Science Data Model v2

François Viallefond
Observatoire de Paris - LERMA

Laboratoire d'Étude du Rayonnement et de la Matière en Astrophysique

Steve Torchinsky
Observatoire de Paris - USN

Science Data Model history from MS to SDMv2

- Measurement Set (MSv2)
  - Used offline with AIPS++ (now casacore)
  - Evolved to become the ALMA SDM (ASDMv0)

Science Data Model history from MS to SDMv2

- Measurement Set (MSv2)
  - Used offline with AIPS++ (now casacore)
  - Evolved to become the ALMA SDM (ASDMv0)

ASDMv1

  - Real time data capture
  - ALMA and EVLA
  - ~50 tables
  - This is the current version used for ALMA & EVLA

Science Data Model v2 - Goals

- Generalised model for any radio telescope
  - Single dish
  - Aperture synthesis
    - (eg. ALMA, EVLA, PdB, ...)
  - VLBI
  - Aperture plane phased arrays (eg. EMBRACE)
  - Focal Plane Arrays
    - Feed clusters (eg. Parkes, Arecibo, Effelsberg, ...)
    - Phased array feeds (eg. ASKAP, APERTIF)

Science Data Model v2 - Overview

- Enum
  - Categorised enumerators
- PhysQuan (Physical Quantities)
  - Collections of physical quantities (eg. Temperatures) and measures (eg. QDirections)
  - Virtual Quantities (eg. Phase_Dir, Staircase for OTF)
  - Values with uncertainties (eg. Position, Length)
- SDM
  - Ensemble of Tables
    - Relations within tables, and between tables
  - Normalisation
    - Allows queries in a database
  - Set/subset based

Science Data Model v2 - components

- Meta data
  - eg. definition of a configuration
- Auxilliary data
  - eg. instrument diagnostics (temperatures, encoder values, etc)
- Data
  - eg. visibilities, autocorrelations

Objects and Types persist and remain accessible

Example: EMBRACE@Nançay

EMBRACE@Nançay

- 5760 antenna elements
- Hierarchical beam forming
- Organised into tilesets
  - 20 tilesets
  - Each tileset is 4 tiles of 72 elements each
  - “statistics” include
    - Cross correlations between tilesets (imaging)
- Lots of flexibility regarding bandpass, pointing, number of pointings

EMBRACE@Nançay

Hierarchical beamforming

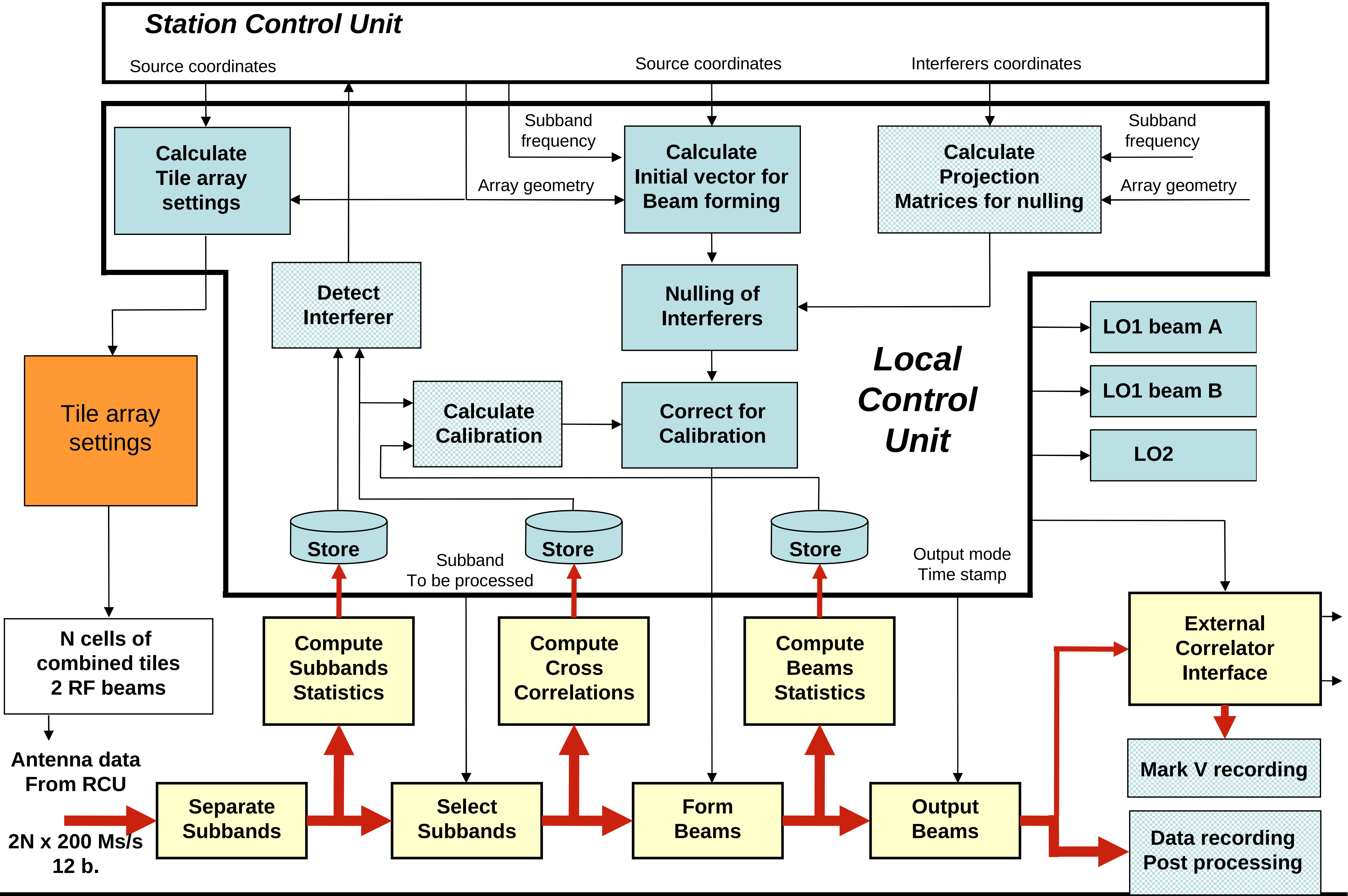
4 elements phased
Together in the
Beamformer
microchip

