Audio Engineering Society

Convention e-Brief 605

Presented at the 148th Convention
2020 June 2 – 5, Online

Trans-Europe Express Audio: testing 1000 mile low-latency uncompressed audio between Edinburgh and Berlin using GPS-derived word clock, first with jacktrip then with Dante.

Paul Ferguson[1], Chris Chafe[2,3], and Simon Gapp[3]

[1] *Edinburgh Napier University, Edinburgh, Scotland*

[2] *CCRMA, Stanford University, Stanford, USA*

[3] *Technische Universität Berlin, Berlin, Germany*

correspondence should be addressed to Paul Ferguson: (P.Ferguson@napier.ac.uk)

ABSTRACT

For nearly two decades, networked audio research using jacktrip has shown that multi-channel uncompressed audio was possible over a National Research and Education Network (NREN) and was now becoming viable over some public network connections. There was however, a 'Dirty Secret', in the absence of any synchronisation between transmitting and receiving word clocks, periodic audio loss due to data over-run or over-run was a certainty. The authors describe a low-cost GPS-derived clocking solution for jacktrip and then apply it to off-the-shelf Dante equipment and Dante Domain Manager for the world's first long-distance Dante audio over standard academic networks.

1. INTRODUCTION

Jacktrip is a linux, macOS, bsd and Windows multi-machine audio system used for network music performance over the internet [2]. It features uncompressed audio streamed bidirectionally with minimal latency between connected machines using UDP. Audio can be multi-channel (with very large channel count possible). Multi-site connections schemes are supported in a number of ways (hub mode, and many-to-many peer-to-peer modes). The basic use of jacktrip is via command line operation. Jacktrip is a free, open source program authored by Chris Chafe and Juan Pablo Caceres at Stanford University.

2. GPS-DERIVED AUDIO SYNC.

A goal of this project was low-cost synchronization. To this end, the authors have utilized general purpose GPS timing modules and modified a budget USB audio interface to add the required word clock input not usually found on budget devices.

2.1. Deriving word clock from GPS

The U-Blox NEO M8T [3] is a timing module using the Global Navigation Satellite System (GNSS) constellation of satellites which includes GPS along with Galileo, GLONASS, and BeiDou satellites. The module has two time pulse outputs that can be configured to generate satellite-derived frequencies between 0.25Hz and 10MHz. U-Box EVK-M8T evaluation kits were used in Edinburgh and Berlin to generate 48KHz word clocks to synchronize the audio interfaces.

2.2. Adding external word clock

The Behringer UM2 is a 2in/2out, 16bit/48kHz USB audio interface using a Cool Audio C2902 Codec clocked by a single 12 MHz crystal. All other required clocks are generated internally, this simplifies the task as the incoming word clock only needs to be upscaled once to 12MHz. The design depicted in Fig. 1 makes the following assumptions:

- The user selects between internal crystal or external word clock via a hardware switch
- The input frequency from an external word clock is fixed at 48 kHz

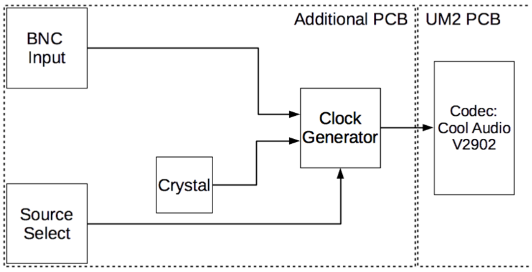

Fig. 1: modified UM2 with word clock input

The external word clock PCB is designed to be powered directly from the on-board 3.3V power supply and requires an additional 25mA. The clock generator is the Cirrus Logic CS2300 and is programmed via I^2C. In order to generate a 12 MHz clock for the codec, the incoming 48 KHz clock needs to be multiplied with the factor 250. The CS2300 offers a high-resolution multiplication format which requires the factor to be in 12.20 format. The 12 Msbs represent the integer binary portion while the remaining 20 Lsbs represent the fractional binary portion.

The final PCB contains the BNC word clock input and the 12 MHz crystal which was originally soldered directly to the UM2 PCB. The output of the newly designed word clock PCB is directly soldered to the codec's external oscillator input. The result is depicted in Fig. 2.

