INTERMODAL HUBS IN THE PHYSICAL INTERNET

Nadeepa Wickramage  
Clemson University

Bill Ferrell  
Clemson University

Abstract
Distribution of finished goods is currently an effective but inefficient process that consumes significant quantities of fossil fuel to move empty assets. This results in increased costs that are passed to the consumer and unnecessarily increased carbon emissions. The Physical Internet (PI) is focused on shared logistics that could prove to be an important element of next-generation logistics systems. The idea is to store and transport goods in anonymous standard-sized containers so that transportation and warehousing can be efficiently shared by many companies including competitors. If implemented, this idea has the potential of dramatically increasing efficiency thereby reducing fuel consumption and decreasing costs as well as emissions. This paper focuses on one aspect of the PI, intermodal hubs. More importantly, the key difference between the PI hub as imagined in this research and a transhipment facility or breakbulk terminal is that the control is decentralized; hence, this research explores decentralized control of a PI hub through experiments using scenarios and heuristics in an effort to gain some understanding how design and operations impact performance.

1 Introduction

Effectively moving freight from manufacturers to customers is a critical facet of our everyday life as well as the backbone for many of the world’s economies. Many companies, large and small, participate in this business and the quest for improved efficiency has spawned technological and strategic innovations through the years. Today, a significant new challenge is emerging due to a confluence of factors. Inefficiency in the current transportation system is colliding with disruptive trends - dramatic increase in demand to ship smaller packages, customers requiring ever-shortening lead times between placing an order and having it delivered, and more densely populated urban areas - in a way that is stressing our logistics systems into becoming what Montreuil [1] described as “unsustainable economically, environmentally, and socially.”

One vision for a new logistics system that begins to address these issues is the Physical Internet [1]. The Physical Internet is based on a vision of an interconnected
logistics system where goods are handled, stored, and transported across a common network that is shared by all users. It is enabled, thus, enabled by different companies – some of whom are direct competitor – sharing elements like warehouses and trailer space as well as agreeing to abide by certain protocols like using standard sized containers. Physical Internet research to date has explored a number of aspects related to this system including estimating the potential improvement associated with shared resources, determining the best sizes for modular containers, building physical prototypes of containers to test the feasibility for them to support demands of the Physical Internet, identifying business models that would motivate participation, and designing facilities and hubs at a conceptual level. The results are promising. For example, Meller, Ellis, and Loftis [2] used data from the consumer packaged goods industry and showed that if the Physical Internet were in place, trailer fullness can be increased more than 30% and cost per load reduced by more than 25%. Landschützer, Ehrentraut, and Jodin [3] designed and tested modular boxes that meet customer requirements for things like safety and cost as well as Physical Internet enabling characteristics like flexibility and interlocking stackability. This research focused on another operational aspect of the Physical Internet, design and operation of the intermodal hubs.

The work is based on a generic interpretation of the Physical Internet in which modular boxes are the smallest units shipped. For transport and handling, these are assembled into “unit loads” and the unit loads are assembled into “transportation containers.” Hence, it is possible that a single shipper can completely fill a transportation container with modular boxes that are to be delivered to a single customer. Alternately, a shipping container can contain modular boxes from many shippers and these modular boxes can be destined for a large number of customers in a highly dispersed geographic area. Regardless, the transportation containers move through the network by passing from one intermodal hub to the next. At each hub, one of two things happens to them. One is that the transportation containers are passed along to another hub intact. This would occur when the constituent modular boxes have the same destination or they are reasonably close, at least from the perspective of the hub. (e.g., modular boxes with destinations in southern California are likely considered “very close” when decisions are made at a hub near New York.) On the other hand, the transportation container might need to be broken down so the unit loads can be redirected. It is also possible that the unit loads must be disaggregated so that modular containers can be redirected or for last-mile delivery. Regardless, aggregating module containers into unit loads and unit loads into transportation containers is a function of the hubs. Clearly, there are a number of important operational decisions that must be made to efficiently handle inbound and outbound freight. Further, these decisions are closely linked to design issues like amount of storage and layout at a hub.

