Herzberg Extensible Adaptive Real-time Toolkit (HEART) software architecture

Dan Kerley[a], Malcolm Smith[a], Jean-Pierre Véran[a], Jennifer Dunn[a], and Niloofar Gharavi[a]

[a]National Research Council of Canada - Herzberg, 5071 West Saanich Rd, Victoria, Canada

ABSTRACT

The National Research Council of Canada Herzberg (NRC-H) is developing the Herzberg Extensible Adaptive Real-time Toolkit (HEART) for the next generation of Adaptive Optics (AO) systems. HEART is a distributed and scalable software framework that includes a collection of libraries, tools and other software that can be assembled to construct a wide variety of real-time AO control systems. HEART is currently under development, implemented in C/C++, and is intended to run on off-the-shelf CPUs; HEART has evolved from the NRC-H design of the Narrow Field Infrared Adaptive Optics System (NFIRAOS) Real Time Controller (RTC) for the Thirty Meter Telescope (TMT), which passed its final design review. HEART is principally designed for Laser Guide Star (LGS) Multi-Conjugate Adaptive Optics (MCAO), however the modular architecture of HEART lends itself to be restructured for different forms of AO systems, such as: Single Conjugate Adaptive Optics (SCAO), Laser Tomography Adaptive Optics (LTAO), Ground Layer Adaptive Optics (GLAO), and Multi-Object Adaptive Optics (MOAO). The primary goal of HEART is to reduce the development time and cost of next generation AO control systems, while providing more reliable and robust control software that minimizes maintenance effort and promotes code reusability.

Keywords: HEART, RTC, AO, MCAO, MOAO, real-time, adaptive optics

1. INTRODUCTION

The Herzberg Extensible Adaptive Real-time Toolkit (HEART) is a CPU-based software framework for the development of a Real-Time Controller (RTC) for a generalized Adaptive Optics (AO) system. This approach reduces development effort due to use of a familiar CPU platform and leveraging generalized modules which can be configured for a variety of different AO systems including Single Conjugate Adaptive Optics (SCAO), Laser Tomography Adaptive Optics (LTAO), Ground Layer Adaptive Optics (GLAO), and Multi-Object Adaptive Optics (MOAO) as well as support for both Laser Guide Star (LGS) and Natural Guide Star (NGS) AO correction. One of the main strengths of HEART is its distributed and scalable architecture. Since the next generation of AO systems for ELTs will become more complex and demanding in the future, it is critical not only that HEART allows for scalability through its distributed structures, but that this is at the core of its design.

HEART is heavily based off the design of the Narrow Field Infrared Adaptive Optics System (NFIRAOS)[1] Real Time Controller (RTC), which is the first-light AO system for the Thirty Meter Telescope (TMT).[2] The NFIRAOS RTC[3] passed its final design review in Spring of 2018 and will be starting construction phase, using the HEART framework, in early 2020.

The HEART design is also highly modular and provides a clear separation between the code and the physical devices by defined generalized interfaces. This design choice was made largely to accommodate the interface for future Wavefront Sensors (WFSs) and Deformable Mirrors (DMs) without requiring substantial changes to the core HEART code. However, this separation from hardware also extends to the servers on which the HEART RTC code runs. This is achieved by defining all real-time internal and external communication via UDP sockets over high-speed Ethernet links. Internally the HEART interfaces allow various part of the real-time pipeline to run on different servers depending on configuration. For external interfaces with AO components (e.g. WFSs and DMs) communication with the RTC is established via dedicated driver software that adapts the native component interface to the HEART universal internal socket interface.

Further author information: (Send correspondence to Dan Kerley)
Dan Kerley: E-mail: dan.kerley@nrc-cnrc.gc.ca, Telephone: 1-250-363-8107

2. SOFTWARE ARCHITECTURE

While a RTC can be thought of as single unified system, the HEART architecture breaks the system down into a collection of 'roles'. For larger AO systems, multiple servers would be used, each being assigned nominally one role per server, while for smaller AO systems, a single server could be used to run all the roles. In either case the collection of server roles work in parallel to achieve the desired performance.

Within the HEART architecture, each server in the system has at least one of the following four roles:

- High-Order Processing (HOP)
- Wavefront Corrector Controller (WCC)
- Telemetry Engineering Display (TED)
- Persistent Telemetry Storage (PTS)

The High-Order Processing (HOP) role accepts high order WFS pixels and processes them into reconstructed High Order (HO) DM command vectors. There is one HOP role for each high order WFS in the system. The Wavefront Corrector Controller (WCC) role accepts HO vectors from all the HOP roles, as well as pixels from all low order WFSs. The WCC processes these inputs to generate commands for the wavefront corrector elements, such as DMs. It is possible that there are multiple WCC roles, each receiving HO vectors, particularly in the case of an MOAO system, however in the standard HEART implementation, as for a GLAO or MCAO system, only a single WCC role is required. Telemetry Engineering Display (TED) role acts as a gateway to the RTC, handling commands, publish and subscribing to events and telemetry, and well as hosting web-based Graphical User Interfaces (GUIs). The TED role will require the most customization, as it must fit within the constrains of external subsystems and interface standards of the particular telescope and instrument. Finally the Persistent Telemetry Storage (PTS) role stores data to disk for engineering purposes and post-processing such as point-spread function reconstruction. Depending of the system dimensions and rate of telemetry data there may need to be multiple PTS roles. Together the HOP and WCC roles form the real-time pipeline of the RTC, while the TED and PTS roles perform non-real-time support activities.

To achieve the distributed nature of HEART, internal real-time communication is based on UDP sockets over high-speed Ethernet links. It's useful to note that the overhead of using UDP messages locally on a server as a trigger signal or for synchronization, only adds a trivial amount of latency when compared to using semaphores, while gaining some useful advantages. The primary advantage is the ability to move processes between servers without needing to modify code, only configuration; while a useful secondary advantage is that meta data can also be passed via the UDP datagram which can simplify code, particularly in the case of synchronizing multiple data streams. Additionally using Ethernet sockets allows a test program to easily act both as a data source and sink, which facilitates simulated operations of the RTC without requiring the use of the physical hardware such as WFSs and DMs. This also makes testing and benchmarking of the RTC much simpler since test programs can be constructed to interact directly with sub-components of the RTC, easily testing blocks of the real-time pipeline in isolation.

