Crowd Supported Maximum Flow from Sender to Receiver

Regular paper

Elizabeth Varki
University of New Hampshire
email: varki@cs.unh.edu

Abstract

It takes an instant to transmit a file over the Internet, but it could take hours, even days, to transmit a large file - in the Terabytes range - from one Internet user to another. The reason for the delay is the non-availability of large bandwidth for individual users at end networks. End networks (Tier 3) purchase fixed Internet bandwidth. The network's users share bandwidth with other users of the network; during the day when Internet usage is high, each user gets proportionally little bandwidth. During early morning hours, when Internet traffic is low, more bandwidth is available to individual users, so large file transmission should be scheduled for these low traffic times. When the sender and receiver are in different time zones, their high bandwidth availability times are asynchronous. End-to-end transmission rate is determined by the smaller of the two bandwidth capacities, so transmission rates are low when the sender and receiver are separated by several time zones.

The transmission can be sped up by collaborating with users at other end networks located in time zones closer to the sender or the receiver. The sender divides the file into segments and transmits segments to users at transit end networks; these transit users later transmit to other end networks or to the receiver. Large segments require large bandwidth at the transit end networks. It is possible to transmit a large file quickly without using large bandwidth at the transit end networks, if the sender divides the file into 1-10 MB sized micro segments; the micro segments are transmitted to a crowd of end networks, one micro segment per crowd end network.

This paper evaluates crowd supported large file transmission. The paper formulates the problem as a maximum flow problem, so it can be solved by a maximum flow algorithm. The outputs from a maximum flow algorithm are the number and sizes of micro segments, the transmission start times for each micro segment, the number of crowd nodes, the time zone location of each crowd node, the bandwidth capacity of each crowd node, and the transmission path - sender, crowd hops, receiver - for each micro segment. This paper does not develop a new maximum flow algorithm; this paper develops the flow network that is input to the maximum flow algorithm.

The only input parameter values provided by the problem are the sender and receiver's end network bandwidth distributions. The flow network for crowd supported flow models thousands of crowd end networks and the flow capacities between them. To construct the flow network, the significant Internet parameters are extracted and then mapped to graph parameters. The Internet has a dense, mesh structure with thousands of end networks that span the globe. The flow network developed in this paper, called GPSnet, captures the size, span, and inter-connectivity of the Internet. The contribution of this paper is GPSnet, a flow network for crowd supported Internet flow; the complexity of GPSnet is O(T), where T is the flow duration.

1 Introduction

This paper evaluates how a sender transmits a large file to a receiver under the constraint that the sender and receiver have bandwidth at different times. For example, a sender in Chicago wants to transmit a 1 TB file to a receiver in Japan; the sender has 1 Gb/s bandwidth from 06:00-09:00 UTC and the receiver has 5 Gb/s bandwidth from 15:00-20:00 UTC. In this scenario, it is impossible to transmit the file directly from Chicago to Japan. It is possible to transmit the file indirectly from Chicago to Japan via one or more intermediate - transit - hops. The sender transmits the file from Chicago to a transit end network during 06:00-09:00 UTC. If a single transit end network does not have sufficient bandwidth, then the sender divides the file into segments and transmits segments to several end networks; the transit end networks store segments until 15:00 UTC and then transmit to the receiver in Japan.

The above problem has practical applications. End networks purchase fixed Internet bandwidth, so bandwidth is a limited resource, which is shared by the end network's users. Internet traffic at an end network follows the diurnal sleep-wake cycle of its users; network traffic has a wave distribution, which crests around 8:00 PM local time and troughs around 2:00 AM local time [9]; more bandwidth is available during early morning hours when fewer Internet applications are running. Network administrators are likely to permit bandwidth intensive applications scheduled for early morning hours when there are fewer Internet applications whose performances could be disrupted. Bandwidth intensive applications, such as large file transmission, have access to maximum "free" bandwidth during early morning hours when few users are on the network. When the sender and receiver are in different time zones, their times of high bandwidth are mismatched. This paper develops a scheduler that ensures transmission from the sender during early morning hours and transmission into the receiver during early morning hours, regardless of the geographic location of the sender and receiver.

The goal is to transmit a large data set from the sender's network to the receiver's network cheaply and quickly. We analyze large file transmission from the perspective of end users of the Internet. The sender and receiver have access to existing network infrastructure and file transmission protocols such as ftp, HTTP, gridFTP. End users do not have permission to route packets, so this paper does not evaluate packet routing. End users, however, may divide the file into segments and transmit segments to users at various end networks; these transit users, in turn, transmit to other transit users or to the receiver.

Large file transmission is a bandwidth intensive application requiring large bandwidth for extended periods. The sender and receiver either purchase bandwidth and/or get permission from their network administrators. If there are few transit end networks, then the segment sizes are also large, requiring considerable bandwidth at the transit end networks. It is unlikely that end network administrators would permit hops in-and-out of their networks without payment, even if the transmissions are scheduled for early morning hours of high bandwidth availability. Therefore, we propose dividing the file into *micro segments* that are 1-10 MB in size. Instead of a few transit end networks, a *crowd* of transit end networks participate in the transmission; each end network in the crowd is a transit hop for at most one micro-segment. By dividing the file into micro segments, the file is transmitted by reserving large bandwidth only at the sender and receiver's networks.

This paper develops a scheduler that determines the locations of the crowd of end networks, the number and sizes of micro segments, the transit hops in a micro segment's path from sender to receiver, and the transmission times at the sender and at each hop. We refer to the crowd flow scheduler as **Flowes**: **Flow** of **E**lectronic **S**egments. Flowes solves the logistical problem of user-controlled, crowd-supported maximum flow from a sender's network to a receiver's network. The maximum flow is a popular, thoroughly researched problem [5]. The search complexity of a maximum flow algorithm depends on the number of nodes and edges in the underlying flow network. We do not develop a new maximum flow algorithm for Flowes; we use the Edmonds Karp maximum flow algorithm [6]. We develop the input to the maximum flow algorithm - the flow network - a directed graph modeling user-controlled Internet flow. The complexity of our flow network is O(T) where T is the flow duration. The contribution of this paper is the maximum flow model of the Internet for crowd Flowes.

2 Related research

We are the first to evaluate the problem of crowd supported, user controlled, maximum flow on the Internet when the end networks have bandwidth at different times. We do not develop a new algorithm to solve the problem since there are several maximum flow algorithms [5]. We develop the flow network for crowd supported maximum flow. The Internet is the "network of networks": it consists of thousands of end networks, each end network has thousands of users; the Internet spans the globe and has a mesh structure with several paths between end networks. This paper develops a scalable model of Internet flow.

We are not the first to propose transmitting large files during early morning hours when more bandwidth is available. Prior research on Internet flow has focused on store-and-forward transmission protocols. Store and forward bulk transmission protocols are delay tolerant and transmit packets when large bandwidth is available [7, 11]. The delay tolerant protocols [10] rely on the sleep-wake Internet traffic pattern [9]. Packets from bulk transmission are forwarded during early morning hours when traffic is low and high bandwidth is available.

The fundamental difference between our research and prior store-and-forward research is crowd supported flow. We are the first to propose crowd supported flow where large bandwidth is reserved only at the sender and receiver end networks. Bulk transmission protocols [10] include few transit end networks in the transmission, but crowd supported flow includes several thousands of transit end networks.

