

Clock synchronization in Cesnet monitoring infrastructure

Vladimir Smotlacha

31.12.2005

1 Introduction

Monitoring is a great challenge of the Cesnet 2 network in order to provide detailed information for an effective network control and measurement. We have designed and installed a monitoring infrastructure in main network nodes, i.e. nodes connected by two and more backbone lines. The infrastructure consists of set of PC boxes designed to provide both active and passive measurement. Boxes can host PCI-X passive monitoring adapters Combo6 (or alternatively DAG) in order to provide full line speed monitoring.

2 Matter of clock synchronization

Measurement of some network parameters depends on clocks of sites of measurement. An example is one-way delay: the measurement is simple in principle, as it is enough to know sending time and receiving time of the packet. However, the transport capacity of links and network nodes is permanently increasing and the value of one-way delay is decreasing. Currently, one-way delay is in the order of hundreds of microseconds in LANs and several milliseconds across the Cesnet 2 network. Meaningful measurement of one-way delay requires an evaluation of timestamps with the accuracy for one order in magnitude better than the expected delay. The accuracy of measurement depends on relative accuracy of both clocks that assign timestamps. Such dependence on the clock is called 'time sensitivity'

Another reason for providing an accurate time is the assigning of unique timestamps to captured packets.

3 Methods of clock synchronization

The simplest method is to utilize NTP via network. Despite the accuracy, which is in the order of hundreds of microseconds, such solution has a drawback:

when the clock of box located in the backbone node A is synchronized by a NTP server located in another node B, the synchronization is influenced by the backbone path between A and B. As a result, any measurement depending on synchronized clocks A and B, can't give objective information about the path (or its part) between A and B.

To avoid this problem (which is a NTP native feature), we have to provide all boxes by an accurate local source of time. We decided to install in each node a GPS receiver disciplining clock of the measurement box. According to our previous experience, we installed Garmin receivers in majority of localities. The receiver has a PPS (pulse per second) signal output and the producer declares, that the PPS signal accuracy is better than 1 microsecond.

4 Utilized hardware

4.1 Measurement boxes

All measurement boxes that are currently in operational status are identical:

- mainboard Supermicro,
- PCI-X bus,
- Xeon 3 GHz,
- 1 GB RAM,
- 70 GB SCSI disk.

Such hardware configuration is powerful enough for installation of Combo6 or DAG cards up to 10 Gbps. Depending on the required capacity and throughput of disk system, more disks can be installed.

4.2 GPS receiver options

Depending on local condition, we used several solutions how to provide PPS signal for measurement box. In most localities, a NMEA signal (containing the label of current second) is also available.

Praha: the measurement box is provided by the signal of high-quality Trimble Acutime 2000 which is already distributed in the computer room and the laboratory. We have two distribution units in cascade, each with 8 RS-232 outputs (Figure 1).

Brno, Plzen, Ceske Budejovice and Olomouc: the GPS receiver has been installed on the roof in each locality. The solution is described in more detail in a later chapter.

