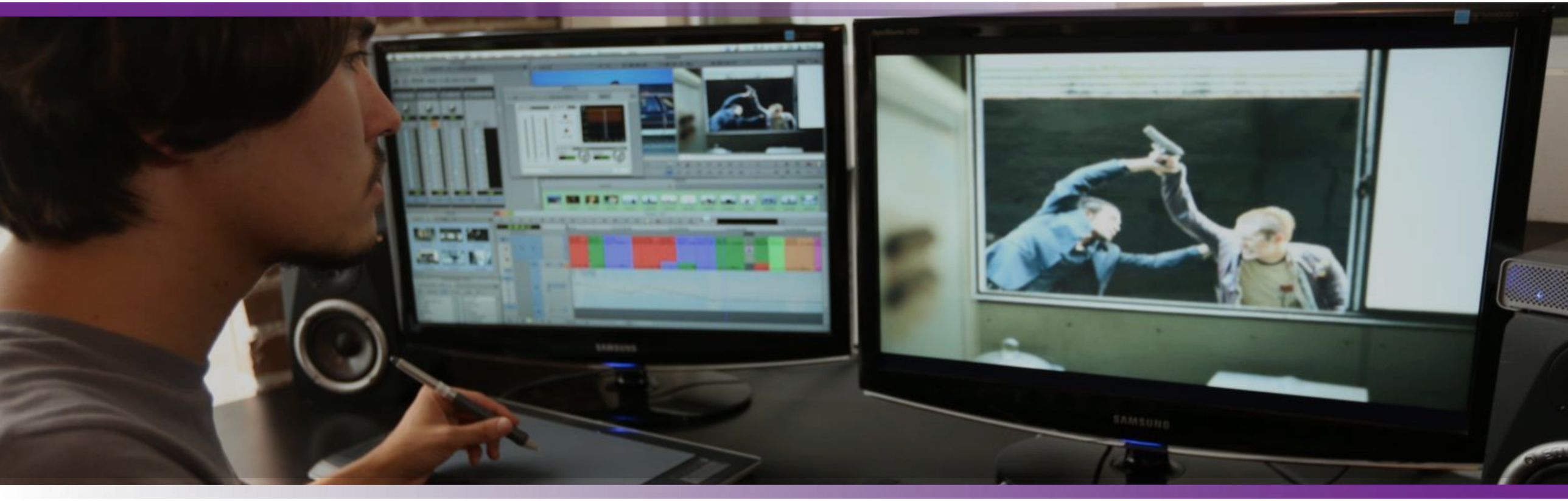


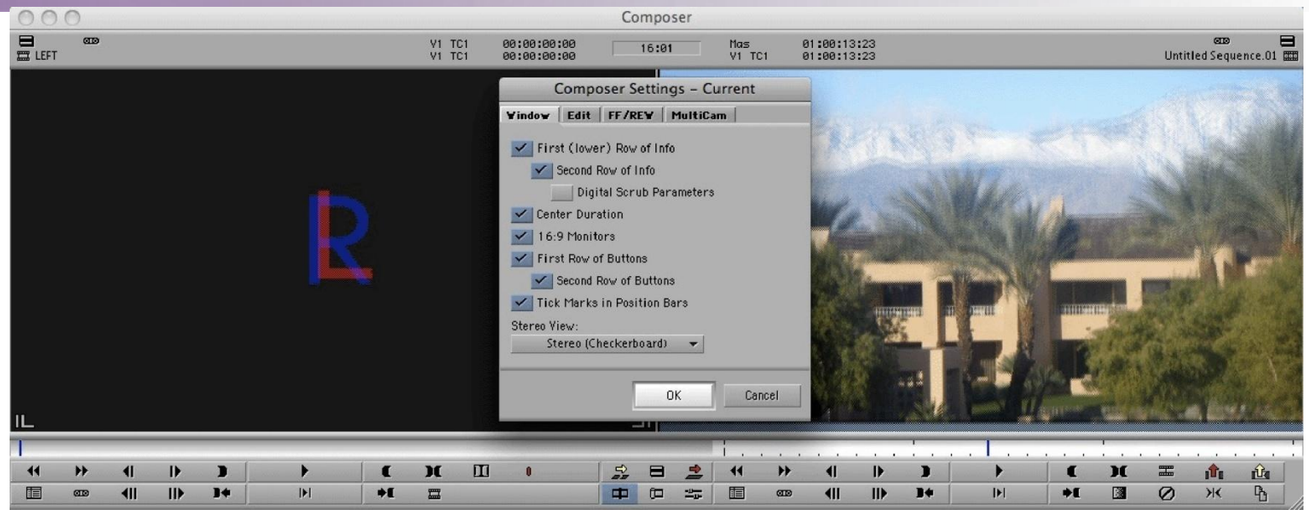
Avid's Stereoscopic Editorial Architecture

Light Fields, Intelligent Computing and Beyond

Shailendra Mathur, Video Chief Architect | March 19th 2013

Topics

- Stereo 3D data handling
- Avid Intelligent Compute Architecture



Avid introduction

- Tools for Film, Broadcast and Live productions
- Video and Audio
- Focus today on Film editorial editing tool – Media Composer
- Our customers are Academy, Grammies and Emmy award winners,
  - … and so are we!