Hierarchical beamforming

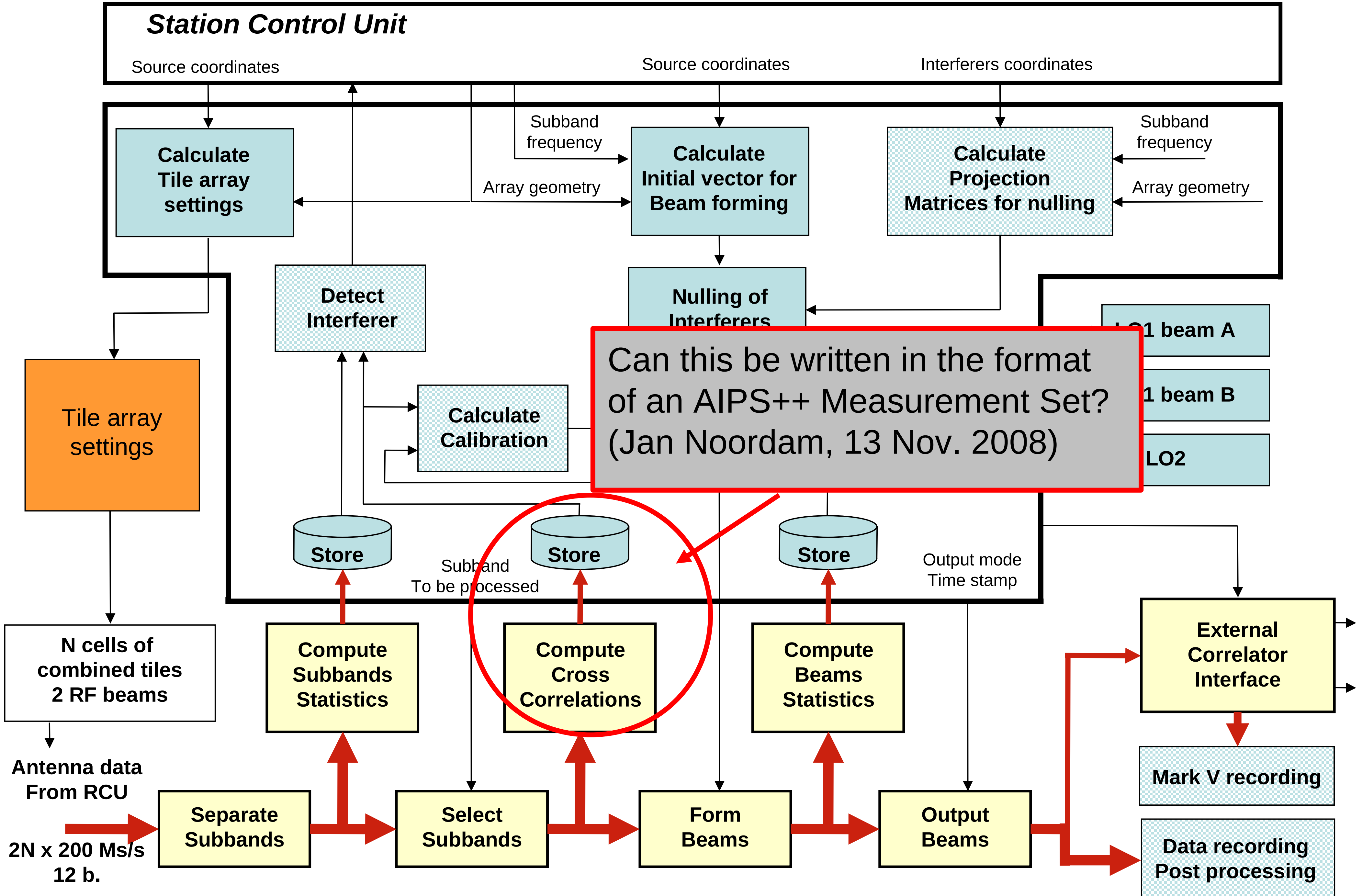
Hierarchical beamforming

Hierarchical beamforming

EMBRACE Station Digital Processing



EMBRACE Station Digital Processing



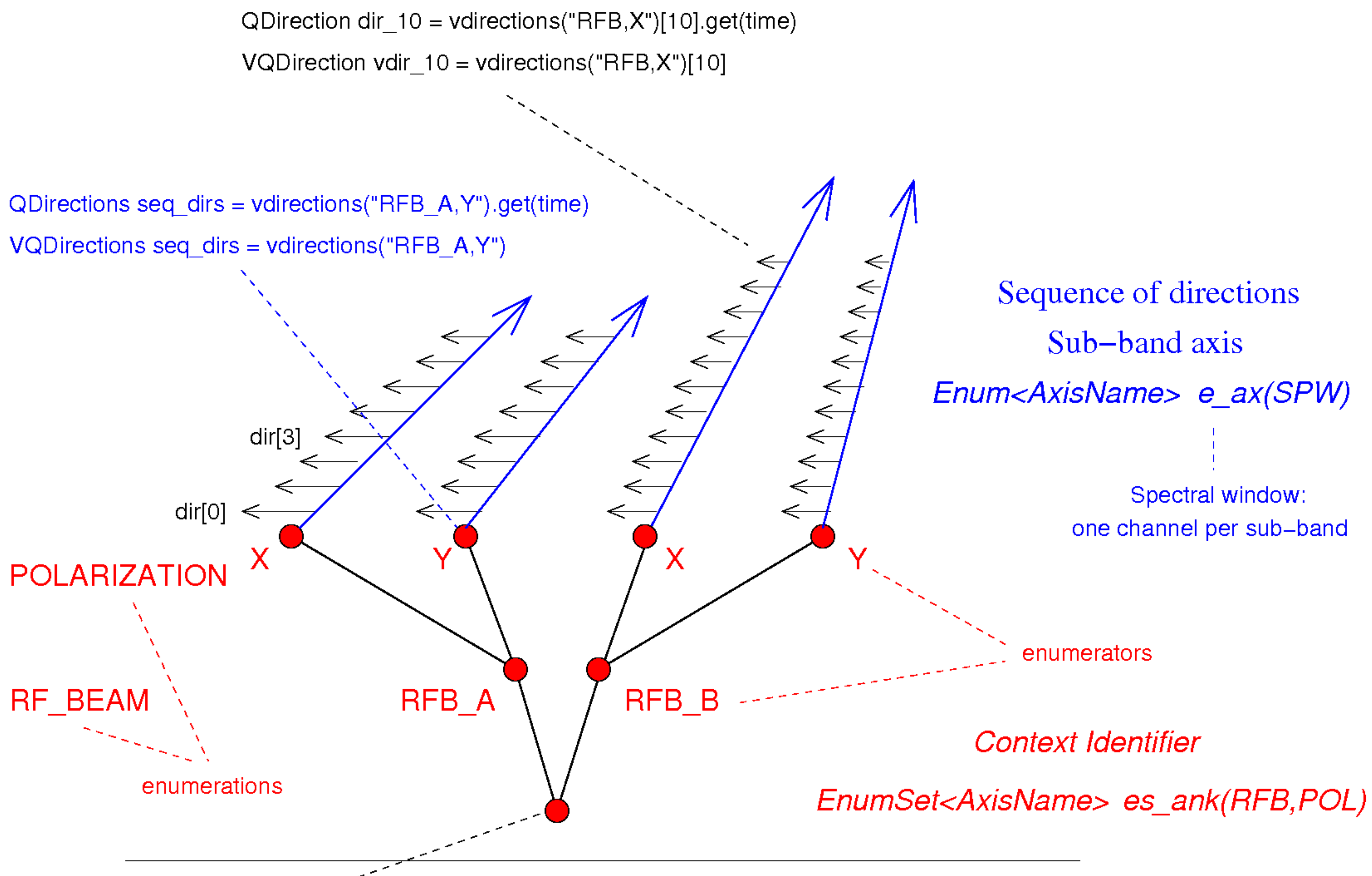
SDMv2 Main Features

- Can be easily pruned for a specific instrument
- Can be applied to any instrument of any complexity or simplicity
  - eg. future radio telescopes: ASKAP, APERTIF, EMBRACE, ALMA, EVLA, VLBI, etc
- eg. for SKA: It will be composed of multiple technologies, but all can produce data in SDMv2

Example of a Collection of Measures

FIELD::DELAY_DIR VQDirections

Object in one cell of the DELAY_DIR column in the FIELD table



VQDirections vdirections: Sequence of sequences of virtual directions

Examples of Table Dependencies

SDMv2 for EMBRACE

- 3 fundamental axes
  - Time
    - scan, subscan, integration, subintegration
  - Spectral
    - baseband, subband, spectral window
  - Aperture
    - station, tileset, tile, feed element (Vivaldi)
- Table
  - Set of sections
    - Key (one or ordered sequence of fields)
    - Data description (meta data)
    - Data (contains primary key, secondary key, etc)

