High-Performance Benchmarking of the European NG-MEDIUM FPGA

G. Lentaris, (glentaris@microlab.ntua.gr)
K. Maragos,
V. Leon,
D. Soudris, David Gonzalez-Arjona

ECE, NTUA, GREECE

GMV, SPAIN

Put into Context

- major task of *QUEENS-FPGA* project (ESA, 2017-2018)
  - "Quality Evaluation of European New SW for BRAVE"
  - goals: assessment + improvement of programming tools

- "high-performance" → primary purpose of BRAVE
  - high-density FPGA developed for demanding algorithms
  - must be tested with typical DSP benchmarks for space applications (e.g., in payload processing, category "DC3")

- intensive beta testing, based on methodology
  - almost completed (preliminary results today)

Contents

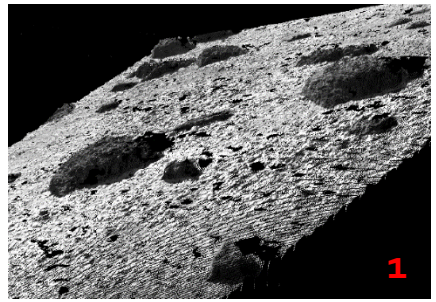
1. benchmark selection
2. synthesis assessment
3. place & route assessment
4. general comments

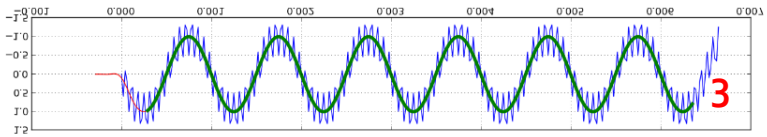
High-Performance Benchmarks

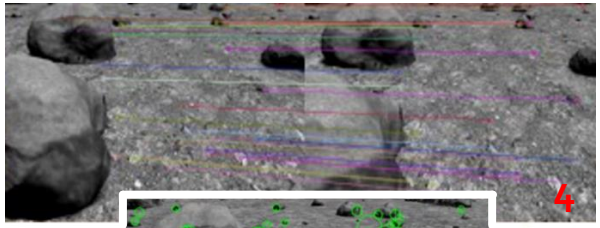
- many available from past ESA activities (in-house)
  - mostly image processing for navigation
  - parametric VHDL, initially on Xilinx FPGAs

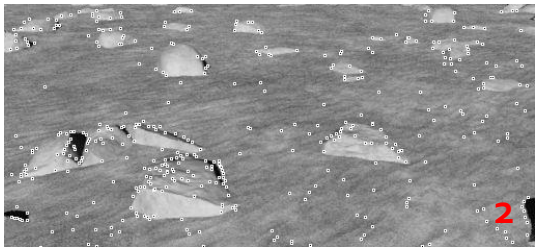
1. depth extraction
2. corner detection
3. filtering
4. feature matching
5. blob extraction
6. edge detection
7. feature description

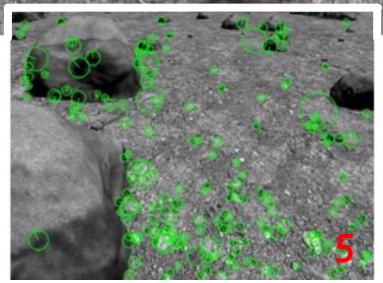
etc. (~15 total)

