The Stateful Landscape: The Then and Now and the Future

Alex Chircop - Founder & CEO StorageOS
Erin Boyd - Cloud Engineering - Apple
Quinton Hoole - Production Engineering - Facebook

Agenda

- Overview of the SIG, how to join and how to help
- Overview of storage projects in the CNCF
- Projects that are currently being being reviewed
- Areas we would like to see, and gaps in the landscape
- Overview of the CNCF Storage Landscape document
- Overview of the Performance and Benchmarking document

Overview

Meetings are on the 2nd and 4th Wednesday of every month at 8AM PT (USA Pacific)

- **Home**: https://github.com/cncf/sig-storage
- **Conf call:** http://bit.ly/cncf-sig-storage-call
- **Agenda:** https://bit.ly/cncf-storage-sig-minutes
- **Recordings:** https://bit.ly/cncf-storage-sig-recordings
- **Mail list:** https://lists.cncf.io/g/cncf-sig-storage

Our calls and membership are open!

SIG Storage

Who we are...

- We are a diverse set of users & developers of Cloud Native technologies with a storage focus
- We are leaders & early adopters

- We are organized as:
  - **Co-Leads**
    - *Alex Chircop*
    - *Erin Boyd*
    - *Quinton Hoole*

  - **Tech Leads**
    - *Xing Yang*
    - *Luis Pabón*
    - *Sugu Sougoumarane*
    - *Saad Ali (also a TOC member)*

What we do

“Scale contributions by the CNCF technical and user community, while retaining integrity and increasing quality in support of the CNCF mission (to make cloud native computing ubiquitous).”

...this means we

- Educate
- Review Storage Project Proposals
- Engage with the user community
- Work directly with the TOC & other SIGs as subject matter experts

Education

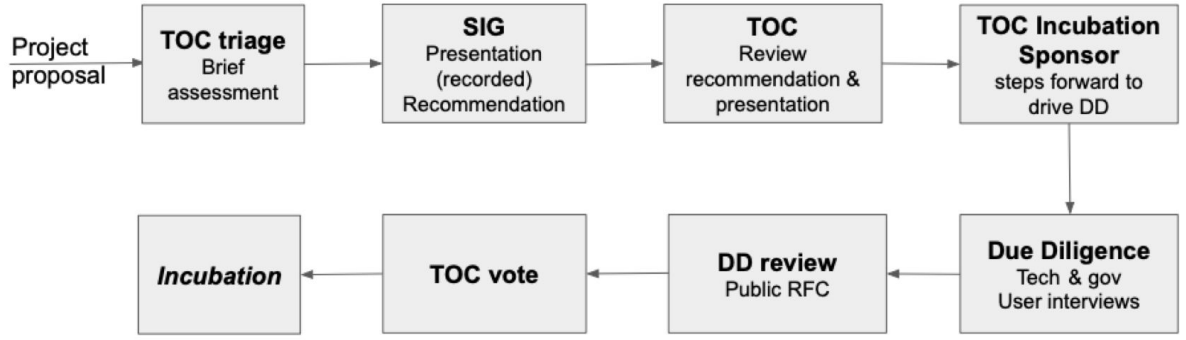
End User Education

- Provide up-to-date, high quality, unbiased and easy-to-consume material to help end users to understand and effectively adopt cloud-native technologies and practises within the SIG's area, for example:
  - White papers, presentations, videos, or other forms of training clarifying terminology, comparisons of different approaches, available projects or products, common or recommended practices, trends, illustrative successes and failures, etc.
    i. CNCF Storage White Paper (https://bit.ly/cncf-storage-whitepaperV2)
    ii. SIG Storage Performance & Benchmarking Whitepaper
  - As far as possible, information should be based on research and fact gathering

Review

Project Review

- Understand and **document** a high level roadmap of projects within this space, including CNCF and non-CNCF projects. Identify gaps in **CNCF project portfolio**.
- For projects that fall within the CNCF, perform **health checks**.
- Perform **discovery** of and **outreach** to candidate projects.
- **Help candidate projects** prepare for presentation to the TOC.
- Every CNCF project will be assigned to one suitable SIG by the TOC.



