PARALLEL AND DISTRIBUTED COMPUTING ISSUES IN CYBER-PHYSICAL SYSTEMS AND THE FOURTH INDUSTRIAL REVOLUTION

Doctor Engineer Didier El Baz
LAAS-CNRS,
Toulouse France
&
International Lab on Security of
Cyber-physical Systems
ITMO University

LAAS CNRS | ITMO UNIVERSITY | NVIDIA.

Outline

- 1. Introduction
- 2. AI, parallel and distributed computing
- 3. Our contributions
- 4. Conclusions

1. Introduction

- The Fourth industrial revolution has started.
- Fusion of physical, digital world & the Internet.
- From cyber-physical systems to smart systems and smart world.
- Autonomy, collaboration, factory of the future.

Fig.1.1 Collaboration of man & robot AUDI, Ingolstadt, Germany

1.1 The fourth industrial revolution

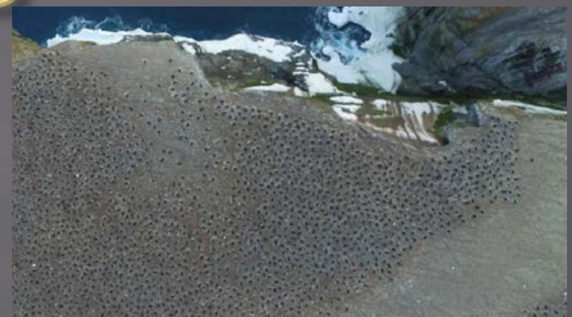
- AI comeback
- AI today → Autonomy.

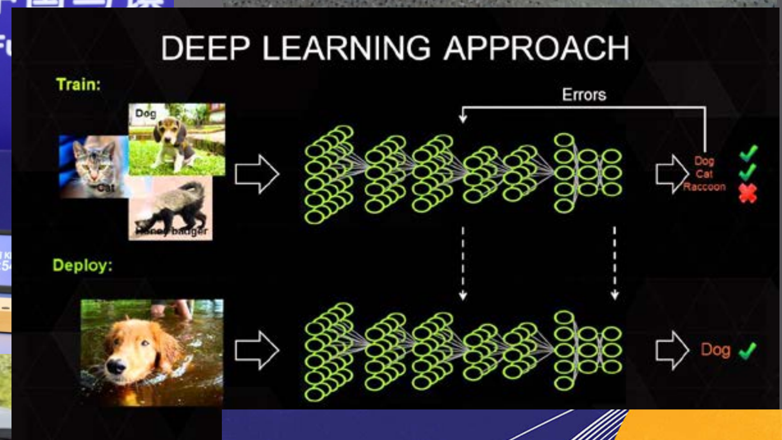
Fig.1.2 Google car, USA

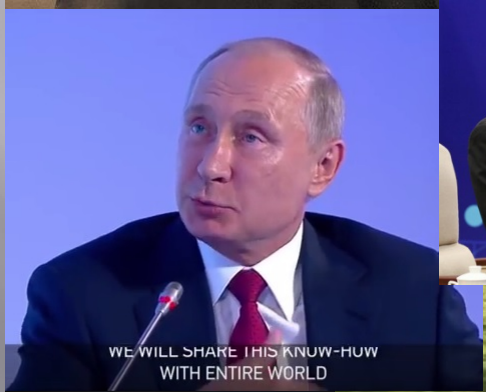
1.2. AI

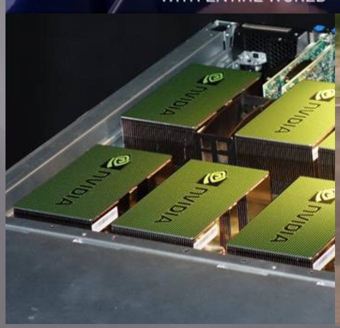
- AI: Tremendous opportunities in transport and manufacturing industry.
- Smart systems, smart cities, smart world
- Huge impact on society.
- Less urbanization?

2. AI, parallel and distributed computing









2.1 Amazing AI

Fig. 2.1 The next Rembrandt
TU Delft, Microsoft, ING, Mauritshuis
2016.
Cloud platform: Microsoft Azure VM
Parallelism (up to 1000 servers).

2.1 Amazing AI

Fig. 2.1 The next Rembrandt

2.1 Amazing AI

Fig. 2.1 The next Rembrandt

Fig. 2.2 The next GO game champions :
Alphago (Google) ;
earlier Deep Blue vs G. Kasparov (1997).

2.2 AI for sciences

- Counting Adélie penguin via AI (751,527 pairs)
- Deep Neural Network (DetectNet) to analyze photo collage & counting penguin nests.

Fig. 2.3 Penguin nests in Danger Island, Antartica February 2018

2.2 AI for sciences

- Looking for Planet 9 via Data Mining, AI, HPC.
- Passing all detections through a machine-learning system trained to catch and reject artifacts: satellite trails, hot pixels, cosmic rays.
- Cori supercomputer,
  Cray XC40 rank 8 of Top500
  14 Pflops, Xeon Phi.

