MPSOC 2011 BEAUNE, FRANCE

imec

BOADRES: A SCALABLE BASEBAND PROCESSOR TEMPLATE FOR Gbps RADIOS

RUDY LAUWEREINS

**VICE PRESIDENT, CHAIRMAN OF THE TECHNOLOGY OFFICE
PROFESSOR AT THE KATHOLIEKE UNIVERSITEIT LEUVEN**

imec

STATUS SDR BASEBAND MODEM IN 2008

BEAR platform
- TSMC 90nm GP
- 3x Digital front end (VGA, filter, synchronization ASIP)
- 2x ADRES CGA instance for inner modem
- 1x ARM for inter-tile comm. and control
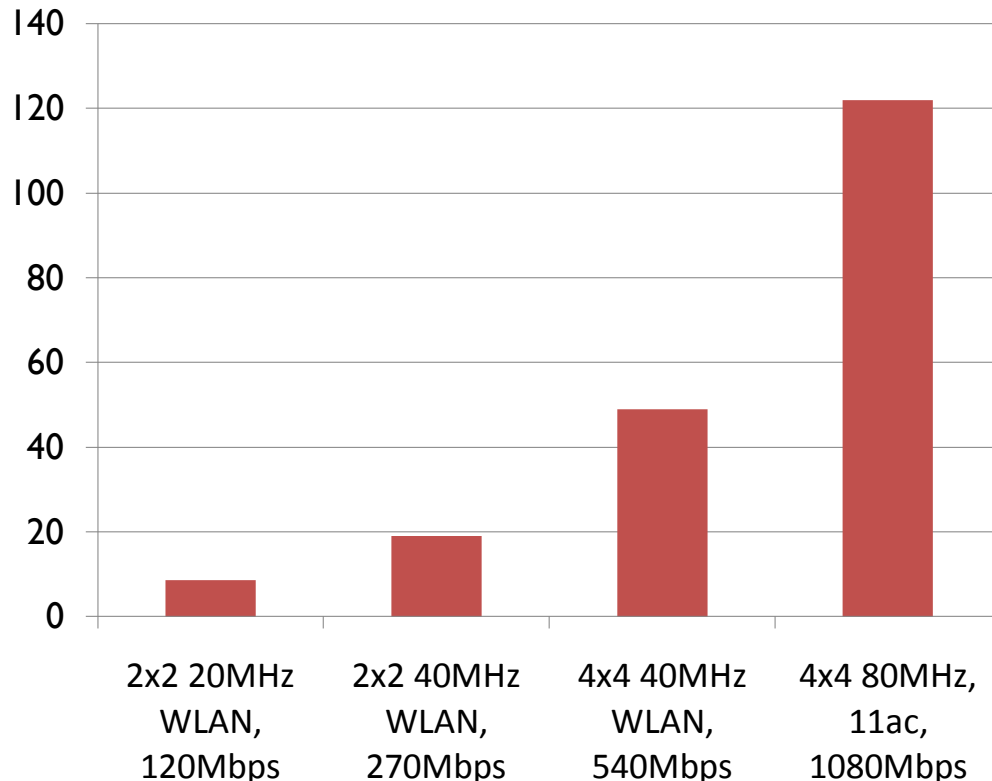- Supports
    - 11n 2x2 40MHz (>200Mbps)
    - LTE 2x2 5MHz

ADRES instance
- Single thread
- 1 3-issue VLIW
- 4x4 FUs
- 64-bit FU supporting SIMD-4 (i.e. 4x 16-bit real)
- 400MHz clock

HOW MUCH COMPUTE POWER IS NEEDED TO SUPPORT INNER MODEM OF NEW STANDARDS?

m-GOPS = mapped GOPS,
not peak GOPS

Estimated 16-bit m-Gops for inner modem data processing

Category	m-Gops
2x2 20MHz WLAN, 120Mbps	~9
2x2 40MHz WLAN, 270Mbps	~19
4x4 40MHz WLAN, 540Mbps	~49
4x4 80MHz, 11ac, 1080Mbps	~122

More than linear with bit-rate

Similar story for LTE

How to increase compute power?

