# FlexPath – Dynamic Reconfigurable Processing Paths in NPUs

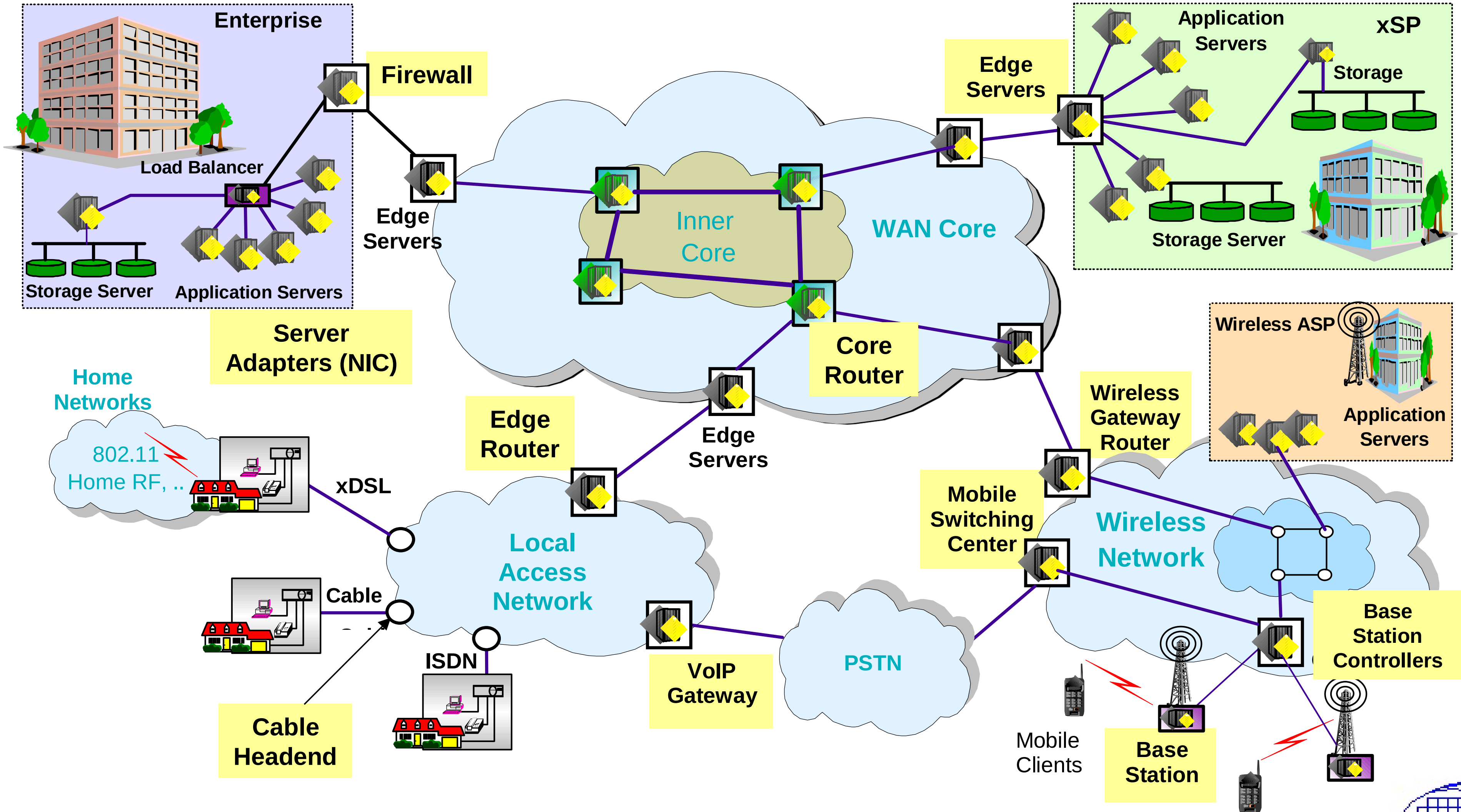
A. Herkersdorf, M. Meitinger, R. Ohlendorf, T. Wild

