

ASIP Designerによる RISC-VベースのAIアクセラレータ設計の実例

RISC-V Day Tokyo 2021

日本シノプシス合同会社
シニア・アプリケーションエンジニア
伴野 充

2021/4/22, 23



Synopsys Today: シリコンからソフトウェアへ



FY20 売り上げ:
~\$3.7B



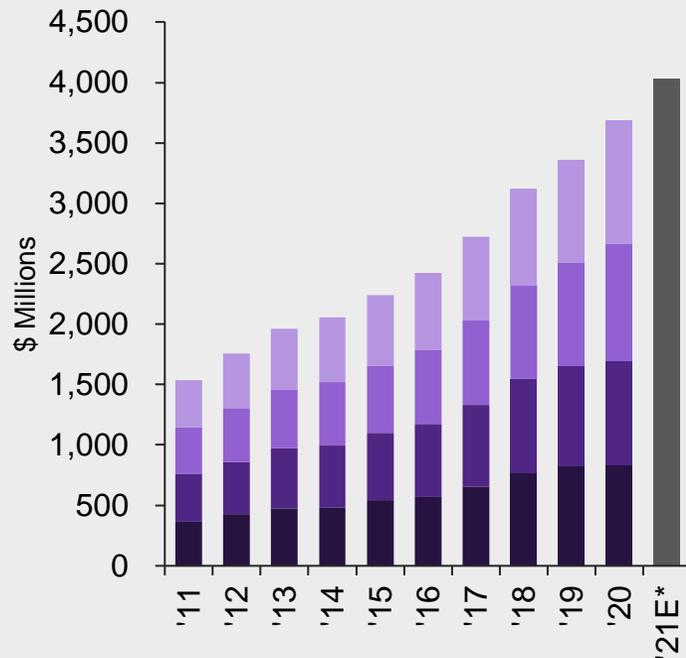
従業員数:
~15,000



取得特許:
~3,300



オフィス拠点:
~125



業界ナンバーワンの (EDA)電子設計
自動化ツールとサービス

幅広いIPポートフォリオと#1の
インターフェースIP、アナログ、
組込みメモリ、物理IPおよびプロセッサ

ガードナーのアプリケーションセキュリ
ティ・テスト向け“Magic Quadrant” の
リーダー

ドメインに特化したプロセッサの時代

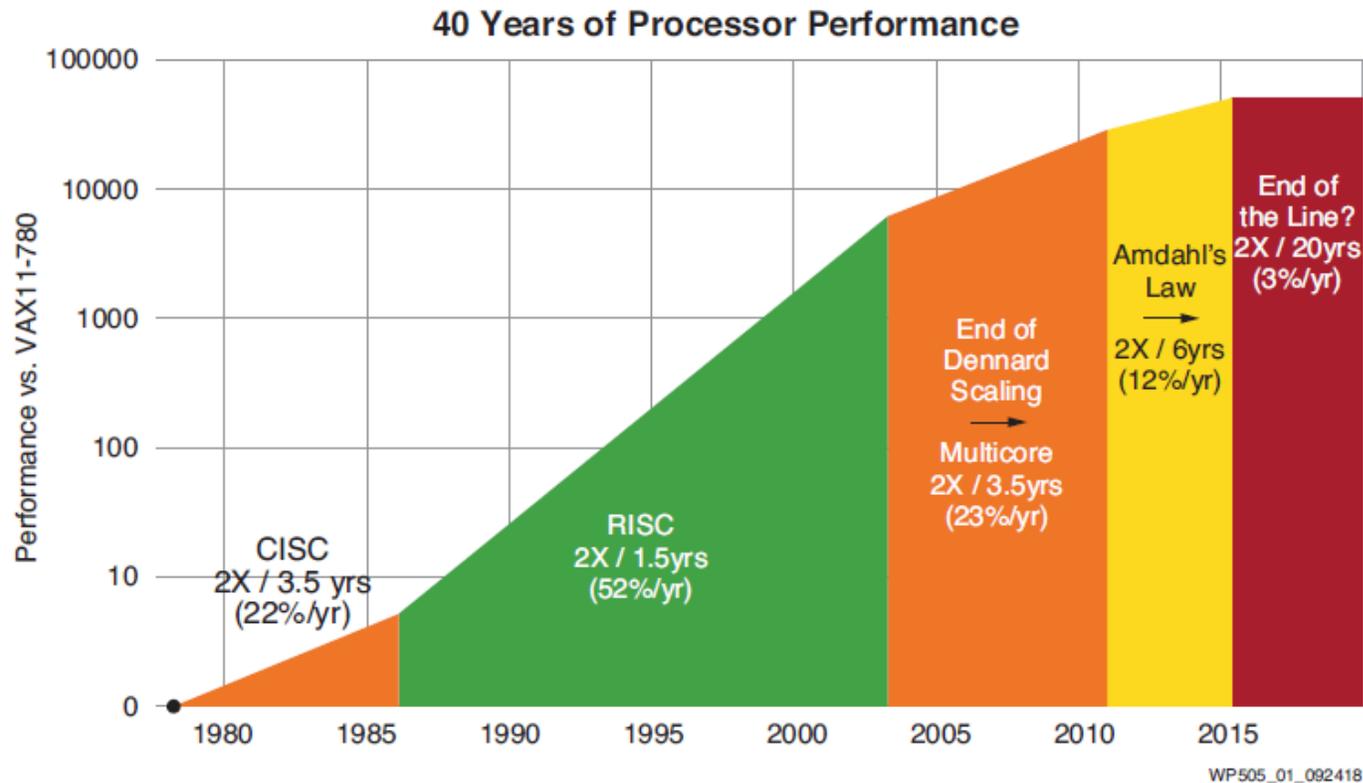


