

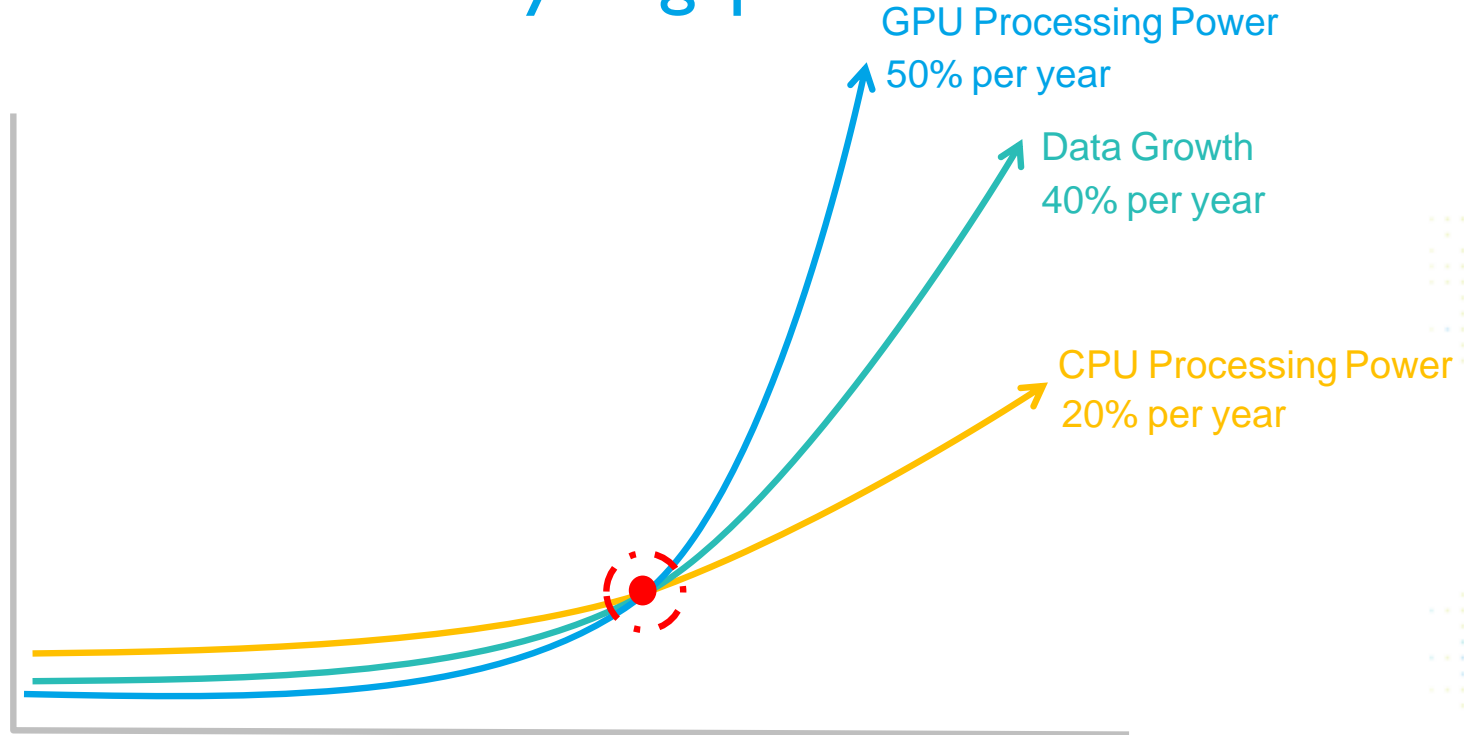


The Role of GPUs in Geovisualization

Todd Mostak, CEO

Oct 12, 2017

GPUs close the analytic gap



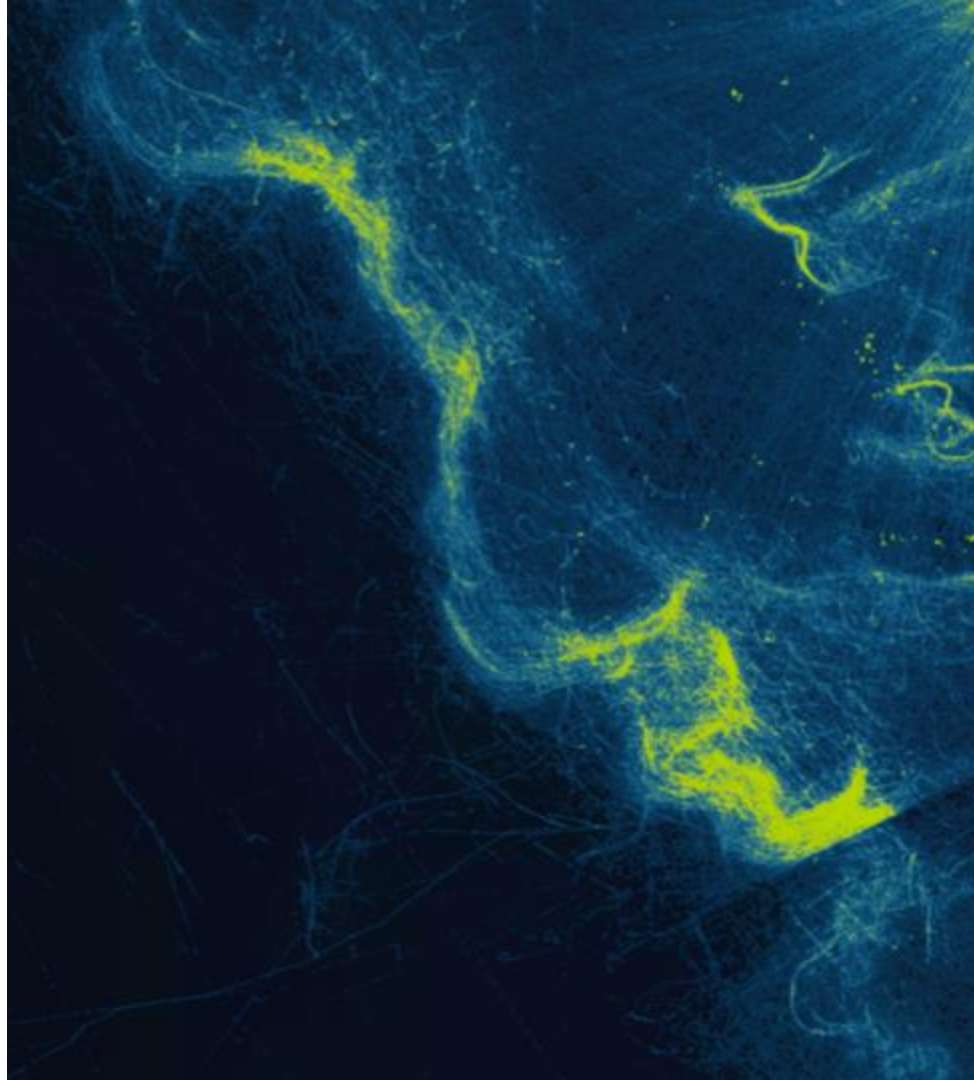
MapD Core

The fastest
analytic database,
architected on GPUs

```
...
(select count(*) as rowsInTable from taxi_weather_tract
returned.
on time: 21 ms, Total time: 22 ms
SELECT
CASE WHEN conditions IN ('Clear', 'Haze', 'Shallow Fog',
WHEN conditions IN ('Light Rain', 'Light Rain Showers
WHEN conditions IN ('Heavy Rain', 'Light Thunderstoms
'Light Freezing Drizzle', 'Ice Pellets', 'Blowin
Thunderstorms and Rain', 'Thunderstorm', 'Squal
ELSE 'Unknown' END AS WeatherConditions,
COUNT(*) as RidesCount,
AVG(trip_distance) as AverageTripDistance,
