

cISP: A Speed-of-Light Internet Service Provider

Debopam Bhattacharjee¹, Sangeetha Abdu Jyothi², Ilker Nadi Bozkurt³, Muhammad Tirmazi⁴, Waqar Aqeel³, Anthony Aguirre⁵, Balakrishnan Chandrasekaran⁶, P. Brighten Godfrey², Gregory Laughlin⁷, Bruce Maggs^{3,8}, Ankit Singla¹

¹ETH Zürich, ²UIUC, ³Duke University, ⁴LUMS, ⁵UCSC, ⁶MPI Saarbrücken, ⁷Yale University, ⁸Akamai Technologies

ABSTRACT

Low latency is a requirement for a variety of interactive network applications. The Internet, however, is not optimized for latency. We thus explore the design of cost-effective wide-area networks that move data over paths very close to great-circle paths, at speeds very close to the speed of light in vacuum. Our **cISP** design augments the Internet’s fiber with free-space wireless connectivity. **cISP** addresses the fundamental challenge of simultaneously providing low latency and scalable bandwidth, while accounting for numerous practical factors ranging from transmission tower availability to packet queuing. We show that instantiations of **cISP** across the contiguous United States and Europe would achieve mean latencies within 5% of that achievable using great-circle paths at the speed of light, over medium and long distances. Further, we estimate that the economic value from such networks would substantially exceed their expense.

1 INTRODUCTION

User experience in many interactive network applications depends crucially on achieving low latency. Even seemingly small increases in latency can negatively impact user experience, and, subsequently, revenue for the service providers: Google, for example, quantified the impact of an additional 400 ms of latency in search results as 0.7% fewer searches per user [13]. Further, wide-area latency is often the bottleneck, as Facebook’s analysis of over a million requests found [16]. Indeed, content delivery networks present latency reduction and its associated increase in conversion rates as one of the key value propositions of their services, citing, *e.g.*, a 1% loss in sales per 100 ms of latency for Amazon [2]. In spite of the significant impact of latency on performance and user experience, the Internet is not designed to treat low latency as a primary objective. This is the problem we address: reducing latencies over the Internet to the lowest possible.

The best achievable latency between two points along the surface of the Earth is determined by their geodesic distance divided by the speed of light, c . Latencies over the Internet, however, are usually much larger than this minimal “ c -latency”: recent measurement work found that fetching even small amounts of data over the Internet typically takes $37\times$ longer than the c -latency, and often, more than $100\times$ longer [12]. This delay comes from the many round-trips between the communicating endpoints, due to inefficiencies in the transport and application layer protocols, and

from each round-trip itself taking $3\text{-}4\times$ longer than the c -latency [12]. Given the approximately multiplicative role of network round-trip times (RTTs) (when bandwidth is not the main bottleneck), eliminating inflation in Internet RTTs can potentially translate to up to $3\text{-}4\times$ speedup, even without any protocol changes. Further, as protocol stack improvements get closer to their ideal efficiency of one RTT for small amounts of data, the RTT becomes the singular network bottleneck. Similarly, for well-designed applications dependent on persistent connectivity between two fixed locations, such as gaming, *nothing* other than resolving this $3\text{-}4\times$ “infrastructural inefficiency” can improve latency substantially.

Thus, beyond the networking research community’s focus on protocol efficiency, reducing the Internet infrastructure’s latency inflation is the next frontier in research on latency. While academic research has typically treated infrastructural latency inflation as an unresolvable given, we argue that this is a high-value opportunity, and is much more tractable than may be evident at first. It is even plausible that infrastructural improvements will be easier and faster to deploy than our ongoing decades-long efforts towards new protocols.

What are the root causes of the Internet’s infrastructural inefficiency, and how do we ameliorate them? Large latencies are partly explained by poor use of existing fiber infrastructure: two communicating sites often use a longer, indirect route because their service providers do not peer over the shortest fiber connectivity between their locations. We find, nevertheless, that even latency-optimal use of *all* known fiber conduits, computed via shortest paths in the recent InterTubes dataset [25], would leave us $1.93\times$ away from c -latency. This gap stems from the speed of light in fiber being $\sim\frac{2}{3}c$, and the unavoidable circuitousness of fiber routes due to topographic and economic constraints of buried conduits.

We thus explore the design of **cISP**, an Internet Service Provider that provides nearly speed-of-light latency by exploiting wireless electromagnetic transmissions, which can be realized with point-to-point microwave antennae mounted on towers. This approach holds promise for overcoming both the aforementioned shortcomings fundamental to today’s fiber-based networks: the transmission speed in air is essentially equal to c , and the richness of existing tower infrastructure makes more direct paths possible. Nevertheless, it also presents several new challenges, including:

- overcoming numerous practical constraints, including tower availability, line-of-sight requirements, and the impact of weather on performance;

- coping with the limited wireless bandwidth;
- solving a large-scale cost-optimal network design problem, which is NP-hard; and
- addressing switching and queuing delays, which are more prominent with the smaller propagation delays.

To meet these challenges, we propose a hybrid design that augments the Internet’s fiber connectivity with nearly straight-line wireless links. These low-latency links are used judiciously where they provide the maximum latency benefit, and only for the small proportion of Internet traffic that is latency-sensitive. We design a simple heuristic that achieves near-optimal results for the network design problem. Our approach is flexible and enables network design for a variety of deployment scenarios; in particular, we show that cISP’s design for interconnecting large population centers in the contiguous U.S. and Europe can achieve mean latencies as low as $1.05\times c$ -latency at a cost of under \$1 per gigabyte (GB). We show through simulation that such networks can be operated at high utilization without excessive queuing.

To address the practical concerns, we use fine-grained geographic data and the relevant physical constraints to determine where the needed wireless connectivity would be feasible to deploy, and assess our design under a variety of scenarios with respect to budget, tower height and availability, antenna range, and traffic matrices. We also use a year’s worth of meteorological data to assess the network’s performance during weather disturbances, showing that most of cISP’s latency benefits remain intact throughout the year.

Our weather simulation and an animation showing how the hybrid network evolves from mostly-fiber to mostly-wireless with increasing budget are available online; see [18] and [20].

Lastly, we explore the application-level benefits for Web browsing and gaming, and present estimates showing that the utility of cISP vastly exceeds its cost.

2 TECHNOLOGY BACKGROUND

At the highest level, our approach involves using free-space communication between transmitters mounted at a suitable height, *e.g.*, using dedicated towers or existing buildings, and separated from each other by at most a certain limiting distance. Network links longer than this range require a series of such transmitters. Typically, even after accounting for terrain, such network links can be built close to the shortest path on the Earth’s surface between the two end points. Further, the speed of light in air is essentially the same as that in vacuum, c . These properties make this approach attractive for the design of (nearly) c -latency networks.

Technology choices. Several physical layer technologies are amenable for use in our design, including free-space optics (FSO), microwave (MW), and millimeter wave (MMW). At present, we believe MW provides the best combination of range, resilience, throughput, and cost. Future advances in

any of these technologies, however, can be easily rolled into our design, and can only improve our cost-benefit analysis.

While hollow fiber [24] could, in the future, also provide c -latency, it would still suffer from the circuitousness of today’s fiber conduits. Low-Earth-orbit satellites may also help, but their connectivity fundamentally varies over time, necessitating extremely high density to provide latencies similar to those achievable with a terrestrial MW network.

Switching latency. While long-haul MW networks have existed since the 1940s [23], their recent use in high-frequency trading has driven innovation in radios so that each MW retransmission only takes a few μs . Thus, even wide-area links with many retransmissions incur negligible switching latency. As an example, the HFT industry operates a MW relay between New Jersey and Chicago comprising ~ 20 line-of-sight links that operates within 1% of c -latency end-to-end at the *application* layer [48].

Packet loss. Packet loss occurs for several reasons, including, notably, weather disruption, and intermittent multi-path fading, especially over bodies of water.

To examine loss, we obtained network performance data for an FCC-licensed, operational Chicago-to-New-Jersey MW relay link from an ultra-low latency wireless provider. The data comprise 2,743 distinct one-minute intervals between 10/22/2012 through 11/01/2012, spanning the intervals 9:30AM – 4:00PM EDT when both the futures and equity markets were open and trading in Aurora, IL, and Carteret, NJ, respectively. This period provides an extreme test: Hurricane Sandy caused broad, serious disruption in the NJ area for 4 days within this period. These data thus show a high average packet loss rate of 16.1%. Even for this period, the median loss rate is much smaller at 1.4%. In Section 6.1 we present a broader analysis of the impact of diverting traffic to alternate (fiber or MW) routes during inclement weather using a year’s worth of weather data.

Note also that the IL-NJ link was designed to absolutely minimize latency for the HFT industry. Hence forward error correction spanning multiple packets was used minimally or not at all. A less aggressive design using such techniques, together with shorter tower-to-tower distances is likely to further reduce loss rates.

Spectrum and licensing. We propose the use of MW communication in the 6-18 GHz frequency range. These frequencies are not very crowded, and licensing is generally not very competitive, except at 6 GHz in cities, and along certain routes, like the above mentioned HFT corridor. The licenses are given on a first-come, first-served basis, recorded in a public database, and they protect against the deployment of other links that would interfere with licensed links.

Line-of-sight and range. Successive MW towers need line-of-sight visibility, accounting for the Earth’s curvature, terrain, trees, buildings, and other obstructions, and atmospheric refraction. Attenuation also limits range. A maximum range

of around 100 km is practicable, but we show results with maximum allowed range varying between 60-100 km (§6.5).

Bandwidth. Between any two towers, using very efficient encoding (256 QAM or higher), wide frequency channels, and radio multiplexing, a data rate of about 1 Gbps is achievable [34]. This bandwidth is vastly smaller than for fiber, and necessitates a hybrid design using fiber and MW.

Geographic coverage. Connecting individual homes directly to such a MW network would be cost-prohibitive. To maximize cost-efficiency, we focus on long-haul connectivity, with the last mile being traditional fiber. At short distances, fiber’s circuitousness and refraction are small overheads.

