



新浪CDN自动化运维

研发中心-平台架构部-刘宇

2013年03月30日

I

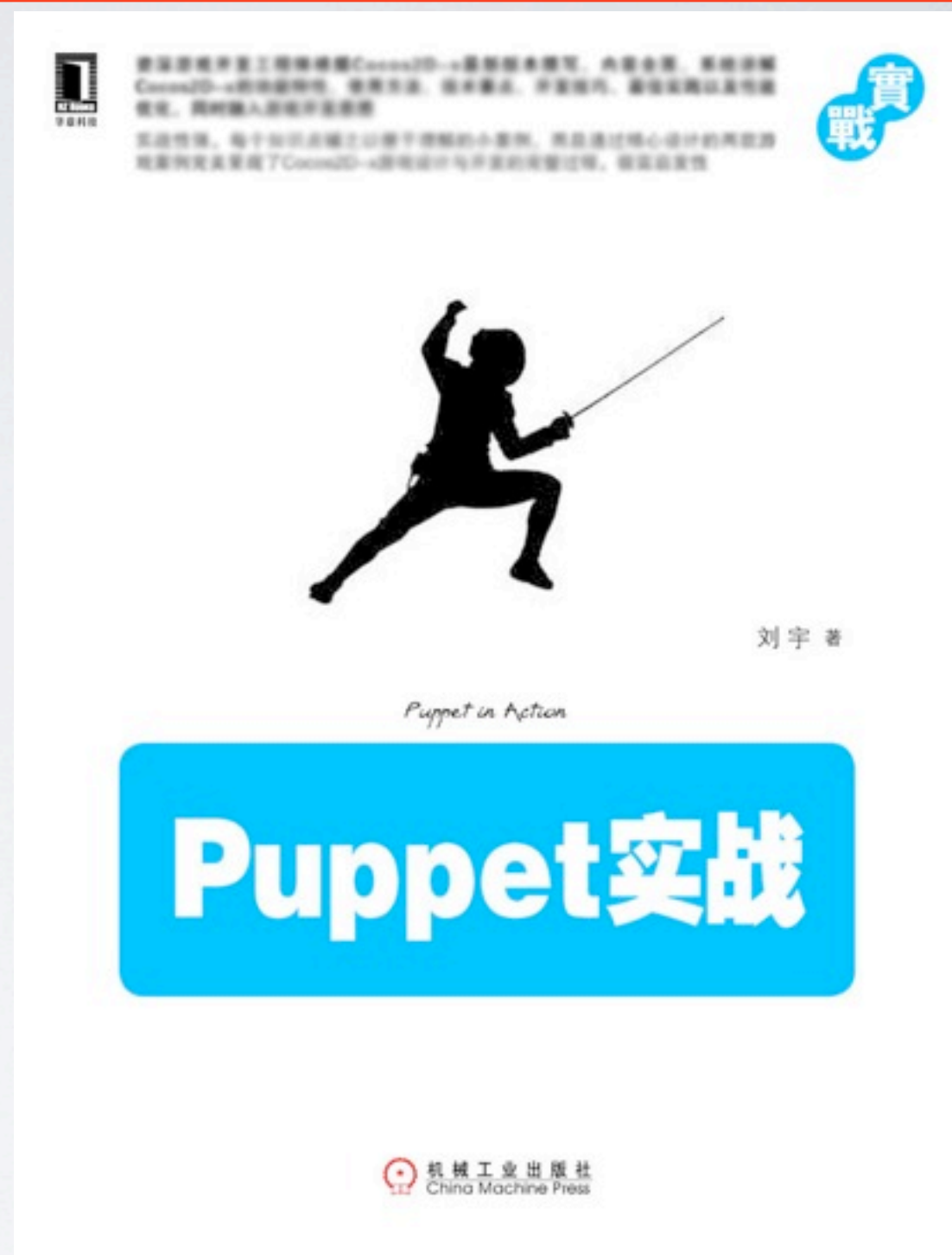
个人介绍

·刘宇 at 新浪

·码 《Puppet实战》 中

·社区: LinuxTone.org

·微博: @守住每一天



我叫刘宇，就职于新浪。去年受华章图片杨福川主编邀请，正在撰写《Puppet实战》一书，预计今年7月份将由机械工业出版社出版。

个人利用业余时间和朋友一起创办Linuxtone.org社区，致力推广“学习是一种信仰，分享是一种快乐！”也希望更多的人参与到我们社区。

欢迎大家通过新浪微博@守住每一天 与我进行线上交流。

平台介绍



新浪内容加速平台简称新浪CDN，平台诞生的目的不是为了替代商业CDN，而是为了和商业CDN进行更完美的整合，更适应公司的业务特征。与商业CDN相互补充，做到成本与质量最优化。

目前平台所提供的加速有静态元素加速、流媒体直播加速以及高速下载。

从所周知，CDN即Content Delivery Network，即内容分发网络，基本思路就是避开互联网上有可能影响数据传输速度和稳定性的瓶颈和环节，其特点是节点分散，跨多个运营商。

为此平台规模承载了300G带宽，覆盖全国7个ISP 40多个IDC，服务器的规模也有上千台。

平台介绍

· 静态元素加速



新浪内容加速平台简称新浪CDN，平台诞生的目的不是为了替代商业CDN，而是为了和商业CDN进行更完美的整合，更适应公司的业务特征。与商业CDN相互补充，做到成本与质量最优化。

目前平台所提供的加速有静态元素加速、流媒体直播加速以及高速下载。

从所周知，CDN即Content Delivery Network，即内容分发网络，基本思路就是避开互联网上有可能影响数据传输速度和稳定性的瓶颈和环节，其特点是节点分散，跨多个运营商。

为此平台规模承载了300G带宽，覆盖全国7个ISP 40多个IDC，服务器的规模也有上千台。

平台介绍

- 静态元素加速
- 流媒体直播加速



新浪内容加速平台简称新浪CDN，平台诞生的目的不是为了替代商业CDN，而是为了和商业CDN进行更完美的整合，更适应公司的业务特征。与商业CDN相互补充，做到成本与质量最优化。

目前平台所提供的加速有静态元素加速、流媒体直播加速以及高速下载。

从所周知，CDN即Content Delivery Network，即内容分发网络，基本思路就是避开互联网上有可能影响数据传输速度和稳定性的瓶颈和环节，其特点是节点分散，跨多个运营商。

为此平台规模承载了300G带宽，覆盖全国7个ISP 40多个IDC，服务器的规模也有上千台。

平台介绍

- 静态元素加速
- 流媒体直播加速
- 高速下载



新浪内容加速平台简称新浪CDN，平台诞生的目的不是为了替代商业CDN，而是为了和商业CDN进行更完美的整合，更适应公司的业务特征。与商业CDN相互补充，做到成本与质量最优化。

目前平台所提供的加速有静态元素加速、流媒体直播加速以及高速下载。

从所周知，CDN即Content Delivery Network，即内容分发网络，基本思路就是避开互联网上有可能影响数据传输速度和稳定性的瓶颈和环节，其特点是节点分散，跨多个运营商。

为此平台规模承载了300G带宽，覆盖全国7个ISP 40多个IDC，服务器的规模也有上千台。

平台介绍

- 静态元素加速
- 流媒体直播加速
- 高速下载



新浪内容加速平台简称新浪CDN，平台诞生的目的不是为了替代商业CDN，而是为了和商业CDN进行更完美的整合，更适应公司的业务特征。与商业CDN相互补充，做到成本与质量最优化。

目前平台所提供的加速有静态元素加速、流媒体直播加速以及高速下载。

从所周知，CDN即Content Delivery Network，即内容分发网络，基本思路就是避开互联网上有可能影响数据传输速度和稳定性的瓶颈和环节，其特点是节点分散，跨多个运营商。

为此平台规模承载了300G带宽，覆盖全国7个ISP 40多个IDC，服务器的规模也有上千台。

平台介绍

- 静态元素加速
- 流媒体直播加速
- 高速下载



- 带宽300Gbps



新浪内容加速平台简称新浪CDN，平台诞生的目的不是为了替代商业CDN，而是为了和商业CDN进行更完美的整合，更适应公司的业务特征。与商业CDN相互补充，做到成本与质量最优化。

目前平台所提供的加速有静态元素加速、流媒体直播加速以及高速下载。

从所周知，CDN即Content Delivery Network，即内容分发网络，基本思路就是避开互联网上有可能影响数据传输速度和稳定性的瓶颈和环节，其特点是节点分散，跨多个运营商。

为此平台规模承载了300G带宽，覆盖全国7个ISP 40多个IDC，服务器的规模也有上千台。

平台介绍

- 静态元素加速
- 流媒体直播加速
- 高速下载
- 带宽300Gbps
- 服务器1000+



新浪内容加速平台简称新浪CDN，平台诞生的目的不是为了替代商业CDN，而是为了和商业CDN进行更完美的整合，更适应公司的业务特征。与商业CDN相互补充，做到成本与质量最优化。

目前平台所提供的加速有静态元素加速、流媒体直播加速以及高速下载。

从所周知，CDN即Content Delivery Network，即内容分发网络，基本思路就是避开互联网上有可能影响数据传输速度和稳定性的瓶颈和环节，其特点是节点分散，跨多个运营商。

为此平台规模承载了300G带宽，覆盖全国7个ISP 40多个IDC，服务器的规模也有上千台。

平台介绍

- 静态元素加速
- 流媒体直播加速
- 高速下载
- 带宽300Gbps
- 服务器1000+
- 7+ISP, 40+IDC



新浪内容加速平台简称新浪CDN，平台诞生的目的不是为了替代商业CDN，而是为了和商业CDN进行更完美的整合，更适应公司的业务特征。与商业CDN相互补充，做到成本与质量最优化。

目前平台所提供的加速有静态元素加速、流媒体直播加速以及高速下载。

从所周知，CDN即Content Delivery Network，即内容分发网络，基本思路就是避开互联网上有可能影响数据传输速度和稳定性的瓶颈和环节，其特点是节点分散，跨多个运营商。

