

# PEZY-SC3

高い電力効率を実現するMIMDメニーコアプロセッサ

初田 直也<sup>1</sup> 角田俊太郎<sup>1</sup> 内田広平<sup>1</sup>

石谷太一<sup>1</sup> 塩谷亮太<sup>2</sup> 石井敬<sup>1</sup>

<sup>1</sup>PEZY Computing <sup>2</sup>東京大学

# 目次

- ▶ PEZY-SCxのアーキテクチャ
  - ▶ PEZY-SCxとは
  - ▶ MIMDメニーコア
  - ▶ 細粒度マルチスレッディング
- ▶ PEZY-SC3の実装
  - ▶ LSI実装
  - ▶ 基板・システム実装
- ▶ アプリケーションでの評価
  - ▶ LINPACKベンチマーク
  - ▶ ヒトゲノム解析
- ▶ まとめ

# 目次

- ▶ PEZY-SCxのアーキテクチャ
  - ▶ PEZY-SCxとは
  - ▶ MIMDメニーコア
  - ▶ 細粒度マルチスレッディング
- ▶ PEZY-SC3の実装
  - ▶ LSI実装
  - ▶ 基板・システム実装
- ▶ アプリケーションでの評価
  - ▶ LINPACKベンチマーク
  - ▶ ヒトゲノム解析
- ▶ まとめ

# PEZY-SCxとは

- ▶ PEZY Computingで開発しているスーパーコンピュータ向けプロセッサ
  - ▶ スレッドレベル並列性の高いアプリケーションがターゲット
- ▶ PEZY-SCからPEZY-SC3sまで4世代をリリース

|                  | PEZY-SC       | PEZY-SC2      | PEZY-SC3      | PEZY-SC3s    |
|------------------|---------------|---------------|---------------|--------------|
| プロセッサコア          | 1024 @ 733MHz | 2048 @ 1.0GHz | 4096 @ 1.2GHz | 512 @ 1.0GHz |
| 倍精度演算性能 [TFlops] | 0.75          | 4.1           | 19.7          | 2.0          |
| 外部メモリ帯域 [GB/s]   | 153.6 (DDR4)  | 102.4 (DDR4)  | 1228.8 (HBM2) | 614.4 (HBM2) |
| 製造プロセス           | TSMC 28nm     | TSMC 16nm     | TSMC 7nm      | TSMC 7nm     |
| リリース             | 2014/08       | 2017/04       | 2021/04       | 2022/04      |

# PEZY-SCxの特徴

- ▶ プロセッサのアーキテクチャ
  - ▶ MIMDメニーコア
- ▶ コアのマイクロアーキテクチャ
  - ▶ 細粒度マルチスレッディング
- ▶ 命令セット
  - ▶ RISC系の独自命令セット

# PEZY-SCxの特徴

- ▶ プロセッサのアーキテクチャ
  - ▶ MIMDメニーコア
- ▶ コアのマイクロアーキテクチャ
  - ▶ 細粒度マルチスレッディング
- ▶ 命令セット
  - ▶ RISC系の独自命令セット

# MIMDメニーコア

- ▶ 階層構成
  - ▶ 多数のコアをどのように構造化するか
- ▶ メニーコアの同期制御
  - ▶ 各コアは独立に動作する
  - ▶ 何らかの同期機構が必要

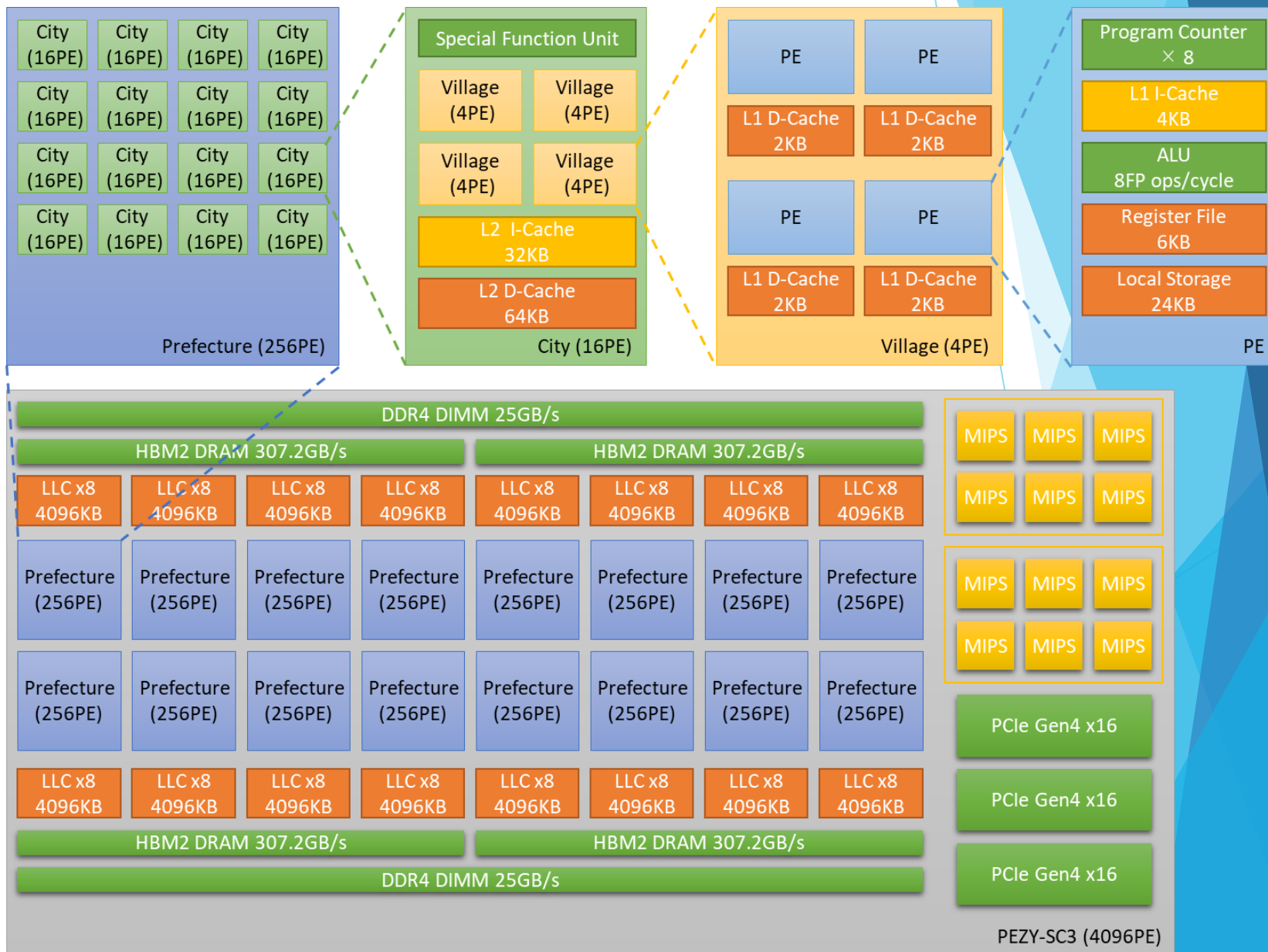
# 階層構成

## ▶ 5階層の階層構造

- ▶ PE (プロセッサコア)
- ▶ Village
- ▶ City
- ▶ Prefecture
- ▶ State (チップ全体)

