



HAL
open science

Mise en place d'une solution de visio-conférence (captation et streaming) polyvalente et modulable basée sur le protocole NDI et l'application OBS

Laurent Sengmanivanh, Philippe Girault

► To cite this version:

Laurent Sengmanivanh, Philippe Girault. Mise en place d'une solution de visio-conférence (captation et streaming) polyvalente et modulable basée sur le protocole NDI et l'application OBS. JRES (Journées réseaux de l'enseignement et de la recherche) 2024, Renater, Dec 2024, Rennes, France. hal-04894002

HAL Id: hal-04894002

<https://hal.science/hal-04894002v1>

Submitted on 17 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Mise en place d'une solution de visioconférence (captation et streaming) polyvalente et modulable basée sur le protocole NDI® et l'application OBS

Laurent Sengmanivanh

CEA DRF/IPhT

Orme des Merisiers

91191 Gif-sur-Yvette Cedex

Philippe Girault

CEA DRF/IPhT

Orme des Merisiers

91191 Gif-sur-Yvette Cedex

Résumé

À la suite de la pandémie et face au besoin de relancer la capture vidéo des cours, conférences et séminaires tenus dans notre établissement (Institut de Physique Théorique, CEA Saclay), ainsi qu'à l'obsolescence de la solution de captation précédente, nous avons dû trouver une solution alternative.

La nouvelle solution proposée par notre fournisseur n'étant plus adaptée à nos besoins, notre choix s'est porté sur une plateforme ad hoc basée sur l'utilisation du logiciel de streaming OBS et du protocole NDI. L'architecture est modulable et permet l'utilisation de périphériques non propriétaires. Sa modularité lui permet de répondre à différentes situations telles qu'avoir un orateur distant ou local et peut s'interfacer avec tout système de visioconférence (Zoom, BBB, Teams, etc.)

OBS est compatible avec le protocole NDI. Ce protocole présente de multiples avantages, le support utilisé étant l'Ethernet, la contrainte pour la longueur des câbles utilisés est moins stricte que pour des connexions USB ou HDMI. Couplé au PoE, le NDI simplifie le câblage des périphériques. De plus, grâce à la popularité d'OBS dans le monde de la captation vidéo, la prise en main et l'apprentissage de cette solution peut être très simple et rapide, même pour des utilisateurs non habitués aux outils informatiques.

Nous expliquerons dans cet article les besoins qui nous ont amenés à la solution et détaillerons ses composantes. Nous verrons enfin les futures briques qui viendront enrichir les fonctionnalités.

Mots-clefs

NDI, PTZ, Visio, Zoom, BBB, OBS, WebRTC

1 Introduction

L'IPhT (l'Institut de Physique Théorique) utilisait initialement une solution propriétaire fournie par la société Ubcast pour enregistrer ses cours, conférences et séminaires. Cette solution comprenait une station de travail pour la capture, un média serveur pour le stockage des vidéos et utilisait une caméra IP avec un flux RTSP (Real Time Streaming Protocol). Une fois montées, les vidéos enregistrées étaient transmises vers la chaîne YouTube de l'Institut^[1].

L'après Covid 19 a changé l'organisation des cours et des conférences. La capacité des salles a été réduite de moitié et les participants en distanciel souhaitent maintenir la possibilité d'interagir. La solution Ubcast ne répondait donc plus à ces nouvelles contraintes. Bien qu'un nouveau contrat avec de nouvelles technologies et une nouvelle infrastructure ait été proposé il ne correspondait plus aux besoins de l'Institut et n'a pas été reconduit.

Des visites d'autres instituts ont permis de découvrir différentes approches. L'un d'eux a choisi une solution réalisée et faisant intervenir un prestataire extérieur pour la captation et le montage post-conférence. La diffusion en direct n'étant pas une priorité pour eux, l'enregistrement était suffisant. Un autre a opté pour une solution clé en main, Zoom Rooms, qui exclue la connexion à d'autres systèmes de visioconférence. La maintenance de ces deux solutions est assurée par l'intégrateur.

N'ayant pas trouvé de solution qui permette à la fois le live streaming et une liberté dans le choix de l'outil de visioconférence, l'Institut s'est résolu à construire une solution reposant sur des logiciels open source et personnalisables.

2 Besoins de l'Institut

2.1 Interactivité

L'IPhT, un institut de recherche fondamentale du CEA, regroupe une soixantaine de chercheurs permanents et accueille de jeunes doctorants, de post-doctorants et des professeurs invités. Il organise régulièrement des conférences, des séminaires et des cours.

