**1. Redesign the post process logic**

In the original design, the DLA computes the post process, like bias, batch normalization, element-wise addition and multiplication, after the convolution or FC produced. The PE array (for convolution and FC) and post-process elements are pipelined for real-time ASR. However, this design requires sufficient sources.

In wavenet, the number of multiplication is detailed in Eq.1

 (1)

Where Wout is the length of output time sequence. K is the Kernel Size. Cin and Cout are the numbers of input and output channel, respectively.

However, the number of multiplication in BN or Element-wise operation or Activation Function, such as sigmoid and tanh, is calculated in Eq.2 as

 (2)

Hence, the computation in post process only accounts for a relatively modest portion of the whole layer.

Similar result can be obtained in FC/LSTM.

Therefore, I make a compromise between the one-time pipeline logic and one-after-another operation flow in reference [1]. The bias or BN can be computed at once when the convolution or FC is finished. However, the features should be accessed again from the Feature Buffer to do the element-wise operation. In this way, 32 x 2 floating-point multipliers and adders are saved and the post-processed logic are fully loaded when convolution or FC is under calculation.

Reference:

[1] Laika: A 5uW programmable LSTM accelerator for always-on keyword spotting in 65nm CMOS.

**2. Transmit the DLA SoC slave interface to AXI bus and the DLA DDR master interface to PHI bus for DDR.**

This work is undergoing. As my current laptop can not P&R the relative Xilinx IPs because of the meager CPU and DDR capacity, I am reviewing the previous work in Tanji3. I will get a more power computer in Friday. Hope it can run the Vivado P&R.

**3. If anyone can enter the lab, please reset my computer to UNIX and start the teamvierwe.**