The flow network developed in prior papers is a complete graph [4, 10] and models the mesh structure of the Internet; when there are n transit nodes, the number of edges is O(n^2). When bandwidth for maximum flow is only available during early morning hours, it may take days to transmit a large file. Therefore, another relevant parameter in maximum Internet flow is the flow duration T. The complexity of prior Internet flow networks is O(n^2T) where n is the number of transit end networks and T is the flow duration. For crowd supported flow, the value of n ranges from several thousands to million. Modeling crowd supported flow with existing O(n^2T) flow networks is impractical - the scale of the input model would overwhelm the maximum flow algorithm. We develop a new flow model - with O(T) complexity - for crowd supported flow.

3 Direct flow from sender to receiver

We first develop the flow network for the simplest case of direct flow from sender to receiver. In a graph, the sender and receiver are modeled by nodes and their bandwidth is modeled by the capacity of the edges connecting the nodes. The edge capacity is the minimum of the sender's bandwidth and the receiver's bandwidth. This section highlights: 1) the system parameters relevant to Internet flow, 2) the mapping from system parameters to graph parameters, and 3) the necessity of including other end networks in the flow.

The sender and receiver are end network ASs (Autonomous Systems). The relevant system parameter is the bandwidth available to the sender and receiver for maximum flow. Let the sender be represented by s and the receiver by r. The sender has uplink, ul, to the Internet and the receiver has downlink, dl, from the Internet. The uplink/downlink capacity of an end network is the minimum of the network's LAN capacity and its backbone Internet capacity. The values of ul(s) and dl(r) are usually public.

The bandwidth capacity between s and r is determined not only by the uplink from s and downlink to r but also by the links between them. The sender and receiver ASs are end network LANs connected via backbone ASs (Wide Area Networks - WANs); there are several paths between s and r. Let WAN(s,r) represent the total flow capacity from s to r along backbone networks. It is difficult to compute WAN(s,r) since backbone Internet

Service Providers (ISPs) conceal network details. End network customers pay for 95^{th} percentile backbone bandwidth usage; this business policy, in theory, ensures that end networks have access to the backbone capacity they purchase. If WAN details are unavailable, then it is assumed that WAN(s,r) is not the bottleneck. This is a reasonable assumption since the architectural and business features of the Internet ensure that the backbone links are not usually a bottleneck.

3.1 Constant bandwidth:

We first develop the flow model when the sender and receiver have constant bandwidth for maximum flow. The capacity, c, of the arc, (s,r), is the minimum of the sender's uplink, the receiver's downlink, and the total backbone links between the nodes. $c(s,r) = \text{minimum}\{\mathsf{ul}(s), WAN(s,r), \mathsf{dl}(r)\}$. The flow network G=(V,E) has node set V=$\{s,\ r\}$ and arc set E=$\{$(s,r)$\}$) where arc (s,r) has capacity function c(s,r). A flow in G defined by function f: (s,r) $\longrightarrow \mathbb{N}_0$ that satisfies the capacity constraint f(s,r) $\leq$ c(s,r). ($\mathbb{N}_0 \equiv \mathbb{N} \cup \{0\}$.) The largest file transferred from s to r is the maximum flow value $\mid f_{max} \mid = c(s,r)$. Example 1 in Appendix A computes flow when the sender and receiver have constant bandwidth at all times.

3.2 Variable bandwidth

Consider the realistic scenario of a Flowes user sharing bandwidth with the rest of an end network's users. End network administrators purchase bandwidth by estimating peak usage at their networks. A Flowes user has access to the end network's available bandwidth, which is the purchased bandwidth that remains unused. The distribution of this remaining bandwidth depends on the Internet traffic distribution at the end network. The Internet traffic at end networks has a predictable diurnal wave distribution [9] where traffic increases gradually during the day with peak bandwidth usage between 6:00 PM and 10:00 PM; the traffic drops off sharply after midnight.

With variable bandwidth, a new system parameter - UTC time - becomes relevant. The UTC time, which is a system parameter, must be mapped to a graph parameter. Flow networks that model time are called *dynamic* (as opposed to the *static* flow network developed in the previous subsection). In a dynamic flow network, the parameter that models time is the flow instant. We map UTC instants to flow instants; we then develop the dynamic flow network, which models Flowes when the sender and receiver have variable bandwidth.

The available uplink capacity ul and the available downlink capacity dl of end networks are a function of time of day. The WAN(s,r) bandwidth also varies with time. Let τ represent UTC time. This paper models time in discrete increments, $\tau = 0, 1, 2, ..., \Gamma - 1$, where Γ is the total number of time instants in a day. The value of Γ depends on the time unit; $\tau = 0$ refers to the UTC time interval starting at 00:00 UTC, and $\tau = \Gamma - 1$ refers to the last UTC time interval ending at 00:00 UTC. For example, if time unit is an hour, then $\Gamma = 24$ and link capacities are defined for each hour in [00:00-23:00]; $\tau = 0$ represents UTC [00:00-01:00), $\tau = 1$ represents UTC [01:00-02:00), ..., and $\tau = 23$ represents UTC [23:00-00:00). If time unit is 5 minutes, $\Gamma = 288$, $\tau = 0$ represents UTC [00:00-00:05), $\tau = 1$ represents UTC [00:05-00:10), ..., and $\tau = 287$ represents UTC [23:55-00:00). The link capacities, ul(s,τ), dl(r,τ), and WAN(s,r,τ), are defined for each UTC instant. For example, suppose time unit is 1 hour: if uplink for s is 10 Gb/s from UTC[03:00- 04:00), then ul(s,3)=4500 GB. To improve clarity, we drop units such as MB, GB, Gb when specifying bandwidth capacity per instant, so ul(s,3)=4500. Example 2 of Appendix A shows a mapping of end network distribution against UTC time instants.

The UTC time zone (offset from local time) specifies the geographic positions of s and r; thus, the global span of the Internet is incorporated in this model. Let θ represent the UTC instant at which Flowes is initiated. In Example 2 (Appendix A), if Flowes is initiated at 09:00 UTC ($\tau = 3$) then $\theta = 3$. Suppose Flowes duration is 6 time instants; the UTC instants relating to Flowes duration are: τ=θ=3, τ=θ+1=4, τ=θ+2=5, τ=θ+3=6, τ=θ+4=7, τ=θ+5=0. Thus, Flowes runs from 09:00 UTC to 03:00 UTC the next day.

Graph model - mapping UTC instant τ to flow instant t:

In graph theory, a dynamic flow network models flows over time instants t=0,1,2,...,T-1 [8]. At flow instant t=0, the flow starts, and at flow instant t=T-1, the flow ends. Thus, t=0 of the flow graph corresponds to UTC instant τ=θ and t=T-1 of the flow graph corresponds to UTC instant τ=θ+T-1 in modulo Γ arithmetic. The modulo Γ arithmetic is a consequence of the periodicity of UTC (system) time, where period is one day. The flow network models flow instants t=0, 1, 2, ..., T-1 which corresponds to UTC instants $\tau = \theta, \theta+1, ..., \theta+T-1$ modulo Γ.

