

# Near real-time focusing of Sentinel-1 IW data with GPGPU

Achille Peternier, Paolo Pasquali  
sarmap SA, Switzerland

Raffaele Vitulli  
ESA/ESTEC, The Netherlands

*ESA project 3-13467/12/NL/FF/fk*

## Remote Sensing: the Big Data issue

- New sensors such as Sentinel-1 or ALOS-2:
  - High resolution.
  - Short revisit time.
  - Free availability (e.g., through ESA's SciHub).
- New algorithms and techniques:
  - Multi-temporal Analysis (SBAS, PSInSAR):
    - Using dozens of high-resolution images.
  - Larger series of images.

 **Big Data requires big computational power**

## Project goals

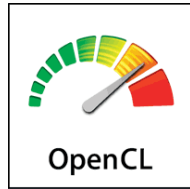
- The SARIPA project (**SARscape Image Processing Accelerator**): focusing of ENVISAT and Sentinel-1 images in near real-time.
- Given a high-end reference machine, migrate the SARscape ENVISAT and Sentinel-1 (to be created) focusing pipelines to GPGPU for better performance.
- Time to beat:
  - 1x ENVISAT frame -> **less than 15 s**
  - 100 km of Sentinel data -> **less than 120 s**

## Project goals

- Reference machine is a high-end server-class computer with:
  - 2x Intel Xeon 8 core CPUs
  - 64 of RAM (DDR3)
  - 4x NVidia Tesla K20 GPUs (5 GB of VRAM each)
  - 2x standard SSD HDs (500 MB/s)
  - 1x high perf. SSD HD (2'000 MB/s)
  - Windows/Linux



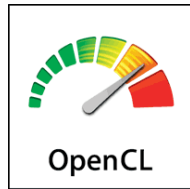
# OpenCL



- Open Computing Language.
- “Write once, run everywhere” philosophy:
  - Runs on CPUs, GPUs, accelerators, FPGAs, etc.
- Writing parallel algorithms is still difficult, but APIs and code-portability are now much easier.

*A. Peternier, M. Defilippi, P. Pasquali, A. Cantone, R. Krause, R. Vitulli, F. Ogushi, and A. Meroni. Performance analysis of GPU-based SAR and Interferometric SAR image processing. In Proc. of the 4th Asia-Pacific Conference on Synthetic Aperture Radar (AP SAR), sep 2013.*

