

高性能浮点DSP助力，加速系统的大数据吞吐率

Harry Xin,
ARC DSP处理器产品营销经理
2020年5月13日



广泛使用的ARC处理器产品方案

第2名

处理器IP出货

250+项

ARC内核授许

5大产品系列

提供业界最具扩展性、兼容性的
处理器产品组合

约26亿

每年交付的带ARC内核的芯片

第一款

ASIL D级
满足认证的处理器

20款

DSP增强处理器

DesignWare ARC处理器IP

EM 系列



- 针对超低功耗物联网进行了优化
- 带高效DSP的3级流水线
- 功耗低至3uW/MHz
- 小至0.01mm²@28HPM

SEM系列



- 用于物联网和移动应用的安全处理器
- 保护硬件、软件和侧通道免受攻击
- 适用于可信执行环境的SecureShield（安全防护）

HS系列



- 迄今，性能最高的ARC CPU
- 带有DSP的高速10级流水线
- SMP Linux支持
- 单、双、四核配置

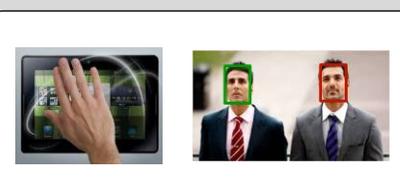
NEW

VPX系列



- 高性能的矢量DSP
- SIMD/VLIW设计，用于大规模并行处理
- 多向量浮点引擎可实现高精度计算

EV系列



- 异构多核，视觉处理
- 最先进的卷积神经网络（CNN）
- 高产出、基于标准的工具套件

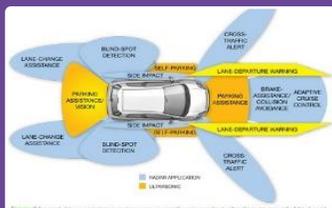
功能安全（FS） 处理器

- ARC EM、HS、VPX和EV处理器系列，集成硬件安全功能
- 加快针对汽车安全SoC的ISO 26262认证

关键应用，推动大规模数字信号处理的使用

传感器融合

- 更大的传感器阵列配置
- 可提高被检测对象质量的交叉计算
