CMA, Yves Pelletier from CMC, representing NAEFS, Steven Worley and Don Middleton from NCAR and Baudouin Raoult, Manuel Fuentes and Laurent Gougeon from ECMWF. They were briefly joined by teleconference by Zoltan Toth from NCEP and Jordan Alpert from EMC during which the highlights of the meeting were given and Jordan gave a presentation on NOMADS.

The group was given a presentation of the TIGGE requirements as described in the report document from the first TIGGE workshop, followed by a presentation of NAEFS. Each partner presented the current infrastructures and applications relevant to the project.

Then the group reviewed the requirements and considered the various issues and solutions. This document presents a first set of recommendations along with an initial summary of issues that need to be resolved. The document is not comprehensive; in particular a more thorough description of user access needs is required. Nevertheless, it should be sufficient to start building the TIGGE database.

It is important to note that the implementation of the TIGGE phase 1 will be done with no extra resources, therefore the group only considered what is feasible within the existing infrastructure, in the desired time constraints, with the manpower available and with a minimum of developments.

3 TIGGE Phases

In phase 1, several Archive Centres, represented in Figure 1 below as the three grey boxes, will collect data from different Data Providers (the coloured arrow). Each

Archive Centre has a copy of all the data from all the Data Providers, which form the TIGGE database (represented by the multicoloured cylinders). A user can extract a selection of products across all Data Providers. The same selection can be requested from any Data Centre.

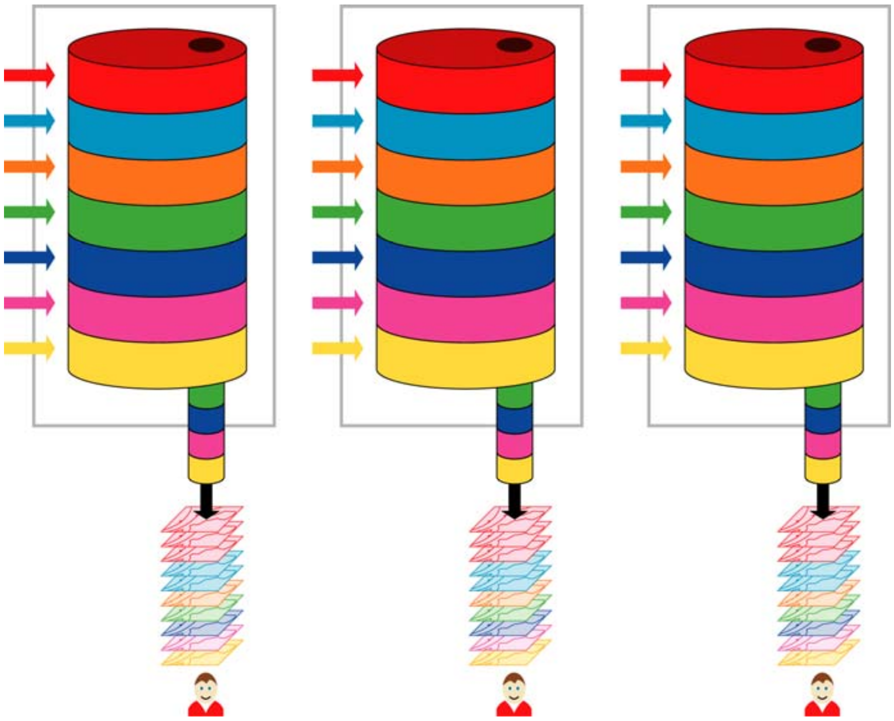


Figure 1: TIGGE Phase 1

In Phase 2, the TIGGE database will be distributed. A user will still be able to retrieve data from different Data Providers, but an interface will have to be developed to virtualise access to data.

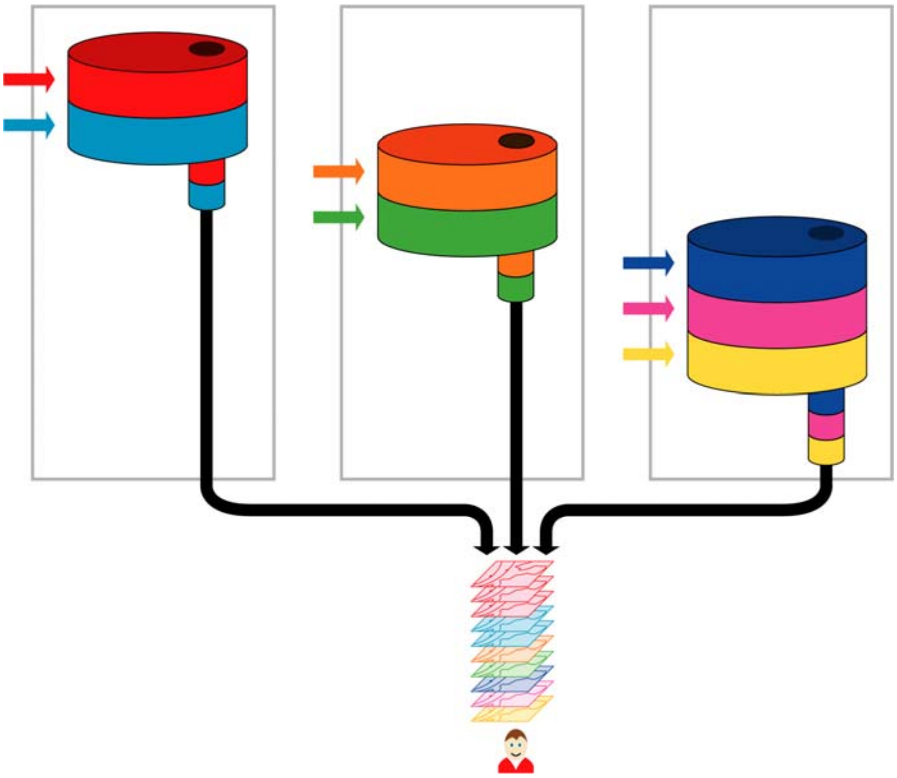


Figure 2: TIGGE Phase II

This document only addresses Phase 1.