Figure 1 shows an example configuration of server roles for an MCAO system such as NFIRAOS, where multiple LGS WFS are used for tomographic reconstruction of the atmospheric turbulence in a number of layers located at different altitudes above the telescope. NFIRAOS uses six LGS WFSs and two deformable mirrors (DMs). The highlight boxes in red show external sub-system or hardware where custom interface code would be required. In this example, LGS WFSs stream pixels to HOP roles, one per WFS. HO DM vectors from the HOP roles are aggregated by the WCC role, which along with low order NGS pixels are used to generated DM and Tip/Tilt Stage (TTS) commands. The TED role interfaces to higher-level instrument software and observatory software. The PTS roles collect telemetry for the Point-Spread Function Reconstructor (PSFR) and system diagnostics; one PTS role is dedicated to storing LGS pixels, while the other PTS role stores all other telemetry. Additionally, a Reconstructor Parameter Generator (RPG) provides updates to the control matrix and other optimized control parameters for the RTC.

Another style of AO system that incorporates multiple WFS are MOAO systems. These systems have several narrow-field science channels positioned over a large field of regard to observe specific science targets. Each channel has a dedicated DM, and these DMs are driven by projecting the turbulence estimate reconstructed by tomography in the exact direction of the science object, potentially providing a higher level of correction than

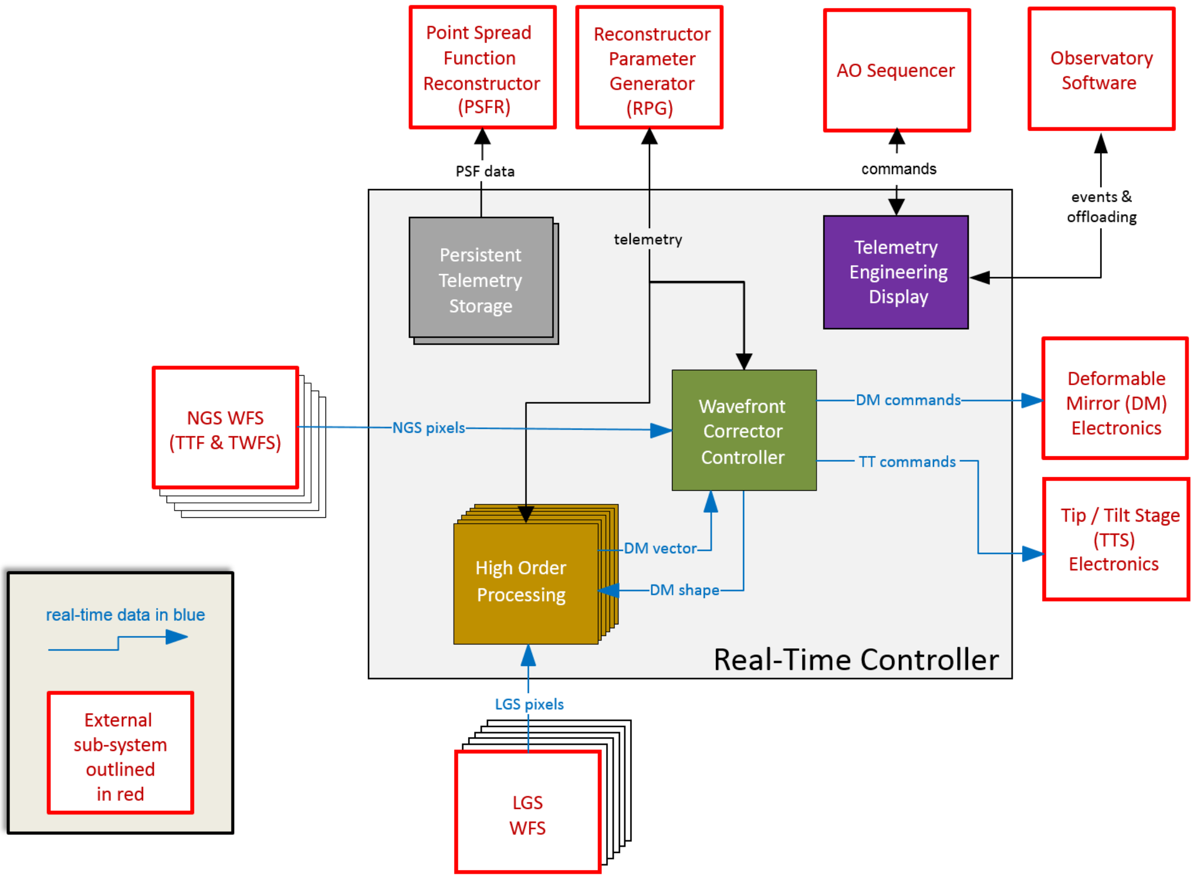

Figure 1. HEART Server Roles - Example LGS MCAO System

MCAO, which must correct over the entire field of regard. These types of AO system are easily supported by HEART, where there would be an additional WCC role for each open loop DM channel.

3. HARDWARE ARCHITECTURE

HEART is designed for large and demanding AO systems, with many measurements (i.e. WFS pixels/gradients) and many degrees of freedom (i.e. DM modes/actuator commands). The scale of the many next generation AO systems requires RTC computations to be spread over several different physical servers or accelerator cards, such as GPUs. However a RTC for many smaller or current AO systems could be run on a single modern multi-core server. Therefore, HEART was designed to be distributed across many processors, while being agnostic as to whether different parts of the real-time pipeline are run on different servers communicating via Ethernet, or simply running on the same server.

The HEART project has decided to focus on CPU-based solutions, using Commercial-Off The Shelf (COTS) servers, running standard Linux distributions with the real-time patch, as we believe this is the simplest, more reliable and most cost-effective approach, when taking development and maintenance effort into consideration. That said, we have left the design open and flexible enough to support the use of GPUs in future development.

The primary reason for the distributed architecture in HEART is to accommodate a very large high-order Matrix Vector Multiply (MVM), such as those required for ELTs. For many AO systems, with existing technology, it is practical to compute a full component of the DM command vector corresponding to a single high order WFS on a single dual-socket server. This modularity in the design makes it straightforward to either increase or decrease the number of servers depending on number of the high order WFSs and the size of the MVM. A partitioned server can then be grouped into server roles, which would run concurrently on a single server with a particular hardware configuration. Therefore, while the HEART design calls for several different server roles, the activities of all of these roles could be executed on a single server for a typical AO system on an eight-meter class telescope.

