La vidéo sur IP en production « live », un univers composite

La vidéo sur IP est devenue une expression banale dans le monde des communications numériques. Dès qu'un flux d'images animées est transporté par un réseau IP, il s'agit de vidéo sur IP. Cette terminologie recouvre une infinité d'usages et de situations, allant de la vidéosurveillance à la production live en passant par le streaming, la visioconférence, la messagerie instantanée, la webTV, l'OTT… et cette énumération est loin d'être exhaustive.

Par Pierre-Antoine Taufour

En réduisant le champ de vision, pour se focaliser sur la production vidéo en direct, nous constatons chaque jour à travers nos contacts avec des fournisseurs, des utilisateurs ou des prestataires que le vocabulaire est loin d'être clair et que derrière cette expression chacun propose sa vision personnelle. D'autant que la dénomination des produits, des standards techniques et des architectures est encore loin d'être stabilisée.

Trois technologies distinctes

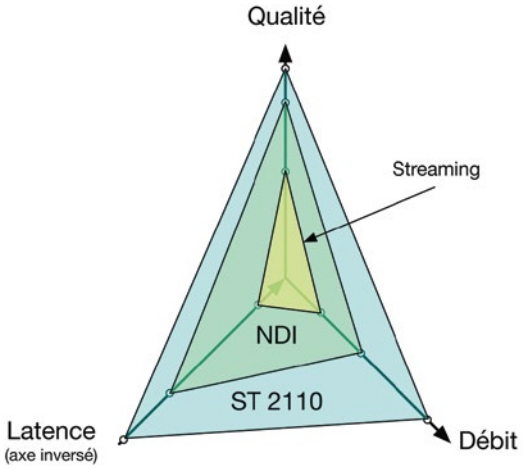
Cette évolution technique, plutôt cette révolution, s'appuie sur le marché existant et sur son découpage en plusieurs gammes ou secteurs bien ancrés dans l'esprit des utilisateurs. Cette segmentation conduit très souvent à une division des catalogues en trois gammes : les équipements très performants pour lesquels aucun compromis sur la qualité n'est toléré, une gamme intermédiaire avec des caractéristiques d'un bon niveau de manière à proposer des produits performants à des coûts conformes avec l'économie de la production et enfin l'entrée de gamme pour lesquels des compromis techniques plus sévères amènent à serrer les prix pour correspondre aux besoins de la production légère ou du marché « corporate ». Pour éviter d'entrer dans des débats interminables, nous laissons le lecteur effectuer sa propre classification selon les moyens qu'il utilise quotidiennement ou ceux dont il rêve en secret.

Pour convertir une régie vidéo au réseau IP, il ne suffit pas de remplacer les coaxiaux SDI par du câble RJ-45 ou de la fibre optique. Les modifications sont beaucoup plus conséquentes et bouleversent toute l'architecture technique.
© Adobe Stock xiaoliangge

Lorsque nous examinons les offres techniques de vidéo sur IP conçues pour un environnement de production, nous constatons qu'elles se répartissent en trois catégories, à la fois en termes de performances et de coûts, mais aussi d'usages. Pour remplacer le câblage SDI d'une régie vidéo traditionnelle par une architecture vidéo sur IP, il y a actuellement trois propositions technologiques qui ne sont pas réellement en concurrence et qui correspondent chacune à l'un de ces trois niveaux de segmentation du marché. Il s'agit, par ordre décroissant de performances et donc de coût, des équipements basés sur le standard SMPTE ST 2110 associé aux spécifications NMOS, puis des infrastructures exploitant le codage NDI de NewTek, et enfin d'outils (ou plutôt de logiciels) exploitant les technologies de streaming en complément des entrées/sorties SDI et ou HDMI.