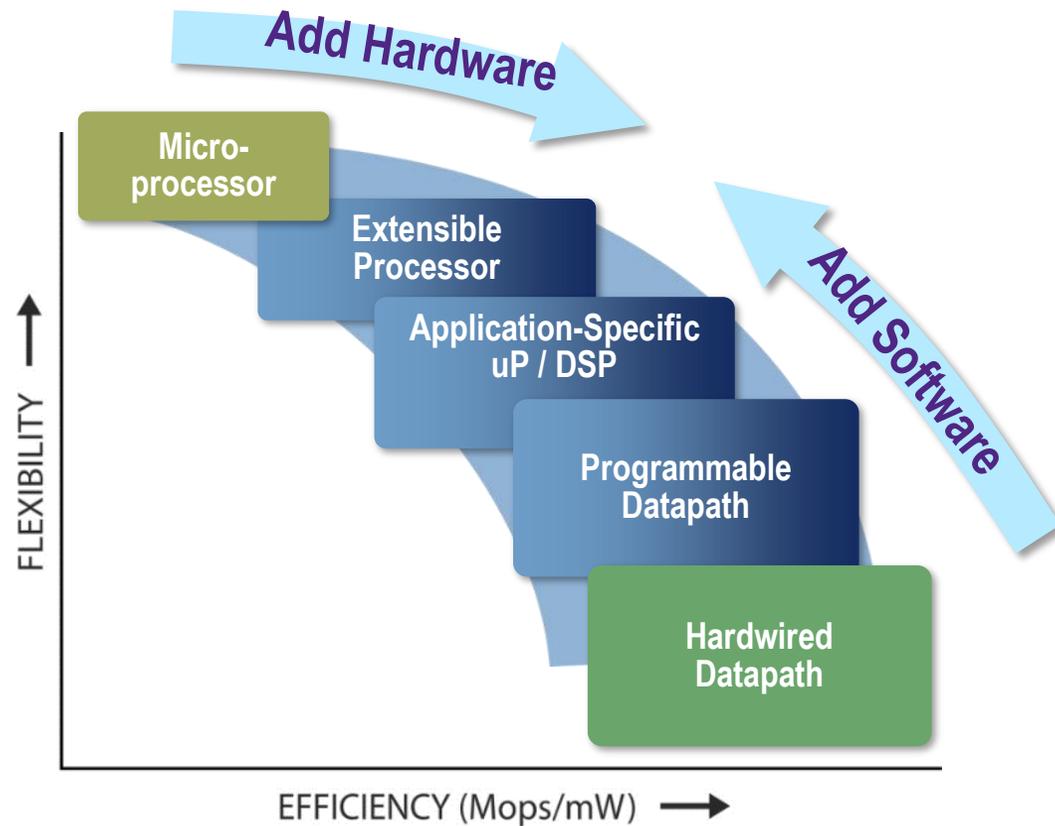
Figure 1: Processor Performance vs. Time

Source: Xilinx Whitepaper 505 – Versal ACAP
(reference to J. Hennessy, D. Patterson, *Computer Architecture: A Quantitative Approach* (6th Edition, 2019))

- 1995 – 2003:
プロセス技術による性能(より高い周波数だが同じ電力消費)およびメモリアクセスの最適化(Dennard Scaling)
- 2003: Multicore
タスクの並列化制限に到達(Amdahlの法則)
- 2015: Domain-Specific Processors
ヘテロジニアス・マルチコア設計への移行

Application-Specific Instruction Set Processors (ASIP)

