Pingin' in the Rain

Aaron Schulman and Neil Spring
University of Maryland
{schulman,nspring}@cs.umd.edu

ABSTRACT

Residential Internet connections are susceptible to weather-caused outages: Lightning and wind cause local power failures, direct lightning strikes destroy equipment, and water in the atmosphere degrades satellite links. Outages caused by severe events such as fires and undersea cable cuts are often reported upon by operators and studied by researchers. In contrast, outages cause by ordinary weather are typically limited in scope, and because of their small scale, there has not been comparable effort to understand how weather affects everyday last-mile Internet connectivity.

We design and deploy a measurement tool called *ThunderPing* that measures the connectivity of residential Internet hosts before, during, and after forecast periods of severe weather. ThunderPing uses weather alerts from the US National Weather Service to choose a set of residential host addresses to ping from several vantage points on the Internet. We then process this ping data to determine when hosts lose connectivity, completely or partially, and categorize whether these failures occur during periods of severe weather or when the skies are clear. In our preliminary results, we find that compared to clear weather, failures are *four times* as likely during thunderstorms and *two times* as likely during rain. We also find that the duration of weather induced outages is relatively small for a satellite provider we focused on.

Categories and Subject Descriptors

C.2.5 [**Computer Communications Networks**]: Local and Wide-Area Networks—*Internet*

General Terms

Experimentation, Measurement

Keywords

ThunderPing, Weather, Outage, Ping

IMC'11, November 2–4, 2011, Berlin, Germany.

1. INTRODUCTION

Last-mile residential Internet links are increasingly relied upon, even for emergency communication. Yet despite their apparent reliability on temperate days, network links are somewhat more prone to failure during times of severe weather. Branches may fall on telephone wires or power lines, and lightning strikes interfere with wireless transmissions and can destroy antenna towers.

Researchers have studied the Internet-wide effects of large scale outages, including undersea cable cuts [1] and terrorist attack [2]. Large-scale failures cause the Internet to fall back to large-scale redundancy: alternate geographic paths that absorb the load. Residential Internet connectivity, in contrast, lacks similar redundancy.

In this paper, we present a preliminary study into the prevalence of weather-related residential Internet outages in the United States. Our broader goal is to understand the factors involved in day-to-day network reliability, including aspects inherent to link type (e.g., wireless or cable), geography (a hurricane zone), or, when possible, the ISP itself. We focus on the US in this work because of the diversity of weather and the availability of a feed of weather alerts and coverage by Internet-accessible weather stations.

We find that compared to clear weather, failures are four times more likely during thunderstorms and two times more likely during rain. We expect that failures during clear weather are a baseline, typically indicating hosts that are powered off by users. We also observed that the failures for a satellite ISP that we studied tend to be shorter during thunderstorms than they are in clear conditions. Our dataset is available on the project website[1].

There are limitations to our study. First, we do not (yet) isolate weather-related power failures from network failures. Although the ISP should not be blamed for the failure, the effect (to us and to the user) is likely the same. Second, we do not (robustly) distinguish between actions of individuals from network failures that affect many hosts: For example, people may turn off their computers (and home routers). Further, they may be more likely to power down if they believe doing so will protect them from unstable power during a lightning storm. Third, we use scalable, but imprecise reports of weather and of host location to find correlations between both. As such, this work is preliminary, but shows promise in understanding the susceptibility of various networks to varied weather events in practice.

This paper is organized as follows. In Section 2, we describe the design of our measurement tool, ThunderPing,

[1] http://www.cs.umd.edu/~schulman/thunderping.html

```
<title>Severe Weather Statement issued May 12 at 4:46PM CDT
 expiring May 12 at 5:15PM CDT by NWS GreenBay
 http://www.crh.noaa.gov/grb/</title>
<summary>...A SEVERE THUNDERSTORM WARNING REMAINS IN EFFECT
 FOR CENTRAL WAUPACA AND NORTHWESTERN OUTAGAMIE COUNTIES
 UNTIL 515 PM CDT... AT 443 PM CDT...NATIONAL WEATHER
 SERVICE DOPPLER RADAR INDICATED A SEVERE THUNDERSTORM
 CAPABLE OF PRODUCING QUARTER SIZE HAIL...AND DAMAGING
 WINDS IN EXCESS OF 60 MPH. THIS STORM WAS LOCATED 7 MILES
 NORTH OF NEW LONDON...OR 20 MILES NORTHEAST OF
 WAUPACA...MOVING</summary>
<cap:effective>2011-05-12T16:46:00-05:00</cap:effective>
<cap:expires>2011-05-12T17:15:00-05:00</cap:expires>
<cap:urgency>Immediate</cap:urgency>
<cap:severity>Severe</cap:severity>
<cap:certainty>Observed</cap:certainty>
<cap:geocode><valueName>FIPS6</valueName>
<value>055087 055135</value></cap:geocode>
```

Figure 1: Example XML entry for a weather alert for two counties in Wisconsin. Some XML entries omitted for brevity.

which collects reachability samples for hosts before, during, and after severe weather events. Section 3 describes our analysis of the data to observe failures. Section 4 presents our preliminary results. We conclude with a discussion of the implications of this work and describe our future work in Section 5.