The flow network G=(V,E) has node set V={s, r} and arc set E={(s,r)} in which arc (s,r) has capacity function c(s,r,t) defined by: $c(s, r, t) = \text{minimum}\{\mathsf{ul}(s, \tau), \text{WAN}(s, r, \tau), \mathsf{dl}(r, \tau)\}$ where $0 \leq \theta < \Gamma$ is the UTC instant when flow is initiated, $0 \leq t < T$ is the flow instant and T is the flow duration. The system time parameter τ is mapped to the graph time parameter t by $\tau = \theta + t$ in modulo Γ arithmetic. A flow in G is a function $f : (s, r) \times [0, T) \longrightarrow \mathbb{N}_0$ that satisfies the capacity constraint $f(s, r, t) \leq c(s, r, t)$. The largest file that can be transferred from s to r is given by the maximum flow value $\mid f_{max} \mid = \sum_{t=0}^{T-1} c(s, r, t)$.

Examples 3 and 4 of Appendix A compute maximum flow from UK to Japan at different start instants. In Example 3, maximum file size transmitted from UK to Japan is 8 when Flowes is initiated at θ=6 (UTC 18:00-21:0); in Example 4, maximum file size transmitted from UK to Japan is 0 when Flowes is initiated at θ=1 (UTC 03:00-06:00).

A consequence of variable bandwidth in the Flowes model is that a new parameter, namely, start time, becomes significant to maximum flow computation. Another consequence of variable bandwidth is that sender to receiver flow is not necessarily continuous. After start time, if there are time instants at which the sender or the receiver have no bandwidth, then flow stops, before resuming at another instant. Thus, the file is divided into segments, and each segment takes exactly one instant to flow from sender to receiver.

4 Indirect flow from sender to receiver via crowd

When the sender and receiver are in different time zones, their high capacity times are mismatched. In Example 2 of Appendix A, maximum flow from UK to Japan is 8 even though each end network has 56 bandwidth units per day. A maximum of 56 units can be transmitted from UK to Japan if a suitable "transit" user is found. Suppose a transit user in Germany has access to 20 units/3 hours for the entire day. UK starts transmitting at UTC instant 0 when UK has uplink capacity 10 and Japan has downlink capacity 8. At UTC 0, UK transmits 8 units to Japan and the remaining 2 units to Germany. During the next 3 instants, UK transmits to Germany since Japan has 0 bandwidth. When bandwidth becomes available in Japan, Germany transmits to Japan. In the previous section, we constructed the flow network with only two nodes, namely, the sender and receiver, Here, we construct the flow network with a sender, receiver, and one or more transit end networks.

When other networks are included in the maximum flow transmission, the flow from sender to receiver is not continuous. The sender divides the file into segments; each segment follows a different path from the sender to receiver; some paths are direct, while other paths include one or more transit hops. If the flow from sender to receiver is direct, then the segment flow takes one time instant (end-to-end flow). If the flow is indirect, then the segment flow takes several time instants; a segment is transmitted to a transit hop where it is stored for one or more time instants before being transmitted to the receiver or the next transit hop.

When the Flowes transmission includes few transit end networks, then the segment sizes are large. Consequently, large bandwidth is required not only at the sender and receiver's networks but also at the transit end networks. It is unlikely that transit end network administrators would allow such large transmissions to-and-from their networks without payment. Therefore, we evaluate the scenario where heavy bandwidth usage is restricted to just the sender and the receiver. The sender divides the file into 1 MB to 10 MB micro segments, which are typical of file sizes in a single ftp or HTTP transmission. The micro segments are transmitted to a crowd of transit users' end networks. A crowd end network transmits its micro segment, on schedule, to another crowd end network or to the receiver. When all the micro segments arrive at the receiver, the file is reassembled. The transmission of a single micro segment uses little bandwidth. The advantage of transmitting micro segments to a crowd of end networks is that bandwidth need not be purchased at transit end networks. The large file is transmitted indirectly by using large bandwidth only at the sender and receiver.

Input parameters

The crowd Flowes model includes a crowd of end networks. The Internet contains several thousands (millions) of end networks, so the challenge is to construct a scalable flow model. For crowd nodes, the following issues are relevant when constructing the model:

1. how many crowd nodes should be modeled?

   The Internet has several thousands of end networks. A crowd Flowes network should model the infinity of crowd nodes.

2. where should crowd nodes be located - uniformly across the globe or clustered around the sender and receiver?

3. how should the location of a crowd node be modeled?

   UTC time zones capture global positioning, so UTC time zones should be modeled.

4. how to model the constraint that a crowd end network should handle the hop of at most one micro segment?

   At every flow instant, the bandwidth capacity of a crowd node should not exceed the bandwidth to transmit a single micro segment. Let b represent the capacity to transmit a micro segment, and let u represent a crowd node. For all UTC instants τ, $\mathsf{ul}(u, \tau) = \mathsf{dl}(u, \tau) = b, \quad \forall 0 \leq \tau < \Gamma$.

5. should WAN link capacities between crowd nodes be modeled?

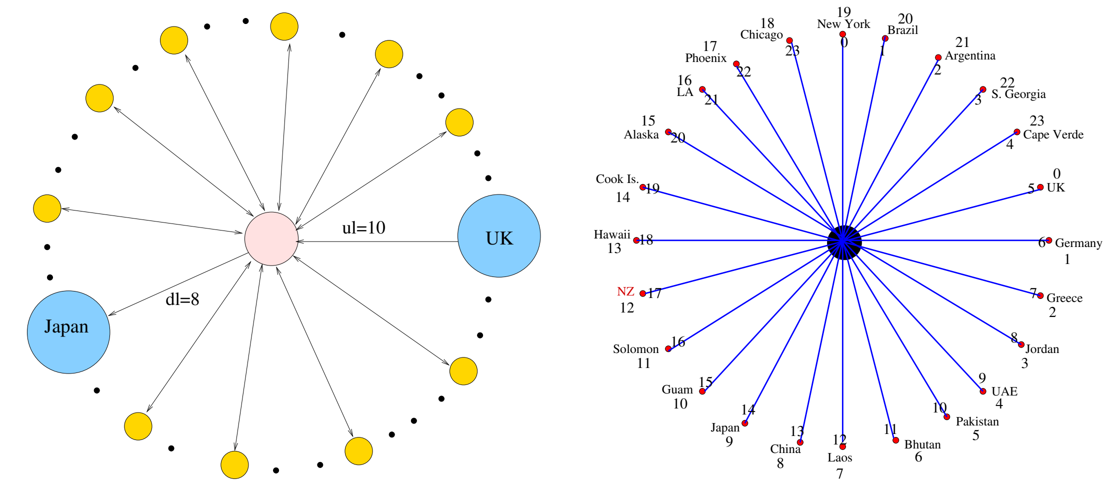
   No, since end networks transmit at most one micro segment.

6. should WAN link capacities in regions be modeled?

   Yes, since a crowd of end networks transmitting micro segments stresses Internet exchanges and long haul networks of that region.

7. what is the optimal flow time unit for crowd Flowes?

   For direct sender to receiver flow, the flow time unit is assumed to be small enough to capture the variance of bandwidth distribution. A time unit of 5 minutes is sufficient to capture variance in bandwidth [11]. For crowd supported maximum flow, 5 minutes is too large since a micro segment is transmitted in less than a second. For crowd Flowes, the time unit should be small enough to capture end-to-end flow of a single micro segment. A realistic time unit is 1 second, 1/2 second, or 1/4 second, or smaller (depending on the bandwidth capacity of the sender). When time unit is a second and flow duration is a day, $\Gamma = 86400$ and T=86400; when flow duration is 2 days, T=172800. Thus, for crowd Flowes, the number of flow instants is in the thousands.



