

## How to process all of this data?

### Dealing with the "Compute Gap"

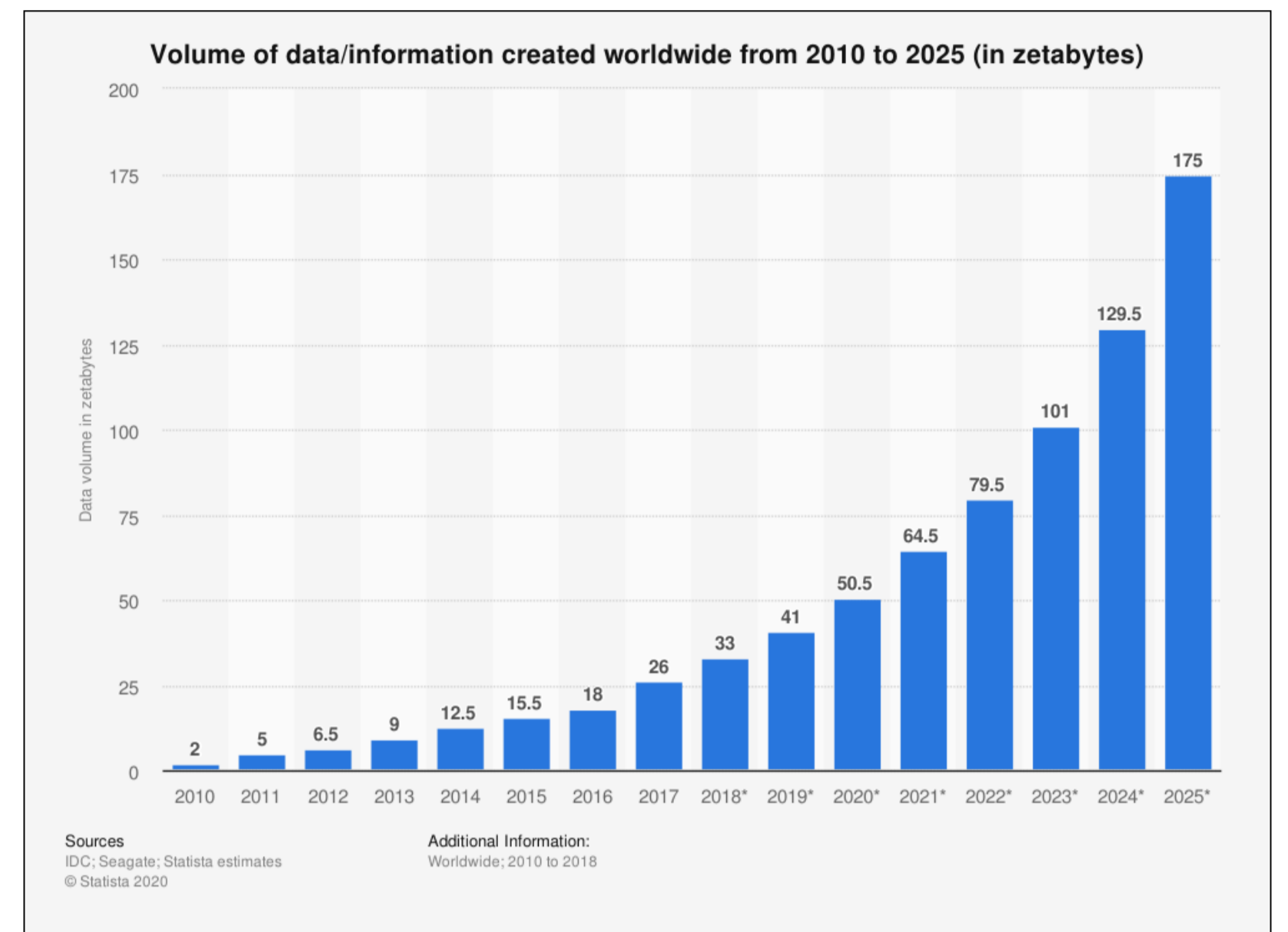
As you can see from the graph to right there is an increasing amount of data created every year. But with Moore's Law being dead to many how is the industry supposed to keep up with the data growth? Let's take a more in-depth look at database management systems (DBMS).