L'Institut a très rapidement initié l'enregistrement de ses événements pour les diffuser sur sa chaîne YouTube IPHT-TV <https://www.youtube.com/@IPhTTV>.

Après la pandémie du Covid 19, l'Institut souhaitait reprendre l'organisation de ces événements et maintenir l'interactivité de ses conférences via ses diffusions en direct.

2.2 Modularité

Les cours de l'Institut sont réalisés suivants deux modalités dans une salle disposant d'un tableau noir et de deux écrans de projection :

- En présentiel uniquement : intervention au tableau noir et/ou projection de slides.
- En hybride : visioconférence avec enregistrement des intervenants qu'ils soient présents ou à distance.

3 Choix des outils et matériels

L'étude des besoins a mis en avant les critères suivants :

- Maintenabilité : supportée par une grande communauté et compatible avec divers constructeurs ;
- Modularité : facilement adaptable à la nature de l'événement ;
- Évolutivité : possibilité d'ajouter des fonctionnalités ;
- Pérennité : durable dans le temps ;
- Économie : utilisation des COTS (*component on the shelf*) et de logiciels libres ;
- Simplicité d'usage : permettre aux chercheurs motivés d'être autonomes avec la solution.

Le choix s'est rapidement porté sur **OBS** ^[2](Open Broadcaster Software) qui est un logiciel open source, largement utilisé par les créateurs de contenus sur YouTube. La communauté derrière ce projet reste très active et propose régulièrement des évolutions.

Lors des premiers essais, la caméra RTSP a montré une latence de 2 à 3 secondes entre l'image captée et l'image transmise rendant impossible une diffusion en direct. L'usage d'une caméra USB PTZ présentait des contraintes de câblage notamment avec une longueur de câble limitée à 5 mètres et une alimentation externe. Les caméras compatibles avec le protocole **NDI®**^[3] se sont révélées plus adaptées à nos attentes ; elles permettaient une latence faible, une connexion Ethernet et une prise en charge du PoE.

Les micros de l'ancienne solution fonctionnaient bien, un micro cravate pour l'orateur et un micro main pour le président de session suffisaient. Cependant lors des conférences ou séminaires où les orateurs se succèdent, le micro cravate devait être ré-installé et testé pour chaque intervenant. Après une étude des solutions disponibles sur le marché, la barre de son **NUREVA HDL 300**^[4] a été sélectionnée pour ses performances acoustiques et techniques.

3.1 Le protocole NDI® et son implémentation

Le protocole NDI®^[5] (Network Device Interface), développé par NewTek en 2015, permet l'identification et le transport de flux vidéo de qualité non compressés avec une latence extrêmement faible sur les réseaux IP et Ethernet, ce qui le rend idéal pour les diffusions en direct. Bien que propriétaire, les spécifications du protocole NDI® restent publiques.

Reposant sur les couches TCP/IP et Ethernet, NDI® bénéficie de l'ensemble des fonctionnalités d'une application réseau et permet de s'affranchir des limitations usuelles de distance entre sources et récepteurs. Il permet d'utiliser les outils réseaux classiques tel que Wireshark pour diagnostiquer d'éventuels problèmes. Le débit moyen d'une source provenant d'une caméra NDI® Full HD étant d'environ 100 Mbps, un câblage supportant le Gigabit est requis.

Les périphériques NDI® s'annoncent sur le réseau via le protocole mDNS (multicast Domain Name System), facilitant une identification rapide des sources par les applications. Si le mDNS n'est pas possible, par exemple dans le cas où les dispositifs ne se trouvent pas sur le même VLAN, NDI® prévoit l'utilisation d'un Discovery Server qui centralise toutes les sources NDI. Enfin, il est tout à fait possible de déclarer chaque source NDI manuellement. Certaines sources NDI® autorisent également une diffusion en multicast.

NDI® simplifie la connexion entre les équipements vidéo, en s'appuyant sur les infrastructures réseaux existantes. Grâce au câblage Ethernet, il est possible de transporter simultanément les données des flux audio et vidéo, les commandes de contrôle des caméras PTZ et l'alimentation des équipements.

NDI® permet également le transport de l'audio. Cette fonctionnalité n'est pas exploitée par notre solution. Les flux vidéo n'étant pas chiffrés, un VLAN dédié a été mis en place pour le transport des flux NDI®, pour assurer la sécurité et s'affranchir des perturbations éventuelles des broadcasts sur le réseau utilisateur.

