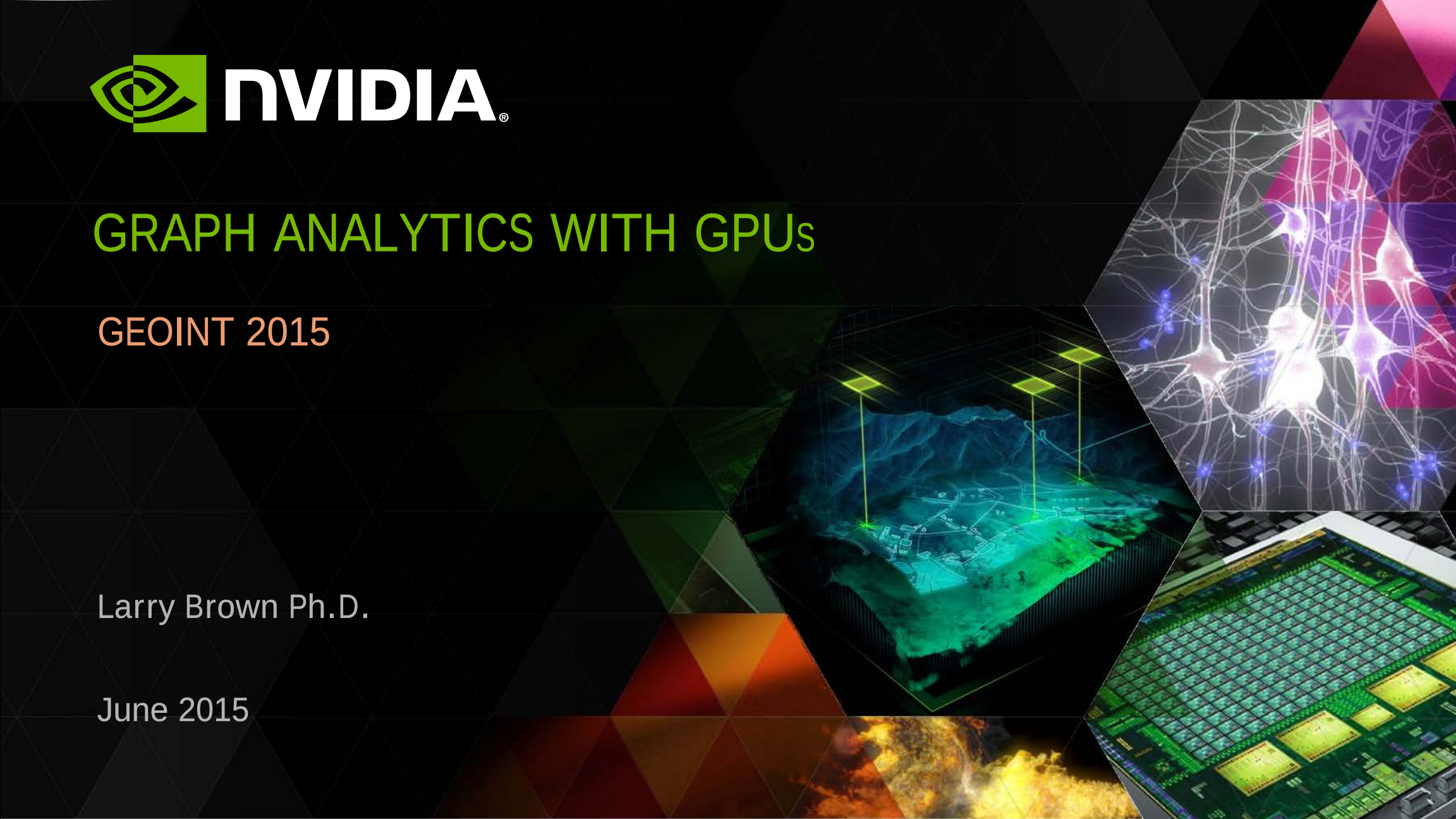


GRAPH ANALYTICS WITH GPUs

GEOINT 2015

Larry Brown Ph.D.

June 2015

AGENDA

1 What is Graph Analytics?
2 Graph Analytics for GEOINT
3 GPUs and Graph Analytics
4 GPU Roadmap for Graph Analytics

...and a video!

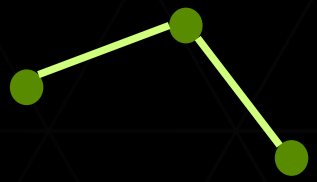
What is Graph Analytics?

GRAPH ANALYTICS - INTRODUCTION

Any analytics performed on a Graph

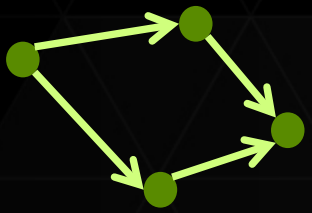
- Graph is “just another data structure”, like a tree or an array.
- Focus is on pairwise relationship between two objects at a time.

- Edges and Nodes
- Edges can be Directed, or Undirected.



