

# 방송 아카이브 기술 개요

□ 김현규, 김정만 / 한국편마이크로시스템즈, \*\*페타데이터

## 1. 서론

최근 방송 분야에서의 화두는 단연 디지털이다. 2012년 디지털 방송 전환을 앞두고 벌써부터 많은 준비를 해왔고, 앞으로도 많은 관련 업체들이 이 분야에 더 많은 관심을 기울일 것은 분명한 사실이다. “디지털과 아날로그” 어느 것이 사람이 살아 가는데 행복을 더 가져다 주는 지에 대한 해답은 내릴 순 없지만, 필요에 의해서 또한 편리함을 위해서 많은 정보들이 디지털 도구를 통해 저장되고 있다. 방송 분야에서도 제작에서 송출에 이르기 까지 빠른 속도로 디지털화하고 있는데, 이는 예전에 비해 갈수록 방대해지는 방송 영상 자료의 양적, 질적 고도화에 이바지할 것으로 믿는다.

디지털 시대에 정보가 대량으로 생산되면서, 정보의 보존 및 재활용을 경제적으로 하기 위한 아카이브 기술도 덩달아 많은 발전을 이루고 있다. “백

업”에 비해 “아카이브”가 저장 정보의 손상 복구보다는 정보 재활용 측면에서 강조된 기술이라면, “아날로그 아카이브”에 비해 “디지털 아카이브”는 정보를 분류해서 저장하고 다시 활용하는 기술이, 저장 대상의 조작 편의성으로 인해 혁신적으로 향상되었다고 말할 수 있다. 그리고, 단순 저장 보관보다는 활용 측면이 더 강조되면서, 다양한 정보의 유형 및 특성을 고려할 수 많은 아카이브 솔루션이 존재하고 있다. 예를 들면, 파일 아카이브, 이메일 아카이브, 데이터베이스 아카이브 솔루션 등이 있고, 파일 아카이브 솔루션 중에서도 단순 파일 특성이 아닌 이미지, 동영상 파일 중심의 아카이브 솔루션 등으로 전문화 또는 세분화 되어 있다. 하나의 솔루션으로 모든 정보 유형에 대한 아카이브 적용을 바라는 사용자의 요구도 있지만, 개별 정보에 대한 특성, 처리 및 활용 방안이 다르기 때문에, 특정 정보 유형에 맞는 전문 아카이브 솔루션들이 존재

한다.

참고적으로, 방송 아카이브 솔루션의 경우에는 다른 정보에 비해 다음과 같은 데이터 특성을 잘 다룰 수 있는 필수 요소 기술을 필요로 한다. 첫째, 물리적으로 다른 데이터에 비해 거대한 저장 용량이 필요하다. 이로 인해, 데이터 전송 및 처리 시간 단축 등을 위해 데이터 압축 및 부호화 기술을 기본적으로 사용한다. 둘째, 문자 데이터는 정적인 반면에 방송 영상 데이터는 시간의 흐름에 따라 사용자에게 자연스럽게 보여야 한다. 즉, 실시간 처리를 위한 시간적 특성이 존재한다. 그러므로, 시간적 제약에 따른 방송 영상 데이터를 전달하는 통신 채널당 상당한 대역폭을 필요로 하게 된다. 셋째, 방송 영상 데이터는 시청각적 특성을 가지고 있기 때문에, 단순 검색어로 정확하게 표현할 수 없는 비정형 데이터이다. 따라서, 내용 기반 검색을 위해 완전 일치가 아닌 오차

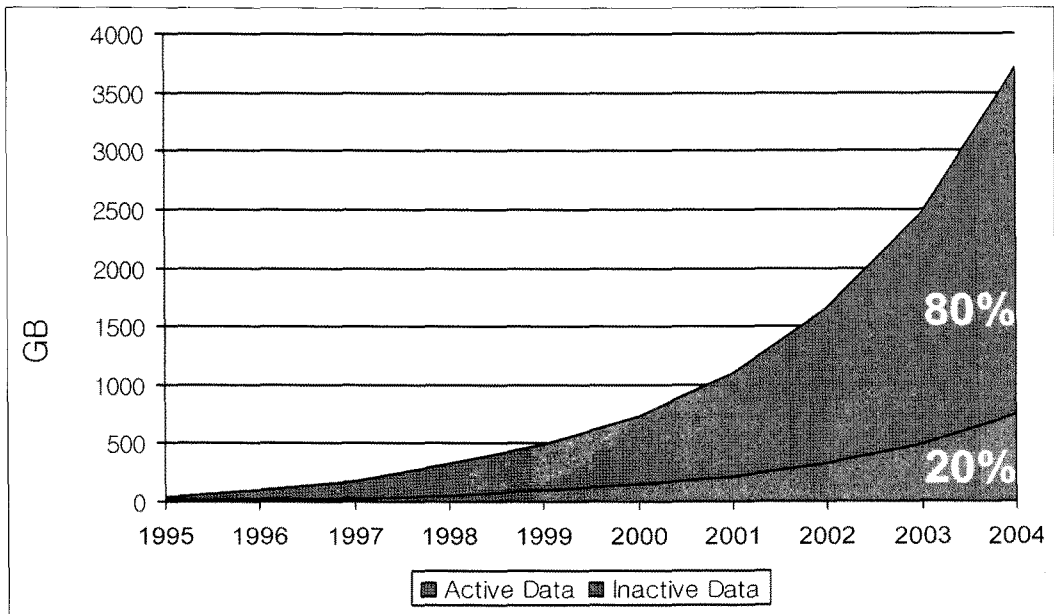
를 가지는 유사 일치 검색이 사용되기 때문에, 영상 검색 정확도를 높이는 기술이 요구된다.

