Tab Butler
MLB Network
Director Media Management & Post Production

From DIAMOND to Archive:

File Based Workflows… Keeping Up With the Technologies

“If you build it, they will come…”

MLB Network’s Studio 3 & Studio 42

Wildcard Wednesday

September 28th 2011

MLB Network Fun Facts

- Largest debut in cable TV history

- 50 Million Households on Launch – January 2009
- 72 Million Households – September 2012
- 2,000 hours per week of HD content recorded in 2010 season
- 2,500 hours per week of HD content recorded in 2011 season
- 3,100 hours per week of HD content recorded in 2012 season
- 3,700 LTO-4 Tapes consumed for Archive in 2010 season
- 4,600 LTO-4 Tapes consumed for Archive in 2011 season
- 6,100 LTO-4 Tapes consumed for Archive in 2012 season

MEDIA MANAGEMENT – THE WORKFLOW

The Architecture

Application Level

DIAMOND
Asset Management System

Aurora & Final Cut Pro
Edit Systems

Digitized Content
DIAMOND

Infrastructure Level

Live Feeds (Game Broadcasts)

Grass Valley
Digital Production System

HD Video Storage

Browser Quality Video Storage

HD Play out to Control Rooms, other Facility areas...

Finished Programs for MLB Network

Omneon
Play-to-Air

HD Broadcast Server

Commercial Content

On Air Playout

MLB Network

Archive Level

Front Porch Digital
DIVA

SUN SL-8500

HD Video Stored On Data Tape

Warehouse

Off Site
Data Tape Storage

MEDIA MANAGEMENT – THE WORKFLOW

The Architecture

Application Level

Infrastructure Level

Archive Level

DIAMOND
Asset Management System

Aurora & Final Cut Pro
Edit Systems

Grass Valley
Digital Production System

HD Video Storage

Browser Quality Video Storage

HD Play out to Control Rooms, other Facility areas...

Live Feeds (Game Broadcasts)

File Feeds (Game Records)

