

КОНФЕРЕНЦИЯ **АРХИТЕКТУРА@**  
**ПРОГРАМ\*НЫХ СИСТЕМ**

**/ 7 февраля**



## Василенко Алексей, Неврев Семен

- Старший архитектор, разработчик
- Центр R&D БИТ
- ПАО МТС

Блокчейн в кармане. Выбор лучшего проекта RnD совета 2019

В докладе мы расскажем о реализации надежного голосования на мобильных устройствах используя современный алгоритм распределенного консенсуса.



# Выбор лучшего проекта RnD совета 2019

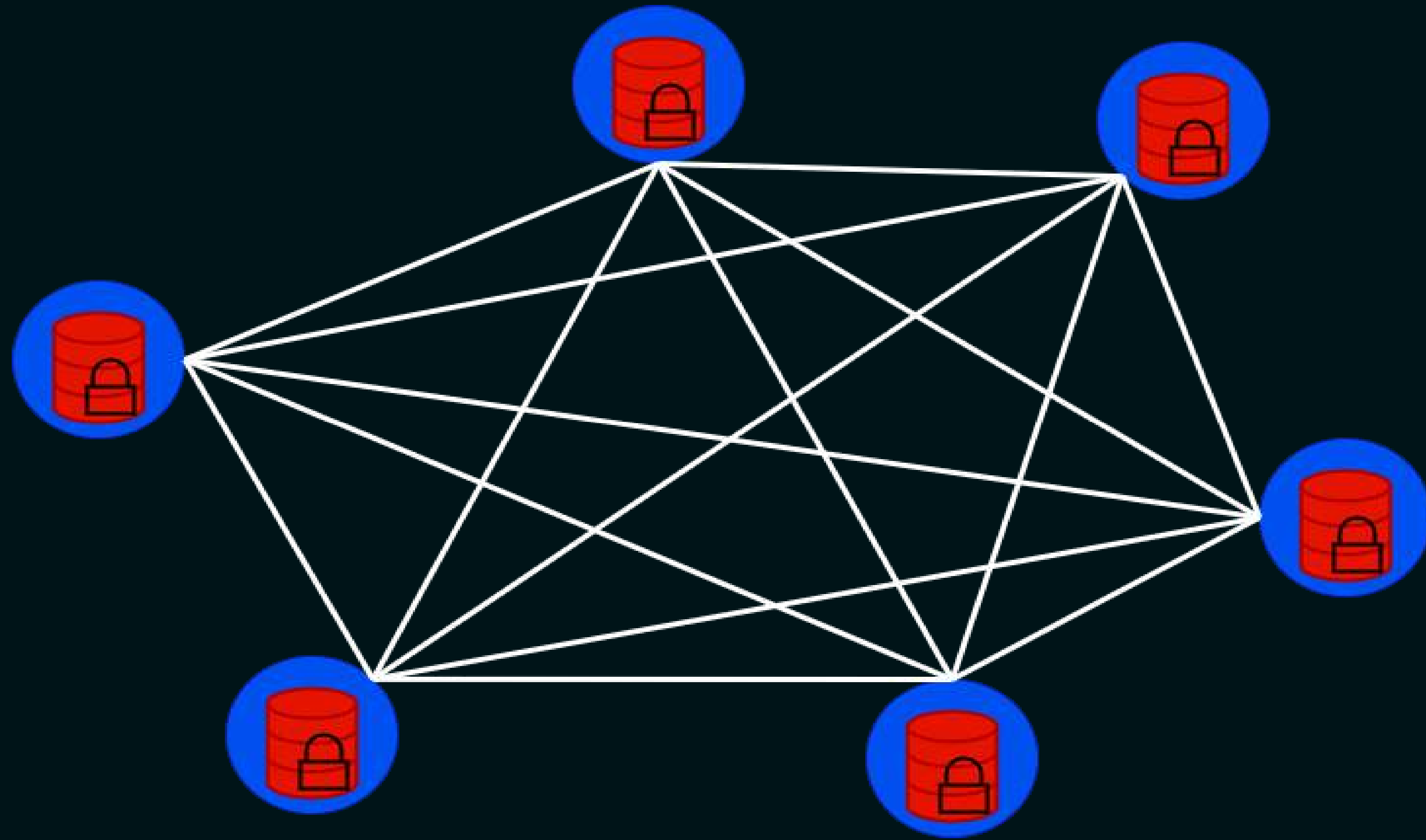
## *Необходимо:*

- Решить какой проект является лучшим в 2019 году
- Выбор должен быть наиболее честным и независимым
- Процесс выбора должен быть прозрачным и понятным как избирателям, так и наблюдателям

## *Что решили:*

- Воспользоваться голосованием
- Для прозрачности и честности реализовать процесс через приложение используя перспективную технологию распределенного реестра

# Технологии. Распределенный реестр



# Технологии. RAFT

## How does RAFT WORK?

RAFT IS A CONSENSUS ALGORITHM

used in many projects in the real world - ex. etcd

a process for getting multiple machines to "agree"!

**WHY?** Any part of a distributed system can fail! Consensus across several nodes gives us backups of data.

### WHAT ABOUT NETWORK PARTITIONS?!

In a partition, if the leader falls on the smaller side, a new leader is elected within the majority side. When the partition ends, messages received by the minority side are discarded, and those nodes converge their state to match the majority's.

### WHEN A NEW WRITE HAPPENS:



Any node can accept writes!



- RAFT doesn't stand for anything. It's a bunch of logs tied together.

Message is queued...

COPY me!



Message forwarded to leader node



ACK!

