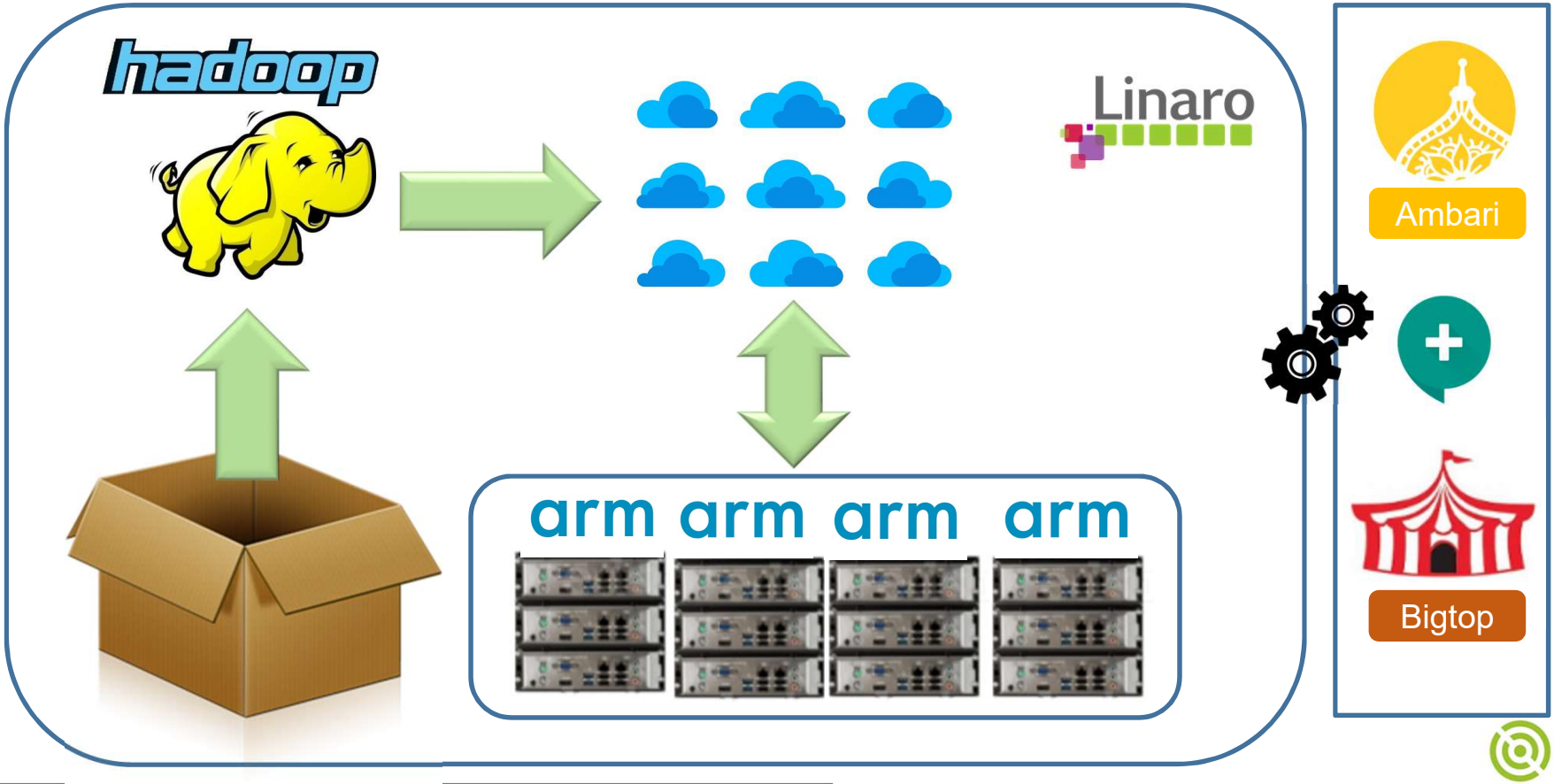
How to lift elephant with Arm

Yuqi Gu
yuqi.gu@arm.com
yuqi.gu@linaro.org

Linaro connect
San Diego 2019

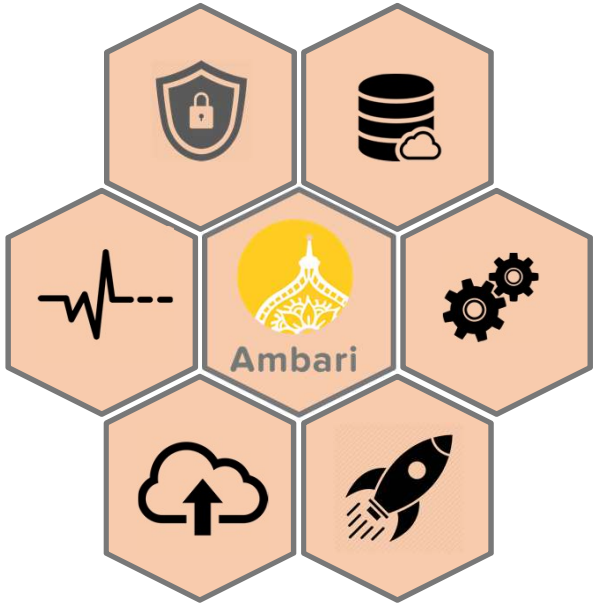
How? Key takeaways



Agenda

- Ambari and Mpack Introduction
- Bigtop Overview
- Bigtop-Mpack Definition
- What We have Done
- Future Work
- Demo

