Vidyo™
Personal Telepresence

Scalable Video Coding and Videoconferencing

Danny Hong

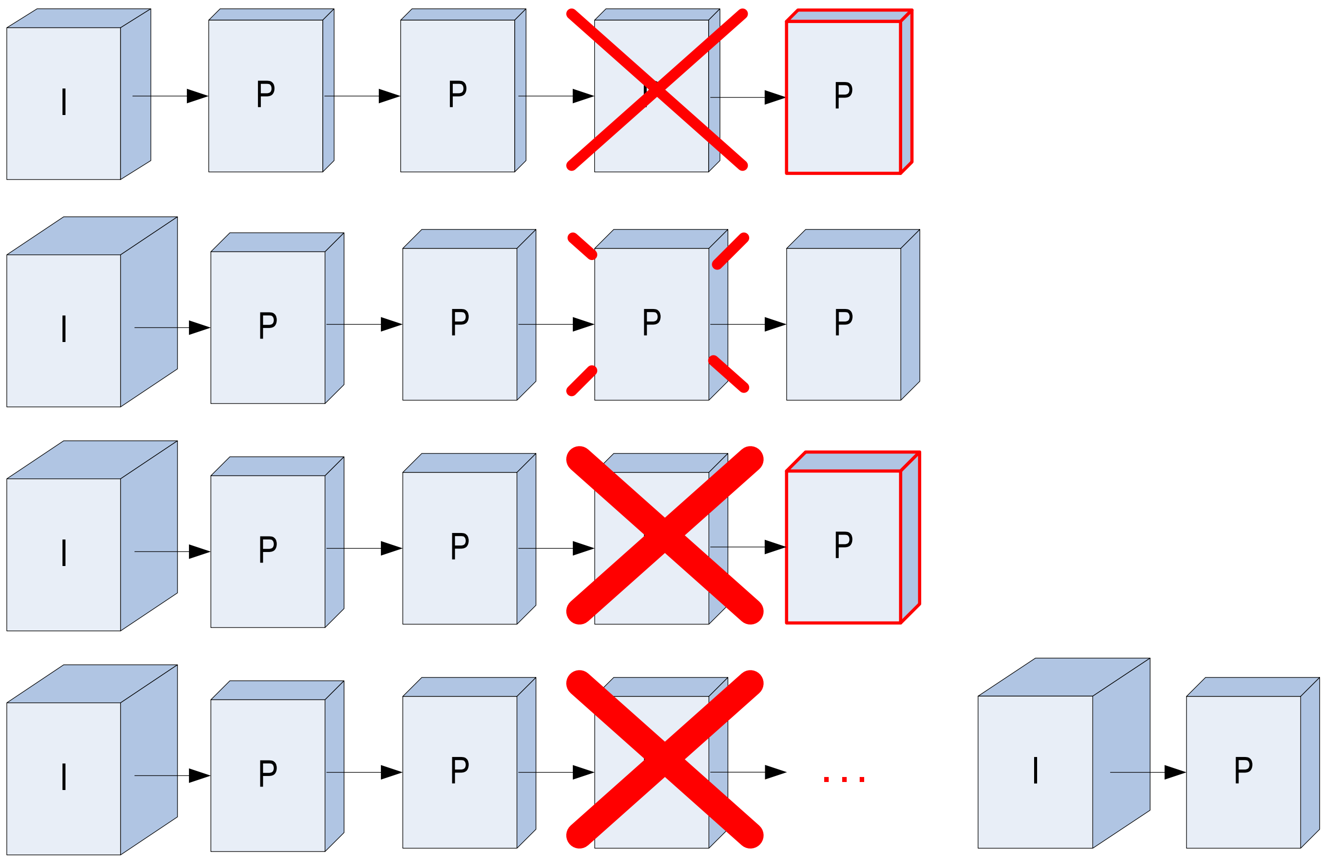
Videoconferencing over General IP Networks

Challenges to Videoconferencing

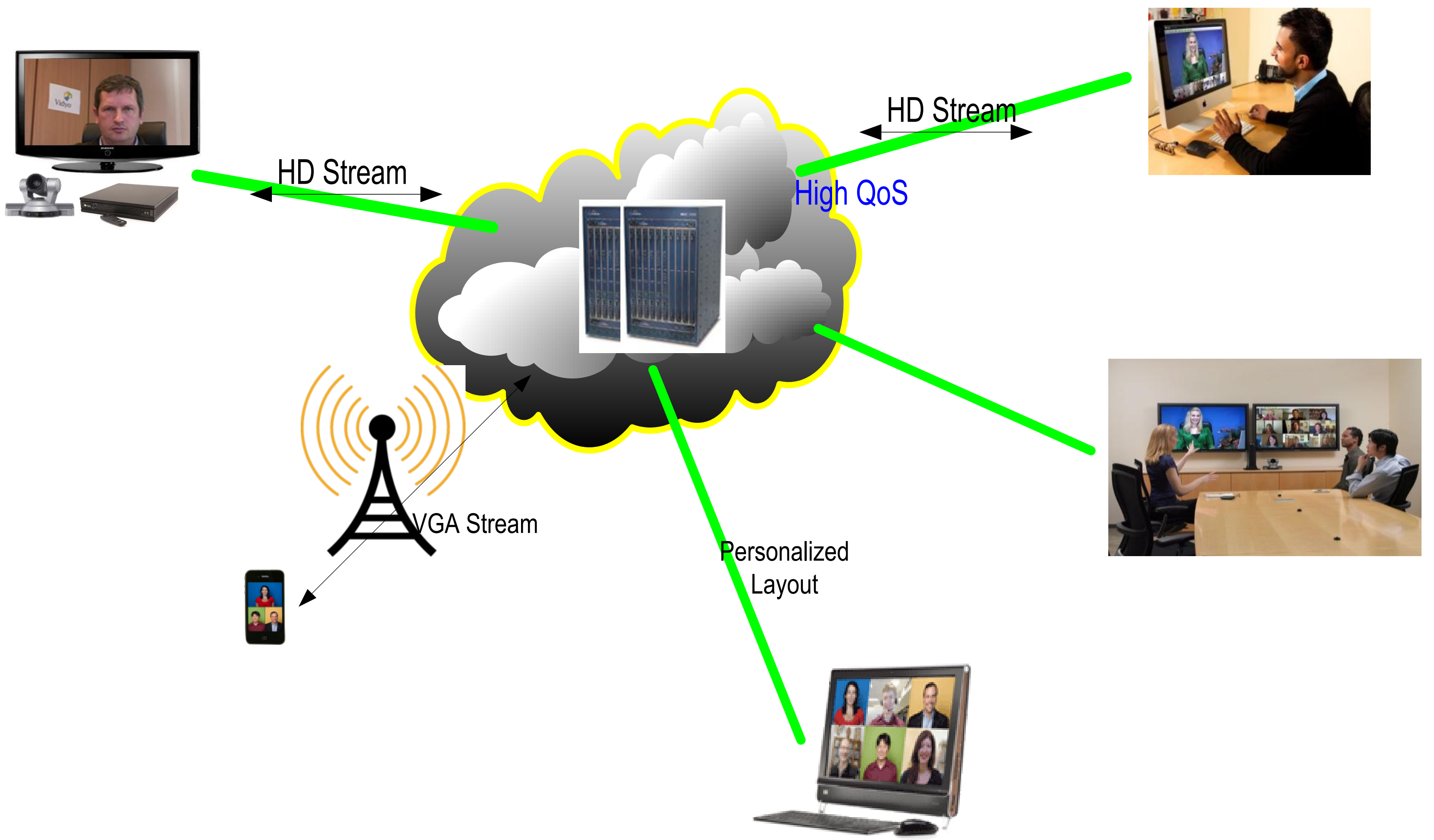
Vidyo™ Personal Telepresence

- High bandwidth associated with high quality video
- Fluctuating network bandwidth
- Packet loss and jitter
- Users may access videoconferencing services over channels that have very different bandwidths (e.g., DSL vs. Ethernet)
- Different endpoints with different capabilities
- Traditional video codecs are designed to provide a single bitstream at a specified bit-rate for a target resolution
- Need to maintain acceptable total delay for natural communication