AVG(total_amount) as AverageTripAmount,
sum(trip_distance) / (sum(extract(epoch from dropoff_da
FROM taxi_weather_tracts_factual
WHERE distance_in_meters(-73.9772,40.7527,pickup_longitude
AND dropoff_datetime > pickup_datetime
ORDER BY WeatherConditions;
Conditions | RidesCount | AverageTripDistance | AverageTrip
recipitation | 98019 | 2.381121 | 11.
recipitation | 27034 | 2.182159 | 11.
recipitation | 3714531 | 3.294560 | 11.
recipitation | 36305282 | 5.461787 | 11.
returned.
on time: 86 ms, Total time: 90 ms
```

MapD Immerse

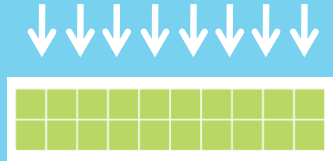
A visualization client,
accelerating interaction
with billion-row datasets



DEMO

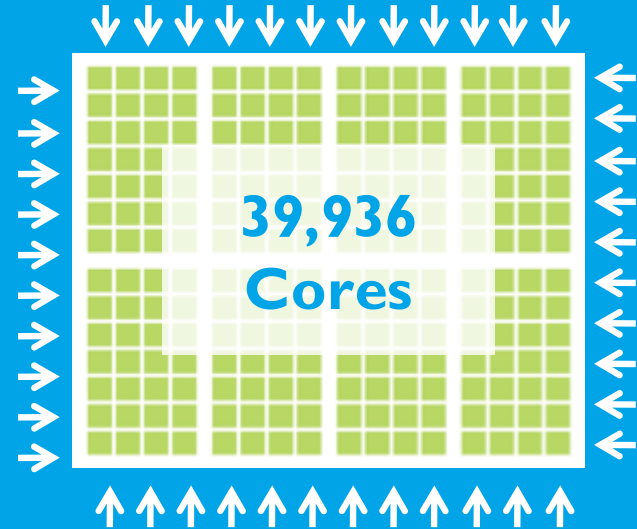
GPU speed: core density and parallel processing

