Trident 2.1 Concept: Updates to the CSP Reference Design

Omar Yeste Ojeda

September 11, 2020

Executive Summary

The Integrated Receiver and Digitizer (IRD) [1] and the Central Signal Processor (CSP) [2] described in the Next Generation Very Large Array (ngVLA) reference design are based on independent designs, neither of which were originally conceived for the ngVLA. As a result, their interfaces are not compatible with each other. The solution proposed in the reference design relies on inserting the Digital Back-End (DBE) [3] as an adapter between both subsystems. In this arrangement, the DBE must perform part of the CSP functionality in order to comply with preliminary system requirements, but the task division between the DBE and the CSP remained unclear. An improved, more definite solution with a more natural transition from the IRD to the CSP is desired for the conceptual design, which is the purpose of this document.

The proposed modification is summarized as follows:

1. Remove the threefold partition of the CSP.

2. Nominal sub-band bandwidth of 200 MHz (35 sub-bands per 7 GHz of bandwidth). Sub-band oversampling factor of 35/31 (see Section 2).

3. Nominal down-conversion scheme and sub-band assignment according to Table 2.

4. The DBE at the antenna assumes all the functionality (with some exceptions) of the Very Coarse Channelizer (VCC) part of the CSP Trident design, notably the generation of sub-bands. The VCC part is removed from the Trident design. Use the current VLBI Data Interchange Format (VDIF) standard for data transmission (more details in Section 4).

5. Equip the CSP with at least 90 Frequency Slice Processors (FSPs), which satisfies the system requirement of simultaneous processing of 20 GHz of bandwidth[1] in correlation mode and additionally 8 GHz in beamforming mode (see Section 5).

6. Provide the CSP with necessary switched fabric to connect the DBE to the CSP, namely its FSP part, as explained in Section 6.

7. Install new devices on the CSP for the remote antennas, the remote antenna buffering transceivers (RABiTs), which perform data buffering, packet reordering, and bulk delay corrections (see Section 7).

---

[1]Throughout this document, bandwidth, sub-band or any other similar concept, is always expressed on a per polarization basis. Two polarizations are always processed.

Contents

List of Acronyms

CBE	CSP Back End
CSP	Central Signal Processor
DBE	Digital Back-End
DRAM	Dynamic Random-Access Memory
FFT	Fast Fourier Transform
FPGA	Field-Programmable Gate Array
FSA	Frequency Slice Architecture
FSP	Frequency Slice Processor
GbE	Gigabit Ethernet
Gbps	Gigabits (10^9 bits) per second
IF	Intermediate Frequency
IRD	Integrated Receiver and Digitizer
LO	Local Oscillator
ngVLA	Next Generation Very Large Array
NRC	the National Research Council of Canada
RABiT	Remote Antenna Buffer-ing Transceiver
RF	Radio Frequency
RFI	Radio Frequency Interference
SKA	Square Kilometre Array
SoC	System on a Chip
TBD	To Be Defined
VCC	Very Coarse Channelizer
VDIF	VLBI Data Interchange Format
VLBI	Very-Long-Baseline Interferometry

1 Motivation and Design Drivers

In this document, I describe a proposal for the modification of the reference design of the ngVLA CSP. A few reasons motivate this proposal: First, the current reference design does not satisfy system requirements related to commensal data processing requirements. An initial attempt of minimum deviation from the reference design is annexed to this document as [4]. The solution described therein maintains the passive optical fiber interface to distribute spectral sub-band among sub-band processors. Nonetheless, that interface has been identified as a potential limitation to support commensal back-ends in the future. Therefore, the solution proposed herein utilizes an active switched fabric network for the distribution of sub-bands. A second reason for the modification of the reference design concerns the generation of sub-bands. System requirements demand that the sub-band generation process be performed at each antenna station [5, SYS0904]. As a result, the VCC of the Trident architecture become redundant, except for compensating variable data transport latency. The FSPs cannot perform this task for remote antennas due to hardware limitations, given the expected range of transport latency. A third motivation is the harmonization of the frequency plan so that it nicely fits the bandwidth of Trident's frequency slices. The impacted subsystems are the Local Oscillator Reference and Timing subsystem [6] and the IRD. Finally, there exists a new FSP function mode for correlation that trades dynamic range for bandwidth, as compared to the single correlation mode described in the reference design. In this function mode, one FSP can process two independent sub-bands in correlation mode using 8-bit input quantization instead of 16-bit. This allows great savings in the design by reducing the number of FSPs required to process the same bandwidth. Sub-bands containing strong Radio Frequency Interference (RFI) can still be processed using 16-bit quantization. For the sake of brevity, I assume the reader is familiar with the ngVLA reference design. Further details can be found in the bibliography.