본 원고에서는 아카이브 기술에 대한 일반적 개요, 일반 파일과 다른 데이터 특성을 가지는 방송 아카이브 기술에 대한 구성 요소, 적용 사례, 기술 동향 및 향후 전망에 대하여 살펴보기로 한다.

## II. 아카이브 기술 개요

### 1. 아카이브 솔루션 필요성

오늘날 기업에서 수집하고 사용하는 정보가 한정된 예산과는 달리 기하급수적으로 증가하고 규제 요구 사항도 꾸준히 늘어나기 때문에 정보 유지 비용도 자연스럽게 커질 수밖에 없다. 또한, 저장된 정보



<그림 1> 일반적 기업의 데이터 구성비

<표 1> 정보 생명 주기 관리 (ILM) 적용 단계

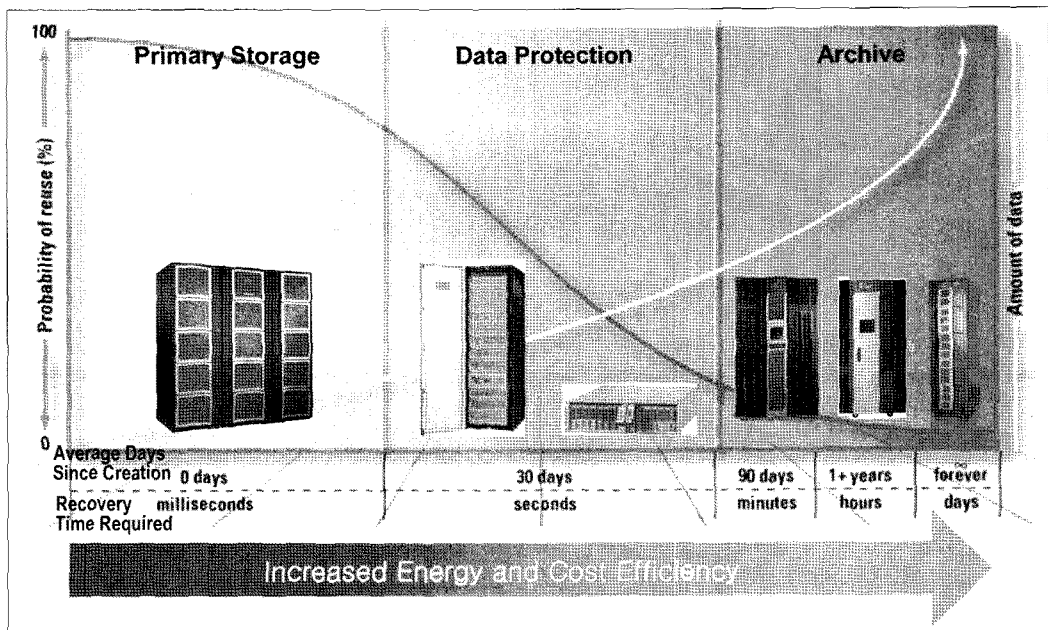
데이터 분석			데이터 분류	스토리지 저장	데이터 이동 및 관리
담당자 면담	데이터 수집	분석 평가	데이터 분류 모델	데이터 분류 결과에 의한 스토리지 재배치	정보 생명 주기 적용 및 정책에 의한 관리

역시 <그림 1> 과 같이 높은 가치를 가지고 계속해서 사용되는 활성 데이터 보다는 시간이 흐를수록 단순히 보존되는 비활성 데이터의 비율이 엄청나게 높다는 것을 알 수 있다.

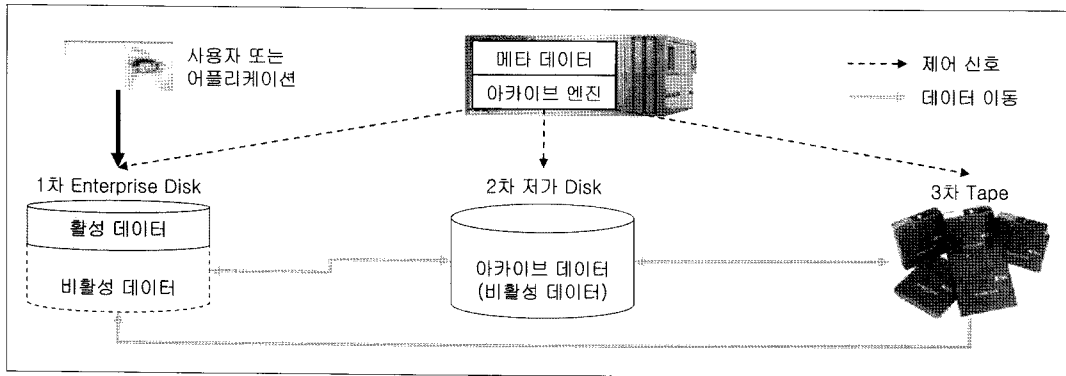
이런 경향은 <표 1> 과 같이 모든 정보를 조직에서 가치와 역할에 따라 분석 및 분류하고, 가치와 역할에 맞는 적절한 계층적 스토리지에 저장하고, 유지 및 관리하는 정보 생명 주기 관리 (Information Lifecycle Management, 이하 ILM) 를 필요로 하게 된다. 즉, 정보 생명 주기 관리 (ILM)는 모든 정보가

동일한 가치를 가지지 않기 때문에, <그림 2> 에서와 같이 “Right Data, Right Device, Right Price” 를 목표로 한다.

