Appendix to **IRS 90940**

2nd edition, September 2022 **Last update 01/07/2022**

Original

Data exchange with Driver Advisory Systems (DAS) following the SFERA protocol

Appendix E

SFERA Data Handbook

This appendix:

- is updated regularly,
- corresponds to the above date of update.

INTERNATIONAL UNION OF RAILWAYS

Table of contents

Appendix E - SFERA Data Handbook

This appendix is intended as a guide for handling SFERA data in the implementation. It begins with a “getting started” section aimed at providing a complete list of SFERA messages and core elements, and some suggestions for a minimum starting point. It then goes in depth into how to model railway infrastructure and train features, and how to put these together to provide advice to trains.

The rest of the appendix offers details on specific data topics such as train positioning, messaging and non-core data. The final sections provide implementation hints and considerations on the choices made while developing SFERA.

1 Getting started with SFERA data

This section is aimed at supporting the implementation of SFERA by providing a reference to the complete list of SFERA messages and core elements, as well as some suggestions for minimum implementation and some guidelines for validating a SFERA “package”.

1.1 SFERA messages and core elements

As described in the chapter “The Journey Profile Model” of the IRS, SFERA contains four core elements in which the information is exchanged. From ground to board, DAS Trackside (DAS-TS) sends Journey Profiles (JPs), Segment Profiles (SPs) and Train Characteristics (TCs) to DAS on board (DAS-OB). From board to ground, DAS-OB sends Status Reports (SRs) to DAS-TS.

These core elements are not sent directly, but through ground-to-board (G2B) and board-to-ground (B2G) messages, which can contain several such core elements, as well as other data and some meta-data (e.g. timestamps and message IDs). As explained in the chapter “Data Structure” there are six possible types of messages in SFERA, each with a header and a payload:

- SFERA_G2B_EventMessage
- SFERA_B2G_EventMessage
- SFERA_G2B_RequestMessage
- SFERA_B2G_RequestMessage
- SFERA_G2B_ReplyMessage
- SFERA_B2G_ReplyMessage

This section gives an overview of the six types of SFERA message and the four core elements. For more details and specific descriptions, please refer to the XML Schema Definition (XSD) in Appendix D.

The explanations of the messages will refer to the SFERA use cases which are described in detail in Appendix A “Use Cases” The explanations mostly refer to DAS-OB and DAS-TS, but sometimes they have to be interpreted as “Provider” and “Receiver” in the sense explained in Chapter 4 of the IRS.

1.1.1 SFERA messages

1.1.1.1 SFERA_G2B_EventMessage

The SFERA ground-to-board event message is mostly used to send JP updates from DAS-TS to DAS-OB. This is covered by Use Case JP3. It is also possible with this message to force a driving mode change (see Use Case DM4), to send C-DAS-C advice updates (see Use Case CDC3), plain text messages (described in Section 8.1 of this appendix) or related train information (see Section 8.3). It is only possible to send one of these different payloads at a time.

Fig. 1 shows the structure of SFERA_G2B_EventMessage.

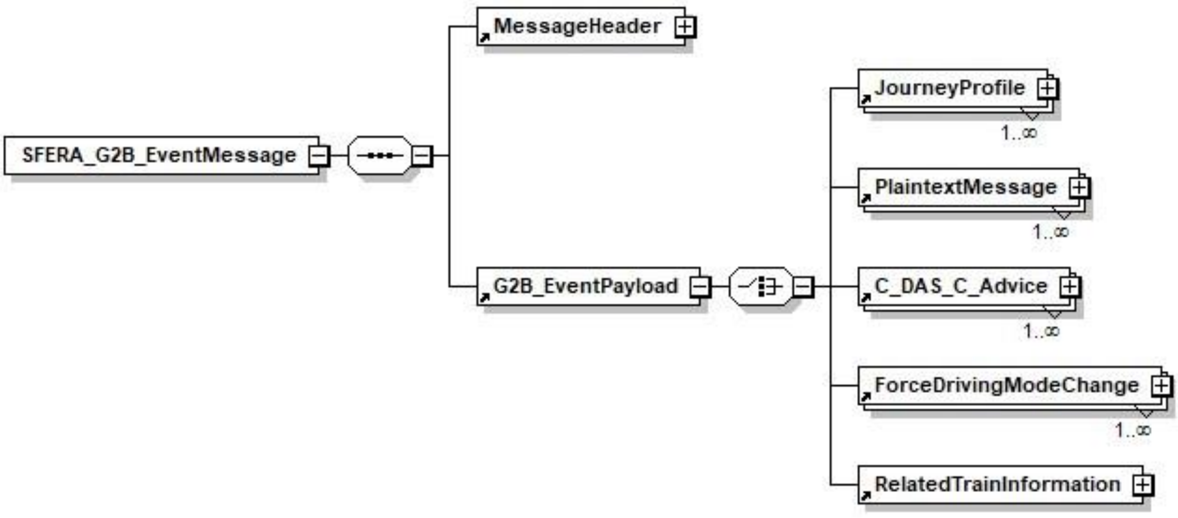


Fig. 1 shows the structure of SFERA_G2B_EventMessage.

1.1.1.2 SFERA_B2G_EventMessage

The SFERA board-to-ground event message is used to send either a notification of Session Termination (see use case DX2) or a Status Report (SR) from DAS-OB to DAS-TS. The SR is described in Section 1.1.2.4 of this appendix.

Fig. 2 shows the structure of SFERA_B2G_EventMessage.

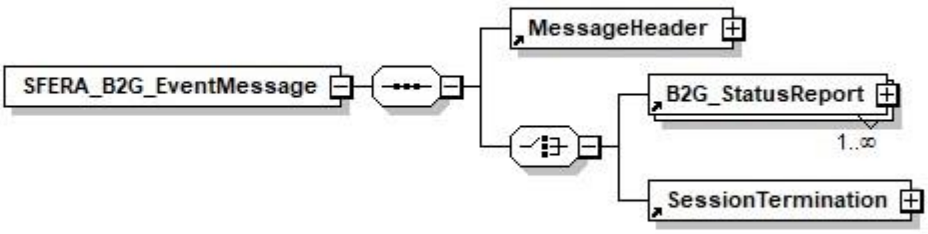


Fig. 2 shows the structure of SFERA_B2G_EventMessage.

1.1.1.3 SFERA_G2B_RequestMessage

The SFERA ground-to-board request message is used to send requests from DAS-TS to DAS-OB. It is possible to request the DAS driving mode, as described in Use Case DM1, as well as to request the termination of the session (see use case DX4). It is also possible to request a plaintext message or a status report from on board.

Fig. 3 shows the structure of SFERA_G2B_RequestMessage.

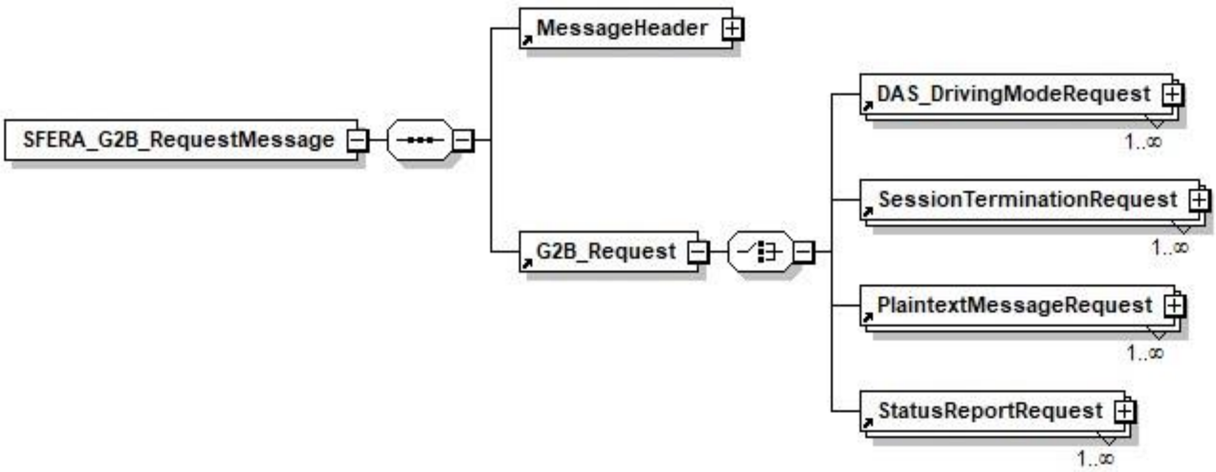


Figure 3: SFERA_G2B_RequestMessage

1.1.1.4 SFERA_B2G_RequestMessage

The SFERA board-to-ground request message can have two different types of payload: either a HandshakeRequest, which is typically used only at the beginning of the communication between board and ground (see Use Case DX1) to present functional parameters to be used in the subsequent use cases (e.g. timeouts and available functions on both sides), or a B2G_Request, which is used to request information from DAS-TS.

In a typical C-DAS-O situation, the B2G_Request is used by DAS-OB to request JPs, SPs and TCs from DAS-TS (see Use Cases JP2 and JP3). In C-DAS-C, it can also be used to request C-DAS-C advice (see Use Case CDC2).

There are also specific situations in which the SFERA_B2G_RequestMessage can be used to request information from the trackside. For instance, as described in Use Case AF1, DAS-OB may need to request its own position from DAS-TS (e.g. in a tunnel, where GNSS is not available). Or after a time of forced disconnection, the DAS-OB may inquire about whether any plaintext messages or status change requests were sent from DAS-TS.

Fig. 4 shows the structure of SFERA_B2G_RequestMessage.

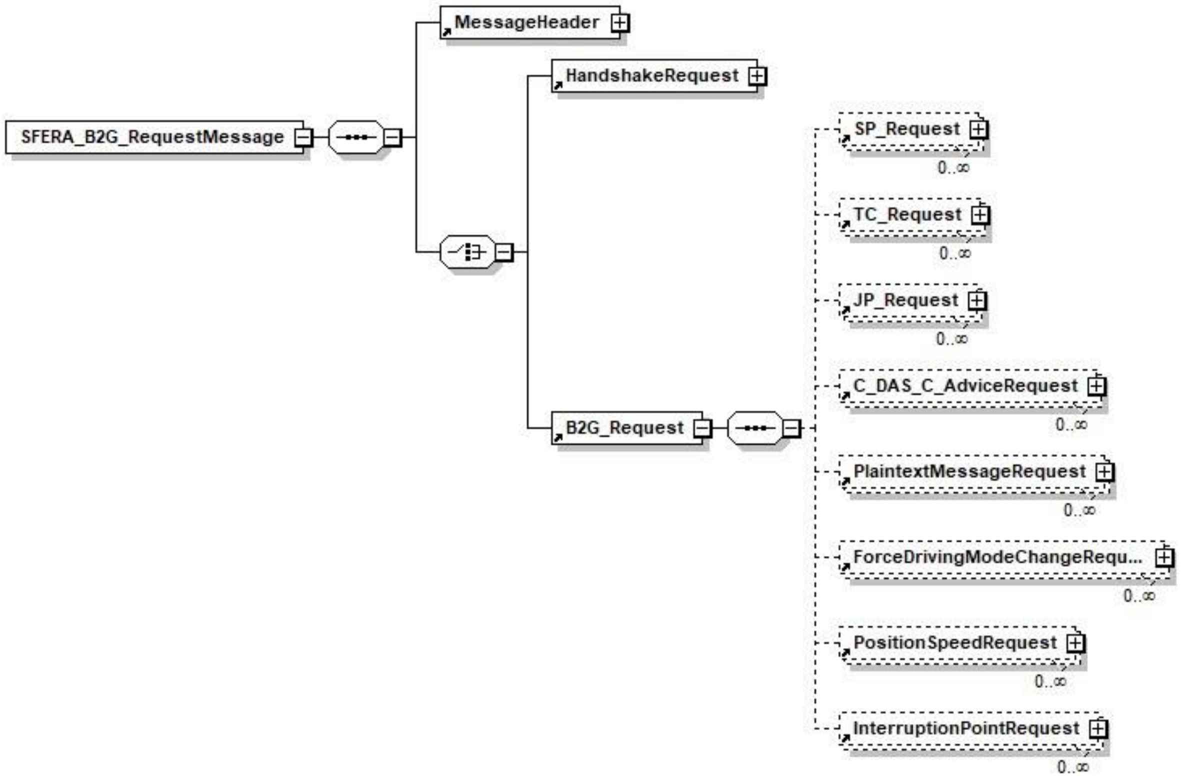


Figure 4: SFERA_B2G_RequestMessage

1.1.1.5 SFERA_G2B_ReplyMessage

The SFERA ground-to-board reply message is used to reply to a B2G_EventMessage or a B2G_RequestMessage. It can contain either HandshakeAcknowledgement/HandshakeReject, which is only used during a connection handshake (see use case DX1) or a G2B_ReplyPayload, which is used to send back information from DAS-TS to DAS-OB. This can be one or more JPs, SPs and TCs, or other information requested with the B2G_RequestMessage.

The reply message can also simply be a G2B_MessageResponse, which is an acknowledgement of the message received. This can be a positive acknowledgement ("OK") or an error. When an error message is sent, this means that the original message has one or more problems. The reply message can then indicate the parts of the original XML message where the errors lie and error codes for each of the problems found. The list of error codes is available in Appendix H "Error Codes in SFERA".

Fig. 5 shows the structure of SFERA_G2B_ReplyMessage.

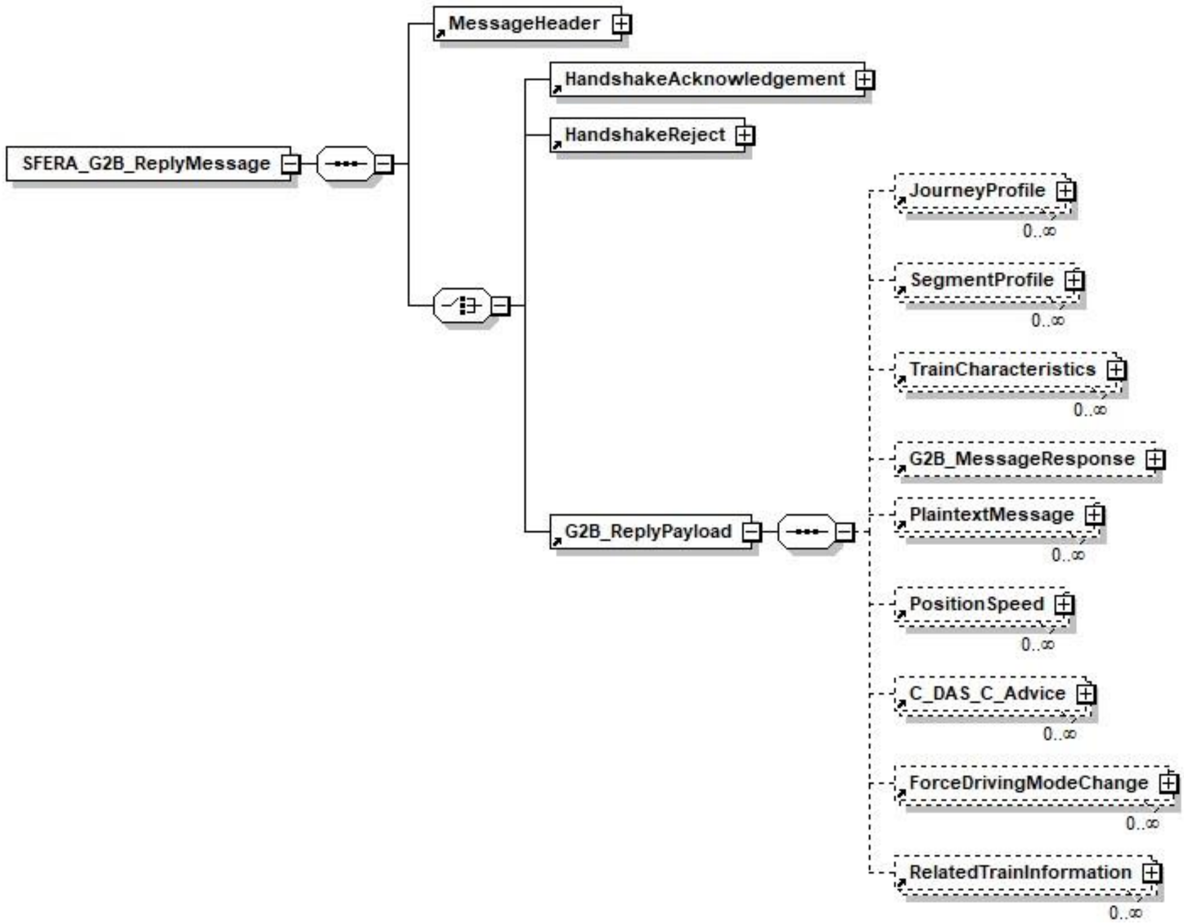


Figure 5: SFERA_G2B_ReplyMessage

1.1.1.6 SFERA_B2G_ReplyMessage

The SFERA board-to-ground reply message is used to reply to a G2B_EventMessage or a G2B_RequestMessage. It can contain the DAS driving mode, as described in Use Case DM1, a plaintext message as described in Section 8.1 of this appendix, a Status Report if requested from ground or a session termination.

As with the ground-to-board reply message, the SFERA_B2G_ReplyMessage can also contain a message response with a positive acknowledgement or an error and the details of all the errors found (location in the original message and error codes).

Fig. 6 shows the structure of SFERA_G2B_ReplyMessage.

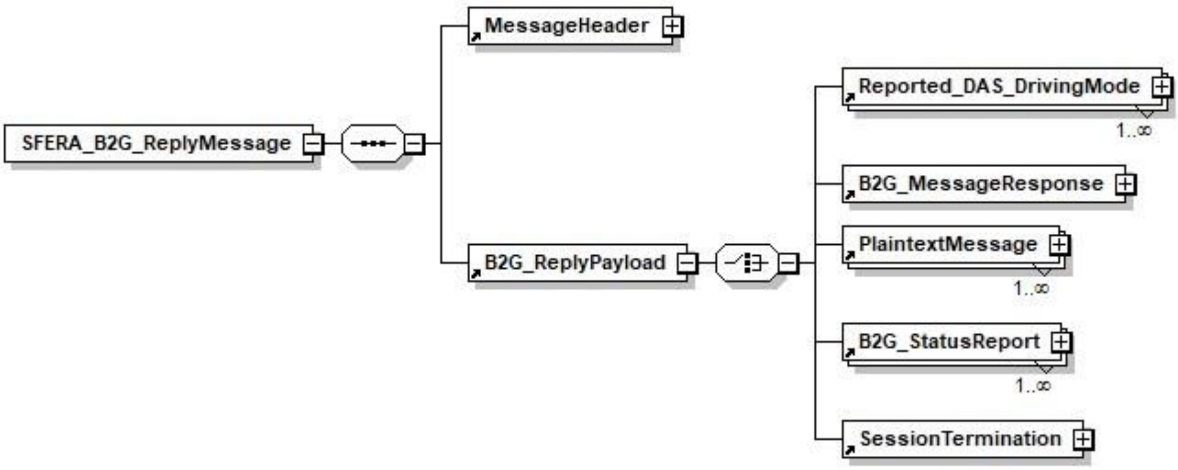


Figure 6: SFERA_B2G_ReplyMessage

1.1.1.7 Messages and possible responses

The following tables present:

- An exhaustive list of the SFERA messages, taking into account the different payloads.
- The responses that can be expected for each of these messages.

Table 1 - Expected answers to B2G events and requests

Message type of sent message	Payload type	Possible content of the SFERA_G2B_ReplyMessage used to answer	Significance of the reply
SFERA_B2G_ EventMessage	B2G_StatusReport	G2B_MessageResponse / result = “OK”	Status Report received
SFERA_B2G_ EventMessage	B2G_StatusReport	G2B_MessageResponse / result = “ERROR” [optional: G2B_MessageResponse / G2B_Error / errorCode]	Error in receiving or reading the StatusReport
SFERA_B2G_ EventMessage	SessionTermination	G2B_MessageResponse / result = “OK”	Session Termination received
SFERA_B2G_ EventMessage	SessionTermination	G2B_MessageResponse / result = “ERROR” [optional: G2B_MessageResponse / G2B_Error / errorCode]	Error in receiving or reading the Session Termination
SFERA_B2G_ RequestMessage	HandshakeRequest	HandshakeAcknowledgement	This is the nominally expected payload
SFERA_B2G_ RequestMessage	HandshakeRequest	HandshakeReject [optional: handshakeRejectReason, ATOTS_ID or SP_Zone]	The ground side of the communication rejects the Handshake. Optionally, a reason can be given as

Message type of sent message	Payload type	Possible content of the SFERA_G2B_ReplyMessage used to answer	Significance of the reply
			well as the indication of another ATOTS or SP_Zone to be contacted.
SFERA_B2G_ RequestMessage	HandshakeRequest	G2B_MessageResponse / result = "ERROR" [optional: G2B_MessageResponse / G2B_Error / errorCode]	Error in receiving or reading the HandshakeRequest
SFERA_B2G_ RequestMessage	B2G_Request / SP_Request	G2B_ReplyPayload / SegmentProfile	This is the nominally expected payload
SFERA_B2G_ RequestMessage	B2G_Request / SP_Request	G2B_MessageResponse / result = "ERROR" [optional: G2B_MessageResponse / G2B_Error / errorCode]	Error in receiving or reading the SP Request
SFERA_B2G_ RequestMessage	B2G_Request / TC_Request	G2B_ReplyPayload / TrainCharacteristics	This is the nominally expected payload
SFERA_B2G_ RequestMessage	B2G_Request / TC_Request	G2B_MessageResponse / result = "ERROR" [optional: G2B_MessageResponse / G2B_Error / errorCode]	Error in receiving or reading the TC Request
SFERA_B2G_ RequestMessage	B2G_Request / JP_Request	G2B_ReplyPayload / JourneyProfile	This is the nominally expected payload
SFERA_B2G_ RequestMessage	B2G_Request / JP_Request	G2B_MessageResponse / result = "OK"	The ground side of the communication does not have a JP available that is more recent than the JP known to the board side of the request.
SFERA_B2G_ RequestMessage	B2G_Request / JP_Request	G2B_MessageResponse / result = "ERROR" G2B_MessageResponse / G2B_Error / dataFirstAvailable	The JP for the train is not available yet. Try again later.
SFERA_B2G_ RequestMessage	B2G_Request / JP_Request	G2B_MessageResponse / result = "ERROR" [optional: G2B_MessageResponse / G2B_Error / errorCode]	Error in receiving or reading the JP Request

Message type of sent message	Payload type	Possible content of the SFERA_G2B_ReplyMessage used to answer	Significance of the reply
SFERA_B2G_ RequestMessage	Multiple B2G_Request / SP_RequestAnd/orMultiple B2G_Request / TC_RequestAnd/orMultiple B2G_Request / JP_Request	Either multiple reply messages containing SegmentProfiles, or a JourneyProfile or TrainCharacteristics each.	This is the nominally expected payload
SFERA_B2G_ RequestMessage	Multiple B2G_Request / SP_RequestAnd/orMultiple B2G_Request / TC_RequestAnd/orMultiple B2G_Request / JP_Request	Or a unique reply message containing all SegmentProfiles and TrainCharacteristics and JourneyProfiles	This is the nominally expected payload
SFERA_B2G_ RequestMessage	Multiple B2G_Request / SP_RequestAnd/orMultiple B2G_Request / TC_RequestAnd/orMultiple B2G_Request / JP_Request	G2B_MessageResponse / result = “ERROR” [optional: G2B_MessageResponse / G2B_Error / errorCode]	Error in receiving or reading the request message
SFERA_B2G_ RequestMessage	Multiple B2G_Request / SP_RequestAnd/orMultiple B2G_Request / TC_RequestAnd/orMultiple B2G_Request / JP_Request	G2B_MessageResponse / result = “ERROR”G2B_MessageResponse / G2B_Error / dataFirstAvailable	The JP for the train is not available yet. Try again later.
SFERA_B2G_ RequestMessage	B2G_Request / C_DAS_C_AdviceRequest	G2B_ReplyPayload / C_DAS_C_Advice	This is the nominally expected payload
SFERA_B2G_ RequestMessage	B2G_Request / C_DAS_C_AdviceRequest	G2B_MessageResponse / result = “OK”	The ground side of the communication does not have C_DAS_C_Advice available that is more recent than the C_DAS_C_Advice known to the board side of the request.
SFERA_B2G_ RequestMessage	B2G_Request / C_DAS_C_AdviceRequest	G2B_MessageResponse / result = “ERROR” [optional: G2B_MessageResponse / G2B_Error / errorCode]	Error in receiving or reading the C_DAS_C_Advice request message
SFERA_B2G_ RequestMessage	B2G_Request / PlaintextMessageRequest	G2B_ReplyPayload / PlaintextMessage	This is the nominally expected payload
SFERA_B2G_ RequestMessage	B2G_Request / PlaintextMessageRequest	G2B_MessageResponse / result = “ERROR” [optional: G2B_MessageResponse / G2B_Error / errorCode]	Error in receiving or reading the C_DAS_C_Advice request message

Message type of sent message	Payload type	Possible content of the SFERA_G2B_ReplyMessage used to answer	Significance of the reply
SFERA_B2G_ RequestMessage	B2G_Request / ForceDrivingModeChangeRequest	G2B_ReplyPayload / ForceDrivingModeChange	This is the nominally expected payload
SFERA_B2G_ RequestMessage	B2G_Request / ForceDrivingModeChangeRequest	G2B_MessageResponse / result = "ERROR" [optional: G2B_MessageResponse / G2B_Error / errorCode]	Error in receiving or reading the ForceDrivingModeChange request message
SFERA_B2G_ RequestMessage	B2G_Request / PositionSpeedRequest	G2B_ReplyPayload / PositionSpeed	This is the nominally expected payload
SFERA_B2G_ RequestMessage	B2G_Request / PositionSpeedRequest	G2B_MessageResponse / result = "ERROR" [optional: G2B_MessageResponse / G2B_Error / errorCode]	Error in receiving or reading the PositionSpeed request

Table 2: Expected answers to G2B events and requests

Message type of sent message	Payload type	Possible content of the SFERA_B2G_ReplyMessage used to answer	Significance of the reply
SFERA_G2B_EventMessage	G2B_EventPayload / JourneyProfile	B2G_ReplyPayload / B2G_MessageResponse / result = "OK"	JP received without error. (The fact that the JP is being used will be confirmed later by a StatusReport stating the JP_Version in use by the algorithm)
SFERA_G2B_EventMessage	G2B_EventPayload / JourneyProfile	B2G_ReplyPayload / B2G_MessageResponse / result= "ERROR" [Optional: B2G_ReplyPayload / B2G_MessageResponse / B2G_Error / errorCode]	Error in receiving or reading the JP
SFERA_G2B_EventMessage	G2B_EventPayload / PlaintextMessage	B2G_ReplyPayload / B2G_MessageResponse / result = "OK"	PlainTextMessage received without error
SFERA_G2B_EventMessage	G2B_EventPayload / PlaintextMessage	B2G_ReplyPayload / PlainTextMessage	PlainTextMessage sent by the driver to the sender of the original G2B message

Message type of sent message	Payload type	Possible content of the SFERA_B2G_ReplyMessage used to answer	Significance of the reply
SFERA_G2B_EventMessage	G2B_EventPayload / PlaintextMessage	B2G_ReplyPayload / B2G_MessageResponse / G2B_Error	Error in receiving or reading the PlainTextMessage message
SFERA_G2B_EventMessage	G2B_EventPayload / C_DAS_C_Advice	B2G_ReplyPayload / B2G_MessageResponse/ result="OK"	C_DAS_C_Advice received without error
SFERA_G2B_EventMessage	G2B_EventPayload / C_DAS_C_Advice	B2G_ReplyPayload / B2G_MessageResponse / result= "ERROR" [Optional: B2G_ReplyPayload / B2G_MessageResponse / B2G_Error / errorCode]	Error in receiving or reading the C_DAS_C_Advice message
SFERA_G2B_EventMessage	G2B_EventPayload / ForceDrivingModeChange	B2G_ReplyPayload / B2G_MessageResponse / result="OK"	ForceDrivingModeChange message received without error.
SFERA_G2B_EventMessage	G2B_EventPayload / ForceDrivingModeChange	B2G_ReplyPayload / B2G_MessageResponse / result= "ERROR" [Optional: B2G_ReplyPayload / B2G_MessageResponse / B2G_Error / errorCode]	Error in receiving or reading the ForceDrivingModeChange message
SFERA_G2B_EventMessage	G2B_EventPayload / RelatedTrainInformation	B2G_ReplyPayload / B2G_MessageResponse / result="OK"	RelatedTrainInformation message received without error
SFERA_G2B_EventMessage	G2B_EventPayload / RelatedTrainInformation	B2G_ReplyPayload / B2G_MessageResponse / result= "ERROR" [Optional: B2G_ReplyPayload / B2G_MessageResponse / B2G_Error / errorCode]	Error in receiving or reading the RelatedTrainInformation message
SFERA_G2B_RequestMessage	G2B_Request / DAS_DrivingModeRequest	B2G_ReplyPayload / Reported_DAS_DrivingMode	Here is a Status Report containing at least my DAS_drivingMode
SFERA_G2B_RequestMessage	G2B_Request / DAS_DrivingModeRequest	B2G_ReplyPayload / B2G_MessageResponse / result= "ERROR" [Optional: B2G_ReplyPayload / B2G_MessageResponse / B2G_Error / errorCode]	Error in receiving or reading the DAS Driving Mode Request

Message type of sent message	Payload type	Possible content of the SFERA_B2G_ReplyMessage used to answer	Significance of the reply
SFERA_G2B_RequestMessage	G2B_Request / SessionTermination Request	B2G_ReplyPayload / SessionTermination	Session Termination request received and confirmed
SFERA_G2B_RequestMessage	G2B_Request / SessionTermination Request	B2G_ReplyPayload / B2G_MessageResponse / result= “ERROR” [Optional: B2G_ReplyPayload / B2G_MessageResponse / B2G_Error / errorCode]	Error in receiving or reading the Session Termination Request
SFERA_G2B_RequestMessage	G2B_Request / PlaintextMessageRequest	B2G_ReplyPayload / PlaintextMessage	This is the nominally expected payload
SFERA_G2B_RequestMessage	G2B_Request / PlaintextMessageRequest	B2G_ReplyPayload / B2G_MessageResponse / result= “ERROR” [Optional: B2G_ReplyPayload / B2G_MessageResponse / B2G_Error / errorCode]	Error in receiving or reading the Plaintext Message Request
SFERA_G2B_RequestMessage	G2B_Request / StatusReportRequest	B2G_ReplyPayload / B2G_StatusReport	This is the nominally expected payload
SFERA_G2B_RequestMessage	G2B_Request / StatusReportRequest	B2G_ReplyPayload / B2G_MessageResponse / result= “ERROR” [Optional: B2G_ReplyPayload / B2G_MessageResponse / B2G_Error / errorCode]	Error in receiving or reading the Status Report Request

1.1.2 SFERA Core elements

1.1.2.1 Segment Profile

As described in Chapter 3 of the IRS, the Segment Profile contains data about the infrastructure. Every SP is identified by an ID and a “Zone”, which can be an IM identifier (that has to be a RICS Code) and/or a NID_C from ERTMS (at least one of the two shall be specified). Inside a Zone, SP_IDs have to be unique.