Cost model. We rely on cost estimates in recent work [45] and based on our conversations with industry participants involved in equipment manufacturing and link provisioning. The cost of installing a bidirectional MW link, on existing towers, is approximately \$75K (\$150K) for 500 Mbps (1 Gbps) bandwidth. The average cost for building a new tower is \$100K, with wide variation by terrain and across cities and rural areas. Any additional towers needed to augment bandwidth for particular links incur this “new tower” cost.

The operational costs comprise several elements, including management and personnel, but the dominant operational expense, by far, is tower rent: \$25 – 50K per year per tower. We estimate cost per GB by amortizing the sum of building costs and operational costs over 5 years.

3 cISP DESIGN

At an abstract level, given the tower and fiber infrastructure, a set of sites (*e.g.*, cities, data centers) to interconnect, and a traffic model between them, we want to select a set of tower-level connections that minimize network-wide latency while adhering to a budget and the constraints outlined in §2. Our approach comprises the following three broad steps.

- (1) Identifying a set of links that are likely to be useful by determining, for each pair of sites (s, d), the best feasible tower-level connectivity, if s and d were to be directly connected by a series of towers.
- (2) Building all $O(n^2)$ direct links, connecting each site to every other, would be prohibitively expensive. Thus, a subset of site-to-site links, together with existing fiber conduits, form our network. Choosing the appropriate subset is the key algorithmic problem.
- (3) Provisioning capacity beyond 1 Gbps along any link involves building additional tower-level links, *e.g.*, by identifying and using links that are also nearly shortest paths, but were omitted in step 1 above.

3.1 Step 1: Feasible hops

We first use line-of-sight and range constraints to decide which tower pairs can be connected. Achievable tower-to-tower hop length is limited primarily by the Earth’s curvature. MW hops must clear this curvature and any obstructions in an ellipsoidal region between the sender and the receiver antennae called the Fresnel zone with width h_{Fres} . The Earth’s curvature can be treated as a “bulge” of height h_{Earth} that a straight-line path must clear. At the midpoint of a hop of length D , using a MW frequency f , we have:

$$h_{\text{Fres}} \approx 8.7 \text{ m} \left(\frac{D}{1 \text{ km}} \right)^{1/2} \left(\frac{f}{1 \text{ GHz}} \right)^{-1/2}$$

$$h_{\text{Earth}} \approx \frac{1 \text{ m}}{50 K} \left(\frac{D}{1 \text{ km}} \right)^2$$

where K accounts for atmospheric refraction [49]. Towers should clear the sum of these heights and any other obstructions. In favorable weather, and with adequately large dish antennae, ranges of up to $D \approx 100$ km are achievable at high availability, provided such line-of-sight clearance [61]. As a specific example, the FCC licensing database [22] indicates that McKay Brothers, LLC (a provider for the financial industry) operated a $D = 96$ km hop from Chicago, IL (lat. 41.88° , lon. -87.62°) to Galien, MI (lat. 41.81° , lon. -86.47°) as part of a 1183 km MW relay. This shows that multipath interference issues (associated in this case with a traversal over Lake Michigan) are not an impediment to hop viability.

We assess hop feasibility between each pair of towers by using terrain data made available by NASA [54], which includes buildings and ground clutter, and effectively incorporates the height of the tree canopy.¹ We also require a fully clear Fresnel zone, and adopt $K = 1.3$ and $f = 11$ GHz in the above formulae. We have used our hop engineering routines to design line-of-sight networks, at least 4 of which are now deployed, including ultra-low latency routes between data centers hosting financial market matching engines. Our methodology has routinely provided correct clearance assessments when the physical paths are flashed. It is relatively rare that the hop feasibility assessment is inaccurate; if a problem arises, it is most likely that the locations themselves are not available to rent. In §6.5, we explore relaxations of the tower rental assumptions.

After identifying feasible tower-to-tower hops, for each pair of sites, we find the shortest path through a graph containing these hops, which we call a *link*. In line with observations from the tower data around major population centers, we assume each site itself hosts enough towers to use as the starting point for connectivity from that site to many others.

¹This NASA data set combines data from the Shuttle Radar Topography Mission (SRTM) [54] and the National Elevation Database (NED) [66], and typically yields acceptably small error (~ 2 m) against reference, high-accuracy LIDAR measurements.

3.2 Step 2: Topology design

Picking a subset of these site-to-site links involves solving a typical network design problem. The Steiner-tree problem [29] can be easily reduced to this problem, thereby establishing hardness. However, standard approximation algorithms, like linear program relaxation and rounding, yield sub-optimal solutions, which although provably within constant factors of optimal, are insufficient in practice. We develop a simple heuristic, which, by exploiting features specific to our problem setting, obtains nearly optimal solutions.

Inputs: Our network design algorithm requires:

- A set of sites to be interconnected, v_1, v_2, \dots, v_n .
- A traffic matrix H specifying the relative traffic volume $h_{ij} \in [0, 1]$ between each pair v_i and v_j .
- The geodesic distance d_{ij} between each v_i and v_j .
- The distance along the shortest, direct MW path between each pair, m_{ij} , as well as its cost, c_{ij} . This is part of the output of step 1.
- The optical fiber distance between each pair, o_{ij} , which we multiply by 1.5 to account for fiber’s higher latency.
- A total budget B limiting the maximum number of bidirectional MW links that can be built.

Expected output: The algorithm must decide which direct MW links to pick, *i.e.*, assign values to the corresponding binary decision variables, x_{ij} , such that the total cost of the picked links fits the budget, *i.e.*, $\sum_{ij} x_{ij} c_{ij} \leq B$. Our objective is to minimize, per unit traffic, the mean stretch, *i.e.*, the ratio of latency to c -latency, where c -latency is the speed-of-light travel time between the source and destination of the traffic.

Problem formulation: Expressing such problems in an optimization framework is non-trivial: we need to express our objective in terms of shortest paths in a graph that will itself be the *result*. We use a formulation based on network flows.

Each pair of sites (v_s, v_t) exchanges h_{st} units of flow. To represent flow routing, for each potential link ℓ , we introduce a binary variable $f_{stij,m}$ which is 1 *iff* the $v_s \rightarrow v_t$ flow is carried over the microwave link $v_i \rightarrow v_j$, and a binary variable $f_{stij,o}$ which is 1 *iff* the same flow is carried over the optical link² $v_i \rightarrow v_j$. The objective function is:

$$\min \sum_{s,t} \frac{h_{st}}{d_{st}} \sum_{i,j} (o_{i,j} f_{stij,o} + m_{i,j} f_{stij,m}) \quad (1)$$

The h_{st} term achieves our goal of optimizing *per unit traffic*. The $\frac{1}{d_{st}}$ term achieves our goal of optimizing the *stretch*.

For brevity, we omit the constraints, which include: flow input and output at sources and sinks; flow conservation; total budget; and the requirement that only links that are built ($x_{ij} = 1$) may carry flow. All variables are binary, so

²A “link” between sites can use multiple physical layer hops, both for MW and fiber. The underlying multi-physical-hop distances are already captured by the inputs o_{ij} and m_{ij} so the optimization views it as a single link.

flows are “unsplittable” (carried along a single path) and the overall problem is an integer linear program (ILP).

Note that we have decomposed the problem so that link capacity is *not* a constraint in this formulation: MW links will be built with sufficient capacity in step 3; fiber links are assumed to have plentiful bandwidth at negligible cost relative to MW costs. As a result, the objective function will guide the optimizer to direct each $v_i \rightarrow v_j$ flow along the shortest path of built links, which is the direct MW link $v_i \rightarrow v_j$ if it happens to be built, or otherwise, a path across some mix of one or more fiber and MW links.

Solution approach: As we shall see, simply handing the ILP to a solver did not scale to beyond medium-sized networks. By exploiting our problem structure, however, we develop a simple heuristic that yields near-optimal results at smaller scales (verified against the exact ILP solution) and can solve the problem at the larger scales of interest.

The first observation we make is that a large number of variables in our formulation will never take non-zero values, allowing us to eliminate them and any resulting null constraints. Roughly stated: if, for a particular (v_s, v_t) pair, a microwave path is of higher latency than a fiber path (which we can always use, at zero expense), then it will never carry $v_s \rightarrow v_t$ flow, though other flows may still traverse it. Similar observations apply to individual “distant, off-path” fiber and MW links. This simple observation substantially reduces the problem size. Note that standard network design problems do not typically have this structure available. This is entirely due to the hybrid design using fiber, which is assumed to be cheap, where available. We benefit, in this case, from having an “oracle” that tells us *a priori* when certain flow assignments are “obviously bad” and will not be useful. Further, carefully defined, such constraints preserve optimality; this part of our solution is not an approximation.

Second, we use a fast greedy heuristic to prune out MW links that are unlikely to be chosen. The heuristic operates using a larger budget ($2\times$ in our implementation) than we are ultimately allowed. In each iteration, we add to the solution the MW city-to-city link that decreases average stretch the most, continuing until the total cost reaches the inflated budget; the chosen links are candidates given to the ILP. Intuitively, the other links are uninteresting – they are unlikely to be picked in the final optimization even when a substantially larger budget is available, and so are not presented as options to the ILP. This approach does not provide any guarantees, but we find that on small problem sizes, where the exact ILP can also be evaluated, it obtains the optimal solution.

3.3 Step 3: Augmenting capacity

In many scenarios, certain links require more capacity than a single MW link provides. For short physical distances, this is a non-issue: the MW link can simply be replaced by fiber

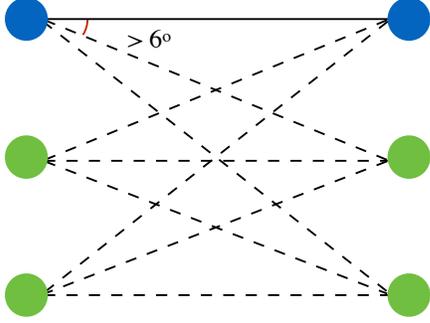


Fig. 1: Bandwidth augmentation: k^2 hops with $O(k)$ new towers.

without a large impact on the network’s average latency. However, for longer distances, this is not acceptable.

