Application of Tesla C2050 to Flow Cytometry and Cancer Detection

By Bob Zigon

Beckman Coulter, Inc

Overview

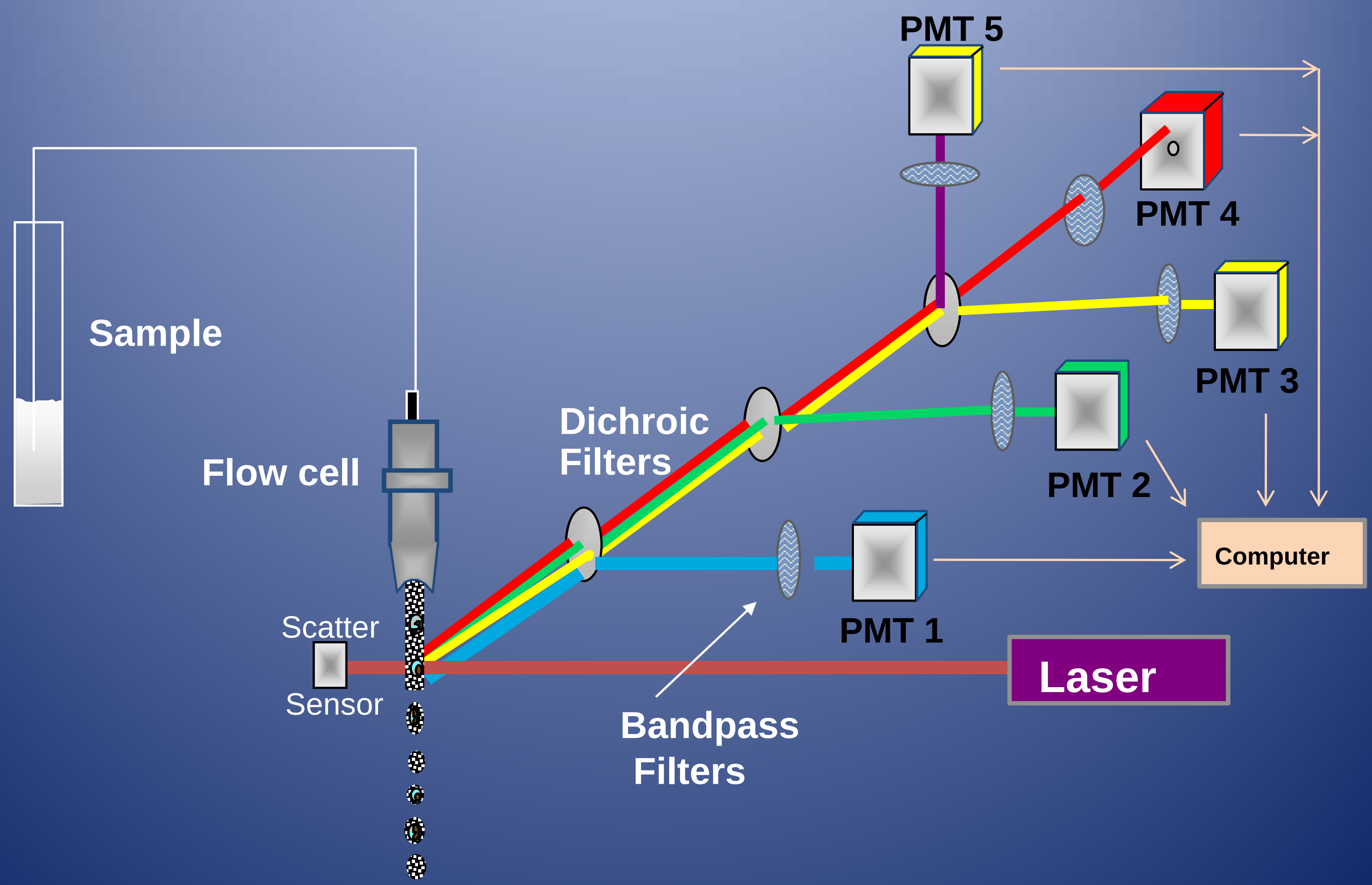
- What is Flow Cytometry?
- What is Kaluza?
- What did Tesla do for us?
- Why is acceleration important?
- Anatomy of the pipeline
- Gigaflop requirements
- Floating point challenges

What is Flow Cytometry?

The use of lasers to interrogate cells delivered by a hydrodynamically focused fluidics system.

