

DC/OS技术的实践-联通CU-DCOS平台

Mesos practice in China Unicom

中国联通研究院

2017 年 06 月

目 · 录

01

联通IT系统面临的压力

02

CU-DCOS平台

03

技术选型

04

广泛的应用案例



联通面临的压力



资源有效
调度与利用
率均衡
需求

缺乏纯物理机资源
粗颗粒度的调度能力

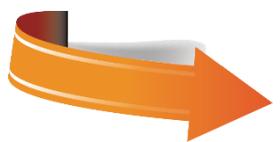
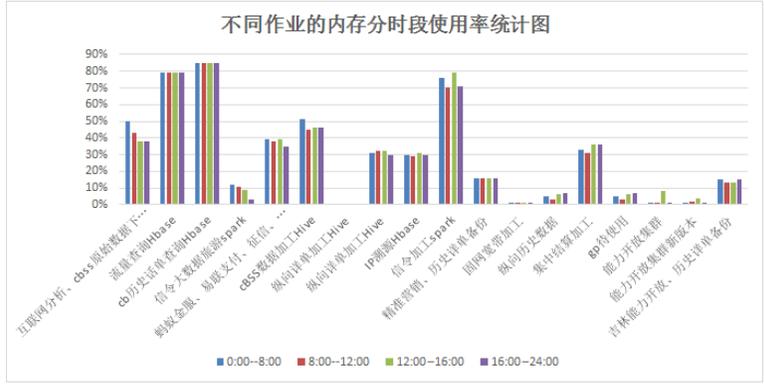
大数据业务的增长
资源相对变得紧致

