Scientific Visualization for the SRCs

Arpan Das

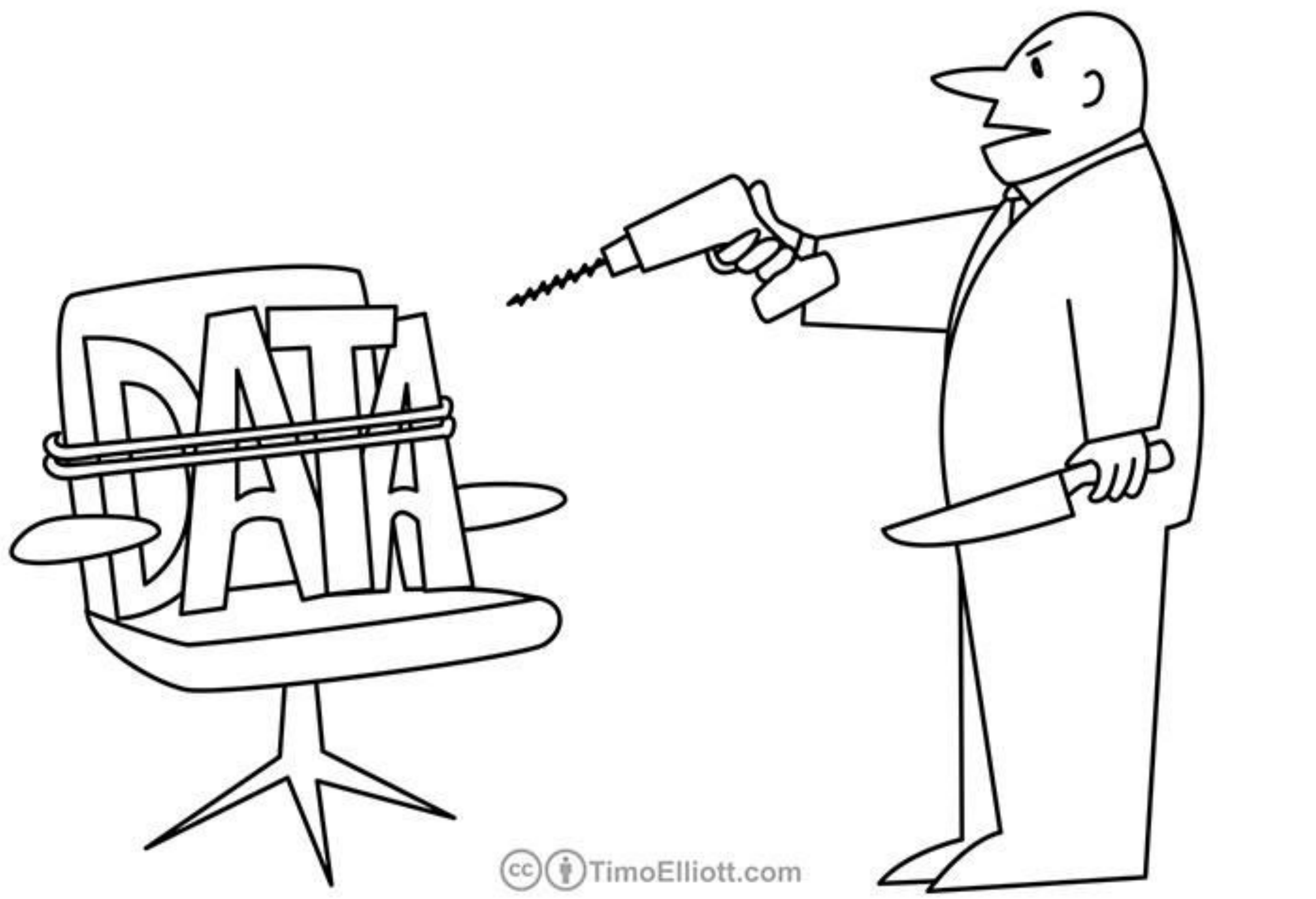
Spring SKACH Consortium, Zurich, 24th May

SKACH

SKAO

EPFL

INFORMATIQUE SCIENTIFIQUE & SUPPORT APPLICATIF
SCITAS
SCIENTIFIC IT AND APPLICATION SUPPORT



"If you don't reveal some insights soon, I'm going to be forced to slice, dice, and drill!"

Data Scientist for Astronomy @ EPFL

PhD in computational Astronomy

(Thesis: Formation of Supermassive Black Holes in the Early Universe)

40% @ Orange SRC team

Visualization of SKA data with high volume of users and high amount of data.