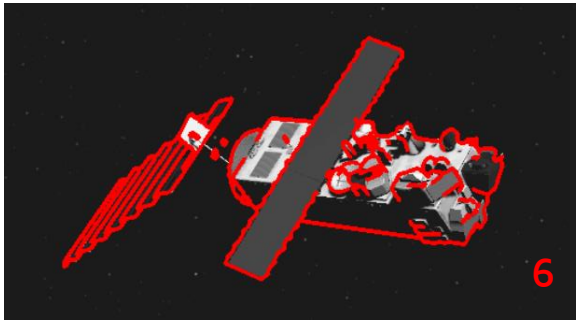


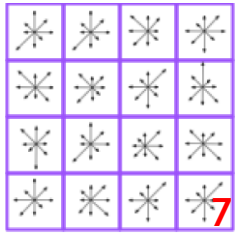


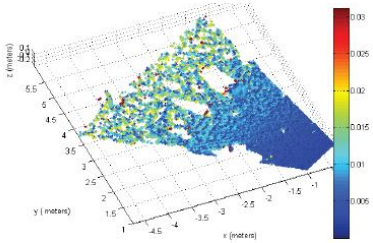






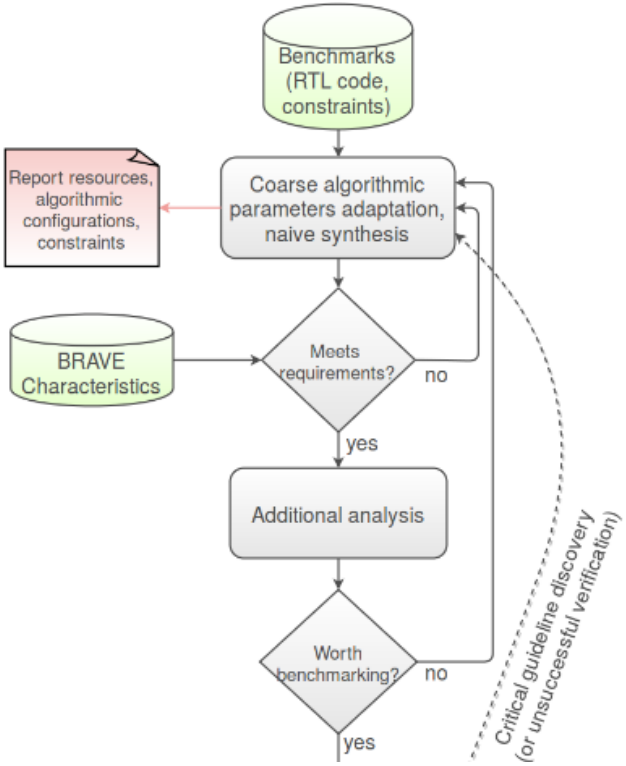






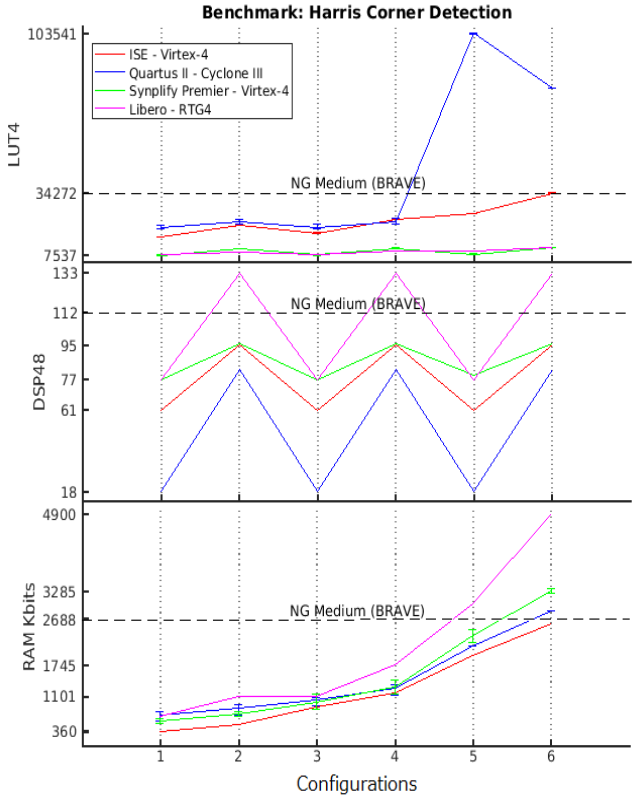
Benchmark Selection

- methodology (fig), to select most suitable for assessing/stressing BRAVE (NG-MEDIUM)
  - start without using *NanoXmap*
  - draw big picture via many 3rd party tools (all major competitors)
- consider multiple criteria
  1. IPs resources vs NGMEDIUM size
  2. parameterization/stressing of IPs
  3. diversity of resource types of IPs
  4. high activity, communication (IO)
  5. complexity of debugging (even at netlist level), demo prospects



