

Requirements for the STAR Trigger Level-0 Data-Processing System Upgrade
H.J. Crawford, J.M. Engelage, E.G. Judd, C.W. Perkins
August 15, 2012

Introduction

The STAR Trigger System performs a quick analysis of data from the fast trigger detectors in order to decide if the slower readout of data from the rest of STAR should be initiated. The current STAR Trigger Level-0 data processing system is based on the Data Storage and Manipulation boards (DSM). The DSM boards run synchronously with the RHIC clock, receiving new data from trigger detectors every tick of the clock, processing it and passing on the results. Each board can only accept a limited number of channels of input so many DSM boards are linked together in a tree structure. This tree funnels the results of the data analysis to the Trigger Control Unit (TCU), which actually makes the trigger decision.

The DSM system is currently the oldest part of the STAR Trigger system and a number of problems have appeared:

- The electronics is so old that replacement parts are no longer available
- The DSM system has grown significantly over the years, including crates of electronics that are not on the platform. This has led to a huge mass of cables that is hard to maintain and debug.
- The limited number of output bits from each DSM board, and the fixed structure of the DSM tree place serious limits on the type of data analysis that can be performed. For example, the ideal algorithm for the Electromagnetic Calorimeter (EMC) is one that would sum the ADC values from the calorimeter's towers in groups to make all possible jet patches, and then compare all of those ADC sums to various thresholds. Unfortunately each existing DSM board can only accept a fraction of the number of towers needed to make even one useful jet patch. The current algorithms therefore calculate partial sums and pass them up the tree to be combined later. There are not enough bits available to pass along all of the necessary partial sums so only a fixed subset of partially-overlapping jet patches are made. This results in a loss of trigger efficiency, and an uneven acceptance that makes subsequent jet analyses more complicated. Another algorithm that is excluded by the current DSM tree is one in which the multiplicity in each Time-of-Flight tray (TOF) is compared to the ADC sum from the EMC towers with which it overlaps. This comparison would make it possible to confirm or exclude the presence of muons, and would therefore make muon triggers much more efficient. Currently the EMC and TOF data are not brought together in the DSM tree until the end, by which time only summary data is available. Fixing this would involve re-cabling the DSM tree, and sacrificing existing, important connections.
- The existing DSM boards are read out in series within each crate over the VME backplane. This relatively slow readout mechanism is now limiting the overall STAR trigger rate.
- After a trigger decision has been made the information has to be distributed to the DSM crates to initiate readout. The current mechanism for distributing the information is not part of the synchronous system so it has a variable latency.

It is now time to upgrade the Level-0 data processing system and replace the existing hardware to solve all of these problems. Figure 1 shows all the inputs and outputs to the current Level-0 data processing system.

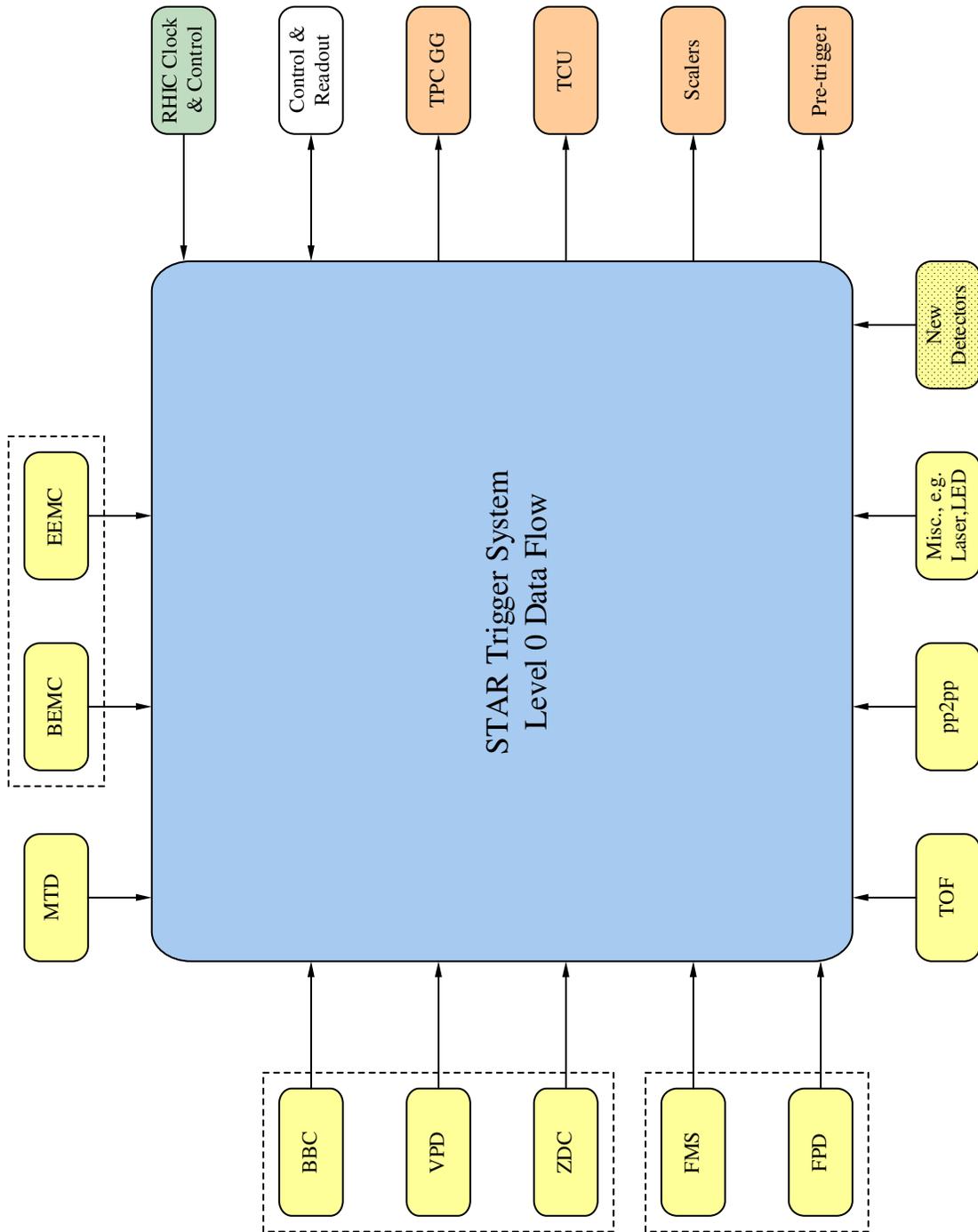


Figure 1: Inputs and Outputs of the Level-0 Data Processing System

Figure 2 shows an example of how much data comes into the system from each detector, and when. The length of each arrow shows how long it takes from the interaction to when that data set reaches the DSM tree. The width of each arrow indicates how much data is available. Based on this information we can define the requirements for the upgraded data processing system.