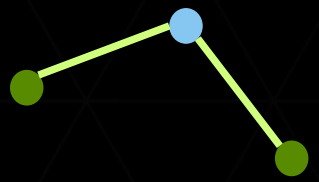
Examples of data well-suited to Graphs

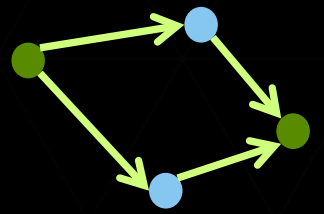
- Road networks
- Communications network
- Social networks
- Web pages & links (directed graph)
- Financial transaction data



KEY GRAPH ALGORITHMS

- Search
- Cutting / Partitioning
- Shortest Path
- Connected components
- Pagerank





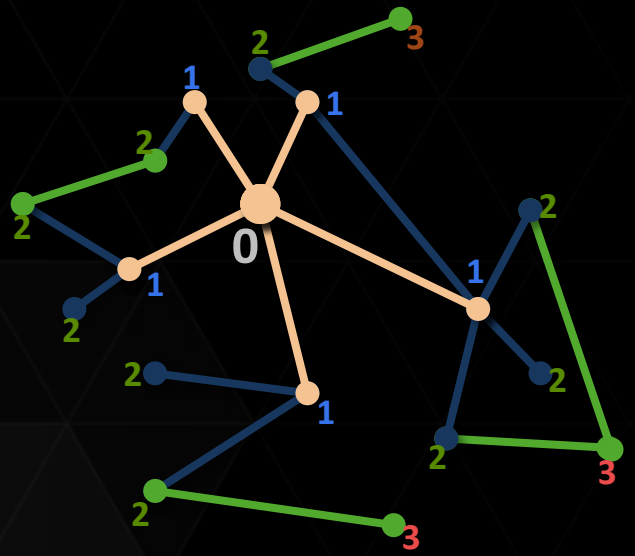
Often tell us something non-obvious about the graph.

Who is the most important person on Twitter?
What websites are most likely to be visited?

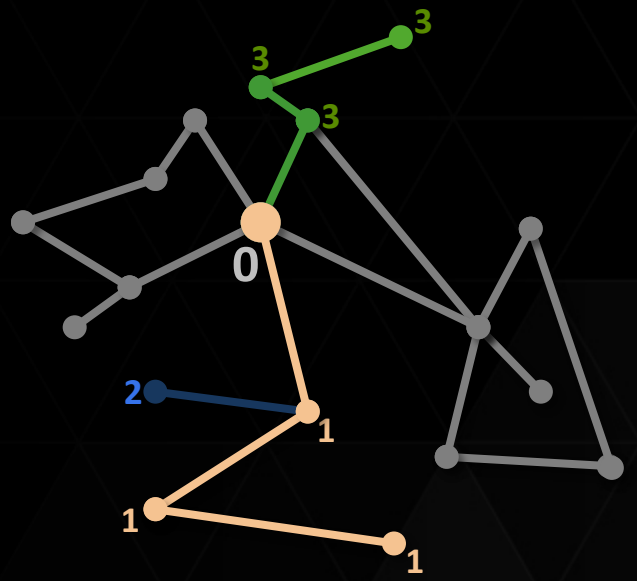
GRAPH ALGORITHMS

- Search
- Cutting / Partitioning
- Shortest Path
- Connected components
- Pagerank

- Breadth First search



- Depth First search

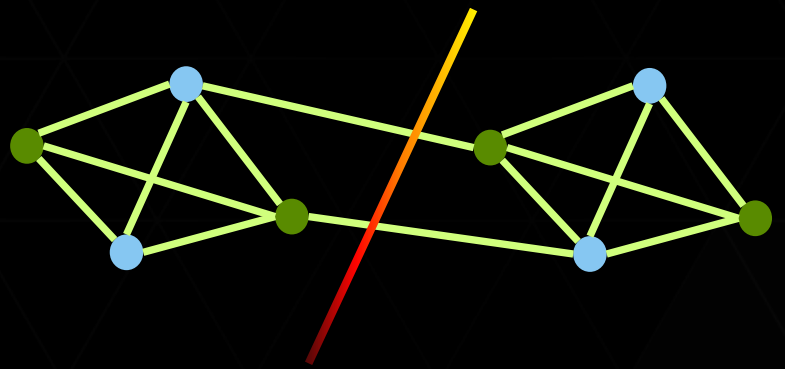


GRAPH ALGORITHMS

- Search
- Cutting / Partitioning
- Shortest Path
- Connected components
- Pagerank

Find the cut with the fewest number of crossing edges.

