Past, Present and Future of Freight Railway Shipment Control Systems

Railway Applications Section
September 2020 Webinar

Carl Van Dyke
TransNetOpt
carl@cvdzone.com
Our focus today is on the systems that are often called “car scheduling” or “trip planning” systems, and their associated impacts on yards, lines, resources, and customers.

The scope of these control systems includes:

- The instructions on how shipments should be routed (blocking)
- The overall mission, routing, workplan and timing of each train
- The itinerary or trip plan for each shipment
- The resultant workloads for yards and trains, and requirements for crews and locomotives

**Agenda**

1) Evolution of existing systems
2) Core logic used by existing systems
3) Fit between existing systems and current railway business mix
4) Potential future directions
5) Examples of potential new capabilities
Information Technology and Freight Railways
How did we get to where we are?

- Information Technology Applications have evolved over time
- Systems are all basically “stove pipe” systems with a batch focused legacy
- Each sub-system was independently designed to address a specific issue

▪ Each step along the evolutionary path represents a new maturity level for railway IT systems, and an increase in quality control and customer focus
- All major railways in North America are at level 5 in the above
- Large railways outside of North America exist across the spectrum from level 2 to 5
  ▪ Local services are often not well represented in systems
  ▪ Green Cargo (Sweden) is the only railroad known to have moved beyond level 5 in a deep and consistent manner

1960’s → 2000’s
Maturity Level 1: Over-the-Road Execution Support Systems

- No focus on customer’s shipments or railcar asset velocity
- No visibility to railcar locations/status
- No control of railcar routing
- No workload forecasts
Maturity Level 2: Point-to-Point Consisting & Yard Control

- Local Yard Inventory System
- Locally Maintained Marshalling System (Primarily for Hump Controller)
- Advanced consists: railcars on trains including railcar destinations
- Locally Maintained Marshalling System (Primarily for Hump Controller)

Customer perspective:
- No central routing control
- Very limited visibility to railcar location/status
- No transit time prediction & monitoring capability
- No workload forecasts beyond timeframe of advanced consists
Maturity Level 3: Electronic Waybills with Railcar Movement Database

- Provides location visibility to customers
- No system level routing control: customer still cannot receive predictions of railcar arrivals
- Still no ability to forecast future workloads

- Permits elimination of paper railcar movement instructions
- Requires disciplined data collection system
- Can be tied to advanced consist and yard inventory systems
Maturity Level 4: System Wide Central Marshalling System

- Ensures consistent routing of all railcars
- Eliminates need for routing clerks at yards (still need some for movement reporting)
- Supports network level planning and control
- Customer still cannot receive predictions of railcar arrivals
- Still no ability to forecast future workloads

This level of sophistication was first achieved by Southern Pacific through its TOP system in the early 1970’s – many other railways were using TOP or variants of TOP in the 1980’s and 1990’s (e.g., MP, CN, BN)
Missouri Pacific was the first to implement these capabilities using SP’s TOP as the foundation in its car scheduling system in the late 1970’s/early 1980’s.

Funded in part by the Federal Railroad Administration, this system became the design standard for all other large N.A. railways, and is still used today (with original code!) by the Union Pacific.

Trip planning was not implemented at all large N.A. railways until the mid to late 1990’s.
Shipment Schedules Are Key to Asset, Operational and Customer Management

The car or shipment schedule is the heart and soul of a “car scheduling” system

- Itineraries can be viewed two ways:
  - A series of train movements (shipment legs)
  
    ![Diagram of train movements](image)

  - A series of yard connections (connection events)

    ![Diagram of yard connections](image)

- Organizing the data by train provides train size projections
- Organizing the data by connection provides yard workload projections & marshalling/train make-up instructions
- Monitoring against shipment schedule focuses on customer service and asset velocity
  - Shipment schedules are also needed to support booking or capacity management systems

- Total train consist views spanning loaded & empty movements may be more appropriate for unit trains
- Intermodal requires separate, but linked views for railcars and trailers/containers
Core Freight Railway Planning and Control System Architecture

While great for carload traffic, this structure contains incomplete planning and execution support for unit trains, intermodal, large block bulk traffic, and in some cases local operations.
The Core Railway Planning and Control Systems in a Broader Context
Over time the railway legacy systems have grown in breadth and complexity, such that they can be used to manage most aspects of the railway

