

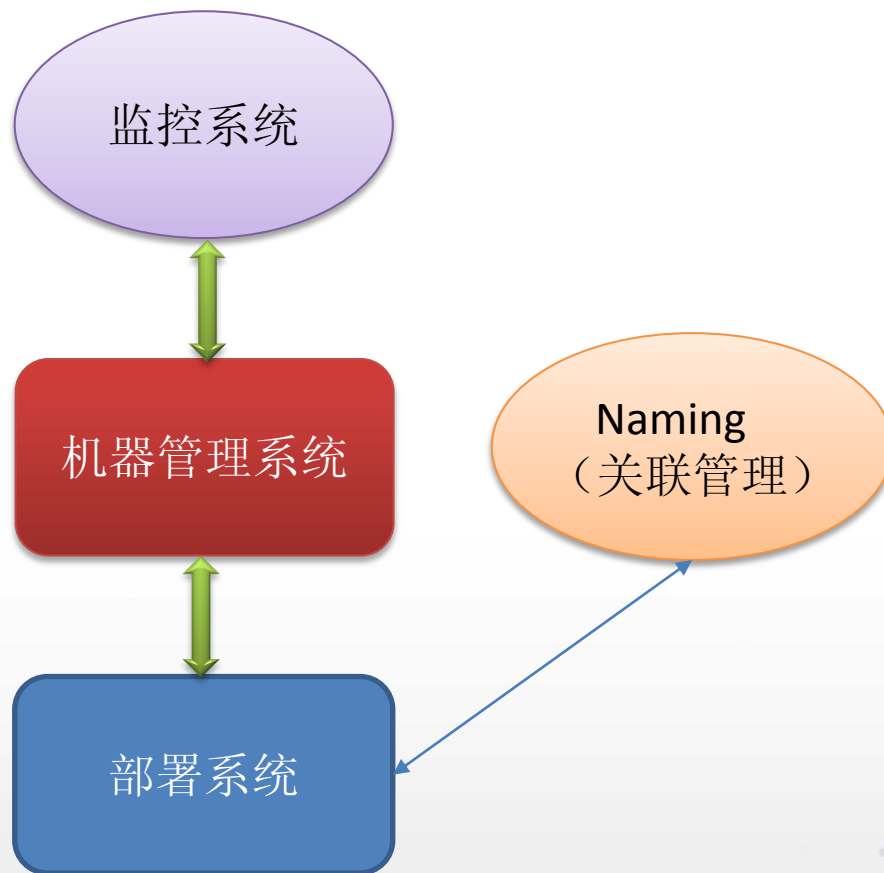
SACC 2014中国系统架构师大会
SYSTEM ARCHITECT CONFERENCE CHINA 2014

