

OPEN POSSIBILITIES.

Why You're Not Running More System Experiments
(And What You Can Do About It)



OCP
GLOBAL
SUMMIT

NOVEMBER 9-10, 2021

Why You're Not Running More System Experiments (And What You Can Do About It)

Thomas Ameling
Software Architect, Keysight

Dan Mihailescu
System Architect, Keysight

Alex Bortok
Solutions Architect, Keysight

Razvan Stan
Software Development Leader, Keysight

OPEN POSSIBILITIES.



OPEN
COMMUNITY®



Agenda

- Why system experiments?
- What stands in the way?
- Demo
- Building a foundation with open models



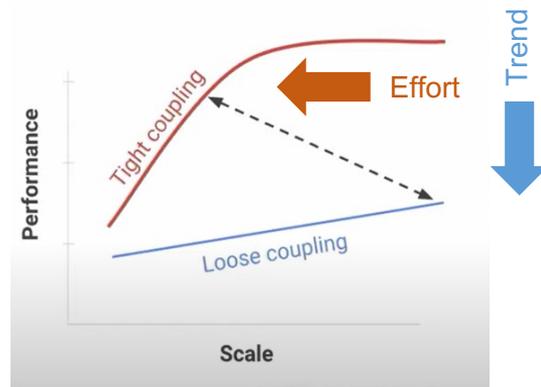
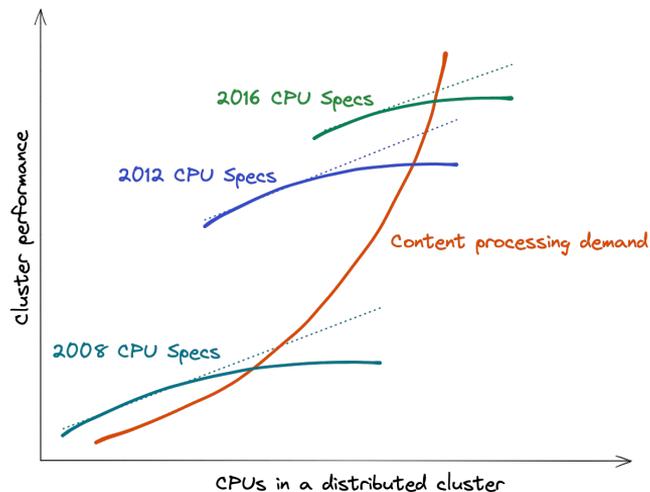
NETWORKING

OPEN POSSIBILITIES.



Why system experiments?

- Scaling-out with generic compute can't keep up with demand
- Cannot ignore 2x-10x gain from application-tailored design
- Need meaningful benchmarks and a lot of measurements



NETWORKING

OPEN POSSIBILITIES.

Amin Vahdat - SIGCOMM Lifetime Achievement Award 2020 Keynote



Example: ML Job – Model Training



NETWORKING

Job Completion Time Factors

- Data Ingestion
- Computation
- **Collective Communications**

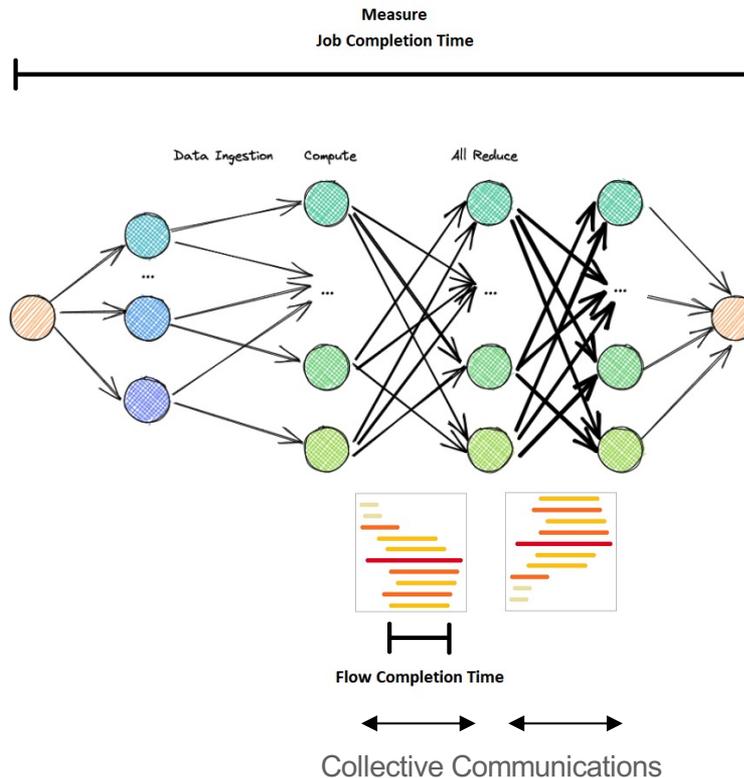
Network matters!

