

EARN Document

Title: Network usage and server compensation  
Author: M. Hebgen  
Date: 1989/11/21  
Committee: Executive  
Document: EXEC152 89 EARN-MIN LISTSERV@UKACRL  
Revision number: 3  
Supersedes:  
Status: Proposal  
Maintainer: M. Hebgen  
Access: Unrestricted

-----  
--

## 1. HISTORY OF EARN FINANCING

---

EARN was initiated in 1984 by IBM which supported EARN financially up to the end of 1987. Since 1988 EARN is self funded by its country members with support from DEC for the OSI transition. The distribution of EARN central costs was/is based in

- o 1988 on EARN/RARE keys,
- o 1989 on the mean values of EARN/RARE keys and GNP values and
- o 1990 on EARN/RARE keys for volume independent and on GNP values for volume dependent budget items,

where GNP is the Gross National Product of a country and EARN/RARE keys are smoothed GNP values - RARE keys for RARE countries and similar EARN keys for non-RARE countries.

## 2. REQUIREMENT

---

The EARN Board of Directors approved at their meeting in June 89 the "Principles for future distribution of EARN costs" and the "EARN statis-

tics directive" - see documents BOD28 89 and BOD33 89 in the appendix.

One of the principles is

Volume dependent costs are based on usage

---

Work should be initiated to extract reliable figures independent of seasonal fluctuations from the accounting statistical data, e.g. using data from the first six months in a year for the budget presented in the fall for the next year. Traffic from file servers should be handled correctly, so that the country operating a server does not pay for traffic initiated by other countries.

Thus this paper will deal with the distribution of EARN central costs based on the usage of the member countries in the past. There is no intention within EARN to introduce volume tariffs or similar things, but only to get a fair distribution acceptable by the majority of the EARN Board of Directors.

### 3. AVAILABLE DATA

---

#### 3.1 Global country data

---

The accounting statistical data as described in BOD33 89 show 3 types of international traffic figures per country on a monthly basis:

- o Data sent to other countries
- o Data received from other countries
- o Data sent to and received from other countries

Today 2 programmes to calculate country statistics derived from standard RSCS accounting records are available and used for the

international  
traffic figures:

- o CNTYACCT written by Udo Meyer
- o CTRYSTAT written by Jose Maria Blasco and maintained by Manfred Bogen  
at DEARN

The programm CTRYSTAT shows extra statistics for LISTSERV, NETSERV and MAILER and can be enhanced to calculate statistics for other servers. These server statistics are based on all accounting records with the name of the server either in the sending or in the receiving address.

### 3.2 Server compensation data

---

Many countries provide additional service to EARN by running NETSERV, LISTSERV and other file servers like TRICKLE, MACSERV or ASTRA. This provision of service causes additional resources in the operating country which should not get the penalty paying for traffic initiated by other countries.

Within EARN we can differentiate between 2 major types of servers

- o Mail servers and
- o File servers

and between 3 types of traffic related to servers

- o Requests sent to servers,
- o Traffic between servers and
- o Results sent back to the requestor.

#### 3.2.1 Mail servers

---

Mail servers are the MAILER and the mail distribution function of LISTSERV. A typical LISTSERV list communication consists of the follow-

ing steps:

- o A user sends mail to the LISTSERV managed list, implemented as a "dummy" user on the VM system.
- o LISTSERV takes the reader files of the list and processes them.
- o LISTSERV then sends the mail result
  - \* for peered lists to other LISTSERVs for further processing,
  - \* for non-peered lists to another LISTSERV if there is one closer to the destination node and
    - \* in all other cases to the MAILER doing the transportation.

Thus the accounting data do NOT indicate LISTSERV, the destination address for sending is the list and the source address for distribution is the MAILER. Furthermore on most IBM VM systems the MAILER is not only used by LISTSERV but also by all users especially in combination with the MAIL/MAILBOOK package.

Neither LISTSERV nor the MAILER today provide additional accounting data to identify the "real" source. For LISTSERV lists it is possible to get a LISTSERV statistic per list and to take this into account - but because statistics can be disabled this is not further considered in this paper.

### 3.2.2 File servers

---

Examples of file servers are NETSERV, MACSERV, KERMSERV, ASTRA, TRICKLE and the file distribution part of LISTSERV. A typical communication with a file server consists of the 2 steps

- o Send a mail or interactive message to the file server
- o which then in turn sends the requested files back.

LISTSERV and TRICKLE may send the requested file as mail to users if they have requested it explicitly or if they are reachable via a

mail gateway. In this case the MAILER is used for transportation.