Archive Content

Finished Programs for MLB Network

Omneon
Play-to-Air

HD Broadcast Server

On Air Playout

MLB Network

Commercial Content

Front Porch Digital
DIVA

SUN SL-8500

HD Video Stored On Data Tape

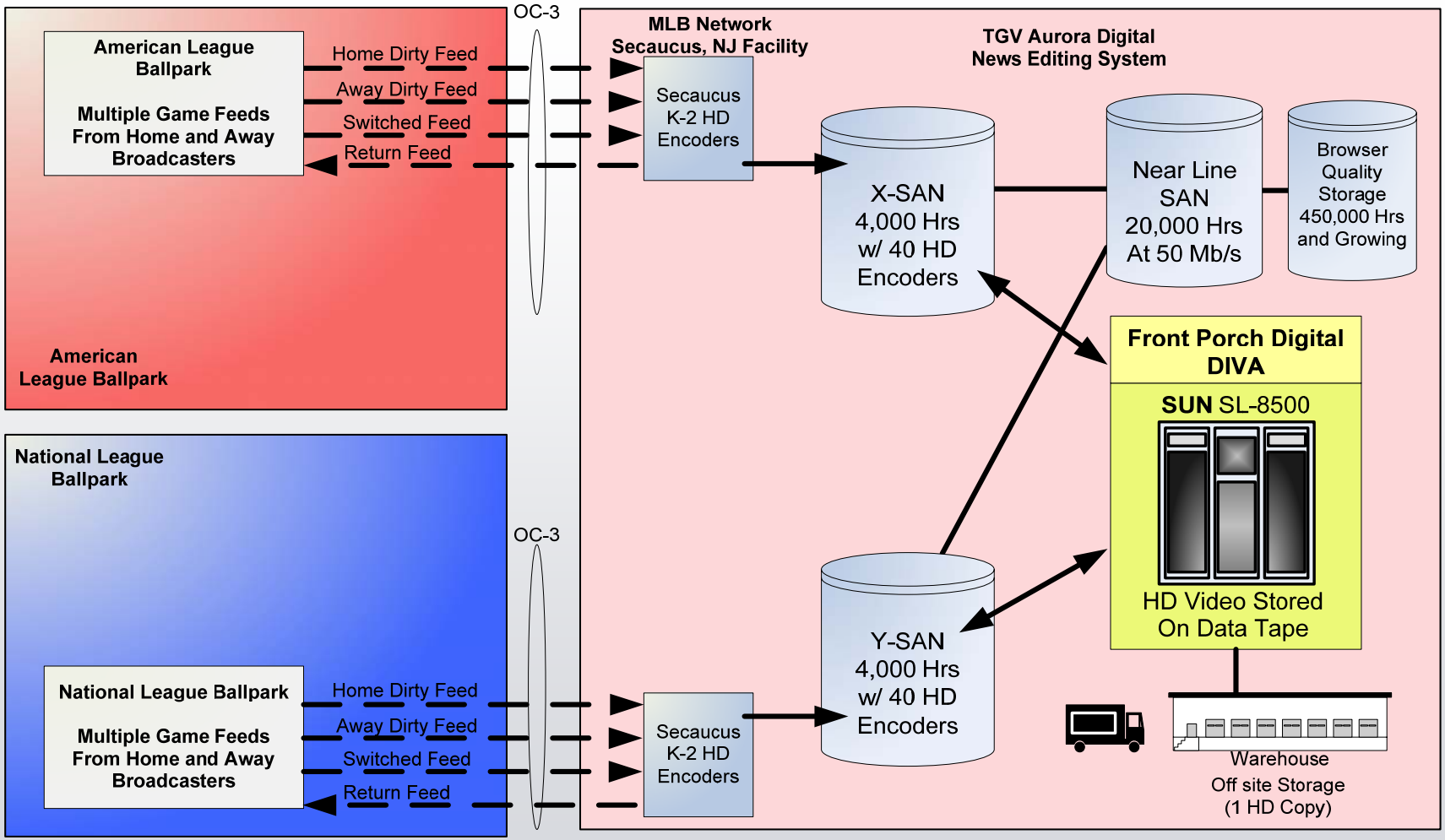
Warehouse

Off Site
Data Tape Storage

MLB BallParkCam – 'Hub and Spoke' Workflows

The Ball Parks

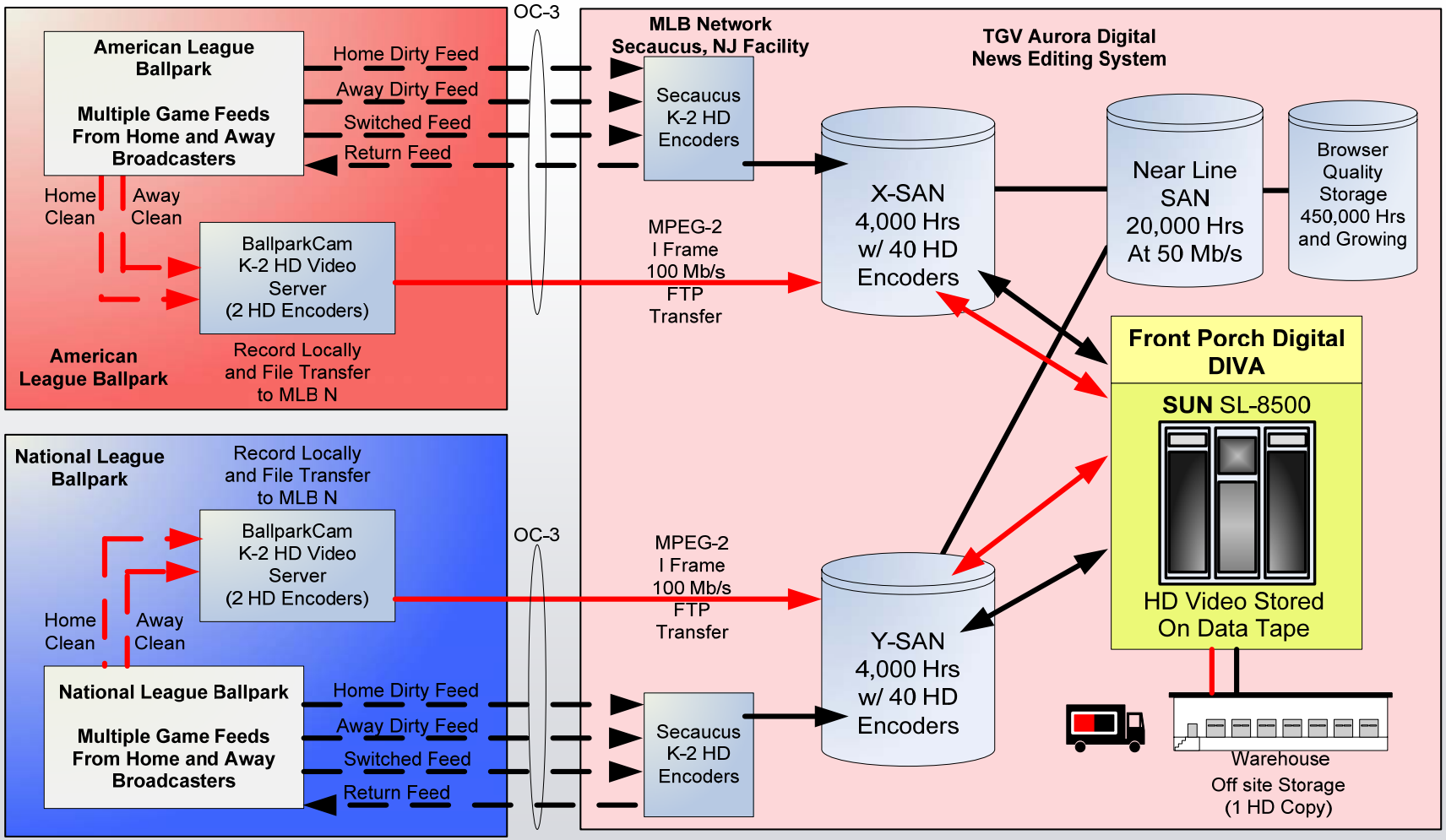
MLB Network Workflow



MLB BallParkCam – 'Hub and Spoke' Workflows

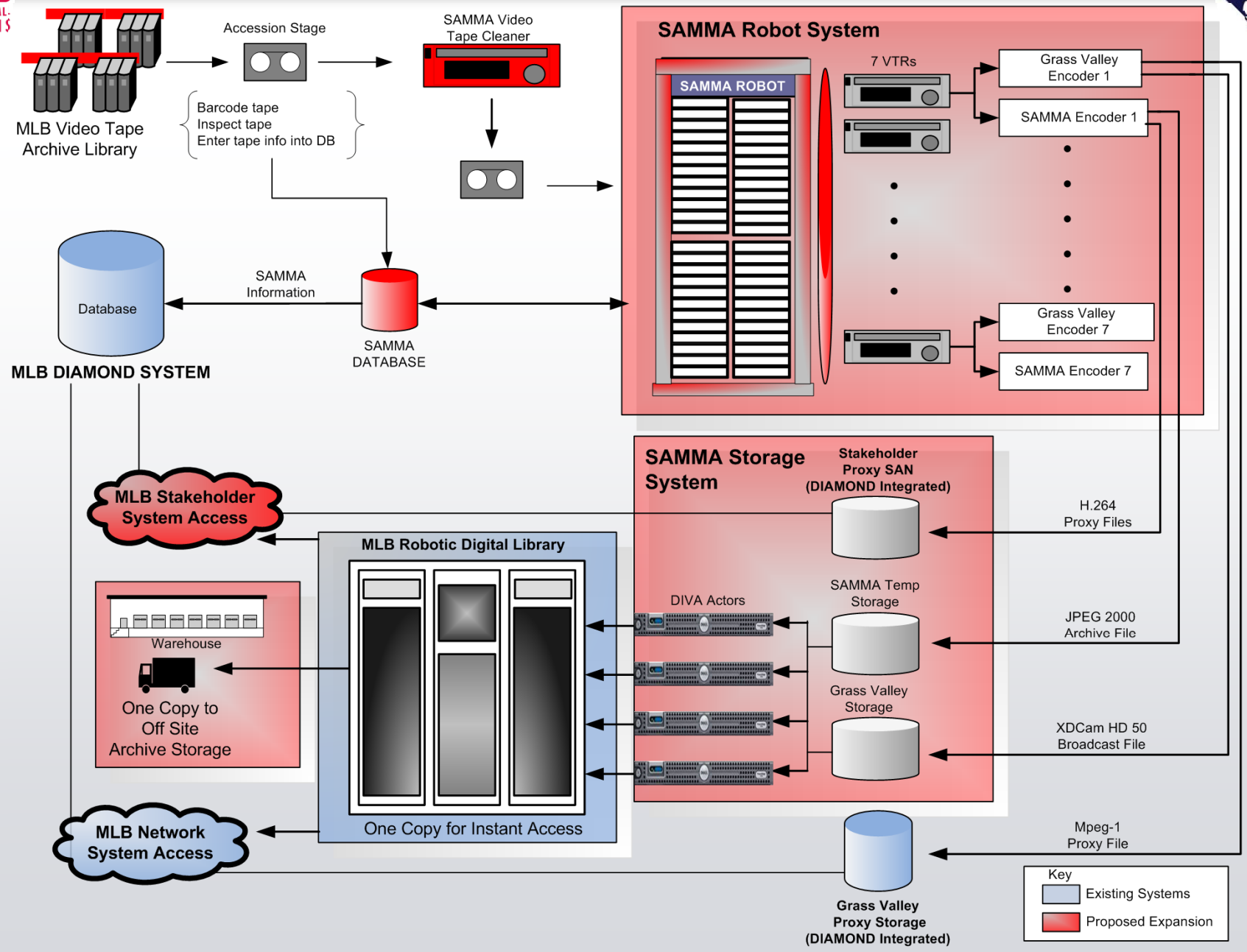
The Ball Parks

MLB Productions Archive

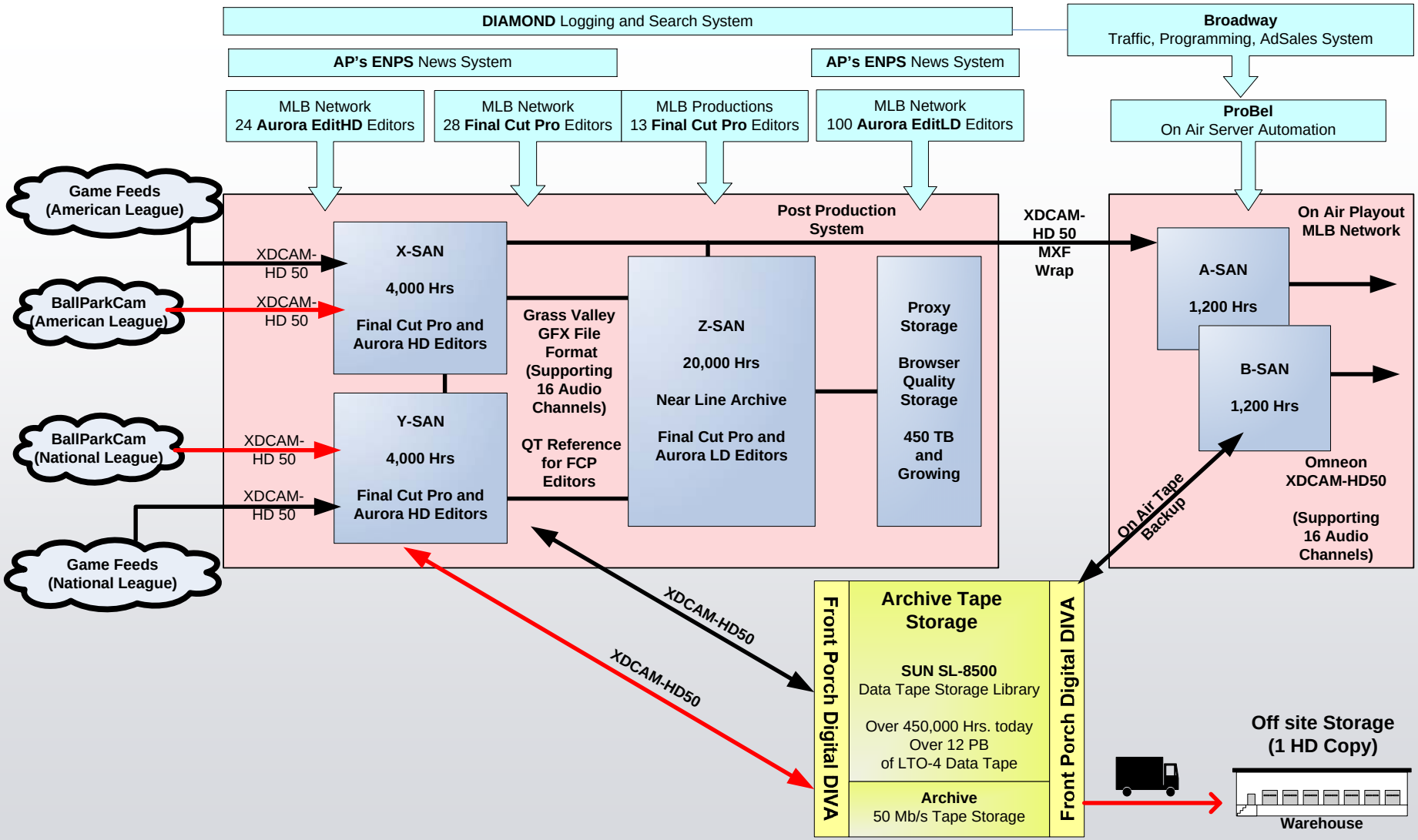


MLB Productions - Archive Digitization Project

MLB Productions - Archive Digitization Project



MLB Facility – File Based Workflow

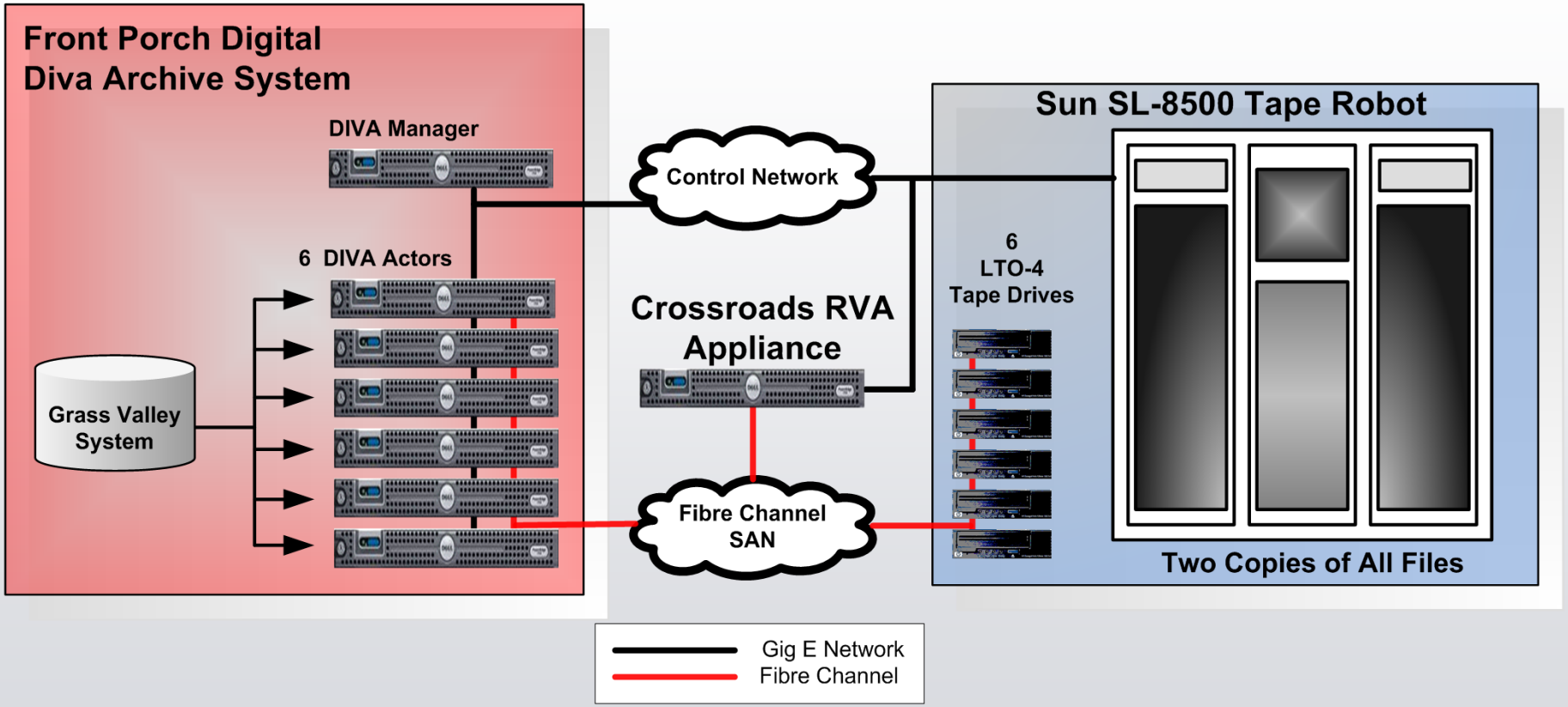


The "Content Wave" Statistics – 2009 Season

- **Dirty Archive Games**
  - 50 Mb/s Video, 720p, XDCamHD50, Long GOP, 16 Audio Tracks
  - 1500 LTO- 4 Tapes for 2009 Season
- **MLB Network Archive - Capture Daily Program Content**
  - 6 – 10 Hours of Content per Day
  - 500 LTO-4 Tapes for 2009 Season
- **Annual 2009 Storage Consumption of LTO-4 = 2,000Tapes**