4 Homogeneity of the TIGGE database

For this project to succeed, it is paramount that the content of the TIGGE database be as homogeneous as possible. This will insure a productive environment that has systematic data management and user access to data from many provider centres. The more consistent the archive the easier it will be to develop applications.

The multi-model seasonal forecast project DEMETER is a successful example where the effort put into creating a homogenous archive led to a variety of useful applications.

There are three facets to homogeneity:

4.1 Common terminology

All the partners must agree on a common way to reference data. Using common dictionaries all fields should be described with the same attributes (dates, level, step, parameter, etc.). This means a common set of metadata descriptors will be used across all data providing partners. This promotes the most rapid data processing, creating uniform reference catalogues in the access portals, and TIGGE-wide accurate search and discovery capability at the repository data access portals. This is developed in more detail in chapter 6.

4.2 Common data format

All partners must agree on a common data format to encode fields. They should also agree to use the same units. This is also developed in chapter 6.

4.3 Definition of a core dataset

A field is uniquely identified within the TIGGE dataset by the following tuple[2]:

(base date, base time, time step, origin centre, ensemble number, level, parameter)

When using fields to create a "grand ensemble", i.e. when considering all members from several origin centres as a super ensemble, we must make sure that they share the same values for the tuple *(base date, base time, time step, level, parameter,).*

As a result, a core dataset must be defined in terms of parameters, base times, time steps and levels. All Data Providers must adhere to the core dataset definition.

The proposal for the core dataset is given in Annex 1.

5 Data transfers

5.1 Network Bandwidth

It is thought that the available bandwidth between Europe and the USA is sufficient to meet the needs of TIGGE, whereas CMA raised concerns about the current bandwidth between China and Europe, as well as between China and the USA (the latter being probably better). The current bandwidth between China and the other partners currently appears to be much lower than projected TIGGE requirements. Nevertheless, the situation may improve by the end of the year, and the Working Group is interacting with various international networking groups to investigate some of the possible network options (e.g. CTSNET, GLORIAD). This is a potential risk for this project.

Bandwidth tests will be performed, firstly between Archive Centres, then between Data Providers and Archive Centres, in order to establish the best route for the data, as well as the optimum network settings (buffer sizes, TCP window size, …).

Tests, using ordinary FTP, between NCAR and ECMWF have established that TCP tuning (matching window/buffer size) improves data transfer rates. Specifics of this testing will be documented and made available for consideration at additional Archive Centres and Data Providers.

Some preliminary test results are given in Annex 2.

5.2 Data flow

Data flow refers to the transfer of model output in near real-time from Data Providers to Archive Centres. Depending on who initiates the transfer, data can flow in two modes: it can either be pulled by Archive Centre from Data Provider, or it can be pushed by Data Provider to the Archive Centre.

[2] To which we should add a version for testing.

The working group assessed several scenarios:

5.2.1 Many to one

In this scenario, each Archive Centre will have a “zone of responsibility”, either defined geographically, or by closeness according to network bandwidth. All Data Providers within this zone will send the data to this Archive Centre. The data will then be distributed between Archive Centres.

Pros:

- Each Data Provider sends its data only once.
- The Archive Centre is responsible to make sure each Data Provider within its zone delivers the data.

Cons:

- If an Archive Centre is down, a complete zone is isolated for the rest of partners.
- Synchronisation between Archive Centres may prove difficult.

5.2.2 Many to many

In this case, all Data Providers will send their data to all Archive Centres.

Pros:

- There is no single point of failure.
- Archive Centres do not have to constantly engage in data synchronization.

Cons:

- Requires more bandwidth from the Data Providers, as there will be 3 times more data flowing from each Data Provider.
- More interactions between Data Providers and Archive Centres.

5.2.3 Identifying data transfer mechanisms

Strategies for dealing with data distribution must address important requirements in a number of areas:

- Software and systems must be easily deployable to heterogeneous sites.
- Performance must be scalable to the levels required for TIGGE, and this may also require parallel transfers.
- It must be possible to monitor, audit, and trouble-shoot any problems that might arise.
- A variety of security/firewall profiles must be accommodated.
- It may need to be possible to configure a mix of fixed and dynamic routes, along with the ability to dynamically modify these configurations as needed.

NCAR noted the existence of IDD/LDM[3] (Internet Data Distribution system, Local Data Manager), an Internet based distribution system that may suit the requirements of TIGGE. This system is fundamentally designed for the distribution of real-time data, and is already running operationally in the UNIDATA community at roughly three hundred sites worldwide. It is thought to be scalable, and provides built-in mechanisms for monitoring and providing statistics on data transfer activity. IDD/LDM is also being employed operationally in the CONDUIT[4] project to provide distribution services for large-volume USA/NCEP forecast products. NCAR will investigate the suitability of IDD/LDM for TIGGE, in particular how it implements security, how it copes with firewalls, how rerouting and failover are dealt with and what level of resources (CPU, disks) are required.

