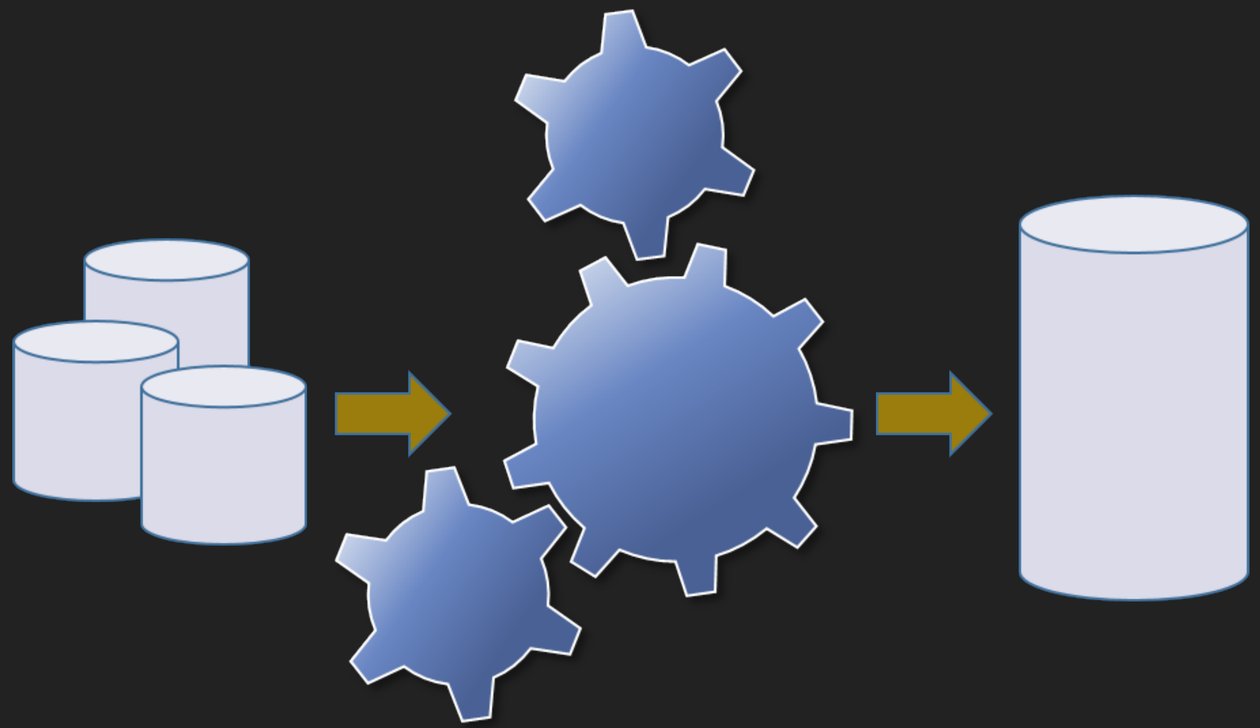
饿了么 | IT大咖说 不止于技术

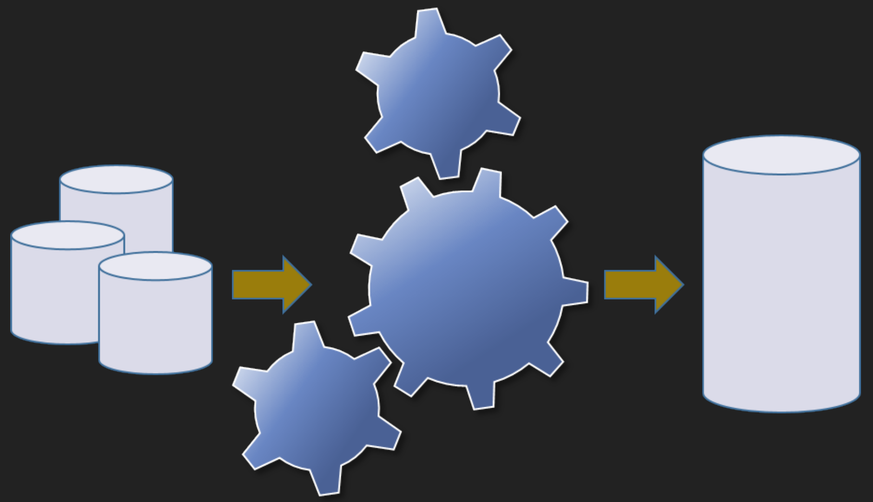
如何构建作业云

ELASTIC JOB CLOUD

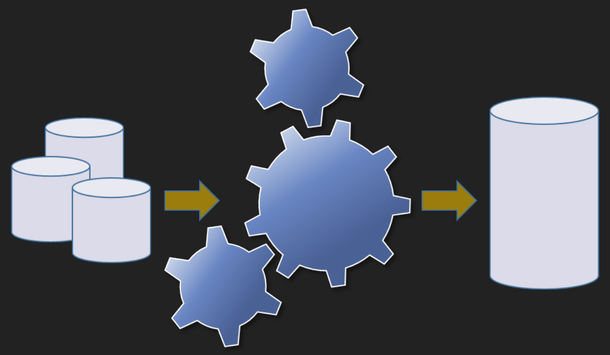
作业简史



Extract → Transform → Load



Extract → Transform → Load


Extract → Transform → Load

Crontab

Crontab

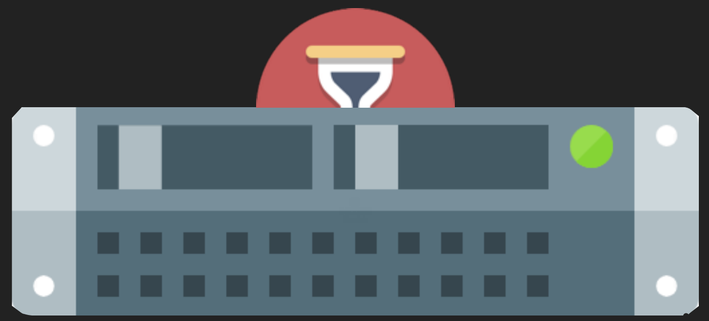
Elastic-Job

Crontab

QUARTZ

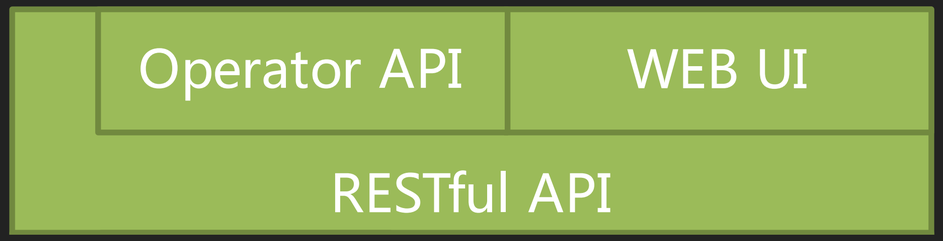
弹性的本质

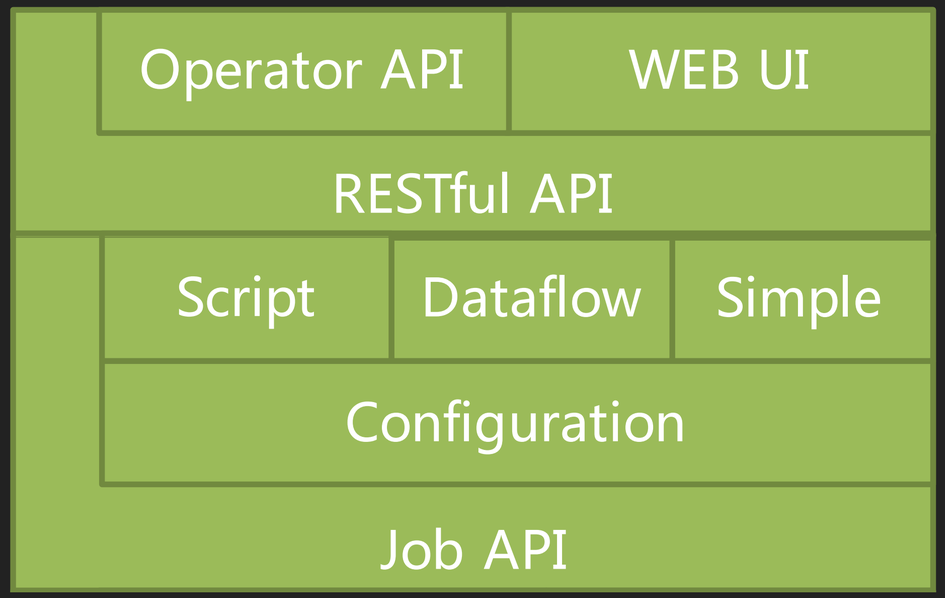
Scale-out

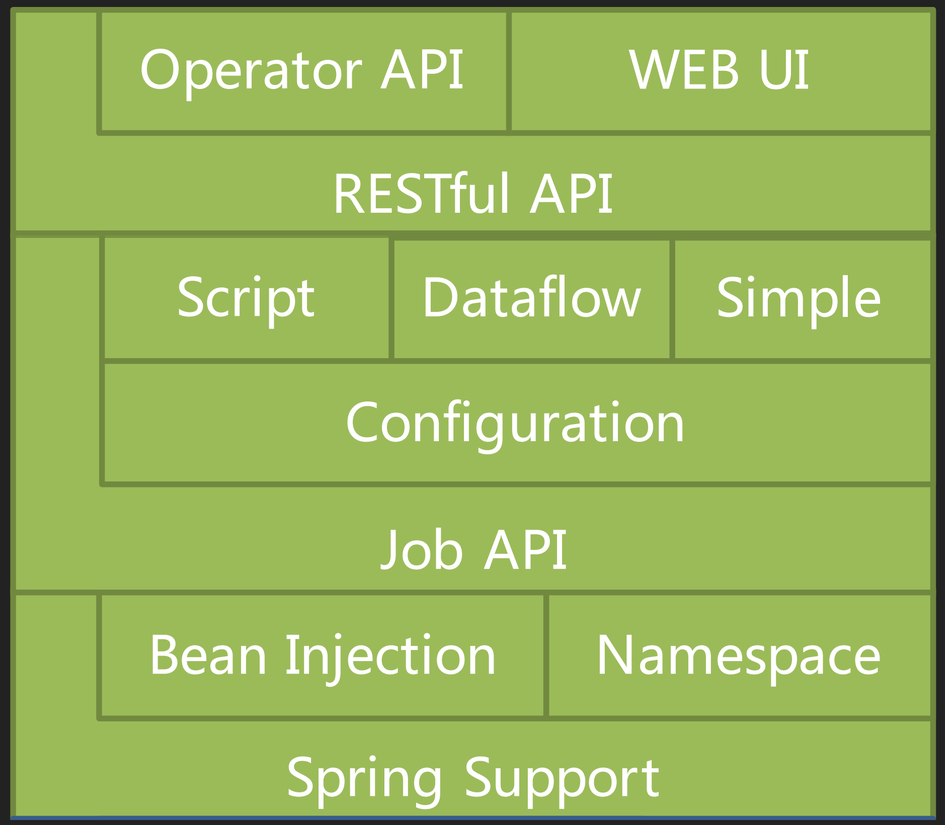


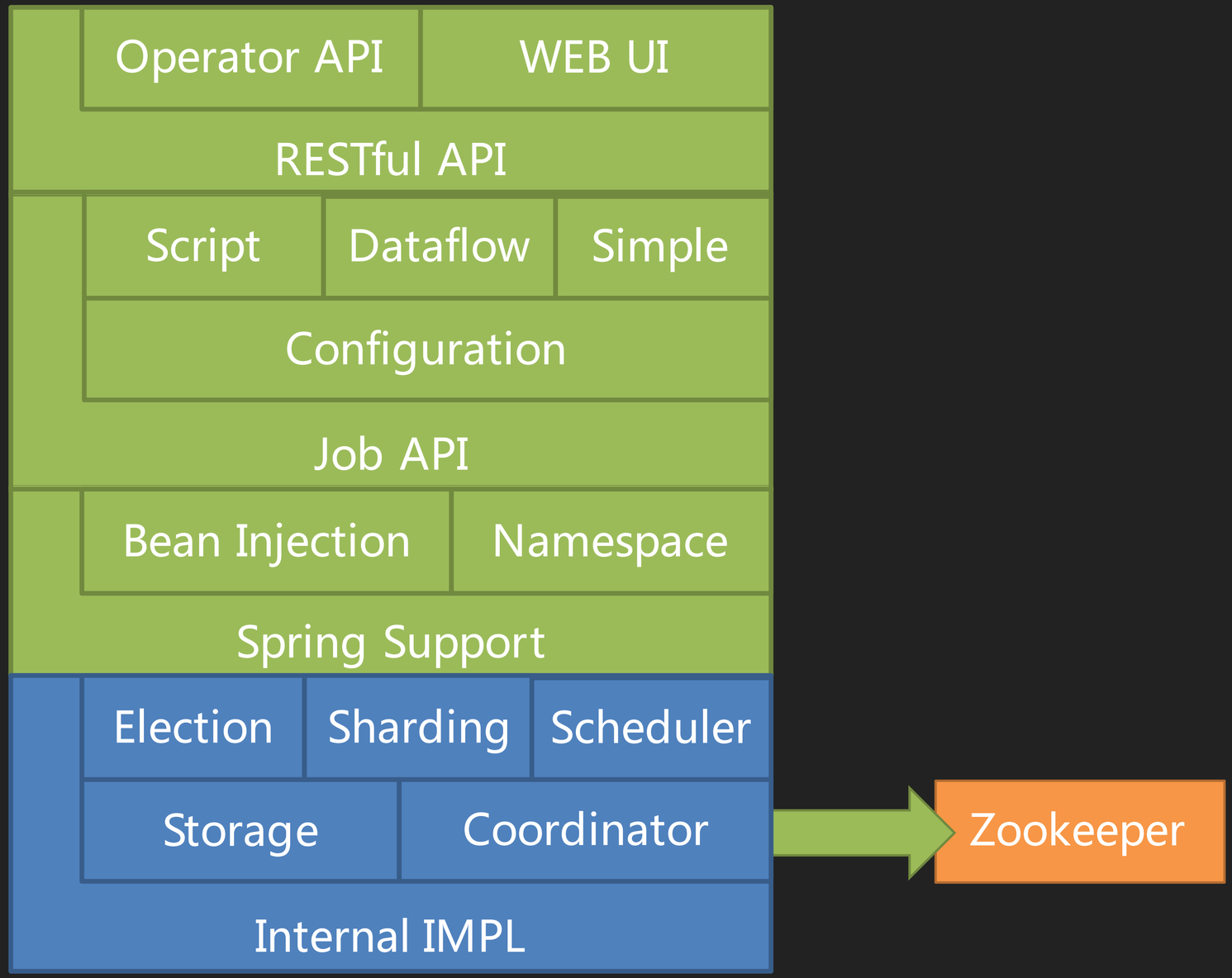
Fail Tolerance

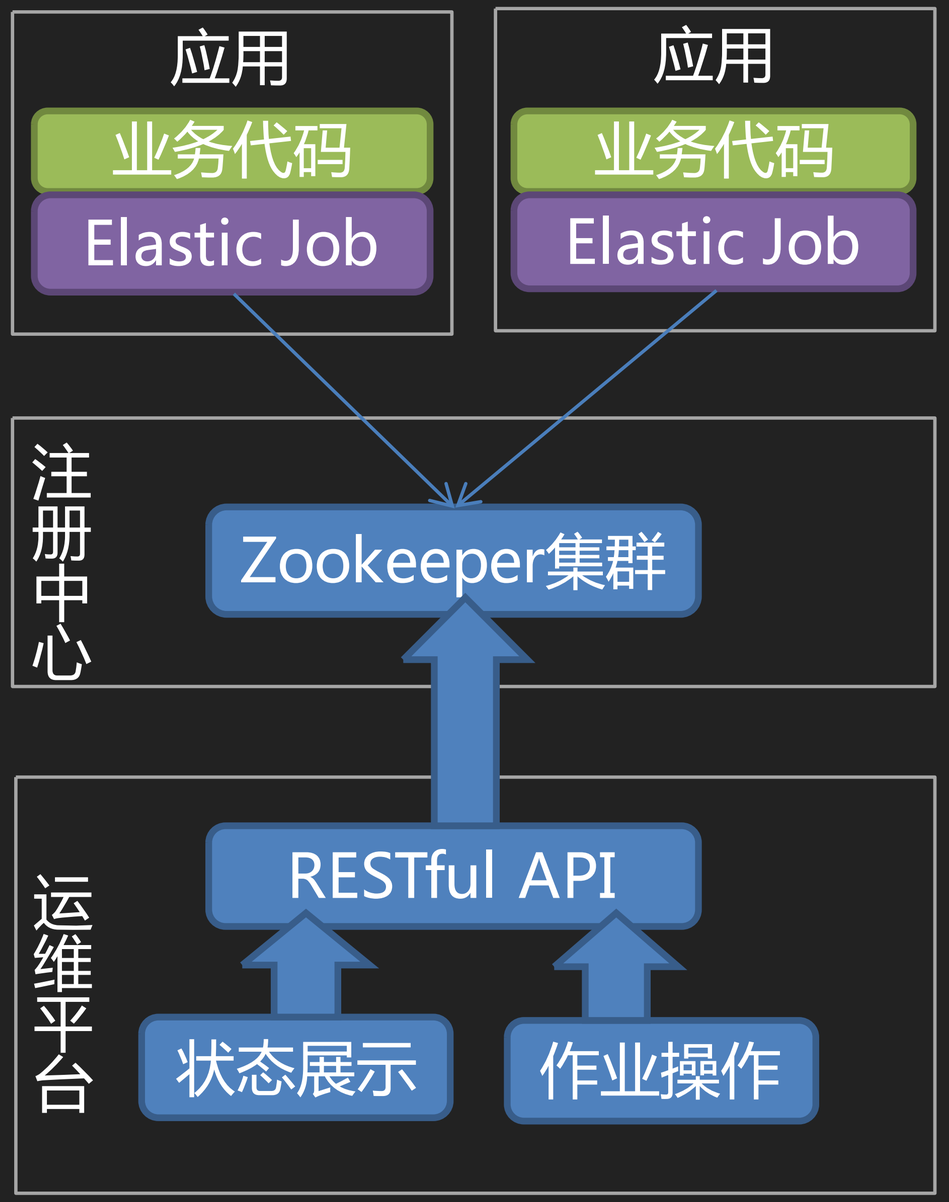
- 作业简史
- 弹性作业


Operator API
WEB UI
RESTful API


Operator API
WEB UI
RESTful API
Script
Dataflow
Simple
Configuration
Job API


Operator API
WEB UI
RESTful API
Script
Dataflow
Simple
Configuration
Job API
Bean Injection
