SAND2016-0093R

IDC RE-ENGINEERING REPORT

SAND20XX-XXXX
Unlimited Release
January 2016

IDC Re-Engineering Phase 2 Iteration E1 Data Model to IDC Schema Mapping

Version 1.0

Ben Hamlet, Mark Montoya, Rudy Sandoval, James Vickers

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

SAND20XX-XXXX
January 2016

Error! Reference source not found.

Error! Reference source not found.
Sandia National Laboratories
P.O. Box 5800
Albuquerque, New Mexico 87185-MS0414

ABSTRACT

This initial draft document contains formative data model content for select areas of Re-Engineering Phase 2 IDC System. The purpose of this document is to facilitate discussion among the stakeholders. It is not intended as a definitive proposal.

This page intentionally left blank.

REVISIONS

Version	Date	Author/Team	Revision Description	Authorized by
0.1	12/2015	IDC Re-engineering Team		M. Harris

Table of Contents

1. Notation:

- A Data Model Class and Attribute that is highlighted green means we have values in the Data Model that are currently not represented in the schema.
- A Data Model Class and Attribute that is TBD means there is a table column with no equivalent value in the Data Model. The decision to add these to the Data Model or remove them/leave as N/A values in the table is to-be-determined.

2. Station Mappings

2.1. site

IDC Database Schema	Data Model	
site Table Column	**Class**	**Attribute**
sta	Station, Site	name, name
ondate	Station, Site	on time, on time
offdate	Station, Site	off time, off time
lat	Station, Site	lat, lat
lon	Station, Site	lon, lon
elev	Station, Site	elevation, elevation
staname	TBD	TBD
statype	Station	station type
refsta	Site	station membership (reference)
dnorth	TBD	TBD
deast	TBD	TBD
lddate	TBD	TBD
	Site	location code

Notes: We can fit the Station and Site classes into the site table, but we lose the clean distinction as well as the relationship between the two entities. We also lose the ability for a site to belong to multiple stations by using only refsta. Ideally, Station and Site would exist as separate tables related by a Station Membership table, in order to allow the intended many-many relationship

2.2. station_alias

IDC Database Schema	Data Model	
Table Column	**Class**	**Attribute**
	Alias	parent station
	Alias	name
	Alias	on time
	Alias	off time

Notes: The Alias class does not map to a current CSS table.

2.3. network

IDC Database Schema	Data Model	
network Table Column	**Class**	**Attribute**
net	Network	name
netname	TBD	TBD
nettype	TBD	TBD
auth	TBD	TBD
commid	TBD	TBD
lddate	TBD	TBD
	Network	monitoring org
	Network	on time
	Network	off time

Notes: Network provenance information, such as on time and off time are not currently captured in CSS

2.4. participation

IDC Database Schema	Data Model	
participation Table Column	**Class**	**Attribute**
net	Affiliation	network (reference)
sta	Affiliation	station (reference)
begin_date	Affiliation	on time
end_date	Affiliation	off time
lddate	TBD	TBD
	Affiliation	author

Notes: The participation table is similar to the affiliation table, but contains more columns such as begin_date and end_date, which are similar to the desired ontime, offtime attributes.

2.5. sitechan

IDC Database Schema	Data Model	
sitechan Table Column	**Class**	**Attribute**
sta	Raw Channel	site (reference)
chan	Channel	name
ondate	Raw Channel	on time
chanid	Raw Channel	instrument (reference)
offdate	Raw Channel	off time
ctype	Raw Channel	channel type
edepth	Raw Channel	depth
hang	Channel	horizontal angle
vang	Channel	vertical angle
descrip	TBD	TBD
lddate	TBD	TBD
	Raw Channel	location code
	Channel	lat
	Channel	lon
	Channel	elevation
	Derived Channel	signal operation (reference)

Notes: Raw Channel and Derived Channel are two distinct classes both extending a base Channel class in the Data Model. Representing these entities in the sitechan table is possible, but leads to some redundant information and loss of clarity.

