DIY Multiview Camera System: Panoptic Studio Teardown

Capture Software, Storage, Calibration, and Multiple Kinect Subsystem

Hanbyul Joo and Tomas Simon

Robotics Institute
Carnegie Mellon University

The Panoptic Studio

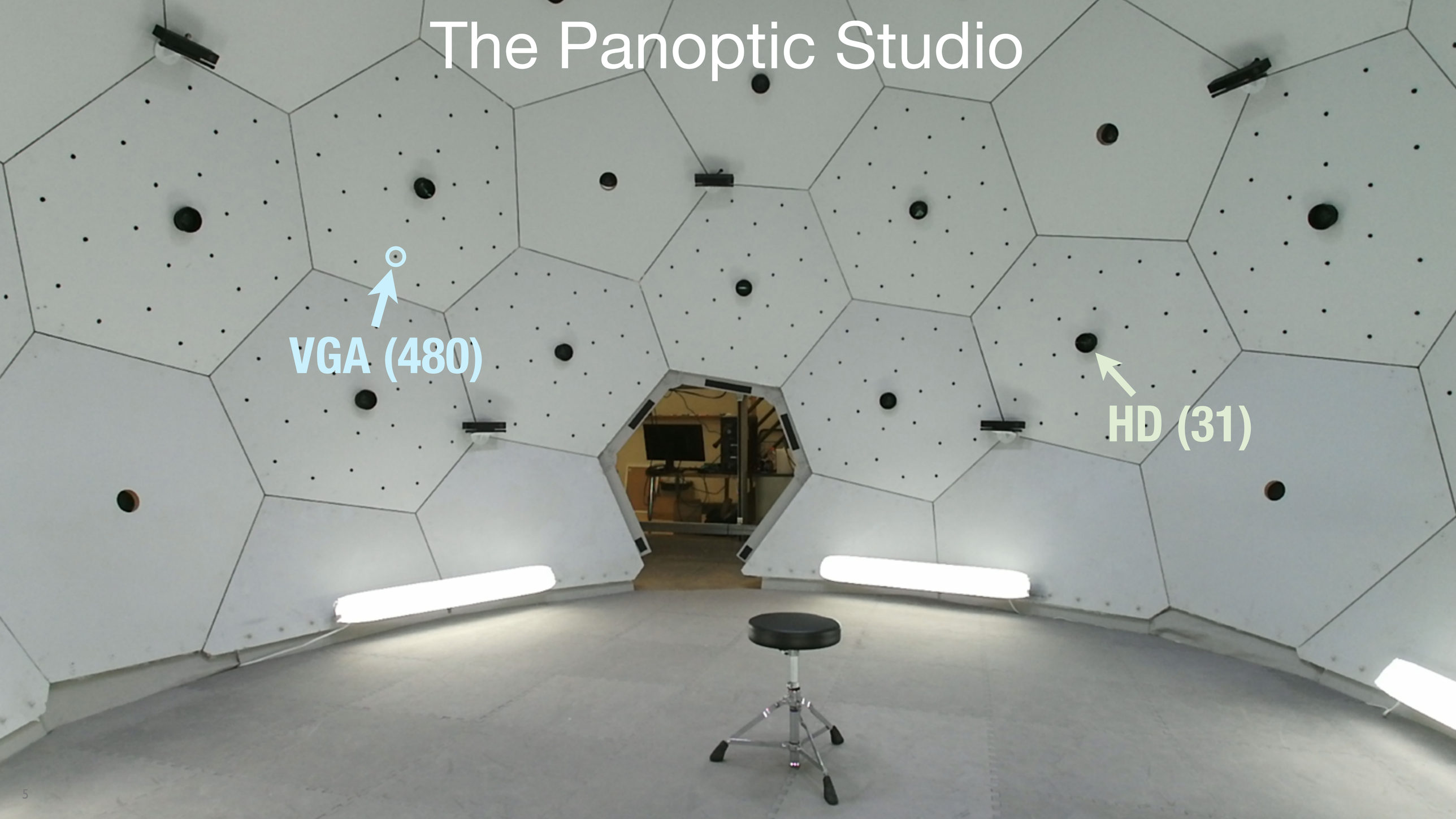
Modularized Design with 20 Panels

480 VGA Cameras