### How to make DBMS faster?

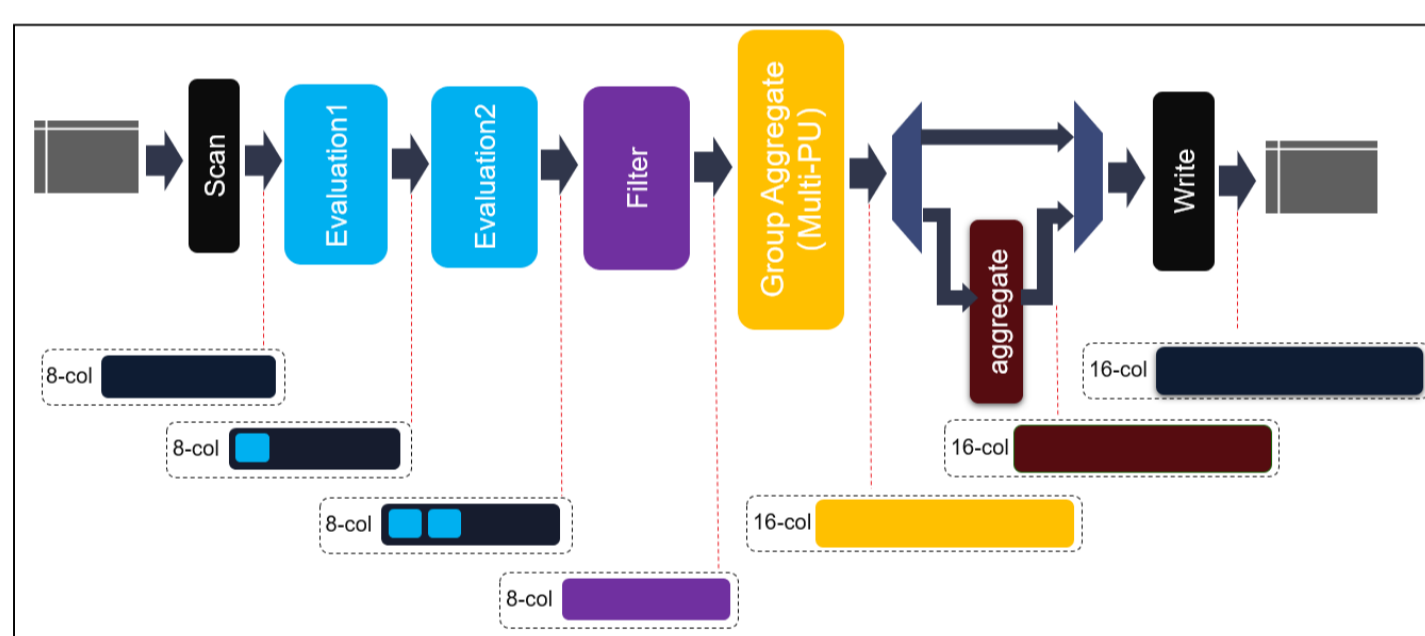
- More CPU Cores – Expensive, Takes power
- GPUs – Worse than in-memory DBs with SIMD vector instructions, No fast DRAM access
- FPGAs – High complexity, Difficult to sell specialised HW packaged to DBMSs
- ASICs – High up-front cost, No flexibility

### Potential with FPGAs

FPGAs have the same problems as GPUs do, but because of the extra flexibility they can be made more valuable and possibly overcome the usefulness of simply adding more CPU cores. The problems mentioned for FPGAs are to do with the inherent complexity of heterogenous systems and using specialised HW. The biggest example is the failed Netezza product which failed to attract enough customers to subscribe to a highly specialised and non flexible service.

