# 注意：data\_reorganize未提供，在Demo中是在arm实现的

# 加速器与HPS连接关系



# 功能简介

CNN卷积模块（cnn\_top）实现普通卷积、dw卷积、relu激活等功能。

数据重排模块（data\_reorganize）实现了输入数据重排功能与输出数据重排功能，原始图像的数据要经过输入重排，将数据排列成CNN加速器卷积需要的数据格式进行卷积计算；当框架要使用卷积结果时，需要将卷积结果进行输出重排，框架使用输出重排的结果进行识别。

# 工作流程

1、Hps通过avalon总线向数据重排模块（data\_reorganize）下发：输入重排模式、通道数、特征图尺寸、输入数据地址、输出数据地址，配置完成后，再下发开始命令；数据重排模块会将原始图像数据排列成CNN加速器需要的数据格式。

2、Hps通过avalon总线向CNN卷积模块（cnn\_top）下发：parameter、scale、input、weight、output在DDR中的基地址，然后在下发开始命令；CNN卷积模块启动对数据进行卷积。

3、卷积完成后，框架要使用卷积结果需要进行输出数据重排，Hps通过avalon总线向数据重排模块（data\_reorganize）下发：输数重排模式、通道数、特征图尺寸、输入数据地址、输出数据地址，配置完成后，再下发开始命令；数据重排模块将数据排列成框架能使用的数据模式。

# 接口说明

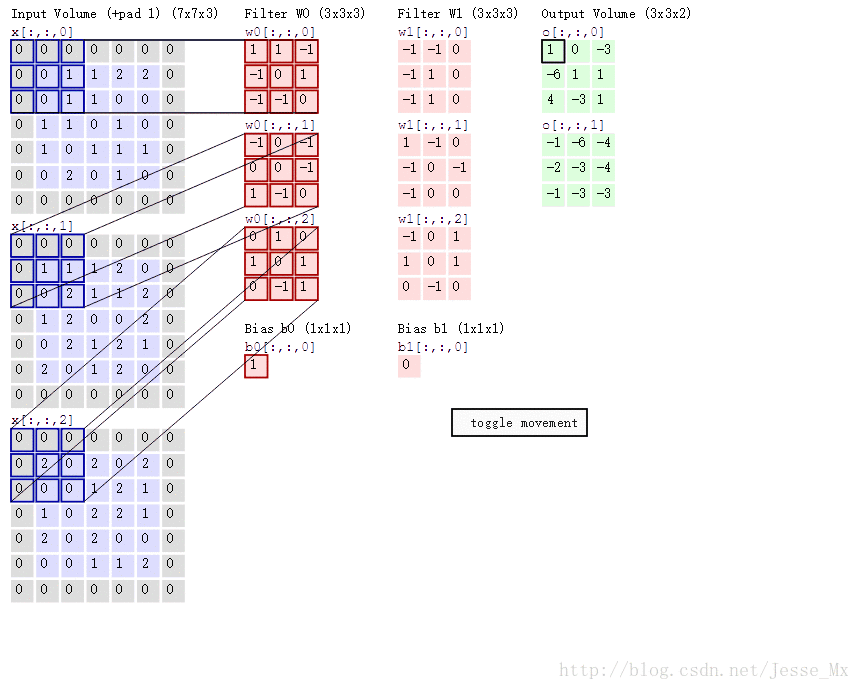
#### CNN接口说明

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 信号名 | 信号位宽 | 端口类型（I/O） | 信号来源/去向 | 详细描述 |
| sysclk | 1 | I | PLL | 系统时钟 |
| rst\_n | 1 | I | external | 系统复位 |
| as\_rd | 1 | I | HPS | 与HPS的HPS连接的avalon总线从机接口，接收parameter、scale、input、weight、output在DDR的基地址以及卷积开始命令 |
| as\_wr | 1 | I |
| as\_addr | 3 | I |
| as\_rd\_data | 32 | O |
| as\_wr\_data | 32 | I |
| as\_data\_waitquest | 1 | O |
| cfg\_avm\_read | 1 | O | DDR | 与HPS的DDR控制器连接的avalon总线，从DDR中读取卷积控制参数 |
| cfg\_avm\_address | 32 | O |
| cfg\_avm\_readdata | 32 | I |
| cfg\_avm\_readdatavalid | 1 | I |
| cfg\_avm\_byteenable | 4 | O |
| cfg\_avm\_burstcount | 3 | O |
| cfg\_avm\_waitrequest | 1 | I |
| load\_avm\_read | 1 | O | DDR | 与HPS的DDR控制器连接的avalon总线，从DDR中读取input、weight |
| load\_avm\_address | 32 | O |
| load\_avm\_readdata | 128 | I |
| load\_avm\_readdatavalid | 1 | I |
| load\_avm\_byteenable | 16 | O |
| load\_avm\_burstcount | 3 | O |
| load\_avm\_waitrequest | 1 | I |
| scale\_avm\_read | 1 | O | DDR | 与HPS的DDR控制器连接的avalon总线，从DDR中读取store模块计算需要的scale、weight\_scale参数 |
| scale\_avm\_address | 32 | O |
| scale\_avm\_readdata | 32 | I |
| scale\_avm\_readdatavalid | 1 | I |
| scale\_avm\_byteenable | 4 | O |
| scale\_avm\_burstcount | 3 | O |
| scale\_avm\_waitrequest | 1 | I |
| output\_avm\_write | 1 | O | DDR | 与HPS的DDR控制器连接的avalon总线，将卷积计算结果写回DDR中 |
| output\_avm\_address | 32 | O |
| output\_avm\_writedata | 128 | O |
| output\_avm\_byteenable | 16 | O |
| output\_avm\_waitrequest | 1 | I |
|  |  |  |  |  |

