

customME

presenta

Progetto DiAA

Digital Audio Archive

Management System of Audio Archive

Indice generale

1. Informazioni preliminari.....	2
1.1. Nota sui diritti di proprietà.....	2
1.2. Nota sui marchi.....	2
1.3. Documentazione e supporto.....	2
2. Introduzione.....	3
2.1. Scopo.....	3
2.2. Destinatari.....	3
2.3. customME Profile.....	3
3. Descrizione del Sistema.....	4
4. Studio d'esempio.....	6
4.1. Descrizione delle specifiche.....	6
4.2. Sorgenti dati del sistema.....	8
4.3. Encoder audio compresso e non compresso.....	8
4.4. Web server (Front End).....	8
4.5. Storage compresso.....	8
4.6. DB Server.....	9
4.7. Storage non compresso.....	9
4.8. Digitalizzazione audio.....	9
5. Architettura del Sistema.....	10
5.1. Hardware.....	10
5.2. Software.....	10
6. Quotazione.....	11
7. Servizi professionali customME.....	12
7.1. Assistenza tecnica customME.....	12

1. Informazioni preliminari

1.1. Nota sui diritti di proprietà

Tutti i diritti riservati. Questo documento è di proprietà di customME - Elvio Magliocca. E' vietato riprodurlo, trasmetterlo, trascriverlo, tradurlo in parte o integralmente o depositarlo in un sistema di archiviazione, con qualsiasi mezzo, elettronico, meccanico, magnetico, ottico, chimico, manuale o qualsivoglia, senza esplicito consenso scritto da Elvio Magliocca.

1.2. Nota sui marchi

I prodotti citati in questo documento sono marchi registrati dei rispettivi fabbricanti.

1.3. Documentazione e supporto

Per quanto il contenuto di questo documento sia stato ampiamente verificato, customME non garantisce che sia privo di errori od omissioni e si riserva il diritto di sottoporlo ad aggiornamenti e revisioni ed apportarvi correzioni o modificare le informazioni in esso contenute. Per ulteriori informazioni contattare:

customME

email: info@elvio.org

2. Introduzione

2.1. Scopo

Scopo del presente documento è fornire una indicazione tecnica/economica della soluzione per la gestione dinamica della memorizzazione e manipolazione di contenuti audio.

2.2. Destinatari

Tutte le aziende che intendo archiviare in formato digitale qualunque documento audio analogico/digitale.

2.3. customME Profile

CustomME è un'unione di professionisti che progetta e realizza differenti tipologie di progetti su specifiche dettate dal cliente. Il nome stesso è sinonimo della libertà di espressione del cliente e di apertura mentale di customME, realizzando qualunque tipo di progetto che si adatta completamente alle richieste del mercato e del cliente, perchè il mercato è il cliente.

CustomME impiega una protezione psicologica contro la pirateria, cioè la divulgazione e copia di materiale di qualunque tipo; questa tipologia di protezione sperimentale, ma funzionale, permette di coinvolgere a pieno il cliente. Così che il medesimo sarà proprietario della tutela dei diritti di copia e di vendita, ed in prima persona si preoccuperà di proteggere il suo investimento.

Investimento, perchè ogni progetto realizzato dalla customME è un investimento da parte del cliente; il quale, non solo può contare sulla realizzazione di un progetto che si adatta pienamente alle specifiche, ma anche di supporto continuo e globale.

CustomME si completa con l'unione di settori dedicati che collaborano insieme per la realizzazione di progetti complessi. Aree di integrazione:

- customSoft: progetti Software
- customNet: progetti Reti, Web e Sistemi Distribuiti
- customMedia: progetti Multimediali
- customTV: progetti Broadcast radio e televisioni

Ulteriori informazioni sono disponibili attraverso le documentazioni delle aree citate.

customME

email: info@elvio.org

3. Descrizione del Sistema

Il sistema è composto da differenti blocchi funzionali basati sul principio di archiviazione e gestione gerarchica (HSM: Hierarchical Storage Management). Come ogni sistema basato su questo tipo di architettura siamo in presenza di due livelli di archiviazione definiti on-line e near-on-line:

- archivio on line: è la cache disk che consente l'accesso immediato ai materiali
- archivio near on line: costituito da una libreria che consente accessi molto rapidi ma non immediati

L'archivio near on line è composto da una libreria, con impiego di supporto di materiale fisico come CD audio, DAT, music cassette, etc.