In view of the outcome of this study, the group will also consider other options, including the AFD[5] (Automatic File Distributor) system, which is an ftp-based tool developed at DWD in Germany and in use at roughly one hundred institutions worldwide. Other ftp-based approaches are possible as well.

If these approaches prove impractical, the Working Group decided to implement data transfers using FTP in a “many to many” fashion; nevertheless, in light of experience and in particular due to bandwidth availability, a hybrid solution may prove more appropriate, in which some of the traffic flows between Data Providers and Archive Centres, and some flows between Archive Centres.

The Working Group debated the two primary modes of transfer, i.e. push and pull, and found that there were advantages and disadvantages with both approaches. Should it be determined that none of the existing available automated systems discussed earlier meet TIGGE requirements, this particular issue will be revisited and analyzed and a new strategy adopted.

6 Data formats

6.1 Data representation

The group studied GRIB, the WMO standard for encoding fields. While GRIB edition 1 is widely used in the meteorological community, it lacks homogeneity: each production centre uses different local extensions for metadata and different code tables for parameters.

The group found several opportunities in using GRIB edition 2 (GRIB2) for encoding model output:

- GRIB2 is the only WMO standard that supports ensemble data without the need of local extensions to represent the different members.
- The NAEFS community is committed to using it.
- TIGGE may be the ideal project to foster the use of GRIB2 by producing a great wealth of data in that format.

---

[3] IDD/LDM - http://www.unidata.ucar.edu/software/idd/

[4] CONDUIT - http://www.unidata.ucar.edu/data/conduit

[5] AFD – http://www.dwd.de/AFD

- There is little experience in using GRIB2, which also means that there is no proliferation of local tables or extensions yet.

It is clear that the more homogeneous the dataset (in data format and content), the more homogeneous the service can be. Therefore, the group proposes the use of GRIB2 as the data format for the TIGGE database. The working group will provide clear guidelines (best practices) on how all TIGGE fields should be coded in GRIB2. In particular, this group will:

- Identify the list of GRIB2 codes, tables and templates to use for each of the fields of the TIGGE database. The partners will agree on the usage of certain GRIB2 attributes that should be used to recognize unconditionally a field from the TIGGE project from any other fields, e.g. by adding to GRIB2 table 1.3 (production status of data) a new entry for THORPEX products.

- Decide which metadata must be stored in the GRIB2 section 2 (also known as local extension). This metadata will contain information required by the Archive Centres in order to classify the data, as well as information used by Data Providers to qualify the data.

- Propose that each field be encoded as a single GRIB2 message and not use the capability to encode more than one field in the same message. Multiple fields in a single message would unnecessarily complicate the user interface. Other proposals will be made, such as encode the control forecast as ensemble member 0.

- Make available examples of EPS model output from one or more centres in GRIB2. The group found this essential in order to ease the task of Data Providers. It will illustrate all the GRIB2 encoding for the message and the message layout within a TIGGE archive file. Non-negligible work is required to establish the GRIB2 codes and the effort needs to be coordinated with those already familiar with the format as well as WMO.

This working group identified several issues regarding data encoding and decoding:

- Data representation should be as common as possible. For example, ECMWF produces a mixture of data representations, namely spherical harmonics for upper-air fields and Gaussian grid (regular or quasi-regular) for surface fields. Other centres may produce regular latitude/longitude fields. It is proposed that data providers encode their data in a single representation, preference given to the simplest to handle. ECMWF may provide all its TIGGE data in regular Gaussian grid.

- Unnecessary format conversions should be avoided, as they may change data values. Thus, data providers shall make their data available in the archive format.

- Data from different Data Providers will be created at different resolutions. Data will have to be interpolated into a common grid before any multi-model comparison study. Different interpolation methods may lead to different results. The Archive Centres request that the Data Providers provide or insure suitable interpolation software is available for the users of their products. This can be available from the Archive Centres or through appropriate links to the Data Providers online information site. It is expected that the Data Providers will provide software support for the users.

It has to be noted that the choice of GRIB2 implies that existing data analysis and visualization tools (GrADS, NCL, IDL, Ferret, Metview, etc…) will have to be adapted to handle GRIB2. Thus, we might expect TIGGE to serve as a valuable catalyst to trigger this needed new development. On the other hand, the time required for the tool-building teams to engineer the new GRIB2 capabilities is unknown. This could lead to pressure to deliver the data in other formats such as GRIB1, which is not viewed as desirable for a variety of reasons, which were discussed earlier.

6.2 File structure

Efficient data ingest at the TIGGE Archive Centres and uniformly format data delivery to the users depends on adherence to an agreed upon file structure. Receipt of data in potentially a different file structure from each Data Provider would result in a data management problem at the Archive Centres. There are no Phase I resources that would help support large-scale data file manipulation and reorganization. This is especially true at NCAR where access is based on file-organized archives and is less significant at ECMWF, which handles all data using the MARS archive system. The CMA is still developing strategies for handling large data collections.

It is acknowledged that recommending a TIGGE file structure for the Data Providers will add development work at each centre. However, it is believed this is the only way to make the data management tractable at the repositories and will serve to benefit the users. Example data files will be developed by the repositories and distributed to the Data Providers.

6.2.1 Recommendations:

File sizes are intended to be optimum for network transfer and large enough for efficient mass storage operations.

For each run (base-date and base-time) there will be two files per forecast time step, one containing upper air fields (pressure levels) and one containing single level fields.

The following table provides a quick estimate of the order of magnitude of the daily production of ECMWF and NCEP.

