

第八届

全国大学生集成电路创新创业大赛

报告类型： 设计、仿真报告

参赛杯赛： 中科芯杯

作品名称： Hydra-高速多端口共享缓存管理模块

队伍编号： CICC1651

团队名称： Hydra

**目录**

[1.概述 4](#_Toc172841190)

[1.1题目再述 4](#_Toc172841191)

[1.2项目简介 4](#_Toc172841192)

[2.基本参数 5](#_Toc172841193)

[3.项目基本框架 6](#_Toc172841194)

[3.1动态缓存管理 6](#_Toc172841195)

[3.1.1 SRAM颗粒（页）划分 6](#_Toc172841196)

[3.1.2内存回收 7](#_Toc172841197)

[3.2数据包管理 8](#_Toc172841198)

[3.2.1链表维护优先级队列中的数据包 8](#_Toc172841199)

[3.2.2严格优先级调度 9](#_Toc172841200)

[3.2.3 WRR调度 9](#_Toc172841201)

[3.3 数据校验 9](#_Toc172841202)

[3.3.1汉明校验 9](#_Toc172841203)

[3.3.2校验信息存储 9](#_Toc172841204)

[4.项目亮点 10](#_Toc172841205)

[4.1结合Crossbar架构与总线架构 10](#_Toc172841206)

[4.2跳转表代替链表维护优先级队列 10](#_Toc172841207)

[4.3为数据包匹配较优SRAM，缓解读出冲突 12](#_Toc172841208)

[4.4设置写入前端缓冲结构，支持断点续传 15](#_Toc172841209)

[4.5基于尾部预测的快速拼接机制，支持单包边读边写 16](#_Toc172841210)

[4.6多匹配模式支持，可设置全动态、半动态、静态模式 17](#_Toc172841211)

[4.7保障保序性，消除传统仲裁的缺陷 17](#_Toc172841212)

[4.8 “安抚”读取机制，降低读出延迟 18](#_Toc172841213)

[4.9基于掩码集的无需复杂计算的快速WRR 19](#_Toc172841214)

[4.10并行汉明校验，实现单周期编码解码 20](#_Toc172841215)

[4.11其他亮点 21](#_Toc172841216)

[5.模块介绍 22](#_Toc172841217)

[5.1各模块说明 22](#_Toc172841218)

[5.1.1 hydra(顶层模块) 22](#_Toc172841219)

[5.1.2 port\_wr\_frontend 23](#_Toc172841220)

[5.1.3 port\_wr\_sram\_matcher 23](#_Toc172841221)

[5.1.4 port\_rd\_frontend 24](#_Toc172841222)

[5.1.5 port\_rd\_dispatch 24](#_Toc172841223)

[5.1.6 sram\_interface 24](#_Toc172841224)

[5.1.7 sram 25](#_Toc172841225)

[5.1.8 ecc\_encoder 25](#_Toc172841226)

[5.1.9 ecc\_decoder 26](#_Toc172841227)

[5.2重要逻辑说明 27](#_Toc172841228)

[5.2.1数据包写入流程 27](#_Toc172841229)

[6.接口与配置 29](#_Toc172841230)

[6.1.写控制IO口 29](#_Toc172841231)

[6.1.1 IO口介绍 29](#_Toc172841232)

[6.1.2 使用方法 29](#_Toc172841233)

[6.2写反馈IO口 29](#_Toc172841234)

[6.2.1 IO口介绍 29](#_Toc172841235)

[6.2.2 使用方法 29](#_Toc172841236)

[6.3读控制IO口 30](#_Toc172841237)

[6.3.1 IO口介绍 30](#_Toc172841238)

[6.3.2 使用方法 30](#_Toc172841239)

[6.4读反馈IO口 30](#_Toc172841240)

[6.4.1 IO口介绍 30](#_Toc172841241)

[6.4.2 使用方法 30](#_Toc172841242)

[6.5配置选项 30](#_Toc172841243)

[7.验证方法 31](#_Toc172841244)

[7.1 RTL/行为级仿真 31](#_Toc172841245)

[7.1.1 RTL级电路图 31](#_Toc172841246)

[7.1.2 Feature验证 31](#_Toc172841247)

[7.1.3压力测试 35](#_Toc172841248)

[7.2硬件实现验证 37](#_Toc172841249)

[7.2.1 验证环境 37](#_Toc172841250)

[7.2.2 综合布线结果 37](#_Toc172841251)

[7.3 FPGA验证 39](#_Toc172841252)

[7.3.1 Ecc校验编解码模块 39](#_Toc172841253)

[8.设计优缺点 40](#_Toc172841254)

[8.1优点 40](#_Toc172841255)

[8.2缺点 40](#_Toc172841256)

[9.后续开发计划 41](#_Toc172841257)

[9.1进一步优化读写延迟 41](#_Toc172841258)

[9.2更全面的压力测试 41](#_Toc172841259)

[9.4继续提升代码质量 41](#_Toc172841260)

# 概述

## 1.1题目再述

赛题要求设计一款可对SRAM进行有效管理的SRAM控制器IP。具体要求如下：

1. 支持管理至少32块256K bit的SRAM单元，总计至少8Mbits存储容量；
2. 支持16个端口同时进行读写操作，每个端口的传输带宽需达到1Gbps；
3. 每个端口支持8个优先级队列，实现按队列进行数据缓存；
4. 支持按数据包进行缓存和调度，数据包长度范围为64到1024字节；
5. 支持多端口、多队列的动态共享缓存；
6. 支持多个端口同时写入和写出数据，每个端口独立操作，相互之间不受影响；
7. 支持数据校验，确保数据传输的准确性和完整性。

## 1.2项目简介

Hydra是一款高速多端口共享缓存管理模块，其基础功能包括但不限于：

1. 管理16个端口独立同时读写32块SRAM；
2. 每端口8个队列按优先级调度数据包；
3. 所有端口所有队列动态分配缓存；
4. 对SRAM存储的数据进行SEC数据校验；
5. 提供严格优先级、WRR两种调度模式。

**图1 模块示意图**

Hydra在实现基础功能的同时，设计亮点包括但不限于：

1. 改进Crossbar架构交换机的诸多缺陷；
2. 引入跳转表的思想，降低维护优先级队列所需的资源；
3. 采用特殊的匹配机制，降低由于读出冲突导致的读延迟；
4. 添加暂停传输信号，支持断点续传；
5. 基于尾部预测的快速拼接，支持单数据包级别的边读边写；
6. 多种匹配模式支持，集合动静态分配功能于一体；
7. 高度保序，消除传统仲裁写入的缺陷；
8. 采用位掩码实现极低复杂度的WRR机制
9. 并行优化的数据校验加码-解码逻辑。

综上所述，Hydra在完全实现题目基本要求的前提下，采用大量设计优化模块空间占用，降低读写延迟，提升模块性能，满足赛题背景中关于网络设备高速数据缓存的需求。

# 2.基本参数

**时钟频率**：25Mhz~400Mhz

**端口数量**：16

**SRAM数量**：32

**单块SRAM规格**：双口256Kbits (16×16384)

**读写带宽（250MHz）**：

写入带宽（理论、测试）：3.52Gbps/端口

读出带宽（理论、测试）：3.52Gbps/端口

总计（理论、测试）：56.3Gbps

**读写延迟**：

按数据包头：8~24周期+匹配阈值（可配置）

按数据包尾：0周期（可提前/立即读出）

**突发传输支持**：

未拉高pause信号时，端口任意时刻均支持突发传输。在保持一直读出数据包的情况下，支持长期突发传输。压力测试已通过>260us的突发传输。

**主要存储资源占用情况**（合计约2Mbits，对应88块片上BRAM）：

跳转表：32\*32Kbits = 1Mbits (32\*36Kbits BRAM)

空闲队列：32\*22Kbits = 704Kbits (32\*36Kbits BRAM)

ECC校验信息：32\*16Kbits = 512Kbits (32\*18Kbits BRAM)

写入前端缓冲区：16\*1Kbits = 16Kbits (16\*18Kbits BRAM)

**功率**：2.354W

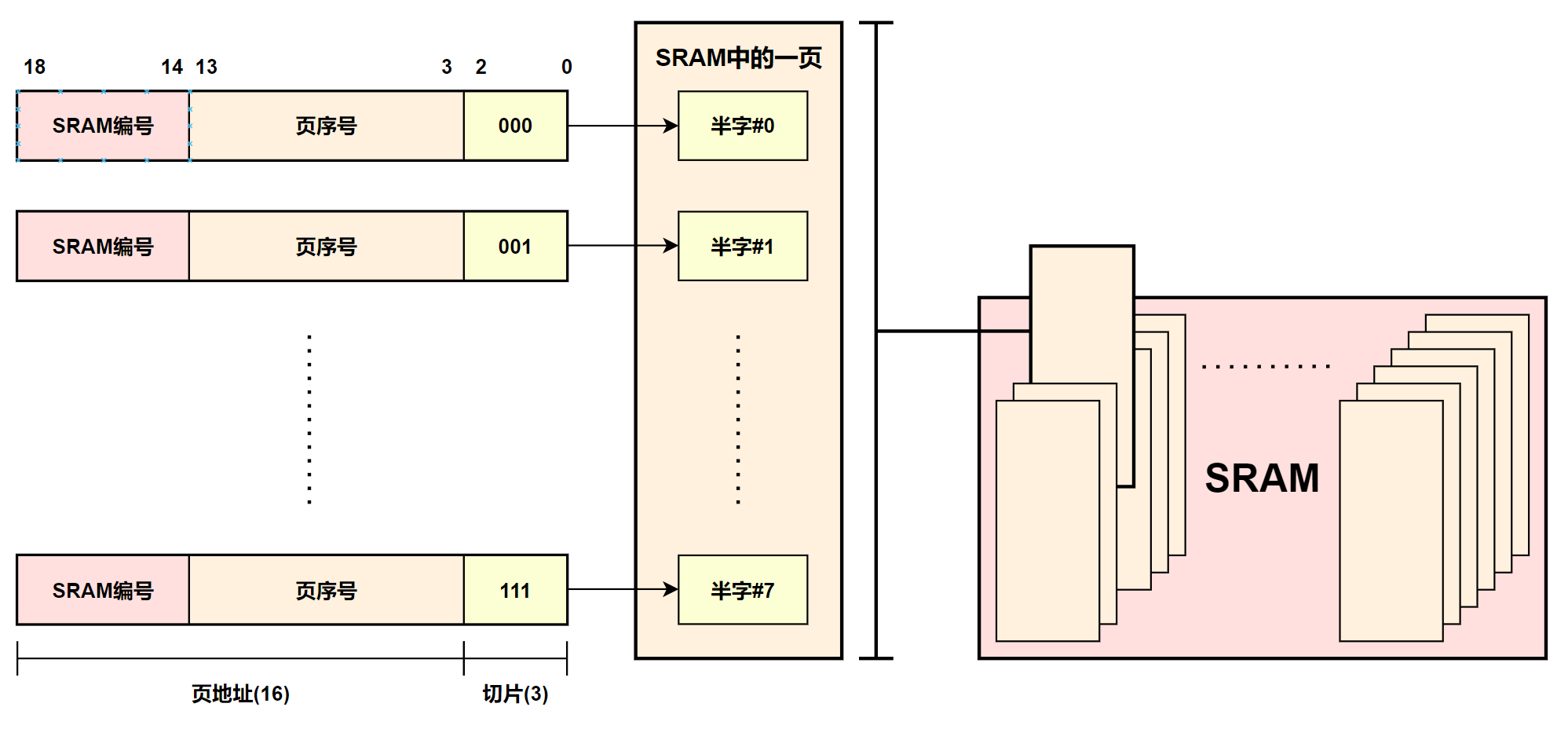
# 3.项目基本框架

## 3.1动态缓存管理

### 3.1.1 SRAM颗粒（页）划分

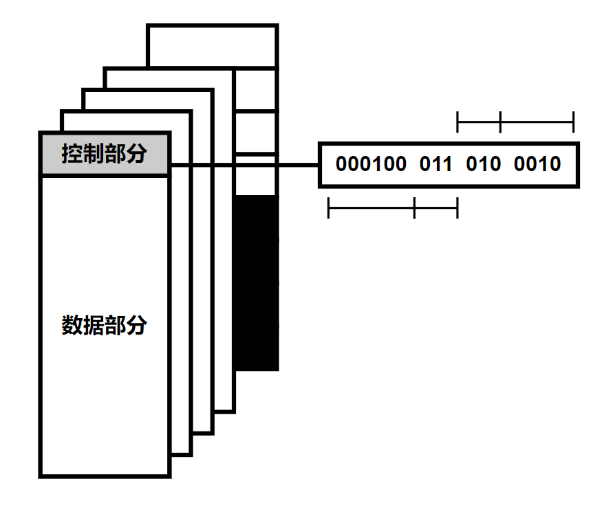
对于每个SRAM，物理地址宽度为14，要选取所有SRAM中一个半字需要5+14=19位宽的地址，不利于存储与管理。若对每半字的数据进行校验运算，由公式可知至少需要5位校验位，所有数据所需的校验资源高达总SRAM容量的5/16。

为了便于存储地址、压缩校验功能所需的资源，Hydra将SRAM的资源划分成128位的小颗粒，每颗粒称为1页，即每8半字为1页。此时指向SRAM中一页的地址被压缩成14-3=11位，选取所有SRAM中一页需要5+11=16位宽的地址，刚好与一半字对齐，便于存储。



**图3.1.1(1) SRAM颗粒划分**

校验也以页为单位，根据公式，每128位生成8位校验码，生成所需时间相比于16-5时间上的差别可忽略不计（3次异或运算的时间），且校验占用的资源大幅下降，只需总SRAM容量的1/16。

 同时注意到一个数据包的长度折合后为4~64页/32~512半字。即只需用9位描述一个数据包的长度，即一个数据包有多少字节。9位长度信息与3位优先级信息、4位目的端口号信息结合，刚好为一个半字（9+3+4=16）。将其作为控制信号的一部分输入，可以使数据包的控制信息与有效数据之间能清晰地划分开来，处理时无需进行二次切割、拼接，复杂度大大下降。

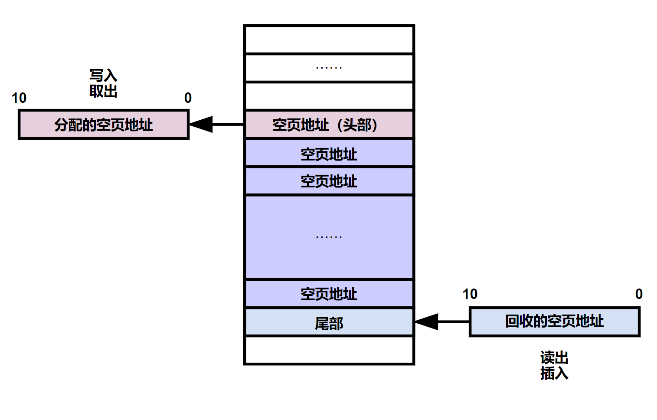
**图3.1.1(2) 数据包控制部分格式**

数据包存储占用的最后一页可能并不会被使用完全，即有几半字的区域无有效数据。为了对齐页的划分，这些空缺的半字并不会被利用（但并非永远浪费，数据包读出，占用的页被回收后，它们仍可以被新的数据包装填），当数据包长度随机时，被浪费的空间小于2%，即使数据包均为不利好页表划分的小数据包，被浪费的空间也小于5%。

由于模块数据传输均以半字为单位，一页的数据交互需要八个周期，故需要3位的批次计数器（batch）记录当前处理到页中的第几个半字的数据。页地址与批次计数器拼接即可得到半字的地址，根据其线性映射关系，我们称页地址为线性地址，线性地址11位拼接后得到物理地址14位，与前面的数据是吻合的。

### 3.1.2内存回收

传统的内存回收策略是为SRAM建立相应长度的位图（bitmap），其中每一位的数据分别对应SRAM中某单位（半字）是否有数据写入，1表示被占用，0表示未被占用，下称空闲。通过在写入数据时置1，读出数据时置0，即可描述SRAM的空闲位置，新来的数据只需直接写入空闲的位置。但是搜索位图中0的位置是一个时序性不良好的操作，即使经过独热码转化后，仍需通过遍历操作才能得到一个空闲位置。若要维护时序性，则需要等待较长的时间才能搜索到空闲位置，这与高速低延迟缓存管理模块的设计理念相悖。

Hydra采用的方案是为每一个SRAM维护一个“空闲队列”，其本质是一个FIFO，存储着空闲的线性地址，在写入数据时只需从队列头取出地址，读出数据时回收页，将地址插入队列尾，即可实现一个时间复杂度为O(1)的回收机制。不过其可观的时序性需要牺牲一定资源。目前采用的空闲队列FIFO深度为2048，宽度为11（线性地址的宽度），32块SRAM共需要32\*22Kb的存储资源，合32块BRAM。

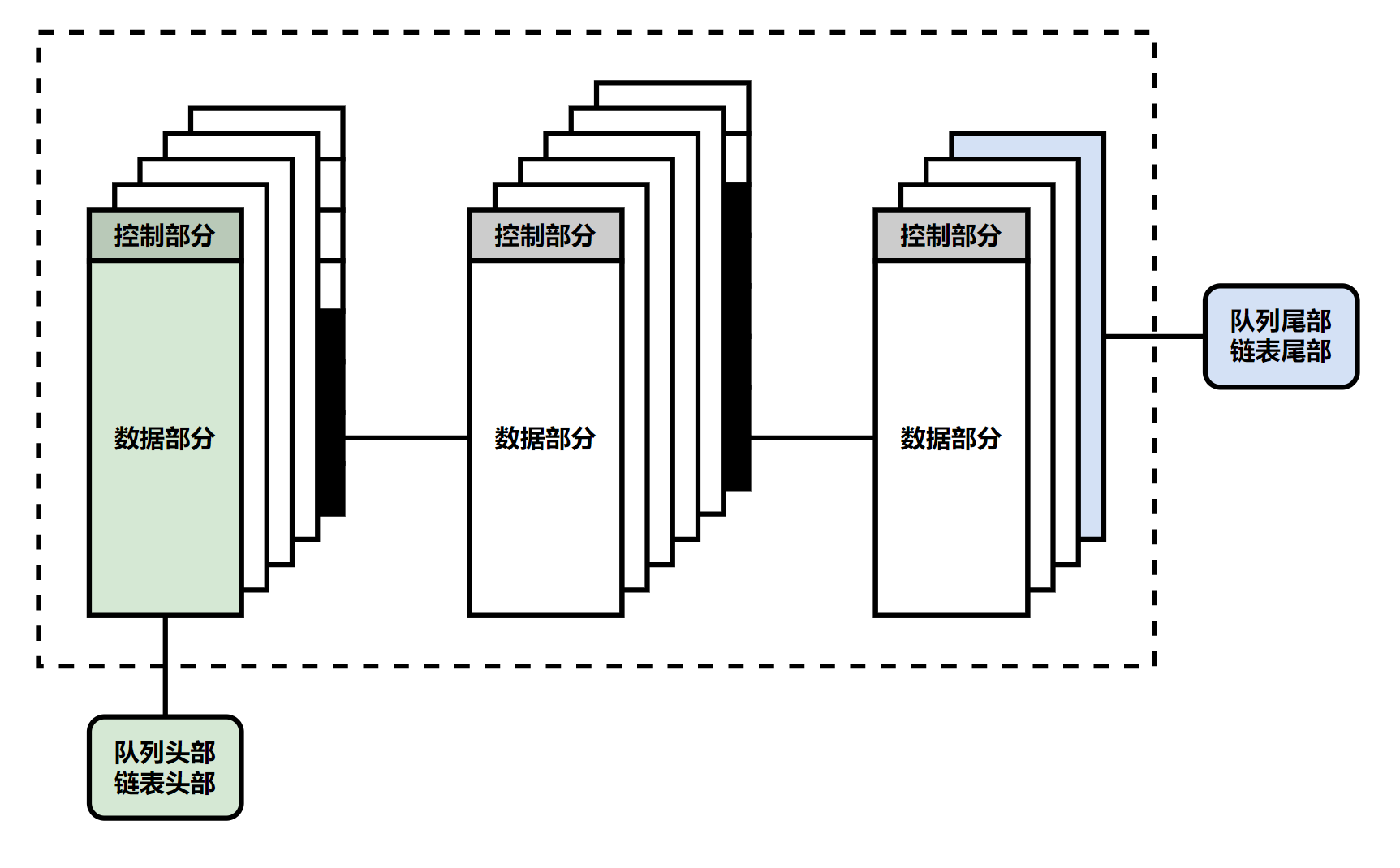
**图3.1.2 空闲队列实现内存回收与分配**

Hydra基于上述空闲队列指导的内存回收，实现了完全动态的空间分配机制，具体表现为可以在任意时刻（有空闲空间的时刻）无延时地申请一块空间，进行实时的数据存储。这使得只要还剩下空闲的空间，任意端口就可以写入数据，与传统的静态分配策略相比，带宽更高且更灵活（例如吞吐量大的端口占有更多的资源）。

## 3.2数据包管理

### 3.2.1链表维护优先级队列中的数据包

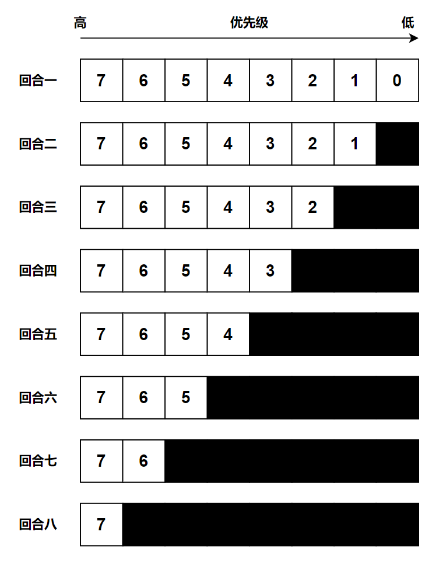
由于Hydra动态地分配内存，数据包每一部分存放的位置可能不连续，因此对于每个数据包，需要维护一个链表，记录每一页的地址。写入时通过空闲队列得到空闲页，写入数据后将其地址加入链表；读出时遍历链表中的节点即可。



**图3.2.1 单向链表维护优先级队列**

由于数据包读出时仅考虑优先级和写入时的先后顺序，因此一个优先级队列就可以用一个单向链表维护，其中存储的数据包是一条条“小链”。单向链表与传统的链表结构相比，每个结点仅存储指向下个结点的指针信息，易于管理与存储。