The mapping of HEART server roles and real-time pipeline blocks to hardware and the details of the server hardware selection are specific to each AO system. However a natural delineation is that a server, with the HOP role, is used to process high order gradients and perform the corresponding partial high order reconstruction for each high order WFS. Each HOP server then sends partial HO vectors to another server, with the WCC role, that aggregates the vectors and executes the rest of the real-time pipeline, including the processing and reconstruction of low order measurements.

Modern servers often include more than one set of memory interfaces. For example, in a dual-socket Xeon server, each CPU will directly access the memory which is physically connected to it, but will need to communicate with the other CPU in order to access the remaining memory. Accessing remote memory is slower than accessing local (directly connected) memory. This architecture is referred to as a non-uniform memory architecture (NUMA). In fact, some modern CPUs contain more than one NUMA region each. When mapping real-time pipeline blocks onto a single server, it is natural to take the NUMA architecture into account. This mapping is straightforward for smaller systems since we simply make sure that each instance of a pipeline block does not span more than one NUMA region. Each CPU (or NUMA region) may however host more than one pipeline block. Restricting pipeline blocks to a single NUMA region reduces the number of remote memory accesses. Software, such a HEART, which takes into account the NUMA architecture is called NUMA aware.

4. REAL-TIME PIPELINE

The HEART modular design is based on a framework of configurable blocks assembled to produce the real-time pipeline for AO correction. A hierarchical architecture has been adopted, in which the generalized HEART pipeline is decomposed into the follow top-level standard super-blocks:

- LGS Processing
- High Order NGS Processing
- Low Order NGS Processing
- Truth Wavefront Sensor (TWFS) Mode Processing
- High Order Reconstructor
- Low Order Reconstructor
- Combination and Temporal Filtering
- Closed Loop Wavefront Corrector Control
- Open Loop Wavefront Corrector Control
- Wavefront Controller
- Telescope Offload

Each super-block represents a collection of related processes common to many RTCs. These super-blocks are shown in Figure 2, represented by round-corner blocks. Square-cornered blocks represent the lower-level decomposition of processes within a super-block. Each lower-level block is then further decomposed into threads, which parallelize the problem and then re-synchronized to produce the block's output streams. Blocks with double-borders signify that there may be multiple instances of the block. As an example for LGS MCAO system, there would be one LGS Processing super-block per LGS WFS. Additionally dashed-line blocks denote custom interfaces where instrument specific code is required to interface with hardware or external sub-systems, while dotted-line block denote the actual hardware or external sub-systems. Finally blocks and connections in red denote processes only active in LGS AO mode while blocks and connections in blue denote those active in NGS AO mode, the remaining blocks are used in either mode. In this context LGS AO mode refers to any mode of operation where high order measurements are from one or more laser guide star, while NGS AO mode refers to high order measurements from natural guide stars.

It is important to keep in mind that all blocks and connections within HEART are optional and may be added or removed based on the specifics of the AO system that is being developed. As an example an Open

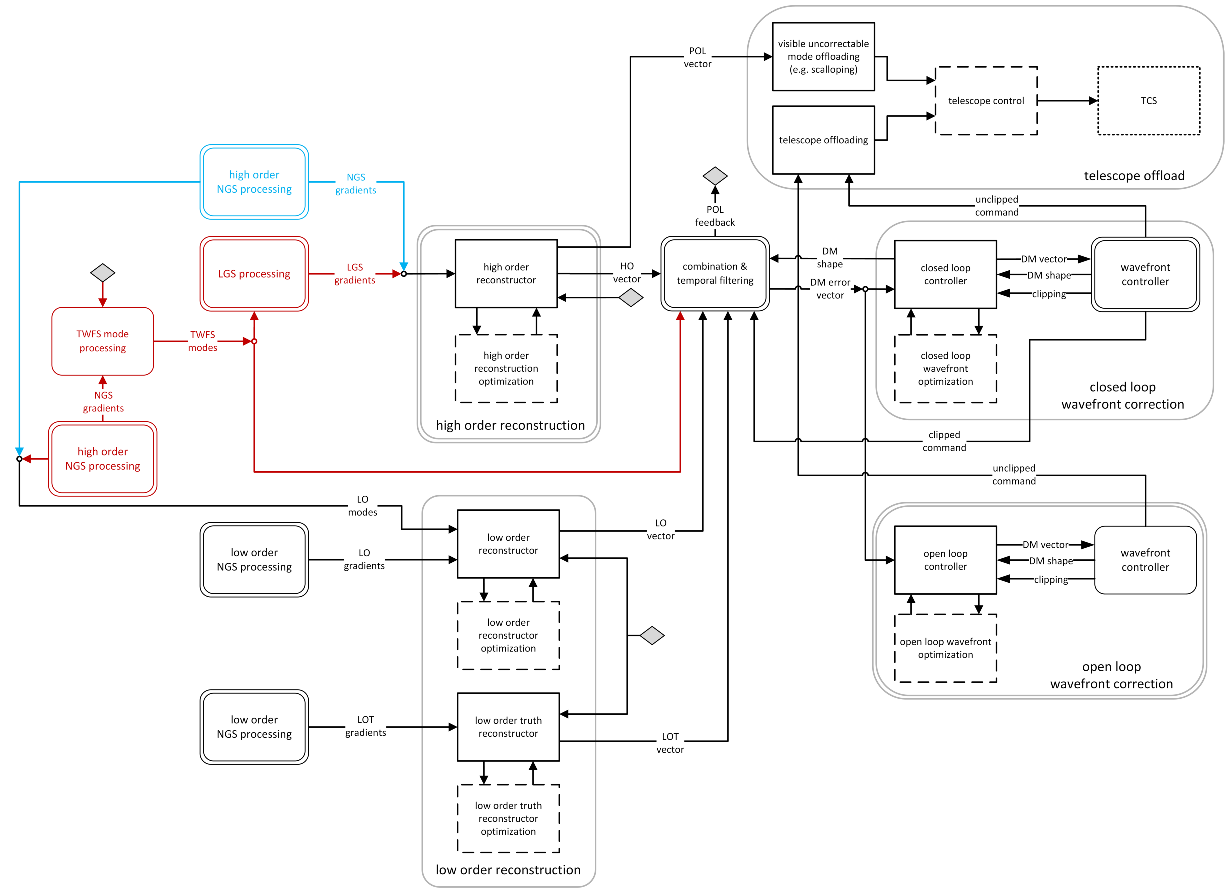

Figure 2. HEART Real-Time Pipeline Top-Level Block Diagram

Loop Wavefront Corrector Control block may only be used in an MOAO system; while other systems may only support NGS AO, therefore the LGS Processing and TWFS Mode Processing blocks would not be used. The interfaces between blocks and super-blocks are generalized such that they can be connected in a variety of configurations. These generalized interfaces allows for two important features of the HEART framework. First is that any block can be replaced with a customized block, that exceeds the configurability of the standard block, to meet the particular needs of an AO system, as long is it conforms to the generalized interface. This custom block can be inserted into the HEART pipeline and work seamlessly with the other standard HEART blocks. The second feature is that a custom block can be inserted in between any two blocks, to perform a specialized operation, assuming it preserves the generalized interface.

