Big Data mais basse consommation
L’apport des processeurs manycore

Laurent Julliard - Kalray

Le potentiel et les défis du Big Data

Séminaire ASPROM – 2 et 3 Juillet 2013

AGILE
PERFORMANCE

Presentation Outline

- Kalray : the company, products and markets
- The rising of Manycore architectures
  - Feartures and benefits
- Using Manycore processors in Big Data
  - Opportunities and customer cases
- Q&A

Kalray at A Glance

- Founded in 2008 – located in Paris, Grenoble (France) & Tokyo (Japan)
- 55+ people
- Independent and unique technology : MPPA® MANYCORE processor (Multi-Purpose Processing Array) and software programming environment
- Targeting the industrial, embedded and computing intensive markets
- Large patent portfolio
- Several awards over the past years
  - Kalray ranked in "EETimes' silicon 60 : Hot start up to watch" in 2012
  - "Best technology award" from "Les Trophées de l'Embarqué" in 2012
  - "Startup of the year" by ElectroniqueS magazine in 2013

First MPPA®-256 Chips with CMOS 28nm TSMC

Released November 2012

- High processing performance
  700 GOPS – 230 GFLOPS
- Low power consumption - 5W
- High execution predictability
- Software programmable

KALRAY, a global solution

MPPA MANYCORE

Powerful, Low Power and Programmable Processors

MPPA ACCESSCORE

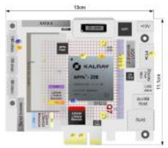
C/C++ based Software Development Kit (SDK) for massively parallel programing

MPPA DEVELOPER

Development platform
Reference Design Board

MPPA BOARDS

Reference Design board
Application specific boards
Multi-MPPA or Single-MPPA boards




Target Application Areas

INTENSIVE COMPUTING

- Finance
- Numerical Simulation
- Geophysics
- Life sciences

IMAGE & VIDEO

- Broadcast
- Medical Imaging
- Digital Cinema
- Augmented reality
- Vision

EMBEDDED SYSTEMS

- Signal Processing
- Aerospace/Defence
- Transport
- Industrial Automation
- Video Protection