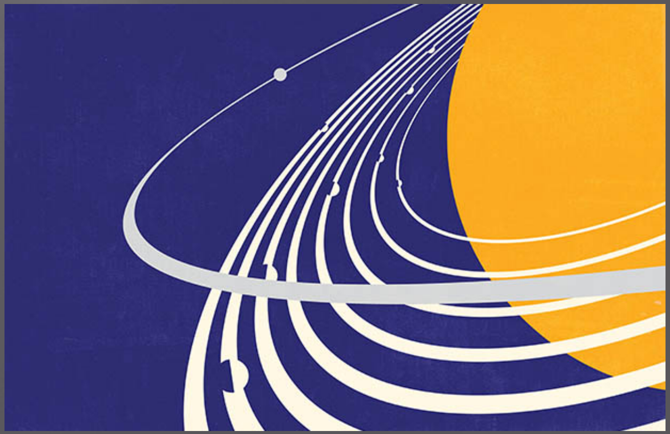
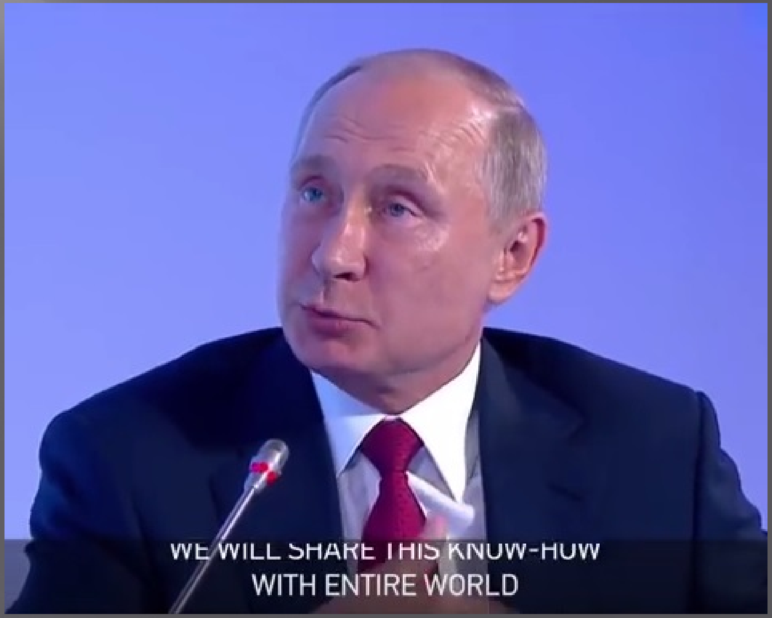


Fig. 2.4 Looking for Planet 9

2.3 Opinions

- Popular topic today
- « Artificial intelligence is the future … it comes with colossal opportunities but also threats that are difficult to predict.»

V. Poutine
Talk to the students
4 Septembre 2017



2.3 Opinions

- « Europe will need to invest a lot in supercomputer infrastructures, a domain where it is in late as compared with China and USA. Europe will need also to

  invest in the semi conductor

  industry. This will be a big deal. »

  C. Villani

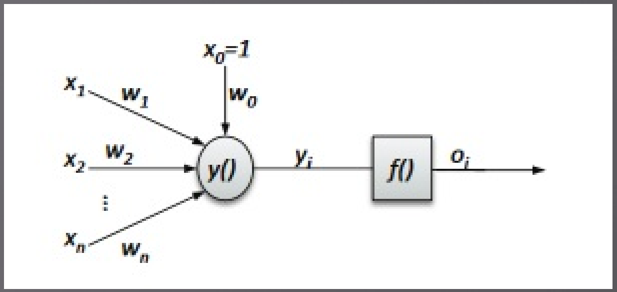
  Fields Medal 2010

  Deputy

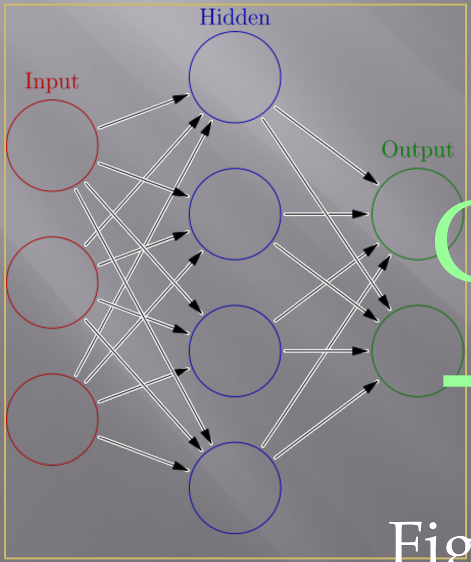
  L'Obs

  28 Février 2018

2.4. How?



- Mid 40s to 2018.
- From Computational models for neural networks to parallel deep Learning methods.
- Artificial Neural Networks (ANN).