过度依赖虚拟化技术

私有云可以纳管物理机并自动化部署、可视化
管理，但不能做到自动化
调度物理资源。都缺乏
纯物理机资源粗颗粒
度的调度能力；

现有机房77.4%为纯物
理机。大数据业务的增长
带来资源紧致。

生产系统中3000+物理
节点资源利用现状：
大规模集群间资源使用率
最多相差75% 不同时段间
同一集群资源使用率最多
相差67%



- 不同应用资源
分配不均衡
- 不同时段资源
利用不均衡

目 · 录

01

联通IT系统面临的压力

02

CU-DCOS平台

03

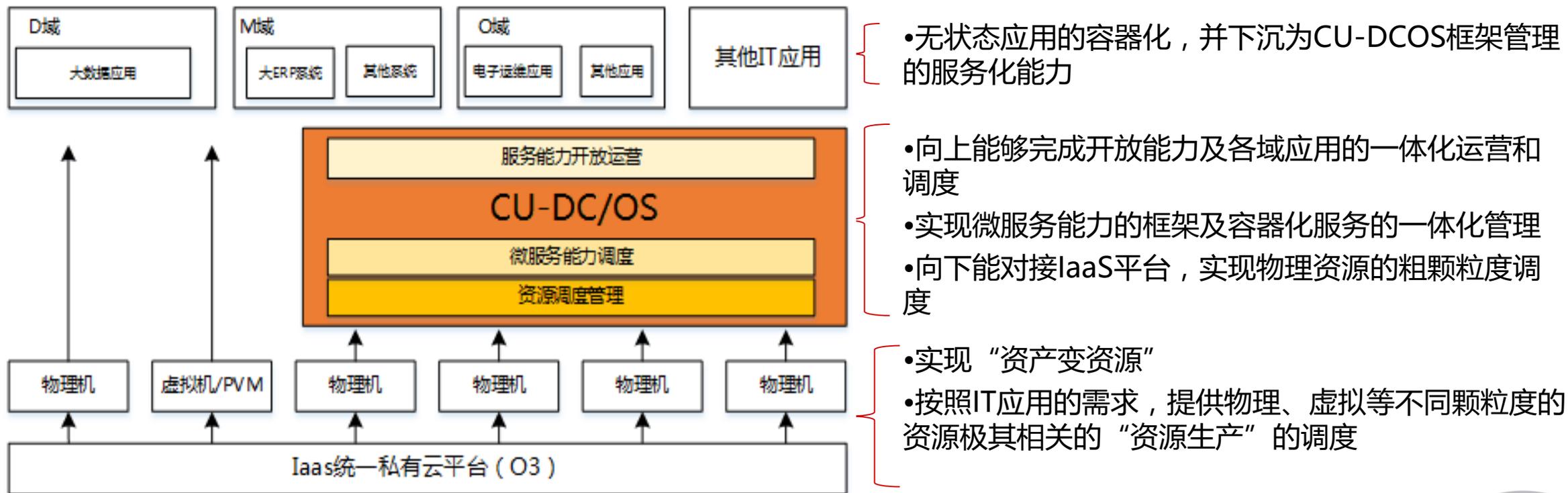
技术选型

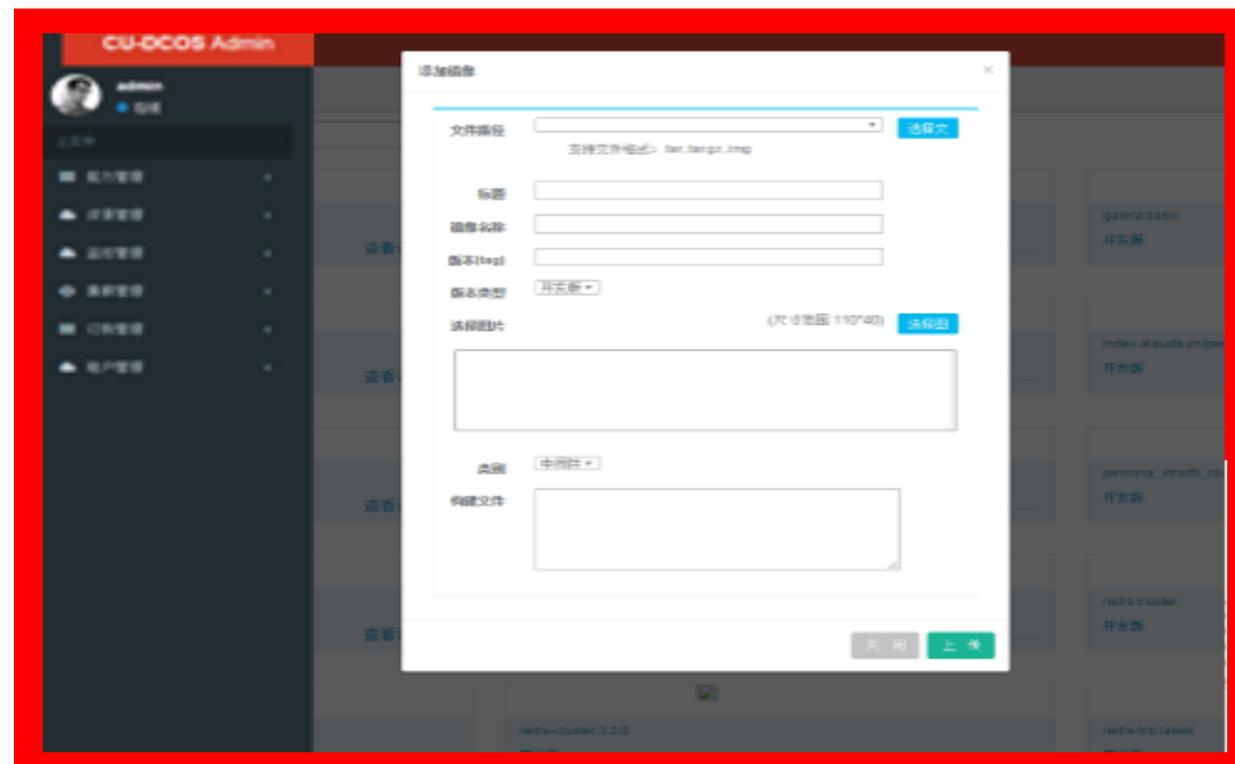
04

广泛的应用案例

来自联通研究院的CU-DCOS

完成IT资源的一体化调度；提供大数据微服务化能力管理、调度和开放化运营的管理框架；承载各IT域无状态应用的容器化微服务能力，开展基于CU-DCOS能力平台的应用的DevOps。





CU-DCOS是中国联通基于mesos开源系统自主研发的未来数据中心操作系统，实现资源的一体化调度、大数据微服务化管理能力，完成IT资源的集中管理；

容器化承载各IT域无状态应用的微服务能力，集成多个容器化大数据能力，支撑基于CU-DCOS能力平台的应用的DevOps。

□ CU-DCOS作为新一代的大数据Pass平台，面向联通“**两地三中心**”的大数据中心提供大数据、数据库以及容器化的服务能力。

跨中心实现统一管理

- 便捷的集群自动化扩缩容

多样化的能力服务

- 面向大数据提供数十种大数据服务能力，同时提供Mysql、Oracle等关系型数据库服务能力；

容器化的敏捷开发

- 丰富的容器应用私有镜像，支持Devops开发，大数据业务快速上线；

便捷的运营管理

- 自服务和管理两个入口；
- 业务运营和后台管理全部采用可视化操作

标准化的资源管理

- 基于mesos技术实现，厚重框架之间的资源调度。
- 基于kubernetes技术实现docker的统一资源管理

高弹性的资源伸缩

- 基于容器技术实现特殊业务需求场景下的高可靠性伸缩
- 应对新增业务上线造成的系统运行压力



CU-DCOS | 服务能力

应用
服务
(10种)



Wordpress



区块链



DokuWiki

Confluence

ETL



Alfresco



Jenkins



TensorFlow

数据库
能力
(5种)



redis



ORACLE



mongoDB



etcd

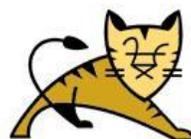


MySQL

中间件
(3种)



NGINX



Apache Tomcat



kafka

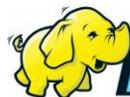
大数据
能力
(7种)



Apache Zookeeper



Spark



hadoop



HIVE



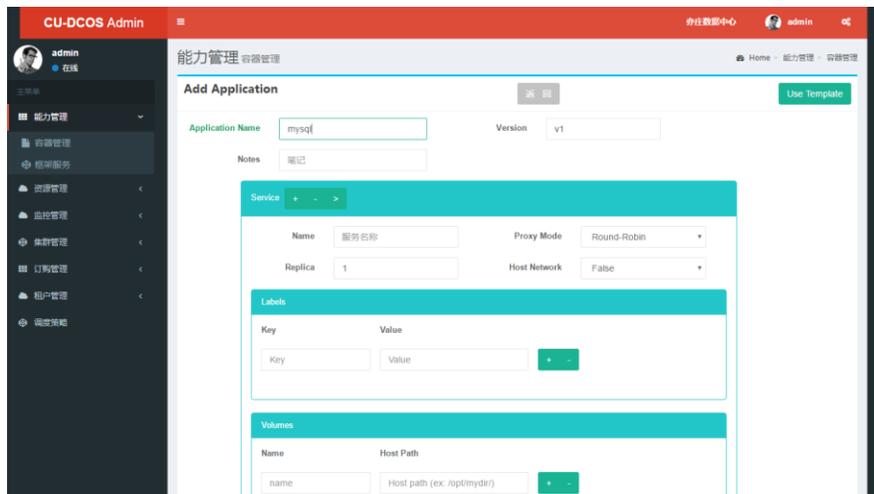
Cassandra



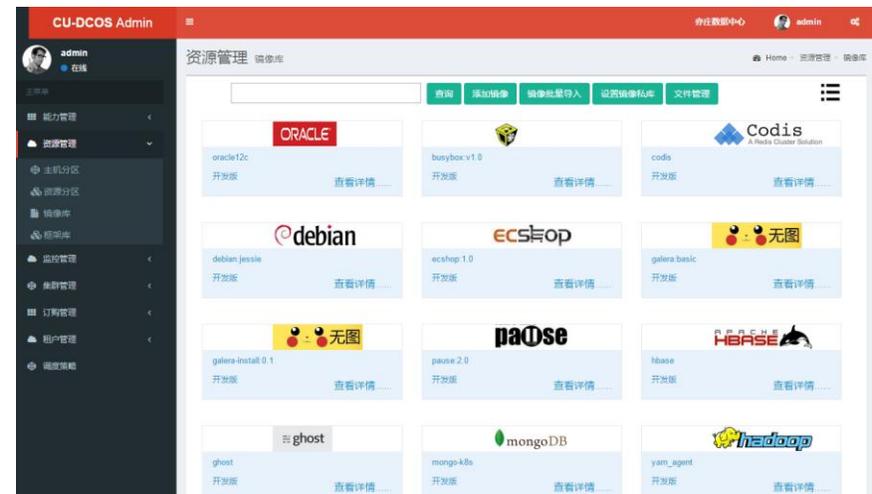
APACHE STORM™
Distributed · Resilient · Real-time



