DSA-2000 Document No. 00018

PSR/FRB search subsystem

Liam Connor, Vikram Ravi, Myles Sherman, Casey Law, Martin Pokorny, Caltech

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  | | --- | --- | --- | --- | | Version: | 1 |  |  | | Version date: | 19 September 2023 |  |  | | Original date: | 19 September 2023 |  |  | | Controlled document: | Yes |  |  | | WBS Level 2: | PT |  |  | | Document type: | DES |  |  | |
| |  |  |  |  | | --- | --- | --- | --- | |  | **Name** | **Signature (PDF)** | **Date** | | **PI** |  |  |  | | **Subsystem Lead** |  |  |  | | **Project Engineer** |  |  |  | | **Project Manager** |  |  |  | |  |

Revision History

| **Ver.** | **Date** | **Sections Affected** | **Reasons / Remarks** | **Author(s)** |
| --- | --- | --- | --- | --- |
| 1 | 2023-09-19 | All | Original | Liam Connor |

Table of Contents

[1 Introduction 4](#_Toc874544560)

[2 Pulsar search 4](#_Toc529370301)

[3 FRB search 5](#_Toc542815544)

[4 Data products 6](#_Toc1907276343)

[5 Requirements 6](#_Toc1985837423)

Abstract

The pulsar and fast radio burst (PSR/FRB) search subsystem is responsible for forming beams from voltage data in the standard packet format produced by RCF, passing those beams through a corner turn to search nodes, and performing single pulse and periodicity searches on those beams. In this document we define the specifications of a standalone PSR/FRB subsystem.