On the surface, this appears to strongly resemble a breakbulk terminal for less-than-truckload operations. Obviously, the flow of packages is similar but the critical key difference is the control. In LTL operations, a large fraction of packages has their routing between origin and destination predetermined; that is, they use centralized control. In our vision of the Physical Internet, control is distributed and each hub
makes dynamic decisions in real time. This difference dramatically changes the picture.

2 Related Literature

In an intermodal network, a shipment uses multiple modes of transportation in its journey from the origin to the destination in a seamless manner through the use of intermodal containers [4]. There are a number of papers that review the application of operations research models and methods in the field of intermodal transportation like [4], [5], and [6]. The conclusion is that much research has addressed intermodal transportation but much if not most is from the perspective of centralized control. Also, collaborative transportation planning for intermodal transportation has been studied where coordination of independent operators of different stages of intermodal transportation has been the focus [7]. The dynamic nature of the envisioned PI hub seems to suggest the approach here must be different from the current literature; scheduling and temporary storage problems (see [8], [9], and [10]). As before, this work explicitly or implicitly assumes centralized information which we currently think changes the fundamental nature of the research.

3 Research Problem

Intermodal hubs are an important part of the PI because they are the places where critical decisions are made and executed regarding all inbound freight. A key change in the PI is that transport from origin to destination (point-to-point) is replaced by distributed multi-segment intermodal transport [1]. Hence, hub operations must handle high volume throughput and make many decisions based on current information like the level of disaggregating and aggregating, and destinations of transportation containers leaving in the next time period. The facility design has an impact on the possible decisions; for example, quantity of storage impacts the trade-off between storing and sends freight to a hub that is not the destination. It is assumed that however, the work of others has definitely influenced how we frame and understand the problem and solution approach.

Indeed, cross-docking operations have similarities with operations at intermodal hubs. Cross-docking is used by many companies as a logistics strategy to transfer incoming shipments directly to outgoing vehicles without storing them in between. The literature addresses aspect like locating cross-docks and network, layout design, pickup and delivery vehicle routing, dock door assignment, inbound and outbound truck accurate knowledge of inbound and outbound transportation containers is available to the intermodal hub controller in real time. For inbound freight, this will include time of arrival and complete information about the modular boxes in each inbound transportation container such as dimensions, weight, destination, and promised delivery time. This information will be known to the intermodal hub controller no later than when the transportation container leaves the previous hub – at least half a day in practice. Some of the destinations for outbound transportation
containers will likely be known well in advance because the demand on routes, say
hubs near two major cities, is reasonably constant so capacity along those routes is
established months in advance. Other outbound lanes will be opened on an as-needed
basis which can only be determined as details of the inbound freight demand are
revealed which could be as short as half a day. The point here is that decisions at the
intermodal hubs are time sensitive and the time to make them can be rather short
considering the actions required like arranging for an extra outbound trailer and driver
or performing an unexpected large amount of disaggregating and aggregating.

On the surface, the PI appears to strongly resemble a breakbulk terminal for less-
than-truckload operations or the basic operation of a third-party logistics provider
(3PL). Obviously, the flow of packages is similar but the critical key difference is the
control. In LTL operations, a large fraction of packages have their routing between
origin and destination predetermined; that is, they use centralized control [11]. In our
vision of the Physical Internet, control is distributed and each hub makes dynamic
decisions in real time. This difference dramatically changes the picture.

3.1 Problem Statement

This research assumes that all transportation containers are the same size. As will be
seen, this is simply for convenience and poses no restriction. The research effort is
aimed at defining some of the key design and operational elements and constructing
simple models to investigate relationships and sensitivity of controllable parameters
and performance measures. These models are based on the assumption that the unit
load, not the module box (i.e., m-box), is the smallest unit that is handled and that
they are homogeneous (i.e., all modular boxes contained in the unit load have the
same destination hub and delivery deadline). This means that transportation
containers can be decomposed into unit loads and the unit loads can be consolidated
into transportation containers (or sent for last-mile delivery) at the hubs but individual
modular boxes are not handled. Note that this is indeed a strong assumption from the
perspective of a practical design if unit loads inbound to a hub can contain m-boxes
destined for different end-hubs because the entire handling process required to
disaggregate and re-aggregate is different and much more time consuming that that
represented in this model.

