Invention Title
Distributed scalable Floor Control Model for Audio & Video Conferencing

Invention Summary
This invention disclosure proposes an improved, distributed and scalable algorithm for determining the loudest N speakers in an A/V conference. We propose multiple scalable approaches for granting so-called floor access to a particular speaker (being the speaker in the A/V conference whose volume control is turned up while others are muted). Our approach is particularly well suited for (but not limited to) conferencing systems that are hosted in cloud computing environments. Furthermore, the distributed approach proposed here facilitates hosting of A/V conferences by different cloud providers more easily than a strict centralized model.

Invention Description
For large audio and video conferences (10's of participants or more), it is desirable to have a system in place that includes Automatic Gain Control (ACG) to adjust the volume of various speakers participating in the conference, so as to even out volume levels among speakers for improved speech intelligibility. Current state of the art enables systems to determine the loudest speakers in an audio/video conference. The conferencing system then turns up the gain control for these loudest speakers, while turning the gain control down or completely off (mute) for everyone else. This approach of granting speaking privilege in a conference call is generally referred to as "floor access". In current systems, the decision process regarding the determination of the loudest N speakers is carried out in a single, centralized location.

As A/V conferencing systems are moving towards cloud-based implementations, a strictly centralized model for determining speaker volume, making decisions about applying AGC and applying mute control is no longer optimal. Cloud-based A/V conferencing systems end up running on a variety of computing resources, either inside a single cloud computing site, multiple sites across a larger geographic area or potentially even in multiple cloud computing sites dispersed across an entire country or among multiple continents. Thus, an efficient and scalable signaling system to aid in making decisions about speaker volume and related aspects of A/V conferencing is an absolute necessity.

For the purpose of improving speech intelligibility, an A/V system typically considers the loudest N speakers in a conference of M participants (N<M). Clearly, N =1 for a single loudest speaker, but typically N<5...10 in most cases).

Please note that speaker groups in the A/V conference call could be hosted by different providers. This approach lends itself very well to a cloud computing-based system.

Case (1) - Conventional centralized floor control model (current state of the art)
The audio/video conferencing system is hosted on a single resource and makes centralized decisions about conferencing floor access. In the drawing below, all participants are connected to the same computing entity, hosted in a single, centralized location. The central location of the hosting entity imposes bandwidth and computational limitations as the number of participants scales up.

Case (2a) - Distributed scalable floor control model (this invention disclosure)
The audio/video conferencing system is hosted on several, potentially distributed resources and makes a distributed decisions about conferencing floor access. In the drawing below, smaller,

decentralized, more agile computing entities are assigned to smaller groups of conference participants. Each computing entity makes its own localized decisions about speaker volume and other decisions about granting floor control access. Once the loudest speaker in each group has been determined, the result is communicated to the level above, where yet another decision is made to determine the loudest speaker from each group.

Case (2b) - Distributed peer-to-peer scalable floor control model (this invention disclosure)
The audio/video conferencing system is hosted on several, potentially distributed resources and makes a distributed decisions about conferencing floor access. Decisions are based on a peering model where an agile, localized controller makes decisions about floor access and exchanges its decision with its nearest neighbor peers. This process is repeated until all peers agree on a speaker to grant floor access to. The advantage of this approach is that no centralized entity is needed to process floor access requests.

Please note that this scheme is extensible in the sense that multiple hierarchies of hosting entities and floor access decision making processes can be stacked on top of each other.

Case (2c) - Distributed hierarchical scalable floor control model (this invention disclosure)
The audio/video conferencing system is hosted on several, potentially distributed resources and makes a distributed decisions about conferencing floor access, as in Case 2b above. In addition, the decision is made in a hierarchical approach by passing peer decisions up to the controller for final floor access arbitration. At the top level, all computing entities determine the loudest N speakers via a peer-to-peer decision mechanism.

The combination of a peer-to-peer model with a hierarchical controller means that fewer, local decisions can be made and the controller can deal with a light data load. As a result, it could be implemented as a lightweight controller requiring few computing resources.

Invention Commercial Value/Customers
Scalable and distributable allocation of computing resources - customer can assign computing resources to best trade off latency, computing power, bandwidth need and system response time.

Audio/ video conferencing host can move & allocate conference-related applications and data usage based on available computing resources in hardware and computing cloud.

Invention Differences
Conventional audio/video conferencing systems make decision processes regarding floor control in a single, centralized location.