SP-725 Perentie Engineering Subarray

Grant Hampson and John Bunton

12th February 2020

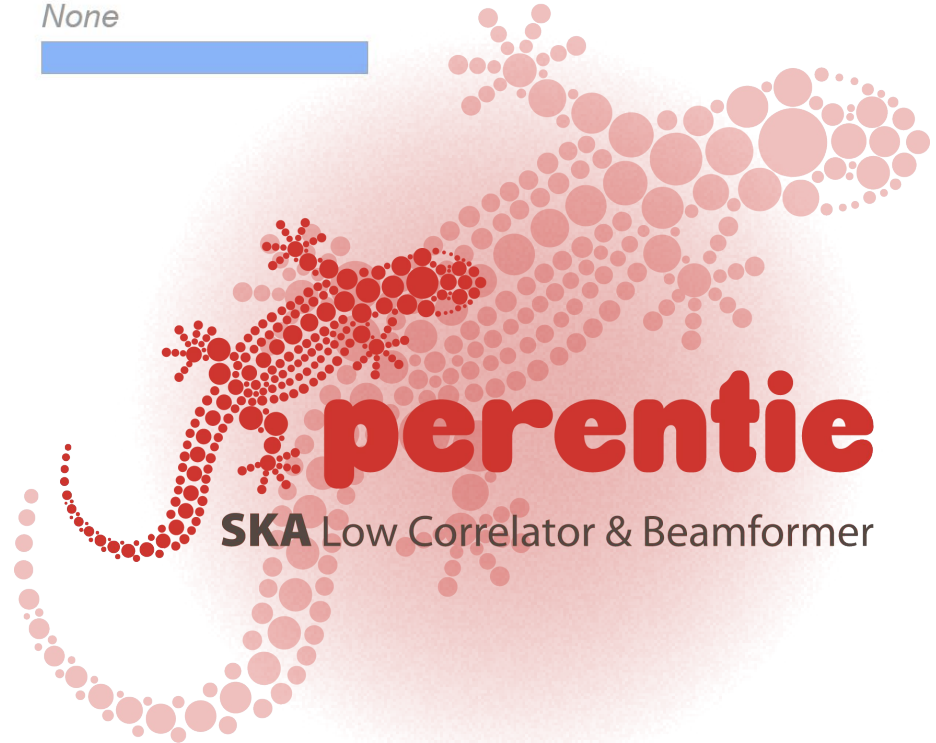SP-725

Perentie Subarray Options

CSP-Low (Perentie) Roadmap
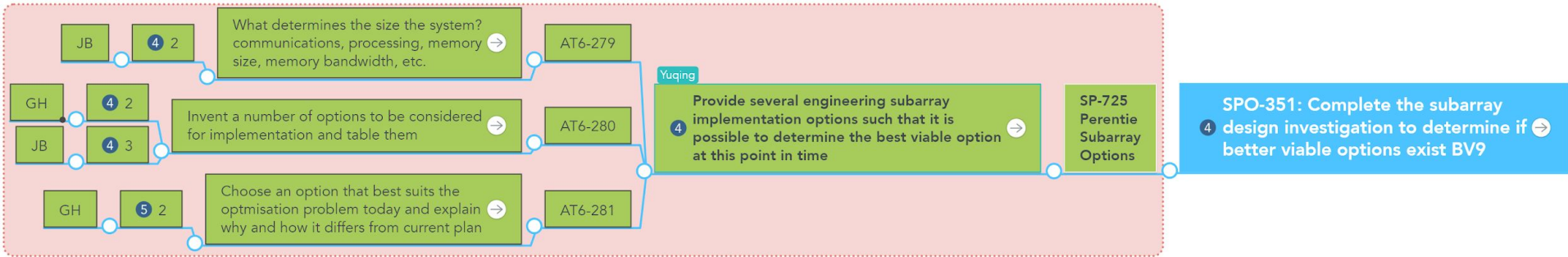
Bridger, Alan

None

perentie

SKA Low Correlator & Beamformer

Feature Context

- https://jira.skatelescope.org/browse/SPO-351
- https://jira.skatelescope.org/browse/SP-725
- https://jira.skatelescope.org/browse/SP-538 (old engineering subarray feature)