ハードウェアとソフトウェアの融合



Application-specific instruction set processor

From Wikipedia, the free encyclopedia

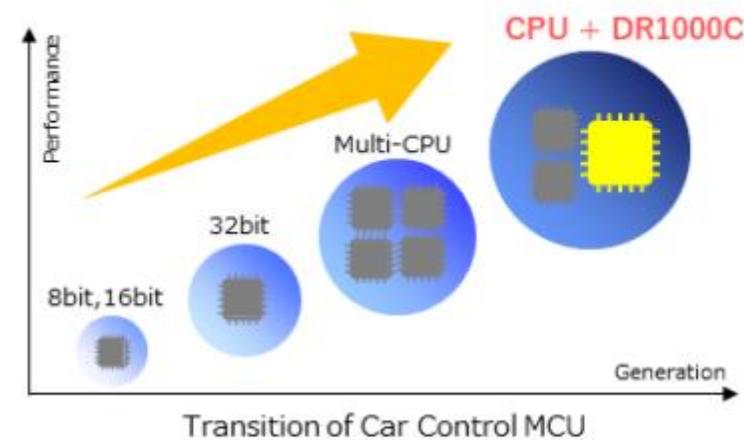
An **application-specific instruction set processor (ASIP)** is a component used in **system-on-a-chip** design. The **instruction set** of an ASIP is tailored to benefit a specific application. This specialization of the core provides a tradeoff between the flexibility of a general purpose CPU and the performance of an ASIC.

2020年9月17日発表：車載用データフロー・プロセッサIP



The screenshot shows the Synopsys website's news release page. The header includes the Synopsys logo and navigation links for '会社' (Company), '投資家情報' (Investor Information), 'コミュニティ' (Community), 'ニュースルーム' (Newsroom), 'リソース' (Resources), and '採用情報' (Careers). The main content area features a breadcrumb trail: 'Home / Japan / ニュースリリース / ニュースリリース - 2020年9月16日'. The title is 'ニュースリリース - 2020年9月16日'. The main text describes the release of the ASIP Designer tool, highlighting its ability to reduce development time for custom processors. A '概要' (Summary) section lists three key points: 1) The tool enables the development of 5 types of custom processors with high computing performance. 2) It allows for the automatic generation of software development kits, enabling development by a limited number of staff. 3) The tool provides SystemC interfaces for processors, simulation, and model export, allowing for the development of a virtual prototype based on a Data Flow Processor (DFP) core.

NSITEXE DFP™ DR1000C



この度発売する「DR1000C」は、RISC-V*1ベースのプロセッサとして世界初となる自動車向け機能安全規格 ISO 26262の安全要求レベルASIL Dに対応します。マルチスレッド機構とベクトル命令*2により、将来の車両制御向けマイコンに要求される高負荷演算処理をオフロードするのに最適な並列プロセッサとなります。車両制御向けマイコンは「DR1000C」を搭載することで、モデル予測制御などの先進的な制御アルゴリズムに対応し、強化される将来の法規制にも対応することが可能となります。