Capable de supporter des résolutions allant jusqu'à 4K, NDI® est aujourd'hui largement adopté sur les plateaux de télévision.

L'implémentation hardware du NDI® est faisable au travers de boîtiers de conversion permettant de transformer une source vidéo HDMI en flux NDI®. Un boîtier BirdDog Mini^[6] est ainsi utilisé pour raccorder les ordinateurs des intervenants qui utilisent leur propre matériel pour leur présentation.

L'implémentation software du NDI® se fait au travers de logiciel comme celui implémenté par la société Newtek qui propose une série de logiciels permettant de capturer la sortie vidéo d'un PC et la diffuser en NDI sur Ethernet : NDI Scan Converter^[7]. Ce logiciel est utilisé sur le PC de présentation mis à disposition pour les conférenciers.

3.2 OBS

OBS Studio (Open Broadcaster Software) est une solution open source de mixage des flux audio et vidéo, très populaire dans la communauté des créateurs de contenus. Elle offre de très nombreuses fonctionnalités nativement : picture in picture, transitions, ajout de texte (third lower) etc. Ses capacités peuvent être enrichies par des plug-ins et des options de configurations puissantes.

OBS découvre les sources NDI via le protocole mDNS.

La solution permet l'enregistrement des flux vidéo au format mkv et la diffusion vers des plateformes de streaming telles que YouTube. Le MP4 est également supporté mais ce format peut devenir inutilisable lors des montages si le fichier est corrompu.

OBS intègre depuis la version 26 la fonctionnalité de caméra virtuelle, permettant de créer un périphérique caméra reconnu par les logiciels de visioconférence comme une webcam. OBS permet ainsi d'envoyer le flux mixé vers ce périphérique virtuel. Deux plug-ins ont été ajoutés à OBS :

- OBS-NDI^[8] : permet d'ajouter des sources de type NDI® à une scène, ce plug-in a été récemment renommé DistroAV ;
- PTZ Controls^[9] : permet le pilotage des caméras et l'ajout de préséglages de positionnement.

OBS se révèle simple à prendre en main, une formation de 30 minutes suffit pour que les chercheurs maîtrisent les bases du logiciel. OBS permet de créer différentes scènes et de basculer d'une scène à l'autre par l'appui d'une simple touche. Les scènes de base suivantes sont proposées :

- Plan large : accueil et présentation de l'intervenant (flux NDI® camera) ;
- Blackboard : plan fixe sur le tableau (flux NDI® camera) ;
- Slides Only : affichage du bureau du PC de présentation (flux NDI®) ;
- Slides + Orateur : affichage du bureau du PC de présentation (flux NDI®) + recadrage du flux caméra.

