



Data Classification Hardware Acceleration

Ivan Manuylov

Senior Software Developer, DLP Research
Kaspersky Lab



Speed



Speed



Speed?



Speed?



Speed?



Speed?



Hardware

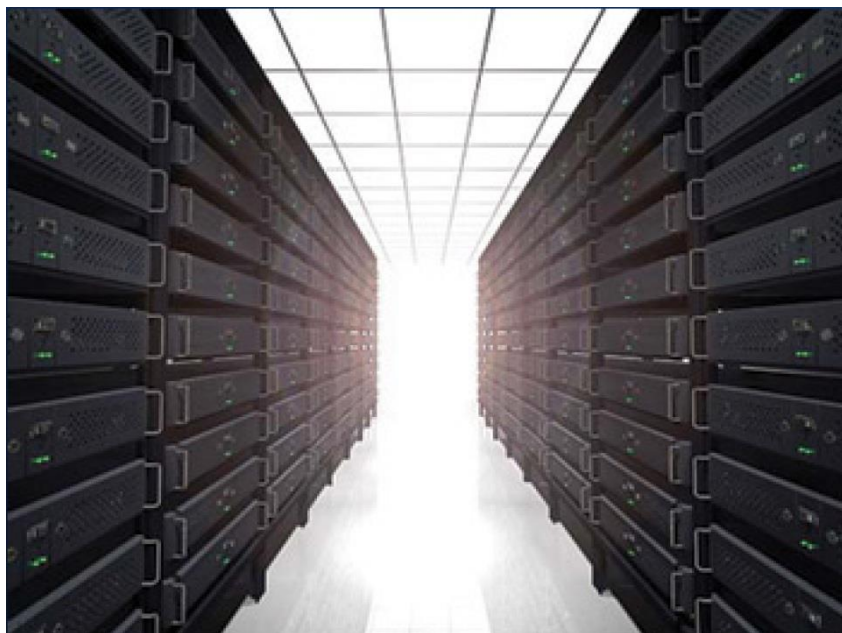


Cores

Who has access to 1000+ cores?

