

Go FaaSter : Serverless 平台冷启动优化

周维跃

TGO 鲲鹏会

汇聚全球科技领导者的高端社群

🌐 全球12大城市

👤 850+ 高端科技领导者

使命
Mission

为社会输送更多优秀的
科技领导者

愿景
Vision

构建全球领先的有技术背景
优秀人才的学习成长平台



扫描二维码，了解更多内容

极客邦科技 会议推荐2019

5月

QCon 北京

全球软件开发大会

大会: 5月6-8日
培训: 5月9-10日

QCon 广州

全球软件开发大会

培训: 5月25-26日
大会: 5月27-28日

6月

GTLC
GLOBAL
TECH LEADERSHIP
CONFERENCE

上海

技术领导力峰会

时间: 6月14-15日

GMTC 北京

全球大前端技术大会

大会: 6月20-21日
培训: 6月22-23日

7月

ArchSummit 深圳

全球架构师峰会

大会: 7月12-13日
培训: 7月14-15日

10月

QCon 上海

全球软件开发大会

大会: 10月17-19日
培训: 10月20-21日

11月

GMTC 深圳

全球大前端技术大会

大会: 11月8-9日
培训: 11月10-11日

AiCon 北京

全球人工智能与机器学习大会

大会: 11月21-22日
培训: 11月23-24日

12月

ArchSummit 北京

全球架构师峰会

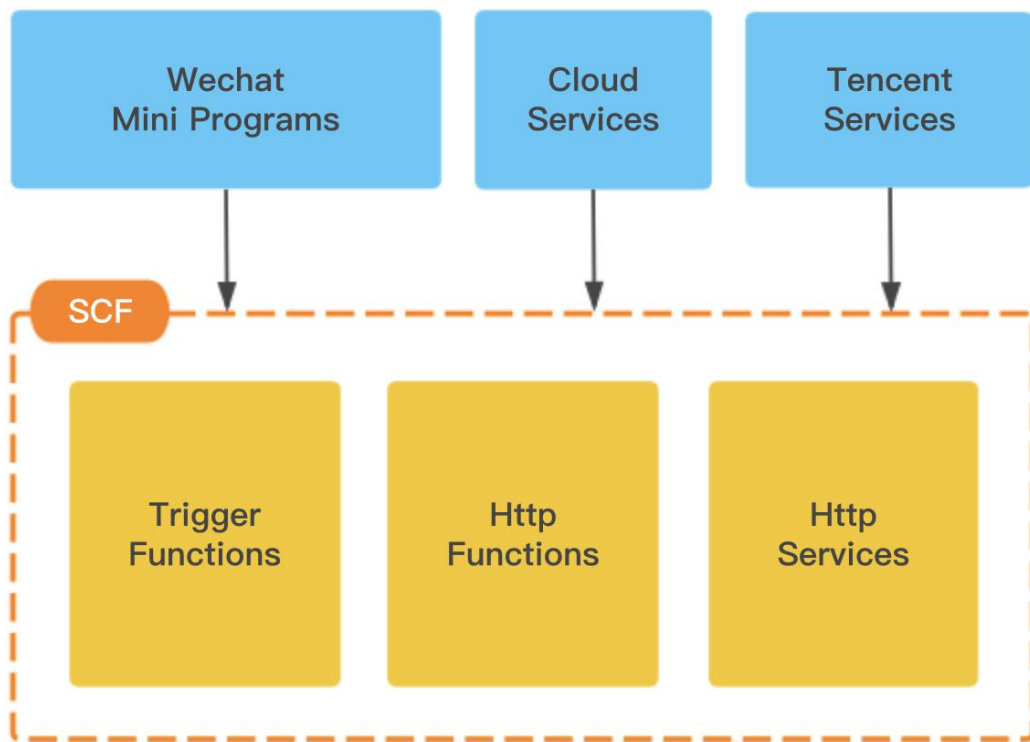
大会: 12月6-7日
培训: 12月8-9日

TABLE OF CONTENTS 大纲

- 腾讯云无服务器云函数介绍
- 什么是函数冷启动？
- 优化函数冷启动的一些方法

1.1 腾讯云无服务器云函数

Serverless Cloud Functions



1. 微信小程序开发

百万级别的小程序开发者

十亿小程序终端用户

2. 公有云的服务

Web服务，API，计划任务，异步触发场景

3. 腾讯自研业务

腾讯内部的社交业务上云，部分直接迁移到函数计算

1.2 遇到的问题和挑战

1. 安全隔离

虚拟机级别的隔离

2. 高并发扩容/缩容函数实例

每秒10,000个实例

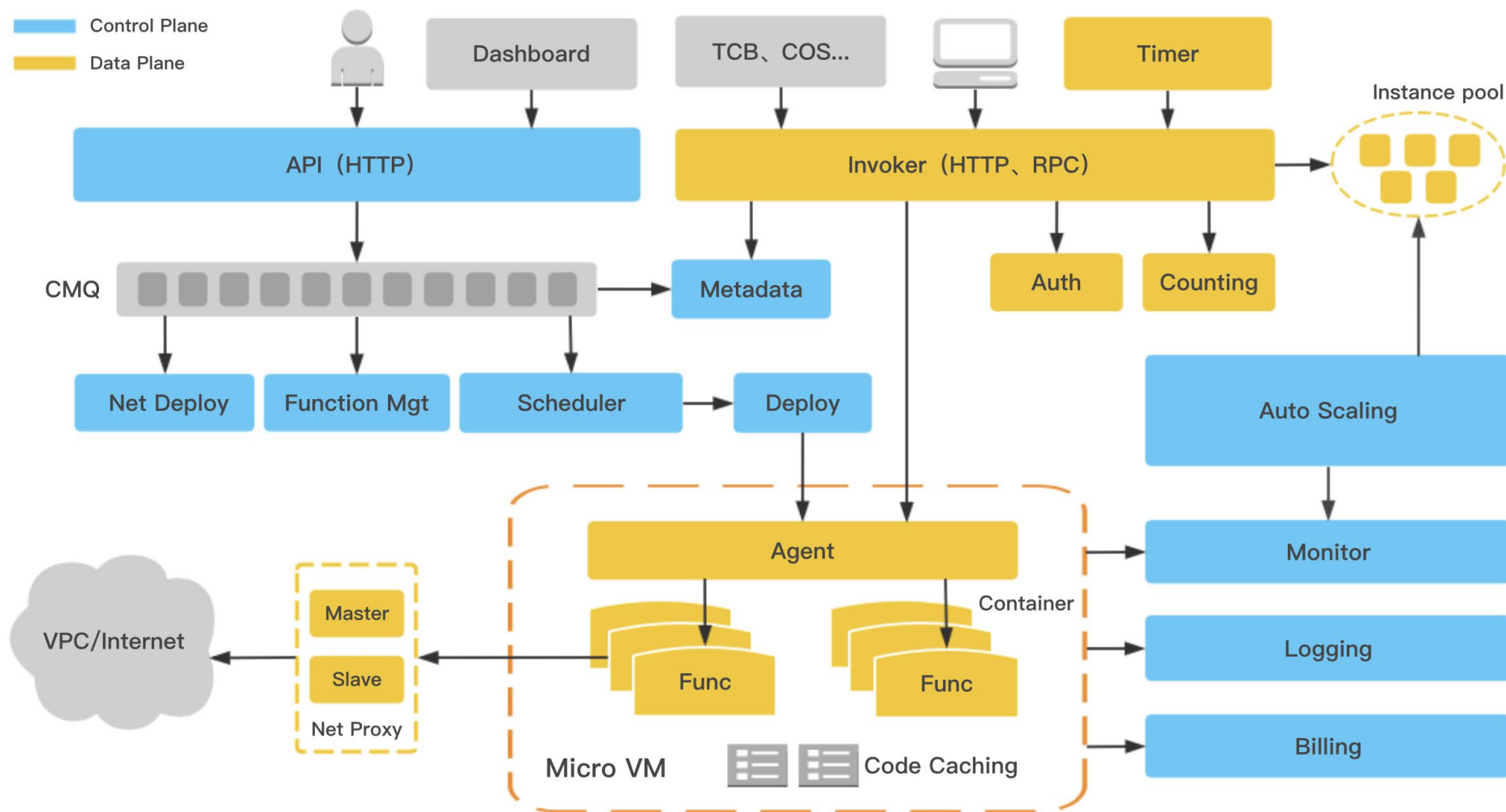
3. 大规模的实例集群管理

百万级别的函数实例

4. 极低的冷启动延时

低于100毫秒，对于活跃函数要做到0%冷启动

1.3 SCF整体架构

