

Serverless Computing: Design, Implementation, and Performance

В данной статье авторы из университета Нотр Дама создают новый движок для выполнения *Serverless*-функций. Эта область является весьма новой как на время написания статьи, так и сейчас. Поэтому авторы описывают решения созданные крупными компаниями поставщиками облачных услуг и прodelывают весьма небольшой анализ предшествующих работ.

Основная новизна их идей заключается в оптимизации поиска свободного места для выполнения функции. Они предлагают идею заключающуюся в создании двух очередей – горячую и холодную. В начале процесс выглядит как обычно – выделить свободное место для контейнера в котором будет выполняться наша функция; после выполнения поместить очередь в горячую. Затем после того как контейнер внутри очереди просрочит таймаут поместить этот контейнер в холодную очередь. Для того, что бы любая функция могла бы взять контейнер из холодной очереди под нее выделяется(или перевыделяется) столько памяти, сколько требуется для самой потребляющей функции. Когда приходит новый запрос память обрезаается под конкретную функцию.

Затем появляется потребность в освобождении памяти от долго неиспользуемых контейнеров. Контейнер может быть удален в двух случаях – неиспользования в течении какого-то времени(15 минут в статье) или при удалении функции.

В части тестов авторы проверяют свою реализацию на нескольких различных бэнчмарках, таких как: количество параллельных вызовов и время между выполнением двух одинаковых функций.

В ходе чтения статьи возникает очевидный вопрос – как справиться с тем, что возможна ситуация когда много контейнеров находятся в состоянии *IDLE*, и при этом не таймаутятся из-за структуры очереди(*FIFO*) и периодических запросов "обновляющих" очередь. Авторы, в части *Future Work*, предлагают использование стека(*LIFO*) вместо очереди. И дальше они предлагают такие улучшения как: асинхронные выполнения, улучшения использования воркера и безопасность.

Личное мнение: статья не лишена элегантности и написана понятным языком, однако очевидно(и сами авторы это признают), что во многих аспектах их метод недоработан и является лишь прототипом(что и отражается в названии их метода).