Objectives for 2022 :

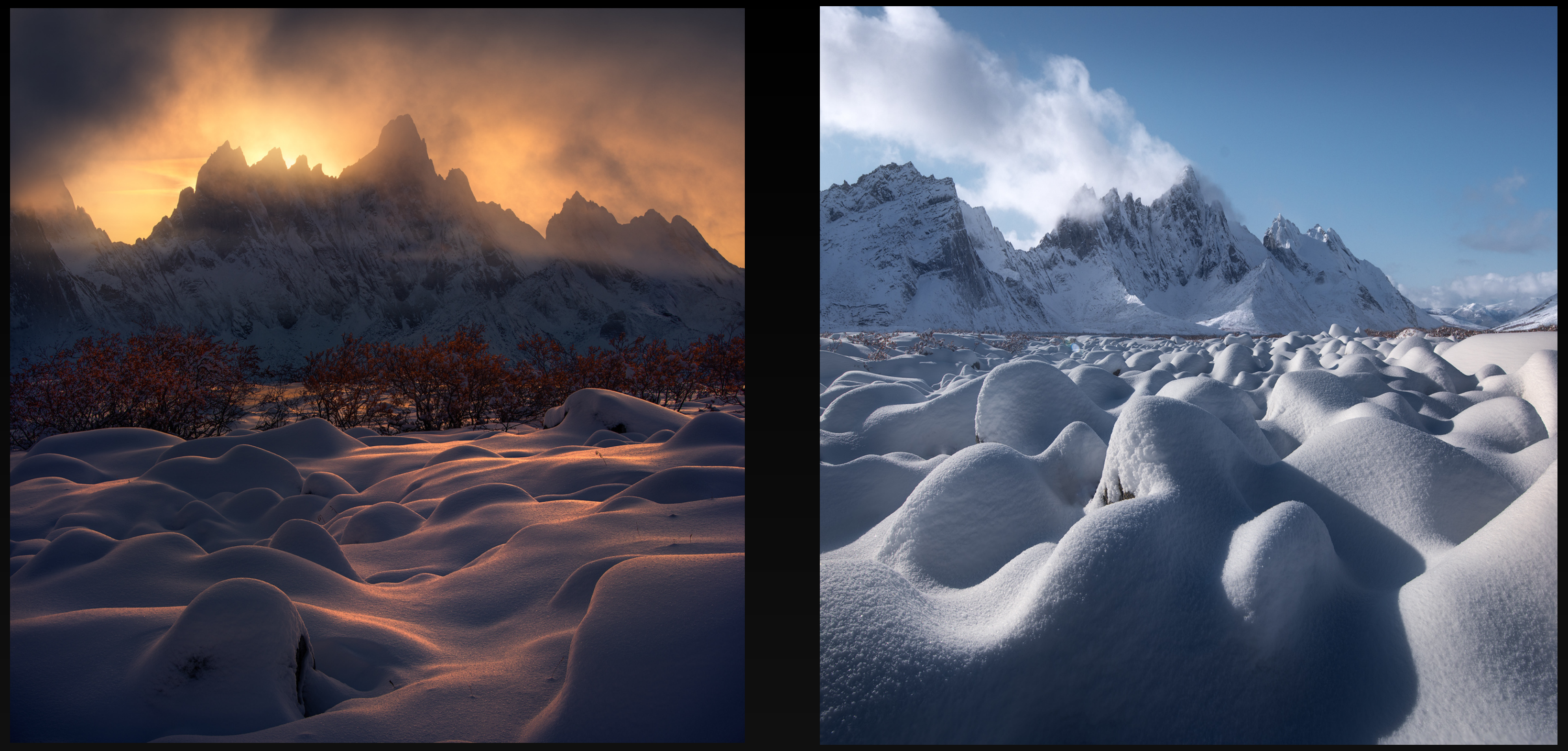
- Identify deliveries and roadmap
- Create backlog
- Identify visualization tools that could be used for SKA data
- Test and compare for SKA data size and scalability
- Test on the federated data network and adapt
- Put code in the common repository
- Create deployable entities

40% @ Tangerine SRC team

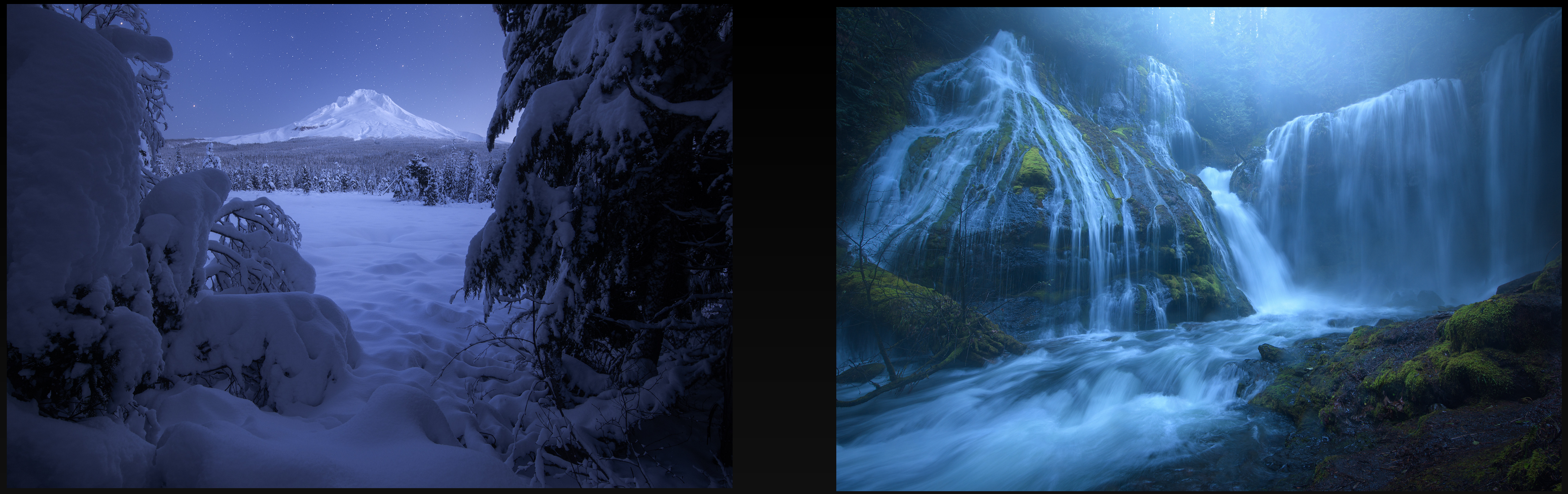
Data processing Notebooks and Distribution of software, tools and services

