UIC CODE 406

1st edition, September 2004

Original

R

Capacity

La Capacité
Kapazität

UNION INTERNATIONALE DES CHEMINS DE FER
INTERNATIONALER EISENBAHNVERBAND
INTERNATIONAL UNION OF RAILWAYS

Leaflet to be classified in Volume:

IV - Operating

Application:

With effect from 1 September 2004
All members of the International Union of Railways

Record of updates

1st edition, September 2004 First issue

The person responsible for this leaflet is named in the UIC Code

Contents

Summary

The rising volume of traffic, with simultaneously increasing demands in terms of quality and quantity, requires a unique, harmonised and generally-valid understanding to be developed as regards available railway-infrastructure capacity.

The method presented in this leaflet enables Infrastructure Managers (IMs) to carry out capacity calculations - following common definitions, criteria and methodologies from an international standpoint - for lines/nodes or corridors based on different criteria such as traffic quality (market needs or company requirements), timetable quality (requirements of timetable compilers) or effective and economical utilisation of infrastructure (IM requirements).

Capacity as such does not exist. Railway infrastructure capacity depends on the way it is utilised. Therefore, the proposed methodology is based on examination of an existing pre-constructed timetable for train operation on the particular infrastructure. This timetable, which expresses market needs, shall be analysed within a line section through compressing timetable train paths in a pre-defined time window. During the compression, it is not just timetabled train paths that are taken into account: any occupation times, even when not related to the train paths, must be incorporated.

The result of the compression process is the so-called infrastructure occupation. The capacity consumption shall be measured by infrastructure occupation, to which is added buffer times for timetable stabilisation and, where necessary, maintenance requirements.

If the determined infrastructure occupation figure is higher than a typical value, infrastructure shall be declared congested. If this typical value is not reached, a leftover capacity shall then become available. The way in which this leftover capacity can be used for additional train paths shall be determined for each single case on the basis of market needs.

The application of the compression methodology on different lines on several European networks showed that it is always possible to insert additional paths, but one must stop at a certain level due to stability requirements specific to the timetable. It appears impossible to assess standard precise values of the infrastructure occupation time, since generally valid determination of the necessary buffer times that are dependent on many factors has not been possible up to now.

However, in the interests of comparability of the results, this approach is intended as a first approximation. For this reason, only some guideline figures are proposed, which must be further developed through complementary analysis.

1 - Introduction

For an improved understanding of the definition of capacity and the method of calculating it, proper account must be taken of external and internal conditions. If these conditions - several of which cannot be influenced by the Infrastructure Managers (IMs) - are unknown, this can result in erroneous interpretations of the term capacity and in misunderstandings of the information contained in this leaflet.

1.1 - Aim of the leaflet

The rising volume of border-crossing traffic in Europe, with simultaneously increasing demands in terms of quality and quantity, requires a unique, harmonised and generally-valid understanding to be developed as regards available railway-infrastructure capacity.

Any railway-infrastructure evaluation, in order to be generally valid, must be underpinned by a common definition of capacity among railway infrastructure managers.

Due to the different concepts and procedures concerning capacity and the resulting calculations applied by IMs, a comparison is not feasible and general conclusions are not possible, which means that a unique procedure must be developed.

The method presented in this leaflet enables IMs to carry out capacity calculations - following common criteria and methodologies from an international standpoint - for lines/nodes or corridors based on different criteria such as traffic quality (market needs or company requirements), timetable quality (requirements of timetable compilers) or effective and economical utilisation of infrastructure (IM requirements).

1.2 - Basic parameters underpinning capacity

Capacity as such does not exist. Railway infrastructure capacity depends on the way it is utilised. The basic parameters underpinning capacity are the infrastructure characteristics themselves and these include the signalling system, the transport schedule and the imposed punctuality level.

On a given infrastructure, capacity is based on the interdependencies existing between:

- the *number of trains* (per time interval, e.g. trains per hour). When train intensity increases, less capacity is left for quality, as expressed in the parameters described below;
- the *average speed*. The braking distance increases proportionally more than the average speed;
- the *stability*. Margins and buffers have to be added to the running time of trains and between train paths to ensure that minor delays are suppressed instead of amplifying and so causing (longer) delays to other trains;
- the *heterogeneity*. When the differences in running time between different train types worked on the same track are great, similarly the capacity consumption of the same number of trains will increase proportionately.

The relation between these parameters is clearly shown in the "capacity balance", as illustrated in Fig. 1 below. In this qualitative model, an axis for each parameter is drawn from a unique origin. A chord links the points on the axes, corresponding to the value of each parameter. The length of the chord represents *the* capacity. Capacity utilisation is defined by the positions of the chord on the four axes. Increasing capacity means increasing the length of the chord.