It is also assumed that the hub controller has knowledge of inbound freight some
time period in advance of arrival; that is, the controller knows the destination and
deadline for each unit load in inbound transportation containers and the time they will
arrive. We assume that at least some outbound capacity is prescheduled. For now,
we are approaching this problem by analysing different scenarios with models
predicated on penalizing shipments to destination hubs that are different from the
correct one and for late shipments. For example, a transportation container in which
all unit loads are destined for hub A is not penalized if it is shipped to A before the
deadline; however, shipping it elsewhere or late is penalized. The penalty can be
based on any number of things; we are currently using a combination of shipping to a
hub that is not the desired destination and lateness. Different scenarios are considered
that specify model details and dictate the type of decisions that must be made. The
simplest scenario is one in which the hub has no storage capability so the decisions are limited to breaking down and reforming transportation containers and if addition outbound capacity is required. As storage capacity is added to the hub, the decisions become more complicated because they must address not only what to store but for how long.

The objective is to explore a few operations and design characteristics of a PI hub. The approach to be taken uses operational heuristics on different scenarios that embody some of the unique operating characteristics of the PI, notably completely decentralized and dynamic decision making regarding on the contents of outbound transportation containers. We submit this simple approach is appropriate for this initial work that has the object of exploring operational features of a PI intermodal hub and the interaction of these on some parameters of the physical layout from a design perspective. This research is not an attempt at developing an optimal design or a methodology that will yield an operational design.

Inbound truck trailers arrive to the intermodal hub from location $n_I (\geq 1)$ and outbound trailers are destined for location $n_O (\geq 1)$. In this problem we limit the planning horizon to $n_T (\geq 1)$ time periods. If an inbound trailer from location $i \in I = \{1, \ldots, I\}$ carries $n$ unit loads destined for location $j \in O = \{1, \ldots, O\}$ in time period $t \in T = \{1, \ldots, T\}$, we denote it by $n_{ij}^t$. The capacity of all trailers is known at the beginning of the time period. Outbound trailer capacity destined for location $j \in O$ in time period $t \in T$ is denoted by $o_j^t$. The duration of a time period $t$ is has maximum $T_{\text{max}}$ minutes.

Unit loads from inbound trucks need to be moved to outbound trucks or storage. If a unit load destined for location $j \in O$ cannot be loaded into a truck at outbound location $j$, it might be temporarily storable at the intermodal hub. We assume that it costs $s$ monetary units to store a unit load for one time period. In addition, the intermodal hub has a maximum storage capacity of $S$ unit loads. Alternatively, if a unit load cannot be loaded on a trailer for the correct destination, it can be shipped to another destination $k \in O, k \neq j$. At $k$, the decision is made regarding the next segment for the unit load – maybe to the final destination and maybe to another hub based on outbound destinations and storage capacity. Shipping to an alternate destination costs $p_{jk}$ monetary units per unit load. We assume that in addition to scheduled outbound shipments, additional shipments can be arranged to a specific location (e.g., a shipment can be arranged from the intermodal hub directly to another intermodal hub) at a cost of $x$ per shipment regardless of destination location.

Intermodal hubs will undoubtedly have many different layouts but in this work we assume the simplest possible intermodal hub layout illustrated in Figure 1. This resembles the common “I” configuration cross-docking facility which is ideal when the number of doors is 150 doors or less [8]. It is also assumed that the intermodal hub has $n_I \geq n_I$ inbound dock doors and $n_M \geq n_O$ outbound dock doors; also, the distance between dock doors, $d_{lm}, (l \in L = \{1, \ldots, n_L\}, m \in M = \{1, \ldots, n_M\}$ is known. Unit load carrying vehicles, such as fork lifts, move the unit loads in the intermodal hub. It is also assumed that each vehicle can carry a maximum of $U (\geq 1)$ unit loads per trip and they travel at $s_f$ feet per minute when they are not loaded and at $s_b$ feet per minute when they are loaded. The time to pick a unit load from an inbound
truck trailer is $t_p$ and time to store a unit load in an outbound truck trailer or in storage is $t_s$.

