



# **A Virtual Reality Camera Design with 16 Full HD Video Inputs Sharing a Single DRAM Chip**

## **Introduction**

The development of Virtual Reality (VR) technologies is advancing fast and a need for small, low power and consumer oriented multi-view cameras is anticipated. Such cameras will be required by the consumer in order to help create their own virtual worlds, as well as recording their experience in high resolution at 360 degrees. Indeed designs of such cameras have been announced by both [Google](#) and [Facebook](#) as well as others.

These cameras, however, come with significant technical challenges, especially if designed for the consumer market where size and power consumption are essential constraints. The challenge, for the designer, is to capture and store the data coming from 8, 16 or more full HD cameras.

Image compression technology is an obvious choice here. However, while intra frame methods like JPEG have low complexity, they still result in high bandwidth, especially for a large number of cameras and/or high frame rate. On the other hand, more advanced compression schemes, such as H.264, require reference frame(s) storage and this, in practice, means DRAM. In fact, due to the large bandwidth required, one chip of DRAM for each camera is often necessary. This raises costs as well as power requirements for consumer applications to prohibitive levels.

This paper will describe a realistic design that allows the real time H.264 compression of 16 full HD (1080p@30) video streams, using both I and P frames, sharing the bandwidth of a single DDR3 DRAM chip with 16 bit data bus. This can be achieved thanks to Ocean Logic's proprietary Compressed Frame Store (CFS) technology that allows perfect reconstruction of the compressed frame store data with compression ratios of 10-20:1.

## **The CFS Technology**

In March 2011 Ocean Logic announced the CFS technology. Key points of this technology are :

- High compression of the reference frame (10:1 or higher)
- No accumulated error/drift when the bitstream produced is decoded by a third party decoder
- Allows extremely low power designs
- Generality of the method that can be applied to multiple video compression algorithms, not just H.264
- No need to change existing video compression standards

In August 2012, after negotiations and evaluation of the technology, a large corporation signed an exclusive agreement to use the CFS technology. This has resulted in a SoC that includes an H.264 1080p@30 video encoder that uses both I and P frames and does not require external DRAM. The reason for the latter is that the reference frame is compressed and stored in internal SRAM thanks to the CFS technology. This SoC will be released soon for a variety of applications.

Ocean Logic has meanwhile received a patent for the CFS technology in China and one in the US is about to be issued. Others countries are pending.

With the exclusivity now expired, Ocean Logic is keen to see applications of this technology in a variety of markets.

## A CFS VR Camera

The proposed design is shown in the figure below. The design comprises 16 Compression Units (CUs) sharing a single DDR3 DRAM chip. For clarity, only some are detailed.