The SP data is considered “mostly static” because changes are obviously possible (e.g. on a powerless section or on a balise), but are unlikely to happen very often. However, if they occur, the SP has mandatory major and minor version attributes that can be used when something actually changes in the SP. Splitting the version into a major and minor part exists for maintenance purposes only. A change of minor and/or major version will have the same functional impact on the on-board device.

Another mandatory attribute of the SP is its length. This is because it is essential for the DAS to be able to locate any point, area or characteristic of the SP. SP length shall not be zero. Elements inside the SP have a location, i.e. the distance of the element from the beginning of the SP. This location can range between 0 (inclusive) and SP length (inclusive).

Many other attributes and elements are contained in the SP, all of which are described in detail in the XSD. Some of them are discussed in this appendix.

An SP is a one-dimensional object that describes a single and consecutive section of track. Whenever this is relevant, the data relates to the centre line between the two rails of a track.

Fig. 7 shows the structure of the Segment Profile element.

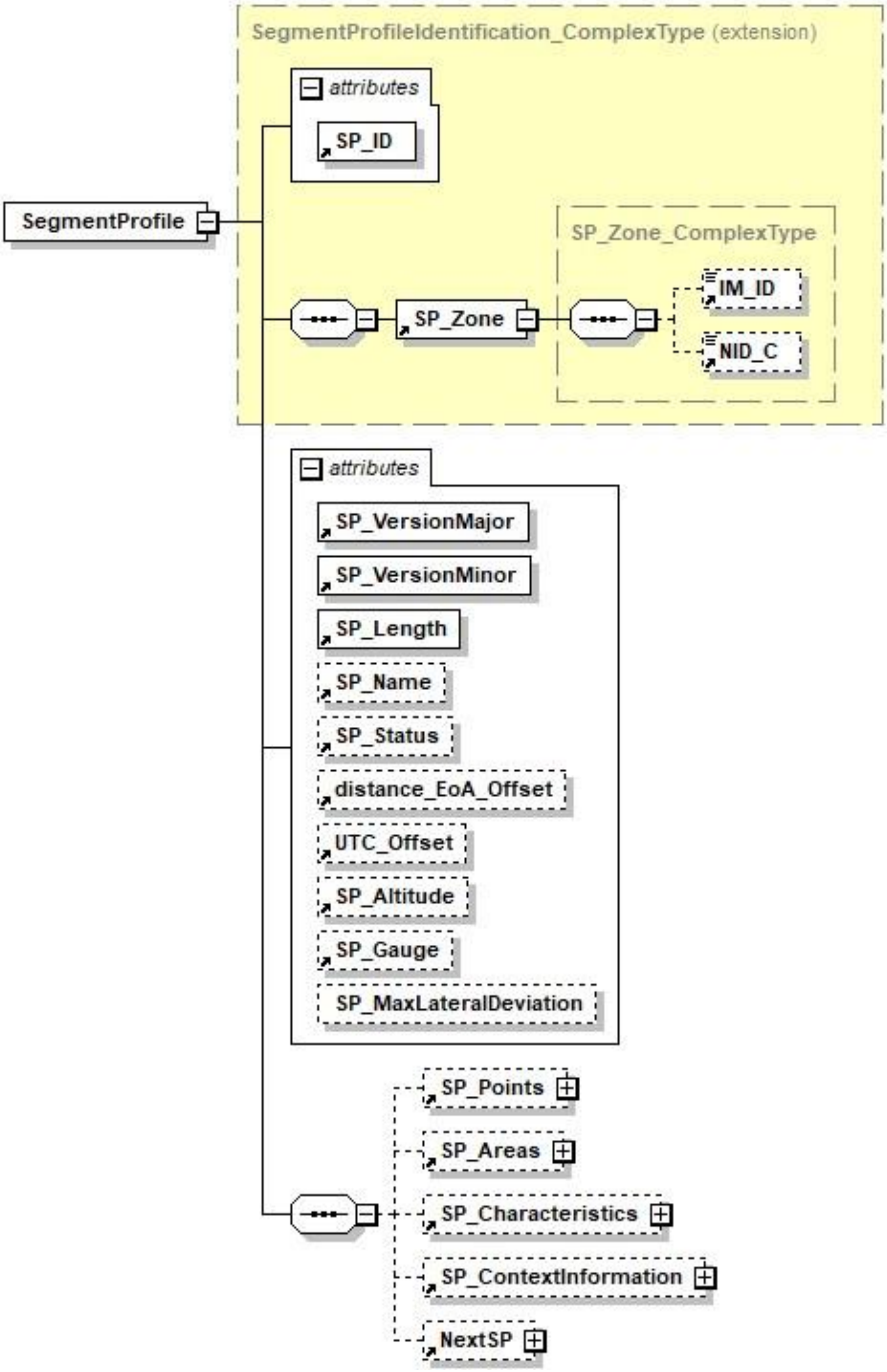


Figure 7: Segment Profile

1.1.2.2 Train Characteristics

The information about the train features are contained in the TC. The TCs have an identifier produced by a TC_ID together with the ID of an RU (a RICS code) or the ID of a VKM. Inside the same RU or VKM, the TC_ID has to be unique.

TCs also have a major and a minor version. The version shall be changed every time there is any change in the TC_Features.

The TC_Features are a host of characteristics for the train (e.g. length, weight, engines, vehicles, etc.) including the ATP systems supported by the train and its traction and braking force curves.

Fig. 8 shows the structure of the Train Characteristics element.

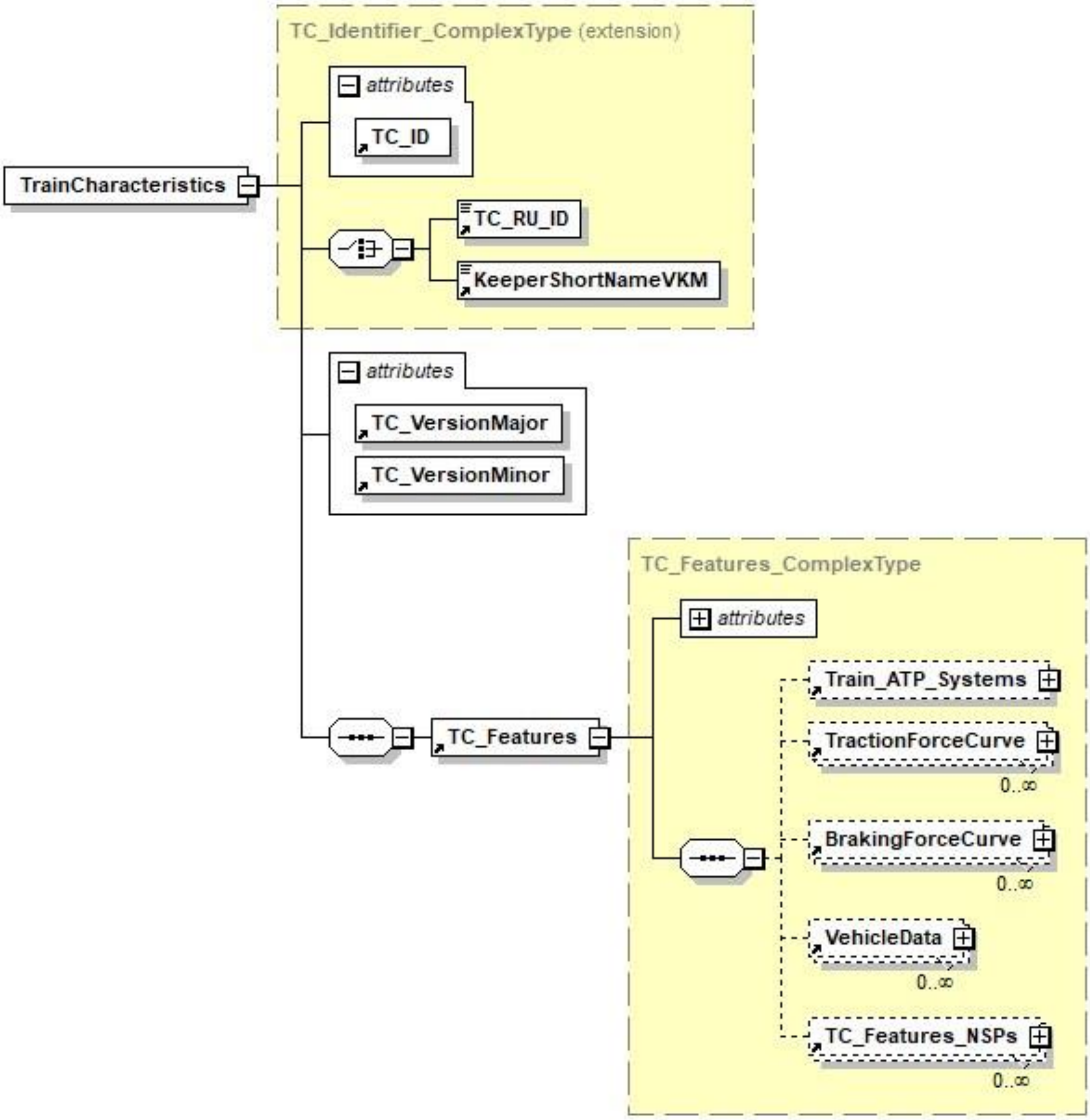


Figure 8: Train Characteristics

1.1.2.3 Journey Profile

As mentioned in Chapter 3 of the IRS, the JP is the element that provides the actual objectives on the journey to the DAS-OB. It contains the identification of the train and a list of references to SPs, with their ID and version. In each SP reference, it is possible to reference different TCs (together with their IDs and versions). It is also possible to include references to Timing Points and a set of temporary constraints for the journey, either regarding signals or other topics (e.g. current, low adhesion, Temporary Speed Restrictions).

All elements are listed in the order the train will traverse them. In ERTMS/ATO it is mandatory to add the SP_Direction (nominal or reverse) for each SP. The SPs references for a journey shall not have any gaps and shall not overlap. This also applies to JP updates. When reaching the end of a train journey, the last Timing Point can have the TP_Information=1 indicating End of Journey (mandatory for ERTMS/ATO) and the SP shall have all sufficient information to ensure the train can achieve the specified alignment at the final Timing Point. Each JP shall have at least one Timing Point.

Fig. 9 shows the structure of the Journey Profile element.

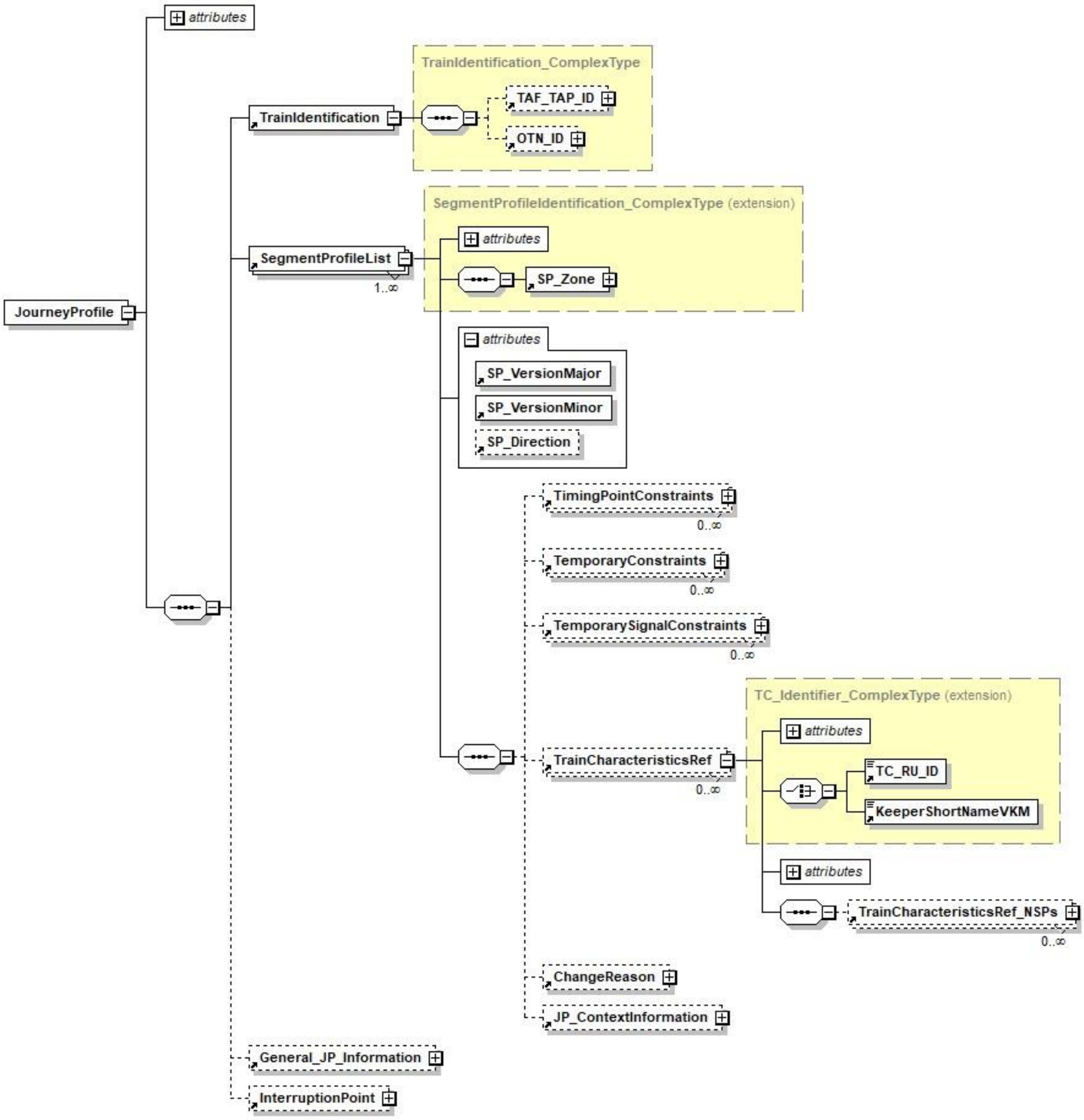


Figure 9: Journey Profile

1.1.2.4 Status Report

The Status Report, sent by the SFERA_B2G_EventMessage, may contain a set of data. For example, the position and speed of the train (see Use Case SR1) and the references to past or upcoming Timing Points, which may indicate whether the train can or cannot respect the time window at a future TP (see Use Case SR2).

The SR may also be used to communicate a change in adhesion conditions found on the track (see Use Case SR3), or the forecast for consumption and regeneration of current (see Use Cases PW3 and PW4). And if the driver or system needs to inform DAS-TS that the characteristics of a train have changed (see Use Cases TC1 and TC2), this is also done through the SR.

The element energyData can be used to include real-time energy-related data. This is useful when the DAS is connected to TCMS. Autonomy and maxPower relate to battery trains and hydrogen trains, for example, that have a limited energy source on board. Maximum power may therefore have to be reduced with less autonomy available.

Several other pieces of information can be sent in an SR: the status and behaviour of DAS-OB, as described in Chapter 3 of the IRS; the current JP in use or the C-DAS-C advice shown to the driver (see Use Cases JP3 and CDC3); a plain text message, as described in Section 8.1 of this appendix; and a number of variables used in the Status Report of SUBSET-126 (for more details and the translation of SUBSET-126 variables to SFERA and vice-versa, please see Appendix G “Correspondence Table SUBSET-126/SFERA”).

Fig. 10 shows the structure of the Status Report element.

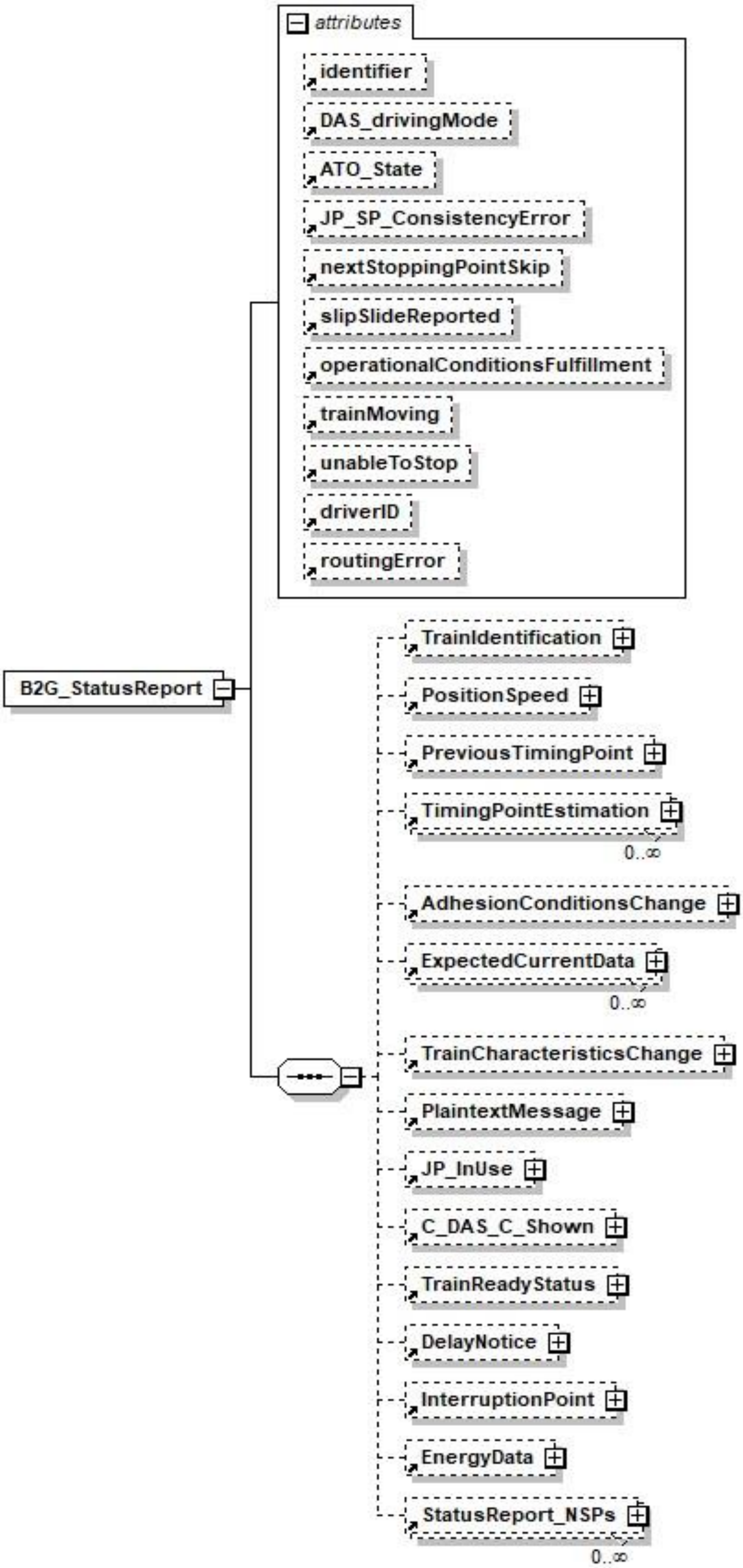


Figure 10: Status Report

1.2 Network-Specific Parameters

In several parts of the data structure, it is possible to include so-called Network-Specific Parameters (NSPs). These are parameters that have not been standardised for interoperable use in SFERA, but may be used in bilateral agreements between trackside and on-board DAS. They can include data, for instance, that is only relevant in a specific IM area, or they can be used to test an additional function before submitting a Change Request to include the relevant data structure in SFERA.

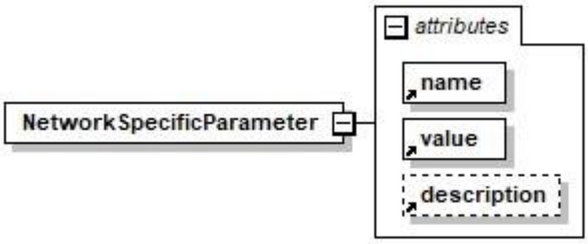


Figure 11: Network-Specific Parameter

A Network-Specific Parameter entry always takes the form of a pair (name, value) with an additional “description” attribute (see Fig. 11). NSPs are integrated in the data structure in the most appropriate way required by the location. For example, it is possible to define Network-Specific Points or Areas in the Segment Profile, Network-Specific Temporary Constraints or Network-Specific features in the Train Characteristics.

1.3 Suggestions for a minimal implementation

This section suggests a minimal set of fields that should be present in any SFERA message. This set should include the most important elements of Journey and Segment Profiles, such as Timing Points and virtual balises. Note that the “minimal set” proposed here does not correspond only to the mandatory elements and attributes to be included in the XML message, because there are multiple ways in which SFERA can be used.

Depending on the ATP system and track layouts, some information on the signals may be required. This signal information will be less relevant on ATP systems with fixed layouts (e.g. PZB) or ERTMS FS. Similarly, gradients are required in all but very flat regions.

For the Timing Points, the departure and/or arrival time (at least one of these) should be present at each stopping point.

Regarding virtual balises, a train should know where it is located on the network (see Section 2.1.1). This can be done by reading external inputs and odometers. However, the primary source for this information on handheld devices will be GNSS. At least the location of first departure should be present; providing locations at each stop is encouraged.

As for Train Characteristics, hardcoding by the RU can be considered for an S-DAS implementation. In that case, sending TCs from ground to board is not needed.

In other cases, the following fields in the TC are generally required for minimal operation:

- TrainLength,
- TrainMaxSpeed,
- TrainWeight,
- EnginePower,
- RolloutCoefficientA,
- RolloutCoefficientB,

- RolloutCoefficientC,
- RollingStockType,
- RotatingMassFactor,
- TractionForceCurve,
- BrakingForceCurve.

Infrastructure Managers should assume that there are DAS that only use this minimal set, to be studied in bilateral agreements.

1.4 Guidelines for validating a SFERA “package” (JP+SP+TC)

Logically, a Journey Profile is a top-level element, referring to objects in Segment Profiles and Train Characteristics. A data inconsistency error shall be reported by the DAS if not all information can be obtained, or if the elements are not consistent. Specifically, the following situations may require attention:

1. Valid versions of all SPs and TCs referred to in the JP shall be obtained by the DAS.
2. All Timing Point identifiers mentioned in the JP shall be defined in the related SPs. The Timing Point ID shall be unique within one SP zone.
3. Contiguous range types for SP_Areas with startEndQualifier different than StartsEnds should merge well.
4. Correlation IDs given in a message header shall refer to valid earlier messages.
5. For changed or cancelled signal aspects in the Temporary Signal Constraints, the identifiers used in the signal shall match the identifiers of at least one signal object in the SP.
6. For the “Timing Point cannot be respected error” it shall be validated that Timing Points can be reached within the traction force curve, braking force curve and speed restrictions set by speed profiles and signal aspects.

1.5 Connection modes between DAS subsystems

Currently, the connection modalities (e.g. IP or DNS address, supported SFERA versions, etc.) between DAS subsystems depend on the implementation.

In future, this section will provide details of how to obtain a parameter file available to all SFERA users and an update process. This would include data such as the identifier of a DAS-TS (IM_ID or NID_C), the DAS-TS Address(es) to connect to, the SFERA versions supported, the supported architectures and the supported setups.

2 Modelling railway infrastructure in SFERA

This section deals with the way railway infrastructure is modelled in SFERA. As mentioned before, most of the railway infrastructure is described in Segment Profiles (SP), and additional elements are contained in the Journey Profile (JP).

2.1 Infrastructure description

2.1.1 Positioning on the infrastructure: Virtual Balises

A DAS needs to know the actual position of the front of the train and its speed in order to give valid advice. The DAS may have one or multiple sources for this information, e.g. GNSS sensor(s), odometer(s), ground radar(s), ETCS balise readers or visual mark readers.

The positioning system of SFERA is designed to be agnostic on the number and type of position and speed sensors. Therefore, all positioning (both for ground-to-board and board-to-ground communication) is done based on a sensor-independent linear coordinate system in each SP.

Some of the sensors only support relative positioning. For these, the passage of fixed points shall be observable. One of the physical marker types are ETCS balises.

Balises of one group (same NID_BG) shall be configured within the same SP. In ERTMS/ATO, balises can also be used to detect routing errors. In order to allow this, all balises with N_PIG=0 which are part of a balise group marked as linked shall be configured in the SP.

When an ETCS balise group is used to synchronise positioning, the on-board device needs to take into account the passing distance travelled during a transmission delay and the speed of the train. In case a second balise group is passed over this passing distance, only the last one is reported to the on-board ATO. The ATO will not be informed on passing the first balise group. This should be taken into account when creating SPs in order to avoid routing errors and to achieve correct synchronisation for positioning.

A virtual balise is a generic fixed point on the line. The identifiers of a virtual balise may be automatically interpreted if they represent an automatically identifiable object. This may be the case, for instance, with visual or RFID markers, or a balise of a Class B system or CBTC. In addition, latitude, longitude and altitude may be added such that GNSS may be used at this location. This is essential if DAS operation is foreseen on mobile devices not connected to the TCMS.

For DAS devices relying solely on a GNSS sensor, which is generally the case for DAS integrated into handheld hardware (tablet or smartphone), the coordinates of the virtual balises are used to construct a geo-localised representation of the track that is to be followed by the train. The DAS will project the coordinates given by the GNSS sensors onto this representation.

2.1.2 Languages for Timing Points and track and station names

Locations along the track can be indicated to the drivers on the DAS. The name of a Timing Point is contained in TimingPoint, under TP_Name. In some countries, different drivers may select the language in which a station is displayed. This may be local languages, but also a more universal language such as English. Therefore, the TP_Name element allows station names to be delivered in several languages. The DAS may then select the appropriate one. For instance, “Brussel Zuid” may be used for Dutch drivers coming from the north, while the same station can be known as “Bruxelles Midi” to French drivers coming from the south. The DAS should know which one is applicable (e.g. by hardcoding, driver selection, region parameters of the OS, etc.).

In many countries, Timing Points also have fixed abbreviations (usually between 1 and 8 characters in length). Each station has only one unique abbreviation. Therefore, TP_abbreviation does not have a language attribute.

A platform or track within a station or a track along the line may be indicated in the stationTrack attribute of a TimingPoint. This should also be universal for all languages. This is for information to drivers on their routes. The platform attribute does not include a name, and only indicates the usable length of a platform. It may be used to disable door opening for trains that exceed the platform length.

2.1.3 Line identifiers and km reference points

Optionally, track kilometres can be transmitted using the structure KilometreReferencePoint. They are transmitted in numeric form, with the indicator ‘ascending’ denoting if the counting runs parallel with the SP direction. This way interpolation between points is possible between the signs in the DAS display. Those with a textual identifier may also be shown. They are often identifiable in this description by a visual indicator along the track.

In each SP, at least one kilometre reference point should be present for each part with the same line and track identifier. Track kilometres shall be assumed to be contiguous within such a segment. A line and track identifier is not required for the use of a kilometre reference point. If the counting changes along a line, a pair of kilometre reference point values shall be given at the boundary. In that case, a distance of one metre in-between shall be given between the two points on either side of the gap.

2.1.4 Gradients

SFERA has three fields for gradients: average, steepest gradient and decisive gradient area. The differences are explained in this section.

Some ATP systems use the gradient for the calculation of braking curves. The steepest downhill gradient for a given part of the track is given to the ATP system. The additional gravitational force is subtracted from the maximum braking force in order to yield a safe braking curve. This curve is stored in SFERA, as the DAS needs to know how the ATP system limits train speed. IMs are generally conservative in terms of safety and send the steepest gradient for extended ranges of track. Thus, ATP gradients will never sum up to absolute height differences of points along a track.

The SFERA average gradient field is used for running time calculations. It affects the train speed in all phases (acceleration, cruising, coasting and braking) and should be filled with the best estimate of the gradient. The ATP gradient may yield too early arrival if used. The average gradient should be averaged over longer lengths, as this gives the correct running time estimate when used outside an area. Specifically, highest and lowest points differing more than a few meters on a route should be connected using gradients that sum up to the correct height differences. Including these gradients near elevations (e.g. bridges and tunnels) will give the DAS the opportunity to yield more stable advice. Otherwise, the DAS will advise speeding up at the top of a bridge (as the speed is a little lower) and then braking once the train is level again, as speeding up at the top may drive the train ahead of schedule. For downhill slopes into tunnels, having the correct average gradients enables the DAS to avoid overspeeding while coasting.

The decisive gradient area is shown to the driver. This is the steepest gradient for an area, uphill and downhill. Both can be shown for exactly the same area. Uphill, it is used by the driver to calculate the towing load, for example, to make sure, with the available engine power, that the locomotive can get away after it stops. Downhill, it is used by the driver for the calculation of either the minimum braking percentage or the speed limit, because in case of reduced braking capacity, the train might not be able to brake properly and the driver must therefore adjust the speed.

2.1.5 Altitude

Altitude can be given as an element at the starting point of each SP. The altitude can also be given on virtual balises. The altitudes are mainly used for the evaluation of the energy efficiency ratios of diesel engines. An ERTMS/ATO requirement is to determine the altitude value at the starting point of each SP.

2.1.6 Signals

The “Signal” elements in SFERA are meant to submit signal aspects that can be expected in the future and affect the way the driver has to drive the train. They are not used for submitting actual signal aspects to the train, as the actual aspect of a signal can differ from the expected aspect. Displaying any of this information to the driver should be avoided – the signal elements are dedicated purely to the calculation of the driving profile in the DAS.

The Signal object is designed with the following objectives in mind:

- to describe the effects on driving and not the way the signal looks in real life (there are too many different signals in the world);
- to focus on possible driver reactions for normal revenue running. Reactions for on-sight running or shunting, for instance, are not included, as neither DAS nor ATO GoA 2 is foreseen to benefit from this;
- to describe the physical characteristics of the signals, which can be useful for implementations relying on video recognition of signals, or for context information for the driver. Note that SFERA gives the general characteristics of the signal, but does not give the expected aspect.

The information displayed by a permanent signal in nominal operation is described in the SP. It can include, for instance, a maximum speed, a speed range to avoid, a power constraint or a traction type transition.