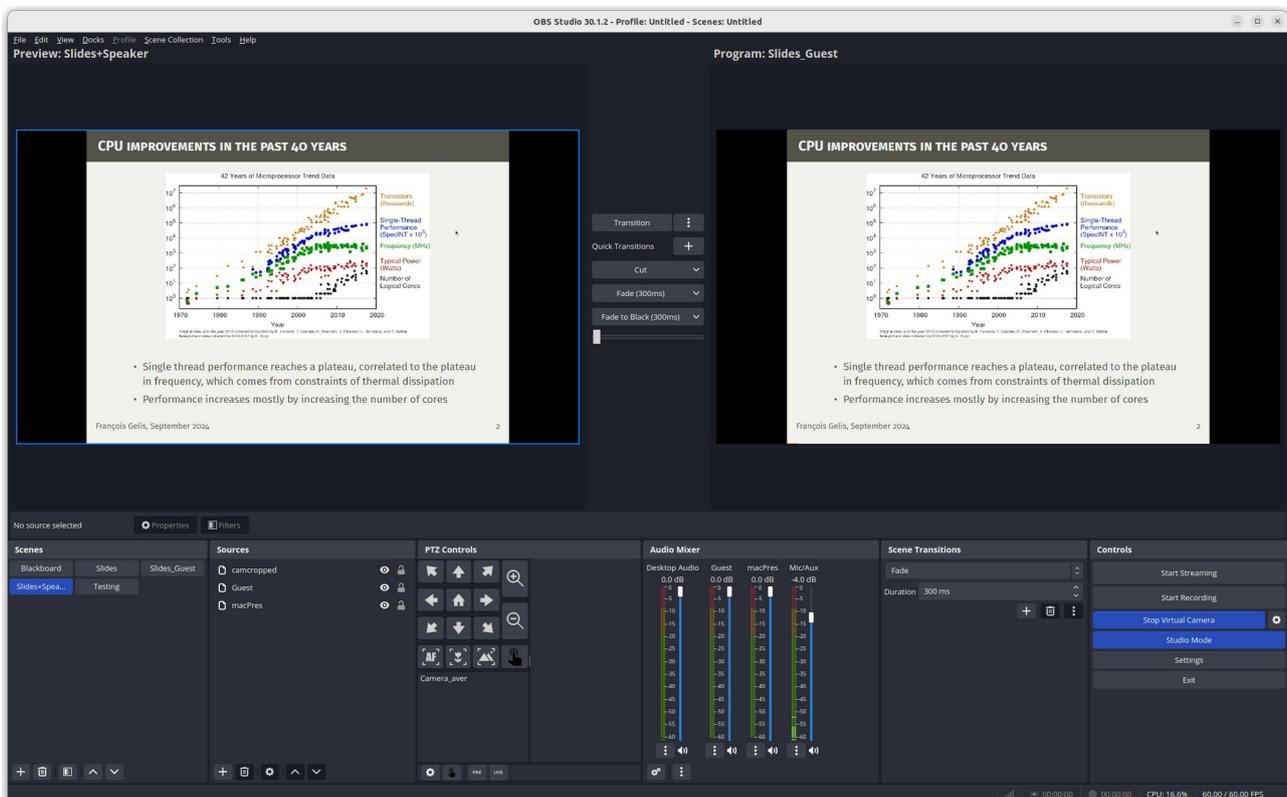


Figure 1: Exemple d'un écran OBS lors d'une conférence Multiview Screen OBS

Sur la Figure 1, la partie gauche de l'écran OBS affiche une prévisualisation. La partie droite correspond au flux transmis.

OBS intègre nativement la fonctionnalité de diffusion en direct sur les plateformes vidéo telles que YouTube.

3.3 Caméra PTZ avec auto-tracking

Une caméra PTZ de marque AVer PTC310HN^[10] dotée de l'auto-tracking de personnes par analyse vidéo, permet de suivre automatiquement l'intervenant, éliminant le besoin d'un opérateur. Ce choix a été réalisé après des tests concluants dans l'Institut. Le tarif proposé était comparable aux autres caméras dotées des fonctionnalités similaires.

3.4 Poste utilisé pour la présentation

Pour faciliter l'utilisation de la solution, un poste de présentation avec tous les outils pré-configurés est mis à disposition de l'orateur qui n'a plus qu'à copier ses documents sur le poste. Afin de supporter un large éventail de formats de fichier (Word, Powerpoint, Excel, Keynote, PDF, Jupyter Notebook, NoteBook Mathematica, Libre Office, etc.) un Apple Macintosh est utilisé.

Le logiciel NDI Scan Converter développé par la société Newtek, installé sur le poste de présentation permet la capture de son affichage et sa conversion en une source NDI®.

L'avantage d'une machine dédiée à la présentation est de réduire le temps lors des changements de conférenciers tout en évitant les problèmes classiques de connectivité (adaptateur USB-C vers HDMI, protection HDCP, configuration de l'affichage, etc.).

Une tablette tactile dotée d'un stylet est mise à disposition des conférenciers. Elle est destinée à remplacer le tableau noir lorsque des explications complémentaires manuscrites sont données. Elle permet au conférencier de rester face à son auditoire. De plus, la captation directe du tableau noir par la caméra peut parfois présenter des reflets gênants pour l'auditoire distant. Malheureusement, peu de conférenciers ont adopté ce dispositif car beaucoup ont conservé le réflexe de prendre la craie !

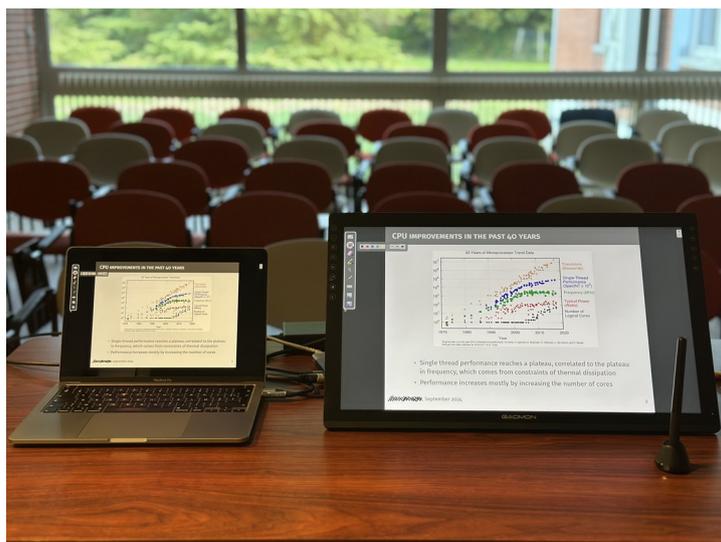


Figure 2: Mac de présentation et tablette tactile

L'utilisation d'un Apple Macintosh permet aux conférenciers de projeter leur présentation depuis une tablette iPad grâce à la fonction AirPlay.

