

ADQ14-FWATD

User Guide

Author(s): Teledyne SP Devices
Document ID: 16-1849
Classification: Public
Revision: PA9
Print date: 2019-10-14

Contents

| | | |
|----------|-----------------------------------------|-----------|
| 1 | Introduction | 2 |
| 1.1 | Definitions and Abbreviations | 2 |
| 2 | System Description | 2 |
| 2.1 | Accumulation Engine | 2 |
| 2.1.1 | Partitioning Algorithm | 3 |
| 2.1.2 | Record Segmentation | 4 |
| 2.1.3 | Repeated Accumulations | 5 |
| 2.2 | Physical Interface | 5 |
| 2.3 | Data Struct | 6 |
| 2.3.1 | Enabling the Struct Format | 7 |
| 2.4 | Collection Modes | 7 |
| 2.4.1 | Single Measurement | 7 |
| 2.4.2 | Streaming | 7 |
| 2.5 | Overflow Behavior | 8 |
| 2.6 | Low Latency Streaming | 10 |
| 2.6.1 | System Configuration | 10 |
| 3 | Discussion | 11 |
| 3.1 | Partitioning Algorithm | 11 |
| 3.2 | Record Length and Data Format | 12 |
| 4 | Future Editions | 16 |

1 Introduction

This document presents the user guide for the advanced time-domain (ATD) firmware option on ADQ14. The outline of this document is presented below.

- In Section 2, the features of ADQ14-FWATD are discussed. High-level descriptions of the different parts of the waveform averaging engine is presented to gain an understanding of the system and develop the basis needed to discuss the limitations.
- In Section 3, the limitations of FWATD are discussed and motivation as to why specific parameters boundaries exist is provided.
- In Section 4, coming additions to this document is listed.

1.1 Definitions and Abbreviations

Table 1 lists the definitions and abbreviations used in this document and provides an explanation for each entry.

Table 1: Definitions and abbreviations used in this document.

| Term | Explanation |
|--------|-------------------------------------------|
| AFE | Analog front-end |
| ATD | Advanced time-domain |
| API | Application programming interface |
| Batch | Accumulated record output by the device |
| FW | Firmware (device FPGA configuration) |
| PCIe | Peripheral Component Interconnect Express |
| USB3.0 | Universal serial bus (version 3.0) |
| WFA | Waveform averaging |

2 System Description

The advanced time-domain firmware for ADQ14 offers the ability to perform hardware-accelerated averaging of waveforms with safe scaling up to 65536 accumulations. The text in this section will refer to functions in the ADQAPI. They are typeset using a typewriter style, e.g. `ATDStartWFA()` and up-to-date information on each function can be found in the ADQAPI reference guide [1].

2.1 Accumulation Engine

This section aims to describe the accumulation engine from a user-perspective. Fig. 1 presents a high-level block diagram of the entire system—from input signal to user-space buffers. The diagram has been

significantly simplified so that the standard functions of the ADQ14 have been bundled together in the blocks labeled 'Acquisition Engine' and 'Transfer Engine'.

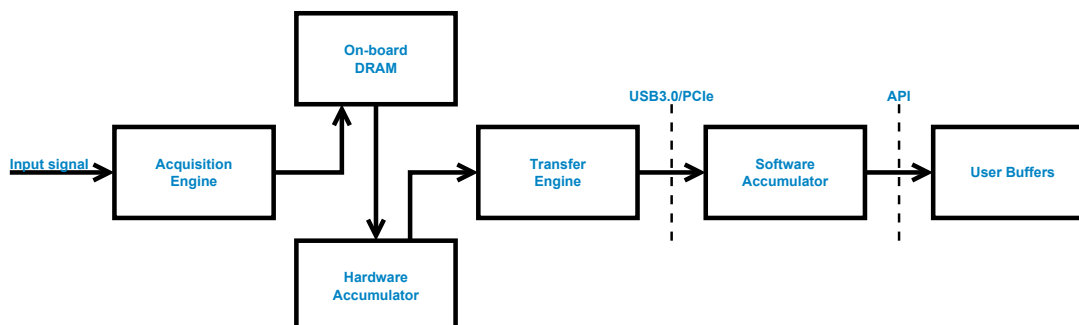


Figure 1: Block diagram of the FWATD averaging system.

The acquisition engine creates records from the input signal which are stored in the on-board DRAM (2 GB capacity). The averaging engine is divided into two parts: one in hardware, and one in software. Thus, the work load is divided between the device and the host computer. This split is determined by a partitioning algorithm applied to the parameter `nof_accumulations` in `ATDSetupWFA()`. The goal of the algorithm is to determine the most suitable factorization of the number of accumulations that the user wishes to perform (refer to Section 2.1.1 for a detailed description of the algorithm).