31 HD Cameras
10 Kinects

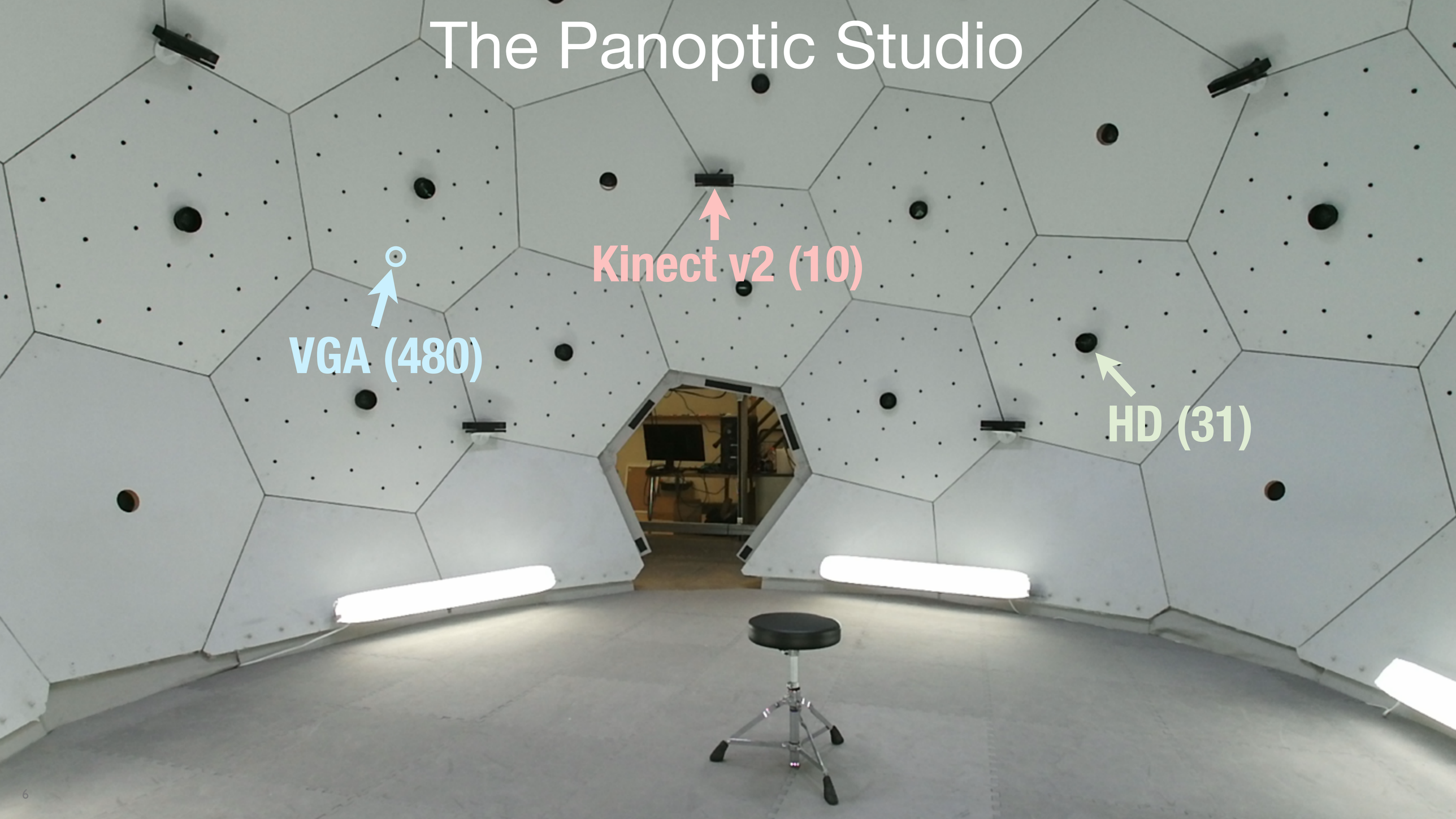
The Panoptic Studio

The Panoptic Studio

The Panoptic Studio

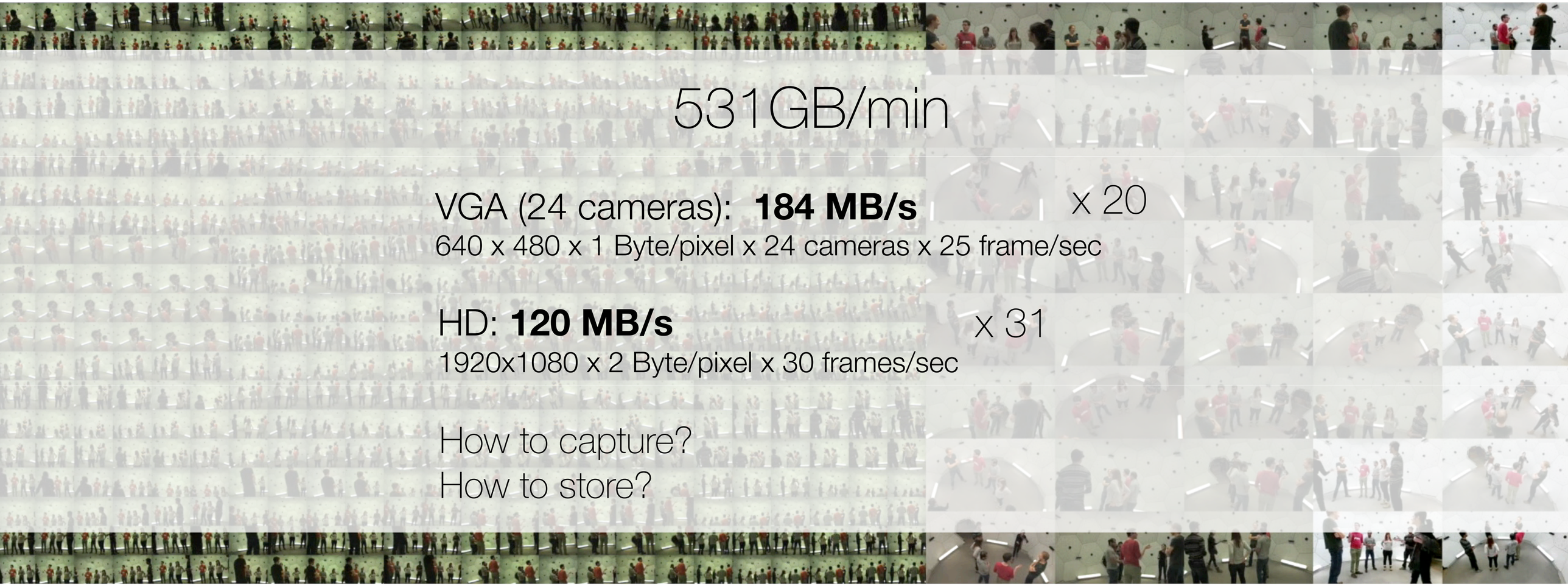


The Panoptic Studio



Synchronized Videos from 521 Views

480 VGAs, 31HDs, and 10 RGB+Ds



531GB/min

VGA (24 cameras): **184 MB/s** x 20