Namespace
Spring Support


Operator API
WEB UI
RESTful API
Script
Dataflow
Simple
Configuration
Job API
Bean Injection
Namespace
Spring Support
Election
Sharding
Scheduler
Storage
Coordinator
Internal IMPL
Zookeeper


应用
业务代码
Elastic Job
应用
业务代码
Elastic Job
注册中心
Zookeeper集群
运维平台
RESTful API
状态展示
作业操作

```
EmbedZookeeperServer.start(EMBED_ZOOKEEPER_PORT);
CoordinatorRegistryCenter regCenter = setUpRegistryCenter();
JobEventConfiguration jobEventConfig = new JobEventRdbConfiguration(setUpEventTraceDataSource());
JobCoreConfiguration coreConfig = JobCoreConfiguration.newBuilder("javaSimpleJob", "0/5 * * * * ?",
3).shardingItemParameters("0=Beijing,1=Shanghai,2=Guangzhou").build();
SimpleJobConfiguration simpleJobConfig = new SimpleJobConfiguration(coreConfig,
JavaSimpleJob.class.getCanonicalName());
new JobScheduler(regCenter, LiteJobConfiguration.newBuilder(simpleJobConfig).build(), jobEventConfig,
new JavaSimpleListener(), new JavaSimpleDistributeListener(1000L, 2000L)).init();
```

```java
public class JavaSimpleJob implements SimpleJob {

    private FooRepository fooRepository = FooRepositoryFactory.getFooRepository();

    @Override
    public void execute(final ShardingContext shardingContext) {
        System.out.println(String.format("------Thread ID: %s, Date: %s, Sharding Context: %s, Action: %s",
Thread.currentThread().getId(), new Date(), shardingContext, "simple job"));
        List<Foo> data = fooRepository.findTodoData(shardingContext.getShardingParameter(), 10);
        for (Foo each : data) {
            fooRepository.setCompleted(each.getId());
        }
    }
}
```