Benchmark Selection, Results

- data from 1000+ syntheses
  - by changing parameters of algorithms & tools (7K results)
  - created comparison tables and plots (per tool & benchmark)
- *3 best* = ***Disparity***, ***Harris***, ***FIR***
  - cover all FPGA resource types
  - NG-MEDIUM utilization 11-98%



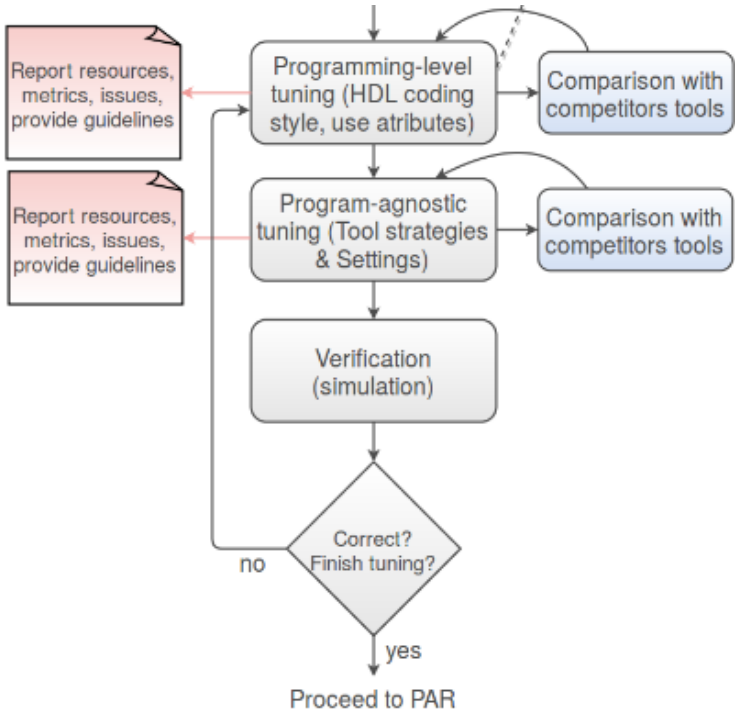
** custom grading 0-3, summarizes all results from 4 tools and 4 FPGAs (3rd party)*

	feasibility	scalability	diversity	throughput	debugging	demo	TOTAL
1. Disparity	**3**	**3**	**2**	**3**	**3**	**3**	**17**
2. Spacesweep	3	2	3	3	1	3	**15**
3. Harris	**3**	**3**	**3**	**3**	**2**	**3**	**17**
4. SURFdet	2	1	2	3	2	3	**13**
5. SIFTdesc	3	1	1	2	1	1	**9**
6. SURFdesc	1	1	3	2	1	1	**9**
7. SIFTmatch	2	1	3	2	3	1	**12**
8. BRIEFmatch	2	1	3	2	3	1	**12**
9. FIR	**3**	**3**	**1**	**3**	**3**	**3**	**16**

Synthesis Assessment

- methodology (fig) to test all *NanoXmap* parameters and compare to 3rd party tools (via CIC abstraction)
- on 6 versions of *NanoXmap* (2.7.1 – 2.8.4, mainly 2.8.0)

- parameter tuning at 2 levels
  - benchmark agnostic (generic)
  - VHDL/algorithm (insightful)
- netlist verification (*QuestaSim*)
- note: besides DC3, it becomes imperative to use small circuits