The signal types are reviewed to be consistent for Germany, France, Sweden, Switzerland, Austria, Norway, the Netherlands and Belgium. The SFERA working group reviewed the signals present in all the countries represented in the working group, considering that this would provide a near exhaustive panel of the different impacts on driving. If implementation of SFERA in another country reveals particularities not covered in the signal object, this can be the subject of a change request in the maintenance phase.

2.1.6.1 Applicability of a signal object

Optionally, the applicability of a signal object can be limited to certain trains on the following criteria:

- ATP system(s),
- train category(ies),
- intinerary(ies).

If one or more “Signal_ATP_System” is declared, only the trains which are considered as using one of these ATPs on this portion of track will take this signal object into account.

If one or more “Signal_TrainCategory” is declared, only the trains matching one of those categories on this portion of track will take this signal object into account.

If one or more “ApplicableIfTrainRunsOnSP” is declared, only the trains which have the referenced SP among the SPs in their Segment Profile List will take this signal object into account. The given SP shall be “downline” relative to the train’s direction of travel from the point where the signal is situated. This possibility can, for example, be used if a speed limit is only applicable when taking a switch in the deviating direction, and this switch is in a different SP.

In the example shown in Fig. 12, a signal is present on SP1, which gives an indication to the driver only if the point at the end of the common part of SP2 and SP3 is in a deviating position. This means that the signal does not give information to a driver going towards SP2. Therefore, the signal will reference SP3 in the “ApplicableIfTrainRunsOnSP”, which means the DAS will only take into account a speed reduction if the train is routed on the deviated track.

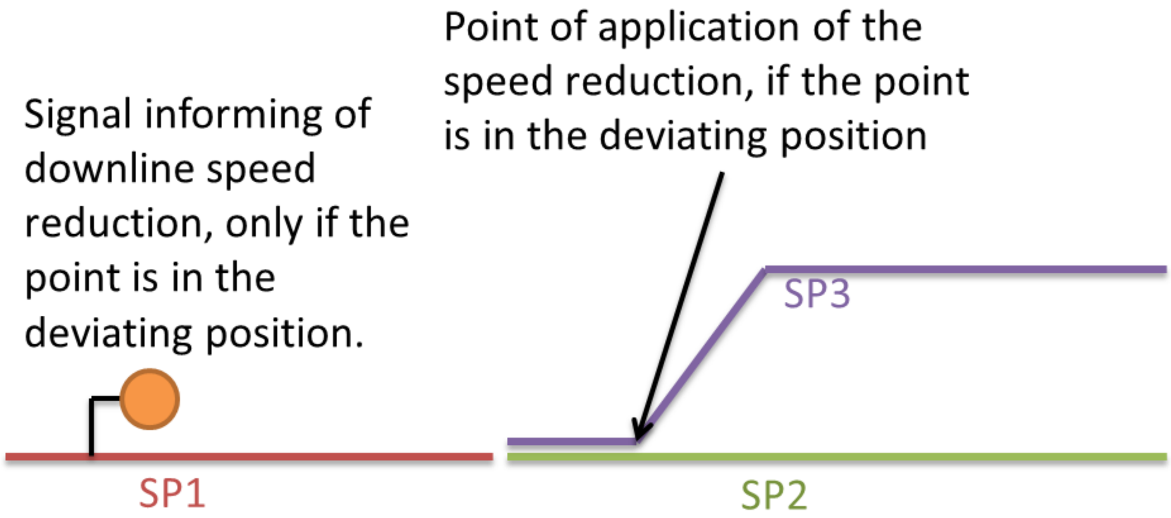


Figure 12: Example of signal application

If none of these elements is declared in a signal object, the “SignalInformation” (description of the effects of the signal on the train) is to be applied to all the trains.

2.1.6.2 Application area

The signal object has been built to be able to represent its effect on the driver's reactions. In many cases, the driver does not apply the restriction at the signal's exact position. This is why the signal has been structured with a separation between:

- the physical position of the signal;
- the application area where the driver is expected to apply the indication of the signal.

For example, in the case of a limited speed zone that is enforced by a distant signal, an execution signal, and an "end of execution" signal, a single signal object can be used:

- The distant signal's real-life position is used for the location information of the signal object;
- The startSignalApplication element will be used to define the point where the driver will be supposed to apply the speed restriction. This point could be:
  - the position of the execution signal;
  - a point *x* metres before the execution signal, in cases where the ATP system forces the driver to reach the specified speed *x* metres before the signal.
- The endSignalApplication element would be defined using the "end of execution" signal location.

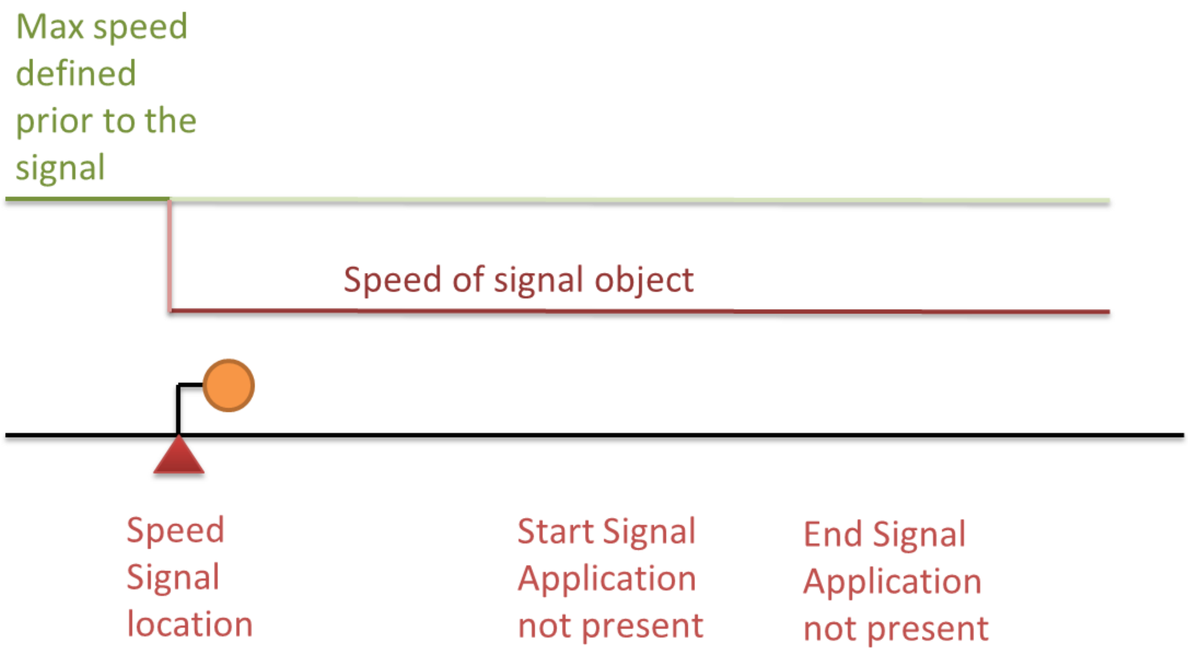
The DAS will consider the location where the restriction given by the signal is to be enforced by the driver by using the optional startSignalApplication and endSignalApplication elements. Even though they are optional, these elements should be provided as often as possible, as they remove the possible ambiguities that could arise by applying the two following default rules:

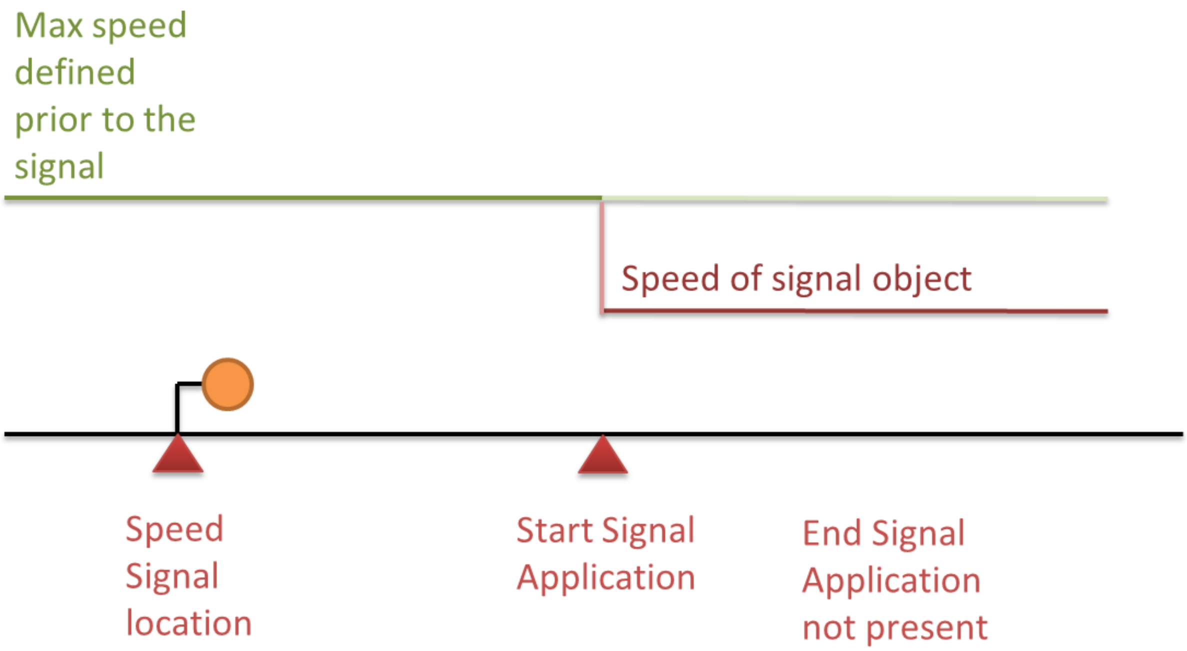
- If the startSignalApplication element is not present, the signal location will be used instead;
- If no endSignalApplication is declared for a MaxSpeed or AvoidSpeedRange signal, the information applies up to the next signal encountered by the train. In some cases, (e.g. StopBeforePassingSignal) the endSignalApplication is ignored. See Table 3 and the following figures for a clarification.

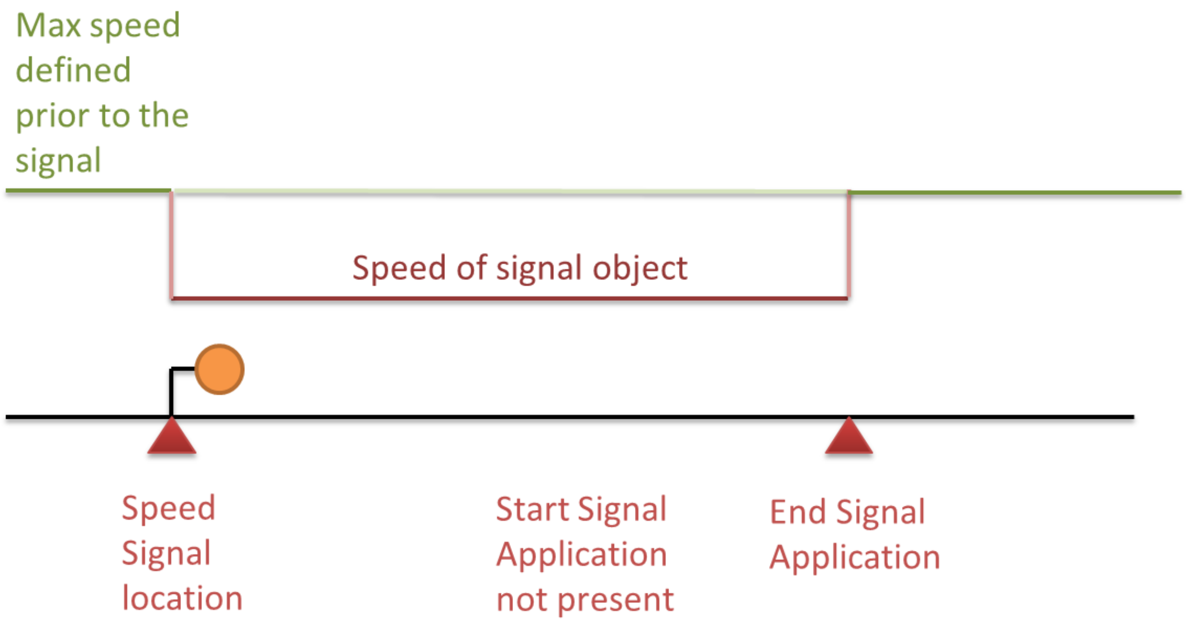
Table 3: Example of signal application for different signal types

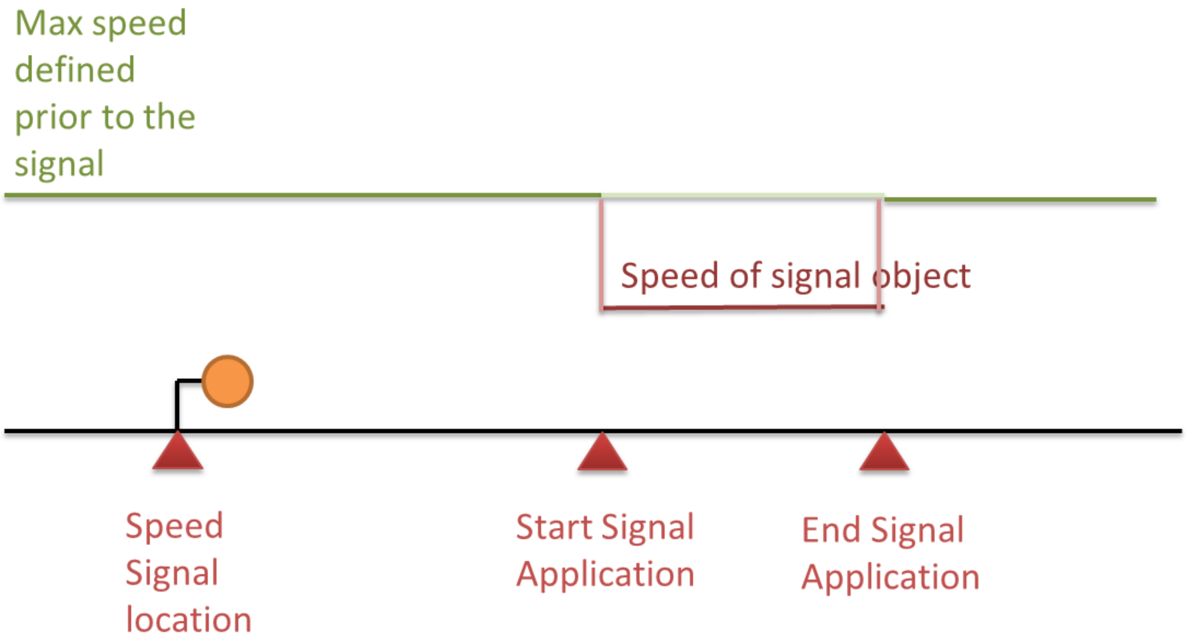
Signal location	Start Signal Application	End Signal Application	MaxSpeed	Avoid Speed Range	Stop Before Passing Signal
Mandatory	Not Present	Not Present	Applied from the signal location of the signal to the next signal location	Applied from the signal location of the signal to the next signal location	Applied at the signal location
Mandatory	Present	Not Present	Applied from the start signal application to the next signal location	Applied from the start signal application to the next signal location	Applied at the start signal application

Signal location	Start Signal Application	End Signal Application	MaxSpeed	Avoid Speed Range	Stop Before Passing Signal
Mandatory	Not Present	Present	Applied from the signal location of the signal to the end signal application	Applied from the signal location of the signal to the end signal application	Applied at the signal location
Mandatory	Present	Present	Applied from the start signal application to the end signal application	Applied from the start signal application to the end signal application	Applied at the start signal application
Mandatory	Present, but after “End signal application”	Present, but before “Start signal application”	This is to cover a “start signal application” located on another path but close enough to affect the speed behaviour of the train on the common path. Train shall slow down to be at signal’s max speed limitation at the “Start signal application”, but the slowing down is revoked by the “End signal application” and the “Start signal application” no longer has an effect on the Journey of the train. The “End signal application” may have a new “Speed signal” in the same place.	See text for MaxSpeed	Applied at the signal location

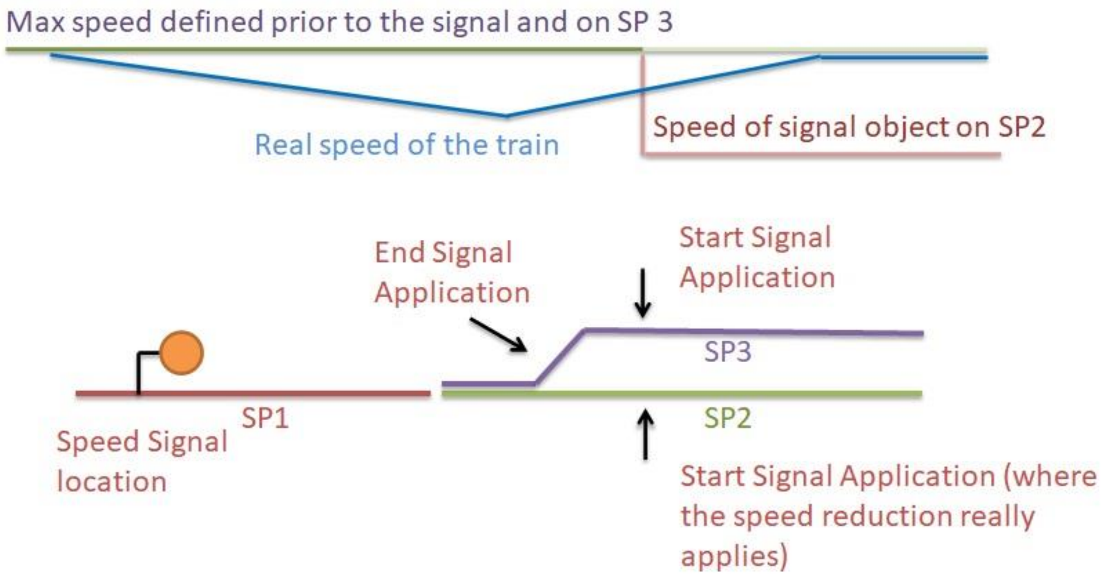

Max speed defined prior to the signal
Speed of signal object
Speed Signal location
Start Signal Application not present
End Signal Application not present


Max speed defined prior to the signal
Speed of signal object
Speed Signal location
Start Signal Application
End Signal Application not present


Max speed defined prior to the signal
Speed of signal object
Speed Signal location
Start Signal Application not present
End Signal Application


Max speed defined prior to the signal
Speed of signal object
Speed Signal location
Start Signal Application
End Signal Application

The speed signal is on the common path SP1. The train will go on SP3, the slow area is on SP2, but close enough to affect the speed on the common path. This is why «Start Signal Application» is on SP3 too, at the same distance as it would be on SP2. On the common path, the train must slow down, as it would aim to reach the max speed of the signal at its «Start Speed Application» point, but the «slowing down» phase is canceled by an early appearance of «End Signal Application» before the «Start Signal Application» as soon the train arrives on SP3, since it is not concerned anymore by the speed limit.



2.1.6.3 Signal represented by multiple signal objects

When a physical signal gives multiple pieces of information to a driver, this will be described using multiple signal objects. To show that they are in fact the same physical signal, a unique physical signal identifier should be used in all the signal objects. Additional signal object identifiers shall be used to distinguish these different signal objects.

2.1.6.4 Physical signal characteristics

As mentioned earlier, the intention of the SFERA standard is not to describe what signals look like in real life. Rather, it is possible in SFERA to express some physical characteristics of the signal, such as type and offset. This information can be used in object recognition or can be shown on a DMI to make the driver more aware of the surroundings and the position of the train. The signal type describes the type of signal and is network-specific (free text) and so the SFERA project has not identified a standardised solution. With the offset it is possible to specify if the signal is on the left or the right side of the track and its height from the track. Note that the horizontal offset should be reversed if the direction of application on an SP is reversed. Also, there is the possiblity of using network-specific parameters to specify further characteristics.

2.1.6.5 Examples of signal coding in the SP

This paragraph contains a few examples of the way the SFERA signal object could be used in the SP.

2.1.6.5.1 Entering a station with a point limited to 30km/h in the deviating position – open exit signal (France)

When a train enters a platform where the speed is limited to 30 km/h, it could encounter the sequence shown in Fig. 13:

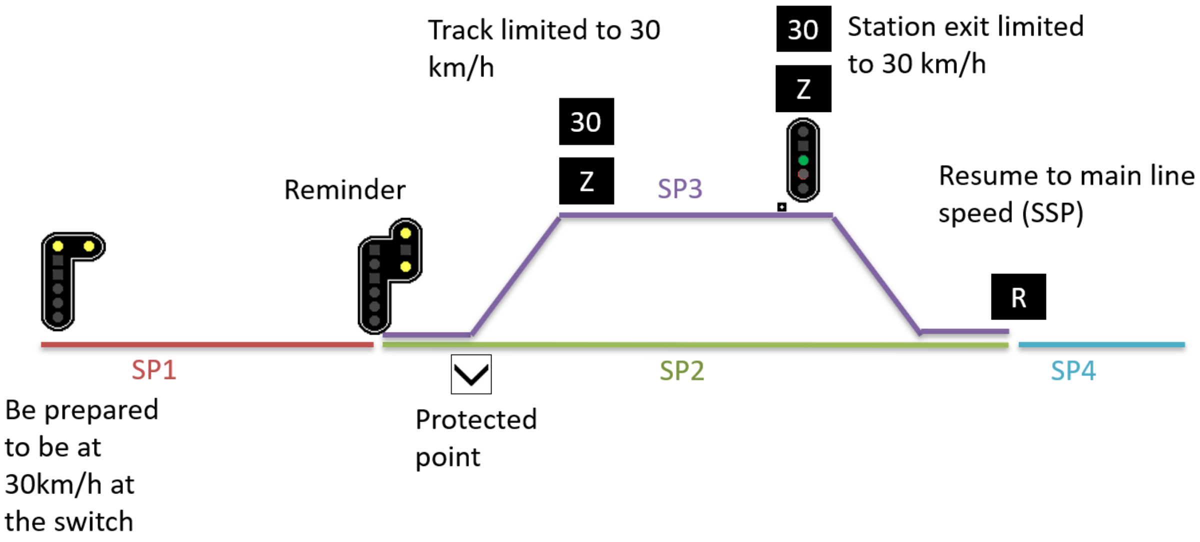


Figure 13: Entering a station with a point limited to 30 km/h in the deviating position – open exit signal

The first sequence to brake down to 30 km/h can be described by a signal object in the SP, as shown in Table 4 (this is the least restrictive aspect that can be seen by the driver on this itinerary). The signal object will take into account the fact that the drivers will respect the 30 km/h speed limit 200m before the protected point, because it is imposed by the KVB.

Table 4: Signal object - Entering a station with a point limited to 30km/h in the deviating position – open exit signal

Signal Object Attributes			**Value**
Physical Signal Identifier			Signal number xx
Signal Object Identifier			-
Location			Signal Location
Signal Application	StartSignalApplication	Distance	Point of application – 200m
Signal Application	StartSignalApplication	startSignalApplicationMustBeClearedBy	Head

Signal Object Attributes			**Value**
Signal Application	EndSignalApplication	Distance	Point of application
Signal Application	EndSignalApplication	endSignalApplicationMustBeClearedBy	TrainEnd
Signal Application	Signal_ATP_Systems		KVB
Signal Application	Signal_TrainCategory		-
Signal Application	Signal_TrainLength		-
Signal Application	ApplicableIfTrainRunsOnSP		SP3
Signal Information	StopBeforePassingSignal		-
Signal Information	MaxSpeed		30 km/h
Signal Information	AvoidSpeedRange		-
Signal Information	TractionTypeTransition		-
Signal Information	PowerConstraint		-
Signal Information	StationAheadSign		-

After the point of application, the SP can describe the 30 km/h track limitation using the SSPs on SP3.

2.1.6.5.2 Alternative current neutral zone (France)

When a train encounters a neutral zone, four fixed signals will be presented to him, as shown in Fig. 14:

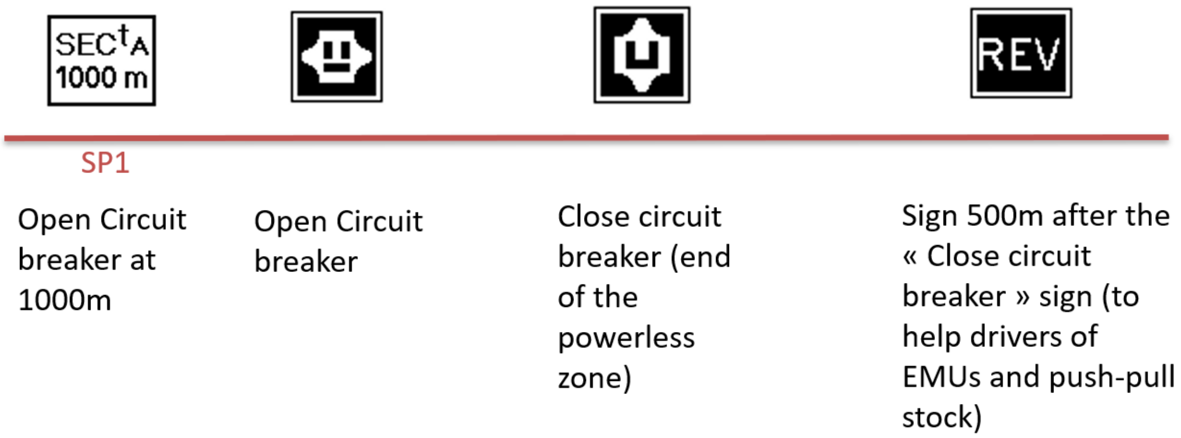


Figure 14: Alternative current neutral zone

An open circuit breaker informs drivers that they have to perform the necessary operations to open the circuit breaker before the “open circuit breaker” signal. The IM and RU need to estimate the distance at which most drivers effectively stop the traction. In this case it will be considered that it is 500m before the open circuit breaker signal.

The driver is allowed to close the circuit breaker when the last pantograph of the train has passed the “close circuit breaker” signal. The IM and RU need to estimate the distance at which most drivers effectively reengage traction (they take a margin after the pantograph passes the sign, and a certain time is needed to restart acceleration). In this case it will be considered that it is 300m after the last pantograph has effectively passed the signal.

This sequence can be described by two signal objects in the SP (shown in Tables 5 and 6), one describing the circuit breaker opening sequence, and the other describing the circuit breaker closing sequence.

Table 5: Alternative current neutral zone – signal for circuit breaker opening sequence

Signal Object Attributes			**Value**
Physical Signal Identifier			Signal number nn
Signal Object Identifier			-
Location			“Sect at 500m” Location
Signal Application	StartSignalApplication	Distance	“Open Circuit Breaker” location – 500m

Signal Object Attributes			**Value**
Signal Application	StartSignalApplication	startSignalApplicationMustBeClearedBy	Head
Signal Application	EndSignalApplication	Distance	-
Signal Application	EndSignalApplication	endSignalApplicationMustBeClearedBy	-
Signal Application	Signal_ATP_Systems		-
Signal Application	Signal_TrainCategory		-
Signal Application	Signal_TrainLength		-
Signal Application	ApplicableIfTrainRunsOnSP		-
Signal Information	StopBeforePassingSignal		-
Signal Information	MaxSpeed		-
Signal Information	AvoidSpeedRange		-
Signal Information	TractionTypeTransition		Powerless
Signal Information	PowerConstraint		-
Signal Information	StationAheadSign		-

Table 6: Alternative current neutral zone – signal for circuit breaker closing sequence

Signal Object Attributes			**Value**
Physical Signal Identifier			Signal number mm
Signal Object Identifier			-
Location			"Close Circuit Breaker" Location
Signal Application	StartSignalApplication	Distance	"Close Circuit Breaker" location + 300m
Signal Application	StartSignalApplication	startSignalApplicationMustBeClearedBy	TrainLastPantograph
Signal Application	EndSignalApplication	Distance	-
Signal Application	EndSignalApplication	endSignalApplicationMustBeClearedBy	-
Signal Application	Signal_ATP_Systems		-
Signal Application	Signal_TrainCategory		-
Signal Application	Signal_TrainLength		-
Signal Application	ApplicableIfTrainRunsOnSP		-
Signal Information	StopBeforePassingSignal		-
Signal Information	MaxSpeed		-
Signal Information	AvoidSpeedRange		-
Signal Information	TractionTypeTransition		25000V AC 50Hz
Signal Information	PowerConstraint		-
Signal Information	StationAheadSign		-

2.1.6.5.3 Transition to an unelectrified track for bimodal trains (France)

When a bimodal train leaves an electrified track, it will encounter two signals. The first one will inform drivers that they must prepare to lower the pantograph, and the second one is the point at which the pantograph shall be lowered (see Fig. 15).

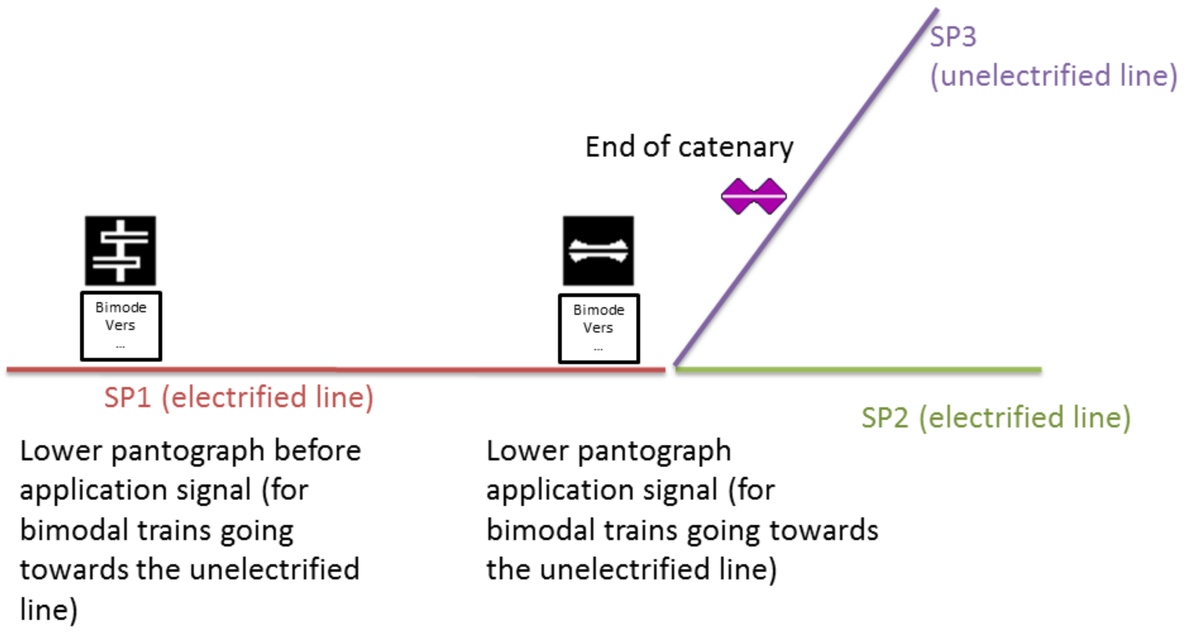


Figure 15: Transition to an unelectrified track for bimodal trains

A signal object in the SP (see Table 7) can be used to describe this sequence. As for the previous example, the driver will have stopped accelerating 200m before the application signal.