640 x 480 x 1 Byte/pixel x 24 cameras x 25 frame/sec

HD: **120 MB/s** x 31

1920x1080 x 2 Byte/pixel x 30 frames/sec

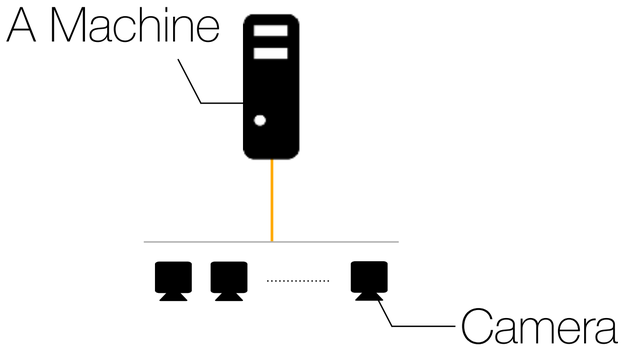
How to capture?

How to store?

Capture Software

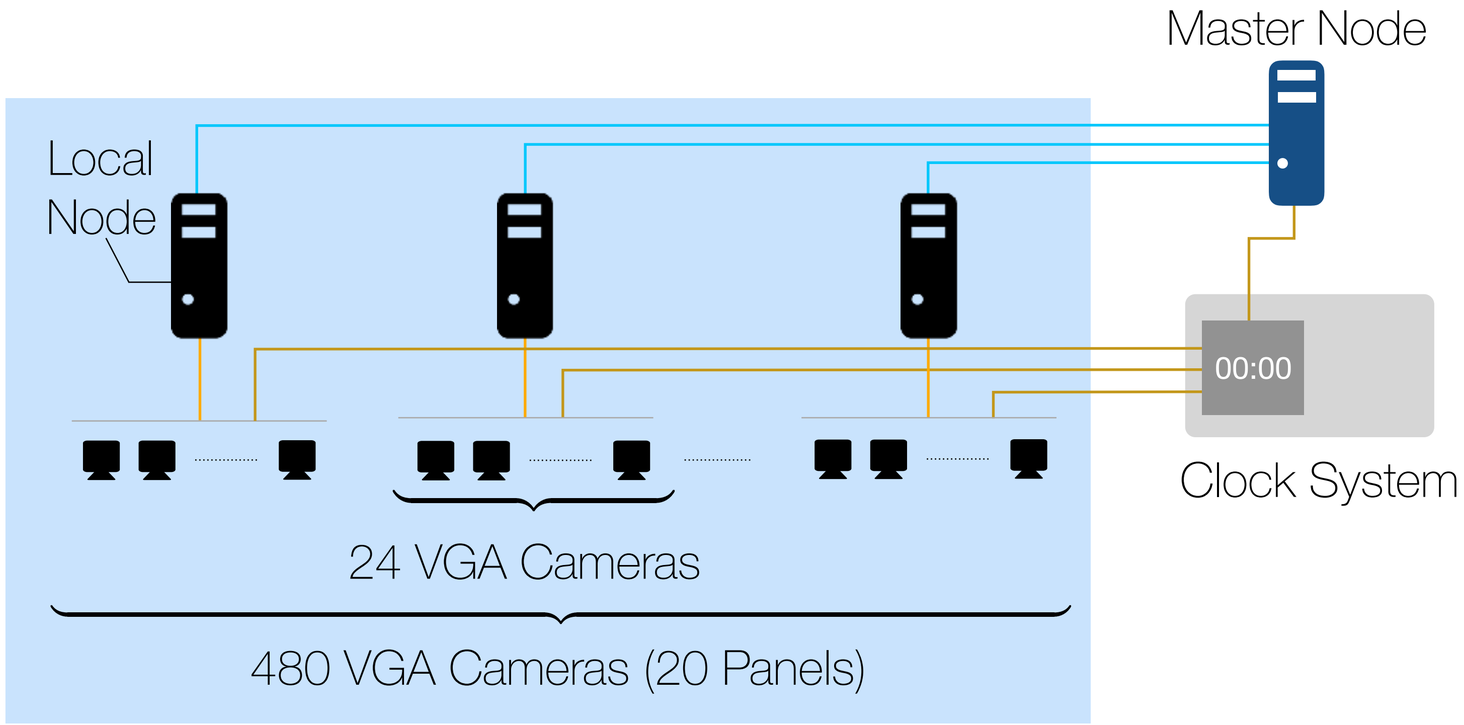
Capture Software

Why Is This Not Straightforward?