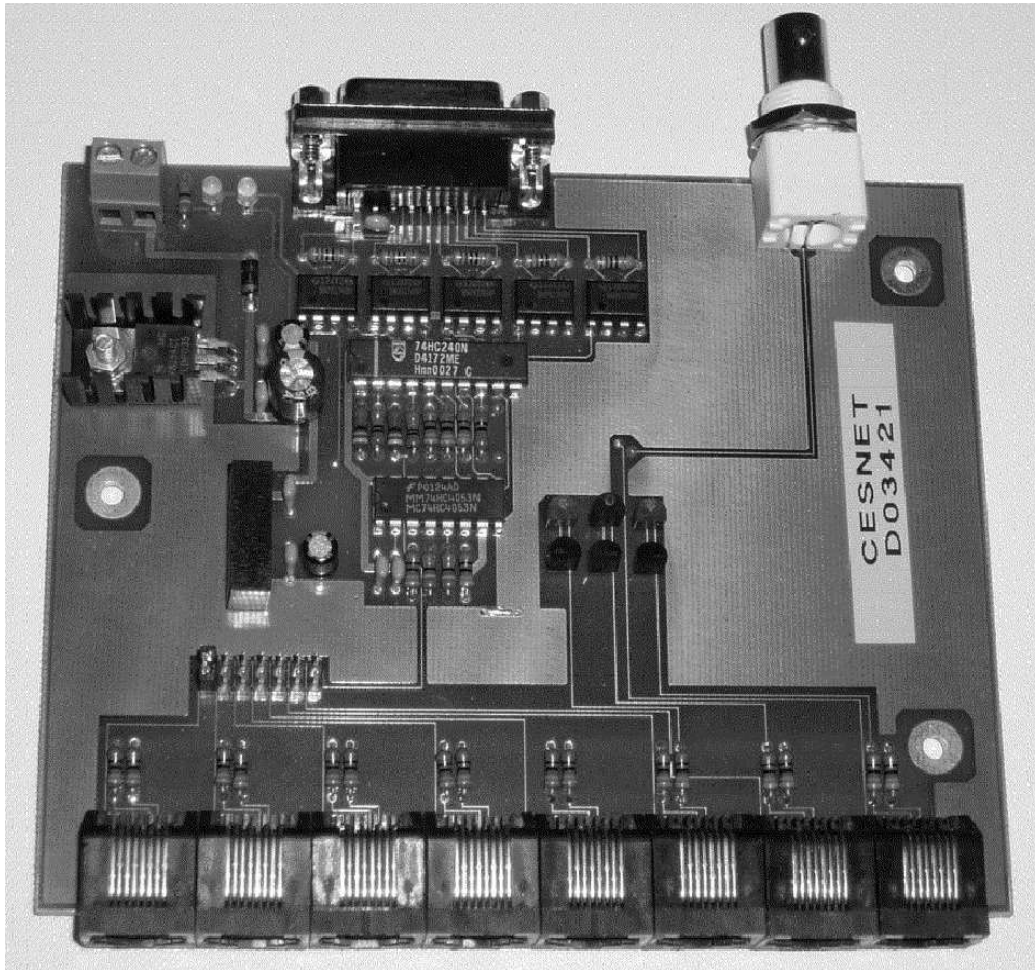


Figure 1: PPS distribution unit

Ostrava: a precise geodetic GPS receiver Topcon GB-1000 has been already installed in computer room in Ostrava. Therefore, we decided to utilize its PPS signal for the clock synchronization.

Liberec: a specific situation in Liberec node resulted in impossibility of GPS receiver installation on the roof. The measurement box is synchronized only by a NTP server located in Praha.

5 GPS receivers installation

There was unavoidable to install new GPS receivers in Brno, Plzen, Ceske Budejovice, and Olomouc. We used Garmin GPS 35 or GPS 18, that provide signals (PPS and serial) in the RS-232 data format. However, RS-232 is unsuitable for many tens of meters long cable so some kind of conversion is necessary. We tried to find a professional solution to prolong serial line up to hundreds of meters and we chose adapters LD232 manufactured by Papouch s.r.o. The conversion unit also supplies power for GPS receiver and provides galvanic separation between measurement box and the receiver. Picture 2 shows all installation set: the receiver Garmin GPS 18, the first adapter in weather-proof box, and the second adapter. Both adapters are connected by standard UTP(STP) cable with 4 pairs of wires. When a structural cabling system is available, it can be utilized.

