

July, 10th
2014

From exotics to vanillas with GPU
© Murex 2014



Selected Industry Recognition and Rankings 2013-2014



Structured products
 TECHNOLOGY **RANKINGS 2014**
 TRADING SYSTEMS
 CROSS-ASSET
1 TRADING SYSTEMS
 COMMODITIES
 RISK MANAGEMENT
 COLLATERAL MANAGEMENT

OVERALL #1
TOP
TECHNOLOGY
VENDOR

Risk
 TECHNOLOGY RANKINGS
2013

- #1 Trading Systems
- #1 Pricing & Risk Analytics
- #1 Enterprise-wide Credit Risk Management
- #1 Limit Checking
- #1 System Implementation Efficiency

AsiaRisk
 TECHNOLOGY
 RANKINGS 2013

- #1 Trading Systems
- #1 Pricing & Risk Analytics
- #1 Enterprise-wide Credit & Liquidity Risk Management
- #1 Limit Checking
- #1 System Implementation Efficiency
- #1 After Sales Service



READERS' CHOICE AWARDS

BEST RISK MANAGEMENT
 PRODUCT OR SERVICE

Murex

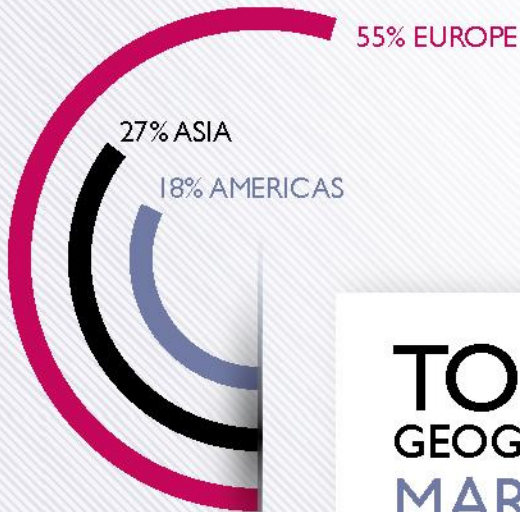
RiskTech
100

2014 Category Winner
 Capital Markets – Sell-Side

#1
 Best Selling Wholesale
 Banking System
 Murex
 MX.3

IBS journal
 Sales
 League
 Table
 2014

A worldwide presence



**TOP 10
GEOGRAPHIC
MARKETS**

- UK
- SOUTH EAST ASIA
- USA
- GERMANY
- AMERICA Excl. USA
- BENELUX
- ITALY
- SPAIN
- FAR EAST ASIA
- Other European Countries

GPU or equivalent – Xeon Phi – are here to stay

2008
Experimental

2009
OpenCL

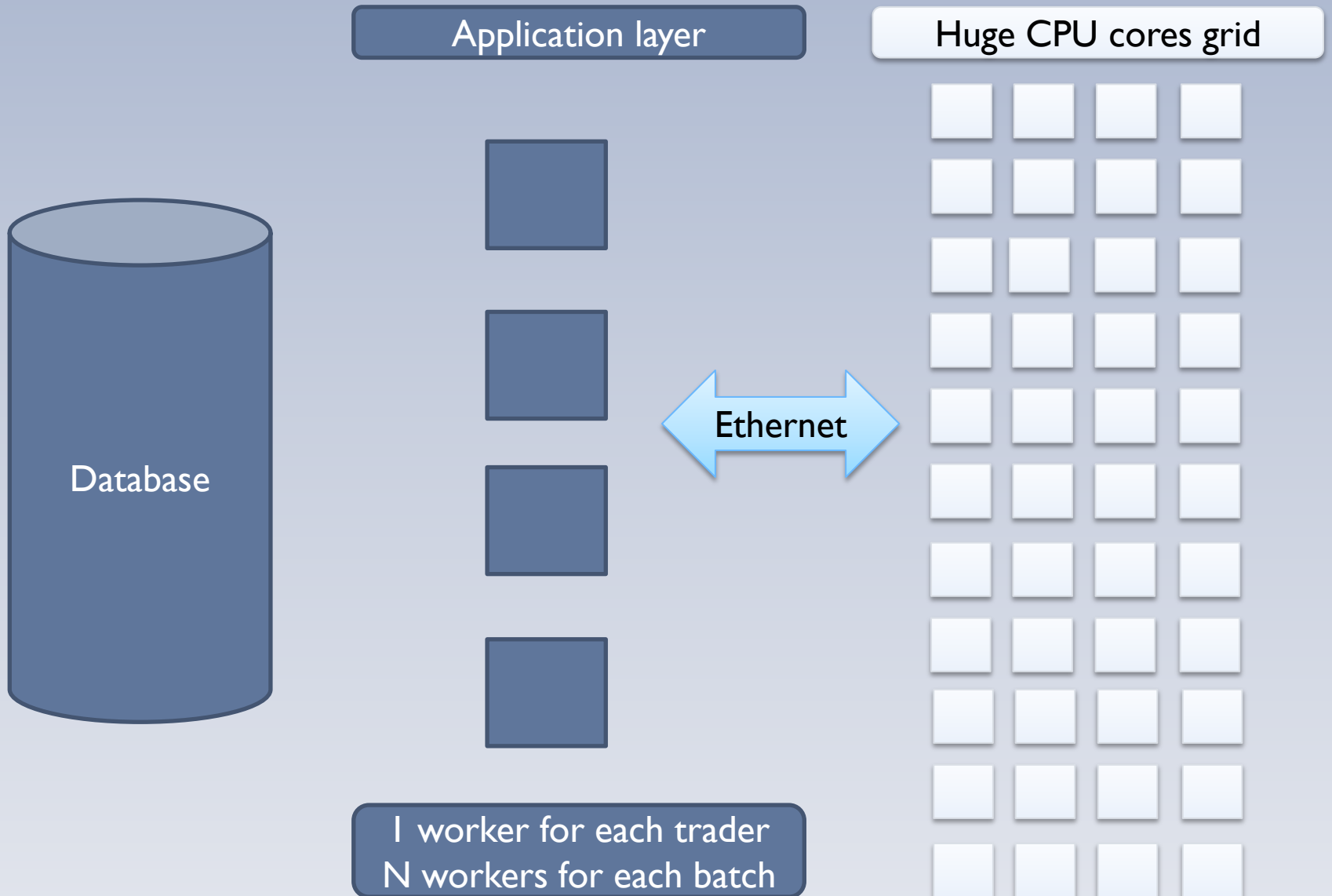
2011-2012
Murex in
production for
exotic options
live risk
management
with GPU

