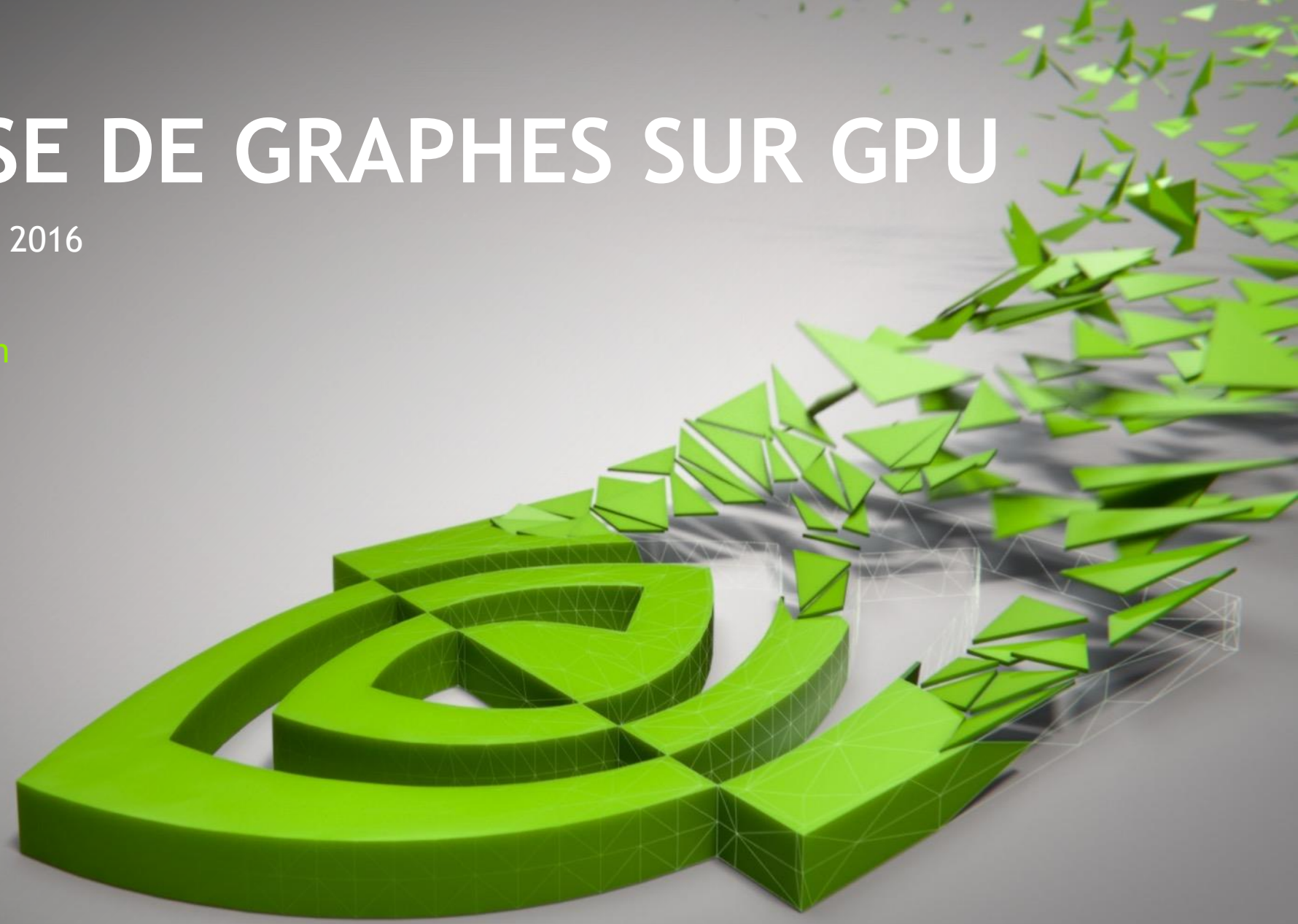
ANALYSE DE GRAPHES SUR GPU

11e Forum Teratec - 2016

Alexandre Fender

afender@nvidia.com

NVIDIA



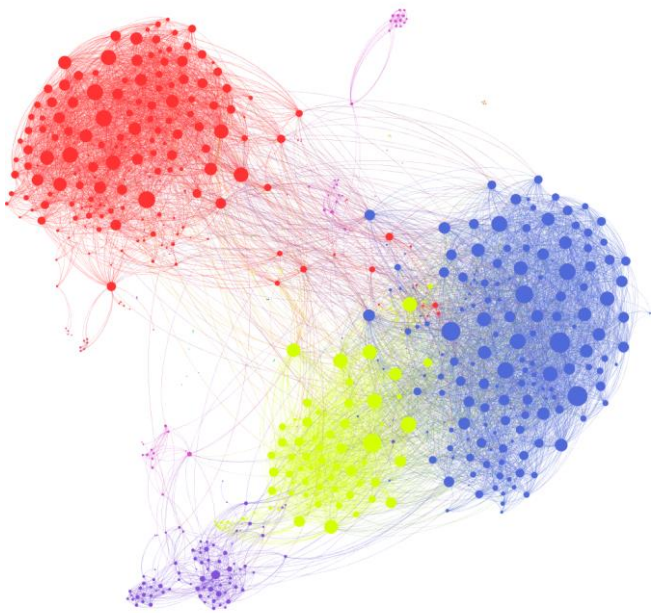
Accelerated graph analytics