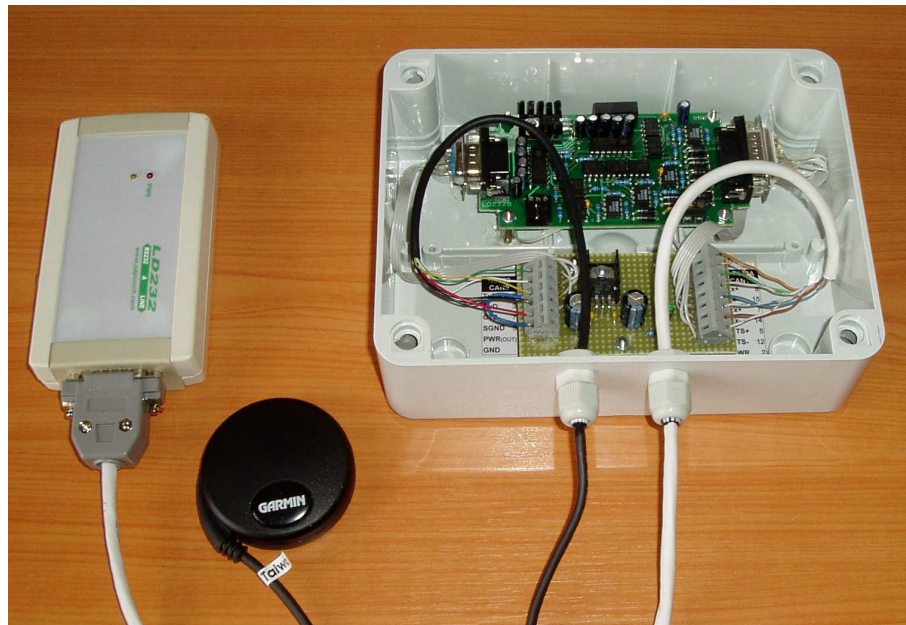


Figure 2: GPS receiver and link link adapter

6 Software issues

Linux is our primary choice of operating system. Precise clock synchronization by a PPS signal is not possible with neither 2.4.x nor 2.6.x standard kernel as it requires two kernel hacks: the PPS API and nanosecond resolution of internal clock. As nanosecond clock resolution has not yet been implemented in 2.6.x kernels, we utilize 2.4.29 kernel with the 'nanokernel' patch.

We installed NTP 4.2.0, the last stable version. Garmin receivers as well as Acumatic 2000 are configured to operate in NMEA output mode, which is processed by a NTP package driver. The NMEA format provides label of current second and state of the PPS signal. In Ostrava, no label information is provided by the GPS receiver, so the label is determined by a main Cesnet NTP server.

Specification of measurement software is out of scope of this report.

7 Accuracy of clocks

Following graphs show typical 24-hour record of the synchronized clock offset of all sites of monitoring infrastructure. The offset was reported by the 'ntpd' process. We can conclude, that only exceptional offset values exceed the the interval +/- 30 us. A special case is Liberec, when we did not succeed to install the GPS receiver and therefore the box is synchronized only via the network.

Also or monitoring boxes are identical, we can see differences of clock offset values. The clock offset is influenced by:

- parameters of PPS signal,
- parameters of crystal oscillator,
- variation of ambient temperature (cooling system, air condition, ...)

As declared PPS signal accuracy of Garmin GPS is better than 1 us, the influence should be negligible. However, we see similar graphs with offset spikes in all localities with Garmin GPS 18 or GPS 35 (i.e. Brno, Plzen, Ceske Budejovice, and Olomouc). In case of Trimble Acumatic 2000 (Praha) or Topcon GB-1000 (Ostrava) are these spikes significantly smaller or missing at all. This phenomenon should be verified and than clarified.

8 Ongoing tasks

We are now ready to install two other boxes in Usti nad Labem and Hradec Kralove but we are waiting for the manufacturing of LD232 adapters that was delayed for several months. We also have to improve accuracy of clock synchronization in Liberec.