- Benefits:
  - To be sure Perentie is implementing an architecture than implements the required engineering subarray functionality in construction
  - Full compliance to the requirements (instead of partial)
  - Easier upgrade and testing path for future versions of software and firmware

Documentation

- ## Rewritten Chapter 9: Dynamic View of the Low.CBF DDD

  - https://docs.google.com/document/d/12GQaN2LIqUr2dSnQE0IGAMe5RXBFrPfj/edit#

7   DYNAMIC VIEW

The Low.CBF system consists of control computers, network switches, a VLBI server plus Perentie rack units, each containing up to 12 processing FPGAs. The Perentie rack units implement the signal processing task while the computers and switches provide monitor and control as well as distribute the wall clock time. The Perentie rack units are grouped into three main systems that implement either station-based processing, correlation, or beamforming. The station-based processing is implemented in 43 Multi Station Processors (MSPs), with each processing data from 12 LFAA stations. Correlation is implemented in 12 Correlation Slice Processor (CSPs) and beamforming in 12 Beamformer Slice Processors (BSP).   The MSPs connected to the CSPs and BSps via passive optical cross connects. The VLBI server ingests pulsar timing beams from the BSPs and reformats them for VLBI.
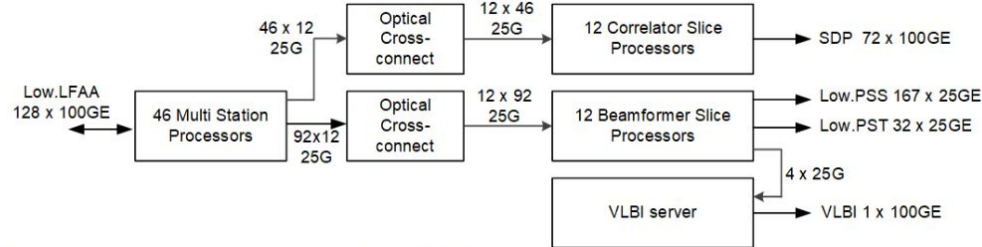

*Figure 1 Overview of the correlator and beamformer (REDO diagram)*

- ## The following is a short summary

Existing Perentie System

- Uniform array of FPGAs of size 6x6x8=288.
- Each FPGA contains all three major functions; filterbanks, beamformers and correlators.
- The 300MHz (384 coarse channels) is evenly distributed and processed across the entire 288-FPGA array.
- If distributed evenly each FPGA processes 3456/288=12 fine frequency channels (226Hz) per coarse channel. Pre Filterbank routing raises this to 192 (44kHz out of 781kHz)
- Removing an FPGA results in missing fine channels in many coarse channels
- There are also further data losses as the FPGAs also act as data switches
- Creating an engineering subarray of any size results in missing frequency channels scattered across the entire 300MHz.

