SODA Open Data Framework Update

Key Features, Releases, Landscape ….

Noel McLoughlin

Najmudheen CT

Saurabh Gupta

Ashit Kumar

Agenda

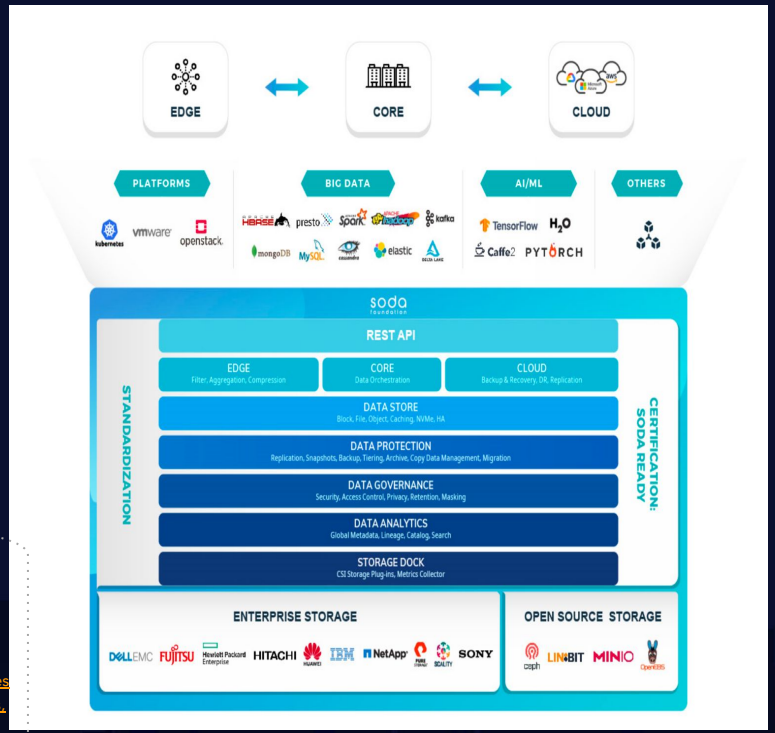
- Introduction to SODA Projects
  - SODA API, Controller and Dock [Ashit]
  - SODA multi-cloud [Ashit]
  - SODA NBP [Naju]
  - SODA Delfin [Naju]
- SODA Experience
  - SODA Installer [Noel]
- Use case and End-User Adoption [Saurabh]
- Roadmap and Session Close [Ashit]
- QnA

SODA Framework

SODA is a single data framework connecting disparate solutions into seamless end to end solutions. This framework is open source, allowing any developer, vendor, or end user to build and extend upon it.

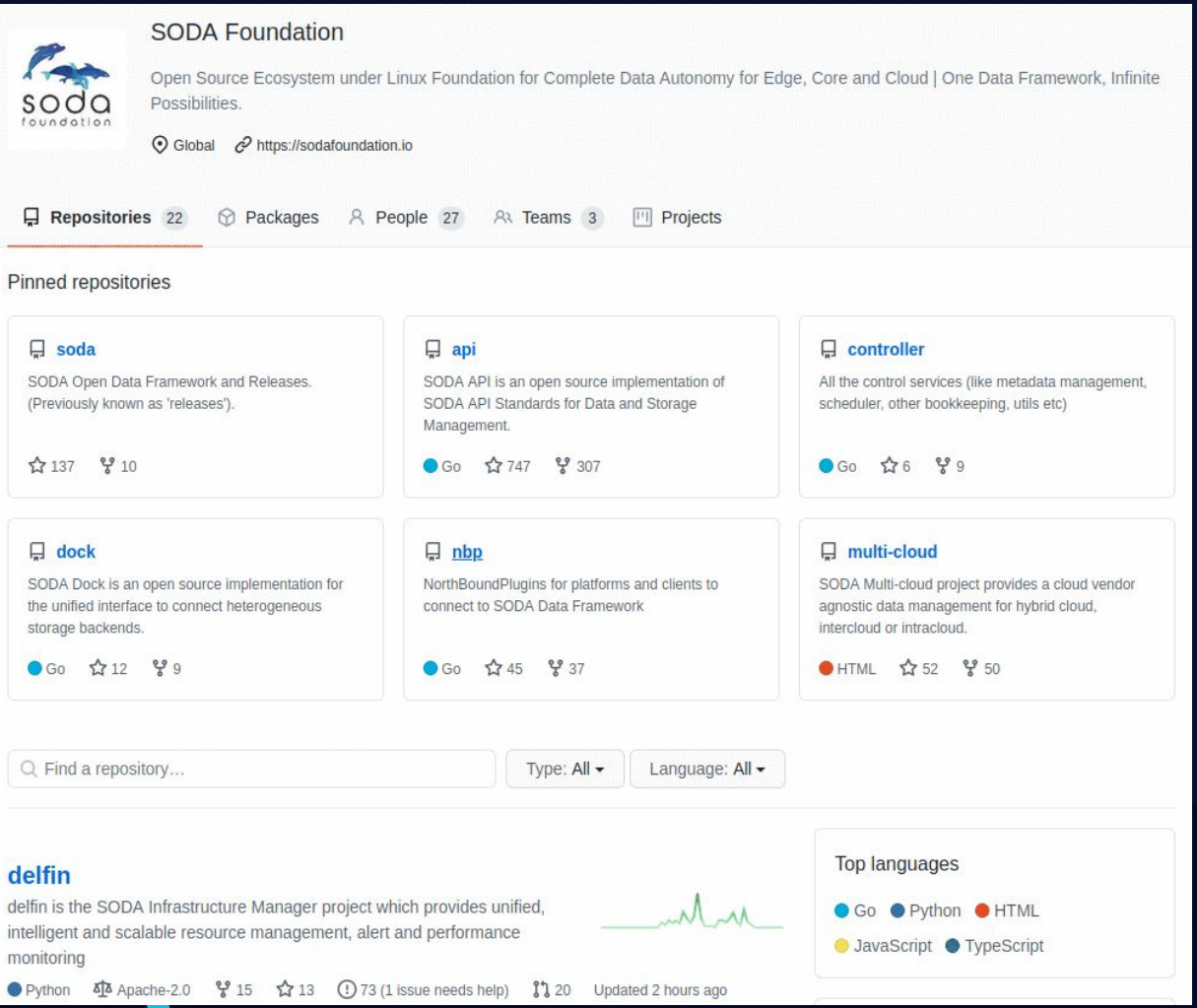
The value propositions of this framework are:

- data mobility
- standardized interfaces
- interoperable solutions
- optimized transfers
- autonomous operations
- high scalability



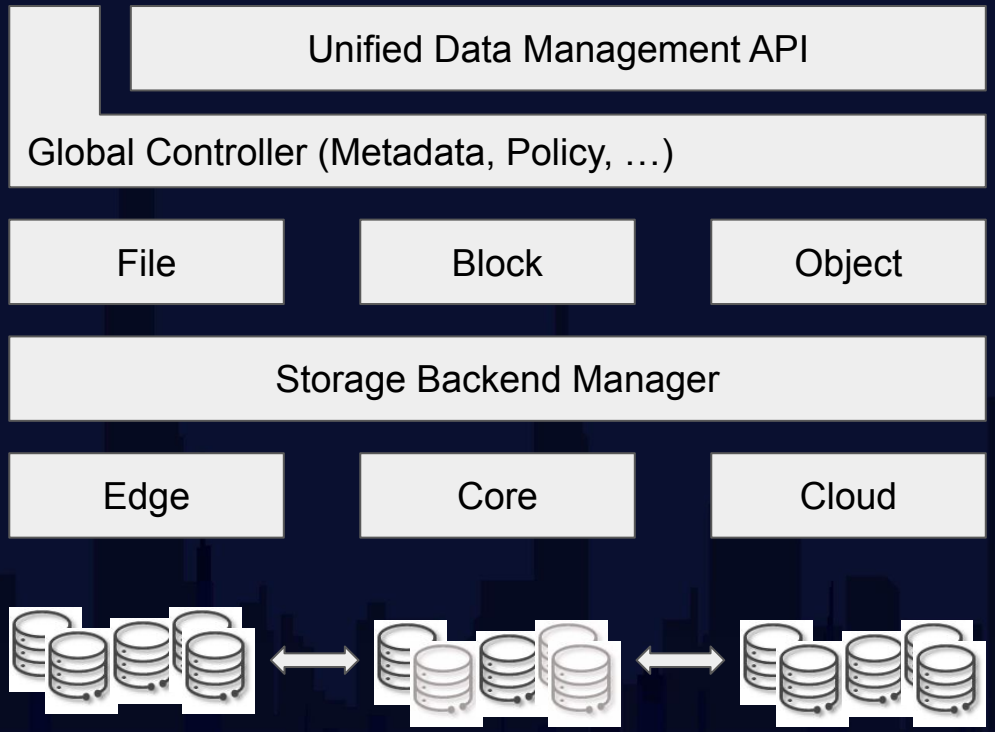
HIGHLIGHTS

- open source - data and storage projects
- standardization – data and storage management
- ecosystem – hardware, software, solutions, services
- certification – ecosystem components, developers, operators