- **Median** vs **Tail** latency

Contributors

- TCP Stack | Comm Library
- DPU (NIC)
- Data Center Fabric
- ...

OPEN POSSIBILITIES.



Why are you not experimenting enough?

Re-running a real job

- **Complicated** tooling
- **Expensive** clusters
- **Uncontrollable** variations

Need **tools** for experimentation

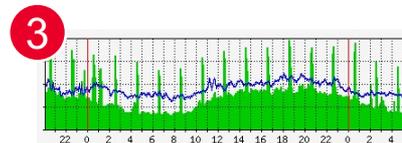
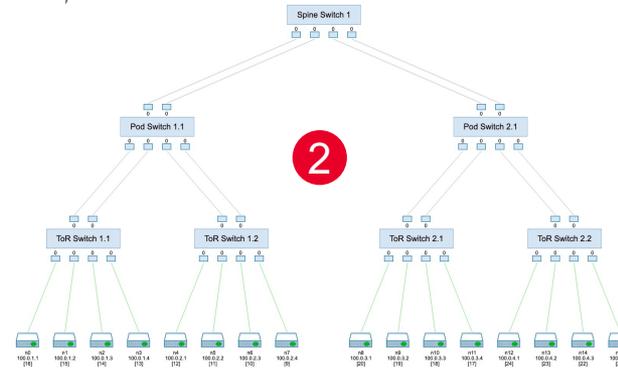
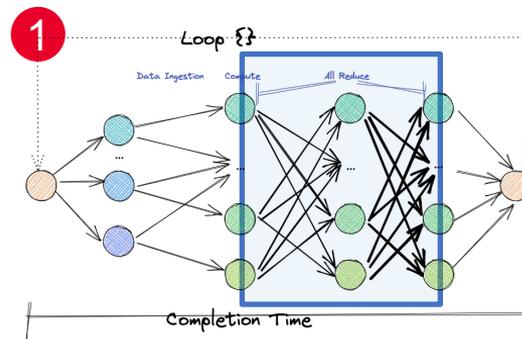
1. Reproduce job communications
2. Emulate fabric on OCP switch
3. Introduce controllable chaos

And **standardized API**

OPEN POSSIBILITIES.