# OpenCL

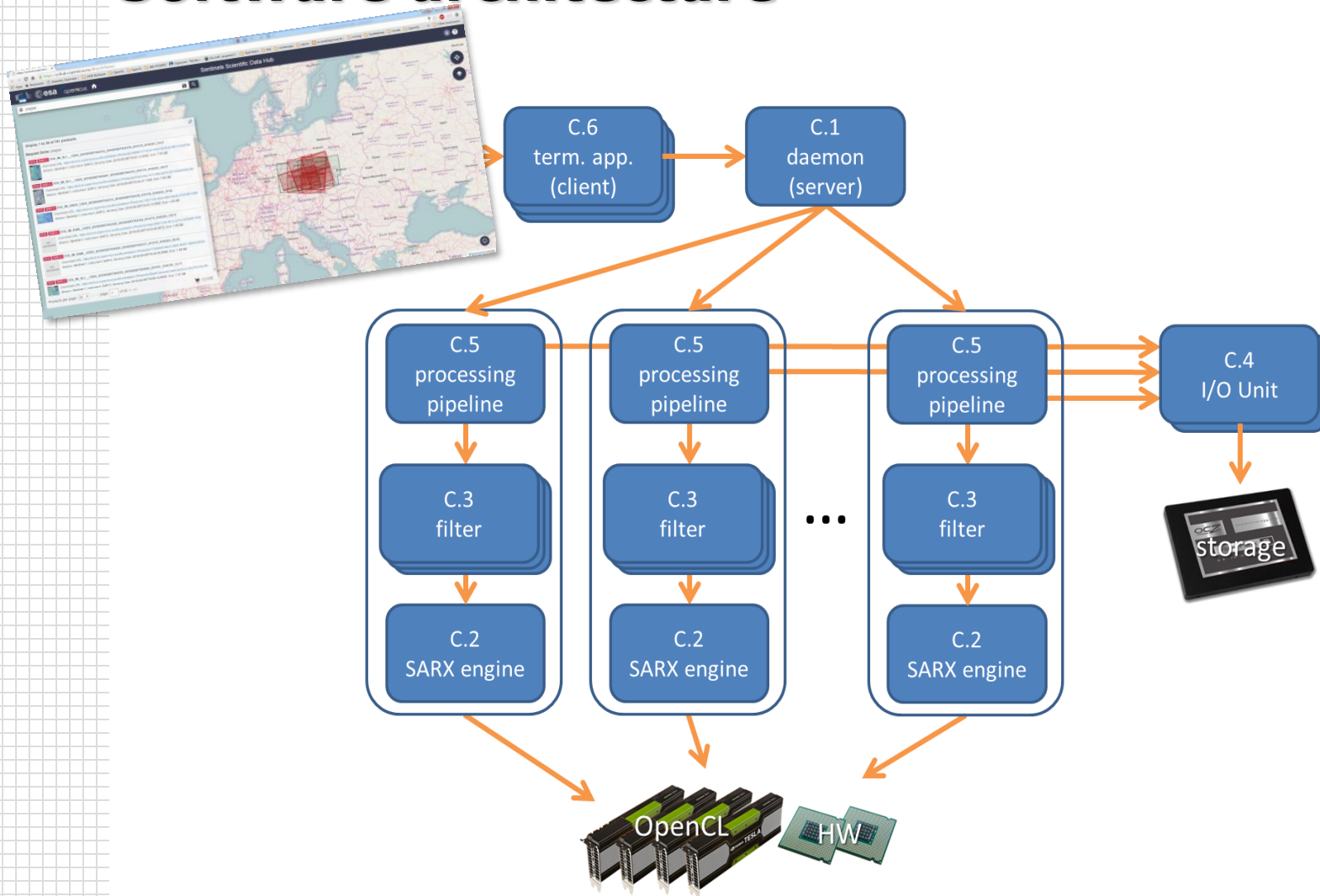


- Mature ecosystem (now):
  - Documentation, tutorials, examples, books, ...
  - Development tools (debuggers, profilers, ...)
  - External libraries (BLAS, FFT, abstraction layers, ...)
- Limitations (of GPGPU in general):
  - Moving data to/from system/OpenCL is expensive.
  - OpenCL memory is a fraction of the system memory.
  - Requires an OpenCL platform to be first installed on the target system.
  - Code must be ported or rewritten.

## Software architecture

- SARIPA is a compact local-cluster abstraction layer on top of OpenCL, Boost, and in-house libraries.
- Automatic features:
  - Multiple parallel pipelines.
  - Hardware awareness:
    - Hardware discovery and auto-configuration, etc.
  - Self-tuning (e.g., according to available memory and number of cores).
  - Deferred I/O operations.

# Software architecture





# Software architecture

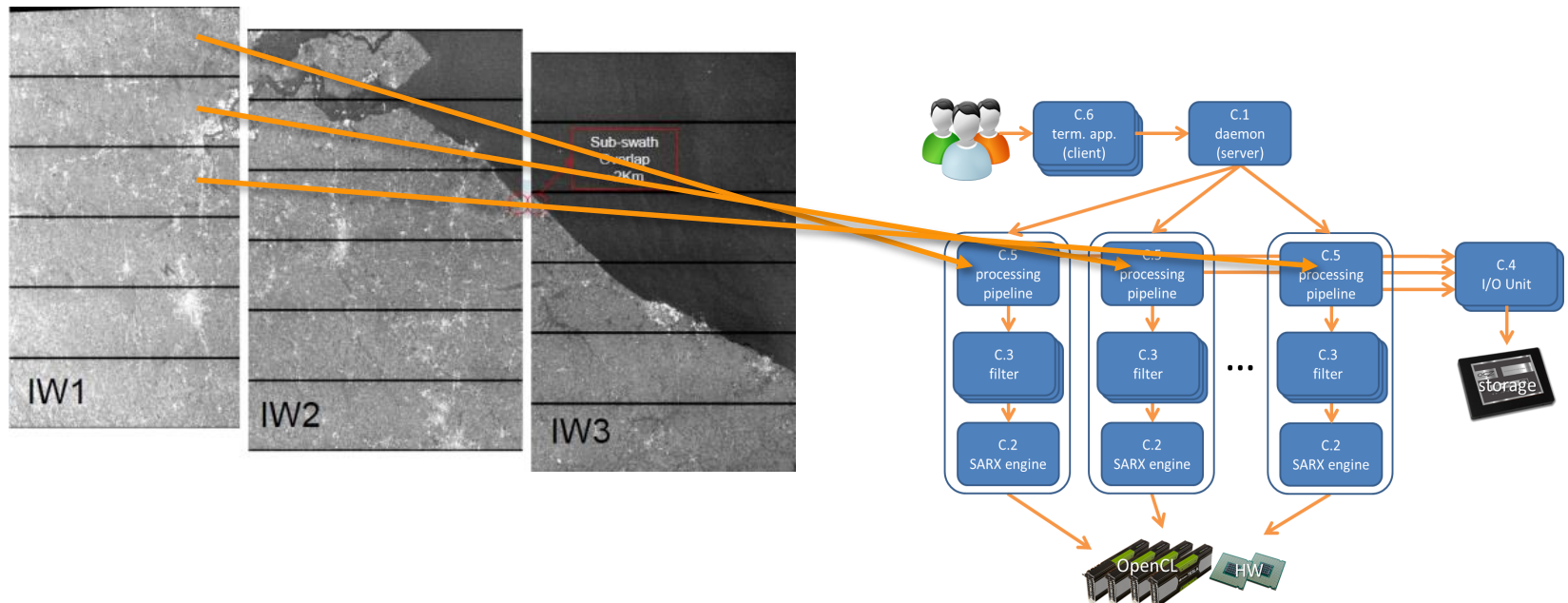
- Two processing pipeline programs:
  - RAW file decoding:
    - Input: **L0 raw file**
    - Output: **N raw bursts** (as float complex buffers)
  - Burst file focusing:
    - Input: **1 raw burst data**
    - Output: **focused burst**
    - Operations:
      - Effective velocity estimation
      - Range compression and DC estimation
      - Pre-filter with Unfolding and Resampling (UFR)
      - Omega-k Azimuth compression
      - Post-filter (UFR)
      - File writing

