Fraunhofer IIS

Fraunhofer Institute for Integrated Circuits IIS

STXMOD – SoC Overview

Norbert Schuhmann

Fraunhofer IIS - Fraunhofer ITWM

STXMOD Application

The Problem

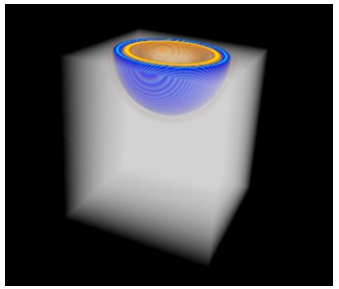
- Since 2005 processor development has an energy problem
  - Multicore Chips
  - today > 64 cores
- Commodity processor bear a heavy legacy load of instructions
  - Inceased complexity – more transistors - high power consumption ….
- Current high end HPC systems are using Accelerators NVIDIA / AMD GP-GPUs
  - Ende of Moores law 7nm, 5nm, 3nm -> increasingly expensive with little gain
  - NVIDIA A100 : about 250Watt, NVIDIA Hopper about 700Watt (3x Performance)
  - Costs are increasing -> single Hopper Chip about 15,000 US$

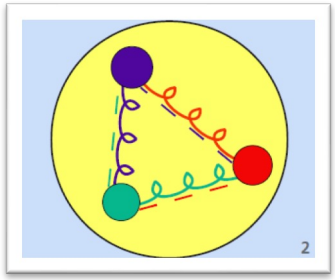
-> Need for new energy efficient & cost effective chip architectures

STXMOD Application

The Answer: STX -> Stencil Tensor aXelerator (Stencil processing unit = SPU)

- Stencils are repetitive applied regular local pattern
- Explicit Simulation Methods
  - Wave Propagation – Seismic, Fluids
  - Fluid Dynamics - Turbulence
  - Weather & Climate Simulation
  - Lattice-Boltzmann
- Convolutional Neural Networks (CNN)
  - Latest developments in AI change the pattern
- Quantum chromodynamics
- Image Analysis