#sodacon2020

One Data Store



The Journey of SODA Projects

Cloud

SODA
Multi-cloud

On-Prem

API,
Controller,
Dock,
Delfin

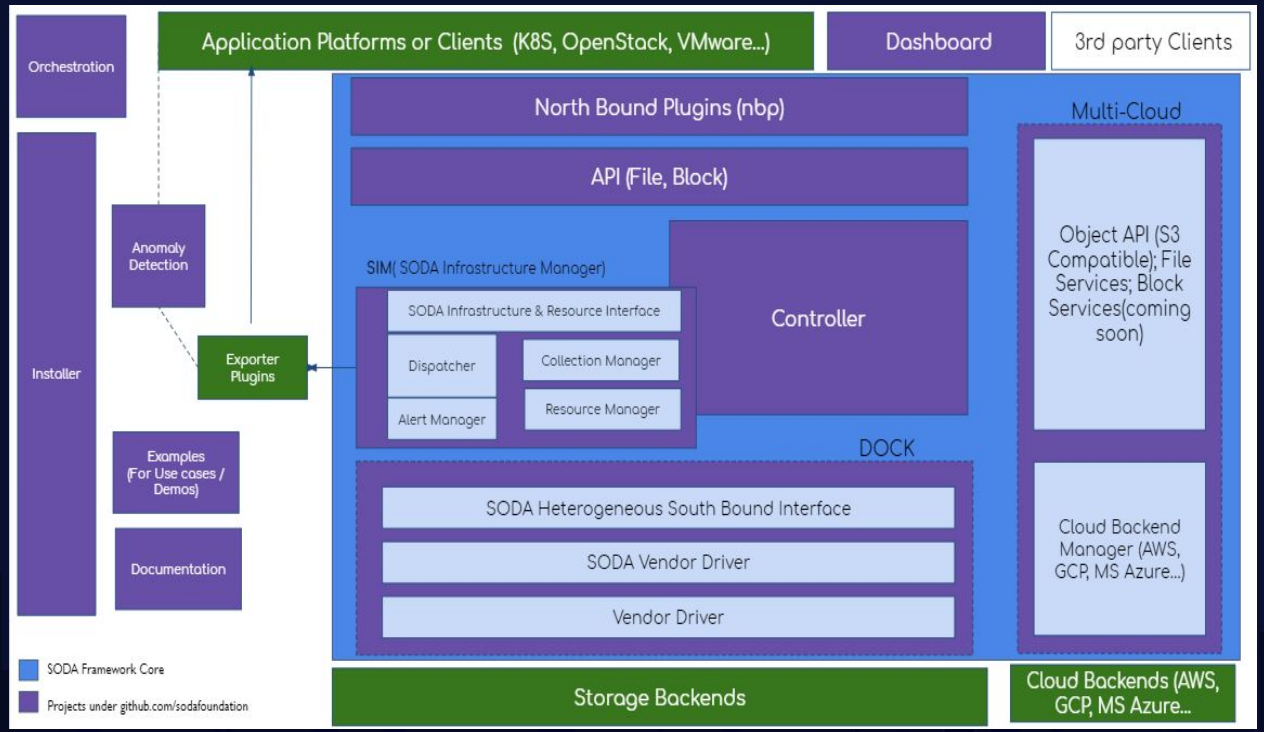
Edge

SODA-Edge

SODA Experience and Experiments

Dashboard
Installer
Documentation
Orchestration and Anomaly Detection

With On-prem and Cloud



1

Core (On-prem) Data Management Across Heterogeneous Storages

2

Core (On-prem) Storage Monitoring Across Heterogeneous Storages

3

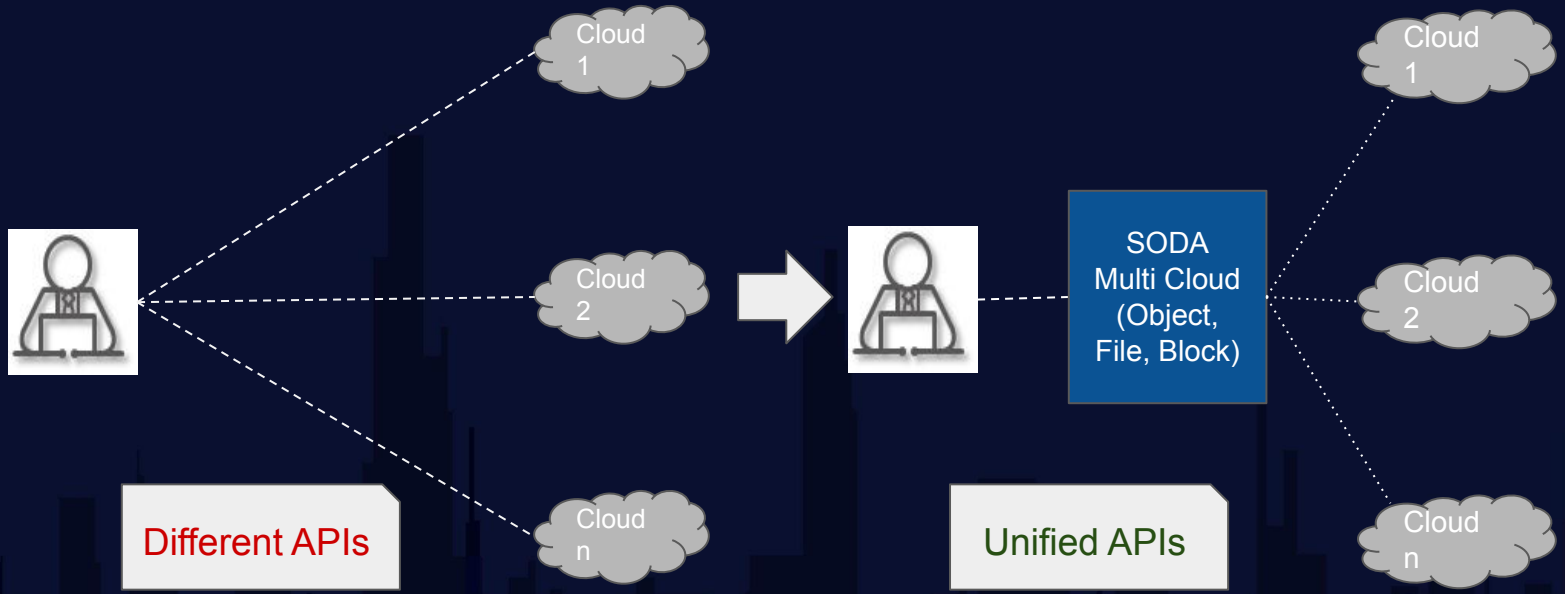
Multi Cloud Data Management Across Heterogeneous Cloud Vendors (and on prem)

On-Prem: Where we are?