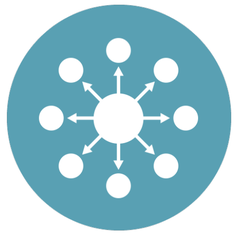
Capture Software

Why Is This Not Straightforward?

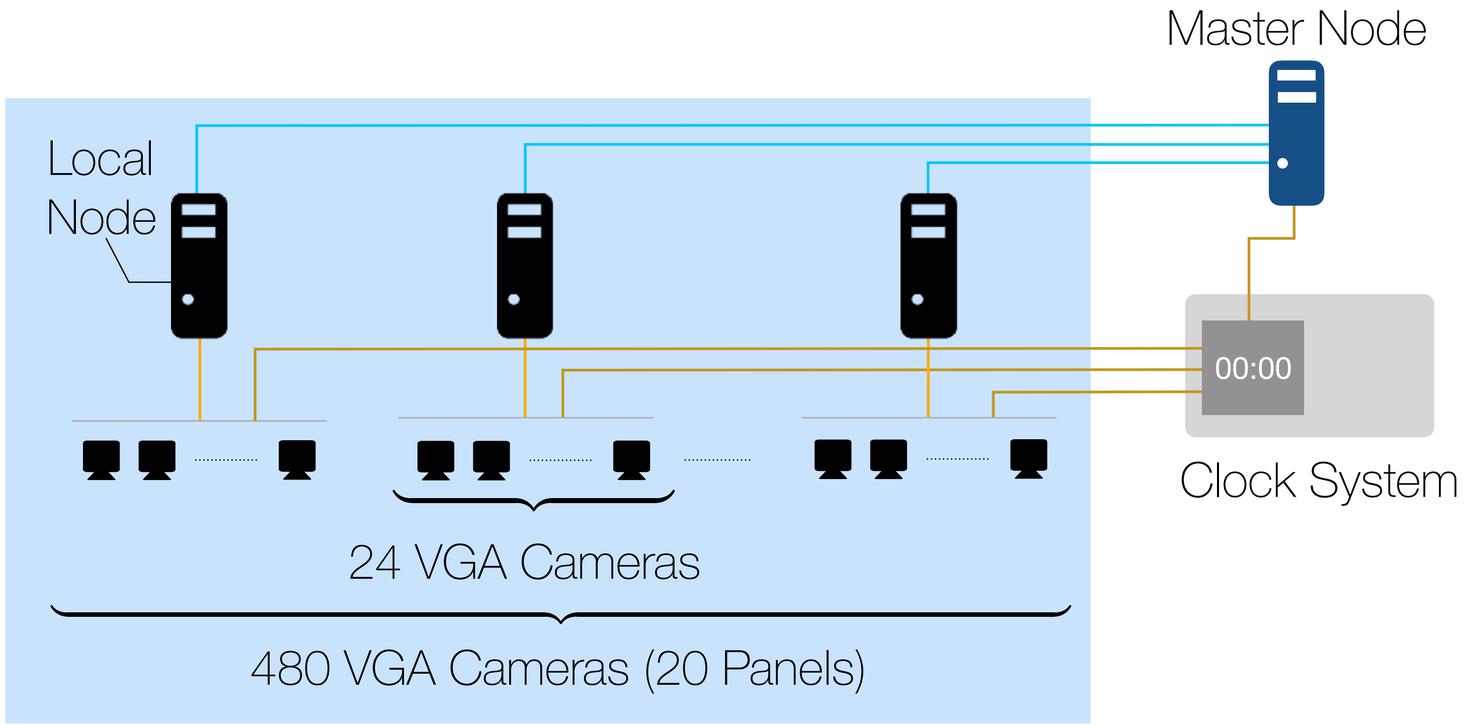


Capture Software

Why Is This Not Straightforward?