2. MEASURING THE RESPONSIVENESS OF INTERNET HOSTS DURING WEATHER

We developed ThunderPing to test our hypothesis that weather affects the performance of residential Internet connections, measured by connectivity, loss rate, and latency. In this section, we describe the design of ThunderPing. At a high level, ThunderPing listens for severe weather alerts issued by the US National Weather Service (NWS). When one is issued, ThunderPing finds IP addresses in the area covered by the alert and uses ten geographically distributed PlanetLab hosts to ping those addresses every eleven minutes for up to six hours before, during, and six hours after the alert. Because we are interested in distinguishing network failure from network congestion, ThunderPing retries a lost ping at most ten times, if it has seen a successful ping in a prior interval.

2.1 Finding IP addresses subject to weather

The first problem to address is to find residential network IP addresses in a geographic region that can be matched to a US National Weather Service alert. We select IP addresses by a scan of the reverse DNS space, classify each IP address as residential by DNS suffix (domain), and determine their approximate location by the MaxMind GeoIP database.

The focused scan of reverse DNS records proceeds as follows. First we choose three IP addresses, ending in .1, .44, or .133 from every possible /24 block, and query for the name of each. If any of the three have a name matching a well-known US residential ISP, such as comcast.net or verizon.net, we determine all the names of all the IP addresses in the block and include the addresses with matching names. This approach is comparable to that used to study residential Internet connections in prior work [3, 13]. From this method, we discovered 100,799,297 US residential IP addresses.

The US National Weather Service provides an XML feed of the latest severe weather alerts for regions in the US [7]. An example alert appears in Figure 1. The regions under alert are listed by FIPS code, which is a numeric code for each county in the US. The FIPS code for Los Angeles, for example, is 06037. We consider all weather alerts including "watches," which indicate conditions conducive to severe weather, and "warnings," which indicate that severe weather has been observed.

To link IP addresses to the FIPS codes used in weather alerts requires IP geolocation. We used MaxMind's GeoIP [5] database to determine an estimate of the latitude and longitude of each IP, and the US Census Bureau's county border data file[2] to determine the FIPS county location for any IP address.

We use MaxMind's database because of its availability and the potential to determine the location of every possible residential IP address. Researchers have questioned its accuracy [9], and have developed probing-based methods for positioning Internet hosts [10, 12] that seem impractical for locating 100 million hosts. Clearly, improved IP geolocation methods would yield more precision to the location and might lend more accuracy to our analysis. We expect, however, that precision in geolocation would have limited benefit because weather alerts are provided on the scale of a county and because weather does not respect city or county boundaries.

After an alert comes in, we pick 100 IP addresses from every provider and link type (when embedded in the DNS name) in each FIPS-coded region in the alert. We identify a provider and link type by the DNS name without numbers (e.g., pool----.sangtx.dsl-w.verizon.net).

2.2 Pinging (residential IPs) in the rain

Testing our hypothesis that weather affects the Internet is difficult because weather's effect on the connectivity of Internet hosts may be hidden by congestion, outages at the source, or other network events.

2.2.1 Ping infrequently

Internet measurement traffic has a tendency to generate reports of network abuse from recipients of unsolicited traffic. We send typical ICMP echo messages with an identifying payload as infrequently as possible.

We follow the inter-ping interval chosen by Heidemann et al. in their Internet census. [4]. They surveyed the occupancy of IP addresses on the Internet on the scale of tens of minutes. They reported that they could send pings at an interval of 5 minutes without generating any abuse reports. For their surveys they pinged IP addresses for several weeks at an 11 minute interval without generating many abuse reports, so we do the same.

2.2.2 Omit needless pings

In addition to sending more probes to determine if a host is down, ThunderPing must cull the set of observed hosts during a weather alert to include only those that respond to pings. Otherwise, the pinger would waste time pinging addresses that either are not assigned to a host or have a host that is not awake for the weather event. We implement a simple timeout: If after an hour (five pings from ten vantage points) a response is not heard from the host, then it is no longer pinged for that weather alert.

[2] http://www.census.gov/geo/cob/bdy/co/co00ascii/co99_d00_ascii.zip

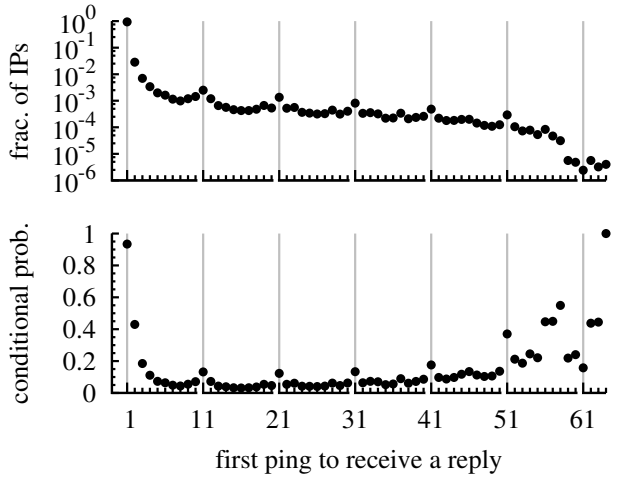

Figure 2: Hosts may not respond to the first ping. The lower graph shows the conditional probability that a host first replies to ping x, given that it does not reply to any of the pings preceding x. The upper graph shows the fraction of hosts (in log scale) that first reply to ping x. The gray lines indicate the ping interval when all ten vantage points are operational.

2.2.3 *One vantage point is not enough*

ThunderPing distinguishes between faults in the middle and faults at the endpoint, the observed host, by simultaneously pinging from several vantage points. The responsiveness of the host is not determined by any individual vantage point, but by agreement between the vantage points. For the experiment in this paper we used ten PlanetLab machines as vantage points: Ten permitted some to fail occasionally while giving each IP address a good chance of demonstrating availability even during partial outages.