- Train Calling & Dispatching Systems
- Locomotive Management Systems
- Crew Management Systems
- Empty Railcar Distribution Systems
- Waybill Systems
- Revenue Management
- Car Hire / Interline Settlement
- Yard Inventory Projection & Workload Management
- Dated Train Databases with Train-Blocks
- Movement Reporting Systems (incl. AEI)
- Train Consist Projection Systems
- Operating Plan Design
- Shipment Scheduling & Monitoring Systems
- Train Profile Databases with Train-Blocks
- Blocking & Classification Databases
- Railcar Movement Tracking Databases w/ Shipment Schedules
- Reference Files
- Equipment Maintenance
- Line-of-Road Maintenance Management
- Marketing Support & Forecasting
- Handheld/Remote Device Mgmt
- Intermodal Systems
- Corporate Financial Systems
- Shipment Costing Systems
- Interline Communications
- HR Systems
- T & E Payroll
- Equipment Maintenance
- Line-of-Road Maintenance Management
- Marketing Support & Forecasting
- Handheld/Remote Device Mgmt
- Intermodal Systems
- Corporate Financial Systems
- Shipment Costing Systems
- Interline Communications
- HR Systems
- T & E Payroll

© Carl Van Dyke/Oliver Wyman
2. Core logic used by existing systems
Traditional car scheduling process
The generation of car schedules (trip plans) based on the operating plan is a well-established process that is used by major railroads in North America, Europe, and elsewhere

An iterative process is followed:

Select block
At current shipment location, given shipment destination and other characteristics, which block should unit take from location?

Select train
Based on block assignment, what currently planned trains are available to carry block, and based on business rules, which is the right choice?

Advance shipment
Advance shipment to location where selected train sets off block

Is the shipment at destination?
YES
Trip plan complete
NO

• Keys to above process are:
  – Well-defined blocking plan that can be interpreted accurately by computer, where plan is followed by field
  – Well-defined train schedule, including rules for which blocks can (will) go on train, where plan is followed by the field
  – No imposition of capacity constraints allowing every shipment to be processed independently

• The train schedules can still be somewhat dynamically created/changed, as long as an up-to-date, complete, forward view of the plan is maintained in the computer system with a 7 to 14 day planning horizon

• The system can self-correct when cars take alternate routes by comparing plan to car movement reports and regenerating plan when they differ
Traditional Car Scheduling Process:
Train Selection for Car Schedules if Often Strictly Time Based

• If one knows the block a traffic record will use, then all of the trains that can move that block can be identified from the block-to-train assignments

• Most car scheduling systems select the next train that carries the block once some minimum processing time has elapsed

• Some railroads instead use an environment where all of the possibilities are examined, and the “best” choice is selected, which sometimes is not the first available train

• When cars are left behind because a train is full, this is “discovered” by the system rather than predicted or managed by the system and the car is “rolled” to the next train
The blocking problem

**Definition of a block**: a group of diverse railcars with various origins and destinations that are moved together as a group from one location to another.

![Diagram of blocking plan]

<table>
<thead>
<tr>
<th>Blocking Location</th>
<th>Block Destination</th>
<th>Traffic Destinations</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>C</td>
<td>B, C, D, E, F</td>
</tr>
<tr>
<td>C</td>
<td>D</td>
<td>D, E, F</td>
</tr>
<tr>
<td>D</td>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>
The Efficiency of a railroad’s carload network is determined in large part by the blocking plan

- Two key blocking decision variables:
  - What blocks should be made at each location
  - What traffic should be placed in each block
- This determines routing of each shipment, number of handlings, yard workloads, and ultimately the transit times and reliability of shipments

- When combined with trains, the itinerary or trip plan for the shipment is determined