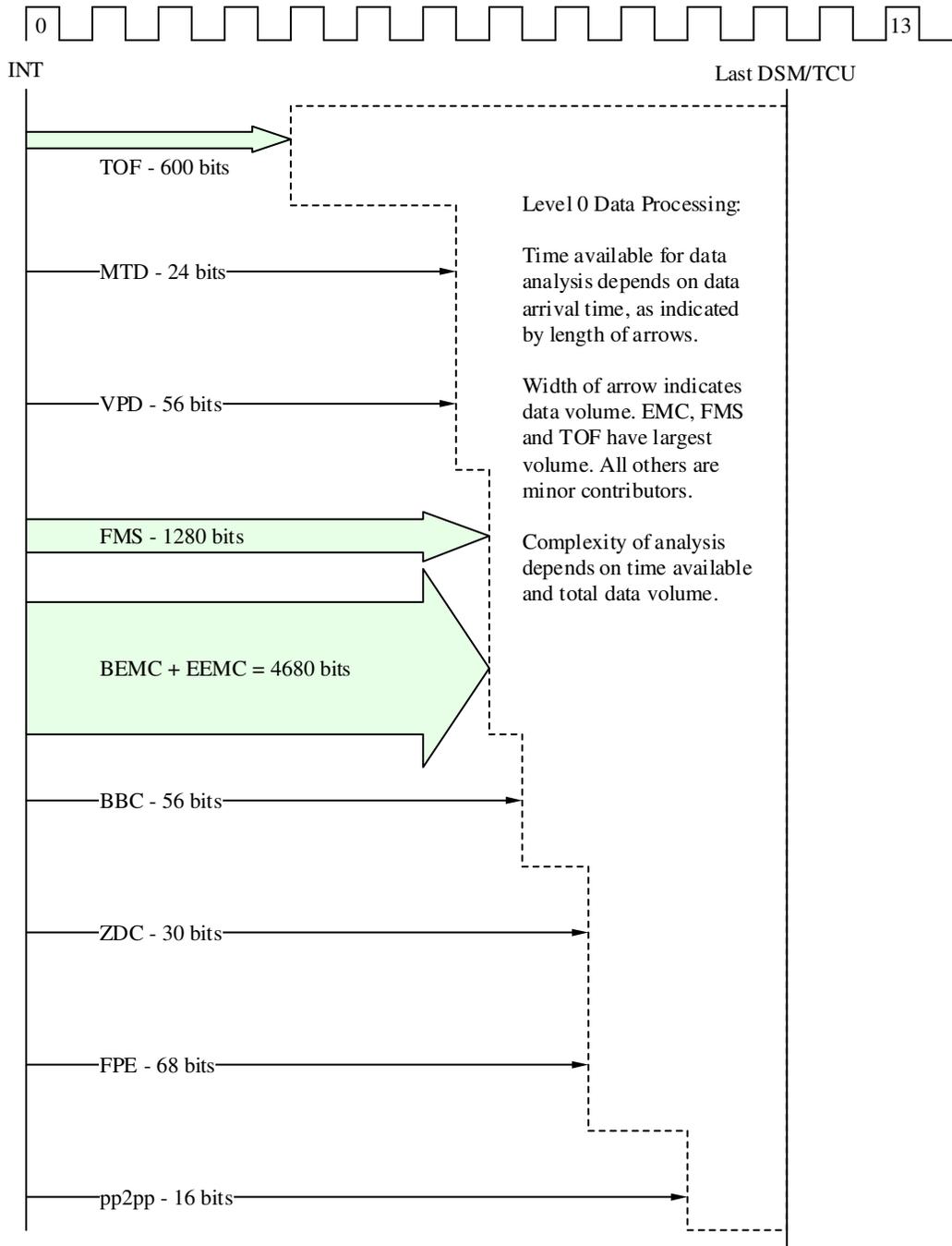


Figure 2: Arrival Time and Volume of Trigger Detector Data entering the existing Level-0 Data Processing System. This example was taken from 2009 Tier1 file settings.

Level-0 Data Processing System Requirements

Synchronous Data Flow

- 1) Requirement: **Current Data**. The new system must receive all the detector data that the existing system receives
Justification: The new system has to replace the old one and fulfill all its existing functionality.
- 2) Requirement: **New Data**. The new system must be easily expandable to receive data from new detectors
Justification: STAR is already planning upgrades, and some new detectors will provide data to the trigger.
- 3) Requirement: **Data Form Factors**. The new system must be able to receive data in multiple form factors including, but not limited to, PECL parallel bit streams on copper cables.
Justification: The existing detectors all send their data to the trigger system as parallel PECL bits on copper twist-and-flat cable. However the new detectors may use a different scheme.
- 4) Requirement: **Synchronous Data Input**. The new system must receive new data every tick of the RHIC clock
Justification: RHIC is a synchronous system. There is a bunch crossing every tick of the RHIC clock, and the trigger system must examine them all
- 5) Requirement: **Configurable Data Format**. The connections that receive or transmit data must connect directly to a configurable processor, i.e. there must be no hardware that is built on assumptions about the size, structure or timing of the incoming data.
Justification: This will make it possible to upgrade the data format in the future without having to make hardware changes. Failure to meet to this requirement led to a problem in the existing system. Data is currently latched onto each DSM board on the rising edge of the RHIC clock. The source of this RHIC clock is the primary clock distribution network on the DSM board, i.e. it does not come from an FPGA. As a result it is impossible to change the frequency of the input latch without affecting the rest of the board. This meant it was impossible to accommodate a user who wanted to multiplex extra data onto one channel by using a faster clock.
- 6) Requirement: **Early Data Delay**. The Level-0 data processing system must be capable of delaying early data to wait for the later data.
Justification: Data from different detector systems arrive at the Level-0 system at different times, depending on the particle flight, detector response function and the electronics chain. For any RHIC crossing the final trigger decision is based on combining the data from that crossing generated by many different detectors.

