



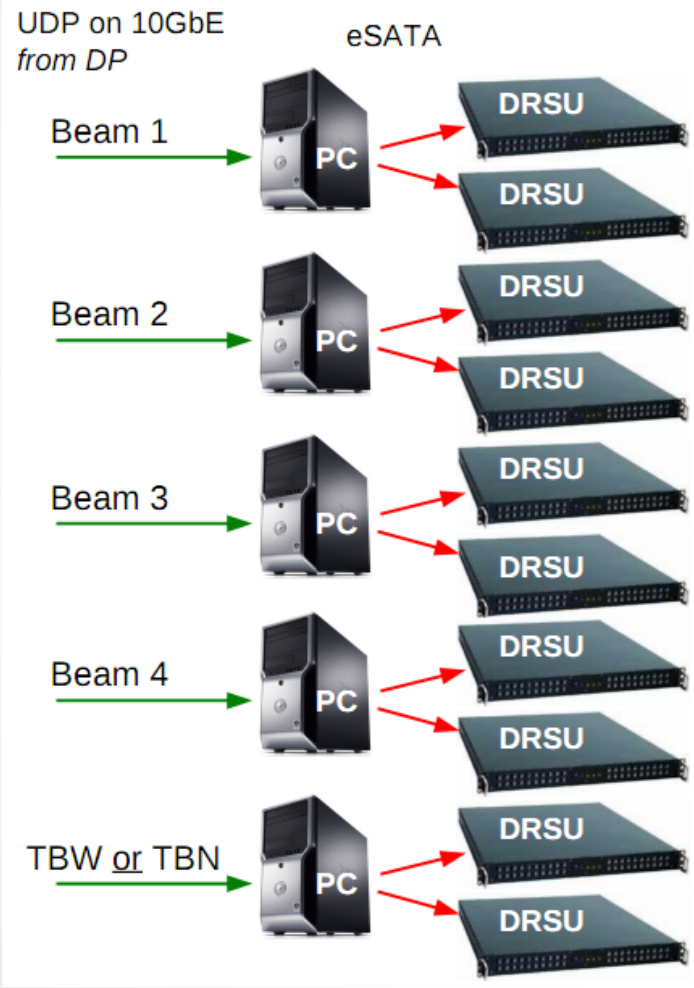
# Real-Time Data Analysis Using LWA-1 Data Recorders

Christopher Wolfe

# Topics

- Overview of LWA-1 Data Recording (DR) System
- Current Observation Data Flow and Limitations
- Proposed DR enhancements
- Strategies and Limitations

# Data Recorder Overview



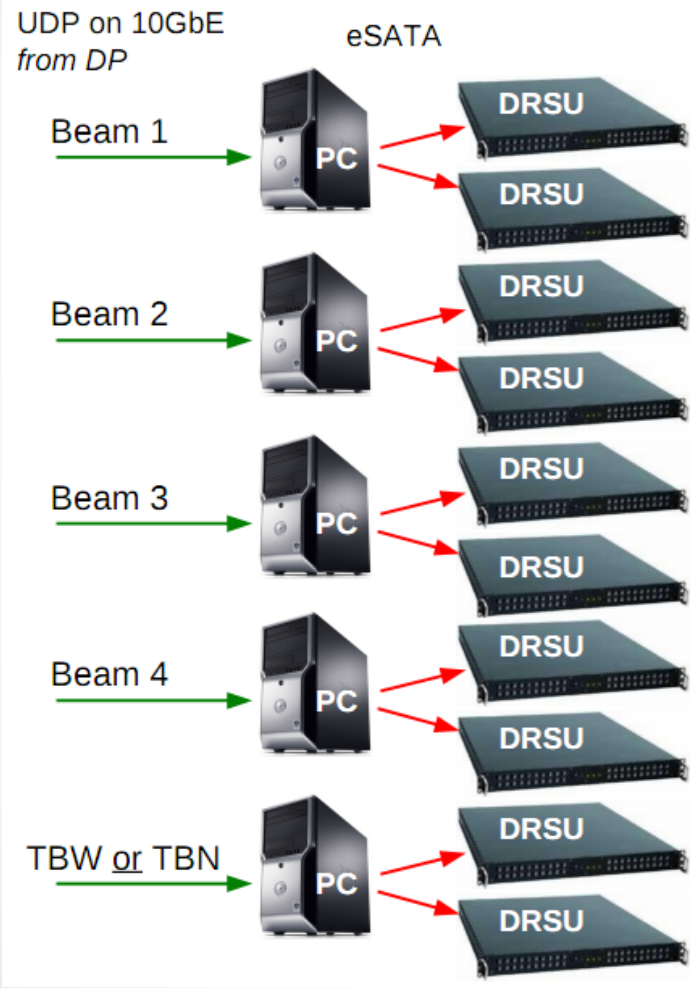

## UDP Packet Recorder

- Five Data Recorders per LWA station
- Each composed of a PC (DRPC) and 1 or 2 storage units (DRSUs)
- No packet reordering or processing performed

- Performance limitations
  - DRSU R/W Bandwidth
  - Network I/O ops/s

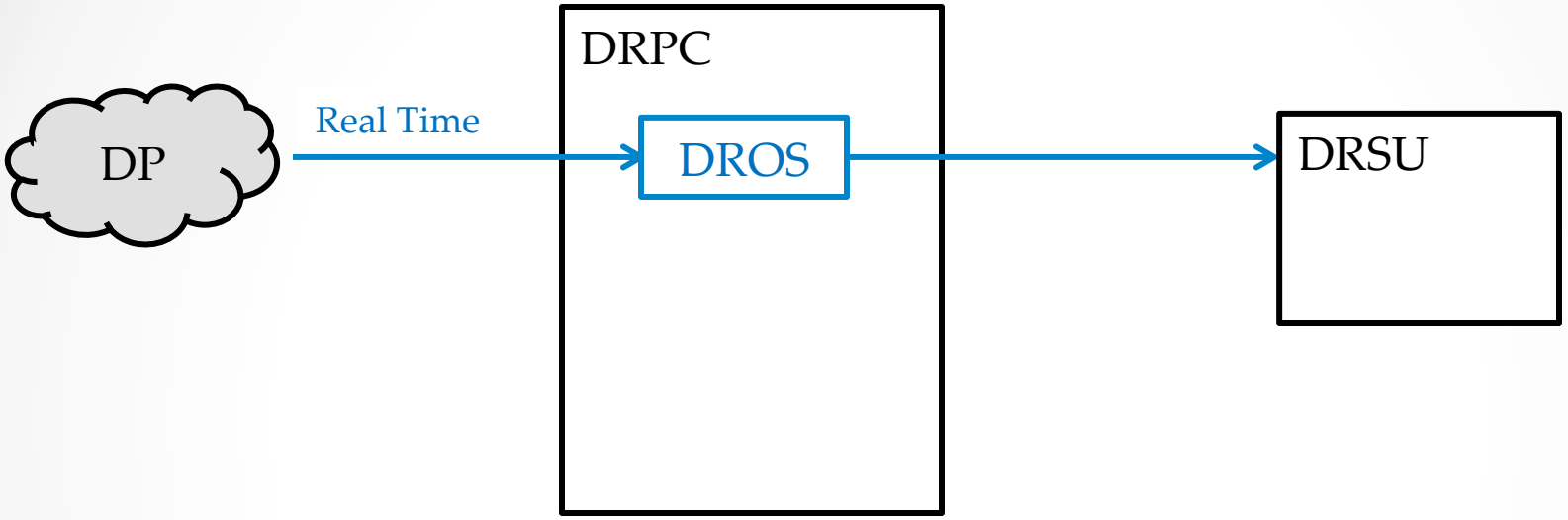
- Can record any UDP stream
  - up to 112 MiB/s
  - Up to ~150,000 IO/s