Technical University of Munich
Theresienstr. 90

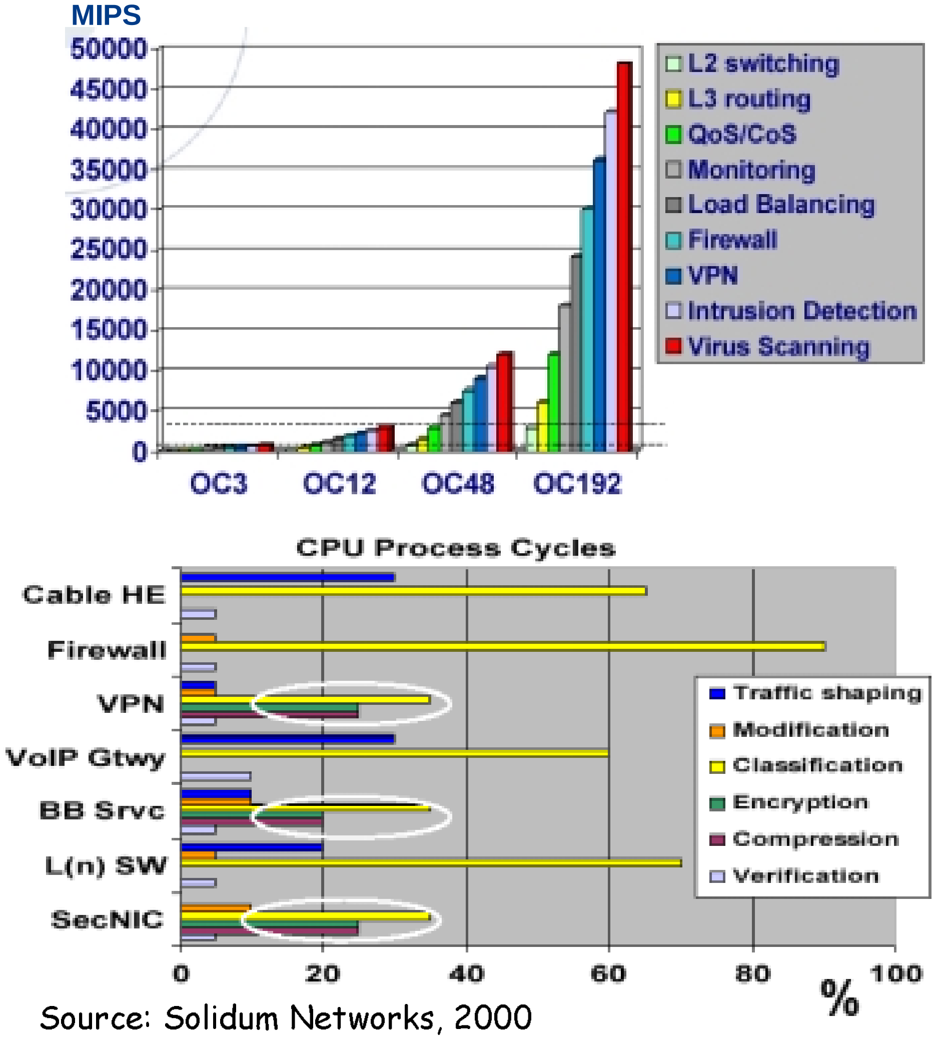
www.lis.ei.tum.de

TECHNISCHE
UNIVERSITÄT
MÜNCHEN

# Network Processor Application Areas



# NPU Challenge: Processing Performance

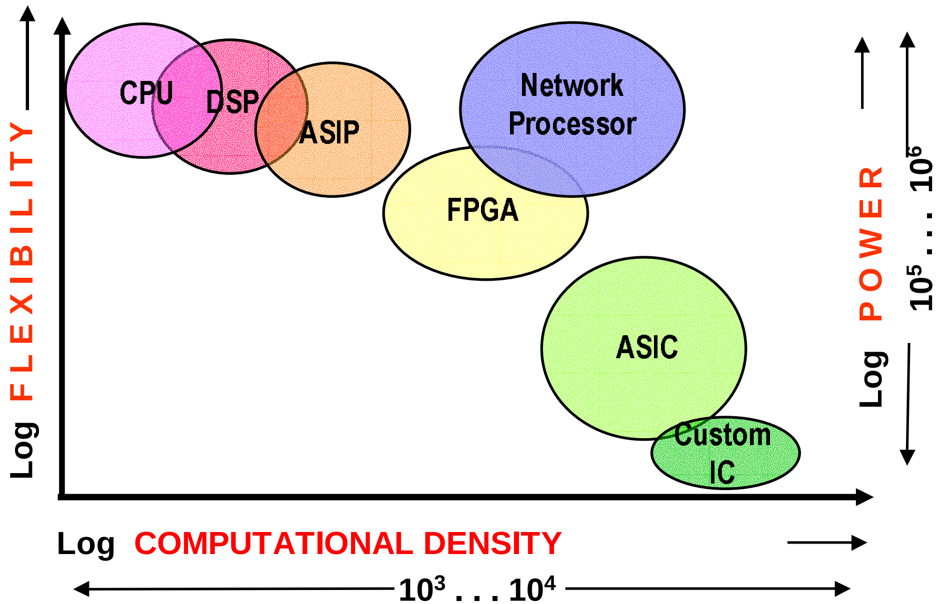



Source: Solidum Networks, 2000

- Wide spread of required MIPS between different applications
  - ~200 – 3000 instr. / packet
- High-end parallel CPU's clusters already for "low-end" applications
  - Equivalent of 3-4 Pentium III (1GHz) for OC-48 CoS forwarding
- Imbalance between NP functions
  - Classification, security and traffic management consume most MIPS

➔ Off-load CPU from processing intensive task with flexible HW assists and coprocessors