(i) Star graph - models end network bottlenecks (ii) WANnet - models WAN bottlenecks

Figure 1: Flow networks that model some characteristics of crowd Flowes

4.1 Graph configurations

We evaluate different graph structures for crowd Flowes. We start with intuitive, naive graphs; these models are unsuitable since they are too large or they only capture some of the essential features of crowd Flowes. We end with a scalable model of crowd Flowes that encapsulates all its essential characteristics.

Complete graph

The complete graph models the flow network underlying bulk transmission protocols [10]. The Internet is a global, dense network with several paths between any two nodes. Thus, the Internet has a mesh structure and is naturally modeled by a complete graph. For crowd Flowes, the nodes represent end networks - sender, receiver, and crowd; the edges represent flow between the end networks. The complete graph models the mesh architecture of the Internet, but the graph does not model several parameters relevant to crowd Flowes: end network bottlenecks, WAN bottlenecks, and crowd network locations, to name a few.

Star graph

Next, we model crowd Flowes with a star graph shown in Figure 1(i). Each crowd end network is represented by a leaf node, with arcs to-and-from the hub node. The arc from a leaf to the hub represents the network's uplink and the arc from a hub to a leaf node represents the network's downlink. The hub node represents an Internet eXchange (IX) where networks connect. The star graph models end network bottlenecks and the connectivity of the Internet, but the star graph does not provide a suitable framework to model crowd nodes. The model does not provide insight into the number of crowd nodes to be modeled. Moreover, the positioning of a crowd node is not modeled.

Crowded WANnet graph

There are two bottlenecks in crowd Flowes - the end network bottleneck and the WAN bottleneck. The traffic pattern at Internet exchanges (such as DE-CIX [2], AMS-IX [1], and LINX [3], to name a few), follow the sleep wake cycle, similar to that at end networks. Therefore, WAN bottlenecks are modeled for each time zone. The WAN bottleneck is equal to the available capacity of IXs or long haul networks in the region. The WAN bottleneck is modeled as a star graph with 24 leaf nodes, representing the 24 standard time zones. Each time zone (IX) node has 2 edges connecting it to the hub - one edge represents the WAN uplink and the other

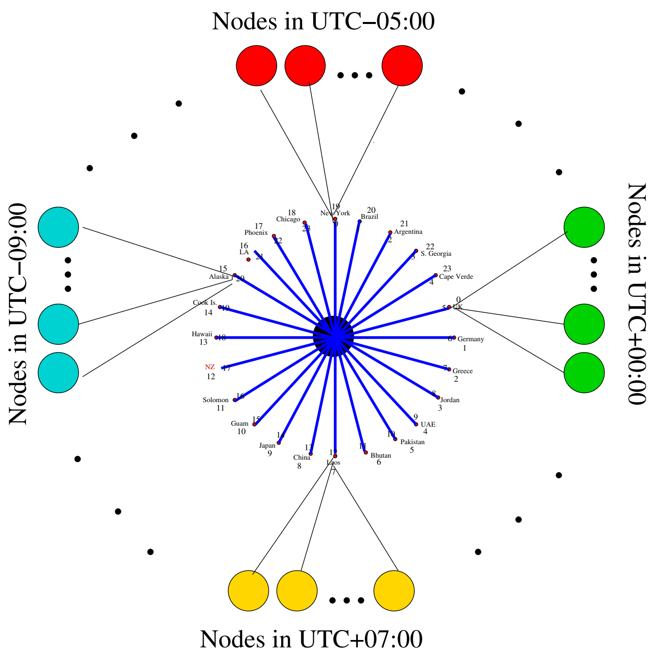


Figure 2: Crowded WANnet graph

edge represents the WAN downlink. Figure 1(ii) shows the star graph representing WAN bottlenecks; we refer to this graph as the WANnet. In the Figure, the edges representing uplinks and downlinks are shown as a single edge with no arrows; each time zone node (IX node) is labeled by: the UTC zone number, the name of a country/city in the region, and the local time in that region at the flow instant represented. (The only necessary label is the UTC zone number; the other labels add meaning by pinpointing geographic positioning.)

Figure 2 presents the flow network with crowd nodes and the WANnet. Every WANnet leaf (IX node) is linked to several crowd nodes. The sender and receiver nodes are each connected to the WANnet leaf corresponding to their regions. An issue is the number of crowd nodes modeled. A back-of-the-envelope estimate is that the number of crowd nodes attached to an IX node is the smallest number such that the total capacity of all the crowd nodes matches (or just exceeds) the WAN capacity of the region for the duration of the flow. The end networks store micro segments, so they have storage capacity, while the IX nodes and the hub have no storage.

The crowded WANnet graph models all the relevant parameters, which are: the number of crowd nodes, the geographic positioning of crowd nodes, the end network bottlenecks, and the WAN bottlenecks. The downside of the crowded WANnet graph is its size - determined by the number of nodes (and edges), n, and the number of flow instants, T. The crowded WANnet flow network models every crowd node explicitly. The value of n lies in the thousands to million range. Since time unit is less than a second and flow duration lasts for one or more days, the value of T is several thousands. The time expanded crowded WANnet flow network, which models flow over instants 0, 1, 2, ..., T-1 contains O(nT) nodes. It is impractical to compute maximum flow with a flow network of this size.

4.2 Crowd Flowes graph: GPS flow network

The defining characteristic of the crowded star graph is size; even though the flow network captures the essential features of crowd Flowes, the size of the graph would overwhelm a maximum flow algorithm. This section

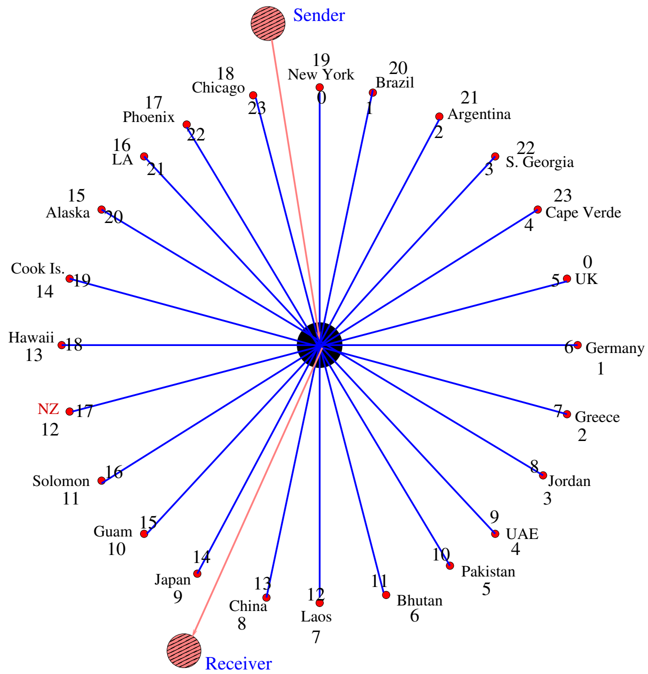


Figure 3: GPSnet: Flow network for crowd Flowes: Each node has 3 labels: the outer number represents UTC zone number and is the only necessary label. The country/city label adds meaning by signifying global positioning of the node. The inner number represents local time.