One approach to resolving this problem is simply to build multiple parallel MW links, over multiple series of towers. While tower siting is often a challenging practical problem, with individual sites valued by the HFT industry at as much as \$14 million [10], in the cISP context, there is a much larger “tolerance” than in HFT, where firms compete for fractions of microseconds. For a 500 km long cISP link, the midpoint diverging 10 km from the geodesic would increase latency by a negligible 0.2%. Thus, the problem of tower siting is substantially simpler. Also, in many cases, tower infrastructure is dense enough already to allow multiple parallel links. For instance, the HFT industry operates nearly 20 parallel networks in the New York-Chicago corridor [45].

We can also employ a simple trick to enhance the effectiveness of parallel series of towers, as shown in Fig. 1. Instead of k parallel series of towers providing merely a $k\times$ bandwidth improvement, connecting multiple antennae on each tower to other towers, we can obtain a $k^2\times$ improvement. Using antennae with overlapping frequencies requires an angular separation of 6° [49], as shown in Fig. 1. Again, the stretch caused by the resulting gap between parallel series of towers is small. For a tower-tower hop distance of 100 km, the minimum distance between two parallel towers should be $100 \cdot \tan(6^\circ) = 10.6$ km, which, as noted above, has a small effect on end-to-end latency for long links.

This approach implies that for site-to-site bandwidths under 1 Gbps, we need just one series of towers; for bandwidths between 1-4 Gbps, we need 2 series; for 4-9 Gbps, 3; etc. While tower siting circumstances are often unique, we are aided by two observations: (a) there is substantial redundancy in existing tower infrastructure, and we can often find existing towers for parallel connections (see Fig. 4(b) and the related text in §4); and (b) when new towers are needed, there is substantial tolerance in where they are sited, as noted above. Bandwidth may potentially be increased even further through spatial diversity techniques, whereby multiple antennae are placed appropriately on the same tower such that they can adaptively cancel interference by multiple transmission streams within the same frequency channel [67].

3.4 Generality

Note that the above outlined approach applies broadly across other line-of-sight media, such as free-space optics and millimeter wave networking. Multiple technologies, beyond the mix of fiber and MW that we consider, can be easily incorporated into this framework, preserving relevance with technology evolution. Such line-of-sight free-space networking seems, for the near future, to be the only cost-effective solution for achieving nearly c -latency on the Internet.

4 A cISP FOR THE UNITED STATES

We now support the above discussion of our abstract framework with a concrete instantiation: designing a cISP for the U.S. mainland. To assess line-of-sight connectivity between existing towers, we use fine-grained data on tower infrastructure, buildings, terrain, and tree canopy. The fiber conduit data is available from past work [25].

Defining the sites and traffic model: To maximize utility while keeping costs low, we connect only the 200 most populous cities in the contiguous United States. In addition, we coalesce suburbs and cities within 50 km of each other, ending up with 120 population centers. (Henceforth, when we refer to “cities”, we refer to these population centers.) Based on population data for 2010 [15], we calculate that 85% of the US population lives within 100 km of these 120 cities. For the traffic matrix H , we set h_{ij} proportional to the product of the populations of cities v_i and v_j .

Step 1: Which city-city links are feasible? We use existing towers listed in FCC’s Antenna Structure Registration [26] and databases from American Towers, Crown Castle, and several other tower companies for which we were able to download data. We cull these rather large databases of MW towers to a subset of 12,080 towers as follows: Towers from rental companies are typically suitable for use. From the FCC database, we only use towers over 100 m height. When tower-density exceeds 50 towers per 0.5° square grid cell, we randomly sample towers. (Using all towers could only improve our results, but increases compute time.)

Evaluating link feasibility across tower pairs within range of each other using the aforementioned NASA data [54], we find 261,019 tower-tower hops that satisfy line-of-sight constraints. We find that each city itself has large numbers of suitable towers in its vicinity. We run a shortest path computation on a graph comprising the cities and towers and city-tower and tower-tower hops to find the shortest city-city MW links. This yields both the cost (*i.e.*, number of towers) and latency (*i.e.*, distance along the chosen series of towers) for each city-city link.

For fiber distances, we compute the shortest paths over the InterTubes [25] dataset on US fiber conduits.

Step 2: What subset of links should we build? With the inputs now ready, we can run the algorithms of §3.2 for any

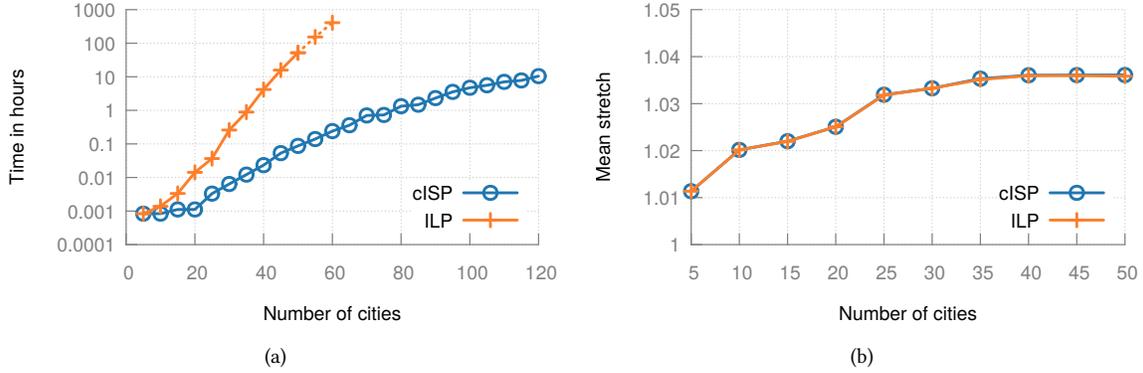


Fig. 2: *cISP's* design method is fast-enough and near-optimal: (a) *cISP* generates an optimized topology within hours for 120 cities while the ILP does not yield a result even after 2 days for more than 50 cities. For the ILP, runtimes for 50+ cities are extrapolated by curve fitting. (b) The stretch achieved by *cISP* matches that of the ILP to two decimal places for instances that can be optimized by the ILP.

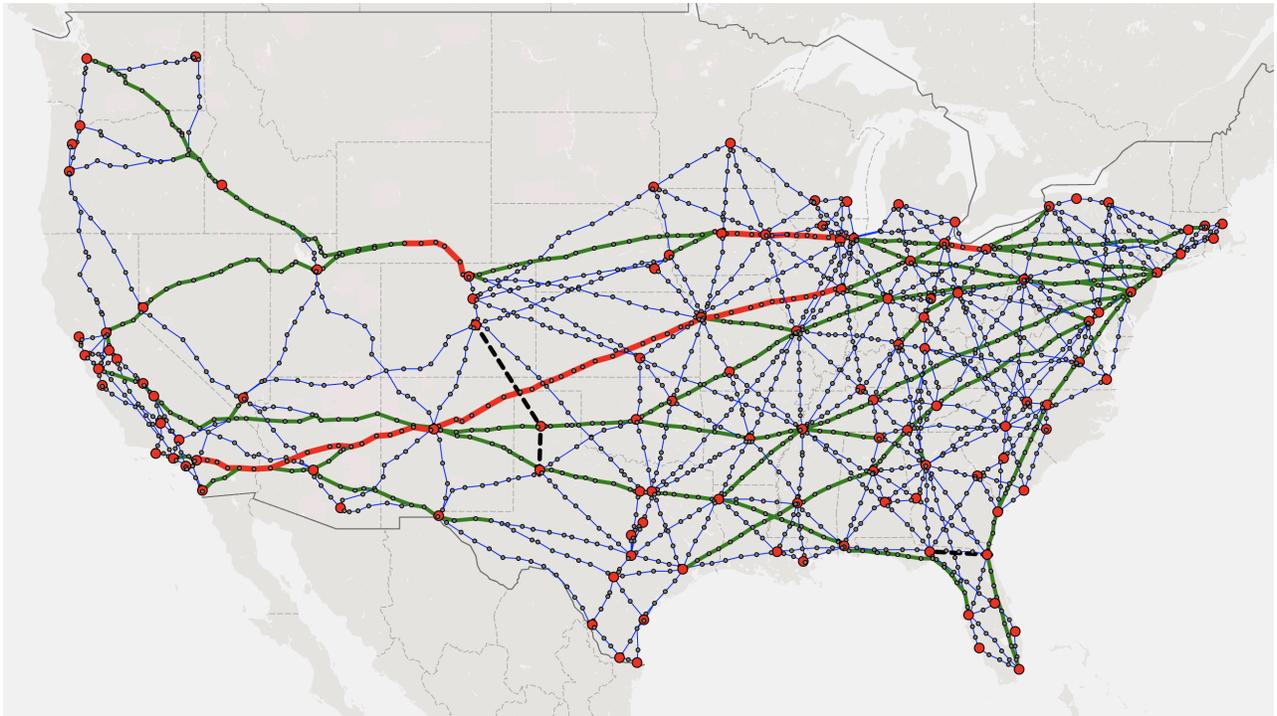


Fig. 3: A 100 Gbps, 1.05 \times stretch network across 120 population centers (big, red) in the US. Blue links (thin) need no additional towers. Green links (thicker) and red links (thickest) need 1 and 2 series of additional towers, respectively. The black dashed links represent fiber paths.