20% @ EPFL/SCITAS

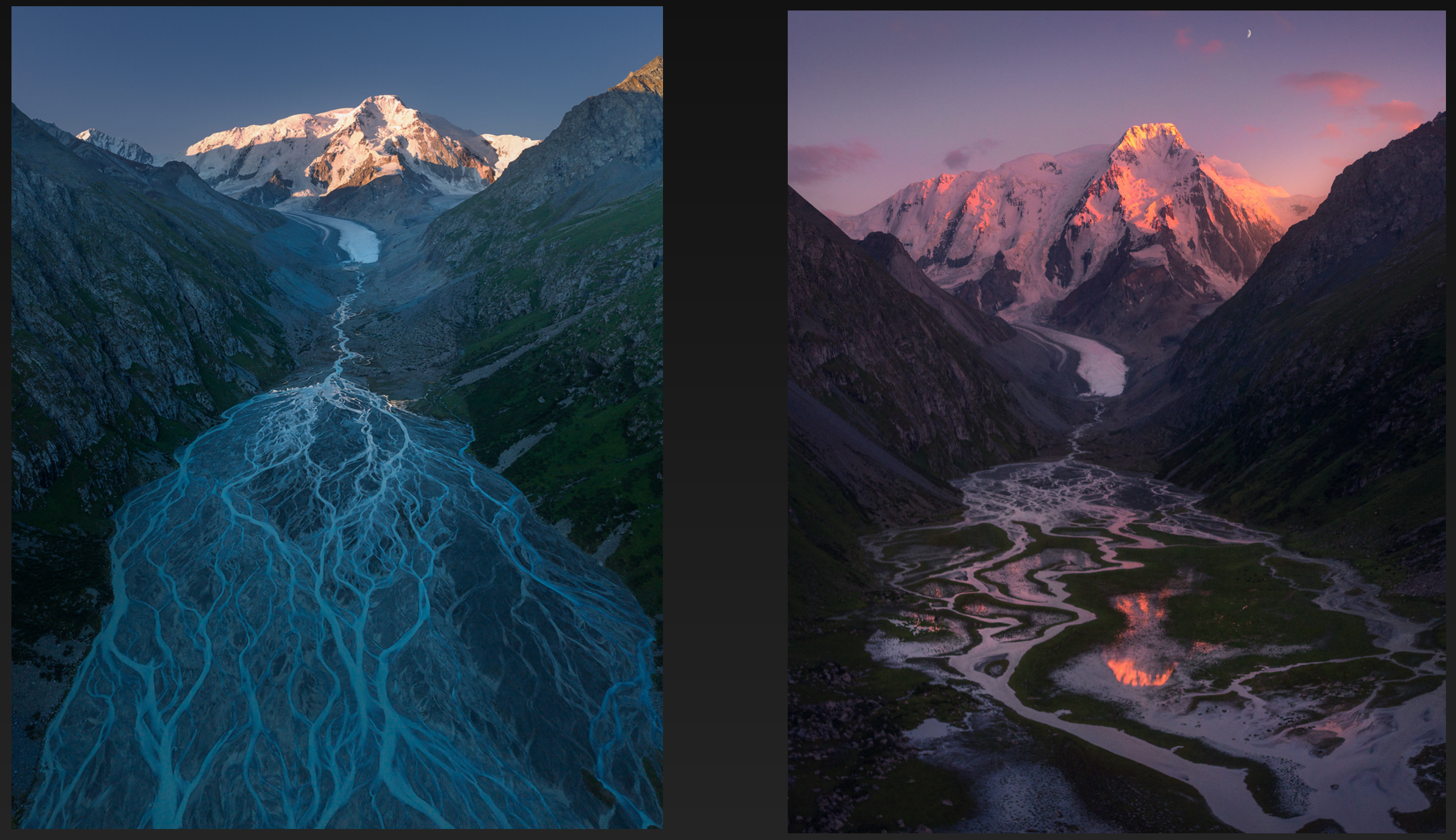
Replace the MeasurementSet and processing infrastructure of CASA with GPUs
(Nvidia Rapid libraries)