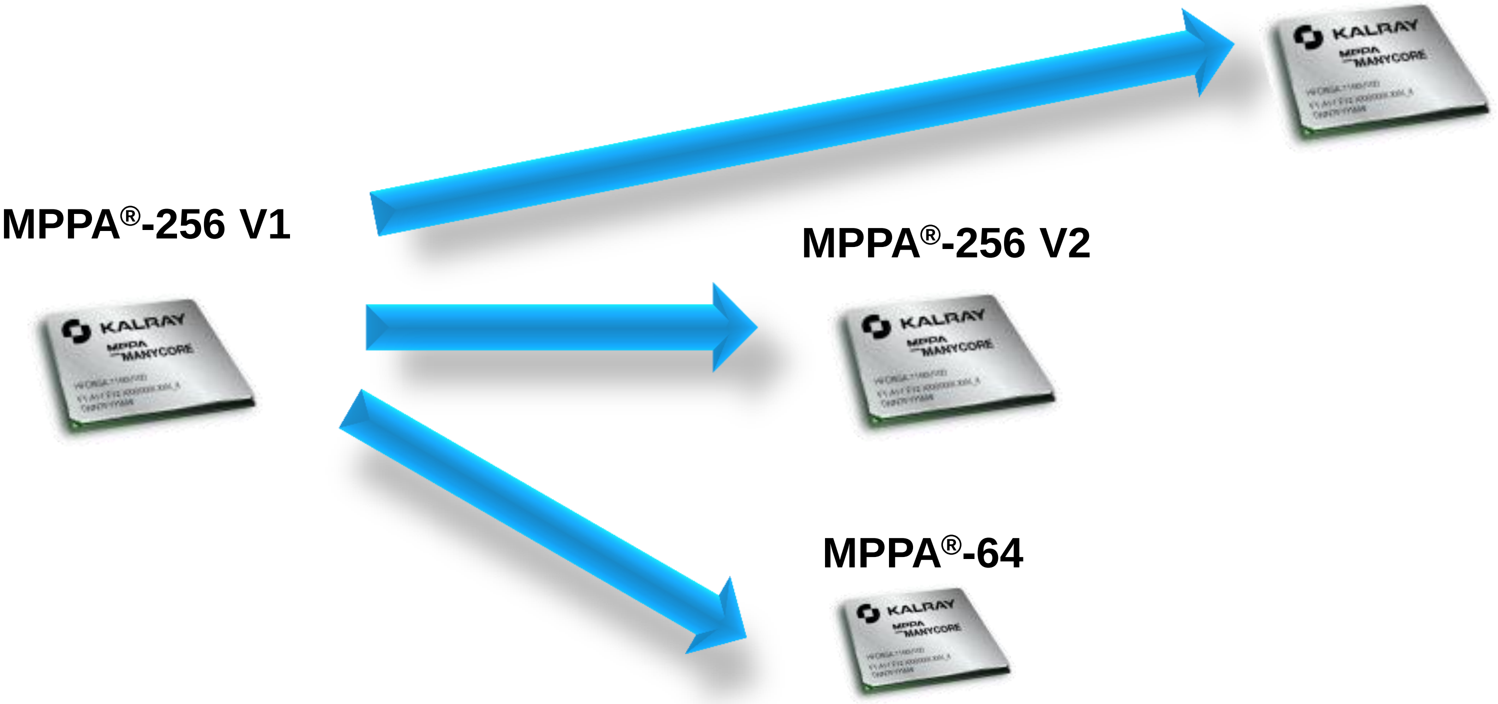
TELECOM / NETWORKING

- Packet Switching
- Network Optimisation
- Security Services
- Software Defined Radio
- Software Defined Network