# Data Recorder Overview



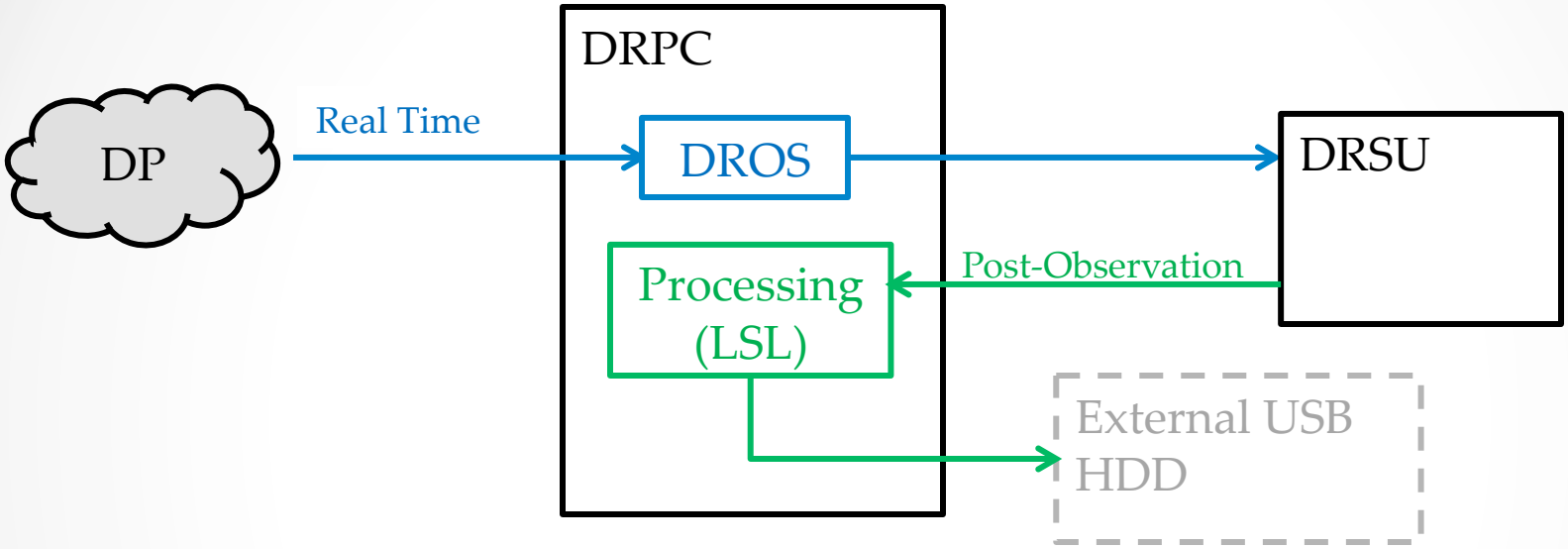

- up to 112 MiB/s
  - Det. by DRSU (120 MiB/s)
  - Flash-based drives overcome limitation but are prohibitively expensive ($15,000/ DRSU)
- Up to ~150,000 IO/s
  - Det. by DP data format
  - Could reduce burden by aggregating up to jumbo frame size, but would require changes to ICD, programming

# Current Observation Data Flow




- Data recorded in real-time to DRSU

# Current Observation Data Flow



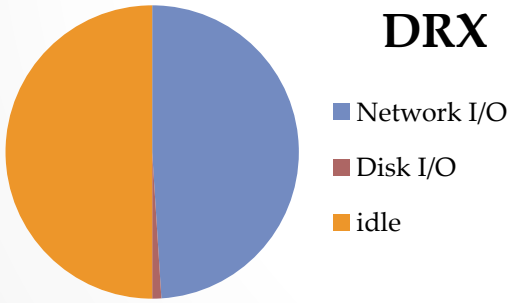

- Recorded data is transferred to external drive via copy
- (optionally) processed with LWA Software Library (LSL)
  - Channelization
  - Integration
- Duty Cycle limited to $\frac{t_{obs.}}{t_{obs.}+t_{proc.}} < 14\%$, (21% with no processing) [2]
- Long delay between observation and meaningful feedback
- Tedious management problem for operator / UNM

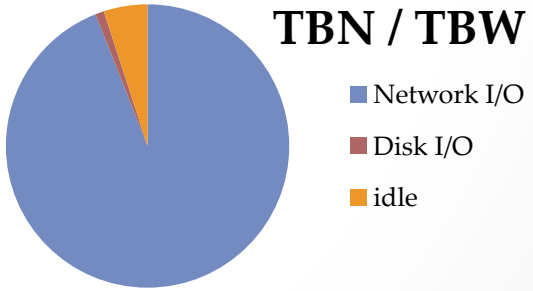
# Inspiration

## Underutilized CPU

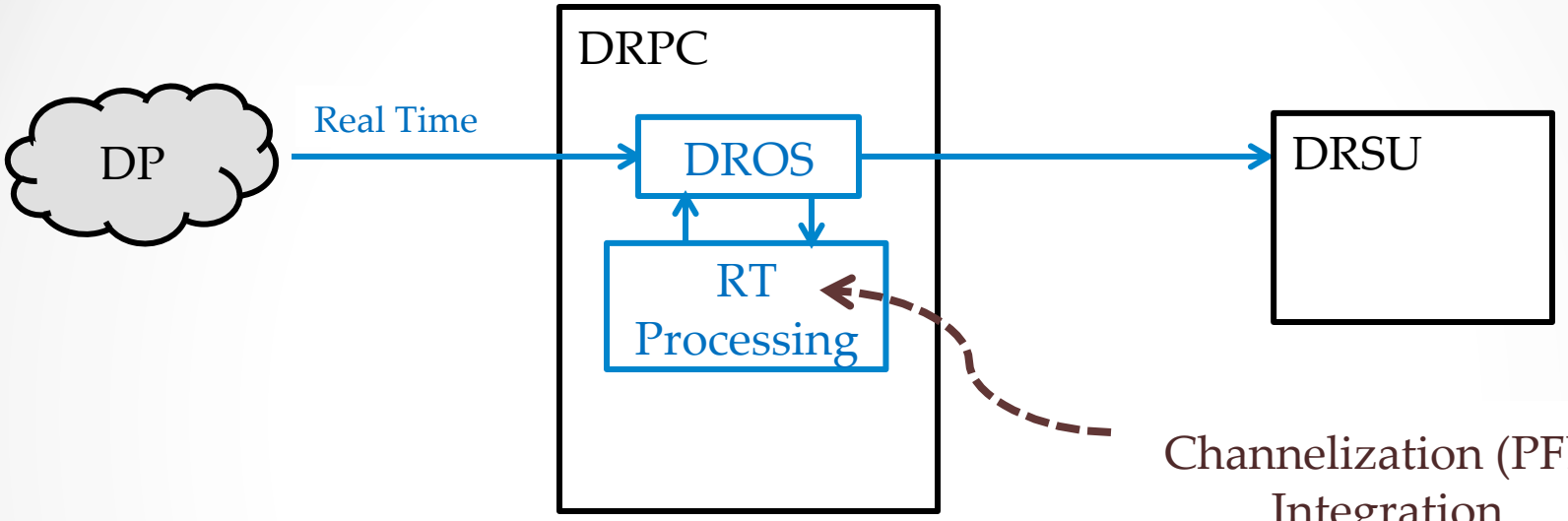
~35-50% CPU time is spent waiting for write completion in DRX mode (only about 5% for TBN)