Another possibility is that a user has registered himself on LISTSERV, NETSERV or TRICKLE as interested in new versions of a file or a file package. When the file will be changed by the owner the registered users will get automatically either a file update information or the file itself.

#### 4. TECHNICAL IMPLICATIONS

---

Under some known circumstances the accounting statistical data or global country data are not valid, e.g. in case of

- o re-routing, i.e. sending the data a second time on a different route,
  - to solve intermediate network problems,
- o files which have been received but may never be sent because of their
  - filesize,
- o files which have been received but may never be sent because of sys-tem failure (cold start etc.),
- o traffic from unknown nodes because of overlapped updating of routing tables,
- o data from systems emulating the IBM NJE protocol only partially,
  - especially when the number of records indicated in the header is not equal to the real number of records sent,
- o error loops between servers mainly caused by unsynchronised updates

Furthermore part of the traffic is not included or can not be related to a country like

- o interactive messages are not counted at all,
- o traffic to and from gateways is counted to the country operating the

gateway, there is no information to describe the "real" originating or receiving country.

See also the initial draft of an EARN technical document in the appendix.

Nevertheless it is assumed for the purpose of this paper that errors in the accounting statistical data are fairly distributed and that these data are reasonable in the overall view and can be taken as a basis for distribution of the costs by usage.

## 5. RELIABLE FIGURES

---

According to BOD28 89 reliable figures should be extracted independent of seasonal fluctuations from the accounting statistical data.

A higher independency is of course caused by a longer time period of data collection. Taking into account that the EARN BOD meeting in the fall of a year decides on the budget for the next year, the latest month in the year can be June for data collection and preparation of the budget figures.

For the first year of using usage data in the budget at least 6 month (from January to June) are required. For all following years 12 month of data starting in July up to June of the then current year should be used for the calculation.

## 6. ALTERNATIVES

---

## 6.1 Global country data

---

For the global country data only the data sent to other countries should be used to describe the "usage" of EARN by a country because they are caused by people in the sending country. The data received from other countries can not be used for calculating the usage of EARN by a country because the country can not influence it.

## 6.2 Use server accounting data

---

For the server compensation data (to correct the country's global value) one attempt can be to use as much as possible server accounting data. The guiding principle for this approach is to count the traffic to the initiating country wherever possible.

This implies technical changes to most of the servers and to the gateways to provide user accounting records which indicate the "real" source or destination. Furthermore for LISTSERV discussions a relationship of the answer to the question is needed, but this is not possible at all.

The compensation for LISTSERV traffic will result in charging the "sender" double - once for sending the posting and once for each receipt (or interListserv copy sent). But if we charge the "sender" twice, most lists will turn into question forums with no people to answer! Why should someone post an answer to a list, where he is doing the "questioner" a favor by helping him out, when he will be billed also for the outgoing copies?

The problem with the compensation for File server traffic will

be encouraging users NOT to use network file servers like NETSERV and KERMSERV since they will incur a double bill. They will be billed for sending the request to the server and will be billed for receiving the file in order to compensate the network server. This will encourage each country to create a mirror copy of all data available in the network and to keep it available on a national level. This will also cause people to set up alternate file servers in BITNET with funny names like GLUMP or FIZZ. Since BITNET doesn't have volume accounting, every European user will go to the FIZZ server in the USA and therefore will only be billed once and not double.

The impact of such an approach would be counterproductive to the attractiveness of EARN. With this method EARN will be killing off the cooperation that has been fostered by the lists and by the file servers. No one will want to use an EARN file servers and people will not bother to answer people posing questions in lists.

### 6.3 Not to use server accounting data

---

Another alternative for the server compensation data (to correct the country's global value) can be to minimize the use of server accounting data. All RSCS accounting records that contain server userids (i.e. LISTSERV, NETSERV, etc.) irregardless if the data is being sent or received are NOT to be used in the calculation of a country's global value.

This way, countries that have many servers in their country will not

be penalized for sending out data to users and EARN in addition eliminates the entire problem "how to account for Listserv traffic". What is left, are just pure records going from user to user or user to gateway and among those records we will only use the "sent" records and not received records.

The one qualification is that the traffic can be related to the userid LISTSERV or TRICKLE, and not to the MAILER, which seems to be technically possible by

1. either running a national LISTSERV in each EARN country

Countries that do not have at least one LISTSERV will probably have the accounting records show the data as travelling from MAILER to MAILER. As of today Luxemburg, Yugoslavia, Egypt and Sweden do not have a national LISTSERV.