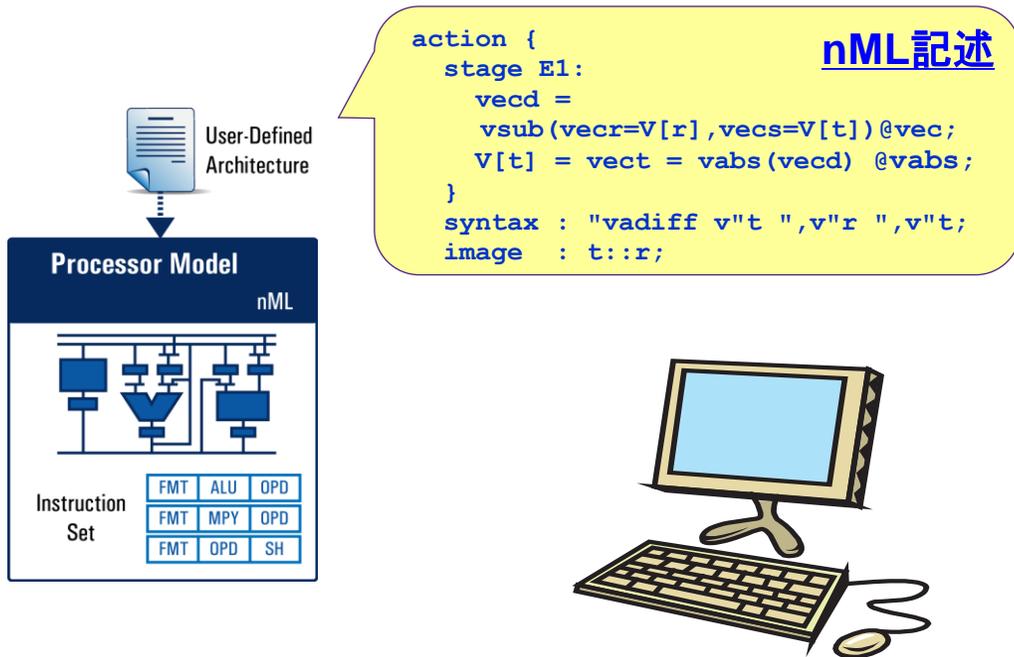
出典: NSITEXE社HP

ASIP Designer

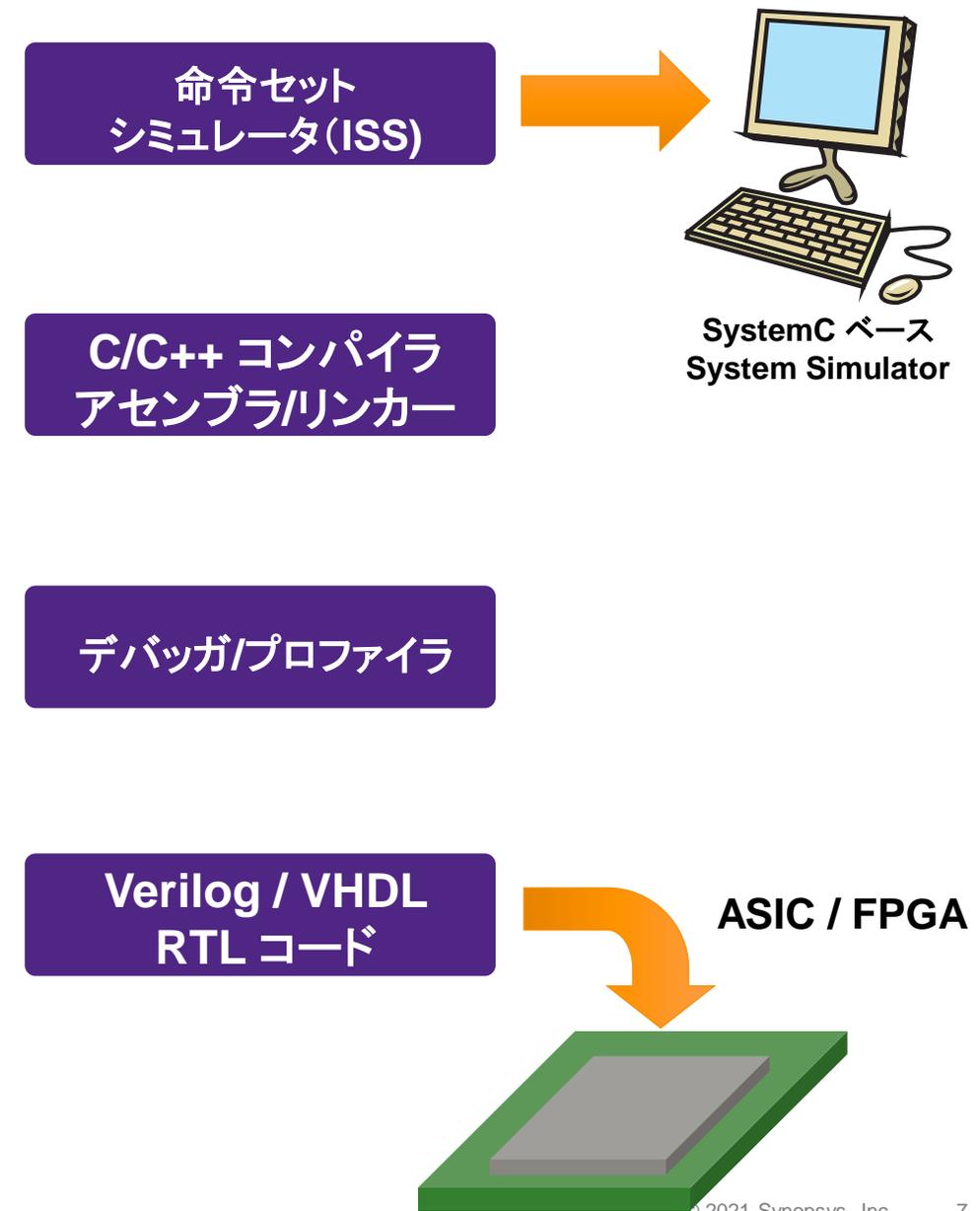
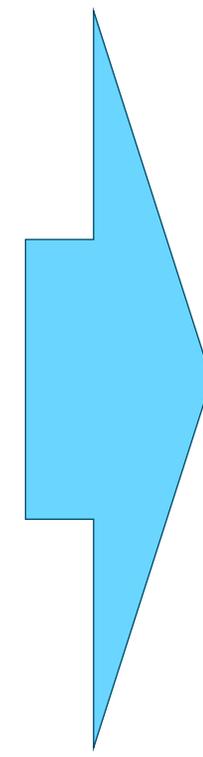


ASIP Designer

アーキテクチャ構造記述言語(nML)により、
アプリケーションに特化した命令セットを持つ
専用プロセッサ/DSPを自動生成する設計環境



ASIP設計開発ツール
"ASIP Designer"