Engage

End User Input Gathering

- Gather useful end user input and feedback regarding expectations, pain points, primary use cases, etc.
- Compile this into easily consumable reports and/or presentations to assist projects with feature design, prioritization, UX, etc.

Community Enablement

- SIGs are open organizations with meetings, meeting agendas and notes, mailing lists, and other communications in the open.
- The mailing list, SIG meeting calendar, and other communication documents of the SIG will be openly published and maintained.

Trusted Expert Advisors to the TOC

- Perform **technical due diligence** on new and graduating projects, and advise TOC on findings.
- Be involved with, or **periodically check** in with projects in their areas, and advise TOC on health, status and proposed actions (if any) as necessary or on request.

Community

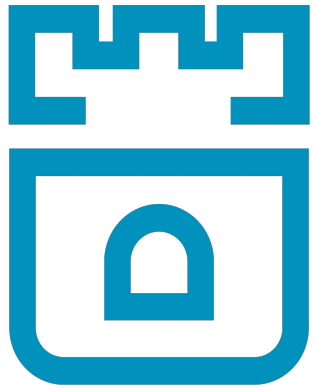
- How you can get involved?
  - Join our meeting
    - 2nd & 4th Wednesday each month
  - Submit and help review projects for consideration
    - https://github.com/cncf/toc/tree/master/process

- We value community presentations of projects in the cloud native storage space including, but not limited to: *management frameworks, block stores, filesystems, object stores, key-value stores and databases*

- Several projects have presented to the SIG such as: CSI, Rook, REX-Ray, TiKV, Dotmesh, Yugabyte, OpenEBS, Open Services Broker, Vitess, Minio, OpenSDS, Redfish/Swordfish, ChubaoFS, Longhorn, Dragonfly, Harbor, Pravega, Piraeus, Dataset Lifecycle Framework, Linstor

CNCF Storage Projects

Graduated



etcd

TiKV

Incubating

Dragonfly

CNCF Storage Projects

The CNCF Sandbox is the entry point for early stage projects.

