AsTree: An Audio Subscription Architecture Enabling Massive-Scale Multi-Party Conferencing

(Operational System Track)

Tong Meng   Wenfeng Li   Chao Yuan   Changqing Yan   Le Zhang

Presenter: Tong Meng

usenix THE ADVANCED COMPUTING SYSTEMS ASSOCIATION
nsdi '25
ByteDance 字节跳动

To Start

Say multi-party conferencing is important ...

(Don't take more than ~~10~~ 15 seconds)

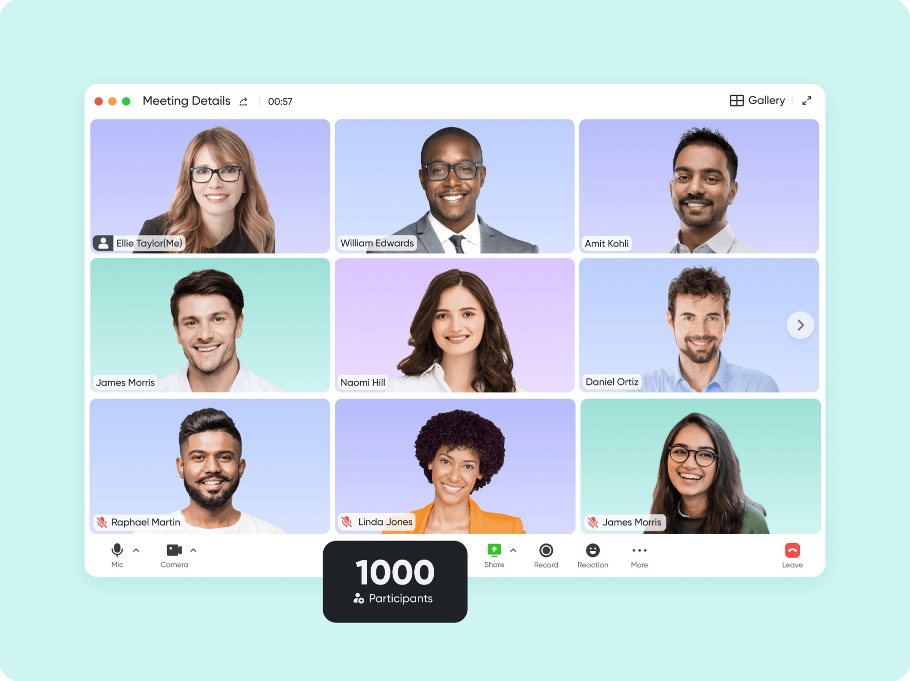
We are VeRTC

Volcano Engine RTC

…… ……

抖音 PICO Lark

- 80+ billion monthly call minutes
- 3000+ nodes worldwide
- < 400ms e2e latency in over 99.5% cases



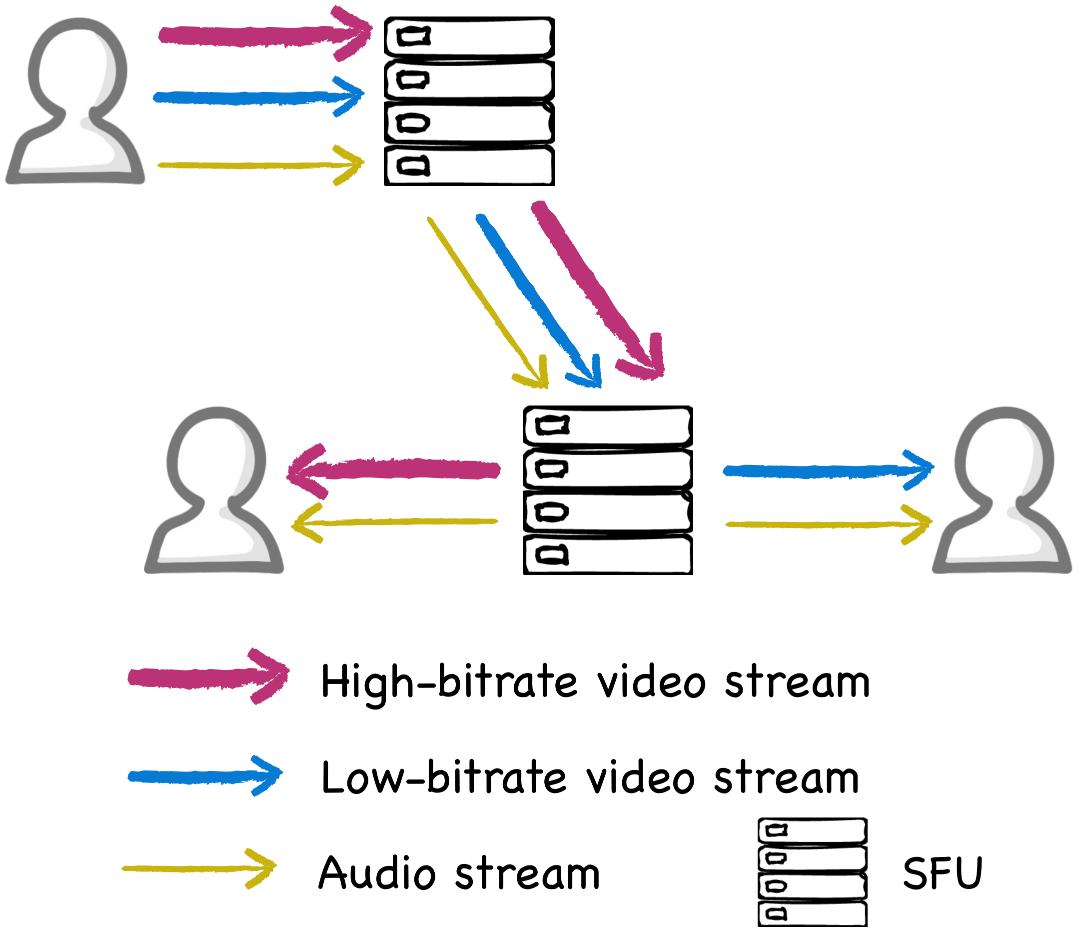
- 1000 active participants per room, and more for large-scale webinars

Roadmap

- Background
  - Multi-Party Conference Based on Selective Forwarding Units (SFUs)
- Motivation
  - Scalability Issues of Audio Conferencing
- Design
  - AsTree Architecture
  - AsTree Cascading Topology
  - Audio Selection
- Performance

SFU-Based Multi-Party Conferencing

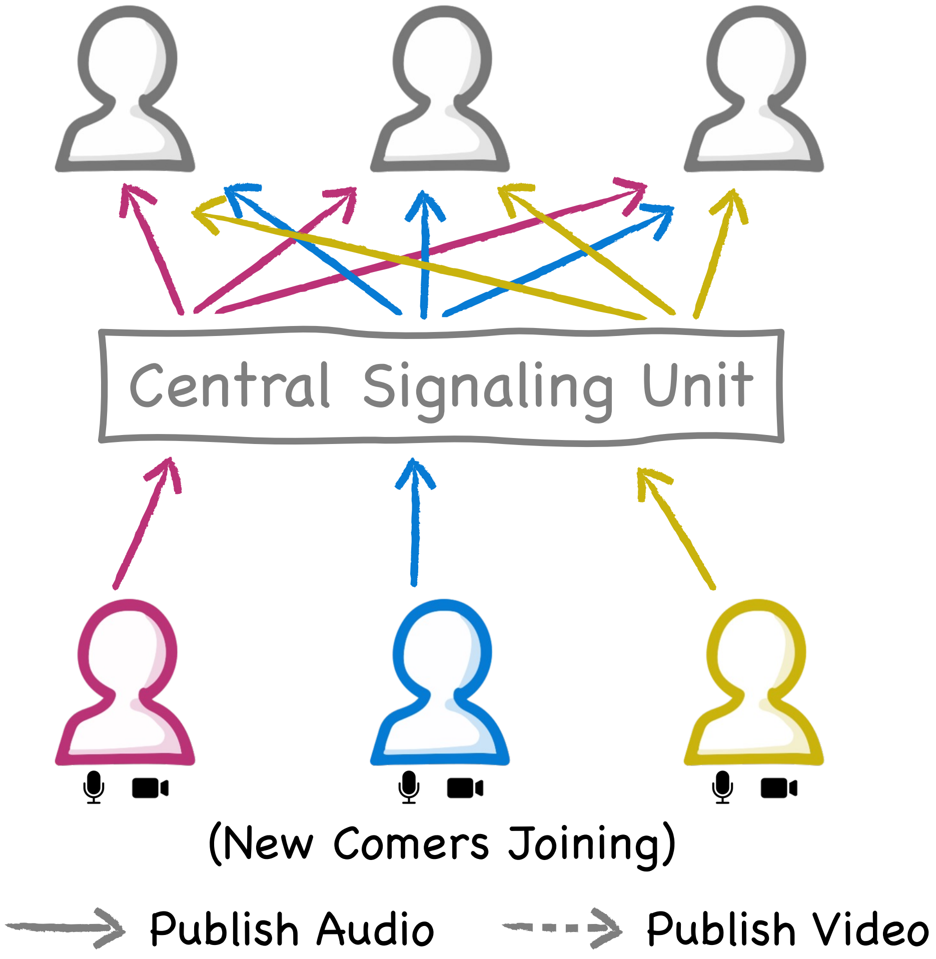
- SFU selectively forwards media streams between participants without decoding them (*e.g.*, Simulcast for video)
  - Server-side cost effectiveness
- Multiple SFUs cascaded to connect distributed participants
  - Low client-side first-hop latency



Scalability Issues of Audio Subscription

Signaling Storm

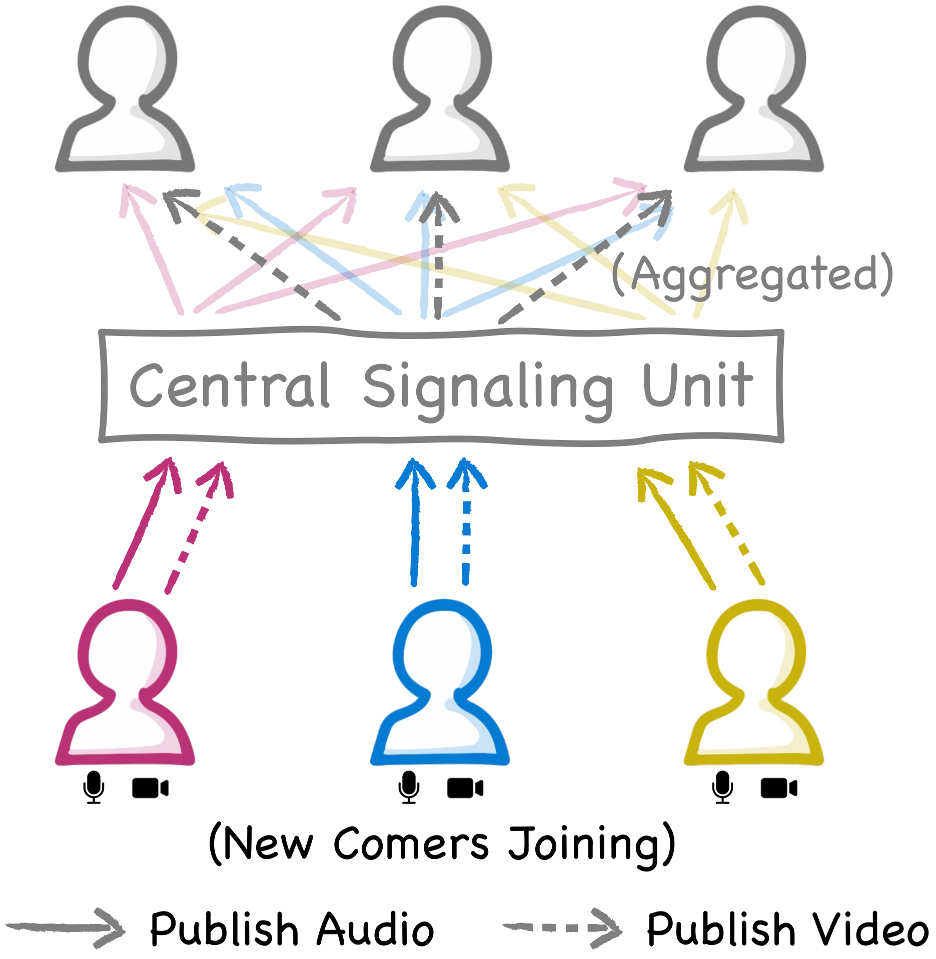
- $O(N)$ signaling messages per participant
- Combined to $O(N^2)$ signaling messages per room



Scalability Issues of Audio Subscription

Signaling Storm

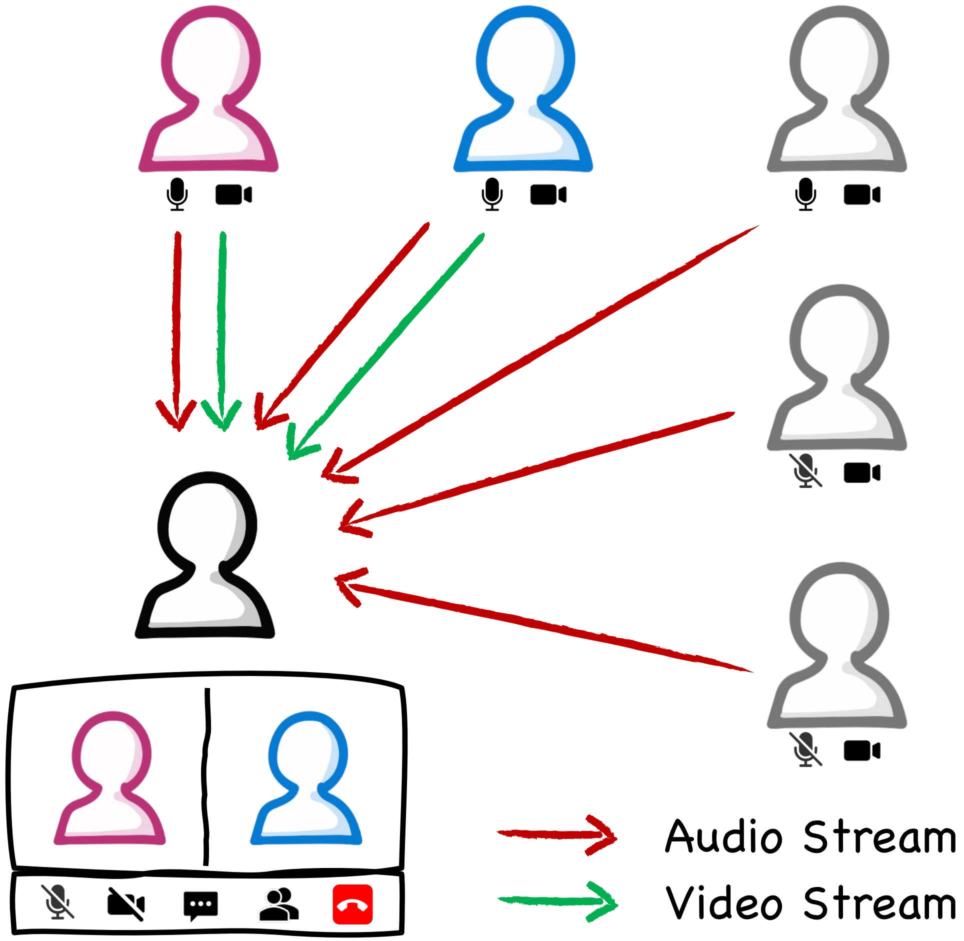
- $O(N)$ signaling messages per participant
- Combined to $O(N^2)$ signaling messages per room
- Cannot be aggregated as in video (due to tighter latency requirements)



Scalability Issues of Audio Subscription

Indefinite Overheads

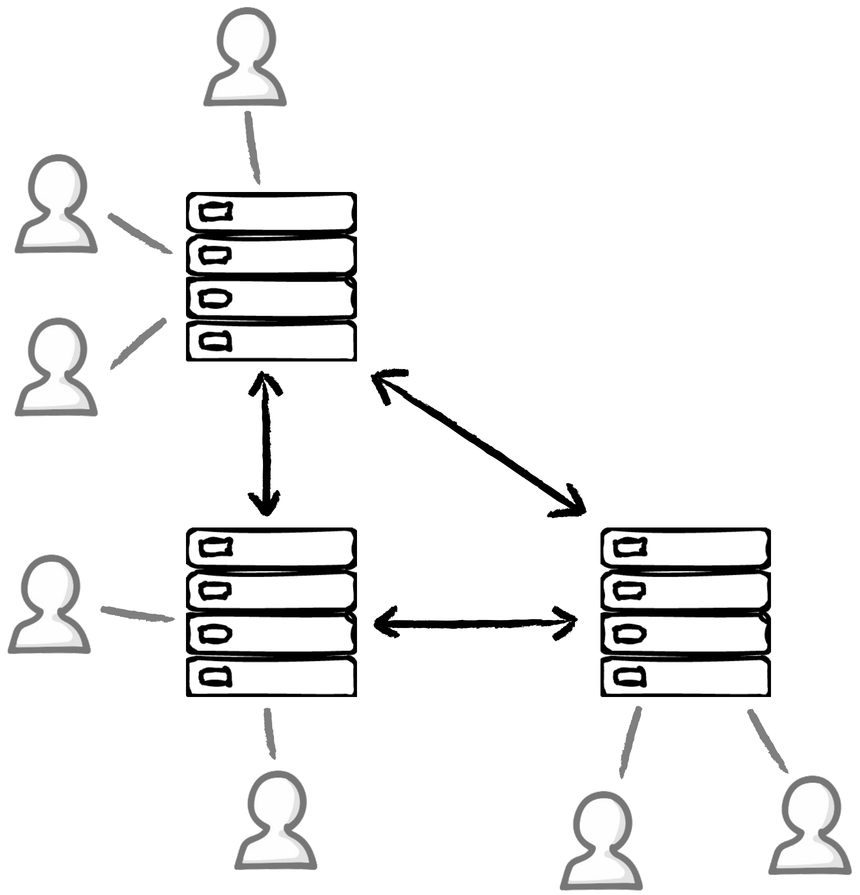
- "Subscribe to all" for audio, while video limited by user interface
- Linear memory and bandwidth overheads on client-side
- Combined to $O(N^2)$ server-side egress bandwidth per room