Figure 1: A single-stage intermodal hub layout

4 Approach and Results

Since this research is at a rather early stage and our goal is to understand basic relationships, the current approach is very simple - use several seed heuristics on different scenarios and look for trends. There are two basic scenarios: 1) The outbound capacity is unlimited to predetermined destination hubs, and 2) The outbound capacity exceeds that inbound but is limited in amount of excess or destinations or both.

Commands to the material handling devices are determined based on several different approaches but all heuristics track a number of performance metrics including time and cost. Time is monitored because in many of the scenarios, as well as in the real world, the material handling devices are capable to moving multiple unit loads that are stacked and including time to manoeuvre the unit loads is included. At this point in the research, cost is a universal measure only meaningful on a relative basis because it accumulates the impact of time, penalties, and storage.

A simple example is an adaptation of a dual-command heuristic previously developed for unit load warehouses when unit loads are allowed to be stacked two high. All routes start at a depot with the material handling device empty. Commands
initially have the device pick one or two unit loads from an inbound transportation container and/or storage location. From there, commands are sent to pick or store until all inbound transportation containers are empty. The simplest version of this basic heuristic assigns the next action randomly. Obviously, a bit of intelligence can be easily added to improve performance.

The results below are some of the preliminary overall trends that we have found and that are directing the current research.

4.1 Unlimited outbound capacity

The base case scenario explores how material handling capacity influences intermodal hub performance when outbound truck trailer capacity exceeds the outbound demand so the only performance impact is due to material handling capacity that is represented by the number of vehicles available. The performance measures are the total time, maximum number of units stored, the total trips made by the material handling devices, and the total operation time. Thirty different replications of this scenario were randomly generated and each heuristics was applied to each replication of each scenario. Typical results for the means of several performance measures are presented in Figure 2.

![Figure 2: Mean Results - Unlimited Outbound Capacity](image)

These are completely predictable. The maximum number of unit loads stored at the end of time horizon decreases linearly with the number of material handling vehicles. The cumulative time of trips by all vehicles reaches a plateau when the number of vehicles used is increased. When the cumulative time to move the unit loads to outbound truck trailers exceeds the time limit, unit loads not on outbound trailers are stored until the next period. An increased number of vehicles mean more unit loads can be moved to outbound trailers within the time limit. If all inbound unit loads can be moved to outbound trailers, the number stored is zero. Results like these are useful to confirm the heuristics are performing as expected. For some of the more
“sophisticated” heuristics that add details based on storage location and are multi-period, they might provide some insight for dynamic control of the facility because they solve the problem so quickly. For example, given details of unit loads in storage (i.e., destination and deadline), the inbound freight characteristics over a planning period, and the scheduled outbound capacity, these types of heuristics can help determine the number of material handling vehicles that would need to be in service for the desired state at the end of the horizon.

4.1.1 Other trends with unlimited outbound capacity

Since there is unlimited outbound capacity, the trends depicted here have analogues in the other parameters as well. Figures 3 shows the mean of the storage cost plus the penalty cost associated with sending a unit load to different hubs when the total storage limit is varied.

![Figure 3: Mean Storage & Penalty Cost](image)

Again, these results are from one of the heuristics but are typical of all. The mean performance is as expected – the cost is greatest when storage is most limited so unit loads must be shipped to destination hubs that are not the final destination for the unit loads and, therefore, incur a penalty. The standard deviation across the replications is a bit more interesting as illustrated in Figure 4.
Figure 4: Standard Deviation of Storage & Penalty Cost

While the mean storage plus penalty cost show a linear trend, there is larger standard deviation across the replications than might be expected. The overall trend is understandable. It is possible for storage to be increased to the point that all of the unit loads can either be sent to the correct destination hub or stored for all replications. When storage is limited, specifics of the scenario are important. The maximum cost replication minus the minimum cost replication shown in Figure 5 reinforces this observation. Figures 4 and 5 also show what could be an interesting pattern.