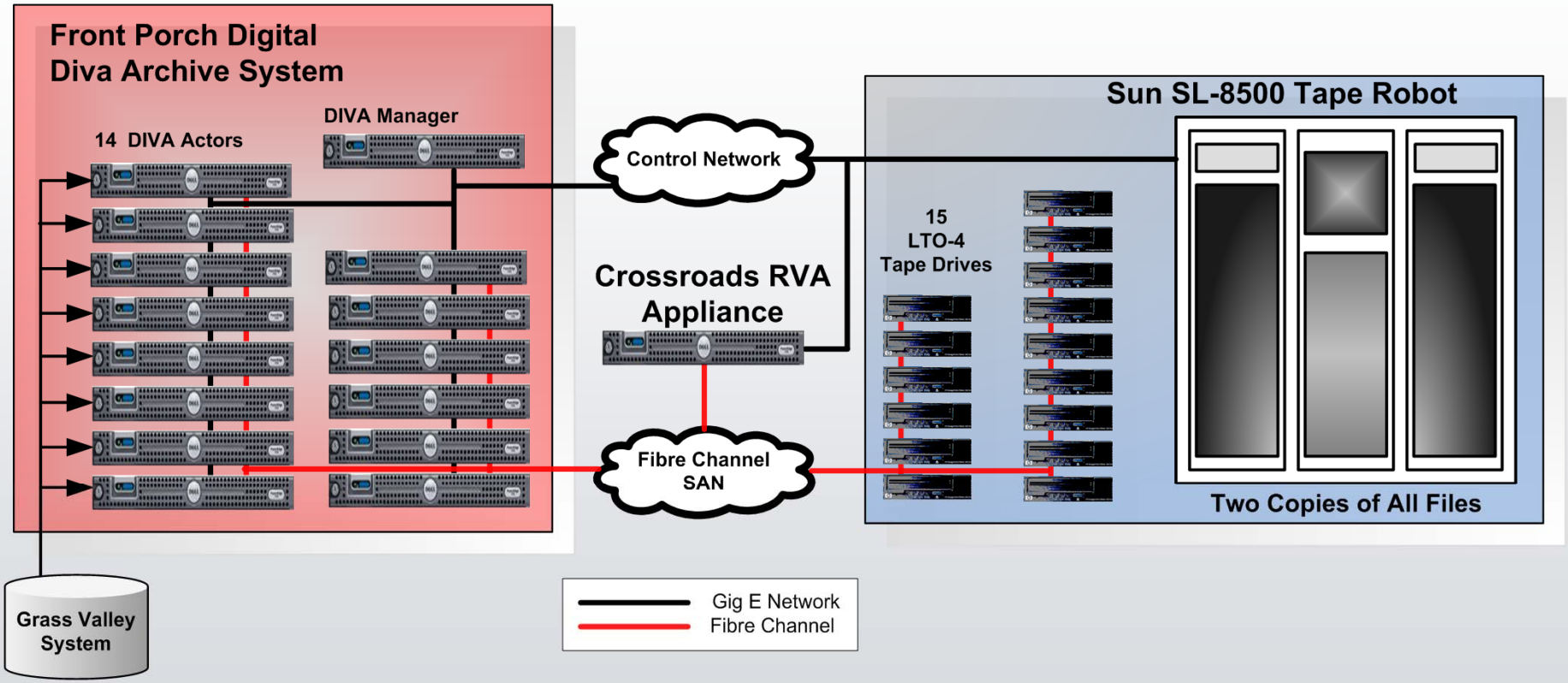
MLB BallParkCam – File Based Workflow

The "Content Wave" Statistics – 2009 Season



MLB BallParkCam – File Based Workflow

The "Content Wave" Statistics – 2010 Season



The "Content Wave" Statistics – 2010 Season

- **Clean Archive Games**
  - 100 Mb/s Video, 720p, MPEG-2, I-Frame Only, 16 Audio Tracks
  - 1500 LTO- 4 Tapes Per Year
- **Dirty Archive Games**
  - 50 Mb/s Video, 720p, XDCamHD50, Long GOP, 16 Audio Tracks
  - 1500 LTO- 4 Tapes Per Year
- **MLB Productions Archive - Capture up to 100 + Hours of content Per Day**
  - Four Encoder Versions of Every Recording
  - 200 LTO- 4 Tapes Per Year
- **MLB Network Archive - Capture an additional 20 + Hours of content Per Day**
  - 500 LTO-4 Tapes Per Year
- **Annual Storage Consumption of LTO-4 = 3,700 Tapes**

The "Content Wave" Statistics – 2011 Season

- **Clean Archive Games**
  - 50 Mb/s Video, 720p, XDCamHD50, Long GOP, 16 Audio Tracks
  - 1500 LTO- 4 Tapes Per Year
- **Dirty Archive Games**
  - 50 Mb/s Video, 720p, XDCamHD50, Long GOP, 16 Audio Tracks
  - 1500 LTO- 4 Tapes Per Year
- **MLB Productions Archive - Capture up to 100 + Hours of content Per Day**
  - Four Encoder Versions of Every Recording
  - 800 LTO- 4 Tapes Per Year
- **MLB Network Archive - Capture an additional 40 + Hours of content Per Day**
  - 800 LTO-4 Tapes Per Year
- **Annual Storage Consumption of LTO-4 = 4,600 Tapes**

The "Content Wave" Statistics – 2012 Season

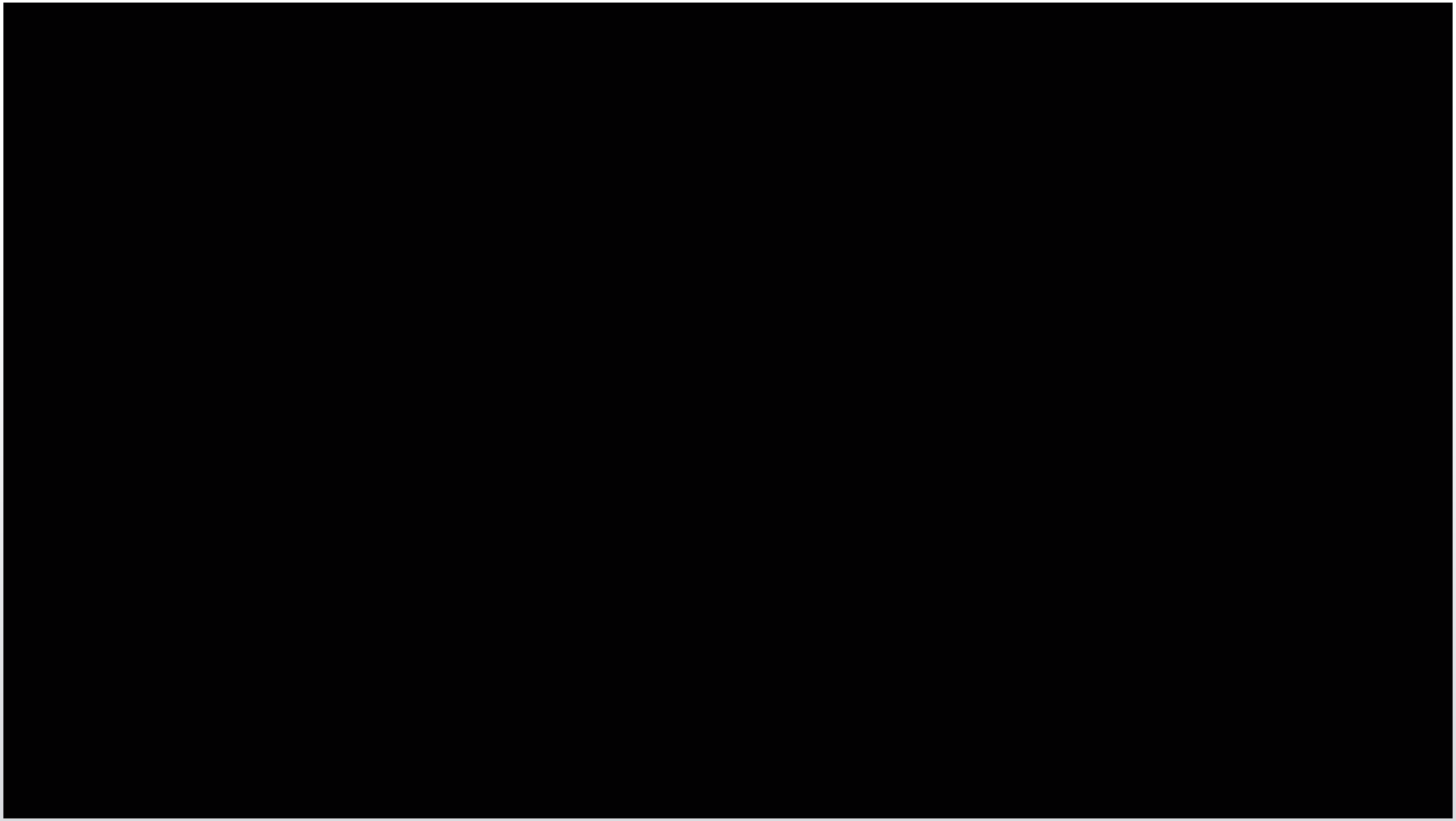
- **Clean Archive Games**
  - 50 Mb/s Video, 720p, XDCamHD50, Long GOP, 16 Audio Tracks
  - 1,500 LTO- 4 Tapes Per Year
- **Dirty Archive Games (Pre-Game & Post-Game Shows included)**
  - 50 Mb/s Video, 720p, XDCamHD50, Long GOP, 16 Audio Tracks
  - 1,800 LTO- 4 Tapes Per Year
- **MLB Productions Archive - Capture up to 100 + Hours of content Per Day**
  - 800 LTO- 4 Tapes Per Year
- **MLB Network Archive - Capture an additional 120 + Hours of content Per Day**
  - Expanded 'Live' Broadcast day to 14 hours in March 2012 …
  - Expanding 'Live' Broadcast day again in October by adding a 4 hour Morning Show
  - 2,000 LTO-4 Tapes Per Year
- **Annual Storage Consumption of LTO-4 = 6,100 Tapes**

Mining the MLB Archives

Digitized
Industry
Assets
Managed
Optimally
for
Networked
Distribution

No baseball game is a single event, but rather a collection of thousands of events . . . . DIAMOND allows us to efficiently capture, thoroughly catalog, and visually publish the audio-visual and statistical history of the game.