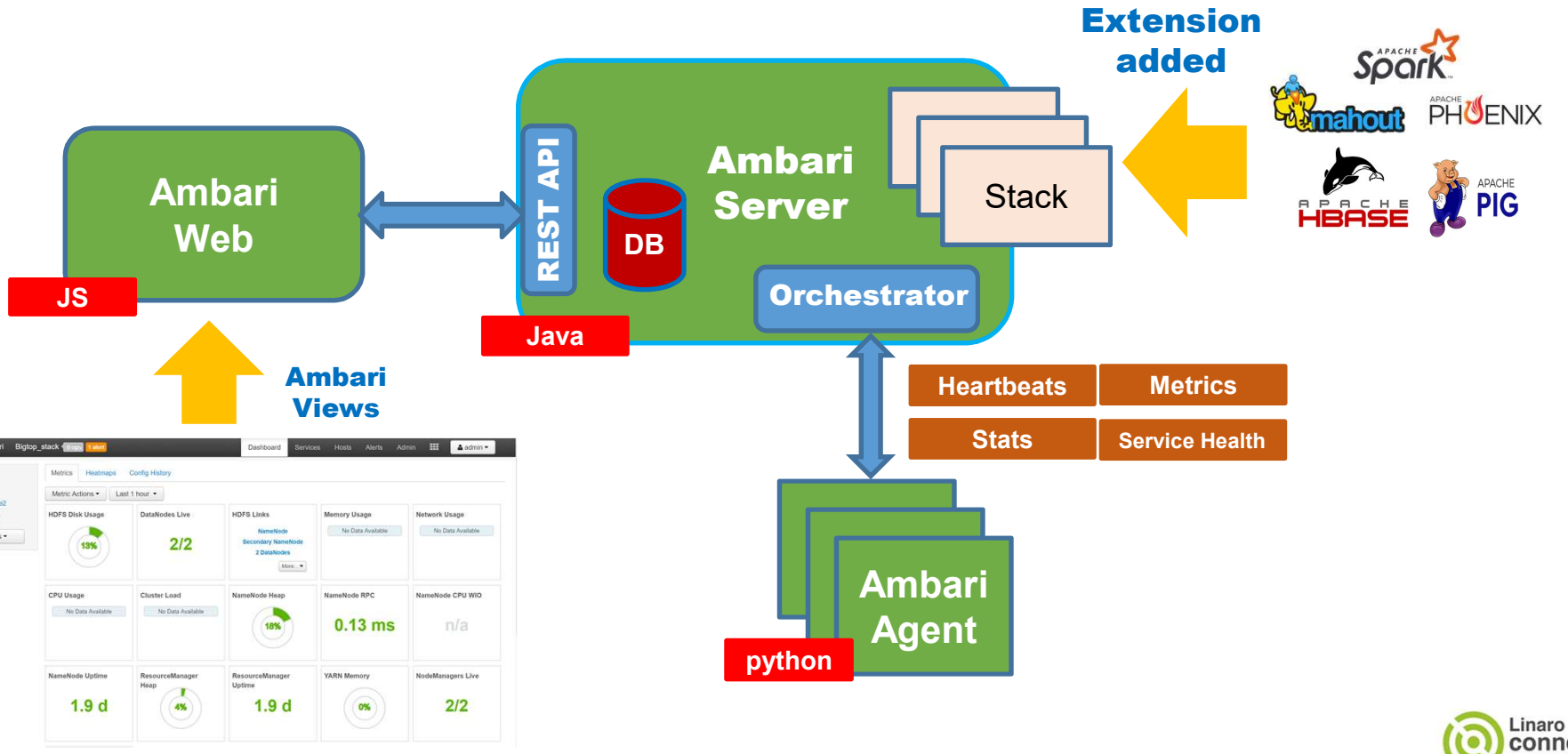
What Is Apache Ambari?



An open-source central **management platform** for **provisioning**, **managing**, **monitoring** and **securing** Hadoop clusters.

