Where on Earth are the Best-50 Time Servers?

Yi Cao and Darryl Veitch

School of Electrical and Data Engineering
University of Technology Sydney, Australia
{Yi.Cao, Darryl.Veitch}@uts.edu.au

Abstract. We present a list of the Best-50 public IPv4 time servers by mining a high-resolution dataset of Stratum-1 servers for Availability, Stratum Constancy, Leap Performance, and Clock Error, broken down by continent. We find that a server with ideal leap performance, high availability, and low stratum variation is often clock error-free, but this is no guarantee. We discuss the relevance and lifetime of our findings, the scalability of our approach, and implications for load balancing and server ranking.

Keywords: Leap second · NTP · Stratum-1 server · Network measurement · LI bits · UTC · load balancing · clock synchronization

1 Introduction

A high proportion of the global computer population achieves its time synchronization via public time servers accessed by the NTP protocol. Such servers are hierarchical in that a *Stratum-s* (or S-s) timeserver itself synchronizes to a *Stratum* $s-1$ server. Anchoring the system are the *Stratum-1* time servers, which have local access to reference hardware.

Clients rely on their server's notion of time, however, as we describe below, server quality varies in important ways, often with no warning being delivered to clients. It would clearly be of interest to map out server quality across the Internet, both for its own sake, and also to inform client server selection. However, it is not immediately clear how this could be achieved at scale, and reliably, across the latency noise of the Internet.

Recently the problem of server health monitoring has begun to receive attention, in particular regarding the small but critical Stratum-1 class. Techniques, described in [18,5], have been developed for the unambiguous detection of errors in server clock timestamps, even from vantage points where the path to the server is both long in terms of Round Trip Time (RTT), and noisy. In [18], studying around 100 servers, it was found that significant errors are not rare, being found in a surprisingly high proportion of popular public servers, including many from National Laboratories. Errors can be both large in magnitude (10's to 100's of milliseconds and even beyond) and long lasting (from hours to days and even continuously over months), or both. In [17] a similar server set was analyzed with respect to their leap second performance, and recently [5], using

a new and much larger data set, looked at both server clock error and protocol failures during the end-2016 leap second. In these servers, which include all those Stratum-1 servers employed in the widely used NTP Pool service [11], only 37.3% were found to perform adequately.

In this paper we mine the IPv4 data set, available at [4], used in [5]. We evaluate quality according to four dimensions: server Availability, behaviour surrounding a Leap Second (a stress test for both NTP protocol compliance and clock behaviour), Stratum Constancy, and finally, severity of server Clock Errors. We limit our list to 50 members, and within this group servers are not explicitly ranked. Instead, because of the importance to clients of the RTT to its server, a key factor in synchronization performance in practice (though not necessarily in theory, see [16]) due to its correlation with path asymmetry, congestion and loss, we structure our results in a per-continent then per-country breakdown.

There are a number of arguments for a 'Best-50'. One is for direct use by measurement specialists, in particular operators of measurement infrastructures [14,1,2], who require servers of both high availability and high accuracy. Another is to highlight the server health issue. Quantifying best practice increases awareness of ongoing problems, and provides the context (and an incentive) for efforts to improve the system and to track performance over time. A third goal is to explore concretely a number of quality metrics, and how they relate to actual, verifiable errors in server timing. Although there have been some papers surveying network timing performance [6,9,8,10,7], we believe this is the first attempt to accurately identify the best servers, using diverse metrics.

After providing background in Section 2 and an overview in Section 3, the main results are presented in Section 4. Section 5 discusses their significance, limitations, and implications for the definition and use of a server quality rank, with reference to load balancing services including NTP Pool. We conclude in Section 6.

2 Background

We summarize the experimental setup, data set and server list (see [5] for full details). We then summarize the operation of the NTP Pool service.

2.1 The Experiment

The experiment covered a 64 day period from Nov. 16 2016 to Feb. 2 2017, including the end-2016 leap second. For each server in a target server list in parallel, an independent instance of a request–response exchange daemon, using a per-server customized polling period as close to $\tau = 1$ seconds as possible, was launched.

For an NTP packet i which successfully completes its round-trip from the client to server and back, a 4-tuple *stamp* $\{T_{a,i}, T_{b,i}, T_{e,i}, T_{f,i}\}$ of timestamps is recorded. Here $T_{b,i}, T_{e,i}$ are the (incoming and outgoing respectively) UTC timestamps made by the server. These are extracted from the returning NTP packet header, along with the Leap Indicator (LI) bits and the server Stratum

field. The timestamps $T_{a,i}, T_{f,i}$ are of passively tapped NTP packets, hardware timestamped using high performance Endace DAG 7.5G4 capture cards, whose hardware clocks are disciplined to a rubidium atomic clock, itself locked to a roof mounted GPS receiver. The error in the client side timestamps measurement is therefore sub-microsecond and is ignored here.

The IPv4 servers studied came from five sources:

Org: the public S-1 URL list maintained at *ntp.org*
Pool: S-1 servers participating in the NTP Pool Project
LBL: S-1 servers caught at the Lawrence Berkeley Laboratory border router
Au: the set of Australian public facing S-1 servers (plus 6 private)
Misc: miscellaneous servers of interest.