Mining the MLB Archives



Digitized
Industry
Assets
Managed
Optimally
for
Networked
Distribution

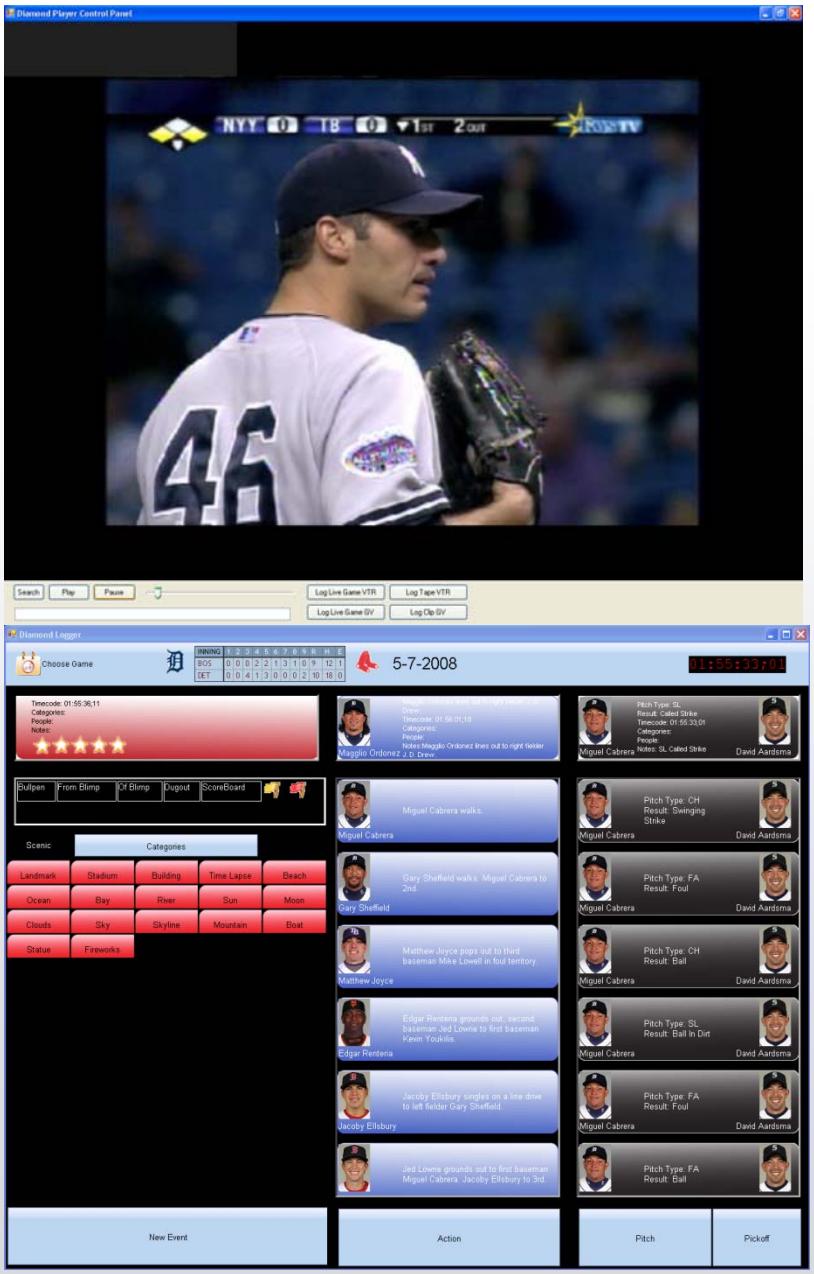
DIAMOND
Asset Management System

Mining the MLB Archives
The Power of DIAMOND

20 DIAMOND Logger Workstations

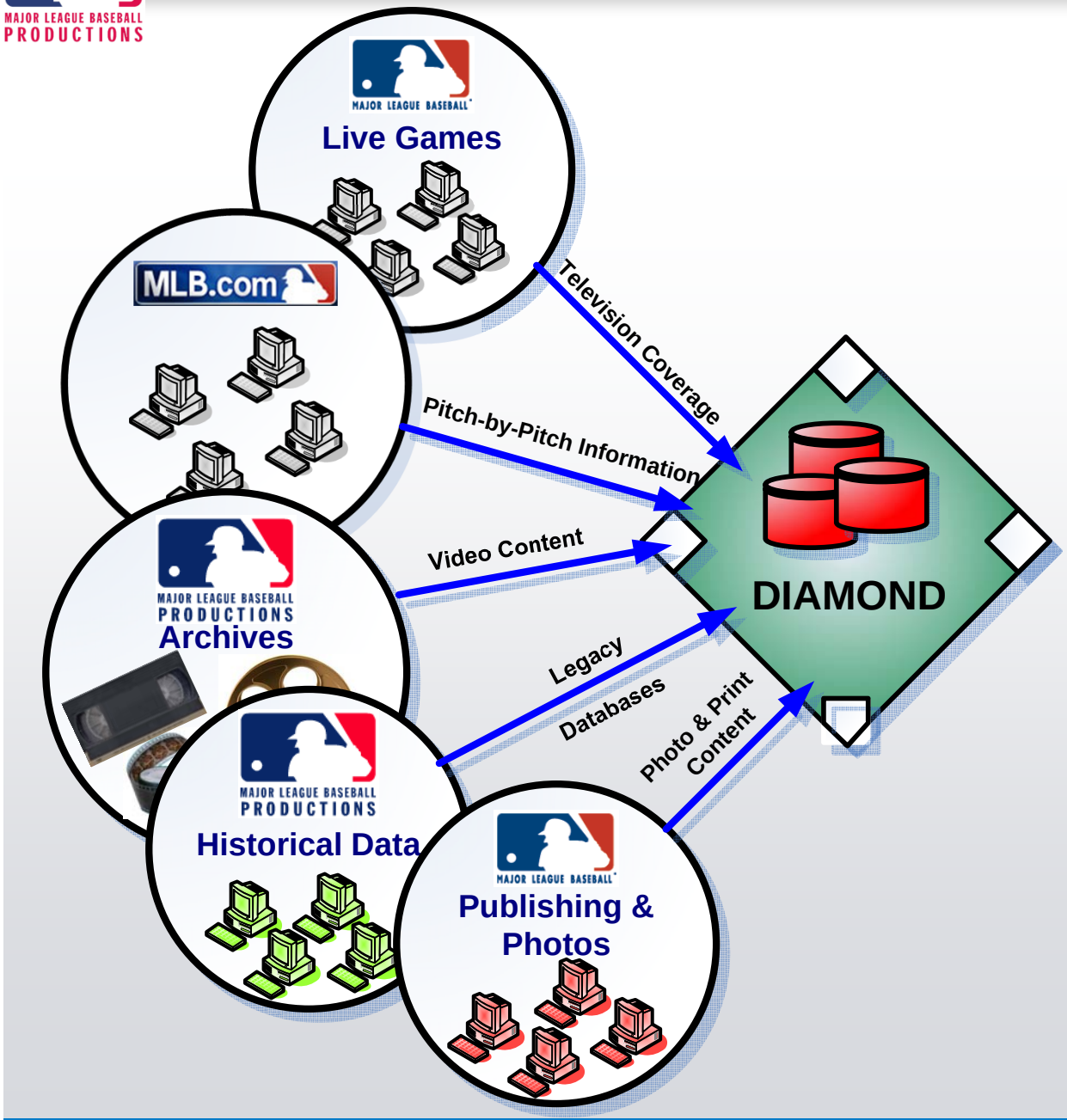
Mining the MLB Archives
The Power of DIAMOND

20 DIAMOND Logger Workstations

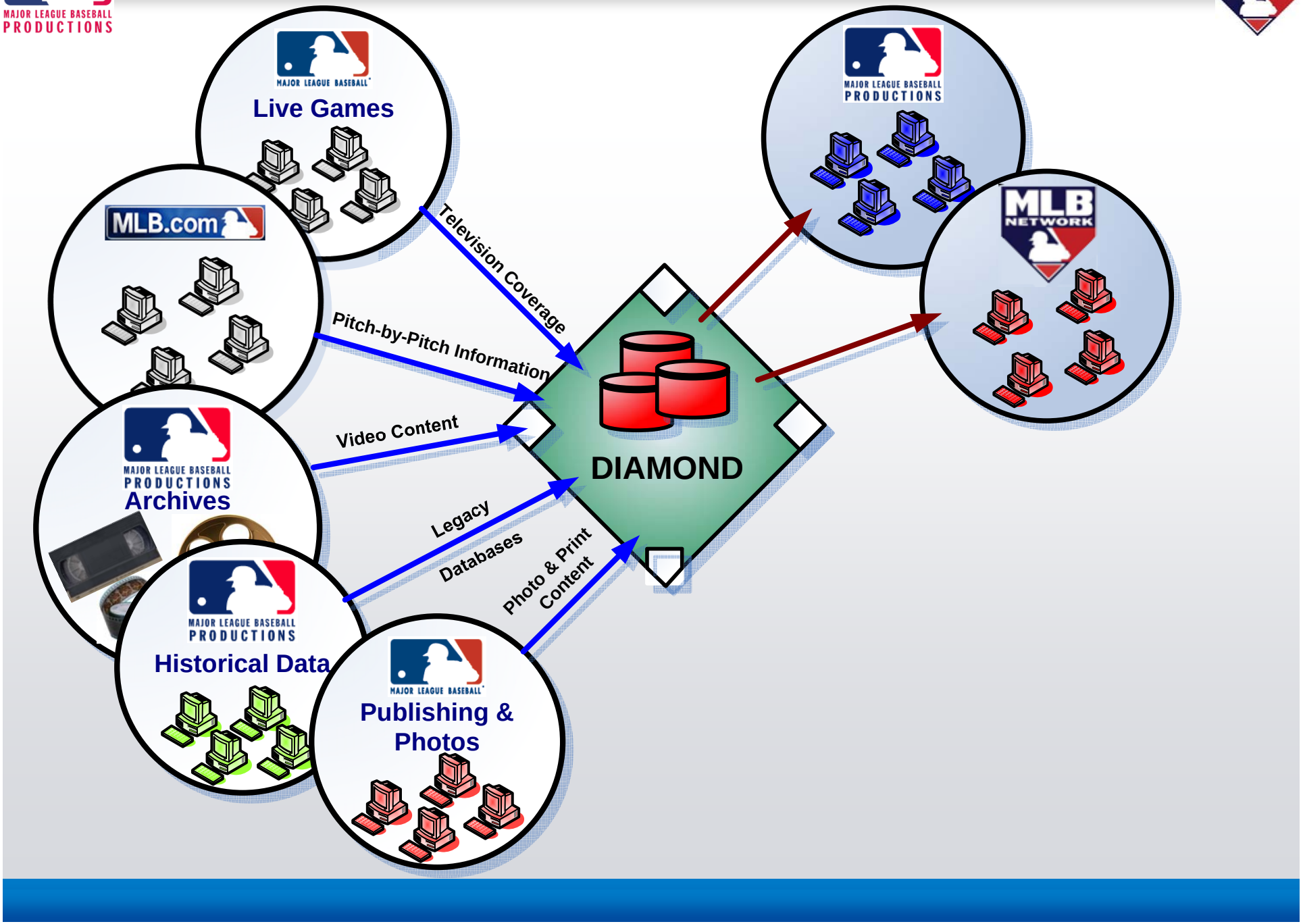

Diamond Player Control Panel
Search
Play
Pause
Log Live Game VTR
Log Tape VTR
Log Live Game SV
Log Clip SV
Diamond Logger
Choose Game
5-7-2008
Bullpen
From Blimp
Of Blimp
Dugout
ScoreBoard
Scenic
Categories
Landmark
Stadium
Building
Time Lapse
Beach
Ocean
Bay
River
Sun
Moon
Clouds
Sky
Skyline
Mountain
Boat
Statue
Fireworks
New Event
Action
Pitch
Pickoff

