doi:10.3969/j.issn.0253 - 4967.2009.02.015

基于 SeisComP 大型地震台阵的数据管理

吴明恕 刘 治 陈九辉*
(中国地震局地质研究所,北京 100029)

摘 要 随着中国科学探测台阵观测系统的建成,大规模流动地震台阵观测项目不断开展。其中,川滇地区安装的由约 300 台宽频带地震仪组成的地震台阵支持下的“活动地块边界带动力过程与强震预测”探测项目和中国地震局地球物理研究所在首都圈地区进行的约 250 台地震仪台阵实验探测项目最具代表性,这对流动数字地震台阵数据的有效管理和自动处理提出了前所未有的迫切要求。文中介绍了可用于流动地震数据管理和自动预处理的 SeisComP 软件及其数据的归档结构,针对台阵仪器众多、参数配置繁琐和数据处理环节复杂的特点,给出了基于 SeisComP 软件包台阵数据的一种高效管理方案,整合了数据格式转换和参数修正等环节,实现了台阵数据的自动归档和地震信息的自动提取,可以提高数据的管理效率,对台阵地震信息的及时发布和波形数据共享提供了良好的科研条件。

关键词 地震台阵 数据管理 SeisComP REFTEK SeedLink ArcLink
中图分类号: P315.63 **文献标识码**:A **文章编号**:0253 - 4967(2009)02 - 0341 - 08

0 引言

地震台阵是现代地震学研究的一个有力工具,利用地震台阵较强的获取信号能力可以为我们研究地球内部的介质结构和动力学提供有效的数据质量保证。宽频带流动地震台阵观测技术在揭示地震形成的环境条件方面可以发挥重要作用,配合相应的数值模拟研究,宽频带流动地震台阵观测研究的结果将对地震形成机理的研究及地震预测发挥重要的作用,宽频带流动地震台阵日益成为地球动力学和构造物理学研究的重要手段之一(刘启元,2000)。

大规模流动地震观测采集的数据量会随着时间不断膨胀,地震台阵数据的传输汇集、海量存储和高效管理便成为台阵研究面临的首要问题。以中国地震局地质研究所为执行国家重点基础研究发展计划项目开展“活动地块边界带动力过程与强震预测”深部探测研究,在川滇地区安装架设的由 297 台宽频带地震仪组成的大规模流动地震台阵(刘启元等,2008)探测为例,其仪器规模与目前的国家测震台网数量可相比拟,台阵观测一年的原始数据产出量超过 4TB。这对流动数字地震台阵数据的有效管理和自动处理提出了前所未有的迫切要求。

目前数字测震台网的数据管理和处理已经比较成熟。其中以美国 IRIS 的 NetDC、欧洲测震台网的 SeisComP[①]、中国国家测震台网的数据管理系统为代表,已经实现了测震台网数据的汇集管理、存储共享和一定程度上的自动处理。在流动地震台阵观测数据管理方面,美国

〔收稿日期〕 2008 - 05 - 19 收稿,2008 - 07 - 30 改回。

* 通讯作者:陈九辉,cjh@sac.ac.cn。

① What is SeisComP, http://www.gfz-potsdam.de/geofon/seiscomp.

PASSCAL 数据中心采用了价格昂贵的 Antelop 软件包,欧洲则以仅限于数据归档和汇集功能的 NetDC 为主,并不包含地震数据处理功能。

流动数字地震仪器采集的数据分散存储在众多存储介质中,要对数据进行进一步的集中保管、处理和利用,就需要把所有数据按照统一的格式要求进行归档,并且能够方便高效地查询和提取数据,为数据的进一步处理提供良好的数据资源。地震台阵有其流动性、密集性和数据关联性的独特特点,数据管理上既要保证高效管理众多台阵仪器的相关参数,又要保证方便地修正、转换观测记录的数据格式并汇集到数据中心,地震台网的数据管理方案不能完全满足实际需要。SeisComP 是为适应台网管理的需要开发的,在管理台阵方面有不少地方满足不了实际需要,但其中的数据通讯、归档结构和处理模块等部分对台阵管理是通用的。在 SeisComP 基础上作者做了修改以适应台阵的处理要求,文章给出了基于 SeisComP 地震台阵的数据管理与发布共享方案。