<table>
<thead>
<tr>
<th>Activity</th>
<th>Location</th>
<th>Time Impact</th>
<th>Driver</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shipper Release at A</td>
<td>A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Pick-up car from shipper</td>
<td>A</td>
<td>Train Schedule</td>
<td></td>
</tr>
<tr>
<td>Process car into block to B</td>
<td>A</td>
<td>Yard Processing</td>
<td></td>
</tr>
<tr>
<td>Wait for train to B to depart</td>
<td>A</td>
<td>Train Frequency</td>
<td></td>
</tr>
<tr>
<td>Arrive at B</td>
<td>B</td>
<td>Train Schedule</td>
<td></td>
</tr>
<tr>
<td>Process car into block to C</td>
<td>B</td>
<td>Yard Processing</td>
<td></td>
</tr>
<tr>
<td>Wait for train to C to depart</td>
<td>B</td>
<td>Train Frequency</td>
<td></td>
</tr>
<tr>
<td>Arrive at C</td>
<td>C</td>
<td>Train Schedule</td>
<td></td>
</tr>
<tr>
<td>Process car into block to D</td>
<td>C</td>
<td>Yard Processing</td>
<td></td>
</tr>
<tr>
<td>Wait for train to D to depart</td>
<td>C</td>
<td>Train Frequency</td>
<td></td>
</tr>
<tr>
<td>Arrive at D</td>
<td>D</td>
<td>Train Schedule</td>
<td></td>
</tr>
<tr>
<td>Process car for delivery</td>
<td>D</td>
<td>Yard Processing</td>
<td></td>
</tr>
<tr>
<td>Wait for delivery time to customer</td>
<td>D</td>
<td>Train Frequency</td>
<td></td>
</tr>
<tr>
<td>Deliver to customer</td>
<td>D</td>
<td>Train Schedule</td>
<td></td>
</tr>
</tbody>
</table>
Reliability and transit times for carload is driven by the success or failure of yard connections

- Railcars are usually handled at a minimum of three yards, and often even more
- In N.A. railcars lose about one day of transit time for every yard visit
- At many railways the average dwell time for hump yards can exceed 24 hours
  - (One train a day departure rate, with 12 hour minimum connection, and random arrivals, yields an average dwell time of 24 hours)
- In N.A. connection rates (right block, right train, right date) can be as low as 70%

Transit time across a railroad is more dependent on yard and local operations than on road train operations.

Majority of cycle time in yards – not in trains or at customers

Reported Dwell of Merchandise Cars

- Customer 18%
- Train 18%
- Yard 64%
Blocking rules: Traditional look-up process is based on rule ordering. Rules are manually maintained by location to determine the block assignments.

If last rule reached with no match, either provide default block assignment (root record) or declare traffic stranded and generate alert.
Algorithmic blocking is an alternative to traditional class tables: The blocking plan forms its own network of nodes and arcs

- Traditionally, cars are assigned to blocks based on a set of rules tables
- However, by forming a network where the arcs are blocks, one can instead employ a weighted shortest path to find the “best” block sequences
  - Still need rules and cost weightings to capture the “mission” of each block and determine its eligibility to carry a specific shipment

Do not confuse blocks with trains:
- Blocks have no departure or arrival times
- Blocks have no day-of-week frequency

First implemented at Norfolk Southern, now also used by Canadian Pacific and CSX
How the Algorithm Routes Traffic

First, which blocks can the traffic take?

**Network Restrictions** and **Rules** on yard-blocks determine which blocks are allowed for a given traffic movement. For example, if high multi-level cars are being sequenced from O to D, and yard-block CD1 has a clearance restriction, that yard-block is ruled out “//”.

Second, which block sequence is the best choice?

Each block has a **Cost Value** in terms of distance (which can include yard costs). The best sequence has the least cost. The best sequence for the multi-level traffic is OA1, AD1.

* Network restrictions can be inherited from the routing, which might be an attribute of the block, determined by an algorithm, or come from the associated trains.
3. Fit between existing systems and current railway business mix
Today’s Railroads Typically Have Four to Five Primary Lines of Business

- Each line of business has a different operating model – and thus any comprehensive “car scheduling” system must meet the unique needs of each within a unified support framework

![Share of Revenue by Line of Business - 2008](image)

Source: 2008 annual reports & investor fact books produced by each railroad

* “Coal” may include coke and iron ore in some cases; “Agricultural” may include some finished or consumer products; “Industrial” may include some bulk products