Stereo 3D Data Handling

Trends

- Increasing granularity (resolution) of video
  - High frame rate usage
  - High spatial resolution
  - High bit-depth
- Increasing number of views
  - Stereo shooting - Mono and S3D deliverables
  - Plenoptic/Light Field Cameras
  - Camera Grids

Our Vision!

“Super-sample the work, edit it later!”

- The Story-teller editing the film wins!

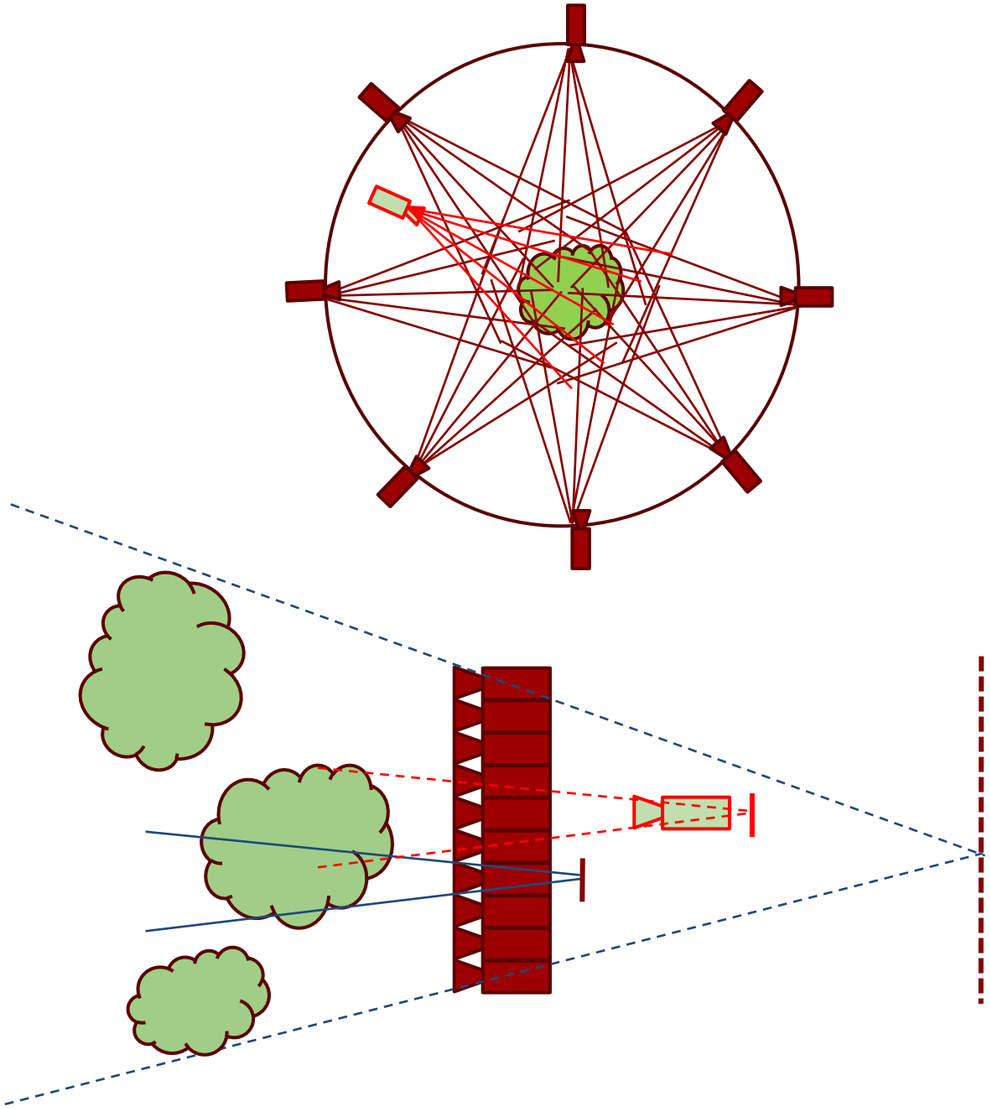
- More elements of the scene to work with
  - Zoom-in,
  - Pan and scan
  - Slow-down,
  - Speed-up
  - Increase contrast,
  - Expose new dynamic ranges

…without losing display precision!

Problems to solve

- Need a data model for managing multiple views and resolutions
- Support common editing and data management functions and workflows