## ▶ キャッシュ階層に対応

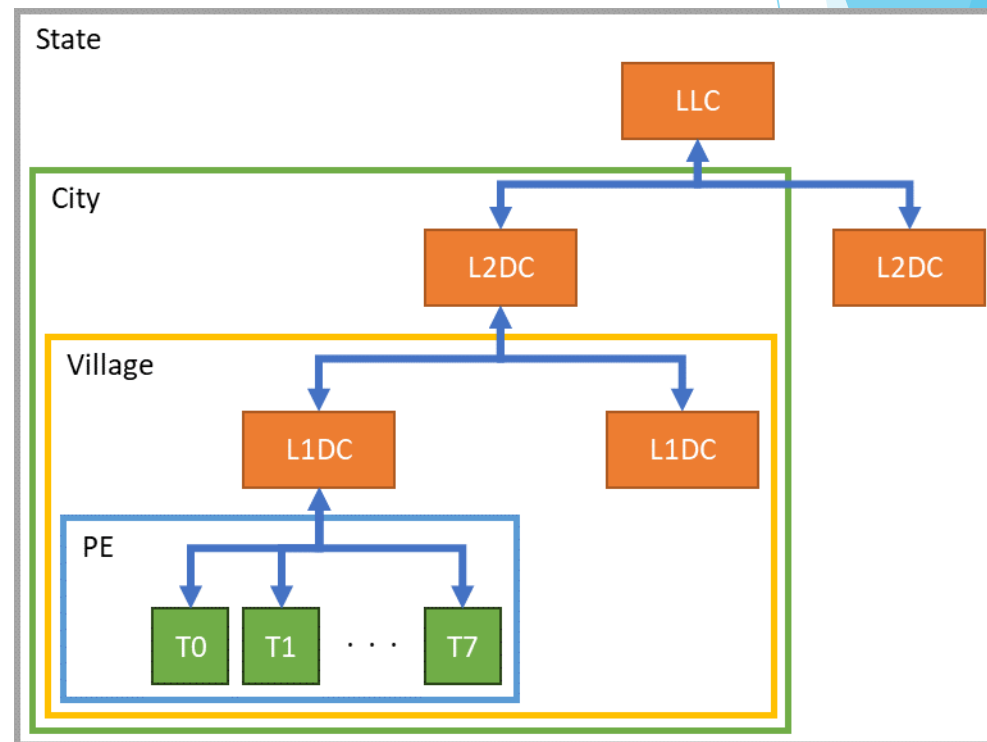
- ▶ Village: L1
- ▶ City : L2
- ▶ State : LLC





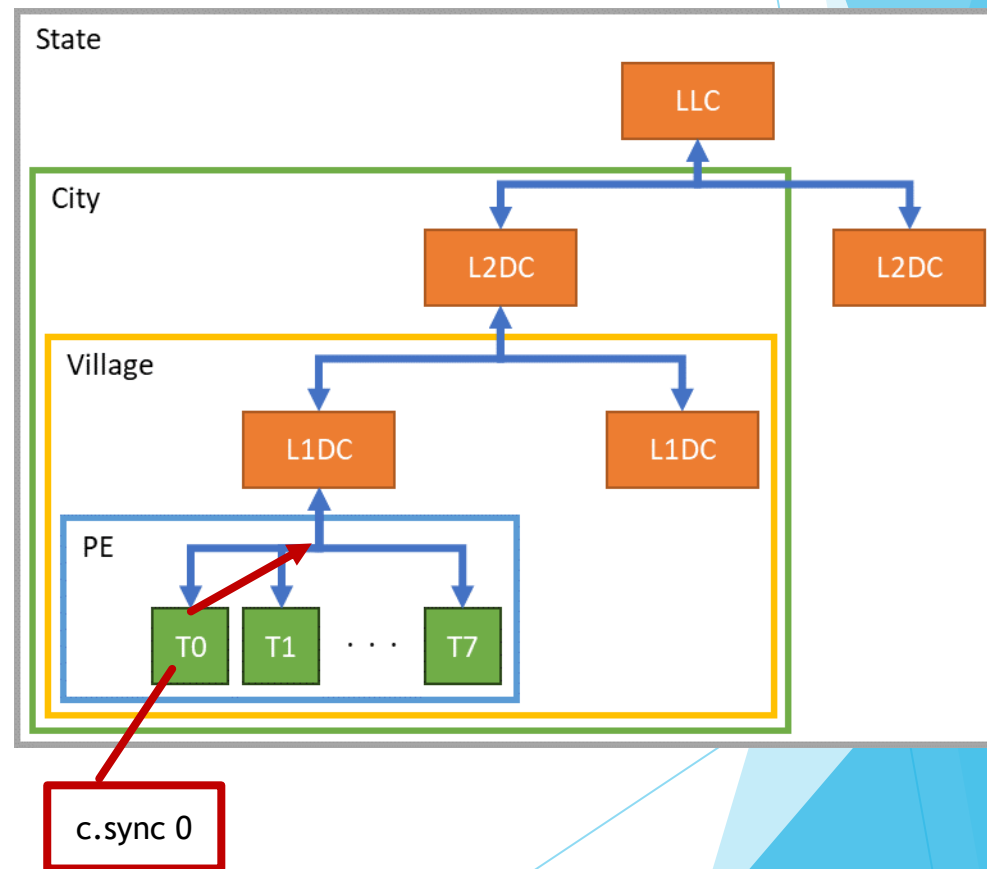
# 同期制御

- ▶ 同期専用のツリー型ネットワーク
  - ▶ メモリバスとは独立した専用ネットワーク
    - ▶ 隣接コア間で高速に同期を取れるよう最適化
  - ▶ 専用命令で同期範囲を指定