Synthesis Assessment, Results (½)

- collected details/data from 500 runs of *NanoXmap*
- tested parameters: *Mapping Effort*, *MaxRegisterCount*, *Timing Driven*, *MergeRegisterToPad*, *DefaultROMMapping*, *DefaultRAMMapping*, *LessThanTo DSPMapThreshold*, *DefaultFSMEncoding*, *AdderToDSPMapThd*, *MultiplierTo DSPMapThreshold*, *LessThanToDSPMapThreshold*, *AddMappingDirective*

Synthesis Assessment, Results (½)

- collected details/data from 500 runs of *NanoXmap*
- tested parameters: *Mapping Effort*, *MaxRegisterCount*, *Timing Driven*, *MergeRegisterToPad*, *DefaultROMMapping*, *DefaultRAMMapping*, *LessThanTo DSPMapThreshold*, *DefaultFSMEncoding*, *AdderToDSPMapThd*, *MultiplierTo DSPMapThreshold*, *LessThanToDSPMapThreshold*, *AddMappingDirective*

- few were non-responsive (still under development)
  - ROM mapping, ADD/MULT mapping per component, …
  - MAC on 1-DSP, big MULT on 1-DSP, few RAMB config., …
- many allowed the user to drive synthesis correctly
  - mapping all to RAM or Register File, FSM encoding, …
  - reasonable map on LUT/CARRY/RF (w.r.t. architecture)

Synthesis Assessment, Results (2/2)

- comparison to 3rd party tools: "*averaged*" vs "*tailored*"
  - *averaging* on 3rd tools for abstraction, single reference
  - *tailoring* on NanoXmap to balance the FPGA resources
    - e.g., decrease ~5x LUT/CARRY (overutilized) via 40-70 DSP

Synthesis Assessment, Results (2/2)

- comparison to 3rd party tools: "*averaged*" vs "*tailored*"
  - *averaging* on 3rd tools for abstraction, single reference
  - *tailoring* on NanoXmap to balance the FPGA resources
    - e.g., decrease ~5x LUT/CARRY (overutilized) via 40-70 DSP

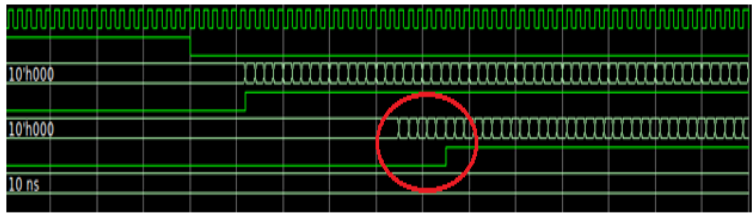
- competitive w.r.t. resources, room for improvement
  - logic: **+5% LUTs** (+123% incl. pass-through), **+27% DSPs**
  - memory: **+67% RAM bits** (fewer blocks!), **+42% DFFs**
  - but with spikes: e.g., +128% DFF, +275% DSP, +465% LUT

Synthesis Assessment, Results (2/2)

- comparison to 3rd party tools: "*averaged*" vs "*tailored*"
  - *averaging* on 3rd tools for abstraction, single reference
  - *tailoring* on NanoXmap to balance the FPGA resources
    - e.g., decrease ~5x LUT/CARRY (overutilized) via 40-70 DSP

- competitive w.r.t. resources, room for improvement
  - logic: **+5% LUTs** (+123% incl. pass-through), **+27% DSPs**
  - memory: **+67% RAM bits** (fewer blocks!), **+42% DFFs**
  - but with spikes: e.g., +128% DFF, +275% DSP, +465% LUT