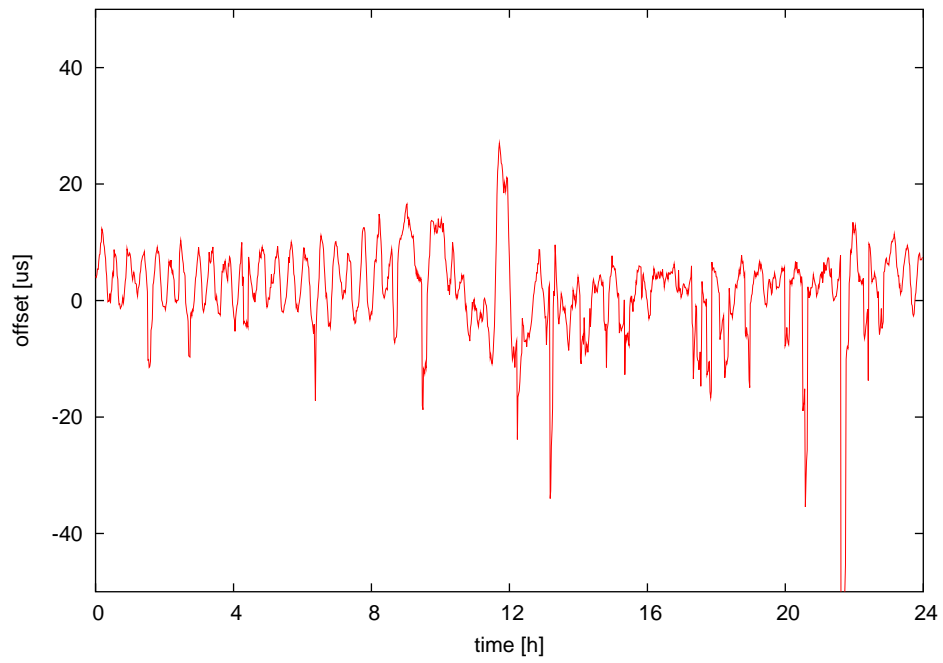


Figure 3: Clock offset - Prague

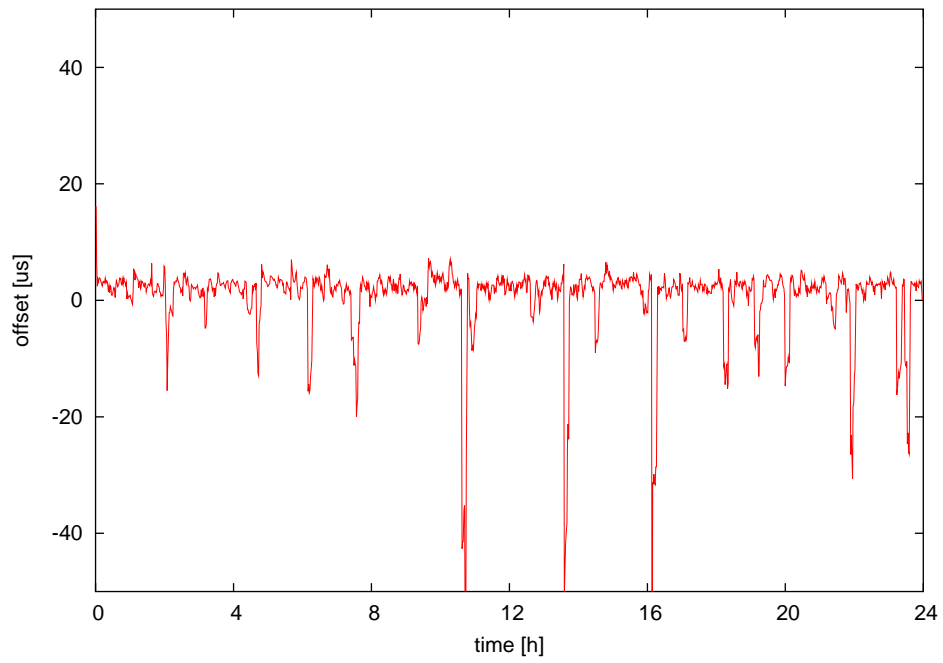


Figure 4: Clock offset - Brno

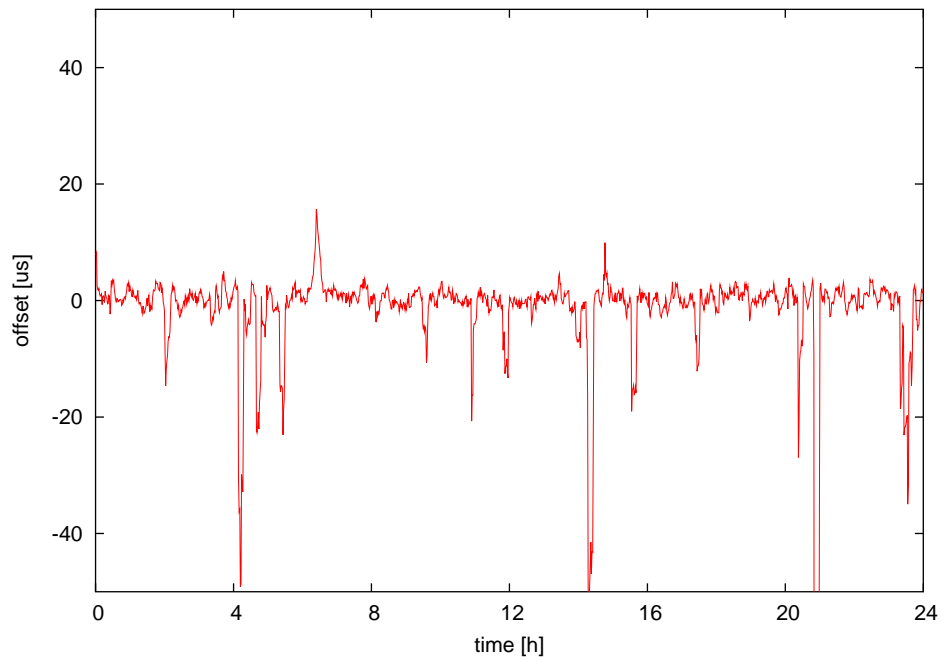


