

# **F10: Parallel Computing for Actuaries**

Matthew Evans - EMC Actuarial and Analytics Jan Iwanik - Iwanik Limited



#### and Faculty of Actuaries

# **F10: Parallel Computing for Actuaries**

#### From Excel to GPU Clusters

17 October 2018

## Agenda

- Introduction
  - What is parallel processing and which processes are susceptible?
  - Why are we talking about this now?
- Performance gains
  - Excel, R, to C++ on GPUs
- Example Pricing Applications
  - Personal Lines Impact Analysis
  - Commercial Lines Claim Simulation
- Concluding remarks



#### Introduction

- Parallel processing in CPU cores and GPU cores
  - A CPU has a handful of cores for processing multiple *independent* threads
  - A GPU has hundreds or cores for processing multiple *similar* threads in parallel





#### Introduction

- Susceptible Processes
  - Cannot be "sequential"
  - Must permit separate calculation of elements, then later combination
- Examples for actuaries
  - Monte Carlo, including Cat modelling and other loss simulation
  - GLM fitting
  - Cross-Validation for testing model fits
  - Machine Learning



#### Introduction

• Why are we talking about this now?







#### **Performance Gains**

- These days, most computers have 4 or more CPU processor cores
- They also have additional GPU processor cores

10<sup>7</sup> Transistors (thousands) 10<sup>6</sup> Single-Thread 10<sup>5</sup> Performance (SpecINT x 10<sup>3</sup>)  $10^{4}$ Frequency (MHz) 10<sup>3</sup> **Typical Power**  $10^{2}$ (Watts) Number of  $10^{1}$ Logical Cores 10<sup>0</sup> 1970 1980 1990 2000 2010 2020 Year

40 Years of Microprocessor Trend Data

Original data up to the year 2010 collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond, and C. Batten New plot and data collected for 2010-2015 by K. Rupp

#### **Parallel calculations in Excel**

- Excel always attempts to use all available CPU cores...
- ...so you don't have to do anything to benefit

- However, when Excel reports utilising all CPU cores...
- ... it may not be true!



#### How about more advanced tools?





• The objective – predict the distribution of possible outcomes after a hypothetical rate change.

Method – GLM-free prediction methodology (GIRO 2015, PS 2017).

 Parallel component – simulate by bootstrapping a portfolio of 262,144 motor policies. Then apply the predictive calculation.



- We test the following approaches:
  - Excel
  - R
  - C++
  - C++ with BLAST
  - C++ with mining rig





ΤοοΙ	Cores utilised	Time	
Excel, disable multi-threaded calculations	1 CPU	59 s	
Excel, default settings	4 CPU	51 s	
R, clean code using vectors	1 CPU	171 ms	x300 faster
C++	1 CPU	160 ms	
C++, BLAST routines	1 CPU	41 ms	v1000 factor
C++, OpenCL	4 CPU + 4 GPU	7 ms	x1000 faster
C++, OpenCL on mining rig	2 CPU + 6 x 36 GPU = 218	0.16 ms	





#### Calculations per second

- Note the speed increases once multi-threading is deployed at scale
- Difference between mining-rig and typical (good) office PC using R is still a factor of 1000.



#### **Commercial Lines Pricing Example: Claim Simulations**

- Monte Carlo claims simulation for aggregate layer pricing
- Our C# system takes 120 seconds to simulate 1,000,000,000 Generalised Pareto Distribution claims and put through a layer structure
- Using C++ on laptop's 4 CPU cores gives a factor of 6 improvement in speed (18 seconds)
- Processing on laptop's 20 GPU cores gives a further factor of 1.5 improvement (12 seconds)
- Mining rig gives further factor of 5 (2.5 seconds)





#### **Commercial Lines Pricing Example: Claim Simulations**

• A full mining rig may not be practical...

• ...but material improvements are possible with standard, cheap GPU hardware, that in many cases is already present.

• This approach offers a way to improve speed, price higher layers more accurately, and extend the life of computer hardware.



## **Concluding Remarks**

- Computer graphics technology is a gift to actuaries and data scientists
  - We can test more interactions, run more simulations, include more data...
- Limitations of the approach
  - Only some processes are susceptible
  - GLMs show only marginal benefits with R multi-threading packages
  - Interactions (3 way for 10/10 is 720,000 models) still numerous..!
  - Specialist Technical Expertise required
- Quick wins: Monte Carlo; Cross-validation; R & Python tools.
- Slow wins: Exhaustive model interaction testing; Any others..?



## **Concluding Remarks**

- A note on crypto-currencies
- BTC cannot be mined profitably with GPUs
- GPU miners choose ETH or others
- Price has collapsed!
- But good for gamers, data scientists and actuaries?





## **About Us**

- Jan Iwanik
  - Pricing
  - Free Pricing Tools
  - GPU Programming



- Matthew Evans
  - Pricing
  - Reserving
  - Data Science
  - InsurTech



- jan.iwanik@gmail.com
- Iwanik Limited

- <u>matthew@emc-actuarial.com</u>
- EMC Actuarial & Analytics





The views expressed in this [publication/presentation] are those of invited contributors and not necessarily those of the IFoA. The IFoA do not endorse any of the views stated, nor any claims or representations made in this [publication/presentation] and accept no responsibility or liability to any person for loss or damage suffered as a consequence of their placing reliance upon any view, claim or representation made in this [publication/presentation].

The information and expressions of opinion contained in this publication are not intended to be a comprehensive study, nor to provide actuarial advice or advice of any nature and should not be treated as a substitute for specific advice concerning individual situations. On no account may any part of this [publication/presentation] be reproduced without the written permission of the IFoA [or authors, in the case of non-IFoA research].