...while followers copy it to their logs

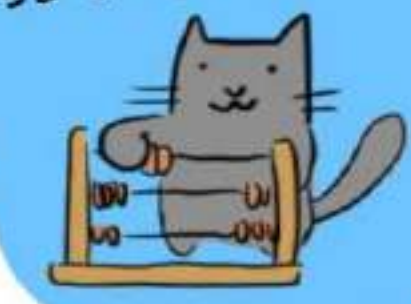


ACK

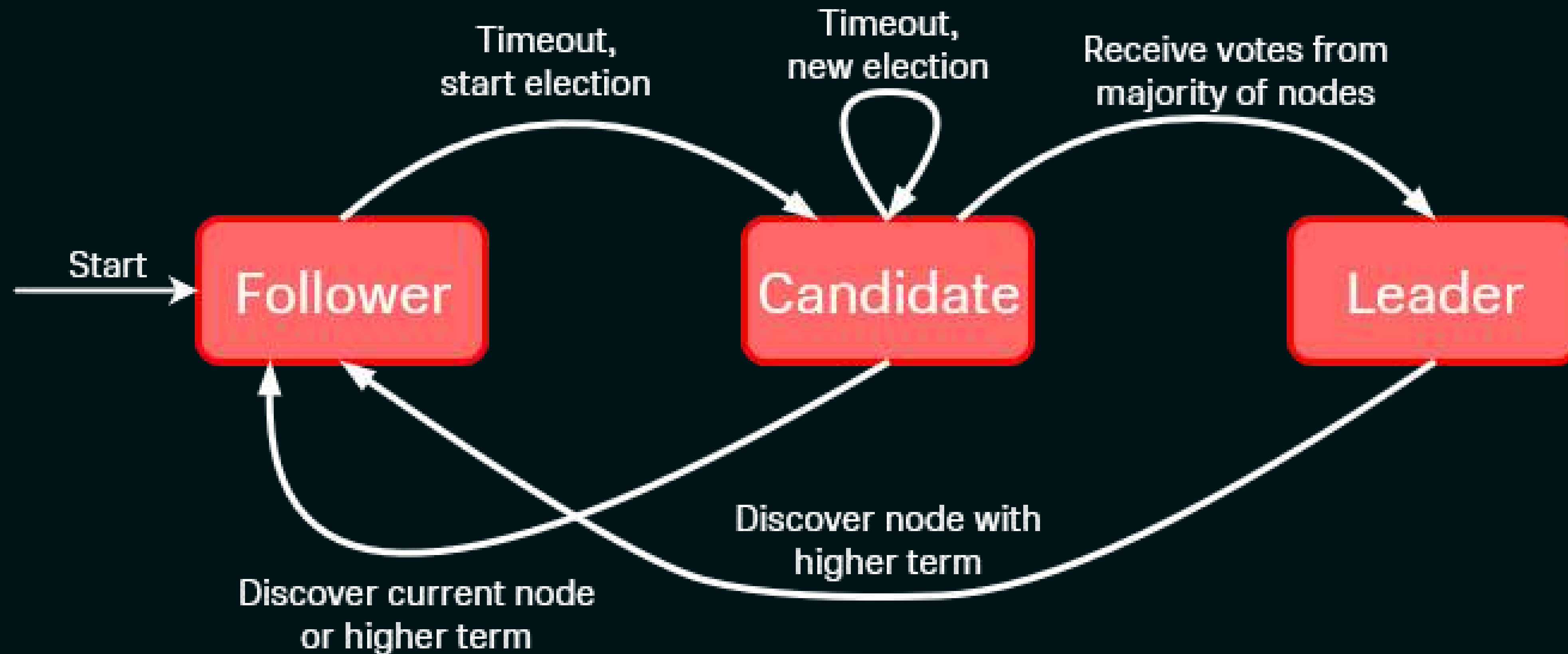


When a quorum have copied the message, leader "commits" the new data & will now return it on read

Followers also commit so they return new data!

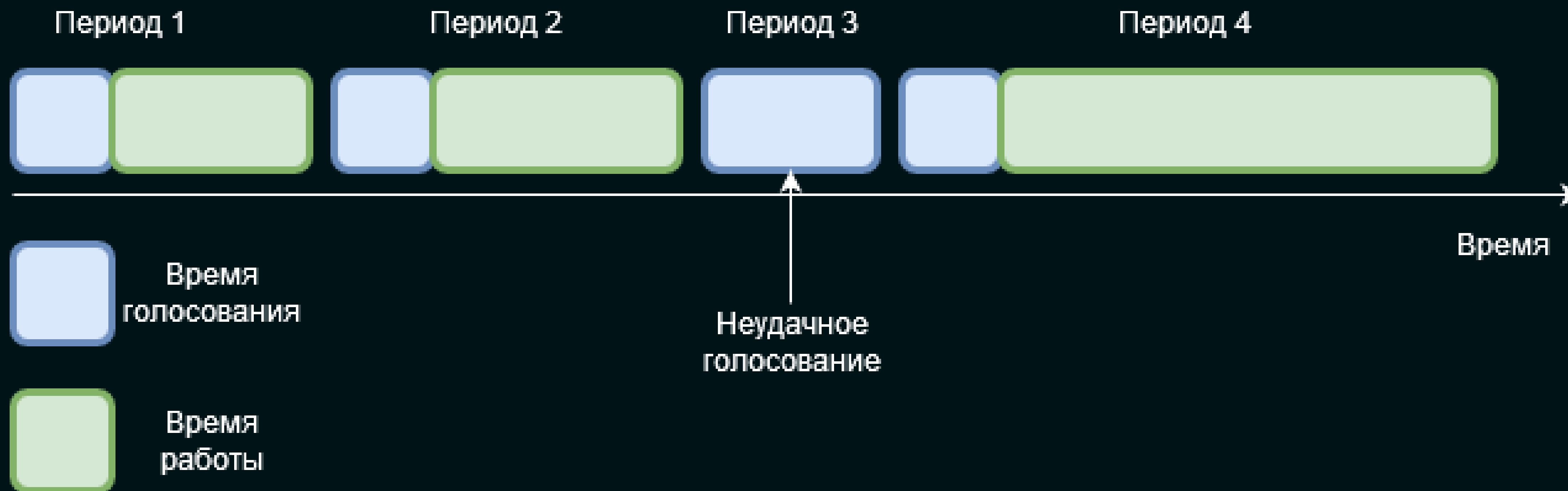


# Состояния узла



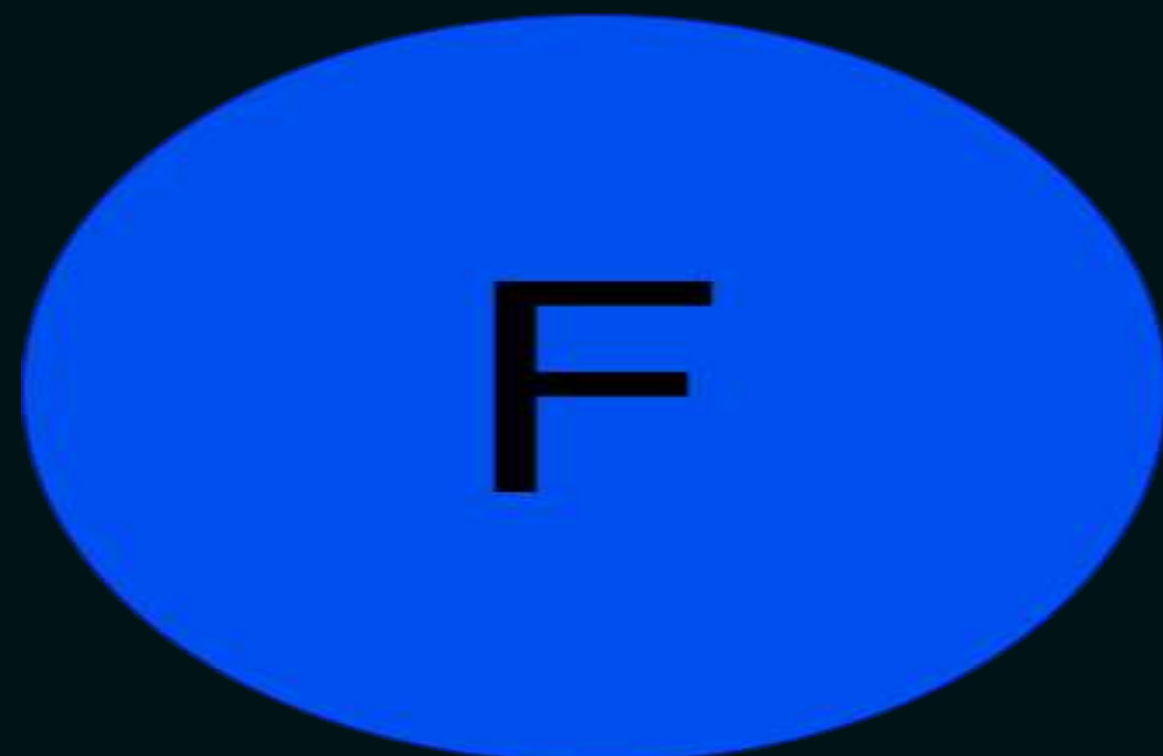
# Состояния системы

Время работы системы разделено на периоды избрания лидера (голосования) и его правления (непосредственной работы).

