

Μεταφορά/διανομή και Αρχειοθέτηση τηλεοπτικού περιεχομένου

Επικαιροποίηση της Ενότητας 3.6 + Μελέτη τεχνικών και
προτύπων που χρησιμοποιούνται στην πράξη από μεγάλους
τηλεοπτικούς σταθμούς(BBC)

Αθανάσιος Μπίμης

08/09/2020

Πίνακας περιεχομένων

1. Σύγχρονες τάσεις στη μεταφορά/διανομή και παρουσίαση πολυμεσικού περιεχομένου	2
1.1 Streaming με χρήση του πρωτοκόλλου HTTP	2
1.1.1 Apple's Low-Latency HTTP Live Streaming (Low-Latency HLS) ..	2
1.1.2 Web Real-Time Communication (WebRTC)	3
1.1.3 Secure Reliable Transport (SRT)	4
1.1.4 Real-Time Messaging Protocol (RTMP)	4
1.2 Low Latency MPEG-DASH πάνω απο HTTP 2.0 με WebSocket	5
1.3 Διανομή CMAF περιεχομένου με Δυναμικές Ροές πάνω από HTTP (DASH)...	8
2. Μελέτη τεχνικών για την αρχειοθέτηση τηλεοπτικού περιεχομένου από μεγάλους τηλεοπτικούς σταθμούς	9
2.1 Ανάγκη για αρχειοθέτηση	9
2.2 Συστημα Ψηφιοποίησης Ingex(BBC R&D).....	10
2.3 Material Exchange Format.....	11
2.4 Τεχνικές αναζήτησης αρχείου μέσω Μέτα - δεδομένων.....	12
2.4.1 Σύστημα ABC-IP(Automatic-Broadcast-Content-Interlinking Project).....	13
2.4.2 Speech-to-text Συστήματα - Σύστημα COMMA (BBC R&D).....	13
2.4.3 Δημιουργία Μετα-Δεδομένων βάσει Διάθεσης (Moodata).....	14
Βιβλιογραφία – Αναφορές.....	15

1. Σύγχρονες τάσεις στη μεταφορά/διανομή και παρουσίαση πολυμεσικού περιεχομένου(Επικαιροποίηση Ενότητας 3.6 Επικοινωνίες Πολυμέσων)

Έπειτα από χρόνια καθιέρωσης του προτύπου MPEG-2 στην ψηφιακή τηλεόραση, αρχίζει μια ανανέωση που συμβαδίζει με την ραγδαία ανάπτυξη της τεχνολογίας. Πέρα από τα τηλεοπτικά προγράμματα, το real-time streaming πλέον χρησιμοποιείται από πολλαπλές εφαρμογές για μεταφορά περιεχομένου(Netflix, Youtube, Facebook Live).

Οι τεχνικές που έχουμε ήδη δει(MPEG-DASH, HLS)[[SODA2011](#)][[SODA2012](#)][[www.1-1](#)] έχουν αρχίσει να ανανεώνονται, καθώς με την ανάπτυξη ταχύτατων δικτύων παγκοσμίως (WiFi, 5G, VDSL,...) χρειαζόμαστε ταχύτερη κωδικοποίηση με μικρότερο latency και κοινό τύπο δομών, για την ευρεία χρήση live IPTV και υψηλής ανάλυσης On – Demand περιεχομένου.

Στην συνέχεια γίνεται μια σύντομη αναφορά στις εξελίξεις και στα κύρια πρότυπα υλοποίησης των παραπάνω μοντέλων όσον αφορά την διακίνηση πολυμεσικού περιεχομένου, με σκοπό την επικαιροποίηση της παραγράφου 3.6 του βιβλίου των Επικοινωνιών Πολυμέσων[[ΠΟΛΥΜ](#)].

1.1 Streaming με χρήση του πρωτοκόλλου HTTP

1.1.1 Apple's Low-Latency HTTP Live Streaming (Low-Latency HLS)

Έπειτα από την αρχική της πρόταση για το HTTP Live Streaming(HLS), η Apple έρχεται να συμβαδίσει με τις ανάγκες του σήμερα όσον αφορά το χαμηλό latency, ανακοινώνοντας το Low-Latency HTTP Live Streaming(Low-Latency HLS) στην WWDC τον Ιούνιο του 2019[[www.1.1-1](#)]. Το νέο πρωτόκολλο φτάνει τον χρόνο του latency για public δίκτυα σε αριθμούς τηλεοπτικών μεταδόσεων. Το Low-Latency HLS είναι μια επέκταση του ήδη υπάρχοντος HLS πρωτοκόλλου[[www11](#)] προσφέροντας την σημαντικότερη λειτουργικότητα στην παραγωγή μικρότερων από πριν(HLS) κομματιών(Partial Segments).

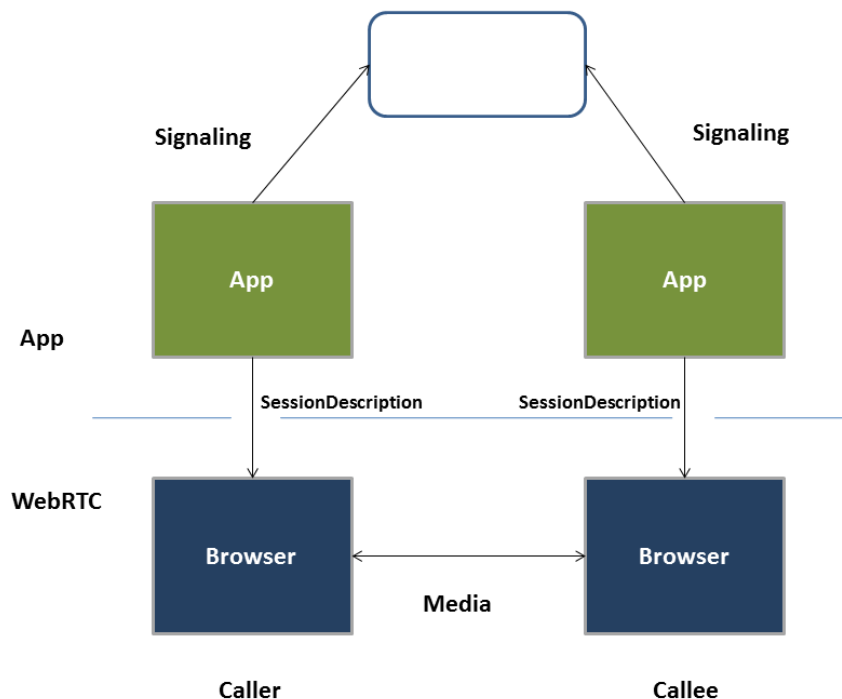
Το Low-Latency HLS δίνει ένα παράλληλο κανάλι για μετάδοση live περιεχομένου, όπου το περιεχόμενο σπάει σε ένα πολύ μεγαλύτερο αριθμό κομματιών(CMAF chunks[[CMAFakam](#)]). Τα κομμάτια αυτά ονομάζονται HLS Partial Segments και έχουν την δυνατότητα να αποθηκεύονται, πακετάρονται πολύ πιο γρήγορα στην Media Playlist από το Parent Segment. Δηλαδή ενώ τα Media Segments είναι 6s το καθένα, τα νέα segments μπορούν να φτάσουν τα 200 ms.

Το πρώτο Partial Segment γίνεται publish μόνο 200ms μετά το προηγούμενο δημοσιευμένο Segment, ακολουθούμενο από 29 από τα αδερφάκια του και ένα 6s Media Segment. Το media segment περιέχει το ίδιο περιεχόμενο με το σύνολο των 30 partial segments. Οπότε για να μειωθεί ο φόρτος του πρόσθετου αρχείου-καταλόγου(playlist ή index file), τα partial segments αφαιρούνται μόλις γίνουν μεγαλύτερα από 3 target durations.