2.2.4 *One ping is not enough*

Pings can be lost: A single failed ping is not convincing evidence of a host being unavailable. Clearly, then, a failed ping must be followed up to test whether the failure is transient or persistent. ThunderPing vantage points retry a failed ping up to ten times, doubling the interval between each attempt to avoid flooding the host.

The short-term retry of a failed ping is important beyond the simple reaction to background loss: because of a nuance in ARP, at least a second ping is *necessary* to determine whether a host is up. When a packet reaches a router that does not have an ARP table entry for the next hop (in our case, likely the destination) IP address, the ARP RFC [8] states that the router should drop the packet and send an ARP request instead. We confirmed that Cisco's IOS follows this behavior. This means if you only send one ping to an IP, if there are any expired ARP entries along the path, the packet will be dropped. However, we also observed that the default behavior of Cisco IOS is to attempt to refresh an ARP entry as soon as it expires, which mitigates the problem.

In our dataset, if a host does not respond to the first ping, the first reply is most likely to come from the second ping. For the 1,240,114 hosts that reply to pings within one hour of receiving pings, the bottom of Figure 2 shows the conditional probability that a host first replies to ping request x, given that it does not reply to any of the ping requests preceding x. The conditional probability for the second ping is at least 2.3x the conditional probability for all pings preceding ping 41. For the remaining pings (41-64), fewer than 0.5% of the hosts that respond in one hour remain. The expected conditional probability at the right edge of the graph, when there are few hosts remaining, is to increase until the 64th ping where all remaining hosts reply. This is because the graph is constrained to show only those addresses that responded to at least one of the 64 pings.

The hosts that do not reply to the first few pings, but still reply within the first hour, appear to come online at a random time during the hour. The data indicate this because the conditional probability is consistent after the third ping and it peaks every ping interval. When all ten (occasionally 11) vantage points are operational, all ten ping once every 11 minutes: the first vantage point to ping during each epoch is more likely to discover a newly-online host.

The apparent importance of the second ping might be an artifact of our measurement setup. However, we examined the data to filter out hosts that might be biased by the measurement setup. In particular, we noticed that as a source sends a long sequence of pings to various hosts, a suffix of this sequence sometimes receives no responses. Pings in that suffix may fail to be transmitted due to overload at the PlanetLab source, or fail to traverse the network past some buffer. In response, we observe whether the sequence of pings has a failed suffix and filter out hosts pinged in that suffix for this analysis.

2.3 Potential sources of error

A source of error for our probing would be when a host appears to have failed, but in reality, its DHCP lease just ran out and it was given a new address. From correlating Hotmail identities to IP addresses for one month of data, Xie et al. [13] report that for SBC, one of the largest DSL providers in the US, most users retained the same IP for one to three days. For Comcast, one of the largest cable providers in the US, they report that 70% of IP addresses do not change for users within a month-long trace. This stability suggests that the addresses of responsive hosts will not be reassigned in a way that would suggest failure during weather events.

3. ANALYZING THE PINGS

The severe weather forecasts and geolocation to IP addresses derived in the previous section permit ThunderPing to observe residential hosts before, during, and after weather events. There are two remaining tasks, however: first to determine the actual weather during the instrumentation, and second to interpret the ping data to decide whether a host or set of hosts failed. We discuss each in turn.

3.1 The weather at a host during a ping

The NWS and Federal Aviation Administration (FAA) administer approximately 900 Automated Surface Observing System (ASOS)[3] weather stations at airports in the US. These stations provide hourly weather measurements, primarily for pilots, in METAR format. Beyond the basic wind, pressure, and rainfall sensors, ASOS stations include a Light

[3] http://www.weather.gov/asos/

	Cable			DSL						Satellite	Fiber
	Charter	Comcast	Cox	Ameritech	CenturyLink	MegaPath	Speakeasy	Windstream	Verizon DSL	WildBlue	Verizon FiOS
pinged	321,705	746,051	161,756	156,327	167,347	23,076	110,795	370,018	213,737	395,810	86,344
alive	54,277	100,627	31,918	8,553	76,705	3,771	7,869	113,541	47,839	52,979	36,143
UP to DOWN	3,689	6,769	2,552	742	12,746	216	418	8,988	5,376	17,825	805
UP to HOSED	1,227	2,111	644	240	5,385	107	163	3,358	2,565	13,807	211
airports	237	320	130	113	203	129	229	186	189	424	145
Clear	17.5	14.0	57.4	21.3	30.3	51.3	39.8	22.5	20.4	33.3	13.3
Cloudy	10.9	9.9	28.9	15.5	13.7	39.3	30.5	9.4	17.0	17.4	10.1
Fog	1.8	1.5	6.4	1.5	2.0	10.8	5.4	1.9	1.8	2.3	2.8
Rain	1.6	1.6	3.7	2.7	2.2	5.2	3.6	1.6	4.3	3.4	1.2
Thunderstorm	0.7	0.6	1.6	1.0	1.0	1.4	1.4	0.7	1.0	1.1	0.3

Table 1: Summary of a small portion of the data collected by ThunderPing. For a sample of providers, the number of IP addresses pinged, the number that are ever seen alive, as well as the number of IP addresses that transition at least once from UP to DOWN and UP to HOSED. The next row is the number of airports that the alive IPs map to. The final five rows are the average time (in hours) ThunderPing pings an IP address during each weather condition.