Example: The Feed Table

- Key Section
  - Antenna Id
  - Spectral Window Id
  - Time interval
  - Feed Id

The 3 fundamental axes

- Data Description Section
  - Number of receptors
- Data Section
  - eg. Beam offset, Phased Feed Id, RF Beam Id, etc

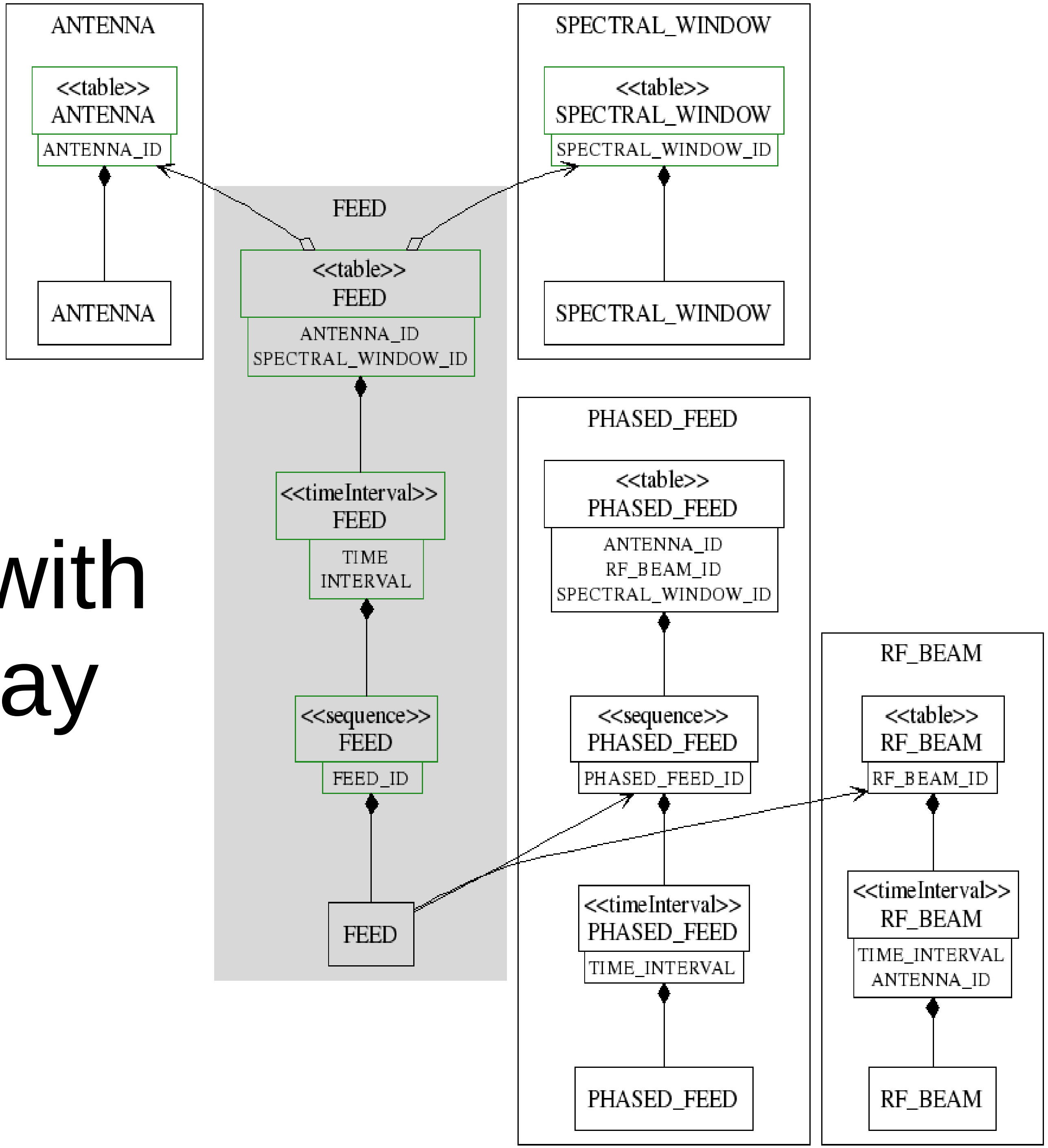
Example: The Feed Table (converted to a Measurement Set)