2 Sub-band Bandwidth

The first parameter that requires careful system-wide harmonization is the sub-band (or 'frequency slice', as referred in the Frequency Slice Architecture (FSA)) bandwidth. This is the information bandwidth contained in each of the sub-bands, as opposed to the sub-band sampling frequency, which is higher due to channelization constraints. In the Trident architecture, each FSP receives (at least) one sub-band from every antenna and processes it according to its functional mode. The reference design assumes an approximate bandwidth of 200 MHz, which is the bandwidth the FSP has been originally designed for.

Assumption 1 *– The optimum sub-band bandwidth is 200 MHz. Higher values exceed the capabilities of a single FSP, while smaller values require more FSP hardware for the same overall bandwidth.*

In principle, the CSP should allow any bandwidth under the above constraint, but the implementation of the sub-band channelizer is much simpler when an integer number of uniformly spaced sub-bands are generated across the digitized bandwidth. The IRD modules can be configured to sample at two rates, 14 GS/s (giga-samples per second) and 7 GS/s. The sampled signal can then be split in 200-MHz wide sub-bands, yielding 70 or 35 sub-bands depending on which sampling rate was used.

The next possible value for the sub-band bandwidth is 194.4 MHz, which produces 72 and 36 frequency slices at 14 and 7 GS/s, respectively. In [4], I opted for this option because the prime factorization of 72 and 36 allows a simpler implementation of the Fast Fourier Transform (FFT) utilized for the sub-band channelizer. On the other hand, the smaller sub-band bandwidth implies an 3% increase in FSP hardware cost to process the same amount

of overall bandwidth. However, this option ultimately led an even number of sub-bands between adjacent Local Oscillator (LO) frequencies which turned problematic, and hence I propose using 200 MHz instead. I will elaborate on the problems caused by an even number of sub-bands in the next section, as the issue persists in Band 6 because the separation between adjacent LO frequencies doubles with respect to lower frequency bands. Under the above assumption, smaller bandwidths are possible, but they would require even more additional hardware while there is no clear benefit justifying that choice.

Sub-bands are oversampled at generation to attain high spectral selectivity while relaxing the prototype filter design. Oversampling also allows extending optimal filter performance over the full information bandwidth, effectively eliminating any performance degradation or coverage gaps at the sub-band edges. The oversampling factor is defined as the ratio between the sampling frequency and the channel spacing or sub-band bandwidth determined above. In the reference design, this factor is assumed approximately 10/9, or 11% excess bandwidth, which is the value in the original design of the FSP.

Assumption 2 – *The frequency slice oversampling factor must follow the expression $\frac{M}{N}$, where M is the number of frequency slices and N is any integer $0 < N < M$.*

The above assumption implies generating one sample for each of the *M* sub-bands every *N* input samples. This notably simplifies the filter bank design as no additional resampler is needed. When generating 35 sub-bands, all valid oversampling factors follow the expression $\frac{35}{N}$, and are equally valid for generating 70 sub-bands. I propose using an oversampling factor 35/31, with a 12.9% excess bandwidth, which is the minimum valid value greater than the original one, 10/9.

3 Frequency Plan

The downconversion frequency plan described in the IRD reference design utilizes fixed LO frequencies, except for slight LO offsets from antenna to antenna (detailed offset plan is TBD) [1]. These nominal LO frequencies are multiple of 2.9 GHz to serve two purposes: simplicity of the generation scheme through harmonic generation and relaxed anti-aliasing filter requirements by providing enough spectral overlap between adjacent IRD bands. Setting the LO fundamental as a function of the sub-band bandwidth will additionally allow seamless transition across IRD bands and minimum redundancy between adjacent sub-bands at IRD band edges, that is, the last sub-band of one IRD band, and the first sub-band of the following IRD band. The amount of non-redundant Radio Frequency (RF) bandwidth covered by a given number of sub-bands can be maximized by minimizing the spectral overlap between them. Setting the LO frequency spacing as a multiple of the sub-band bandwidth guarantees that sub-bands become uniformly spaced even across IRD bands, providing a seamless transition from one IRD band to another. Because the LO spacing in the reference design, i.e. 5.8 GHz, already comprises an integer (29) number of sub-bands, there is no need to modify the fundamental 2.9 GHz, which is 14.5 times the sub-band bandwidth.