ご提供可能なExampleプロセッサ

種類	説明
マイクロコントローラ	
Tnano	16-bit microcontroller, lightweight and configurable
Tmicro	16-bit microcontroller, fully featured
DLX (family)	Variants of 32-bit microcontroller
Tmcu	32-bit microcontroller
Trv32 (family)	32-bit microcontroller featuring RISC-V ISA, 3 or 5 pipeline stages
Trv64 (family)	64-bit microcontroller featuring RISC-V ISA, 3 or 5 pipeline stages
PD_Triop	32-bit microcontroller with 64-bit address spaces
DSP及び並列演算エンジン	
Tdsp	16/32-bit DSP
Tvec (family)	Variants of SIMD processor
Tvliw (family)	Variants of VLIW processor
暗号処理エンジン	
Tsec	High throughput SHA/RSA/AES accelerator
Tcript	High throughput AES accelerator
アプリケーション特化	
Tmoby	Accelerator for AI application, featuring MobileNet V3 graph
Tvox	Accelerator for SLAM (simultaneous localization and mapping)
MMSE	Minimum Mean Square Error Equalization
Tgauss	Vectorization and memory management for image processing
Tmotion	Accelerator of motion estimation kernel
Tcom8	SIMD processor optimized for some communication kernels
FFTcore	Scalar implementation of complex FFT

green =
released in 2020

nML記述
ソースコード
ご提供

豊富なExample

- マイクロコントローラ
- DSPs
- SIMD、VLIW
- マルチスレッド
- セキュリティ(AES、SHA等)
- メモリI/F、AXIバス等

Example使用のメリット

- 基本的な機能は既に設計開発済み
- 動作可能な例を参照可能
- モデリングの概念を効率的に習得
- nMLソースコードで提供されるため、ユーザーが自由に追加変更可能

Trv(RISC-V ISA)モデル

内容

- RISC-V ISAに最適化されたモデル
- ISA拡張のスターティングポイントとしての活用 (ASIP設計)

	32-bit データパス	64-bit データパス
3ステージ パイプライン	Trv32p3 Trv32p3x	Trv64p3 Trv64p3x
5ステージ パイプライン	Trv32p5 Trv32p5x	Trv64p5 Trv64p5x

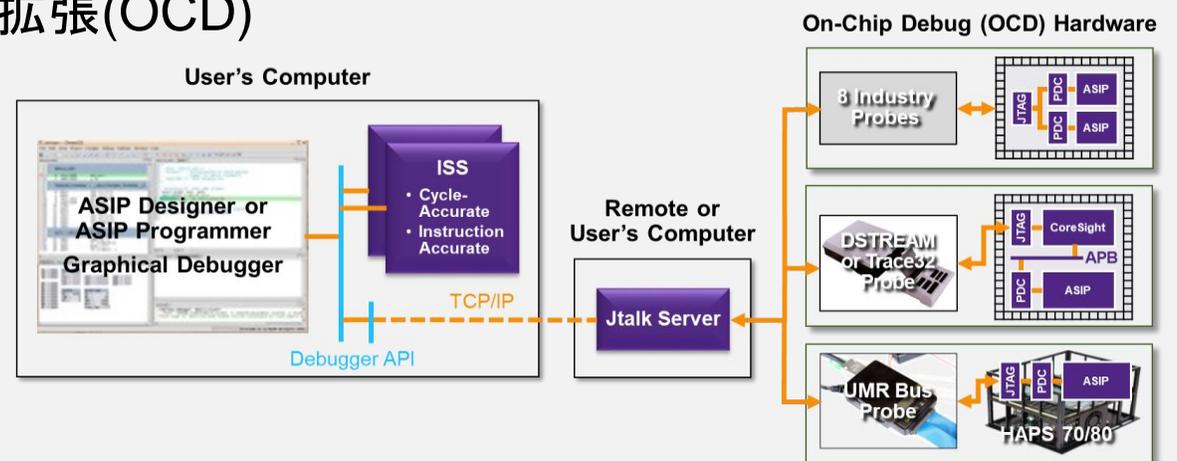
機能 (2020.03)

- 対応ISA: RV64IM, RV32IM
 - 整数命令 (Integer)
 - 乗算命令 (Multiply)
- オプション拡張: Trv<mm>p<n>x
 - (標準) 圧縮命令 (Compressed: IMC)
 - (カスタム) ゼロオーバーヘッドループ
 - (カスタム) アドレス更新後のロード/ストア
- マイクロアーキテクチャの特長
 - 5段ステージパイプライン: IF, ID, EX, ME, WB
 - 3段ステージパイプライン: IF, ID, EX
 - 保護されたパイプライン (Protected pipeline)
 - 可能な場合はレジスタをバイパス
 - バイパスが出来ない時はStall (bubble)
 - ハードウェア乗算器
 - 64x64 → 128 → select 64bit,
 - 32x32 → 64 → select 32bit
 - 繰り返し処理の除算器