이러한 정보 생명 주기 관리 (ILM) 를 현실에서 가장 가깝게 구현한 솔루션이 아카이브 솔루션이다. 아카이브 솔루션은 <그림 2> 에서와 같이 정보의 가치 및 역할에 맞게 1차 저장된 후, 시간이 지날수록 가치가 변함에 따라, 상위 계층의 스토리지로부터 하위 계층의 스토리지로 사전에 정의된 정책에 의해 자동적으로 이동되고 혹, 하위 계층에 존재하는 정



<그림 2> 정보 생명 주기 관리 (ILM) 모델



<그림 3> 아카이브 솔루션 기본 구성도

보 중에서 필요한 것을 다시 상위 계층의 스토리지로 옮기는 역할을 한다.

## 2. 아카이브 솔루션 기본 개념

일반적으로 아카이브 솔루션은 아카이브를 구현하는 엔진 소프트웨어와 엔진이 탑재되는 서버, 데이터를 저장하는 스토리지(디스크 어레이, 테이프 라이브러리 등) 하드웨어로 구성된다. <그림 3>에서와 같이 아카이브 엔진이 저장된 모든 데이터에 대해 저장 위치, 액세스 시간, 보존 기간 등 관리에 필요한 모든 정보를 메타 데이터로 관리하고 있어, 사용자 입장에서는 필요한 정보에 대해 저장된 장소에 구애받지 않고 평소 사용하던 방식대로 활용할 수 있다. 다만, 하위 스토리지에 아카이브된 데이터는 저장 장치의 액세스 시간에 따라 시간 지연이 발생된다.

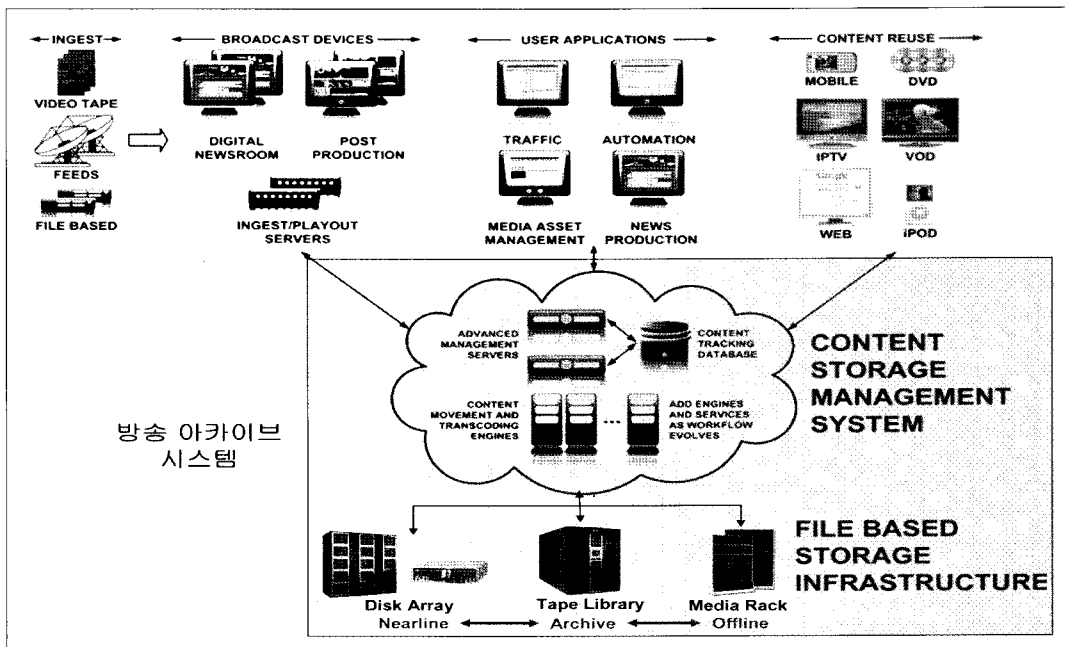
하는 경우가 많으나, 두 솔루션은 원천적으로 구현 목적이 서로 다르고 또한, 두 솔루션이 하나의 솔루션을 완벽하게 대체하지 못한다. 따라서, 필요사항을 면밀하게 분석하여 더 적합한 솔루션을 사용하여야 한다. 완벽한 데이터 보호와 효율적인 정보 생명주기 관리를 위해선 두 솔루션을 함께 사용할 수도 있다.