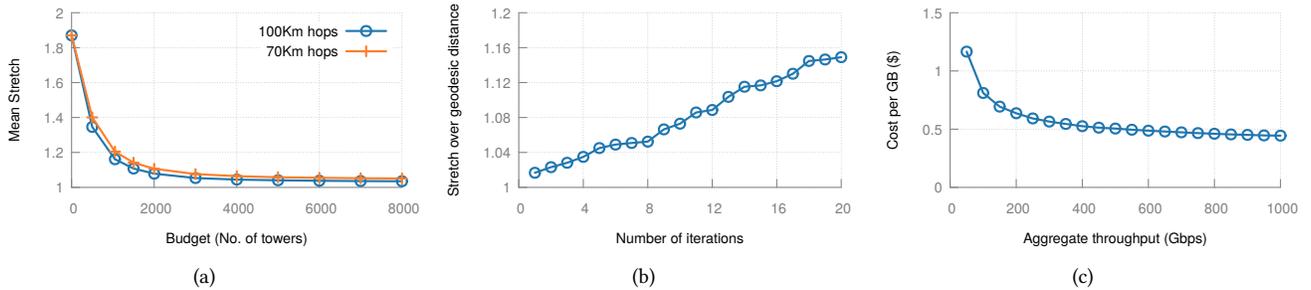


Fig. 4: (a) Network stretch reduces as we add more MW towers. (b) Stretch for 20 shortest tower-disjoint purely MW paths along the long red Illinois-California link in Fig. 3. (c) Cost per GB for the city-city traffic model decreases with increasing aggregate throughput.

given budget to obtain a set of city-city MW links to build. We use the Gurobi solver [32] for this purpose.

First, as we show in Fig. 2(a), the exact ILP, without using our observations on the problem structure, is too computationally inefficient to scale to this scenario. We use subsets of all 120 cities to assess scalability, with the budget proportional to the number of cities in each test, with a budget of 6,000 towers at the largest scale. Even after 2 days of compute, the exact ILP was unable to obtain a result for sets of cities larger than 50. In contrast, our cISP design heuristic is able to solve the problem at the full scale. Second, as Fig. 2(b) shows, at small scales, where we can also run the exact ILP, our heuristic yields the optimal result. We also tested a linear program rounding approach, but even the naive LP relaxation followed by rounding did not scale beyond 60 cities, and gave results worse than optimal.

Fig. 3 shows an example network. Designed with a budget of 3,000 towers and maximum hop length of 100 Km, its average latency is $1.05\times c$ -latency. Fig. 4(a) shows the reduction of the network’s stretch with increases in budget for maximum hop lengths of 70 and 100 Kms. Given the similarities with 70 and 100 Kms, hereon, we only present results for the latter. An animation, showing how the network structure evolves from mostly-fiber to mostly-MW as the budget increases, is available online [20].

Step 3: Augmenting capacity: We produce a target aggregate demand (*i.e.*, the sum of all site-site traffic demands) by scaling the traffic matrix H . Then, each tower-tower MW hop that would be over-utilized (given the routing of §3.2 and the 1 Gbps capacity from §2) is augmented with additional towers at each end, as described in §3.3. Fig. 3’s topology, when provisioned for an aggregate throughput of 100 Gbps, has 1,660 tower-tower hops that use only already built towers seen in tower databases, while 552 hops need one additional new tower at each end, and 86 hops need 2 additional towers at each end. Using the cost model described in §2, we find that the cost per GB for this topology, with latency within $1.05\times$ and 100 Gbps throughput, is \$0.81. For some context, this is $\sim 10\times$ the cost per GB for content delivery networks [51].

Provisioning even more bandwidth would require more new towers. For 1 Tbps, some tower-tower hops would need as many as 8 additional towers at each end. This is not infeasible – latency would not be inflated excessively, and towers could be found or built. In fact, for the long red link in the map in Fig. 3, which spans 2,700 km from Illinois to California, we find that the *longest* of these 8 additional series of towers would be only 5% longer than the shortest MW path, incurring a stretch of 1.07, instead of 1.02.

We can extend this argument even further: for the same Illinois to California link, we compute tower-disjoint shortest paths, *i.e.*, after finding the shortest path, we remove all towers used by it, find the next-shortest tower-path, etc. In this process, we use only existing towers from our databases, and adhere to the same link feasibility constraints. Fig. 4(b)

shows that stretch increases gradually as we keep eliminating towers; nevertheless, even after 20 such iterations, stretch is much smaller (1.15) than with the existing fiber conduit (1.75). Note that this route runs through the Rocky mountains and other areas of low tower density. Thus, in accounting for the cost of bandwidth augmentation entirely using the (higher) cost estimates for building new towers, we are substantially overestimating the expense.

There is also another reason our costs are over-estimates: at sufficiently high bandwidth, there is a better option than building many parallel long-distance MW links: one could use the same number of towers to construct a single line of towers with shorter tower-tower distances. This can make shorter-range, but higher-bandwidth technologies like MMW or free-space optics, more cost-effective.

Despite the above two factors, we use parallel MW towers, with all the required additional towers accounted for as new towers, to provide conservative cost-estimates as aggregate bandwidth increases in Fig. 4(c).

5 ROUTING & QUEUING

The HFT industry’s point-to-point MW deployments demonstrate end-to-end application layer latencies within 1% of c -latency, after accounting for all delays in microwave radios, interfacing with switching equipment and servers, and application stacks. Such low latencies across point-to-point long-distance links place sharp focus on any latencies introduced at routers for switching, queuing, and transmission.

Internet routers can forward packets in a few tens of microseconds, and specialized hardware can hit $100\times$ smaller latencies [39]. Transmitting 1500 B frames at 1 Gbps takes $12\ \mu s$. Thus forwarding and transmission even across many long-distance links incur negligible latency. Longer routes and queuing delays, however, can have substantial impact.

To assess the impact of routing and queuing in cISP, we use ns-3 [58]. We use UDP traffic with a uniform packet size of 500 bytes. We use the built-in FlowMonitor [14] to measure delay and loss rate, and add a new monitoring module to track link-level utilization. All experiments simulate 100 Gbps of network traffic for one second of simulated time. An experiment takes approximately 10 hours to complete on a single core of a 3.1 GHz processor. Even achieving this running time requires some compromises: we aggregate the bandwidth of parallel links and remove the individual tower hops to focus on network links between the routing sites.

Routing schemes: Besides ns-3’s default shortest path routing, we implement two other schemes – throughput optimal routing, and routing that minimizes the maximum link utilization, a scheme commonly employed by ISPs [42].

Results: When the traffic and routing match the design target, *i.e.*, the population-product traffic routed over shortest

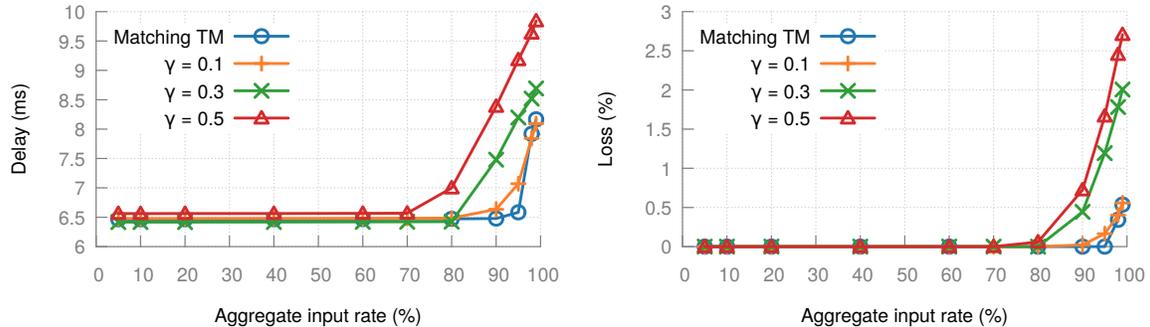


Fig. 5: Average delay (left) and loss rate (right) remain consistent across perturbations of the city-city traffic model, except under heavy load.

paths, we find that the network can be driven to high utilization (95%) with near-zero queuing and loss. Non-shortest-path routing schemes needlessly compromise on latency in such scenarios. (Plots for this easy scenario are omitted.)

We also test the network’s behavior under deviations from the designed-for traffic model. We emulate scenarios where a city produces more or less traffic than expected by allowing, for each city, a “population perturbation” — each city’s population is re-weighted by a factor drawn from the uniform distribution $U[1 - \gamma, 1 + \gamma]$ for a chosen $\gamma \in [0, 1]$.

Fig. 5 shows the results for $\gamma \in \{0.1, 0.3, 0.5\}$. Even for large perturbations, the mean delay does not increase by more than 0.1 ms and the loss rate is zero up to an aggregate load of 70% of the capacity designed for, even with just shortest path routing. Other routing schemes are indeed more resilient to higher load, achieving virtually zero loss and queuing delay even at high utilization, but at the cost of latency. For the tested topology, both the alternative routing schemes incur 10% higher latency on average (not shown in the plots). These results indicate there would be significant value in work that reduces the amount of over-provisioning required by making modest compromises on latency on some routes, e.g., as in [33].

Speed mismatch: The bandwidth disparity between the network core and edge for cISP may seem atypical, in the sense that in most settings, the core has higher bandwidth links compared to the edge, while in cISP, edge links (such as those at large data center end points) may often have much higher line rates when they feed their outgoing traffic into cISP. Thus, we also evaluate if this “speed mismatch” causes persistent congestion at cISP’s ingress.

We run ns-3 simulations with several sources (S_i) connected to a sink (D) through the same intermediate node (M). The M - D link rate is fixed at 100 Mbps. We then evaluate settings with every S_i - M link being either 100 Mbps or 10 Gbps. The former is the control, and the latter is the setting with a speed mismatch. M has an unbounded queue. Ten sources send 100 KB TCP flows (small, as is expected in cISP) to the sink, D . The arrival of these TCP flows follows a Poisson process, consuming on average 70% of the I - D link’s

bandwidth. Each simulation run lasts 10 s and we conduct 100 such runs. We test TCP both with and without pacing.

Fig. 6(a) shows that the median queue occupancy at M is higher without pacing, especially at the 95th percentile. However with pacing, queueing behavior is nearly the same. The median flow completion times (Fig. 6(b)) are unaffected both with and without pacing.