### 3.2.2严格优先级调度

 采用严格优先级调度时，端口先按队列优先级，再按写入时的先后顺序读出数据包。

### 3.2.3 WRR调度

采用WRR调度时，端口引入了回合权重，不完全按队列优先级读出数据包。Hydra支持一种经典的WRR调度，回合制处理读出队列中的数据包。对于每个端口，第一回合所有队列从高到低轮流读出数据包；第二回合优先度较高的七个队列从高到低轮流读出数据包；……以此类推，八个回合结束后从第一回合重新开始，在保证优先级高的队列享有较大读出带宽的同时，解决了低优先级被高优先级阻塞的问题。

**图3.2.3 WRR调度示意图**

## 3.3 数据校验

### 3.3.1汉明校验

Hydra采用了(136,128)汉明校验，支持SEC（single error correction，单错误纠错）。



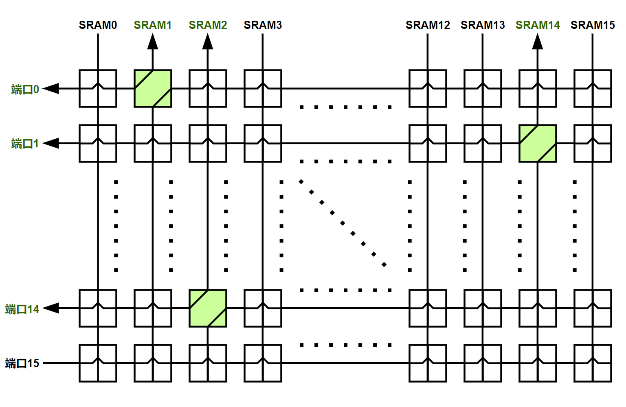
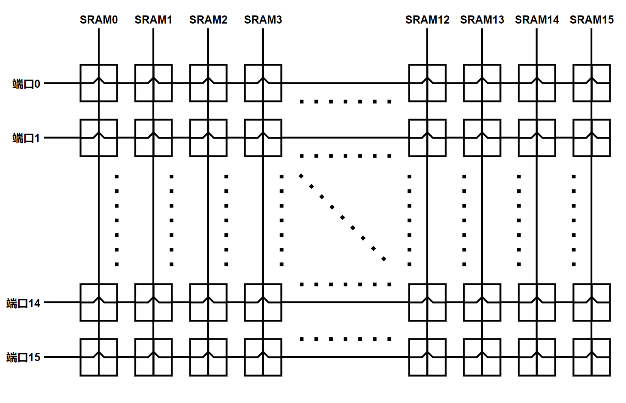
**图3.3.1 (136,128)汉明校验机制示意图（H-矩阵可视化）**

### 3.3.2校验信息存储

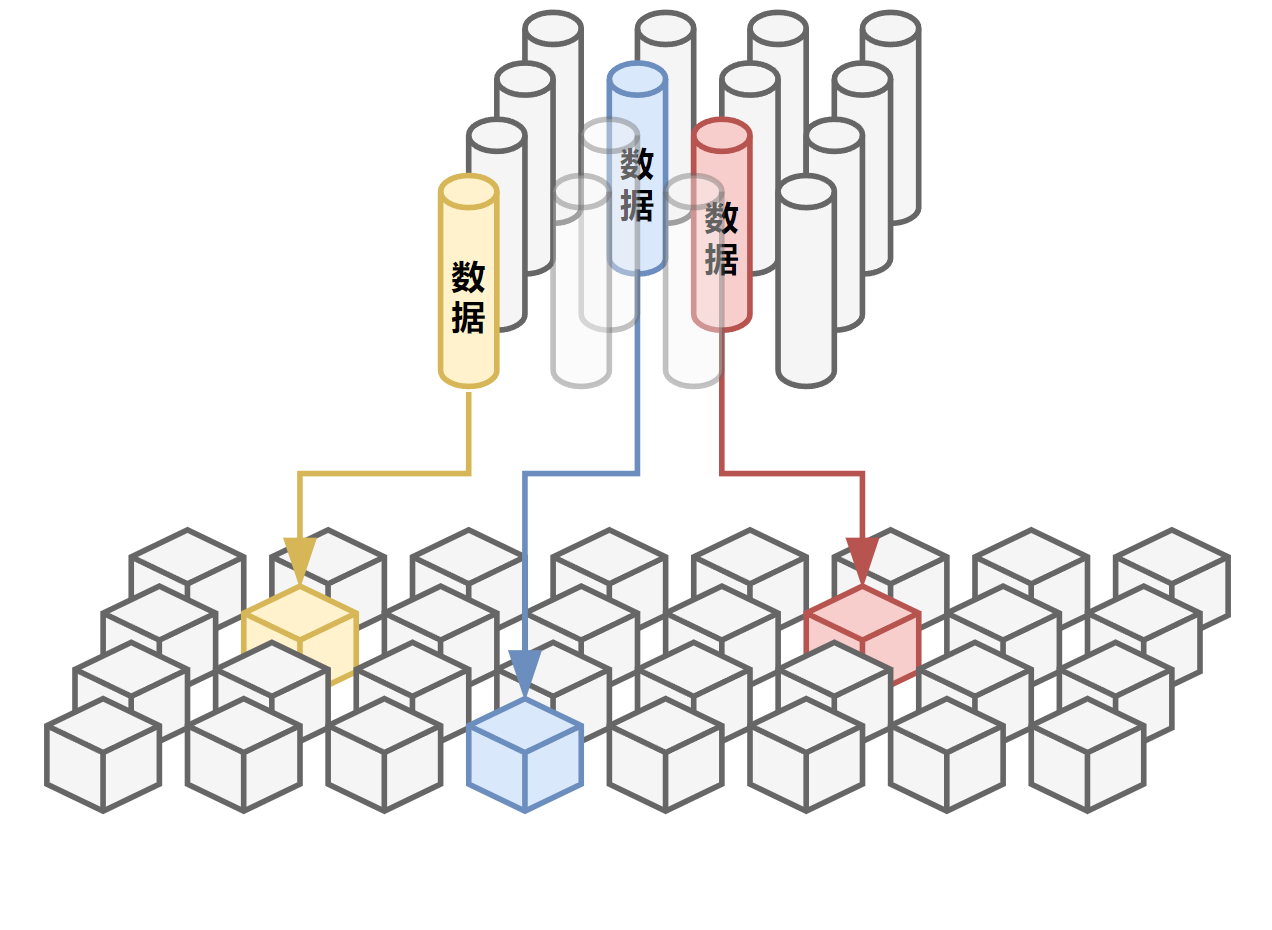
由于校验以页为单位，故校验信息的存储也以页为单位，每一个页地址指向了一个8位的校验码。每个SRAM都有一个校验存储空间，大小为2048\*8=2KB。由于同时只会有一个端口与SRAM进行交互，所以同时也只会有一个端口与校验存储空间进行交互，故校验存储可被置于片上资源的Block RAM。

# 4.项目亮点

## 4.1结合Crossbar架构与总线架构



**图4.1(1) 经典Crossbar电路实现端口与SRAM间的数据传输**

Hydra借鉴了现代路由器微架构，采用**Crossbar架构**（见上图）实现了端口与SRAM之间的数据传输机制（见右图16个端口和SRAM之间自由建立的数据通道示意图），同时对Crossbar结构进行了一定的改进，并结合了**总线架构的优点**，包括但不限于：

1. 端口与SRAM之间可以点对点传输数据而不相互阻塞，支持各端口**独立无限制突发传输数据**直到无法找到可用的SRAM；

**图4.1(2) 端口与SRAM间**

**自由建立的数据通道**

1. 完全避免Crossbar和总线结构都具有的写入仲裁缺陷，同时有数据包写入时，Hydra可**并行处理并传输**，无需建立仲裁机制逐个处理；
2. 点对点传输的建立也能具有较为**复杂的策略**：Hydra在数据包写入前设置了较为复杂的匹配机制，在不影响点对点传输的前提下，缓解了读出冲突现象。

以上机制的具体实现见下文。

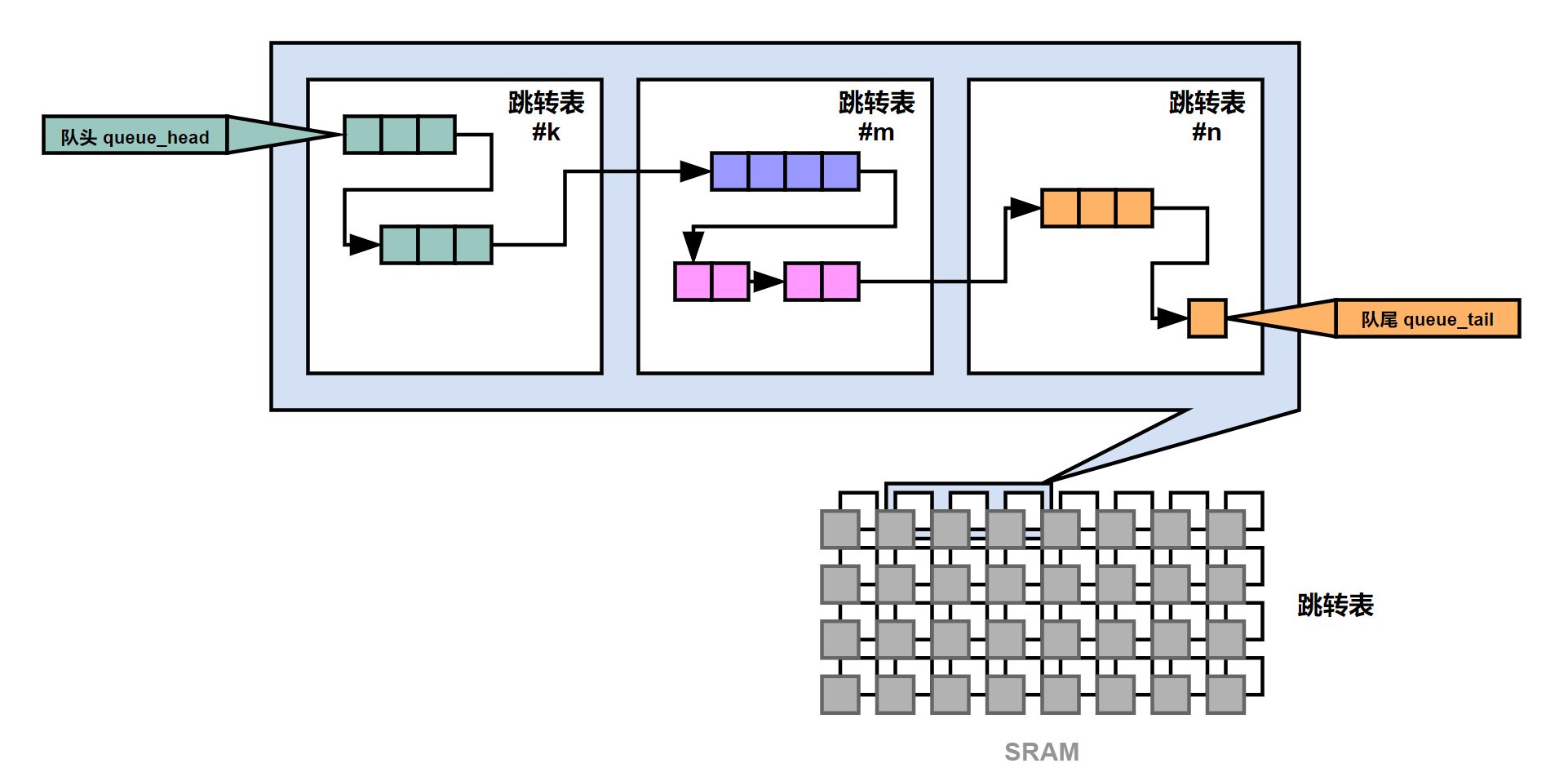
## 4.2跳转表代替链表维护优先级队列

由于Hydra采用动态分配缓存的策略，故一个优先级队列中的数据包可能存放在不相邻的地方，因此需要专门存储所有数据包页地址。由于同一个优先级队列遵循先进先出的原则，可使用FIFO进行存储。

上述基础框架中提到的传统方案是给每个优先级队列建立一个FIFO管理队列中数据包的页地址，写入新的页时向FIFO末端插入页地址；读出页时从FIFO首端弹出页地址。但是由于动态分配的不平衡性，可能会存在一个优先级队列数据包极多，但是其他优先级队列几乎没有数据包的情况。若要支持最极端的条件（即所有空间都被一个优先级队列的数据包占用），每个队列的FIFO的深度需要65536，每个元素宽16位，记载了一个带SRAM编号的线性地址（5+11）。所有FIFO的存储资源共128\*65536\*16bits=**16MB**，这是令人无法接受的。若酌情减少FIFO的深度，则会导致队列**有数据量限制**，**无法做到完全的动态分配**。

Hydra采用的方案是为每个SRAM建立一个**“跳转表”**，基于链表数据结构的思想进一步改进传统方案中的单向链表：将队列中的页地址都视为一个结点，每个结点存储了下一结点的地址信息，这样就可以将**所有优先级队列存储在一起**，称为“跳转表”（因为每个结点都存储了跳转到下一个结点的信息）。读取队列中数据时，只需查找跳转表中当前页地址对应的内容，即可得到下一页的地址，再根据跳转表中下一页地址对应的内容，即可得到下下页的地址，以此类推，对于每个优先级队列，只需维护其队头的页地址，即可以按顺序访问队中的所有元素；向队列中新增数据时，只需将原来队尾页的跳转信息指向新的页地址，将新的页地址设置为新的末端即可。

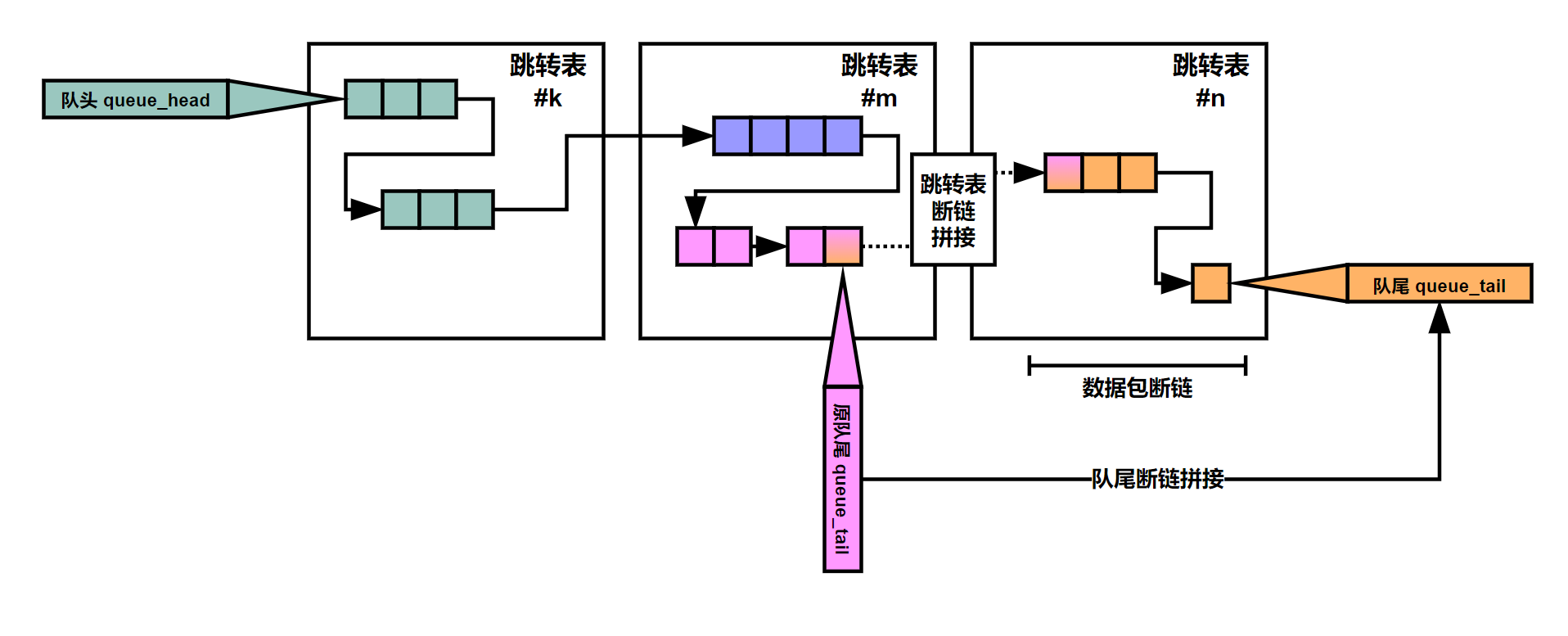
从另一个角度看，单向链表与跳转表方案的区别在于维护队列的主体，前者是端口，后者是SRAM，由于SRAM的空间固定，因此跳转表没有在极端情况下出现极高占用的弊端。



**图4.2(1) 跳转表维护队列示意图**

值得注意的是，即使各端口同时独立地与SRAM进行高速数据交互，并不会存在同时访问/修改同一个SRAM的跳转表的问题，这是因为在Crossbar架构下，每个SRAM同时只会与一个端口进行交互。所以跳转表具有RAM的结构，从而占用非常小的电路体积。

在多个端口同时向一个优先级队列末端插入数据时，可能会有冲突的情况，因此对于每个数据包，写入第一页时，暂时不和队列末端拼接，但其后的所有页的跳转表信息正常更新，在数据包最后一页写完之后，数据包已经在跳转表中呈现为一条**“断链”**，只需将“断链”与队列末端拼接即可，原来队尾对应的页地址的跳转信息更新为“断链”首页地址，Hydra还设计了对该入队、拼接过程加速的方案，具体请见4.7。



**图4.2(2) 新数据包写入时入队&拼接的过程**

利用跳转表，Hydra可以在不增加读入、写出页地址所需时间复杂度的情况下，解放队列的数据量限制，大大压缩地址管理所占用的资源。当前Hydra中跳转表占用资源为32\*2048\*16=32\*32Kbit，合32块BRAM。

## 4.3为数据包匹配较优SRAM，缓解读出冲突

一个端口写入数据到SRAM时，一般有多个SRAM可用，若随机写入其中一个，随着存入缓存的数据量增多，会导致一个SRAM里会有去往多个端口的数据包，若这些端口同时发起读出请求，在访问该SRAM时会发生冲突，一些端口的请求需要等待别的端口请求完数据才被受理，从建立请求到真正获得数据，读出延迟极高（少则几百个周期多则几千个周期）。

动态分配在该方面的缺陷无法避免，端口占用资源的不平衡总会使一个SRAM里有不同端口的数据。即使如此，我们仍能缓解过高的读出延迟，将数据包开始存入SRAM前，Hydra会为其**匹配一个较优的SRAM**，具体的匹配规则如下：

**硬性要求**（不满足该规则的SRAM将会被忽略）

1、SRAM容量充足：Hydra规定一个数据包不得拆散在不同SRAM中，即SRAM剩余空闲空间必须大于等于新写入数据包的长度。

2、SRAM未被其他端口锁定，即SRAM未正在被写且未被其他端口搜索过程匹配，前者是因为伪双口SRAM无法同时进行多次写操作，后者是因为若无此机制，不同端口可能同时认为一个SRAM“最优”，导致写冲突，需要复杂的仲裁逻辑修正。

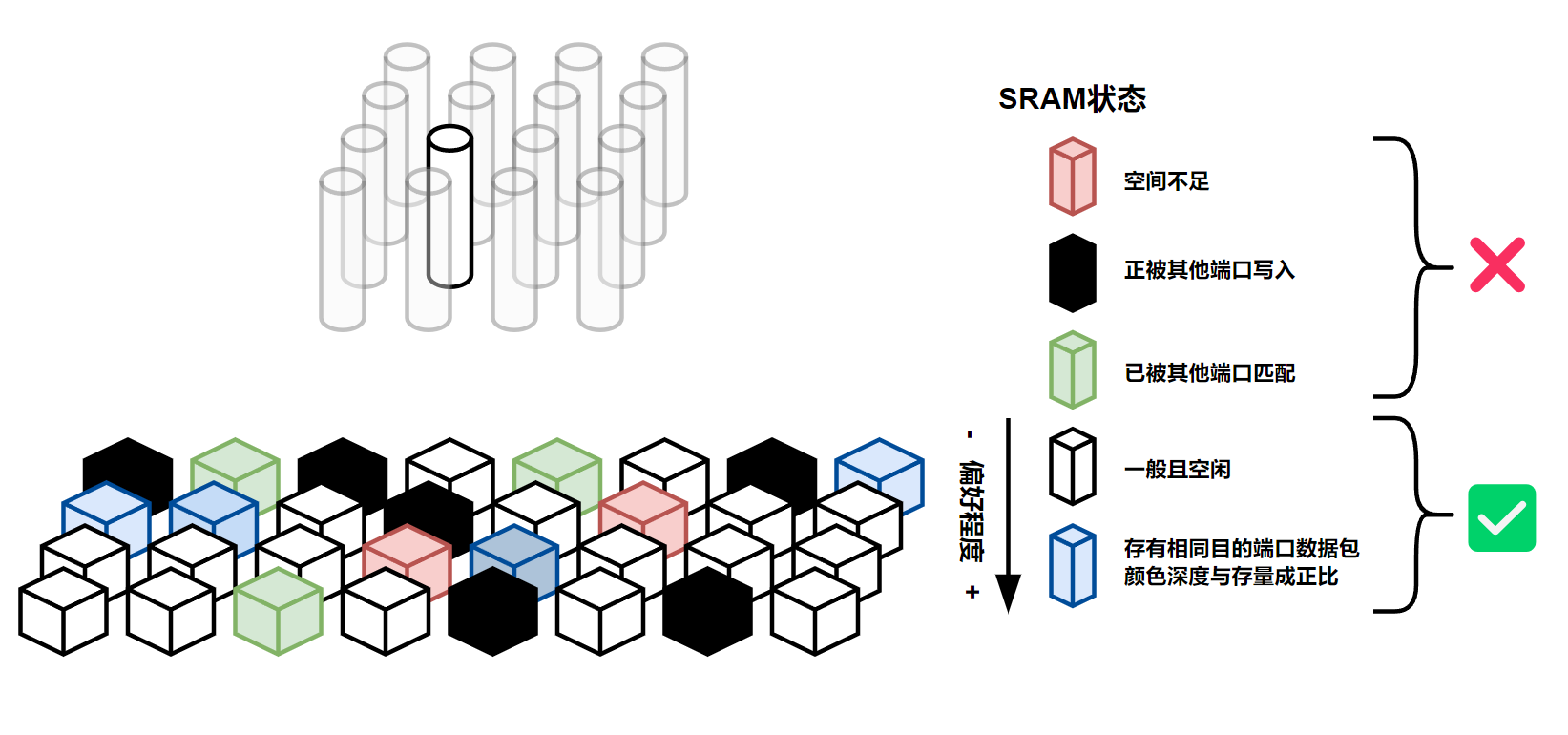
**软性要求**（越满足该规则的SRAM越被认为合适）

1、包含数据包目的端口的其他数据包较多；

2、包含的数据包对应目的端口的总数较少。

软性要求的目的在于尽可能把从同一端口读出的数据聚集在一起，降低一块SRAM极多端口数据混杂的可能性，从而缓解读取冲突。

下图是一个端口匹配的例子，红、黑、绿色SRAM为不符合硬性要求的SRAM，匹配时会直接跳过，白、蓝色SRAM为可匹配的SRAM，根据上述软性规则，该端口对蓝色最深的SRAM偏好程度最高。