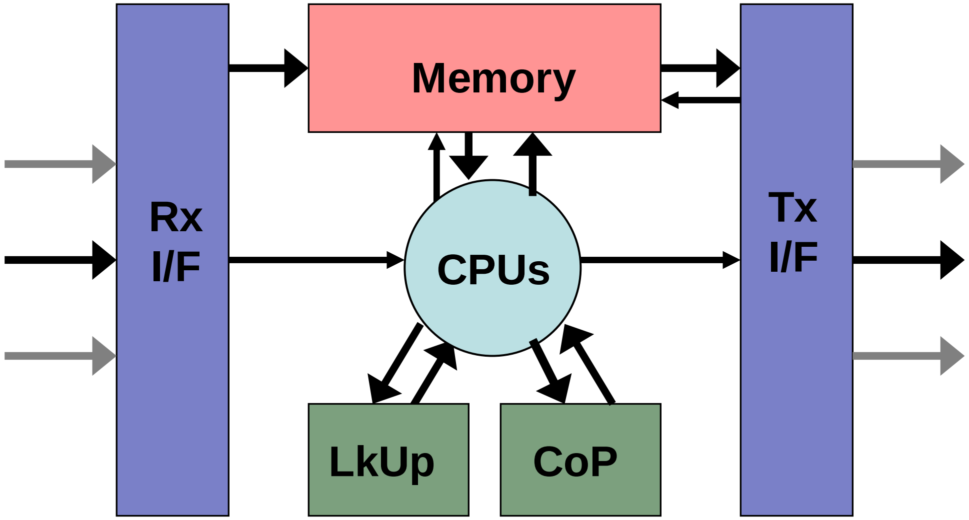
# NPU Challenge: Processing Performance




Source: T. Noll, RWTH Aachen

- Balanced partitioning of total MIPS in CPU-MIPS and "HW-MIPS"
- Focus on avoiding:
  - flexibility loss through "ASIC dominance"
  - Performance loss through "CPU bottleneck"

# Conventional NPU Architecture




Processor-centric NPU Architectures:

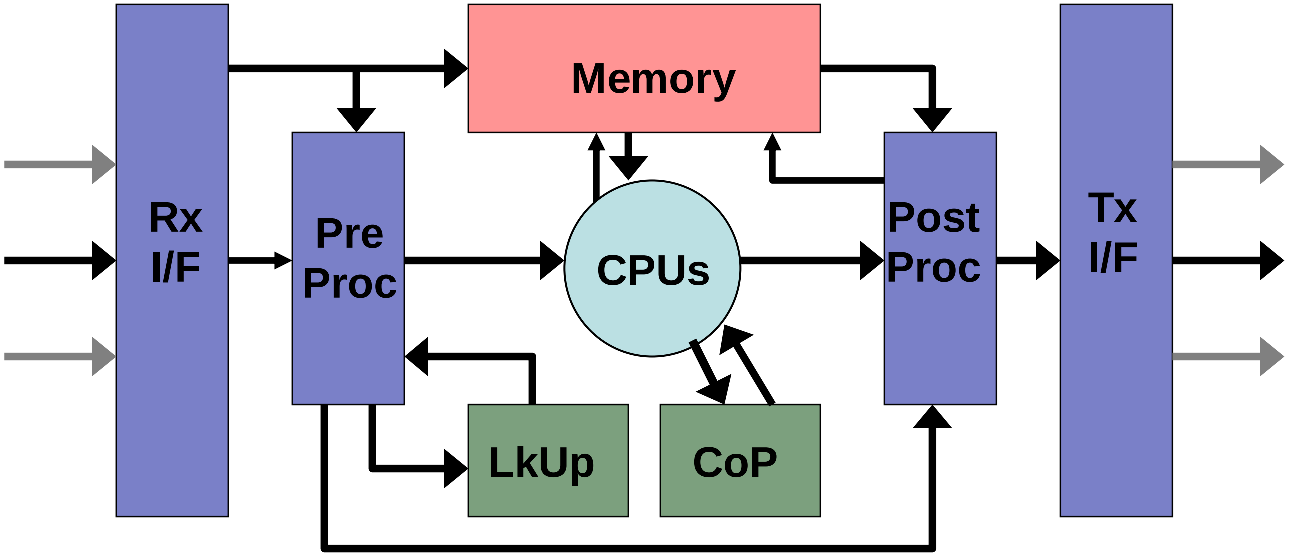
- Each packet traverses CPUs (multiple times)
- … independent of processing demand
- CPU has to perform all bit-level packet analysis and modifications

Challenge:

- Number of events / packets to be handled by CPU(s)
  - CPU clock cycle budget per packet

| | OC-3 | OC-12 | OC-48 |
|---|---|---|---|
| Mbps | 155.5 | 622.1 | 2488 |
| pkt/s | 304 K | 1.21 M | 4.86 M |
| s/pkt | 3.29 μ | 823 n | 206 n |

# LIS FlexPath NPU Architecture




FlexPath NPU:

- PreProc performs hdr parsing, bit-level analysis
- PreProc calls address LkUp coprocessor
- … and pass results to CPU
- … or bypasses CPU "where applicable" (e.g. TCP/IP ACKs)
- PostProc performs packet modifications and issues transmission