I driver di questa libreria sono lettori appositi, collegati a digital audio workstation, impiegate come grabber ed encoder dei differenti formati audio.

Del sistema fa ulteriormente parte un server per il controllo remoto della libreria, comunicazione e monitoraggio degli encoder e del grab dell'audio, gestore delle richieste di grab e codifica provenienti dal sistema di web.

Esempio di strutturazione del sistema con cd-audio

L'audio sorgente si trova codificato in CD audio che sono fisicamente posizionati all'interno della libreria.

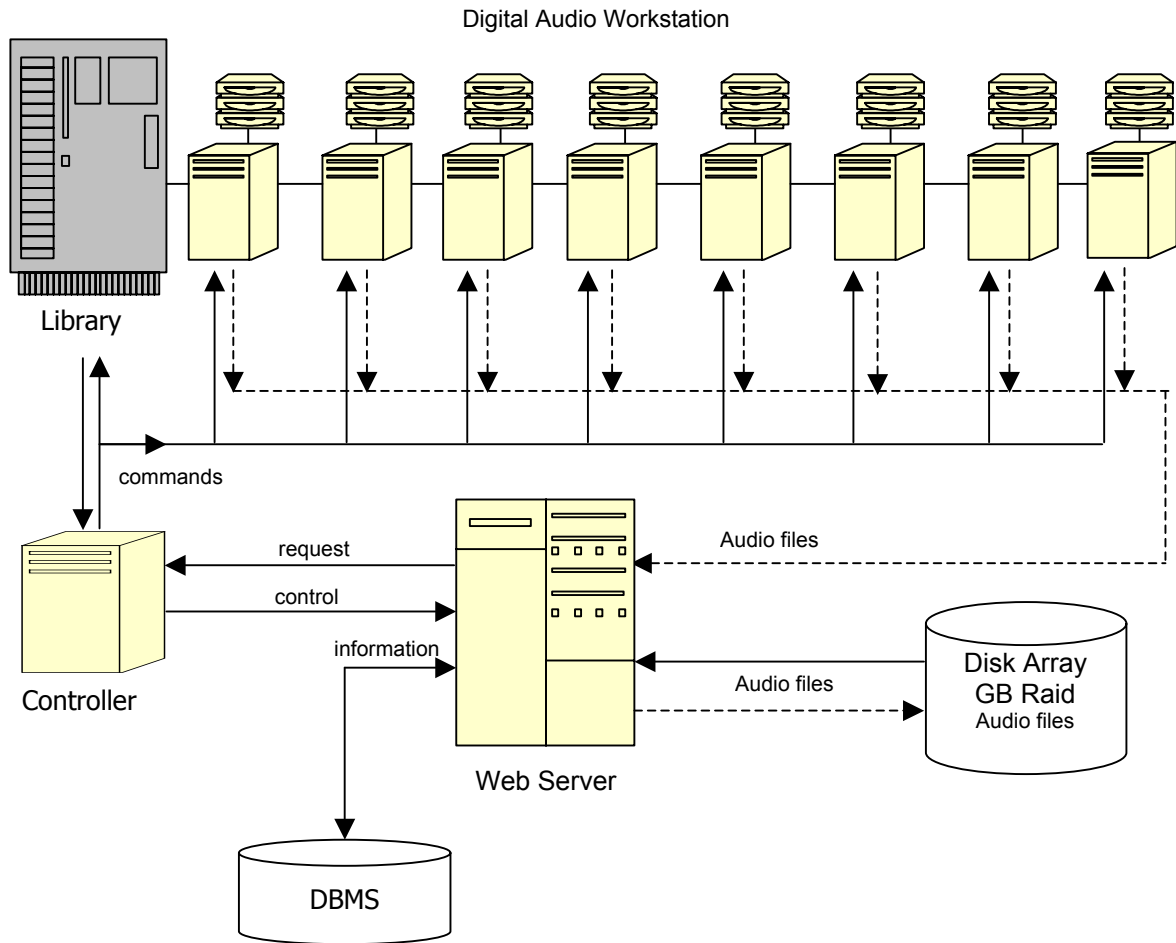
Dai CD viene estratto l'audio e salvato in formato wave su hard disk dei singoli encoder, che fungono da cache on-line insieme ad un raid collegato al web server. Quindi, si hanno una serie di file dove ad ognuno corrisponde una traccia di un determinato CD. Durante l'estrazione e la formattazione in file tipo non compresso (es. Aiff/Wave) avviene la codifica in formato compresso HQ (es. mp3, mpeg Layer III, ad un bit rate di 196kbps e 44.1kHz stereo). Esso viene fisicamente scritto all'interno della cache on-line come file e mantiene il medesimo nome usato per il file tipo non compresso, naturalmente con estensione omonima (es.mp3), in modo che siano direttamente correlati tra loro.