Obviously, the above discussion becomes meaningless if LO offsets are set arbitrarily. However, there are at least two ways of maintaining sub-bands uniformly spaced across IRD bands and still introduce LO offsets. One possibility is to introduce the same LO offset to all the IRD bands, as LO offsets need to be different only among antennas. Other method would be introducing the offsets in the central building to the frequency reference sent to each antenna [6]. In any case, further discussing LO offsetting is beyond the scope of this report. On the other hand, it is very likely that seamless channel transitions across sub-bands will still need fine tuning of these parameters, specially if an integer number of samples per second is desired to facilitate pulsar timing observations. Alternatively, additional corrections at the phase tracking stage can also achieve that goal.

RF Band	LO Frequency (GHz)	RF Range (GHz)	# of Sub-bands	Sub-band Range
1	N/A	1.1 - 3.5	12	6 - 17
2	5.8	3.5 - 8.7	26	18 - 43
	11.6	8.7 - 12.3	18	44 - 61
3	14.5	12.2 - 17.4	26	61.5 - 86.5
	20.3	17.4 - 20.6	16	87.5 - 102.5
4	23.20	20.5 - 26.1	28	103 - 130
	29.0	26.1 - 31.9	29	131 - 159
	34.8	31.9 - 34.1	11	160 - 170
5	31.9	30.4 - 34.8	22	152.5 - 173.5
	37.7	34.8 - 40.6	29	174.5 - 202.5
	43.5	40.6 - 46.4	29	203.5 - 231.5
	49.3	46.4 - 50.6	21	232.5 - 252.5
6	75.4	69.9 - 81.3	57	350 - 406
	87.0	81.3 - 92.9	58	407 - 464
	98.6	92.9 - 104.5	58	465 - 522
	110.2	104.5 - 116.1	58	523 - 580

Table 2: Proposed nominal down-conversion and sub-band allocation plan.

Table 2 summarizes the resulting nominal frequency plan, which is essentially the same as proposed in the reference design [7]. Because LO frequencies are not widely tunable and all sub-bands within an IRD band are generated uniformly spaced, it makes sense to arrange sub-bands across the RF spectrum. The numbers in the sub-band range column indicate the covered sub-bands when numerated from the origin, where the sub-band number multiplied by 200 MHz (the sub-band bandwidth) yields the RF corresponding to the sub-band center frequency. Note that the sub-bands numbers are not necessarily integers, although their fractional part is the same across an entire RF band to avoid unnecessary overlaps.

The separation between LO frequencies within an RF band is an odd number in all ngVLA frequency bands except for Band 6. An even number of sub-bands between LO frequencies becomes an issue when a sub-band is centered around zero Intermediate Frequency (IF) because that will result in another sub-band centered around the edge frequency between consecutive IRD bands. To fully cover the edge sub-band, either the receiver bandwidth has to be increased, or additional CSP resources must to be used to processes the edge sub-band twice, each time generated from each IRD band. An alternative solution would be shifting the digitized spectrum before sub-band generation enough to minimize the portion of the edge sub-band beyond the receiver's fully compliant passband while still allowing fine digital sideband separation if needed. This has to be coordinated with the specific digital sideband separation technique implemented by the DBE. Even doing nothing could be an option, assuming the performance degradation at the frequencies out the receiver bandwidth is acceptable, or if those frequencies can be

discarded. Note that this issue does not generate a coverage gap, but in the worst case two additional FSPs could be required to process both halves of all the edge sub-bands at the optimal receiver's performance.