Trvモデル拡張 (2021.03)

F-クラスサポート, オンチップデバッグ

- Trv32モデルにオプションのF命令クラスを拡張
F命令クラス
 - IEEE754 単精度浮動小数点のサポート
(round to nearest-even mode)
 - レジスタファイル
 - FADD, FSUB, FMUL, FCMP (EQ/LT/LE), FMIN, FMAX, FCVT → DesignWareモジュール
 - FDIV, FSQRT → 繰り返し処理の実装
- Trvモデルにオプションのオンチップデバッグを拡張(OCD)
 - ASIP DesignerのOCD自動生成機能を使用
 - 自動的にISAの拡張に適応



SDX: シンプルデータパス拡張 (Simple Datapath eXtensions)

コンセプト

- SDXは(RISC-V)プロセッサモデルに単純な拡張命令を追加するメカニズム
 - 拡張命令スケルトンは、プロセッサモデルの一部として提供
 - ユーザーは特別なnML記述を学ぶ必要がない
 - ユーザーが拡張機能の動作をPDGとしてコーディング
 - PDGは、データパス関数の動作を定義するために使用されるCのような言語
 - 拡張命令をターゲットとするコンパイラ組み込み関数を提供
- RISC-Vの“custom-2”命令としてエンコード
- 命令バリエーション
 - 3つのreg 32ビットおよび64ビット、累積、追加のシングルレジスタ入力および出力等々, ...
- 例: SHA256 (オペレーション融合), FFT (コンプレックス処理), CNN (パッケージ化されたSIMD)



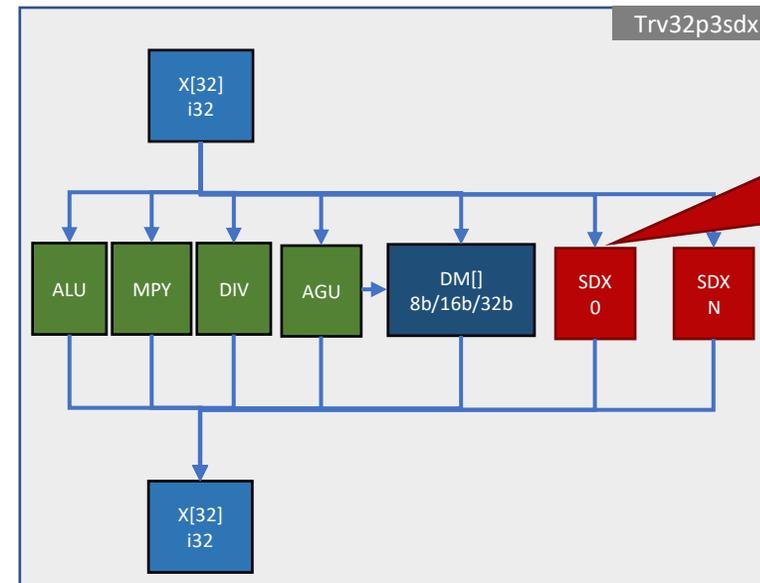
Behavioral model (PDG)

```
w32 sdx0(w32 a, w32 b)
{
  w32 r;
  r[15:0] = a[15:0] + b[15:0];
  r[31:16] = a[31:16] + b[31:16];
  return r;
}
```

Compiler intrinsics

```
int sdx0(int,int);
int add2(int,int);
```

User can assign an intuitive name



SDX Examples

ASIP Designerに付属

FFT

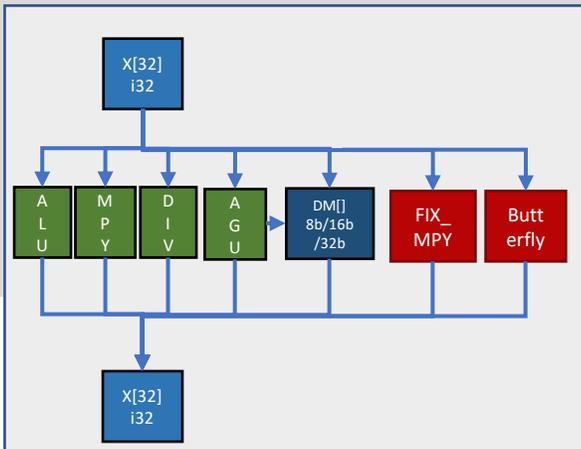
SDX instructions accelerating

- Complex fixed-point multiplication & scaling
`sdx1 rd,rs1,rs2`
- ABS(x) function: `sdx2 rd,rs1,rs2(x0)`
- FFT Butterfly: `sdx5 rd,rs1,rs1`