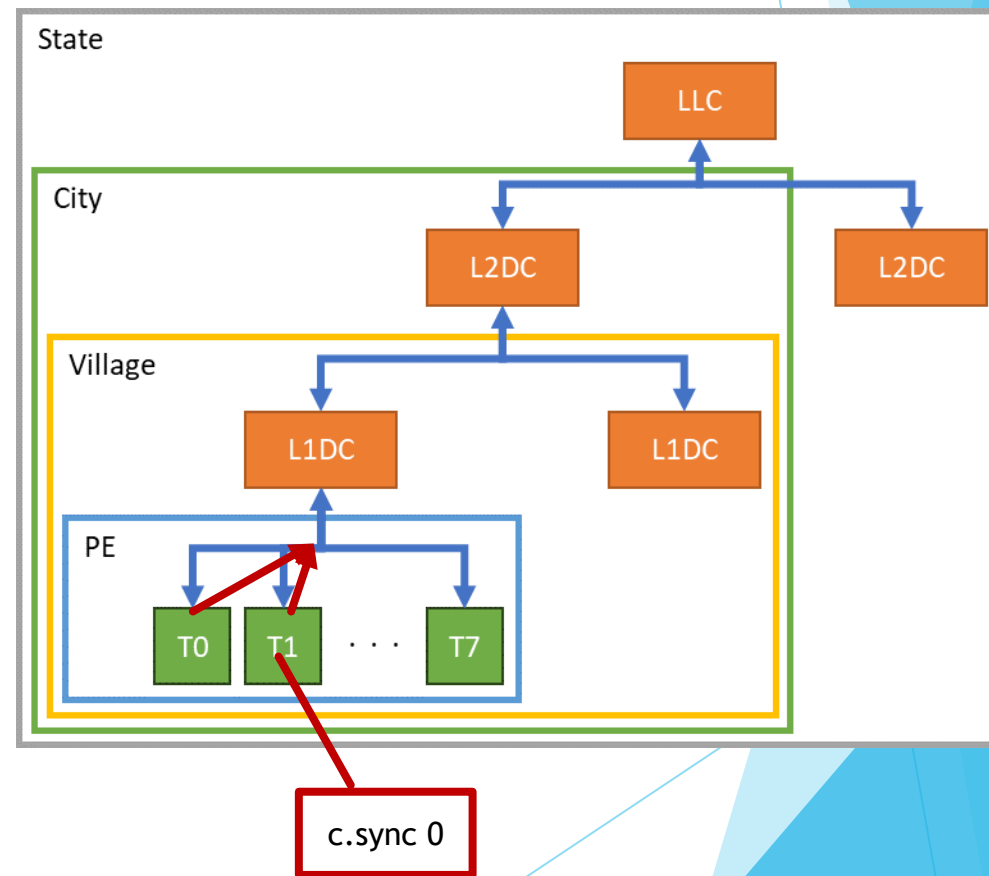
# 同期制御

- ▶ 同期専用のツリー型ネットワーク
  - ▶ メモリバスとは独立した専用ネットワーク
    - ▶ 隣接コア間で高速に同期を取れるよう最適化
  - ▶ 専用命令で同期範囲を指定



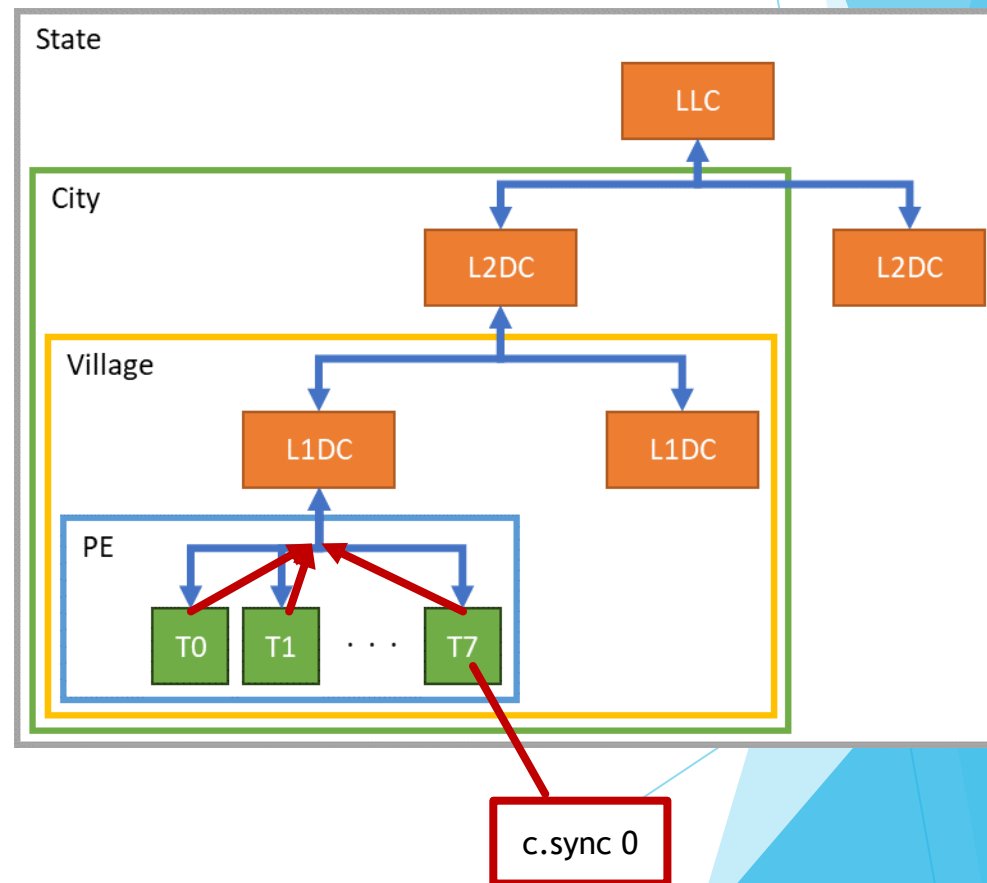
# 同期制御

- ▶ 同期専用のツリー型ネットワーク
  - ▶ メモリバスとは独立した専用ネットワーク
    - ▶ 隣接コア間で高速に同期を取れるよう最適化
  - ▶ 専用命令で同期範囲を指定



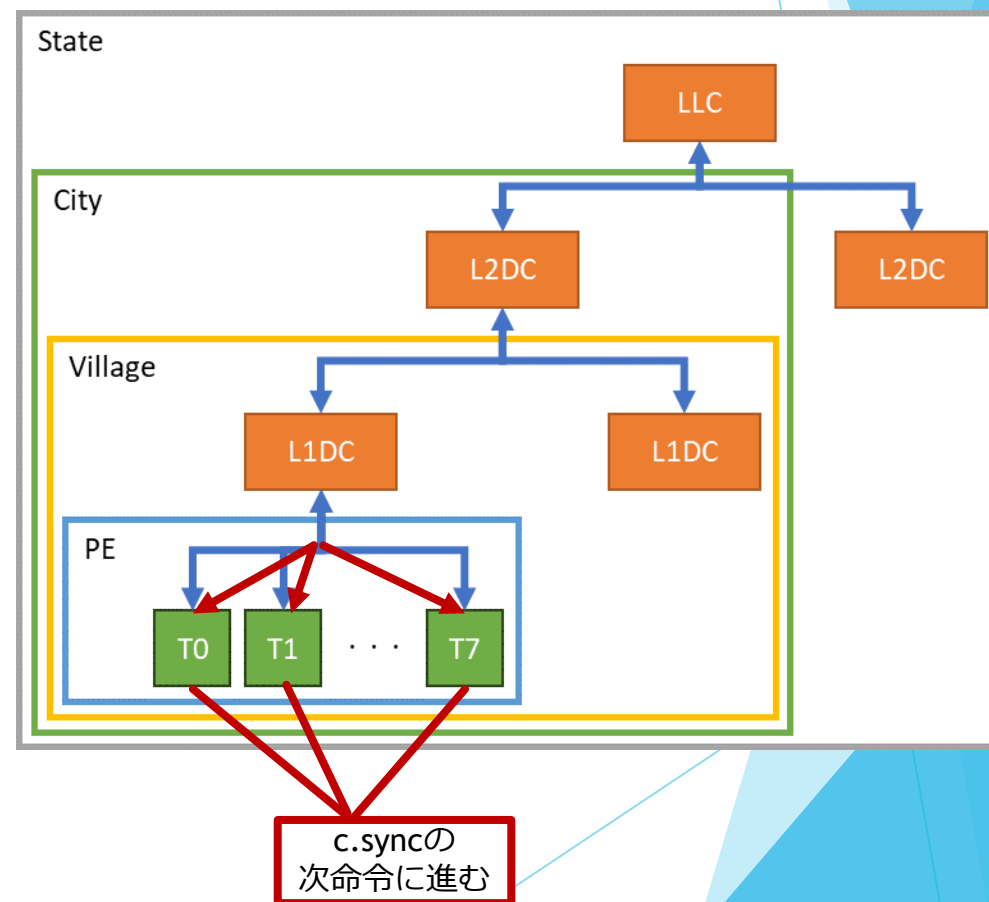
# 同期制御

- ▶ 同期専用のツリー型ネットワーク
  - ▶ メモリバスとは独立した専用ネットワーク
    - ▶ 隣接コア間で高速に同期を取れるよう最適化
  - ▶ 専用命令で同期範囲を指定



