Zuzlewski Stéphane November 1998

Waveforms Schema

Volume IV Version 1.2

Table of Contents

1.- Introduction :

This paper is about the design of the NCEDC schema for the waveforms. This schema was designed to be able to store different types of waveforms as well as their associations with events .

There are 3 relations in the schema. The information stored in these relations includes:

- Waveform information.
- Association of waveforms and events.
- Files information.

The remainder of this volume consists of three chapters:

Chapter 2	*Database structure* Each relation is defined.
Chapter 3	*Database relations* Each relation is described to identify the key fields and the links among the relations.
Chapter 4	*Database attributes* Each attribute is described.

2.- Database structure :

This chapter defines the physical structure of each table, as it exists within the DBMS data dictionary. The name of the relation appears in **bold** print at the top left and a description of the table at the top right. **Bold** attributes represent primary keys, *italic* attributes represent foreign keys, and underlined attributes stand for CSS attributes.

Waveform		Summary information on a waveform
Attribute	***Type***	***Description***
wfid	int	Waveform identifier
net	char(8)	Network code
sta	char(6)	Station code
auth	char(15)	Source/originator
subsource	char(8)	Subsystem source
channel	char(3)	Channel name
channelsrc	char(8)	Domain of channel name
seedchan	char(3)	SEED channel name
location	char(2)	SEED location
archive	char(8)	Location of seismogram archive
datetime_on	float	Start time
datetime_off	float	End time
samprate	float	Sampling rate (samples/sec)
wavetype	char(1)	Type of waveform (cont., triggered, …)
fileid	int	Filename identifier
foff	int	Pointer to simple data
nbytes	int	Number of bytes of data portion in trace
traceoff	int	Pointer to the beginning of the trace
tracelen	int	Length in bytes of the entire trace
status	char(1)	Status (Empty, Temp, Archive)
wave_fmt	int	Waveform format type
format_id	int	Format of data (Steim-1, …)
wordorder	int	Word order
recordsize	int	Record size for data
locevid	int	Local event identifier
lddate	date	Load date

AssocWaE		Data associating waveforms with events
Attribute	***Type***	***Description***
wfid	int	Waveform identifier
evid	int	Event identifier
datetime_on	float	Start time of waveform association
datetime_off	float	End time of waveform association
lddate	date	Load date

Filename		Summary information on a waveform filenames
Attribute	***Type***	***Description***
fileid	int	Filename identifier
dfile	char(32)	Data file
datetime_on	float	Start time
datetime_off	float	End time
nbytes	int	File size
lddate	date	Load date

3.- Database relations :

This chapter describes the DBMS relations that comprise the version 1.2 schema. The information given here, along with that in Chapter 4, *Database attributes*, constitutes the data dictionary. There's an entry for each relation. Within the entry, the relation's name appears first, followed by a list of its attributes. A brief description completes the entry.

Name:	**Waveform**	
Keys:		
	Primary:	*wfid*
	Foreign:	*fileid*
Data:	*sta, net, auth, subsource, channel, channelsrc, seedchan, location, archive, date_time_on, datetime_off, dt, wavetype, foff, nbytes, traceoff, tracelen, status, wave_fmt, format_id, wordorder, recordsize, locevid, lddate*	
Description:	Waveform header file and descriptive information. This relation provides a pointer (or index) to waveforms.	

Name:	**AssocWaE**	
Keys:		
	Primary:	*wavid, evid*
	Foreign:	
Data:	*datetime_on, datetime_off, lddate*	
Description:	Waveform mapping file. This relation links event identifiers to waveform identifiers.	

Name:	**Filename**	
Keys:		
	Primary:	*fileid*
	Foreign:	
Data:	*dfile, datetime_on, datetime_off, nbytes, lddate*	
Description:	Filename information. This relation provides information about the physical file where the seismogram is stored.	

4.- Database attributes :

This chapter describes each of the attributes used in the version 1.2 schema. Descriptions of the relations are found in Chapter 3, *Database relations*.