<표 2> 아카이브 솔루션과 백업 솔루션의 차이점

기능 및 특징	아카이브 솔루션	백업 솔루션
구축 목적	저장 장치의 용량 폭주, 성능 및 경제상의 문제로 원본 또는 복사본을 계층형 스토리지로 이동 보관하는 것	원본 데이터 손상 또는 과거 데이터 필요 시, 복구를 전제로 복사본을 보관하는 것
데이터 사용 방법	일반적으로, 기존 원본 데이터 사용 방법과 동일	백업 어플리케이션을 통한 접근
데이터 검색 시간	빠름	느림

## 3. 아카이브 솔루션과 백업 솔루션 차이점

아카이브 솔루션과 백업 솔루션을 혼동하여 사용



<그림 4> 방송 아카이브 시스템 기본 구성도

### III. 방송 아카이브 기술 구성 요소

#### 1. 방송 아카이브 시스템 구성 요소

방송 아카이브 시스템은 일반적인 아카이브 솔루션과 마찬가지로 <그림 4> 에서와 같이, 아카이브 관리 및 제어하는 Content Storage Management (이하 CSM) 시스템과 영상 자료를 저장하는 하드웨어인 디스크 어레이 및 테이프 라이브러리 등으로 이루어진다.

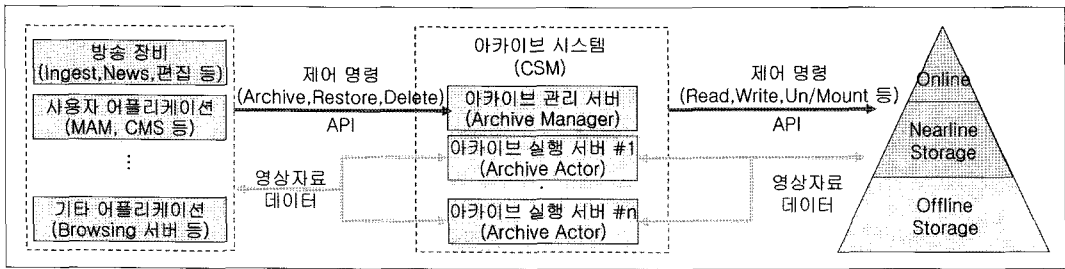
#### 2. 방송 아카이브 소프트웨어- Content Storage Management (CSM) 시스템

CSM 시스템은 계층적으로 사용자 어플리케이션 Media Asset Management (이하 MAM) 및 Content

Management System (이하 CMS)와 아카이브 스토리지 사이에 존재한다. CSM 시스템은 사용자 어플리케이션인 MAM 이나 CMS로부터 요청을 받아, 아카이브 스토리지에 영상 자료를 저장하거나 반대로 저장된 영상 자료를 회수한다.

CSM 시스템과 방송 장비 (비디오 서버 등) 및 사용자 어플리케이션 (MAM 및 CMS) 간은 CSM 소프트웨어가 제공하는 Application Programming Interface (이하 API) 에 의해 연동된다. 또한, CSM 시스템과 하드웨어 저장 장치 간의 연동은, 디스크 어레이의 경우 표준 명령어를 사용하고, 테이프 라이브러리의 경우 테이프 라이브러리 제조사에서 제공하는 API를 사용한다.

또한, CSM 시스템은 기본적으로 <그림 5> 와 같이 분산 병렬 처리가 가능한 아키텍처를 채택하여야 한다. 왜냐하면, 도입 후 지속적으로 사용량이 늘어



<그림 5> 방송 아카이브 (CSM) 시스템 분산 병렬 아키텍처

<표 3> 방송 아카이브 (CSM) 시스템 주요 구성 요소

구성 요소	역 할	주요 기능
아카이브 관리 서버 (Archive Manager)	<ul style="list-style-type: none"> <li>CSM 시스템 전체 관리 및 제어 모듈</li> <li>안정성 위해 Active/Standby 구성</li> </ul>	<ul style="list-style-type: none"> <li>사용자 요구 관리</li> <li>아카이브 영상 관리</li> <li>주요 자원 할당</li> <li>자동/수동 Repack (De-fragmentation) 제어</li> <li>테이프 라이브러리 관리</li> </ul>
데이터베이스	<ul style="list-style-type: none"> <li>아카이브 관리 서버가 전체 시스템을 관리하기 위한 각종 데이터 보관</li> <li>성능 및 안정성 검증된 RDBMS 사용</li> </ul>	<ul style="list-style-type: none"> <li>시스템 구성 데이터 저장</li> <li>오브젝트 카탈로그 저장</li> <li>처리 및 이벤트 기록 저장</li> </ul>
아카이브 실행 서버 (Archive Actor)	<ul style="list-style-type: none"> <li>아카이브 관리 서버의 지시를 받아, 실 데이터 I/O 작업을 수행</li> </ul>	<ul style="list-style-type: none"> <li>Archive, Restore, Partial Restore, Delete, Repack 등 작업 수행</li> </ul>
관리/제어 GUI Tool	<ul style="list-style-type: none"> <li>아카이브 관리 서버 기능 원격 제어 및 모니터링</li> <li>아카이브 시스템 구성 관리</li> <li>장애 발생시 긴급 사용 유틸리티 제공</li> </ul>	<ul style="list-style-type: none"> <li>아카이브 모니터링</li> <li>예러 및 이벤트 리포팅</li> <li>Online/Offline에서 구성 설정, 변경, 삭제</li> </ul>
API	<ul style="list-style-type: none"> <li>CMS, MAM 등과 연동 위한 인터페이스</li> </ul>	

날 것에 대비해 아카이브 용량뿐만 아니라 처리 성능도 자연스럽게 늘릴 수 있어야 하고, 기존 투자를 보호할 수 있는 분산 병렬 아키텍처가 최적이기 때문이다. 분산 병렬 아키텍처 방식의 CSM 시스템의 주요 구성 요소는 <표 3>과 같다.