Enfin, pour les intervenants souhaitant utiliser leur propre ordinateur, une connexion HDMI est mise à disposition. Cette dernière est raccordée à un boîtier convertisseur HDMI vers NDI® BirdDog Mini. Cette solution permet d'uniformiser les sources vidéo dans OBS.

3.5 Capture audio

Le NDI supporte également les flux audio, mais les équipements compatibles restent onéreux.

Notre solution audio initiale, basée sur un micro cravate pour l'intervenant et un micro main pour les questions de l'auditoire, n'a pas donné satisfaction. La spontanéité des échanges, permise avec un auditoire réduit, était perdue du fait du délai requis pour passer le microphone.

L'étude et les tests de quelques produits du marché ont conduit à choisir la barre de son HDL300 de la société NUREVA.

Cette barre de son simule 8192 micros virtuels répartis dans une matrice en 3 dimensions (Microphone Mist Technology). Elle intègre également des hauts-parleurs pour diffuser le son. Les questions des personnes en distanciel sont également retransmises dans la salle.

Cette solution s'est révélée pleinement satisfaisante et facilite le démarrage des conférences puisqu'il n'est plus nécessaire d'équiper le conférencier d'un microphone.

L'installation de la barre est très simple, il suffit d'une connectique USB pour la relier au PC de captation. Un logiciel de calibrage est fourni pour l'initialiser avant la première utilisation.

4 Mise en place de la solution

4.1 Organisation de la salle

La figure ci-dessous illustre l'agencement des principaux éléments de la solution.

Lors d'un événement, l'organisateur s'installe devant le PC OBS afin de surveiller le bon déroulement de la diffusion vidéo.

