



SMAC

SMART MULTIMEDIA ARCHIVE FOR CONFERENCES

Didier von Rotz, David Bourillot, Omar Abou Khaled & Rudolf Scheurer, Ecole d'Ingénieurs et d'Architectes Fribourg, Denis Lalanne & Rolf Ingold, Université de Fribourg et Jean-Yves Le Meur & Thomas Baron, CERN



INTRODUCTION

Le projet SMAC, déjà introduit lors d'un précédent article [1], est le fruit d'une collaboration entre le CERN, l'Université de Fribourg (Unifr) et l'École d'Ingénieurs et d'Architectes de Fribourg (EIA-FR). Chaque partenaire a apporté ses compétences qu'il a acquises lors de précédents projets qui ont servi de base pour la réalisation du projet SMAC.

Le CERN, via le projet Indico [2], s'est construit une solide expérience autour des systèmes de gestion de conférences [3], ainsi qu'au niveau de la recherche et de l'indexation de documents scientifiques avec le logiciel CDSWare [4].

L'université de Fribourg, en collaboration avec l'école d'Ingénieurs et d'Architectes de Fribourg, a développé des compétences importantes sur le projet IM2.DI [5][11][12], dans l'intégration des documents dans des applications multimodales.

L'objectif du projet SMAC est le développement d'une plate-forme permettant l'acquisition, l'archivage, l'analyse et la consultation d'archives de conférences multimédias dans un premier temps, puis son intégration avec les logiciels Indico et CDSWare dans un second temps. Cet ensemble permet une gestion complète d'une conférence, de sa création jusqu'à sa consultation à la demande (fig. 1). La validation de cette réalisation s'est faite par l'installation et la mise en service de plusieurs prototypes dans les salles de conférences principales de l'EIA-FR et du CERN.

Le but de cet article est de présenter les résultats du

projet SMAC, de les commenter, et de donner une impulsion pour la suite des développements qui peuvent constituer un intérêt majeur pour les différentes institutions désireuses de se doter d'un tel système pour la gestion de leurs conférences ainsi que pour toute autre application se basant sur l'utilisation de vidéos, de transparents et de documents, le tout étant synchronisé et diffusé selon les besoins de l'utilisateur.

Pour commencer, l'architecture du système sera abordée, puis son déploiement ainsi que son fonctionnement.

ARCHITECTURE

La principale contrainte lors de la conception de SMAC était de pouvoir l'exploiter aussi bien de façon autonome, avec sa propre interface de gestion, mais aussi via Indico. Il a donc fallu développer SMAC de façon à ce qu'il soit facilement interfaçable. Une partie du projet consistait aussi à intégrer CDSWare afin de permettre l'indexation et la recherche de documents liés aux présentations.