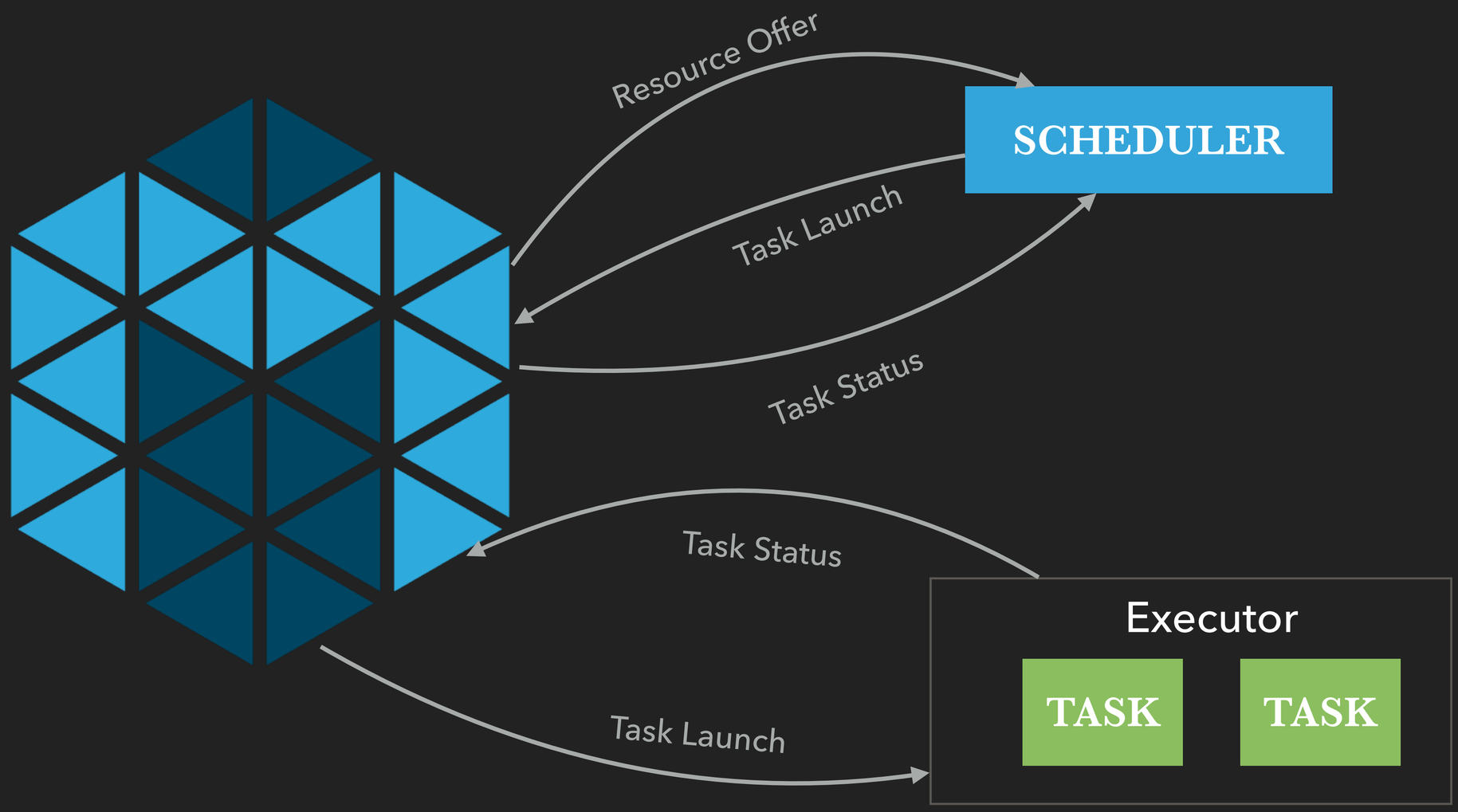
```
simple.id=springSimpleJob
simple.class=com.dangdang.ddframe.job.example.job.simple.SpringSimpleJob
simple.cron=0/5 * * * * ?
simple.shardingTotalCount=3
simple.shardingItemParameters=0=Beijing,1=Shanghai,2=Guangzhou
simple.monitorExecution=false
simple.failover=true
simple.description=\u53EA\u8FD0\u884C\u4E00\u6B21\u7684\u4F5C\u4E1A\u793A\u4F8B
simple.disabled=false
simple.overwrite=true
simple.monitorPort=9888
```

- 作业简史
- 弹性作业
- 弹性作业云

	Framework	Architecture	Resource granularity	Multi-scheduler	Pluggable logic	Priority preemption	Re-scheduling	Oversubscription	Resource estimation	Avoid interference
OPEN	Kubernetes	monolithic	multi-dimensional	N[v1.2, DD, Issue]	Y[DD]	N[Issue]	N[Issue]	Y[DD]	N	N
	Swarm	monolithic	multi-dimensional	N	N	N[Issue]	N	N	N	N
	YARN	monolithic/ two-level	RAM/CPU slots	Y	N[*app-lvl. only*]	N[JIRA]	N	N[JIRA]	N	N
	Mesos	two-level	multi-dimensional	Y	Y[*framework-lvl.*]	N[JIRA]	N	Y[v0.23, Doc]	N	N
	Nomad	shared-state	multi-dimensional	Y	Y	N[Issue]	N[Issue]	N[Issue]	N	N
	Sparrow	fully-distributed	fixed slots	Y	N	N	N	N	N	N
CLOSED	Borg	monolithic[7]	multi-dimensional	N[7]	N[7]	Y	Y	Y	Y	N
	Omega	shared-state	multi-dimensional	Y	Y	Y	Y	Y	Y	N
	Apollo	shared-state	multi-dimensional	Y	Y	Y	Y	N	N	N

Go to mark...

MESOS


Resource Offer
SCHEDULER
Task Launch
Task Status
Task Status
Executor
TASK
TASK
Task Launch