GPUs
→

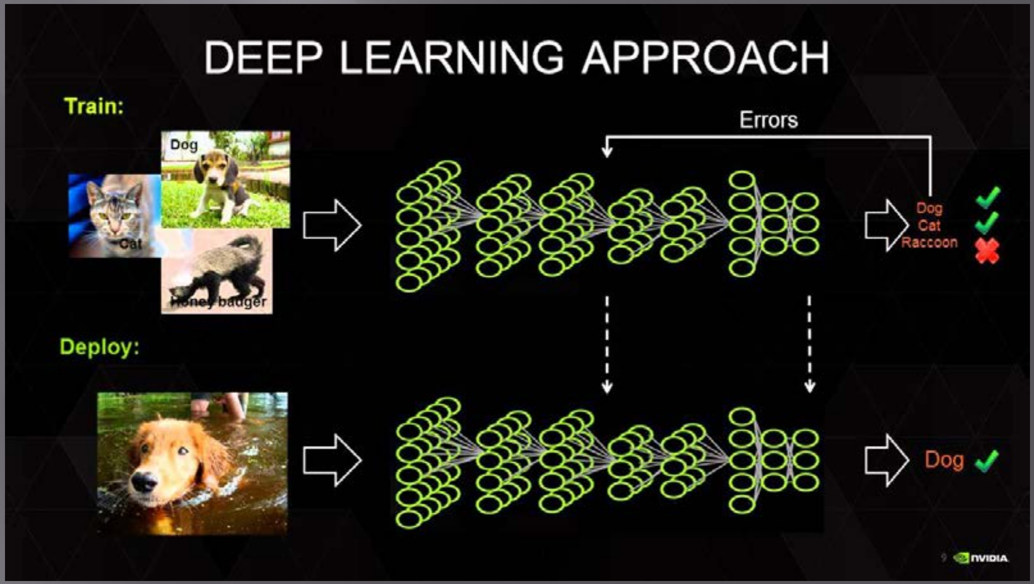


Fig 2.5 ANN

2.4.1 Data & Algorithms

- Importance of Data
- Algorithm that reproduces human / animal decision making.
- Heuristics or metaheuristics like
- genetic algorithm;

  genetic algorithms → distributed computing.
- ant colonies, swarms, flocks, fish school;
- neural networks, deep learning.
- Huge computations → parallelism; GPUs.

  Training ANN.

2.5. Where?

- Where is intelligence?
- Embedded intelligence (in the device);

  cost.
- Distributed intelligence (in the network),

  e.g., modular cyber-physical systems.

  resilience, volume, security issues.
- Hosted intelligence (not local: deported on a server).

  miniaturization, data mining, security issues.

  In a supercomputer.

  solving difficult combinatorial optimization

2.5.1 New computing platforms and AI

- Use massive parallelism of Graphics Processing Units (GPU)

e.g. NVIDIA Jetson TX2 (embedded intelligence)
256 CUDA cores ;

< 10 Watts ;

< $400 ;

> 1 Tera flop (simple précision).
Up to six cameras.

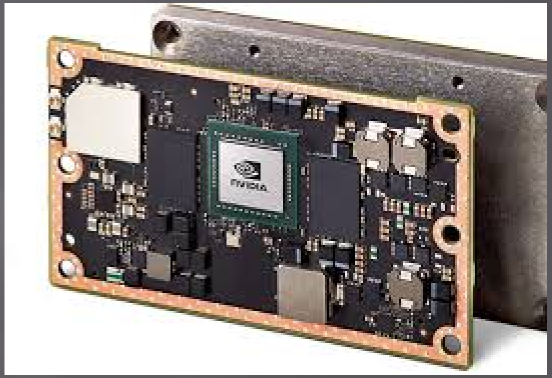


Fig. 2.6 NVIDIA Jetson TX2

2.5.1.1. GPU architecture

- **Three main kinds of cards:**
- Gaming: GeForce and Quadro products;
- HPC: Tesla products;
- Embedded system: Jetson TX2.