6 PRACTICAL CHALLENGES

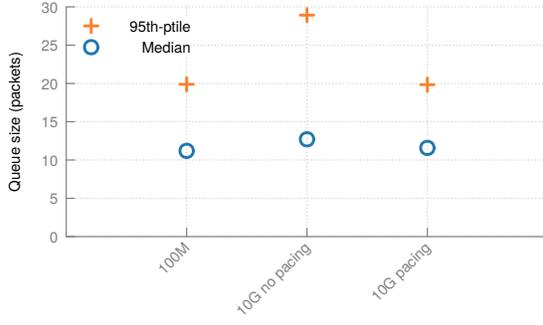
Deploying cISP would involve several practical challenges beyond network design and routing, which we now address.

6.1 Impairments due to weather

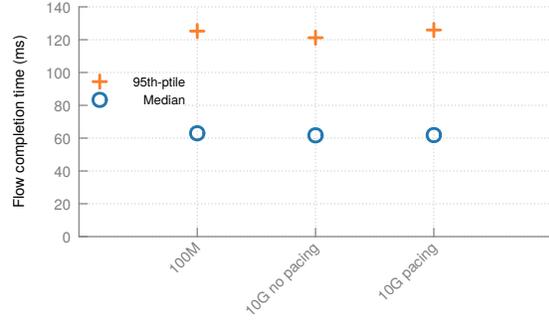
Precipitation causes MW signal attenuation. We use standard equations in MW engineering [38] to calculate attenuation. While the physical layer could trade link bandwidth for higher resilience to weather, we treat the impact of precipitation in a binary manner: if attenuation exceeds a threshold that would degrade bandwidth, we conservatively consider a link to have failed.

We assume that when a link fails, traffic is shifted to the shortest available route, which may use any combination of MW and fiber. The high precipitation that causes failures is easy to predict, especially on the timescale of minutes. Thus, even slow, centralized management would suffice to anticipate failures and reroute accordingly.

We use NASA’s precipitation data [53] to determine which links are down when, and what the impact of such failures is on the network’s latency. For each day over a period of a year (July 2015 - June 2016), we select a 30-minute interval uniformly at random, and identify the links that would fail during it. We then evaluate the latency for each pair of cities end-to-end for each interval. Fig. 7 shows that 99th-percentile latencies are nearly the same as the best, fair-weather latencies. In terms of the median across city-pairs, even the worst latencies over the year are 1.7 times lower than those over fiber. Large increases in latency due to weather typically occur only between nearby city-pairs, the fiber route to which runs through a farther-away city, e.g., in Texas, Austin and Killeen fall back to a fiber route through Fort Worth. A more



(a)



(b)

Fig. 6: TCP pacing addresses the problem of capacity mismatch (a) by reducing persistent queuing (b) without affecting flow completion times.

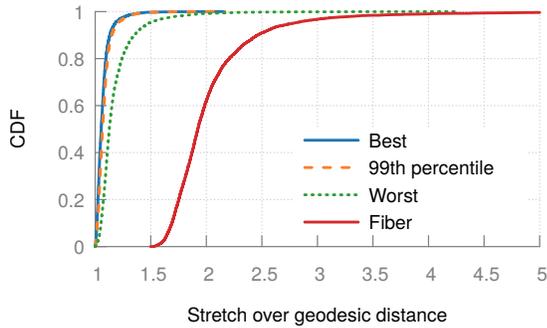


Fig. 7: Stretch across all city-pairs over a year due to precipitation. The 99th-percentile stretch is comparable to the best stretch.

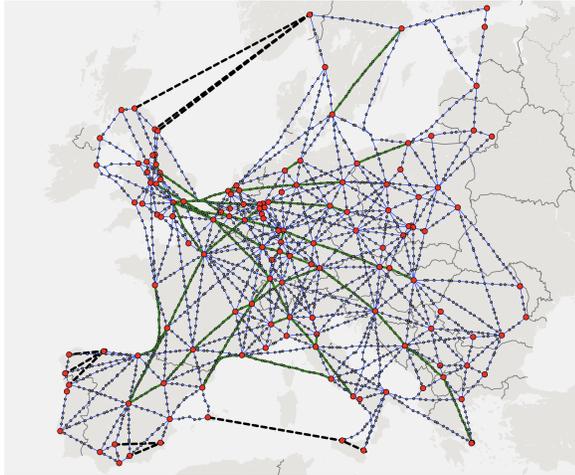


Fig. 8: A 100 Gbps 1.04 \times stretch cISP across Europe. This network uses several fiber connections (dashed, black lines).

sophisticated analysis allowing dynamic link bandwidth adjustment rather than binary failures can only improve these numbers. Thus, even under significantly adverse weather, most of the latency advantage of cISP remains intact.

We have also created an animated visualization of the network’s latency evolving over a year’s weather [18].

6.2 Is the US geography special?

So far, we have limited our analysis to the contiguous U.S. It is reasonable to ask: are the population distribution and geography of the U.S. especially amenable to this approach, or is it applicable more broadly? The availability of high-quality tower data and geographical information systems data for the U.S. enables a thorough analysis. While similar data is, unfortunately, not available to us for other geographies, we can approximately assess the design of a cISP in Europe using public, crowd-sourced data on cellular towers [65]. Lacking fiber conduit data, we assume that fiber distances between cities are inflated over geodesic distance in the same way as in the US ($\sim 1.9\times$). Using our methodology in §3, we design a European cISP of similar geographical scale across cities with population more than 300k, targeting the same aggregate capacity and mean latency ($1.04\times$ here vs. $1.05\times$ for cISP-US). The cost of this design, shown in Fig. 8, is similar as well, with $\sim 3k$ towers. Note that the impact of Europe’s higher population density is not seen here, because we explicitly design for the same aggregate throughput. One could, alternatively, normalize throughput per capita, and compare cost per capita, to obtain similar results.

Admittedly, there is not yet a known approach to bridging large transoceanic distances using MW, limiting our approach to large contiguous land masses that need to be interconnected with fiber. In the distant future, LEO satellite links, hollow-core fiber, or even towers on floating platforms may be of use for such connectivity.

6.3 Is the city-city traffic model special?

So far, our results have used the city-city population-product based traffic model. Ideally, we would be able to use wide-area traffic matrices from some ISP or content provider for modeling. In the absence of such data, we focus on showing that cISP can be tailored to vastly different deployment scenarios and their corresponding traffic models. Apart from the city-city population product model, we use (a) traffic between a provider’s data centers; and (b) traffic between the cities and data centers.

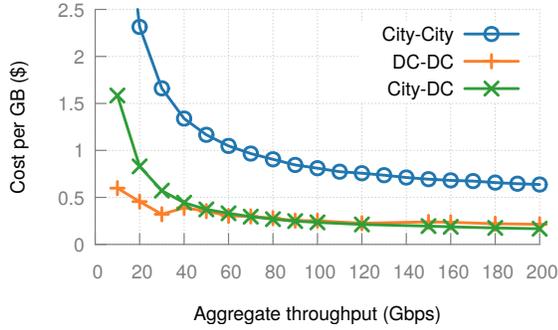


Fig. 9: Cost per GB for different traffic models: the City-City model, discussed in the most detail, is the most expensive.

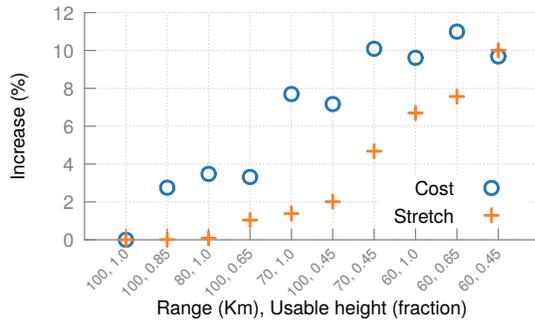


Fig. 10: As constraints on tower space and range become tighter, the network becomes more expensive, and stretch increases.

An inter data center cISP: We use Google data centers as an example, considering all 6 publicly available US locations - Berkeley, SC; Council Bluffs, IA; Douglas County, GA; Lenoir, NC; Mayes County, OK; and The Dalles, OR. In the absence of known inter-data center traffic characteristics, we provision equal capacity between each DC-pair.

Data centers to the edge: We also model a scenario where data centers are to be connected to edge locations in cities. Each of the 120 cities connects to its closest Google data center, with traffic proportional to its population.

We show in Fig. 9 that using the same design approach as in §3, both of the above scenarios result in networks with lower cost than the city-city model. Thus, cISP can be tailored to a variety of use cases and traffic models.

6.4 Traffic model mismatches

A cISP may carry a mix of city-city, inter-DC, and DC-edge traffic. How does its performance degrade as the *proportion* of these traffic types departs from the design assumptions?

We design a cISP to carry an aggregate of 100 Gbps with a city-city : DC-edge : inter-DC traffic proportion of 4:3:3. Using ns-3 simulations similar to those in §5, we then test this network under several traffic mixes different from this designed-for mix — 5:3:3, 4:3:4, and 4:4:3.

Fig. 11 shows that there is a difference of less than 0.05ms in mean delay across different combinations of traffic matrices up to an aggregate load of 70% of the design capacity. Similarly, loss remains nearly 0 until this load. The decrease in delay at high load (4:4:3 for $x > 90$ in Fig. 11) is due to losses, which are likelier on longer, higher-delay paths.

Mean delay depends more on city-city traffic, as expected: city-city traffic requires a wider infrastructure footprint, and deviations from its design parameters have greater impact.

Thus, as discussed in §5, significant traffic model deviations can be absorbed using some over-provisioning, in line with current ISP practices.

6.5 Tower height and availability

Our initial design assumed a MW hop to be feasible if it spans a distance of 100 km or less, and satisfies line-of-sight constraints using the tops of the towers. In practice, however, a tower chosen for a route might not have a free spot for a new antenna at the necessary height, especially at the top, where structural concerns for large parabolic antennae are greatest, and where access and maintenance can be problematic. Further, for smaller antennae, insufficient gain margins can decrease the 100 km maximum range. Hence, we evaluate cost and latency of the network with hop-level restrictions modeling these effects.

