US CONCEPT FOR A CTBT INTERNATIONAL DATA CENTER

STEVEN R. BRATT
ADVANCED RESEARCH PROJECTS AGENCY
NUCLEAR MONITORING RESEARCH OFFICE
e-mail: sbratt@arpa.mil; phone: 703-696-2238; fax: 703-696-2203

ABSTRACT

The Advanced Research Projects Agency (ARPA) has been tasked by the Administration to prototype a Comprehensive Test Ban Treaty (CTBT) International Data Center (IDC) -- a central element of the proposed CTBT verification regime. The IDC would support and make more affordable the verification responsibilities of States Parties by performing the computationally intensive tasks necessary for effective, global monitoring. The IDC would collect and archive large volumes of data from hundreds of worldwide seismic, radionuclide, hydroacoustic and infrasound sensor sites of the CTBT International Monitoring System (IMS). Analyses, performed both by automated systems and human experts, would rapidly reduce this large volume of data by computing key parameters that estimate the existence, location and character of detected events. In the US model, the IDC would stop short of providing a final identification of these events. However, all States Parties would have open and convenient access to all raw data and IDC products, with customized interfaces that allow interested States Parties to apply national criteria to screen for the subsets of data of interest to them.

To meet the requirements for a CTBT IDC, ARPA, together with an international team of developers and operators, is expanding the framework of the Center for Monitoring Research, with an increased focus on data authentication, automated fusion of multi-sensor data, regional knowledge acquisition, reliable and secure distributed processing on UNIX systems, advanced data management technologies, effective data visualization and access, and an open and modular system architecture. The still-evolving prototype IDC is the centerpiece of an on-going Group of Scientific Experts seismic monitoring experiment (called GSETT-3) that began full-scale operations in January 1995. The IDC is producing a daily bulletin containing 50 - 200 seismic events. Radionuclide data are also being collected, processed and disseminated as part of a separate multi-lateral technical experiment. The processing of hydroacoustic and infrasound data will follow later this year. On-going transition of the IDC software to the prototype US National Data Center at the Air Force Technical Applications Center (AFTAC) is expected to improve the US capability to monitor to the low thresholds required to verify a CTBT without increasing the number of human analysts. The US has also offered to transition the product of ARPA's work to the Conference on Disarmament (CD, the body negotiating the CTBT in Geneva) for use within the future international CTBT verification organization. This transition is expected to begin and continue over the next two or three years.

KEY WORDS: automation, data processing, knowledgebased systems, seismology, radionuclide, hydroacoustic, infrasound,

SYNOPSIS

In the view of the US, the fundamental task of the International Data Center (IDC) for the Comprehensive Test Ban Treaty (CTBT) is to provide States Parties with equal, open and convenient access to standard products and services to support their national CTBT verification requirements. In doing so, the IDC would perform a large fraction of the computationally-intensive, technical work required by most States Parties. This would include collecting and archiving data from the International Monitoring System (IMS) sensors and other sources (e.g., associated measures, on-site inspections), performing data analysis to detect and locate events and to compute parameters useful for event identification, and customized provision to States Parties of the subsets of data of interest to them. It would be the responsibility of States Parties to apply national criteria to the technical products from the IDC and other sources to make verification decisions regarding the identification of events, the attribution of those events, and the options for responding within or outside of the CTBT. All products and services would be objective and technical in nature, and would aim to facilitate cost-effective compliance assessments by States Parties. Advanced physical and computer science technologies would be incorporated to allow the IDC to perform its work with as high a degree of security, reliability, cost-efficiency and automation as possible.

The vision for the CTBT IMS is shown in Figure 1. Data from sensors certified as part of the IMS would be transmitted to the IDC either directly or through National Data Centers (NDC) operated by States Parties. It is likely that the IDC would collect and archive a daily volume of up to 10 Gbytes of raw and partially-processed data from over 1,500 channels of data received from 40 - 150 seismic, 50 - 100 radionuclide, 5 - 20 hydroacoustic and 60 - 70 infrasound stations. The IMS sensors would be divided into two categories: primary and auxiliary. The primary sensors would establish a detection threshold by automatically forwarding continuous and/or preprocessed data that would be used to detect and initially locate events. The IDC could automatically or interactively retrieve additional data from auxiliary sensors to improve event location and characterization. Primary, and perhaps auxiliary, sensors would provide digital signatures with each data package which could be authenticated by the IDC or States Parties as a means of increasing confidence in the integrity of the data.