MPPA MANYCORE Roadmap

Architecture scalability for high performances and low power

Q4 2012	Q2 2014	Q2 2015



MPPA®-1024 — Low Power 7 W

MPPA®-256 V1

MPPA®-256 V2 — Low Power 5 W

MPPA®-64 — Very Low Power 75 mW - 1,8 W

1st core generation
50 GFLOPS/W

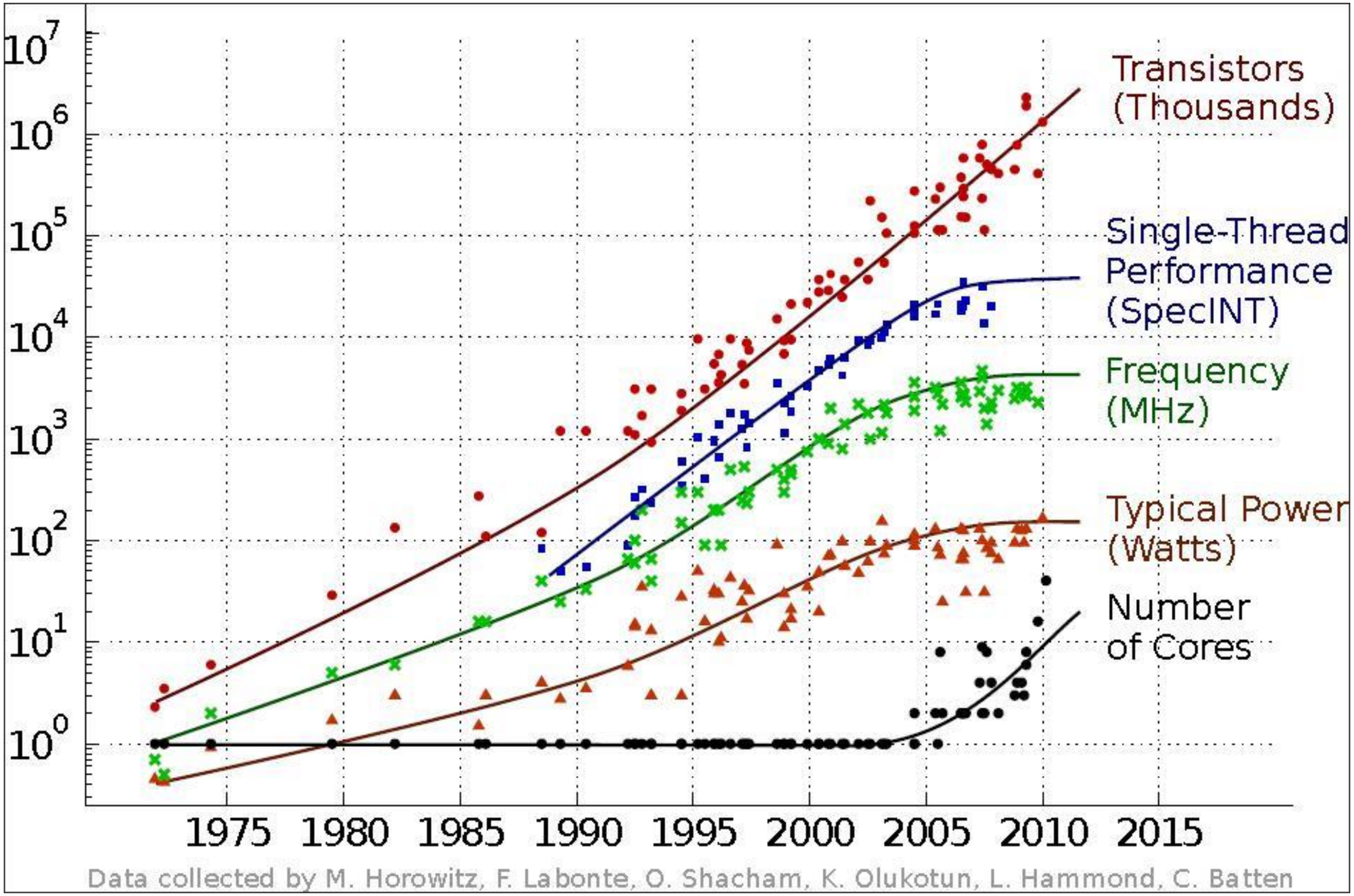
2nd core generation
80 GFLOPS/W

3rd core generation
100 GFLOPS/W

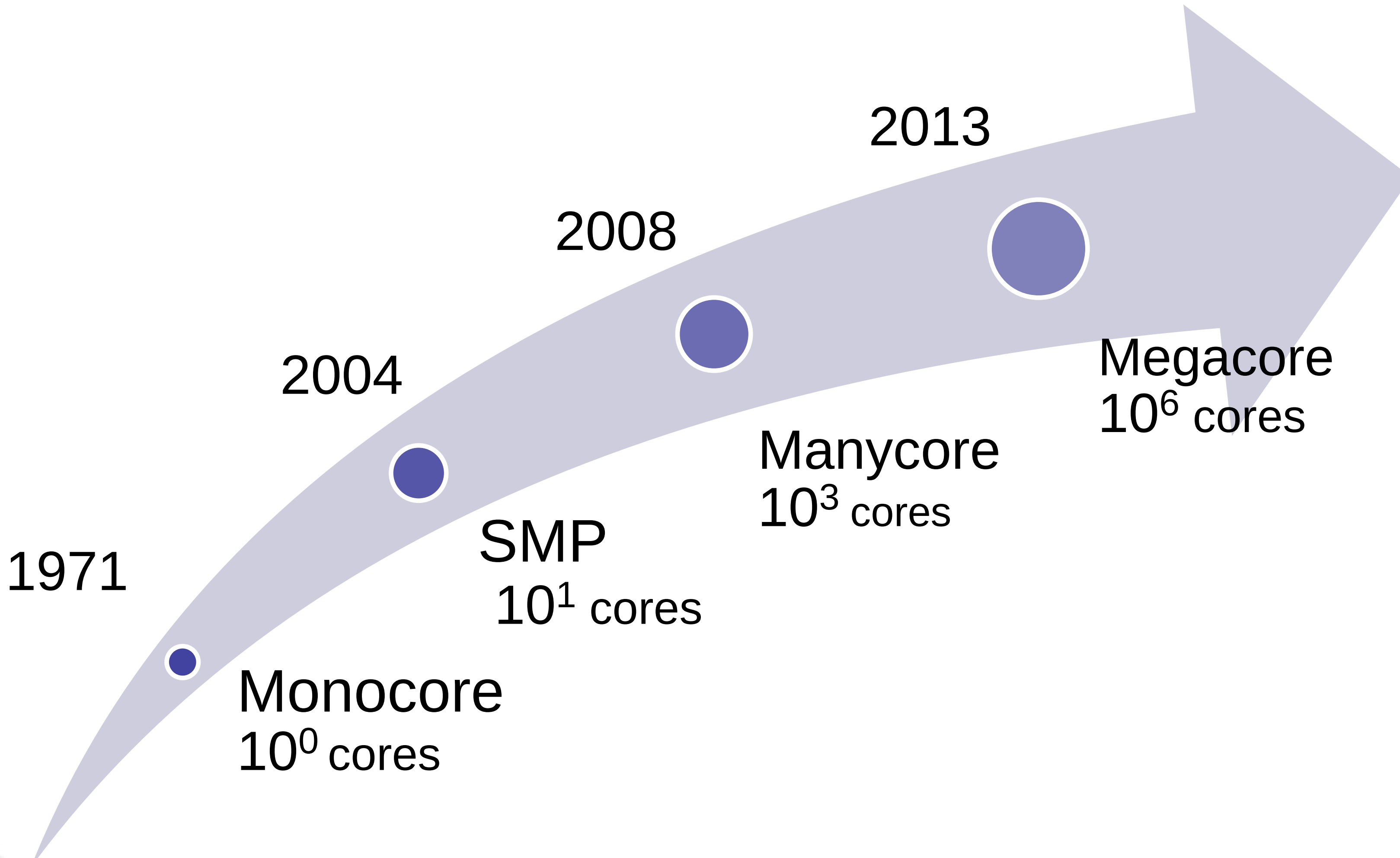
KALRAY

The rising of Manycore architectures

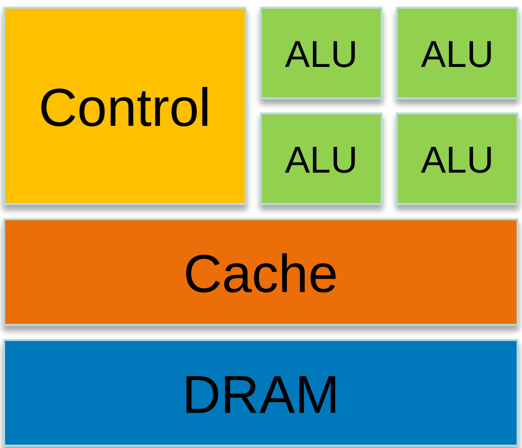
3 decades of microprocessor trend



Even more cores on a single chip…



Multicore CPUs vs GPUs vs Manycore

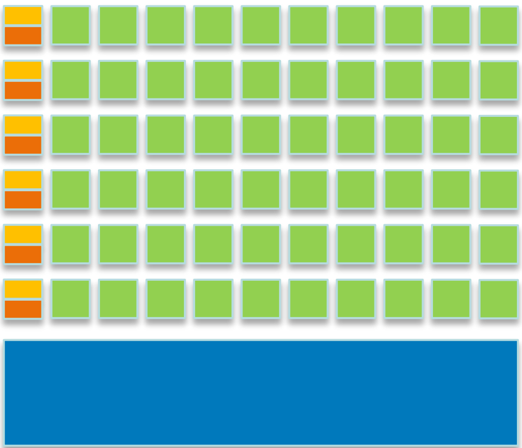


CPU

- CPU are optimized for sequential code performance
- Sophisiticated Control Logic to execute several instruction at the same time