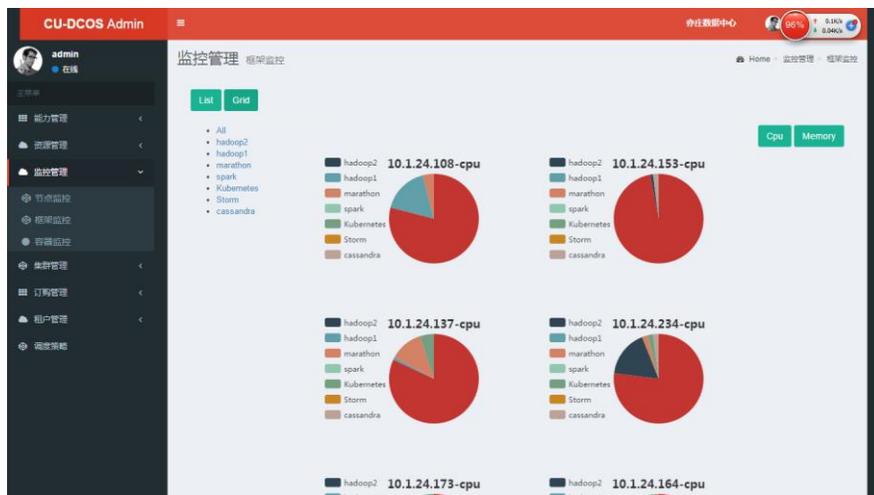
HDFS



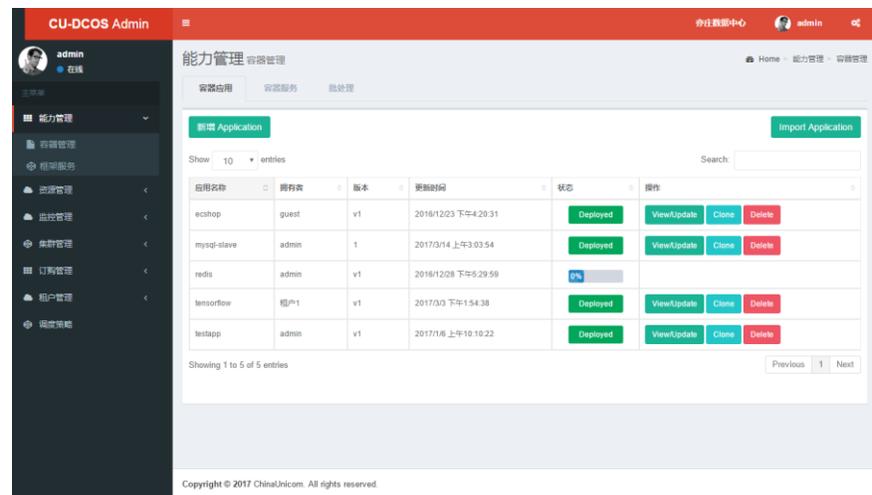
容器管理



镜像管理



资源监控



The screenshot shows the 'Ability Management' page, displaying a table of deployed applications. The table includes columns for application name, owner, version, update time, status, and actions.

应用名称	拥有者	版本	更新时间	状态	操作
ecshop	guest	v1	2016/12/23 下午4:20:31	Deployed	View/Update Clone Delete
mysql-slave	admin	1	2017/3/14 上午9:03:54	Deployed	View/Update Clone Delete
redis	admin	v1	2016/12/28 下午5:29:59	0%	
tensorflow	租户1	v1	2017/3/3 下午1:54:38	Deployed	View/Update Clone Delete
testapp	admin	v1	2017/1/6 上午10:10:22	Deployed	View/Update Clone Delete

能力管理

目 · 录

01

联通IT系统面临的压力

02

CU-DCOS平台

03

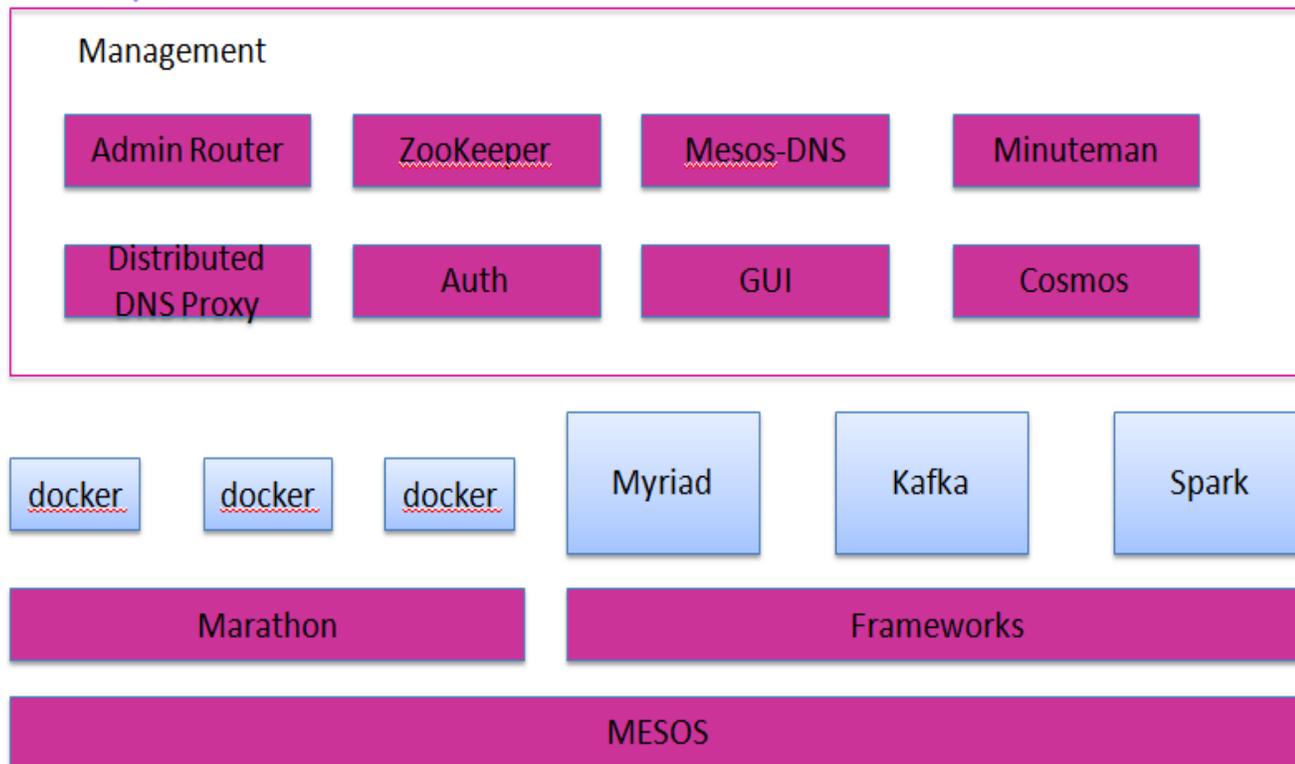
技术选型

04

广泛的应用案例

CU-DCOS | 技术选型Mesosphere DC/OS

Open DC/OS



Open DC/OS 是mesosphere2016年上旬发布的企业数据中心操作系统，主要是围绕开源的Apache Mesos附加了一些面向企业级应用的管理功能，核心内容还是两支，一边是面向微服务治理的marathon，一边是通过framework来管理大数据类的一些应用，其附加功能：

