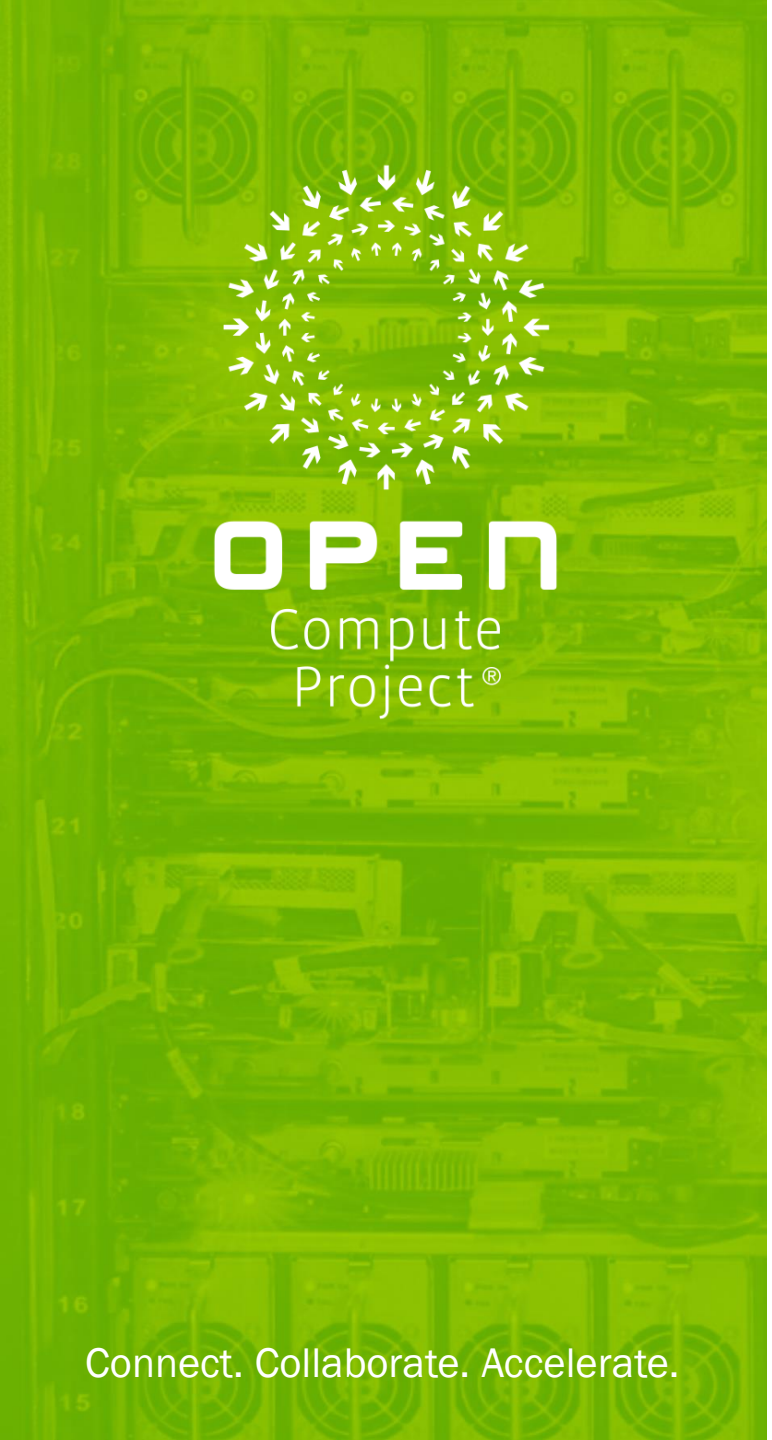


Time Sync in DC What's Next?

Dotan D. Levi

NVIDIA

Outline

- The 10,000 feet Overview:
  - What has changed since TAP established?
- The Next 3 challenges:
  - Extending and Improving the Service
    - Reliable Clock
    - Better Accuracy
  - More Time Sync use cases
  - Component Vendors- can telco product fit in Data center?
  - Call for Action: OCP is all about collaboration, how can you contribute help, make an impact?