OFFICIAL FRAMEWORK

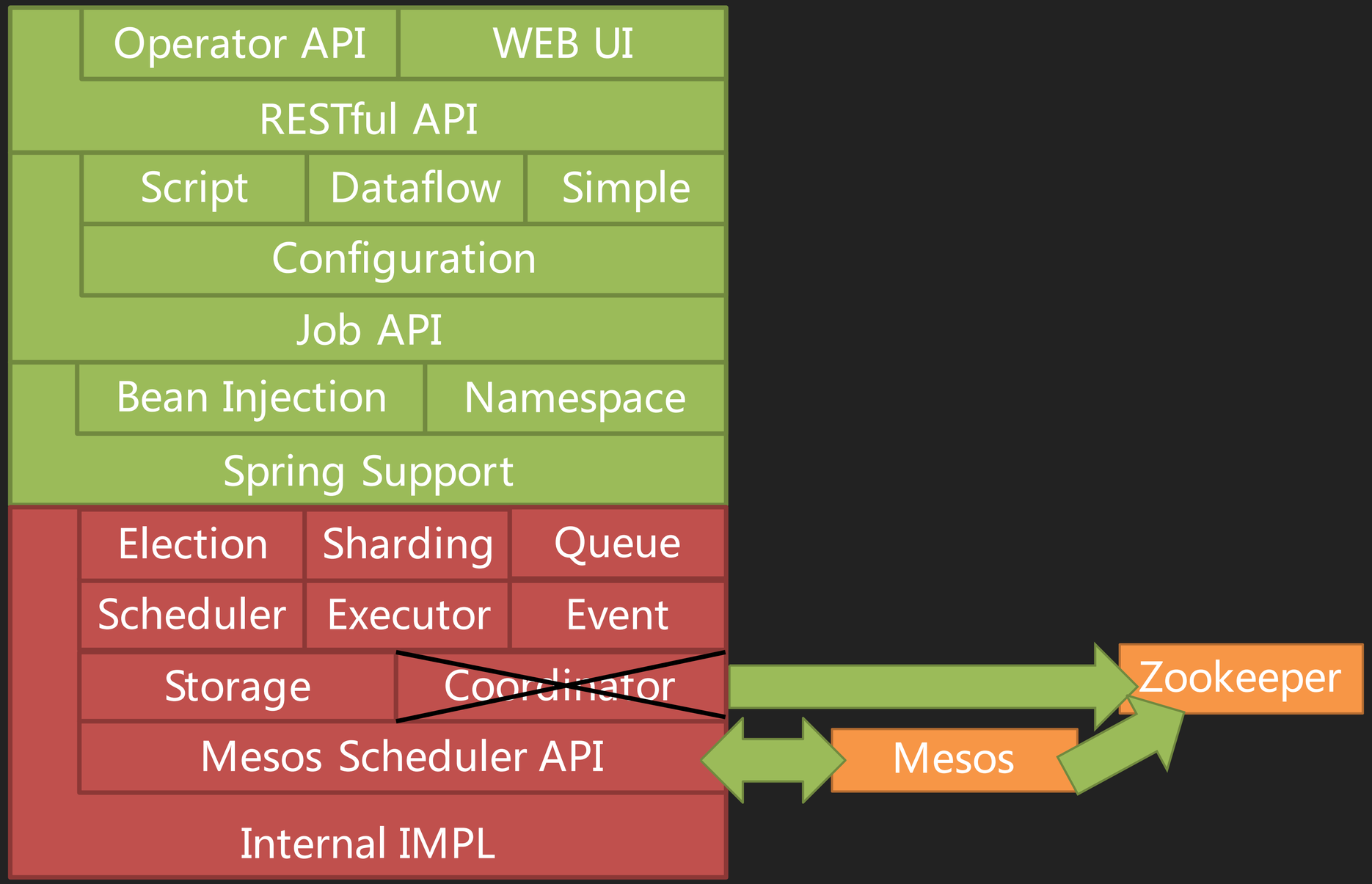
Batch Scheduling

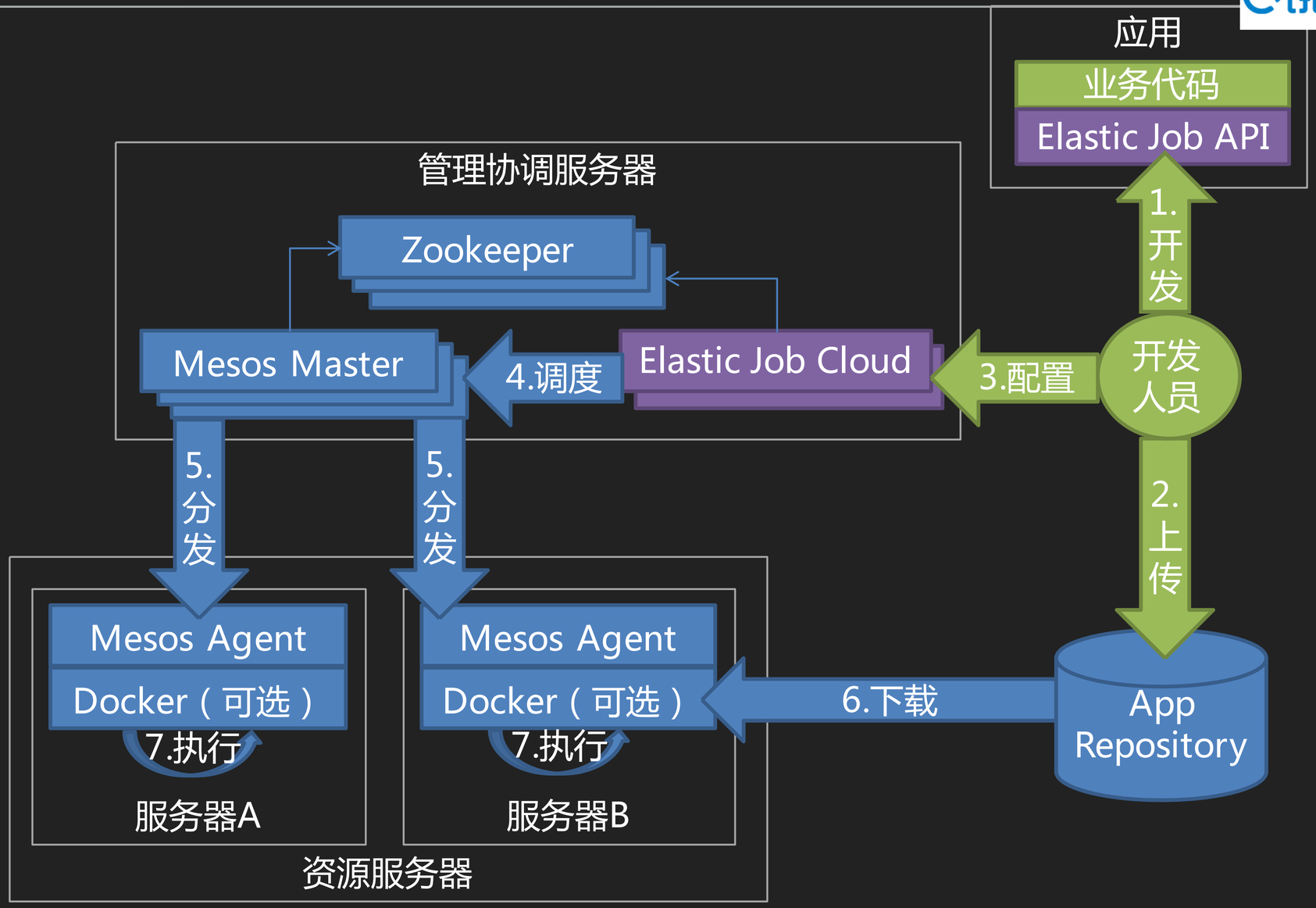
- Chronos is a distributed job scheduler that supports complex job topologies. It can be used as a more fault-tolerant replacement for Cron.
- Cook is a job scheduler like Torque that not only supports individual tasks, but also Spark. Cook provides powerful automatic preemption and multitenancy features for shared clusters, in order to guarantee throughput to all users while allowing individuals to temporarily “burst” t additional resources as needed. Cook provides a simple REST API & Java client for interaction.
- Elastic-Job-Cloud is a distributed scheduled job cloud solution designed with HA and fault-tolerance in mind. It focuses on horizontal scaling, and provides transient and daemon jobs, event and schedule based job triggers, job dependencies, and job history.
- GoDocker is a batch computing job scheduler like SGE, Torque, etc. It schedules batch computing tasks via webui, API or CLI for system or LDAP users, mounting their home directory or other shared resources in a Docker container. It targets scientists, not developers, and provid plugin mechanisms to extend or modify the default behavior.
- Jenkins is a continuous integration server. The mesos-jenkins plugin allows it to dynamically launch workers on a Mesos cluster depending on the workload.
- JobServer is a distributed job scheduler and processor which allows developers to build custom batch processing Tasklets using point and click web UI.