- 更多的传感器输入通道，可达约10GB/s的更高数据速率计算



雷达/激光雷达

- 超过100GB的吞吐量
- 感测距离范围更广
- 提高了准确性，增强检测能力
- 能够检测不可见的目标



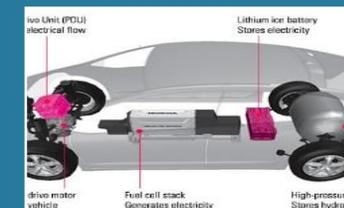
无线通讯

- 高达10Gbps的高数据速率的5G NR
- 信道预测算法，提高复杂性
- 信道条件映射中的机器学习算法



汽车/动力总成

- 多种汽车动力单元切换
- 节能的动力总成和发动机管理的算法提升



越来越多的先进算法，采用浮点计算

复杂算法建模

增加动态范围

上市时间

低延迟

软件算法赋能的产品差异化

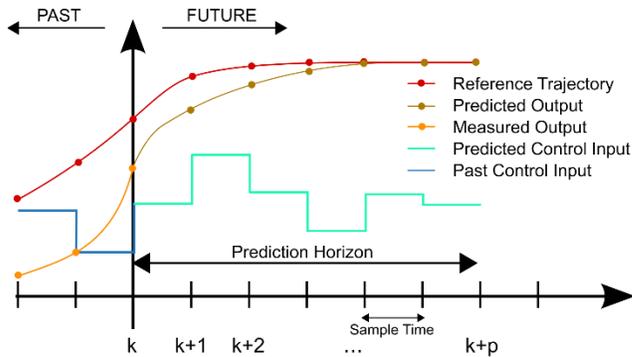
低频率带来的低功耗

最新的应用场景，需要预测性建模

可更快、更流畅地做出响应的真实世界的环境建模和预测

- 反馈控制：使用MPC来建模
- MPC使用系统模型来预测其未来行为
- 建模是线性代数计算

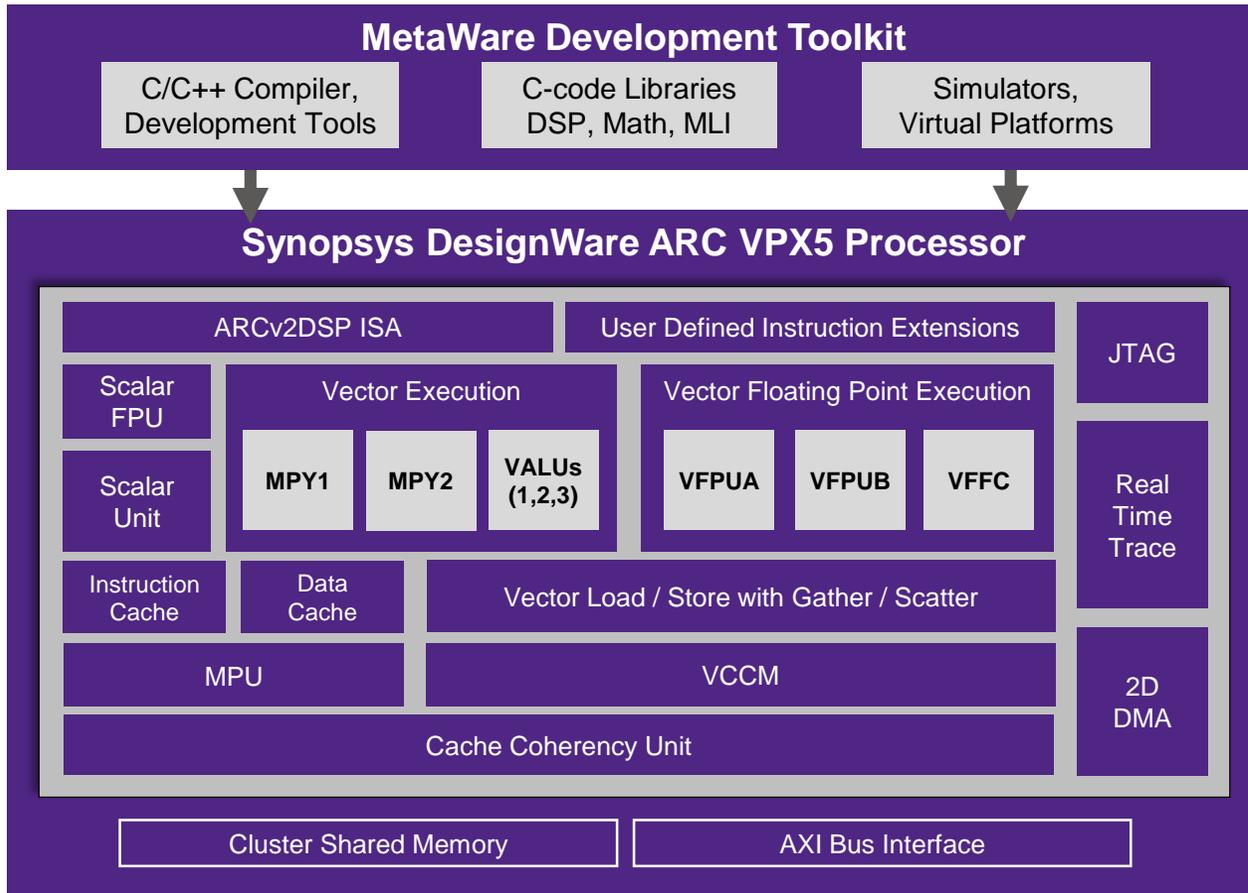
MPC: Model Predict Control



新的DesignWare ARC VPX5 DSP 处理器IP



面向以数据为中心的下一代DSP架构



- 先进的**VLIW/SIMD DSP处理器**可处理各种DSP工作负载，包括汽车ADAS、基带通信和传感器融合
- **超高性能浮点处理**可为最新算法提供高精度结果
- **线性代数硬件加速**，适用于需要高吞吐率数学计算的应用
- **可扩展和可配置**，针对特定的SoC要求调整性能、大小和功耗
- 为**高效机器学习**而优化的架构、数据类型和软件库
- 使用针对DSP处理器硬件优化的ARC MetaWare编译器进行**标准C编程**