Light Fields – a parametric model



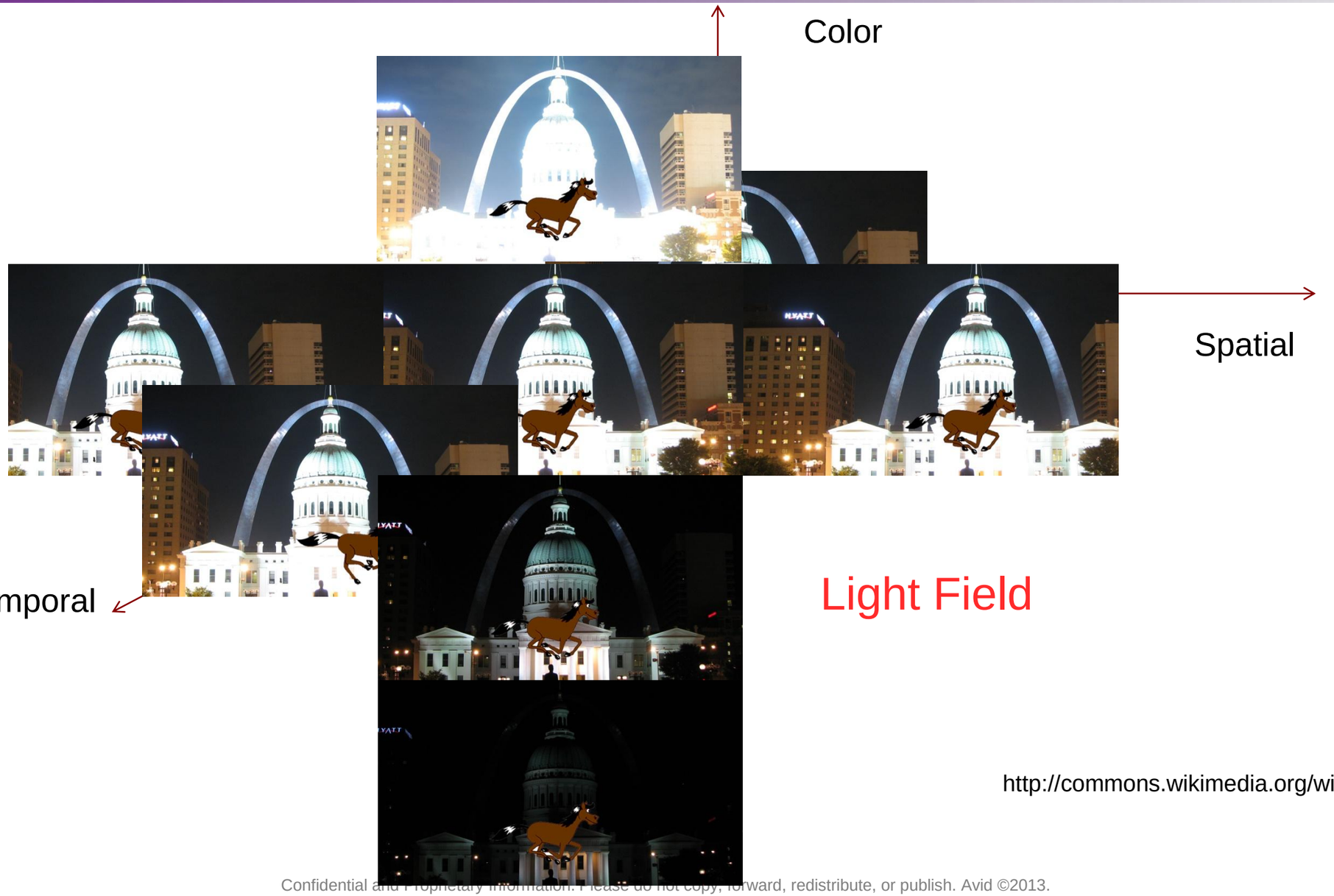
- 7-D Plenoptic Function

$$I = P(\Theta v, \Phi v, Ex, Ey, Ez, t, \lambda)$$

- Camera streams are spatio-temporal-colorimetric view-samples of

  "Radiance received along any direction V arriving at any point E in space, at any time 't' and over any range of wavelength"