1.1.2 Web Real-Time Communication (WebRTC)

Το πρότυπο Web Real-Time Communication (**WebRTC**)[[www1121](#)] είναι ένα δωρεάν ανοιχτού-λογισμικού project που υποστηρίζεται από τις Apple, Google, Microsoft, Mozilla και Opera. Η προσπάθεια του WebRTC τυποποιείται σε επίπεδο API από τον W3C και σε επίπεδο πρωτοκόλλου από τον IETF[[www2422](#)]. Στο Πλαίσιο 1 δείχνεται σχηματικά η ροή περιεχομένου με τη συγκεκριμένη τεχνική. Το WebRTC κάνει δυνατή την ζωντανή peer-to-peer επικοινωνία μεταξύ browsers, χωρίς τη χρήση επιπλέον λογισμικού. Κάθε χρήστης πρώτα βρίσκει την public IP του. Έπειτα, κανάλια δεδομένων δημιουργούνται δυναμικά για να εντοπίσουν άλλους χρήστες και να υποστηρίξουν διαπραγμάτευση χρήστη-με-χρήστη όπως και εδραίωση session. Μόλις δύο ή περισσότεροι χρήστες συνδεθούν στο ίδιο κανάλι, μπορούν να επικοινωνήσουν και να ανταλλάξουν πληροφορίες για το session.

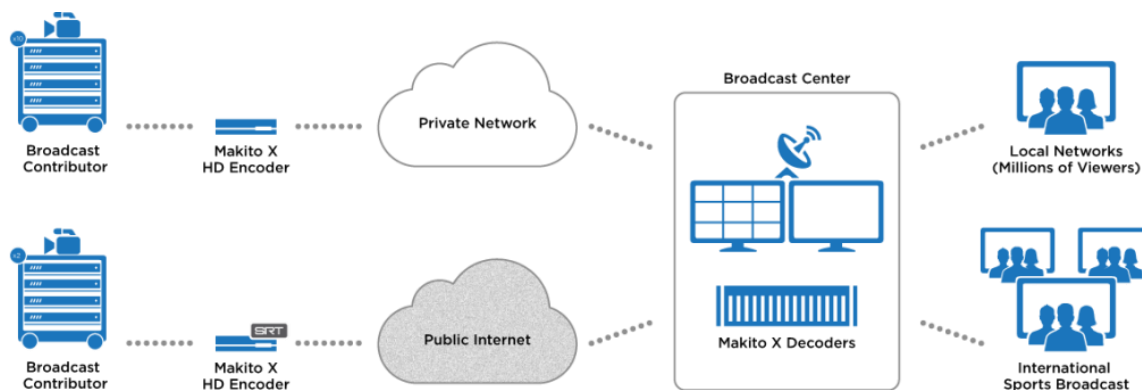
Στην ουσία, ο χρήστης που ξεκινά την σύνδεση στέλνει μια προσφορά κάνοντας χρήση πρωτοκόλλου σήματος Session Initiation Protocol(SIP) και Session Description Protocol(SDP), περιμένοντας απάντηση από οποιονδήποτε χρήστη είναι συνδεδεμένος στο κανάλι. Μόλις λάβει απάντηση, μια διαδικασία ξεκινά ώστε να βρεθεί ο καλύτερος Interactive Connectivity Establishment (ICE) υποψήφιος που έχει κάθε χρήστης. Μόλις οι καλύτεροι υποψήφιοι βρεθούν, όλα τα σχετικά αρχεία metadata, network routing και media information για κάθε χρήστη είναι δεκτά. Τότε το socket session μεταξύ των peers είναι πλήρως εδραιωμένο και ενεργό. Στη συνέχεια, ξεκινά η μετάδοση δεδομένων με τις ροές να δημιουργούνται και από τους δύο χρήστες, με δεδομένα να είναι δυνατό να μεταδοθούν και από τις δύο μεριές. Αν η διαδικασία επιλογής του καλύτερου ICE υποψηφίου αποτύχει, γίνεται χρήση του TURN server ως relay.



Πλαίσιο 1: WebRTC Peer-to-Peer σύνδεση

1.1.3 Secure Reliable Transport (SRT)

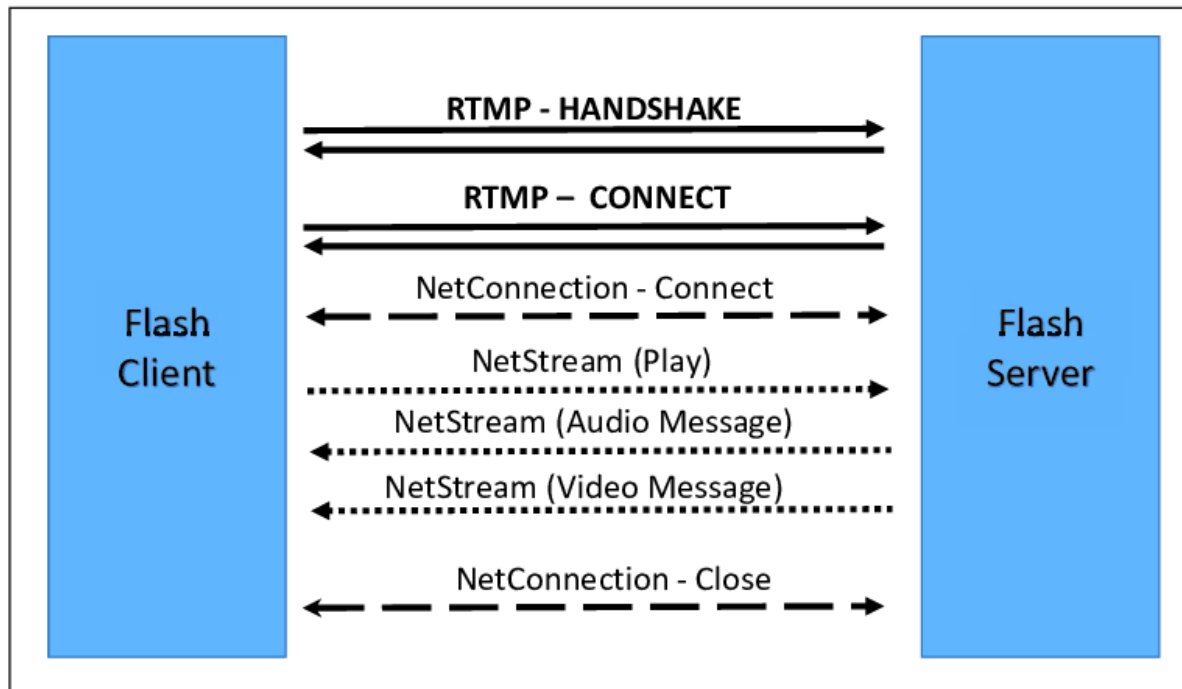
Το πρότυπο **Secure Reliable Transport (SRT)** [www.1.1.3-1] είναι ένα ανοιχτού λογισμικού πρωτόκολλο μεταφοράς βίντεο και technology stack, που έχει αναπτυχθεί από τις Haivision και Wowza. Το SRT κάνει χρήση ασφαλών streams και εύκολων firewall traversal για την βελτιστοποίηση της απόδοσης του streaming και για την διανομή υψηλής ποιότητας βίντεο ακόμη και σε αναξιόπιστα δίκτυα [[SRTspec](#)]. Το εκάστοτε αρχείο βίντεο (MPEG-4) αποστέλλεται στους κωδικοποιητές της wowza Makito X [www1134], αποστέλλεται στο public internet, στους αποκωδικοποιητές και στη συνέχεια στον τελικό χρήστη. Το πρωτόκολλο SRT επιτυγχάνει την άμεση και ασφαλή αποστολή των πακέτων βάζοντας χρονικό όριο αναμετάδοσης κάθε χαμένου πακέτου προσαρμοζόμενο συνέχεια με το latency του δικτύου. Όσον αφορά την ασφάλεια του περιεχομένου, το πρωτόκολλο χρησιμοποιεί 128/256 bit AES κρυπτογράφηση. Στο Πλαίσιο 2 δείχνεται σχηματικά η ροή περιεχομένου με τη συγκεκριμένη τεχνική.



Πλαίσιο 2: Διάγραμμα Secure Reliable Transfer(SRT) [www1133]

1.1.4 Real-Time Messaging Protocol (RTMP)

Το Real-Time Messaging Protocol (RTMP) [[RTMPspec](#)] σχεδιασμένο από την Adobe (Macromedia), σχεδιάστηκε κυρίως για μετάδοση βίντεο, ήχου και δεδομένων υψηλής αποδόσεως, μέσα στην πλατφόρμα Adobe Flash (π.χ. Adobe Flash Player, Adobe AIR) [www.1.1.4-1]. Το πρωτόκολλο σχεδιάστηκε για τη προσπάθεια μείωσης χρόνου απόκρισης και αποκοπή των ροών βίντεο ήχου σε κομμάτια, δίνοντας τους την δυνατότητα να μεταφέρονται σε μια σύνδεση. Πιο συγκεκριμένα, ο Flash player επικοινωνεί με τον Flash Media server και δημιουργείται μια RTMP σύνδεση. Έτσι, η RTMP σύνδεση ξεκινά την χειραψία. Ο Flash Player ζητά συγκεκριμένη ροή. Όταν το request φτάσει στο server, αποστέλλεται στο Flash SWF μέσω της RTMP σύνδεσης. Η ίδια ροή μπορεί να αποσταλεί σε όσους clients το ζητήσουν, αρκεί το bandwidth και από τις δύο μεριές να είναι αρκετό. Στο Πλαίσιο 3 δείχνεται σχηματικά η ροή περιεχομένου με τη συγκεκριμένη τεχνική.