Paramètre	Technologies vidéo/IP		
	SMPTE ST 2110	NDI large bande NewTek	Techniques de streaming
Qualité image	Excellente	Très bonne	Bonne
Débit du flux vidéo	1 à 10 Gb/s	90 à 350 Mb/s	2 à 20 Mb/s
Type de port réseau	10 à 100 Gb/s	1 ou 10 Gb/s	Gigabit (WiFi)
Latence	Quelques lignes	Une image	de 4 à 20 images

Les performances d'une technologie de vidéo sur IP résultent toujours d'un compromis entre la qualité des images, la latence et le débit numérique. Chaque industriel ou organisme de normalisation effectue ses choix en fonction des usages et du marché visé.
© PA Taufour



Pour chacune des trois technologies de vidéo sur IP évoquées dans l'article, la combinaison des trois paramètres montre que leurs performances sont fort diverses et ne couvrent pas les mêmes besoins. L'échelle sur les trois axes n'est pas linéaire.
© PA Taufour

Le standard ST 2110 associé au NMOS

Dès 2007, la SMPTE a publié le standard ST 2022 qui explique comment transporter un signal vidéo non compressé sur un réseau IP. Conçu dans une perspective de liaison de contribution et de distribution, le regroupement des essences (vidéo, audio, données auxiliaires) s'est révélé inadapté aux situations de production. Elle a donc remis son ouvrage sur le métier pour proposer en 2017 le standard ST 2110 mieux adapté aux usages de production. Basé sur un premier travail mené par le VSF (Video Services Forum) qui réunit des utilisateurs et des prestataires, il a pour objectif de définir les spécifications du transport, de la synchronisation et la description des données vidéo et audio en flux séparés dans un environnement de production temps réel. Mais comme cela est détaillé dans l'article suivant à ce sujet, le ST 2110 ne suffit pas à décrire la totalité des fonctionnalités indispensables à la mise en œuvre d'un équipement de production dans une architecture IP. Le ST 2110 est donc complété par les spécifications NMOS développées dans le cadre de l'AMWA (Advanced Media Workflow Association). D'autres organismes ont été associés à ces travaux et sans vouloir leur faire offense on ne peut pas dire que la communication ou le marketing soit au cœur de leurs préoccupations. La conséquence concrète est qu'il n'y a aucun sigle ou dénomination générique explicite pour désigner l'ensemble de ces technologies et standards. Par simplification, nous employons le terme ST 2110, mais ce vocabulaire est trop restrictif. La désignation complète du standard est d'ailleurs « Professional Media over Managed IP Networks », ce qui est loin d'être évident à glisser dans la conversation.

Le NDI, une dénomination claire pour un écosystème en croissance

NewTek est beaucoup plus clair avec la technologie de vidéo sur IP qu'il a développée. Elle s'appelle tout simplement NDI. Elle recouvre de nombreux aspects techniques et ne se limite pas uniquement aux mélangeurs TriCaster de la marque. Dès le départ le constructeur a mis à disposition de l'industrie les spécifications du NDI. Une multitude de produits et de logiciels sont basés sur ce codage et ce protocole, et c'est un véritable écosystème qui s'est développé en quelques années autour de nouveaux modes de transport de la vidéo sur réseau dans un environnement de production. Nous en détaillons les caractéristiques plus loin dans ce dossier et relatons l'avis d'utilisateurs.

Les techniques de streaming sont également de la partie

Les technologies de streaming ont d'abord été développées pour la diffusion de contenus audio et vidéo via Internet, à la fois en VOD, mais aussi en live. Sur un plateau, elles n'apportent pas d'avantage décisif, d'autant qu'elles sont basées sur une forte compression qui induit une latence non négligeable. Par contre pour relier une source éloignée ou utiliser un smartphone comme caméra supplémentaire, elles apportent une solution économique et simple à mettre en œuvre du fait de leur compatibilité avec Internet. Elles trouvent également leur place dans de nombreuses architectures de remote production, à condition de choisir un codec et un protocole de transport maîtrisant la latence. Les premières solutions basées sur le RTSP ou le RTMP laissent la place à des outils mieux adaptés comme le SRT ou le WebRTC ou d'autres solutions propriétaires.