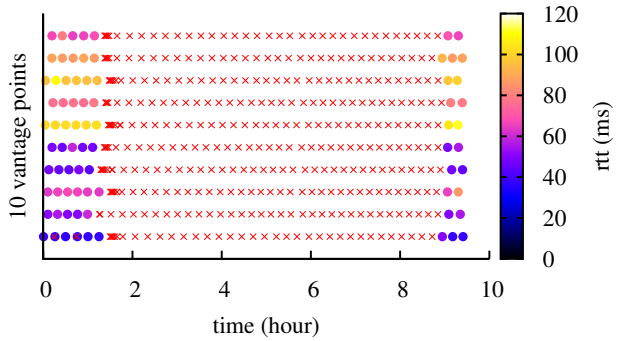

Figure 3: During heavy thunderstorms reported at KTUP airport in Tupelo, Mississippi at 3:16AM on April 27th, 2011, ten vantage points (y-axis) saw Comcast cable customer 75.66.230.135 go down for 7.6 hours. A • represents a ping with a response, a × represents a ping that timed out. At this time, eight other Comcast customers near this airport failed.

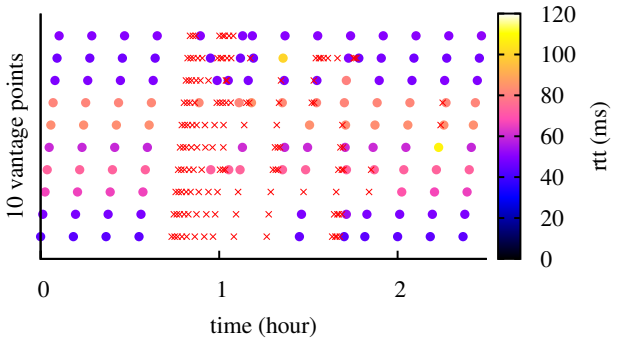

Figure 4: We also observed hosts enter a state of partial responsiveness. In this state, the host intermittently responds to pings. CenturyLink DSL customer 209.206.208.59 goes into a partial responsiveness state for 37 minutes during heavy thunderstorms reported at KSTP in St Paul, Minnesota at 6:30AM on May 9th, 2011.

Emitting Diode Weather Indicator (LEDWI) that measures the type (Rain, Hail, Snow) and density of precipitation. Most stations also have antennas that measure lightning strikes. Weather Underground provides an archive of the hourly METAR readings from many ASOS stations [11]. We do not use the larger network of older AWOS weather stations because some do not have the LEDWI sensor, which provides data we consider important.

We interpret the ping data to classify a host into three states: UP, where most pings receive responses, HOSED where few pings receive responses, and DOWN where none of the pings receive responses. The task is to identify when hosts transition between these states and to correlate those transitions with the current weather conditions for the IP address.

The transition rules are as follows. UP to DOWN occurs as soon as all vantage points report failure (Figure 3). We intend for DOWN to represent a complete lack of responsiveness, and we do not assume a short interval of failed pings represents a failed host. A DOWN to UP transition occurs as soon as all active vantage points agree that the host is up.

These are the simple transitions, but these rules permit an intermediate state. The HOSED state represents a situation in which fewer than half of the pings see responses: The host is not entirely DOWN, but there is *something* going on (Figure 4). To identify these transitions, we slide a window of three pings from every vantage point (30 pings total) across time. To enter HOSED, more than half (but not all) of the pings in the sliding window must fail. To leave HOSED, more than half of the pings must succeed. The sliding window prevents short periods of UP-like responsiveness from splitting an interval of HOSED-state behavior. Because ThunderPing retries pings that fail, the HOSED state will be found quickly: much more quickly than three ping intervals. We have not observed oscillations between UP and HOSED when loss rates are near 50%, perhaps because this is rare.

To handle vantage point failure, we exclude those that fail from the unanimous voting for UP or DOWN. In our experiments, when vantage points fail they are automatically replaced, though we developed the analysis to work despite up to four failures at once.

4. RESULTS

We collected pings with ThunderPing for 66 days, between April 27th, 2011 and August 1st, 2011. Table 1 shows a summary of the observations. ThunderPing pings IP addresses at providers only if a weather alert is issued for an area in which the provider has customers. Because both weather and ISPs are regional, the IP addresses of some providers will spend different amounts of time in each condition.

Although we probed customers of many providers, in this analysis we focus on relatively large providers where the cus-

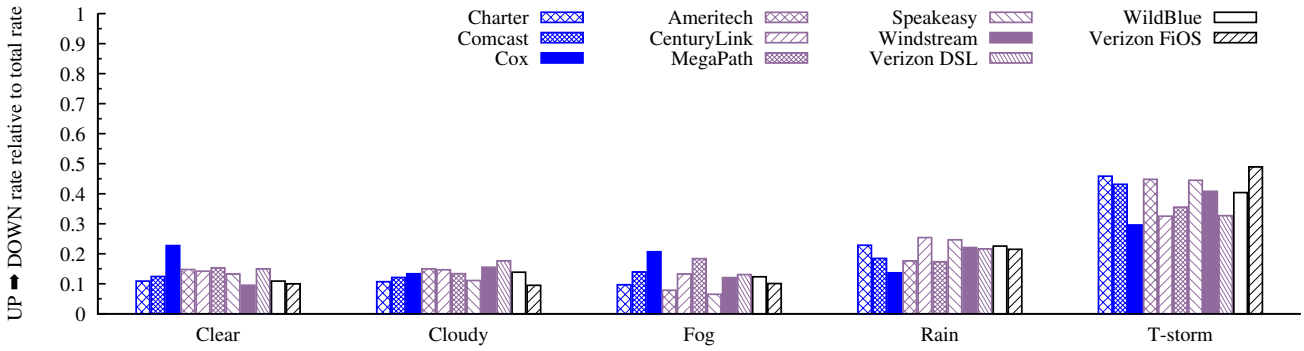

Figure 5: The rate of UP to DOWN failures relative to the total failure rate for each provider. Weather affects the failure rate for all types of links, and for all selected providers.