2.5.1.1. GPU architecture

- Streaming Multiprocessor (SM)-based GPU architecture

Figure 2.7 NVIDIA Kepler GK110 architecture

2.5.1.1. GPU architecture

- Streaming Multiprocessor architecture

SMX — Streaming Multiprocessor

Core — Single Precision cuda core

DP Unit — Douple preci-sion unit

LD/ST — Loading and Storing unit

SFU — Special function unit

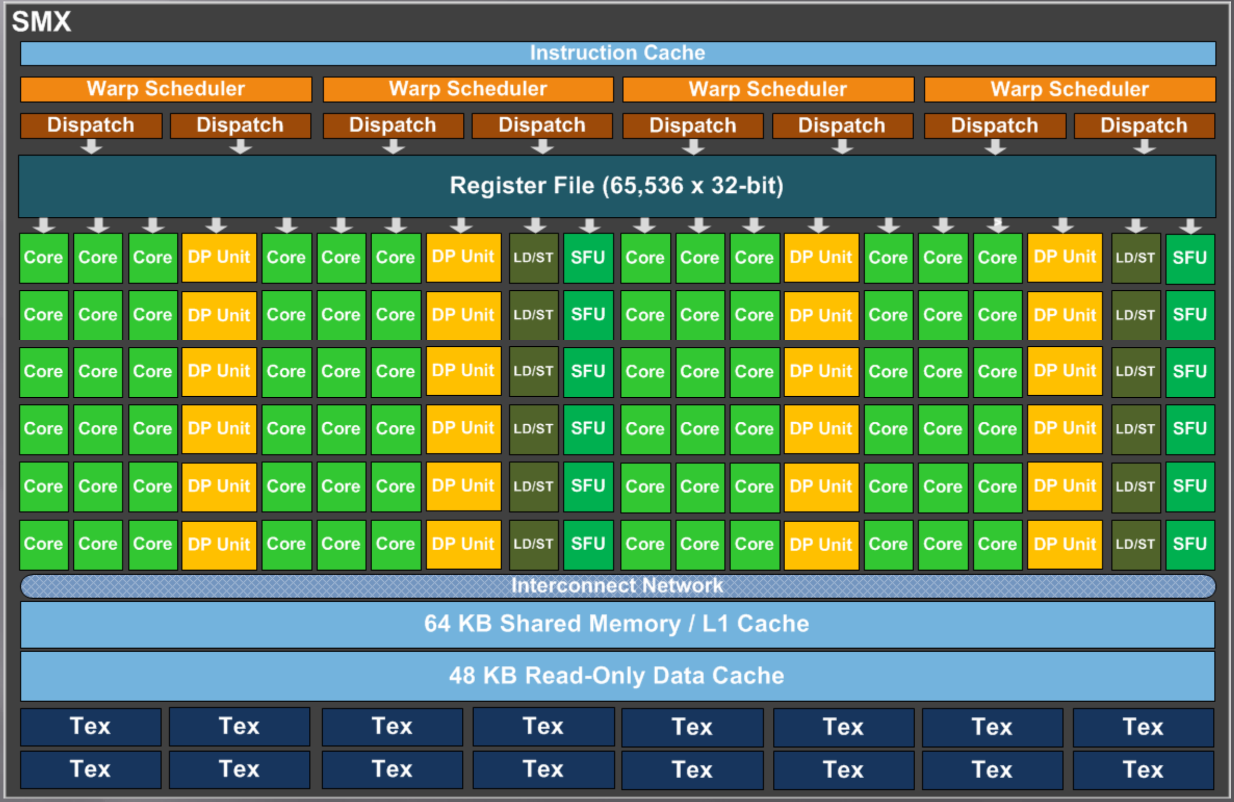


Figure 2.8. SM architecture (modified)

2.5.1.1. GPU architecture

- GPUs are powerful accelerators featuring thousands of computing cores;
- GPUs are widely available;
- GPUs are relatively cheap devices;
- GPUs are compact devices;
- GPUs accelerators require less energy than CPU.

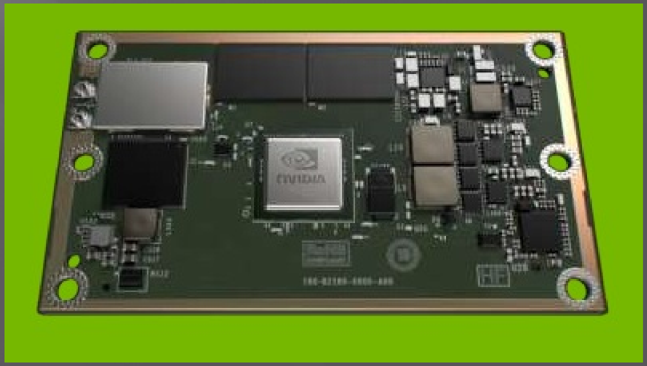


Figure 2.9 Jetson TX1

2.5.1.2. GPU synthesis

- GPUs are massively parallel computing accelerators.
  - Thousands of CUDA cores.