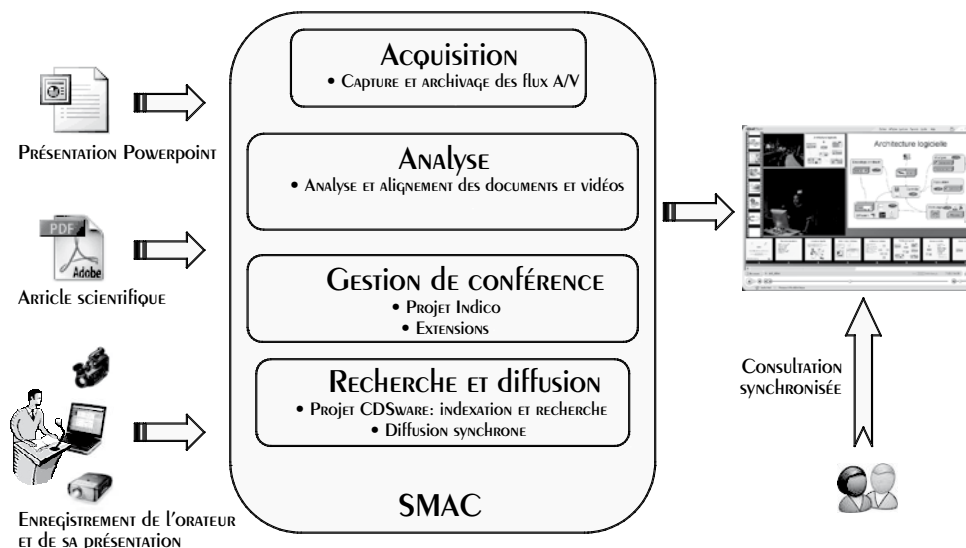


fig. 1 – Principales fonctionnalités de SMAC

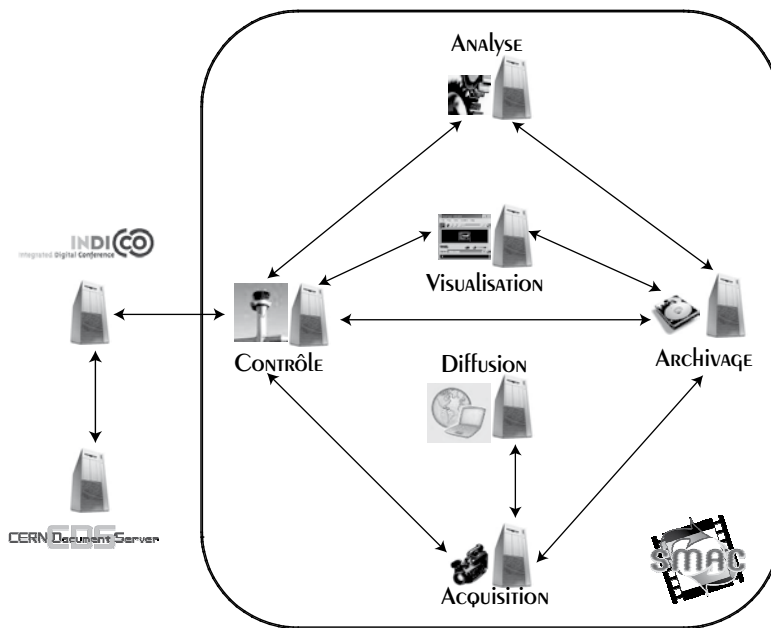


fig. 2 – ARCHITECTURE MODULAIRE DE SMAC

SMAC

L'architecture SMAC a été conçue selon une approche modulaire (fig. 2), les principales tâches ont été affectées à des modules indépendants et distribués. Ceci a permis une grande souplesse et une maintenance facilitée, grâce aux interfaces standardisées entre chaque module.

Ces modules ont été conçus avec les technologies et les langages les plus appropriés à leurs fonctions. La communication inter-modules est assurée par le protocole XML-RPC. Celui-ci permet à diverses applications s'exécutant sur divers systèmes d'exploitation ou différentes machines de communiquer entre elles. Cette communication se fait via des appels de procédures à distance. Également, XML-RPC assure la compatibilité grâce à l'utilisation du format XML sur une couche de transport HTTP; ces deux technologies étant disponibles dans la plupart des langages et environnements.

Une couche d'abstraction supplémentaire a été ajoutée par le biais d'une bibliothèque de programmes clients centralisant les accès aux fonctions disponibles pour chaque module.

Les modules de la figure 2 sont détaillés dans les paragraphes suivants.

Module de contrôle

Ce module gère l'état des contributions de toutes les conférences. Il permet aussi de contrôler les différentes phases de traitements, comme l'acquisition, la diffusion en direct ou l'analyse de la présentation.

Ses tâches principales sont l'automatisation des processus à lancer sur les différents modules, ainsi que le stockage des métadonnées résultantes des analyses des différents médias d'une contribution (vidéos, documents PDF ou PowerPoint).

Ces fonctionnalités sont assurées par un logiciel développé en Python associé à une base de données MySQL.

Le module de contrôle fournit aussi une interface Web pour la gestion du système et la soumission des documents électroniques (fig. 3). Les technologies utilisées pour cette interface sont des scripts en Python s'exécutant sur un serveur Apache.

Module d'acquisition

Lors d'une présentation, il est nécessaire de capturer les différents flux audio/vidéo provenant des caméras filmant l'orateur et/ou la salle, ainsi que le diaporama projeté, de façon synchrone. La synchronisation de ces flux est primordiale pour leurs futures analyses. En effet, les résultats de celles-ci permettront la rediffusion correcte de la présentation.

La numérisation d'une vidéo brute demande beaucoup d'espace de stockage. Il est donc nécessaire de les compresser. Ceci consommant beaucoup de puissance de calcul, la solution a été d'utiliser une carte permettant l'acquisition et la compression à la volée de plusieurs flux audio et vidéo en parallèle.

Une telle carte ne prend en charge que l'acquisition et la compression, de façon hardware. La synchronisation, par contre, nécessite une partie logicielle. Pour ceci, **Microsoft Directshow** est la meilleure solution, pour autant que les *drivers* de cette carte le permettent.

Directshow est une interface de programmation permettant de rejouer ou capturer des flux audio et vidéo à partir d'entrées analogiques ou digitales. Elle permet en particulier de piloter et synchroniser la capture des flux audio et vidéo des caméras et microphones du système d'enregistrement de SMAC. Le principe de Directshow est basé sur un graphe de capture, constitués de filtres, reliés entre eux. **GraphEdit** est une application, fournie avec Directshow, qui permet de simuler la construction et l'exécution d'un tel graphe. La figure 4 représente un graphe permettant de capturer