Increase TLP
- More ADRES cores
- More threads per core

Increase DLP
- Wider SIMD

Increase ILP
- More FUs
- More and more powerful intrinsics
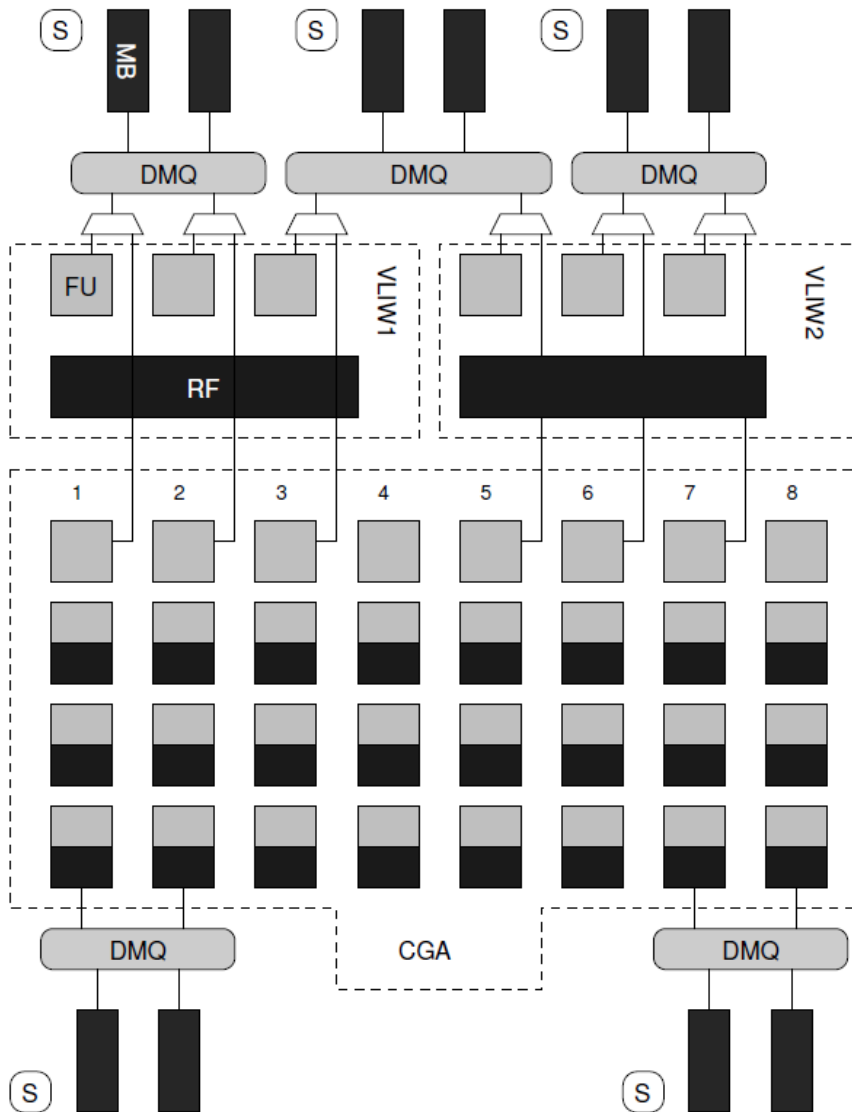
But how much of each?

LEVEL OF AMBITION

Develop a template that scales to at least 1 Gbps, i.e. 11ac and LTE-advanced

Implement a first instance in 40nm technology that supports
- WLAN: 11n 4x4 40 MHz
- Cellular: LTE 2x2 20 MHz
- Fast handover between WLAN and LTE: support two concurrent standards at the same time

EXTEND THE ADRES TEMPLATE


Modifications to old ADRES template:

- Multi-threading: x VLIWs

- Multi-threading: private and shared memory access

- Multi-threading: flexible CGA sub-array allocation to VLIW

- Variable SIMD: not all FU same SIMD

- Variable SIMD: Shuffle FU (pack, unpack, repack)
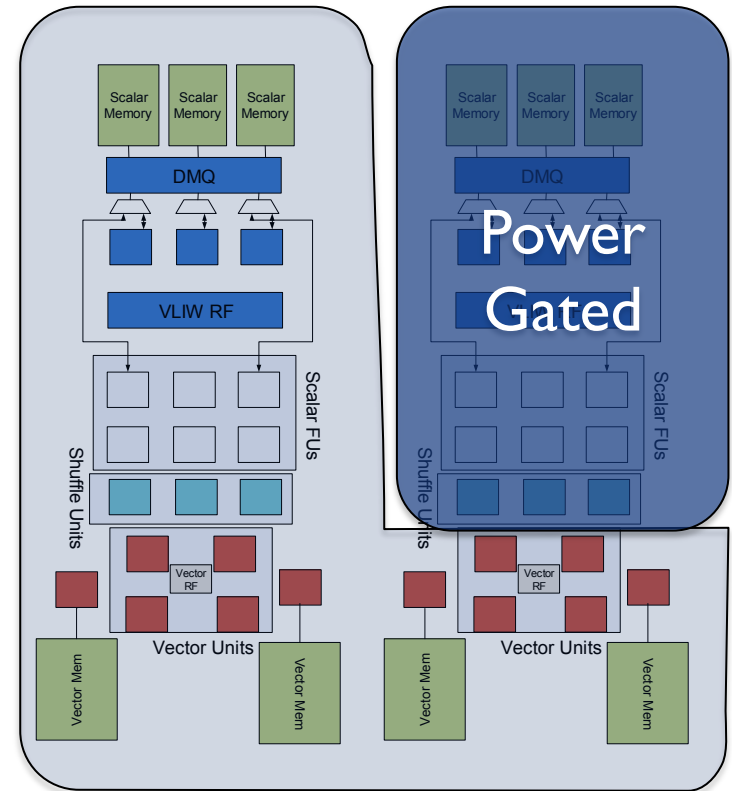
- Dedicated LD/ST FU

