

Система распределённого оптического распознавания символов **Nebulous**

Спецификация

Техническое и функциональное описание системы.

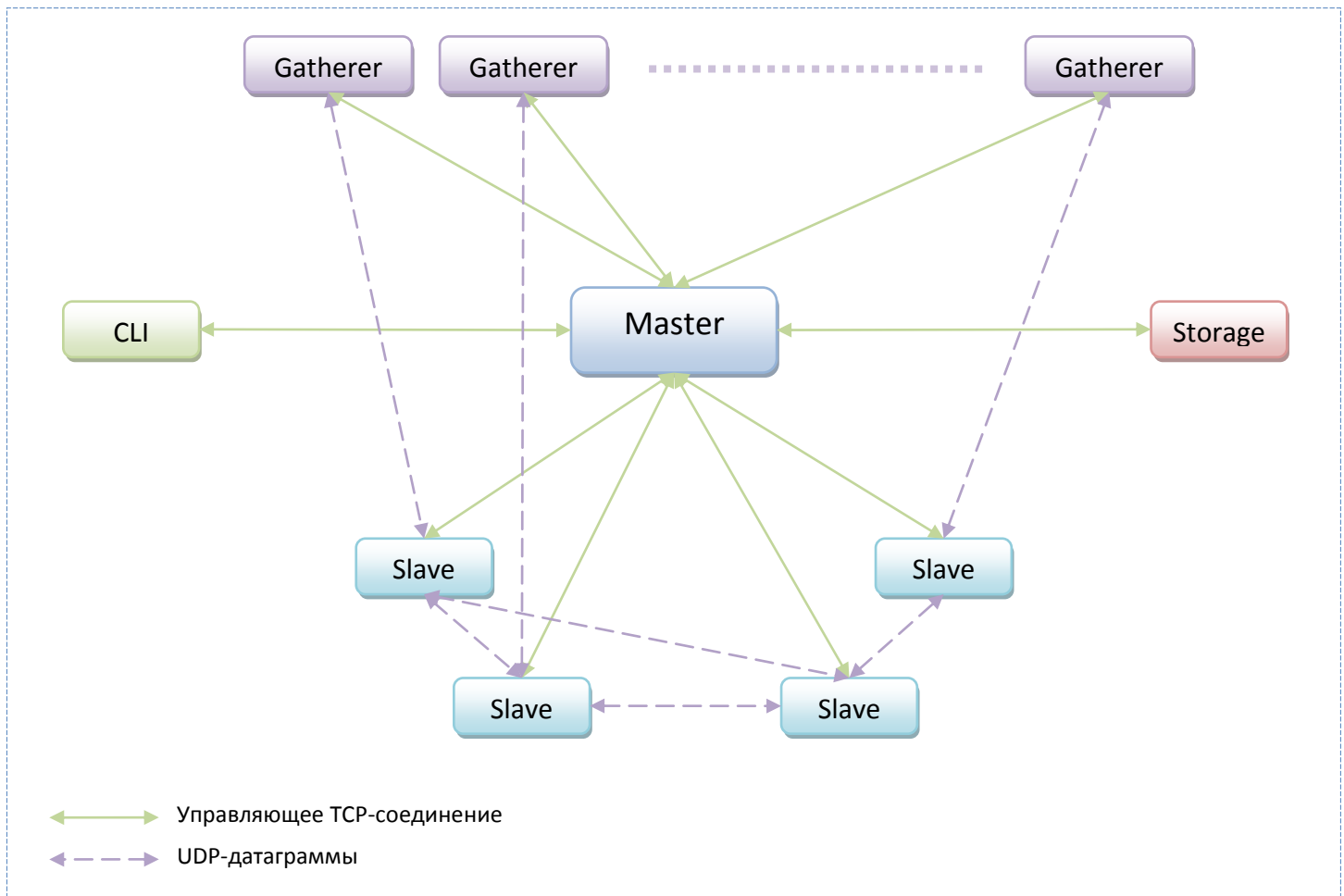
Игорь Калашников, Павел Пересторонин
20.10.2011

1 Введение

Целью данного проекта является проектирование и реализации оптического распознавания символов на большом потоке растровых изображений. Построенная система должна принимать на вход несколько потоков растровых изображений, выделять и распознавать символы на них и сохранять результаты распознавания вместе с остальными метаданными в централизованном хранилище.

2 Обзор архитектуры

Компоненты системы взаимодействуют между собой следующим образом:



Все данные между всеми узлами передаются в формате [Protocol Buffers](#) с помощью собственной библиотеки [Cloudy Message Passing Library](#), которая, кроме того, отвечает за поддержание работы распределённой системы в целом, абстрагируясь от бизнес-логики конкретного приложения.

2.1 Master

Задачи Master'a:

- Отслеживать появление и исчезновение узлов в сети. Вести списки узлов и групп узлов.
- Поддерживать топологию сети.
- Выполнять аутентификацию узлов.
- Обеспечивать возможность peer-to-peer соединений между узлами, находящимися за NAT и Firewall'ами.

2.2 Slave

Вычислительный узел.

2.3 Gatherer

Задача Gatherer'a – передача изображений из внешней среды в систему.

2.4 Storage

Хранилище данных о результатах распознавания.

2.5 CLI (Command-Line Interface)

Консоль для управления настройками системы и контроля над её работой.