Perentie "Frequency Slice Processor" Concept

- Similar concept to Talon FSP concept to enable an Engineering Subarray (as well as many other system advantages)
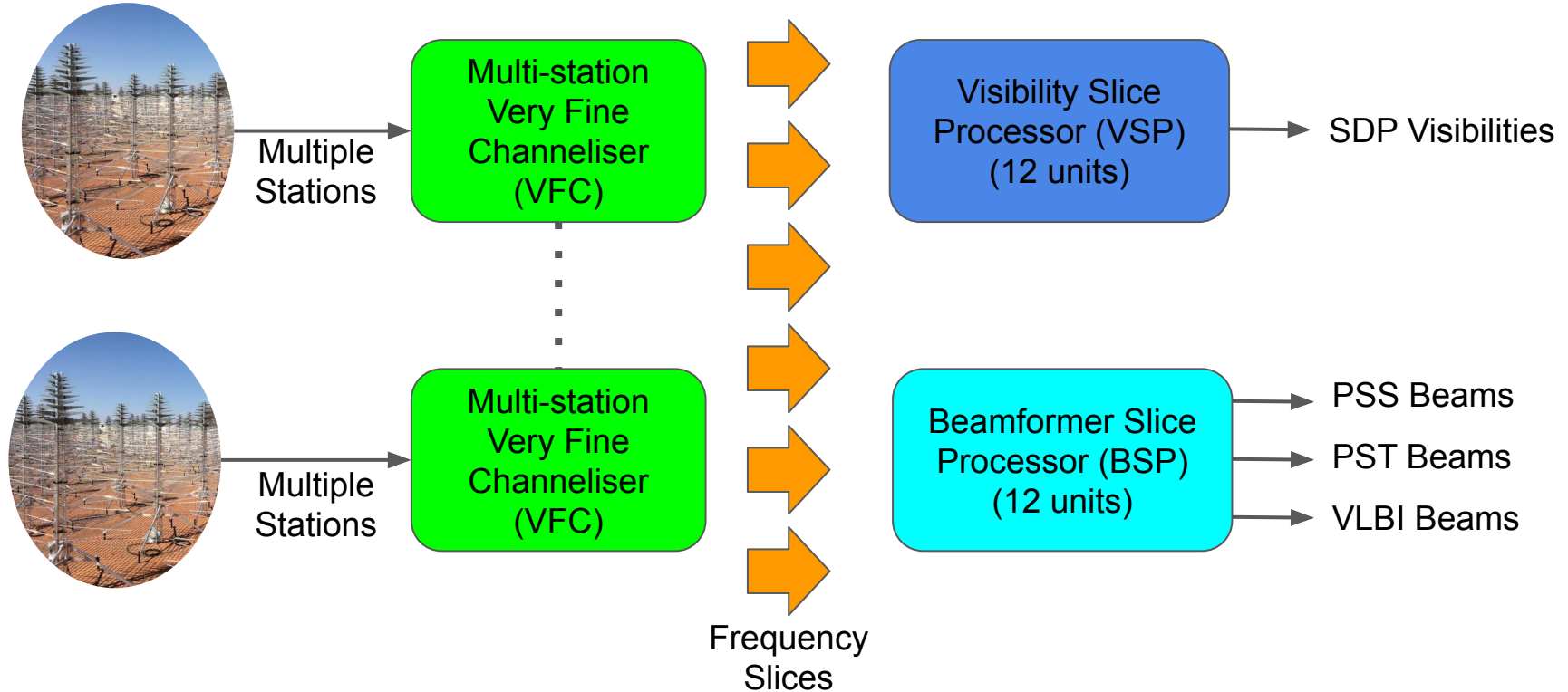
Perentie "FSP" Advantages

- Improved Engineering subarray functionality
  - Each VFC/VSP/BSP can become an engineer subarray
- Fine frequency striping failure effect is completely removed

- Can individually add capacity to all processing stages (VFC, VSP, BSP)
- Scaling up during construction (12-station and 25MHz increments)
- Can build as little or as much of the system as required independently

- Improved reliability in Correlator and Beamformer (N+1)
- Gemini LRU spares are stored in the system
- Small reduction in the number of Gemini possible

Basic Comparison to Mid.CBF

Low has stations with multiple beams

Similar optical cross connects

Low has sufficient hardware - no output switching

Low has very fine channelisers (coarse in LFAA) and multi-stations per VFC

Low has 12 Visibility and Beamformer Slices, each 25MHz

Band-specific Very Coarse Channelizer (VCC) 1

Band-specific Very Coarse Channelizer (VCC) 2

FSP-N

Function-configurable Band-invariant Frequency Slice Processor-1 (FSP-1)

Function-configurable Band-invariant Frequency Slice Processor-2 (FSP-2)

FSP-1 Data products

FSP-2 Data products

FSP-N Data products

Frequency Slices

Output Switch

Visibilities to SDP

PSS Beams

PST Beams

VLBI Beams

VLBI visibilities

Low and Mid comparison

- Mid 192 antennas, Low 512 stations
  - So Low is 512x512/197x197 = ~7 times larger
- Mid 4GHz bandwidth, Low 300MHz
  - So Mid is 4000/300 = ~13 times larger
- So overall Mid Correlator is about 2 times larger than Low Correlator
- Mid PSS Beamformer is also 3 larger at 1500 beams compared to 500

- 20 VCC units - each 5 LRUs, each 2 FPGAs = 200 FPGAs
- 27 FSP units - each 10 LRUs, each 2 FPGAs = 540 FPGAs
  - Total of 740 Mid FPGAs
- Low is 288 FPGAs, or 2.6 times smaller
  - But Mid can't do full BW Correlation and Beamforming in parallel? Maybe FPGA capacity is different?