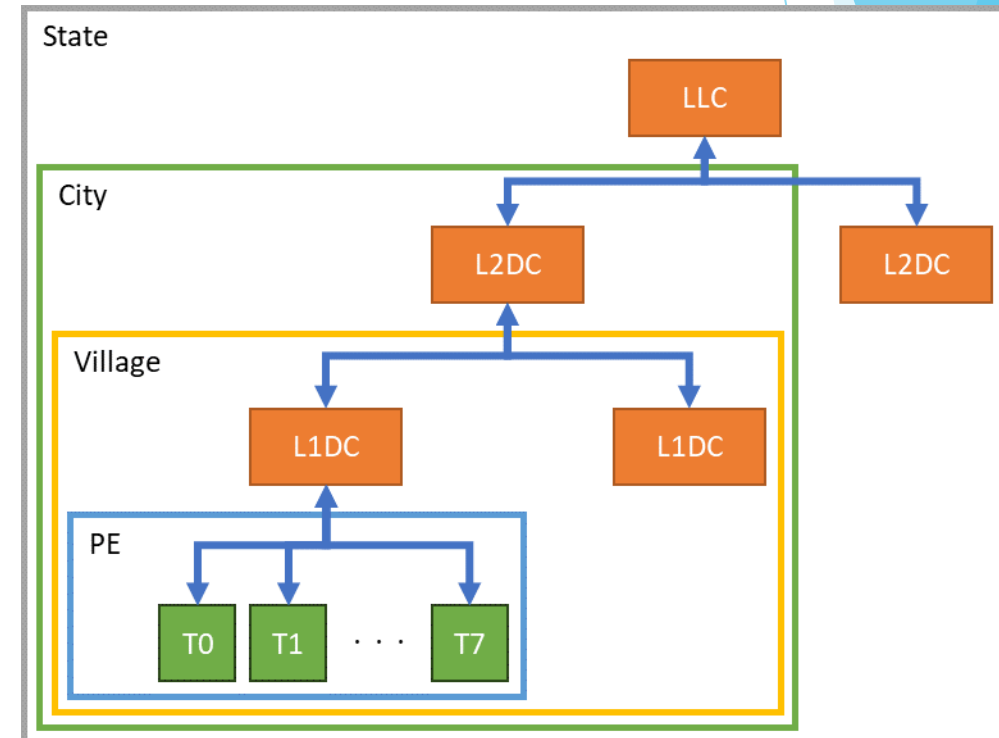
# 同期制御

- ▶ 同期専用のツリー型ネットワーク
  - ▶ メモリバスとは独立した専用ネットワーク
    - ▶ 隣接コア間で高速に同期を取れるよう最適化
  - ▶ 専用命令で同期範囲を指定