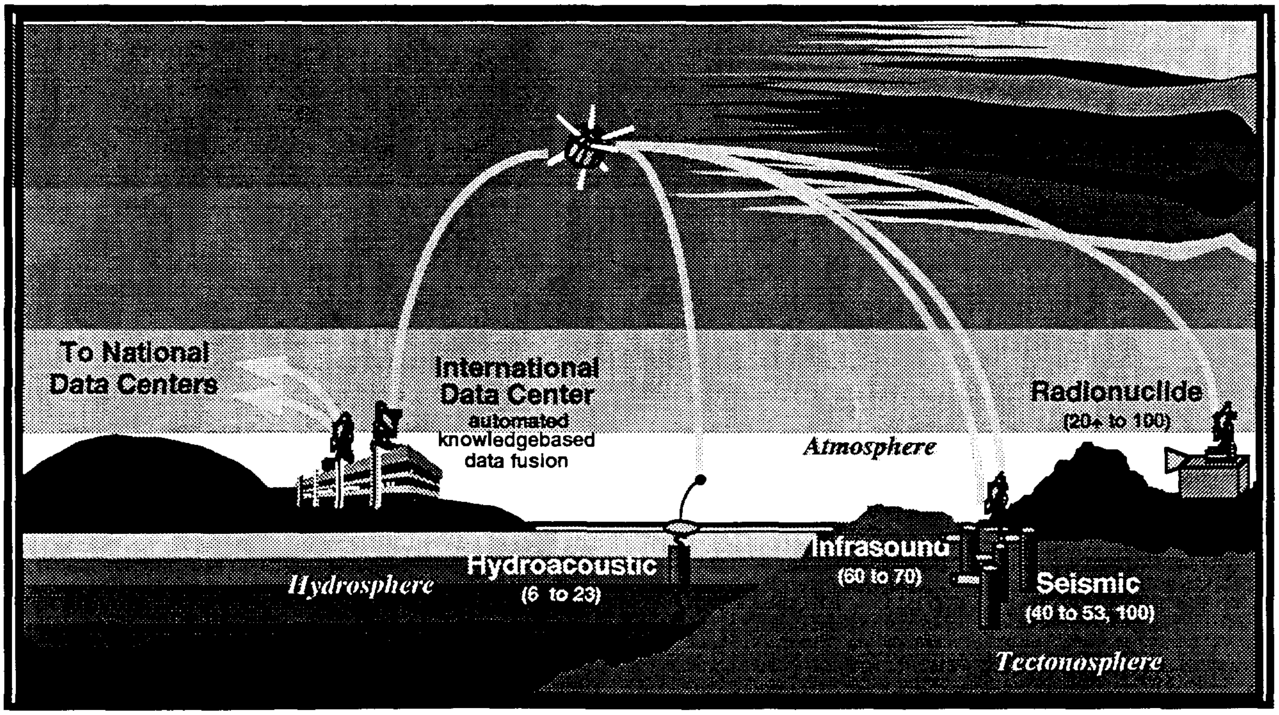

Figure 1. Elements of the proposed CTBT International Monitoring System.