In the generalized LGS mode, pixels from each LGS WFS are streamed into separate LGS Processing blocks, each block produces LGS gradients corresponding to the input WFS. These gradients are in turn streamed to one or more High Order Reconstruction blocks. Additionally in LGS mode, one or more High Order NGS WFS will stream pixels into High Order NGS Processing block to produce a stream of NGS gradients. These gradients feed the TWFS Mode Processing block to produce TWFS modes which are used as a reference signal that can be added into the LGS Processing block or alternatively added after high order reconstruction in the Combination and Temporal Filtering block. Optionally, the NGS gradients can also be projected into low order modes and used as part of the Low Order Reconstruction block.

In the generalized NGS mode, pixels from each NGS WFS are streamed into a separate High Order NGS Processing block. Similar to LGS mode, these gradients are streamed to High Order Reconstruction blocks and optionally the Low Order Reconstruction block.

In the generalized HEART pipeline, high order gradients are streamed to one or more High Order Reconstruction block to produce HO DM command vectors. Depending on the size and dimensions of reconstruction problem, the High Order Reconstruction blocks may be configured in different ways, effectively dividing the problem base on either WFSs or Wavefront Correctors or both. Depending on this configuration the Combination and Temporal Filter blocks will aggregate the HO vectors differently, either concatenating or summing the vectors. Also note the reconstruction can be done either zonally or modally, changing the nature of the HO vector. Pseudo Open Loop (POL) feedback is also available, where state information from the wavefront corrector is projected into gradient-space to produce pseudo open loop gradient. Similarly Low Order (LO) WFSs and Low Order Truth (LOT) WFSs stream pixels into Low Order Processing blocks to produce LO and LOT gradients. A LOT WFS, typically located near or in the focal plane, provides quasi-static tip, tilt and optionally focus information to compensated for flexure and other slow-moving effects. These LO and LOT gradients, along with any projected LO mode from the high order processing and POL feedback, are used to produce LO and LOT mode vectors. By definition, the LOT vectors are updated at slower rate than the LO vectors, while either HO vector or the LO vector may be updated faster than the other.

The Combination and Temporal Filter block accepts HO, LO and LOT vectors, potentially all at different rates and combines them together. This block synchronizes these vectors, temporally blending the input streams together while executing at the fastest of the three input rates, as well as handling any missing or late data. This block produces DM error vectors for any open loop DMs, each one handled by an Open Loop Wavefront Correction block, as well as a single DM error vector handled by the Close Loop Wavefront Correction block. Close Loop Wavefront Correction block integrates the error signal and composes the integrated DM vector into virtual DMs. Each virtual DM vector sent to a Wavefront Controller block is in turn decomposed into one or more sets of physical wavefront controller commands, such as temporally and spatially splitting the virtual DM vector between a physical DM and a TTS. Feedback from the Wavefront Controller blocks can be sent to the Telescope Offload block to produce commands or events that can be offloaded to the Telescope Control System (TCS).

4.1 LGS Processing

Figure 3 show the expanded LGS Processing super-block, consists of several distinct internal processing blocks which includes: reading LGS pixels; computing LGS gradients using a choice of algorithms, including center of gravity, or matched-filter; offloading tip, tilt and focus to stabilize the LGS spots; and subtracting a reference vector based on TWFS modes. More complex gradient computation algorithms such as matched-filters can be optimized in real-time, e.g. via a dithering process. This block accepts a pixel stream from a single LGS WFS to produce a stream of LGS gradients that are fed to a High Order Reconstruction block. If there are multiple LGS WFSs then multiple instances of this block are used. It can also accept TWFS modes, from the TWFS Mode Processing block to update its LGS reference vector. The dashed blocks on the left of the diagram contain custom code that must be made to correspond to the details of the particular AO system, e.g. tip and tilt can be offloaded to a fast steering mirror either in the laser launch system, or in the AO system itself. The remaining blocks can be configured via external parameters for normal use. Each of the blocks within the diagram below can be modified if required, e.g. other gradient computation algorithms could be implemented as AO science progresses.

AO parameters for configuring the LGS Processing block include items such as: number of LGS WFSs, LGS WFS geometry, number of LGS WFS subapertures, mask of physical subapertures and mask of illuminated subapertures, number of pixels per LGS WFS subaperture, number of TWFS modes, offloading control parameters, and gradient computation method. While computed or measured AO parameters required by the LGS Processing include items such as: matched filters for gradient computation, gradients to tip/tilt/focus transformation matrix, focus to gradients transformation matrix, TWFS modes to LGS WFS gradients transformation matrix, LGS WFS detector flats, darks and biases.

The pixel reading and pixel processing for each LGS WFS can be performed independently. For large systems, each LGS WFS may be processed by a separate server. In smaller systems, each LGS WFS may be processed by a separate core in order to reduce latency. The mapping from LGS processing tasks to computer hardware will be specified in a configuration file. In order to make the mapping from tasks to hardware more clear, a single hardware configuration file will contain all of the mappings for a given AO system.