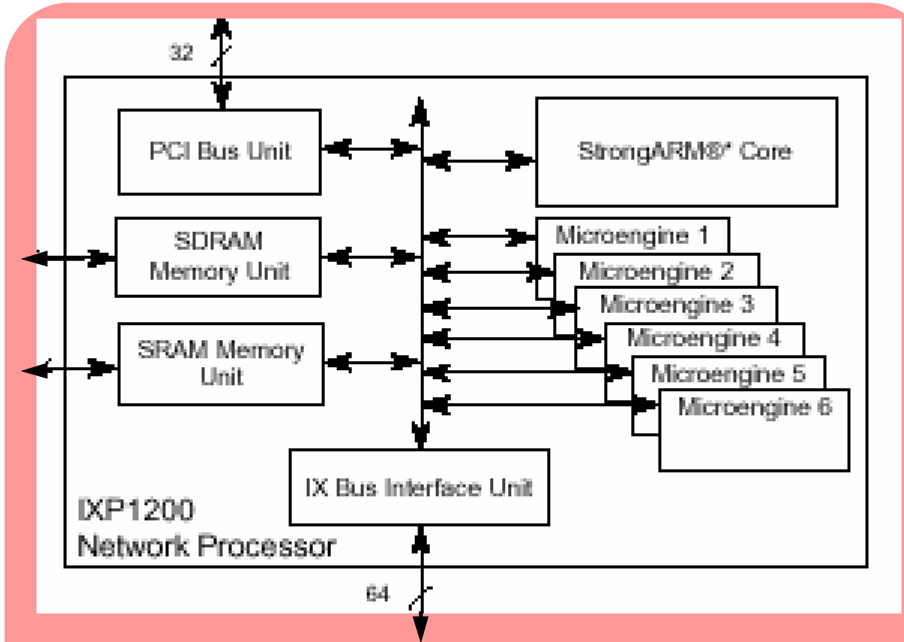
Optimized function traversal:

- Off-loads bit-level manipulations from CPU
- Reduces absolutes events (packets) traversing CPUs

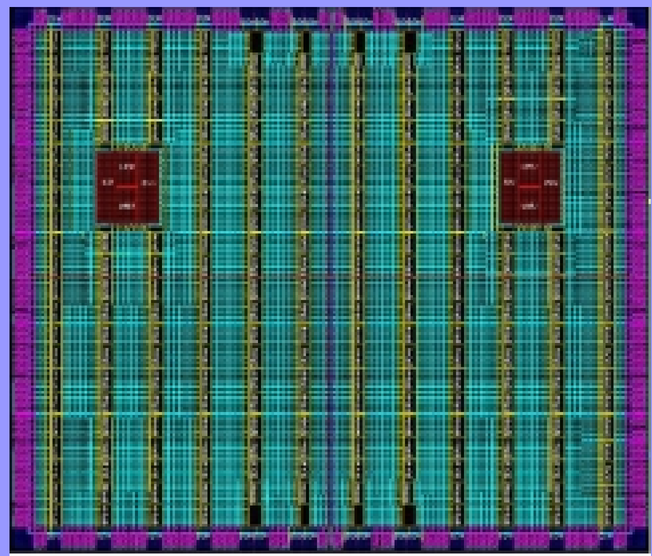
Rationality:

- High speed links carry heterogeneous traffic mix serving variety of applications
  - Processing requirements in NPU vary substantially on a per "micro flow" basis

# Quantifying FlexPath Concept Leverage

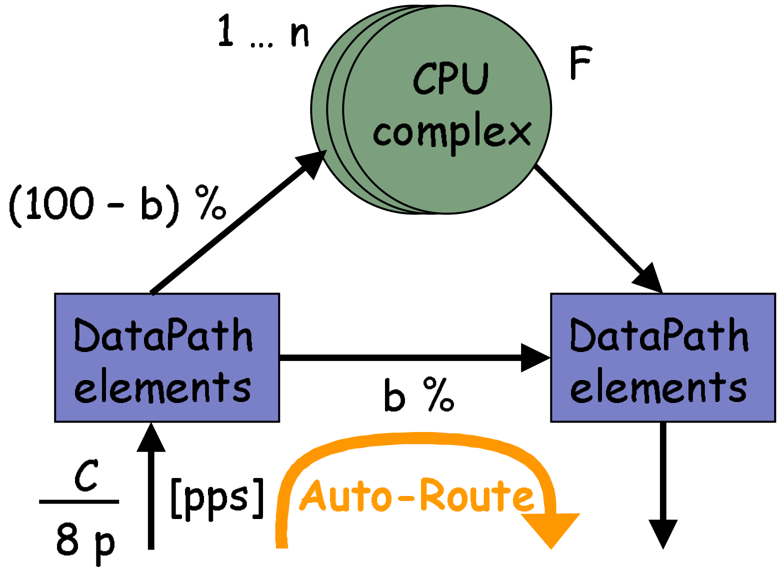



- Intel IXP 1200 NPU product
- Aquired through LevelOne multi-Billion $ acquisition
- 6 multi-threaded ASIPs
  - 996 MIPS (nominal)