	ECMWF	NCEP
Number of fields per file: Upper air	9 levels * 5 variables * 51 members = 2295 fields	9 levels * 5 variables * 11 members = 495 fields
Number of fields per file: Surface	19 variables * 51 members = 969 fields	19 variables * 11 members = 209 fields
Number of steps	Each forecast to 360 hours with output every 6 hours = 60 steps	Each forecast to 384 hours with output every 6 hours = 64 steps
Number of runs	2 forecast runs per day	4 forecast runs per day
Number of files per day per level	60 steps * 2 forecasts/day = 120 files	64 steps * 4 forecasts/day = 256 files
Daily number of	2295 fields (upper-air) * 120 files + 969	495 fields (upper-air) * 256 files + 209

fields	fields (single level) * 120 files = 391680 fields per day	fields (surface) * 256 files = 180 224 fields per day
Average field size (Gaussian grid)	T399 (d0-d7), T255 (d8-d15): 303720 bytes	T126 (d0-d7), T62 (d8-d16): 41160 bytes
Daily data to transfer	391680 fields * 303720 bytes/field = 110 GBytes	180 224 fields * 41160 bytes/field = 6.9 GBytes

The field hierarchy, highest to lowest order, in a file for both delivery and receipt is ensemble member number, level, and parameter. For example in the form of pseudo-language code the input or output loop is,

(outer loop) **ensemble member**

Level (for multi-level variables)

Parameter (inner loop)

Where the loops iterate the values in ascending orders.

Regarding filenames, the group propose to use the WMO file naming convention (a detailed description can be found on the WMO web site[6]):

`Z_TIGGE_C_`*`CCCC_yyyyMMddhhmmss_SSSS_LL_VVVV`*`.bin[.compression]`

where:

CCCC	Originator, the standard *CCCC* country code
yyyyMMdd	*yyyyMMdd*, is the base date of the run
hhmmss	*hhmmss*, is the base time of the run (UTC)
SSSS	*SSSS*, forecast time-step in hours (0006, 0012, 0240, 0360, …)
VVVV	*VVVV* is a version, numerical identifier (`0001` for operational)
LL	*LL*: type of level, `sl` (single level), `pl` (pressure level), `ml` (model level)
compression	An optional suffix referring to a standard public data compression algorithms (e.g. *.gz* for *gzip*) that has been applied to the data file.

[6] http://www.wmo.int/web/www/WDM/ET-IDM-3/Doc-4(1).doc)

Following is an illustration on how to name the file containing time-step 24 from NCEP for the run of 1st September 2005 at 12Z for their operational version, single level parameters, all ensemble members, including the control, in GRIB edition 2 format without compression:

```
Z_TIGGE_C_KWBC_20050901120000_0024_sl_0001.bin
```

Files will be exclusively GRIB messages. No record padding, or supplementary headers or trailers should be added.

Non-compliant files will not be accepted by the Archive Centres.

6.3 Metadata and naming conventions

In order to create a homogenous catalogue and promote user data discovery, common metadata information, i.e. a common way of describing data, is of the utmost importance. This metadata will be used to specify retrieval requests.

Most of the Data Providers have their own ways to name variables, use different abbreviation and GRIB code tables.

The TIGGE partners will have to agree on a common vocabulary of terms. This vocabulary should use existing names defined by WMO, as well as the work done by the group that created the Climate and Forecast (CF) metadata convention[7].

Data Providers may want to store site-specific information associated to their data, and may want to preserve it in the TIGGE catalogue. Also, it may be necessary to store other information in order to make possible the interoperability of TIGGE with other disciplines, like Geographical Information Systems (GIS).

Such other metadata will be stored in section 2 of the GRIB message. It is necessary to establish the structure of such metadata in order to enable the development of common software that can understand output from different sources.

The working group will develop the TIGGE metadata as efforts begin to establish the sample data files. In this process the technical contacts from the Data Providers will be informed and asked for opinion, discuss, and agreement. Data Providers will be requested to map locally used metadata into the TIGGE standard.

7 Organisation of the collaboration

TIGGE is a collaborative project with the focal point for data exchange at the Archive Centres during Phase I. The success of TIGGE is directly linked to the degree of commitment of the partners, and the ability of the partners and Centres to work together. From the Archive Centre's view the following are important aspects for successful operations.

As the Archive Centres will have a global view of the data production, it is proposed that they provide the project technical coordination and take on the responsibility of defining the necessary procedures.

[7] http://www.cgd.ucar.edu/cms/eaton/cf-metadata/

7.1 General organisation

Each partner will nominate two contact points:

- A technical contact point, which will be able to address operational and technical issues, such as troubleshooting, networking or timeliness of delivery.
- A scientific contact point, which will be able to address issues such as forecast performances or numerical errors.

When these contact points are unavailable the partners must nominate alternate contact persons.

The communication will be established through a series of mailing lists, collaborative tools (e.g., Wiki) and a web site, which ECMWF offers to host and will be mutually cross linked to the Archive Centres.

7.2 Quality assurance

7.2.1 Data integrity and backup

It takes only one bad piece of data to invalidate a large archive. Because large amounts of data will be moved across different media (memory, disks, network, tapes), data corruption will be unavoidable. The group recommends using checksums to assure data integrity.

There are two approaches:

- Each file is sent with its corresponding checksum.
- Each field has a checksum embedded in the local extension.

The first solution is easier to implement, whereas the second will provide more security as the checksum is computed at the very creation of the field and can be checked at any time in the future, when the field is in isolation. The group will investigate the possibility of adding checksum in the GRIB message, which will require the definition of a common algorithm to compute checksums that would be provided to the partners.