The servers which returned useful data, 459 in total, are broken down by source in Table 1 (the sets overlap). Of the AU servers, 6 are in fact private and will be excluded from the final results. Table 2 provides a geographical breakdown. The low values for AF, AN and SA reflect the immaturity of Internet timing infrastructure across these continents.

Population	Org	Pool	LBL	Au	Misc
#	197	258	257	14	10
%	43	56	56	3	2

Table 1. Server Source breakdown.

Population	AF	AN	AS	EU	NA	OC	SA
#	1	0	50	203	169	29	7
%	0.2	0	0.9	44.2	36.8	6.3	1.5

Table 2. Continental breakdown of servers.

2.2 NTP Pool

The NTP Pool Project [11] provides a load balancing and convenient configuration service for millions of NTP clients, by supplying a set of URLs resolved via a tailored DNS server, to members of a pool of participating volunteer NTP servers of various strata.

Users can access at *pool.ntp.org* the complete worldwide pool, or subsets thereof at *#.pool.ntp.org*, where # is one of {0,1,2,3}. These subsets are influenced by client geo-location but otherwise random, and refresh every hour [12]. The full details of how server subsets are selected is not documented.

A degree of client-control is supported via *CONT.pool.ntp.org*: continental zone pools where CONT is one of {africa, antarctica, asia, europe, north-america, oceania, south-america}, and CY-coded country pools at *CY.pool.ntp.org*, and *#.* prefixed subsets of these [13].

For the pool associated to a given client at a particular time, the system uses DNS round robin to resolve URL queries to the IP address of a server in that pool. NTP Pool includes a monitoring system which queries the pool servers, scoring their performance based in NTP packet fields including {offset, stratum, LI, RTT, noresponse}. Servers are evaluated periodically and only those with a *score* above 10 are made available.

3 Server Characterization

We characterize servers according to the following four criteria or dimensions.

Availability This simple but critical criterion is measured by the ratio of response packets received to request packets sent. This will underestimate the true availability, because of packet loss and reachability failure in the network.

Stratum Constancy Possible stratum values range from $S = 0$ (unsynchronized), to $S = 1, 2 \ldots 16$. A Stratum-1 server may change stratum if its hardware reference has a problem, if the system has a reboot, or if its synchronization daemon/algorithm decides it would prefer an remote reference, and stratum values of 0, 2, 3 or even higher could result. We measure the 'Stratum-1 downtime' (S1Downtime) as the proportion of response packets which report a stratum other than 1. Values of S1Downtime close to zero suggest a well managed Stratum-1 server in a stable environment. We also record the list of all stratum values ever seen.

Leap Performance Leap Second events are a stress test for servers, both in terms of the detailed clock performance (does it jump cleanly by exactly 1 second at exactly the right time, and nothing else?) and protocol compliance (does it set the LI bits in accordance with the standard?). This question was studied in detail for each server in the list in [5]. Here we classify servers according to a subset of the characterization defined there, as:

Ideal: no observed clock error linked to the leap second, ideal protocol behaviour;
Adequate: no clock error, compliant protocol behaviour;
Clock Good: no evidence of clock error about the leap,

where Ideal $\subset$ Adequate $\subset$ Clock Good $\subset$ All. For convenience, we add two more classes by set difference:

Clock Good Only (CGO): Clock-Good$\setminus$Adequate;
Clock Not Good (CNG): All$\setminus$Clock-Good.

Although leap seconds are rare, they occur regularly. If a server handles them poorly, the impact can be severe, for example taking weeks to jump, or never.

Clock Errors/Anomalies Our approach is based on the methodology we pioneered in [18] for the remote detection and measurement of server errors. It uses baseline analysis of the RTT timeseries to identify changes in the 'Error' time series $E_i = (D_i^{\uparrow} - D_i^{\downarrow})/2$ due to server errors, rather than the alternatives of path routing changes and/or congestion. Here $D_i^{\uparrow} = T_{b,i} - T_{a,i}$ and $D_i^{\downarrow} = T_{f,i} - T_{e,i}$

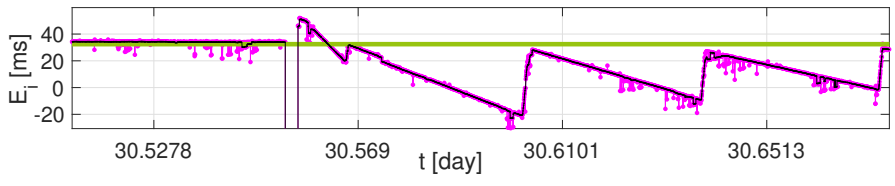

Fig. 1. Server errors cause $E(i)$ to deviate from its true underlying value (green line).