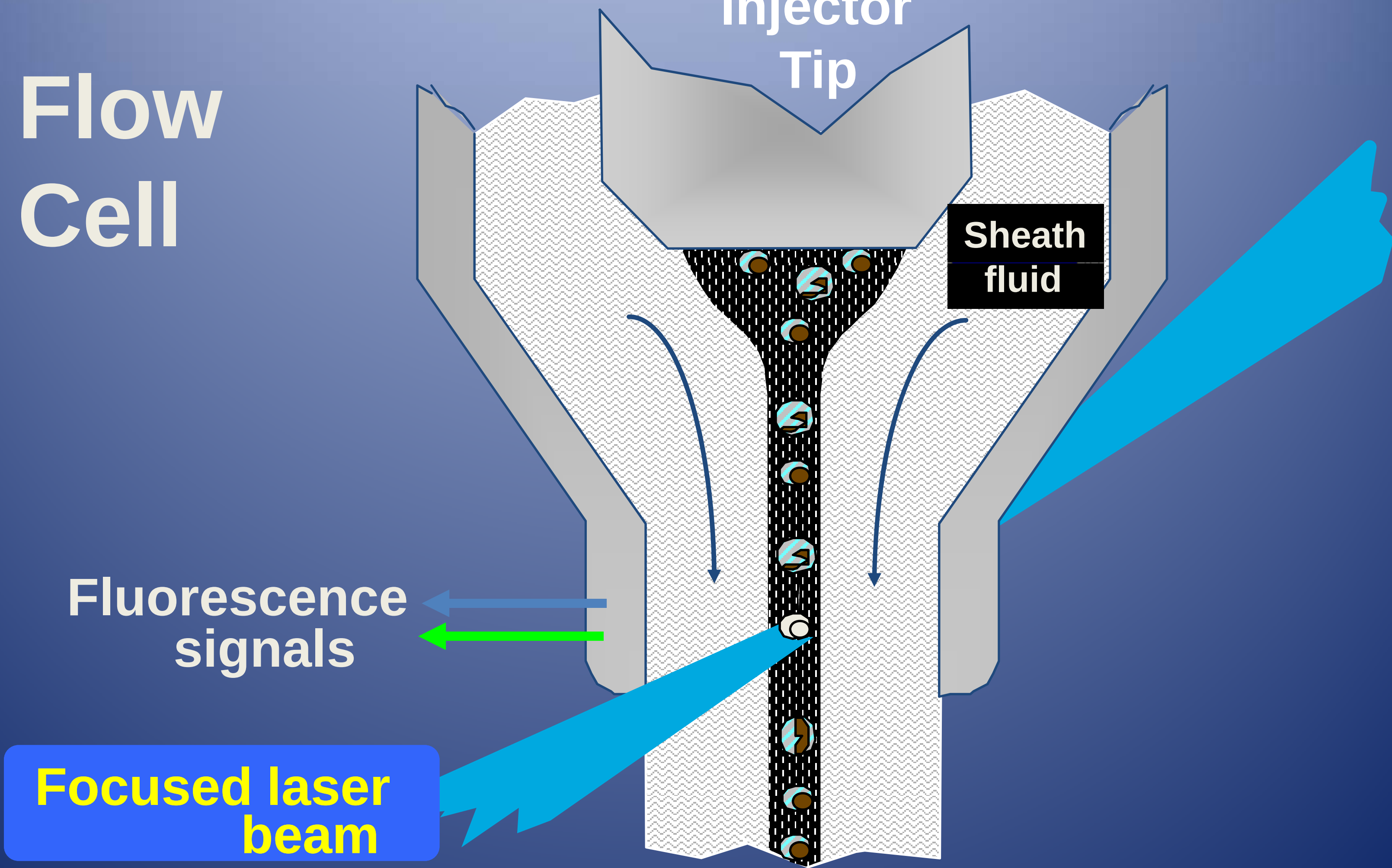
We can measure cell size, internal complexity, and the absence/presence of proteins.