1 SeisComP 软件

1.1 SeisComP 概述

SeisComP(Seismological Communication Processor)作为一种地震数据通讯处理系统的开源软件,起初是为德国 GEOFON 台站网络开发的,后又在 MEREDIAN 项目中做了进一步的扩展和加强。SeisComP 通过插件可以连接多种仪器和网络数据,从而汇集各个台站或数据中心的波形数据,实现不同数据中心或节点之间(近)实时波形数据的汇集和交换共享,为业务应用(如实时处理等)提供波形数据准实时服务。通过 SeisComP 也可实现台站部分元数据的传送和交换。该软件提供了 Mini - SEED 数据格式的支持,另一方面还包含了震相拾取、地震定位、波形图绘制和数据分析等数据常规处理功能的额外的软件包。SeisComP 对单台仪器的配置并不算复杂,但若对上百台仪器而言,相似的重复工作是低效的也是不宜采取的。

SeedLink 协议是 SeisComP 软件数据获取子包的核心部分,是数据(近)实时传输遵循的协议。在客户端和服务器端接口均支持 TCP/IP 协议和 PPP 协议。SeedLink 协议数据是以 512B 大小的 Mini - SEED 格式的数据包形式发送的,其中包括 8B 大小的 SeedLink 头段数据。SeedLink 支持不可靠连接,最近请求的数据包保留在主存里面可以高效连接,旧的数据包则被暂时存储在硬盘缓存里,主存和硬盘缓存大小均可由用户指定(Hanka *et al.*,2000)。

1.2 SeisComP 的参数配置

SeisComP 的参数由软件目录下 key 和 config 文件夹下的文件设置。可以直接修改这些 key 和 config 下的文件配置软件,但一般不建议这样做。所有参数的文件都是由模板文件根据交互的一些参数自动替换变量创建并全部重写的,之前的所有配置文件将被覆盖更新,欲对某个参数修改时可以先修改模板文件再进行配置。

模板文件中的参数部分均是通过命令交互配置的,对于有限几个台阵仪器来说,交互配置并不算繁琐,尤其是已预先配置好了针对各个处理模块包的参数文件。而对于多台特别是大型台阵仪器而言,这种参数配置会非常繁琐,这就需要一个适合需求的方便的配置解决方案。我们在 SeisComP 交互配置的基础上做了进一步的开发和改造,把台阵布设的仪器参数写到 1 个文件中,让 SeisComP 自动读取文件并实现绝大多数参数的自动配置,同时实现了对各个处理模块软件的自动配置,从而方便地配置了众多复杂的参数,使得服务器能够确认数据的来源,读入

数据并做归档，相应的数据处理也可随之进行。这样，绝大多数台阵相关的参数可以实现自动配置，从而方便了软件的参数管理，提高了台阵数据的管理效率。

1.3　地震数据归档结构

SeisComP 的数据归档结构定义为 SDS（SeisComP Data Structure），SDS 规定了波形数据文件的目录结构，这为数据服务器、数据分析包及其它需要获取分析数据文件的软件提供了数据的存储标准，可以方便地查询和使用。SDS 的结构是 < SDSdir > /Year/NET/STA/CHAN. TYPE/NET. STA. LOC. CHAN. TYPE. YEAR. DAY[①]，这种 SDS 的数据结构可以方便地根据时间、台站号和通道号进行较高效的查询。

2　台阵数据管理

2.1　数据的汇集

观测项目中的地震台阵全部采用由美国 Reftek 公司生产的 130 型号数字地震记录仪和英国 Guralp 公司生产的地震传感器组成。RFRETEK130 型号 24 位数字地震仪是一种高精度、宽频带、低能耗的三分量地震数据采集系统，是目前国际上较为先进的地震仪。仪器的多种通讯接口给我们的数据传输提供了多种支持模式，我们可以根据具体的条件构建台阵数据的传输方式。台阵仪器全部由相同性能参数的宽频带仪器组成，这为我们数据的处理带来了很大的便利，同时也会提高数据的处理精度。REF TEK130 系列的数据采集器（DAS）记录的数据格式是 PASSCAL 记录格式，目前还没有直接记录为 Mini – SEED 格式，但提供了相应的转换工具。数据的记录格式还可以转换成其他格式，包括 SEG – Y（Society of Exploration Geophysicists）、SUDS（Seismic Unified Data System）和 ASCII 格式。