Number of VFC FPGAs

- Completely dependent on LFAA interface data rate

- Currently 256 x 40GbE links
  - 2 stations per 40GbE link
  - Occupancy is ~22Gbps, or 11Gbps per station

- Why do Perentie care?
  - Each Gemini card has 4 QSFP interfaces (can be equipped with 40 or 100GbE links)
    - 512/8 is 64 Gemini cards
  - If 100GbE is used then could have 8 stations per link - one quarter of the links
    - So 64/4=16 Gemini cards, which is significantly smaller in size and lower cost
  - Suggestion is 12 stations per Gemini using 3x40GbE (2 stations each) + 1x100GbE (6 stations)
    - Requires 512/12=46 Gemini cards (8 are 4x40GbE), only remote are 100GbE

- There are some memory and communication limitations for filterbank processing, but in general these limits are not reached

- Like to encourage LFAA to use higher speed links
  - Nx10G is ageing technology, whereas Nx25 is modern technology
  - Links don't have to be 100% loaded (just needs to be efficient and have some forward capacity)

VFC FPGA Comparison to CDR

- More filterbanks per FPGA - but still lots of compute within FPGA though
- HBM is the critical path FPGA resource - memory requirements have increased
  - 12-stations x 2-pol x 384-coarse x 2x8-bits x 926kHz = 137Gbps = 17GBytes/second
  - Store in double buffered 0.1s blocks = 3.4GBytes (half the 8GB HBM memory size)
  - Filterbanks "may" clock faster as filterbank priming takes greater % of time
    (clocks "may" also reduce as we can use more filterbanks lowering the clock)
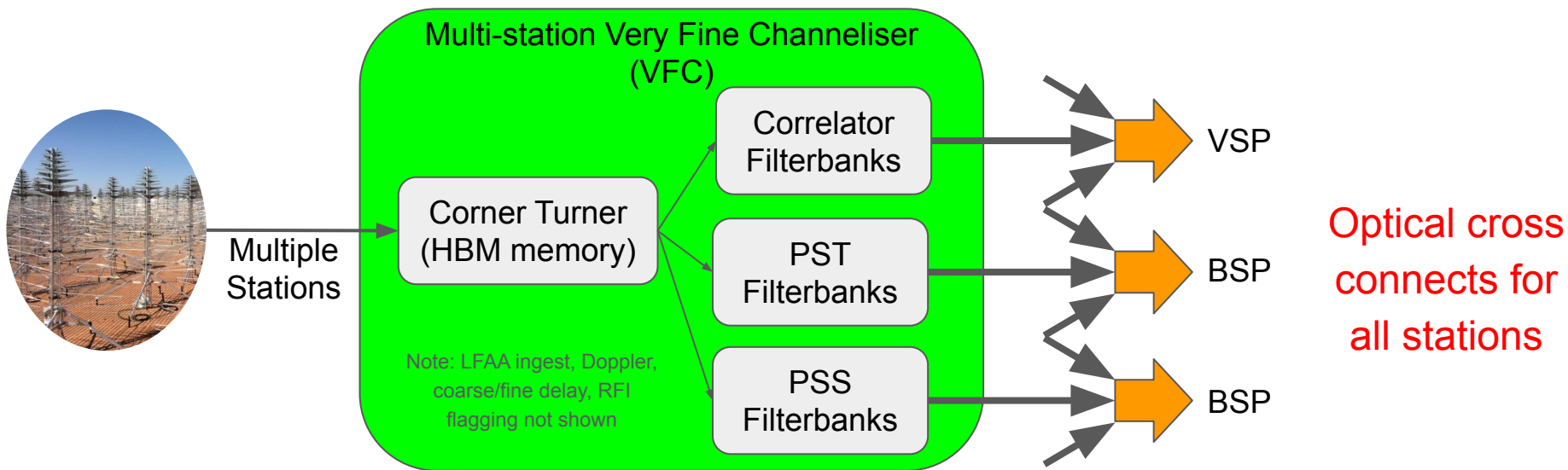
Number of VSP/BSP FPGAs