DECIDE ON TLP

Stay with 1 core as long as you can:
- Compiler supports all parallelism within 1 core
- For multi-core: you are on your own... With limited tool support

Two concurrent threads to support two concurrent standards with flexible resource allocation
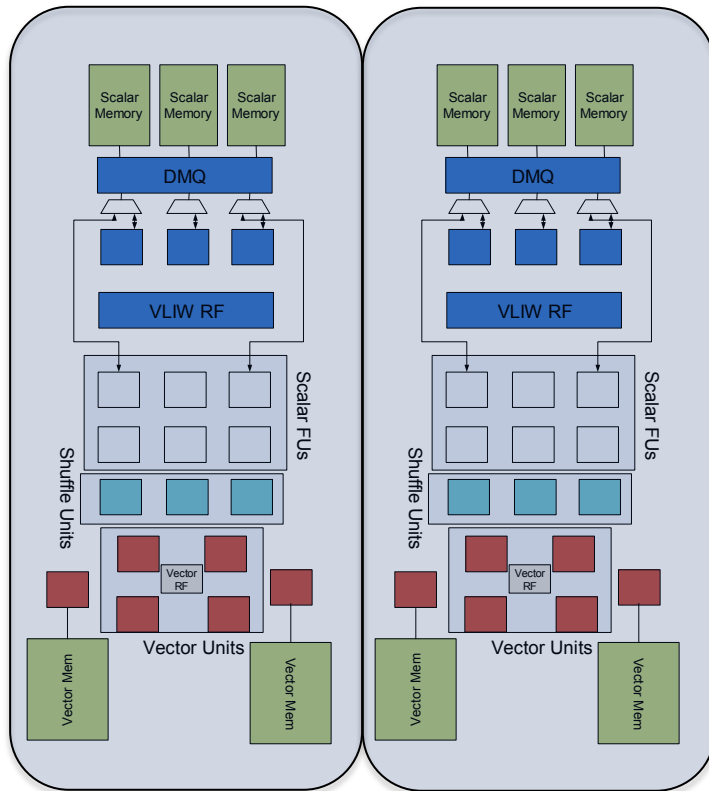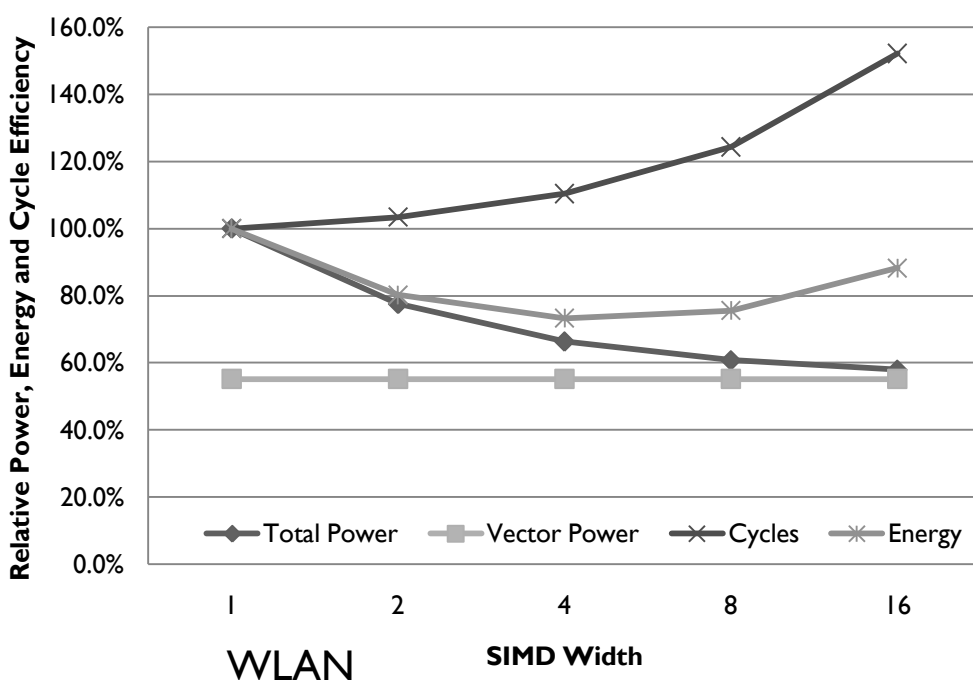