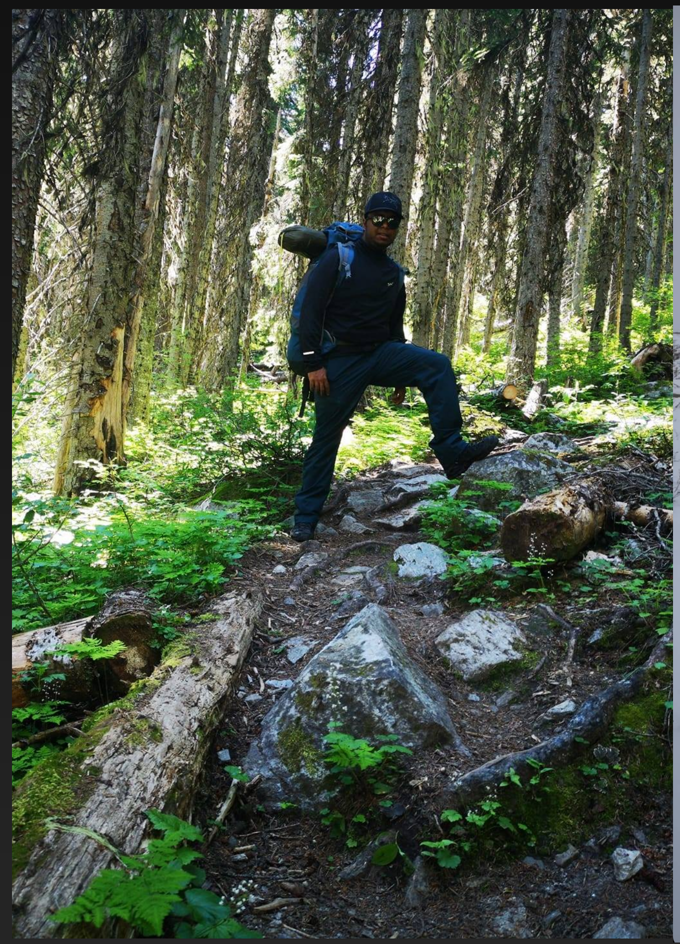
Yukon Territory, Canada



Pacific Northwest, USA

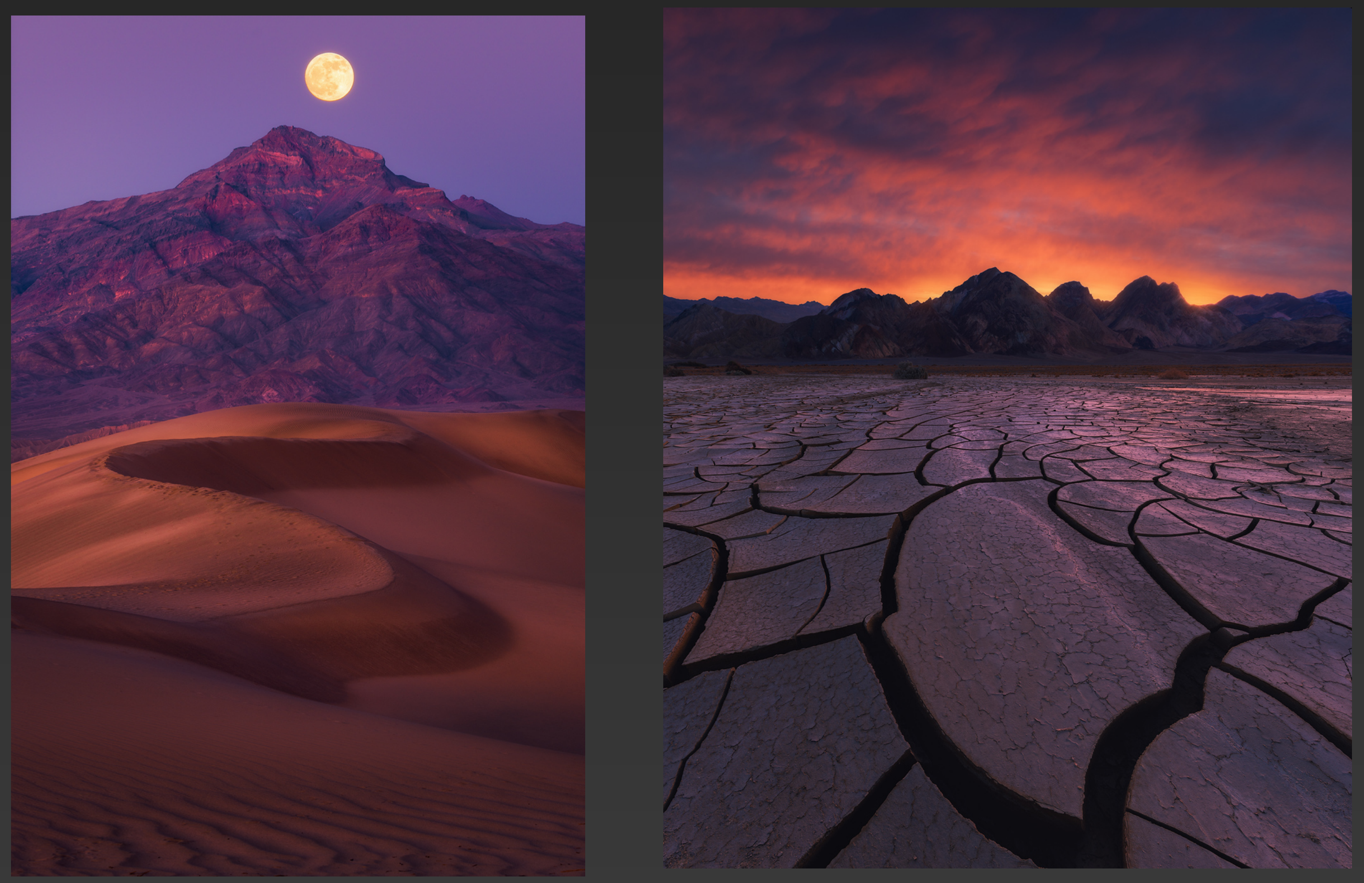


Kyrgyzstan



British Columbia, Canada

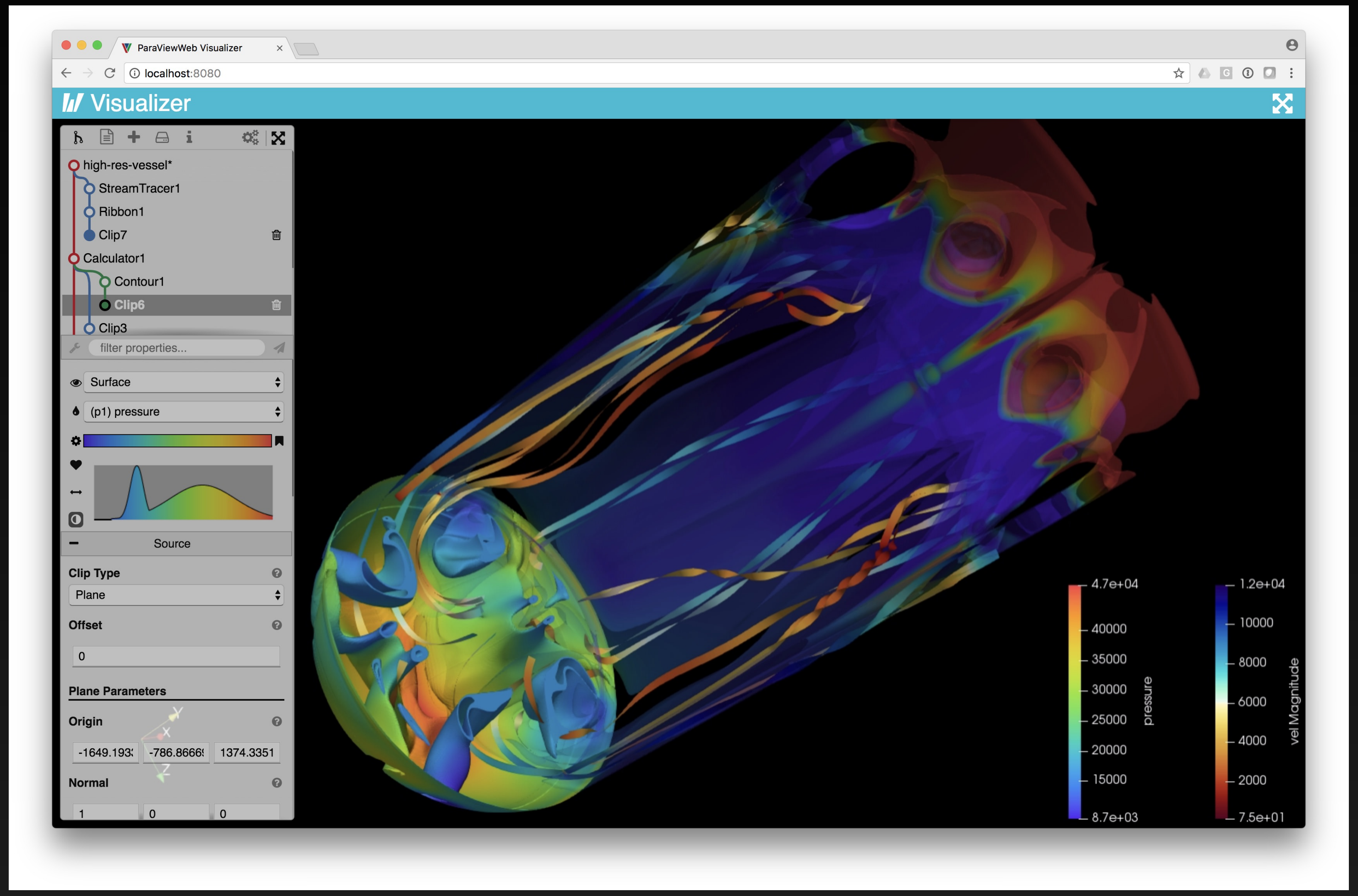