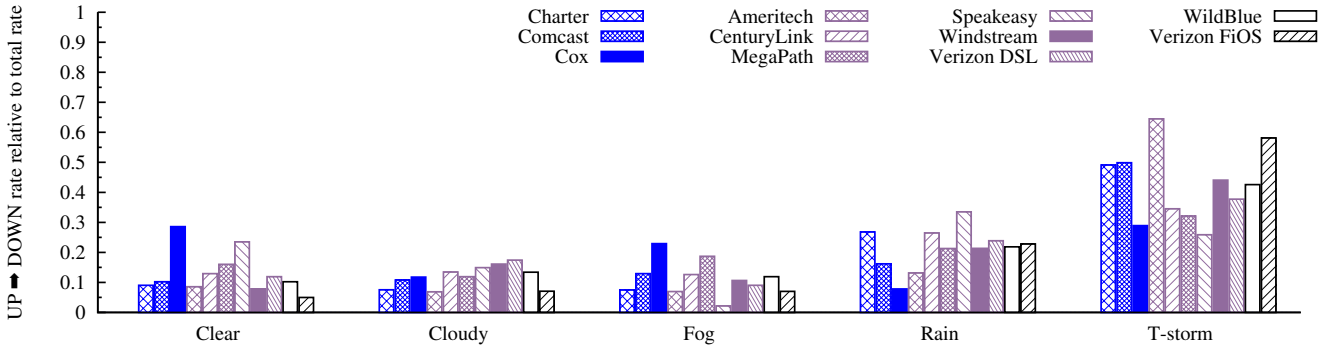

Figure 6: The normalized rate of correlated UP to DOWN failures. Thunderstorms appear to affect the failure rate for all types of links and providers.

tomer's link type is known and easy to identify in their reverse DNS names. We avoid providers that offer various link types but do not include that information in their DNS names. We chose three prominent cable providers: Charter, Comcast, and Cox; six DSL providers: Ameritech, CenturyLink, MegaPath, Speakeasy, Windstream and Verizon DSL; and two uncommon link types: the rural satellite provider WildBlue, and Verizon's fiber optic service.

4.1 Weather affects residential Internet

To understand the effect of weather on the Internet we found the transitions from UP to DOWN for all IP addresses broken down by provider. To isolate failures that are severe (longer than a few minutes) and not caused by human behavior (e.g., turning a computer off for at least 8 hours at night[4]) we plot transitions where the IP address is unresponsive for more than 30 minutes but less than 8 hours. Then, we total up the time that each provider experiences each type of weather. The *failure rate* is the number of UP to DOWN transitions divided by the total time each IP address observed each weather condition.

Figure 5 depicts the failure rate for each weather condition, relative to the total failure rate for that provider. This normalization allows us to observe the effect that weather has on the different providers, and compare the effect across providers. During thunderstorms the failure rate more than doubles compared to clear for most providers across all link types. For example, the failure rate for Comcast (cable) during thunderstorms is 3.58 times the failure rate in clear weather; for the least affected DSL provider the failure rate is 3.21 times higher than clear. As expected, due to rain fade of their Ka band (30 GHz) signal, rain appears to double the failure of the satellite provider WildBlue, but surprisingly an increased failure rate appears across providers. We address the outlying reliability of Cox during thunderstorms in the following section.

In Figure 6, we perform the same analysis for correlated failures. We define a correlated failure as a failure where at least one other host from the same provider, located near the same airport where we measure the weather, 5.5 minutes before and after the failure (one ping interval). Selecting failures that are correlated in time attempts to choose those that are caused by factors beyond the user's control: individual people may turn off their computer independently, but are unlikely to do so simultaneously, while a network outage may affect more than one customer. In this plot, the effect of weather is even more evident than when looking across all failures. During thunderstorms, several providers have about five times the clear correlated failure rate.

4.2 Duration of failures

Failures of the rural satellite provider WildBlue are relatively short compared to the other providers we observed

[4] In our dataset, most UP to DOWN transitions occur at 10 PM local time, with an average DOWN time of 12 hours.

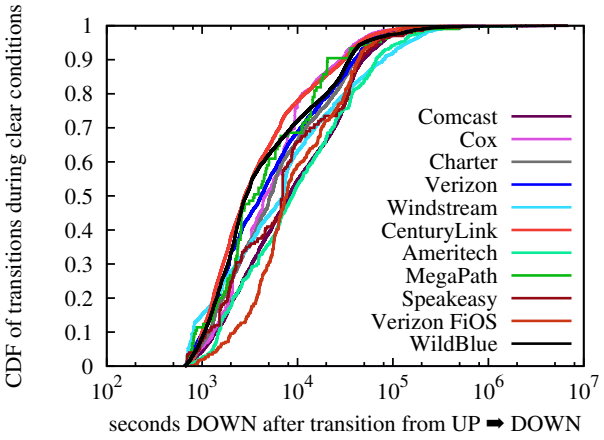

Figure 7: CDF of the time DOWN after transitioning from UP in clear conditions. Some providers stay DOWN longer than others and it does not appear to be related to link type.