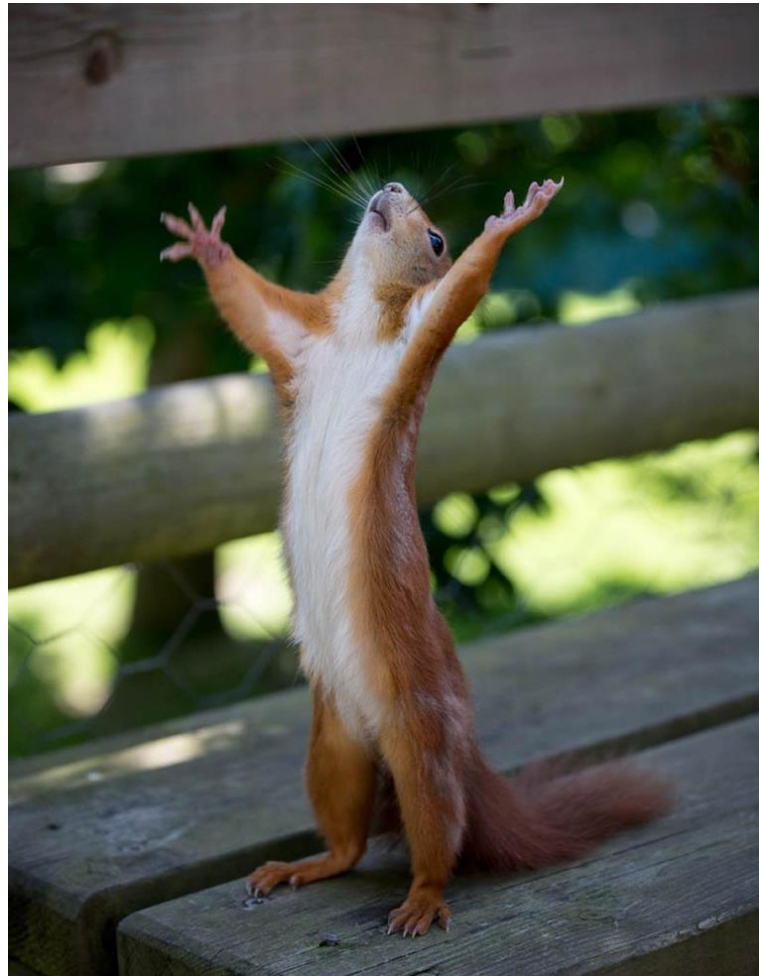
Google

you



Purpose of GPU

General Purpose computation on GPU



How to use GPGPU

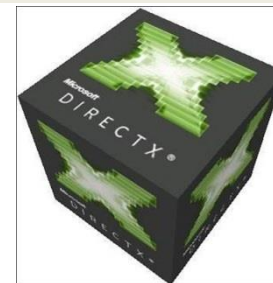


OpenCL

- Cross platform
- Multi vendor
- C kernel
- Languages:
 - C/C++
 - .NET
 - Java
 - Ruby
 - Python



- Cross platform
- NVIDIA
- C/C++ kernel
- Languages:
 - C/C++
 - .NET
 - Java
 - Fortran
 - Perl
 - Python



Direct Compute

- Windows
- Multi vendor
- HLSL kernel
- Languages:
 - C++
 - C++ AMP
 - .NET

Kernel code

```
__kernel void CalculateKGramms(
    unsigned int kFactor,
    unsigned int wFactor,
    unsigned int overlappingSize,
    __global unsigned char* data,
    unsigned int dataSize,
    __global unsigned int* outHashes)
{
    size_t globalIndex = get_group_id(0) * (get_local_size(0) - overlappingSize) + get_local_id(0);
    size_t localId = get_local_id(0);

    __local unsigned int buffer1[SHARED_BUFFER_SIZE];
    __local unsigned int buffer2[SHARED_BUFFER_SIZE];
    if (globalIndex < dataSize)
    {
        buffer1[localId] = data[globalIndex];
    }
    barrier(CLK_LOCAL_MEM_FENCE);

    if ( globalIndex < (dataSize - kFactor + 1) && localId < (SHARED_BUFFER_SIZE - kFactor + 1))
    {
        unsigned int crc32 = 0;
        for (int i = 0; i < kFactor; ++i)
        {
            unsigned char idx = (unsigned char)crc32 ^ (unsigned char)buffer1[localId + i];
            int c = g_deviceCrcPolynomialMap[idx];
            crc32 = crc32 >> 8;
            crc32 = crc32 ^ c;
        }
        buffer2[localId] = crc32;
    }
    barrier(CLK_LOCAL_MEM_FENCE);
}
```

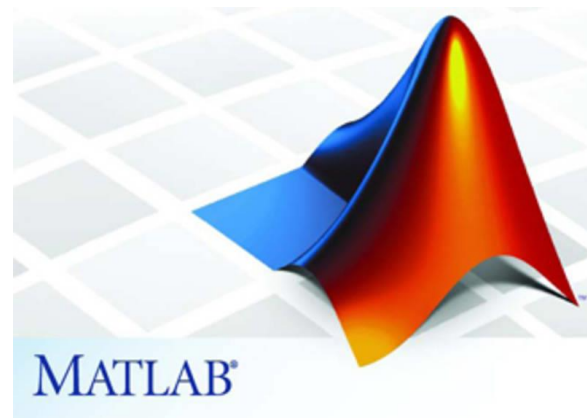
Using of Hardware Acceleration



Using of Hardware Acceleration



ELCOMSOFT
PROACTIVE SOFTWARE



Research



Hardware

CPU

- Intel Core i5 – 750
- 2.66 GHz
- 4 cores
- DDR3:
 - 667 MHz
 - 12 Gb/s

GPU

- NVIDIA GeForce GTS 450
- 1.6 GHz
- 192 cores
- GDDR5:
 - 1.8 GHz
 - 58 Gb/s

Fingerprinting



Fingerprinting

Guidelines and Specifications for Flood Hazard Mapping Partners [November 2004]

D.4.14 Glossary

Most of the coastal engineering terms in this glossary are from the *Shore Protection Manual* (USACE, 1984) and *Coastal Engineering Manual* (USACE, 2002) and are supplemented with additional terms relevant to hazard mapping. FEMA has an extensive glossary posted on the FEMA website at http://www.fema.gov/fhm/dl_cgs.shtml <Glossary>.