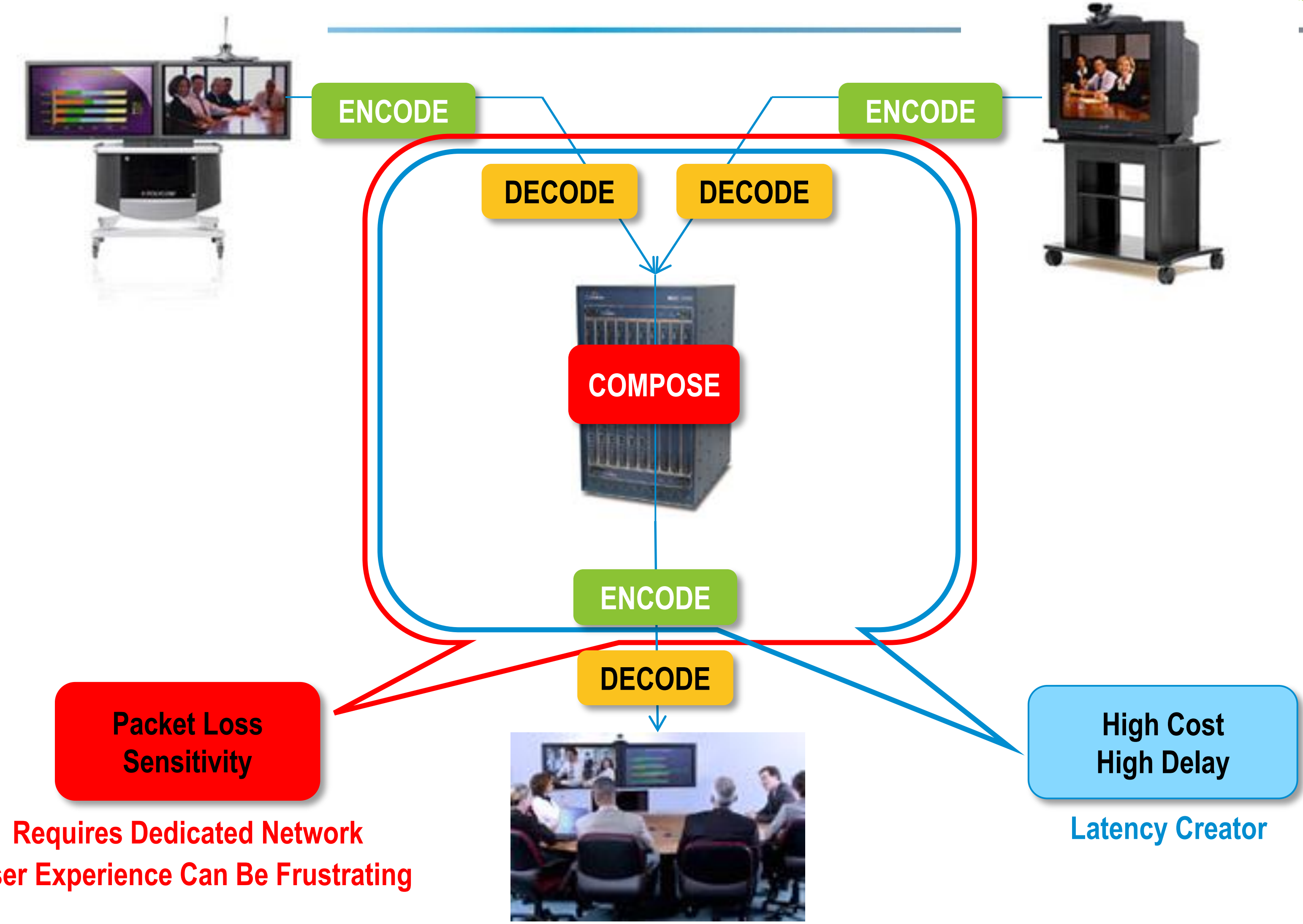
Traditional, Non-scalable Video Coding



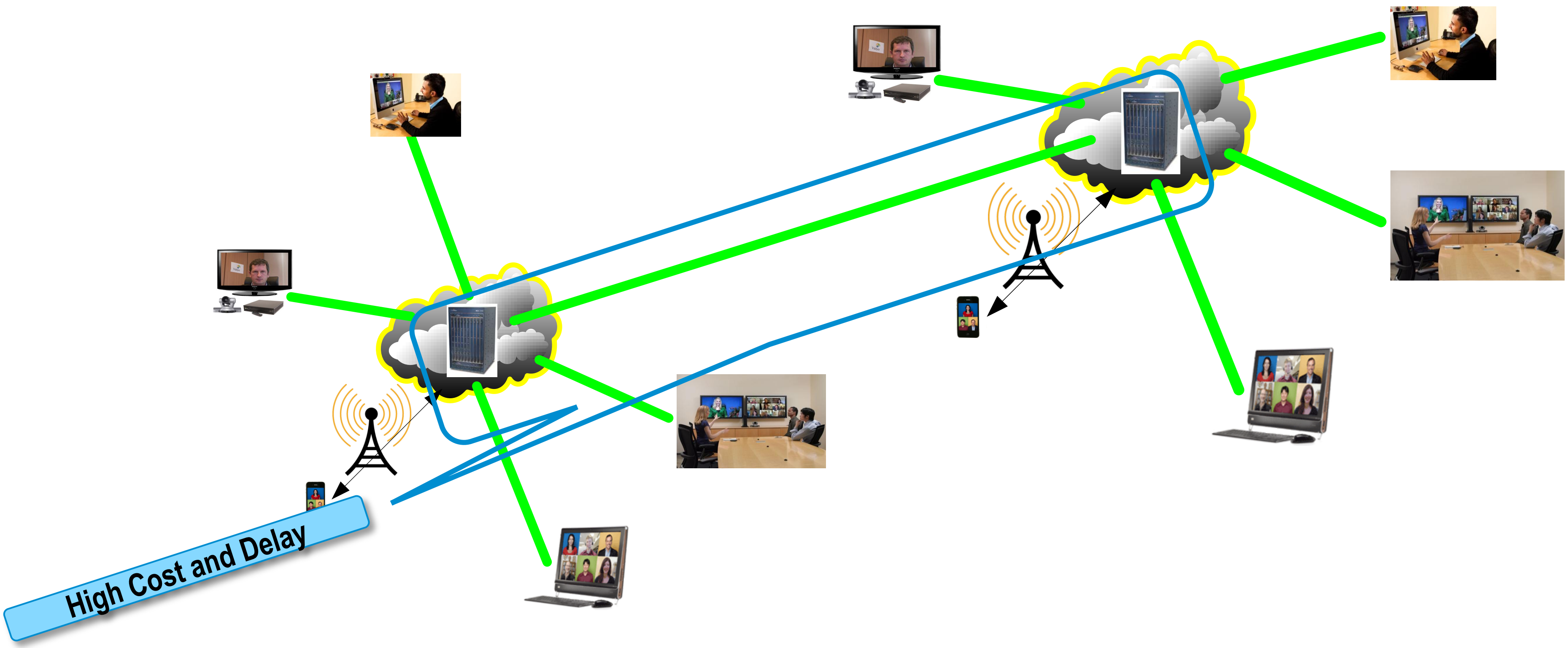
Traditional Multipoint Support (Use MCU)