- 7) Requirement: **TCU Data Summary**. The new system must provide a 128-bit summary of the detector data for each RHIC crossing to the Trigger Control Unit (TCU). Those bits must be driven in parallel with each other, not as a serial bitstream. Justification: The TCU will use that summary data to make the trigger decision for each RHIC crossing. The TCU expects to receive 128 bits in parallel every tick of the RHIC clock.
- 8) Requirement: **Fixed Transit Time**. The new system must have a known, static, total transit time from the beam interaction to when the data is sent to the TCU. Justification: Some downstream detector systems are pipelined and when a trigger is issued they need to know which element of their pipeline should be read-out.
- 9) Requirement: **Short Transit Time**. The total transit time through the new system from the beam interaction to when the data is sent to the TCU must be similar ($\sim 1.2\mu\text{s}$) to the current value. Justification: Some existing downstream detector systems are pipelined and the depths of the pipes were determined by the expected latency between the interaction and the receipt of the trigger. If the transit time becomes too long then the data from those detectors will have left the pipe before the trigger is received. Another detector, the TPC, doesn't start digitizing data until it receives the trigger. Data that passes through the endcap before digitization starts is lost.
- 10) Requirement: **Scaler Data**. The new system must provide data from each RHIC crossing to the Scaler System. This data must be in the form of parallel bits, not a serial bitstream. Justification: The scalers are used to monitor the performance of the triggers, the behavior of the accelerator and the luminosity of the beam. The current Scaler System expects to receive bits in parallel every tick of the RHIC clock.
- 11) Requirement: **Upgradable Interfaces**. The interfaces between the new system, the TCU and the Scaler System must be implemented in such a way that if either the TCU or the Scaler System is upgraded to receive data in a different form (e.g. serial bitstream instead of parallel bits, different form factor, etc...) then the Level-0 system can accommodate that with only minor hardware changes. For example, the relevant output bits could be routed from the processor to a daughter card. The first version of the daughter card would drive those bits, in parallel, as 5V PECL signals since that is what the current TCU and Scaler Systems require. When either system is upgraded all that would be necessary is to upgrade the daughter card. Justification: The TCU and the Scaler System will both need to be upgraded at some point. Those upgrades should not result in large changes to the Level-0 data processing system.
- 12) Requirement: **Lemo Outputs**. The new system must provide programmable lemo outputs. Justification: Some downstream detectors require a fast pre-trigger that is issued before the final trigger decision.

13) Requirement: **Data Analysis**. The new Level-0 data processing system must analyze the data received from all the detectors in order to make the summary that is sent to the TCU. The new processors must be able to perform all calculations that are done now. In addition each processor must be capable of dealing with at least 390 channels instead of the current maximum of 20 channels.

Justification: One of the problems with the existing Level-0 data processing system is that the individual processors can only accept a small fraction of the data channels. This creates various problems as described in the Introduction. For example when analyzing the EMC data the ideal algorithm would be one that combined channels in many different combinations to make all possible jet patches. The limitations of the current system mean that just a subset of partially-overlapping patches can be constructed. The EMC currently sends the largest number of channels, 390, to the existing Level-0 data processing system. If the new processors can accept at least 390 channels then one processor could accept all the EMC data and therefore could implement the ideal algorithm.

14) Requirement: **Data Routing**. The new system must be capable of making at least four copies of each incoming data set and routing them to different processors.

Justification: In the current system the processors are linked together in a fixed tree structure that is defined by the cabling between DSM boards. Each DSM board has a very limited set of output bits. Since each piece of the tree needs to pass data downstream, and also to the Scaler System, there is essentially no ability to send multiple copies of one data set to different places for analysis with other sets of detector data. For example, it is not possible to send one copy of the BEMC data to be combined with the EEMC data while another copy goes to a different processor to be combined with the TOF data. This has limited the number and type of triggers that can currently be implemented. The new system should at least reduce this limitation. Making four copies of each data set would enable one set to go to two processors and the Scaler System, and still leave one spare copy for future expansion.

15) Requirement: **Test Facility**. The new system must have built-in testing facilities that make it possible to test each piece in stand-alone mode and combined into a system but without necessarily requiring an external detector to provide data.

Justification: The system must be regularly checked to determine if it is working properly. It is not always possible or desirable to put together a full system, including an external detector, to do these tests.

Synchronous Control

16) Requirement: **RCC2 Interface**. The new system must receive the 9.38 MHz RHIC clock from the RCC2 system.

Justification: RHIC is a synchronous system; the RHIC clock defines when bunches will cross in the interaction region. The RCC2 system is used to distribute it around STAR.

- 17) Requirement: **RHIC Clock Frequency**. The new system must be capable of operating at RHIC clock frequency that is three times the current value, i.e. 28.14 MHz, with at most minor hardware and firmware alterations.
Justification: As part of its plan to increase luminosity, RHIC is planning to increase the RHIC clock frequency (most likely by a factor of 3) in the next 5-10 years.
- 18) Requirement: **Synchronous Control**. The new system must also receive the synchronous STAR control signals from the RCC2 system and use them, with the clock, to control storing detector data and sending data to the TCU and Scaler System.
Justification: The synchronous STAR control signals are used to define when STAR is taking data. These signals are also used for test and monitoring purposes.

Data Storage and Readout

- 19) Requirement: **Readout Rate**. The minimum steady-state rate at which events can be readout must be at least 20 kHz. Those parts of the readout process that can be made to go faster than this should in fact operate as fast as possible.
Justification: STAR is currently in the process of upgrading various systems with the aim of increasing the overall event rate to 10-20 kHz. So long as the Level-0 Data Processing System can be readout at 20 kHz it will avoid becoming the bottleneck that limits the overall STAR event rate. The readout will involve several stages including distributing the command to initiate readout, extracting the relevant data from storage, collecting the data in one place and sending it for further analysis. Some of these stages will run much faster than 20 kHz but one of them will be the slowest and it is that piece that will limit the overall rate. If the stages that can go faster than 20 kHz are implemented to run as fast as possible then that will help prevent those stages from automatically becoming the bottleneck when the slowest stage is finally sped up.
- 20) Requirement: **Raw Data Storage**. The new system must store all detector data for a time period long enough to cover the maximum time from data storage to readout. The amount of memory necessary for this will vary between different parts of the system depending on how much data is stored there and how much time the data spends there. The early parts of the new system must store all the data from every tick of the RHIC clock while the trigger decision is made. Later parts of the system only need to store the data from those ticks of the clock where a trigger was issued. Currently this is approximately 5.5 kB per tick. The time that the data spends in the slowest part of the system must be no more than 205ms. The detailed requirement for each part of the new system will be discussed in the board-specific sections later on in this document.
Justification: The decision to issue a trigger, and therefore to read out the data, is made downstream of the Level-0 data processing system by the TCU. There is therefore a delay between when data is received by the Level-0 data processing system and when it is read out. The overall readout rate requirement is a steady-state minimum of 20 kHz. In order to achieve this it must take a maximum of 50 μ s to fully

read out each event. Resource management within the trigger system is token-based. A trigger is only issued if tokens are available. There are a maximum of 4095 tokens so it is possible to have a burst of triggers issued in 4095 consecutive RHIC clock ticks. If it takes 50 μ s to read out each event then it would take \sim 205 ms to read out the longest burst of events. The slowest piece of the system must therefore have enough memory to store data for 205ms, however faster pieces of the system will need much less.