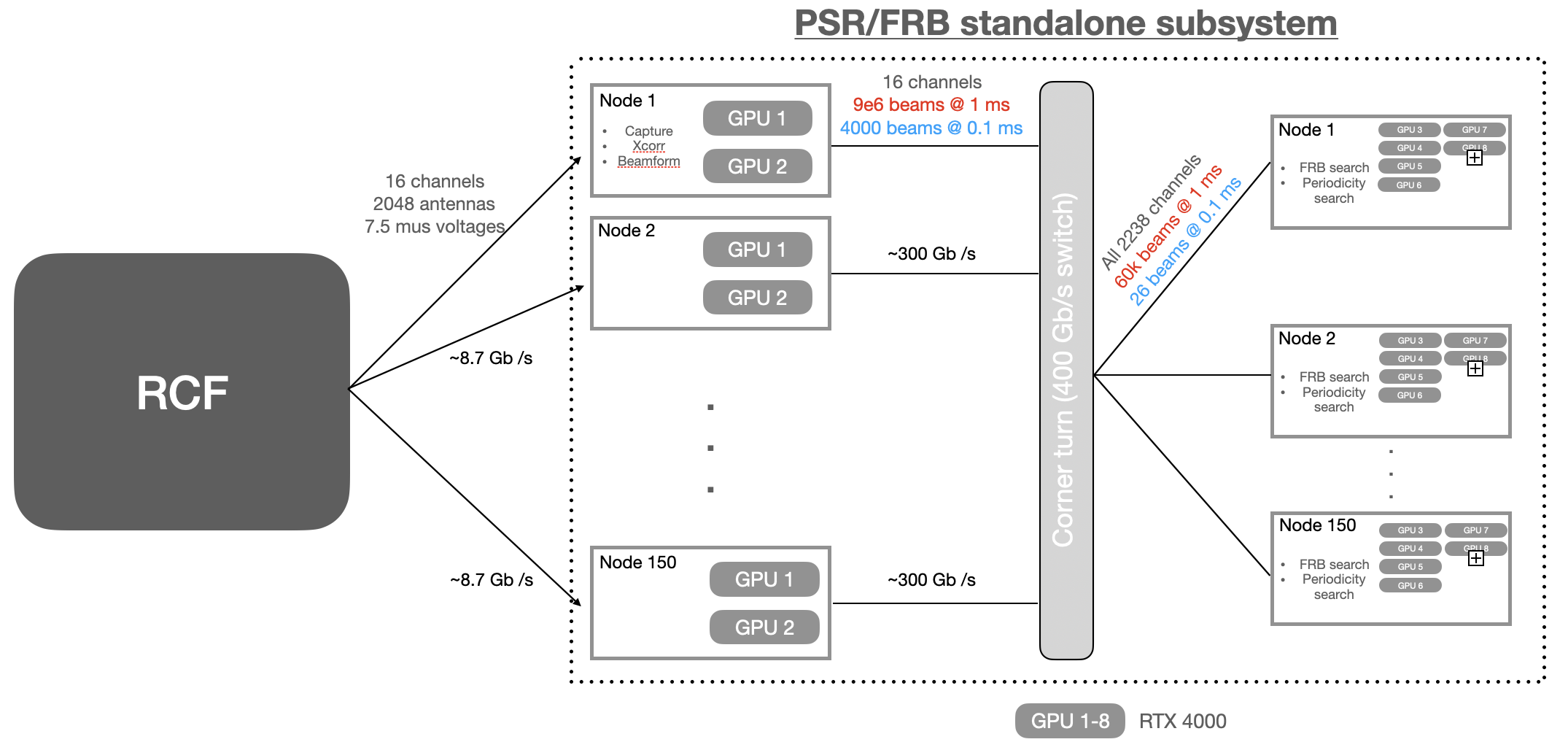
# Introduction

The DSA-2000 is poised to discover a large number of FRBs and pulsars. This is highlighted as Key Science Projects (KSPs) 5 and 6 in the MSRI pre-proposal, with significant impact on KSP1. However, the array’s filling factor of ~1e7 requires that we form and search an enormous number of beams to take advantage of DSA-2000's large field of view. We have developed solutions for both pulsar and FRB search that will reduce the complexity of this subsystem, which we outline in Section 3. and 4. respectively. The specifications of the PSR/FRB subsystem are given in Table 1. In brief, we will have 150 compute nodes, each with 8 x NVIDIA Ada generation RTX 4000 GPUs and ~2 TB of RAM. Two of the eight GPUs will be devoted to beamforming, three will be used for pulsar search, and the remaining three will do the FRB search. We also require 400 GbE interconnect between these nodes.

**Input data:** The PSR/FRB subsystem will receive packets directly from the RCF, form high-time resolution visibilities, and generate beams to be searched for single pulses as well as periodic signals. Rather than receive a copy of the output of the RCP, we elect to run our own instance of the correlator with high time resolution output. Each of the 150 PSR/FRB nodes will receive 7.5 microsecond voltage data for all antennas but just 16 frequency channels. The packet format can be identical to the data input to the imaging nodes.

**Beamforming:** From this voltage data, we will form all 2e6 visibilities in the tensorcore correlator with 0.96 ms integration time, which will be beamformed to produce 9e6 FRB search beams. Roughly 4,000 pulsar search beams will be formed directly from the voltage data with an output time resolution of 0.1 ms. We will buffer voltage data to enable saving formed voltage beams around the location of an FRB candidate.

**Corner turn:** Data must be swapped from a subset of frequency channels and all beams on a given server to a subset of beams and all frequency channels. This is the corner turn. The data rate coming out of each beamforming node is roughly 320 Gb/s. We therefore require 400 Gbe interconnect for the corner turn. From here, 6e4 beams and all 2400 frequency channels at 0.96 ms are sent to the search machines. At most 4,000 pulsar search beams with 0.1 ms sampling will be sent as well.



# Pulsar search

In order to make pulsar detection tractable, we wish to search significantly fewer beams than the ~1e7 required to blindly search the full primary beam. This can be done by placing our pulsar search beams on continuum candidates detected in the first year’s CASS. Candidates will be selected based on spectral index, their likeness to a point-source, polarization, scintillation, as well as through anti-coincidencing with known extragalactic sources. We estimate that this will reduce the number of candidates, and therefore pulsar beams, to no more than ~4,000 per 618 second pointing. We plan to search periods between 1ms and 10s and accelerations between +- 500 m/s^2.