Admin Router：一个反向代理用于引导用户请求到相应的服务，服务引流的功能。

Zookeeper：或者是Exhibitor，为DCOS其上的分布式系统提供协调服务

Mesos-DNS：服务注册、服务发现

Minuteman：内部网络第四层负载均衡

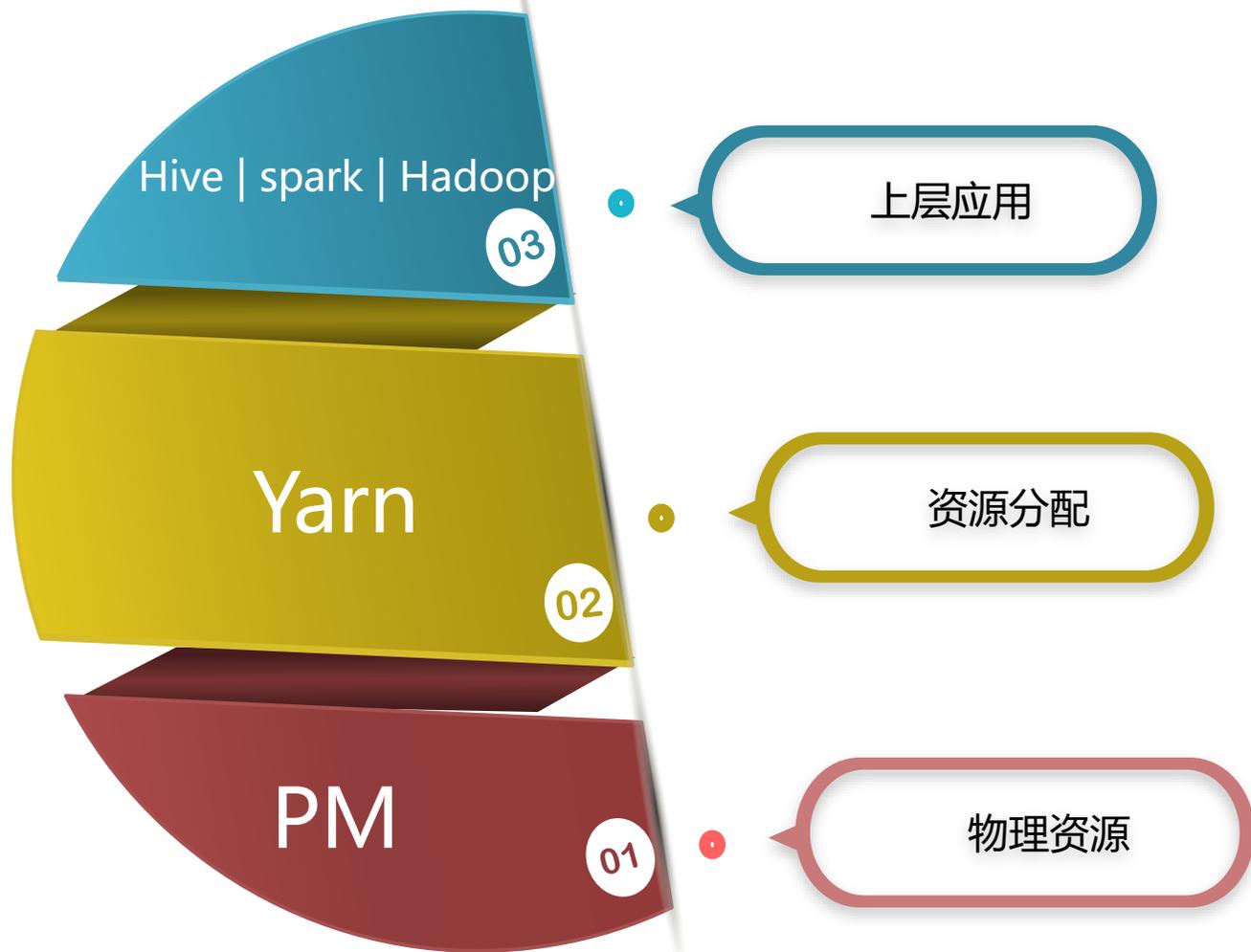
Distributed DNS Proxy：分布式DNS代理

Auth：用户权限管理

GUI：图形化的为监控、配置界面

Cosmos：框架安装包私库

Myriad | 为什么选择它？

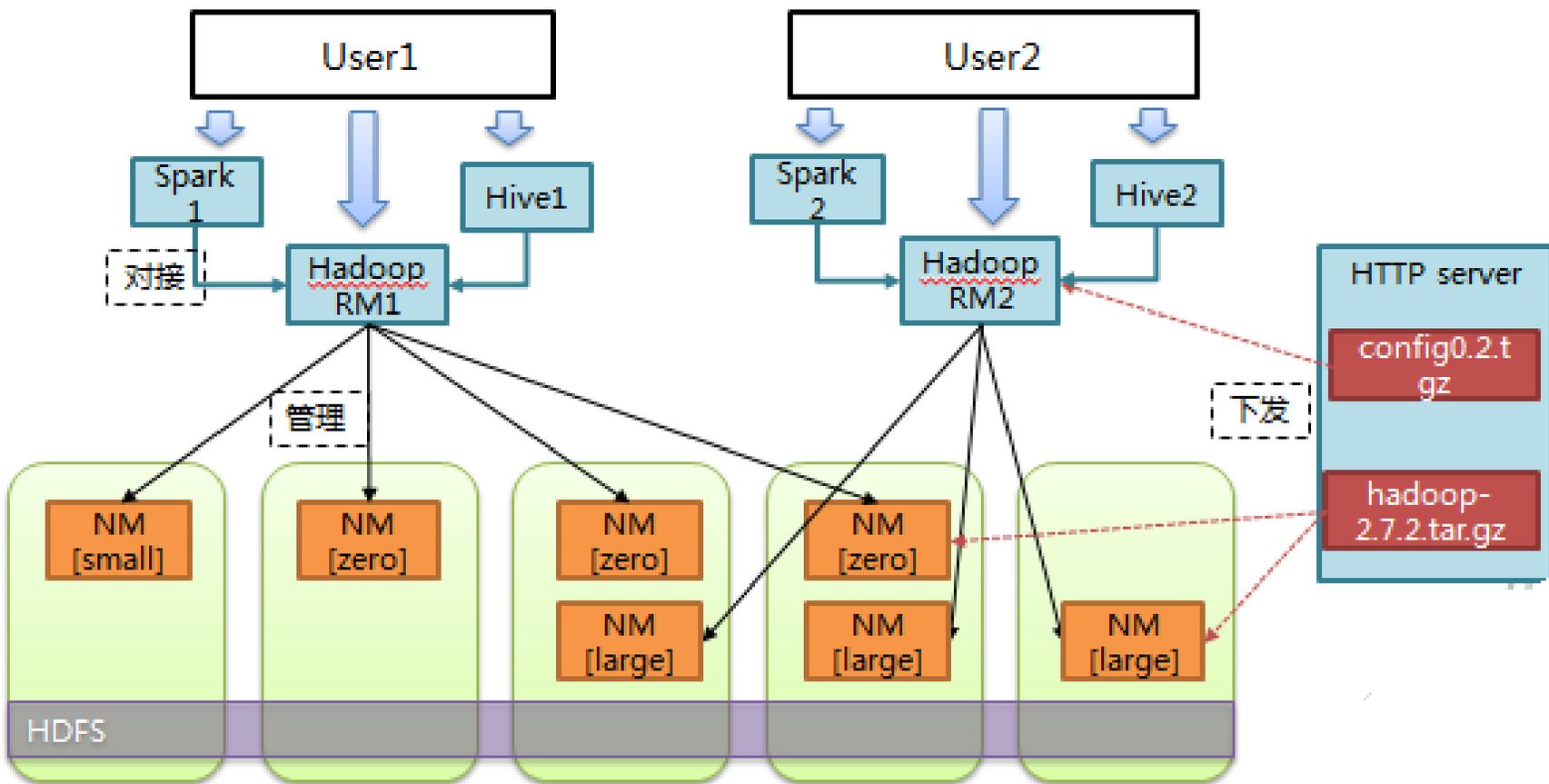


打通原有资源隔离，构建统一的大数据和容器化服务资源管理调度平台

- 针对原有情况进行转变

针对现有系统yarn体系，进行升级。保留现有的大数据任务，对现网影响降到最小

自动部署的Hadoop



01

配置文件

配置文件存储在服务器端，以HTTPserver的方式下发

02

构建时下发

创建集群时下发配置，无需事先部署hadoop

03

HDFS部署位置

