System Architecture of Seismic Networks and its Implications to Network Automatization

M. Joswig
Institute for Geophysics, Ruhr-University Bochum, Germany

Introduction

The last 30 years have seen great improvements in earthquake measurement and interpretation methods due to the introduction of digital data acquisition. However, the variety of geographical layouts with different optimization constraints and the rapid progress of computer equipment has produced a large number of solutions. Different in more or less significant features, they can't be overseen by an individuum any more. In order to allow for concise discussions, we must classify this variety. As important as the pure classification, we must point out how the different technical solutions will interact with the possible approaches of automated seismogram analysis. The number of 20.000 events per year, contained in Terabyte of raw data and achieved today by arrays like NORESS or GERESS, already indicates that any further progress in network sensitivity can be utilized only if seismology can rely on methods to unload the human from the majority of observatory routine works.

Network generations

For now, let's take any spatial distribution of connected, equally equipped stations as a seismic network - the necessary distinction of arrays will be introduced later. Then we are able to summarize all the individual solutions of digital seismic networks into 3 different generations. They follow up the technical possibilities and represent - each at its time - the achievable state-of-the-art in earthquake measurement.

The first generation marks the beginning of digital networks at all. Computers were expensive at these times, so the most affordable solution was to place it in the hub of the observatory - see Fig. 1. All the signals were transmitted analog, then digitized and archived. In many cases, the same computer served as processing tool for the human analysis as well. Examples for this architecture date back to the mid-60ies, most notably the USGS-network in California with hunderts of stations (Lee & Stewart, 1981). Even today this principle is in wide use all over the world, an example for a just recently installed network is the ISSN in Israel (Shapira, 1990). The transmission of analog signals is performed by FM either on leased phone lines or via telemetry and is, in fact, the most significant disadvantadge of this concept. Even in good conditions, the dynamic range is limited to some 40 dB; in practise we must add many noise bursts and spikes that tend to disturb any digital filtering routine or autoscaling of seismogram plots.

These restrictions led to the development of the second generation of digital networks. Once the computers got cheaper, it was possible to place them outside too to perform the analog to digital conversion on site. This allowed for a dynamic range that is limited only by the technology of ADC circuitry and the availabe bandwidth of the communication lines. This enhancement in dynamic range was most urgent for broad-band seismometry, so one of the earliest implementations of a second generation network was the GRF-array in Germany in 1975, equipped with gain-ranged 15 Bit ADC's and a mini-computer (!) at each site (Harjes & Seidl, 1978; Henger & Stork, 1986). The achieved dynamic range was already 24 bit, only the resolution was compromised to 12 bit. In common with the networks of the first generation, the data transmission was still on-line and demanded dedicated lines. Thus the most easy way of time synchronization was still to perform it by one central clock, when all the data had arrived in the hub.

The introduction of the third and most recent generation of seismic networks once again reflects the progess in computer technology. Microprocessor boards with large RAM are cheap now and deliver significant processing power at each site which is best used for more independent remote units. In full realization, this set-up consists of the real-time data acquisition with local time synchronization, a rudimentary data base system to hold the seismograms, a single-trace detector like STA/LTA and the communications frontend for the dialog with the observatory hub (see Fig. 2). While these functional units are realized in any digital field box for refraction measurement at date, it is important to stress that for the continuous data acquistition in seismology, they must be realized in a multi-tasking environment. This excludes some of the current refraction units like Reftek 72A or Geotech PDAS 100 which can either measure or communicate. On the other hand, some more sophisticated field stations like Geotech RDAS 200 or Lennartz MARS 88 are available off-shelf to build up a third generation network.

