Solving Lip Sync Issues for Remotely Produced Multi-language Sports Events Over IP at Minimal Cost
Why Choose Grass Valley?

- We offer a unique solution that's unmatched in the industry
- It saves customers cost in staff travel, accommodation, network bandwidths and system complexities
- Our Media Assurance Technology combines several Grass Valley products for maximum efficiency
- Increases the quality of output (puts all video and audio tracks fully in sync)
- Fully scalable to cope with the burgeoning sports market

In the past few years, live broadcasting — specifically live sports — has fostered the growth of remote broadcasting. Though applications vary, remote broadcasting essentially relies on the use of IP connectivity to reduce production costs in a variety of ways. This includes not having to send staff on location, potentially a huge saving.

When it comes to major international sporting events such as the Olympics or international soccer championships, media companies buy the rights to air these events in real time (or close to it). Viewership is, of course, comprised of fans around the world that speak different languages, which presents a challenge for rights holders to ensure a consistent quality of experience. Often, they have to send staff to these events to commentate in a nation’s specific language, which is expensive and time consuming; there’s usually a need for additional equipment and telecoms/IP trucks. Plus, translating broadcasts into different languages can sometimes delay the production process.

However, not sending those staff has its own challenges. Take audio, for example. Sending sound and pictures from the event location — say Moscow to London — to add the live commentary, then back to their original location with the new audio track takes time. Back at the live event location, the new audio commentary track won’t be synced with the original. To line it up correctly with the original content means that the original video and effects audio need to be delayed to match — you can’t move the new audio forward in time! If you’re in a situation where you have 10 different commentary teams translating to 10 different languages, issues will undoubtedly arise. A delay in video and audio will be continuously variable due to the vagaries of IP networks, plus it will be different for each language due to the varying transmission distances and routing complexity.

Help is at hand. Grass Valley’s Media Assurance Technology makes the necessary adjustments to the video and audio signals automatically and continuously to millisecond accuracy. This means that a highly complex problem requiring very tricky continuous adjustment can be managed without human intervention, saving on equipment costs and delivering a higher quality of service without lip sync errors.

How It’s Done

The original stadium/background/effects audio is put through a Media Biometric generator to produce a signature. This signature is created using the IQSAM card — which is part of the Modular IQ family and a key component of Grass Valley’s Media Assurance technology. A low-bandwidth video (good enough to view and commentate on, saving bandwidth and costs in carriage) and a background audio track is then sent via an IP contribution network to the distant country. From there, a real-time voiceover track is created in the local language. After return to the geographic source another signature is generated at the host location using the returned audio tracks (without need to return the video). These two Media Biometric signatures are then sent to a media assurance point in the IQSAM00 card where the two signatures are compared.

Through the comparison engine, the time difference between the original and the new audio signal is calculated to provide an accurate delay between the source signals and the new voiceover signals. This delay for each and every target country and language is automatically applied to the original video and local audio prior to the video, effects and new commentary tracks being combined and sent onward to the receiving countries, entirely in sync. The continuously variable logic, calculations and change instructions required are managed by the logic engines commonly used throughout Grass Valley’s control systems.

Grass Valley products that make this possible include:

- IQSAM00 Card — Grass Valley’s modular card offering two Media Biometric generators and one assurance point on a single board
- Sirius 850 AHP Router supporting Advanced Hybrid Processing and used for audio delay adjustment for all audio tracks
- Grass Valley’s Control and Monitoring “Logic Engine” to allow operational control of audio track mapping differences between nations. This logic engine could be anywhere in Grass Valley’s control and monitoring solution where a logic engine can be hosted — contact Grass Valley for more information
- Grass Valley Control and Monitoring GUIs allowing operator monitoring and manual intervention
- IQDLY30 card — Grass Valley’s modular card allowing video delay adjustments
SOLVING LIP SYNC ISSUES FOR REMOTELY PRODUCED MULTI-LANGUAGE SPORTS EVENTS OVER IP AT MINIMAL COST

1. Event video and audio to Sirius 800 AHP, IQSAM00 generates signatures
2. Low-res video and background audio sent overseas, VO created
3. New VO and background audio sent back, another signature created
4. Signatures from two background tracks compared in IQSAM00
   Automatic adjustments calculated in logic engines and applied in IQDLY30 and Sirius 800 AHP hybrid router
5. Final video and effects and new VO sent in sync to destination