数字信号处理应用

VPX在Synopsys处理器产品组合中的定位

*VPX5是VPX DSP 产品系列中的首款

VPX5*

512位向量DSP



- 通信基带处理
- 传感器计算 (雷达/激光雷达)
- 动力总成/发动机管理
- 高端物联网边缘节点
- 下一代面向对象的音频

HS4xD

带DSP增强功能的高性能控制器



- 组网
- 存储
- 通信控制器
- 高端音频

EMxD

具有DSP增强功能的超低功耗、小尺寸控制器



- 窄带物联网 (NB-IoT)
- 物联网边缘节点
- Always on sensing
- 音频和语音

DSP性能要求

ARC VPX5: 具有独特特性, 满足最新应用对DSP的需求



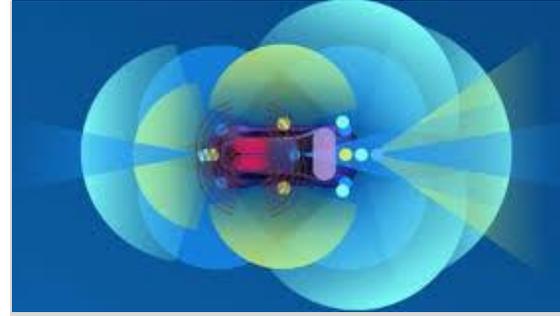
激光雷达

- 丰富的浮点资源可提高鉴别目标时的准确性
- 对线性代数和机器学习的硬件和软件支持, 可实现更小的传感器阵列尺寸、降低成本



雷达

- 高带宽向量浮点提供更高的目标检测精度, 能够检测静止目标周边情况
- 先进的线性代数计算, 可改善目标跟踪, 实现更及时的碰撞预警



传感器融合

- 线性代数和AI算法加速, 可减少物理传感器的需求
- 更高精度的浮点结果, 可提供更佳准确性, 以更好地制定决策



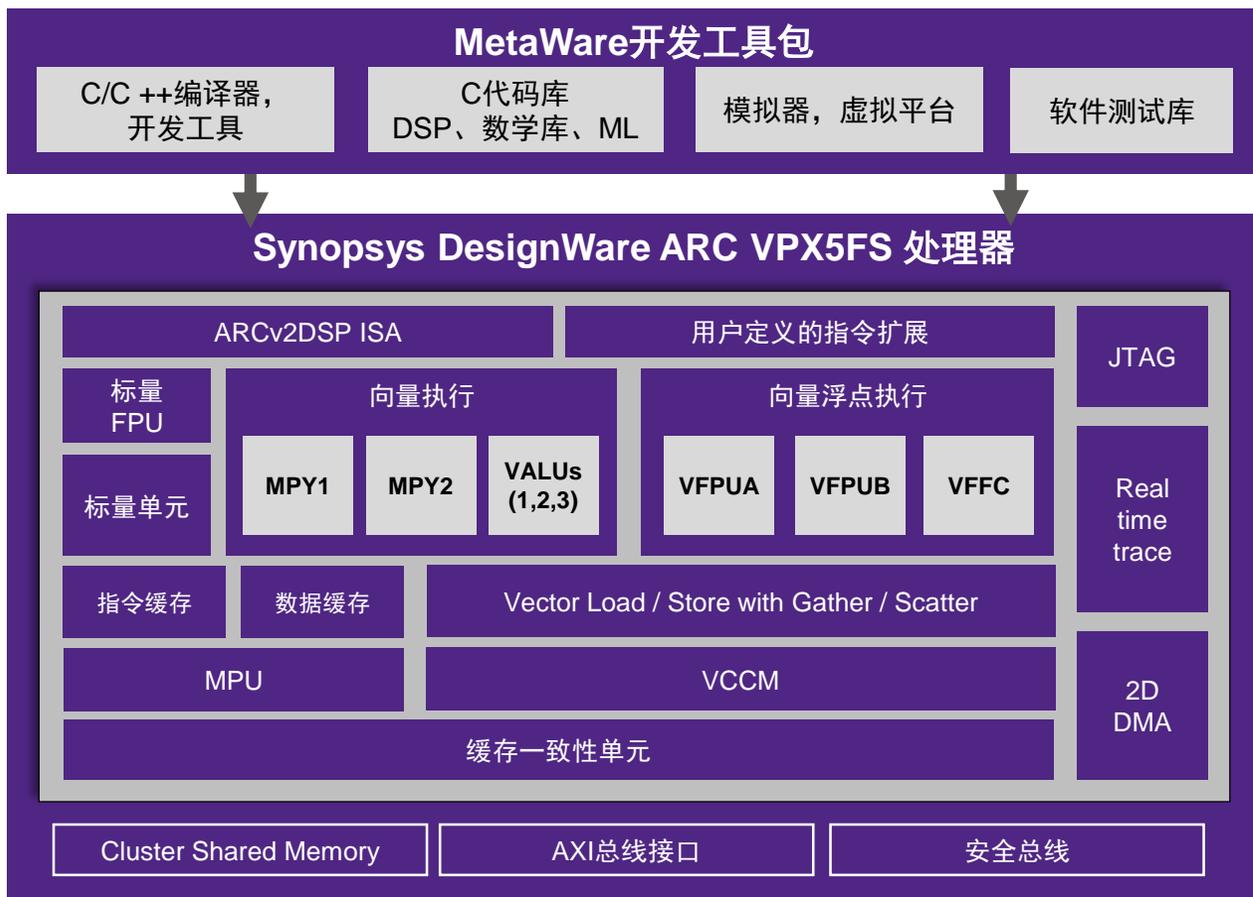
5G通信

- 向量浮点吞吐率, 支持5G数据速率
- 向量硬件的灵活使用, 可实现高效的信号处理, 并针对低功耗进行优化

DesignWare ARC VPX5FS DSP处理器IP



符合ISO 26262 ASIL B和ASIL D标准的功能安全DSP



- 解决了广泛的汽车应用需求，如ADAS传感器（激光雷达、雷达），动力总成、传感器融合等
- 集成了安全关键硬件特性
 - 所有存储器和接口的ECC保护
 - 安全监控器和安全总线
 - 锁步功能
- ARC MetaWare安全开发工具包可加快符合ISO 26262的软件开发速度
 - 包括不需要硬件冗余时符合ASIL-B的软件测试库
- 安全文档：FMEDA报告和安全手册可加快功能安全评估的速度

