INES DISTRIBUTION SYSTEM: STRUCTURE AND REQUIREMENTS

J.D. Ponz
ESA-D/TOS, Villafranca, Apartado 50727, 28080 Madrid, Spain

1 Introduction

The IUE project at Villafranca has dedicated a special effort to the distribution of the observations. Along the history of the project several versions of the archive have been generated:

- Uniform low dispersion archive (ULDA/USSP). Distribution of low resolution data via wide area networks. Access to the archive is done with a special purpose software using a catalogue subset.
- Final Archive product distribution with high quality catalogue and uniform reduction system (NEWSIPS). Standard file formats (FITS).

The new distribution system, INES, is the continuation of the two previous activities, to distribute IUE data in a very efficient mode. The purpose of the project is to reach the maximum number of scientists, with a very simple distribution scheme, and to provide spectra in a form that does not require detailed knowledge on the instruments.

The main concepts come from the lessons learn with ULDA/USSP. Its implementation uses new technology based on the WWW architecture, with minimal requirements on the user node.

The system is designed according with the client/server model, in which the client queries the archive and requests data, and the server provides the information. The Principal center and National hosts are servers, while the client is at the end user node. Servers communicate with clients using the http protocol on the Internet.

2 On-line Information Items

The on-line information items have been defined to simplify the format and reduce the size of the distributed set. The new spectra contain the essential information to be used for scientific purposes, so that instrument details – for instance, the spectral order number – are not required.

Low-dispersion spectra have been re-extracted from the master archive to improve on NEWSIPS algorithms for certain cases, like emission line objects or spectra with low exposure levels. To this aim, the extraction algorithm, uses a correct noise model extrapolation, and a new background estimator has been implemented. Other improvements include estimation and removal of solar contamination in long-wavelength spectra, and uniform sampling domain for LWP and LWR cameras.

In addition, high-dispersion spectra are sampled at low-dispersion steps, including order concatenation and new error estimation. Finally, high-dispersion spectra have been concatenated.

The complete on-line archive set consists of the following information items:

Access Catalogue Contains the information required to query the archive and evaluate the quality of the spectra. Entries in this catalogue are defined by **camera**, **image**, **dispersion** and **aperture**. The number of entries is estimated in 110,000. This catalogue is generated from the Master Archive Catalogue.

Publications Catalogue Contains reference to publications as defined in NASA Abstract Data System (ADS). Entries in this catalogue are defined by **camera** and **image**. The estimated number of entries is 45,800.

Low-dispersion 1-D spectra (`ARC_LOW1D`) Contains the flux, associated error and quality flags for low-dispersion spectra, and high-dispersion spectra converted to low-dispersion sampling domain.

High-dispersion 1-D spectra (`ARC_HIGH1D`) Contains flux, associated error and quality flags for high-dispersion spectra, with orders concatenated.

Low-dispersion 2-D spectra (`ARC_LOW2D`) Contains the line by line spectra for low-dispersion.

Documentation Project documents describing the Final Archive (IUEFA) and the Newly Extracted Spectra (INES).

Filenames for 1-D extracted spectra are defined as

`<camera><image><dispersion><aperture>.FITS`

where

`<camera>` is the three letter camera name (`SWP`, `LWP`, `LWR`).

`<image>` is a 5 digit image number,

`<dispersion>` single letter identifying the dispersion as `L` low, `H` high and `R` high converted to low-dispersion.

`<aperture>` single letter identifying the (extracted) aperture as `L` large, `S` small.

Archive data files are stored in a directory hierarchy defined by camera and image numbers in groups of thousand. Estimated size for each archive data set is given below.

	File size (Kbytes/sp)	Total size (Gbytes)
ARC_LOW1D (low)	17.3	1.2
ARC_LOW1D (high)	17.3	0.6
ARC_LOW2D	242.0	17.0
ARC_HIGH1D	340.0	12.5