Table 7: Signal object – Transition to an unelectrified track for bimodal trains

Signal Object Attributes			**Value**
Physical Signal Identifier			Signal number zz
Signal Object Identifier			-
Location			Signal Location
Signal Application	StartSignalApplication	Distance	Application signal – 200m
Signal Application	StartSignalApplication	startSignalApplicationMustBeClearedBy	TrainLastPantograph
Signal Application	EndSignalApplication	Distance	-
Signal Application	EndSignalApplication	endSignalApplicationMustBeClearedBy	-

Signal Object Attributes		**Value**
Signal Application	Signal_ATP_Systems	-
Signal Application	Signal_TrainCategory	-
Signal Application	Signal_TrainLength	-
Signal Application	ApplicableIfTrainRunsOnSP	SP3
Signal Information	StopBeforePassingSignal	-
Signal Information	MaxSpeed	-
Signal Information	AvoidSpeedRange	-
Signal Information	TractionTypeTransition	Self-propelled
Signal Information	PowerConstraint	-
Signal Information	StationAheadSign	-

The signal object does not limit its application to bimodal trains because:

- self-propelled trains will disregard this traction type transition;
- electric trains should not be routed on the unelectrified line.

2.1.6.6 CancelPreviousSignal Element

A signal's state may change and for that reason, the element "CancelPreviousSignal" has been introduced.

In the following example, S1 indicates that at the S2 position, the signal is closed.

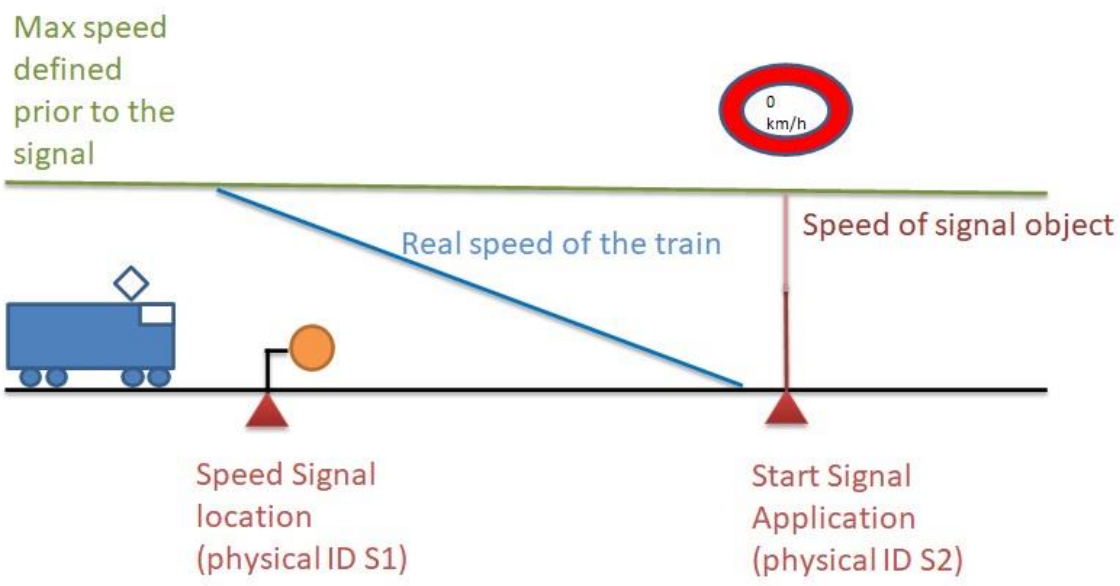


Figure 16: Signal S2 closed

This may be indicated with a signal complex type instance having Signal_ID.signal_ID_Physical set to S1, StartSignalApplication set to the location of S2 and SignalInformation set at MaxSpeed = 0.

As shown in the following figure, after the train passes S1, the state of S2 changes: it is open, and the speed is set to 60 km/h.

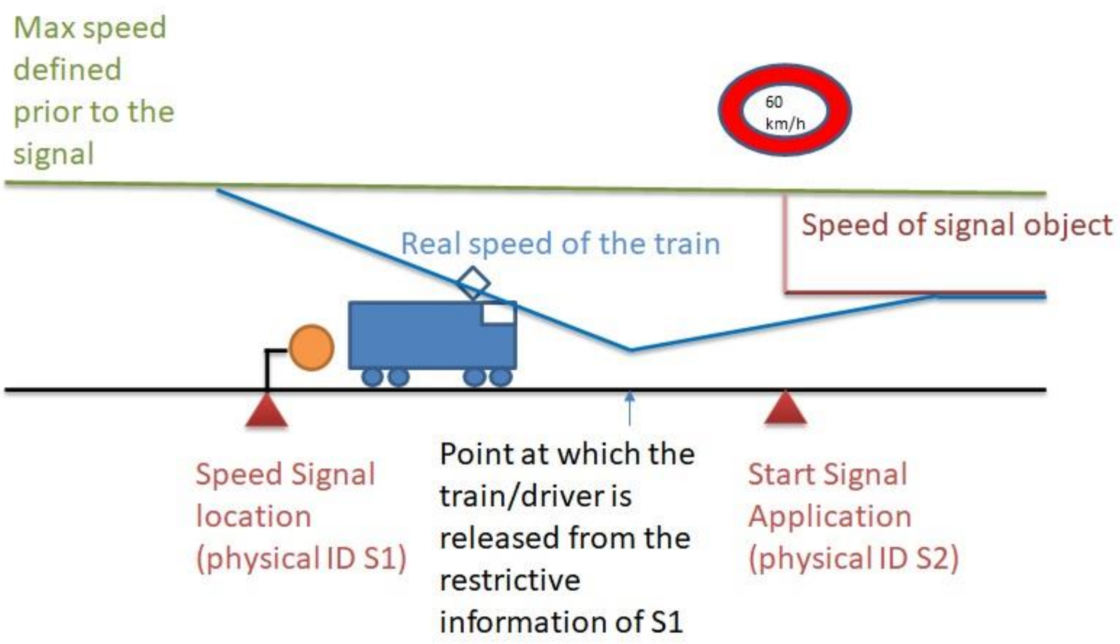


Figure 17: Signal S2 changed from closed to 60

The initial information displayed by S1 is no longer valid. For this reason, new Signal information shall be sent. Signal_ID.signal_ID_Physical would be set to S2, SignalInformation would be set to 'CancelPreviousSignal', whereby the attribute signal_ID_Physical should be set to S1 (the signal whose prior indication was cancelled). In some situations, however, the train might not be authorised to accelerate before it reaches a certain distance before S2 (for example, if this signal is not in the visual contact field of the driver) and, in that case, the attribute 'distanceBeforeSignal' can be set accordingly.

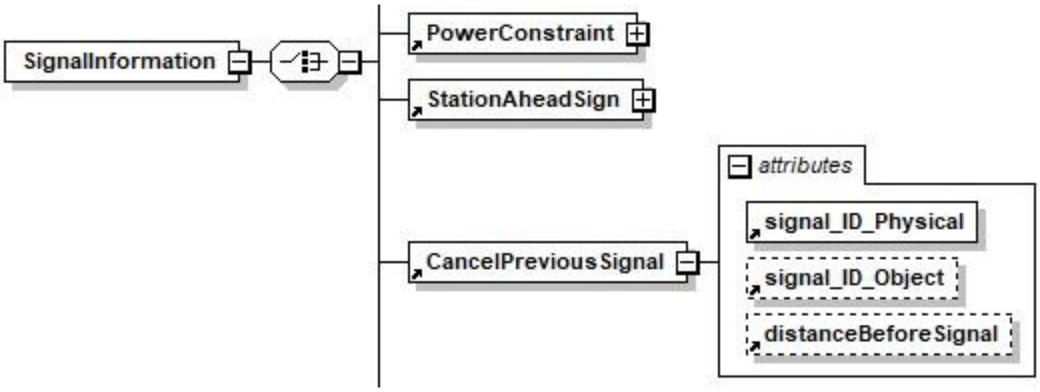


Figure 18: CancelPreviousSignal Element of SignalInformation in the Complex Signal Type

Of course the new speed information of 60 km/h shall also be set. This is possible according to Section 2.1.6.3 (Signal represented by multiple signal objects) by declaring an additional signal object for S2 having the element MaxSpeed set to 60.

2.1.7 Technical identifiers

2.1.7.1 SP_Zone: NID_C and RICS

SFERA uses two identifiers to define zones, with two main objectives:

- Helping to determine the entity in charge of providing SFERA information on a given SP, and the SP that immediately follows the given SP.
- Uniquely identifying elements in those given zones.

The two identifiers are:

- The NID_C, which is an ERTMS variable used to define a country's network, or a subdivision of this network (usually one or more lines in a given country). The NID_C is only defined on ERTMS tracks, and is used in SUBSET-126 for ERTMS/ATO. ERA is in charge of maintaining this list to keep values unique;
- The IM_ID, which uses the RICS code to define the IM in charge of the zone. This identifier has been added to SFERA to cover the lines that do not have a NID_C identifier. The UIC is in charge of maintaining this list to keep values unique. All IMs (even if not affiliated to the UIC) can obtain a RICS code.

In the SFERA XSD, the NID_C and the IM_ID are optional. However, the following rules and guidelines apply:

- At least one of the two zone identifiers (NID_C or IM_ID) shall be defined for every SP;
- In an ERTMS zone, it is recommended to define the NID_C for cross-compatibility with SUBSET-126;
- It is recommended to systematically provide the IM_ID, even if a NID_C is also defined for this zone. By doing this, the RU will only need to maintain a correspondence table linking RICS codes to the addresses providing SFERA data on those networks. Otherwise, the RU would also need to add a NID_C/SFERA address correspondence table;

- For the Segment Profiles, Timing Points and balises (elements from SUBSET-126), virtual balises, signals, kilometre reference points and unprotected level crossing stops (SFERA specific elements), the identifier:
  - shall be unique within a given NID_C;
  - shall be unique within a given IM_ID for the Segment Profiles not covered by a NID_C.
- For virtual balises, signal objects, kilometre reference points and unprotected level crossing stops, the identifiers shall be unique in the SP of their location.

2.1.7.2 Coding of borders in Segment Profiles

Each SP describes the track within one IM or NID_C. In the SP, it is possible to indicate in the Next_SP_Zone field that the next segment will be provided by another IM (or contained in another NID_C). This triggers the request of both a new Journey Profile and segment data from the next IM in the first sequence of the handover between two DAS-TS (see Use Case DX3). The JPs obtained for all IMs that are traversed should be joined to create a full JP for the DAS. Only after passing the border, can the connection to the first IM be dropped. IMs usually have arrangements in place already to provide timings for the first Timing Point past the border. These mechanisms should be used to update the timing windows in upcoming countries.

For the SPs closest to the border, the IMs shall implement periodic checks that the data fits exactly up to the border and there are overlapped regions. Furthermore, care should be given to the coding of extended objects such as tunnels.

This is valid for C-DAS-O and C-DAS-C modes. It is to be noted that the SP is optional and not essential for C-DAS-C operations. If the IM does not deliver the SP in this mode, the next IM shall be obtained by the DAS by other means (e.g. manual selection).

2.1.8 Valid and invalid SPs

An SP can be “valid” or “invalid”. These definitions can be found in SUBSET-125, 10.1.3.23. In SUBSET-125 it is explained that:

“*10.1.3.23 The Segment Profile shall indicate its status to the ATO-OB (Segment Profile Status):*

a) Valid: specifies that the data sent correspond to a SPReq sent;

b) Invalid: the requested Segment Profile is not available in the ATO-TS.”

The usage of major and minor versions of an SP shall be the same as in ERTMS/ATO.

2.1.9 Direction of SP

A portion of track can obviously be traversed in two directions. To model a portion of track, it is possible to create two Segment Profiles (the first for one direction, and the second for the opposite direction). But it is also possible to create a single SP, and to specify in the Journey Profile (with the attribute SP_Direction under SegmentProfileList) if the SP is traversed in the “Nominal” direction or in the “Reverse” direction. In the reverse direction, the DAS shall make sure that it considers that locations are relative to the nominal entry point.

Some SP objects are valid independently of the direction of the train (e.g. virtual balises). Some objects, however, may only be valid for one direction. Therefore, if the intention is to use a single SP to model both directions, for some elements the direction of application (nominal, reverse or both) shall be specified. This applies to Signal and to UnprotectedLevelCrossingStop: in these cases, the attribute directionOfApplicationOnSP is used. Also, some objects are relevant to both directions, but the effect on the calculations will be different depending on the direction (for example, the gradient has to be considered as uphill in one direction and downhill in the other direction). The DAS-OB implementation needs to take these cases into consideration.

2.1.10 "Empty" Segment Profiles

It is possible to include "Empty" Segment Profiles, i.e. Segment Profiles that only contain mandatory data (SP_ID, SP_Zone, SP_VersionMajor, SP_VersionMinor and SP_Length). The length of the SP can also be estimated if no precise value is available.

Empty SPs can be useful to "Fill" the data for parts of a journey where no ATO or C-DAS-O operation is possible because there is not enough data; in these parts of the journey, only DAS-C, "Timetable" driving mode or no DAS at all can be used.

To provide consistent functionalities of DAS-O up to the boundaries of this SP, care shall be taken regarding the virtual balises provided. There should be virtual balises on the extremities of either:

- the "Empty" SP(s) (preferable if multiple empty SP(s) are linked in the undescribed zones);
- the SP(s) surrounding the "Empty" SP(s).

Additional data associated with virtual balises could be added in these SPs to provide assistance to the driver, without permitting full C-DAS-O information (as this would not be sufficient to fulfill the minimum conditions expressed in Section 9.9). For example, Timing Points can be included, and time objectives defined in the Journey Profile. The DAS could therefore calculate the delay of the train in stations along the route.

On those SPs provided with minimum data, care shall be taken by the IM to document the DAS operation modes that are unavailable, using the "DAS_ModesUnavailability" type SP_Areas. In the general case, a restriction of the "DAS not connected to ATP" and "GoAx" driving modes will be used. Note that declaring the unavailability of the "BoardAdviceCalculation" (DAS-O) would prevent "Timetable" Driving mode from being used.

2.2 Choices in the arrangement of the data

2.2.1 Choice of segments

A journey in SFERA is split into several segments. In the implementation, the IM needs to choose how to split up the data of the country's rail system. This section gives considerations for choosing an implementation.

Section 2.1.7.1 explains how SPs are identified using the NID_C, IM_ID and SP_ID variables. The length of a segment is open for implementation. The choice will depend on a number of factors.

In ERTMS/ATO, all information on the segments is directly coupled to a piece of track. In SFERA, this is not the case, as Class B systems and usual signal aspects are supported: a piece of track may have different expected signal values, depending on the remainder of the route; a red exit signal, or different routes through switches may trigger different usual entry signal aspects.

For S-DAS, the segments can be used to efficiently transmit many JPs, while only sending track information once. C-DAS may be optimised to limit changes if Timing Points or routes are changed. The choice of SPs will be reflected by these possibilities.

One can, for example, do the following:

1. Split segments into parts at each switch, with separate segment IDs for different expected signals and stopping positions;
2. Split segments into parts at each switch, with different expected signals referencing the different downline SPs;
3. Split segments into parts for free lines, and approaches and exits to each platform at a station, with separate segment IDs of different expected signals and stopping positions;
4. Split segments into parts at each named Timing Point, with segments connecting different platforms via several routes using the usual signal aspects for such a route;

5. Split segments into parts per stop at each service, with segments connecting different platforms via several routes using the usual signal aspects for such a route.

The list above is ordered on the expected average segment length (shortest first). With many short segments, the probability for re-use for different journeys is highest. This will stress the SP cache. Profiles at each stop in a service may be a good starting point if precision and bandwidth is limited.

As a variation on the four given main choices, an IM can also use segments for the least restrictive signal aspects for a route and use TemporarySignalConstraints in the JP for more restrictive signals that will be passed before and on deviating routes.

The SFERA protocol itself, once again, leaves these choices as implementation freedom, based on the precision, update frequency and bandwidth.

ERTMS/ATO mentions three reasons that make it mandatory to use a different SP:

- change of timing zone (i.e. different value for UTC_Offset);
- boundary between two adjacent ATO-TS;
- overflow of iterators.

The first two reasons are self-evident. The last reason needs further consideration. SUBSET-126 uses 5 bits for the iterators. This means that an SP to be translated to SUBSET-126 can only have 32 different values (e.g. for curves or gradients). This limit is not applicable to elements not existing in SUBSET-126, such as the virtual balises.

2.2.2 Estimated number of segments and data sizes

The number of segments in a journey and their sizes vary based on many choices. SFERA implementations should be able to handle the amount of data in the countries where the DAS is implemented.

Implementations of S-DAS and use of low-bandwidth communication channels can define boundaries. In the worst-case scenario, an IM could split up the journey at each sign, signal or switch. In most countries, this would lead to a few segments per kilometre outside stations and 40 segments per kilometre inside station areas. The fetching of segments should be organised so that it is efficient for the longest train journey in a country.

The amount of data, on the other hand, depends more on the distance between virtual balises (see Section 2.1.1), changes in gradients and changes in curvature. The amount of data can grow quickly without a noticeable gain in precision of the DAS advice. For the IM´s reference, Section 9.1 gives some relationships. Designers, however, should be able to handle very detailed profiles as well.

2.2.3 Splitting segments inside stations

Splitting segments within stations may be done at a Timing Point if the stopping position does not depend on the train length. In that case, the location of the Timing Point may be given in both the arrival segment and the departure segment. This means that the Timing Point information should be included twice in the JP, as the data is linked to two distinct SPs. The IM shall ensure consistency of the Timing Point in that case. If the DAS needs platform information from both SPs, the DAS may join the information to obtain information on the whole platform.

Alternatively, an SP may be created for part of a platform. In this case, the time of the Timing Point can be provided and the stopping position may be train-length-dependent.

2.2.4 How many virtual balises: The choice of SP_MaxLateralDeviation

SP_MaxLateralDeviation is the maximum allowed lateral deviation of any middle point on the used track to straight lines connecting the virtual balises.

The choice of SP_MaxLateralDeviation for a country or region should be made following data availability and bandwidth. Infrastructure Managers can usually at least provide GNSS positions with a resolution of a few hundred meters at Timing Points. These points then represent the centre of the stations, and stopping positions can be derived relative to this centre.

Alternatively, Infrastructure Managers may provide positions of hectometre indicators beside the track or even at the centre of the track for each of the separate tracks.

Some Infrastructure Managers may even provide data every 10 metres. This will increase the amount of data transmitted, often with a small increase in DAS precision.

The density of virtual balises is mainly of importance in curves, as the interpolation may yield small differences in lateral and longitudinal positioning. Along straight sections it does not affect the quality of the advice that DAS gives.

The SFERA protocol does not prescribe what resolution should be used, and implementations may vary per Infrastructure Manager.

3 Modelling train features in SFERA

In addition to the railway infrastructure, the DAS needs to know the features of the train running the service. This is mostly defined in the Train Characteristics (TC) object.

3.1 Operational management of Train Characteristics

In the SFERA protocol, the RU is responsible for providing correct TCs. Each set within an RU will be identified by a TC_ID. TC_IDs only need to be unique within an RU, which, however, can get the information from another source (e.g. a leasing company). In that case, it is possible to refer to the keeper’s VKM instead. This avoids duplicating data in case a train is shared by different RUs. The TC_ID shall be unique within the context of an RU or a VKM.

For passenger operation, a full list of possible train consists will usually be delivered, possibly with some variants for degraded operation. The RU can provide the DAS with the TC_IDs of all train consists that would be likely for a given route. On-screen selection can be done using values from displayName.

For freight trains, there will be many different values for different lengths and weights. Each of them shall have a different TC_ID which is unique within an RU.

Some specific requirements apply to the data for the TC use cases (Train Characteristics Change), see Appendix A.

If a change is required (e.g. different trainMaxSpeed or braked weight percentage due to operational constraints or trainWeight or trainCategoryCode due to loading for freight trains), a trainCharacteristicsChange is sent with the original TC_ID and all changed fields known to the DAS. The DAS can directly start processing with this new set.

The DAS-TS shall use this to compile a new set of TC and a new unique TC_ID. This new TC_ID will be sent to the DAS as an acknowledgement. This acknowledgement will force read-only DAS to be synchronised. If changes are reverted, the original TC_ID may be sent (without changed parameters) in order to let the DAS-TS (and the other read-only DAS) know the new state.

There is a specific case in which trainCharacteristicsChange is used even though it is not an actual “change” of characteristics. It takes place when a SUBSET-126 Status Report containing the variable L_TRAIN (i.e. the length of the train) is sent from board to ground. That variable is translated with trainLength inside trainCharacteristicsChange (see Appendix G).

The following paragraphs detail which fields require updates to be sent, as they will have a functional impact on the TMS.

3.1.1 Attributes modified in failure modes

Some failures in the train yield to degraded performance of the train. Some examples are presented in the table below. Generally, maximumSpeed, enginePower, brakingForceCurve, and tractionForceCurve will be updated to reflect the actual performance. The TrainCharacteristicsChange element in the SR is used to notify the TMS of this behaviour, so that the planning may be updated to reflect this degraded performance.

Table 8: Examples of failure modes

Failure	Description	Action
ATP system failure or change	In case of an ATP system failure, it might be possible to switch to another ATP system, or the train shall proceed without ATP system.	ATP system shall be updated, eventually maximumSpeed shall be reduced according to the alternative ATP system or in case of missing ATP system. This may also affect brakingForceCurve.
Reduced braking percentage	In case of a brake failure on some of the rolling stock, the braking capabilities become reduced.	brakingForceCurve shall be recalculated, maximumSpeed may also be reduced according to national regulations.
Reduced maximum speed	Other reasons can also result in a reduced maximum speed, for example a minor failure of one of the wagons.	maximumSpeed shall be reduced.
Locomotive or unit failure	In case of multiple locomotives or multiple units, one or more of them can be shut off. If at least one locomotive or unit is able to produce traction, the train can in some cases proceed with its journey, but it will be less powerful.	The number of engines and enginePower is reduced, tractionForceCurve and brakingForceCurve shall be recalculated.
Engine failure	If in case of an engine failure, the locomotive or multiple unit has at least one engine left and can proceed with its journey, the train can produce only part of its power.	enginePower is reduced, tractionForceCurve and brakingForceCurve shall be recalculated.
Pantograph failure	If a pantograph has been taken out of service but the train can still run on its route (e.g. a bimodal train can travel in self-propelled mode).	All TractionForceCurves of the traction mode(s) needing this pantograph are deleted.

3.1.2 Adapting the rollout coefficients

If only part of a train can proceed with its journey, the following elements may be adapted:

- trainLength;
- trainWeight;
- enginePower;

- rolloutCoefficientA, rolloutCoefficientB, rolloutCoefficientC;
- brakedWeightPercentage;
- rotatingMassFactor;
- [number of] engines;
- [number of] wagons;
- axleLoadCategory.
-

3.2 Formulas to be used in TCs

The running resistance *R* of the train is calculated with the formula

$$R = A + B \cdot v + C \cdot v^2$$

where

- rollout coefficients of the train
- v speed of the train (in m/s: this is the speed in km/h divided by 3.6)

In the coefficients A, B and C, the rollout characteristics of the engines and the wagons have to be combined.

For the additional tunnel resistance R_T, the Filipovic formula is used:

$$R_T = \left[\frac{2 + n_e}{3} \frac{a_e}{Q^{b_e}} + n_w \frac{a_w}{Q^{b_w}}\right] \cdot v^2$$

where

- Q aerodynamic cross section of the tunnel (taken from the SP)
- v speed of the train (in m/s)
- n_e number of engines
- n_w number of wagons
- a_e, b_e coefficients for engines
- a_w, b_w coefficients for wagons

The following typical values of the coefficients can be used, unless otherwise stated in the TCs:

- a_e = 1140 [$kg \cdot m^{(2b_e-1)}$]
- b_e = 1,48 [-]
- a_w = 662 [$kg \cdot m^{(2b_w-1)}$]
- b_w = 1,75 [-]

3.3 Processes between RUs and IMs for the IM to know the TCs

3.3.1 Data consistency of TCs with TAF/TAP

Several fields in the TCs are also present in TAF/TAP messages for the train run. These values are safety critical: for instance, the length of a train will determine if it fits a certain track. Decisions were made in the past based on assumptions for shorter trains that have contributed to accidents[1]. Similarly, the contents (and thus the weight) of a freight train need to be fully declared.

Therefore, any update in length and weight should come from the RU via the TAF/TAP process. The IM should provide the data defined in the TAF/TAP message in the JP, including updates of rolling coefficients and traction/braking curves. This is similar to the process for TrainCharacteristicsChange (see the TC use cases).

The driver can change the length or weight of a train in SFERA. This may be required if the TAF/TAP update in the RU takes some time. There should be sufficient flexibility to allow the correct DAS advice to be set immediately. All processes within the RU should be set up to consolidate the DAS and TAF/TAP data as quickly as possible. Once the TAF/TAP is updated, the IM shall transmit the new length, weight, optionally train type for EMU-DMUs, rolling coefficients and traction/braking curves as soon as possible. The driver should then verify that the data in all systems is consistent with the actual values for the train.

Note that SFERA has many train features that determine the running times but are not present in TAF/TAP. These include fields that may be adjusted due to degraded operation, as described in Section 3.1.1.

3.3.2 Data exchanged when a driver selects a new rolling stock type from a list

When the driver selects a new type of rolling stock from a list (to declare that the TC associated with the JP corresponds to a different train type), the DAS will send a TrainCharacteristicsChange element that only includes “rollingStockType”.

This will work on two conditions:

- The IM and the RU will need to have predefined a list of rolling stock types.
- The rolling stock types effectively define the rolling stock. This is fairly simple for multiple units, but for locomotive-hauled stock a choice has to be made between the following options:
  - have rolling stock types that include their length, which could be practical for operations where the rakes are of fixed lengths;
  - ask the driver to enter a length in addition to the rolling stock type and send both to the RU.
  -

3.4 Interpretation for selecting a train type on the DAS

The train type in the TCs is specifically meant to allow a driver to quickly select a generic train type. This section guides RUs in making the list of available entries for train types.

For passenger trains, it is best practice to show the possible combinations of EMUs. For freight trains, presets for train types may be used. Length and weight will be more variable for freight however.

The RU can take several approaches to train type. For DMUs and EMUs, it is reasonable to give each consist combination a specific train type (e.g. VIRM-4, VIRM-6, VIRM-4+4). If degraded modes occur regularly, these may be preset as well (e.g. VIRM-4+ VIRM-4 without traction). For hauled passenger trains, it makes sense to preset several numbers of carriages. For unlisted lengths, the length and weight may be specified. For freight trains, presets on number and type of locomotives, lengths and (empty or loaded) weights make sense. The length and weight will be adjusted regularly. Integration with a Train

[1] See for example https://erail.era.europa.eu/occurrence/NL-4737-50-1/Trains-collision.

Control Management System (TCMS) makes sense in order to load this data for freight trains.

For unlisted compositions, length and weight may be specified. Roll-out coefficients are specified as a force. The DAS should adjust these as well (e.g. proportionally) when a length or weight is changed. Note that the traction and brake force curves may depend on the train length as well. SFERA implementations themselves should not change the roll-out coefficients if the length and weight are adjusted. This may give rise to unphysical train models. There are no SFERA TCs to automatically scale roll-out coefficients with length and weight.

4 Providing advice to trains

This section provides details and considerations on the advice sent to the train in order to determine the driving profile. It is the most sizeable section of this appendix for the obvious reason that the main objective of DAS is to provide advice to trains and SFERA has been built with a wide array of features to make this process as flexible as possible.

The section will start with the main elements affecting the driving profile of a train, i.e. the speed constraints, the traction and braking information and the power elements. It continues with notes on how to handle some special JP elements and it ends with considerations on S-DAS and on ATP systems.

4.1 Speed constraints

The objective of this section is to explain the different elements to take into account when defining the maximum speed authorised for a given train on a given portion of infrastructure. The way the deceleration is calculated to respect these speeds is not covered in this section.

In SFERA, the speed range(s) authorised for a train journey are to be defined using the following elements:

- The TCs, including the train's maximum speed, the train category(ies), and the ATP systems available.
- The static speed profiles (SSP) (included in the SP), which give the maximum speed of a portion of track, with the possibility of having different values for different train categories.
- The axle load speed profile (included in the SP).
- Additional speed restrictions (included in the JP).
- The signal elements (included in the SP and the JP), which can limit the speed, exclude a speed range, or impose a stop before passing a signal. Regarding the SP speed profiles, they also have the added value of permitting the point where the signal is located to be separated from the area where the speed restriction shall be applied.

The following paragraphs will describe the way the individual elements are to be read, and then how they interact with each other.

4.1.1 TCs impacting speed

4.1.1.1 Train Maximum Speed

The TCs (under TC_Features) contain a trainMaxSpeed attribute. This specifies a speed which shall be respected at all times on the parts of the JP to which this TC element is applicable.

4.1.1.2 Train Categories

Train categories related to speed constraints can be expressed via four attributes:

- NC_Train: this train category is defined by ERTMS SUBSET-026 and divides trains into three categories:
  - Passenger Train,
  - Freight Train in "P" braking mode,
  - Freight Train in "G" braking mode.
- NC_CDTrain: this characterises trains per cant deficiency, as defined by ERTMS SUBSET-026;
- trainCategoryCode: this code is specific to SFERA. It is used for countries where the categorisation of trains regarding speed and signals is finer-tuned than that permitted by the NC_Train and NC_CDTrain attributes. This attribute is a freeform string to take into account the variety of categories found in different IM regulations.
- AxleLoadCategory: this train category is defined by ERTMS SUBSET-026 (M_AXLELOADCAT).

4.1.1.3 ATP Systems

The TCs contain a list of all the ATP systems that are installed on the train. The ATP systems that are usable on a portion of a train's journey are deduced by defining the intersection between the list of ATP systems in the TC and the list of ATP systems supported on this portion of track given by the ATP_Systems element in the SP_Characteristics.

