

Storage для Lake'a

Alexander Kazanskiy

План

- Что за Lake
- Что за серh
- Как это работает
- Мифы
- Как ломается
- Всё



Эволюция архитектуры

Эволюция дата архитектур

Data Warehouse

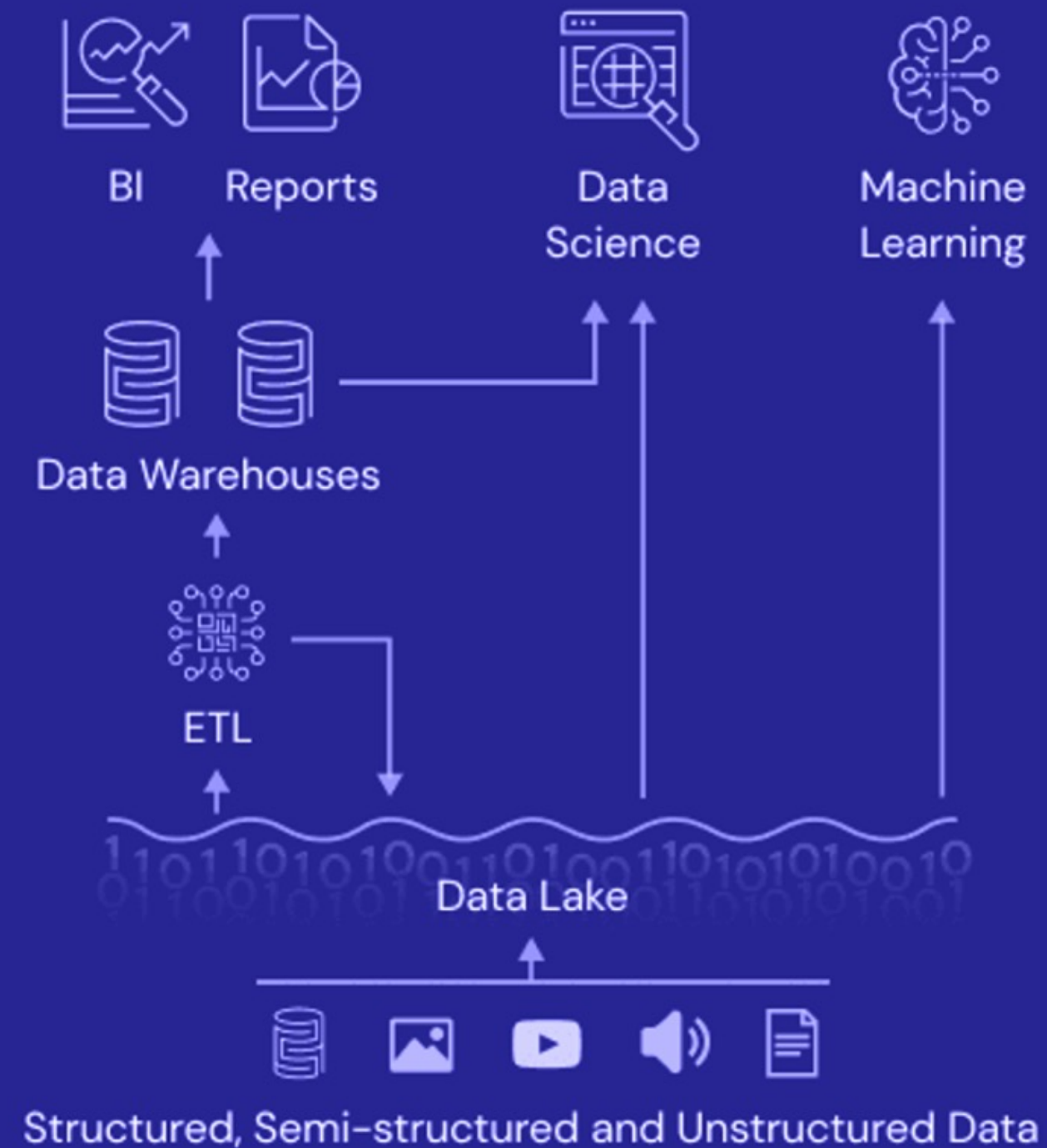


Эволюция дата архитектур

Data Warehouse



Data Lake