Scalability Issues of Audio Subscription

Mesh Cascading

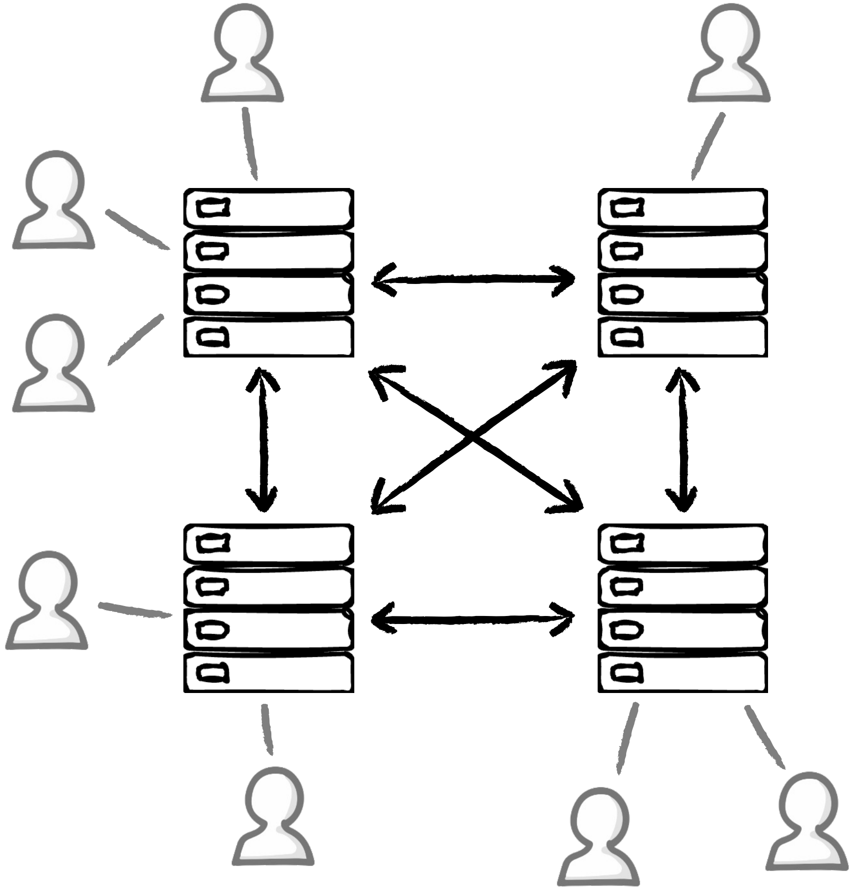
- Each room leads to an overlay mesh between involved SFU media servers, regardless of video subscription relationship



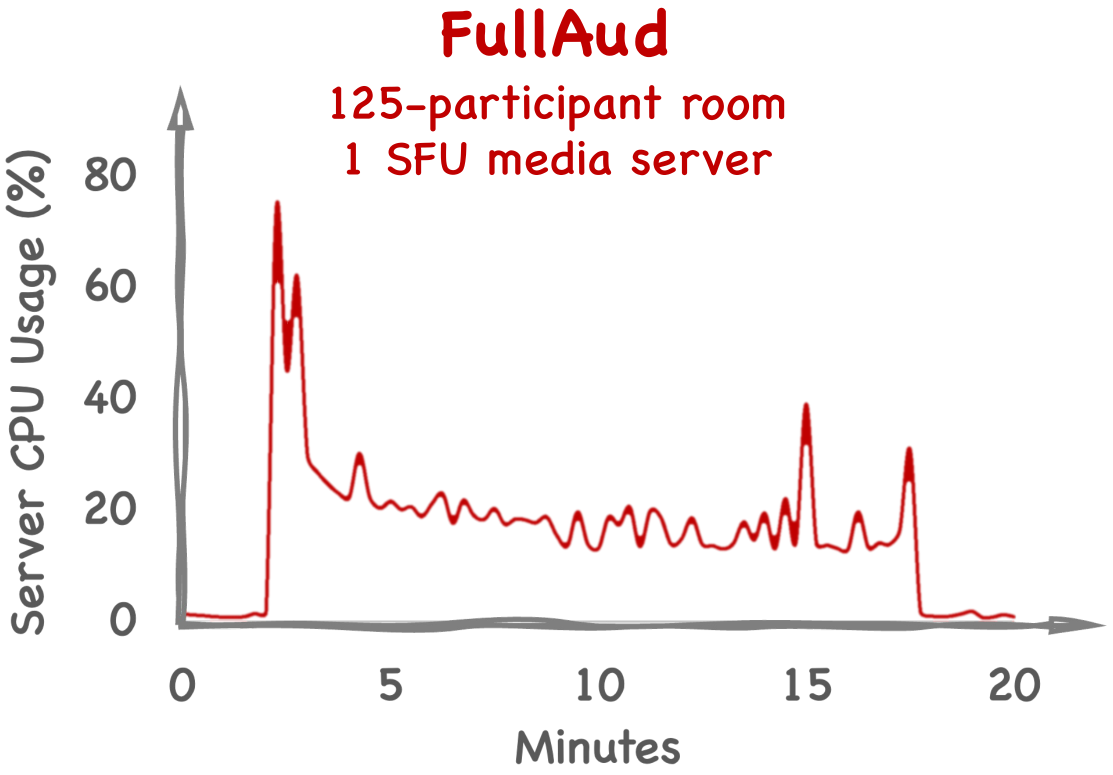
Scalability Issues of Audio Subscription

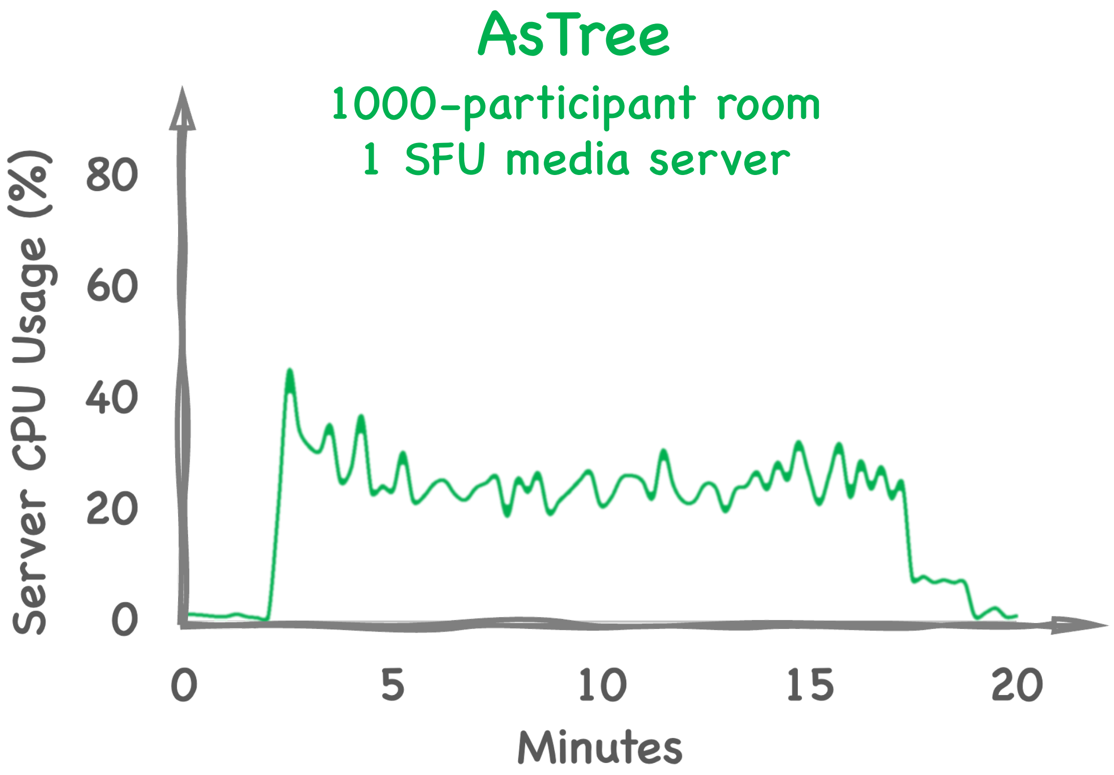
Mesh Cascading

- Each room leads to an overlay mesh between involved SFU media servers, regardless of video subscription relationship
- High complexity to maintain and update