The analysis objective of the IDC would be to produce global bulletins and other products that support the CTBT verification objectives of States Parties. The differences between the techniques would typically be related to the schedules for processing, the particular algorithms applied during processing, and whether or not there is a need for both primary and auxiliary data processing. Seismic, hydroacoustic and infrasound are timeseries data that are similar enough that they could be analyzed with variants of the same software. However, radionuclide analyses diverge somewhat from the others, in that the propagation of radionuclides from the causative event to the sensors can take days, location of events requires temporally-varying meteorological models, there may be some degree of in-field processing, physical filter and gas samples may need archiving and additional analyses, and coordination with certified labs may be required. In general, an initial, automatically-computed event list would be available as rapidly as possible, depending on the monitoring technique, after the event's occurrence. An analyst-reviewed bulletin for each technique would be available hours or days later, again depending on the technique. In addition to the single technique products, “fused” event lists, which contain correlated signals and located events based on different techniques, would be compiled automatically and updated over time (Figure 2). The final product would be an analyzed, "fused" event bulletin. The processing and analysis procedures at the IDC would be scientifically validated and documented in the Operational Manual for the International Data Center. This manual would provide for IDC procedures to meet changing requirements and to improve operational quality and efficiency.

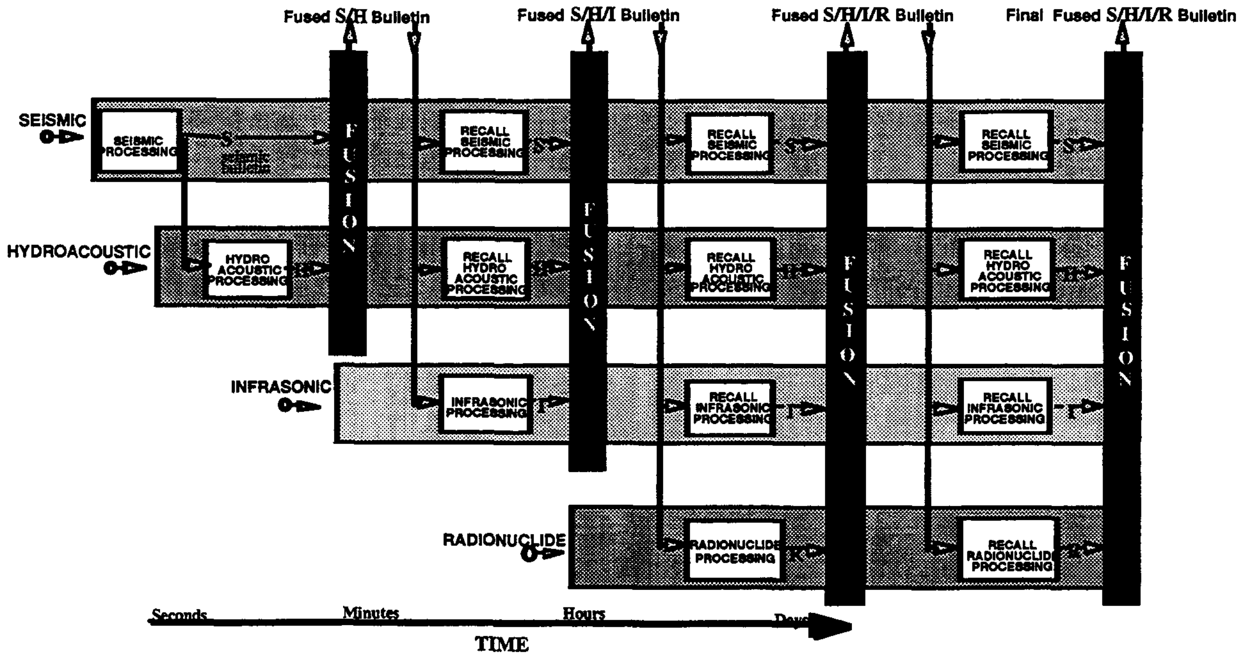

Figure 2. IDC fusion architecture and data flow.

The IDC would provide data access services that make it convenient for States Parties to find and retrieve, only those data of interest to them. Services would range from the re-transmission of large volumes of continuous, raw data, to provision of select, relatively small, subsets of data and products. For the latter, the IDC would allow States Parties to establish “subscriptions” that apply their own national criteria to screen for and automatically forward the subsets of data necessary to meet national verification needs. These tools would generally use the parameters computed from the raw data by the IDC as a basis for the screening.