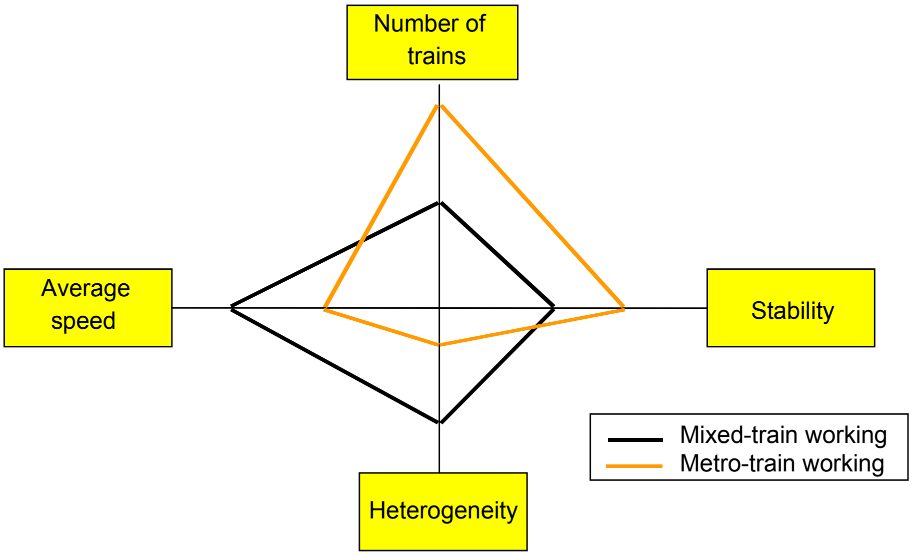


Fig. 1 - Capacity balance

1.3 - Different views of capacity

Due to differences in requirements, capacity is viewed differently from the position of the market, infrastructure planning, timetabling and operations. These are summarised in Fig. 2.

Market (customer needs)	Infrastructure planning	Timetable planning	Operations
expected number of train paths (peak)	expected number of train paths (average)	requested number of train paths	actual number of trains
expected mix of traffic and speed (peak)	expected mix of traffic and speed (average)	requested mix of traffic and speed	actual mix of traffic and speed
infrastructure quality need	expected conditions of infrastructure	existing conditions of infrastructure	actual conditions of infrastructure
journey times as short as possible	time supplements for expected disruptions	time supplements for expected disruptions	delays caused by operational disruptions
translation of all short and long-term market-induced demands to reach optimised load	maintenance strategies	time supplements for maintenance	delays caused by track works
		connecting services in stations	delays caused by missed connections
		requests out of regular interval timetables (system times, train stops, ...)	additional capacity by time supplements not needed

Fig. 2 - Different views of capacity

While capacity demands from the market standpoint are oriented towards satisfying peak values, infrastructure planning by contrast must be oriented towards a definition of capacity that, on average, guarantees profitable utilisation of the infrastructure. From a timetable point of view, capacity considerations bring together a given infrastructure and existing demands for train paths. From an operational point of view, on the other hand, the capacity situation is always in flux and depends on current infrastructure availability, progress with building measures, delays, diversions and the number of additional trains.

Each of these different situations is correct in terms of its own specific background. The parameters mean, however, that each viewpoint leads to a different capacity-requirement result.

1.4 - Capacity-relevant constraints

Alone, each set of circumstances mentioned below can lead to considerable capacity limitations. If, however, several points come together, this translates into virtually insoluble capacity problems.

1.4.1 - Priority

Existing priority regulations influence available capacity, through which the mix ratios can be specified or influenced. Different priority regulations can have a more restrictive effect on cross-border traffic than capacity restrictions for other reasons. IMs cannot influence priority regulations directly but can give advice about how to flesh them out.

1.4.2 - Timetable structure

Small and large-scale traffic networking, e.g. integrated regular-interval timetables for local transport, line systems for long-distance passenger trains, create a variety of constraints both on routes and at junctions. Line and integrated timetable systems are relatively rigid and result in infrastructure occupation with little train-path planning flexibility. Particularly critical is the superimposition of several different systems on the same infrastructure, which then leaves hardly any room at all for other traffic.

1.4.3 - Capacity allocation process

Different schedule chains and methodologies for capacity allocation constitute constraints, especially in the case of cross-border train pathing. Any conflicts arising are dealt with in different ways in accordance with different rules whereby significant differences exist in capacity allocation for freight traffic at short notice.

1.4.4 - Design rules

Differences exist in the design method, in the determination of the possible physical journey time, in the treatment of time allowances to ensure the individual train-path timetables plus timetable stability, as well as in the consideration given to building and maintenance measures. All of these factors can translate into restrictions for cross-border traffic.

1.4.5 - Environmental protection

Rules for environmental protection, e.g. those relating to noise emissions, have a restrictive impact not only on the volume of all traffic but also on special, noise-intensive train movements, e.g. at certain times of the day. Prohibitions on hazardous loads exclude the use of certain routes and force traffic onto predefined routes and junctions.

1.4.6 - Safety aspects and technical constraints

Passing prohibitions imposed on certain types of traffic for a variety of reasons are highly restrictive on, for instance, routes with numerous tunnels or high-speed lines. Only defined routes may be used for certain types of traffic which, if moved by other routes, can result in considerable restrictions in terms of capacity. This applies equally to the use of certain types of braking systems. The power supply can also have a limiting effect on capacity, namely in respect of the type of power supply and the power available.

1.4.7 - Theoretical capacity

A theoretical maximum capacity expressed in terms of the maximum number of trains can be calculated by defining ideal circumstances. In this connection, a scenario is proposed that will never arise if account is taken of the points mentioned above, as well as of realistic demands on the mixing ratios. Accordingly, the value calculated can only be a theoretical one. To calculate theoretical capacity, the following conditions must be assumed:

- absolute train-path harmony (same parameters);
- shortest possible spacing for all trains;
- incorporation of the normative section on the corridor under the constraints of national quality conditions.

This already reflects the theoretical nature of the results. Conditions such as these can only be found on routes that are dedicated to no more than one type of traffic, e.g. metro lines. Therefore, it is not possible to determine a theoretical capacity within the framework of a generally-valid definition and method of calculation.