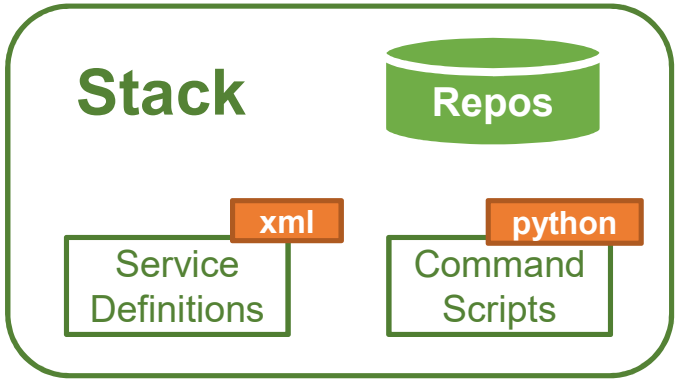
- Simplified Installation, Configuration and Management.
- Centralized Security Setup.
- Highly Extensible and Customizable.

Ambari Architecture



Ambari Stack

- **Stacks define Services + Repos**
  What is in the Stack, and where to get the bits
- **Each Service has a definition**
  What Components are part of the Service
- **Each Service has defined lifecycle commands scripts**
  start, stop, status, install, configure



Limitations

Stack definitions are bundled within Ambari releases

A new version of a stack is released;
Enforce having to do an Ambari release.

There is no release vehicle that can be used to ship add-on services (Custom service).

Linaro connect San Diego 2019

Management Packs (Mpacks)

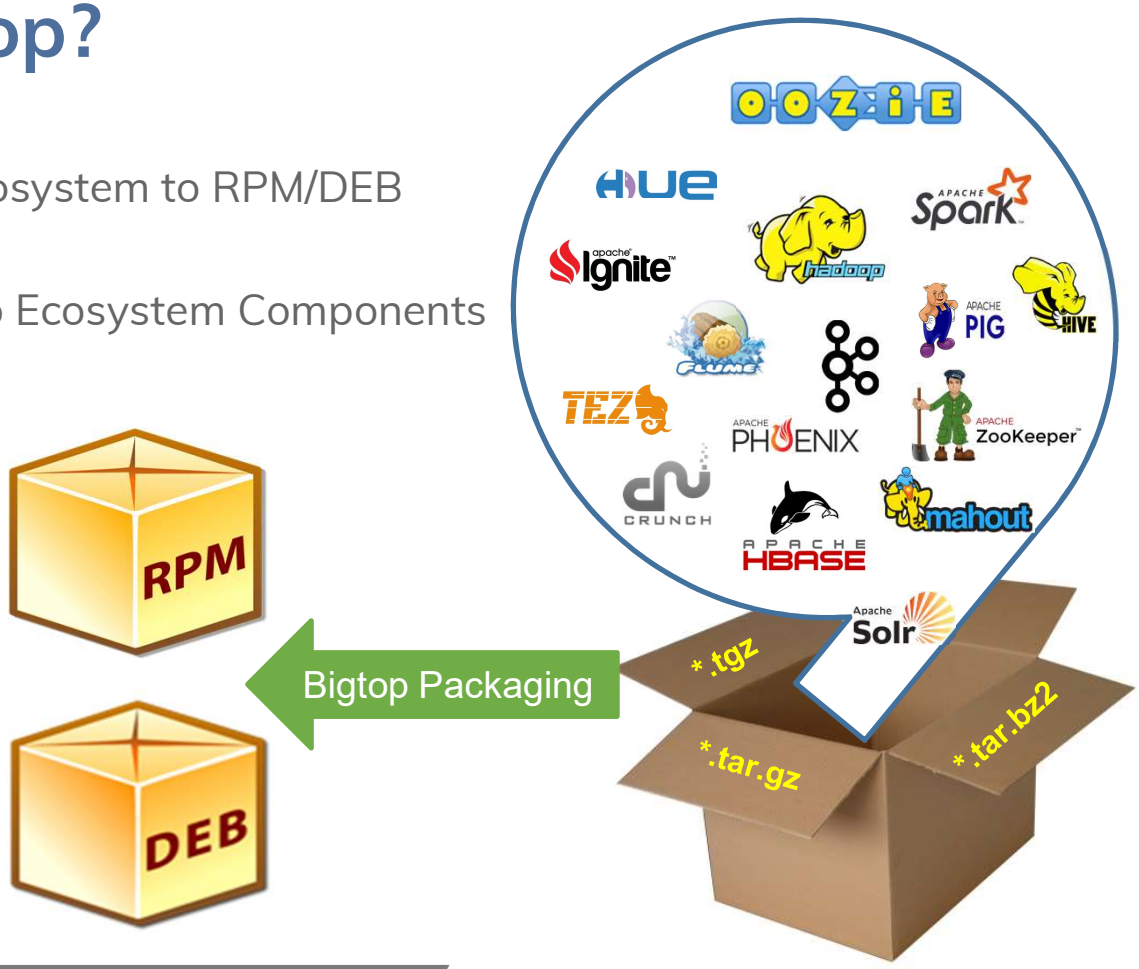
A high-level design can bundle multiple stack definitions, views and Add-On services.

- Improved Release Management:
  - Decoupling stacks from Ambari core
  - Allow stack definition to be a deploy time artifact, not compile time

- Support Add-on Services:
  - Release vehicle for 3rd party services, views
  - Self-contained release artifacts

What is Bigtop?