HDFS是不能轻易伸缩的，尤其是不能缩。大数据集群计算节点需要相对固定在HDFS节点，预留节点

Myriad on mesos带来的好处

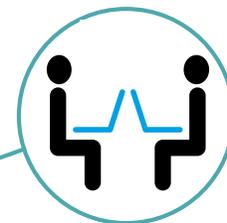
资源配额

保留资源允许框架和集群操作员在需要时确保有足够的资源可用。通常进行预订以保证在峰值负载下有足够的资源。



资源抢占/归还

集群资源空闲时，资源告急集群瞬间吸收空闲资源，被吸收集群需要时，自动释放资源进行归还



动态预留

为关键工作负载预留资源，同时支持资源的超额预订和动态预留

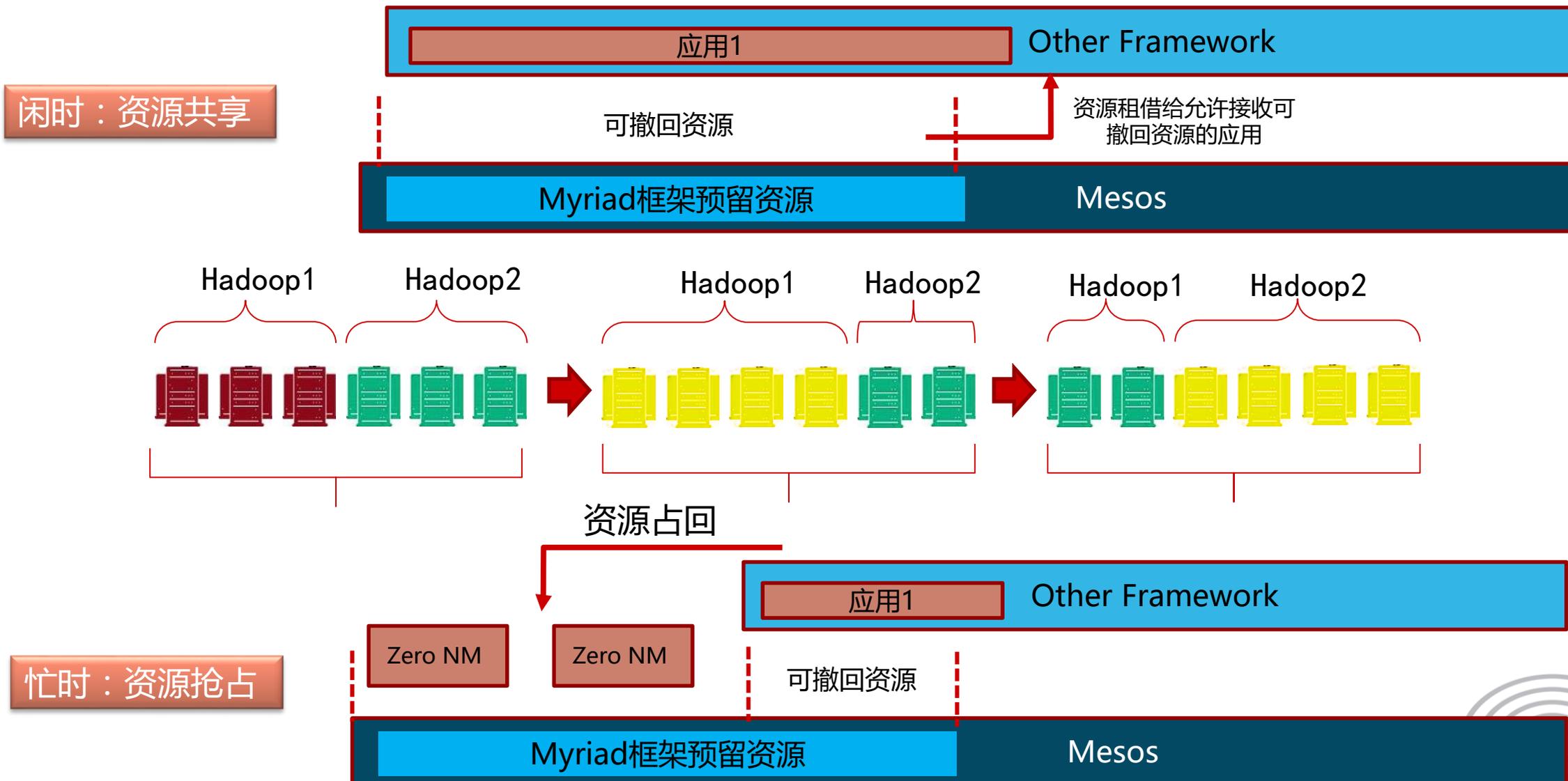


粗/细粒度调度

大数据不同进程CPU，内存使用多样化调度。不同的大数据作业和资源有效匹配。

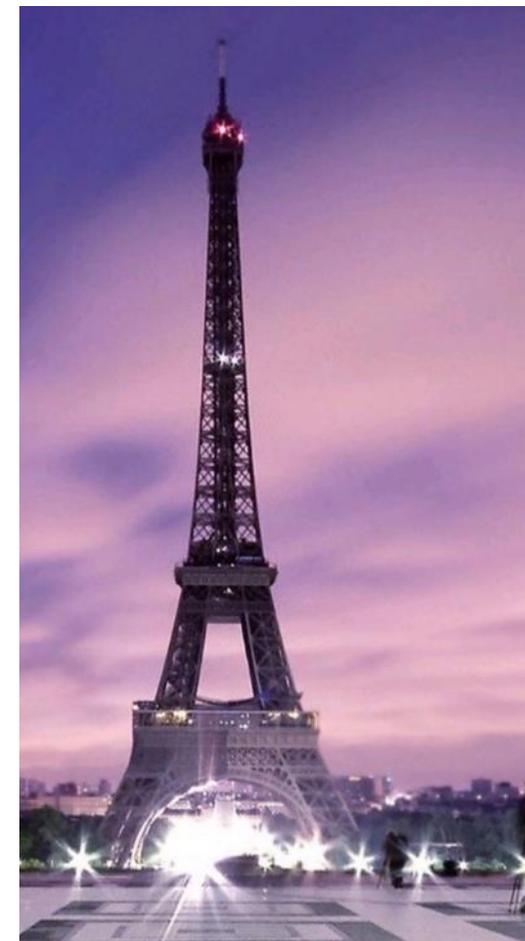
Hadoop on mesos: 盘活物理资源

- 利用细颗粒度调度和事件性、周期性调度策略解决数据中心大数据资源调度、共享问题



Myriad 的不足与缺陷

序号	描述