2 - Definitions

2.1 - Prerequisite

A unique, true definition of capacity is impossible.

As mentioned already in point 1 - page 2, there are different views of capacity. Because of these differences and given the various consequences of capacity-relevant constraints, a generally-applicable definition is not appropriate.

Thus, the definition presented below (point 2.2) is supposed to work for as broad a spectrum as possible.

2.2 - Definition of capacity

The capacity of any railway infrastructure is:

- the total number of possible paths in a defined time window, considering the actual path mix or known developments respectively and the IM's own assumptions;

- in nodes, individual lines or part of the network;

- with market-oriented quality.

This must also take account of the IM's own requirements.

2.3 - Additional definitions

2.3.1 - Corridor

All possible journey routes (main route or alternative routes), according to market needs, between a defined source and target.

2.3.2 - Route

Consecutive lines and nodes as a whole, between a defined source and target.

2.3.3 - Line

A link between two large nodes and usually the sum of more than one line section.

2.3.4 - Nodes

Points of a network in which at least two lines converge. Nodes can be stations or junctions. They can be differently-sized, depending on the number of converging lines and their tasks.

2.3.5 - Stations

Points of a network where overtaking, crossing or direction reversals are possible, including marshalling yards.

2.3.6 - Junctions

Point of a network in which at least two lines converge and neither overtaking, crossing nor direction reversals are possible.

2.3.7 - Line sections

The part of a line, in which:

- the traffic mix and/or the number of trains,

- the infrastructure and signalling conditions

do not change fundamentally.

It consists of one or more coherent sections, which are limited by two neighbouring stations or nodes.

2.3.8 - Part of a network

Specific combination of nodes and lines to fulfil specific tasks and controls.

2.3.9 - Relevant block section

Block section within the chosen line section, which determines the minimum headway along the entire chosen line section.

2.4 - Market needs

Infrastructure capacity is related to the ability of the particular infrastructure to offer train paths in accordance to market needs as represented by customer requirements.

The IM customer is the Railway Undertaking (RU) which expresses its requirements in the form of path requests. These path requests result from the combination of:

- end customers' (passengers, carriers) needs and requirements,

- the needs of RUs themselves (e. g. type, utilisation or maintenance of rolling stock).

The path request is related to two parameters:

- a typical running time (depending on the performances of the rolling stock),

- a departure time and/or arrival time.

The IM must comply with the requirements of these parameters.

During the negotiation process in respect of capacity allocation, an iteration process occurs between RU requirements and IM offers.

As long as the alteration following the iteration is accepted, the market needs are fulfilled. The infrastructure is not to be considered as saturated.

Requirements for punctuality (time supplements and buffer time) are incorporated in any case.

3 - Calculation of capacity consumption

3.1 - Approach

The proposed methodology is closely related to the definition of capacity. It takes into account the fact that the level of capacity consumption is the only value that can be measured objectively by reference to all parameters mentioned in the definition. Capacity itself cannot be measured because of the reasons described in point 1 - page 2.

The basic necessary condition for capacity examination shall be the existing pre-constructed timetable (a real operational one or a case study) for train operation on the particular infrastructure. This timetable, which expresses market needs, must be analysed from the viewpoint of capacity consumption. Only scheduled trains or trains that are highly likely to run need to be incorporated.

Capacity consumption shall be analysed within a line section through compressing timetable train paths in a pre-defined time window. The effects of the compression on neighbouring line sections are not taken into account. This is acceptable because the analysis must be done for the limiting section of the line, and no conclusions concerning the timetable feasibility on neighbouring line sections shall be derived from this analysis.

If the result of the capacity examination is that there is some capacity left, this unused capacity can be either unusable or can be assigned to some additional train path(s). This depends on possible market needs concerning the type and exact time position of additional train paths. The existence of capacity leftover itself does not automatically mean the possibility of inserting additional train paths into a timetable.

3.2 - Suitable areas

Calculation of capacity consumption using the compression methodology must be applied on a single line section.

In order to assess capacity and bottlenecks for a line or a whole route, the capacity consumption of every single line section shall be calculated. The highest value of capacity consumption on a line section shall determine the capacity consumption along the whole line or route. In order to assess capacity along a corridor, an analysis shall be carried out for every possible route along the corridor (e. g. on alternative routes as well as on the main route).

3.3 - Suitable time windows

Capacity consumption varies according to time of day, weekday and season.

The basis for the compression is one representative day (Thursday for example) over a peak period at least two hours long.

Factors to be taken into account:

- all customer-requested paths insofar as they are part of the yearly timetable;
- all paths pre-planned by the IM without any customer requests (e. g. Freight Freeways); in this case, an indication of the average utilisation of the paths concerned shall be given;
- shunting movements in nodes and marshalling yards;
- track occupation by rolling-stock (vehicle stabling on the track for longer than commercially needed).

In general the time windows shall be fixed by reference to the purpose of the analysis. Recommended time windows shall be:

- peak hours (at least one hour);
- 24 hour daily total (determination of a representative weekday), although it should be noted that a full 24 hours is never available due to reasons such as maintenance work.

3.4 - Compression

Delimitation of the area (point 3.2 - page 9) and of the time window (point 3.3 - page 9) without neglecting the peculiarities of single track lines and the definition of that block section at the beginning of which the infrastructure occupation time must be measured. On single-track lines the measuring points must be installed in crossing stations.