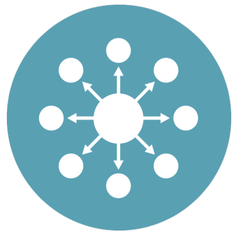


Scalability

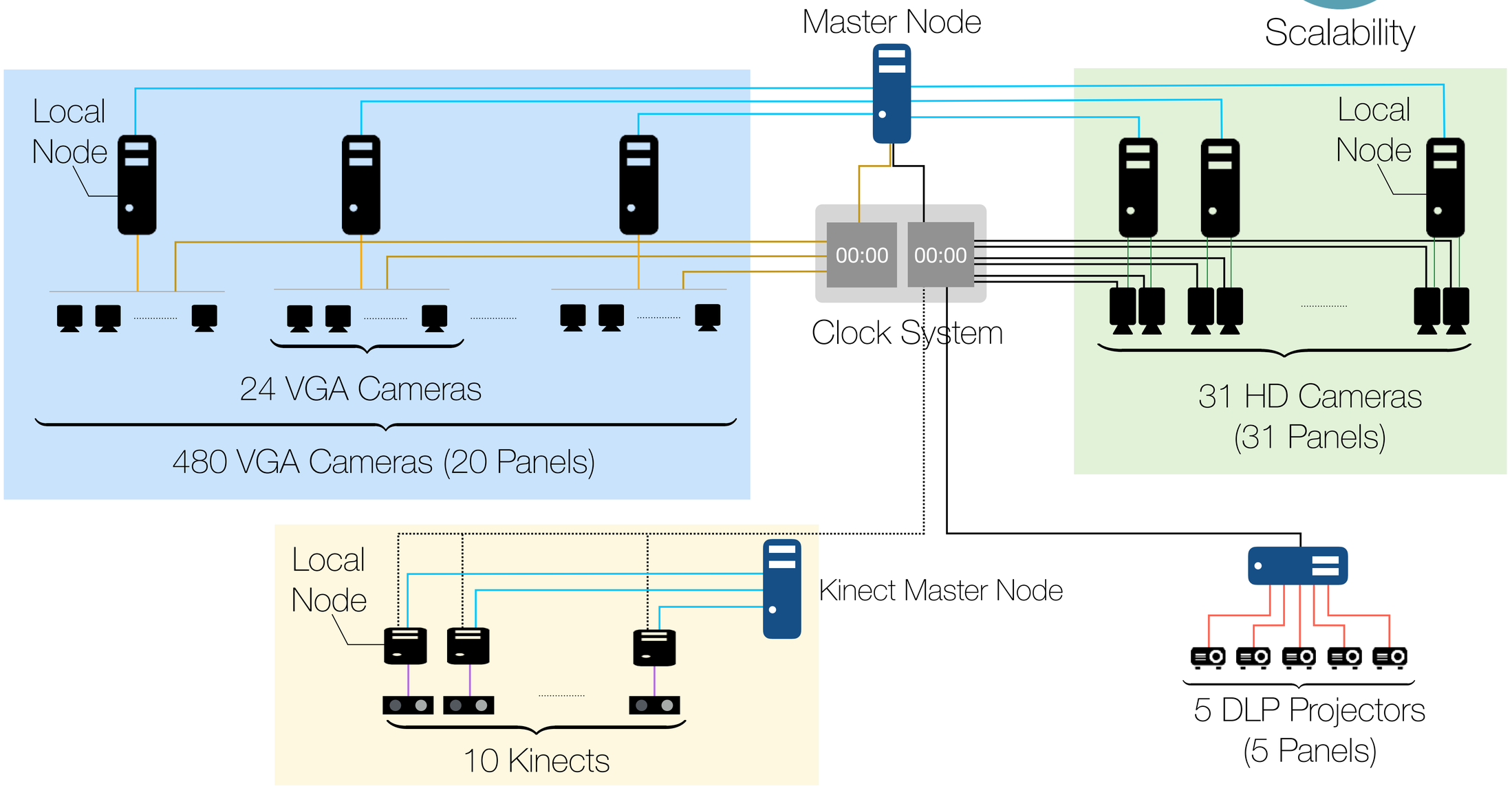


Capture Software

Why Is This Not Straightforward?