台阵的数据是由布设在各个地点的仪器采集的，由于目前我们并不具备条件实现数据的实时传输，我们把所有台阵仪器的数据适时不定期地打包汇集上报到数据中心，根据 REFTEK 格式归档，并建立 NFS（Network File System）服务器，一方面作为采集数据的归档入口，另一方面为数据的后续处理提供数据接口。REFTEK 仪器公司提供了相应的数据操作软件。其中用到了 REFTEK 仪器公司提供的建立归档结构和写数据的 2 个软件工具。

此时建立的 REFTEK 格式的文件目录数据只是实验的源数据，为了利于数据的进一步使用，还需把数据格式进行转换，我们采取了转换为 Mini – SEED 的数据格式。另外为了统一管理，还需要对源数据的数据所属网络号、台站号进行修正。

2.2　数据的归档处理流程

若要把地震数据读入 SeedLink 服务器，首先需要把数据所属网络号、台站号配置到服务器端，一般只需运行配置脚本进行一次设置。然后我们从上面制作好的 NFS 服务器中获取数据，对数据的选取、格式转换、参数修正等数据处理部分作者编写了相应的处理脚本，并能自动顺序调用。从 REFTEK 实验数据 NFS 服务器读取数据时可以根据预先设定的数据起始时间和时间段来选定与获取数据，并能通过设置参数获取跨日期的连续数据。数据的格式转换是把获取的源数据转换成 Mini – SEED 格式，而参数修正则是用来修正 Mini – SEED 格式数据中头段台站标识码、台站网络标识码和数据通道号定义。经过脚本处理之后的数据通过命名管道被读入

① SeisComP Data Structure（SDS）definition，http：//www. gfz-potsdam. de/geofon/seiscomp/SDS. html.

SeedLink 服务器,使用专门针对 SeisComP SeedLink 服务器端的一个数据读入接口插件,可以对选取的数据以先进先出的模式传送到 SeedLink 服务器端。数据的归档是使用归档数据的一个客户端工具实现的,它从服务器端获取数据并以 SDS 目录格式进行归档。数据的归档处理流程如图 1 所示。

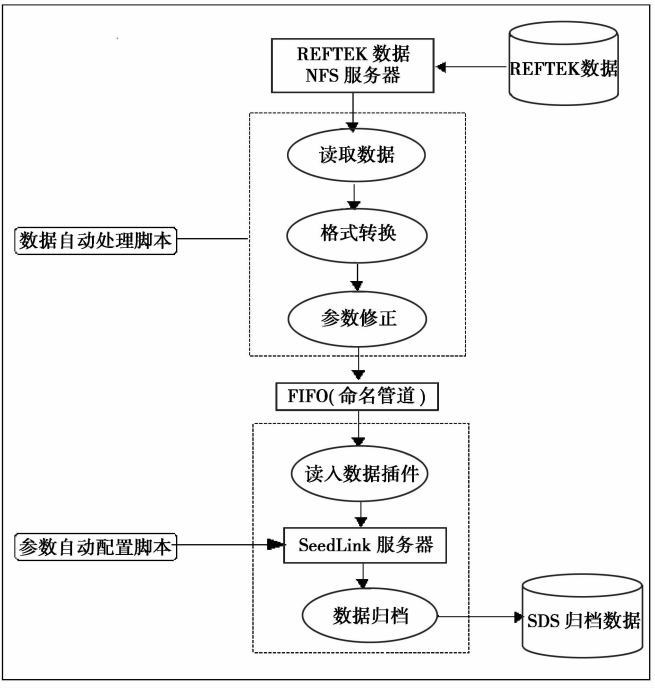

图 1 数据的归档处理流程图

Fig. 1 Flow chart of seismic data archiving and processing.

2.3 台阵数据定位

这里我们台阵数据的定位目的主要是针对近震进行触发,从而获得近震的大致分布。SeisComP 软件包中地震数据的定位实现过程如图 2 所示。震相拾取由自动拾取软件包通过读取数据的客户端读取服务器端的数据,按照 STA/LTA 事件检测算法把拾取的震相结果源源不断地追加在 1 个震相拾取的文件中。而地震定位由自动定位软件包相应地不断读取震相拾取文件的最后一页震相拾取信息,然后对符合预定条件的台站记录逐个进行定位,生成随震相拾取不断追加而刷新的事件文件。SeisComP 中自动定位的定位程序包是使用 LocSAT 定位软件和基于 IASP91 地球模型的走时表(Kennett *et al.*,1991)对拾取的地震进行定位的,LocSAT 在经典的 Geiger 定位法基础上增加使用了方位角数据,使用了到时与反方位角残差方程最小二乘反演迭代算法。我们从最终确定的事件文件中编写了触发脚本,用来自动生成地震目录及表达地震信息的网页。