3 Architecture Overview

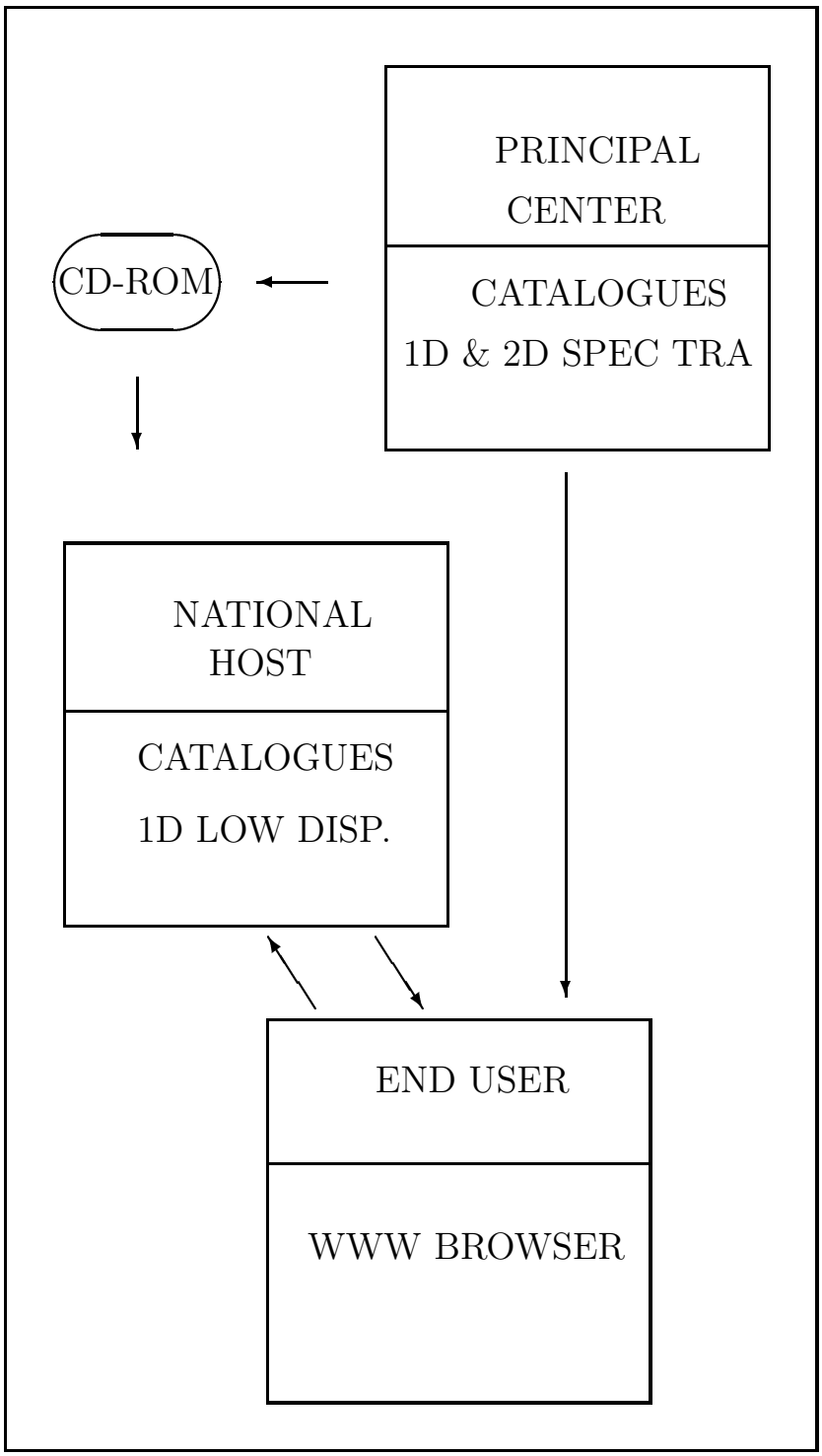
The distribution system is structured in three levels corresponding to a single **Principal Center**, several **National Hosts** and unlimited **End User Nodes**. Each level has the following information items:

Principal Center Contains the complete database and archive server. It is the core of the distribution system. It contains all on-line archive sets and provides the access method for information not available at the National hosts.

National Hosts National Hosts are located in different countries to provide easy access to the archive, possibly implementing different versions of the software to access the archive. National Hosts contain the Access and Publications Catalogues and the 1D low-dispersion spectral set: low-dispersion spectra and rebinned high-dispersion spectra. Requests to retrieve spectra are resolved locally or forwarded to the Principal Node. Archive data, catalogues and a basic package to implement the access methods are provided to National hosts in a set of CD-ROMs that also includes selected project documentation.

End User Nodes Unrestricted set of nodes that access the National Hosts and/or the Principal Center via the World-wide Web. No special software is required to access the archives other than standard network support packages.

This architecture allows to easily inter-operate with other data archives so that, for instance, related bibliographical references can be retrieved from NASA Abstract Data System.


PRINCIPAL
CENTER
CATALOGUES
1D & 2D SPEC TRA
CD-ROM
NATIONAL
HOST
CATALOGUES
1D LOW DISP.
END USER
WWW BROWSER

4 Basic National Host Requirements

As indicated in the previous section, National Hosts will receive a set of CD-ROMs containing access and publications catalogues, low-dispersion spectra and re-sampled high-dispersion spectra, selected project documents and a basic software package to access the archive and retrieve spectra.

Baseline hardware configuration is an Intel PC with 32 to 64 Mbytes RAM memory and 3 to 4 Gbytes disk storage. In addition, a CD-ROM reader and Internet connectivity are required.

Linux has been selected as operating system to implement access and distribution procedures. This selection is based on the cost, after careful testing of the system during an extensive time period. The proposed relational database handling the catalogues is MySQL, a licensed product offered at no cost for scientific applications.

The list of INES National Hosts, based on the expressed interest by the Institutes associated with the previous working arrangements for ULDA/USSP, is indicated in the table below; the map illustrates the locations of these Institutes.

