



Использование Elliptics в Яндексе

Антон Картунов

Я.Субботник в Санкт-Петербурге, 1 декабря 2012

Что такое Elliptics

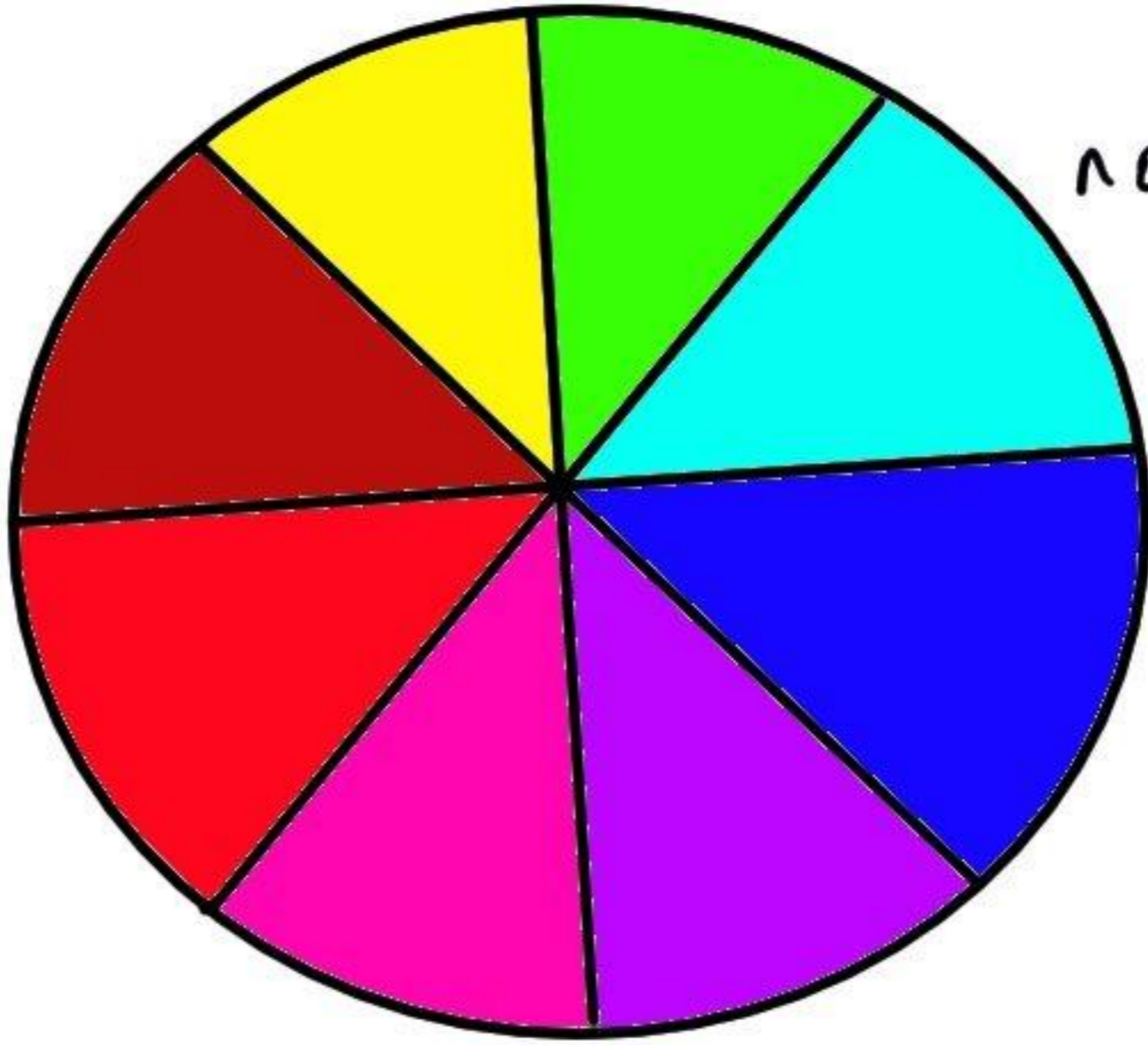
Elliptics network is a fault tolerant key/value storage. With default key generation policy it implements distributed hash table object storage. The network does not use dedicated servers to maintain the metadata information, it supports redundant objects storage. Small to medium sized write benchmarks can be found on eblob page and in the appropriate blog section.