We test the impact of restricting usable height on towers to three levels, as a fraction of tower height: 0.85, 0.65, and 0.45. Testing for line-of-sight visibility with these restrictions eliminates more towers than using tower tops. We also vary the maximum range, which can necessitate the use of a larger number of towers, thus increasing the cost and potentially making some city-pairs infeasible to connect using MW.

We assess the *percentage increase* in cost and stretch values compared to the baseline values with 100 km range and using the tower tops, *i.e.*, height fraction = 1. Fig. 10 shows the results for different combinations of the range and antenna-height constraints, sorted by lowest to highest stretch. The maximum increase in cost is 11% (with the absolute cost per GB under these constraints being \$0.90), while the maximum increase in stretch is 10% (with the absolute stretch compared to the geodesic being 1.16). Thus, even substantial potential problems in tower siting and mounting antennae do not change our overall conclusions about the viability of cISP.

In our experience, assessments like those in this work yield accurate estimates of the latency (especially for tolerances larger than in the HFT industry) and the number of tower-tower hops that will ultimately be used to connect two sites (and hence accurate cost estimates). The precise set of towers often differs based on real-world constraints, particularly tower unavailability for structural and rental-related reasons. Thus, while accurate in terms of cost and latency, this work does not provide fully engineered routes. In practice, to improve accuracy in preparation for building a MW route, we assign each tower in a swathe connecting the

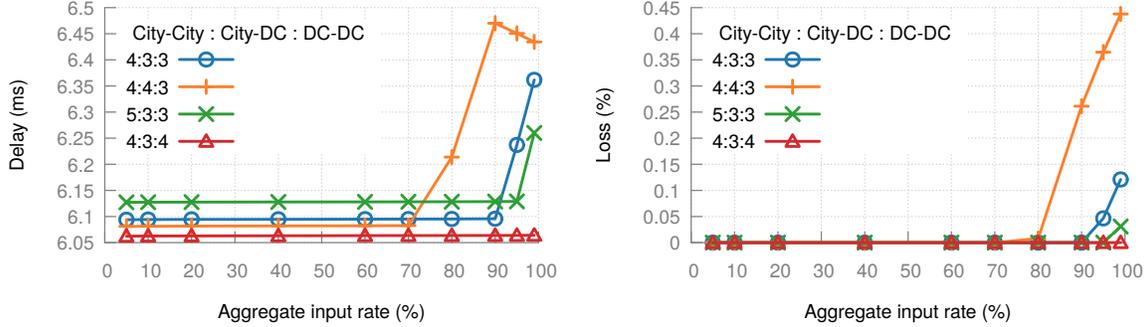


Fig. 11: Average delay (left) and loss rate (right) remain consistent across deviations from the designed-for traffic mix, except under heavy load.

sites an acquisition probability, which depends on a number of factors (e.g., tower type, ownership, location). Further, for towers that can be acquired, we use a uniform distribution to model height at which space for antennae is available. With this probabilistic model, we compute thousands of candidate MW paths between site pairs, with refinements as acquisitions and height availabilities are confirmed. We make available in video form [19] an example of such refinement.

6.6 Integration into the Internet

We next discuss potential problems cISP may face in terms of integration into the present Internet ecosystem.

Low-hanging fruit: The easiest deployment scenarios involve one entity operating a significant network backbone:

- A CDN could use cISP for “back-office” traffic between its locations and content origins, which is often part of latency-sensitive user-facing interactions [60].
- Content-providers like Google and Facebook can benefit from cISP – such WAN designs already accommodate distinctions between latency-sensitive and background traffic [37, 40].
- Purpose-built networks such as for gaming [28] can easily use cISP between their edge locations and servers.

All of these are interesting and economically viable use cases with minimal deployment barriers, and each *alone* may justify a design like cISP. For instance, while it is tempting to dismiss gaming as a niche, it is a large and growing market: the Steam gaming platform claims up to 16 million players Worldwide, with 17% of their traffic being US-based [64]. At a 10 Kbps rate per player,³ this aggregates to 27 Gbps – enough to make cISP viable in this setting. (We present cost-benefit estimates, including for gaming, in §8.)

User-facing deployment: Access ISPs may use cISP as an additional provider, and incorporate a low-latency service into their broadband plans.⁴ Utilizing cISP in this manner

³See measurements for popular online games in [21].

⁴While large last-mile latencies can overshadow cISP’s low latency, this is an entirely orthogonal problem, on which significant progress is being made – 5G prototypes are already showing off sub-millisecond latencies [35].

can help ISPs to provide and meet the requirements of demanding Service Level Agreements, the case for which was made in recent work [9]. ISPs may use heuristics to classify latency-sensitive traffic and transit it using cISP. Alternatively, software at the user-side may make more informed decisions about which traffic may use the fast-path exposed by the ISP. While this would require significant user-side changes, note that many of today’s applications already manage multi-modal WiFi and cellular connectivity.

7 A FEW POTENTIAL APPLICATIONS

Several applications require low latency over the wide area-network. Applications focused on user interactivity, such as augmented and virtual reality, tele-presence and tele-surgery, music collaboration over long-distances, etc., can all benefit from low-latency network connectivity. Likewise, less visible and user-centric applications, such as real-time bidding for Web page advertisements and block propagation in block-chains, would also benefit from a network like cISP. While it is beyond the scope of this paper to analyze this in significant detail, we assess, in simplified environments, the improvements cISP could achieve for two application areas.

7.1 Online gaming

We discuss cISP’s benefits for both models of online gaming: **thin-client** (where each client essentially streams everything in real-time from a server) and **fat-client** (where the client has an installation of the game, performs computations, etc., and only relies on the server for updates on the game state based on other players’ actions).

Fat-clients are dominant today, and are easy to tackle: communication is almost entirely composed of latency-sensitive player actions and game-state changes, and is low-volume, typically a few Kbps per client for popular games [21]. It can all be transferred over the low-latency network, reducing latency by 3-4× compared to today’s Internet.

Thin-client gaming is still in its infancy, as it depends heavily on the network, with data rates in Mbps. We explore the potential of a speculative execution approach: the server speculates on the game state and sends data for multiple

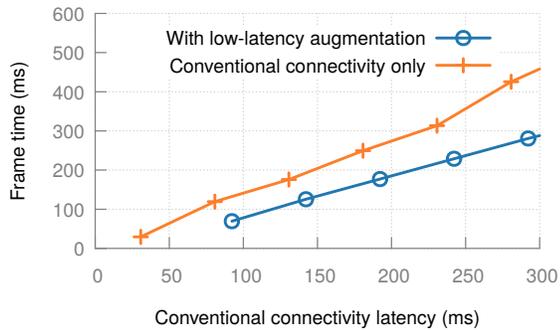


Fig. 12: A substantial reduction in frame time can be obtained by the use of a parallel low-latency augmentation to the present Internet.

speculated scenarios in advance over fiber, and then issues messages indicating which scenario occurred on the low-latency network. Such speculation has already shown success for rich games like “Doom 3” in prior work [46].

We use a toy thin-client for a multi-player Pacman variant to explore the latency benefit. Our rudimentary implementation speculates on all 4 movement directions possible as user input. In line with the online-gaming literature, we measure “frame-time”, which “corresponds to the delay between a user’s input and the observed output” [46]. We evaluate frame-time as latency over conventional connectivity increases (emulated by adding latency in software), and for a low-latency network always incurring 1/3 of the latency of the corresponding conventional network.

As Fig. 12 shows, the speculative approach enabled by the low-latency network augmentation substantially reduces frame-time. This comparison would improve further if non-network overheads from processing and rendering in our naive implementation were smaller. We do not use any significant graphics on which to evaluate the additional bandwidth overhead on fiber, but even in the sophisticated scenarios examined by prior work [46], this bandwidth overhead can be contained to 2-4.5 \times .

7.2 Web Browsing

We evaluate the potential impact of cISP’s latency improvement on Web page load times (PLTs) (based on the *onLoad* event [41]) using Mahimahi [56]. Our experiments use a uniform random sample of 80 Web sites from Alexa’s list of popular Web sites [6]. We replay each page with unmodified latencies (as a baseline) and with latencies reduced to 0.33 \times their original values. No bandwidth limitations are imposed.

Fig. 13(a) shows the results. Compared to the baseline, a 66% reduction in latencies (‘cISP’-line) results in a 31% decrease (an absolute decrease of 302 ms) in the median PLT. This PLT reduction is less than the 66% reduction in RTT, because loading a Web page also involves significant non-network activity. For fetching the individual objects comprising these pages, these overheads are smaller, and cISP’s improvements are larger. As shown in Fig. 13(b), for

the same 66% reduction in RTT, object load times decrease by 49%. Small objects (under 1460 bytes) show a reduction of 59%. Thus, with a faster network like cISP, the bottlenecks shift to the protocols and Web page design.

While Web-browsing traffic comprises only a small fraction of total Internet traffic,⁵ we can further reduce the load by carrying only client-to-server traffic on cISP. Hence, we extend Mahimahi to enable *selective* manipulation of RTTs in the replay, such that some traffic sees lower RTTs than other traffic. We then emulate scenarios where only client-to-server traffic is sent over cISP at a reduced latency, e.g., “cISP-selective” implies that only the client-to-server latencies are adjusted, and set to 0.33 \times the recorded latency. We assume that the unadjusted latencies are symmetrical in each direction. This approach yields a median improvement of 27% (265 ms) and requires sending only 8.5% of the bytes over cISP.

8 COST-BENEFIT ANALYSIS

The value of reducing Internet latencies is reflected in industry investments in this direction: Riot Games is operating its own wide-area backbone [28]; Zayo acquired faster fiber routes previously used exclusively for HFT, for broader use by “content, media and cloud providers” [36]; and CDNs routinely use overlay routing to cut latency for dynamic, non-cacheable content, for which edge replication is difficult or ineffective [3]. We nevertheless present *quantitative* lower-bound estimates of cISP’s value per bit in a variety of contexts and assess whether its expense is justified.