Sans oublier les logiciels de mixage vidéo

Les technologies de streaming sont privilégiées aussi par toutes les solutions logicielles de mixage vidéo qui se sont développées ces dernières années. Même si les logiciels mélangeurs vidéo ne sont pas directement liés au transport de la vidéo sur IP, puisque la plupart reçoivent les signaux vidéo par des cartes SDI ou HDMI ou via des boîtiers de capture, les ordinateurs sur lesquels ils tournent sont tous munis d'une ou plusieurs interfaces réseau qui offrent une entrée privilégiée pour des flux de vidéo sur IP, qu'ils soient ST 2110, NDI ou codés en streaming. Ces logiciels dont les plus connus sont OBS (open broadcaster software), vMix, Studio 6 de Livestream ou encore Wirecast de Telestream proposent une ergonomie assez différente des pupitres traditionnels des mélangeurs. Mais ils sont plébiscités par le monde du gaming et de nombreuses webTV. Plusieurs logiciels de mixage vidéo sont également proposés sur tablettes. Il ne faut pas oublier non plus les services de mixage vidéo dans le cloud comme Virtual Production de Sony, ceux de Dazzl et Vodalys ou encore Easylive, qui eux sont intégralement basés sur le traitement de flux de streaming. Nous reviendrons sur ces solutions alternatives dans les prochains numéros de *Mediakwest*. ■

Merci à Pascal Souclier et à Emanuele Di Mauro, respectivement directeur et responsable de projet à l'institut de formation IIFA, pour l'aide apportée à la rédaction de ce dossier.

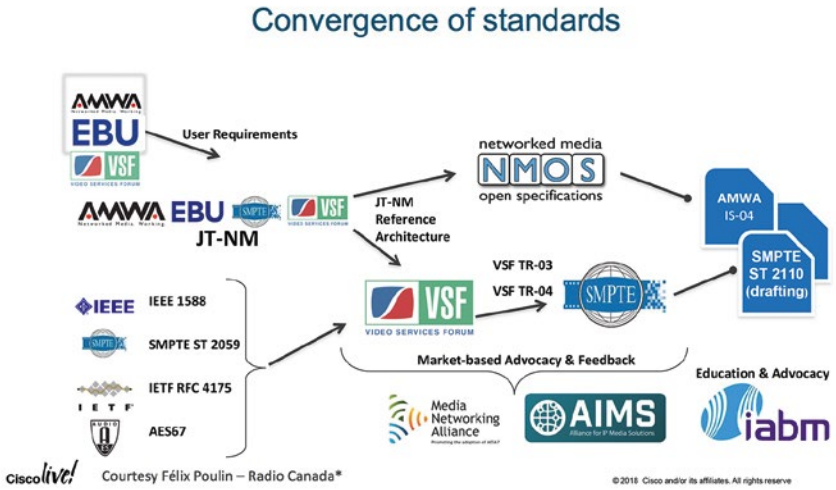
L'univers du ST 2110 et du NMOS, un si long chemin…

Comme dans tous les domaines de l'audiovisuel, le monde du broadcast a entrepris sa mutation vers les réseaux IP. Pour la production live, la SMPTE a démarré il y a plusieurs années un important travail de normalisation, concrétisé par la publication des standards SMPTE ST 2110. Autour du même sujet, est apparu le sigle NMOS avec ses spécifications IS. Ce travail normatif de tous les éléments de la vidéo IP dans un environnement broadcast est encore loin d'être achevé.

Par Pierre-Antoine Taufour

Le transport d'un signal vidéo de qualité broadcast sur un réseau IP dans un environnement de production live combine un nombre élevé de paramètres, des mécanismes complexes et des configurations largement plus nombreuses que pour une simple architecture SDI. La production des signaux live exige de s'appuyer sur quelques principes intangibles comme l'intersynchronisation des sources, l'absence de compression (quoique sur ce point les choses évoluent avec le Jpeg XS), une commutation propre lors du passage d'une source à l'autre dans une structure en paquets totalement décorrélée du balayage vidéo et la préservation d'une synchronisation entre l'image et le son. Par ailleurs le passage au tout IP ne doit pas bouleverser les habitudes de travail des équipes techniques et il faut leur garder les interfaces habituelles de commande.