Figure 5: Clock offset - Plzen

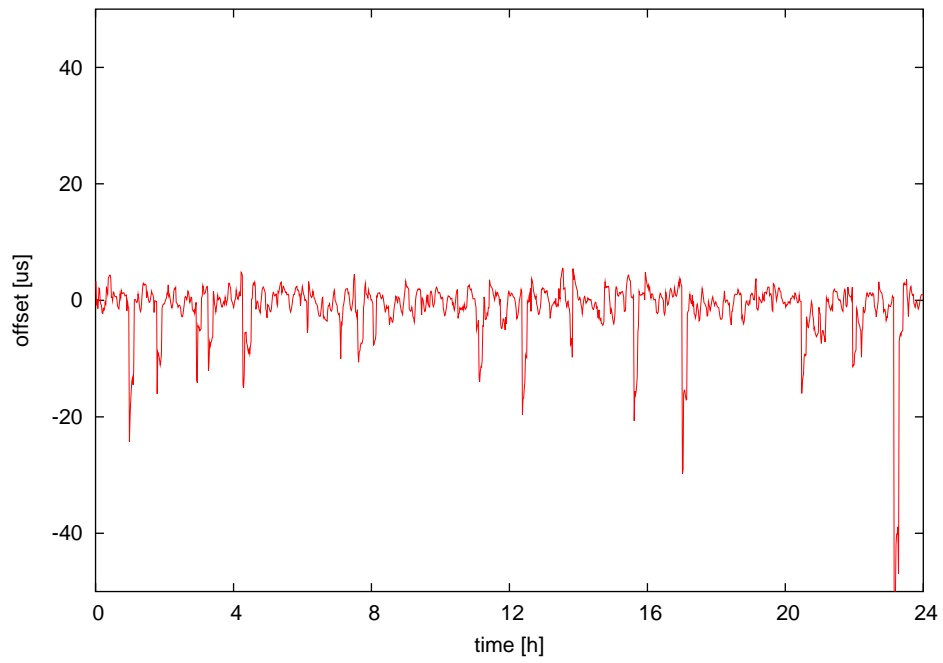


Figure 6: Clock offset - Ceske Budejovice

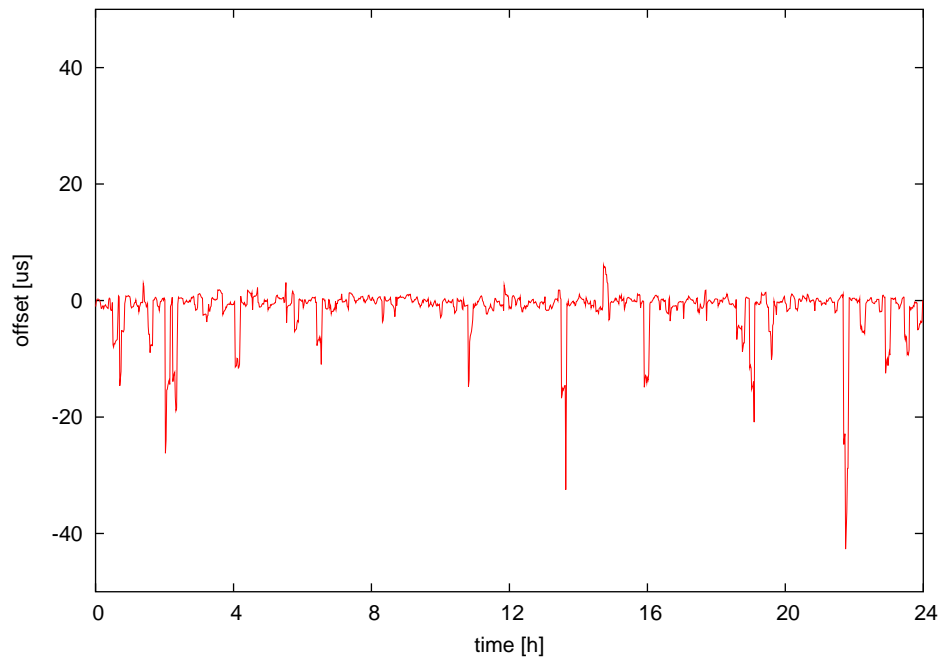