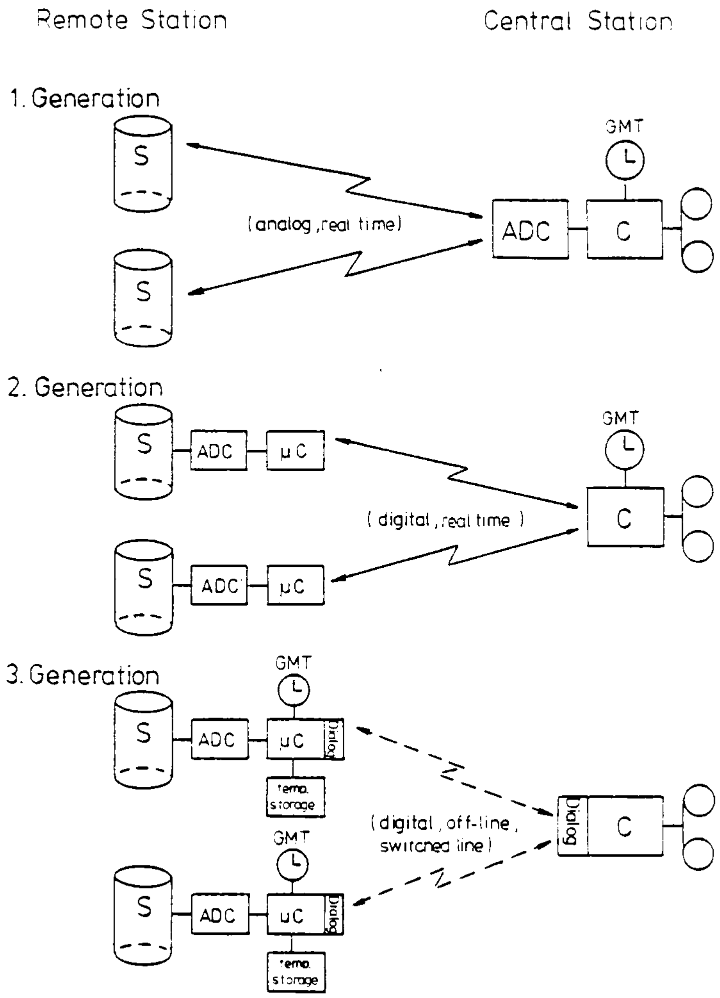

Fig. 1 Generations of digital seismic networks

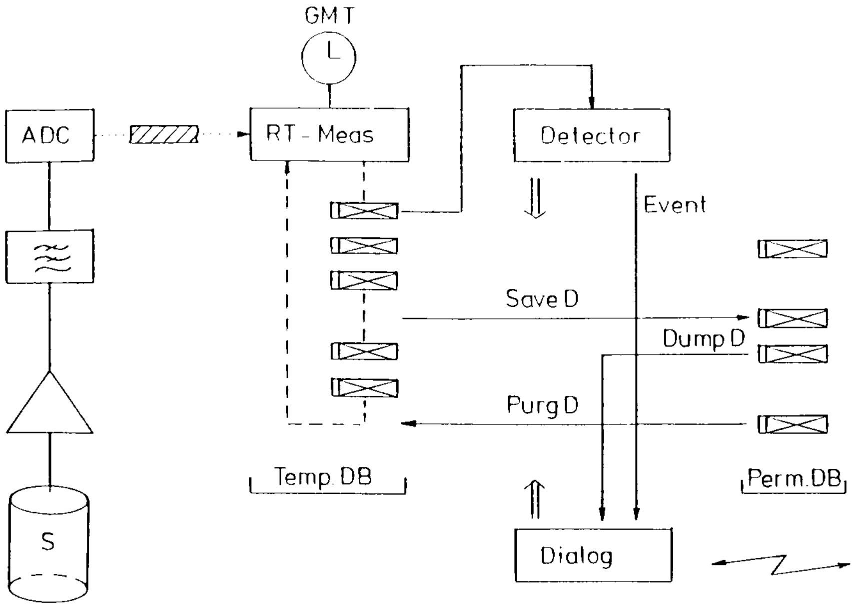

Fig. 2 Functional elements of 3. generation field station

The first realization, to our knowledge, was the Bochum University Germany (BUG) network that was designed 1983 and is in operation since end of 1986 (Joswig & Harjes, 1986; Joswig, 1987). At that time, all electronics and software must be custom made, so it took us 4 years to finalize. An additional aspect of the BUG network was its software approach that introduced the ISO/OSI concept of a layered computer network to seismology (see Fig. 3). The advantadge of clearly separated layers is that the network design can be handled twice: (I) bottom-up with all the details in ADC, real time microcomputer programming and data communication protocols for the view of technicians as well as (II) top-down. The latter view is the one of science since, seismologists will deal with event collection (automated by detectors), interactive phase picking and location algorithms in an analysis program only. So this turns out to be the more important view - on condition that the basic layers are engineered well in the one or other way. For the BUG network, we introduced the concept of a network-wide data base layer that satisfies the actual data requests of the seismologist regardless where the data are stored physically. This transparency demands a high level command shell ('CI') that allows for simple syntax with wildcards for station, component and datatype options.

The obvious difference between networks of 1. or 2. versus 3. generation is in the data communications part. Dedicated real-time connections can be replaced by any off-line medium, even swichted connections by X.25 or telephone modems are possible. The off-line criterion allows for error-checked protocols with retransmission like TCP/IP or MNP5; on the other hand, it demands a time receiver at each remote station. Instead of simplex telemetry in the former networks, now we need duplex communications that allow for data flow from stations to hub and command transmission vice versa.

The most discussed impact, however, is the chance of restrictions in the stream of data transmission. The local detectors could be used to perform a preselection of those seismogram segments that are worth any further evaluation while the rest of data is treated as environmental ground noise that can be purged on site. This concept is not agreed on by all seismologists, it also depends on the scientific task of the network as will be seen later. In all those cases where we choose this way for a given situation, then the size of local storage capacity can be used as a second order diversative for third generation networks. In the first case, we reley completely on automated processing, so the local FIFO storage is just sufficient to reflect the time delays by some coinicidence decision in the network hub, the transmission time and some spare for retransmission or short term network breakdown. The overall time results in some hours and can be realized by solid state memory to give very rugged, low power field stations. The other alternative is that the seismologist can specify any request for data within some reasonable delay which will yield storage times of some weeks. Then we need disc space on site and consequently a more sheltered environment for our stations.