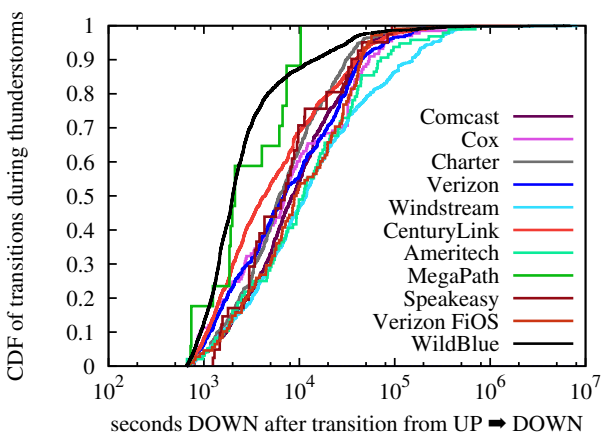

Figure 8: CDF of the time DOWN after transitioning from UP during thunderstorms. Surprisingly, the satellite provider's failures are shorter than in clear conditions.

(Figure 7). A likely explanation for this is that bursty outages caused by weather and atmospheric conditions are a common failure for Ka band satellite networks [6]. Our dataset confirms this hypothesis, WildBlue has the shortest median duration of failures as well as the shortest median time between failures.

As expected, in clear conditions, failures of a satellite provider are bursty, but in a thunderstorm, failures could be much longer. Figure 8 shows the distribution of the time hosts are DOWN after a failure during thunderstorms. Surprisingly, however, the time between failures for WildBlue appear to be shorter in thunderstorms than they are in clear conditions. Long failures also do not occur for a large portion of the WildBlue IPs observed. Of the 2,833 WildBlue IPs that fail at least once in thunderstorms, only 14% fail at least once for at least 3 hours, compared to 54% for the 1,121 Windstream IPs.

In the failure rate plot, Figure 6, Cox has correlated failures that do not seem to be related to weather. Cox has a significant fraction of transitions that last 0.8 hours. We observed a large correlated failure of 68 Cox customers with the reverse name mta*.mm.br.br.cox.net where the numbers appear to be a MAC address. The MAC address manufacturer is Arris equipment, and their E-MTA product is a cable modem and VoIP endpoint. This model of modems has a battery backup to provide voice service even if the power goes out. Considering that all 68 failed at night (12:40AM) at the same time, during clear weather at their closest airport KDTS, this seems likely to be a provider caused outage, possibly a firmware update.

5. DISCUSSION AND FUTURE WORK

We presented a preliminary study of failures during fairly typical weather events. Despite the short measurement duration of 66 days, for the large providers discussed in the results we observed the failures of 65,529 distinct IP addresses (23,957 more went HOSED) and observed significant correlation between severe weather and unreachability, even when looking only for simultaneous failures of nearby hosts. We see this as a promising preliminary analysis into the typical causes of everyday Internet unavailability.

We plan to continue this study by addressing sources of potential inaccuracy. We chose our location database for its ability to quickly position 100 million host IP addresses so that we might determine which to ping, but after data collection, we might need accurate positions of only the 129,631 that showed faults. Such may permit the use of more precise methods after-the-fact. Increased precision in IP positioning would demand increased precision in weather observations. The weather observed at an airport might not be observed just a few miles away. We would like to augment these surface measurements with more precise weather location and precipitation density information from Doppler radar. (Radar images on television that include precipitation type are a similar synthesis of surface observations of the precipitation type and the radar image showing precipitation location, but are not provided by NWS.)

We would also like to identify whether observed failures are indeed in the "last mile" network or are instead in a backhaul network. We expect that intermittent TTL-limited probing may help identify at least some upstream faults, depending on the network technology deployed.

We expect that additional data will enable study of how weather-induced outages depend, not just on the currently active condition, but also on measurable quantities, such as the duration of a storm, the inches of precipitation, or wind speed.

Finally, we would like to expand the study beyond the United States; the key stumbling point is in finding standardized, location-specific forecast data like that provided by NWS and a public network of standardized sensors, comparable to ASOS, that includes the precipitation type sensor.

6. ACKNOWLEDGMENTS

Thanks to Patrick Shoemaker and Dave Levin for helpful discussions. Also thanks to the anonymous reviewers and our shepherd Fabián Bustamante for their comments. This work was supported by NSF-0643443, NSF-0917098 and NSF-0626629.

7. REFERENCES

[1] E. W. W. Chan, X. Luo, W. W. T. Fok, W. Li, , and R. K. C. Chang. Non-cooperative diagnosis of submarine cable faults. In *Passive and Active Measurement Conference (PAM)*, 2011.

[2] Committee on the Internet Under Crisis Conditions: Learning from the Impact of September 11. *The Internet Under Crisis Conditions Learning from September 11*. National Academies Press, 2003.

[3] M. Dischinger, A. Haeberlen, K. P. Gummadi, and S. Saroiu. Characterizing residential broadband networks. In *Proceedings of the ACM Internet Measurement Conference (IMC)*, 2007.