Una volta estratto l'audio e codificato viene convertito in un ulteriore formato compresso LQ a scelta, (es. Wma, Windows Media Audio, a 20kbps e 32Khz mono), questo per poter essere trasferito attraverso reti ad ampiezza di banda limitata; il formato compresso LQ viene anch'esso scritto direttamente all'interno della cache on-line nel raid come file e mantiene lo stesso nome del file non compresso con estensione omonima (es. Wma).

Dei formati codificati, i formati compressi HQ e LQ possono essere suonati da utente. I formati non compressi e compressi HQ possono inoltre essere scaricati dall'utente dal quale viene effettuato l'accesso a sistema.

Nel caso l'utente richieda di scaricare una traccia in formato non compresso o compresso HQ e questa non esiste più, allora il flusso dei dati viene ripercorso, ricreando così il file di tipo richiesto che manterrà il nome che aveva in precedenza.

Di seguito si può vedere uno schema rappresentativo del sistema nel suo complesso.



4. Studio d'esempio

4.1. Descrizione delle specifiche

Le presenti specifiche sviluppano uno studio tipico relativo alla possibilità di archiviare in modalità digitale documentazione della seguente tipologia: audio analogico, audio digitale.

Il materiale sorgente è supportato in tutte le sue forme analogiche e digitali: cd-audio, music cassette, dischi vinile, dat audio, e molti altri.

Il dimensionamento dell'archivio è dinamicamente implementabile con modeste ripercussioni tecniche ed economiche sul progetto.

Questo archivio è stimato, al momento in circa 100.000/120.000 file che saranno archiviati in forma digitale. Viene quindi fornito il sistemi di digitalizzazione per l'archiviazione. Si vuole supportare la digitalizzazione della documentazione in due modalità:

- "High quality Compress" formato compresso mp3 bassa risoluzione per streaming via web ed ascolto veloce
- "High Quality Not Compress" formato non compresso wave/aiff ad alta risoluzione per la parte conservativa

Entrambe le codifiche saranno memorizzate in archivio digitale per consentire:

- per il materiale HQ Compress, la possibilità di essere richiamato dall'utente per una visione rapida
- per il materiale HQ Not Compress, la possibilità di essere richiamato in modalità on-demand in donwload per un utilizzo/ascolto qualitativo del materiale

Esempio valutazione spazio necessario

Spazio necessario per singolo file audio stereo, non compresso e compresso ad alta qualità.

Resolution	Bit-rate	N° canali	N° bit	frequenza	Spazio MB	Tempo
Non Compresso	~1500Kbps	Stereo	16	44.100	~ 50	~ 5'
Compresso mp3	196Kbps	Stereo	16	44.100	~ 7	~ 5'

N° documenti stimati = 100.000

% variazione = 20 %

Totale = N° documenti stimati + % variazione = 120.000

Stima esemplificativa spazio necessario audio non compresso

Non compresso => 50 MB * N° documenti stimati = 5.000.000 MB = 4.880 GB

Non Compresso => 50 MB * Totale = 6.000.000 MB = 5.860 GB

Stima esemplificativa spazio necessario audio compresso alta qualità

Compresso => 5 MB * N° documenti stimati = 500.000 MB = 490 GB
 Compresso => 5 MB * Totale = 600.000 MB = 586 GB

Lo spazio di storage necessario, ipotizzato per la digitalizzazione di 100.000 file più un 20% di possibile range di variazione della quantità è quindi il seguente:

- HQ Compress: 586GB (mp3 stereo 196kbps)
- HQ Not compress: 5.860 GB (wave/aiff stereo 44.100Hz 16 bit)

Come si può vedere lo spazio necessario di archiviazione è notevole, per ovviare a questo inconveniente è possibile attivare un sistema di gestione temporale dei file archiviati. Al superamento di una soglia prefissata di occupazione dello storage on-line, il sistema si occuperà di cancellare quei file che hanno una data antecedente agli altri, fino alla liberazione di una quantità di storage prescelto. E' opzionale la possibilità di gestire la cancellazione di file per attributi differenti, come: i meno utilizzati, i più grandi di dimensione o altro.

