Aids to the Presentation and Analysis of WSPR Spots: TimescaleDB database and Grafana

Gwyn Griffiths, G3ZIL and Rob Robinett, AI6VN

For more information visit wsprdaemon.org
or email gwyn@autonomousanalytics.com

With enormous thanks to Gary, Corrie and the team providing wsprnet.org

Speaking notes for talk at TAPR/ARRL Digital Communications Conference, Sept 2020.

Slide 1

Hello, I'm Gwyn Griffiths. Rob Robinett and I would like to acknowledge, at the outset, the sterling efforts of Gary, Corrie and the team that delivers the wsprnet.org service. Our work builds on their solid foundation.

This talk will show, with examples, how a modern time-series database and visualization using Grafana can be most helpful to the amateur community using the WSPR digital mode.

Slide 2

I'll cover the following topics: our motivation for this work and how it has changed over the last year since Rob presented early ideas at the 2019 DCC at Detroit. I'll summarize our use of the Timescale database and our current hardware and then delve into several examples of uses, which after all, is why we have created this tool. And I'll end with some thoughts on next steps, recalling that this is essentially a two-person part-time effort.

Slide 3

The initial idea from Rob was for a robust software program - WsprDaemon - to report WSPR spots from one or more multichannel KiwiSDR receivers - initially up to 8 channels per receiver and now 14. Glenn Elmore and others saw the advantages of estimating noise at the same time that WSPR estimates SNR to get at the signal level. I joined the team, contributed and tested two noise estimation algorithms and by the 2019 DCC Rob's friend Tommy Nourse had implemented an Influx database for noise with Grafana providing a simple one-panel graph.

With that success under our belt we thought it useful to add spots to an Influx database and learn how to do it ourselves so that we could add fields not available from wsprnet.org, such as azimuth at the receiver and latitude and longitude as numbers.

The story of why we had to move from Influx to Timescale is set out in our paper. In summary, Influx's indexing method results in long query times for WSPR data. With the move to Timescale completed Ron invested in the hardware needed to gather spots from all users of wsprnet - not just users of WsprDaemon - in part to help our own long-term studies but also to help reduce the load on wsprnet.org from 3rd party apps.

Slide 4 (edited out of the video)

This is a graphical summary of where we were in September 2019. WsprDaemon was up and running. It sent data from users to wsprnet.org via HTTP PUT and sent noise estimate data to Tommy's cloud-hosted Influx database with a single Grafana panel.

Slide 5

This is a summary of today's WsprDaemon. The route for WsprDaemon users' data to wsprnet.org is essentially unchanged. Now, their spot data to the WsprDaemon server is cached locally - and uploaded later if an outage happens at either end - the dataset now includes all the parameters from the Franke Taylor WSPR decoder program wsprd.

A separate database table receives the noise estimates and another table holds all the spot data obtained via an API from wsprnet.org, with derived variables added.

An FTP mirror sends WsprDaemon data to a backup server, which has its own API path to wsprnet.org.

The main server allows client connections to users and third-party apps.

We also bring in additional data, e.g. the Kp geomagnetic disturbance index from NOAA's Space Weather Prediction Center.

Slide 6 (edited out of the video)

As some of our studies need access to on-line data from further back than the two-weeks available at wsprnet.org Rob has made the necessary investments in hardware. Timescale is essentially a set of extensions to the very well established postgreSQL open source database.

Timescale uses the concept of chunks of time - and a chunk should not occupy more than 25% of available RAM. This works out at 48GB for the 192GB in the WsprDaemon server. This can easily cope with chunks of 30 days without needing to page to/from disk.

Older chunks are stored on a Solid-State Disk up to a year - after that we will migrate data to a 7.3TB RAID5 that should hold data for 11 years - a whole sunspot cycle.

Slide 7

Currently, and to our knowledge, we are serving data to two well-known apps. First, the website of VK7JJ, Phil Barnard, where he has added options to construct advanced searches - essentially a form-filling query builder - to take full advantage of a direct postgreSQL interface to WsprDaemon. In turn, Phil has several hundred regular users - now hitting WsprDaemon rather than wsprnet.org.

Another app is WSPRwatch from VK2TPM, Peter Marks, available for iOS with a nice range of presentation graphics.

Slide 8

Let us move on to Grafana so that we can look at data we've collected. To be honest, I do not find Grafana's to be the easiest of the most intuitive of user interfaces, but it has got better with the recent release of Version 7. I've written a guide available on the WsprDaemon website at http://wsprdaemon.org/grafana.html. It does not help that is uses a very small font.

Here are the elements you would encounter if you were about to construct a new 'Dashboard' - Grafana jargon for a collection of 'Panels' that comprise graphics of one type or another.

Bottom left is a form-filling query builder that knows the syntax for your data source - or you can select text entry and type in raw postgreSQL. You can add additional queries to Panels, and you can add additional Panels to your Dashboard.

On the right is where you customize your graphs, e.g. select dots rather than lines, set limits for your y-axes and choose the type of graphic.

Top right, Save is all to close to Discard!

Slide 9

And so to data ... On the basis that plots in academic journals are an indication of the form of useful data - let us see if we can obtain an equivalent from our Timescale database and Grafana.