- Package Hadoop ecosystem to RPM/DEB artifacts
- Support 25+ Hadoop Ecosystem Components

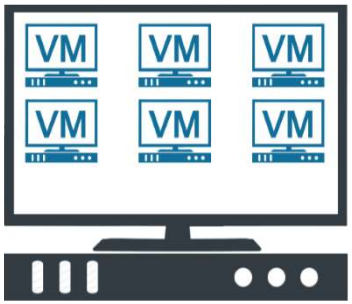


Linaro connect San Diego 2019

Bigtop Feature Snapshot

Packaging

Virtualization



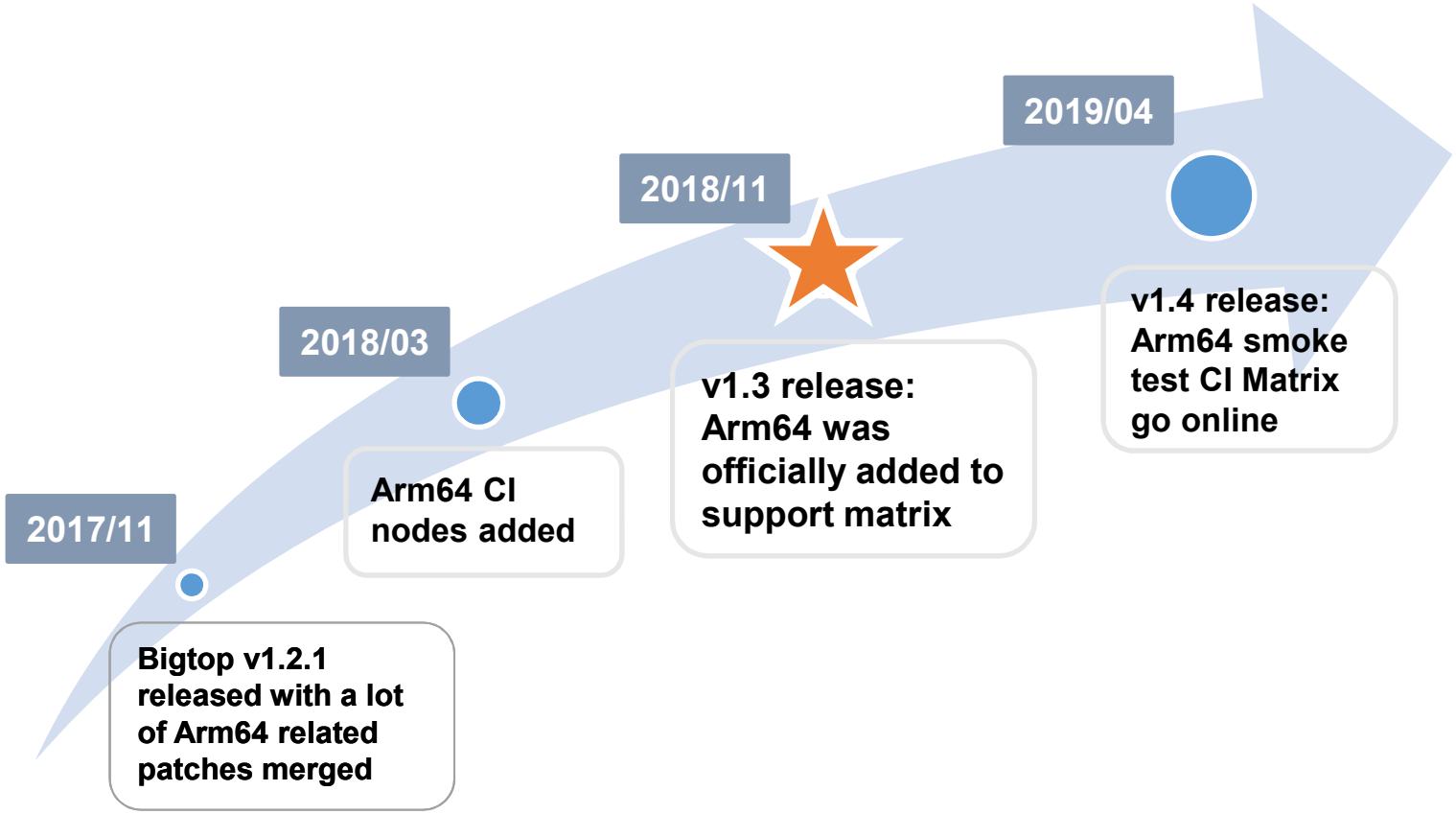
Deployment

Testing

Total solution to build Big Data Stack

arm

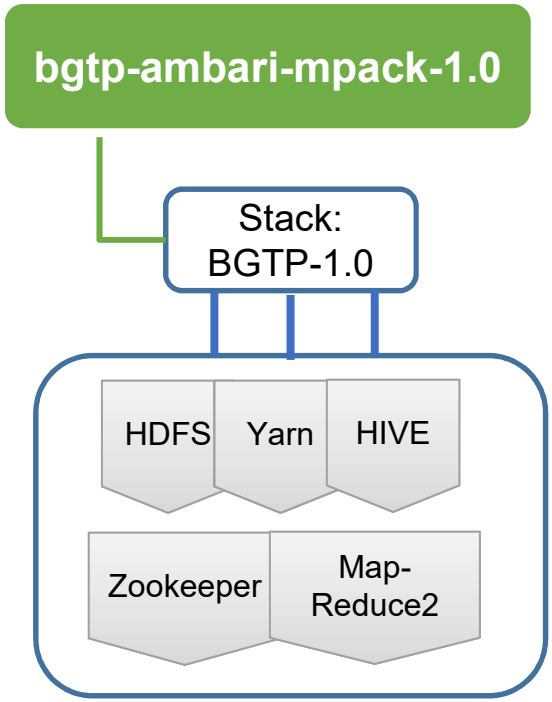