2014 NVIDIA
GPU powered
grid is the
default
configuration
for exotics
evaluation &
risk
management
with Murex

Goal

Show how new market needs and GPU experience have changed the financial software architectural design and are raising the need of other high performance computing technics

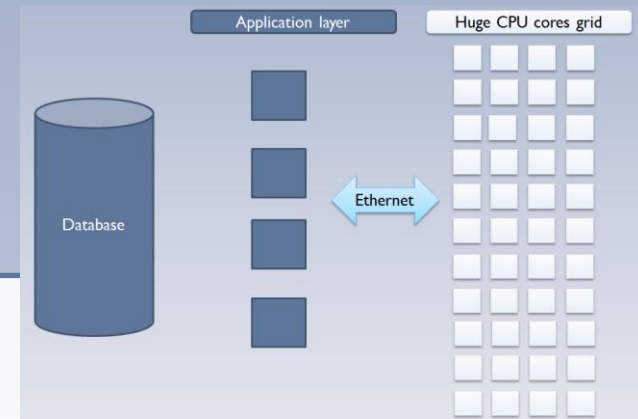
The good old financial trading software design



The good old financial trading software

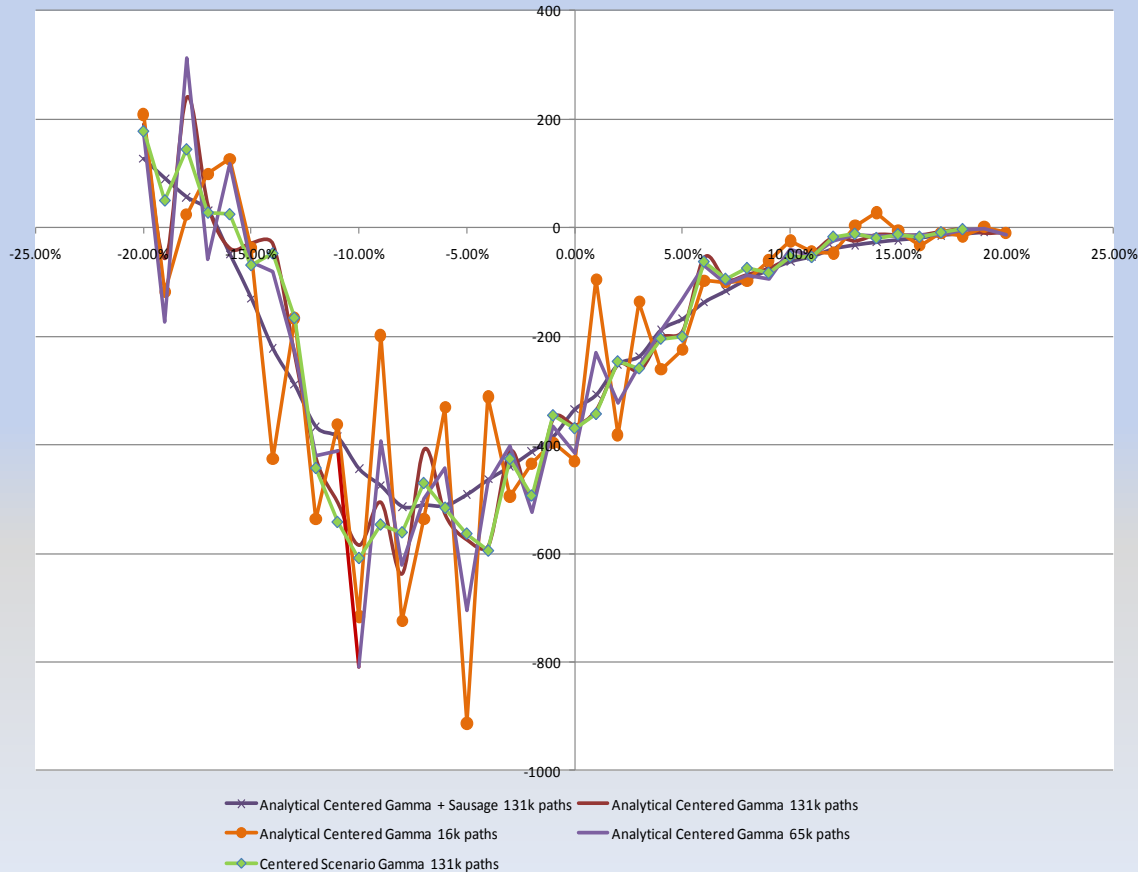
Build having in mind

- Simplicity
- Flexibility
 - Time to market for new products
 - Quants should only implement the PV algorithm and the system should do the rest
 - Hardware agnostic
- Not much hardware cost constraints
- Analytics for vanillas are extremely fast and can run in a single x86 process
- Analytics for exotics are extremely slow and need a grid
- Aggregation is not a problem – so few data -