Two concurrent standards

One medium demanding standard

DECIDE ON DLP

WLAN · SIMD Width

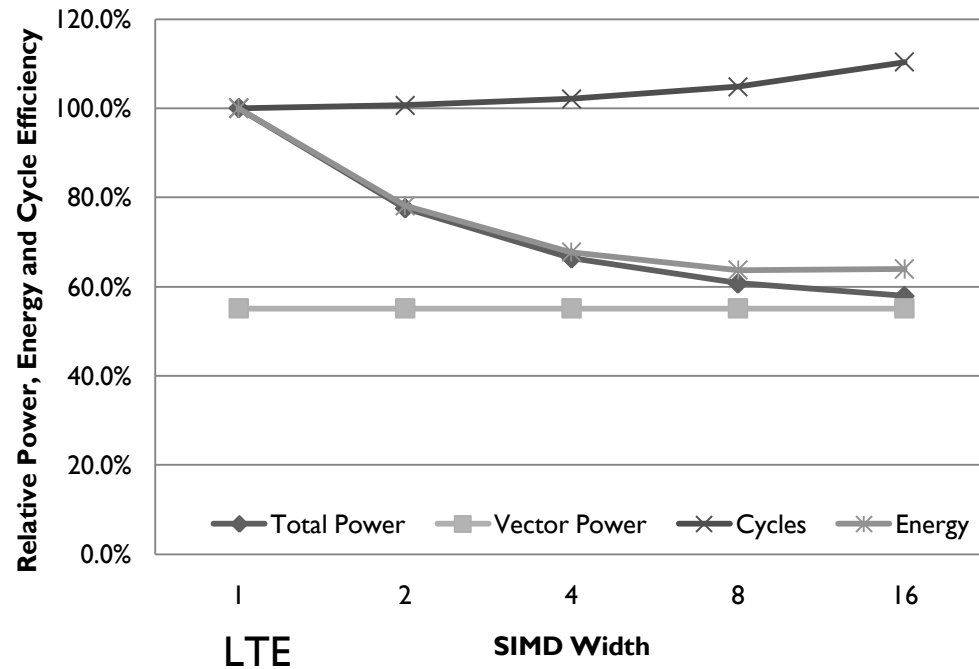Power goes down if vector width increases

Cycles go up due to less efficient software pipelining: less iterations when vectors become wider
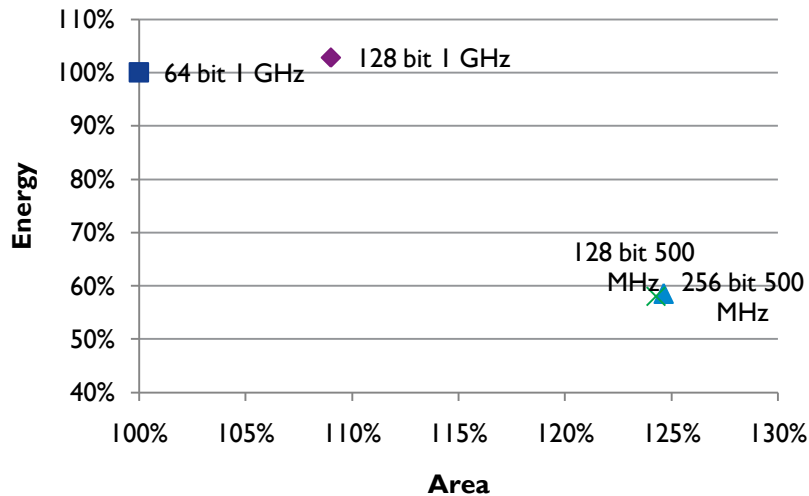
Optimum energy for WLAN: 4

Optimum for LTE: 8 because more loop iterations (1024 for LTE, 128 for WLAN)

Decision: 8
- Optimal for LTE
- Almost optimal for WLAN

LTE · SIMD Width

DECIDE ON FREQUENCY

Energy

110%
100%   ■ 64 bit 1 GHz   ◆ 128 bit 1 GHz
90%
80%
70%
60%   128 bit 500
                                         MHz ▲ 256 bit 500
50%   MHz
40%
     100%   105%   110%   115%   120%   125%   130%

Area

Frequency has a big impact on energy per bit, less on area per bit

SIMD width has no impact on energy per bit, limited impact on area per bit

Energy and area per bit for a pipelined multiplier
- Post physical synthesis, without full place and route
- Pipeline depth adapted to frequency

Decision:
- 500MHz because energy is our primary concern
- 256 bit (SIMD-8 for 2x16-bit complex numbers) because this has no impact on area and energy