CPU Processing (Traditional)



20 Cores

GPU Processing



MapD queries are orders of magnitude faster

Blogger Mark Litwintchik benchmarked MapD on a billion-row taxi data set and found it to be **up to 12,500x faster than the fastest CPU databases**

Query 1	Query 2	Query 3	Query 4	Setup
0.021	0.053	0.165	0.51	MapD & 8 Nvidia Pascal Titan Xs
0.027	0.083	0.163	0.891	MapD & 8 Nvidia Tesla K80s
0.028	0.2	0.237	0.578	MapD & 4-node g2.8xlarge cluster
0.034	0.061	0.178	0.498	MapD & 2-node p2.8xlarge cluster
0.036	0.131	0.439	0.964	MapD & 4 Nvidia Titan Xs
0.051	0.146	0.047	0.794	kdb+/q & 4 Intel Xeon Phi 7210 CPUs
0.762	2.472	4.131	6.041	BrytlytDB & 2-node p2.16xlarge cluster
1.034	3.058	5.354	12.748	ClickHouse, Intel Core i5 4670K
1.56	1.25	2.25	2.97	Redshift, 6-node ds2.8xlarge cluster
2	2	1	3	BigQuery
4	4	10	21	Presto, 50-node n1-standard-4 cluster
6.41	6.19	6.09	6.63	Amazon Athena
8.1	18.18	n/a	n/a	Elasticsearch (heavily tuned)
10.19	8.134	19.624	85.942	Spark 2.1, 11 x m3.xlarge cluster w/ HDFS
11	10	21	31	Presto, 10-node n1-standard-4 cluster
14.389	32.148	33.448	67.312	Vertica, Intel Core i5 4670K
34.48	63.3	n/a	n/a	Elasticsearch (lightly tuned)
35	39	64	81	Presto, 5-node m3.xlarge cluster w/ HDFS
43	45	27	44	Presto, 50-node m3.xlarge cluster w/ S3
152	175	235	368	PostgreSQL 9.5 & cstore_fdw
264	313	620	961	Spark 1.6, 5-node m3.xlarge cluster w/ S3

Source: <http://tech.marksblog.com/benchmarks.html>

MapD Core: Comparative Query Acceleration*				
System	Query 1	Query 2	Query 3	Query 4
BrytlytDB & 2-node p2.16xlarge cluster	36x	47x	25x	12x
ClickHouse, Intel Core i5 4670K	49x	58x	32x	25x
Redshift, 6-node ds2.8xlarge cluster	74x	24x	14x	6x
BigQuery	95x	38x	6x	6x
Presto, 50-node n1-standard-4 cluster	190x	75x	61x	41x
Amazon Athena	305x	117x	37x	13x
Elasticsearch (heavily tuned)	386x	343x	n/a	n/a
Spark 2.1, 11 x m3.xlarge cluster w/ HDFS	485x	153x	119x	169x
Presto, 10-node n1-standard-4 cluster	524x	189x	127x	61x
Vertica, Intel Core i5 4670K	685x	607x	203x	132x
Elasticsearch (lightly tuned)	1,642x	1,194x	n/a	n/a
Presto, 5-node m3.xlarge cluster w/ HDFS	1,667x	735x	388x	159x
Presto, 50-node m3.xlarge cluster w/ S3	2,048x	849x	164x	86x
PostgreSQL 9.5 & cstore_fdw	7,238x	3,302x	1,424x	722x
Spark 1.6, 5-node m3.xlarge cluster w/ S3	12,571x	5,906x	3,758x	1,884x

*All speed comparisons are to the "MapD & 8 NVIDIA Pascal Titan Xs" benchmark



Common use cases

Powering analytics beyond the limits of CPUs



FINANCIAL SERVICES

Trading model acceleration
Loss prevention
Fraud anomaly detection



TELECOMMUNICATIONS

Network optimization
Know your customer
NOC/SOC monitoring



FEDERAL

Geospatial intelligence (GEOINT)
Cyber security
Disaster command center



ENERGY

Well log analytics
Smart meter analytics
Field service management



TRANSPORTATION

Real-time fleet monitoring
Incentive-based insurance
Smart cities



ADTECH

Segmentation analytics
Ad pricing and placement
Behavioral insights



About MapD

Investors

\$25 million Series B in
March 2017



Customers, Team and Community

- May 2017 – MapD Core open sourced
- April 2017 – MapD 3.0 with distributed core and HA
- Founding member of GPU Open Analytics Initiative (GOAI)
- Available on AWS, Azure and GCP public cloud platforms





Q & A