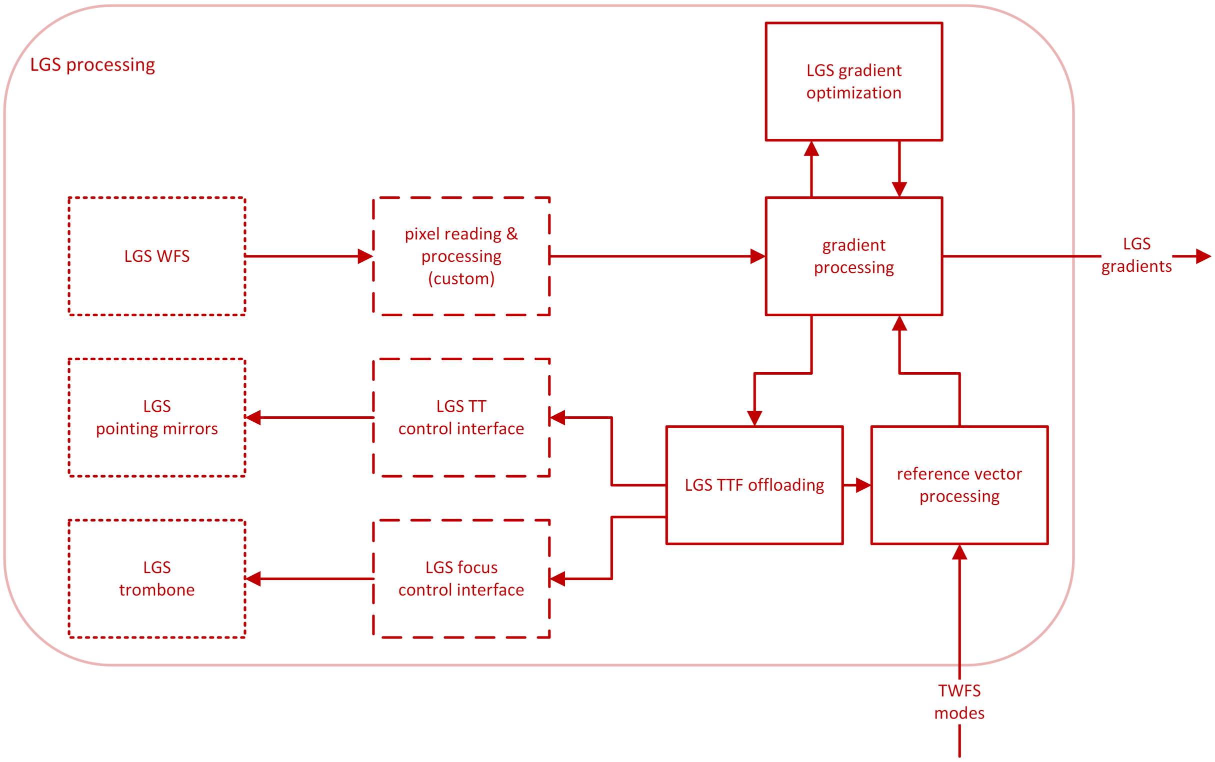

Figure 3. HEART LGS Processing Super-Block Diagram

For some systems, such as NFIRAOS, the LGS pixel reading corresponds to reading UDP datagrams containing LGS pixels from an Ethernet port. On other systems, the pixel reading may consist of reading pixels from a buffer populated by a WFS controller board. In either case, the pixel reading blocks produce a stream of LGS pixels which are then processed by the LGS pixel processing blocks. The processing starts as soon as the first pixels arrive.

4.2 High Order NGS Processing

The High Order NGS Processing super-block, as shown in Figure 4, consists of distinct processing stages which includes: reading NGS pixels, computing NGS gradients, high-order NGS gradient optimization, projecting NGS gradients to low-order modes, and high-order NGS steering. This block accepts a pixel stream from a single high order WFS to produce a stream of NGS gradients that are fed to a High Order Reconstruction block in NGS mode or to a TWFS Mode Processing block in LGS mode. Additionally high order gradients can be projected into low order modes to feed the Low Order Reconstruction block. If there are multiple high order WFSs then multiple instances of this block are used. The dashed boxes on the left of the diagram contain custom code that must be made to correspond to the details of the particular AO system. The remaining blocks can be configured via external parameters for normal use.

AO parameters for configuring the High Order NGS Processing block include items such as: number of high-order NGS WFSs, NGS WFS geometry, number of NGS WFS subapertures, number of pixels per NGS WFS subaperture, mask of illuminated subapertures, number of low-order modes, and gradient computation method. While computed AO parameters required for High Order NGS Processing block include items such as: matched filters for gradient computation, NGS WFS gradients to low-order modes transformation matrix, NGS WFS detector flats, darks and biases.

Pixel reading and processing is handled similarly to the LGS Processing block, as discussed in Section 4.1.