为此平台规模承载了300G带宽，覆盖全国7个ISP 40多个IDC，服务器的规模也有上千台。

平台挑战

平台挑战

- 快速扩张（每次上百台）
- 快速迭代更新（每天3-5次，每次5分钟）

平台挑战

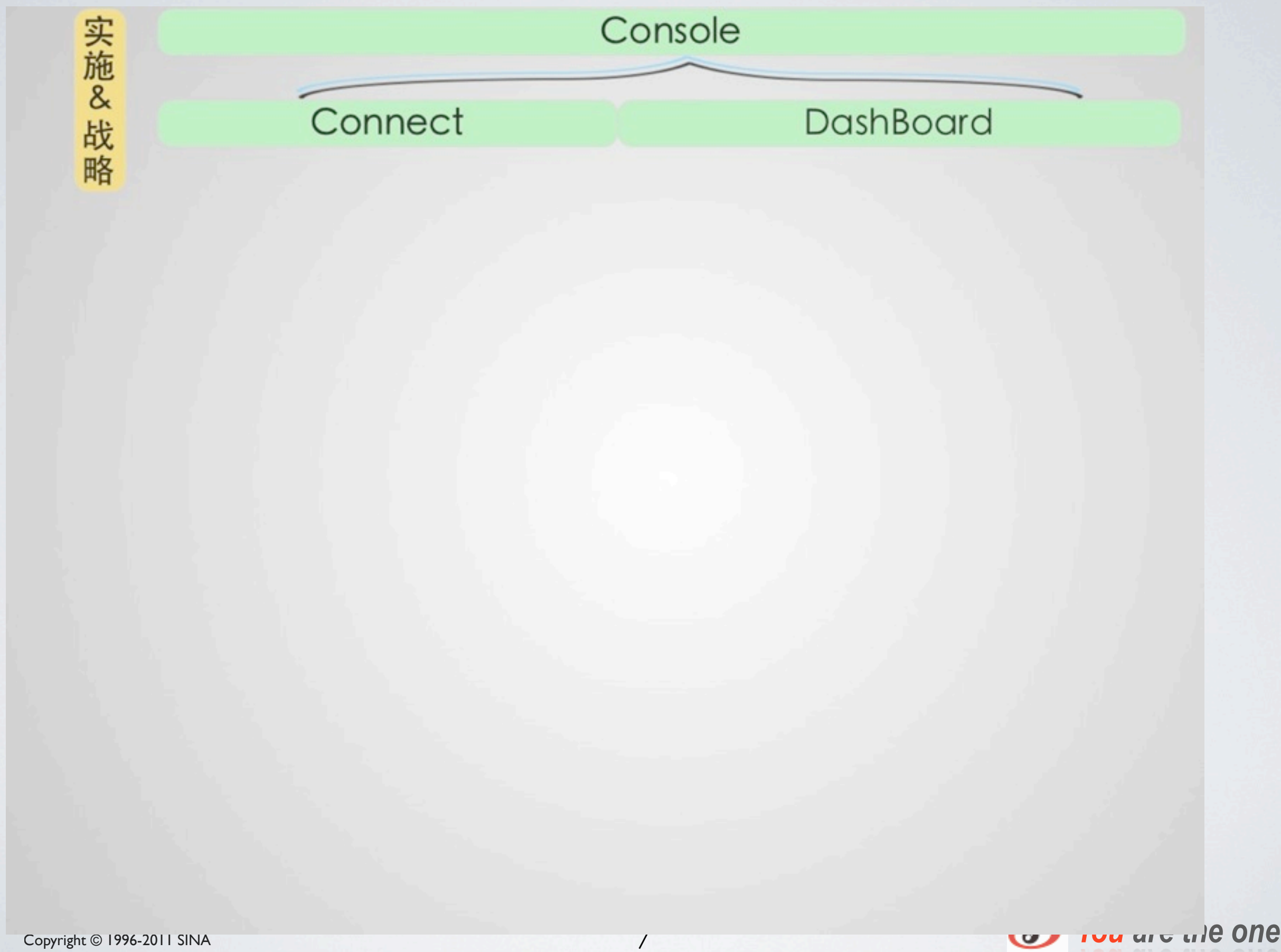
平台挑战

- 运维效率、质量及成本（较少的人工干预）
- 监控完善与故障的快速响应（多路探测机制，架构容错）



Puppet

自动化

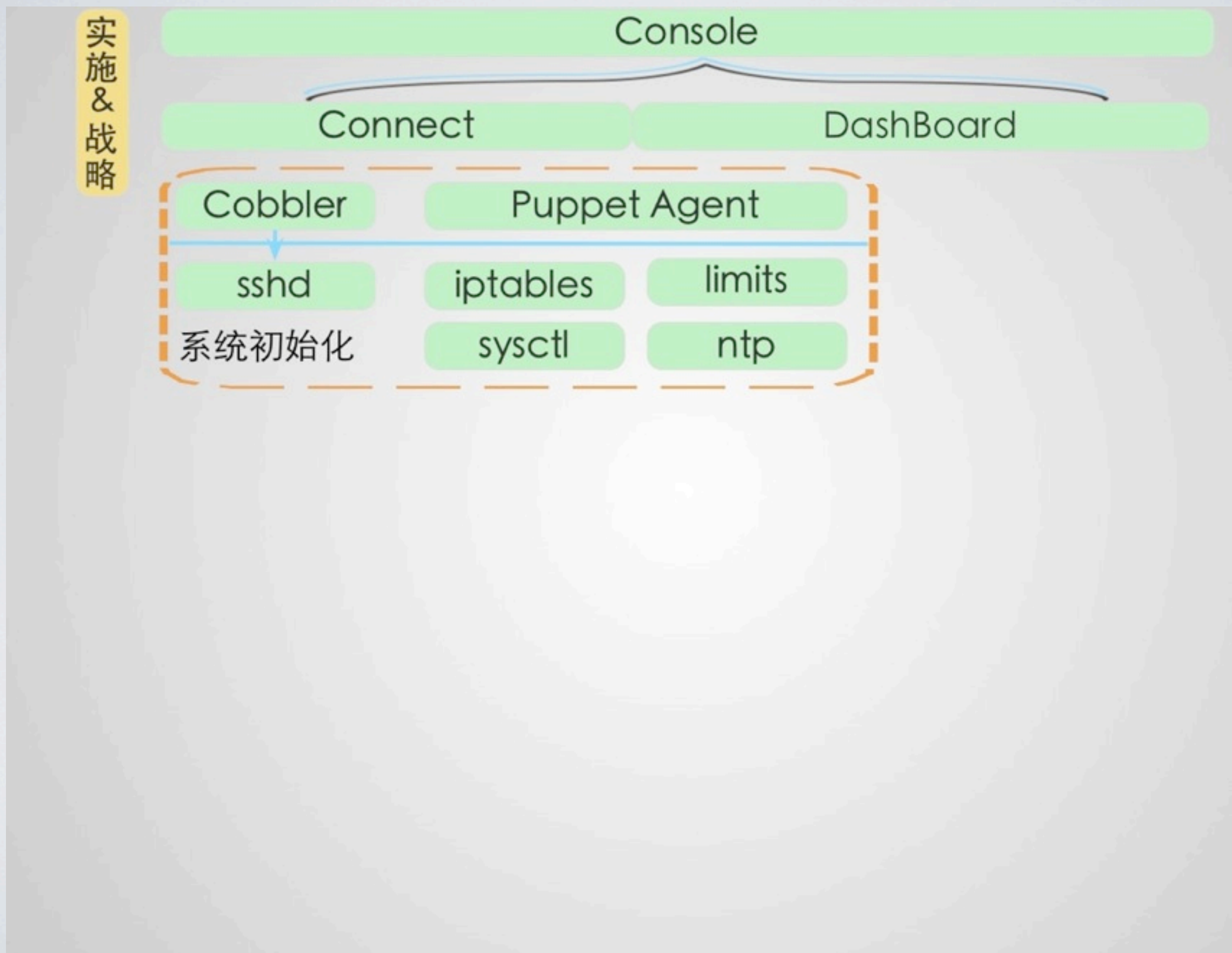


13年4月22日星期一

在自动化的实施与战略方面，我们有一个集中的Console来管理，它主要分为二块，Connect 关系型管理，DashBoard报告管理。

通过系统安装时安装好Puppet客户端，在Console中对主机，或一组主机进行关系定义后，自动完成系统初始化包括ssh iptables ntp等。客户端根据配置文件定义的角色，加载不同的模块完成软件的安装和配置，然后将结果汇报给Console。像监控及其它的部署也将通过Puppet来实成。

自动化



Copyright © 1996-2011 SINA

/

 You are the one

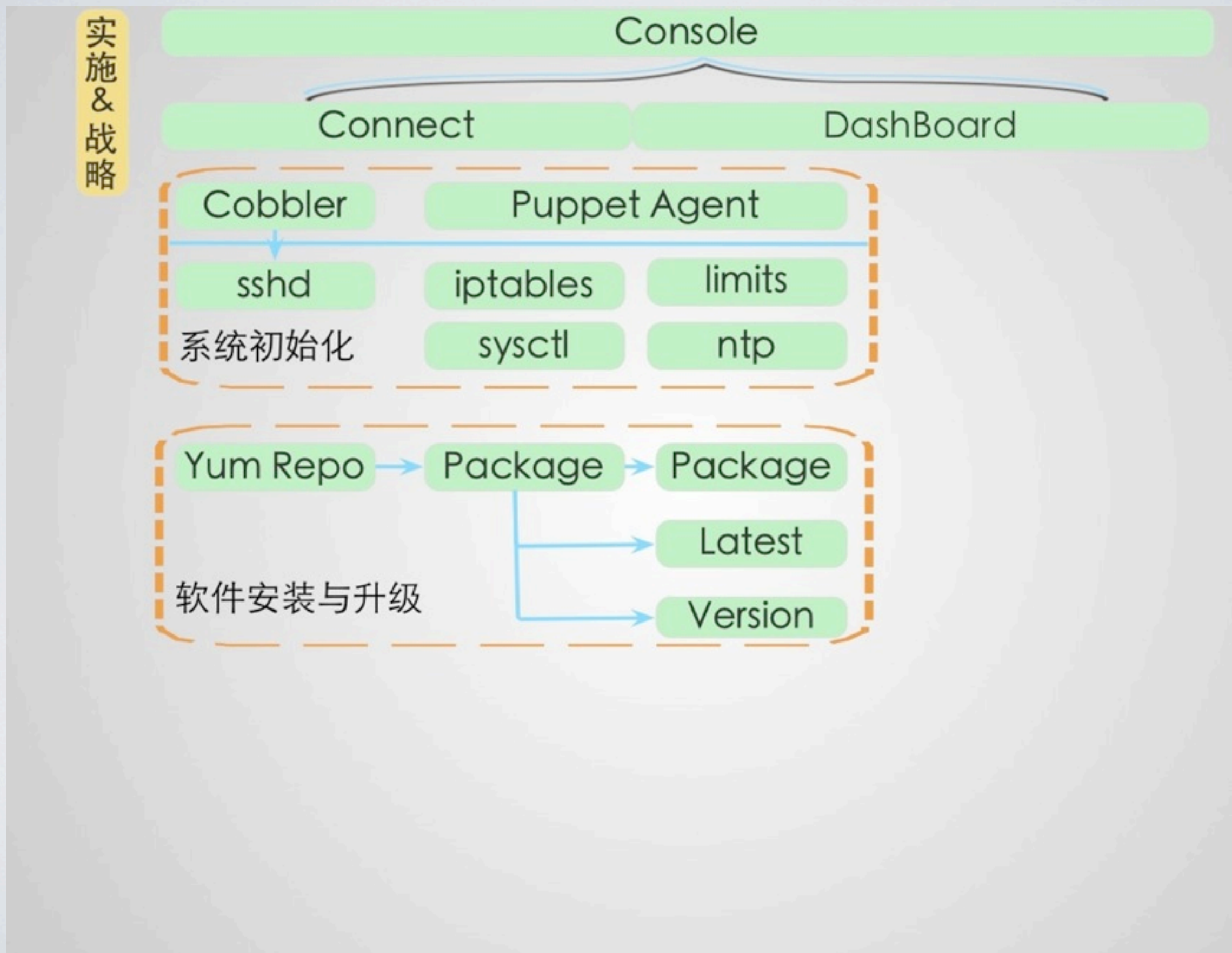
13年4月22日星期一

7

在自动化的实施与战略方面，我们有一个集中的Console来管理，它主要分为二块，Connect 关系型管理，DashBoard报告管理。

通过系统安装时安装好Puppet客户端，在Console中对主机，或一组主机进行关系定义后，自动完成系统初始化包括ssh iptables ntp等。客户端根据配置文件定义的角色，加载不同的模块完成软件的安装和配置，然后将结果汇报给Console。像监控及其它的部署也将通过Puppet来实成。