Define a *batch* as the resulting record output by the device after performing its assigned number of accumulations. The batch is transferred over the physical interface (ideally USB3.0 or PCIe) to the host computer where a collection thread is performing the last accumulation step by combining a sequence of batches into one record which is delivered to the user buffers.

Since the on-board DRAM has been re-tasked with realizing part of the WFA algorithm, the device-to-host link is much more sensitive to stalls by the host computer as the only buffers available are small (compared to the DRAM) due to residing in the FPGA block RAM.

Note

In ADQ14-FWATD, the DRAM no longer functions as a large FIFO able to buffer large amounts of data in the event that the host computer is busy.

2.1.1 Partitioning Algorithm

Suppose the number of accumulations, N_A , is factorized into

$$N_A = N_{ADQ} \cdot N_{PC} - R, \quad (1)$$

where N_{ADQ} and N_{PC} is the number of averaging steps performed by the device and host computer, respectively. The remainder R is required to allow N_A to assume all integer values in the supported range. It follows that the product $N_{ADQ} \cdot N_{PC}$ will always be equal to or larger than the user-specified value. Following the earlier definition, a batch may be identified as a record averaged N_{ADQ} times and N_{PC} as the number of batches. The device handles the remainder R by only performing $N_{ADQ} - R$ averages in the final batch.

Note

The factorization result of `nof_accumulations` is output in the trace log file.

The goal of the partitioning algorithm is to select an optimal integer value for N_{ADQ} constrained by the boundaries N_L and N_U such that

$$N_{ADQ} \in [N_L, N_U]. \quad (2)$$

These boundaries can be set and read by the user by using the two functions `ATDSetWFAPartitionBoundaries()` and `ATDGetWFAPartitionBoundaries()`, respectively. However, performing changes to these boundaries may cause the system to no longer fulfill the standard specification [2] with regards to the maximum record length and/or duty cycle. The default boundary values can be restored by calling `ATDSetWFAPartitionBoundariesDefault()`. Please refer to Section 3.1 for an in-depth discussion on the effects of changing these parameters.

Note

Altering the factorization boundaries will cause the WFA performance to change with regards to maximum supported record length and/or duty cycle. Use with caution.

Additionally, this partitioned approach does not result in a data rate reduction by a factor of N_A over the physical interface, but rather the factor $N_{ADQ}/2$, as determined by the partitioning algorithm.

Note

The raw data rate is not reduced by a factor equal to the number of accumulations over the device-to-host interface.

2.1.2 Record Segmentation

To manage accumulation of large records, the record is split into segments of more manageable size. Due to reasons explained in Section 3.2, the record is required to be a multiple of a certain number of samples. Define this number as the *length factor*.

Table 2: Record length factor for all ADQ14 variants.

| Device | Length factor |
|-----------------------|---------------|
| ADQ14-2A and ADQ14-4A | 36 |
| ADQ14-2C and ADQ14-4C | 72 |
| ADQ14-2X and ADQ14-1X | 144 |

How the record is split into these segments is also determined by a partitioning algorithm, this time applied to the `record_length` parameter in `ATDSetupWFA()`. Due to reasons which fall outside the scope of this document, certain values of the record length will work better than others in terms of allowing a higher duty cycle. In these cases, the partitioning algorithm is able to find an efficient segmentation of the record, allowing the hardware accumulator to work with minimal overhead. These optimal values are in no way intuitive which is why the function `ATDGetAdjustedRecordLength()` exists. This function uses

the input record length and a search direction to return an adjusted record length which is able to fulfill the duty cycle specification [2]. If this function is used, it is important to allocate the user-space buffers using this adjusted value. Otherwise, segmentation faults may occur.

Note

The function `ATDGetAdjustedRecordLength()` should be used to find a record length which is optimal from a hardware perspective. Any allocation of user-space target buffers should use the adjusted value to avoid segmentation faults.

2.1.3 Repeated Accumulations

A common wish is to be able to perform multiple waveform accumulations back-to-back, i.e., the first N_A triggers resulting in one record in the user-space, the next N_A triggers generating a second record, and so on. The parameter `nof_repeats`, N_R , in `ATDSetupWFA()` exists for this purpose. The accumulation engine uses this parameter to determine the number of user-space records to produce. The accumulation process will continue until the completion criteria is met, i.e. the hardware has captured $N_R \cdot N_A$ triggers and they have all been transferred to the host computer, or the user aborts the process by calling `ATDStopWFA()`¹.

Additionally, if there is a wish to generate a constant stream of averaged records, the number of repeats should be set to $2^{32} - 1$ or `UINT_MAX`. This mode requires the user to implement the *streaming* data collection mode detailed in Section 2.4.2 and will require calls to `ATDStopWFA()` followed by `ATDWaitForWFACompletion()` to correctly end the collection loop.