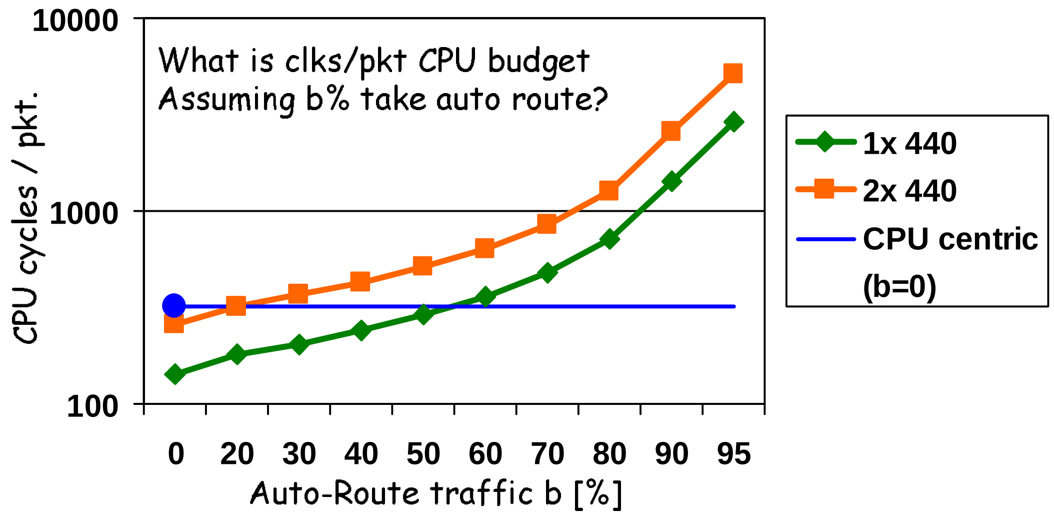


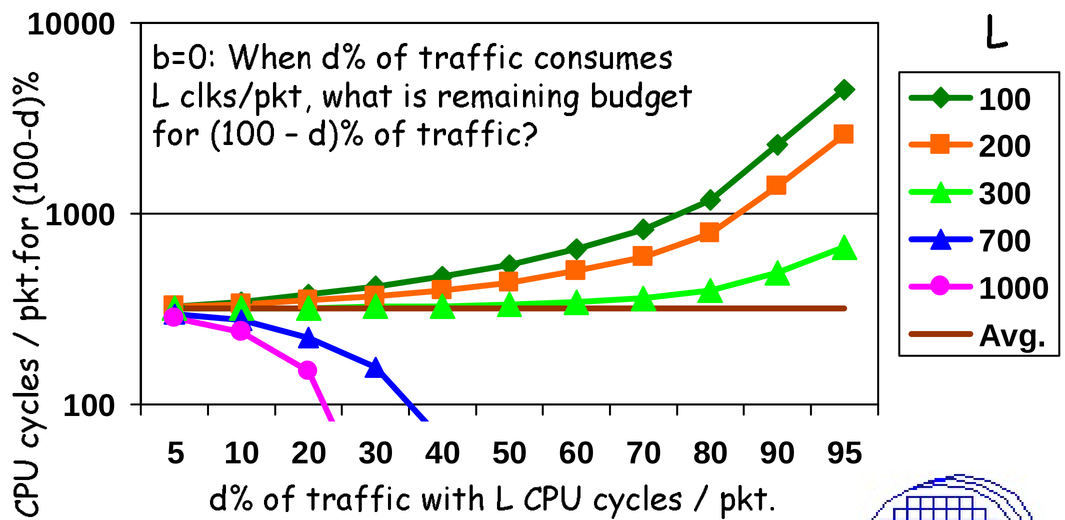
- XILINX Virtex II Pro FPGA
- 2 std. PowerPC 405 RISCs
  - 2 x 200 MIPS (nominal)
- LIS-FlexPath NPU architecture

# CPU Bypass Gain Potential




| | 1x 440 | 2x 440 | centric |
|---|---|---|---|
| Processor clock: F [MHz] | 400 | 400 | 166 |
| Packet size: p [B] | 80 | 80 | 80 |
| Aggregate NP capacity: C [Mbps] | 2000 | 2000 | 2000 |
| Number of parallel processor: n | 1 | 2 | 6 |
| Nominal MIPS (CPI = 1) | 400 | 800 | 996 |
| Avg. clks / pkt ( b = 0%) | 130 | 260 | 320 |