2. or to run 2 MAILERS, one for personal mail and one for mail caused by servers.

Thus the outgoing traffic caused by servers like LISTSERV or TRICKLE will have the unique name of the mailer for servers in the accounting records. The major difference to the national LISTSERV approach is that the "last country in the queue" running LISTSERV has to run 2 MAILERS for a country not having a national LISTSERV.

The guiding principle for this approach is to declare the traffic caused by servers as an EARN infrastructure and not to charge the "sender". A separate evaluation of the server traffic as available with the CTRYSTAT program is desirable to see how much load is caused by servers

compared  
to individual traffic.

From a technical standpoint this will be the easiest and perhaps the solution that EARN will be able to implement, the only remaining task is to establish an agreed list of Names of Servers within EARN.

## 7. POLITICAL IMPLICATIONS

---

Following is a summary of comments I have received during the EARN Board of Directors meeting in fall 1989 which are of a more general nature and should be considered in a principle discussion on distribution of EARN central costs.

A "usage" based distribution is dangerous because

- o there is a risk of major changes in the country contribution of 50 percent or more - major changes of this size have already taken place in the past, but always on the same few countries,
- o it is a "floating" base and every change has an impact on all other countries,
- o it might cause countries to install "private" international lines to reduce the traffic on "EARN" lines with the risk of driving traffic away from EARN or losing member countries,
- o EARN runs into competition with cheaper sponsored alternatives like EASINET or IXI,
- o it tends to minimize the use of EARN lines where a distribution model based on fixed numbers like EARN/RARE keys tends to maximize the usage.

## 8. PROPOSAL

---

Out of the listing of available data and the discussion above on alternatives it is proposed for calculating a country's usage of EARN

- +-----  
-+
1. To use the data sent to other countries as the global country data.
  2. To correct this global value by server compensation values for servers like LISTSERV, NETSERV, MACSERV, KERMSERV, ASTRA, TRICKLE etc.
    - 2.1. by not using all server RSCS accounting records except for the MAILER when calculating the global country data.
    - 2.2. by directing all EARN countries to operate a national LISTSERV - if this is technically not possible the neighbouring countries are requested to run 2 MAILERS, one for personal mail and one for mail caused by servers.
  3. To use the month June as a deadline for data collection for the next years budget calculation.

4. To use at least 6 month (January to June) data collection in the first and 12 month (July to June) in all following years.

---

## 9. ACKNOWLEDGEMENTS

Thanks to the EARN Network Operation Group (NOG) - especially Manfred Bogen, Dominique Dumas, Turgut Kalfaoglu, Hank Nussbacher and Berthold Pasch - for providing the technical input and helping me in the discussion of possible alternatives.

---

## APPENDIX A: BOD28 89

BOD28 89  
revision of BOD27 89

EARN BOARD OF DIRECTORS

Principles for future distribution of EARN costs

issued by  
F Greisen  
June 6, 1989

---

1 Volume dependent costs are based on usage

Work should be initiated to extract reliable figures independent of seasonal fluctuations from the accounting statistical data, e.g. using data from the first six months in a year for the budget presented in the fall for the next year. Traffic from file servers should be handled correctly, so that the country operating a server does not pay for traffic initiated by other countries.

Until agreed reliable statistics can be produced, GNP ratios are used. Preliminary studies show a reasonable correlation.

The budget items in the volume dependent category are 3 "Staff", 5 "Inter-continental lines", 6 "Development", and 7 "Contingency fund".

2 Volume independent costs are based on RARE keys

For non-RARE countries the key for a RARE-country is the same GNP class is used. These items are 1 "President's office", 2 "EARN office", and 4 "Other expenses".

3 International links are funded by countries

Each country still pays its connection to the network.

4 Countries on the EARN X.25 with EARN funded lines contribute to EARN central funds

Countries connected to the network through EARN paid lines contribute the equivalent of the cost of a line which would otherwise be chosen to central funds.

5 Co-operating countries contribute to the items they use and influence

The contribution to 5 "Inter-continental lines" is according to usage (or GNP) and the contribution to 1 "President's office" and 2 "EARN office" is according to (extended) RARE keys.

APPENDIX B: BOD33 89

---

BOD33 89  
Revised BOD26 89  
and EXEC71 89

EARN BOARD OF DIRECTORS

EARN statistics "directive"

Approved at the Board of Directors Meeting June 1/2, 1989  
and revised under delegated powers by the Executive.

issued by  
A Auroux  
June 1, 1989

---

## 1 Requirement

EARN now collects traffic figures on international links. These are required to:

- show the loading on lines with a view to indicating:
  - desirable topology changes
  - desirable line upgrading
  - desirable relocation of servers.
  