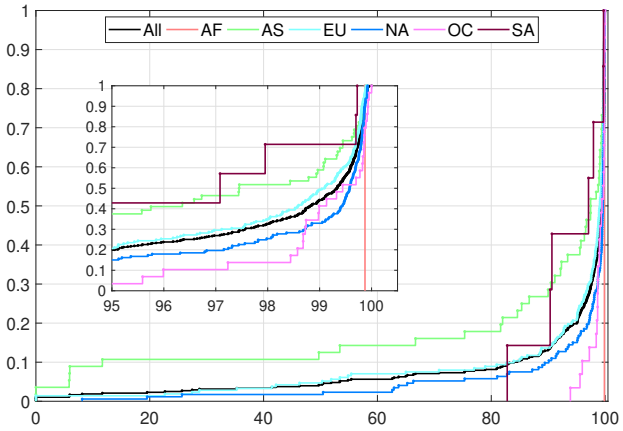

Fig. 2. CDF of Availability (in %) over all servers (black), and per-continent.

are the empirical outgoing and incoming delays to the server. An example of a server error zone, beginning at around $t = 30.544$ days, is given in Figure 1.

We have improved the methodology of [18] by (i) replacing non-linear filtering based congestion suppression (which can be fooled in certain circumstances) with strict RTT bounding, (ii) systematically recording not only error sizes but also the precise locations of all error zones, (iii) increasing the granularity of error frequency reporting: we classify servers according to the number of errors as: **G**ood: no errors; **R**are: less than one error per week; **C**ommon: more than one error per week, but not **H**igh; and **H**igh: continuous stretches of error covering at least 25% of the trace. In [18] **R** and **C** were combined into **R**.

Since the selection of error zones is performed manually (due to the need to disambiguate from complex routing, congestion and error scenarios), the detection process is very labor intensive. It is essential however for our purposes here where, unlike [18], we evaluate not only error presence and representative size but also how often the server is in error (see *Errtime* below).

3.1 Server Overview

We provide some context by examining the first three of the above dimensions over all servers.

Figure 2 shows the Cumulative Distribution Function (CDF) of availability for all servers. Availability is good overall, with 80% of servers having values exceeding 95%, and over half exceeding 99%. The per-continent results show lower availability for regions further from the testbed in Sydney, Oceania. This can be explained through a measurement bias due to higher loss rates over longer paths leading to lower apparent availability.

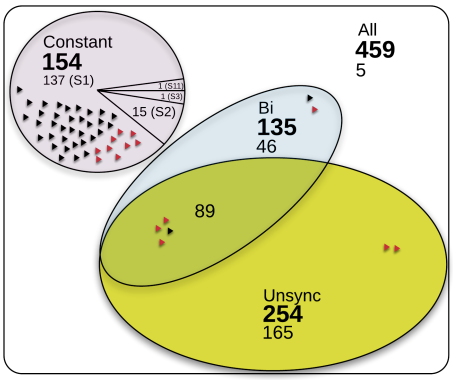

Fig. 3. Relationship between the Stratum classes. Symbols denote servers in the Best-50, red symbols denote those with server errors.

The leap performance results over all servers appear in Table 3. Only 37% exhibit Adequate behavior, necessary to allow their clients to navigate a leap second without incident.

	All	CGO	CNG	Clock Good	Adequate	Ideal
#	459	134	154	305	171	36
%	100	29	34	66	37	8

Table 3. Leap Performance summary.

Figure 3 provides a pertinent classification of servers according to strata. In the Constant class only one stratum value is ever seen (not always Stratum-1!), in Bi only two, and in Unsync at least one response carries Stratum-0. We see that 154 servers (34%) have constant strata, and the majority of the 305 that do not, 254 or 83%, announced themselves as unsynchronized at least once.

Overall 137 servers (30%) announce themselves as Stratum-1 in each and every response. This appears as a discrete mass of weight 0.3 at the origin in the S1Downtime CDF in Figure 4, which shows that servers which are not Constant have a wide variety of S1Downtime values.

4 The Best-50 Servers

What we would ideally like is clear: to find servers that are always available, and that have no detectable clock errors. However, to determine the latter implies a prior detailed examination, which is too labour intensive using our server error methodology and tools to deal with 459 servers, each with up to 2 months of high resolution data, each with potentially a large number of errors.

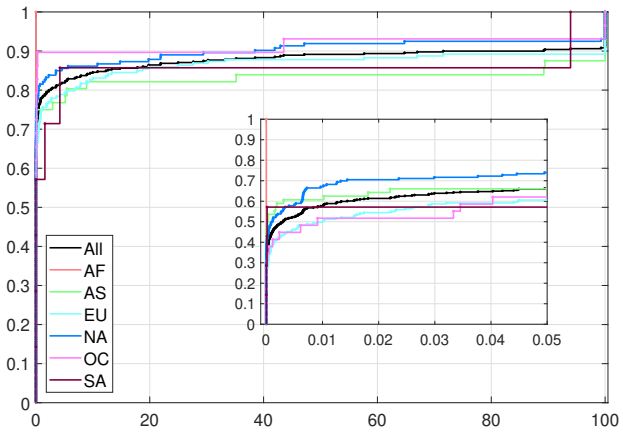

Fig. 4. CDF of S1Downtime (in %) over all servers (black), and per-continent.

Accordingly, our approach is to first assemble a list of ostensibly high quality servers using the dimensions of Availability and Stratum Constancy that are readily calculated, and Leap Performance, available from prior work, and to apply the Clock Error analysis on this much smaller number of servers, which moreover are likely to be simpler to analyse. In this way we approximate the ideal above in a scalable way (see Section 5), with a practically appropriate bias toward servers with stable management (high Stratum Constancy) and competent configuration and performance during high stress (Leap Performance).