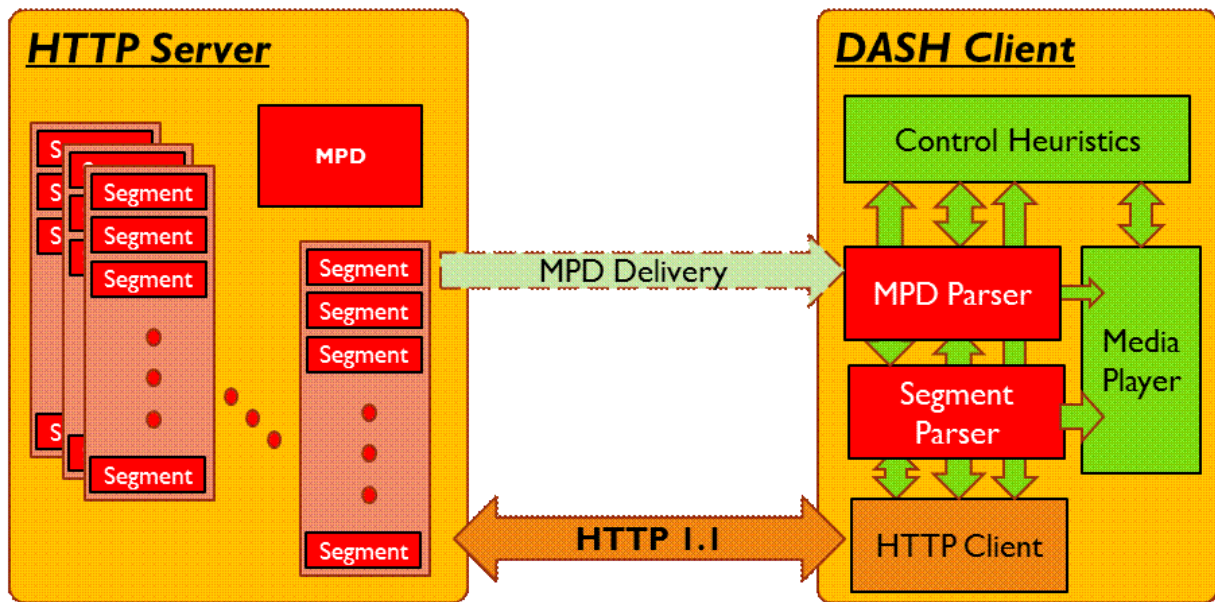
Πλαίσιο 3: Διάγραμμα Real-Time Messaging Protocol (RTMP) [\[www1142\]](http://www1142)

1.2 Low Latency MPEG-DASH σύστημα πάνω απο HTTP 2.0 με WebSocket

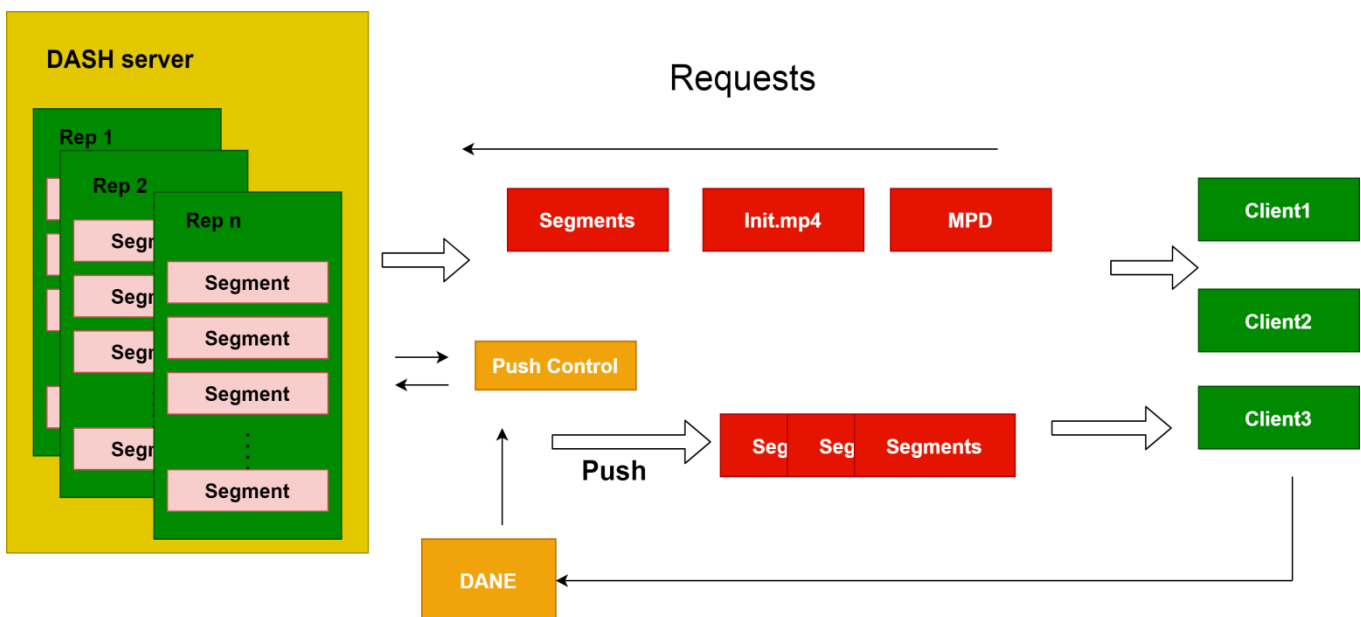
Η τεχνολογία DASH αναπτύχθηκε από την ομάδα MPEG και έγινε διεθνές πρότυπο [\[SODA2012\]](http://SODA2012) το 2012. Όπως φανερώνει το όνομά του υποστηρίζει δυναμικά προσαρμοζόμενες ροές περιεχομένου στο Διαδίκτυο με χρήση του πρωτοκόλλου HTTP και τείνει να εξοβελίσει τις διάφορες συναφείς εμπορικές τεχνικές όπως την IIS Smooth Streaming της Microsoft [\[ZAMB2009\]](http://ZAMB2009), την HLS (HTTP Live Streaming) της Apple, την HDS (HTTP Dynamic Streaming) της Adobe [\[www121\]](http://www121).

Πλέον το πρωτόκολλο αυτό έχει ανανεωθεί στέλνοντας μηνύματα [SAND](http://ISO/IEC_23009-5:2017) [\[ISO/IEC 23009-5:2017\]](http://ISO/IEC_23009-5:2017) ανάλογα με το bitrate του δικτύου, μέσω ενός WebSocket συνεχόμενα, για real-time υπολογισμό των νέων καταστάσεων του δικτύου, αλλάζοντας το bitrate.

Όπως βλέπουμε στο Πλαίσιο 4, το πολυμεσικό περιεχόμενο αποτελείται από MPD αρχεία και DASH segments, τα οποία αποθηκεύονται στον webserver μόλις δημιουργούνται. Τότε ο client απαιτεί τα MPD αρχεία, παίρνει το path των target segments σύμφωνα με την κατάσταση του δικτύου ζητώντας τα ένα-ένα στο παραδοσιακό σύστημα ροών. Στο σύστημά μας, εκμεταλλευόμαστε την push στο FDH πάνω από HTTP 2.0 με SAND messages για βελτίωση απόδοσης. Το σύστημα πλέον θα διαμορφωθεί ανάλογα με τα SAND μηνύματα που δέχεται από τον client, αλλάζοντας τα bitrates, όπως φαίνεται στο Πλαίσιο 5.