自动化



Copyright © 1996-2011 SINA

/

 You are the one

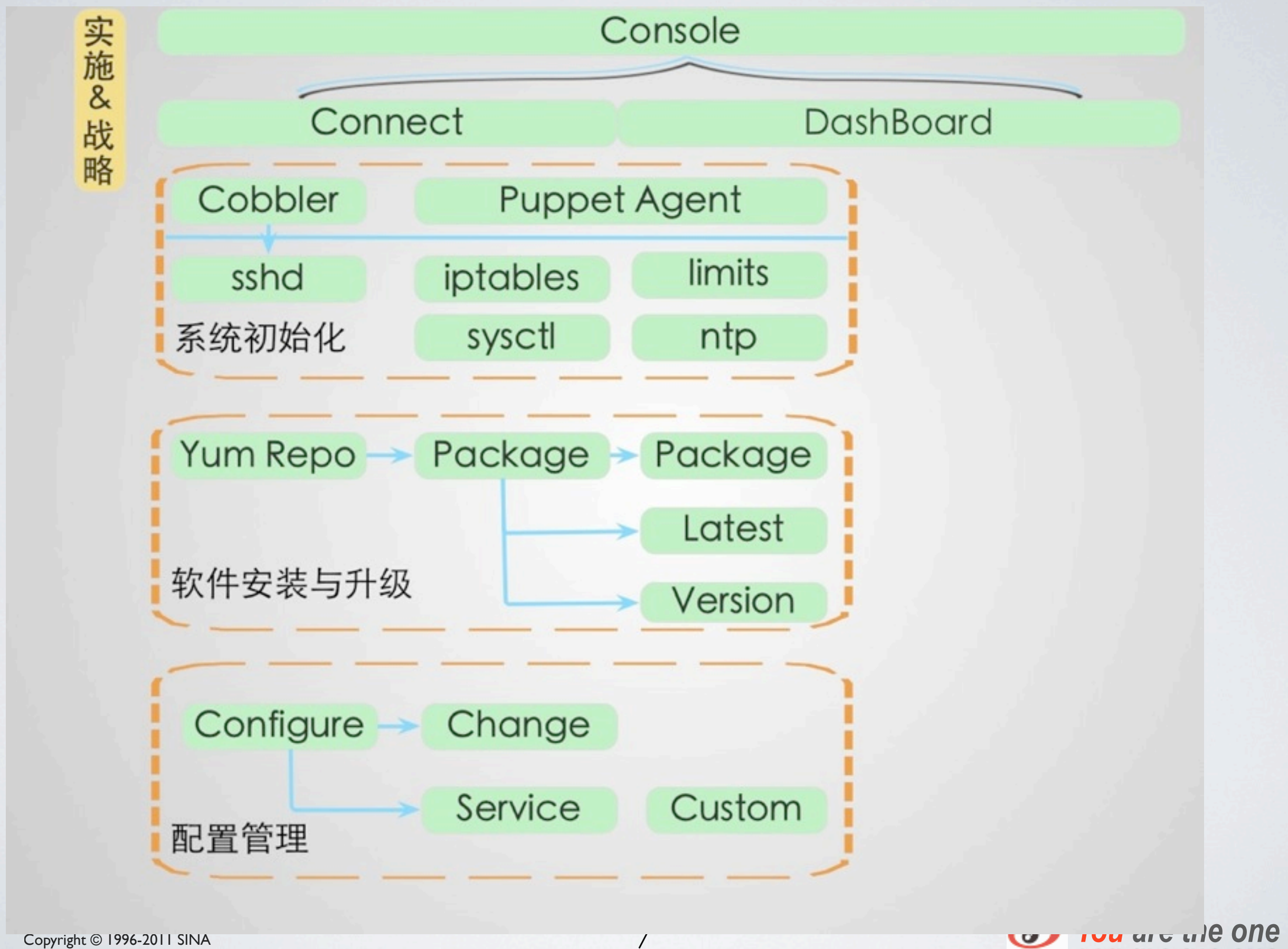
13年4月22日星期一

7

在自动化的实施与战略方面，我们有一个集中的Console来管理，它主要分为二块，Connect 关系型管理，DashBoard报告管理。

通过系统安装时安装好Puppet客户端，在Console中对主机，或一组主机进行关系定义后，自动完成系统初始化包括ssh iptables ntp等。客户端根据配置文件定义的角色，加载不同的模块完成软件的安装和配置，然后将结果汇报给Console。像监控及其它的部署也将通过Puppet来实成。

自动化

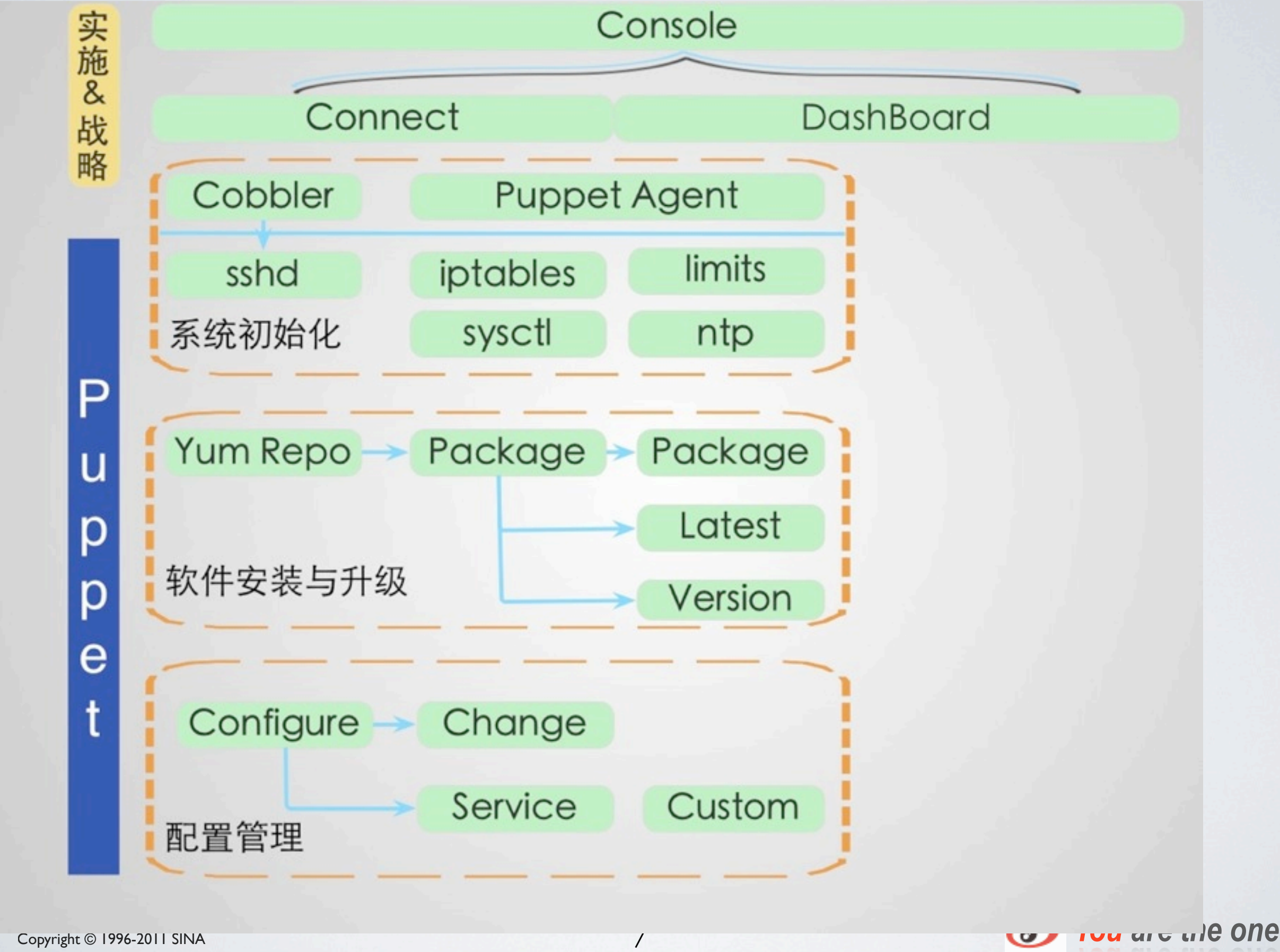


13年4月22日星期一

在自动化的实施与战略方面，我们有一个集中的Console来管理，它主要分为二块，Connect 关系型管理，DashBoard报告管理。

通过系统安装时安装好Puppet客户端，在Console中对主机，或一组主机进行关系定义后，自动完成系统初始化包括ssh iptables ntp等。客户端根据配置文件定义的角色，加载不同的模块完成软件的安装和配置，然后将结果汇报给Console。像监控及其它的部署也将通过Puppet来实成。

自动化

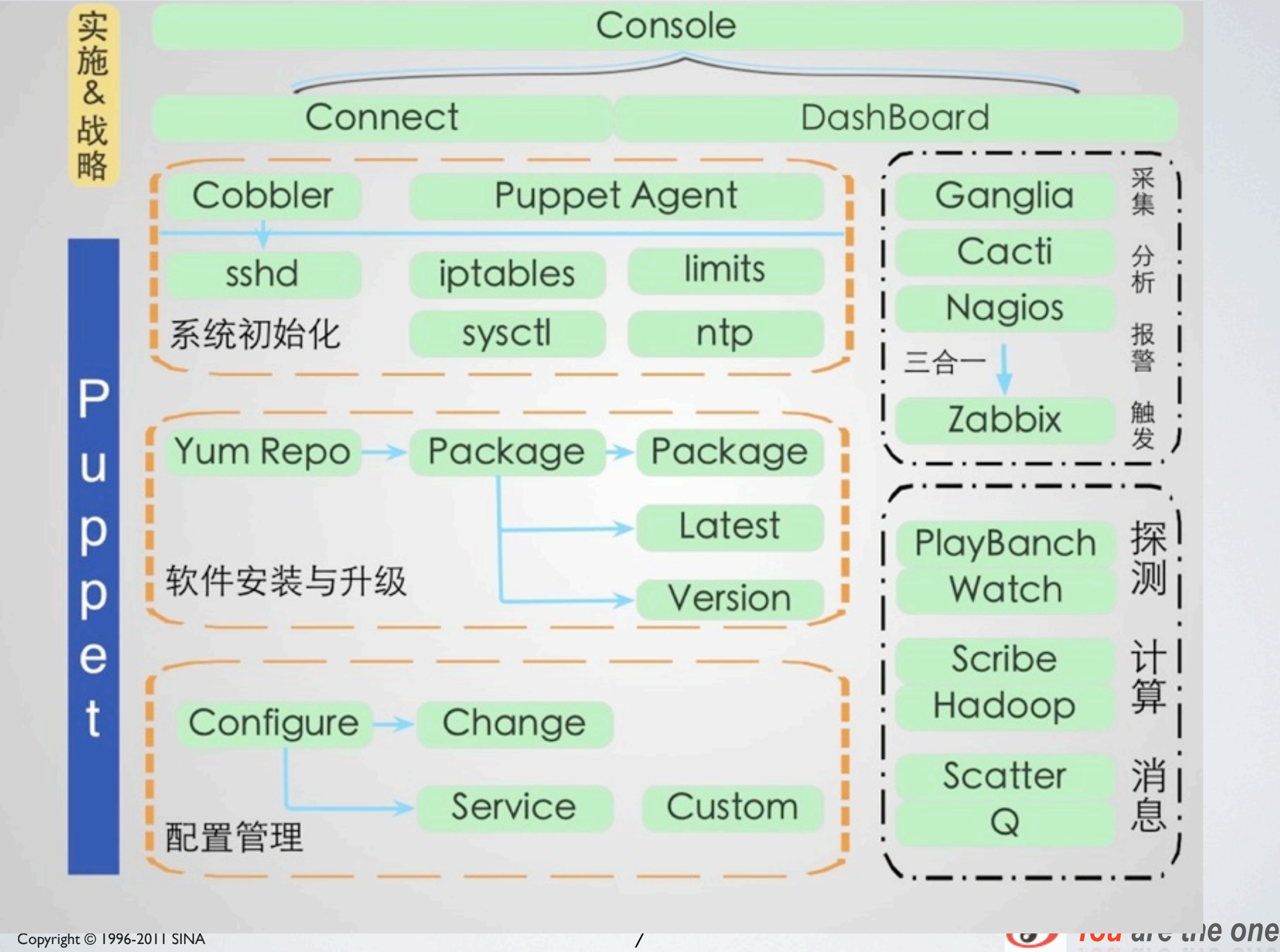


13年4月22日星期一

在自动化的实施与战略方面，我们有一个集中的Console来管理，它主要分为二块，Connect 关系型管理，DashBoard报告管理。

通过系统安装时安装好Puppet客户端，在Console中对主机，或一组主机进行关系定义后，自动完成系统初始化包括ssh iptables ntp等。客户端根据配置文件定义的角色，加载不同的模块完成软件的安装和配置，然后将结果汇报给Console。像监控及其它的部署也将通过Puppet来实成。

自动化



在自动化的实施与战略方面，我们有一个集中的Console来管理，它主要分为二块，Connect 关系型管理，DashBoard报告管理。

通过系统安装时安装好Puppet客户端，在Console中对主机，或一组主机进行关系定义后，自动完成系统初始化包括ssh iptables ntp等。客户端根据配置文件定义的角色，加载不同的模块完成软件的安装和配置，然后将结果汇报给Console。像监控及其它的部署也将通过Puppet来实成。

主机名规范

在定义这些关系时，主要通过主机名规范来实现。通过dns反解，直接使用域名登录，报警时直接知道业务情况，值班可根据主机名相关规则决定是不是叫你起床。

可以根据任意关键字进行匹配，在Puppet 节点定义与判断时非常有用

主机名规范

• 当前主机名：

- ISP.IDC.Group.Role.ip4.domain.com

在定义这些关系时，主要通过主机名规范来实现。通过dns反解，直接使用域名登录，报警时直接知道业务情况，值班可根据主机名相关规则决定是不是叫你起床。

可以根据任意关键字进行匹配，在Puppet 节点定义与判断时非常有用

主机名规范

• 当前主机名：

- ISP.IDC.Group.Role.ip4.domain.com

• 将变更为：

- ISP.IDC.City.Group.Role.Product.domain.com

主机名规范.与-的区别

• PS1 提示符:

- \H ISP.IDC.City.Group.Role.Product.domain.com
- \h ISP

Puppet配置

```
file { ["/home/${username}/.bash_profile":  
  source => "puppet:///modules/user/${username}-bashprofile",  
  owner  => $username,  
  group  => $username,  
]
```

```
[root@cnc.tc.bj.test.ab.cdn.domain.com ~]  
0> █
```

单台主机

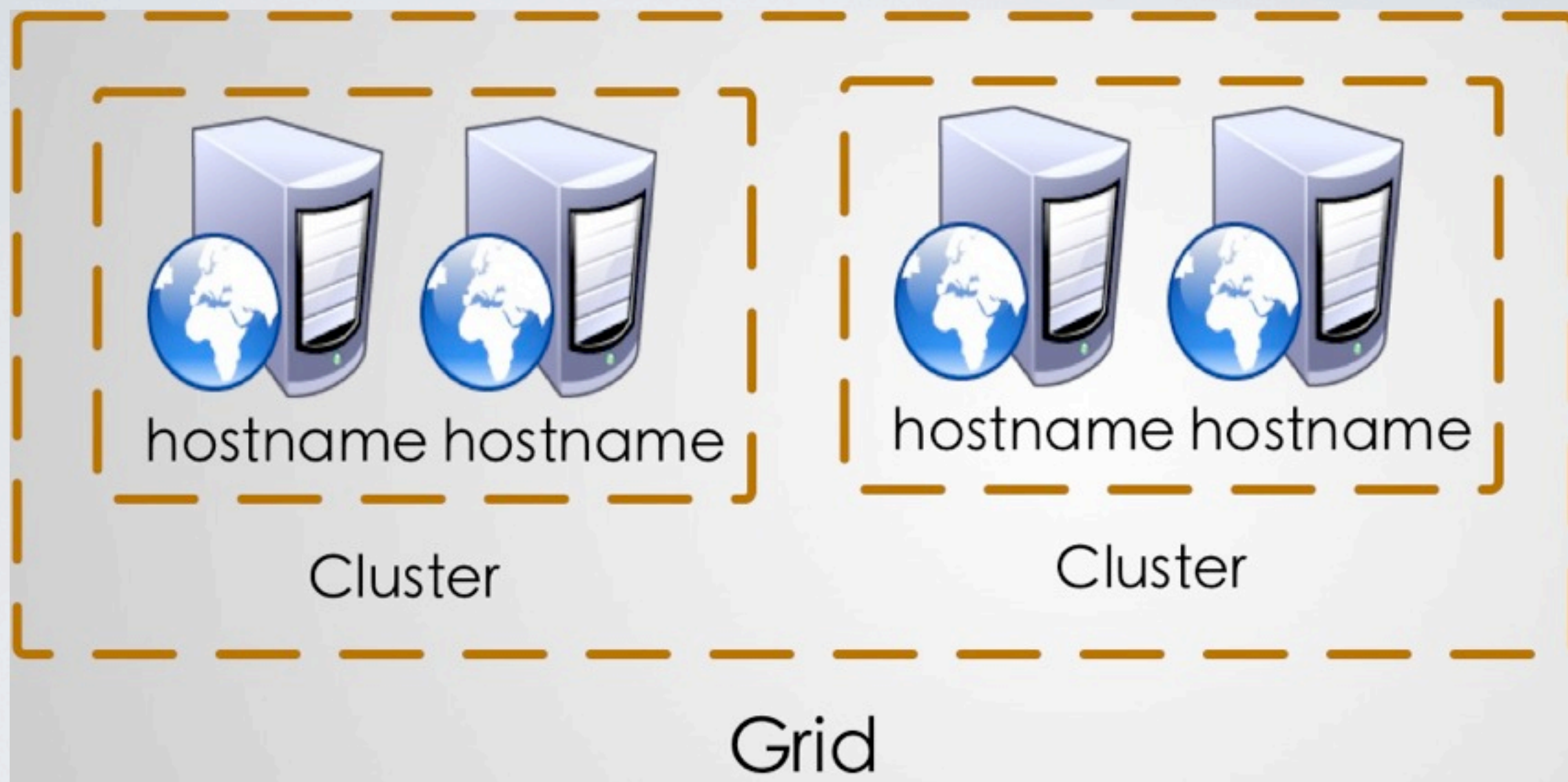


ISP.IDC.City.Group.Role.Product.domain.com

集群



全局定义



变更与发布

Puppet中服务配置

•  /etc/puppet/modules

- systemg
 - sshd
 - iptables
 - sysctl
 - ntp
- nginx
- zabbix
- scribe

Puppet中节点定义

Puppet中节点定义

•🔗• `/etc/puppet/manifests/nodes`

Puppet中节点定义

-  /etc/puppet/manifests/nodes
 - Grid

Puppet中节点定义

-  /etc/puppet/manifests/nodes
 - Grid
 - Cluster

Puppet中节点定义

•  /etc/puppet/manifests/nodes

- Grid
 - Cluster
 - Node1

Puppet中节点定义

•  /etc/puppet/manifests/nodes

- Grid
 - Cluster
 - Node1