Diamond Logger
Choose Game
INNING 1 2 3 4 5 6 7 8 9 R H E
BOS 0 0 0 2 2 1 3 1 0 9 12 1
DET 0 0 4 1 3 0 0 0 2 10 18 0
5-7-2008
01:55:33;01
Timecode: 01:55:36;11
Categories:
People:
Notes:
Magglio Ordonez lines out to right fielder J. D. Drew.
Timecode: 01:56:01;18
Categories:
People:
Notes:Magglio Ordonez lines out to right fielder J. D. Drew.
Magglio Ordonez
Pitch Type: SL
Result: Called Strike
Timecode: 01:55:33;01
Categories:
People:
Notes: SL Called Strike
Miguel Cabrera
David Aardsma
Bullpen
From Blimp
Of Blimp
Dugout
ScoreBoard
Scenic
Categories
Landmark
Stadium
Building
Time Lapse
Beach
Ocean
Bay
River
Sun
Moon
Clouds
Sky
Skyline
Mountain
Boat
Statue
Fireworks
Miguel Cabrera walks.
Miguel Cabrera
Pitch Type: CH
Result: Swinging Strike
Miguel Cabrera
David Aardsma
Gary Sheffield walks. Miguel Cabrera to 2nd.
Gary Sheffield
Pitch Type: FA
Result: Foul
Miguel Cabrera
David Aardsma
Matthew Joyce pops out to third baseman Mike Lowell in foul territory.
Matthew Joyce
Pitch Type: CH
Result: Ball
Miguel Cabrera
David Aardsma
Edgar Renteria grounds out, second baseman Jed Lowrie to first baseman Kevin Youkilis.
Edgar Renteria
Pitch Type: SL
Result: Ball In Dirt
Miguel Cabrera
David Aardsma
Jacoby Ellsbury singles on a line drive to left fielder Gary Sheffield.
Jacoby Ellsbury
Pitch Type: FA
Result: Foul
Miguel Cabrera
David Aardsma
Jed Lowrie grounds out to first baseman Miguel Cabrera. Jacoby Ellsbury to 3rd.
Pitch Type: FA
Result: Ball
New Event
Action
Pitch
Pickoff

