

The Compute and Control for Adaptive Optics (cacao) real-time control software package

Arnaud Sevin,
Julien Bernard,
Damien Gradatour



Observatoire de Paris,
(France)

**Software engineering &
development**
**Integration with COMPASS
simulation environment, RT
hardware (GPU, FPGA) and RT
processes management/monitoring**

SCEAO Subaru Coronagraphic
Extreme Adaptive Optics

Olivier Guyon, Julien Lozi,
Nour Skaf

Subaru Telescope / SCEAO
(US, Japan)

**cacao development and
On-sky testing for ExAO
applications**

جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology



Hatem Ltaief & Dalal Sukkari
KAUST (Saudi Arabia)

HPC expertise
**Develop and provide linear
algebra libraries for Machine
Learning**

Users / co-developers

Keck Observatory
Sylvain Cetre



**Facility-class AO
(NGS & LGS)**

Kernel project/OCA
Frantz Martinache



Focal plane WFS/C

MagAO-X
Jared Males



Extreme-AO

Subaru Telescope
Christophe Clergeon



**Facility-class AO
(NGS & LGS)**



THE UNIVERSITY OF
SYDNEY

Alison Wong & Barnaby Norris

**Machine learning / AI
Neural Networks**

cacao's goals

Support today & tomorrow's off-the-shelf powerful computing hardware

- manycore systems (CPU, GPU) and FPGAs
- high performance computing engine to requirements of large scale AO systems

Support and enable advanced AO systems and algorithms

- flexible modular software architecture
- scalable solution
- built-in (and growing) machine learning support for predictive control, sensor fusion
- facilitates asynchronous links between sensors and loops
- full speed telemetry to disk for post-processing

Foreseen applications : Extreme-AO, MCAO, Tomographic AO ...

Community effort, fully open-source

- no proprietary / closed source roadblocks
- enables easy/quick implementation of new algorithms
- adopts standard data stream format

Easy to adopt and use: short path from ideas to real-time implementation

- abstract away / hide HPC complexity
- manage challenging real-time timing constraints for the user
- provide high-level GPU/CPU configuration

Current Status

Provides low-latency to run control loops

→ Use mixed CPU & GPU resources, configured to RTC computer system

On SCExAO, control matrix is $14,000 \times 2000$. Matrix-vector computed in $\sim 100\mu\text{s}$ using 15% of RTC resources @ 3kHz

Portable, open source, modular, COTS hardware

→ No closed-source driver

→ std Linux install (no need for real-time OS)

→ using NVIDIA GPUs, also working on FPGA use

→ All code on github: <https://github.com/cacao-org/cacao>

Easy for collaborators to improve/add processes

→ Hooks to data streams in Python or C

→ Template code, easy to adapt and implement new algorithms

→ Provide abstraction of link between loops

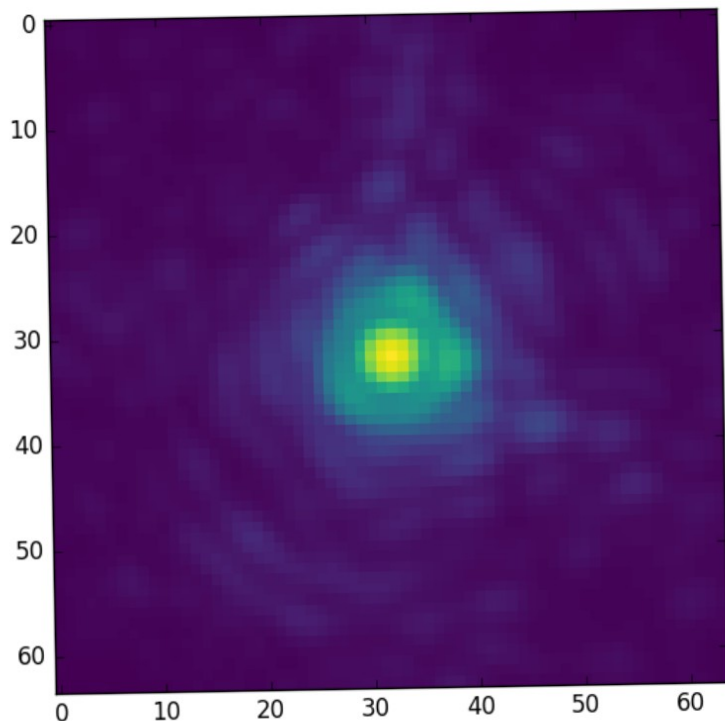
→ Toolkit includes viewers, data logger, low-latency TCP transfer of streams

cacao RTC: current status & on-sky performance for ExAO

**Supports high frame rate
conventional ExAO
operation**

*SCExAO: 1200 modes corrected at 3.5 kHz
(input: 120x120 Pyramid image,
output: 2000 actuator DM)*

On-sky visible image (750nm), log scale
(VAMPIRES)



**Supports advanced operation
modes for enhanced science
performance ... Why ?**

Mixed CPU/GPU solution provides ample computing power and also supports advanced operation modes:

- **Model-free predictive control** using machine learning approach → **deeper contrast, fainter stars**
- **On-sky response matrix acquisition** while loop is running (<2 sec to acquire 2000-mode response matrix) → improved calibration → **higher performance control**
- **Real-time links between control loops, sensor fusion** → **speckle control, low-order modes correction**

Main Guiding Principle: data owns IPC

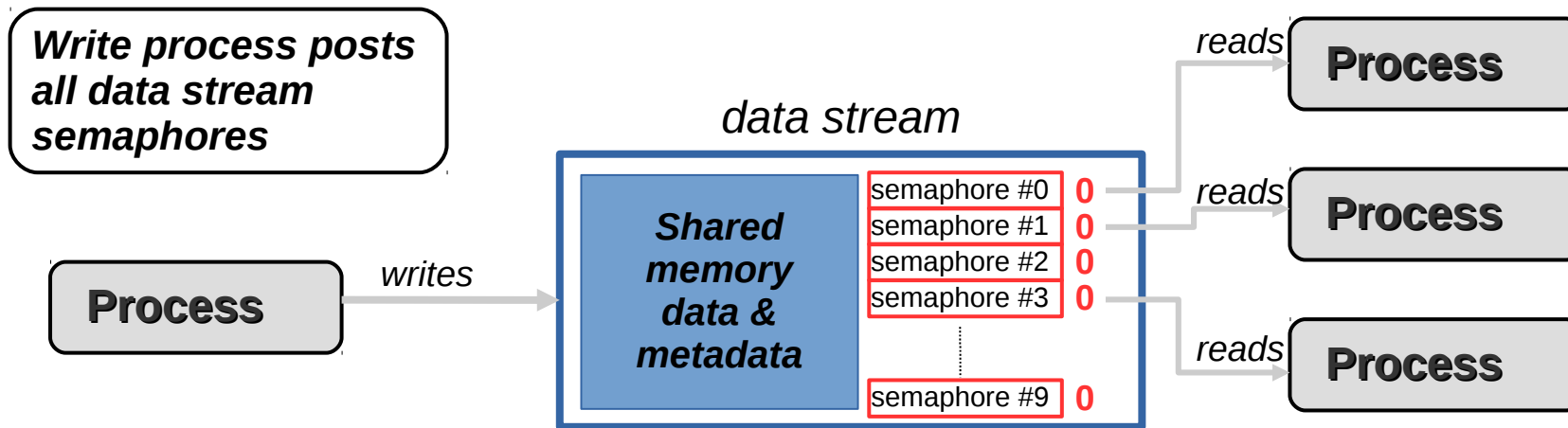
Processes sign up to semaphores

Build AO control loop as a finite set of CPU-managed processes.

Processes can manage computations on GPGPU(s)

Interprocess communication (IPC) is contained in data: cacao streams are held in shared memory and contain semaphores

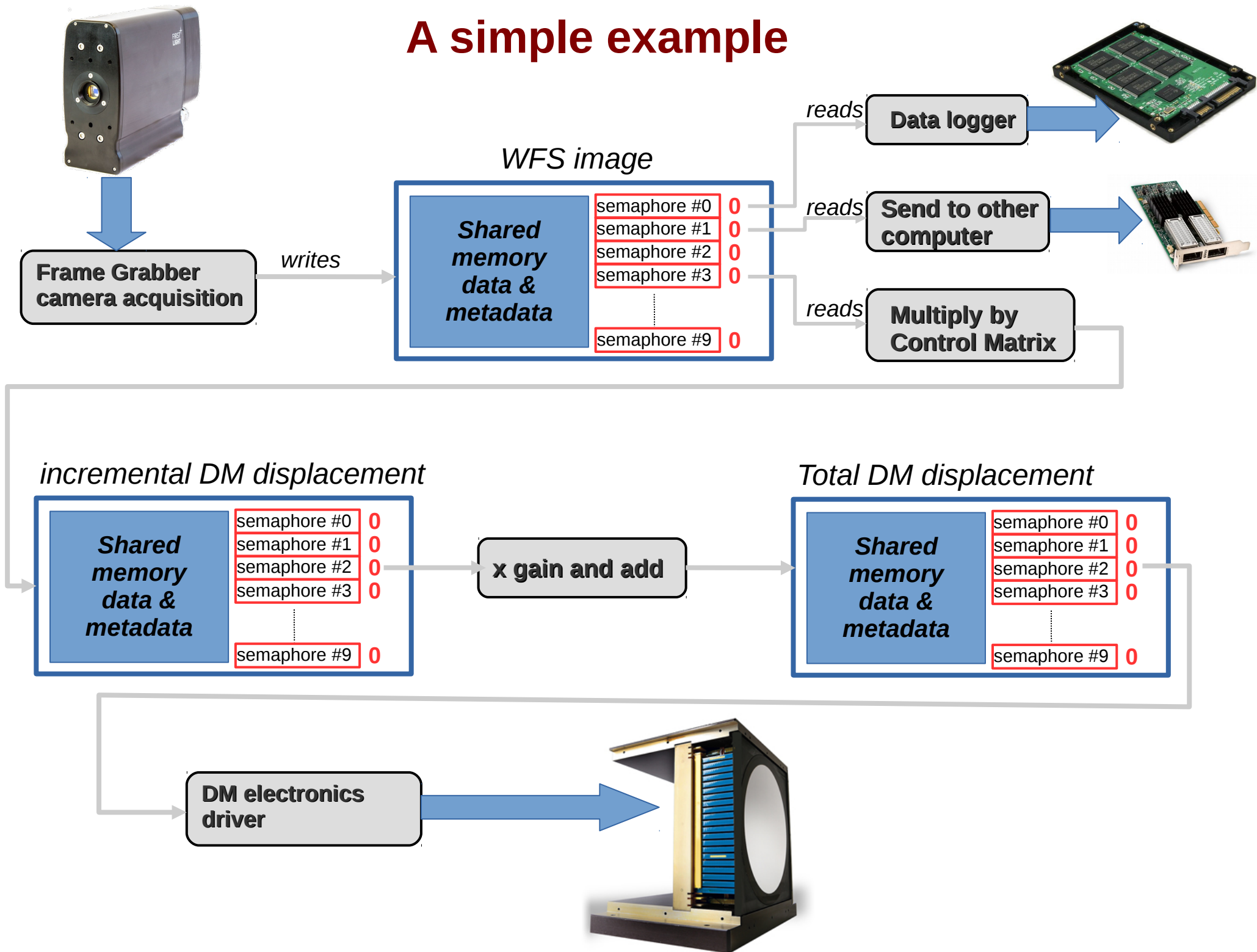
→ Complexities of IPC are handled by compiler and Kernel (semaphores, shared memory)

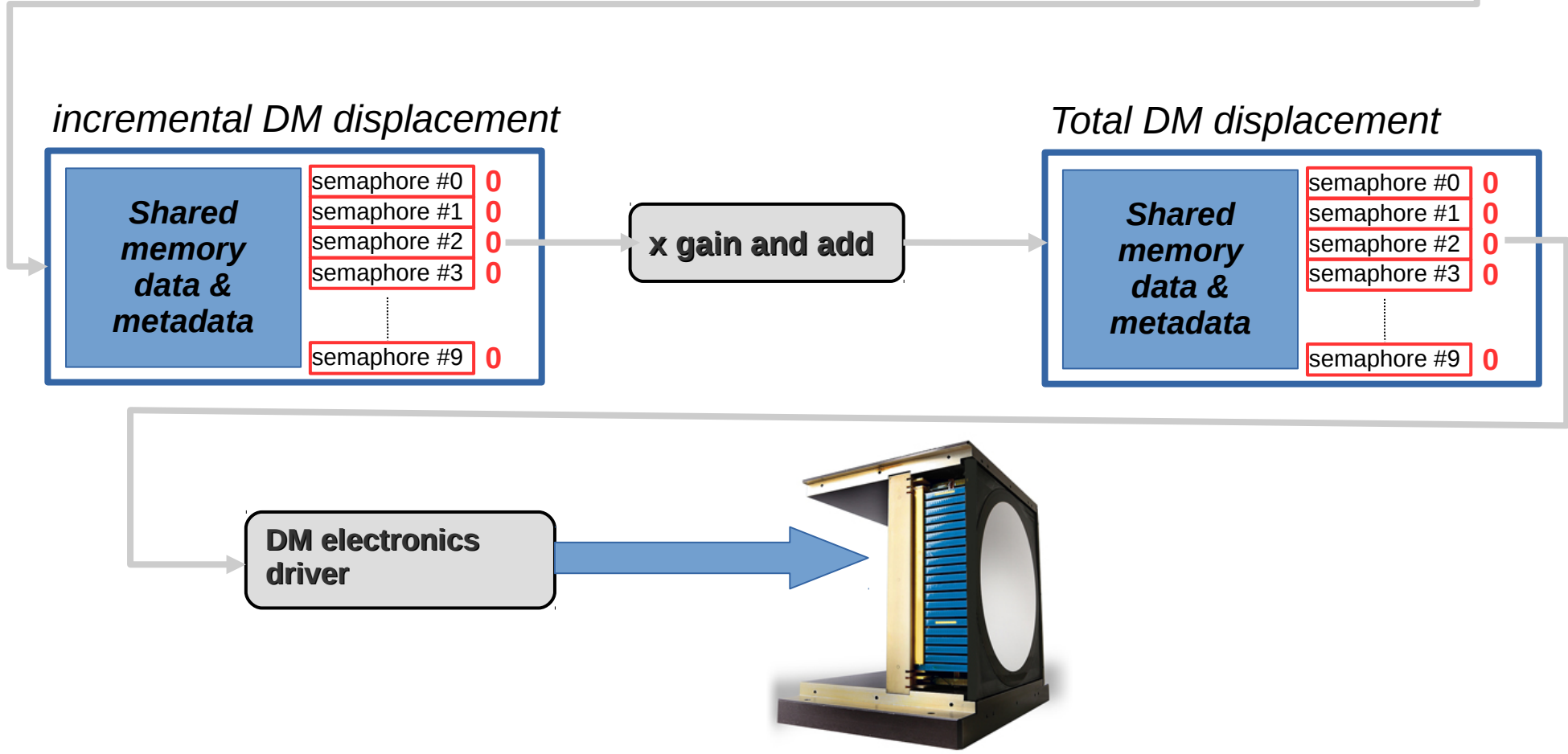
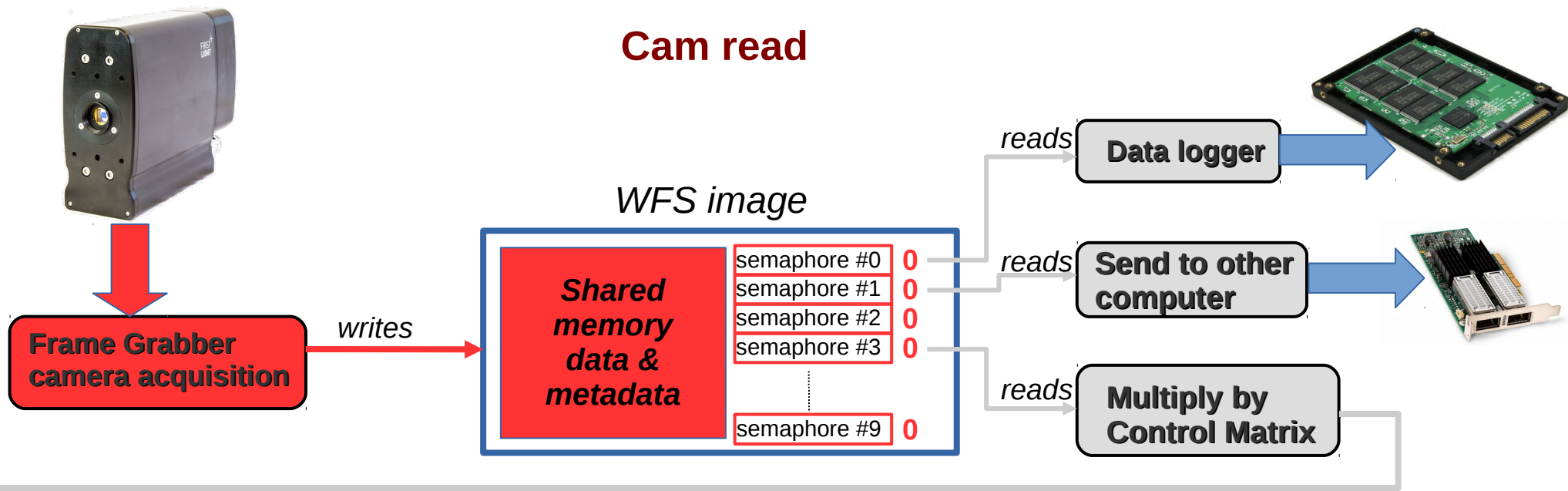


Each of these processes waits on a single semaphore

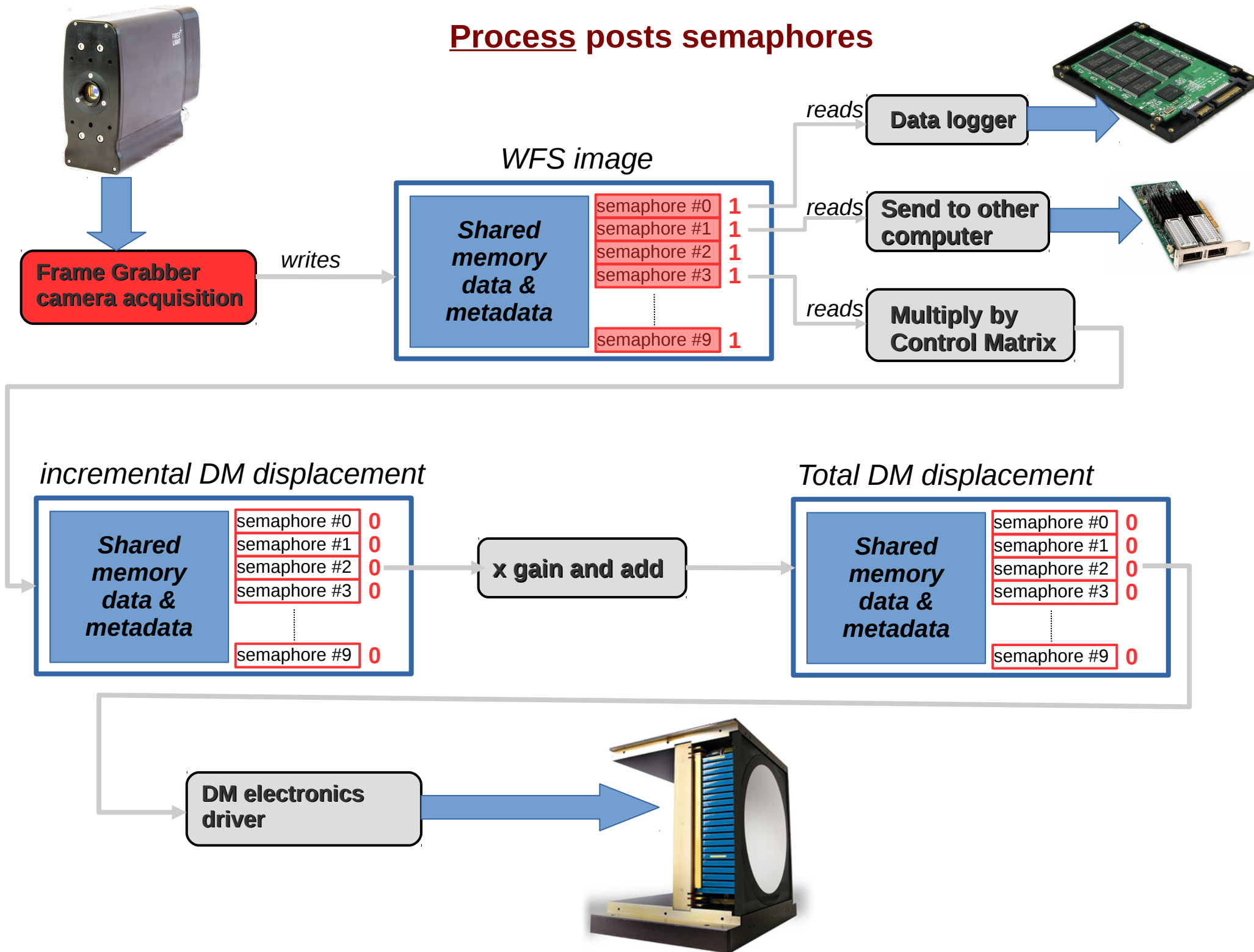
Semaphore-based IPC has us-level latency

A simple example

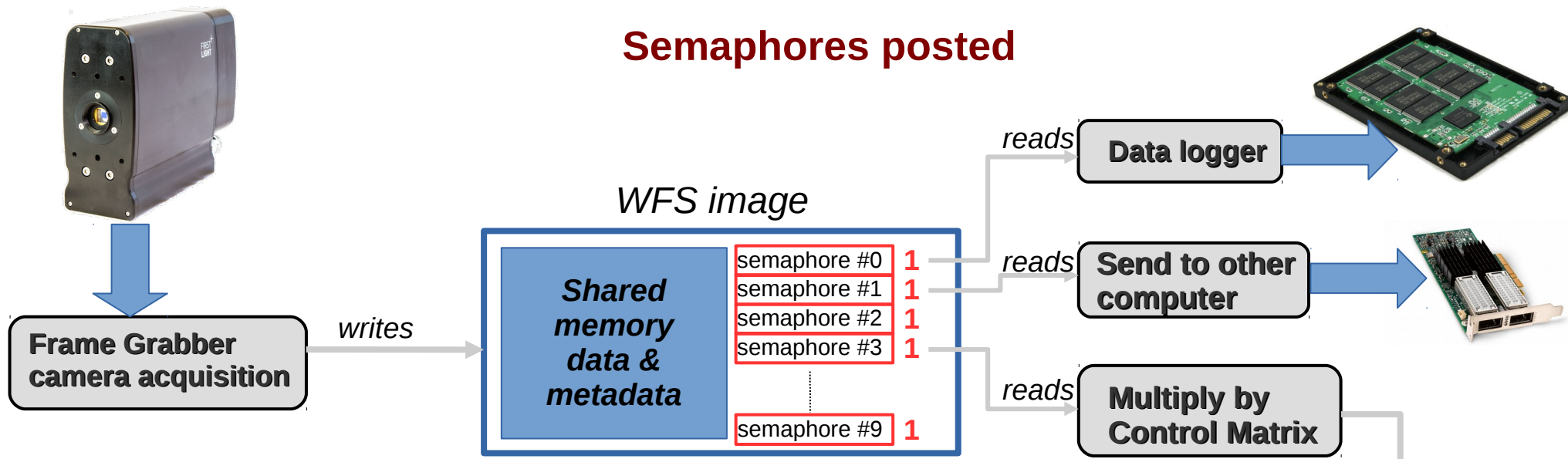




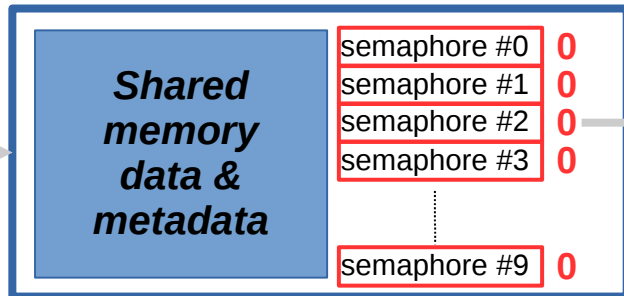
Process posts semaphores



Semaphores posted

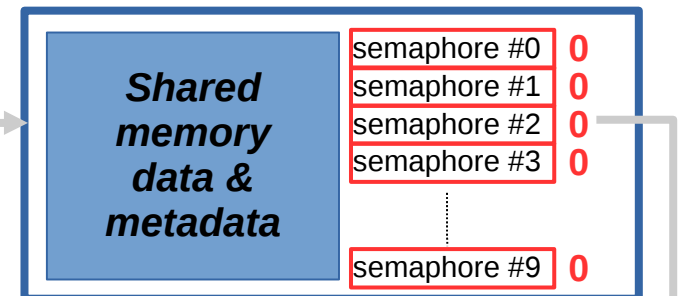


incremental DM displacement

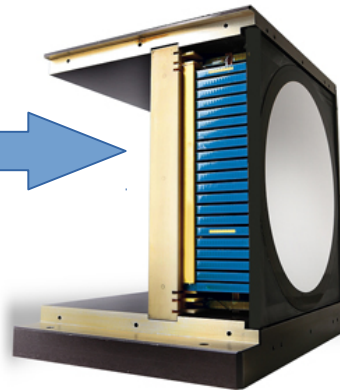


x gain and add

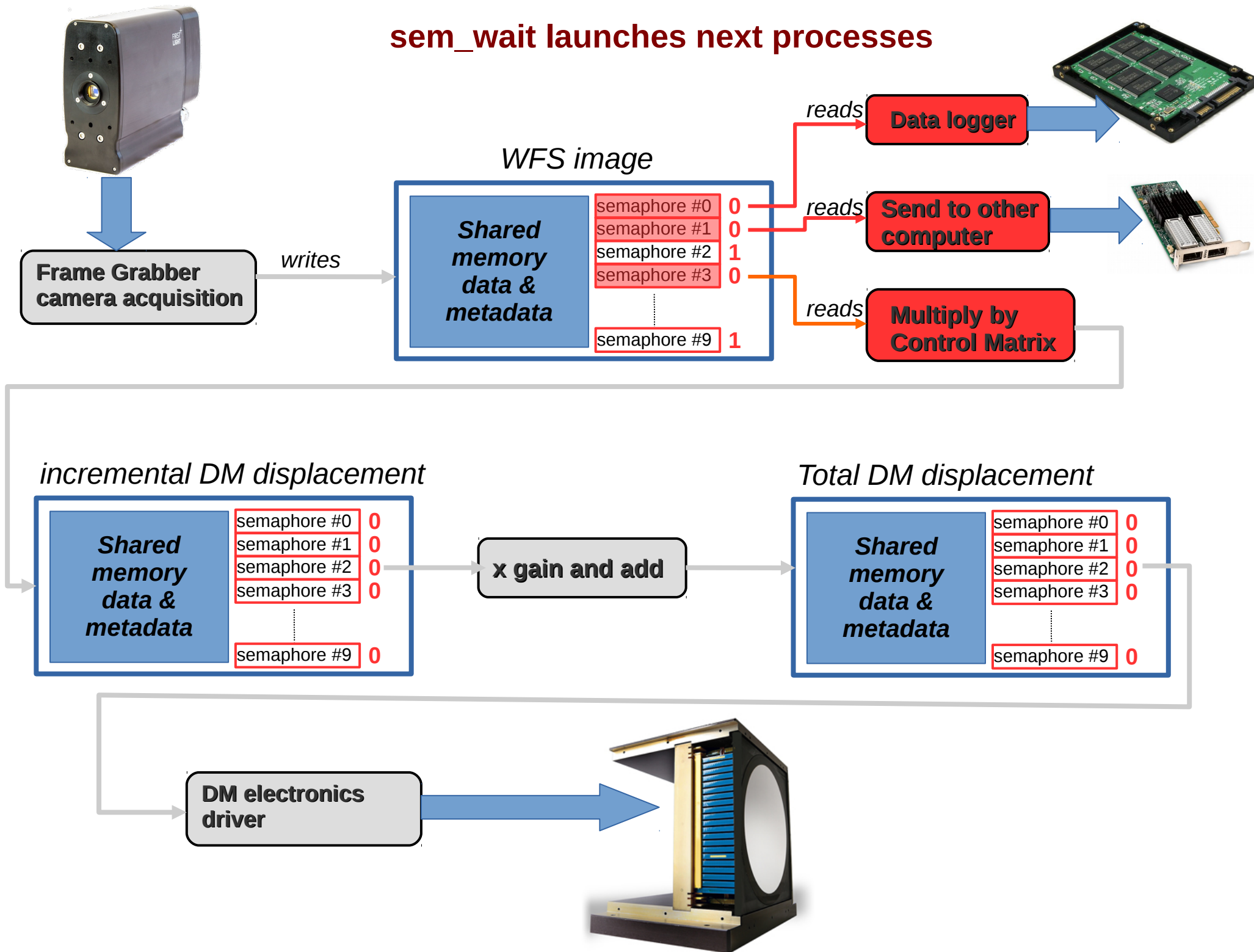
Total DM displacement



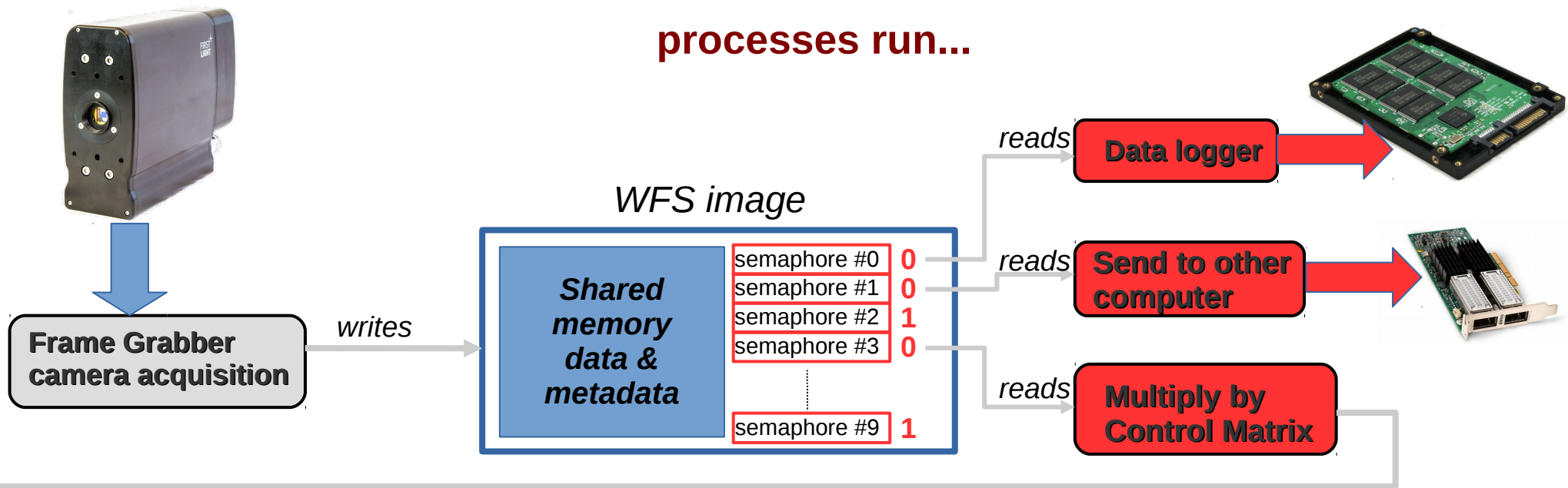
DM electronics driver



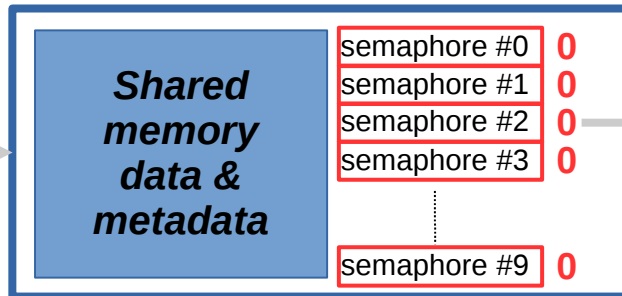
sem_wait launches next processes



processes run...