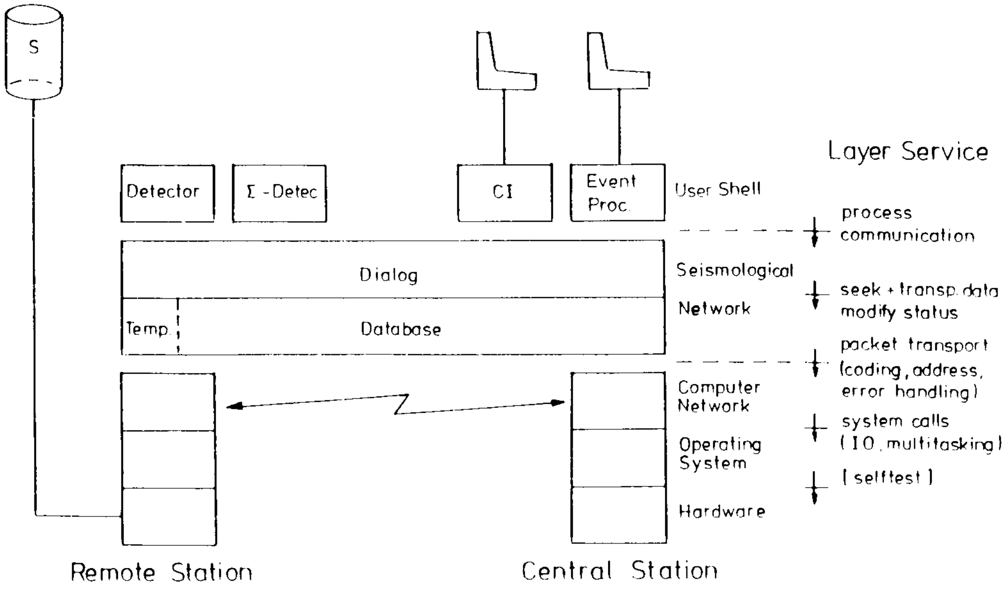

Fig. 3 System layers of a seismic network

Layout of seismic networks

Independent of the technical generations, seismic networks can be distinguished by their geographical layout. If our task is to monitor a given region - or the whole world - for the actual seismicity, we can choose a station distribution that covers the whole area (see Fig. 4). This is, in a more strict sense, the common understanding of a seismic network. The principal alternative is the seismic array with many stations on small aperture. If we neglect performance criteria for a moment, both set-ups can be realized by many stations in the dense network or array configurations and by a minimum of 3 stations for the sparse arrays or networks; the future will also see networks of array sites. For the arrays, we sample the spatial distribution of a plane wavefield and thus can utilize the inter-station signal coherence in *integral processes* like beam-forming, fk-analysis (Capon, 1969) or vespagram calculation (Davies et al., 1971) - see Fig. 5. These approaches are, in general, more powerful than any