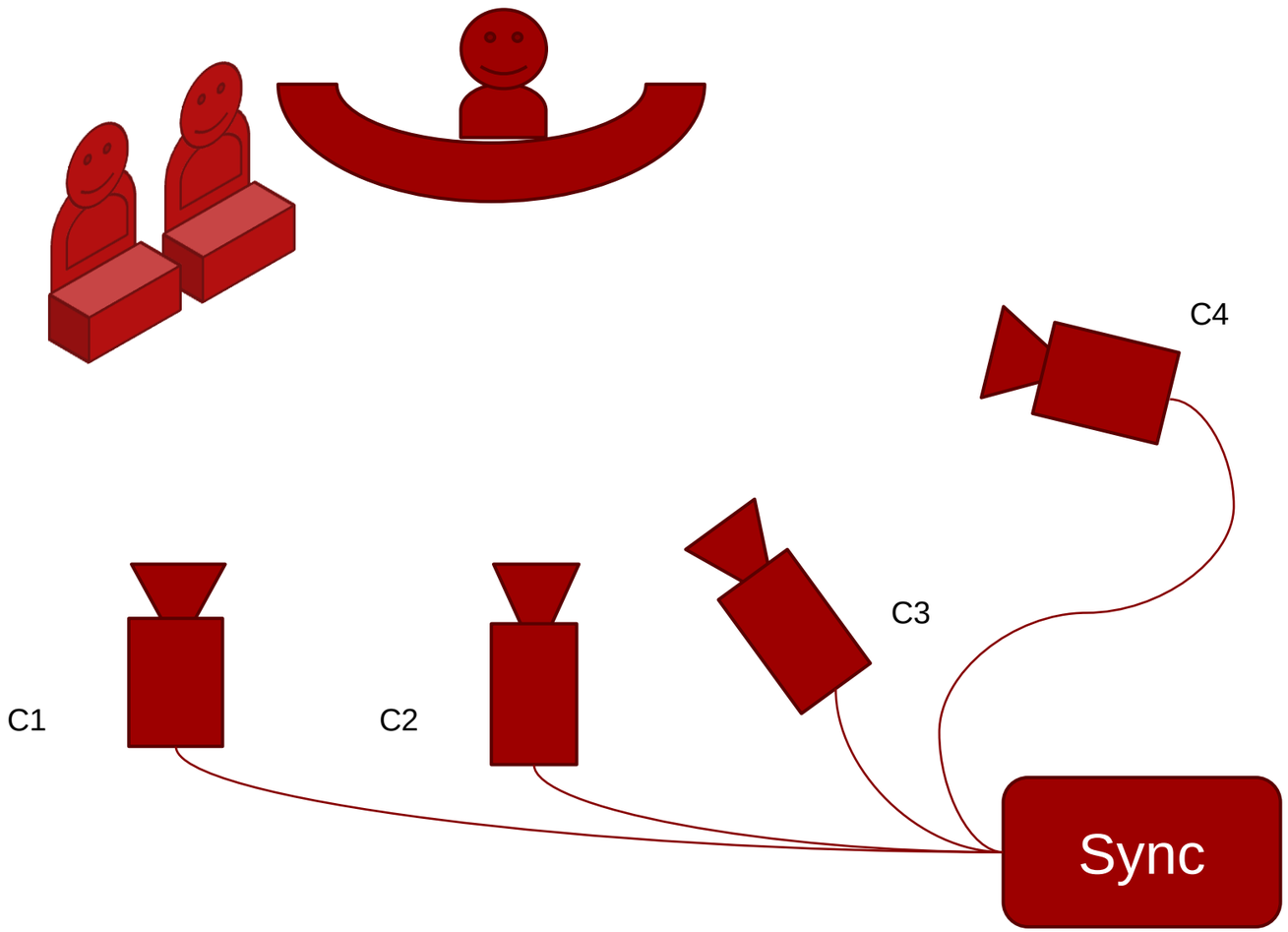
Multi-view sampling of a light field



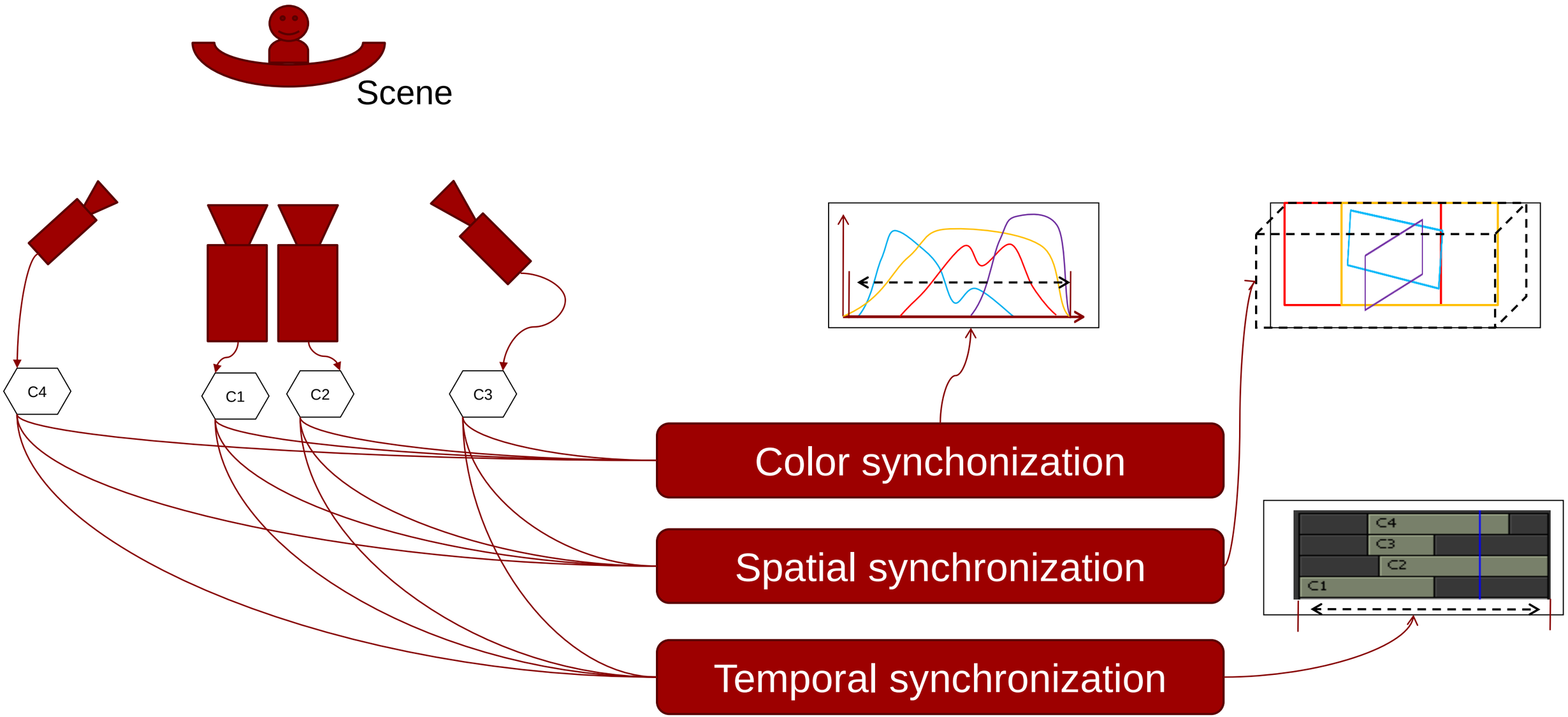
http://commons.wikimedia.org/wiki/Main_Page

