

Три режима: 50 / 500 / 5 000 с одного взгляда

СВОЙСТВО	≤ 50	~500	5 000+
Архитектура	SFU, всё	SFU, активные	MCU-микс; каскад
Декодирует аудио?	Никогда	Никогда	Только активных
Слушатель получает	Каждый голос	Активных	Один микс
Главная стоимость	Декод клиента	Веер сервера	CPU микса
Потоков скачивает	1 на говорящего	1 на говорящего	Ровно 1 (микс)
Задержка	Минимум	Низкая	Выше (микс)
Контроль голоса	Полный	Полный	Теряется
Запись файлом	Отдельный микс	Отдельный микс	Микс уже есть

Почему взрывается: $K \times (K-1)$ доставок голоса

5 человек	$5 \times 4 = 20$ доставок голоса.
50 человек	$50 \times 49 = 2\,450$ — примерно в 122 раза больше, чем на 5.
500 человек	$500 \times 499 = 249\,500$ доставок, если пересылать всё.
5 000 человек	$5\,000 \times 4\,999 = 24\,995\,000$ — квадрат: удвоение $\approx \times 4$ работы.

Правила, делающие масштаб возможным

Аудио $\neq$ видео	Last-N пересылает только верхние видео; ВСЁ аудио идёт всем в малых/средних комнатах.
DTX	Молчащий Opus падает с ~24-32 kbit/s до ~1 kbit/s — стоимость следует за говорящими.
RFC 6464	Громкость в каждом пакете (0-127 -dBov); сервер читает её без декодирования.
Микс N-1	Каждый говорящий слышит всех КРОМЕ себя; смешивайте только 3-6 громких каналов.
Каскад / гибрид	За пределом сервера: ретрансляция по сетке или стрим микса по LL-HLS пассивным.

Запомнить

- ☐ ≤ 50 человек: SFU пересылает всё аудио всем; DTX держит молчащих почти бесплатно.
- ☐ 50-500 человек: пересылайте только активных говорящих по уровням громкости RFC 6464.
- ☐ 5 000+ человек: смешивайте 3-6 самых громких в ОДИН поток; никогда не складывайте все.
- ☐ За пределом одного сервера: каскад по сетке или гибрид (микс + LL-HLS пассивным).
- ☐ Проектируйте под ПИК одновременных говорящих, а не под среднее — спайк «есть вопросы?».
- ☐ Лимит одновременных говорящих 3-6; подкрепите его UI «поднять руку» или сценой.
- ☐ UI говорящего мигает на паузах? Подозревайте разрывы Opus DTX, а не алгоритм.
- ☐ Если почти никому не нужно отвечать — это трансляция; раздайте пассивных стримингом.