<http://www.ioremap.net/projects/elliptics>

<https://github.com/reverbrain/elliptics/>

node 0

node 1

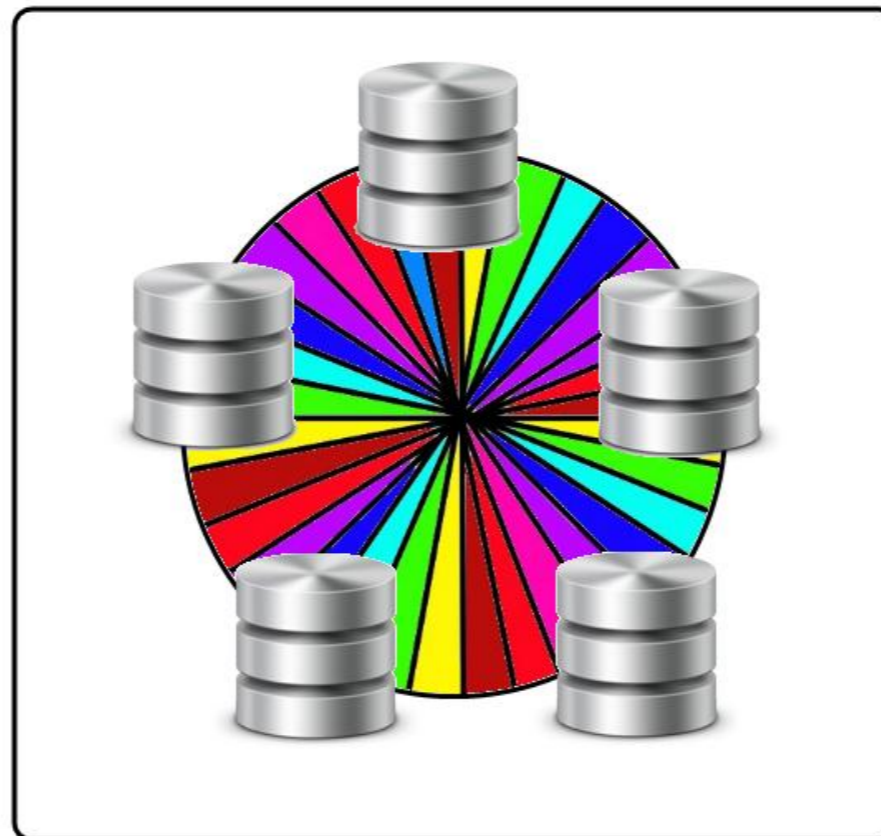
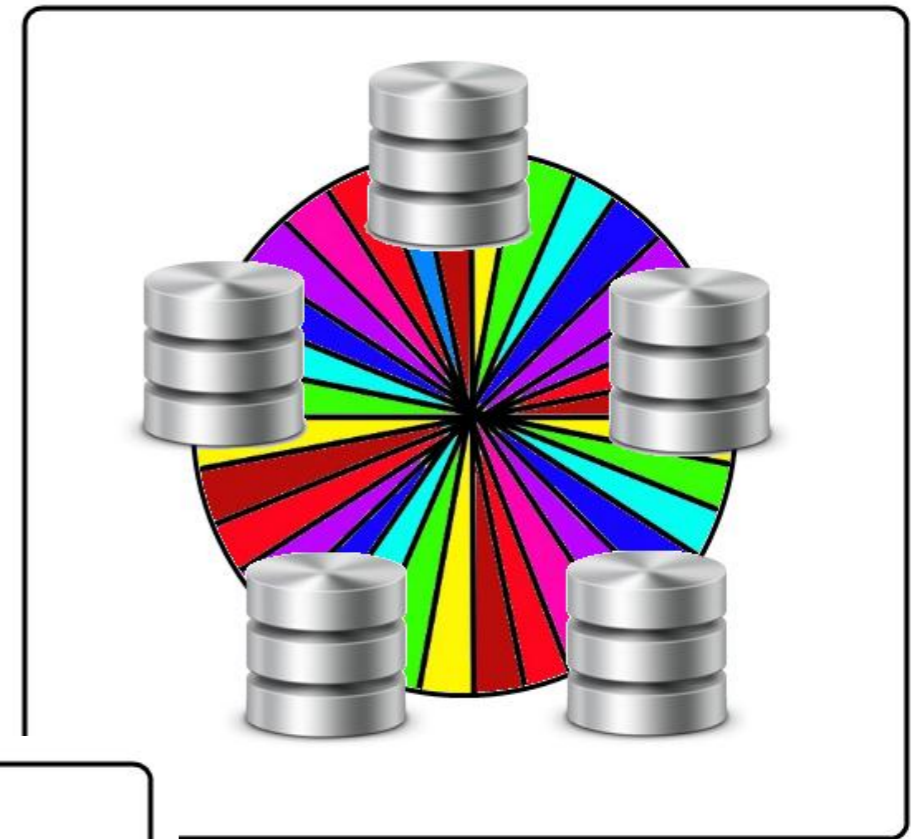