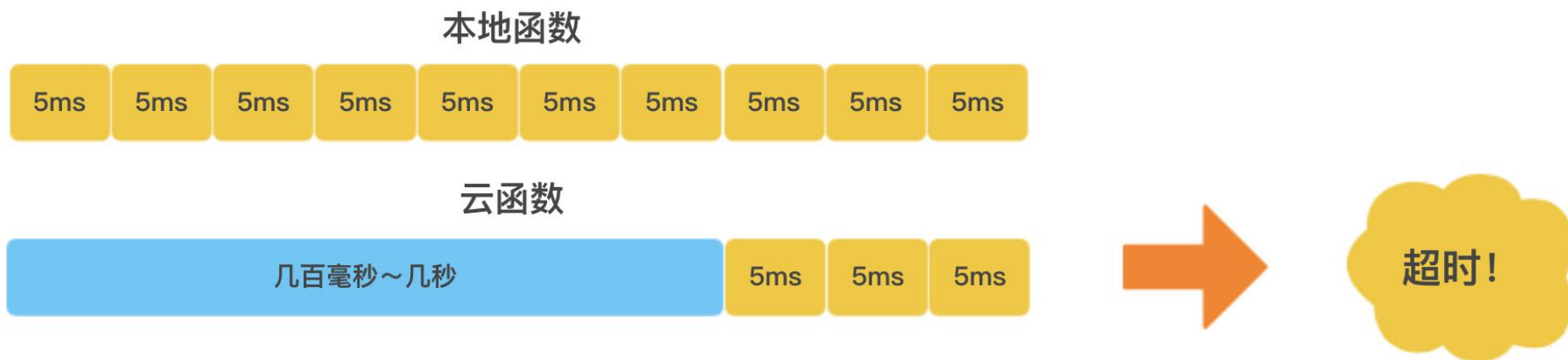


2.1 什么是函数冷启动

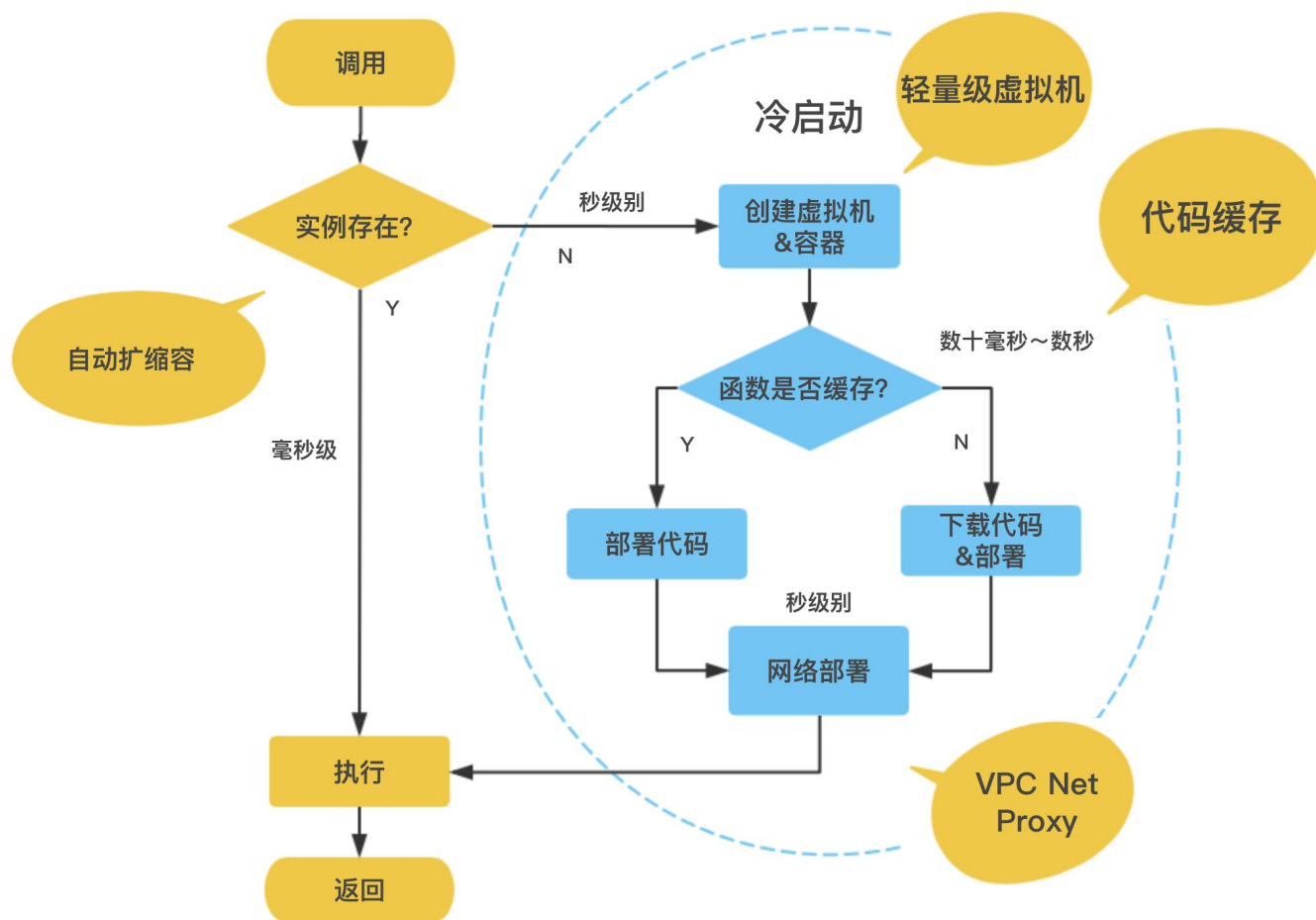
函数冷启动就是第一次部署函数运行实例的过程！

万事皆有冷启动？

为什么如此关键？



2.2 函数冷启动的过程



阶段一：创建虚拟机或容器

虚拟机创建耗时分钟级别

容器创建耗时秒级别

阶段二：函数代码包下载

耗时主要取决于代码包的大小

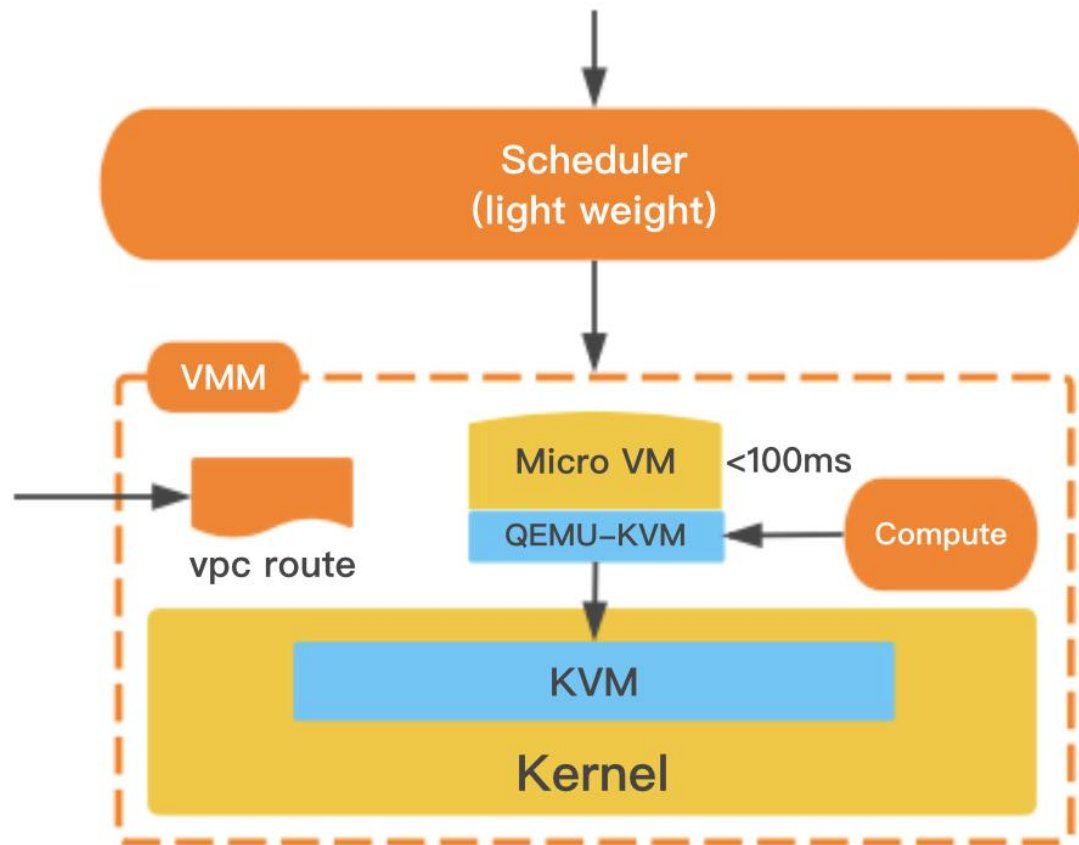
阶段三：打通VPC网络

部署弹性网卡通常需要秒级别

3 优化冷启动的一些方法

1. 轻量级虚拟机系统
2. 代码缓存
3. VPC网络代理
4. 自动扩缩容
5. 用户可以做些什么？

3.1 轻量级虚拟化系统



1. 调度方面

轻量化，降低调度复杂度

2. 网络方面

提前预下发

3. 虚拟化方面

轻量级虚拟机，极速启动

3.1.1 调度流程

原有的调度模块需要考虑的维度很多

数十种虚拟机的配置，CPU/内存/存储/网络等

亲和性和反亲和性

部署组的需求：跨宿主机，交换机，机架

资源利用率，碎片填充

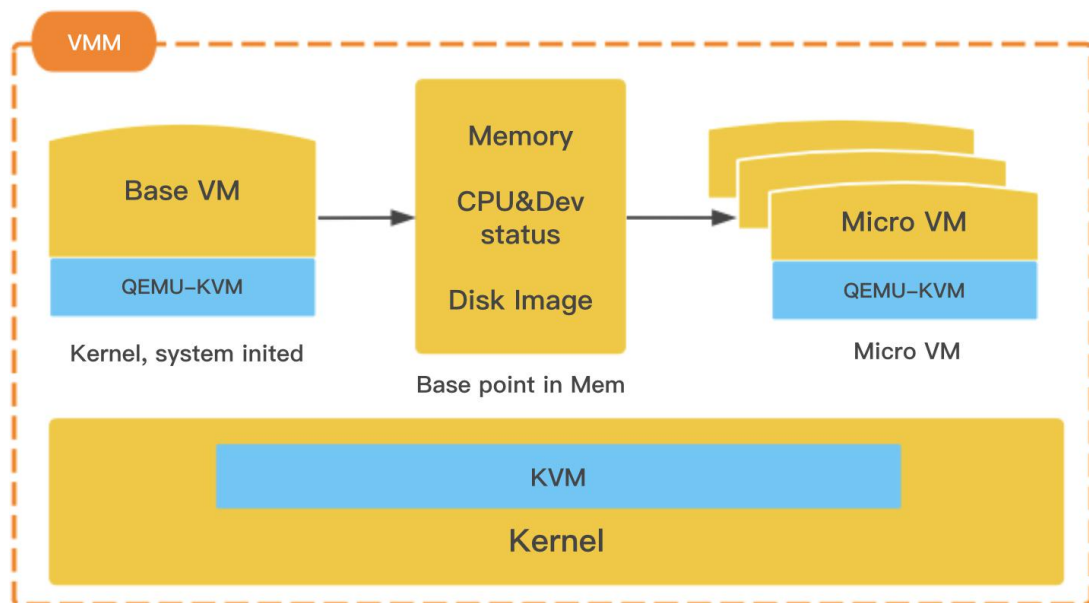
...