More precisely we proceed as follows. For Availability, we seek servers that are almost always available, with due allowance for measurement bias due to packet loss. Based on Figure 2 we believe a cutoff of 97% is safe. For Leap Performance, we insist that servers are in the Adequate class. Next, we use S1Downtime to order the servers that pass the above two criteria. Our Best-50 servers are then defined as the first 50 servers in this ordering (starting from the zero S1Downtime end) whose Clock Error class is either **G** or **R**.

Server errors in a given server are further quantified through the metrics of **Size** (the median over all error zones of the error range over that zone), and **Errtime** (the proportion of the trace taken up by error zones).

The resulting Best-50 servers are given in Table 4. Within each continent group, servers are ordered according to country code first, and then lexicographically according to their URL. The mapping from URL to IP address is provided in the Appendix.

Beyond the identities of the servers themselves and their geographical breakdown, the most important observation from the table is the fact that even excellent performance under each of Availability, Stratum Constancy and Leap Performance does not mean that the server is error free. Indeed, out of 15 servers with detected server errors, 9 give no warning of this with a S1Downtime of zero, yet have Sizes ranging from 2.1 to 1000 ms, albeit with Errtime being generally low (0.9 seconds in the hour on average in the worst case of 0.025%). The worst

CONT	URL	CY	Strata		Server Error			Avail. (%)	Leap Perf
			List	S1Down time(%)	Class	Size [ms]	Errtime (%)		
AF	stratum1.neology.co.za	ZA	{1}	0	R	2.1	7.0e-5	99.87	Adeq.
AN	–	–	–	–	–	–	–	–	–
OC	ntp1.net.monash.edu.au	AU	{1}	0	R	180	1.4e-4	99.86	Adeq.
EU	ntp1.oma.be	BE	{0,1}	2.9e-4	R	28	0.032	99.04	Adeq.
	ntp.freestone.net	CH	{1}	0	G	–	–	99.80	Ideal
	netopyr.hanacke.net	CZ	{1}	0	G	–	–	99.25	Ideal
	ntp.nic.cz	CZ	{1}	0	G	–	–	99.86	Adeq.
	ptbtime1.ptb.de	DE	{0,1}	2.9e-4	R	1.1	2.9e-4	99.78	Adeq.
	ptbtime3.ptb.de	DE	{1}	0	R	5.46	0.014	99.78	Ideal
	hora.roa.es	ES	{0,1,2}	2.9e-4	R	120	5.8e-3	99.40	Adeq.
	ntp.i2t.ehu.es	ES	{1}	0	G	–	–	98.94	Ideal
	unknown1	GB	{1}	0	G	–	–	99.71	Ideal
	unknown2	GB	{1}	0	G	–	–	99.71	Ideal
	ntp2.litnet.lt	LT	{1}	0	G	–	–	99.87	Ideal
	metronoom.dmz.cs.uu.nl	NL	{1}	0	G	–	–	99.66	Ideal
	unknown3	NO	{1}	0	G	–	–	98.88	Ideal
	goblin.nask.net.pl	PL	{1}	0	G	–	–	99.79	Ideal
	ntp.certum.pl	PL	{1}	0	R	7.0	0.025	97.55	Adeq.
	ntp.fizyka.umk.pl	PL	{1}	0	G	–	–	99.45	Ideal
	time.assecobs.pl	PL	{1}	0	G	–	–	99.10	Ideal
	ntp1.niiftri.irkutsk.ru	RU	{1}	0	G	–	–	98.83	Ideal
	ntp2.niiftri.irkutsk.ru	RU	{1}	0	G	–	–	98.94	Ideal
	ntp1.gbg.netnod.se	SE	{1}	0	R	1000	1.8e-5	99.89	Adeq.
	ntp2.gbg.netnod.se	SE	{1}	0	R	1000	1.8e-5	99.89	Adeq.
	ntp1.mmo.netnod.se	SE	{1}	0	R	1000	3.6e-5	99.87	Adeq.
	ntp2.mmo.netnod.se	SE	{1}	0	G	–	–	99.88	Adeq.
	ntp1.sth.netnod.se	SE	{1}	0	G	–	–	99.82	Adeq.
	ntp2.sth.netnod.se	SE	{1}	0	R	1000	8.8e-4	99.81	Adeq.
NA	istntpprd–02.corenet.ualberta.ca	CA	{1}	0	G	–	–	99.89	Ideal
	tick.usask.ca	CA	{1}	0	G	–	–	99.86	Adeq.
	tock.usask.ca	CA	{1}	0	R	17	2.5e-4	99.58	Adeq.
	clepsydra.dec.com	US	{1}	0	G	–	–	97.82	Ideal
	m4c2236d0.tmodns.net	US	{1}	0	G	–	–	99.87	Ideal
	m4d2236d0.tmodns.net	US	{1}	0	G	–	–	99.88	Ideal
	montpelier.ilan.caltech.edu	US	{1}	0	G	–	–	99.76	Ideal
	navobs1.gatech.edu	US	{1}	0	G	–	–	99.70	Adeq.
	ntp.colby.edu	US	{1}	0	G	–	–	99.71	Ideal
	ntp1.digitalwest.net	US	{1}	0	G	–	–	99.82	Ideal
	tick.ucla.edu	US	{1,2}	2.6e-4	G	–	–	99.50	Adeq.
	time–a.netgear.com	US	{1}	0	G	–	–	99.78	Ideal
	time–a.stanford.edu	US	{1}	0	G	–	–	99.92	Adeq.
	tock.phyber.com	US	{1}	0	G	–	–	99.87	Adeq.
	usatl4-ntp-002.aaplimg.com	US	{0,1,2}	5.7e-5	R	1.5	0.063	99.83	Adeq.
	usno.hpl.hp.com	US	{1}	0	G	–	–	97.82	Ideal
	usnyc3-ntp-003.aaplimg.com	US	{0,1}	1.8e-3	R	6.4	0.052	99.85	Adeq.
AS	f2.kns1.eonet.ne.jp	JP	{0,1}	2.8e-4	G	–	–	99.83	Adeq.
	jptyo5-ntp-001.aaplimg.com	JP	{1,2}	2.3e-4	R	39	0.029	99.11	Adeq.
	ntp1.noc.titech.ac.jp	JP	{1}	0	G	–	–	99.82	Adeq.
	ntp-b2.nict.go.jp	JP	{1}	0	G	–	–	99.90	Ideal
	unknown4	SG	{1}	0	G	–	–	99.91	Ideal
SA	ntp.shoa.cl	CL	{1}	0	G	–	–	99.70	Ideal