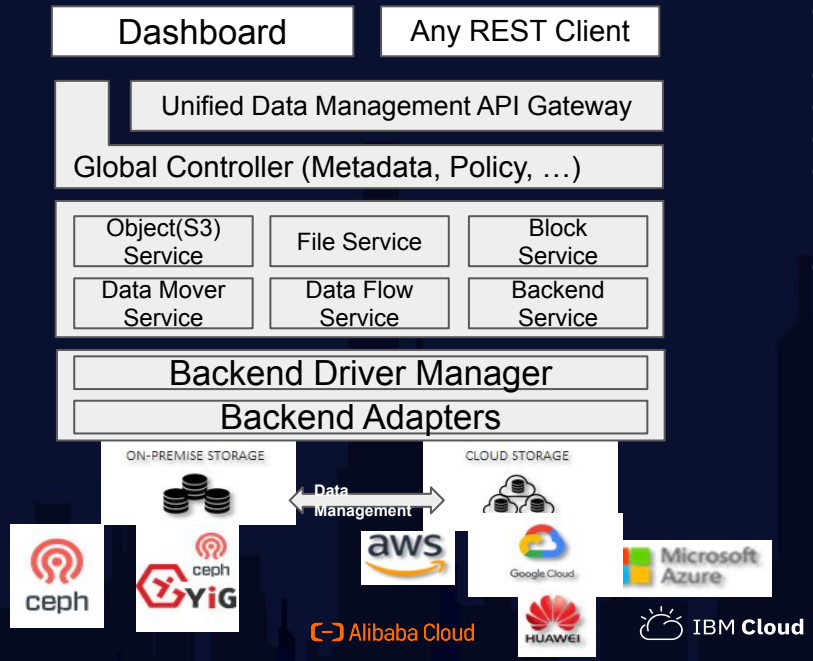
- Unified API for File and Block Management
- Heterogeneous storage support
  - Supported Storage Drivers
    - Ceph
    - DRBD
    - Fujitsu
    - HPE Nimble
    - Huawei Oceanstor [SAN and NAS] and FusionStorage
    - IBM Spectrum Scale [Block]
    - NetApp [SAN and NAS]
    - Scutech
    - Linux Native LVM
    - ChubaoFS
    - OpenStack Manila and Cinder
- Distributed Metadata management

SODA Multicloud

Unified APIs for Data Management Across Multiple Cloud Vendors



Multi-cloud

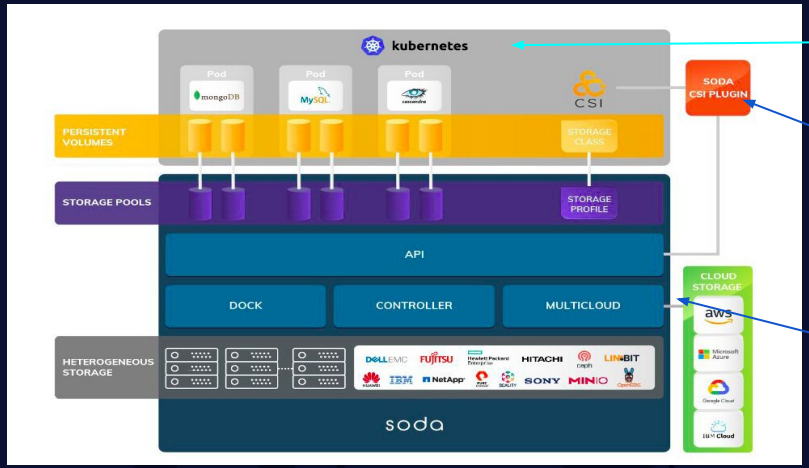


- Provides a cloud vendor agnostic data management for hybrid cloud, intercloud or intracloud.
- REST based interface
- Global Configurations
- Supports Object, File and Block
- Policy Based Data Management:
  - Data Migration
  - Data Lifecycle Management
- Heterogeneous On Prem:
  - Ceph and YIG Ceph
  - SODA On Prem Connectivity for heterogeneous storages (in progress)
- S3 Compatible Object API
- Pluggable Adapter Model for Backends (Easy to add backends)
- Security (Enhancing..)
- Supported Cloud Vendors
  - AWS, Azure, GCS, IBM, Huawei, Ceph, YiG, Alibaba
- Archival and Retrieval (Upcoming…)

SODA Plugin (NBP Project)