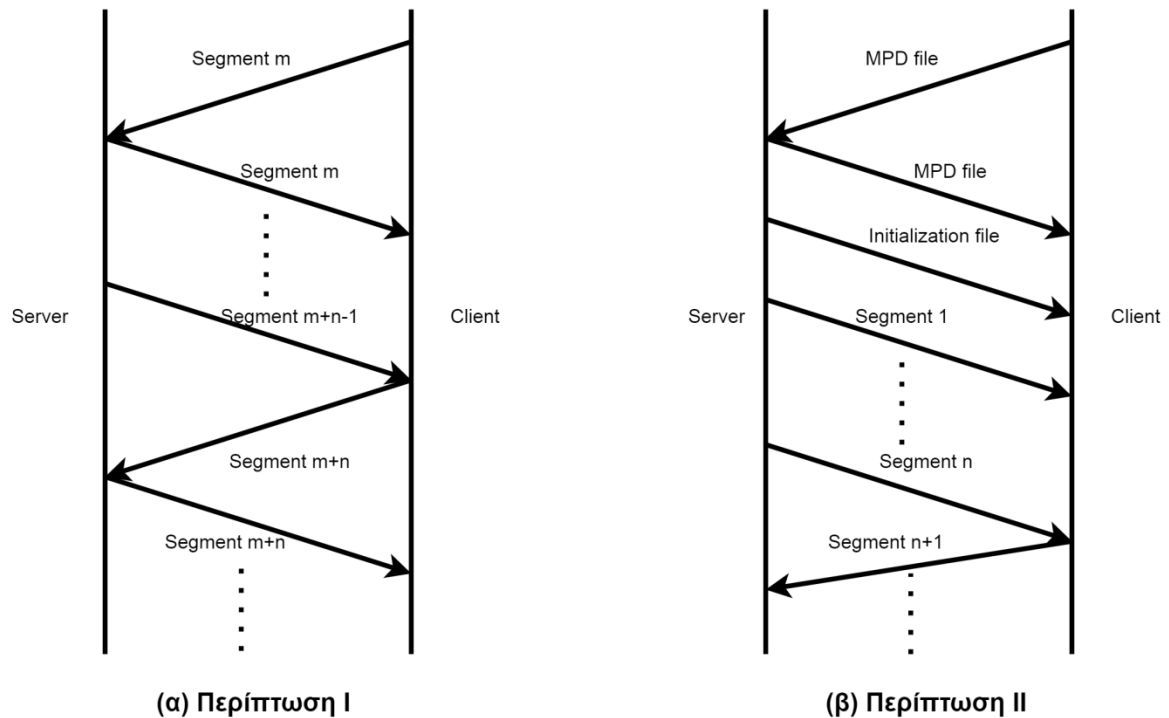
Πλαίσιο 4: Παραδοσιακό MPEG – DASH[SODA2012]



Πλαίσιο 5: Low – Latency MPEG-DASH με WebSocket[DASHoverWebSocket]

Ο server τότε θα κάνει push(HTTP 2.0)[www.1.2-2] τις πηγές στον client πριν στείλει τις απαιτήσεις για συγκεκριμένα segments, πραγματοποιώντας μείωση στο ζωντανό και start-up delay. Με αυτή τη μέθοδο είναι δυνατό να μειώσει τον αριθμό των συνδέσεων στο server και το μέγεθος του header που αποστέλλεται από τον client.

Για την μείωση του live latency, πρέπει να κάνουμε push ορισμένα segments ξανά και ξανά, με σκοπό την μείωση χρόνου στην πραγματοποίηση σύνδεσης όπως βλέπουμε στο Πλαίσιο 6a. Για το start-up delay, είναι απαραίτητο να γίνουν κάποιες αλλαγές στην αρχή του playback, που τις βλέπουμε στο Πλαίσιο 6b.



Πλαίσιο 6: (α) Μείωση live latency, (β) Μείωση startup delay [DASH over WebSocket]

Αφού λάβει το Manifest, ο πελάτης αναλύει το αρχείο και τα αιτήματα για τα κομμάτια σε σωστά bitrates. Ο διακομιστής ιστού θα ωθήσει όχι μόνο την αιτούμενη, αλλά και έναν ορισμένο αριθμό επόμενων τμημάτων με την ίδια ποιότητα. Μετά την ανάγνωση από το buffer, ο πελάτης θα επαναλάβει τη διαδικασία αίτησης. Αυτό μπορεί προφανώς να μειώσει τον χρόνο αναμετάδοσης (RTT) σε ζωντανή ροή. Λαμβάνοντας υπόψη ότι η push στρατηγική θα χρησιμοποιηθεί ως επί το πλείστον σε ζωντανή ροή, όπου οι χρήστες παρακολουθούν συνεχώς βίντεο. Συνεπώς, δεν θα υπάρξει πλεονασμός δεδομένων, διότι θα διατεθούν όλοι οι πόροι.

Για να μειωθεί η καθυστέρηση έναρξης, μπορούμε να κάνουμε push τις πηγές μας, εφόσον ο server λάβει αίτηση για Media Presentation Description (MPD). Μόλις ο διακομιστής web ανιχνεύσει το αίτημα, ωθεί το αρχείο mp4 αρχικοποίησης και μερικά τμήματα βίντεο και ο πελάτης μπορεί να αρχίσει να αναπαράγει το βίντεο αμέσως. Παρόλο που δεν μπορούμε να παίξουμε τα τμήματα σε άριστη ποιότητα ώστε να χρησιμοποιήσουμε στο έπακρο το εύρος ζώνης, μειώνουμε την καθυστέρηση εκκίνησης για να έχουμε καλύτερη εμπειρία για τους χρήστες. Εκμεταλλευόμενος την τεχνολογία pull, ο client θα διαλέξει την καλύτερη ποιότητα κομματιών ανάλογα με το bandwidth. Έπειτα γίνεται χρήση της τεχνολογίας push και πρέπει να επιλέξουμε κατάλληλη ποιότητα για να μεταφέρουμε το μήνυμα. Διαλέγουμε το WebSocket καθώς έχει υψηλή απόδοση σε real-time μεταφορά των SAND messages.

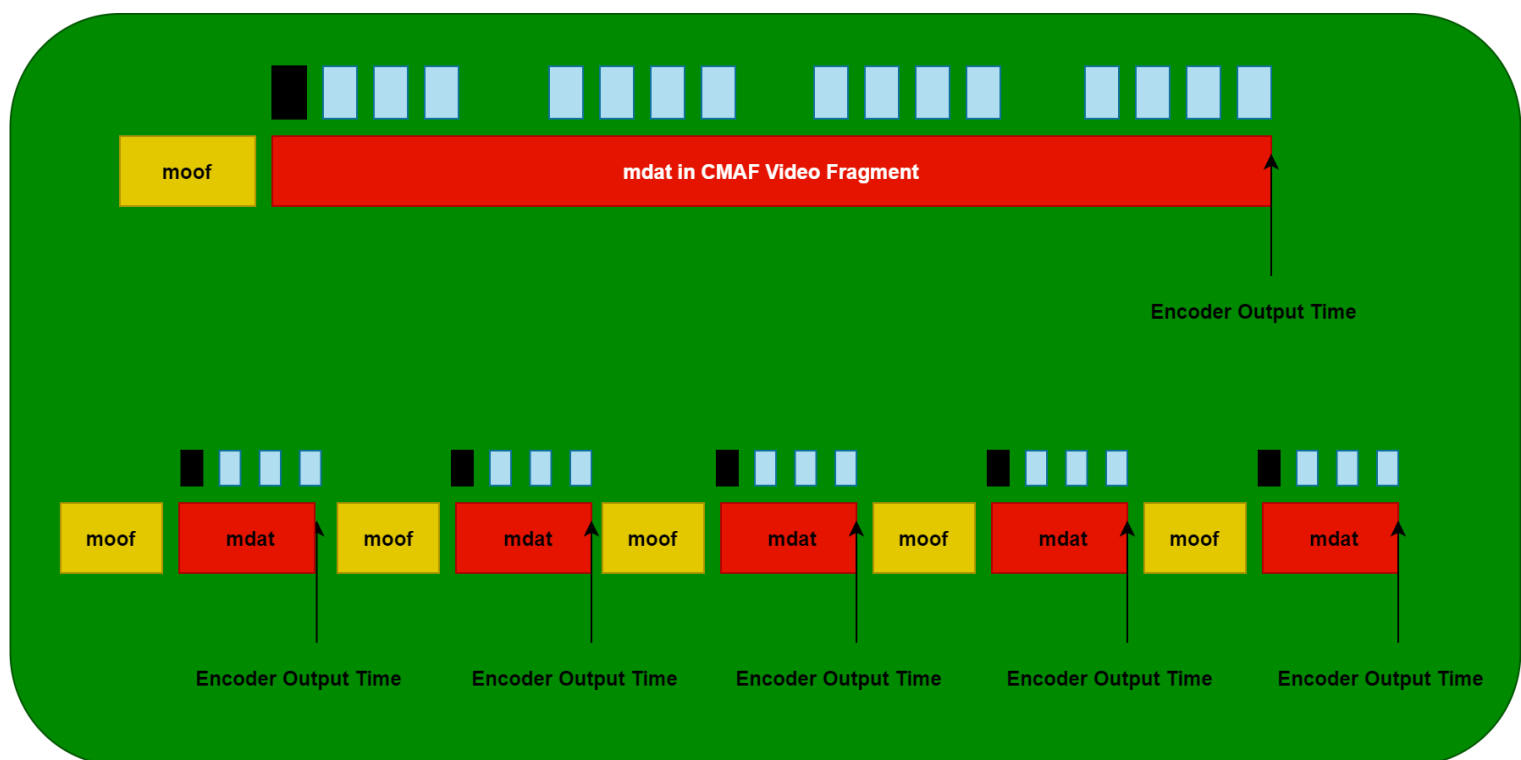
1.3 Common Media Application Format (CMAF) περιεχομένου με Δυναμικές Ροές πάνω από HTTP(DASH)