- distribute part of the EARN costs from financial year 1991 onwards.

To achieve these requirements traffic figures must be collected by each international node. To this end this EARN "directive" is being issued.

The definition of an EARN directive can be found in EXEC34 89. It is mandatory for relevant nodes to implement directives.

## 2 Directive

Each international node is required to collect traffic data on EARN international links for traffic to and from every other country. Data is collected on a calendar monthly basis and sent to the "traffic data" co-ordinator before the 10th of the following month.

The format of the data and the destination address for the data are determined by the Network Operations Group.

Suitable code for collecting and submitting data exists or is being developed.

This directive must be implemented by September 1, 1989 in order to have

complete traffic data starting in September 1989.

### 3 Dispensation

International nodes not using IBM VM or MVS will only be required to submit figures when suitable software has been developed.

### 4 Technical annex (not part of the directive)

The file which contains the traffic data must currently be sent to Dominique Dumas (BRUCH@FRMOP11) with the name "ccyymm DATA" (cc=Country Code, yy=year, mm=month).

The file may be produced by:

- the Udo Mayer program
- the Jose-Maria program
- any other program producing the same output.

Dominique Dumas should be contacted for details of the format required for submitted data.

### 4 Results (not part of the directive)

The results of the analysis of the data are stored on LISTSERV@DEARN with names "STATyyymm DATA". Other types of analysis will be developed as required.

## APPENDIX C: INITIAL DRAFT OF AN EARN TECHNICAL DOCUMENT

---

From: Peter Sylvester +49 228 8199645 <GRZ027@DBNGMD21>  
Subject: Statistics considered harm/use+ful/less

During the NOG meeting last week I offered to create an initial list with problems of the EARN statistics. The list will contain some effects that I have seen in our environment. Completeness is not my goal. None of the problems is new.

EARN technical document number nnnnn status: initial draft

The statistics collection in EARN is based on RSCS accounting records. For RSCS these data are created by the native IBM code. For JES2 system a modification and exit is available to create creates SMF records with RSCS accounting records as data.

The main usage of the data is to calculate the international traffic. A matrix of country-to-country traffic and line utilization data are derived. The data are collected at international nodes. At least two programs are used for the statistics calculation.

There are several sources of problems that can lead to wrong raw data, or wrong statistic maps. Some of these reasons cannot be easily circumvented.

- The data do not contain an identification of the line that is used. In order to derive the line that has been used the origin and target node names and a topology information must be used to find the line. This process can be incorrect in some cases:

When there are two international nodes in one country, the traffic is collected on both nodes, and when there is backup traffic between these nodes data may be counted twice.

When the topology changes and/or when backup international paths are used the situation is similar.

For MVS batch job output origination and target are switched. In general the origin node of the file may sometimes be the origin node of the job, and this is normally identical to the receiving node of the output. This means that the origin/target node show a national traffic where in fact there may be international traffic.

It is necessary that the point of entry and exit into and from the international backbone is clearly identified. This cannot be safely done without having the name of the link in accounting data (except by restricting the topology in sometimes undesirable ways). Whether this is a sufficient mean is for futher study.

- The number of records that are shown in the RSCS accounting records are not correct.  
The field in the RSCS/NJE headers that is used to record

the number of data records in the file being transmitted is supposed to be created at the originating system. Some NJE emulations create invalid data in these fields.

We have seen that some older version of the ANJE package always set a value of 256 into this field.

RJE output that has been originated on some Cray station seems to show only half of the actual number of records. This was seen on a JES2 node with the \$QNET package installed. \$QNET counts the actual number of records and shows both the expected records and actual records.

Recently we have seen negative data. The reason seems to be that system (either the originating or some intermediate one) uses an invalid offset into the NJE records.

The assumption that the originating system sets correct fields can lead to arbitrary results. Any modification of the data during the transit is possible.

- The number of records is not a good measure for the number of bytes that are transmitted. The number of bytes per records and the way how records are compressed is not clear. The current assumption is that a records contains 80 bytes. It is not known whether this assumption is fair or not. If the data pattern used in all countries is very similar (and this is expected for electronic mail) than the error is failrly distributed but the absolute number of bytes that are transferred may differ from the data that are recorded in the accounting data.
- Interactive messages are not recorded.

Please feel free to send me additional items or clarifications.

Mentioning problems in other systems does not necessarily mean that these systems need to be banned from the network because of protocol violations or else.