NETWORKING



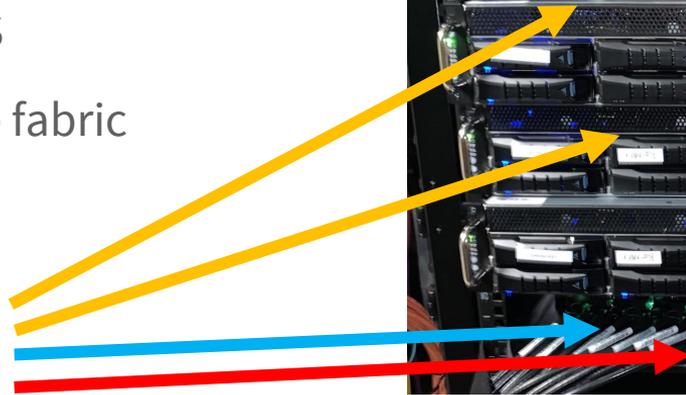
Tools that we built – Keysight DISE



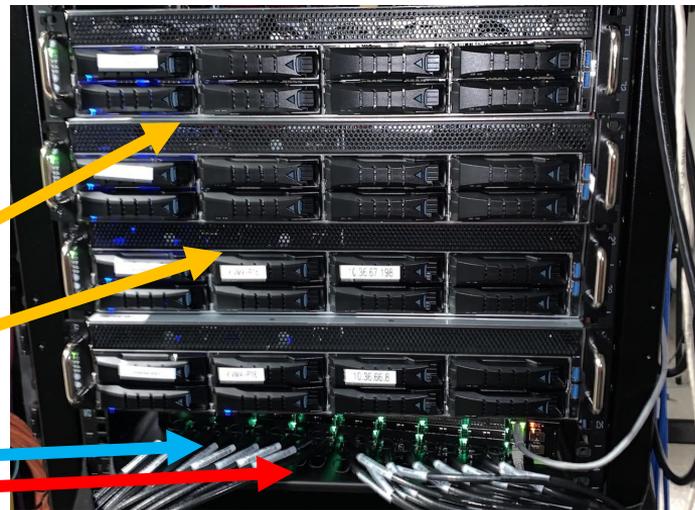
NETWORKING

Tools we are going to use in the demo

- Data Flow Emulator software
- Fabric Emulator NOS
- Chaos module inside fabric
- Jupyter notebook



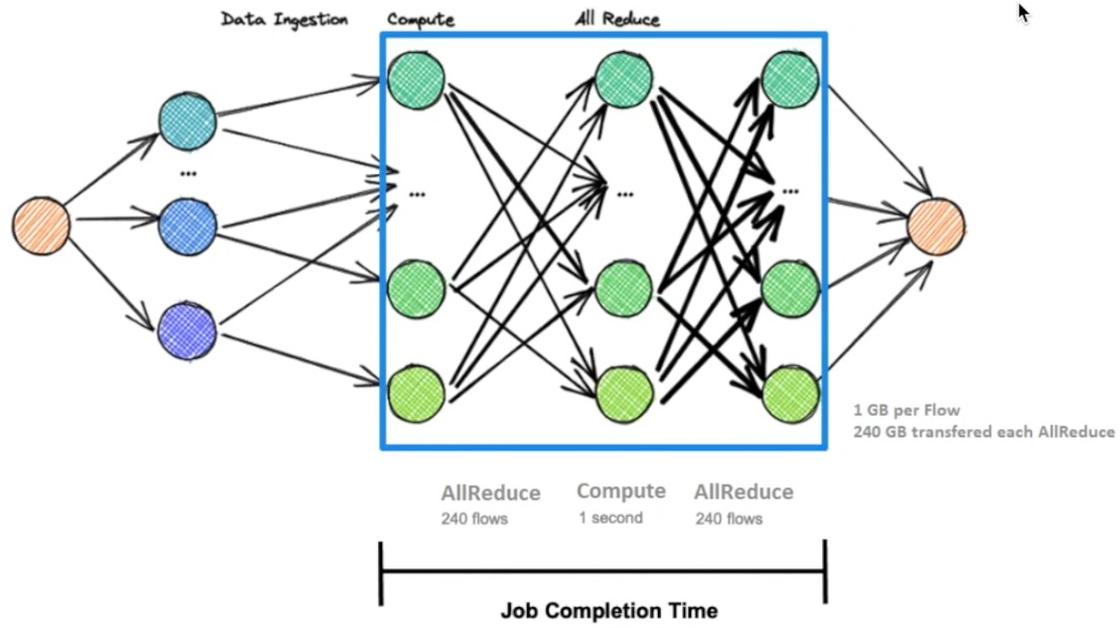
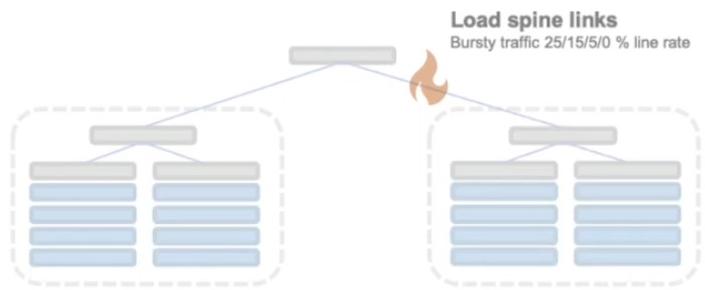
4 compute nodes, 4x100GE NICs / node