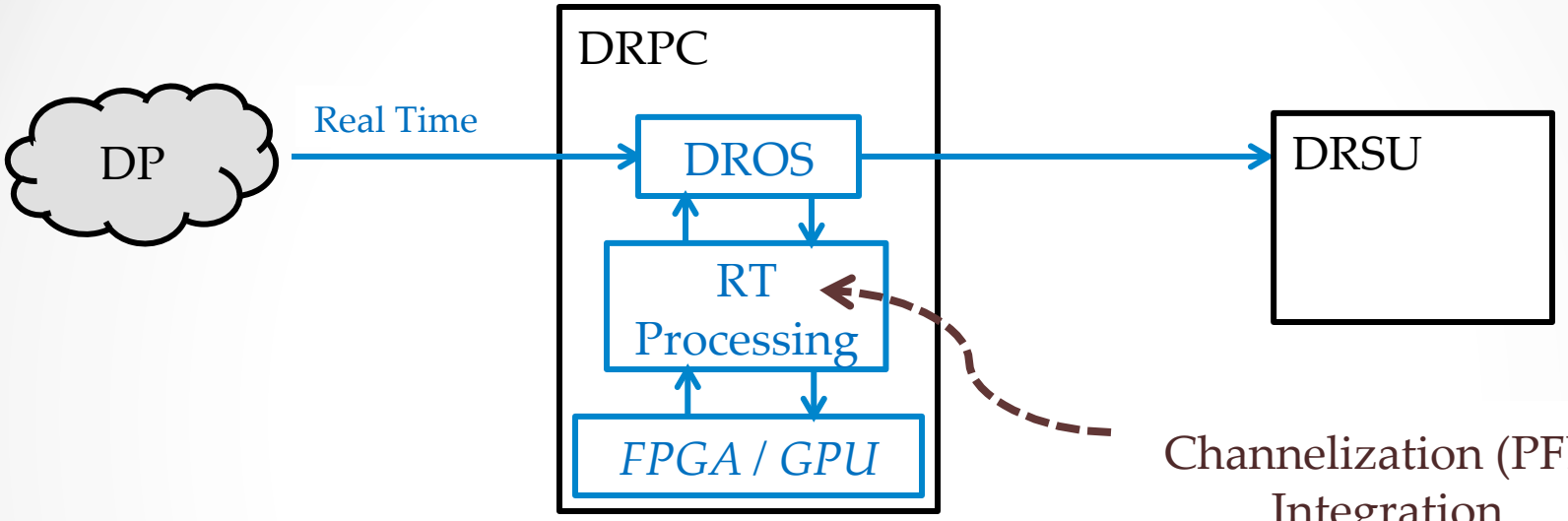



# Proposed Observation Data Flow (A)



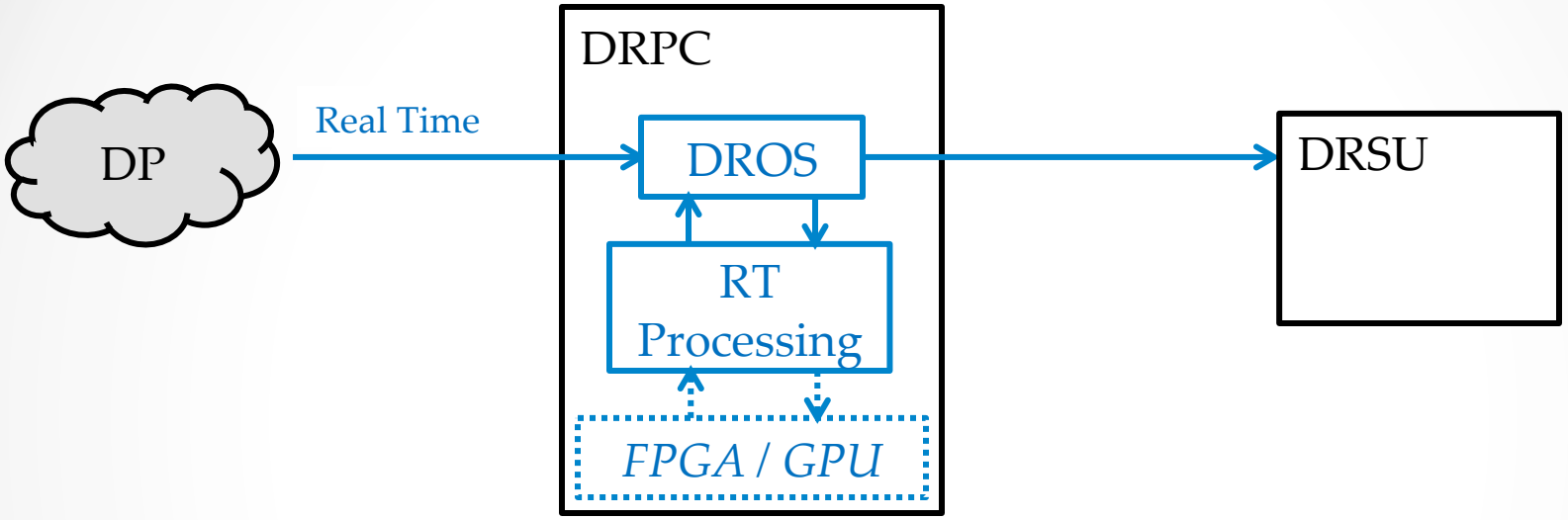

- Purely software option

# Proposed Observation Data Flow (B/C)




- Hardware accelerated option

# Proposed Observation Data Flow




- Duty Cycle limited by # DRSUs, and effective compression ratio of processing
  - (@ 30:1 compression, 100% duty for 2.11 weeks on a single DRSU)
- Comparable HDD Cost
  - $96/TB DRSUx2, ($147/TB for DRSU)
  - $50-$80/TB USB external
- Eases data management efforts
- Analysis available sooner, etc.