Note

The number of repeats determines how many user-space records are generated. A value of $2^{32} - 1$ enables infinite streaming, requiring the user to implement the *streaming* data collection mode (see Section 2.4.2).

2.2 Physical Interface

Transferring data from the device to the host computer is done over a physical interface. ADQ14 supports USB3.0 (USB2.0² fallback), PXIe, PCIe and MTCA, although not every variant of ADQ14 supports every interface. For example, an ADQ14-PXIe device cannot be used as a PCIe device and an ADQ14-USB cannot be used as a PXIe device with regard to how it is connected to the host PC.

Note

PXIe and MTCA devices use the PCIe protocol to transfer data.

These interfaces have different specifications on the maximum transfer rate with USB3.0 being the slowest at a theoretical limit³ of roughly 500 MB/s (effective data rate). However, achieving these transfer rates is dependent on how the user sets up the API transfer buffers. In the USB case, failure to allocate

¹A call to `ATDStopWFA()` must always be followed by a call to `ATDWaitForWFACompletion()` to properly close the thread spawned by the API.

²USB2.0 is limited to a transfer rate of roughly 57 MB/s (480 Mbps), ideally. Teledyne SP Devices strongly advises against using the digitizer configured for FWATD with this interface. Refer to the discussion in Section 3.1 for more details.

enough memory will cause significant overhead, causing the effective transfer rate to drop. PCIe is more tolerant of smaller transfer buffer sizes.

The transfer buffers should be set up in the user code by calling `SetTransferBuffers()`. This function allows the user to specify the number of transfer buffers to allocate in parallel, and their size. Once a buffer is filled, the API will parse the contents, rendering records which are later presented to the user. While this process is ongoing, the buffer is marked as full and no new transactions on the physical interface can target this buffer. In order not to block the flow of data, there is a need for several buffers in parallel. Naturally, the minimum requirement is two buffers but allocating a several is recommended.

Additionally, large buffers result in a device-to-host interface able to reach higher transfer speeds. The allocated size is critical to achieving high transfer rates over USB, less so for PCIe. Transfer buffers in the range 1 MB – 4 MB are recommended. The API will attempt to allocate memory linearly for each transfer buffer. Should this operation fail, the call to `SetTransferBuffers()` will return 0 and a message will be visible in the trace log file.

Note

Allocating several transfer buffers in parallel is needed in order not to block the flow of data over the physical interface. The recommended size is in the range 1 MB – 4 MB. Memory is allocated linearly.

2.3 Data Struct

Starting from release R191014 the advanced time-domain firmware and API supports acquiring data using a struct format instead of the raw data format. This enables transferring useful metadata to the user space, in addition to the accumulated waveform data. The struct `ATDWFABufferStruct` is defined in the header file `ADQAPI.h` and is shared between ADQ14 and ADQ7. The struct consists of the fields presented in Table 3 where the first table entry starts at byte zero.

Table 3: The fields in the `ATDWFABufferStruct`. Note that not all fields are applicable to ADQ14.

| Field name | Type | Description |
|--------------------|-----------|--------------------------------------------------------------------------------------------------|
| Timestamp | uint64_t | The timestamp of the first raw record in the user space record. |
| Data | int32_t * | Pointer to memory where the record data is located. |
| RecordNumber | uint32_t | The record number—the value starts at zero and is incremented by one for each user space record. |
| Status | uint32_t | N/A (will always read as zero) |
| RecordsAccumulated | uint32_t | The number of accumulated records reflected in the data. |
| Channel | uint8_t | The originating channel. |

³This speed is not guaranteed by ADQ14-USB devices.

2.3.1 Enabling the Struct Format

To maintain backwards compatibility, the API assumes that the acquisition should be carried out using the raw data format by default. The user has to specifically request that the struct format should be used by calling `ATDSetWFABufferFormat()`. For example, using the C API:

```
ADQ_ATDSetWFABufferFormat(adq_cu, adq_num, ATD_WFA_BUFFER_FORMAT_STRUCT);
```

This function will return an error if the current firmware or API does not support the operation.

2.4 Collection Modes

This section presents an overview of the two data collection modes available to the user. The streaming mode (Section 2.4.2) is more general than the single measurement mode (Section 2.4.1) and is the recommended mode of operation.

Note

Starting from release R191014 the advanced time-domain firmware and API supports acquisition using a struct format of the WFA buffers. Refer to Section 2.3 for details. The acquisition process implemented in the example in C employs this mode.

2.4.1 Single Measurement

The single measurement mode is well suited to observing short events and is simple with respect to the required user code. The mode is activated by specifying an array of valid target buffer pointers to `ATDStartWFA()`. The available memory pointed to must be at least

$$4N_R L_R \text{ bytes,}$$

where N_R is the number of repeats and L_R is the record length in samples. The channel mask will determine which channels are active and thus which array entries are valid pointers.