## Software architecture: L0-decoding

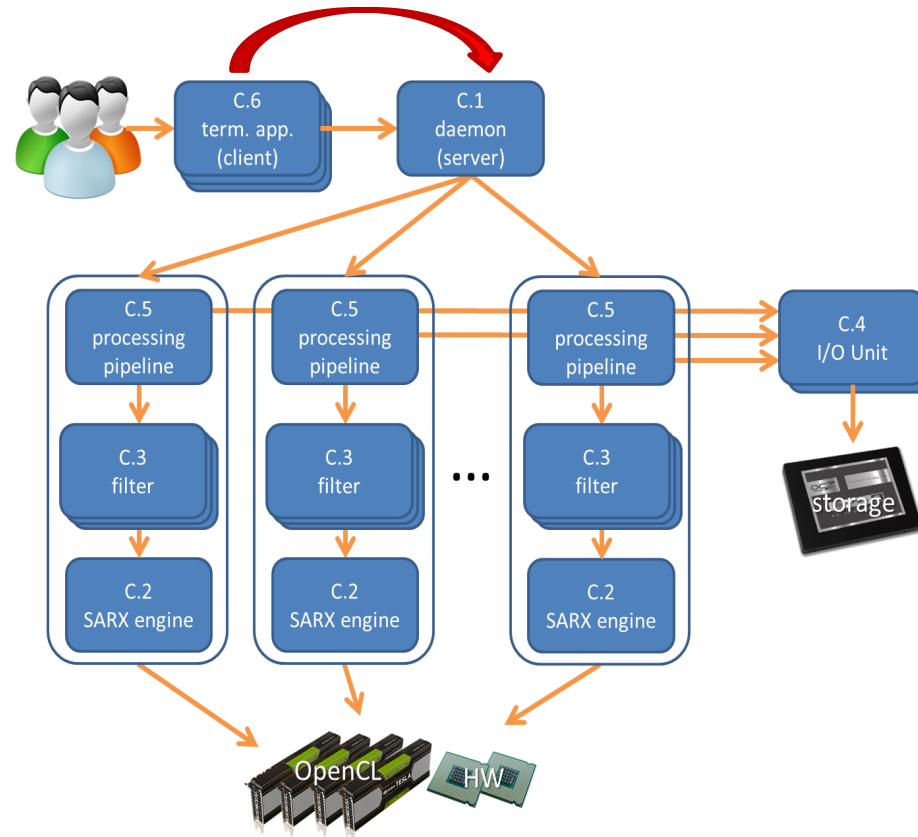
- Written from scratch based solely on the specifications and sample data.
- Parallel CPU-only version:
  - Intensive string processing and bitwise operations.
  - Better memory caching (CPU L2-L3).
- 1 burst per thread, 1 thread per core.
- Compact: ~2000 lines of code.
- Fast: **less than 12 s** to generate **8 GB** of data (the size of a DVD movie!)

# Software architecture: burst focusing

- Structured into swaths and bursts:
  - Easy macroscopic parallelism level: one burst per processing pipeline.