The MCU-based Architecture



The MCU-based Solution Does Not Scale



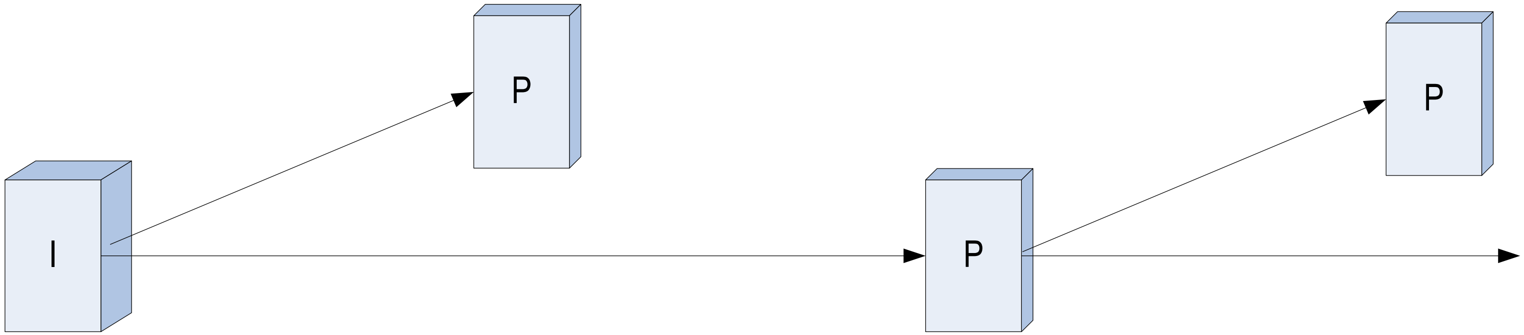
Traditional Solutions to Videoconferencing

- Dedicated networks are used – very costly
- Error resilient algorithms:
  - Forward Error Correction (FEC) – significant overhead
  - Automatic Repeat reQuest (ARQ) – increases delay
- For multipoint calls, a Multipoint Conferencing Unit (MCU) is needed for transcoding (and compositing) video
  - Introduces delay; decoding, compositing, and re-encoding time
  - Reduces video quality; repeated decoding/encoding is known to reduce video quality by 0.5-1.5 dB
  - Requires a lot of compute cycles
  - Does not scale

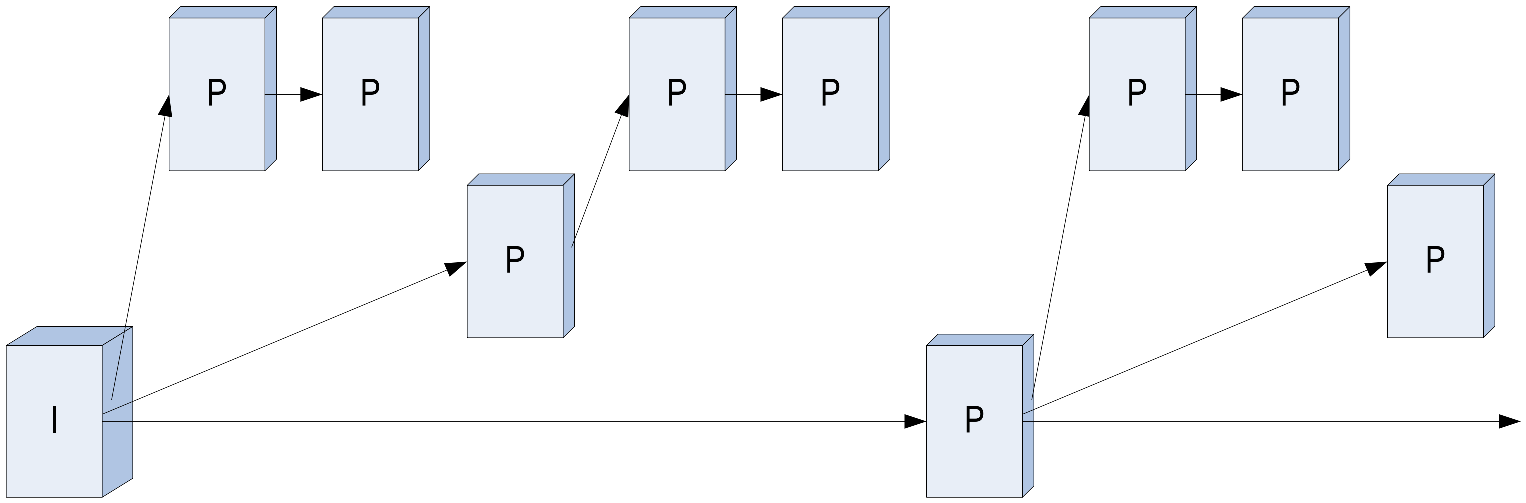
Scalable Video Coding

- Encoding of video signal with different resolution scales
- A decoder can decode selectively only part of the coded bitstream
- **Temporal scalability**: temporal subsampling
- Frequency scalability: sharpness reduction
- Quality scalability (SNR scalability): coding noise insertion (high QP)
- **Spatial scalability**: spatial subsampling
- Fine grain scalability (FGS)
- Complexity scalability
- Content related scalability: selection of content
- Object scalability