Since it is required to allocate all the memory needed for the measurement at once, this mode is not well suited for long records combined with a high number of repeats, instead the streaming mode (Section 2.4.2) should be used.

Note

The FWATD example in C demonstrates how the streaming mode is implemented. Follow the program flow with the define `SINGLE_SHOT_COLLECTION` *active*.

2.4.2 Streaming

The streaming mode is based around a buffer queue system controlled by the API functions `ATDRegisterWFABuffer()` and `ATDWaitForWFABuffer()`. The mode is activated by specifying `NULL` as the value of the target buffer array in `ATDStartWFA()`.

Each buffer represents a record and before calling `ATDStartWFA()` a sufficiently large number of buffers should be registered for each channel intended to be used. Completed records are delivered to the user via the function `ATDWaitForWFABuffer()` by reporting a location in the memory (same as the

previously registered area) where the latest completed record can be read. The user is responsible for memory allocated for the buffer queue and the queue is reset in the call to `ATDStopWFA()`.

Note

The user is responsible for the allocated memory at all times. The clearing of the queue in `ATDStopWFA()` does not free the memory.

The minimum number of buffers needed to be registered will depend on multiple factors such as the transfer buffer size, the data throughput and the number of repeats, i.e. varying with the use-case. The goal is to keep the system balanced so that there always is a buffer available to place the accumulated record. In short, the queue must never be empty while the measurement is ongoing. Since large transfer buffers are recommended (see Section 2.2), the parsing of this buffer may generate several user-space records at once, thus requiring the queue of empty target buffers to be filled to a certain threshold. Generally, a large number of target buffers (≥ 50) is recommended. However, if the memory can be afforded, an even larger number may have additional benefits.

Note

The queue system has to be balanced such that empty target buffers are available whenever the need arises.

Once the user has processed the accumulated record, e.g. by writing the data to disk, the target buffer should be recycled as an empty buffer by calling `ATDRegisterWFABuffer()`.

Note

The FWATD example in C demonstrates how the streaming mode is implemented. Follow the program flow with the define `SINGLE_SHOT_COLLECTION` *inactive*.

Data Collection Flowchart

Fig. 2 presents a flowchart of the data collection loop.

2.5 Overflow Behavior

In this section, the term 'overflow' refers to data being discarded due to a bandwidth mismatch in the physical interface or due to a violation of the maximum supported duty cycle.

ADQ14-FWATD does not keep track of where data has been discarded in the stream of records, only that this has occurred. The user is notified of this event through the WFA status, read by calling `ATDGetWFAStatus()`. The status is *sticky*, meaning it reflects all the status codes that have occurred during the measurement. Configuring the WFA for a new measurement (`ATDSetupWFA()`) clears the status value.

In the event of an overflow, it is imperative that the WFA engine is shut down in a well-defined manner to ensure correct behavior once it is restarted. Two operations are required from a user perspective: detecting the overflow condition and closing the internal streaming thread.

For a minimal flow control, no polling mechanism is required. Instead, the status is indirectly provided through the return value of `ATDWaitForWFABuffer()` together with the address of the returned buffer. If an overflow is detected by the internal streaming thread, the buffer queue is placed into a state in which