### 3. 방송 아카이브 하드웨어 - 디스크 어레이, 테이프 라이브러리

아카이브 시스템에 사용되는 하드웨어는 데이터

를 안정적으로 저장 및 회수하기 위한 요구 성능 및 저장 용량으로 모델 및 사양을 결정하게 된다. <표 4>는 일반적으로 아카이브 시스템에서 사용되는 저장 장치 사양을 표시한 것이다.

일반적으로 <표 4> 와 같이 저장 매체 종류별로 권장 사용 기한과 교체 주기가 존재하기 때문에, 데이터의 보존 기간에 알맞은 매체를 선택하여야 한다. 만약, 저장 매체의 권장 사용 기간 및 교체 주기가 데이터 보존 기간보다 짧거나 보존 기간이 영구적일 경우, 가급적 기존 매체의 데이터를 신규 매체

<표 4> 아카이브 저장 장치 일반 사양

스토리지 구분	스토리지 종류	요구 성능(Access Time)	저장 용량	데이터 보존 평균 기간	매체 교체 평균 주기
Online	Disk Array (FC, SATA HDD)	milliseconds ~ seconds	~ Tera Bytes	~ 1 month	~ 5 years
Nearline	Tape Library (Tape Media)	minutes	Peta Bytes ~	months ~ years	15 years ~ 30 years
Offline	Tape Media	hours	∞	∞	15 years ~ 30 years

로 옮기기 쉬운 매체 종류를 선택하는 것이 바람직하다. 또한, CSM 시스템과 마찬가지로 데이터 용량이 급속하게 증가할 것을 대비하여, 처리 성능 및 저장 용량에 대한 확장성이 최초 도입 모델에서 충분히 제공될 수 있는 하드웨어 선택이 필요하다.

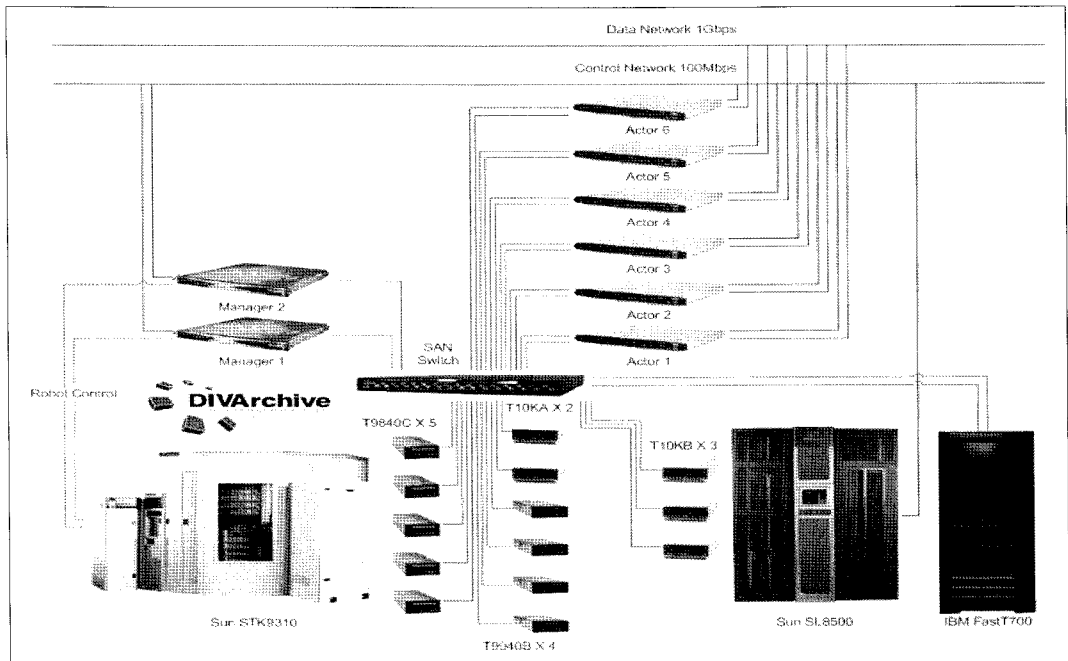
방송 영상 자료의 경우 보존 기간이 일반 데이터에 비해서는 매우 길고, 데이터를 손실할 경우 재생산이 불가능하므로 안정성을 확보할 수 있게 여러벌의 복제본 생성, 분산 보관, 저장 시스템 이중화,

재해 복구 대책이 필수적으로 마련되어야 한다.

## IV. 방송 아카이브 적용 사례

### 1. 국내 뉴스룸 아카이브 사례 - 공중파 방송사

뉴스 아카이브는 평소에는 아카이브 작업이 주어



<그림 6> 국내 뉴스룸 아카이브 구성 “예”

지만, 뉴스 프로그램을 앞두고 리스토어 및 파셜 리스토어(Partial Restore)가 급증하는 특성이 있다. CSM 소프트웨어와 아카이브 스토리지가 이에 대응할 수 있는 기능을 갖춰야 한다.