CSI Plugin

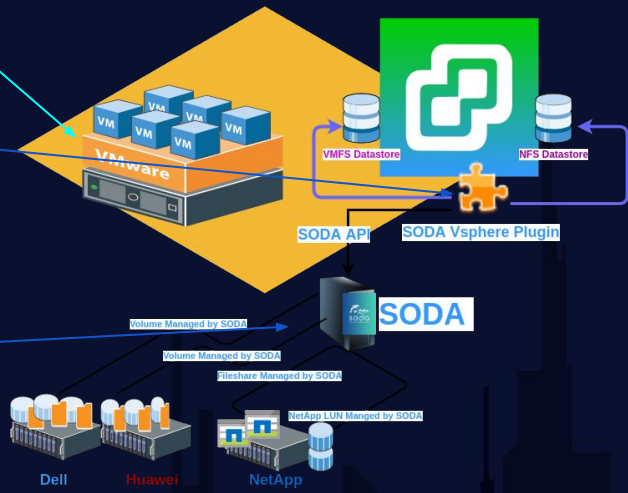
VMWare Plugins



PLATFORMS

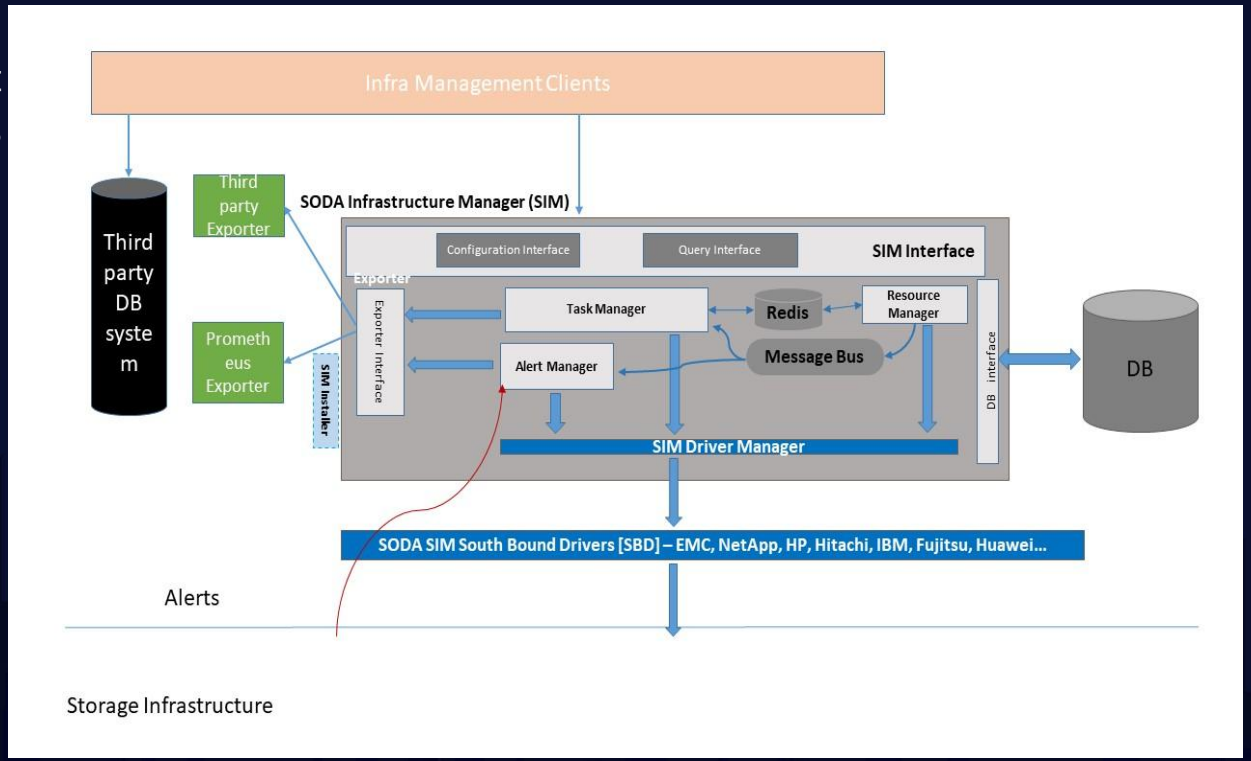
PLUGINS

SODA CORE



Intelligent Storage Monitoring and Management (Delfin)

- SODA foundation project for unified heterogeneous resource and performance monitoring and alerting
- Drivers to collect data
- Exporters to push data
- A scalable architecture to add drivers and exporters.



SODA Experience (Installer)

- *Installer*
  - Single Installer for different SODA projects - **Ansible** (official), **Salt** (experimental)
  - Challenge: Fragmentation / Lack of interest in installers in FOSS / Repeatable
  - Competencies: K8S/Ansible, Salt; yaml, cuelang, json; Microservices; CNCF

- *Orchestration*
  - **Stackstorm**, **Apache-Airflow** (experimental)
  - Problem: Fragmented from Installer / Integration with Soda/User TechStack

- **Redesign the SODA Experience - New SODA Installer(s)**
  - Software **Upgrades** / Configuration **Updates**; Models / Deployments / Configurations.
  - Models - Component Fragment model => Soda Model => Deployment Models
  - Models - Merge, Validate, Iteration, Consume (YAML + JINJA + CUELANG)
  - Model scaling: Fragment -> Project -> Deployment (software, config, version)
  - Pattern merge, technology, use cases, Engineering, CI/CD
    - YAML modelling, Cuelang validation, Jinja2 Templating, Orchestration of Data Mobility

Requested Features for SODA installer #408

Open **noelmcloughlin** opened this issue 23 days ago · 0 comments

noelmcloughlin commented 23 days ago • edited Member

Issue/Feature Description:

These requirements are derived from current open issues for features which should be working.