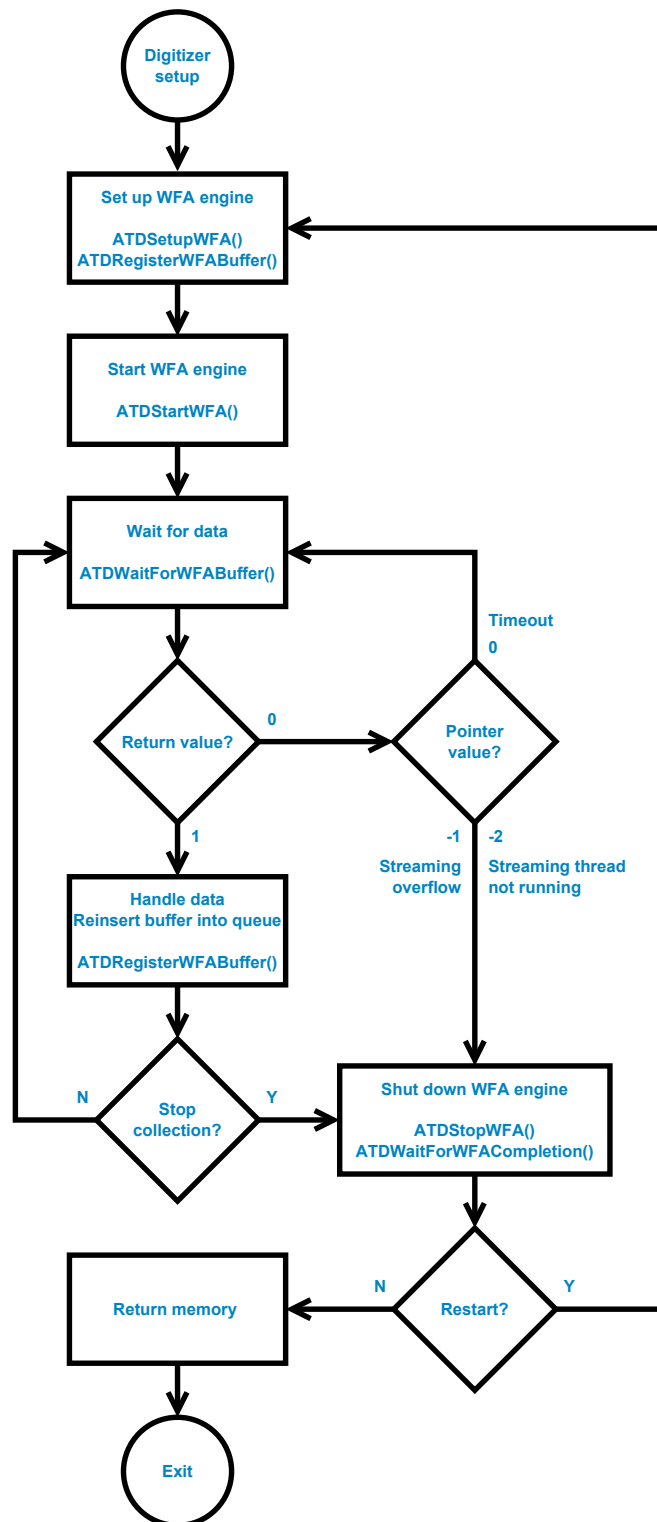


Figure 2: Data collection flowchart for the streaming mode.

subsequent calls to `ATDWaitForWFABuffer()` will return the value '0' with the pointer value '-1'.

2.6 Low Latency Streaming

Some applications may require that the data buffer arrives in user space as soon as possible after its corresponding raw data has been captured. By default, ADQ14-FWATD premier high throughput by recommending large transfer buffers and allowing the data to drive the output rate.

While setting up smaller transfer buffers would reduce the latency, the action on its own is not enough to solve the problem and still keep the system robust. Since the data drives the output (low overhead), there are no guarantees that the data representing one user space record constitutes an even number of transfer buffers. This mismatch creates suboptimal transfer performance when combined with small buffer sizes.

For each use case, the solution to the problem involves

- minimizing the required bandwidth of the device-to-host interface,
- activating an internal flush mechanism and
- computing and tuning the optimal transfer buffer size.

The internal flush mechanism intentionally adds overhead to ensure that every user space record consists of a fixed number of *packets*. Packets represent abstracted data and are what constitutes the data in a transfer buffer.

Note

Activating the internal flush mechanism requires ADQAPI r35550 or later.

The resulting system will be responsive and perform well for the attuned use case, but will be less tolerant of other stimuli patterns.

2.6.1 System Configuration

The following steps outline the process of configuring low latency data acquisition for a particular use case. A *use case* is characterized by its requirements on the record length, R_L , the number of accumulations N_A , the trigger period T_P , the maximum tolerable latency $T_{lat, max}$ and the device sample rate. The record length R_L is assumed to fulfill the requirements outlined in Sections 2.1.2 and 3.2.

1. Optimize the WFA partition boundaries to relax the requirements on the device-to-host interface. Increase the upper bound to (3) for -C and -X devices (1 GSPS and 2 GSPS base sample rate, respectively) and to (4) for -A devices (500 MSPS base sample rate).

$$N_U = \left\lceil \frac{1024^3}{\frac{8}{9} \cdot 8 \cdot R_L} \right\rceil \quad (3)$$

$$N_U = \left\lceil \frac{1024^3}{\frac{8}{9} \cdot 16 \cdot R_L} \right\rceil \quad (4)$$

by calling `ATDSetWFAPartitionBoundaries()`. Refer to Sections 2.1.1 and 3.1 for details on this equation.

2. Compute the required transfer buffer size to fit exactly one record. Equation (5) gives the size in bytes for a –C device and (6) gives the size in bytes for an –X device. Equation (5) is also valid for –A devices.

$$TB = 1024 \left(1 + \left\lceil \frac{R_L \cdot 4 - 1000}{1016} \right\rceil \right) \quad (5)$$

$$TB = 1024 \left(1 + \left\lceil \frac{R_L \cdot 4 - 992}{1008} \right\rceil \right) \quad (6)$$

3. Activate the internal flush mechanism by calling `FlushPacketOnRecordStop()` (requires ADQAPI r35550 or later).
4. (OPTIONAL) If the device-to-host bandwidth is insufficient, the transfer buffer size may be increased to a multiple of the unit size computed in step 2. Define the integer k as this factor. The worst-case latency defines an upper bound on its value as

$$0 < k < \frac{T_{lat, max}}{T_P \cdot N_A} \quad (7)$$

The user has some freedom in choosing the value of k , unless the upper bound is less than 1, in which case this method is not a valid option. Depending on the record size, a lower k may be desired to keep the transfer buffer size below e.g. 4 MB.