DECIDE ON ILP: NUMBER AND TYPE OF FUs
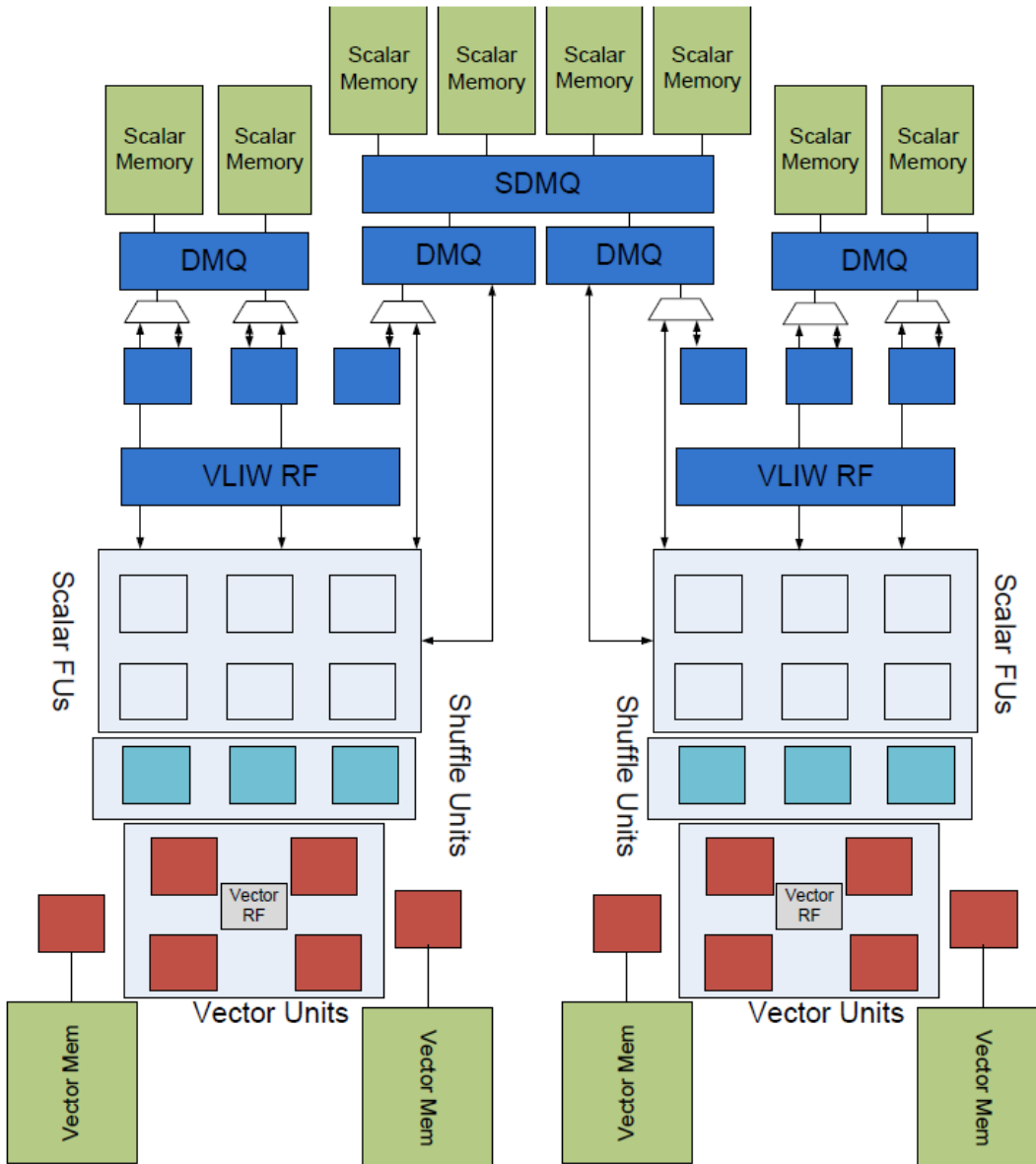
- ▸ Memories
  - 2 scalar (2 banks each)
  - 1 global scalar (4 banks)
  - 4 vector (1 bank each)
- ▸ Units
  - 2x 3-issue VLIW
  - 2x6 scalar: negligible hence no detailed analysis
  - 2x3 pack
  - 2x4 vector: meet performance
 - 2x2 LD/ST: meet data access needs of vector
- ▸ Registers
  - 2 64-deep VLIW
  - 2x2 8-deep vector