Multi-cam productions

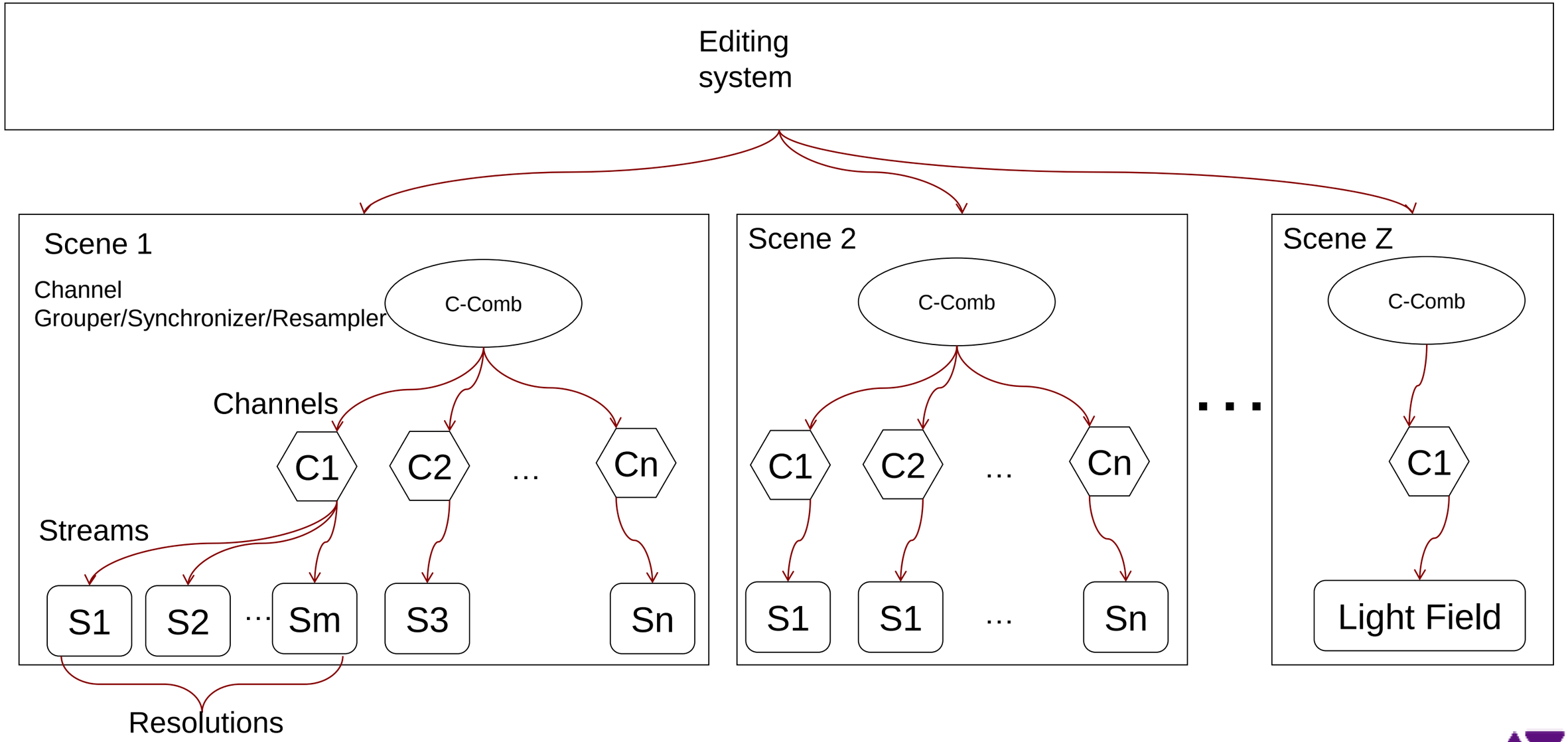
Temporally aligned multi-cam group



Registered multi-cams

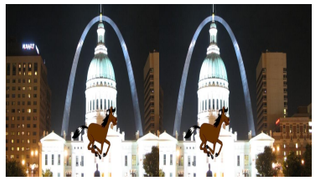


Multi-channel video data-model



Proxy Editing and Multi-Mastering

Transcode to side by side S3D



Stage1

Output context:
SidebySide, L|R,
DNX36,
L|R audio

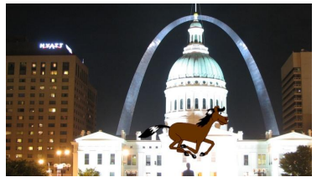
Stage2

Output context:
Full Frame, L|R,
uncompressed,
L|R audio

Stage3

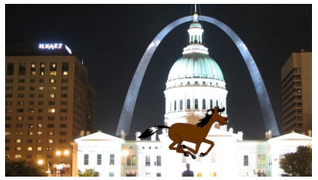
Output context:
Full Frame, Leading Eye,
DNX444,
L|R audio

Full Frame S3D Master output



Channel Combiner

Mono Master output using leading eye



Capture full frame uncompressed S3D left and right views

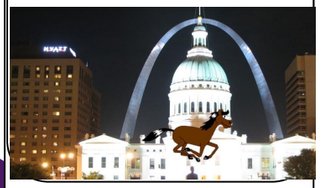
C1:
Left Eye
(Leading)