----- A -----

ACCRETION May be either natural or artificial. Natural accretion is the buildup of land, solely by the action of the forces of nature, on a beach by deposition of water- or airborne material. Artificial accretion is a similar buildup of land by reason of an act of man, such as the accretion formed by a GROIN, BREAKWATER, or beach fill deposited by mechanical means. Also AGGRADATION.

ADJUSTABLE GROIN A GROIN whose permeability can be changed, usually with gates or removable sections.

ADVANCE (of a beach) (1) A continuing seaward movement of the shoreline. (2) A net seaward movement of the shoreline over a specified time. Also PROGRESSION.

AEOLIAN See EOLIAN.

ALIGNMENT The course along which the center line of a channel, canal, or drain is located.

ALLUVIAL DEPOSITS Detrital material which is transported by a river and deposited usually temporarily at points along the flood plain of a river. Commonly composed of sands and gravels.

ALLUVIAL PLANE A plain bordering a river, formed by the deposition of material eroded from areas of higher elevation.

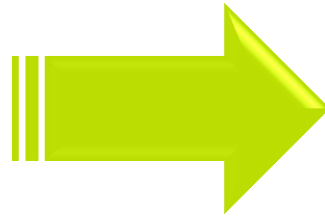
ALLUVIUM Soil (sand, mud, or similar detrital material) deposited by streams, or the deposits formed.

ALONGSHORE Parallel to and near the shoreline: LONGSHORE.