https://www.cncf.io/sandbox-projects/

Current Projects in Review

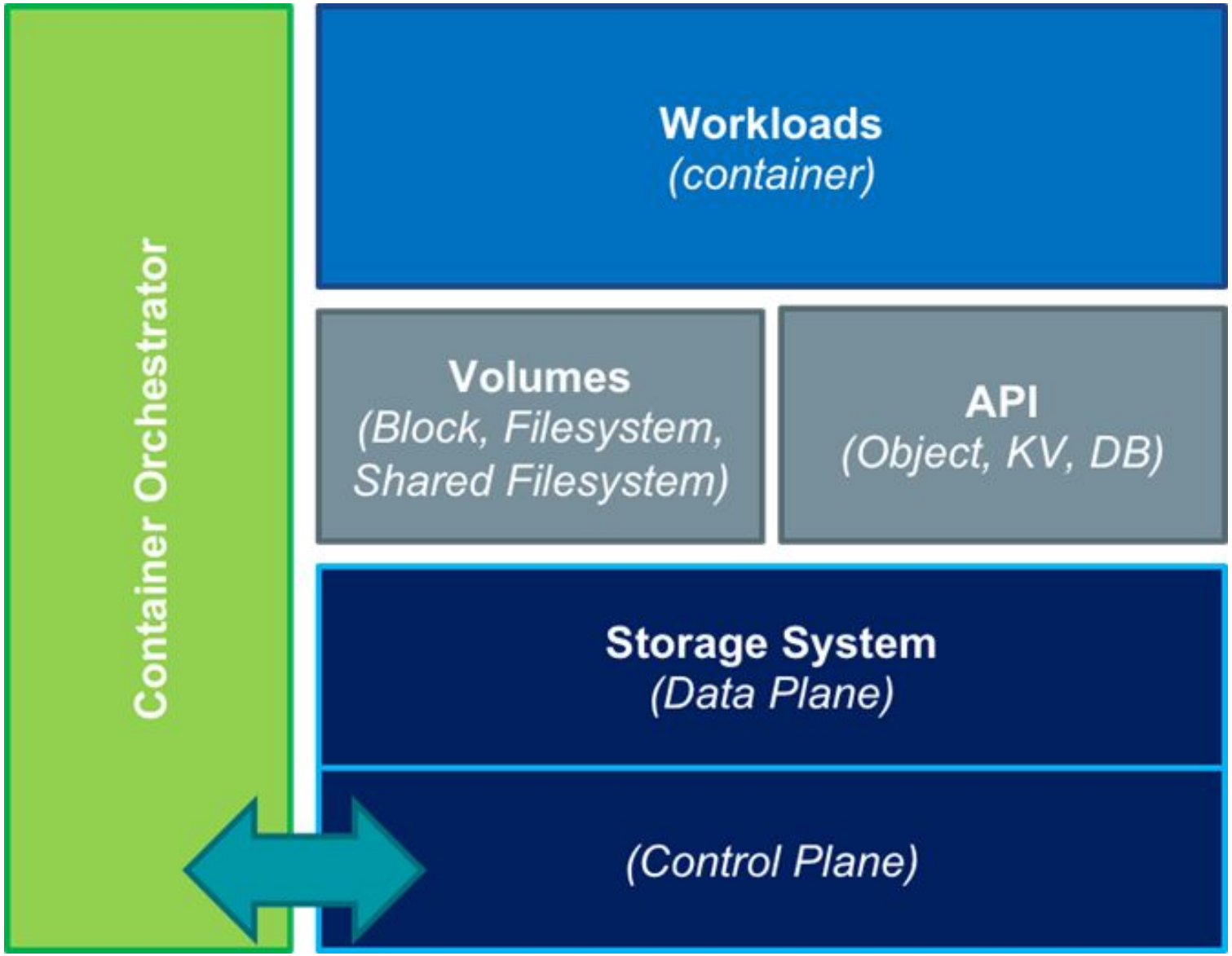
- **Incubation**
  - Pravega
    - https://github.com/pravega/pravega

- **Sandbox → Incubation**
  - OpenEBS
    - https://github.com/openebs/openebs

CNCF Storage Whitepaper

- **Whitepaper: https://bit.ly/cncf-storage-whitepaperV2**
- **Definition of the attributes of a storage system**
- **Definition of the layers in a storage solution with a focus on terminology and how they impact the attributes**
- **Definition of the data access interfaces in terms of volumes and application APIs**
- **Definition of the management interfaces**