轻量化调度模块

少数的虚拟机配置

宿主机的可用资源离线计算

3.1.2 轻量级虚拟化



1. 提前创建虚拟机模版文件

运行一个基础虚拟机，通过QA检测启动状态

保存虚拟机的内存数据、CPU状态、设备状态等信息到
共享内存

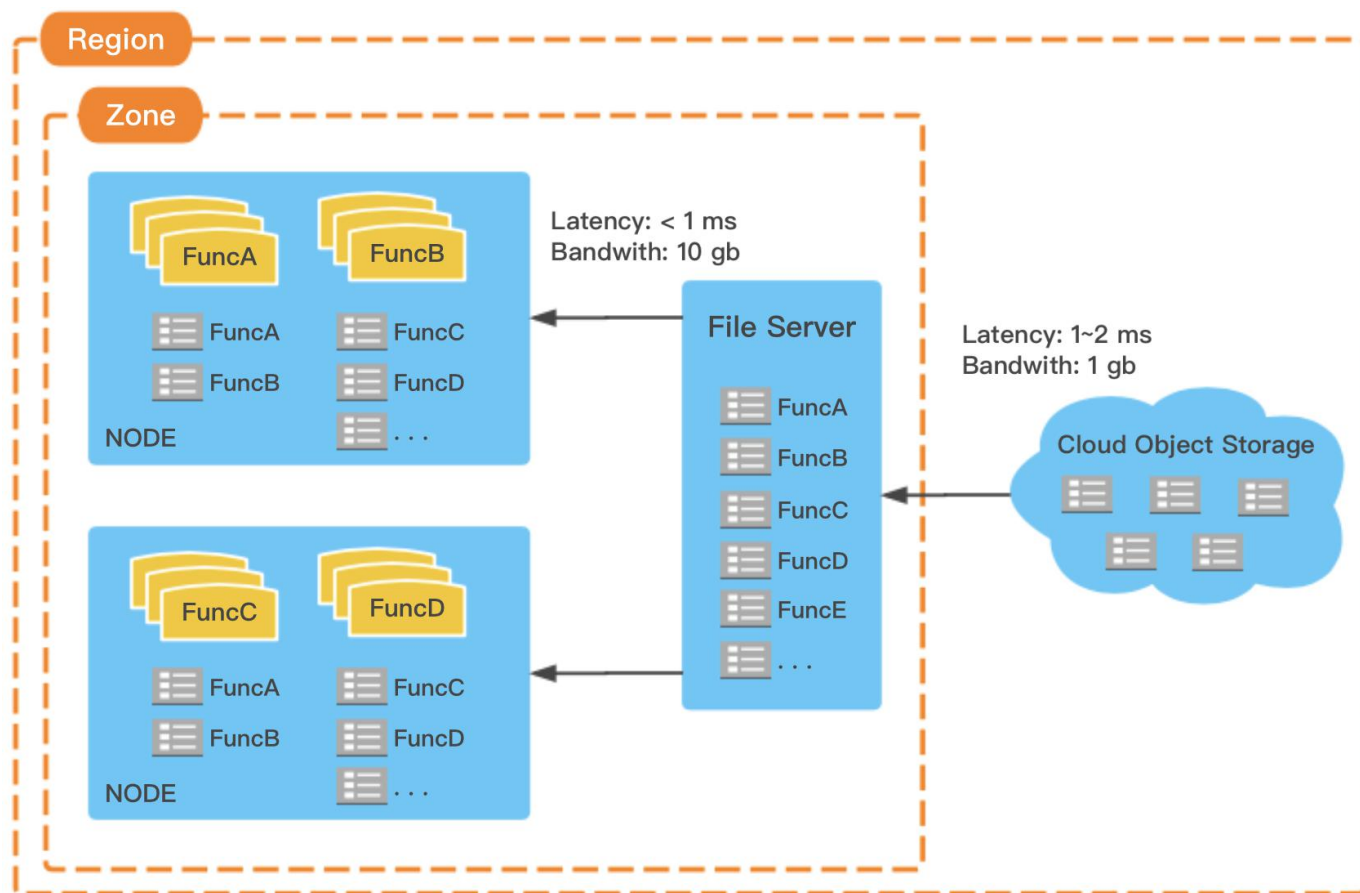
2. 基于虚拟机模版文件启动轻量级虚拟机

-incoming exec:cat /dev/shm/vm_template

系统盘: qcow2 snapshot, 3层镜像

修改Mac, IP, hostname等唯一信息

3.2 代码缓存



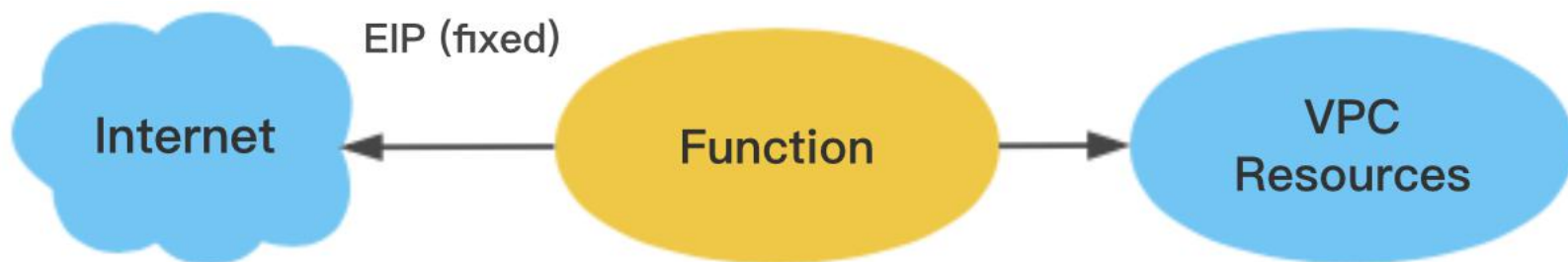
L1计算节点本地缓存：

同一个开发者的所有函数代码全部缓存
缓存命中率100%

L2可用区本地缓存：

代码仓库所有代码缓存到可用区本地
缓存命中率100%

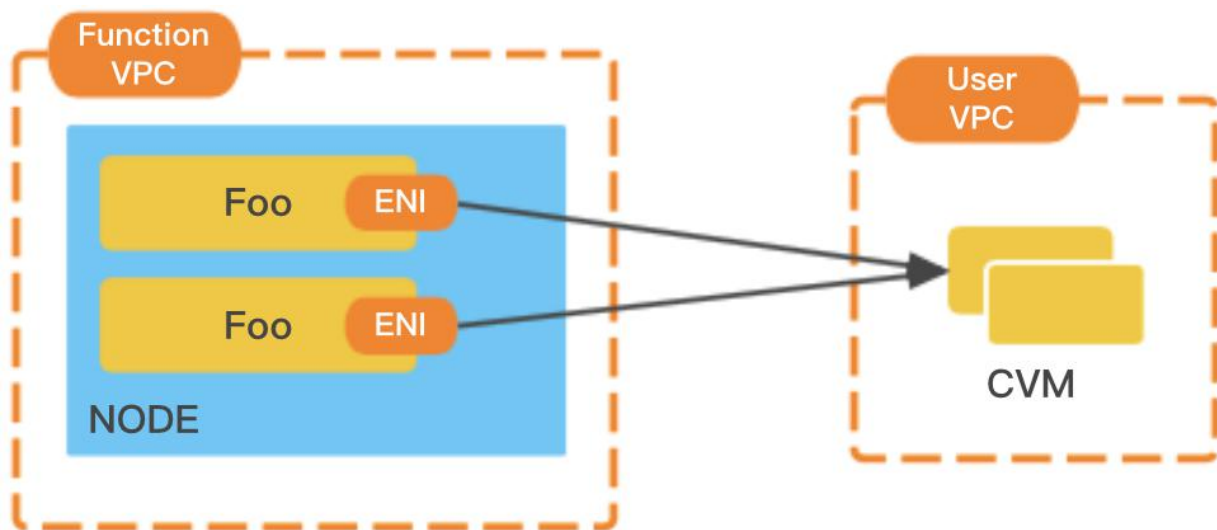
3.3 网络访问模型