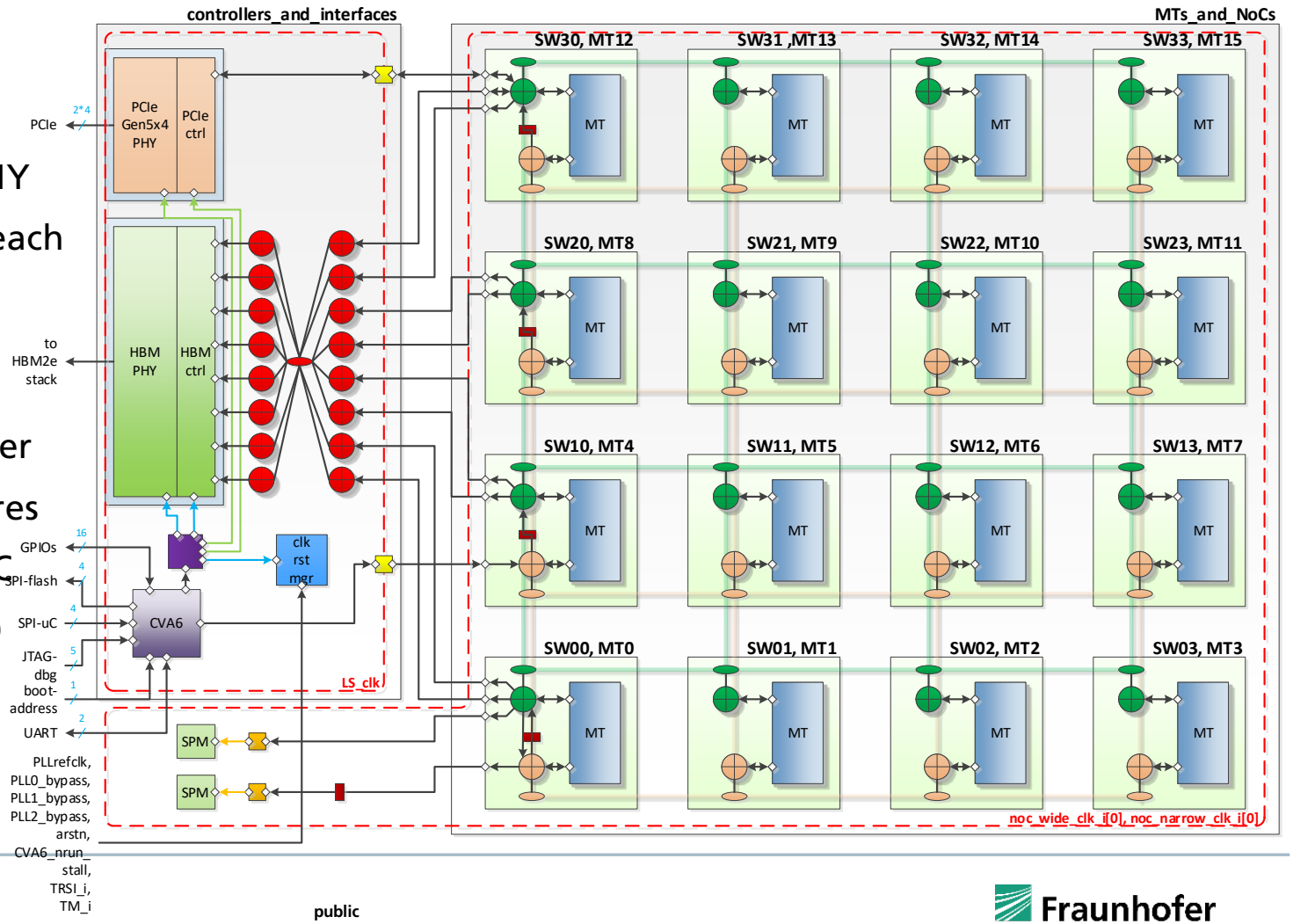


=> The solution: Consequent Hardware-Software Co-Design

STXMOD ASIC

Architectural overview

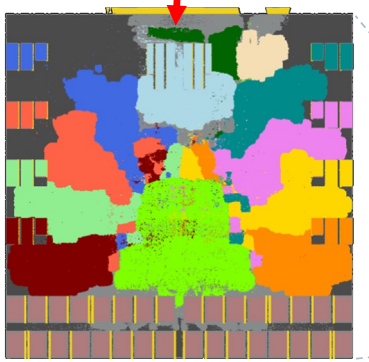
- HBM2 Controller + PHY
- PCIe Gen5x4 Controller + PHY
- 16x Microtiles w/ 4 clusters each
- SPU Stencil Processing Unit
  - the processing core -
- 8 RISC-V w/ SPU
  + 1 RISC-V w/o SPU per cluster
- Total of 512 parallel SPU cores
- Wide-,narrow-, memory NoC
- NoC (512 Bit, 64 Bit, 128 Bit)
- PVT-timing closure @1+Ghz,
  Die size: 180 mm^2
  @ GF12LPP

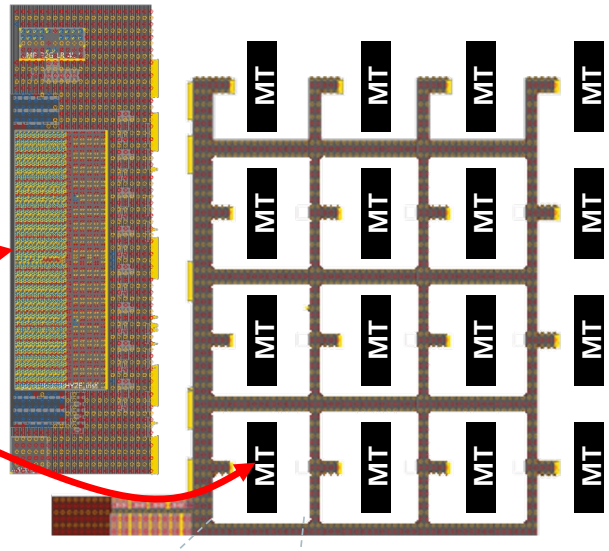


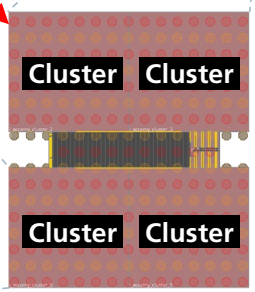
STXMOD ASIC Backend

Partitions

- Partitions
  - Top
    - Left side
    - NoC partition
    - MT Micro Tile
      - Cluster (8 SPUs)