C2:
Right Eye

C3:
Left Audio

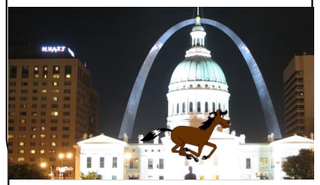
C4:
Right Audio

S4 L DNX444

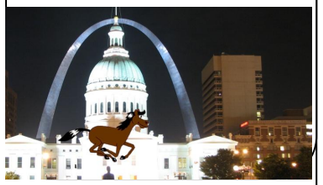


Resolution 3

S1 L 1:1

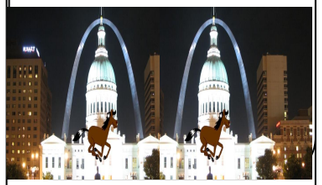


S2 R 1:1



Resolution 1

S3 R|L 1:1



Resolution 2

S4 with
Left audio

S5 with
Right Audio

Data Model – Problem solved.

- Method of grouping multiple views
- Spatial, temporal and Color synchronization of views
- Method of arranging resolutions under the views
- Resampling of scene based on output context.

Avid Intelligent Compute Architecture

Performance Issues

- Multiple views => more processing
  - codecs,
  - visual effects
  - graphics,
  - visualization
- Higher resolution, frame rates and bit-depth

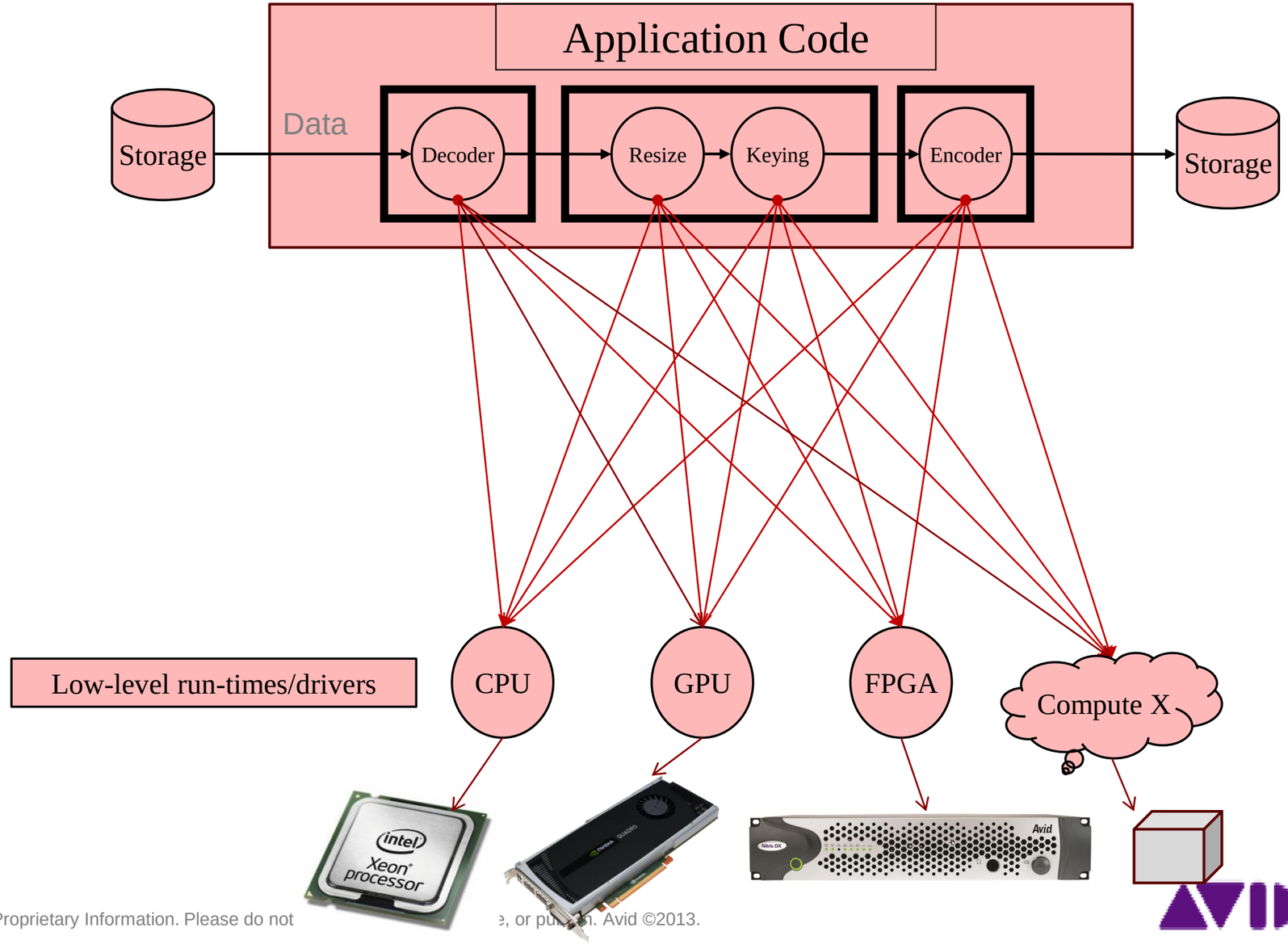
Avid Intelligent Compute Architecture

- Heterogeneous compute Hardware
  - Intel multi-core CPUs (HP Z800/820, Apple Mac Pro)
  - Nitris DX w/ dual DNxHD or AVC-' codec acceleration modules
  - NVIDIA GPU
  - NVIDIA 3DVision support