Figure 7: Clock offset - Olomouc

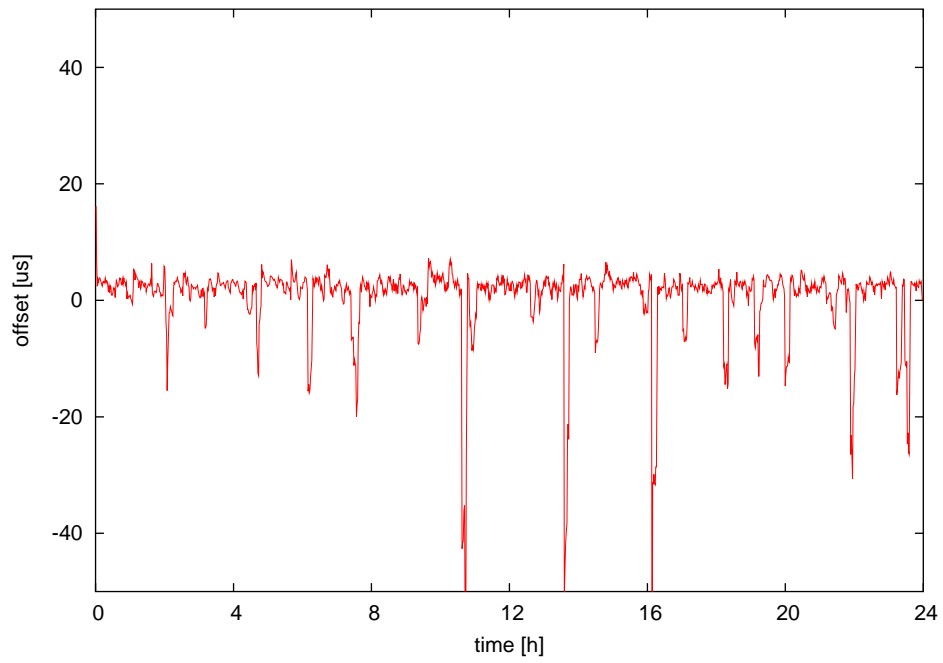


Figure 8: Clock offset - Ostrava

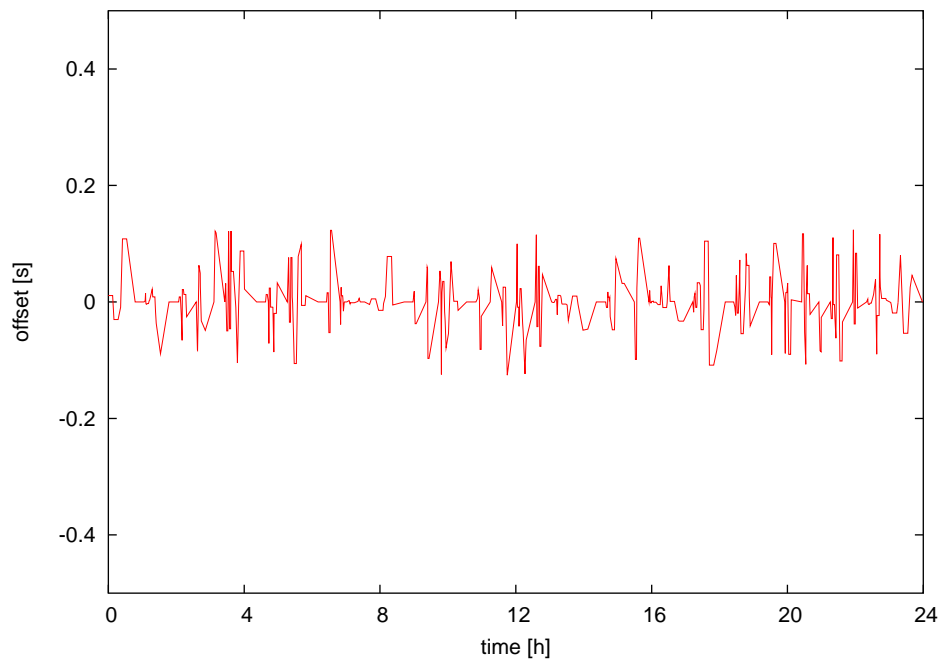


Figure 9: Clock offset - Liberec