Web search: Putting together Google’s quantification of the impact of latency in search [13], their estimated search revenue restricted to the US [50], their search volume [31], estimated data transferred per search,⁶ and estimated cost per search [43], we estimate that speeding up page load times for 12 Gbps of their US search traffic by only 200 ms (400 ms) would yield an additional yearly profit of \$87 (\$177) million. This translates to an added value of \$1.84 (\$3.74) per GB.

E-commerce: Combining Amazon.com’s estimated number of visits, page fetches per visit, percentage of US traffic [62], and estimated page size, we arrive at an estimate of 483 PB of traffic per year. Using their US sales estimate [52] and North America profit margin of 4% [11] results in an estimated \$7.9 billion in profits per year. Estimates for the dependence of conversion rate on e-commerce Web sites on PLT vary from 1% [47] to 2.4% (on desktop) and 7% (on mobile) per 100 ms of additional latency [4]. Thus, a speedup of 100 ms could increase profits by \$78.7-\$551 million. If we can save 200 ms by sending less than 10% of the data over cISP (\$7.2), this translates to \$3.26-\$22.82 per GB.

⁵Cisco’s 2018 estimate puts “Web/Data traffic” at 13% [17] including non-latency sensitive traffic like software updates and some file transfers.

⁶From Firefox desktop’s network tools; mobile responses would be smaller.

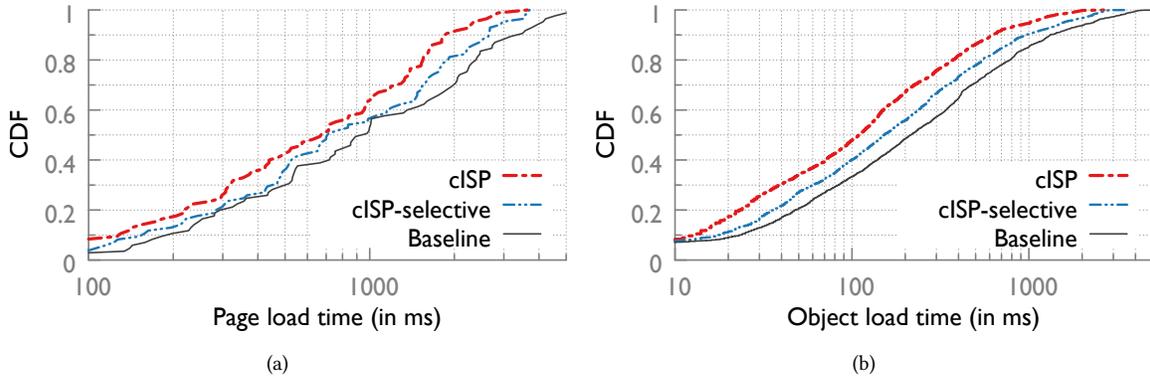


Fig. 13: Impact of latency reductions on (a) Web page load times and (b) individual (Web page) object load times.

Gaming: Online gamers often pay for “accelerated VPNs”, which promise to lower network latency (perhaps using overlays). Such services cost \$4-\$10 per client per month [1, 8, 59]. Full-time gaming at 8 hours a day at a 10 Kbps rate (as before in §6.6) translates to 1.08 GB / month. Thus, if cISP were priced like a cheap accelerated VPN service at \$4 / mo, this would translate to a value of at least \$3.7 / GB. A less aggressive model than “full-time gaming” would only improve cISP’s value. Note that cISP’s latency benefits are likely to be substantially larger than such VPN services.

Another indicator of latency’s value in gaming is the market for gaming monitors with high screen-refresh rates: the 6-10 ms of latency advantage is valued at over \$50 by many gamers, estimated from the pricing of monitors which are exactly the same except in terms of refresh rate [7].

The value per GB obtained from cISP’s latency reduction in above cases – \$1.84-\$3.74, \$3.26-\$22.82, and over \$3.7 – substantially exceeds its cost estimate of \$0.81 per GB. Even accounting for substantial over-provisioning leaves intact a clear economic argument for designs like ours. Upcoming application areas like virtual and augmented reality can only make this case stronger. We expect cISP’s most valuable impact to be in breaking new ground on user interactivity over the Internet, as explored in some depth in prior work [63].

9 RELATED WORK

While networking research has made significant progress in measuring latency, as well as improving it through transport, routing, and application-layer changes, the underlying infrastructure’s latency inflation has received little attention, and has been assumed to be an unresolvable given. This work proposes and analyzes a nearly speed-of-light ISP, demonstrating that this is far from the case.

There are several ongoing high-profile Internet infrastructure efforts, including Google’s Loon project [68], Facebook’s drones [44], and the satellite Internet push by OneWeb and WorldVu [5, 30]. These, however, are all addressing a different problem – expanding the Internet’s coverage. One

particularly noteworthy effort, from Alphabet’s X moonshot factory, is a network under deployment in an Indian state, based on free-space optics, and described as “a cost effective way to connect rural and remote areas across the state” [27]. Free-space networks of this type will likely become more commonplace in the future, and this work is further evidence that many of the concerns with line-of-sight networking can indeed be addressed with careful planning. Further, cISP’s design approach is flexible enough to incorporate a variety of media (fiber, MW, MMW, free-space optics, etc.) as the technology landscape changes.

To the best of our knowledge, the only efforts focused on wide-area latency reduction through infrastructural improvements are in niche scenarios, such as the point-to-point links for financial markets [45], and isolated submarine cable projects aimed at shortening specific Internet routes [55, 57].

10 CONCLUSION

A speed-of-light Internet not only promises significant benefits for present-day applications, but also opens the door to new possibilities, such as *eliminating the perception of wait time* in our interactions over the Internet [12].

We thus present a design approach for building wide-area networks that operate nearly at *c*-latency. Our solution integrates line-of-sight wireless networking with the Internet’s fiber infrastructure to achieve both low latency and high bandwidth. We use data on existing towers, and terrain and tree canopy, and a cost model reflective of current practice in engineering such networks to inform our design.

Apart from providing a near-optimal solution to the underlying network design problem, we also address numerous practical challenges, such as the availability of antenna space on towers, and assess latency degradation due to adverse weather, and deviations from the designed-for traffic model.

Lastly, our design’s value far exceeds its cost for applications we could compute estimates for. Thus, greatly reducing

the Internet’s infrastructural latency is not only tractable, but surprisingly cost-effective, and an exciting opportunity.

