### 弹性伸缩算法初步调研

目前，弹性伸缩算法的一个总体思路是采用MAPE循环。其中MAPE的含义如下：M表示监视(Monitoring)、A表示分析(Analyzing)、P表示计划(Planning)、E表示执行(Executing)。首先是监视系统收集有关系统和应用程序的状态信息，例如CPU和内存的使用情况。弹性伸缩器涵盖了分析和计划阶段：使用检索到的信息来估计未来的资源利用和需求，然后计划适当的资源修改操作，例如删除微服务实例或添加一些额外的内存。然后执行其操作。

同时，弹性伸缩的算法有五类：

* 基于阈值的弹性伸缩算法
* 基于强化学习的弹性伸缩算法
* 基于排队理论的弹性伸缩算法
* 基于控制理论的弹性伸缩算法
* 基于时间序列分析的弹性伸缩算法

基于阈值的弹性伸缩算法在云服务提供商中得到了广泛的应用。其核心思路是通过设定最大值和最小值，在某一个特定的时刻对资源进行扫描，然后根据条件语句判断分配资源。然后，经过设置的间隔，重复上述操作；基于强化学习的弹性伸缩算法，本质上是让弹性伸缩器从弹性应用之间的直接交互中来进行学习。弹性伸缩将根据经验学习最佳伸缩操作，具体取决于输入负载，性能或其他变量集所给出的当前状态； 基于排队理论的弹性伸缩算法是将运用于运筹学的排队理论迁移到了弹性伸缩。用于弹性应用程序MAPE循环的分析阶段，以估算性能指标，例如队列长度或请求的平均等待时间；基于控制理论的弹性伸缩是将应用于控制学的闭环系统迁移到了弹性伸缩， 通过反馈机制改进弹性伸缩策略；基于时间序列分析的弹性伸缩算法是对资源未来的使用量进行预测。在金融、工程、经济学和生物信息学，时间序列通常表示度量随时间的变化。以便弹性伸缩器在输入工作负载中找到重复的模式或尝试预测将来的值。