**图4.3(1) 端口匹配SRAM时，SRAM状态影响着匹配结果**

端口匹配SRAM的过程中，每个周期轮流询问一个SRAM，并建立一个中间寄存器保存已经匹配到的最优的SRAM。每次询问时，若硬性要求满足，则与当前已经匹配到的最优的SRAM对比软性要求，若新匹配的SRAM更优，则其成为新的最优的SRAM。匹配完成后，可以得到较优的SRAM，接着再启动SRAM的写入，具体机制见4.4。

在保证较高的匹配效率的同时，为了维护良好的时序性，Hydra采用**错位轮询匹配**的方式，经过巧妙的设计，Hydra支持所有端口**无需仲裁地同时匹配SRAM**，通过设置错位偏移量为2，为数据选择器设置缓冲，降低匹配的组合逻辑。如下列轮询匹配中端口0、1轮流匹配SRAM#3的情况，可提前一周期缓存好匹配所需的SRAM#3有关的信号（MUX周期），简化比较时（COM周期）的组合逻辑。



**图4.3(2) 全动态分配模式下错位轮询匹配的一个回合**

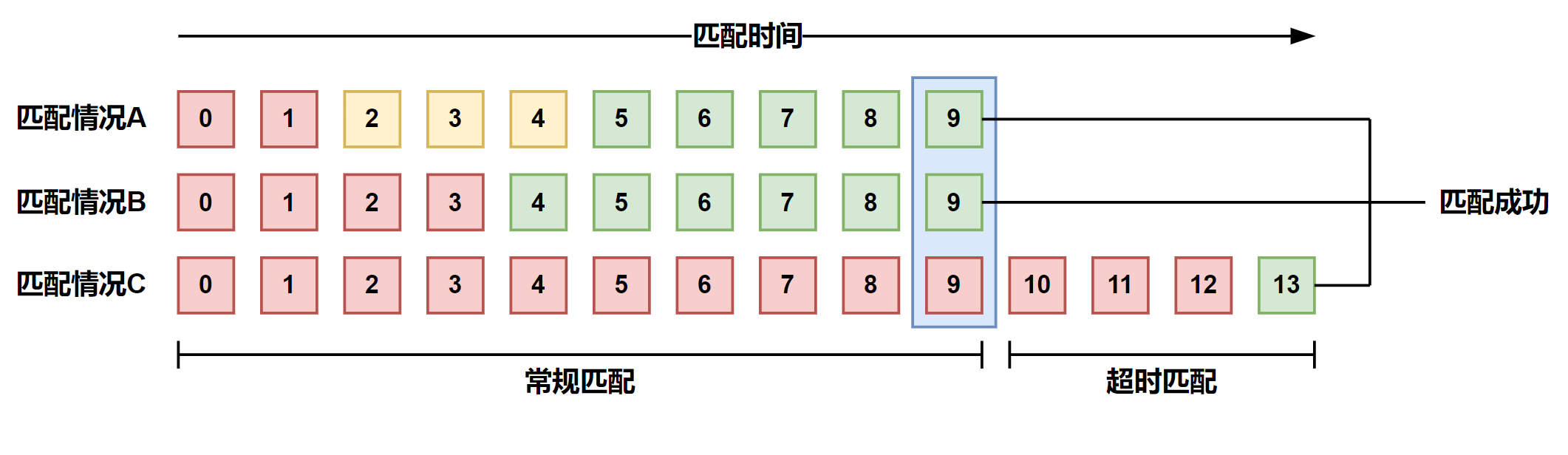


**图4.3(3) 半动态分配模式下错位轮询匹配的一个回合**

同时Hydra支持用户**自定义匹配算法的激进程度**，即可以自定义数据包匹配SRAM过程的时间长度，称为**匹配阈值**。端口匹配SRAM过程中，每周期比较本次匹配持续时长和阈值，若时长达到阈值，且已有匹配结果，则完成匹配；如果无匹配结果，则进入超时匹配阶段，之后一旦有结果则立即完成匹配。

简单推理可知，匹配阈值较低时，平均匹配用时较少，写入延迟降低，但是可能无法得到最优的SRAM，导致对读取冲突的缓解程度下降；匹配阈值较高时，平均匹配用时较长，写入延迟升高，但是可以尝试匹配更多的SRAM，使得匹配结果质量上升，对读取冲突的缓解程度上升。从激进程度的角度考虑，**低阈值对应激进策略，高阈值对应保守策略**。

下图是匹配阈值等于9时可能出现的匹配情况，红色代表没有匹配结果，黄色、绿色代表有结果，绿色为最终匹配结果。情况A中，端口在第2周期先匹配到了某个SRAM，后来在第5周期的匹配中，经过比对找到了更优的结果，最终在匹配时间达到阈值时结束匹配，发送匹配成功信号；情况B与A类似；情况C在达到阈值时仍未找到可用的SRAM，进入超时匹配阶段，在第13周期匹配到了可用的SRAM，同时立即结束匹配，触发匹配成功信号。



**图4.3(4) 匹配阈值等于9时可能出现的匹配情况举例**

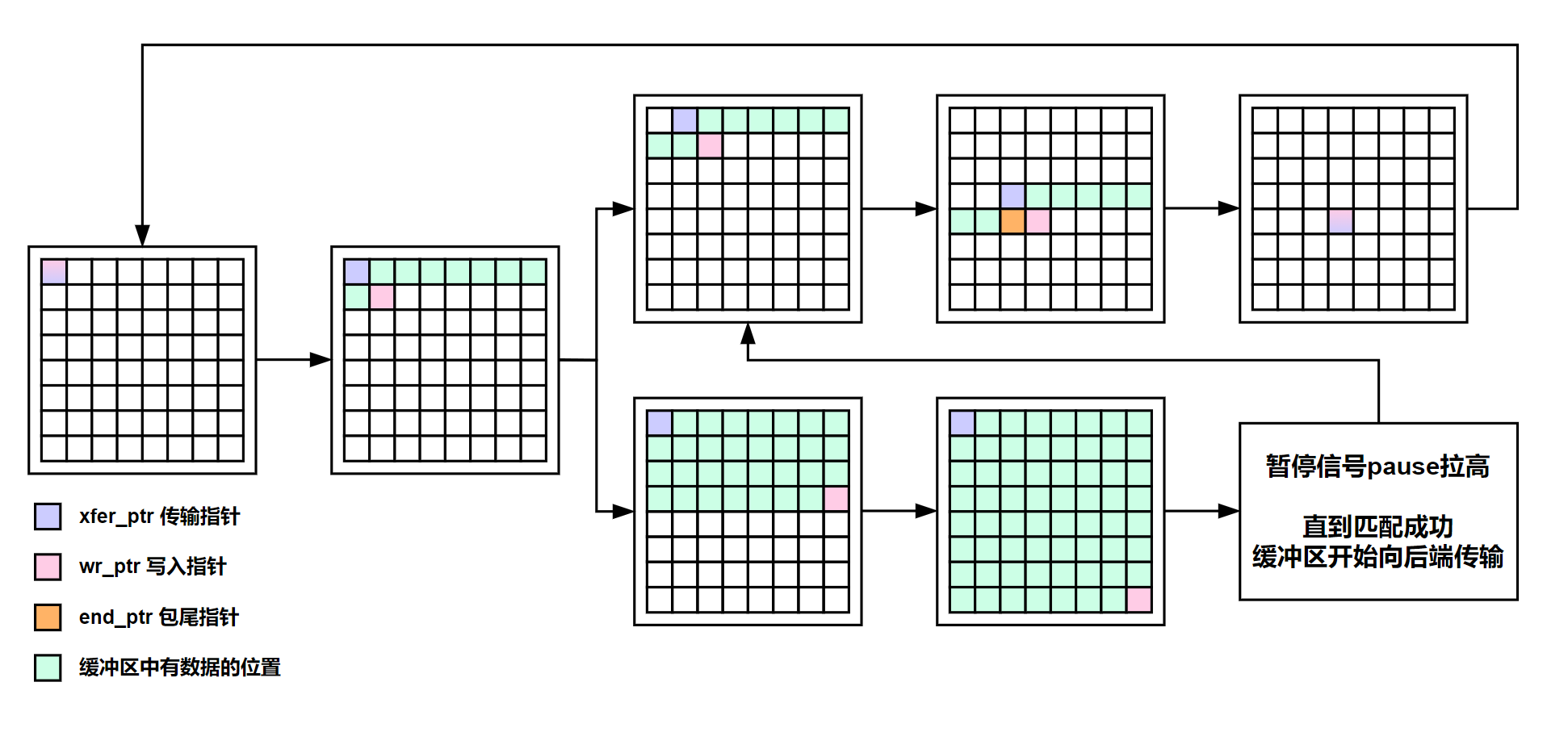
上述匹配机制仅需消耗数个周期即可缓解读取冲突问题，降低数百、数千周期的读取延迟，间接地大幅提升读取带宽。对于该机制造成的写入延迟（即匹配消耗的时间，一般是匹配阈值），Hydra设置了基于尾部预测的快速拼接机制**抵消**了它，具体请见4.5。

## 4.4设置写入前端缓冲结构，支持断点续传

由于匹配机制，数据包进入Hydra一段时间后才能真正开始被写入SRAM，此时需要一个缓冲区用于存放匹配未完成时，已经进入模块的数据。

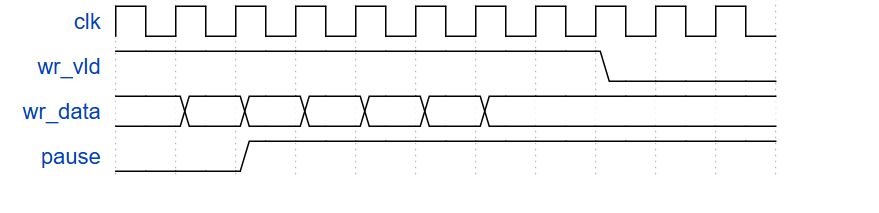
Hydra将该缓冲区独立成模块，称为**“写入前端”**，前端与外界交互，并将数据暂时存放于缓冲区；通过建立与匹配模块的交互，在匹配结束的时启动向主模块（“后端”）的数据传输。缓冲区采用FIFO的结构，Hydra使用三个指针维护缓冲区中数据的读写：xfer\_ptr和wr\_ptr为FIFO的头部和尾部，end\_ptr则指向数据包结尾半字在FIFO中的位置，方便传输时切割不同的数据包。

下图是三个指针维护缓冲区的具体机制，数据包进入模块时，数据写入缓冲区，即图中绿色的位置，紫色xfer\_ptr指向队头，粉色wr\_ptr指向队尾，若匹配完成，启动前端向后端的传输过程，则对应上面一行三张图的情况，此时数据包前面已经写入缓冲区的部分随着xfer\_ptr的移动被逐个传输至后端，后续还未写入部分同时有条不紊地进入缓冲区。数据包写入缓冲区完成后，将会留下橙色end\_ptr，当xfer\_ptr与end\_ptr重合时，前端将会发出数据包终止信号，并回到初始态。



**图4.4(1) 三指针维护写入前端缓冲区的机制**

当匹配超时较长时间，向后端的数据传输一直未能启动，数据将会在缓冲区中堆积，见图中第二行两张图的情况，缓冲区快满时，会对外界发送写入暂停的信号，即pause信号。值得一提的是，pause信号会在完全暂停前若干拍提前拉高（见下方波形图），便于外界设备反应后及时暂停，且不用将已发送的数据冲刷/收回，可见Hydra具有不凡的用户友好性。pause信号的设计使得Hydra支持**断点传输**：外界设备等待匹配完成，前端进入正常传输流程，缓冲区腾出空闲空间，pause信号拉低后仍然可以继续数据包的传输，无需重新发送整个包。



**图4.4(2) pause信号提前拉高保证外接设备响应时数据无需冲刷/收回**

## 4.5基于尾部预测的快速拼接机制，支持单包边读边写

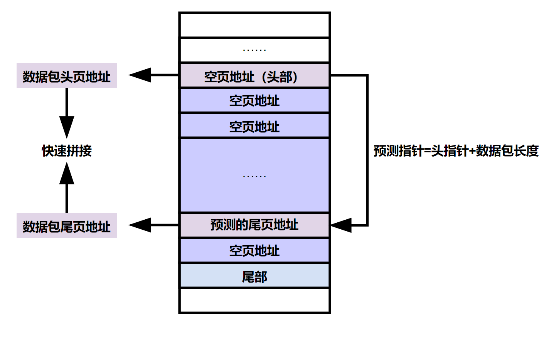
由于匹配机制，使得数据包从外界写入到前端缓冲区，到传输至后端，真正写入SRAM，再到入队、跳转表拼接过程完成后，变为可读取的状态之间存在若干周期的延迟，这使得一个数据包写入Hydra需要等待较长时间才能被读出。

详细分析上述过程每一步的延迟，可知从数据包最初写入到可以读出的时刻之间相距的总延时Delay=传输延迟XD+匹配阈值T+数据包长度L+入队拼接C个周期，为了缓解该延迟，Hydra包含了**基于尾部预测的快速拼接机制**，可以**将Delay降低60%~98%**。

以火车进入隧道比喻上述过程，数据包长度L为火车长度，传输延迟XD+匹配阈值T+入队拼接C为隧道长度，我们要得知火车的信息，实际上并不需要等待火车完全通过隧道，而只需要火车刚从隧道的另一边冒出车头时，查看车头的型号即可知道火车的信息。数据包也是一样，在数据包第一半字（控制部分）被传输至后端时，就可得知数据包的长度、优先级、目标端口。



**图4.5(1) 火车进入隧道比喻Hydra如何降低传统数据包调度的读写延迟**

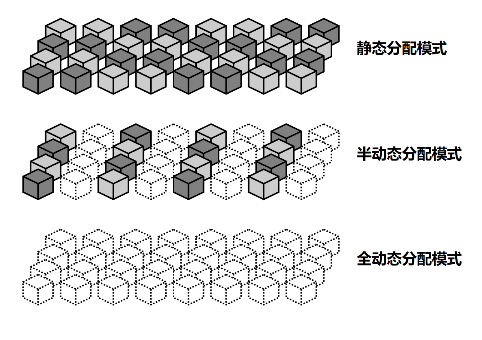
又因为Hydra采用空闲队列回收与再分配内存（见3.1.2），因此只需在数据包第一页写入时，将其在空闲队列中的地址与数据包长度相加即可得到数据包最后一页在空闲队列中的地址，以此提前预测数据包传输完毕后尾页地址，如右图。有已知的头和预测的尾即可完成数据包的入队、拼接操作，使其在完全写入完毕前就可被提前读取。

**图4.5(2) 尾部预测的具体实现**

通过基于尾部预测的快速拼接机制，Hydra直接消除了上述写入-读出延迟中的L，而XD+T+C一般远小于L。在匹配阈值小于24周期的情况下，几乎所有数据包都可以在外部设备写入完毕之前入队、拼接，即外部设备**写入完成后可立即读出**。Hydra**不仅抵消了匹配SRAM过程额外的时间开销**，**还使得数据包的读写更加灵活**。

## 4.6多匹配模式支持，可设置全动态、半动态、静态模式

Hydra不仅支持用户自定义匹配策略的激进程度，还支持用户定义匹配模式。目前有三种模式可用：全动态、半动态、全静态模式。

**静态模式**：16个端口分别与2块SRAM绑定，不共享任何SRAM，适用于中低速数据交换，读取延迟最低；

**半动态模式**：16个端口分别与1块SRAM绑定，并共享剩余的16块SRAM，适用于极端情况较少，需读取延迟较低的情况；

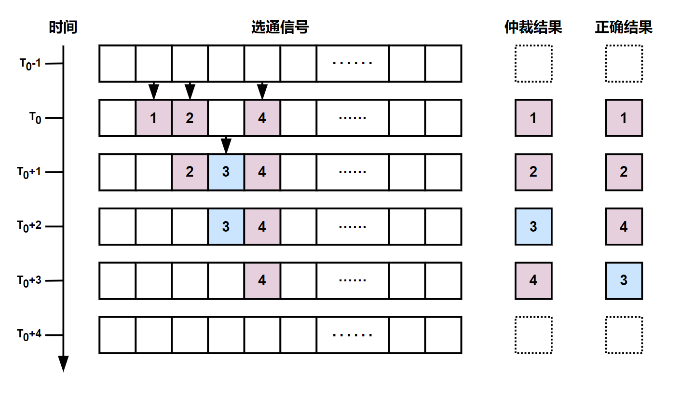
**全动态模式**：16个端口完全共享32块SRAM，适用于可能会出现极端情况的高速数据交换。

**图4.6 三种模式下缓存分配情况**

值得注意的是，无论是哪种模式，Hydra设计下的轮询匹配保证每个SRAM不会连续两个的周期被匹配，这使得每个端口可以提前从SRAM处抓取匹配过程需要的信号，降低匹配周期组合逻辑的复杂性。

## 4.7保障保序性，消除传统仲裁的缺陷

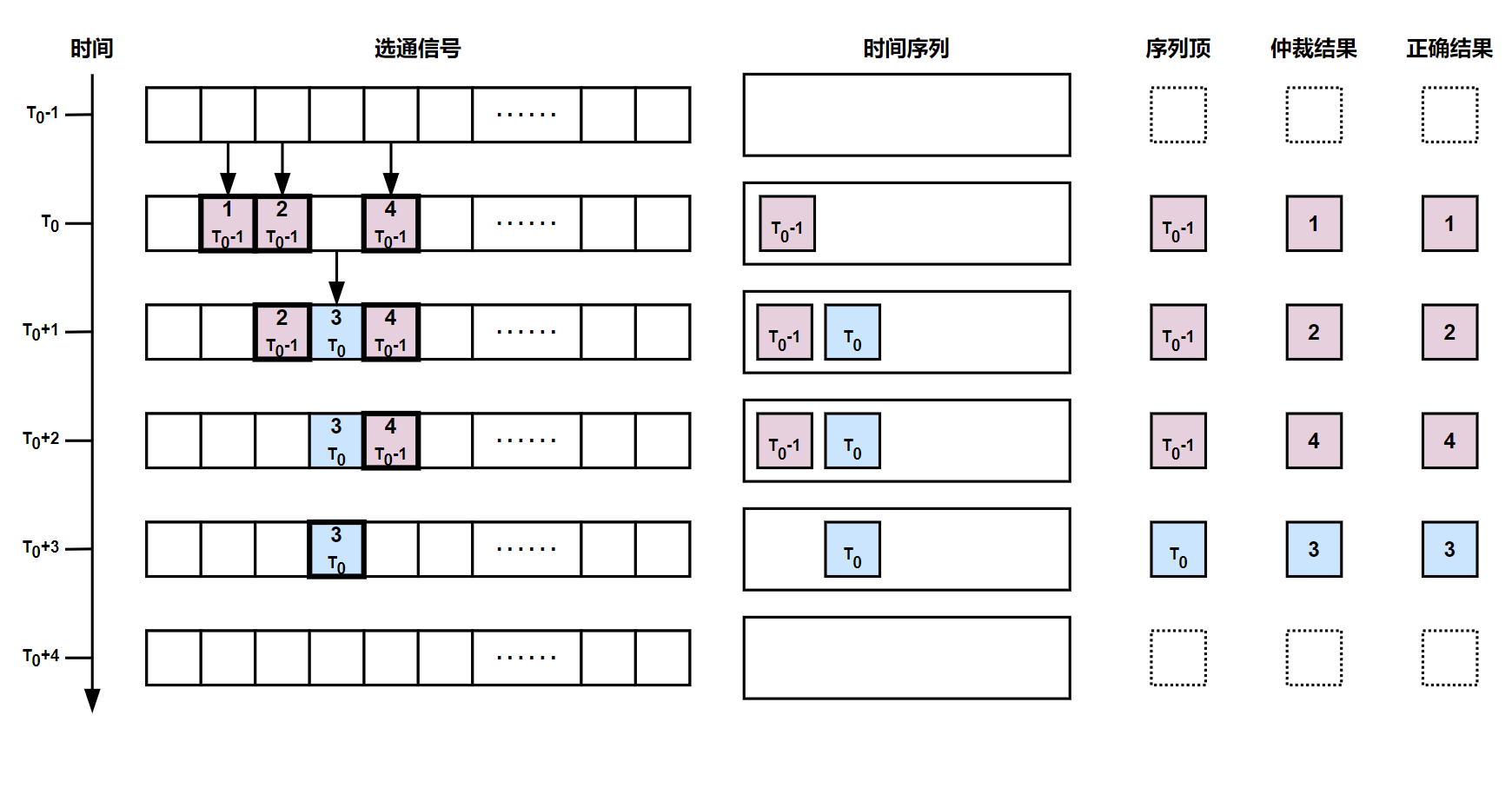
当多个端口同时写入相同目的端口相同队列的数据包时，会同时向发出多个竞争的入队请求，此时需要仲裁模块逐个处理。传统仲裁具有一定缺陷，即无法判别请求的先后性。

举一个简单的例子，如右图，在某一周期端口1、2、4分别写入了目的端口9优先级0的数据包，第下一周期端口3写入了目的端口、优先级相同的数据包，假如第二周期处理完了端口1的请求，则此后仲裁选通信号中2、3、4位为高，接下来三周期依次处理2、3、4的请求。这时仲裁就出现了错误：本应后入队的来自端口3的数据包比端口4的数据包先被处理了，导致读出顺序错乱。