Specialization:

- Fractional data types
- Complex numbers (16bit/16bit -> 32bit register)

スピードアップ: 280% エリア増: 31%



SHA256

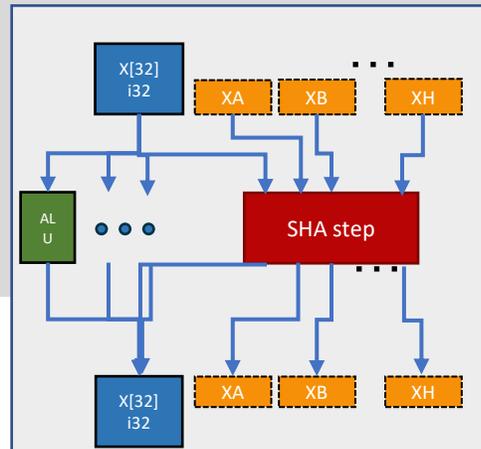
Computes a hash of message W using bitwise AND, OR, XOR operations, shift operations and additions

Custom data path is ideal to implement the complex hash function in one instruction

Additional state of the hash functions (8 state variables) require an SDX variant that supports **8 additional register reads and writes**

```
sdx7 rd,rs1,rs2,x24,x25,...,x32
```

スピードアップ: 270% エリア増: 16%



Keyword Spotting

Based on small sized Neural Network (3.3M MACs)

SDX architecture feature: **packed SIMD**

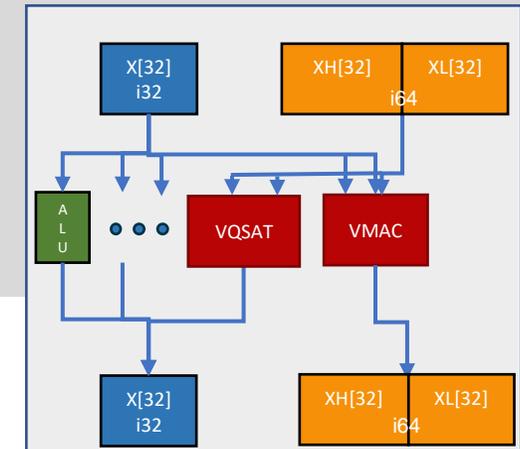
32-bit register contains vector of 4x 8-bit values



Use of register pairs, enabling 64-bit access

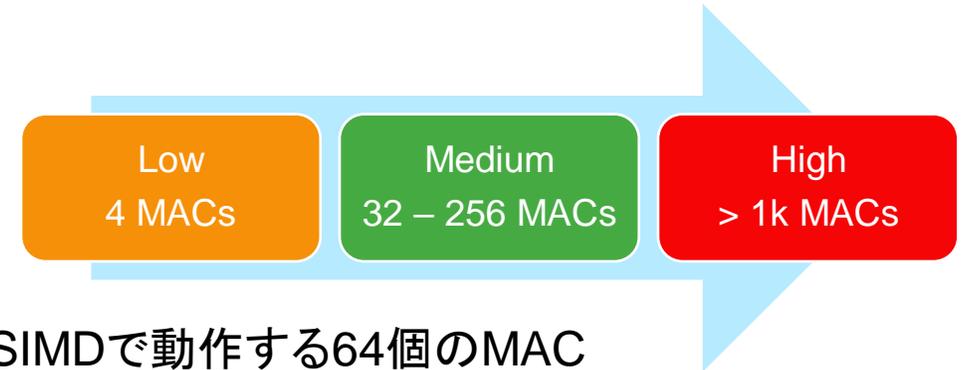
```
sdx4a_dr dd,rs1,rs2,mode // vmac
sdx0_dd rd,ds1,ds2 // vqsat
```

スピードアップ: 1160% エリア増: 16%

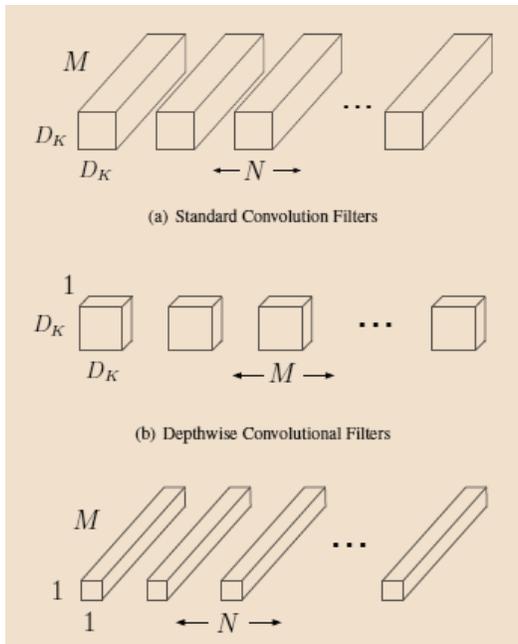


Tmoby: MobileNet向けASIP

- MobileNetはモバイルや組み込みビジョンアプリケーションなどで用いられるCNN
 - 深さ方向に分離可能な畳み込み演算が基本
- Tmobyは中速のスループット要件に最適化