Although the First Workshop on TIGGE report states that each data producer is responsible for the backup of its own data, the group suggests that the Archive Centres backup one another. This implies that each Archive Centre should archive or be able to re-create a file (including checksum) in form and content as it was originally delivered by the Data Provider.

7.2.2 Issue of completeness

The objective is to have 100% complete data at the Archive Centres. In real world operational modelling with data collection at Archive Centres this may not be achieved for two reasons: the transfer of the data from the partner to the Archive Centre fails, and operational activities at the partner are interrupted and back filling past runs is impractical.

Sometimes the data will be produced but transmissions will fail. In this case the Archive Centres must make sure that they eventually receive the data. If the network outage is too long the Data Provider will need to hold a local copy until the outage is repaired and the data are copied. Data Providers should take this into account when planning and building their TIGGE infrastructure.

Sometimes a Data Provider will suffer a major failure with the consequence of data outage for a model cycle. In this case, as Data Providers work under tight real-time constraints, they may decide not to run this particular cycle. In this case, no data will be available for the TIGGE database.

Our goal is to minimize incompleteness, but recognize it will occur. Unfortunately, an incomplete dataset is often difficult to use. Most of the current tools used for ensemble data assume a fixed number of members from day to day. These tools will have to be rewritten. In addition, tools producing derived products will need to know when all data is available in order to compute such products.

The problem of incomplete data raises several questions:

- How do we define completeness and to what extent is it an issue? Is having data from only one provider sufficient for TIGGE?
- How do we document and track the difference between "data that has not arrived yet" vs. "data will never arrive"? In the first case, Archive Centres may still try to acquire the data.
- TIGGE has real-time and delayed mode objectives. When it comes to a real-time use of the TIGGE data, a "cut-off" time will have to be defined. This may be difficult as all Data Providers have their own operational schedules when running their models and may have different uploading capabilities to the Archive Centres.
- How do we make sure that all Data Providers deliver their data, especially after a major incident? Who is responsible for making sure the data eventually reach the Archive Centres?

Data Providers must endeavour to send missing data, whenever possible, to the Archive Centres, even if this means rerunning a forecast cycle.

7.2.3 Procedures to assure completeness

Many operating details that will help maximize completeness are yet to be worked out, but a few essential ones are:

- All files will have some type of checksum capability. A common tool to compute a checksum will be provided to the Data Providers.
- All files that do not adhere to the agreed upon format, content and structure will be rejected at the Archive Centres
- Data receipt inventories will be created, kept up to date, and used to document consistent delivery from the Data Providers.

7.3 Operations

7.3.1 Day to day operations

Tools must be built to monitor the data transfer within the system. Transfer statistics are required to quantify variations in the performance of the Internet.

Each Archive Centre will set up a web page showing volumes, date of data and date of reception for each Data Provider. This information will be used to cross-validate the content at the three archives.

When problems arise that prevent data delivery to the Archive Centres, the Data Provider will be responsible to notify all the Archive Centres, e.g. by sending an email to the appropriate TIGGE email list.

When an Archive Centre does not receive the expected data from a Data Provider, or if the data are incomplete or corrupted, it will first check with other Archive Centres and determine if the failure is an isolated case. If it is an isolated case recovery will be initiated between Archive Centres, if not the Data Provider must re-initiate the data delivery. In any case, the incidents must be investigated and documented. The use of a trouble ticketing system will be investigated to facilitate tracking problems.

7.3.2 Long term operations

The Archive Centres have agreed to define and collect common metrics that can be used to create combined TIGGE-wide reports. This information will be used for future evolutions of the system. Participation in TIGGE must not interfere with the operational activities of Data Providers, i.e. they should be able to upgrade models, introduce higher resolutions, and make all customary changes as needed. Mechanisms should exist that allow new products from the Providers to be easily integrated into the TIGGE Archive Centres. These procedures need to be established and will include ways to test delivery of new products and will likely require version number control, to name just a few features.

On the other hand, Data Providers must take into account their participation into TIGGE when planning changes to their forecasting systems, and must inform Archive Centres accordingly.

8 User access

8.1 Registration

The TIGGE dataset is available to the research community. A “registration authority” must be set up that should verify each registration request and establish if they are linked to a genuine research activity. Access to valid data is not permitted apart from specific field experiments.

“Valid data” is data for which the value of *base date* + *base time* + *time step* + 24 hours is greater than or equal to the current date, i.e. any data in the future.

8.2 Data retrieval

The group considered the possibility that each Archive Centre would provide an identical retrieval interface to the TIGGE database.

After noting that each centre are using very different technologies, have different user communities and different purpose, it was established that such a unified interface was not doable without significant development and would require extra resources.

However, the Archive Centre will guarantee that user interfaces will present the same information (e.g. same variable names), and that similar requests, although expressed differently, should return identical results.

ECMWF will utilise the MARS system initially, and NCAR will build upon its Research Data Archive and Community Data Portal environments in order to serve their respective user communities. CMA is still in the development process of their data delivery system. Over time and with additional project support, it is expected that there will be opportunities to further unify the user interface by leveraging developments from the WMO Information System (WIS) effort.

For multi-field requests, the retrieved fields will be delivered back to the user in the following order:

(outer loop) **Date**

Time

Step

Origin

Ensemble member

Level (for multi-level variables)

Parameter (inner loop)

This allows users to process fields from different origins as if they were members of the same ensemble.