Table 4. Best-50 public timeservers organised by continent, country, and URL. Cyan URLs marks National Laboratory servers.

CONT	URL	CY	Strata		Server Error			Avail. (%)	Leap Perf
			List	S1Down time(%)	Class	Size [ms]	Errtime (%)		
OC	ntp10.net.monash.edu.au	AU	{1}	0	**C**	18.46	0.002	99.86	Adeq.
NA	time-a.timefreq.bldrdoc.gov	US	{1}	0	**H**	23.16	100	99.47	Adeq.
NA	time-c.timefreq.bldrdoc.gov	US	{1}	0	**H**	8.98	100	99.69	Adeq.
OC	ntp.waia.asn.au	AU	{0,1,3}	**0.040**	R	700	0.128	99.44	Adeq.
EU	ntp1.fau.de	DE	{1,2}	**0.381**	R	1.76	0.628	99.70	Adeq.
NA	srcf-ntp.stanford.edu	US	{1}	0	G	–	–	99.93	**CGO**
SA	a.st1.ntp.br	BR	{0,1}	1.1e-4	G	–	–	99.72	**CGO**
EU	ntp1.vniiftri.ru	RU	{0-3,12}	**0.029**	R	2.30	1.852	98.05	**CNG**
EU	ntp3.fau.de	DE	{1,2}	**0.401**	**H**	6.3	100	99.69	Adeq.
NA	ntp.myfloridacity.us	US	{0,1}	3.9e-4	**H**	14.61	100	98.73	**CNG**
NA	time-b.nist.gov	US	{1}	0	**C**	2.10	0.254	**63.73**	Adeq.
NA	t2.timegps.net	US	{0,1,2}	**0.011**	R	333.50	0.043	99.59	**CGO**
EU	rustime01.rus.uni-stuttgart.de	DE	{1,2}	**0.380**	R	4.50	3.485	**95.05**	**CGO**
EU	ntp2.usv.ro	RO	{0,1}	**0.003**	G	–	–	**96.70**	**CNG**

Table 5. Five categories of examples of servers outside the Best-50 in one or more criteria. Bold column entries mark failed criteria.

S1Downtime in the table, NA server *usnyc3-ntp-003.aaplimg.com*, which is also an **R** server, only drops from Stratum-1 (to Stratum-0 in this case) 0.0018% of the time. This is 29 times less often than its Errtime at 0.052%. Thus for this server, error is a more serious concern than stratum stability.

The Best-50 are marked via symbols within Figure 3, where certain observations are more immediate. For example we clearly see that 9 of the Constant S1 servers in the Best-50 have clock errors, and that only 2 in the Best-50 take 3 or more stratum values.

Another observation of note is that, with the exception of *ptbtime3.ptb.de*, servers with Ideal Leap Performance and zero S1Downtime enjoy Server Error ratings of **G**, suggesting that this pair could serve as a useful indicator of an exceptionally well managed server, and hence be predictive of exemplary Error behaviour. Useful does not mean foolproof however: in addition to the exception above the two NIST servers in Table 5 provide sobering counter-examples.

The server list contains 35 servers from Apple's *17.253* domain. Three of these make it into the Best-50, though all exhibit server errors with relatively large Errtime values. Finally, it is worth noting that despite having 66 servers from National Laboratories in the list, only 12, those colored cyan, make it into the Best-50 (an additional 5 from the NMI in Australia are excluded as they are not publicly accessible).

Because the criteria of entry into the Best-50 are so strict, there is a limit to what one can say about these servers: they are indeed very well behaved. However, if one relaxes the criteria in different dimensions, a much wider variety of behaviour is quickly revealed. To make this concrete, and to indicate what could have been included in the Best-50 had things been a little different, a number of contrasting examples are provided in Table 5, separated into five

categories. For each server bolded column entries mark the criteria which did not meet the Best-50 standard.