Name:	*archive*
Relations:	**Waveform**
Description:	This field specifies the place where the seismogram is archived.
Datatype:	***char(8)***
NULL Value:	***No***
Range:	Any free-format string up to 8 characters

Name:	*auth*
Relations:	**Waveform**
Description:	The *auth* field specifies the source of the information. This may be a network identifier (FDSN or otherwise), or any other string indicating the source of the information in the record.
Datatype:	***char(15)***
NULL Value:	***No***
Range:	Any free-format string up to 15 characters

Name:	*channel*
Relations:	**Waveform**
Description:	Channel name.
Datatype:	***char(3)***
NULL Value:	***Yes***
Range:	Any valid channel name, which may be network-specific. Networks using SEED channel names should use the same value for *channel* and *seedchan*

Name:	*channelsrc*
Relations:	**Waveform**
Description:	Domain for channel. This specifies what naming convention is used for the channel name (i.e. SEED, USGS, etc.).
Datatype:	***char(8)***
NULL Value:	***Yes***
Range:	Any free-format string up to 8 characters long

Name:	*datetime_off*
Relations:	**Waveform, AssocWaE, Filename**
Description:	Time of last datum. In **Filename**, this attribute is the time of the last sample in the waveform file.
Datatype:	***float***
NULL Value:	**No**
Range:	Any valid float. *datetime_off* = 0.0 on **January 01, 1970** at **00:00:00.0000**
Units:	***Seconds***

Name:	*datetime_on*
Relations:	**Waveform, AssocWaE, Filename**
Description:	Time of first datum. In **Filename**, this attribute is the time of the first sample in the waveform file.
Datatype:	***float***
NULL Value:	**No**
Range:	Any valid float. *datetime_on* = 0.0 on **January 01, 1970** at **00:00:00.0000**
Units:	***Seconds***

Name:	<u>*dfile*</u>
Relations:	**Filename**
Description:	Data file. This is the file name of a disk-based waveform file.
Datatype:	***char(32)***
NULL value:	***No***
Range:	Any free-format string up to 2 characters long

Name:	<u>*evid*</u>
Relations:	**AssocWaE**
Description:	Event identifier. Each event is assigned a unique integer which identifies it in a database. It is possible for several records in the **Origin** relation to have the same *evid*. This indicates there are several opinions about the location of the event. The *evid* does not have to be identical to the local network's event identifier, which is stored in the **Origin** table. This allows to merge data for a single event from multiple sources without insisting that database *evid* is identical to the source's local event identifier.
Datatype:	***int***
NULL Value:	***No***
Range:	*evid* > 0

Name:	*fileid*
Relations:	**Waveform, Filename**
Description:	Filename identifier. The key field is a unique identifier for a physical file on a disk.
Datatype:	***int***
NULL Value:	***No***
Range:	*fileid* > 0

Name:	*foff*
Relations:	**Waveform**
Description:	File offset. This is the byte offset of a waveform segment within a data file.
Datatype:	***int***
NULL Value:	***Yes***
Range:	$foff \geq 0$

Name: *format_id*

Relations: **D_Format**

Description: This attribute denotes a code indicating the encoding format. This number is assigned by the FDSN Data Exchange Working Group.

Datatype: ***int***

NULL Value: ***Yes***

Range: *format_id* ≥ 1

CODES 1-9	GENERAL
1	16 bit integers
2	24 bit integers
3	32 bit integers
4	IEEE floating point
5	IEEE double precision floating point
CODES 10-29	FDSN Networks
10	STEIM (1) Compression
11	STEIM (2) Compression
12	GEOSCOPE Multiplexed Format 24 bit integer
13	GEOSCOPE Multiplexed Format 16 bit gain ranged, 3 bit exponent
14	GEOSCOPE Multiplexed Format 16 bit gain ranged, 4 bit exponent
15	US National Network compression
16	CDSN 16 bit gain ranged
17	Graefenberg 16 bit gain ranged
18	IPG – Strasbourg 16 bit gain ranged
CODES 30-39	OLDER NETWORKS
30	SRO Format
31	HGLP Format
32	DWWSSN Gain Ranged Format
33	RSTN 16 bit gain ranged

Name:	<u>*lddate*</u>
Relations:	**Waveform, AssocWaE, Filename**
Description:	Load date. Date and time that the record was created or last modified, in Oracle ***date*** datatype.
Datatype:	***date***
NULL Value:	***No***
Range:	Any valid date between **January 01, 4712 BC** and **January 01, 4712 AD**
Units:	***YYYY/MM/DD HH24:MI:SS***

Name:	*location*
Relations:	**Waveform**
Description:	Describes the individual sites on an array station, operated by the same network operator.
Datatype:	***char(2)***
NULL Value:	***Yes***
Range:	Any free-format string up to 2 characters long

Name:	*locevid*
Relations:	**Waveform**
Description:	Local event identifier. This attribute describes the original event identifier of the event. In combination with *auth* and *subsource*, it can be used to uniquely identify the original event identifier from that source. For example, waveforms from the original source may be associated with this identifier.
Datatype:	***int***
NULL Value:	***Yes***
Range:	Any valid integer

Name:	*nbytes*
Relations:	**Waveform, Filename**
Description:	Number of bytes. This quantity is the number of bytes in a waveform segment or a file.
Datatype:	***int***
NULL Value:	***Yes***
Range:	*nbytes* ≥ 0