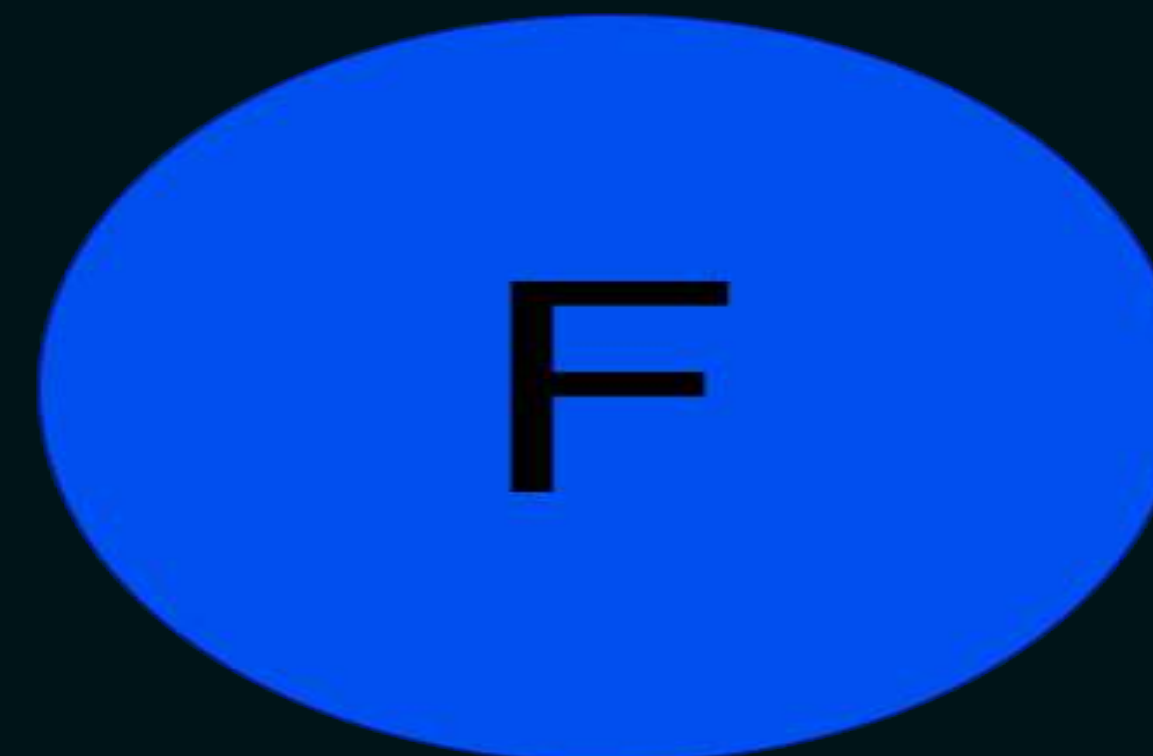


# Выбор лидера

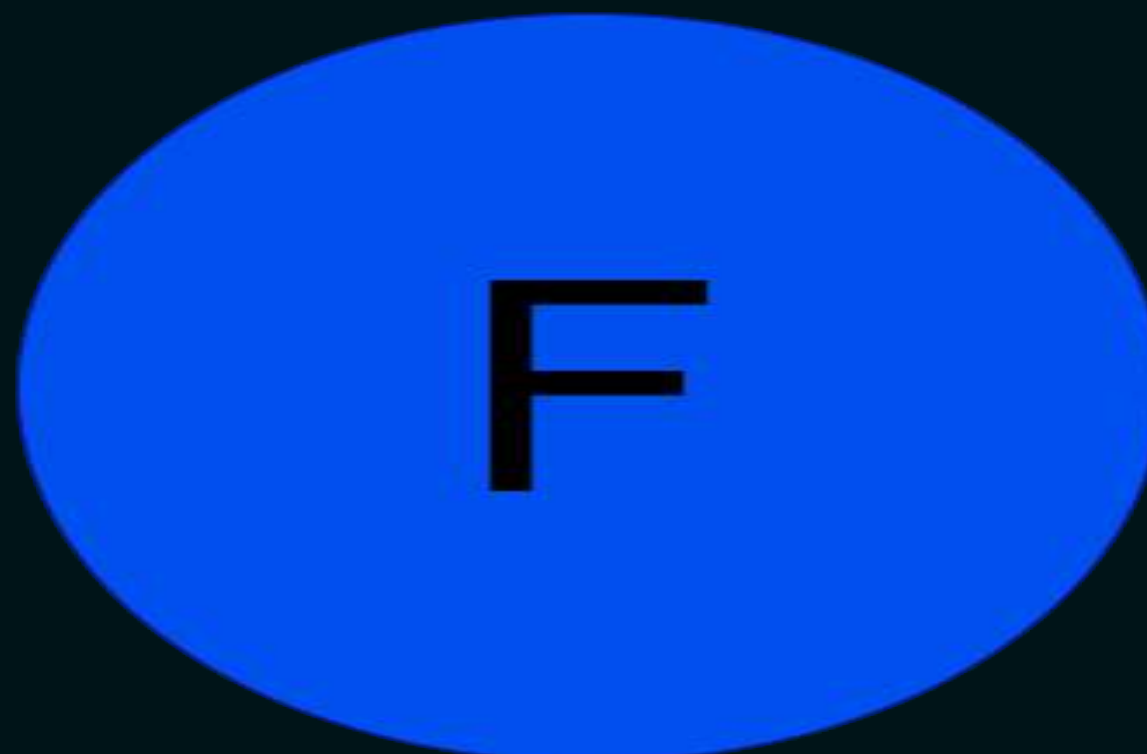
Node: A  
Term: 0  
LogSize: 0  
ElectionTimeout: 5