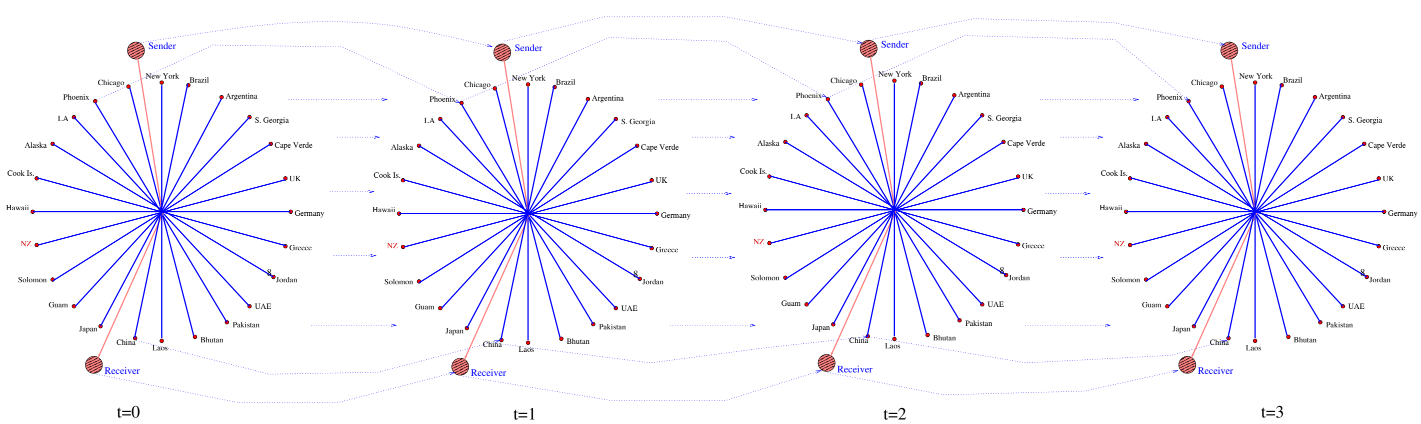


Figure 4: Time expanded GPSnet: 4 sub graphs representing flow instants t = 0, 1, 2, 3.

presents **GPSnet**, a scalable flow network for crowd Flowes.

Figure 3 presents GPSnet; it consists of the WANnet graph and the sender and receiver nodes. The sender node has uplink to the hub and the receiver has downlink from the hub; all the other nodes have uplink and downlink edges to-and-from the hub. (In the Figure, the two edges are not shown explicitly, but are implicitly represented by a single edge with no arrows.) The capacities of the sender and receiver edges are set to the available bandwidth of their respective end networks. The WANnet nodes represent timezone IXs and their edges have total WAN uplink/downlink available capacity per region, per time instant.

The flow network, GPSnet, has a total of 24 + 2 leaf nodes representing the 24 standard time zones and the sender/receiver; there is a single hub node. The leaf nodes have storage capacity; the hub node has no storage capacity. This ensures that GPSnet models end-to-end flow; a flow initiated from leaf node u at time t will arrive at a leaf node v at time t (or it could stay in node u until next flow instant t+1). When flow is initiated from leaf nodes u to leaf node v, the flow cannot exceed the minimum of ul(u) and dl(v).

The question is: how does GPSnet model crowd Flowes when there are no crowd nodes in the network? The crowd nodes are modeled implicitly by time zone nodes and by setting flow time unit to be small enough to explicitly model the transmission of a single micro segment. Each WANnet time zone IX node represents all the crowd nodes of the region. The uplink and downlink capacities represent the maximum in-flow and out-flow for the region. The flow instant is such that the sender node can transmit at most a single micro segment at each instant. By setting small flow instants, all flow from the sender to the receiver will be restricted to micro segments. Each micro segment that arrives at a time zone IX is assumed to be directed to a unique end network in the region. The time zone IX nodes and the flow time units implicitly model the crowd nodes participating in the flow. Hence, we have modeled crowd Flowes with a O(T) graph where T is the total flow duration.

Definition For a flow duration t=0 until t=T-1 (UTC instants $\tau = \theta$ until τ = θ+T-1, $0 \leq \tau, \theta < \Gamma$), the crowd Flowes network (GPSnet) $G^{\star}$=(V$\cup x$, E,T) where

- V is the set of 26 nodes representing the the sender s, the receiver r, and the 24 time zone IXs.
- x represents a single hub node; and
- E is the set of arcs $\{(s,x),(x,r)\} \cup \{\ (v,x),(x,v)\ \mid\ \forall v \in$ V-s-r $\}$.

$G^{\star}$ is a star network where the leaves are in node set V and the hub is node x. The nodes in V have infinite storage capacity, and the node x has zero storage capacity. The arcs in E have bandwidth capacity:

c(v,x,t) = ul(v,τ) $\forall$ (v,x) $\in$ E

c(x,v,t) = dl(v,τ) $\forall$ (x,v) $\in$ E

$\forall t = 0, 1, ..., T-1$ where τ = θ + t in modulo Γ arithmetic. □

The majority of maximum flow algorithms ignore time by assuming instantaneous flow from sender to receiver; the corresponding flow network is static. Crowd Flowes is modeled by GPSnet flow network where arc capacity varies with time; such a flow network is dynamic. Ford and Fulkerson [8] proposed the following solution for dynamic flows: convert a dynamic flow to a static flow using a *time expanded* flow network. The time expanded flow network is a static network in which there is a copy of the graph for each time instant $0 \leq t <$ T. Figure 4 shows the time expanded version of Figure 3 when duration is 4 time units. In this Figure, there are 4 sub-graphs, each representing an instant of time. The dashed lines are called *holdover* arcs and they represent storage capacity at leaf nodes. For clarity, we have not shown all the holdover arcs, but have implicitly shown them by the dashed arcs between sub graphs. The holdover arc (u_t, u_{t+1}), $u \in V$, models u at t storing micro segments for transmission at $t + 1$ Storage capacity at end networks is not a bottleneck, so capacity of holdover arcs is set to infinity.

Definition The dynamic Flowes network $G^{\star}$=(V $\cup$ x,E,T) transforms to the static crowd Flowes network G=($V_T \cup X_T$, $E_T \cup H_T$) where

- V_T is the set of end nodes u_t, $\forall u \in V$ and t=0,1,...,T-1;
- X_T is the set of hub nodes x_t, t=0,1,...,T-1;
- E_T is the set of arcs (u_t, x_t), (x_t, u_t), $\forall$ (u, x), $(x, u) \in E$ and t=0,1,...,T-1;
- H_T is the set of holdover arcs (u_t, u_{t+1}), $\forall$ $u \in V$ and t=0,1,...,T-2.

The arcs have capacity:
c(u_t,x_t) = c(u,x,t), c(x_t,u_t) = c(x,u,t), and c(u_t,u_{t+1}) = ∞.

□

The dynamic flow from s to r is equivalent to a corresponding static flow from s_0 to r_{T-1} [8]. Therefore, finding maximum flow in the dynamic network can be solved by finding maximum flow in the corresponding static time expanded network.