- Intelligent Media Player

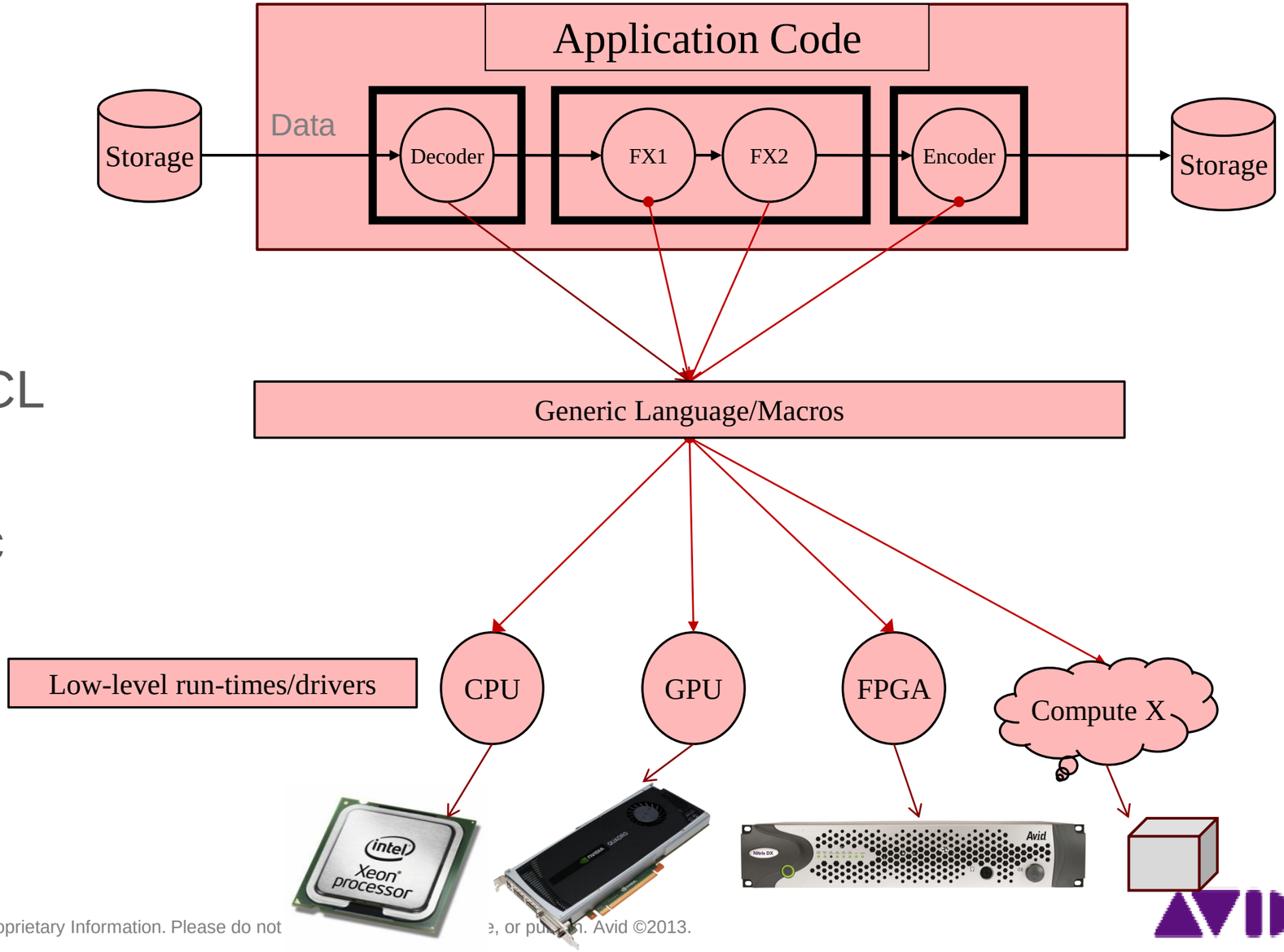
Problem with heterogeneous compute

- Multiple pieces of code for each compute domain
- Balancing across multiple compute domains dynamically
- Optimal transfers between hardware?
- Dynamically adding a new hardware processing domains