# Real-time Processing Feasibility

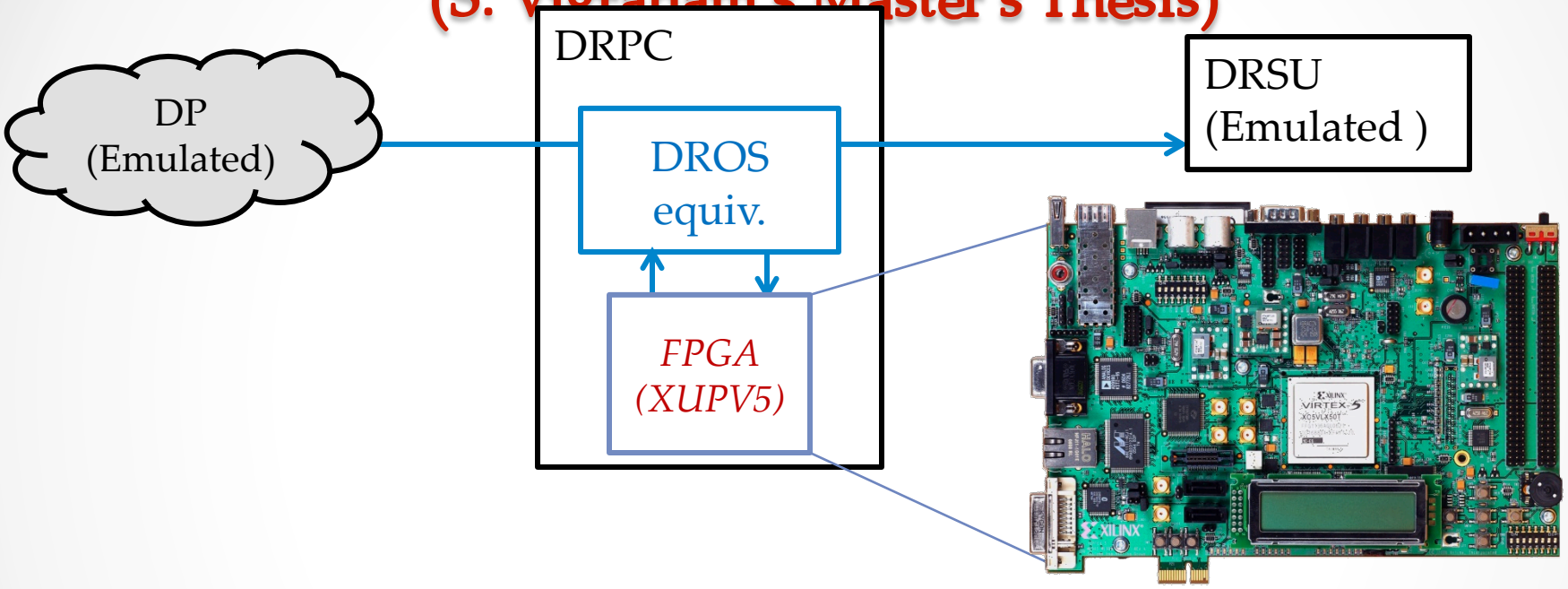
- Underutilized CPU time
  - $T_{f,DRX}$ ≅35% DRX ~ 21k IOPS
  - $T_{f,TBW}$ ≅10% TBW ~ 84k IOPS
  - $T_{f,TBN}$ ≅5% TBN ~ 112k IOPS
- For DRX, rate-limited by DRSU (~120-130MiB/s)
  - Data is internally moved @ ~465MiB/s
- Characterize potential
  - CPU free-time($T_f$) vs. time writing($T_w$)
  - Upper Bound for $T_f$=0.72 $T_w$ (w/ no overhead, $10^6$:1 reduction)
- Implication
  - processing must consume less than $T_f/T_w$% CPU time on a ***single core***

# Implementation Options

| | CPU-only | GPU add-in | FPGA add-in |
|---|---|---|---|
| ***Reference*** | *DRPC* | *nVidia GeForce 2xx, 4xx, 5xx* | *Virtex 6/7 (*: XC6VLX240T)* |
| **Additional HW Costs** | 0 | $200/DR | ***$1800/DR*** |
| **Additional Power** | 5% | ***20-120%*** | 5% |
| **GFLOPS** | 2 to 20 (dep. on # cores used) | 60 to1800 (dep. on GPU model) | 200* to1145 (dep. on FPGA model) |
| **Watts/GFLOP** | ~7 | ~1.5 to ~0.2 | ~0.7* |
| **RT-feasible Complexity** | ≤O(N) | $O(N\log_2 N)$; N≤256k, $O(N^2)$; N≤64 | $O(N\log_2 N)$; N≤256k, $O(N^2)$; N≤64 |
| ***Largest PFB No Integration*** | ***8 ch, 2 taps*** | ***256k ch, 128 taps*** | ***512k ch, 32 taps*** |
| ***Largest PFB 40:1 Integration*** | ***1k ch, 2 taps*** | | |

# FPGA Proof of Concept

## (S. Vigraham's Master's Thesis)




- Demonstrated sustained throughput of 160MiB/s
  - Real world: would be rate-limited to about 120-130 MiB/s due to DRSU
- Xilinx XUPV5 (Virtex 5, LX110T; **Single lane PCI-E bus**)
- Key result: any on-the-fly processing is limited only by the size of the FPGA
- Useful as a "Second Chance" DP stage

# PFB size estimates: CPU Only

| M Channels \ T Taps/Channel | 1 | 2 | 4 | 8 | 16 | 32 | 64 |
|---|---|---|---|---|---|---|---|
| 2 | | | | | | | |
| 4 | | | | | | | |
| 8 | | | | | | | |
| 16 | | | | | | | |
| 32 | | | | | | | |
| 64 | | | | | | | |
| 128 | | | | | | | |
| 256 | | | | | | | |
| 512 | | | | | | | |
| 1024 | | | | | | | |
| 2048 | | | | | | | |
| 4096 | | | | | | | |
| 8192 | | | | | | | |
| 16384 | | | | | | | |
| 32768 | | | | | | | |
| 65536 | | | | | | | |
| 131072 | | | | | | | |

Legend

| CPU Availability Class |
|---|
| Real-time TBN |
| Real-time, DRX |
| Offline, Single-threaded, 1:1 |
| Offline, Multi-threaded, 1:1 |
| Worse than 1:1 |

Red line indicates approximate upper limit of Real Time PFB implementation

# PFB size estimates: FPGA add-in

| M Channels \ T Taps/Channel | 2 | 4 | 8 | 16 | 32 | 64 | 128 | 256 | 512 | 1024 | 2048 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | | | | | | | | | | 1031 | |
| 4 | | | | | | | | | 1031 | | |
| 8 | | | | | | | | 773 | | | |
| 16 | | | | | | | | 1031 | | | |
| 32 | | | | | | | 644 | | | | |
| 64 | | | | | | | 773 | | | | |
| 128 | | | | | | | 902 | | | | |
| 256 | | | | | | | 1031 | | | | |
| 512 | | | | | | | | | | | |
| 1024 | | | | | | | | | | | |
| 2048 | | | | | | | | | | | |
| 4096 | | | | | | | | | | | |
| 8192 | | | | | | | | | | | |
| 16384 | | | | | | | | | | | |
| 32768 | | | | | | | | | | | |
| 65536 | | | | | | | | | | | |
| 131072 | | | | | | | | | | | |

Legend

| FPGA Class Required | Approx. Price |
|---|---|
| < High end Virtex 5 | $1,800 |
| Low end Virtex 6 | $2,793 |
| High end Virtex 6 | $9,055 |
| Low end Virtex 7 | not available |
| High end Virtex 7 | not available |

# PFB size estimates: GPU add-in

| M Channels \ T Taps/Channel | 2 | 4 | 8 | 16 | 32 | 64 | 128 | 256 | 512 | 1024 | 2048 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 2 | 3 | 5 | 9 | 17 | 33 | 65 | 129 | 257 | 513 | 1025 |
| 4 | 3 | 5 | 9 | 17 | 33 | 65 | 129 | 257 | 513 | 1025 | 2049 |
| 8 | 5 | 9 | 17 | 33 | 65 | 129 | 257 | 513 | 1025 | 2049 | 4097 |
| 16 | 9 | 17 | 33 | 65 | 129 | 257 | 513 | 1025 | 2049 | 4097 | 8193 |
| 32 | 17 | 33 | 65 | 129 | 257 | 513 | 1025 | 2049 | 4097 | 8193 | 16385 |
| 64 | 33 | 65 | 129 | 257 | 513 | 1025 | 2049 | 4097 | 8193 | 16385 | 32769 |
| 128 | 65 | 129 | 257 | 513 | 1025 | 2049 | 4097 | 8193 | 16385 | 32769 | 65537 |
| 256 | 129 | 257 | 513 | 1025 | 2049 | 4097 | 8193 | 16385 | 32769 | 65537 | 131073 |
| 512 | 257 | 513 | 1025 | 2049 | 4097 | 8193 | 16385 | 32769 | 65537 | 131073 | 262145 |
| 1024 | 513 | 1025 | 2049 | 4097 | 8193 | 16385 | 32769 | 65537 | 131073 | 262145 | 524289 |
| 2048 | 1025 | 2049 | 4097 | 8193 | 16385 | 32769 | 65537 | 131073 | 262145 | 524289 | 1048577 |
| 4096 | 2049 | 4097 | 8193 | 16385 | 32769 | 65537 | 131073 | 262145 | 524289 | 1048577 | 2097153 |
| 8192 | 4097 | 8193 | 16385 | 32769 | 65537 | 131073 | 262145 | 524289 | 1048577 | 2097153 | 4194305 |
| 16384 | 8193 | 16385 | 32769 | 65537 | 131073 | 262145 | 524289 | 1048577 | 2097153 | 4194305 | 8388609 |
| 32768 | 16385 | 32769 | 65537 | 131073 | 262145 | 524289 | 1048577 | 2097153 | 4194305 | 8388609 | 16777217 |
| 65536 | 32769 | 65537 | 131073 | 262145 | 524289 | 1048577 | 2097153 | 4194305 | 8388609 | 16777217 | 33554433 |
| 131072 | 65537 | 131073 | 262145 | 524289 | 1048577 | 2097153 | 4194305 | 8388609 | 16777217 | 33554433 | 67108865 |