Definition A flow in time expanded G is a function f: $(V_T \cup X_T) \times (V_T \cup X_T) \longrightarrow \mathbb{N}_0$ satisfying the property:
$\forall$ nodes $u_t \in V_T, x_t \in X_T, t \in$[0,T), f($u_t$,$x_t$) $\leq$ c(u_t,x_t) and f(u_t,u_{t+1}) $\leq$ c(u_t,u_{t+1}).
The value $\mid f \mid$ of a flow f in G is $\mid f \mid$ = f(s_0,x_0) + f(s_0,s_1) = f(x_{T-1},r_{T-1}) + f(r_{T-2},r_{T-1})
The function, f, satisfies capacity constraint, skew symmetry, and flow conservation [5]. □

The GPSnet is a star graph with global positioning nodes that model how bandwidth capacity of networks change with time-of-day. For each node, the expanded graph shows how the bandwidth capacity of the node changes with every flow instant. When GPSnet is input to a maximum flow algorithm, it generates the maximum flow from sender to receiver along with the corresponding Flowes schedule. The Flowes schedule lists the times at which micro segments should be transmitted to-and-from various nodes starting from the sender and ending at the receiver. In Appendix B, we present a motivating example.

5 Conclusions

The paper formulates the problem of crowd supported maximum flow. We identify the relevant system parameters and map them to graph parameters. We then develop the flow network, GPSnet, which models Internet flow. The GPSnet time expanded graph is input to the Edmonds Karp algorithm. With each change in input parameter values, such as start time, a new time expanded graph must be generated (as demonstrated in Appendix B). We use GPSnet to generate the time expanded graphs from input files that store bandwidth distribution by local time.

The GPSnet is not just a star graph, it is a data structure that concisely captures global positioning (UTC), time-of-day (local time), and flow instants. The Edmonds Karp algorithm views the GPSnet simply as a directed graph. While searching for flow paths, the algorithm does not use any of the features of GPSnet nor does it use any of the characteristics of the problem. The defining feature of GPSnet is the global positioning of nodes. The defining characteristic of the problem is the dependence of bandwidth on the sleep wake cycle, which is periodic (with a period of a day). A search algorithm that views the logistical problem of crowd Flowes as a time synchronization problem - not merely as a capacity synchronization problem - would be more efficient than existing maximum flow algorithms.

References

[1] Ams ix amsterdam internet exchange; statistics, https://www.ams-ix.net, 2015.

[2] De-cix where networks meet; statistics, http://www.de-cix.net/about/statistics/, 2015.

[3] Linx london internet exchange, https://www.ams-ix.net, 2015.

[4] Parminder Chhabra, Vijay Erramilli, Nikolaos Laoutaris, Ravi Sundaram, and Pablo Rodriguez. Algorithms for constrained bulk-transfer of delay-tolerant data. In *ICC'10*, pages 1–5, 2010.

[5] Thomas H. Cormen, Charles E. Leiserson, and Ronald L. Rivest. *Introduction to Algorithms*. The MIT Press, 1 edition, June 1990.

[6] Jack Edmonds and Richard M. Karp. Theoretical improvements in algorithmic efficiency for network flow problems. *J. ACM*, 19(2):248–264, April 1972.

[7] Sushant Jain, Kevin Fall, and Rabin Patra. Routing in a delay tolerant network. *SIGCOMM Comput. Commun. Rev.*, 34(4):145–158, August 2004.

[8] L.R. Ford Jr. and D.R. Fulkerson. Constructing maximal dynamic flows from static flows. *Operations Research*, 6:419–433, 1958.

[9] Anukool Lakhina, Konstantina Papagiannaki, Mark Crovella, Christophe Diot, Eric D. Kolaczyk, and Nina Taft. Structural analysis of network traffic flows. In *SIGMETRICS'04*, pages 61–72, 2004.

[10] Nikolaos Laoutaris, Michael Sirivianos, Xiaoyuan Yang, and Pablo Rodriguez. Inter-datacenter bulk transfers with netstitcher. In *Proceedings of the ACM SIGCOMM 2011*.

[11] Nikolaos Laoutaris, Georgios Smaragdakis, Pablo Rodriguez, and Ravi Sundaram. Delay tolerant bulk data transfers on the internet. In *SIGMETRICS*, pages 229–238, 2009.

A Appendix

Example 1: *Constant bandwidth at end networks:* "What is the maximum number of bytes that can be transmitted from UK to Japan during 12 hours given that the sender's end network has bandwidth capacity of 10 Gb/s, the receiver has bandwidth capacity of 20 Gb/s and the backbone links between the end networks have a total capacity of 25 Gb/s."
For the example, ul(s)=54 TB, dl(r)=108 TB, WAN(s,r)= 135 TB, so c(s,r)=minimum$\{$54, 108, 135$\}$ = 54 TB. The maximum file size that can be transmitted from the end network in Chicago to the end network in Japan for a transmission duration of 12 hours is 54 TB. □

Example 2 - mapping bandwidth distribution: Suppose the sender's network in UK and the receiver's network in Japan permit Flowes transmission from local time [12:00 AM-12:00 PM). The available bandwidth distribution by local time is provided below:

[12:00 AM-3:00 AM): 10 units/3hours;
[3:00 AM-6:00 AM): 20 units/3hours;
[6:00 AM-9:00 AM): 18 units/3hours;
[9:00 AM-12:00 PM): 8 units/3hours.

The bandwidth distribution is specified in units/3hours for computational simplicity. Each time instant is equivalent to 3 hours, so $\Gamma = 8$. $\tau = 0$ represents UTC interval [00:00-03:00), $\tau = 1$ represents UTC interval [03:00-06:00), $\tau = 2$ represents UTC interval [06:00-09:00), ..., $\tau = 6$ represents UTC interval [18:00-21:00), $\tau = 7$ represents UTC interval [21:00-00:00).
The sender and receiver have identical distribution by local time, but they are in different time zones, so their bandwidth distribution by UTC differ.
UK is in UTC + 00:00 (zone 0), so local time equals UTC time. Therefore, the uplink, ul(s,τ), from UK is given by:

ul(s,0)=10; ul(s,1)=20; ul(s,2)=18; ul(s,3)=8; ul(s,4)=0; ul(s,5)=0; ul(s,6)=0; ul(s,7)=0.

Japan is in UTC + 09:00 (zone 9), so the local time is 9 hours ahead of UTC time (*i.e.,* 3 time units ahead). Therefore the downlink, dl(r,τ), to Japan is given by:

dl(r,0)=8; dl(r,1)=0; dl(r,2)=0; dl(r,3)=0; dl(r,4)=0; dl(r,5)=10; and dl(r,6)=20; dl(r,7)=18. □

Example 3: Consider the setup of Example 1. Suppose Flowes is initiated from UK to Japan at $\theta = 6$ for 4 times instants; $\tau = 6, 7, 0, 1$. The values of ul(s,τ) and dl(r,τ) (as computed in Example 1) are: ul(s,6)=0, ul(s,7)=0, ul(s,0)=10, ul(s,1)=20; dl(r,6)=20, dl(r,7)=18, dl(r,0)=8, dl(r,1)=0. Assume that WAN(s,r) is not the bottleneck.
Mapping system parameters to graph parameters: T=4 and t=0, 1, 2, 3, so values of c(s,r,t) are

c(s,r,0)= minimum{ul(s,6), dl(r,6)} = 0,
c(s,r,1) = minimum{ul(s,7), dl(r,7)} = 0,
c(s,r,2) = minimum{ul(s,0), dl(r,0)} = 8,
c(s,r,3) = minimum{ul(s,1), dl(r,1)} = 0.