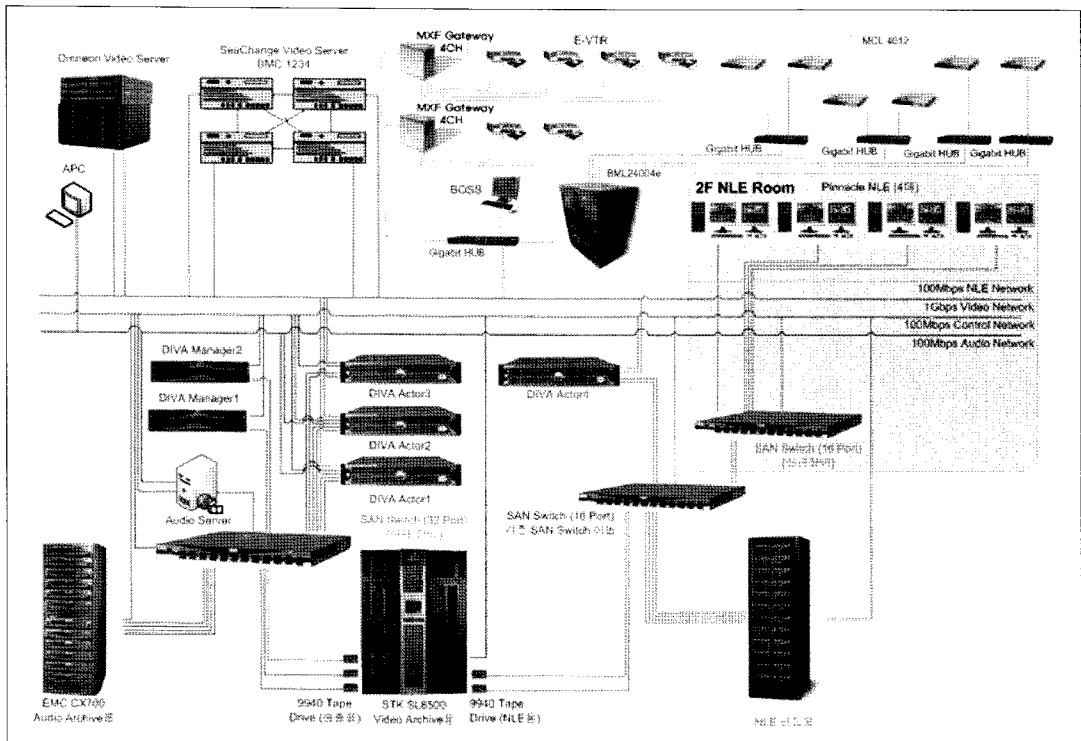
## 2. 국내 송출용 아카이브 구성 사례 - 전문 송출사

송출 아카이브는 송출 채널 수가 많아 평소 지속적으로 아카이브/리스토어 작업이 일어난다. 물론 어느 시스템이나 안정성이 중요하지만, 송출 아카이브 시스템의 경우 24x365 가용성이 중요하므로 주요 모듈 및 하드웨어의 이중화 및 가용성 보장이 때

우 중요하다.

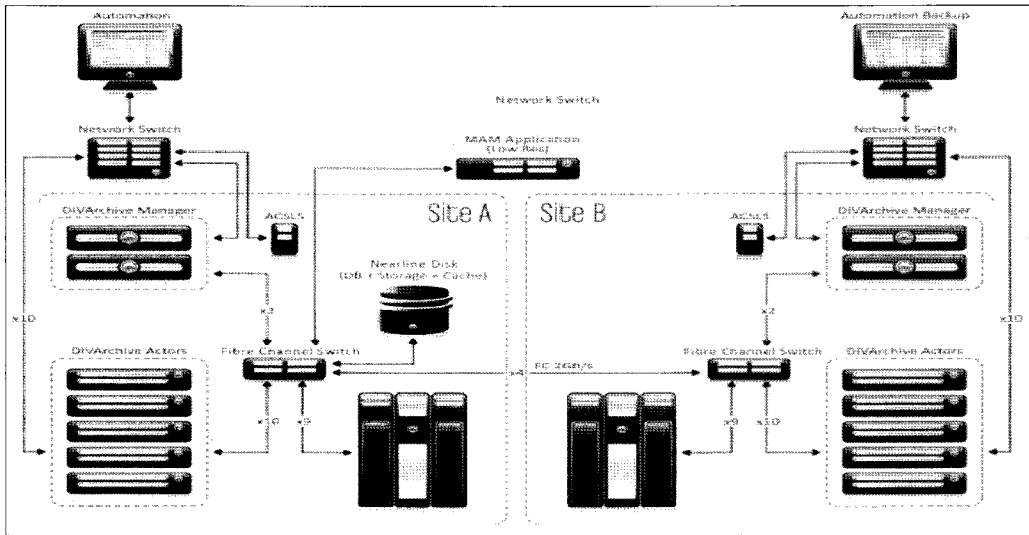
## 3. 해외 종합 방송 아카이브 구성 사례- Bayerischer Rundfunk

Bayerischer Rundfunk(BR)은 독일의 공영 방송으로 5개의 라디오 채널, 하나의 디지털 오디오 방송 채널, 24시간 방송하는 2개의 채널을 보유하고 있다. 2002년에 독일에서 처음으로 아카이브 시스템을 구축하였다. 현재는 하루에 15시간 분량의 미디어를 아카이브 하고 있으며 금년 말까지 하루 40시간 이상의 아카이브를 예상하고 있다. 또한, 이 방송국은 아카이브 재난 복구(Disaster Recovery) 시