1	为保证生产机群计算安全对于配置环境采用多个Resource Manager的方式，Myriad框架目前暂时无法支持。
2	Myriad框架对运行Node Manager的机器列表可以进行限制策略，但必须设定一个固定的实例，且第一个启动无法限制部署机器。
3	Myriad框架对内存、CPU等资源需要在配置文件预先进行配置，即使设置zero模式，也必须启动一个固定的实例才行。
4	Myraid框架对hadoop版本支持有限制，目前只能支持hadoop2.6以上版本。
5	测试中发现myraid框架不稳定的情况，偶尔出现提交任务不报错但也不执行的情况。
6	Myriad框架代码与hadoop代码捆绑太紧，编译时对版本的要求导致支持的版本有限制，应该实现在hadoop外层进行包装，只负责启动hadoop服务。



Mesos 多租户与资源的问题



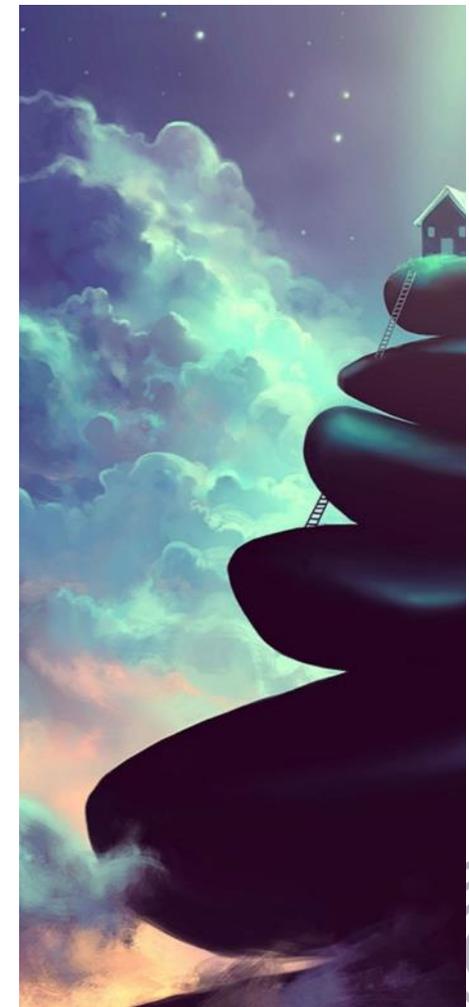
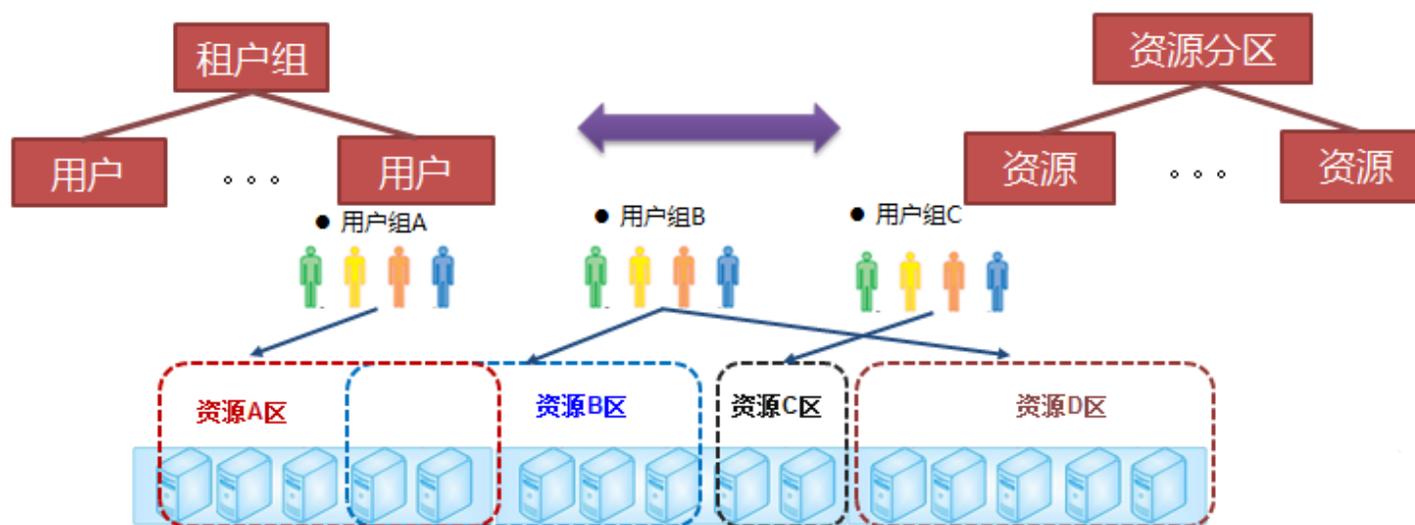
DCOS能力平台通过对主机的物理划分和对用户组权限的逻辑划分，实现了多租户的能力。实现了对每个用户所拥有的权限进行自定义。