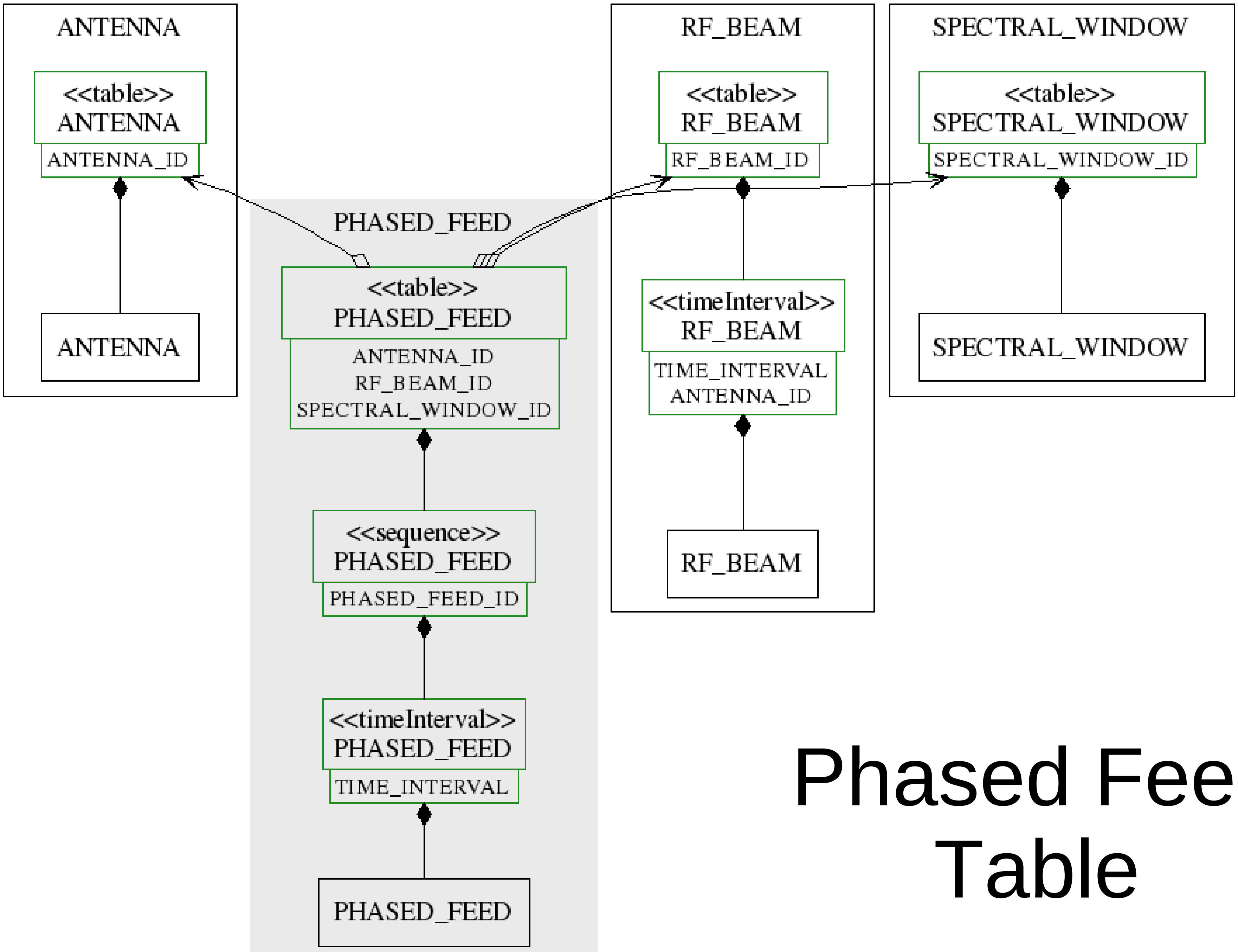
Table 7: Feed characteristics

FEED			
Name	**Format**	**Measure**	**Comments**
Key			
ANTENNA_ID	Int		Antenna id
SPECTRAL_WINDOW_ID	Int		Spectral window id.
TIME	Double	EPOCH	Interval midpoint
INTERVAL	Double		Time interval
FEED_ID	Int		Feed id
Data description			
NUM_RECEPTORS	Int		Number of receptors on this feed
Data			
BEAM_OFFSET	Double	DIRECTION	Beam position offset (on sky but in antenna reference frame).
(FOCUS_LENGTH)	Double		Focus length
(PHASED_FEED_ID)	Int		Phased feed
POLARIZATION_TYPE	String[N_r]		Type of polarization to which a given RECEPTOR responds.
POL_RESPONSE	Complex[N_r][N_r]		Feed polzn. response
(RF_BEAM_ID)	Int		RF beam