- 12 frequency slices in both VSP/BSP
  - 300MHz / 12 = 25MHz each slice (or subrack)
- So 25MHz or 384/12=32 coarse frequency channels sent to each VSP/BSP
  - One whole frequency channel processed by a VSP/BSP - not a small part of it
  - Each LRU within a VSP/BSP processes a fraction of the coarse channel
- Full freedom in allocating coarse frequency channels to FSP
  - Enables zooms to be distributed over available VSP resources
  - Could send same channel to every Gemini and compare all outputs (good test)
  - Enables an engineering subarray to be allocated continuous spectrum
- How to handle Visibility zooms?
  - Each subarray can still have four zooms
  - Coarse channel distribution based on zooms to even out VSP processing load
- Change from CDR design
  - Grouping of the hardware and the distribution of data
 - The core processing itself is unchanged

VSP FPGA

- Input data is received (1-fibre = 12-stations, 2-pol, 32-coarse)
  - e.g., 512/12=43-fibres total   (4 or 5 per FPGA)
- Distribute fine channels within subrack over 10-Gemini
  - Full 12x12 cross connect installed (2 spare slots only 10 active)
- Fine channels are processed in Correlator and sent to SDP
  - HBM size is 512-stations x 2-pol x 32-coarse x 3456 x 2x8-bits x 226Hz x 0.9s /10 = 4.3GB
  - Integration time for Correlator Visibilities can be an integer subset of 0.9s (0.45s, 0.3s, ...)
  - Two Gemini cards per SDP output - so combiner required
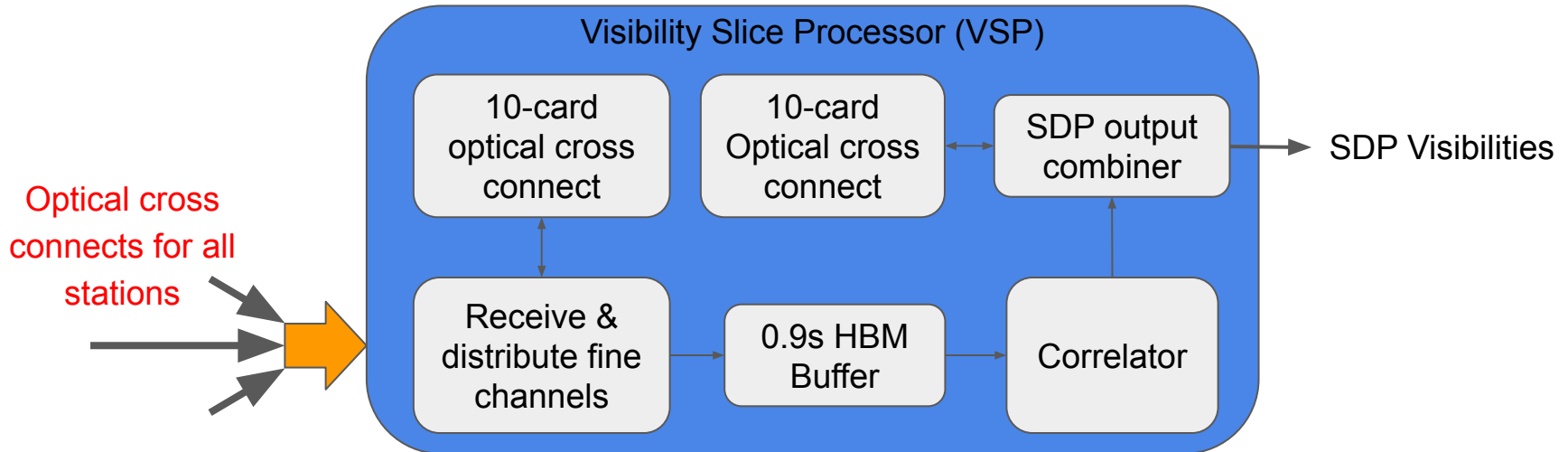
BSP FPGA

