

# Satellite definition

	Satellite	
PK FK	ParentHashKey	References the parent's primary key. Together with LoadDate they provide the context and time of the change.
PK	LoadDate	Same as in Hub definition. This is the time at which the data warehouse saw the record for the first time.
	LoadEndDate	Indicates the date and time when the satellite entry becomes invalid. It is the only attribute that is updated in a satellite. The update occurs once a new entry is loaded from the source system. While the new entry has a current LoadDate, the last valid entry's LoadEndDate is updated to match the new entry's LoadDate.
	RecordSource	Same as in Hub definition.
Optional	HashDiff	It is a hash value of all descriptive data of a satellite's entry. Having this hash values facilitates the comparison process between rows to identify if a new record is really unique or to identify if a record changed. The hash can be calculated using all descriptive attributes or some attributes. It depends on attribute changes are required to track.
Optional	ExtractDate	Captures the date and time when the data was extracted from the source system. It relies on the format and configuration of the source system, which indicates low reliability.
	Attributes **	The rest of descriptive attributes.

## Satellite characteristics:

- Stores the descriptive data of a business object, relationship or transaction.
- Add context at a given time or over a time period to hubs and links.
- Tracks the changes in data over time.
- One satellite is attached to only one link, hub or satellite. Therefore it is identified with parent's ID and the timestamp of the change.

- The Satellite is the only Data Vault Entity that stores the end date of a relationship.
- Not all descriptive data will be stored in the same satellite, it will be categorized and split into multiple satellites if necessary.

#### Warning

Because the history of data needs to be preserved, you are not allowed to update or attribute the data in the satellite. The only exception to this rule is the `LoadEndDate` attribute of the previous version of the data.

## Splitting satellites.

#### Info

Always create one satellite per source.

#### 1. It is recommended to **split first by source system**:

- Splitting by source systems allows to improve parallel loading (multiple systems at a time), better adapt the definition to different requirements of different sources, allows to add new sources without changing existing satellites entities, etc.
- Even when satellites are separated based on source system, it is needed to have the `RecordSource` to track specifics like `SAP.FINANCE` and `SAP.MARKETING`.

#### 2. It is recommended to split **second by rate of change**:

- If there are more than two frequencies, there should be more than two satellites.
- One satellite for non changing attributes, one for those attributes that change monthly, one for those that change daily.