Legend

| GPU Class Required | Approx. Price |
|---|---|
| GTX 9 series or better | $50 |
| Any nVidia 200 GPU | $150 |
| GT 230 or better | $300 |
| GTX 280 or better | $550 |
| Not possible with current offerings | N/A |

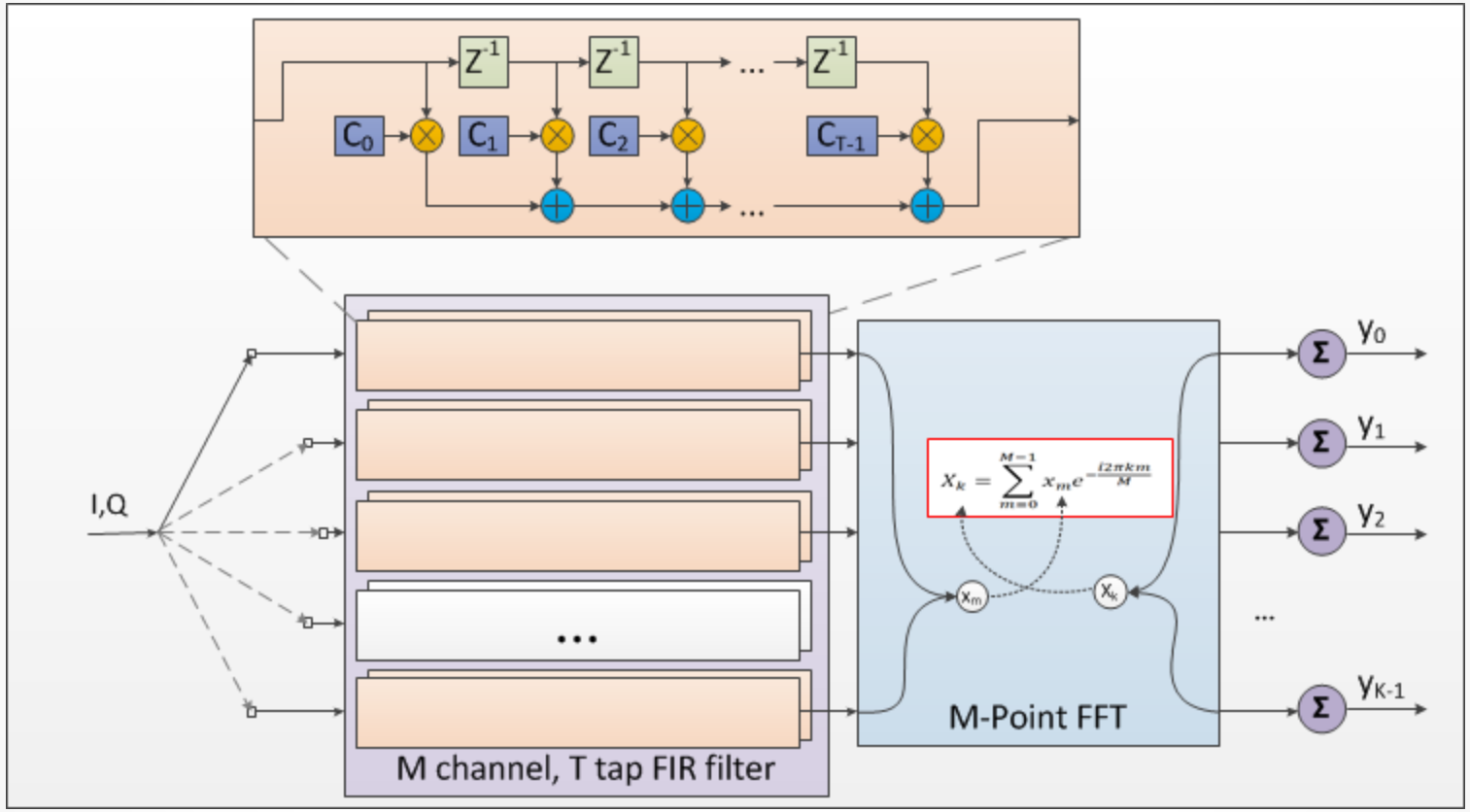
# Concluding Remarks

- Current LWA observation data flow
  - limits system availability
  - poses a tedious data management issue for both UNM and operator
- SW only option
  - $0 additional cost
  - ~5% additional power usage
  - Significantly limited processing capability (limited by DR)
- FPGA add-in option
  - $2000 or more additional cost
  - ~5% additional power usage
  - Extensive processing capability (limited by FPGA, not DR)
- GPU option
  - Less than $500 in additional HW cost
  - 20-120% additional power usage
  - Extensive processing capability (limited by GPU, not DR)

# References


1. http://www.ece.vt.edu/swe/mypubs/110105_URSI_LWA_Ellingson.pdf
2. http://www.ece.vt.edu/swe/lwa/memo/lwa0177a.pdf
3. http://www.xilinx.com/support/documentation/data_sheets/ds202.pdf
4. http://www.xilinx.com/support/documentation/data_sheets/ds100.pdf
5. http://press.xilinx.com/phoenix.zhtml?c=212763&p=irol-newsArticle&ID=1107663&highlight
6. http://h10025.www1.hp.com/ewfrf/wc/document?docname=c01901210&lc=en&dlc=en&cc=us&lang=en&product=4242509&key=null&site=null#N342
7. http://support.dell.com/support/edocs/systems/435mt/en/SG/P368KA00MR.pdf
8. http://www.dell.com/us/business/p/precision-t1500/pd
9. http://www.xilinx.com/images/boards/ml505/ml505_front.jpg
10. http://www.xilinx.com/support/documentation/boards_and_kits/ug534.pdf
11. http://www.xilinx.com/support/documentation/white_papers/wp387_TeraFLOPs_Performance_and_HLS.pdf

# PFB Reference Implementation



# PFB Reference Implementation

Fundamentally $O(M(\alpha T+\beta \log_2 M))$

T=number of filter taps

M=number of channels

Specifically: (per polarization, beam, tuning)

| | PFB-FIR | PFB-FFT | PFB-INT | Totals |
|---|---|---|---|---|
| Complex Multiply | 2M(T) | $\frac{M\log_2 M}{2}$ | 0 | $\frac{M(4T+\log_2 M)}{2}$ |
| Complex Add | 2M(T-1) | $M\log_2 M$ | M | $\frac{M(4(T-1)+2\log_2 M)}{2}$ |
| Total | | | | $\frac{M(8T+3\log_2 M-4)}{2}$ |