Node: B  
Term: 0  
LogSize: 0  
ElectionTimeout: 4



Node: C  
Term: 0  
LogSize: 0  
ElectionTimeout: 3





# Выбор лидера

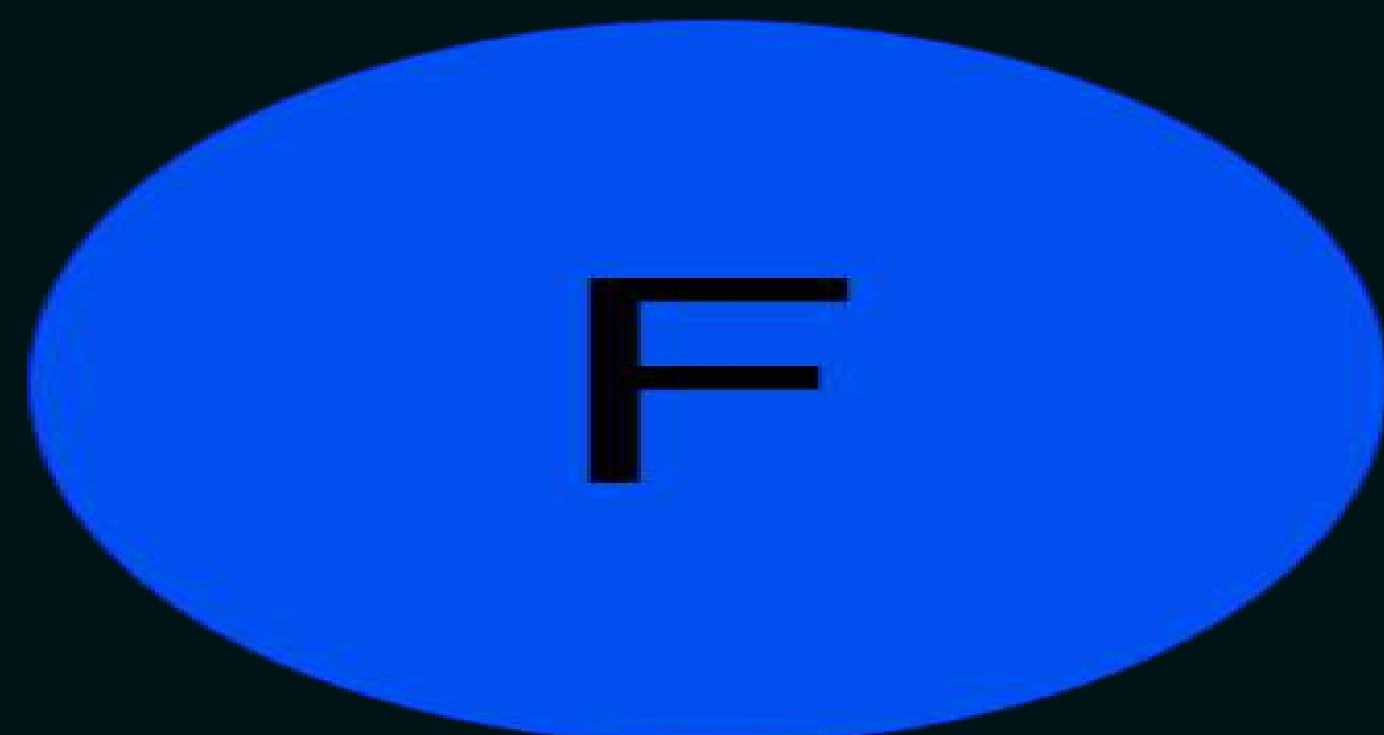
Node: A

*Term: 0*

**LogSize: 0**

ElectionTimeout: 2

VotedFor: C



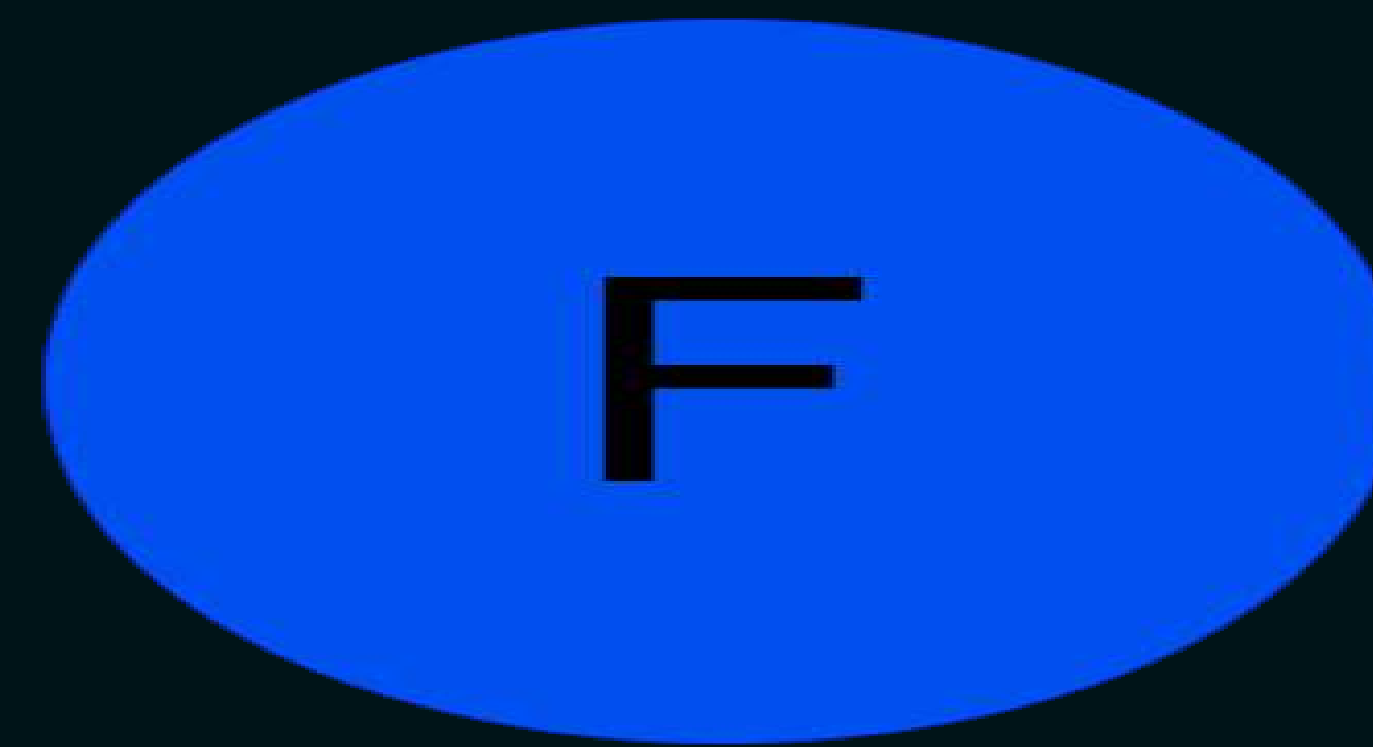
Node: B

*Term: 0*

**LogSize: 0**

ElectionTimeout: 1