À ces impératifs s'ajoutent aussi le souhait des chaînes et des prestataires de ne pas être enfermés dans des choix technologiques et des protocoles propriétaires propres à chaque constructeur. Dans les diverses instances qui réunissent tous les acteurs concernés : organismes techniques qui éditent des spécifications comme la SMPTE ou l'EBU, mais aussi des groupes d'utilisateurs comme le VSF, l'AIMS ou l'AMWA, deux exigences ont été exprimées très vite : définir des protocoles et des architectures ouvertes (ou open), disponibles de manière gratuite pour garantir l'interopérabilité et faciliter les évolutions dans le futur et s'appuyer au maximum sur les protocoles et standards d'Internet. Dans leurs réflexions, les initiateurs de cette révolution pensent à des architectures flexibles et réparties dans lesquelles les équipements sont organisés entre divers pôles techniques reliés par des faisceaux de fibres optiques et auxquels les équipes de production accèdent à distance, dans des configurations évolutives et redimensionnables en permanence. Cette notion d'interopérabilité est donc essentielle.



De nombreux organismes ont participé à l'élaboration des standards de la vidéo live sur IP, ce qui rend leur généalogie un peu complexe à comprendre. © Cisco

Une première étape avec le standard SMPTE ST 2110

Vu l'ampleur de la tâche, plusieurs organismes se sont attelés à ce titanesque travail de standardisation en le répartissant par grandes fonctions. La SMPTE s'est centrée sur deux grands domaines : la référence horaire et l'intersynchronisation d'un côté, le transport des contenus de l'autre. Elle avait déjà mené un premier travail qui avait conduit au standard SMPTE ST 2022 centré sur le transport des contenus dans une perspective de contribution ou de distribution avec une organisation des signaux dans un flux unique. Cela reste peu adapté à un travail de production où les diverses « essences » vidéo et audio doivent être traitées de manière différenciée. La SMPTE 2110 part sur un principe différent en les séparant clairement. La vidéo, l'audio et les données auxiliaires font l'objet chacun d'un chapitre distinct.

Tous ces éléments sont maintenant stabilisés et ont été repris par de nombreux constructeurs. Leur interopérabilité a été validée lors de multiples programmes de tests. Deux évolutions récentes sont à noter : la compression vidéo avec l'adoption de la norme Jpeg XS

Standards ST 2110	Intitulé	Description	N° couche pyramide EBU
ST 2110-10	System and timing	Distribution horloge PTP, protocole transport RTP, description session SDP	
ST 2110-20	Uncompressed video	Transport vidéo non compressée	1
ST 2110-21	Video Stream Packet Shaping	Mise en forme des paquets du flux vidéo	1
ST 2110-22	Constant Bit Rate Compressed Video	Vidéo compressée en mode CBR	
ST 2110-23	Single Essence Video Transport over Multiple ST 2110-20 Streams	Transport d'un flux unique pour signaux vidéo multiple (par ex. UHD en mode 4 x 3G SDI)	
ST 2110-24	Transport of SD signals		
ST 2110-30	Uncompressed audio	Format des données audio linéaires PCM - Basé sur AES67	1
ST 2110-31	AES3 Audio Streams	Transport audio compressé en AES3	
ST 2110-40	Ancillary Data	Métode de transport des données auxiliaires (Time code, Ss/Titres, format image,…)	
ST 2110-41	Extensible Fast Metadata Transport	Transport de méta données dynamiques	
ST 2022-7:2018	Seamless Protection Switching	Commutation automatique d'une liaison redondante	1
ST 2022-8	Timing of ST 2022-6 streams in ST 2110-10 systems	Référence temporelle pour transport flux ST 2022-6 dans ST 2110 *(en cours d'élaboration)*	
ST 2059v2	Timing and clock signals	Horloge de référence, basée sur IEEE 1588-2008	2

Le standard SMPTE ST 2110 est organisé en diverses parties dont la liste s'enrichit régulièrement. En fin de liste les autres standards qui leur sont associés. La colonne de droite renvoie vers le niveau des couches de la pyramide EBU.
© PA Taufour