- Very large cache to reduce access time to instruction and data of complex applications
- Clock frequency limits reached long ago
- 50 to 150 Watt

CONSTRAINED BY

- Power consumption , Complexity, Scalability

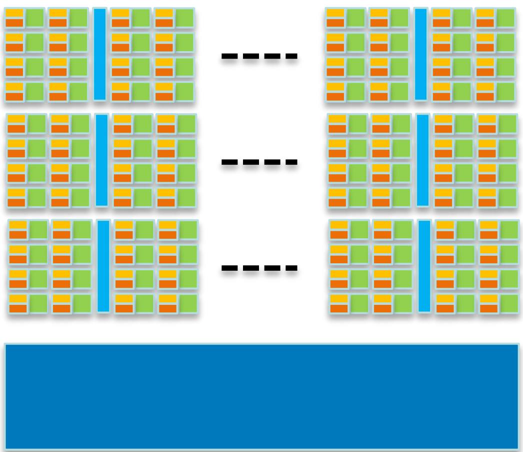


GPU

- Originates from the video game industry
- Very simple control units and huge number of floating point units working in parallel
- SIMD processing model
- Smaller cache than CPUs
- 20 to 300 Watt

CONSTRAINED BY

- Power consumption , Programming model, Communication overhead



Manycore

- Rich CPU-like Control units + FPU
- Cluster of processors share a local memory
- Clusters communicate through a high speed low latency network on chip
- MIMD paradigm (Multiple instructions Multiple Data) rather than MIMD
- 5 to 10 Watt