Note: Revenue tends to overstate some lines of business in terms of cars handled, and understate others. This difference can be up to +/- 30%; that is the market share numbers for a product such as coal could be up to 30% higher when expressed in carloads, and for manifest up to 30% lower.
Implications of Business Mix: Operating strategies and resulting systems / business process requirements vary significantly by line of business

<table>
<thead>
<tr>
<th>Line of Business</th>
<th>Type of Operation</th>
<th>Fundamental Planning Unit</th>
<th>Fit with traditional carload model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermodal</td>
<td>Expedited / premium / international</td>
<td>Container / trailer</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Dedicated trains – multi-block</td>
<td>Railcar / vehicle</td>
<td>Strong</td>
</tr>
<tr>
<td>Automotive</td>
<td>Solid empty trains</td>
<td>Railcar / train</td>
<td>Weak</td>
</tr>
<tr>
<td></td>
<td>Empty gathering process</td>
<td>Railcar</td>
<td>Moderate</td>
</tr>
<tr>
<td>Grain</td>
<td>Shuttle train</td>
<td>Train set</td>
<td>Weak</td>
</tr>
<tr>
<td></td>
<td>Ad hoc unit</td>
<td>Train / car lot</td>
<td>Weak</td>
</tr>
<tr>
<td></td>
<td>Solid train</td>
<td>Block / car lot</td>
<td>Weak</td>
</tr>
<tr>
<td></td>
<td>Local distribution &amp; gathering</td>
<td>Railcar / car lot</td>
<td>Moderate</td>
</tr>
<tr>
<td>Coal</td>
<td>Shuttle / pipeline trains</td>
<td>Train set</td>
<td>Weak</td>
</tr>
<tr>
<td></td>
<td>Ad hoc unit</td>
<td>Train / car lot</td>
<td>Weak</td>
</tr>
<tr>
<td>Manifest &amp; Bulk</td>
<td>Carload road trains (including some grain &amp; auto)</td>
<td>Railcar</td>
<td>Strong</td>
</tr>
<tr>
<td></td>
<td>Local gathering and distribution</td>
<td>Railcar</td>
<td>Strong</td>
</tr>
<tr>
<td></td>
<td>Ad Hoc Unit (steel, some rock, potash, etc.)</td>
<td>Train / car lot</td>
<td>Weak</td>
</tr>
<tr>
<td></td>
<td>Shuttle trains (ethanol, rock, crude)</td>
<td>Train set</td>
<td>Weak</td>
</tr>
</tbody>
</table>
Is it time to rethink core system functionality?
With the rise of unit trains and intermodal, most core transportation systems are capable of truly meeting the needs of less than 50% of a railway’s traffic

• When originally created, these “car scheduling” systems were focused on the general carload business. As a result they did little to support either unit trains or intermodal, which now represents between 30% and 70% of a railway’s traffic
  – Example: The classification and trip planning functions of most systems do not apply to unit train traffic. Result is that this traffic must be handled manually, and there is limited support for managing the cycling of train sets. Customer ETAs for unit trains are often either not provided or left to the customer to estimate based on reported train locations
  – Example: Many of these systems are also incomplete with respect to local pick-up and delivery services, requiring both guesswork and manual intervention to plan and manage these activities
  – Numerous other examples exist

• The legacy of the blocking and trip planning logic is that of a carload railroad – even with the development of algorithmic blocking, this legacy continues to hold undue influence over the design of new solutions
Evolving Focus = Evolving System Requirements
Effectively meeting these evolving objectives requires a suite of planning and control system capabilities that go beyond the current system capabilities.

Elements of the Operating Plan/Car Scheduling Control System

- Classification strategies
- Train routing and block-to-train assignments
- Train schedules or scheduling guidelines
- Empty railcar distribution policy
- Yard performance standards
- Unit train planning
- Gathering network/solid train management
- Capacity Management

- Local pickup and delivery standards
- Intermodal/automotive operations management
- Origin-to-destination trip time standards
- Resource management issues:
  - Locomotive assignments
  - Assignments of yard, local, and road crews
  - Support personnel scheduling

Understanding the evolving needs of major railways represents a critical issue.
4. Potential future directions
Future Directions: Examples (1 of 3)
There are many areas that need exploration when creating next generation freight railway management systems