If the TC and the SP have multiple ATP systems in common on a portion of track, the ATP system to be considered is defined by a priority configuration included in the DAS device.

4.1.2 Determining the Static Speed Profile (SSP)

4.1.2.1 ATP Systems

In contrast to ETCS SUBSET-126, multiple SSPs can be described for an SP. This possibility has been provided to cover situations where different speeds are allowed for different ATP systems.

This shall be reflected in the description of the SP in one of the following ways:

- The SP includes an SSP without an ATP system, and optionally additional SSPs for the ATP systems with different speed characteristics. If this is the case, the DAS will consider the basic SSP for any ATP system not in the optional SSPs;
- The SP does not include an SSP without an ATP system. If this is the case, the IM shall make sure that all the ATP systems available on the SP (ATP_Systems in the SP Characteristics) have an associated SSP. If this is not the case, the DAS device will not be able to calculate an SSP for an undeclared ATP System.

An ATP system can be associated with one - and only one - SSP.

4.1.2.2 Train Categories

Once the SSP to consider has been determined, the DAS will need to determine the maximum speed on the SP regarding the different train categories. These are given by the StaticSpeedProfileStart and StaticSpeedProfileChange elements (under SP_Characteristics / StaticSpeedProfile) which are structured identically.

A basic static speed profile (SSP_Speed, equivalent to V_Static from SUBSET-126) is present, and will be used for a train on every portion where no specific SSP applies.

Specific SSPs can be added if on a portion of track some categories of trains shall respect a lower or higher maximum speed. These specific SSPs can be of two types:

- Cant Deficiency SSP: A given SSP is applicable to the train if the cant deficiency category matches the cant deficiency category in the TC;
- Other Specific SSP: In SFERA, two fields are present to determine the train category: SSP_NC_DIFF and trainCategoryCode. SSP_NC_DIFF comes from the ERTMS/ATO SUBSET-126, while trainCategoryCode is specific to SFERA. At least one of those attributes shall be present in the SSP description. The SSP is applicable to a given train if all the attributes present in the given SSP match the NC_Train and trainCategoryCode of the TC.

If a train is simultaneously affected by a cant deficiency SSP and “Other specific SSP”, the priorities from SUBSET-026 apply.

4.1.2.3 SSP_Front

As defined for ETCS, the DAS shall take into account the SSP_Front attribute for all SSPs (basic or specific). This parameter defines if the train shall respect a speed limit when the front of the train passes the point of application or liberation, or if it shall wait for the end of the train to have passed that point.

4.1.3 Axle Load Speed Profile (ALSP)

An ALSP shall be respected if the axleLoadCategory included in the TC is equal or superior to the axleLoadCategory present in the ALSP, as defined for ETCS (M_AXLELOADCAT).

4.1.4 Additional Speed Restrictions

An additional speed restriction can be detailed in the Temporary Constraints of the JP, as defined for ETCS.

4.1.5 Signals with an impact on speed

Section 2.1.4 describes the way signal elements are defined in SFERA. As mentioned earlier, signals contain a part with information (e.g. a maximum speed) and an optional part with the applicability of the signal, which may include a start and/or an end for the signal application.

In the cases where the “startSignalApplication” is present, and is different from the signal location, the DAS will consider that the signal location is the position of a distant signal. Consistent with the RU and IM regulations and the applicable ATP system, this element could be taken into account by the DAS implementation in defining the moment when braking starts by adopting one of the following strategies (see Fig. 19):

- start braking exactly at the signal location and stop braking at the application point;
- start braking before the signal location and stop braking at the application point;
- start braking exactly at the signal location and brake as defined in the TC, even if the speed limit is attained long before the start signal application point;
- start braking before the signal location and brake as defined in the TC, even if the speed limit is attained long before the start signal application point;
- just in time to respect the speed at the start signal application point;
- etc.

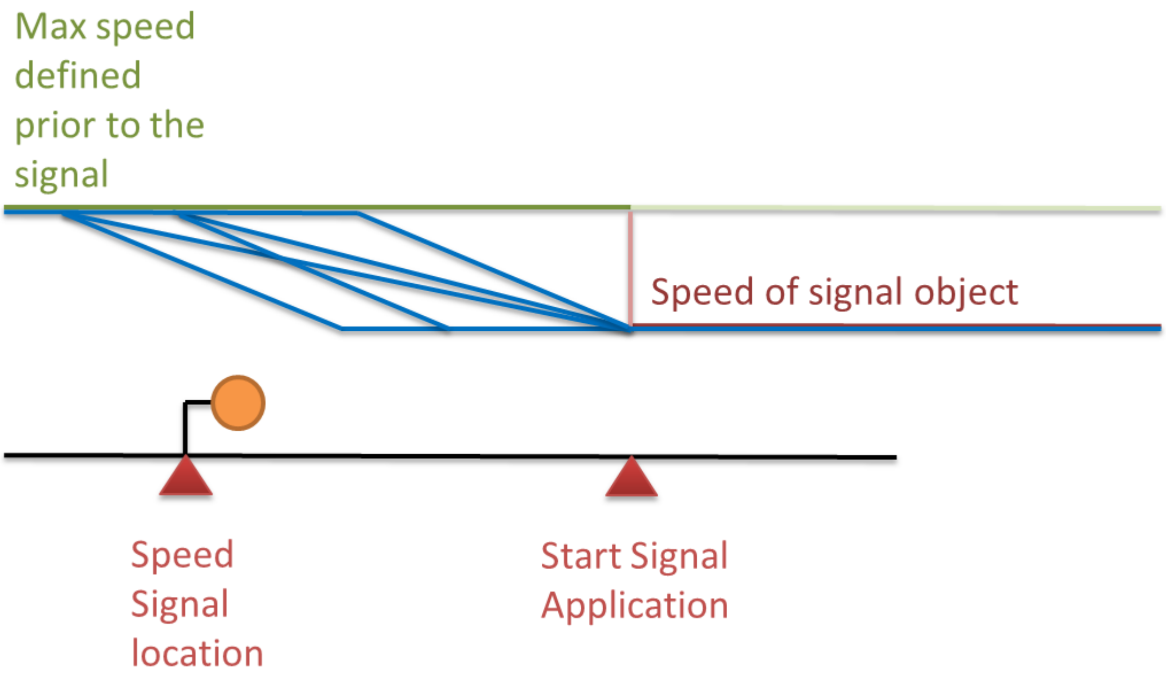


Figure 19: Possible braking strategies between the signal location and start signal application point

4.1.5.1 Changes to permanent signals

If the TMS has anticipated an operational situation (e.g. a conflict with another train), where the information displayed to the driver would be different from a permanent signal described in an SP, signal objects can be introduced in the JP with the "TemporarySignalConstraints" attribute set to "ExistingSignalAspectModification". Also, if the IM has temporarily cancelled a signal, signal objects can be introduced in the JP with the "TemporarySignalConstraints" attribute set to "CancelledSignal".

In both cases, if the signal in the JP:

- only has the physical signal identifier, all signals in the SP with the same physical signal identifier are disregarded (this means that the signal objects from the SP which are still valid would need to be resent in the JP);
- has both identifiers (signal_ID_Physical and signal_ID_Object), only the signal object in the SP having this pair of identifiers will be disregarded.

4.1.5.2 Temporary signals

Temporary signals are additional signals installed in the track, usually in situations where there is work being done on the track, and are declared by using TemporarySignalConstraints in the JP and defining the "temporarySignalConstraintType" as "TemporarySignal".

The physical signal identifier used for a temporary signal shall be different from the other physical signal identifiers used by other signals on this SP.

4.1.5.3 Examples of signal coding in the JP

This paragraph contains a few examples of the way the SFERA signal object could be used in the JP.

4.1.5.3.1 Braking to a scheduled stop in a station with closed exit signal (Germany)

In front of a scheduled stop, the PZB (Punktförmige Zugbeeinflussung) has an effect on the braking curve if the exit signal of the station is predicted to be closed at the moment the train is passing the distant signal before the stop. The speed is restricted to a certain speed after a certain period of time after passing the distant signal. The speed is dependent on the train type.

In this case, temporary signals will need to be described in the JP (see Table 9) to explain the impact of the closed exit signal on the driver. Each signal will give a different speed according to the different categories.

Table 9: Signal object – Braking to a scheduled stop in a station with closed exit signal

Signal object attributes			**Value**
Physical signal identifier			Identifier of the fixed signal in the SP if it exists
Signal object identifier			To be used if different effects apply to different trains (ATP system, train category, etc.)
Location			Signal location
Signal application	StartSignalApplication	Distance	Point of application of the speed restriction imposed by the PZB
Signal application	StartSignalApplication	startSignalApplicationMustBeClearedBy	TrainHead
Signal application	EndSignalApplication	Distance	Stopping point
Signal application	EndSignalApplication	endSignalApplicationMustBeClearedBy	TrainHead
Signal application	Signal_ATP_Systems		PZB
Signal application	Signal_TrainCategory		Different per signal object
Signal application	Signal_TrainLength		-
Signal application	ApplicableIfTrainRunsOnSP		Could be necessary depending on

Signal object attributes		**Value**
		the SP segmentation.
Signal information	StopBeforePassingSignal	-
Signal information	MaxSpeed	Speed imposed by PZB for a given train type
Signal information	AvoidSpeedRange	-
Signal information	TractionTypeTransition	-
Signal information	PowerConstraint	-
Signal information	StationAheadSign	-

4.1.5.3.2 Acceleration after a scheduled stop in a station with closed exit signal (Germany)

After the stop, the PZB prevents the driver from accelerating against the closed exit signal (even if the signal has been cleared in the meantime). In this case the speed is controlled to be under the "restrictive speed" for a certain distance. This restrictive speed is normally 25 km/h. So, in case the exit signal is prognosed to be closed, a signal constraint should be sent to the train indicating the reduced speed from the stopping point to the end of the restriction.

In this case, a temporary signal will need to be described in the JP (see Table 10) to explain the impact on the acceleration of the driver.

Table 10: Signal object – Acceleration after a scheduled stop in a station with closed exit signal

Signal object attributes		**Value**
Physical signal identifier		Identifier of the fixed signal in the SP if it exists
Signal object identifier		To be used if different effects apply to different trains (ATP system, train category, etc.)

Signal object attributes			**Value**
Location			Signal location
Signal application	StartSignalApplication	Distance	Stopping point position
Signal application	StartSignalApplication	startSignalApplicationMustBeClearedBy	TrainHead
Signal application	EndSignalApplication	Distance	End of “restrictive speed”
Signal application	EndSignalApplication	endSignalApplicationMustBeClearedBy	TrainHead
Signal application	Signal_ATP_Systems		PZB
Signal application	Signal_TrainCategory		-
Signal application	Signal_TrainLength		-
Signal application	ApplicableIfTrainRunsOnSP		Could be necessary depending on the SP segmentation.
Signal information	StopBeforePassingSignal		-
Signal information	MaxSpeed		25 km/h
Signal information	AvoidSpeedRange		-
Signal information	TractionTypeTransition		-
Signal information	PowerConstraint		-
Signal information	StationAheadSign		-

4.1.5.3.3 Stop before a faulty level crossing (Germany)

Sometimes the secured state of the level crossing is not correctly given back to the signal box, so that the main signal before the level crossing cannot be cleared. In this case, the train shall stop in front of the level crossing. If the level crossing is secured and free of road vehicles, the train can pass and accelerate immediately.

A signal can be declared in the JP to describe this constraint (see Table 11).

Note that the safety-relevant operational information shall be given to the driver by separate means ("Befehl" in Germany).

Table 11: Signal object – Stop before a faulty level crossing

Signal object attributes			**Value**
Physical signal identifier			Identifier of the fixed signal in the SP if it exists
Signal object identifier			To be used if different effects apply to different trains (ATP system, train category, etc.)
Location			Signal location
Signal application	StartSignalApplication	Distance	Level crossing location
Signal application	StartSignalApplication	startSignalApplicationMustBeClearedBy	Head
Signal application	EndSignalApplication	Distance	-
Signal application	EndSignalApplication	endSignalApplicationMustBeClearedBy	Head
Signal application	Signal_ATP_Systems		-
Signal application	Signal_TrainCategory		-
Signal application	Signal_TrainLength		-

Signal object attributes		**Value**
Signal application	ApplicableIfTrainRunsOnSP	-
Signal information	StopBeforePassingSignal	0 min (driver should depart immediately if the level crossing is secured and free of road vehicles)
Signal information	MaxSpeed	-
Signal information	AvoidSpeedRange	-
Signal information	TractionTypeTransition	-
Signal information	PowerConstraint	-
Signal information	StationAheadSign	-

4.1.5.3.4 Entering a station with a point limited to 30km/h in the deviating position – closed exit signal (France)

This example uses the same track described in Section 2.1.6.5.1. When entering that track, the exit signal could be closed, for example because another train is leaving the station from another track. In that case, the driver will encounter a “caution” signal (one yellow light), which is added to the 30 km/h reminder (see Fig. 20).

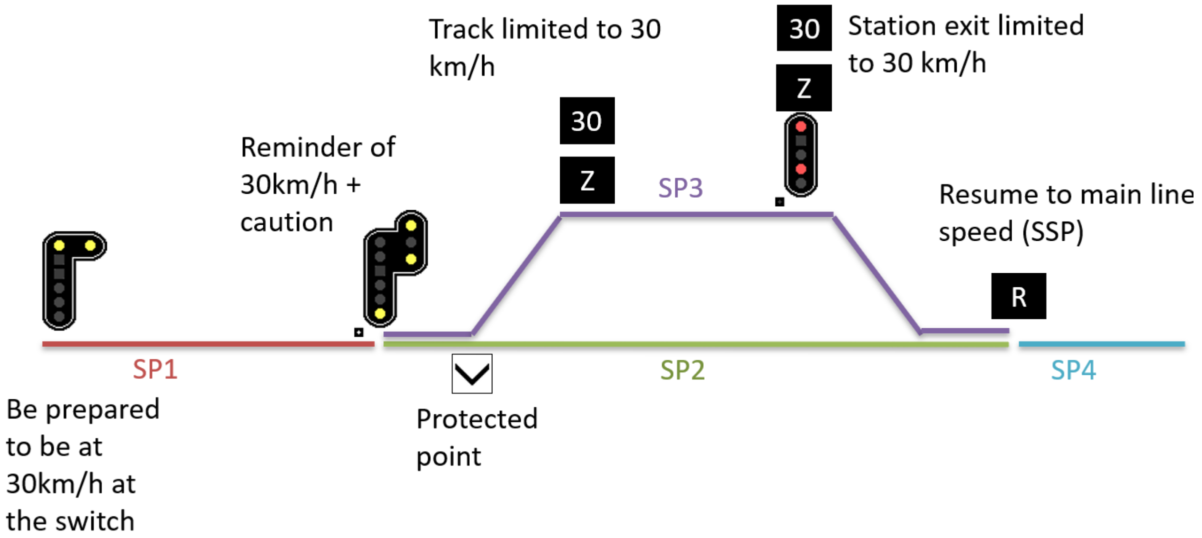


Figure 20: Entering a station with a point limited to 30km/h in the deviating position – closed exit signal

In this case, the distance between the “danger” signal (two vertical red lights: mandatory stop) and the exit point is short. A 10 km/h approach speed will be imposed at a distance estimated in this case at 200m before the “danger” signal. It will be applicable until the train passes the “danger” signal (which should show a “clear” aspect when the train leaves the station).

A temporary signal object will need to be created in the JP (see Table 12) to describe the additional constraints on driving imposed by the restrictive driving sequence enforced by the “caution” signal.

Table 12: Signal object – Entering a station with a point limited to 30km/h in the deviating position – closed exit signal

Signal object attributes			**Value**
Physical signal identifier			Signal number yy
Signal object identifier			To be used if different effects apply to different trains (ATP system, train category, etc)
Location			“Reminder + caution” Signal location
Signal application	StartSignalApplication	Distance	Location of “danger” signal – 200m

Signal object attributes			**Value**
Signal application	StartSignalApplication	startSignalApplicationMustBeClearedBy	TrainHead
Signal application	EndSignalApplication	Distance	Location of “danger” signal
Signal application	EndSignalApplication	endSignalApplicationMustBeClearedBy	TrainHead
Signal application	Signal_ATP_Systems		KVB
Signal application	Signal_TrainCategory		-
Signal application	Signal_TrainLength		-
Signal application	ApplicableIfTrainRunsOnSP		-
Signal information	StopBeforePassingSignal		-
Signal information	MaxSpeed		10 km/h
Signal information	AvoidSpeedRange		-
Signal information	TractionTypeTransition		-
Signal information	PowerConstraint		-
Signal information	StationAheadSign		-

4.1.6 Priority of elements in calculating the speed constraints

By following the different rules in the preceding paragraphs, the DAS will have been able to define the following elements that apply to a given train journey:

- The maximum speed allowed for the rolling stock,
- A maximum speed profile that combines the applicable SSPs on the different portions of track, and if applicable to the train journey,
- Additional speed restrictions,
- Axle load speed profiles,

- The combined effects on speed of the signals.

At each point of the train's journey, the DAS will consider advising a speed that is simultaneously compatible with all these constraints.

It is possible that the application areas of several signal objects are in intersection. In this case the following rules should be used for the intersecting areas:

- if multiple "MaxSpeed" signal objects are applicable, the lowest speed is taken into account;
- if multiple "AvoidSpeedRange" signal objects are applicable, the ranges are added;
- if multiple "StopBeforePassingSignal" signal objects are applicable at the same location, the longest "stopTime" is taken into account.

In Fig. 21, the coloured zones make explicit the speed ranges in which the speed profile calculated by the DAS shall be maintained.

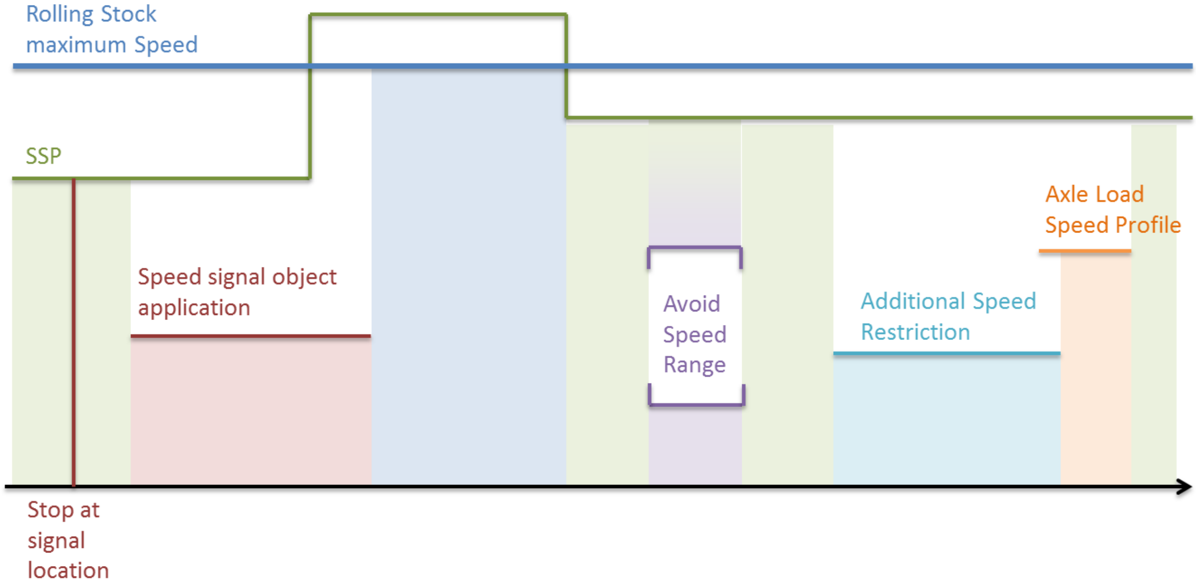


Figure 21: Interactions between different elements affecting speed

4.1.7 Advised speed from ground to board

Occasionally, an advised speed over a certain area can be transmitted from ground to board as a temporary "AdvisedSpeed" constraint. This speed, if compatible (and it should be) with the speed constraints presented in Section 4.1.6, should be included in the driving profile.

4.2 Traction and braking

This section describes how to handle traction type constraints and traction/braking force curves in SFERA.

4.2.1 Traction type constraints

This paragraph describes the elements to be taken into account to determine the traction type to be applied on the different parts of the journey, and the elements that can impose that no traction power can be used.

4.2.1.1 SFERA elements that help determine the traction type

The traction type to be applied on different parts of the train journey is deduced by the DAS by combining the different information provided by:

- The Train Characteristics:
  - available traction types for the rolling stock;
- The Segment Profile:
  - The type of track electrification system,
  - The description of signals that implement points of transition between traction types that are not directly linked to the points where electrification changes. For example, bimodal trains do not necessarily transition between self-propelled and catenary/third rail operation at the exact point where a non-electrified line leaves an electrified line.
- The Journey Profile:
  - signals for traction type transitions that are added or modified for a short period of time.

In addition, some elements will be used to define sections of the track where no power can be applied by trains running in electric mode. These are provided by:

- The Train Characteristics:
  - The position of the last point of the train that shall clear the powerless section before the driver can finish taking into account a power constraint (main switch on/off, lower/raise pantograph, etc.).
- The Segment Profile:
  - The powerless sections on this track to be applied by trains running under ETCS;
  - The signals that order power constraints for the drivers of non-ETCS trains, as non-ETCS trains usually apply the restriction on areas larger than the pure powerless section of the infrastructure declared in the SP_Characteristics as the part under CurrentLimitation with maxCurValue equal to 0;
  - Current limitations.
- The Journey Profile:
  - Temporary current limitations;
  - Temporary signals used to describe temporary power constraints.

4.2.1.2 Data necessary to define the traction type

This section presents the data to be taken into account to define the traction type to be applied on the different parts of a train's journey.

TC: Available traction types

The traction types that a given type of rolling stock can operate under are deduced by listing all the traction types that appear in at least one of the traction force curves available in "TC_Features".

A train that has traction force curves for the self-propelled mode and at least one electrification type shall be considered as a bimodal train in the following paragraphs.

SP: Type of electrification

On an electrified line, the DAS can identify the type of electrification of the infrastructure by using the "RatedVoltage" element in the "SP_Characteristics" (this element corresponds to the "Power Voltage Start" element of ERTMS/ATO SUBSET-126). This element uses the "RatedVoltageStart" and

“RatedVoltageChange” elements, which give the rated voltage value, and the additional NID_CTRACTION attribute, which gives more precision on the traction system (AC/DC, frequency, etc.).

On parts of the SPs where the voltageValue in the “RatedVoltageStart” or “RatedVoltageChange” element is set to 0, only rolling stock with a self-propelled traction source can be taken into account by the DAS calculations.

SP and JP: Traction Type Transition Signal

SFERA includes the option of implementing “TractionTypeTransition” signals in the SP and the JP. Its main aim is to be able to specify transitions that are imposed to drivers/trains at points of the infrastructure that are different from the “RatedVoltage” transitions of the SP.

As specified in Section 2.1.6.1, the location of the signal and the application area can be differentiated.

If the “StartSignalApplication” is not present, the signal location is used to define the point where the transition occurs (see Table 13).

The “EndSignalApplication” is not considered for this type of signal.

Table 13: Signal application for TractionTypeTransition signal

Signal location	Start signal application	End signal application	Traction type transition signal
mandatory	not present	N/A	applied at the signal location
mandatory	present	N/A	applied at the start signal application

SP and JP signals can be cancelled or modified using the same mechanisms as signals affecting speed, described in Section 4.1.5.1.

If the “TractionTypeTransition” is only applicable for trains going towards a certain line (e.g. bimodal trains leaving an electrified line towards a non-electrified line), the “ApplicableIfTrainRunsOnSP” element can be specified.

4.2.1.3 Data necessary to define powerless sections for trains in electric mode

Once the traction type on the different parts of a journey has been defined, the DAS will need to identify sections of the track where no electric power can be used by the train, although the applicable traction type is electric. These sections can be present at traction type transition points, or at any intermediate point of a type of electrification.

TC: Position of the last pantograph

In some regulations, drivers may be asked to re-engage traction after a certain point of their train has cleared the end of a powerless section.

This point can be described in the TC by using the attribute “PositionOfLastPantograph”.

SP: Powerless sections

Permanent powerless sections on electrified tracks can be described in the SP_Characteristics using the CurrentLimitation element and entering maxCurValue equal to 0. The powerless section is strictly the part of the track where the train is not allowed to draw power.

It is to be noted that this element does not specify the action to be taken by the driver/train (pantograph lowering or main switch opening).

SP and JP: PowerConstraint Signal

The “PowerConstraint” signal object is to be implemented in addition to the powerless section in two cases:

- When the area of application by the driver is larger than the strict area given by the ETCS powerless section. This is often the case in non-ETCS use cases, when drivers can be required to start the procedure at the distant signal. This means that the traction can be cut off a few hundred meters before the powerless section;
- To provide the type of action the driver/ATO shall apply. The information that can be given is:
  - pantograph up/down
  - main switch on/off
  - traction on/off
  - HVAC on/off

As specified in section {sec:applicability-of-a-signal-object}, the location of the signal and the application area can be differentiated. Table 14 shows the different cases of application of the PowerConstraint signal.

If the “StartSignalApplication” is not present, the signal location is used to define the first point of application.

A single PowerConstraint signal can be self-supporting to define a power constraint zone if at least the EndSignalApplication is present and the signal information is “pantograph down”, “main switch off”, “traction off” or “HVAC off”. The “powerConstraintInstruction” is applied from the startSignalApplication (or the signal location by default if the startSignalApplication is not declared).

If the “endSignalApplication” element is not present, two signal objects shall be used: a first one that specifies the beginning of the power constraint (e.g. pantograph down), and a second one that specifies the end of the power constraint (e.g. pantograph up).

Table 14: Application of the PowerConstraint signal

Signal location	**Start signal application**	**End signal application**	**PowerConstraint Signal**
mandatory	not present	not present	Applied from the signal location of the signal to the next PowerConstraint signal of the same type of “powerConstraintInstruction”
mandatory	present	not present	Applied from the start signal application to the next PowerConstraint signal of the same type of “powerConstraintInstruction”
mandatory	not present	present	Applied from the signal location of the signal to the end signal application
mandatory	present	present	Applied from the start signal application to the end signal application

As for the speed-linked signals, the start and end of the signal application shall be defined according to the way the whole system reacts. If, for example, the regulations of a given country impose that the driver closes the main switch 200m before the signal that indicates the beginning of a zone with a power constraint, the start and end signal application elements shall be described accordingly.

Overlapping power constraint zones are to be avoided if possible when constructing SFERA data. If overlapping occurs, the types of power constraints are to be considered separately. This means that a “pantograph down” power constraint is not cancelled if a “switch on” power constraint is encountered further on before a “pantograph up” signal is encountered.

SP and JP: Current limitation

If current limitations declared for the SP and the JP are applicable on the same portion of track, then the lowest value is to be taken into account. The current limitation in the SP is also used to specify the parts of the track where it is not allowed to take or return current to the OCL.

4.2.1.4 Priority rules for traction type and powerless sections

Rules in determining the applicable traction type

Self-propelled trains: For solely self-propelled trains (diesel, hydrogen, battery, etc.), the “RatedVoltage” and “TractionTypeTransition” elements are not useful.

Mono and multi-current electric trains: The DAS will only be able to perform calculations on the sections where the “RatedVoltage” matches one of the traction types of the TC.

Bimodal trains at transitions between electrified and non-electrified tracks: The transition of a multi-current bimodal train between different electrification types is done by the same rules as for multi-current trains (see above).

The SFERA protocol considers the hypothesis that if a bimodal train starts its journey on an electrified track with a compatible voltage, it will start using the electric source.

When a bimodal train transitions from an electrified part of its journey to a non-electrified part, the DAS will consider that the transition to self-propelled mode occurs:

- at the point described by a “self-propelled”/“TractionTypeTransition” signal that is encountered before the “RatedVoltage” transition;
- at the location of the “RatedVoltage” transition, if no “TractionTypeTransition” signal is encountered by the train.

When a bimodal train transitions from a non-electrified part of its journey to an electrified part, the DAS will consider that the transition to electric mode occurs at the first “TractionTypeTransition” signal with the electric traction type encountered by the train. If no signal is present, the train is considered to continue its route in self-propelled mode.

Rules in adding the powerless sections

Once the basic traction type to be applied on different parts of the journey has been determined, the DAS will need to take into account the powerless sections that the driver will encounter.

The DAS will consider the power constraints imposed by signal objects and “0” value current limitations, and will take into account the combination of the effects of these elements to determine the portions of the journey where no traction power can be applied.

4.2.2 Traction Force Curves

4.2.2.1 What Traction Force Curves represent

A traction force curve describes the dependence between traction force and speed of the train under normal operating conditions (e.g. dry and clean track conditions). The traction force should consider all working traction units of the train. Multiple traction force curves can be defined for one train: one for each different traction type for the train (AC/DC with different voltages/current, self-propelled (diesel, battery)). Each possible traction type for the train shall have at least one traction force curve.

Furthermore, the actual voltage on the catenary and the maximum allowed current can influence the maximum traction. Therefore, specific curves for a certain TFC_estimatedVoltage and/or TFC_maxCurrentValue can be given. For diesel trains, the traction may be lower at high altitude. However, this parameter has not been considered in the traction force curve.

Each traction force curve is characterised by an optional identifier, the respective traction type and a list of data points. These data points each contain a pair comprising speed and corresponding traction force.

If a traction force curve is used for calculation on board or on the ground, it is likely that the current speed of the train will not correspond to the exact speed of one data point. In that case, the traction force should be interpolated between the two neighbouring data points.

4.2.2.2 Interactions with voltage and current constraints

The estimated voltage at a certain point and the maximum current on the track can be found in the SP and/or the temporary constraints part of the JP. If there is no suitable traction force curve, the one with the nearest traction type characteristics should be chosen for the calculation. In case the distance to two traction force curves is equal, the one with lower voltage or current should be chosen.

If the combination of current and voltage is not found in one traction force curve, the following rules apply:

1. If a suitable current is found in one or more traction force curves, the nearest voltage is chosen. In case of equal distance to two voltages, the one with lower voltage should be chosen;
2. If a suitable voltage is found in one or more traction force curves, the nearest current is chosen. In case of equal distance to two currents, the one with lower current should be chosen;
3. If neither suitable current nor voltage is found in any traction force curve, the traction force curve with the lowest current and voltage difference rate (in percent) should be chosen.

Therefore, in contrast to the interpolation between speed curves, there should be no interpolation in the voltage or current selection.

A suitable traction force curve should be chosen for the calculation in accordance with the active traction type of the train.