- Example Use cases (Referenece Deployments) [**@rajat-soda**] [#404]
- Single Command Install for SODA with default configs [**@rajat-soda**] [#403]
- SODA orchestrator (soda upgrade paths) [**@noelmcloughlin**] [#400]
- Cleanup installation [**@anvithks**, **@joseph-v**] [#399, #278]
- Ubuntu 20.04 / CentOS / SystemD [**@osiveteam**, **@kumarashit**, **@kumarashit**] [#353, #261, #381]
- CI/CD for installer [**@kumarashit**, **@kumarashit**] [#345, #396]
- SODA Installation with Single Command using defaults [**@rajat-soda**] [#388]
- SODA Installer UI [**@osiveteam**] [#386]
- SODA Idempotence Installer [**@NajmudheenCT**, **@joseph-v** , **@Shruthi-1MN**] [#401, #369, #277]
- SODA (?) Cloud Native Installer [**@joseph-v**, **@sushanthakumar**, **@PravinRanjan10**] [#361, #328, #323]
- SODA Validate Released Artifacts [**@kumarashit**] [#352]
- SODA Pre-install check and configure [**@kumarashit**] [#350]
- CNCF compoents configuration (no hardcoding) [**@himanshuvar**, **@joseph-v**, **@Shruthi-1MN**] [#376, #327, #375, #382]
- CNCF component installers [**@PravinRanjan10**, **@thatsdone**][#326, #270]
- SODA configuration [**@kumarashit**, **@sushanthakumar**, **@NajmudheenCT**] [#318. #306, #301]
- SODA Upgrade/Reinstall, no data loss [**@joseph-v**] [#279]
- YIG installer [**@276**]

Assignees

noelmcloughlin

Labels

None yet

Projects

None yet

Milestone

No milestone

Linked pull requests

Successfully merging a pull
issue.

None yet

Notifications

Unsu

You're receiving notificatio
watching this repository.

SODA Installer Roadmap

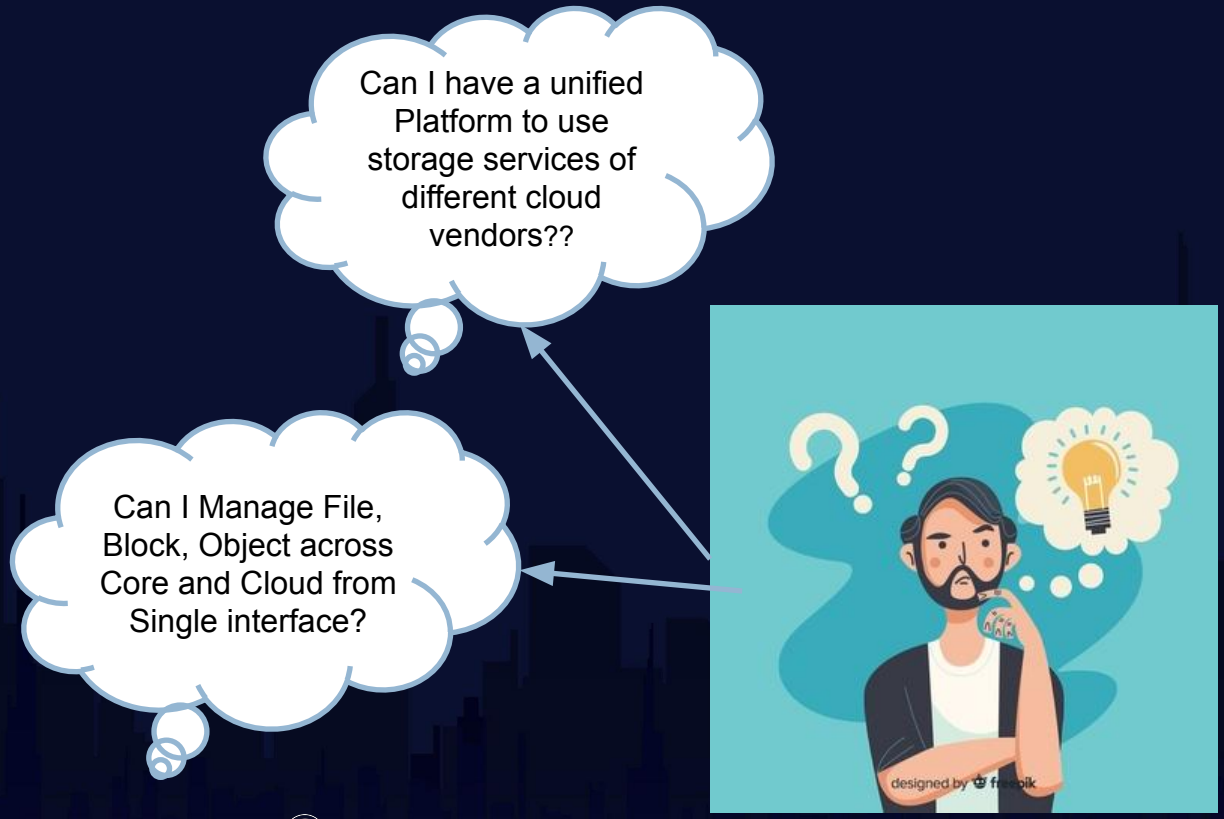
- **Goal: Model Driven Deployments + Ecosystem reuse + No open Issues = Summer 2021**
  - Each CNCF component has **downstream installer** (Ansible Galaxy, SaltStack formulas)
  - Each CNCF component has **downstream CI/CD** (Ansible Galaxy, Saltstack formulas)
  - Each SODA project has **own installer** (scale down)
  - Each SODA project has **own Component model, fragment of Common Model**
  - Each User has own **Sitedata Model** (networking, urls, versions, integrations)
  - Installer merges all model fragments into **Common deployment model (scale up)**
  - Installer applies the **Deployment Description**
- Technologies:
  - Modeling: YAML, JINJA2 Merge, CUELANG Model Validation, K8S Kustomize.
  - Automation: Ansible Playbook, Saltstack Formula, Event Driven Orchestration (Airflow, Stackstorm)
- Collaboration: SNIA, GAIA, IDSA, Multi-Cloud, CNCF, Ansible-Galaxy, Saltstack Formulas.