nvGRAPH, presentation and performances

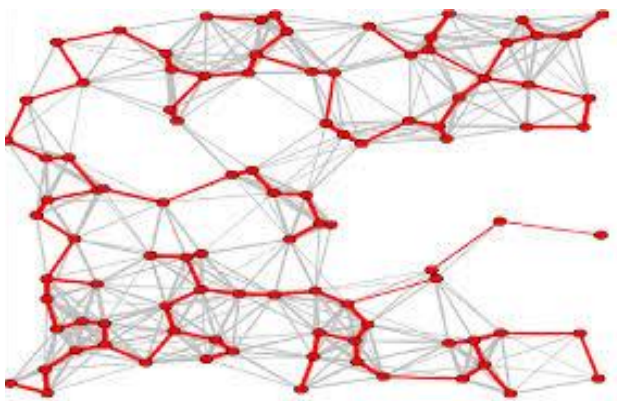
Research in spectral graph analysis

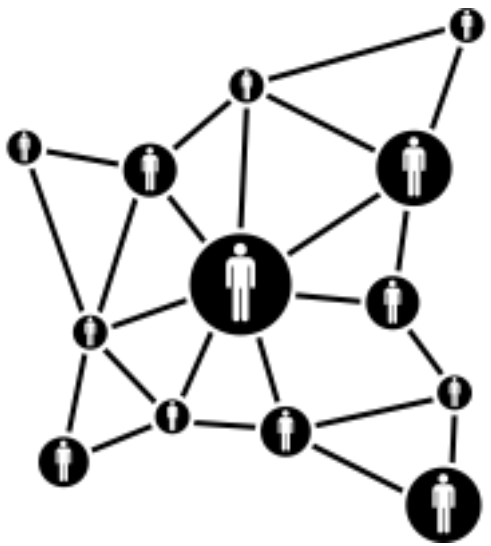
Graph Analytics

Understanding relationships among high volumes of connected data