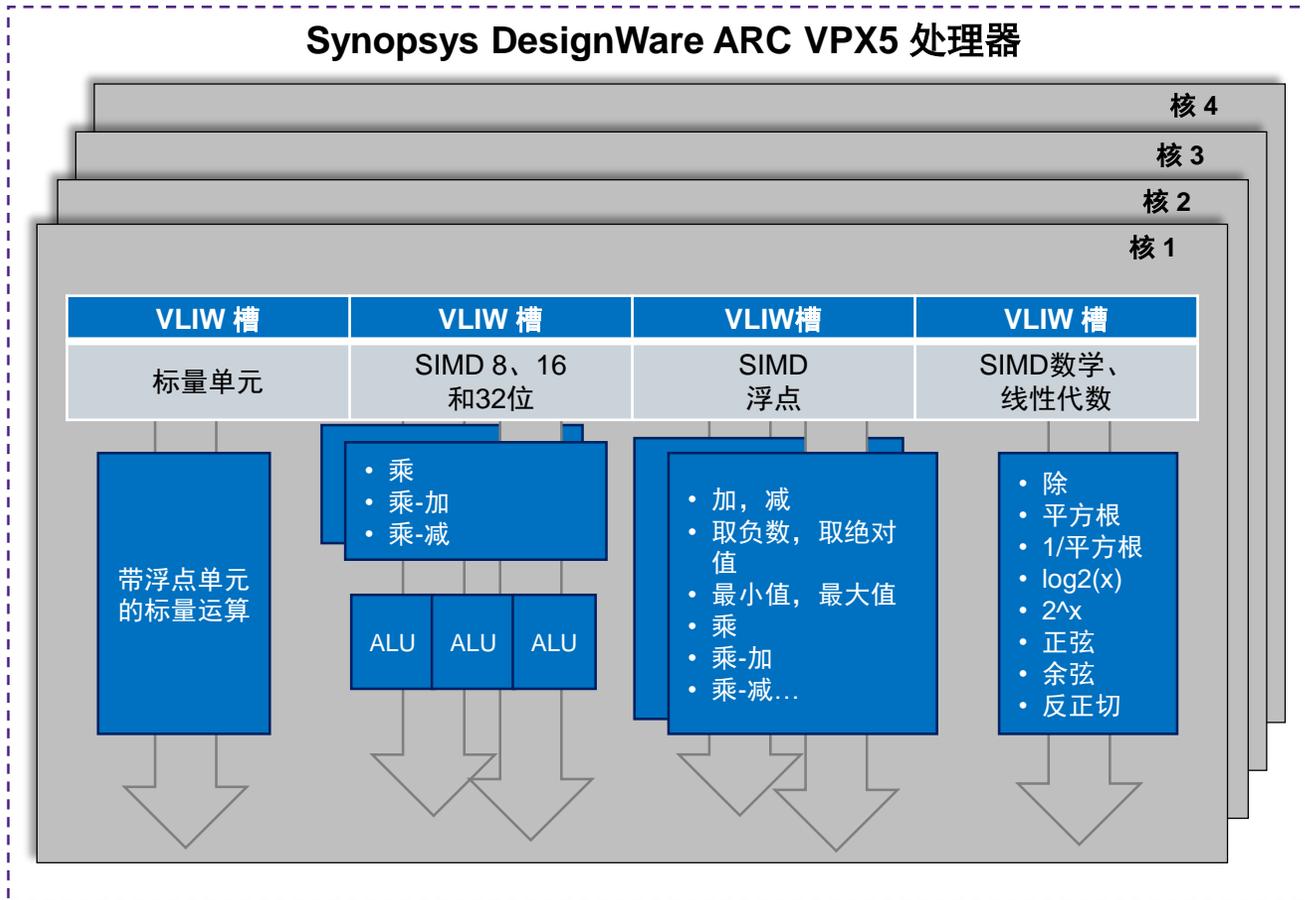
VPX5 DSP处理器规格

Parameter	Specification
Pipeline stages	Base Core Configuration with 14-stages
VLIW	4 slots
Instructions Word Length	Variable, maximum word length 128-bit
Vector Load/Store Width	512-bits
Vector Register Bank	32 by 512-bits
Predicate Register Bank	8 of 64-bit (max)
Data types	Integer 8-bit, 16-bit and 32-bit Floating point, 16-bit (half precision), 32-bit (single precision) and 64-bit (double precision)

高性能架构，大规模并行执行

一个时钟周期内，最多可执行512个乘法运算、512个浮点运算（FLOP）和32个数学浮点运算

Synopsys DesignWare ARC VPX5 处理器



- 512位向量长度可实现超高并行计算
- 三个向量浮点流水线可为对精度敏感的应用提供高精度结果
- 独特的线性代数/数学引擎可对自然现象进行更高效建模，以实现更好的系统响应
- 8位整数支持，可优化机器学习推理计算
- 集成的编译器实现了硬件并行能达到的最大优势
- 单一VPX5处理器可选1至4个内核，因此设计师可将硬件资源与DSP工作需求匹配