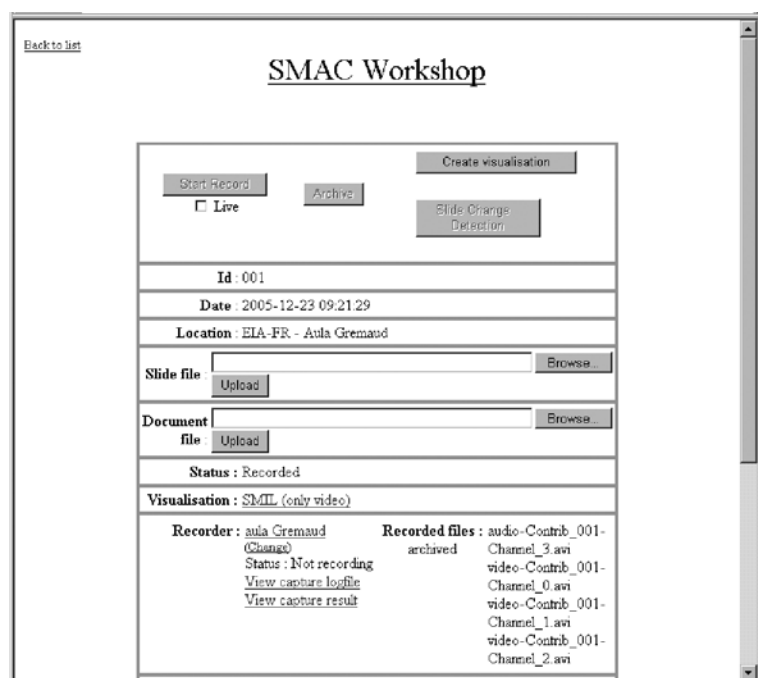


fig. 3 – INTERFACE (TEMPORAIRE) DE GESTION DE CONTRIBUTION

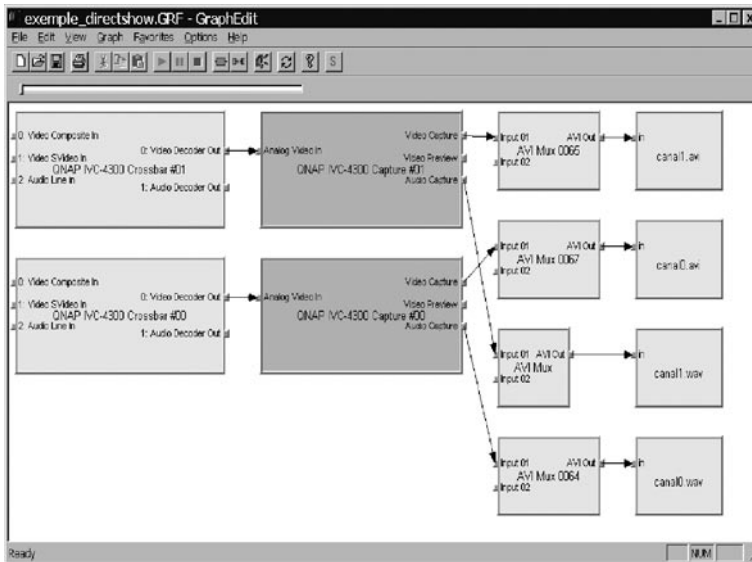


fig. 4 – Exemple de graphe de capture Directshow dans GraphEdit

deux flux audio et vidéo à partir d'une carte d'acquisition. L'exécution de celui-ci va produire deux fichiers vidéo et deux fichiers audio, tous synchrones.

Chaque canal (couple de flux audio et vidéo) est configurable. Il est possible de modifier certains paramètres pour la vidéo (débit, résolution, type de compression, etc.) et l'audio (mono/stéréo, fréquence d'échantillonnage, etc.). Les fichiers résultants sont au format AVI, utilisant le codec Xvid pour la vidéo et au format PCM (non compressé) pour l'audio.

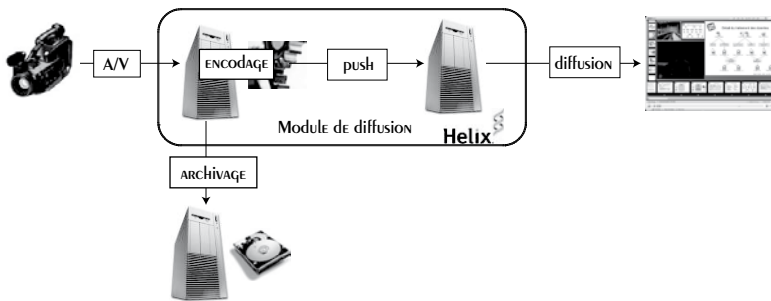


fig. 5: Diffusion de flux A/V EN DIRECT

La carte utilisée possède 4 canaux. Un ordinateur peut supporter jusqu'à 4 cartes, donc 16 canaux.

La partie logicielle de ce module est composée de deux programmes. Le premier, écrit en C++ (Directshow) est responsable du pilotage de la carte d'acquisition. Les principales fonctionnalités sont le paramétrage des canaux, le démarrage et l'arrêt de la capture. Le second, écrit en Python, gère l'envoi des fichiers au module d'archivage et contient l'interface lui permettant de communiquer avec les autres modules.

Un système SMAC est capable de gérer de multiples modules d'acquisition. Il est donc possible d'équiper plusieurs salles et de les gérer de façon centralisée et concurrente.

Module de diffusion

Ce module est responsable de la diffusion en direct d'une présentation.

Il comprend deux parties distinctes, l'encodeur et le diffuseur (fig. 5). L'encodeur (*Real Producer*) se charge de compresser le flux audio/vidéo provenant d'une des caméras de la salle au format *RealMedia* et de le pousser vers le diffuseur (*Helix Server*).

La partie logicielle est composée d'un programme Python qui pilote l'encodeur *Real Producer*. Le démarrage et l'arrêt de la retransmission sont synchronisés avec ceux du module d'acquisition. Le fichier *RealMedia* résultant est ensuite archivé et rendu accessible pour une rediffusion sur demande.

