SonATA Overview

John Dreher

2005 June 23

The SETI on the Allen Telescope Array (SonATA) project is the successor to the current New Search System (NSS). SonATA is intended for large (10^5 to 10^6 star) targeted searches using the Allen Telescope Array (ATA) beginning in 2007. This memo presents a rough attempt at an overall system description – it is meant to clarify the thinking of the Sonata team and to provoke discussion. Nothing is set in stone yet.

The ATA will provide four independent RF tunings, or "IF channels", that may be placed anywhere within the 0.5 to 11.2 GHz frequency coverage of the ATA. Each tuning in turn provides four independent 104 MHz bandwidth dual-polarization beams that can be directed to any point on the sky (although in most applications they would be directed to directions within the primary beam of the ATA elements). The polarizations are orthogonal linear designated "X" and "Y". The ATA Phased Array Network Data Interface (PANDI) receives the sampled voltages from the phased array beam outputs of the ATA Digital IF Processor. PANDI packetizes the 104.16 MHz bandwidth complex-voltage sample streams from these beams, adding time stamps and other information to the packet headers, and places them on a set of 10 Gigabit Ethernet IP links for use by the phased-array backends, including SonATA, or even by external users on the internet, within the constraints of the internet bandwidth available.

Figure 1 shows the subsystems of SonATA and how they relate to the subsystems of the ATA. The data distribution subsystem, DD, consists of a 10 Gb IP network with suitable managed switches. The signal detection subsystem, DX, is made up of a large number of general-purpose servers that execute the SETI detection algorithms in software and, as economically justified, in additional specialized hardware. The results from the detectors are passed to the Search System Executive subsystem, SSE, which implements the automated observing protocol including classification of detected signals and resolution of interesting candidates, manages the SonATA software and hardware, and provides the interface to the ATA control system. SSE runs as a set of software processes on the same cluster as DX. The hardware support subsystem, HS, provides electrical power, cooling, cable management, and racking for the rest of the system.

The present, preliminary plan for the large targeted search will use four beams from one IF channel (tuning) to search four target stars at the same frequencies simultaneously. By comparing the results for each signal detection from one star against those from the other three, we should be able to eliminate almost all RFI. But with four ~100 MHz search units observing, we will certainly have some noise "hits". A second, smaller set of "follow-up" detectors fed from a second IF channel will reobserve these detections, and, if they were just noise, eliminate them. Confirmation strategies for any survivors will be discussed in a later memo. Eventually it would be desirable to install one of our detectors some other large telescope, such as the GBT, so that we can use our previously successful very-long-baseline interferometry to demonstrate a candidate signal source is (or is not) a point coincident with the target star to very high accuracy. Such coordinated confirmation observations will occur only on those rare occasions when we have a real (*i.e.* non-noise) signal that cannot be established as terrestrial in origin by local ATA observations.