 - Node2

Puppet中节点定义

• /etc/puppet/manifests/nodes

- Grid

- Cluster

- Node1

- Node2

• /etc/puppet/manifests/nodes/grid/cluster/node1.pp

Puppet中节点定义

• /etc/puppet/manifests/nodes

- Grid
 - Cluster
 - Node1
 - Node2

• /etc/puppet/manifests/nodes/grid/cluster/node1.pp

- include systemg

Puppet中节点定义

• /etc/puppet/manifests/nodes

- Grid
 - Cluster
 - Node1
 - Node2

• /etc/puppet/manifests/nodes/grid/cluster/node1.pp

- include systemg
- include nginx

Puppet中节点定义

• /etc/puppet/manifests/nodes

- Grid
 - Cluster
 - Node1
 - Node2

• /etc/puppet/manifests/nodes/grid/cluster/node1.pp

- include systemg
- include nginx
- include zabbix

Puppet中节点定义

• /etc/puppet/manifests/nodes

- Grid
 - Cluster
 - Node1
 - Node2

• /etc/puppet/manifests/nodes/grid/cluster/node1.pp

- include systemg
- include nginx
- include zabbix
- include scribe

快速部署

·✎· 软件包发布

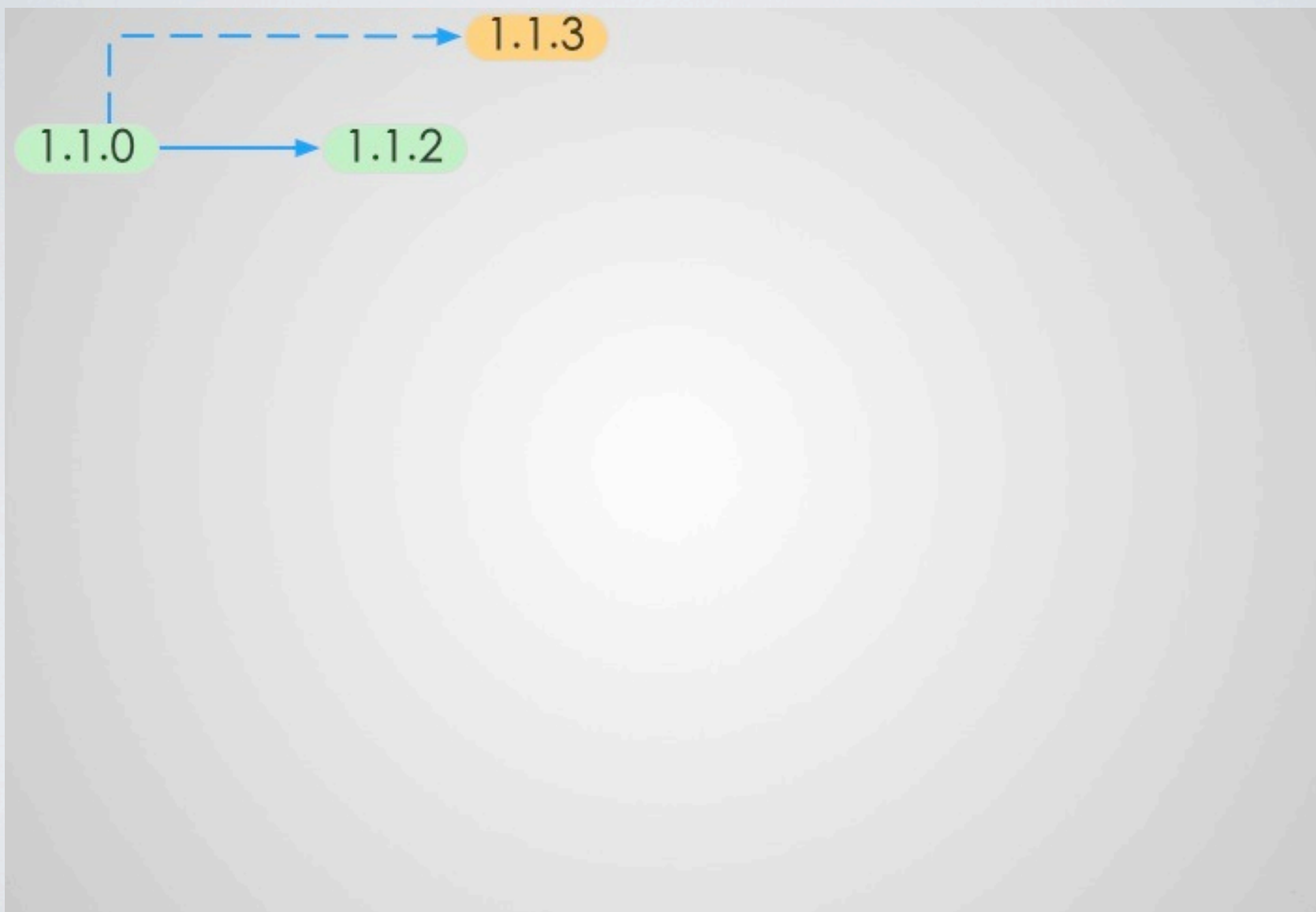
·✎· 配置文件变更

软件包更新，代码的发布都需要通过RPM包来实现，我们不单独针对某个代码进行更新。
配置文件更新主要是新客户的增加。

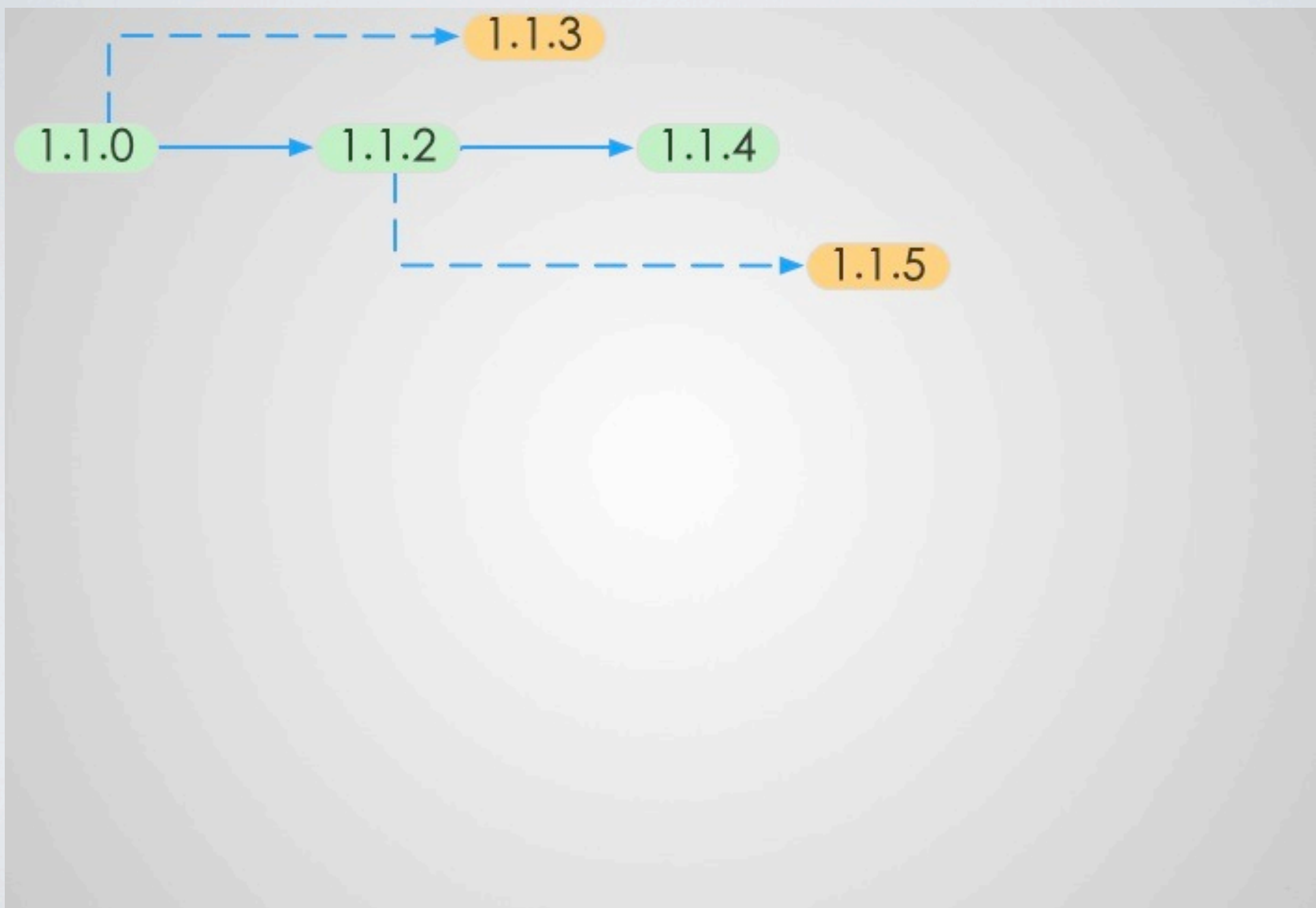
软件包升级/回滚



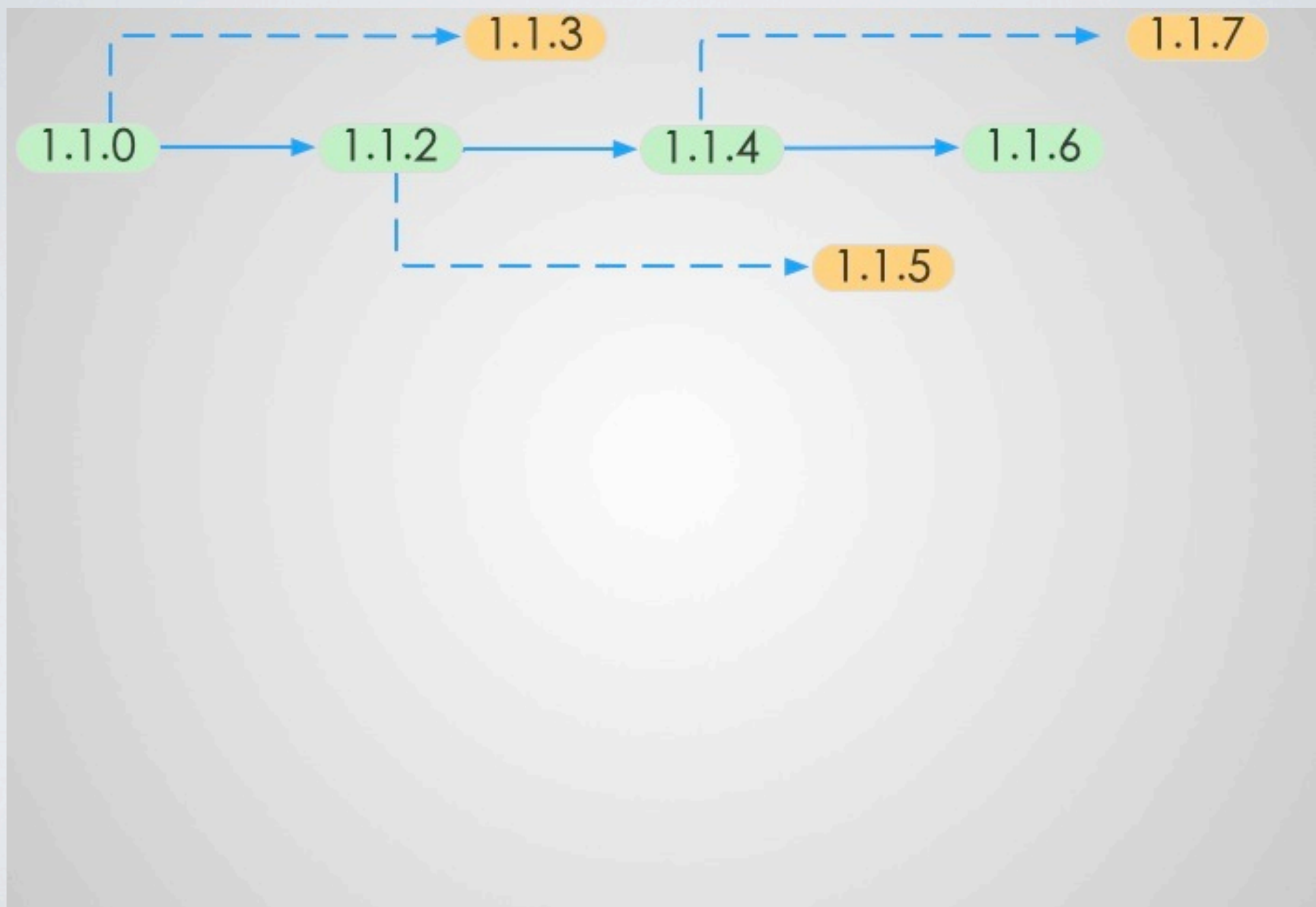
软件包升级/回滚



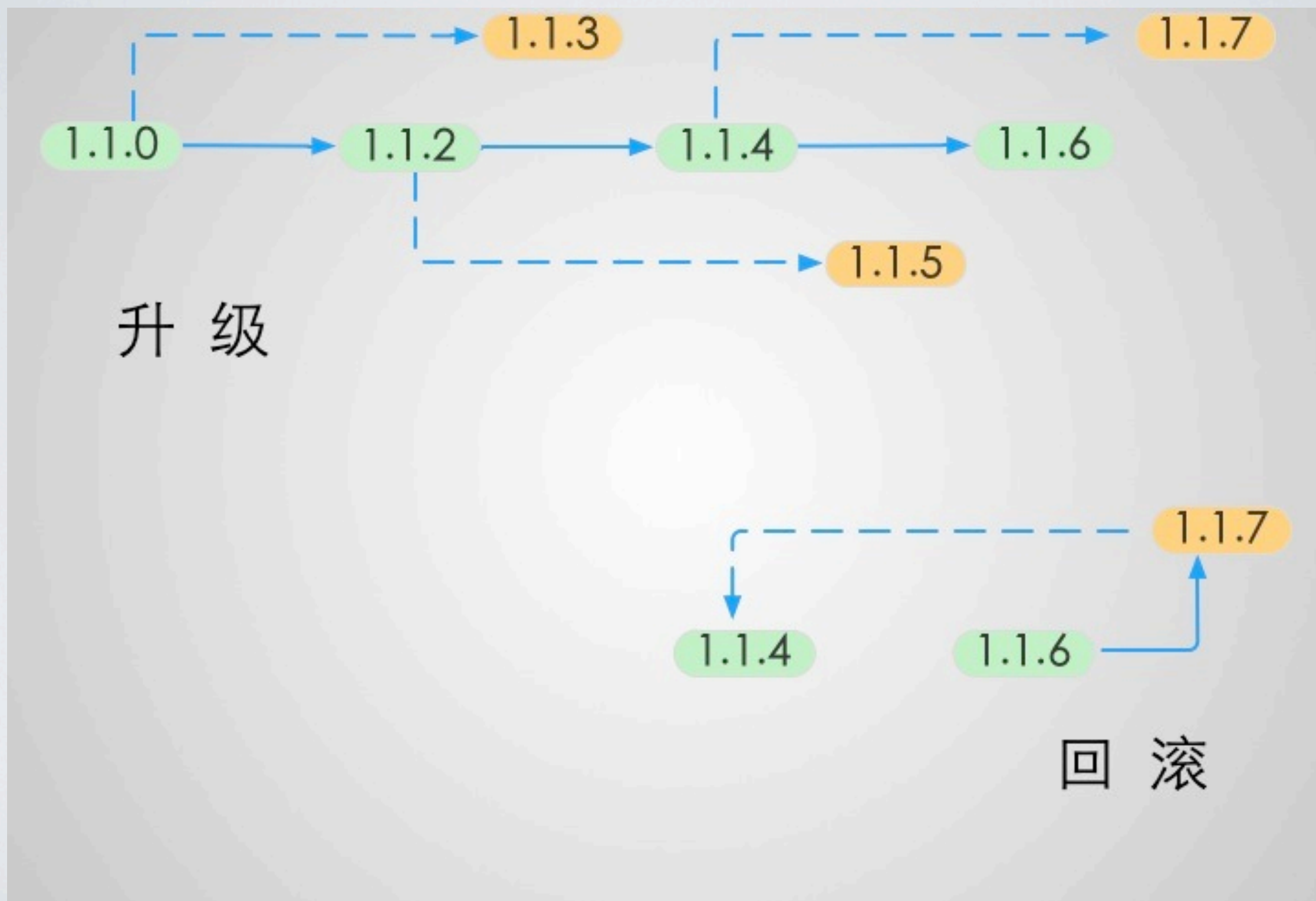
软件包升级/回滚



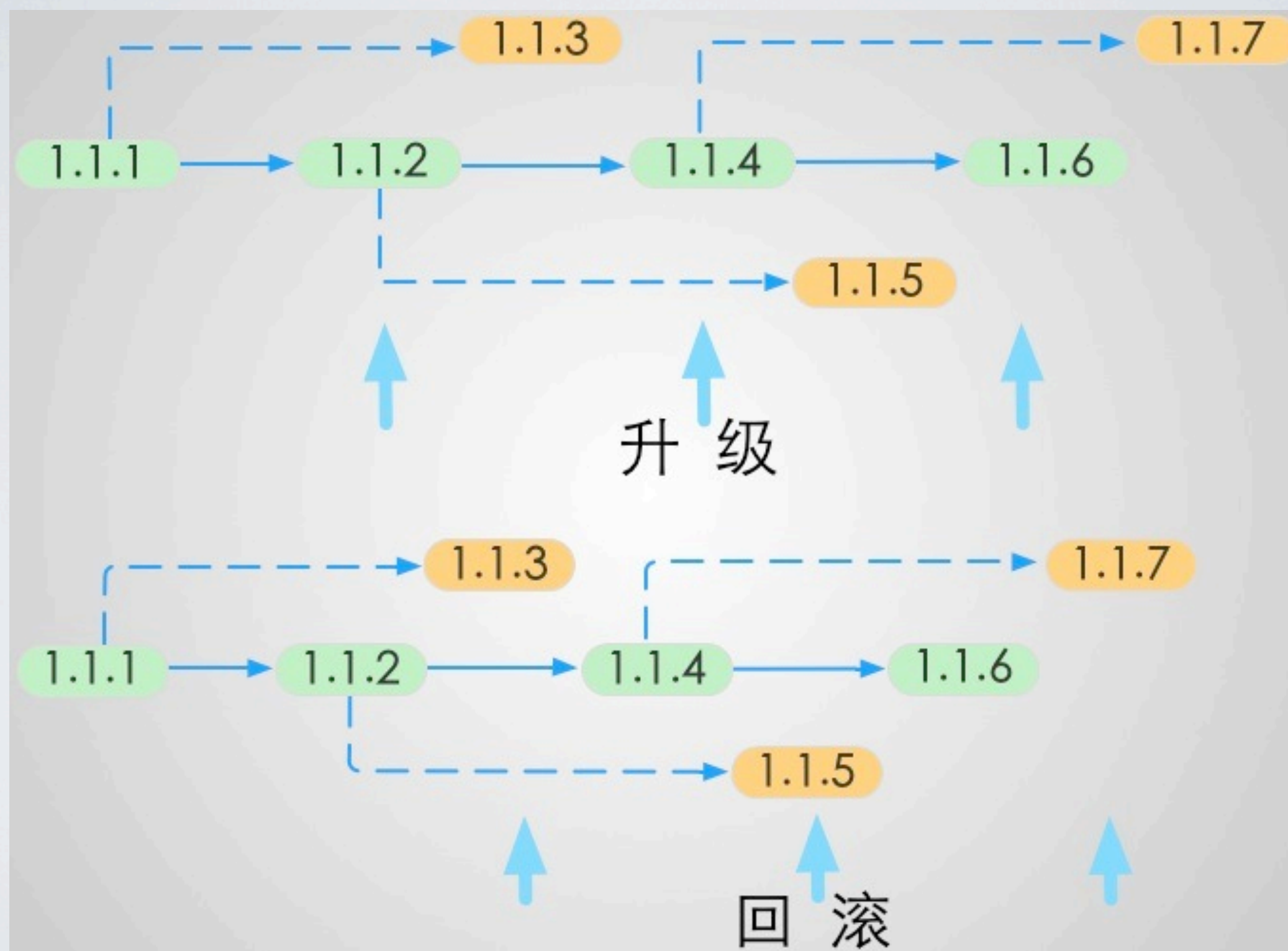
软件包升级/回滚



软件包升级/回滚

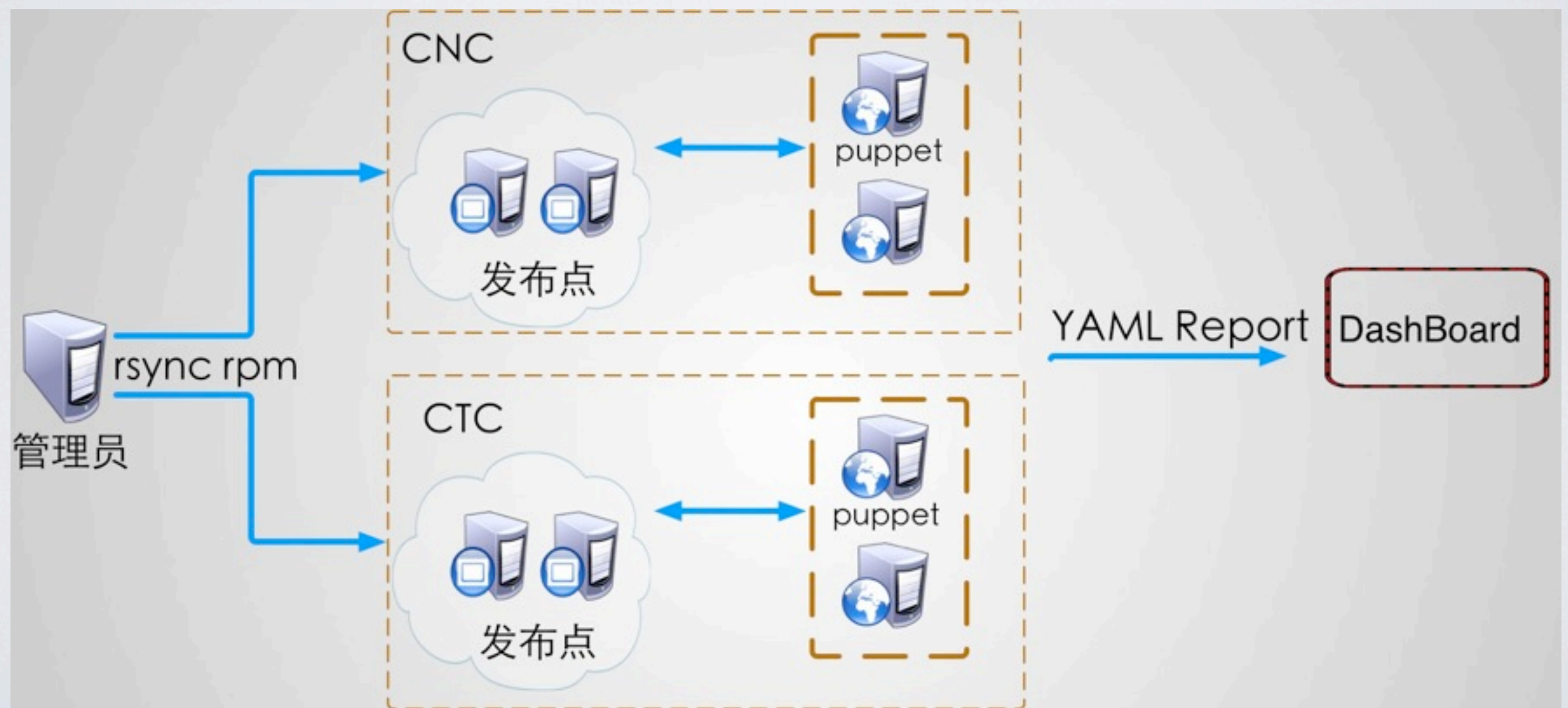


软件包升级/回滚



软件包发布

- nginx-1.1.2
- `fpm -s dir -t rpm -n nginx -v 1.1.2 /tmp/installdir/`
- `package { '$name': ensure => latest }`



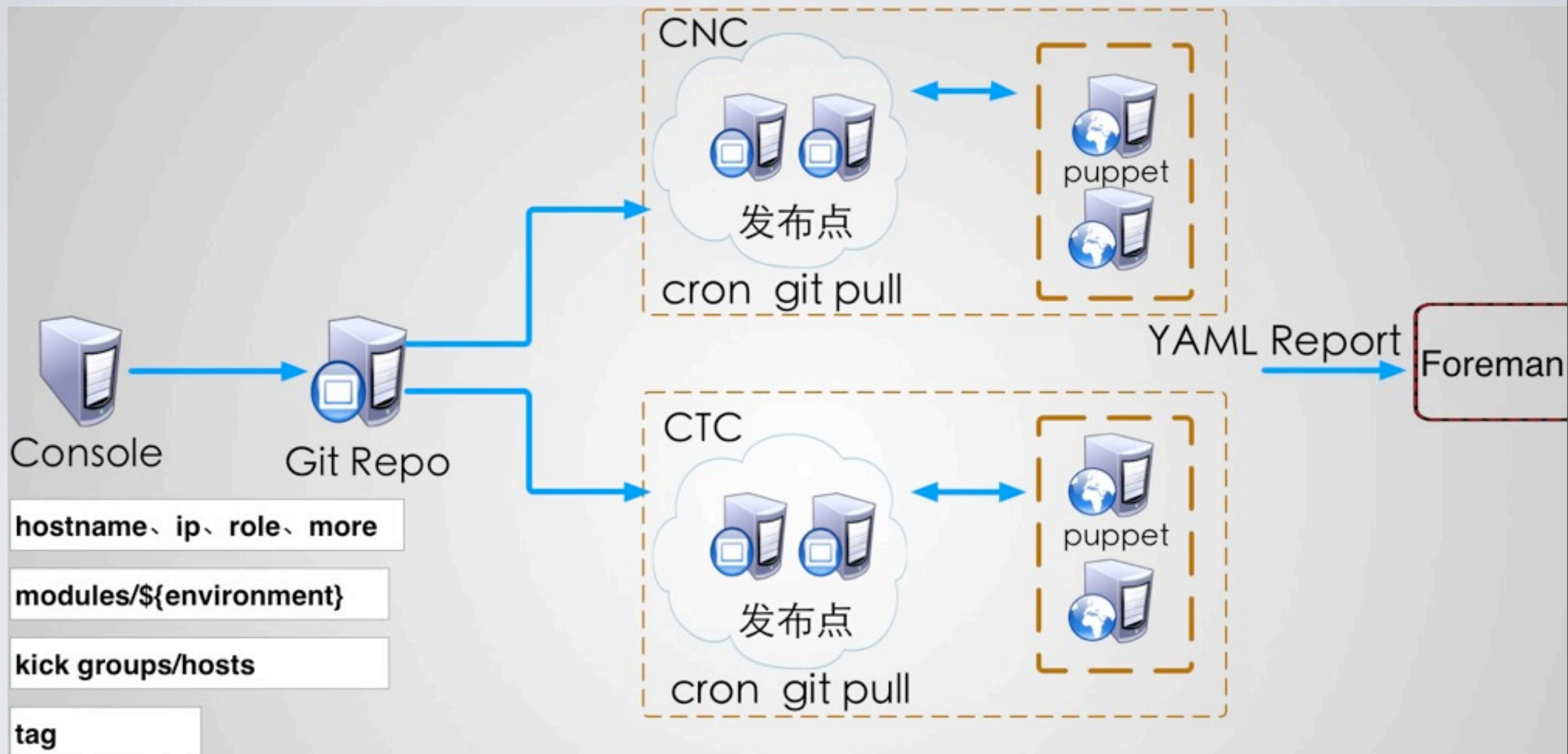
软件包的发布主要是通过将软件打包成rpm。

sina底层操作系统非常统一，这给运维带来了极大的便利

管理员在与线上环境一致的服务器上，将软件采用fpm进行打包，然后rsync到发布点，如果需要更新时采用puppet kick触发，或等puppet agent运行即可。

在puppet中配置package为latest，即为每次安装最新版

配置文件变更



配置文件通过git进行版本控制，默认由puppet agent主动触发更新。
如果有紧急情况时，利用puppet kick命令触发agent更新。
报告通过第三方foreman查看。