- Find weak spots in data/communications networks
- Community detection in social networks

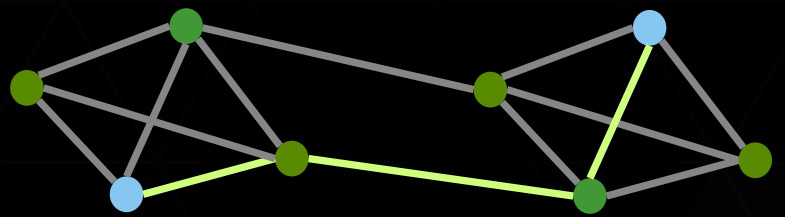


GRAPH ALGORITHMS

- Search
- Cutting / Partitioning
- Shortest Path
- Connected components
- Pagerank

Find the shortest path between two nodes of interest.

- Transportation logistics
- Social network analysis

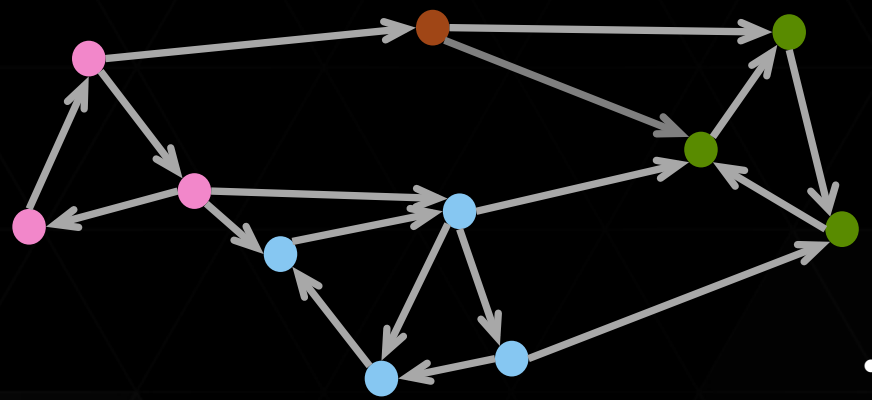


GRAPH ALGORITHMS

- Search
- Cutting / Partitioning
- Shortest Path
- Connected components
- Pagerank

A strongly connected graph is one where you can get to every node in the graph from any starting node.

The Strongly Connected Components (SCC) are the maximal sub-regions of a graph for which each sub-region is strongly connected.

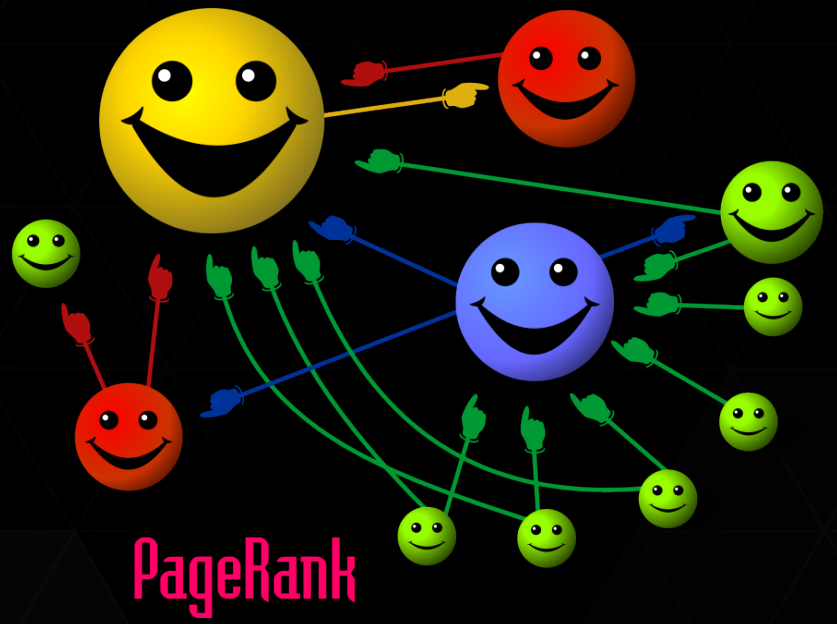


- Social network analysis

GRAPH ALGORITHMS

- Search
- Cutting / Partitioning
- Shortest Path
- Connected components
- Pagerank

Used by internet search for ranking web pages. Invented by Google.
A measure of popularity. Named after Larry Page (not web page).

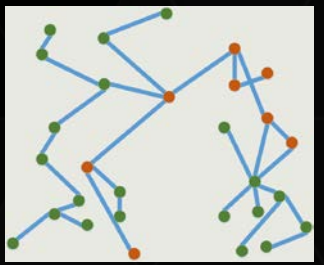


Graph Analytics for GEOINT

GEOSPATIAL GRAPH EXAMPLE

1. Acquire imagery data (GIS, SAR , LiDAR…)
2. Create semantic labels for areas.
   - Use Machine Learning / Deep Learning
   - Building / Trees / Grass / Road / Water
3. Create Graph
   - Node = region identified in step 1, with attributes

     Area, Aspect ratio… Temporal information (last absent, first seen)
   - Edge = distance between regions
4. Perform Analytics
   - Search for patterns of interest using graph node template.
   - “high schools in Anne Arundel County, MD”



First Application of Geospatial Semantic Graphs to SAR Image Data Sandia National Lab 2013

GPUs and Graph Analytics

COMPUTING WITH GRAPHS

What is the nature of graph analytics and traversals?