2.6. instrument

IDC Database Schema	Data Model	
instrument Table Column	**Class**	**Attribute**
inid	TBD	TBD
insname	TBD	TBD
instype	TBD	TBD
band	TBD	TBD
digital	TBD	TBD
samprate	TBD	TBD
ncalib	Instrument	nominal cal ratio
ncalper	Instrument	nominal cal per
dir	TBD	TBD
dfile	TBD	TBD
rsptype	TBD	TBD
lddate	TBD	TBD
	Instrument	raw channel (reference)
	Instrument	manufacturer
	Instrument	model
	Instrument	serial number
	Instrument	on time
	Instrument	off time
	Instrument	nominal tshift
	Instrument	nominal response

2.7. sensor

IDC Database Schema	Data Model	
sensor Table Column	**Class**	**Attribute**
sta	TBD	TBD
chan	Calibration	raw channel (reference)
time	Calibration	calibration time
endtime	TBD	TBD
inid	TBD	TBD
chanid	TBD	TBD
jdate	TBD	TBD
calratio	Calibration	cal factor
calper	Calibration	cal per
tshift	Calibration	tshift
instant	TBD	TBD
lddate	TBD	TBD
	Calibration	duration
	Calibration	response data

3. Signal Detection Mappings

3.1. amplitude

IDC Database Schema	Data Model	
amplitude Table Column	**Class**	**Attribute**
ampid	-	-
arid	Signal Detection Hypothesis Feature Measurement	parent Signal Detection Hypothesis
parid	TBD	TBD
chan	TBD	TBD
amp	Signal Detection Hypothesis Feature Measurement	measurement value
per	TBD	TBD
snr	TBD	TBD
amptime	TBD	TBD
start_time	Signal Detection Hypothesis Feature Measurement	calculation window
duration	Signal Detection Hypothesis Feature Measurement	calculation window
bandw	TBD	TBD
amptype	Signal Detection Hypothesis Feature Measurement	feature type
units	Signal Detection Hypothesis Feature Measurement	units
clip	TBD	TBD
inarrival	TBD	TBD
auth	Signal Detection Hypothesis Feature Measurement	author
lddate	TBD	TBD
uncertainty	Signal Detection Hypothesis Feature Measurement	uncertainty
	Signal Detection Hypothesis Feature Measurement	algorithm

Notes: This table stores the Signal Detection Hypothesis Feature Measurement class in the amplitude table of CSS. Although in CSS this table specifically stores amplitude measurements, here it is being used in a general way to store various measurements.

3.2. arrival

IDC Database Schema	Data Model	
arrival Table Column	**Class**	**Attribute**
sta	TBD	TBD
time	TBD	TBD
arid	TBD	TBD
jdate	TBD	TBD
stassid	TBD	TBD
chanid	TBD	TBD
chan	TBD	TBD
iphase	Signal Detection Hypothesis	phase
stype	TBD	TBD
deltim	TBD	TBD
azimuth	TBD	TBD
delaz	TBD	TBD
slow	TBD	TBD
delslo	TBD	TBD
ema	TBD	TBD
rect	TBD	TBD
amp	TBD	TBD
per	TBD	TBD
logat	TBD	TBD
clip	TBD	TBD
fm	TBD	TBD
snr	TBD	TBD
qual	TBD	TBD
auth	Signal Detection Hypothesis	author
commid	TBD	TBD
lddate	TBD	TBD
	Signal Detection Hypothesis	processing stage
	Signal Detection Hypothesis	rejected
	Signal Detection Hypothesis	parent Signal Detection

Notes: This table stores the Signal Detection Hypothesis class in the arrival table of CSS. The Data Model contains the Signal Detection class which aggregates the Signal

Detection Hypothesis class; this structure is not present in CSS. Some of the other information in this this table is duplicated from amplitude; however, repeating it here is perhaps confusing when similar (or the same) information is presented in the amplitude table.

3.3. sigdet

Column	Class	Attribute
	Signal Detection	monitoring organization

Notes: This table stores information for the Signal Detection class of the data model. It is not present in CSS. Alternatively, this information could be considered as a missing column of the arrival table.