www.arpandas.com

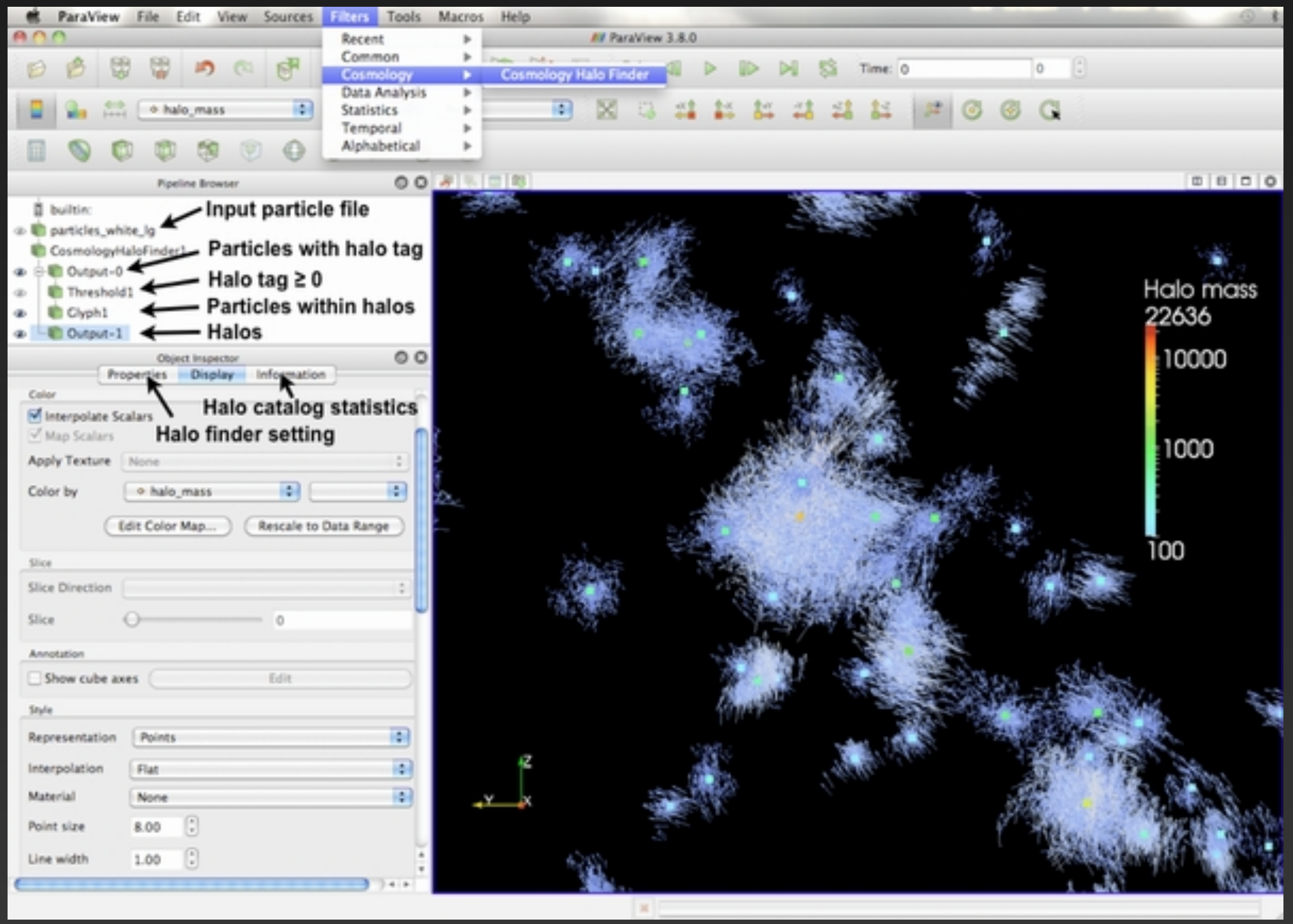


Death Valley NP, USA

ParaView / High-Performance Post-Processing

- Open-source, multi-platform, data analysis and visualization application
- Analysis of extremely large datasets using distributed memory computing resources
- Over 20 years of development
- Contributions from over 270 developers
- Over 1.6 million lines of code
- Over 150k yearly downloads
- www.paraview.org