In the first category we give 3 of the 5 servers (of which {2,3} were rated {**C**,**H**} respectively) that failed to make the Best-50 because of excessive server errors. By definition, and as noted earlier, such servers illustrate the fact that the (Availability, Stratum, Leap) three-tuple is not sufficient to predict the absence or otherwise of clock errors, nor their severity in terms of Size or Errtime. Particularly noteworthy is the fact that **H** servers, which by definition have an Errtime over 25%, and typically have Errtime of a dramatic 100%! can and do appear. The second category exhibits two examples of servers that failed only due to being too low in the S1Downtime ranking, one of which has Size of 700 ms and Errtime three times higher than its S1Downtime. The third category gives examples failing only the Leap criterion, that are exemplary in other respects. There were no examples of servers which failed in Availability only. The fourth category includes five diverse examples where two criteria were not met. Finally, the fifth category includes servers that are still generally respectable despite failing in three criteria.

5 Discussion

We discuss the limitations, implications and future of our work.

Source Coverage Because of the widespread usage of the Pool service, and the high profile of the Org list, we expect the server list to contain most of the widely used public S-1 servers, but how representative are they of the (unknown) complete set? There is in fact a high degree of overlap, 50% or more, between each of the three main sources: Org, Pool and LBL, leading to speculation in [5] that the server list contains a significant percentage of the global public facing Stratum-1 server population. We now consider how to evaluate this claim.

Population estimation based on re-sampling a marked sub-population is known as the *capture-recapture* problem in statistics. To fit within this framework, it is natural to group the Org and Pool sources together as they are both community based, and have a strong, non-random relationship. Thus we have $n = |\mathbf{Org} \cup \mathbf{Pool}| = 356$ servers which represent a 'marked' sample of the total unknown population N. The LBL source now represents a random sample of $K = 257$ servers, of which $k = 175$ lie in $\mathbf{Org} \cup \mathbf{Pool}$, that is they are marked servers that are 'recaptured'. The population can now be estimated from n, K and k. For example the Chapman estimator [3,15], yields $\hat{N} = \lfloor (K+1)(n+1)/(k+1) \rfloor - 1 = 522$. A corresponding (non-symmetric) 95% coverage interval for N is $[497, 562]$. This suggests that our Best-50 is well founded as it is based on a number, 453, being between 80% and 91% of all public servers.

The random sampling assumptions underlying the Chapman estimator do not hold strictly here, so the above estimate can only be viewed as a rough indication. To determine the true value of N a better approach, for IPv4 servers, is simply to exhaustively probe the IPv4 address space. We did not do so here, as that would not have given us the leap second performance information we require.

List Shelf Life As it derives from a static data set, the utility of our Best-50 will decrease over time. Some indication of its expected lifetime can be gained from the longitudinal results in [18], which report on a subset of Org servers using data collected over 151 days in 2011-12 (Exp1), and 124 days in 2014-15 (Exp2). Although Availability, Leap performance, and Errtime are not given, we can compare with respect to Stratum Constancy, and Error Classification.

Of the Best-50 servers, there are 13 which also appear in that study. All 13 (100%) were found to be error-free in each of Exp1 and Exp2, as well as having zero S1Downtime for Exp2 (stratum data was unavailable for Exp1). For the metrics available, this represents perfect agreement.

Of the 14 servers which feature in Table 5, 13 also appear in the study, of which 3 are suitable for direct comparison as they pass our criteria for S1Downtime and have Error class in $\{\mathbf{G}, \mathbf{R}, \mathbf{C}\}$. Of these, all 3 exhibit close agreement, with no detected errors in each of Exp1 and Exp2, and again with zero S1Downtime. Finally, at the other end of the spectrum, of the 4 servers in the continuously errored H class in Table 5, 3 were also classed as H in [18].

Based on the above, we expect that the level of churn in the Best-50 list provided here will be low on useful timescales, for example 5 years. Knowledge of server configuration would be of interest here also to attempt root cause analysis, as would correlating against network failures. We have attempted to contact administrators, however the response rate was minimal.

Measurement Cost The analysis used here requires specialist hardware, techniques, unusual data (leap events), and significant effort. A priori, this does not scale. A goal of future work must be to develop lighter weight approximate techniques and more automated server error detection using standard hardware. The work here can serve to evaluate the effectiveness of such techniques.

Scalability cost divides substantially along criteria lines. Stratum Constancy measurement scales trivially, as it depends neither on special hardware nor the network path. Availability also scales readily, though to remove packet loss bias requires measurement close to the server and/or path diversity, and hence client placement diversity ideally. Leap Performance is inherently difficult as opportunities to measure it occur only every $\approx$2.5 years. On the other hand this also limits the workload, and the protocol aspects are as scalable as Stratum Constancy. Rankings could be defined which exclude leap second criteria for applications where this is not needed, for example Internet measurement campaigns not covering leap events, which are announced months in advance.