VotedFor: C

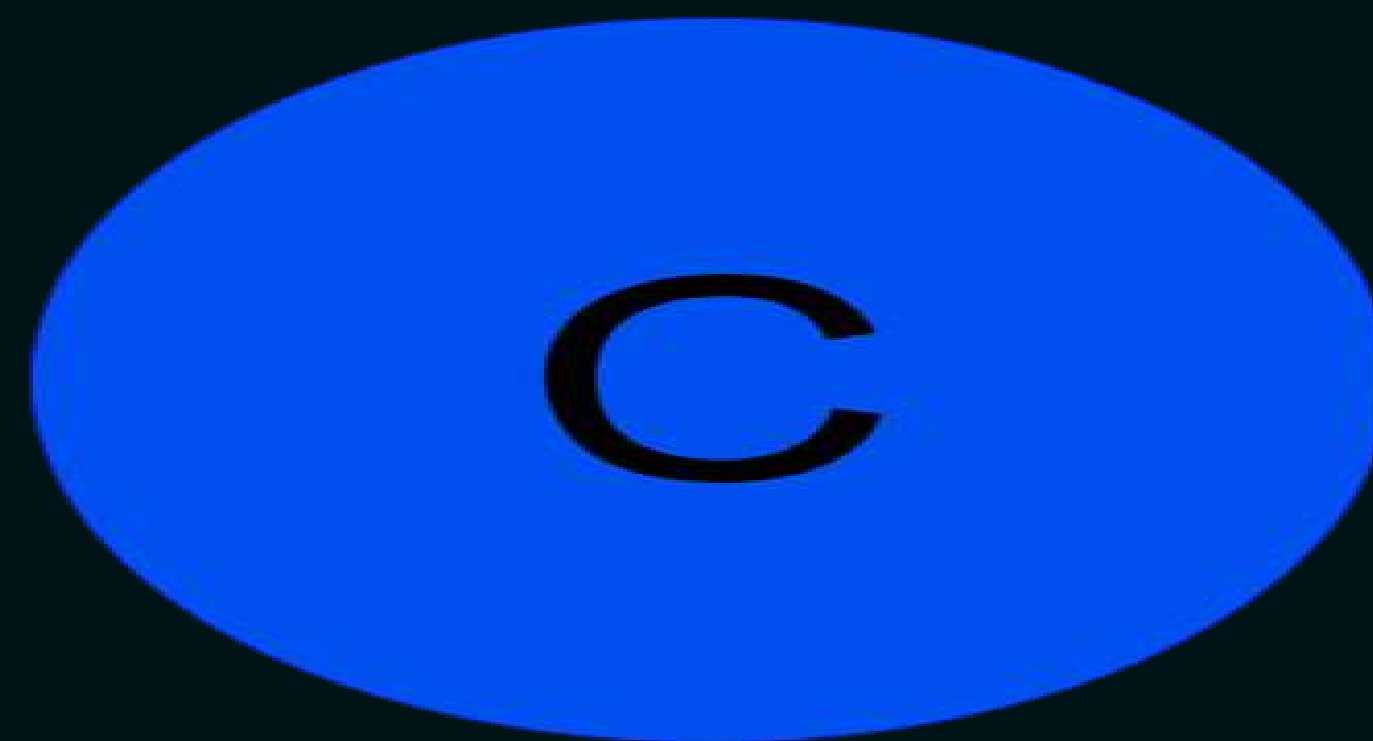


Node: C

*Term: 1*

**LogSize: 0**

VoteCount: 1

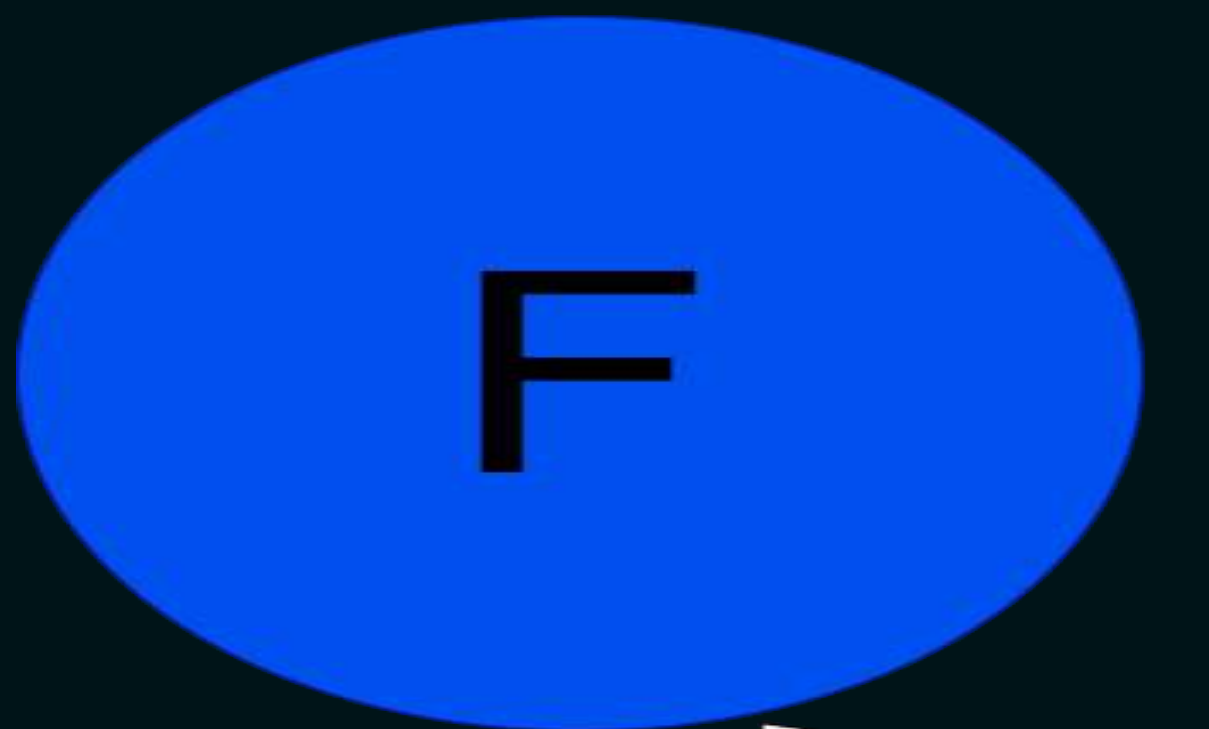


RequestVote

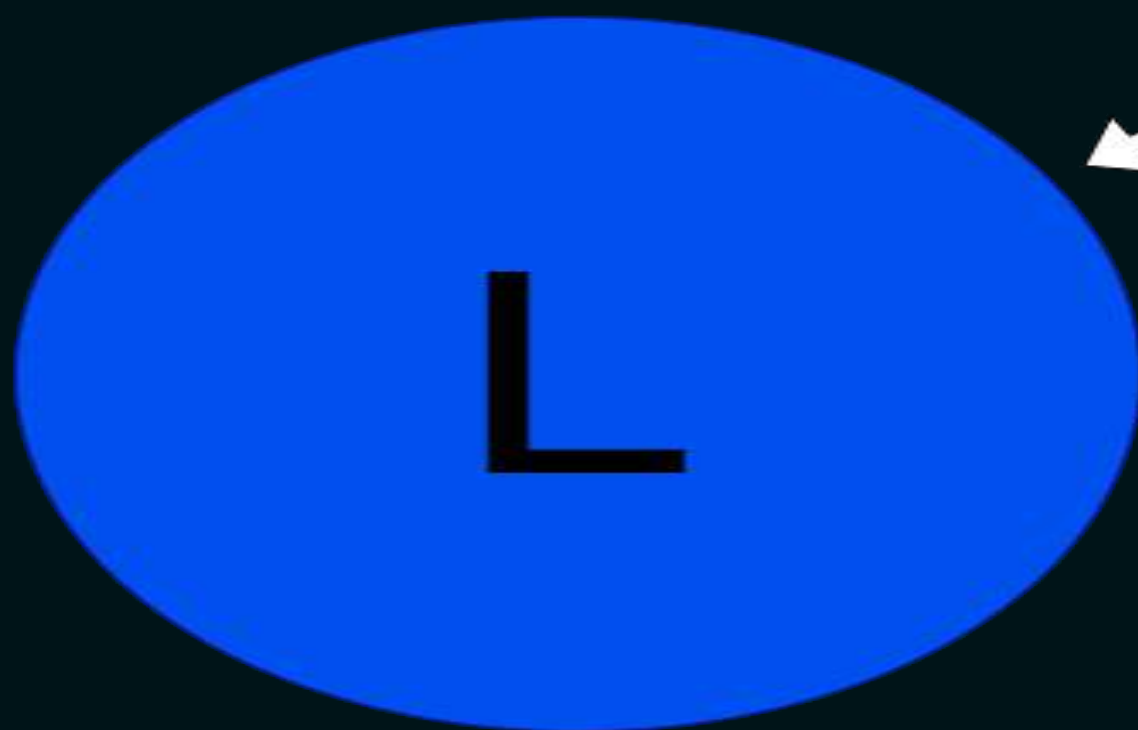
RequestVote

# Работа системы

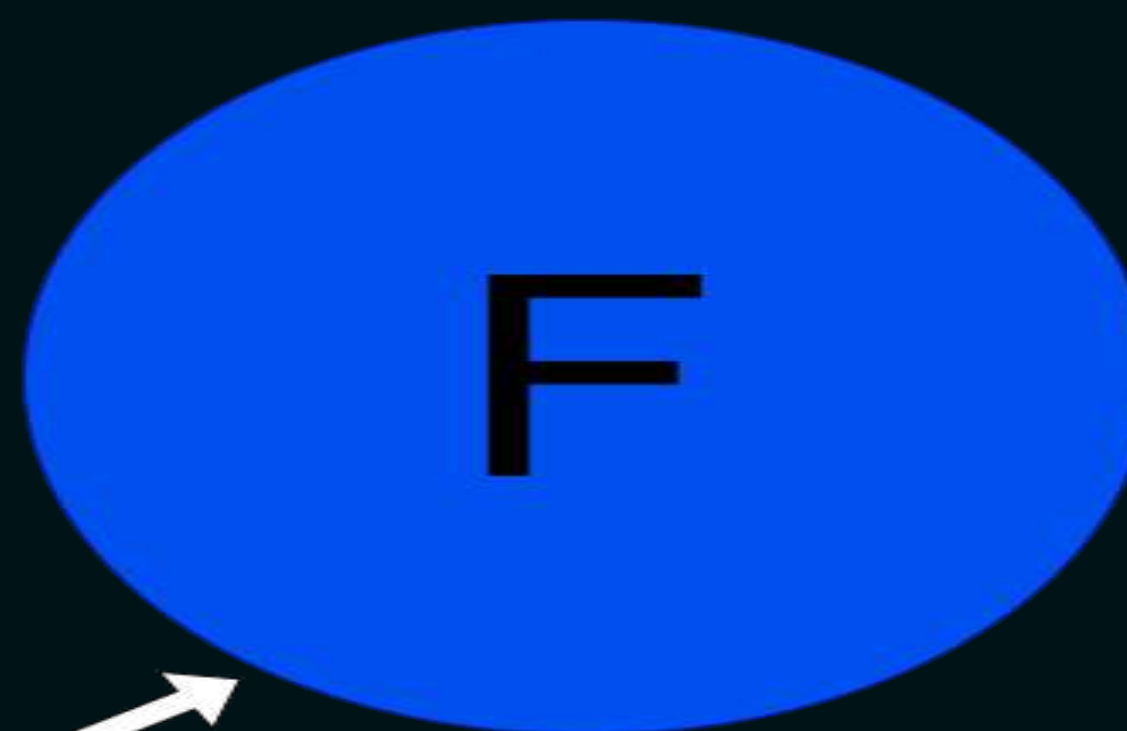
Node: A  
Term: 1  
LogSize: 1  
ElectionTimeout: 5  
Leader: C