The total Flowes flow from UK to Japan, is $\sum_{t=0}^{3} c(s, r, t) = 8$ when Flowes is initiated at $\theta = 6$ for 4 time instants. □

Example 4: Reconsider Example 1, Suppose Flowes is initiated from UK to Japan at $\theta = 1$ for 4 time instants.
c(s,r,0) = minimum{ul(s,1), dl(r,1)} = 0,
c(s,r,1) = minimum{ul(s,2), dl(r,2)} = 0,
c(s,r,2) = minimum{ul(s,3), dl(r,3)} = 0,
c(s,r,3) = minimum{ul(s,4), dl(r,4)} = 0.
The total Flowes flow from UK to Japan is $\sum_{t=0}^{3} c(s, r, t) = 0$ when Flowes is initiated at $\theta = 1$ for 4 time instants. □

B Motivating example

For presentation clarity, this example shows a scaled down version of GPSnet. This GPSnet divides the globe into 8 time zones instead of 24, and we set time unit to 3 hours. With this setting, it is possible to print the output of the maximum flow algorithm (Tables 1, 2) and plot meaningful graphs.

The objective is maximal flow from Chicago (Ch) to Japan (Jp) when the sender in Chicago and the receiver in Japan are in end networks with identical bandwidth distribution by local time. Since Japan is 15 hours ahead of Chicago, the sender and receiver have different distributions by UTC time. Suppose crowd users are located in time zones that are 3 hours apart. Starting from Chicago, the Flowes users, in UTC zone order, are in: Chicago (Ch), Argentina (Ag), United Kingdom (UK), Jordan (Jd), Bhutan (Bh), Japan (Jp), New Zealand (NZ), and Alaska (Ak). (Argentina is 3 hours ahead of Chicago, Alaska is 3 hours behind Chicago.) We analyze Flowes with two bandwidth distributions. (It is assumed that a network's upload and download distributions are identical.)

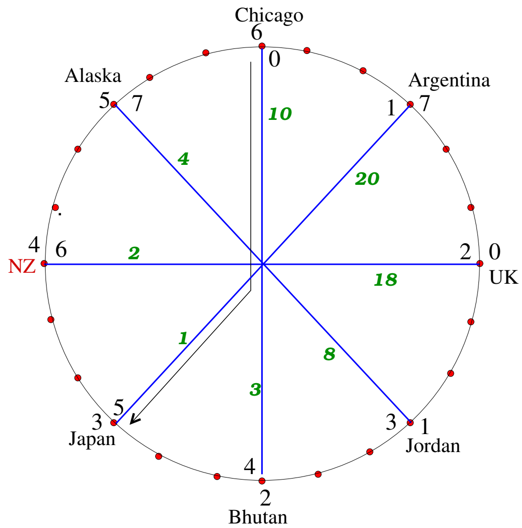


Figure 5: GPSnet showing UTC and local times at each node. The numbers outside the hub represent UTC instants while the numbers just inside the hub represent local time instants. The numbers along the links represent WAN bandwidth (distribution 1) at the corresponding instant.

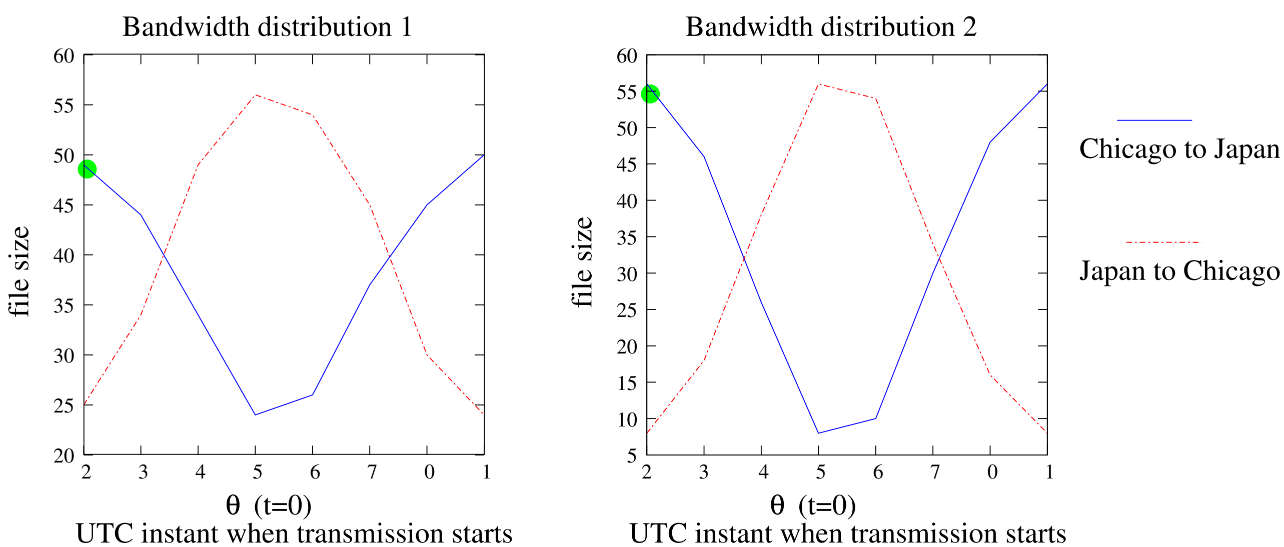


Figure 6: Impact of Flowes start time on size of transmitted file over duration of at most 24 hours ($0 \leq t < 8$). The marked points on the Chicago to Japan plots of graphs 1 and 2 correspond to total flow shown in the last row, first column of Tables 1 and 2, respectively.

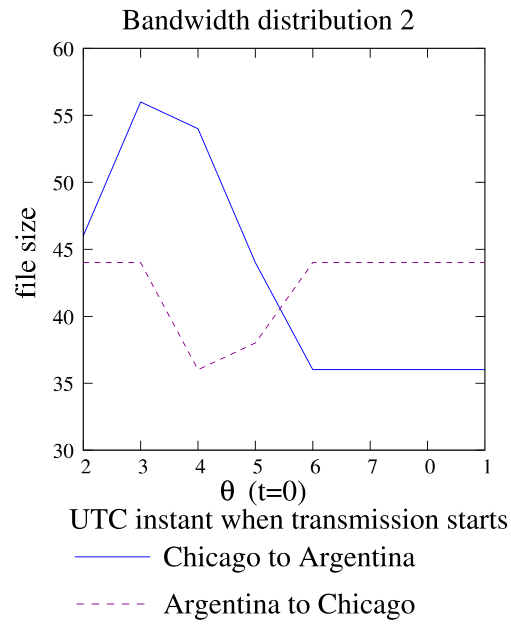


Figure 7: Impact of Flowes start time on size of transmitted file over duration <= 24 hours ($0 \leq t < 8$).