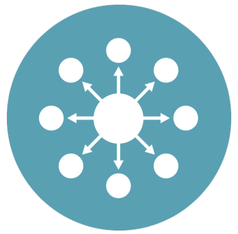


Scalability



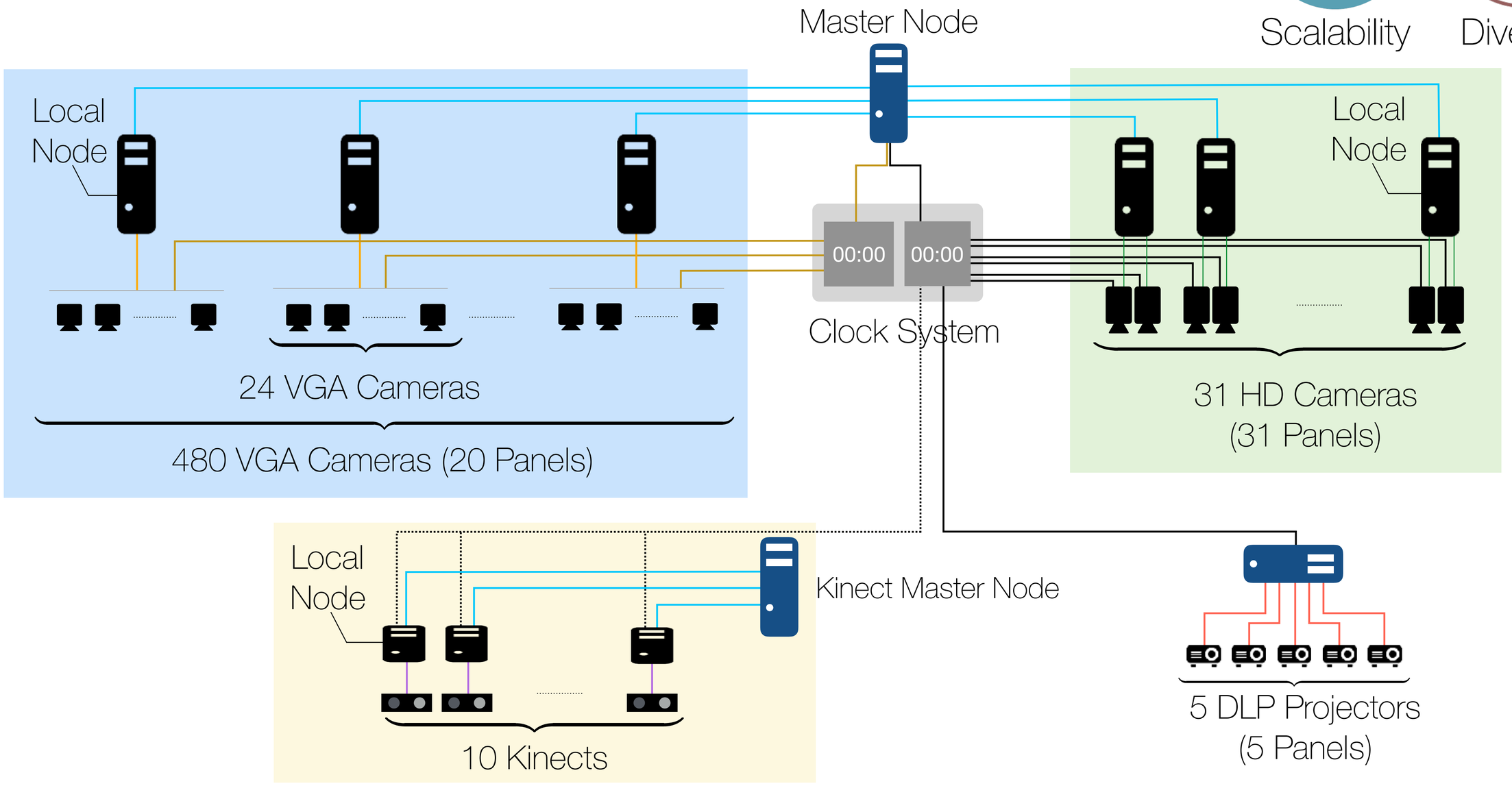
Capture Software

Why Is This Not Straightforward?