**图4.7(1) 传统仲裁的失序错误**

为了维护数据包的有序性（称为**“保序”**），Hydra引入了**时间序列**，一个存储时间戳的FIFO结构。在被写入数据包的SRAM向目的端口发起入队请求时，请求中会附带发起时间戳，主模块在有请求时将时间戳加入时间序列，每一周期根据时间序列中的顶部时间戳，生成对应的选通掩码，与选通信号经过简单逻辑仲裁得到应该处理的入队请求编号。

仍然使用刚刚的例子，第一周期由于存在请求，时间戳T0-1被送入时间序列，第二周期同理，但在端口1、2、4的请求被处理完成前，时间序列顶一直都是T0-1，经过简单的判等和按位与操作，选通信号中端口3的请求被掩盖，从而搁置直到请求1、2、4处理完毕，时间序列弹出T0-1。

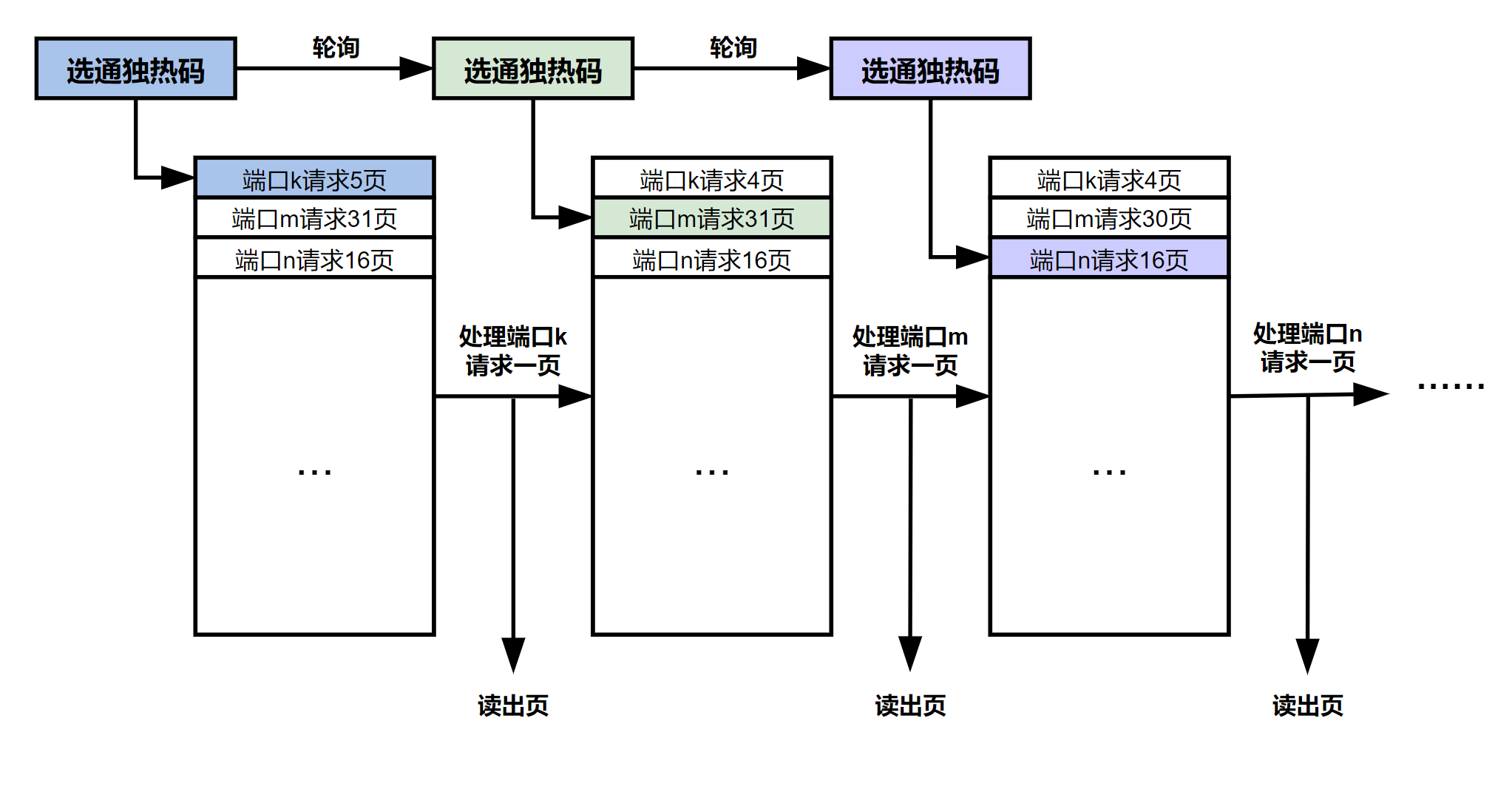


**图4.7(2) 时间序列保序机制运作原理举例**

通过引入时间戳和时间序列，Hydra保障了数据包读出的先后顺序。

## 4.8 “安抚”读取机制，降低读出延迟

虽然Hydra引入了匹配机制、基于尾部预测的快速拼接机制，但是无法完全避免读取冲突。当读取冲突发生时，传统的方案是仲裁阻塞读取，即按某种顺序处理读出请求，后续处理的请求必须等待先处理的请求完全结束之后才能开始。这会导致读取冲突时，被仲裁靠后读出的数据包，其读出延迟取决于前面读出数据包的长度，从请求刚建立（ready拉高）到真正读出（rd\_vld）延后了数百甚至上千周期。



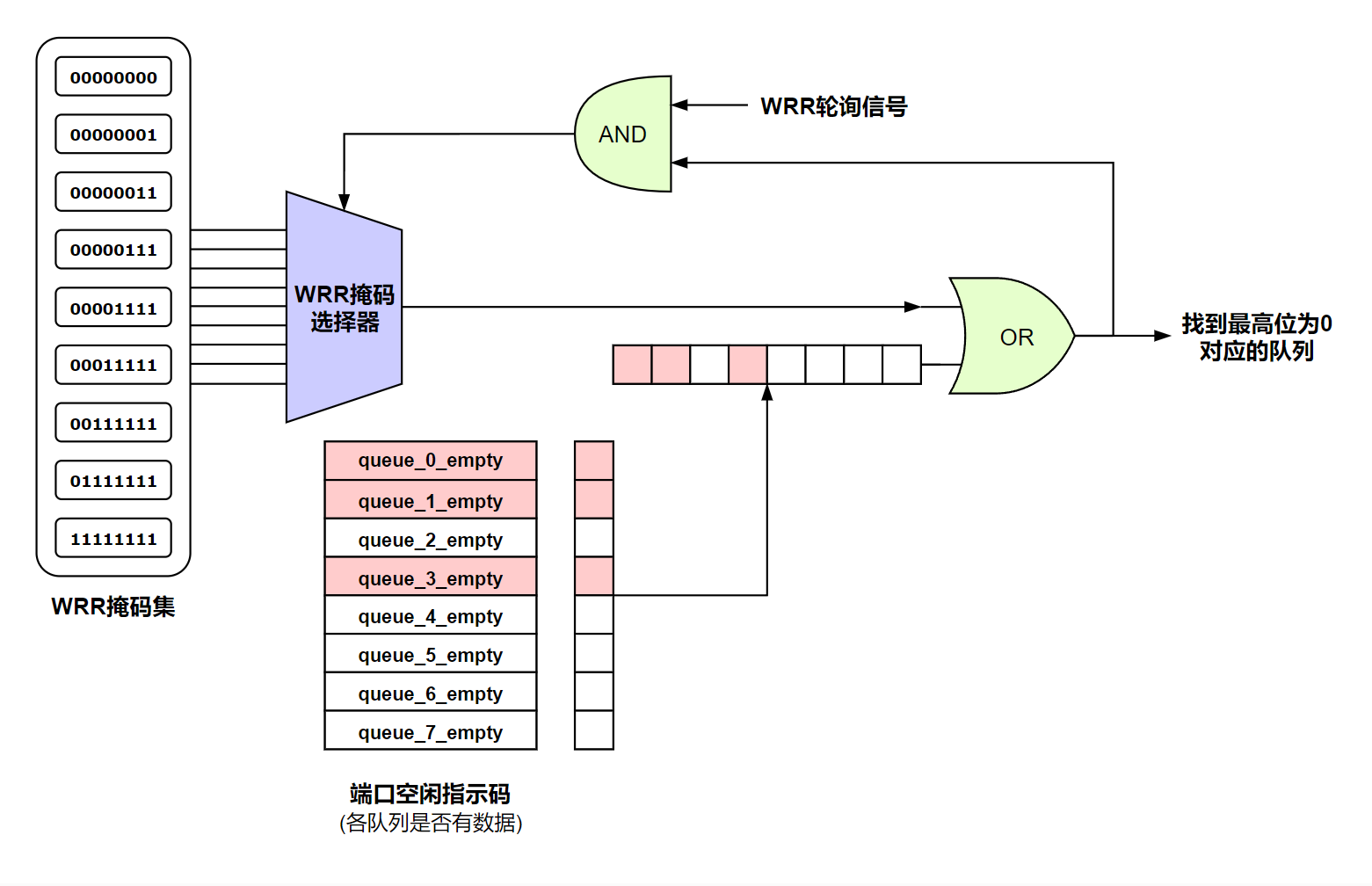
**图4.8 “安抚”读取机制——轮流满足读出冲突的k、m、n端口**

为了缓解这种情况，Hydra会**轮询处理读取请求**，即出现读取冲突时，每个请求**轮流处理一页**。这样每个端口将平均地享有读出带宽，从请求刚建立（ready拉高）到真正读出（rd\_vld）的延迟一般不多于32周期，最多仅有128周期（最极端情况，实际上在匹配策略下不可能出现）。由于该轮询机制类似于在“安抚”冲突时“不耐烦”的请求，故称之为**“安抚”读取机制**。

## 4.9基于掩码集的无需复杂计算的快速WRR

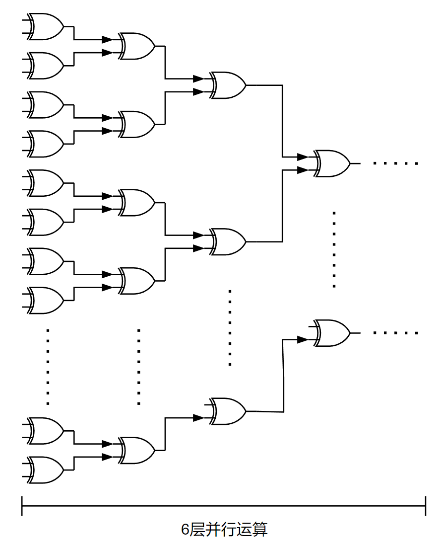
WRR机制常常伴随着移位运算出现，为了避免体积庞大的移位电路，Hydra预生成了所有可能出现的WRR掩码情况（称为WRR掩码集）。在处理请求时，端口会生成WRR轮询信号，将下一次使用的WRR掩码切换到阵列中的指定位置，并与队列的非空信号进行简单的逻辑运算预生成下一次应当被读取的队列编号。实现了无需复杂计算的WRR机制。

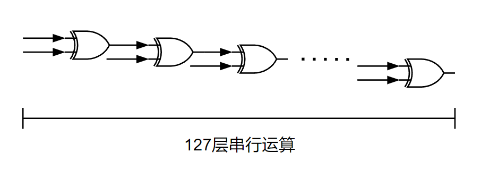
由于每次读取的队列编号都是预生成的，故Hydra能够在ready信号拉高一周期后立刻响应并发起数据包的读出请求。



**图4.9 基于掩码集的快速WRR**

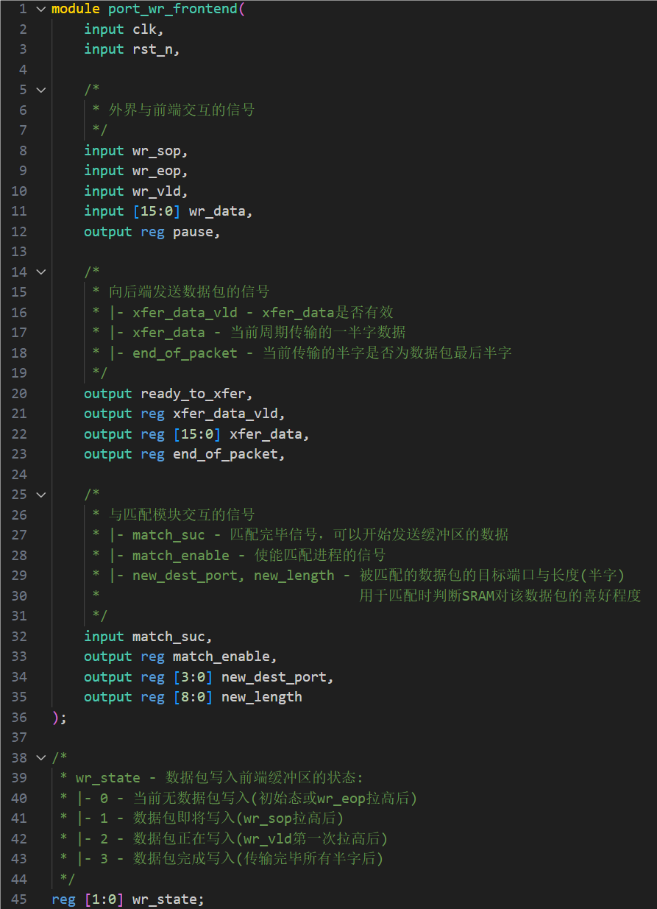
## 4.10并行汉明校验，实现单周期编码解码

Hydra根据异或操作的对称性，将串行异或运算转化为了并行异或运算，每一层执行多次异或运算并一层一层汇总，最终生成(136,128)SEC编码。编码过程在RTL电路层面仅需6层XOR门，可在单周期内完成编解码。