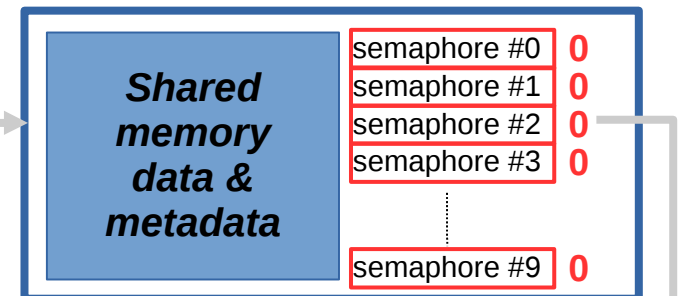


incremental DM displacement

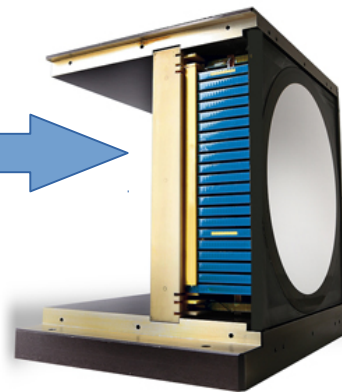


x gain and add

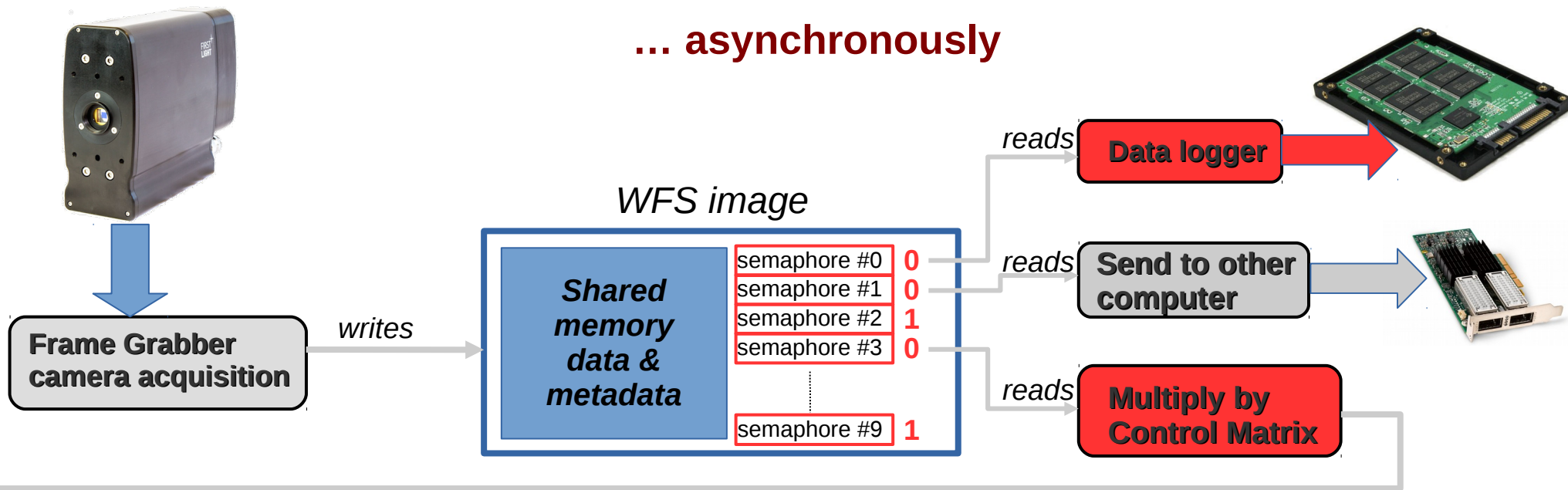
Total DM displacement



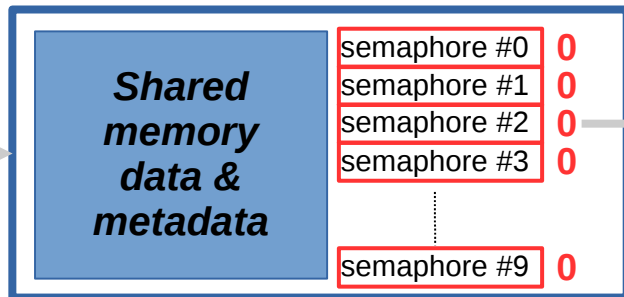
DM electronics driver



... asynchronously

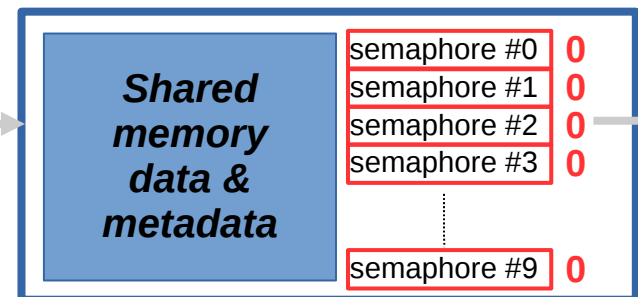


incremental DM displacement

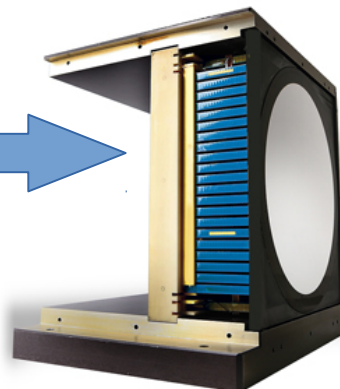


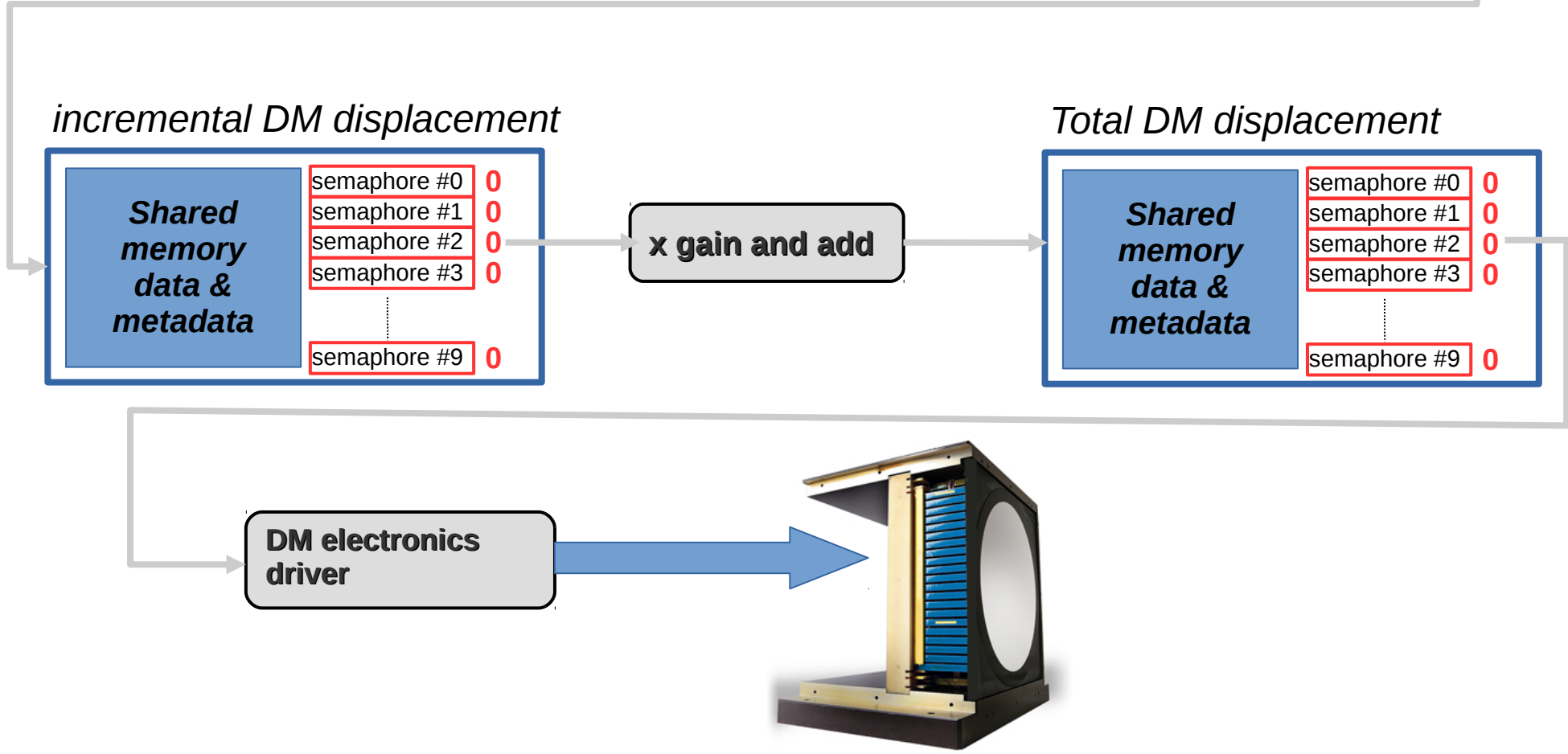
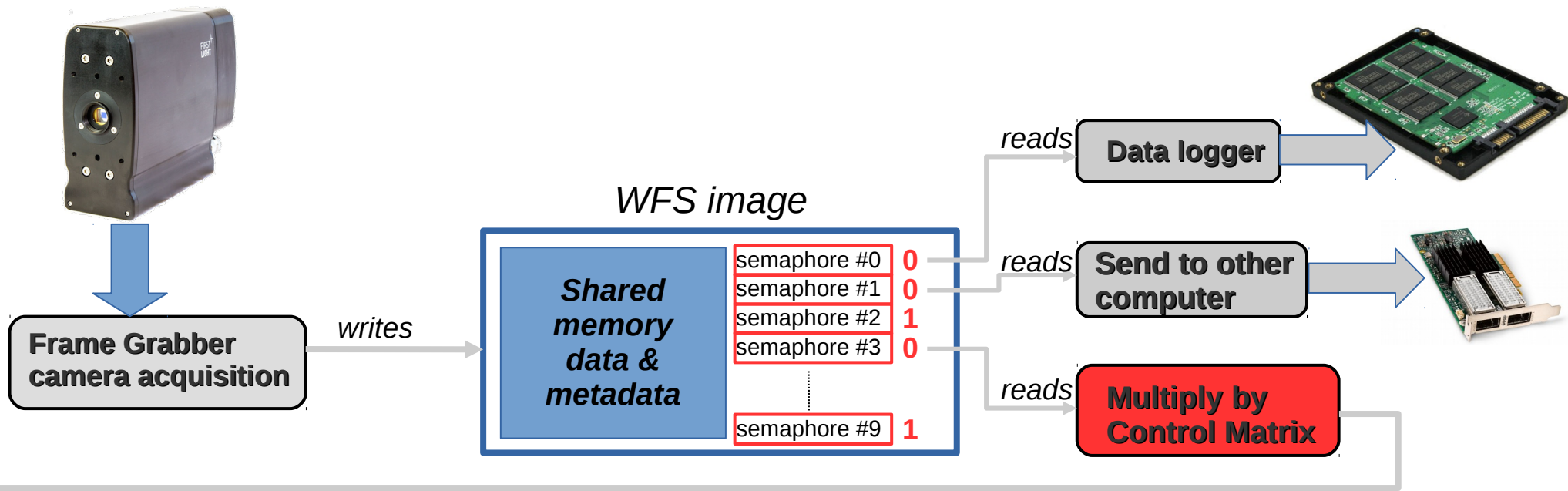
x gain and add

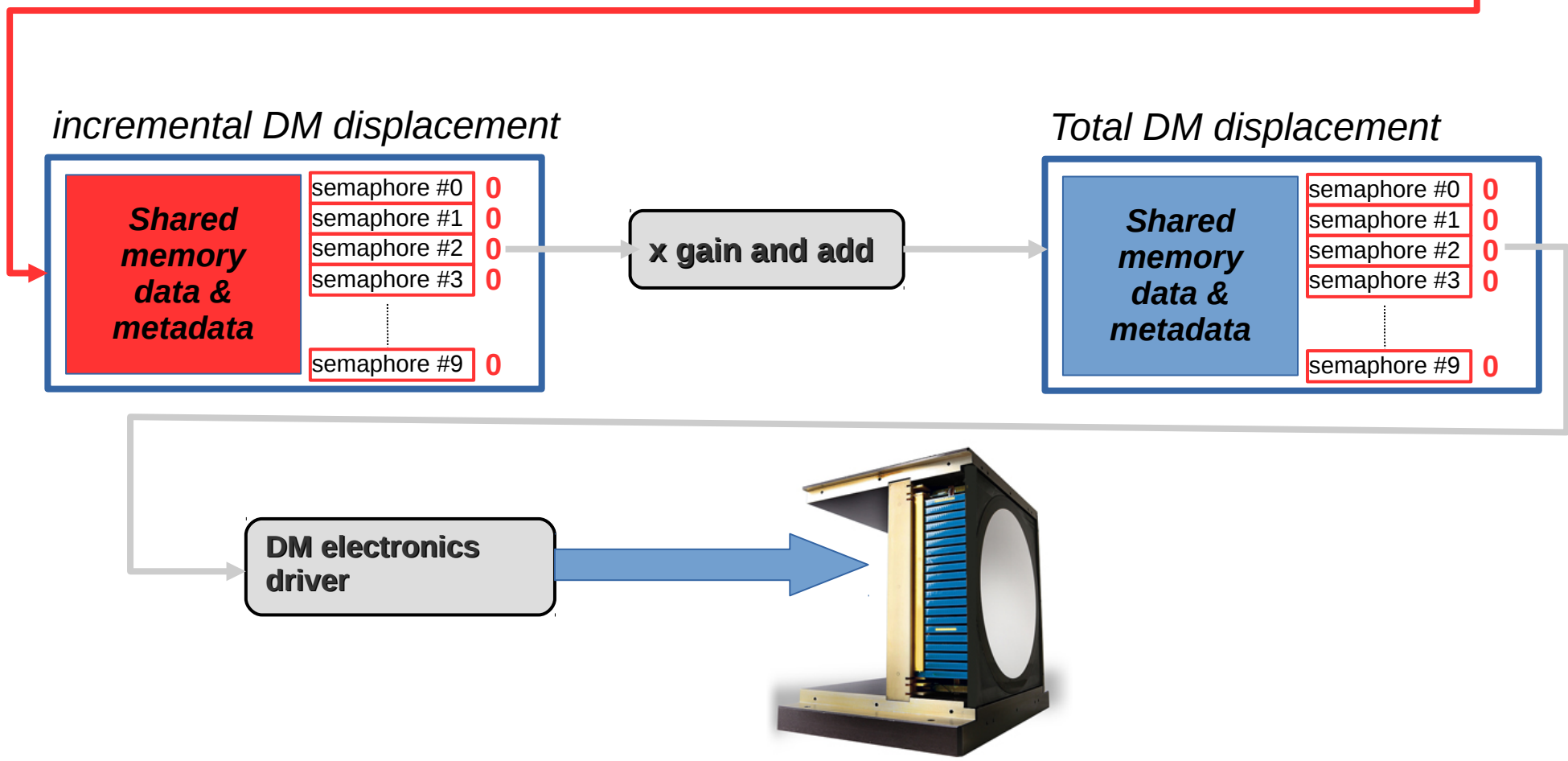
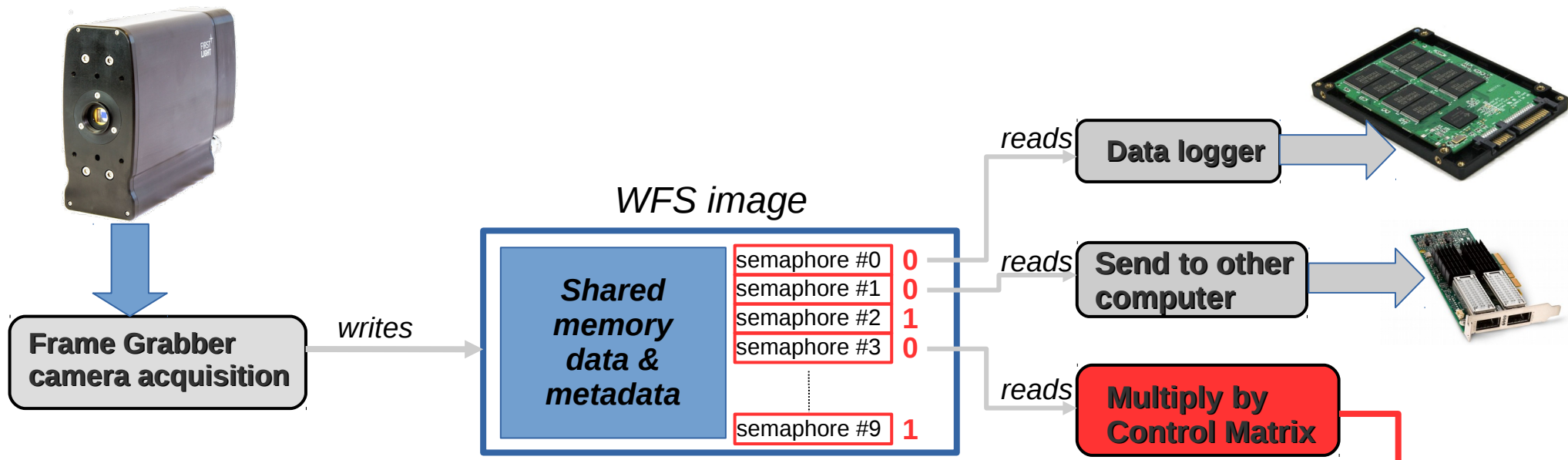
Total DM displacement

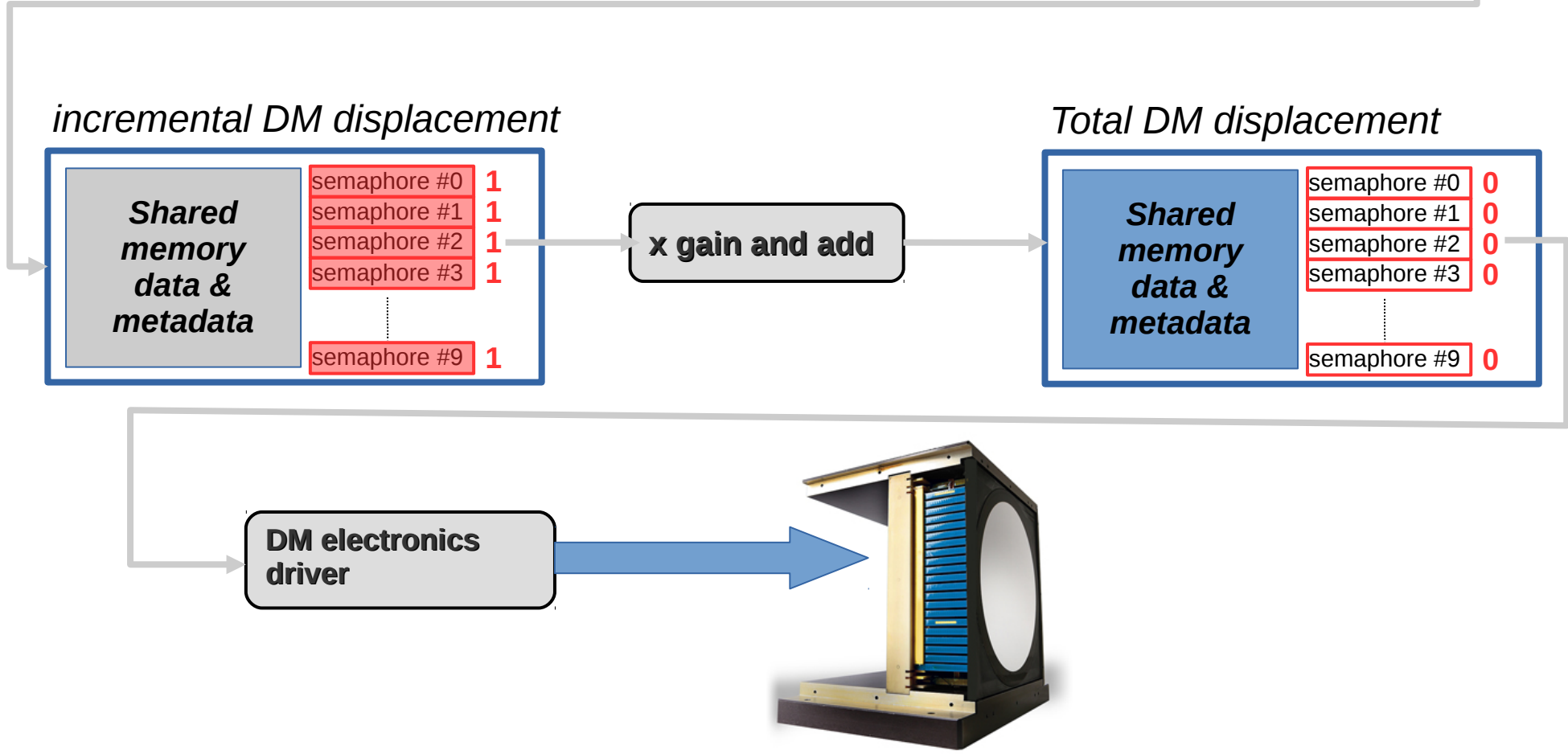
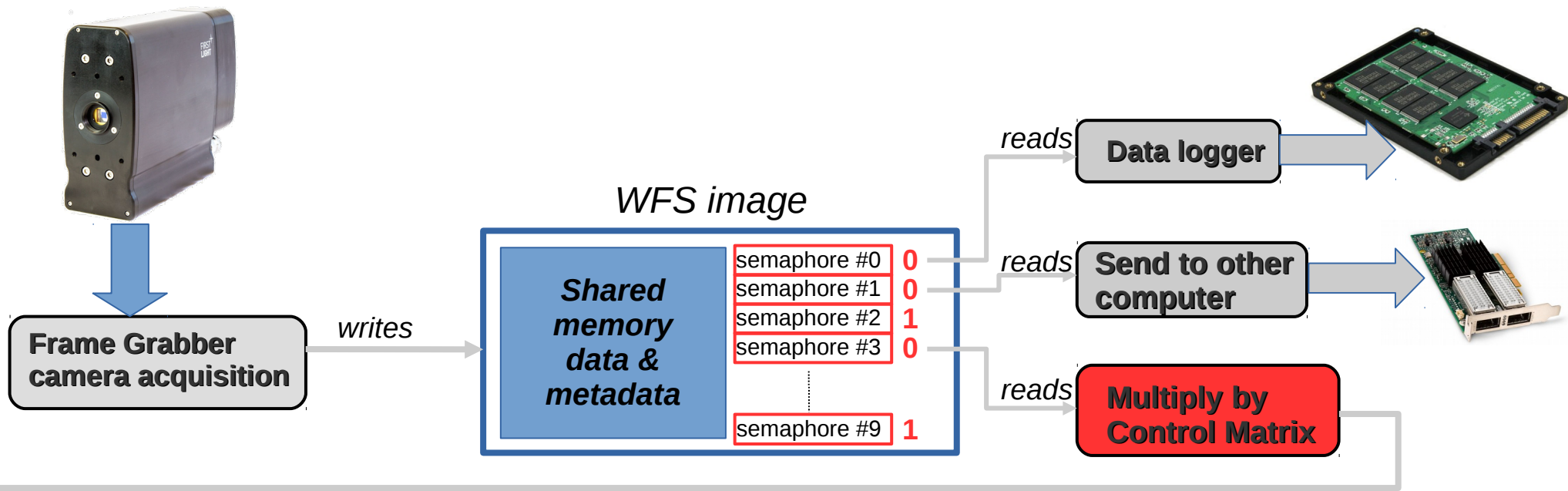


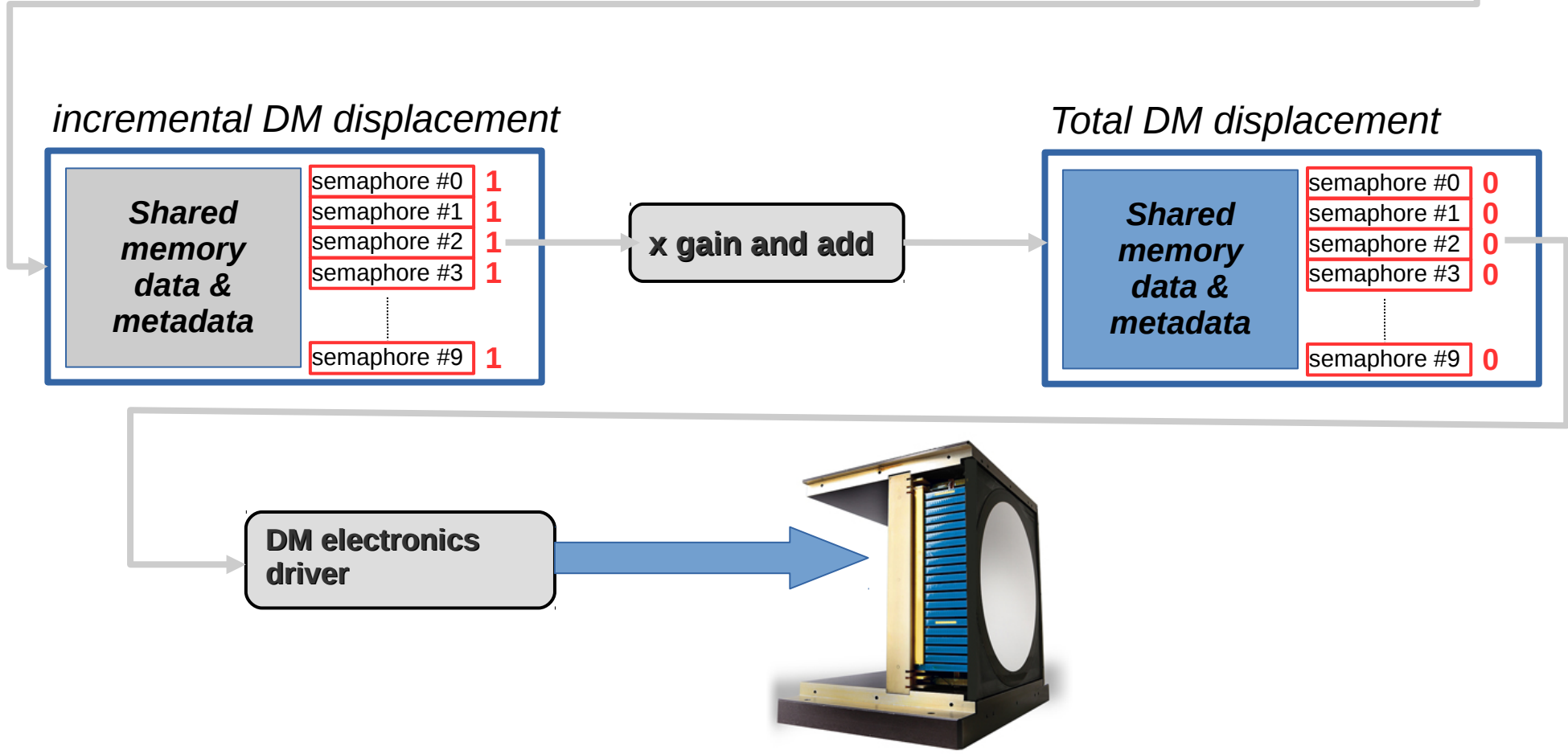
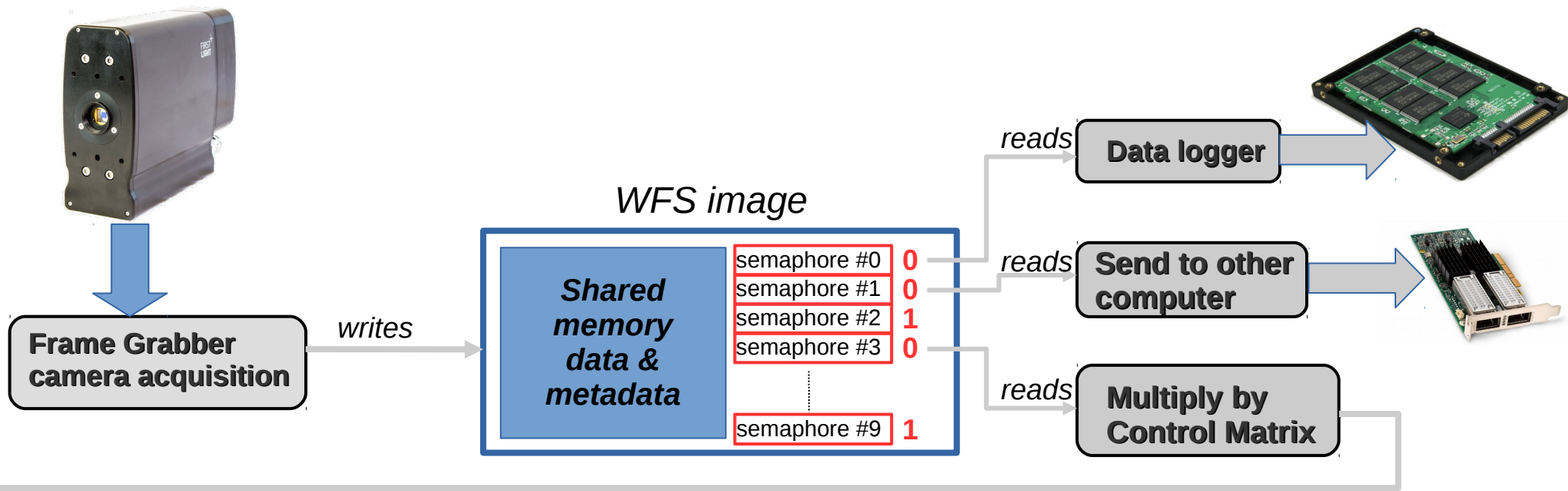
DM electronics driver

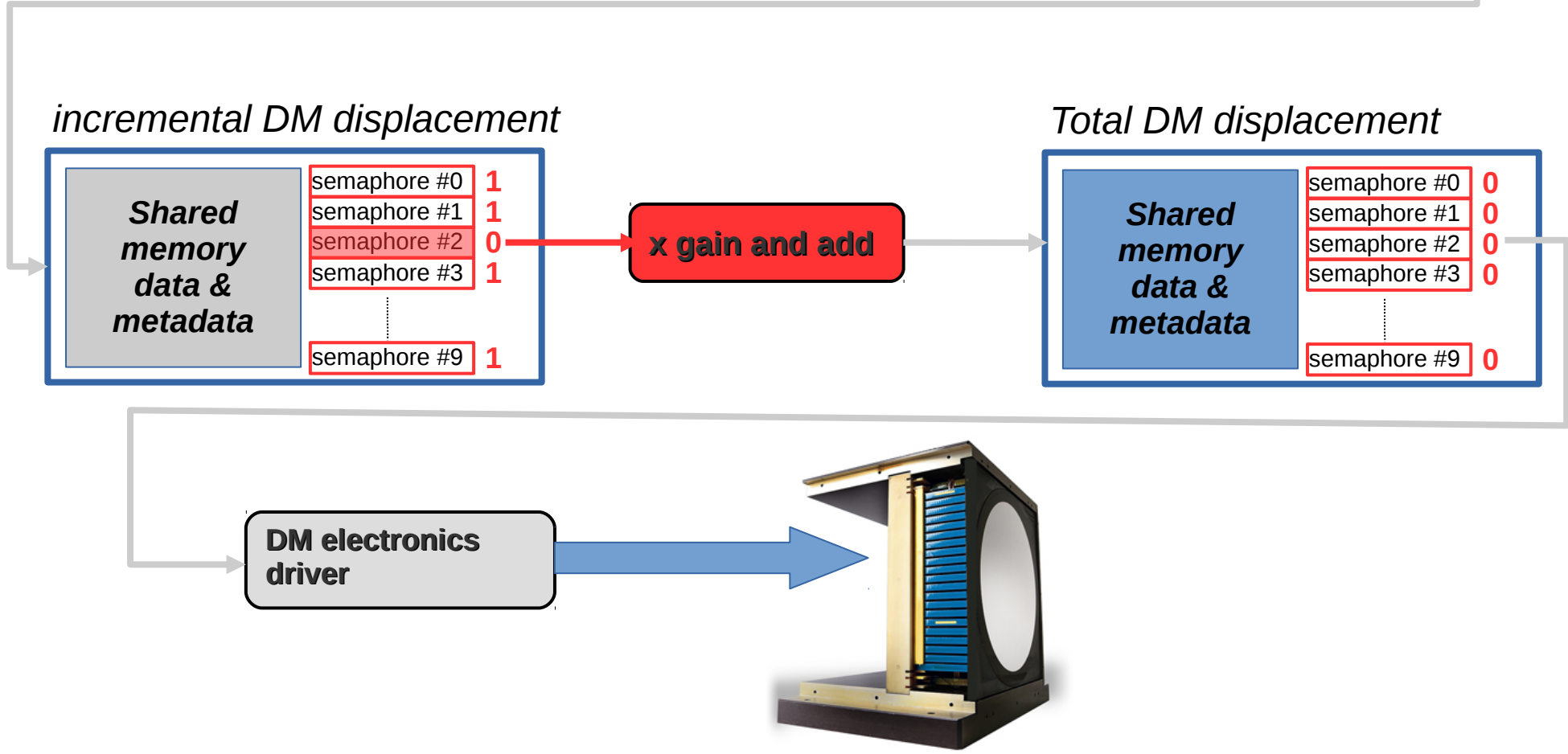
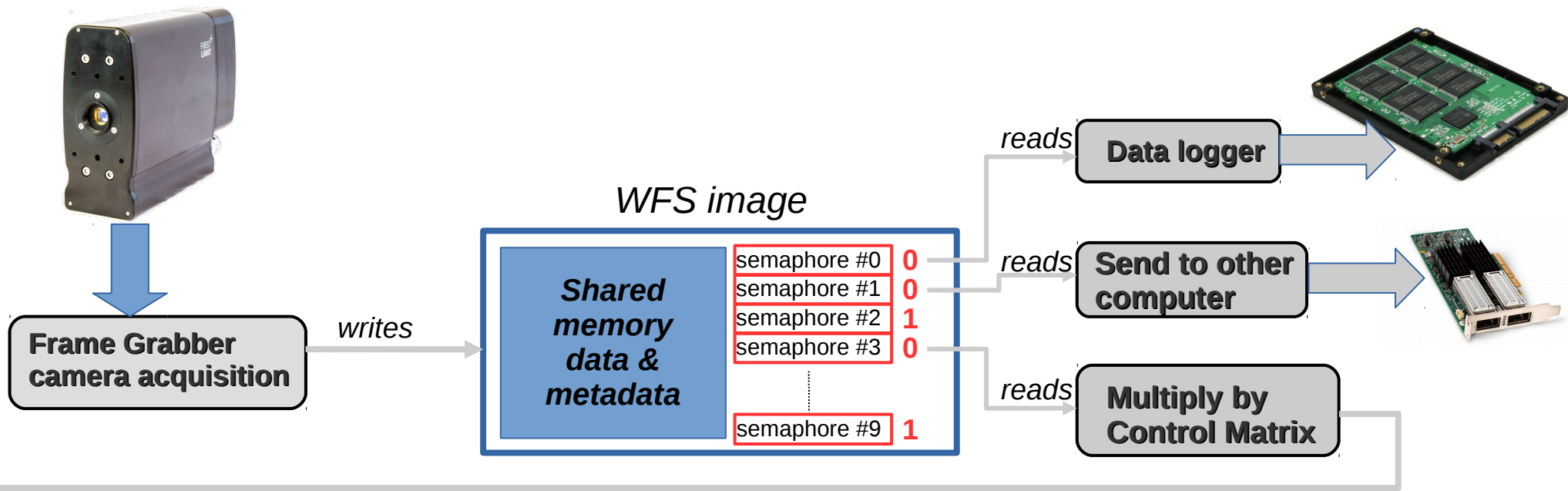


