|  |
| --- |
| **Система распределённого оптического распознавания символов Nebulous** |
| Спецификация |
|  |
| Техническое и функциональное описание системы. |
|  |
| **Игорь Калашников, Павел Пересторонин** |
| **20.10.2011** |
|  |

# 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.

# Обзор архитектуры

Компоненты системы взаимодействуют между собой следующим образом:

Master

Slave

Storage

CLI

Slave

Slave

Gatherer

Slave

Gatherer

Gatherer

Управляющее TCP-соединение

UDP-датаграммы

Все данные между всеми узлами передаются в формате [Protocol Buffers](http://code.google.com/p/protobuf/) с помощью собственной библиотеки [Cloudy Message Passing Library](https://github.com/eigenein/cloudy), которая, кроме того, отвечает за поддерживание работы распределённой системы в целом, абстрагируясь от бизнес-логики конкретного приложения.

## Master

Задачи Master'а:

* 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.

## Slave

Вычислительный узел.

## Gatherer

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

## Storage

Хранилище данных о результатах распознавания.

## CLI (Command-Line Interface)

Консоль для управления настройками системы и контроля над её работой.