发现架构之美

部署系统的前世今生

小米 方少森 北京

运维核心系统



部署系统的几个阶段

业务多样性 – 规范化阶段

运维复杂性 – 以动制动阶段

彻底解放 – 完全动态阶段

1.1 业务多样性

- 早期业务发展快
 - 30+开发团队
- SRE人手不足
 - 50:1以上
- 服务管理、部署奔放
 - 环境复杂
 - 缺少服务守护
 - 编译方式杂



1.2 规范一：环境依赖



- ✓ 系统环境收敛
- ✓ 服务自描述runtime依赖
- ✓ 支持不同服务的runtime依赖隔离

1.3 规范二：进程管理方式

Daemon服务
Nginx、resin等

Nohup
Java等前台服务

Cron
定时守护任务

Supervisord
管理前台启动服务

GOD

A Process Monitoring Framework in Ruby

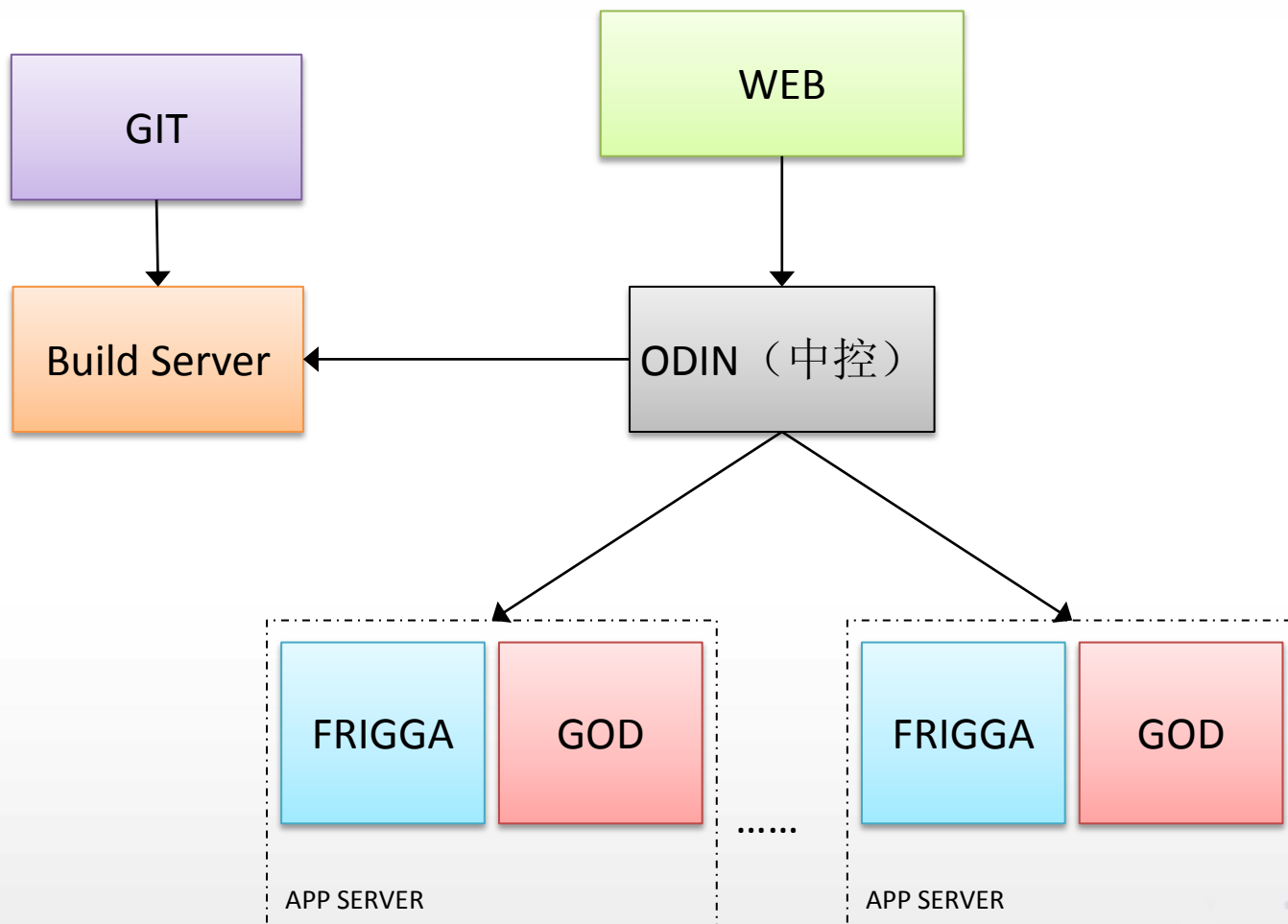
god load service
god start service
god stop service
god nuke service
god log service
god cron service

1.4 规范三：编译部署联动

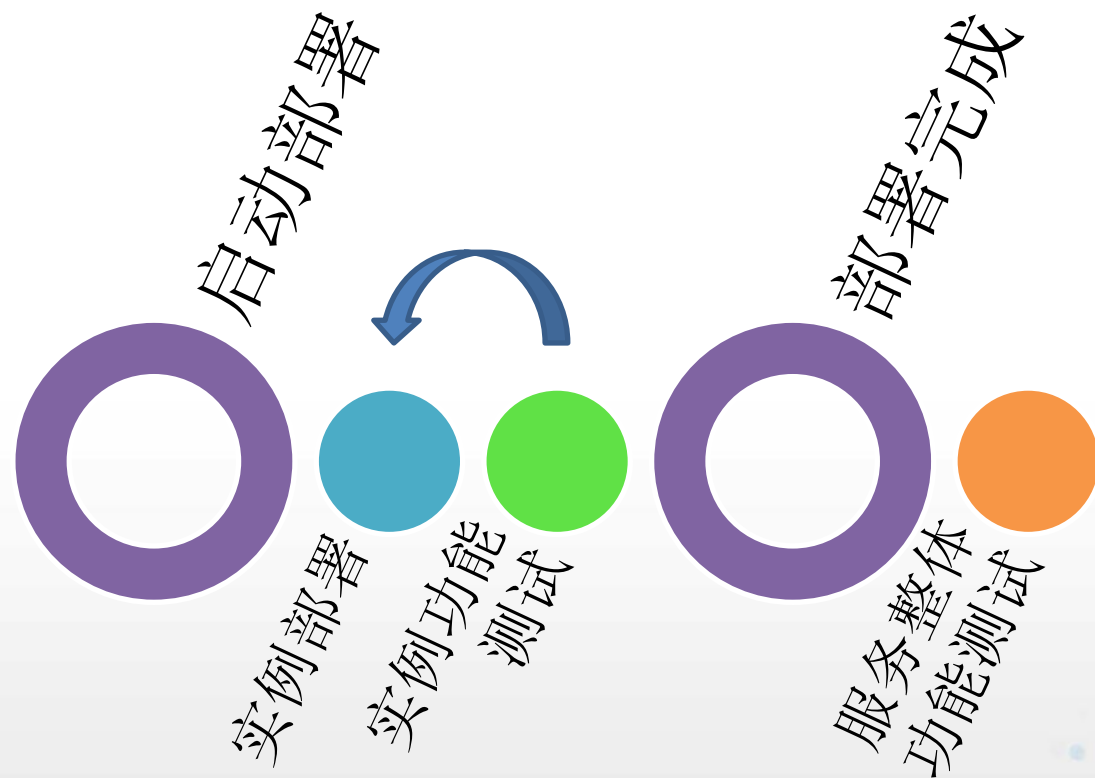
- ✓ 发布包规范
- ✓ 提供标准的编译框架
- ✓ 可选的一键编译+测试+部署机制