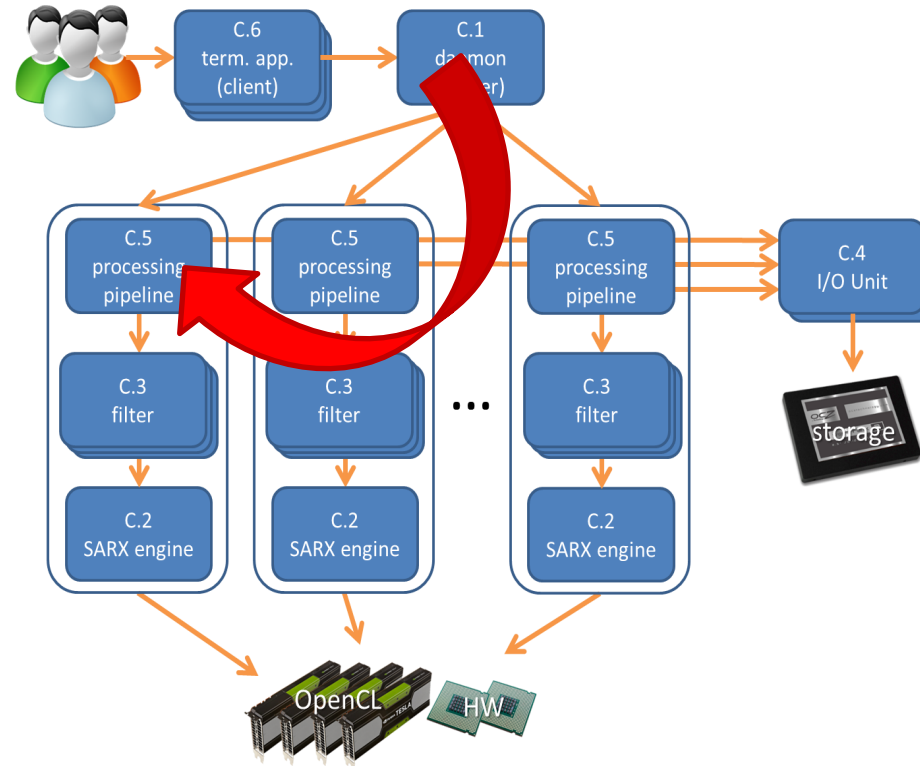


1)

`sarx_console sentinel_1img_n1.txt`

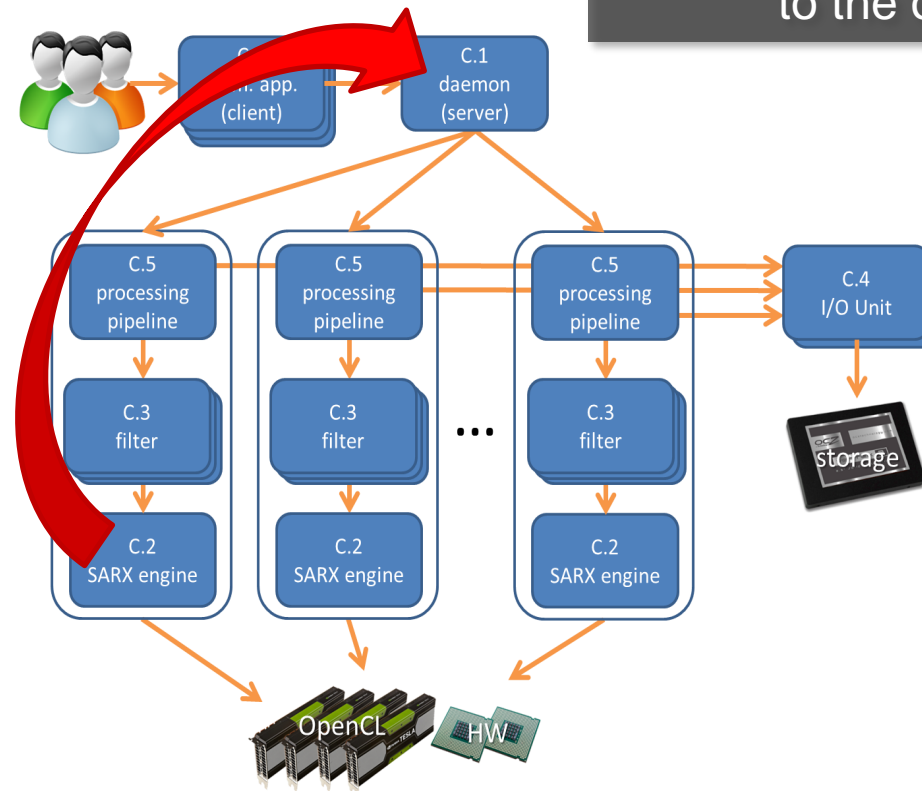
2)

Decoding is sent to the first available pipeline



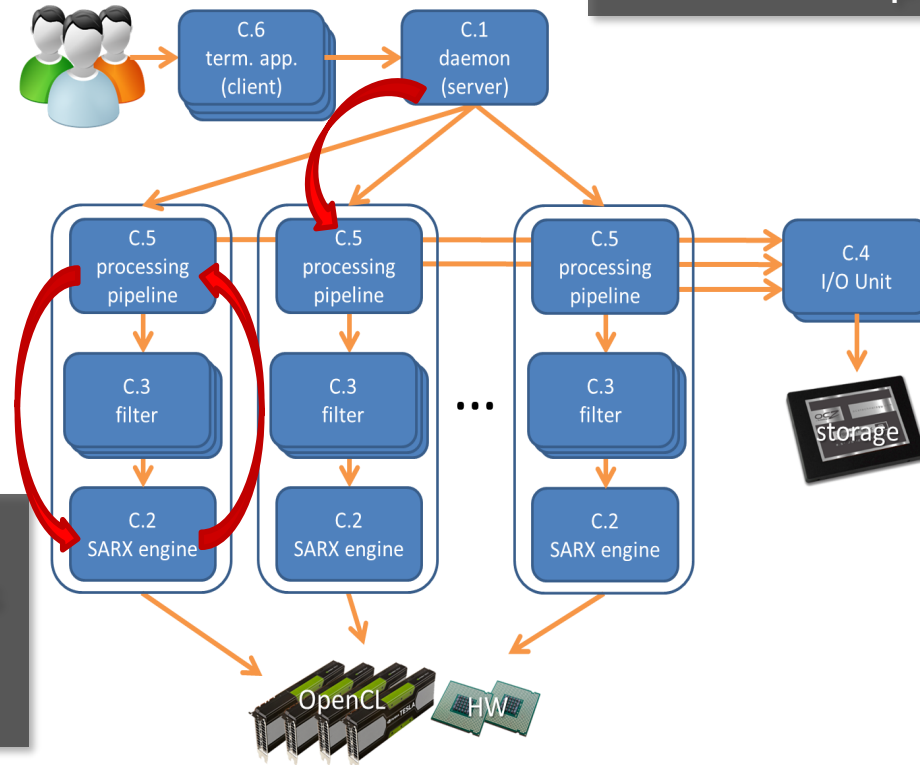
3)