Prototyping and Testing a CTBT IDC

The US is leading an international effort to prototype and test an IDC for processing seismic data during the Group of Scientific Experts Third Technical Test (GSETT-3). GSETT-3 began full-scale operations in January 1995 (CD/1296). The US is building upon the GSETT-3 system to prototype and test a more general IDC for processing data from all CTBT monitoring techniques. The US has offered to transition this prototype to the future CTBT organization (CD/NTB/WP.192). ARPA continues to transfer elements of the IDC to the prototype US National Data Center at AFTAC. The schedule for development, testing and transfer is shown in Figure 3.

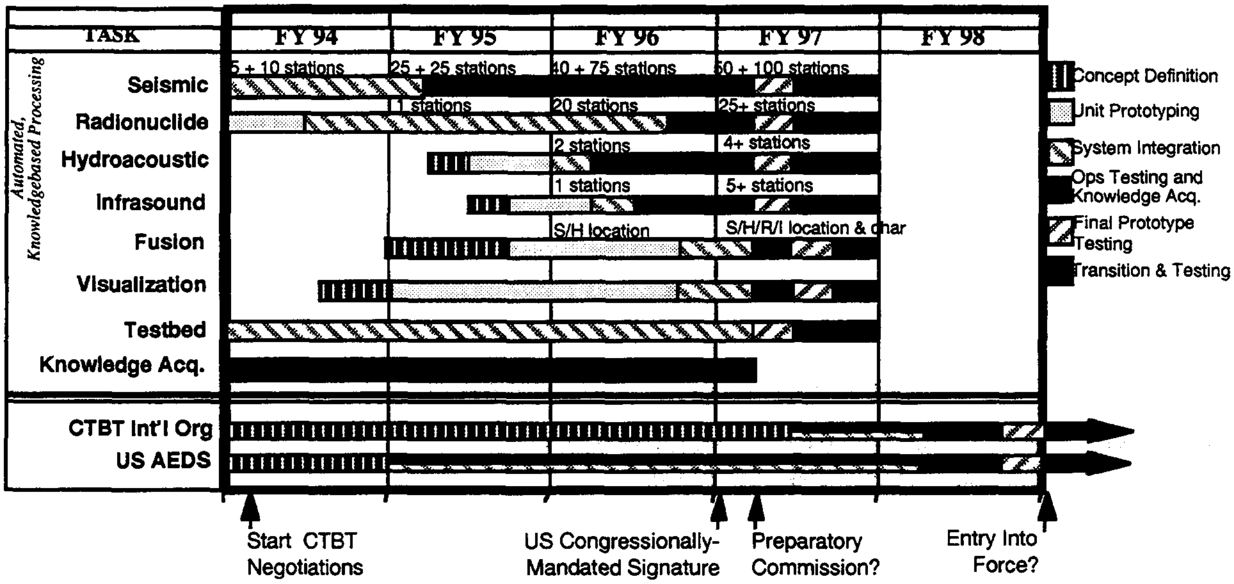

Figure 3. Schedule for development, testing and transition of the prototype IDC.

The IDC is being prototyped and tested at the Center for Monitoring Research (formerly, the Center for Seismic Studies) in Arlington, Virginia. The hardware and software infrastructure has been designed to be flexible enough to support sophisticated, knowledge-driven fusion of data collected from seismic, hydroacoustic, infrasound and radionuclide sensors. The seismic processing capability being developed in support of GSETT-3 is the most mature at this time, and is quite unique. Signal processing and event location systems will incorporate station- and path-specific knowledge (e.g., detection parameters and phase identification parameters, time and amplitude path corrections). The knowledge base and the resulting quality of the products are improving with experience. Signal and event parameters not normally presented in earthquake bulletins (e.g., spectral ratios, spectral variance, complexity, and measures of similarity to previous events) are also routinely computed. The detection capability of the network is permitting calibration of regional magnitude scales around the world. New concepts for maintaining the integrity of the facilities, the data, the products and the systems are being tested. All products are being archived on-line at the IDC and made available to NDCs via the variety of data access services. Seismic bulletins and a wide range of information are also available to the public through the IDC's World Wide Web pages at http://www.cdidc.org/.