Module d'archivage

Ce module a deux tâches principales: le stockage et la conversion des fichiers audio, vidéo, images, PDF et PowerPoint. Ces fichiers sont envoyés par les différents modules au moyen du protocole SFTP (Secure FTP). Ils sont alors référencés avec un identifiant unique. Le contrôleur

sauvegarde cet identifiant parmi les métadonnées du fichier concerné. Les autres modules peuvent alors accéder facilement aux archives, via le client du module d'archivage.

La partie conversion se charge de convertir chaque fichier nouvellement stocké. Le format de la conversion dépend du type de fichier. Les vidéos et pistes son sont converties au format *RealMedia*. Les images bitmap, provenant du module d'analyse, sont converties au format JPEG. Ces différents formats facilitent la diffusion de ces données.

La partie logicielle est composée d'un programme Python et d'un serveur SSH. Le premier gère la structure des archives, ainsi que l'exécution des processus de conversion. Chaque module ayant besoin d'échanger des fichiers avec l'archivage utilise un programme client qui s'occupe des transferts avec le serveur SSH.

La figure 6 illustre le transfert d'un fichier du module d'acquisition au module d'archivage. Le module d'acquisition utilise le client du module d'archivage pour faire une requête d'envoi de fichier. Il reçoit en retour les informations lui permettant de placer le fichier au bon endroit sur le serveur SSH. Il peut donc se connecter au serveur SSH pour y déposer le fichier.

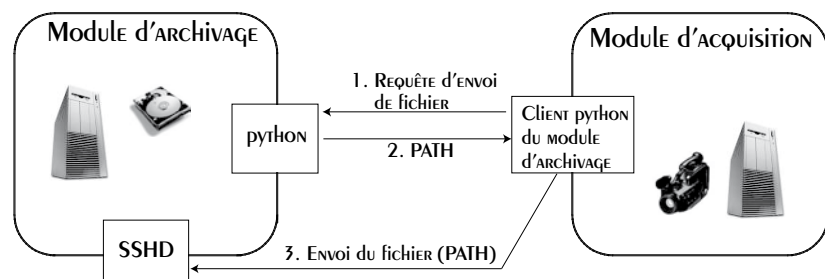


fig. 6 – ARCHIVAGE D'UN FICHIER À PARTIR DU MODULE D'ACQUISITION

Module d'analyse

Ce module est responsable de la création de métadonnées à partir des différents matériels d'une présentation (vidéos, documents, etc.). Celles-ci sont utilisées pour synchroniser les médias d'une présentation, afin de pouvoir créer des liens de navigation lors de la visualisation de celle-ci.

Un programme Python a pour tâche l'exécution des différentes analyses. Toutes les données produites sont alors archivées. Il se charge aussi de mettre à disposition les données à analyser, en allant les chercher sur le module d'archivage.

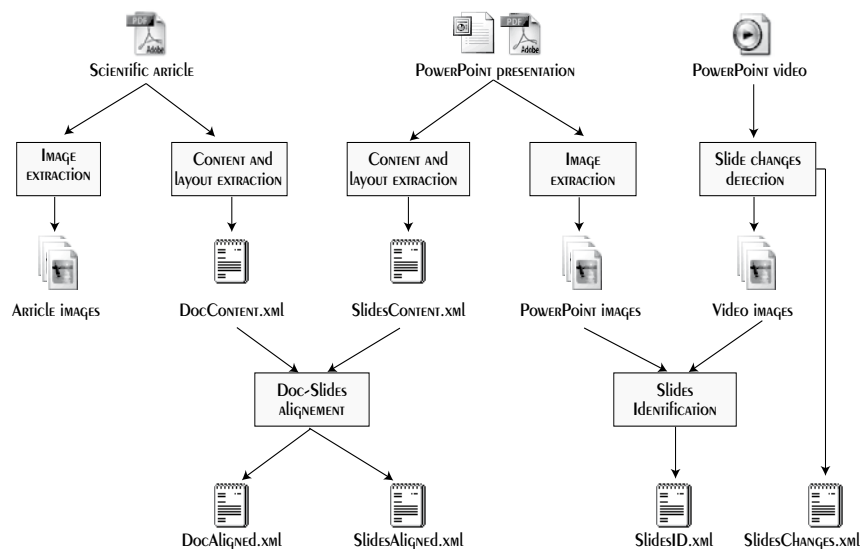


fig. 7 – SÉQUENÇEMENT DES ANALYSES DES DONNÉES

La figure suivante (fig. 7) illustre le séquençement des analyses des différents médias d'une contribution.

Les programmes utilisés pour les analyses de données de conférences sont issus des travaux menés par le groupe DIVA de l'Université de Fribourg [6][11][12].

La première analyse est la détection des changements de diapositive. L'analyse de la vidéo du diaporama fournit un index temporel pour chaque passage stable (plan non variant d'au moins trois secondes) [7]. Une image est extraite pour chacun de ces passages.

Deux types de document sont liés à chaque présentation: le document scientifique discuté, au format PDF, et le diaporama, au format PowerPoint ou PDF. À partir du document scientifique, on extrait les coordonnées et le contenu textuel de chacun des blocs physiques (titres, paragraphes, etc.) de chaque page [8]. Le traitement du document du diaporama est similaire. La différence est que seul le contenu textuel par page est extrait. Une image bitmap est extraite pour chaque page des deux documents.