Optical Design



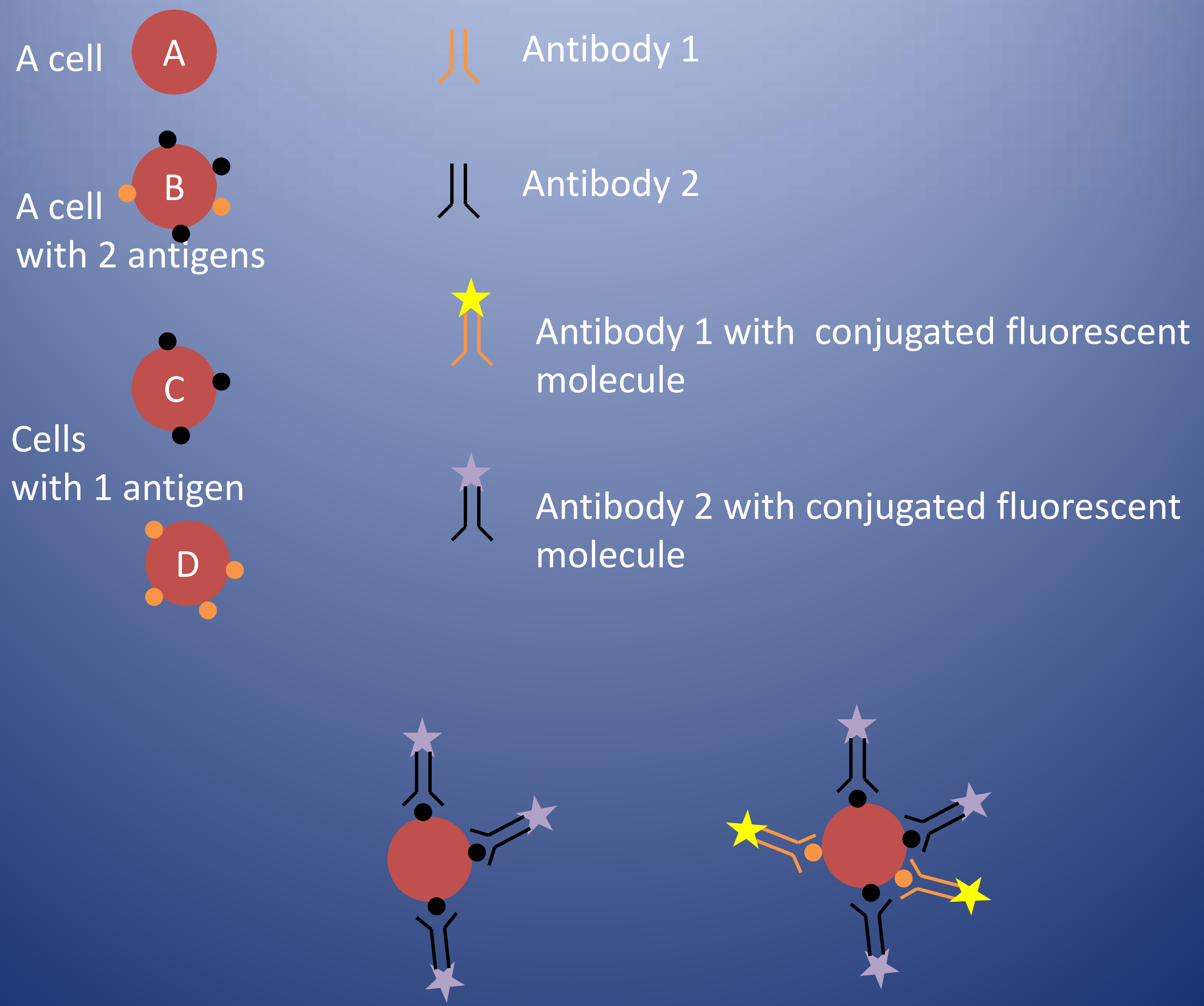
Slide from J. Paul Robinson, Purdue

Hydrodynamic Focusing



Slide from J. Paul Robinson, Purdue

How do we detect the surface proteins?



What does the data look like?

	Measurement 1	Measurement 2	Measurement 3	.	.	.	.	Measurement 10
Cell 1	55	117	280	317	.	.		
Cell 2	0	0	110	9	.			
Cell 3	541	.	.					
.	.							
.								
.								
.								
Cell 10,000,000								

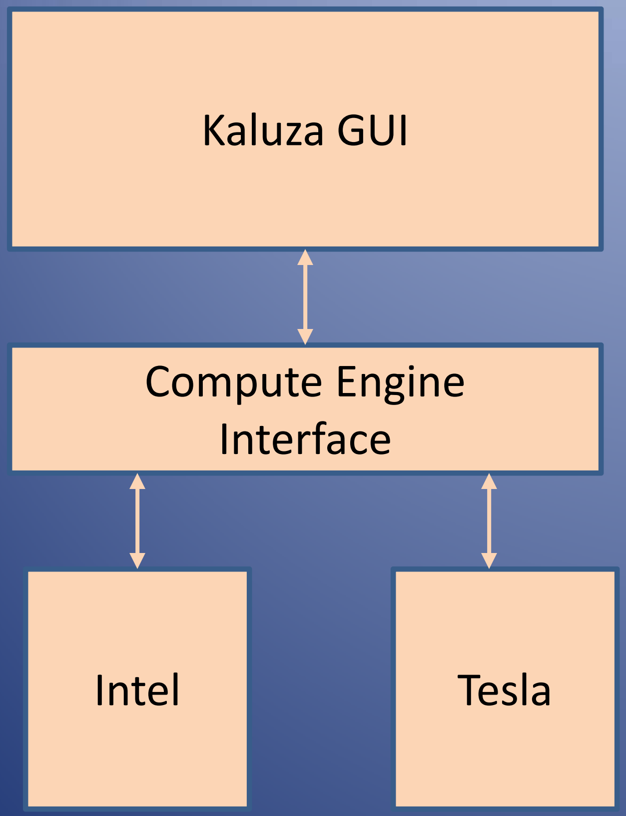
A matrix filled with 100,000,000 cellular measurements consuming 400 megabytes of RAM.