[4] J. Heidemann, Y. Pradkin, R. Govindan, C. Papadopoulos, G. Bartlett, and J. Bannister. Census and survey of the visible Internet. In *Proceedings of the ACM Internet Measurement Conference (IMC)*, 2008.

[5] MaxMind, Inc. GeoIP city database. http://www.maxmind.com/app/city, Jan. 2011.

[6] I. Minei and R. Cohen. High-speed internet access through unidirectional geostationary satellite channels. In *IEEE Journal on Selected Areas in Communications*, 1999.

[7] National Weather Service. NWS public alerts in XML/CAP v1.1 and ATOM formats. http://alerts.weather.gov/.

[8] D. C. Plummer. RFC 826: Ethernet Address Resolution Protocol: Or converting network protocol addresses to 48.bit Ethernet address for transmission on Ethernet hardware, Nov. 1982.

[9] I. Poese, S. Uhlig, M. A. Kaafa, B. Donnet, and B. Gueye. IP geolocation databases: Unreliable? In *ACM SIGCOMM Computer Communication Review (CCR)*, 2011.

[10] Y. Wang, D. Burgener, M. Flores, A. Kuzmanovic, and C. Huang. Towards street-level client-independent IP geolocation. In *USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, 2011.

[11] Weather Underground. Weather history and data archive. http://www.wunderground.com/history/.

[12] B. Wong, I. Stoyanov, and E. G. Sirer. Octant: A comprehensive framework for the geolocalization of Internet hosts. In *USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, 2007.

[13] Y. Xie, F. Yu, K. Achan, E. Gillum, M. Goldszmidt, and T. Wobber. How dynamic are IP addresses? In *Proceedings of the ACM SIGCOMM Conference*, 2007.

Summary Review Documentation for

“Pingin' in the Rain”

Authors: A. Schulman, N. Spring

Reviewer #1

Strengths: Novel, fun idea, interesting results.

Weaknesses: Validation is a little weak.

Comments to Authors: Fun paper. Now I wish I could get that song out of my head...

Clearly a novel idea, and an interesting one at that. Given personal experience of living in different rural areas of the states and dealing with power outages of different lengths during storms, it makes sense that there is some elevation in network outages during bad weather.
I appreciate the discussion of weaknesses and potential sources of error. There are clearly some potential issues there, but it appears that, given the steps taken to mitigate those problems, the measurements are accurate enough to say that there are clearly some network problems correlated with bad weather. Still, coming up with some ways to validate what you are measuring, even for a subset of locations (e.g., select some location(s) in tornado alley, with cooperative ISPs and/or users) would lend much more credence to the work. The couple anecdotes were helpful (Figs 2 and 3), but something a little more systematic would benefit the work.

There was no discussion as to the coverage of the 100k ip addresses. How are these distributed geographically? Are there clumps of addresses in different places, or are they distributed similarly as population? Are all zip codes covered, or what fraction? Put another way, were there weather alerts for which you had few or no addresses to ping?

I wonder whether there are electric grid outage reports that can be correlated with your measurements. It might get to the interesting question you raise regarding whether you are seeing power outages or ISP outages.

It is too bad some measurements were not done in a slightly more intrusive way to try to get more fine grained measurements, especially of outage times.

Reviewer #2

Strengths: Neat, novel idea and pretty good approach.

Weaknesses: Not so many interesting results and relatively poor analysis; quite a few magic numbers (5.5 minutes, 8hrs,...).

Comments to Authors: This is a neat piece, if perhaps a bit premature. The report reads somewhat like a class project, which it probably is, and the analysis is a tad surfacy.

There is a large number of interesting questions unanswered, which perhaps makes sense for this as a short paper.

For instance, how are you going to address the second and third weaknesses in your approach? What is the impact of bad geolocation? How commonly would users turn off their machines during a storm and how could you get rid of those data points?

The definition of a weather event is rather broad, basically independent of the severity of it (wind speed, for instance). Also, I will assume that storm or high-wind prone areas are less susceptible to their common phenomena than other areas.

While I have no problem with the need for more than one ping, the argument is weak - starting from section header like “one ping is not enough” to go down to “2.3% fewer responses than the 2nd ping”.

What do you mean by “to instrument” residential hosts (Sec. 3)?

You state - “to isolate failures that are not caused by human behavior we plot transitions where the IP address is unresponsive for more than 30 minutes but less than 8 hours”; I assume you mean to say “to factor out failures caused by human behavior”; it is unclear, however, what this filtering yields and the logic behind the bounds - why 8hrs?

You aggregate failures for a provider *across* all the events of a certain type; I assume this is due to not having enough data points? Treating all T-storms as a single set, for instance, is a bit too simplistic and may hide some interesting results on the impact of more common weather phenomena.

Why 5.5' in your correlation analysis? How sensitive are your results to this?

This is just a few of my main issues, perhaps none of them totally negative.

Reviewer #3

Strengths: An interesting, enticing hypothesis that is novel and requires methodological innovations to prove or disprove.

The authors have made enough headway to show viability.

Weaknesses: All said and done, the evidence in the paper of higher failure rates during bad weather is circumstantial. (But given its a short paper and a first attempt to do such a thing, we should accept this paper.)

Comments to Authors: This is a great short paper! You have taken an interesting hypothesis and tried to prove it. I've always

wondered about the impact of weather on network connectivity, for both residential connectivity as well as wireless (what with moisture absorbing my transmissions).
It was also interesting to see how you've managed to extract some of the needed signal by correlation different data sources. I think others would pick up these information sources.