Presented in Figure 4 are the primary and auxiliary seismic stations planned for GSETT-3. Figure 5 shows the locations and error ellipses (most too small to be seen) of events from the first almost seven months of the GSETT-3 Reviewed Event Bulletins (REB). These events were analyzed and available within about two days after the end of the data day. There are on average 60 events, and up to 200 events during large aftershock sequences, in the REB each day. Estimates of the 90% threshold for detection by three Alpha stations range from m_b 3.1 in northern Europe up to m_b 4.7 in the southern oceans. Typical location uncertainties for events within 2000 km of the closest station are between 40 and 50 km. Completion of the station network and calibration through the use of regional, rather than globally-averaged travel times, will reduce the location uncertainty. Table 1 demonstrates some of the progress that has been made during the preparation for and conduct of full-scale GSETT-3 operations.

	1 Feb 1994	20 Aug 1994	20 Feb 1995	1 Aug 1995	Plan
Nations	10	13	28	38	>40
Alpha Stations	6	12	32	35	60
Beta Stations	13	21	44	71	~150
Coverage	European	Sparse Global	Global	Global	Global
NDCs -> Gamma data	1	5	11	14	20-30
Data Vol/day to IDC	0.2 Gbytes	1.5 Gbytes	2.5 Gbytes	2.5 Gbytes	>5 Gbytes
Data Days/week	2	3	7	7 (24hrs/day)	7 (24hrs/day)
REB Events/day	20	20	60	60	>100
REB Events Total	1500	2700	6800	11000	
IDC Staff (int'l)	27(0)	27(0)	47(9)	49(10)	50(9)

Table 1. Key statistics for GSETT-3 as a function of time.

Over the past year, increasing effort has been focused on adapting the GSETT-3 IDC infrastructure to support prompt, secure and reliable collection, processing, analysis, storage and provision of data derived from the diverse types of sensors envisaged for the CTBT IMS. Shown in Figure 4 is the suite of hydroacoustic, infrasound and radionuclide sensors expected to be providing realtime data to the prototype IDC during 1995. The first phase of this effort has been to integrate atmospheric radionuclide processing into the IDC. The data flow in this system, at a high-level, is similar to that for seismic, in the sense that signal detection, location, characterization and timely reporting are important requirements. The prototype radionuclide processing system is testing new concepts that focus on automating and accelerating these processes. By the end of 1995, the IDC will likely be collecting and analyzing data from stations in the US, Australia, Canada, Finland, Germany, Kuwait, Russia, South Korea, Sweden and Turkmenistan. Future work will focus on integrating into the prototype IDC innovative concepts for hydroacoustic and infrasound monitoring, and for fusion of data from multiple techniques to improve event detection, location and characterization. It is expected that data from hydroacoustic stations near Wake and Ascension Islands, and from infrasound sensors in New Mexico and Utah will start arriving at the IDC during the fall of 1995.

Credits

The principal contributors to the success of the International Data Center include Science Applications International Corporation, Pacific Sierra Research Corporation, AFTAC, Atlantic Aerospace, Australian National University, Autometric, Canberra, Cornell University, Digital Systems Research, Ensco, Harvard University, Mission Research, Multimax, NOAA, NORSAR, Planing Research Incorporated, Ruhr University, Science Horizons Incorporated, S-CUBED/Maxwell Industries, Southern Methodist University, Teledyne-Brown, Torrey Sciences, Trusted Information Systems, University of California - San Diego, University of Virginia, Woodward-Clyde Federal Services and the many international experts who have provided service, software and analyses.

REFERENCES

References

CD/1296 (1995) *Progress Report to the Conference on Disarmament on the Fortieth Session of the Ad Hoc Group of Scientific Experts to Consider International Cooperative Measures to Detect and Identify Seismic Events*, Conference On Disarmament, Geneva, 17 pp.

CD/NTB/WP.192 (1994) *U.S. Approach for the Comprehensive Test Ban Treaty International Data Center*, Conference On Disarmament, Geneva, 25 pp.