Le but de la prochaine analyse est de faire correspondre une diapositive à un des passages stables de la vidéo du diaporama. Pour ceci, une signature est extraite pour toutes les images (de la vidéo et du diaporama). Ces signatures sont alors comparées et un lien est établi pour les images ayant la plus grande correspondance [9]. Ceci permet de savoir à quel moment a été projetée chaque diapositive. De plus, cela permet de remplacer une image extraite d'une vidéo par son original qui est de qualité supérieure.

La dernière analyse utilise le contenu textuel des deux types de documents, afin de créer des liens thématiques entre les blocs physiques du document scientifique et les pages du diaporama, et inversement [10]. Les résultats de ces analyses sont stockés dans des fichiers XML.

Module de visualisation

Ce module est chargé de créer différentes interfaces de visualisation en fonction des données disponibles, pour une contribution donnée. Pour ceci, les médias et documents de la présentation, ainsi que les métadonnées résultantes des analyses, sont utilisés. Deux types de visualisations sont disponibles: une au format SMIL (Synchronized Multimedia Integration Language) et une interface Web. Toutes deux sont accessibles à distance, via *streaming*.

L'interface SMIL est composée des vidéos de l'orateur, du diaporama et d'aperçus de la salle, de

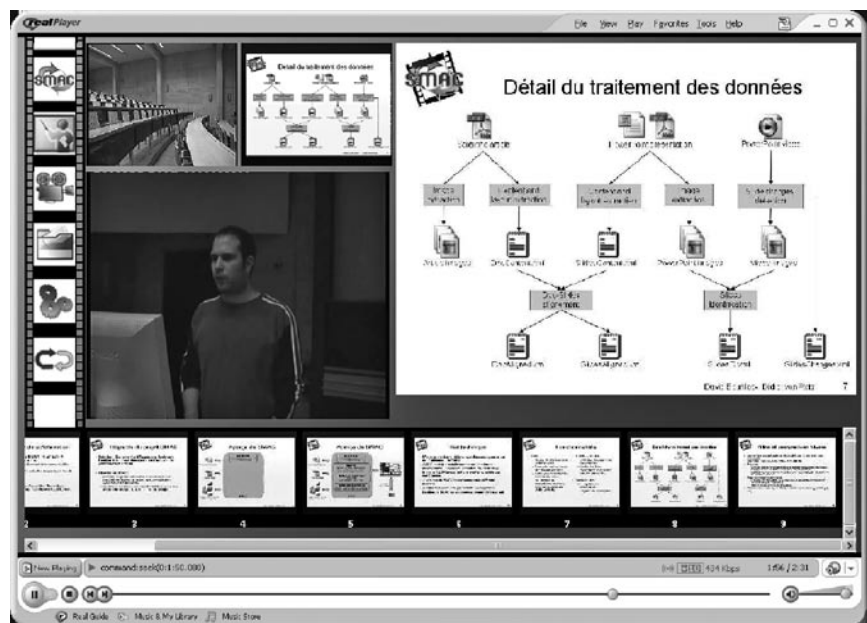


fig. 8 – INTERFACE SMIL POUR LA VISUALISATION D'UNE PRÉSENTATION

l'image de la diapositive courante, ainsi que d'un index des diapositives (fig. 8). Tous ces composants sont synchronisés; en cliquant sur une image de l'index ou en utilisant la barre de navigation, l'utilisateur peut accéder à un moment précis de la présentation.

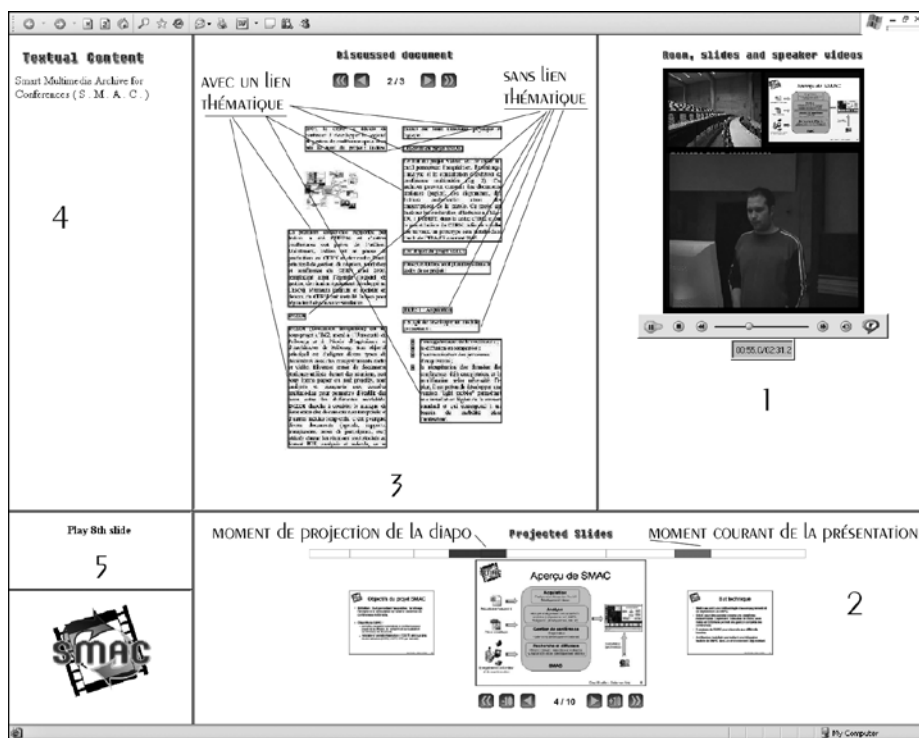


fig. 9: INTERFACE Web POUR LA CONSULTATION D'UNE PRÉSENTATION

La version Web de l'interface est composée des éléments suivants (fig. 9):