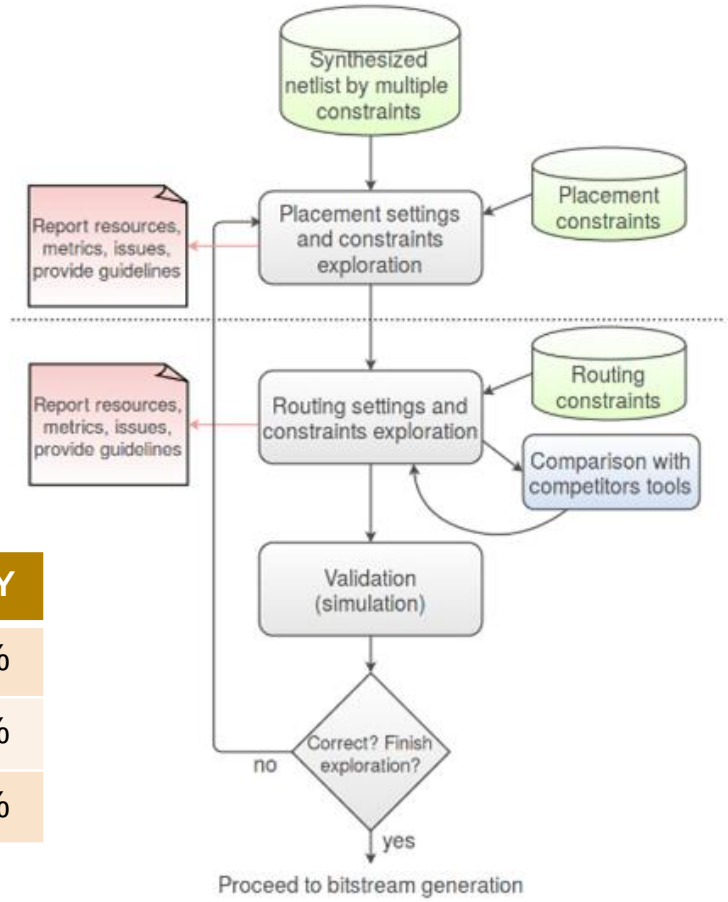
- long test of netlist correctness
  - few problems in simulation
  - synthesizer's bugs corrected



Place & Route Assessment

- similar to synthesis, but fewer *NanoXmap* parameters

- examine separately **Placement** and **Routing** and **Timing** closure
  - fine-tune VHDL to fit in NGMED
  - stress tool with diverse & high utilization, and time constraints



Benchmark	LUTs	DFFs	RAMB	DPSs	CARRY
FIR C2	33,8%	33,7%	0	57,1%	15,6%
Disparity C1	13,7%	10,8%	55,4%	40,2%	19,0%
Harris C3'	54,4%	40,9%	98,2%	65,2%	73,4%

Place & Route Assessment, Results

- almost same resources from synthesis to P&R (small Δ)
  - reasonable: behavior of 3rd party tools is similar (small Δ)