![Graph showing Standard Deviation of Storage & Penalty Cost](image)

Figure 5: Maximum minus Minimum Storage & Penalty Cost

There is little impact in the variability measures until some level of storage is reached. In these experiments, it was in the 130 range for all heuristics. At that point, there is a noticeable but small decrease. At around 170 units of storage, all heuristics could move all unit loads to the lowest cost options so both the mean and variance were small. This trend is being explored further because it could have an impact on

4.2 Limited outbound capacity

We now limit outbound capacity. Note that the average outbound capacity over time must be greater than or equal to the average inbound demand or inventory levels in the hub would grow without bound as in a simple queueing model with intensity greater than one. The experiments are based on varying the ratio of outbound capacity to inbound demand. Other parameters are varied as before like the number of unit load handling vehicles and storage capacity. Outbound capacity to each destination in each period is randomly selected between 0 and 50 and inbound demand is determined by the outbound/inbound capacity ratio. Storage and shipping to a destination that is not the final one for a unit load are allowed but each induces a cost. Finally, if inbound demand over a period exceeds outbound capacity, we
assume sufficient storage is available to store the excess. The remainder of the parameters remain unchanged from the unlimited capacity experiments.

Figure 6 shows results from the simplest case with unlimited storage – the maximum units stored over all replication as a function of outbound/inbound rate and number of material handling vehicles.

![Maximum Storage vs. Number of Vehicles](image)

**Figure 6: Maximum units stored, limited outbound capacity**

The data simple confirms intuition. As before, the number stored decreases as the material handling capacity increases. The required maximum storage capacity for a given number of vehicles is slightly smaller when the outbound/inbound capacity ratio is higher. When outbound/ inbound capacity ratio is higher, more units can be directly shipped instead of storing or reshipping. Therefore, for a given number of vehicles, less storage is required. The amount of decrease between different outbound/inbound levels decreases because of congestion as well as the costs associated with reshipping and the heuristics themselves. In addition, maximum required storage does not reach zero even with a large amount of capacity because outbound capacity might not always exceed outbound demand for a given horizon in the experiment. Also, reshipping the cost balance between storing and reshipping is considered in some of the heuristics. For all of these reasons, some storage is required even when resources to move unit loads is not a limitation.

### 4.2.1 A look at time and storage

From a design perspective, understanding the total time to handle the unit loads that arrive in a given time horizon is another aspect to consider. We have begun to explore this by fixing the material handling capacity at 8 trucks and looking at total time as a function of storage capacity for different inbound/outbound ratios. Results from one of the heuristics that is frequently seen is illustrated in the figures starting with 7. When the storage capacity is small, it is used quickly and in an obvious way for all ratios, namely, for unit loads with a destination that does not have a scheduled
outbound shipment and that is not sufficient large to dynamically schedule a outbound shipment there. As the storage capacity increases, the total time increases as the heuristics use the storage more often to minimize the total cost. At some point, there

![Total Time Chart](image)

Figure 7: Average total cost for 8 vehicles, limited outbound capacity

is excess storage so all moves in and out of storage required to created lowest cost can be accommodated and the total time plateaus. It is interesting to note that each of the inbound/outbound ratios seems to have a different place where the plateau starts and that the lowest total time at larger storages is not associated with the greatest outbound capacity. The latter observation is magnified in Figure 8 with some additional ratios added.

![Total Time Chart](image)

Figure 8: Expanded view of Figure 7
First note that on an absolute scale in a practical context, this isn’t a huge issue – less than 20% difference on inbound/outbound rations that almost double. (We will now focus exclusively when all times have plateaued at a storage capacity of 240.) On the other hand, we submit that it is important enough to merit continued exploration so that design of the hub and the heuristics that will handle this distributed control will promote the most efficient operations possible. The more interesting part of Figure 8 is that the maximum total time is seen with the most congestion when the inbound and outbound are almost equal at 1.15. The average time over all replications becomes slightly less, about 3%, when the ratio increases to 1.35 but then decreases about a 10% decrease with a ratio of 1.55. After that point, the total time begins to increase as the ratio increases. Looking at several of the heuristics across a number of replications, it appears that the sequential nature of the control (i.e., the unit loads in each transportation container are allocated as they arrive to the hub with limited knowledge of future activity) causes this behaviour. When there is extra capacity of material handling equipment and storage, more moves can be made, especially temporarily placing them in storage, to reduce the total cost. When the extra material handling equipment and storage capacity is small, there are fewer options for extra moves to reduce cost so time and cost increase because the unit loads must be moved before the end of the planning horizon. The transition for the heuristics we tested all occur at about a ratio of 1.5. Better understanding the dynamics surrounding this observation is the next phase of this research.