#### DATA\_REORG接口说明

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 信号名 | 信号位宽 | 端口类型（I/O） | 信号来源/去向 | 详细描述 |
| exclk | 1 | I | externa | 外部时钟 |
| rst\_n | 1 | I | external | 外部复位 |
| as\_rd | 1 | I | HPS | 与HPS连接的avalon总线从机接口， |
| as\_wr | 1 | I |
| as\_addr | 8 | I |
| as\_rd\_data | 32 | O |
| as\_wr\_data | 32 | I |
| as\_data\_waitquest | 1 | O |
| r\_avm\_waitrequest | 1 | I | DDR | 与HPS的DDR控制器连接的avalon总线 |
| r\_avm\_read | 1 | O |
| r\_avm\_address | 32 | O |
| r\_avm\_readdata | 64 | I |
| r\_avm\_readdatavalid | 1 | I |
| r\_avm\_byteenable | 8 | O |
| r\_avm\_burstcount | 5 | O |
| w\_avm\_waitrequest | 1 | I | DDR | 与HPS的DDR控制器连接的avalon总线 |
| w\_avm\_write | 1 | O |
| w\_avm\_address | 32 | O |
| w\_avm\_writedata | 64 | O |
| w\_avm\_byteenable | 8 | O |
| w\_avm\_burstcount | 5 | O |

# 数据重排说明



如图展示的是一个3输入通道，2输出通道的卷积处理流程。

FPGA要实现这样的功能是比较简单的，简单计算一下可以得出大概需要：

存储空间：219byte

乘法器：9\*3个

加法器：10个

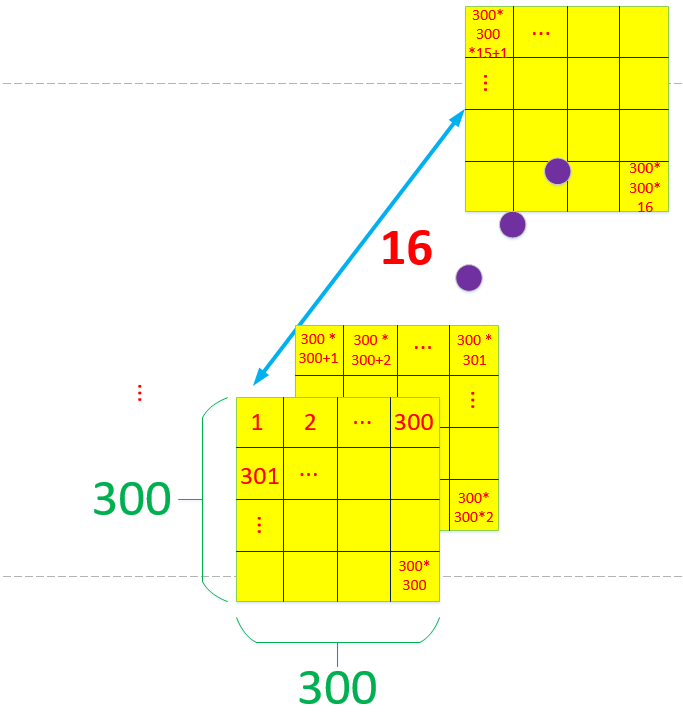
当特征图尺寸变大、输入通道、输出通道变多之后，要实现卷积功能所需要的存储空间、乘法器、加法器数量将以几何倍数增。MobileNetV1模型特征尺寸最大为224\*224，输入输出通道最多为1024，按照上面的卷积方法在FPGA上难以实现



为解决FPGA资源问题进行通道接切片减少乘法器数量，如下图所示，将输入通道和输出通道划分成一个个通道切片，每次对一个输入/输出通道进行卷积。



一个切片的数据排列如下图所示：



这样的排列在CNN加速器读取数据时，因为地址并不连续所以读取效率低下，所以需要数据重新排成下图的数据格式方便CNN加速器读取数据的连续性，提升数据读取的效率。