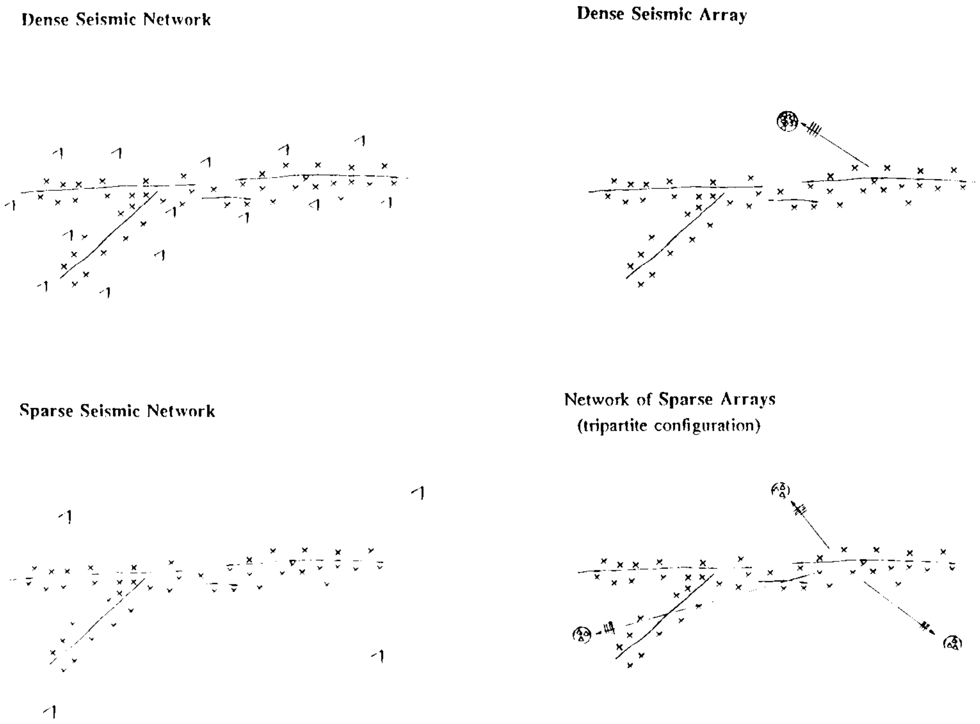

Fig. 4 Seismic monitoring by different geographical layouts

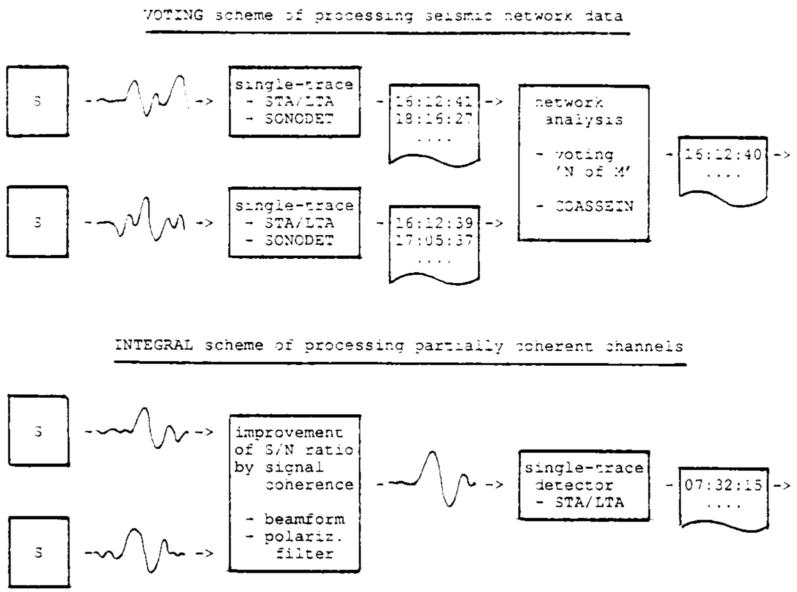

Fig. 5 Signal processing schemes in seismic networks

independent processing of single seismogram traces with subsequent *voting scheme* (Ringdal et al., 1972, 1975; Wirth et al., 1976). In seismic networks, however, we are restricted to the latter because of lack of network-wide signal coherence, so a more exact name could be the *incoherent network*.

Usually the definition of an array also includes some common, on-line data processing and uniform instrumentation (e.g., Ringdal & Husebye, 1982). With sufficient data quality, a seismometer substitution by software is simple nowadays and the common data processing can be achieved even for data of different networks by means of the innovative 'open station' concept that allows for seismogram access via computer network in near real-time (e.g., Kanamori & Hauksson, 1991). The only 'hard' criterion that remains valid for the distinction between arrays and networks is the ratio of the length of signal coherence at f_{max} versus the maximum station aperture. This ratio will change for any fixed site in dependence of the epicenter distances. So it is only consequent that an array for teleseism like GRF with its 100 km length disintegrates into an incoherent network for local events, vice versa any local network yields array traces for teleseism.

Our weaker distinction is based on the analysis methods that can be performed on selected seismograms and allows us to treat, e.g., the NARS configuration appropriately as an array (Dost, 1990). We should call it *off-line array* to dintinguish from the 'real', *on-line array* that must fulfil some additional criteria. Firstly, its station distribution is by no means arbitrary but tuned for an optimum in its spatial transfer function, i.e., small and distinct side lopes outside the wavenumber area of interest (Mykkeltveit et al., 1983). Secondly, the integral processing steps of Fig. 5 are performed on all the data, thus an on-line array excludes any kind of data preselection at its single sites and demands continuous - but not real-time - data flow to the hub.