La quantità di file archiviati crescerà nel tempo, si considera fino dalla prima release la possibilità di espansione del sistema stesso.

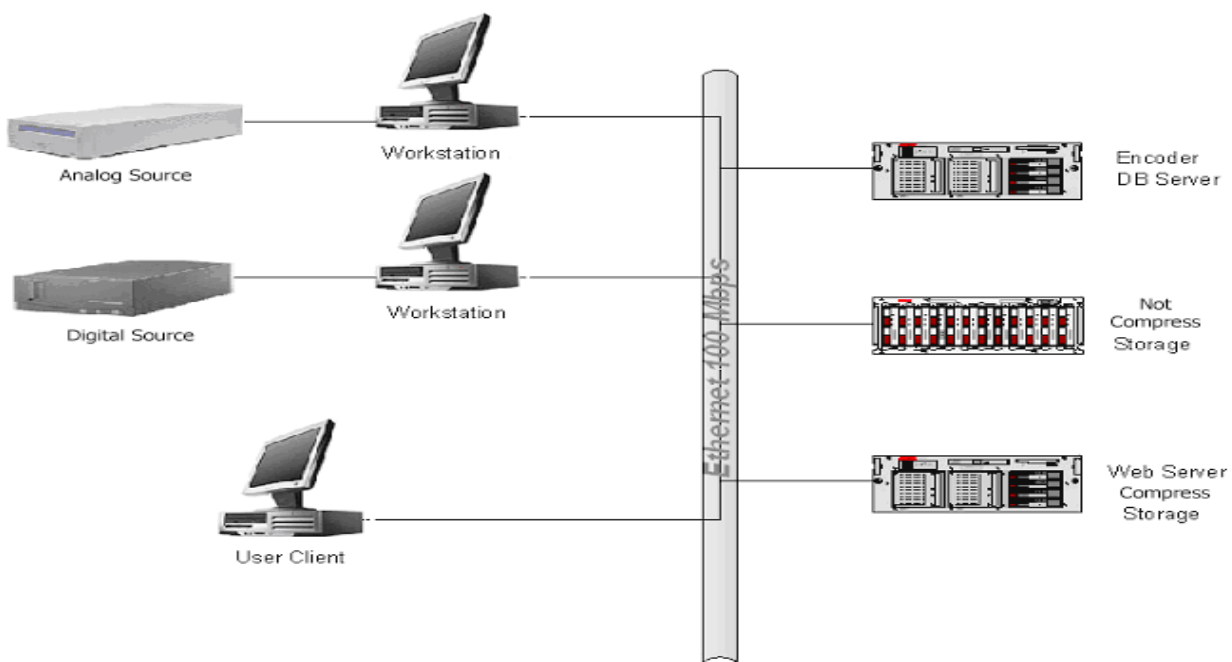
Nei sistemi di archiviazione digitale, vista la quantità di occupazione e le modalità d'utilizzo, i materiali vengono allocati su due tipi di storage:

- storage "on line", è uno storage molto veloce che consente l'accesso immediato al materiale prescelto (generalmente materiale compresso)
- storage "near on line", costituito da un secondo storage che consente accessi rapidi al materiale prescelto (generalmente materiale non compresso o materiale sorgente se già digitalizzato)

All'utente questa suddivisione su due tipologie di memorizzazione è totalmente trasparente. Saranno le applicazioni a rendere disponibile il materiale richiesto indipendentemente dalla sua allocazione.

I file saranno documentati e nominati secondo la metodologia e i tracciati dettati dal cliente.

Di seguito una possibile configurazione rappresentativa del sistema:



Gli elementi che contraddistinguono il sistema sono dunque:

1. Storage materiale non compresso
2. Storage materiale compresso
3. Web Server
4. Encoder audio non compresso e compresso
5. DB server
6. Digital audio workstation
7. Analog Source
8. Digital Source

Il sistema di digitalizzazione file sorgenti verrà proposto in due possibili modalità, come servizio e come sistema hardware.