(avec sa désignation officielle ISO/IEC 21122) et son utilisation dans le cadre du chapitre ST 2110-31. Le Jpeg XS est un codec de compression basé sur les ondelettes, comme le Jpeg 2000 avec un taux de compression similaire mais une très faible latence, inférieure à la milliseconde. Son taux de compression, entre 4 et 10, réduit les débits de l'UHD à ceux du 1080p, soit des valeurs plus adaptées aux ports à 10 Gb/s, moins onéreux que les modèles à 25, 40 ou 100 Gb/s. Le Jpeg XS est issu des travaux qu'a menés Intopix sur le Jpeg 2000 et sur son codec propriétaire Tico.

L'autre évolution concerne l'implémentation du ST 2110 dans les chipsets d'Audinate pour le transport du Dante. Dans un premier temps, la ST 2110 traitait les signaux audios sur IP en conformité avec l'AES67 avec un ensemble de paramètres communs aux divers protocoles propriétaires d'audio sur IP pour remplir le rôle de passerelle entre eux, mais avec des restrictions sur les codages et les formats. Suite aux évolutions du protocole PTP de référence horaire qui est passé en v.2, Audinate a implanté le ST2110 dans son chipset en parallèle à l'AES67, de manière à offrir une interopérabilité complète entre Dante et le ST 2110.

Mais la SMPTE ST 2110 ne fait pas tout

Plusieurs équipements d'envergure comme des cars régies ou des centres de production ont été déployés sur la base des standards SMPTE ST 2110. Mais dans plusieurs webinaires diffusés cet été par l'AIMS, leurs responsables techniques ont fait part de leurs difficultés pour la configuration et le pilotage des équipements. L'un d'eux relate que ses équipes ont dû configurer 500 000 adresses IP de manière manuelle avec juste l'aide de tableaux Excel. Ils constatent que le « data plane » et l'interconnexion des signaux fonctionnent correctement avec le ST 2110, mais que le gros challenge qui reste à relever concerne le « control plane ». Emanuele Di Mauro, responsable de projet à l'IIFA, exprime le même constat avec une métaphore routière : *« On a construit les autoroutes, les camions (les datas) peuvent circuler, mais qui va indiquer au camion son point de départ, sa destination et à quelle heure il doit partir ? C'est là toute la question. »*

Jusqu'à présent chaque concepteur d'orchestrateur (Grass Valley, Imagine Communications, Lawo entre autres) définissait les commandes pour piloter chaque équipement ou applicatif. Si elles n'étaient pas disponibles pour l'un d'eux, il fallait concevoir un

Spécifications NMOS	Intitulé	Description	N° couche pyramide EBU
IS-04	Discovery and registration	Découverte des ressources et constitution d'un registre d'adresses	3
IS-05	Device Connection Management	Établissement et suspension des flux de données	3
IS-06	Network Control	Liaison avec le SDN qui gère les actifs réseau (topologie, bande passante, etc…)	
IS-07	Event & Tally	Données et timing des détecteurs et actionneurs (Tally,…)	3
IS-08	Audio Channel Mapping	Gestion des canaux audio, affectation, groupage, séparation,…	3
IS-09	System *(en cours d'élaboration)*	Gestion des données communes à un système	4
BCP-002-01	Natural Grouping	Regroupement des émetteurs et récepteurs appartenant à un même équipement	3
BCP-003-01	Securing Communications	Sécurisation des communications dans les spécifications NMOS	5
BCP-003-02 (IS-10)	Authorization *(en cours d'élaboration)*	Gestion des autorisations clients au niveau des API NMOS	5

Chaque spécification NMOS est consacrée à un groupe de fonctions de contrôle des équipements et applicatifs. En fin de liste, les BCP (best current practice) sont l'étape initiale à l'élaboration des futures spécifications. La colonne de droite renvoie vers le niveau des couches de la pyramide EBU. © PA Taufour

connecteur « logiciel », petit programme développé sur mesure pour permettre le dialogue entre l'orchestrateur et l'équipement. Vu l'étendue des catalogues, le nombre de combinaisons devient vite énorme. Il est donc indispensable de mettre en place des mécanismes automatisés pour éliminer cette étape chronophage encore au cœur de nombreux projets.