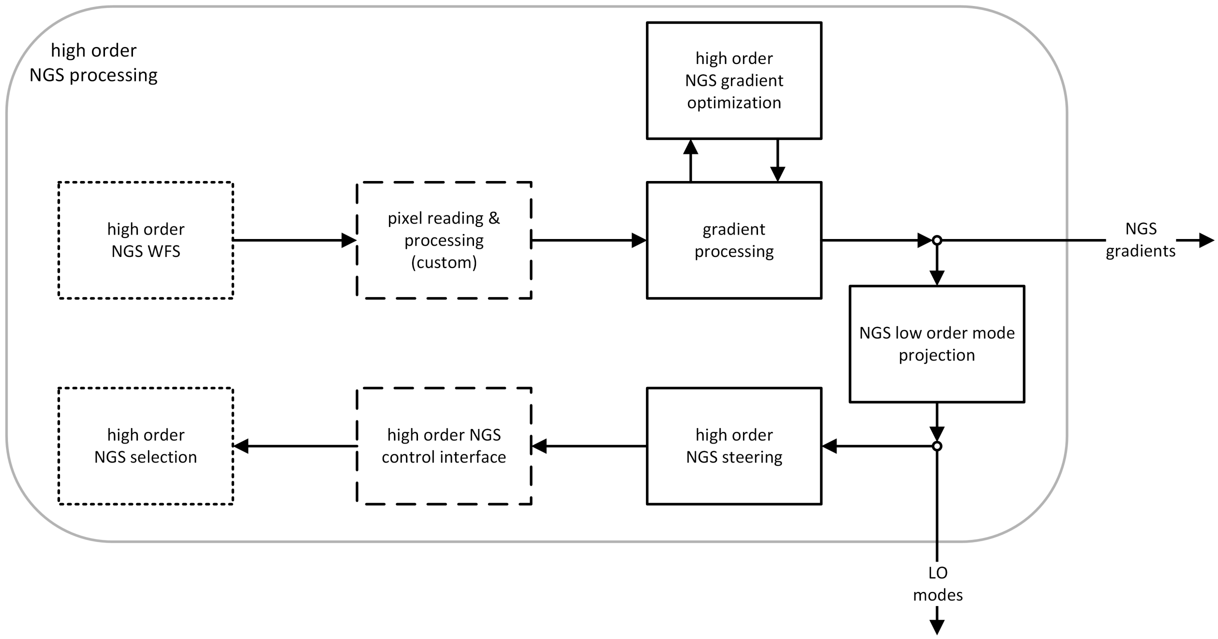

Figure 4. HEART High Order NGS Processing Super-Block Diagram

4.3 Low Order NGS Processing

The Low Order NGS Processing super-block, as shown in Figure 5, consists of distinct processing stages which include: reading low-order NGS pixels, computing low-order NGS gradients, optimizing gradient computation and steering the low-order NGS WFS. This block accepts a pixel stream from a single low order WFS to produce a stream of LO gradients that are fed to the Low Order Reconstruction block. It can be used to process pixels from any Low Order WFS, including an On-Interment Wavefront Sensor (OIWFS) or an On-Detector Guide Window (ODGW), and is used for both LO and LOT gradient processing; the primary difference being rate at which the pixel frames are received. If there are multiple low order WFSs then multiple instances of this block are used. This block is effectively just a more specific form of the generalized High Order NGS Processing super-block.

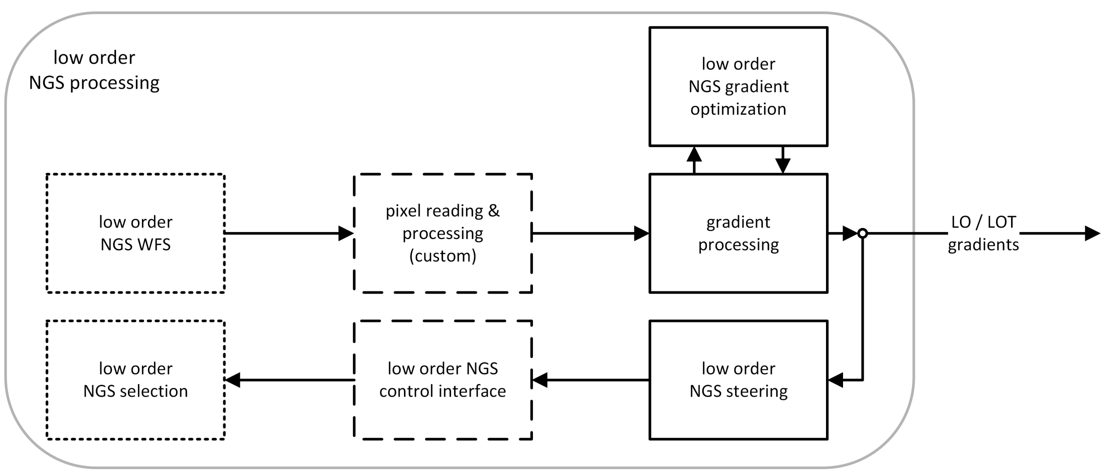

Figure 5. HEART Low Order NGS Processing Super-Block Diagram

AO parameters for configuring the Low Order NGS Processing block include items such as: number of low-order NGS WFSs, NGS WFS frame rate for each detector, size of each low-order detector, and gradient computation method. While computed AO parameters required for Low Order NGS Processing block include items such as: matched filters for gradient computation, NGS WFS detector flats, darks and biases.

Pixel reading and processing is handled similarly to the LGS Processing block, as discussed in Section 4.1

4.4 TWFS Mode Processing

The Truth Wavefront Sensor (TWFS) Mode Processing super-block, as shown in Figure 2, accepts gradient streams from one or more High Order NGS Processing blocks and generated TWFS modes via a reconstruction MVM. These modes are either sent to each of the LGS Processing blocks to be projected into the LGS gradient-space and used update the LGS reference vector, or they are sent to the Combination and Temporal Filter block to be projected into the reconstructed high-order DM vector space used to update HO reference vector. Optionally the TWFS reconstruction process may accept DM shape feedback, which is projected into NGS gradient-space, to produce pseudo open loop (POL) gradients from the NGS measurements.

AO parameters for configuring the High Order NGS Processing block include items such as: number of TWFS modes, number of high order NGS gradients used for TWFS modes, number of LGS WFSs, and whether pseudo open loop reconstruction is used. While computed AO parameters required for TWFS Mode Processing block include items such as: TWFS reconstruction matrix and POL projection matrix.

Sending TWFS modes to the LGS Processing blocks allows each LGS Processing block to project the modes into the gradient-space independently in parallel, typically occurring on different HOP servers. While sending TWFS modes to the Combination and Temporal Filter block allows for the LGS reference vector to be added after high order reconstruction by projecting the modes directly into the HO vector-space. Since both methods have their trade-offs, it is left to the specific implementation of the RTC to configure this option as most appropriate.

4.5 High Order Reconstruction

The High Order Reconstruction super-block, as shown in Figure 2, performs the reconstruction via a Matrix Vector Multiply (MVM). However this block could be modified to implement an iterative algorithm if required. The high order gradients are transformed into either a model or zonal HO vector, depending on the details of the control matrix. The High Order Reconstruction is a parallel operation being performed on multiple CPU cores, possibly across multiple servers. For large AO systems, the MVM is spread across multiple servers and multiple instances of the block, while for smaller AO systems, the MVM is likely spread across multiple cores of one or more physical CPUs. It is assumed that each instance of a super-block is implemented on a single machine, yet smaller AO systems may have many or all of the High Order Reconstruction blocks implemented on a single machine or even as a single instance of the block. Typically there is one High Order Reconstruction block for each set of high order gradient (e.g. one per LGS Processing block), each producing a partial High Order (HO) vector of full size. These HO vectors from each of the High Order Reconstruction blocks would be summed by the Combination and Temporal Filter block to produce a complete HO vector.

Alternatively the block can be configured such that all high order gradients are sent to each High Order Reconstruction block, each producing a complete and independent subset of the HO vector, in which case the Combination and Temporal Filter block would concatenate the HO vector subsets. Or in the case of an MOAO system, each High Order Reconstruction block could send complete HO vector, corresponding in a single channel, to an independent Combination and Temporal Filter block that would merge in any low order information before sending the DM vector to an Open Loop Wavefront Correction block. This configuration choice effectively boils down to choosing whether to parallelize the MVM column-wise or row-wise or even both. In the later case there would be many High Order Reconstruction blocks, each sending information to potentially multiple Combination and Temporal Filter blocks, however this solutions in only suggested if the MVM is extremely large in both directions.