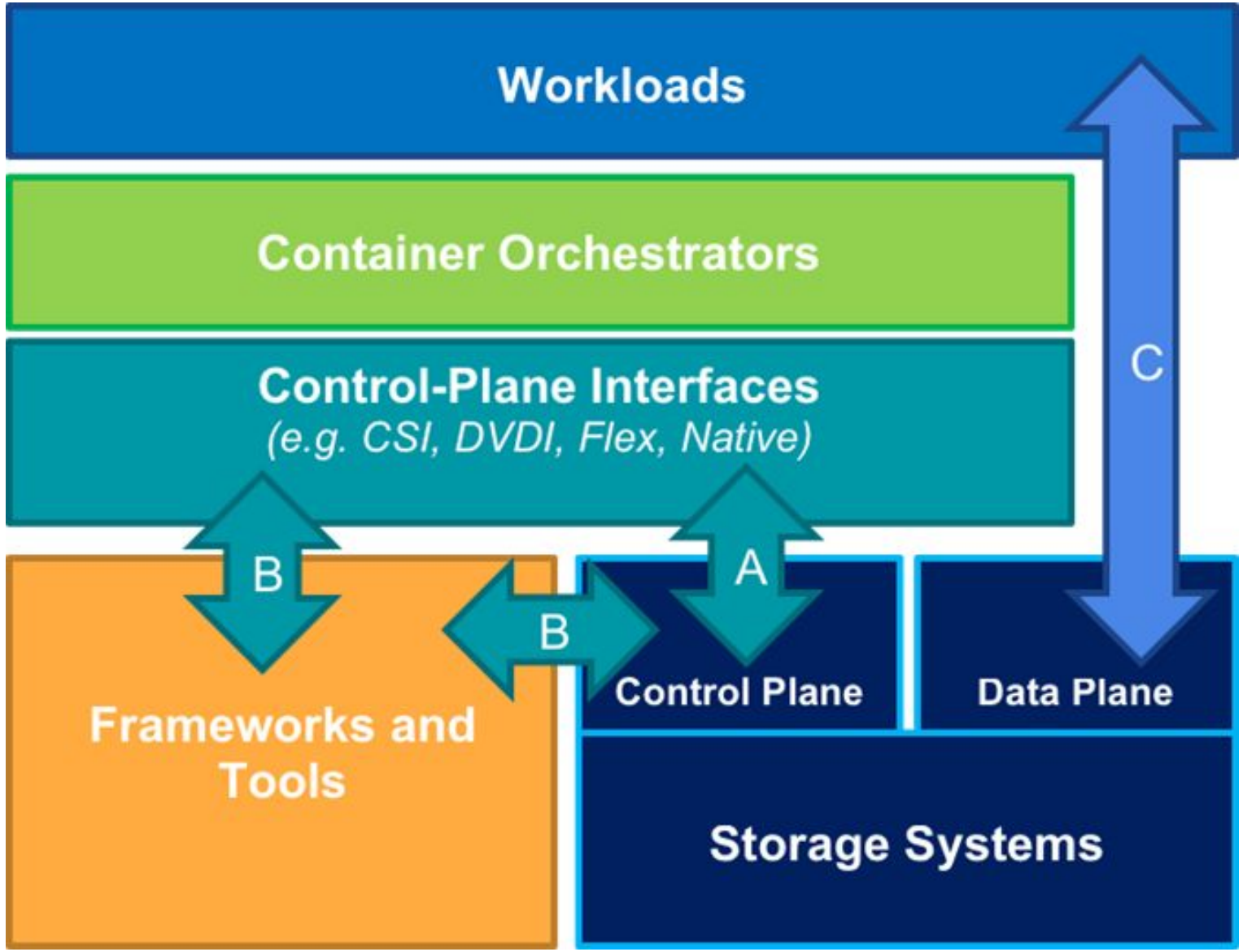
Data Access Interfaces



Storage can be accessed via **Data Access Interfaces**:

- **Volumes** – accessed through a more traditional file interface in a **block** or **filesystem** interface
- **API** – other ways to persist data such as **object stores**, **KV stores** or **databases**

Management Interfaces



Container Orchestration system (**CO**) uses an interface to interact with a storage system

The storage system can:
(A) support control-plane API directly
(B) interact via an API Framework layer or other Tools

Workloads consume **(C)** storage via a data access interface

Storage Attributes

Availability	Scalability	Performance	Consistency	Durability
•Failover	•Clients	•Latency	•Delay to access correct data after a commit	•Data protection
•Moving access between nodes	•Operations	•Operations		•Redundancy
•Redundancy	•Throughput	•Throughput	•Delay between commit and data being committed to non-volatile store	•Bit-Rot
•Data Protection	•Components			

Storage Layers

Orchestrator, Host and Operating System
Storage Topology (centralized, distributed, sharded, hyperconverged)
Data Protection (RAID, Erasure coding, Replicas)
Data Services (Replication, Snapshots, Clones, etc.)
Physical, Non-Volatile Layer

Performance Whitepaper

- **Whitepaper: http://bit.ly/cncf-sig-storage-performance-benchmarking**
- **Definition of common concepts for measuring performance and benchmarking for volumes and databases**
- **Definition of common pitfalls and considerations**
- **Sample tooling for benchmarking**
- **Important takeaway: published results are not useful for making comparisons - it is hard to compare published results without a deep understanding of the test conditions, so it is always important to run your own test**


KEEP CLOUD NATIVE
EVERYWHERE
KubeCon
CloudNativeCon
North America 2020
Virtual