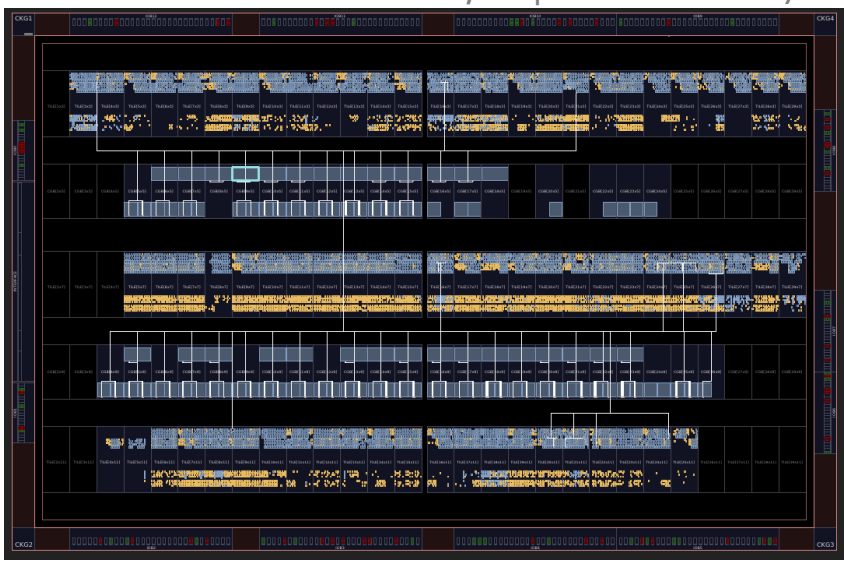
Place & Route Assessment, Results

- almost same resources from synthesis to P&R (small Δ)
  - reasonable: behavior of 3rd party tools is similar (small Δ)
- '*timedriven*': 10-20% f_{clk} boost with same resources on FPGA, but ~60% increase of P&R time
- cmp'd to 3rd tools: less than half *max* f_{clk} (~40%), but in good range (50-110 MHz)

Place & Route Assessment, Results

- almost same resources from synthesis to P&R (small Δ)
  - reasonable: behavior of 3rd party tools is similar (small Δ)
- '*timedriven*': 10-20% f_{clk} boost with same resources on FPGA, but ~60% increase of P&R time
- cmp'd to 3rd tools: less than half *max* f_{clk} (~40%), but in good range (50-110 MHz)
- correctness of netlist
  - FIR @ post-PAR @ 110 MHz
- floorplan view: no control, but helps understanding and guiding synthesis

NG-MEDIUM, floorplan view of "Harris C7"

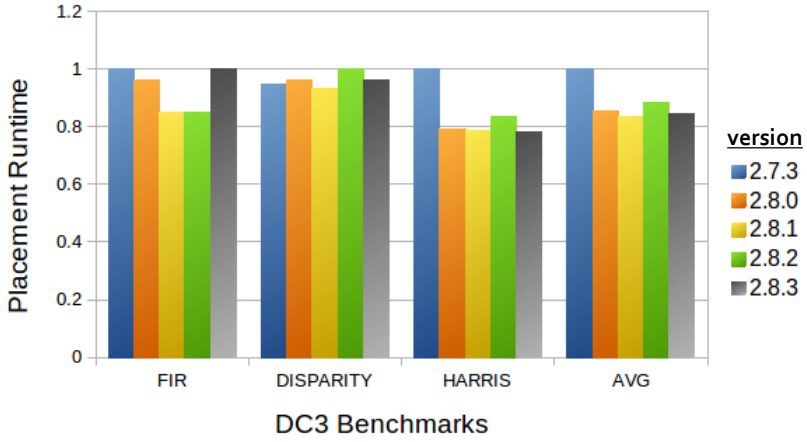


General Comments (1/3)

- very lightweight tool
  - fast: 1–8 minutes total (e.g., *Synth*=1min, *P&R*=3+3min)
    - ~twice faster than competitors
  - low memory footprint (half vs competitors)
  - no installation needed (465MB)
- user-friendly
  - fewer options than 3rd tools (not good for power users)
  - simple GUI, instead provides flexible Python scripting
- useful report files
  - 11 resource types, analyzed (e.g., LUTs due to DFF or CY)
  - info during process (FSM analysis, optimization steps, etc.)

General Comments (2/3)

- considerable progress through *NanoXmap* versions
  - 20 SPRs (most addressed)
  - many issues corrected
    - *Mapping Directives*
      - assign DSPs per component
      - ROM to memory resources
    - *Synthesis* of big memories
  - improved the report files