- Just “visiting” a node. Not much actual mathematical computation.
- Constantly moving around one node at a time, backtracking etc.
- Non-locality is the rule.
- Often have to traverse entire graph (Pagerank, Connected components etc) probably multiple times. Cache thrashing.

Graph analytics stresses communication over computation.

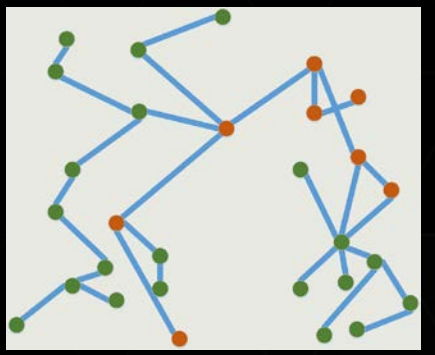
Memory bandwidth more important than Operations/sec.

GRAPH SIZE

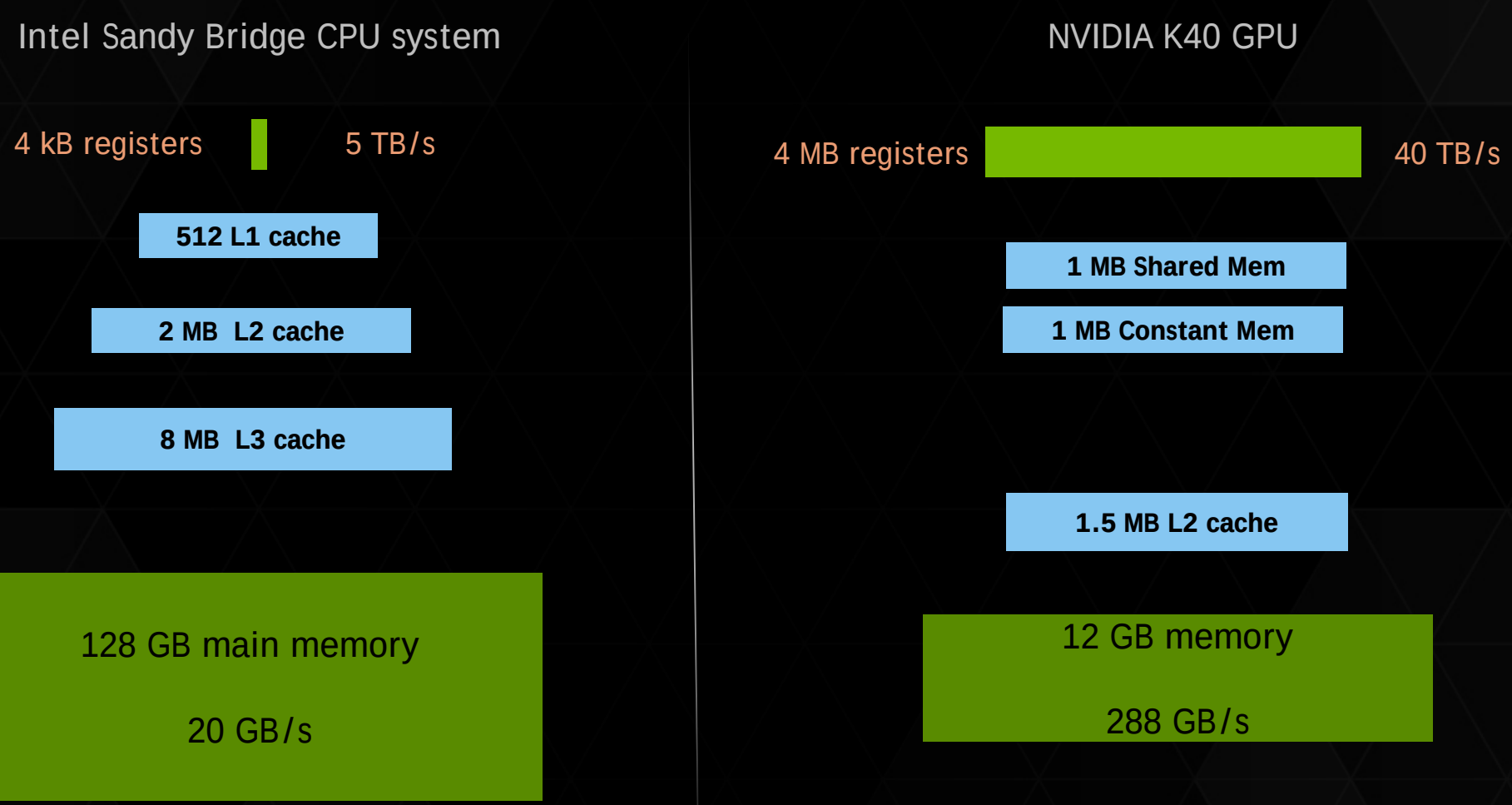
What is the typical size of graphs we are interested in?