Additionally, before applying the matrix, the gradients may be transformed into pseudo open loop (POL) gradients to allow for the application of optimal reconstructors, such as the minimum variance reconstructor (also known as the maximum a posteriori reconstructor).

AO parameters for configuring the High Order Reconstruction block include items such as: number of active DM actuators or modes, number of active HO subapertures, whether pseudo open loop reconstruction is used and whether HO vector output is model or zonal. While computed AO parameters required for High Order Reconstruction block include items such as: high order control matrix, DM shape to high order gradients transformation matrix.

4.6 Low Order Reconstruction

Like the High Order Reconstruction block, the Low Order Reconstruction super-block, as shown in Figure 2, performs the reconstruction of both the Low Order (LO) and Low order Truth (LOT) vectors via MVMs. Similar to the HO path, LO and LOT gradients are transformed into either model or zonal HO vectors, depending on the details of the control matrix. Each LO and LOT reconstruction is perform independently and in parallel, however due to the number of low order modes that are typically reconstructed, (e.g. tip, tilt, focus, and first-order plate scale modes), the MVM for each are not parallelized. The LO and LOT vector are temporally blended together by the Combination and Temporal Filter block.

AO parameters for configuring the Low Order Reconstruction block include items such as: number of NGS high order modes, number LO and LOT WFSs, whether pseudo open loop reconstruction is used, and number of low order modes. While computed AO parameters required for the Low Order Reconstruction block include items such as: gradient to low-order modes reconstruction matrices and DM shape to low order gradients transformation matrices.

4.7 Combination and Temporal Filtering

The Combination and Temporal Filtering super-block, as shown in Figure 6, aggregates and combines the high and low order vectors while applying temporal filtering. Depending on the partitioning of the high order MVM and High Order Reconstruction block, the HO vectors are either summed or concatenated, after which the reference vector based on TWFS modes can be added, for POL reconstruction DM shape information can be subtracted to produce an HO error vector, and for zonal HO vectors, low order modes can be removed. Then HO vector can be temporally filter by a generalized filter. LO and LOT vectors are blended together via complementary high and low pass filters and for POL reconstruction DM shape information can be subtracted to produce a net LO error vector. Additionally focus and rotation can be extracted from the low order paths for offloading purposes. Finally the high and low order paths are combined to be fed to Open or Closed Loop Wavefront Correction blocks.

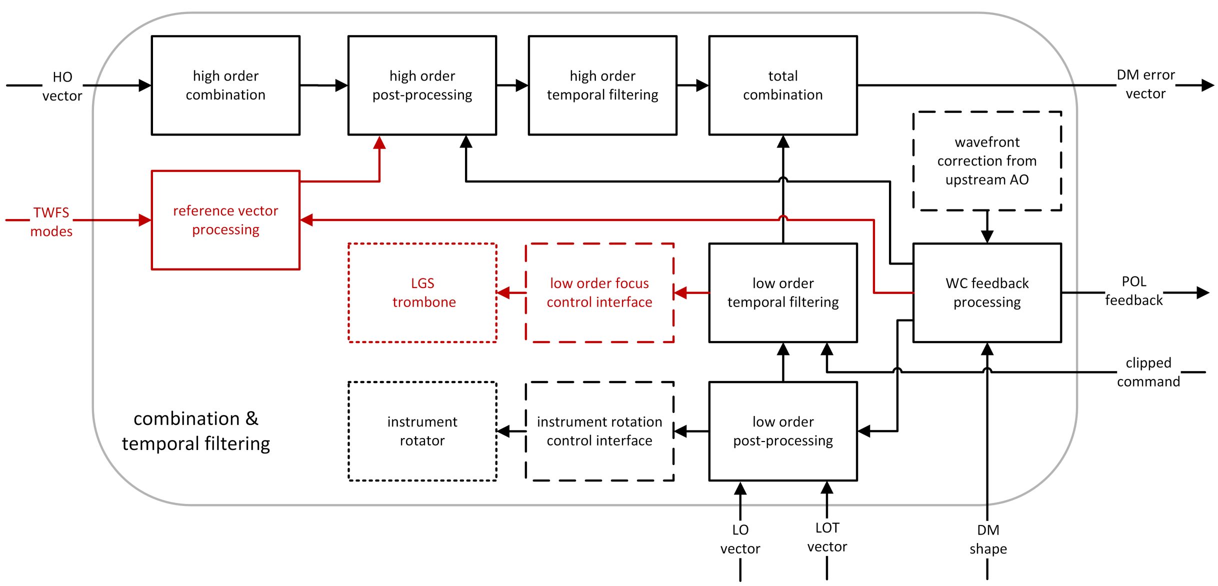

Figure 6. HEART Combination and Temporal Filtering Super-Block Diagram

AO parameters for configuring the Combination and Temporal Filtering block include items such as: zonal or modal HO vector, low order removal matrices, high and low order filtering parameters, low order mode to high order zonal transformation matrix, TWFS to HO vector transformation matrix, and POL projection matrices.

4.8 Closed Loop Wavefront Correction

The Closed Loop Wavefront Correction super-block, as shown in Figure 2, performs the integration of a DM error vector from a Combination and Temporal Filtering block and then decomposes that integrated DM vector into one or more virtual DMs to be send to corresponding Wavefront Controller blocks. As an example in an MCAO system this block would be configured to decompose integrated DM vector into a virtual ground-layer DM and one or more higher altitude DMs, where each of these virtual DMs may be comprised of one or more physical DMs. Additionally this block can be configured to remove uncontrolled modes from the integrator and can also be configured as a leaky integrator. This block uses clipping feedback from the Wavefront Controller blocks to modify the integrator state to prevent integral windup. Optionally DM shape information from each Wavefront Controller block is aggregated and sent back to the Combination and Temporal Filter block for POL feedback.

AO parameters for configuring the Closed Loop Wavefront Correction block include items such as: number of virtual DMs, number of rows in control matrix, total number of actuators or mode per virtual DMs, and mapping from control matrix order to virtual DMs. While computed AO parameters required for the Closed Loop Wavefront Correction block include items such as: main integrator gain, uncontrolled mode removal matrices, virtual DM registration matrix, and virtual DM extrapolation matrix.