1.5 第一阶段系统图



1.6 部署质量保证-部署对接测试



1.7 部署质量保证 – 快速回滚

- ✓ 全量发布
- ✓ 回滚即发布

Top 实例详情

job.deploy-test_service.deploy-test_cluster.production-sd_pdl.com-ovt.miliao_cop.xiaomi

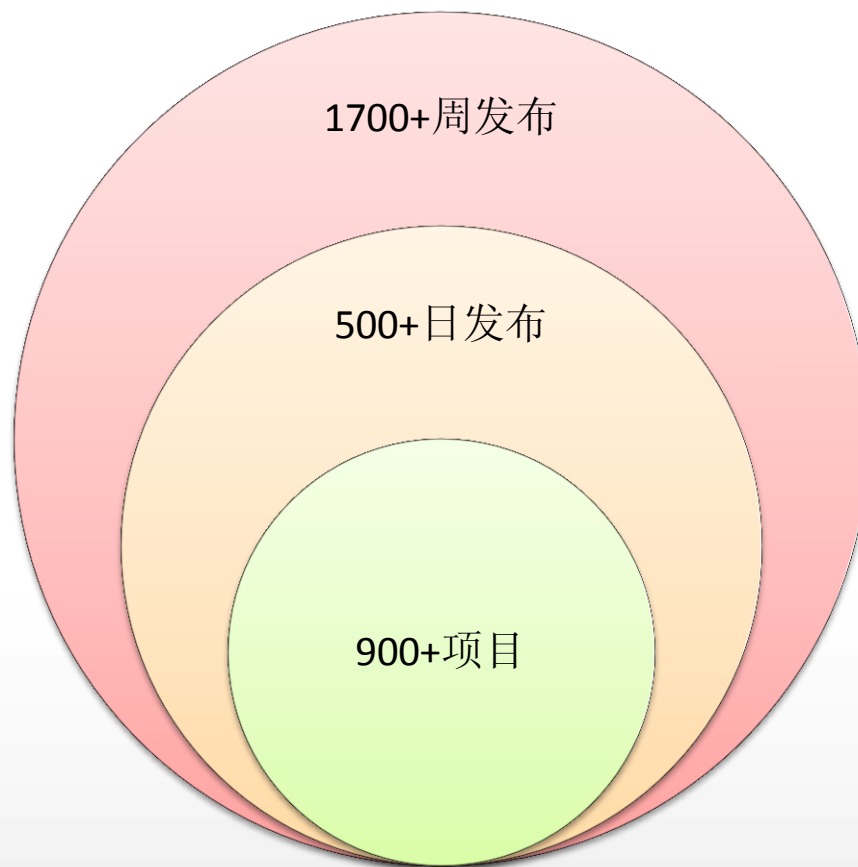
	版本号	发起人	发起时间	机器数	部署详情
回滚至此版本	15b2980f8	fangshaosen	2014-08-15 17:22:49	1	cluster.conf 机器列表
回滚至此版本	15b2980f8	fangshaosen	2014-08-15 16:56:12	1	cluster.conf 机器列表
回滚至此版本	15b2980f8	fangshaosen	2014-08-15 10:47:55	1	cluster.conf 机器列表

1.8 示例：两种部署方式

ID	用户	发起	jobname	version	状态	控制	log
<input type="checkbox"/> 25237	fangshaosen	2014-07-30 15:37:51	job.test_service.deploy_cluster.production-sd	1239a4dd7f58	编译中		查看 操作
任务进度 25237							
<div>0%</div>							
机器名	部署状态 0 0 1		刷新 Top	job日志			
lg-com-dev02.bj				程序日志			

```
bash -x ./build-release.sh -u git@git.n.xiaomi.com:deploy-test.git -t job.test_service.deploy -v 8c5cc2d5e
```