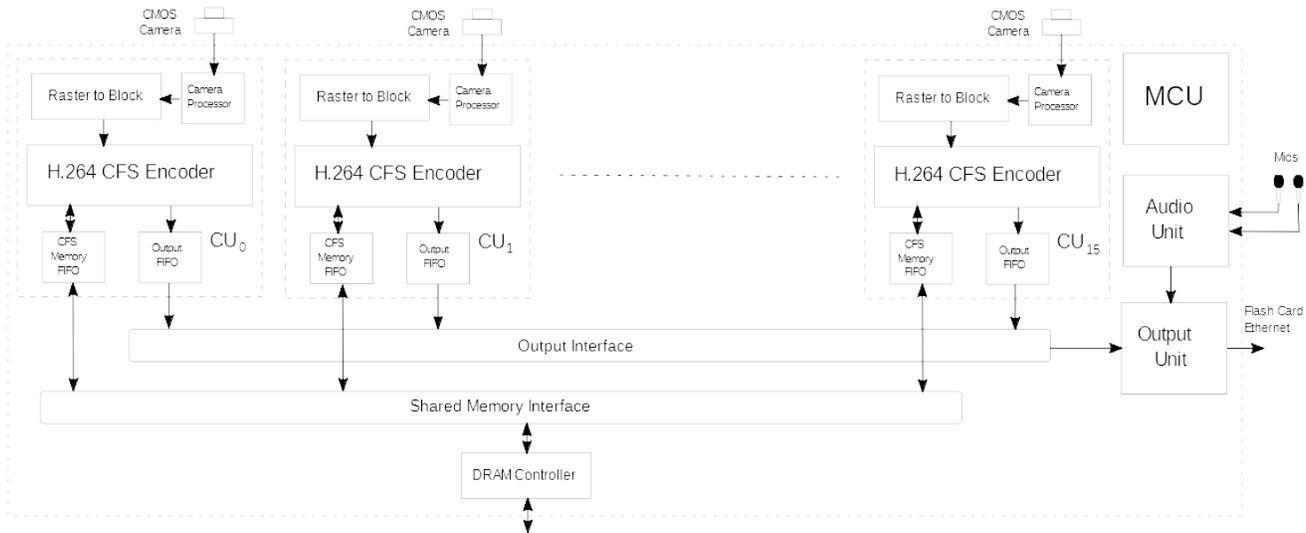


Figure 1 A CFS VR camera design.

Each CU receives video data from a CMOS sensor and outputs a H.264 compliant bitstream. The compressed reference frames each CU needs will be stored on a single chip of DDR3 DRAM, with its access shared amongst all the CUs. Details of the components of each CU, as well as other modules will be given below.

### Camera Processor

The purpose of this module is to interface with the CMOS sensor and process its input. This includes performing Bayer interpolation, white balancing, converting to YUV color space, normalization, gamma correction and sharpening. More complex processors can include lens distortion correction and other functions. This module outputs full HD video as YUV 4:2:0 8 bit pixels in raster order.

### Raster to Block

This module receives the YUV 4:2:0 pixels in raster order and it outputs them as a 16x16 pixels macroblocks consisting of 16x16 luma samples followed by two 8x8 groups of chroma samples which is required by the H.264 CFS encoder.

This block is quite small, but it requires a 16 video line buffer or  $1920 \times 16 \times 1.5 \times 8 = 368,640$  bits of SRAM (for full HD).

### H.264 CFS Encoder

The H.264 CFS encoder processes the YUV 4:2:0 macroblocks from the Raster to Block module and outputs H.264 NALs. This module is the same H.264 CFS encoder that has been proven in silicon with a small modification : instead of the interface to a single port SRAM that normally contains the CFS memory, this core reads and writes compressed frame store data through the FIFOs.

The core takes 1024 cycles to process one macroblock (4 clocks/pixel). This means that, for encoding 1080p@30, the minimum clock frequency is ~250 MHz (~500 MHz for 1080p@60). Both frequencies are achievable in modern ASIC processes.

Its latency (the time from which a macroblock is input to the time the H.264 bitstream that relates to said macroblock) is ~3,000 cycles. However, since it takes 16 video lines to form a macroblock, this is the

dominant latency factor. This is  $\sim 50$   $\mu\text{sec}$  for 1080p@30,  $\sim 25$   $\mu\text{sec}$  for 1080p@60. It is important to note that this time does not include any buffering and transmission delays.

The encoded output is Constant Bit Rate (CBR). This is Hierarchical Reference Decoder (HRD) compliant with the addition of a full-frame encoded data buffer.

This core size is  $\sim 280\text{K}$ gates (1 gate = 1 two inputs NAND) + 217 Kbits of single and dual port SRAM.

### Shared Memory Interface

This is the interface that allows the 16 CUs to share the same DRAM Controller to store their reference frames. Contrary to what one might think, this is not a particularly complicated part of the design. The reason is that the H.264 CFS Encoders access their compressed frame store in an essentially sequential and predictable way. This allows one to statically schedule the memory accesses and assign time slots for each of them, even though 16 distinct processes are competing for the same resource.

### DRAM Controller

This is a run of the mill DDR3 DRAM controller, not particularly fast or sophisticated. It interfaces all the CUs to the DRAM chip that contains their compressed reference frame storage.