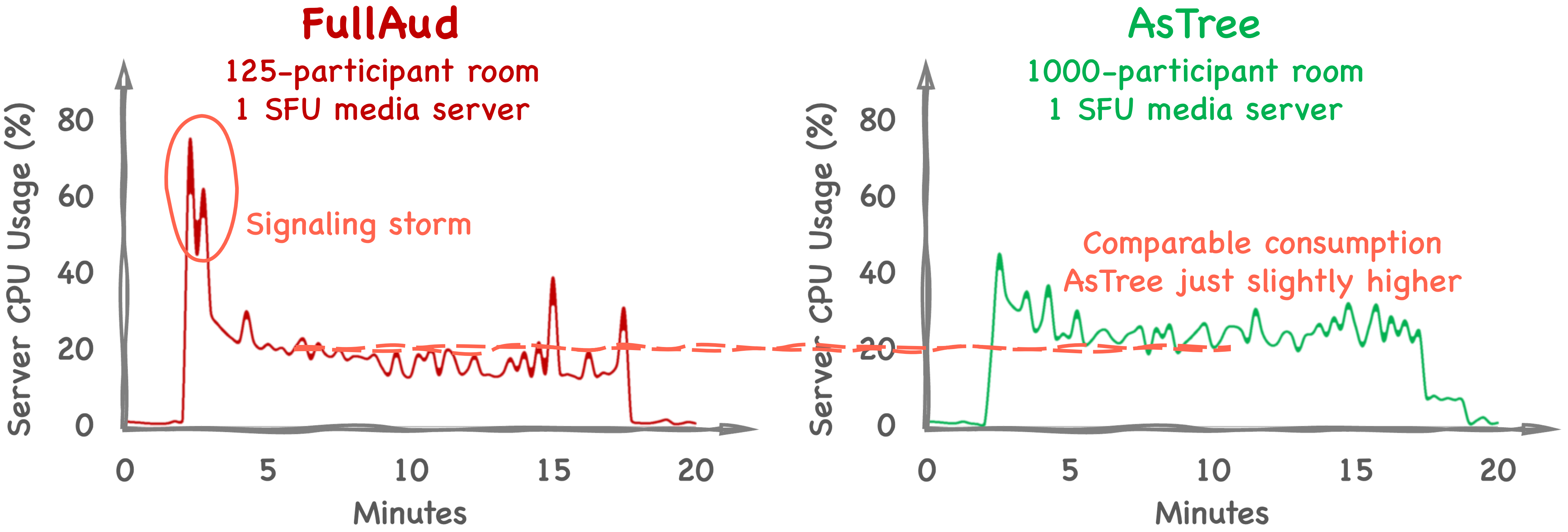


Scalability Issues of Audio Subscription

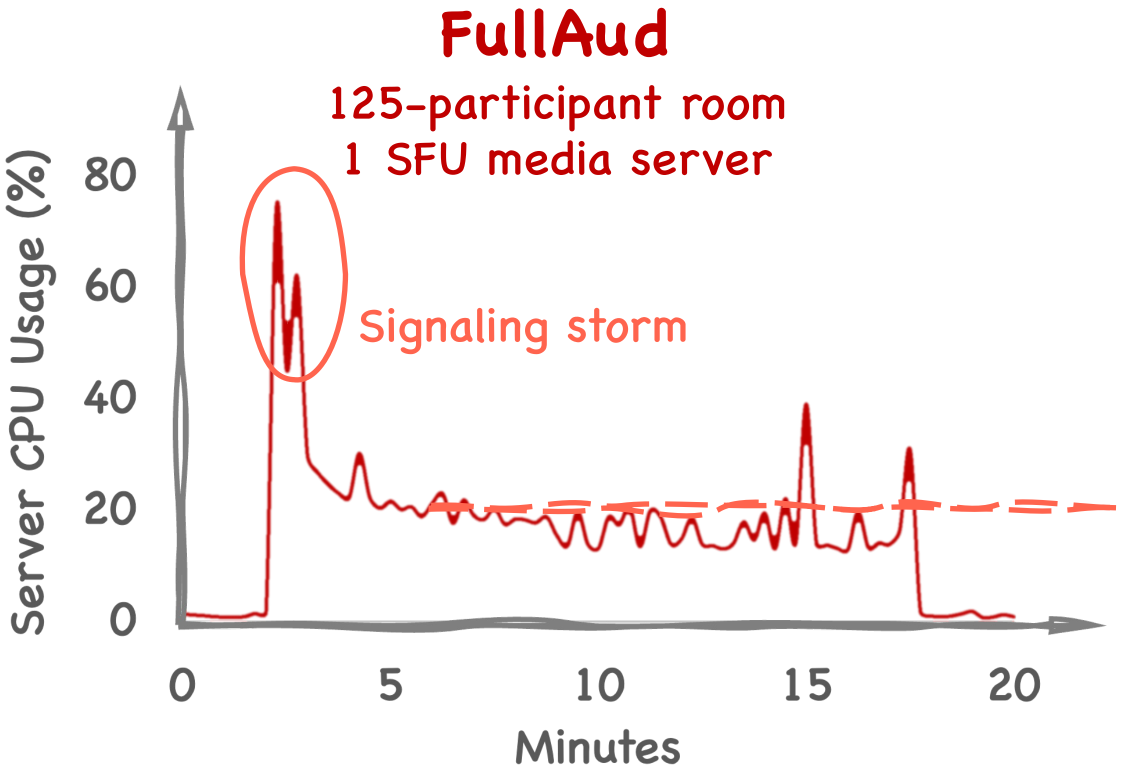


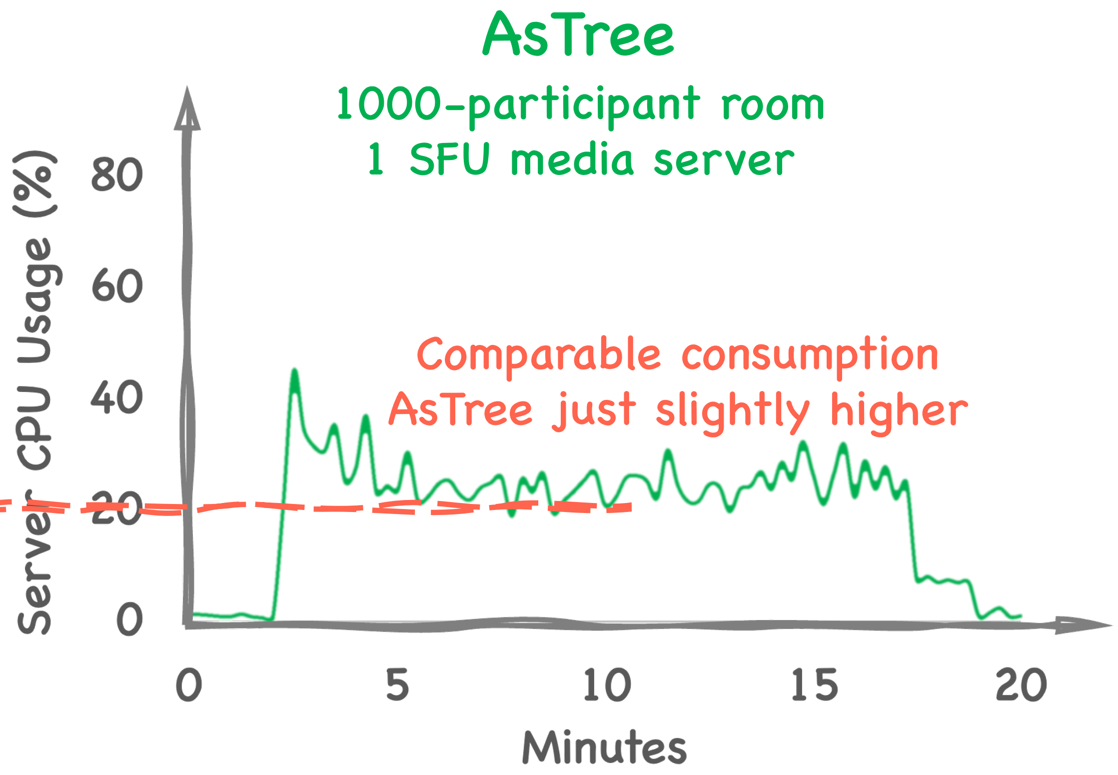


Scalability Issues of Audio Subscription



Scalability Issues of Audio Subscription



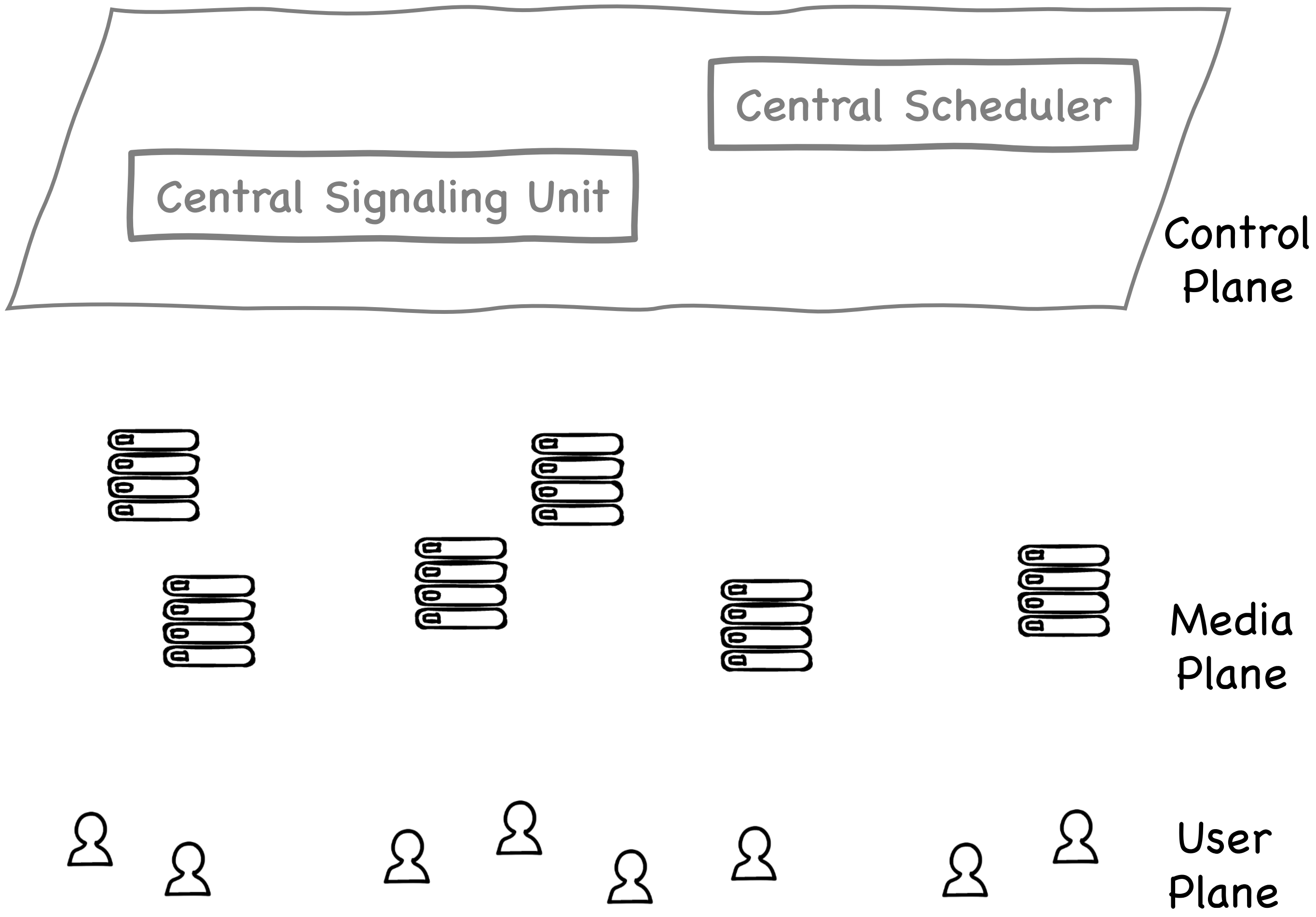


Restricted server capacity due to signaling storm and massive audio subscription

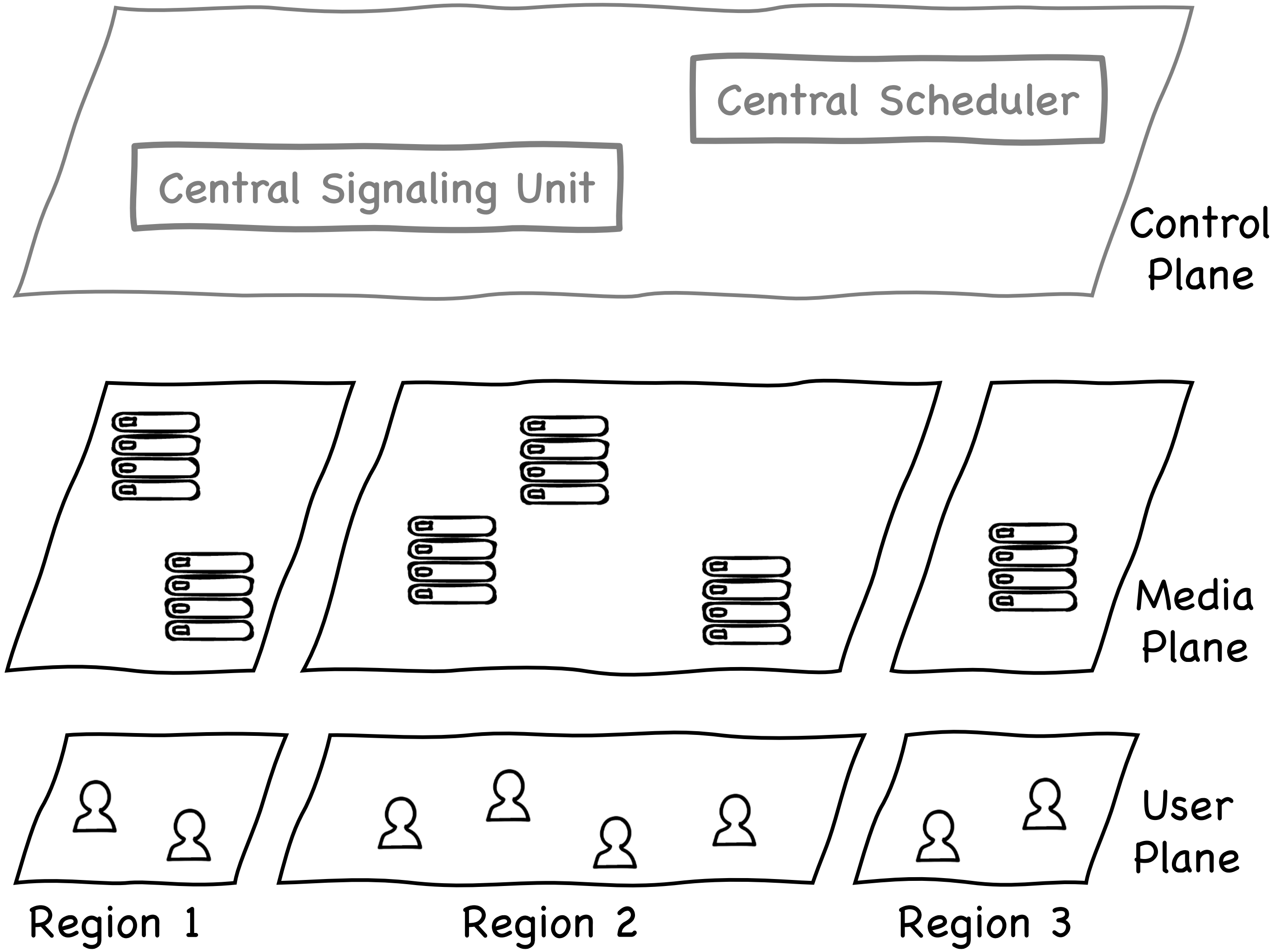
Design Insights

Critical information comes from only a small number of audio streams, by a few loudest speakers

AsTree Architecture

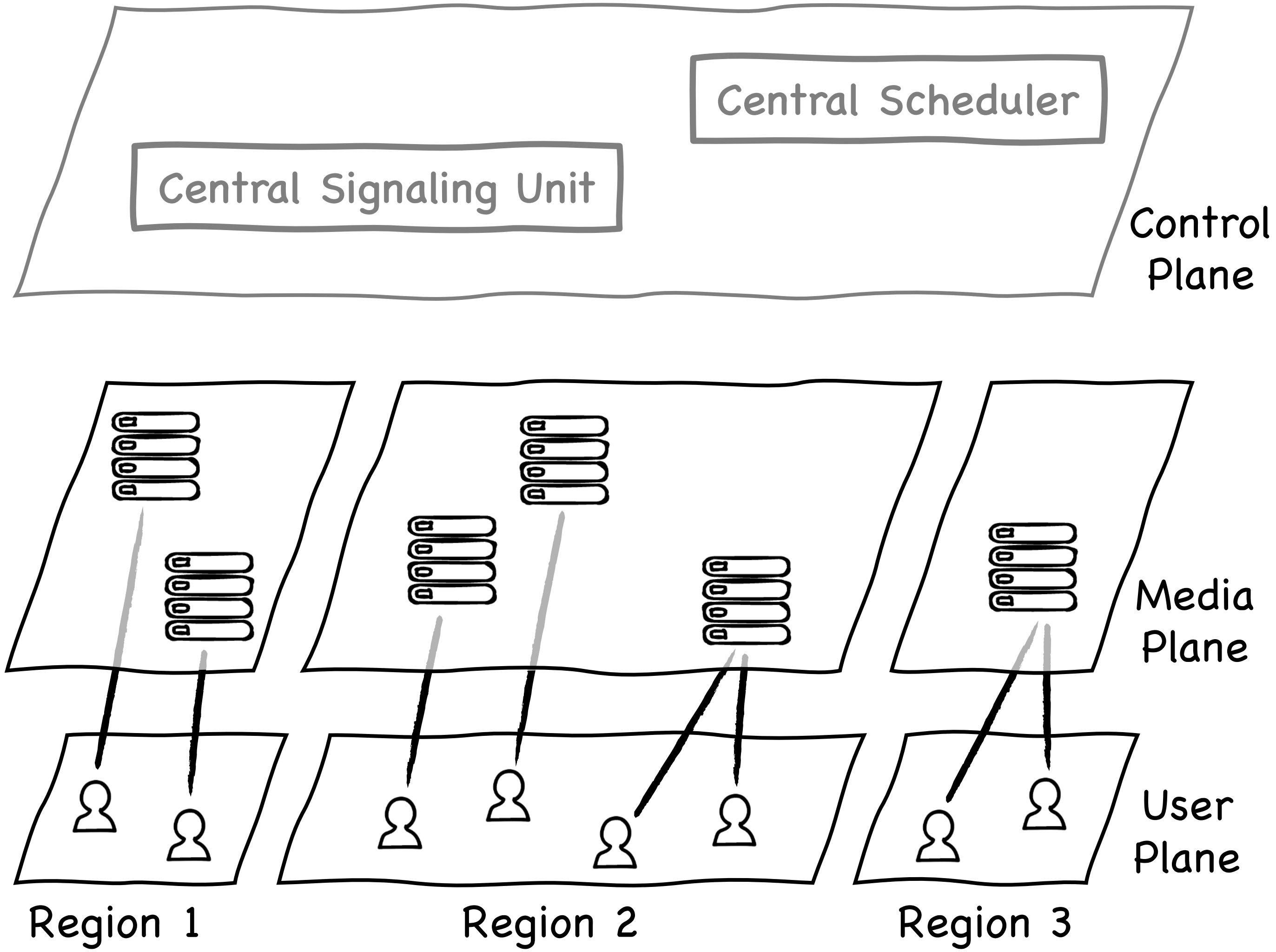


AsTree Architecture



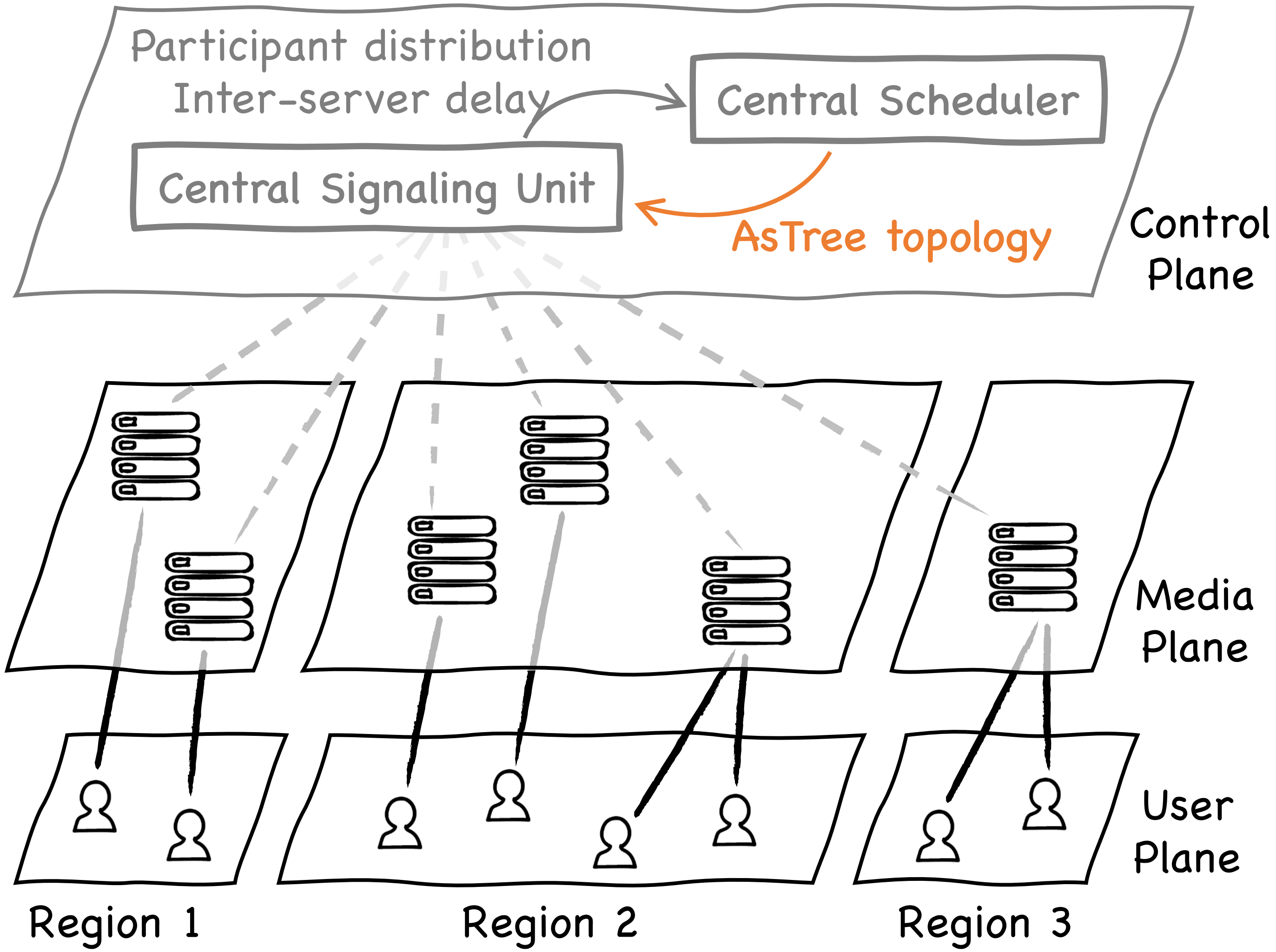
- Geographical regions in user and media plane

AsTree Architecture



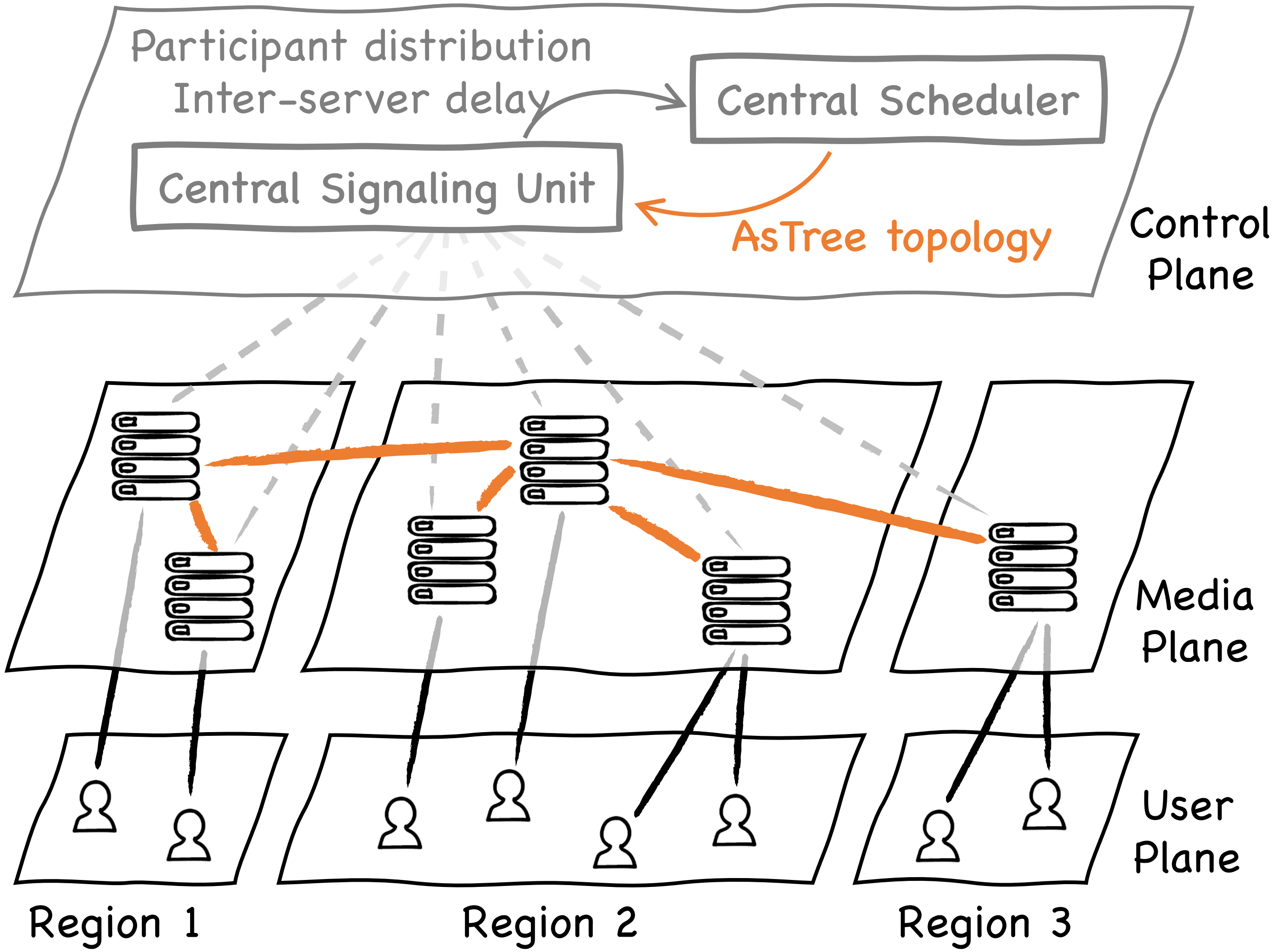
- Geographical regions in user and media plane
- Participants connected to nearby SFU media servers

AsTree Architecture



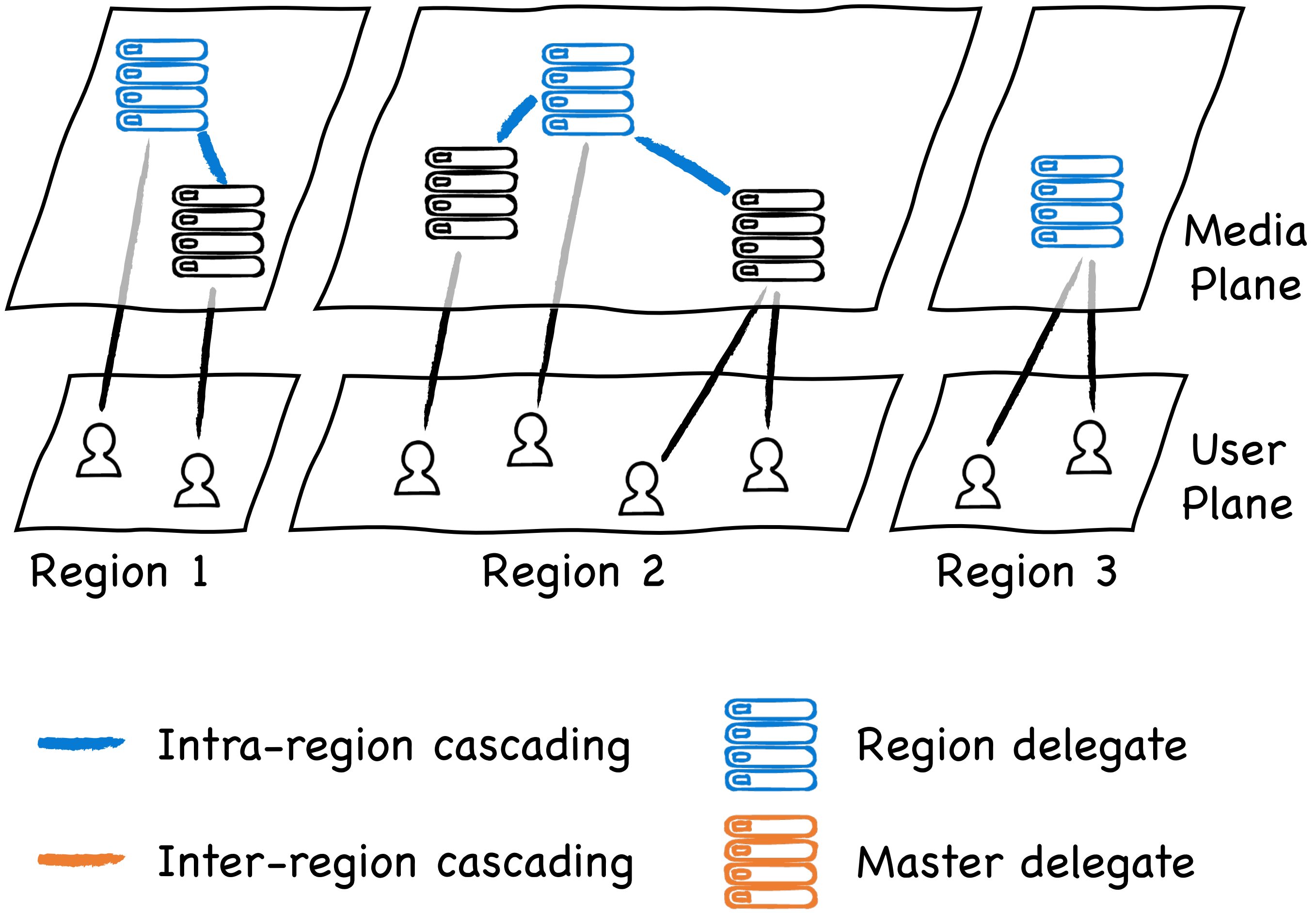
- Geographical regions in user and media plane
- Participants connected to nearby SFU media servers
- AsTree topology calculated at control plane