这里实现的定位是一种全自动的快速定位,对网外的地震和远震的精确定位结果尚需精细地球模型的支持,对网内地震的定位结果仍具有一定的参考价值。我们选取了 25.5°~35°N,98°~106°E 区域内 2006 年 10 月到 2007 年 9 月时间段内台阵记录到的由 10 台仪器数据定位

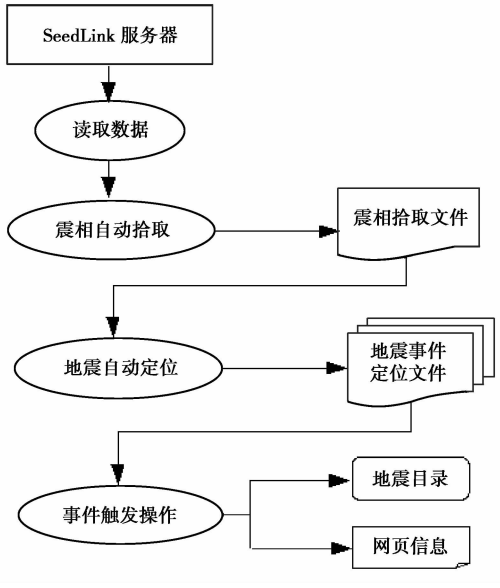

图 2　台阵数据定位过程图

Fig. 2　Process chart of event location using seismic array data.

的 181 个地震事件，如图 3 中的浅色点所示。而这与四川省地震台网 2007 年在该区域监测的 2.9 级以上的部分地震事件的定位结果分布粗略一致，如图 3 中的深色点所示。由于台网仪器布设的局限性，部分地区的地震并未监测到，如青藏高原东缘川滇地区西部断裂带的地震记录呈现空白，而布设的台阵对此有较多的地震记录，可见台阵的监测结果对台网监测是一个很好的完善和补充。台阵对该区域的定位结果分布与对该地区做过的精细重新定位的分布也粗略吻合（朱艾斓等，2005），但此精细定位是在台网监测的基础上做出的，部分地区地震记录仍是空白的。

2.4　台阵地震数据发布

在 SeisComP 中使用了 ArcLink 软件包实现地震数据的发布。ArcLink 包使用的数据传输协议是 ArcLink 协议，也是基于 TCP 协议开发的，可以用来请求归档的非实时波形数据及相关的元数据，是一种适合分布归档数据的存取协议。该协议相对比 SeedLink 协议简单，但通讯处理效率不逊于 SeedLink 的实时通讯协议，而且 ArcLink 可以传输完整的元数据。ArcLink 替代了 SeedLink 通过许多插件支持的不同数据采集系统的做法，ArcLink 请求处理模块支持不同的数据归档结构作为独立的数据源，ArcLink 数据查询的流程如图 4 所示。ArcLink 最初的原型是基于 TCP/IP 处理 WebDC 数据中心间的数据通讯协议，请求处理目前仅支持 SDS（Seiscomp Data Structure）归档结构，其它的归档结构需要单独开发相应的插件。

ArcLink 请求格式类似于 NetDC、AutoDRM 和 breq_fast，请求处理状态可以实时查核，当被请求的数据准备好后就可以下载了。请求提交、状态检查、数据下载都是通过 ArcLink TCP/IP 连接的，而非 e－mail，FTP 等形式。所有的数据都并不需要保留在 RAM 内，所以数据产品可以非常大。数据产品会一直保存在服务器端，直到被客户端请求删除。当发生网络错误时可以重新下载指定段的数据。ArcLink 具有路由能力，用户并不需预先知道需要从哪个服务器请求数