Operational period = M/SampleRate

GFLOPS $\propto$ $(T+\log_2 M)$/SampleRate

Data Rate $\propto$ $(\text{Integration Time})^{-1}$

For M=powers of 4, may reduce to $O(M(\alpha T+\beta \log_4 M))$

# RFI flagging Reference Implementation

Same order of complexity as FFT or PFB less small linear component

Additional Costs

- Basic thresholding
  - 4M complex operations
- $N^{th}$ moment
  - about (6+N)M complex operations

# Data Recorder Storage Unit (DRSU)

Five disk RAID-0 array (capacity, performance)

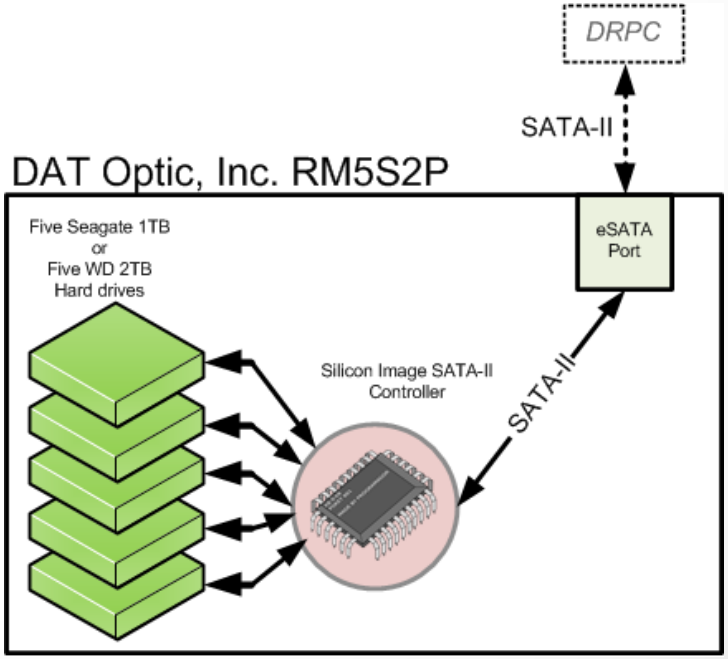
5 TB (DRSU) or 10 TB (DRSUx2) storage capacity per unit

- ~10 h continuous recording @ 112 MiB/s per DRSU
- ~20 h continuous recording @ 112 MiB/s per DRSUx2

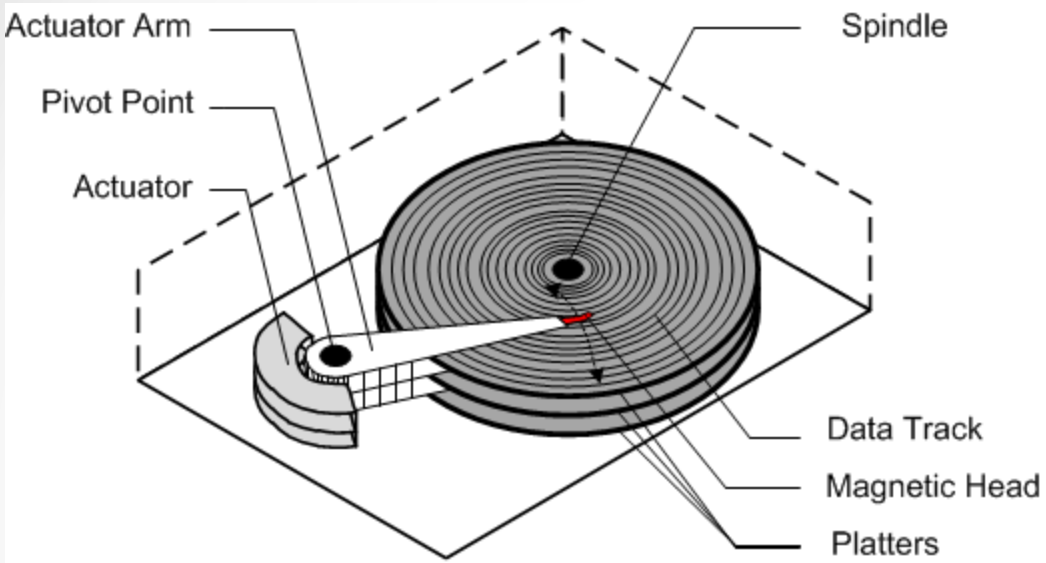
DRSU switching time: ~30-60 s

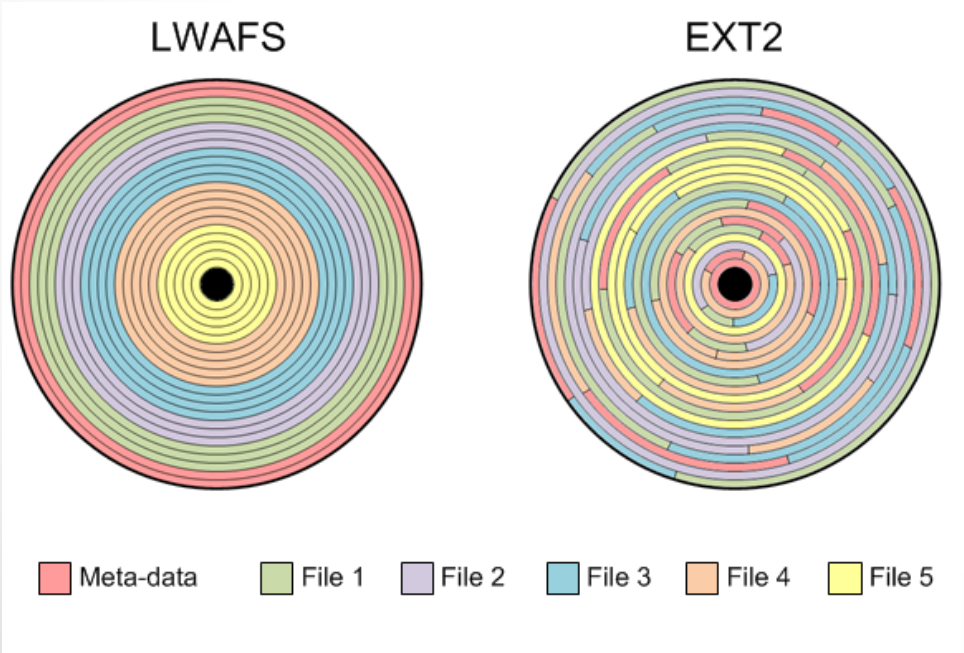
Custom file system

- Overcomes performance limitations of ext2, etc.
- Flat hierarchy
- 1023 files max (customizable)