Bigtop on Arm64



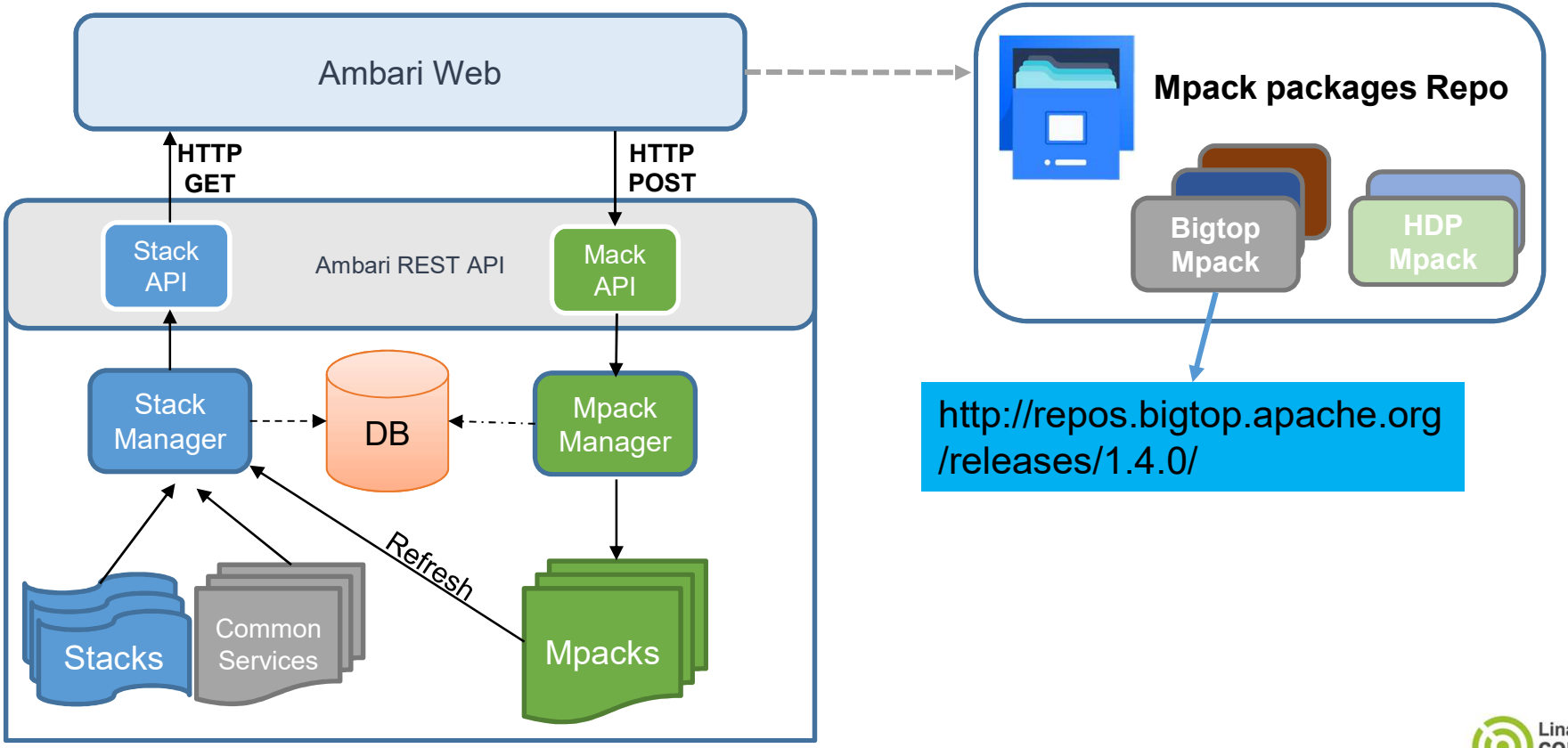
Bigtop Mpack Definition

Bigtop Mpack (BGTP) for Arm64 Servers

BGTP targets to address the complete needs of running big data on Arm64, provide data-at-rest, power real-time customer applications and accelerate decision making on Arm64. (Zookeeper, HDFS, Yarn, MapReduce2)