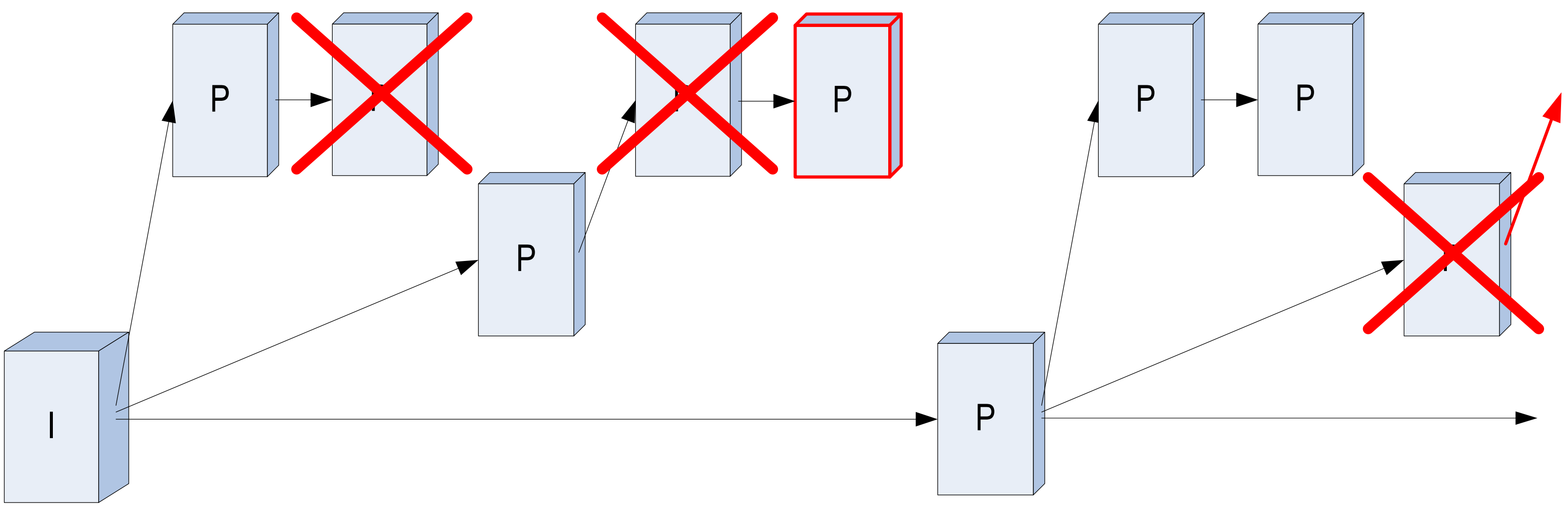
Temporal Scalability I



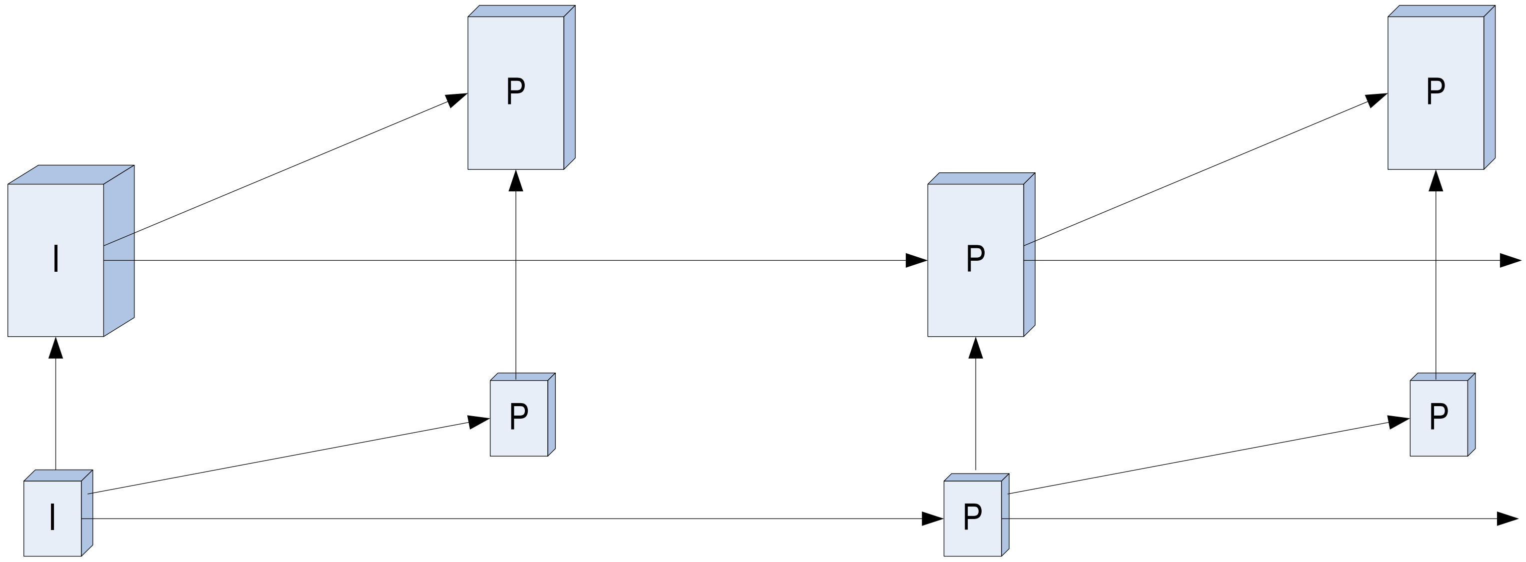
Temporal Scalability II



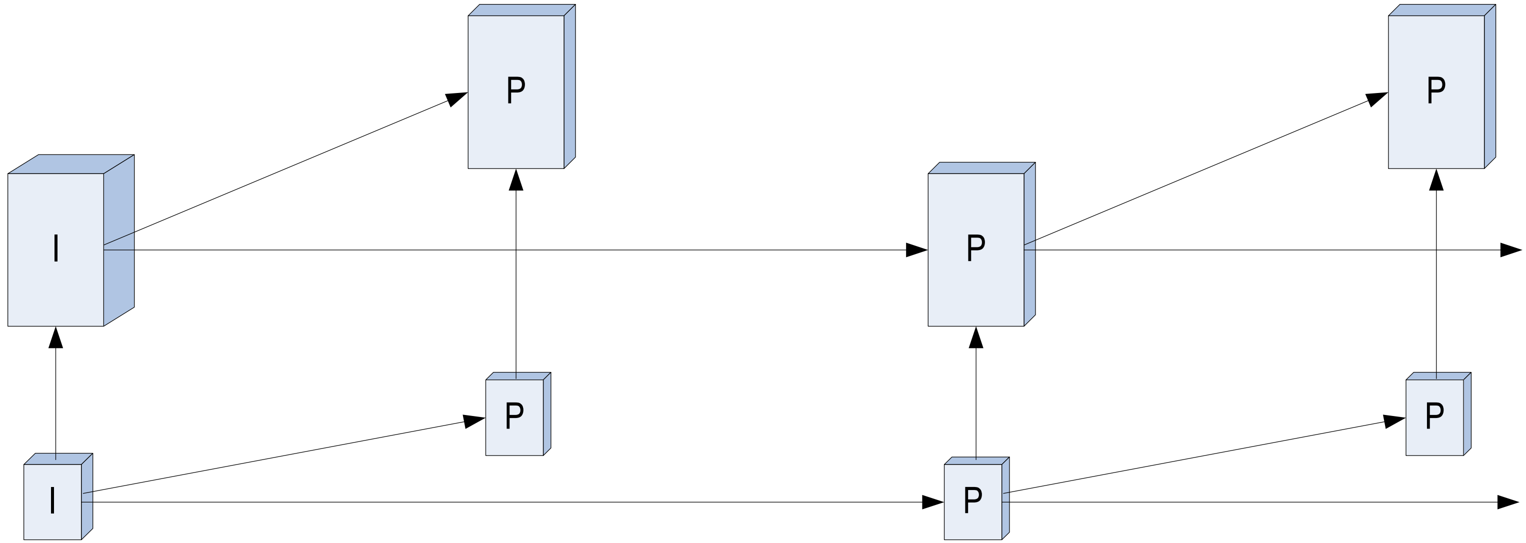
Temporal Scalability and Error Resilience