Input particle file

Particles with halo tag

Halo tag ≥ 0

Particles within halos

Halos

Halo catalog statistics

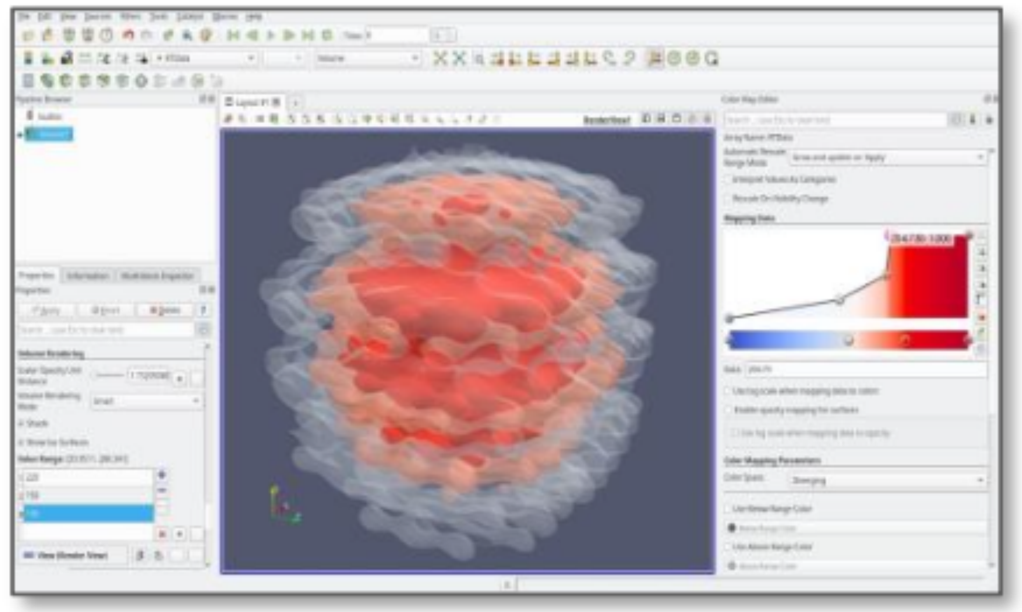
Halo finder setting

ParaView Server

```
10 total, 2 running, 9 stuck, 199 sleeping, 901 threads    23:30:03
40, 1.75, 1.00  CPU usage: 4.15% user, 4.40% sys, 91.44% idle
 1648K resident, 0B data, 0B linkedit.
ns: 31278 total, 1892M resident, 117M private, 564M shared.
ysMem: 5893M used (1191M wired), 10G unused.
 523G vsize, 1026M framework vsize, 0(0) swapins, 0(0) swapouts.
tworks: packets: 12105/8925K in, 11907/1964K out.
isks: 80156/2205M read, 21235/425M written.

PID  COMMAND      %CPU TIME     #TH  #WQ  #PORT MEM    PURG   CMPR PGRP PPID
592  screencaptur 0.0  00:00.02 7    5    55+   1952K+ 20K+   0B   262  262
590  mdworker     0.0  00:00.01 3    0    44    2032K  0B     0B   590  1
589  mdworker     0.0  00:00.01 3    0    44    1572K  0B     0B   589  1
588  top          1.7  00:00.51 1/1  0    22+   2860K  0B     0B   588  584
584  bash         0.0  00:00.00 1    0    15    588K   0B     0B   584  583
583  login        0.0  00:00.01 3    1    28    1228K  0B     0B   583  482
574  auditd       0.0  00:00.00 2    0    25    560K   0B     0B   574  1
567  System Prefe 0.0  00:03.23 3    0    270   39M    8364K  0B   567  1
561  systemstatsd 0.0  00:00.01 2    1    19    1040K  0B     0B   561  1
560  com.apple.We 0.0  00:01.42 9    0    229   25M    0B     0B   560  1
558  com.apple.We 0.0  00:05.07 15   3    224   151M   1716K  0B   558  1
555  bash         0.0  00:00.00 1    0    15    604K   0B     0B   555  554
554  login        0.0  00:00.01 3    1    28    1176K  0B     0B   554  482
550  bash         0.0  00:00.00 1    0    15    608K   0B     0B   550  549
```

Batch mode (ParaView Python)