Exotics first GPU benefit

11HK Spot scenario impact on 11HK Gamma



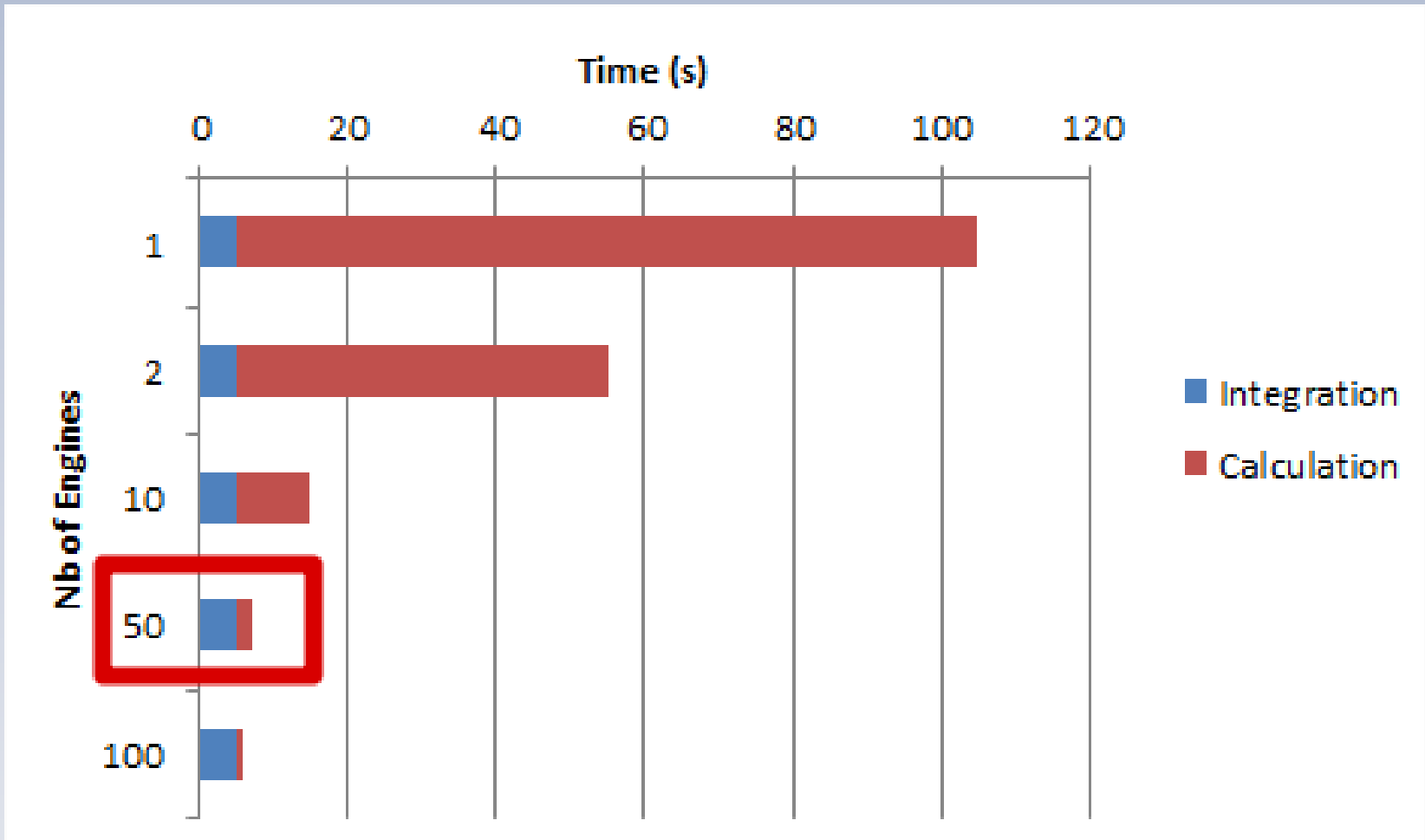
Settings tested

- 8k MC paths
- 16k paths
- 32k paths
- 65k paths
- 131k paths
- 131k paths with Sausage

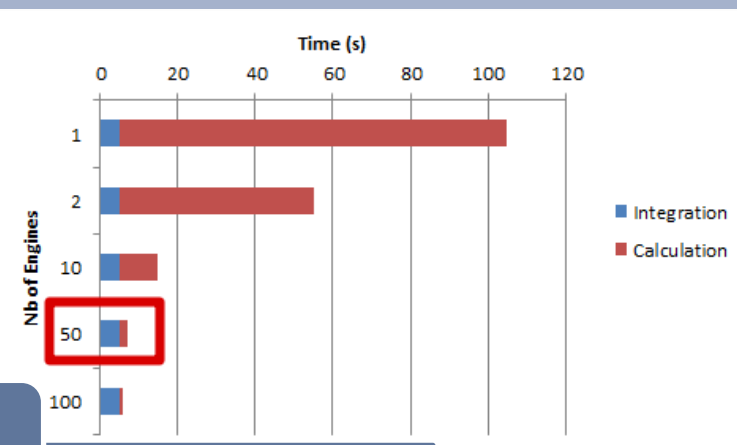
With GPU our customer do not have tradeoff between accuracy and speed and often gain in both

Exotics first GPU & the Amdhal low

Based on a production customer case



Exotics first GPU & the Amdhal low



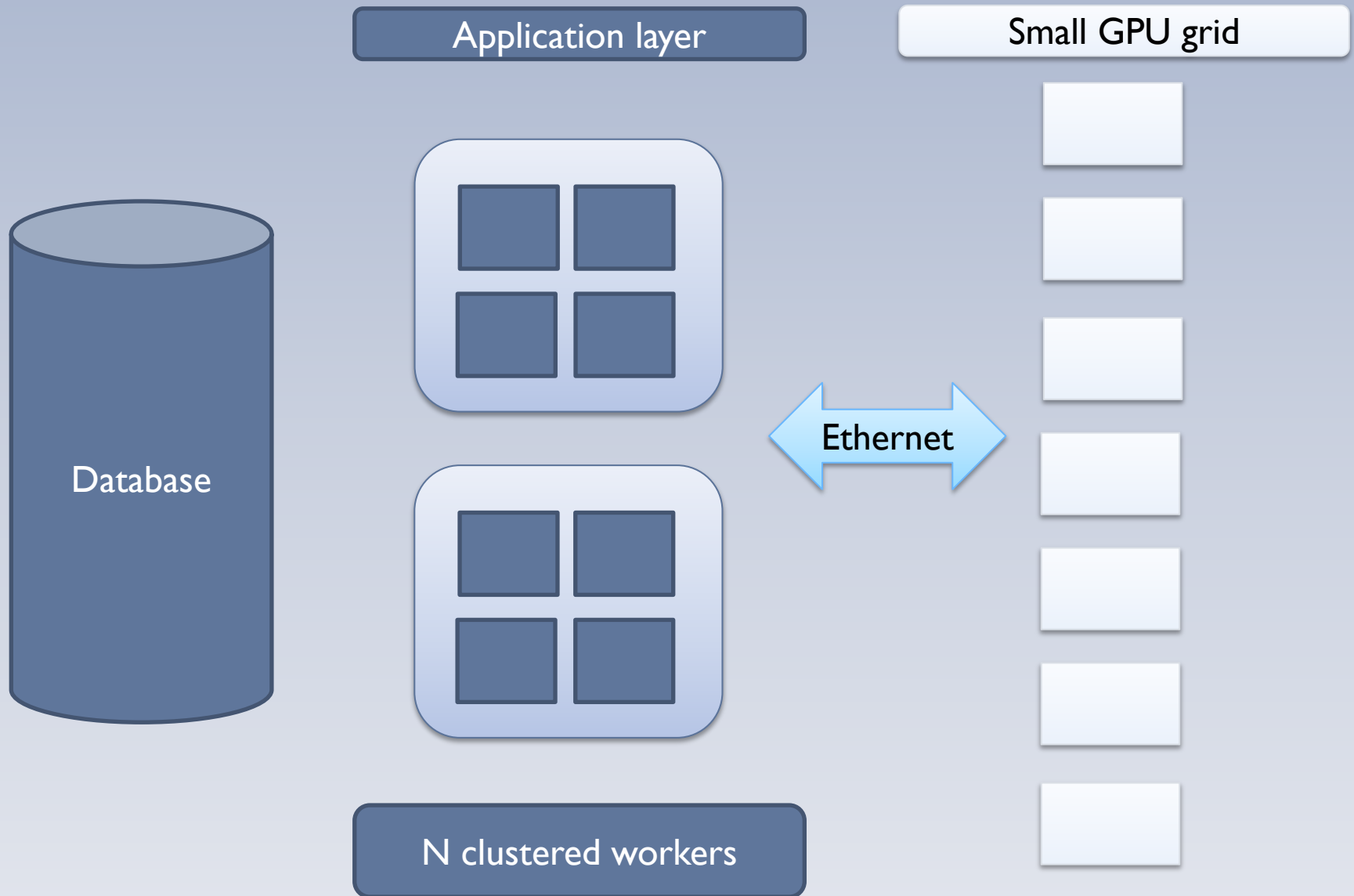
Cannot feed the grid fast enough when using GPU