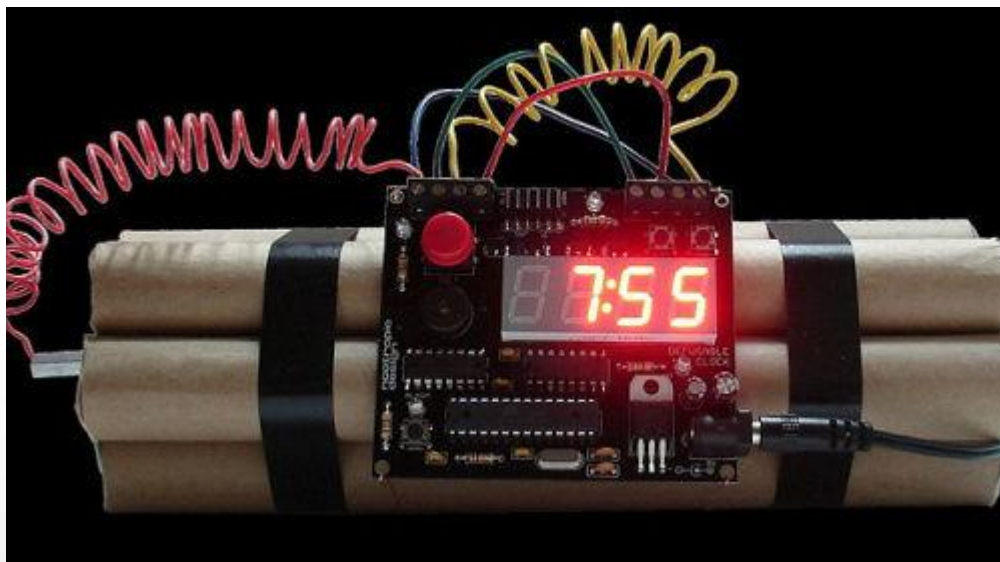
1.9 第一阶段总结



让人欢喜让人忧

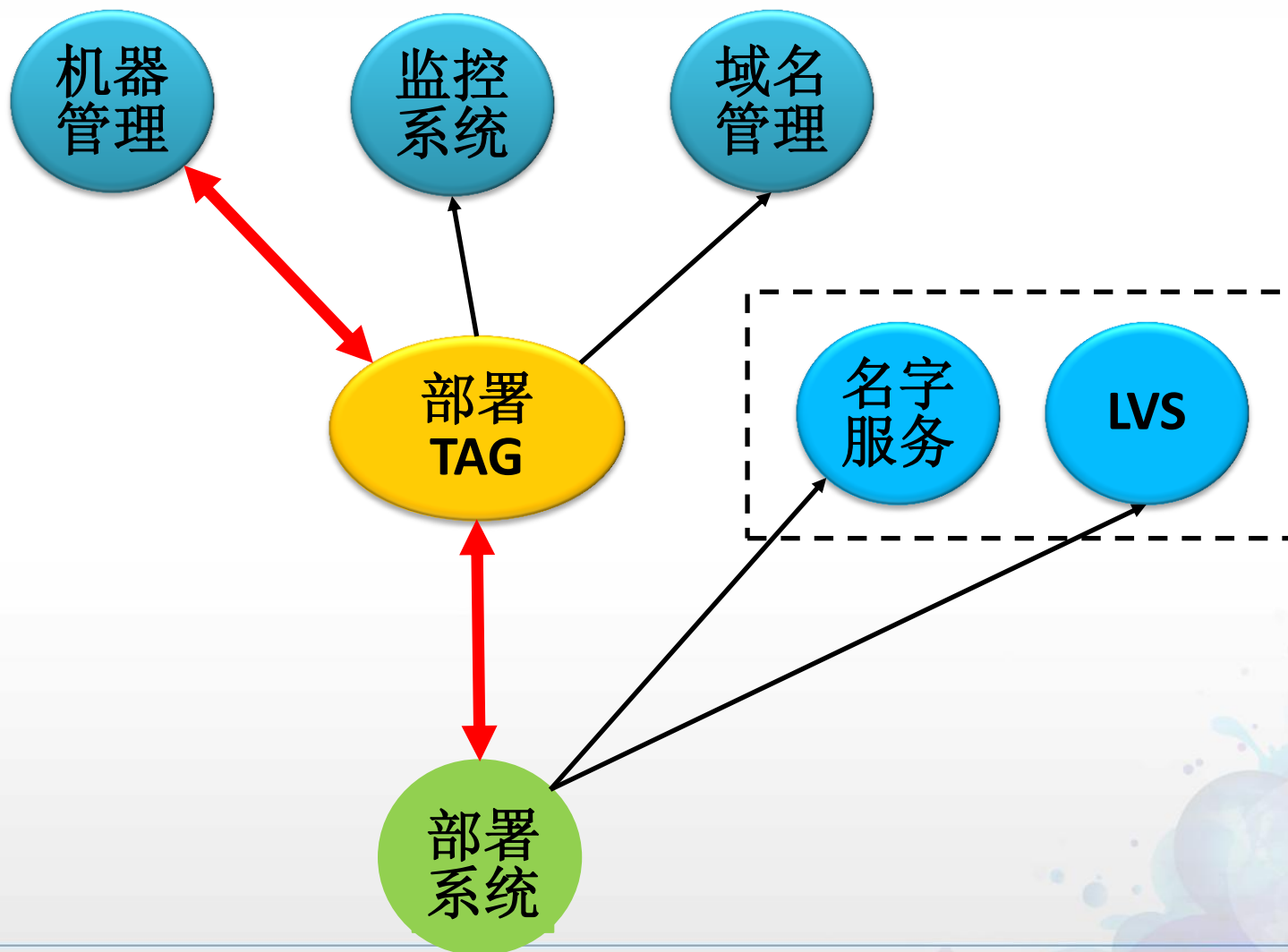
2.1 运维工作复杂性

- ✓ 每次变更手动更新机器管理
- ✓ 每次变更手动更新监控
- ✓ 每次变更手动更新关联

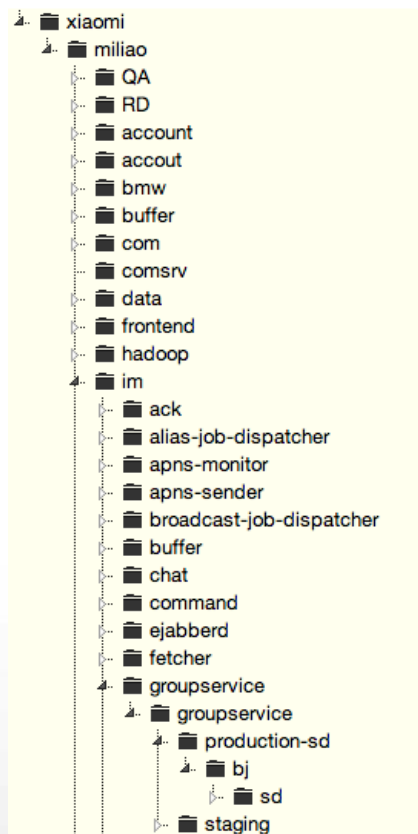
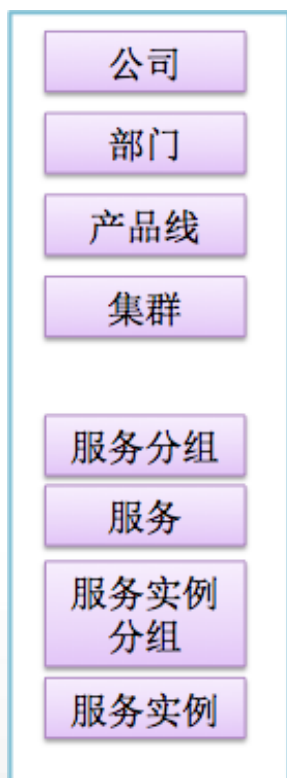


- ✓ 以动制动

2.2 多系统联动



2.3 部署TAG



- ✓ 业务层级关系的可视化展示
- ✓ 业务和机器列表映射关系的可视化展示
- ✓ 维护者：部署系统
- ✓ 使用者：监控系统等

cop.xiaomi_owt.miliao_pdl.im_service.ack
cop.xiaomi_owt.miliao_pdl.im_service.chat