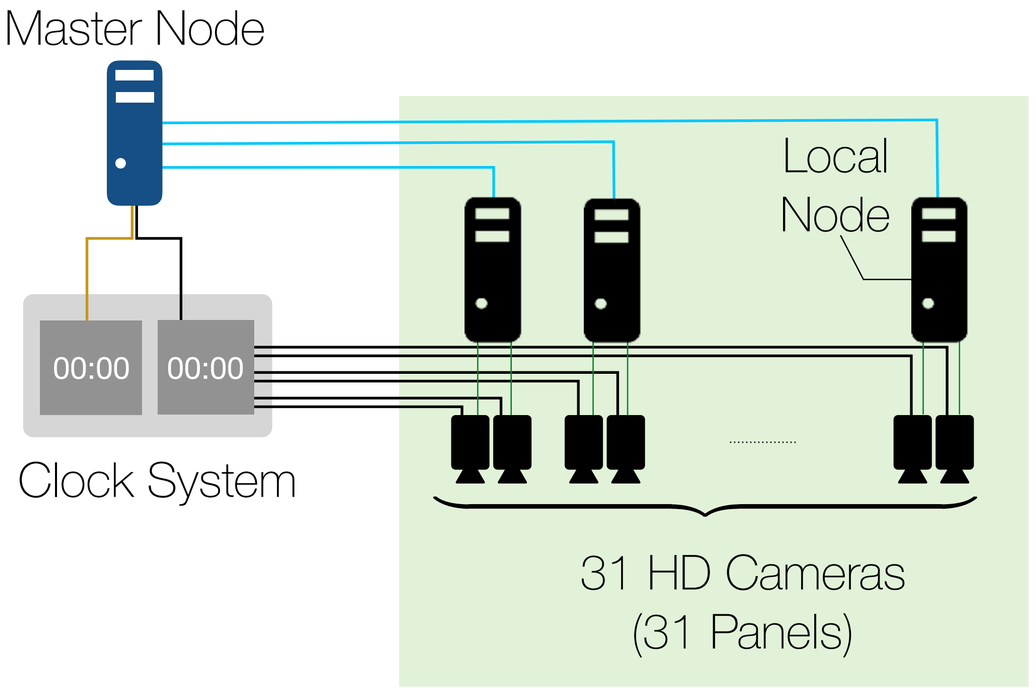
Scalability

Diversity



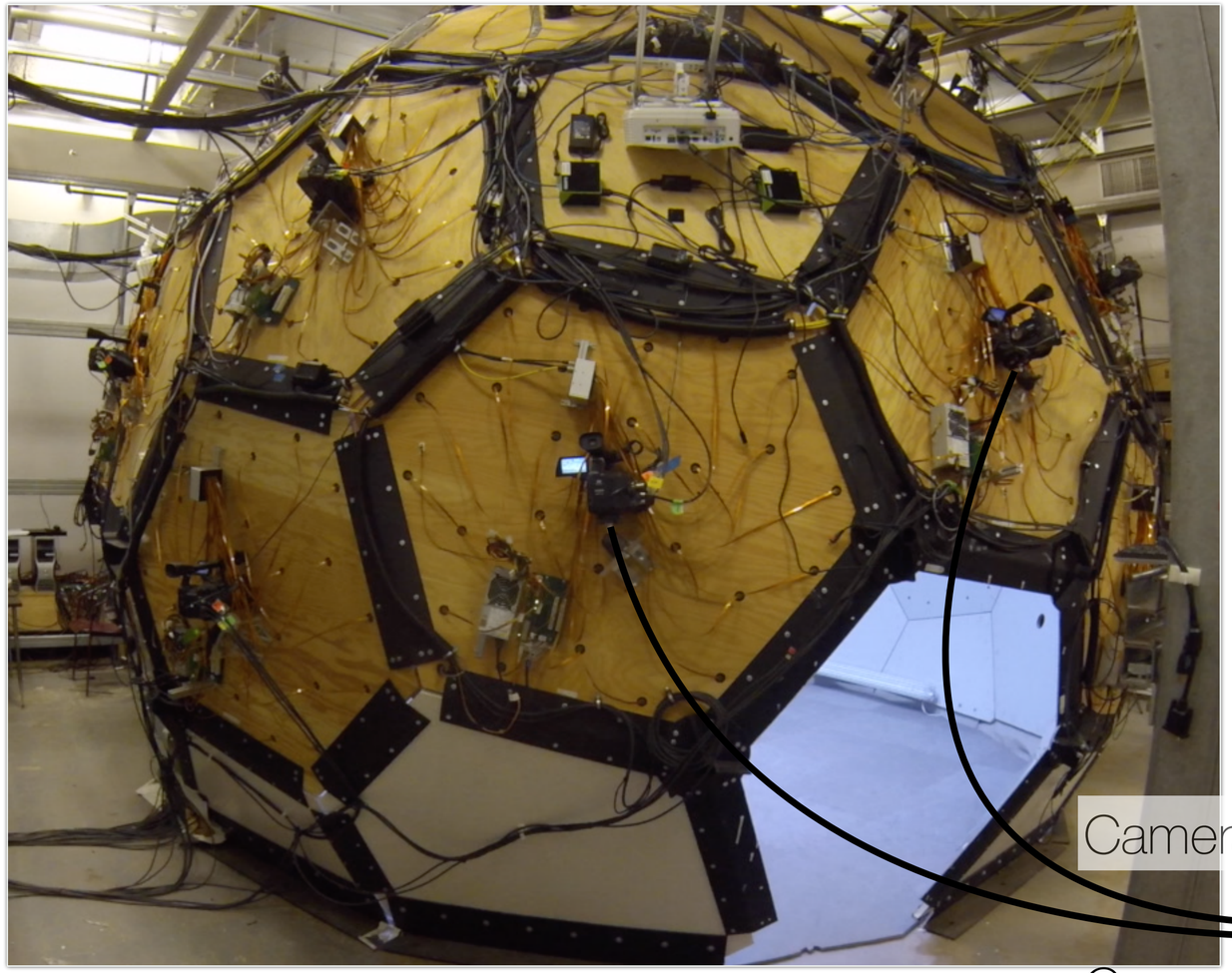
Capture Software

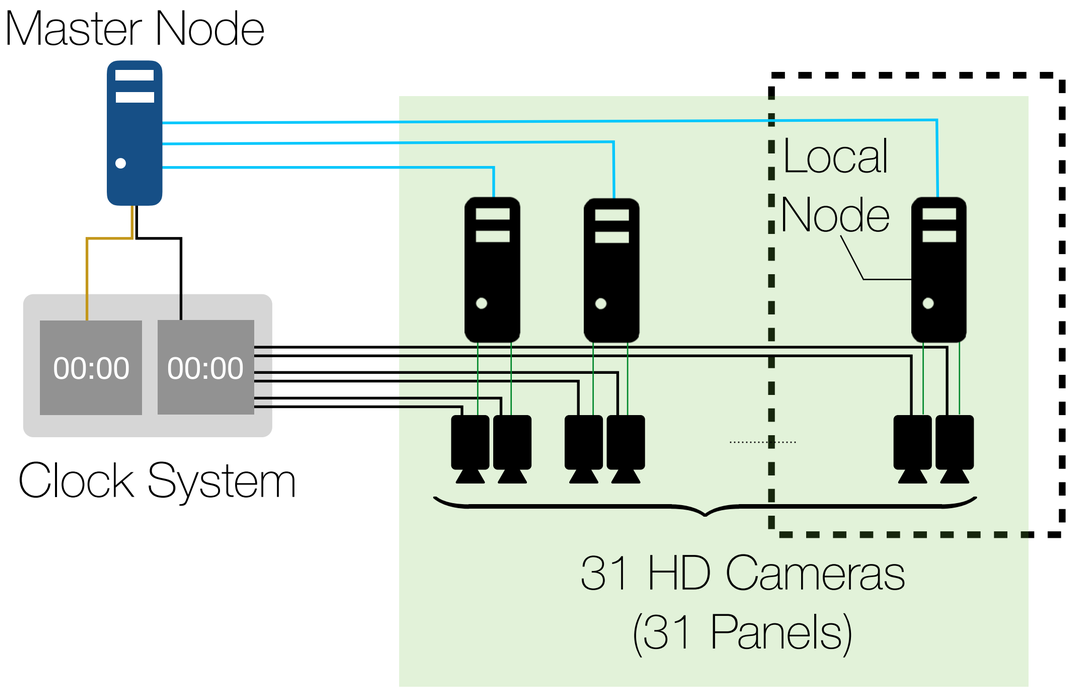
Option 1: Asynchronized Initiation Without Preview