- Facebook has ~ 1 trillion edges in their graph.
- Many web and social graphs are in the 10 GB to 100 GB range, in binary format.
- There are certainly "interesting" graphs that are smaller, and larger, than that.

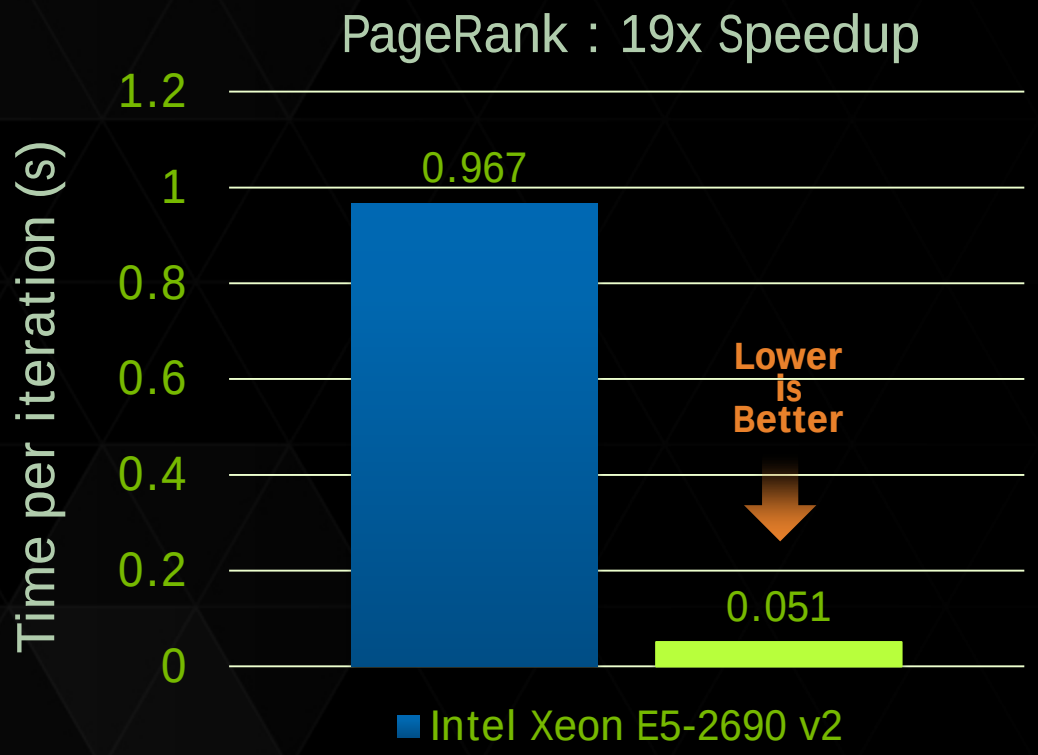
Biological networks



WHERE'S MY MEMORY?