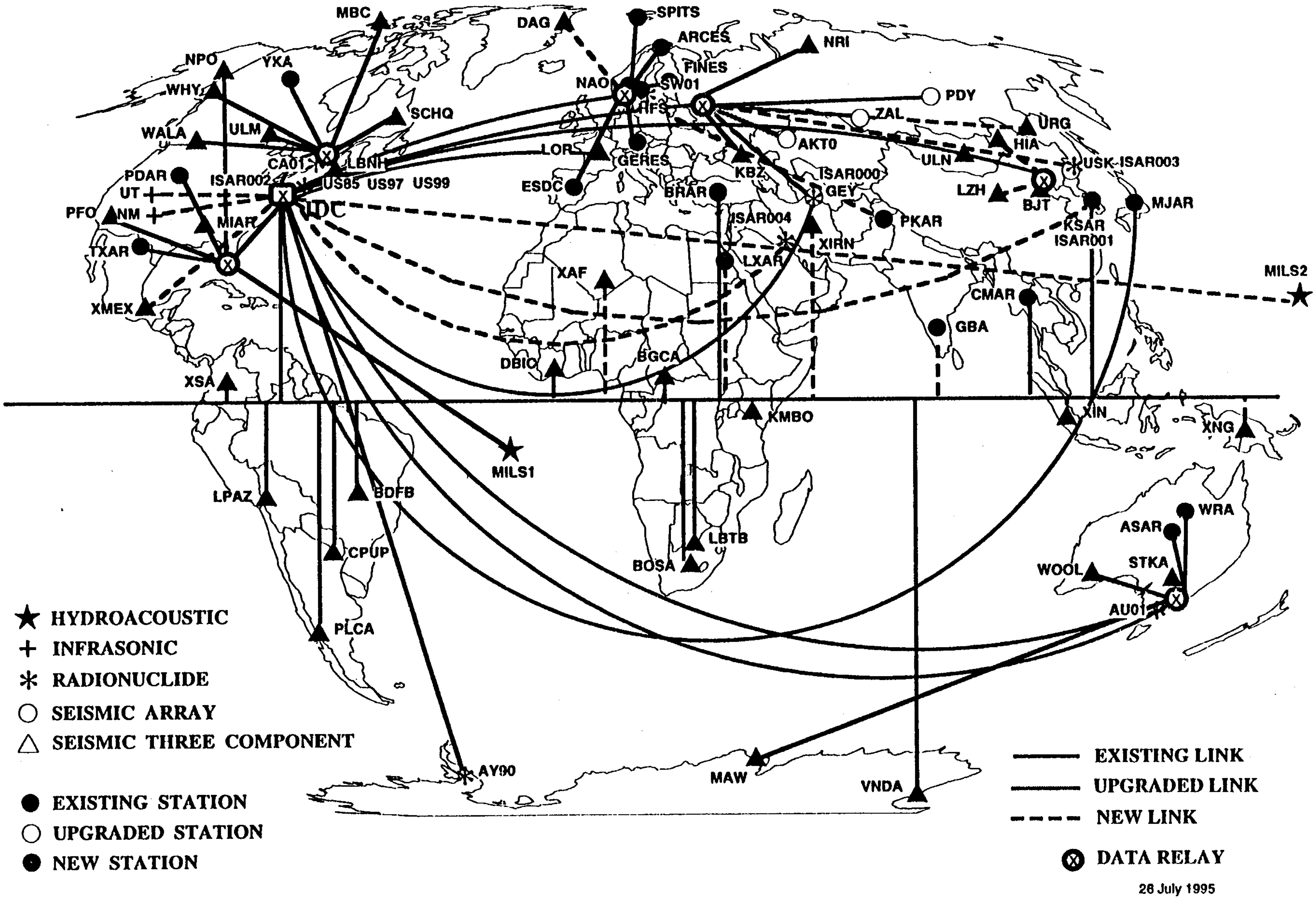

Figure 4. Prototype International Data Center Sensor Network for 1995.