1. les vidéos de l'orateur, de la salle et du diaporama,
2. le diaporama,
3. le document discuté,
4. une zone permettant d'afficher le contenu textuel d'une page du document ou du diaporama,
5. une zone permettant l'affichage de messages d'aide ou d'information.

La zone du diaporama (2) contient les images originales de celui-ci. Il est possible de naviguer de l'une à l'autre grâce à des boutons. La barre en dessus des images permet de visualiser la répartition temporelle des diapositives. Le rectangle rouge représente le moment courant de la présentation; les bleus sont les moments où la diapositive courante est projetée.

Le document discuté est affiché (3) page par page. Les paragraphes sont encadrés par différentes couleurs, suivant qu'ils possèdent un lien thématique (bleu) vers une des diapositives ou non (vert). Il est aussi possible de zoomer sur l'un d'eux ou même d'afficher leur contenu textuel dans la zone prévue à cet effet.

Le document et le diaporama sont liés par des liens thématiques; les vidéos (1) et le diaporama le sont par des liens temporels. Lorsqu'un de ces trois éléments est manipulé, les deux autres sont automatiquement mis à jour.

Du point de vue logiciel, le module de visualisation est composé d'un programme en Python et d'un programme en Java. Le premier se charge de récupérer les données nécessaires de la présentation, c'est-à-dire les fichiers *Real-Media*, les images du diaporama et du document, ainsi que les résultats des analyses sous forme de fichiers XML. Le second programme construit une arborescence de fichiers à l'aide de ces données, pour chaque type de visualisation. Les technologies utilisées sont:

- SMIL, RealPix pour la version SMIL, qui est visualisable grâce au programme RealPlayer.
- SVG, SMIL et JavaScript pour la version Web.

LES MODIFICATIONS APPORTÉES À INDICO

Indico est un logiciel de gestion de conférence. Il permet d'organiser toutes les étapes d'une conférence, depuis l'appel à contribution jusqu'à la publication de ses actes. L'association d'Indico et de SMAC permet une meilleure utilisation de ces deux logiciels. Indico apporte à SMAC une interface complète pour la saisie des métadonnées de chaque contribution ainsi qu'une gestion complète des droits sur le contrôle de l'acquisition et SMAC apporte à Indico un module d'acquisition et de rediffusion des contributions.

Les modifications apportées à Indico se situent au niveau de l'interface Web; l'ajout d'une partie pour la sélection d'un enregistreur SMAC et une autre pour la délégation des droits du contrôle de l'acquisition. Un module dédié à la communication avec SMAC, utilisant le client du module de contrôle de SMAC, a été créé. L'utilisation du client SMAC fait que la seule modification apportée à SMAC a été une fonctionnalité pour l'envoi des liens des visualisations à Indico.

LES MODIFICATIONS APPORTÉES À CDSWARE

CDSWare apporte à Indico un module de recherche sur les données des conférences et contributions. Pour répondre aux demandes de recherche, CDSWare a besoin d'indexer les données d'Indico. Cette opération a été simplifiée par le support des deux logiciels du protocole OAI. Il a donc suffi d'ajouter au serveur OAI d'Indico le format OAI-marc, pour que CDSWare puisse automatiquement et régulièrement récupérer ses données. Il reste alors à configurer CDSWare pour qu'il affiche de la manière souhaitée les résultats des recherches.

DÉPLOIEMENTS DE SMAC À L'EIA-FR ET AU CERN

Deux systèmes SMAC ont pour le moment été installés dans les salles de conférences principales de l'EIA-FR et du CERN. Divers tests ont été effectués en situation réelle lors de conférences au CERN ou lors de présentations du projet SMAC à l'EIA-FR.

