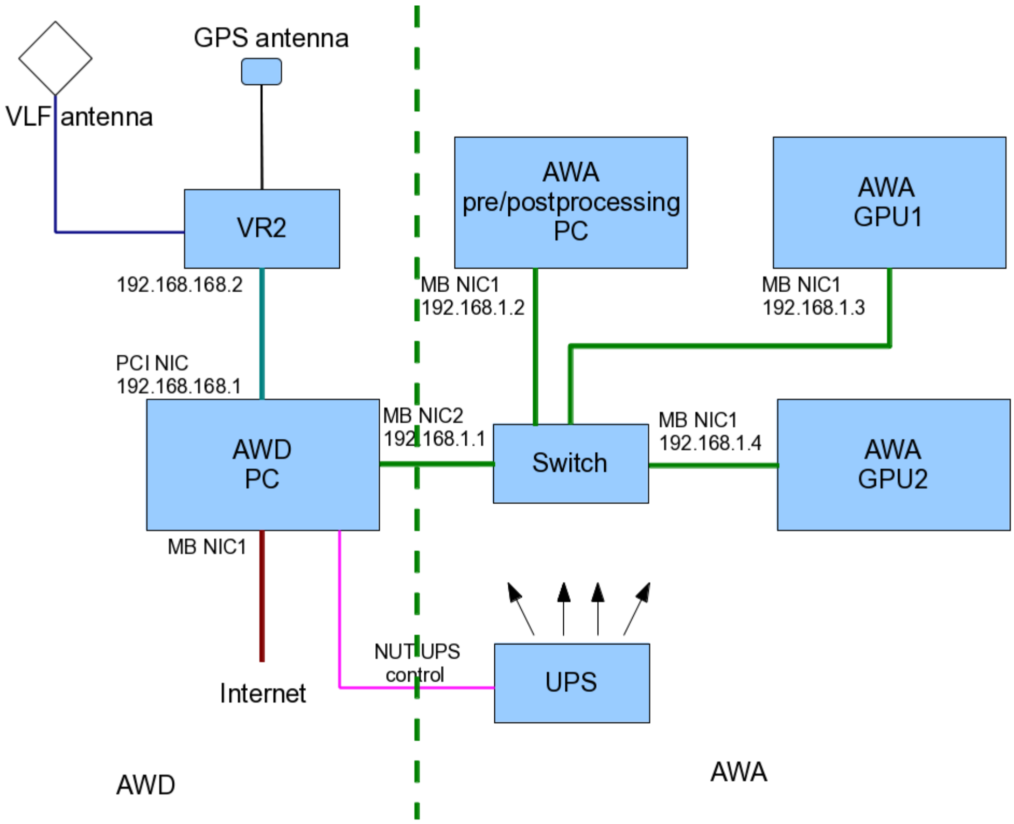
Reorganization
of
Automatic Whistler Detector and Analyzer Network
(AWDANet)
nodes

There are several reasons behind the reorganization of AWDAnet:

1. aging of hardware, including high power consumption of old GPU cards,
2. results on statistical analysis on data volumes, based on two decades AWDANet operation,
3. a recent ESA project on plasmaspheric products, which hopefully leads to an operational service. One of the major input data source of this project is AWDANet.
   This is a special ESA project. ESA has different, sometimes very specific call conditions. This case, the call was exclusively dedicated to Hungarian entities and no entity from other countries are allowed to be joined/financed from the fund. This is why you have not heard about it. And we received the fund exclusively for salaries, no other costs are eligible. This is why the disks I'm planning to send you are purchased from a different grant. The project major goal is to develop plasmaspheric products, 2D density maps and empirical and data assimilative plasmasphere/plasmapuse models for user community (space weather modelers, spacecraft operators and designers). The project ends with a 3 months pilot period, where we provides the products in real time, published on the ESA SWE portal. A successful pilot period (planned in the second half of 2022) hopefully leads to establishing a standard service, where ESA contributes to the operation of AWDANet. All the AWDANet collaborators will be recognized on the SWE portal during the test period - but as no support can go to them it is rather a moral victory…

The hardware and software at AWDANet nodes were purchased and installed in most of the cases during the PLASMON project (2011-2014). The installation consists of the components in the figure below:



There are stations, where there is one or three GPU PCs instead of two and there are places, where AWD and AWA PC are merged into a single one.