For compression purposes, all single train paths are pushed together up to the minimum theoretical headway according to their timetable order, without recommending any buffer time.

This compression can be done:

- by constructing graphical analysis or suitable tools for this case

  or

- by analytical calculation.

Rules of compression

During the compression process, neither the timetable running times, nor the given overtakings, crossings or stopping times (which are requested by railway undertakings), may be changed.

3.5 - Time shares

The basics for the investigation into capacity consumption shall be the time shares of a timetable and of a compressed timetable.

Fig. 3 and 4 - page 12 show the time shares of a non-compressed timetable using the examples of double and single track lines.

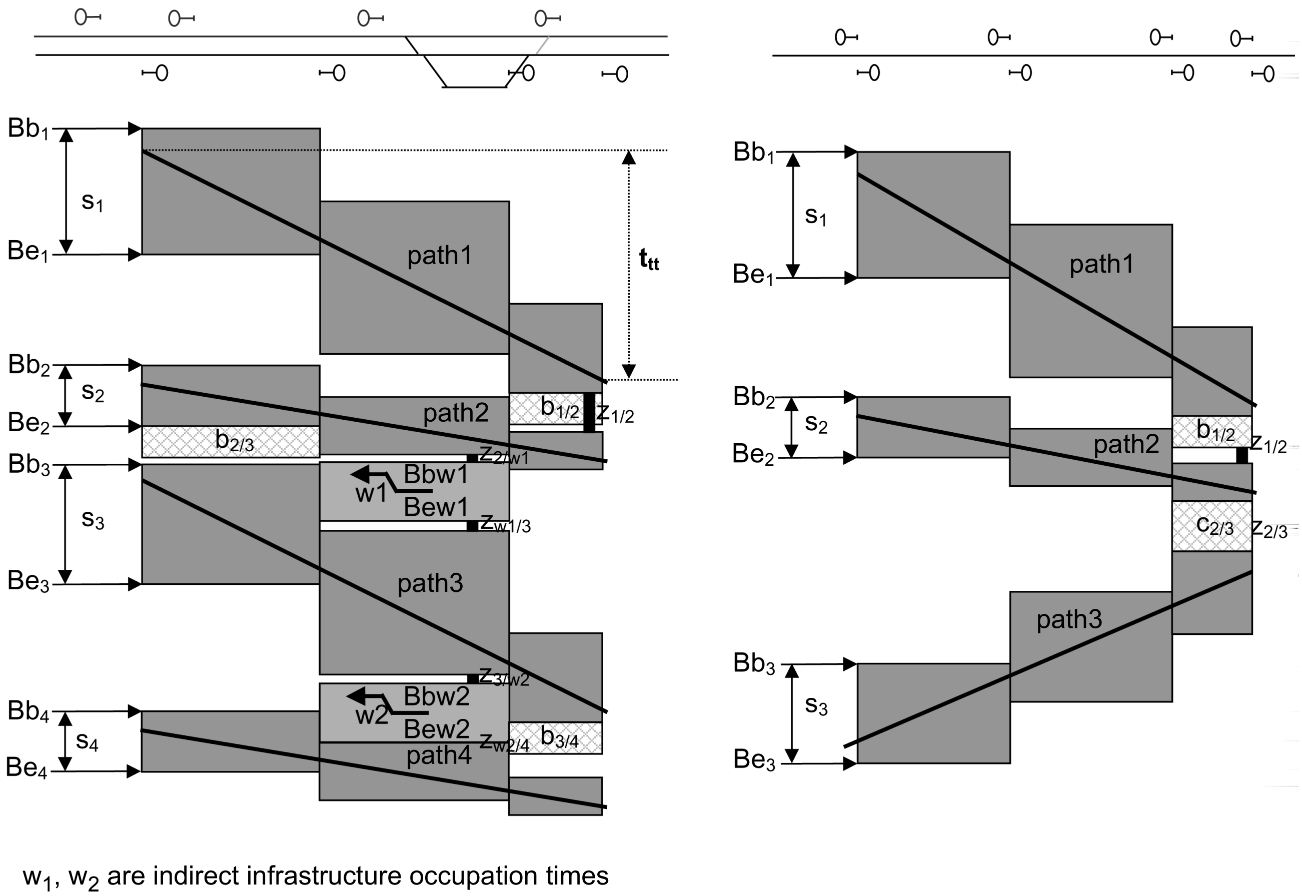
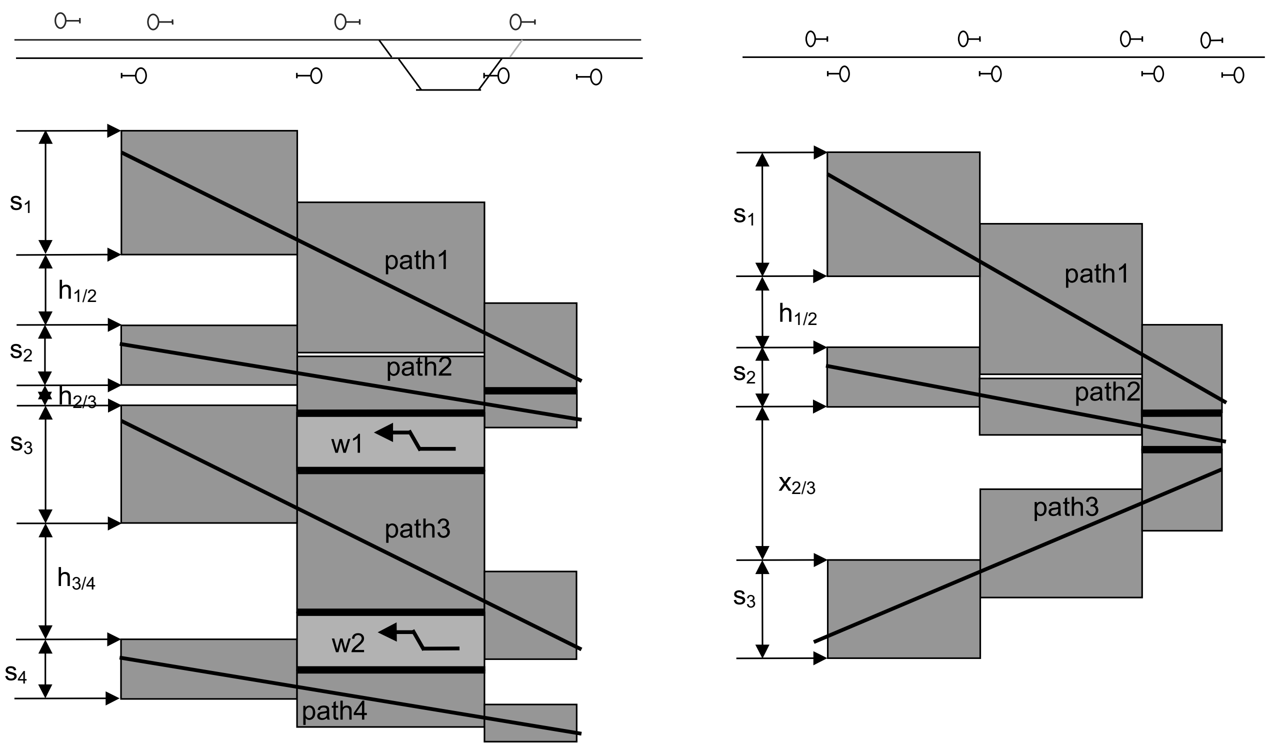