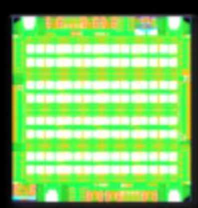
CONSTRAINED BY

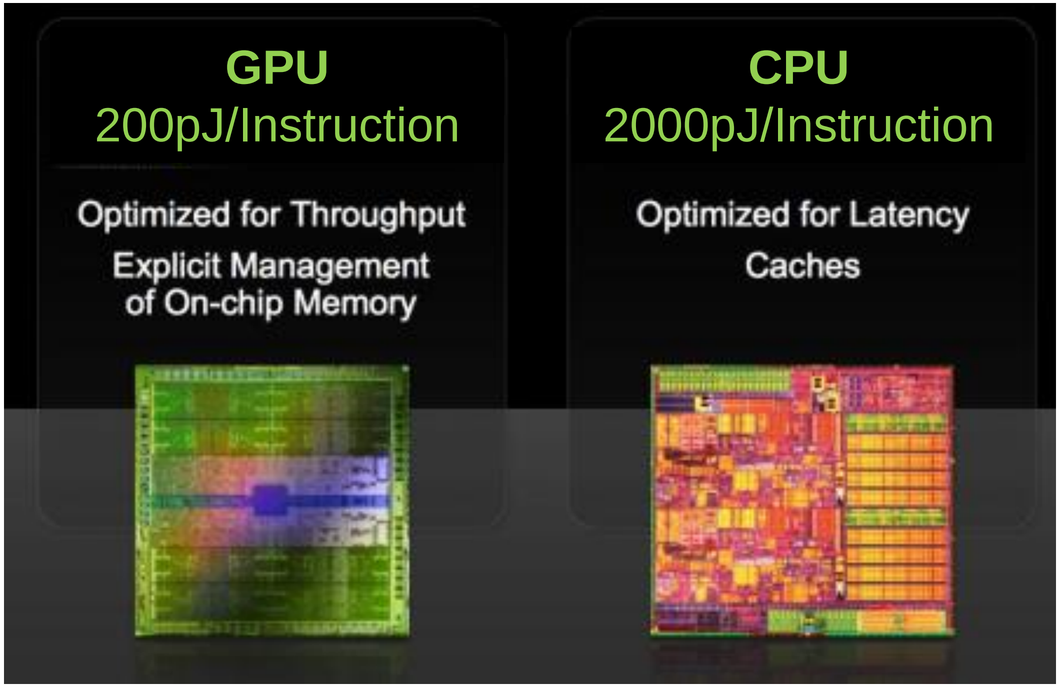
- Disruptive technology adoption curve

MPPA® Technology Compared to GPU & CPU

MPPA-256
20pJ/Instruction

Optimized for GFlops/Watt and $
Distributed Memory model
Low Latency





Source: Bill Dally, "To ExaScale and Beyond" - NVidia

MPPA-256	GPU	CPU
230 GFlops 5 W	3000 GFlops 300 W	700 GFlops 130 W
~50	~10	~5

KALRAY

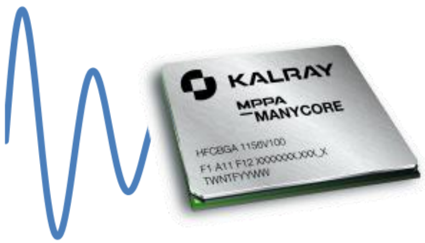
Manycore opportunities in Big Data

Be more efficient remotely

- Computing efficiency (GFlops/Watt)
  - Energy consumption becomes an absolute barrier whether in high-end embedded sytems or data-centers
- Hardware efficiency (GFlops/$)
  - Remove unnecessary hardware overhead
- Bandwidth (MB/s)
  - Bring data in and out fast and avoid bus bottleneck

Do more locally

- Process data at the source
  - Bring more intelligence right next to the data generator
    - Imperatives : more Gflop/W and Gflop/$
  - Use new « online » information processing models
    - Learn from data and predict continuously and in real time (e.g. HTM/CLA)