# LWAFS

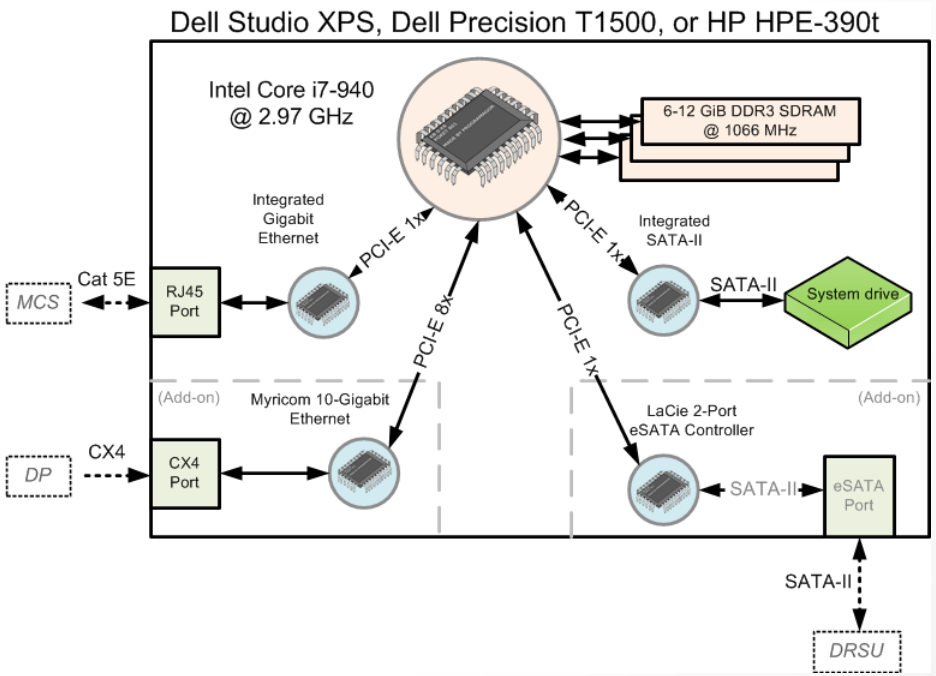

Actuator Arm
Pivot Point
Actuator
Spindle
Data Track
Magnetic Head
Platters


LWAFS
EXT2
Meta-data
File 1
File 2
File 3
File 4
File 5

# Data Recorder PC (DRPC)

C.O.T.S. (PC + NIC + eSATA + Storage)

- Intel Core i7 CPU
- 6-12 GiB DDR3 SDRAM
- Myricom 10 GbE base CX4
- LaCie PCI-E 1x 2-port eSATA



# DRPC Expansion Options

**Prototype + DR1** (Dell Studio XPS 435mt)

One PCI-E 1x slot, Three internal SATA-II, Eight USB

**DR2-DR5** (Dell Precision T1500) *pictured below

One 32-bit 33MHz PCI slot, Two internal SATA-II, Ten USB

**Future DRs** (HP HPE-590t)

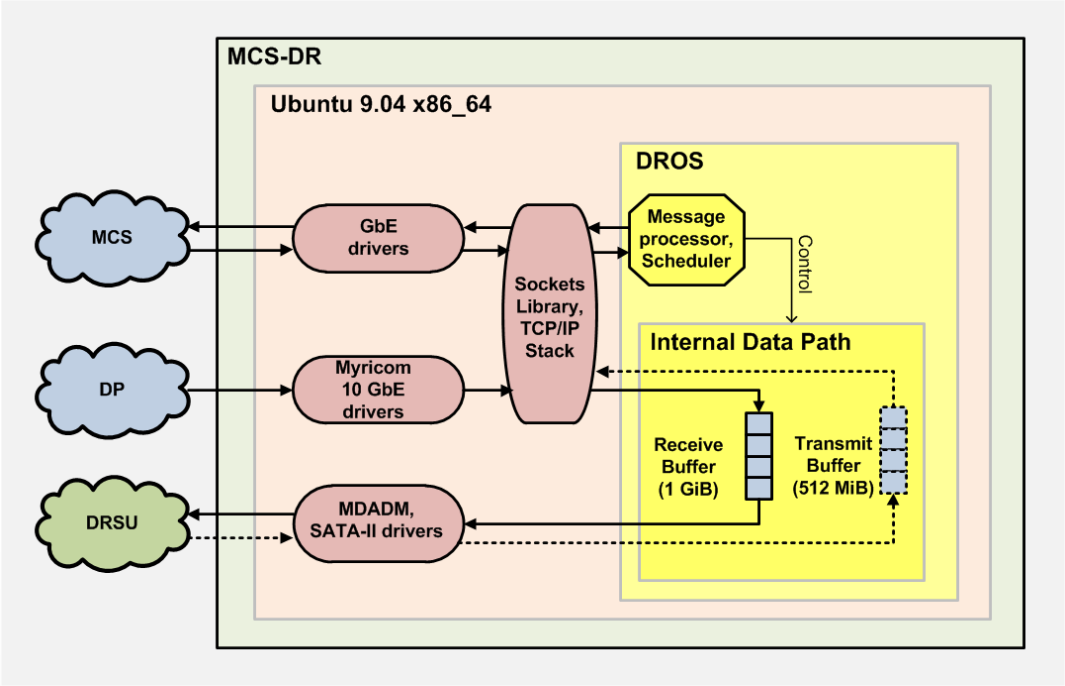
One PCI-E 16x slot, Two internal SATA-II, Five USB

# Data Recorder Operating Software (DROS)

Single, monolithic executable

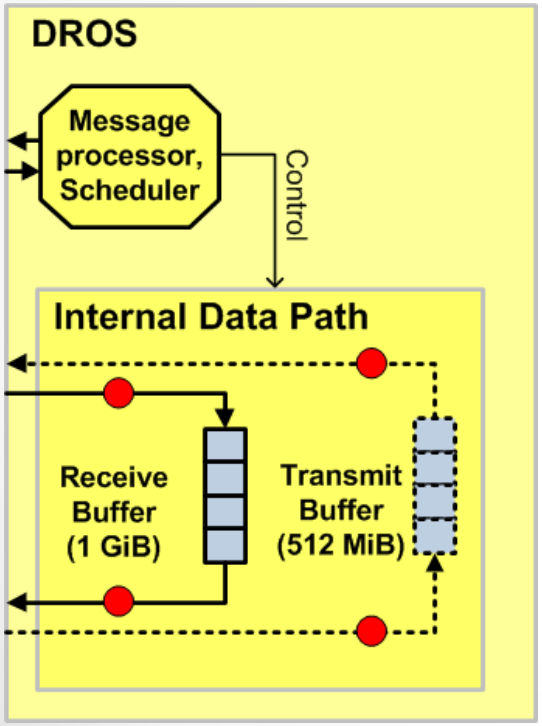
only multi-threading is in POSIX AIO lib

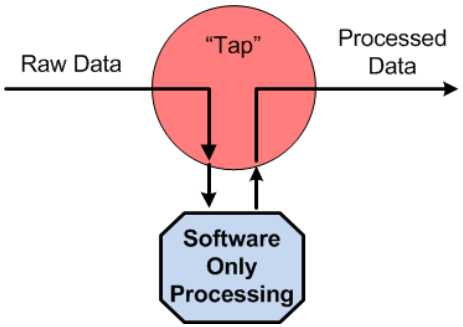
Monitor and Control System as per MCS Common ICD