- 21) Requirement: **Processed Data Storage**. The new system must also store the final results of the data analysis and the results from major intermediate steps for the same time period as it stores the incoming detector data.
Justification: This data is useful for debugging and monitoring the behavior of the algorithms.
- 22) Requirement: **Analysis Results Readout**. It must be possible to read out both detector data and analysis results during normal data taking
Justification: In the current system it is only possible to read out the data received by each processor, not the results of its analysis. In theory every processor passes its results to another, where they are stored and read out. However, sometimes this has not happened and the lack of ability to read out the analysis results directly from the processor doing the analysis has caused problems. This should be fixed in the new system.
- 23) Requirement: **Non-interrupting Readout**. Readout must happen without interrupting the storage of new data
Justification: Any bunch crossing can result in a trigger, so the data from all bunch crossings must be saved.
- 24) Requirement: **Readout Selection**. It must be possible for the user to specify which pieces of the Level-0 Data Processing System must be readout for each trigger type.
Justification: Some physics analyses only require the data from a subset of trigger detectors. Event rates for those triggers can be increased if only the necessary data is readout.
- 25) Requirement: **Readout Network Connection**. Each piece of the new system that is to be readout must have its own connection to the network that distributes the command to initiate readout.
Justification: Speeding up the readout is one of the main reasons to upgrade the current system. All other things being equal, the fastest readout speeds are obtained if the command to initiate readout, and the subsequent readout itself, are all done in parallel.
- 26) Requirement: **Configurable Readout Format**. The connections that receive the readout command, and that transmit the data being readout, must connect directly to a configurable processor, i.e. there must be no hardware that is built on assumptions about the size or structure of the incoming or outgoing packets.

Justification: This will make it possible to upgrade the packets in the future without having to make hardware changes. Failure to meet to this requirement has resulted in an unpleasant feature in the readout of the existing system. Data is currently readout over VME using a 64-bit block transfer. The data storage and transfer is currently implemented in discrete ICs that are controlled by an FPGA, but the data itself does not actually pass through the FPGA. Unfortunately the circuit introduces a twist in the ordering of data blocks that has to be undone later. If the data was transferred through the FPGA, instead of bypassing it, this problem could have been fixed.

Asynchronous Control and Monitoring

- 27) Requirement: **Slow Communications**. Every processor in the new system must have a communications path that allows for remote configuration, control and monitoring. At a minimum this path must include an Ethernet connection and a process capable of opening an Ethernet socket and responding to incoming communications.
Justification: During data taking STAR is not physically accessible so all communication has to be done remotely. Run Control currently communicates with all of STAR's subsystems over Ethernet.
- 28) Requirement: **Communication Independence**. Any clocks used in the slow communications process must come from a local source that is independent of the external RHIC clock (e.g. local oscillator).
Justification: In the current system the processing boards use the RHIC clock to drive their communications logic. This means the boards cannot communicate at all if the RHIC clock drops out, which has caused problems.
- 29) Requirement: **Remote Reset**. Every processor in the new system must have a remotely accessible reset facility that allows the user to reset all logic to the default power-on state without actually cycling power (c.f. VME sysreset).
Justification: Sometimes an error occurs that results in a processor freezing up. Since STAR is not physically accessible during data taking initiating the reset must be done remotely.
- 30) Requirement: **Field Programmable Logic**. All logic must be field-programmable.
Justification: The algorithms that STAR uses to analyze the detector data varies from one type of running to the next so it must be possible to download new algorithms. Also, the structure, size and bit rate of data and communications packets may need to change over time as various other systems are upgraded.
- 31) Requirement: **Configuration Time**. The total time taken to configure the Level-0 Data Processing System must be no more than 15 seconds. This includes the time taken to set all registers and zero all the memory. It does not include the time to download new algorithm logic to data analysis processors because that happens only rarely (beginning of a RHIC run, change of beam species/energy).

Justification: This is a short enough period of time that it will not inconvenience the STAR users. It is also short enough to prevent the trigger system from becoming the slowest system to configure.

New Level-0 Data Processing System Definition

When the current DSM boards were designed the analysis processor that was chosen was the Lucent ORCA OR2C40A FPGA. This was one of the largest FPGAs available at the time. It had 304 I/O pins and could perform useful integer calculations of the type needed by STAR at speeds of a few tens of MHz. However, those 304 pins only supported single-ended signaling standards, not any of the differential standards. This limited the rate at which serial bitstreams could be brought into the FPGA. The rate was low enough that it was not useful for STAR, so instead all the detector information was brought into the FPGA in parallel. Since each FPGA needed to receive input bits, transmit output bits and have enough I/O pins left over available for communication this led to a DSM design with 128 input bits and 32 output bits. Since the STAR trigger has many thousands of input bits the end result was the DSM tree.

Over the last 15 years, since the original DSM boards were designed, FPGAs have evolved considerably. Very large chips (> 1000 I/O pins) are now available with support for differential signaling standards, built-in SERDES capability and the ability to run at much higher clock speeds. At the same time standards and hardware for transmitting data at very high speeds have greatly improved, and very large memory chips have become available and cheap. As a result it is now practical to serialize the trigger detector data and bring many hundreds of channels into one module where that data can be stored, processed and rapidly read out. This scheme will solve all the problems with the existing Level 0 Data Processing System.

The new scheme is shown in Figures 3 and 4. Figure 3 shows a block diagram describing the data flow through the new system. The existing interfaces to the detectors will still provide the input data: DSMI, TDSMI and Tof-DSMI. At the other end the TCU will also initially be the existing model. In between will be two new functional modules: an array of serializers (SER) and a set of data storage and manipulation boards (DSM2). The SER board will receive the existing parallel data from the trigger detectors, duplicate it, serialize all the copies and transmit them onwards. The DSM2 boards will then de-serialize the data, store it, apply the STAR trigger algorithms and send the results to the TCU.

A path for further upgrades is also envisioned. The TCU will probably be upgraded first and may be simply replaced with a modified DSM2 board. As the detectors themselves are upgraded the SER boards may be removed, if the upgraded detector system can drive DSM2 boards directly, or reworked to receive the new input data format. Finally, if the need arises for a more general, configurable, connection between the SER boards and the DSM2 boards then the Trigger Routing Matrix (TRM) could be inserted between them. In its most general form the TRM would receive all the input data from the SER boards and transmit data to every input channel of every DSM2 board. It could copy and route any serialized incoming data stream to any DSM2 board. A partial TRM may actually be implemented by connecting several DSM2 boards.