**图4.10 串并行汉明校验示意图**

## 4.11其他亮点

​1、代码注释齐全，变量命名规范统一，大多数操作状态机化；

​2、占用资源较大的结构全部RAM化，将被综合在BRAM中，降低布线压力；

3、机制流水线化，相邻请求处理可无缝衔接，无需等待信号重置；

4、严格制定每一步的逻辑复杂度，保障良好的时序严谨性；

5、采用feature验方式，一步一步细化进行功能调试与验证，使得Hydra具有良好的鲁棒性。

**图4.11良好的代码规范**

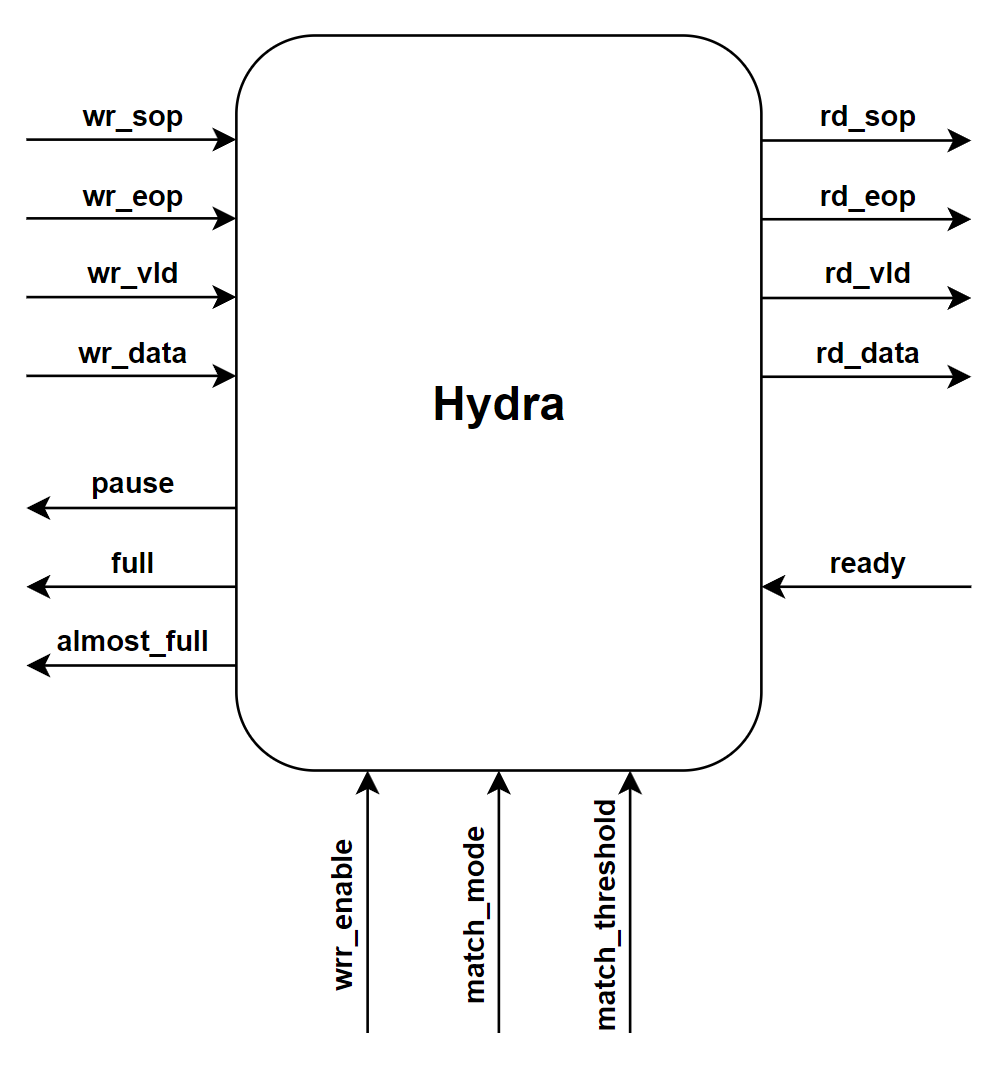
# 5.模块介绍

## 5.1各模块说明

### 5.1.1 hydra(顶层模块)

功能：主控制器模块，连接所有子模块并处理核心逻辑。



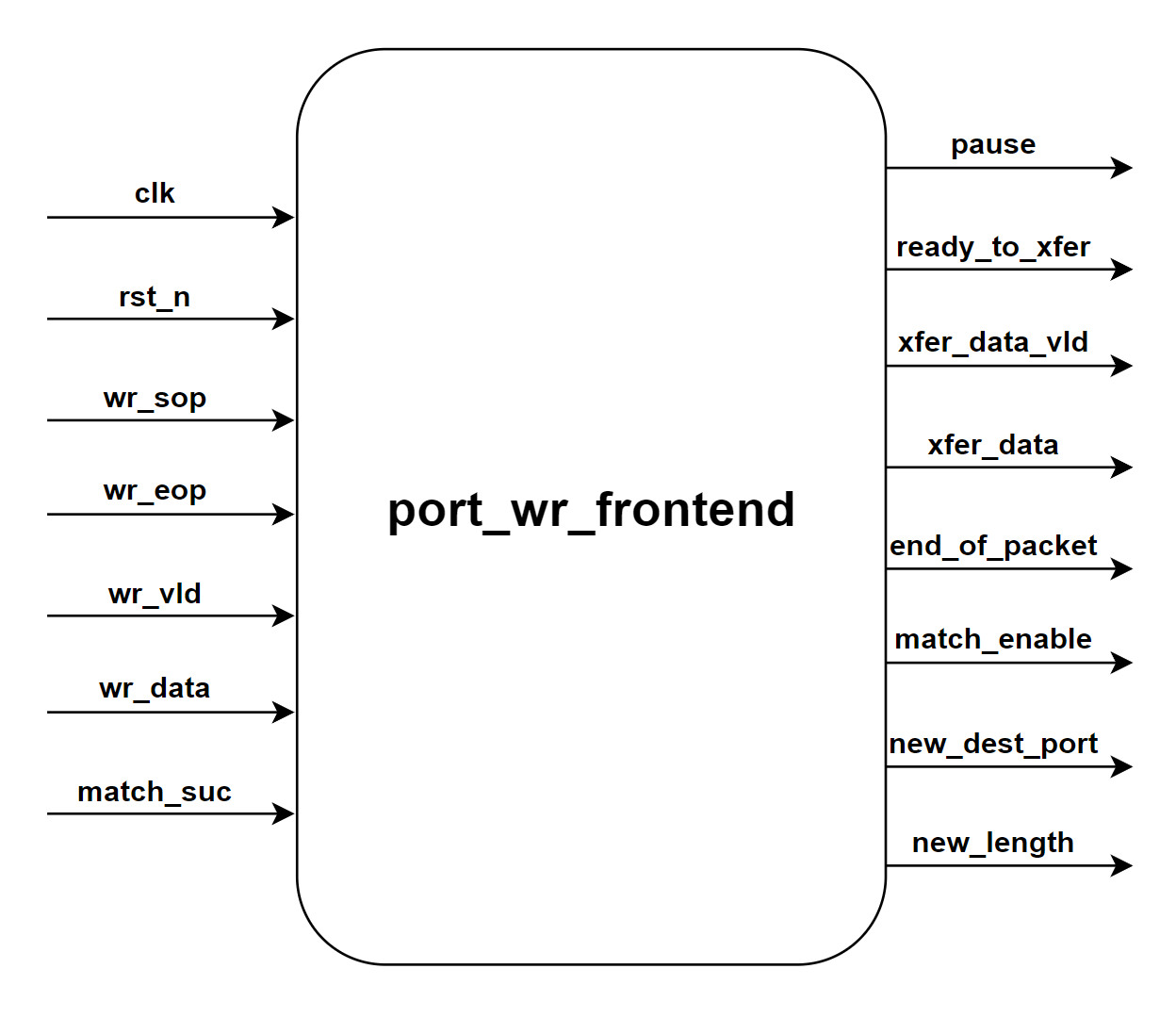


**图5.1.1 hydra模块示意图**



### 5.1.2 port\_wr\_frontend

功能：写入前端缓冲区，对刚写入Hydra的数据进行缓存。



**图5.1.2**

**port\_wr\_frontend模块示意图**

### 5.1.3 port\_wr\_sram\_matcher

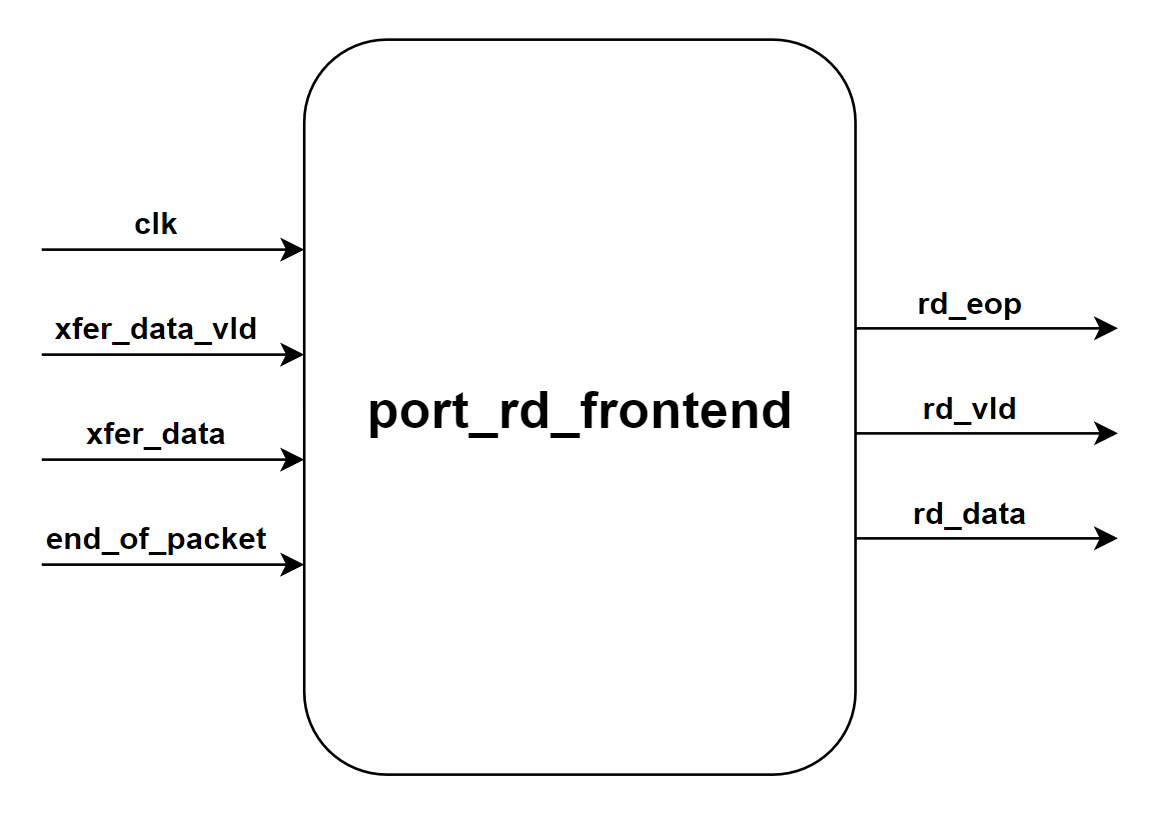
功能：匹配机制核心，与写入前端交互并将最优结果发送至后端



**图5.1.3**

**port\_wr\_sram\_matcher模块示意图**

### 5.1.4 port\_rd\_frontend

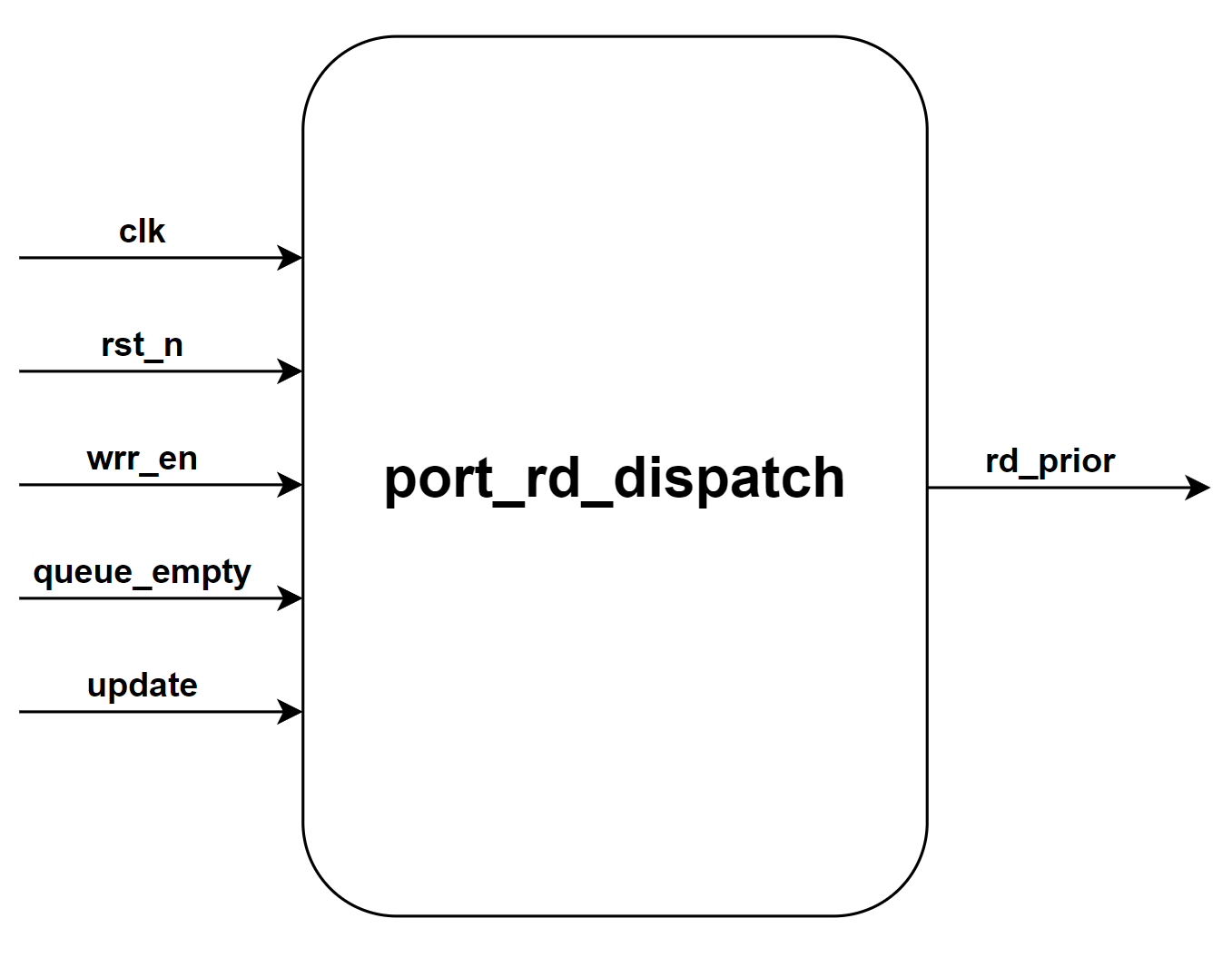
功能：将纠错完毕的数据输出模块。



**图5.1.4**

**port\_rd\_frontend模块示意图**

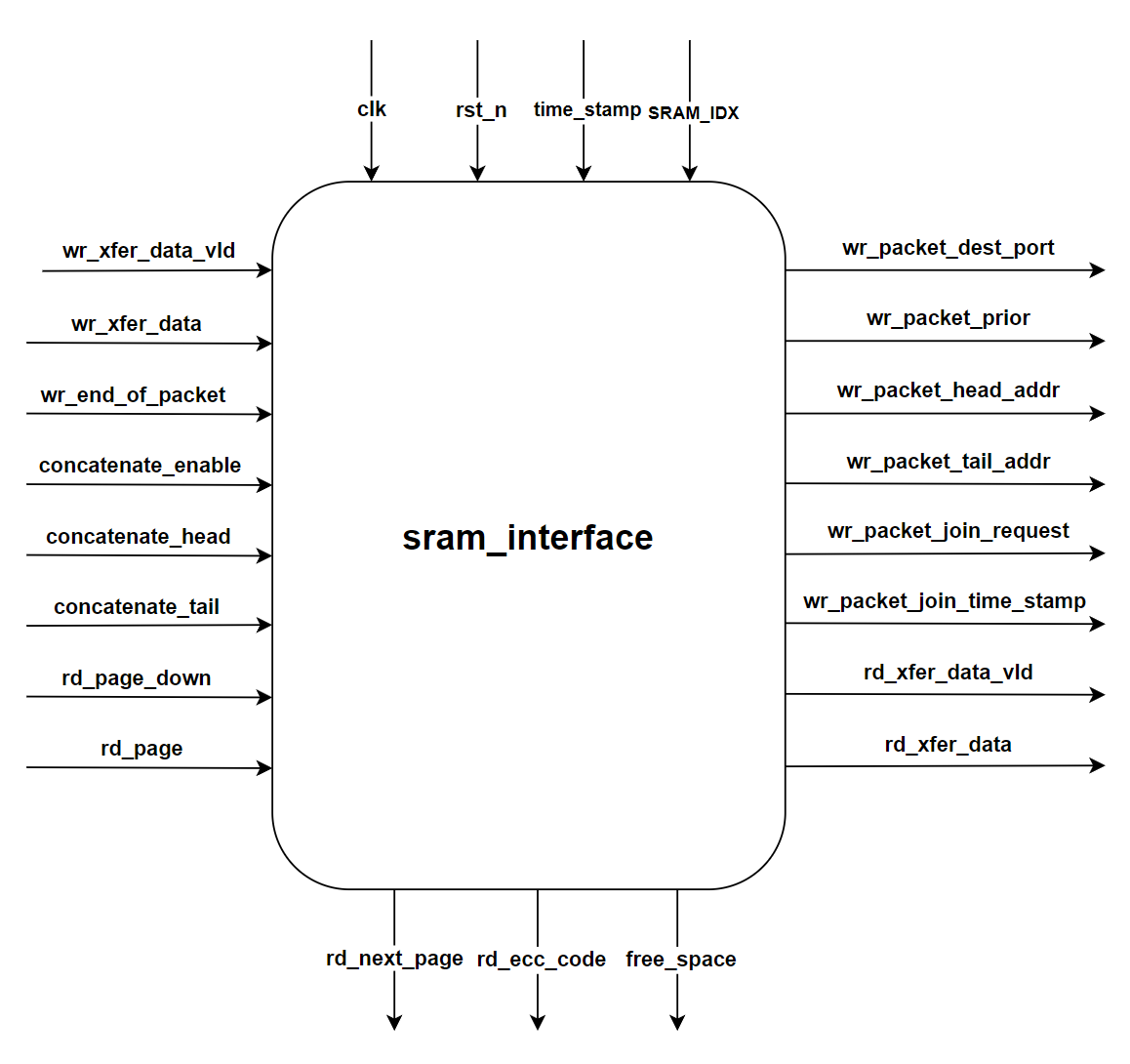
### 5.1.5 port\_rd\_dispatch

功能：端口读出数据包时生成队列调度顺序。

**图5.1.5**

**port\_rd\_dispatch模块示意图**

### 5.1.6 sram\_interface

功能：SRAM封装模块，直接与端口交互，管理ECC存储、跳转表和空闲队列。

**图5.1.6**

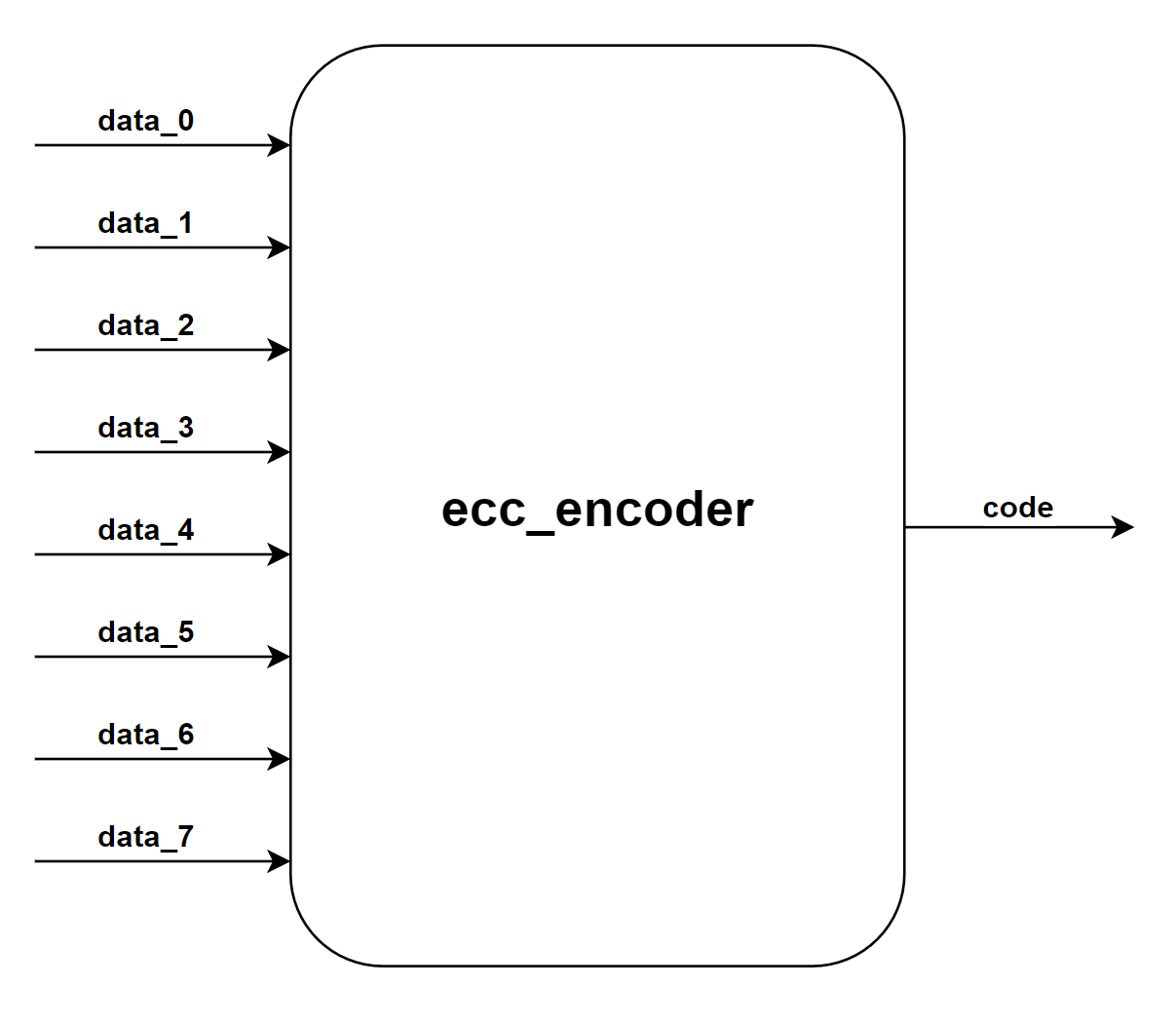
**sram\_interface模块示意图**



### 5.1.7 sram

功能：伪双口SRAM，提供独立工作的读写口各一个。

**图5.1.7 sram模块示意图**

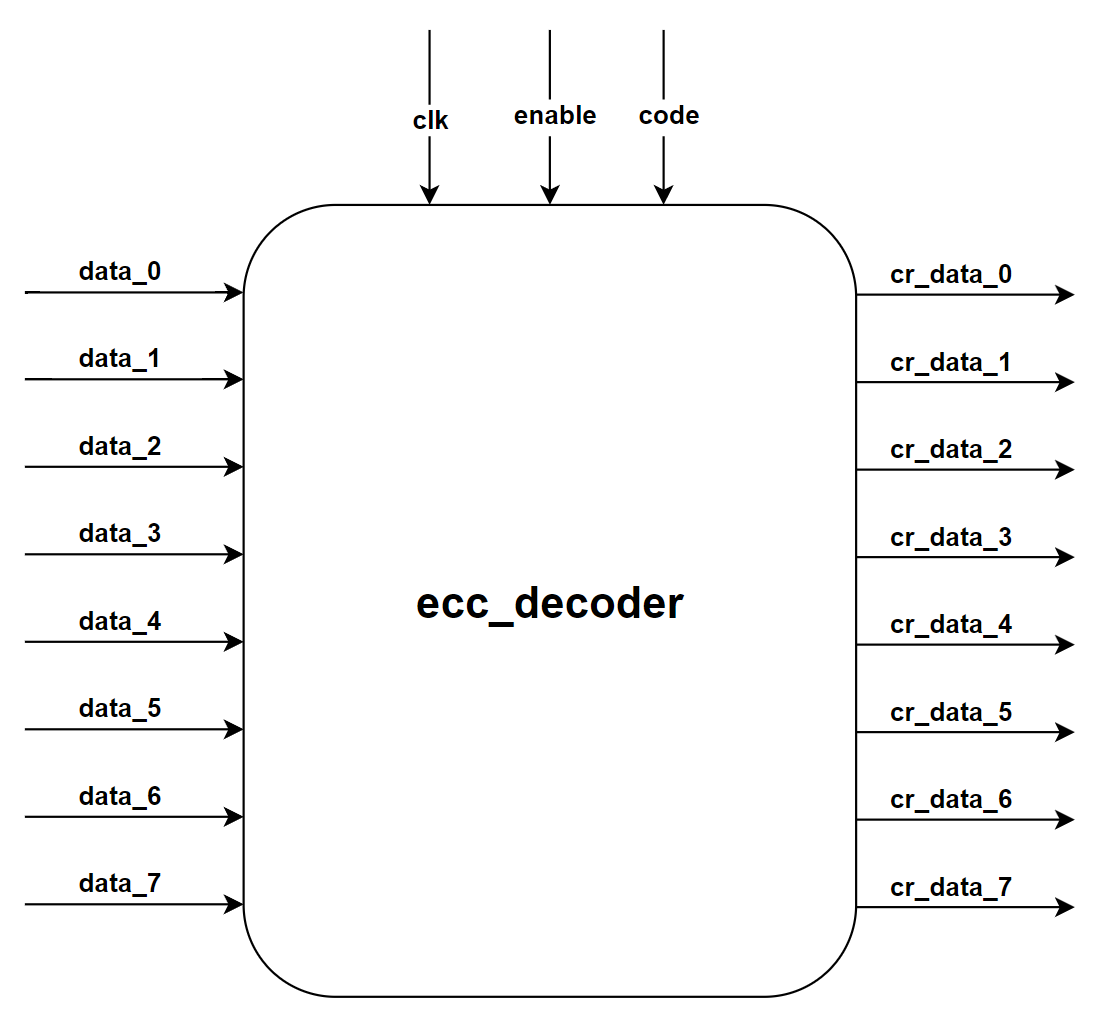


### 5.1.8 ecc\_encoder

功能：根据128位数据生成8位校验码。

**图5.1.8 ecc\_encoder模块示意图**

### 5.1.9 ecc\_decoder

功能：根据8位校验码对一页待纠错数据纠错。

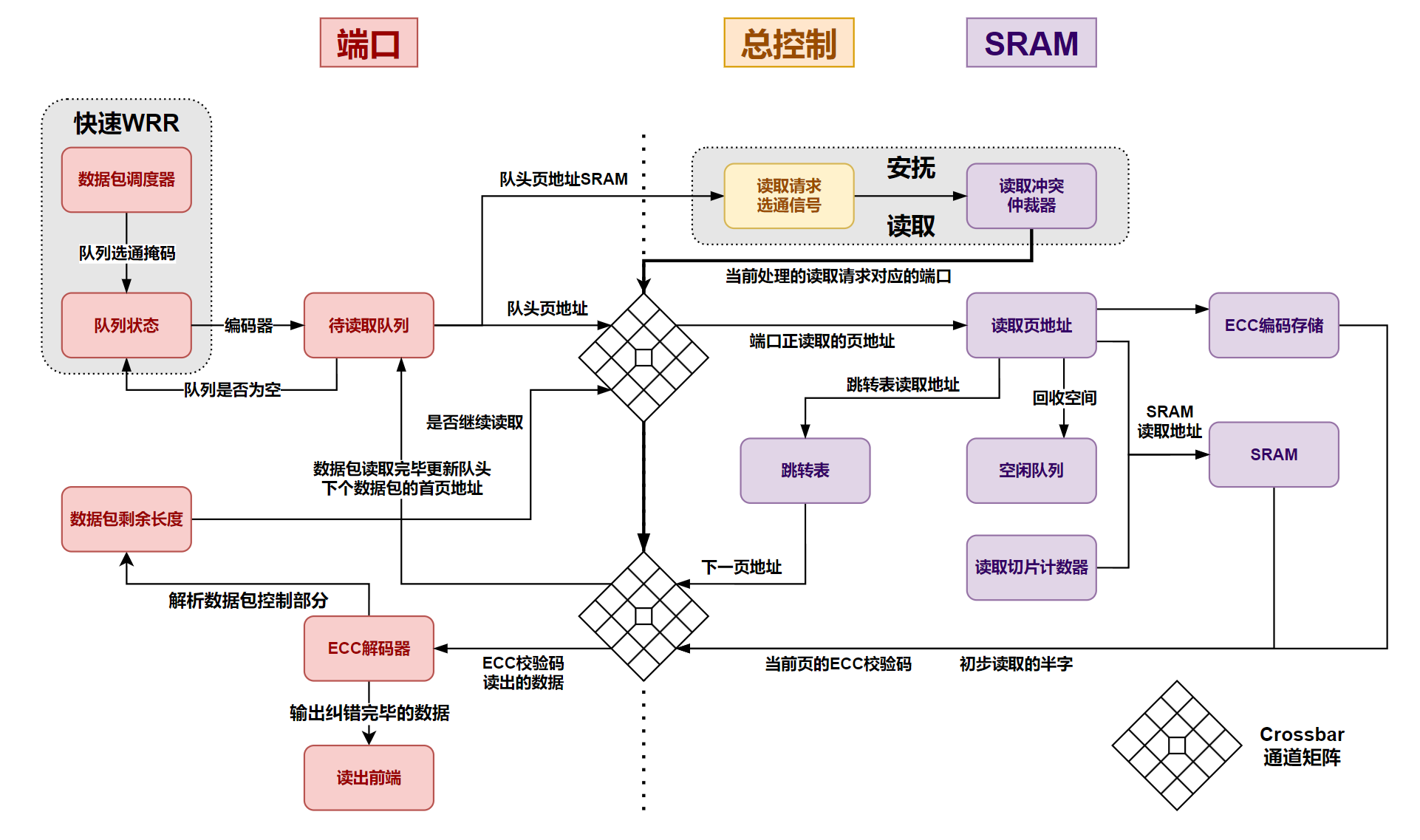
**图5.1.9 ecc\_decoder模块示意图**

## 5.2重要逻辑说明

### 5.2.1数据包写入流程

**5.2.2数据包读出流程**

**图5.2.1 数据包写入流程核心部分示意图**



**图5.2.2 数据包读出流程核心部分示意图**

# 6.接口与配置

## 6.1.写控制IO口

### 6.1.1 IO口介绍

**wr\_sop**：拉高1周期表示开始写入数据包。

**wr\_eop**：拉高1周期表示结束写入数据包。

**wr\_vld**：表示当前通过wr\_data写入的一半字数据有效。

**wr\_data**：写入一半字数据的内容，仅在wr\_vld为高的时候被认为有效。

### 6.1.2 使用方法

**数据包格式**：

控制部分（1半字）和数据部分（31~511半字）。

其中控制部分为16位，描述三个数据包信息：

1）高9位表示数据包的长度（半字）；

2）低4位表示目的端口；

3）剩余3位表示优先级。

**注意事项**：

1）wr\_sop和wr\_eop不能同时拉高。

2）wr\_vld只能在wr\_sop为高后、wr\_eop为高前的一段时间内为高。

## 6.2写反馈IO口

### 6.2.1 IO口介绍

**full**：当所有SRAM被占用/满时拉高。

**almost\_full**：当所有SRAM均满足以下至少一个条件时拉高：

1. 被占用/写满；
2. 剩余空间少于25%。

**pause**：当前端口缓冲区被占满或上个包已写入缓冲区完毕，但仍处于超时匹配时拉高。

### 6.2.2 使用方法

当某个端口的pause拉高时，立即停止数据传输，直到pause拉低即可继续传输。

当almost\_full拉高时，可以考虑拉高ready读出缓存中的数据。

## 6.3读控制IO口

### 6.3.1 IO口介绍

**ready**：请求读出一个数据包的信号。

### 6.3.2 使用方法

ready拉高时，若该端口存在可读取的数据包，rd\_sop将会在下一周期立即拉高，表示开始读出数据包。读出数据包的过程中拉高ready无效，即只有在前一个数据包rd\_eop拉高后，拉高ready才会触发后一个数据包的读出。