This will increase the bandwidth of the device-to-host interface but cause records to have non-uniform (but predictable) latency. For example, for a transfer buffer that fits precisely three user space record, the perceived latency of the first record will be $2T_P N_A$, the second $T_P N_A$ and the last one effectively zero.

3 Discussion

This section aims to provide a discussion on the different aspects of ADQ14-FWATD needed to be considered before and during integration of this product into any system. A discussion on the effects of the partitioning algorithm can be found in Section 3.1, followed by a discussion on the record length and data format in Section 3.2.

3.1 Partitioning Algorithm

Section 2.1.1 described the purpose of the partitioning algorithm used to determine the work load distribution between the hardware and software accumulator. The algorithm determines values for N_{ADQ} ,

N_{PC} and R (indirectly) from a given number of accumulations, N_A , according to (1). Assuming the maximum raw data volume from the ADQ14-FWATD, i.e. ≈ 14.9 GB/s, it is possible to generate the heat maps presented in Fig. 3 for a few values of the partition boundaries N_L and N_U . This figure presents the maximum achievable duty cycle as a function of the host interface throughput and the user-specified number of accumulations.

Another interesting metric is the DRAM utilization efficiency for the last batch due to the remainder R . Let the efficiency η be defined as the relative size between the last batch and the other batches in terms of the number of raw records (triggers) needed, i.e.

$$\eta = \frac{N_{ADQ} - R}{N_{ADQ}}. \quad (8)$$

Fig. 4 presents the corresponding utilization histograms, complementing the data in Fig. 3.

Analyzing the data in Fig. 3, it is evident that tasking the digitizer with performing additional accumulations before transferring the batch alleviates the requirements on the average transfer speed of the host interface. However, if the maximum supported record size is to be maintained, the upper partition boundary must remain. This results in a shrinking set of valid integer values for N_{ADQ} , causing the progressively deteriorating distributions shown in Fig. 4. For low utilization efficiencies, the hardware accumulator will be the bottleneck, potentially causing an overflow due to the large overhead in the memory operations. Therefore, there is a trade-off when only the lower bound is increased where the critical point is moved from the host interface to the hardware accumulator.

Increasing the upper boundary will prevent the set from shrinking and thus avoiding the effects shown in Fig. 4. However, doing so will sacrifice the maximum supported record size. If the use-case is well defined and a record size of 1 MS is not needed, increasing the upper partition boundary may have a beneficial effect on the performance. To compute the maximum value for the upper bound given a certain record length, refer to (9)

$$N_U = \left\lfloor \frac{1024^3}{\frac{8}{9} \cdot 2 \cdot 4 \cdot R_L} \right\rfloor, \quad (9)$$

where R_L is the record length.

The lower bound is determined by the minimum data rate reduction needed to support the physical interface with the lowest average transfer speed. For ADQ14, this is USB3.0 for which an average throughput⁴ of 310 MB/s yields

$$N_L = \left\lceil \frac{16 \cdot 10^9}{310 \cdot 1024^2} \right\rceil = 50. \quad (10)$$

Furthermore, it is clear from Fig. 3 that using ADQ14-FWATD with USB2.0 will not be able to reach high averaging duty cycles unless the device is tasked with performing at least 509 accumulations, assuming an average transfer rate of 30 MB/s. This entails that the upper bound must shift upwards as well, limiting the maximum record size to a few hundred kS.

3.2 Record Length and Data Format

Due to bandwidth requirements, the raw record samples are packed and stored as 14-bit values in the on-board DRAM. This is a difference to the other firmware options on ADQ14 where the samples are

⁴The speed of 310 MB/s is empirically determined from system tests.

16-bit aligned. This packing translates to a compression factor of $\frac{8}{9}$. It is due to this compression that the records have to be a multiple of the length factor, detailed in Table 2.