OCP spec white box switch

OPEN POSSIBILITIES.





```

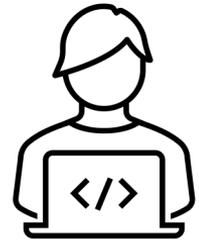
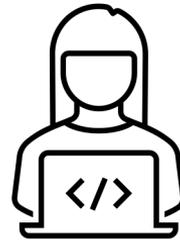
1 # 1. Run Experiments
2 for fabric in FABRIC_SETTINGS:
3
4     # ex: buffer size
5     configure_fabric(fabric)
6
7     for impairment in IMPAIRMENT_SETTINGS:
8
9         # ex: background traffic
10        apply_impairments(impairment)
11
12        for workload in WORKLOAD_SETTINGS:
13
14            # ex: MTU
15            run_workload(workload)
16            save_result()
17
18 # 2. Data Mining
19 # Create Heatmap JCT by Fabric and TCP settings
20 load_all_results()
21 create_jct_heatmaps()
22
    
```

0:00

We built tools, but...

...what if you want to...

- enhance the tools for **new scenarios**?
- **share** your experiments with others?
- **collaborate** with your users?
- run on a **different platform**?
- take your experiments into the **data center**?



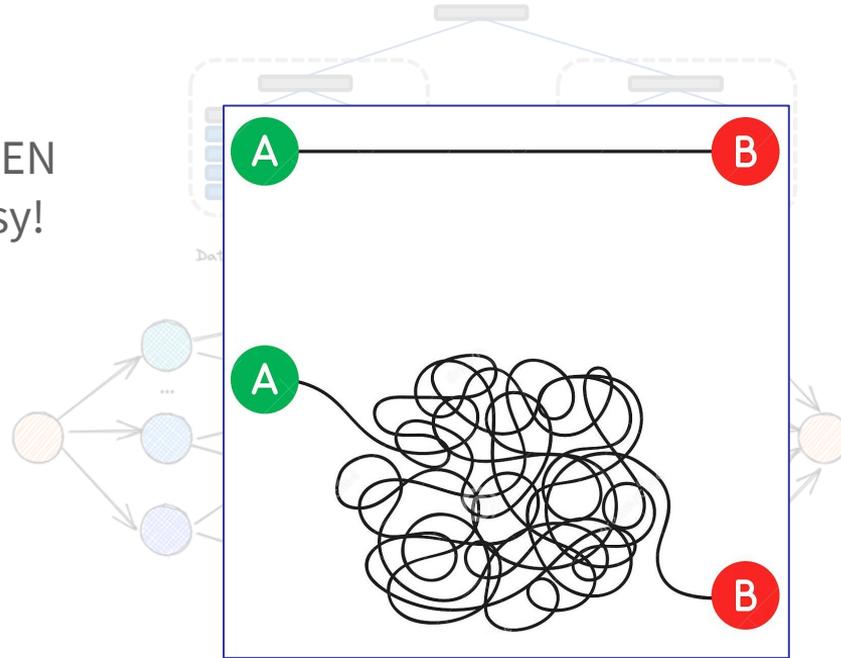
NETWORKING

OPEN POSSIBILITIES.



Engineers unite!

- Let's make this OPEN
- Let's make this easy!