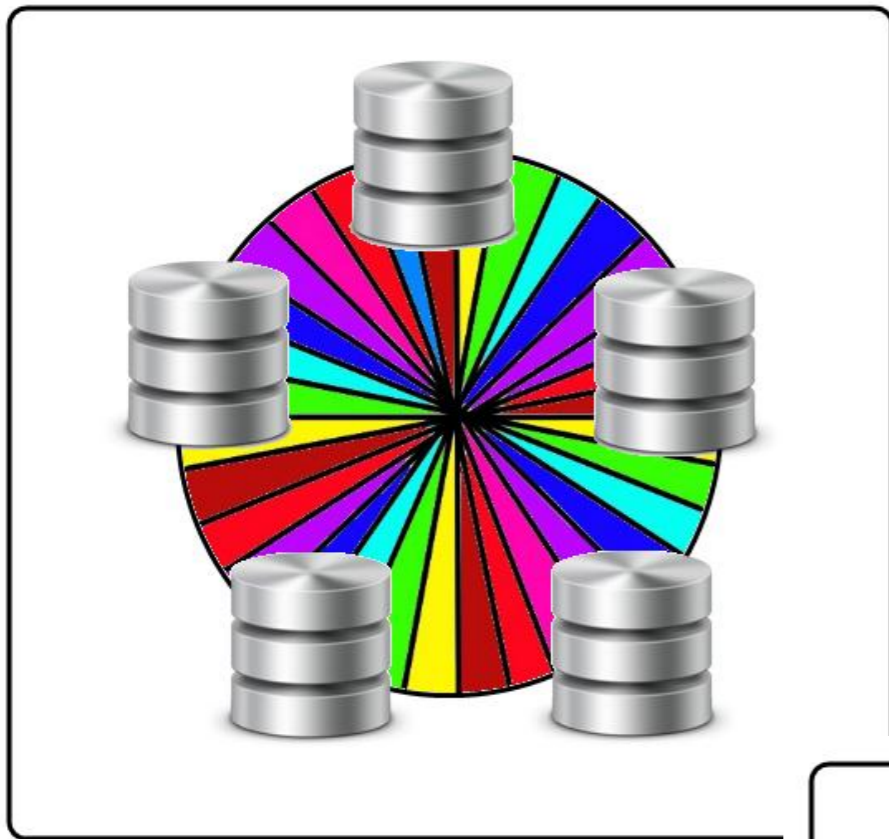


node N

DNТ установка

- Реализует DHT
- Используется в:
 - Яндекс.Музыка
 - Яндекс.Карты
 - Яндекс.Маркет
 - ...
- Доступ через HTTP Proxy