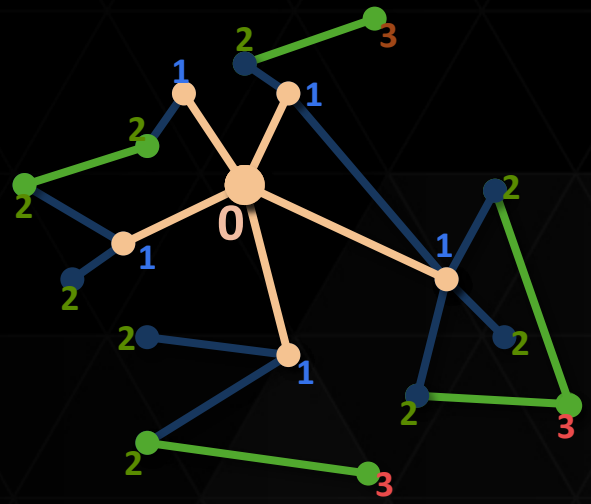


GPU ACCELERATION FOR GRAPH ANALYTICS



1 GPU vs 60 Nodes

280x vs optimized Spark
1440x vs Spark
3420x vs Hadoop



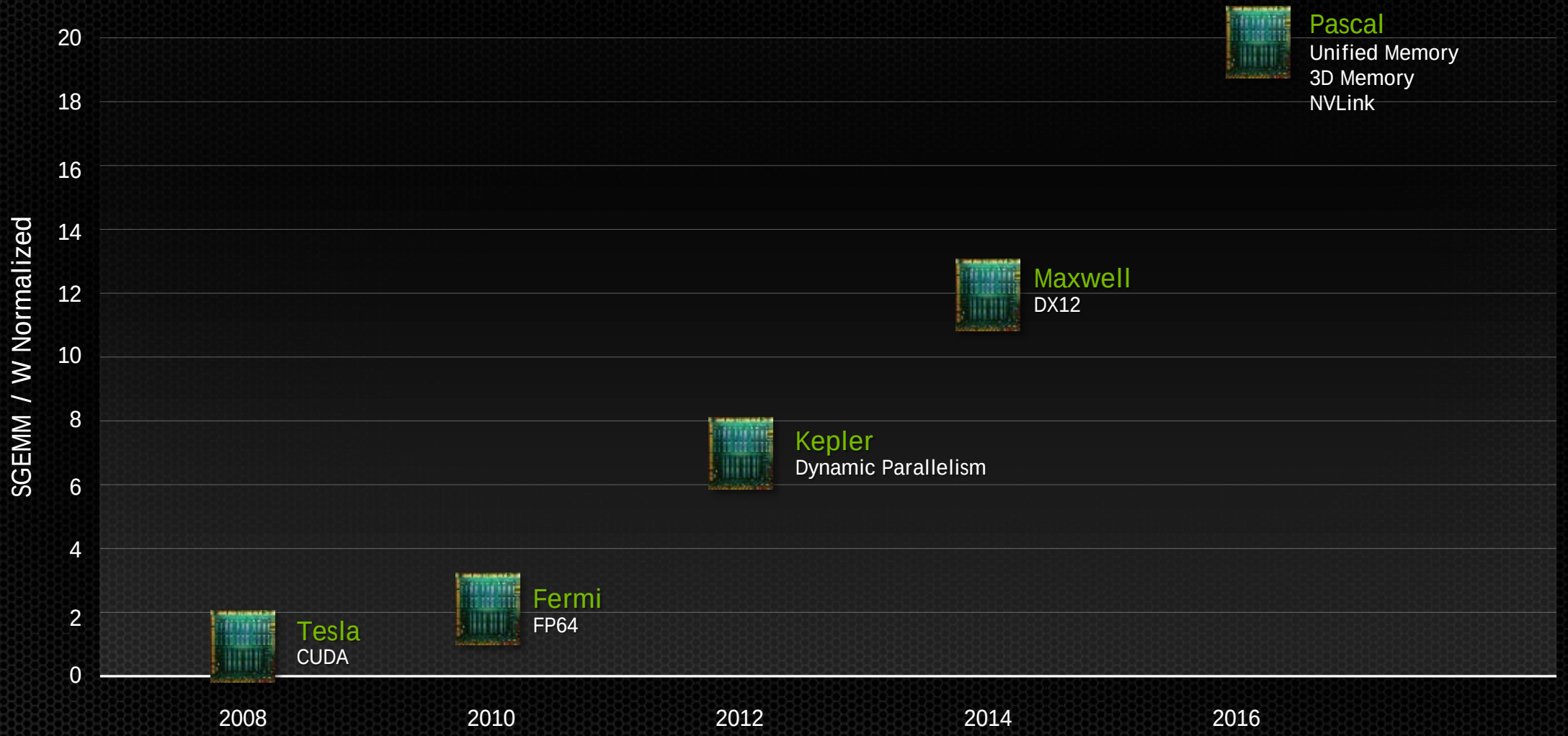
GRAPH TRAVERSAL PERFORMANCE

GTEPS = Giga TEPS
Billions of Traversed Edges Per Second

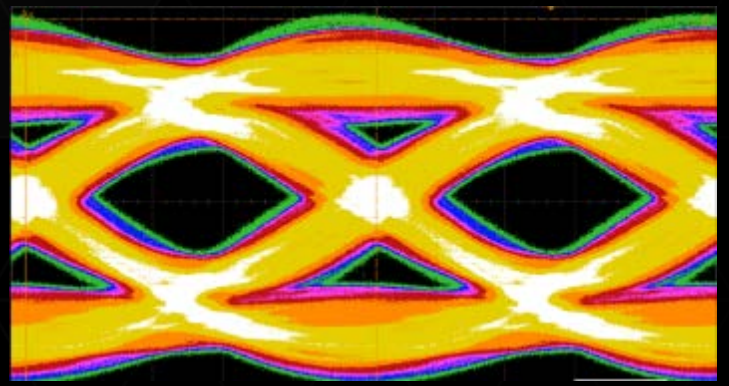
	Configuration	Cost	GTEPS	$/GTEPS
	4-Core CPU	$4,000	0.2	$5,333
	**4-Core CPU + K20 GPU**	**$7,000**	**3.0**	**$2,333**
	XMT-2 (rumored price)	$1,800,000	10.0	$188,000
	**64 GPUs (32 nodes with 2x K20 GPUs per node and InfiniBand DDRx4 – today)**	**$500,000**	**30.0**	**$16,666**
	**16 GPUs (2 nodes with 8x Pascal GPUs per node and InfiniBand DDRx4 – Q1, 2016)**	**$125,000**	**>30.0**	**<$4,166**