NETWORKING

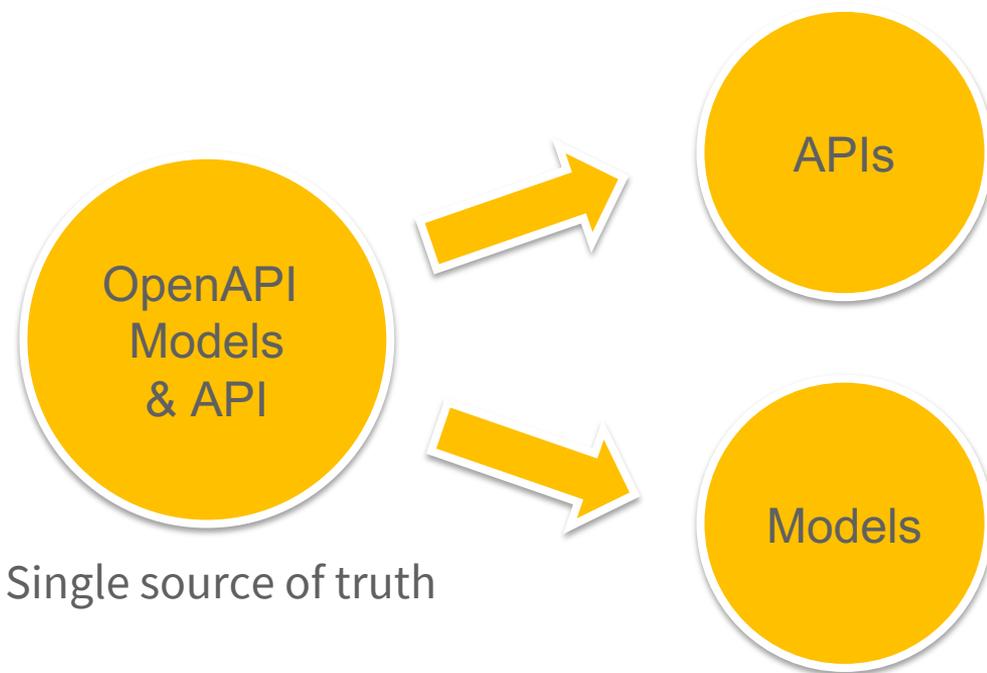
OPEN POSSIBILITIES.



Declarative models, lean APIs



NETWORKING



Single source of truth

Ex: `set_flows()`

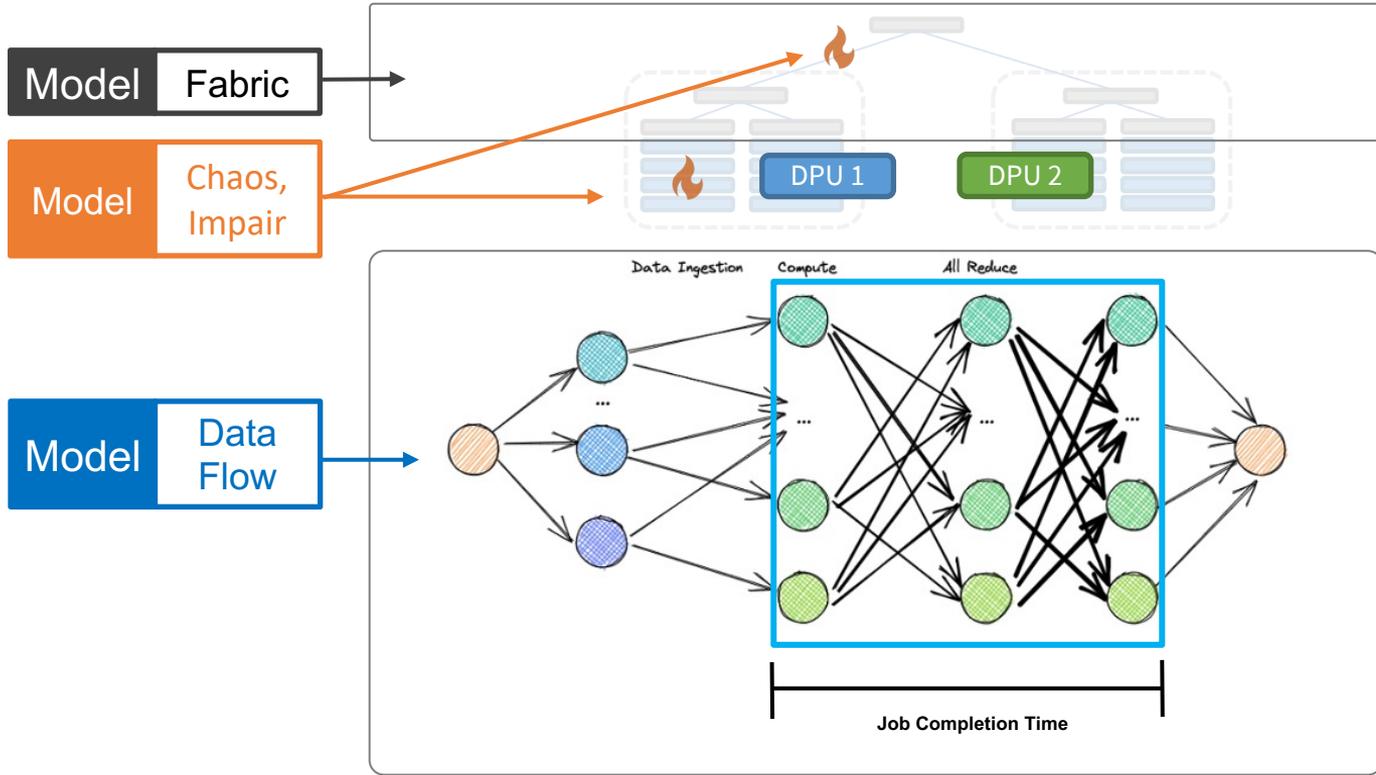
Example:

```
"dataflow": {  
  "steps": [  
    { "name": "digest", "mesh": "full", "hosts": ... },  
    { "name": "compute", ... },  
  ]  
}
```