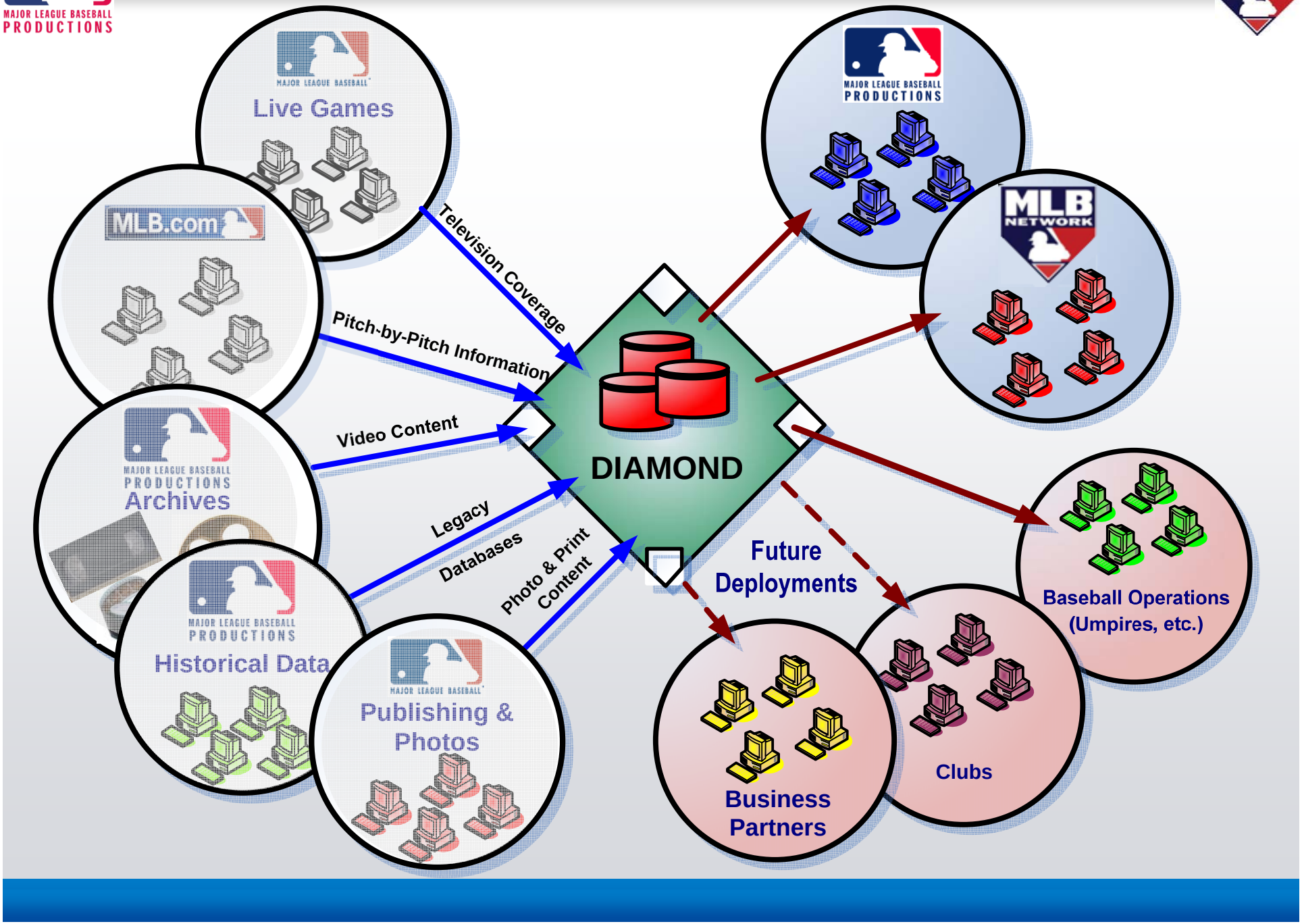
DIAMOND INPUT SOURCES



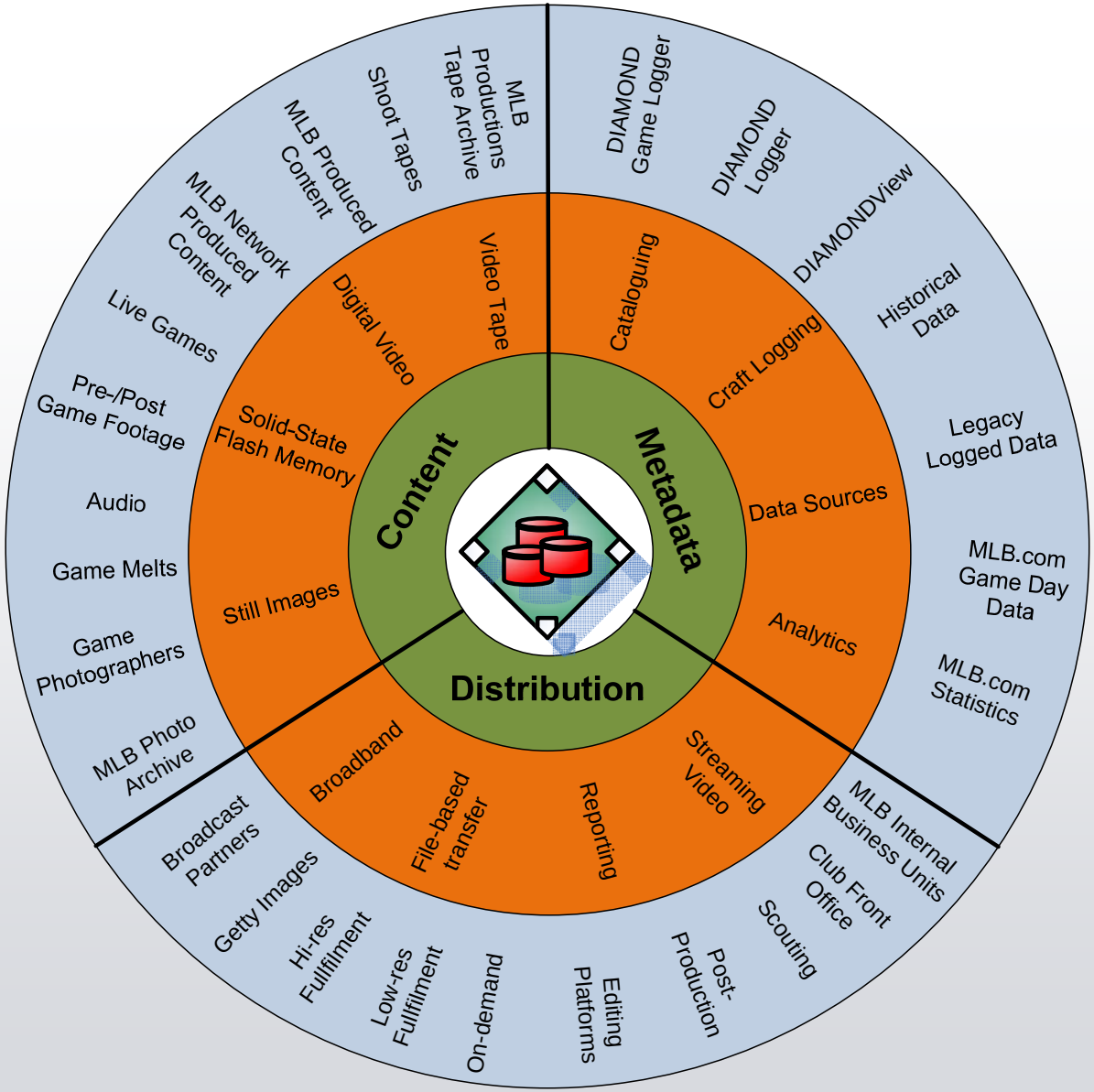
DIAMOND STAKEHOLDERS



DIAMOND STAKEHOLDERS



DIAMOND Sources & Destinations



Mining the MLB Archives

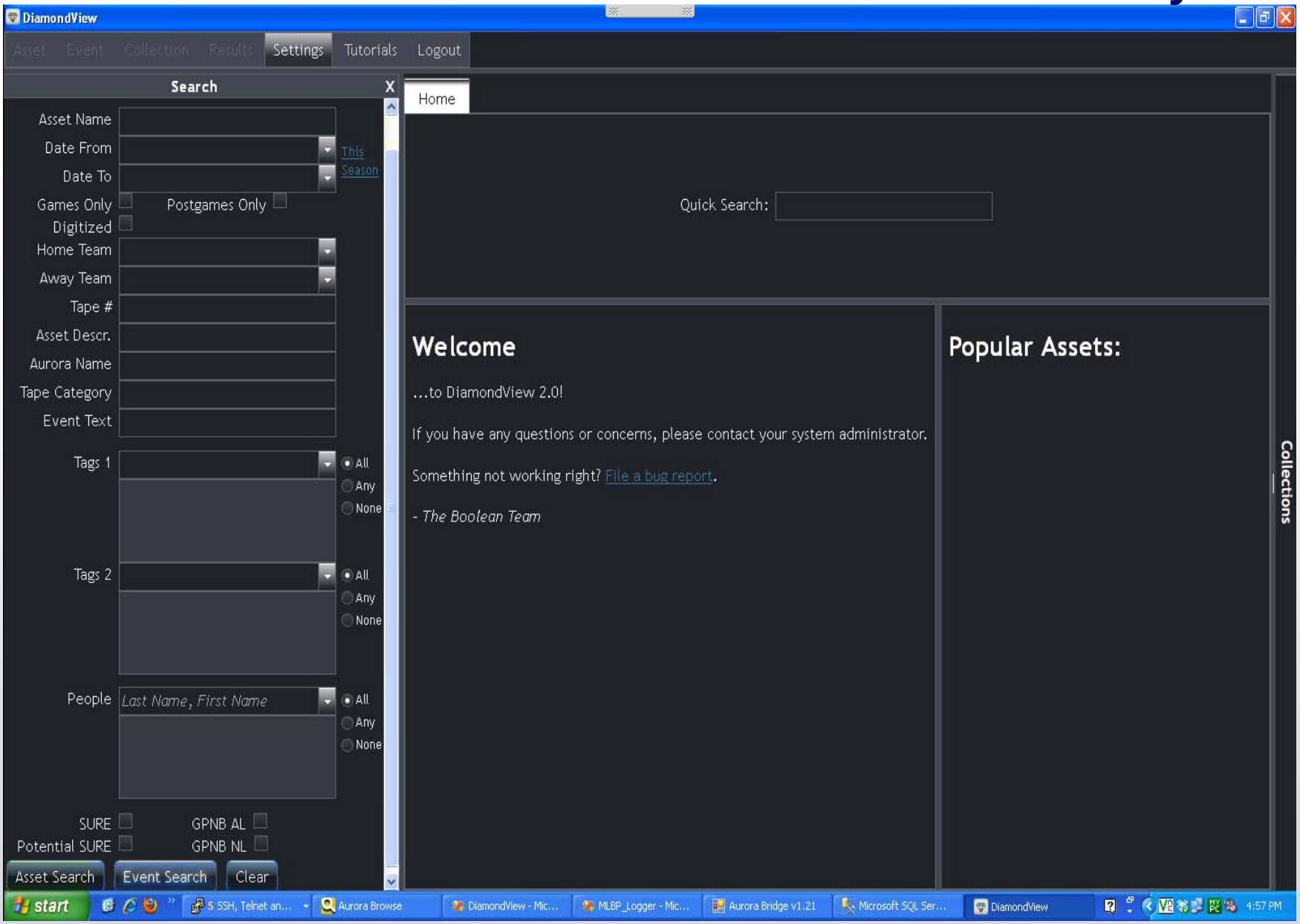
The Power of DIAMOND

- **Capture the Video Assets from Multiple Sources**
  - 'Live' Game Recordings and other Game Feeds (Melts, BallParkCam, etc.)
  - Archive, Interviews, Field Tapes
- **Catalog the Video Content**
  - Across Multiple Databases
  - With Content Expert Loggers
- **Search**
  - Across Multiple Collections of Information
  - Customized for the Individual End User

- **Repurpose and Monetize the Video Content**
  - High-speed Access to Current and Historical Content
  - Repurposing of Archives
    - By the League: MLB Productions Programming, Partner Programming, and 3rd Party Licensing
    - By the Network: In-game Content; Broadcast Highlights
- **Protect the Investment**
  - Store On Site - Instant Access, Quick Turnaround for Production / Broadcast
  - Archive Off Site - Historical Preservation, Disaster Protection