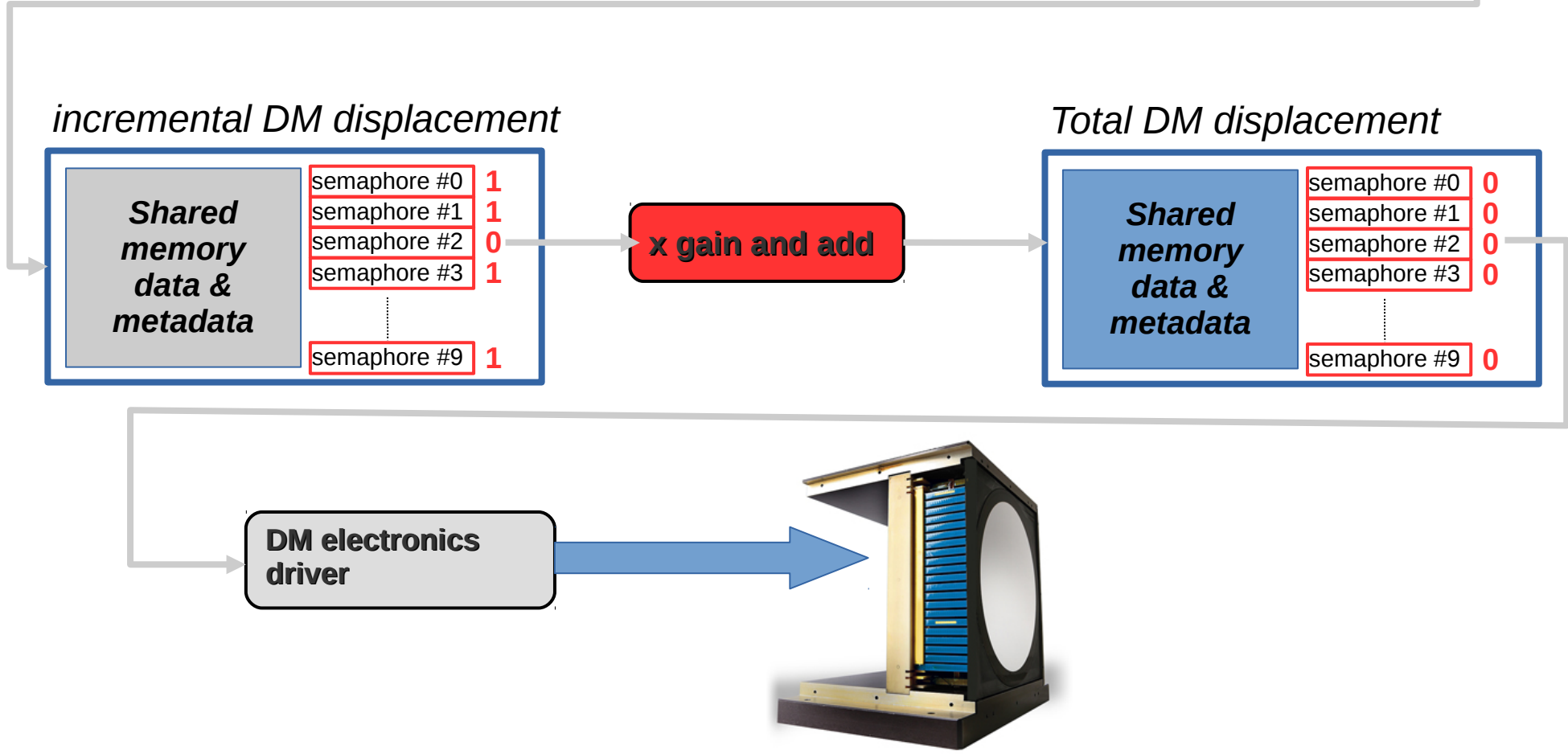
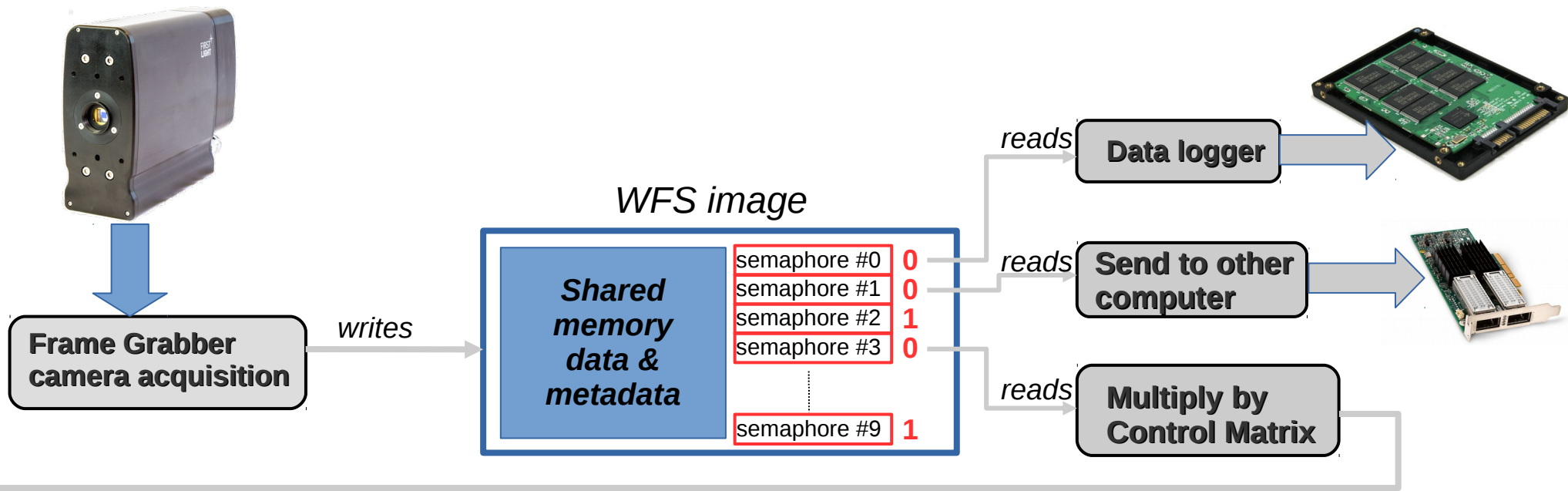


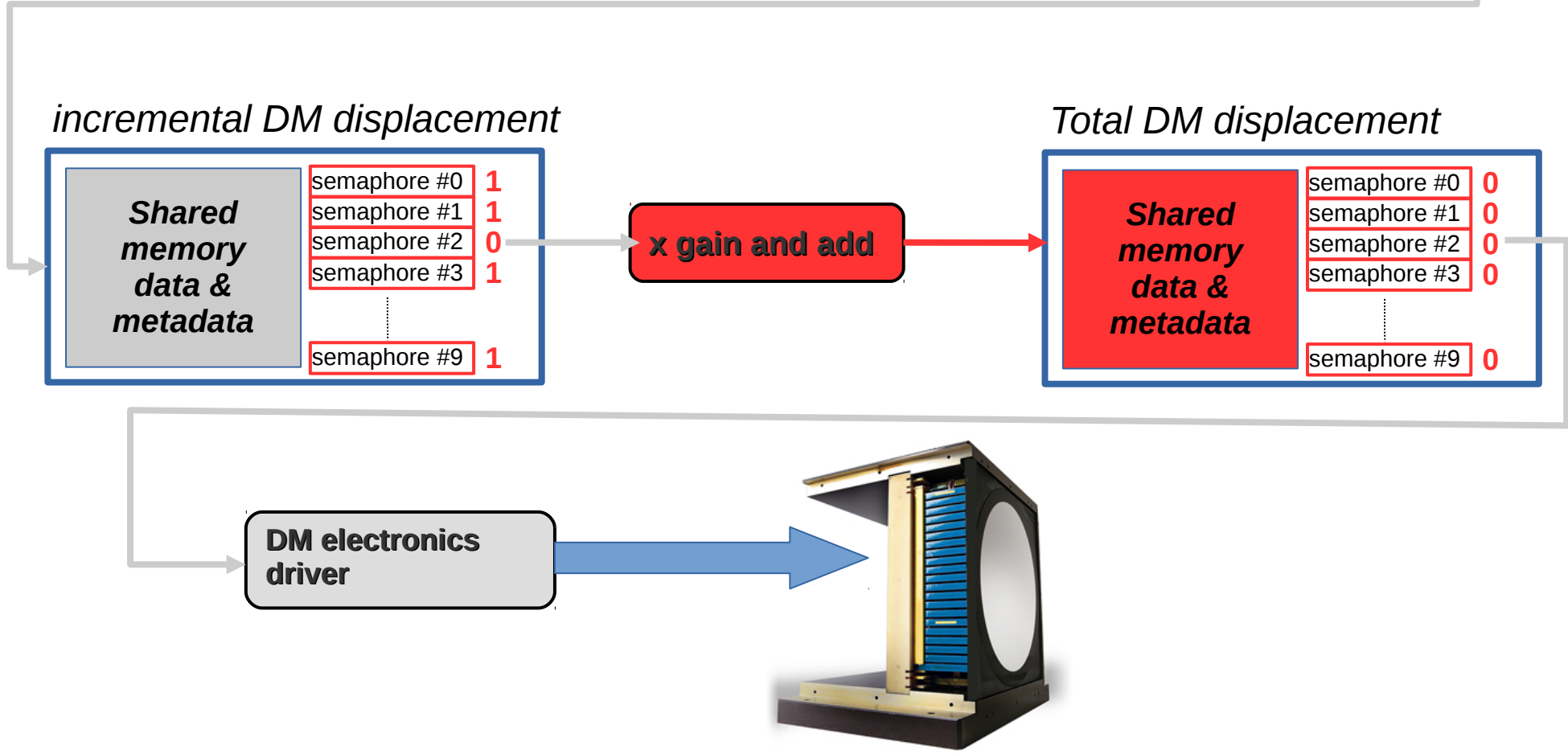
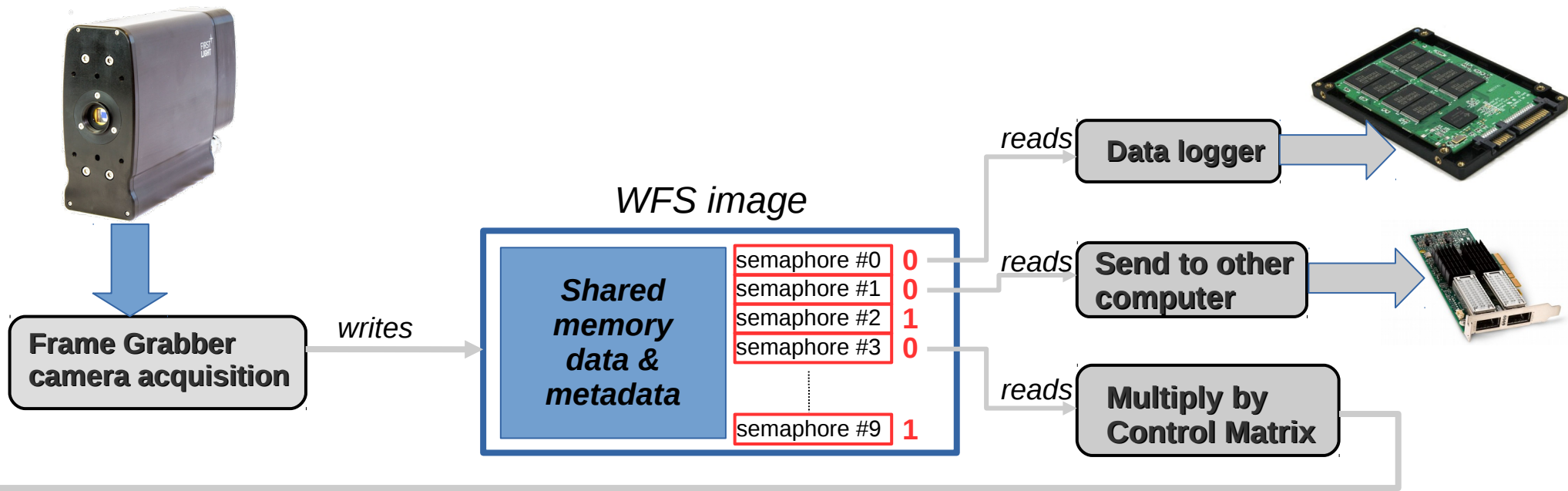


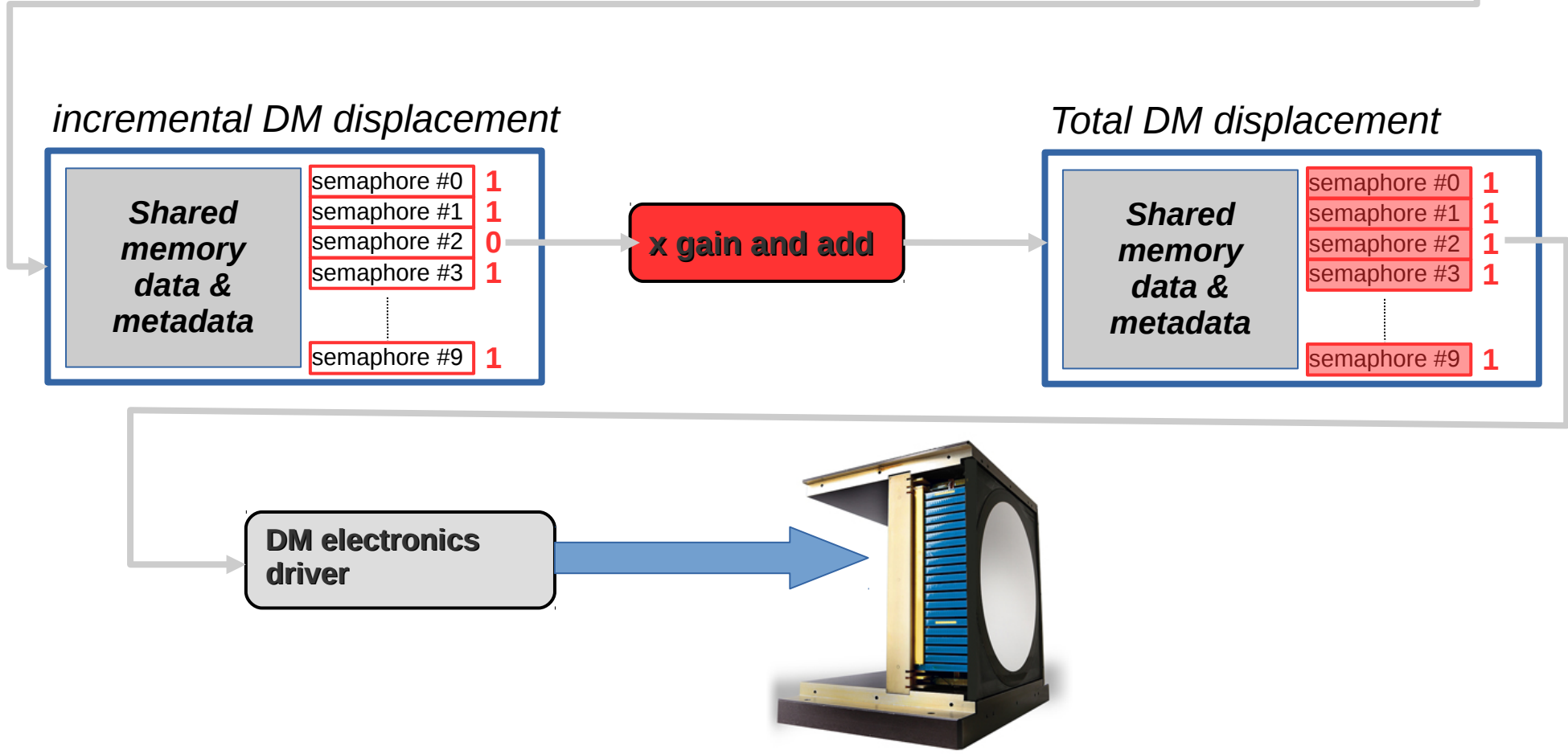
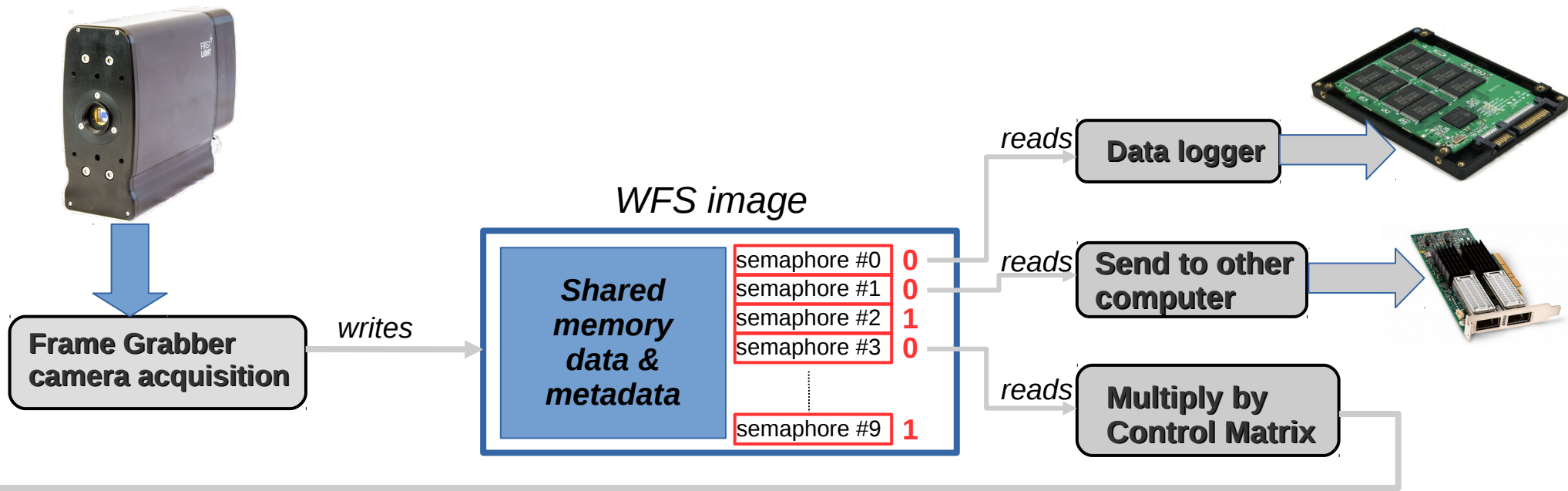


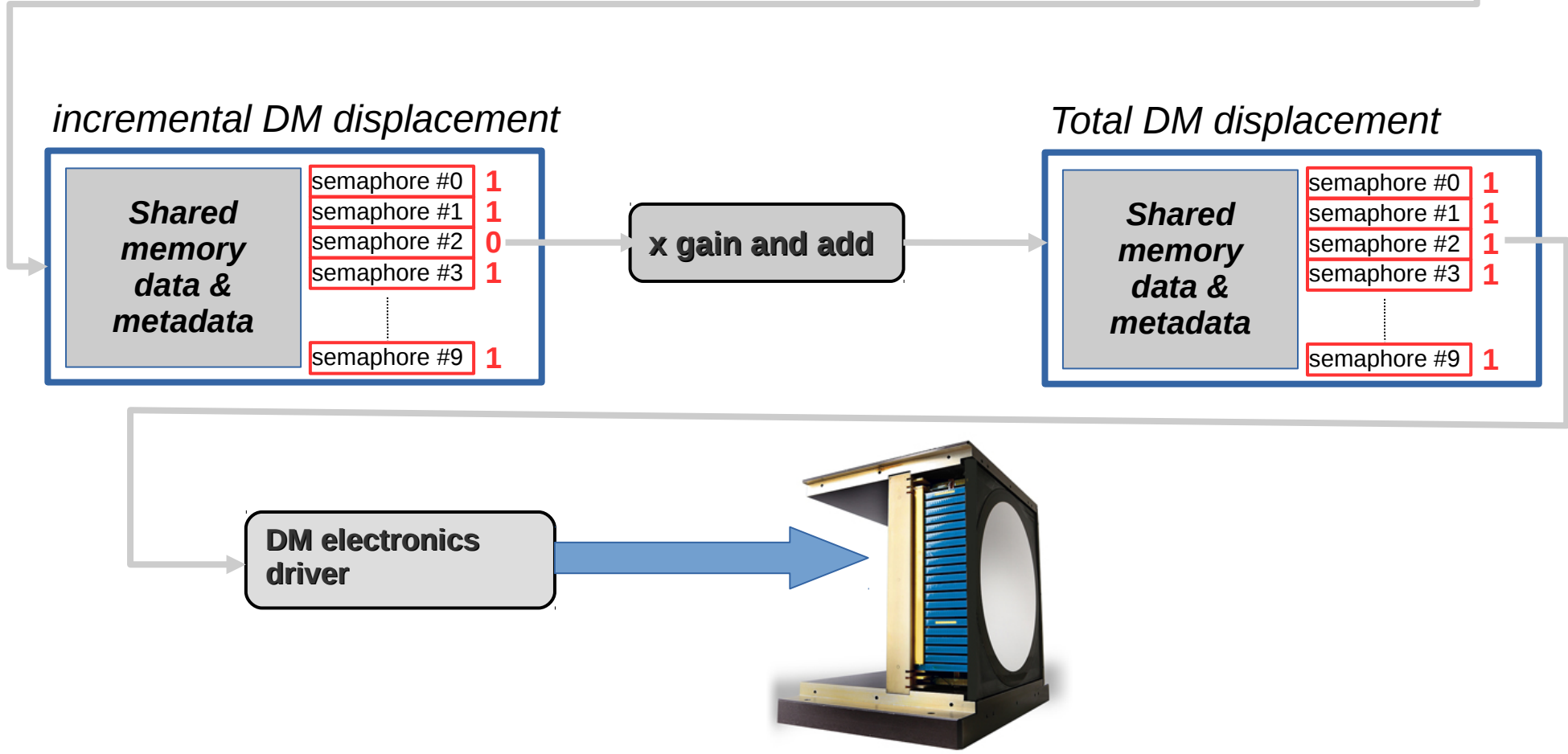
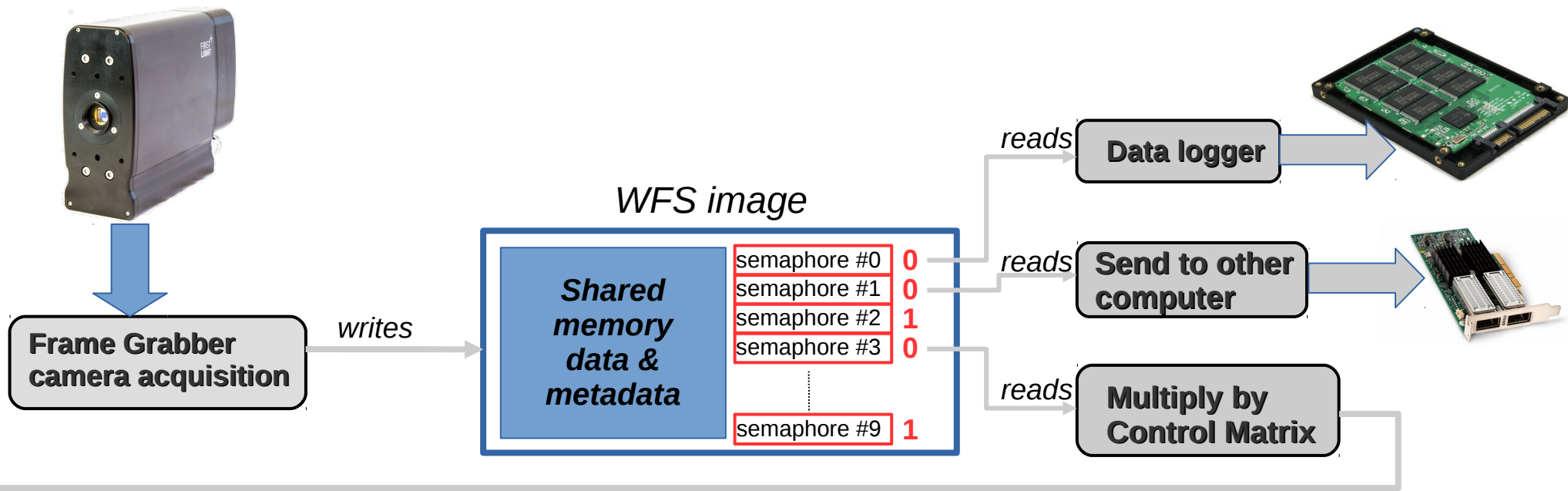


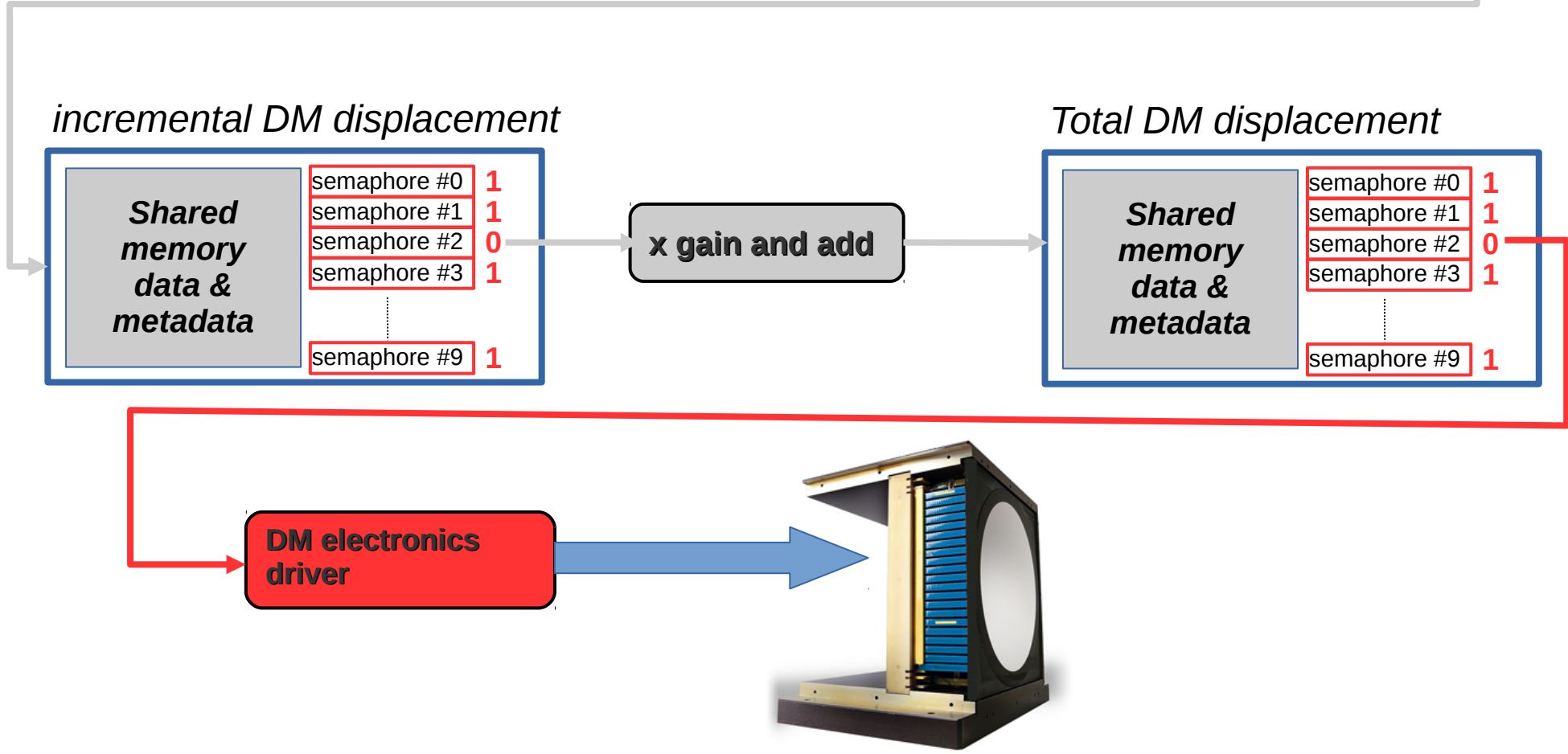
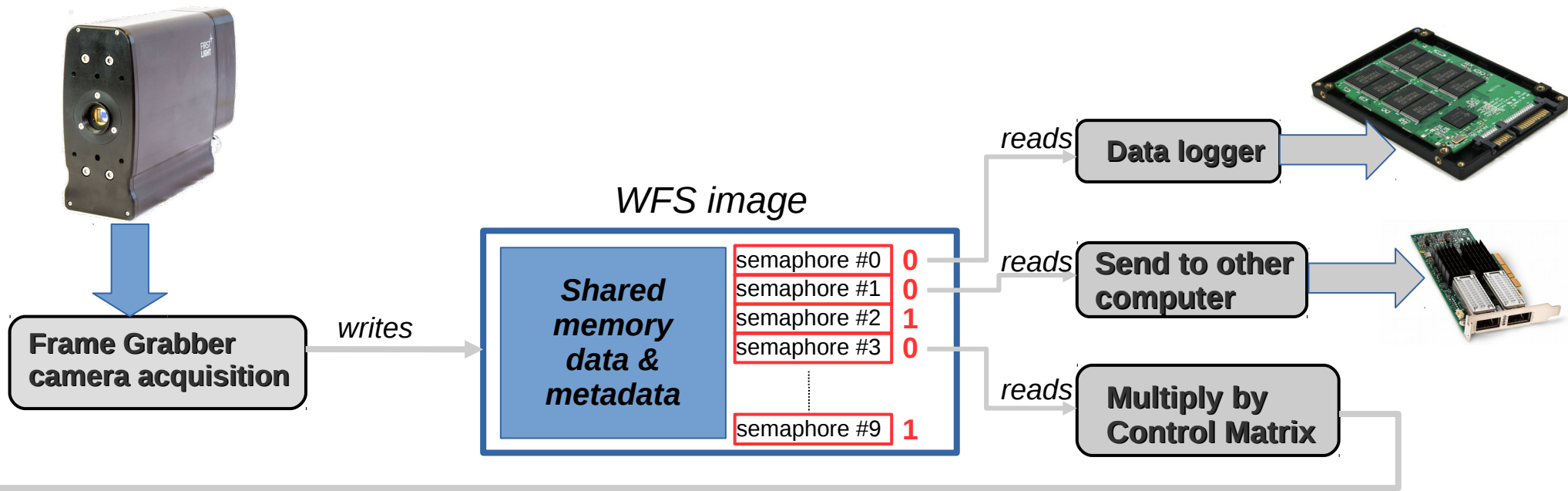


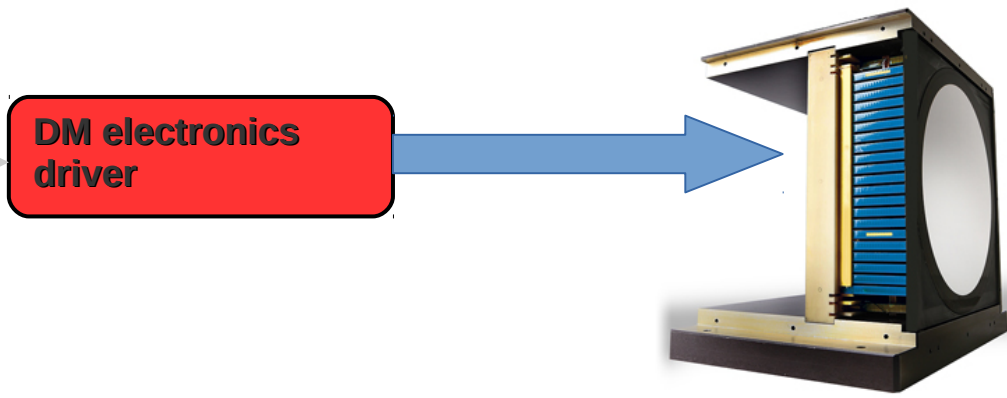
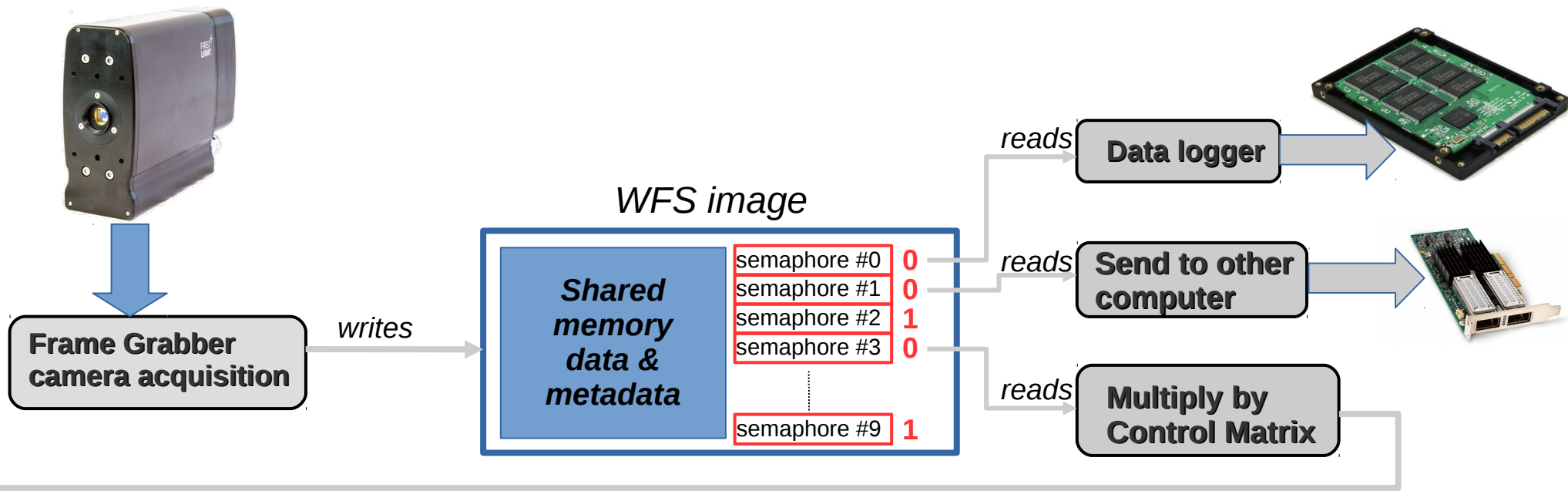