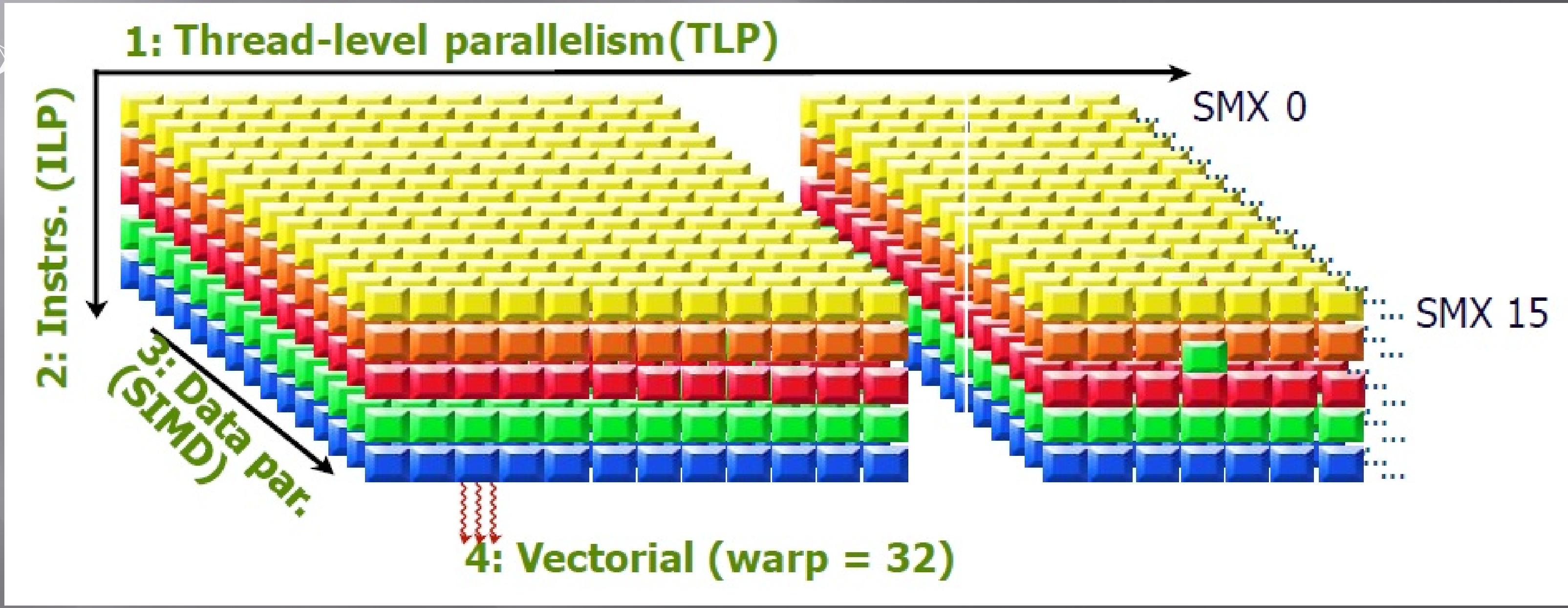


Figure 2.10 Several types of parallelism in Kepler GPUs.

2.5.1.2. GPU synthesis

- If the operations are not data dependent (no thread divergence) then the same instruction is executed inside the same warp of a given kernel.
- Different warps of a given kernel will execute different instructions.

2.5.1.3. Best Practice

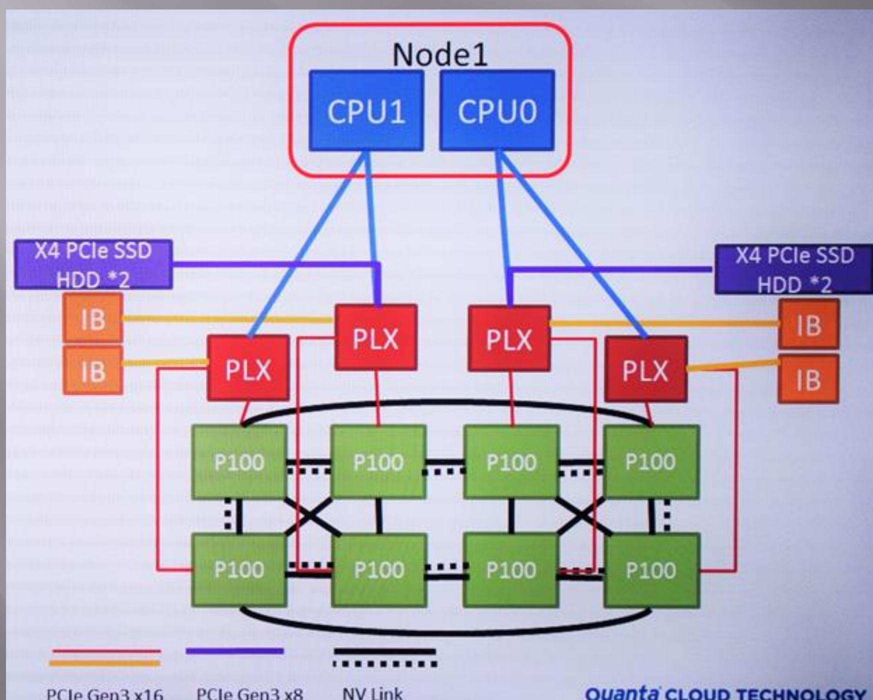
- Best practices:
  - Maximize thread occupancy (provide enough threads).
  - have non divergent threads in the same warp;
  - limit data transfers between CPU and GPU;

2.5.1.3. Best Practice

- Best practices:
  - Store intermediate results in registers instead of global memory.
  - Use shared memory for data frequently used within a thread block.
  - Optimize data locality on the GPU (in order to take benefit of high memory bandwidth.

2.5.2. AI supercomputers

- AI supercomputer: NVIDIA DGX-1
- Eight GPUs: P100 or V100.