VLIW槽分配

标量、向量和浮点数

槽 1	槽2	槽3	槽4
标量	向量		
标量 ALU 标量 FPU 标量加载/存储	VALU1 VMPY1 VFPUA	VALU2 VMPY2* VF PUB* VF FC* PXU	VALU3 VLD/VST Shuffle

*可作为配置选项

可变指令长度，实现最少代码量

- 16到128位指令宽度
- 1个16位或32位标量指令长度
- 1个32位向量指令长度
- 向量+ VL（VLIW槽中已编码的32位立即数）
- 向量+ VK（VLIW槽中已编码的16位立即数）
- 支持任何组合：
 - 1个标量+ 2个向量
 - 2个向量
 - 3个向量+ VL
 - 1个标量+ 1个向量+ VK

业界领先的向量SIMD计算吞吐率

指标	性能 *
8位向量MAC/每周期（整数）	512
16位向量MAC/每周期（整数）	256
32位向量MAC/每周期（整数）	128
½精度浮点运算/每周期	512个浮点运算
单精度浮点运算/每周期	256个浮点运算
数学浮点计算/每周期（VFFC） （除、平方根、1/平方根、log ₂ (x)、2 ^x 、正弦、余弦、反正切）	32个数学浮点运算

假设：

配置了双乘法单元（MPY）的乘法器

2条浮点累加器（FALU）流水线，包含乘+加配置

四核配置

支持向量浮点

向量DSP的配置选项

- 高性能向量FPU
 - 配置选项——每核1或2个向量FPU流水线
 - 高达128G浮点运算/每核（16位浮点乘-加@1GHz）
- 支持32位和16位浮点
- 可选的特殊数学单元
 - 数学函数：除、 \sqrt{x} 、 $\frac{1}{\sqrt{x}}$ 、 $\sin(x)$ 、 $\cos(x)$ 、 $\log_2(x)$ 、 2^x 、 e^x 、 $\text{atan2}(x)$
- 符合IEEE754
 - 16位半精度浮点、radix-2，符合IEEE 754-2008标准
 - 32位单精度浮点、radix-2，符合IEEE 754-2008标准
 - 使用4种舍入模式正确舍入
 - 全精度（0.5ULP，向最近偶数舍入）——也用于FDIV、FSQRT
 - 浮点标志（无效、被零除、上溢、下溢、不精确）
- 符合OpenCL-C 2.0和C语言

向量浮点，业界领先的性能

跨2个向量流水线SIMD FPU引擎的并行计算



编译器可以跨2个
向量流水线进行向
量化处理

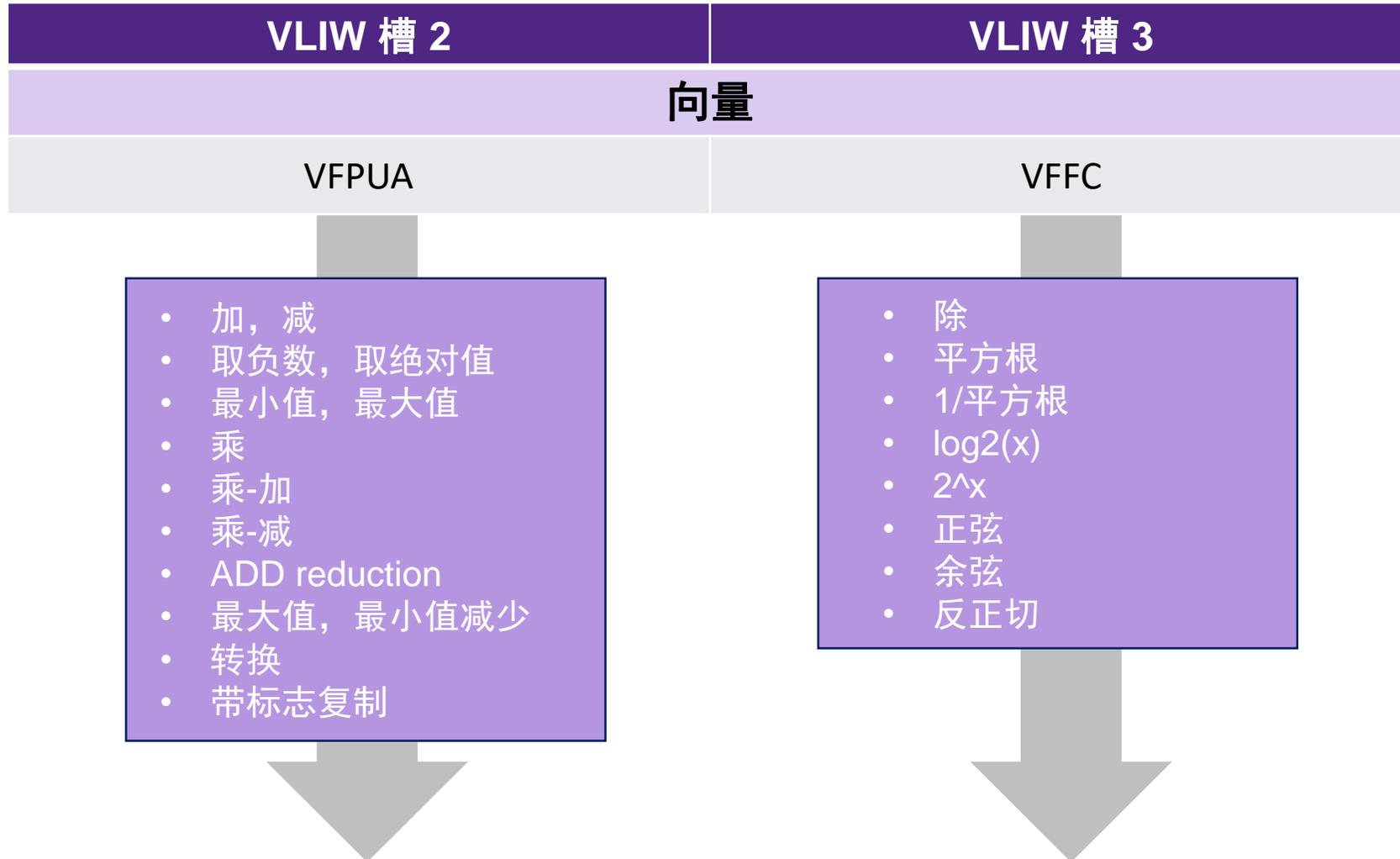
- 加，减
- 取负数，取绝对值
- 最小值，最大值
- 乘
- 乘-加
- 乘-减
- ADD reduction
- 最大值、最小值减少
- 转换
- 带标志复制