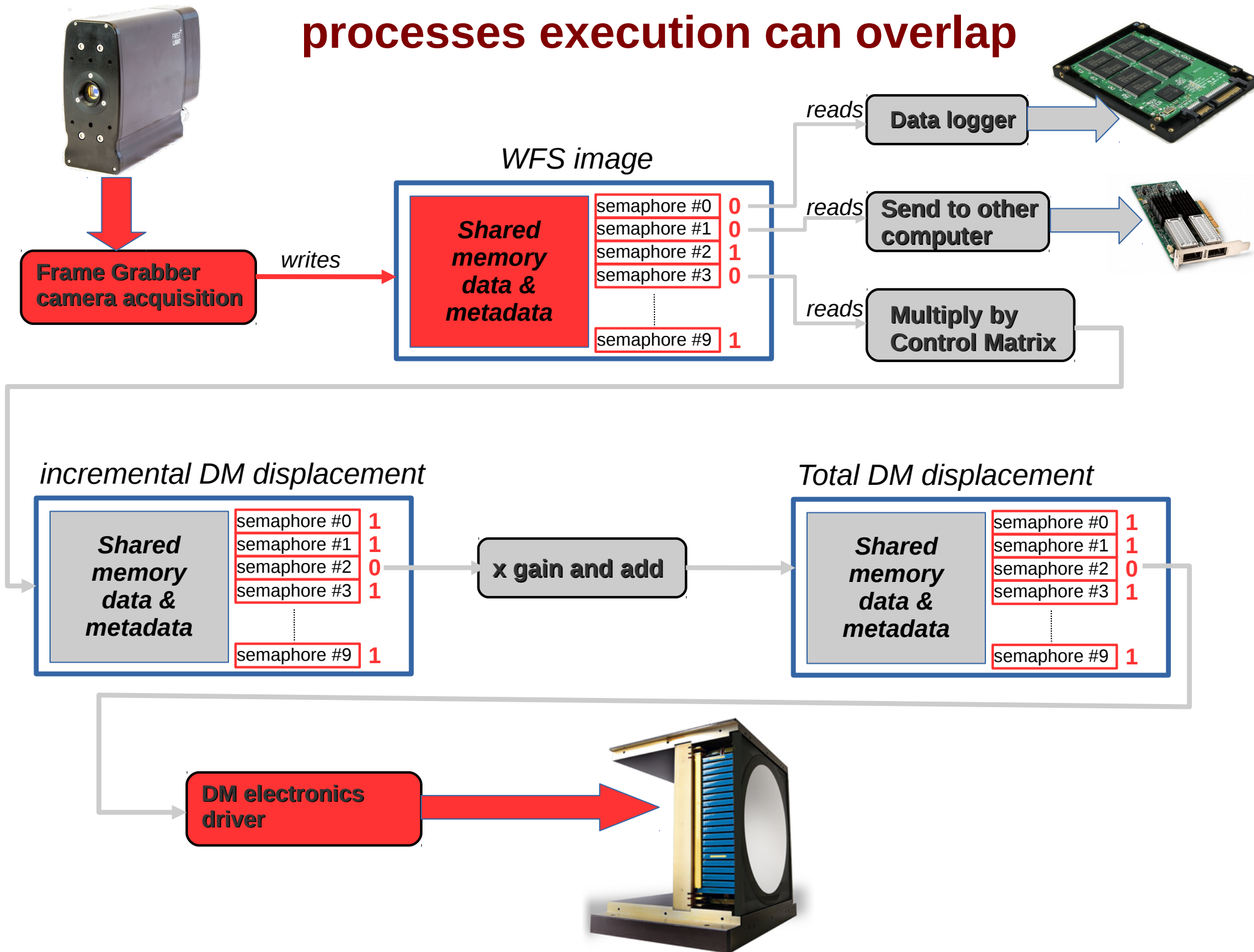


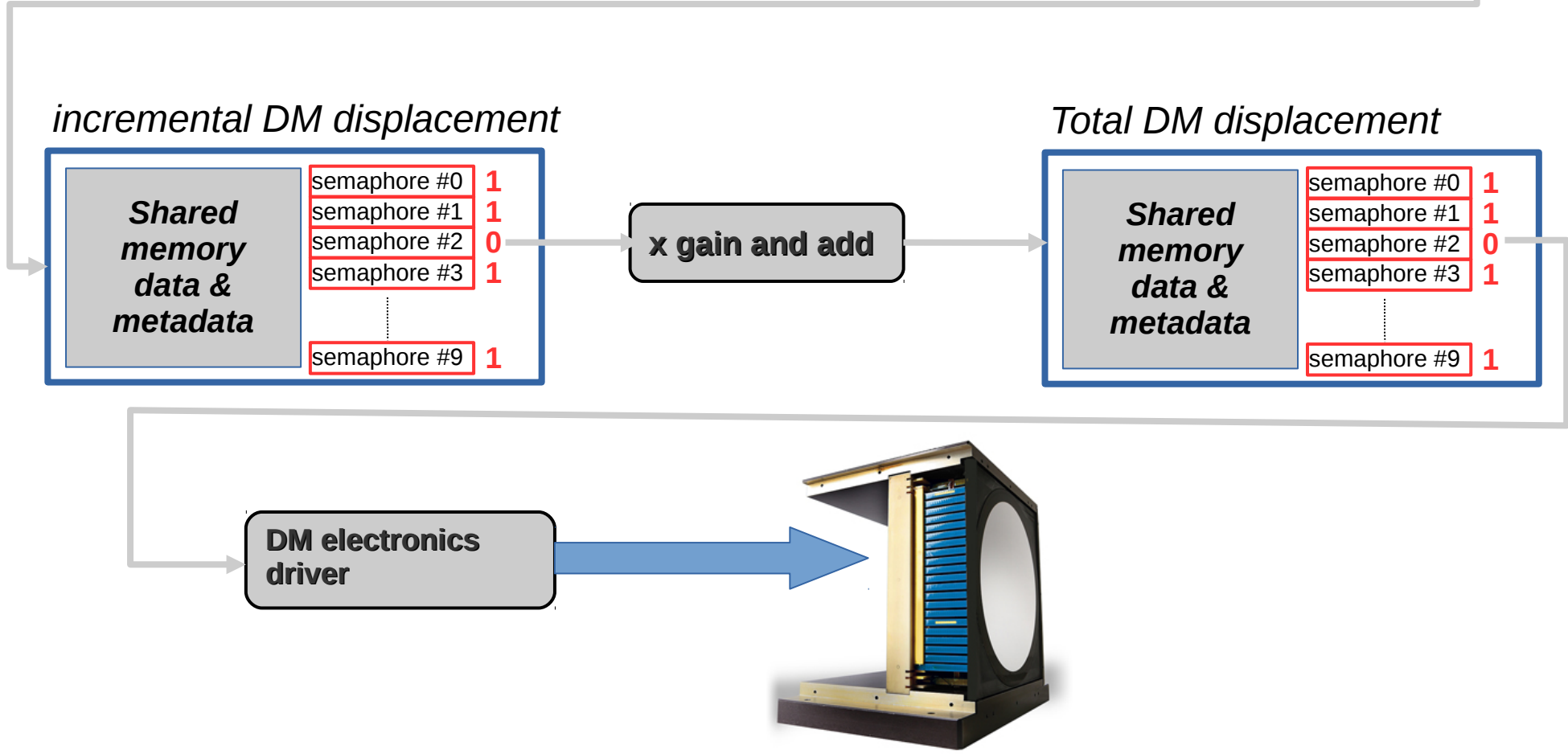
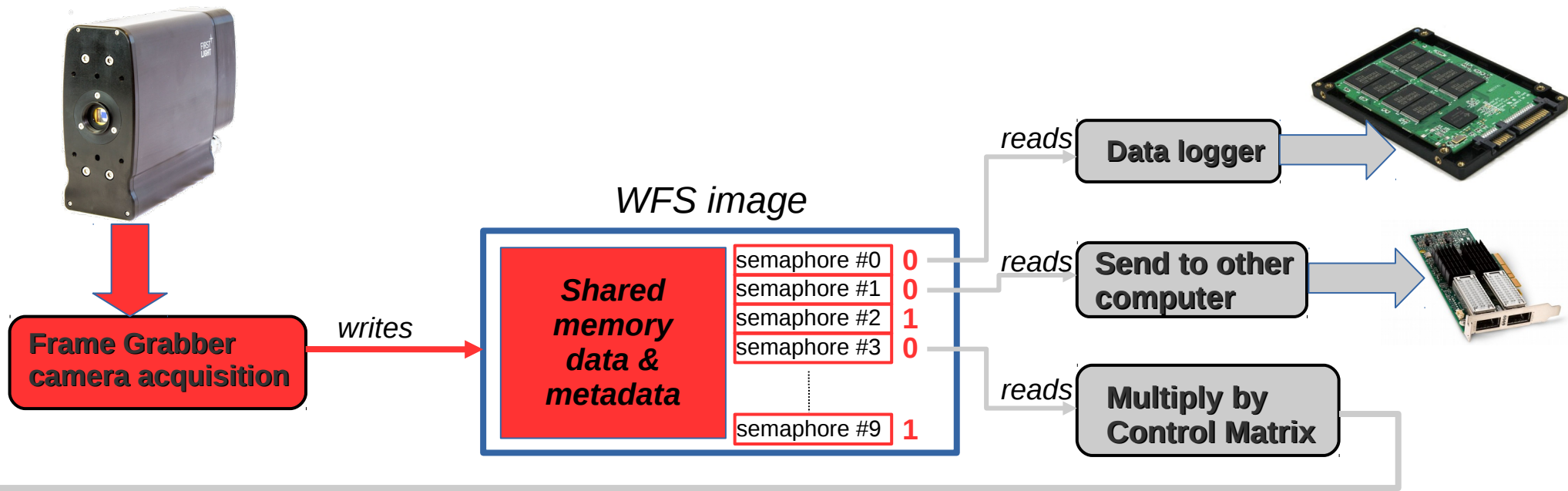




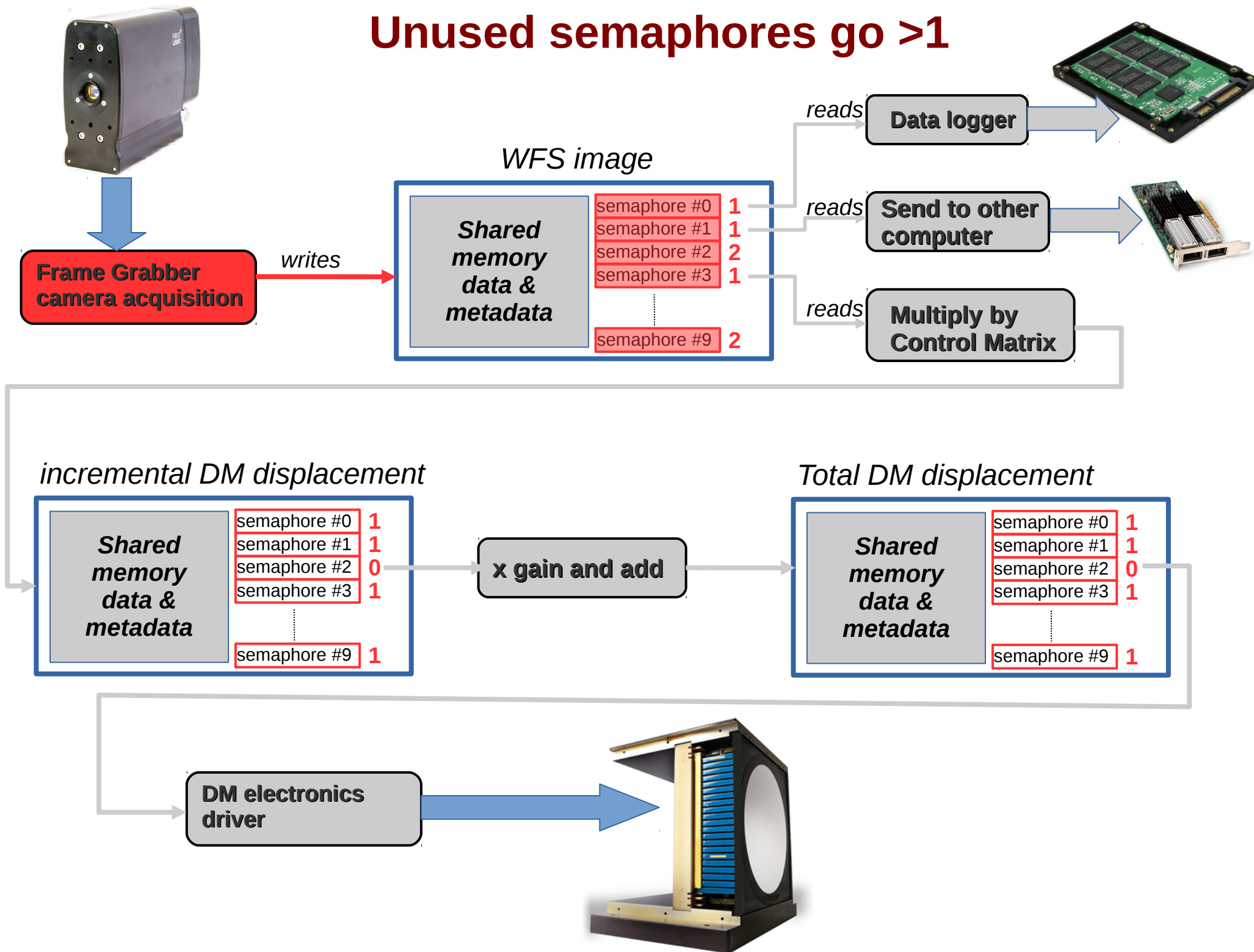


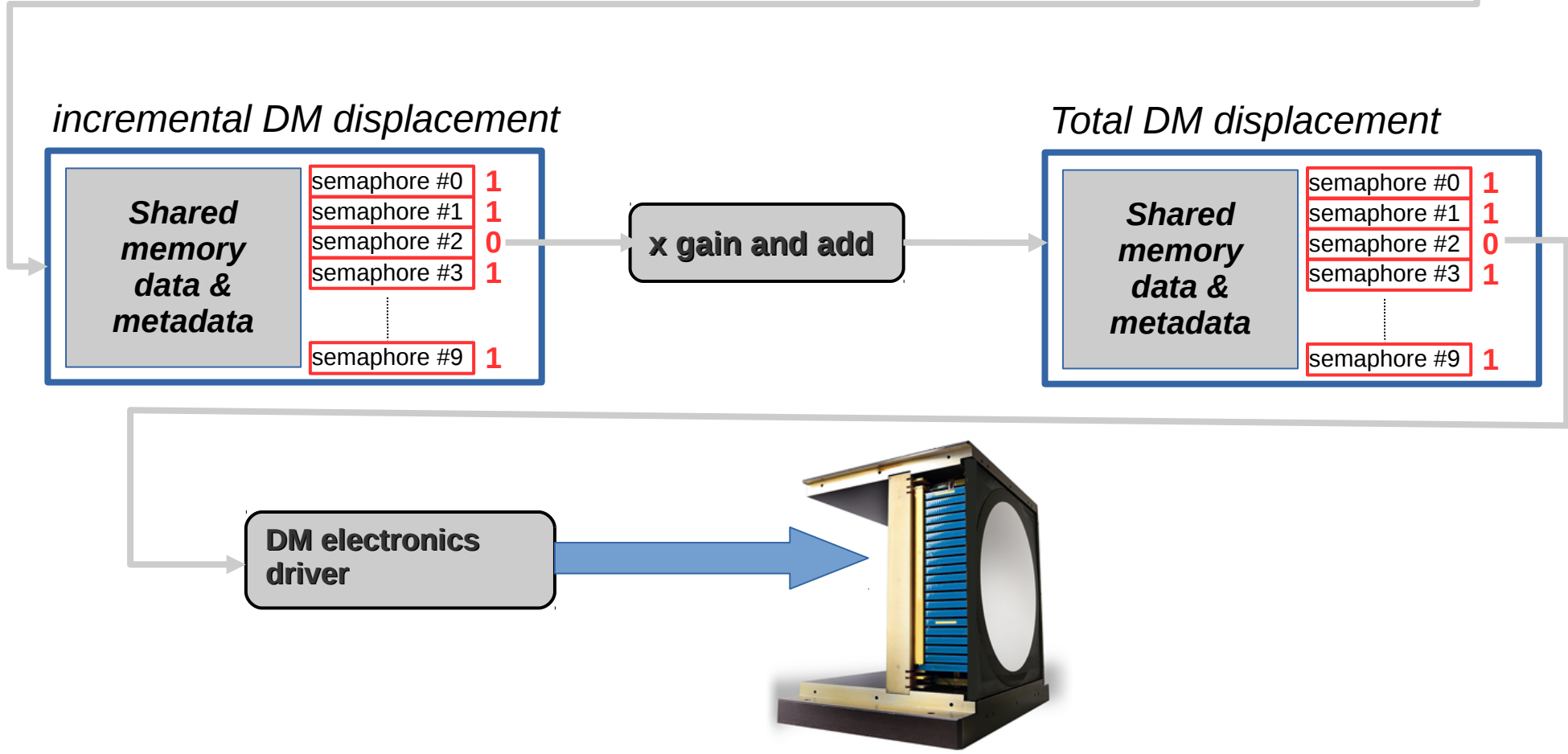
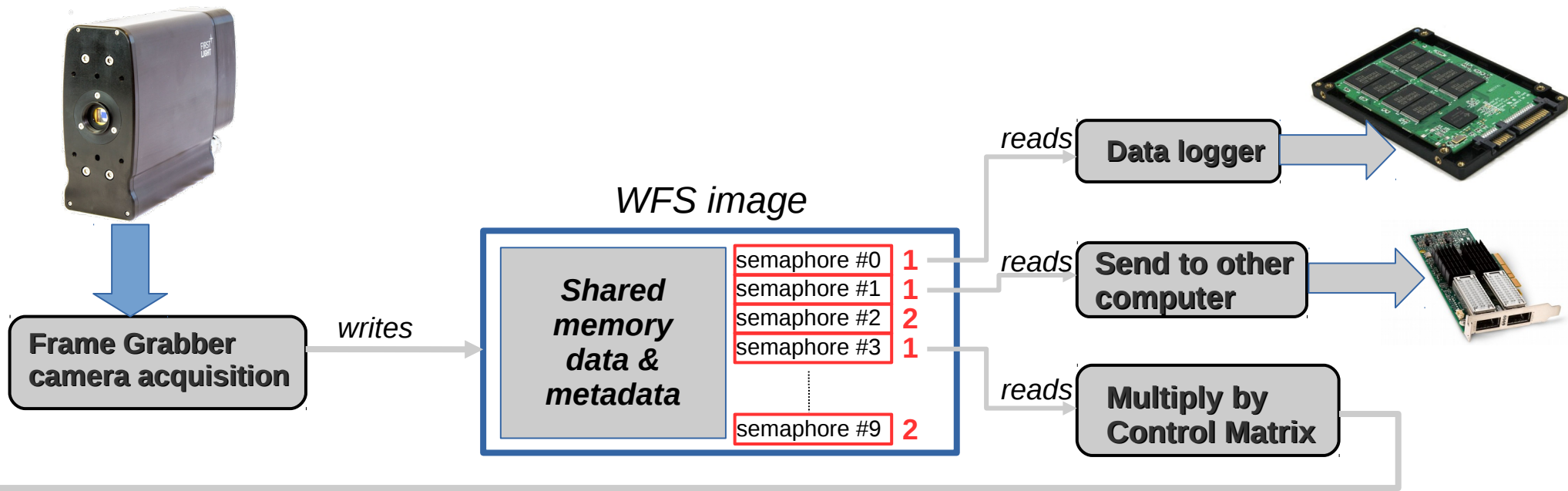
processes execution can overlap



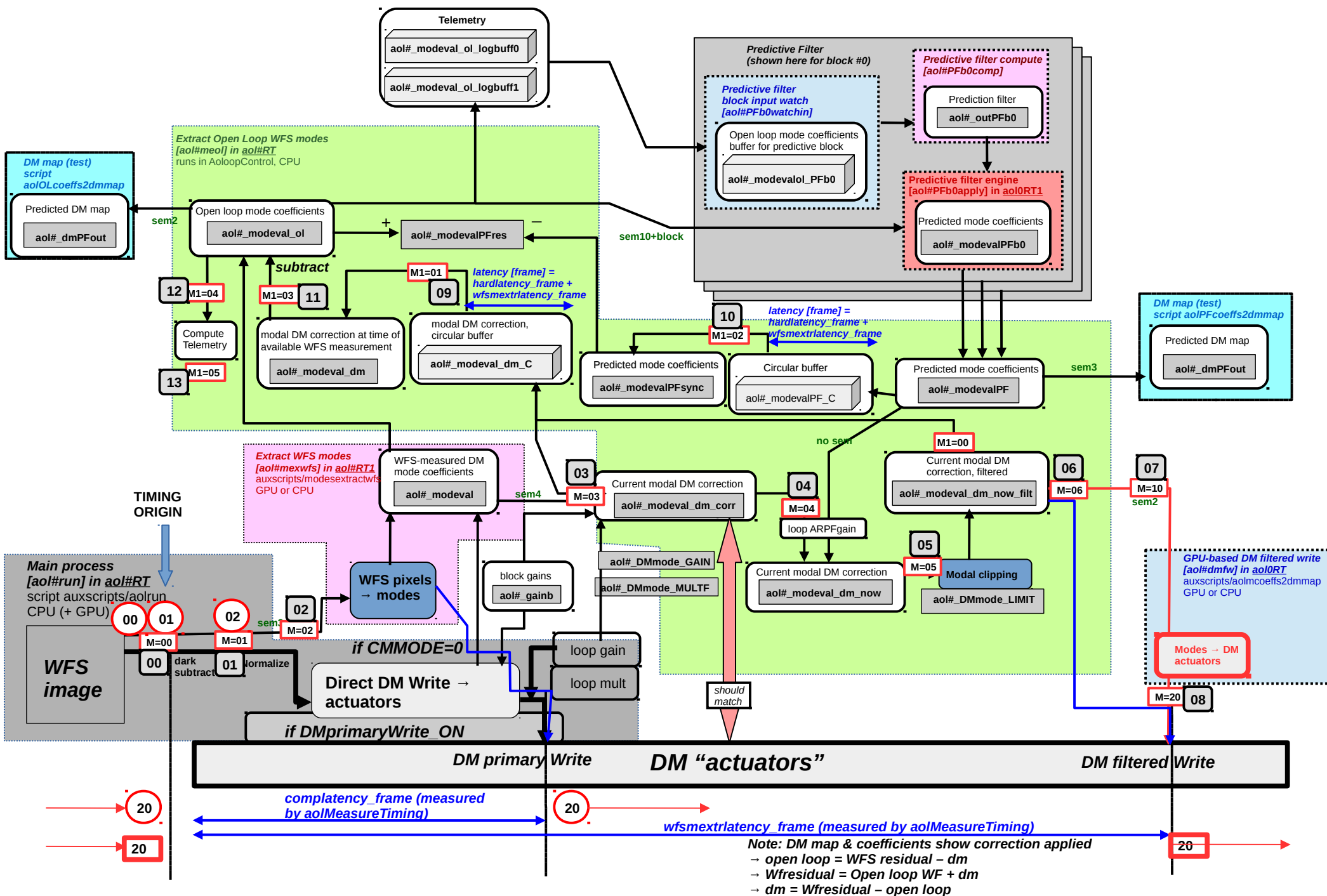


Unused semaphores go >1

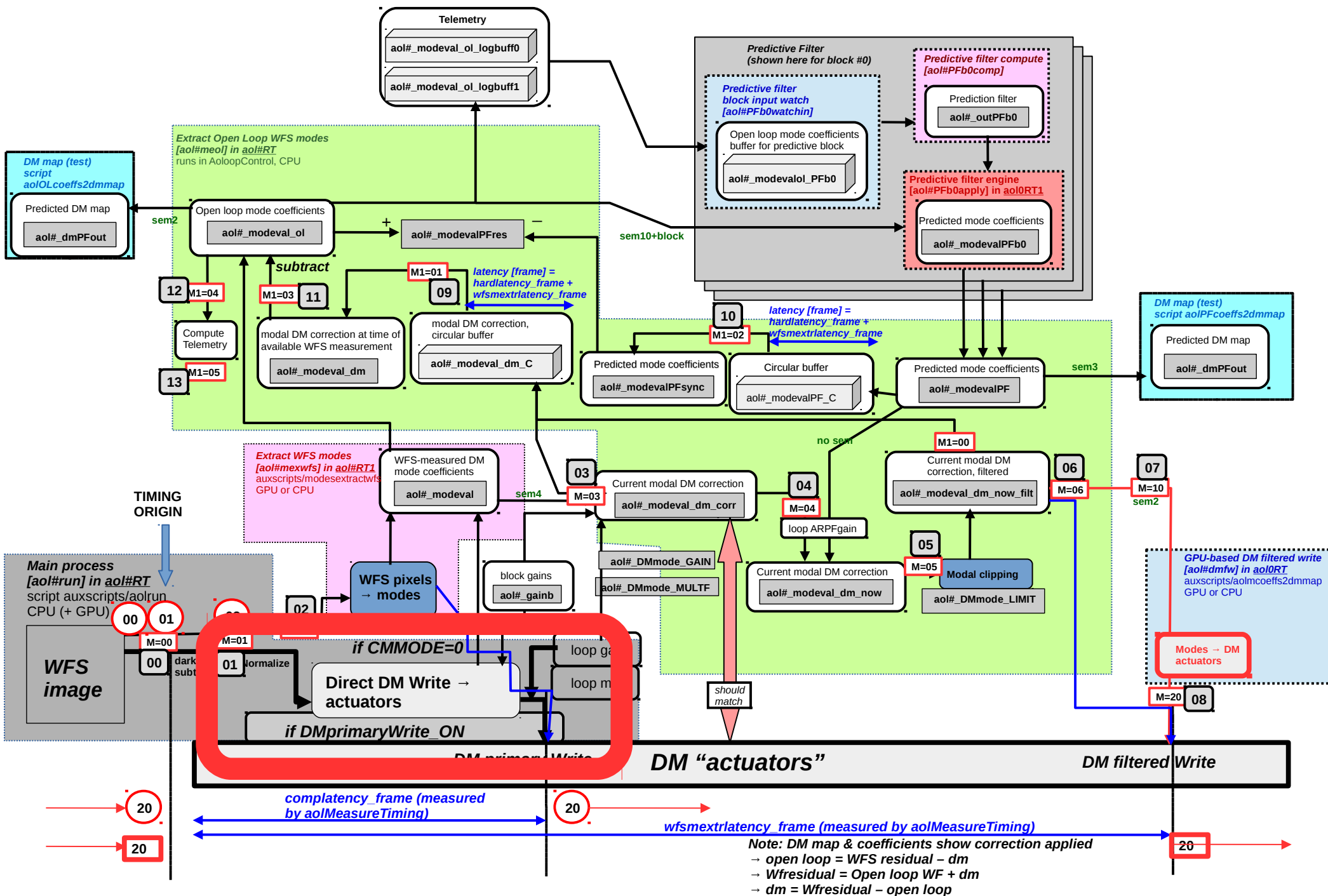




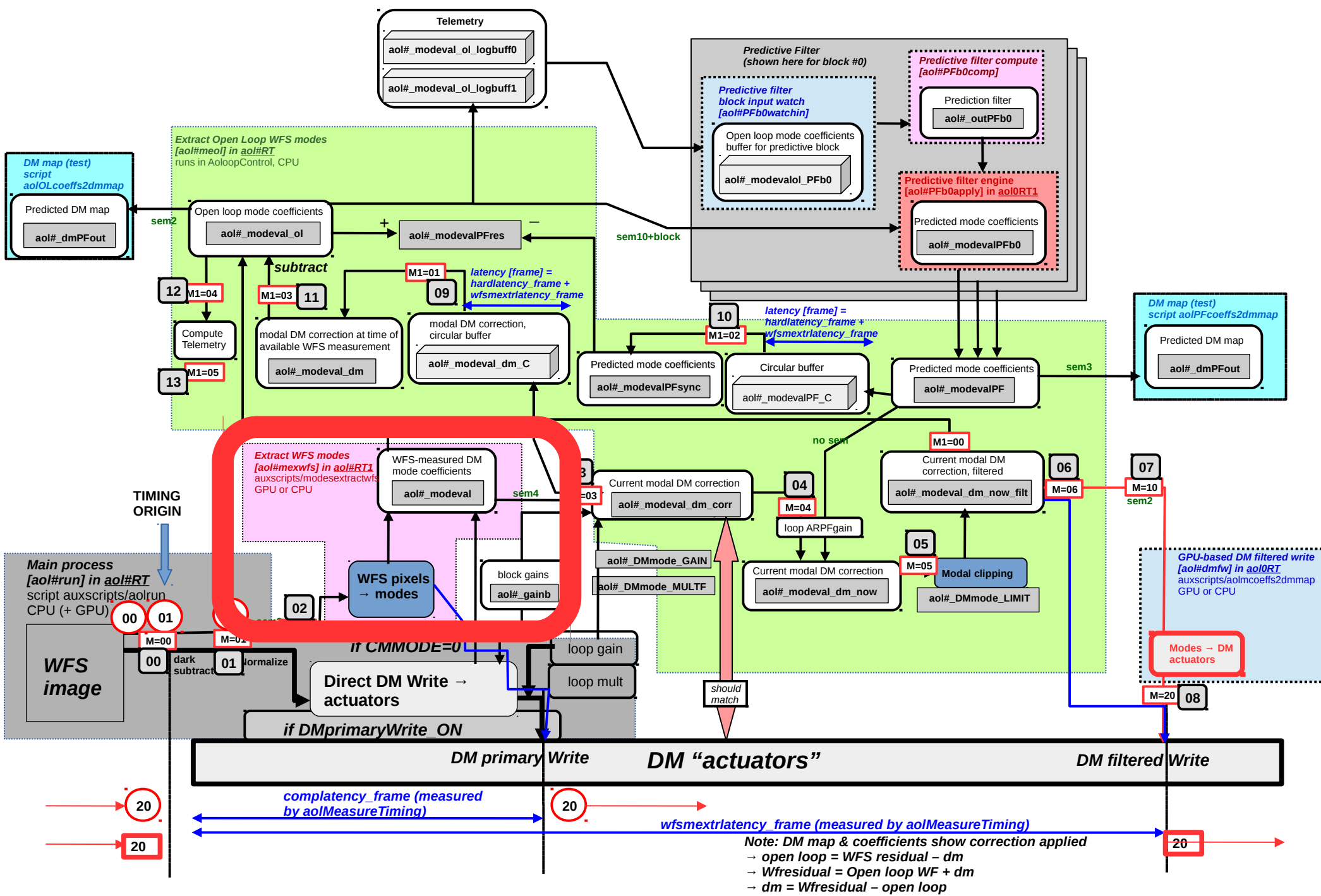
A more powerful example (SCExAO control loop)



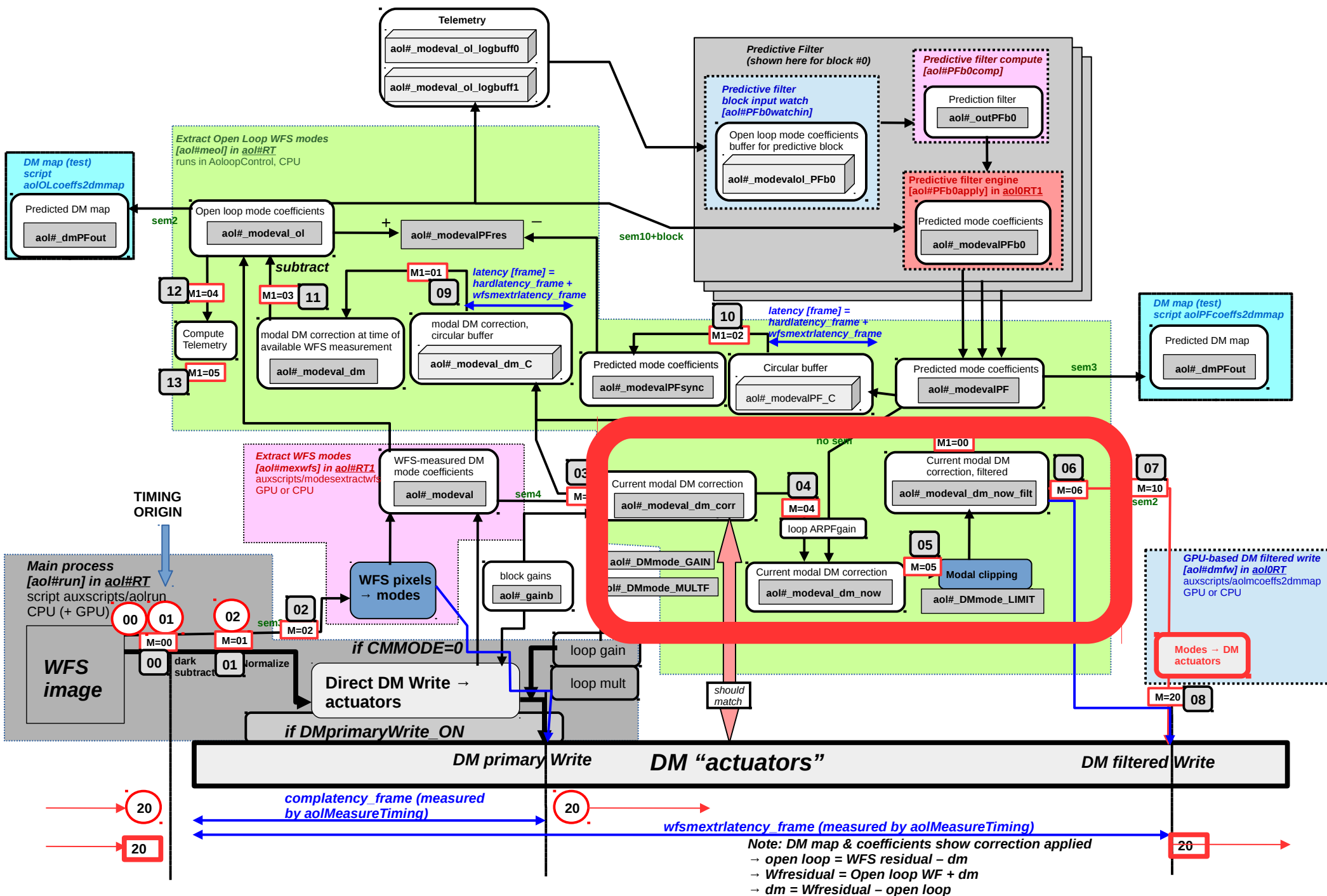
Fastest way to write on DM is to multiply WFS image by control matrix → DM actuators



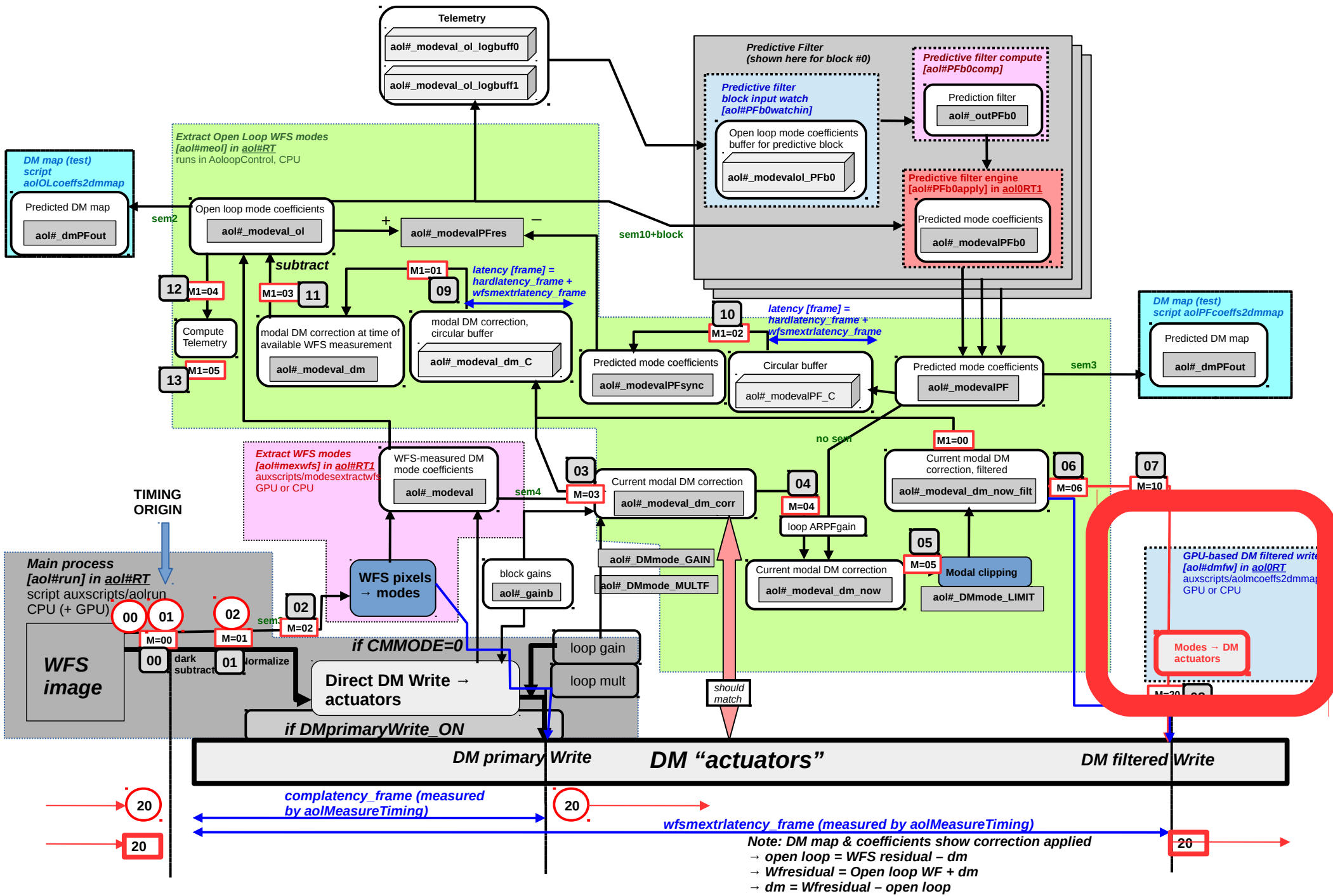
Modal reconstruction: compute WF modes from WFS image (usually GPU-based)