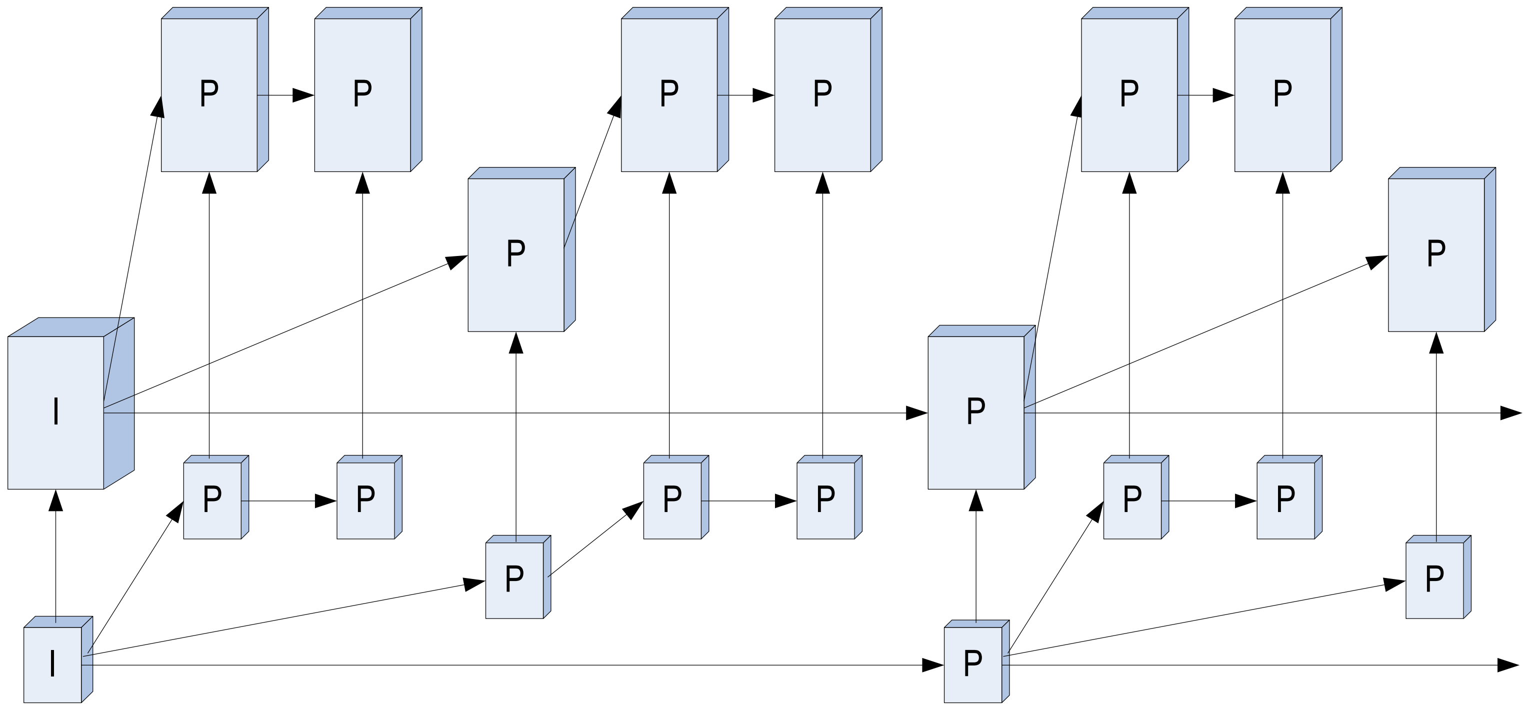
Spatial and Temporal Scalability I



Spatial and Temporal Scalability II

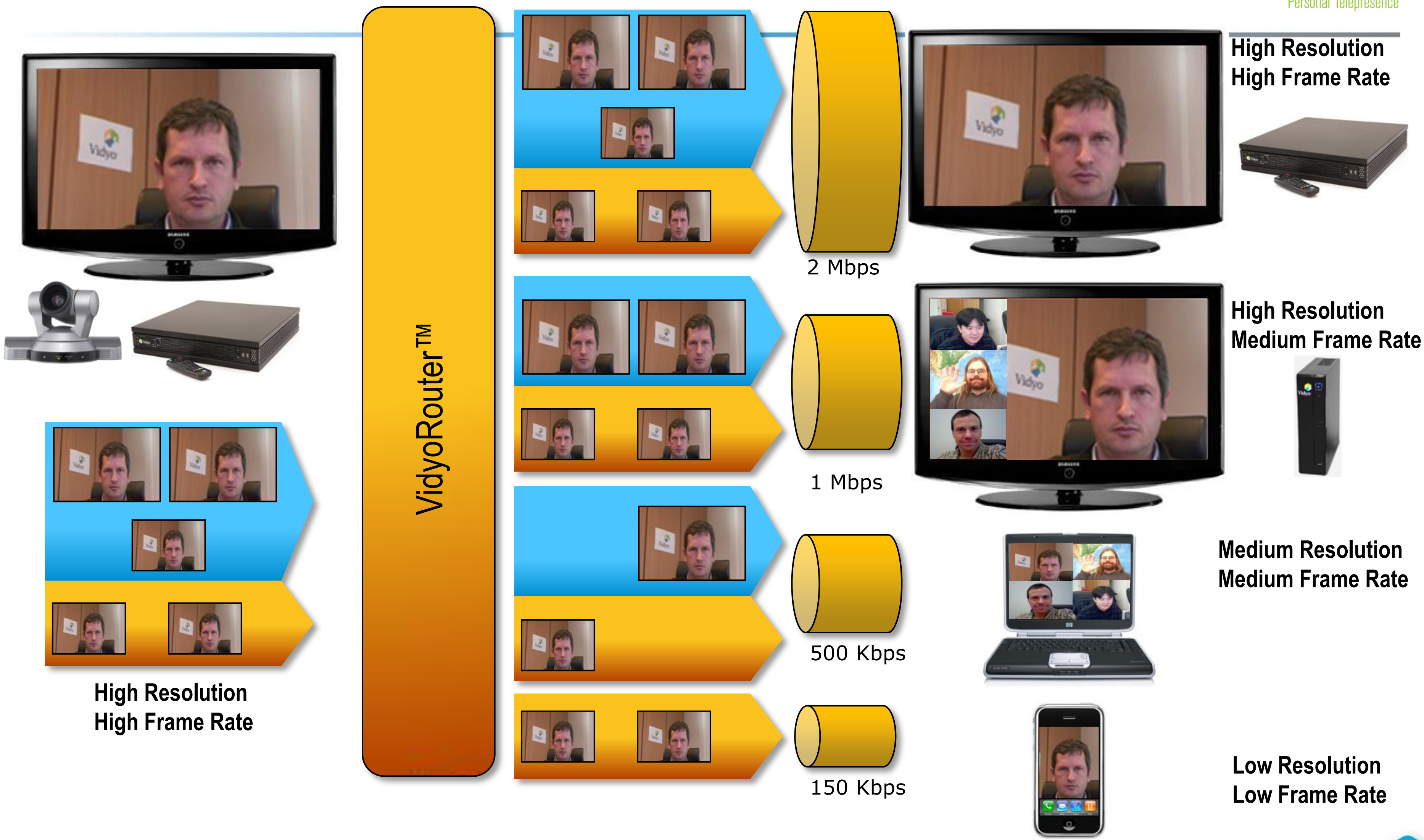


Spatial and Temporal Scalability III

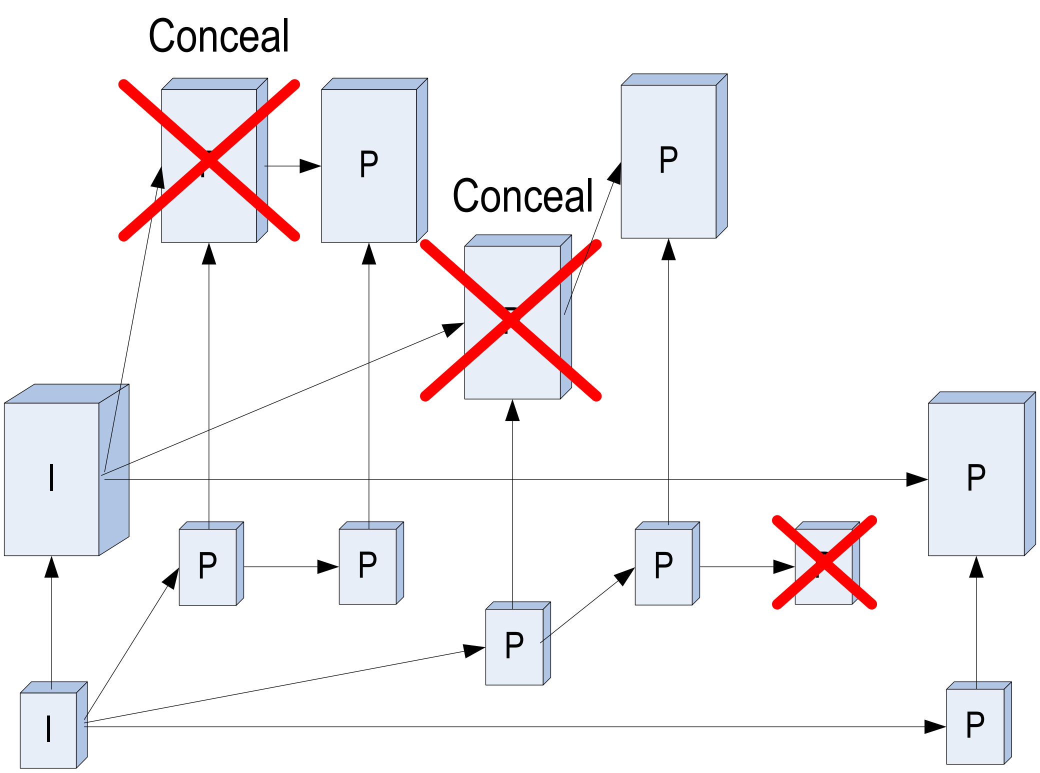


Native Rate Matching & Personal Layout

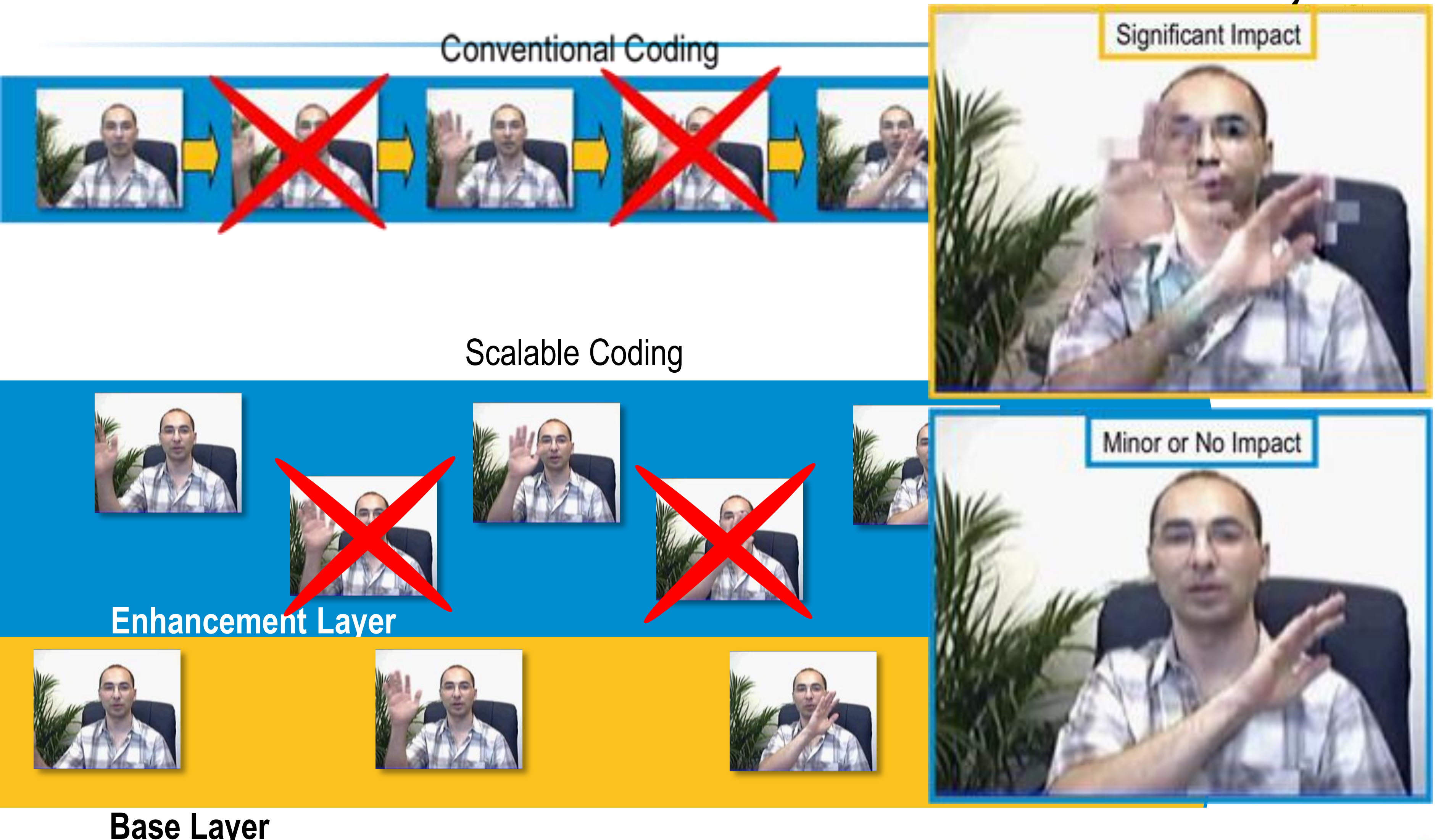
Vidyo™ Personal Telepresence



Spatial and Temporal Scalability and Error Resilience



SVC with a Proper System Eliminates Error Resilience Problem



VidyoRouter™ – A Better Multipoint Architecture