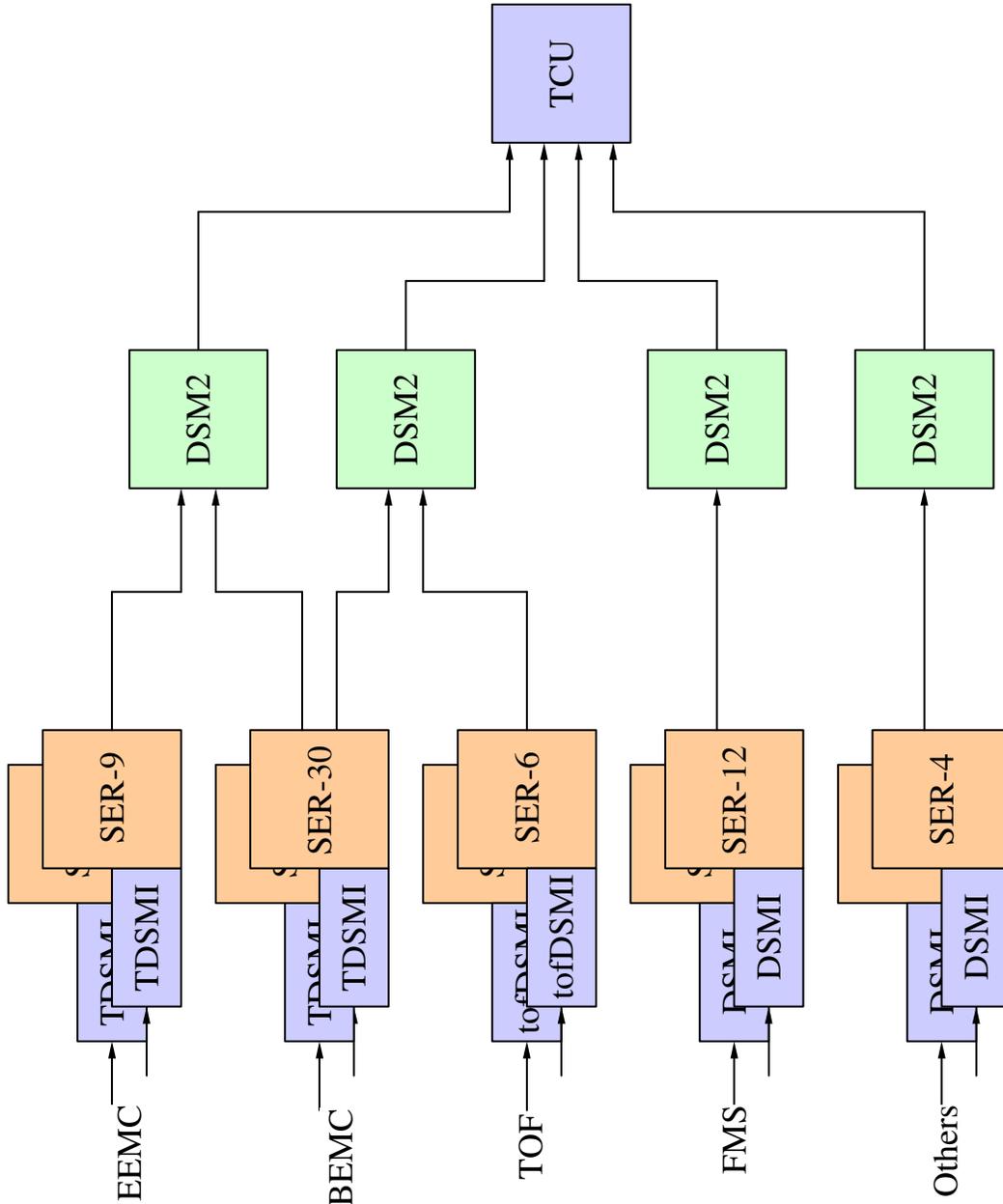


Figure 3: Block diagram showing the data flow through the upgraded STAR Trigger Level-0 Data Processing System.

Figure 4 shows a functional block diagram of the upgraded DSM2 readout scheme. It is a fully parallel system. When the TCU issues a trigger that information is broadcast, in parallel, to all the DSM2 boards over a Trigger Distribution Network (TDN). The DSM2 boards extract the data associated with that trigger from their memories and send it, in parallel, to the Trigger Data Collector (TDC). The TDC gathers the data and dumps it into the memory of the Level 2 Trigger computer for further processing. Currently the TDN function is split between two pieces of hardware: a VME CPU that reads out the TCU and passes the trigger information to an STP

Concentrator, and the STP Concentrator itself which broadcasts the trigger information to the VME CPUS in the DSM crates. The STP Concentrator also receives the existing DSM data from those VME CPUS and acts as the Trigger Data Collector. Ultimately when all the upgrades are complete the DSM2 boards and the TCU will each have a direct connection to the STP Concentrator so the Concentrator will fulfill all the functions of the TDN and TDC. The VME CPUS will no longer be necessary.