## 6.4读反馈IO口

### 6.4.1 IO口介绍

**rd\_sop**：拉高1周期表示开始写入数据包。

**rd\_eop**：拉高1周期表示结束写入数据包。

**rd\_vld**：表示当前通过rd\_data写入的一半字数据有效。

**rd\_data**：写入一半字数据的内容，仅在rd\_vld为高的时候被认为有效。

### 6.4.2 使用方法

数据包格式与sop、eop、vld和data信号线的机制与写入时相同。

## 6.5配置选项

**wrr\_enable**：每个端口各1个，表示是否开启该端口的WRR调度模式。

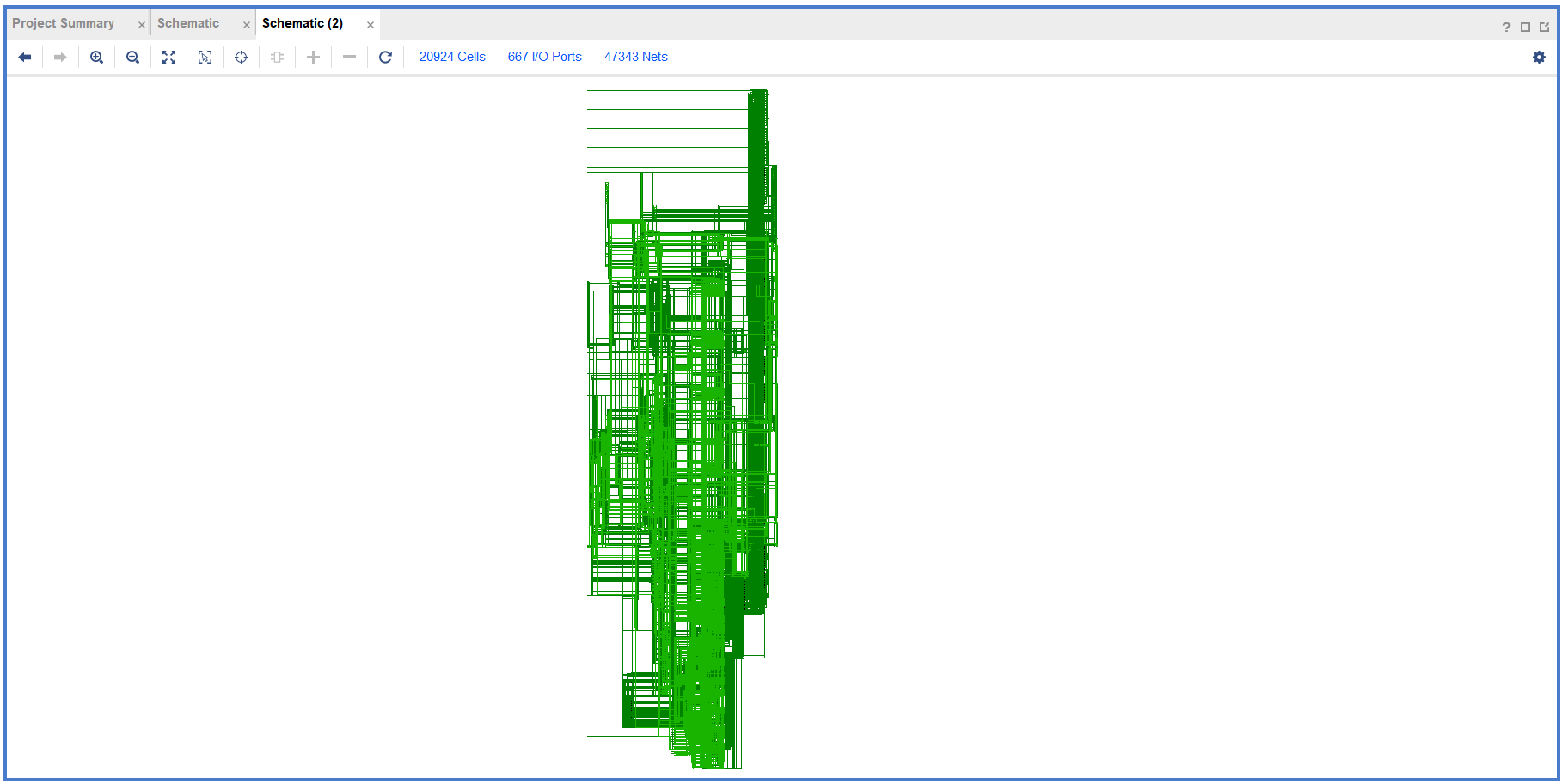
**match\_mode**：**缓存分配模式**，0-静态分配；1-半动态分配；2/3-全动态分配。当需要高速读写且有可能出现极端情况时，建议使用全动态分配模式；需要高速读写且端口使用程度较为平均时，建议使用半动态分配模式；仅需要中低速读写时，建议使用静态分配模式。

**match\_threshold**：**匹配阈值**，当匹配时长超过该值后，一旦检测到可用的SRAM即完成匹配。匹配阈值的大小涉及匹配策略的激进程度，越小匹配所需时间越短，写入延迟降低，但可能降低匹配策略对读取冲突的缓解程度；越大匹配所需时间越长，提升匹配策略对读取冲突的缓解程度，但会使写入延迟提升。静态分配模式下匹配阈值最大为0；半动态分配模式下匹配阈值最大为16；全动态分配模式下匹配阈值最大为30。

# 7.验证方法

## 7.1 RTL/行为级仿真

### 7.1.1 RTL级电路图（Vivado2018.3）

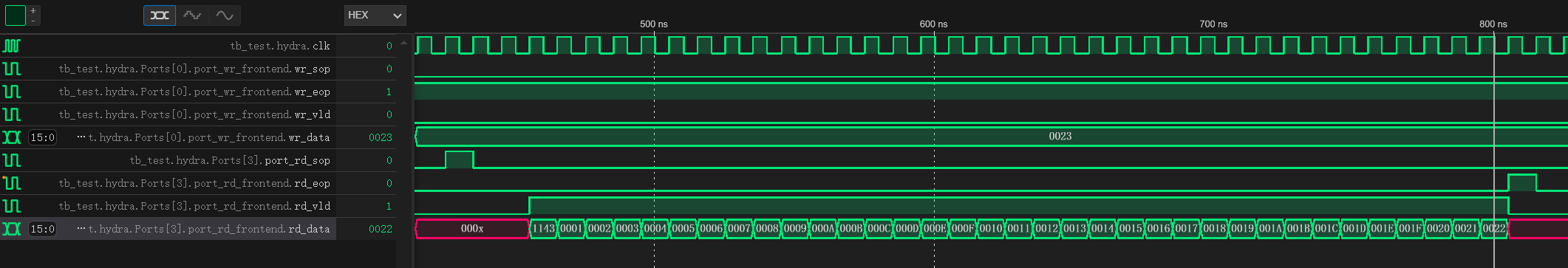


**图7.1.1 RTL级电路图**

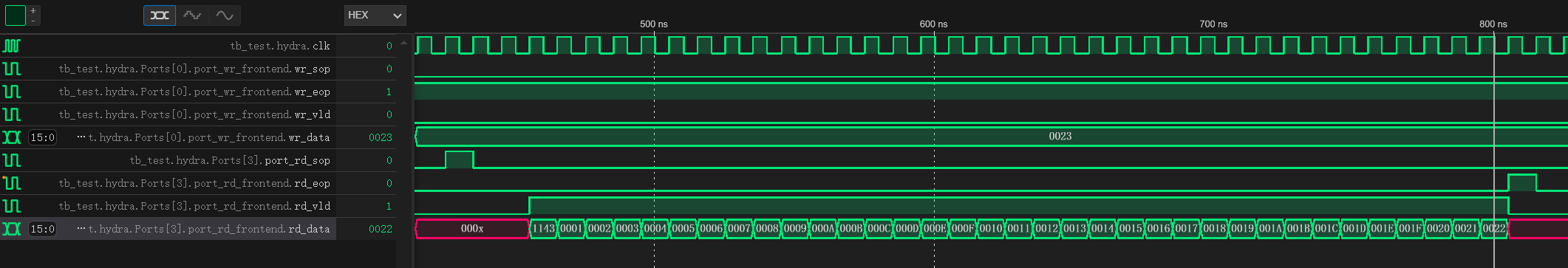
### 7.1.2 Feature验证（VS Code编译+ WaveDrom读取.tb运行生成的.vcd波形文件）

注：Hydra完成了大量Feature Test，因篇幅有限，只展示最重要测试的波形图。

1. **单数据包写入读出**



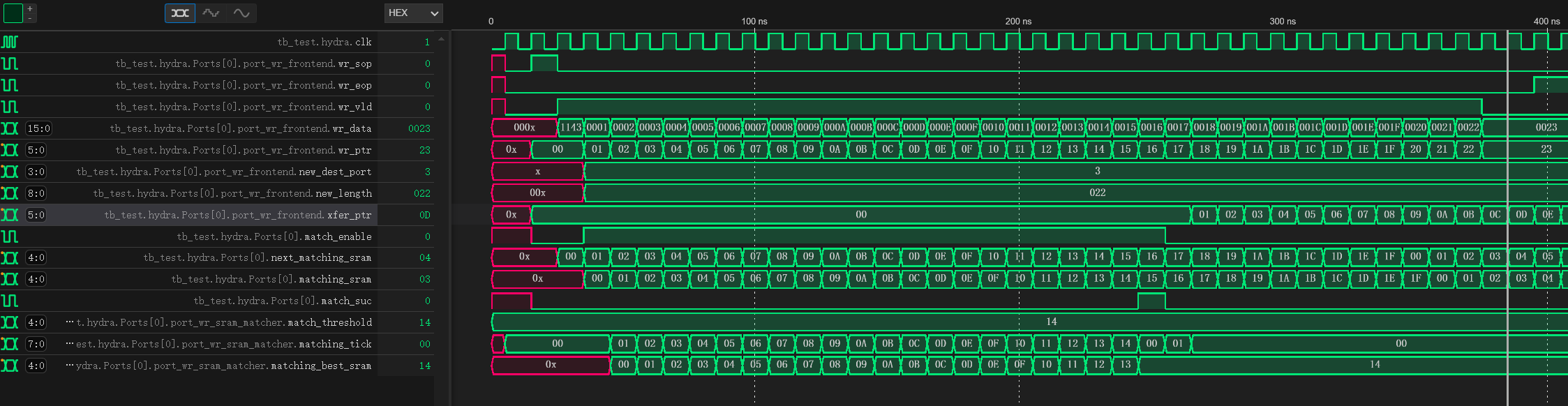
****



**图7.1.2(1) 单数据包写入读出**

上图展示了端口0向端口3发送了优先级为4，长度为34半字的数据包时，两个端口部分IO口的波形图，第一张图为整理情况，第二三张图为写入、读出阶段分别放大后的情况，对比wr\_data和rd\_data信号可验证数据传输的正确性。

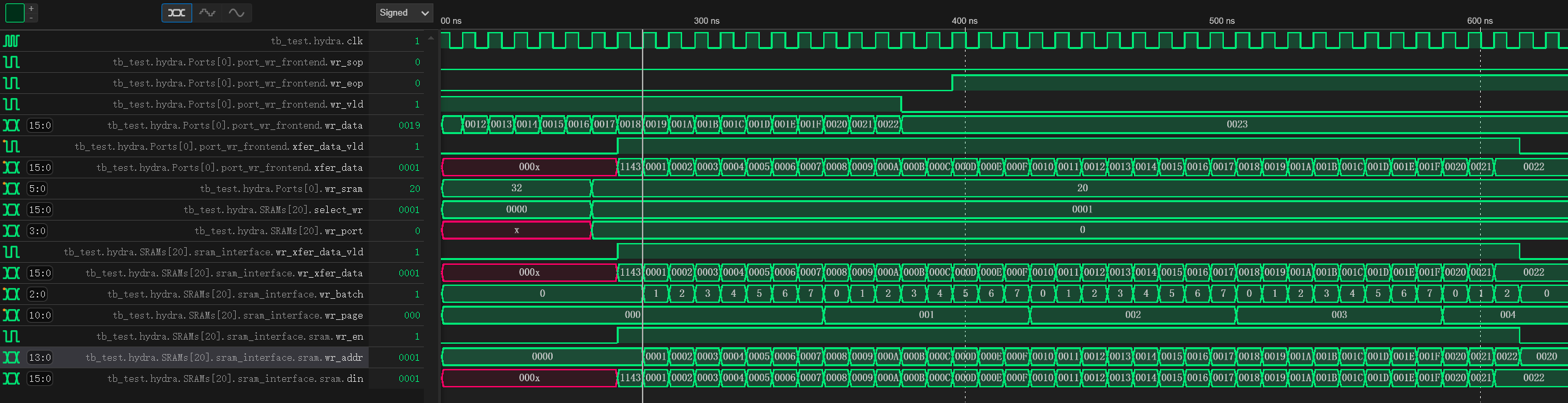
1. **前端缓冲与匹配过程**



**图7.1.2(2) 数据包进入前端后触发匹配过程的波形图**

上图中，端口0向端口3发送了优先级为4，长度为34半字的数据包，数据包先进入前端缓冲区，对应尾指针wr\_ptr不断往后移动。第一半字刚进入时，前端解析其中的控制信息，得到了数据包的目的端口new\_dest\_port和长度new\_length，并拉高启动匹配信号match\_enable。负责匹配的模块收到信号后，每周期匹配相应的SRAM（matching\_sram），且匹配时长matching\_tick不断增加，最终达到匹配阈值matching\_threshold时，发出匹配成功信号match\_suc。前端得知匹配成功后，开始像后端传输信号，对应传输指针xfer\_ptr不断往后移动。数据包最终完整地穿过缓冲区，被传输至后端存入SRAM。

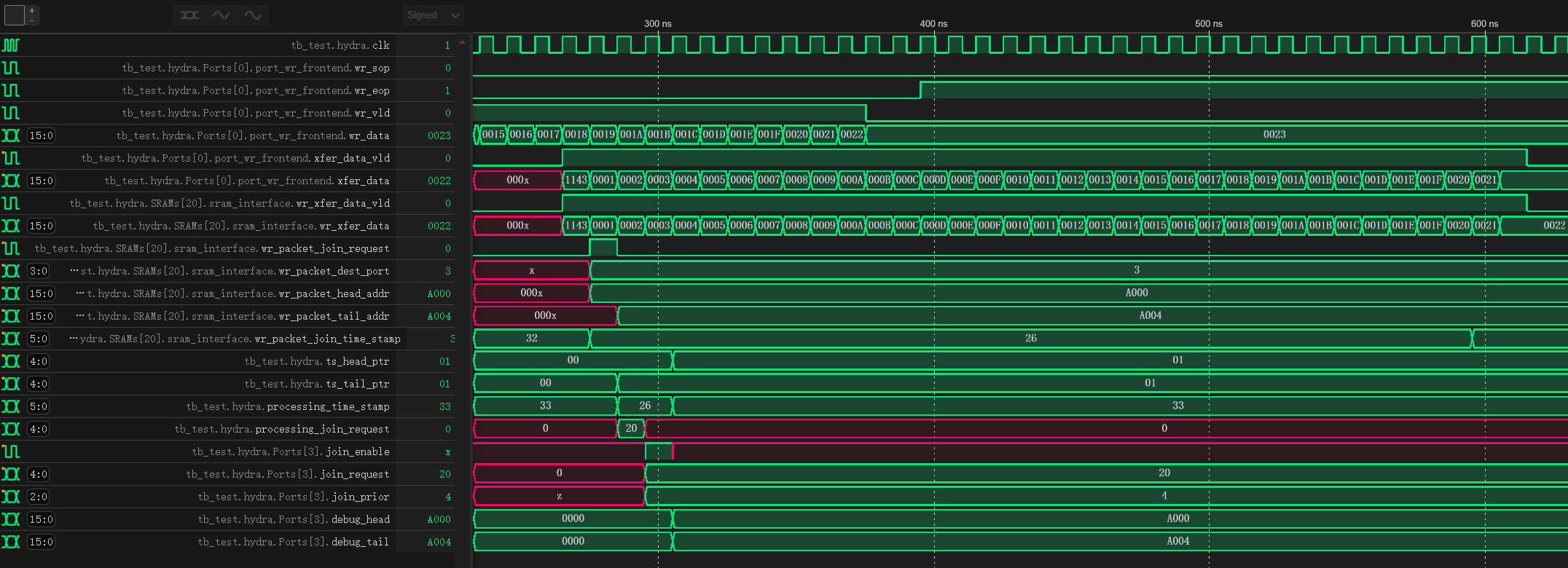
1. **前端传输数据至后端，再进一步被写入SRAM**



**图7.1.2(3) 数据包传输至后端并被存入SRAM**

匹配结束后，Crossbar矩阵中端口和SRAM的结点会被激活，上图中来自端口0的数据包匹配到了SRAM20，选通信号select\_wr的第零位置1，经过编码器，SRAM得到正在向自己传输数据的端口号wr\_port，图中为0，并通过其选择端口数据通道中的数据信号wr\_xfer\_data\_vld和wr\_xfer\_data，它们与前端传输至后端的数据信号同步。同时，SRAM封装模块会生成页地址wr\_page并启动切片计数器wr\_batch，它们组成SRAM的写地址wr\_addr。在数据包结束时，切片计数器清零，并停止写SRAM，对应wr\_en拉低。

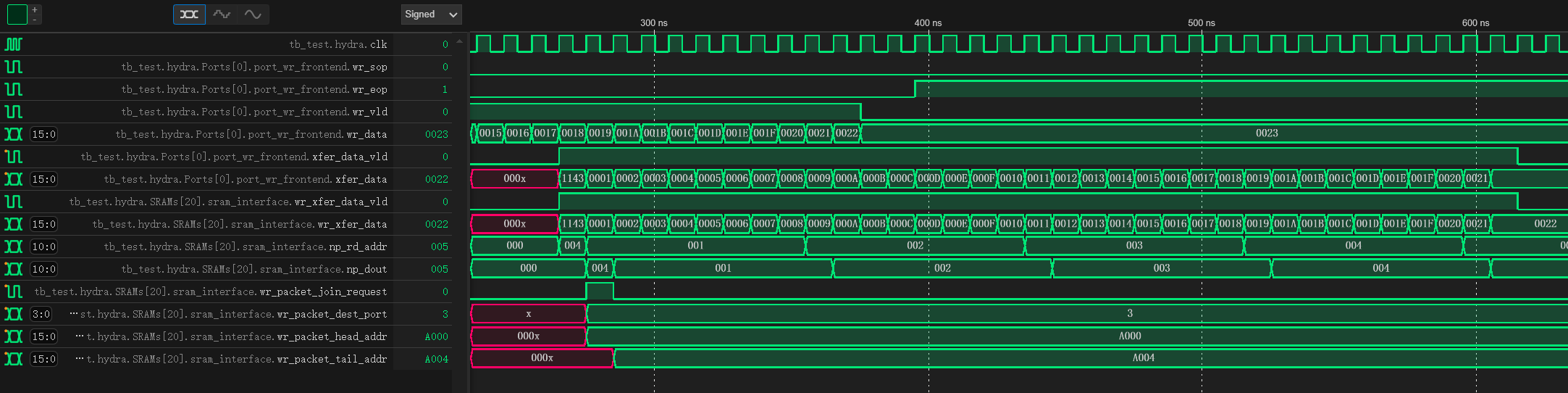
1. **入队请求的生成与时间序列管理**



**图7.1.2(4) SRAM生成入队请求到数据包最终成功入队过程**

数据包传输至后端，被写入SRAM时可得知其首页地址和尾页地址，此时即可建立入队请求，图中控制信号1143解析完毕的下一个周期，拉高入队有效信号wr\_packet\_join\_request，更新数据包的目的端口（3）和头页地址wr\_packet\_head\_addr（A000），同时为请求盖上时间戳wr\_packet\_join\_time\_stamp（26）。主控制器中检测到当前周期有入队请求，向时间序列中加入当前时间戳，对应ts\_tail\_ptr改变，在下一个周期观测时间序列顶部作为正在处理的时间戳processing\_time\_stamp（26），经过逻辑运算和选择编码得到应处理哪个SRAM的请求processing\_join\_request（20），最后由数据包目的端口拾起该请求，刷新请求缓存join\_enable、join\_request、join\_prior，完成数据包断链入队。debug\_head和debug\_tail信号最终分别变为数据包头尾页地址（A000 ~ A004），证明了该机制的正确性。