For example, there could be three different traction force curves available: for currents of 600 A, 900 A and 1200 A. If the traction type states 1000 A, then the traction force curve for 900 A is chosen; if the traction type states 300 A, then the traction force curve for 600 A is chosen.

4.2.2.3 Interactions with Adhesion Constraints

A lower force than that given in the traction force curve may be available mechanically. This is expressed as a low adhesion category. The maximum force available due to adhesion is to be determined by first selecting the traction curve with the highest voltage and/or current available. The adhesion limiting force is a percentage of this maximum available force. The percentages are dependent on the adhesion category and are the best expected values for each of the adhesion categories. When a low adhesion category is used, the range that the coefficient of friction can have according to SUBSET-126 shall be verified. A coefficient of friction of 0.15 corresponds with 100% as low adhesion rate.

The maximum force available for the train is then calculated as the minimum of this adhesion limiting force and the traction force curve for the estimated voltage and current.

4.2.2.4 Traction Force Curve example

An example of a traction force curve is shown in Fig. 22, where the x-axis represents the speed (in km/h) and the y-axis represents the traction force (in kN). The different curves are for the same train type but for a different voltage.

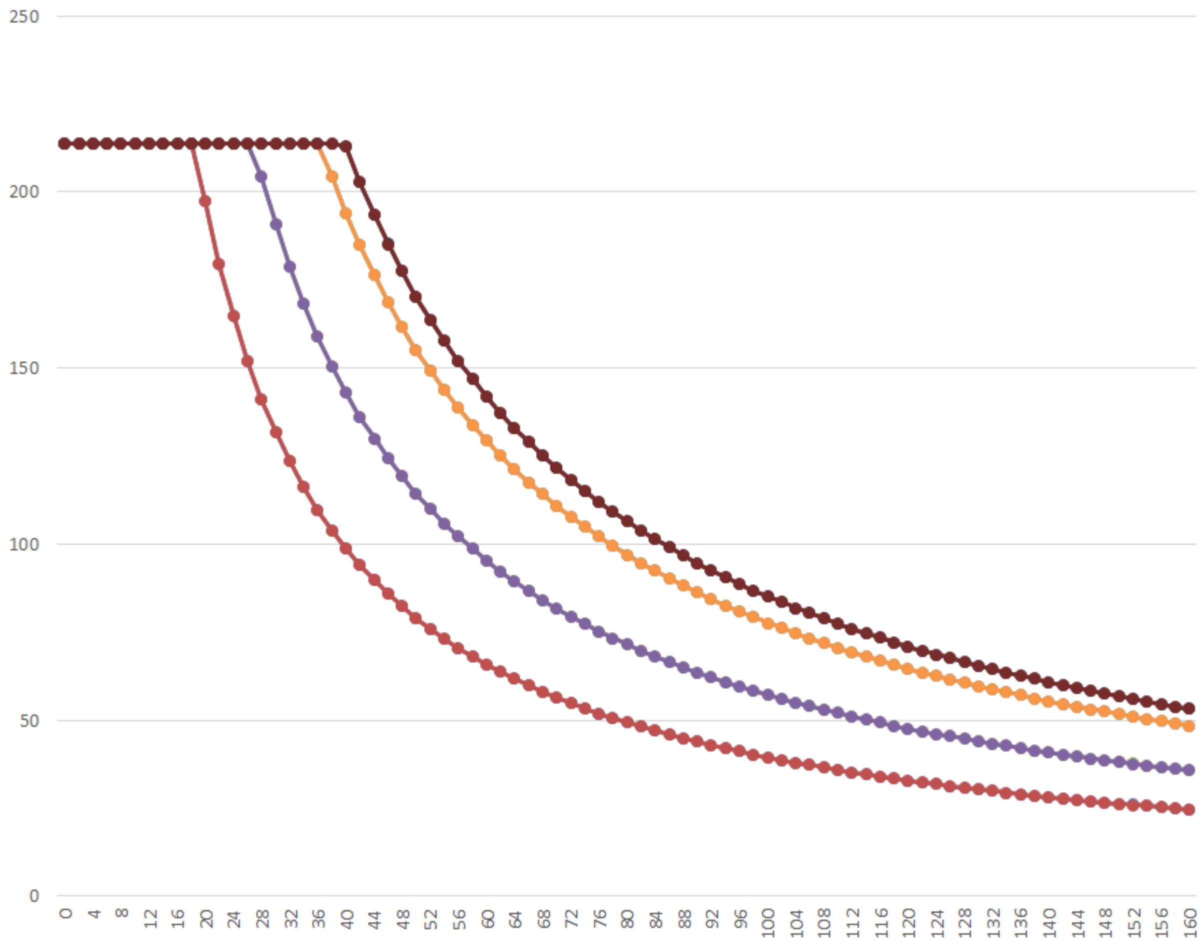


Figure 22: Example of Traction Force Curve

For the example above, Table 15 shows the traction force curve data points for different settings (1200V and 1350V).

Table 15: Data points for traction force curve example

Voltage: 1200V	Voltage: 1200V	Voltage: 1350V	Voltage: 1350V
Speed (km/h)	Traction force (kN)	Speed (km/h)	Traction force (kN)
…	…	…	…
26	152.03	26	213.90
28	141.17	28	204.43
30	131.76	30	190.80
32	123.53	32	178.88
34	116.26	34	168.35
…	…	…	…

4.2.3 Braking force curves

4.2.3.1 What braking force curves represent

A braking force curve describes the dependence of braking force on train speed for usual (comfortable and efficient) braking. This braking force will be much lower than the maximum possible braking force. The braking force should consider all braking units of the train. Multiple braking force curves can be defined for one train for different traction types for that train (AC/DC with different voltages/currents, self-propelled (diesel, battery)) and for different ATP systems on the train. It is recommended that for each possible combination of traction type and ATP system there is a suitable braking force curve. It is also possible that a braking force curve is valid for several traction types and/or several ATP systems.

Each braking force curve is characterised by an identifier, the respective traction type(s), the suitable ATP system(s), the highest adhesion rate and a list of data points. Via the highest adhesion rate, different braking force curves for reduced adhesion can be defined. The data points each contain a pair comprising speed and the corresponding braking force. If the rolling stock of a train can regenerate power, the braking force can be differentiated into total braking force and regenerative braking force.

The regenerative braking force can be used for two things:

- The calculation of a braking phase with only regenerative braking active, in which case only the regenerative braking force is used for the calculation,
- The calculation of the current returned to the overhead line, in which case the total braking force is used for the running time calculation, and the regenerative braking force is used to calculate the regenerated current.

If a braking force curve is used for calculation on board or on the ground, it is likely that the current speed of the train will not correspond to the exact speed of one data point. In that case, the braking force should be interpolated between the two neighbouring data points.

4.2.3.2 Rules for choosing the right braking force curve

A suitable braking force curve should be chosen for the calculation according to the combination of active traction type, active ATP system on the train and the reduced adhesion rate. If no traction type restriction is given in a braking force curve, it can be used for all possible traction types. If no ATP system is given, it can be used for all possible ATP systems. If no maximum adhesion rate is given, it can be used for all adhesion rates.

If no braking force curve for the active traction type is found, the most similar braking force curve should be used (DC, AC, self-propelled). The determination of efficient braking shall take into account the usual amount of regenerative energy that can be returned to the catenary or internal batteries: drivers shall start braking early and lightly if it is possible to optimise regenerative braking.

The braking force curve with an ATP system shall only be used for braking that is enforced by the ATP system (this can be inferred from the signals). This braking curve may have stronger forces than those usually used. For braking to a stop without limiting signal aspects, the braking curve without ATP system shall be used.

If there is no braking force curve with a suitable ATP system, the braking force curve without ATP system should be used (ETCS or national ATP system).

If a braking force curve for only one traction type or for one ATP system is given, this can be used if no similarities are found.

Separate braking curves can be supplied for low adhesion conditions. This reflects the more careful braking by drivers under these conditions. Use the braking force curve with the lowest value for the highest BFC adhesion rate that is higher than the actual low adhesion rate in the temporary constraint. When a low adhesion category is used, verify the range the coefficient of friction can have according to SUBSET-126. A coefficient of friction of 0.15 corresponds to a low adhesion rate of 100%. It is recommended to introduce extra braking force curves for low adhesion rates corresponding with the borders of low adhesion categories in SUBSET-126. Obviously always use the worst conditions, taking into account the braking force curves for traction type, ATP system and low adhesion conditions.

4.2.3.3 Braking force curve examples

Fig. 23 and Fig. 24 are the braking speed force curves of two different trains from SNCF (a high-speed train and a suburban EMU).

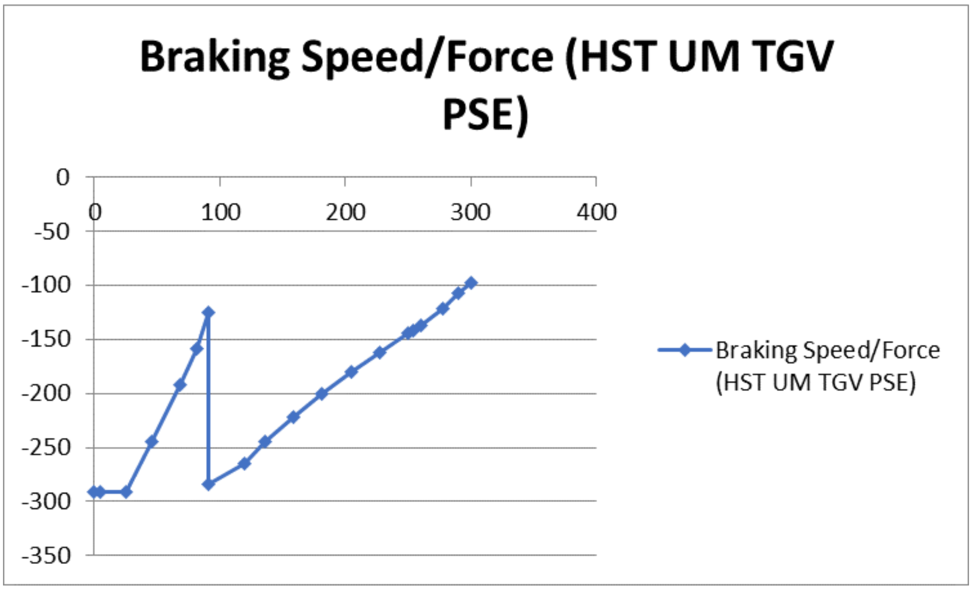


Figure 23: Example of a braking force curve for a type of high-speed train

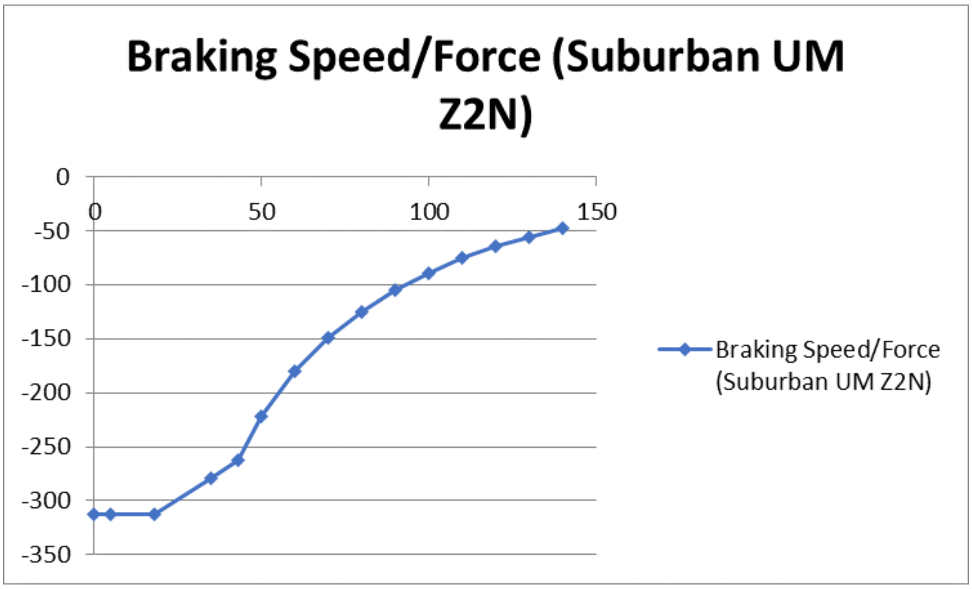


Figure 24: Example of a braking force curve for a type of suburban EMU

4.3 Power elements

This section provides specifics on the operational power management and power status reports in SFERA.

4.3.1 Operational power management

The SP contains the following elements regarding power management:

- Under SP_Points:
  - signals (traction type)
- Under SP_Characteristics:
  - rated voltage
  - estimated voltage
  - current limitation

The values defined in these elements are meant to be static. In the operational context, some of these values can be adapted to take into account the fluctuation of values on a given day, or to document the effect of incidents (e.g. a substation that is shut off).

The DAS can be made aware of such adaptations by using elements in the Journey Profile:

- Temporary signal constraints can be declared to indicate “lower/raise pantograph” or “open/shut main breaker” signals that have been added for an incident or moved for a short period of time.
- Under TemporaryConstraints:
  - EstimatedVoltageConstraints can be declared to send the estimated voltage of the catenary according to the real-time traffic plan, and the current consumption status reports from the DAS of the other trains;
  - TractionTotalCurrent constraints can be declared to adapt the current limitation. This can be used:
    - with a lower value than the SP, for example if a substation risks being overloaded. The driver (or the ATO) then applies this restriction while driving;
    - with a higher value than the SP: if a train is late, and in a time frame where a substation has available capacity, the driver will be authorised to exceed normal restrictions.

Also, temporary power advice constraints can be issued, which can be used to optimise consumption in certain time frames without giving a hard limit that could result in trains not being able to respect their normal Train Path Envelope.

4.3.2 Power status reports

The Status Report message can include different ExpectedCurrentData elements. The aim is to provide this data for a period of time for which the data provided can be sufficiently accurate (e.g. for the following three minutes for a stopping train or for the next 15 minutes for a long-distance train). It should include for different consecutive time periods (from start time to end time) the expected consumption current that will be consumed and the expected regenerative current that will be regenerated. This data will be used in Use Case PW3 to check if there is sufficient power available and in Use Case PW4 to verify if regeneration will be possible.

The data elements provide the total net currents expected at the pantograph (thus including heating and auxiliary power). The position of the train at these times shall be consistent with the JP times. The expected current value averaged over each interval of 15 seconds shall be calculated. The maximum value of these averages in a given period between the start and end times shall be reported. In order to make power management effective, separate start-end-entries shall be given for stationary phases, acceleration phase, cruise phase, coasting phase and braking phase. For all times, the departure time and TP_latestArrivalTime shall be used as a reference. The trackside will add uncertainty margins in time, for example for earlier or later departure or earlier arrival within the chosen window.

Note that while the trackside may make estimates on the acceleration phase, the estimates of the cruising and coasting phases will depend on the optimisation algorithms of the specific DAS and RU and may depend on the outside temperature for instance.

This element is only useful for electric rolling stock.

4.4 Notes on particular JP elements

This section provides some details on a few special elements that can be included in the JP. Some are essential, such as the JP version and the Timing Points, while others are mostly used to provide context information for the driver, such as JP change reasons or stop types.

4.4.1 Determining the latest JP version

The latest JP received is always correct. The time stamp of the SFERA header should make sure that updates are not processed in the wrong order.

The TMS can also use the attribute JP_Version as a unique identifier of the version of the JP. JP_Version is used by the DAS to send a Status Report to the trackside, communicating the JP currently used to generate advice. In the SR, the DAS will also communicate the position within a segment if possible. These SR messages can be used by the TMS to validate that the DAS acts on the most recent information.

4.4.2 Number of Timing Points and Timing Point windows

In the implementation, the IM has to agree with the RUs on how many Timing Points are provided and on the widths of the timing windows. This section contains considerations on how to make this fit the RU's operations.

The timetabling system, the TMS and the DAS all have passing times at each TP. One could put all TPs from the timetabling system or TMS into each JP in TP_latestArrivalTime. In that case, the choice of the width of the arrival window determines the possibilities for energy-efficient and comfortable driving. The DAS can only perform its optimisations if there is sufficient slack time in the timetable, and if the DAS itself can choose how to distribute this over the TPs.

Generally, the DAS advice only requires the TPs where a conflict with other trains lingers if the train gets outside the specific window at that TP. All other TPs should either not be put into the JP, or have a large arrival window (usually several minutes). Note that timetabling data represent nominal times. In order to centre a passing window at a Timing Point, half of the arrival window shall be added to the timetable value for the value of TP_latestArrivalTime.

Advanced TMS can also estimate the time margins between following trains and use this information as a basis to specify TP_latestArrivalTime and the arrival window.

If it is necessary to present timing information at Timing Points without introducing a constraint on the speed profile, TimingPointConstraints could be published with only the TP_PlannedLatestArrivalTime (and not the TP_LatestArrivalTime and arrivalWindow).

4.4.3 Timing Points, Stopping Points and Stopping Zones

In the Segment Profile, it is possible to define at most one Stopping Point Location per Timing Point. This is consistent with SUBSET-126. However, inside a Stopping Point Location it is possible to refer to a "Stopping Zone", and several Stopping Point Locations may belong to the same Stopping Zone.

In the Journey Profile, under Timing Point Constraints, there shall be a reference to either a specific Timing Point ID or to a Stopping Zone.

The goal is to give the option to the TMS to either specify a Timing Point or to give a zone in which to stop, and then it will be up to the DAS-OB (or to the RU DAS-TS, in the IM-RU Setup) to decide the exact Timing

Point. For instance, this can be used when a TMS wants to instruct the train to stop at a platform, but it cannot give a precise location because it doesn't know the length of the train. If all the stopping signs on the platform are coded as Stopping Point Locations which all belong to the same "Stopping Zone" (the platform), the TMS can provide the reference to that Stopping Zone and the DAS-OB will select the proper stopping sign.

Furthermore, in the Stopping Point Location (under Timing Point in the SP), it is possible to add some constraints regarding the Stopping Point Applicability: for example, the minimum/maximum train length, number of cars, types of rolling stock or whether a stopping point is applicable if the train comes from specific preceding SPs.

4.4.4 JP Change reasons

The JP_ChangeReason is a human readable text giving the reason why a JP update is sent (e.g. describing the conflict or conflict resolution). If a reason code is suitable, it is recommended to send the code; it is not necessary to send an additional text. The reason codes to be used are those defined for the C-DAS-C Advice (C_DAS_C_Advice / ReasonCode).

4.4.5 Stop Type

4.4.5.1 Overview

In many countries, the train plan given to drivers does not only contain the stopping and starting times, but also gives additional information regarding the stop. SFERA integrates the possibility of expressing the following information:

- A generic purpose (detailed in the next section).
- A flag indicating if the stop is planned (present in the theoretical timetable) or unplanned (relative to a decision or an event taken after the theoretical timetable of the train has been set). In most S-DAS implementations, all stops will be planned;
- A flag indicating if the stop is mandatory or not. If conditions permit, the driver can pass through without stopping.
- A "Details" text attribute. The three previous pieces of information are expressed in generic form, with a limited number of possibilities. This text has been provided to be able to give more specific elements to the driver.

If a stop is needed for multiple reasons, it is possible to add multiple StopType elements to the StoppingPointInformation element.

4.4.5.2 stopTypePurpose

The purposes for a stop are divided into six generic possibilities:

- Commercial: a stop used by the RU to perform the commercial duties the train was planned for. For a passenger train, this will concern stops where passenger loading or unloading is done. For a freight train, this will concern stops where loading and unloading is done, or when wagons are either added or uncoupled from the train.
- Non-commercial Traffic Management: stops integrated for the purpose of guaranteeing the fluidity of traffic.
- Non-commercial RU: stops used by the RU to fulfil its specific needs. For example, a train driver relay, or a locomotive change could be characterised by this purpose;
- Non-commercial IM: stops necessary for the IM's needs;
- Non-commercial safety: stops used to perform a safety procedure;
- Other: any stop that cannot be categorised using one of the five other types.

4.4.5.3 Examples

Table 16 lists a few examples of how the TMS could use the stop type element to express different situations.

Table 16: Examples of stop type

Example	**Purpose**	**Planned/** unplanned	Mandatory
Normal stop of a passenger train	commercial	planned	yes
Planned stop for driver change	non-commercial RU	planned	usually no (to provide for days where no driver change is needed)
Freight train stop integrated into the theoretical timetable on a passing loop to let an express passenger train overtake	non-commercial regulation	planned	no (if the express passenger train is late, the freight train will not stop)
Train stopped at a signal at a junction to let the previous train through (which is late)	non-commercial regulation	unplanned	no (if the previous train has already cleared the junction when the train arrives at the signal)
The IM delivers a safety order to a driver at a normal stop for a passenger train	commercial	planned	yes
The IM delivers a safety order to a driver at a normal stop for a passenger train	non-commercial safety	unplanned	no
The IM asks the RU to stop a passenger train at a level crossing that is out of order to drop off a maintenance crew	non-commercial IM	unplanned	yes

It is to be noted that the SFERA protocol has not been designed with safety in mind. For “non-commercial safety” stops, the SFERA protocol has not been studied with the safety level to be used to impose the train stop. The IM shall use other means to ensure that the driver does not enter a “danger” zone without having received the proper instructions (for example, they may operate a signal which will stop the train).

4.5 S-DAS

The SFERA communication protocol supports both S-DAS and C-DAS. This section describes some implementation details.

4.5.1 Data in S-DAS

Data for an S-DAS will be loaded from a timetabling system rather than from a TMS. Therefore, data to an S-DAS will only be loaded via a back office in the IM-RU setup.

In the IM-RU setup, the IM needs to provide both SPs (speeds, gradients, GNSS points) and the link from timetabling to the SP IDs. (Note: if information on, for example, GNSS points is not present in the timetabling system, linking this should be done by IM, not by RU).

It is possible to operate an S-DAS with only a few geo points such as stations and significant direction changes in the track, but some implementations can need more than this.

For an S-DAS, all possible TCs should be included in the SFERA data. The driver can then select the appropriate one at the start of the run.

The commonality of SPs for an S-DAS is used to reduce the data to be transmitted. JPs and SPs to alternative routes will not be used in an S-DAS unless the driver has a way to select alternative routes (e.g. routes to alternative platforms).

4.5.2 S-DAS as basis (backup and preloading) for a C-DAS implementation

S-DAS and C-DAS implementations for the same line can use different SP_IDs, as long as the Infrastructure Manager ensures that no doubling in the numbering sequences occurs.

If S-DAS is used as a fallback for C-DAS, stored TCs and SPs may be re-used if the SPs are the same for S-DAS and C-DAS operation. SPs may in that case be preloaded in the DAS.

If an S-DAS and C-DAS implementation share the same SP_ID and TC_IDs, the S-DAS data may be used to preload both JPs and the most-used SPs. If this is done, data transmission usage while running trains will be low in most cases. It is up to the RU to determine which data could be pre-loaded. For instance, SPs for deviation routes might be preloaded if these can be derived from, for example, foreseen alternative platforms in stations or de-rusting schedules.

An RU can use S-DAS functionalities to preload segments. The decision on which segments to preload will depend on what device is used. If a handheld device is used, it will make sense to follow the person's duty. For on-board devices, it may be advisable to follow the likely routes. If this functionality is used for C-DAS preloading, it can be effective to also give a list of the most-used slightly deviating routes (e.g. when different platforms are used) to preload SPs.

The definitions of useful subsets of JPs, SPs and TCs for S-DAS and for caching will depend on the type of device. This will be different for handheld devices and for devices mounted on trains. The SFERA protocol does not describe how to transfer such lists. This is up to the implementers and is foreseen in the initial handshake with the back office. Fetching the JPs, SPs and TCs from the DAS-TS is possible via the SFERA communication setups.

4.6 ATP systems in train and track

SFERA is specifically designed to also be usable for operation on Class B systems. This section provides some considerations on how this is reflected in the SFERA protocol.

4.6.1 DAS and ATP system parameters

Both the TCs and the track can code that one or more ATP systems are present on the train or in the track. In SFERA it is assumed that the DAS can then choose which ATP system (if any) is active and determine the applicable speed and braking regimes.

A number of characteristics in the track data are dependent on the ATP system: the speed profile and the signal variants. In the train data, a braking profile may be specified per ATP system. Note that this functionality is specifically also used in the low adhesion calculations.

Some ATP systems may take into account the following parameters from the TC to calculate braking points and the braking curve: train length, max. train speed, train weight, train category code and braked weight percentage.

It is assumed that the DAS itself has knowledge of other reaction parameters for a specific ATP system (including, for example, speed ranges to which the ATP reacts) from these five parameters. The DAS can either construct a braking curve and send it to the TMS or receive one sent by a TMS.

The max. train speed does not depend on the ATP system. It is assumed that this is coded into the speed profile in the SPs.

4.6.2 Use of the signal type

Many of the signals have no impact on DAS. Some do, however, because they can affect the braking curve. Some “braking curve intelligence” has been put into the placement of signals. The signals are not shown to the driver in the DAS application. However, the information needs to be there to allow systems to use it in the DAS calculations. In different countries (and ATP systems), the SFERA message will be used in different ways.

The signal type shows the locations where the driver will change speed. In this respect, the SFERA signal type holds information for Class B systems similar to the ETCS static speed profile combined with the ETCS Train Characteristics. In many cases, the ATP system will enforce certain actions from a driver. This leads to reasonably predictable speed change patterns. These speed patterns are used in the DAS to calculate a time of arrival. Danger signals only need to be included in the SFERA profiles if the ATP system requires speed reductions at certain distances before such a signal, or for cases when it will be needed to stop a train for a certain period of time at a signal (e.g. delivering a safety order).

Below are four examples of Class B ATP systems and the effect of signal elements on the driving profile.

- Example ATP System 1: PZB-90 (a. o. Germany, Austria). In PZB-90, passing a signal at warning will trigger minimum braking to a certain speed and limit the train speed to that velocity[2]. The positions of the PZB magnets can be inferred from the positions of the signals, as they are always at fixed distances.
- Example ATP System 2: ATB-EG (a. o. the Netherlands). In ATB-EG, the target speed at the next signal is transmitted continuously, with a minimum of 40 km/h. If the actual train speed is higher, minimum braking is enforced. Accelerating above the minimum speed is not possible. Codes change at warning signals, signals that denote a lower maximum speed at the next signal and speed increase signals.
- Example ATP System 3: KVB (a. o. France). In KVB, passing a signal at warning or lower speed announcement will trigger a braking curve towards the goal point.
- Example ATP System 4: TVM (France, Belgium, UK). Current and next speed limits (and in TVM 430 block length) are transmitted[3]. For TVM-300, similar rules are implemented as for ATB-EG, often reducing speed more than necessary. The position of code changes in the SFERA profile is therefore useful. TVM-430 has brake curve control and could in principle also be modelled with driving profiles.

4.6.3 When to update signal aspects in C-DAS operation

The signal type shall show the expected aspect of a signal at the time the train is expected to pass that signal (and explicitly not the **current** aspect at the time data is exchanged). Also, the expected aspect of a signal should not change too often: if a lower speed signal is transmitted for a later part of the journey, the DAS will show faster advice speed in order to make up time.

Too frequent updates will distract the driver, or may force the DAS provider to implement functionalities in their algorithm to avoid showing such frequent changes to the driver. Not updating may mean that the arrival time is less precise, but this may be worthwhile in many cases.

Also, the fixed or temporary expected signal aspect cannot be used to inform the driver on how to avoid signals at warning. This should be done solely by adjusting the timing of points in the JP. A signal at warning in the (fixed or temporary) expected signal aspect shall only be used to take the effect on the

[2] See https://de.wikipedia.org/wiki/Punktförmige_Zugbeeinflussung.
[3] See https://en.wikipedia.org/wiki/Transmission_Voie-Machine.

running time estimate into account.

Rust running: Using deviant routes (e.g. for the Dutch and German practice of cleaning rust from the tracks) needs special attention. Often, these deviant routes are only planned at the last moment (within 10 minutes before the train passes the switch).

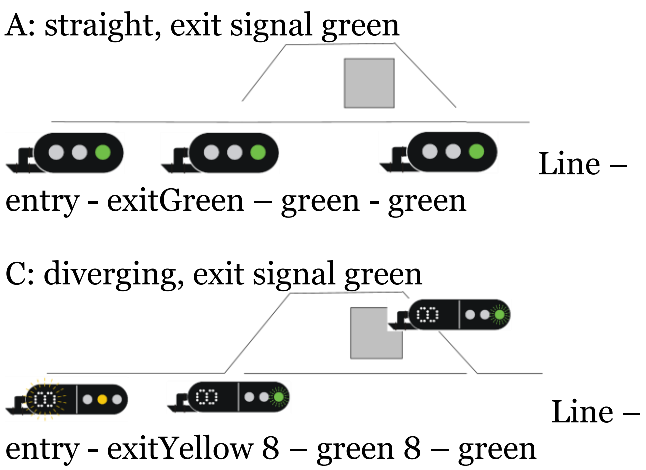
4.6.4 Signals for an example station

For SFERA, the impact of the signals on the running time needs to be right in order to have the train exactly on time. In this section, four examples of expected/usual signal aspects are presented that are common across Europe. This section is intended as an example only to understand generic issues that may arise in specific implementations. The Dutch signal aspects are chosen to denote the situations. The interpretation of these signals is explained in the text.

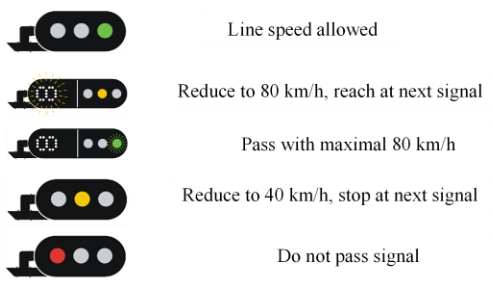
The situation is a station with a platform that can be reached with a switch. In the straight position, the switch can be used at line speed, while for the diverging route the maximal speed is 80 km/h. At the end of the platform there is an exit signal that either shows danger or not.

In Table 17, the four possible approach situations are shown.

Table 17: Signals for an example station



Legend:



In the four situations, the train speeds (and thus the time from the last line signal to arrival) will differ greatly. In situation A, the driver is free to brake at will. In situation B and D, the passing of the signal at danger will trigger an ATP reaction. Therefore, the time will be longer. In situations C and D, the driver has reduced the speed before the entry signal. This takes even more time.

The appropriate time is determined by the platform assignment in the timetable, and the wish to release the exit signal early, or not.

The best practice for this situation is to use temporary signal constraints in the JP for these restrictive signals. See the relevant section. Alternatively, an IM can also use four different SPs to code for the four situations.