- Exotic products evaluation stay slow with GPU but by a far smaller order of magnitude
- Sequentially preparing data, serializing, sending messages and aggregating results becomes the bottleneck when a customer has a reasonable grid

Solution

- Reduce sequential part & messaging by computing sensitivities & price on the same grid engine
- Even group trades together when they are too fast
- Parallelize again the residual part

Mandatory parallelization of the application layer to feed the GPU grid



Vanillas

A disruptive problem

- When pricing a trade I would like to choose if I go through a CCP or through a bilateral agreement taking into account my incremental CVA

Meaning

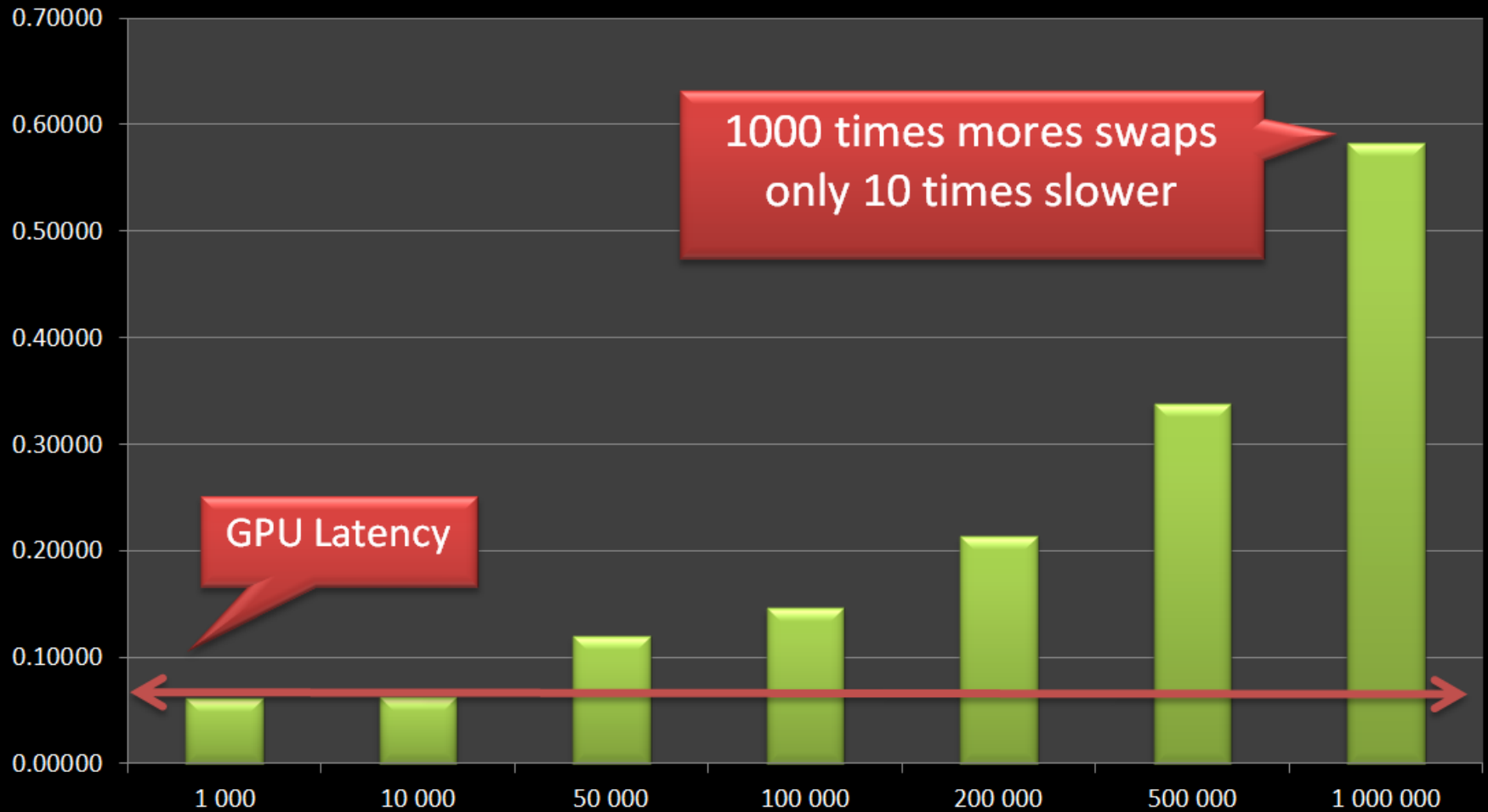
- Evaluate something like 100 000 transactions simply for the pricing of a basic trade
- Far more when we want to see live the market risk profile of our CVA desk

A change of magnitude

- So big that there is no hope that basic horizontal scalability will solve the problem

Vanillas GPU evaluation time

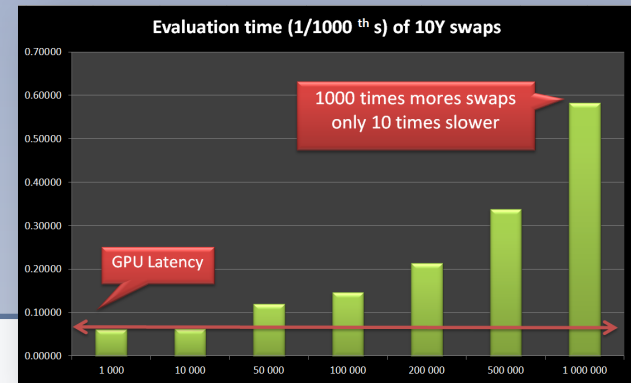
Evaluation time ($1/1000^{\text{th}}$ s) of 10Y swaps