End User Use-cases

- Heterogeneous Storage Monitoring [On-Prem]
- Multi-cloud Data Management
- Heterogeneous Object storage Management [On-Prem/Hybrid Cloud/Public Cloud]
- Workflow based Orchestration and Automation
- Intelligent Storage Monitoring [On-Prem/Hybrid Cloud/Edge]

End-Users: Trials and adoption

NTT
Yahoo! Japan
Huawei
Toyota
Click2Cloud



How Soda Multi-cloud helped Cloudsbrain

- **SODA multi-cloud Provides CloudsBrain with Cloud Storage Solution which includes :-**

❖ Object Operation and management
- ✔ Upload Object
- ✔ Upload Multi-part Object
- ✔ Upload multiple object at once
- ✔ Download Object
- ✔ Delete Object

❖ Object Lifecycle Management
- ✔ In-cloud Lifecycle transition
- ✔ Cross-cloud Lifecycle transition

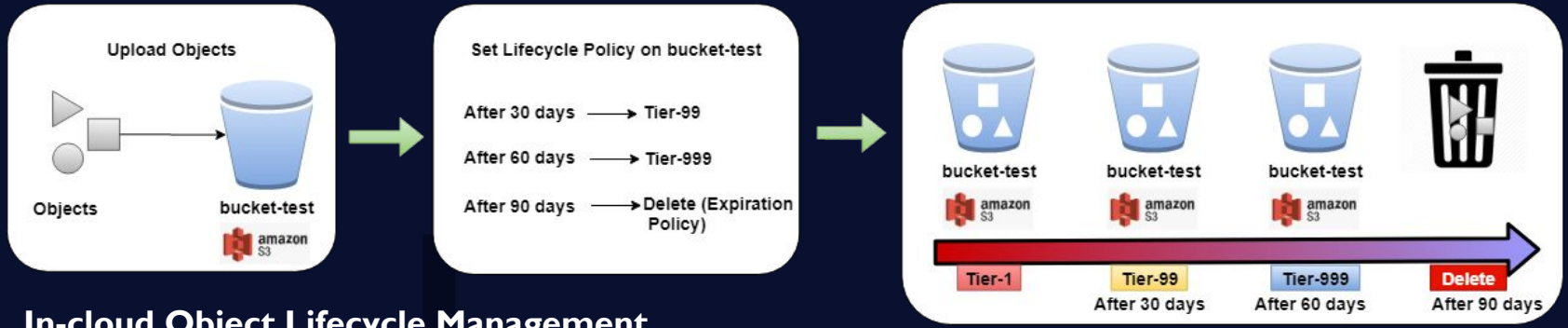
❖ Migration feature
- ✔ Cross-cloud Migration
- ✔ Scheduled Migration

SODA Multi-Cloud: Object Life Cycle Management

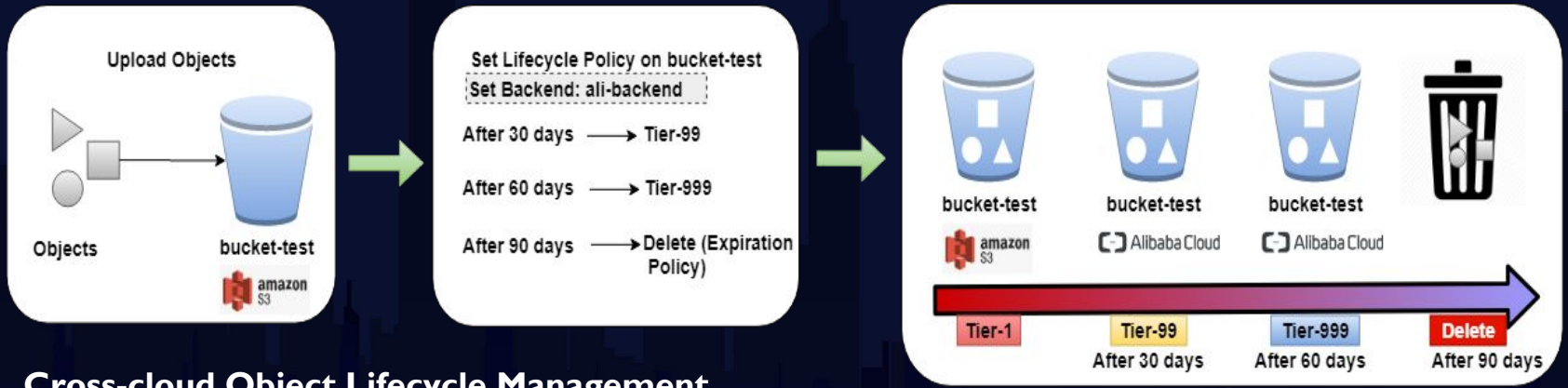
- SODA combines different storage resources on-premise, and across multiple clouds to build tiers.
- In SODA, we have generalized storage class
  - If frequently used: **Tier-1**
  - If used infrequently: **Tier-99**
  - If not in use but still due to data policy, we store it for future purpose : **Tier-999**
- Using data lifecycle policies, data that meets the lifecycle conditions is moved to the next tier, allowing data to be stored efficiently throughout their lifecycle.