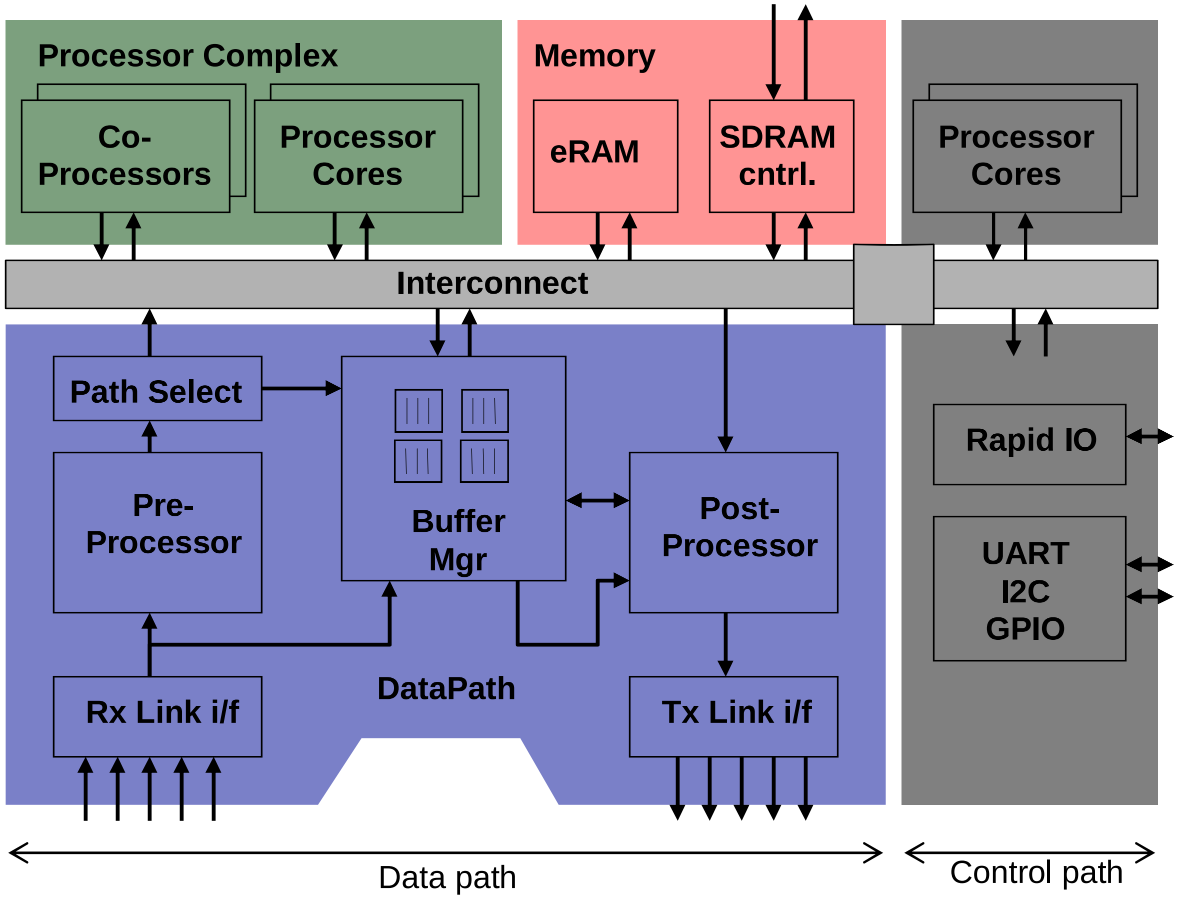



# FlexPath Network Processor SoPC

- Majority of building blocks recruited from IP libraries
- Only NP specific functions developed

Objectives:

- Improve Network-**Processor** performance without touching the **Processor** Complex
  - Pre-/PostProcessor CPU off-load
  - Packet Buffer Manager
- Achieve commercial NPU throughput with conventional RISC CPU (PowerPC) on FPGA evaluation platform



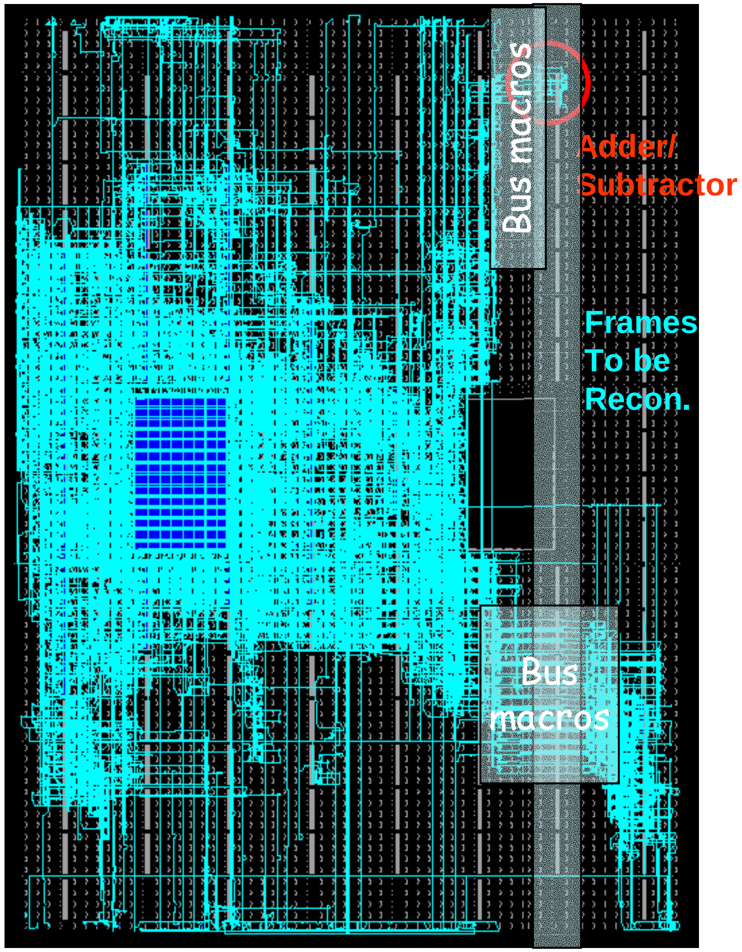
# Reconfigurable Functions / Reconfigurable Interconnect

Do we sufficiently care about reconfigurable interconnects, or is this a "no issue"?