4 Very Coarse Channelizer Incorporation into the Digital Back-End

In the Frequency Slice Architecture [8], the VCCs are responsible for the generation of sub-bands or 'frequency slices'. They are located in the CSP building and connect to every FSP in the same 'tine' of the Trident through a passive optical circuit. This solution has proven problematic for meeting system requirements. Specifically, it conflicts with SYS0904: "If the digitized bandwidth exceeds the instantaneous transmitted and processed bandwidth, the system shall separate the digitized bandwidth into sub-bands for bandwidth selection, transmission and processing." This imposes sub-band generation at the antenna sites, which is reasonable for remote stations although it might not be necessarily the optimal solution for the antennas closest to the CSP building. Other conflicting requirement is SYS5601: "It is desirable to provide interfaces to enable commensal processing of the time-voltage stream from each antenna at the granularity of a digitized sub-band or smaller unit of bandwidth." However, the passive optical circuitry at the VCC output is based on point to point connections and thus is not suitable when a scalable solution is desired.

In the following, I explore a solution to the above issues based on generating sub-bands at the antenna stations. Sub-band generation at the CSP building was considered in the minimum delta solution described in [4]. Because the DBE now generates the sub-bands needed by the FSPs, ideally all the VCC functionality should be translated to the antenna as well, so that the whole VCC part is removed and the FSPs directly process the data streamed from the DBE. In addition to sub-band generation, the DBE must apply digital sideband separation corrections, search for and flag data samples corrupted by RFI, and format the data (including requantization) for its transmission over the corresponding fiber-optic network. In order to support the most distant LBA antenna stations, the ngVLA project should leverage successful e-VLBI projects and adopt similar strategies. One of the biggest successes of the Very-Long-Baseline Interferometry (VLBI) community is the development and adoption of the VDIF standard globally, and therefore, VDIF should be the DBE output format of choice. It has previously been noted [9] that the use of VDIF has many advantages. Just to mention a few, it enables early involvement of VLBI community experts and other stakeholders. It also facilitates the use of an existing (proven) software correlator during the early installation of the system, allowing to run in parallel similar activities for the hardware correlator, thus mitigating risks and critical paths.

Notwithstanding the above, there is one VCC function that cannot be carried out at the antenna. That is the required data buffering to support widely variable packet latency and rearranging packets at reception. The VCC incorporate very long signal buffers for coarse compensation of the bulk delay. Conversations with the National Research Council of Canada (NRC) reveal that the limitations of the current FSP hardware does not allow transferring these long buffers to the FSP part. The reference design assumes these buffers would additionally serve to allow variable packet latency of the data transport network between the antennas and the CSP.

At the time of writing, it is not clear what range of packet latency needs to be supported when using a commercial network. For the antennas in the core and the plains, it seems reasonable assuming that the data transport delay is relatively constant and can be accommodated by the short input buffers of the FSP. Therefore, any bulk delay correction can be timely performed by the DBE, which already uses timing information for timestamping the received data. However, this might not be the case for the most remote antennas, where differences in packet delays can easily exceed what is practically rectifiable. For this reason, I propose the following decision-tree solution for the bulk delay/variable latency compensation:

1. If future FSP hardware allows it, bulk delay and variable latency should be compensated by the FSP part; otherwise, and in the meantime,

2. Bulk delay will be compensated by the DBE, and where needed, additional hardware at the FSP input could be used to compensate for the bulk delay and variable latency.

5   Number of Frequency Slice Processors

The Trident solution described in the ngVLA CSP reference design [2] includes 150 FSPs, distributed in three sets of 50. This number of FSPs was selected so that the CSP comprises enough resources to process 20 GHz of instantaneous bandwidth in correlation mode, in addition to 8 GHz in one beamforming mode. The solution is based on each FSP FPGA processing 10 sub-bands or 'frequency slices', quantized using 16 bits, each sub-band corresponding to a different antenna. The use of 16-bit quantization was motivated by the RFI environment where the Square Kilometre Array (SKA) Mid-Array is located. However, it is not clear that the ngVLA RFI environment will require such dynamic range, at least at the higher frequency bands, [2] which according to NRC would potentially allow a significant reduction in equipment needs for the CSP. This result is based on the assumption that a new FSP operation mode for correlation that uses 8-bit sub-band quantization can be designed to handle two sub-bands per antenna, i.e., 20 sub-bands per FPGA or twice the bandwidth per FSP.