Because the distinctions by technical generation and geographical layout are completely independent, we find any combination of them. At date, the most advanced installations are, of course, all of 3. generation. In this technology, an example for the dense array set-up is GERESS in SE-Germany (Harjes, 1990); a network realization is the GRN with two weeks of local data storage (Hanka, 1991) while a sparse, tripartite array is contained in the BUG network.

The need of automatization

Any improvement in the measurement techniques yields more data that must be transmitted, processed and archived. The main reason for this effect is in the nature of seismicity itself: The number of recognized events depends exponentially on a lowered magnitude threshold. The other cause of more data is the increase in sampling rate, dynamic range (i.e., bytes per sample), number of stations and three-component sites; all these effects result in a larger set of data per single event.

The need for automatization is twofold as well. Firstly, we reach an amount of data that can not be handled by humans any more. In Bochum, the observatory must deal with 1 GB of data per day - 600 MB from GERESS and 400 MB by the local BUG network. Secondly, the communication costs tend to dominate the total budget of running and even purchasing a seismic network. This is especially annoying when most of the money is wasted for data segments with ground noise only. The general solution for both problems is in preselecting the data by detector approaches. For the reduction in communication costs, this preselection must already be performed at the remote sites.

When a few scientists started the development of automated seismogram processing some 20 years ago, they focussed on many-station arrays and high frequency local networks. Here the need of automatization was so urgent that seismologists were willing to accept the achieved methods even despite of poor performance. Today, the initial signal processing methods are complemented by Expert Systems, Pattern Recognition and Artificial Neural Networks. So with matured approaches and even more data than years ago, we once again face the questions: Where is the actual borderline between automated processing and human analysis? Are humans still better than computers in all tasks and under every condition? Where should we - forced by sheer necessity - and where can we trust automated seismogram interpretation?