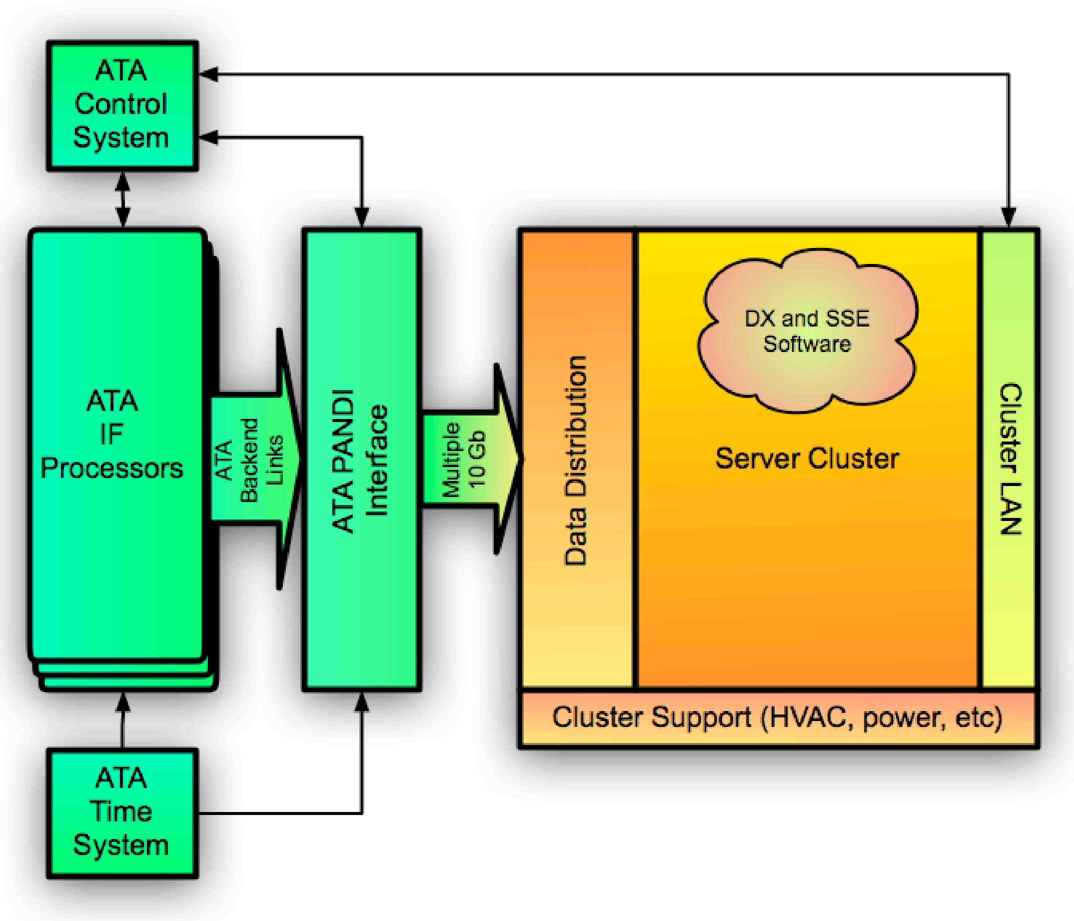

Figure 1. Block Diagram of the SonATA system and its relation to the ATA signal processing.

Data Distribution (DD)

The ATA Digital IF Processor provides a correctly delayed and phased summation of the complex voltages of the ATA elements as a stream of complex voltage samples on IP packets sent over a 10 Gb IP network. At present it appears that 10G Ethernet will be the most practical implementation, but we will evaluate alternatives when the time comes to procure the hardware. Each polarization of each beam will produce 3.33 Gb/s of samples (16 bits Re, 16 bits Im). The definition of the packet format is still in progress; our preliminary design has 1439 byte packets (including 38 bytes of headers) at a rate of 297,600 packets/sec for a total of 3.42 Gb/s.

In addition to accepting the time series data from the ATA over 16 IP links (8 for each of the two tunings used for SETI), the DD needs to transport data between the first stage of the DX processing (a DFB as described in the next section) to numerous other DX processors, while merging the two polarizations for each beam. This task can be most flexibly achieved by using 10G networks and managed switches for all interconnects; however, once the first stage of DX processing has subdivided the bandwidth per beam it would be possible for most of the DX servers to used 1G links, which might reduce costs.

The results of the DX processing will also be sent to the SSE by DD, however the volume of data is much less and can be handled easily by a single 1G Ethernet, although for future flexibility 10G would be preferred.