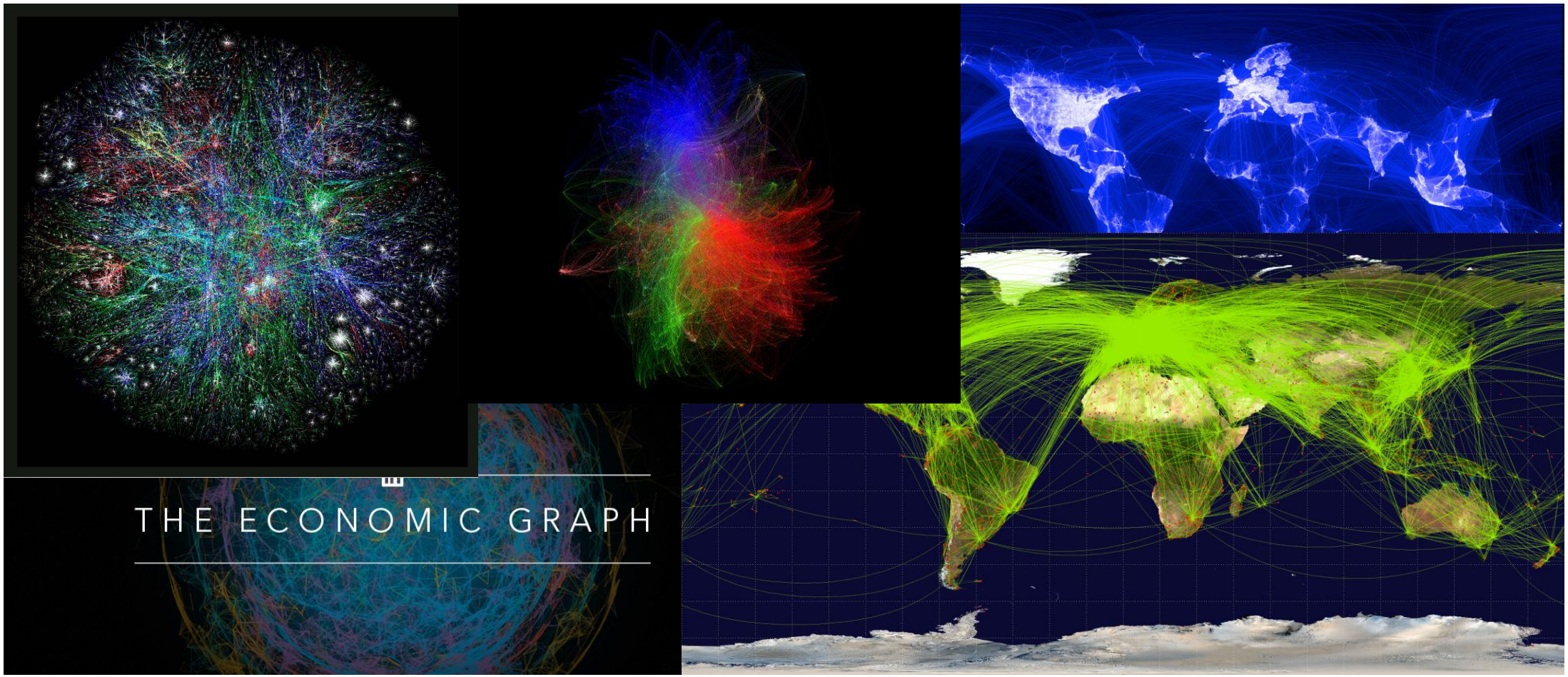


Social Network Analysis
Recommender systems
Parallel computing
Cyber security/Network analytics
Natural language processing
Neural networks
…