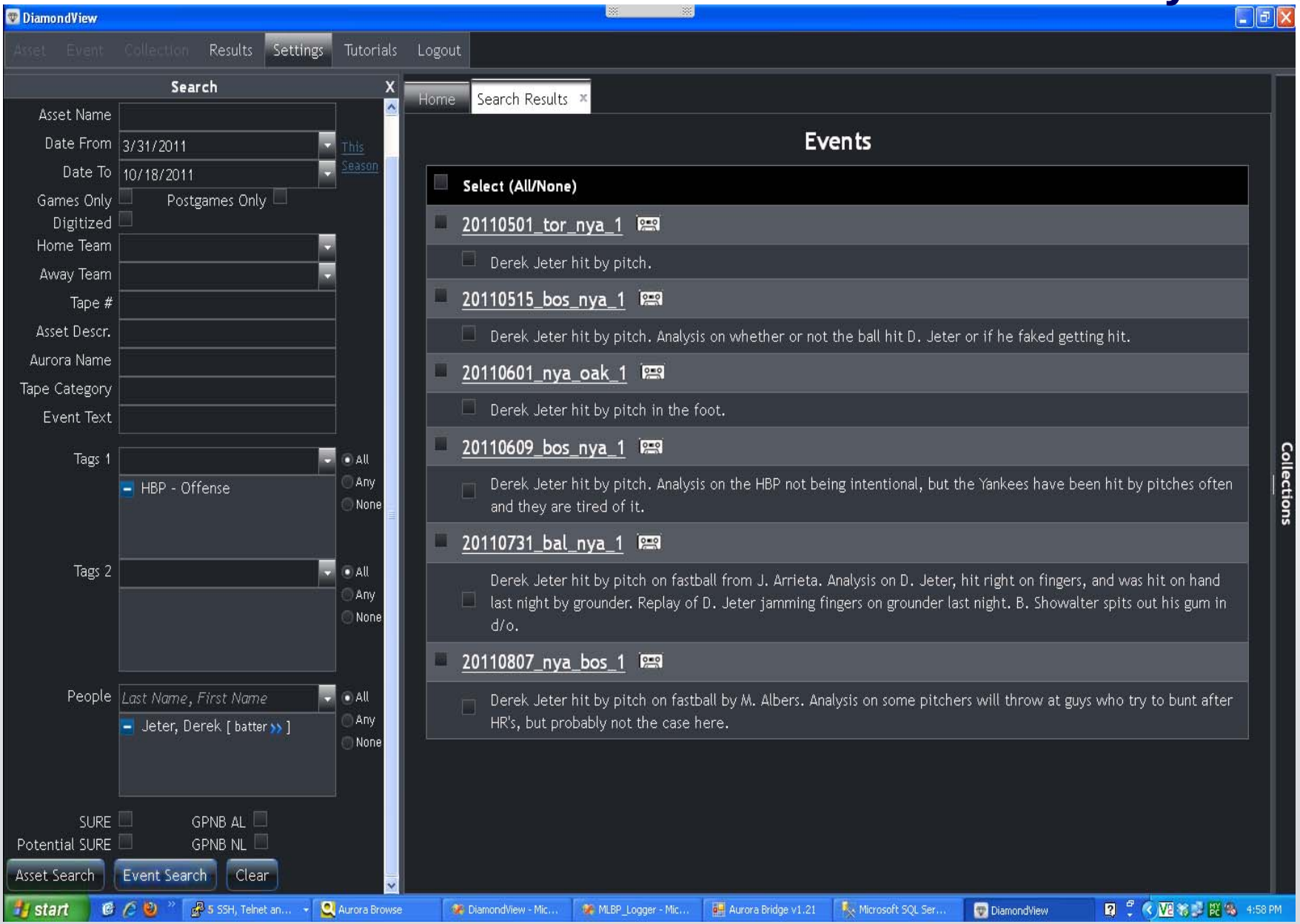
DIAMONDView

Web Based User Search Interface and Video Player



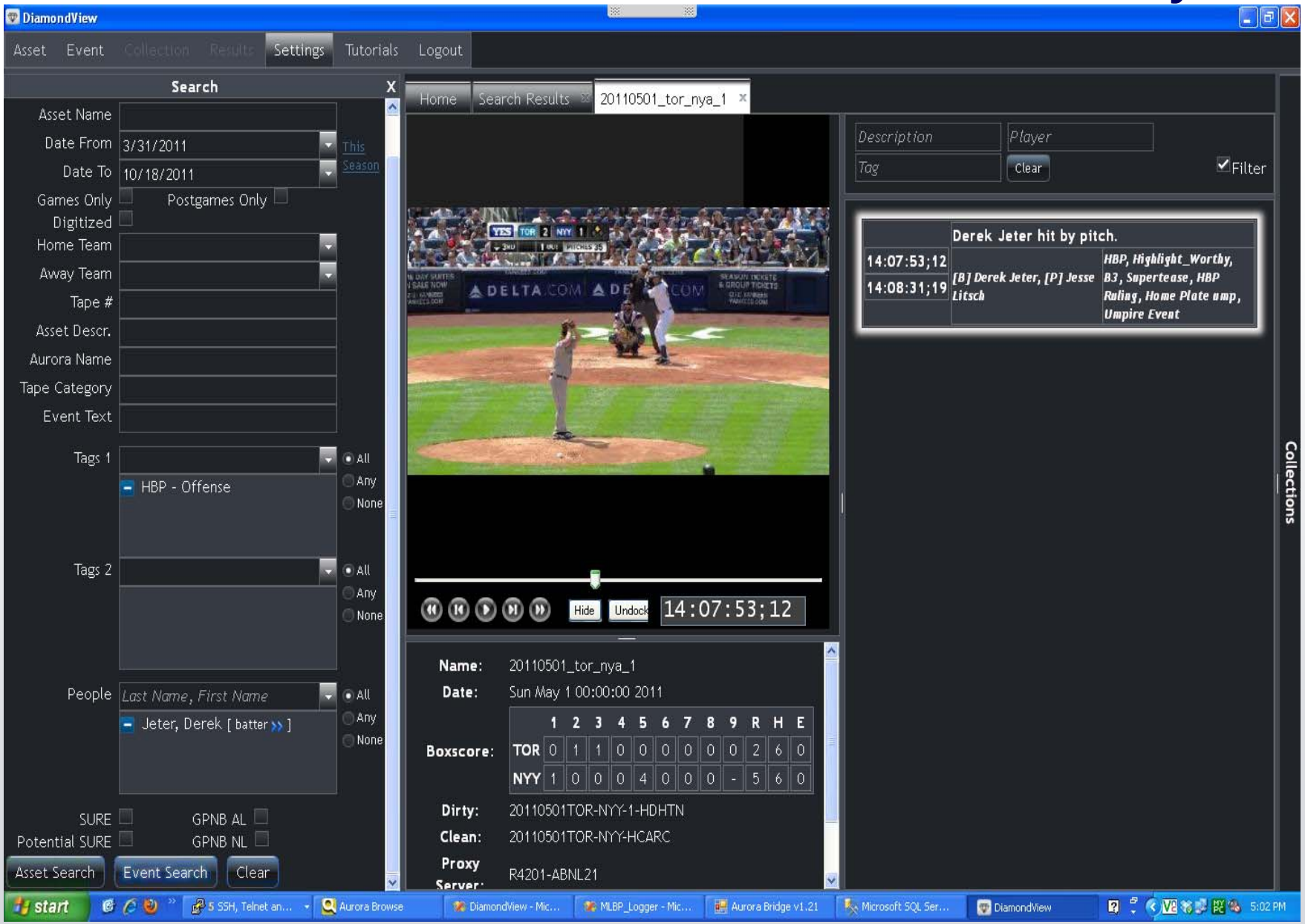
DIAMONDView

Web Based User Search Interface and Video Player



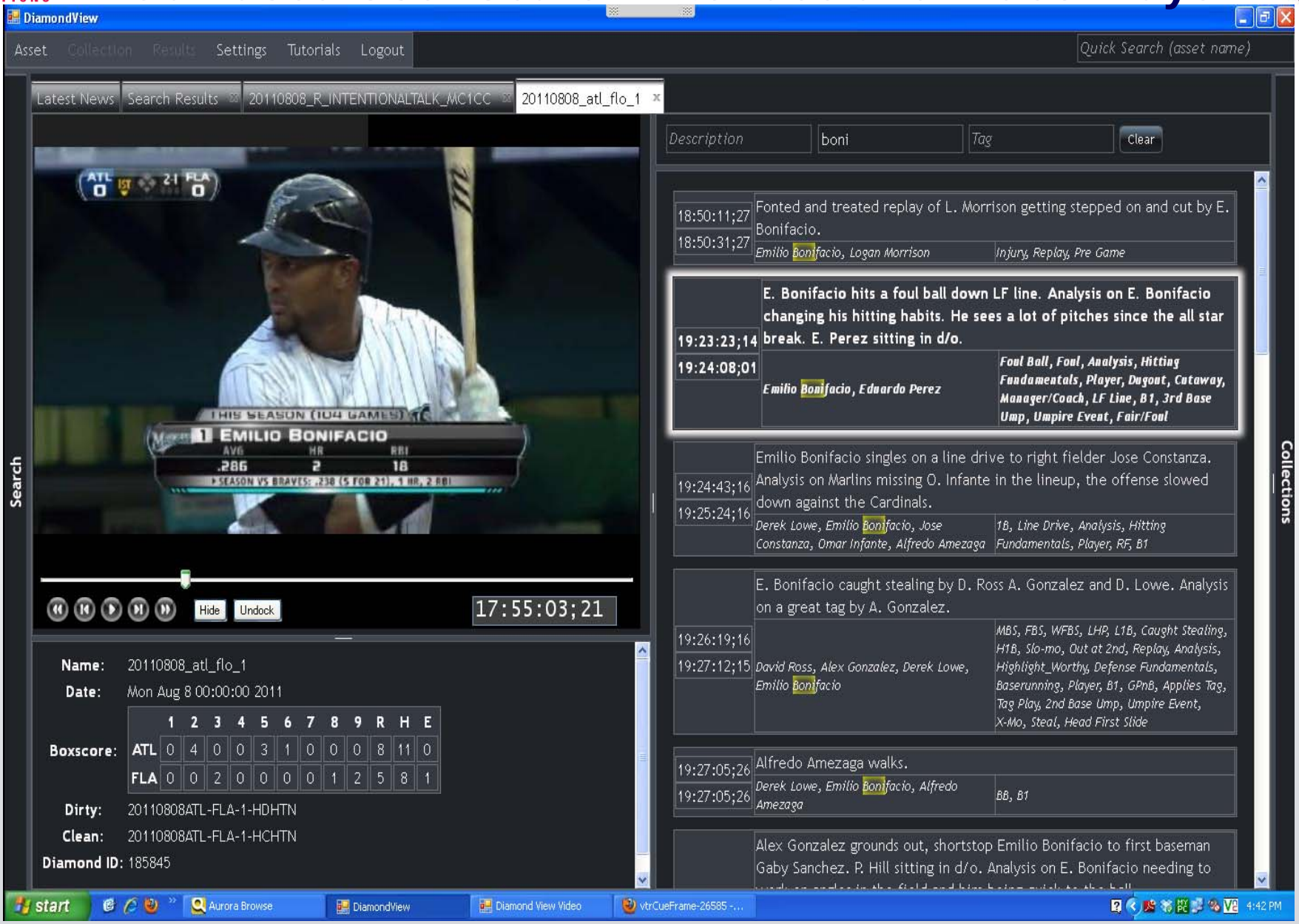
DIAMONDView

Web Based User Search Interface and Video Player



DIAMONDView

Web Based User Search Interface and Video Player



Mining the MLB Archives

The Statistics of DIAMOND

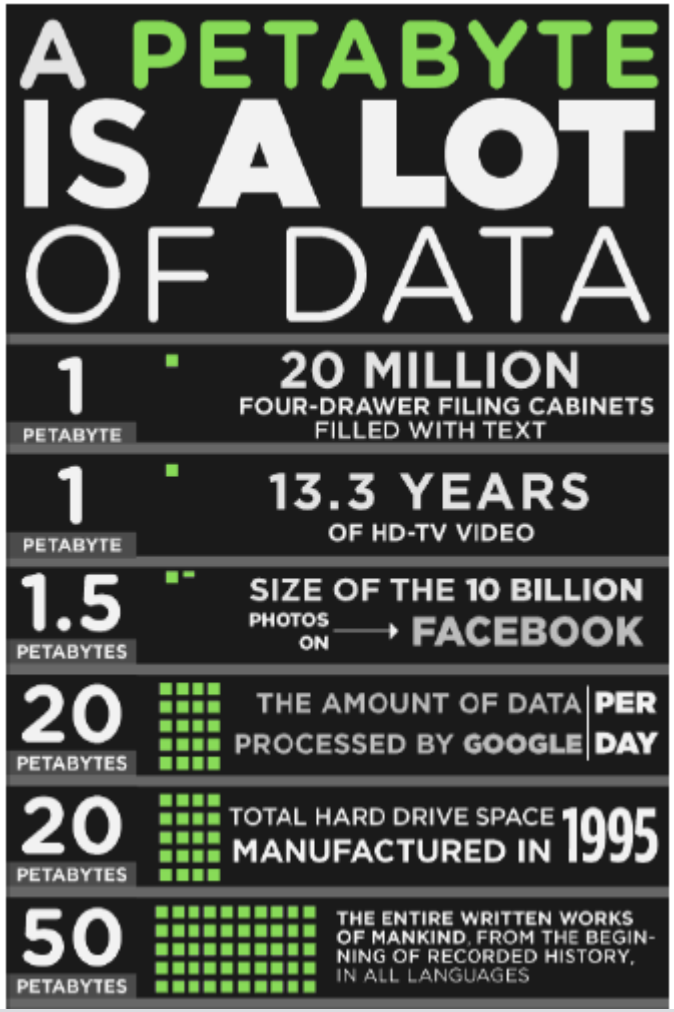
- **Total Assets (Video Tapes & File Records)** — **192,536**
- **Total Hours of Content Logged (Approx.)** — **300,000**

- **Total Individual Events Cataloged** — **4,088,805**
- **Total Individual 'People' Identifiers** — **5,767,117**

- **Total Categorizations** — **12,583,277**

Keeping Up With the Technology…

IBM's Approach to BIG Data



Source: IBM Research Presentation

MLB Network's Approach to BIG Data

- 1 Petabyte of Game Footage was Stored in our GV Library every 10 Weeks during 2011…
- 3.6 Petabytes of Data was Added to our GV Library in 2011…
- Over 12 Petabytes of Data is currently in the GV Library Today…
- In 2012, we will add 4.8+ Petabytes of Data to the GV Library…
- 20-30% increase in amount of Content added to the Library, year over year… and not slowing down…
- By the end of 2013 Season, the GV library will be approaching 20+ Petabytes of Data…

AREA OF DISCUSSIONS – Grass Valley & Archive

The Architecture

Application Level

Infrastructure Level

Live Feeds (Game Broadcasts)

DIAMOND
Asset Management System

Aurora & Final Cut Pro
Edit Systems

Digitized Content
DIAMOND

Grass Valley
Digital Production System

HD Video Storage

Browser Quality Video Storage

HD Play out to Control Rooms, other Facility areas...