Assumption 3 – *8-bit quantization provides enough sub-band dynamic range for ngVLA Bands 5 and 6 to achieve the required RFI tolerance.*

Given the above assumption and the sub-band bandwidth in previous sections, only 50 FSPs are needed for processing 20 GHz of instantaneous bandwidth using 8-bit sub-band quantization. That is the maximum bandwidth per antenna that the CSP must process for ngVLA Bands 5 and 6. Namely, 50 FSPs will process 100 sub-bands, 200 MHz wide each, for 20 GHz in total. Just as a reference, 8-bit quantization provides at least 42 dBc spurious free dynamic range for a tone at half full-scale amplitude (90 dBc in case of 16-bit quantization). Bands 1 to 3 each generate less than 50 sub-bands (see Table 2) and can therefore be fully processed using 16-bit sub-band quantization. I will discuss Band 4 after the next paragraph.

Simultaneously to 20 GHz of instantaneous bandwidth in correlation mode, system requirements demand that the CSP process 8 GHz of bandwidth in a different observing mode, especially pulsar modes. Note that processing two sub-bands at a lower quantization resolution is, in principle, not available to any of the beamforming observing modes. Hence, at least additional 40 FSPs are needed for processing 8-GHz in these modes. As a result, I propose changing the number of equipped FSPs from 150 in the reference design to only 90. Note that this figure is not final whatsoever. Additional FSPs can be installed as needed, since the FSP part of the FSA is fully scalable in number of FSPs. There is just one caveat: At least both the CSP Back End (CBE) and the input network fabric (see Section 6) must be scalable as well.

As any subarray can make use of any number of available FSPs as needed, the proposed 90 FSPs allow full coverage of Band 4 at 16-bit sub-band quantization.[3] That would occupy 68 FSPs, still leaving other 22 FSPs (4.4 GHz at one sub-band per antenna per FSP) available for other subarrays or FSP modes. On the other hand, only

---

[2]Please note that I am referring to the dynamic range required for sub-band re-quantization. This dynamic range will not depend on the digitizer bandwidth.

[3]There are also some bandwidths considerations that may limit the number of sub-bands processed with 16-bit quantization. See Section 6.

34 FSPs are needed when 8-bit quantization is used throughout the entire Band 4, which would allow 11.2 GHz for simultaneous observations using 16-bit quantization. As stated before, more FSPs could be installed if it became necessary.

Finally, it is also worthwhile noting that the 16-bit and 8-bit sub-band quantization are provided through different FSP modes. This is important for simultaneous subarray operation, as sub-bands requiring 16-bit quantization cannot be processed by FSPs operating in 8-bit correlation mode. The other way around is possible, but only at one sub-band per antenna per FSP. Thus, it becomes hard to justify why using 8-bit quantization instead of 16 bits for those subarrays too, except for a lack of available FSPs or due to data rate limitations, as explained in the next section.

6   Switched Fabric for the DBE-FSP Interface

A necessary condition to achieve the required flexibility in commensal observations is the capability of routing any of the generated sub-bands (see Table 2) to any FSP. The use of switched fabric to interconnect DBEs and FSPs readily solves the problem and offers a scalable solution for the future. In the following, I describe a straw-man design based on such a switched fabric. As an alternative, [4] describes a less flexible but power-efficient solution that keeps the reference design optical fiber mesh at the FSP input.

Table 3 gathers the maximum data rates that the switched fabric would need to support. I propose supporting at least Bands 5 and 6 maximum data rates wherever is practical, which should cover at least the antennas in the Plains. The remote antennas will likely have to trade number of bits for bandwidth in many cases, as the maximum data rates in the table seem out of reach, at least in the medium term. Note that the DBE will allow choosing the quantization on a per sub-band basis, anywhere between 2 and 16 bits. The number of bits in the table indicate the maximum number of bits supported for full-bandwidth transmission of an RF band. As regards Band 4, I do not think it needs full-band 16-bit coverage. Instead, either a combination of 16 and 8-bit quantization can be used, or full-band 12-bit quantization would fit within the proposed supported data-rate.

Using a conservative 27% headroom for overheads, I will assume a data link capacity of 1000 Gbps between the antennas and the CSP for the straw-man design. Since data is fed into each FSP FPGA through a set of 12 GXT transceivers (up to 28.3 Gbps) [8, 10], I will base this straw-man design on 200GbE links. Each of these FPGAs receives one sub-band from 10 antennas in 16-bit quantization modes, or one pair of sub-bands from 10 antennas in 8-bit quantization mode. In either case, the input data rate is 145 Gbps.