5 Conclusion

This experimental investigation aimed at gaining a rudimentary understanding of PI intermodal hubs is in the early stages but some potentially interesting results have already been found. At the outset, we submit that the key contextual element here is not that this is a PI hub but that the control of unit loads is distributed and not centralized. This means that decisions must be made dynamically and in real time with limited and changing information. This research simplified the scenarios by preserved the sequential nature of the decision making. Different outcomes for the unit loads were assigned costs and several simple seed heuristics were developed to find low cost solutions and key parameters like total material handling time and total number of unit loads shipped to destinations other than the correct one were monitored. Trends were plotted by varying some parameters of the scenarios and hubs like ratio of outbound to inbound freight flow and storage capacity. Moving forward, this experimental investigation is continuing with added detail to scenario and models like different arrangements of the storage at the hub that impacts the travel and handling time. Longer horizons are being considered so the amount of information available to the decentralized controller can be varied. Finally, we hope to develop an optimization model to gain insight into the layout of the hub, especially the location of the storage and the aisle configuration.
6 Bibliography


Acknowledgements
This material is based upon work supported by the National Science Foundation under Grant No. NSF-IIP-1266013.
Appendix

Random location route construction heuristic

**INPUT**: Inbound capacity, Outbound capacity and destination parameters for time window t

1: for y = 1 To \( y_{max} \) do  
   ▶ Iterate for a pre-determined number of iterations
2:   Assign inbound dock-doors randomly. Assign outbound dock-doors based on least distance between largest outbound demand to outbound container.
3: for i = 1 To \( n_I \) do
4:   for j = 1 To \( n_O \) do
5:     Determine direct outbound assigning quantities.
6:     Generate routes. Calculate Total elapsed time and Cost.
7:     \( NotShipped_{ij} = \max(n_{ij} - CapacityLeft_{ij}, 0) \).
8: for i = 1 To \( n_I \) do  
   ▶ Reship or Store not shipped unitloads
9:   for j = 1 To \( n_O \) do
10:  if \( NotShipped_{ij} > 0 \) then
11:    Choose Storage first or Reship first
12:   if Storage first then
13:      Stored = min(NotShipped_{ij}, S - UsedStorage).
14:      NotShipped_{ij} = NotShipped_{ij} - Stored
15:      UsedStorage = UsedStorage + Stored
16:   if \( NotShipped_{ij} > 0 \) then
17:      Find alternate location k
18:      Reship_{ij} = min(NotShipped_{ij}, CapacityLeft_{k})
19:      NotShipped_{ij} = NotShipped_{ij} - Reship_{ij}
20:      CapacityLeft_{k} = CapacityLeft_{k} - Reship_{ij}
21:      Generate routes. Calculate Total elapsed time and Cost.
22:     if \( NotShipped_{ij} > 0 \) then  
       ▶ If not stored or reshipped
23:        create a special shipment.
24:   else  
       ▶ Reship First
25:      Calculate minimum cost for shipping to alternate location k
26:   if Alternate Shipping Cost < Storage Cost then
27:      Reship_{ij} = min(NotShipped_{ij}, CapacityLeft_{k})
28:      NotShipped_{ij} = NotShipped_{ij} - Reship_{ij}
29:      CapacityLeft_{k} = CapacityLeft_{k} - Reship_{ij}
30:      Generate routes. Calculate Total elapsed time and Cost.
31:   if \( NotShipped_{ij} > 0 \) then
32:      Stored = min(NotShipped_{ij}, S - UsedStorage).
33:      NotShipped_{ij} = NotShipped_{ij} - Stored
34:      UsedStorage = UsedStorage + Stored
35:   if \( NotShipped_{ij} > 0 \) then  
       ▶ If not stored or reshipped
36:      Create a special shipment.
37:   Calculate Total Cost
38: Choose iteration with lowest Total Cost = \( BestSolution \)