Finished Programs for MLB Network

Omneon
Play-to-Air

HD Broadcast Server

On Air Playout

MLB Network

Commercial Content

Archive Level

Front Porch Digital
DIVA

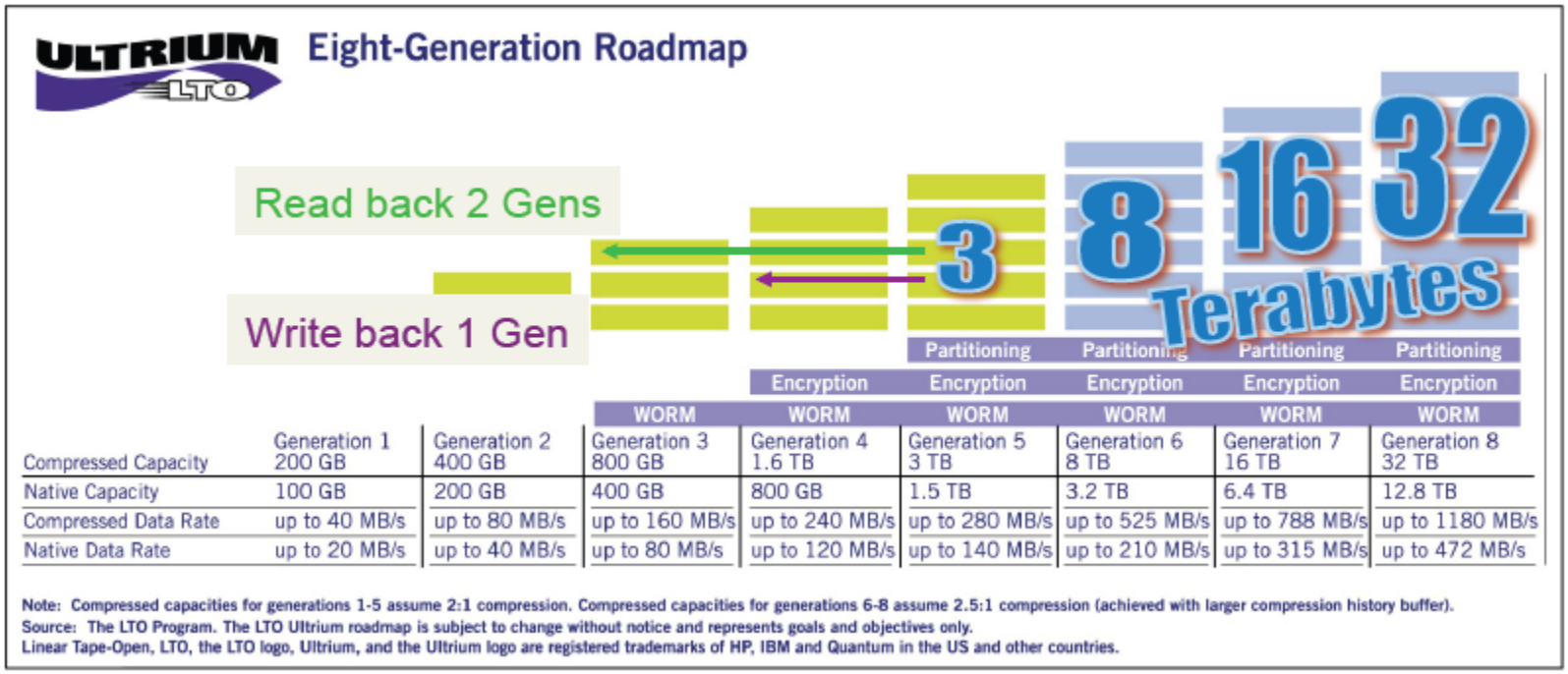
SUN SL-8500

HD Video Stored On Data Tape

Warehouse

Off Site
Data Tape Storage

LTO Span of Data Tape Formats



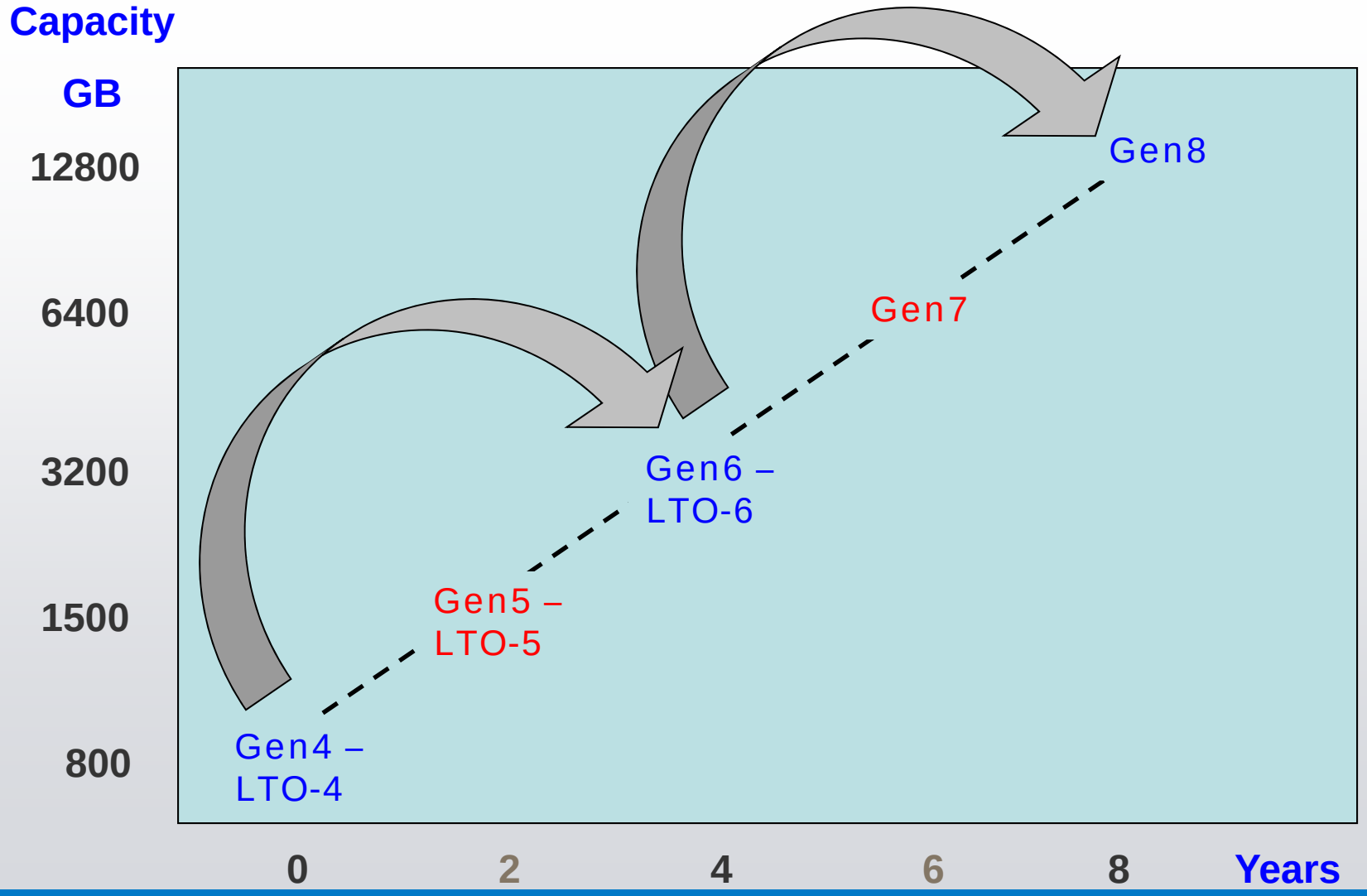
Compressed Capacity	Generation 1 200 GB	Generation 2 400 GB	Generation 3 800 GB	Generation 4 1.6 TB	Generation 5 3 TB	Generation 6 8 TB	Generation 7 16 TB	Generation 8 32 TB
Native Capacity	100 GB	200 GB	400 GB	800 GB	1.5 TB	3.2 TB	6.4 TB	12.8 TB
Compressed Data Rate	up to 40 MB/s	up to 80 MB/s	up to 160 MB/s	up to 240 MB/s	up to 280 MB/s	up to 525 MB/s	up to 788 MB/s	up to 1180 MB/s
Native Data Rate	up to 20 MB/s	up to 40 MB/s	up to 80 MB/s	up to 120 MB/s	up to 140 MB/s	up to 210 MB/s	up to 315 MB/s	up to 472 MB/s

Note: Compressed capacities for generations 1-5 assume 2:1 compression. Compressed capacities for generations 6-8 assume 2.5:1 compression (achieved with larger compression history buffer).
Source: The LTO Program. The LTO Ultrium roadmap is subject to change without notice and represents goals and objectives only.
Linear Tape-Open, LTO, the LTO logo, Ultrium, and the Ultrium logo are registered trademarks of HP, IBM and Quantum in the US and other countries.

Source: LTO Technology Presentation

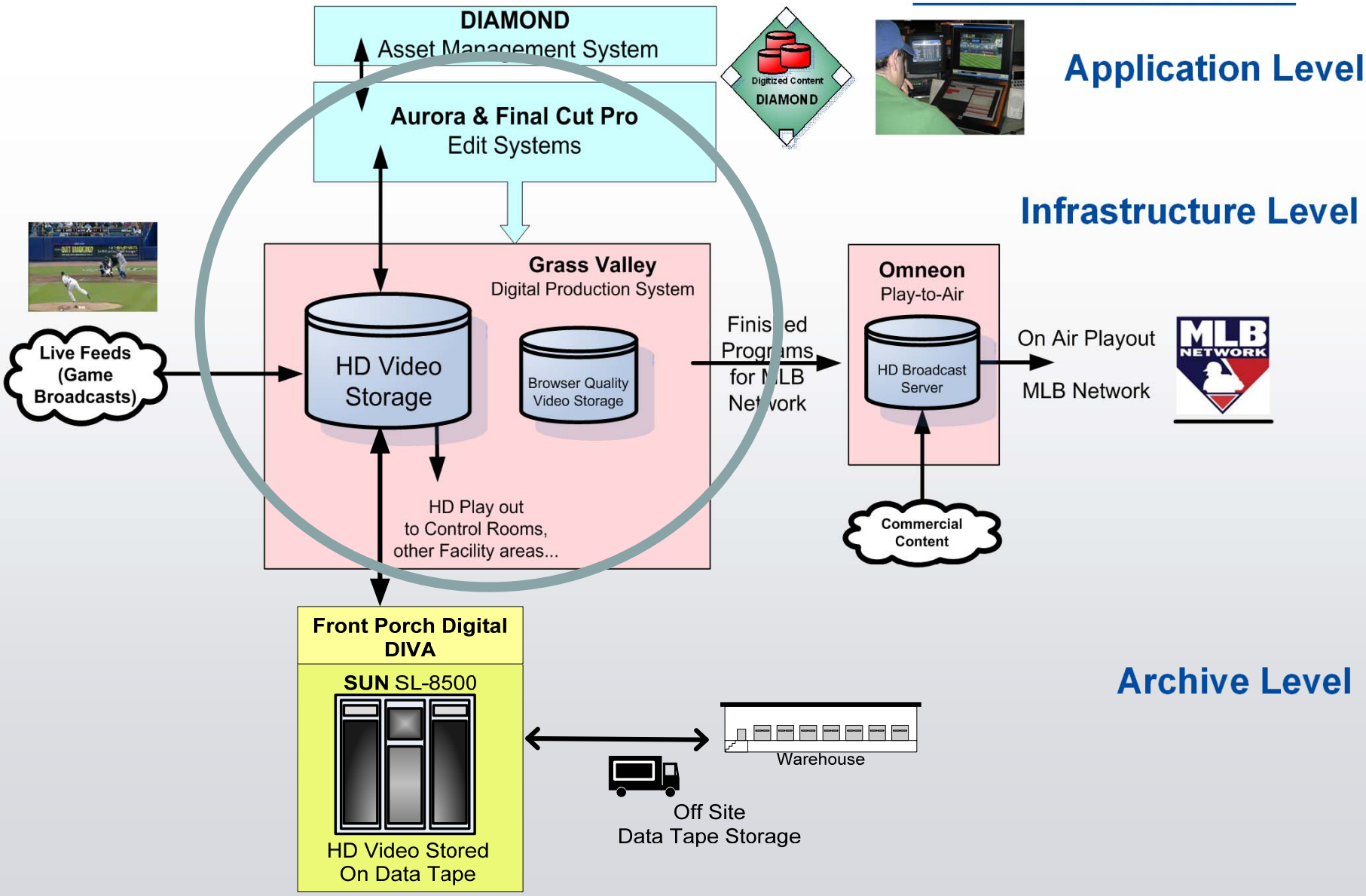
LTO Generation Skipping

Avg. 5-6 year rolling window



AREA OF DISCUSSIONS – Grass Valley & Archive

The Architecture



Testing Workflows in the Broadcast Environment

❖ Final Cut Pro Project Archiving

- ❖ Developed a testing plan and were successful in deploying an 'Operator controlled' project archive workflow within Final Cut Pro Edit Environment.

❖ Proxy Storage for Grass Valley

- ❖ Currently in development and testing of Proxy 'On-Line' disk environment being extended into 'On-Line' mixed Disk / Tape environment for 'Long-Tail Proxy Content'

STRONGBOX T1

- *1 U server*
- *Supports 200 million files*
- *Maximum File transfer Rate: 160 MB/s*

STRONGBOX T3

- *3 U server*
- *Supports up to 5 billion files*
- *Maximum File transfer Rate: 600 MB/s*

Emerging LTO Tape Technologies

Significant Cost Reducing Technologies for Big Media…

LTFS

Linear Tape File System

ULTRIUM LTO

IBM

hp

- Self describing file system
- Faster access to data
- Simple drag and drop
- Compatibility across environment
- Increased data mobility
- A single storage media standard

The major (and I would almost classify it as a breakthrough) advance that LTFS provides is that LTFS changes tape from application dependent to application independent.

Jerome Wendt DCIG

MAJOR LEAGUE BASEBALL PRODUCTIONS

MLB NETWORK

Discuss:

1. Proxy Challenge
2. Editor Challenge
3. DIAMOND Integration Challenge

The MLB Network

MAJOR LEAGUE BASEBALL PRODUCTIONS

Questions

The MLB Network

Thank you

The MLB Network