SODA Storage Tier	AWS	Azure	Alibaba	Huawei
Tier-1	STANDARD	HOT	Standard	STANDARD
Tier-99	STANDARD_IA	COOL	IA	WARM
Tier-999	GLACIER	ARCHIVE	Archive	COLD

SODA MULTI-CLOUD: OBJECT LIFE CYCLE MANAGEMENT

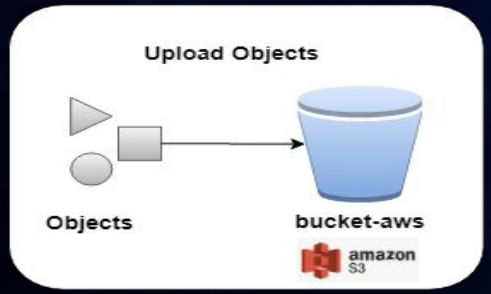


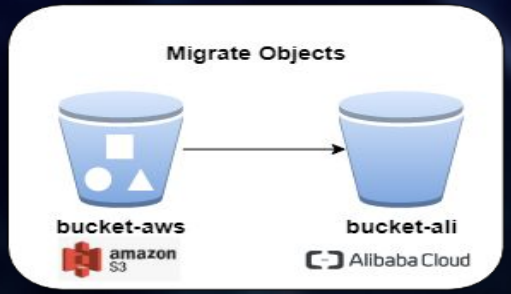
In-cloud Object Lifecycle Management

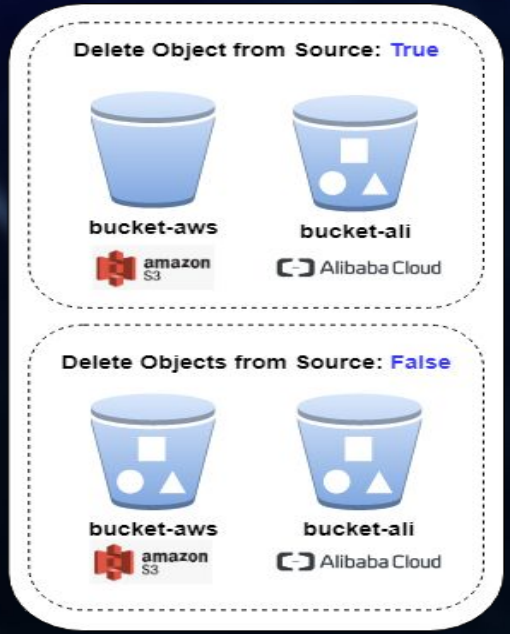


Cross-cloud Object Lifecycle Management

SODA MULTI-CLOUD: BUCKET/OBJECT MIGRATION







Pre-requisite:

- Register Backend (required AK, SK, region, endpoint, cloud-bucket name)
- Create a bucket on registered backend (eg. **bucket-aws**)
- Register another backend and create bucket (eg. **bucket-ali**)
- **Upload Objects**
- Upload multiple objects to bucket-aws
- User can upload object in multi-part also

Migrate Objects

- User can migrate data across the cloud.
- Supported storages: AWS-S3, Azure-Blob, Huawei, Alibaba OSS, Ceph, IBM, GCP, YIG
- User can migrate immediately as well as can schedule the migration to execute later.
- User have option to delete objects from source after successful migration to destination cloud.

After successful migration

- User can perform object operation and utilize objects of destination bucket for further use.

SODA Releases and Roadmap

Greenland
SODA Foundation release v1.1.0

Faroe
SODA Foundation release v1.0.0

Elba
First SODA Foundation release v0.20.0

Daito (v0.10.1)
This is the v0.10.1 release of OpenSDS

Capri (v0.6.1)
This is the v0.6.1 release of OpenSDS

Bali (v0.4.0)
This is the v0.4.0 release of OpenSDS

Aruba (v0.2.0)
This is the v0.2.0 release of OpenSDS

Zealand (v0.1.0)
This is the v0.1.0 release of OpenSDS

Next

Hawaii

K8S DR | EMC-VMAX | EMC-VNX | IBM-SVC | HDS-VSP | Cloud Archive & Retrieval | Performance Metrics | Anomaly Detection | EDGE Big Data

Next

Edge
Data Energy
AI/ML
Automation
API Standards
Data Protection
Data Lifecycle

Thank You!!

Let's Collaborate!! Stay Tuned…….

sodacon 2020

DATA CONNECTED

#sodacon2020