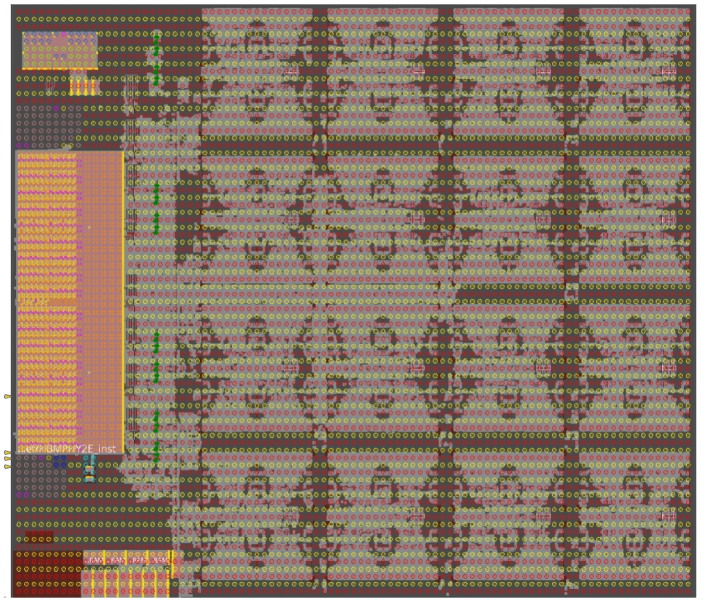






TCDM

All together: TOP Layout

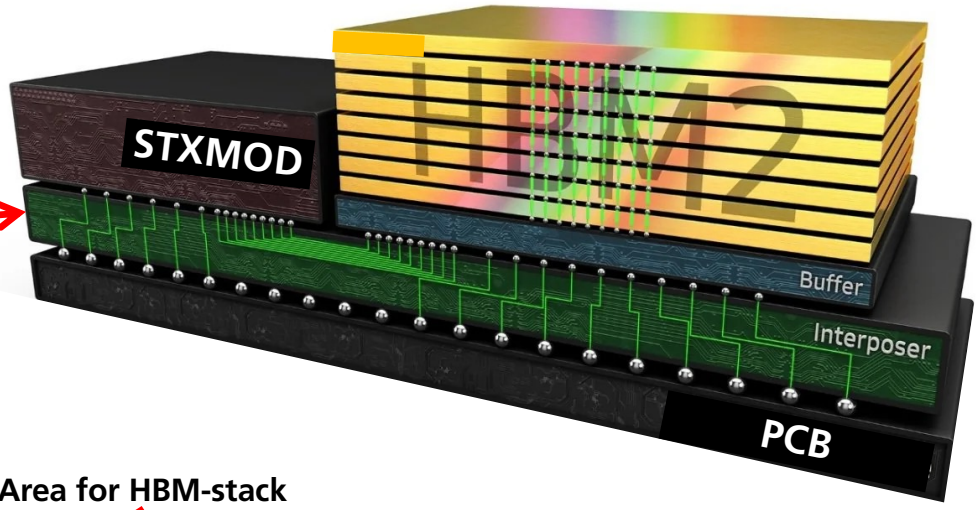


Interposer

2.5D integration

- Additional Silicon (reduced production process)
- Additional IC-design step incl. routing and bump placement
- Additional Signal- and power-integrity analysis and design
- Additional level of functional verification