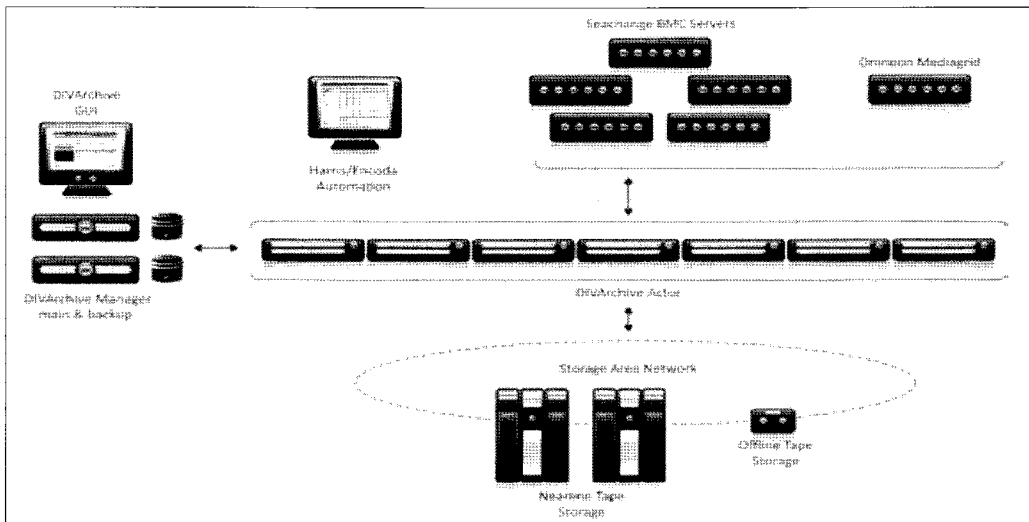


<그림 7> 국내 송출 아카이브 구성 “예”





<그림 8> 해외 방송 아카이브 구성 “예”



<그림 9> 해외 송출용 아카이브 구성 “예”

시스템을 구축하여 가용성을 극대화하였다.

#### 4. 해외 송출용 아카이브 구성 사례 - ESPN Star Sports

ESPN Star Sports는 아시아에서 24개국에 17개

의 채널을 송출하며, 3억 명 이상의 시청자를 보유한 24시간 스포츠 중계 채널이다.

#### V. 방송 아카이브 기술 동향 및 향후 전망

이미 수 년 전부터 해외는 물론 국내 방송 업계에

서도 아카이브 시스템 도입이 늘고 있다. KBS 는 일부 업무이긴 하지만 국제 뉴스 아카이브 시스템을 운영 중이고, 전사 아카이브를 위한 프로젝트 팀이 가동 중에 있다. MBC 는 자료 아카이브 시스템을 여 의도 본사와 일산 드림 센터에 이중화 구성을 하여 운영 중이고, 보도 아카이브 시스템을 준비 중이다. SBS 의 경우 뉴스 아카이브 시스템 및 자료 아카이브 시스템 도입이 완료되어 운영 중이다. 송출 대행사인 디지털 온미디어와 CJ파워퀘스트는 역시 아카이브를 통해 수십 개 채널의 송출 대행 업무를 운영 중에 있다. 뒤를 이어 방송통신대학TV, YTN 뉴스 아카이브 및 EBS 아카이브 프로젝트가 진행 완료되었거나 진행 중에 있다. NLE 와 온라인 스토리지를 이용한 네트워크 편집이 보편화되고 있어 아카이브 시스템 도입은 더 이상 미룰 수 없는 일이 되고 있다. 빠르고 용량이 큰 테이프 장비 출시 및 관련 솔루션들의 가격이 수년 전과 달리 많이 하락하여, 방송 업체들의 아카이브 도입을 부채질하고 있다. CSM 소프트웨어는 단순한 아카이브/리스토어 기능을 넘어 장에 예측을 위한 기능, 다양한 기기종 테이프 스토

리지 동시 지원 및 재난 복구 솔루션 등 다양한 기능을 내장하기 시작했고 앞으로도 계속적으로 발전해 나갈 것이다.

또한, 방송 아카이브 하드웨어 중에서 디스크 스토리지는 여러 가지 새로운 변화 요인으로 인하여 중대한 기로에서 있는 것이 현실이다. 첫째, 기존에 비용 문제로 사용이 어려웠던 Solid State Disk (이하 SSD) 가 최근 성능 및 비용 효율적인 측면에서 다시 고려 대상이 되고 있다. 둘째, SSD 와 Hard Disk Drive (이하 HDD) 를 함께 사용하는 하이브리드 스토리지가 나타났다. 셋째, 서버의 CPU 성능이 순수 계산 기능에도 못 미쳐 I/O 에 대한 부담을 덜고자 나날이 고비용 구조로 강화되던 디스크 어레이가, 최근 서버 CPU 성능의 여유가 생기면서 디스크 컨트롤러의 역할을 서버로 옮긴 오픈 스토리지가 출현함으로써, 비용 효율적인 측면에서 대용량 저장 장치 분야의 새로운 패러다임이 되고 있다.