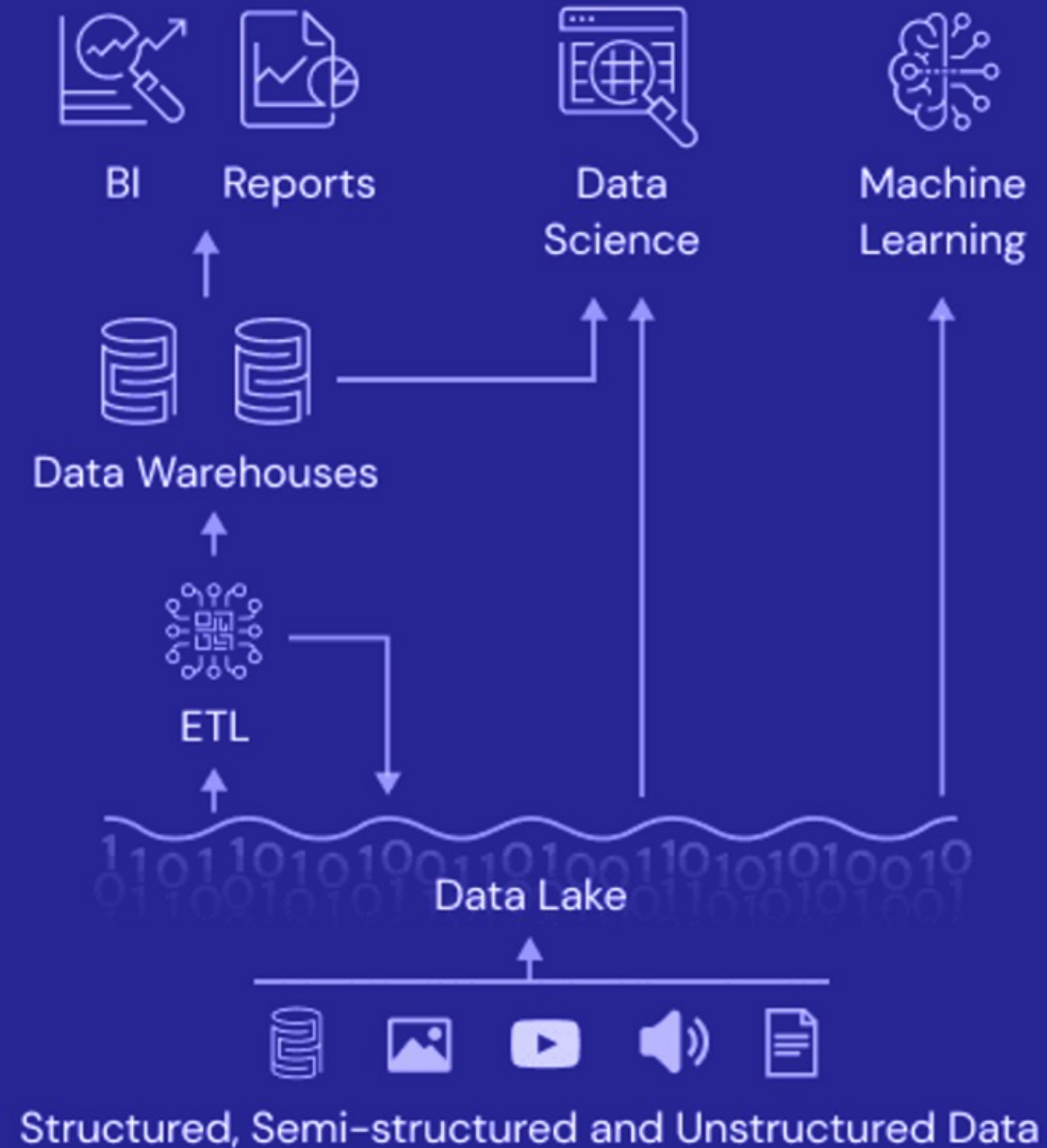


Эволюция дата архитектур

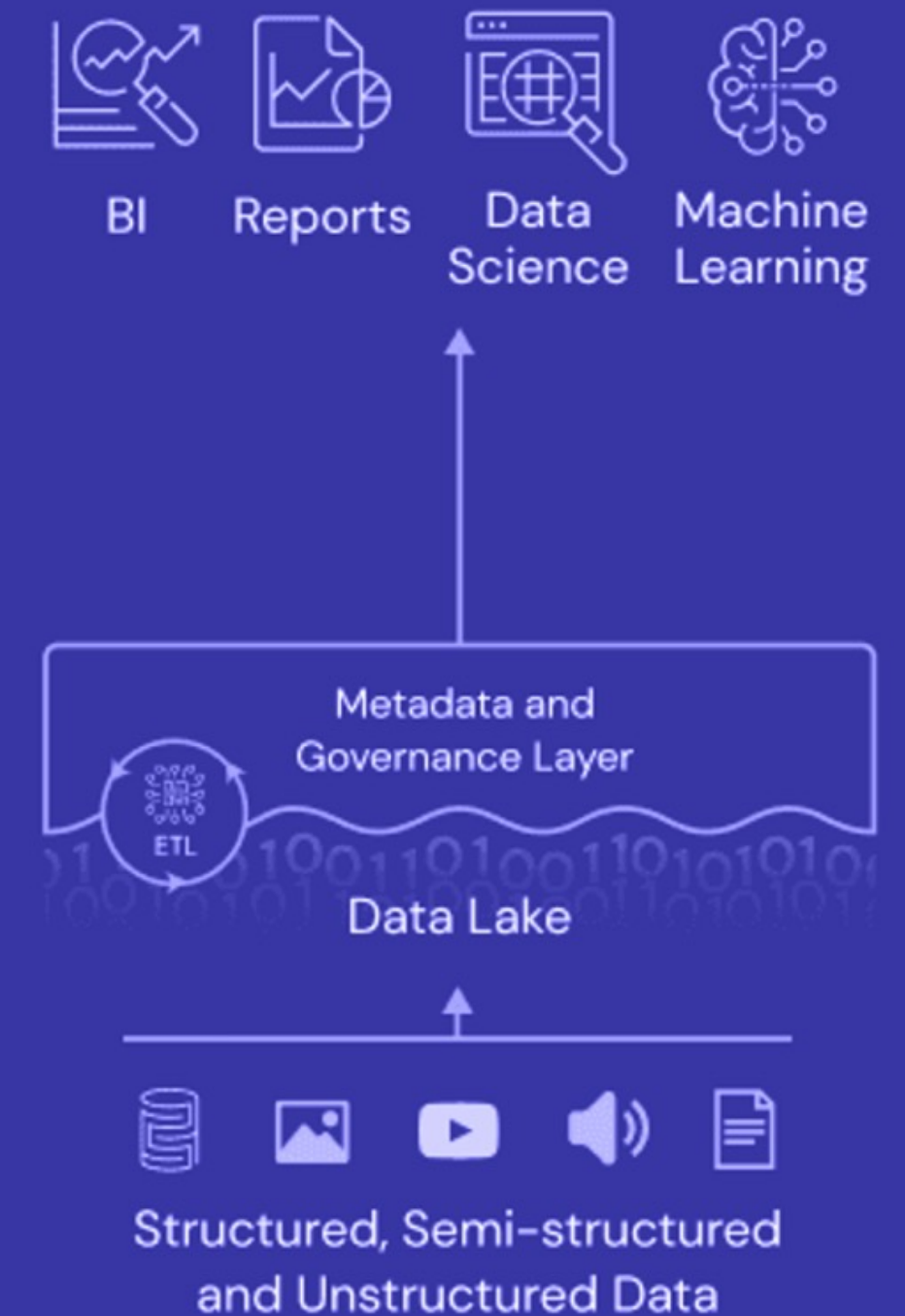
Data Warehouse



Data Lake



Data Lakehouse



Что такое серп



Популярные мифы

Мифы

01

Надёжность

С ним легко потерять
данные

02

Сложность

Нам нужно много
крутых спецов чтобы
его поддерживать

03

Масштабируемость