Because users' requests will generate high data volumes, Archive Centres will prioritize and limit requests according to size. This will make sure that no user can monopolise the system by submitting an unreasonable request.

Very large requests may require delivery by tape media.

9 Risk assessment

The following risks have been identified

Risk	Description	Level	Mitigation
Use of the Internet	Building an operational system on the Internet may be difficult, as we have little control over it.	Low	Careful monitoring and continuous tuning may be necessary. Procedures must be in place to resend the data after outages, to guarantee completeness of the TIGGE database.
File structure	Imposing a file structure may put too much burden on Data Providers and may discourage them.	Low	ECMWF to build tools to convert/organise GRIB fields. These tools may be used by either Data Provider or Archive Centres.
IDD-LDM	This dissemination system may not work with firewalls, or may not be suitable for operations (traceability of problems, timeliness, …)	Medium	Ad-hoc transfer streams will be setup between Data Providers and Archive Centres using standard ftp.

Low bandwidth	As volumes exchanged within the TIGGE partners are very large, availability of network bandwidth is an issue. Preliminary tests between NCAR and ECMWF show that the bandwidth is sufficient, whereas the bandwidth between ECMWF and CMA is too low to transfer the required volumes. This issue may also exist with some of the Data Providers	Medium	Reduce the number of fields in the core dataset, for example by reducing the number of levels. If data cannot be received by CMA during the early phase of the project, delivery of data by tape will have to be considered.
GRIB2	There are currently very few tools able to handle GRIB2, in particular EPS data. Migrating existing tools to use GRIB2 may be too lengthy and may delay the implementation of TIGGE.	High	One solution will be to start the project with GRIB1, and define a TIGGE specific local extension that will be used by all partners. Unfortunately, this will means that the TIGGE database will contain both formats, which may be an issue for end users. In this case, we may decide to stick to GRIB1 through out Phase I of the project, or re-archive older data in GRIB2 once the tools are ready. Re-archiving will have significant impact (cost) on Providers or Archive Centres.

10 Implementation plan

The proposed implementation plan is as follows:

Evaluation of the data transport:

- Test transfer rates between Archive Centres: NCAR, ECMWF and CMA. Find best buffer sizes.
- Investigate other candidates for data transport: IDD/LDM, AFD, sftp
- Test transfer rates between Data Providers (e.g. NCEP) and Archive Centres: NCAR, ECMWF and CMA.

GRIB2 definitions:

- ECMWF will consult with NCEP and WMO in order to make sure we agree on the proper encoding of the fields in GRIB2.
- ECMWF will provide a sample model output to the Data Providers, according to the specification in Section 6, above.
- ECMWF will provide a series of example programs to create these files. It must be noted that these tools may have to be adapted by Data Providers in order to handle their own data and metadata mapping.

Establish archive management communications:

- Mailing lists, web sites and collaborative tools

- Collect list of contact points

Start feeding the TIGGE database with ECMWF, UKMO and NCEP model outputs and possibly other models base on voluntary participation from Data Providers.

- NCEP to send model output to NCAR, ECMWF and CMA[8].
- ECMWF to send ECMWF and UKMO model to NCAR and CMA.

11 Summary of decisions and recommendations

- Investigate IDD-LDM as a way to transfer data between Data Providers and Archive Centres.
- If IDD-LDM is not a suitable solution, data will be transferred between Archive Centres and Data Providers using FTP. A many-to-many data flow will be implemented whenever possible. Otherwise, the routes providing the best bandwidth will be used.
- All data must be accompanied with a checksum.
- Need the definition of a core dataset to ensure the homogeneity of the TIGGE database.
- Use GRIB2 as the common data format. Agree on variables codes and common terminology.
- The structure of the files to be transferred between Data Providers and Archive Centres will be in a TIGGE standard format.
- With the current resources, it is not possible to build a common user interface to the data. The Archive Centre interfaces will be built on existing technologies that are easy to use. The resulting products will be in the TIGGE standard format.
- Examples of properly encoded model output will be provided, together with sample programs
- Website and mailing lists will be created
- Archive Centres will act as technical coordinators.
- Data Providers must endeavour to send missing data, whenever possible, to the Archive Centres, even if this means rerunning a forecast cycle.
- Data providers will provide scientific and technical contact points.
- Data providers will provide software to support their products

[8] Data will be sent to CMA if bandwidth permits. If not, as soon as CMA has the necessary bandwidth with ECMWF or NCAR, we will try to send the content of the TIGGE database, either using the Internet as a background activity, or by sending magnetic tapes.

12 Conclusion

The success of this project will depend greatly on:

- The commitment of each partners
- The establishment of a collaboration methods
- The availability of sufficient network bandwidth
- The homogeneity of the catalogue that is built from standard metadata
- Adherence to a standard GRIB message layout in the data files

The work proposed here will have to be reviewed in the light of experience. The first step will be to establish the collection of data, then have some test users to start using the TIGGE database.

Annex 1. Proposed list of parameters

This is the list given in the report pf the first TIGGE workshop.

1 Single level fields:

It is proposed to archive 19 (+2) single level fields.