DECIDE ON INTRINSICS FOR THE VECTOR UNITS

Table 1: Important extended instructions

Functionality	Targeted signal processing blocks
Complex exp	SCO and CFO compensation,
Angle estimation	SCO and CFO estimation
Reciprocal	Matrix inversion, SCO and CFO estimation, LLR coeficient calculation, etc.
Reciprocal sqrt	Matrix inversion, SCO and CFO estimation, etc.
Soft demapping	Generating soft information
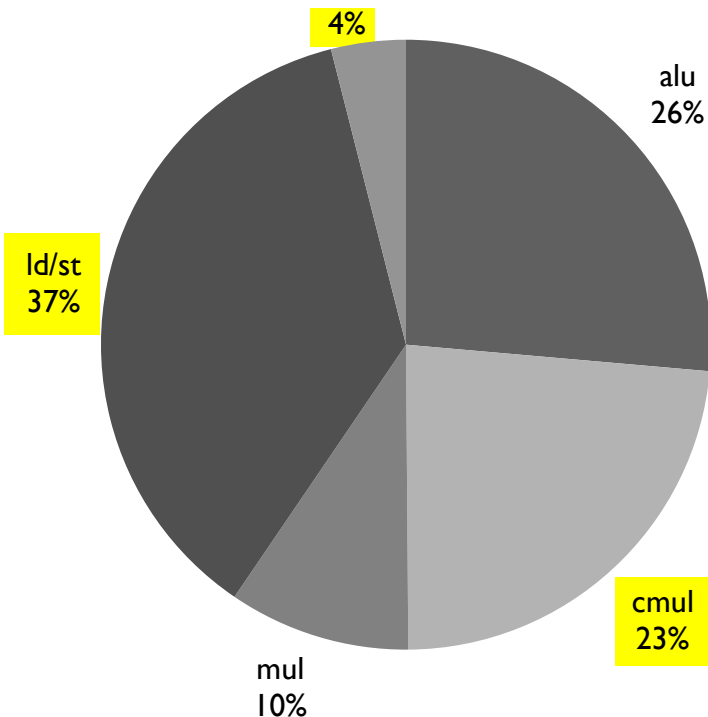
LD/ST is 1/3 of all operations hence 1 LD/ST unit per 2 vector units

Complex multiply is the most important single instruction: was hence already foreseen in the old ADRES as an intrinsic

Only 4%, but:
- In critical path
- Expanding in normal operations would lead to high performance and energy penalty: much higher frequency needed
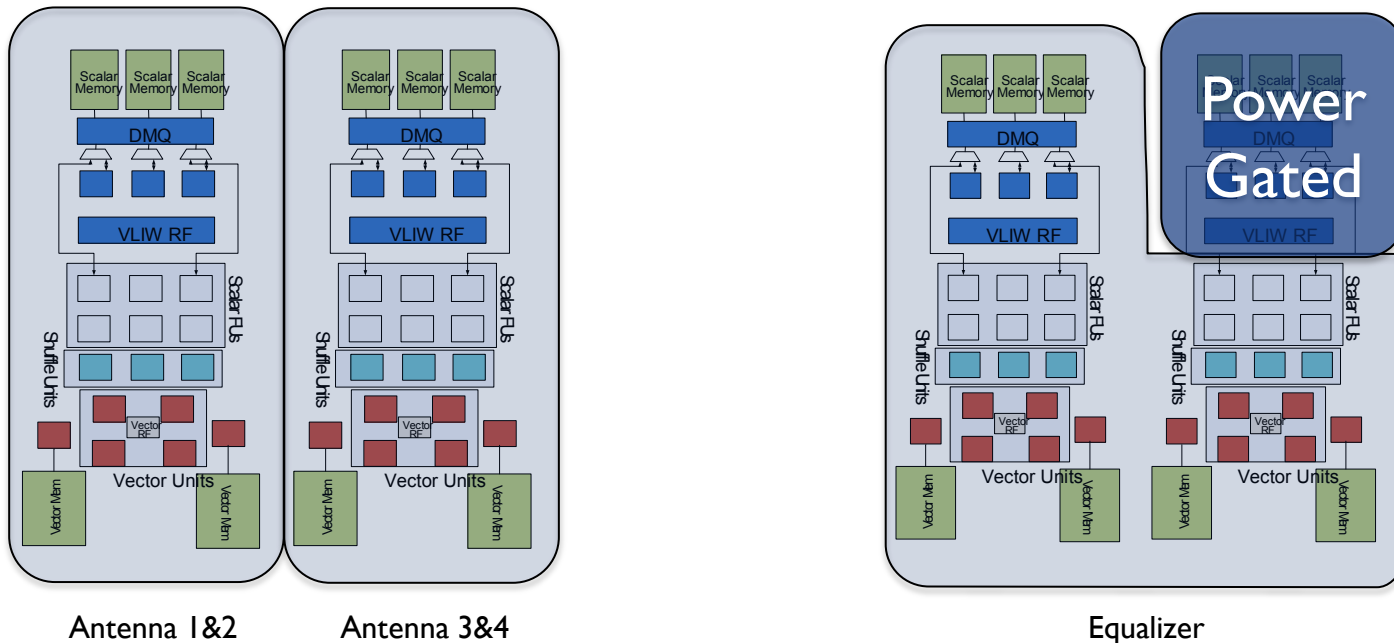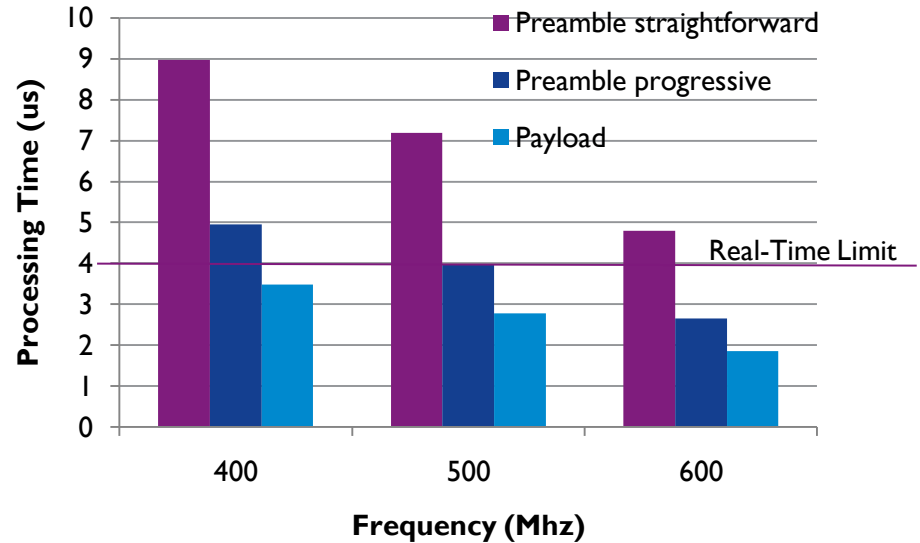
PERFORMANCE FOR WLAN