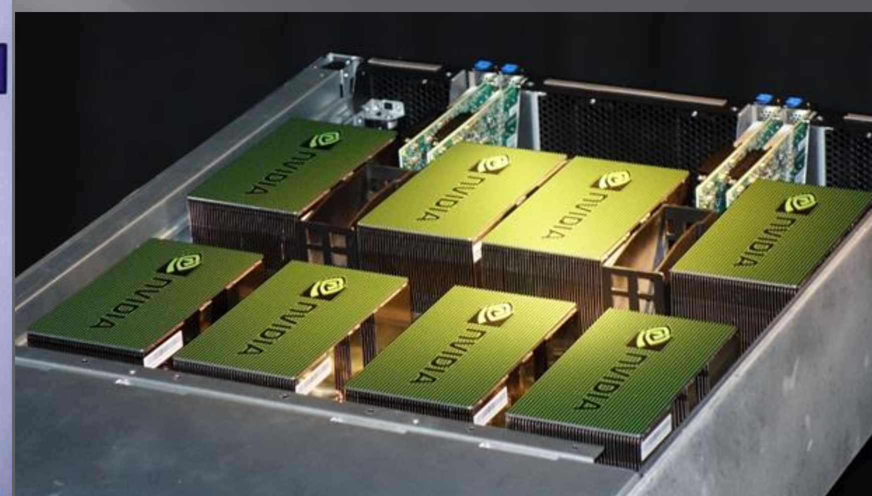


Fig. 2.10 NVIDIA DGX-1 supercomputer

2.5.2 AI supercomputers

- AI supercomputers

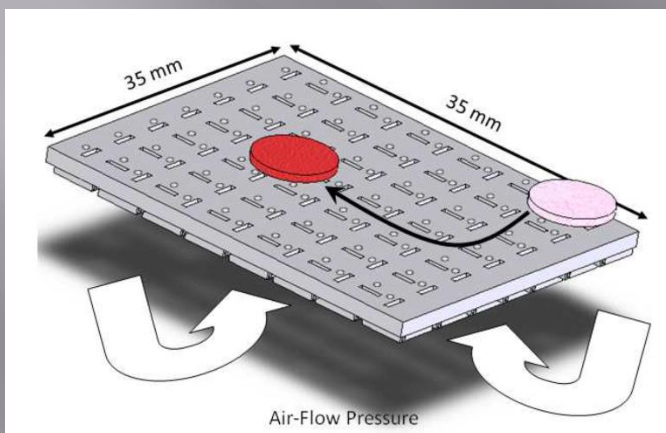
- From 170 TFLOPS up to 1 PFLOPS.
- From 28,672 up to 40,960 CUDA cores.
- Tensorflow.
- 96 x Faster Training than with Dual Xeon E5 -2699 (SIMT model).

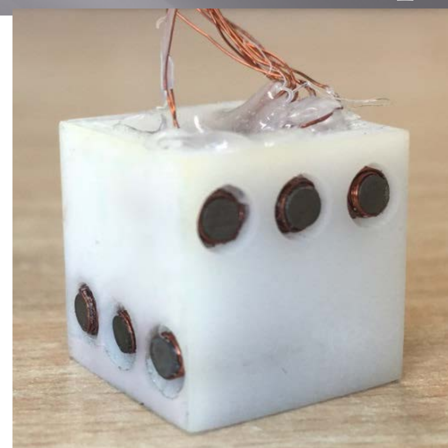
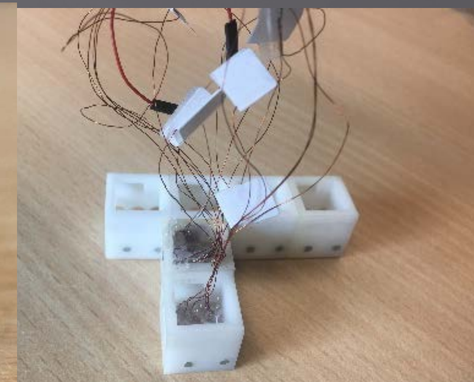
2.5.3 Other devices

- Intel Xeon Phi
  - Knights Corner, Knights Landing;
  - Parallel processor with vector processing units.
- FPGA
  - Reconfigurable logical network.

3. Our contributions

- Reconfigurable distributed smart conveyors.
- Parallel algorithms for planning
- Parallel or distributed metaheuristics for manufacturing
- Training many neural networks in parallel on GPUs.



3.1 Smart Surface

- The Smart Surface conveyor in manufacturing industry.
- ANR 06 ROBO 0009, 2007 – 2010.
- FEMTO-ST, LAAS, LIMMS.
- Distributed part differentiation.

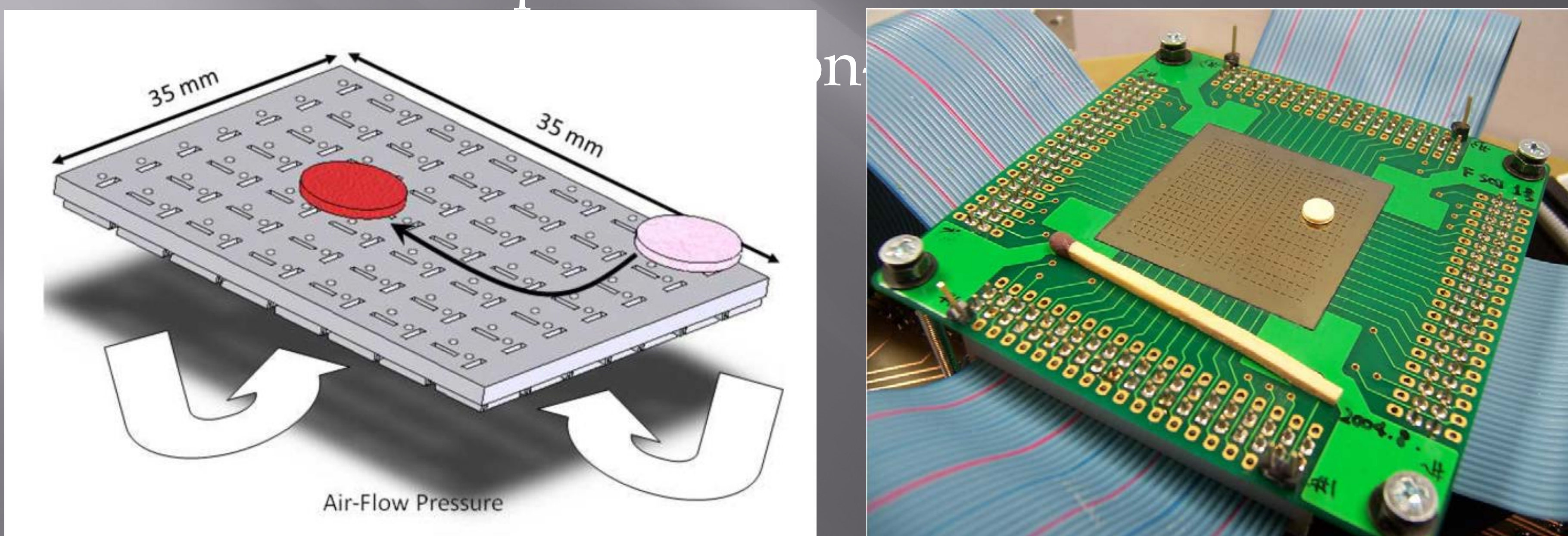
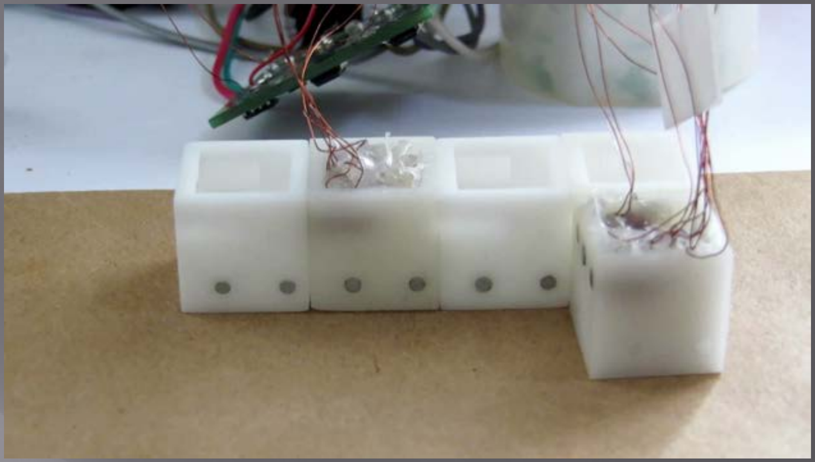