Τι είναι το CMAF[ISO/IEC 23000-19]:

Το Common Media Application Format (CMAF)[[ISO/IEC 23009-5:2017](#)] ορίζει το container που περιέχει τον ήχο και το βίντεο. Δεν είναι ακόμη ένα μέσο παρουσίασης, όπως η Δυναμική Προσαρμοσμένη ροή μέσω HTTP(DASH) ή το HTTP Live Streaming (HLS). Το CMAF είναι μια κωδικοποίηση και τυποποίηση του υπάρχοντος κατατμημένου MP4, με ορισμένες βελτιώσεις. Το CMAF είναι μια προφανής ανεξάρτητη κωδικοποίηση κατάλληλη για DASH και HLS, και έχει σχεδιαστεί τόσο για on-demand αλλά και live streaming στο μυαλό.

Μετά από συνεργασία 18 μηνών, Microsoft και Apple μοιράστηκαν με τον κόσμο την ιδέα τους για κοινό media format μεταξύ DASH και HLS. Το νέο αυτό format Προτάθηκε στην 114^η συνάντηση του MPEG, τον Φεβρουάριο του 2016.

Το πρώτο απαραίτητο στοιχείο για το CMAF είναι το chunk encoding “chunk ορίζεται ως η πιο μικρή μονάδα που περιέχει ένα moof και ένα mdat άτομο – όπως φαίνεται στο πλαίσιο αποκάτω”. Ένα ή περισσότερα κομμάτια συνδυάζονται για την δημιουργία ενός fragment και πολλαπλά fragment για την δημιουργία ενός segment. Φυσικά, ένα ξεχωριστό header απαιτείται για την αρχικοποίηση του playback από τον αποκωδικοποιητή.



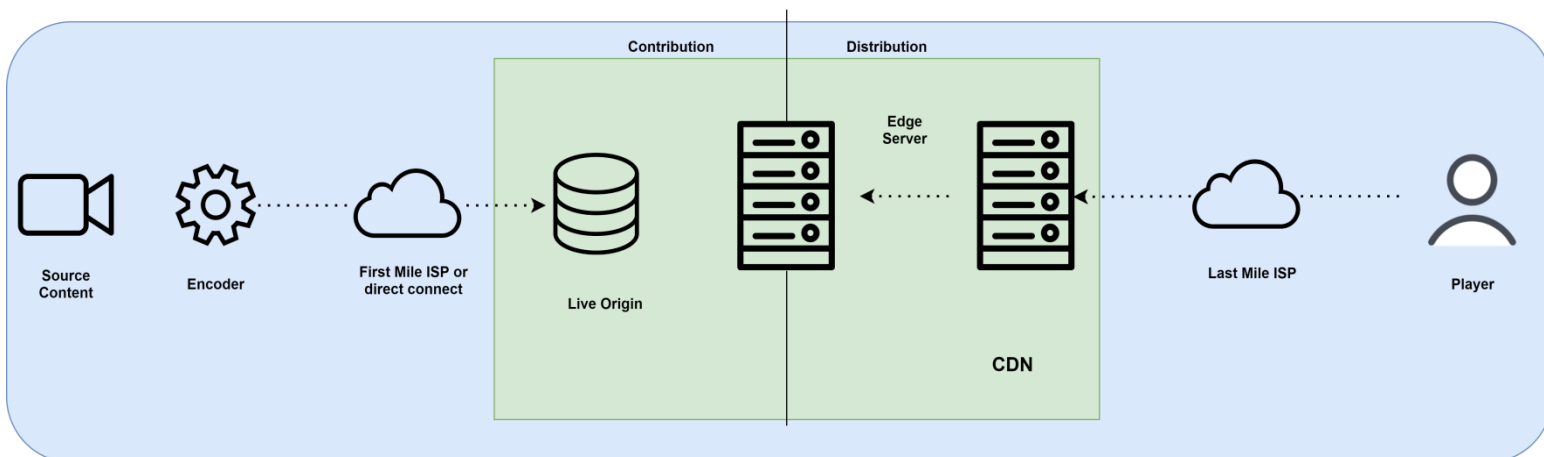
Πλαίσιο: CMAF fragment

Έπειτα από την δημιουργία του segment, ο κωδικοποιητής δεν γνωρίζει το τελικό μέγεθος που θα παραχθεί. Έτσι, ο κωδικοποιητής κάνει HTTP POST, το segment σε έναν ingest server. Το ingest επίπεδο του server (CDN) λαμβάνει το περιεχόμενο και περιέχει και ένα mid-tier επίπεδο που μπορεί να διανέμει το περιεχόμενο.

Στη μεριά του distribution, ένας media player “τραβάει” περιεχόμενο (HTTP GET) από έναν edge server, ο οποίος με την σειρά του το τραβάει από τον αρχικό server. Τέλος, ο κωδικοποιητής θα χρησιμοποιήσει το chunked transfer encoding (HTTP 1.1) για να στείλει το κωδικοποιημένο CMAF chunk στον αρχικό server για αναδιανομή.

Πλαίσιο: CDN για CMAF[CMAFakam]

2. Μελέτη τεχνικών για την αρχειοθέτηση τηλεοπτικού περιεχομένου από μεγάλους τηλεοπτικούς σταθμούς



Την τελευταία δεκαετία το τηλεοπτικό περιεχόμενο έχει κάνει δυνατή την μετάβαση από το αναλογικό στο ψηφιακό σύστημα. Οι τηλεοπτικοί σταθμοί θέλοντας να συμβαδίσουν με τις ταχύτητες που κινείται το online streaming, και η iptv, δεν επαναπαύθηκαν με την ψηφιοποίηση του περιεχομένου και την χρήση υψηλής ανάλυσης - ποιότητας βίντεο και ήχου, αλλά έκαναν δυνατή την προβολή του περιεχομένου τους κατά την επιθυμία των χρηστών.

Η Video On – Demand εποχή έχει φτάσει, κάνοντας δυνατή την προσωποποίηση επιλογών προβολής για κάθε χρήση, εκμεταλλευόμενοι την ευρεία χρήση των metadata όπως και τεχνικών αναζήτησης βάση ηλικίας, διάθεσης και παλιότερων αναζητήσεων.

Οι δομές που επιτρέπουν αυτή την ευελιξία, είναι σε εξέλιξη αυτή τη στιγμή, βελτιώνοντας τα αποτελέσματά τους με τη βοήθεια τεχνικών machine learning(mood metadata)[WHP231], speech-to-text recognition(COMMA, ABC-IP)[www.2-1][www.2-2][WHP231].