AsTree Architecture



- Geographical regions in user and media plane
- Participants connected to nearby SFU media servers
- AsTree topology calculated at control plane
- SFUs cascaded into a tree topology (only for audio subscription)

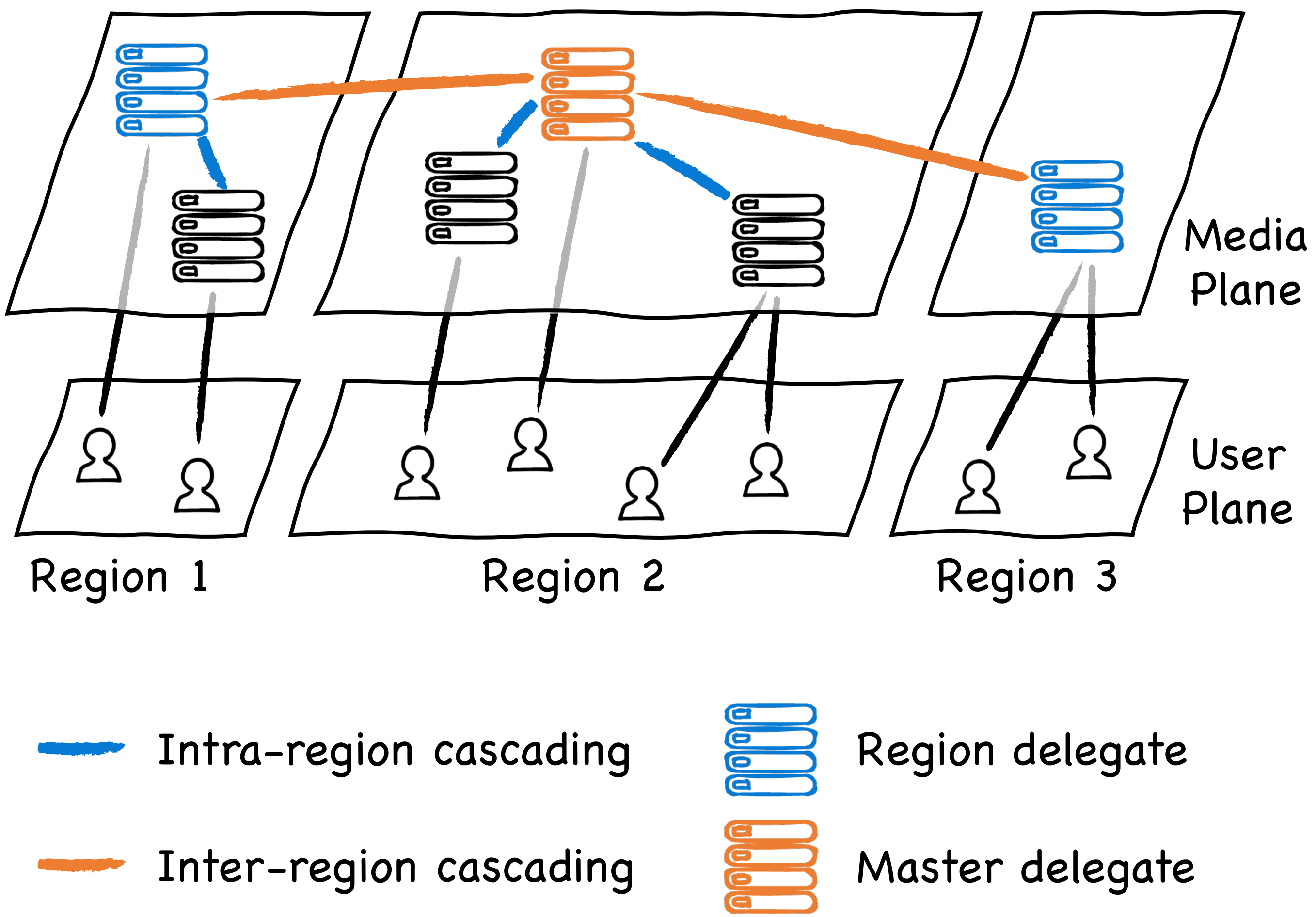
AsTree Cascading Tree: Two Hierarchies



Intra-Region

- All SFUs cascaded to a **region delegate**

AsTree Cascading Tree: Two Hierarchies



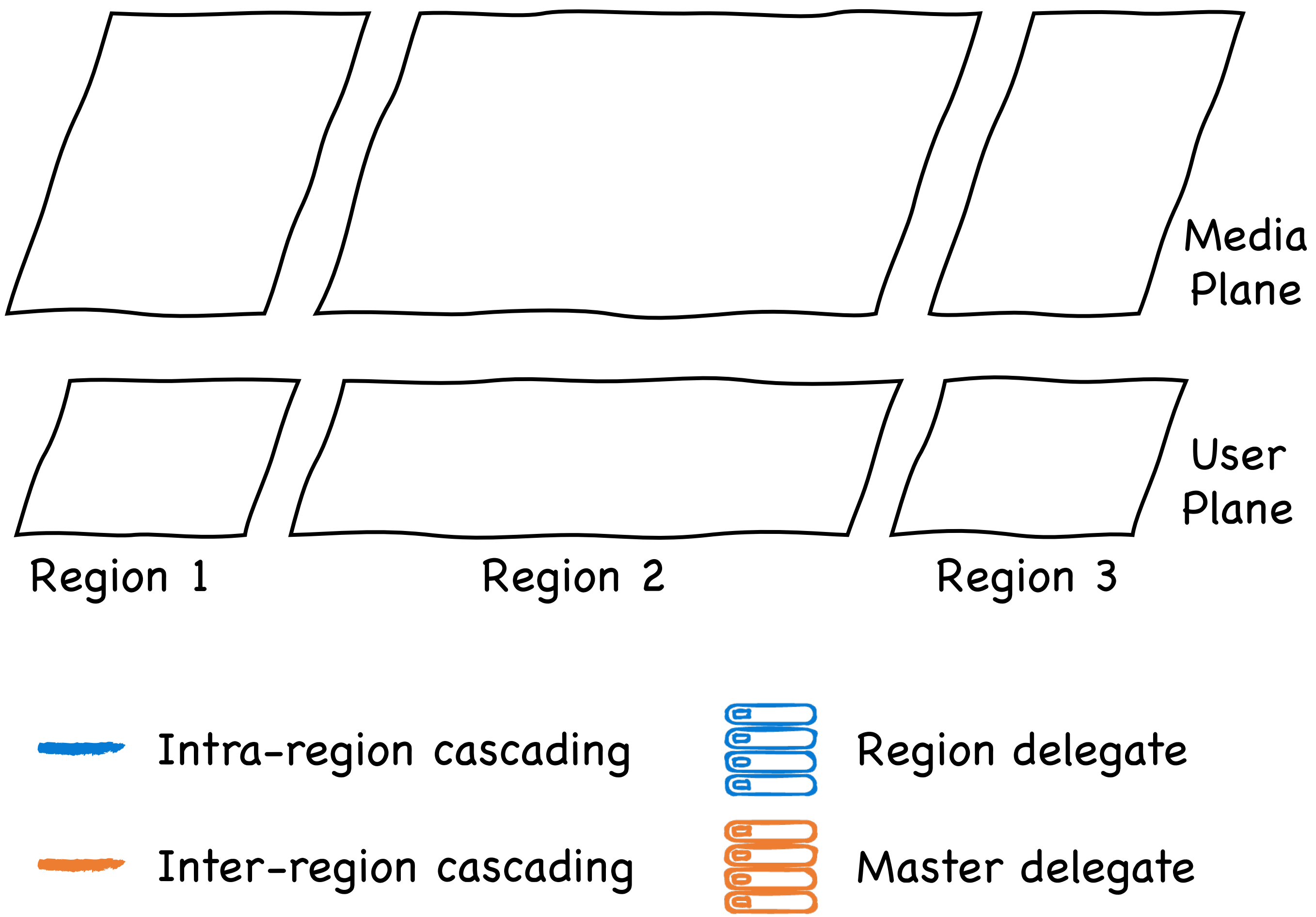
Intra-Region

- All SFUs cascaded to a **region delegate**

Inter-Region

- All region delegates cascaded to a **master delegate**

AsTree Cascading Tree: Incremental Construction



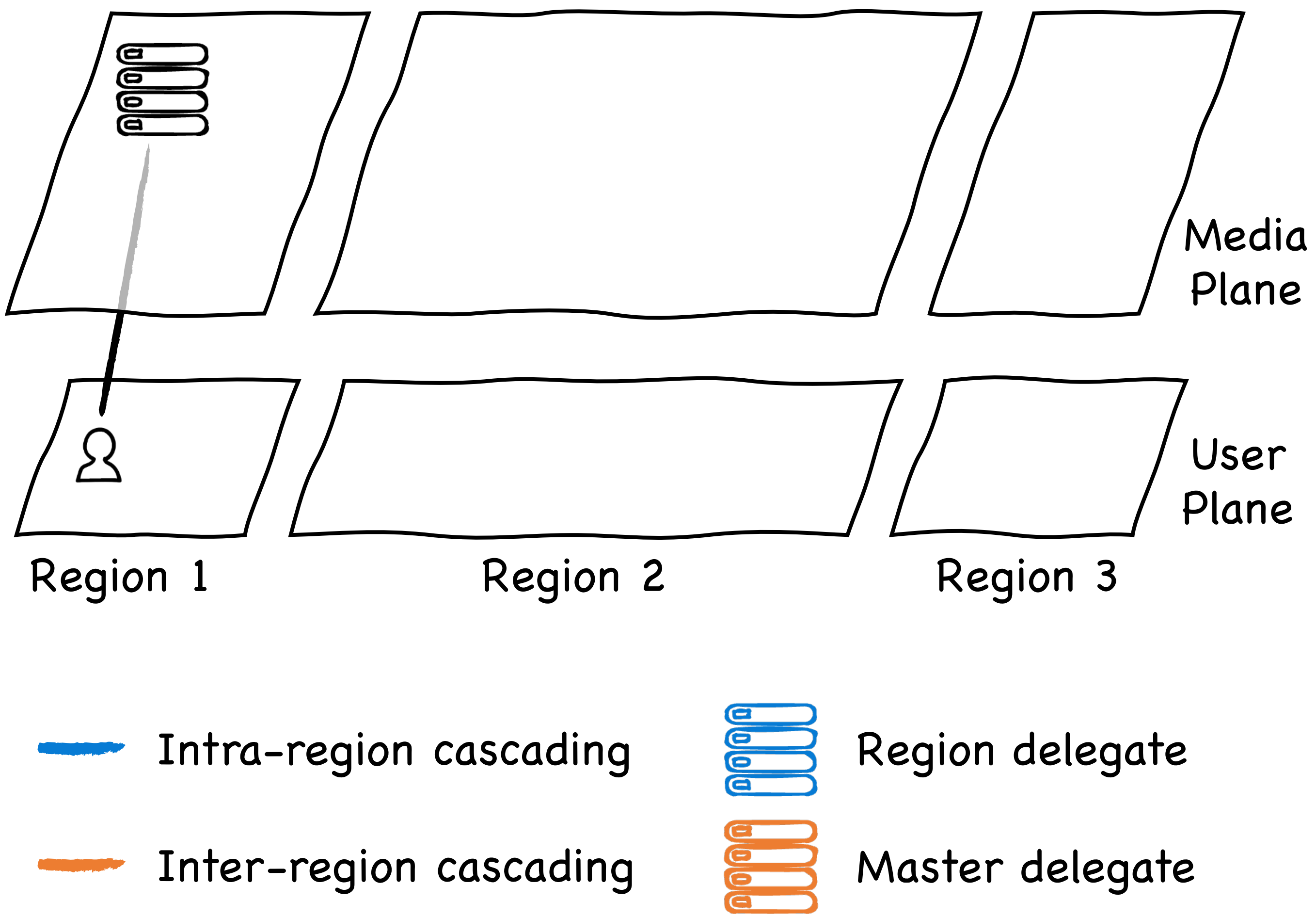
Intra-Region Delegate

- First-comer elected

Master Delegate

- Minimized longest cascading path RTT

AsTree Cascading Tree: Incremental Construction



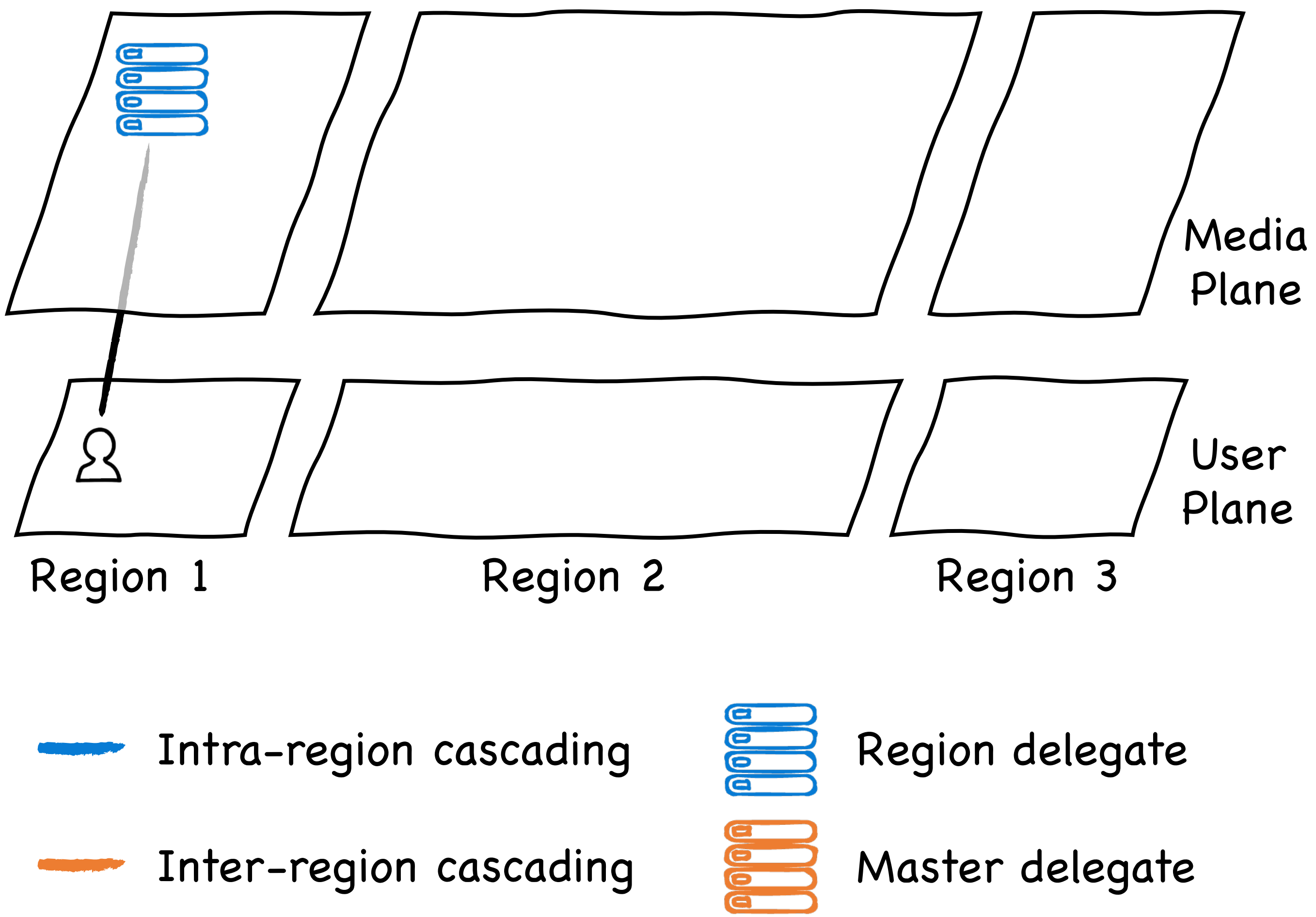
Intra-Region Delegate

- First-comer elected

Master Delegate

- Minimized longest cascading path RTT

AsTree Cascading Tree: Incremental Construction