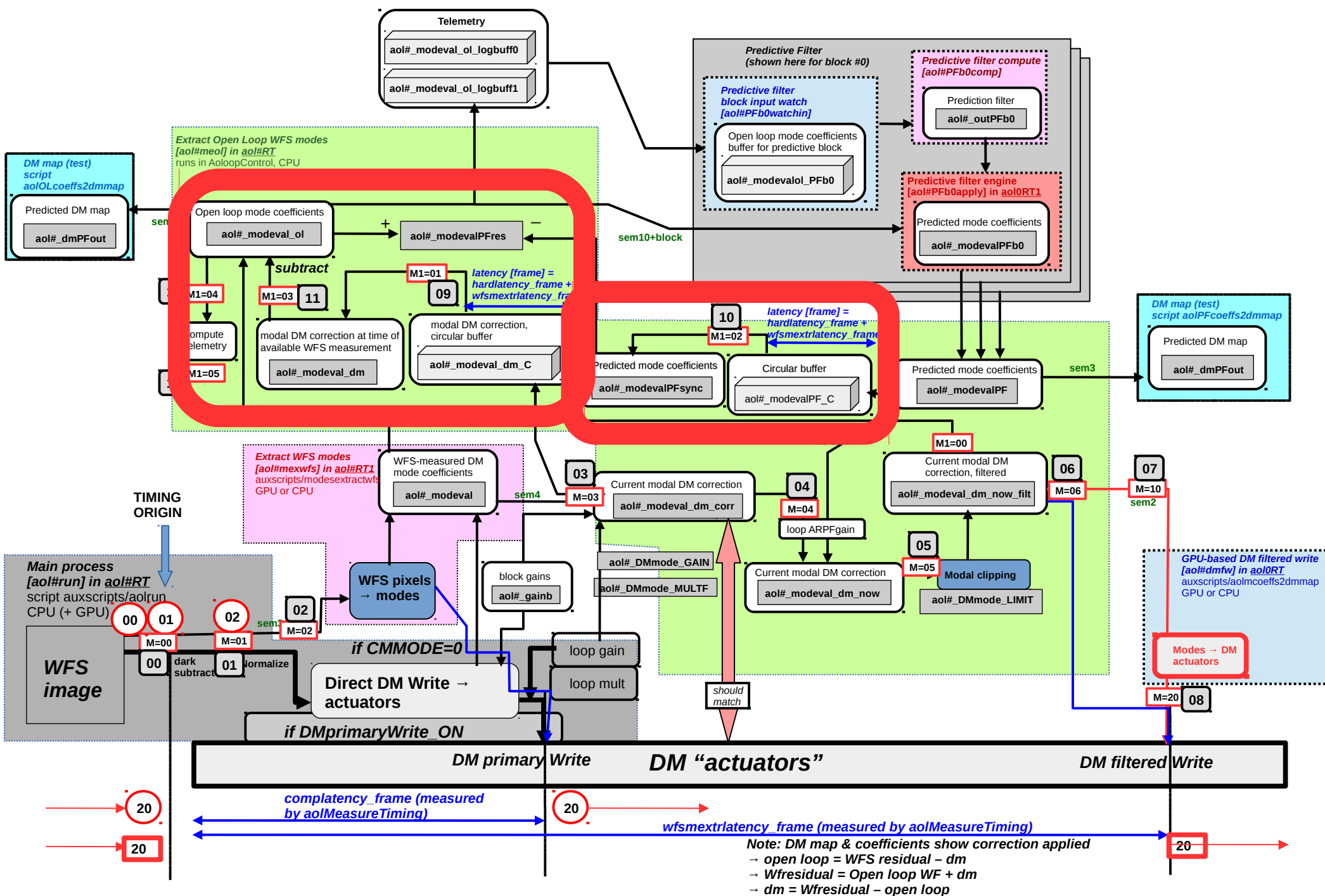
Modal filtering



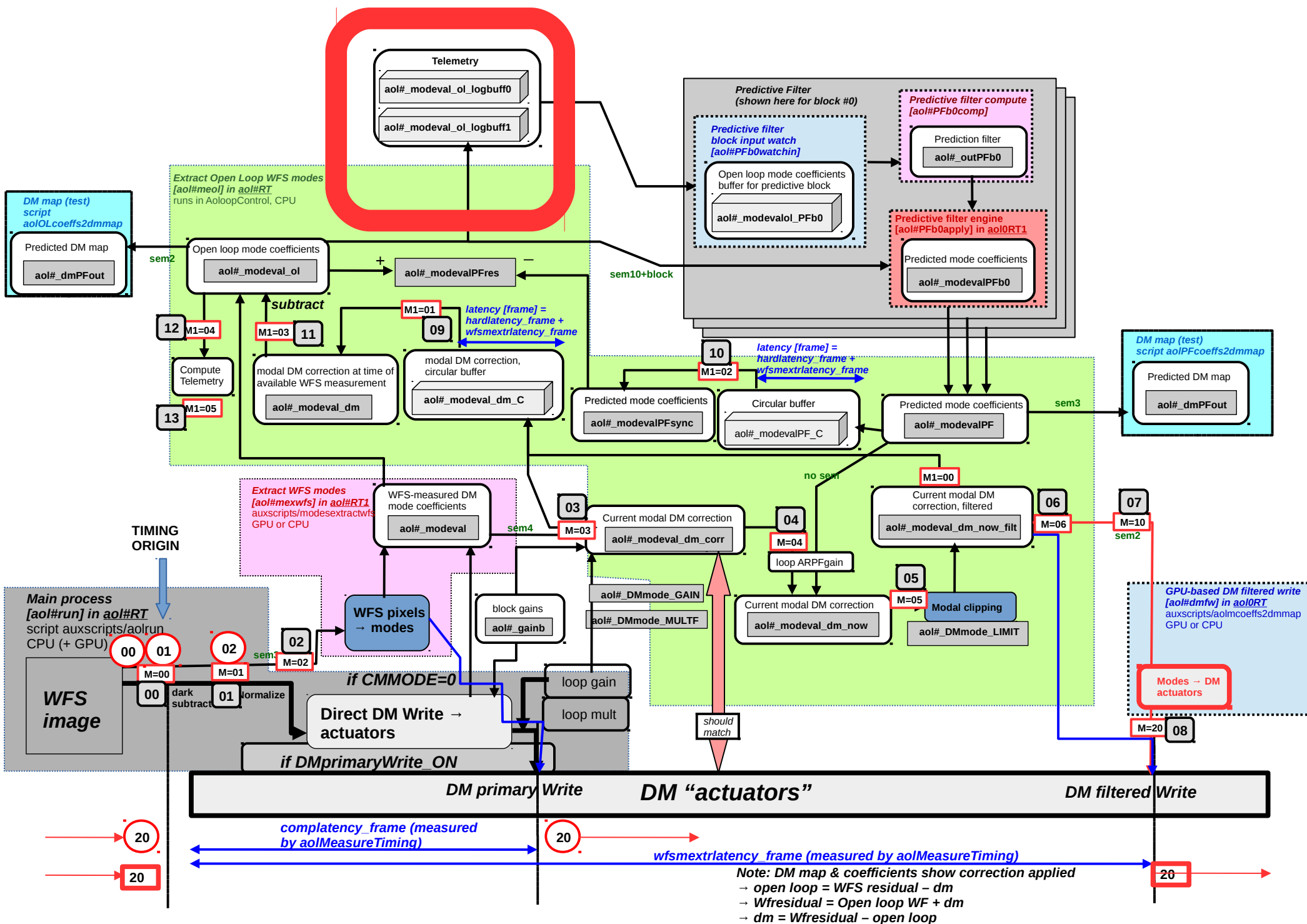
Modes → DM actuators (MVM operation)



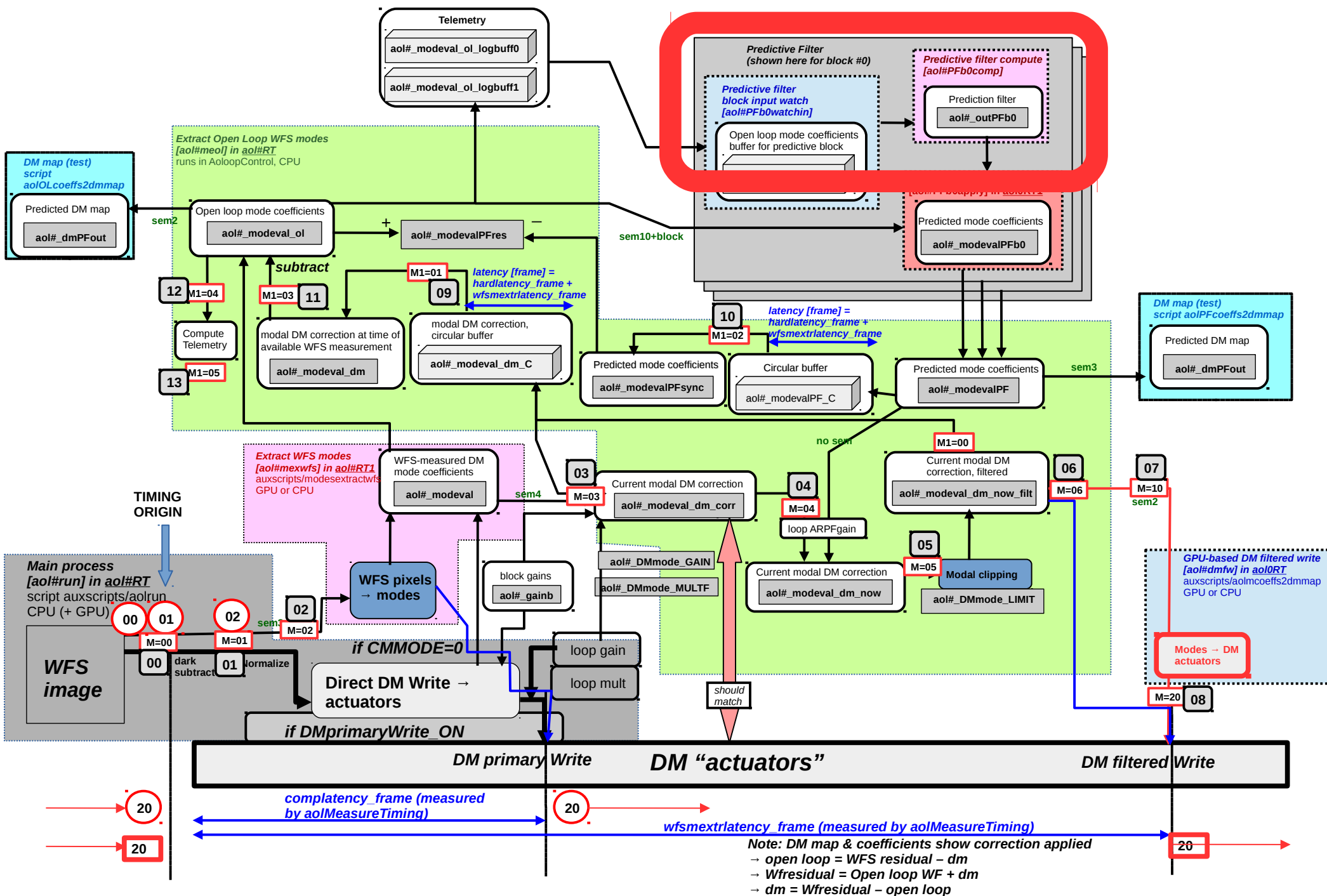
Modal pseudo-open loop reconstruction



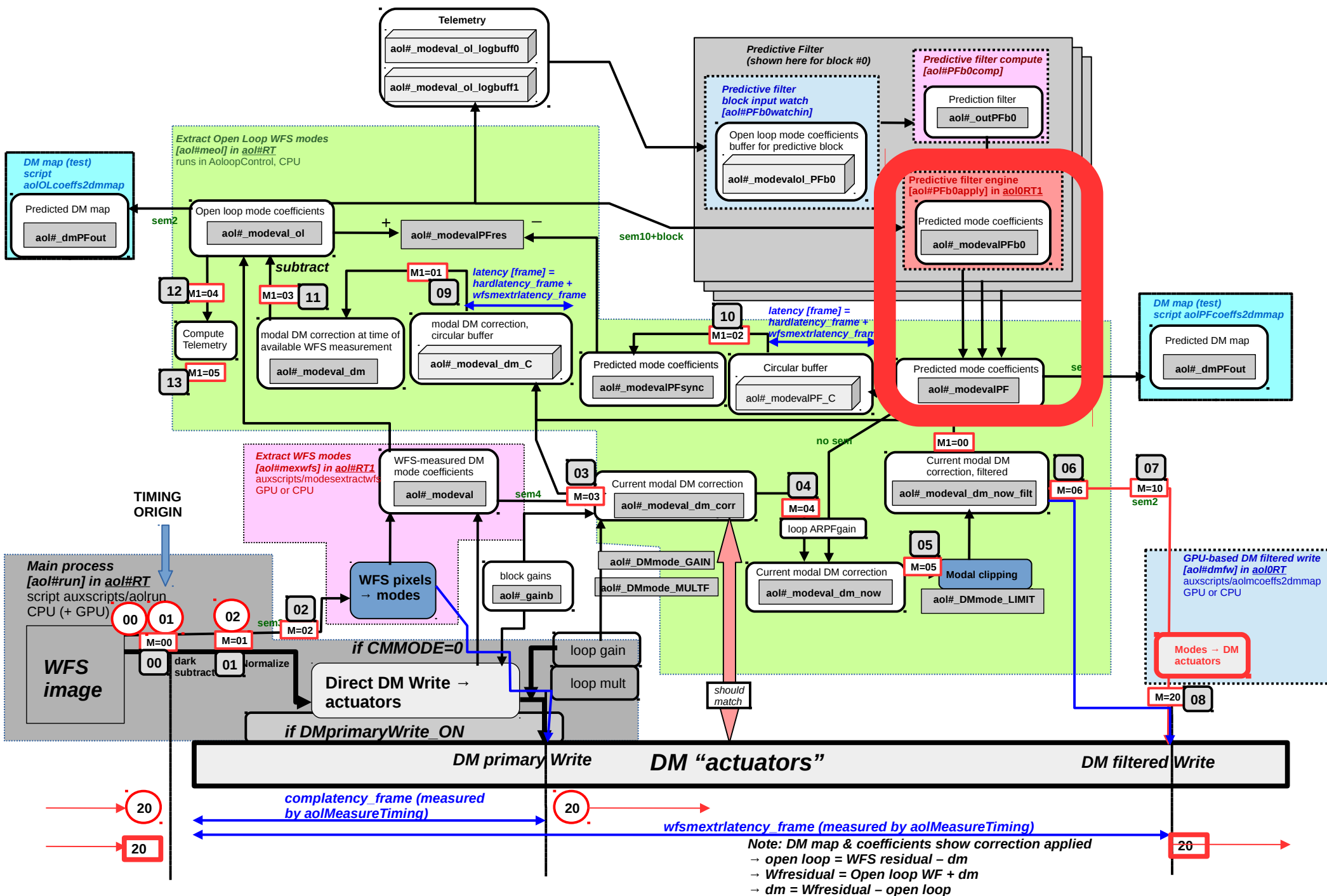
Write modal telemetry buffers



Compute Predictive Filter (soft real-time)



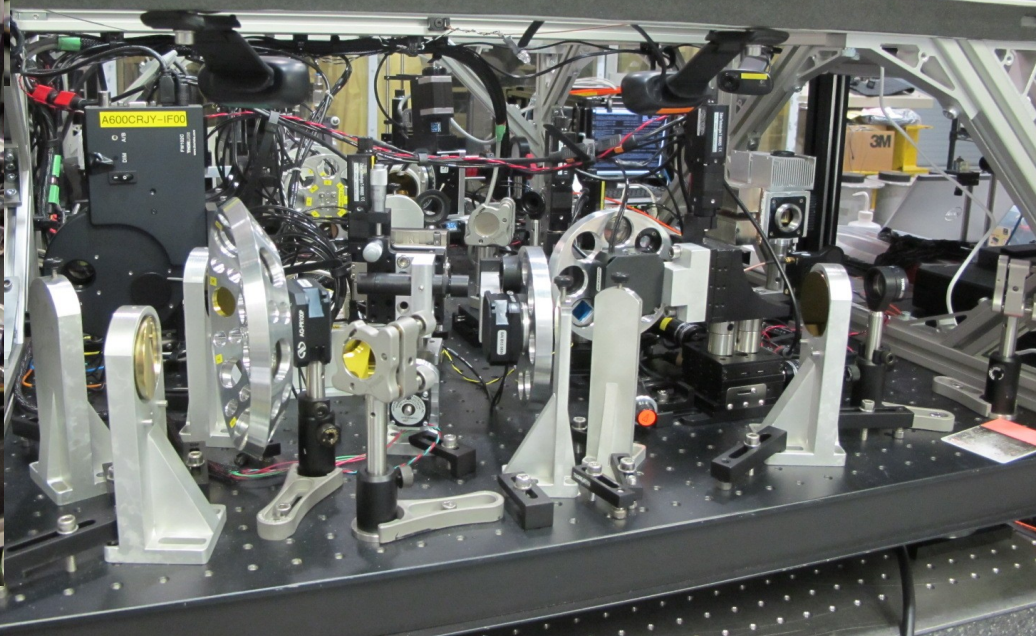
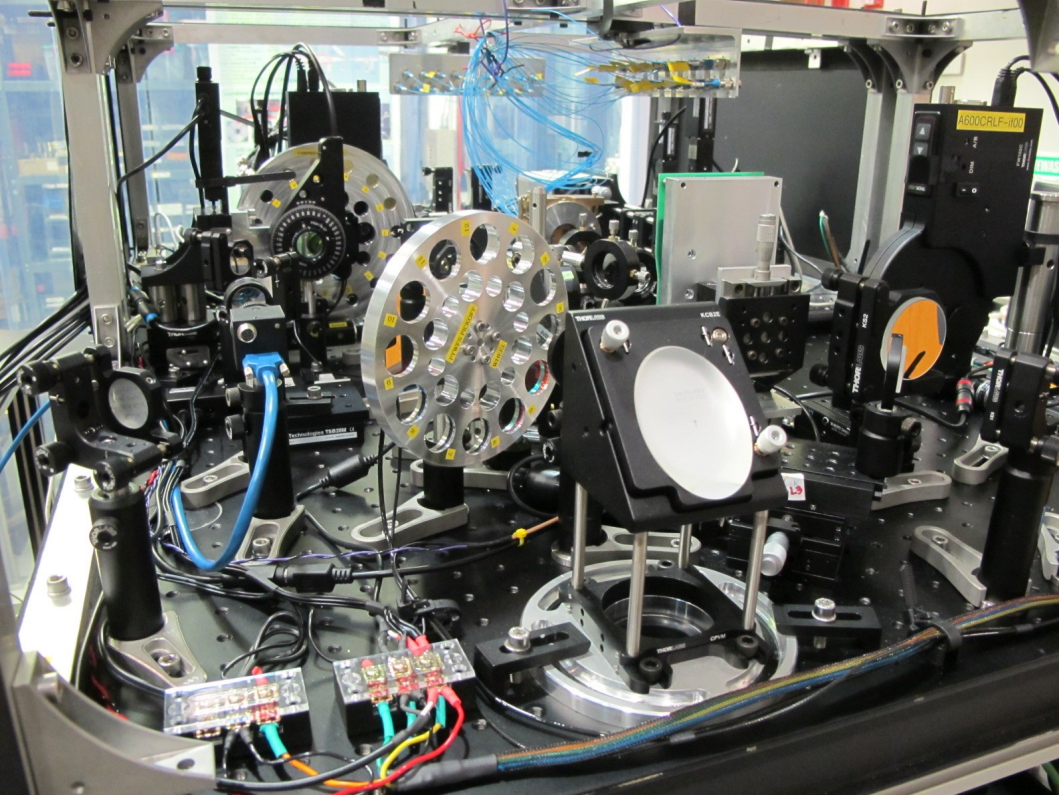
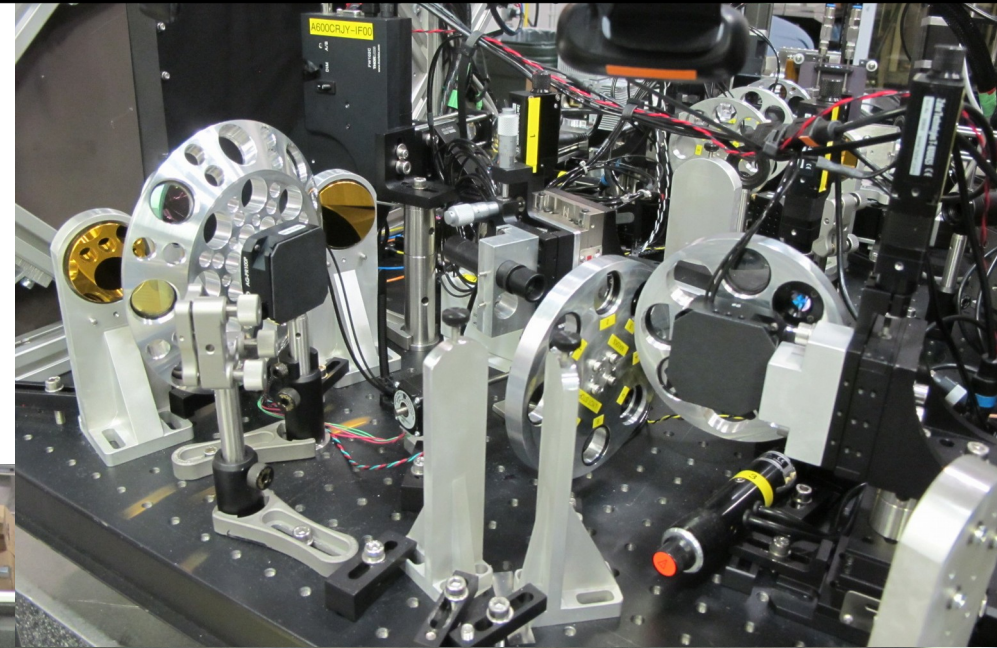
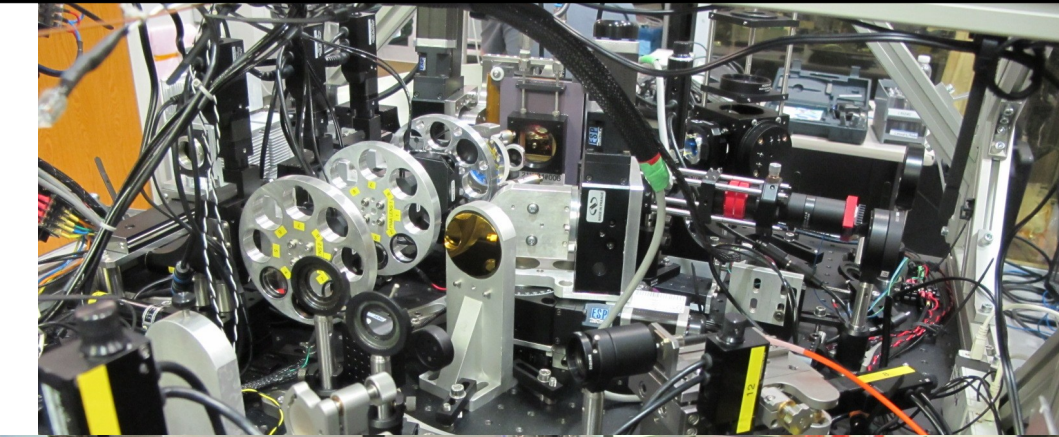
Apply Predictive Filter



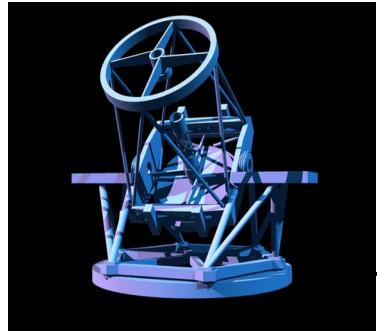
Linking loops / handling multiple WFSs and DMs



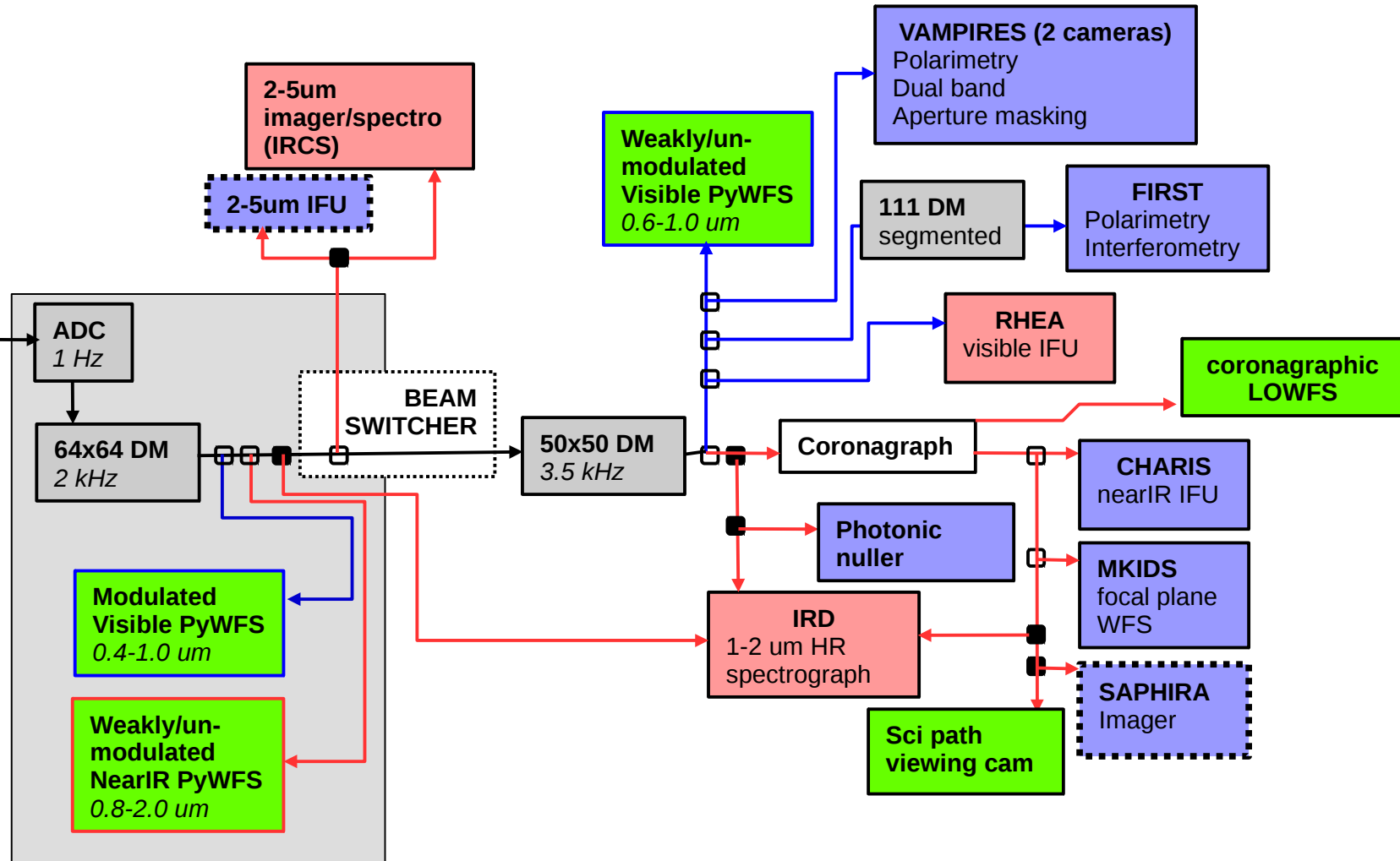
Subaru Coronagraphic Extreme Adaptive Optics



SCExAO Light path



Facility AO



Active WF correction

Dedicated science instrument

Mixed science/WFS

Dedicated WFS

Visitor port

□ *dichroic*

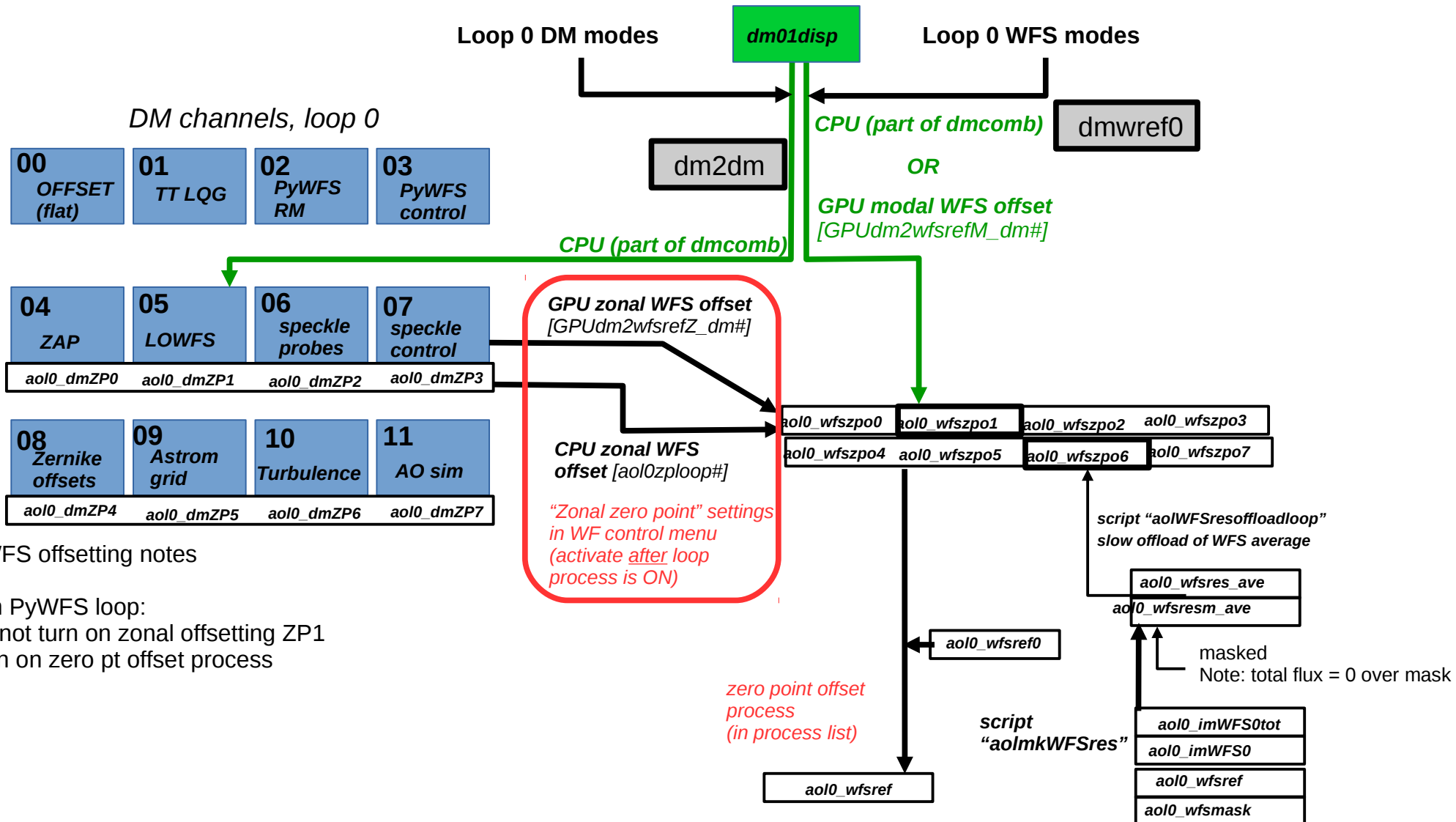
■ *beam switch*

Linking multiple control loops (zero point offsetting)

A control loop can offset the convergence point of another loop @> kHz (GPU or CPU)

Example: speckle control, LOWFS need to offset pyramid control loop

THIS IS DONE TRANSPARENTLY FOR USER → don't pay attention to the diagram below !