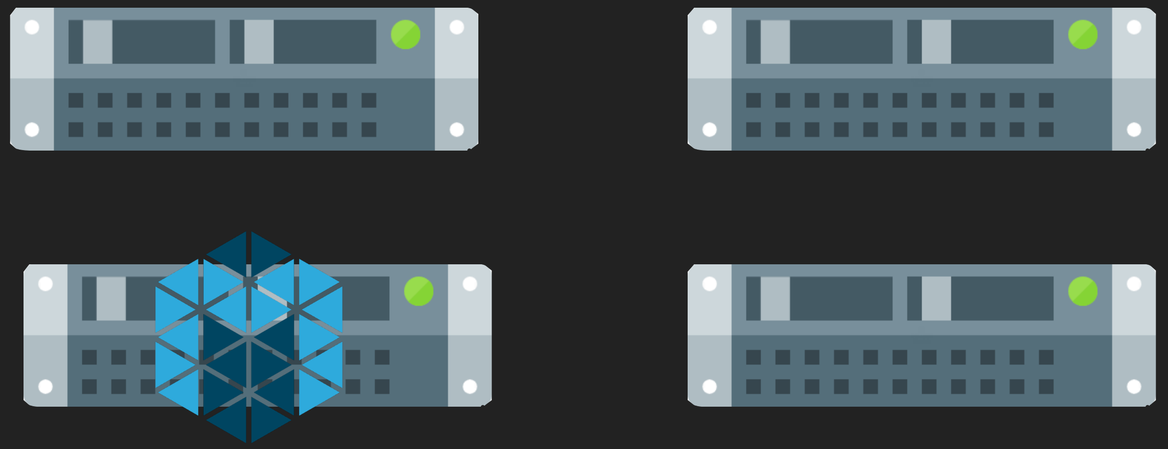
OFFICIAL FRAMEWORK

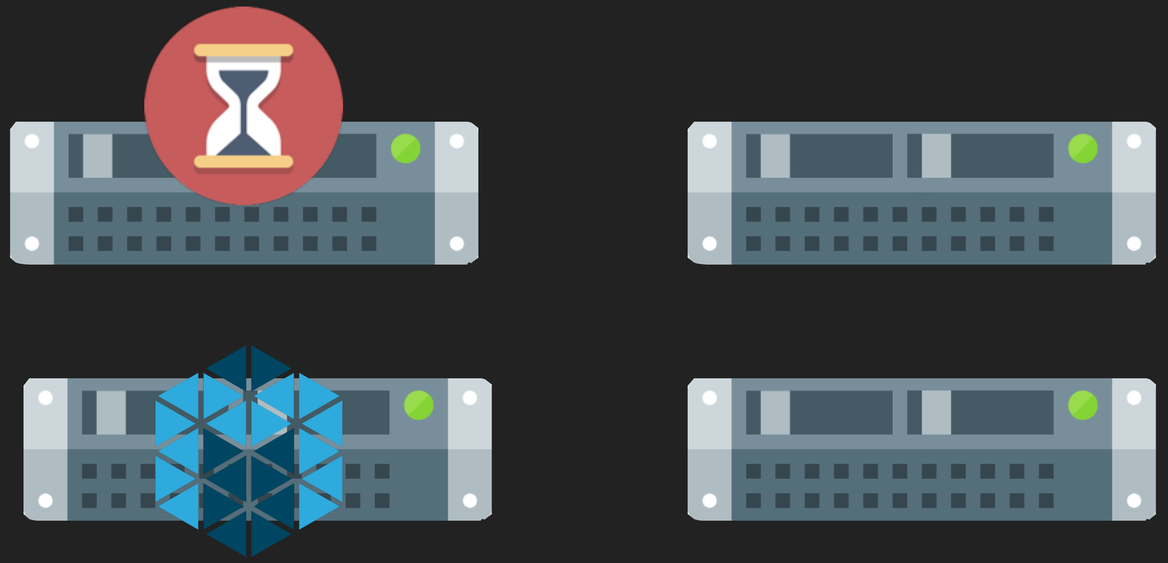
Batch Scheduling

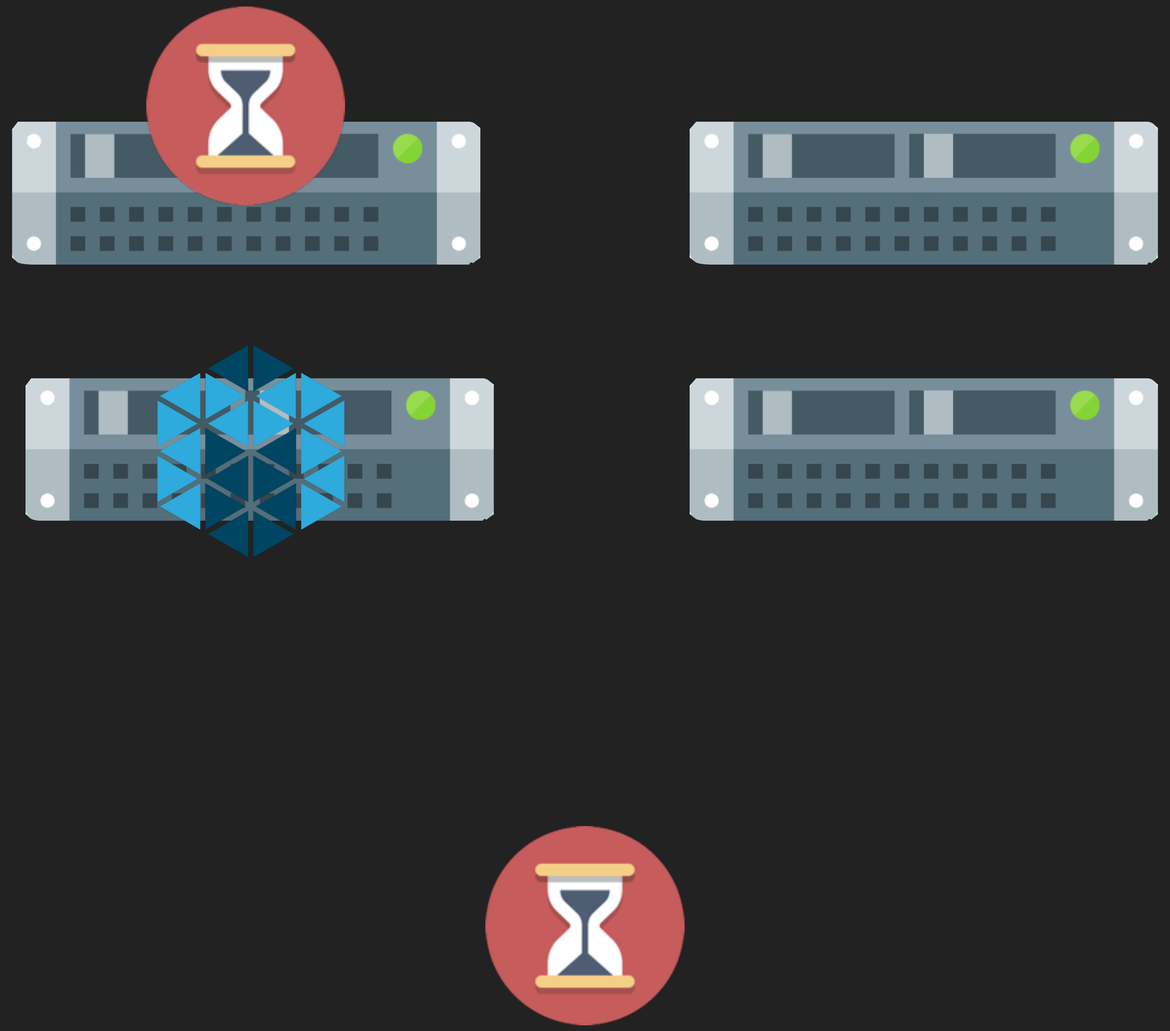
- Chronos is a distributed job scheduler that supports complex job topologies. It can be used as a more fault-tolerant replacement for Cron.
- Cook is a job scheduler like Torque that not only supports individual tasks, but also Spark. Cook provides powerful automatic preemption and multitenancy features for shared clusters, in order to guarantee throughput to all users while allowing individuals to temporarily “burst” t additional resources as needed. Cook provides a simple REST API & Java client for interaction.
- Elastic-Job-Cloud is a distributed scheduled job cloud solution designed with HA and fault-tolerance in mind. It focuses on horizontal scaling, and provides transient and daemon jobs, event and schedule based job triggers, job dependencies, and job history.
- GoDocker is a batch computing job scheduler like SGE, Torque, etc. It schedules batch computing tasks via webui, API or CLI for system or LDAP users, mounting their home directory or other shared resources in a Docker container. It targets scientists, not developers, and provid plugin mechanisms to extend or modify the default behavior.
- Jenkins is a continuous integration server. The mesos-jenkins plugin allows it to dynamically launch workers on a Mesos cluster depending on the workload.
- JobServer is a distributed job scheduler and processor which allows developers to build custom batch processing Tasklets using point and click web UI.