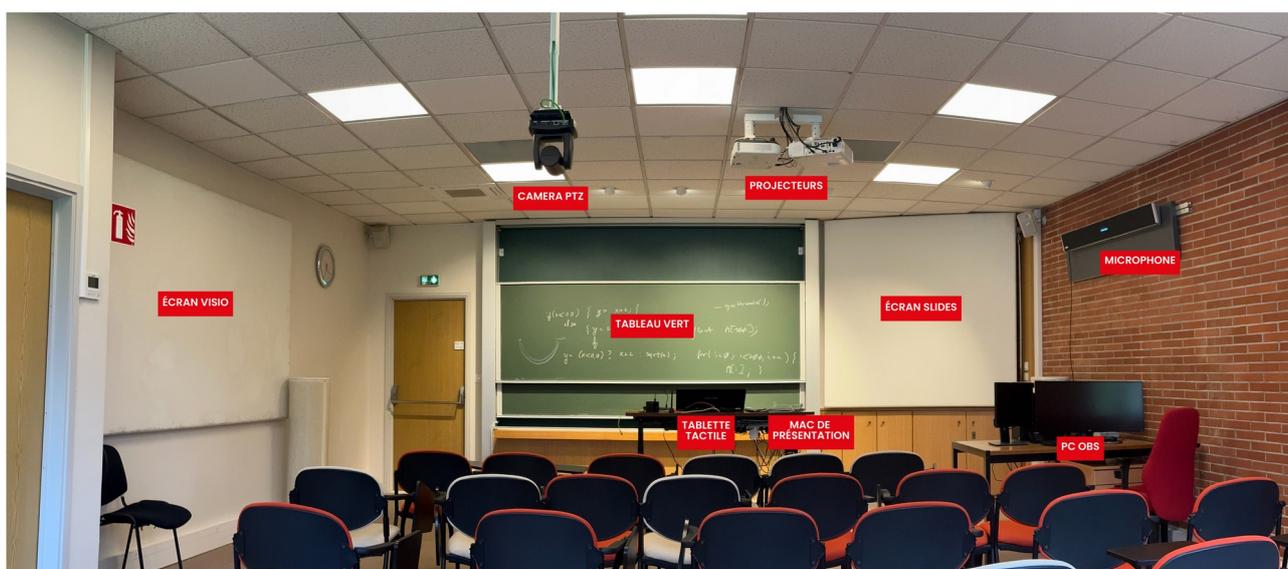


Figure 3: Salle de cours

4.2 Mise en place et fonctionnement

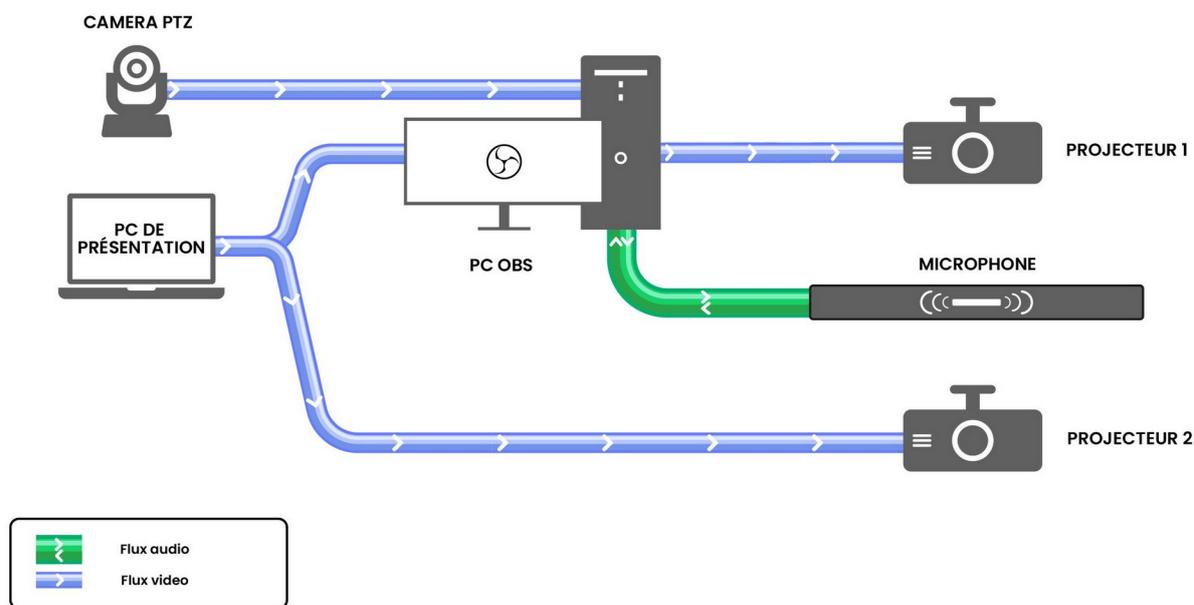


Figure 4: Schéma des flux audio et vidéo

Lors d'une conférence, une à deux sources sont utilisées et transmises au PC OBS (Figure 4) :

- Caméra PTZ : filme le conférencier (flux NDI) ;
- Slides : provenant du Mac de présentation et transmis simultanément au projecteur 2 en HDMI et au PC OBS via NDI avec le logiciel *Scan Converter*.

Cette solution permet une utilisation basique : capture et projection des slides lorsque l'organisateur ne souhaite pas utiliser l'ensemble du dispositif.

La salle virtuelle de la visioconférence, ouverte sur le PC OBS, est affichée sur l'écran visio (Figure 3) via le projecteur 1.

À l'initiation de la session, l'organisateur se connecte sur le PC OBS et démarre les composants nécessaires à l'aide d'un *stream deck* ^[11] (Figure 5). Depuis l'interface d'OBS, il sélectionne la scène à diffuser aux participants distants à partir d'une prévisualisation, il peut ainsi choisir celle la plus adaptée : slides, conférencier ou tableau noir. Pour la diffusion du flux sur la chaîne YouTube, une clé de transmission est requise, elle est saisie une seule fois dans OBS.

L'activation de la fonction caméra virtuelle d'OBS permet d'utiliser le flux de la scène sélectionnée comme source vidéo pour les outils de visioconférence. L'audio quant à lui (émission/réception) est assuré par la barre de son NUREVA. Les caméras des participants distants sont affichés via le second projecteur visio à droite du conférencier (Figure 3), lui permettant ainsi de voir l'auditoire local face à lui et l'auditoire distant à sa droite.

L'ensemble de la session est enregistré et, après un montage succinct (suppression des temps morts du début, de la pause et de la fin de session, etc.), il est mis en ligne sur la chaîne YouTube.