4. Event Class Mappings

4.1. event

IDC Database Schema	Data Model	
event Table Column	**Class**	**Attribute**
evid	-	-
evname	TBD	TBD
prefor	Event	preferred event hypothesis
auth	TBD	TBD
commid	TBD	TBD
lddate	TBD	TBD
	Event	monitoring organization

Notes: The Data Model class structure allows an Event to have many Event Hypotheses and allows each Event Hypothesis to have many Location Solutions. CSS does not support this concept because the origin table combines the information from both the Event Hypothesis and Location Solution classes. Because of this, the event table's foreign key to the origin table refers to both a preferred Event Hypothesis as well as a preferred Location Solution. This is different than the Data Model which represents both the preferred Event Hypothesis of an Event as well as the preferred Location Solution of an Event Hypothesis.

4.2. origin

IDC Database Schema	Data Model	
origin Table Column	**Class**	**Attribute**
lat	Location Solution	lat
lon	Location Solution	lon
depth	Location Solution	depth
time	Location Solution	time
orid	TBD	TBD
evid	Event Hypothesis	parent event
jdate	TBD	TBD
nass	TBD	TBD
ndef	TBD	TBD
ndp	TBD	TBD
grn	TBD	TBD
srn	TBD	TBD
etype	Event Hypothesis	source type
depdp	TBD	TBD
dtype	TBD	TBD
mb	Magnitude Solution	mag
mbid	TBD	TBD
ms	Magnitude Solution	mag
msid	TBD	TBD
ml	Magnitude Solution	mag
mlid	TBD	TBD
algorithm	Location Solution	algorithm
auth	Location Solution	author
commid	TBD	TBD
lddate	TBD	TBD
	Location Solution	restraints
	Location Solution	manual solution

Notes: The Data Model class structure allows an Event to have many Event Hypotheses and allows each Event Hypothesis to have many Location Solutions. CSS does not support this concept because the origin table combines the information from both the Event Hypothesis and Location Solution classes. Because of this, using CSS to represent multiple Location Solutions of an Event Hypothesis requires creating multiple origin rows. Since CSS uses the assoc table to associate origins to arrivals, each origin row representing Location Solutions of a common Event Hypothesis must be related via assoc rows to the same set of arrival rows. This ensures all of the Location Solutions for an Event Hypothesis are related to the same Signal Detection Hypotheses.

4.3. origin_ext

Column	Class	Attribute
	Event Hypothesis	author
	Event Hypothesis	algorithm
	Event Hypothesis	processing stage
	Event Hypothesis	preferred for stage
	Event Hypothesis	rejected
	Event Hypothesis	preferred location solution
	Event Hypothesis	parent event hypothesis

Notes: origin_ext represents additional Event Hypothesis information stored in the Data Model that is not stored by the CSS 3.0 tables. This information could be stored as new rows in the origin table, but since CSS uses the origin table to represent both multiple Event Hypotheses of an Event and multiple Location Solutions of an Event Hypothesis that would result in the information being duplicated for each of an Event Hypothesis' Location Solutions.

4.4. assoc

IDC Database Schema	Data Model	
assoc Table Column	**Class**	**Attribute**
arid	Association	signal detection hypothesis
orid	Association	event hypothesis
sta	TBD	TBD
phase	Signal Detection Hypothesis	phase
belief	TBD	TBD
delta	TBD	TBD
seaz	TBD	TBD
esaz	TBD	TBD
timeres	Prediction Information	residual
timedef	Location Defining Behavior	defining
	Prediction Information	weight
	Prediction Information	uncertainty
	Prediction Information	predictor
azres	Prediction Information	residual
azdef	Location Defining Behavior	defining
	Prediction Information	weight
	Prediction Information	uncertainty
	Prediction Information	predictor
slores	Prediction Information	residual
slodef	Location Defining Behavior	defining
	Prediction Information	weight
	Prediction Information	uncertainty
	Prediction Information	predictor
emares	TBD	TBD
wgt	TBD	TBD
vmodel	TBD	TBD
commid	TBD	TBD
lddate	TBD	TBD
	Association	rejected
	Association	author
	Association	algorithm

Notes: CSS stores a single factor, wgt, to represent an arrival's weighting in a location solution. CSS also stores a reference to a single velocity model used to predict the arrival's travel time, azimuth, and slowness during location calculations. The Prediction Information class in the Data Model extends this concept by representing separate weights and earth models for each feature measurement type (travel time, azimuth, slowness, etc.).