Note that for ETCS tracks, only two driving profiles are required: one for the straight route and one for the deviating route. The green or red exit signal is of less importance, as there is no obligatory braking at the entry signal.

4.6.5 Dual signalling with Class B and ERTMS

If TVM and ETCS are implemented on the same lines, the static speed profiles for both can differ[4].

In the Netherlands, for instance, “dual signalling” tracks exist. The speed limit for trains with ERTMS is given in the static speed profile and is usually 160 or 200 km/h. For Class B systems, the maximum speed is 140 km/h. Also, for Class B systems, speed limits change at signs or signals, while for ERTMS, the speed limit can be projected onto the location (e.g. switch) which is actually restricting. Having the ERTMS static speed profile separate gives more flexibility.

The SFERA protocol is designed so that both data can be transferred. Several fields have the option to add a version that depends on the ATP system. The least-restrictive speed curve is then given unconditionally (in this case, the ETCS speed curve). For the more restrictive ATB system, a separate speed curve can be included. Furthermore, signals can be selected based on an ATP system.

4.6.6 Speed profiles versus signals

The SFERA messages contain a StaticSpeedProfile (and possibly also AdditionalSpeedRestrictions). This SFERA speed profile is designed to work in the same way as in ETCS: the maximum speed at which a train is allowed to run at that moment in time. For Class B systems, this should be the line speed, or a locally restricted speed. However, Class B systems will usually be more restrictive. For example, the ATB and TGV ATP systems will force a train to brake towards the speed at the next restriction. This is not the same as a lowered speed in the static speed profile, as the permitted speed will depend on the passing speed of the announcement of the restriction. Similarly, passing a non-green signal with a KVB or PZB balise will force a certain amount of braking. These forced brake applications should not be included in the static speed profile; they should only be expressed in the signal element of SFERA.

Another advantage of this way of coding is that the static speed profiles will be the same for both ATP systems in overlay situations with ERTMS.

4.6.7 Traction signs versus ETCS traction status

This section explains why SFERA codes traction signs along the track rather than changes in the power system.

In the ERTMS specification (SUBSET-026 and thus SUBSET-126), it is assumed that the train knows how far ahead traction cut-off needs to be applied to be in time. The train knows its length and the response time of the traction system.

This is Point C in the ERTMS specs on http://www.era.europa.eu/Core-Activities/ERTMS/Pages/Set-of-specifications-2.aspx.

In ERTMS, the train will calculate this point, and SUBSET-026/126 will only include the real powerless section. However, a DAS or Class B system will not know this response time for the train.

On non-ERTMS lines, signals are presented to drivers to indicate where they have to take specific action. Often, these signs depend on the train length.

[4] Source: https://www.railwaygazette.com/news/single-view/view/integrating-etcs-level-2-with-tvm430-on-tgv-est.html.

Therefore, the trackside sign that pantographs need to be lowered / raised needs to be included in SFERA. For the ERTMS part (SUBSET-026/126), the real position needs to be relayed. For non-ERTMS lines, the announcement location. The DAS can then calculate the effect on the running time of the length of powerless running.

5 Train positioning

This section contains some notes on train positioning and how it is measured and communicated in SFERA.

5.1 PositionSpeed object

The DAS on board can send the current position and speed of a train to the TMS. The TMS may use this to augment data already available in the TMS on the train's position. The sole use of this DAS data without other inputs is discouraged, as the positioning given by the DAS is relative to the SPs referenced by the JP, but not necessarily the real route taken by the train (route changed, but not known to the TMS, and therefore the DAS). Any position report will include SP_ID, SP_Zone and the location of the head of the train and the speed at the time of the report. The DAS will use all available sensors to give the best estimate for position at the time of the report. Optionally, the width of the 99% confidence interval of the position may be given using location_PositiveErrorMargin and location_NegativeErrorMargin.

If one or more GNSS sensors in the train are active, the positions of these sensors can optionally be sent as well. This may be useful for debugging purposes (e.g. for the verification of the positions of virtual balises), but also if the IM wants to re-project the position on the track (the DAS projects the GNSS position on the SPs which are theoretical, but may not correspond to the actual route that the signalman has set for the train). All measured GNSS information points can be sent: latitude, longitude, altitude, heading, GNSS_precision. As GNSS measurements may be less frequent than other position updates, a separate GNSS timestamp should be sent. Furthermore, the devices will not be mounted at the front of the train. The antennaDistanceFromFrontOfTrain attribute indicates the distance from the front of the train. The SP data shall be used to determine the direction of the offset in the verification. In the case of a mobile device, the device is assumed to be present in the cabin, where its position is known with a precision of less than two meters.

Note that for certain DAS, for example in metros/suburban trains, it would be unpractical to use GNSS positioning. Furthermore, GNSS measurements may not be present in tunnels, for instance, while reasonably precise positioning is still available from odometers. However, SFERA integrates the messages that permit a DAS to request the last-known position from the ground if the device's main source for positioning is unavailable.

5.2 The use of odometers and GNSS for positioning and status updates

In SFERA, the chosen route is presented as a one-directional track. The most likely implementations will use GNSS locations with virtual balises and/or odometry, at least in tunnels. For longer stretches of tunnel, ERTMS or other balises may be used to synchronise the odometry. This application therefore may require some sensor fusion to deduct the actual location. For metro-like implementations, it is also possible to use SFERA without GNSS. Therefore, it is chosen (as in ERTMS/ATO) to report the position in a segment as the current front-of-train position, rather than a GNSS location. Furthermore, this makes calculations simpler if the GNSS antenna is mounted at a fixed distance from the train front end.

However, IMs shall account for DAS implementations that rely on mobile hardware (smartphone or tablet), which solely rely on GNSS positioning.

6 Messaging details

This section contains some considerations relating to SFERA messaging, in particular the contents of the message header.

6.1 TrainIdentification

The TrainIdentification in the MessageHeader may be used for context purposes, but is not necessary for functional purposes. In the IM-Train setup, there is no need to use TrainIdentification, as the connection between TMS and train is direct; in the IM-RU setup, each message may contain information related to several trains, therefore TrainIdentification is included in the relevant payloads.

If it is necessary to present the train number to the driver on the DMI, the SFERA JP provides the JP_ContextInformation/Applicable_OTN element.

6.2 Relations between TrainIdentification and NID_Operational

In ETCS, the numerical NID_OPERATIONAL identifies the number used on the entirety or a section of a train run, and is entered by the driver in the ETCS systems in real time. For SFERA, the TrainIdentification_ComplexType (which can be an OTN or an identifier from TAF/TAP) is used to identify a journey. But in ERTMS/ATO communication in SFERA, the field NID_OPERATIONAL also has to be present in the message header. In many cases, NID_OPERATIONAL can be filled with the numerical elements from the Core (TAF/TAP) or TrainNumber (OTN) element of TrainIdentification_ComplexType. However, information from the other fields may be needed in this conversion. The RU shall provide this logic.

6.3 Relationship between TrainIdentification and the Journey Profile Object

Depending on the TrainIdentification object that is used, the rules to identify the scope of a Journey Profile object may vary.

TAF/TAP TrainID: This identifier, normalised at the European level, is defined by the RU to refer to the whole service of a train seen from their business perspective. As such, it is not supposed to vary on the perimeter of an IM. This means that the IM will systematically communicate a Journey Profile that will extend to the last point of responsibility of this IM. When available, it is preferable to use this identifier. It is foreseen that in the medium-long term, only this type of identifier will be used.

OTN_ID: Depending on the IM, this identifier can refer to two scopes for a train service, especially in the case where the OTN changes for a train service on the network of a given IM:

- First possibility: the IM considers all the parts of a train service, even if the OTN used in the OTN_ID is only used on a part of the train service, and the Journey Profile will extend to the last point of responsibility of this IM.
- Second possibility: the IM considers only the part of the train service using this OTN. In this case, the DAS-OB will need to establish a connection and handshake for every section of the train service that uses a different OTN.

6.4 Relationship between NID_ENGINE and European Vehicle Number (EVN)

Under ERTMS/ATO, NID_ENGINE uniquely identifies an EVC (on-board computer system of ETCS). Usually, an EVC is present per cab on multiple units, but a single one is provided on locomotives (used by both cabs). The NID_ENGINE can therefore be used to identify the head trainset of a train. This field is required for SUBSET-126 communication. Therefore, NID_ENGINE shall be included in each message in the ERTMS/ATO setup.

On a non-ETCS train, the EVN field in the SFERA header can be used to uniquely identify the rolling stock in the messages. This should be the EVN of the car where the driver is situated (locomotive or powerless driving car at the front of the train, first EMU/DMU). This would help the RU to identify errors between their roster and the train in which the drivers identify themselves.

The EuropeanVehicleNumber field in the SFERA message header is optional. It is foreseen that this data will not be manually entered by the driver for handheld DAS. In other setups (integrated DAS), it is advised to include the EuropeanVehicleNumber field at least during the handshake or if the EVN changes.

An EVN may be identified by one or more NID_ENGINE IDs. Furthermore, each train consist will often have both a different EVN and NID_ENGINE per travelling direction.

6.5 Single object for request/reply messages

In the IM-Train setup, it is recommended that the communication event and reply messages only contain a single object. This is because the guaranteed delivery of MQTT (see Appendix B – Communication Protocol) is most efficient when using small messages. This also applies to ERTMS/ATO, as SUBSET-126 only sends single objects. This constraint has not been set in the XSD, as it is still possible that sending multiple objects per request can be beneficial for an IM-RU setup due to a different communication layer, the S-DAS use case or communication in SFERA format between an IM and RU.

In particular, in the IM-Train setup it is not possible to request multiple Journey Profiles, as the communication layer only supports a JWT with authorisation for a single train (see Appendix B for details).

6.6 destinationDevice

The destinationDevice attribute in the header is optional. In ground-to-board communications, it is only used in cases where an Event message is to be acted on by a single DAS on-board device (in cases where multiple devices are active for the same train), identified by the destinationDevice attribute. An example for this is the use case DM4 (see Appendix A) “Force Driving Mode Change of DAS Device”, which is typically used to force a single device to become “readonly”.

If the attribute is omitted, then the message is addressed to all devices in a particular train.

The attribute can be safely ignored in board-to-ground communications (it is assumed that communications are already established with the correct ground server).

6.7 SFERA version and MessageReference

In the SFERA message header, the SFERA version used shall be included inside the attribute SFERA_version. The format of the version has to be <majorVersion.minorVersion> and correspond to what is indicated in the “version” attribute of the <xs.schema> tag within the SFERA XSD used (see Appendix D). As mentioned in Chapter 7 of the IRS, all messages with the same major version shall be backwards compatible.

The SFERA header also includes the optional sub-element MessageReference, which shall be included when the Common Interface is used for communication (see Chapter 6 of the IRS and Appendix B for details). Inside MessageReference, the sub-element MessageTypeVersion shall have the same content of the SFERA_version attribute.

7 Data for Incident Situations:

Information exchanges between the Driver, the RU back office and the IM are very important for incident situations. Up-to-date and reliable information is important to help resolve the situation, inform actors and clients and prepare for post-operational analysis. Given the privileged link that SFERA is helping to build between the driver and the ground, payloads have been added to exchange information regarding delays. The TAF/TAP TSI specifies the information exchanged and the processes linked to these situations for the IM/RU back office link. The payloads in SFERA were directly inspired by the payloads of TAF/TAP messages, to facilitate the translation and use of the data in the IM and RU IT systems.

7.1 Board-to-ground payloads

From board to ground, three payloads are provided:

- Train Ready Status;
- Interruption Point;
- Delay Notice.

7.1.1 Train Ready Status

The TAF/TAP TSI defines a list of information to be exchanged regarding the Ready/Non Ready status of a train. This information shall be provided to the IM in the following situations:

- Train start at origin;
- Composition change;
- RU change.

The TAF/TAP TSI provides a standard XSD definition for the exchange of this information between RU and IM back offices, but alternative means are authorised if a bilateral agreement exists (GSM-R, button on the platform, telephone, etc.).

SFERA's primary aim is to help the RU to obtain information from the driver regarding the Ready/Not Ready status in the IM-RU setup. Once the RU back office has received this information, the RU will have the following responsibilities:

- If relevant for their processes, associate the information with information coming from other actors (operational centre, conductor, agent in charge of assembling the train, etc.);
- Transmit an assembled message to the IM backoffice:
    - Through the standard TAF/TAP data exchange. In this case, the status reports sent from the RU DAS-TS to the IM DAS-TS will not include the TrainReadyStatus payload sent from the DAS-OB;
    - In the case of a bilateral agreement between the IM and the RU, the RU DAS-TS could maintain the TrainReadyStatus in the StatusReport sent to the IM DAS-TS.

The payload is based on the TAF/TAP Train Ready Message payload, with the following changes:

- addition of a free text "PreciseDelayCause";
- no support for the "TrainStartTime" attribute, as it is considered to be covered by the "PositionSpeed" sub-element of the Status Report.

The TrainLocation will be defined according to the TAF/TAP locations included in the Segment Profile. The DAS-OB will first compare its current location with the TAF/TAP locations in the Segment Profile:

- If the DAS-OB is situated on a single Subsidiary location, the DAS-OB will document this subsidiary location in the message;
- If the DAS-OB is situated on multiple Subsidiary locations, the DAS-OB will propose by default the nearest subsidiary location and propose the other near subsidiary locations;
- If the DAS-OB is situated on a Primary Location, the DAS-OB will identify the nearest subsidiary locations related to this primary location and ask the driver to confirm;
- If the DAS-OB is neither situated in a Primary Location, nor in a Subsidiary Location, the DAS-OB will identify the nearest primary location and ask the driver to confirm, and then identify if subsidiary locations related to this primary location are in the vicinity.

Subsidiary locations are of different types, but the TAF/TAP TSI does not define the types specifically used for Train Ready Messages. At this point, the IM and RU will need to agree if the subsidiary locations used shall be restricted to certain types (e.g. restrict train ready messages to signals).

The TrainReadyStatus will be present in the Status Report payload until the train has departed from the location to which the payload applies.

7.1.2 Interruption Point

The TAF/TAP TSI defines a list of information to be given by the RU when a train is stopped outside a planned stop, or the duration of a stop is prolonged. Additionally, this is applicable to stops out of scope for the Train Ready message.

The TAF/TAP TSI provides a standard XSD definition for the exchange of this information between RU and IM back offices, but alternative means are authorised if a bilateral agreement exists (GSM-R, button on the platform, telephone, etc.).

SFERA's primary aim is to help the RU to obtain information from the driver regarding an interrupted status in the IM-RU setup. Once the RU back office has received this information, the RU will have the following responsibilities:

- If relevant for their processes, associate the information with information coming from other actors (operating centre, conductor, agent in charge of assembling the train, etc.);
- Transmit an assembled message to the IM back office:
  - Through the standard TAF/TAP data exchange. In this case, the status reports sent from the RU DAS-TS to the IM DAS-TS will not include the InterruptionPoint payload sent from the DAS-OB;
  - In the case of a bilateral agreement between the IM and the RU, the RU DAS-TS could maintain the InterruptionPoint in the StatusReport sent to the IM DAS-TS.

The payload is based on the TAF/TAP Train Running Interruption Message payload.

The TrainLocation will be defined according to the TAF/TAP locations included in the Segment Profile. The DAS-OB will first compare its current location with the position of the TAF/TAP locations in the Segment Profile:

- If the DAS-OB is situated on a single Subsidiary location, the DAS-OB will document this subsidiary location in the message;
- If the DAS-OB is situated on multiple Subsidiary locations, the DAS-OB will propose by default the nearest subsidiary location and propose the other near subsidiary locations;
- If the DAS-OB is situated in a Primary Location, the DAS-OB will identify the nearest subsidiary locations related to this primary location and ask the driver to confirm;

- If the DAS-OB is neither situated in a Primary Location, nor a Subsidiary Location, the DAS-OB will identify the nearest primary location and ask the driver to confirm, and then identify if subsidiary locations related to this primary location are in the vicinity.

Subsidiary locations are of different types, but the TAF/TAP TSI does not define the types specifically used for Train Running Interruption Messages. At this point, the IM and RU will need to agree if the subsidiary locations used shall be restricted to certain types (e.g. restrict train ready messages to signals).

The interruptionPoint payload shall be present in the Status Report for the duration of the interruption. The interruption will be considered as lifted when the first Status Report is received without this payload.

7.1.3 DelayNotice

The TrainReadyStatus and Interruption payloads for SFERA can contain the cause of delays in their respective use cases, which is very useful for the RU and IM to make informed decisions regarding the train and to feed post-operational analysis. However, they do not cover all use cases, so this payload has been defined for the driver to send the cause for delays.

The Train Location shall be the Location where the delay occurred and not the position of the DAS-OB when the driver documented the cause.

Multiple delayNotices can be given at the same location if the delay was caused by multiple events.

Note that no message is currently provided in the TAF/TAP TSI for this case, so no translation to TAF/TAP messages is foreseen for the RU DAS-TS. It will therefore systematically be sent in the status report to the IM.

The delay notice only needs to be sent once in the status report.

7.2 Ground to board payload: Interruption Point

The TAF/TAP TSI defines that the IM is responsible for informing the RU of decisions to interrupt a train run. If provided by the IM, an interruption means that the IM is unable to generate a reliable forecast from the point of interruption, mainly linked to the fact that the duration of the stop is not precisely known. SFERA provides the means of transmitting this interruption through the Journey Profile. When received by the DAS-OB in the Journey Profile, this means that the train will be stopped at a given location for an unknown amount of time, and that driving advice calculated downline from the interruption point is not to be considered as reliable.

The interruption is considered as lifted from the moment an update of the Journey Profile is received that does not include the InterruptionPoint payload.

8 Non-core data

This section documents some features of SFERA that were considered “non-core” for the main SFERA objectives. However, these features were deemed useful in the SFERA protocol and they have therefore been included. The features are the text messages between ground and board and the advice messages for C-DAS-C.

8.1 Text messages

The text message is a way to secure communication between the driver and the TMS. The text message uses the SFERA communication layer but does not use JP or SP information. A PlaintextMessage can be sent using SFERA B2G and G2B Event and Reply messages.

There is no acknowledgement, cancellation, lifetime management or prioritisation for the PlaintextMessage. Therefore, this function should explicitly not be used for:

- explanations of reasons for changed profiles (the ChangeReason element can be used in the ground-to-board message changing the JP);

- feedback from the driver to traffic management that data was received (no user interaction should be requested while driving);
- feedback from the driver to traffic management (e.g. unplanned stops): standard communication channel (e.g. phone) should be used to secure a single process;
- an escape channel for computer-interpreted information – this could lead to loss of interoperability and to security holes in the operation.

It is also explicitly recommended not to misuse text messages, in particular to implement new undocumented features in SFERA. Any request for additional features in SFERA shall be made through the process described in the “Maintenance of SFERA” chapter of the IRS.

8.2 C-DAS-C

For those DAS that are not able to calculate advice on board, separate C-DAS-C messages may be provided. As mentioned in the IRS, this feature was provided largely to support the smooth transition to C-DAS-O for railways that currently employ C-DAS-C.

Only the DAS with behaviour setting C-DAS-C_Only may display these advice messages. Therefore, there are no provisions for transitions between C-DAS-C mode and other modes. None of the data from C-DAS-C will be displayed in any other mode than in C-DAS-C mode of a C-DAS-C_Only device. Localisation for C-DAS-C is provided by measures deployed by the Infrastructure Manager only. C-DAS-C does not reference JPs or TCs. SPs may be referenced.

Multiple C-DAS-C advice messages may be sent in order to set up a journey. The DAS shall keep all messages with end validity in the future.

8.2.1 Identifier

The unique identifier of the C-DAS-C message is the Message_ID of the SFERA message. This identifier can also be used in board-to-ground Status Reports to communicate the C-DAS-C advice currently being shown to the driver.

8.2.2 Advice type

The type of advice can be one of the following:

- stop advice: start braking at start validity, and be at zero speed before end validity;
- acceleration advice (percentage field optional);
- constant speed advice (optimal speed or delta speed field required);
- coasting advice;
- operational braking advice (percentage field optional);
- electrical braking advice (percentage field optional);
- end of advice: clear C-DAS-C display;
- departure advice (earliest and latest time fields optional);
- text advice;
- delete advice (revoking the advice given earlier with the Identifier given in the header field Correlation_ID).

All combinations of advice messages, such as “coast and proceed at reduced speed” shall be sent as a series of advice messages. Acceleration and braking advice messages are not mandatory if the train should alter its speed. It is sufficient to send constant speed advice. The driver then takes the decision on how to reach the advised speed.

8.2.3 startValidity and endValidity

Start validity of an advice message is the earliest time that such advice may be shown. End validity shall be after start validity.

8.2.4 Optimal speed

This is the optimal speed given to the driver in km/h. It is recommended to fill the parameter optimal speed according to the following rules, depending on the type of advice being sent:

- stop advice: optimal speed shall be null;
- acceleration advice: optimal speed is the target speed that shall be reached at the end of the acceleration phase;
- constant speed advice: optimal speed is the target speed for the cruising phase;
- coasting advice: optimal speed is the target speed that is estimated to be reached at the end of the coasting phase, or special value 999 if not defined;
- operational/electrical braking advice: optimal speed is the target speed that shall be reached at the end of the braking phase;
- departure advice: optimal speed is either null or the target speed that shall be reached after departing at the end of the acceleration phase.

The special value 999 indicates “maximum allowed track speed/most restrictive speed profile” (all but coasting advice).

8.2.5 Delta speed

The delta speed is the difference between the most restrictive speed profile and the optimal speed (in km/h). The most restrictive speed profile is the minimum of: maximum track speed, maximum train speed and other local speed restrictions. Value “0” indicates “most restrictive speed profile”. Values above zero indicate the difference in km/h between the most restrictive speed profile and the speed advice (optimal speed).

8.2.6 Location text

This is a human-readable text of the location information where the advice becomes valid and/or ends, which will be displayed on the DAS on board. The location should be easy to understand and applicable for the driver. If no location is given, the advice is valid within the start and end validity times.

This field should be used as little as possible. Instead of the entire text, the elements of the location should be used in order to let the RU decide what to display to the driver (see detailed location elements, Section 8.2.7). The TMS has predicted journey times, and start validity and end validity should preferably be used in these calculations. Location shall only be used when the start validity and end validity are insufficiently precise.

It is recommended to fill the parameter location according to the following rules, depending on the advice type being sent (this is also valid for the detailed elements of the location in Section 8.2.7):

- stop advice: location is the station the train is stopping at (e.g. train approaching a scheduled stop);
- acceleration advice: location is the location (kilometre/landmark etc.) where the acceleration phase starts;

- constant speed advice: location is the location (kilometre/landmark etc.) where the cruising phase starts and/or ends;
- operational/electrical braking advice: location is the location (kilometre/landmark etc.) where the braking phase starts;
- end of advice: location is null;
- departure advice: location is the station the train is departing from.

8.2.7 Detailed location elements

The locations should be sent in the form of single elements to increase the options for the RU to display them to the driver. For the start and the end location of the advice, the following elements can be sent:

- long and short name (abbreviation) of the station;
- line number and kilometre of the location;
- the name of the signal, if applicable.

8.2.8 Reason text and reason code

The reason text is a human-readable text giving the reason why an advice message was sent (e.g. describing the conflict or conflict resolution). It is possible to use only one of the parameters “reason code” and “reason text” or to send them in combination.

The reason code offers coding for standard situations. If the reason code is null, the last reason code received is still valid. It is possible to use only one of the parameters “reason code” and “reason text” or to send them in combination. A set of predefined reason codes is defined in the SFERA XSD.

8.2.9 Percentage

The percentage is given as a number between 0 and 1. Its meaning depends on the type of advice being sent. It is only valid with certain types of advice:

- acceleration advice: percentage of maximum acceleration;
- operational braking advice: percentage of maximum operational braking;
- electrical braking advice: percentage of maximum electrical braking.

8.2.10 earliestDepartureAdviceTime and latestDepartureAdviceTime

These are the lower and upper bound of the time window for train departure, only to be used if departure advice is sent. If the earliest departure time is not specified, the advice should be obeyed at any time earlier than the latest departure time. If latest departure time is not specified, the advice should be obeyed at any time later than the earliest departure time. If neither attribute is specified, the advice should be obeyed immediately.

8.2.11 Special elements to use when displaying absolute speed mixed with speed differences

When the speed difference is used, the advice may be renewed at any location where the maximum speed changes. Special cases occur when there are mixed advice messages on speed differences and advice messages on absolute speed on the same train run. In order to allow the train driver to know where the new absolute speed starts, some additions shall be made to the first advice message with a new absolute speed.

In C-DAS-C mode, there is typically no SP reference system or geolocation. Therefore, either a station name, a signal number or a track km position and line number (note the difference to LineAndTrackIdentifier, which is a string) can be given. Both the advice before and after this location shall be shown so that drivers can determine their behaviour both before and after passing. As these locations have no natural order, the order of possible advice messages shall be derived from the start validity field.

This advice shall refer to the end location of the last advice message with this absolute speed. Therefore, the end location (endStationShortAbs, endStationLongAbs, endKmAbs, endLineAbs, endSignalAbs) and the end time (endvalidityAbs) of the absolute speed shall be given. Every following advice message until this end shall refer to the ID of the first advice message with this absolute speed (referenceIDAbs).

8.3 Related Train Information

The Related Train Information is the way to exchange trackside and on-board information about the occupation and status of other trains which are on the path of the ‘own train’ and which are running in front or behind and are using the same tracks as the ‘own train’.

This makes it possible to share the position of these trains and obtain additional information about them without requesting their JP and SP (for example, to obtain the next station or current speed). A DAS may use this information to update its driving advice, or drivers can use their expertise to interpret the displayed information. Also, a driver can inform travellers about possible delays that may occur because there is another train ahead.

The message has the following structure (the term ‘own train’ is used here to identify the train that will receive the Related Train Information):

- Own train: Position of the ‘own train’ in terms of segments (SP_ID) and the position within the segments
- Trains in front or behind the ‘own train’: Train that shares parts of the same train path as the ‘own train’, which will use the same infrastructure as the ‘own train’. This part of the structure holds extra information such as the next or passed stations, passed signals, etc. This can be shown on the driver’s DMI.

The most important parts of the data concerning the related trains are:

- Segment Claim: Gives the position of a train in terms of segment (SP_ID) and the position within this segment. A train can claim several segments (or part of the segments) and the claim type may differ over the different segments or within the segment.
- Claim Type: Gives an indication of the status of route setting for this train on the given piece of the segment. Possible values are:
  - Occupied: a train is actually present on a given segment.
  - Reserved: the route and - probably - the interlocking for this train is set, but the train is not there yet.
  - Reserved implicit: the route is fixed, but it is unknown if it is actually reserved (e.g., because the train is in an area with permissive signals and the trackside does not know the state for these signals).
  - Planned: the route setting will be done in time before the train enters the respective part of the route, but it can still change over time.
  - Unknown: the status is unknown.
- Leaving Entering Status: Gives an indication if a train is entering or leaving the train path of the own train at a certain location.
- Same Direction: Only for trains in front of the ‘own train’, gives an indication if the train is going in the same direction (true) as the ‘own train’ or in the opposite direction (false).
- Delay: Delay of the train in time

- Delay Category: Possible values are:
  - Measured,
  - Expected: i.e., to be given before departure,
  - Upcoming: announcement that the delay will be published,
  - Unknown

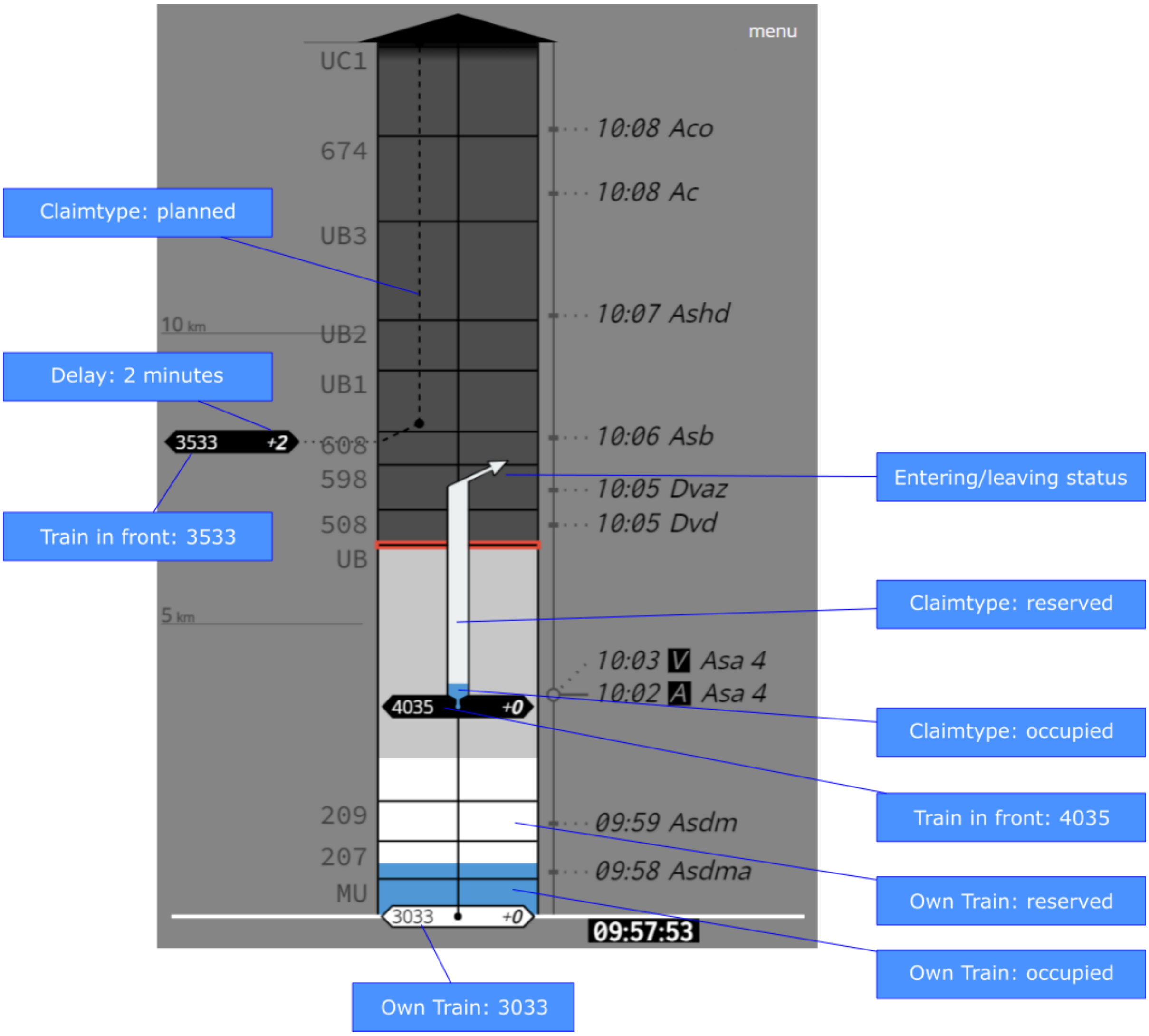


Figure 25: RouteLint: Example from ProRail

In the example in Fig. 25, the ‘own train’ number is 4035. The DMI shows what is happening within the next 15 kilometres. In this example there is one train in front of the ‘own train’: number 3533. It will use the same infrastructure as the ‘own train’. The blue part identifies the “occupied” claim type for the ‘own train’ and the white part above the train number has a “reserved” claim type. The train 3533 with the dotted line is not on the same tracks yet and has a “planned” claim type.