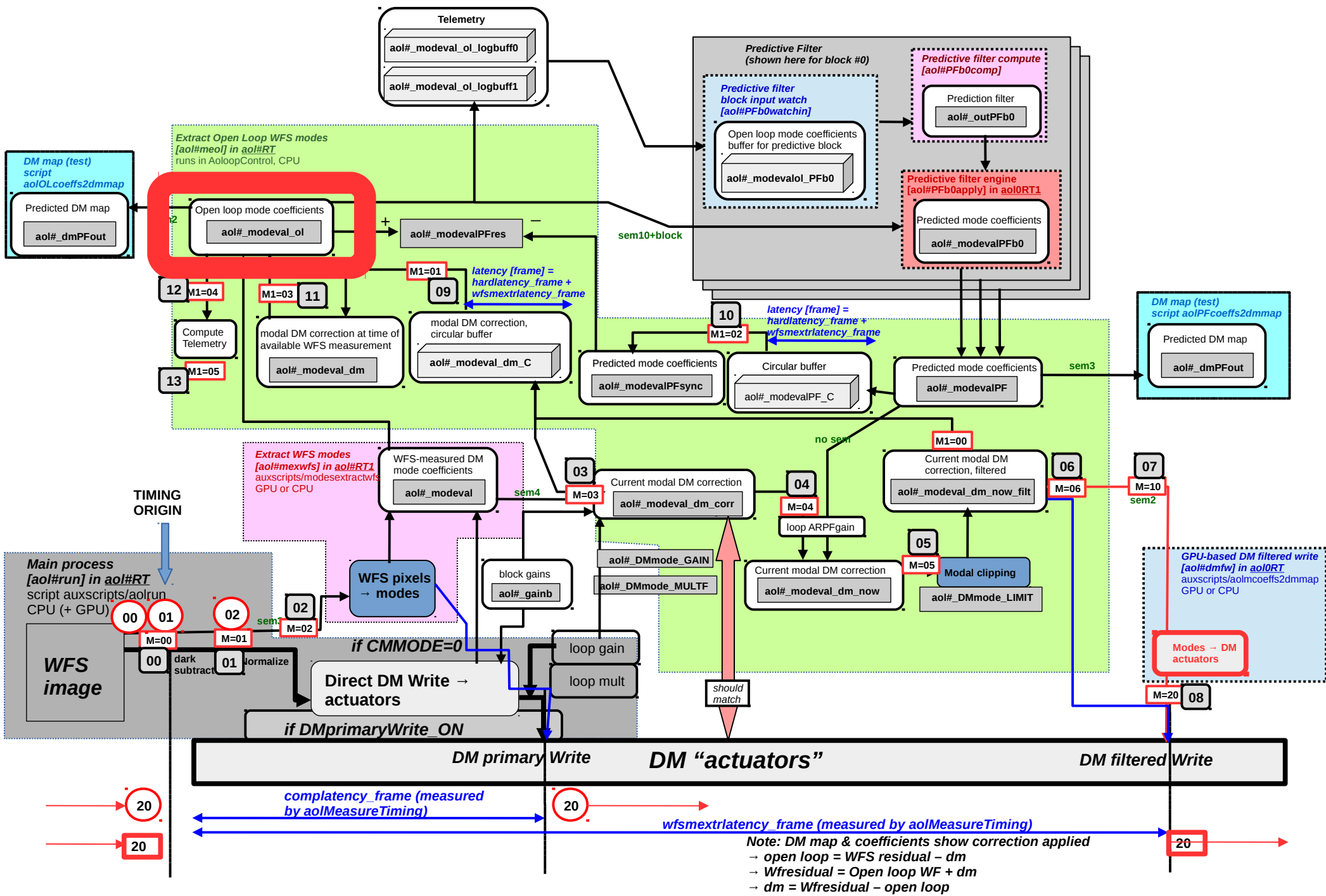


Timing is everything

Good timing knowledge and stability is essential for:

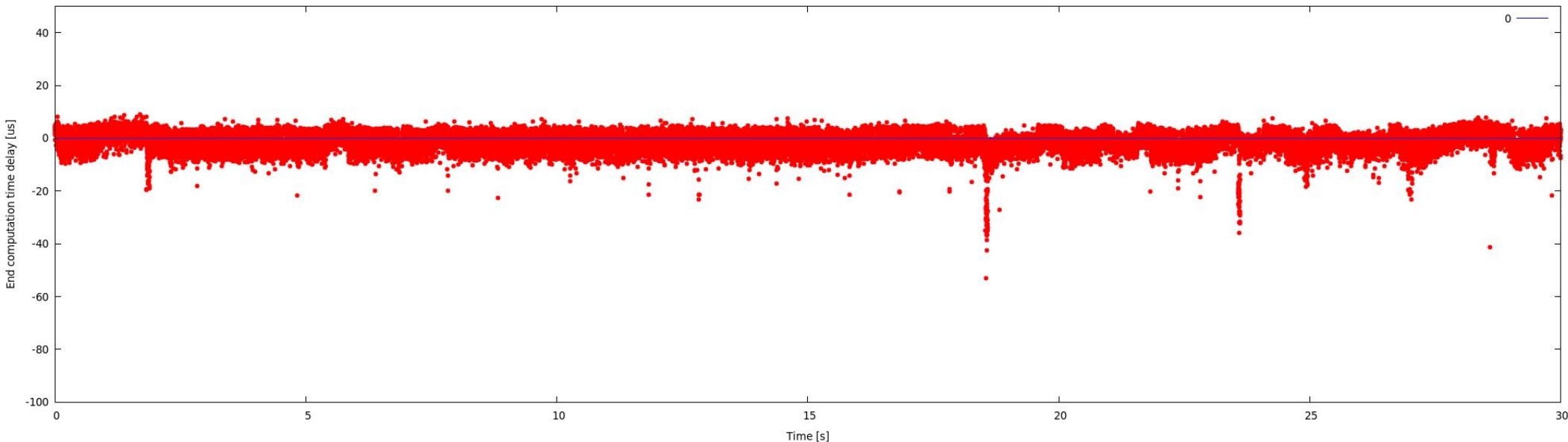
- Pseudo-open loop reconstruction
- Fast response matrix acquisition
- Predictive control

End of real-time computation processes



End-to-end Timing Jitter

RMS < 5 us, max delay ~50us

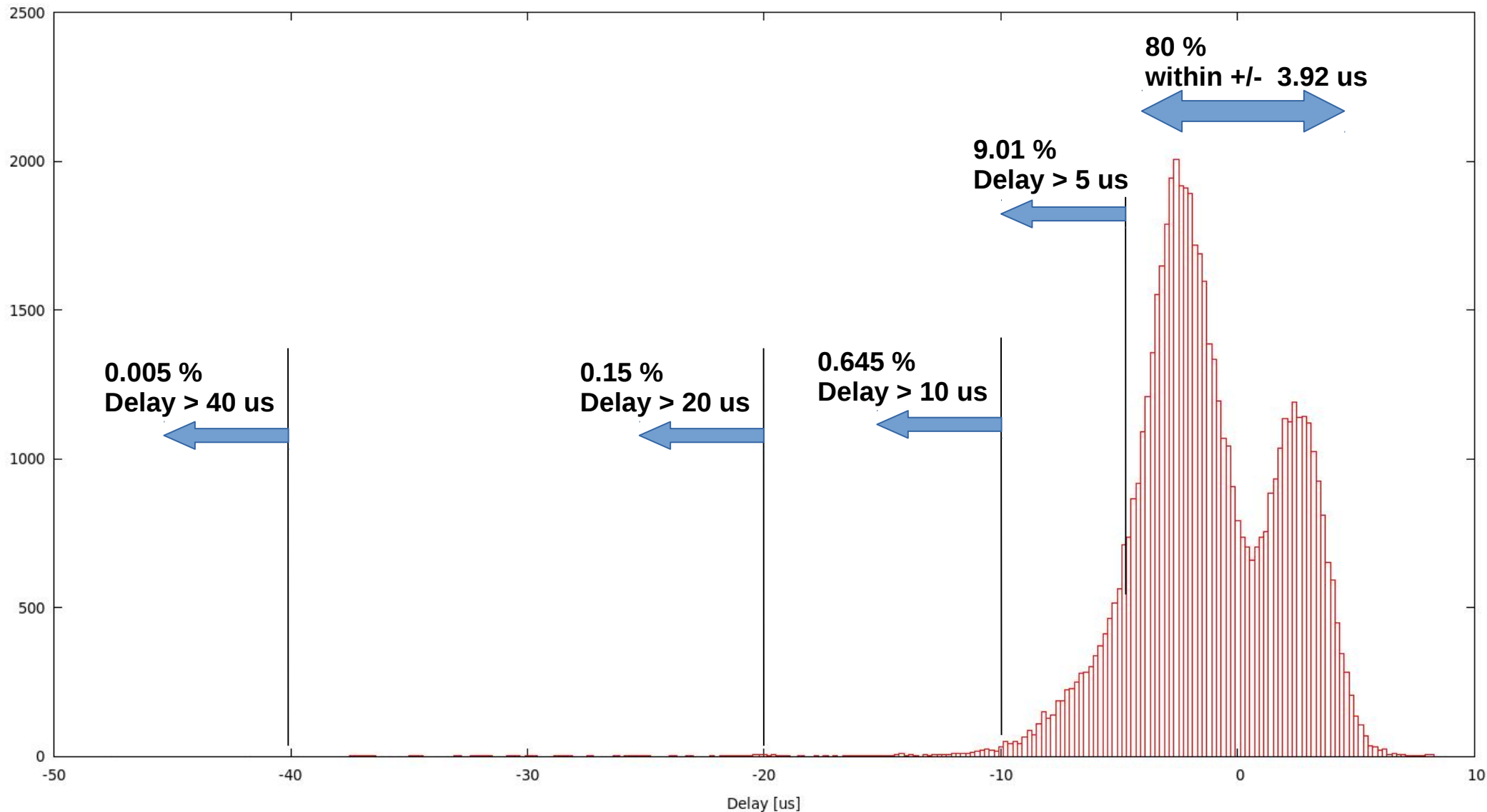


End-to-end timing jitter measured by monitoring completion time of last real-time stream: modal pseudo-open loop coefficients.

Jitter includes following components:

- Hardware synchronization (PyWFS tip-tilt mirror)
- Camera readout
- Data transfer over TCP link
- All real-time computations, CPU and GPU
- Time measurement errors

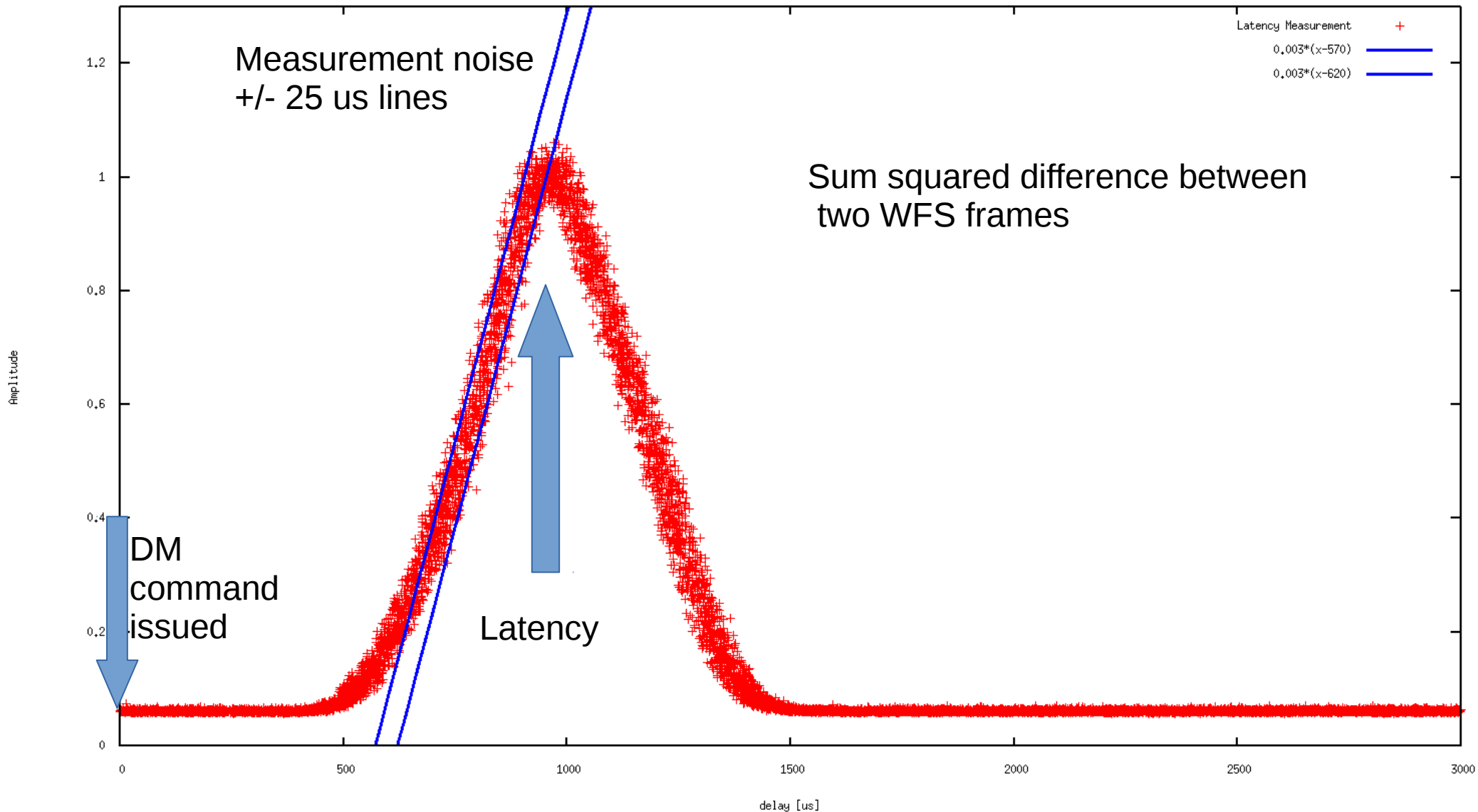
End-to-end Timing Jitter Histogram, measured @ 2kHz



10% of loop iteration @ 2 kHz

Hardware Latency measured on SCExAO

Time between DM command issued and corresponding WFS signal observed
(Camera readout + TCP transfer + processing + DM electronics)



Hardware Latency

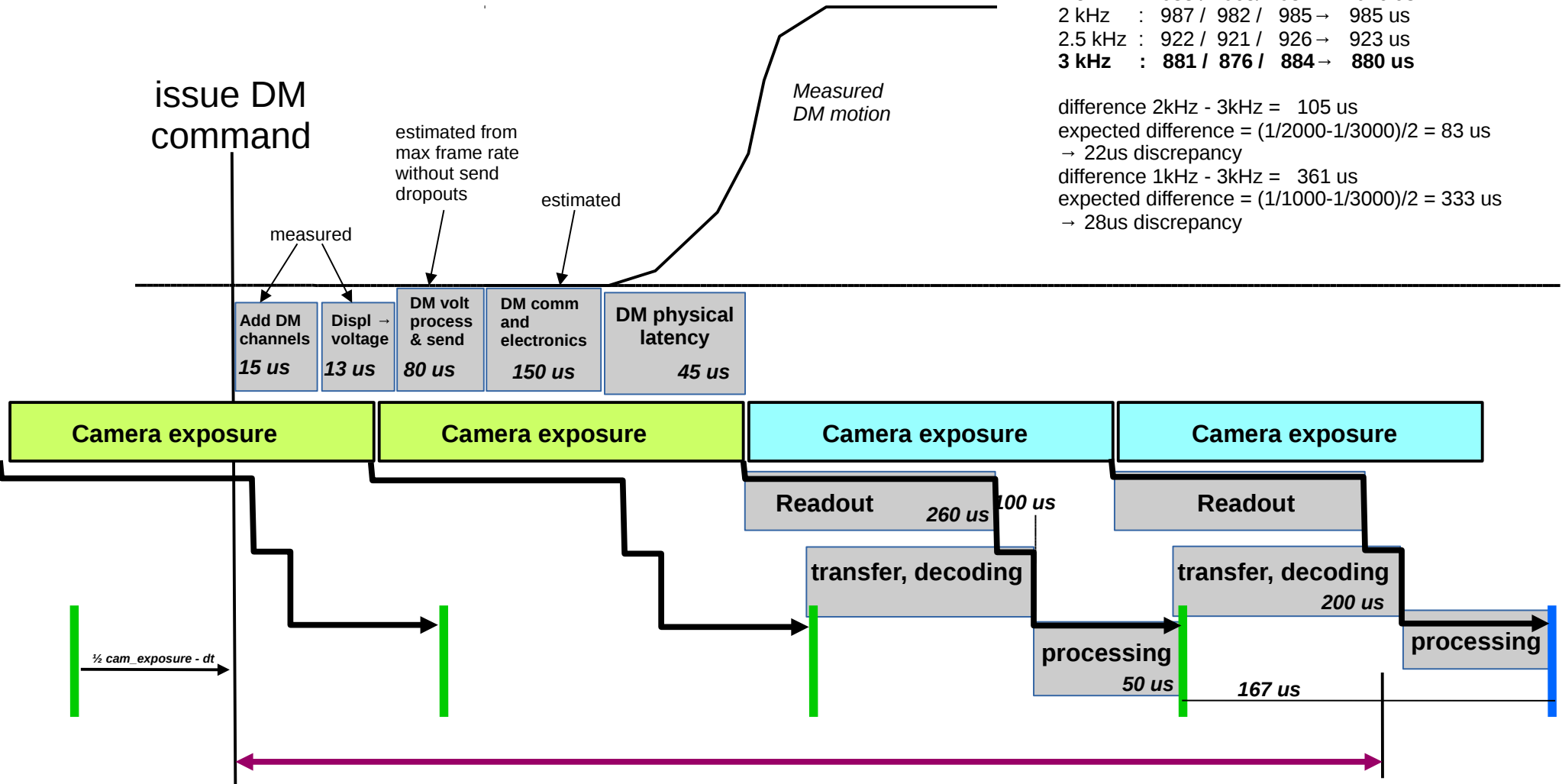
Definition:

*Time offset between **DM command issued**, and **mid-point between 2 consecutive WFS frames with largest difference***

SCEXAO measured hardware latencies:

1 kHz : 1253 / 1260 / 1269 → 1261 us
 1.5 kHz : 1083 / 1065 / 1081 → 1076 us
 2 kHz : 987 / 982 / 985 → 985 us
 2.5 kHz : 922 / 921 / 926 → 923 us
3 kHz : 881 / 876 / 884 → 880 us

difference 2kHz - 3kHz = 105 us
 expected difference = $(1/2000 - 1/3000)/2 = 83$ us
 → 22us discrepancy
 difference 1kHz - 3kHz = 361 us
 expected difference = $(1/1000 - 1/3000)/2 = 333$ us
 → 28us discrepancy



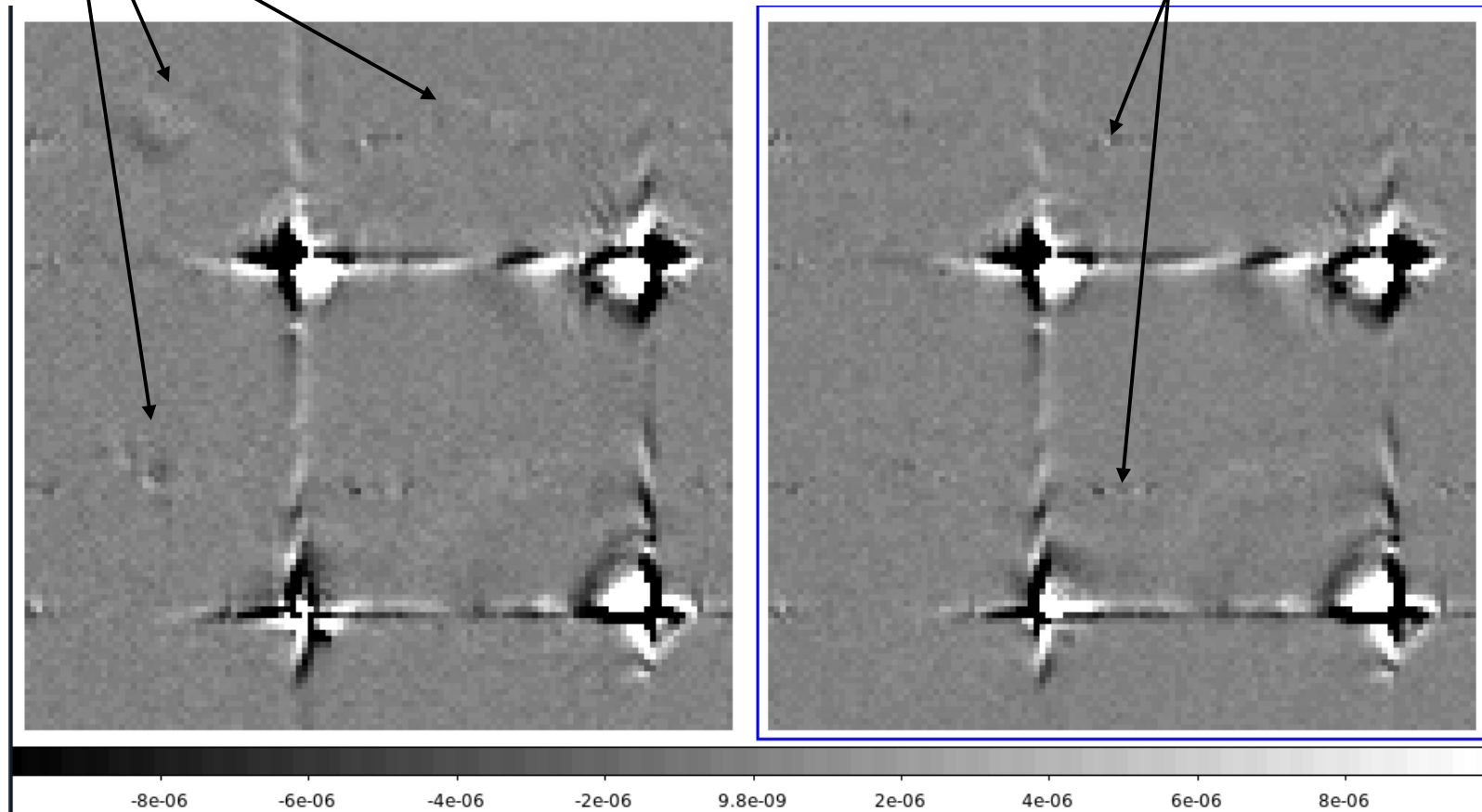
HardwareLatency = DM soft + DM elec + DM phys + CAM readout/transfer + CAM processing + $\frac{1}{2}$ exposure time

HardwareLatency = $N \times \text{cam_exposure} + dt$

**Fast RM acquisition (4000 Hadamard pokes in 2s @ 2 kHz)
+ Removing temporal DM response from response matrix by
using two poke sequences**

Temporal bleeding from previous poke pattern (should be removed from RM)

**Camera readout RF coupling
between pixels
~1% electronic ghosts at 2kHz
frame rate
Needs to be kept in RM**



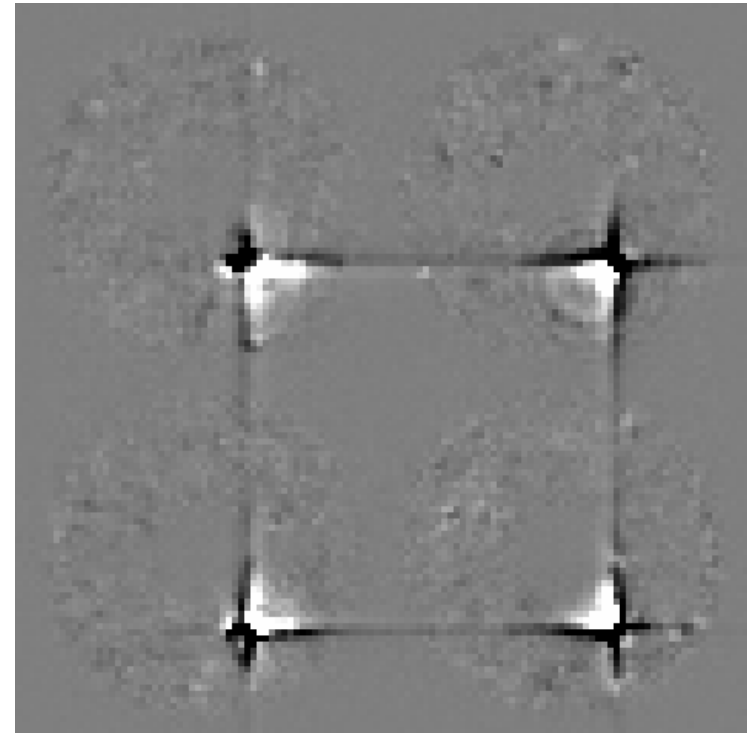
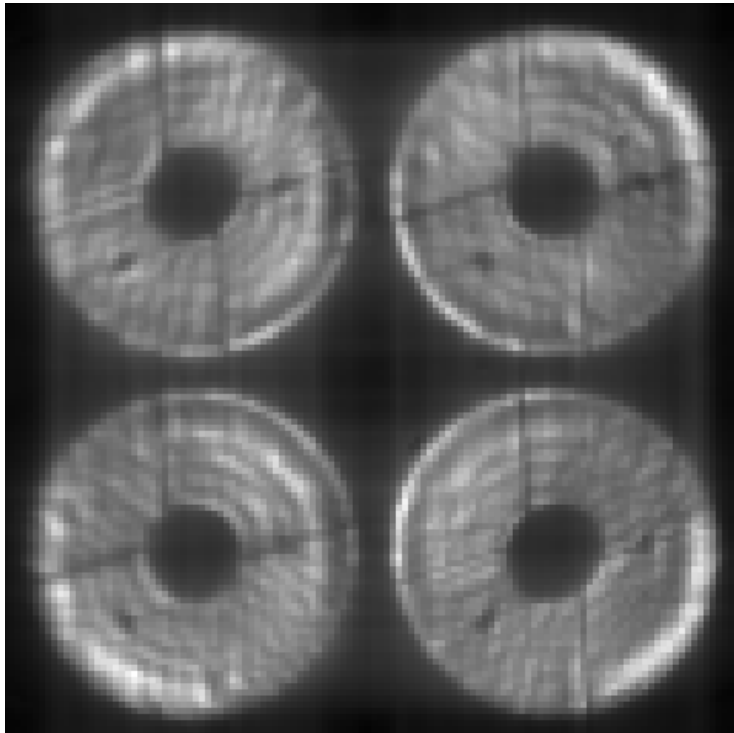
RM assembled from single poke sequence:
+- +- +- +-

RM assembled from average of two poke sequences:
 +- +- +- +-
 +- -+ +- -+

RMs reconstructed from Hadamard pokes, 2kHz modulation (DM moves during EMCCD frame transfer)

Multi-channel DM virtualization & timing knowledge/stability

→ on-sky response matrix acquisition, while ExAO loop running



Left: WFS reference

Right: Response to single actuator poke (one of 2000)

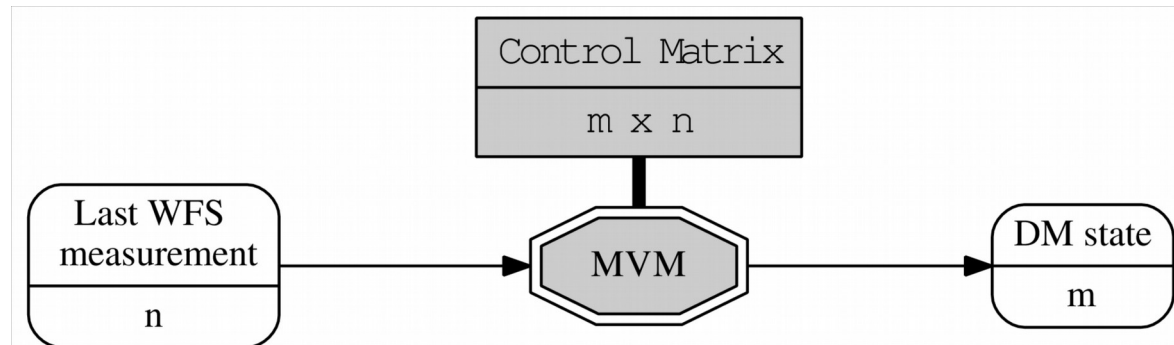
RM measurement @ 2kHz takes 4000 pokes = 2 sec

Multiple RMs averaged to increase SNR

Self-learning AO control

Conventional AO:

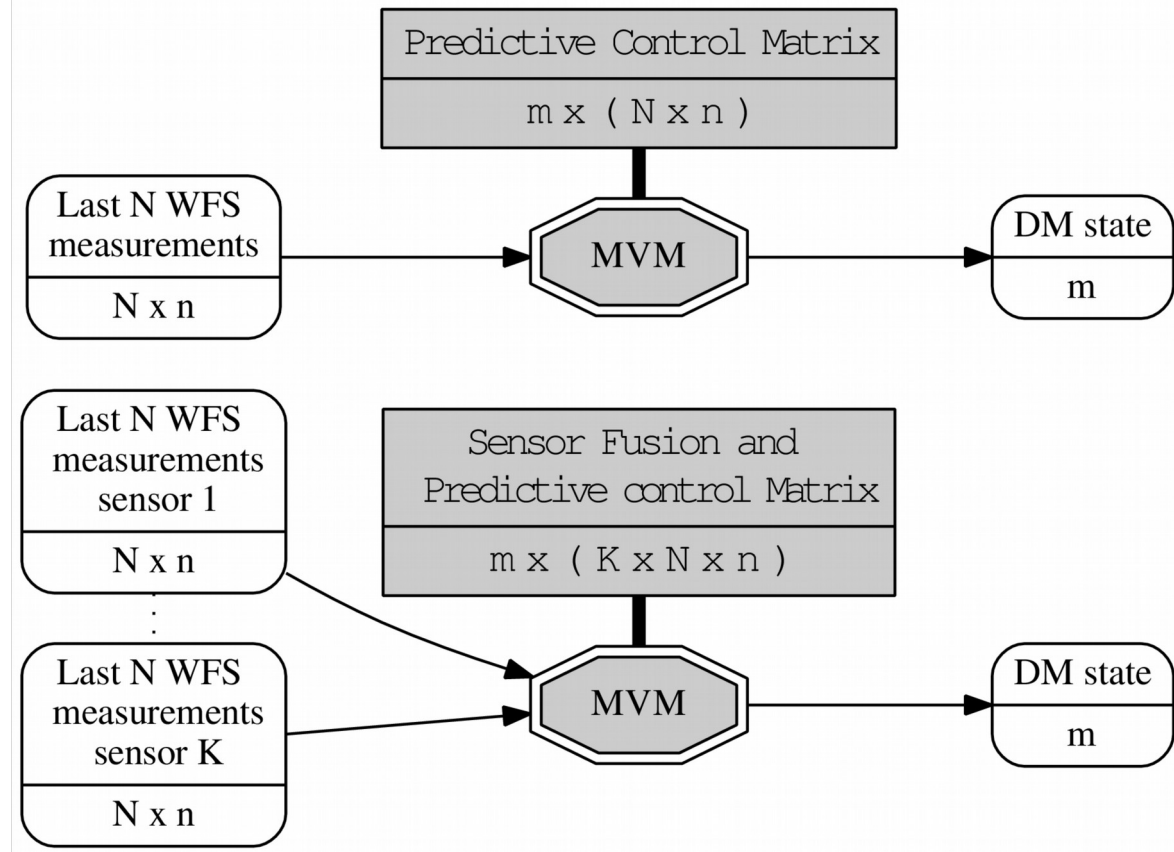
Resp Matrix is measured
CM computed as pseudo-inverse of RM



Self-learning AO control:

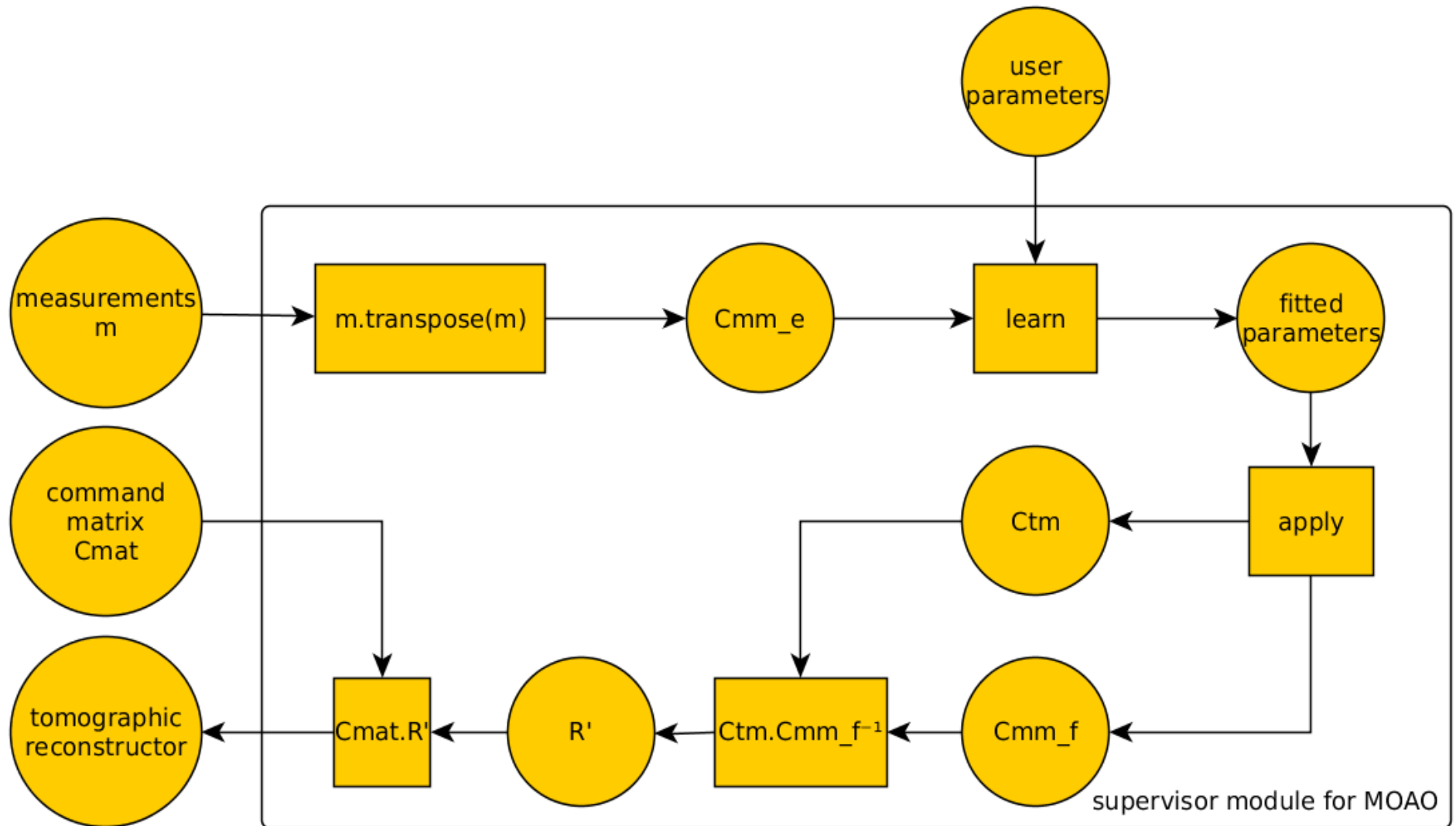
Optimally use recent (predictive control) and auxiliary (sensor fusion) measurements → control matrix is very big, and usually impossible to measure

CM is derived from WFS(s) telemetry using machine learning approaches



Loop supervision module

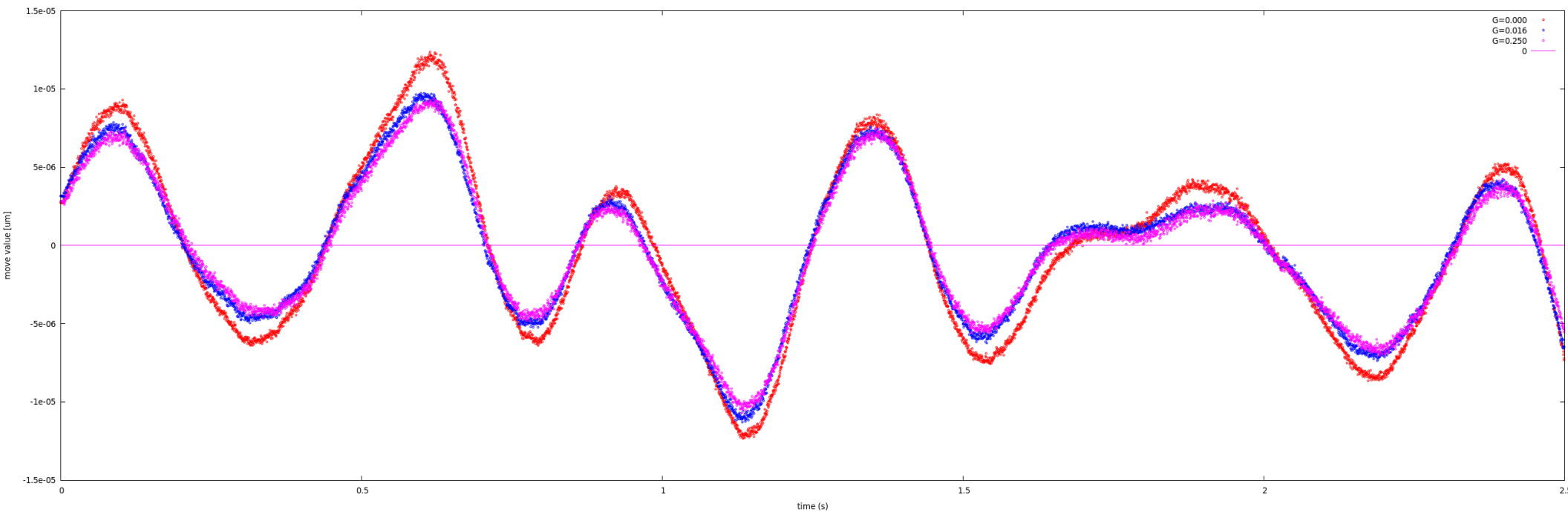
Mix of cost function optimization for parameters identification (“Learn” process) and linear algebra for reconstructor matrix computation (“apply” process)