Detectors (DX)

Since the bandwidth of an ATA beam is greater than can be fully processed on a low-cost server, the first step for DX will be to break the 104.16 MHz bandwidth down into chunks of order 0.1 to 1 MHz, using a Digital Filter Bank (DFB, also known as a polyphase filter bank). For the purposes of this memo I will designate the complex voltage samples representing these narrower chunks as *filaments*.[†] The ATA provides about 50 dB of antialiasing in the digitizers and about 40 dB of image suppression in the quadrature downconverters (after digital tweaking), so it seems reasonable that the DFB be specified to have out of band rejection of 50 dB or better; this will require a preconvolver of length 8 or greater. Note that the signal bandwidth of each filament will need to be oversampled by ~20% in order to prevent aliasing. This first stage of DX should also allow for selection of a subset of the filaments, so that we are not forced to send the aliased band edges (only ~90% of the Nyquist bandwidth of the ATA output is alias free) and the DC band, and for choice of sample requantization. Each DFB will require about 12 Gop/s sustained per beam per polarization, and 16 DFB processes will be required for 2 tunings, 4 beams, and 2 polarizations. Current (2005) servers typically achieve ~3 Gflops per CPU; we anticipate that moderately priced 2-processor servers that can achieve 12 Gflop/s performance will be available in ~2007. (As alternatives, an FPGA-based board with IP interfaces such as the BEE board that as been developed at UCB, or a higher-performance server could be substituted.)

The rest of DX subsystem does the actual work of statistical detection. In the earliest incarnation of SonATA, we expect that the tasks of each module of the detectors will be similar to those of the NSS PDM (to minimize NRE) with the following differences: the voltage samples arrive over the IP network and are placed into a circular buffer in memory. The “brickwall filter” task of the current DDC and all the spectral analysis and formatting tasks of the “Blue Wave” DSP board migrate to software implementations running on the host servers. We also need to place a high priority on developing ~100-second coherent non-drifting CW and pulse detectors to comply with the recommendations contained in SETI 2020. An important open question is whether “non-drifting” must include some search range in drift and/or curvature. A slightly lower priority should be given to the current NSS detectors, including multi-resolution pulse detection. Development of CW detectors for higher drift rates is worthy, but of lower

---

[†] The more ‘natural’ term might be *thread*, but that term, as well as *stream* are already in widespread use in computing. Similarly, the detectors already use the term *slice*. A more elegant and/or shorter designation would be nice: ideas welcomed

priority yet. Minimal computing requirements will be of order 1 to 3 Gflop/s per MHz of dual-polarization bandwidth processed (0.5 to 1.5 Tflop/s aggregate). In order to maximize throughput a large amount (4-12 GByte/MHz) of RAM would be preferred, but it is not clear whether we can afford that; alternatives would be 1) the use of disk as in our current implementation (with possible adverse effects on the bandwidth per server) or 2) restricting the maximum observing length until RAM is cheaper. In the medium term we plan to reduce our observation pipeline length from two stages (data collection/DFB, signal detection) to one stage, which will simplify the observing logic and improve efficiency and response time for follow-up.

The following table estimates the minimum hardware requirements for DD and DX, assuming that 25% of the total bandwidth of the second tuning needs to be analyzed for follow-up.

Stage	tunings	beams	pols	per MHz			aggregate			
				I/O Mb/s	CPU Mflop/s	RAM GBytes	BW MHz	I/O Gb/s	CPU Tflop/s	RAM TBytes
ATA output	2	4	2	32			1667	53		
DX DFB	2	4	2	64	115		1800	115	0.2	
DX detection	1.25	4	2	32	1000	4.0	937	30	0.9	3.7
Total								199	1.1	3.7

The algorithms used for SETI detection are well suited for parallel processing, either using SIMD or thread-level parallelism. The current NSS code base makes use of both features. The current industry trend towards increased CPU parallelism will be easy for us to exploit. We anticipate that our algorithms will continue to evolve – the above estimates apply only to our plans for the immediate future. SETI signal detection can continue to benefit from increased I/O, memory, and computation power as more of these become available per dollar.

The overwhelming driver is cost reduction. The hardware should be viewed as an expendable with a lifetime of 3-5 years. In order to keep NRE under control, we need to consider the relative costs of optimizing code versus just using more hardware. In order to maximize the possibilities for donations of the server "farm" that will form the core of DX, we should also strive to use generic, straight-from-the-box hardware from large vendors rather than the cheapest hardware. For the same reason we should, as we have done in the NSS, stick with common operating systems and vendor supplied processor-dependent libraries wherever possible. The best payback from optimization will be to focus on "generic" optimizations such as multi-threading and memory localization.