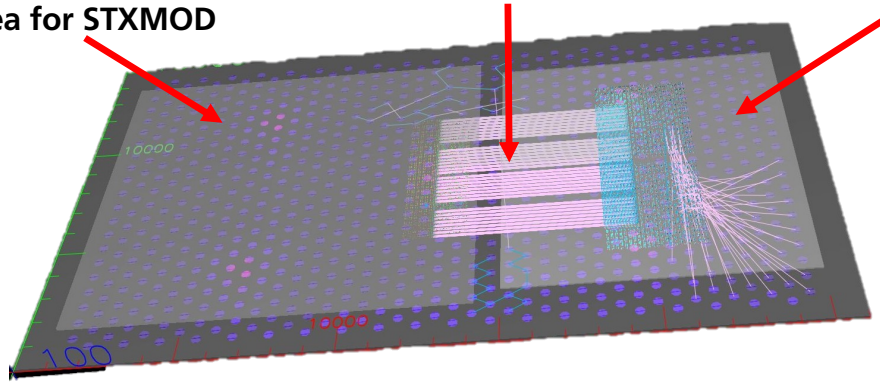
Interposer example



~2.5 k connections between STXMOD and HBM-stack

Area for STXMOD

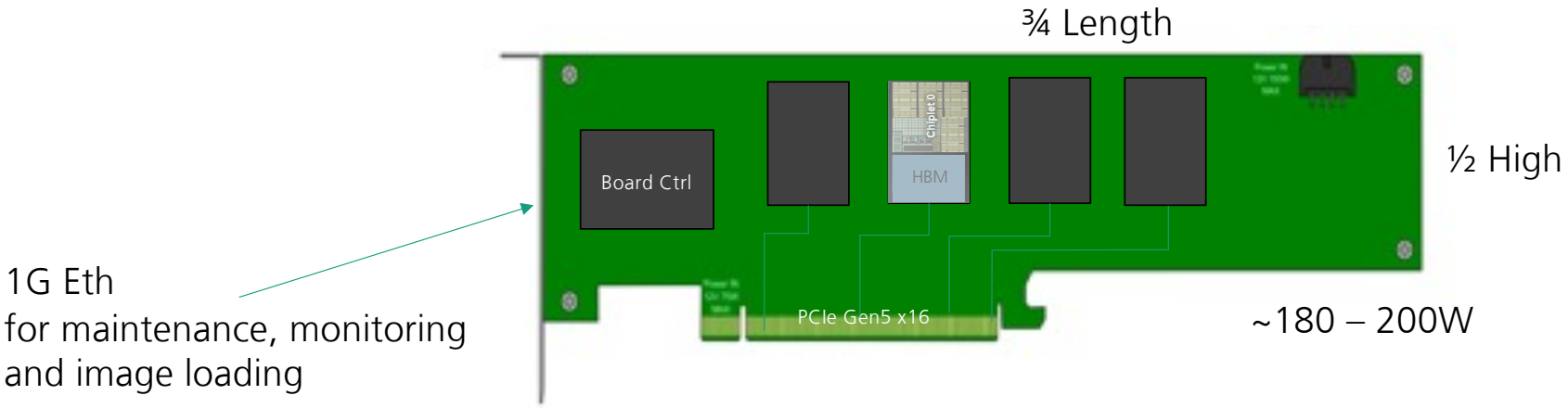
Area for HBM-stack



PCB

STX-Card

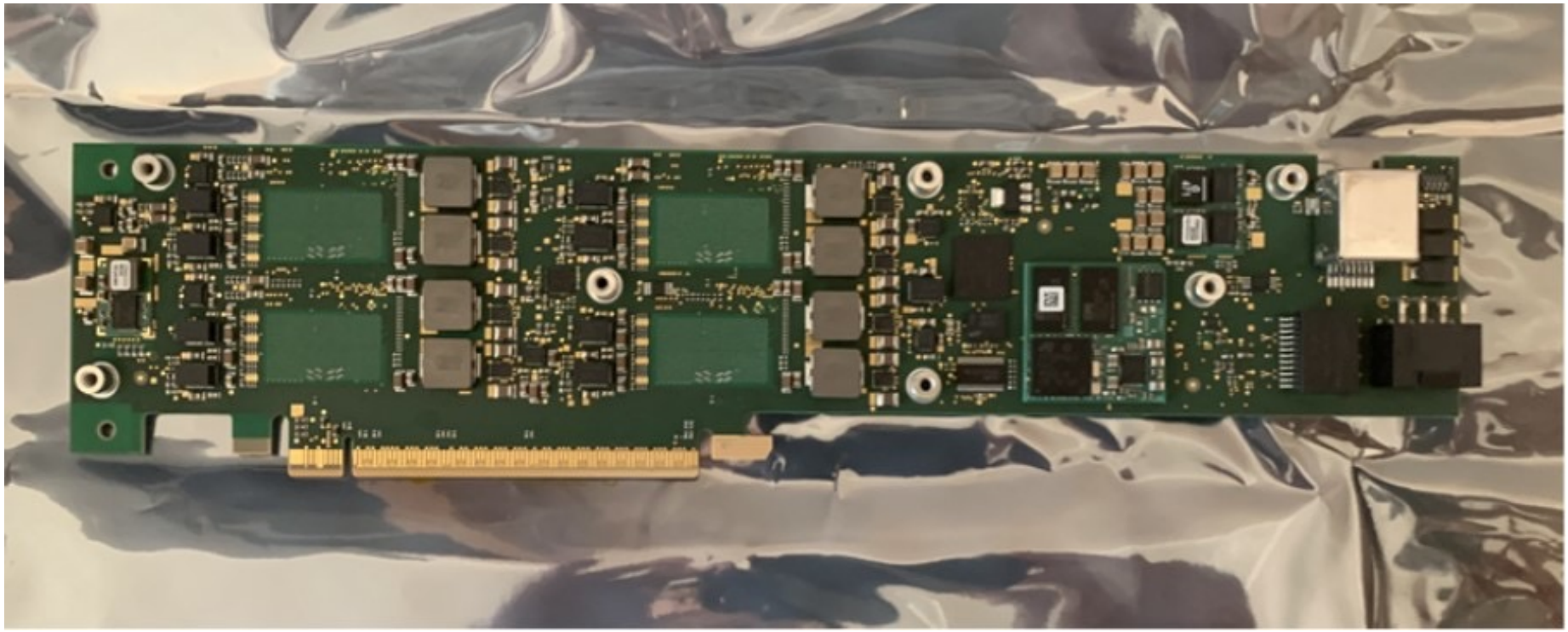
- 4x STXMOD (STXMOD-SoC + HMB-stack on 65 nm silicon interposer)
- Board controller for maintenance + supervision and SW provision
- PMU Power Management Unit
- PCIe gen5 4x as data interface to 4 STXMODs
- Prototype board for feasibility and AVT
- Product version dev starting asap



Fraunhofer IIS

PCB

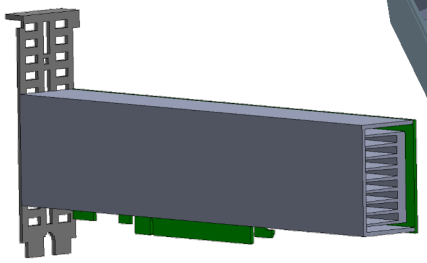
STX-Card

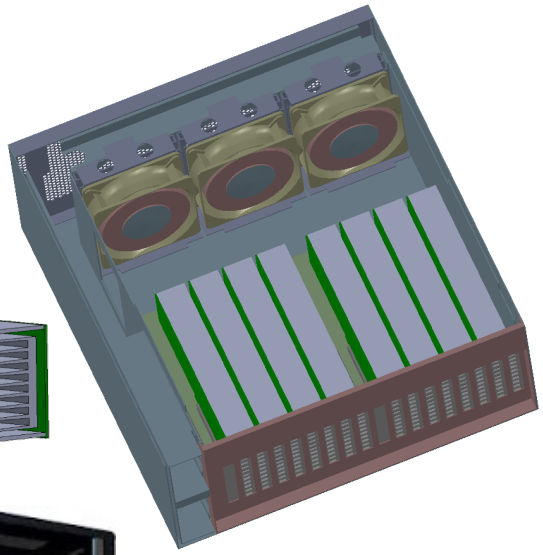


Products

Shelfs / Racks Example

- Air cooled prototype system
  - with 8 STX cards -> 32 STXMODs (==16,384 SPUs)
  - Up to 4 x 400 Gbit/s host channel adapter
- Liquid cooled system
  - 2U enclosure
  - with 16 STX cards -> 64 STXMODs
  - 20 systems per rack
    (== 655,360 SPUs)





USP

Typical HPC Rack Example

- Single Rack:
  - Price: about 1,6 M€
  - Contains: 320 Boards, ~60 kW power
  - Compares for RTM with 600 NVIDIA A100 at about 5 M€ and 150kW

- Our USPs versus NVIDIA:
  - Cheaper: 3x
  - Energy efficiency: 3x
  - Easier to program
  - Smaller in size

- What next: Improved STX generation in 4 nm technology in progress …