2.1 Ανάγκη για αρχειοθέτηση

Η αρχειοθέτηση περιεχομένου πολυμέσων, όσο σημαντική είναι για μια προσωπική βιβλιοθήκη που έχει στην χειρότερη των περιπτώσεων ελάχιστους διαφορετικούς τύπους πολυμέσων(Βιντεοκασέτα, Compact-Disk, DVD, Blue-Ray, Digital Content), τόσο σημαντικότερη είναι για τηλεοπτικούς σταθμούς που έχουν μια τεράστια βιβλιοθήκη κασετών(tapes) βίντεο και ήχου (όπως το BBC), που έχουν περιεχόμενο

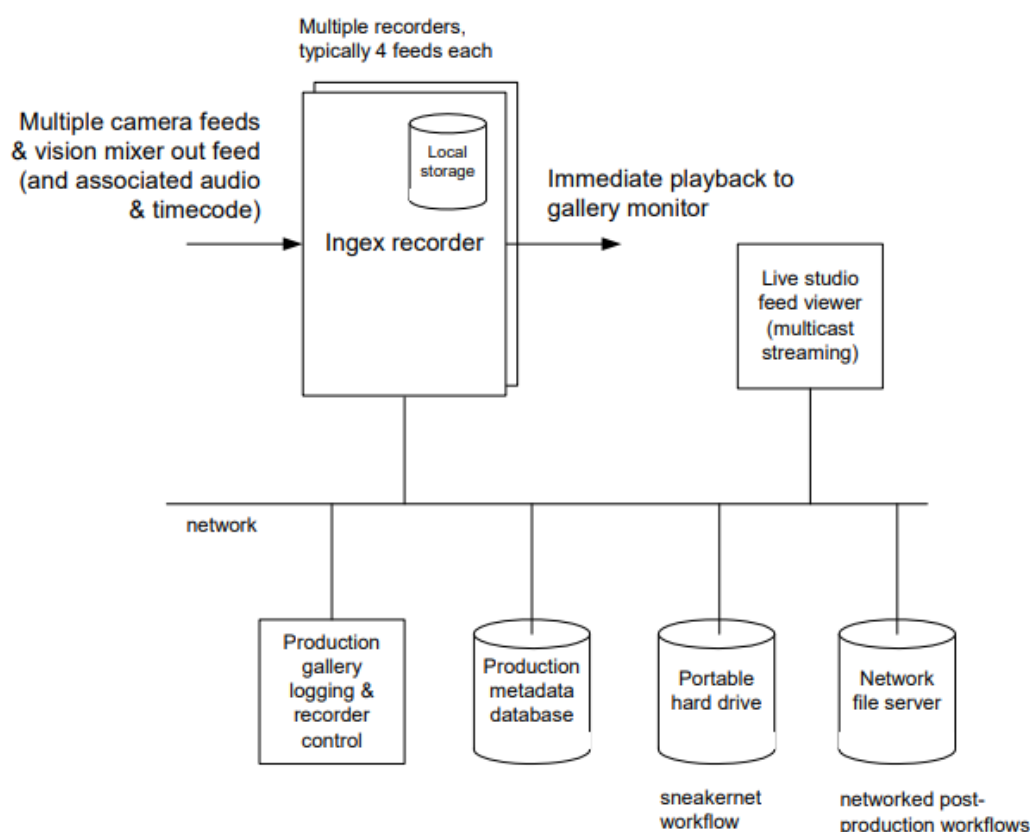
χιλιάδων ωρών και πρέπει με κάποιο τρόπο να δημιουργήσει μια συνέχεια για τα χρόνια που έρχονται και την λεγόμενη «επανάσταση» της ψηφιοποίησης.

Το BBC ξεκινώντας με περίπου 650χιλιάδες ώρες βίντεο και 350 χιλιάδες ώρες ήχου κατάφερε να ψηφιοποιήσει συστηματικά το περιεχόμενό της δημιουργώντας το σύστημα IngeX.

2.2 Σύστημα Ψηφιοποίησης IngeX(BBC R&D)

Το BBC R&D IngeX[[www.2.2-1](#)]είναι ένα σύστημα εγγραφής «χωρίς κασέτα» που αναπτύχθηκε το 2005, για πολυκάμερες παραγωγές τηλεόρασης, που συνδυάζει υψηλής ποιότητας πολυκάμερη εγγραφή με μια σειρά αυτοματοποιημένων λειτουργιών, έτσι ώστε να επιταχύνει χρονοβόρες εργασίες για τατωρινά δεδομένα των τηλεοπτικών σταθμών[[WHP186](#)].

Τα σημαντικότερα σημεία του IngeX φαίνονται στο παρακάτω πλαίσιο:



Πλαίσιο 7: Δομή συστήματος IngeX[[WHP186](#)]

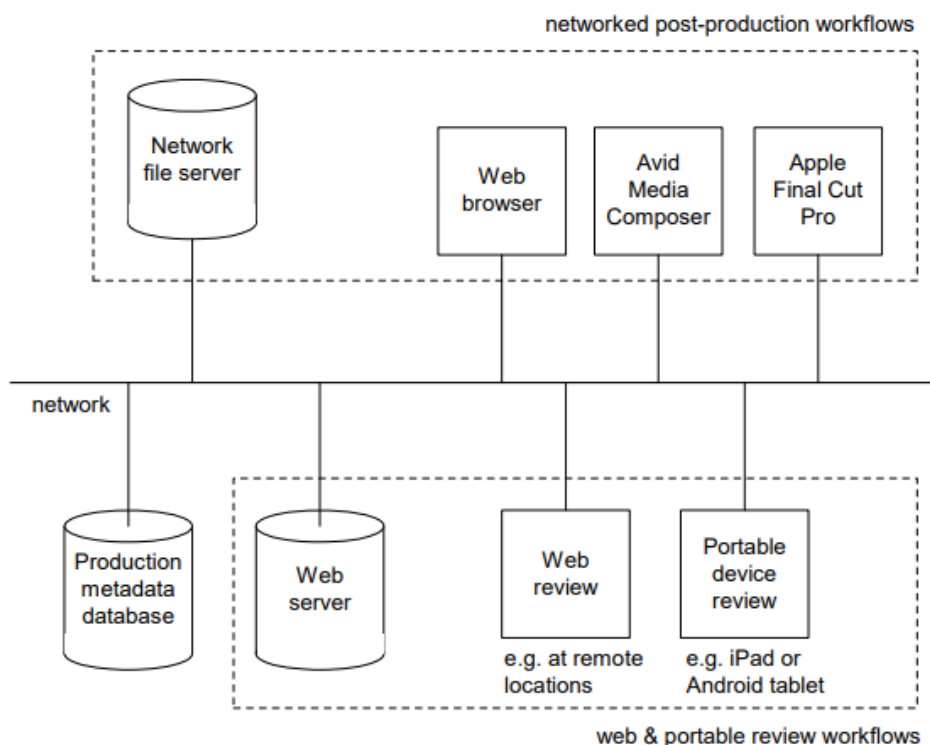
Όπως είναι φανερό από το σχήμα, το σημαντικότερο σημείο είναι ο εγγραφέας IngeX, που εγγράφει κατευθείαν στο δίσκο, εκμεταλλευόμενος την ιπποδύναμη PC με πολυπύρρηνο επεξεργαστή, πολλαπλές SDI/HD-SDI κάρτες εγγραφής και του λογισμικού IngeX. Η δουλειά του IngeX είναι να κάνει εγγραφή του περιεχομένου στην

μέγιστη δυνατή ποιότητας και να τα κωδικοποιεί «ζωντανά» σε formats όπως το MXF που θα δούμε αναλυτικά παρακάτω.

Το σύστημα Ingexέρχεται μαζί με τις εφαρμογές Ingex Player(λογισμικό αναπαραγωγής MXF αρχείων), Ingex Archive(σύστημα διαχείρισης βίντεο κασετών), Media Harmony(Module που επιτρέπει) και libXMF(C βιβλιοθήκη για δημιουργία MXF αρχείων).

Ακόμη, δίνει την δυνατότητα μετά την εγγραφή, και στην φάση του post-production να μπορεί να επικοινωνεί με το υπόλοιπο δίκτυο, διαμοιράζοντας τα MXF αρχεία και δίνοντας την δυνατότητα σε ολόκληρο το δίκτυο του studio να μπορεί να κάνει χρήση του αρχείου και των metadata.

Παρακάτω φαίνεται η δικτυακή δομή του περιεχομένου σύμφωνα με το σύστημα Ingex



Πλαίσιο 8: Δομή διαμοιρασμού αρχείων στο σύστημα Ingex[[WHP186](#)]

Μπορεί το BBC να έχει τα πρωτεία και τις γνώσεις από τη δημιουργία του Ingex συστήματος, όμως δεν διστάζει να διαμοιράσει την τεχνογνωσία κάνοντας το λογισμικό ανοιχτού κώδικα. Πλέον πολλές παραγωγές και οργανισμοί έχουν χτίσει το σύστημά τους σύμφωνα με το Ingex, κινητοποιημένοι από το χαμηλό κόστος.