SSE

The interfaces to the units of the DX subsystem will be via another (possibly slower) Ethernet link from each host. Management of the DX units should be very similar to the tasks of the current SSE software system, extended for the new detector classes. The interface between the SonATA system and the ATA control system is currently being developed (and tested) in conjunction with the Science team and the ATA software team so that we will have suitable observing protocols both for the situation in which SonATA is selecting targets as the primary observer and for the case in which SonATA is running in a complementary mode with an astronomy program (probably a survey) as the primary observer. At first the latter case will be facilitated by mutual design of the two observing programs by the relevant science teams; in future we may evolve the ATA control software to act as a broker in a "market oriented" resource allocation system. Another long term goal will be to incorporate the (extensive) control logic of the SSE into a framework that will allow algorithm developers and SETI scientists to build and use new types of SETI detectors with relative ease.

The user interface to the SonATA control system will need to be used readily over the network, since we do not plan to have a permanent observing staff at Hat Creek. Indeed, our goal is to have the observing proceed with no continuous human supervision in order to reduce the operating cost of the SETI effort. The ATA control system will provide a set of suitable Java-based applications and applets that can be used for secure control of the array over the Internet – these will probably be the basis of the user interfaces to SonATA.

Hardware Support (HS)

The SonATA cluster will probably need of order 100 servers spread over several racks. The hardware support system needs to provide, monitor, and control the distribution of electricity and cooling to these servers. A rack that dissipates 10 kW will need to exhaust about 1000 cfm of hot air (to an exhaust plenum) and pull in the same volume of cold air from the equipment room. The SonATA system will be one of many users of the digital shielded room at the ATA; its overall power and thermal requirements have already been incorporated in the ATA requirements. Although simple compared to modern web-server "farms" we will still need to pay care to managing the network cabling so that maintenance is simplified. We expect that the vendor for the server cluster can provide a good cluster operations monitoring software system, as is becoming standard practice; this too will reduce the amount of labor needed to keep the cluster running. Our "24/7" uptime requirements are minimal compared to most commercial applications – but our maintenance budget is also likely to be minimal, so we must exercise due care. Probably the best way to reduce maintenance is to ensure that the equipment is kept quite cool and that the AC power system protects the equipment from spikes and/or brownouts. UPS protection may not be necessary for most of the servers, since when the mains power is down the ATA cannot observe in any case. Physical security and fire suppression will be the responsibility of the ATA.

Open SonATA

One of the goals of the SonATA effort is to open up the development and use of our software and hardware for SETI work by a broader community and to benefit from the ideas and work of that community. Advice from people familiar with open source development suggests that the “mission oriented” tasks of getting a large targeted search underway will in all likelihood need to be performed by our own, in-house staff. However, we can benefit by making our efforts available as open source software; in this manner 1) we can expect external review and advice on our efforts from the technical community, 2) the assumptions in and accuracy of our work will be exposed to the scientific community in detail, and 3) we will perforce bring our documentation up to a higher level than otherwise would be likely. In some cases we may even be able to obtain specific optimizations or algorithmic improvements that we lacked the expertise or inspiration to produce on our own. Open source *development* of SETI software is likely to occur independent of (but benefiting from) our efforts and explore avenues outside our science team’s present interests. Since the ATA is receiving operating funds from the NSF, a portion of the observing time will be available for outside observers; this time could include such independent SETI research, possibly using the processing hardware in new ways. Also, the outputs of the ATA beamformers could be made available to external researchers for various purposes at any time, subject to the internet bandwidth available for transporting these data. We will seek to have the ATA participate in Internet2 so that this bandwidth can be maximized. As internet bandwidth becomes more available, the ATA could become an “open telescope” producing useful data for clever projects that can exploit serendipitous co-observing. The recent decision to provide the outputs of all the beamformers in the form of industry-standard IP streams was primarily driven by these “open” considerations.