资源和租户分离管理

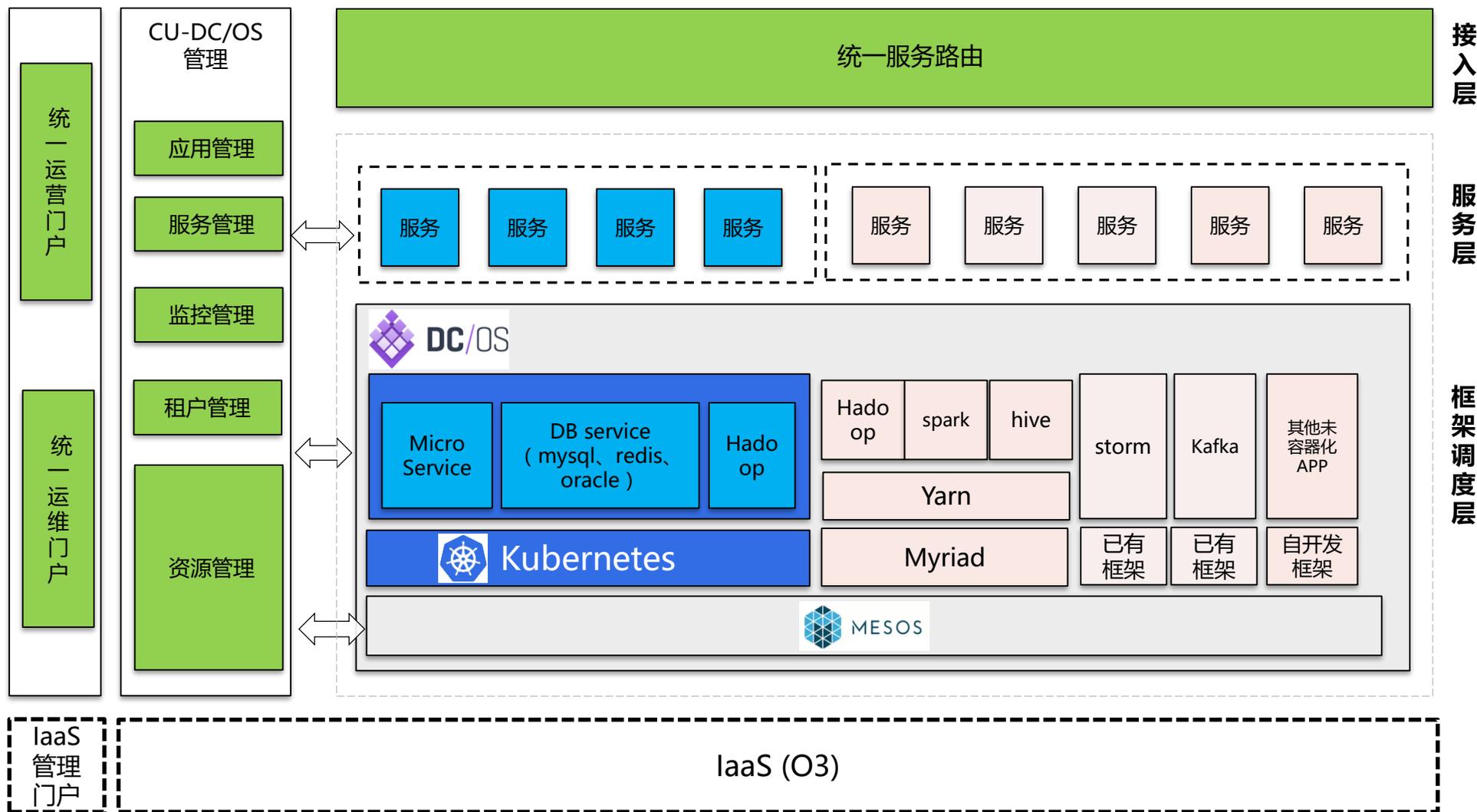


资源和租户相互匹配

用户之间在共享主机资源的基础上实现了业务层面的隔离，使用时匹配，既保证了数据的相对独立，安全，又解决了资源均衡的问题。



CU-DCOS | 技术架构



接入层

服务层

框架调度层

接入层

- ✓ 提供统一的服务接入能力；
- ✓ 实现统一的服务路由能力；
- ✓ 提供应用负载均衡能力；
- ✓ 提供自服务和后台管理

服务层

- ✓ 基于容器编排框架实现应用的容器化封装；
- ✓ 实现大数据能力提供服务；
- ✓ 实现Oracle、MPP等数据库能力服务

框架调度层

- ✓ 实现Kubernetes框架在Mesos承载；
- ✓ 实现Myriad在Mesos承载；
- ✓ 利用已有框架和自主研发框架，实现storm、Kafka等在mesos承载