2.3 Material Exchange Format (MXF)

Για τις ανάγκες ψηφιοποίησης του τηλεοπτικού περιεχομένου, το BBC (BBC R&D) ξεκίνησε στις αρχές του 2007, την δημιουργία του Material eXchange Format (MXF), μέσω του Ingex συστήματος για την ψηφιοποίηση του τηλεοπτικού περιεχομένου, όπως παρατέθηκε παραπάνω. Τα MXF αρχεία δημιουργήθηκαν έτσι ώστε να απλοποιηθεί η μεταφορά αρχείων βίντεο και ήχου, μεταξύ συσκευών και οργανισμών.

Τα αρχεία MXF που παράγονται από το σύστημα αρχειοθέτησης Ingex ακολουθώντας προδιαγραφές του αρχείου BBC Archive MXF, οι οποίες περιγράφονται λεπτομερώς από τα White Papers της BBC R&D [[WHP241](#),[WHP096](#)]. Στην ιστοσελίδα του Ingex αρχείου μπορείτε να δείτε ένα παράδειγμα αρχείου και άλλες πληροφορίες [[www221](#)].

Περιεχόμενο και δομή MXF

Το αρχείο ως RAW βίντεο δεδομένα είναι φυσικό να είναι μεγάλο. Για να δώσουμε ένα παράδειγμα, μια ώρα ενός BBC Archive MXF αρχείου ισοδυναμεί περίπου με 75GB, αν περιέχει 8-bit κομμάτια (8-bit κομμάτια-> σημαίνει δεδομένα βίντεο 720x576, 4:2:2).

Ένας σημαντικός παράγοντας των MXF αρχείων είναι τα metadata, που θα κάνουν δυνατή την εύρεση των σχετικών αρχείων και την κατηγοριοποίηση του στους server του τηλεοπτικού σταθμού.

Αναφορικά τα πεδία Μέτα - δεδομένων που προσκολλούνται στα MXF αρχεία είναι:

• Μορφή	Ημερομηνία Αποθεματοποίησης
• Τίτλος Προγράμματος	Περιγραφέας spool
• Τίτλος Επεισοδίου	Υπενθύμιση
• Ημερομηνία TX	Διάρκεια Προγράμματος
• Πρόθεμα Χωρητικότητας	Αριθμός Spool
• Αριθμός Προγράμματος	Αριθμός Πρόσβασης
• Κωδικός Παραγωγής	Λεπτομέρειες Καταλόγου
• Κατάσταση Spool	Αριθμός Αντικειμένου

2.4 Τεχνικές αναζήτησης αρχείου μέσω Μέτα -δεδομένων (αναφορικά - BBC)

Μπορεί, μέσω της διαδικασίας του Ingex και άλλων συστημάτων να επιτυγχάνουμε την αρχειοθέτηση και αρχειοποίηση των διάφορων τηλεοπτικών παραγωγών, όμως για να είμαστε σε θέση να ανακαλύψουμε το επιθυμητό περιεχόμενο που έχει

αποθηκευτεί στους servers, είναι απαραίτητο να έχουμε κάποιο από τα αναπτυσσόμενα συστήματα ταξινόμησης.

2.4.1 Σύστημα Automatic Broadcast Content Interlinking Project (ABC-IP)

Το BBC σε συνεργασία με το MetaBroadcast, ανέπτυξε μια εξελιγμένη τεχνική επεξεργασίας κειμένου, έτσι ώστε να συνδέσει διάφορες πηγές Μετα-Δεδομένων από μεγάλες συλλογές βίντεο και ήχου. Το project είναι εν μέρει χρηματοδοτούμενο από το Technology Strategy Board (TSB).

Το σημείο εκκίνησης ήταν η τεράστια βιβλιοθήκη ήχου του World Service. Το World Service μεταδίδει από το 1932, οπότε πήρε την πλούσια βιβλιοθήκη ανθρώπων, αντικειμένων και οργανισμών, μέσα σε ένα μεγάλο εύρος κατηγοριών.

Μία από τις πρώτες καινοτομίες του έργου ήταν να βελτιωθεί ο τρόπος με τον οποίο μπορούν να αντληθούν θεματικές ετικέτες από αυτοματοποιημένες μεταγραφές speech-to-text, κάτι που έδωσε ένα εντελώς νέο σύνολο μεταδεδομένων. Ακόμη, έγινε βελτιστοποίηση σε διάφορους αλγορίθμους για να είναι δυνατή η ανάλυση μεταγραφών με υψηλό ρυθμό σφάλματος λέξεων που η αναγνώριση ομιλίας δημιουργεί.

2.4.2 Speech-to-text Συστήματα - Σύστημα COMMA (BBC R&D)

Παίρνοντας ως παράδειγμα για τα speech-to-text συστήματα, αναλύουμε την πλατφόρμα COMMA. Το COMMA [www21] κατασκευάζεται από μηχανικούς της ομάδας BBC R&D, του Internet Research&Future Services. Η τεχνολογία και η σκέψη πίσω από αυτήν οδηγούν επί του παρόντος αρκετές πρωτοβουλίες του BBC, όπως το Πρωτότυπο του BBC WorldService.

Το COMMA βασίζεται σε ένα σύνολο αλγορίθμων επεξεργασίας ομιλίας που ονομάζεται Kiwi. Το 2011, η BBC R&D έλαβε πρόσβαση σε ένα πολύ μεγάλο ραδιοφωνικό αρχείο, το αρχείο του BBC World Service, το οποίο είχε τότε πολύ λίγα μεταδεδομένα. Προκειμένου να οικοδομηθεί το πρωτότυπο γύρω από αυτό το αρχείο, αναπτύχθηκε ένας αλγόριθμος επεξεργασίας ομιλίας, επαναχρησιμοποιώντας δομικά στοιχεία ανοιχτού κώδικα όπου είναι δυνατόν. Στη συνέχεια, δημιουργήθηκε η ακόλουθη ροή εργασίας από αυτούς τους αλγορίθμους:

- Διαχωρισμός ομιλητών, με ταυτοποίηση και ανίχνευση φύλου (με τη χρήση εργαλειοθήκης LIUM diarization toolkit, diarize-jruby και ruby-lsh). Αυτή η διαδικασία είναι επίσης γνωστή ως διάρθρωση. Ουσιαστικά ένα αρχείο ήχου χωρίζεται αυτόματα σε τμήματα ανάλογα με την ταυτότητα του ομιλητή. Ο αλγόριθμος μπορεί να μας δείξει ποιος μιλάει και σε ποιο σημείο του κλιπ ήχου.
- Speech-to-text για τα τμήματα ομιλίας που ανιχνεύθηκαν (χρησιμοποιώντας CMF Sphinx). Σε αυτό το σημείο ο προφορικός ήχος μεταφράζεται όσο το δυνατόν ακριβέστερα σε αναγνώσιμο κείμενο. Αυτός ο αλγόριθμος

χρησιμοποιεί μοντέλα κατασκευασμένα από ένα ευρύ φάσμα δεδομένων του BBC.

- Αυτόματη προσθήκη ετικετών με αναγνωριστικά DBpedia. Το DBpedia είναι μια μεγάλη βάση δεδομένων με δομημένα δεδομένα που προέρχονται από τη Wikipedia. Η διαδικασία αυτόματης προσθήκης ετικετών δημιουργεί τα μεταδεδομένα που μπορούν να αναζητηθούν και τα οποία τελικά μας επιτρέπουν να έχουμε πρόσβαση στα αρχεία πολύ πιο εύκολα. Αυτή η διαδικασία χρησιμοποιεί ένα εργαλείο με την ονομασία «Mango»[[www2422](#)].