Цеф сложно
масштабировать

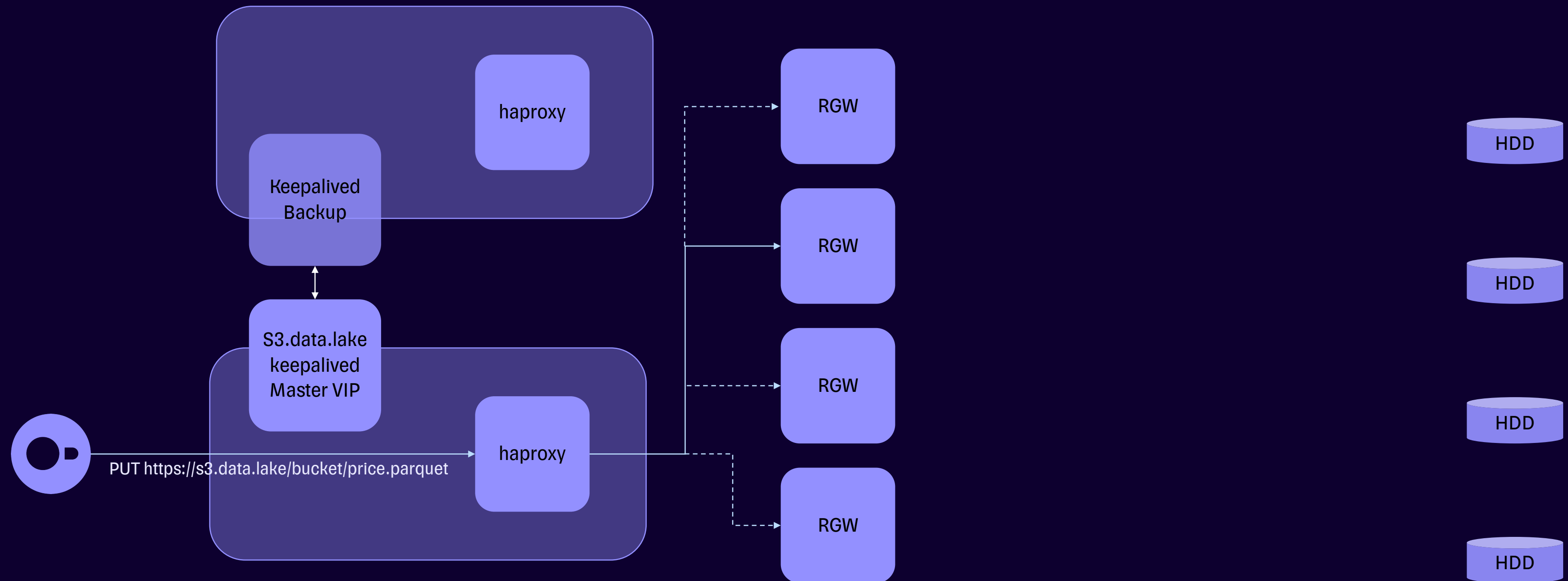
Движение запроса через систему



Движение запроса через систему

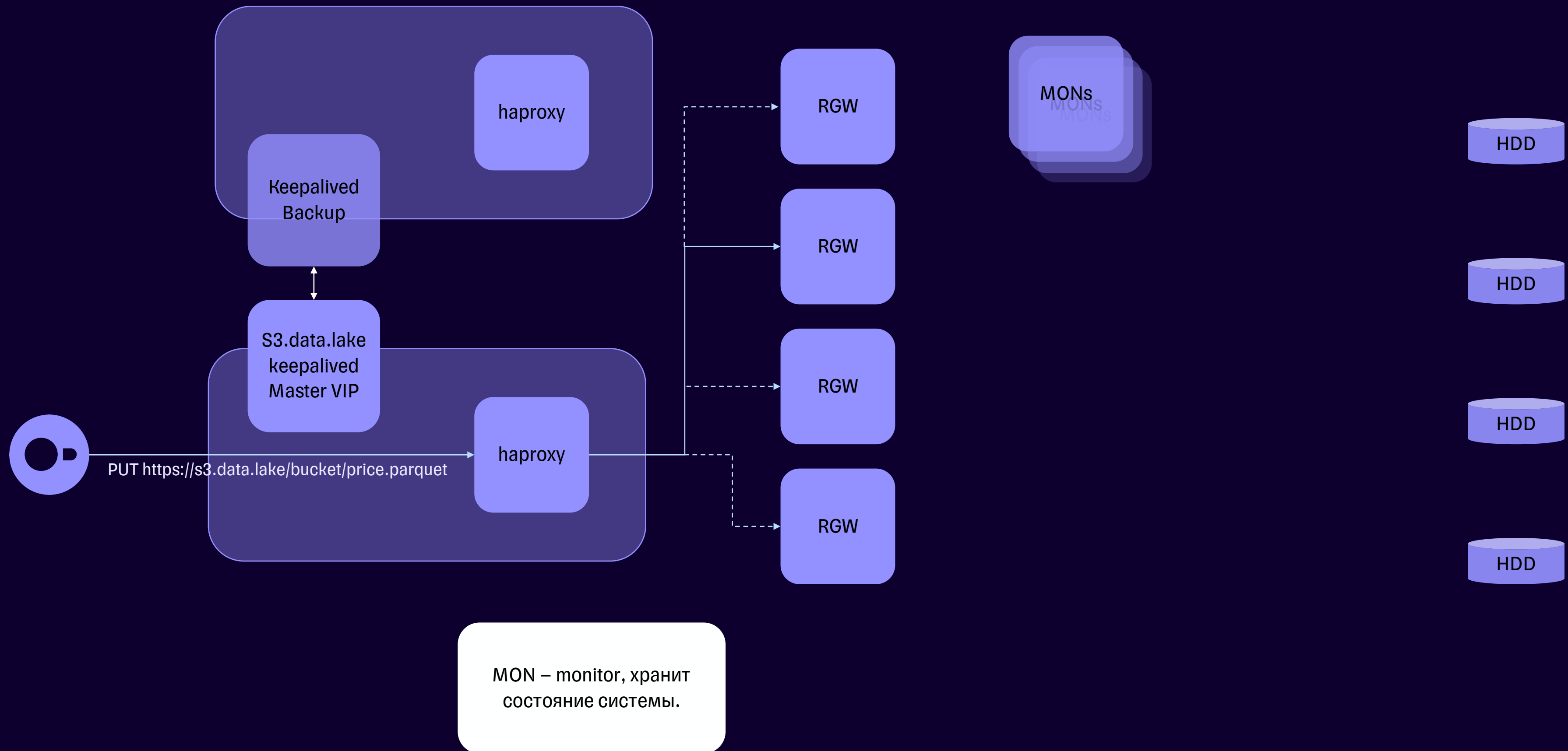


Движение запроса через систему

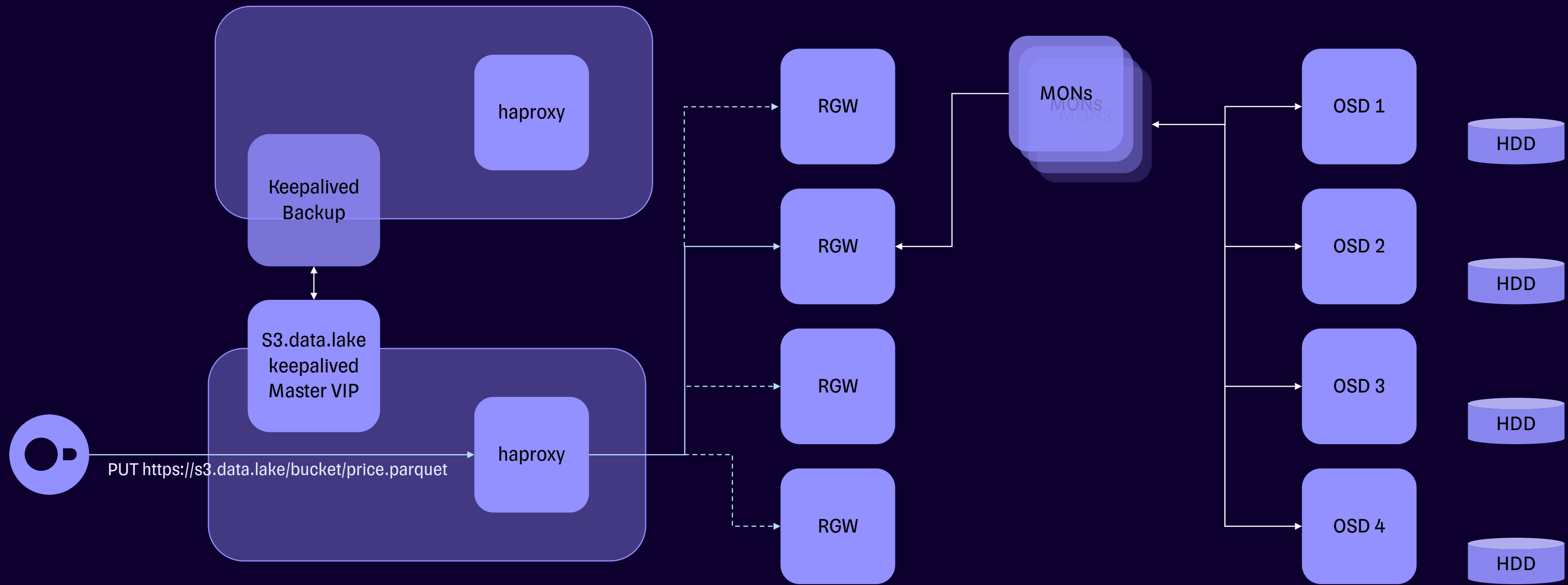


RGW – rados gateway, демон реализующий restful api.