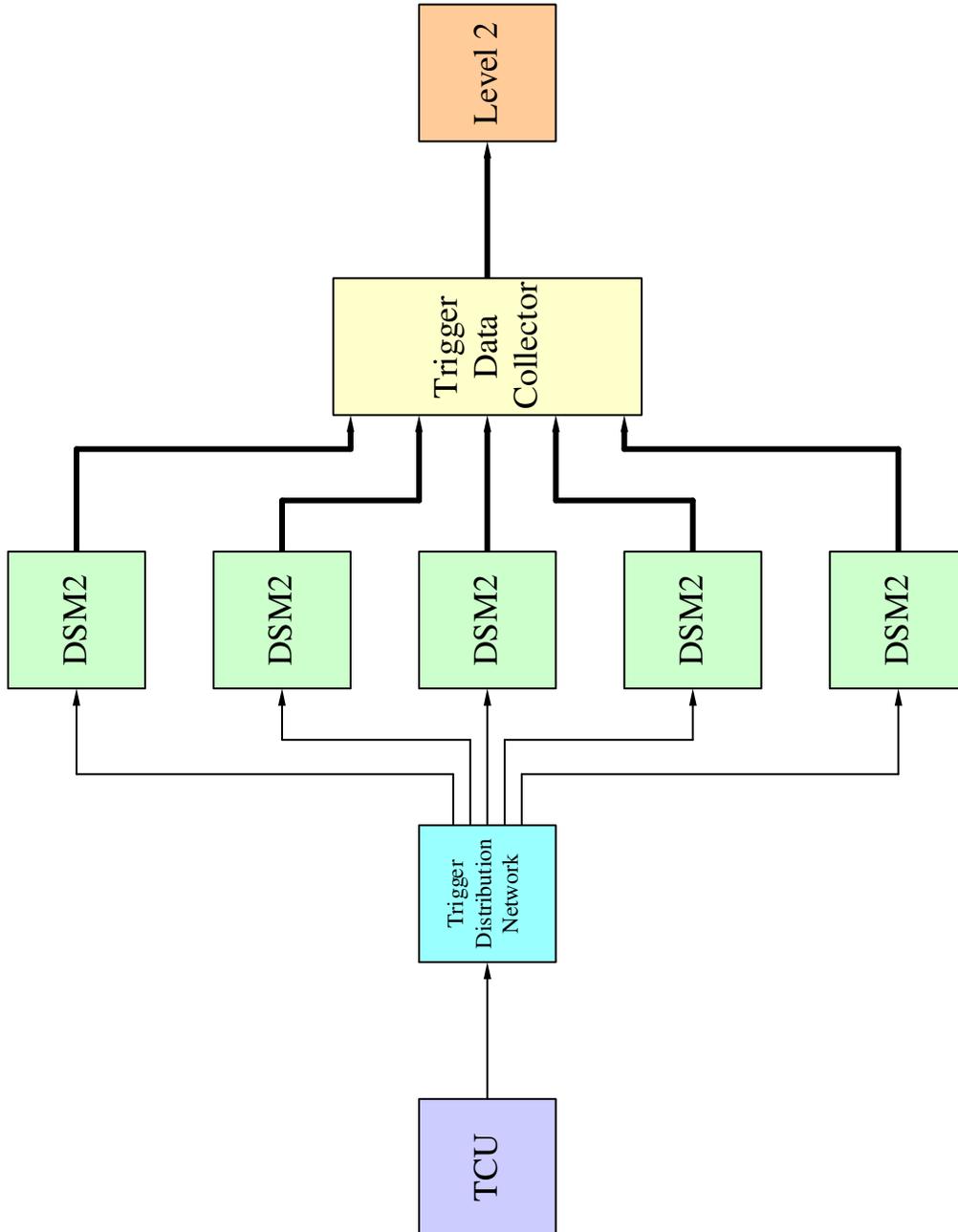


Figure 4: Block diagram showing the fully parallel readout path for the upgraded STAR Trigger Level-0 Data Processing System

Serialization Board – SER (First Version which connects to existing DSMI/TDSMI/TofDSMI boards)

- 1) Requirement: **Board Form Factor**. The first version of the Serialization Board needs to have the 9U VME Form Factor.
Justification: These boards need to plug into the existing VME crates and connect to the existing DSMI/TDSMI/TofDSMI boards.
- 2) Requirement: **Data Form Factor**. The SER must receive 5V TTL bits through the P3 backplane
Justification: This is how the existing DSMI/TDSMI/TofDSMI boards present their data.
- 3) Requirement: **Clock/Control Form Factor**. The SER must receive the RHIC clock and STAR's synchronous control signals through P2 as 5V PECL
Justification: This is how the existing DSMI/TDSMI/TofDSMI boards present their clock and control signals.
- 4) Requirement: **Output Form Factor**. The SER must transmit serialized output data using a signaling standard TBD on a form factor TBD. The signaling standard and form factor must be one that is suitable for transmitting signals over a distance of at least 300 feet.
Justification: 300 feet is the maximum distance from the STAR electronics platform to anywhere in the STAR hall.
NOTE: Need to choose a serial form factor.
- 5) Requirement: **Data Latch and Fanout**. The SER must latch incoming data once per tick of RHIC clock and then make four copies of the data.
Justification: Multiple copies of each data set are sometimes needed: two copies to go to different DSM2 boards, one to go to the Scaler System and a spare.
- 6) Requirement: **Data Delay**. The SER must delay each of the four data streams by a user-specified number of RHIC clock ticks. The maximum delay is 10 RHIC clock ticks.
Justification: It was shown in Figure 2 that data from different detectors arrives at the Level-0 Data Processing System at different times. Data from early detectors must be delayed to wait for data from later detectors. Depending on the circumstances it may be preferable to delay the data in the SER, or the DSM2. The SER therefore needs to have the delay capability. Setting the delay on each of the four channels individually will allow the user the maximum flexibility in the timing. In the existing DSM system no data is delayed by more than a few RHIC clock ticks, so a maximum of 10 ticks will cover all existing cases and allow for the addition of significantly faster detectors in the future.
- 7) Requirement: **Data Serialization**. The SER must serialize each of the fanned out, delayed data streams and send them out.
Justification: The primary function of the SER board is to serialize the incoming parallel data stream to reduce the number of bits needed to get the data into one processor.
- 8) Requirement: **RHIC Clock Frequency**. The SER must be capable of operating with a RHIC clock frequency that is 3 times the current value with minimal hardware and firmware changes.
Justification: As part of its plan to increase luminosity, RHIC is planning to increase the RHIC clock frequency (most likely by a factor of 3) in the next 5-10 years.
- 9) Requirement: **Field-Programmable Logic**. The input latch, data copy, data delay and SERDES must be in reconfigurable logic.

Justification: The precise details of the SER logic and input/output timing may need to be customized for some detectors, and will probably need to be modified over time.

- 10) Requirement: **Data Storage**. The SER needs to have enough memory to store the parallel data as it comes in off the cable for the same length of time as it is stored on the DSM2 boards. This should be ~ 1.4 MB.

Justification: During normal data taking it will probably not be necessary to read out data from the SER boards, since they should just pass the data through to the DSM2 boards. However, for debugging purposes it may sometimes be necessary to read out the raw data coming from the detectors, and in that case the data has to be stored before a decision is made to read it out. Each SER board will replace an existing DSM board, and will therefore receive 16 bytes of data. Data is stored on each DSM2 board for ~ 82k RHIC clock ticks (see DSM2 Fast Readout requirement) so the SER also need to store data for that long. Storing 16 bytes/tick for 82k ticks requires ~ 1.4 MB of memory.

- 11) Requirement: **Test Mode**. The SER board should be configurable so that data can be stored in the memory as it comes off the input cables, or can be read from the memory and sent through the serialization logic and on downstream.

Justification: It is necessary to periodically test the functionality of each SER board, and the Level-0 data processing system as a whole, without the presence of an external detector. This feature would allow simulated detector data to flow through the system as part of the testing procedure.