REFERENCES

- [1] AAA Internet Publishing, Inc. 2009. WTFast. <https://www.wtfast.com/Subscription/Index/new>. (2009). Last accessed: January 26,2017.
- [2] Akamai. 2015. Akamai “10For10”. <https://www.akamai.com/us/en/multimedia/documents/brochure/akamai-10for10-brochure.pdf>. (July 2015).
- [3] Akamai. 2017. SureRoute. <https://developer.akamai.com/learn/Optimization/SureRoute.html>. (2017).
- [4] Akamai. 2017. The State of Online Retail Performance. <https://goo.gl/5dvc9D>. (2017).
- [5] Nick Allen. 2015. Elon Musk announces ‘space Internet’ plan. <http://www.telegraph.co.uk/news/worldnews/northamerica/usa/11353782/Elon-Musk-announces-space-Internet-plan.html>. (January 2015).
- [6] Amazon Web Services, Inc. 2018. Alexa Top Sites. <https://aws.amazon.com/alexa-top-sites/>. (2018). [Online; accessed 07-March-2017].
- [7] amazon.com. 2018. ASUS VG248QE Gaming Monitor. <https://goo.gl/gnFnPv>. (2018).
- [8] Battleping. 2010. Info on our lower ping service. <http://www.battleping.com/info.php>. (2010). Last accessed: January 26,2017.
- [9] Zachary S. Bischof, Fabián E. Bustamante, and Rade Stanojevic. 2017. The Utility Argument - Making a Case for Broadband SLAs. In *Passive and Active Measurement - 18th International Conference, PAM 2017, Sydney, NSW, Australia, March 30-31, 2017, Proceedings*. 156–169.
- [10] Brian Louis / Bloomberg. 2017. Trading Fortunes Depend on a Mysterious Antenna in an Empty Field. <https://goo.gl/82kzXd>. (2017).
- [11] Jeremy Bowman. 2016. Why It’s So Hard to Make a Profit in E-Commerce. <https://goo.gl/EAUDuy>. (2016).
- [12] Ilker Nadi Bozkurt, Anthony Aguirre, Balakrishnan Chandrasekaran, Brighten Godfrey, Gregory Laughlin, Bruce M. Maggs, and Ankit Singla. 2017. Why Is the Internet so Slow?!. In *Passive and Active Measurement - 18th International Conference, PAM 2017, Sydney, NSW, Australia, March 30-31, 2017, Proceedings*. 173–187.
- [13] Jake Brutlag. 2009. Speed Matters for Google Web Search. <http://goo.gl/vJq1lx>. (2009).
- [14] Gustavo Carneiro, Pedro Fortuna, and Manuel Ricardo. 2009. FlowMonitor: A Network Monitoring Framework for the Network Simulator 3 (NS-3). In *Proceedings of the Fourth International ICST Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS ’09)*. Article 1, 10 pages.
- [15] Center for International Earth Science Information Network (CIESIN), Columbia University; United Nations Food and Agriculture Programme (FAO); and Centro Internacional de Agricultura Tropical (CIAT). 2005. Gridded Population of the World: Future Estimates (GPWFE). <http://sedac.ciesin.columbia.edu/gpw>. (2005). Accessed: 2014-01-12.
- [16] Michael Chow, David Meisner, Jason Flinn, Daniel Peek, and Thomas F. Wenisch. 2014. The Mystery Machine: End-to-end Performance Analysis of Large-scale Internet Services. In *11th USENIX Symposium on Operating Systems Design and Implementation (OSDI 14)*. USENIX Association, Broomfield, CO, 217–231. <https://www.usenix.org/conference/osdi14/technical-sessions/presentation/chow>
- [17] Cisco. 2017. Cisco Visual Networking Index: Forecast and Methodology. <https://goo.gl/SrpKbL>. (2017).
- [18] cISP authors. 2018. Impact of rainfall on cISP for a period of 1 year. <https://goo.gl/uVDxcm>. (2018). [Online; accessed 31-May-2018].
- [19] cISP authors. 2018. MW path refining. <https://goo.gl/LwYB5Z>. (2018). [Online; accessed 31-May-2018].
- [20] cISP authors. 2018. The MW+fiber hybrid network evolves with budget. <https://goo.gl/3kDN6H>. (2018). [Online; accessed 31-May-2018].
- [21] Mark Claypool, David LaPoint, and Josh Winslow. 2003. Network Analysis of Counter-Strike and Starcraft. In *IEEE Performance, Computing, and Communications Conference*.
- [22] Federal Communications Commission. [n. d.]. Universal Licensing System. <http://wireless2.fcc.gov/UlsApp/UlsSearch/searchLicense.jsp>. ([n. d.]).
- [23] AT&T Corporation. 2003. AT&T Long Lines Routes March 1960. <http://long-lines.net/places-routes/maps/MW6003.html>. (2003).
- [24] DARPA. 2013. Novel Hollow-Core Optical Fiber to Enable High-Power Military Sensors. <http://www.darpa.mil/news-events/2013-07-17>. (2013).
- [25] Ramakrishnan Durairajan, Paul Barford, Joel Sommers, and Walter Willinger. 2015. InterTubes: A Study of the US Long-haul Fiber-optic Infrastructure. In *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication (SIGCOMM ’15)*. 565–578.
- [26] Federal Communications Commission. 2018. Antenna Structure Registration Database. <http://goo.gl/3OIFDT>. (2018).
- [27] Baris Erkmen for X Company. 2017. Exploring a new approach to connectivity. <https://goo.gl/vHycHi>. (2017).
- [28] Riot Games. 2016. Fixing the Internet for Real-time Applications. (2016). <https://goo.gl/SEoxW2>.
- [29] M. R. Garey and D. S. Johnson. 1977. The Rectilinear Steiner Tree Problem is NP-Complete. *SIAM J. Appl. Math.* 32, 4 (1977), 826–834. <https://doi.org/10.1137/0132071>
- [30] Megan Geuss. 2015. Satellite Internet: meet the hip new investment for Richard Branson, Elon Musk. <https://tinyurl.com/jaqlust>. (January 2015).
- [31] Google. 2015. 100B searches are conducted on Google each month. <https://goo.gl/2z6tqe>. (2015).
- [32] Inc. Gurobi Optimization. 2016. Gurobi Optimizer Reference Manual. (2016). <http://www.gurobi.com>
- [33] Nikola Gvozdiev, Stefano Vissicchio, Brad Karp, and Mark Handley. 2017. Low-Latency Routing on Mesh-Like Backbones. *ACM HotNets* (2017).
- [34] Jonas Hansryd and Jonas Edstam. 2011. Microwave capacity evolution. *Ericsson review* 1 (2011), 22–27.
- [35] Devindra Hardawar. 2016. Samsung proves why 5G is necessary with a robot arm. <https://goo.gl/3gZTn8>. (2016).
- [36] Zayo Group Holdings. 2018. Zayo closes acquisition of Spread Networks. (2018). <https://www.zayo.com/news/zayo-closes-acquisition-spread-networks/>.
- [37] Chi-Yao Hong, Srikanth Kandula, Ratul Mahajan, Ming Zhang, Vijay Gill, Mohan Nanduri, and Roger Wattenhofer. 2013. Achieving high utilization with software-driven WAN. In *ACM SIGCOMM 2013 Conference, SIGCOMM’13, Hong Kong, China, August 12-16, 2013*. 15–26.
- [38] ITU. 2005. Specific attenuation model for rain for use in prediction methods. http://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.838-3-200503-I!!PDF-E.pdf. (2005). [Online; accessed 26-January-2018].
- [39] Ixia. 2012. Measuring Latency in Equity Transactions. <https://goo.gl/RkiBhG>. (2012).
- [40] Sushant Jain, Alok Kumar, Subhasree Mandal, Joon Ong, Leon Poutievski, Arjun Singh, Subbaiah Venkata, Jim Wanderer, Junlan Zhou, Min Zhu, Jon Zolla, Urs Hölzle, Stephen Stuart, and Amin Vahdat. 2013. B4: experience with a globally-deployed software defined wan. In *ACM SIGCOMM 2013 Conference, SIGCOMM’13, Hong Kong, China, August 12-16, 2013*. 3–14.
- [41] Jan Odvarko. 2007. HAR 1.2 Spec. <http://www.softwareishard.com/blog/har-12-spec>. (2007).
- [42] Srikanth Kandula, Dina Katabi, Bruce Davie, and Anna Charny. 2005. Walking the Tightrope: Responsive Yet Stable Traffic Engineering. In *Proceedings of the 2005 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM ’05)*. 253–264.

- [43] Kevin Kelly. 2007. How much does one search cost? <http://kk.org/thetechnium/how-much-does-o/>. (2007).
- [44] Issie Lapowsky. 2014. Facebook Lays Out Its Roadmap for Creating Internet-Connected Drones. <http://www.wired.com/2014/09/facebook-drones-2/>. (September 2014).
- [45] Gregory Laughlin, Anthony Aguirre, and Joseph Grundfest. 2014. Information Transmission between Financial Markets in Chicago and New York. *Financial Review* (2014).
- [46] Kyungmin Lee, David Chu, Eduardo Cuervo, Johannes Kopf, Yury Degtyarev, Sergey Grizan, Alec Wolman, and Jason Flinn. 2015. Outatime: Using speculation to enable low-latency continuous interaction for mobile cloud gaming. In *ACM MobiSys*.
- [47] Greg Linden. 2006. Make Data Useful. <https://goo.gl/eqb3p8>. (2006).
- [48] McKay Brothers LLC. 2017. Quincy Extreme Data Latencies. <http://www.quincy-data.com/product-page/#latencies>. (2017).
- [49] Trevor Manning. 2009. *Microwave Radio Transmission Design Guide*. Artech House.
- [50] Ginny Marvin. 2017. Report: Google earns 78% of \$36.7B US search ad revenues, soon to be 80%. <https://goo.gl/kp4L5X>. (2017).
- [51] Microsoft Azure. 2018. Content Delivery Network pricing. <https://azure.microsoft.com/en-us/pricing/details/cdn/>. (2018). [Online; accessed 30-January-2018].
- [52] Rani Molla. 2017. Amazon could be responsible for nearly half of U.S. e-commerce sales in 2017. <https://goo.gl/QqAYCv>. (2017).
- [53] NASA. 2015. Precipitation Processing System Data Ordering Interface for TRMM and GPM (STORM). <https://storm.pps.eosdis.nasa.gov/storm/>. (2015). [Online; accessed 26-January-2018].
- [54] NASA Jet Propulsion Laboratory. 2015. U.S. Releases Enhanced Shuttle Land Elevation Data. <https://www2.jpl.nasa.gov/srtm/>. (2015). [Online; accessed 28-January-2018].
- [55] NEC. 2014. SEA-US: Global Consortium to Build Cable System Connecting Indonesia, the Philippines, and the United States. <https://tinyurl.com/ybj9nhp3>. (August 2014).
- [56] Ravi Netravali, Anirudh Sivaraman, Somak Das, Ameesh Goyal, Keith Winstein, James Mickens, and Hari Balakrishnan. 2015. Mahimahi: Accurate Record-and-Replay for HTTP. In *USENIX ATC*. USENIX.
- [57] A. Nordrum. 2015. Fiber optics for the far North [News]. *IEEE Spectrum* 52, 1 (January 2015), 11–13.
- [58] ns-3 community. 2011. Network simulator ns-3. <https://www.nsnam.org>. (2011).
- [59] Pingzapper. 2018. Pingzapper Pricing. <https://pingzapper.com/plans>. (2018). Last accessed: January 26, 2017.
- [60] Enric Pujol, Philipp Richter, Balakrishnan Chandrasekaran, Georgios Smaragdakis, Anja Feldmann, Bruce MacDowell Maggs, and Keung-Chi Ng. 2014. Back-office web traffic on the internet. In *ACM IMC*.
- [61] Shkilko, A. and Sokolov, K. 2016. Every Cloud Has a Silver Lining: Fast Trading, Microwave Connectivity and Trading Costs. <https://ssrn.com/abstract=2848562>. (2016).
- [62] SimilarWeb. 2017. Overview: amazon.com. <https://www.similarweb.com/website/amazon.com>. (2017).
- [63] Ankit Singla, Balakrishnan Chandrasekaran, P. Brighten Godfrey, and Bruce Maggs. 2014. The Internet at the Speed of Light. In *Proceedings of the 13th ACM Workshop on Hot Topics in Networks (HotNets-XIII)*. ACM, New York, NY, USA, Article 1, 7 pages. <https://doi.org/10.1145/2670518.2673876>
- [64] Steam. 2017. Steam & Game Stats. (2017). <http://store.steampowered.com/stats/>.
- [65] Unwired Labs. 2018. OpenCellID Tower Database. <https://opencellid.org/>. (2018).
- [66] USGS. 2018. National Elevation Dataset (NED). <https://lta.cr.usgs.gov/NED>. (2018). [Online; accessed 31-May-2018].
- [67] J. H. Winters, J. Salz, and R. D. Gitlin. 1994. The Impact of Antenna Diversity on the Capacity of Wireless Communication Systems. *IEEE Transactions on Communications* 42, 2/3/4 (Feb/Mar/Apr 1994), 1740–1751.
- [68] X Development LLC. 2017. Project Loon. <https://www.solveforx.com/loon/>. (2017).