4.9 Open Loop Wavefront Correction

The Open Loop Wavefront Correction super-block, as shown in Figure 2, is similar to the Closed Loop Wavefront Correction block, with the exception that DM error vectors from the Combination and Temporal Filtering block are not integrated and are instead sent to an Wavefront Controller block that represents an open loop controlled virtual DM. This block is intended to be used with an MOAO system, where there would be one Open Loop Wavefront Correction block per science channel. This block may be used in conjugation with a Closed Loop Wavefront Correction, if the RTC is also responsible for the closed loops aspects of the AO system, or on its own if an up-steam AO system is providing a sharpened image to the MOAO system.

4.10 Wavefront Controller

The Wavefront Controller super-block, as shown in Figure 7, is a sub-block of the Open Loop and Closed Loops Wavefront Correction blocks that accepts virtual DM command vectors from the open or closed loop controllers, as shown in Figure 2. These virtual DM commands are spatially and temporally decomposed into commands for the physical wavefront correctors (e.g. DMs and TTSs). A virtual DM may be as simple as a single DM or may consist of more than one DM, such as in a woofer/tweeter configuration, or may include a single DM on a TTS where low-frequency tip/tilt is apply to the TTS and high-frequency tip/tilt, along with all other modes, are applied to the DM. This block also supports feedback from a figure sensor to drive a DM to the requested shape; this is primarily intended for used with an 'open loop' DM in an MOAO system. Due to the variability of system configurations, this block may need to be customized to fit the particular needs of the AO system, if the standard configurations are not sufficient.

AO parameters for configuring the Wavefront Controller block include items such as: arrangement of physical correctors, spatial and temporal filters for each physical corrector, figure sensor control parameters, command destinations for physical corrector, and command ordering and degrees of freedom.

4.11 Telescope Offload

The Telescope Offload super-block, as shown in Figure 2, projects DM shape feedback into telescope offload modes, filters the modes and temporally down-samples them to be applied by the telescope control system. This block also contains projections, temporal filtering and down-sampling for primary mirror scalloping effects based on HO POL gradients.

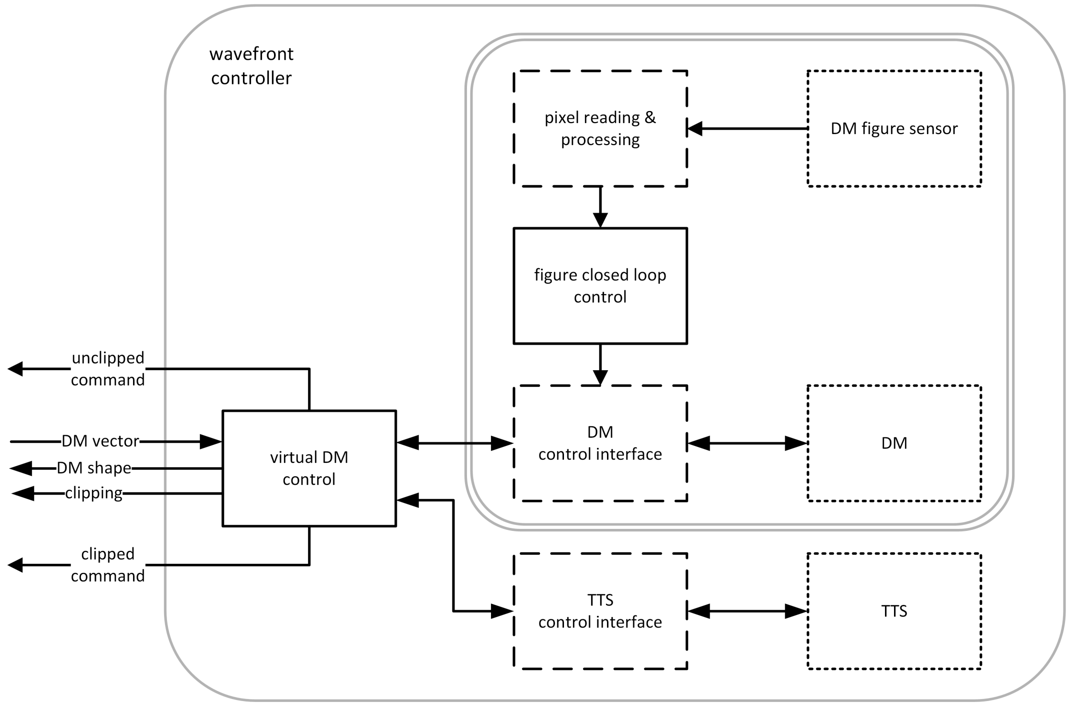

Figure 7. HEART Wavefront Controller Super-Block Diagram

AO parameters for configuring the Telescope Offload block include items such as: offload rates, anti-aliasing filter parameters, DM shape to telescope mode projection matrix, and HO gradients to scalloping mode projection matrix.

5. CONCLUSION

The HEART software framework, in development at NRC-H, provides a flexible and scalable C/C++ platform, on standard CPU-based Linux servers, for the development of a wide variety of RTCs for the next generation of large ELT AO systems. The use of UDP sockets in HEART makes it easy to develop smaller AO systems on a single server, or scale up to larger systems running across several servers. This scalability and modular configuration is baked into the architecture, resulting in useful and configurable AO building blocks. This package will reduce the cost, schedule and risk of RTC development by re-configuring generalized AO blocks and assembling them together in a distributed manner.

REFERENCES

[1] Crane, J., Herriot, G., Andersen, D., Atwood, J., Byrnes, P., Densmore, A., Dunn, J., Fitzsimmons, J., Hardy, T., Hoff, B., Jackson, K., Kerley, D., Lardiére, O., Smith, M., Stocks, J., Véran, J.-P., Boyer, C., Wang, L., Trancho, G., and Trubey, M., "NFIRAOS adaptive optics for the Thirty Meter Telescope," in [*Adaptive Optics Systems VI*], *Proc. SPIE* **10703** (2018).

[2] Boyer, C., "Adaptive optics program at TMT," in [*Adaptive Optics Systems VI*], *Proc. SPIE* **10703** (2018).

[3] Dunn, J., Kerley, D., Smith, M., Chapin, E., Herriot, G., Véran, J.-P., Boyer, C., Wang, L., Gilles, L., and Trancho, G., "The Real-time Controller (RTC) for the Narrow Field Infrared Adaptive Optics System (NFIRAOS) for TMT Final Design," in [*Adaptive Optics Systems VI*], *Proc. SPIE* **10703** (2018).