Vanillas GPU evaluation time

In theory

- We could evaluate the CVA of a portfolio 150 000 swaps, 8000 paths with 120 horizon dates and 100 sensitivities in less than one hour using only two K40 GPUS



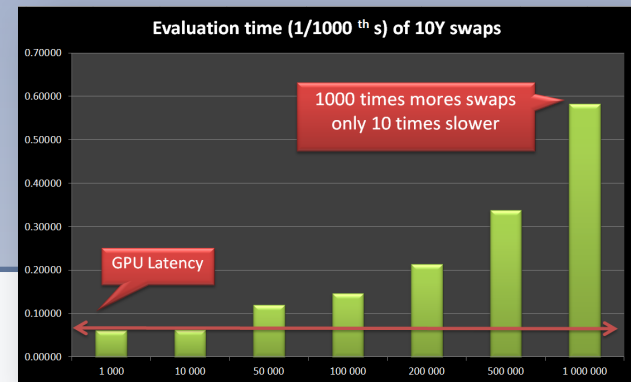
But in practice

- We are not efficient when evaluating only 8000 times the same swap at the same date

Look for GPU efficiency

Solution 1

- Evaluate all the dates at the same time
 - Payoff are not equivalent @ each evaluation date meaning that the code will show some slowness linked to branches in a way or another
 - Need to be able to keep in GPU memory a lot of paths to be efficient ... and it is not possible when having many risk factors or a complex model like BGM



Solution 2 : preferred incremental XVA

- Evaluate PV & sensitivities/scenarios at the same time as for exotics to be on the right part of the graph
 - Works well but do not solve the case where we only need the CVA or only a few sensitivities

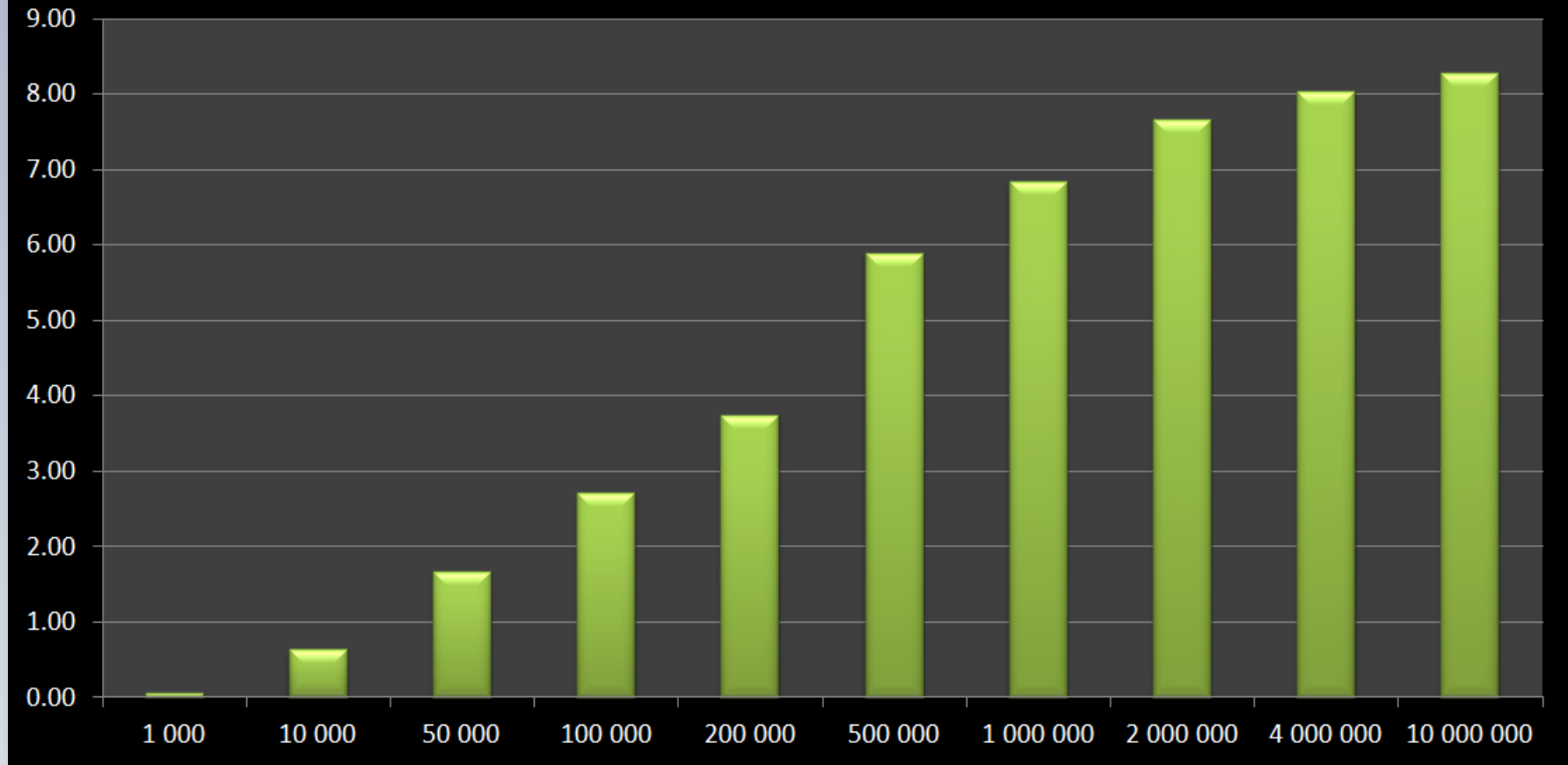
Preferred solution 3 : preferred for batch XVA evaluation

- Treat several swaps at the same time for each horizon date to feed the GPU

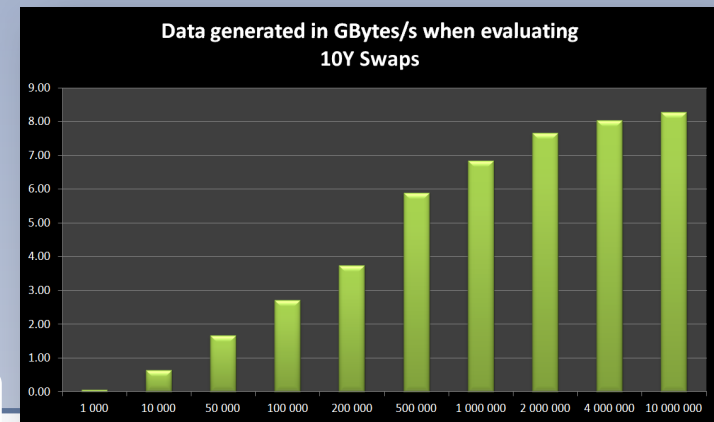
Mix both solutions to be ready for next generations of GPU