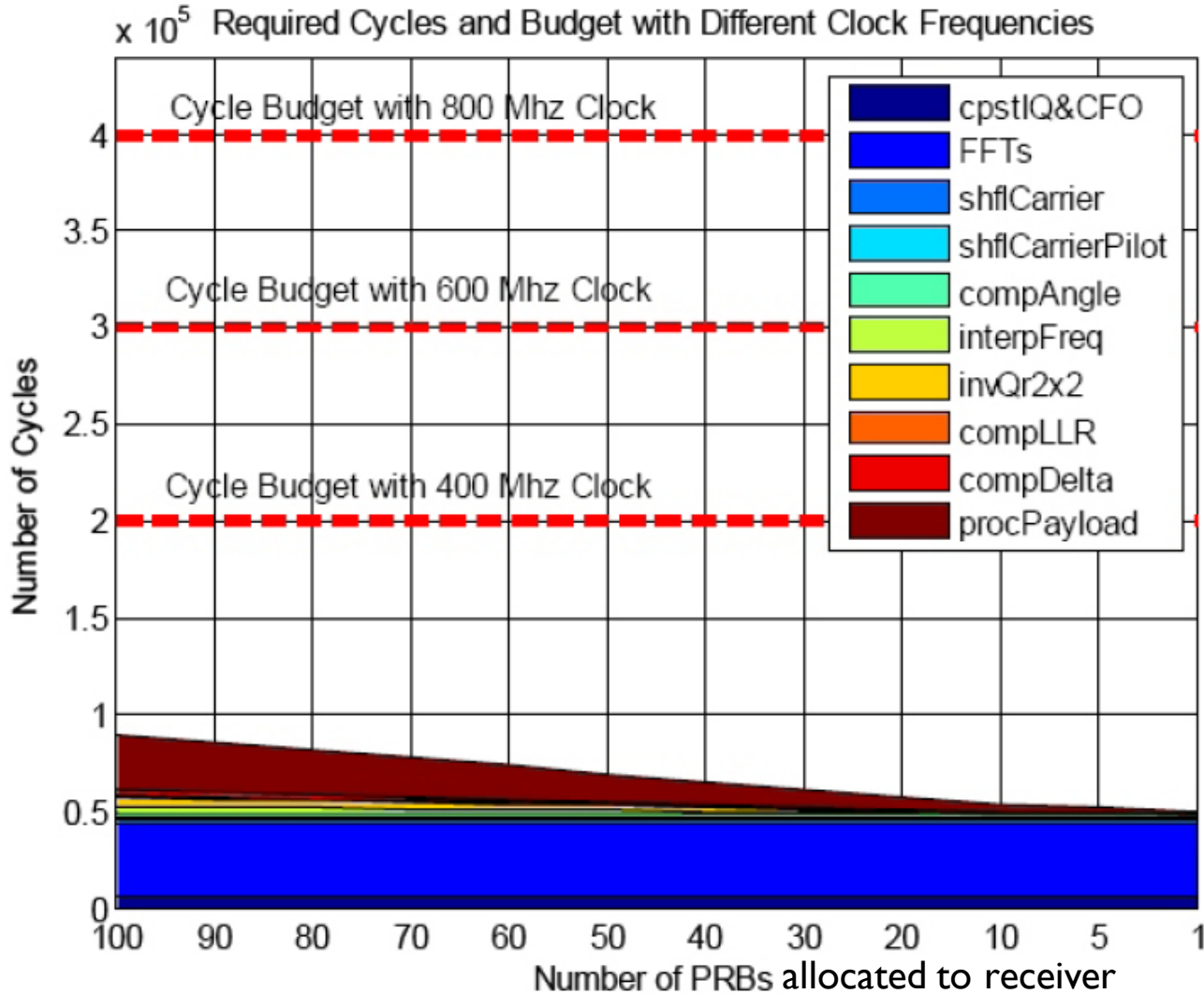Critical timing: 16 μs to process 4 Long Training Fields of preamble