The disput on these topics has not even been started but some preliminary statements can already be made. The task of automatization can be subdivided into the initial parameter extraction from the seismograms and some subsequent reasoning on these parameter data (Joswig, 1991a). The most urgent benefit of automatization would be to offload the seismologist from routine processing which means the majority of small and noisy events that just contribute to bulletin statistics. In general, the decrease in signal to noise ratio will impose great difficulties on every numerical algorithm but in seismology we have the one advantage that the fine structure differences of earthquake signals will decrease as well. For the fewer large events, each one is an individuum and demands some unique analysis for the fault plane solution and investigations on source processes or details of the propagation path. The manifold of weak events, however, appears in clusters, aftershock sequences and repitive patterns from quarry blasts. This situation is ideal for robust but fuzzy recognition techniques based on a limited number of types. Our interest is limited to hypocenter and magnitude distributions so the derived parameters can be restricted as well to onset times, azimuths, maximum amplitudes and cluster associations. To sum it up, the necessity and chance of seismology lies in the automated processing of weak events while the more extended analysis of important earthquakes should be left to humans.

Implications of network architecture

Two consequences of network architecture for automatization and network operation have already been derived in the previous chapters: (I) the third generation networks allow for on-site preselection of data segments and (II) on-line arrays need all the data in the hub to apply some integral precessing schemes. The other consequences could be characterized by the phrase "big efforts pay". Once we invest in many stations for the dense network or array concept, the automated parameter extraction gets easy. For the arrays, beamforming improves the S/N ratio and the fk-analysis yields a unique criterion for the phase association of detected onset times by the appearent velocity (Bache et al., 1990; Mykkeltveit & Bungum, 1984). The evaluation in seismic networks can always be guided by the nearest stations, their good S/N ratio allows for simple detectors (Chiaruttini et al., 1989; Roberto & Chiaruttini, 1991). However, a drawback from the primitive waveform evaluation is that the subsequent rule-based system must cover a lot of exceptions and obvious errors by explicit coding.

In sparse configurations, the parameter extraction is far more critical because no redundancy in the data would allow for later correction by explicit rules. Instead, wrong parameters yield wrong, but uncommented event determinations. This is worst of all for an automated system since it undermines the confidence in any reported result. So sparse arrays and networks must perform some smart, knowledge-based parameter extraction with human-like performance in detection, phase picking and cluster association. On the other hand, the more powerful parameters and the limited number of station combinations result in a much simpler rule-based system that must follow for reasoning on the parameters.

What we have seen when reducing the number of stations in a given configuration of array or incoherent network from dense to sparse, i.e., below some 5 sites, is a change of quantitiy to quality which affects the fundamental choice of appropriate methods for automatization. For the tripartite BUG array, we have developed some of these sophisticated waveform analysis tools like the Sonogram-detector (Joswig, 1990, 1991b), a S-phase picker based on polarization images (Klumpen & Joswig, 1991) and a non-linear crosscorrelation by Dynamic Waveform Matching to perform cluster association and azimuth determination (Joswig & Schulte-Theis, 1991). They fit into a scheme of subsequent, hypothesis-guided refinement of parameters for the routine event determination in Fig. 6 to produce an automated bulletin.

An additional feature of sparse arrays is that its beam has commonly worse S/N ratio than the best single station and that the spatial coverage is not sufficient for fk-analysis. So any integral processing does not make sense in the on-line task of event detection, instead we follow the scheme of voting for incoherent networks. Then a data preselection can be performed already on-site and we can benefit from reduced communication costs. In the long term run, networks of sparse arrays could emanzipate themselfs as the 'best buy' for the ratio of monitoring performance versus site and maintenance costs once a full scheme of reliable automatization is established.