Continued on next page

Feed Table with Phased Array



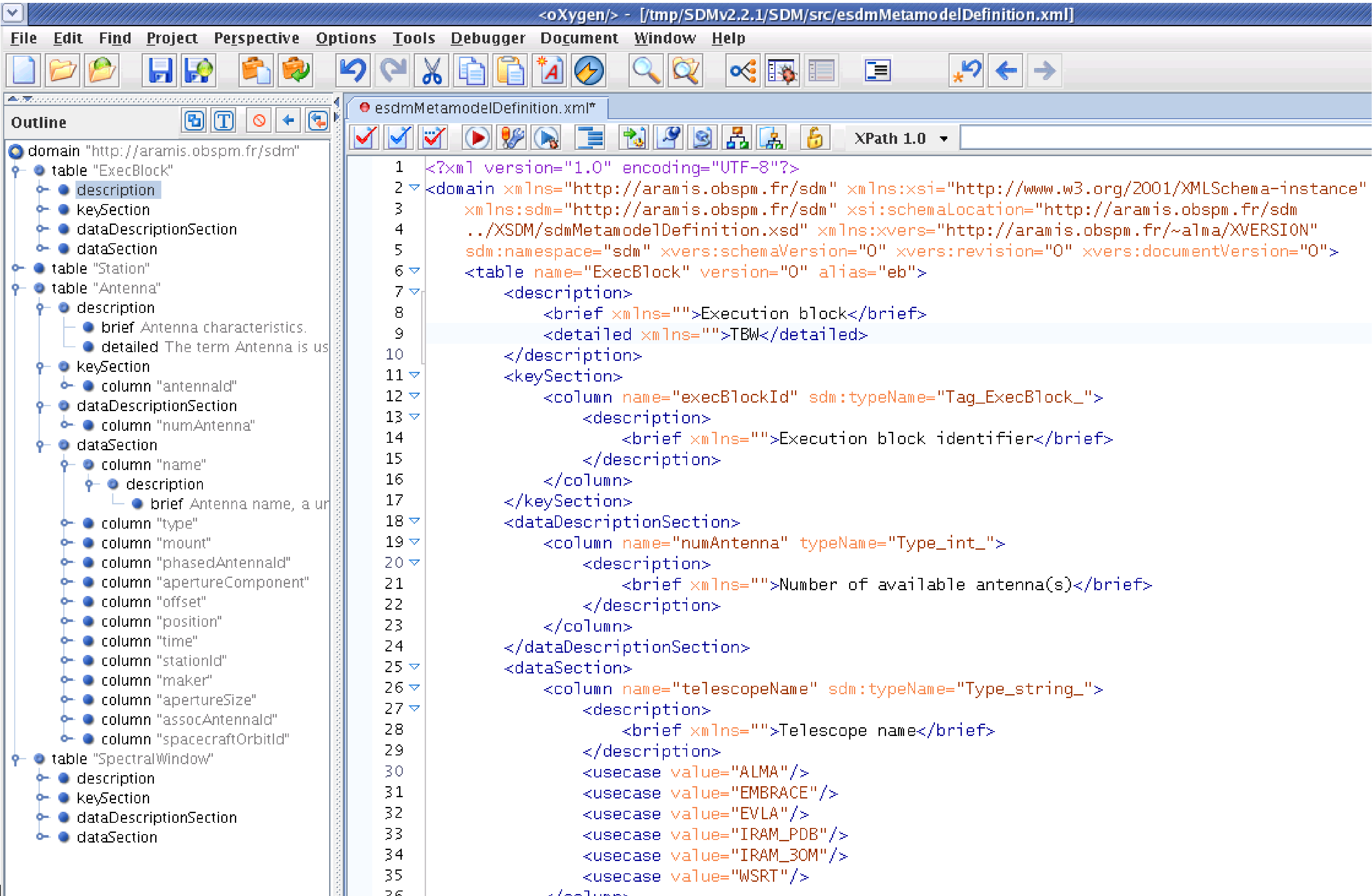


Phased Feed Table

SDMv2 - Implementation

- 3 levels
  - Meta model: generic parameters (XML Schema)
  - Model: instance of the meta model (XML)
  - Data set: instance of the model (XML + binaries)
- Code generation from the model (c++ template)
  - c++ classes
  - XML Schema (Type definitions)
    - eg. Dataset<Sdm>, Dataset<Sdm,EMBRACE>
  - Python interface