The Clock Error criteria is the expensive one, and the most critical. The hardware cost could be avoided by using a robust clock synchronization and timestamping approach such as RADclock [16] as a Stratum-2, with its Stratum-1 server selected from the Best-50 provided here. Although timestamping errors would of course be higher, they would still be well below server error sizes in most cases. In terms of the error analysis itself, it is feasible, albeit non-trivial, to automate this to a good level of accuracy, and this is a direction for our future work. Such a capability would enable, for example, ongoing monitoring and error querying for important servers. However, this is not essential for the purpose of

maintaining the Best-50 as we have defined it here, as the construction of the list, combined with its expected low churn, implies that only a small number of high quality servers (which are faster to process) would have to be evaluated from scratch each year to keep it current. Those remaining would also have to be re-evaluated, but this is less onerous when they have been seen before.

Server Ranking From the quality dimensions we have considered various rankings could be defined. An obvious way to rank the Best-50 is the S1Downtime ordering employed in the list construction, however this cannot be extended over all servers, as many will not satisfy the minimum requirements in other criteria. A candidate which avoids this problem is *Badtime*, defined to be the sum of Errtime and 1−Availability, being the proportion of time a server should be avoided. This should suit contexts where leap second performance is not critical.

Great care must be taken in how any ranking is used, to prevent high ranking servers from receiving high loads. It would be a mistake (and is not the intent of this paper!) to recommend that clients make use of the Best-50 en masse. Instead, server rank should be used within broader systems designed to tradeoff load balancing and server quality appropriately. Indeed, NTP Pool's *score* is an attempt to do this (Section 2), however it is not grounded in knowledge of actual server error. The larger problem is that NTP Pool breaks NTP's inherent load balancing mechanism, namely the server hierarchy, while simultaneously preferencing its own load balancing over server quality. Thus pools contain servers of mixed strata, and clients are given different servers over time with quality which may vary enormously. Instead, we argue that the hierarchy needs to be enforced, and within that, well defined notions of rank given higher prominence.

Client Impacts Finally, a separate, but natural question to ask is, how important is it for a client to select a server of Best-50 calibre? The client impact will depend strongly on many factors including the robustness of the clock synchronization algorithm in use, the policy regarding back-up servers and if they are available, the size of server errors, their duration, the length of non-availability periods, the stratum of the client, the characteristics of the path to the server, and whether a leap second is involved. Potential errors can range from negligible ($< 10\,\mu s$) and short-term (few seconds) at one extreme, to permanent (until server change) and extreme (10's of ms to seconds or well beyond plus high variability) at the other. The onus on the Stratum-1 server is to show near perfect behaviour to anchor and lift performance across the timing system. This is possible, as many in the Best-50 demonstrate.

6 Conclusion

Our Best-50 list is not definitive. It is however the first serious attempt to quantify timeserver best practice that we are aware of. We believe that it will be of use for a number of years at least, by which time the methodology could be improved to make such a list more comprehensive, dynamic and less expensive to generate. It is in any event, feasible to maintain it even with current technology.

Acknowledgment

Partially supported by Australian Research Council's Discovery Projects funding scheme #DP170100451.

References

1. Archipelago monitor locations. http://www.caida.org/projects/ark/locations/.
2. V. Bajpai and J. Schnwlder. A survey on internet performance measurement platforms and related standardization efforts. *IEEE Communications Surveys Tutorials*, 17(3):1313–1341, thirdquarter 2015.
3. S. Brittain and D. Böhning. Estimators in capture–recapture studies with two sources. *AStA Advances in statistical analysis*, 93(1):23–47, 2009.
4. Y. Cao and D. Veitch. TimeServer Dataset 2016-2017. *https://data.research.uts.edu.au/public/DVTSD/*. Per-server results also available.
5. Y. Cao and D. Veitch. Network Timing, Weathering the 2016 Leap Second. In *Proc. of IEEE INFOCOM 2018*, Honolulu, USA, April 15-19 2018.
6. J. D. Guyton and M. F. Schwartz. Experiences with a Survey Tool for Discovering Network Time Protocol Servers, 1994. [Online; accessed 31-July-2015].
7. C.-Y. Hong, C.-C. Lin, and M. Caesar. Clockscalpel: Understanding root causes of Internet clock synchronization inaccuracy. In N. Spring and G. Riley, editors, *Proceedings of the 12th international conference on Passive and active measurement*, volume 6579 of *PAM'11*, pages 204–213, Berlin, Heidelberg, 2011. Springer-Verlag.
8. D. Malone. The Leap Second Behaviour of NTP Servers. In *Proceedings of the Traffic Monitoring and Analysis workshop*. IFIP Digital Library, April 7-8 2016. http://tma.ifip.org/2016/#program.
9. N. Minar. A Survey of the NTP Network, 1999. http://xenia.media.mit.edu/ nelson/research/ntp-survey99/ntp-survey99-minar.ps.
10. C. Murta, P. Torres, and P. Mohapatra. Characterizing quality of time and topology in a time synchronization network. In *Global Telecommunications Conference, 2006. GLOBECOM '06. IEEE*, pages 1–5, Nov 2006.
11. ntp.org. NTP Pool Project, 2018. [Online; accessed 2-May-2018].
12. pool.ntp.org. How do I use pool.net.org? http://www.pool.ntp.org/en/use.html.
13. pool.ntp.org. NTP Pool server selection. http://support.ntp.org/bin/view/Servers/NTPPoolServers.
14. R. N. Staff. Ripe atlas: A global internet measurement network. *Internet Protocol Journal*, 18(3), 2015.
15. W. J. Sutherland. *Ecological census techniques: a handbook*. Cambridge University Press, 2006.
16. D. Veitch, J. Ridoux, and S. B. Korada. Robust Synchronization of Absolute and Difference Clocks over Networks. *IEEE/ACM Transactions on Networking*, 17(2):417–430, April 2009.
17. D. Veitch and K. Vijayalayan. Network Timing and the 2015 Leap Second. In *Proc. of PAM 2016*, Heraklion, Crete, Greece, March 31 - April 1 2016.
18. K. Vijayalayan and D. Veitch. Rot at the Roots? Examining Public Timing Infrastructure. In *Proc. of IEEE INFOCOM 2016*, San Francisco, CA, USA, April 10-15 2016.