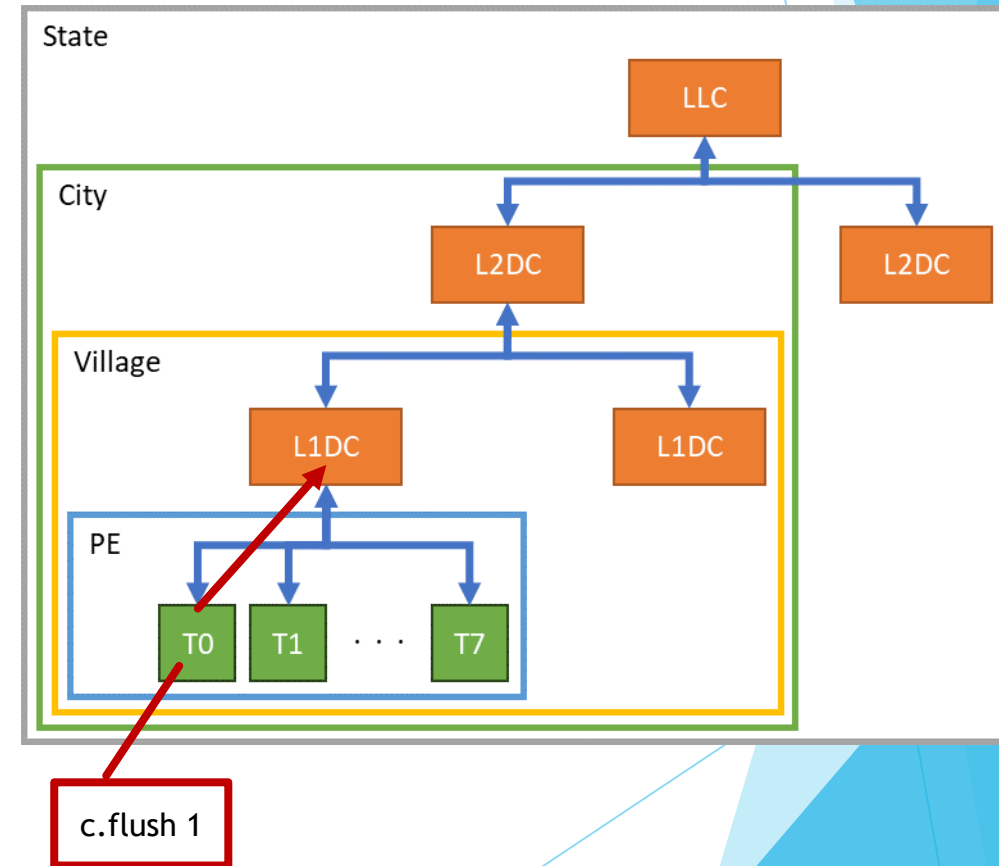
# 同期制御

- ▶ ノンコヒーレントキャッシュ
  - ▶ ソフトウェアによる明示的なキャッシュ同期
  - ▶ 同期ネットワークの1階層として実装



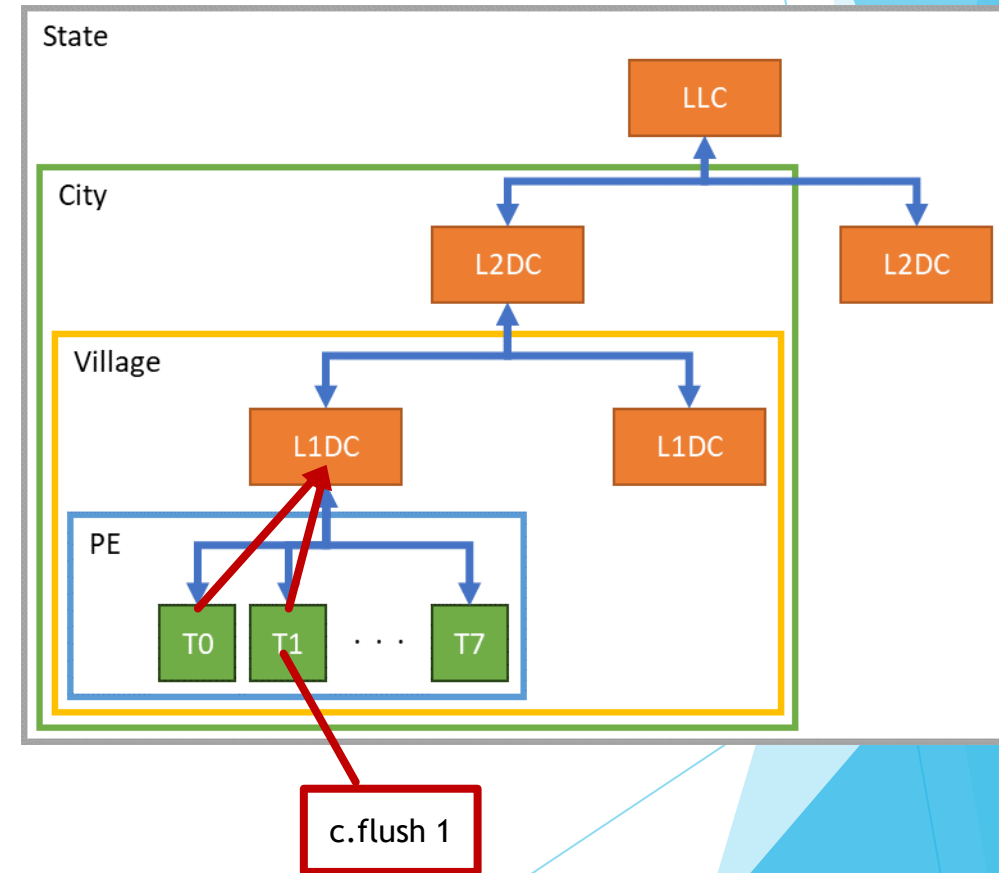
# 同期制御

- ▶ ノンコヒーレントキャッシュ
  - ▶ ソフトウェアによる明示的なキャッシュ同期
  - ▶ 同期ネットワークの1階層として実装



# 同期制御

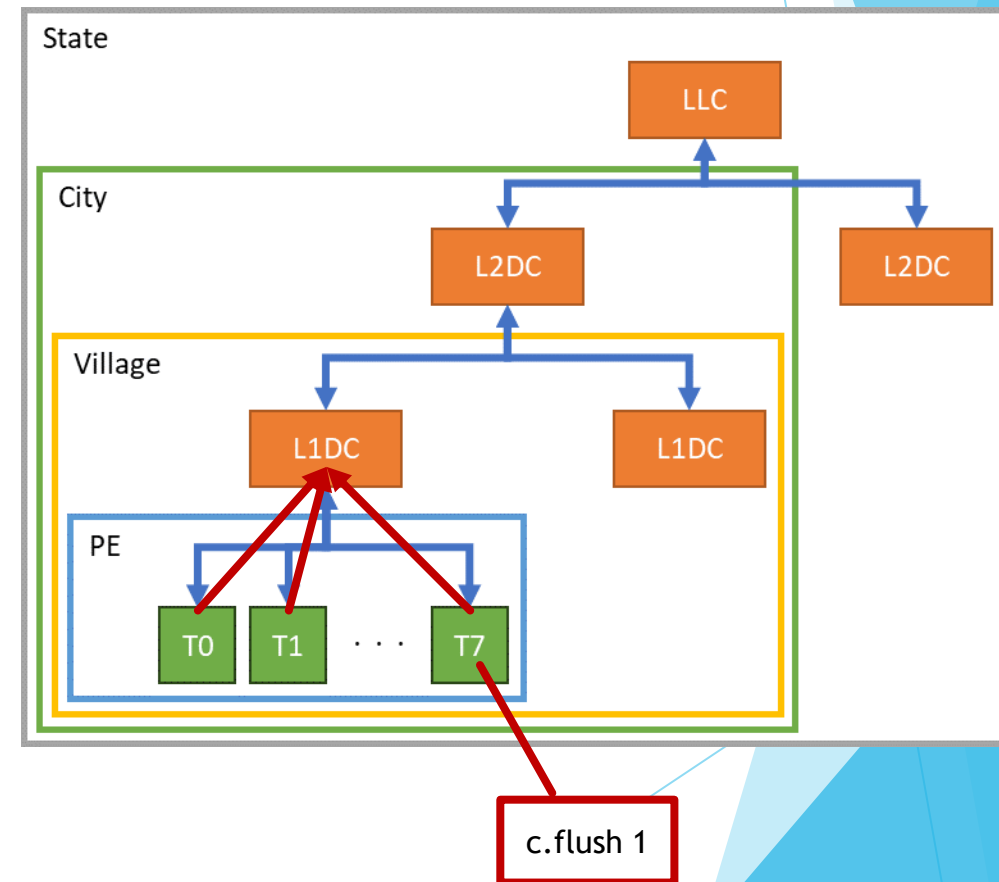
- ▶ ノンコヒーレントキャッシュ
  - ▶ ソフトウェアによる明示的なキャッシュ同期
  - ▶ 同期ネットワークの1階層として実装