Полные реплики в разных ДЦ



Уже работает

- Elliptics как хранилище в CDN
- PONDMEFS – POSIX файловая система как фронтэнд к хранилищу
- Grape – realtime вычислительная платформа
- Бэкенды: eblob, leveldb, smack

Достоинства

- Линейная масштабируемость
- Отсутствие мастер-нод
- Простота маршрутизации: поиск ноды за $O(1)$

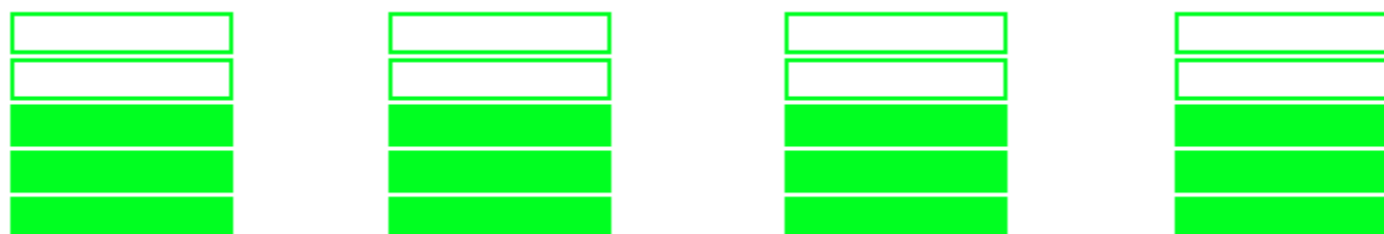
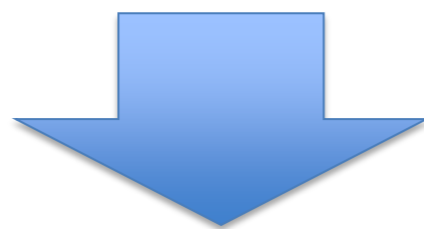
Серебряной пули не бывает



Недостатки

- Сложно добавлять машины
- Невозможность управления нагрузкой на машины
- Сложно восстанавливать

Добавление машины



Управляемый Elliptics

Устройство

- Mastermind – распределённый балансировщик нагрузки
- Работает под Cocaine
- Много маленьких групп вместо больших
- Куда писать – решает балансировщик, учитывая нагрузку и ресурсы нод

Достоинства

- Легко и просто восстанавливать
- Подходит для бесконечных хранилищ UGC
- Нет мастер-нод

Серебряной пули не бывает



Недостатки

- Надо где-то хранить номера групп
- У каждого сервиса должна быть метабазы
- Метабазу можно поднять и в Elliptics, но доступ будет за 2 хита

Эволюция Elliptics

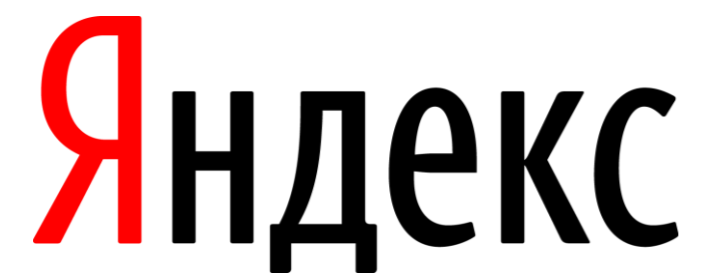
DHT инсталляция

Управляемый Elliptics



Где почитать

- <https://github.com/reverbrain/elliptics/>
- <http://elliptics.ru/> (есть форум)
- <http://www.ioremap.net/>



Антон КОРТУНОВ

Разработчик
распределённого хранилища

toshik@yandex-team.ru