Data Model classes do not currently contain attributes corresponding to the assoc columns delta, seaz, and esaz since this information is readily derived from the event location and the station location. If appropriate, the Data Model could be extended to have attributes holding these values.

4.5. assoc_ext

Column	Class	Attribute
	Prediction Information	residual
	Prediction Information	uncertainty
	Prediction Information	weight
	Prediction Information	predictor
	Prediction Information	signal detection hypotheses feature measurement

Notes: assoc_ext represents additional feature predictions beyond travel time, azimuth, and slowness predictions stored in the CSS 3.0 assoc table.

4.6. origerr

IDC Database Schema	Data Model	
origerr Table Column	**Class**	**Attribute**
orid	TBD	TBD
sxx, syy, szz, stt, sxy, sxz, syx, stx, sty, stz	Location Uncertainty	4D covariance matrix
sdobs	TBD	TBD
smajax	Ellipse	semi major axis
sminax	Ellipse	semi minor axis
strike	Ellipse	major axis azimuth
sdepth	Ellipse	depth uncertainty
stime	Ellipse	time uncertainty
conf	Ellipse	confidence level
commid	TBD	TBD
lddate	TBD	TBD
	Ellipse	scaling factor type
	Ellipsoid	semi intermediate axis
	Ellipsoid	major axis plunge
	Ellipsoid	major axis rotation

Notes: CSS 3.0 uses orid to associate Location Uncertainty information stored in origerr with the Location Solution information stored in origin. The Data Model classes allow multiple Location Uncertainties for each Location Solution. Since CSS 3.0 uses orid as

origerr's primary key and as the link to origin, CSS 3.0 cannot represent multiple origerrs for a single origin unless a new primary key is defined for origerr.
If origerr is used to store multiple Location Uncertainties for a single origin then each of the origerr rows must store the same values for sdobs and the covariance matrix (i.e. sxx, syy, etc.). This is because each Ellipse and Ellipsoid for that origin's location are computed from the same distribution.

5. Magnitude Mappings

5.1. stamag

IDC Database Schema	Data Model	
stamag Table Column	**Class**	**Attribute**
magid	TBD	TBD
ampid	Station Magnitude Solution	signal detection hypothesis feature measurement
sta	TBD	TBD
arid	TBD	TBD
orid	Magnitude Solution	parent location solution
evid	TBD	TBD
phase	TBD	TBD
delta	TBD	TBD
magtype	Magnitude Solution	mag type
magnitude	Magnitude Solution	mag
uncertainty	Magnitude Solution	uncertainty
magres	TBD	TBD
magdef	Magnitude Defining Behavior	defining
mmodel	TBD	TBD
auth	Magnitude Solution	author
commid	TBD	TBD
lddate	TBD	TBD
	Magnitude Defining Behavior	author

Notes: The amplitude table is now represents all different Signal Detection Hypothesis Feature Measurements. The Data Model allows Station Magnitude Solutions to reference multiple Signal Detection Hypothesis Feature Measurements. The current schema would require duplication of rows with a different ampid to cover this.

5.2. netmag

IDC Database Schema	Data Model	
netmag Table Column	**Class**	**Attribute**
magid	TBD	TBD

net	TBD	TBD
orid	Magnitude Solution	parent location solution
evid	TBD	TBD
magtype	Magnitude Solution	mag type
nsta	TBD	TBD
magnitude	Magnitude Solution	mag
uncertainty	Magnitude Solution	uncertainty
auth	Magnitude Solution	author
commid	TBD	TBD
lddate	TBD	TBD