Fingerprinting



100 Kb



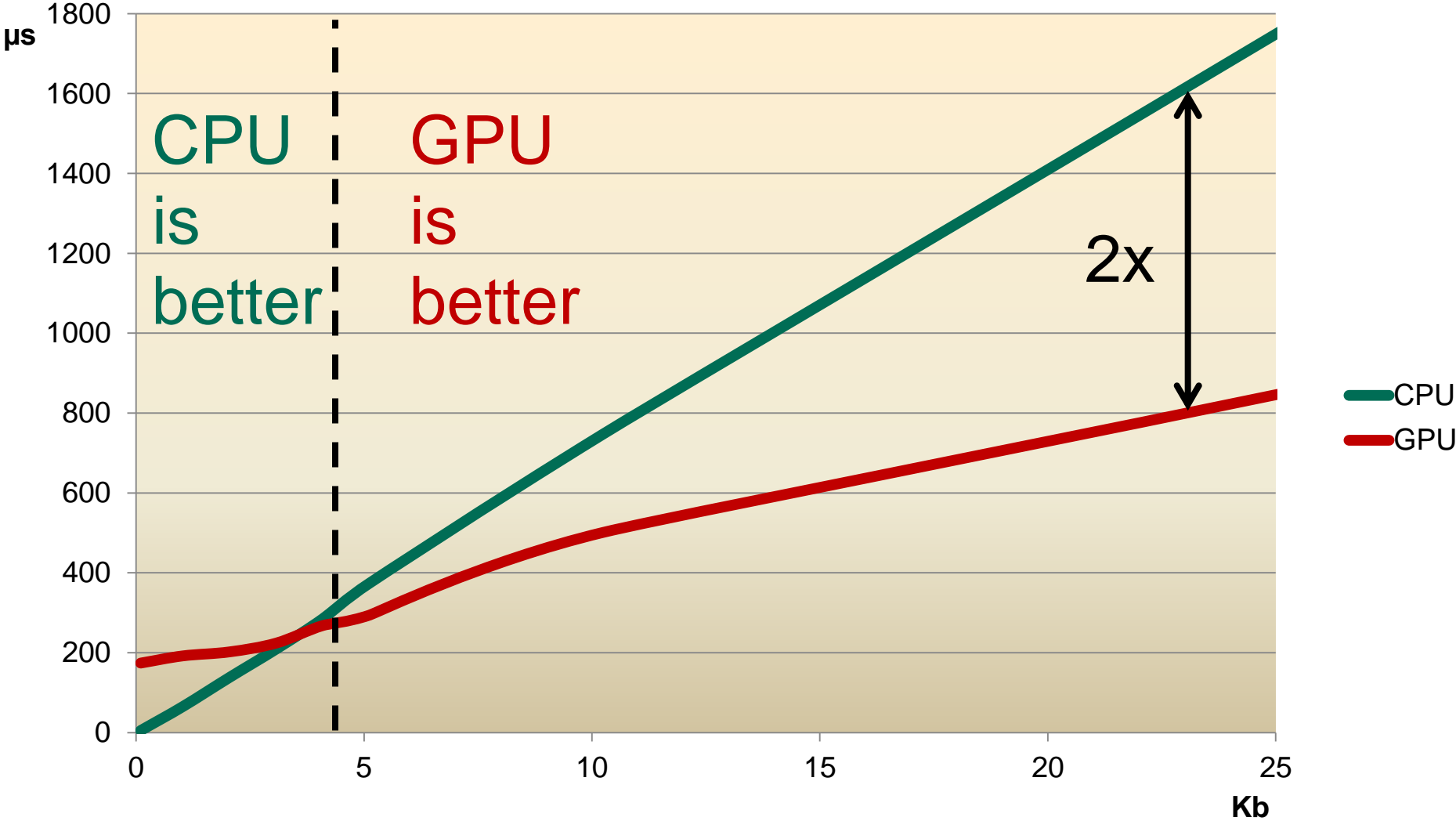
100 000 hashes

Fingerprinting

2x faster



Fingerprinting



Keywords



Keywords

Guidelines and Specifications for Flood Hazard Mapping Partners [November 2004]

D.4.14 Glossary

Most of the coastal engineering terms in this glossary are from the [Shore Protection Manual](#) (USACE, 1984) and [Coastal Engineering Manual](#) (USACE 2002) and are supplemented with additional terms relevant to hazard mapping. FEMA has an extensive glossary posted on the FEMA website at http://www.fema.gov/fhm/dl_cgs.shtm *Glossary*.

----- A -----

ACCRETION May be either natural or artificial. Natural accretion is the buildup of land, solely by the action of the forces of nature, on a beach by deposition of water- or airborne material. [Artificial accretion](#) is a similar buildup of land by reason of an act of man, such as the accretion formed by a GROIN, BREAKWATER, or beach fill deposited by mechanical means. Also AGGRADATION.

ADJUSTABLE GROIN A GROIN whose permeability can be changed, usually with gates or removable sections.

ADVANCE (of a beach) (1) A continuing seaward movement of the shoreline. (2) A net seaward movement of the shoreline over a specified time. Also PROGRESSION.

AEOLIAN See EOLIAN.

ALIGNMENT The course along which the center line of a channel, canal or drain is located.

ALLUVIAL DEPOSITS Detrital material which is transported by a river and deposited usually temporarily at points along the flood plain of a river. Commonly composed of sands and gravels.

ALLUVIAL PLANE A plain bordering a river, formed by the deposition of material eroded from areas of higher elevation.

ALLUVIUM Soil (sand, mud, or similar detrital material) deposited by streams, or the deposits formed.

ALONGSHORE Parallel to and near the shoreline: LONGSHORE.



[Shore Protection Manual](#)
Artificial Intelligence
Compensation
Stock Option
401k
401(k)
401-k
Vesting Period
Salary
wages
Base wage rate
Base rate
BFOQ
career ladder
Compa-ratio
Cost of Living adjustment
hiring rate
hourly rate
incentive plan
[Artificial Accretion](#)
knowledge base pay
lump sum increase
pay adjustment
pay plan
payment plan
pay range
accrual leave plans
apprentice rate
back pay
pension
compensatory leave
signing bonus
copayment
deductible
[USACE](#)
deferred earnings

Keywords

What is the performance boost?