Le rôle des spécifications NMOS

Les spécifications NMOS (networked media open specifications) ont été développées pour fournir une couche de gestion et de contrôle des équipements (aussi bien hardware que logiciels) en complément à la couche de transport définie par le standard SMPTE ST 2110. L'objectif est de fournir des outils pour faciliter l'interopérabilité entre les équipements provenant d'une large gamme de constructeurs et les outils d'orchestration.

La famille des spécifications NMOS est disponible gratuitement à la fois pour les fournisseurs d'équipements et les utilisateurs pour un déploiement de leurs infrastructures. Ces spécifications s'appuient en majorité sur les standards et les technologies d'Internet.

Le développement de NMOS est soutenu par le JT-NM (joint task force of networked media) qui associe pour ces travaux quatre organismes : l'AMWA, l'EBU, le SMPTE et le VSF

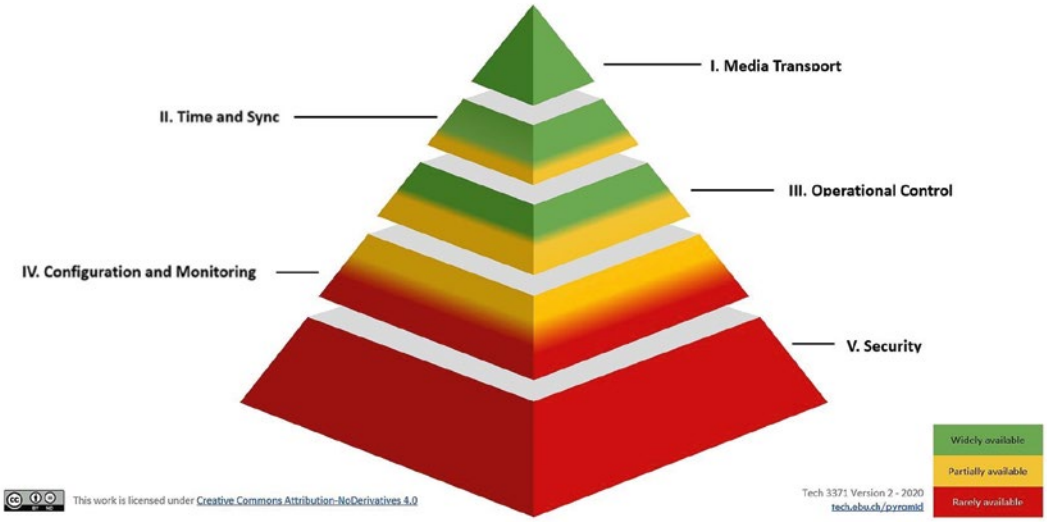
Plus de 70 membres de l'AMWA, fournisseurs et utilisateurs, participent à ces travaux dans divers groupes de travail. Le JT-NM travaille en étroite collaboration avec l'AIMS (Alliance for IP Media Solutions) dont l'objectif de favoriser et de promouvoir l'introduction des technologies IP dans l'industrie des médias.

Les spécifications NMOS sont publiées sous forme de recommandations dénommées IS (interface specifications) avec un numéro et qui correspondent à un groupe de fonctionnalités. Elles servent de base à la programmation d'API par les constructeurs ou éditeurs de services de manière à garantir l'interopérabilité entre les équipements et divers logiciels. Cette liste n'est pas figée car d'autres sujets sont en cours d'examen et font l'objet de BCP (best current practice), étape préalable à l'élaboration d'une spécification NMOS.