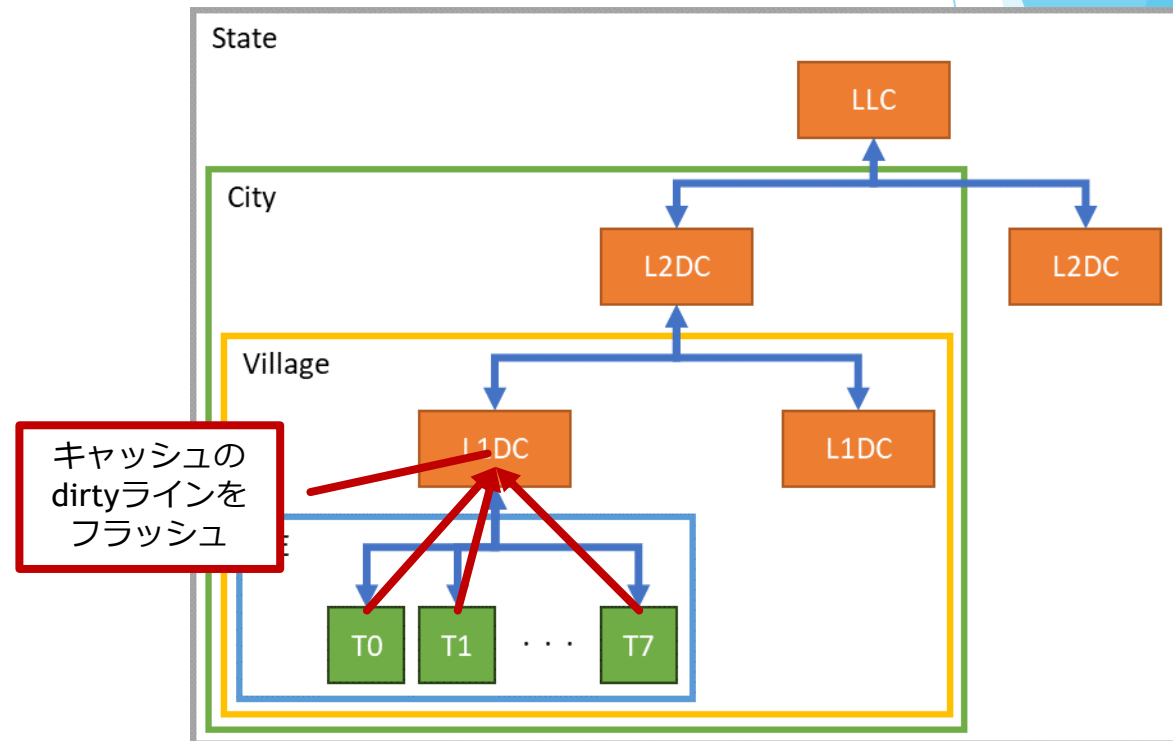
# 同期制御

- ▶ ノンコヒーレントキャッシュ
  - ▶ ソフトウェアによる明示的なキャッシュ同期
  - ▶ 同期ネットワークの1階層として実装



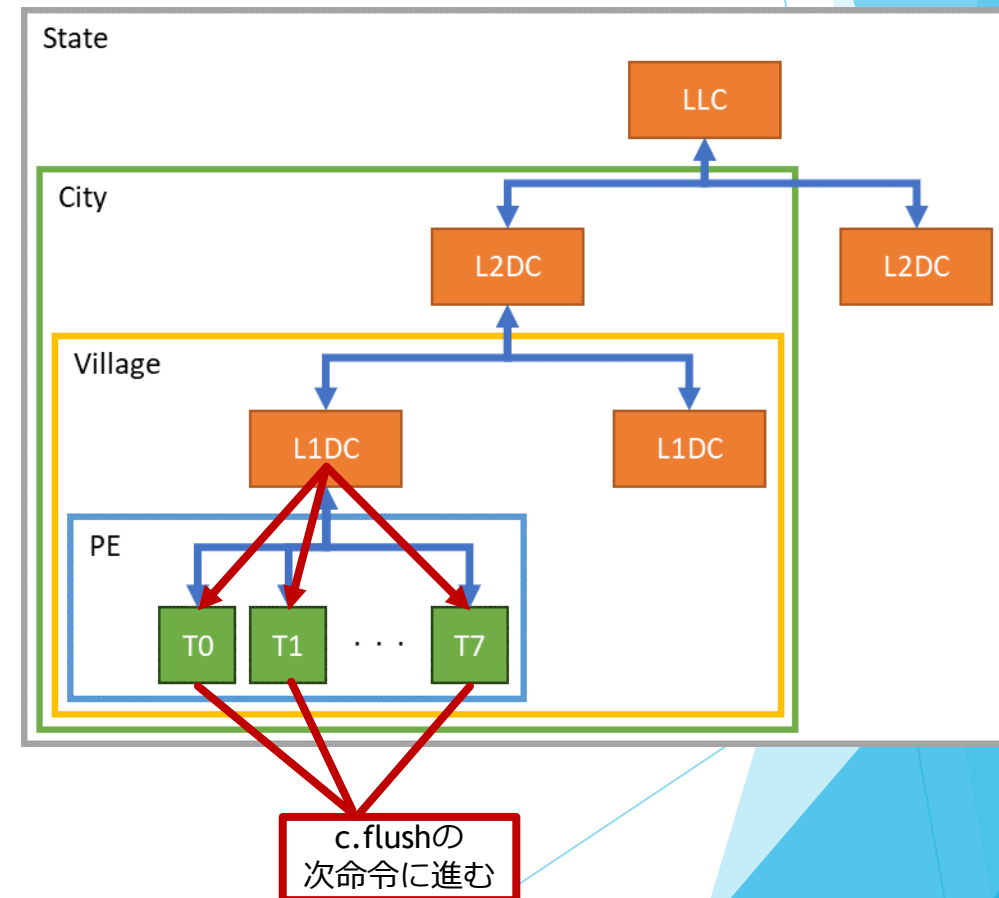
# 同期制御

- ▶ ノンコヒーレントキャッシュ
  - ▶ ソフトウェアによる明示的なキャッシュ同期
  - ▶ 同期ネットワークの1階層として実装



# 同期制御

- ▶ ノンコヒーレントキャッシュ
  - ▶ ソフトウェアによる明示的なキャッシュ同期
  - ▶ 同期ネットワークの1階層として実装

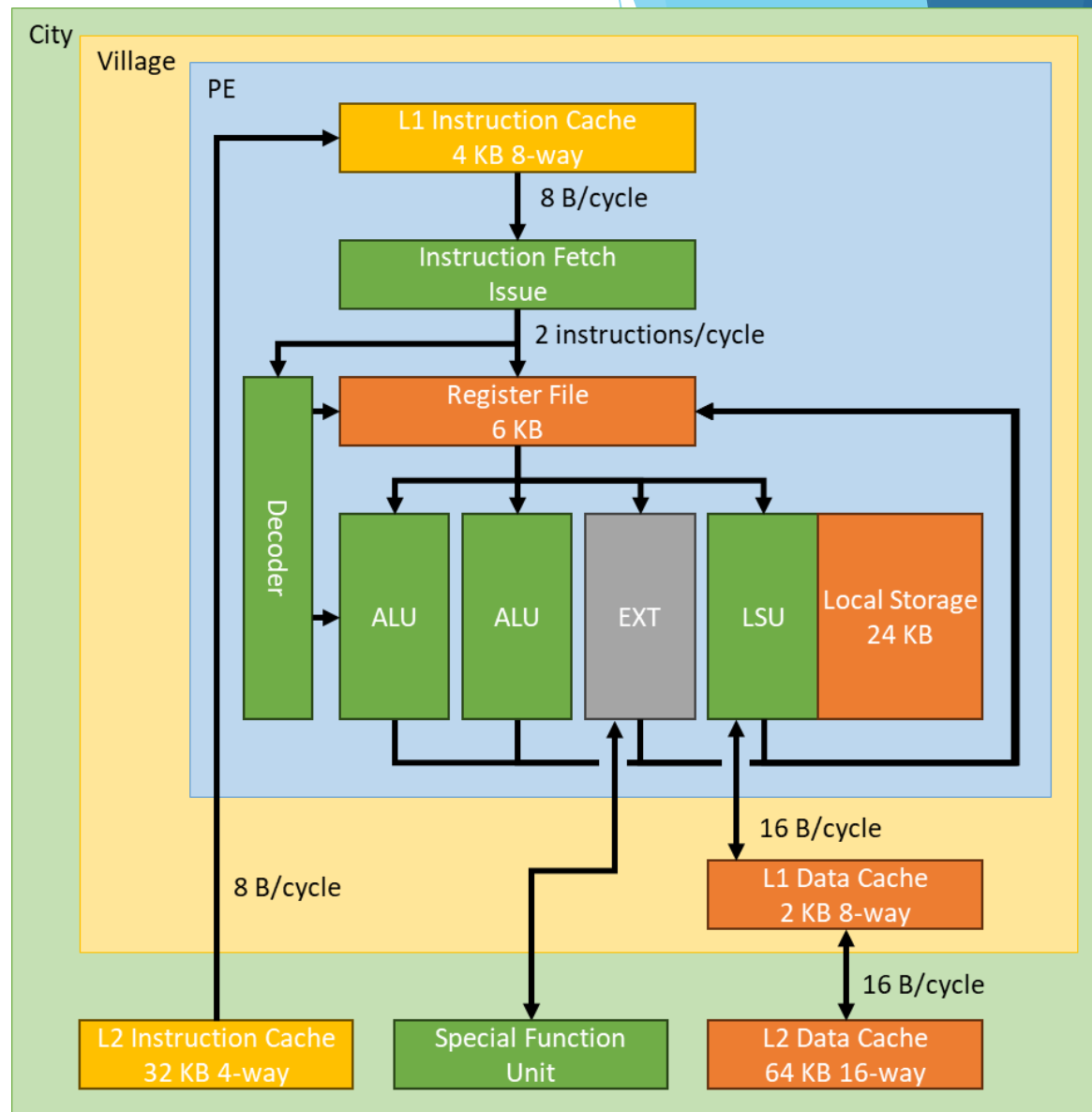


# PEZY-SCxの特徴

- ▶ プロセッサのアーキテクチャ
  - ▶ MIMDメニーコア
- ▶ コアのマイクロアーキテクチャ
  - ▶ 細粒度マルチスレッディング
- ▶ 命令セット
  - ▶ RISC系の独自命令セット