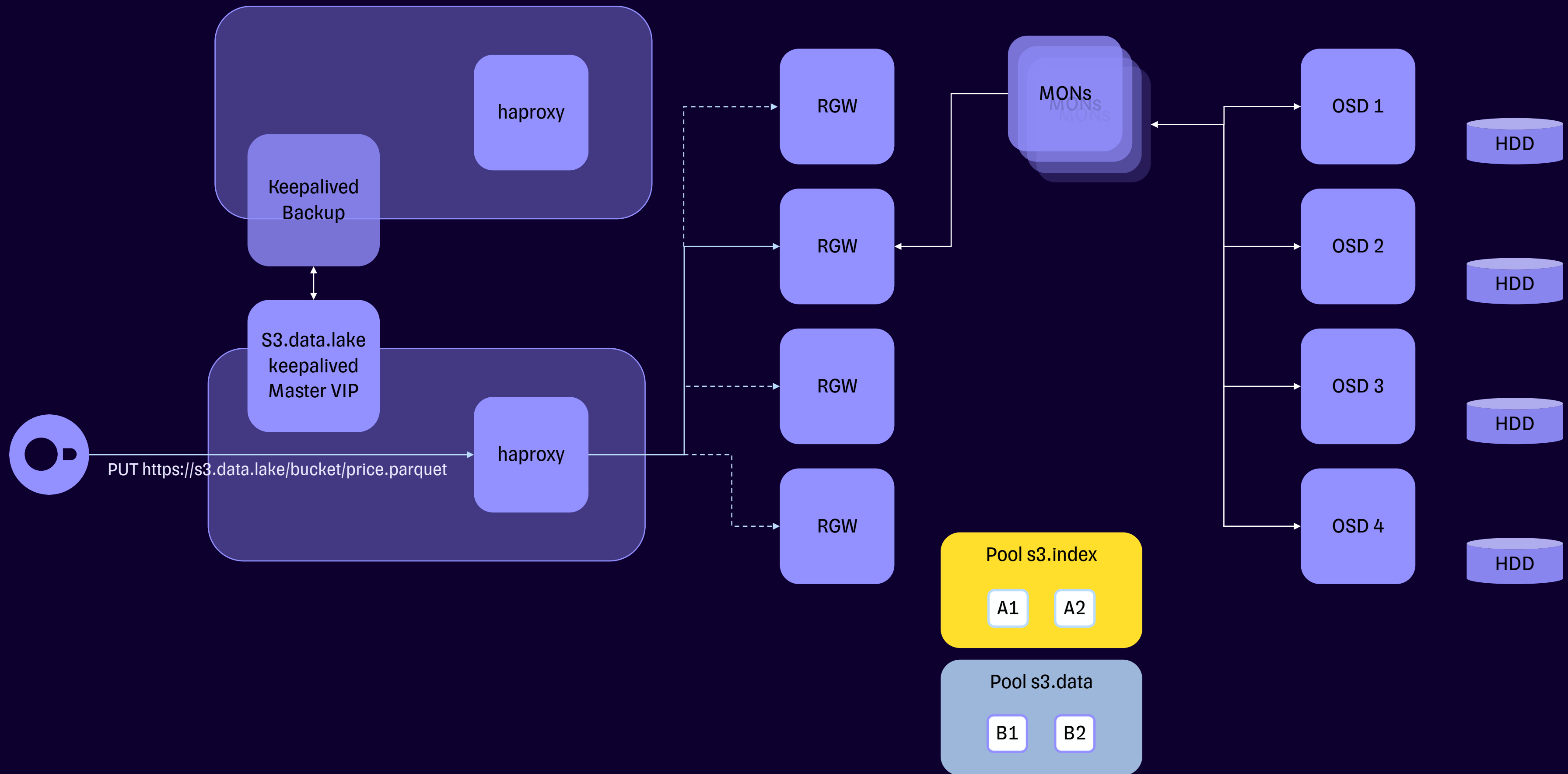
Движение запроса через систему



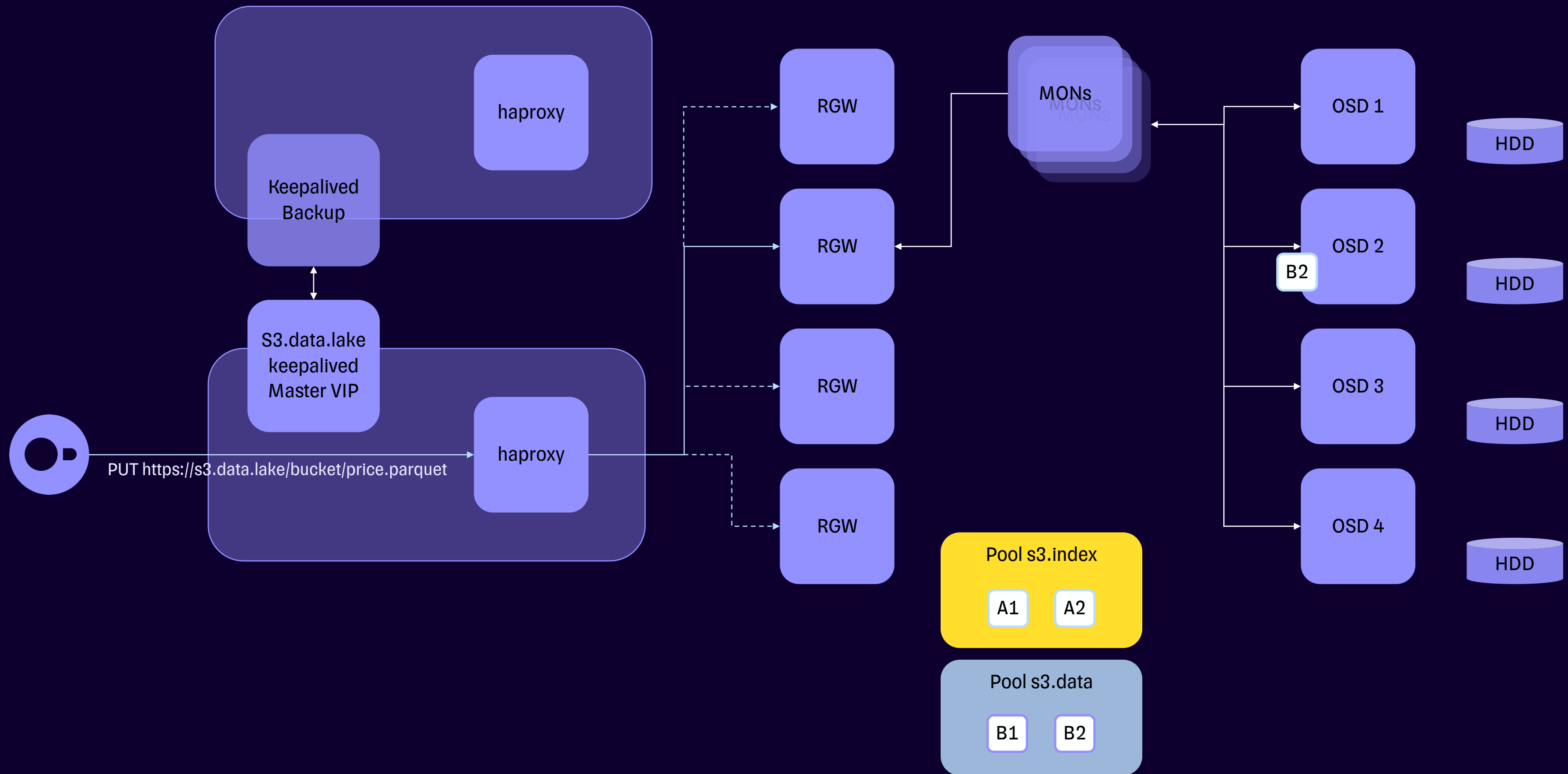
Движение запроса через систему



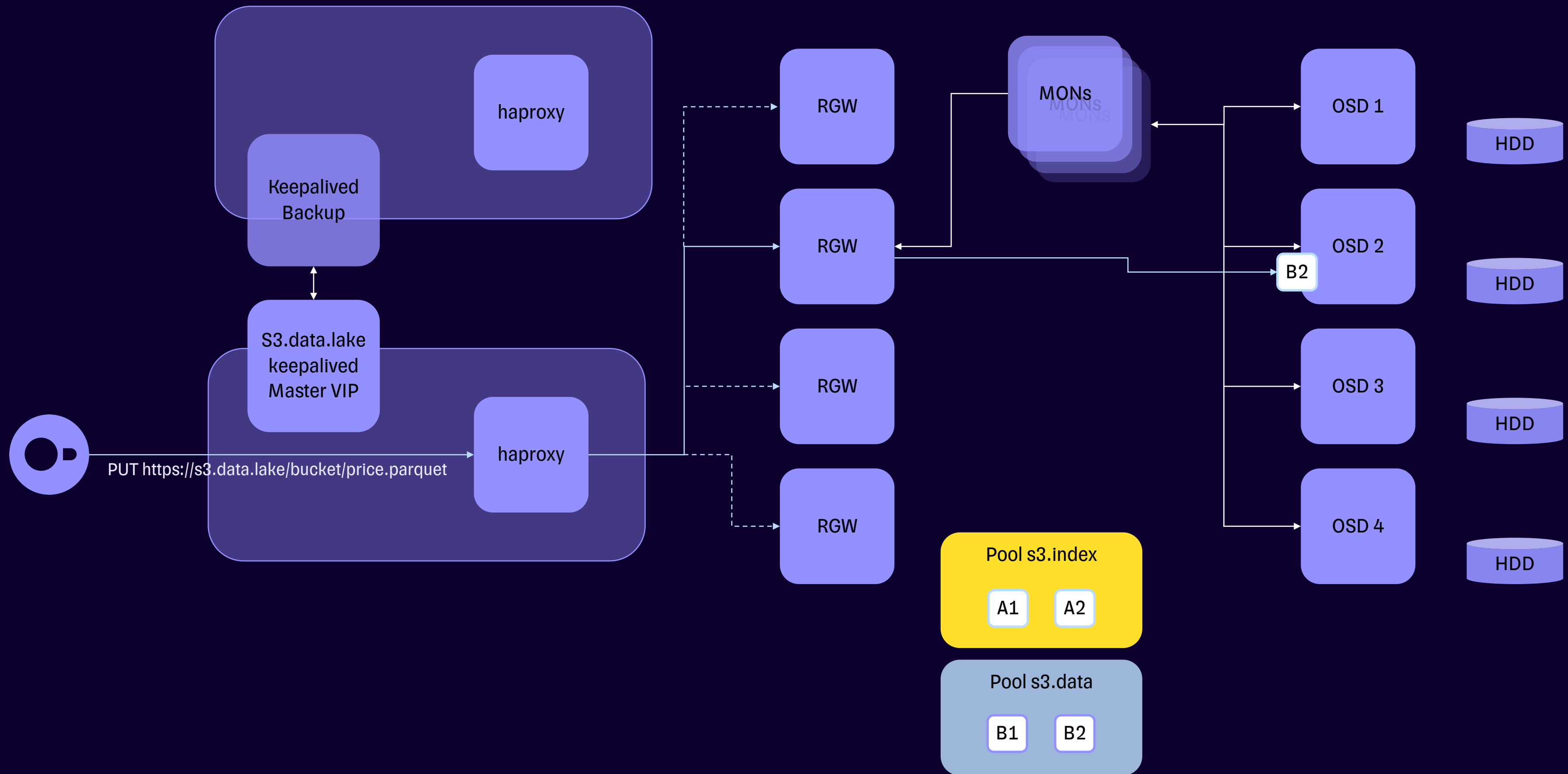
Движение запроса через систему



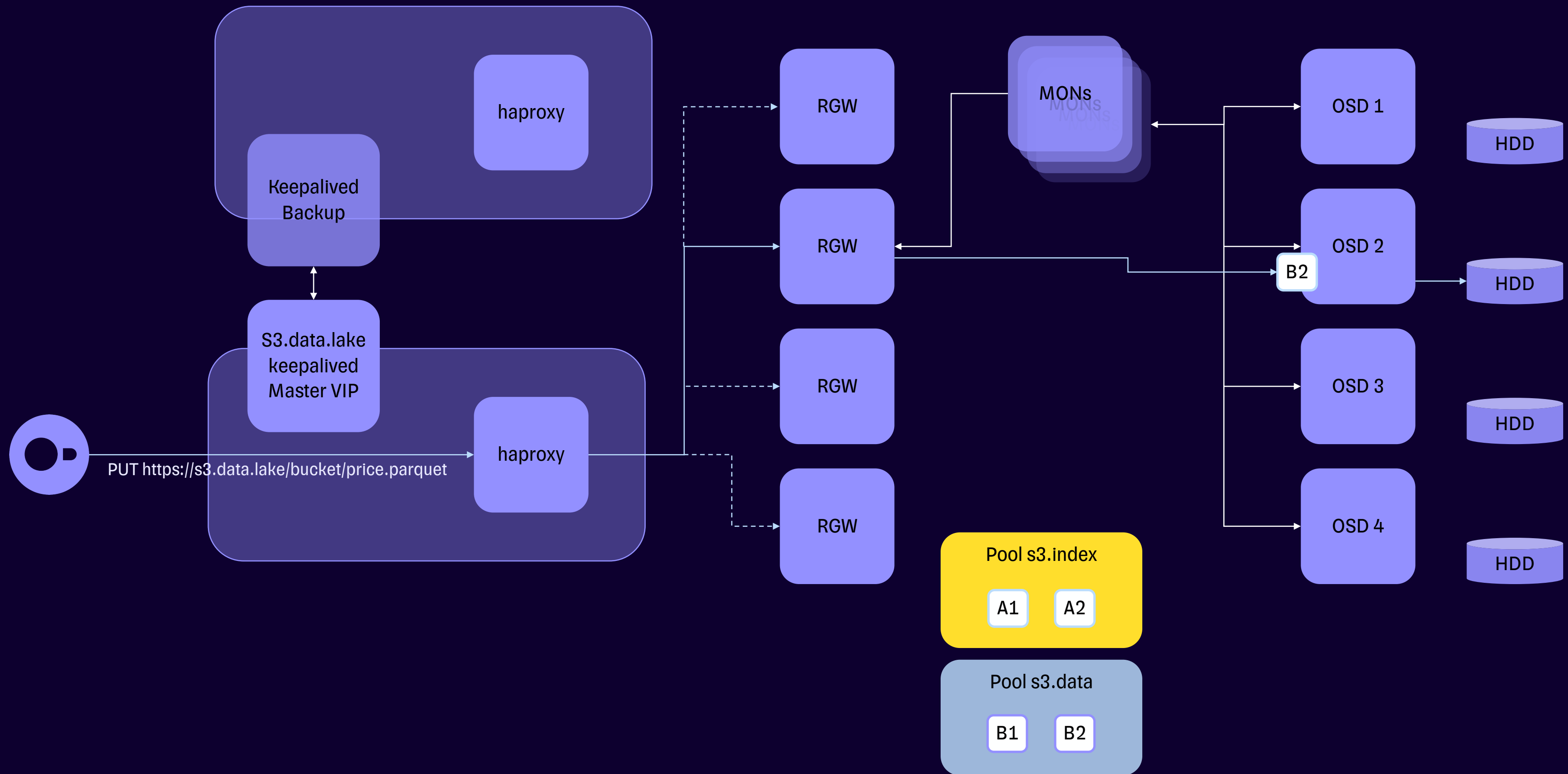
Движение запроса через систему



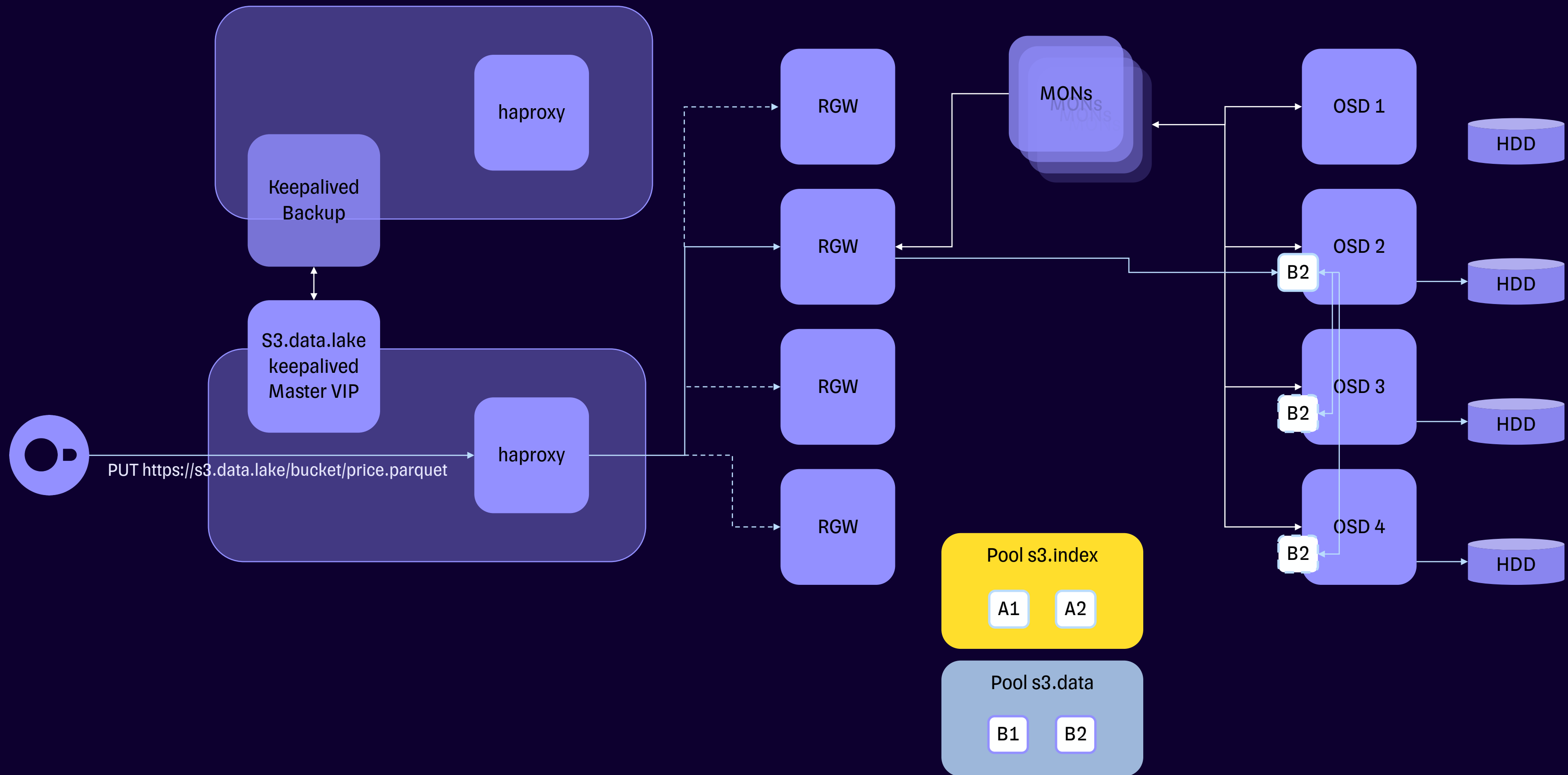
Движение запроса через систему



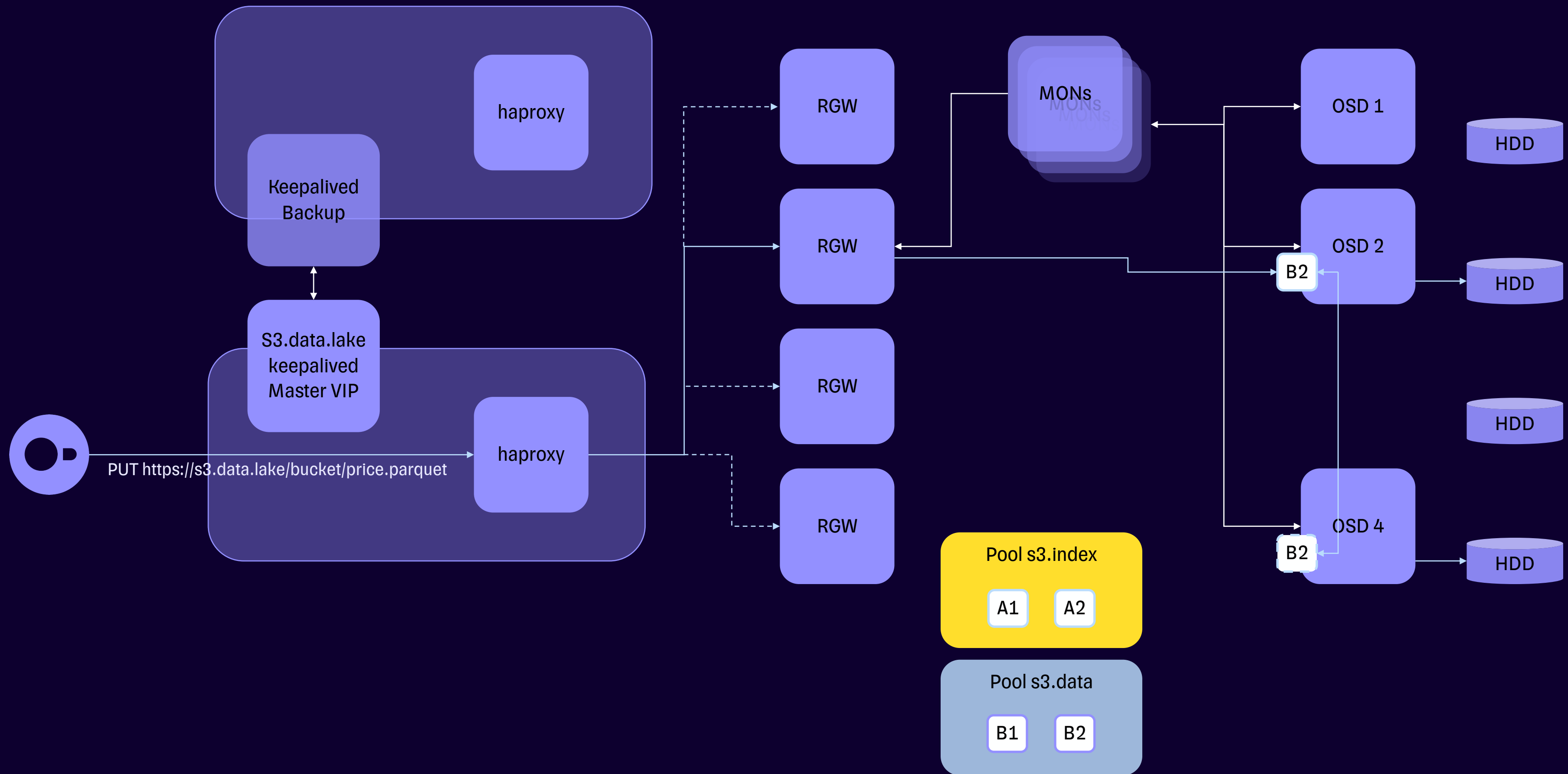
Движение запроса через систему



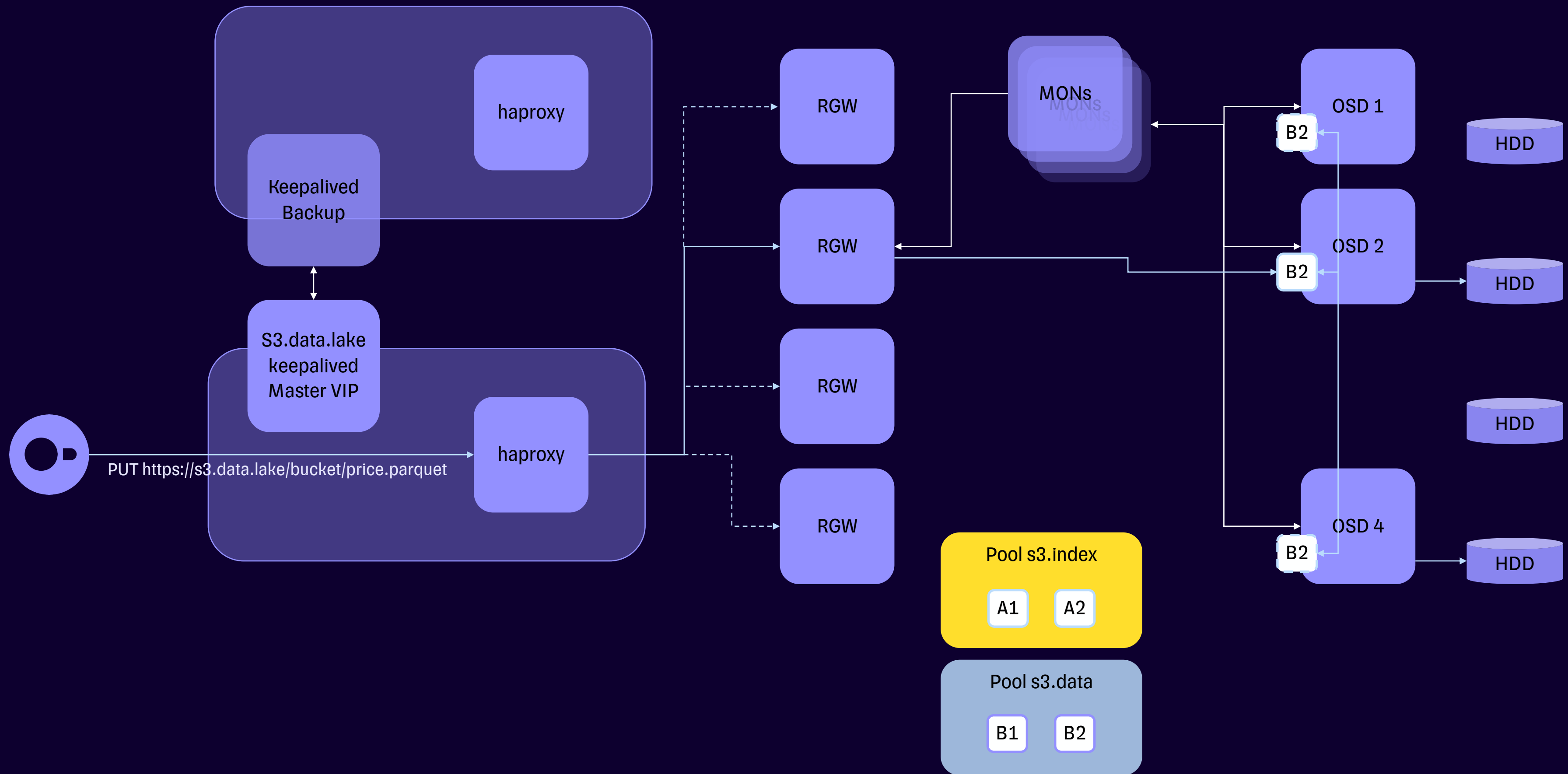
Движение запроса через систему



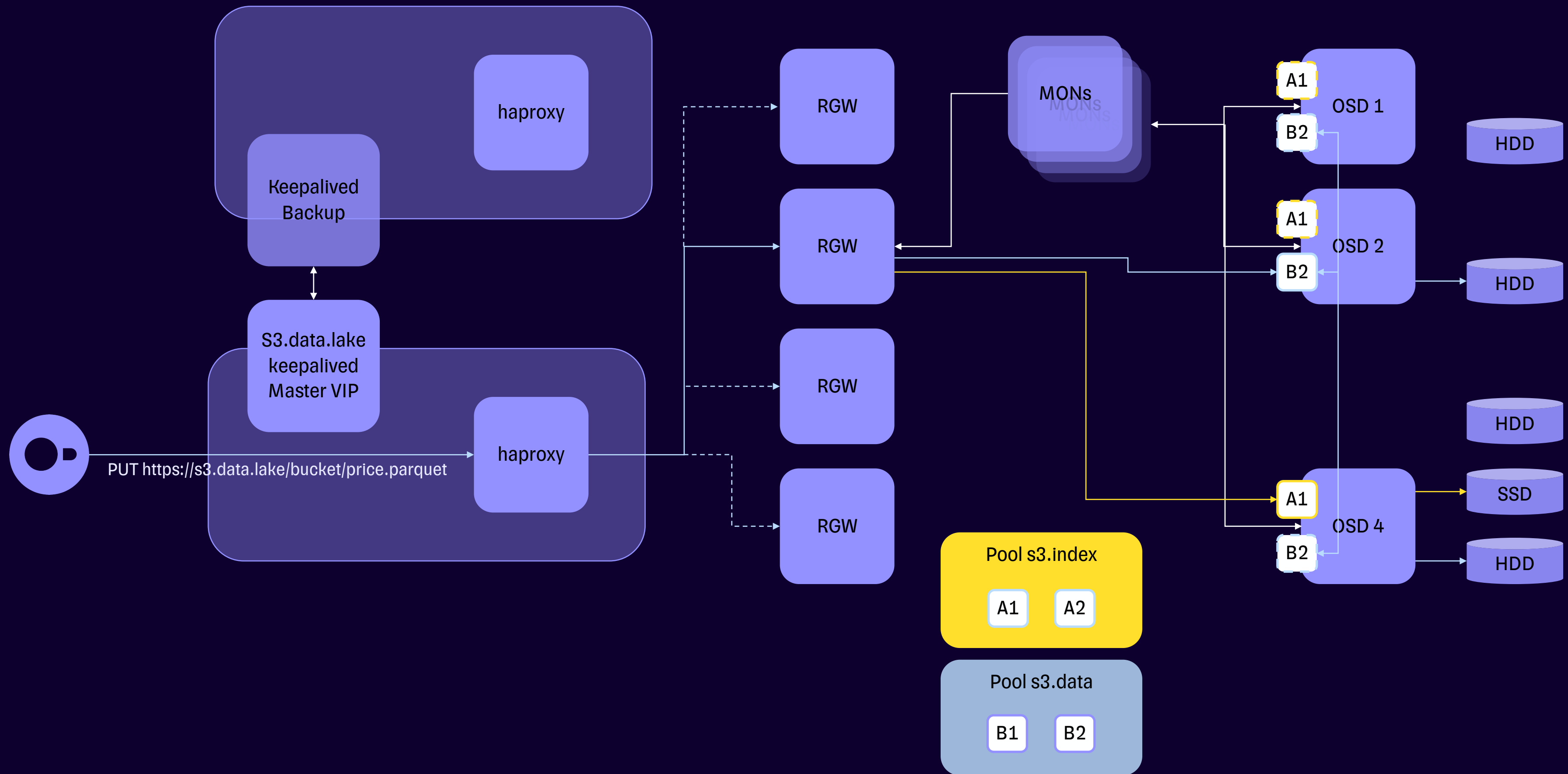
Движение запроса через систему



Движение запроса через систему



Движение запроса через систему





Масштабируемость

Масштабируемость

Масштабируемость



Гибко масштабируемая архитектура

Хотите горизонтально

Хотите вертикально

Масштабируемость



Гибко масштабируемая архитектура

Хотите горизонтально

Хотите вертикально



Нет зафиксированных лимитов