Le *stream deck* (Figure 5) facilite la phase de démarrage du dispositif (caméra, vidéo projecteurs, convertisseur HDMI/NDI, OBS, Zoom, BBB, etc.). Chaque bouton du *stream deck* ^[12] est configuré pour lancer un script ad hoc, par exemple une commande SNMP ou Pjlink.

Exemple de commande SNMP :

```
#!/bin/sh
# Demarre Camera AVER
# Port 3 du switch PoE
/usr/bin/snmpset -m /var/lib/snmp/mibs/ietf/POWER-ETHERNET-MIB -v2c -c <rw_community> <ip_switch>
1.3.6.1.2.1.105.1.1.1.3.1.3 i 1
```



Figure 5: Stream deck

D'autres boutons peuvent être configurés pour le pilotage d'OBS (démarrer/stopper l'enregistrement, sélectionner la scène, etc.). Des logos simples et intuitifs sont implémentés pour rappeler la fonctionnalité de chaque bouton du *stream deck*.

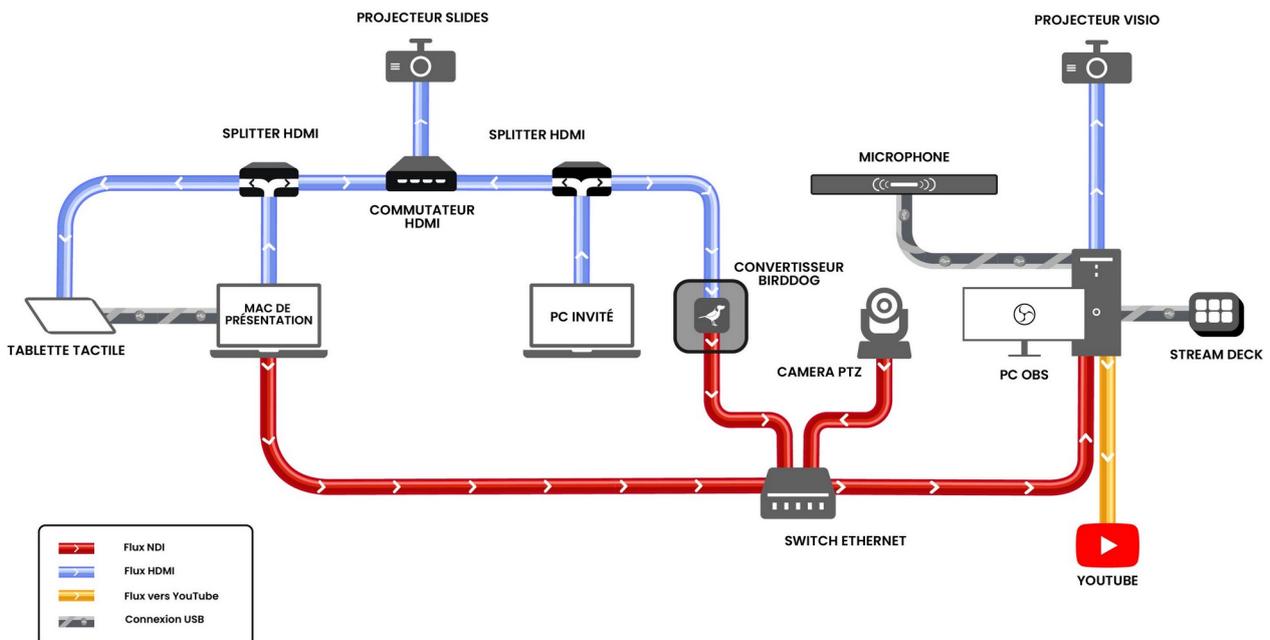


Figure 6: Synoptique de la solution

4.3 Contrainte de l'intégration des produits

La solution n'était pas optimale à sa mise en place, mais des améliorations ont été apportées et continuent de l'être. Par exemple, si la session Zoom était démarrée avant l'activation de la caméra virtuelle sur OBS alors les API WebRTC du navigateur web ne reconnaissaient pas la caméra virtuelle, et il fallait alors relancer la session pour voir réapparaître la caméra. Pour corriger ce désagrément, il suffit de configurer l'activation automatique de la caméra virtuelle au lancement d'OBS. Comme toute nouvelle solution, nous avons procédé à de multiples réglages et corrections afin d'avoir le résultat souhaité. C'est un problème inhérent à l'utilisation de produits qui n'ont pas été configurés pour fonctionner ensemble.