What is Kaluza?

- A flow cytometry analysis application that allows scientists to identify cellular subpopulations and the related disease state.

Brief Demo

What did Tesla do for us?



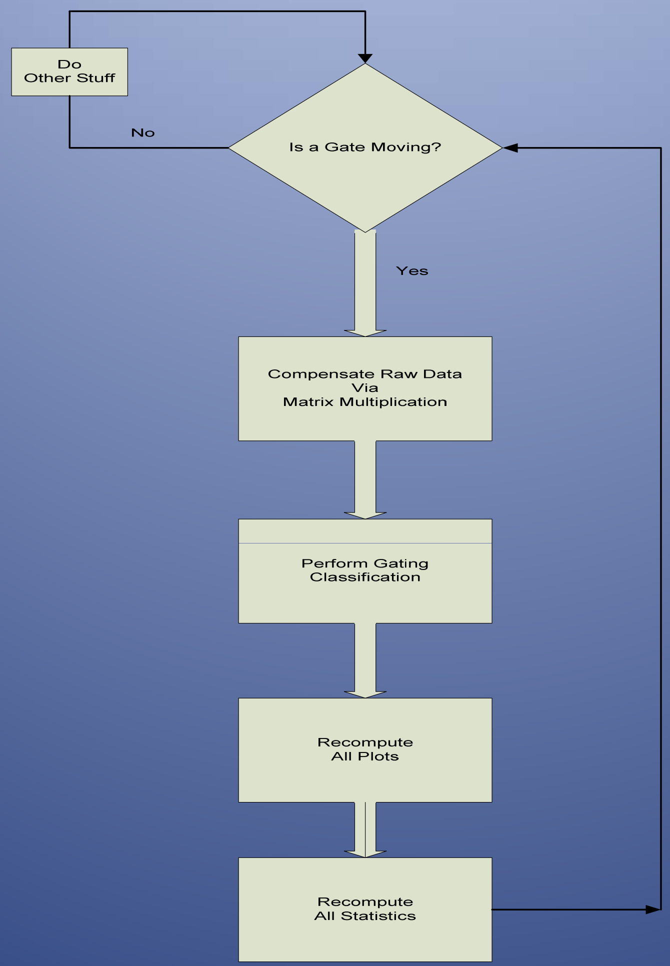
- Tesla provided us with a 400 fold acceleration over our competitors when manipulating monster data sets!

Why is acceleration important?

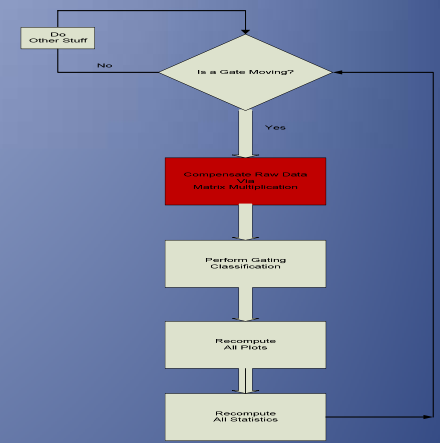
- Interactivity aids in the discovery of subtle patterns
- Magnifying glass vs Electron microscope

- Application : Minimal Residual Disease & Remission

Anatomy of the pipeline



Compensation of raw data



$$D' = D * C$$

Where …

D = Raw Measurement Data (R rows & C columns)

C = Compensation matrix (C rows & C columns)

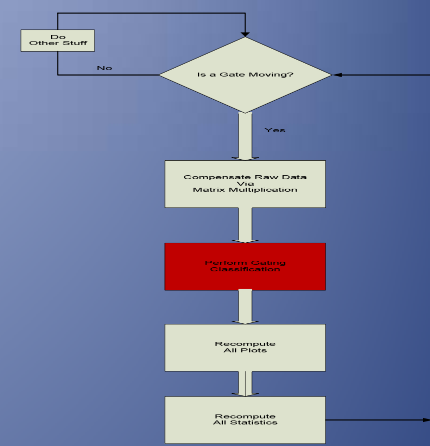
D' = Compensated Measurement Data (R rows & C columns)