Getting more attention



Accelerated graph analytics

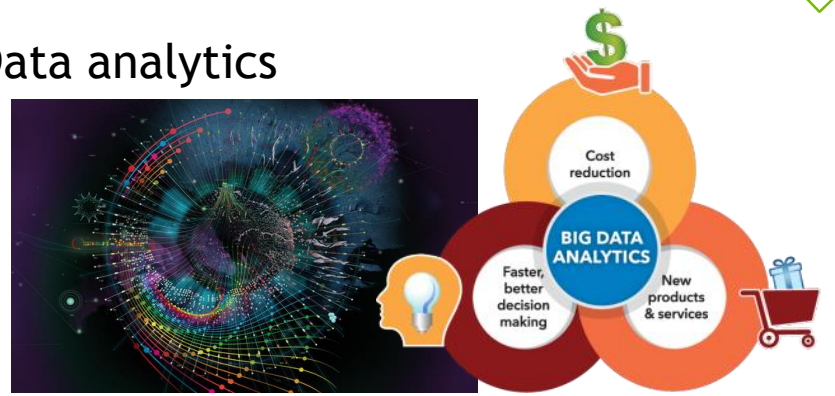
Connecting the dots

Graphs

Linear Algebra

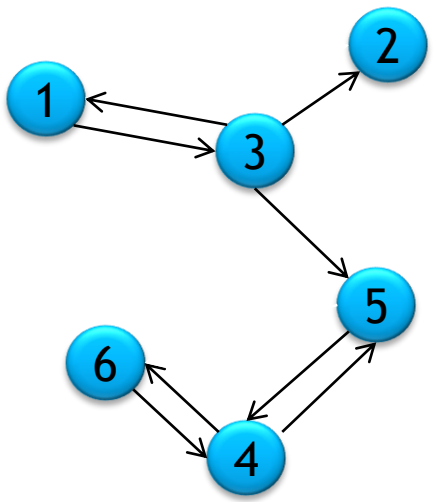
$$Av = \lambda v$$

Data analytics



GPU

Graphs as Matrices

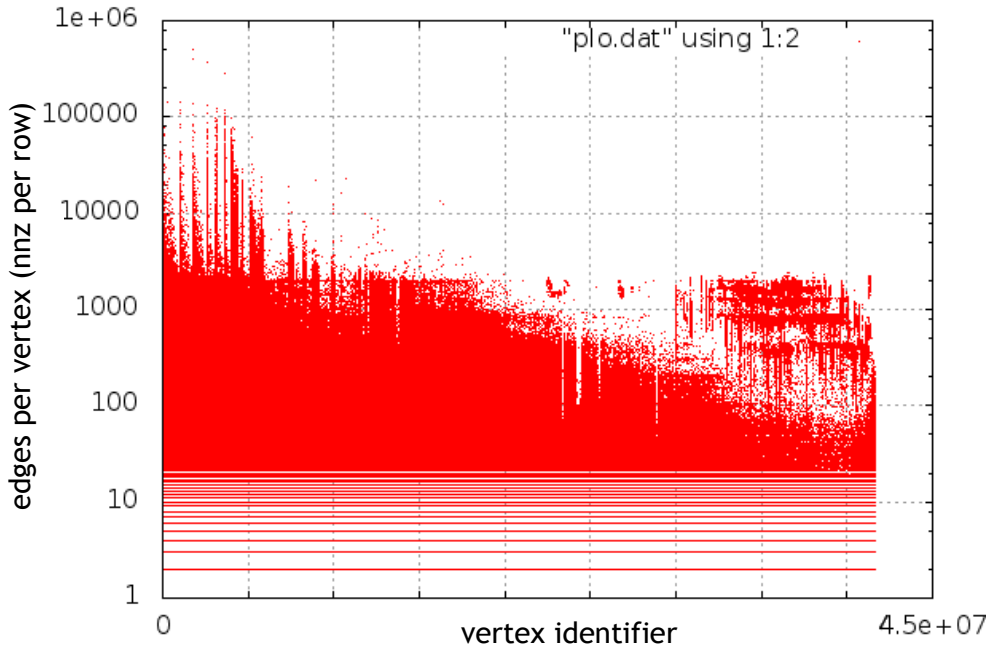


	1	2	3	4	5	6
1	0	1	1	0	0	0
2	0	0	0	0	0	0
3	1	1	0	0	1	0
4	0	0	0	0	1	1
5	0	0	0	1	0	1
6	0	0	0	1	0	0

The input

Twitter data set (2011)

Number of vertices (n)	41,652,230
Number of edges (nnz)	1,468,365,182
Sparsity (nnz per row)	35.25
Minimum nnz per row	0
Maximum nnz per row	770,155
Standard deviation	354



The largest vertex (row) has almost 1 million edges

50% of vertices (rows) have less than 5 edges (non-zero)

nvGRAPH

nvGRAPH

Accelerated Graph Analytics

Part of the CUDA Toolkit 8.0, with a C API.