函数需要访问用户VPC中的资源，比如CVM, CDB...

函数需要访问公网、自建数据中心的资源，有些场景下需要固定公网IP地址

3.3.1 传统网络-访问VPC



方案（一）

ENI（弹性网卡）绑定在Pod上，访问VPC资源

ENI的绑定过程发生在函数调用冷启动的阶段

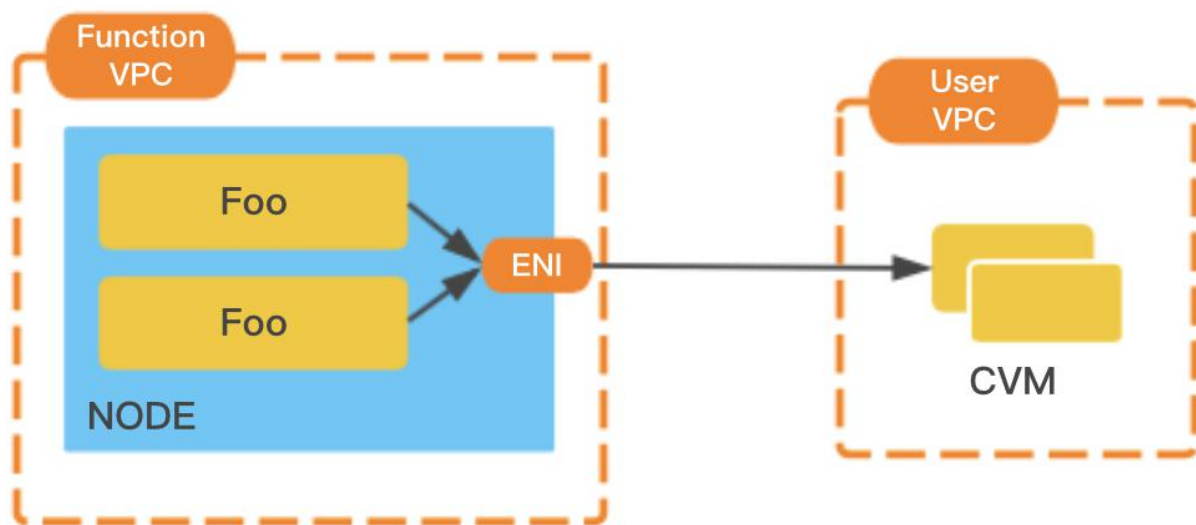
ENI在同一个Pod内多容器共享，在同一个Pod内再次创建函数实例不需要重新绑定ENI

问题：

导致函数冷启动过程中比较高的延时

对用户VPC内的弹性网卡消耗巨大，如果弹性网卡消耗完，会导致函数并发提升受限

3.3.1 传统网络-访问VPC



方案（二）

ENI（弹性网卡）绑定在Node上，访问VPC资源

ENI的绑定过程发生在函数调用冷启动的阶段

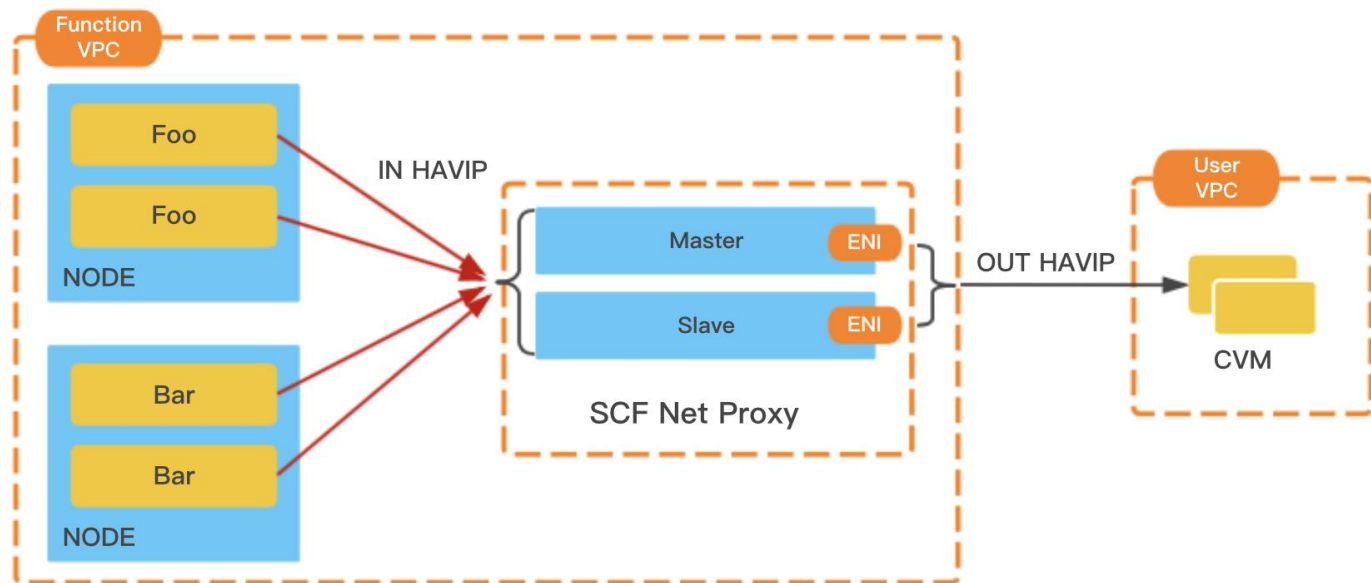
ENI在同一个节点内多容器共享，在同一个节点内再次创建函数实例不需要重新绑定ENI