Η πλατφόρμα COMMA στην πράξη

Επί του παρόντος, το COMMA χρησιμοποιείται εσωτερικά από το BBC. Όπως και με το εκτεταμένο αρχείο της Παγκόσμιας Υπηρεσίας, οι δημοσιογράφοι στην αίθουσα ειδήσεων έχουν μια τεράστια, διαρκώς αυξανόμενη βάση δεδομένων με υλικό που έχουν στη διάθεσή τους. Πολλά από αυτά μπορεί να μην έχουν το είδος των χρήσιμων μεταδεδομένων που καθιστούν εύκολη την εύρεση και χρήση τους, ειδικά όταν εκδηλώνονται ειδήσεις ζωντανά. Το COMMA είναι σε θέση να λαμβάνει αυτόματα τον ήχο από το βίντεοκλιπ, να αναγνωρίζει τα διάφορα ηχεία και να παρέχει μια μεταγραφή των ειδήσεων. Και πάλι, το εργαλείο Mango στη συνέχεια προσθέτει τα σημαντικά μετα-δεδομένα στο μέσο μέσα στη βάση δεδομένων. Η COMMA συμβάλλει στην όσο το δυνατόν πιο αποτελεσματική ροή εργασίας στον ειδησεογραφικό χώρο, φέρνοντας το πιο σχετικό υλικό στους δημοσιογράφους και όσο το δυνατόν γρηγορότερα.

2.4.3 Δημιουργία Μετα-Δεδομένων βάσει Διάθεσης (Mood Metadata)

Το πρότυπο Μέτα-Δεδομένων βάσει διάθεσης(Mood Metadata), αναπτύσσεται και εξελίσσεται ακόμη και σήμερα από την ομάδα του BBC[[www.2.4.3-1](#),[www.2.4.3-1](#),[WHP231](#)] R&D και κατηγοριοποιεί τα πρόγραμμα του τηλεοπτικού σταθμού για ποιοτικότερη και ευκολότερη χρήση από τους χρήστες του iPlayer.

Η ομάδα ανάπτυξης έχει δημιουργήσει αλγορίθμους που αναλύουν τον ήχο και την εικόνα του προγράμματος, και δημιουργούν αυτόματα mood metadata για κάθε πρόγραμμα.

Για να γίνει δυνατός αυτός ο τρόπος εξερεύνησης περιεχομένου, χρειάζονται υψηλής ποιότητας metadata, που θα είναι χρήσιμα και ευκατανόητα από το κοινό. Καθώς το μέγεθος του περιεχομένου σε έναν τηλεοπτικό σταθμό όπως το BBC είναι μεγάλο, θα χρειαστεί μια αυτοματοποιημένη τεχνική για την παραγωγή υψηλής ποιότητας Μέτα – Δεδομένων.

Πως λειτουργεί

Ο αλγόριθμος εξάγει υψηλής ποιότητας Μέτα – Δεδομένων, με τεχνικές ανάλυση σημάτων χαμηλής ποιότητας ήχου και βίντεο και κατηγοριοποιώντας τα με χρήση machine learning.

Για να ξεκινήσει η εξαγωγή συμπερασμάτων με machine learning, χρειάστηκε μια αρχική βάση με αποτελέσματα. Η αρχική βάση αποτελεσμάτων εξήχθη από μια

μεγάλη έρευνα από 200 άτομα, οι οποίοι είδαν μικρά κομμάτια τηλεοπτικών προγραμμάτων, και έπειτα τα αξιολόγησαν βάση διάθεσης.

Στα τελικά αποτελέσματα, τα moods μειώθηκαν σε δύο βασικές διαστάσεις. Η πρώτη σχετίζεται με τη **σοβαρότητα/κωμικότητα του προγράμματος**, και η δεύτερη περιγράφει τον ρυθμό(**χαμηλού ρυθμού/γρήγορου ρυθμού**).

Με τη βοήθεια των τεχνικών machine learning, οι δύο διαστάσεις αξιολογήθηκαν με βαθμό υψηλής ακρίβειας, αξιολογώντας πάνω από 95% των προγραμμάτων με καθαρή διάθεση(mood).

Βιβλιογραφία – Αναφορές

[www.1-1] <https://developer.apple.com/streaming/> (τελευταία πρόσβαση: Σεπτέμβριος 2019)

[www.1.1.1-1] https://developer.apple.com/documentation/http_live_streaming/protocol_extension_for_low-latency_hls_preliminary_specification (τελευταία πρόσβαση: Σεπτέμβριος 2019)

[SODA2011] IrajSodagar, “[Overview of MPEG DASH](#)”, 2011 (τελευταία πρόσβαση: Σεπτέμβριος 2019)

[SODA2012] IrajSodagar, “[MPEG-DASH: The Standard for Multimedia Streaming Over Internet](#)”, White paper on MPEG-DASH Standard, April 2012 (τελευταία πρόσβαση: Σεπτέμβριος 2019)

[DASHoverWebSocket] <https://www.semanticscholar.org/paper/Low-Latency-MPEG-DASH-System-Over-HTTP-2.0-and-Wu-Zhao/4808706e51e22947be7ae71dc0dccbaec6a3b45d>

[CMAFakam] <https://www.akamai.com/us/en/multimedia/documents/white-paper/low-latency-streaming-cmaf-whitepaper.pdf> ,White paper on Low-Latency streaming over DASH

[www.1.1.2-1] <https://webrtc.org/> , Πληροφορίες για το περιεχόμενο του WebRTC από την επίσημη ιστοσελίδα

[www.1.1.3-1] <https://www.haivision.com/resources/streaming-video-definitions/srt/> , (Τελευταία επίσκεψη Σεπτέμβριος 2019)

[www.1.1.4-1] <https://www.adobe.com/devnet/rtmp.html> ,(Τελευταία επίσκεψη Σεπτέμβριος 2019)

[www.1.1.4-2] https://www.researchgate.net/figure/RTMP-Streaming-Session_fig1_272376917

[www.1.1.3-2] <https://www.srtalliance.org/> ,(Τελευταία επίσκεψη Σεπτέμβριος 2019)

[www.1.1.3-3] <http://masterdigitalmedia.com/industries/media-and-broadcast.html>

[www.1.1.3-4] https://www.haivision.com/wp-content/uploads/DS_Makito_X-Harsh_17.2.1.4.pdf

[SRTspec]

https://github.com/Haivision/srt/files/2489142/SRT_Protocol_TechnicalOverview_DR_AFT_2018-10-17.pdf

[RTMPspec]

http://www.images.adobe.com/content/dam/Adobe/en/devnet/rtmp/pdf/rtmp_specification_1.0.pdf, Adobe's Real Time Messaging Protocol

[www.1.2-1] <https://www.adobe.com/devnet/hds.html>, (Τελευταία επίσκεψη Σεπτέμβριος 2019)

[ISO/IEC 23009-5:2017] Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 5: Server and network assisted DASH (SAND)

[www.1.2-2] <https://developers.google.com/web/fundamentals/performance/http2>, (Τελευταία επίσκεψη Σεπτέμβριος 2019)

[www.2.1-1] <https://www.bbc.co.uk/rd/blog/2011-11-abc-ip-and-work-on-audio-archi>, (Τελευταία επίσκεψη Σεπτέμβριος 2019)

[www.2-1] <https://www.bbc.co.uk/rd/projects/comma>, (Τελευταία επίσκεψη Σεπτέμβριος 2019)

[www.2.2-1] <http://ingex.sourceforge.net/>, (Τελευταία επίσκεψη Σεπτέμβριος 2019)

[www.2.4.3-1] <http://moods.ch.bbc.co.uk/>, (Τελευταία επίσκεψη Σεπτέμβριος 2019)

[www.2.4.3-2] <https://www.bbc.co.uk/rd/projects/mood-metadata>, (Τελευταία επίσκεψη Σεπτέμβριος 2019)

[www.2.4.2-2] <https://www.w3.org/TR/webRTC/>

[WHP155] White Paper about File-based Production: Making It Work In Practice Stuart Cunningham and Philip de Nier

[WHP241] White Paper about A Guide to Understanding BBC Archive MXF Files M. Glanville and T. Heritage

[WHP096] White Paper about A Guide to Understanding BBC Archive MXF Files M. Glanville and T. Heritage

[WHP186] White Paper about Developments in Automated Tapeless Production for Multi-Camera Programmes P.N. Tudor, J.S. McKinnell, N.P. Pinks, S. Casey

[WHP231] A Pilot Study for Mood-based Classification of TV Programmes Jana Eggink, Penelope Allen, Denise Bland

[ZAMB2009] Alex Zambelli "[IIS Smooth Streaming Technical Overview](#)", 2009 (τελευταία πρόσβαση Σεπτέμβριος 2019)

[[ΠΟΛΥΜ](#)] Σημειώσεις Επικοινωνίες Πολυμέσων Μέρος Ι