Fig. 3 - Time shares within a timetable



$h_{1/2}$, $h_{2/3}$, $h_{3/4}$ are heterogeneity times

Fig. 4 - Time shares after compression

3.5.1 - Journey time

Typical journey time

The typical journey time shall consist of:

- a basic time without time supplements, which is calculated according to line and rolling stock characteristics;

- a time supplement which is assigned depending on the type of route (route for passenger or freight trains) and which is used to compensate for smaller-scale irregularities in the course of train travel (smaller-scale delays, exceeding of stopping times, etc.). These time supplements are also used to cover the usual speed reductions due to maintenance work. The value of this time supplement is about 5% of journey time.

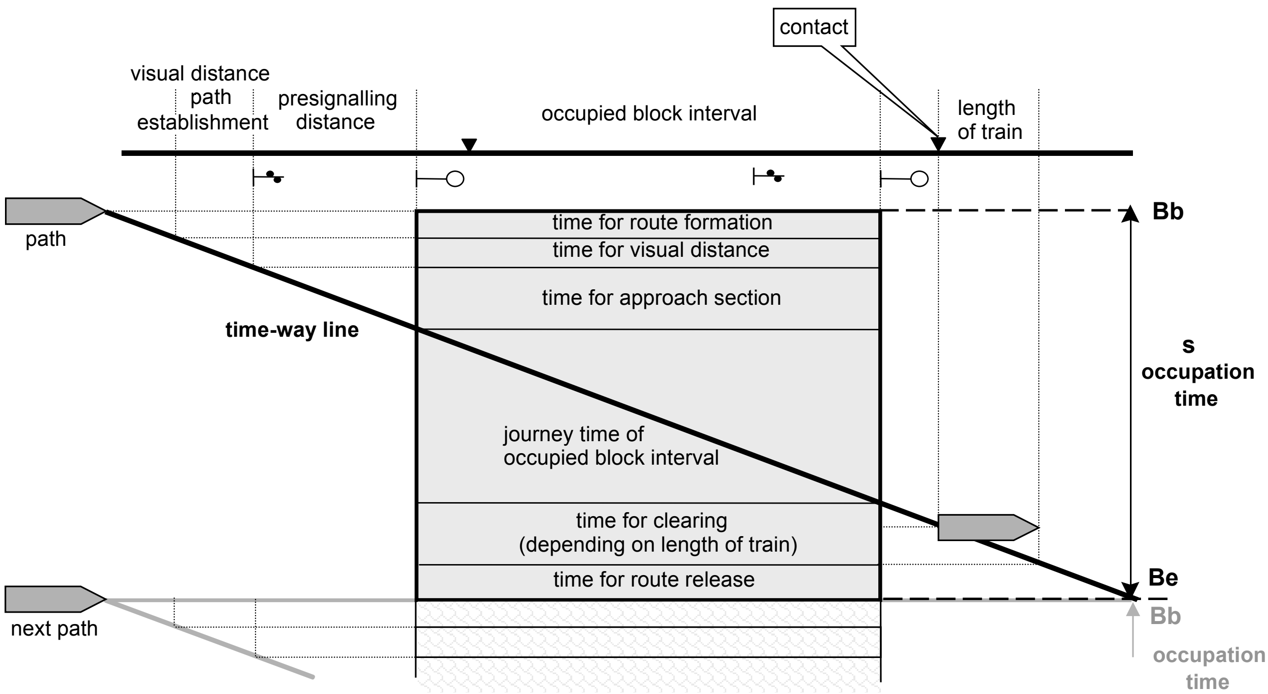
Timetable journey time (t_{tt})

Timetable journey time shall be the sum of:

- the typical journey time;
- additional times, which result from market requirements (e.g. prolonged travelling times of night trains, synchronisation times out of regular-interval timetables);
- additional times which result from timetable construction constraints.