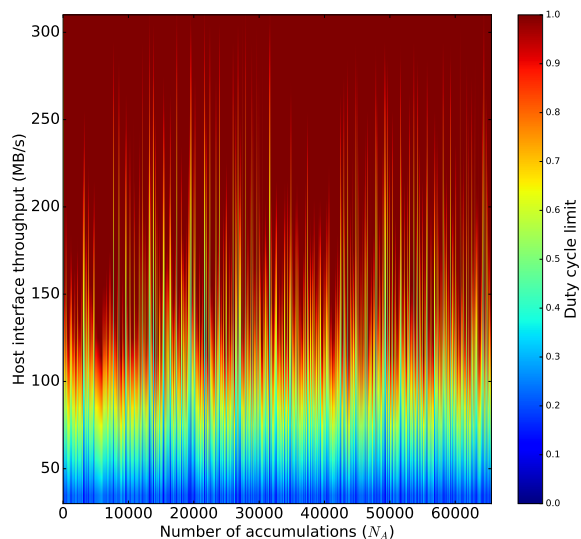
The accumulator uses 32-bit samples to guarantee safe scaling for up to 65536 accumulations. This means that the minimum ADC code, -2^{15} , can be accumulated 65536 times without an arithmetic overflow. The same is true for the maximum ADC code, $2^{15} - 1$. The data type is converted in the hardware accumulator which reads out the 14-bit samples from DRAM and accumulates each sample using a 32-bit accumulator. Therefore, ADQ14-FWATD uses 4 bytes per sample when the data is transferred to the host computer.

If the target signal does not utilize the full input range of the AFE, the upper bound for the number of accumulations where safe scaling is guaranteed increases.

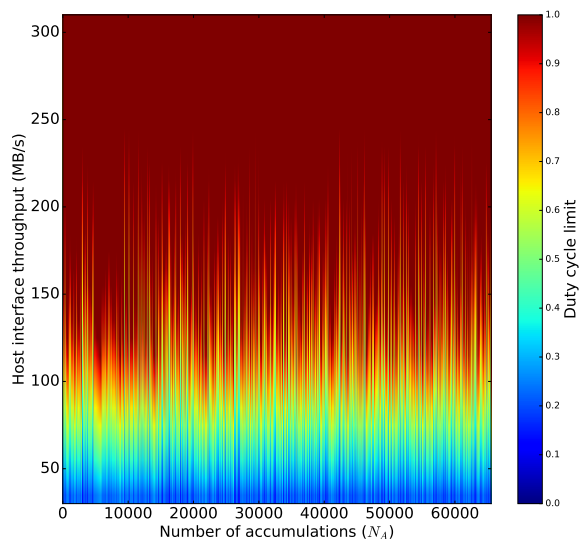
Note

An accumulation overflow is reported by the corresponding bit in the `wfa_status` parameter after calling `ATDGetWFASStatus()`.

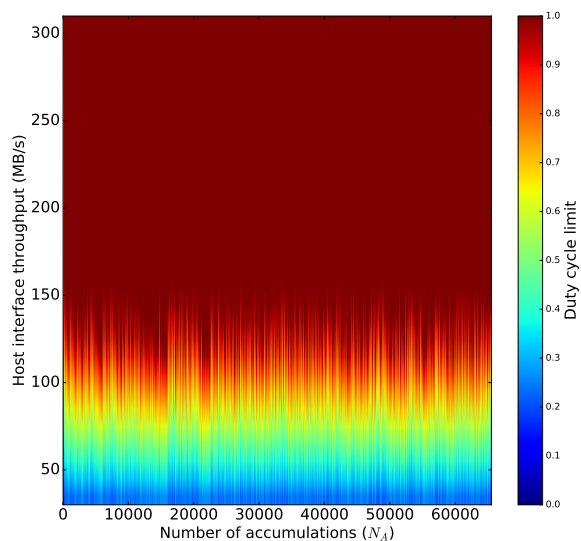
⁵A duty cycle of 100% is not possible in ADQ14-FWATD. Following the end of a record, a re-arm time of 20 ns is required before a new record can begin.



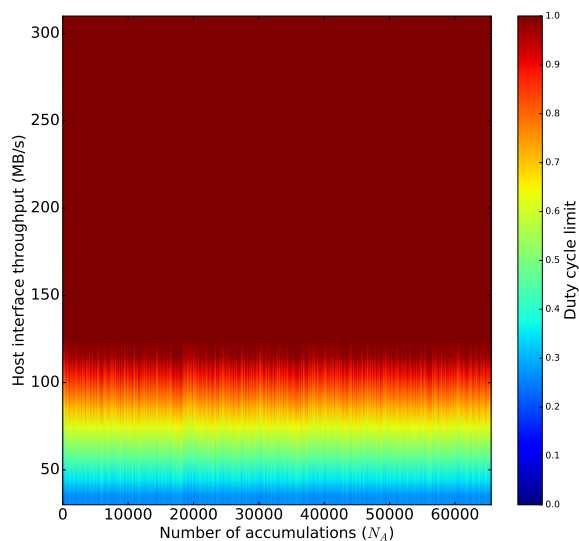
(a) $N_L = 50$, $N_U = 144$.