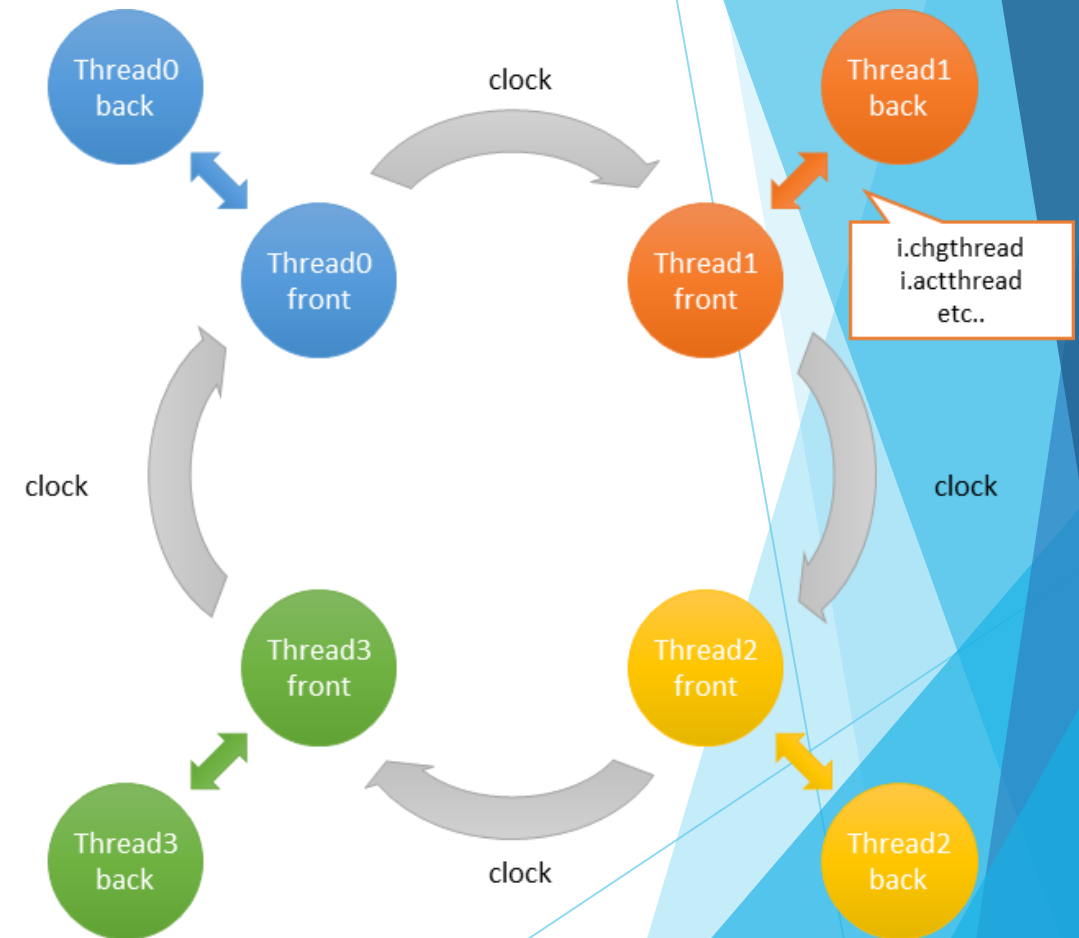
# PEの構造

- ▶ PEのコンセプト
  - ▶ 小さいコアを効率よく動かしたい
  - ▶ 「効率よく」 = 「パイプラインをストールさせない」
- ▶ 細粒度マルチスレッディング
  - ▶ 深いパイプラインは異スレッドで埋めればよい
  - ▶ パイプラインストールを防ぐ複雑な機構が不要
    - ▶ In-Order発行・Out-of-Order完了
    - ▶ 分岐予測なし
- ▶ ローカルストレージ
  - ▶ LSU直結のストレージ (≠キャッシュ)
  - ▶ ストールなしでロードストア可能



# スレッド実行モデル

- ▶ 4組のスレッドペア（計8スレッド）を実行可能
  - ▶ 各ペアの片側が実行状態、もう片側が待機状態
  - ▶ 実行状態の4スレッドをクロック毎に切り替え
- ▶ 明示的な命令でスレッドの実行・待機を切り替え
  - ▶ 長いメモリアクセスレイテンシを隠蔽する



# 目次

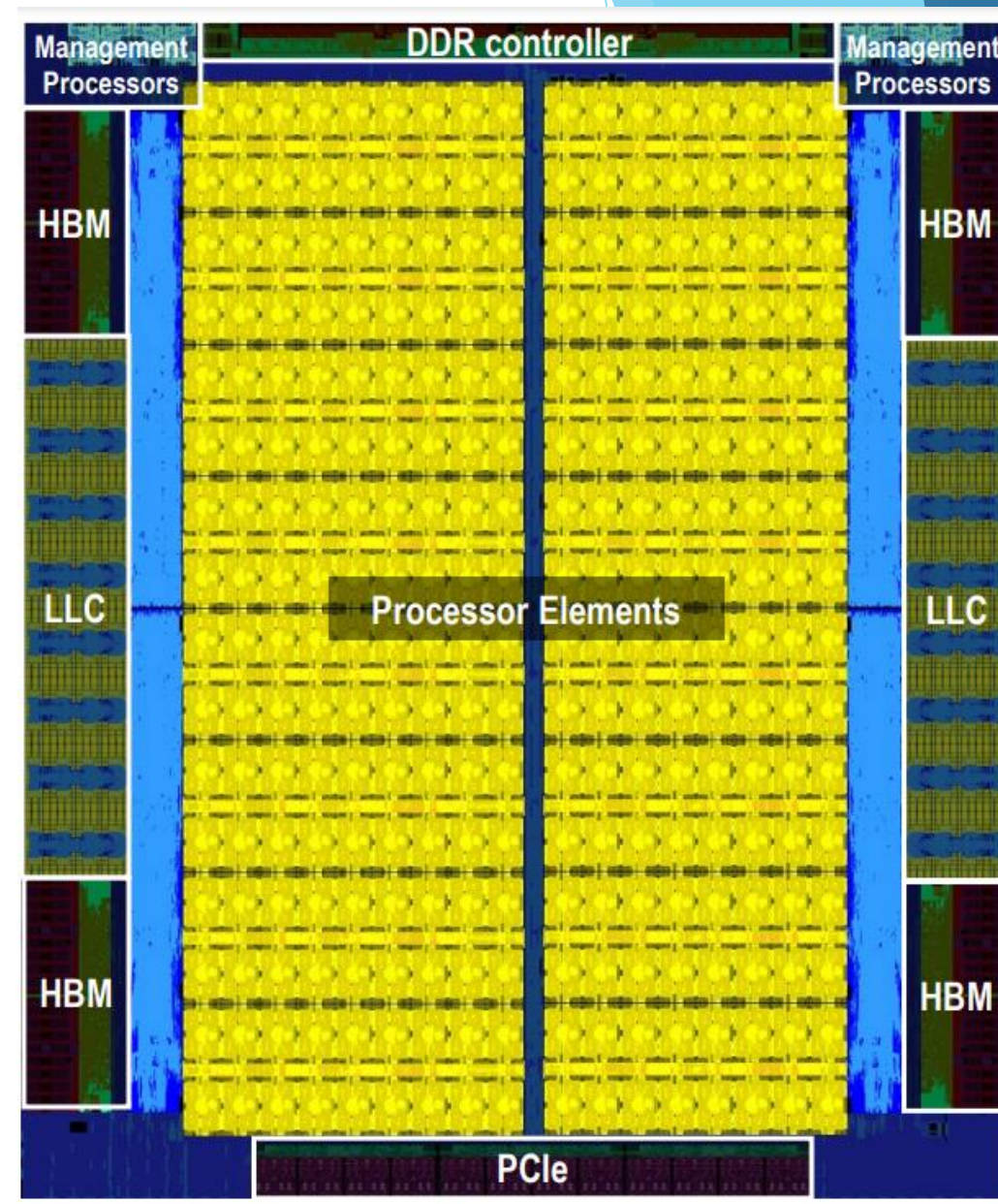
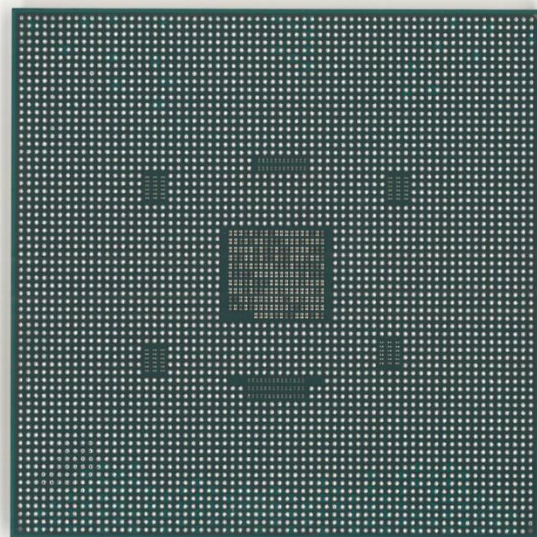
- ▶ PEZY-SCxのアーキテクチャ
  - ▶ PEZY-SCxとは
  - ▶ MIMDメニーコア
  - ▶ 細粒度マルチスレッディング
- ▶ PEZY-SC3の実装
  - ▶ LSI実装
  - ▶ 基板・システム実装
- ▶ アプリケーションでの評価
  - ▶ LINPACKベンチマーク
  - ▶ ヒトゲノム解析
- ▶ まとめ