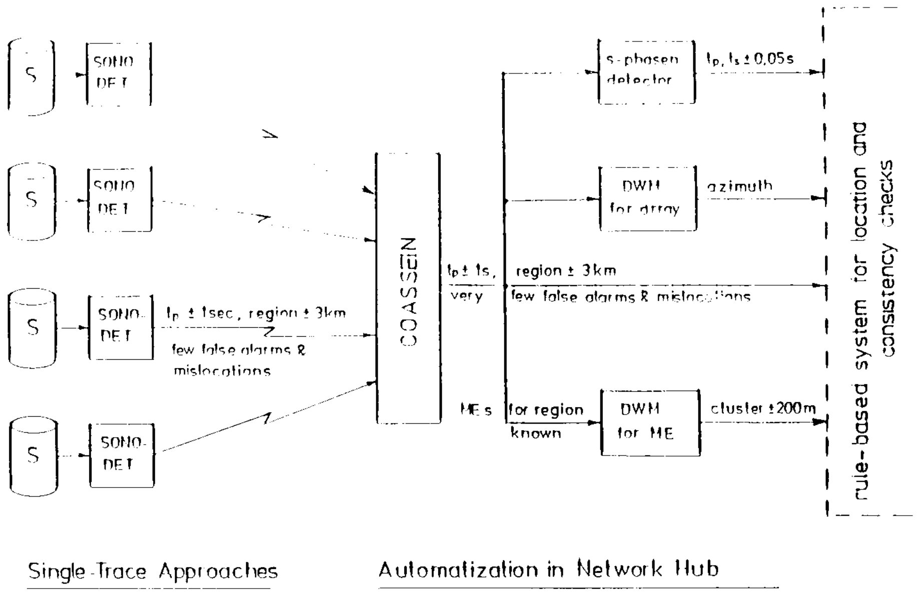

Fig. 6 The automated processing for the BUG sparse array

Conclusions

The third generation of digital seismic networks consists of remote stations with ADC, time synchronization, data base management, detector and dialog to the network hub for data and command transmission. The local storage can either be tuned for automated decision by a FIFO length of some hours or for human interaction by storage capacity of some weeks.

The network can be described best top-down in a layer model to isolate the technical constraints from the principle scientific demands of transparent access and processing of seismograms. With the third generation equipment, any layout of sparse or dense station distribution can be realized. The distinction between arrays and incoherent networks that can perform a voting scheme only, should base on the ratio of signal coherence length versus station aperture.

The increase in raw data forces automatization of routine works, both to unload humans and to reduce the communication costs. This automatization subdivides into parameter extraction and subsequent reasoning. The high efforts in dense station distribution pay today in more simple approaches for parameter extraction, while the future will see emanzipation of the sparse station layouts by more sophisticated seismogram processing utilizing knowledge-based approaches.

Literature

Bache, T. C., S. R. Bratt, J. Wang, R. M. Fung, C. Kobryn, and J. W. Given, "The intelligent monitoring system", *Bull. Seism. Soc. Am.*, vol. 80, no. B6, pp. 1833-1851, Dec. 1990.

Capon, J., "High-resolution frequency-wavenumber spectrum analysis", *Proc. IEEE*, vol. 57, pp. 1408-1418, 1969.

Chiaruttini, C., V. Roberto, and F. Saitta, "Artificial intelligence techniques in seismic signal interpretation", *Geophys. J. Int.*, vol. 98, no. 2, pp. 223-232, 1989.

Davies, D., E. J. Kelly, and J. R. Filson, "The vespa process for the analysis of seismic signals", *Nature*, vol. 232, pp. 8-13, 1971.

Dost, B., "Upper mantle structure under western Europe from fundamental and higher mode surface waves using the NARS array", *J. Geophys. Int.*, vol. 100, pp. 131-151, 1990.

Hanka, W., "The German regional broadband seismic network (GRN) project", in *Workshop on MedNet - The broad-band seismic network for the Mediterranean*, ed. E. Boschi, D. Giadimi, A. Morelli, pp. 83-95, Il Cigno Galileo Galilei Edizioni di Arte e Scienza, Roma, 1991.