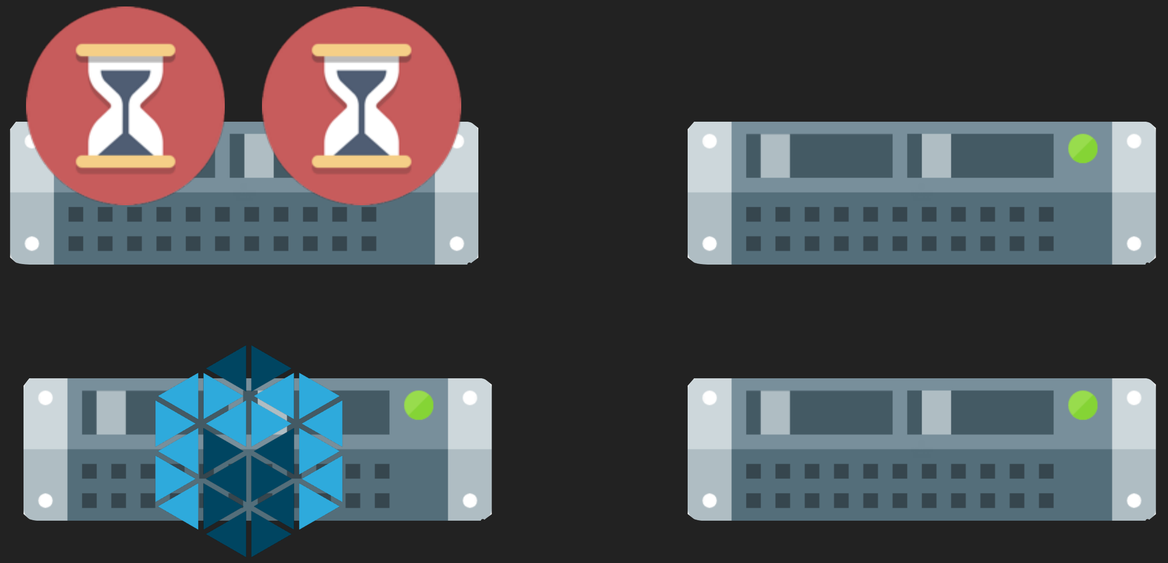

Operator API
WEB UI
RESTful API
Script
Dataflow
Simple
Configuration
Job API
Bean Injection
Namespace
Spring Support
Election
Sharding
Queue
Scheduler
Executor
Event
Storage
Coordinator
Mesos Scheduler API
Internal IMPL
Zookeeper
Mesos

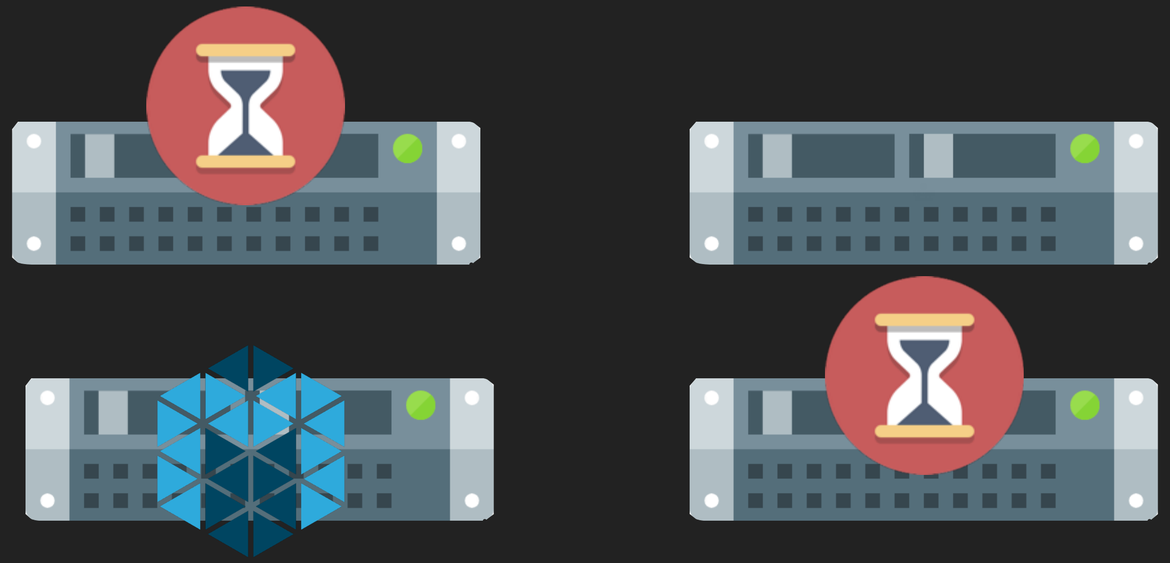

应用
业务代码
Elastic Job API
管理协调服务器
Zookeeper
Mesos Master
4.调度
Elastic Job Cloud
3.配置
开发人员
1.开发
2.上传
5.分发
5.分发
Mesos Agent
Docker（可选）
7.执行
服务器A
Mesos Agent
Docker（可选）
7.执行
服务器B
资源服务器
6.下载
App Repository