To simplify the terminology, sub-bands quantized using 16-bit resolution will be referred in the following as high-res sub-bands, while a low-res pair refers to two sub-bands sampled using 8-bit resolution. On the DBE side, the output data can be distributed among the five 200GbE links using at most 10 high-res sub-bands per link for Bands 1 through 3, or 10 low-res pairs (22 sub-bands) for Bands 5 and 6. The resulting data rate in both cases is 145 Gbps. Obviously, in this straw-man design full high-res cannot be supported throughout Band 4 and some concessions must be made. One solution to keep the same data rate could be distributing the 10 slots per link as 6 high-res sub-bands and 4 low-res pairs. That amounts to 30 high-res sub-bands (slightly less than half Band 4 bandwidth) plus 20 low-res pairs (40 sub-bands). An alternative solution would be using 12-bit quantization in all Band 4 sub-bands and then transmit up to 14 of these sub-bands through each link. In this case, the resulting data rate per link would be 152 Gbps. [4]

---

[4]Or 165 Gbps if 13-bit quantization becomes possible. 14-bit quantization, which yields 178 Gbps, would probably exceed the practical limit of the link capacity.

RF Band	Sub-band Quantization	# of Sub-bands	Max Data Rate
1	16 bits	12	174 Gbps
2	16 bits	44	636 Gbps
3	16 bits	42	607 Gbps
4	16 bits	68	983 Gbps
	Hybrid 16/8 bits		709 Gbps
	12 bits		738 Gbps
5	8 bits	100	723 Gbps
6	8 bits	100	723 Gbps

Table 3: Maximum data rates per RF band. The hybrid mode for Band 4 assumes 30 sub-bands with 16-bit quantization and 38 sub-bands with 8-bit quantization. Both Bands 5 and 6 actually consist of more than 100 sub-bands, but 100 sub-bands (20 GHz) is the maximum required (as a goal) for data transmission [5].

As every single FSP FGPA processes a set of 10 antennas, the connectivity problem can be simplified as how to connect each antenna set to one FPGA of every FSP. Then, the chosen solution can be replicated 27 times for full support of the 263 antennas in the array (or any other number, should the array configuration change). This topology, as the one in the reference design, allows full flexibility of subarray configuration, that is, which antennas are in which subarray. Using a single Ethernet switch per antenna set fully constrains the switching problem (or bandwidth-antenna transposition) within the switched network. The number of 200GbE ports per switch would be 50 ports for the DBE part (5 links times 10 antennas), plus the total number of FSPs, which is 90 as described above. That results in a single switch with 140 ports, which might be slightly superior to what is commercially available now, but likely available in a near future.

In case the required number of ports per switch exceeded technical capabilities in the future, maybe due in part to the addition of new FSPs or custom back-ends, a divide-and-conquer strategy can be used to solve the problem. For example, one could use two switches per antenna set, each one connecting the DBEs to only one half of the FSPs. In order to balance the workload, it would be preferable using 100GbE data links, that is 10 links per DBE output, and 2 links per FSP FPGA input. As a result, each switch requires again 140 ports (50 for the DBE part, 90 for the FSP part), but 100GbE ports. That is equivalent to half the number of ports that was required above,[5] which leaves plenty of room for custom back-ends. Nonetheless, in addition to power and cost considerations, there are other factors to take into account when deciding the best solution. For instance, one factor to be considered is the granularity of the transmitted data when balancing the transmitted data across the different links. It might happen that some links must carry one additional sub-band when the number of sub-bands is not a multiple of the number of links. This will produce a excess data rate in those links, which is more significant the more links are used for the DBE output, and eventually some additional capacity could be needed if too many links are used.

Finally, the proposed solution requires certain coordination or switching capability among the different devices comprising the DBE. The reference design [3] accounts for one or more FPGAs. Using the Trident design as a reference, up to four FPGAs might be needed, one per active IRD module. Note from Table 2 that a single IRD module (or FPGA) can lead to more than half the transmitted data rate, 420 Gbps, which equals 29 high-res sub-bands or 58 low-res pairs. Therefore, in case of a DBE subsystem composed of multiple FPGAs, some additional

---

[5]Many commercial switches allow splitting one 400GbE port into two 200GbE ports, or four 100GbE ports.

switched fabric would be necessary to combine and/or distribute the data among the different links.