As soon as one burst is decoded, a focusing command is sent back to the orchestrator



4)

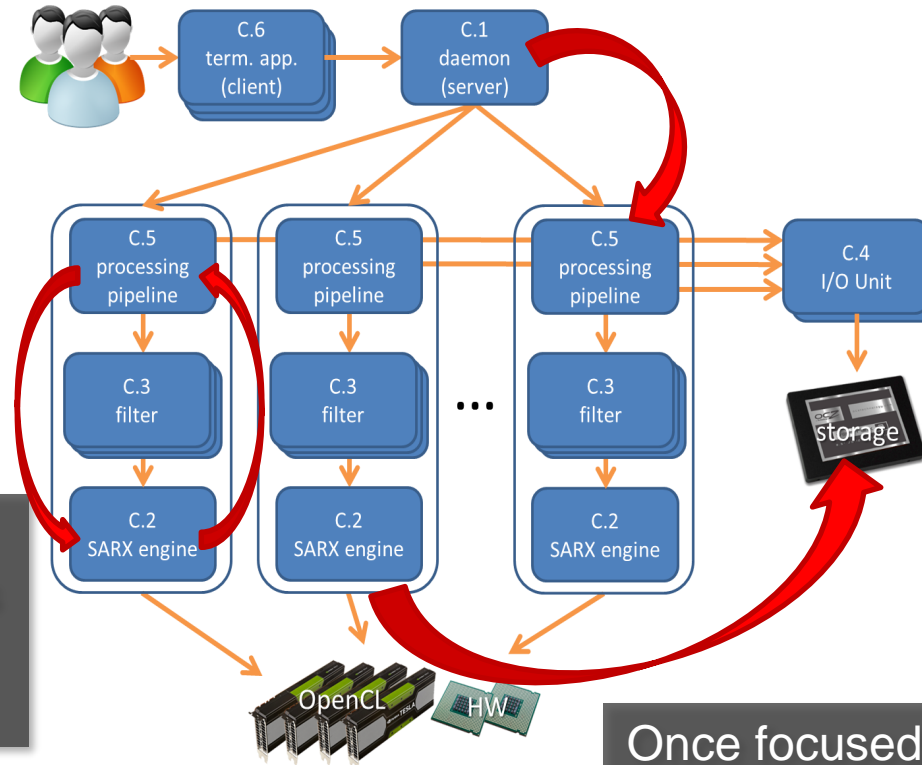
The first burst focusing request is queued to the next available pipeline



In the meanwhile, the first pipeline continues to decode and queue bursts

5)