- 12) Requirement: **Memory Control**. During data taking the SER must use the STAR synchronous run/stop signal to control when access to the memory (reading or writing depending on the mode) should begin. It must also respond to the STAR synchronous “latch address” command and have the potential to respond to what is currently the spare command.

Justification: Every board in the Trigger System has an address counter that increments by one every tick of the RHIC clock. This counter determines where to store the current data in the memory. The run/stop command is distributed simultaneously throughout Trigger System by the RCC2. All the boards use this signal in the same way to start incrementing their address counters and this ensures that the address counters remain synchronized to each other. Data from a specific RHIC bunch crossing will be stored in the same memory location across all the boards. This information is used to determine which data to read out once a trigger has been issued. The “latch address” command is used for monitoring purposes, to check that all the boards are still synchronized. The spare may be used in the future.

- 13) Requirement: **Slow Communications**. A slow communications path is needed for remote reconfiguration of the SERDES logic, control and monitoring of board operations (including status of RCC2 control signals) and access to the memory. At a minimum this path must include an Ethernet connection and a process capable of opening an Ethernet socket and responding to incoming communications NOTE: There is no requirement for the SER boards to have a fast memory read out path. Data that is stored in this memory is for debug purposes only and does not need to be read out during data taking.

Justification: During data taking STAR is not physically accessible so all communication must be done remotely. Run Control currently communicates with all of STAR’s subsystems over Ethernet.

- 14) Requirement: **Configuration Time**. The total time to configure the SER must be less than 15 seconds. This includes the time to set all registers and zero the memory. It does not include the time taken to download a new configuration for the SERDES logic since that happens very rarely, e.g. during RHIC startup or when the beam species/energy change. It also does not include the time to download simulated or test data to the memory since that only happens in test mode.
Justification: The total time to configure the whole Level-0 Data Processing System for normal STAR data taking is required to be no more than 15s.
- 15) Requirement: **Remote Reset**. The SER must have a remotely-accessible reset path to reset all logic to the power-on state.
Justification: Sometimes an error occurs that results in a processor freezing up. Since STAR is not physically accessible during data taking initiating the reset must be done remotely.
- 16) Requirement: **Front Panel Monitors**. The SER boards must each drive the RHIC clock and one data bit to output connectors located on the front panel.
Justification: This is necessary to simply check that the timing is correct and there are no setup-and-hold timing violations.
- 17) Requirement: **Front Panel Indicators**. The SER boards need front panel LEDs to monitor fuse, FPGA configuration state, clock source (RCC2 or local oscillator), RCC2 control signals (run/stop, spare) and memory mode (record/play).
Justification: During testing and debugging this is the easiest way to monitor these features.
- 18) Requirements: **Test Points**. The SER boards need test points connected to GND (for scope probes) and each power supply (to check voltages)
Justification: This makes testing and debugging much simpler.

Upgraded Data Storage and Manipulation Board – DSM2

- 1) Requirement: **Board Form Factor**. The DSM2 must have a form factor that fits into the existing STAR infrastructure, i.e. either VME or pizza boxes.
Justification: The DSM2 will fit into STAR
- 2) Requirement: **Input Data Format**. The DSM2 must receive serial data streams on a physical form factor TBD from either SER boards or the TRM.
Justification: This is how it gets its input.
- 3) Requirement: **Clock and Control**. The DSM2 must receive the RHIC clock and STAR's synchronous control signals from the RCC2.
Justification: This is how the clock and control signals are distributed around the system.
- 4) Requirement: **Serial Output Data**. The DSM2 must produce serial output streams on the same form factor as it receives the input data
Justification: This will allow us to chain DSM2 boards together, either for future expansion or to construct a TRM. Using the same form factor for all these connections will also simplify development, maintenance and purchasing.
- 5) Requirement: **Parallel Output Data**. The DSM2 must be able to drive at least 32 parallel output bits on 5V PECL cables.
Justification: This will allow the DSM2 boards to pass data to the existing TCU and the existing Scaler System. If a TCU upgrade is implemented as a DSM2 board with

modified firmware then this feature would also allow the upgraded TCU to communicate with the existing TCD system.

NOTE: It would be nice, but it is not a requirement, for the DSM2 to have 48 or even 64 parallel output bits. All existing DSM boards have 32 parallel output bits on two 16-bit cables. Those DSMS that feed the TCU are limited to sending it 16 bits because the other set of 16 bits is sent to the Scaler System. In a few cases this has resulted in the need for multiple DSM algorithms with just slightly varying output bits. More output bits would reduce this problem.

- 6) Requirement: **Data Analysis**. The DSM2 must analyze all the incoming data using user-provided algorithms.

Justification: These algorithms check to see if an interesting interaction occurred, and that information will be used to make the trigger decision.

- 7) Requirement: **Data Delay**. It must be possible for the DSM2 to delay both incoming and outgoing data by an integral number of RHIC clock ticks. The maximum delay is 10 RHIC clock ticks.

Justification: It was shown in Figure 2 that data from different detectors arrives at the Level-0 Data Processing System at different times. Data from early detectors must be delayed to wait for data from later detectors. Depending on the circumstances it can be necessary to delay either or both of the incoming and outgoing data. In the existing DSM system no data is delayed by more than a few RHIC clock ticks, so a maximum of 10 ticks will cover all existing cases and allow for the addition of significantly faster detectors in the future.

- 8) Requirement: **Field-Programmable Logic**. The de-serialization of the incoming data, the implementation of the algorithms and the data delay (both input and output) must all be accomplished in reconfigurable logic

Justification: The format of the incoming serial data will be specific to each detector. Some detectors will send ADC and TAC values. Others will send the results of pre-processing the raw data. These formats may change over time as the detector subsystems are upgraded. The users customize the algorithms according to the energy and particle type of the colliding beams. New algorithms are loaded on a regular basis. On the current DSM boards the data delay is implemented using discreet FIFO ICs that are separate from the programmable logic unit. Modern FPGAs are large enough that this is no longer necessary; the FIFO functionality can be implemented inside the programmable logic unit, which will simplify the board design.

NOTE: There is no longer any need for the hardware look-up tables that exist on the current DSM boards separate from the programmable logic unit. Those LUTs were implemented in discreet IC memory chips and originally used for applying gain and pedestal corrections to raw ADC values. That functionality is now implemented in the QT boards so it is no longer needed on the DSM boards. The only remaining function for the current LUTs is to zero out noisy channels. In the DSM2 boards this function can be integrated into the algorithm logic. There is therefore no need for separate, stand-alone LUTs.