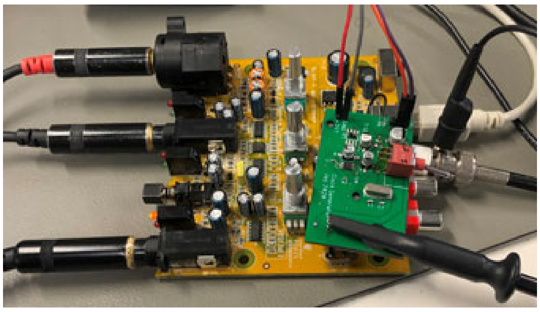

Fig. 2: UM2 with new word clock PCB

3. JACKTRIP GRAPHING & RESULTS

Jacktrip stamps its outgoing packets with sequence numbers. Clock skew between two jacktrip hosts is measured by the delta (or offset) between outgoing and incoming sequence numbers. When a packet is emitted onto the network, the number which has been recorded from the most recent incoming packet is compared to the emitted packet's number. Clock skew is observed as a gradually increasing or decreasing delta quantity. We graphed our observation for 50 minutes as shown in Fig. 3.

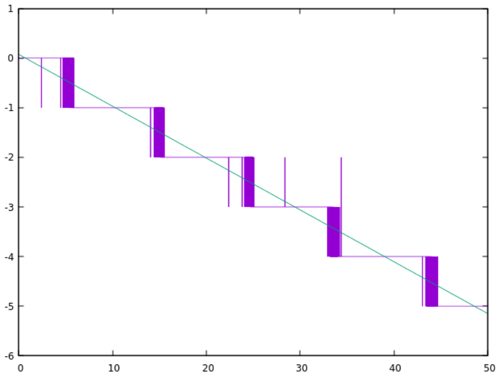

Fig. 3: Asynchronous jacktrip

Measurements were recorded by the jacktrip hub server instance running in Berlin at the Technische Universität (running the Fedora 30 linux operating system on a Lenovo T51 laptop). The jacktrip server code was customized to log delta values each time a packet was emitted. The Jack sample rate was 48000.0 and the frames per period (the quantity also known as the audio buffer size and equivalent to the jacktrip packet size) was set to 256 frames (emitting 187.5 packets per second). A stereo connection was received by the server from a jacktrip hub client instance (running on a macOS host) at Edinburgh Napier University with the same audio settings. The Berlin-side interface was a Hammerfall Fireface UFX (in USB compliant mode) and the Edinburgh interface was a Focusrite Red4Pre connected via Thunderbolt.

The green regression line in Fig. 3 corresponds to a clock skew (from the point of view of Berlin) of -9 ppm. Given this magnitude of difference in the separate server and client word clocks there will be an underrun condition every 10 minutes (approximately 10k packets). This occurs every time the regression passes another through another unit step.

Momentary deviations in the delta value are due to network jitter. The playback buffers in both hosts' jacktrips were sized to maintain an even (and decreasing because of clock skew) offset in their flows despite jitter in inter-packet arrival times. The playback buffer size in jacktrip is set via the "input queue size" parameter (the -q parameter) and in this test was 4 packets.

For the second run, both audio interfaces were slaved to external word clocks from the GPS system described. The desired constant delta was obtained as shown in another 50 minute plot, Fig. 4. As before, jitter-related excursions are evident but accommodated by the playback buffers. A recording-quality audio signal was obtained on each host.

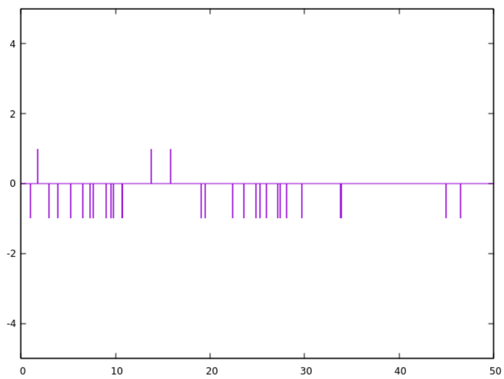

Fig. 4: synchronous jacktrip

ICMP pings from Berlin to Edinburgh averaged 40 ms (RTT). A traceroute in the same direction revealed 16 hops in the path as the traffic traversed Erlangen, Frankfurt, Amsterdam, London, Manchester, Glasgow via the Géant and Janet networks.

4. LONG-DISTANCE DANTE TESTS

Audinate's Dante media networking technology uses the 2002 version of the IEEE1588 Precision Time Protocol (commonly known as PTPv1) to achieve word clock synchronisation between devices on a local network. Fig. 5 shows the network-based sync flow in a typical Dante system where devices have IP addresses within the same subnet (for example: 192.168.1.1 and 192.168.2). One device has been elected Grandmaster and is sending clock to slave Dante devices in the form of multicast PTPv1 messages.