3.5.2 - Occupation time (s) in a single block section

According to the timetable journey time, the route for each path shall be calculated in the block sections. The start time for creating depends on block systems, signalling systems and safety technology. The block section is occupied by a path as long as the point behind the block section, which must be cleared for safety reasons, is passed by the end of the train and the route is cleared.



s	: occupation time of line section for each train path		s = Be - Bb
Bb	: beginning of occupation		
Be	: end of occupation		

Fig. 5 - Elementary occupation time (schematic)

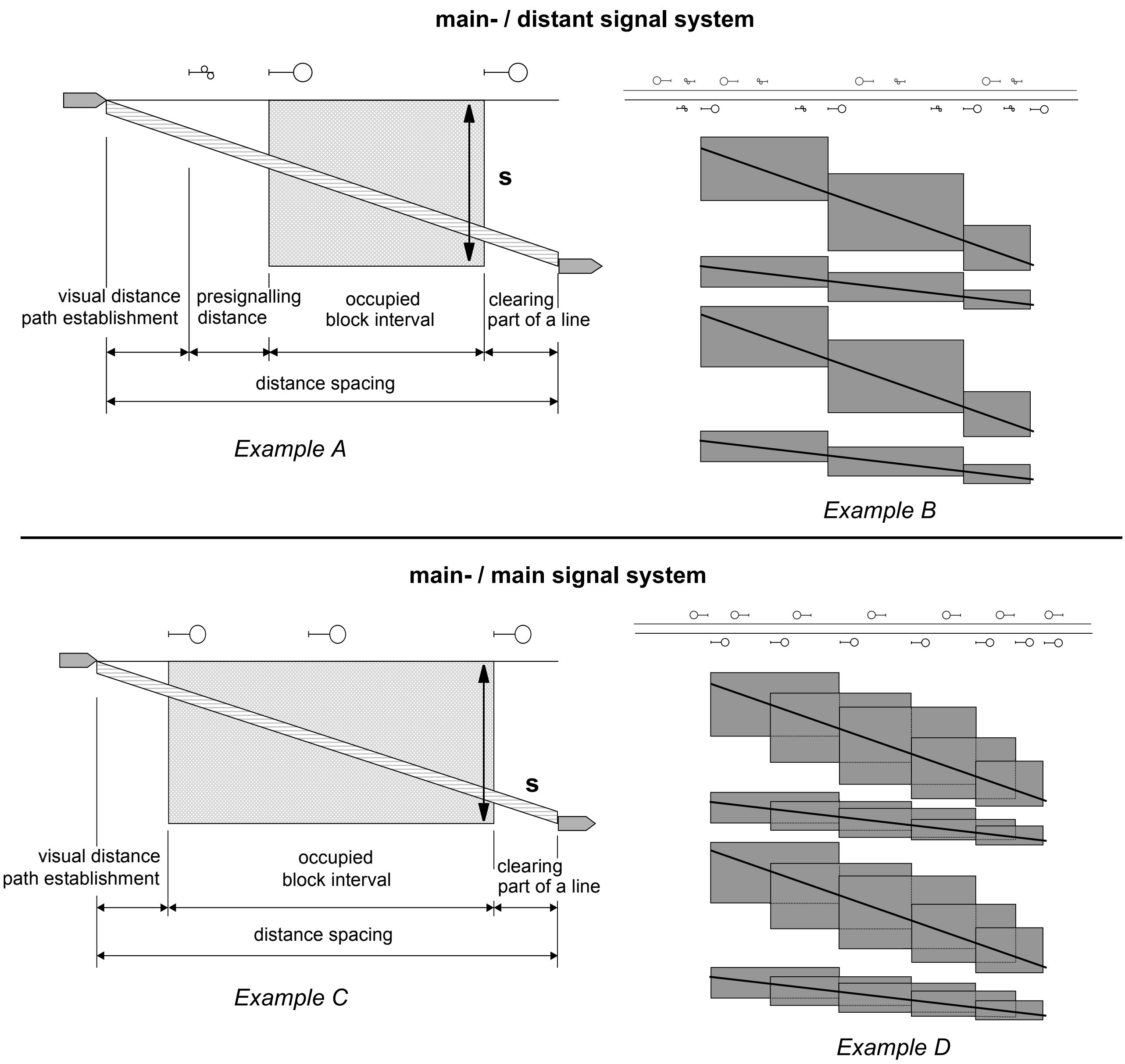


Fig. 6 - Examples of occupation times in different signalling systems

The compression-based methodology does not solely incorporate timetabled train paths on the whole line section of the domain or on a part of the line section. Any occupation times, even when not related to the train paths, must be incorporated. Those times, which are included into the occupation time, are called indirect occupation times.

Indirect occupation times are those times during which one track in a node - in few cases within a line section - is occupied and not available for further train paths. These indirect occupation times can either be connected (stabling of additional wagons, locomotive changeovers, etc.) or can be without any connection (crossing train from other line sections, shunting movements at stations with no specific marshalling tracks, etc.) to timetable train paths.