- **More sophisticated car scheduling logic**
  - Algorithmic blocking and time-space oriented logic

- **Capacity management and yield management**
  - Tracking of capacities, with either simple rolling to later trains, or alternate route generation using time-space type algorithms (network formed from train-blocks)

- **Intermodal shipment and ramp management systems**
  - Scheduling of both railcars and intermodal units
  - Stronger understanding of packer vs. rail switching, and use of rubber moves

- **Unit train equipment cycle management**
  - Tracking and scheduling the use of entire train consists across multiple loaded and empty moves
  - Moving away from car scheduling logic to train-based logic to support customer ETAs
  - Assigning/scheduling cars to trains not based on blocking/connection logic, but on the assignment of unit train waybill sets to date specific instances of trains
Future Directions: Examples (2 of 3)

There are many areas that need exploration when creating next generation freight railway management systems

- **Tactical decision support systems**
  - Evaluating the impact of tactical changes to trains and blocking, determining which are the best choices – or suggesting tactical changes
  - Support for “dynamic by-pass blocking” – for example, upon receipt of a large group of cars going to the same location, seek ways to move the cars direct to destination on existing trains using blockswaps if necessary (eliminating classifications)
    - Examples – interchange receipt of 20 potash cars with the same destination, or release of 20 grain cars from an elevator for the same destination

- **Automotive empty railcar management**
  - Connecting the scheduling of empty multi-levels to the TTX re-load program, and incorporating the associated accumulation processes

- **Tactical train management**
  - Better manage train schedules to make them more realistic, thus driving better crew, locomotive and yard management decisions, better customer ETAs
  - Take real-time line segment and yard capacity into account in setting train schedules
  - Take network capacity into account when launching unit trains
Future Directions: Examples (3 of 3)
There are many areas that need exploration when creating next generation freight railway management systems

▪ **Gathering and distribution network support**
  – Example: scheduling support to identify locations where grain cars should be sent based on origin, to be held until solid trains can be built
  – Would need rules on when to put cars into general manifest service, and when multi-block solid trains can be built

▪ **Empty equipment management and customer order management**
  – What is the next step beyond the current distribution systems?
  – Can the process be more tightly bound to customer orders within the scheduling system?
  – How should intermodal, bulk, grain, and automotive cars fit into this process?

▪ **Product catalog concepts**
  – Product catalogs show the service offer for carload
  – What can be done for other lines of business?
  – How does one tie this to pricing and contribution during planning and execution?
5. Examples of potential new capabilities
Example: Unit train trip planning
The carload trip planning model does not fit well with unit trains, resulting in most railroads not relying on these systems to support unit operations & often leaving customers in the dark with respect to the status of their shipments

• Basic trip planning logic does not bind specific shipments to a specific train
  – The cars are assigned to a block, which provides a set of candidate trains
  – The specific train used is then selected based on the timing of the traffic and the cutoff values
• However, a unit train exists for the sole purpose of moving a specific set of shipments, and no other traffic
  – Conventional trip planning logic cannot guarantee that the right traffic will be assigned to each date specific unit train
  – To succeed with conventional trip planning one has to set the release times and connection standards in a way to try to ensure that the right traffic goes on each unit train
  – Such strategies may not prevent other traffic from also ending up on the unit train, or some unit train traffic failing to be assigned to the right unit train run (or even showing up in the carload traffic projections)
• Trip planning for unit trains could likely be significantly improved and simplified by using a mechanism that ties the shipments to the appropriate, date specific train
  – For example, putting the train set and trip number on the waybill and the train might be a way to glue the traffic and train together and guarantee the process works accurately
  – Block attributes that clearly tie the block to the right unit traffic (though not date specific) are also important – or one could consider eliminating the classic blocking process for unit trains
Example: Setting the date specific timing/scheduling for a unit train

• Currently unit train schedules are largely set by fixing the origin departure date/time, and allowing the remaining time values to be simple offsets based on a set of standardized running time

• Except on a limited basis, there are no network capacity checks when the trains are scheduled

• This could result in the unit train attempting to traverse network segments at times when these segments are congested