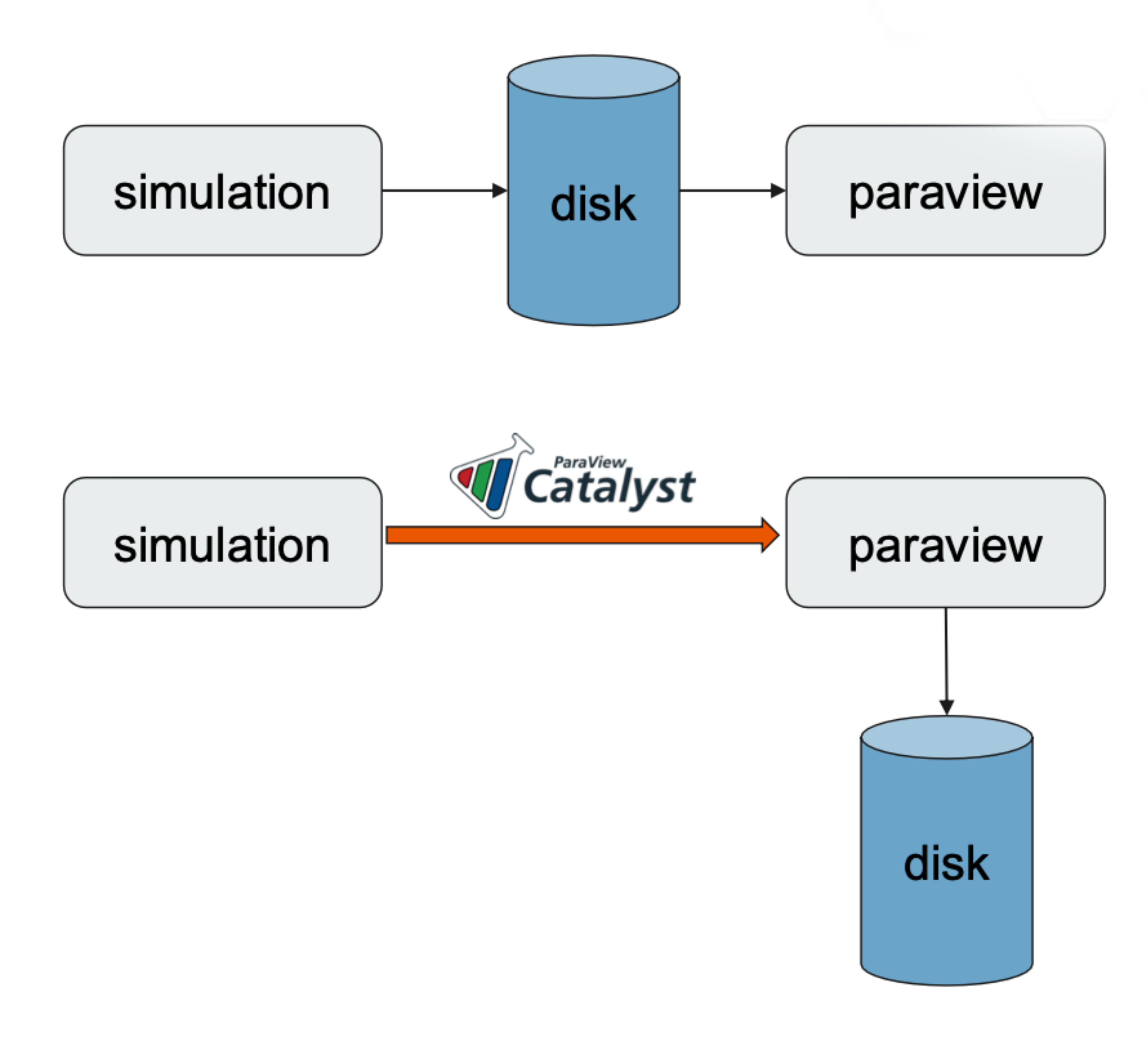


ParaView Client

ParaView Web

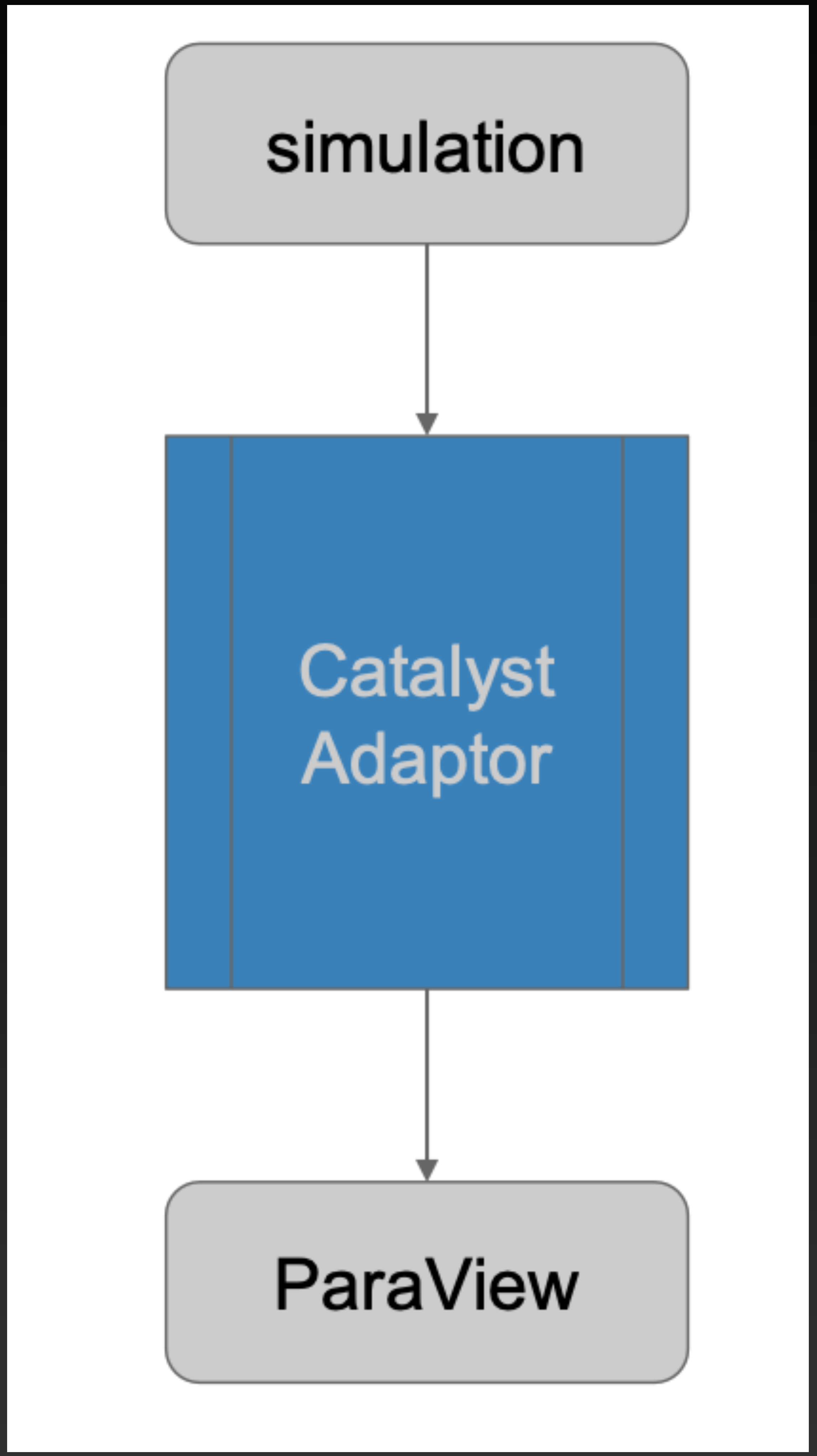
An in-situ framework embedded with ParaView

- in-situ analysis: uses ParaView analysis capabilities
- live visualization: uses ParaView to connect to the simulation
- easy configuration: uses ParaView to generate scripts
- open source



Catalyst 2 :

- Make it easy to develop
  - No need to understand VTK
  - No data conversion
- Make it easy to deploy
  - Several catalyst implementations deployment
  - Not link to specific ParaView version
- Make it easy to build
  - Few dependencies using catalyst-stub
  - No CMake
- Make it easy to maintain and upgrade
  - Stable C API/ABI
  - No need to rebuild simulation for new ParaView version



- Avoid need to understand VTK data model
- Provide mechanism to provide data with zero-copy & meta-data to interpret it

Other visualization packages into consideration

- IT SRC: VisIVO framework

- ZA SRC: Carta

- FR SRC: Aladin, Aladin Lite, HiPS, MOC, TMOC