Open loop reconstruction

Comparison between gain values

$G=0.000 \rightarrow$ over-estimates OL values
All $G>0.0$ reconstructions match at %-level



$G=0.000$ test relies entirely on WF residuals for OL estimation
 $G>0.000$ tests rely mostly on DM values for OL estimation

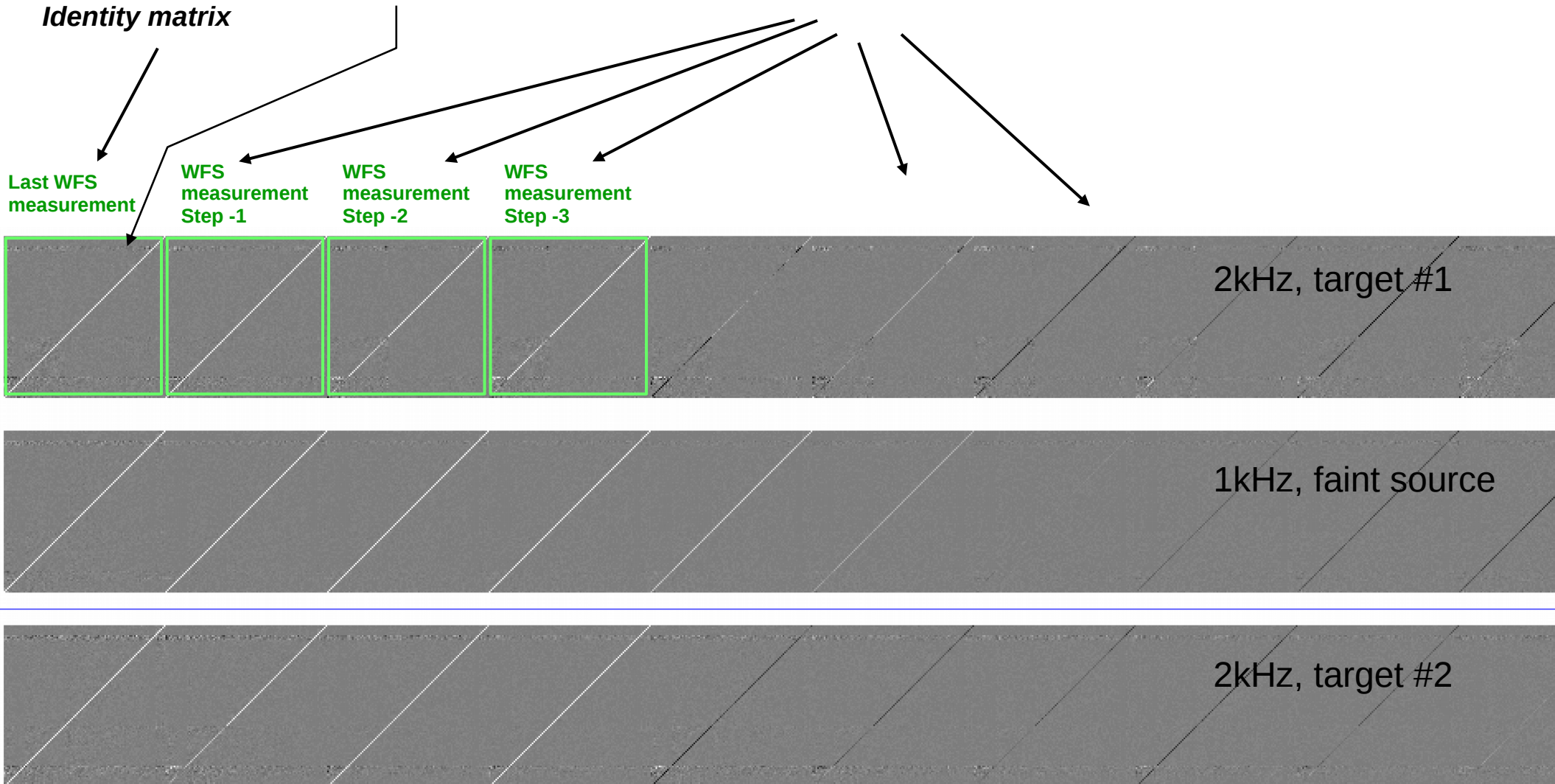
Test shown here uses full speed RM acquisition which underestimates RM by ~15% due to DM time-of-motion $\rightarrow$ reconstructed WFs from WFS are over-estimated by ~15%

On-sky predictive control matrix (modal representation, 100 modes shown)

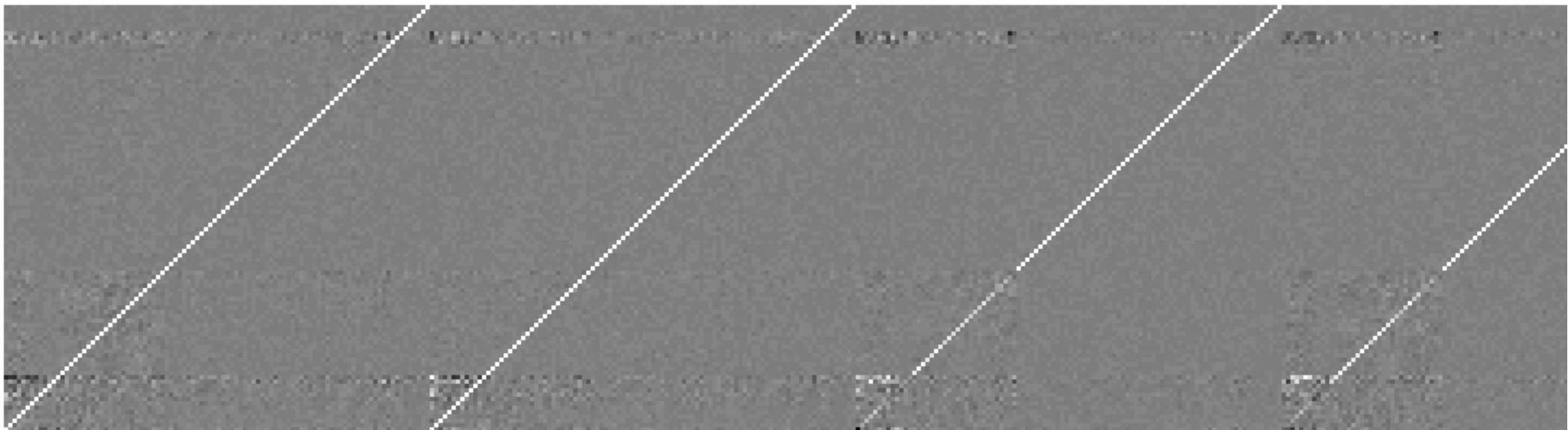
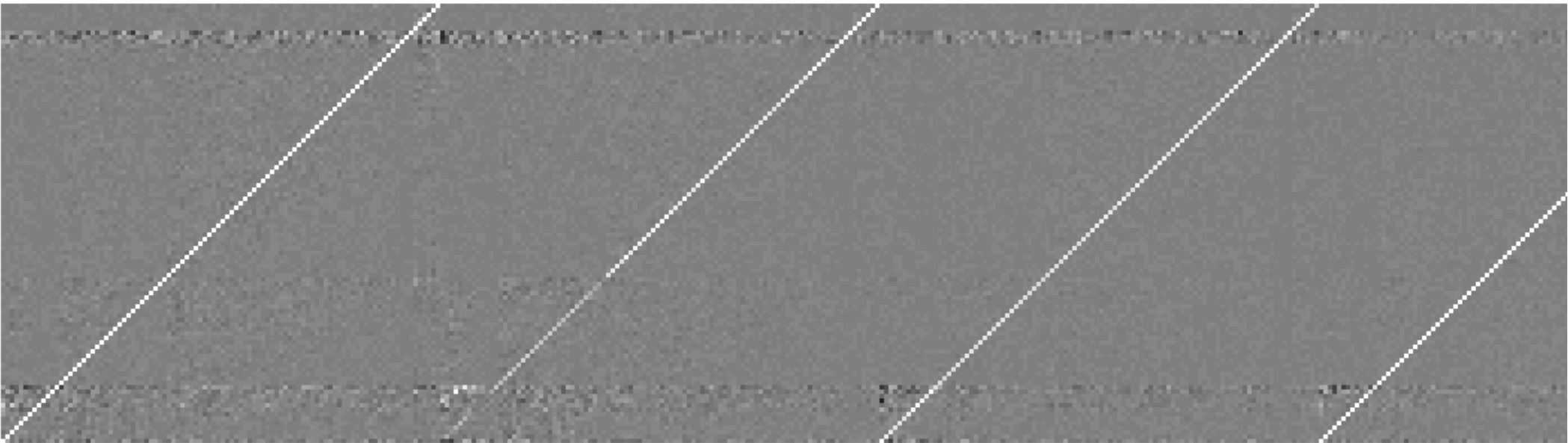
Conventional AO would
have control matrix
100 x 100 elements
Identity matrix

Optimal control adds
elements outside of
diagonal

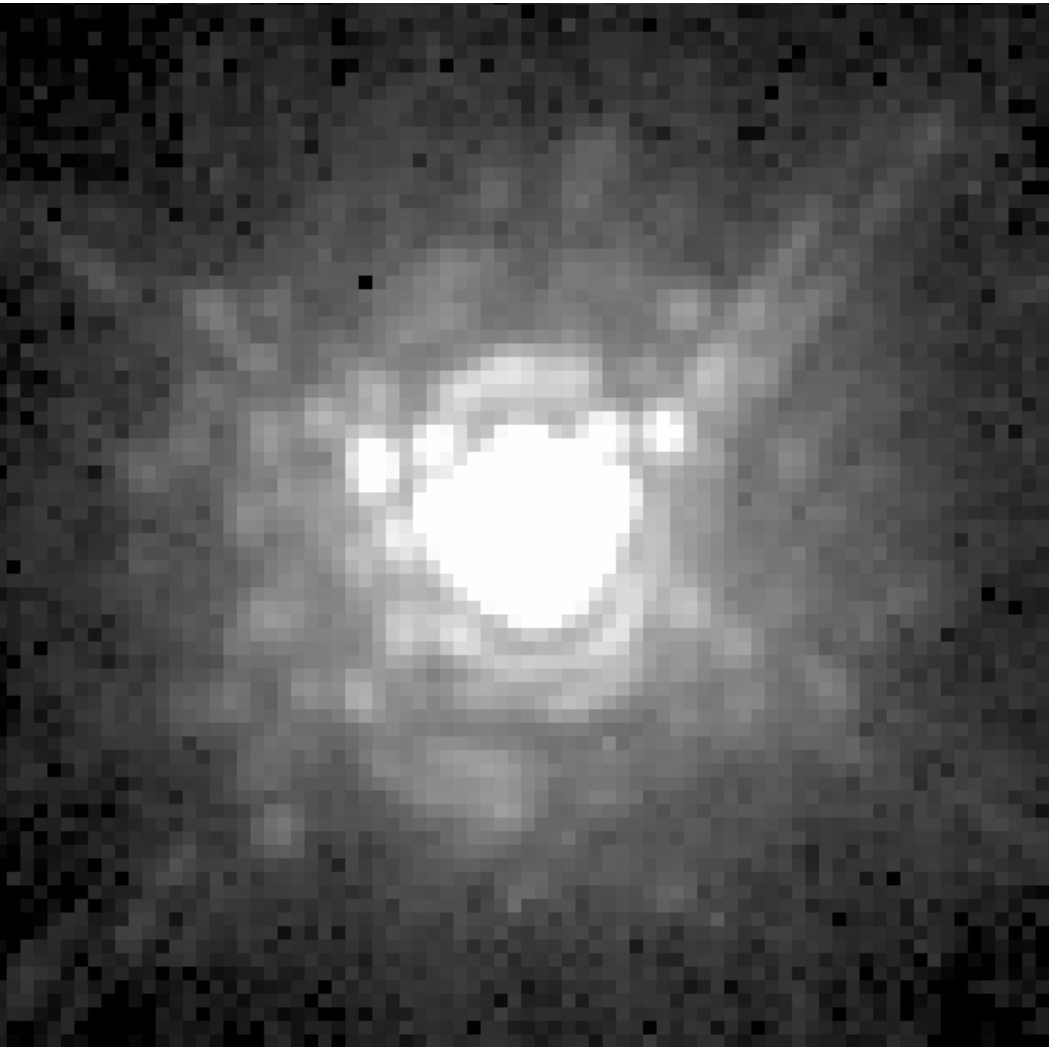
Predictive control adds
these blocks to control
matrix



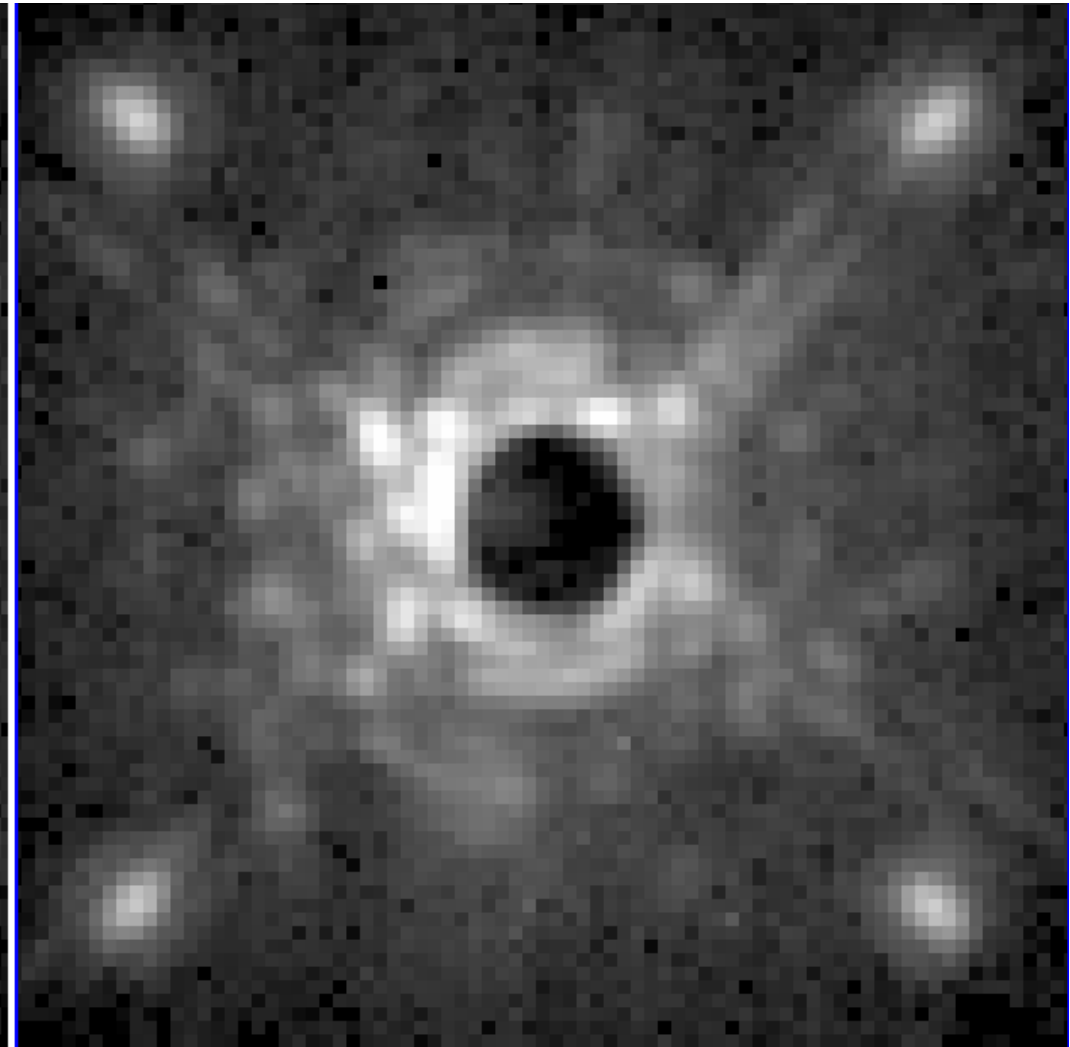
Prediction control matrix



On-sky results (2 kHz, 50 sec update)



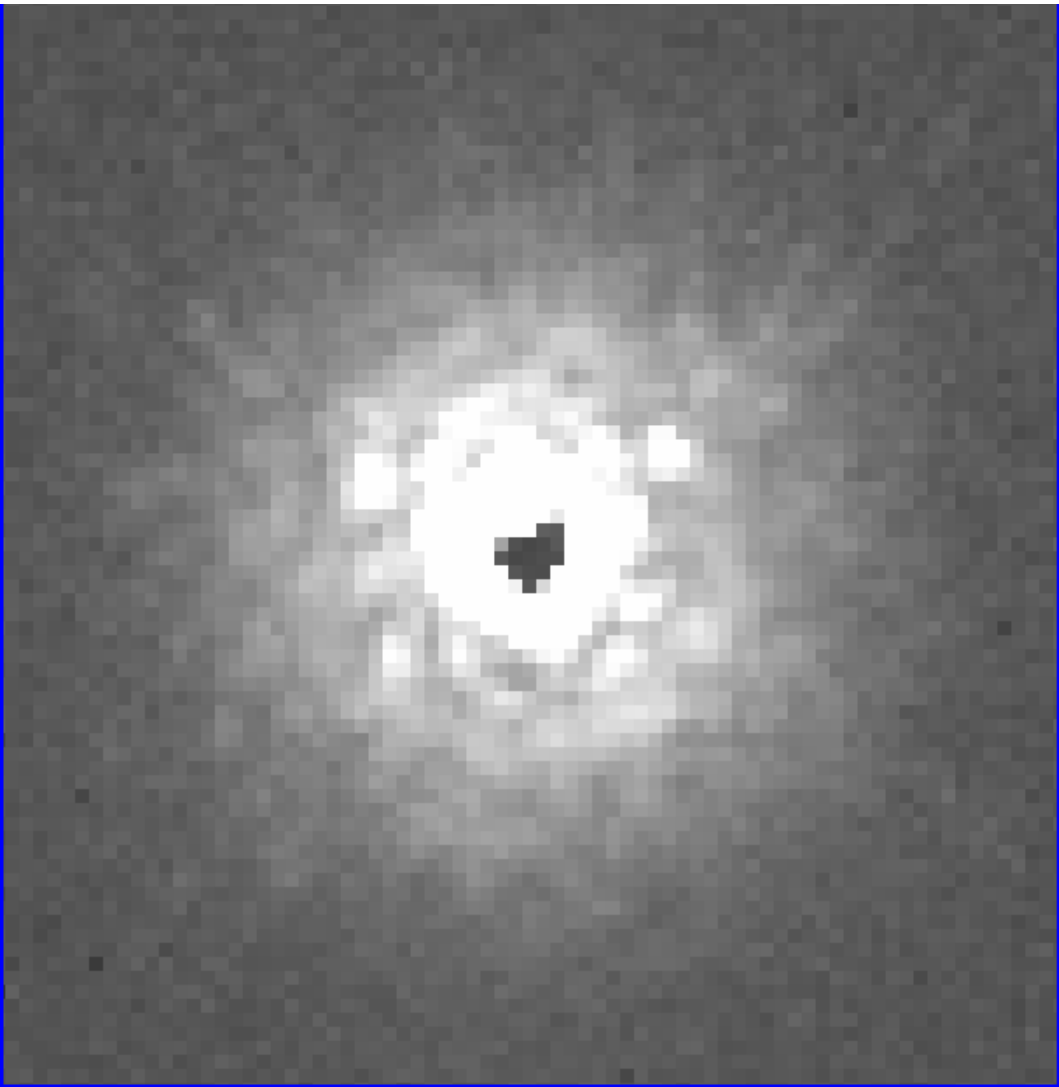
OFF (integrator, gain=0.2)



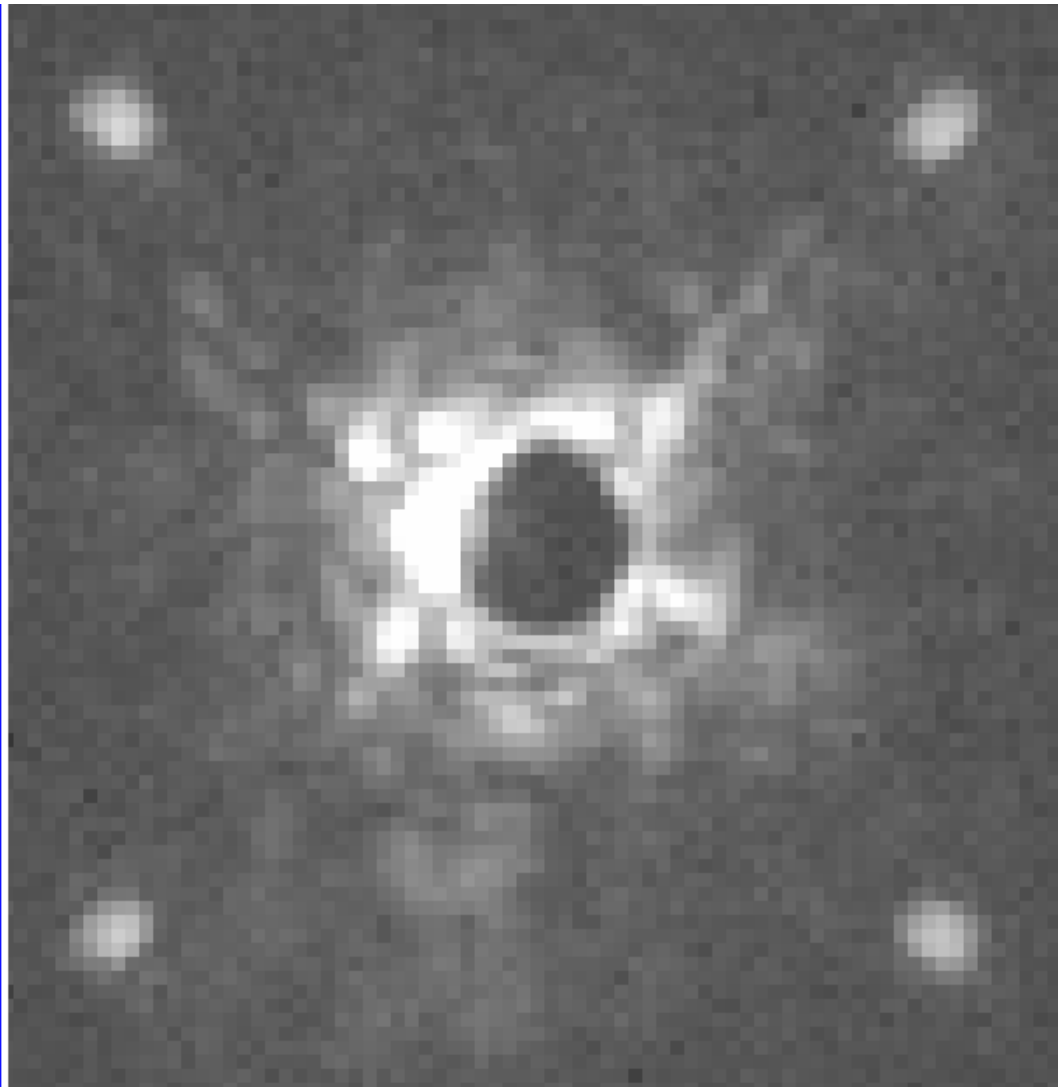
ON

Average of 54 consecutives 0.5s images (26 sec exposure), 3 mn apart
Same star, same exposure time, same intensity scale

Standard deviation improved by 2.5x



OFF (integrator, gain=0.2)

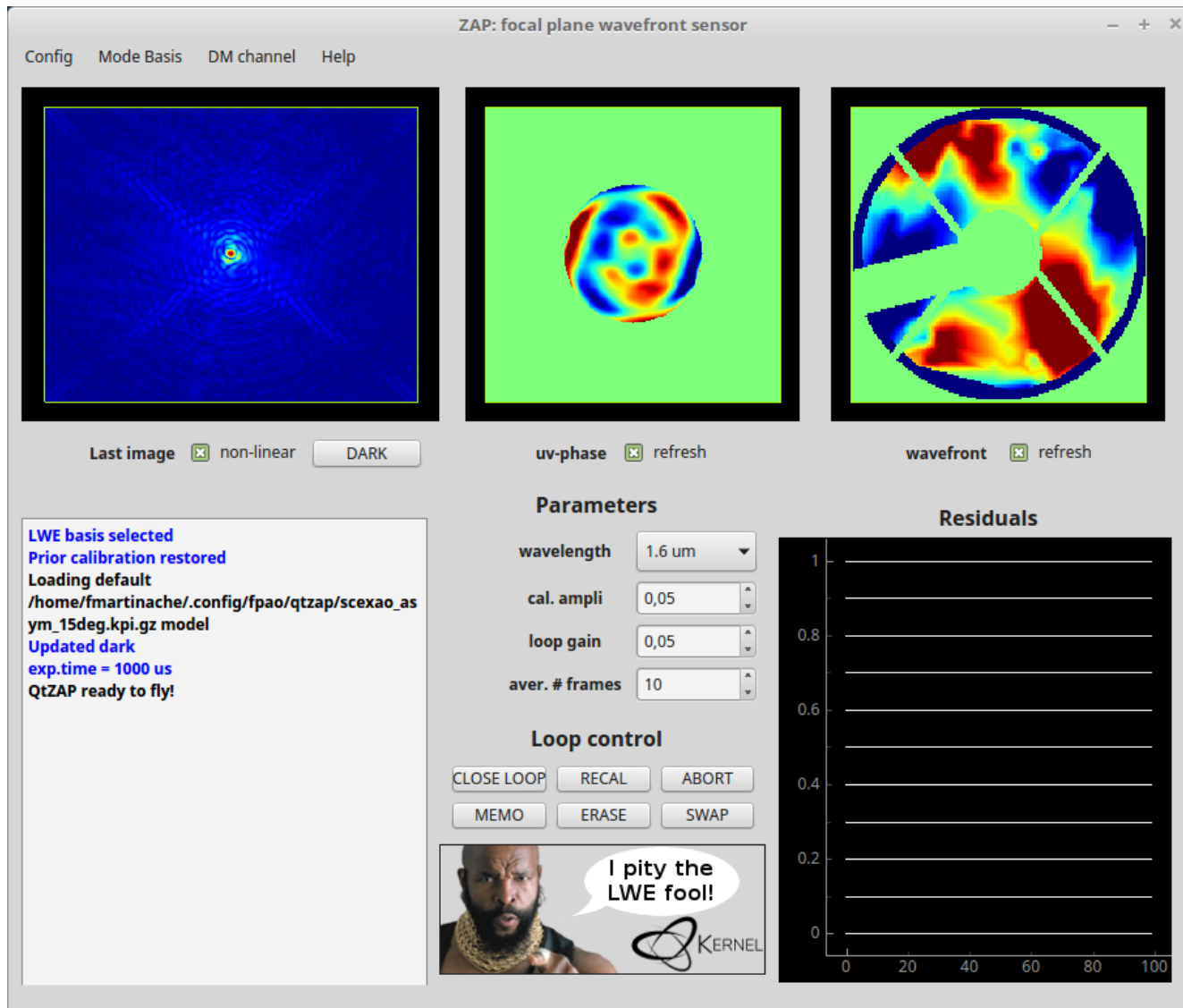


ON

Standard deviation of 54 consecutives 0.5s images (26 sec exposure), 3 mn apart
Same star, same exposure time, same intensity scale

Focal Plane WFS/C

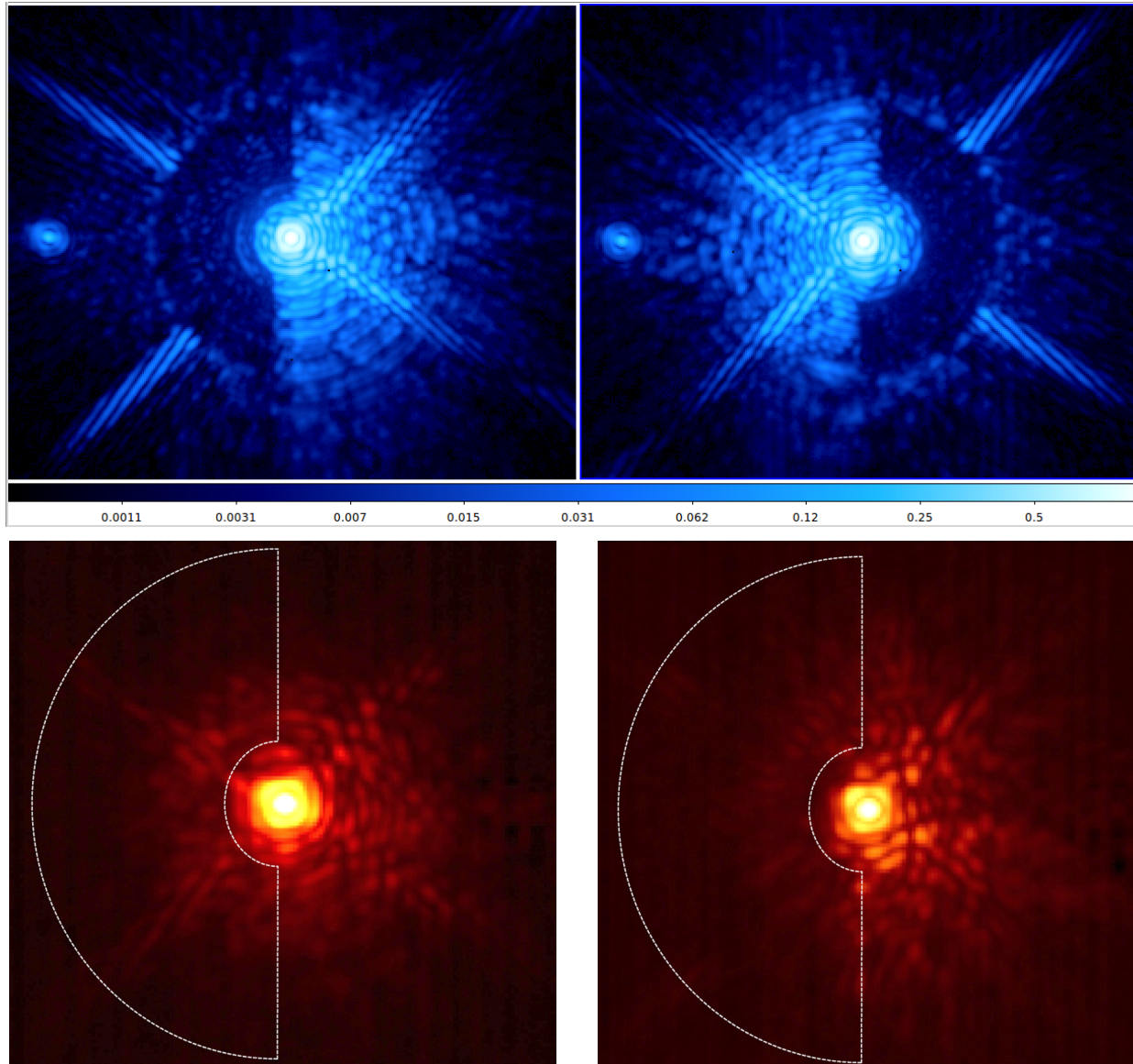
OCA/KERNEL – developed software



- Address NCPA
- Asymmetric mask (pupil)
- On-sky closed-loop control
- Focal plane based WFS
Low-order (Zernike and LWE) modes.
- mode compatible with coronagraphy in development



Speckle Control



Speckle nulling, in the lab and on-sky (no XAO).

Experience limited by detector readout noise and speed.

KERNEL project: C-RED-ONE camera.