- Servizio : i file sorgenti, in formato non compresso, verranno digitalizzati presso la nostra sede
- Sistema hardware : sistema completo di digitalizzazione file audio

4.2. Sorgenti dati del sistema

Le sorgenti di dati in input del sistema sono nei seguenti formati:

- Analogico: music cassetta, dischi vinile, etc.
- Digitale: CD audio, DAT digital audio tape, etc.

La rappresentazione dei dati in output :

- file audio formato compresso: mp3 (196kbps stereo, 44100Khz 16 bit)
- file audio formato non compresso: aiff (stereo, 44100Khz 16 bit)

4.3. Encoder audio compresso e non compresso

Verrà impiegato un encoder software per la codifica dal formato compresso mp3 e non compresso aiff. La codifica dei formati non compressi e dei formati compressi verrà effettuata in automatico durante l'acquisizione dei dati dalle sorgenti in input, in forma parallela; cioè, durante la creazione del file non compresso verrà realizzata la codifica del file compresso.

Sarà inoltre possibile effettuare la singola codifica del formato non compresso o del formato compresso su richiesta dell'utente.

4.4. Web server (Front End)

Verrà qui impiegato un server per la gestione dei servizi richieste dagli utenti, il web server sarà quindi il front end del sistema.

4.5. Storage compresso

Essendo lo spazio occupato dal formato mp3 ridotto si è pensato di utilizzare lo stesso web server per lo storage di queste immagini, le ragioni sono che si evitano trasferimenti di files da una memoria ad un'altra e di riempire la banda disponibile con inutile scambio di dati; inoltre le richieste di browsing del medesimo saranno totalmente gestite dal web server con accesso diretto ai dati.

4.6. DB Server

Le informazioni testuali sono correlati ai rispettive file audio, per tale gestione si utilizzerà un database (es. mySQL). Esso potrà risiedere direttamente sul web server o sulla workstation che ospita gli encoder. I dati testuali non occupano molto volume e verranno archiviati sulla memoria disponibile sul web server.

4.7. Storage non compresso

Lo spazio occupato dall'audio non compresso è in ogni caso significativo e può crescere proporzionalmente nei tempi futuri. Per questa ragione si è scelto uno storage dedicato facilmente espandibile per futuri dimensionamenti e integrabile nell'intero sistema. Il sistema di storage è composto da un disk array, in modalità RAID e quindi file safe, questo permette una produttività molto elevata del sistema nella necessità di leggere/scrivere i file audio contenuti.

4.8. Digitalizzazione audio

Il sistema di digitalizzazione sarà definito alla conclusione di un ulteriore ciclo di approfondimento.

Si possono prevedere due situazioni :

1. Un set hardware da scegliere ed installare ad hoc.
Composto da:
 - 1.1 digital audio workstation con input line analogici e digitali, per convertire l'audio sorgente analogico in forma digitale e per catturare direttamente l'audio in forma digitale.
 - 1.2 strumenti di lettura per cd audio, dat, dischi vinile, music cassette e altro.

Le apparecchiature saranno interfacciate alle workstation dimensionate per all'acquisizione dell'audio in formato non compresso.

2. Il servizio di acquisizione per il materiale audio non compresso sarà effettuato presso la sede di customME. Il prodotto di questo processo sarà costituito da files di dati audio e relativi "meta-dati" per l'opportuna archiviazione.

5. Architettura del Sistema

5.1. Hardware

La configurazione Hardware del sistema prevede diversi blocchi funzionali:

1. Web server
2. Database server
3. Encoder formato compresso
4. Digitalizzazione audio in formato non compresso
5. Storage per materiale compresso
6. Storage per materiale non compresso
7. Gli apparati di rete
8. Workstation dedicate all'acquisizione
9. Dispositivi di lettura

Verranno presentate diverse soluzioni a partire dal semplice storage dell'audio fino ad arrivare ad un sistema completo in versione cluster e non, inoltre sono stati aggiunti anche gli strumenti opzionali per l'acquisizione dei sorgenti.