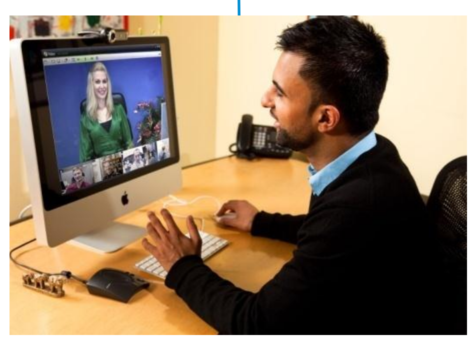
Vidyo™

ENCODE

ENCODE

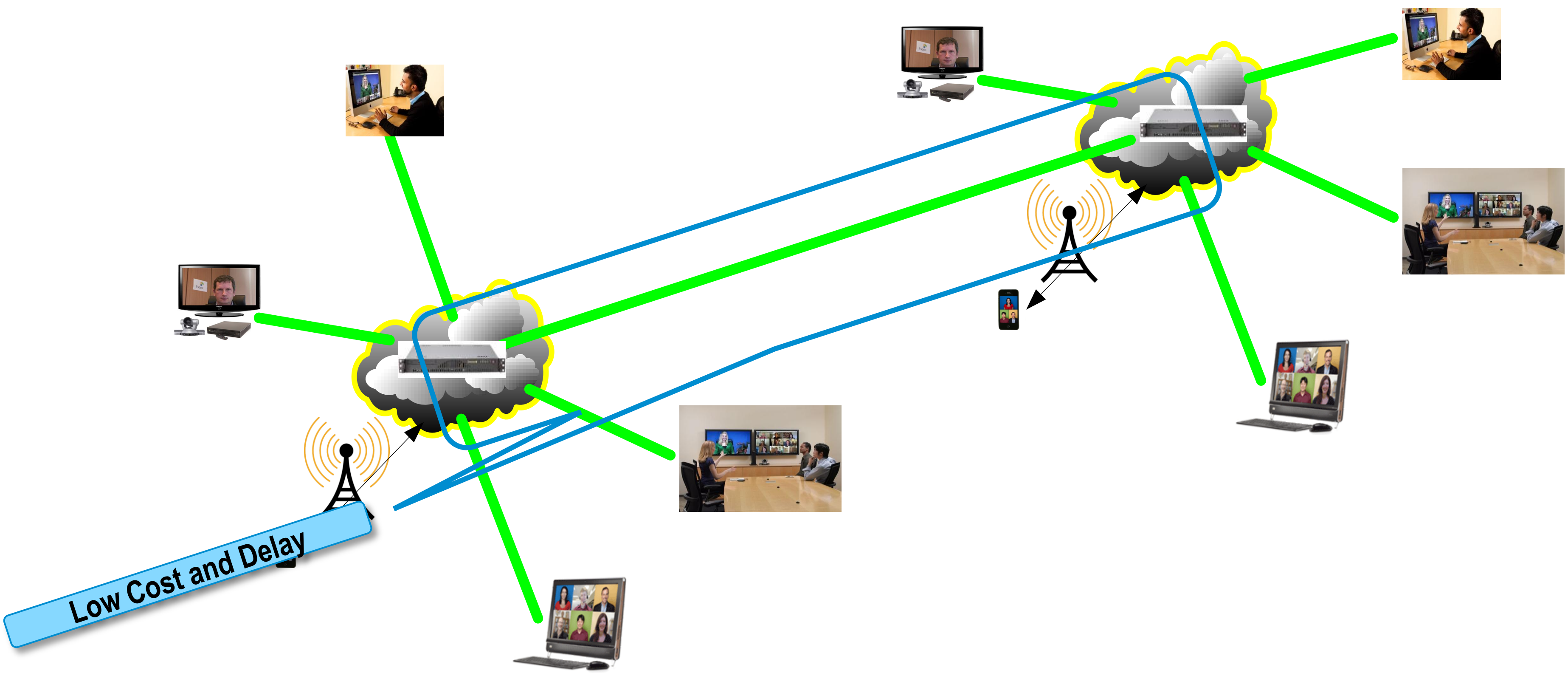
VidyoRouter™

DECODE



- **Multipoint conferencing** – WITHOUT loss/latency inducing transcoding
  - **Rate matching** – send video to multiple participants at different bandwidths
  - **Resolution matching** – send video to multiple endpoints that have different resolution capabilities
- **Error localization** – Individual client network errors do not affect other conference participants