A. Herkersdorf, C. Claus, M. Meitinger, R. Ohlendorf, T. Wild

TECHNISCHE UNIVERSITÄT MÜNCHEN

# Reconfiguration of Functionality



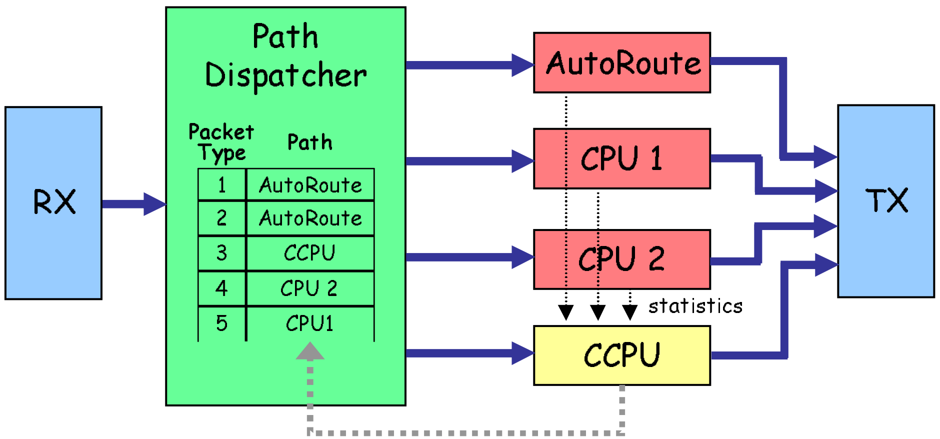

- Change of functionality is achieved by modifying logical elements (CLBs, Routing resources, BRAM)
- Structure of function is modified, communication to the function remains the same (realized through bus macros)
- Certain number of frames in the bit stream changes the configuration memory.
- Adder/Subtractor Design

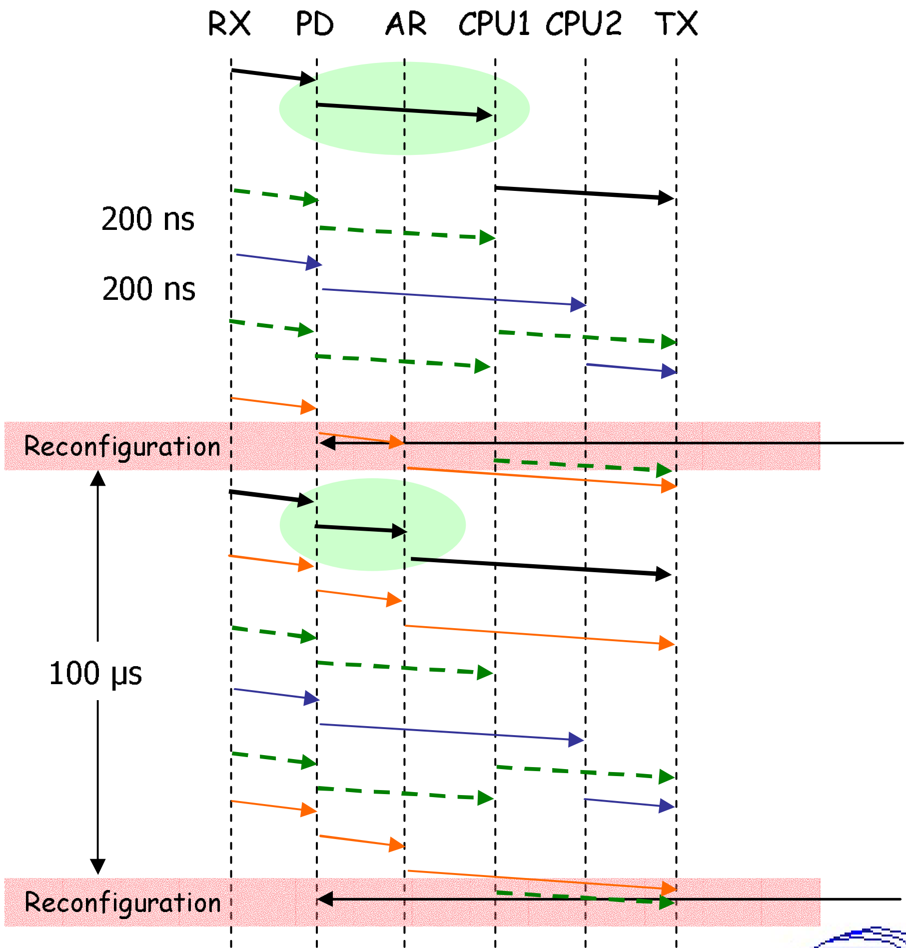
| | |
|---|---|
| Bit stream size: | 20306 Byte |
| # of Frames to be reconfig.: | 22 |
| Reconfiguration time: | 19 ms |

- Aim: to reduce this reconfiguration time to less than 1 ms

[Becker et al., Stechele et al.]