A new type of problem

Data generated in GBytes/s when evaluating 10Y Swaps



A new type of problem



Data generation speed

- Less than GPU or CPU memory speed
- Comparable to PCI express speed.
- Faster than Ethernet speed
- Comparable to IB speed
- Faster than disk storage speed

Use disk as a temporary buffer before flexible aggregation

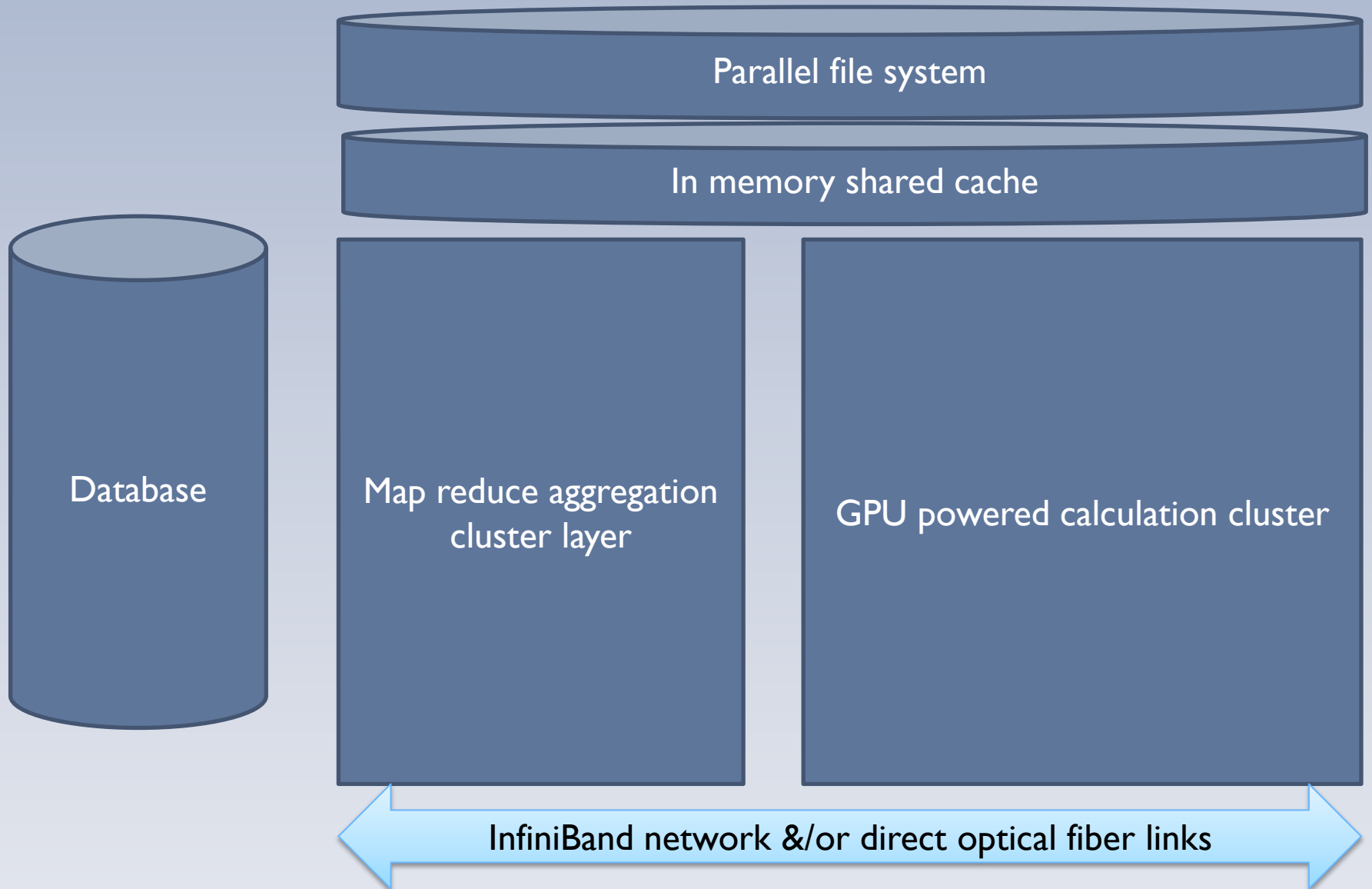
- Need double buffering meaning store while computing the next trades
- Need IB
- Need a parallel file system with SSD & the help of an in memory cache to reduce disk demand for reusable data
- Higher cost but with full flexibility enabling several different aggregations

Aggregate at the level of the CPU or GPU memory before any storage

- Loss of detailed results and per trade contribution
- Force a per netting set treatment of the trades
- CVA sensitivities computation may lead to re-compute useless trades
- Impossible to change the aggregation criteria on demand without re-computation

5 to 10
Tera-Bytes
of data

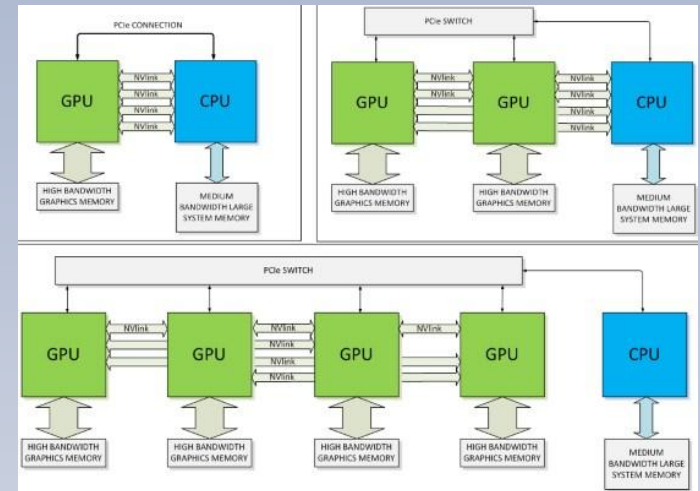
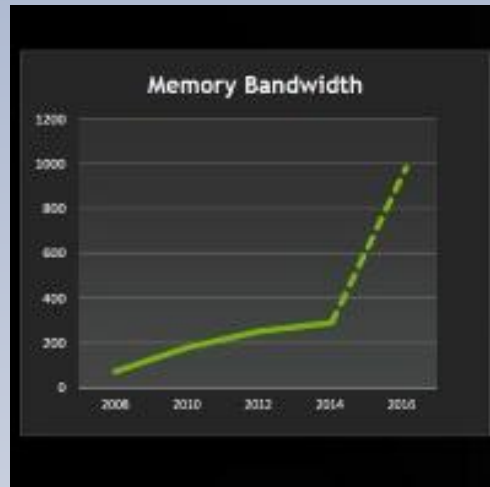
The new normal financial software architecture design