# PEZY-SC3の実装

## ▶ LSI実装

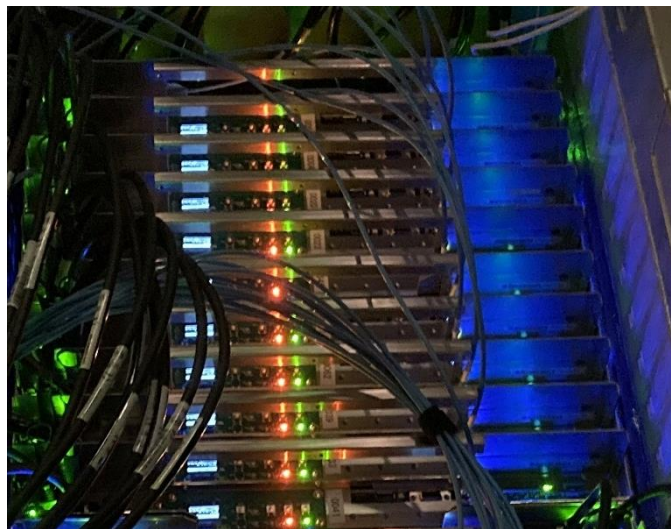
- ▶ プロセス : TSMC 7nm
- ▶ ダイサイズ : 25.7mm x 30.6mm
- ▶ トランジスタ数 : 408億トランジスタ
- ▶ 消費電力 : 470W (max)





# PEZY-SC3の実装

- ▶ ノード実装
  - ▶ ホストプロセッサ : AMD EPYC 7702P x 1
  - ▶ プロセッサ : PEZY-SC3 x 4
  - ▶ インターコネクト : InfiniBand EDR
- ▶ NA-J2システム
  - ▶ ノード数 : 70
  - ▶ ピーク性能 : 2.03PFlops/s
  - ▶ 冷却方式 : 液浸冷却
- ▶ ラックマウントモデル
  - ▶ ノード数 : 1
  - ▶ ピーク性能 : 78.8TFlops/s
  - ▶ 冷却方式 : 空冷



# 目次

- ▶ PEZY-SCxのアーキテクチャ
  - ▶ PEZY-SCxとは
  - ▶ MIMDメニーコア
  - ▶ 細粒度マルチスレッディング
- ▶ PEZY-SC3の実装
  - ▶ LSI実装
  - ▶ 基板・システム実装
- ▶ アプリケーションでの評価
  - ▶ LINPACKベンチマーク
  - ▶ ヒトゲノム解析
- ▶ まとめ

# アプリケーションでの評価

- ▶ LINPACKベンチマーク
- ▶ ヒトゲノム解析

# LINPACKベンチマーク

- ▶ 連立一次方程式の求解ベンチマーク
  - ▶ スーパーコンピュータのTOP500/Green500ランキングで 사용되는
- ▶ 実行結果 (NA-J2)
  - ▶ 演算性能 : 1.95PFlops/s
    - ▶ 481位 @ TOP500 JUNE 2023
  - ▶ 消費電力効率 : 21.892GFlops/W
    - ▶ 39位 @ Green500 JUNE 2023
- ▶ Green500ランキング上位は専用の行列積演算器を持つプロセッサが占めている
  - ▶ NVIDIA H100/A100, AMD MI250X, PFN MN-Core
- ▶ 専用演算器を持たないプロセッサとしては最も消費電力効率が高い
  - ▶ 幅広いアプリケーションで高い電力効率を発揮できる

# ヒトゲノム解析

## ▶ GATK Best Practice

- ▶ ゲノム解析で一般的に使用されている解析ワークフロー
- ▶ ゲノム配列の検索・編集距離計算・尤度計算

## ▶ 実行結果（ラックマウントモデル）

- ▶ 実行時間：15分
  - ▶ 一般的にはCPUで2000分、GPUで50分程度
- ▶ 精度：99.90% (SNP) 99.67% (Indel)
  - ▶ 一般的な解析ソフトウェアと同等以上

## ▶ MIMDメニーコアの特性を活かしたアプリケーション

- ▶ データの並列性が高く、各データ独立に探索する必要があるためMIMDに向いている

# 目次

- ▶ PEZY-SCxのアーキテクチャ
  - ▶ PEZY-SCxとは
  - ▶ MIMDメニーコア
  - ▶ 細粒度マルチスレッディング
- ▶ PEZY-SC3の実装
  - ▶ LSI実装
  - ▶ 基板・システム実装
- ▶ アプリケーションでの評価
  - ▶ LINPACKベンチマーク
  - ▶ ヒトゲノム解析
- ▶ まとめ

# まとめ

- ▶ PEZY-SCxのアーキテクチャ
  - ▶ MIMDメニーコア
  - ▶ 細粒度マルチスレッディング
- ▶ PEZY-SC3の実装と評価
  - ▶ スパコンシステムとラックマウントの筐体を実装
  - ▶ いくつかのアプリケーションで良好な結果を得た
    - ▶ LINPACKベンチマーク
    - ▶ ヒトゲノム解析
- ▶ 今後の予定
  - ▶ 次世代機の物理設計中
  - ▶ 次々世代機の論理設計中