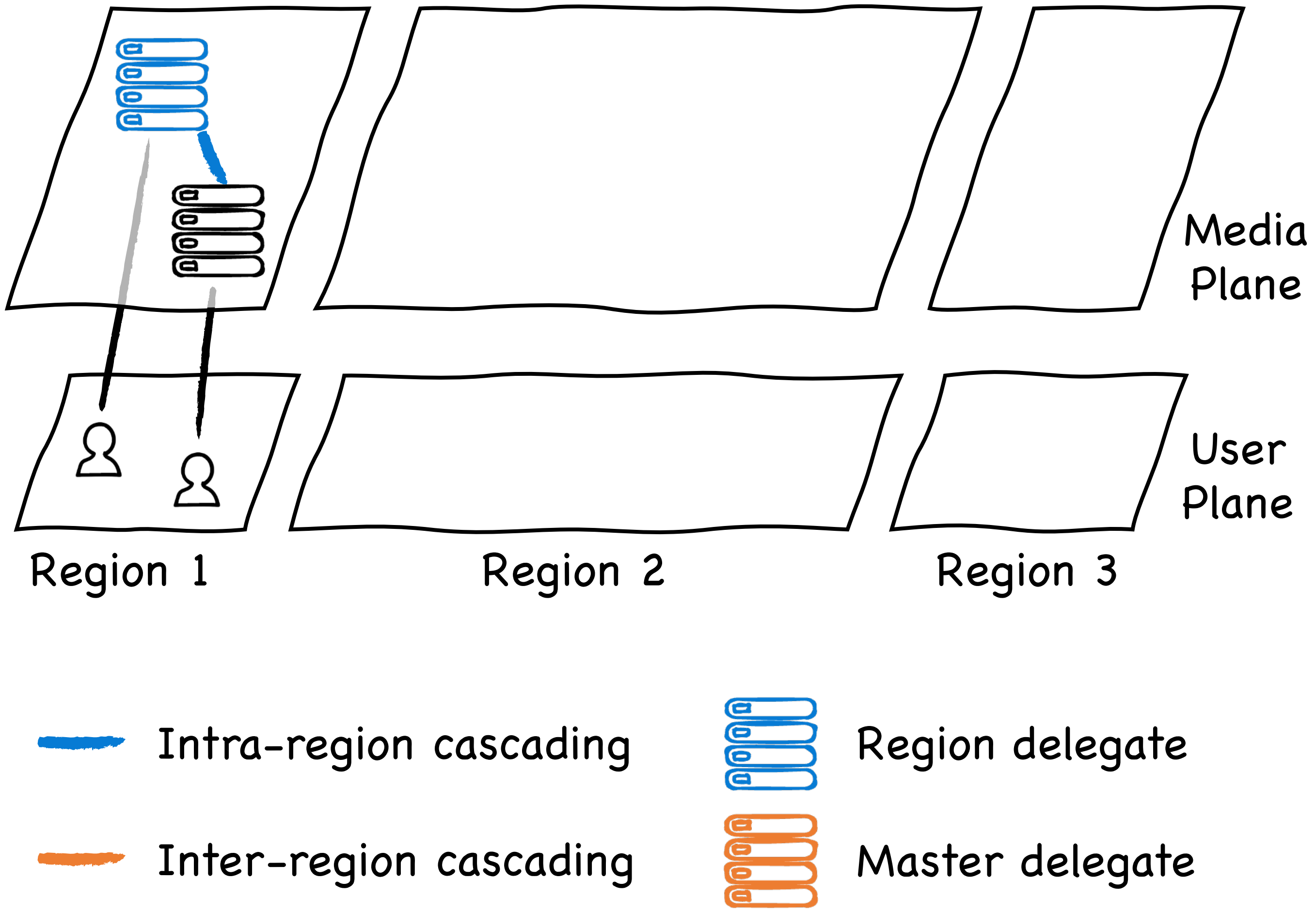
Intra-Region Delegate

- First-comer elected

Master Delegate

- Minimized longest cascading path RTT

AsTree Cascading Tree: Incremental Construction



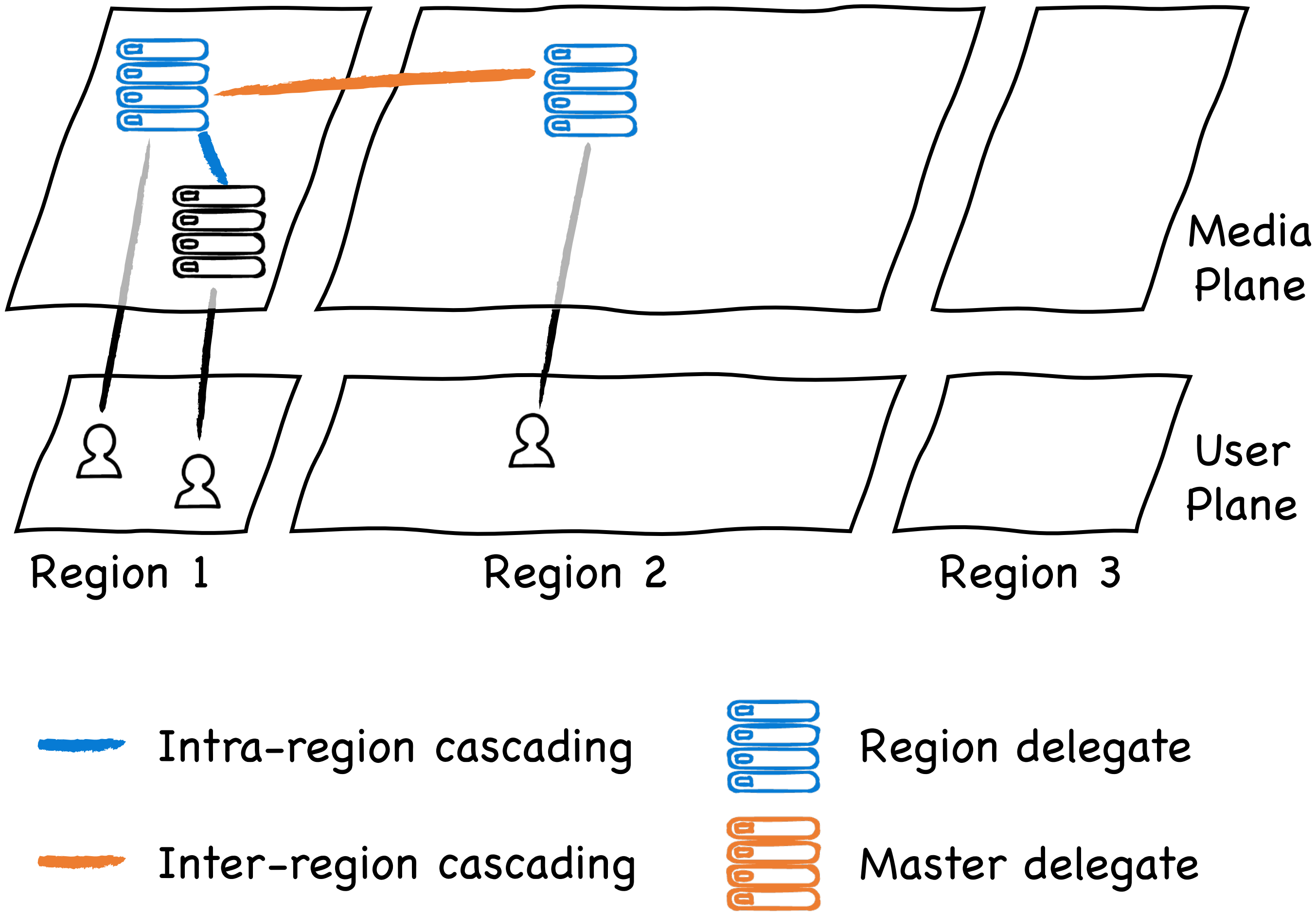
Intra-Region Delegate

- First-comer elected

Master Delegate

- Minimized longest cascading path RTT

AsTree Cascading Tree: Incremental Construction



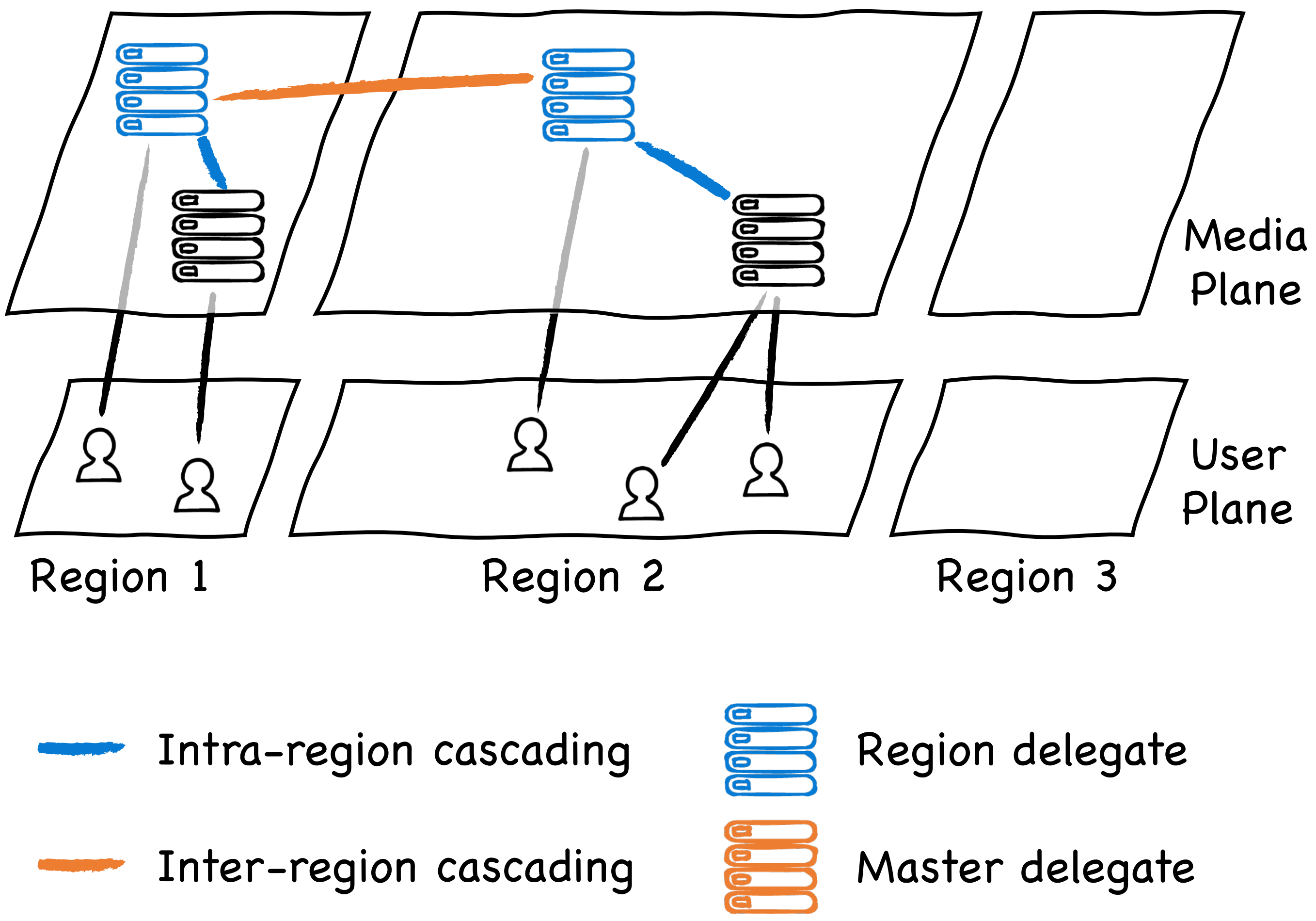
Intra-Region Delegate

- First-comer elected

Master Delegate

- Minimized longest cascading path RTT

AsTree Cascading Tree: Incremental Construction



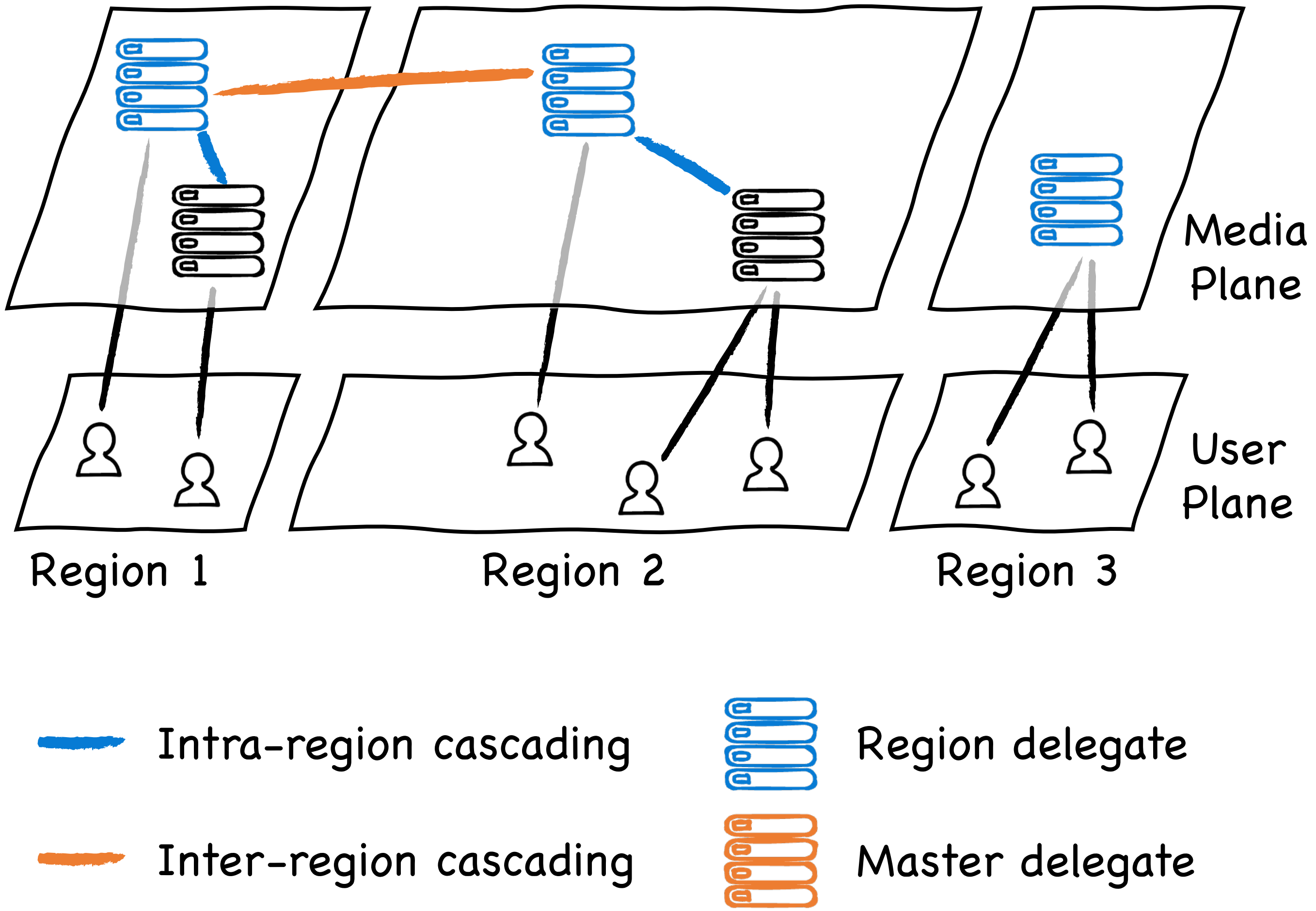
Intra-Region Delegate

- First-comer elected

Master Delegate

- Minimized longest cascading path RTT

AsTree Cascading Tree: Incremental Construction



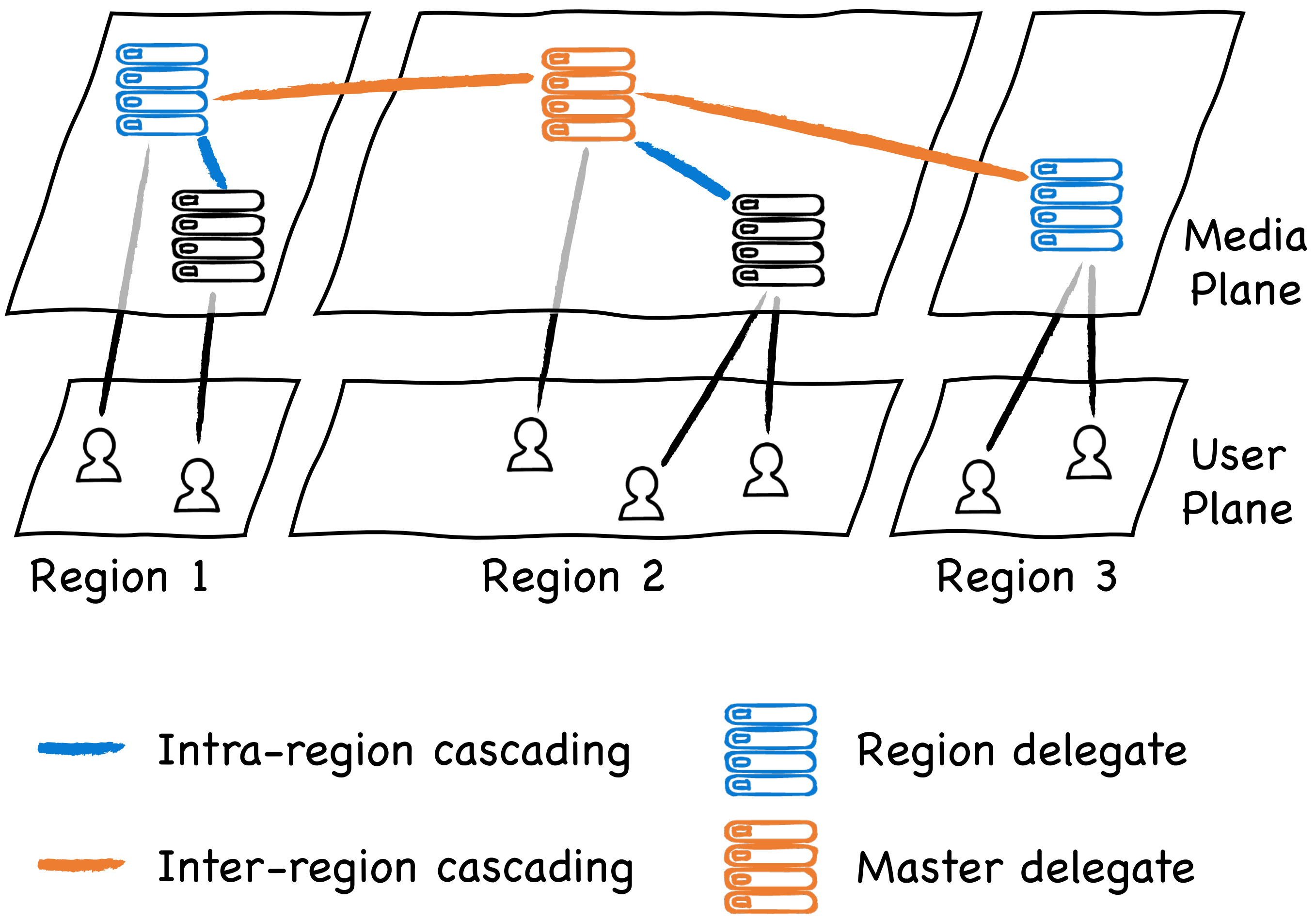
Intra-Region Delegate

- First-comer elected

Master Delegate

- Minimized longest cascading path RTT

AsTree Cascading Tree: Incremental Construction



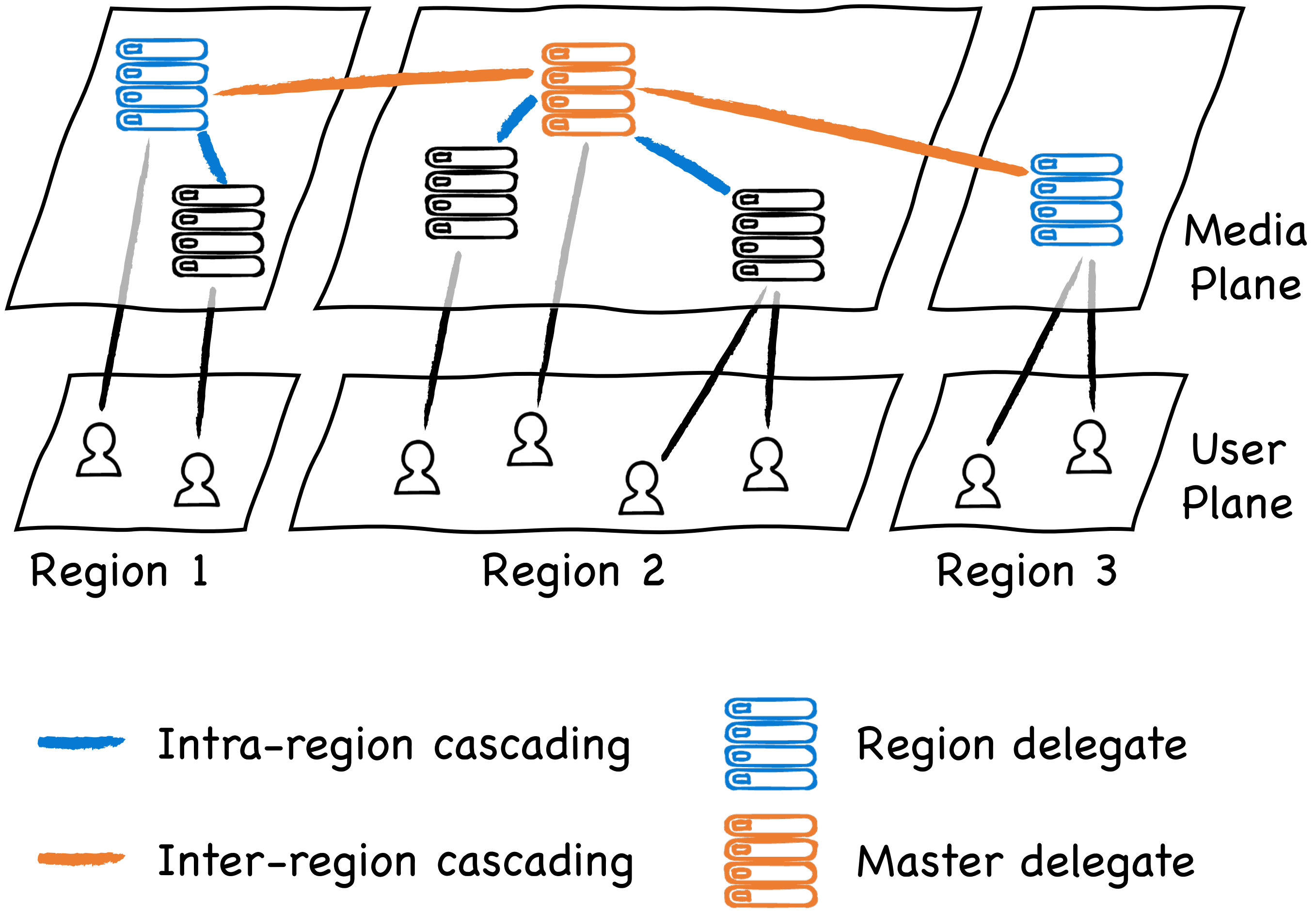
Intra-Region Delegate

- First-comer elected

