



ЦИФРОВОЙ СУВЕРЕНИТЕТ

Разработка программных решений

Разработка универсального парсера логов

Sitronics Group, до 180 баллов

Эксперт — Александр Дворянский

Анализ логов — весьма часто встречающаяся задача при разработке коммуникационных продуктов.

Парсер логов позволяет автоматизировать рутинные операции этой задачи и в несколько раз увеличить скорость анализа, а также выявить сложные закономерности и связи в логируемой информации.

Необходимо разработать парсер, который запускается из командной строки и поддерживает следующие режимы:

1. Выбирает из входного файла лога те строки, в которых есть вхождение одного или нескольких ключевых слов и сохраняет выбранные строки лога в выходной файл.
2. Выбирает из входного файла лога те строки, в которых есть вхождение одного или нескольких ключевых слов и сохраняет выбранные строки лога в несколько выходных файлов, объединяя их по значению уникального идентификатора. В базовом варианте решения, уникальный идентификатор связан с ключевым словом через знак минуса. В расширенном варианте решения положение уникального идентификатора в строке задается с помощью шаблона.
3. Выбирает из входного файла лога те строки, в которых есть вхождение одного или нескольких ключевых слов и сохраняет выбранные строки лога в выходной файл. При сохранении к строкам применяется шаблон, с помощью которого можно вырезать из строки необходимую для анализа часть.

Шаблоны и ключевые слова могут быть заданы в аргументах командной строки или в виде входного файла. Пример файла лога прилагается.

Примеры задач на выборку:

- Выбрать из файла строки содержащие ключи XIMSS или CALLLEG.
- Выбрать из файла строки содержащие ключи XIMSS и сгруппировать по идентификаторам, связанным с XIMSS, например, XIMSS-000003 – в один выходной файл, XIMSS-000005 - в другой выходной файл.