테이프 라이브러리도 디스크 기반 스토리지의 가격이 하락하면서 대용량 저장 장치로서의 자리를 위협받았지만, 급속히 증가하는 대용량의 데이터를 저

<표 5> 대용량 스토리지 로드맵

구분		단위	2007	2009	2011	2017	2019
HDD	Form Factor	inch	3.5, 2.5, 1.8, 1.3, 1.0	3.5, 2.5, 1.8, 1.3	3.5, 2.5, 1.8, 1.3	2.5, 1.8	2.5, 1.8
	Capacity	GB, TB	8GB ~ 1TB	60GB ~ 2TB	120GB ~ 4TB	700GB ~ 30TB	1TB ~ 50TB
테이프	Form Factor	inch	5.25FH, 5.25HH, 3.5	5.25FH, 5.25HH, 3.5	5.25FH, 5.25HH, 3.5	5.25HH, 3.5	5.25HH, 3.5
	Capacity	GB, TB/media (2:1 기본압축)	1.6TB	2.4TB ~ 3.2TB	6TB ~ 8TB	24TB ~ 32TB	96TB ~ 128TB
	Data Rate	MB/sec/driver (2:1 기본압축)	240	320~360	400~560	800~1600	1600~3200

장할 경우 디스크 대비 여전히 비용 효율적이고, 방송 영상 자료와 같이 영구 보관을 요구하는 데이터의 장기 보관이 용이하고, 저장 매체 교체 주기가 길고, 운반이 편리하며, 위험한 환경에서 격리되어 손쉽게 보호될 수 있기 때문에, 데이터 보호의 최후 보루로서 많은 데이터가 테이프 스토리지에 저장되고 있다.

사실 모든 스토리지 매체가 고유의 특성을 가지고 있기 때문에, 성능 및 저장 용량, 비용, 안정성, 이식성, 매체 수명 등을 종합적으로 고려하여 필요에 맞게 선택하여야 한다.

## VI. 결론

기존 아날로그 방송 영상 자료의 경우 매체가 가지고 있는 물리적 특성상 열화 및 손/망실의 우려와 접근적 한계성 때문에 디지털화를 통한 새로운 보존 및 관리 방식이 필요하다. 또한, 앞으로는 제작에서

송출 단계에 이르기까지 모든 영상 자료가 기본적으로 디지털화 된다. 그러므로, 새로운 디지털 방송 시대의 패러다임에 맞는 디지털 방송 아카이브 시스템 구축을 통해 장기 보존, 복원, 접근의 용이성을 제공해야 한다는 것은 분명한 사실이다.

국내의 디지털 방송 아카이브 구축 현황은 공중파 방송사와 주요 송출사를 위주로 구축되었거나 구축되고 있는 실정이다. 이마저도 자체 제작된 모든 방송 영상 자료를 통합 관리하는 기구 또는 체제를 가지고 있는 곳은 거의 드물다. 하지만 궁극적으로는 국가 차원에서 디지털 아카이브 시스템을 구축하여, 공익을 위해 방송 영상 자료를 비롯한 모든 정보를 쉽게 공유 및 활용할 수 있어야 한다고 본다.

본 원고에서는 디지털 방송 아카이브 구축을 위한 아카이브 기술에 대한 기본 사항과 방송 영상 자료의 특성에 맞는 아카이브 시스템이 가져야 할 구성 요소, 고려 사항, 구축 사례에 대하여 설명함으로써, 디지털 방송 아카이브 시스템의 개념적 모형을 제공하고자 한다.

### 참고 문헌

- [1] 방송 영상 산업 진흥 5개년 계획 (2008 ~ 2012), 한국방송영상산업진흥원, 2006. 12. 26
- [2] 방송 영상 프로그램 제작 기반 조성을 위한 디지털 아카이브 구축방안 연구, 한국무선국관리사업단, 2000. 8. 14
- [3] 디지털 방송 영상 아카이브 구축 방안에 관한 연구, 유은혜, 2006
- [4] Mass Data Storage Trends: HDD Technology, Roger F. Hoyt and Tom Coughlin, Coughlin and Associates, 2008. 11. 18
- [5] iNEMI 2009 Roadmap, www.inemi.org, 2009. 4. 1
- [6] Bi-annual iNEMI Mass Storage Report for 2008, www.inemi.org, 2008
- [7] Disk and Tape Square Off Again - Tape Remains King of the Hill with LTO-4, David Reine and Mike Kahn, 2008. 2. 13
- [8] Tape The Digital Curator of the Information Age, Fred Moore - President Horison, Inc., 2008. 7. 17

## 필자소개



김현규

- 1993년 : 부산대학교 컴퓨터공학과, 학사
- 1993년 1월 ~ 2001년 7월 : 포스데이터, 시스템프로그래머/컨설턴트
- 2001년 8월 ~ 2002년 1월 : 삼일 새미스, ERP 컨설턴트
- 2002년 1월 ~ 2003년 6월 : 원솔, 기술지원팀장
- 2003년 7월 ~ 현재 : 한국썬마이크로시스템즈, Data Management Ambassador
- 주관심 분야 : 데이터 보호, 아카이브, 시스템 아키텍처



김정만

- 1991년 : 성균관대학교 정보공학과, 학사
- 1991년 ~ 1994년 : LG소프트웨어, UNIX팀
- 1994년 ~ 2000년 : 동양시스템즈, 기술연구소
- 2000년 ~ 2001년 : 아이큐브, 기술연구소
- 2001년 ~ 2009년 : 아이앤에스 시스템즈, 아카이브 솔루션팀
- 2009년 ~ 현재 : ㈜페타데이터, 디지털 아카이브 솔루션팀
- 주관심 분야 : 디지털 방송 아카이브 기술, 네트워크 편집 스토리지 기술