- SIMDで動作する64個のMAC
- 並列ロード/ストアと同時に次アドレスを計算
- ローカルメモリ/バックグラウンドメモリ+ DMA



• 演算のコスト

- 標準的な3D畳み込み処理

$$D_K \cdot D_K \cdot M \cdot N \cdot D_F \cdot D_F$$

- 深さ方向に分離可能な畳み込み処理

$$D_K \cdot D_K \cdot M \cdot D_F \cdot D_F \\ + M \cdot N \cdot D_F \cdot D_F$$

Tmoby アーキテクチャ

アーキテクチャの特長

- ISA: 4並列命令
- ベクタデータパス
 - SIMD64
 - MAC 8x8→32
- メモリ
 - VM: フューチャー
 - WM: ウェイト
 - ベクタアドレッシング

アプリケーション: MobileNet V3

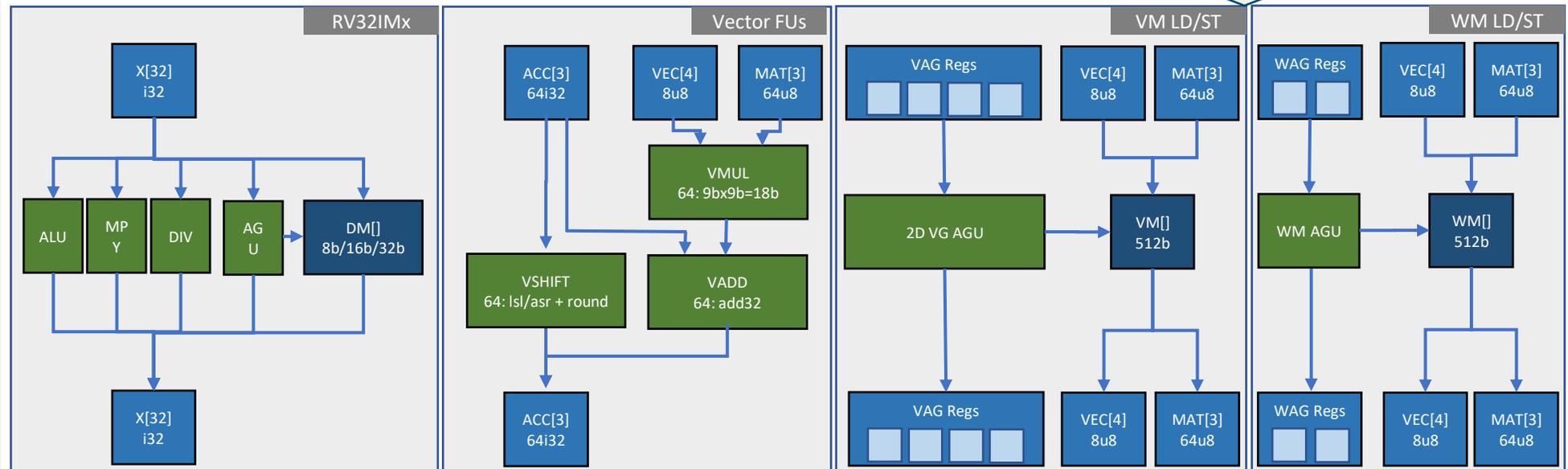
- TensorFlowコードを使用
- Tmoby固有のベクトル組込み関数を用いてCコードに変換

主な特長

- Graphの実装
- メモリコピー
- カーネル

カーネル

- 2D畳込み (point-wise)
- 2DのDepth-Wise
- 加算
- 2Dのアベレージプーリング
- ソフトマックス



まとめ



ASIP Designer — まとめ

- 差別化実現の鍵 — 独自プロセッサ/DSP設計開発
 - nML言語によるトップダウン設計手法で効率的なASIP開発を実現
 - 複数アルゴリズムや機能をASIPに統合することでゲート数を削減
 - 必要最低限な構成と命令で超低消費電力プロセッサを実現
- ASIP Designer(AD)によるASIP開発のメリット
 - 差別化技術を含むロイヤリティフリーなRISC-Vプロセッサを実現
 - 設計開発後期のアルゴリズム変更など急な仕様変更への対応
 - 製品出荷後でもプログラムによるアップデートが可能
 - 言語設計の為、次世代の機種展開が容易(高い設計流用性)
 - 汎用DSP/プロセッサベンダによる生産中止トラブルを回避

Thank You