Master Delegate

- Minimized longest cascading path RTT

AsTree Cascading Tree: Incremental Construction



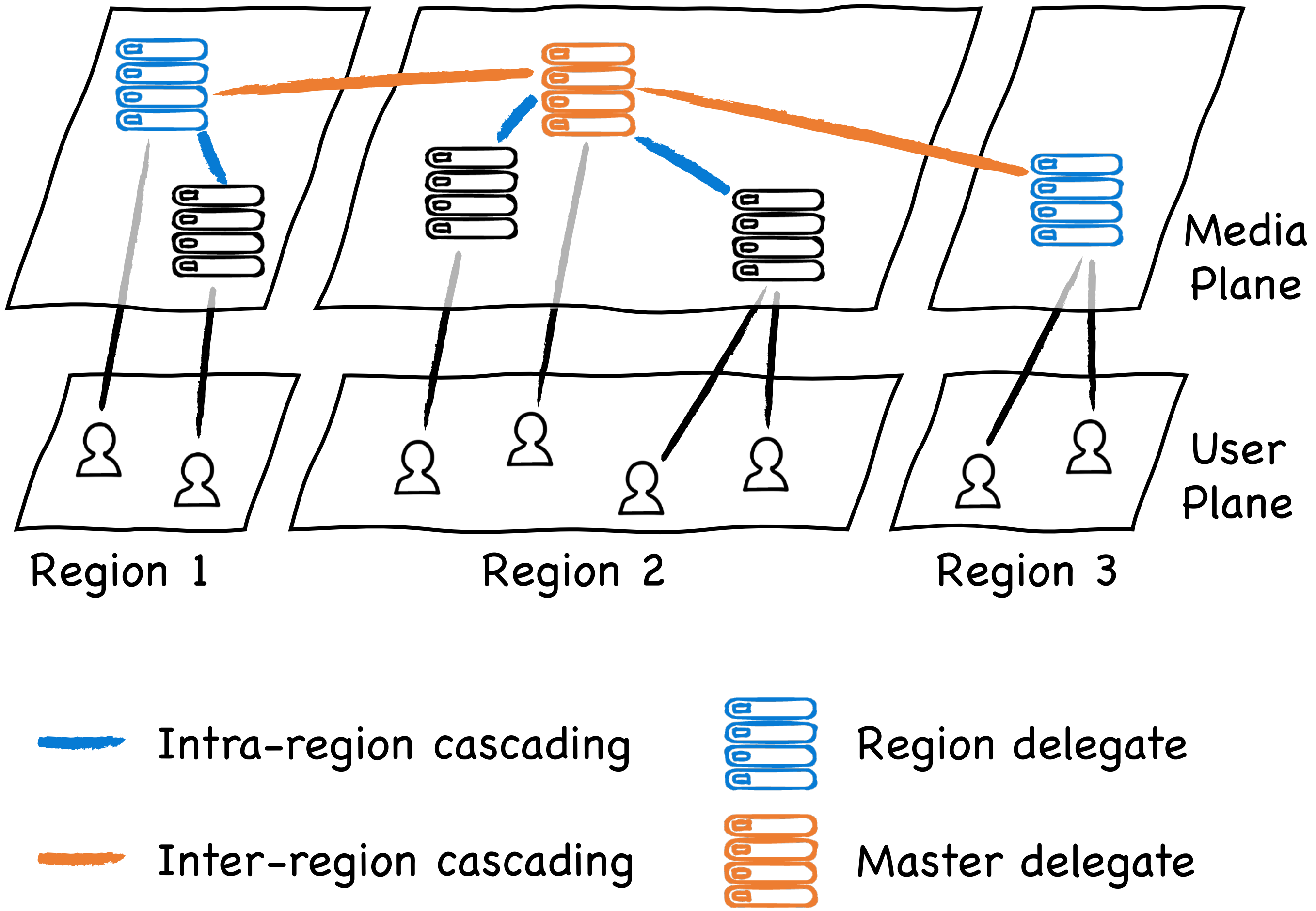
Intra-Region Delegate

- First-comer elected

Master Delegate

- Minimized longest cascading path RTT

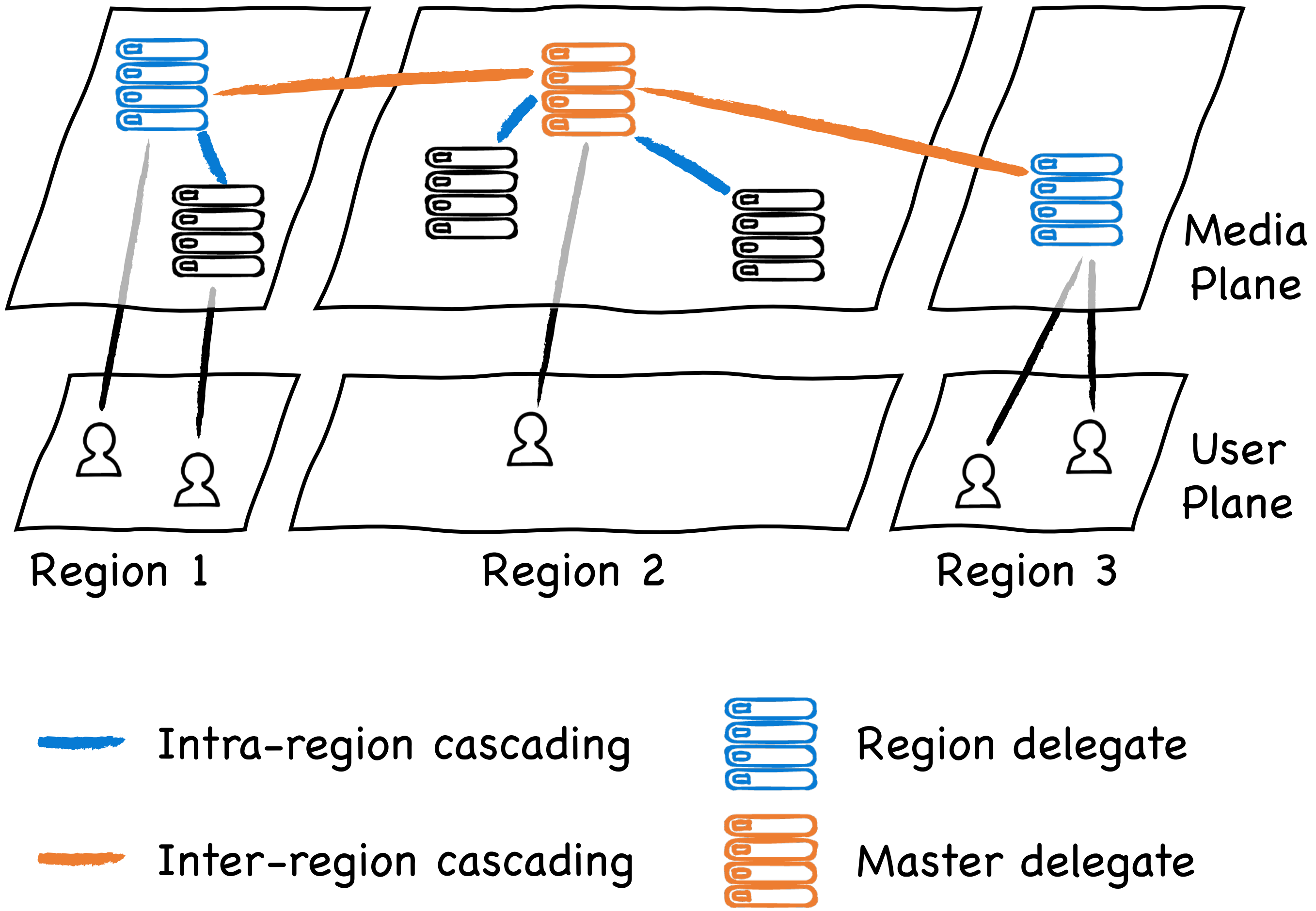
AsTree Cascading Tree: Destruction



No Delegate Re-Election

- Remove an SFU only when all participants in the subtree rooted at it have left

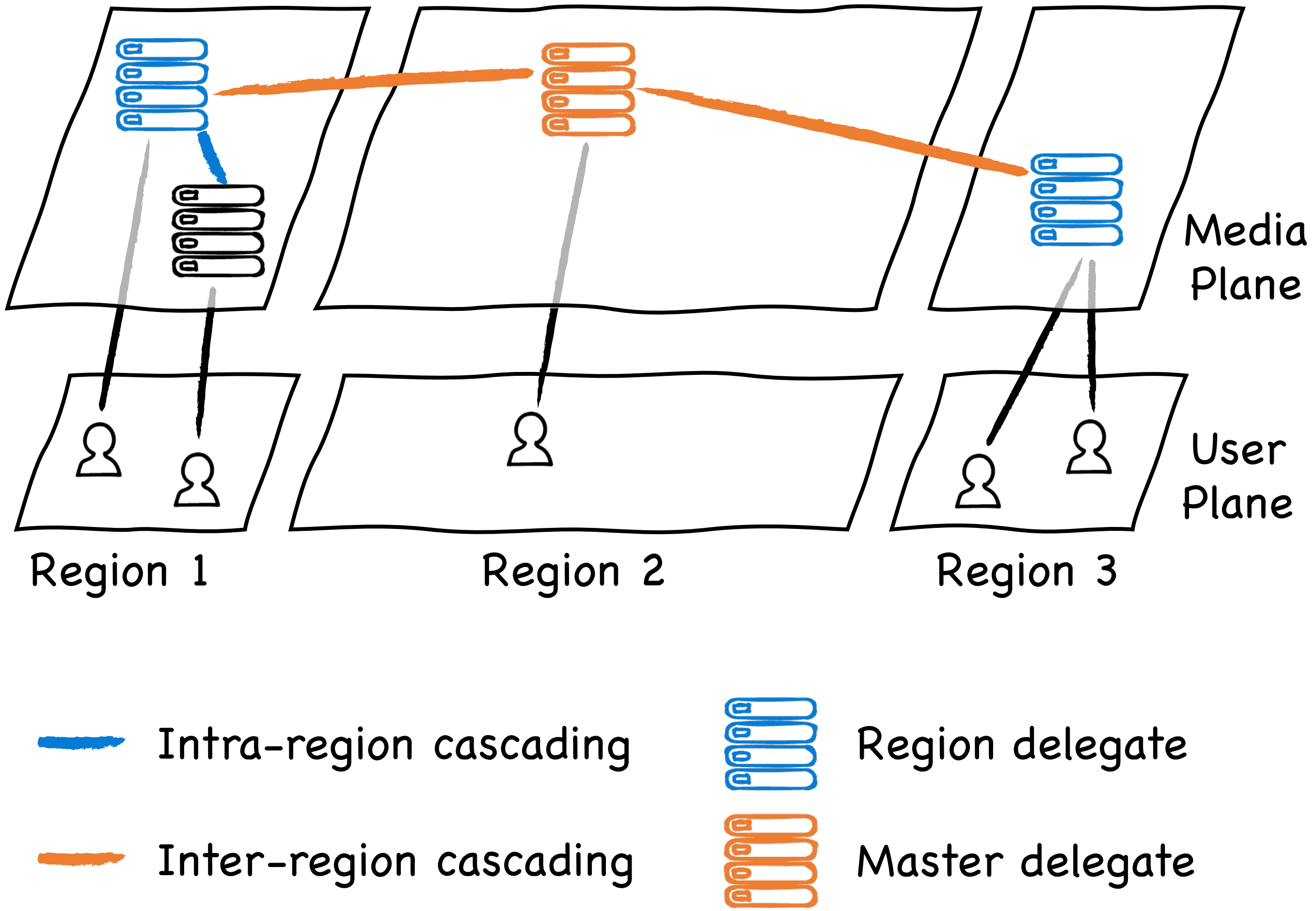
AsTree Cascading Tree: Destruction



No Delegate Re-Election

- Remove an SFU only when all participants in the subtree rooted at it have left

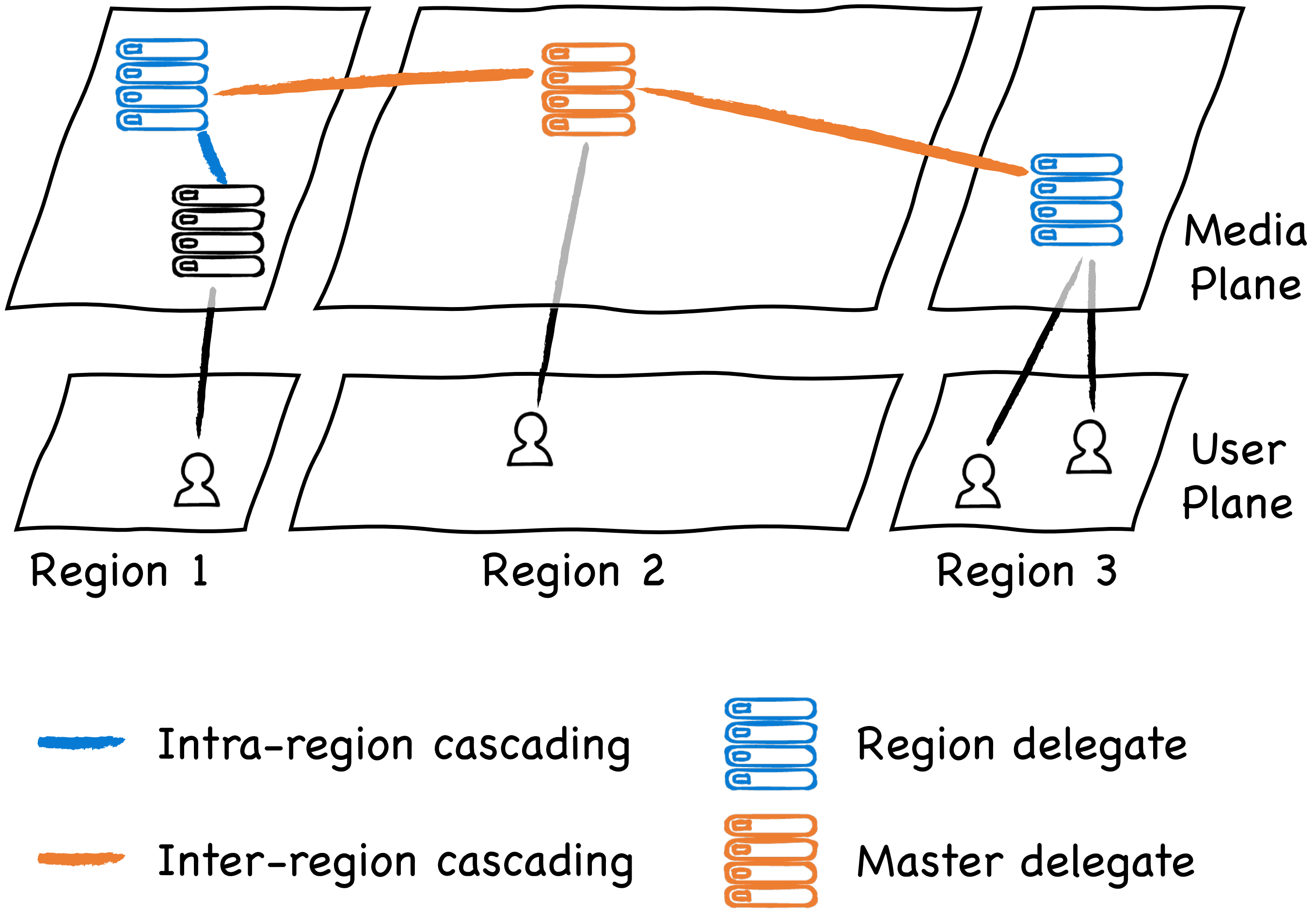
AsTree Cascading Tree: Destruction



No Delegate Re-Election

- Remove an SFU only when all participants in the subtree rooted at it have left

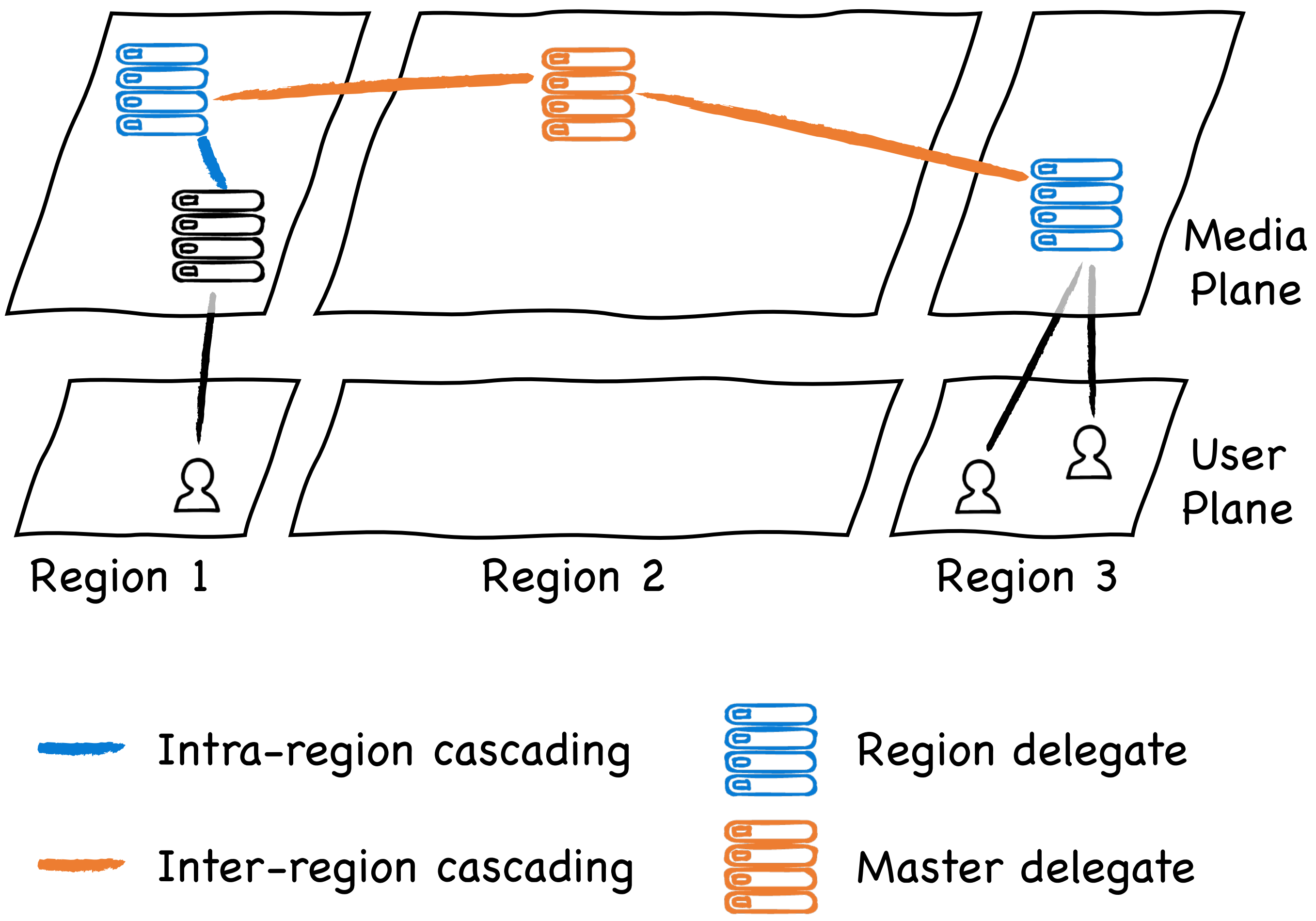
AsTree Cascading Tree: Destruction



No Delegate Re-Election

- Remove an SFU only when all participants in the subtree rooted at it have left

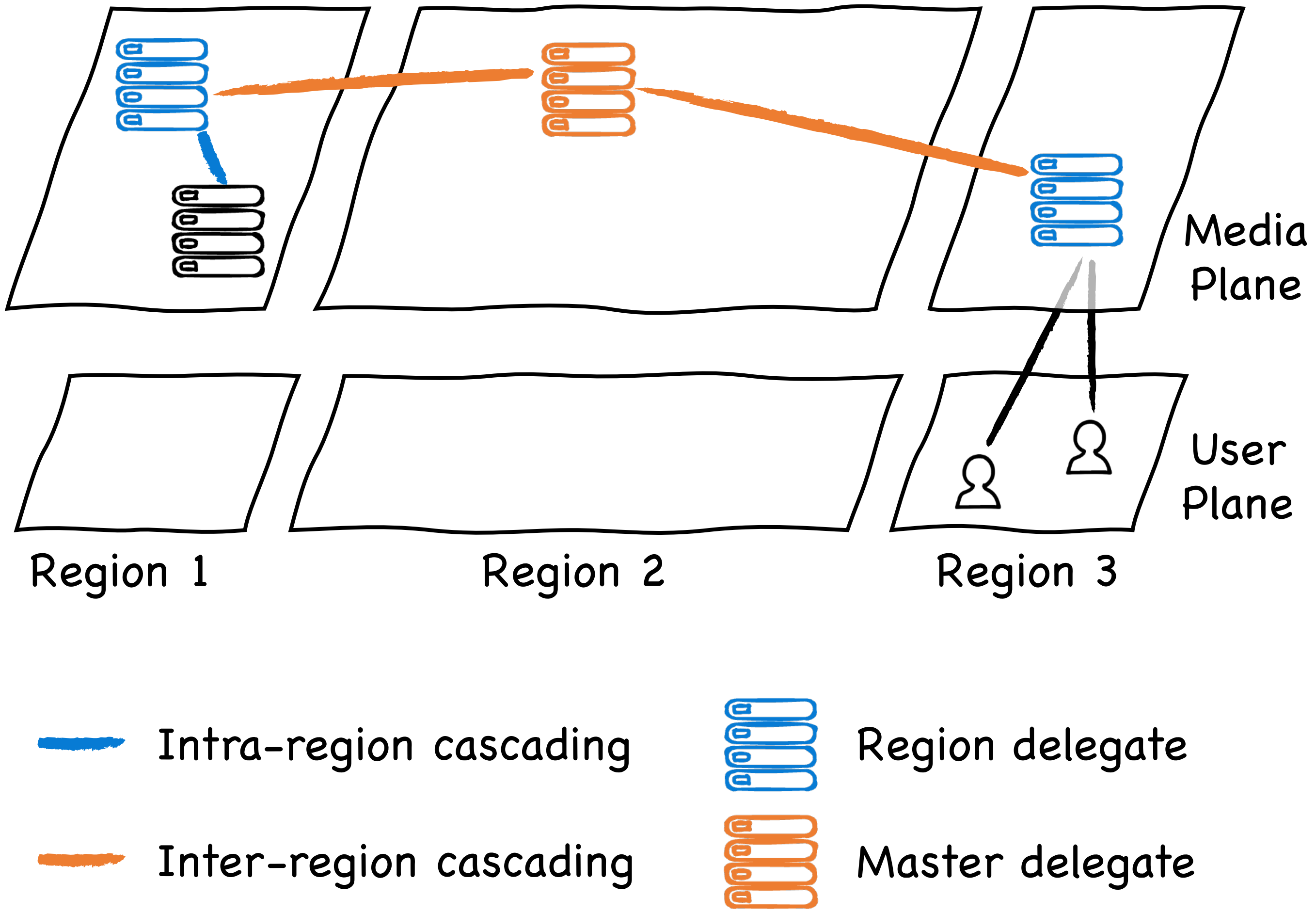
AsTree Cascading Tree: Destruction