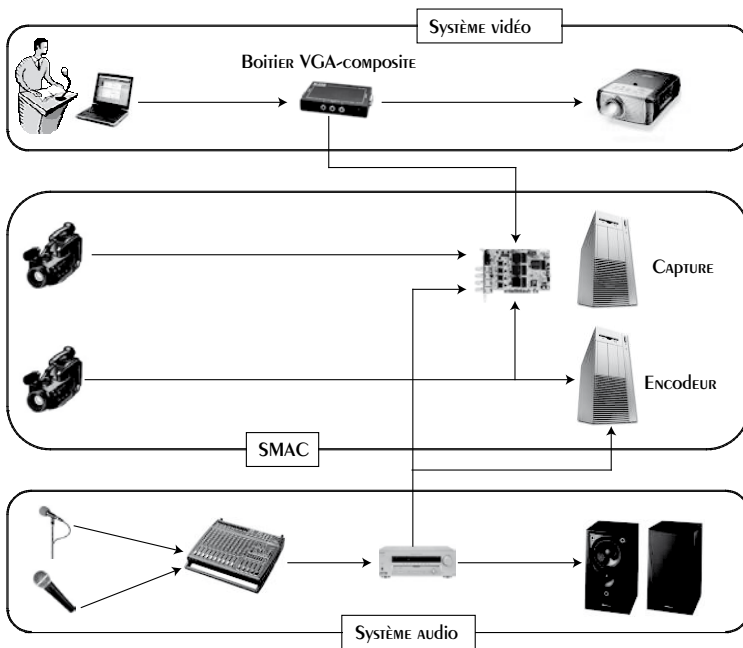


fig. 10 – Système audio et vidéo de la salle de conférences de l'EIA-FR,

DÉPLOIEMENTS DE SMAC À L'EIA-FR

Le prototype installé à l'EIA-FR utilise les systèmes audio et vidéo déjà présents dans la salle (fig. 10). Le système audio de la salle est composé d'un ensemble de micros avec et sans fil, branchés sur une table de mixage. La sortie de cette

table est alors envoyée sur un amplificateur. La prise de son pour SMAC se fait sur une sortie de l'amplificateur par l'intermédiaire de la carte IVC4300 du module d'acquisition. La partie vidéo est quant à elle constituée d'un boîtier de conversion VGA vers composite, qui est relié à la carte du module d'acquisition. Deux caméras ont été ajoutées pour filmer l'orateur et une vue d'ensemble de la salle. Elles sont aussi connectées à la même carte. Une d'elles est aussi utilisée pour la diffusion en direct. Elle est donc reliée à l'encodeur du module de diffusion. Les modules de SMAC ont été installés sur deux machines. Le module de diffusion a été installé seul sur une machine, car il nécessite beaucoup de puissance de calcul pour l'encodage de la vidéo.

Ce système est simple à intégrer dans une salle de conférence. La même topologie a été aisément utilisée au CERN. SMAC peut aussi être utilisé dans d'autres types de salle comme les salles de cours, salles de réunion, etc.

EXPLOITATION AISÉE

SMAC permet la gestion de A à Z d'une conférence, de sa création à la consultation. Le schéma ci-dessous représente les principales étapes d'un tel processus. Indico et CDSWare sont utilisés dans ce séquençement. SMAC peut aussi être exploité sans ces deux interfaces. Les seules fonctionnalités manquantes seront la gestion de conférences et l'indexation et recherche des données. Le module de contrôle de SMAC possède sa propre interface de manipulation de contribution, qui permet de créer des contributions, de les enregistrer, d'ajouter des documents et d'accéder aux interfaces de visualisation.

Les chapitres suivants décrivent les cinq principales étapes (fig. 11) d'une exploitation typique du système SMAC

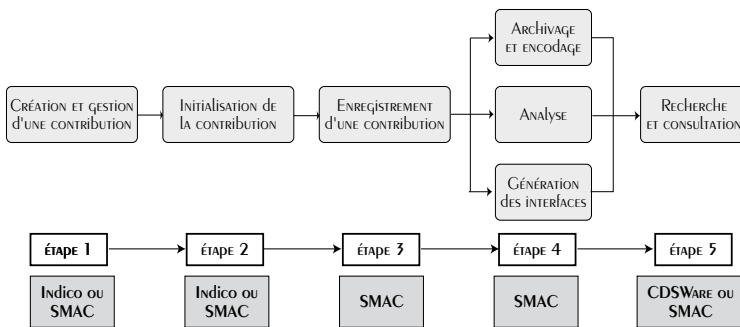


fig. 11 – Séquençement des tâches pour la gestion d'une contribution

ÉTAPE 1 – CRÉATION ET GESTION D'UNE CONFÉRENCE

L'utilisateur final (administrateur, responsable, etc.) commence par la création d'une conférence en ajoutant un événement dans Indico. Les données de base comme le titre, le lieu et la date sont alors saisies (fig. 12). Les contributions sont ensuite créées, soit en entrant directement leurs informations, soit en utilisant le module d'appel à contribution. Enfin, les contributions sont organisées dans le temps et l'espace, en leur affectant des horaires et des salles.

Les auteurs ont la possibilité de déposer leurs documents liés à la présentation, comme le diaporama au format PowerPoint ou l'article scientifique en PDF.