Algorithms : Pagerank, Semi-ring SpMV, SSSP, SSWP, Graph extraction

Solve graphs with up to 2 Billion edges on a single GPU (M40-24GB)

Efficient on networks with power law distribution

developer.nvidia.com/nvgraph

Pagerank

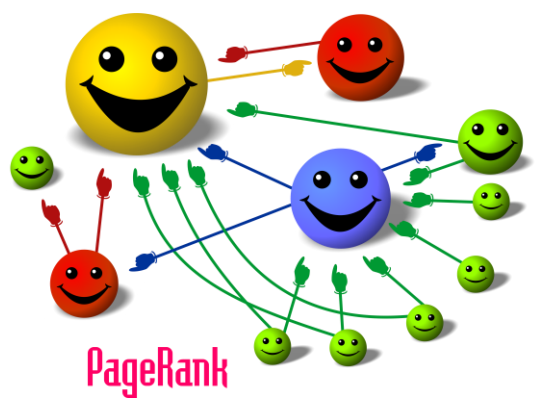
Measure the importance of elements based on the overall topology

Iterative approximation of the largest eigenvector x_{max}

$$x_{k+1} = \frac{G*xk}{||G*xk||}$$

$$G^*x_k = \underbrace{\alpha H^T x_k}_{\text{SpMV}} + \overbrace{\underbrace{b(\alpha a+(1-\alpha)e^T)x_k}_{\text{Dot product}}}^{\text{Constant}}$$

G : Google matrix
H : Markov chain transition matrix
Method : Power Iteration

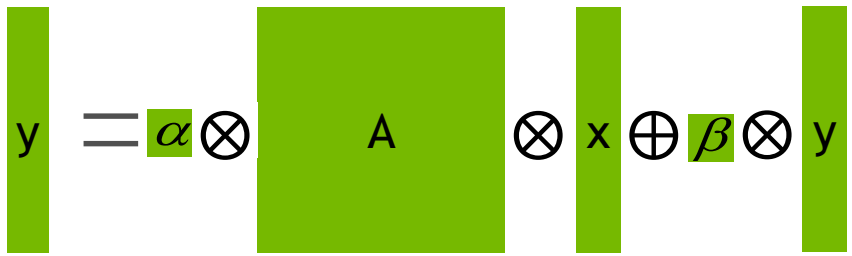


Semi-Ring Matrix Vector Multiplication

Overload of Mat-Vec operators

$$y = \alpha \otimes A \otimes x \oplus \beta \otimes y$$

Where $\otimes$ and $\oplus$ are binary operators satisfying semi-ring properties



SSSP example

Tropical semiring (min, +)

Addition of *a* and *b* is *min* (*a*,*b*)

Multiplication is regular addition

Elementwise Operations in row *i*: $y = min(A^T_{ij_0} + x_{j_0}, A^T_{ij_1} + x_{j_1}, \ldots)$