### Audio Module

This module takes the input of two microphones and directs digital audio to the Output Unit.

### Shared Output Interface

This is the interface that allows the 16 CUs and the Audio Unit to share the same Output Unit.

### Output Unit

This unit could consist of a flash memory card interface and/or a Gigabit Ethernet core to store and/or stream the H.264 bitstreams and audio data.

### Micro-Controller

Even though the H.264 CFS encoders do not require any CPU assistance during encoding, a simple, small micro-controller might be necessary for house-keeping and initialization of the configuration registers.

## DRAM Bandwidth

The CFS technology allows for 10-20:1 compression of the reference frame, depending on the desired quality. For full HD and good quality, let's assume a compression ratio of 10:1. Therefore, a 1080p YUV 4:2:0 frame that would normally require  $1920 \times 1080 \times 1.5 \times 8$  bits =  $\sim 24.9$  Mbits, will now require  $\sim 2.5$  Mbits. We immediately notice that 16 compressed frames easily fit in a single DRAM chip.

Also, in order to avoid large internal buffering during motion estimation, each CU reads its compressed frame store 5 times for every frame it processes. So, for 16 CUs operating simultaneously at 30 fps, the total bandwidth will be  $\sim 16 \times 2.5 \times 5 \times 30 = \sim 6,000$  Mbit/s =  $\sim 6$  Gbit/s ( $\sim 12$  Gbit/s for 60 fps).

If we now consider the slowest DDR3 DRAM chip (400 MHz, 800 MT/s) with a 16 bit data bus, we arrive at the purely theoretical bandwidth of 12.8 Gbit/s. Now, even taking into account DRAM inefficiencies, it seems that the bandwidth should be sufficient to service the 16 CUs, especially considering their essentially sequential access to the compressed frame store. In any case, since we considered the slowest DDR3 case, there is room to maneuver by raising its clocking frequency.

It follows that the same design, for 1080p@60, should work with a 32 bit bus DRAM chip or doubling the DRAM frequency (800 MHz, 1600 MT/s).

## Output Bandwidth

This design aims for a reasonably good video quality (i.e.  $\sim 18$  Mbit/s for each 1080p@30 stream). This would mean a total of  $\sim 16 \times 18 = 288$  Mbit/s. The contribution of two audio streams should be negligible to the total output bandwidth. Flash cards capable of supporting this bandwidth are certainly available. The output bandwidth is also compatible with streaming through Gigabit Ethernet.

One hour of video recording from the 16 full HD cameras should fit in  $\sim 128$  Gbyte of flash memory.

## Resources Summary

This is just a blueprint for a design and it is therefore difficult to give a precise summary of the resources required. However, it is clear that the bulk of it is constituted by the 16 H.264 CFS encoders (total of  $\sim 4.5$  Mbytes + 3.5 Mbits of single and dual port RAMs) and the 16 raster to block SRAMs (total of  $\sim 6$  Mbits of single port RAM). The remaining logic is unlikely to be more than 10-15% of the whole design.

It is also worth noting the low number of pins required. In fact, most of the I/Os would consist of the 16 camera interfaces and the DDR3 DRAM interface with 16 bit data bus.

## Conclusions

This white paper has described the use of the CFS technology for a VR camera recording 16 full HD (1080p@30) streams simultaneously. The heavy compression of the reference frame store allows a single DDR3 DRAM chip to be shared amongst the 16 H.264 CFS Encoders.

An equivalent design that does not use CFS technology and uses only a moderate level of compression (1.5-2x) of the reference frame or no compression at all, would necessitate an external DRAM chip for each encoder, resulting in a substantially higher power requirement, in the order of one order of a magnitude or more.

### **Ocean Logic Pty Ltd**

PO BOX 768 - Manly NSW 1655 - Australia

E-Mail: [contact@ocean-logic.com](mailto:contact@ocean-logic.com) URL : <http://www.ocean-logic.com/>