The VidyoRouter-based Solution Scales Very Well



Scalability in Existing Video Standards

- H.261 – no scalability
- H.263 (base mode of MPEG-4) – no scalability
- H.263+ – B-frame temporal scalability
- MPEG-1 – no scalability
- MPEG-2 (H.262) – data partitioning, temporal, spatial, SNR scalability
- MPEG-4 – most flexible scalability provided: temporal, spatial, SNR, FGS, object scalability

What Is SVC?

- An extension of AVC (H.264 or MPEG-4 Video Part 10), jointly developed by ITU’s VCEG and ISO’s MPEG through a joint group called JVT
  - Annex G of AVC
- SVC lets encoding of video signal as a set of layers, where higher layers depend on lower layers
- A particular layer, together with the layers it depends on, provides the information necessary to decode the video signal at a particular fidelity
  - Fidelity is comprised of one or more of spatial resolution, temporal resolution, or signal-to-noise ratio (SNR)
- The lowest layer, base layer, is compatible with the non-scalable AVC

Why Use SVC?

- Provides a representation of the video signal that enables easy adaptation, without having to decode, process, and re-encode
  - Can change the temporal frame rate, picture resolution, quality, and bit-rate by simple elimination of parts of the encoded bitstream
  - This corresponds to simple elimination of network packets
- Offers graceful degradation – extremely resistant to packet loss
  - Can save network bandwidth by protecting only the base layer
- MCUs can be replaced by an application-level routers (VidyoRouter) that selectively forward scalable layer packets to each endpoints
  - High-end features such as rate matching and personalized layout become routing decisions, rather than transcoding
- Cascading is possible for low-delay interactive video communication

Isn't Single Layer AVC More Efficient than SVC?

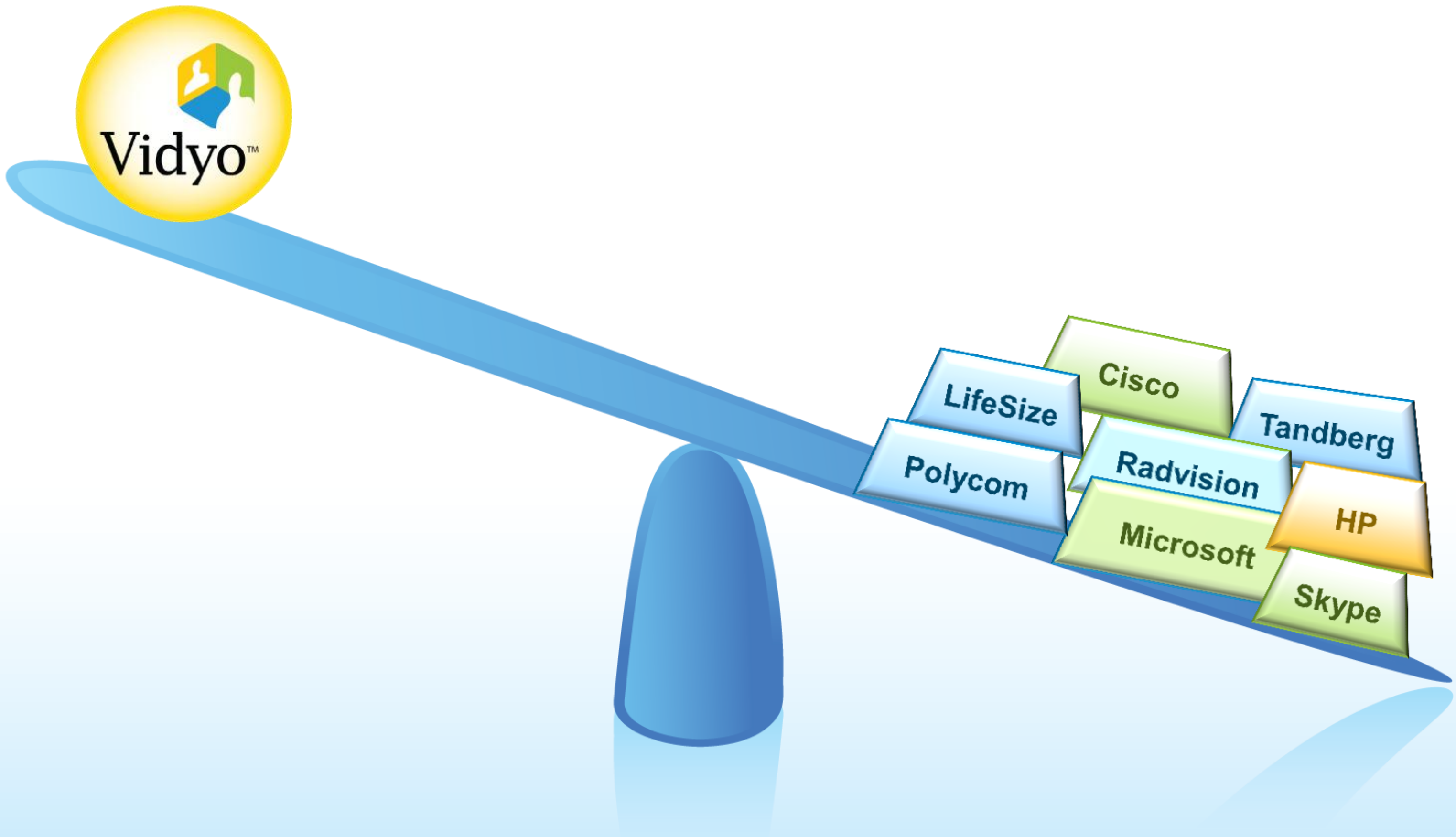
- When conditions allow (extremely high reliable network), SVC can be compacted to a single layer AVC
- In other conditions, we get low end-to-end delay and tremendous error robustness
- If we take into the quality loss caused by lost packets, then AVC is much less efficient than SVC

Conclusion

- SVC-based system brings the complexity of videoconferencing systems down to the level of any other network application
- SVC-based system brings the quality of the experience for the user to a level that is excellent even when operated on the public Internet
- End result is a new system that merges the consumer and professional videoconferencing worlds
- An effective SVC-based system CANNOT be accomplished by simply replacing a legacy AVC codec with a new SVC codec; the whole architecture must be changed to take advantage of the bitstream generated by an SVC codec
- http://www.youtube.com/user/vidyoinc
- http://www.vidyo.com