The preamplifier and VR2 data logger can work for decades unless a nearby lightning stroke damage them. The vulnerable components here are the PSU ‘bricks’.

The main components of PCs were good quality ones at the time of purchase and – even after ten years – the motherboard, CPU and RAM are still working. The failing components include power supply units and hard disks.

The – at the time of purchase – high end GPU cards (dual GTX590 cards) now are outdated, considering both the speed and power consumption. The latter can be a problem at remote sites.

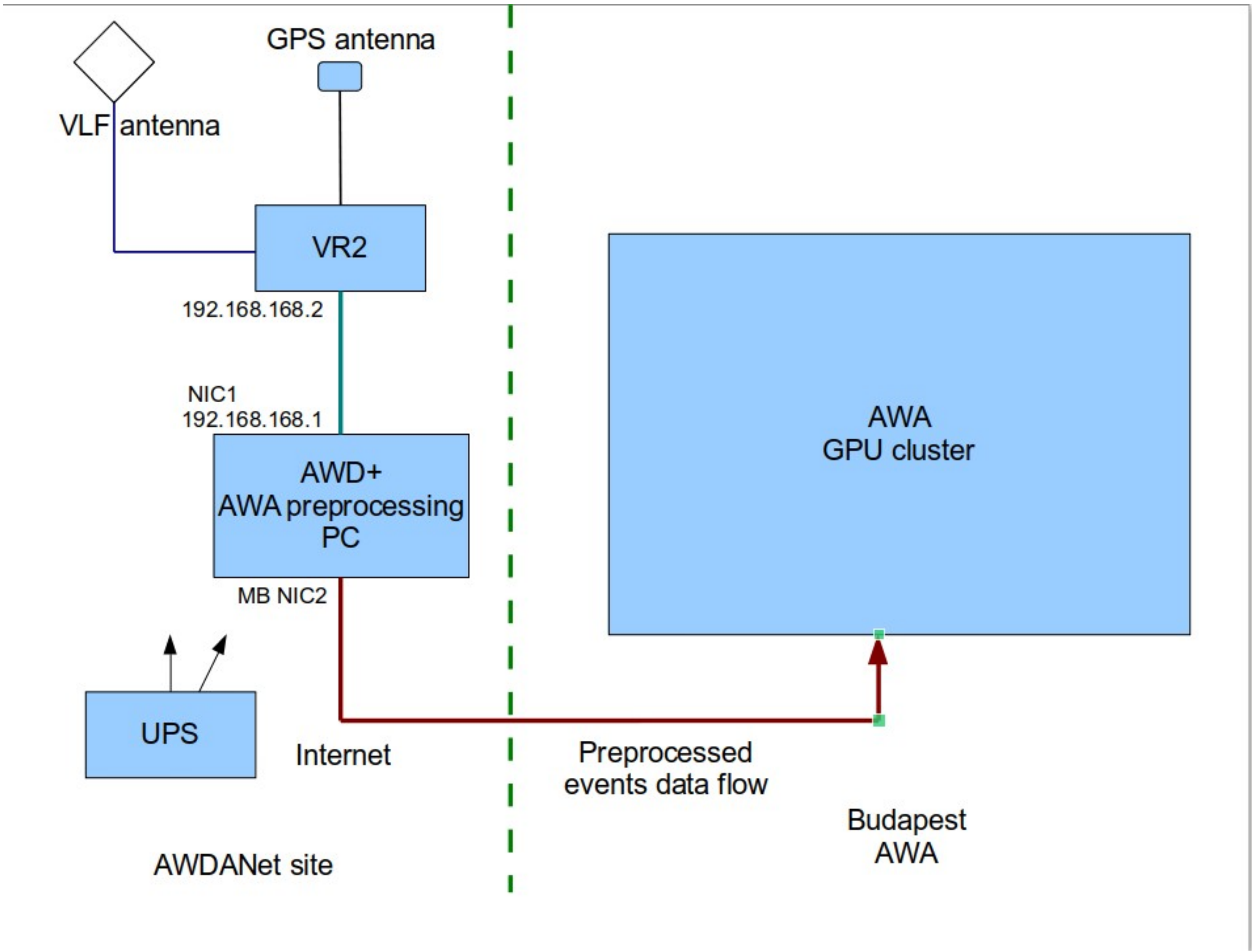
These circumstances led to rethinking the processing chain, AWD and AWA algorithms. At the moment all elements of the chain are executed on-site. The major reason behind this architecture was the high raw data amount and the low available network bandwidth at remote sites.