SSSP formula becomes : $x_{i+1} = \min(A^T x_i, x_i)$

Semi-rings

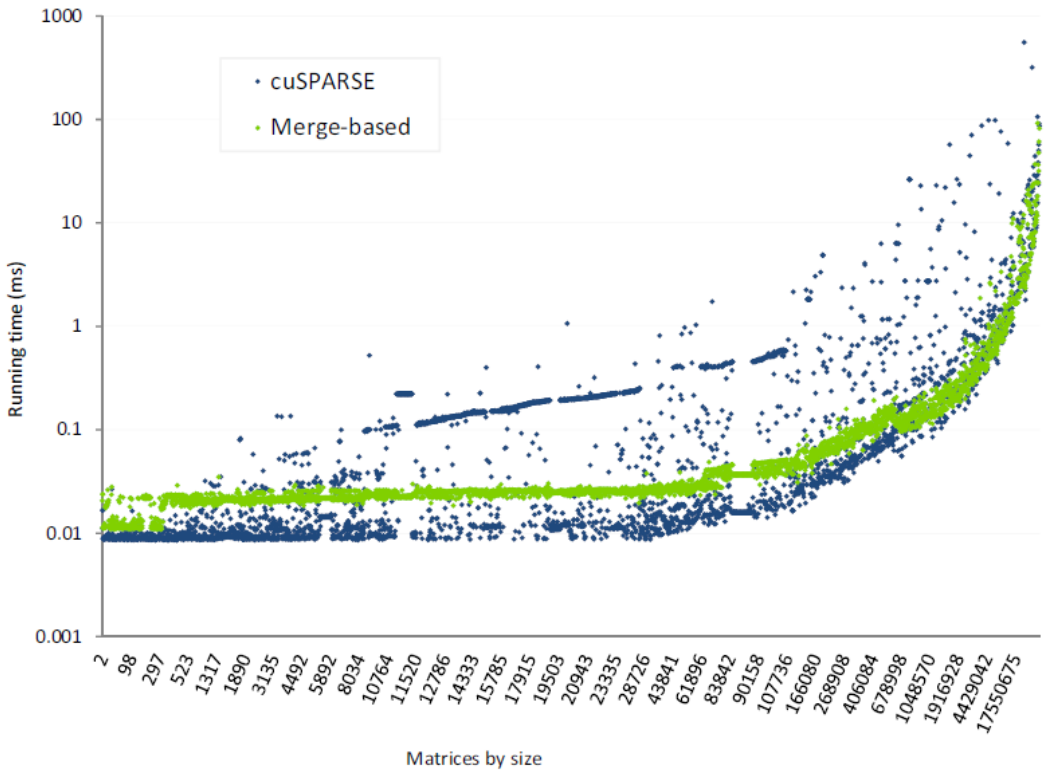
Examples

SEMIRING	SET	PLUS	TIMES	0	1
Real	$\mathbb{R}$	+	*	0	1
MinPlus	$\mathbb{R} \cup \{-\infty, \infty\}$	min	+	∞	0
MaxMin	$\mathbb{R} \cup \{-\infty, \infty\}$	max	min	-∞	∞
Boolean	$\{0, 1\}$	$\vee$	$\wedge$	0	1

nvGRAPH’s performances

General SPMV-MP performance

The entire Florida Sparse Matrix Collection: 4.2K datasets (K40, fp64)



Higher correlation of runtime to problem size

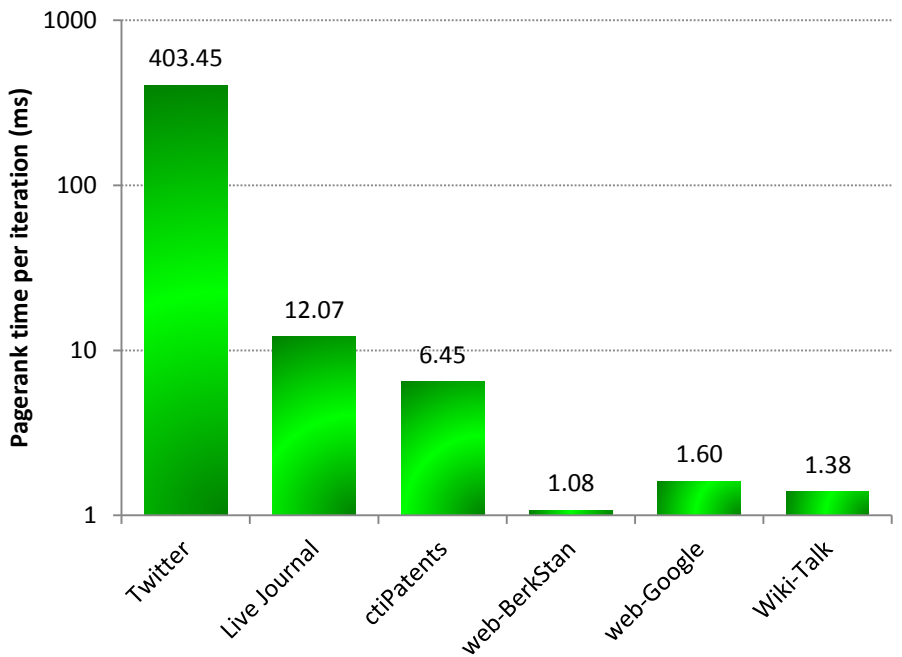
Lower correlation of FLOPS to row-length variation

Merrill and Garland. Merge-based Parallel Sparse Matrix-Vector Multiplication using the CSR Storage Format. Tech. Rep. NVR-2015-002, NVIDIA Corp. 2015