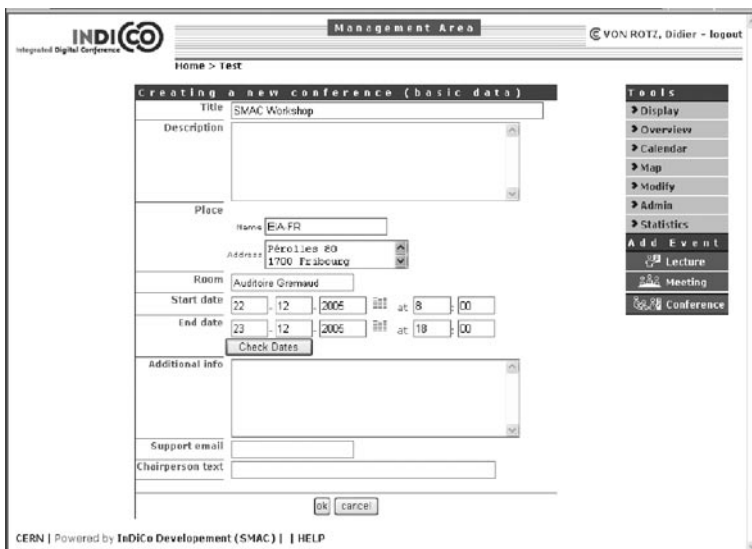


fig. 12 – Interface de création d'événement d'Indico

ETAPE 2 – INITIALISATION DE LA CONFÉRENCE

Pour pouvoir utiliser SMAC pour enregistrer les contributions, il est nécessaire de réserver les systèmes d'acquisition associés aux salles où elles se dérouleront. Les réservations peuvent s'effectuer au niveau de la conférence, des sessions, ou de chaque contribution. Un système est automatiquement proposé si la salle en est équipée. Le droit de démarrer et d'arrêter l'acquisition peut être accordé à des personnes spécifiques, et cela, pour les différents niveaux de la conférence (fig. 13).

ETAPE 3 – ENREGISTREMENT D'UNE CONTRIBUTION

Utilisation

Le démarrage et l'arrêt de la capture d'une contribution se font aussi à partir de l'interface Indico (fig. 14). Seules les personnes ayant les droits nécessaires ont accès à cette fonctionnalité. Une fois la capture achevée, tous les processus suivants sont automatisés.

Fonctionnement

Lorsque l'opérateur démarre l'enregistrement en cliquant sur le bouton de l'interface, le module de contrôle relaie cette commande au module d'acquisition, qui a été auparavant sélectionné et réservé. Les différents flux audio et vidéo sont capturés et stockés en local sur la machine d'acquisition, en tenant compte des paramètres présélectionnés. L'arrêt se déroule de façon similaire. L'archivage des données est automatiquement lancé à la fin de cette étape.

ETAPE 4 – TRAITEMENT DES DONNÉES

Utilisation

Cette étape contient plusieurs processus distincts: l'archivage des données, l'analyse de celles-ci et la génération des interfaces pour la visualisation de la présentation (fig. 7). Elle est totalement transparente pour l'utilisateur.

Fonctionnement

Ces tâches s'exécutent en parallèle. Le principe est le suivant: dès qu'une nouvelle donnée (média, document ou métadonnée résultante d'une analyse) est archivée, un nouveau processus est lancé sur le module d'analyse ou de



fig. 13 – Attribution d'un enregistreur SMAC à une contribution, à partir de l'interface Indico

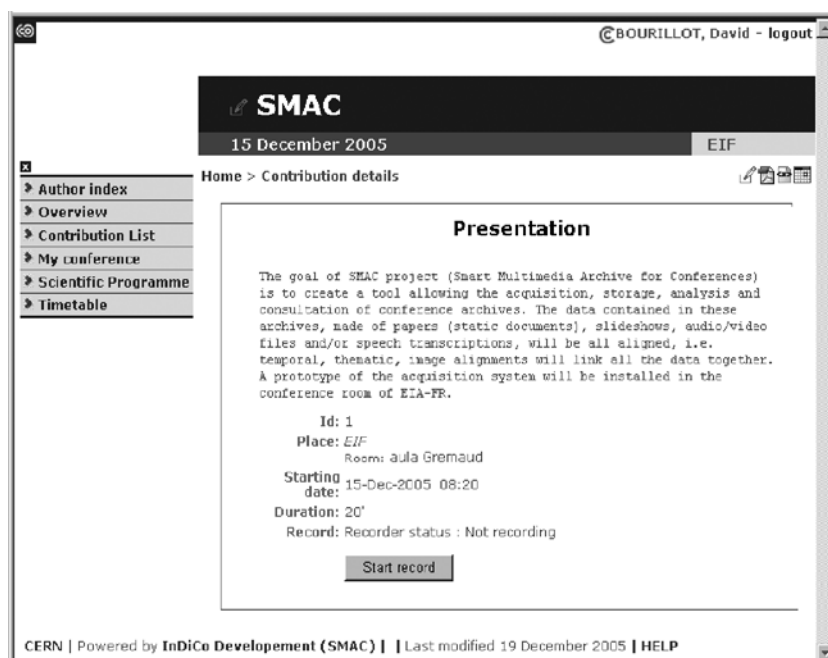


fig. 14 – Démarrage et arrêt de la capture d'une présentation à partir de l'interface Indico

visualisation. Par exemple, dès que la vidéo du diaporama est archivée, elle est automatiquement analysée, pour générer le timing des changements de diapositive. La génération des interfaces de visualisation est aussi relancée à chaque ajout de donnée, par exemple les changements de diapositive ou l'alignement du document et des diapositives.

Des liens vers les différentes interfaces sont automatiquement ajoutés sur la page de la contribution.

ETAPE 5 – RECHERCHE ET VISUALISATION D'UNE CONTRIBUTION

CDSWare indexe régulièrement les données des contributions d'Indico. Il est donc possible grâce à l'interface de recherche de CDSWare de retrouver les présentations qui