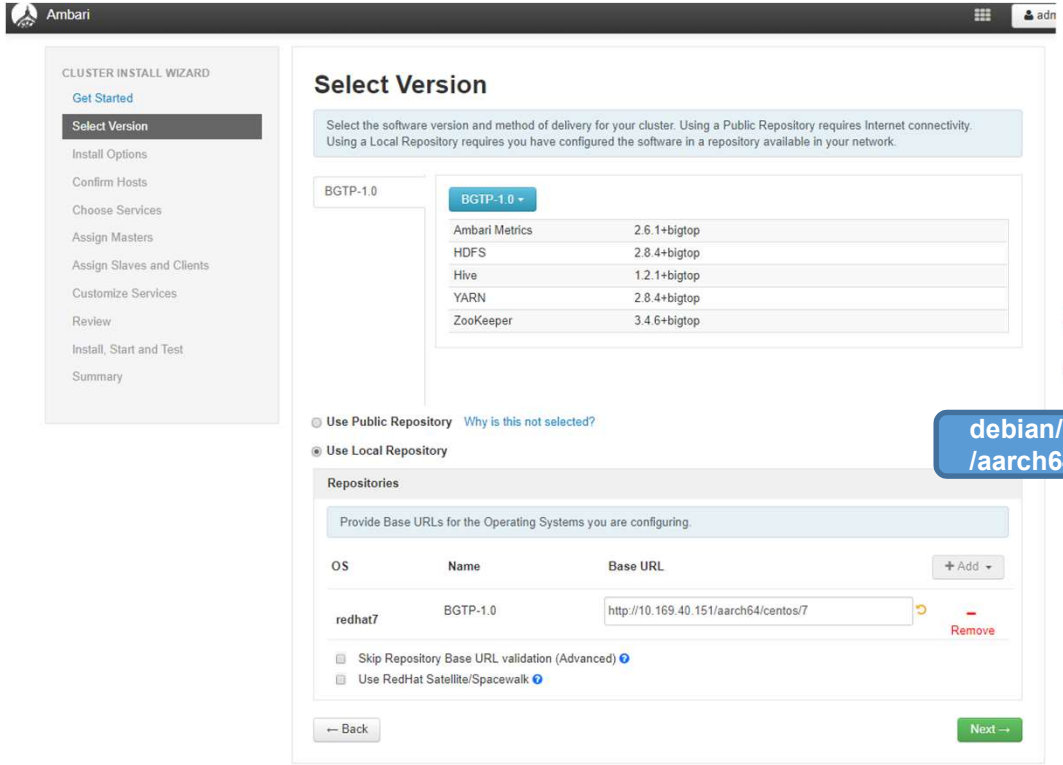


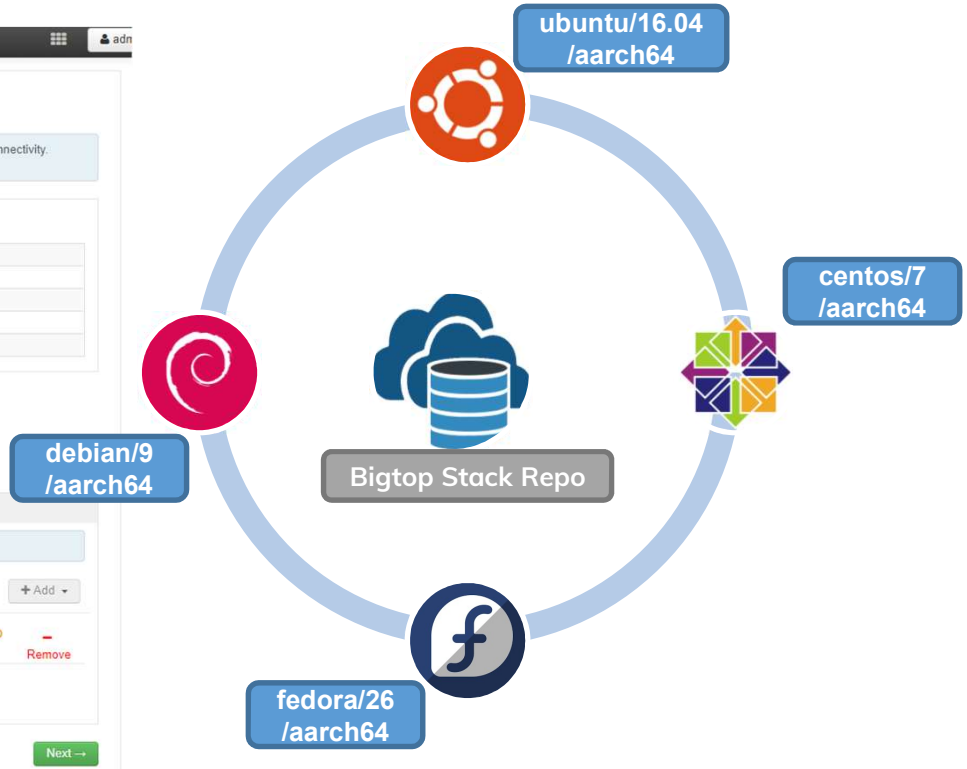
BGTP Mpack Processing Overview



Linaro connect San Diego 2019

BGTP-1.0 Stack Snapshot

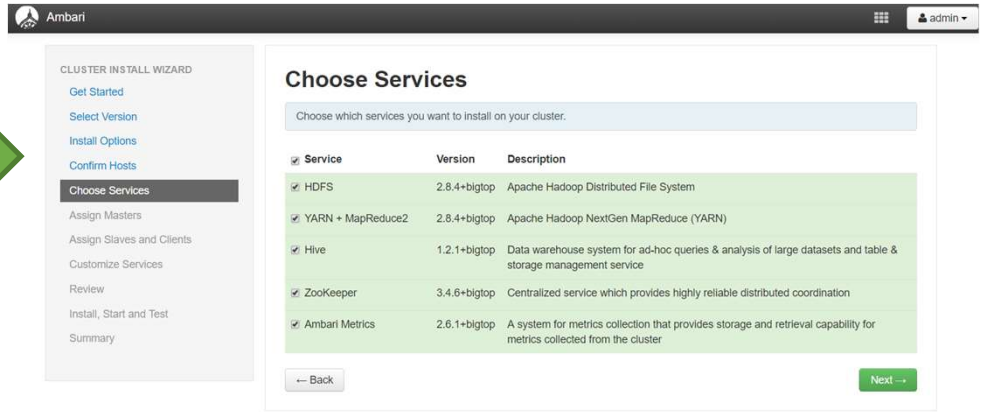




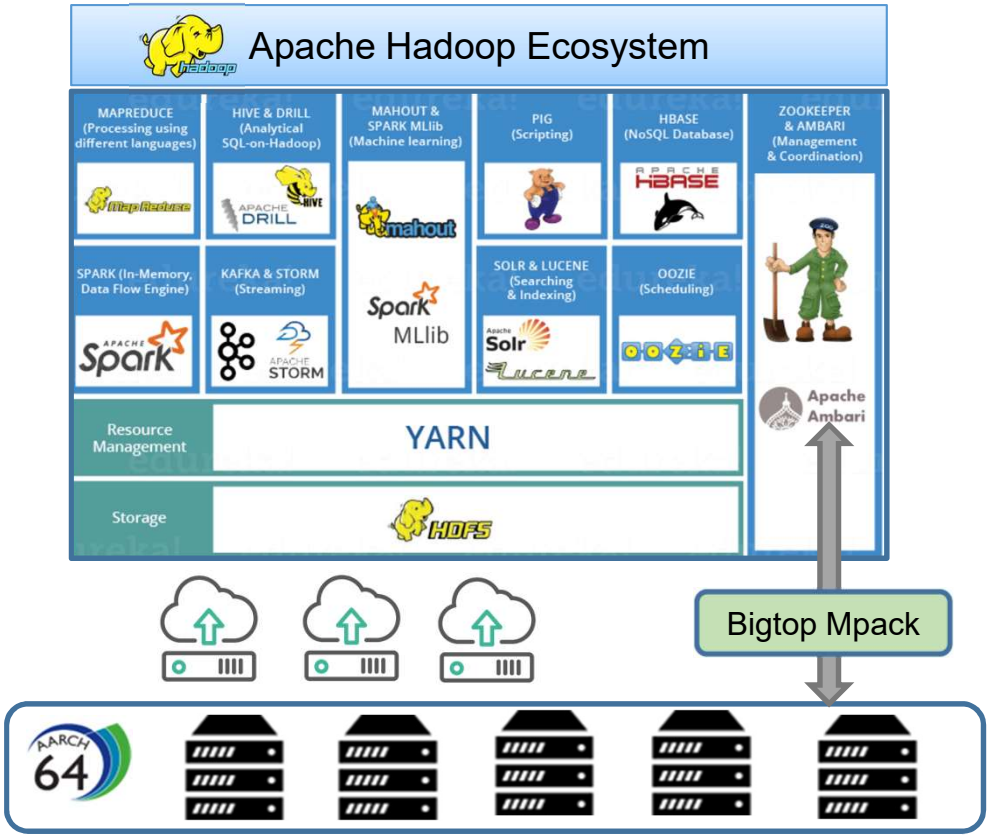
Linaro connect San Diego 2019

Bigtop Mpack Files Structure

```
└── BGTP
    └── 1.0
        ├── blueprints
        ├── configuration
        ├── hooks
        │   ├── after-INSTALL
        │   │   └── scripts
        │   ├── before-ANY
        │   │   ├── files
        │   │   └── scripts
        │   ├── before-INSTALL
        │   │   └── scripts
        │   ├── before-RESTART
        │   │   └── scripts
        │   └── before-START
        │       ├── files
        │       ├── scripts
        │       └── templates
        ├── properties
        ├── repos
        └── services
            ├── AMBARI_METRICS
            ├── HDFS
            ├── HIVE
            │   ├── configuration
            │   ├── etc
            │   └── package
            │       ├── alerts
            │       ├── etc
            │       ├── files