5 Déploiement de la solution dans l'amphithéâtre



Figure 7: l'amphithéâtre

La solution mise en place dans cette salle ayant donné entière satisfaction, l'amphithéâtre de 150 places est actuellement en cours d'aménagement avec un dispositif audio-vidéo reposant sur les mêmes principes : caméra PTZ avec auto-tracking, NDI®, *stream deck*, OBS.

Deux caméras PTZ sont utilisées, une avec suivi automatique pour le conférencier (AVer PTC310HN), l'autre (Birddog P100^[13]) pour filmer le public. Ce dispositif permet aux personnes à distance de voir également le public lors des séances de questions-réponses. L'installation de caméras spécifiques est également envisagée pour filmer un jury de thèse.

Pour la partie audio, une solution basée sur des micros cravate et des micros main sans fil est utilisée. Un micro cube molletonné^[14] (Figure 8) lançable pour les questions a été ajouté, ce qui réduit les délais lors de la prise de parole pour les questions et ajoute une dimension « ludique ».



Figure 8: micro cube molletonné lançable personnalisé aux couleurs de l'Institut

Une table de mixage analogique est utilisée pour gérer les différentes sources audio.

6 Évolutions possibles

La solution mise en place offre une grande modularité. L'ajout d'extension et l'usage de protocoles standards permettent d'enrichir le dispositif avec de nouvelles fonctionnalités. Très prochainement, il est envisagé de compléter la solution de l'amphithéâtre par :

- l'ajout d'une caméra pour filmer le jury local lors de soutenances de thèse en mode hybride ;
- la personnalisation des boutons du *stream deck* pour le basculement des scènes d'OBS ;
- l'utilisation de websocket pour le pilotage d'OBS par une tablette ou un smartphone ;
- l'ajout d'un agenda dynamique lors des conférences (digital signage) grâce à l'application Ontime^[15]. Les modèles de visualisation proposés par cette application accessibles via HTTP peuvent être intégrés comme source dans OBS et être utilisés pour afficher automatiquement et dynamiquement le nom du conférencier.

NDI transportant facilement des flux audio et vidéo, il serait tout à fait possible d'utiliser des caméras sans fil. Des applications permettent d'utiliser la caméra d'un smartphone comme source NDI, pour des captations hors des salles équipées, par exemple pour filmer des expériences scientifiques en extérieur.

Un prompteur pourrait également être installé pour les discours.

7 Conclusion

Grâce à l'utilisation de technologies robustes et efficaces (caméra PTZ auto-tracking d'AVer, micro ambiant de NUREVA, OBS, NDI® de Newtek), les premiers tests ont pu démarrer et la solution a été mise en place rapidement. Ces technologies reposent sur un socle informatique bien connu des administrateurs systèmes et réseaux, permettant, en cas de problèmes, le diagnostic par l'analyse des logs ou des trames réseaux. Du fait de l'intégration de la solution par l'Institut, cela a permis la connaissance de la solution de bout en bout. La liberté dans nos choix a permis de construire une solution agnostique, compatible avec une grande variété de dispositif de visioconférence et de constructeurs.

8 Remerciements

Nous remercions la direction de l'IPhT pour la confiance qu'elle nous a accordée et le soutien sans faille qu'elle nous a apporté tout au long du projet.

Nous tenons à remercier très chaleureusement Riccardo, un chercheur exigeant, perfectionniste, et enthousiaste pour les nombreux échanges sur la solution et sa forte implication pour les tests.

Nous remercions aussi Anne, Dorothée, Julien, Lukas pour leurs remarques.

Et enfin, un grand merci aux JRES pour avoir accepté notre poster et à Laurence et Olivier pour la relecture et les nombreuses corrections de l'article.

Références

- 1: <https://www.youtube.com/@IPhTTV>
- 2: <https://obsproject.com/>
- 3: <https://docs.ndi.video/docs/white-paper/introduction>
- 4: <https://www.nureva.com/audio-conferencing/hdl300>
- 5: https://www.mylumens.com/fr/Blog_detail/117/NDI®-Transmission-Protocol
- 6: <https://birddog.tv/mini-overview/>
- 7: <https://ndi.video/tools/>
- 8: <https://github.com/DistroAV/DistroAV>
- 9: <https://github.com/glikely/obs-ptz>
- 10: <https://presentation.aver.com/model/ptc310hn>
- 11: <https://www.elgato.com/fr/fr/p/stream-deck-mk2-black>
- 12: <https://timothycrosley.github.io/streamdeck-ui/>
- 13: <https://birddog.tv/p100-overview/>
- 14: <https://www.catchbox.com>
- 15: <https://github.com/cpvalente/ontime>