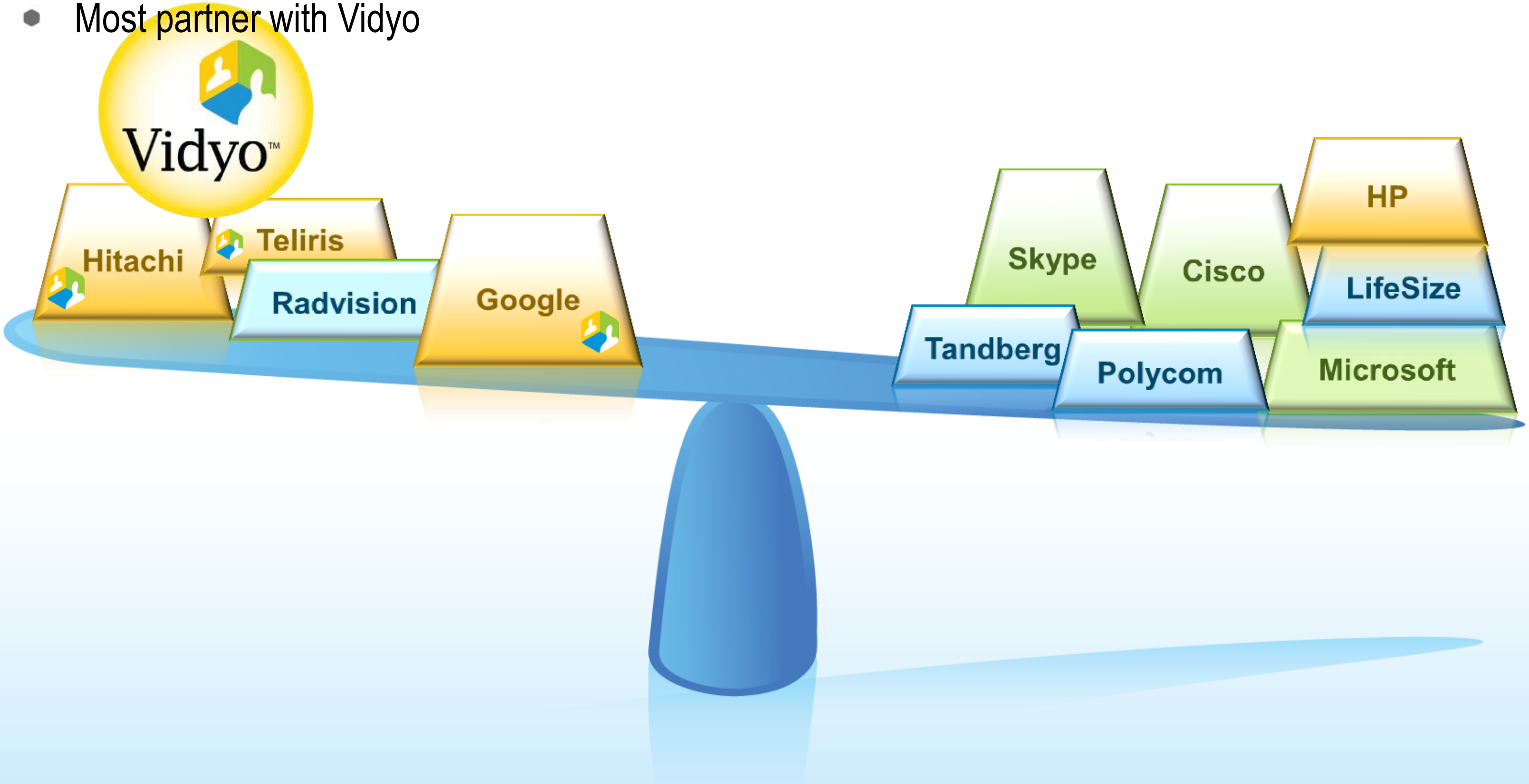
Market Adoption of SVC – 2007

Vendors Have Different Telepresence Architectures



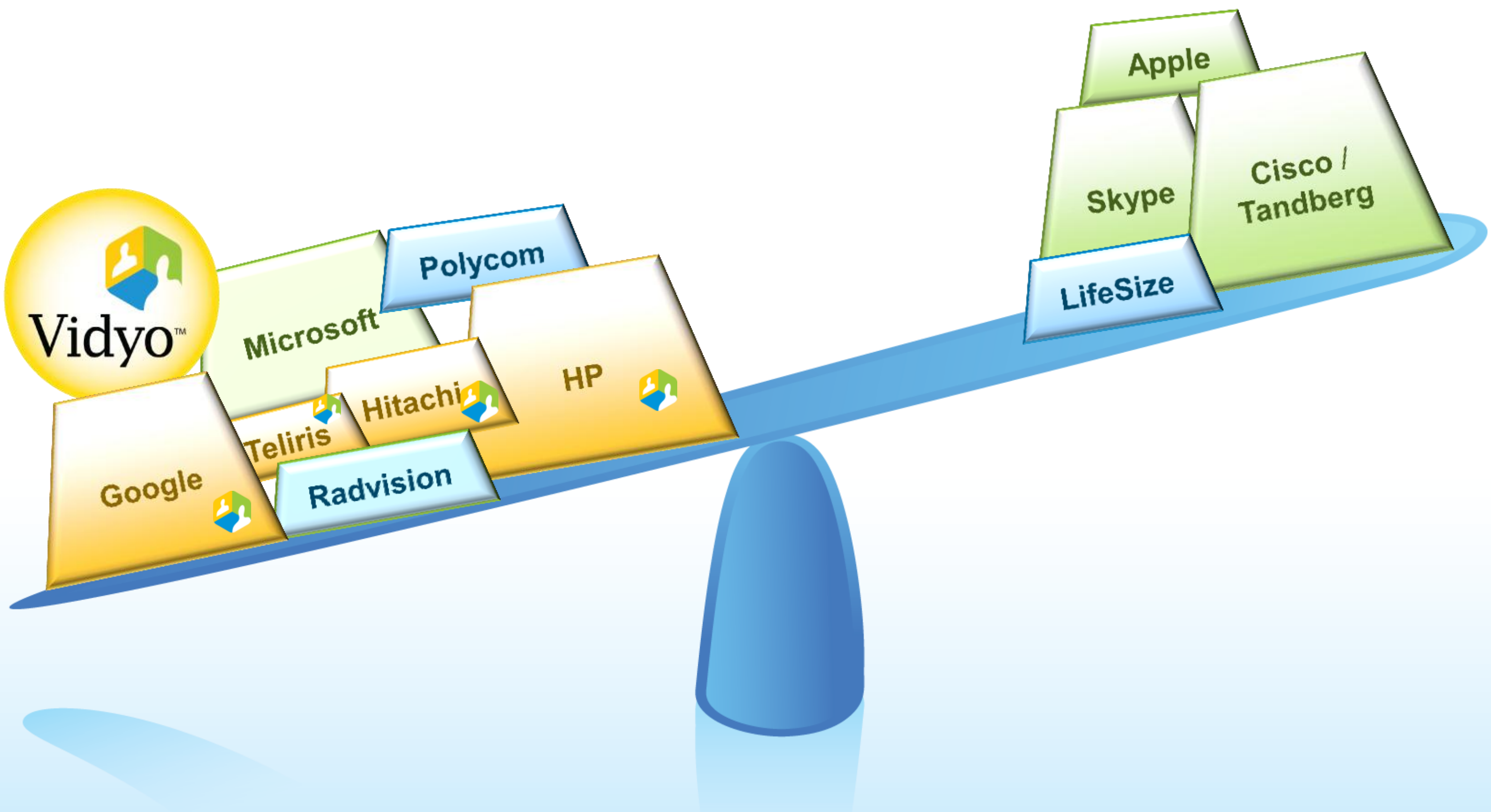
Market Adoption of SVC – 2009

- **Innovative vendors realize SVC is the route for cost effective telepresence**
  - Most partner with Vidyo



Tipping Point to SVC – Nov' 2010

Critical Mass to SVC Adoption and to Vidyo's Architecture



Applications

- Telecommuting
- Education
- Legal services
- “Tele-justice” and “tele-visits”
- Telemedicine
- Etc…