What TAP have Accomplished

- Scale up the PTP service
  - Meta posses the largest known PTP deployment
  - Single GM sync millions of nodes (existing technology was 1000s)
  - All Transparent clock, no BCs
    - All measurable, less noise
- Operational
  - Tear up and down using DC tools
  - Monitoring (need more work there)
- Recognition
  - SW stack leveraging time sync – saving $$$

The Next big 3 things

1. Extend the service- we need a Reliable clock
   - Why? / how ?
2. Leverage more applications
   - Huge money saving, in many domains and dicpilne. This is the next X-factor
   - Which apps? What Factors?
3. Transferring clocking Vendors into DC market
   - Its not trivial to use devices that were set to a different market
   - What is missing? And why?

Reliable Clock –Why do we need it?

- When distributed APP is exploiting the time service, it rely on the correctness of the timestamp
- A bad timestamp could lead to a silent data corruption
- Concept was well presented in this paper:
- **“Sundial: Fault-tolerant Clock Synchronization for Datacenters”** *Yuliang Li et al*
  - https://www.usenix.org/conference/osdi20/presentation/li-yuliang

Problem statement is clear and accepted

but proposed solution doesn’t fit DC scale.

Reliable Clock –How?

- First Thing we did is to establish TAP workstream on Clock interface
  - Extending the clock API to include parameters, that will allow SW to calculate the error bound, (dynamically) and the confidence level of the time
  - Please Join!
- We also wish to Explore a more reliable service
  - Use big data
  - Use physical layer assistance
  - Extending PTP
  - Might create a workstream on this next year

Time Sync Use-Cases

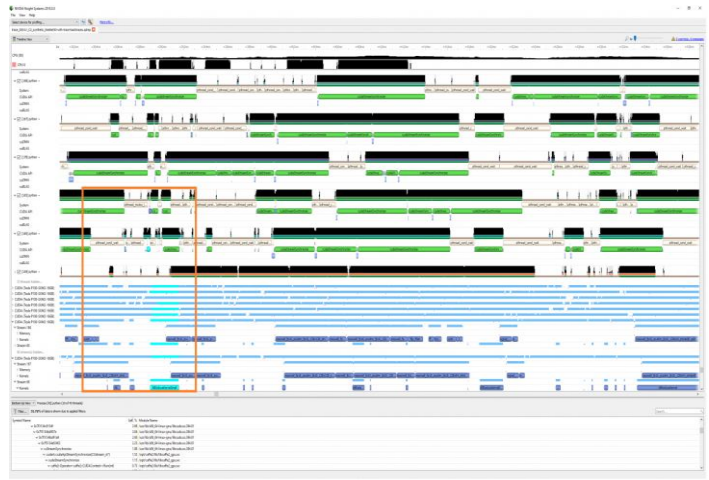
Globally Sync Database

Avoid using commit wait,
reduce mutex stalls
X3 accelration

TAAS

For any cloud Application
E-Commerce, FSI, etc…

Cluster Debug



Sync all trace and event on
a single coherent time axis

Connect. Collaborate. Accelerate.

Time Sync Use-Cases

5G / EDGE

Enabling 5G edge softwarization, enabling 100mhz cell from a SW system

HPC

Accelerate HPC collective Offloads by Time-load-balance the inner networking

Enterprise Service

Broadcasting, M&E, Healthcare, Metaverse/Omniverse

Connect. Collaborate. Accelerate.

Time Sync use Cases

Call for Action:

Application Architects: engage with our expert to explore your options to accelerate

Vendor's wake up call

- Time sync offers an opportunity for clocking vendor to penetrate the datacenter market, as advance clocking solution is needed

- DC is a different market that the traditional market, and require some more capabilities

What is Needed? (2 examples)

- Pinout computability
  - From supply chain issues, each component must have an alternate solution
  - Innovation and quality will not be able to justify vendor locking
  - We will strive to define a packaging standard for DC clocking solution
- Fail safe Programming
  - Programable device must guarantee that the programming flow is fail safe
  - DC operations run fleet commands on 1,000,000 of nodes at once….

Summary – Call for Action

- Join TAP!
  - Contribute, make an impact, we need your Brain and POV
  - Existing workstreams needs more support
- Application architect : explore leveraging your APP using time sync
  - We will publish a “cook-book” soon
- Component Vendors: reach out to find how to add robustness for your components.
- Help us establish the Clocking device workstream to align pinout and architecture

Thank You

Connect. Collaborate. Accelerate.