Model defined with valid XML



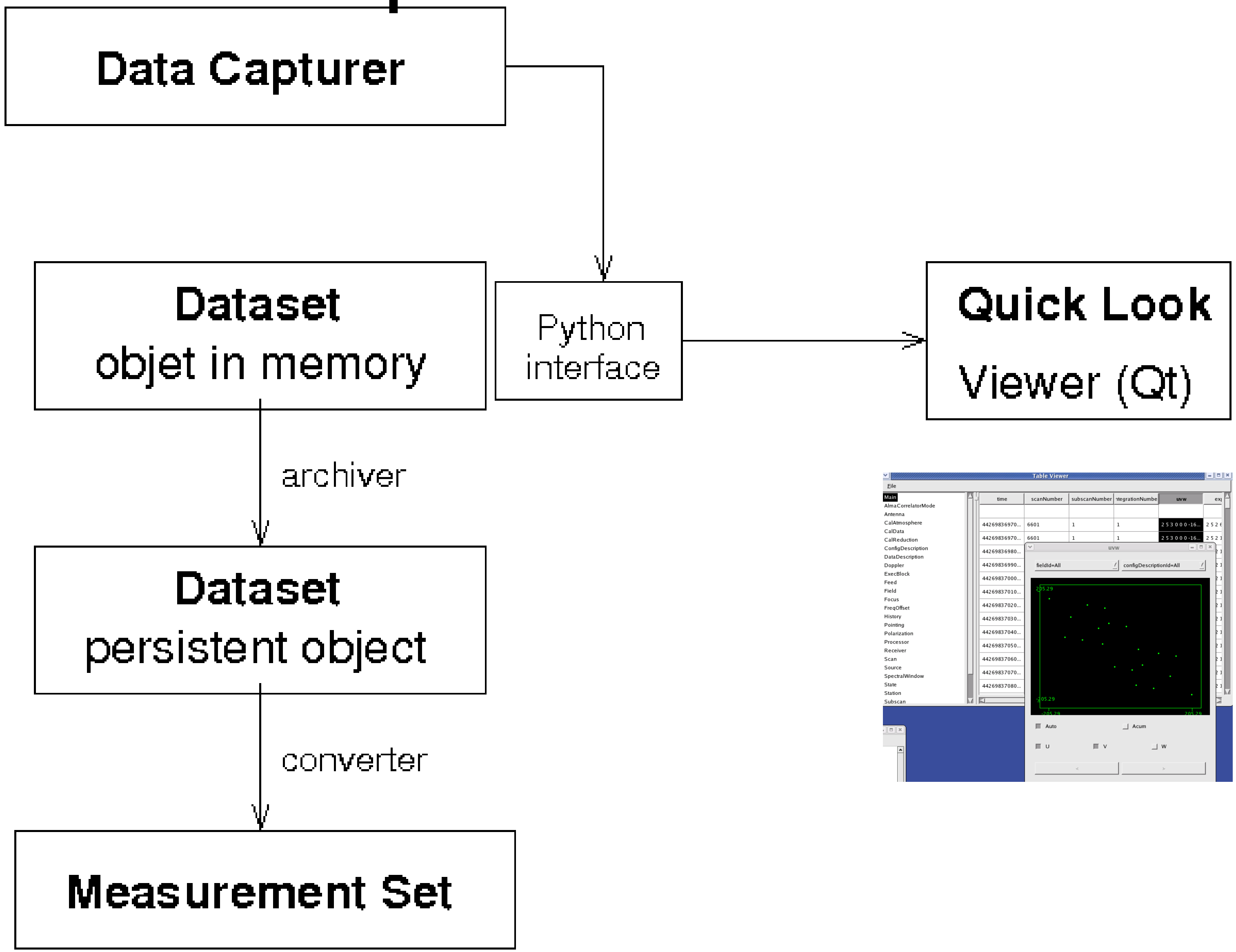
Example: Antenna table defined with valid XML

```xml
<table name="Antenna" version="0" alias="ant">
    <description>
        <brief xmlns="">Antenna characteristics.</brief>
        <detailed xmlns="">The term Antenna is used to define a generic aperture element. This
            element may be either a compound of smaller apertures or the component of a larger
            aperture, or both simultaneously in a hierarchy. The constraint is that all these
            elements must respond in phase to wave-fronts coming from at least one direction. If
            the antenna is a compound of smaller aperture elements, these do not require to form
            a monolithic aperture. With this definition a co-phased aperture synthesis array,
            e.g. the WSRT as a whole, can be one entry (one row) in this Antenna table, a
            typical use-case for VLBI.</detailed>
    </description>
    <keySection>
        <column name="antennaId" sdm:typeName="Tag_Antenna_">
            <description>
                <brief xmlns="">Antenna identifier,</brief>
            </description>
        </column>
    </keySection>
    <dataDescriptionSection>
        <column name="numAntenna" typeName="Type_int_">
            <description>
                <brief xmlns="">Number of co-phased components</brief>
            </description>
        </column>
    </dataDescriptionSection>
```

specialisation with usecases

```xml
<column name="phasedAntennaId" sdm:localKeyRefs="Tags_Antenna_" sizeCol1="numAntenna"
    optional="true">
    <description>
        <brief xmlns="">References to the co-phased aperture components</brief>
    </description>
    <usecase context="APERTURE_ARRAY"/>
</column>
<column name="apertureComponent" sdm:typeName="Enum_ApertureComponent_" optional="true">
    <description>
        <brief xmlns="">Aperture component this list of phased antenna references
            corresponds to.</brief>
    </description>
    <usecase context="APERTURE_ARRAY"/>
</column>
<column name="offset" ns="pq" sdm:typeName="QPosition" refcode="YOKE" refixed="true"
    offset="position">
    <description>
        <brief xmlns="">Phase reference position. Dish antenna: in the Yoke, relative to
            the antenna pedestal. Aperture array: TBD </brief>
    </description>
    <usecase context="APERTURE_ARRAY" value="ANTENNA_STATION"/>
    <usecase context="ALT_AZ" value="YOKE" columnAttribute="refcode"/>
</column>
<column name="position" ns="pq" sdm:typeName="QPosition" refcode="ANTENNA_STATION"
    refixed="true" sdm:ref="Station/position" sdm:keyid="2">
    <description>
        <brief xmlns="">Position of the antenna pedestal reference point, relative to
            the station reference point</brief>
    </description>
</column>
```