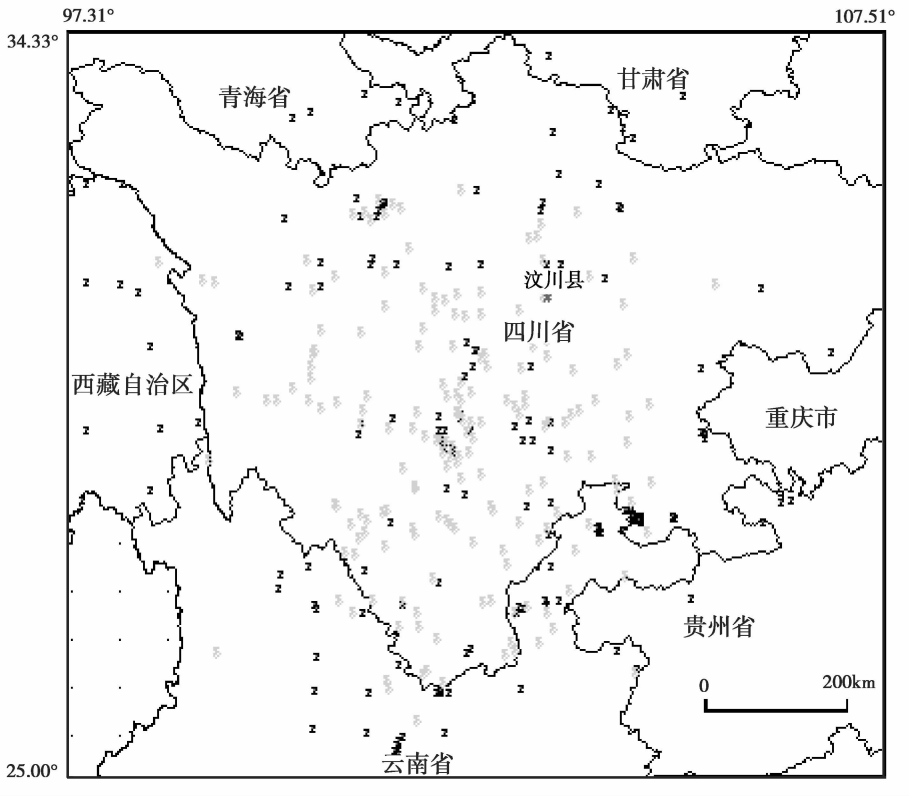

图 3 软件快速自动定位及四川省台网监测的定位结果

Fig. 3 Results of quick auto-location vs. the location result of Sichun seismic network.

浅色点为自动定位结果,深色点为台网监测结果

据。TCP 连接的传输时间不会影响处理时间,获取的台站元数据与通过 FTP/HTTP 方式下载一个压缩的 XML 文件类似,这通常比基于 CORBA 的协议(像 DHI①)快得多。

3 结论与讨论

基于 SeisComP 开源软件包可以根据数据处理要求方便地进行扩展和修改,针对管理台阵数据的要求在软件配置及数据传输、归档和发布方面我们基本上实现了自动作业,较快地实现了台阵数据的归档和初步定位,根据初步的定位结果,可以为数据的后续精细处理在数据选取和处理方面提供参考。由于密集台阵记录的精敏性,台阵的地震目录可以作为台网记录的一个很好的补充。尽管目前仍未实现全部仪器数据的实时传输,但当条件成熟时,我们可以方便地调整部署处理实时的数据,并能协调台网进行监测。

上面介绍的台阵数据管理方式是基于单纯的文件目录进行数据存储和读取的,由于目录格式是根据数据属性标识建立的,读写和查询效率已能较好地满足使用要求,国际上有多个大型地震数据中心也是基于 SDS 这种文件目录结构进行管理和维护的。由于数据库形式的数据存

① DHI(Data Handling Interface) Main Page, http://www.iris.edu/DHI.