New burst focusing requests are queued to the other pipelines

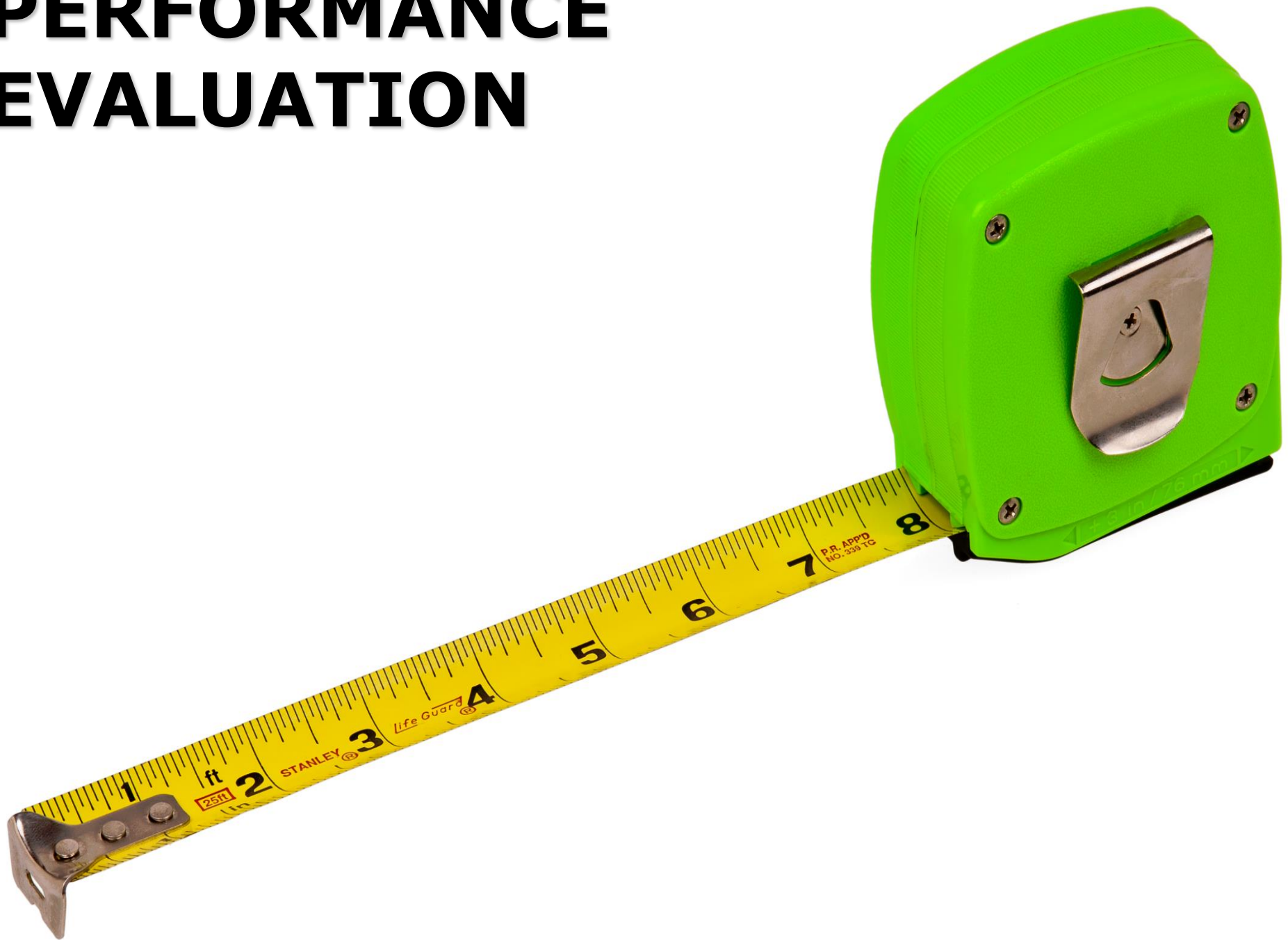


In the meanwhile, the first pipeline continues to decode and queue bursts

Once focused, the burst is directly stored to the disk



# PERFORMANCE EVALUATION

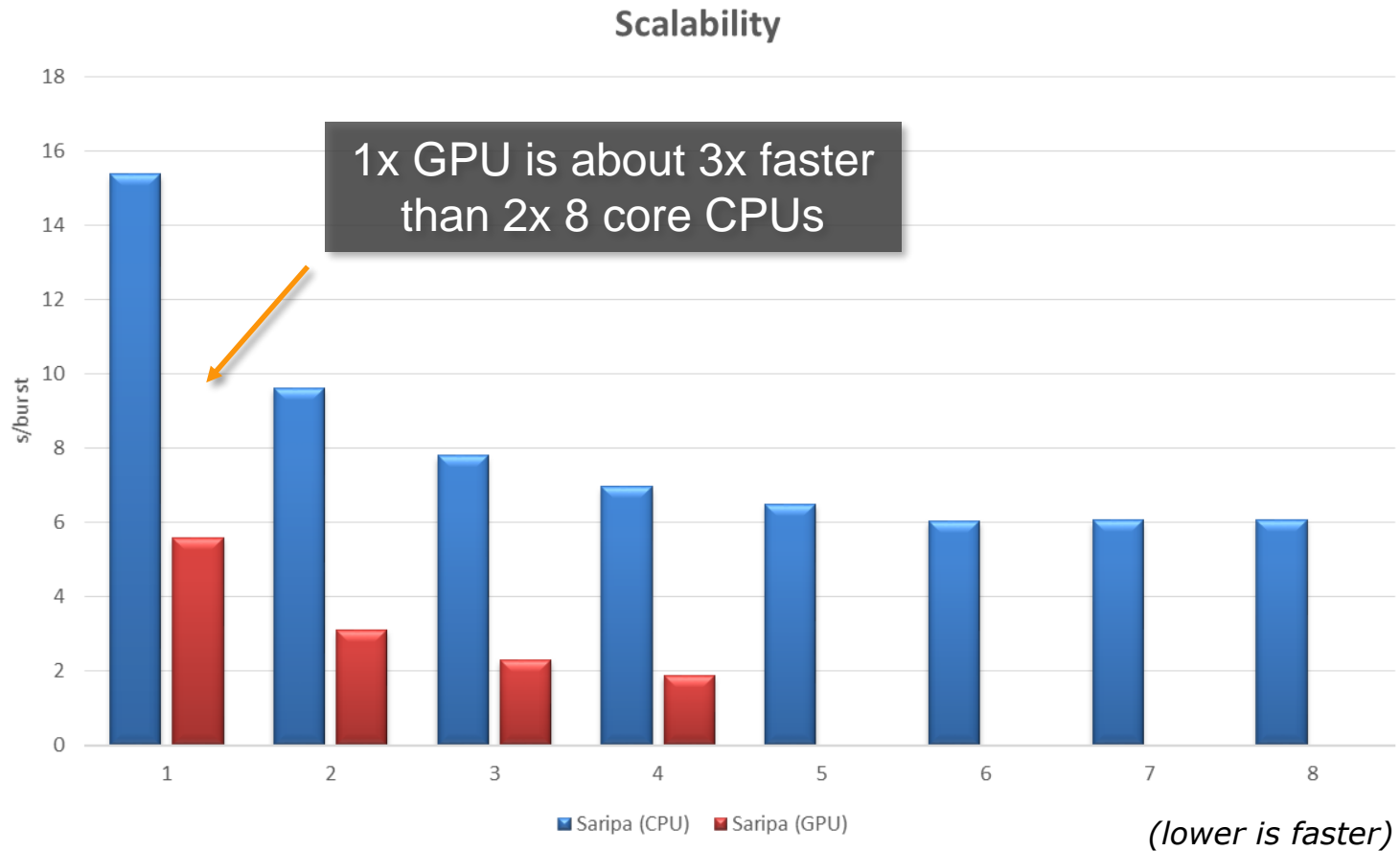


## Time to beat (100 km = 120 s)

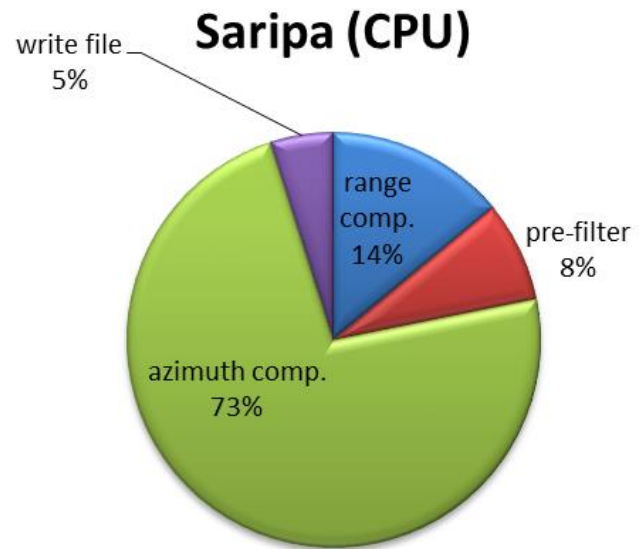
- On the reference machine:
  - 1 polarization full-frame ( $\sim 200$  km) focused in **about 65 s** (4 pipelines on 4 GPUs) and **211 s** (6 pipelines on 2 CPUs):
    - The requirement was to stay under 240 s.

# Focusing scalability

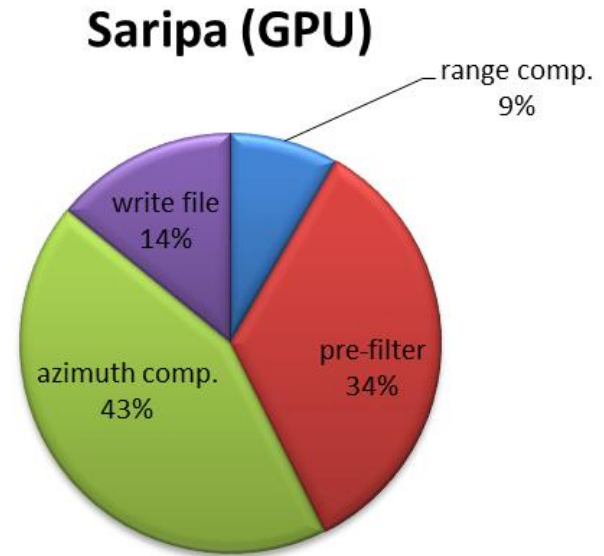
(time to focus 1 burst)



# Focusing steps



(tot. time = 15.5 s)

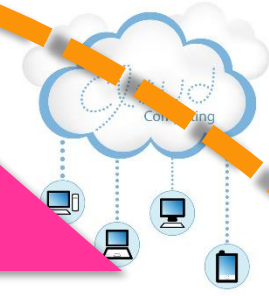


(tot. time = 5.8 s)

## Conclusions

- GPU-based SAR processing provides a significantly faster performance when compared to high-end alternate options such as dual-CPU computers or small, network-connected clusters:
  - Ideal match for embarrassingly parallel problems.
  - Zero-copy (local) VS network (distributed).
- GPGPU works at best when you can migrate/refactor the whole software processing pipeline to run on the GPU:
  - Frequent data ping-pong and code patchwork will kill performance.
- Sentinel-1 focusing will be available in SARscape 5.3.