Data Capture for EMBRACE



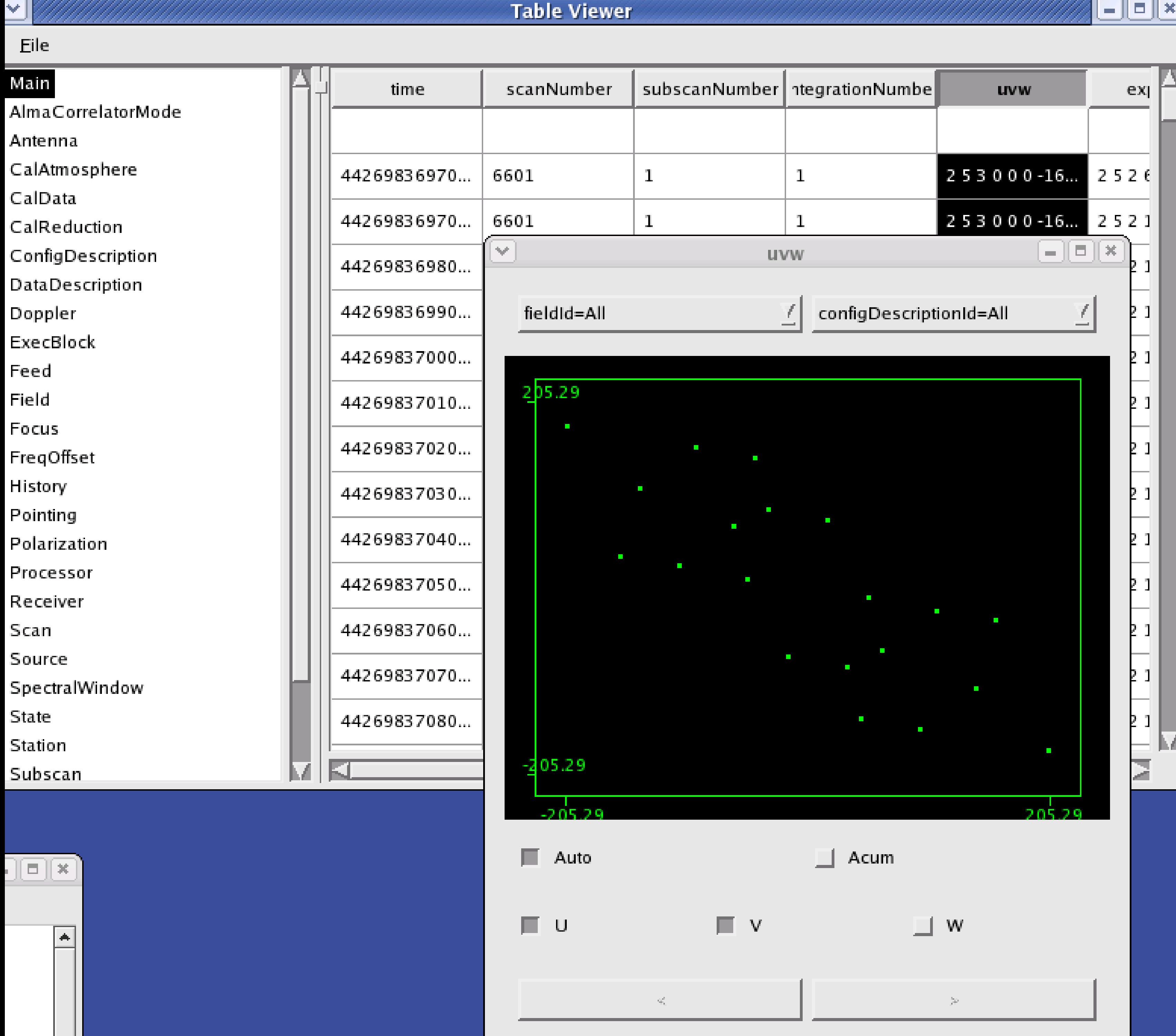


Table Viewer
File
Main
AlmaCorrelatorMode
Antenna
CalAtmosphere
CalData
CalReduction
ConfigDescription
DataDescription
Doppler
ExecBlock
Feed
Field
Focus
FreqOffset
History
Pointing
Polarization
Processor
Receiver
Scan
Source
SpectralWindow
State
Station
Subscan
time
scanNumber
subscanNumber
ntegrationNumbe
uvw
ex
44269836970...
6601
1
1
2 5 3 0 0 0 -16...
44269836970...
6601
1
1
2 5 3 0 0 0 -16...
44269836980...
44269836990...
44269837000...
44269837010...
44269837020...
44269837030...
44269837040...
44269837050...
44269837060...
44269837070...
44269837080...
uvw
fieldId=All
configDescriptionId=All
205.29
-205.29
-205.29
205.29
Auto
Acum
U
V
W

SDMv2 Current Status

- SDMv1 in operation for ALMA & EVLA
- SDMv2 Implementation for EMBRACE nearly complete.
- Testing with EMBRACE@Nançay in September

SDM Collaboration

- ALMA SDM
  - Robert Lucas (ALMA)
- EMBRACE SDM
  - Steve Torchinsky (Obs de Paris - Nançay)
  - Frédéric Badia (Obs de Paris - Nançay)
  - Jean-Michel Martin (Obs de Paris - GEPI)
  - Henrik Olofsson (Onsala Space Observatory)
  - Philippe Picard (Obs de Paris - Nançay)