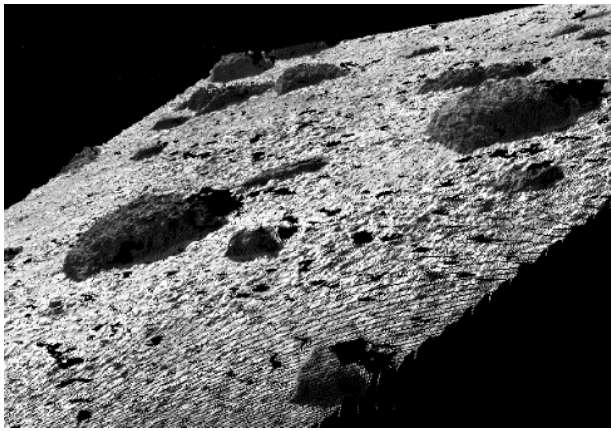


- preparing/assisting in guidelines for 'best practice'
  - e.g., VHDL style in '*read+write' RAMB* (not concurrently)
  - e.g., VHDL style for *multiplications* in corner cases

General Comments (3/3)

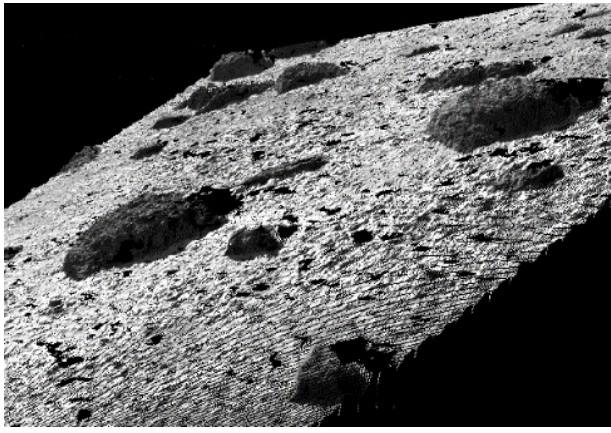
- so, even today, with NG-MEDIUM
  - can accelerate *stereo* (most intensive)
  - the European FPGA would improve *Disparity* of rovers now on Mars
    - **2x** w.r.t. accuracy
    - **10x** w.r.t. speed

    (rough estimation)



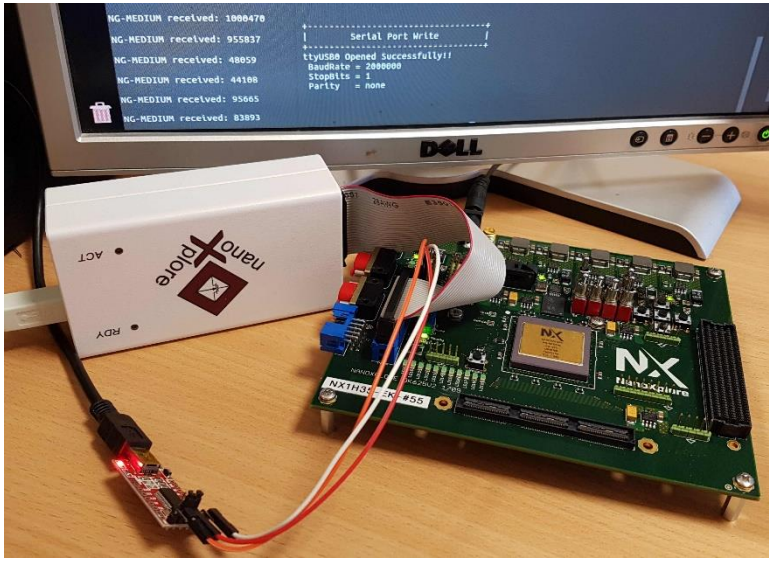
General Comments (3/3)

- so, even today, with NG-MEDIUM
  - can accelerate *stereo* (most intensive)
  - the European FPGA would improve *Disparity* of rovers now on Mars
    - **2x** w.r.t. accuracy
    - **10x** w.r.t. speed



- ongoing work
  - assess HW of BRAVE (as user)
  - develop serial/parallel IO's
  - connect to PC for HW/SW co-processing & DC3 demos
    - FIR already working @2Mbps

➢ entire toolchain verified now!



Thank You!
Questions?

George Lentaris, ECE, NTUA, Greece. glentaris@microlab.ntua.gr