A bit of perspective

3D MEMORY

3D Chip-on-Wafer integration
Many X bandwidth
2.5X capacity
4X energy efficiency



The Pascal GPU should solve by 2016 our two residual concerns

- Memory bandwidth which is often the limitation factor inside the Monte-Carlo thanks to stacked memory
- The PCI express bottleneck replaced by NVLINK which will enable us to extract far more data from the GPU memory to follow the trend of XVA computation

THANK YOU

July, 10th
2014

MUREX NVIDIA GPU COMPUTING IN A NUTSHELL
© Murex 2014



Selected Industry Recognition and Rankings 2013-2014



Structured products
 TECHNOLOGY **RANKINGS 2014**
 TRADING SYSTEMS
 CROSS-ASSET
1 TRADING SYSTEMS
 COMMODITIES
 RISK MANAGEMENT
 COLLATERAL MANAGEMENT

OVERALL #1
TOP
TECHNOLOGY
VENDOR

Risk
 TECHNOLOGY RANKINGS
2013

- #1 Trading Systems
- #1 Pricing & Risk Analytics
- #1 Enterprise-wide Credit Risk Management
- #1 Limit Checking
- #1 System Implementation Efficiency

AsiaRisk
 TECHNOLOGY
 RANKINGS 2013

- #1 Trading Systems
- #1 Pricing & Risk Analytics
- #1 Enterprise-wide Credit & Liquidity Risk Management
- #1 Limit Checking
- #1 System Implementation Efficiency
- #1 After Sales Service



READERS' CHOICE AWARDS
 BEST RISK MANAGEMENT
 PRODUCT OR SERVICE
Murex

RiskTech
100[®]

2014 Category Winner
 Capital Markets – Sell-Side

#1
 Best Selling Wholesale
 Banking System
 Murex
 MX.3

IBS journal
 Sales
 League
 Table
 2014

A worldwide presence



Murex

A mature company in constant revolution

Top Gartner quadrant integrated risk & trading solution

Only global vendor of our type focusing also on analytics

We start NVIDIA GPU programming in 2008

We are live @ several sites with GPU with

- Less hardware
- Faster & more accurate results

Disruptive solution resulting in more

- Advanced models usage
- Volume – x 3 in a few months for our first customer -
- Control

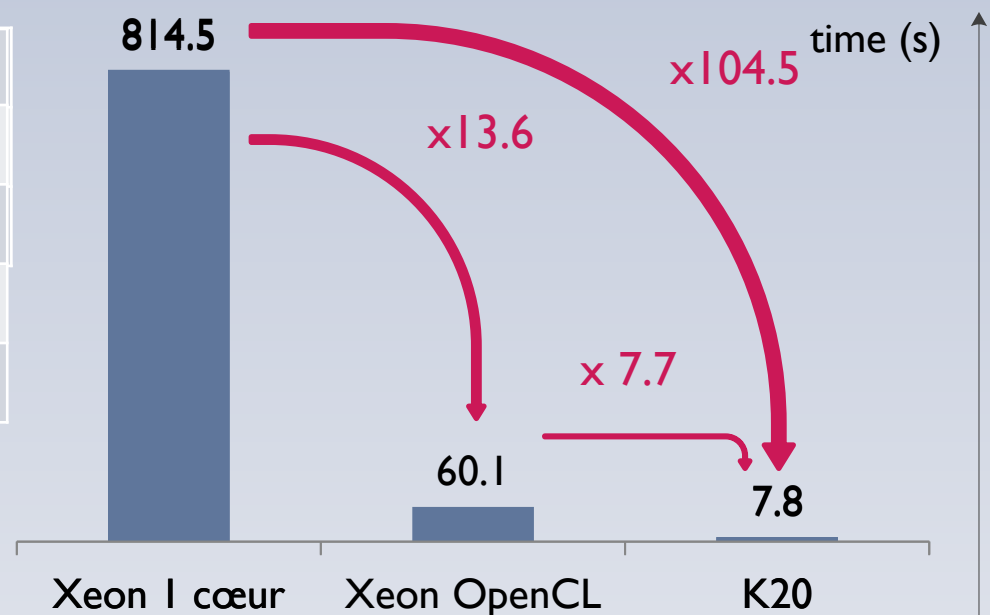
GPU for exotic options valuation and risk management

Performance benchmark

20 different exotic products on a mono-threaded CPU then on GPU.
For each valuation, we compute hundreds of thousands of path to find a fair price.

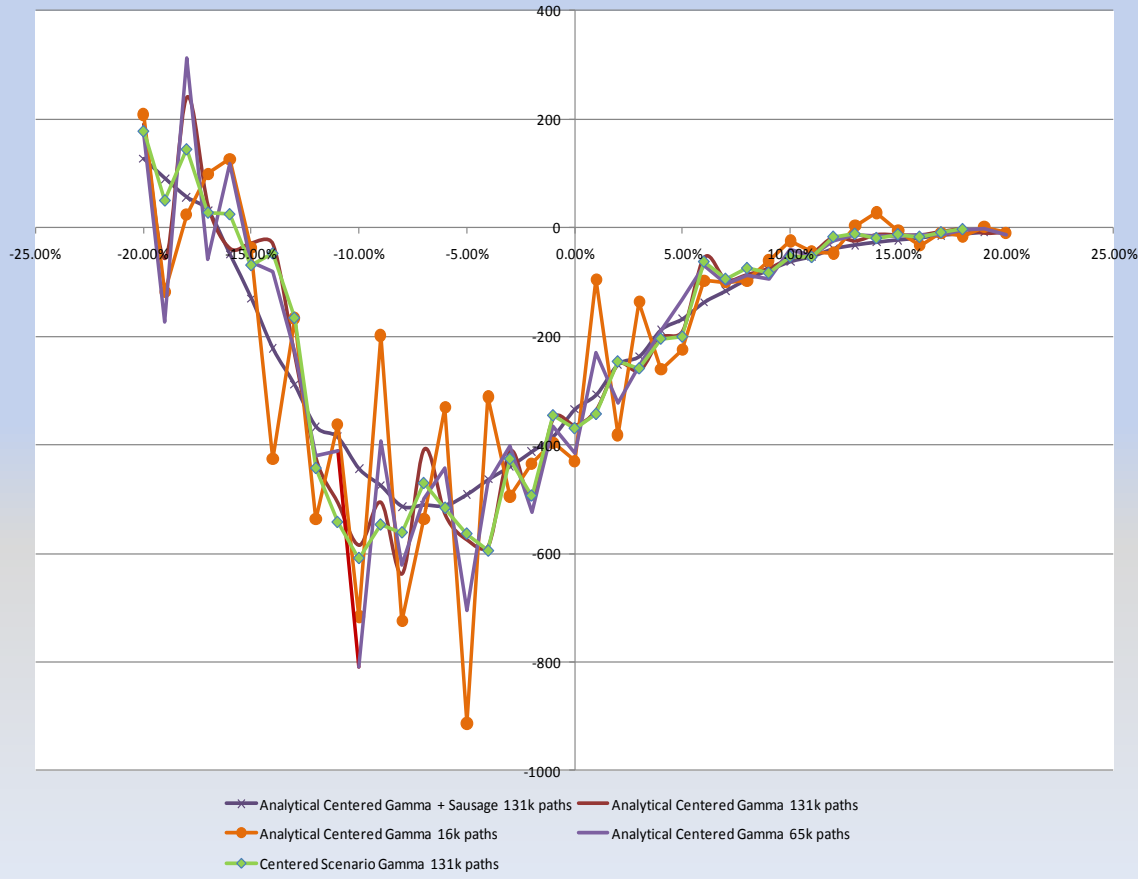
- GPU : NVIDIA Tesla K20c (13x192 SPs, 705 Mhz)
- CPU : Intel Xeon W3690 (one core for this test, 3.46 Ghz)