GPU Roadmap for Graph Analytics

GPU ROADMAP

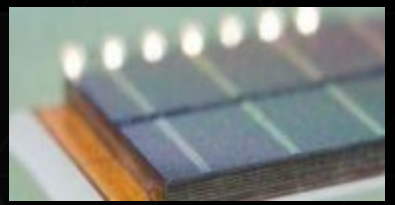

SGEMM / W Normalized
20
18
16
14
12
10
8
6
4
2
0
2008
2010
2012
2014
2016
Tesla
CUDA
Fermi
FP64
Kepler
Dynamic Parallelism
Maxwell
DX12
Pascal
Unified Memory
3D Memory
NVLink

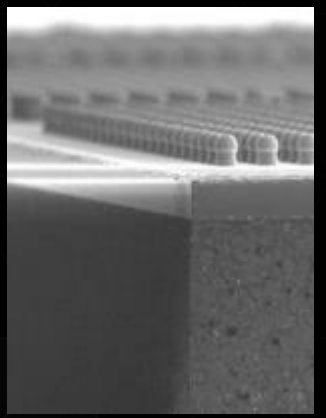
PASCAL GPU FEATURES
NVLINK AND STACKED MEMORY



NVLINK

- GPU high speed interconnect
- 80-200 GB/s

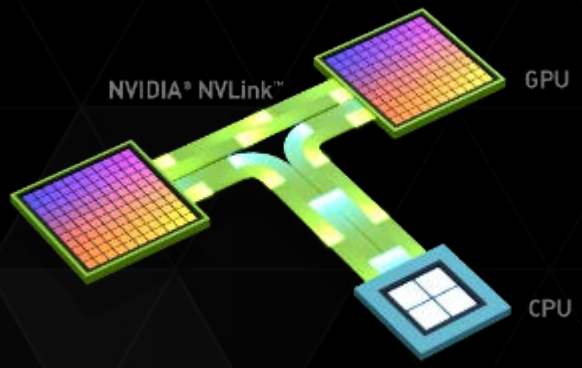


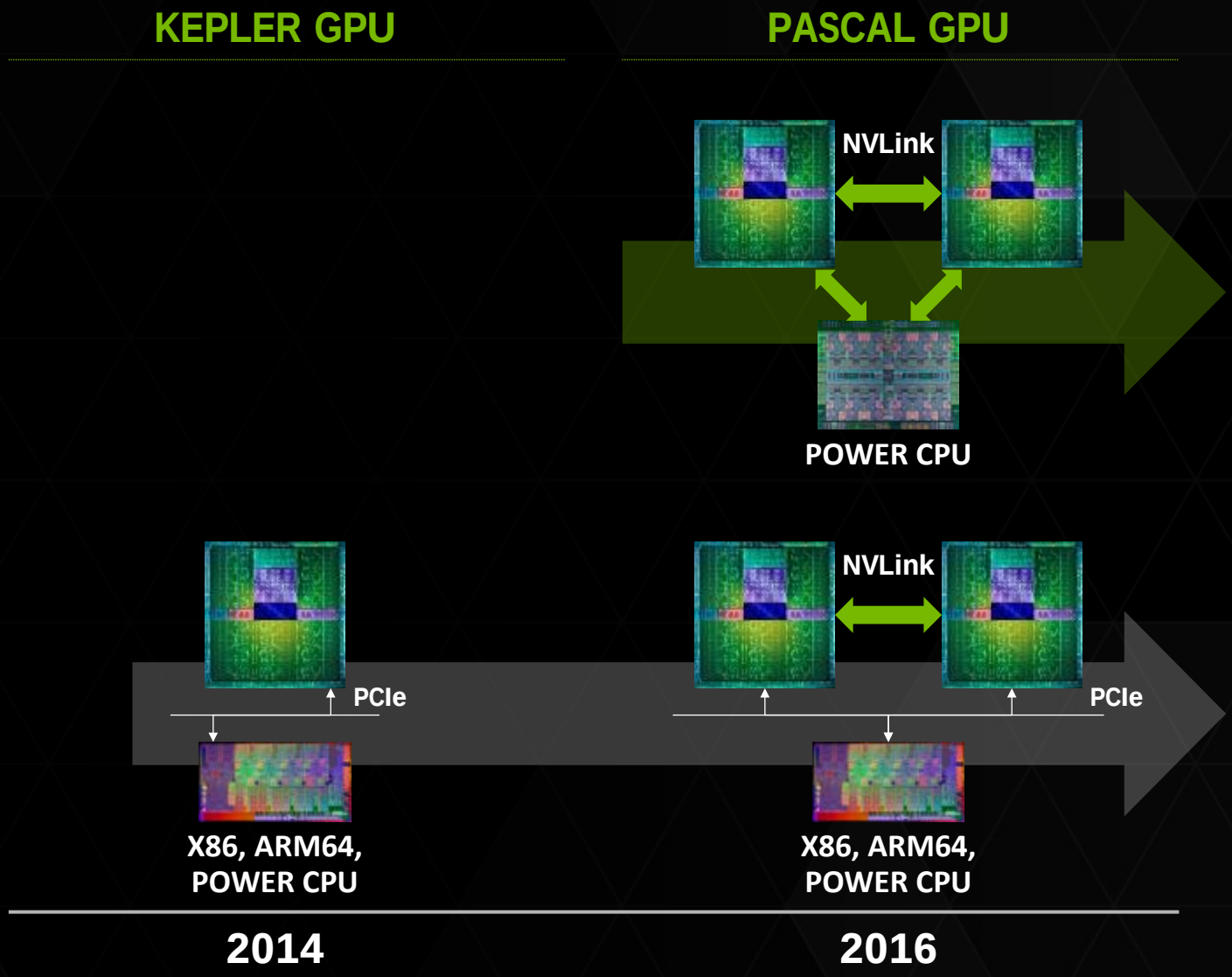