Parameter	Abbreviation	Level	Unit	Output frequency	Comment
Mean sea level pressure	MSL	MSL	Pa	6h	inst[9]
Surface Pressure	SP	surface	Pa	6h	inst
10m U-velocity	10U	10m	m s**-1	6h	inst
10m V-velocity	10V	10m	m s**-1	6h	inst
2m temperature	2T	2m	K	6h	inst
2m dew point temperature	2D	2m	K	6h	inst
2m max temperature	MX2T	2m	K	6h	det_lo
2m min temperature	MN2T	2m	K	6h	det_lo[10]
Total precipitation (liquid + frozen)	TP	surface	m	6h	acc_st[11]
Snow fall	SF	surface	m of water equivalent	6h	acc_st
Snow depth	SD	surface	m of water equivalent	6h	inst
Total cloud cover	TCC	surface	0-1	6h	inst
Total column water	TCW	surface	kg m**-2	6h	inst
Surface latent heat flux	SLHF	surface	W m**-2 s	6h	acc_st
Surface sensible heat flux	SSHF	surface	W m**-2 s	6h	acc_st
Surface solar radiation	SSR	surface	W m**-2 s	6h	acc_st
Surface thermal radiation	STR	surface	W m**-2 s	6h	acc_st
Sunshine duration	SUND	surface	s	6h	acc_st
Convective available potential energy	CAPE	surface	J kg**-1	6h	inst
Orography (Geopotential at the surface)	GH	surface	m**2s**-2	t_0 (and possibly $t_{resolution\text{-}change}$)	inst
Land-sea mask	LSM	surface	0-1	t_0 (and possibly $t_{resolution\text{-}change}$)	inst

Note: Orography and Land Sea Mask will be archived with the control forecast, unless these parameters are perturbed.

[9] Instantaneous output.

[10] Determined over period from last output time to current output time.

[11] Accumulated over period from start of forecast to current output time (or alternatively accumulated from last output time to current output time; to be decided).

2 Upper air data:

It is proposed to archive 5 parameters on 9 pressure levels, i.e. 45 fields.

The 9 levels are 1000, 925, 850, 700, 600, 500, 300, 250 and 200 hPa.

Parameter	**Abbreviation**	**Level**	**Unit**	**Output frequency**	**Comments**
Temperature	T	L9	K	6h	inst
Geopotential	G	L9	m**2 s**-2	6h	inst
U-velocity	U	L9	m s**-1	6h	inst
V-velocity	V	L9	m s**-1	6h	inst
Specific Humidity	Q	L9	kg kg**-1	6h	inst