This plot at top left is from McNamara et al. in Radio Science in 2008. It shows the signal power in dB Watts of the Ottawa, Canada time signal station CHU, then on 7.335MHz, received at Boston, a path of 476km, over a 24 period on 1 September 2003. Time is in UTC.

We see a nighttime trough, but with nice examples of sporadic E lifting the signal level. McNamara commented on the daytime minimum around local noon, 1700UTC, from D layer absorption, and the deep fading during the day - polarisation fading.

Our top Grafana panel shows WSPR SNR at KD2OM upstate New York from transmissions at 7.04MHz from N1ZPY in Maine, a path of 798km. Here we show two days, 24 and 25 August 2020. We've also caught some night time sporadic E on both days.

For the middle panel we've performed a join between spots and noise tables to estimate the signal level, the Y axis between -160 and -60dBm. We also see a minimum in signal level around 1700UTC due to the peak in absorption.

The bottom panel shows the noise estimates we have used in going from SNR to signal level.

Slide 10

Here are two panels that seek to reproduce a figure similar to one in a HamSci presentation by Nathan Frissell, W2NAF.

Our postgreSQL query selected all receivers in grid squares starting with FN13 (New York State). Actually, the grid square is a Grafana template variable - we can select any grid with reporters.

Next we set a centre bearing with reference to the receiver - here it is 270° - and a beamwidth - here 40°.

And so, we'll look at data from the segment shown on the map at right, courtesy of Tom, NS6T. The top panel is for 14MHz, the bottom for 7MHz. Using Grafana's Heatmap graphic the colours from blue to red represent the number of spots received in each 200km range ring in 20-minute time intervals over two days. The y axes are range from 0 to 4000km.

Of course we have to keep in mind that the distribution of WSPR transmitters over the US is far from even, and the area covered by the 200kmrange rings expands with range.

Nevertheless, these two panels show nicely the contrasting diurnal patters of propagation on the two bands. And of course, we can select different headings for our "telescope" to study propagation along different directions.

Slide 11

The combination of Timescale and Grafana can also help us understand conditions at our own station. This is data from my own station in a smallish suburban garden in Southampton, UK. I can just about fit in a pair of 40m inverted V dipoles, which I have connected to a homebrew adaptive noise cancellation system. Thst's G3ZIL. G3ZIL/P is an active vertical dipole (2m tip-to-tip) it has a fully differential high impedance buffer and a line driver designed and built by Glenn Elmore, N6GN.

Using postgreSQL postgreSQL we calculate the SNR difference for each exactly coincident WSPR spot pair from the two systems - the 4-day point cloud in the upper panel. It is of limited use. Of more use are the median and quartiles in the middle panel. There's a clear daily variation and well as day-to-day differences. One lesson is to put little credence on a claim made, even with an hour of data, that antenna A is better by x dB than antenna B.

As well as its own graphics Grafana offers 3rd party plug-ins. In the bottom panel I have used plotly to plot SNR differences against heading at the receiver. What we see here is a broad minimum at around 150°. That's where my noise cancellation system is steering a null - luckily for me that direction has few WSPR transmitters.

Slide 12

So far we have looked at a few days of data at a time. Here let us take a 50-day look. To help us comprehend that quantity of data I'm using an hourly heatmap plug-in for Grafana. I'm binning data into 30minute intervals and each 30 minutes through a day is shown on the y-axis, from 0000UTC at the top to 2359UTC at the bottom. Each column is one day.

All data is from the Northern Utah SDR site at 14MHz courtesy of Clint Turner, KA7OEI. The top panel is noise level in dBm in 1Hz from -147 in blue to -130 in red. The noise level is highest in the evening local time, with a sudden decrease as the sun sets. There's a slower rise at dawn, and if you squint you might just convince yourself that the duration of the nighttime minimum increases over time.

The second panel is over the same time interval and shown the number of spots received, from 1 in blue to 350 in red in each 30 minute interval. Here there is a very definite change in the duration of the minimum going from mid July to end August. We see most spots in the morning and afternoon, coinciding with the times of lowest noise, and fewer spots in the evenings when the noise is higher, and as the band closes for stations from the east.

The bottom panel is the mean distance of spots received, but one would really need to look at the median and quartiles of distance and the raw data. The peak in distance just before the band closes is due to receiving stations from Australia and New Zealand.

Slide 13

The Franke Taylor WSPR decoding program wsprd produces a number of internal parameters that are not sent to wsprnet.org. In case they are of interest to some WSPR high priests or wizards Rob logs them to the WsprDaemon spots table.

Here we show scatterplots of Metric on the left and Sync_Quality on the right, both against SNR on the X-axis. All I can say about Metric is to quote that it represents, "*closeness of a path to the received sequence*". The peak at 810 at an SNR range of -32 to -20dB is when the primary decoder gives up and hands over to an Ordered Statistics Decoder that always produces a result, but it may be wrong. It is only logged if the callsign was previously in your hashtable.

Both of these scatterplots show the usefulness of non-parametric density contours to pick out trends, a feature sadly not available in any Grafana plug-in to date.

Slide 14

For the immediate future these are our next steps - Rob dearly wants to find the time to extend WsprDaemon to handle SDRs with a Soapy interface. And I'll concentrate on extending our portfolio of Grafana Dashboards and try to interpret the results.

We'd both welcome other App developers considering using WsprDaemon and for those with an interest in analysing WSPR data to get in touch.

Thank you.

The end