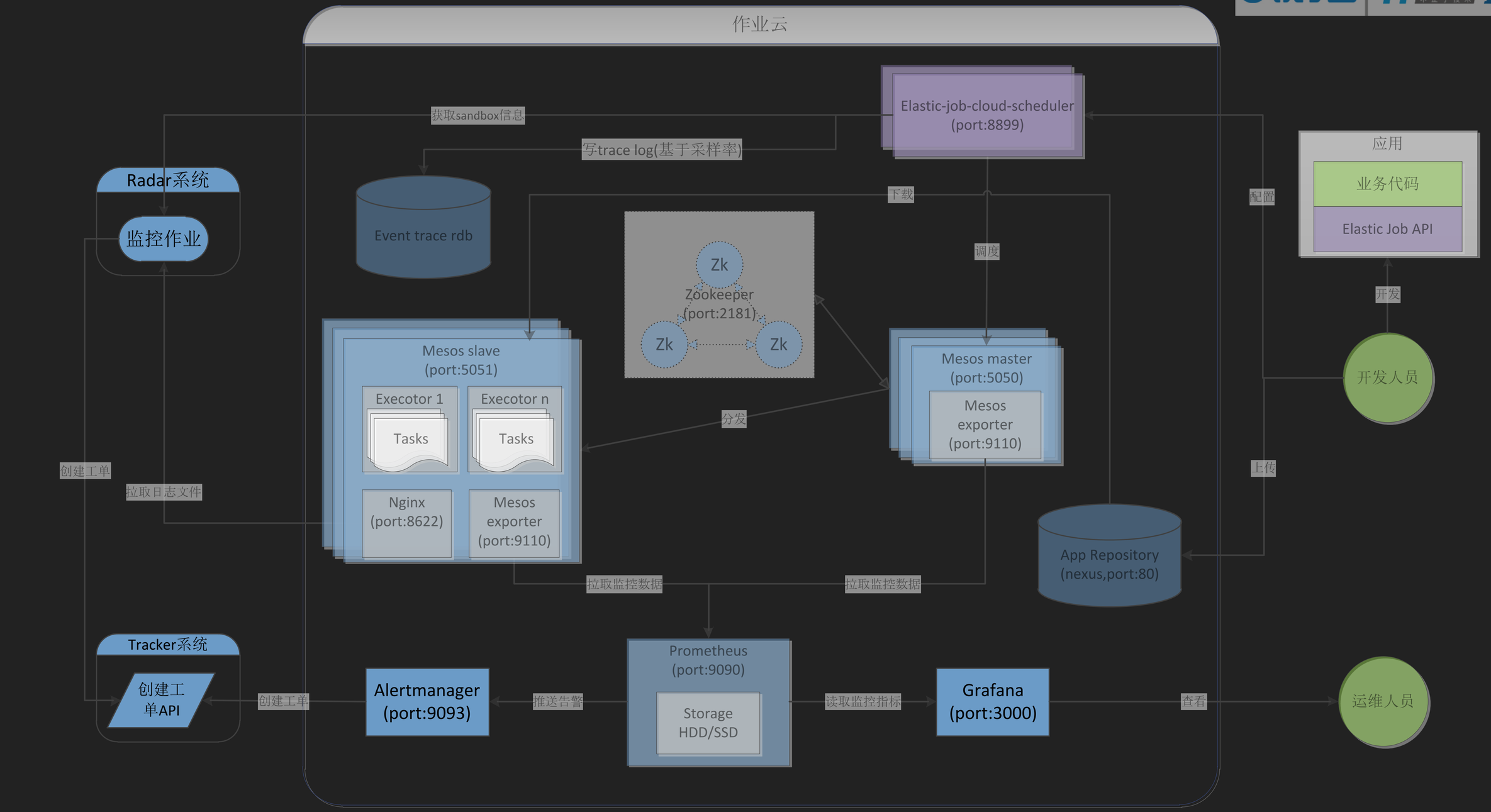








- 作业简史
- 弹性作业
- 弹性作业云
- 弹性作业云最佳实践

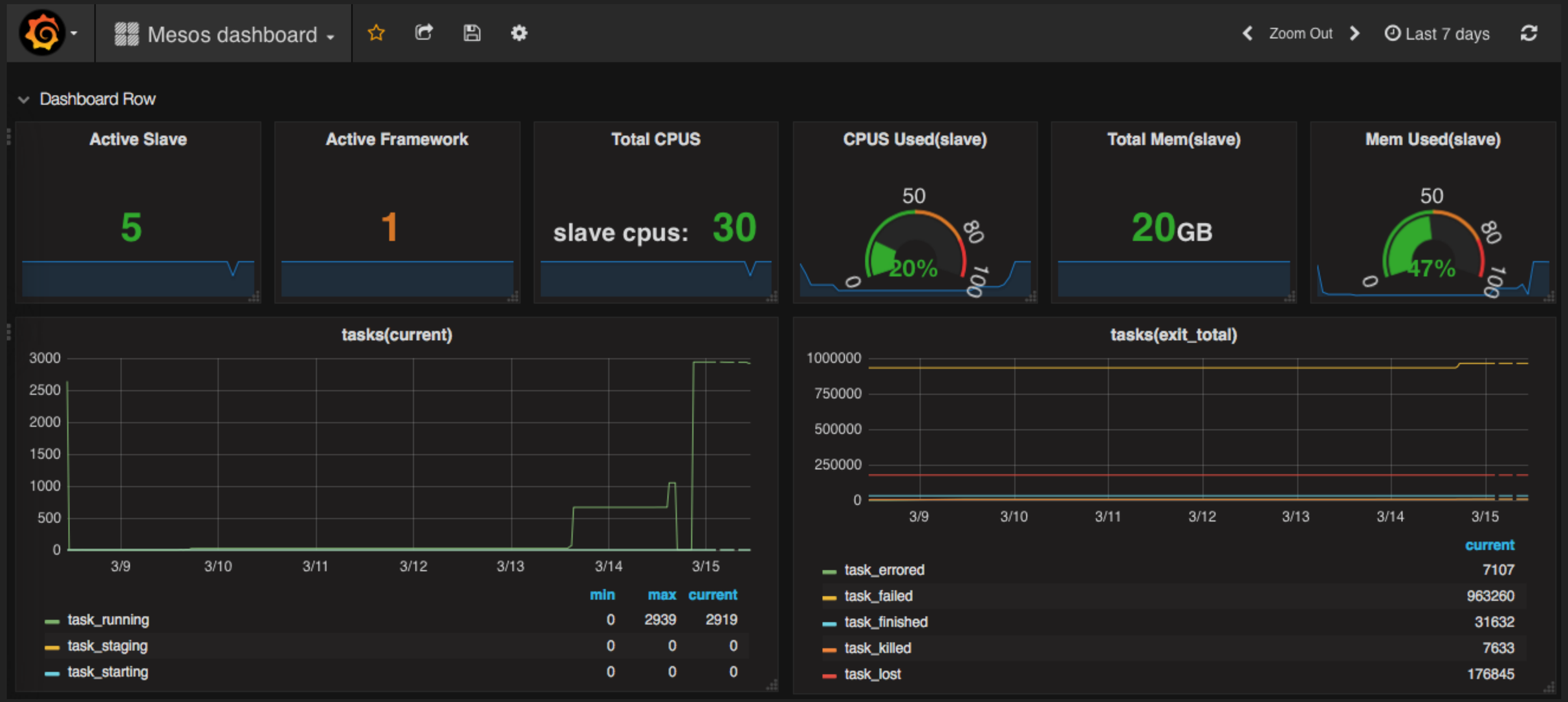

作业云
Elastic-job-cloud-scheduler
(port:8899)
获取sandbox信息
写trace log(基于采样率)
Radar系统
监控作业
Event trace rdb
下载
调度
Zk
Zookeeper
(port:2181)
Zk
Zk
应用
业务代码
Elastic Job API
配置
开发
开发人员
Mesos slave
(port:5051)
Execotor 1
Tasks
Execotor n
Tasks
Nginx
(port:8622)
Mesos
exporter
(port:9110)
分发
Mesos master
(port:5050)
Mesos
exporter
(port:9110)
创建工单
拉取日志文件
上传
App Repository
(nexus,port:80)
拉取监控数据
拉取监控数据
Tracker系统
创建工
单API
创建工单
Alertmanager
(port:9093)
推送告警
Prometheus
(port:9090)
Storage
HDD/SSD
读取监控指标
Grafana
(port:3000)
查看
运维人员

SCHEDULER

STANDBY SCHEDULER

~~SCHEDULER~~

SCHEDULER


Mesos dashboard
Zoom Out
Last 7 days
Dashboard Row
Active Slave
5
Active Framework
1
Total CPUS
slave cpus: 30
CPUS Used(slave)
50
80
100
0
20%
Total Mem(slave)
20GB
Mem Used(slave)
50
80
100
0
47%
tasks(current)
3000
2500
2000
1500
1000
500
0
3/9
3/10
3/11
3/12
3/13
3/14
3/15
min max current
task_running 0 2939 2919
task_staging 0 0 0
task_starting 0 0 0
tasks(exit_total)
1000000
750000
500000
250000
0
3/9
3/10
3/11
3/12
3/13
3/14
3/15
current
task_errored 7107
task_failed 963260
task_finished 31632
task_killed 7633
task_lost 176845