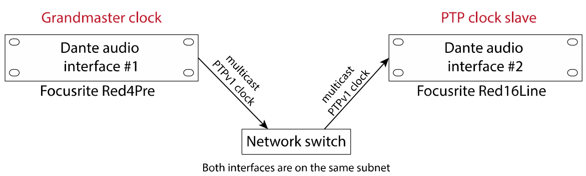

Fig. 5: Single subnet clocking using multicast PTPv1

There are two barriers to using this clocking scheme over long distance:

i. The network switch will normally ensure that multicast messages are limited to within a subnet and will not be passed to a remote location.

ii. Any network jitter will affect the timing of the PTP clock mechanism and result in clock jitter. This may ultimately cause the receiving device's phased lock loop to unlock and the device to mute its audio.

The first barrier can be addressed using Dante Domain Manager (DDM). DDM makes use of the boundary clock addition to the 2008 revision of the IEEE1588 protocol (commonly known as PTPv2) and uses unicast PTPv2 clock messages to bridge two subnets.

Although unicast PTPv2 gives a solution within a building or campus, those unicast PTPv2 messages will still be susceptible to network jitter in a long-distance application. To address the second barrier, DDM 1.1 allows subnets and their Dante devices to be assigned geographically separated 'zones'. In each zone, one device on the network derives its clock from GPS satellite and acts as the Grandmaster clock for that subnet.

This method was successfully demonstrated during the Innovation in Music 2019 keynote performance [1] using Sonifex AVN-GMCS PTPv2 grandmaster clocks in Edinburgh and London. To embrace the low-cost spirit of this trans-Europe test, Fig. 6 shows the use of two U-Blox GPS receivers to externally clock Dante devices in Edinburgh and Berlin.

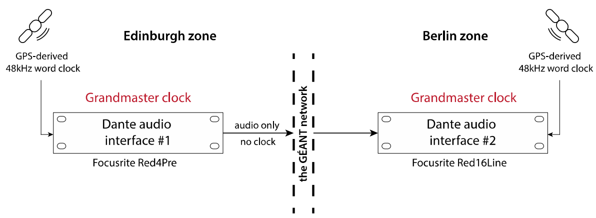

Fig. 6: Two Dante devices enrolled into DDM, one in an 'Edinburgh Zone', the other in a 'Berlin Zone'.

Fig. 7 shows the Clock Status Monitor window within Dante Controller software. In the left window you can see the parts-per-million frequency offset between the Edinburgh Red4Pre's internal oscillator and the U-Blox GPS-derived clock connected to its external word clock input. To demonstrate that the Red4Pre responds to variations in the external clock the right-hand window shows the effect of changing the U-Blox output frequency. The three peaks correspond to 47999Hz, 48000Hz and 48001Hz, all derived from GPS.

Fig. 7: Dante Controller Clock Status Monitor showing GPS-derived external word clock.

4.1. Edinburgh to Berlin Dante results

Users of Dante devices will be familiar with a buffer setting that determines the playout latency. As an example, the buffer in the Brooklyn II-based Red4Pre interface can be set between 250us and 4ms and this choice is influenced by the network performance and the number of switch hops. Implicit in the use of DDM zoning described earlier is geographical separation and, therefore, larger network latency due to distance. As a result, enrolling a Dante device into DDM enables a larger buffer to be selected. In the case of a Brooklyn II-based Dante device the maximum becomes 40ms.

Fig. 8 shows a Dante Controller latency window during a 45 minute Edinburgh to Berlin test. The receiving Dante device indicates low jitter and no packets exceeding the 40ms Brooklyn II buffer. No clock errors or mutes/unmutes were reported by Dante Controller during this time.

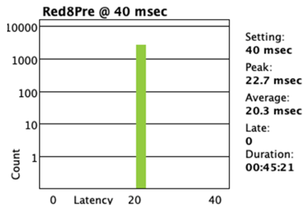

Fig. 8: Latency between Red16Line (Berlin) and Red8Pre (Edinburgh).

5. CONCLUSIONS

These tests successfully demonstrate that both jacktrip and Dante can be synchronised over long distances using low cost GPS-derived word clocks and the Janet/Géant academic networks. Although originally developed for use over high-speed academic networks, jacktrip it is currently finding use home-to-home during Covid-19 lockdown and benefiting from commodity cloud services networks.

6. ACKNOWLEDGEMENTS

Thanks to Marc Voigt, Paul Schuladen, Nils Tonnätt, Henrik von Coler and Orchisama Das,

7. REFERENCES

[1] Focusrite (2020). RedNet Enables 'World-First' Cross-Border Interactive Performance. https://pro.focusrite.com/case-studies/live-sound/rednet-enables-'worldfirst'-crossborder-interactive-performance

[2] Jacktrip - A System for High-Quality Audio Network Performance over the Internet https://ccrma.stanford.edu/software/jacktrip/

[3] UBLOX (2020). UBX-15025193 NEO/LEA-M8T concurrent GNSS timing modules data sheet.