            │       ├── scripts
            │       └── templates
            ├── KERBEROS
            ├── YARN
            │   ├── configuration
            │   ├── configuration-mapred
            │   └── package
            │       ├── alerts
            │       ├── files
            │       ├── scripts
            │       └── templates
            └── ZOOKEEPER
```



BGTP Mpack - Big Picture



Key Takeaways

- BGTP Mpack makes Hadoop-ecosystem deployment and management simpler on Arm64 servers.
- BGTP Mpack is extensively extendable and has an option to customize.

What We have Done

Bigtop:

- Roll out Bigtop 1.3.0 release: https://www.apache.org/dyn/closer.lua/bigtop/bigtop-1.3.0/

  Made Arm officially the first-class citizen with full Arm64 support from the most advanced Big Data stack to date.
- Contribute code to Bigtop and make it happen

  [BIGTOP-2723] [BIGTOP-3088] [BIGTOP-3089] [BIGTOP-3090] [BIGTOP-3091] [BIGTOP-3086]
  [BIGTOP-3135] [BIGTOP-3137] [BIGTOP-3125] [BIGTOP-3145] [BIGTOP-3151] [BIGTOP-3216]
  [BIGTOP-3152] [BIGTOP-3168] [BIGTOP-2858] [BIGTOP-2860] [BIGTOP-3074] [BIGTOP-3092]