灰度发布

灰度发布



灰度发布主要涉及到版本的发布与配置的更新 版本发布主要是：开发层面，新功能上线，bugfix，

配置更新主要是：新应用增长，来新业务了

增加地区的发布策略：紧急发布5分钟内完成、超时2秒、失败的重试3次、重试失败的先忽略，待完成后统一处理

灰度发布

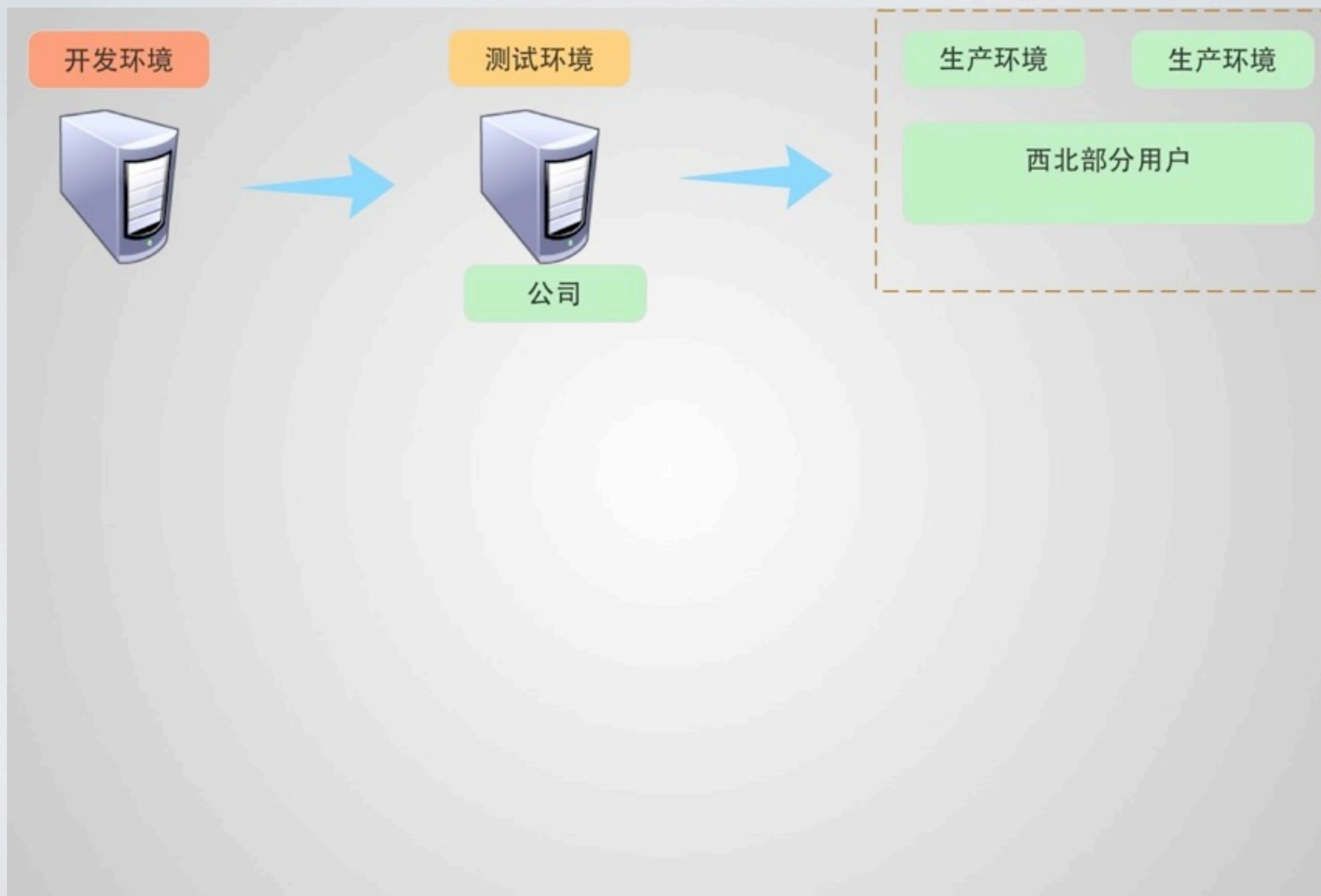


灰度发布主要涉及到版本的发布与配置的更新 版本发布主要是：开发层面，新功能上线，bugfix，

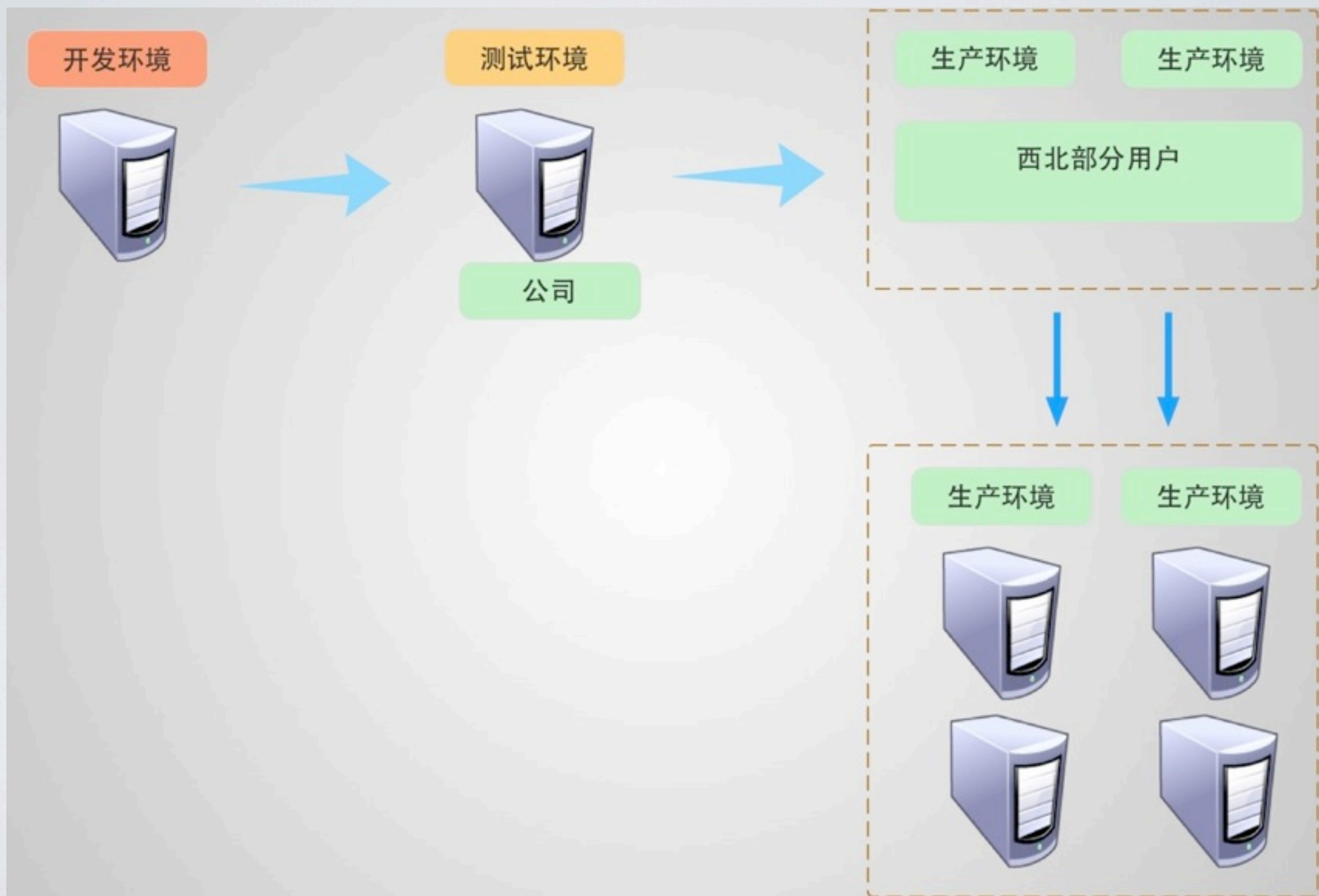
配置更新主要是：新应用增长，来新业务了

增加地区的发布策略：紧急发布5分钟内完成、超时2秒、失败的重试3次、重试失败的先忽略，待完成后统一处理

灰度发布



灰度发布

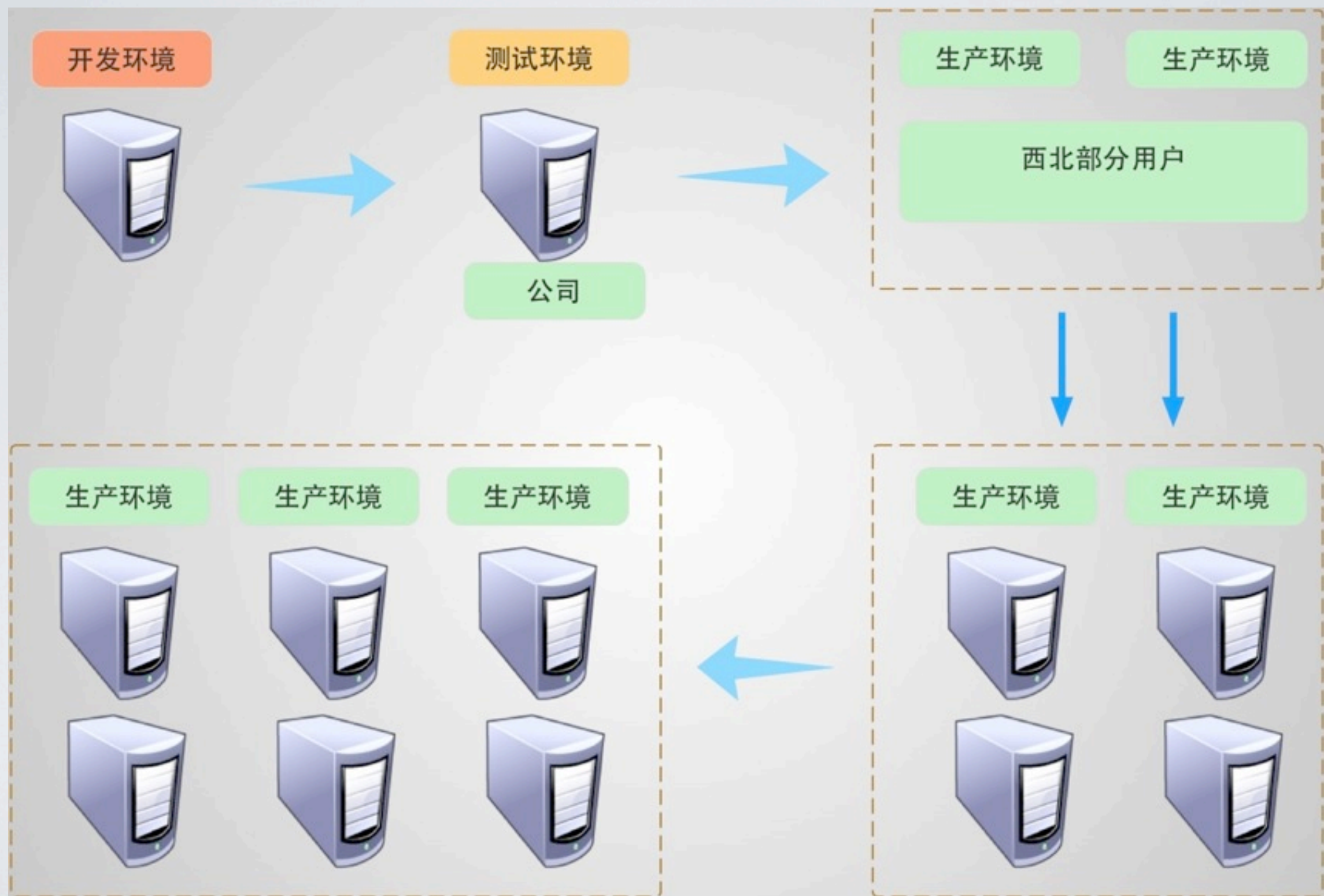


灰度发布主要涉及到版本的发布与配置的更新 版本发布主要是：开发层面，新功能上线，bugfix，

配置更新主要是：新应用增长，来新业务了

增加地区的发布策略：紧急发布5分钟内完成、超时2秒、失败的重试3次、重试失败的先忽略，待完成后统一处理

灰度发布



灰度发布主要涉及到版本的发布与配置的更新 版本发布主要是：开发层面，新功能上线，bugfix，

配置更新主要是：新应用增长，来新业务了

增加地区的发布策略：紧急发布5分钟内完成、超时2秒、失败的重试3次、重试失败的先忽略，待完成后统一处理

灰度发布-区域规则

Location:

Available location

Q理想

Choose all

Chosen location

QSelect your choice(s) and click →

中国-网通-理想国际

Clear all

Records

记录类型	Content	Ttl	权重	备注
A		60	100	测试环境

13年4月22日星期一
选择让公司的人一起做测试，将公司出口IP设计至测试环境

灰度发布-区域规则

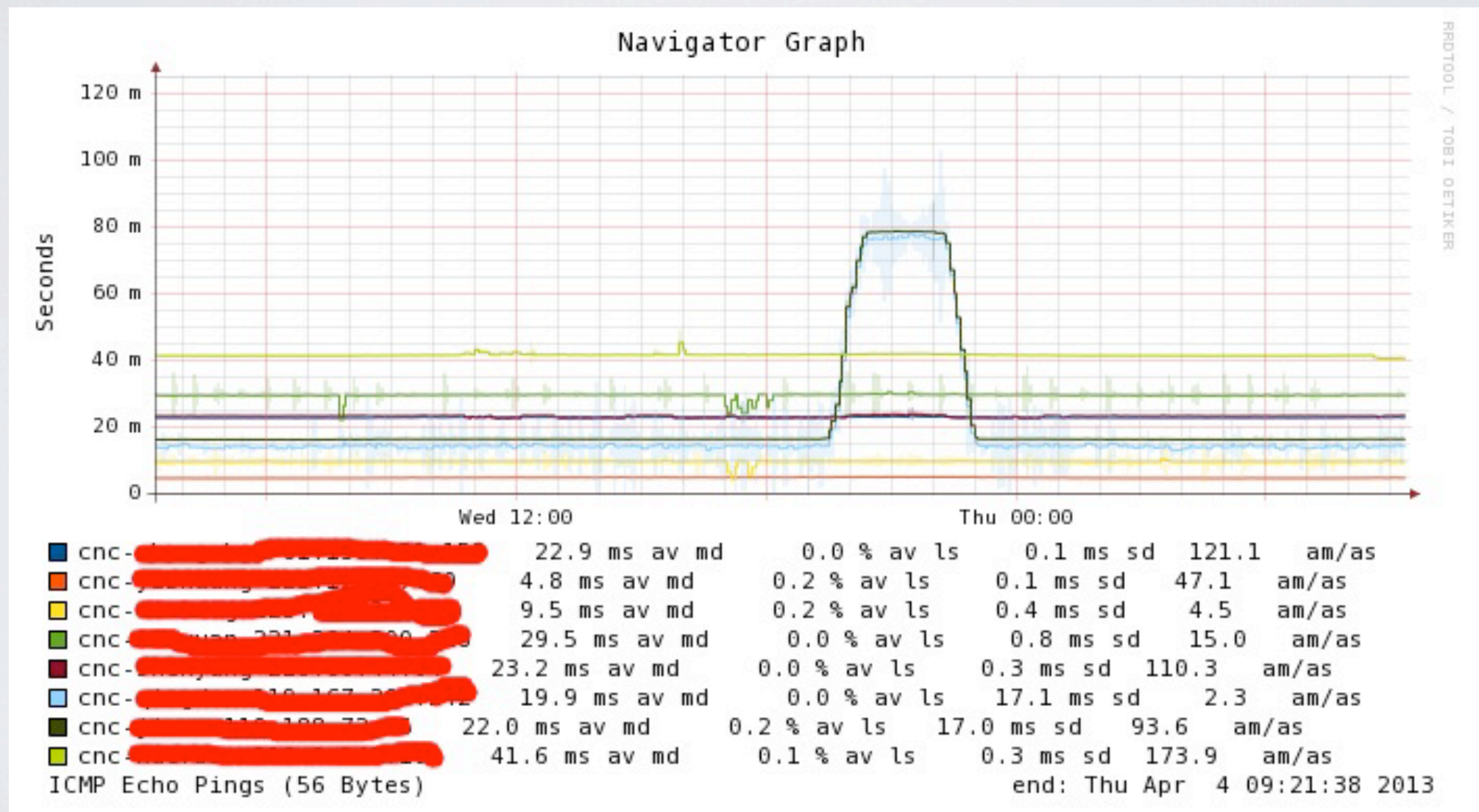
```
if ($fqdn == 'alpha_hostname')
{
file { 'puppet.conf':
    name => "/etc/puppet/puppet.conf",
    require => Package["puppet"],
    source => "puppet://$fileservers/files/puppet/puppet.conf.alpha"
}
}
elsif ($fqdn == 'beta_hostname') { ... }
```

```
puppet_agent.conf
[agent]
environment = alpha
```