Name:	*net*
Relations:	**Waveform**
Description:	Unique network identifier. This character string is the name of a seismic network.
Datatype:	***char(8)***
NULL Value:	***No***
Range:	Any free-format string up to 8 characters long

Name:	*recordsize*
Relations:	**Waveform**
Description:	Record size for data. For MiniSEED, the value is the MiniSEED record size, eg. 512, 4096, etc.
Datatype:	***int***
NULL Value:	***Yes***
Range:	*recordsize* ≥ 0 0 = Not applicable

Name:	*samprate*
Relations:	**Waveform**
Description:	Sampling rate. This attribute is the sample rate in samples/second.
Datatype:	***float***
NULL Value:	***No***
Range:	*samprate* > 0.0
Units:	***1/seconds***

Name: *seedchan*

Relations: **Waveform**

Description: SEED channel name. The first character denotes the band code, the second for the instrument code and the third represents the component code.

Datatype: ***char(3)***

NULL Value: ***No***

Range: *seedchan* ∈ {E | S | H | B | M | L | V | U | R} ∪
{A | B | D | F | G | H | I | K | L | M | P | R | S | V | T | W} ∪
{Z | N | E | A | B | C | T | R | 1 | 2 | 3 | U | V | W}

Band code:

	Sample rate (Hz)	Corner period (sec)
E = extremely short period	≥ 80	< 10
S = short period	≥ 10 to < 80	< 10
H = high broad band	≥ 80	≥ 10
B = broad band	≥ 10 to < 80	≥ 10
M = mid period	> 1 to < 10	
L = long period	~ 1	
V = very long period	~ 0.1	
U = ultra long period	~ 0.01	
R = extremely long period	~ 0.001	

Instrument code:

A = tiltmeter	B = creepmeter
D = pressure	F = magnetometer
G = gravimeter/accelerometer	H = high gain seismometer
I = humidity	K = temperature
L = low gain seismometer	M = mass position seismometer
P = geophone	R = rainfall
S = linear strainmeter	V = volumetric strain
T = tide meter	W = wind

Component code:

N = north-south	E = east-west
T = transverse	R = radial

Z = vertical
A, B, C = triaxial
1, 2, 3 = orthogonal components but non-traditional orientations
U, V, W = optional components

Name:	<u>*sta*</u>
Relations:	**Waveform**
Description:	Station code. This is the common code-name of a seismic observatory. Generally only three or four characters are used.
Datatype:	***char(6)***
NULL Value:	***No***
Range:	Any free-format string up to 6 characters long

Name:	*status*
Relations:	**Waveform**
Description:	This attribute denotes the status of the waveform in the file.
Datatype:	***char(1)***
NULL Value:	***No***
Range:	*status* $\in$ { E, T, A } E = empty T = temporary A = archive

Name:	*subsource*
Relations:	**Waveform**
Description:	A second identifier to specify a unique source within the domain specified by *auth*.
Datatype:	***char(8)***
NULL Value:	***Yes***
Range:	Any free-format string up to 8 characters

Name:	*tracelen*
Relations:	**Waveform**
Description:	Number of bytes. This quantity represents the length of the entire trace.
Datatype:	***int***
NULL Value:	***Yes***
Range:	*tracelen* ≥ 0

Name:	*traceoff*
Relations:	**Waveform**
Description:	File offset. This is the byte offset of the entire trace.
Datatype:	***int***
NULL Value:	***Yes***
Range:	*traceoff* ≥ 0

Name:	*wavetype*
Relations:	**Waveform**
Description:	This attribute denotes the type of waveform.
Datatype:	***char(1)***
NULL Value:	***Yes***
Range:	*wavetype* ∈ { C, T } C = continuous T = triggered

Name:	*wave_fmt*
Relations:	**Waveform**
Description:	Waveform type.
Datatype:	***int***
NULL Value:	***Yes***
Range:	*wave_fmt* ≥ 1 1 = uncompressed binary data 2 = MiniSEED 3 = PASSCAL SEGY 4 = UCB SEGY …

Name:	*wfid*
Relations:	**Waveform, AssocWaE**
Description:	Waveform identifier. The key field is a unique identifier for a segment of digital waveform data.
Datatype:	***int***
NULL Value:	***No***
Range:	*wfid* > 0

Name:	*wordorder*
Relations:	**Waveform**
Description:	Word order.
Datatype:	***int***
NULL value:	***Yes***
Range:	*wordorder* ≥ 0 0 = little endian (VAX, INTEL) 1 = big endian (SPARC, 68000)