

Три ветки записи кратко

КРИТЕРИЙ	КЛИЕНТ	ПО УЧАСТНИКАМ	В МИКСЕ
Где идёт захват	Устройство	Сервер (SFU)	Сервер (MCU)
Затраты CPU сервера	Нет	Низкие (декод)	Высокие (декод+микс)
Форма хранения	1 на устройство	1 на говорящего	1 сведённый
Знает кто говорил	Только себя	Да, встроено	Нет, угадывать
Переживает закрытие	Нет	Да	Да
Играет в любом плеере	Да	Не напрямую	Да
Лучшее для транскрипции	Слабо	Сильнее всего	Слабее всего
Пересвести / редакт потом	Нет	Да	Нет

Путь аудио в ASR (звук → признаки, односторонне)

Ресемплинг	Opus идёт на 48 kHz; ASR опускает до 16 kHz — речь живёт ниже 8 kHz.
Кадр	Режется на окна 25 ms, сдвигаемые на 10 ms (они перекрываются).
Лог-мел	Каждый кадр → 80 мел-полос, громкость лог-сжата. Whisper берёт 80 полос.
Односторонне	30 с → сетка 80×3000 (240 000 чисел). Нельзя вернуть в аудио.

Инструменты, которые встретятся

MediaRecorder	Рекордер на клиенте (W3C). timeslice → куски dataavailable; WebM/Opus.
LiveKit Egress	Запись на сервере: скрытый участник; room-composite = headless Chrome+GStreamer.
Whisper	Вход ASR — лог-мел, не ваша запись. 16 kHz, 80 полос, куски по 30 с.
Диаризация	ruanote/NeMo → RTTM (начало, длит., говорящий); WhisperX сливает с тайм-кодами слов.
Живые субтитры	Частичные за 200–500 ms; финальные после паузы. Компромисс скорость vs устойчивость.

До того, как аудио покинет звонок

- ☐ Живой звонок и его запись — разные задачи; настройте под одну — пострадает другая.
- ☐ Микширование — дверь в одну сторону: теряете «кто что сказал» и возможность редакта одного.
- ☐ Запись по участникам — сильнейший вход для транскрипции и чистейший для соответствия.
- ☐ Тишина DTX + наивный микшер = записи, что пропадают. Держите уровень / добавьте комфортный шум.
- ☐ ASR выбрасывает звук: 16 kHz, 80 лог-мел-полос, односторонне — транскрипт это не запись.
- ☐ По участникам = нет проблемы диаризации — каждый файл уже один говорящий.
- ☐ Кормите субтитратор чистой декодированной дорожкой, а не прореженными DTX живыми пакетами.
- ☐ Хранимый голос — регулируемые данные: согласие, удаление по GDPR, шифрование + 6 лет по HIPAA.