- 2 lots of input data is received (1-fibre = 12-stations, 2-pol, 32-coarse)
  - PSS and PST data (86 fibres across 9 FPGA - 9 or 10 per FPGA)
- Distribute fine channels within subrack over 9-Gemini
  - Full 12x12 cross connect installed (3 spare slots only 9 active)
- Fine channels are processed in Beamformers and sent to PSS/PST
  - HBM size is 512-stations x 2-pol x 32-coarse x 3456 x 2x8-bits x 226Hz x 0.1s /10 = 0.43GB
  - Need output cross connected between BSPs to accumulate all frequency channels
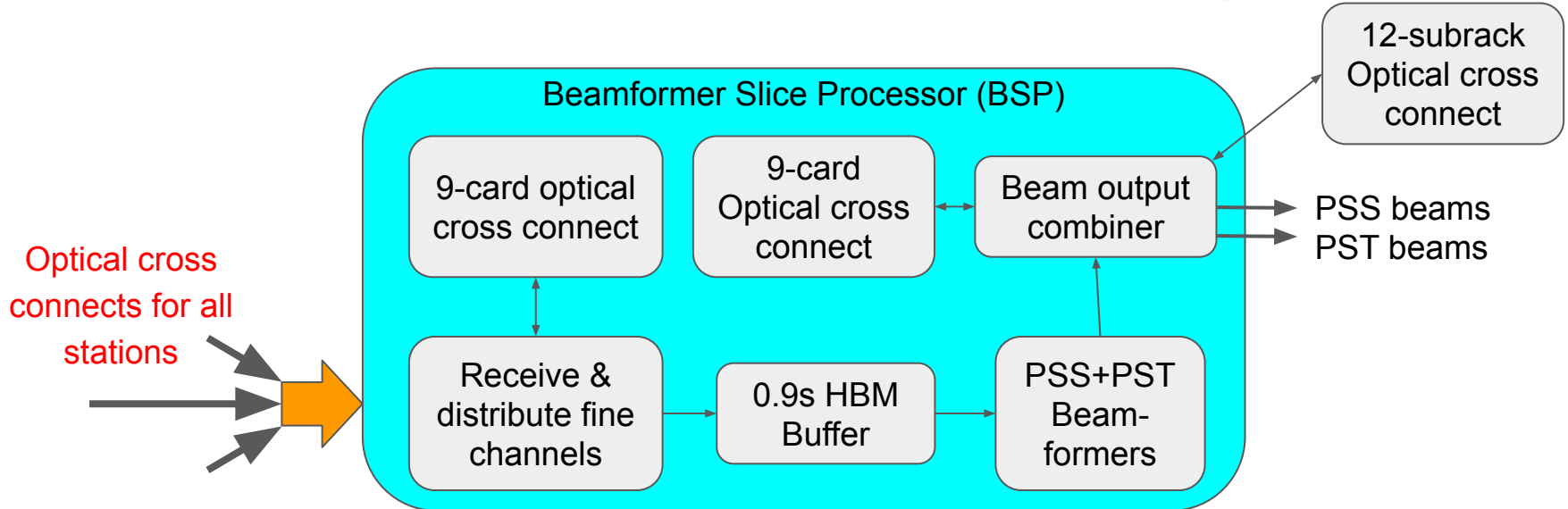
Engineering Subarrays

- Any VSP/BSP slice can become an engineering subarray
  - The whole 25MHz subrack resource can be allocated to the engineering subarray
  - 1 out of 12 = 8%
  - Or multiple subracks can be allocated to an engineering subarray

- The VFC can also form an engineering subarray
  - Whole subracks cannot be allocated to an engineering subarray as this would be a significant number of stations
  - In this case each individual Gemini can be allocated to an engineering subarray
  - 1 Gemini = 12 stations = 2.3% of the total array
    - Downside: choice of stations is fixed by connections to Gemini
  - The only resource being controlled is the Gemini itself