3D Stacked Memory

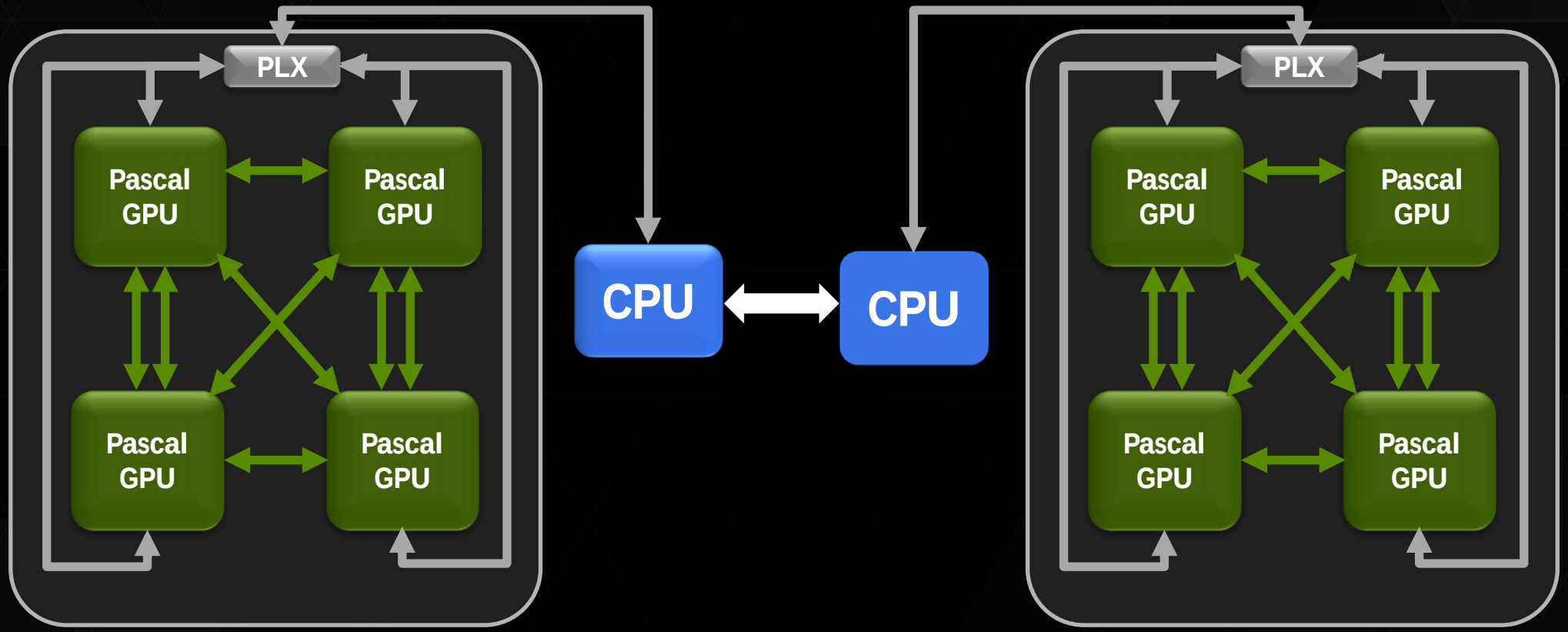
- 4x Higher Bandwidth (~1 TB/s)
- 3x Larger Capacity
- 4x More Energy Efficient per bit

NVLINK
HIGH-SPEED GPU INTERCONNECT





EXAMPLE: 8-GPU SERVER WITH NVLINK

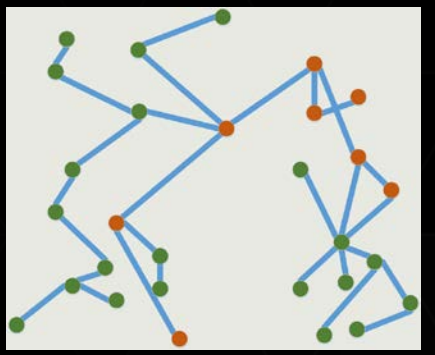


GPU ROADMAP FOR GRAPH ANALYTICS

What are the implications….

- 10X reduction in communication costs
- 2X or greater speedup for BFS with Pascal
- Much larger problem sizes

Biological networks




NVIDIA
Thank you!