It is necessary to declare the intention to exchange related train information between board and ground during the handshake. The DAS-OB can use the relatedTrainRequest attribute of HandshakeRequest, where it is possible to ask the DAS-TS to get information about the ‘own train’, the related trains or both. The DAS-TS will reply in the HandshakeAcknowledgement by mentioning what related train information is supported.

8.4 Context information

The protocol makes it possible to transmit information that is not used to define the driving profile but is only used to be presented to the driver. This context information can be sent both for Journey Profile and Segment Profile. This information can be integrated in the DMI towards the driver.

Journey Profile context information includes the applicable OTN (which is not the train identifier: it is only used for informational purposes to the driver, as it can change along a journey).

Segment Profile context information includes line and track identifiers, kilometre reference, the communication network and a list of contacts for a certain area, level crossings and connecting tracks.

The following figure represents a possible representation of the different types of connecting tracks.

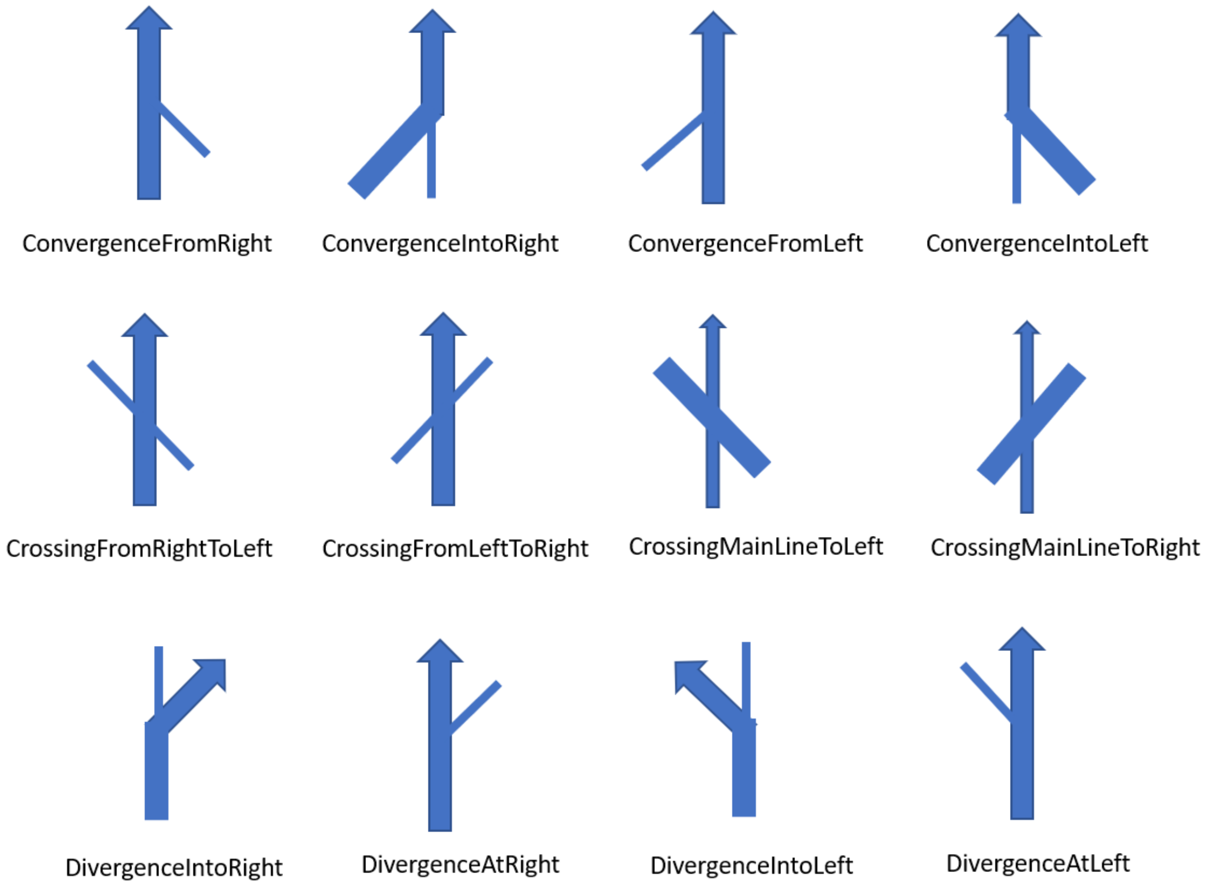


Figure 26: Different types of connecting tracks

9 Implementation hints

This section includes a set of miscellaneous implementation hints for those who wish to use the SFERA protocol.

9.1 Levels of precision

During the introduction of a DAS, the required precision in arrival times and positions and the stability of the advice should be determined. For higher precision, more data elements are needed.

A strategy to start with limited precision and to add fields as they become available might be a feasible phasing. The suggested minimal set is given in Section 1.3.

All other elements will improve performance in terms of timing accuracy and stability of the advice.

9.1.1 Default values if information is not given

The curve radius only has an effect on the rolling coefficients. This will only improve advice messages if Railway Undertakings have very precise rolling coefficients.

The current defaults should be used when no information is given:

1. No gradients means flat.
2. No curves means roll out as straight.

9.1.2 Coding of tunnels and bridges

Tunnels and bridges may be coded. The most important data elements for these are the gradients, as trains will accelerate and decelerate on the ramps. The tunnel area only has an effect on the rolling coefficients. This data will only improve advice messages if Railway Undertakings have very precise rolling coefficients.

9.1.3 Coding of powerless sections

Powerless sections are of importance, as the trains can only coast and brake (without regeneration) and cannot cruise or accelerate in these areas. The DAS should be aware of them. Only use the element CurrentLimitation with maxCurValue equal to 0 and keep the line voltage of an adjacent section in the RatedVoltage and EstimateVoltage fields.

9.1.4 Unavailability of DAS Operating Modes

Section 3.6 of the IRS describes DAS Operating Modes comprising three different components (Driving Mode, Architecture and Connectivity). At the Handshake level, the DAS-OB will send all the triplets supported, and the DAS-TS will choose the pair [Architecture, Connectivity] selected.

For the DAS-TS, it is possible to include a set of unavailable operating modes both in the SP and the JP. For the SP, this is done by defining DAS_ModesUnavailabilityAreas under SP_Areas; for the JP, Temporary Constraints can include Unavailable_DAS_OperatingModes. In both cases, it is possible either to express a triplet [Driving Mode, Architecture, Connectivity] which is unavailable, or to only express one or two of the three attributes; in that case, the values expressed are unavailable independently of the values of the missing attributes.

For example, if the GoA4 driving mode is unavailable in a specific area, it is sufficient to list that value under Unavailable_DAS_OperatingModes; it will then be considered unavailable, independently of architecture and connectivity.

9.2 Aspects for data confidentiality

In SFERA, there are several fields that contain commercially sensitive data (e.g. train weight and length). SFERA data will therefore only be delivered to the RU responsible for the train.

On the other hand, DAS devices may be mounted on a train that runs for several RUs, as it might be owned by a train leasing company, and handheld devices may belong to a driver that runs trains for several RUs.

The SFERA protocol defines the data being exchanged and the circumstances under which this takes place. The SFERA protocol cannot define the elements to be implemented to guarantee the different RUs requirements regarding the confidentiality of their data. The RUs could, for example, require that the data of an integrated DAS shall be deleted when the rolling stock changes RU.

9.3 Changing of train identification

The TrainIdentification field uniquely identifies a train run in SFERA. However, during a run, part of the train identification may change. Known reasons are:

1. At stop (turning, coupled services [many countries]).
2. While the train is running (e.g. in France, passing Paris: odd or even numbering changes).
3. Day boundaries (e.g. in Sweden).
4. Trains that have two numbers, determined by the front engine. The number changes at stations where train turns around.
5. Prefix or postfix change (e.g. in Belgium, when composition or purpose is changed).

The handling of these changes is left to the DAS implementation. SFERA just assumes that the TrainIdentification field is unique. Implementations should not assume that the TrainIdentification field stays constant while the train is running.

9.4 Validation and tests

During the first implementation of a new SFERA version for an RU or IM, it is recommended to validate each message against the XSD. This ensures basic protocol adherence. Validation of each message (also in the production phase) of each message against the XSD is recommended.

In this first phase it is also useful to plot the positions of the virtual balises along a track on a GNSS map and on a GNSS trace from a train run along the tested route. A visual check simplifies the identification of several inaccuracies of the SPs.

9.5 Implementation of blanking near stations and with ERTMS/ATO

In the implementation, care should be taken that drivers are not provided with confusing information.

Current DAS implementations show no advice within 6 km (TGV) or 2 km (other trains) of the approach to a station: the braking phase is calculated, but not shown. This value is appropriate for every train type.

Similarly, if the ERTMS/ATO is active on a cab screen, the ERTMS/ATO display should be the driver screen for SFERA. All other SFERA systems should then be considered in non-driver mode. There are two lines of thought on displaying DAS information in this case: the DAS screen could be blanked to avoid double information. On the other hand, gaining situational awareness when ATO is switched off may be easier if DAS information continues to be displayed. Also, SFERA transmits context information to the driver that is not available in ERTMS/ATO.

9.6 Restriction of braking to optimise regenerative braking

The braking curve for a train can be calculated allowing for regenerative braking only, yielding better energy efficiency. This is done through the TCs. During regenerative braking, slipping rarely occurs.

9.7 How to implement train reordering from the TMS in SFERA

SFERA does not contain a flag “run as quickly as possible” in the SFERA JP. This section describes the handling of two frequent use cases from a TMS and train viewpoint.

Use Case 1: Consider a track A/B–C–D, where two trains run within a few minutes of each other. Train *x* comes from B, arrives in D at 14:19. Train *y* comes from A, passes C, and arrives in D at 14:22. Train *x* waits for *y* to arrive in D and then departs at 14:24 after taking passengers across the platform from *y*.

Now, if *x* is delayed by a few minutes, then *y* could run ahead of *x*. One solution would be to use the flag “run as quickly as possible”. However, *y* might run too fast. As long as the cross-platform connection is maintained, train *y* does not need to run more than the conflict-free time ahead of *x*. The TMS should be able to find the minimal running time of train *x* minus the conflict-free follow-up time.

Conclusion: The TMS just communicates a new arrival time window for the first train. Energy-efficient driving will lead the train to the end of that window.

Use Case 2: Trying to clear a line because of danger. The current status is that no safety-related communication should be reported and that SFERA communication should not replace other communication means. In this use case, other communication would be required.

9.8 When to use aimEarlyArrival

This section explains the use of the field aimEarlyArrival. As seen in Section 9.7, the TMS will be in control in situations where two trains are re-ordered. The situation is different, however, if there are two or more trains waiting behind each other on a line, for example after a blockade. In that case, the aim should be to clear that line as quickly as possible in order to get back to regular operation as soon as possible. For this, the attribute aimEarlyArrival (“be as near as possible to the beginning of the arrival window”) is included in TimingPointConstraints. This can be used if a group of trains is expected to be delayed and the delay for the entire queue should be minimised. The advice to the driver will not be different from driving to the timetable, as all of these trains are delayed.

9.9 Minimum data for “Timetable mode” operation

Before implementing the prerequisites for C-DAS-O or C-DAS-C operation (TMS and/or data description), data is necessary for drivers to give them their timetable and provide the information that is historically provided “on paper” (time objectives in the different stations). This helps provide minimum information to drivers, essentially enabling them to calculate the current delay. The “Timetable” driving mode can be used for this purpose.

The following is considered to be the minimum data necessary for “Timetable mode” operation:

- Segment Profile:
  - Administrative elements in the SP (ID, Versioning, Status).
  - Length of the SP.
  - Timing Points in the points in the network where timetable information will be given, including TP_Name and location and identifier in “StoppingPointLocation”.
  - Balises and Virtual Balises associated with those timing points, for the C-DAS to calculate delays (Virtual Balises for C-DAS tablets, Balises for C-DAS receiving odometry from the TCMS).

- Balises and Virtual Balises at the points of transition between parts of the network with only Timetable information.
- Restrictions saying that only the “Timetable”, “Read-only” and “Inactive” driving modes are possible on this part of the journey.
- Next_SP_Zone Information.

Other elements can be added if available, if they are useful for the driver. Sufficient balises/virtual balises shall be provided around the beginning and end of those elements for the C-DAS to know how to locate the elements on the DMI.

- Journey Profile:
  - Administrative elements in the JP (TrainIdentification, JP_Status, JP_Version).
  - Applicable OTN.
  - TimingPointConstraints with
    - TP_ID, TP_PlannedLatestArrivalTime, daylight saving
    - StoppingPointInformation:
      - Stop type (Departure Time, PlannedDepartureTime, daylight saving)

Other elements can be added if available, if they are useful for the driver. Sufficient balises/virtual balises shall be provided around the beginning and end of those elements for the C-DAS to know how to locate the elements on the DMI. This might prove tricky when frequent coordinates cannot be determined by the IM for temporary constraints which will usually start/end at random points of the track.

9.10 Push/Pull Strategy

Two major strategies exist for communicating data between two systems:

- Push: the data provider generates a message towards the data receiver as soon as new data is available;
- Pull: the data receiver interrogates the data provider when it needs the data, and the data provider responds with the payload if new information is available, or answers “no new data” if this is not the case. In this strategy, the interrogation is usually cyclic, with an adapted interrogation frequency regarding the maximum acceptable latency needed by the data receiver.

In SFERA, the push strategy is implemented using the B2G Event messages and G2B Event messages. To permit the adoption of a pull strategy, all the payloads of the event messages can be provided to the data receiver through “request+reply”. For example, the G2B Request message structure includes “xxx_Request” payloads for all payloads present in the B2G event message, and the B2G Reply message includes all the payloads from the B2G event message.

The pull strategy has been included in the SFERA protocol mainly for the IM DAS-TS/RU DAS-TS link, where it is intended to facilitate implementation regarding degraded states: in loss of connectivity or availability situations for example, the data receiver knows its state best (what is the last message received; did the degraded state lead to losing the previously cached data), which simplifies the functions to implement.

Intuitively, a pull strategy leads to latency in comparison with the push strategy (the maximum being roughly the time between pull requests from the data receiver). However, this latency can be adapted to an acceptable level if the interrogation period is reduced to a value inferior to the acceptable latency. Again, this can intuitively lead to a message overhead if the interrogation period is reduced to low values (e.g. an interrogation every second or hundreds of milliseconds). This can be mitigated using the possibilities offered by the SFERA XSD by mutualising the interrogations. For example, regarding JP updates for the RU: a single request message can contain JP requests for multiple trains. Using this possibility, the RU can send periodic messages containing the requests for all trains in operation, or being prepared for operation (and the latest known version of the JP for each train). For each message sent, the IM can provide two answers:

- A reply message containing the Journey Profile of all the trains that have a new version of their Journey Profile;
- A reply message containing a simple “OK” response which indicates that no new Journey Profile has been published.

10 Considerations on choices made for SFERA

When developing the SFERA protocol, it was necessary to make a number of choices. This section explains the choices that were made in the course of the development work.

10.1 Generic background

10.1.1 Adhesion management

Lowered adhesion may lead to the slipping of trains if the driver tries to accelerate or brake too forcibly. For DAS operation, lower adhesion will limit the acceleration of trains and trigger earlier braking in order to reduce the risks. This affects the running times of trains (a known effect on punctuality in autumn in many countries). Reduced adhesion depends on the fraction of axles that are powered and on the axle load of the powered and braked axles. The effect of reduced adhesion will therefore vary by train type.

SFERA assumes that Infrastructure Managers can make an estimate on the adhesion situation (e.g. at the beginning of the day an estimate per train type related to the weather). Based on the feedback per train run, it can then be evaluated if the estimate per train type should be raised or lowered.

The running times in the DAS will be adjusted if the adhesion estimate is lowered. For DAS with a connection to the TCMS, SFERA messages may be used to indicate back to the IM if the estimate is accurate. For drivers on trains where the DAS is not connected to the TCMS, the same IM rules on reporting changes in adhesion status will be used as for trains without DAS.

Feedback on adhesion can only be given if brake/traction is near the limit of the expected adhesion, or if an unexpected slip occurs. The service brake is usually applied below this limit. Traction usually yields more data points, as more axles are usually braked than tractioned. The models in the trackside should take this information into account per train type. Do not include fields on adhesion feedback in SFERA feedback messages unless feedback can be given.

No information is given in SFERA for scheduling for dynamic brake tests; the regulations vary too much by country. For example in Switzerland, “Bremsprobe auf Wirkung” (basically: test if brakes work) are required once for all trains after their initial departure. When a train has cleared the departing station, the driver is obliged to operate the brakes and check if the train reacts. In France, this running brake test exists for freight trains only, when they can perform it, if signalling and profile authorise its execution. There is no distance or time limit for its performance.

DAS advice messages shall take these additional running times for the running brake test into consideration for situations that most commonly occur, as in the Swiss and French situation. No information is provided to update the running time estimates for an additional test in snowfall, etc. (e.g. as required in the UK). Adhesion reporting during dynamic brake tests can only be given if the brake application is near expected adhesion limits, or on unexpected slips.

SFERA permits the use of a low adhesion rate (value of the traction and braking capacity remaining, in %) or the low adhesion categories as defined in SUBSET-126. Extra information on low adhesion can also been found in EN 15595.

10.1.2 Reasons for power management

At most railways, the catenary power system is dimensioned for absolute worst-case situations (i.e. many spontaneously accelerating trains with all heaters on). This leads to power stations rarely being used at peak. Instead, it is possible to curb the peaks with limited impact on punctual driving. The worst cases can be local and limited by both power availability and rail-to-earth potential, or national, if a limited number of feeding points are used.

SFERA provides the communication toolbox to enable the more effective use of a limited power supply.

10.1.3 Cant and cant deficiency

For the uninitiated, the terms cant, super-elevation and cant deficiency may be challenging. This section presents an introduction and some references. Cant (also called super-elevation) is the height difference between the outer rail and the inner rail in a curve, and is usually defined in mm[5].

For a given train speed and super-elevation, cant deficiency (CD) is defined as

$$CD = \frac{gauge}{\sqrt{1 + \frac{R^2 g^2}{\left(\frac{v}{3.6}\right)^4}}} - superelevation$$

Where

- $gauge$ is the distance between two rails (in mm)
- R is the radius of curvature (in m)
- g is the gravitational constant (m/s²)
- v is the train speed (in km/h)[6]
- $superelevation$ is in mm.

Note that the first term is small for a large radius of curvature and low speeds (value in the denominator is large). In that case, the cant deficiency is negative. For high speeds, the CD will become positive. Each train type has a maximum CD. The maximum CD is set mainly for passenger comfort reasons. The maximum CD can be increased using tilting mechanisms on the train as long as safety constraints are fulfilled (no risk of derailment).

Some tracks have speed profiles that depend on the CD of a train (and thus also on the tilting mechanism being operable). Therefore, ETCS and modern ATP systems such as the Swedish ETC have speed profiles that can be set for a certain maximum CD and where the train selects the correct curve based on operational parameters. The DAS needs to be aware of the maximum CD of a train. Wear on rails and wheels is lowest when CD = 0. Also, rolling resistance is lowest when CD=0. Thus, there may be some advantage in not running at very low speed in bends. Furthermore, passenger comfort is optimal at CD=0. For these reasons, SFERA provides the opportunity to transmit the cant along a track.

For rolling resistance, there is no agreed formula for additional rolling resistance due to CD not being equal to 0, even though curve resistance[7] is described.

For ERTMS trains, the gauge is transmitted. Super-elevation is not necessary for safety aspects, and therefore not included. For SFERA it is optional to transmit gauge and super-elevation in order to calculate rolling resistance.

[5] See https://en.wikipedia.org/wiki/Cant_(road/rail).
[6] See https://en.wikipedia.org/wiki/Cant_deficiency.
[7] See https://en.wikipedia.org/wiki/Curve_resistance_(railroad).

Per train type, a maximum CD can be set by the train owner. ERTMS trackside holds maximum speeds (calculated by the formulas above) and the lowest super-elevation of the track in that segment. SFERA on an ERTMS track needs to know both maximum CD (in TC) and the maximum speeds per CD category from the track. The latter is already available in SUBSET-126.

Note that both the speed profile per CD category and super-elevation will vary along the length of a segment. Therefore, both are structured to vary inside the SP.

10.1.4 Relationship of SUBSET-126 / SUBSET-026 / SUBSET-125 with definitions in SFERA

Some readers may not be familiar with the documentation structure of ERTMS and ERTMS/ATO. This section gives a short primer. ERTMS/ETCS is a signalling standard that is being rolled out in Europe. The definitions are maintained by ERA and are publicly available. The ERTMS/ETCS definitions are split up into “SUBSETS” per aspect. For SFERA, the most important SUBSETS are given in Table 18.

Table 18: Overview of ERTMS/ETCS SUBSETS relevant for SFERA

SUBSET #	Contents
126	Functional data definition of ERTMS/ATO. Defines Journey Profiles, Segment Profiles and Status Reports. Part of the Segment Profile is an extension of SUBSET-026 data past the current movement authority.
125	The system specification of ERTMS/ATO. It defines the requirements on which SUBSET-126 is based
139	ATO train interface. In ATO, it is assumed the DAS can get its information on the train from the Train Control Management System (TCMS). SFERA provides this information via Train Characteristics.
026	System requirements specification, including Chapter 7 ERTMS/ETCS language
ERA_ERTMS_040001	Assignment of values to ETCS variables: defines many numerical values used in ETCS.

As ERTMS/ETCS is a signalling standard, SUBSET-026 will not provide information past the first signal at stop. For DAS it is the information up to the next stopping point of the train that enables optimisation. SUBSET-126 extends the data between the first signal at stop and the next stopping point of the train. This is done using the same terms as in SUBSET-026. Therefore, the DAS/ATO system can replace the original (possibly less granular) data from the Segment Profile with actual data up to the current movement authority.

SFERA enables the transmission of more detailed data than SUBSET-126. The design of SFERA is such that all data in SUBSET-126 can be expressed easily in SFERA. The ways to express driving profiles, gradients and virtual balises may not be intuitive for DAS designers. They are specifically designed this way for the purpose of compatibility.

10.1.5 Relationship of RailML, RailTopoModel and EULYNX with definitions in SFERA

Other standards were considered for the definitions that SFERA uses.

RailML and RailTopoModel were considered for Segment Profiles. However, they do not meet the requirement of compatibility with SUBSET-126.

EULYNX was considered for the signals. However, EULYNX communicates only on the shape of the signal, not on its function for the driver. Therefore, each country has several codes in EULYNX for the same function for the driver. Furthermore, many EULYNX codes are more or less country-specific.

10.1.6 Choice of XML in SFERA

Several other standards were considered for the definitions SFERA uses. It was decided to use a text-based format (rather than binary, as in SUBSET-126), mainly for ease of extension.

The decision to use XML rather than JSON was made because of its use in the rail industry (e.g. TAF/TAP, RailML) and the range of validation tools.

10.1.7 Choice of XSD field names

The XSD elements are named using the following rules:

- upper camel case for elements;
- lower camel case for attributes;
- all uppercase for abbreviations within identifiers (e.g. JP, SP, TC);
- underscores are used when there are acronyms (e.g. DAS_StatusReport).

For elements that resemble a SUBSET-126 element, a name corresponding to the SUBSET-126 identifier is chosen. Elements that exactly match the SUBSET-126 definition have a tag [SUBSET-126] in the XSD annotations to indicate elements/attributes/text directly quoted from SUBSET-126.

10.1.8 Choice for message signing and message encryption

SFERA does not include separate encryption and signing on a payload level.

- In IM-RU setup, IM and RU jointly define the protocols and decide whether to use encryption. It is possible that parts of the communication between IM and train (through back office RU) happens with legacy protocols. The RU back office can also enhance messages in both directions.
- In the IM-Train setup, encryption and signing is already handled at protocol level. MQTT already uses TLS encryption and verification to make sure it will only communicate with a verified service.
- In ERTMS/ATO mode, any additional encryption and signing between the IM trackside and train is useless, as it will pass through a SUBSET-126 translation which implements its own security.

10.2 Considerations for RAMS (Reliability, Availability, Maintainability and Safety) aspects

10.2.1 Aspects for design for safety and reliability

Systems using SFERA shall not be detrimental for safety and reliability of the railway system. SFERA is not designed as a safety-critical system. Therefore, its use should be strictly limited to the purpose for which it was designed: data exchange for DAS and ATO GoA 2 on Class B systems.

Some functions are required for any DAS using SFERA:

- Blanking DAS information in DAS prohibition areas (e.g. close to stations).

SFERA explicitly **does not** include the following functions:

- Feedback from the driver to traffic management that data was received: no user interaction should be requested while driving;
- Feedback from the driver to traffic management (e.g. unplanned stops): standard communication channel (e.g. phone) should be used to secure a single process;
- Safety-critical adhesion reporting: this should be done outside SFERA. The adhesion function in SFERA is only for performance tuning (mainly during acceleration).

Furthermore, the use of SFERA data for the following purposes is strongly discouraged:

- Showing speed advice messages that could conflict with restrictions (showing speed differences against line speed with coasting advice, unless there are technical means for the ATP system to enforce all restrictions);
- Showing Temporary Speed Restrictions in a way that the driver will start to rely on it;
- Showing when to start braking and how much, otherwise the driver will start to rely on it (has an impact especially in degraded situations, e.g. unforeseen low adhesion);
- Showing expected lineside signal aspects on a DAS display (drivers could get confused);
- Use changes in TC to change TAF/TAP information on trains (the protocol is not designed to be that reliable, and train lengths are safety-critical).
- Use SFERA position reports as sole position information for traffic management (the protocol is not designed to be that reliable).

10.2.2 Power demand

SFERA power management use cases and data structure are advanced applications and are expected to be used experimentally in the first years. This section gives an introduction into this functionality.

SFERA implements a way to communicate limits in power uptake, and feedback on expected power usage. SFERA assumes that the IM implements stable algorithms for regulation. Simple feedback systems that sum the total required power and throttle back in retrospect should be avoided. Lessons should be learned in designing this kind of feedback loops from the implementation of EN 50163, where direct feedback mechanisms could lead to oscillations[8]. Instead, algorithms should be used that forecast the power uptake based on the Train Path Envelope and TC. The percentage of trains being receptive to power limitations may be derived from the SFERA status. Based on the feedback on power usage from the DAS, the IM can validate that its effect will be displayed to the driver.

On lines equipped with ETCS Level 2, it is preferable to send power limitation messages using packet 40 Track Condition Change of allowed current consumption. Using ETCS messages for power control delivers them in a more secure fashion.

10.2.3 Selecting train identification is done outside SFERA

The TrainIdentification field uniquely identifies a train run. SFERA does not provide the means to transmit train identification to a DAS device. The reason for this is that such selection is highly dependent on the situation and processes of the RU. The information in routing of personnel and rolling stock is usually not shared with the IM. With a train-mounted DAS, the train roster might define the runs of the train. With a DAS in a personal device for the driver, their roster may define the order. For integrated devices, the RU may have other means to pre-fill TrainIdentification, as many other systems will require this number as well. Therefore, the SFERA protocol leaves it up to the RU to implement this in their fashion.

10.2.4 Continuous feedback is done outside SFERA

The DAS can provide feedback on many of the SFERA elements. Gradients, radius of curvature, positions and distances of virtual balises, signal positions (including reactions from ATP systems), speed profiles, traction and braking profiles can be validated on consistency. The fact that the SFERA protocol does not provide a means to provide this feedback is intentional. Such feedback should be processed by the RU and provided in a consolidated form to the IM. It was decided when designing SFERA that real-time feedback to the IM on individual events in each DAS would only confuse the interpretation of the feedback.

[8] See Buhrkall, L & Danielsen, S & Eisele, A & Bergman, M & Galic, J. (2010). *Low-frequency oscillations in Scandinavian railway power supply - Part I: Basic considerations*. eb - Elektrische Bahnen. 108. 56-64. https://www.researchgate.net/publication/289661505_Low-frequency_oscillations_in_Scandinavian_railway_power_supply_-_Part_I_Basic_considerations.

10.2.5 Time synchronisation is done outside SFERA

Driver Advisory Systems need a time that is synchronised within a few seconds. The fact that the SFERA protocol does not specify how time synchronisation should be performed is intentional, as various effective methods exist. Synchronisation with other systems on board the train might set stricter requirements than the operation of the DAS. Furthermore, most DAS have a data connection. Several protocols (e.g. NTP, GNSS or GSM network time or time from GSM antenna) are available to synchronise the time within the required range.

There are many possible time bases (GNSS measurement time, DAS system time at sending message, DAS system time at measurement). Depending on the performance requirements of the RU and IM, these may or may not be important. Accuracy should be included in the RU-IM agreement.

10.2.6 Latency requirements

Latency of data transfer is most visible once a new train running number is selected. Not only a new JP needs to be downloaded, but also new SPs and TCs. Acceptance by the driver may depend on the responsiveness of the DAS.

Drivers may also want C-DAS to be updated to new situations within a few seconds once they become aware by other means of new situations arising. For example, a platform change or change in arrival time given by the travel information system may also lead to new driving advice (see Section 4.6.3). Similarly, a train dispatcher may require a driver to respond quickly (within seconds) to a change in train order by changing its speed. On the other hand, latency may not be an issue in many situations: a change in schedule further ahead or downloading SPs for caching are not critical and should be allowed to last longer. Furthermore, there may be longer stretches with no connectivity where the DAS should continue functioning even though not all the latest updates have been processed.

There are no fixed requirements on latency in the SFERA standard. For the IM-RU setup, latency will usually be determined by the RU back office to DAS communication. For the ERTMS/ATO setup, the ERTMS/ATO standard sets limits. For the IM-Train setup, the infrastructure set up by the IM will be the limiting factor.