Xeon 1 core No SSE	814.5 s
K20	7.8 s
K20 / Xeon 1 core	104.5
Xeon OpenCL SSE	60.1
K20 / Xeon OpenCL	7.7



GPU for exotic options valuation and risk management

11HK Spot scenario impact on 11HK Gamma

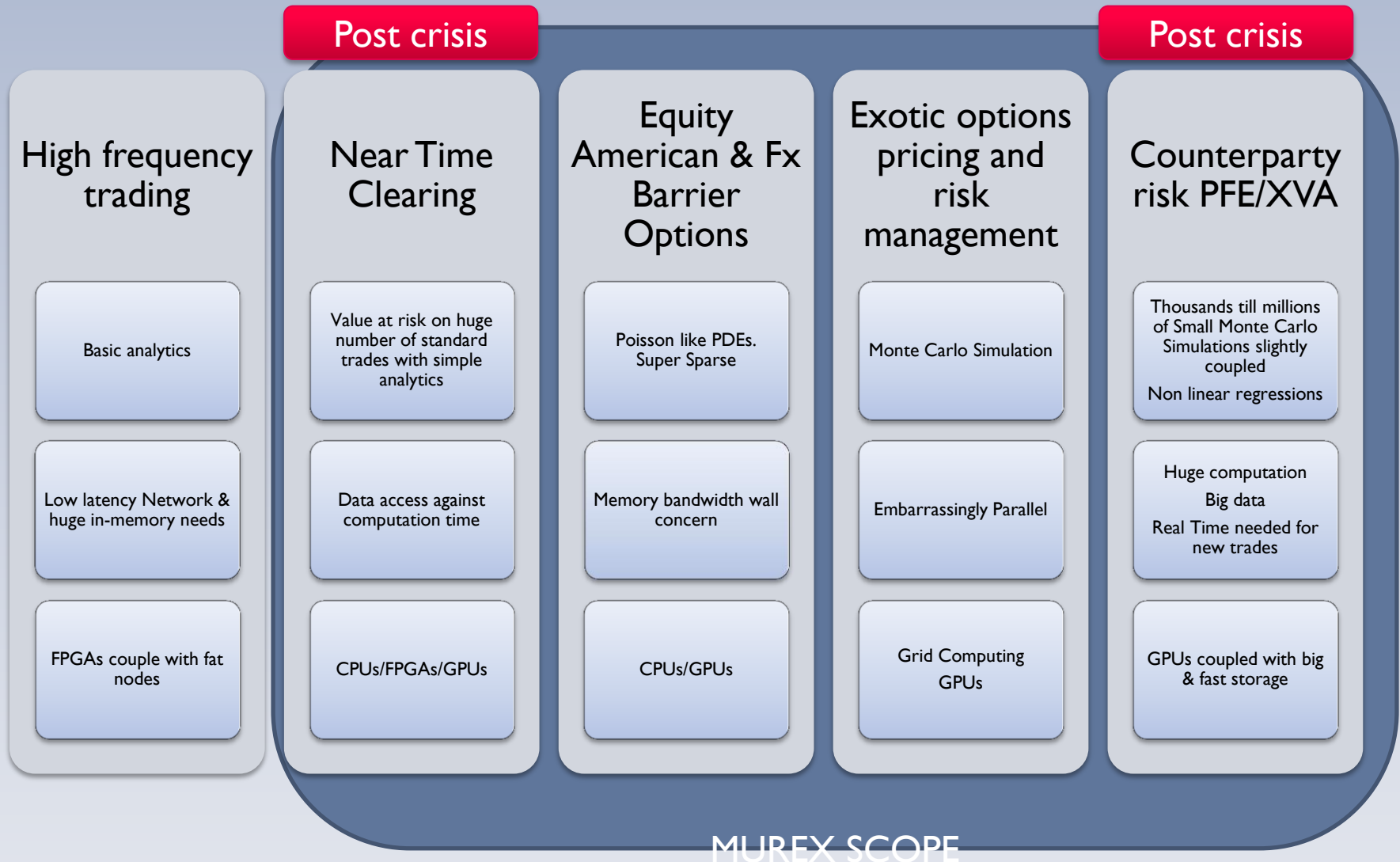


Settings tested

- 8k MC paths
- 16k paths
- 32k paths
- 65k paths
- 131k paths
- 131k paths with Sausage

With GPU our customer do not have tradeoff between accuracy and speed and often gain in both

FINANCIAL PROBLEMS/HPC Mapping



MUREX SCOPE

Our vision

In the post crisis world real time will become mandatory everywhere and HPC is needed

GPU is the easiest & cheapest way to enter the HPC market

GPU & High performance computing is changing the way a system is built

- Take time to invest
- Do not port but rethink & rewrite

Gains are huge and sustainable whatever the new hardware will be

Do not be afraid of so called disruptive technologies & Be ready before they become common place

THANK YOU