	Unified 288 Design	Separate Function Design
Lack of interaction between functions during debug	No, Possible interaction with each recompilation	Yes
Independent work on functions	No	Yes
Independent VHDL compilation	No	Yes
Regular internal FPGA structure	Poor	Better
Compatible with existing LFAA links	Yes	Yes
Upgradeable to fully 100G LFAA links	No	Yes
Easily increased compute capacity for any function	No	Yes
Engineering subarray for station based processing	Yes	Yes
Engineering subarray for correlator (continuous bandwidth for astronomy)	No	Yes

Acceptance

- A report has been written that contains a number of options and recommendation
  - This presentation
  - Rewrite of DDD Section 7
 - PI#5 documents

- Circulate the report to relevant stakeholders and request feedback
  - Feedback received in PI#5
  - This is a solidification of the Option 3 shown in PI#5
    - which has been further improved upon with the 25MHz per subrack

Key Results

- KR: provide a document describing a number of options to make the LOW CBF improve compliance with the engineering subarray requirements.
  - Provided in PI#5
- The options shall contain design, estimated cost and impact on schedule and interfaces.
  - Design has been shown - see documentation
  - Estimated cost - no cost increase, possibly small cost reduction in hardware
    - increase in 1 rack required, reduction in Gemini by 288-274=14 (easily pays for rack)
  - Impact on schedule - same schedule predicted
  - Interfaces - change in LFAA interface
    - 40GbE interfaces still exist
    - **However, 100GbE links from RPFs (with 6-stations) go direct to Low.CBF and do not get broken up into 40GbE links**

Risks

- Is there is something that we have overlooked?
  - The team have looked over the concept and couldn't find anything (the proof of the pudding is in the eating)
- Changes to the baseline CDR design are not agreed by everyone
  - Please speak now or forever hold your peace
- May add costs to construction
  - No added costs to construction
- LFAA can't provide 6-stations on one 100GbE link from RPF's?
  - **Need SKAO to check this**
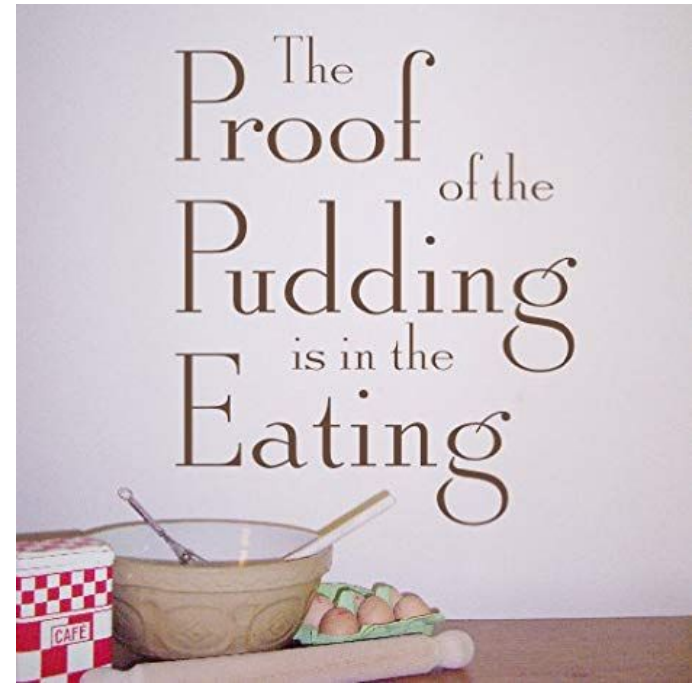
Summary

- There are no apparent shortcomings to the design
- Its has far greater reliability and availability
- There is improved compliance compared to current design
- All requirements are now compliant
- The new design has a smoother rollout
- Cost can be reduced by reducing stations or bandwidth
  - The schedule is still the same!
- It is possible to create an engineering subarray with little impact on the system (and possibly none - TBD)
- There is a reduction in number of Gemini = 12 x (10+9) + 46 = 274
- Will proceed with ECP in PI#6