(b) $N_L = 64$, $N_U = 144$.



(c) $N_L = 100$, $N_U = 144$.



(d) $N_L = 128$, $N_U = 144$.

Figure 3: Illustration of maximum duty cycle⁵ as a function of the host interface throughput and the user-specified number of accumulations factorized using different values for the partition boundaries.

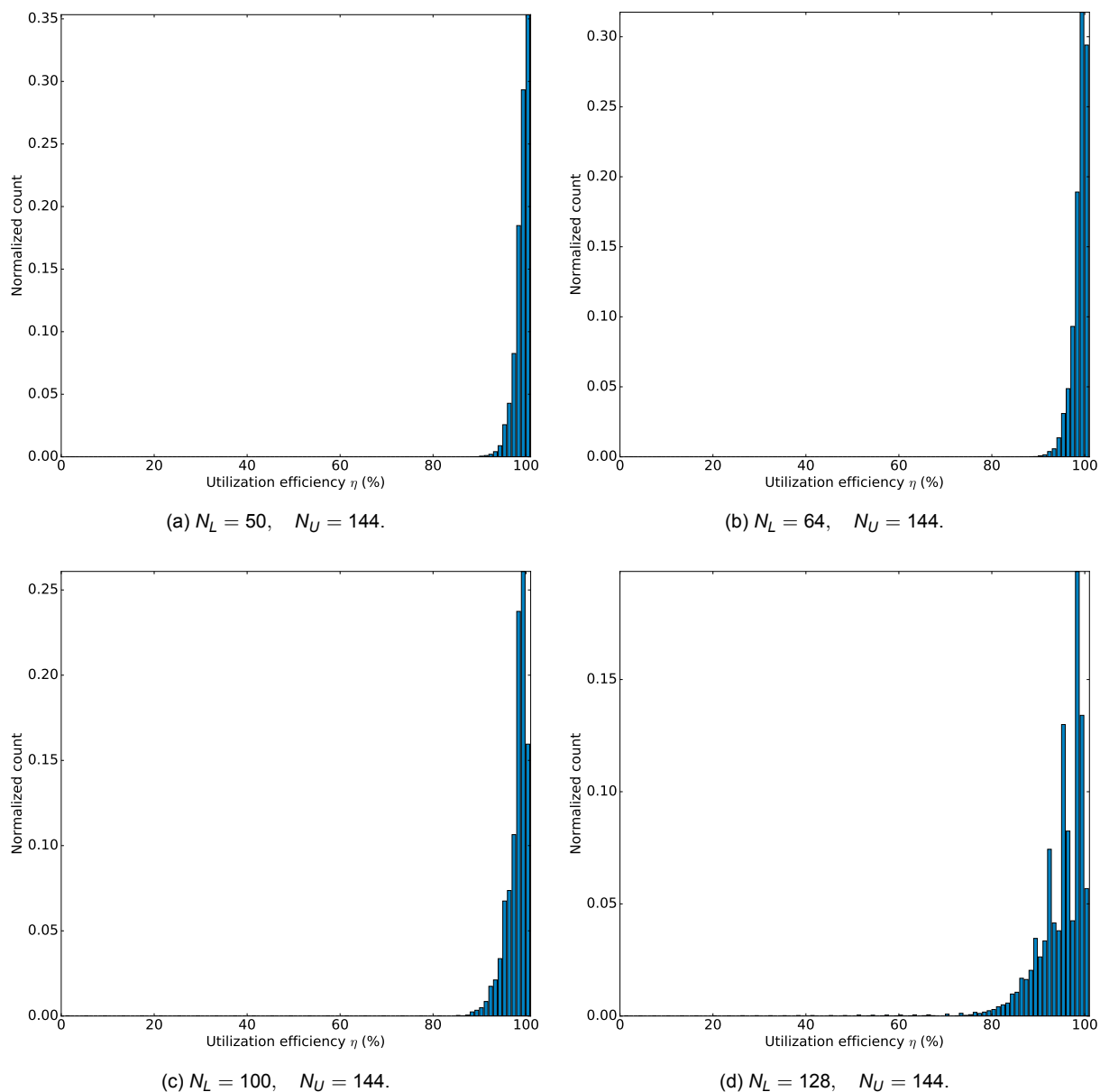


Figure 4: Illustration of DRAM utilization efficiency with a few different values on the partition boundaries.

4 Future Editions

Future editions of this document will include

- a section describing the use of the thresholding filter.

References

- [1] Signal Processing Devices Sweden AB, *14-1351 ADQAPI Reference Guide*, Dec. 2016. Technical Manual.
- [2] Signal Processing Devices Sweden AB, *14-1397 PA8 ADQ14-FWATD Datasheet*, Nov. 2015. Technical Manual.