After the corner turn, 27 pulsar beams will be sent to each search node. A full 618 s pointing requires 107 GB of memory which must be split across 3 GPUs. Assuming RTX 4000 cards with 20 GB of GPU memory, we cannot fit all 33 beams on the 3 GPUs. We must therefore process the beams serially in chunks. This is possible because the system is limited by global memory bandwidth and not compute limited.

Table 1. Parameters of the PSR/FRB search subsystem, including both beamforming and the timedomain searches.

|  |  |  |
| --- | --- | --- |
|  | Pulsar | FRB |
| Number of beams | ~4,000 | ~9,000,000 |
| Radio band | 700—1100 MHz | 700—1100 MHz |
| Bandwidth | 325 MHz | 325 MHz |
| Sampling time | 0.1 ms | 0.96 ms |
| Number of channels | 2400 | 2400 |
| Channel width | 130.2 kHz | 130.2 kHz |
| Max DM | Variable | 10,000 pc cm\*\*-3 |
| Number of DM trials | <1000 | ~2800 |
| Compute nodes | 150 | 150 |
| Pointing time | 660 s | N/A |

# FRB search

As with the pulsar search, we seek to reduce the computational complexity of our single-pulse search without sacrificing appreciably in detection rate. The all-sky FRB rate appears to have a relatively flat spectrum, which means most FRBs detected by DSA-2000 will be found at the bottom of our band, thanks to the larger field of view at low frequencies. This, combined with the relatively weak impact of integrated bandwidth on sensitivity, allows us to search only the bottom quarter of our band without losing many sources. We therefore plan to search just ~325 MHz of total bandwidth between 700—1100 MHz, selecting only RFI-clean channels. Extrapolating from the detection rate of CHIME/FRB, we expect to detect and localize many thousands of FRBs per year.

The FRB search will operate on ~9 million beams at 0.96 ms temporal resolution and 134 kHz channel width. Searching to a maximum DM of 10,000 pc cm\*\*-3, we need just ~2800 DM trials. We will split data across the three remaining GPUs by DM range: 0—1000 pc cm\*\*-3 will be searched on GPU6, 1000-2000 pc cm\*\*-3 will be searched on GPU7, and >2000 pc cm\*\*-3 will be processed on GPU8.

Searching at high DMs requires a large amount of GPU memory because of the large dispersion delay (DM=5000 pc cm\*\*-3 corresponds to ~20 seconds of delay and 720 GB of data). We can mitigate the memory issue by “scrunching” in time for high DM trials, effectively downsampling to match the timescale of intrachannel dispersion smearing. This will reduce the memory load by a factor of 3-15 for DMs 2000—5000 pc cm\*\*-3. Still, we will need to process chunks of beams in serial in order to fit the dataset in memory.

When an FRB candidate is detected, a trigger will be sent to dump buffered beamformed voltage data to disk. The latency of this system will be shorter than the voltage beamformer, as is the case with DSA-110.

# Data products

**Pulsar search:** Metadata for pulsar candidates. DM, period, width, best bet psrchiv data cube.

**FRB search:** We plan to preserve 64 voltage beams for each candidate in the vicinity of the FRB’s sky position. This will allow improved offline localization precision, coherent dedispersion, and high time/frequency resolution studies of FRB pulse profiles. Additionally, we will preserve a snapshot image cube in Stokes I---in other words, dedispersed intensity dynamic spectra for a subset of formed beams---during at time of the burst. If we save 100,000 of the nearest intensity beams with 500 ms of 2b data, this is 30GB of data per candidate and roughly 1 TB per day, assuming a low false-positive rate.

# Requirements

1. We require 400 Gb ethernet to enable our large data transfer rates
2. Enough CPU RAM for three large buffers: a 30-second intensity data buffer for the FRB beams (~1 TB), a 600-second intensity buffer for pulsar search beams (~120 GB), and a one-minute voltage buffer (65 GB per node) at capture. Total ~1.2 TB of RAM, based on which we conservatively request 2 TB of RAM.
3. ~1,200 GPUs spread out over 150 PSR/FRB nodes with PCIe 5x16
4. After data capture, we have an I/O requirement of ~4.3 GB/s per GPU