Pagerank performance

Large real networks with power law distribution

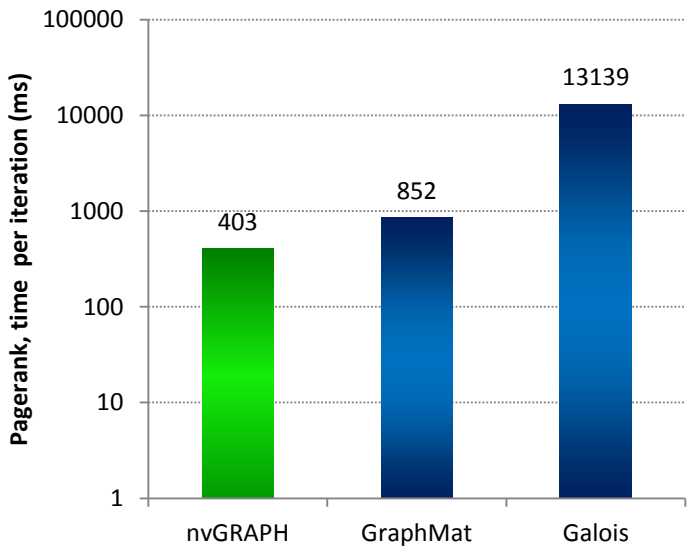
Dataset	Vertices	Edges
Twitter	41652230	1468365182
Live Journal	5363260	79023142
ctiPatents	3774768	16518948
web-BerkStan	685230	7600595
web-Google	916428	5105039
Wiki-Talk	2394385	5021410



Tesla M40
damping factor = 0,85

Pagerank performance

How does it compare to others?



Speedup from 2x to 20x
1 GPU (Tesla M40) vs 2 CPU (32 cores)

Dataset : Twitter, damping factor = 0,85

nvGraph on Tesla M40

GraphMat and Gallois on 2 socket, 32 core, Xeon E5-2698 v3@2.3GHz 3.6GHz Turbo (Haswell) HT off

Internal benchmark

See also:

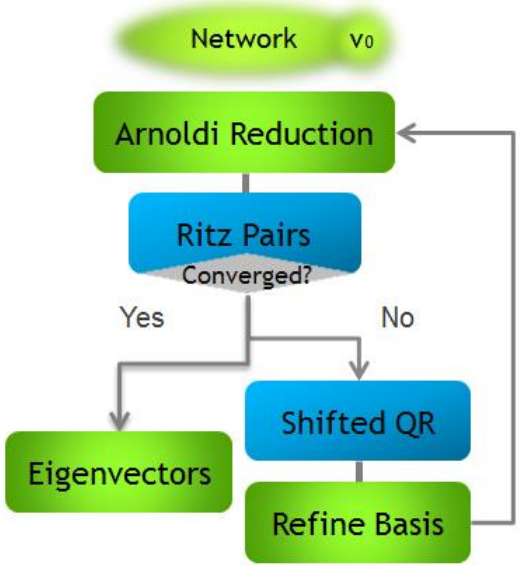
GraphMat: High performance graph analytics made productive, Narayanan Sundaram et. al., Tech. Rep. INTEL Corp, 2015