Ceph delivers extraordinary scalability—thousands of clients accessing petabytes to exabytes of data.

Масштабируемость



Гибко масштабируемая архитектура

Хотите горизонтально
Хотите вертикально



Нет зафиксированных лимитов

Ceph delivers extraordinary scalability—thousands of clients accessing petabytes to exabytes of data.



Публичные референсы

Мой опыт - 20+ PiB
CERN 100+ PiB,
200PiB DigitalOcean, и десятки миллиардов объектов

Масштабируемость



Гибко масштабируемая архитектура

Хотите горизонтально
Хотите вертикально



Нет зафиксированных ЛИМИТОВ

Ceph delivers extraordinary scalability—thousands of clients accessing petabytes to exabytes of data.



Публичные референсы

Мой опыт - 20+ PiB
CERN 100+ PiB,
200PiB DigitalOcean, и десятки миллиардов объектов



Нет известных экзабайтных кластеров

Возможно потому, что пока никому не нужно :peka:



СЛОЖНОСТЬ

Циф сложный

Цеф сложный



**Цеф это как кубер –
всего 5 бинарей**

И у вас готовый кластер S3 :река:

Цеф сложный



**Цеф это как кубер –
всего 5 бинарей**

И у вас готовый кластер S3 :река:



Фокус разработки на упрощение

Разработчики тратят много сил
на то чтоб сделать серh проще

Цеф сложный



**Цеф это как кубер –
всего 5 бинарей**

И у вас готовый кластер S3 :река:



Фокус разработки на упрощении

Разработчики тратят много сил
на то чтоб сделать серh проще



**Крутая документация
и не одна**

Разработчики вкладываются как в офф-доку, так
и в собственные документации, KB, живая
рассылка, посты в блогах и т.д.

Цеф сложный



**Цеф это как кубер –
всего 5 бинарей**

И у вас готовый кластер S3 :река:



Фокус разработки на упрощении

Разработчики тратят много сил
на то чтоб сделать серh проще



**Крутая документация
и не одна**

Разработчики вкладываются как в офф-доку, так
и в собственные документации, KB, живая
рассылка, посты в блогах и т.д.



Цеф сложный

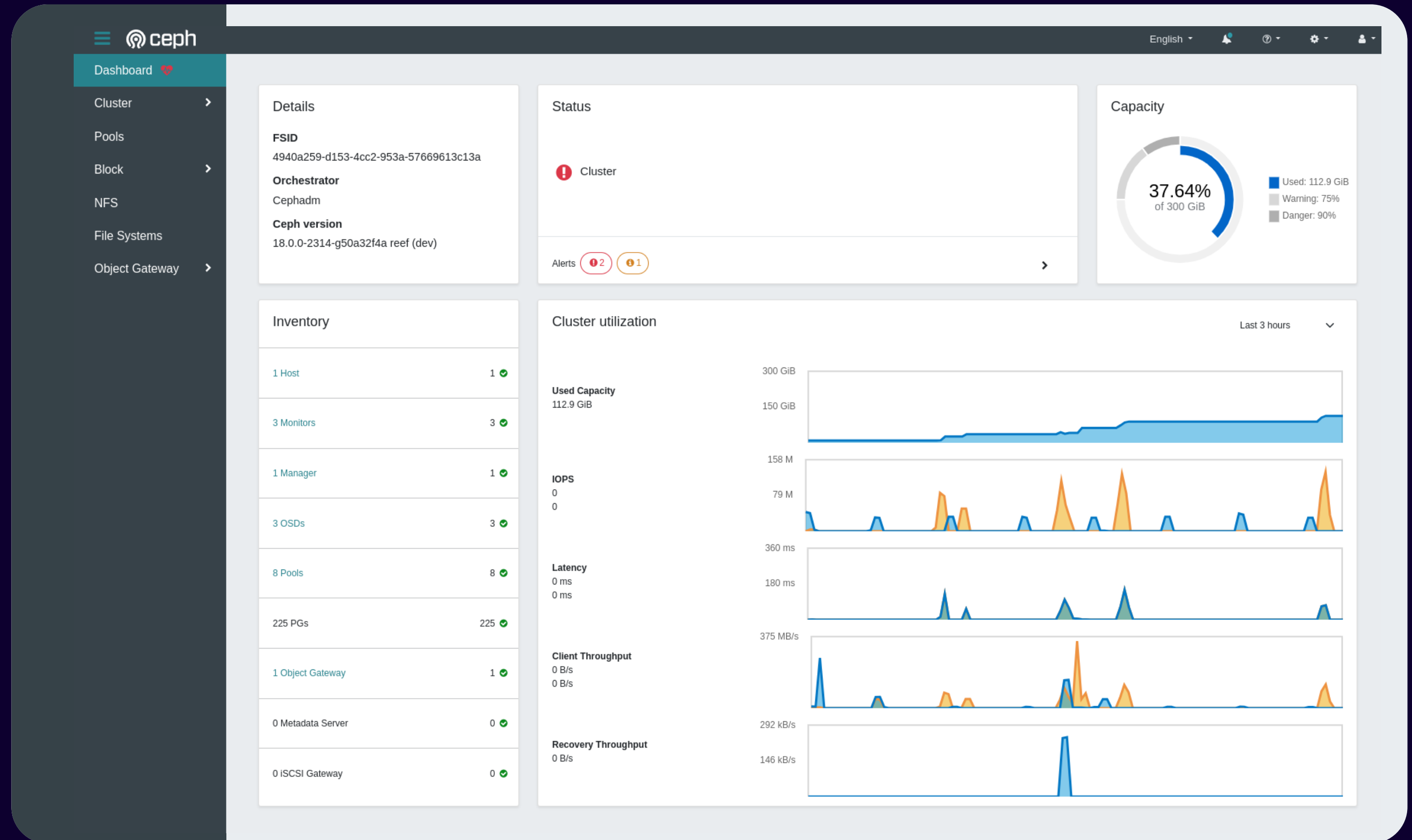
Чем глубже погружаешься тем сложнее
становится

ceph manager daemon

- Installation and Configuration
- Writing modules
- Writing orchestrator plugins
- Dashboard module
- Ceph RESTful API
- Alerts module
- DiskPrediction module
- Local pool module
- RESTful module
- Zabbix module
- Prometheus module
- Influx module
- Hello module
- Telegraf module
- Telemetry module
- Iostat module
- Crash module
- Insights module
- Orchestrator module
- Rook module
- RGW module
- MDS Autoscaler module
- NFS module
- SMB module
- Progress Module
- CLI API Commands module

ceph manager daemon

- Installation and Configuration
- Writing modules
- Writing orchestrator plugins
- Dashboard module
- Ceph RESTful API
- Alerts module
- DiskPrediction module
- Local pool module
- RESTful module
- Zabbix module
- Prometheus module
- Influx module
- Hello module
- Telegraf module
- Telemetry module
- Iostat module
- Crash module
- Insights module
- Orchestrator module
- Rook module
- RGW module
- MDS Autoscaler module
- NFS module
- SMB module
- Progress Module
- CLI API Commands module