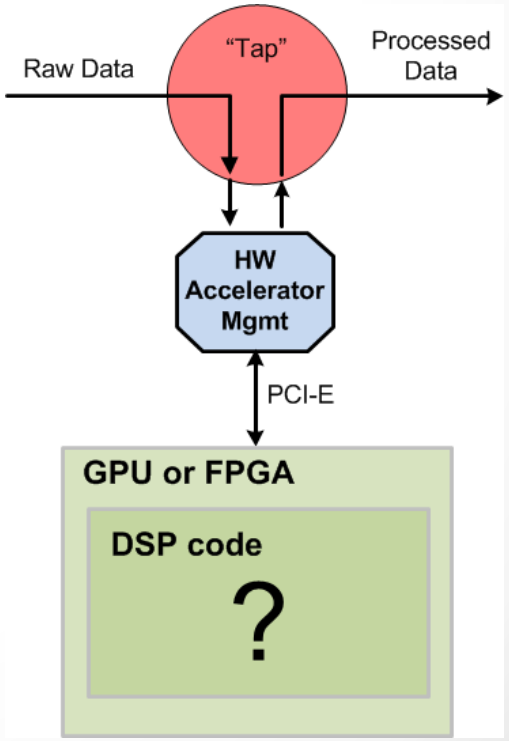


# DROS Extension Points

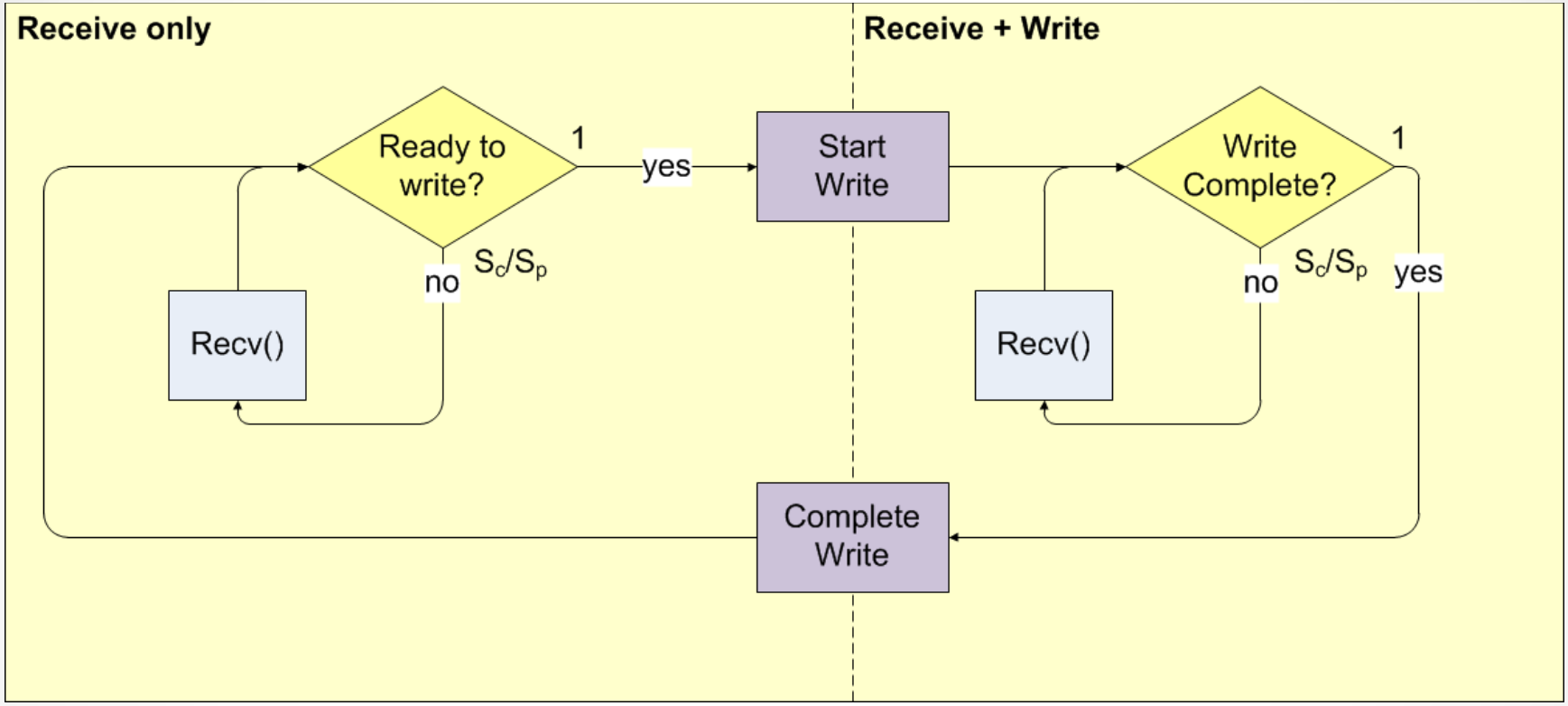
accessible “taps” in internal data path



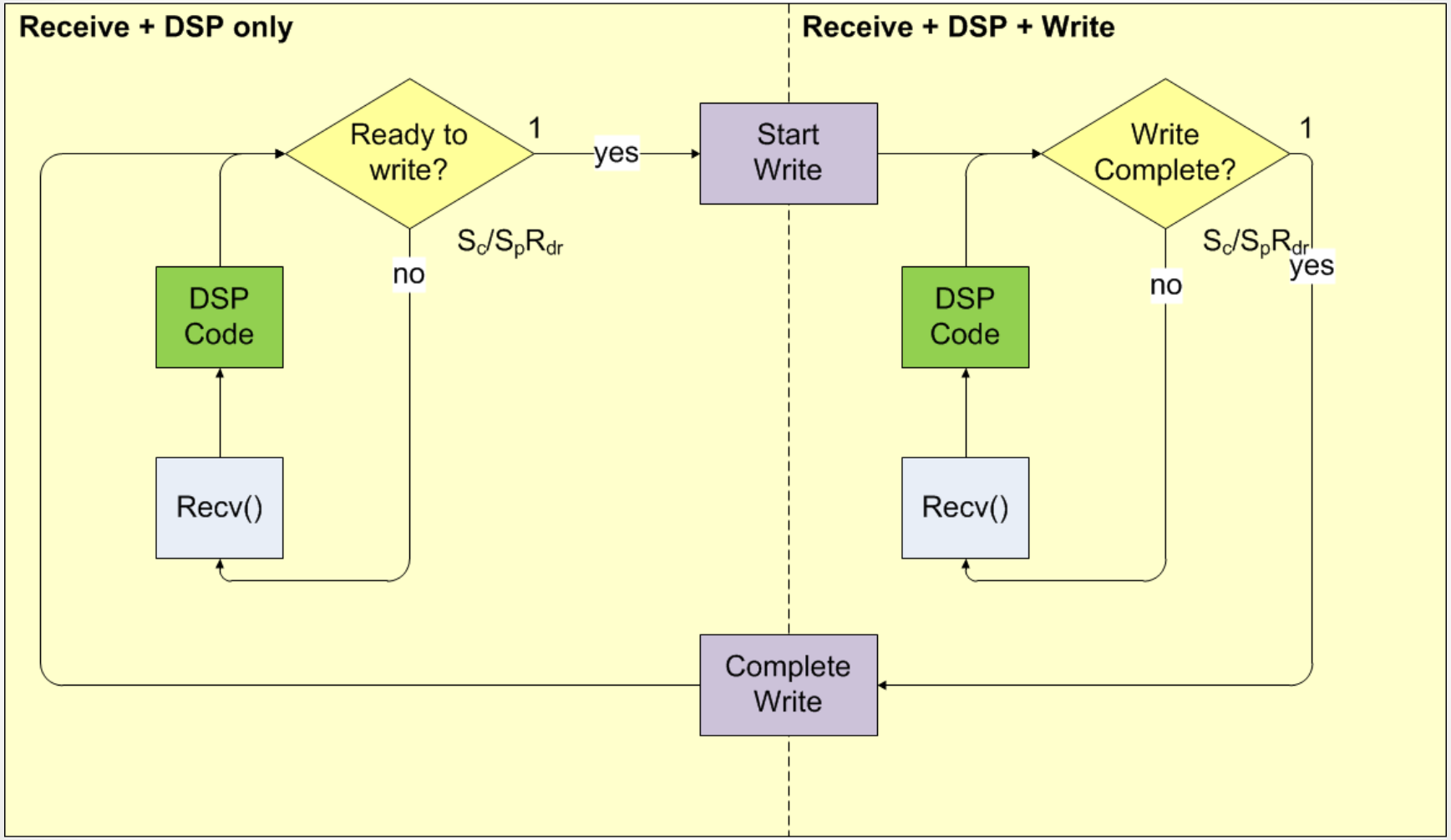



# Default Pipeline



# DSP Pipeline



# Real-time Processing Feasibility

- IOPS
  - DRX ~ 21k IOPS
  - TBW ~ 84k IOPS
  - TBN ~ 112k IOPS