Fig. 3.1 Smart Surface

3.2 Smart Blocks

Distributed autonomous modular system;
Reconfigurable conveyor;
Cyber-physical systems;
ANR-2011-BS03-005, 2011 – 2015,
Huge cmputations.



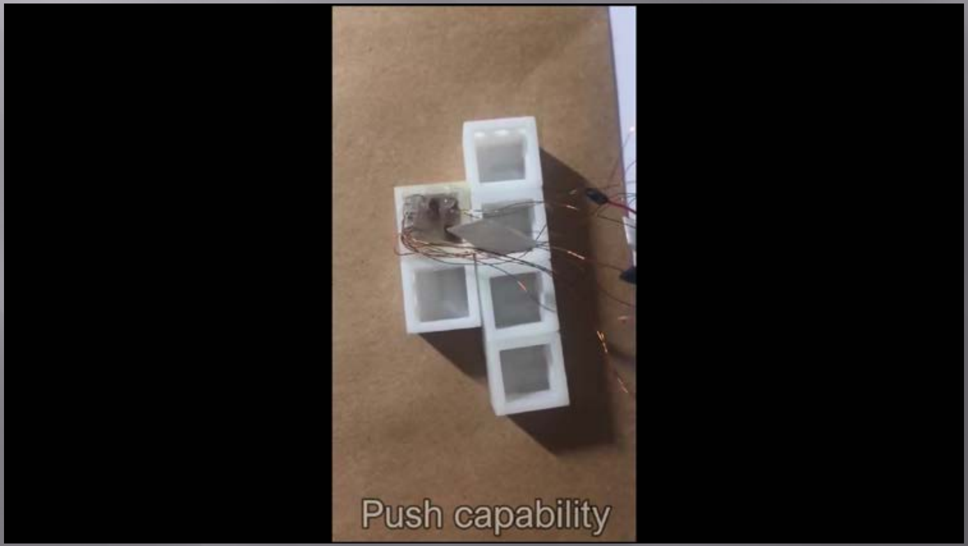


Fig. 3.2 distributed algorithm for reconfiguring distributed smart conveyors

3.3 Scheduling and Parallel Metaheuristics

- Scheduling problems (very difficult)
  - Energy Efficient Dynamic Flexible Flow Shop Scheduling.
  - Energy Efficient Dynamic Flexible Job Shop Scheduling.
  - GPU-based Parallel Genetic Algorithm; K40 GPU.
  - Find solutions previously unknown, improves solution quality and reduce computing time.

Fig. 3.3 Production line in car manufacturing

3.4 ANN Training and GPUs

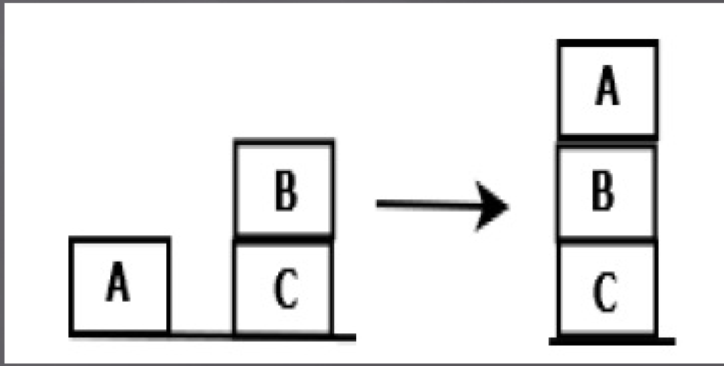
- Predict roduct demand, brewery company (real data).
- Training many ANN in parallel via back-propagation, K20 & K40 GPUs.