No Delegate Re-Election

- Remove an SFU only when all participants in the subtree rooted at it have left

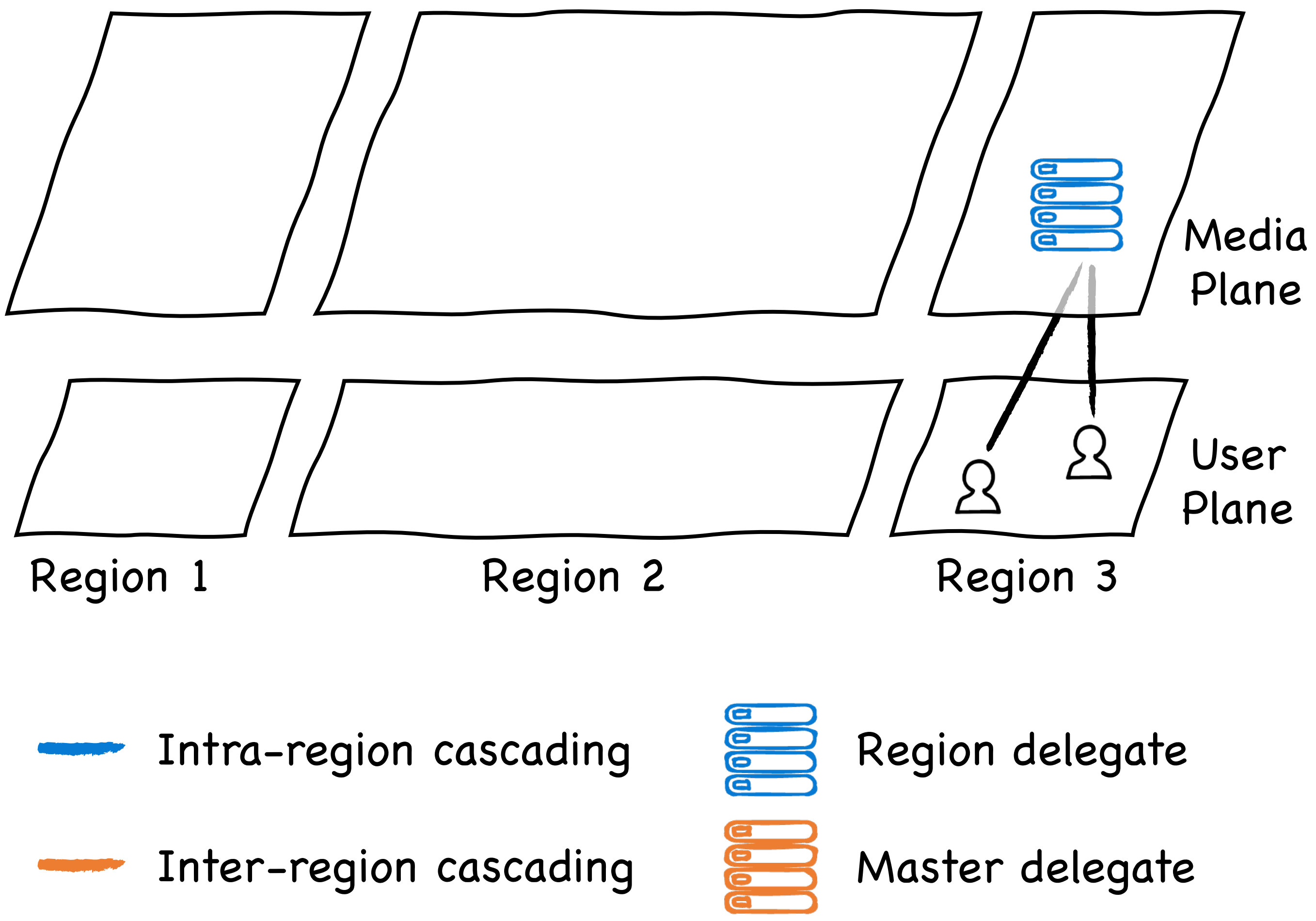
AsTree Cascading Tree: Destruction



No Delegate Re-Election

- Remove an SFU only when all participants in the subtree rooted at it have left

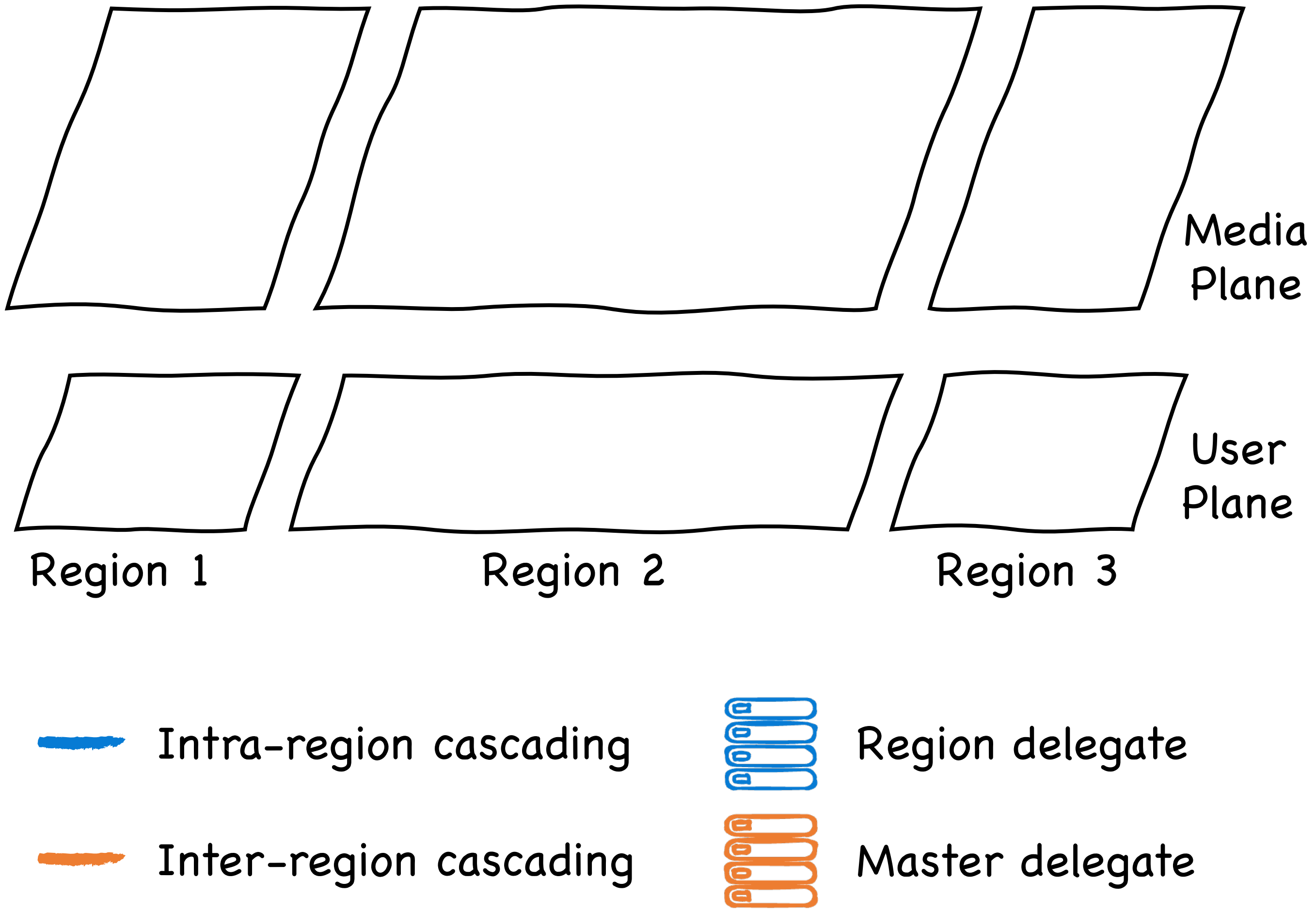
AsTree Cascading Tree: Destruction



No Delegate Re-Election

- Remove an SFU only when all participants in the subtree rooted at it have left

AsTree Cascading Tree: Destruction

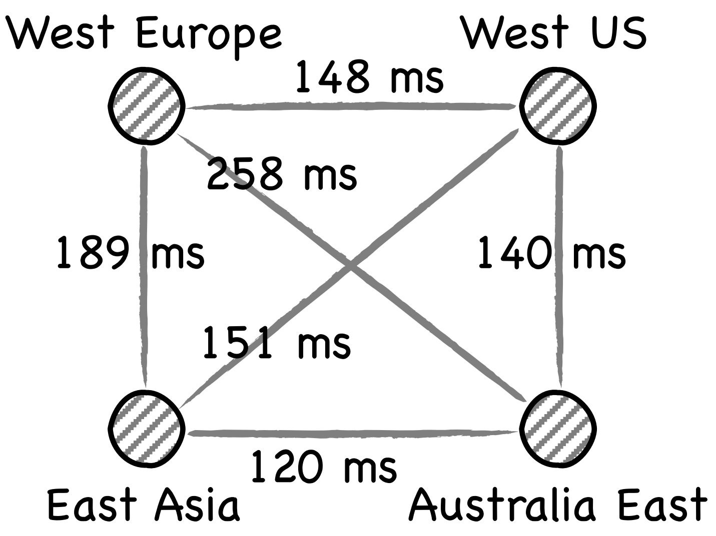


No Delegate Re-Election

- Remove an SFU only when all participants in the subtree rooted at it have left

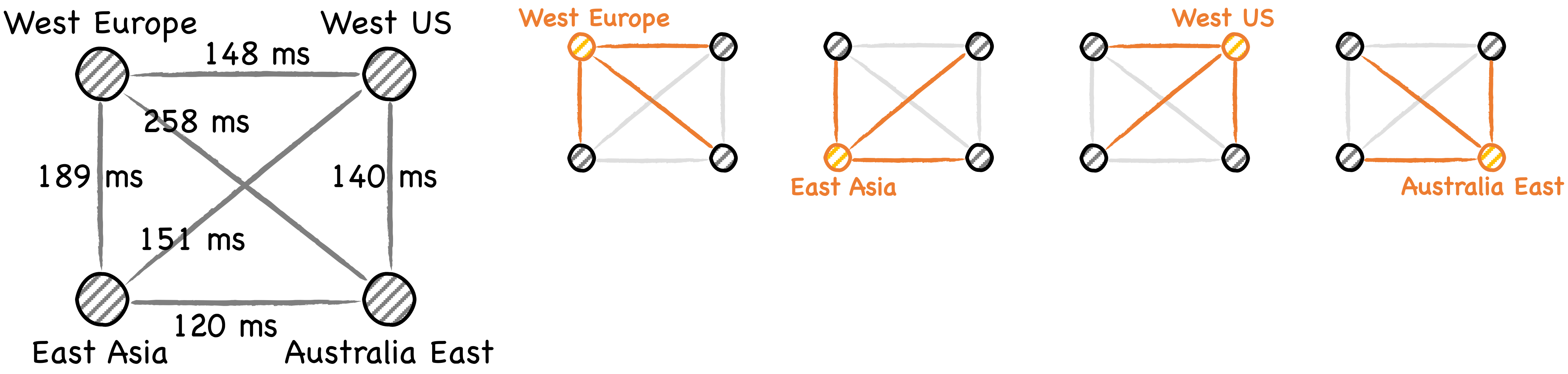
Master Delegate Election

- Can be formulated into a spanning tree problem
- We adopt the objective of minimizing longest cascading path RTT



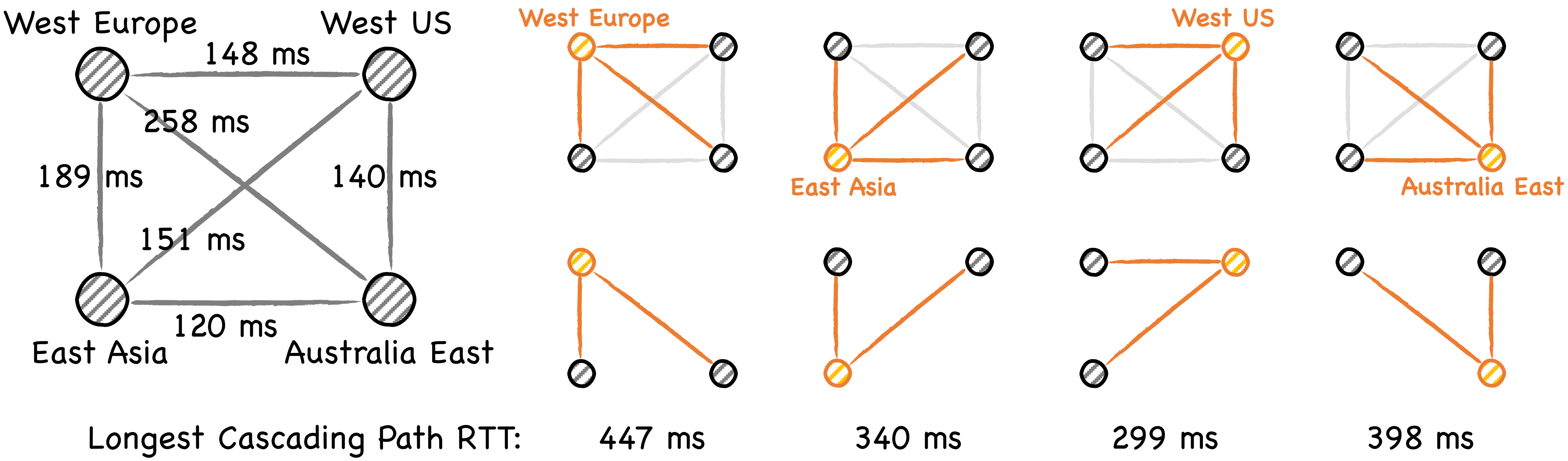
Master Delegate Election

- Can be formulated into a spanning tree problem
- We adopt the objective of minimizing longest cascading path RTT