Node: C  
Term: 1  
LogSize: 1  
HBeatTimeout: 4.5



Node: B  
Term: 1  
LogSize: 1  
ElectionTimeout: 4  
Leader: C



AppendEntries



AppendEntries

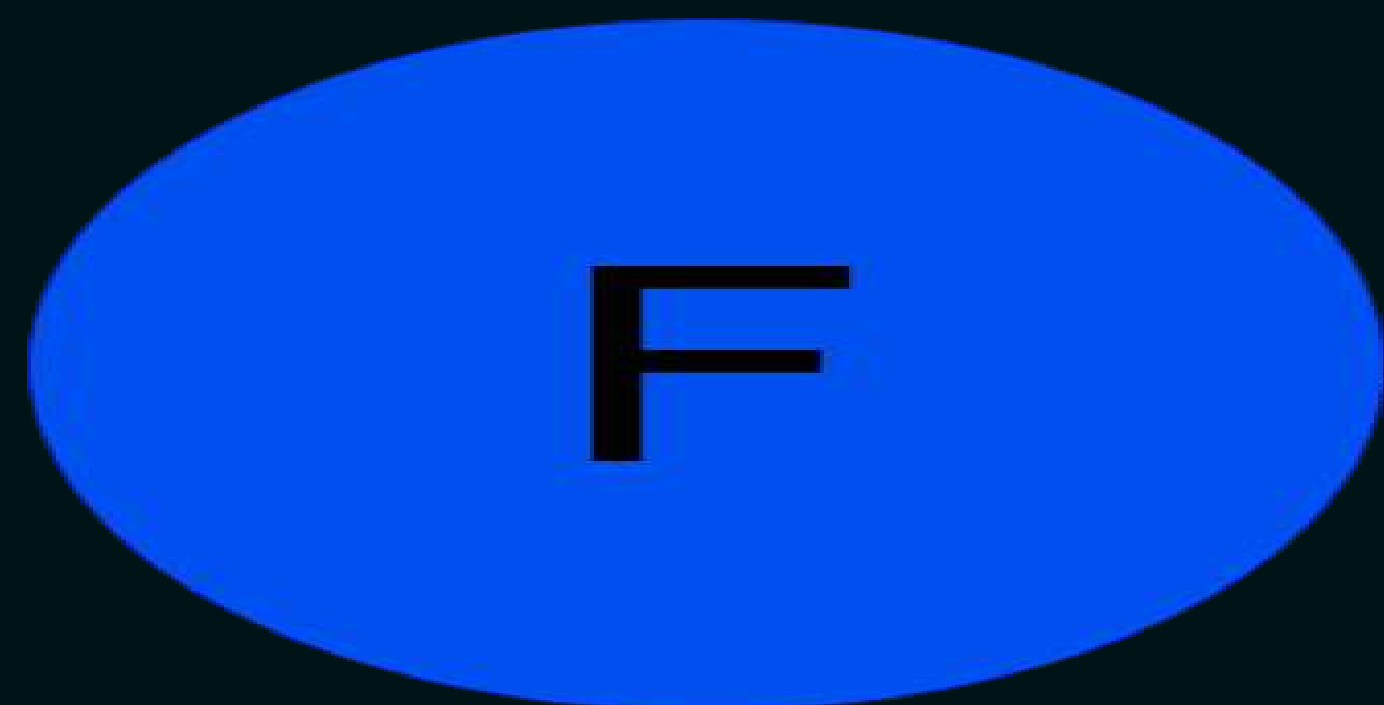


SetData

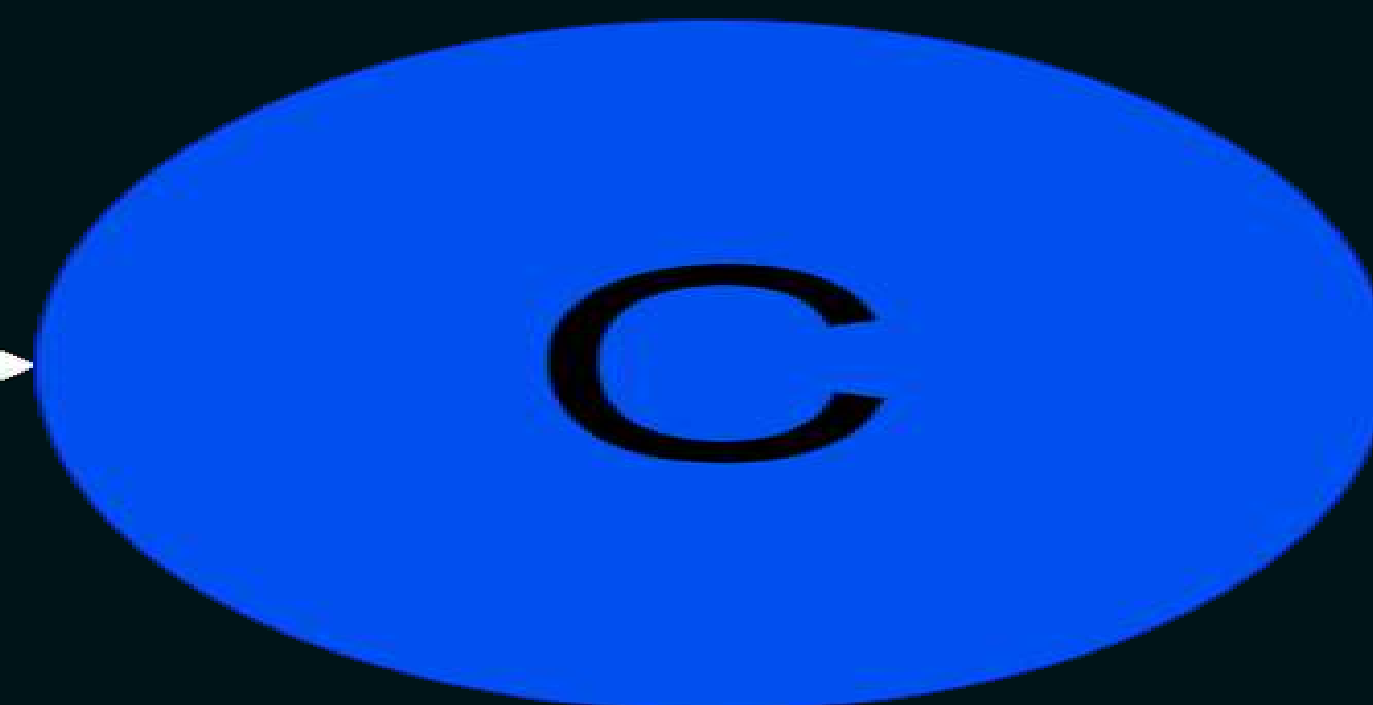


# Переизбрание лидера

Node: A  
*Term: 1*  
LogSize: 1  
VotedFor: B