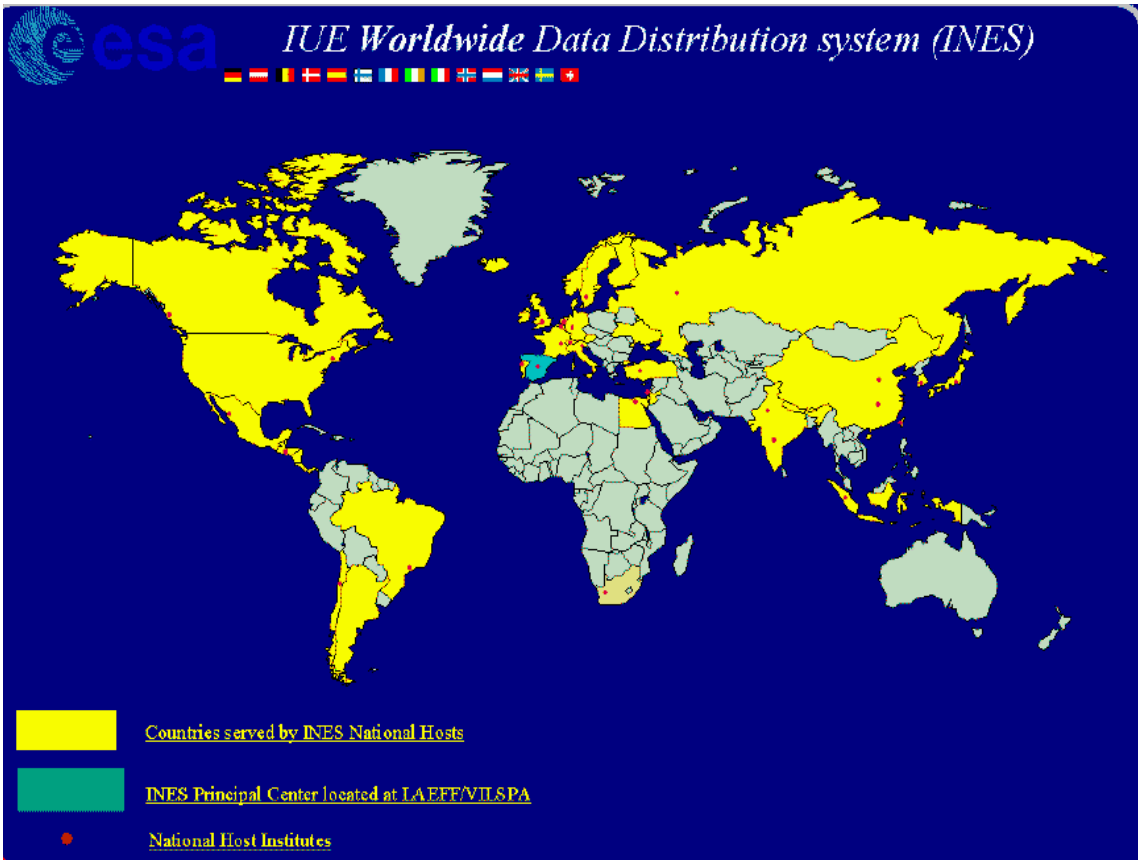


Figure 1: Locations of INES National Hosts

National Hosts	
Austria	Kuffner-Sternwarte, Wien
Belgium	Royal Observatory of Belgium, Brussels
Brazil	Instituto Astronomico e Geofisico, Sao Paulo
Canada	CADC/DAO, Victoria B. C.
Chile	AURA/CTIO, La Serena
China, P.R.	Centre for Astrophysics - USTC, Hefei
Costa Rica	University of Costa Rica, San Jose
Egypt	NRIAG - Helwan Observatory, Cairo
France	CDS - Observatoire de Strasbourg, Strasbourg
Germany	AIT, Tübingen
India	Space Science Data Centre - ISRO HQ, Bangalore
Israel	Wise Observatory, Tel Aviv University, Tel Aviv
Italy	Osservatorio Astronomico di Trieste, Trieste
Japan	National Astronomical Observatory, Tokio
Korea	Department of Astronomy and Space Science, Chungbuk
Mexico	INAOE, Puebla
Netherlands	Sterrenkundig Instituut, Utrecht
Nordic countries	Uppsala Astronomical Observatory, Uppsala
Portugal	Centro de Astrofisica da Universidade do Porto, Porto
Russia	Institute of Astronomy of Russian Acad. Sci., Moscow
South Africa	South African Astronomical Observatory, Sutherland
Spain	LAEFF/VILSPA, Madrid
Switzerland	Inst. d'Astronomie de Universite de Lausanne, Chavannes-des-bois
Taiwan	Inst. of Physics and Astronomy, Chung-Li
Turkey	Physics Department - METU, Ankara
United Kingdom	Rutherford Appleton Laboratory, Chilton
USA	STScI, Baltimore
Local use	ST-ECF/ESO, Garching
Local use	Indian Institute of Astrophysics - VBO, Alangayam