Harjes, H.-P. and D.Seidl, "Digital recording and analysis of broad-band seismic data at the Gräfenberg (GRF)-array", *J. Geophys.*, vol. 44, pp. 511-524, 1978.

Harjes, H.-P., "Design and siting of a new regional seismic array in central Europe", *Bull. Seism. Soc. Am.*, vol. 80, no. 6B, pp. 1801-1817, 1990.

Henger, M. and B. Stork, "Digital acquisition, transmission and recording of array data", *Geol. Jahrbuch*, vol. Reihe E, no. 35, pp. 20-27, Hannover, 1986.

Joswig, M., *Methoden zur automatischen Erfassung und Auswertung von Erdbeben in seismischen Netzen und ihre Realisierung beim Aufbau des lokalen 'BOCHUM UNIVERSITY GERMANY' - Netzes*, Wissenschaftl. Veröffentl. (Reihe A), pp. 1-124, Inst. f. Geophysik, Ruhr-Univ., Bochum, 1987.

Joswig, M. and H.-P. Harjes, "A seismic network with loosely coupled remote stations", in *Proc. XX. General Assembly of ESC 1986 in Kiel*, ed. D. Mayer-Rosa, J. M. Van Gils, H. Stiller, pp. 117-127, Fed. Inst. Techn., No. 101, Zürich, 1987.

Joswig, M., "Pattern recognition for earthquake detection", *Bull. Seism. Soc. Am.*, vol. 90, no. 1, pp. 170-186, 1990.

Joswig, M. and H. Schulte-Theis, "Master event correlation of weak local earthquakes by Dynamik Waveform Matching", *Geophys. J. Int.*, (submitted), 1991.

Joswig, M., "Rule-based coincidence evaluation of earthquake detections in the BUG small array", *Bull. Seism. Soc. Am.*, (submitted), 1991a.

Joswig, M., "Earthquake recognition by the Sonogram-detector with master event adaptation", *Bull. Seism. Soc. Am.*, (submitted), 1991b.

Kanamori, H. and E. Hauksson, "Terrascope", *IRIS newsletter*, vol. 10, no. 2, pp. 1-3, 1991.

Klumpen, E. and M. Joswig, "High-frequency S-phase discrimination for local earthquakes", in *Proc. XXII. General Assembly of ESC 1990 in Barcelona*, (in press), 1991.

Lee, W. H. K. and S. W. Stewart, *Principles and applications of microearthquake networks*, Academic Press, New York, 1981.

Mykkeltveit, S., K. Astebol, D. J. Doornbos, and E. S. Husebye, "Seismic array configuration optimization", *Bull. Seism. Soc. Am.*, vol. 73, pp. 173-186, 1983.

Mykkeltveit, S. and H. Bungum, "Processing of regional seismic events using data from small-aperture arrays", *Bull. Seism. Soc. Am.*, vol. 74, no. 6, pp. 2313-2333, 1984.

Ringdal, F., E. S. Husebye, and A. Dahle, "Event detection problems using a partially coherent seismic array", NORSAR Technical Report 45, Kjeller, Norway, 1972.

Ringdal, F., E. S. Husebye, and A. Dahle, "P-wave envelope representation in event detection using array data", in *Exploitation of Seismograph Networks*, ed. K.G. Beauchamp, pp. 353-372, Noordhoff, Leiden, 1975.

Ringdal, F. and E. S. Husebye, "Application of arrays in the detection, location, and identification of seismic events", *Bull. Seism. Soc. Am.*, vol. 72, pp. S201-S224, 1982.

Roberto, V. and C. Chiaruttini, "Seismic signal understanding: a knowledge-based recognition system", *(submitted for publication)*, 1991.

Shapira, A., "Seismological Bulletin No. 9", Inst. f. Petrol. Research, Holon, Israel, 1991.

Wirth, M. H., R. R. Blandford, and R. H. Shumway, "Automatic seismic array and network detection", *Bull. Seism. Soc. Am.*, vol. 66, pp. 1375-1380, 1976.