7 Remote Antennas

Herein, remote antennas are those connected through a network fabric whose latency exceeds the buffering capabilities of the FSP. This is the usual case when using commercial networks. Data packets can be lost, arrive in different order and/or with very different delays. In that case, trying to correct the bulk delay at the DBE becomes useless. In addition, the DBE subsystem should support flexible sub-band quantization, anywhere between 2 and 16 bits, in order to provide the user with the freedom of trading in number of bits for more bandwidth. Note that would imply a relaxation of current digital efficiency requirements. Any sub-band pair with quantization of 8 bits or less, can be processed in 8-bit mode, and all sub-bands can be processed in any 16-bit mode via proper zero-padding.

From conversations with NRC, coarse delay corrections are applied by the VCC as the FSP cannot perform this task due to limited memory resources. Whereas transport latency is quite variable for the remote antennas, a signal buffer is required at the CSP end, which needs to be implemented by additional hardware other than the FSP. I suggest developing new hardware to carry out this task, as the TALON hardware seems to me quite an expensive option for this simple task. A first cut design would be based on a small FPGA or System on a Chip (SoC), which would receive data addressed to a specific FSP FGPA from the switched network, reorder data packets through an external DRAM-based data buffer, apply a bulk delay correction, and feed the data into the corresponding FPGA. For that reason, I call such a device a Remote Antenna Buffer-ing Transceiver (RABiT), and it would be inserted between the FSP FPGA and the network switch.

8 Conclusion

This proposal of changes to the reference design focuses on the most notable open issues of the CSP. First, any redundancies between the CSP and the DBE is removed by clearly assigning the generation and selection of sub-bands to the DBE at the antenna. Second, the utilization of switched fabric at the CSP input guarantees a scalable long-term solution which removes unnecessary internal partitions and serves as a gateway for custom back-ends. This is a working document open to discussion, and better alternatives could be possible. As compared to the reference design, the proposed changes imply the removal of nearly 50% of the CSP hardware. That is more than sufficient to compensate for the cost of the added elements such as the network switches or the RABiTs.

9 Acknowledgments

References

[1] M. Morgan and S. Durand. *Integrated Downconverters and Digitizers Design Description*. Tech. rep. 020.30. 15.00.00-0002-DSN-A. ngVLA Reference Design, 2019. URL: https://ngvla.nrao.edu.

[2] O. Yeste Ojeda. *Central Signal Processor: Preliminary Reference Design*. Tech. rep. 020.40.00.00.00-0002-DSN-A. ngVLA Reference Design, 2019. URL: https://ngvla.nrao.edu.

[3] J. Jackson et al. *Digital Back End/Data Transmission System: Reference Design Description*. Tech. rep. 020.30.
25.00.00-0002-DSN-A. ngVLA Reference Design, 2019. URL: https://ngvla.nrao.edu.

[4] O. Yeste Ojeda. *ngVLA Electronics Memo No. 4. Trident 2.0 Concept: A Minimum Delta Update to the Central Signal Processor Reference Design*. Tech. rep. 2020. URL: https://ngvla.nrao.edu.

[5] R. Selina et al. *System Requirements*. Tech. rep. 020.10.15.10.00-0003-REQ-B. ngVLA Reference Design, 2020. URL: https://ngvla.nrao.edu.

[6] W. Shillue. *Local Oscillator Reference and Timing Design Description*. Tech. rep. 020.35.00.00.00-0002-DSN-A. ngVLA Reference Design, 2019. URL: https://ngvla.nrao.edu.

[7] M. Morgan. "IRD and Front-end Interfaces". ngVLA Antenna Electronics Workshop. Aug. 2019.

[8] B. Carlson and M. Pleasance. "Trident Correlator-Beamformer for the ngVLA: Preliminary Design Specification". 2019. URL: https://ngvla.nrao.edu.

[9] V. Dhawan. private communication. 2019.

[10] *Intel® Stratix® 10 SX SoC FPGAs*. https://www.intel.com/content/www/us/en/products/programmable/soc/stratix-10.html. Accessed: 2020-08-28.