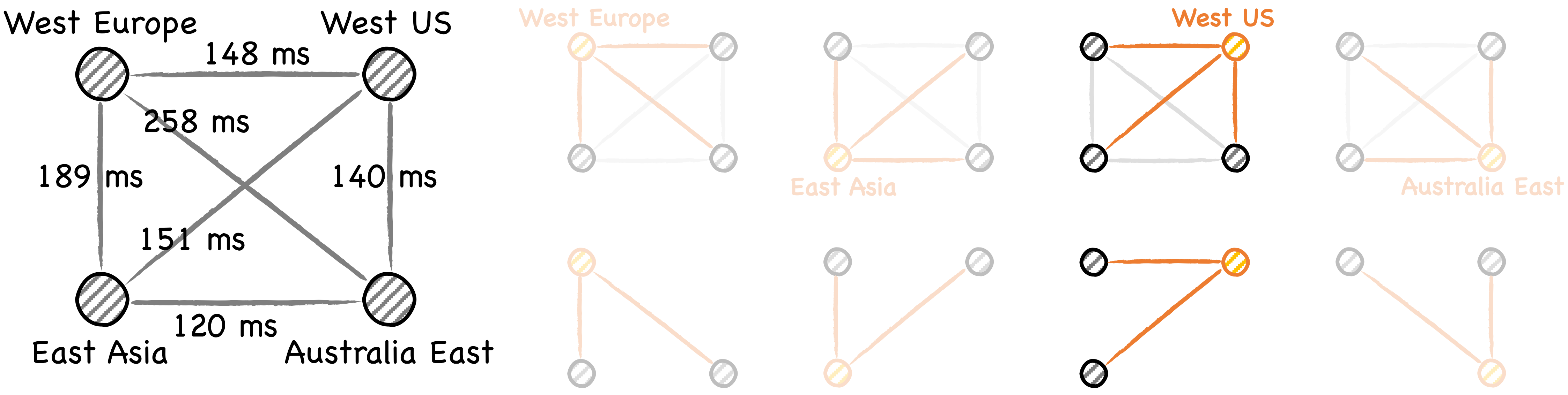
Master Delegate Election

- Can be formulated into a spanning tree problem
- We adopt the objective of minimizing longest cascading path RTT



Master Delegate Election

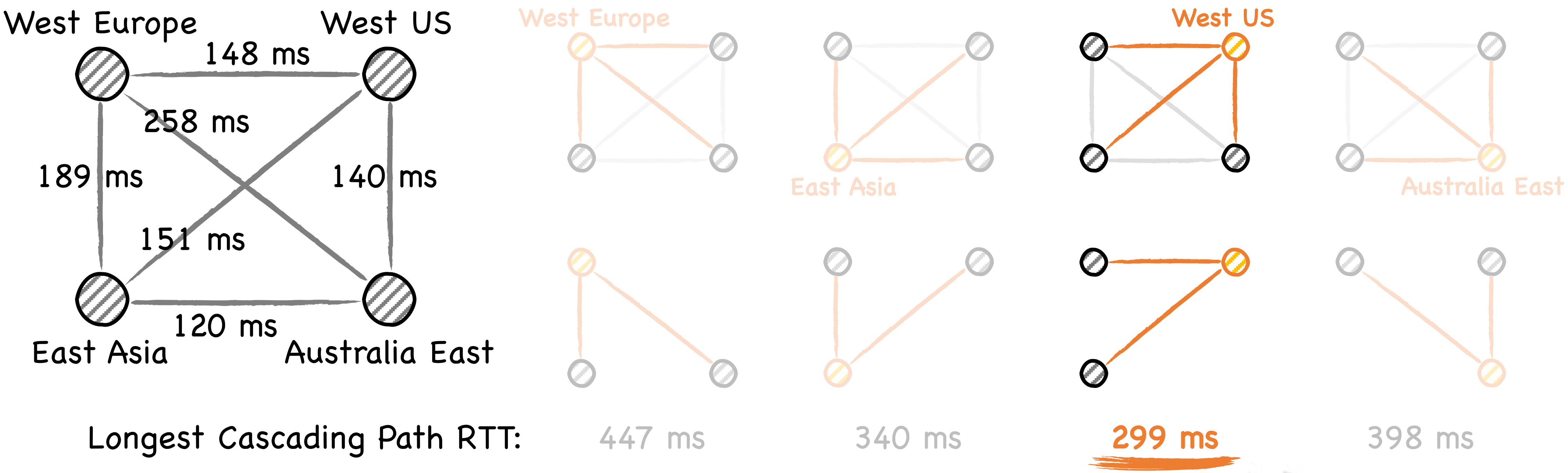
- Can be formulated into a spanning tree problem
- We adopt the objective of minimizing longest cascading path RTT



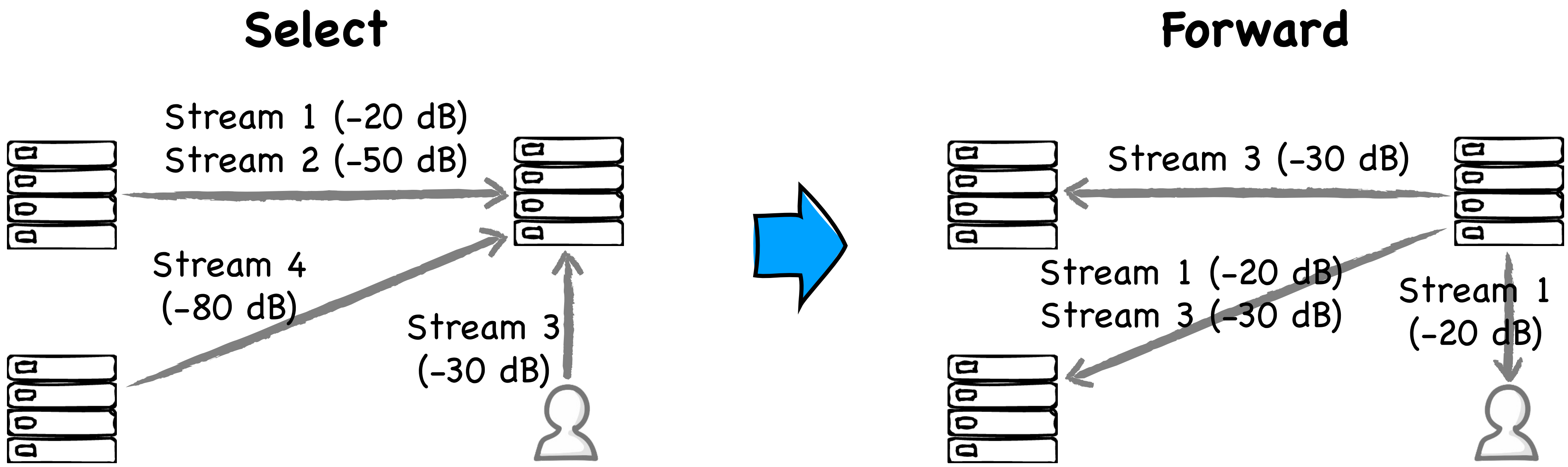
Longest Cascading Path RTT: 447 ms | 340 ms | 299 ms | 398 ms

Master Delegate Election

- Can be formulated into a spanning tree problem
- We adopt the objective of minimizing longest cascading path RTT
- Optimal topology not guaranteed during incremental construction



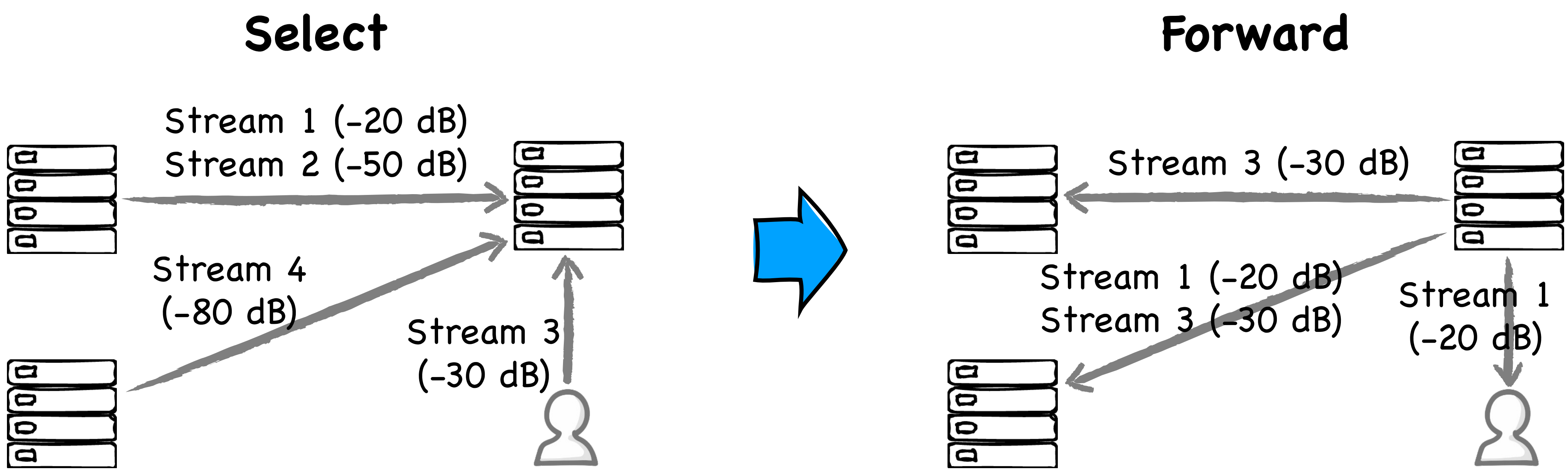
Select-Before-Forward Audio Selection



- Select streams from cascaded SFUs and connected participants
- Audio level carried in RTP header extension, averaged per stream

- Do not forward an audio stream to where it is received

Select-Before-Forward Audio Selection



- Select streams from cascaded SFUs and connected participants
- Do not forward an audio stream to where it is received

Please refer to our paper for dominant speaker identification algorithm details, such as how to avoid frequent changes in the selected set of active speakers

More Details in Paper

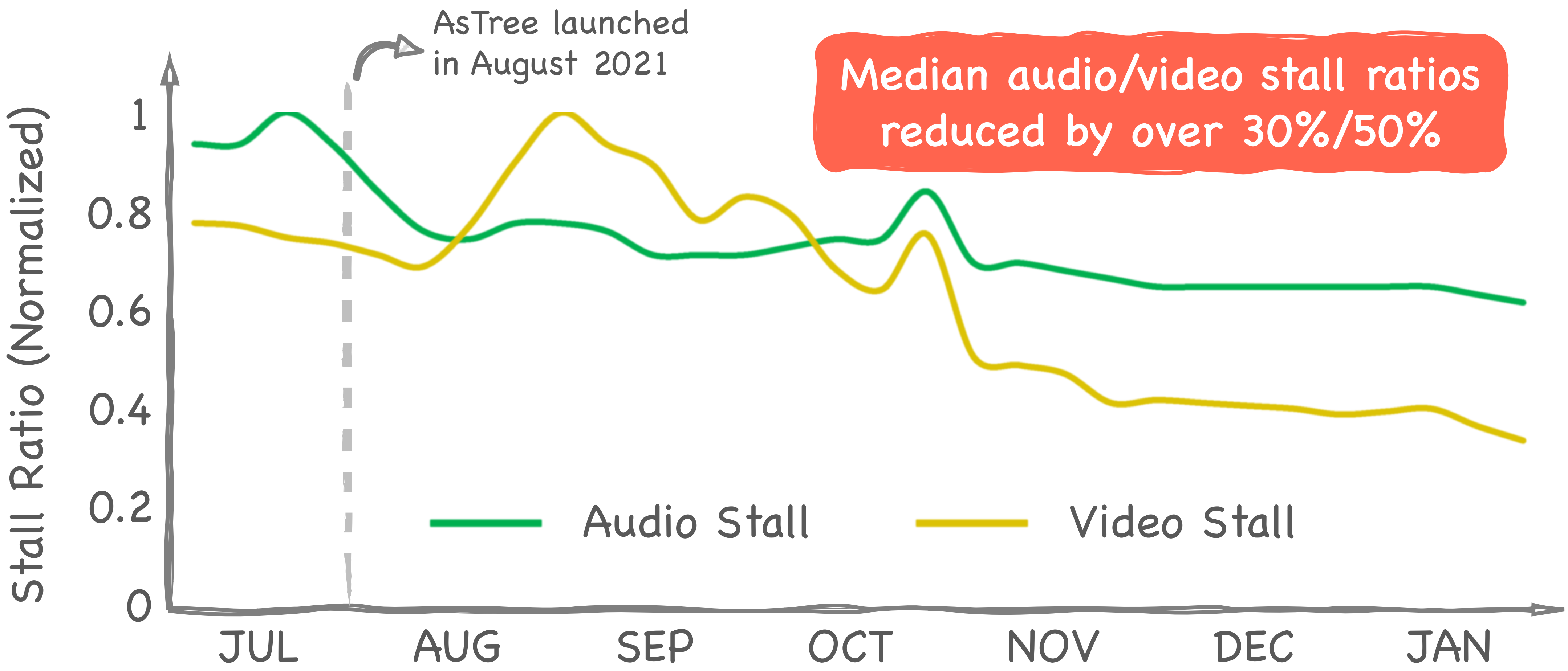
- Implementation
  - Participant aggregation to lower cascading complexity
  - Decoupled video and audio subscription
    - Relies on client-side jitter buffer for audio-to-video synchronization
  - Proactive subscription
    - SDN answer update triggered by SFU notifies identity of newly selected speaker

- Fault tolerance
  - Routing connectivity managed by SDN controller
  - Connection and cascading re-establishment in case of faulty SFUs

Evaluation

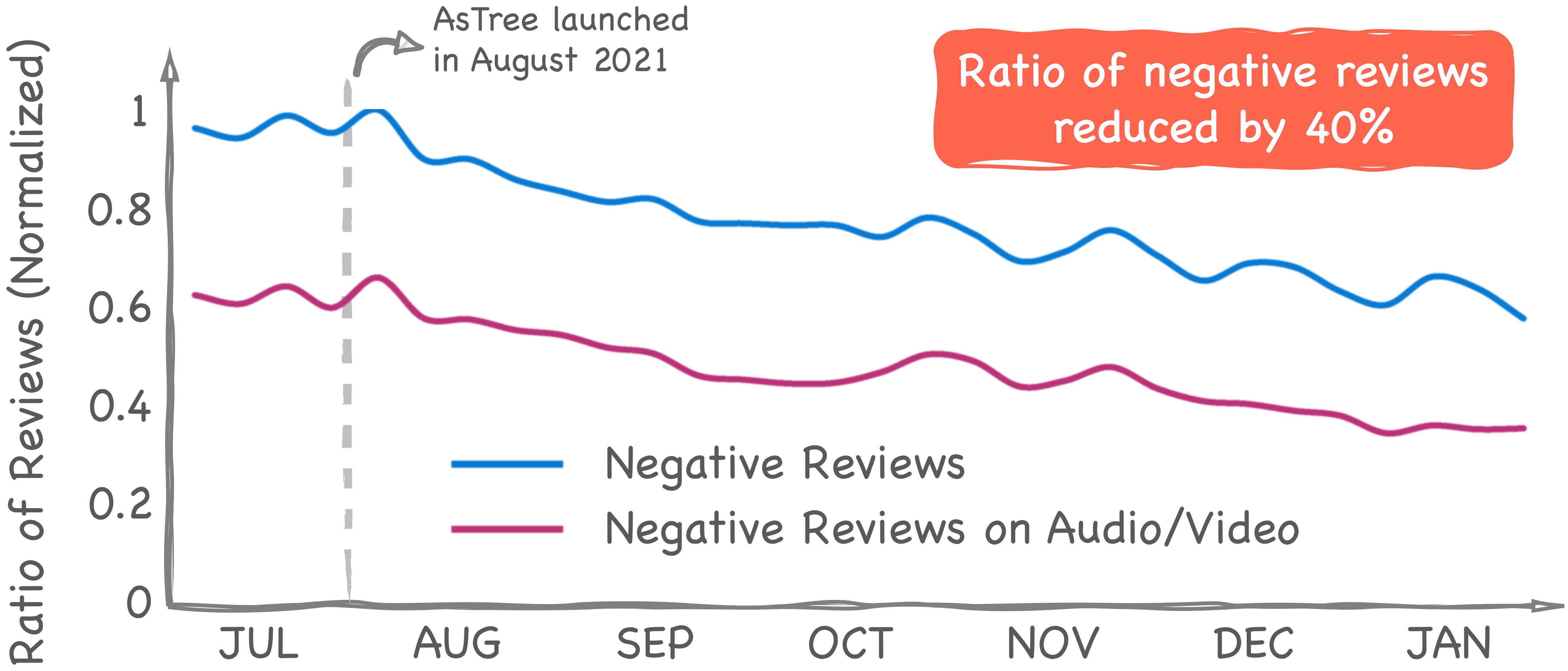
- Controlled experiments in test environment
  - Actual media servers in real-world clusters
  - Isolated from production environment
  - Virtual clients running Linux SDK
- Results Highlights
  - Consumes 80%-90% less CPU and memory on server side with 100+ clients per room
  - Increases server capacity (*i.e.*, participants per room per server) by at least 8x
  - Lowers audio/video rebuffering ratios by orders of magnitude with 100+ clients per room

Deployment



AsTree launched
in August 2021
Median audio/video stall ratios
reduced by over 30%/50%
Stall Ratio (Normalized)
1
0.8
0.6
0.4
0.2
0
JUL
AUG
SEP
OCT
NOV
DEC
JAN
Audio Stall
Video Stall

Deployment



AsTree launched
in August 2021
Ratio of negative reviews
reduced by 40%
Ratio of Reviews (Normalized)
1
0.8
0.6
0.4
0.2
0
Negative Reviews
Negative Reviews on Audio/Video
JUL
AUG
SEP
OCT
NOV
DEC
JAN

Wrap-up

- **AsTree** to scale audio subscription in multi-party conferencing
  - Two-hierarchy cascading tree, based on geographical regions
  - Select dominant speakers based on per-stream audio levels

- Open Questions
  - More sophisticated optimization objective for AsTree construction
  - Maintain topology optimality, e.g., with dynamic cascading deformation

THANKS

ByteDance字节跳动