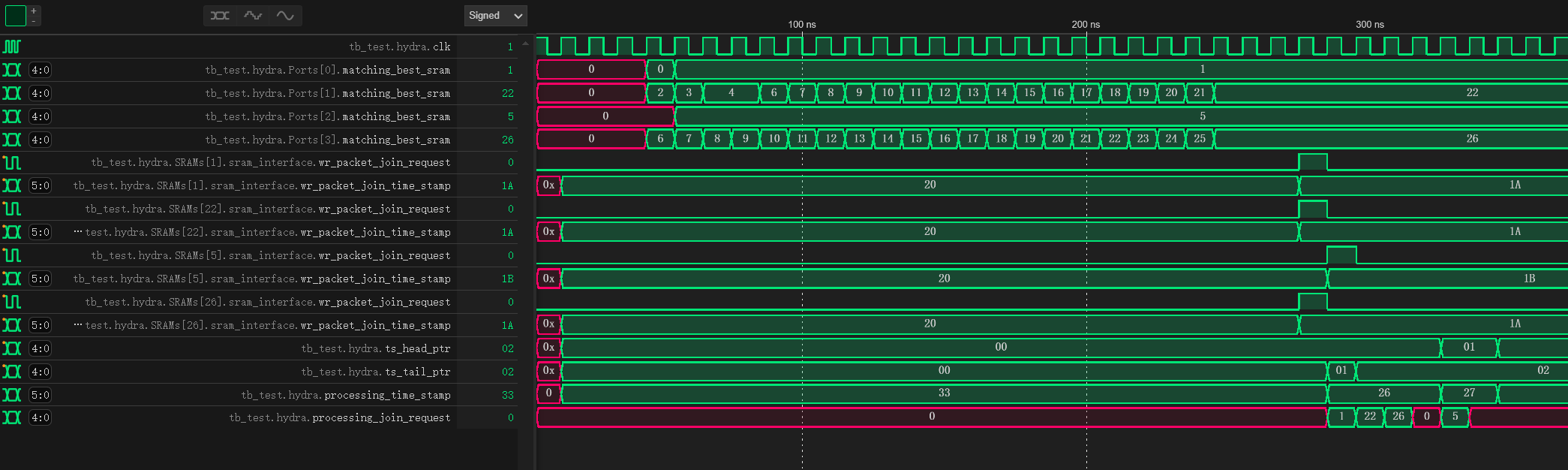
1. **基于尾部预测的快速拼接**



**图7.1.2(5) 基于尾部预测的快速拼接过程**

数据包刚进入SRAM时，可通过其长度和头页地址计算得到指向空闲队列中数据包尾页的指针，此时更新空闲队列的读取地址np\_rd\_addr（004），即可在下个周期预知数据包尾部地址，并将其当作入队请求的尾部wr\_packet\_tail\_addr（A004）。经过该过程，Hydra得以在数据包刚被写入时就启动入队、拼接过程，大大降低读取延迟。

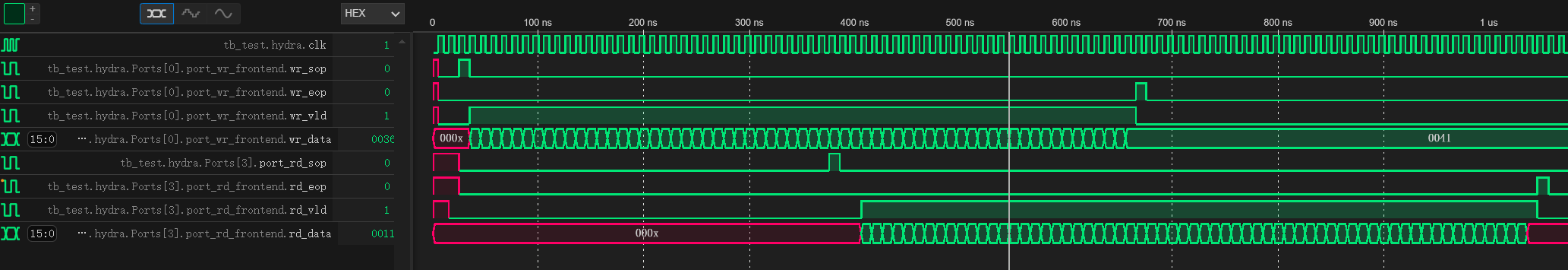
1. **多数据包的保序性**



**图7.1.2(6) 一组入队请求的保序处理**

上图是一个经典的容易在保序方面出错的入队请求组合，SRAM1、22、26同时发起入队请求（数据包来自端口0、1、3），SRAM5在后一个周期发起入队请求（数据包来自端口2），由于时间序列在接下来的三个周期只处理时间戳为26的请求，故SRAM5的请求会被遮盖，不会被处理直到主控制模块将时间戳为26的请求全部处理完毕，开始处理时间戳为27的请求。图中processing\_join\_request即为入队请求处理顺序（1、22、26、5）。

1. **基于快速拼接的无延迟/提前读取**

****

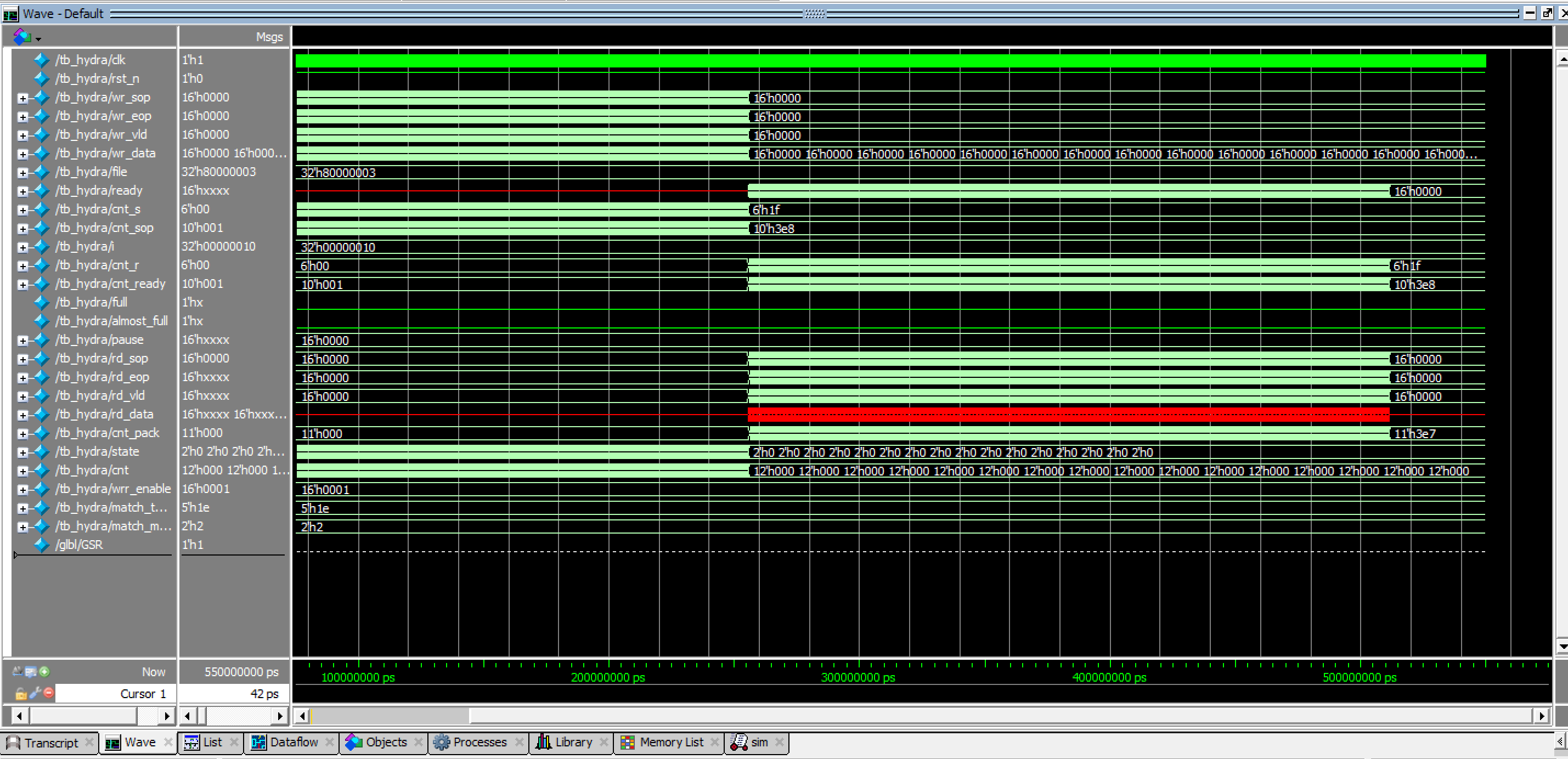
**图7.1.2(7) 提前读取数据包**

上图是提前读取数据包的案例，在外界设备还没完成写入时，Hydra就已经允许读出该数据包了。这样可以消除平均80%的读出延迟，因为可读出实际不再需要被数据包的长度影响。