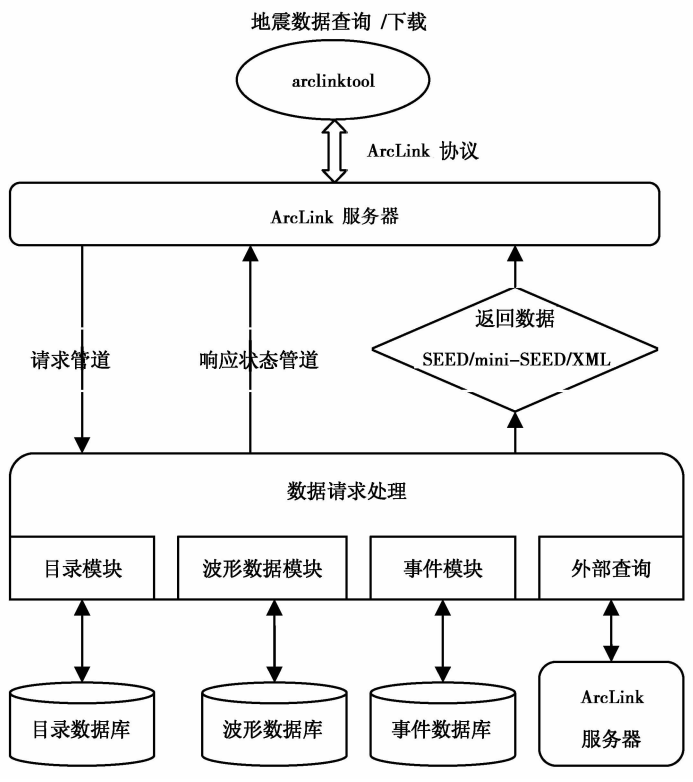

图 4　ArcLink 数据查询流程图

Fig. 4　Flow chart of data query by ArcLink.

储在管理上、使用上和维护上会有更大的优势，下一步我们会采取数据库形式的数据存储管理。

参　考　文　献

刘启元. 2000. 高分辨率地震成像研究：21 世纪地震学发展的一个重要趋势[J]. 国际地震动态，(4)：9—11.

LIU Qi-yuan. 2000. Research on high resolution seismic imaging: an important development trend of seismology in the 21st century [J]. Recent Developments in World Seismology, (4): 9—11 (in Chinese).

刘启元，陈九辉，李顺成，等. 2008. 汶川 M_S 8.0 地震：川西流动地震台阵观测数据的初步分析[J]. 地震地质，30(3)：584—596.

LIU Qi-yuan, CHEN Jiu-hui, LI Shun-cheng, *et al*. 2008. The M_S8.0 Wenchuan earthquake: Preliminary results from the western Sichuan mobile seismic array observations [J]. Seismology and Geology, 30(3): 584—596 (in Chinese).

朱艾斓，徐锡伟，周永胜，等. 2005. 川西地区小震重新定位及其活动构造意义[J]. 地球物理学报，48(3)：629—636.

ZHU Ai-lan, XU xi-wei, ZHOU Yong-sheng, *et al*. 2005. Relocation of small earthquakes in western Sichuan, China and its implications for active tectonics [J]. Chinese Journal of Geophysics, 48(3): 629—636 (in Chinese).

Kennett B L N, Engdahl E R. 1991. Travel times for global earthquake location and phase identification [J]. Geophys J Int, 105: 429—465.

Hanka W, Heinloo A, Jaeckel K H. 2000. Networked seismographs: GEOFON real-time data distribution [J]. ORFEUS Newsletter, 2(3): 24.

THE DATA MANAGEMENT OF LARGE-SCALE SEISMIC ARRAY BASED ON SEISCOMP

WU Ming-shu LIU Zhi CHEN Jiu-hui

(*Institute of Geology, China Earthquake Administration, Beijing* 100029, *China*)

Abstract

With the establishment of China Seismic Array for Scientific Research, large-scale mobile seismic array observation projects have been carried out. One of the observation projects, named Dynamical Process of Active Block Boundary Belts and Strong Earthquake Prediction, consists of about 300 sets of broadband seismic instruments in Sichuan and Yunnan. And another seismic array for experimental survey equipped with about 250 sets of seismic instruments has been in operation in the capital area of China. As representative projects, an unprecedented urgent request was raised on the effective management and automatic processing for mobile digital seismic array data. The data acquired from mobile digital seismic instruments is stored in many scattered storage medium. For better storing, processing and use of the data, all the data should be archived in accordance with the requirements of a unified format. Seismic array data has unique characteristics of liquidity, density and relevance. It is necessary to ensure efficient management of many related parameters of array equipment. Ensuring an easy way to amend and change the observation data and pooling them to data center are also needed. Data management solutions to seismic network can not fully meet the actual needs. In this paper, SeisComp software which can be used for seismic array data management and preprocessing is introduced together with its data archive characteristic. A high efficient algorithm for seismic array data management is proposed based on SeisComP packages, which conforms data format conversion and parameter modification and realizes the automatic data archiving and pick-up. The method improves the efficiency of data management, and meanwhile, facilitates the data information distribution and waveform data sharing.

Key words seismic array, data management, SeisComP, REFTEK, SeedLink, ArcLink

〔**作者简介**〕 吴明恕,男,1983 年出生,2008 年毕业于中国地震局地质研究所,获硕士学位,助理馆员,从事信息共享技术研究,电话:010 - 62004028, E - mail: mingshu5@ gmail. com。