问题：

导致函数冷启动过程中比较高的延时

对用户VPC内的ENI消耗巨大，如果弹性网卡消耗完，会导致函数并发提升受限

3.3.2 SCF网络-访问VPC



优势：

冷启动时间从秒级下降到毫秒级

仅消耗一对ENI

主备节点支持高可用，秒级故障切换

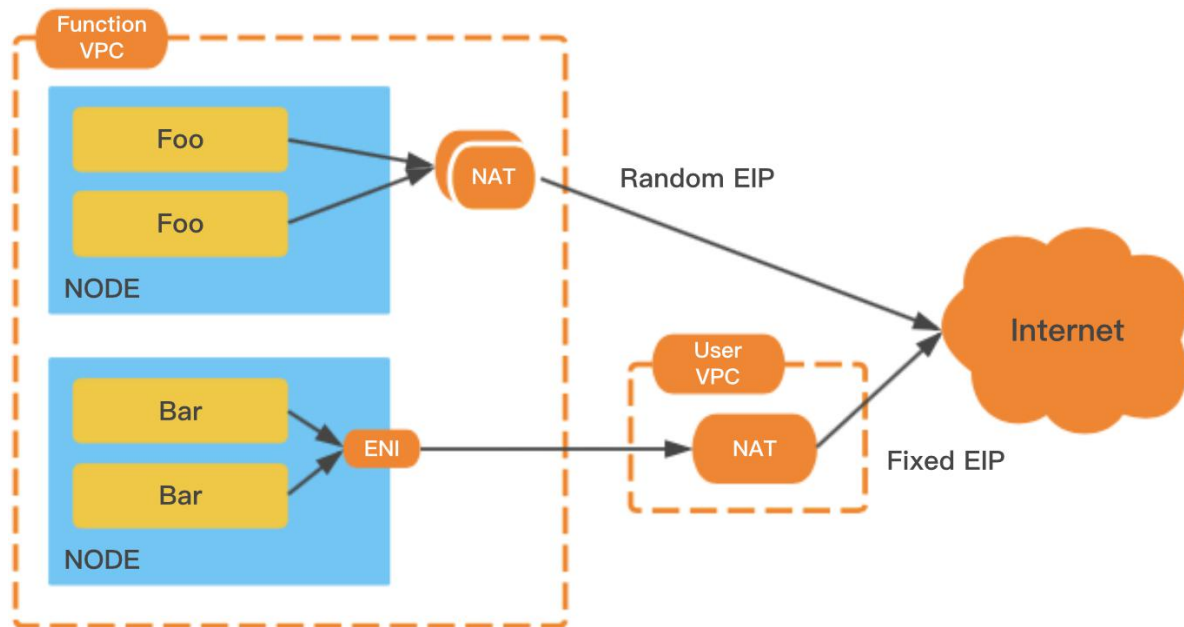
转发节点的带宽可自动扩缩容

函数访问VPC资源：

通过SCF Net Proxy转发流量

在函数的创建过程中将ENI绑定到Proxy的转发节点上，从函数调用的数据流转移到控制流的操作

3.3.3 传统网络-访问公网



需要固定公网IP:

- 1、在用户VPC中部署一个NAT实例
- 2、函数访问VPC的NAT实例出公网

问题：

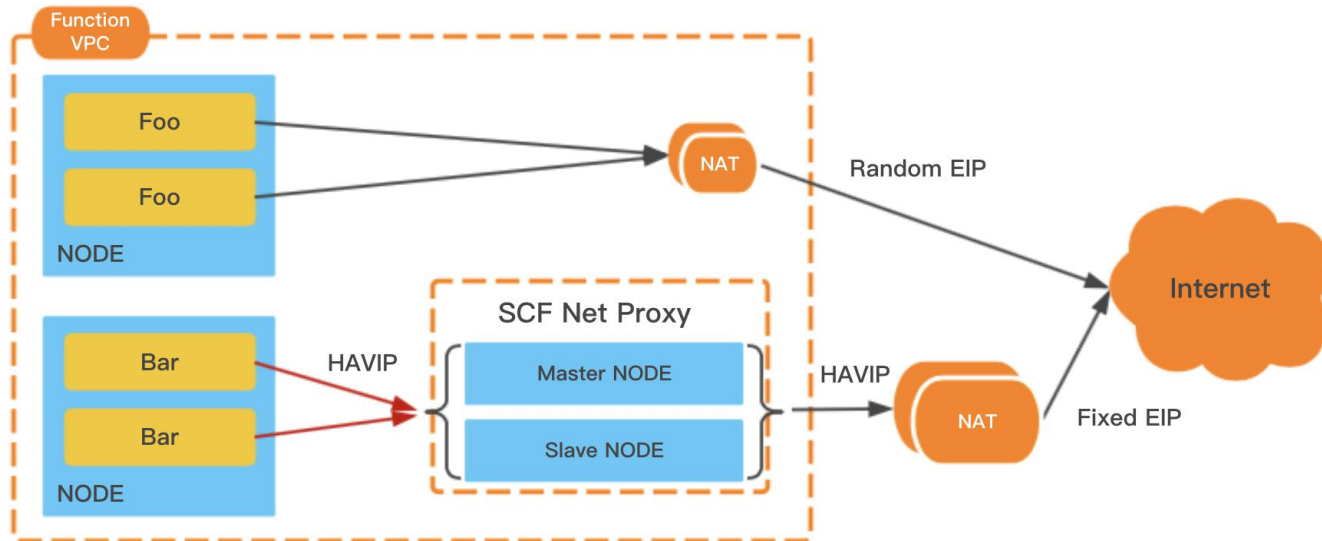
对中小企业、个人开发者来说门槛过高，配置过于复杂

需要为NAT支付固定的月度费用

实例部署需要绑定ENI，冷启动耗时增加

实例部署需要绑定ENI，消耗大量ENI

3.3.4 SCF网络-访问公网



优势:

冷启动时间从秒级下降到毫秒级

仅消耗一对ENI

主备节点支持高可用，秒级故障切换

转发节点的带宽可自动扩缩容

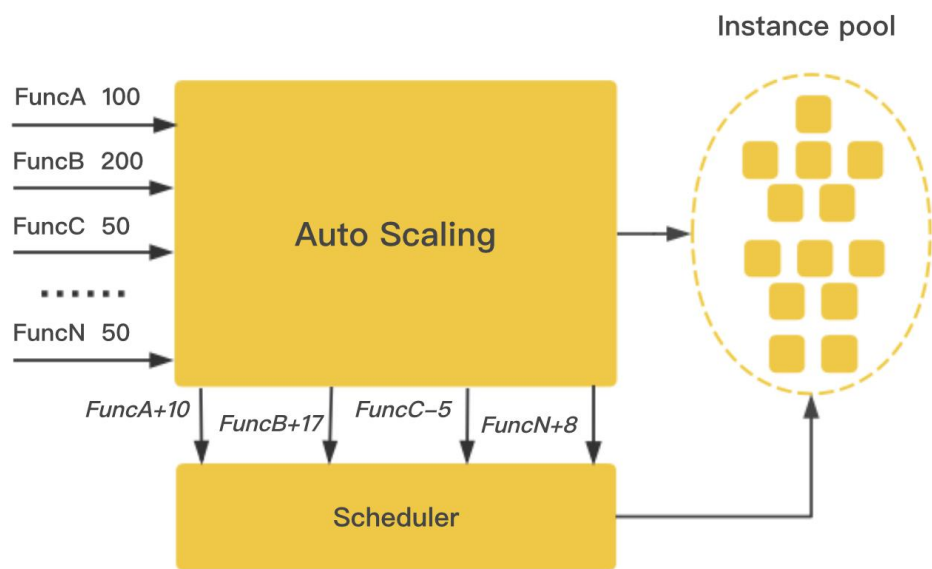
用户不需要理解、配置NAT，无需支付固定费用

需要固定公网IP:

通过SCF Net Proxy转发流量

在函数的创建过程中将ENI绑定到Proxy的转发节点上，从函数调用的数据流转移到控制流的操作

3.4.1 自动扩缩容

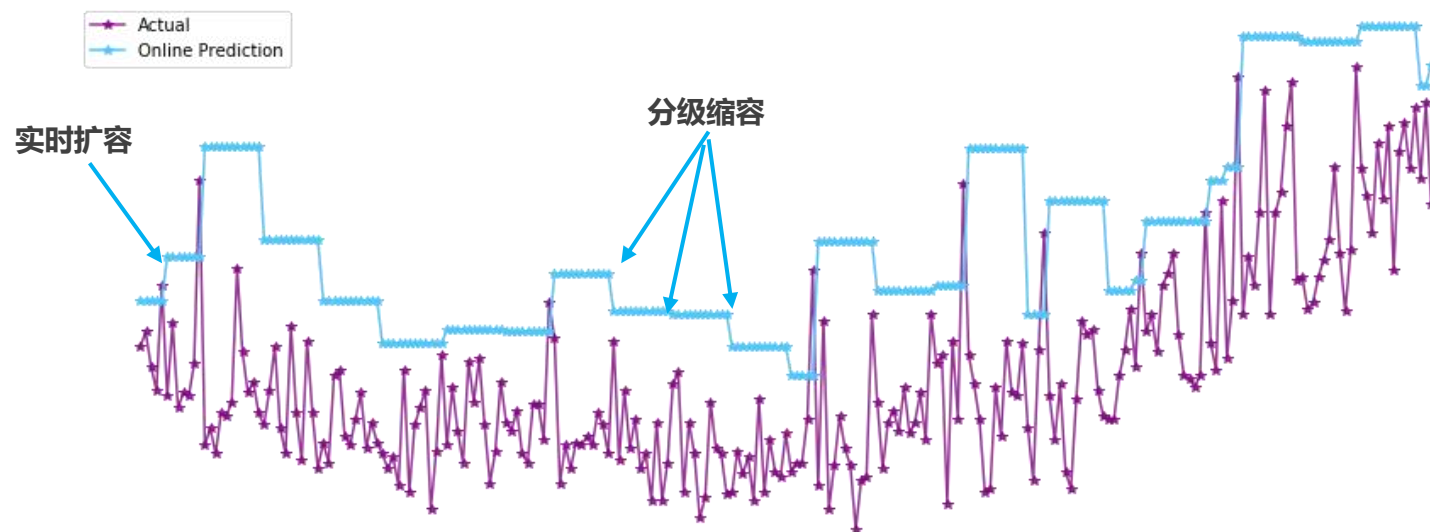


实时计算扩缩容

秒级实时扩容，预留buffer

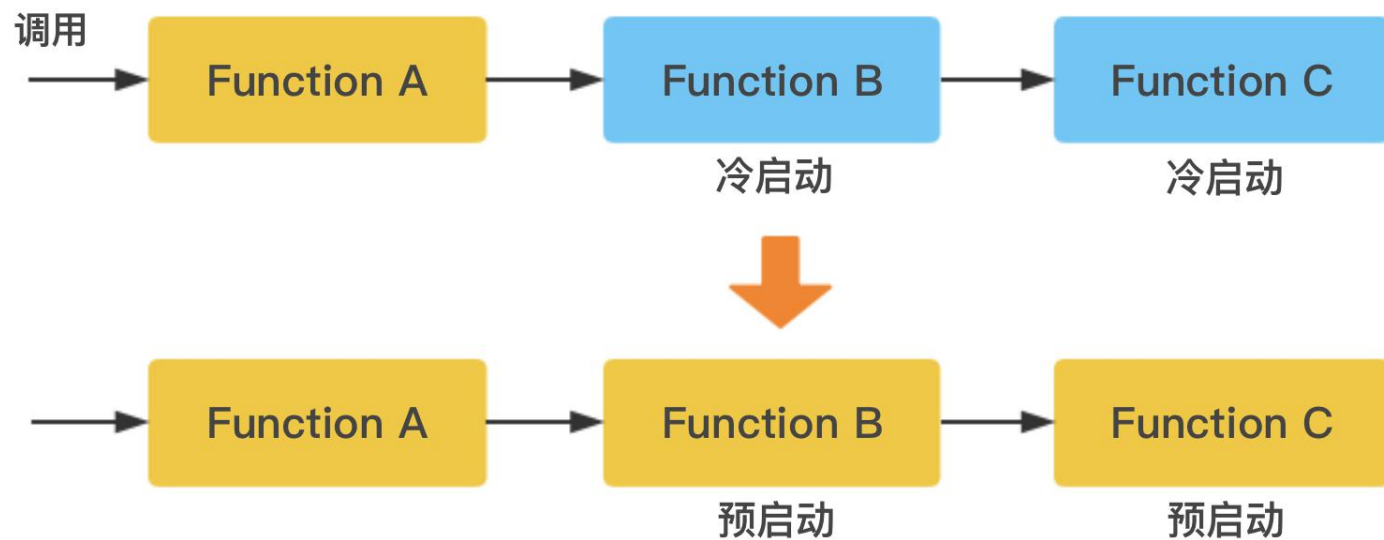
冷却时间，分级缩容

Concurrent Request Time Series Prediction

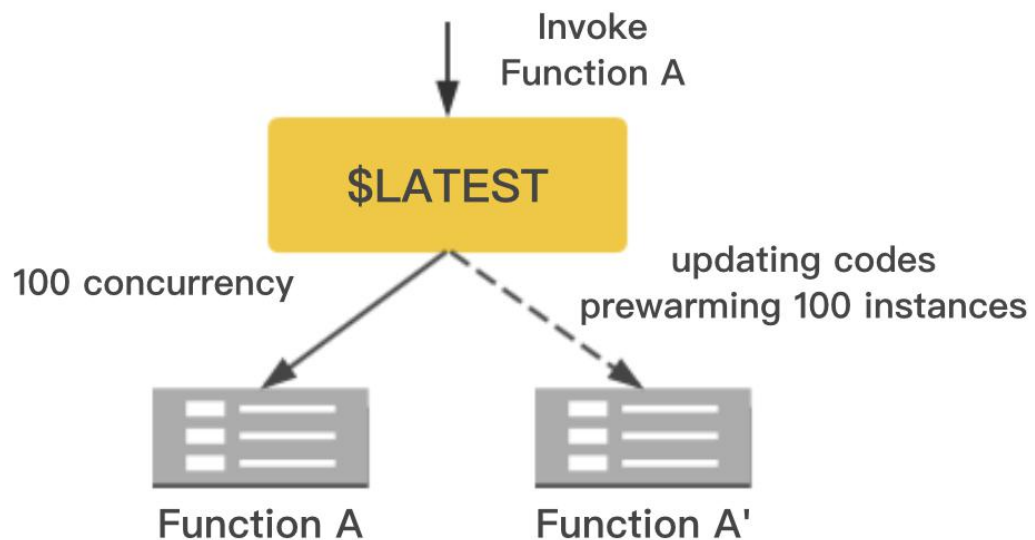


3.4.2 可预测的扩容（一）

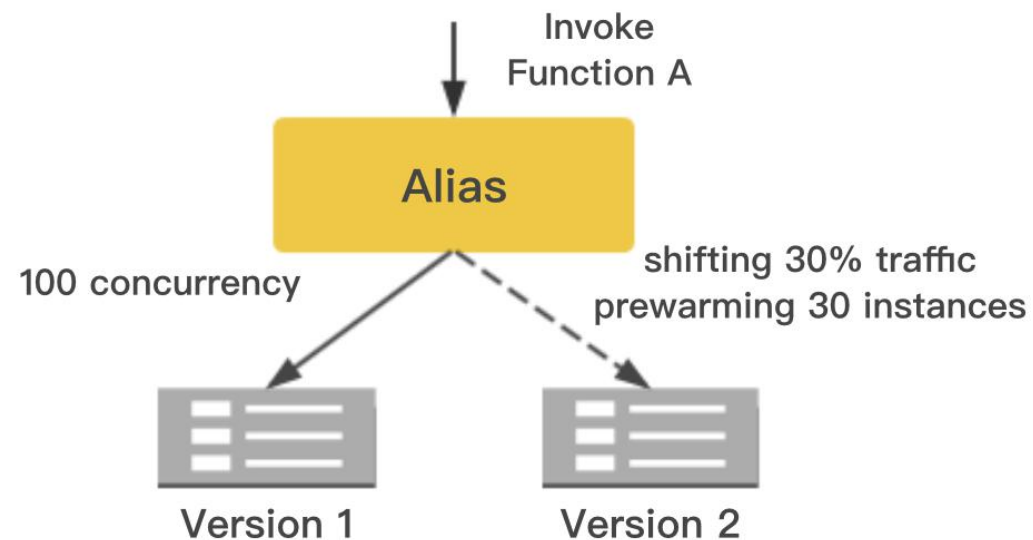
函数A调用B，B调用C



3.4.2 可预测的扩容（二）



1. 更新\$LATEST版本



2. 灰度切换流量到另外一个版本

3.5 用户可以做些什么？

- 1、保持代码精简，减小体积
- 2、全局变量，资源复用
- 3、定时触发保活

总结

- 1、从函数架构层面优化，支持更大规模的集群管理和大并发实例部署
- 2、优化轻量级虚拟机系统，降低虚拟机创建耗时
- 3、优化VPC网络转发模块，降低弹性网卡部署的耗时和资源消耗
- 4、实时自动扩缩容，避免冷启动
- 5、用户裁剪代码、资源复用、保活策略

谢谢！

Q&A



云原生公众号



SCF用户交流群

极客时间全部课程任学 喊老板来买单！

- ✔ 精选 13+ 热门职位的学习路径，包括架构、运维、前端工程师等
- ✔ 根据不同技术岗位能力模型匹配合适的课程
- ✔ 一键设置购买条件，成员按需选课，自主制定学习计划
- ✔ 享充值满赠优惠，帮老板省钱，团队免费学习



立即申请



InfoQ官网 全新改版上线

促进软件开发领域知识与创新的传播



关注InfoQ网站
第一时间浏览原创IT新闻资讯



免费下载迷你书
阅读一线开发者的技术干货

About The **SPEAKER**

周维跃 (Scott Zhou)

目前担任腾讯云Serverless平台研发负责人

2008年加入腾讯，曾经负责过消息队列中间件、域名&DNSPod，虚拟机等产品。其中，通过虚拟机大规模热迁移调度技术，实现了毛利率50%以上的增长

近期聚焦在Serverless架构中函数冷启动优化、自动扩缩容、函数与云资源互联互通等核心能力建设

平时喜欢长途自驾穿越，旅行摄影