3.5.3 - Buffer time (b)

Buffer times shall be times that are inserted between train paths in addition to the minimum interval between trains that arises depending on the signal systems. They serve to reduce transfer of delays from one train to the next. They shall be introduced either between each path, or globally, provided that their total amount is sufficient.

3.5.4 - Crossing buffer time (c)

On single track lines, buffer times can be used as crossing buffer time placed between two train paths running in alternate directions. As for buffer times, they can be introduced between each crossing path or globally, provided that their total amount is sufficient.

3.5.5 - Heterogeneity time (h)

Heterogeneity time shall be related to different speeds between trains. It is the minimum time between two train paths as a result of different speeds of these train paths and the block system.

3.5.6 - Crossing time (x)

Crossing time shall be the time between two train paths, one going in one direction, the second going in the other. The duration of the crossing time between two train paths (x) shall depend on their timetable transportation time (t_{tt}) and the relevant block section (z).

3.6 - Calculation method

This point describes the methodology for determining capacity consumption.

Capacity consumption

Capacity consumption shall be measured by infrastructure occupation in a defined time position, to which is added time supplements for timetable stabilisation and, where necessary, maintenance requirements.

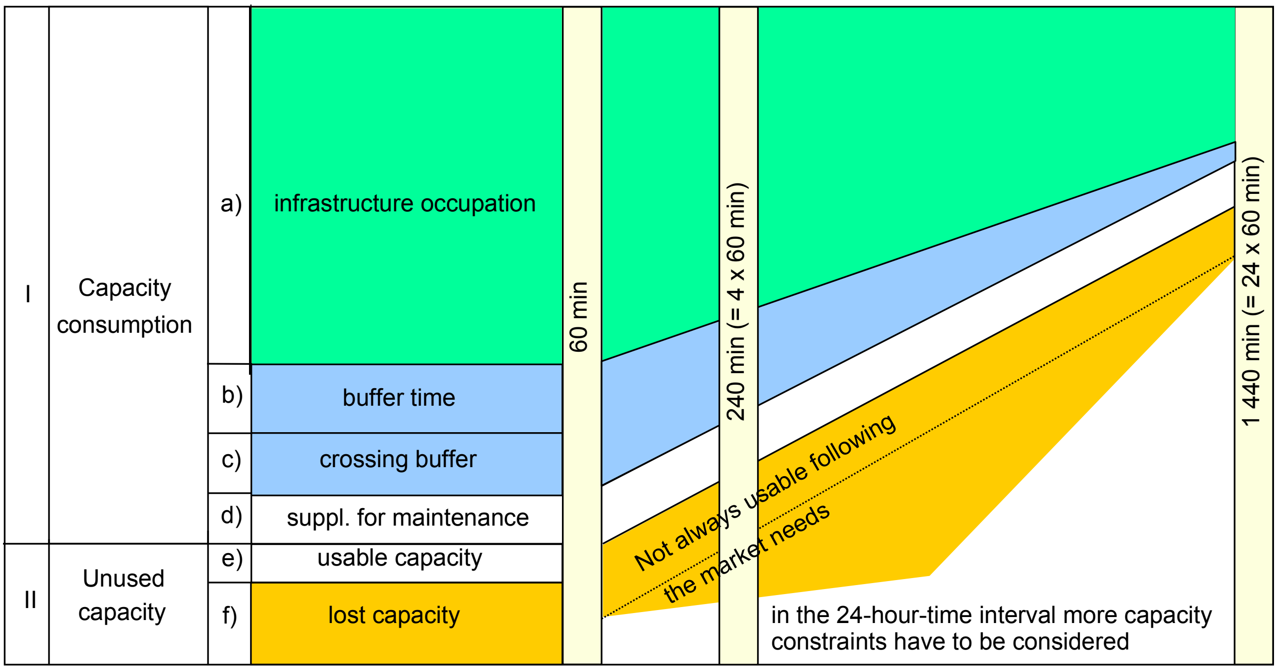


Fig. 7 - Determination of capacity consumption

The formula for determining capacity consumption shall be as follows:

$$k = A + B + C + D$$

k: total consumption time [min]

A: infrastructure occupation [min]

B: buffer time [min]

C: supplement for single-track lines [min]

D: supplements for maintenance [min]

$$K = k \cdot 100 / U$$

K: capacity consumption [%]

U: chosen time window [min]

1. *Infrastructure occupation*
   Infrastructure occupation shall be the result of the compression process and shall be measured at the beginning of the first block section within the line section.
   This procedure works for single track lines as well as for multiple-track lines (when evaluating each single track).

2. *Supplement for maintenance*
   The supplement for maintenance may either be part of infrastructure occupation or may be shown as an additional supplement.
   The supplement for maintenance shall be part of capacity consumption.

3. *Unused capacity* (see fig. 7 - page 16)
   The difference between capacity consumption and chosen time-window shall be called "unused capacity". The amount of unused capacity is determined by the possibly "usable capacity" and "lost capacity".

4. *Usable capacity*
   "Usable capacity" shall exist if unused capacity can possibly be used for additional train paths, providing they meet the customer requirements (typical characteristics of the paths) for the area considered.
   The additional train paths shall be incorporated into the original timetable before compression. Afterwards, a new analysis of capacity consumption is necessary.

5. *Lost capacity*
   Ultimately, there will come a time when no further train paths can be added.
   This time shall be called "lost capacity".

3.7 - Congested infrastructure

If the determined infrastructure occupation figure is higher than a typical value, infrastructure shall be declared congested.