- 加，减
- 取负数，取绝对值
- 最小值，最大值
- 乘
- 乘-加
- 乘-减
- ADD reduction
- 最大值、最小值减少
- 转换
- 带标志复制

(128个半精度浮点运算)
(64个单精度浮点运算)

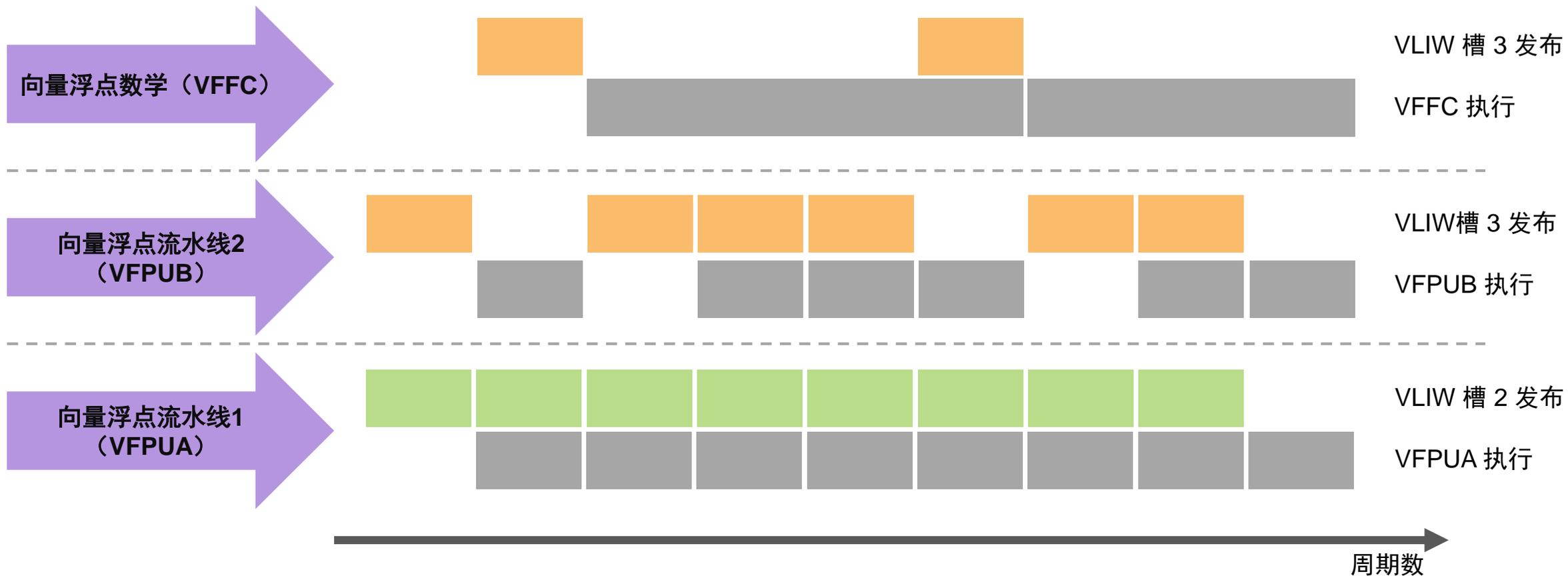
向量浮点，业界领先的性能

VFPU和VFPU数学（线性代数）的并行计算

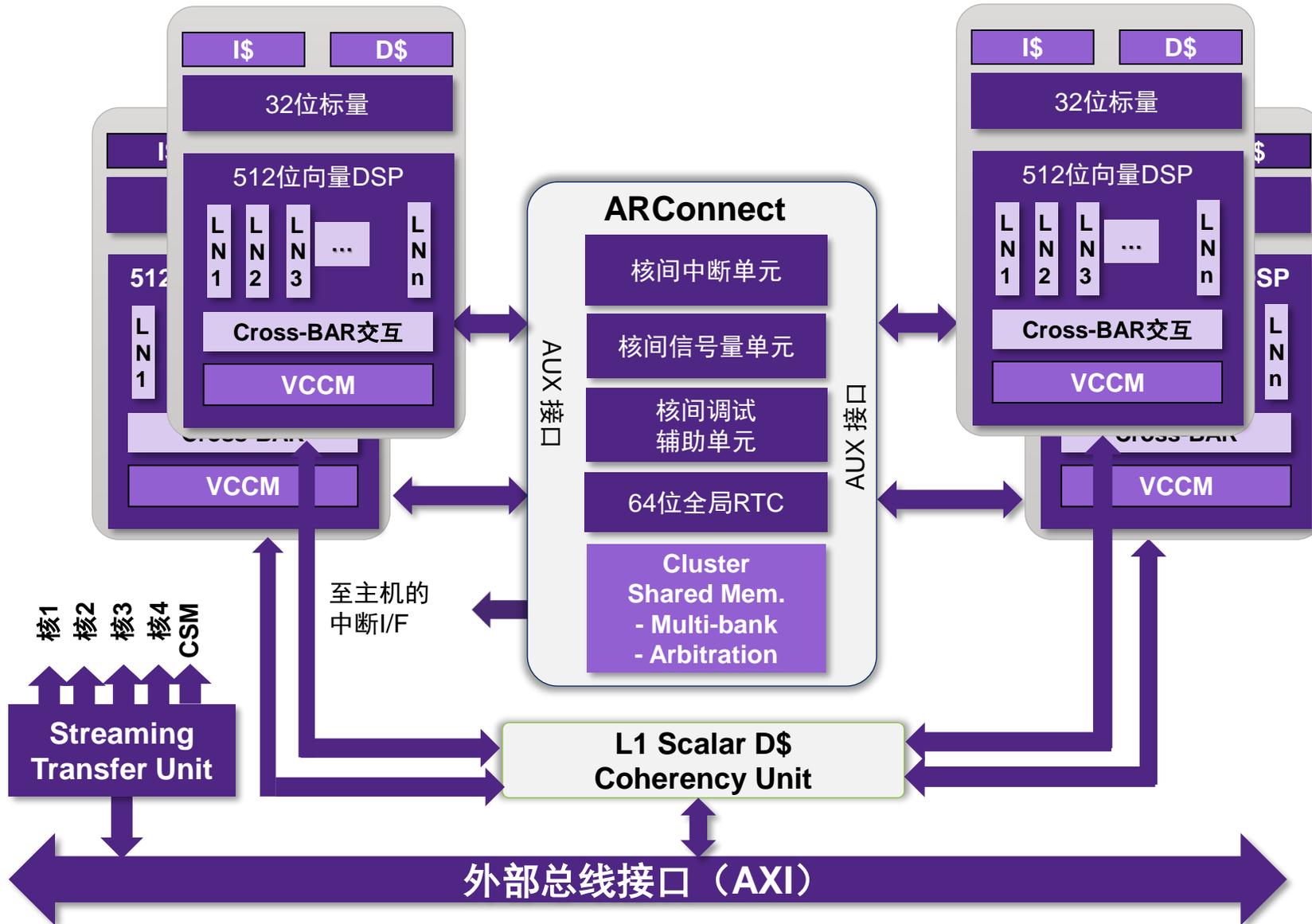


半精度和单精度

用于3个并行向量FPU执行的编译器分配



VPX5 DSP的可扩展性，可实现超高性能



- 显式管理的内存，用于高性能像素处理
- 支持高效的多核同步和数据通信
- 带有一致性的L1存储器，可实现高效的控制代码

使用ARC MetaWare工具优化的DSP软件

快速的软件开发

- 先进的自动向量化C编译器
- 高效映射到VLIW槽，以实现跨向量单元的最大化并行执行
- 完全支持单核、双核和四核配置
- 强大的DSP、线性代数和机器学习推理（MLI）软件库

综合调试器

- 广泛的分析功能，可精确定位热点，包括实时跟踪
- 由领先的第三方调试技术供应商（如Ashling、Lauterbach、Digilent）提供支持
- ARC nSIM Pro指令集模拟器