Waiting till all have been received, before starting to process is no option

Progressive computation requires 1.5% more operations, but can meet real-time

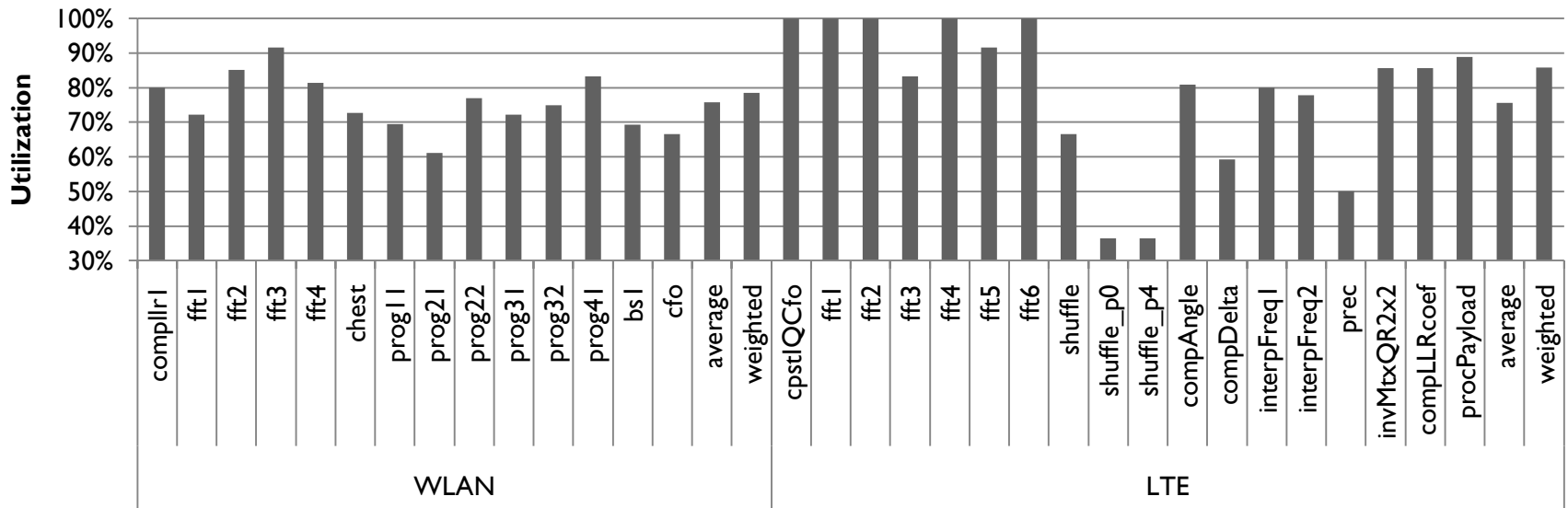Apply TLP and flexible resource allocation

Frequency (Mhz)

Antenna 1&2   Antenna 3&4

Equalizer

PERFORMANCE FOR LTE

Piece of cake, even when the full bandwidth of the base station is allocated to a single receiver

UTILIZATION OF THE VECTOR RESOURCES

Weighted average:
- WLAN: 78%
- LTE: 86%

Balanced architecture
Excellent extraction of parallelism by compiler

CONCLUSION (FIGURES RELATIVE TO 90NM ADRES 2008)

	ADRES 2008 @ 90nm	ADRES 2008 @ 40nm	ST-BOADRES @ 40nm	MT-BOADRES @ 40nm
m-GOPS	**1.0** m-GOPS @ 400 MHz	1.0 m-GOPS @ 400MHz / 1.76 m-GOPS @ 700MHz	4.0 m-GOPS @ 700MHz	8.0 m-GOPS @ 700MHz
Core Area (sqmm)	**1.00** sqmm	0.29sqmm	0.30sqmm (est.)	0.61sqmm (est.)
Target WLAN Standards	2x2 20MHz ~150Mbps	2x2 40MHz ~300 Mbps	4x4 40MHz ~600Mbps	4X4 80MHz ~1000Mbps
Area Efficiency (m-GOPS/sq mm)	**1.0**	6.0	13.2	13.2

8x more powerful
1.7x due to technology
4.7x due to architecture

13x more area efficient
6.0x due to technology
2.2x due to architecture

**ASPIRE
INVENT
ACHIEVE**

imec