We have performed a statistical analysis of processed data at all sites. The figures showing the results for your site(s) attached separately. We have calculated the maximum/average hourly data volume for whistler events. The original AWA preprocessing algorithm assigns a quality factor (qf) to each detected whistler events on a 00-15 scale. According to the information gained during the two decades of AWDANet operation, the candidates of successful whistler inversion fall into qf=00-04 range. The raw VLF data volume for whistler events falling into this range is a small fraction of the original raw data volume, therefore it is probably feasible for most (hopefully all) of the sites to transfer this data over internet in real time to a central location (Budapest) for further processing.

Therefore we decided to split the AWDA (=AWD+AWA) procedure into two sections:

1. AWD algorithm and AWA preprocessing modules: they remain at AWDANet sites. The selected whistler events are cut from raw data flow, compressed (this can lead to further 25-30% data volume reduction) and copied to Budapest,
2. Main AWA module: running on local CPU+GPU cluster at our lab. The AWA algorithm will also be upgraded.

This the block diagram of the new setup:



The critical point in this plan is of course the network. Analyzing the statistical data, we think, it is feasible and will not cause network overload problem. But of course, it has to be discussed by the local IT guys. As our original processing concept in PLASMON was ‘quasi real time’, even limited bandwidth can work at sites, where the bandwidth and load ratio requires it. Further, in the case of high data traffic, the data volume to be transferred can be dynamically reduced shrinking the qf range from 00-04 to e.g 00-02, halving the data volume.

The original processing flow works this way:

1. The full inversion process (including AWA pre-and post-processing) takes 60-100 sec for an event.
2. At t=0, we take the best event (with smallest qf) occurring at t=-100 - 0 sec and process it.

3. After AWA process is finished (it takes places at t=+60- +100 sec), we take again the best (and latest) event in the period of t= 0-+100 sec. If no events, the flow idles for another 100 sec.
4. Steps 2. and 3. are repeated, advancing the time by 100 sec.

The only exception, when 'no events' do not happen is Rothera in July-August, when literally there are whistler event in every or every other second. This would mean (changing to 'local' to 'central' processing method) to transfer all raw data from Rothera to Budapest (560Mbyte/hour). However, we do not need a plasmaspheric density in every second as plasmasphere varies much slowly. Thus, transferring an event in every 60-100 second is enough and it reduces the data volume to ~10Mbyte per hour at the highest occurrence periods. The data rates at other sites and other periods are (sometimes significantly) smaller than this.

As I mentioned above, the major components of local hardware can still work for a while, therefore there could be two scenarios for the upgrade:

1. a completely new hardware replacing the existing one. That is, the AWD, AWA and GPU PCs are replaced by a single PC. This is certainly the best solution, but the cost is high compared to the available resources at the moment.
2. keeping one PC from AWA and GPU PCs – basically they are identical except the GPU cards – replacing the HDDs with a new SSD system disk and two 16TByte (10TByte for sites, where the sample rate is 20kHz, like SANAE and Halley) SATA HDDs for data storage and backup. The two large HDDs allow to store the full raw VLF data at each sites for ~3 years. This can be a huge advantage for analyzing events, e.g looking for chorus, hisses or other VLF emissions simultaneously with satellite data (the present AWD saves and stores only a fraction of raw data). To decrease the costs, 12TByte HDDs can also be considered, they can store raw data for 2 years. As the parts are identical, the non-used PCs can serve as a spare part "banks" for the operating one. Of course, this is a temporary solution, the old PC components need to be replaced sooner or later – hopefully for that time, when it will be necessary, we will have resources for that (e.g from ESA...)

The advantage of the reorganization for the AWDANet nodes is a much simpler local installation (a single PC instead of 3-4-5 ones), much smaller power consumption – the single PC consumes 75-105W (based on tests), while the present GPU PC with the dual GTX590s can consume 1kW on high load.

Another advantage is that no need to transport high volume and weight to the remote sites, the new central processing unit in Budapest can consist of fewer PCs and GPU cards as the activity of AWDANet sites on Northern and Southern hemispheres complement each other leading to less investment and operational costs.

The disadvantage is the higher network load, but as you see form the attached plots, in general it is not very high, except a few periods, but it hopefully manageable.
Please note, that the numbers in the plots need to be divided by 60-100 as we process a single event in every 60-100 sec – even if there are more events.