Fig. 3.4 Forecasting product demand in brewery

3.5 Parallel algorithms for planning

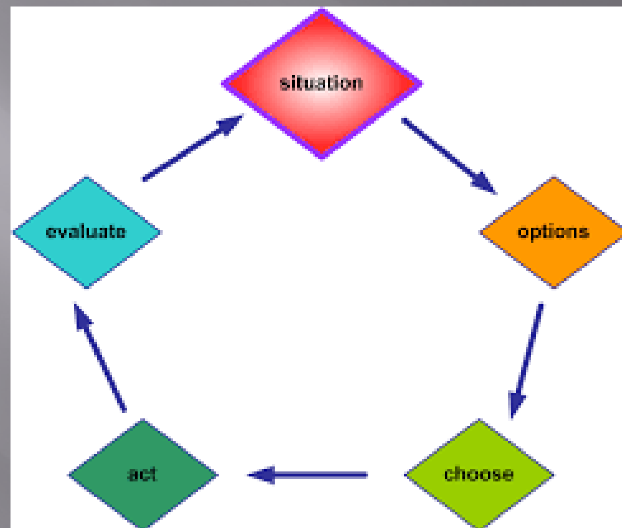
- Parallel best first search algorithms
  - Grid computing
  - GPU computing
- Application to airport tasks on planes, satellites task planning, oil industry,…

Fig. 3.5 Planning actions on blocks so as to obtain a given pattern and ordering Of blocks



3.6 Intelligent Flying Machines

Deep Learning

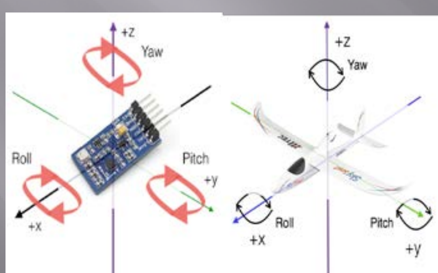


NonMonotonic Reasoning for Uncertain Situations

Decision Making

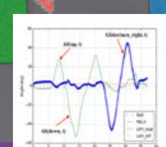
UAV testbed at Luminy, JL Vilchis Ph.D. student

AI is driving UAV Intelligence



Input

Pre-processing



Action chosen

Output

Set of facts [W]

Calculation of extensions

4. Conclusions

- 4th Industrial Revolution.
  - Convergence of many domains like Robotics, AI, Data Mining, Paralel or Distributed Computing /HPC.
  - In particular, design of parallel or distributed AI algorithms is a hot topic.

Publications

- D. El Baz et al. Distributed part differentiation in a smart surface, Mechatronics, Vol. 22, Issue 5, 2012, p. 522-530.
- J. Luo, S. Fujimura, D. El Baz, B. Plazolles, GPU based parallel genetic algorithm for solving an energy efficient dynamic flexible flow shop scheduling problem, Journal of Parallel and Distributed Computing, Journal of Parallel and Distributed Computing, 2018.
- Adel Dabah, Ahcène Bendjoudi, Abdelhakim AitZai, Didier El Baz, Nadia Nouali Taboudjemat, Hybrid Multi-core CPU and GPU-based B&B Approaches for the Blocking Job Shop Scheduling Problem, Journal of Parallel and Distributed Computing, 2018, 117, 73-86.
- D. El Baz, Cyber-physical systems and various computer science issues in smart distributed autonomous robots, IM & CTCPA 2017, Saint Petersburg Russia, 19 Décember 2017
- D. El Baz, Smart Systems, the Fourth Industrial Revolution and New Challenges in Distributed Computing, International Conference Parallel Computing, ParCo2017, Bologne Italy, 12-15 September 2017.

Publications

- D. El Baz, Challenges in Computing Accelerators and Heterogeneous Computing, 25th International Conference on Parallel Distributed and networked based Processing (PDP 2017) Saint Petersburg Russie, 6 au 8 Mars 2017.
- D. El Baz, IoT and the Need for High Performance Computing, in Proceedings of the International Conference on Identification, Information and Knowledge in The Internet of Things (IIKI2014), 17-18 Octobre 2014, Pékin Chine, p. 1-6, IEEE CPS.
- J. Cruz-Lopez, V. Boyer, D. El Baz, Training Many Neural Networks in Parallel via Back-Propagation in Proceedings of the 27th IEEE Symposium IPDPSW 2017 / PDCO'17, Orlando USA, 29 May 2 June 2017.
- Didier El Baz, Mhand Hifi, Lei Wu, Xiaochuan Shi, A Parallel Ant Colony Optimization for the Maximum-Weight Clique Problem in Proceedings of the 30th IEEE Symposium IPDPSW 2016 / PCO 2016, Chicago 2016, 23-27 May 2016, p. 796 - 800.

Publications

- J.L. Velchis, P. Siegel, A. Doncescu, “ Autonomous Aerial Vehicle Based on Non-Monotonic Logic”, VEHITS 2017, Porto, Portugal, 22-24 Avril, 2017.