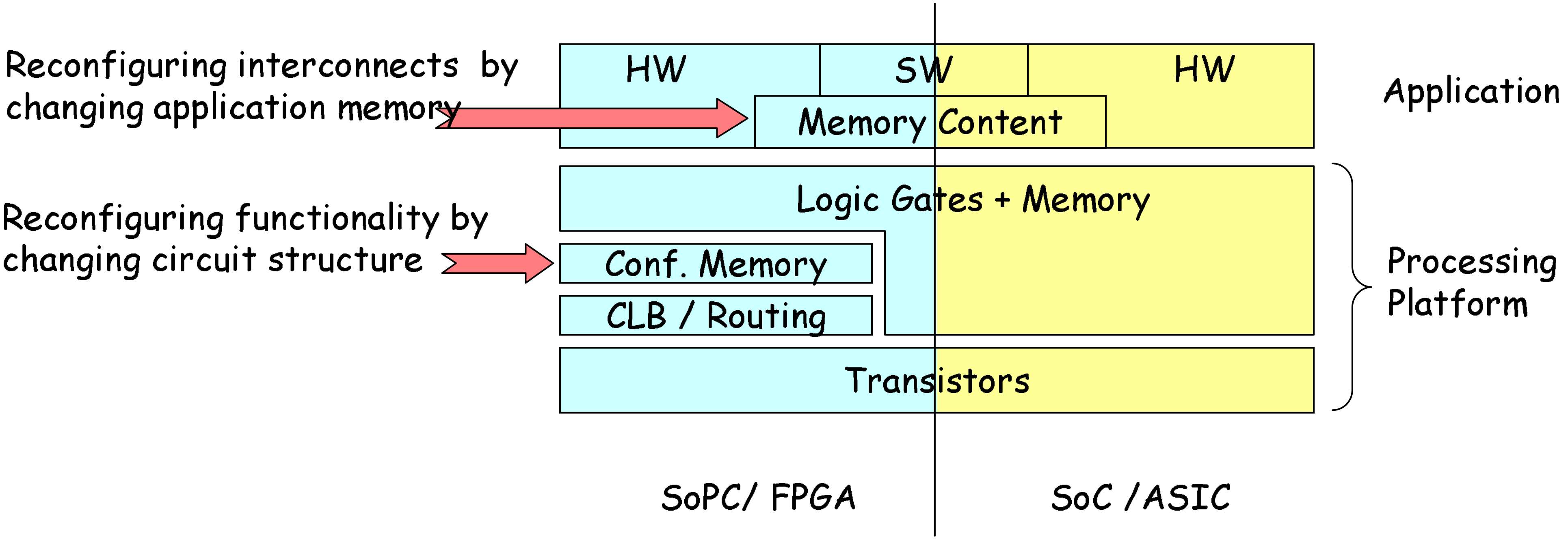
# Reconfiguration of Interconnect




- Functionality of modules is fixed, communication paths through system change
- PD contains rule base, which determines the path of different packet types after passing the PD
- Rule base lookup rates ~ 200 ns, upper bound for reconfiguration time
- Rule base update times ~ 100 µs during running system

# Reconfigurable Architectures - Layers



# Reconfigurable Architectures - Comparison

| Characteristics | Reconfiguration of Functionality | Reconfiguration of Interconnect |
|---|---|---|
| Reconfiguration time | ~ 10-100ms | ~ 200ns |
| Real-time requirements | More relaxed (~ 10-100ms), circuit can be (partially) stopped during reconfiguration | Very hard (~ 200ns), circuit must remain fully operational during reconfiguration |
| Generator of reconfigur-ation information | Pre-calculated bit streams, online modified bit streams | Online control point SW |
| What is reconfigured | HW circuit: CLBs, routing resources, BRAM content | Memory (lookup) contents |
| Possible HW platforms | FPGA | FPGA, ASIC |
| Initiator of Reconfiguration | Control Plane Software | |

# Multiway SoC Interconnect Taxonomy

| | Bus | Crossbar | Ring | NoC |
|---|---|---|---|---|
| Topology | Shared media | NxM switched | Add/drop psp segments | Meshed, random switch / router |
| Capacity | 1 (normalized) | close to M (< N) | ≤ N | close to N |
| Multicast | easy | costly | easy | costly |
| Complexity | low | medium | low | high |
| Congestion resolution | Arbitration | Buffering, flow control | Connection setup | Buffering, FC, setup |
| Path selection | Out of band addressing | In/out of band addressing | In band addressing | In band addressing |
| Additional potential / constraints from reconfiguration | • Is there a technical/economical need for additional reconfiguration flexibility beyond addressing?<br>• What are implications for the application from reconfiguration at interconnect level? | | | |