01110
10100
11011

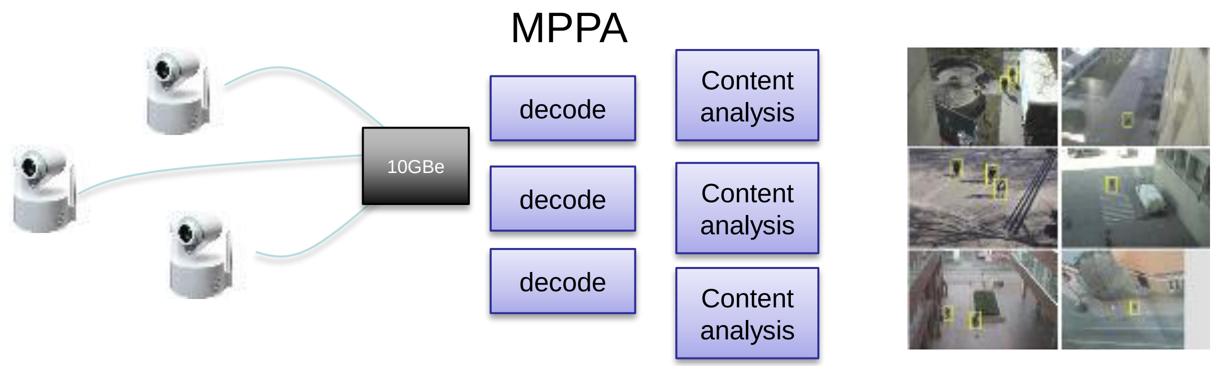
- Benefits
  - Save on the global system cost (e.g. lighter network infrastructure)
  - Save on energy : transmitting and processing date remotely takes 100 to 1000 times more energy than doing it locally
  - Adapt local acquisition and screening strategies

MPPA for video protection

- Improved Content analysis
  - High resolution camera / low false detection rate
  - Robust algorithms ➔ high performance computing of MPPA
  - Real Time detection
  - More simple infrastructure ➔ Compute power at the source

- System integration: Ethernet input / decode / content analysis / encode

- Multi camera system on single chip



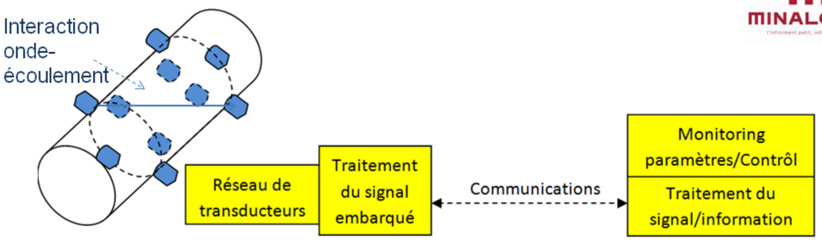
3. Projet en cours – interface avec MPPA

Projet « Smart Hydro Monitoring »

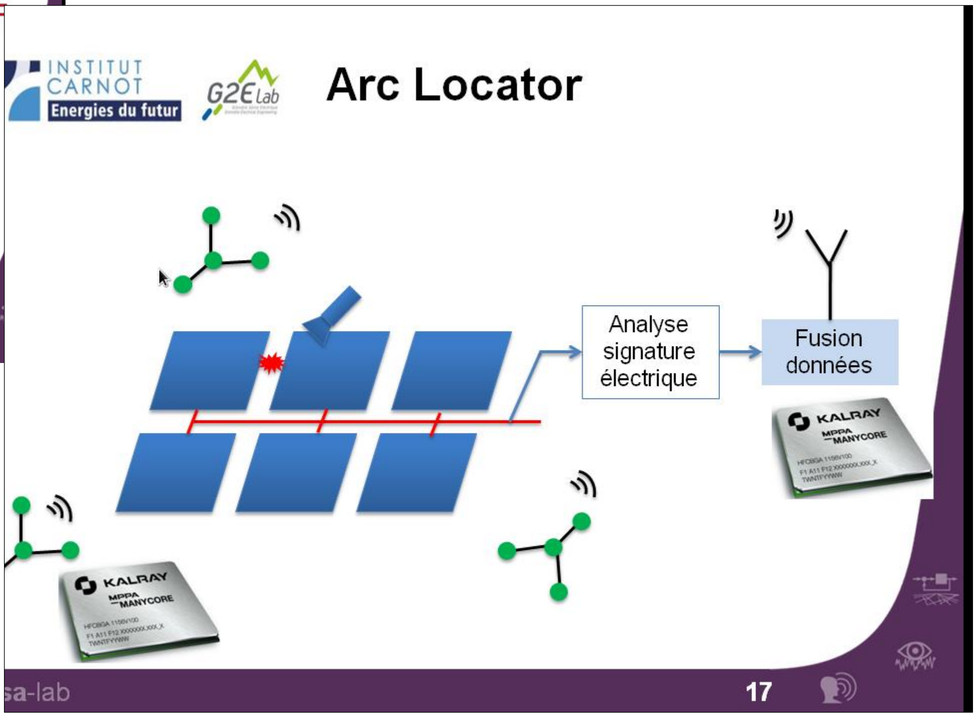
Monitoring en temps réel du profil d'écoulement

tenerrdis

MINALOGIC



gipsa-lab 16



Q&A

www.kalray.eu - info@kalray.eu