# Thanks!



DISTRIBUTED  
COMPUTING

$$\begin{aligned}
 & \langle \bar{R}_N + \frac{Z_N}{2} n_t | \hat{\rho}(t) | \bar{R}_N - \frac{Z_N}{2} n'_t \rangle \\
 &= \sum_{n_0, n'_0} \int d\bar{R}_0 dq_0 dp_0 dq'_0 dp'_0 \\
 & \times \frac{1}{4} (q_{n_t} + ip_{n_t})(q_{n'_0} + ip_{n'_0}) \\
 & \times \int \prod_{k=1}^{N-1} d\bar{R}_k d\bar{q}_k d\bar{p}_k d\bar{q}'_k d\bar{p}'_k \\
 & \times \prod_{k=1}^{N-1} \dots
 \end{aligned}$$

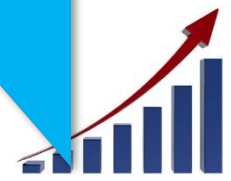
NEW ALGORITHMS AND  
MORE EFFICIENT CODE

next-gen  
SARscape

MULTICORE, GPGPU AND  
PARALLEL COMPUTING



OpenCL



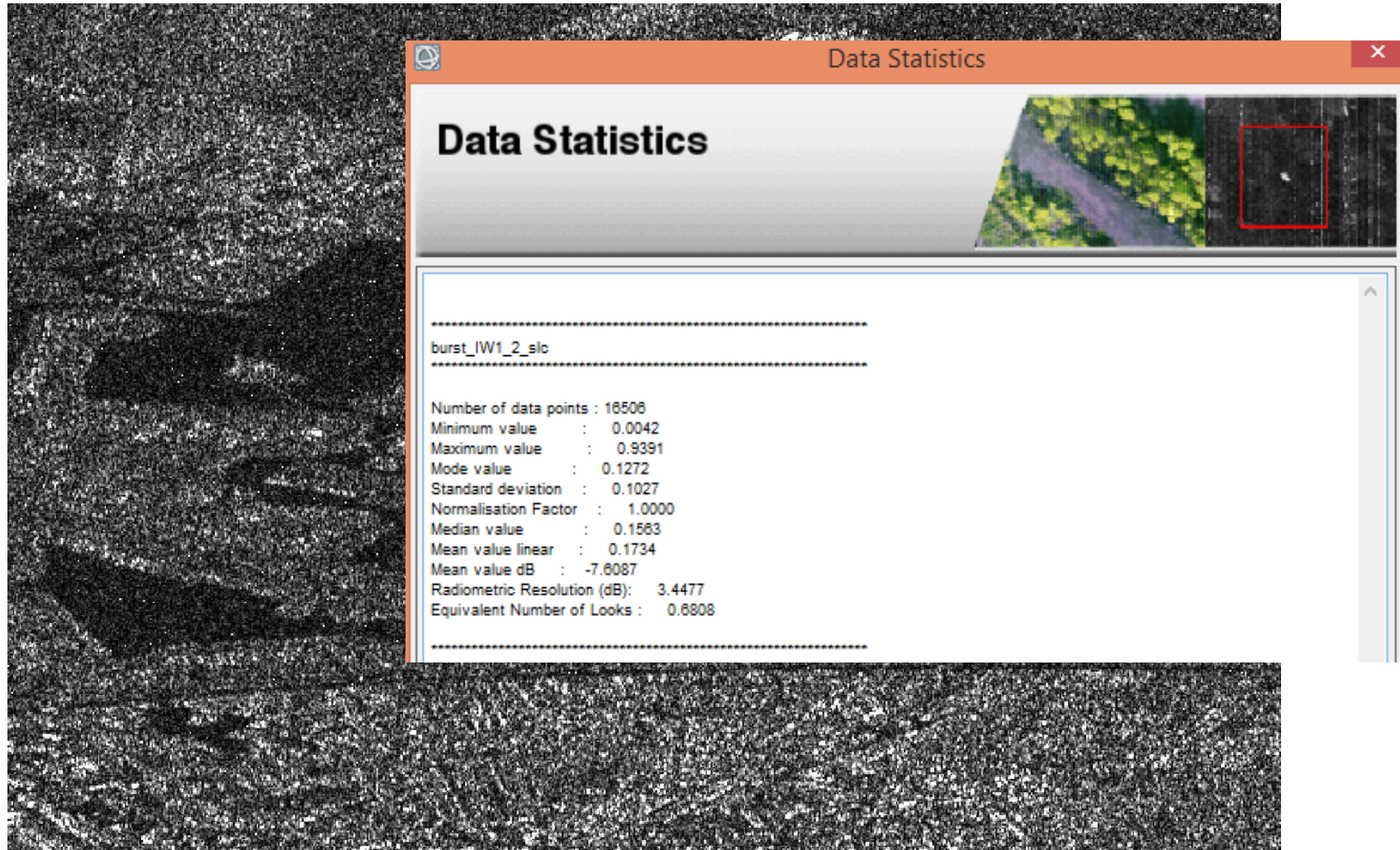
## Reminder: Time to beat (ENVISAT = 15 s)

- 1x ENVISAT frame focused in **about 8 s**
- On the demo machine, even in **less than 7 s**

*A. Peternier, M. Defilippi, A. Cantone, and P. Pasquali.  
Computationally efficient and operational solutions for the processing  
of very large SAR datasets based on modern multicore and GPGPU  
technologies. Presentation at the International Geoscience and  
Remote Sensing Symposium (IGARSS), Milan, Italy, 2015.*



# ESA Sentinel-1 SLC



# SARIPA Sentinel-1 SLC