Research in spectral graph analysis

More advanced algorithms for spectral analysis

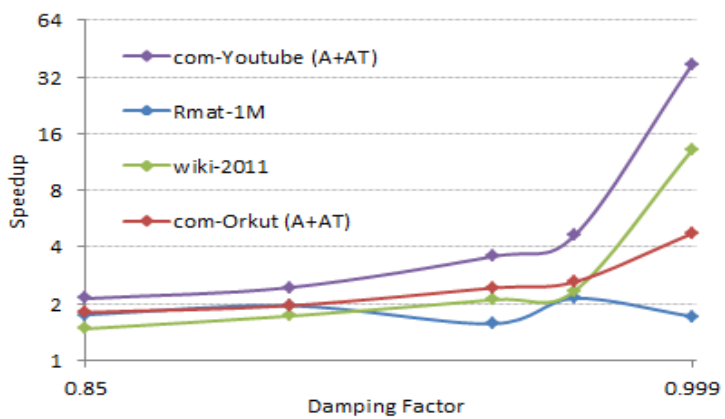
Ongoing research

Implicitly restarted Krylov methods

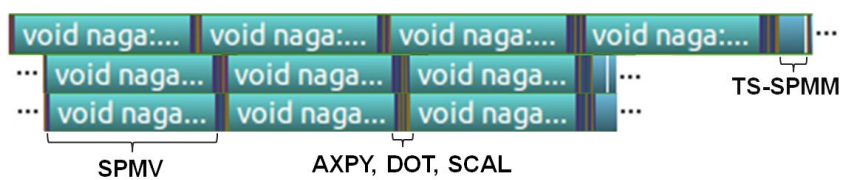


Green = Device - Blue = Host

Hybrid (GPU/CPU) algorithm
Average speedup (over power iteration on GPU) from 1.8x to 14.4x

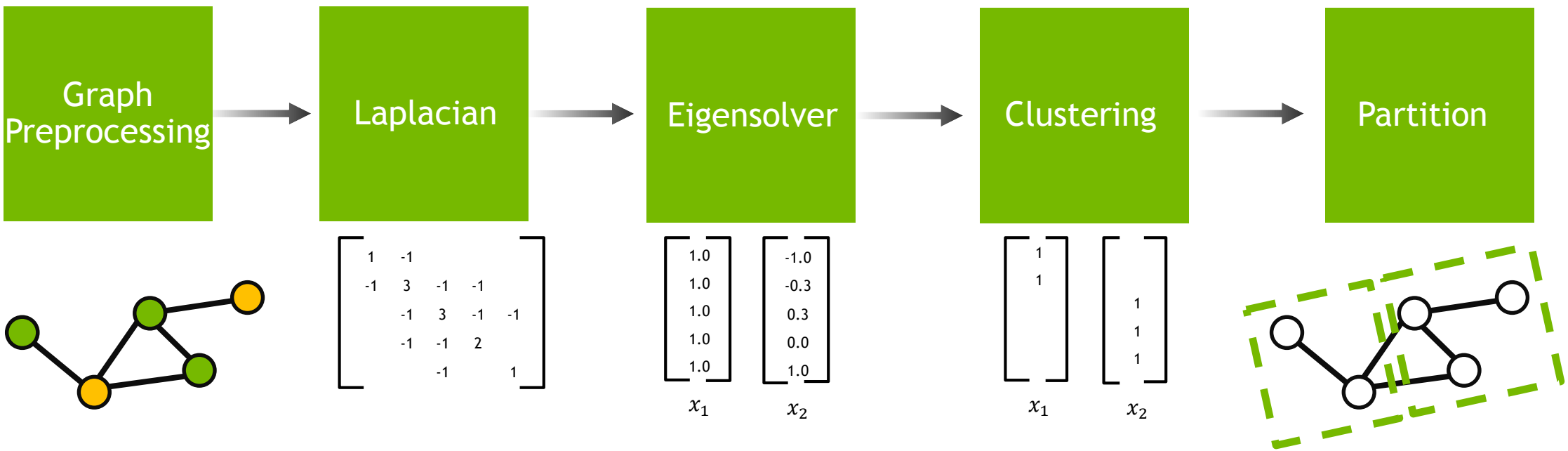


Profile



Fender, Emad, Eaton, Petiton. 2016. “Accelerated Hybrid Approach for Spectral Problems Arising in Graph Analytics.” Procedia Computer Science 80:2338–47.

Spectral partitioning /clustering



Naumov and Moon. Parallel Spectral Graph Partitioning. Tech. Rep. NVIDIA Corp. 2016

Thank you!

nvGRAPH is freely available as part of CUDA 8.0 toolkit

www.nvidia.com/getcuda

EA feedback

“Shows lots of promise, looks like it is going to be a great library”

“Pagerank's performance is awesome”

“Looking forward keep using it”

ANALYSE DE GRAPHES SUR GPU

11e Forum Teratec - 2016

Alexandre Fender

afender@nvidia.com

NVIDIA