Computational complexity O(R*C*C)

There may be 10 colors of light we are searching with …

There may be 10,000,000 cells we are examining …

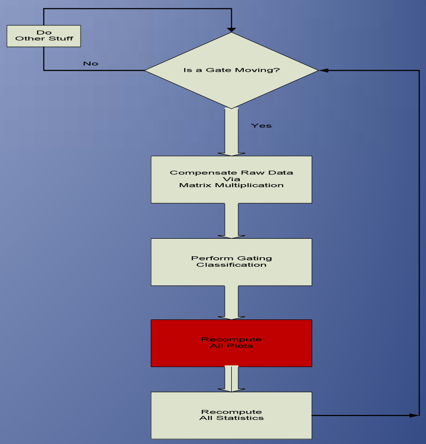
Gating classification




```
For each gate G[i] that changed
{
        For each event E[j]
        {
          Perform Linear or Log transform on E[j]

          Classify E[j] with respect to G[i]
        }
}
```

There may be 20 to 40 gates defined …
There may be 10,000,000 cells we are examining …

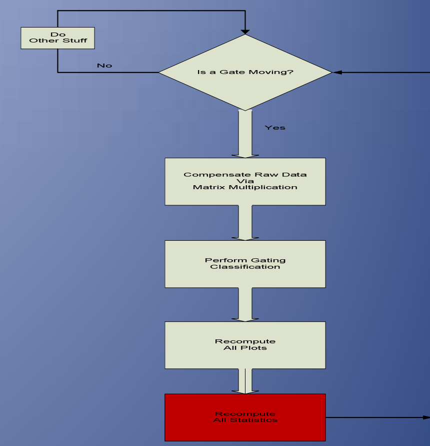
Recompute all plots




```
For each plot P[i]
{
        For each event E[j]
        {
          Perform Linear or Log transform on E[j]

          If E[j] is in the Gate on P[i] then
                Plot E[j] on P[i]
        }
}
```

There may be 30 to 50 plots defined …
There may be 10,000,000 cells we are examining …

Recompute all statistics



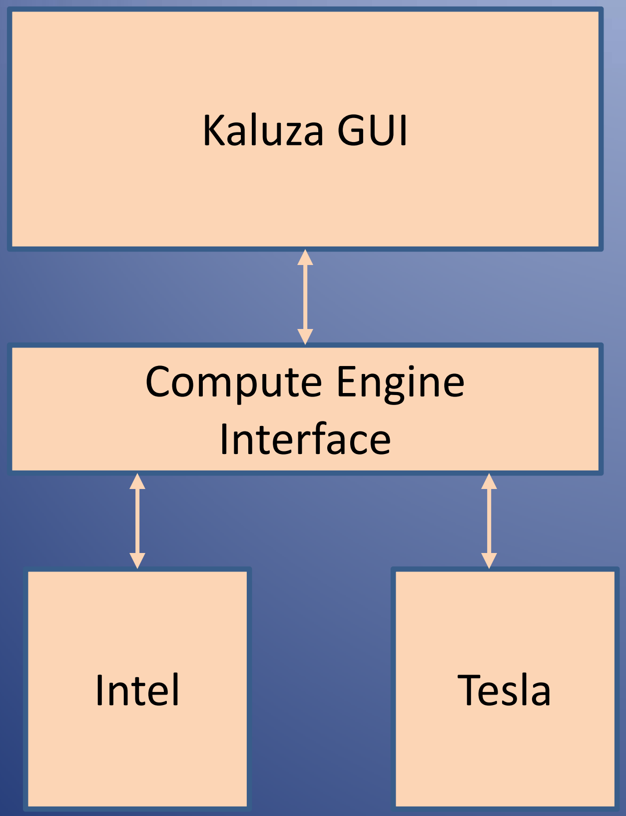

```
For each statistic S[i]
{
        For each event E[j]
        {
          Perform Linear or Log transform on E[j]

          Compute statistic S[i] on E[j]
        }
}
```

There may be 30 to 50 plots defined …
There may be 10,000,000 cells we are examining …

Gigaflop requirements for 15 frames per second

Logarithmic Transform (30 flops)
10 million cells
8 fluorescent measurements
28 bivariate plots 8 histogram plots
65 gigaflops without matrix comp
215 gigaflops with matrix comp

Floating point challenges



- Intel is IEEE 754 compliant
- Tesla C1060 is not.
- Tesla C2050 is.
- How do you generate results that are identical at the binary level?

Thank you

robert.zigon@beckman.com