• Using a capacity model one could conceive of a process where a capacity check was done when the unit train is proposed, and the timing of the train was adjusted so that there would be capacity for the train on each network segment the train was to traverse
  – These timing adjustments would include holding of the train at origin, as well as potentially holding the train at intermediate points when the train might otherwise enter a network segment when it was at or above capacity

• A core element of this process is knowing what other trains are already scheduled to use each network segment
  – Our assumption would generally be that the scheduled trains would be assigned to the network first, and that service design would set the timing of these trains to ensure they were all feasible (and the schedules for these trains would be respected by operations)
  – The ad hoc trains would then be scheduled as they became identified, fitting into the remaining capacity on each segment
  – The trains already assigned to each segment would need to be tracked by the system
Example: Capacity constraints
The basic blocking/trip planning model is uncapacitated and timing insensitive – yet we live in a capacity constrained world

• Capacities come in many flavors:
  – Specific train size constraints
  – Facility constraints (classification yards, intermodal ramps, fueling facilities, mines, ports, etc.)
  – Over-the-road limits on trains operated
  – Equipment availability (and linkage to velocity)

• One concept is to use “dynamic car scheduling” to address at least the train size constraints and the facility constraints
  – Unfortunately, railroad operations may be too unreliable, and advance order information to sketchy, to produce reliable views of volumes more than a few hours in advance

• Another is to actively manage what traffic goes on a specific train in the trip planning process

• Other capacity constraints need to be addressed through other strategies that may fall outside of the traditional blocking and car scheduling environment
  – Unit train scheduling taking facility, network, and equipment into account in determining when to run a train
  – Load acceptance concepts
    - Intermodal ramps
    - At mines, ports & industrial plants
    - Based on equipment availability of network fluidity
  – Bidding systems for allocating scarce grain equipment
Example – Capacity Reservations Systems and Yield Management

While it is straightforward to design such systems, creating a system that can actually work is another matter given the variability in day-to-day operations.

I believe Green Cargo in Sweden to be the only example of a railway running a fully reserved, capacity managed carload operation – they are greatly benefited by their ability to produce a carload product with a greater than 95% reliability record (using a zero tolerance measure for late deliveries).
Example: Dynamic scheduling
While an interesting concept, it is a high risk proposition in an unreliable operational environment, and would only apply to a limited set of traffic such as selected carload and intermodal movements

• Two fundamental ideas:
  1. Traffic routings (block & train sequences) can change by day of week and time of day in response to variations in the structure of the operating plan (primarily train schedules) – examples include:
   - Allow the operating plan to vary by time-of-day or day-of-week to better suit volume variations
   - Allow the scheduling of a car to react in (semi-) real time to changes in the train schedules or yard conditions; essentially “trip-plan repair” – could include “dynamic detours” if plan elements are unexpectedly eliminated (line closures, yard closures, train annulments, etc.)
  2. Traffic routings can change due to a lack of carrying capacity on specific trains or through specific yards on a date/time specific basis, and alternate routings might be more than just delaying to a later train, but instead might mean taking a different set of blocks & trains
   - This is usually considered in two parts: (1) given the priority (and perhaps revenue) of the car, which set of train-blocks should it take, and (2) how to adjust in semi-real time the set of trains/train-blocks available to move the traffic
   - Could even “generate” extra trains or annul trains based on capacity projections in some of the most automated visions for the system

• Not aware of any railway currently doing this operationally
Example: Train-based capacity management
General concept is to set up rules for how much traffic should be put on a train for each train block in a capacity constrained situation

- Reflect the relative importance of different blocks in trading off between multiple blocks picked up at the same station
- Protect capacity for pick-ups at downstream locations
- Capacities come in several flavors:
  - Physical Infrastructure related
    - Line-of-road maximum train size, train length
    - Length restrictions based on characteristics of locations where trains must do work
  - Asset based – pulling power of locomotives
  - Design based
    - Allocation of capacity among competing blocks on the same train
    - Size limits due to train performance and other considerations
- Train schedules should support detailed specification of train capacities by train, train-route, and block – with the option of deriving route-based limits from the infrastructure
- Take these capacities and allocation rules into account when deciding which traffic is placed on each train through the trip planning process
- This approach is followed at Green Cargo in Sweden
Questions?