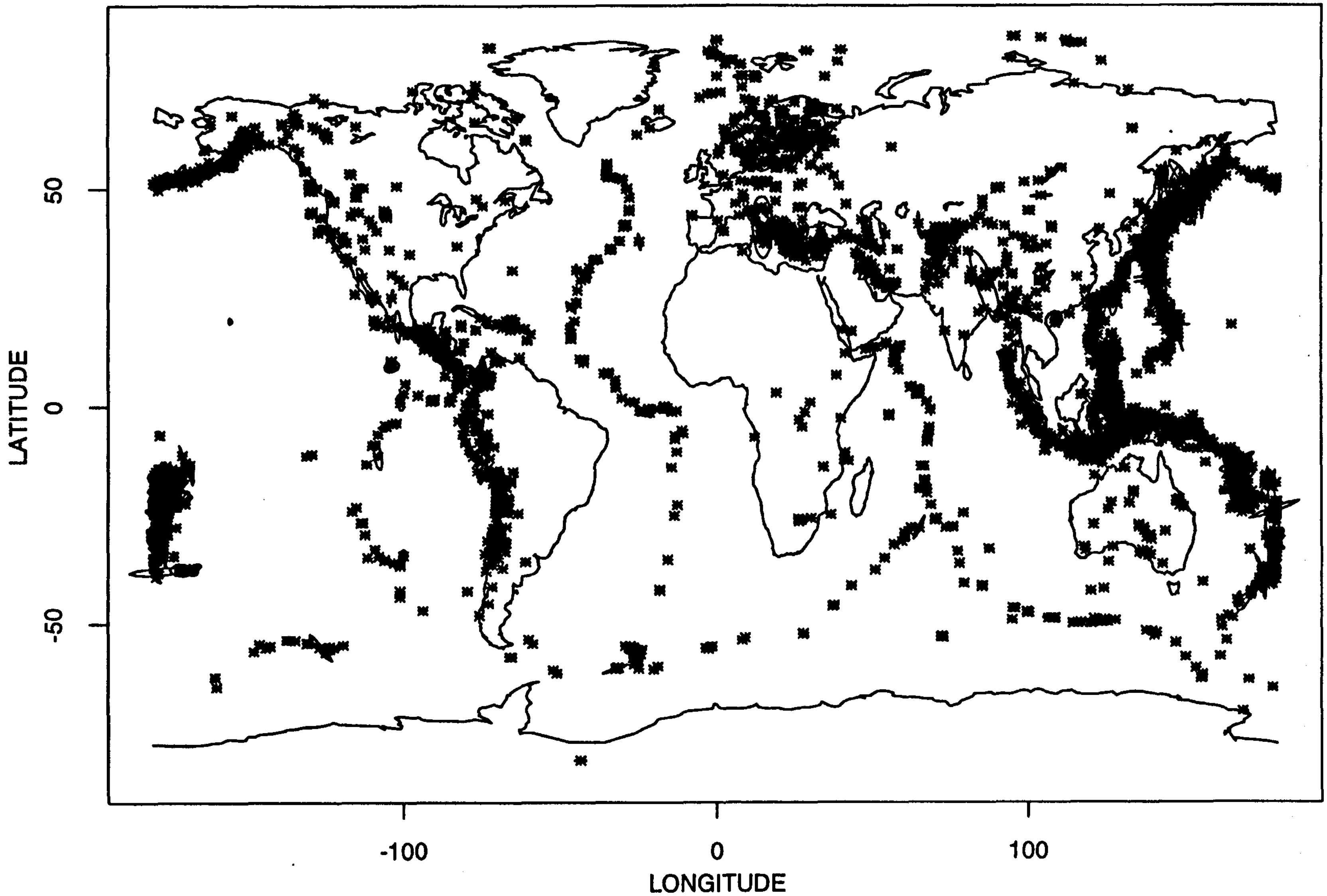

Figure 5. GSETT-3 Reviewed Event Bulletin events from January 1 through July 23, 1995, with at least 6 defining phases.