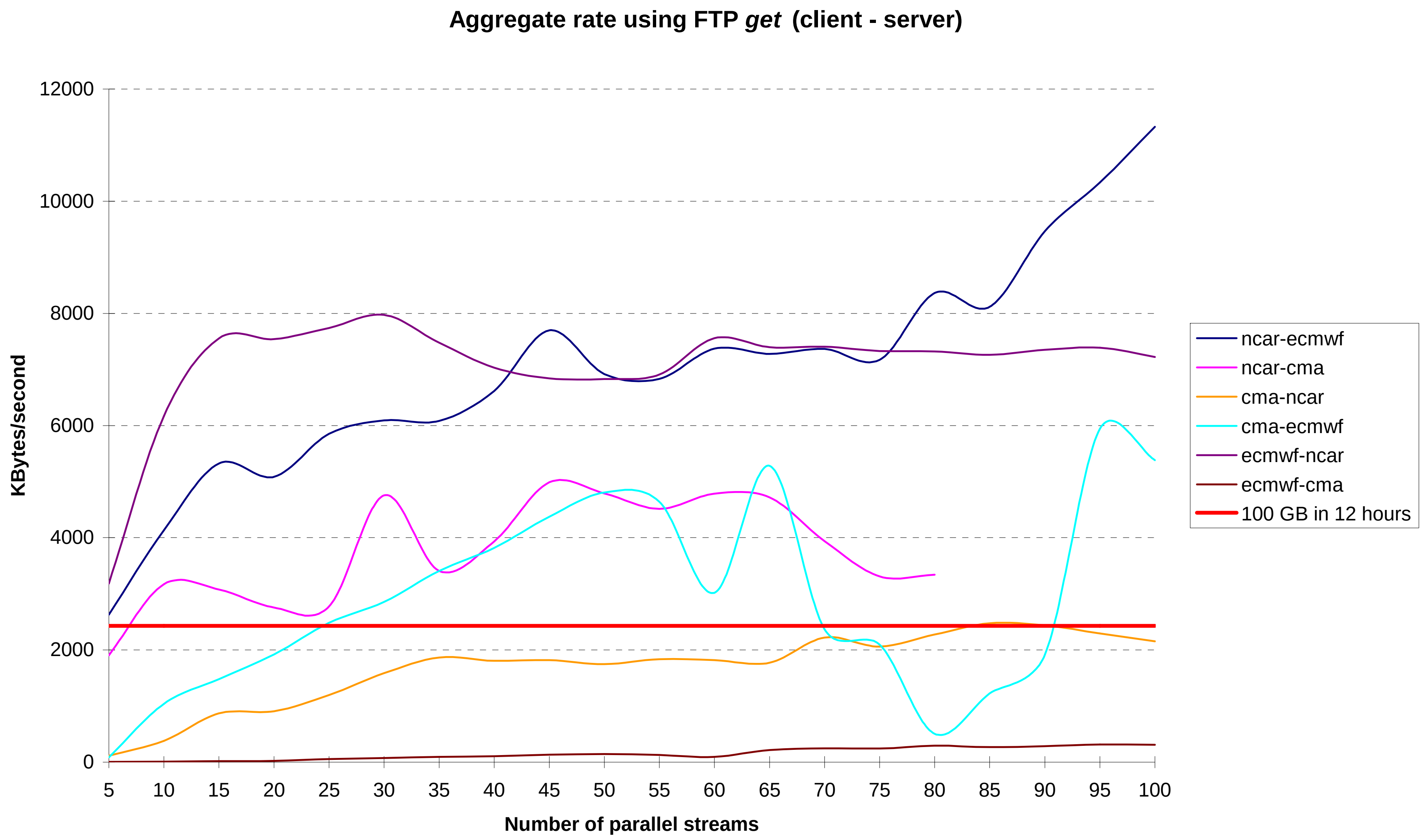
Annex 2. Initial transfer test results

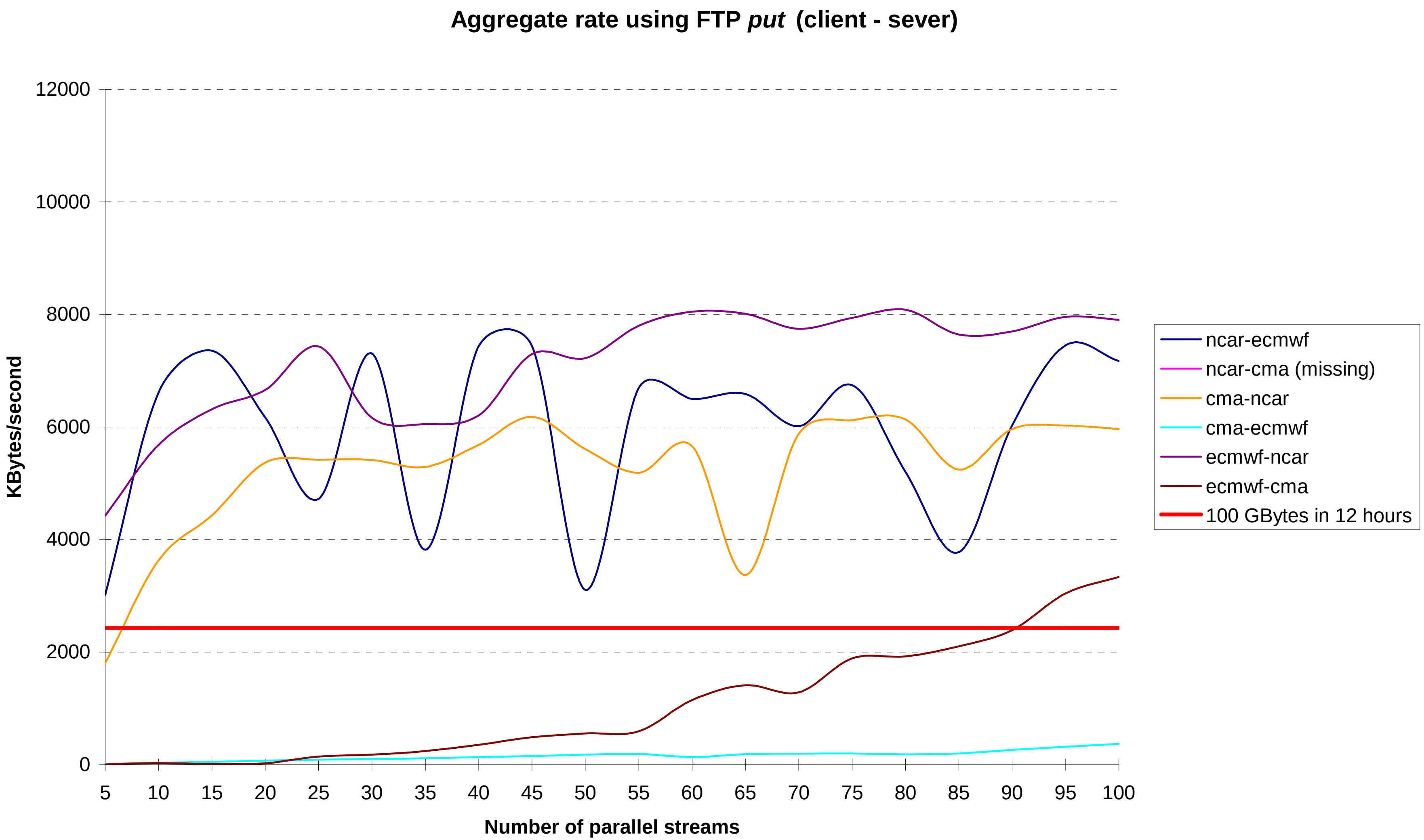
The following plots show aggregate transfer rates between CMA, NCAR and ECMWF. Every possible transfer has been tested.

The 100 GB/12 hours threshold shows a rate we should target. This is based on the fact that ECMWF model output is expected to be in the order of 100 GB, and that it should be sent into a maximum of 12 hours, giving us the chance to retransmit it once on a given day.

According to the graphs, the 100 GB per 12 hours goals can achieve, but a lot of testing, tuning and tweaking will be necessary. Performances depend on who initiate the transfer (get or put), buffer sizes, TCP window size….

We will have to regularly run these tests to make sure we are using the best possible options.


Aggregate rate using FTP *get* (client - server)
KBytes/second
Number of parallel streams
0
2000
4000
6000
8000
10000
12000
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100
ncar-ecmwf
ncar-cma
cma-ncar
cma-ecmwf
ecmwf-ncar
ecmwf-cma
100 GB in 12 hours



Aggregate rate using FTP *put* (client - sever)
KBytes/second
12000
10000
8000
6000
4000
2000
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100
Number of parallel streams
ncar-ecmwf
ncar-cma (missing)
cma-ncar
cma-ecmwf
ecmwf-ncar
ecmwf-cma
100 GBytes in 12 hours