Node: B  
*Term: 2*  
LogSize: 1  
VoteCount: 1



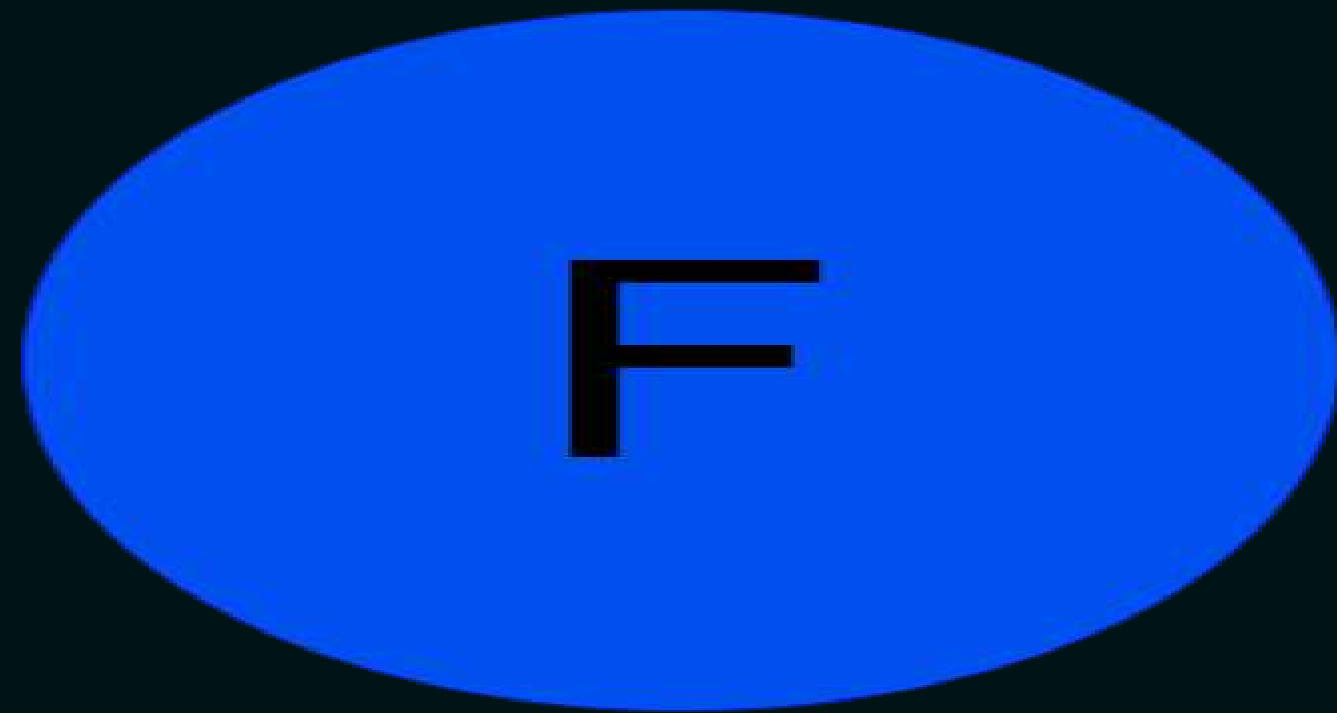
RequestVote

On fire or  
disconnected

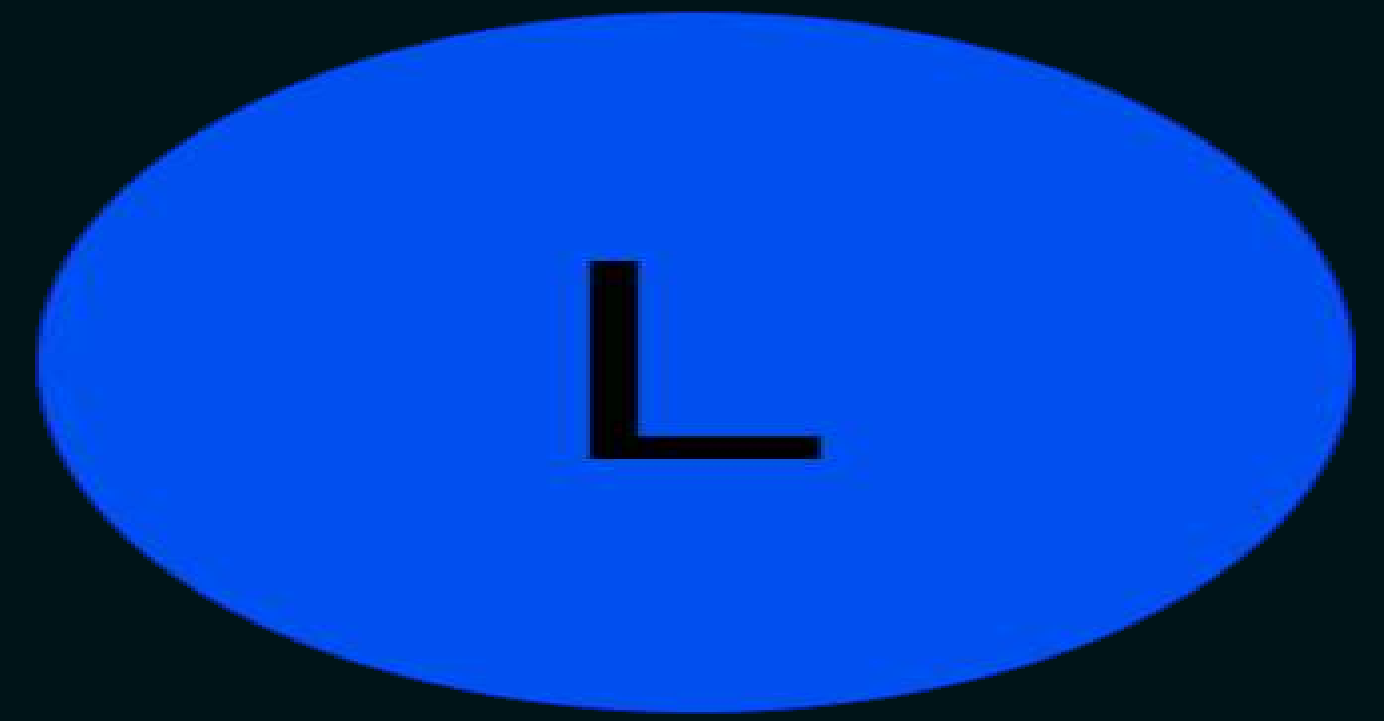


# RAFT

Node: A  
Term: 2  
LogSize: 1  
ElectionTimeout: 5



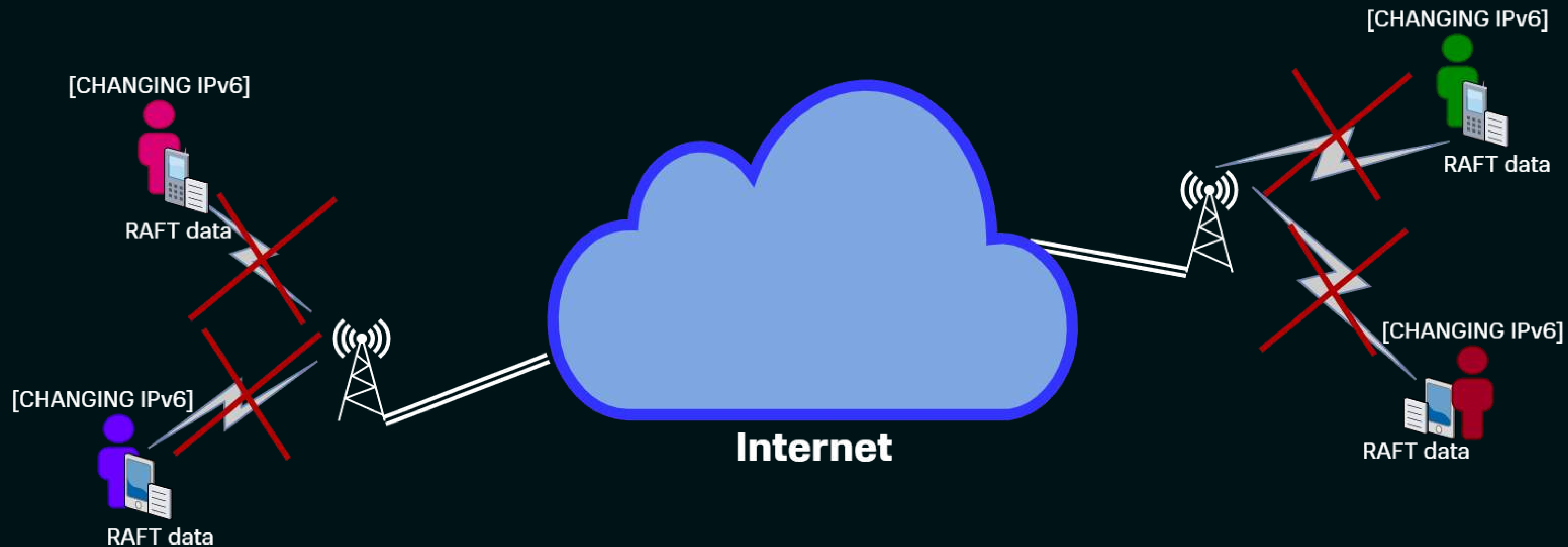
Node: B  
Term: 2  
LogSize: 1  
HBeatTimeout: 4