If this typical value is not reached, a leftover capacity shall then become available. The way in which this leftover capacity can be used for additional train paths shall be determined for each single case on the basis of market needs.

Bottleneck and/or congested infrastructure

An infrastructure shall be congested, if:

- infrastructure occupation > "typical value"
  (no additional route can be inserted without changes to the infrastructure or shifting of routes)

  or

- shifting of routes is so extensive that market requirements can no longer be met.

4 - Application

4.1 - Application process

Capacity consumption shall be analysed in several steps:

1. In the first step, compression of timetabled train paths must be carried out.

2. The value of the infrastructure occupation [% of time-window] as a result of this compression must be compared with certain typical value(s) corresponding to the type of track.

3. If the infrastructure occupation [%] is higher than or equal to this certain typical value, the analysed line section shall then be called congested infrastructure and no more additional train paths may be added to the timetable.

4. If infrastructure occupation [%] is lower than this certain typical value, the capacity analysis must then be developed further.

5. Attempts must be made to incorporate further additional train paths of the type corresponding to the particular area into the timetable concerned.

6. If this incorporation is not possible, the leftover capacity is lost capacity which cannot be used any longer.

7. If this incorporation is possible a certain part of the leftover capacity shall then become usable capacity. In this case further analysis must be carried out, beginning with the compression of the timetabled train paths including the additional train paths.

8. This procedure shall be repeated until either the infrastructure occupation reaches congestion level (see paragraph 3) or no more train paths can be incorporated into the timetable (see paragraph 6).

On the condition that the buffer times are sufficient for timetable stability, the capacity consumption can theoretically reach 100% of the time-window considered. If it is greater, the infrastructure is congested.

4.2 - Standard values

In order to assess the standard value of the infrastructure occupation time, the compression methodology was applied on about 3 000 km of lines on several European networks.

A first result of these applications is that the limiting factor of the occupation time does not derive from the difficulty of drawing new paths, but from the level of buffer time (stability requirements).

In other words, possibilities for drawing additional paths always exist, but one must stop at a certain level due to stability requirements specific to the timetable.

Having said that, the remaining question is the value of infrastructure occupation time (% of time window) that must not be exceeded.

It then appears impossible to assess standard precise values.

As a matter of fact, several parameters must be taken into account, including the following:

- infrastructure reliability,
- rolling stock reliability,
- the interdependency of the line section with other line sections (that is if frequent delays to trains coming from converging line sections should occur, and vice versa if frequent delays to trains switched onto diverging line sections are prohibited),
- the level of quality required by RUs - for example, one can imagine that in case of delays, train cancellation is agreed, in which case the buffer time can be reduced,
- the margin on journey time,
- the number of trains per hour,
- the length of the line section and the possibility to organise overtaking or crossings within it.

For this reason, only the status of "recommended value" can be given as a guideline, which must be further developed through complementary analysis taking into account the above-mentioned criteria. These complementary analyses require high level expertise.

On the basis of current practices among European IMs, the following guideline figures can be proposed:

Type of line	Peak hour	Daily period	Comment
Dedicated suburban passenger traffic	85%	70%	The possibility to cancel some services in case of delays allows for high levels of capacity utilisation.
Dedicated high-speed line	75%	60%	
Mixed-traffic lines	75%	60%	Can be higher when number of trains is low (smaller than 5 per hour) with strong heterogeneity.

4.3 - Length of line section

The choice of line section on which the compression methodology will be applied is a key issue. In fact, the results may differ sharply depending on whether a long or a short line section is chosen.

The following recommendations are made:

- use the compression mechanism on short line sections,
- apply the enrichment process (that is the process to try and draw new paths) on longer line sections.

Compression of short line sections:

Ideally, the line section used for compression should be reduced to the line section between two neighbouring stations (without overtaking or crossing possibilities).

These overtaking or crossing possibilities are those that are commonly used during operation, for example:

- sidings used for overtaking freight trains,
- stations in which a regional passenger train can be overtaken in case of delay by a fast passenger train.

It must be ensured that the planned timetable makes a sufficient number of these facilities available for use.

Enrichment process on longer line section:

In this situation, account must be taken of the type of traffic considered.

In the case of international long-distance freight train paths, it is recommended to take a line section between two large freight nodes where a significant number of freight trains terminate or originate. This means a freight path ending at the end of a line section will find a connecting path on the following line section.

If it is only a matter of regional passengers services, then the line section will be much shorter (length of service area).

List of abbreviations

A	infrastructure occupation
b	buffer time
B	buffer time [min]
Bb	beginning of occupation
Be	end of occupation
c	crossing buffer time
C	supplement for single-track lines
D	supplements for maintenance
h	heterogeneity time
H	sum of heterogeneity times at measuring point
IM	Infrastructure Manager
k	total consumption time
K	capacity consumption
L	lost capacity
min	minute
N	unused capacity
P	usable capacity
RU	Railway Undertaking
s	occupation time
S	sum of occupation times at measuring point
t_{tt}	timetable running time
U	chosen time window
w	indirect consumption time
x	crossing time
X	sum of crossing times at measuring point
z	relevant block section

Warning

Printed by the International Union of Railways (UIC)
16, rue Jean Rey 75015 Paris - France, September 2004
Dépôt Légal September 2004

ISBN 2-7461-0800-3 (French version)
ISBN 2-7461-0801-1 (German version)
ISBN 2-7461-0802-X (English version)