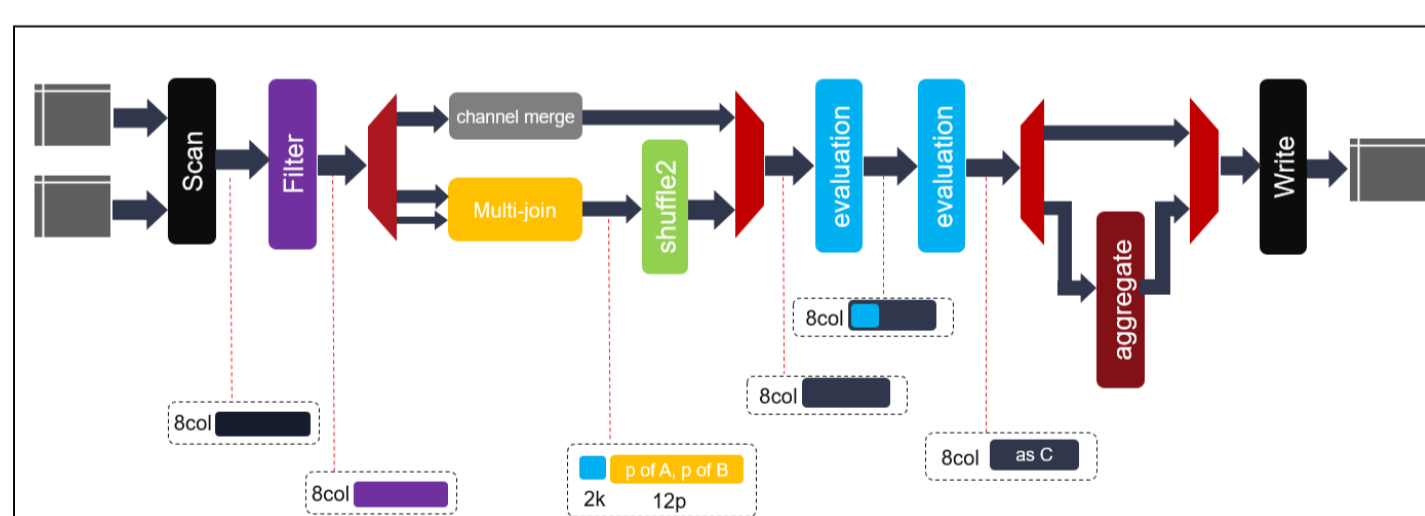


## Making DBMS acceleration easy!



Let the already existing query optimisers do their work!

Need to offer additional options. Optimisers use cost functions for different operations and their parameters. These can be expanded by an extension which then allows additional FPGA resources to be used.



Because of the added latency of moving data online analytical processing is the target application rather than online transaction processing

OLAP

~~OLTP~~

### What is out there?

- Operator accelerators – Different levels as below
- Kernels and frameworks – Overlays, working systems
- Dynamic kernels – Systems with partial reconfiguration
- Transparent DBMS extensions which use FPGAs

### How much? - Drawbacks?

- A lot – Needs a system around to be used. Difficult
- Few – Not flexible. Need to know the workload
- Very few – Scalability problems, research area
- None – How to make?

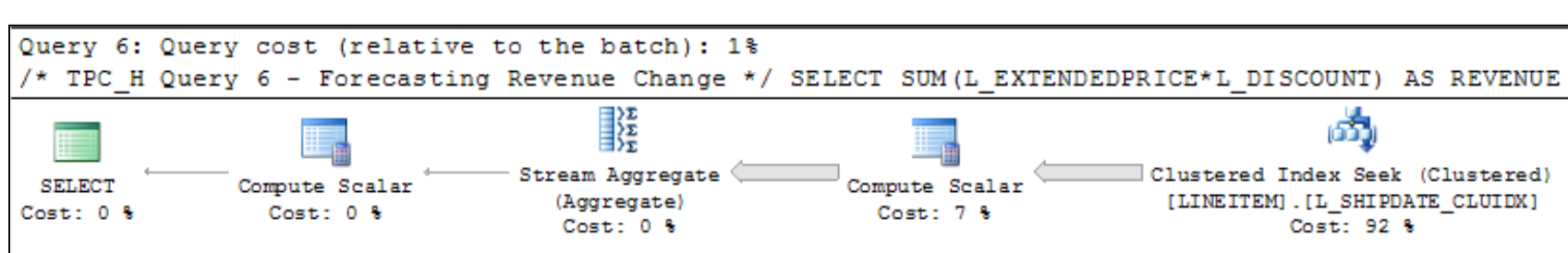
| String matching | Join     | Stochastic gradient descent |
|-----------------|----------|-----------------------------|
| Sorting         | Select   | Skyline                     |
| Filtering       | Where    | K-means                     |
| Arithmetic      | Group by | Decision tree ensembles     |

## How would this be done?

### Improve benchmark times! - TPC-H

#### TPC\_H Query 6 - Forecasting Revenue Change

```
SELECT SUM(L_EXTENDEDPRICE*L_DISCOUNT) AS REVENUE
FROM LINEITEM
WHERE L_SHIPDATE >= '1994-01-01' AND L_SHIPDATE < dateadd(yy, 1, cast('1994-01-01' as date))
AND L_DISCOUNT BETWEEN .06 - 0.01 AND .06 + 0.01 AND L_QUANTITY < 24
```



### Off to the races!

Read about the acceleration options at different abstraction levels & learn about FPGAs and security: FPGA '20 Invited Tutorial: FPGA Hardware Security for Datacenters and Beyond