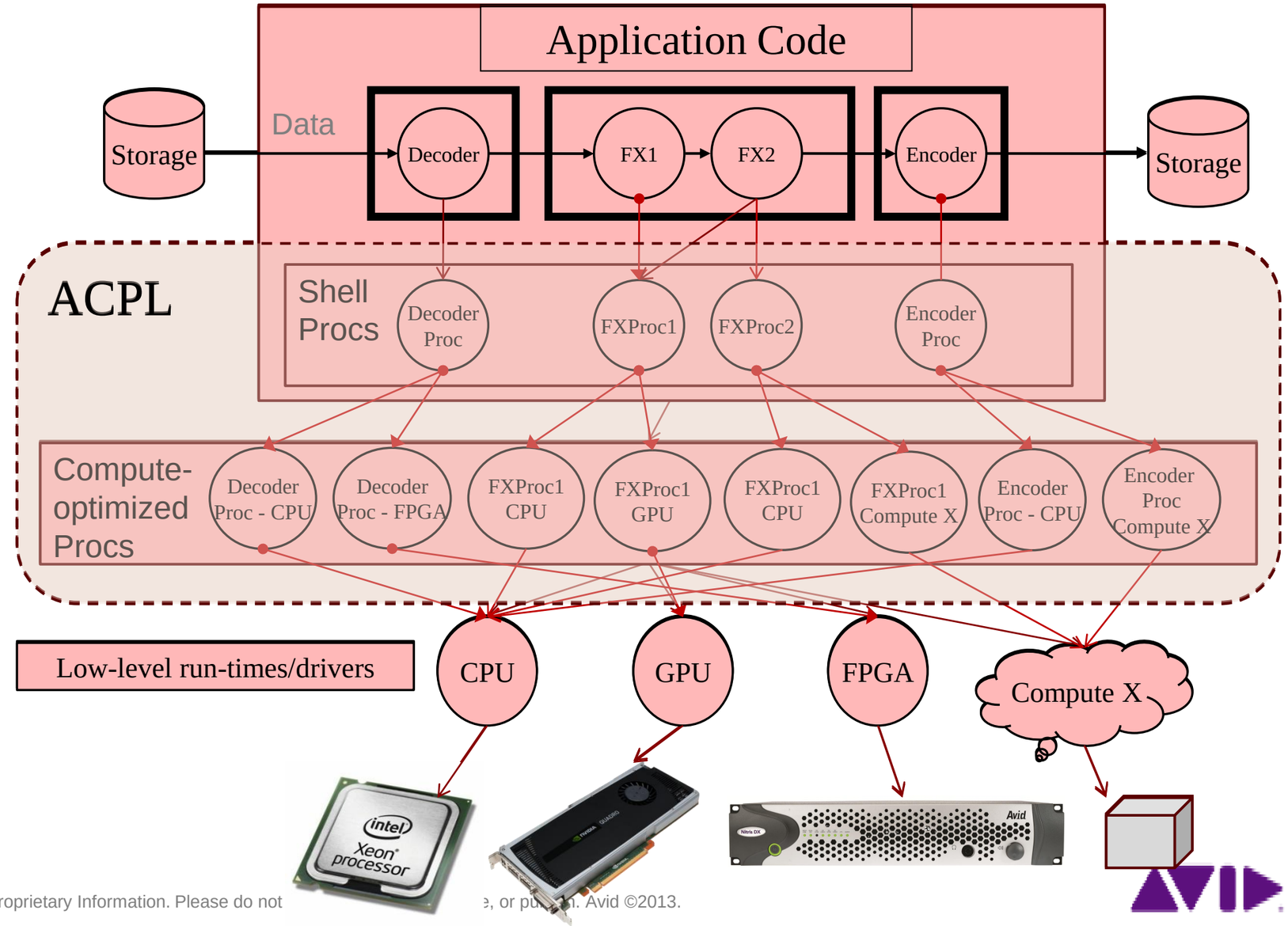
How others have solved it

- Main philosophy - "Write once, support many"
  - Generic Language with HW specific compilation, e.g. OCL
  - Code generation for specific compute models, e.g. Cg



How we solved it

- Philosophy - "Optimize Many, Host once"
  - write compute domain specific optimized code
  - single API contract with client application.
  - Domain-abstracted player



Plug’N Play Architecture

- Home grown component model
  - Similar to COM

- Plug and play
  - Compute domains
  - Processors
  - Data types and formats
  - Converters

Plugging-in NVIDIA GPU - Benchmark results

- System
  - Avid Symphony 6.0
  - HP Z800 Workstation
  - Windows 7 64-bit
  - 6 GB RAM
  - NVIDIA Quadro 4000
  - Avid Nitris DX
  - Avid VideoRaid storage

- *Material*
  - AVC-Intra 100 sources rendered to Avid DNxHD 185x
  - 1080i/50 10-bit
  - Processing:
  - Full resolution HD
  - Full-Frame processing of Stereo 3D streams 16 bit processing precision
  - Render quality: best (advanced polyphase)

Task	How much faster with NVIDIA GPU?
4 streams with title	2.6x
4 color corrections, 3 picture-in-picture with soft border, 1 title	
4 streams with title	7.3x
4 color corrections, 3 perspective (3D Warp) with soft border, 1 title	
2 advanced effect streams	1.8x
1 color correction + stabiliser + spectramatte key + resize + Symphony advanced color correction	
1 FullFrame stereo effect stream	1.4x
color alignment + positional alignment + S3D rotational alignment + vergence + zoom + floating window	
2 FullFrame stereo effect streams	1.9 x
2 color alignments + 2 positional alignments + 2 S3D rotational alignments + Depth transition + Dissolve	

Example

AVC-I Decode + color alignments + 2 positional alignments + 2 S3D rotational alignments + Depth transition + Dissolve + DN220x encode

- Stereo Full Frame AVC-I decode
- color alignments per stereo pair
- positional alignments per stereo pair
- S3D rotational alignments per stereo Pair
- Depth transition
- Dissolve
- DNx220x 10bit Encode