Cephadm



MicroCeph

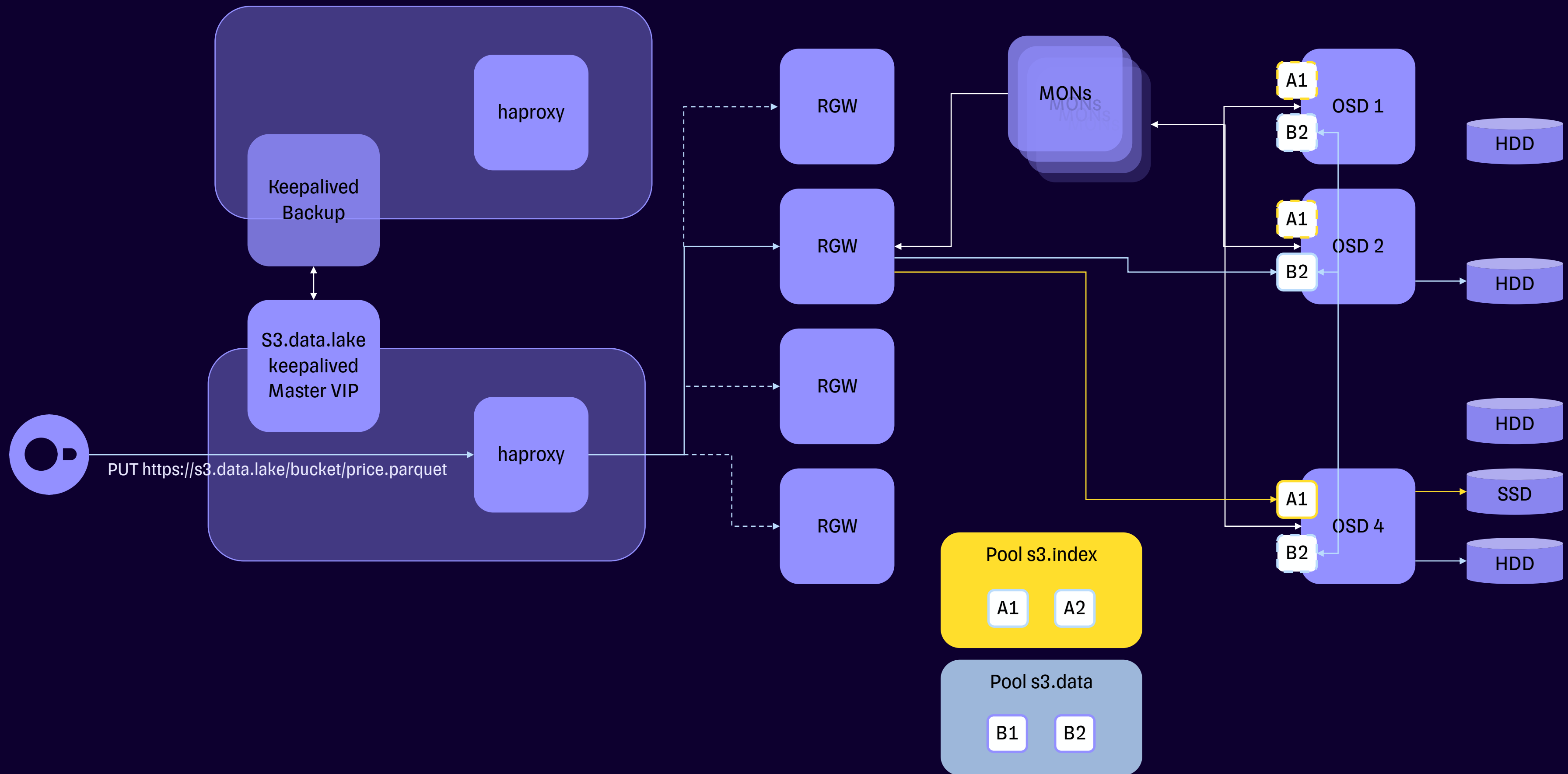
MicroCeph

```
sudo snap install microceph --channel quincy/edge  
sudo snap refresh --hold microceph  
sudo microceph cluster bootstrap  
sudo microceph disk add loop,4G,3  
sudo ceph status
```



Надёжность

Движение запроса через систему



Нет дыма без 🔥

Нет дыма без 🔥

Безумная
экономика



Нет дыма без 🔥

Безумная
экономика



Отсутствие
проектирования



Нет дыма без 🔥

Безумная
экономика



Отсутствие
проектирования



Поспешишь - комьюнити
насмешишь



Выводы



Серh – это SDS общего назначения



Надёжный

Обеспечивает надёжное, отказоустойчивое хранение данных и умеет обрабатывать различные отказы, и самовосстанавливаться



Масштабируемость

Теоретически безграничное горизонтальное масштабирование



Функциональный

Широкая функциональность, серh это не только крутой S3.



Opensource и комьюнити

Открытый код, много разработчиков, много потребителей, большое активное комьюнити.

Открытость и распространённость

Разрабатывают



ORACLE



ubuntu[®]
Delivered by Canonical

FUJITSU



Используют

Bloomberg[®]

DigitalOcean

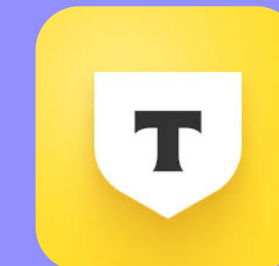
OVHcloud



Selectel



BOSTON
UNIVERSITY



WB



RU КОМЬЮНИТИ В ТЕЛЕГЕ

https://t.me/serph_ru



Выводы

**серh – это крутой продукт,
который обязательно стоит
рассматривать при строительстве
больших хранилищ**

Выводы

**серh – это крутой продукт,
который обязательно стоит
рассматривать при строительстве
больших хранилищ**



Не так страшен серh как его малюют

Цеф может быть сложен для глубокого понимания всех нюансов системы, но для большинства задач он не сложнее, а где-то и проще, других СХД

Выводы

серh – это крутой продукт, который обязательно стоит рассматривать при строительстве больших хранилищ

Не пренебрегайте проектированием

Когда вы строите что-то большое и сложное, посчитав и прописав всё на старте, вы скорее всего сможете сэкономить время и деньги. Если вы получили неожиданные результаты при расчётах разберитесь почему, не забывайте включать критическое мышление и не верить в сказки



Не так страшен серh как его малюют

Цеф может быть сложен для глубокого понимания всех нюансов системы, но для большинства задач он не сложнее, а где-то и проще, других СХД

Выводы

серh – это крутой продукт, который обязательно стоит рассматривать при строительстве больших хранилищ

 **Не пренебрегайте проектированием**

Когда вы строите что-то большое и сложное, посчитав и прописав всё на старте, вы скорее всего сможете сэкономить время и деньги. Если вы получили неожиданные результаты при расчётах разберитесь почему, не забывайте включать критическое мышление и не верить в сказки



Не так страшен серh как его малюют

Цеф может быть сложен для глубокого понимания всех нюансов системы, но для большинства задач он не сложнее, а где-то и проще, других СХД



Если вы не можете найти бородатого цефовода, наймите консалтера

Многие крутые ребята оказывают услуги консультации и поддержки, если у вас пока нет сильной команды они могут помочь сэкономить много сил и времени на разных этапах проекта



One more thing...

**Надёжность системы
не отменяет
необходимость делать
резервные копии
ваших данных**



Спасибо!