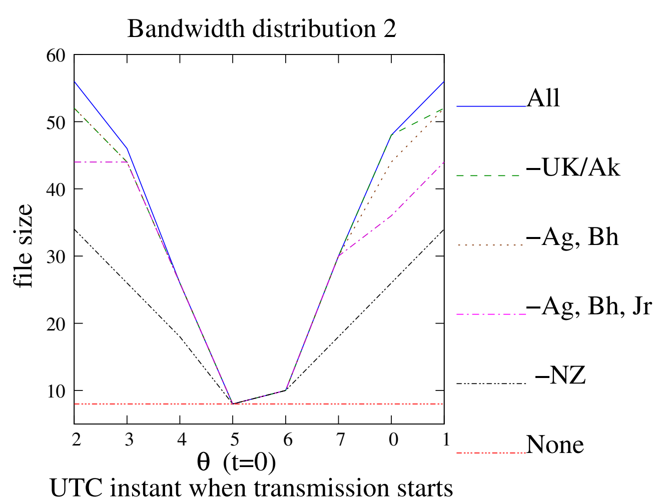


Figure 8: Impact of positioning of crowd nodes on size of transmitted file: Flowes is not permitted in some regions. All: with all zones; -UK/Ak: minus either UK or Alaska; -Ag,Bh: minus both Argentina and Bhutan; -Ag,Bh,Jr: minus Ag, Bh, and Jordan; -NZ: minus NZ; None: direct flow from sender to receiver.

Table 1: Bandwidth distribution 1: Flowes scheduler output for Chicago to Japan transmission with start time UTC 06:00 (local time in Chicago 12:00 AM) and Flowes duration of 24 hours. Each row presents a micro segment flow path from Chicago to Japan via one or more crowd nodes. The first column of the last row presents the total units transmitted from t=0 until t=7; other columns in the last row present the units transmitted to Japan at the corresponding time instant.

Seg size	UTC 06:00 (τ=2,t=0)	UTC 09:00 (τ=3,t=1)	UTC 12:00 (τ=4,t=2)	UTC 15:00 (τ=5,t=3)	UTC 18:00 (τ=6,t=4)	UTC 21:00 (τ=7,t=5)	UTC 00:00 (τ=0,t=6)	UTC 03:00 (τ=1,t=7)
1	Ch-to-Jp							
2		Ch-to-Jp						
4			Ch-to-Jp					
8				Ch-to-Jp				
2		Ch-to-Ak		Ak-to-Jp				
3					Ch-to-Jp			
3		Ch-to-Ak			Ak-to-Jp			
5			Ch-to-NZ		NZ-to-Jp			
2			Ch-to-Ak		Ak-to-Jp			
1						Ch-to-Jp		
5	Ch-to-Jd					Jd-to-Jp		
2							Ch-to-Jp	
4	Ch-to-UK						UK-to-Jp	
1		Ch-to-UK					UK-to-Jp	
3			Ch-to-Ak	Ak-to-NZ		NZ-to-Jp		
3								Ch-to-Jp
49*	1	2	4	10	13	9	7	3

Bandwidth distribution 1: The bandwidth distribution, by local time, at Chicago and Japan is as follows: 12:00 AM-3:00 AM: 10 units; 3:00 AM-6:00 AM: 20 units; 6:00 AM-9:00 AM: 18 units; 9:00 AM-12:00 PM: 8 units; 12:00 PM-3:00 PM: 3 units; 3:00 PM-6:00 PM: 1 unit; 6:00 PM-9:00 PM: 2 units; 9:00 PM-12:00 AM: 4 units. (We have moved away from bandwidth units such as b/s and simply give bandwidth as flow values.) The time zone IXs have the following distribution by local time: 12:00 AM-3:00 AM: 5 units; 3:00 AM-6:00 AM: 5 units; 6:00 AM-9:00 AM: 5 units; 9:00 AM-12:00 PM: 5 units; 0 units for the remaining time. Thus, WAN networks (other than WAN in the sender and receiver's zones) only permit Flowes from midnight until noon and have far less bandwidth than either the sender or the receiver. ☐

Table 2: Flowes scheduler output: Chicago to Japan for bandwidth distribution 2 with Flowes start time at UTC 06:00 and Flowes duration of 21 hours. Each row represents a micro segment flow path from sender to receiver via one or more crowd nodes.

Segment size	UTC 06:00 (τ=2,t=0)	UTC 09:00 (τ=3,t=1)	UTC 12:00 (τ=4,t=2)	UTC 15:00 (τ=5,t=3)	UTC 18:00 (τ=6,t=4)	UTC 21:00 (τ=7,t=5)	UTC 00:00 (τ=0,t=6)
8				Ch-to-Jp			
2		Ch-to-Ak		Ak-to-Jp			
8		Ch-to-Ak			Ak-to-Jp		
10			Ch-to-NZ		NZ-to-Jp		
8	Ch-to-Jd					Jd-to-Jp	
2			Ch-to-Ak	Ak-to-NZ	NZ-to-Jp		
2	Ch-to-UK						UK-to-Jp
6		Ch-to-UK					UK-to-Jp
6			Ch-to-Ak	Ak-to-NZ		NZ-to-Jp	
2		Ch-to-Ag	Ag-to-Ak	Ak-to-NZ		NZ-to-Jp	
2		Ch-to-Ag	Ag-to-Ak	Ak-to-NZ	NZ-to-Bh	Bh-to-Jp	
56*	0	0	0	10	20	18	8

Bandwidth distribution 2: Suppose all networks - sender, receiver, and IXs - have identical bandwidth distribution by local time. All networks are only allowed to transmit from midnight until noon. The bandwidth distribution, by local time, is as follows: 12:00 AM-3:00 AM: 10 units; 3:00 AM-6:00 AM: 20 units; 6:00 AM-9:00 AM: 18 units; 9:00 AM-12:00 PM: 8 units; 12:00 PM-3:00 PM: 0; 3:00 PM-6:00 PM: 0; 6:00 PM-9:00 PM: 0; 9:00 PM-12:00 AM: 0. ☐

Both bandwidth distributions model the sleep-wake diurnal cycle with more bandwidth being available during the early morning hours. The fundamental difference between the two distributions is the bandwidth at the IXs - in distribution 1, the IXs have small bandwidth for Flowes. Figure 5 shows the GPSnet flow graph for the example's infrastructure.

The objective is to transmit maximum flow from Chicago to Japan in 24 hours. Mapping to the graph model: since time unit is 3 hours, there are 8 GPSnet sub-graphs in the time expanded graph (Figure 4), where each sub-graph represents flow time t, $0 \leq t < 8$. For each change in input parameter values, a new expanded GPSnet is constructed and the maximum flow algorithm is run.

Tables 1 and 2 list all the maximum flow paths from Chicago to Japan which are output by Edmonds Karp algorithm when transmission starts at midnight, Chicago time. The transmission schedules show the hops along a Flowes path. Argentina and Bhutan do not appear on the schedule in Table 1. Summing column 1 of Tables 1 and 2 gives maximum flow of 49 and 56 units, respectively. The maximum flow of 49 units takes 24 hours while the maximum flow of 56 units takes 21 hours.

Figure 6 plots the file sizes transmitted from sender to receiver as starting time instant is varied. Flowes duration (for each start instant) is at most 24 hours; each time instant represents 3 hours. These graphs show the impact of start time on the transmitted file size. The first graph assumes distribution 1 and the second graph assumes distribution 2. The solid line references Chicago to Japan transmission and the dashed line references Japan to Chicago transmission. The sender and receiver have more bandwidth with distribution 1, but the transit networks are the bottleneck. Consequently, more data are transmitted with distribution 2. Figure 7 plots file sizes as start time is varied when transmissions are scheduled from Chicago to Argentina and vice versa.

Figure 8 shows the impact of removal of one or more time zone IXs. When flow is direct from sender to receiver, the start time has no effect on maximum flow since 8 units are transmitted over 24 hours with each start instant. The removal of New Zealand has a major negative impact on maximum flow. The removal of one or more of the other zones does not have significant impact on maximum flow.