目 · 录

01

联通IT系统面临的压力

02

CU-DCOS平台

03

技术选型

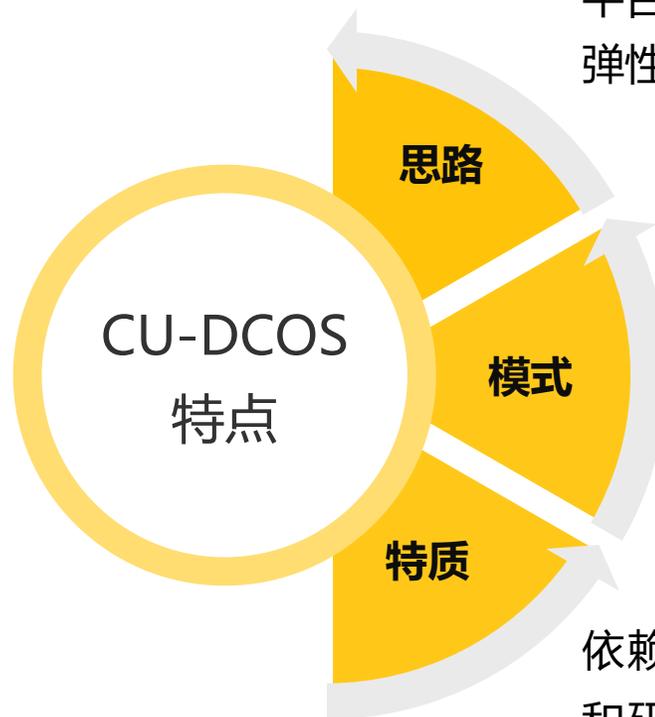
04

广泛的应用案例

联通PaaS平台的升级改造

联通PaaS平台作为联通资源集中运维管理系统，具有虚拟化资源运维管理的能力，**缺乏微服务能力化，标准化运营，高弹性的业务伸缩，峰值应对等能力。**

平台自身能力容器化，依靠CU-DCOS实现弹性伸缩，统一运营的能力。

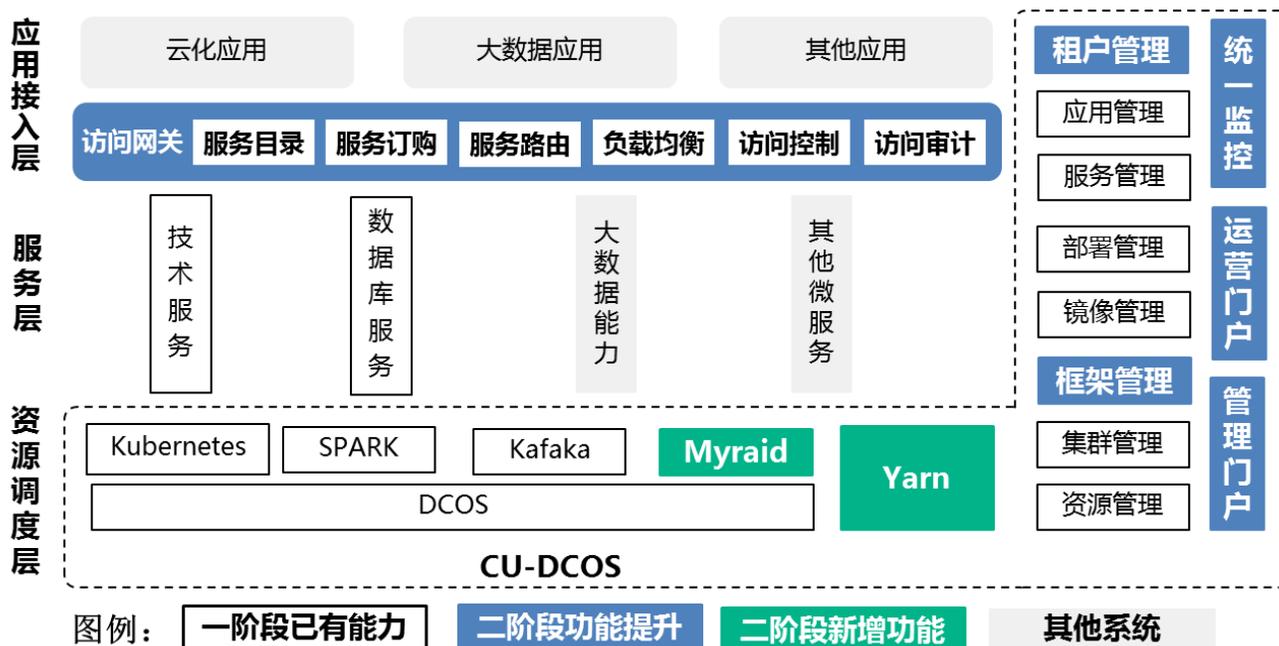


应用框架的容器化技术在生产运维过程中进行快速部署
通过统一微服务网关实现多地一体化开放运营

依赖于开源软件，通过集成和研发，具备生产运营的条件。

联通PaaS平台与CU-DCOS的融合

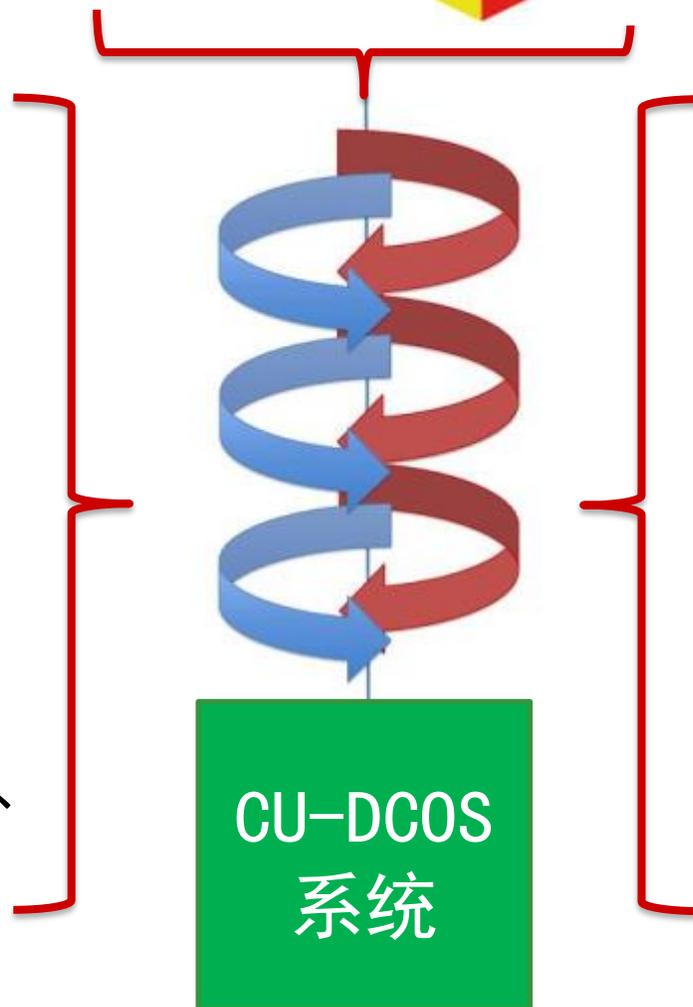
- 联通PaaS平台通过借助CU-DCOS平台的能力，进行升级融合。把**缓存，消息，日志，任务，通知，4A，数据库，文件，报表，搜索**等能力，进行解耦合，把能力容器化微服务化，通过订阅能力，进行更为高效的集中运维管理
- 相对于传统平台，与CU-DCOS融合的联通PaaS平台可以实现传统业务的**秒级支撑，多样化能力提供，对于大数据能力的强力支撑**



联通牛人部落&联通DAS大数据孵化平台的实践



- ◆ 多样化的大数据能力提供
 - ✓ 1 提供hadoop、hive、spark等主流能力
 - ✓ 2 解决大数据不同进程CPU，内存使用不均衡的问题。不同的大数据作业和资源有效匹配。
- ◆ 平台本身的容器化快速、灵活、轻量级部署



- ◆ 解决大数据应用自动化DevOps能力
 - ✓ 快速迭代开发，灰度发布，滚动升级
- ◆ 实现能力下沉、运营集中的PaaS服务能力

CUDCOS | 产生的成果

目前已完成了全部**15种PaaS能力的封装**，可对外提供服务。
PaaS平台上的**数十个O域、M域应用已经完成CU-DCOS整体迁移**，并且运行稳定

中国联通公共创新大数据能力开放平台

该平台提供大数据、数据库以及容器化的服务能力，已成为公司**开源技术与业务转型相结合的创新型示范项目**，在资源利用率提升的同时，节省建设及运营成本，降低能耗。相比传统的基于虚拟机、物理机分配、申请、部署、测试，**基于容器软件部署时间节省了80%以上**，**容器部署均为自动化驱动**，可靠性大幅度提升。

中国联通PaaS平台

CU-DCOS

牛人部落

支撑中国联通牛人部落实验室的基础架构，目前已应用于**100多台IT设备并利用CU-DCOS进行统一的部署和资源调度**。实现了静默安装、可视化配置、自定义配置等多种部署方案；支持服务实例的实时弹性扩展和动态部署。成功搭建了开放式实验环境，满足中国联通IT实验室的需要。

谢谢