5.2. Software

Viene delineata la realizzazione di tale progetto attraverso portale di tipo web (es. PHPNuke con web server Apache, database di tipo MySQL e/o superiori, sistema OsX server/Unix/Linux o altri). Il portale sebbene riunito in un sito unico esporrà diverse funzionalità principali:

- 1) immissione dati: audio e relativi metadati secondo maschera di inserimento;
- 2) visualizzazione archivio: con form di ricerca e consultazione;
- 3) ascolto on line dell'audio in formato compresso;
- 4) download dell'audio in formato compresso e non compresso;
- 5) amministrazione e manutenzione;

Il workflow tipo dell'operatore sarà:

1. l'operatore si logga sul portale
2. l'operatore acquisisce l'audio;
3. l'operatore apre il form di inserimento dati, individua la traccia audio precedentemente acquisita sfogliando il file system (la directory dove vengono posizionati);
4. inserisce i metadati e pubblica tutte le informazioni attraverso validazione; la pubblicazione rende disponibile a tutti gli utenti l'audio digitalizzato.
5. il nostro sistema per ogni file audio non compresso, produce un file audio in formato compresso che sarà pubblicato sul portale; inoltre l'utente potrà richiedere la codifica di qualsiasi file non compresso nel formato compresso, nel caso di corruzione di file compressi precedentemente archiviati.
6. l'audio sarà collegabile sul portale attraverso thumbnails (icone) che ne permettono l'ascolto immediato. Su richiesta dell'utente sarà possibile il download del formato non compresso.

6. Quotazione

La quotazione viene realizzata sulla base di specifiche cliente, essa include le seguenti risorse:

- Risorse fisiche: hardware necessario all'implementazione del progetto (non necessariamente fornito da customME, ma acquistabile direttamente dall'utente, sotto precise indicazioni di customME)
- Risorse umane: personale qualificato per l'installazione e manutenzione delle risorse fisiche, personale qualificato per lo sviluppo ed integrazione dei software per l'amministrazione del sistema.
- Servizi: servizio di acquisizione e digitalizzazione dell'audio da supporto, con rilascio del materiale digitalizzato e corrispettivi meta-data per l'opportuna archiviazione.

7. Servizi professionali customME

System Integration: Installazione, collaudo, manualistica e addestramento sono inclusi i seguenti servizi:

- L'installazione software e hardware dei server
- Installazione software e hardware degli encoder
- Integrazione dei sistemi già installati
- Tuning di tutti i sistemi
- Produzione di adeguata manualistica e documentazione dei flussi di lavoro
- Addestramento del personale del cliente
- Baby sitting allo start-up dei sistemi

Le fasi di lavoro sono così riassunte:

- Sopralluogo del sito per identificare con il cliente le attività di pre-installazione
- Determinare la schedulazione delle attività ed i periodi di esecuzione
- Preparazione del sito per l'installazione, identificazione della localizzazione dell'impianto, alimentazione, cavettistica, apparati di rete e relative canalizzazioni
- Pre-installazione in-house dei nuovi sistemi
- Installazione dei sistemi pre-installati (e relativo rack standard nel caso degli encoder e server)
- Addestramento del personale del cliente all'uso dei sistemi e relativo baby sitting
- Collaudo e chiusura progetto

La partecipazione e la responsabilità del Cliente riguardano le seguenti aree:

- Predisposizione dell'ambiente per l'installazione, configurazione e testing concordando con customME le fasi delle attività.
- Messa a disposizione di almeno un contatto tecnico, con responsabilità di system administration ed appropriati livelli di accesso, ed informazioni necessarie per l'installazione e la configurazione dei sistemi.

7.1. Assistenza tecnica customME

CustomME prevede diversi livelli di assistenza tecnica sugli apparati proposti, che sono coperti da garanzia. Attualmente customME eroga servizi che vanno dalla riparazione in-house dei sistemi all'intervento on-site. I servizi vengono concordati con il cliente in funzione delle sue esigenze specifiche di ripristino. Alcune attività, quali gli aggiornamenti software e il monitoring dei sistemi, possono essere effettuate tramite accesso remoto. In questo caso è richiesta una linea dedicata. Nel caso di intervento on-site, le richieste vengono inoltrate tramite email su apposita casella postale presidiata, o tramite numeri telefonici dedicati per l'assistenza. Nell'ambito del servizio on-site sono previsti anche interventi di manutenzione preventiva a cadenza temporale da definire con il cliente.