### 7.1.3压力测试（Java生成输入数据，ModelSim10.5生成.do和.wlf波形图）

**压力测试方案**：模拟突发传输写入大量数据，写完后每隔500周期拉高各端口的ready（方便调试与数据统计），写入的均为最短数据包。

**1、单端口线性映射写入1000个64字节数据包并完全读出**

****

**写入用时**：256us **间隔**：64周期

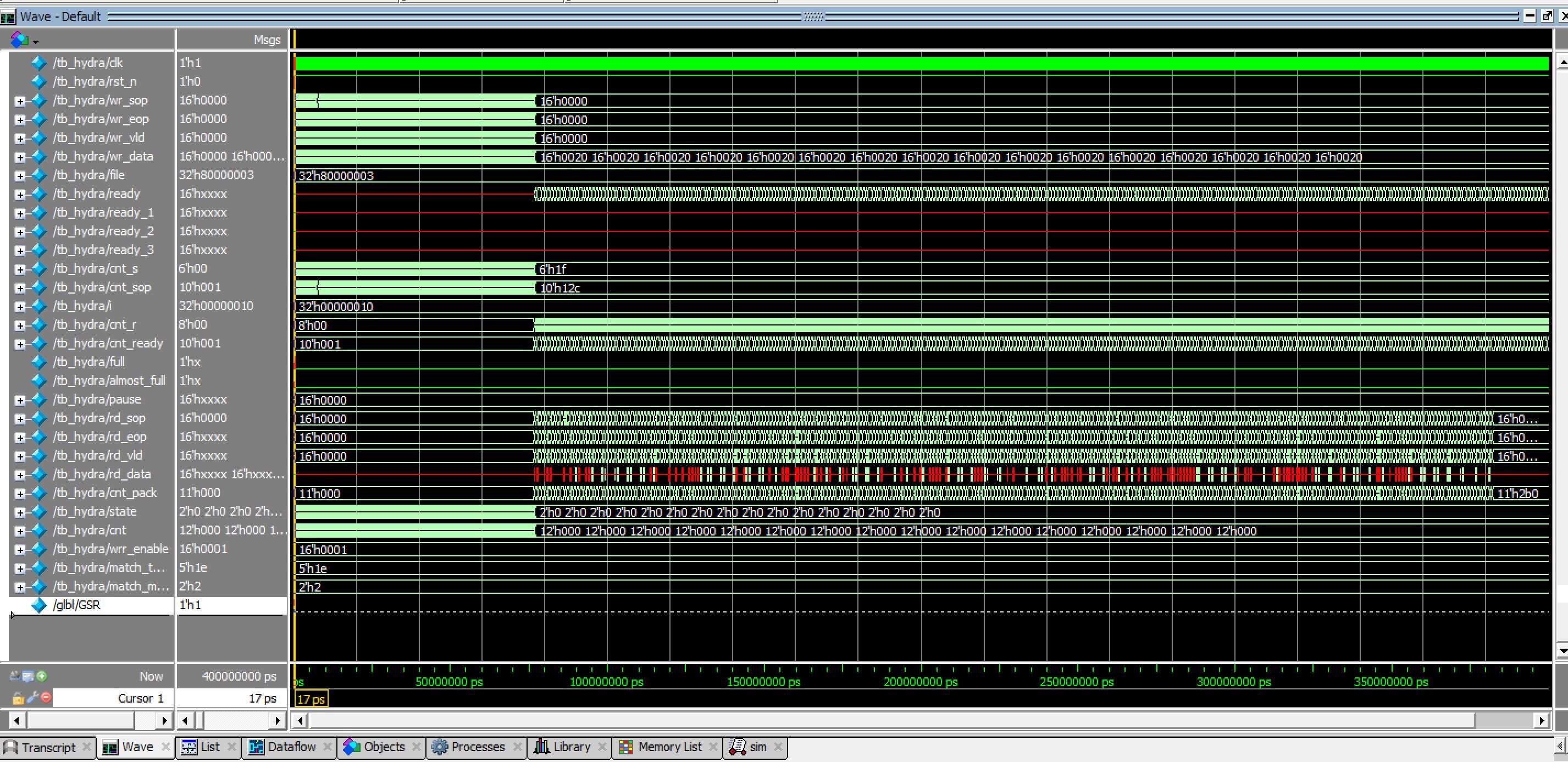
**读出用时**：256us **间隔**：64周期

**交换数据量**：512000bits

**测试写入速度**：

**测试读出速度**：

**2、16个端口同时线性映射写入共4800个64字节数据包并完全读出**



**写入用时**：77us **间隔**：64周期

**读出用时**：305us **间隔**：256周期

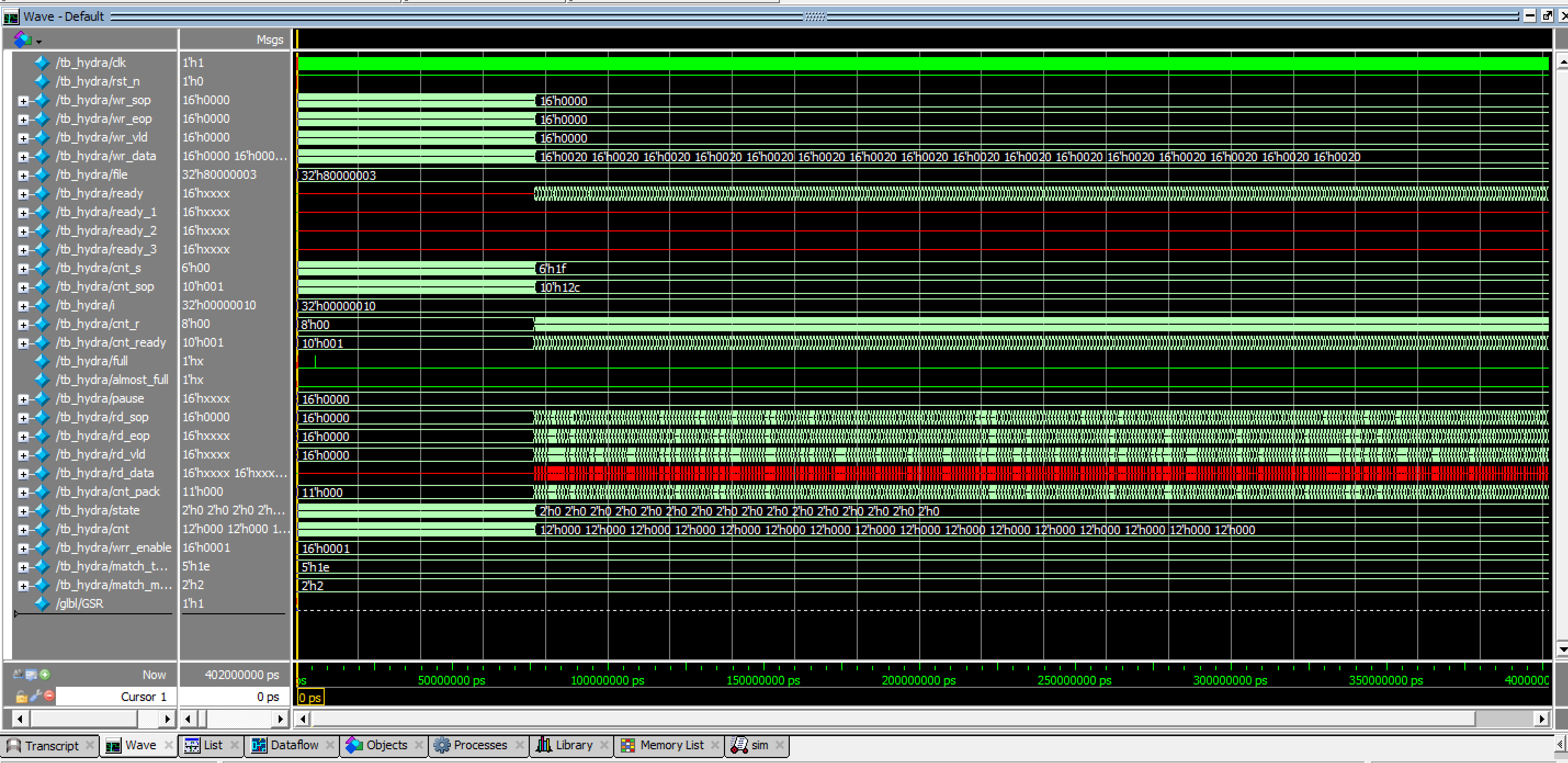
**交换数据量**：2457600bits

**测试写入速度**：

**测试读出速度**：

**读出冲突**：最大3，平均1.03

**3、16个端口同时随机映射写入共4800个64字节数据包并完全读出**



**写入用时**：77us **间隔**：64周期

**读出用时**：325us **间隔**：256周期

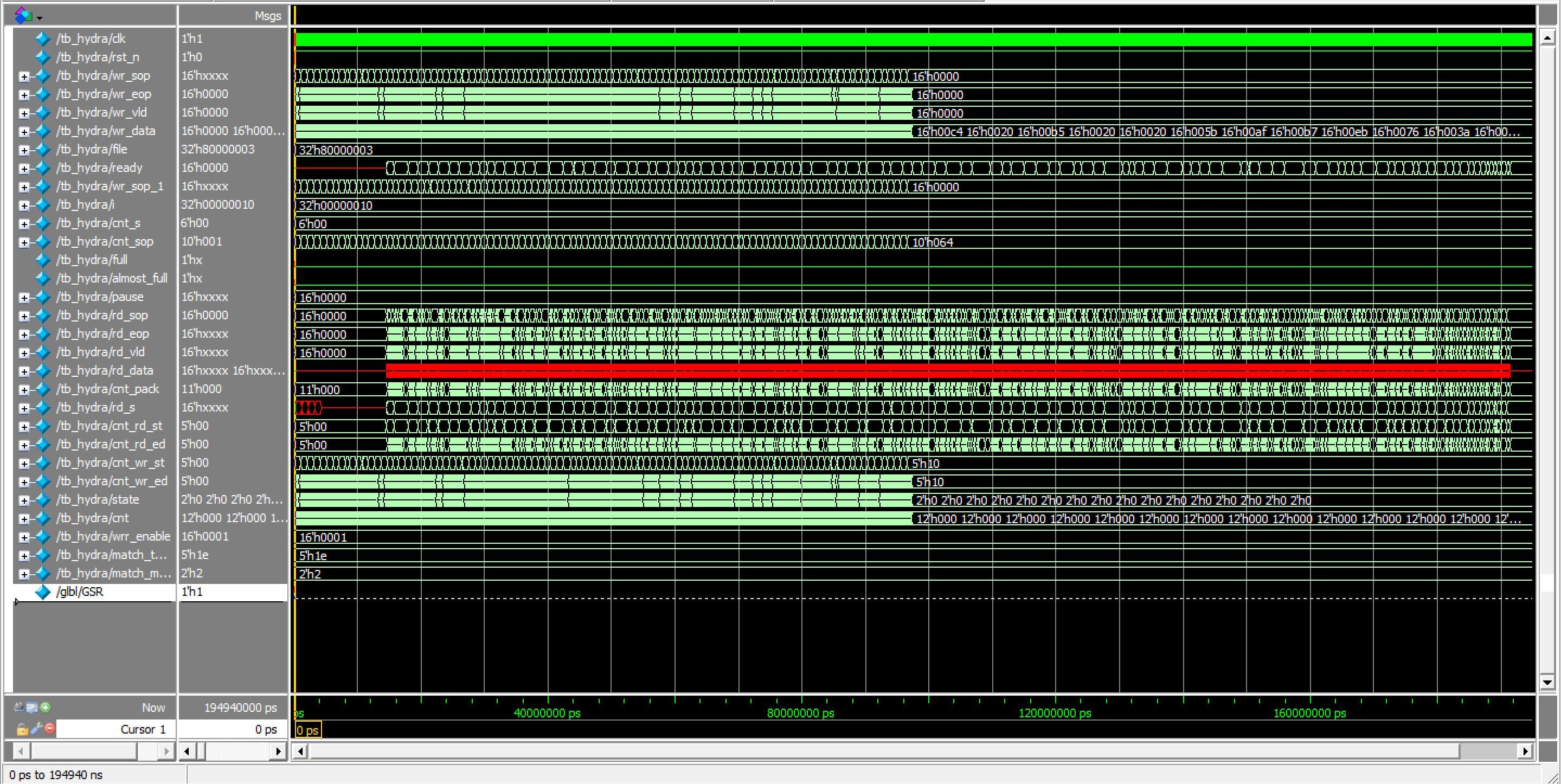
**交换数据量**：2457600bits

**测试写入速度**：

**测试读出速度**：

**读出冲突**：最大6，平均1.385

**4、16个端口同时随机映射写入共1600个随机长度数据包并完全读出**



**写入用时**：95us **间隔**：1周期

**读出用时**：177us **间隔**：同时拉高ready

**交换数据量**：3246576bits

**测试写入速度**：

**测试读出速度**：

**读出冲突**：最大4，平均1.3

## 7.2硬件实现验证

### 7.2.1 验证环境

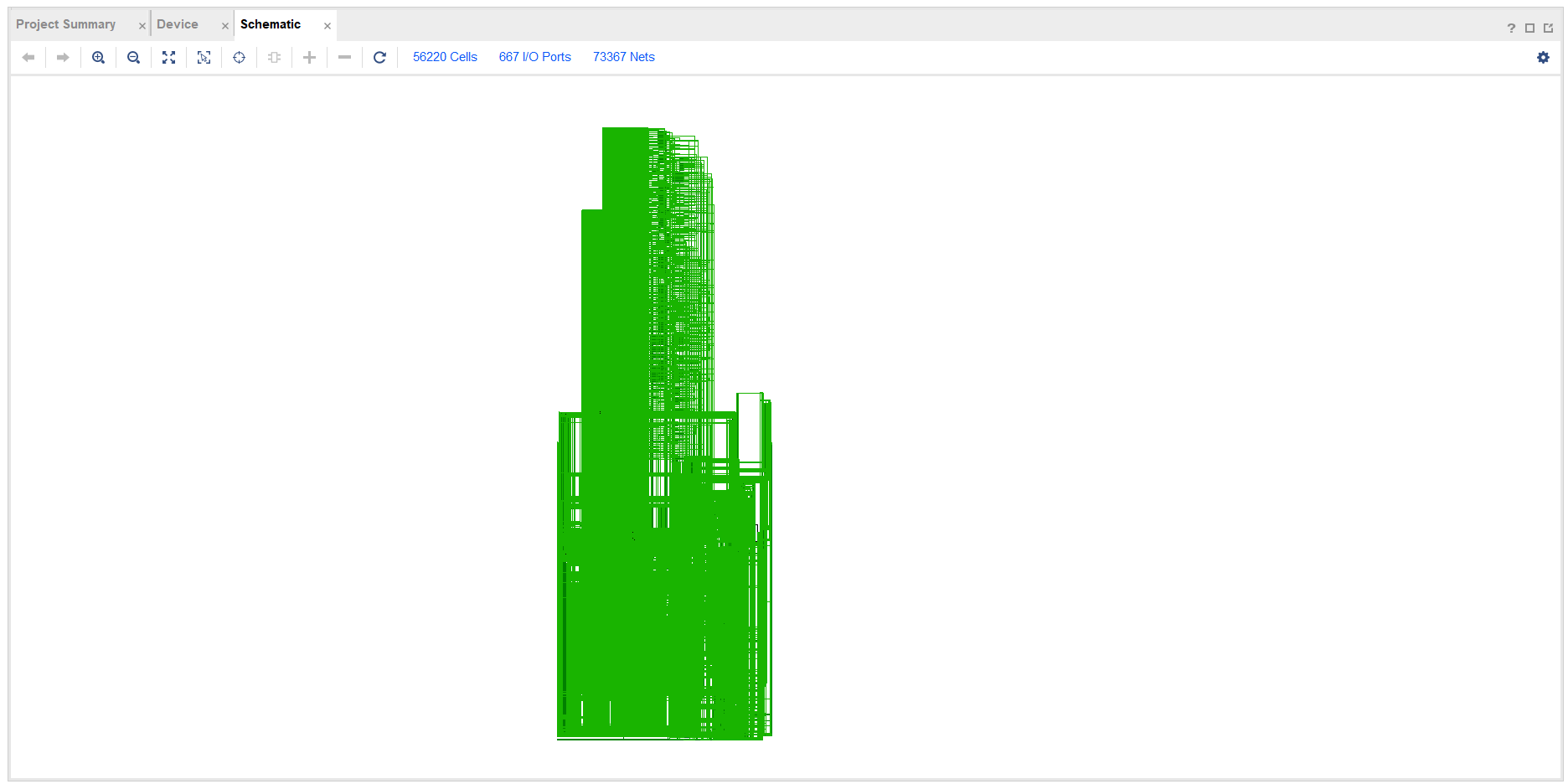
**Vivado版本**：2018.3

**芯片型号**：xcku115-flvb2104-2-e

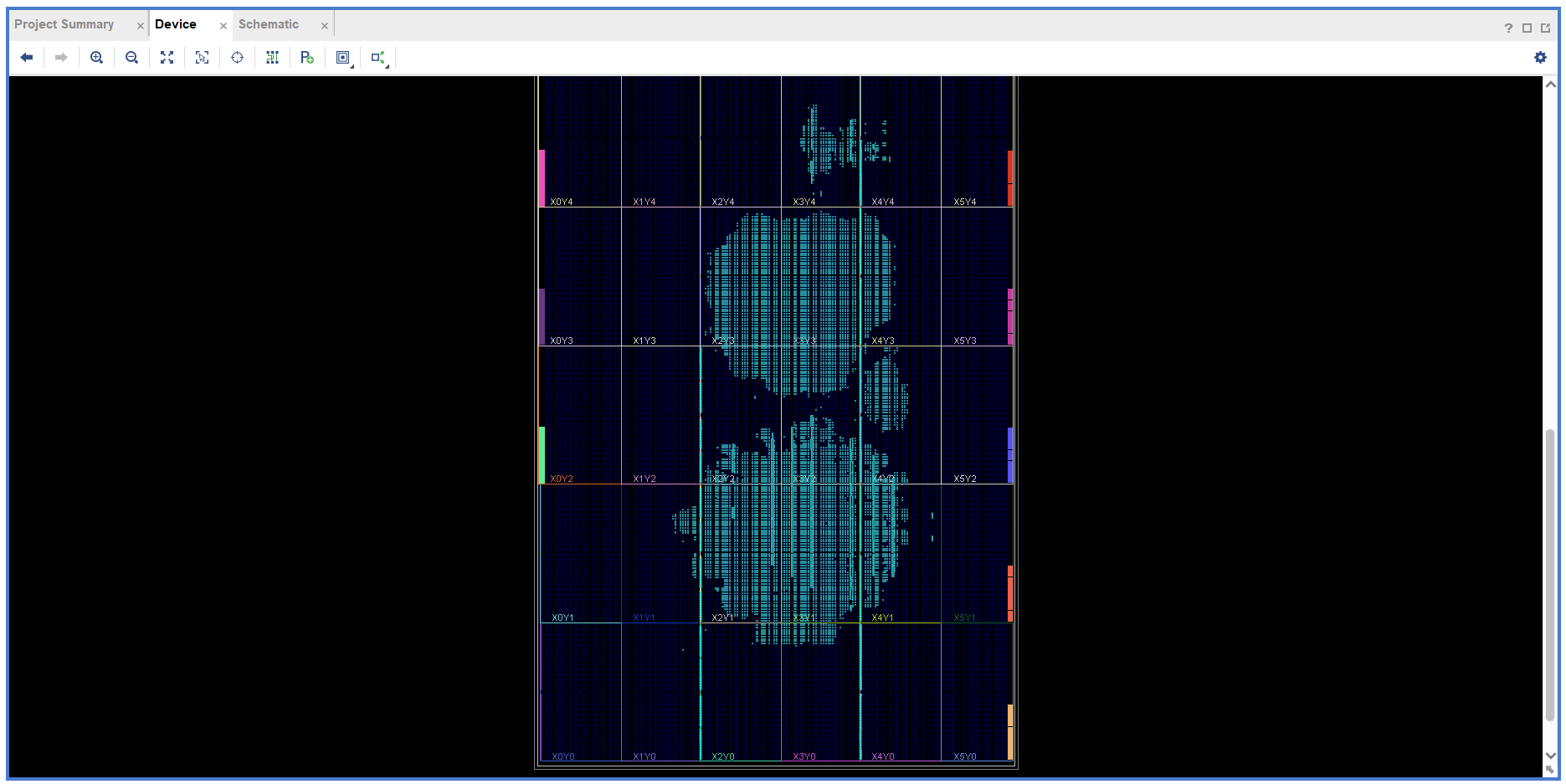
**综合策略**：Flow\_AreaOptimized\_high

**布线策略**：Performance\_NetDelay\_high

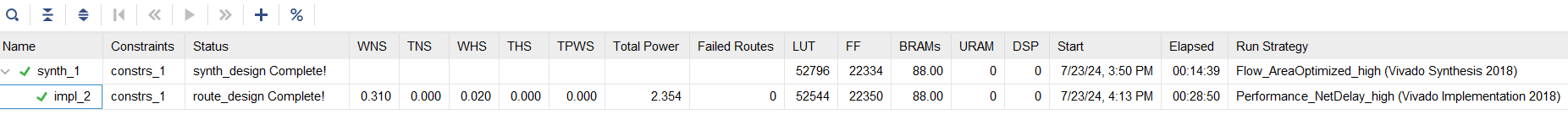
### 7.2.2 综合布线结果



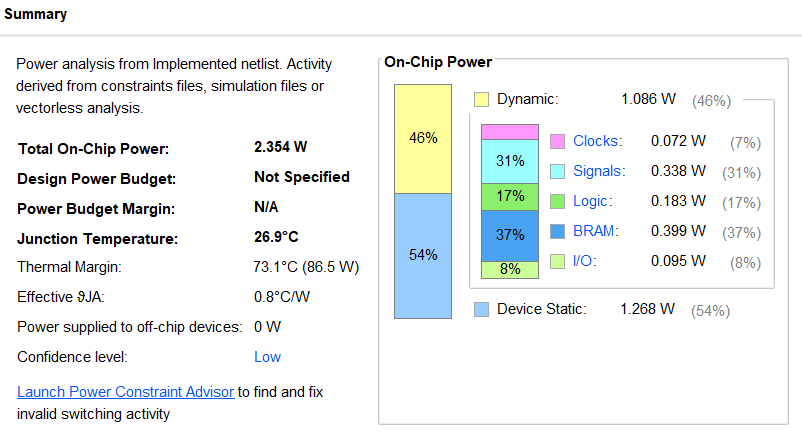
**图7.2.2(1) 布线结果电路图**



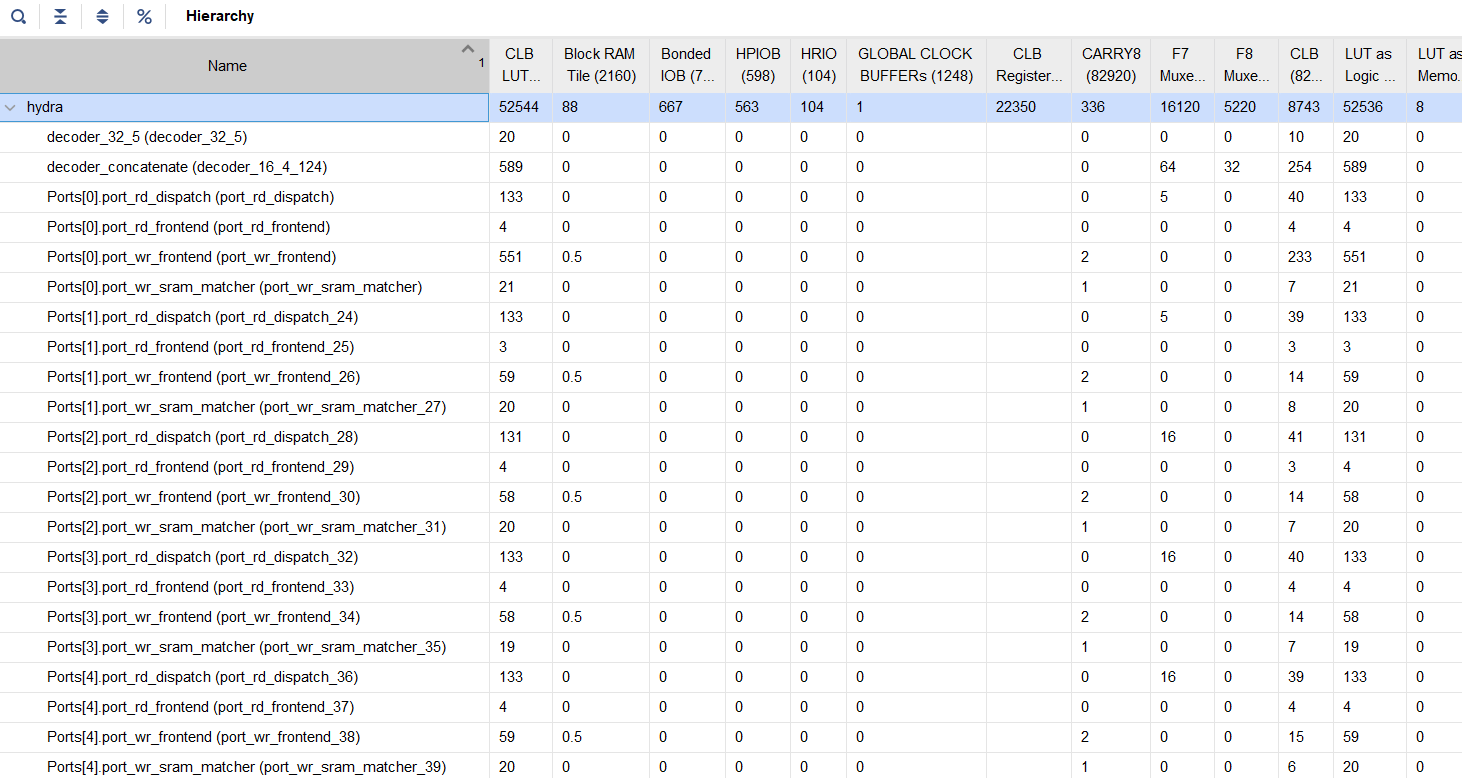
**图7.2.2(2) Device（布线阶段）**



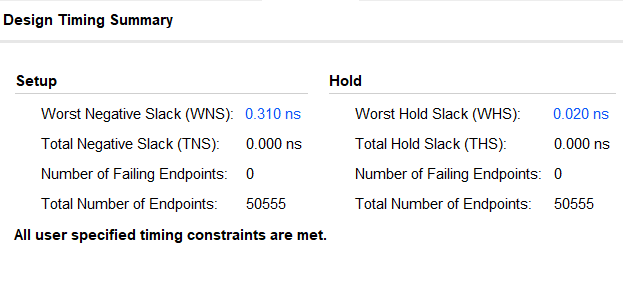
**图7.2.2(3) 综合、布线结果**



**图7.2.2(4) Power Summary（布线阶段）**



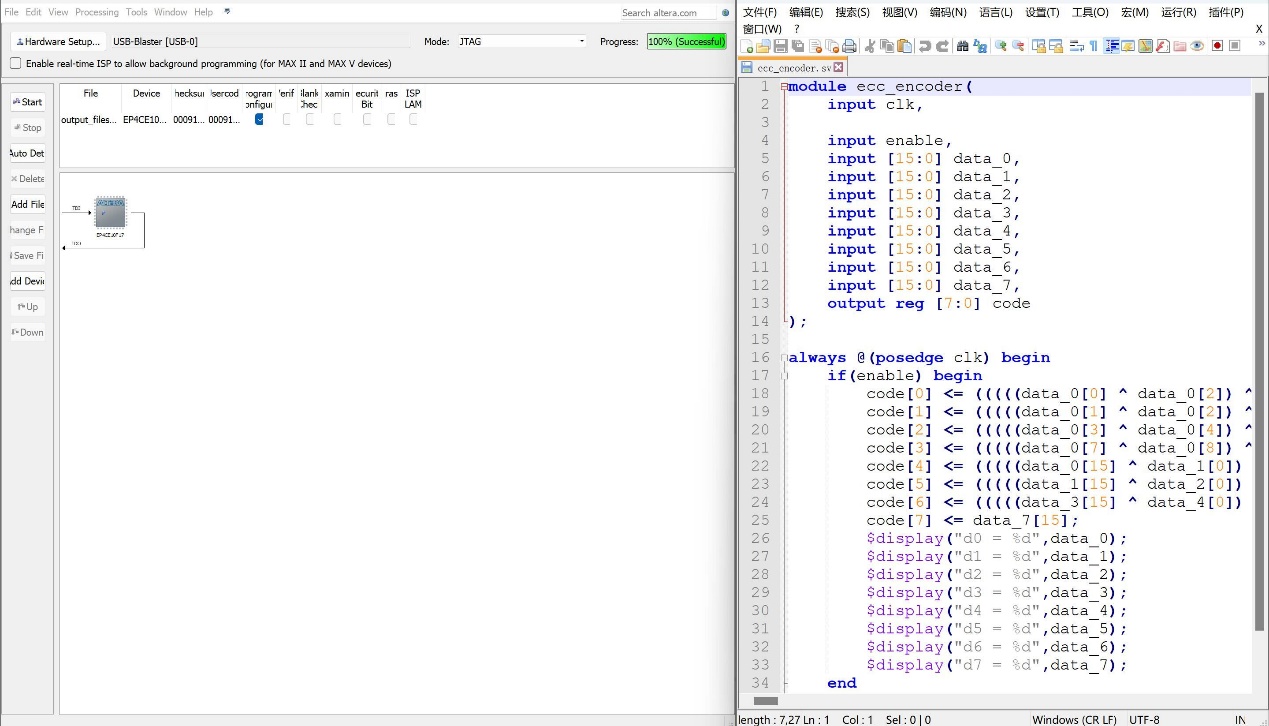
**图7.2.2(5) Utilization（布线阶段）**



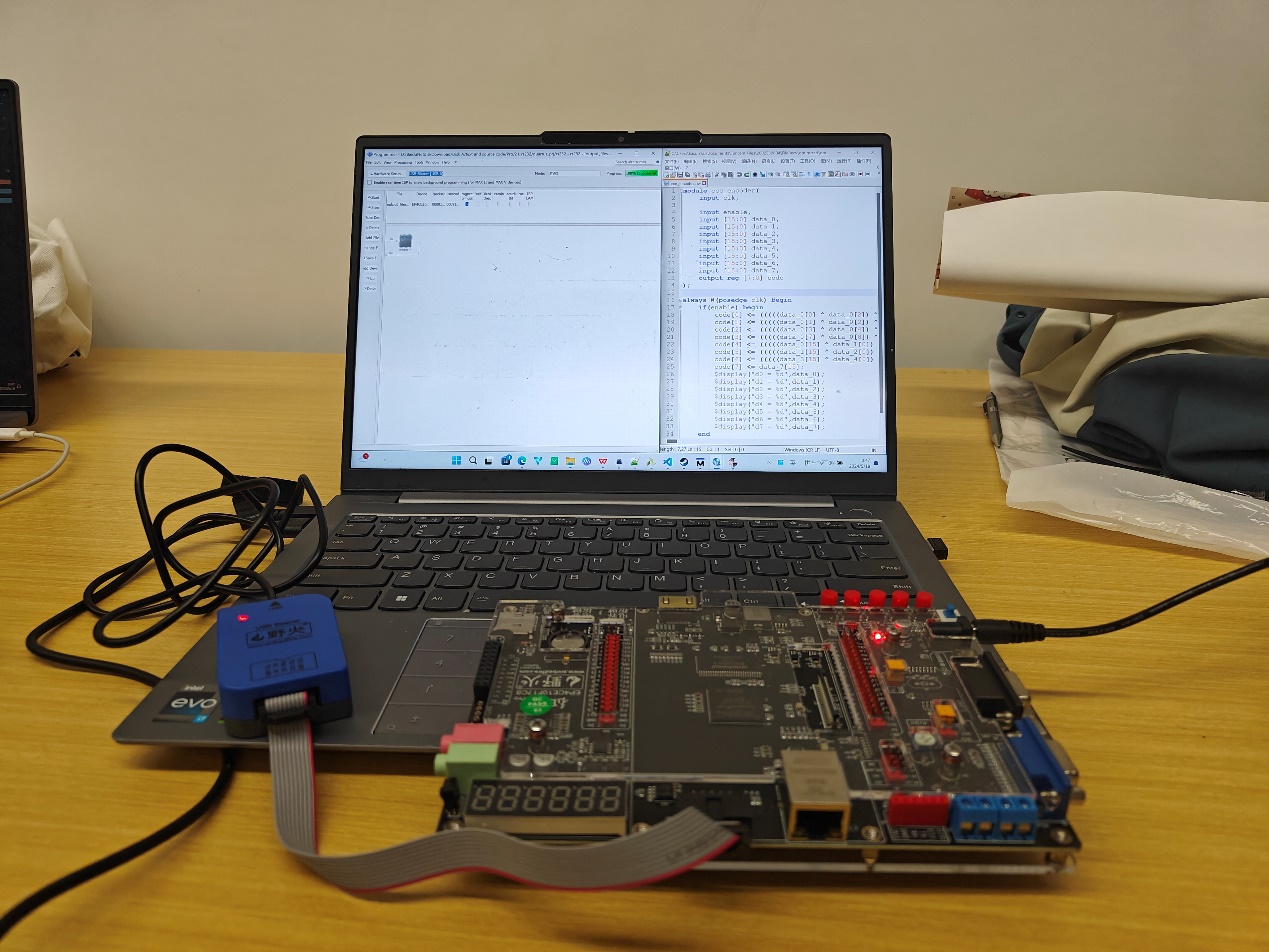
**图7.2.2(6) Timing Summary（布线阶段）**

## 7.3 FPGA验证

### 7.3.1 Ecc校验编解码模块



**图7.3.1(1) 程序烧录截图**



**图7.3.1(2) FPGA板运行图**

# 8.设计优缺点

## 8.1优点

1）完全动态分配内存，各端口、队列数据量无限制；

2）引入数据结构与算法思想，降低逻辑复杂度；

3）引入页表管理的思想，进一步划分SRAM，减少了存储时需要的控制信息；

4）时序性良好，使用跳转表替代时序性差的位图，各模块机制流水化；

5）偏好性存储降低读取延迟，使得大多数情况下端口读取带宽可以达到最大值；

6）在有可用SRAM时保证端口写入带宽总为最大值；

7）极大压缩模块的存储资源占用（相较于传统方案而言）；

8）极低的读写延迟，在写入未完毕时即可开始读出；

9）代码风格良好，严格遵循命名规范，注释清晰。

## 8.2缺点

1）Crossbar架构对模块布线压力稍大。

注：若压力稍大，可以选择降频。由于Crossbar架构的优越性，Hydra在低频率下带宽仍能远超总线高频率下的带宽，具体数据见下表。



**图8.2 SRAM宽度为16时两种架构在不同频率下的理论带宽**

# 9.后续开发计划

## 9.1进一步优化读写延迟

继续简化读写逻辑，降低传输延迟XD。

改良快速拼接机制，降低入队拼接延迟C。

## 9.2更全面的压力测试

模拟更加极端的情况，测试时间更长，吞吐量更高的情况。

编写全自动压力测试脚本，高级语言按照测试策略随机生成输入数据，testbench读取输入数据并保存模块输出至文件，高级语言再读取输出文件进行对拍。

**9.3模拟现实应用中的文件传输**

使用模块同时传输多个大文件，并比对前后传输结果。

设置系统仪表盘，更加方便地观测SRAM、端口内部状态和彼此连接关系。

## 9.4继续提升代码质量

进一步整理代码，丰富注释。