From:

- 114 e- RON
- 170 Hz frame rate

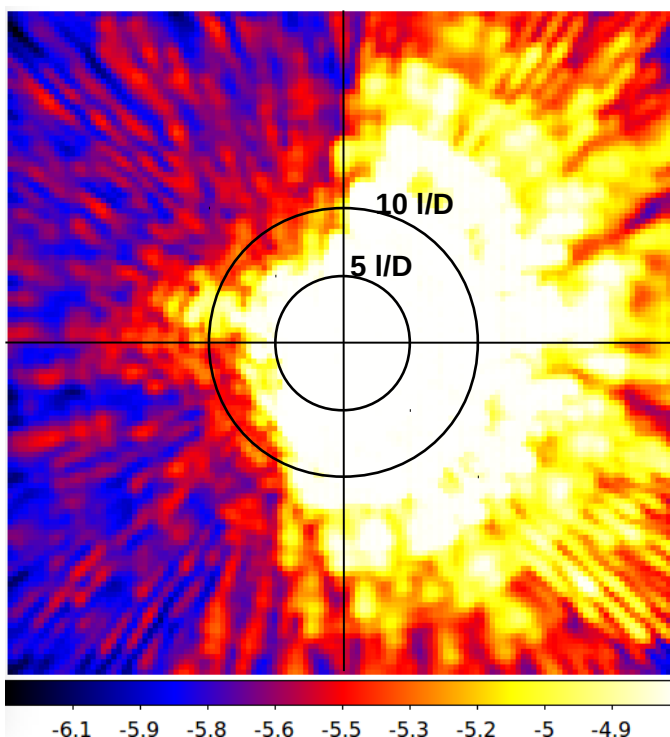
To:

- 0.8 e- RON
- 3500 Hz frame rate

Expect some updates

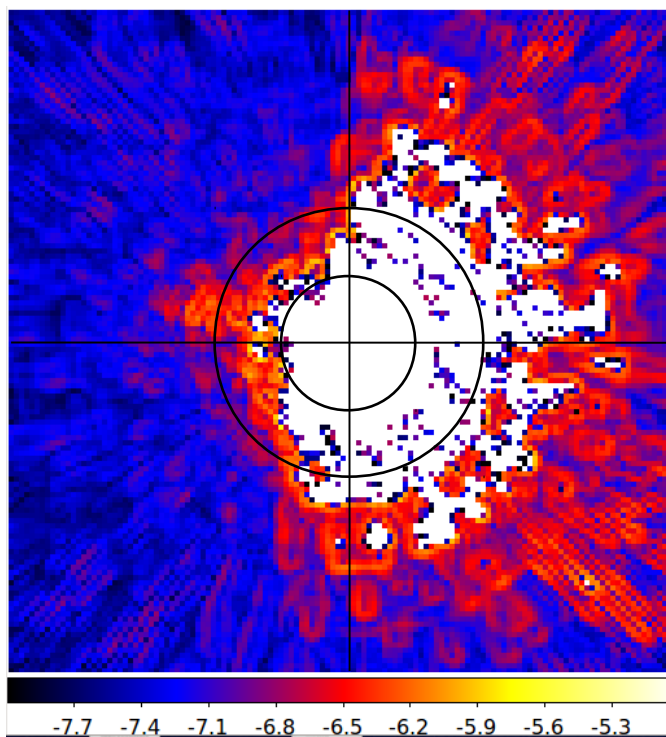
Conventional Lyot Coronagraph, Broadband light: 0.9-1.7 μm (62% wide band)

Raw Contrast



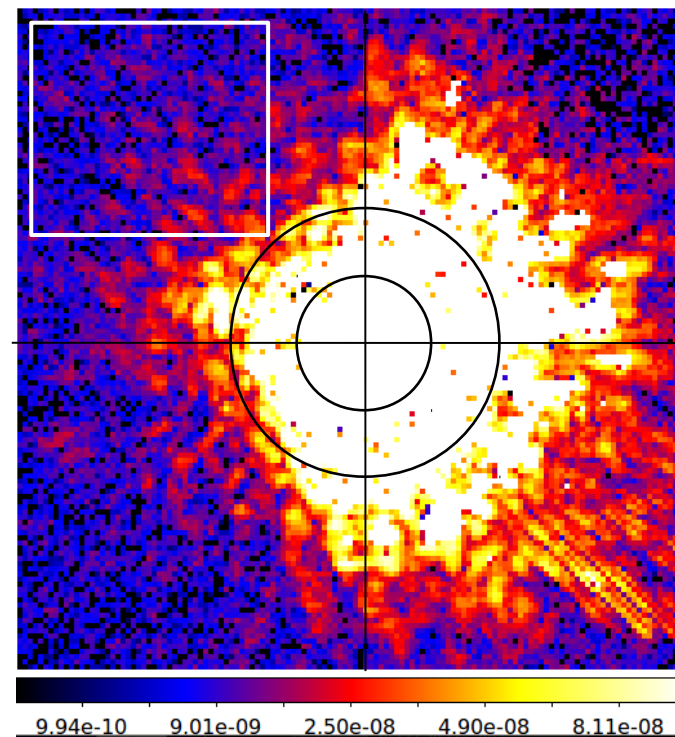
Average raw contrast [15-20 I/D]
= $2.3\text{e-}6$

Open-loop contrast stability

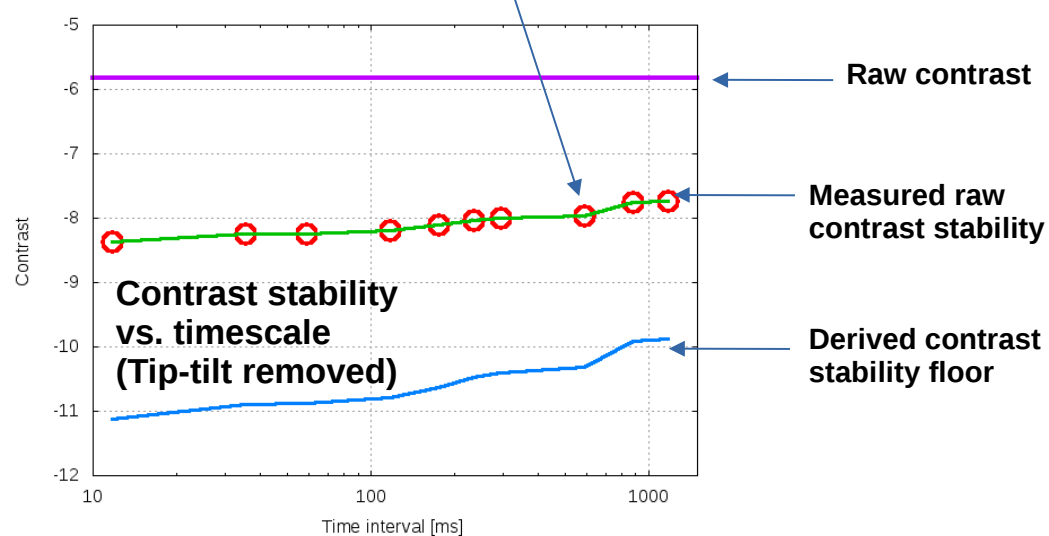
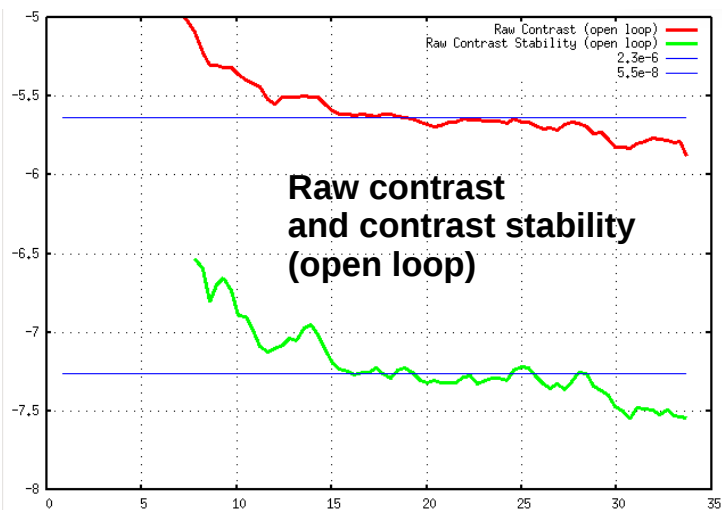


Average raw contrast stability [15-20 I/D]
= $5.5\text{e-}8$

Contrast stability over 0.6 sec



Average raw contrast stability [11-34 I/D]
= $1.1\text{e-}8$ (averaged value within white box)

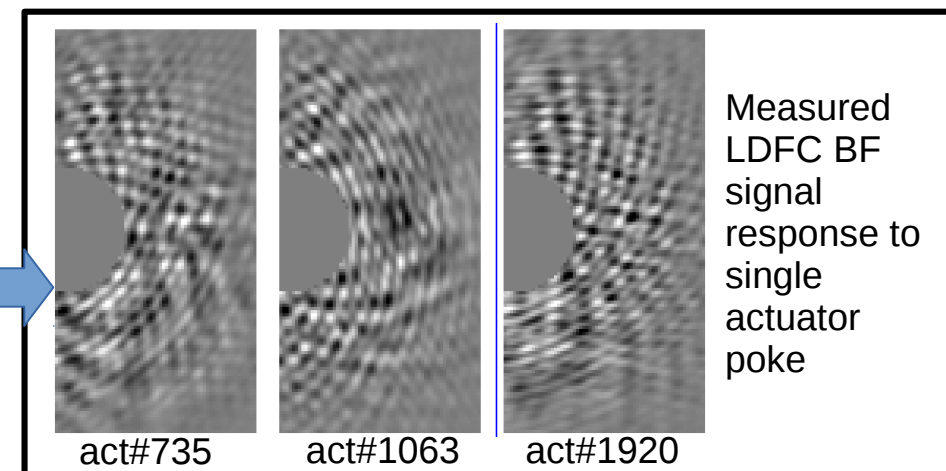
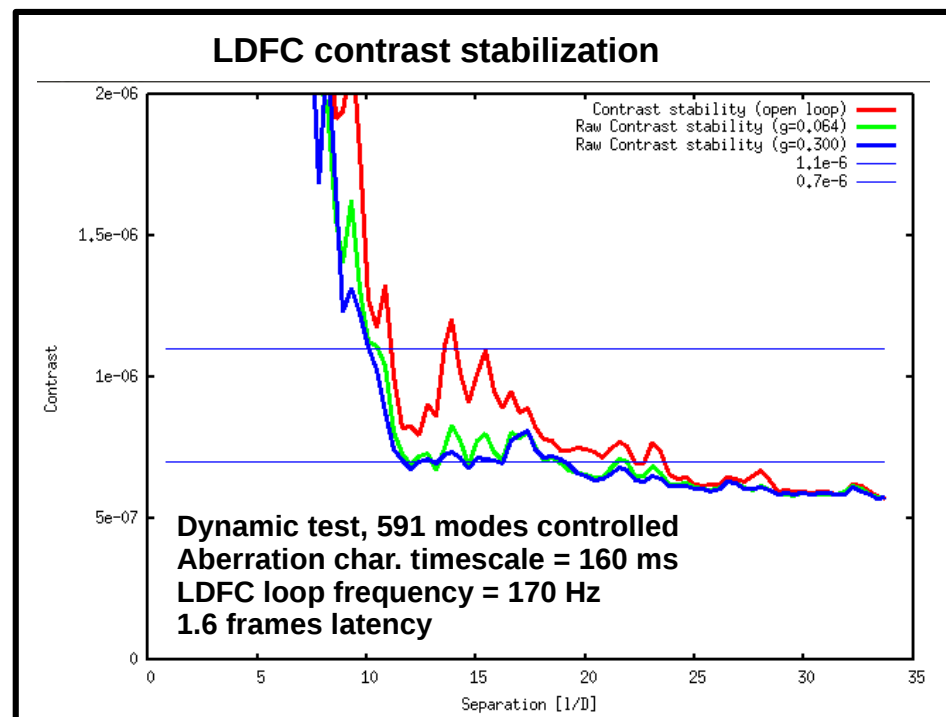
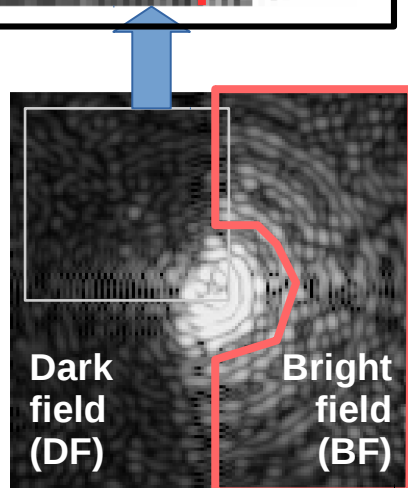
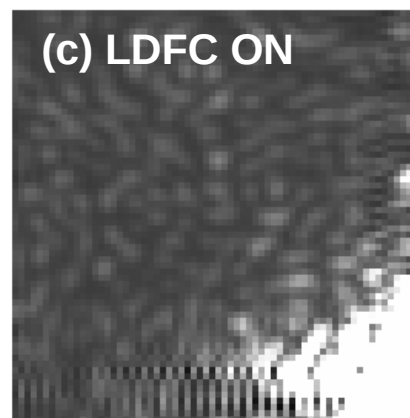
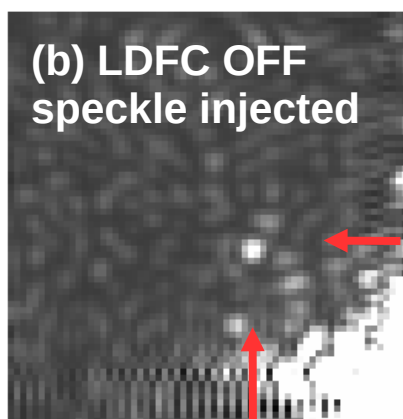
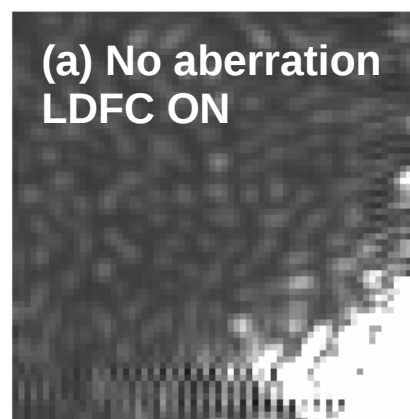


Linear Dark Field Control (internal source)

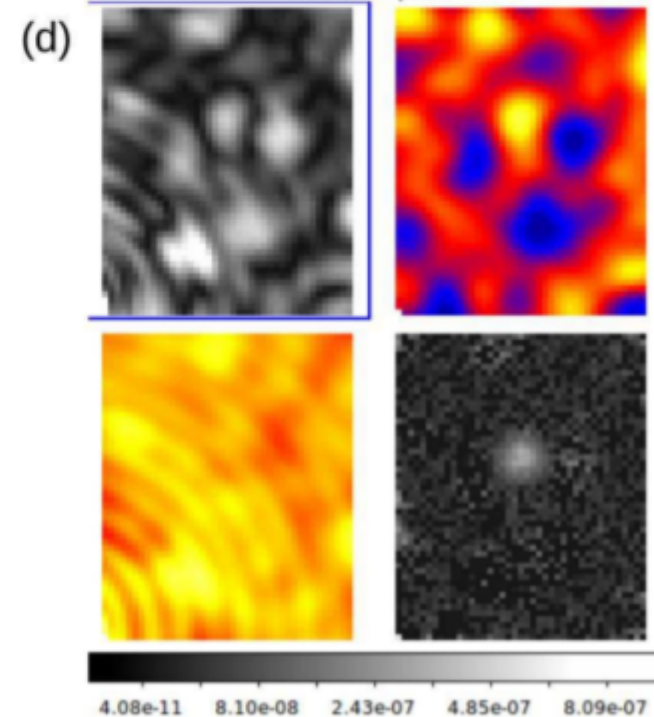
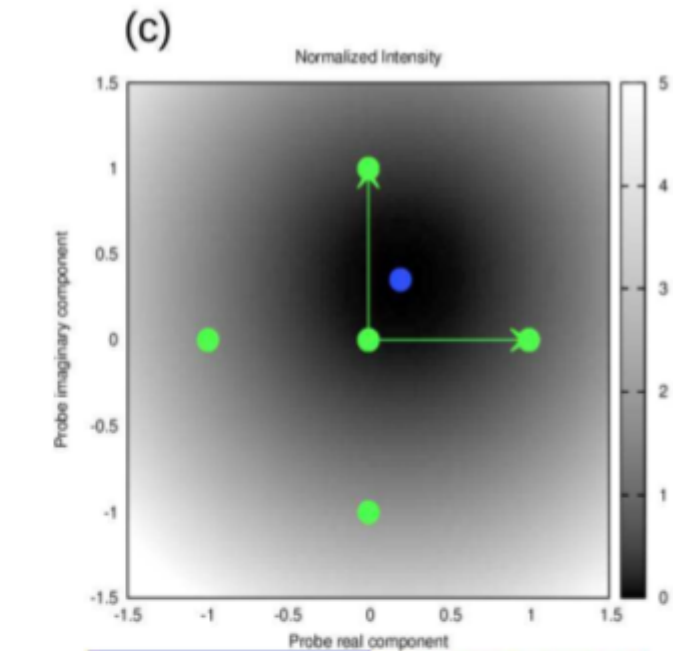
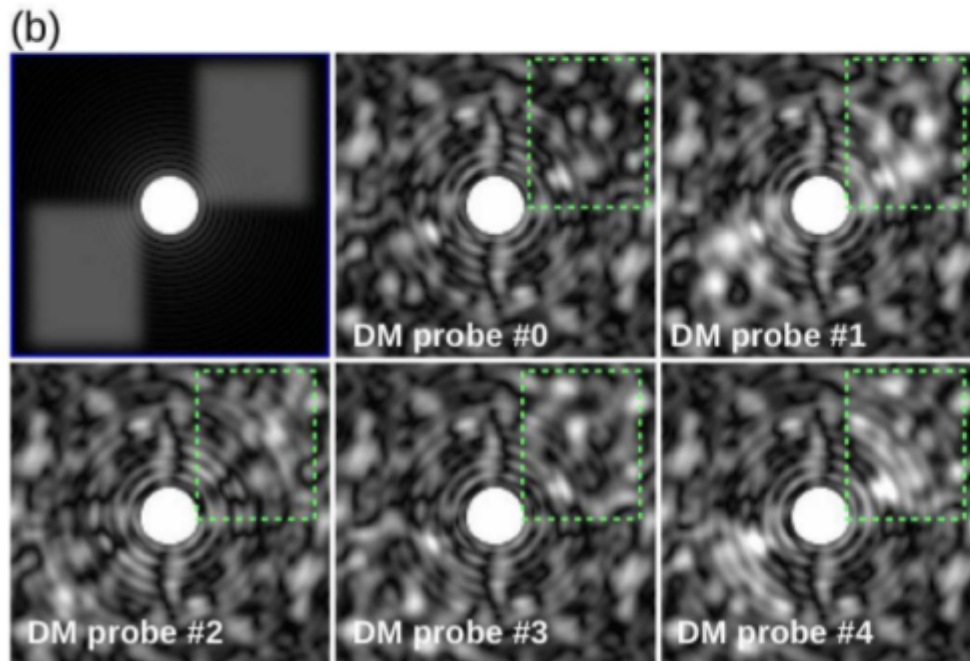
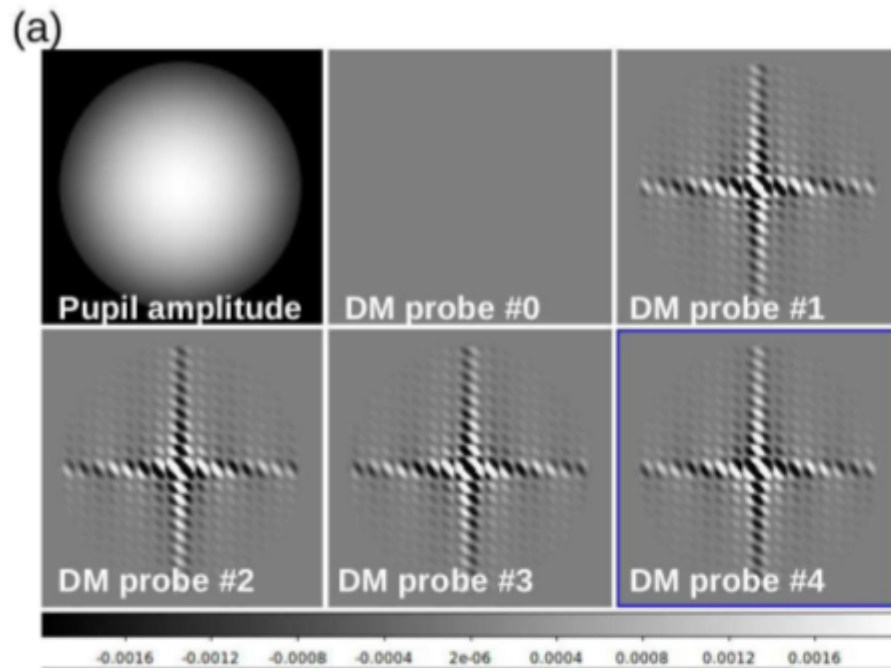
Near-IR spatial LDFC validation @ SCExAO

Frame rate = 170 Hz, Lyot coronagraph in near-IR
(1.55 μ m, 50nm wide band)

$C=1.25e-5$ speckle at 13 λ/D separation



Coherent Speckle Differential Imaging



User Interface and programming / scripting

Top-level ASCII-based menus
launch cacao programs in
tmux sessions

bash scripts assist with:

- sequencing between tasks
- hardware resources allocation
(CPU cores, GPUs, memory)



To be upgraded by more capable
KRAKENS python-based tool
(currently in development)

Users can code new
processes in C or Python



Developing well-documented
examples users can modify

Example control GUI (bash scripts)

```
TOP MENU
[Active conf = ] [ Mon Apr 10 12:48:10 UTC 2017 ]

DM CHANNELS AND OUTPUT (dcomb process)

S [00] Set DM index
dmax [50] Set DM x size (if modal control, = number of modes)
dmax [50] Set DM y size (if modal control)

noLink Auto-configure: main DM (no link) -> DM actuators are physical actuators
dnoLink Auto-configure: DM output linked to other loop -> DM actuators represent modes

DmodeM DM is ZONAL Modes constructed from spatial DM actuators (select to toggle to MODAL)

d2dmModel [ OFF ] DM-to-DM is OFF (select to activate virtual (modal) DM to physical DM mode)

dmuref1 [ OFF ] CPU-based dcomb output WFS ref is OFF (select for DM output applied as WFS offset)
dmurefM [ MISSING ] WFS Resp Matrix aol0_dmurefM -> empty
dmurefD [ MISSING ] WFS zp output stream aol0_dmurefD -> empty

dmvlt0 [ ON ] De-activate DM volt output [-> dmvolt]

dcombave [0] DM combination averaging mode

setDMdelayval [0] Set DM delay value [us]
setDMdelayDM [DM delay is OFF] press to toggle DM delay to ON state

initDM (re)-START DM comb process (-> dm0disp00..07 dm0disp)

2 -> AD CONFIGURE AND CONTROL

C0 CALIBRATE SYSTEM (CPAmax = 22.0) RM, CH -> staged (compute masks)
C01 CALIBRATE SYSTEM (CPAmax = 22.0) RM -> staged (Re-use masks)
C01 RM -> CH (staged)
AD ADOPT CALIBRATION: staged -> conf, SharedMem

M load all (Memory)
C (Configure/Link AO loop
CH Modes and Control Matrix

L Control AO (Loop)

3 -> PREDICTIVE CONTROL

P Predictive Control
F Filtering

4 -> TEST AND MONITOR

L List running AO processes, locks
T Test mode: simulated AO system
V View / monitor

5 -> DATA LOGGING / ANALYSIS

R Record / analyze

6 -> CUSTOM EXTERNAL SCRIPTS

A Align
MC Hardware Control

< Select > < Exit >
```

```
ALIGNMENT - LOOP SCEAOPyWFS (0)

TT loop is : OFF
Pcam loop is : OFF
Pyw Filter : 3

1 -> Pyramid modulation

pyfr05 freq = 0.5 kHz
pyfr10 freq = 1.0 kHz
pyfr15 freq = 1.5 kHz
pyfr20 freq = 2.0 kHz
pyfr25 freq = 2.5 kHz
pyfr30 freq = 3.0 kHz
pyfr35 freq = 3.5 kHz

pymoda05 modulation amplitude = 0.05 (modulation radius = 12.5 mas)
pymoda10 modulation amplitude = 0.10 (modulation radius = 25.0 mas)
pymoda15 modulation amplitude = 0.15 (modulation radius = 37.5 mas)
pymoda20 modulation amplitude = 0.20 (modulation radius = 50.0 mas)
pymoda25 modulation amplitude = 0.25 (modulation radius = 62.5 mas)
pymoda30 modulation amplitude = 0.30 (modulation radius = 75.0 mas)
pymoda35 modulation amplitude = 0.35 (modulation radius = 87.5 mas)
pymoda40 modulation amplitude = 0.40 (modulation radius = 100.0 mas)
pymoda45 modulation amplitude = 0.45 (modulation radius = 112.5 mas)
pymoda50 modulation amplitude = 0.50 (modulation radius = 125.0 mas)
pymoda55 modulation amplitude = 0.55 (modulation radius = 137.5 mas)
pymoda60 modulation amplitude = 0.60 (modulation radius = 150.0 mas)
pymoda65 modulation amplitude = 0.65 (modulation radius = 162.5 mas)
pymoda70 modulation amplitude = 0.70 (modulation radius = 175.0 mas)
pymoda75 modulation amplitude = 0.75 (modulation radius = 187.5 mas)
pymoda80 modulation amplitude = 0.80 (modulation radius = 200.0 mas)
pymoda85 modulation amplitude = 0.85 (modulation radius = 212.5 mas)
pymoda90 modulation amplitude = 0.90 (modulation radius = 225.0 mas)
pymoda95 modulation amplitude = 0.95 (modulation radius = 237.5 mas)
pymoda100 modulation amplitude = 1.00 (modulation radius = 250.0 mas)

pyfilt1 PyWFS filter 1 (Open)
pyfilt2 PyWFS filter 2 (700 nm BW)
pyfilt3 PyWFS filter 3 (850 nm BW)
pyfilt4 PyWFS filter 4 (750 nm, 50 nm BW)
pyfilt5 PyWFS filter 5 (850 nm, 25 nm BW)
pyfilt6 PyWFS filter 6 (850 nm, 40 nm BW)

pypick01 PyWFS pickoff 01 (Open)
pypick02 PyWFS pickoff 02 (Silver mirror)
pypick03 PyWFS pickoff 03 (50/50 splitter)
pypick04 PyWFS pickoff 04 (650 nm SP)
pypick05 PyWFS pickoff 05 (700 nm SP)
pypick06 PyWFS pickoff 06 (750 nm SP)
pypick07 PyWFS pickoff 07 (800 nm SP)
pypick08 PyWFS pickoff 08 (850 nm SP)
pypick09 PyWFS pickoff 09 (750 nm LP)
pypick10 PyWFS pickoff 10 (800 nm LP)
pypick11 PyWFS pickoff 11 (850 nm LP)
pypick12 PyWFS pickoff 12 (Open)

2 -> Pyramid TT align ( 90.3 mas/V )

Current position ( scale = 90.3 mas/V ) = -5.623 -4.403
tz Zero TT align (-5.0 -5.0)
ttr Move to TT reference position [ -5.268 -4.801 ]
tts Store current position as reference
tst0 alignment step = 0.05
tst1 alignment step = 0.1
tst2 alignment step = 0.2
tst3 alignment step = 0.5
tst4 TT x -0.2 (PyWFS bottom left)
tst5 TT x +0.2 (PyWFS top right)
tst6 TT y -0.2 (PyWFS top left)
tst7 TT y +0.2 (PyWFS bottom right)
tst8 ===== Start TT align =====
tst9
tst0 py TT loop gain = 0.1
tst1 Monitor TT align tmax session

3 -> Pyramid Camera Align ( 5925 steps / pix )

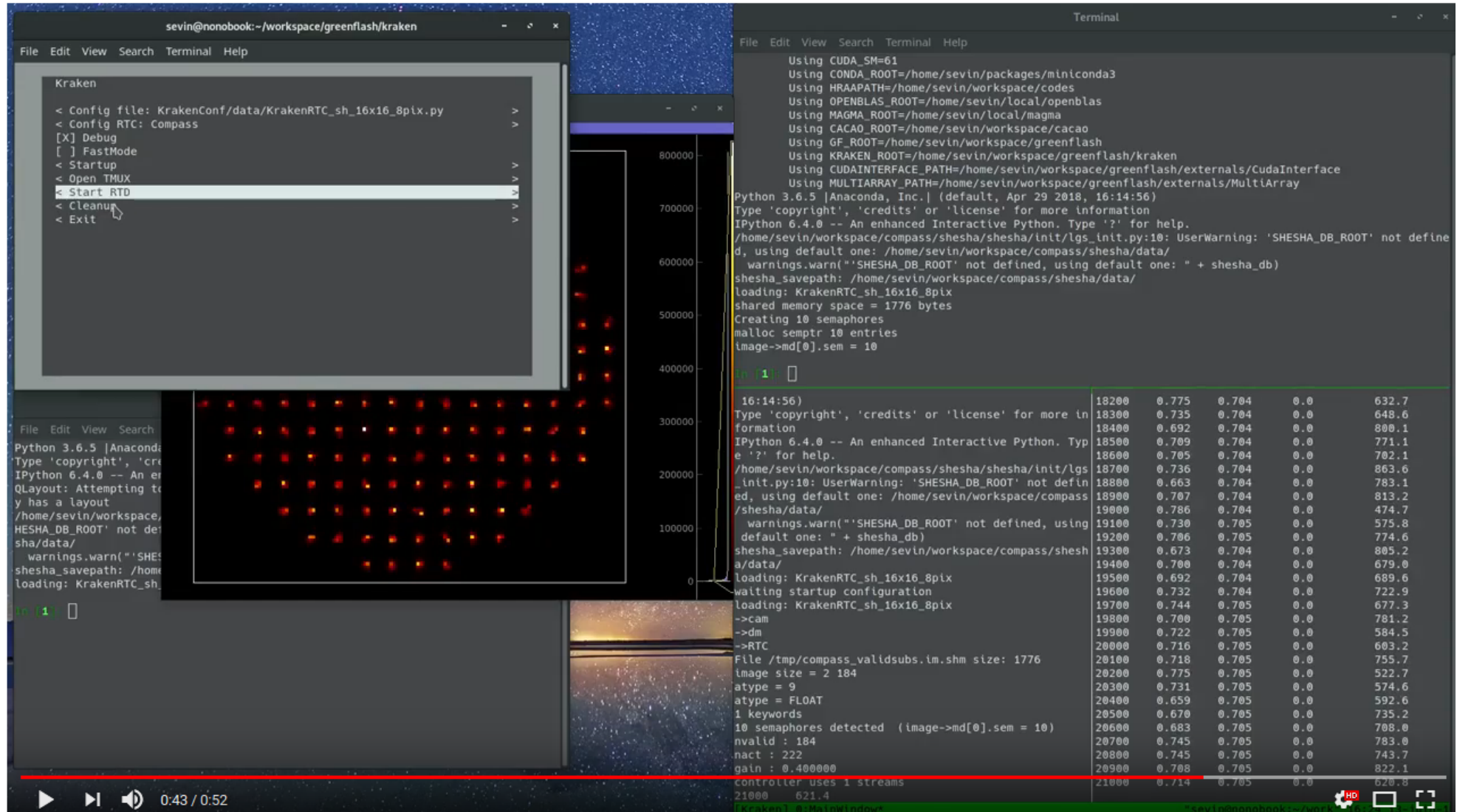
pz Zero Pcam align ( 143736 60198 )
pz0 alignment step = 50000
pz1 alignment step = 10000
pz2 alignment step = 3000
pz3 alignment step = 1000
pzn Pcam x -10000 (right)
pzn Pcam x +10000 (left)
pzn Pcam y -10000 (top)
pzn Pcam y +10000 (bottom)
pz1 ===== Start Pcam align =====
pz2
pz3 Pcam loop gain = 0.4
pzn Monitor Pcam align tmax session

4 -> DM Flatten

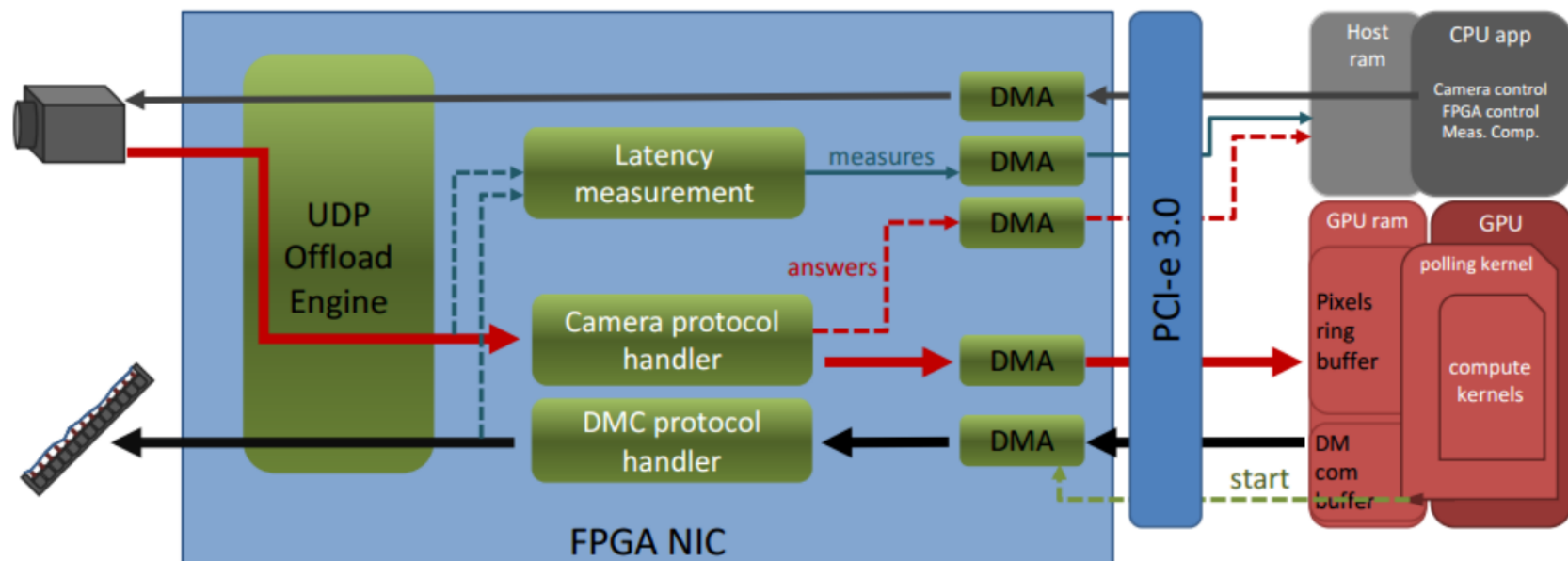
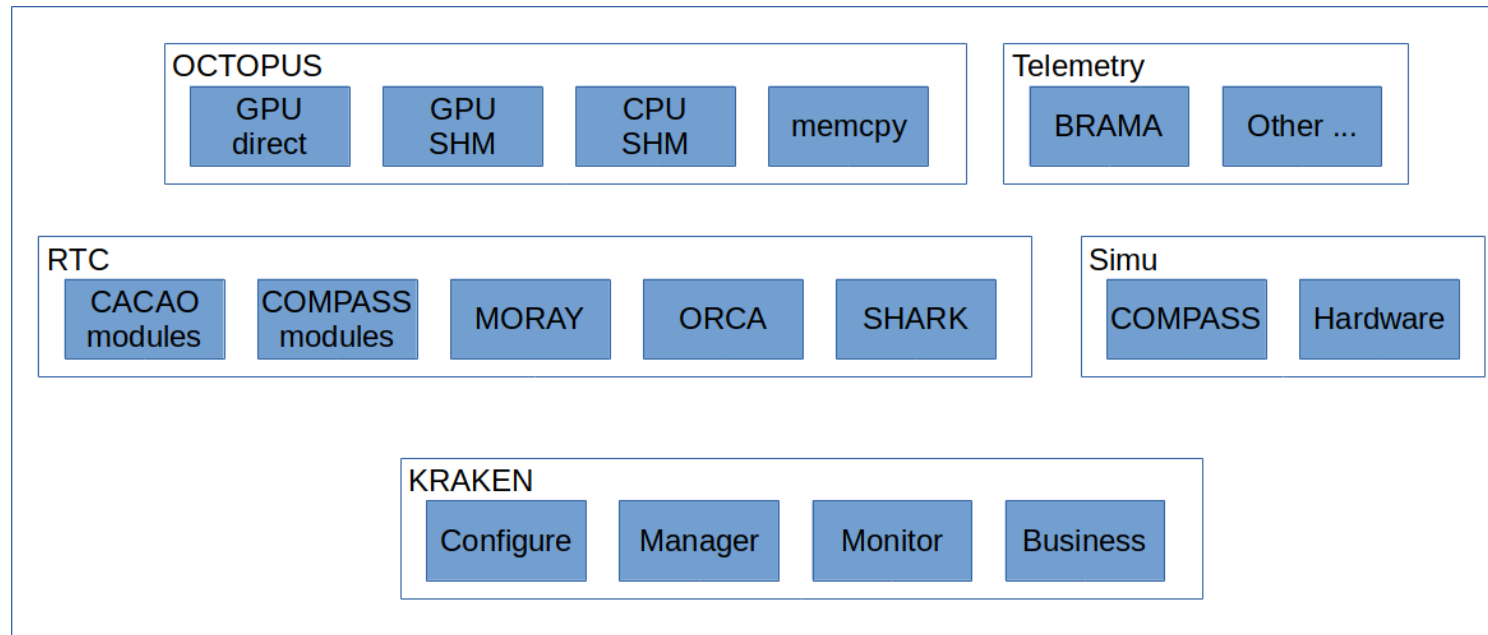
f1 Flatten DM for pyWFS
f1x End Flatten DM process
f1x Remove Flatten DM solution
f1x Apply flatten DM solution
f1x Monitor DM flatten tmax session

< Top > < Exit >
```

Software ecosystem integration & High performance hardware abstraction



Software ecosystem integration & High performance hardware abstraction



Thank you !

<https://github.com/cacao-org/cacao>