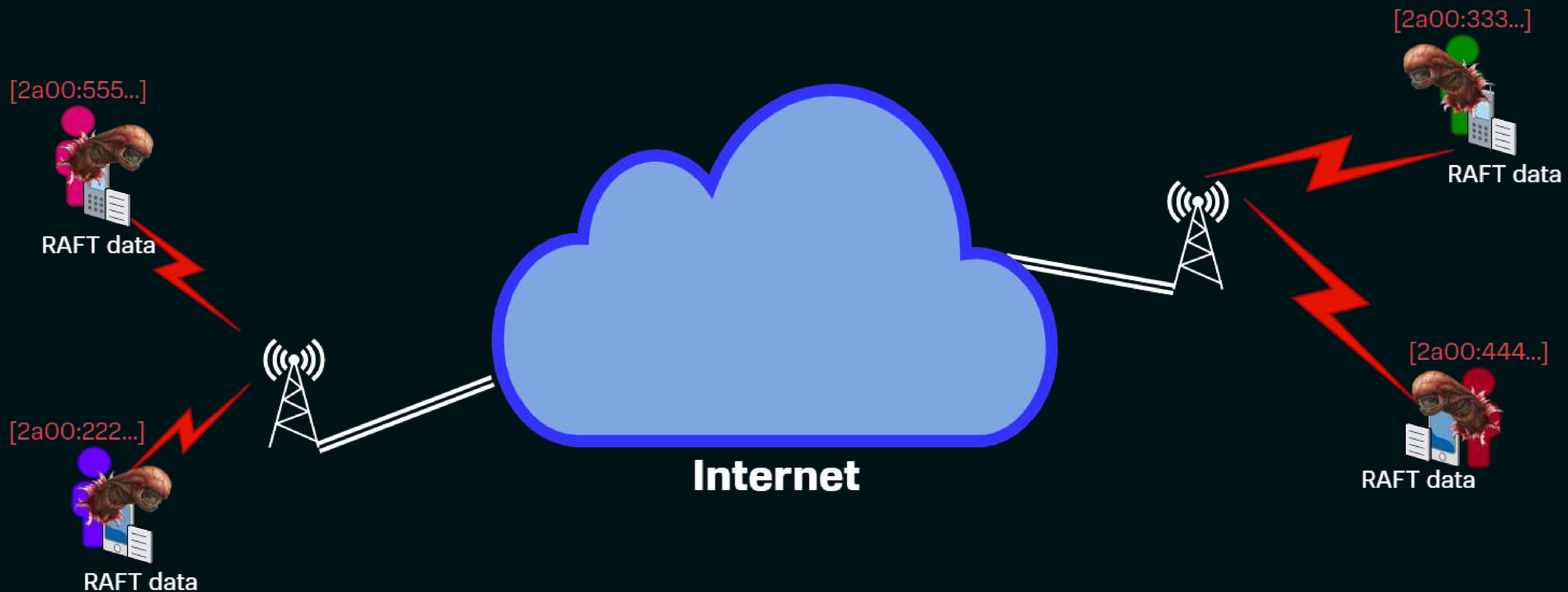
# Реализация v1



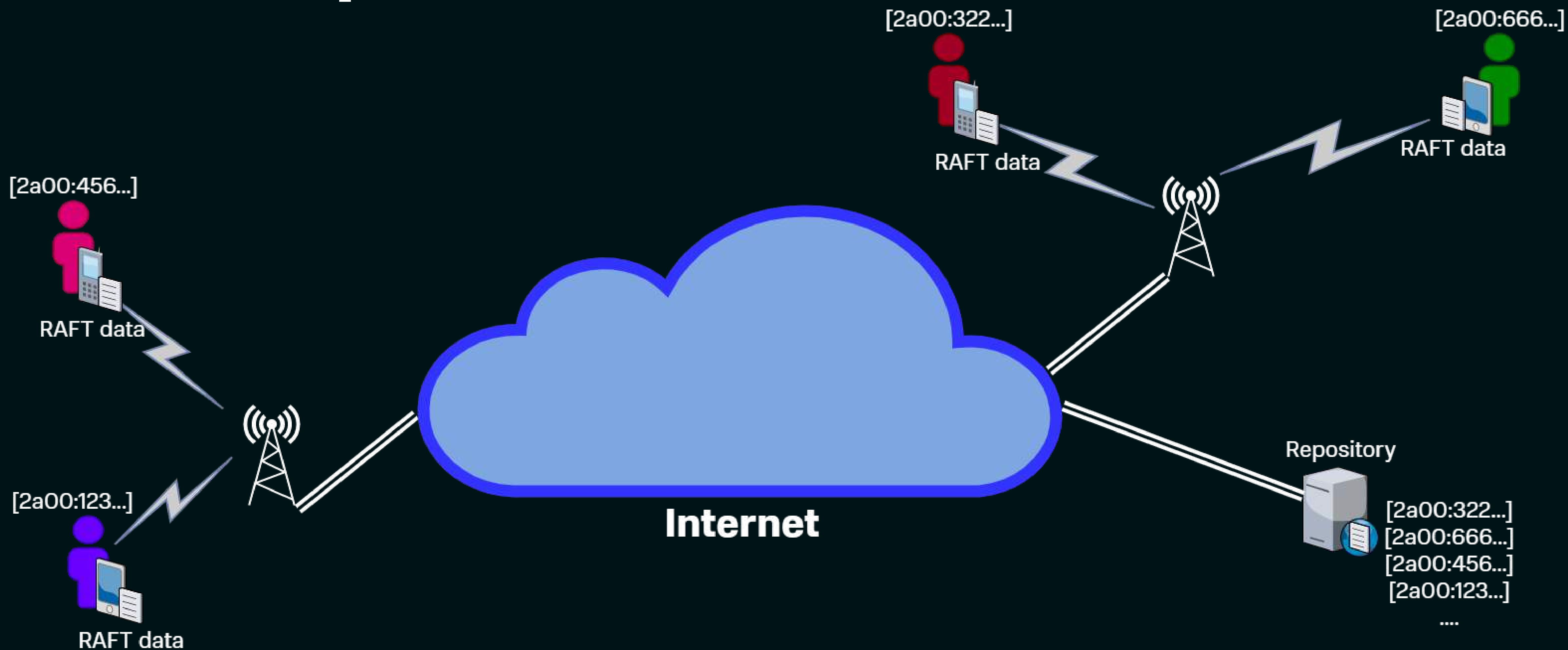
# Реализация v1



# Реализация v1



# Реализация v2





# Результаты

- Процедура оценки производилась в течение суток. Более 20 участников произвели оценку через приложение
- На основе оценок выбран лучший проект
- Проведен анализ результатов, подведение статистики, и в рассылке предоставлена детальная информация об оценках

# Результаты

- Максимальное количество участников «онлайн»: 12
- Количество полученных оценок: 147
- Максимальный интервал репликации сообщений: 6 секунд
- Интервал выбора нового лидера: 8 секунд
- Количество записей журнала в сообщении репликации: 10

# Дальнейшее развитие

1. Использование более легкого протокола UDP, динамическое формирование таймаутов и размера сообщений в зависимости от общего состояния кластера
2. Разработка мультиплатформенного фреймворка для обеспечения переносимости функционала между платформами.



# Спасибо! Вопросы?

Лаборатория архитектурной экспертизы

07.02.2020



# Победители

- Проект: Разработка технического решения по видеоаналитике

- Команда проекта:

- ❖ Алекс Шершебнев
- ❖ Алексей Семенюк
- ❖ Анна Кащеева
- ❖ Арсений Павленко
- ❖ Григорий Погорелов
- ❖ Дарья Кувшинова
- ❖ Дмитрий Зиганшин

- ❖ Екатерина Ванская
- ❖ Елизавета Сукач
- ❖ Иван Копылов
- ❖ Кирилл Овчинников
- ❖ Кирилл Подберезин
- ❖ Наталья Григорьева
- ❖ Олег Саитов
- ❖ Павел Елисеев