  [BIGTOP-3075] [BIGTOP-3077] [BIGTOP-3076] [BIGTOP-3078] [BIGTOP-2101] [BIGTOP-3055]
  [BIGTOP-3068] [BIGTOP-3067] [BIGTOP-2858] [BIGTOP-3064] [BIGTOP-3062] [BIGTOP-3065]

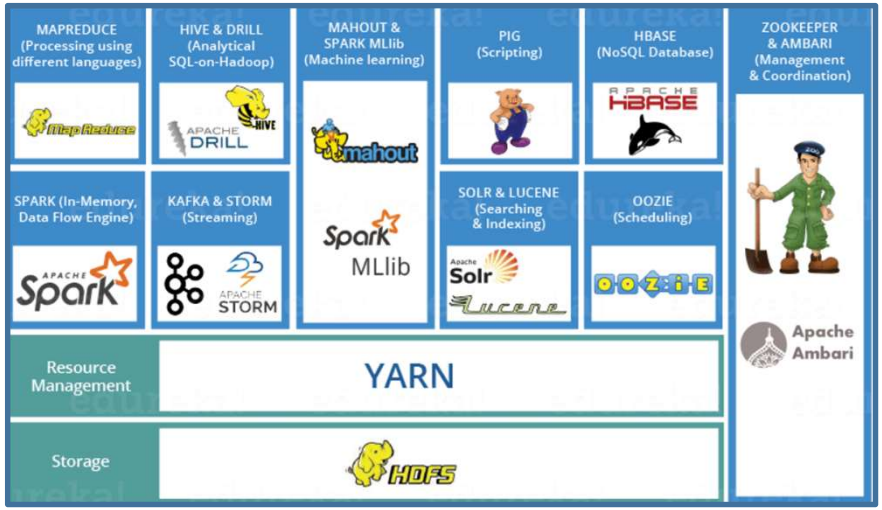
Ambari:

- RFC for adding Bigtop Management Packs:
  https://issues.apache.org/jira/secure/attachment/12967363/RFCs%20for%20Bigtop-mpack.pdf
- Propose the Bigtop mpack definition: AMBARI-25366

  PR: https://github.com/apache/ambari/pull/3073

Future Work

- Add more services in BGTP mpack .
- Upgrade the Ambari dependency from 2.6.1 to 2.7.3.
- Add BGTP-mpack to Bigtop release matrix.



Demo

Questions ?

Contact Us:

Yuqi Gu

yuqi.gu@arm.com

yuqi.gu@linaro.org

Jun He

jun.he@arm.com

jun.he@linaro.org

Thanks to Linaro Team

Ganesh Raju

Naresh Bhat

Guodong Xu

Bigtop Mailing lists

Bigtop CI page

Bigtop Jira issues

Bigtop Wiki

Ambari Development lists

Ambari Wiki

Mpack RFC

Ambari Bigtop-Mpack PR

Linaro Collaborate Page

Thank you

Join Linaro to accelerate deployment of your Arm-based solutions through collaboration

contactus@linaro.org

Linaro connect
San Diego 2019