2.4 监控自动维护



基础监控

- 自动上报
- 策略随TAG自动更新

服务监控

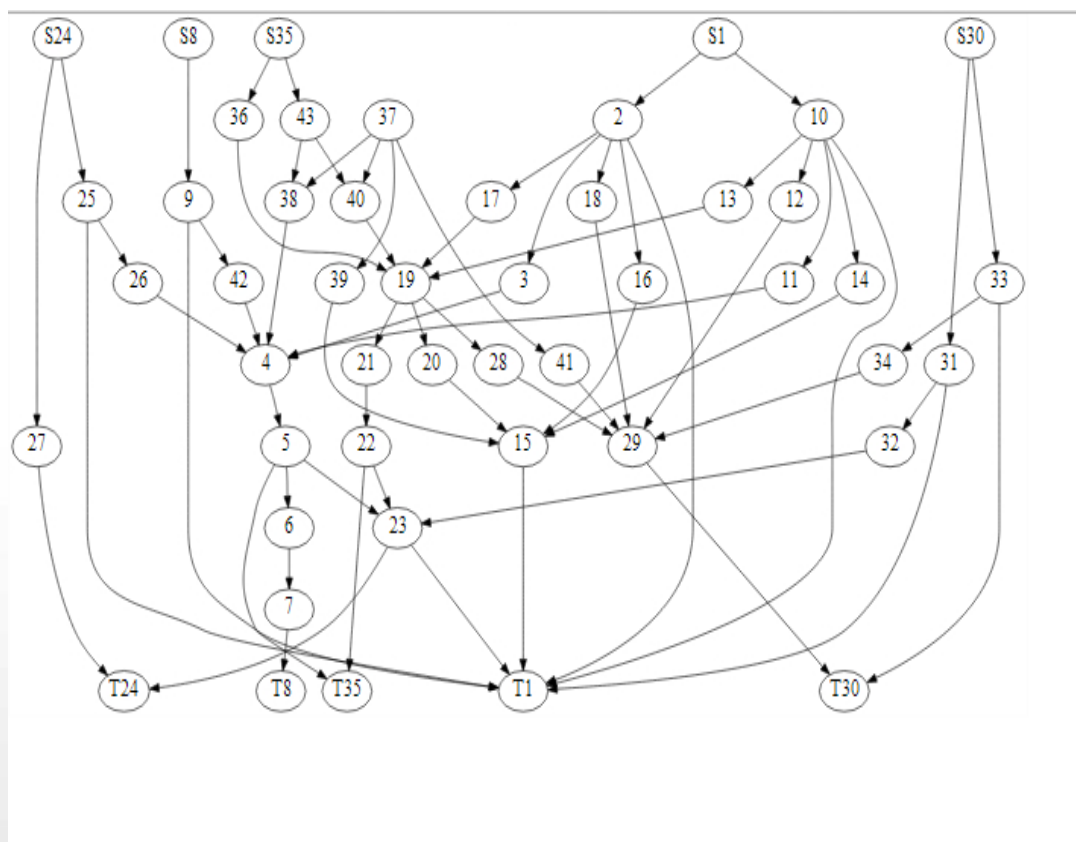
- God自动上报
- 策略随TAG自动更新

2.5 监控自动维护实例

理论上可以做到一条策略监控全公司所有实例的运行情况

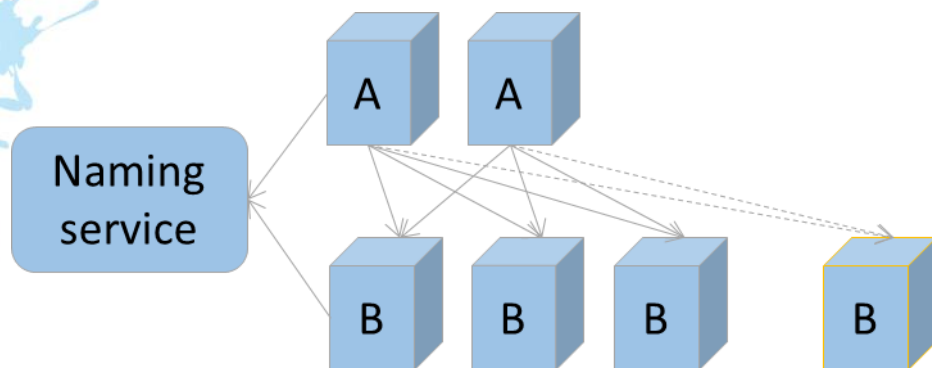
metric	tags	func	op	右值	报警次数	P级	note	生效开始	生效结束
common.servicerunning	cop=xiaomi,owt=miliao,pdl=account	all(#3)	=	0	99	2			