- 9) Requirement: **RHIC Clock Frequency**. The DSM2 must be capable of operating with a RHIC clock frequency that is 3 times higher than the current value with minimal hardware and firmware changes

Justification: As part of its plan to increase luminosity, RHIC is planning to increase the RHIC clock frequency (most likely by a factor of 3) in the next 5-10 years.

- 10) Requirement: **Fast Readout**. The DSM2 must have a readout path so data stored in the memory can be readout during data taking. Reading out data must not interrupt storage of new data. The readout path must operate as fast as possible. At a minimum it must be fast enough to maintain a steady-state event rate of significantly more than 20 kHz.

Justification: The data that led to a trigger must be read out as soon as possible in order for the full event to be built and analyzed. The aim is to make a decision whether to abort the event or save it. Any bunch crossing can result in a trigger, so the data from all bunch crossings must be saved even if readout of a previous bunch crossing is in progress.

STAR is currently in the process of upgrading various systems with the aim of increasing the overall event rate to 10-20 kHz. So long as the DSM2 can be readout significantly faster than 20 kHz it will avoid becoming the bottleneck that limits the overall STAR event rate. In addition, the amount of memory that is needed to store the data will depend on the maximum delay between when data is stored and when it is read out. Minimizing the readout time will minimize the memory requirements, and therefore simplify the memory implementation.

- 11) Requirement: **Readout Selection**. The DSM2 must be capable of interpreting a mask that will be sent as part of the readout command to determine which, if any, of its data needs to be read out.

Justification: Some physics analyses only require the data from a subset of trigger detectors. Event rates for those triggers can be increased if only the necessary data is readout. In addition, some data sets stored in the DSM2 boards are expected to be duplicates, e.g. the BEMC data will most likely be sent to two DSM2 boards (see Figure 3). This is a large data volume (see Figure 2) and there is no need to read out duplicate sets. The user will therefore need to specify which data (raw and/or processed) should be read from which DSM2 boards for each trigger type. The appropriate mask will be distributed along with the readout command every time a trigger is issued.

- 12) Requirement: **Data Storage**. The DSM2 must store all detector data for a time period long enough to cover the maximum time from data storage to readout. The memory access time must be short enough that it is possible to make at least two accesses during each RHIC clock tick.

Justification: The decision to issue a trigger, and therefore to read out the data, is made downstream of the DSM2 by the TCU. There is therefore a delay between when data is received by the DSM2 and when it is read out. The maximum delay occurs when a burst of triggers is issued in consecutive RHIC clock ticks. The maximum burst length is 4095 triggers because triggers are only issued if a token is available, and STAR has a maximum of 4095 tokens. The new system must therefore store all data from every RHIC clock tick long enough for a burst of 4095 events to be readout. This will guarantee that data from triggered events is always available to be read out. Once the data has been readout the memory can be re-used for new data, so a circular buffer can be used. During normal data taking new data will be stored at the next location in the circular buffer every tick of the RHIC clock. At random times it will also be necessary to read out data from a random location in the buffer when a readout command has been received. The memory access time must therefore be short enough to allow at least two accesses during each RHIC clock tick.

NOTE: A more precise definition of this requirement will be based on the results of testing a prototype readout system, and on a survey of what memory chips are available.

- 13) Requirement: **Test Mode**. The DSM2 must be configurable so that data can be stored in the memory or can be read from the memory and sent on downstream.

Justification: It is necessary to periodically test the functionality of the DSM2, and the Level-0 data processing system as a whole, without the presence of an external detector. This feature would allow simulated data to flow through the system as part of the testing procedure

- 14) Requirement: **Memory Control**. During data taking the DSM2 must use the STAR synchronous run/stop signal to control when access to the memory (reading or writing depending on the mode) should begin. The user must be able to specify the starting address for that memory access. The DSM2 must also respond to the STAR synchronous “latch address” command and have the potential to respond to what is currently the spare command.

Justification: Every board in the Trigger System has an address counter that starts from a user-specified value and increments by one every tick of the RHIC clock. This counter determines where to store the current data in the memory. The run/stop command is distributed simultaneously throughout Trigger System by the RCC2. All the boards then use this signal in the same way to start incrementing their address counters and this ensures that the address counters remain synchronized to each other. The users adjust the starting address, based on the arrival time of the data and how much delay is added to ensure that data from a specific RHIC bunch crossing will be stored in the same memory location across all the boards. This information is used to determine which data to read out once a trigger has been issued. The “latch address” command is used for monitoring purposes, to check that all the boards are still synchronized. The spare may be used in the future.

- 15) Requirement: **Slow Communications**. The DSM2 must have a slow communications path for remote reconfiguration of algorithms, control and monitoring of board operations (including status of RCC2 control signals) and non-time-critical access to the memory. At a minimum this path must include an Ethernet connection and a process capable of opening an Ethernet socket and responding to incoming communications.

Justification: During data taking STAR is not physically accessible so all communication must be done remotely. Run Control currently communicates with all of STAR’s subsystems over Ethernet.

- 16) Requirement: **Configuration Time**. The total time to configure the DSM2 must be less than 15 seconds. This includes the time to set all registers and zero the memory. It does not include the time taken to download new algorithm logic since that happens very rarely, e.g. during RHIC startup or when the beam species/energy change. It also does not include the time to download simulated or test data to the memory since that only happens in test mode.

Justification: The total time to configure the whole Level-0 Data Processing System for normal STAR data taking is required to be no more than 15s.

- 17) Requirement: **Remote Reset**. The DSM2 must have a remotely-accessible reset path to reset all logic to power-on state.

Justification: Sometimes an error occurs that results in a processor freezing up. Since STAR is not physically accessible during data taking initiating the reset must be done remotely.

- 18) Requirement: **Lemo Outputs**. The DSM2 needs a front panel Lemo connector driven from the data analysis processor.

Justification: Some downstream detectors require a fast pre-trigger that is issued before the final trigger decision. This feature could also be useful for debugging purposes.

- 19) Requirement: **Front Panel Monitors**. The DSM2 must drive the RHIC clock, one serial I/O stream and one parallel I/O stream to output connectors on the front panel.

Justification: This is necessary to simply check that the timing is correct and there are no setup-and-hold timing violations.

- 20) Requirement: **Front Panel Indicators**. The DSM2 needs front panel LEDs to monitor the fuse, FPGA configuration state, clock source (RCC2 or local oscillator), RCC2 control signals (run/stop, spare) and memory mode (record/play).

Justification: During testing and debugging this is the easiest way to monitor these features.

- 21) Requirement: **Test Points**. The DSM2 needs test points connected to GND (for scope probes) and each power supply (to check voltages)

Justification: This makes testing and debugging much simpler.