The weakest part of the paper, and I suspect that you realize it already, is how do you validate such a thing. I am sympathetic here because I realize the difficulty of doing so. But are there other data sources you can pull on?

A couple of things that come to mind: publicly available information sources on failures (e.g., outages list or FCC filings); or private sources (e.g., internal ticket databases) acquired through collaboration with one of the providers.

Also, I hope that some of the issues with evidence can be handled through more sophisticated data analysis, which might require a lot more data. That would eliminate anomalies like Speakeasy and Megapath that you have to explain away. In any case, you should figure out a way to add confidence intervals to these findings, so we can judge which ones are statistically significant. (Your attempts at pulling signal out of highly noisy numbers reminds me of what sociologists and economists do often; consider talking to someone or looking up their methods.)

I wondered if Figure 4 is better plotted by providers (a group of adjacent bars per provider), rather than by weather condition. In the end, your primary goal is to enable us to visually compare the impact of weather within each provider, rather than compare the failure rate for different providers in a given weather condition.

From the introduction, why were you surprised that the median failure time for the satellite ISP was lower? Also, how (statistically) significant is that finding? (You need a lot of failures and, depending on the statistical tools used, some assumptions about the distribution of failure time.)

I did not understand why you could not use the older AWOS stations. What exact information that you need is missing? They perhaps enable a coarser (but more reliable) analysis of more data?

Reviewer #4

Strengths: It is a really fun topic, and, more importantly, it is true as well. My AT&T connection dies regularly during periods of heavy rain.

Weaknesses: There is plenty of room for improvement in the paper, which the authors acknowledge from the start. The biggest problem for me is the uncertainty as to the reason of *why* a host goes offline, which could include powering down for briefer periods. The fact that your best shot at identifying weather-related outage is the fact that bad weather was around and the outage was brief is not that great (again judging from personal experience, where the time to recovery can vary substantially, up to days). The dataset is also too thin to draw conclusions for some ISPs..

Comments to Authors: Very fun topic - I really enjoyed it.

It seems that one approach to distinguish regular machine uptime from bad weather effects would be time of day, as it is less likely that a computer remains powered up during the night but powered down for a few hours during the day.

In Table 1 I cannot distinguish the boundaries of connectivity type. Either add horizontal lines left and right of “Cable”, “DSL”, etc in order to indicate (preferred), or add vertical lines that separate the categories.

I encourage you to broaden the study to other countries. I personally have multi-year experience of three countries on two continents and can tell you that to me the sensitivity of Internet connectivity during bad weather is a uniquely American problem.

Finally, I thought I would share a weather-related story. A few years ago rain-induced outages were common enough for me that I called my ISP and asked for an engineer to come out. So one did, on a perfectly sunny day, and he promptly found nothing wrong with my DSL connection at all. He suggested I make an appointment with him for a day when the weather forecast is heavy rain... yeah right.

Reviewer #5

Strengths: To my knowledge, this is the first paper to provide quantitative evidence that weather affects residential connectivity.

Weaknesses: I do not think this is a good topic for a short paper..

Comments to Authors: My only negative comment is that this should have been a full paper: It does not introduce any new and/or interesting measurement technique. The conclusion (that weather affects connectivity) is not surprising, hence, I don't quite see the point of publishing preliminary evidence to support it. I think it would have made more sense to publish a full paper with a complete analysis of weather-induced network disruptions, most importantly try to identify the reason behind weather-induced network problems.

That said, what constitutes a good short paper is, admittedly, a subjective issue. I do acknowledge that this is the first paper to look at the impact of weather on residential connectivity, and I am sure it will create interesting discussion at the conference.

Response from the Authors

Comments addressed in the camera ready version of the paper:

We added a geographic distribution statistic for the providers that we focused on. There is a new row in the dataset description table (Table 1) that shows for each provider, the number of airports that customer IPs map to.

We clarified the reasoning behind removing failures shorter than 30 minutes and longer than 8 hours. To justify the 8 hour cutoff we included measurements of the diurnal failure pattern in the dataset.

We added two plots and text to indicate why we were surprised that the median failure time for a satellite ISP was relatively low.

In summary, we expected satellite provider failures would be longer during thunderstorms than in clear weather.

We strengthened the argument for more than one ping by providing a graph of the effectiveness of each ping and several other measurements.

Comments that we do not address in the paper:

Validating the cause of outages is a future goal for the project. Early on we tried to collect all reported power outages. However, we encountered two problems. (1) In the US, only a few power companies publish live power outage information on their website. (2) In the US it is difficult to map a geographic location to the power company that serves that location. In future work we may be monitoring social networks for reports of power outages. We also plan to correlate outages with the outages.org mailing list.

We binned weather conditions broadly because we did not have enough samples of specific conditions. The weather stations do report intensity, for example, the ASOS station reported both "heavy thunderstorms" and "light thunderstorms," for Ameritech customers. However, there were only 205 Ameritech failures observed during all thunderstorm intensities, and only 54 were during heavy thunderstorms.

We did not compute confidence intervals for the failure rate plots because we have not determined the dependence between failure rate samples; failures may be related by network links, geography and time. Although we do not quantify the confidence that failures are four times more likely in thunderstorms and two times more likely in rain, these observations are consistent across providers.

We would like to see how weather affects residential Internet connections outside of the US. We can add countries to the study only if they provide a feed of weather alerts and way to collect measurements from weather stations with precipitation type sensors.