好处

- C编程模型使代码更易于更新、高度可移植
- 旨在为以最佳性能快速实现而设计
- 自动向量化将顺序代码转换为向量运算，以实现最大吞吐率
- 通过ISO 26262认证的工具可开发符合安全标准的代码

DSP库

- DSP库随MetaWare Development Toolkit一起提供
- 为所有ARC DSP内核进行了移植和优化
- 支持定点和浮点DSP函数
- 丰富的数学和矩阵计算函数

DSP库摘要

丰富的通用DSP功能库

类别	Functions
标量数学	正弦、余弦、平方根、 2^x 、log2、除、 $1/x$ 、atan2
向量数学	绝对值、加、减、乘、向量点积、取负数、偏置、缩放、移位、最小值、最大值
复杂向量数学	共轭、向量点积、幅度、幅度平方、complex_mul_complex、complex_mul_real
矩阵函数	加、减、乘、缩放、转置、倒数
IIR 滤波器	双二阶级联IIR DF1、双二阶级联IIR DF2、lattice IIR
FIR 滤波器	实数FIR、复数FIR、卷积、相关、FIR抽取器、FIR内插器、LMS
插值	线性插值、双线性插值、多项式插值
变换	CFFT、CiFFT、实数FFT、实数iFFT、DCT2、DCT3、DCT4、MDCT、iMDCT

LAPACK和特征值功能

程序类型	操作	实数	复数
线性方程组	求解线性方程组的一般系统 $AX = B$	✓	✓
标准特征值和奇异值问题	计算一般矩阵的特征值和左右特征向量	✓	✓
标准特征值和奇异值问题	计算一般矩形矩阵的奇异值分解 (SVD)	✓	✓
广义特征值和奇异值问题	计算一般矩阵的特征值以及左和/或右特征向量	✓	✓
线性方程组	一般的 $M \times N$ 矩阵的LU分解	✓	✓
线性方程组	利用因式分解求解	✓	✓
线性方程组	利用因式分解求倒数	✓	✓
线性方程组	因式分解	✓	✓
线性方程组	利用因式分解求解	✓	✓
正交分解	枢轴旋转QR分解	✓	✓
正交分解	QR分解, 无枢轴旋转	✓	✓
正交分解	生成Q	✓	✓
对称特征值问题	借助QR的特征值/特征向量	✓	✓

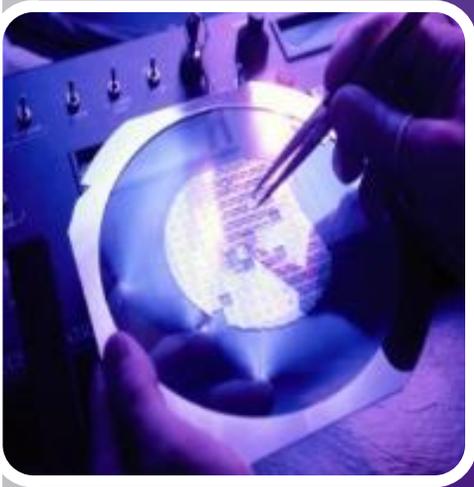
embARC MLI S/W库v.1.0

针对ARC EMxD和HS4xD也进行了优化

组	函数	简短描述
主要操作	<ul style="list-style-type: none"> • 2D 卷积 • 深度 (Depth wise) 2D卷积 • LSTM • 简单 RNN • 完全连接 	将输入要素与一组经过训练的权重进行卷积
池化 (pooling)	<ul style="list-style-type: none"> • 平均池化 • 最大池化 	用函数池化输入特征
变换/激活函数	<ul style="list-style-type: none"> • ReLU • SoftMax • Leaky ReLU • Sigmoid • TanH 	根据特定函数变换输入数据的每个元素
数据路由操作	<ul style="list-style-type: none"> • 填充 (Padding) • 转置 (Transpose) • 级联 (Concatenation) 	按指定模式移动输入数据
元素操作	<ul style="list-style-type: none"> • 向量算术运算 	将多操作数函数逐元素应用到多个输入

- 针对ARC 各种 DSP内核上ML推理的软件库
 - 内核函数库：有助于中小型机器学习模型的推理计算
- 支持轻松实现分层NN图拓扑
 - 为实现多种NN层类型而优化的库
 - 高效、小体积内核实现
 - C风格的 API
- 使用模型：
 - 短期：用户可调用（手动图形映射）
 - 长期：自动图形映射（Caffe、TensorFlow Lite 等）
- embARC.org上的开源资源

VPX，面向以数据为中心的世界的下一代DSP架构



- 浮点应用的强烈增长趋势
- 向量浮点计算的迅速增加
- VPX5 DSP 提供领先水平的浮点计算
 - 宽SIMD 向量和 VLIW并行执行
 - 数学代数专用的向量 FPU 单元
 - 多核产品可提供超级性能
- 成熟的软件工具链和开发环境
- 完备的DSP、线性代数和机器学习库

Thank You