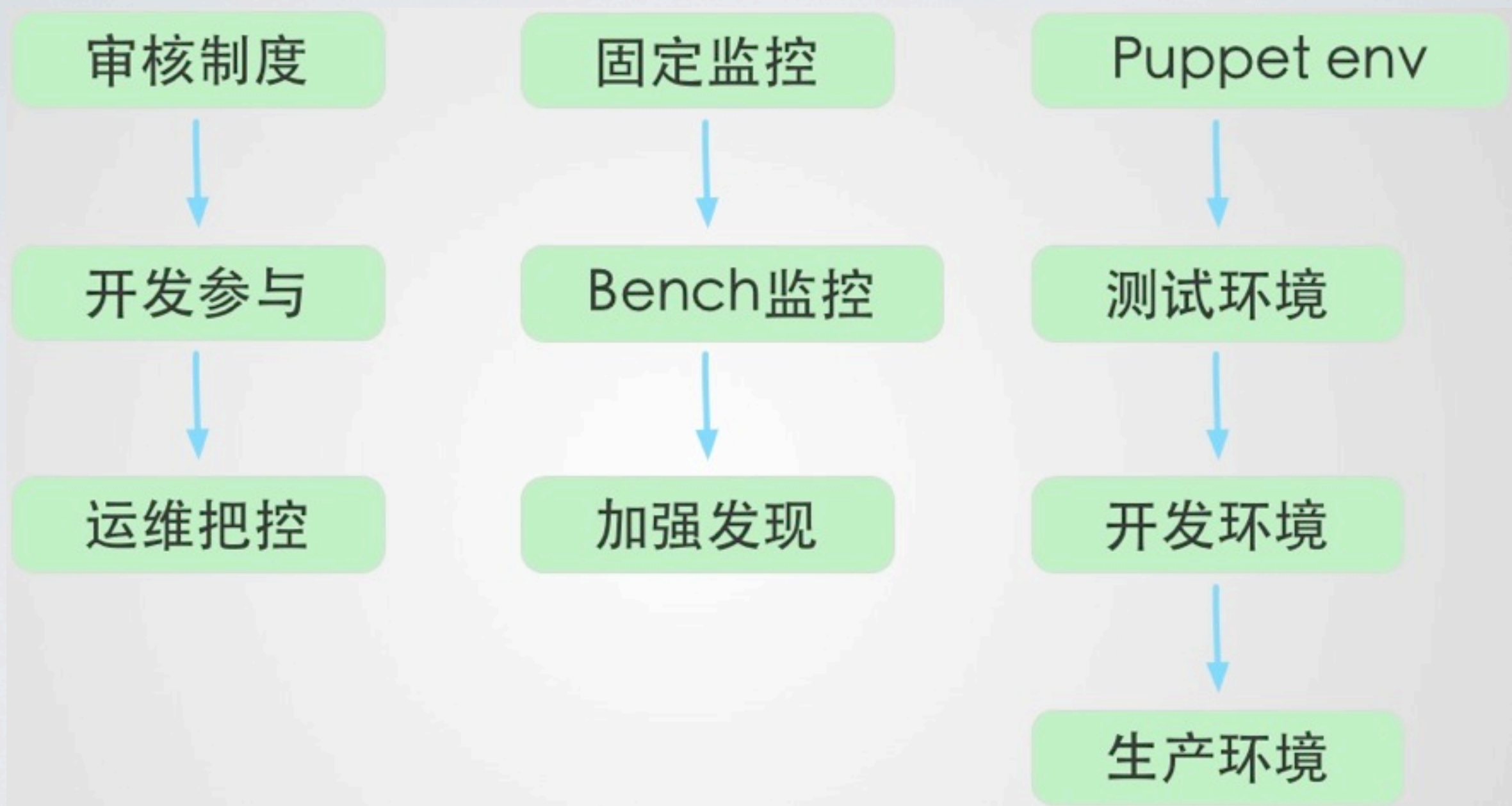
puppet 指定hosts的环境，通过这些达到灰度发布的目的。
缺点是每次都是固定的，如果需要变更时，需要再次操作。

灰度发布-发布点

- 收集全网延迟+丢包数据
- 及时更换发布点



灰度发布-错误检测



13年4月22日星期一

根据不同的业务，建议长期的探测机制，一旦有故障，及时报警。机制存在灰度的每个环节发布前全面审核，细到每个步骤与细节，可能对系统或业务产生的影响
交叉审核、开发全面参与灰度的每个阶段，减少故障的发生、运维人员必须全程跟进，

灰度发布-实践

- 🌀· 新功能上线时遗漏测试小应用
- 🌀· 2小时应用故障

灰度发布-实践

- 新功能上线时透露测试小应用
- 2小时应用故障
- 固定监控项也有失灵的时候
- 开发所有应用测试程序
- 延长重大变更测试周期

元自动化

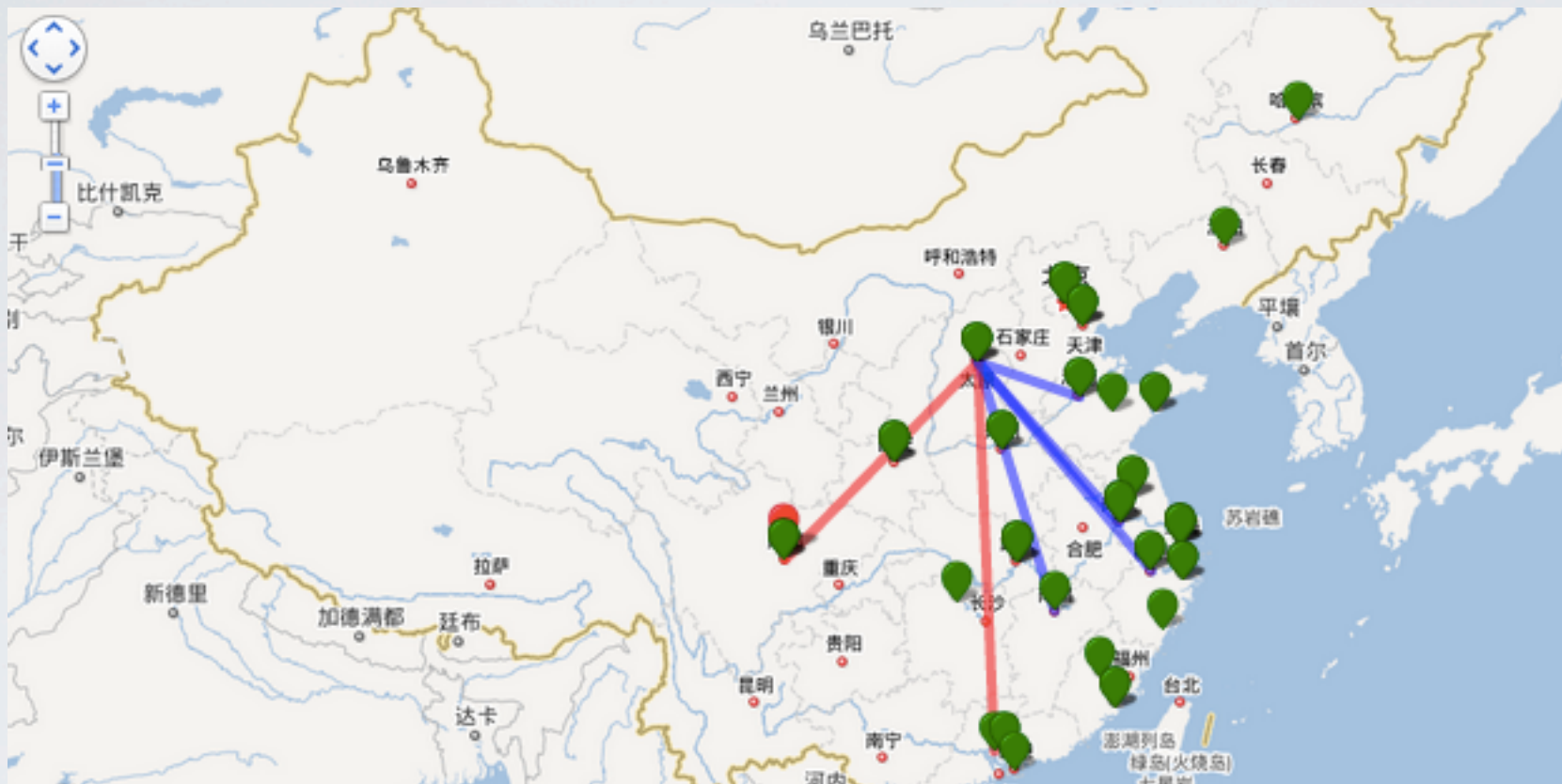
Console

- Django+Bootstrap
- 不仅仅是自动化

国家	省份	ISP	访问数
中国	广东	电信	1006746
中国	北京	联通	369116
中国	上海	电信	376765
中国	江苏	电信	352865
中国	浙江	电信	320236
中国	河北	联通	316408
中国	广西	电信	313626
中国	山东	联通	309178
中国	辽宁	联通	292036
中国	河南	联通	267144
中国	天津	联通	249886
中国	四川	电信	244466
中国	江西	电信	212511
中国	山西	联通	207994
中国	湖北	电信	198132
中国	广东	联通	186052

不能只是做简单的工具自动化，还需要做更多的集成工作。分析工作等

多路探测



SRE&DevOps

SRE

我们也在实践与探索之中，欢迎交流

SRE

·🌀· Site Reliability Engineering

SRE

·🔗· Site Reliability Engineering

SRE

· Site Reliability Engineering

· 职能：

- 具有对系统的全局把握能力
- 保障团队的工作时间与质量
- 随叫随到，解决任何问题
- 记录未自动化事宜，提出改进

DevOps

更多的说的是运维+开发之间的事

DevOps

·🔗· Development & Operations

更多的说的是运维+开发之间的事

DevOps

·🔗· Development & Operations

更多的说的是运维+开发之间的事

DevOps

·🌀· Development & Operations

·🌀· 职能：

- 自动化工具开发
- 解决SRE分配的问题
- 数据分析及监控工具的开发

SRE&DevOps关键点

- 🌀· 文化、思想观念的改变
- 🌀· 统一的标准化流程
- 🌀· 统一的工具
- 🌀· 统一的版本控制库
- 🌀· 清晰的依赖关系
- 🌀· 尽可能一自动化

Q&A

招牛人，招怪兽，招浪人！

Pythoner

DevOps

c (SinaStorage)

liuyu1 #staff.sina.com.cn

更多职位欢迎咨询～我还是人口贩子哦！