2.6 关联关系的维护



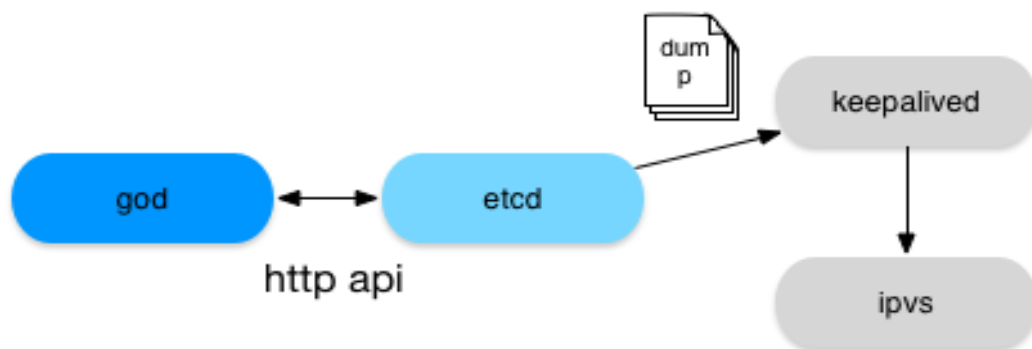
- ✓ 传统方式存在的问题
 - ✓ 更新不及时
 - ✓ 使用方不清晰

2.6 关联关系的维护

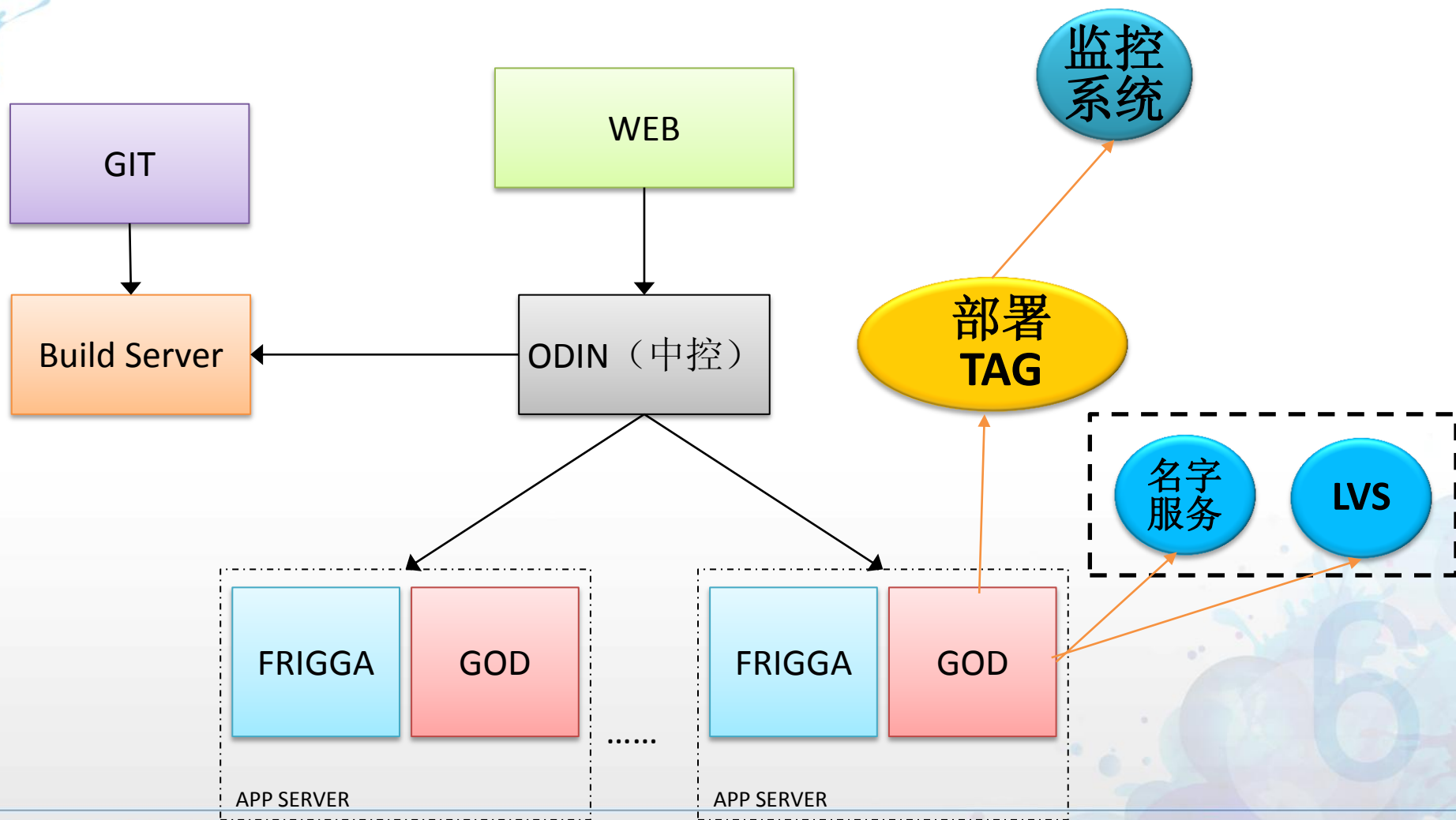


✓ 服务关联

✓ LVS注册



2.7 第二阶段总结



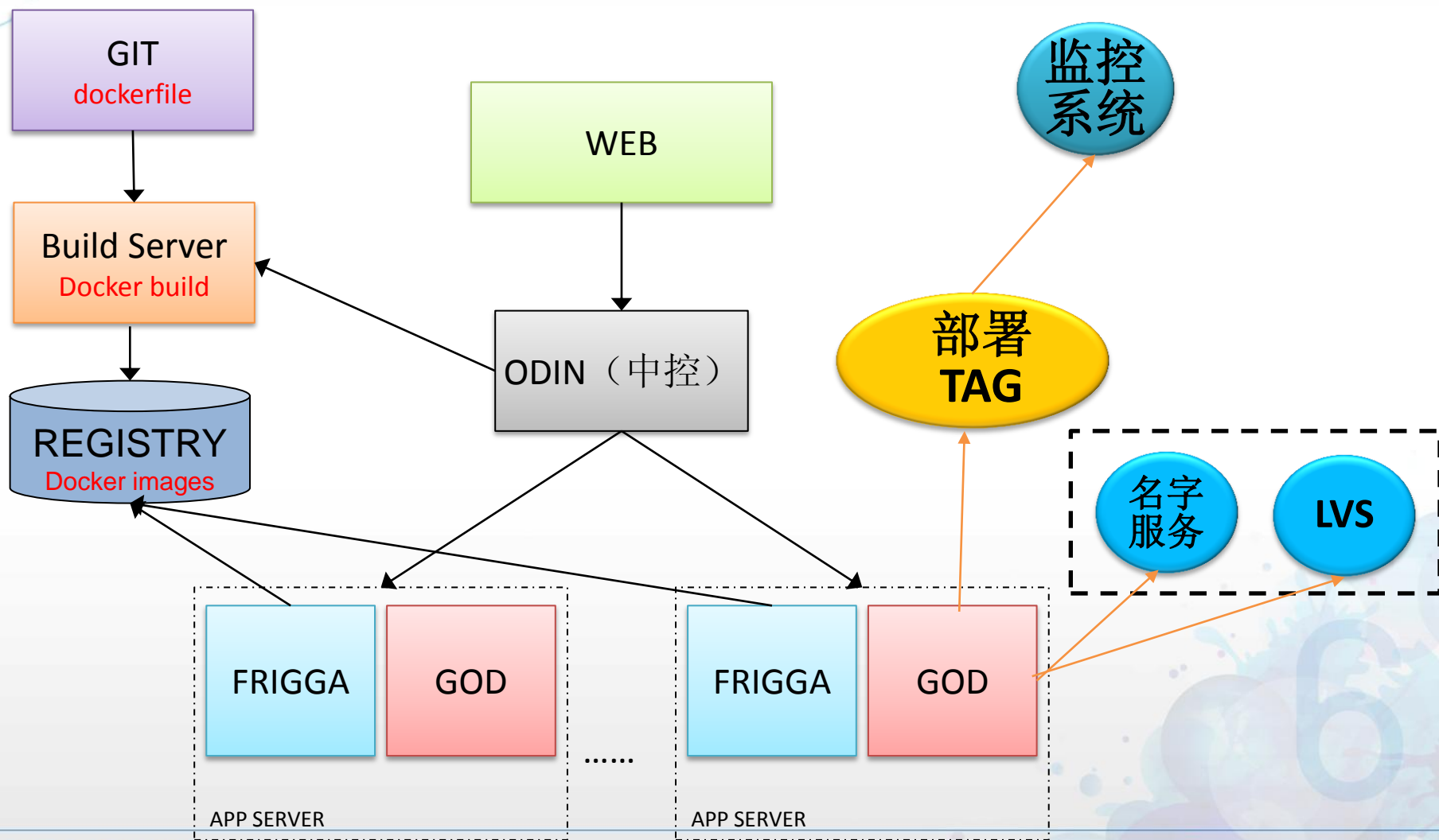