Faire le lien avec le réseau en dialoguant avec le SDN

Maintenant les principaux orchestrateurs du marché ont intégré les protocoles NMOS. Des difficultés subsistent encore au niveau des endpoints (sources et destination des signaux). Ils doivent être rendus compatibles NMOS. Or beaucoup d'équipements ne le sont pas encore, même s'ils sont annoncés comme conformes. D'autre part certains matériels n'offrent pas toute la souplesse permettant de les modifier à distance et sans délais. Il faut parfois changer manuellement leur configuration d'exploitation et, seulement après un reboot, il est alors possible de le piloter à distance. NMOS ne remplit pas encore toutes les fonctions dont les équipes ont besoin pour une vraie configuration « plug and play ». À terme il faut pouvoir repérer un équipement avec un libellé unique. Il y a juste à modifier son affectation par exemple « Camera 1 /Studio 3 », et tous les mécanismes de configuration s'adaptent de manière automatique. Il faut aussi pouvoir gérer facilement un parc d'équipements non permanents exploités en pool ou loués à l'extérieur.

Emanuele Di Mauro constate que certains orchestrateurs restent dans une logique de contrôle applicatif et que cela pose des difficultés : *« Si je demande à un récepteur d'aller récupérer un flux*



L'EBU publie régulièrement une synthèse de la disponibilité des divers standards et recommandations concernant la vidéo live sur IP, répartis sur cinq grandes couches fonctionnelles. Le vert correspond à des outils largement disponibles, tandis que le rouge indique des fonctions encore rarement prises en compte. Version de mars 2020. © EBU

Un dernier domaine dans lequel les travaux de normalisation et d'interopérabilité débutent seulement, concerne tous les aspects liés à la sécurité. Et dans un monde totalement interconnecté, nous savons qu'ils revêtent une importance capitale.

supplémentaire, je ne suis pas sûr que le réseau soit capable d'absorber ce débit. Même chose si j'ajoute un équipement ou bien si je bascule l'infrastructure d'un mode 50i à du 50p. Avec ce type d'orchestrateur je reste externe au réseau. D'autres orchestrateurs utilisent la technologie SDN (software defined network) qui sert à contrôler et paramétrer automatiquement le réseau et à rediriger les flux. Il avertit l'administrateur en cas de dépassement des capacités du réseau en étant toujours "non blocking". Il est important d'ajouter à l'orchestrateur une couche SDN pour qu'un dialogue s'établisse entre les deux entités. Dans le cas d'un simple contrôle applicatif, je devrais au moins vérifier que le réseau reste dans un état "non blocking" si on lance toutes les sources vers toutes les destinations et que le réseau n'est pas saturé. »

La gestion de l'état « blocking » et « non blocking » du réseau devient primordial et doit être remontée de manière explicite à l'orchestrateur. Si ce n'est pas le cas, la saturation d'un lien peut conduire à des dégradations des images et des sons. Si le SDN refuse l'action demandée, il doit en avertir l'orchestrateur pour qu'un message explicite soit renvoyé à l'exploitant vidéo pour lui annoncer que l'action demandée n'est pas possible. Le SDN, en particulier dans le cas d'une topologie de type « spine leaf » redondée, pourra rediriger les flux supplémentaires vers un lien moins chargé et optimiser la charge globale du réseau. Ces mécanismes vont devenir incontournables dans des architectures de remote production dans laquelle les liens intersites sont plus limités en termes de bande passante. Si le SDN n'est pas embarqué dans l'orchestrateur, on peut recourir à celui fourni par le constructeur des actifs réseaux. Dans ce cas, l'interaction entre l'orchestrateur et le SDN peut se gérer avec la spécification IS-06 (network control) du NMOS.

Un dernier domaine dans lequel les travaux de normalisation et d'interopérabilité débutent seulement, concerne tous les aspects liés à la sécurité. Et dans un monde totalement interconnecté, nous savons qu'ils revêtent une importance capitale.

Ces quelques exemples d'interdépendance entre les équipements, les flux, les actifs réseaux montrent qu'un important travail d'élaboration de spécifications, standards ou autres recommandations est encore nécessaire avant de déployer des infrastructures basées sur le standard SMPTE ST 2110 et le NMOS dans un environnement fiable et totalement « plug and play ». ■