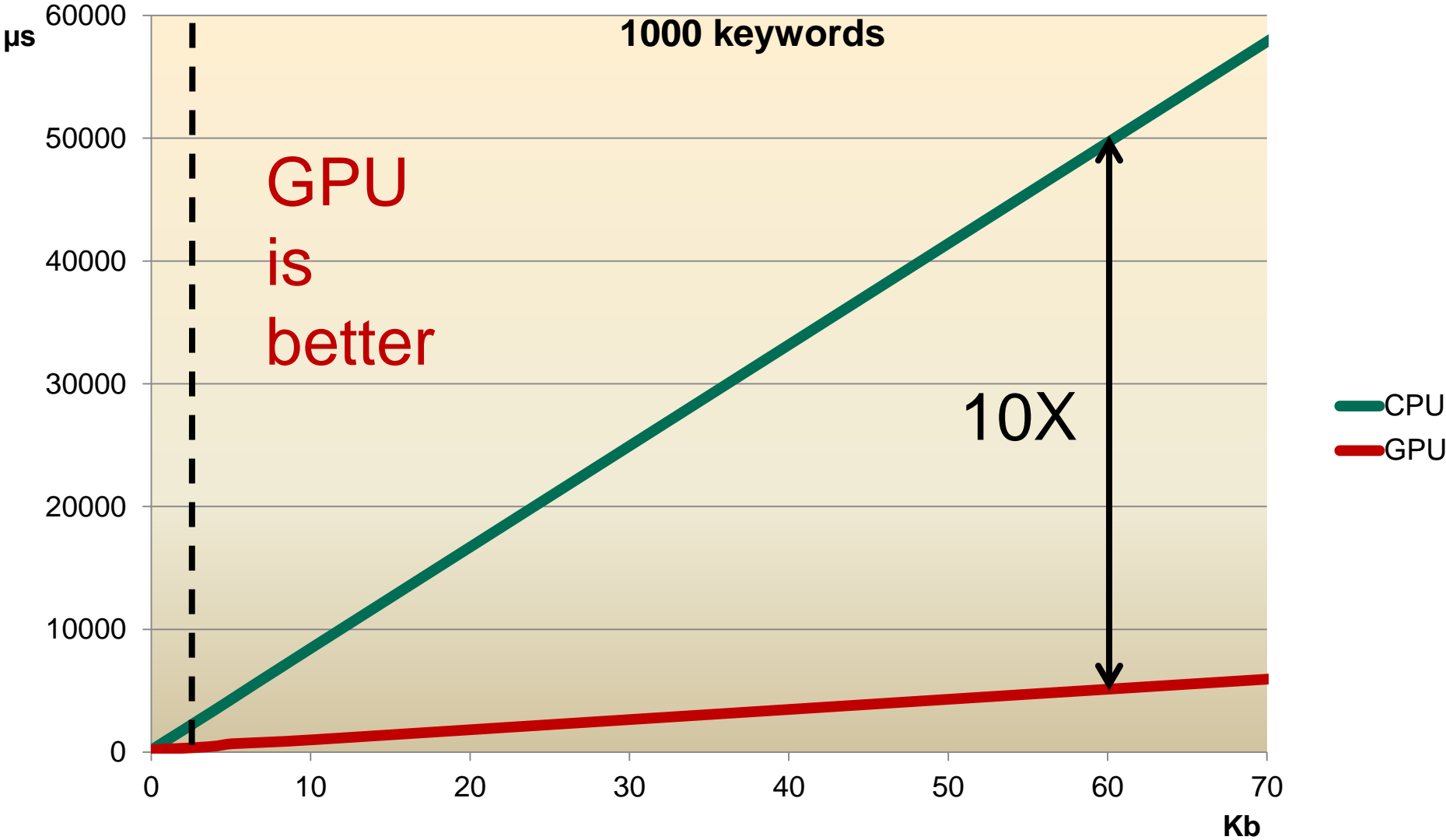


Keywords

What is the performance boost?

10+
times

Keywords



Epilogue



Summary



Fingerprints search **2x**



Keywords search **10x**

We can do it!



Data Classification Hardware Acceleration

Thank you!

Ivan Manuylov

Senior Software Developer, DLP Research,

Kaspersky Lab

Ivan.Manuylov@kaspersky.com