3.1 进阶问题

- 故障节点动态迁移
- 资源利用率提升

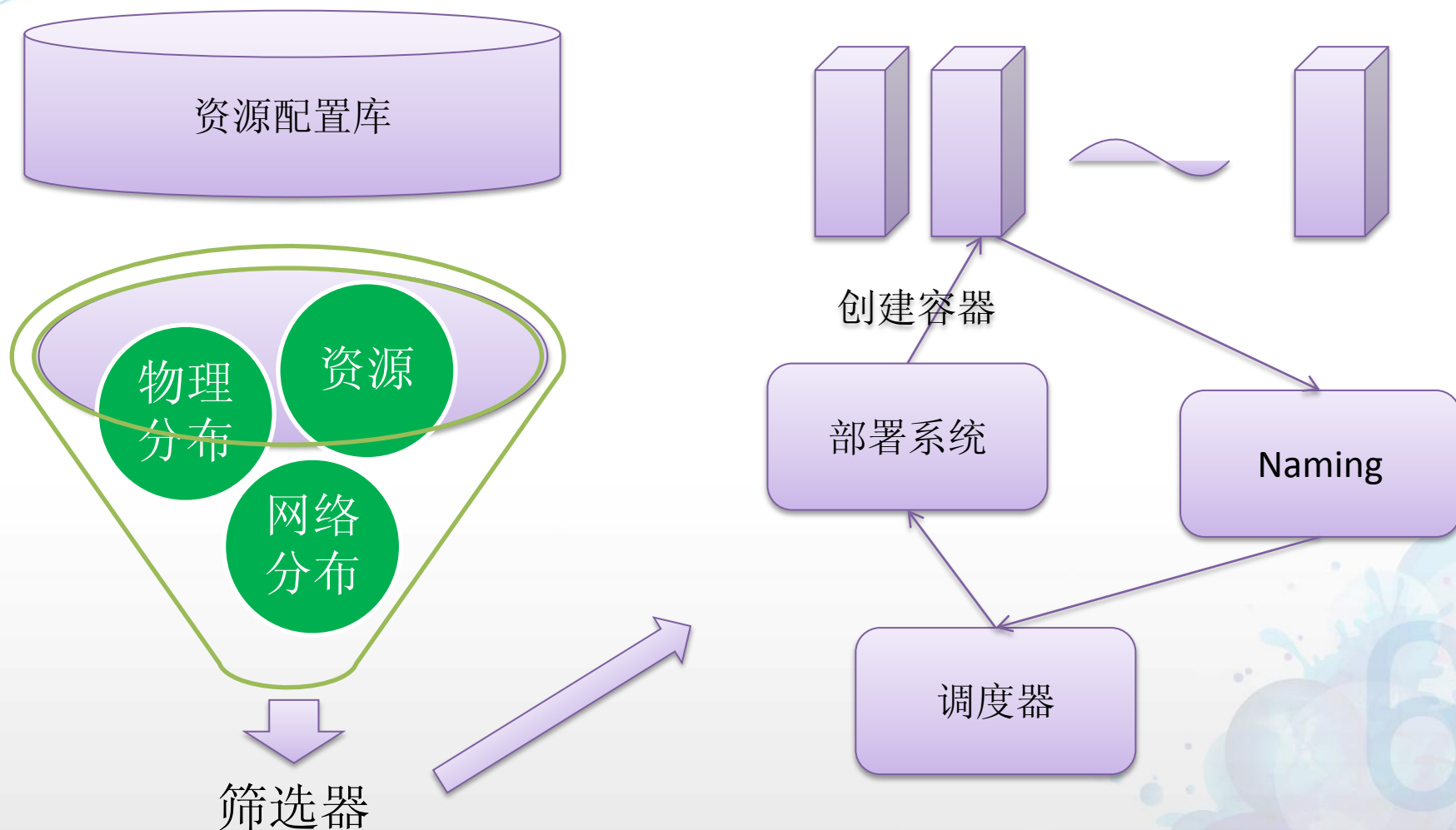
3.2 资源隔离 - docker

- 选型
 - 快速构建、定制化
 - 轻量化、快速发布
 - 资源消耗少
 - 兼容原有部署系统
- 优化
 - IO性能问题
 - Volume方案
 - 网络性能问题
 - 动态IP方案

3.3 docker部署



3.4 动态调度



总结

- ✓ 不同阶段的目标
 - ✓ 第一阶段：运维刚介入，规范化
 - ✓ 第二阶段：运维自身解放
 - ✓ 第三阶段：提升

Our goal is noops!!!

技术博客: <http://www.noops.me>

Q&A

THANKS

SequeMedia
盛拓传媒

IT168.com
www.it168.com

ChinaUnix

ITPUB