Appendix

CONT	URL	IP	CY
AF	stratum1.neology.co.za	41.73.40.11	ZA
AN	–	–	–
OC	ntp1.net.monash.edu.au	130.194.1.96	AU
EU	ntp1.oma.be	193.190.230.65	BE
	ntp.freestone.net	193.5.68.2	CH
	netopyr.hanacke.net	94.124.107.190	CZ
	ntp.nic.cz	217.31.202.100	CZ
	ptbtime1.ptb.de	192.53.103.108	DE
	ptbtime3.ptb.de	192.53.103.103	DE
	hora.roa.es	150.214.94.5	ES
	ntp.i2t.ehu.es	158.227.98.15	ES
	unknown1	188.39.213.7	GB
	unknown2	81.187.202.142	GB
	ntp2.litnet.lt	193.219.61.120	LT
	metronoom.dmz.cs.uu.nl	131.211.8.244	NL
	unknown3	148.252.105.132	NO
	goblin.nask.net.pl	195.187.245.55	PL
	ntp.certum.pl	213.222.200.99	PL
	ntp.fizyka.umk.pl	158.75.5.245	PL
	time.assecobs.pl	195.189.85.132	PL
	ntp1.niiftri.irkutsk.ru	46.254.241.74	RU
	ntp2.niiftri.irkutsk.ru	46.254.241.75	RU
	ntp1.gbg.netnod.se	192.36.133.17	SE
	ntp2.gbg.netnod.se	192.36.133.25	SE
	ntp1.mmo.netnod.se	192.36.134.17	SE
	ntp2.mmo.netnod.se	192.36.134.25	SE
	ntp1.sth.netnod.se	192.36.144.22	SE
	ntp2.sth.netnod.se	192.36.144.23	SE
NA	istntpprd–02.corenet.ualberta.ca	129.128.5.211	CA
	tick.usask.ca	128.233.154.245	CA
	tock.usask.ca	128.233.150.93	CA
	clepsydra.dec.com	204.123.2.5	US
	m4c2236d0.tmodns.net	208.54.34.76	US
	m4d2236d0.tmodns.net	208.54.34.77	US
	montpelier.ilan.caltech.edu	192.12.19.20	US
	navobs1.gatech.edu	130.207.244.240	US
	ntp.colby.edu	137.146.28.85	US
	ntp1.digitalwest.net	72.29.161.5	US
	tick.ucla.edu	164.67.62.194	US
	time–a.netgear.com	209.249.181.52	US
	time–a.stanford.edu	171.64.7.105	US
	tock.phyber.com	207.171.30.106	US
	usatl4-ntp-002.aaplimg.com	17.253.6.253	US
	usno.hpl.hp.com	204.123.2.72	US
	usnyc3-ntp-003.aaplimg.com	17.253.14.123	US
AS	f2.kns1.eonet.ne.jp	60.56.214.78	JP
	jptyo5-ntp-001.aaplimg.com	17.253.68.125	JP
	ntp1.noc.titech.ac.jp	131.112.125.48	JP
	ntp-b2.nict.go.jp	133.243.238.163	JP
	unknown4	210.23.25.77	SG
SA	ntp.shoa.cl	200.54.149.24	CL

Table 6. URL to IP mapping of the servers in Table 4.

CONT	URL	IP	CY
OC	ntp10.net.monash.edu.au	130.194.10.150	AU
NA	time-a.timefreq.bldrdoc.gov	132.163.4.101	US
NA	time-c.timefreq.bldrdoc.gov	132.163.4.103	US
OC	ntp.waia.asn.au	218.100.43.70	AU
EU	ntp1.fau.de	131.188.3.221	DE
NA	srcf-ntp.stanford.edu	171.66.97.126	US
SA	a.st1.ntp.br	200.160.7.186	BR
EU	ntp1.vniiftri.ru	89.109.251.21	RU
EU	ntp3.fau.de	131.188.3.223	DE
NA	ntp.myfloridacity.us	71.40.128.146	US
NA	time-b.nist.gov	129.6.15.29	US
NA	t2.timegps.net	69.75.229.43	US
EU	rustime01.rus.uni-stuttgart.de	129.69.1.153	DE
EU	ntp2.usv.ro	80.96.120.252	RO

Table 7. URL to IP mapping of the servers in Table 5.