OPEN POSSIBILITIES.



Independent models



NETWORKING

OPEN POSSIBILITIES.



Demo example config



NETWORKING

```
hosts = [
  { "name": "n1", "address": "1.1.1.1", "tor": 1 },
  { "name": "n2", "address": "1.1.2.1", "tor": 2 }, ...
]
dataflow = {
  "dataflow": {
    "steps": [
      { "name": "digest", "mesh": "full", "hosts": hosts },
      { "name": "compute", "delay": 1 }, ...
    ]
  },
  "tcp": {
    "congestionAlgorithm": "dctcp"
  }
}
fabric_config = {
  "topology": {
    "spine": {
      "count": 1
    },
    "pod": {
      "count": 2,
    }, ...
  }
}
chaos_config = {
  "loss": {
    "link": "Link S.1/1",
    "type": "PERCENTAGE",
    "percentage": 1
  }, ...
}
```

Data Flow Emulator software

On 4 compute nodes, 4x100GE NICs / node



Fabric Emulator

OCP spec white box switch

OPEN POSSIBILITIES.

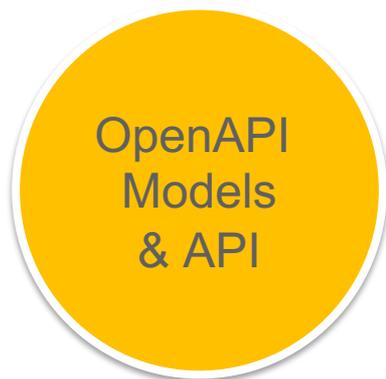


Coexisting implementations



NETWORKING

Open source, vendor-agnostic API,
Rich tooling support



model data

SDK

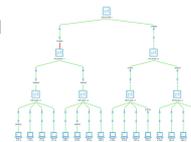
REST

Same script works with any implementation

Vendor-specific
implementations



SIMULATION



EMULATION

TEST
PODS



OPEN POSSIBILITIES.



Acknowledgements



NETWORKING

Ideation:

Matt Bergeron, Ariel Hendel, Winston Liu

Implementation:

Chris Sommers, Vlad Ilie, Jon Stroud, Diana Galan,
Russil Wvong, Lyle Thompson, Andy Balogh

Talks we learned from:

[Bringing the F16 Network into the Lab, OCP Summit 2020](#)
[SIGCOMM Lifetime Achievement Award 2020 Keynote](#)

OPEN POSSIBILITIES.



One last thing...



NETWORKING



JOIN THE TEAM
We're hiring!

Find us at booth **C33**

OPEN POSSIBILITIES.



Call to action



NETWORKING

Engineers unite!

- Let's build good open models
- Contribute your use cases & expand the models
- Compare experimental results in different environments
- As we get traction, form a Sub-Project on the model specs

Where to learn more

- Open Network Experiments

<https://github.com/open-network-experiments>

OPEN POSSIBILITIES.



Open Discussion



OCP
GLOBAL
SUMMIT

NOVEMBER 9-10, 2021