Capture Software

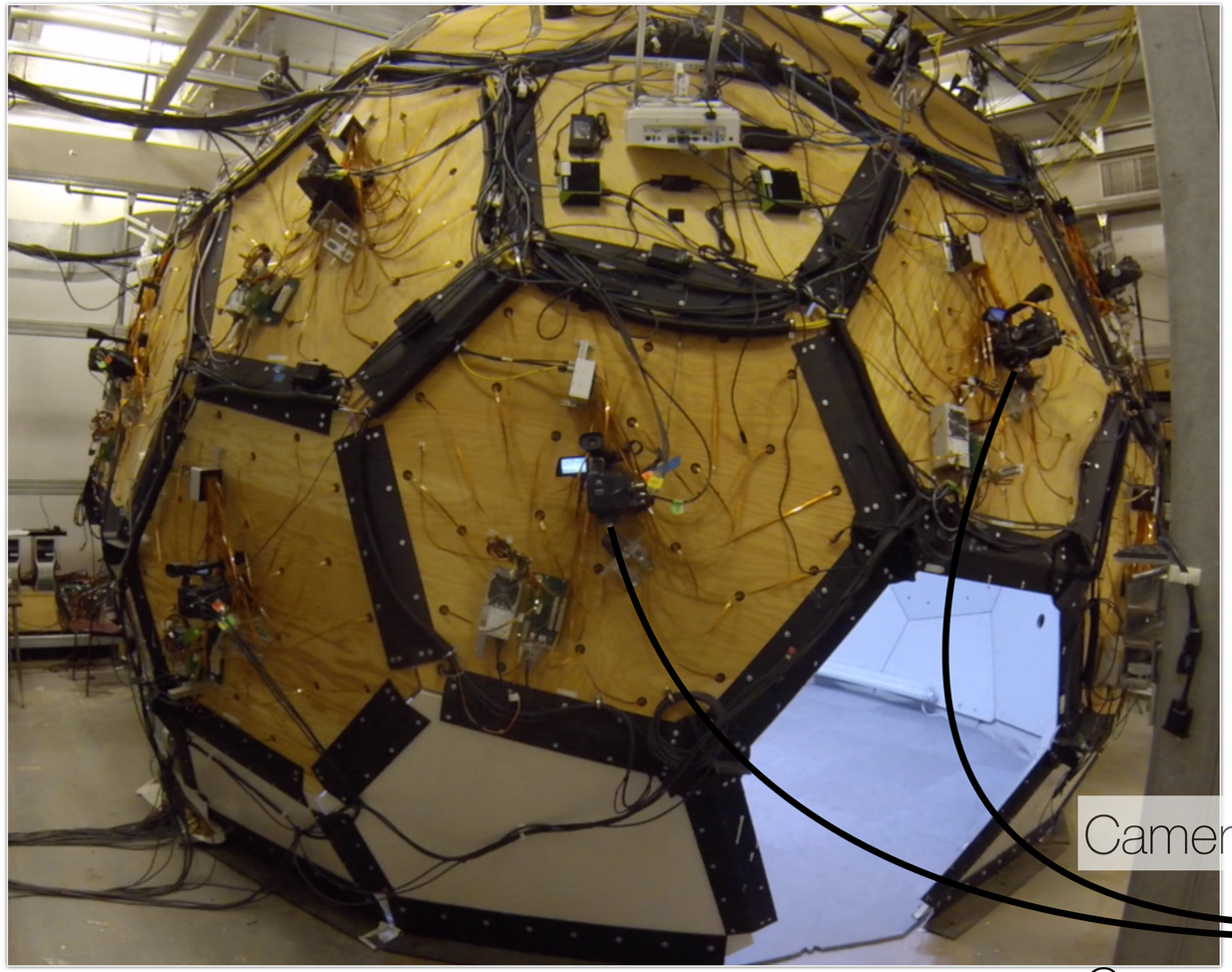
Option 1: Asynchronized Initiation Without Preview

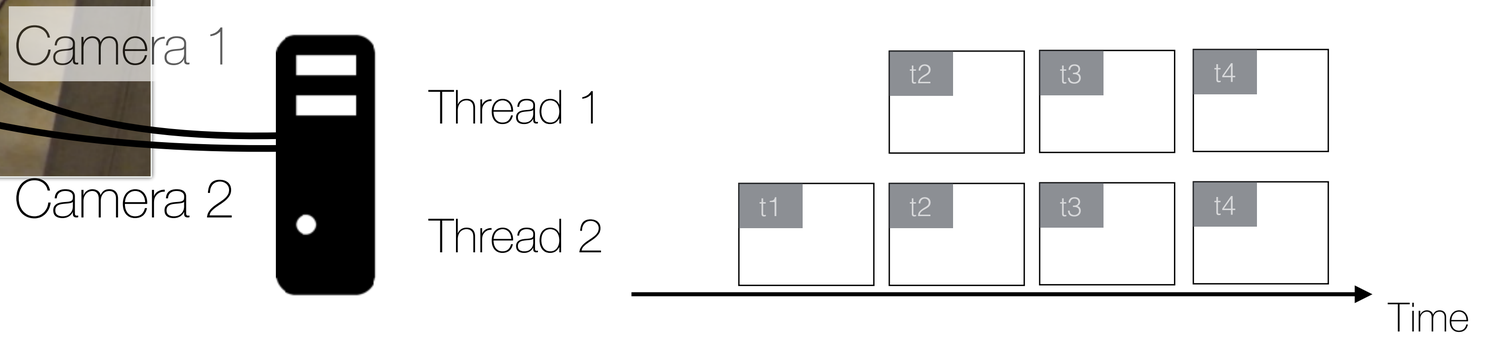




Capture Software

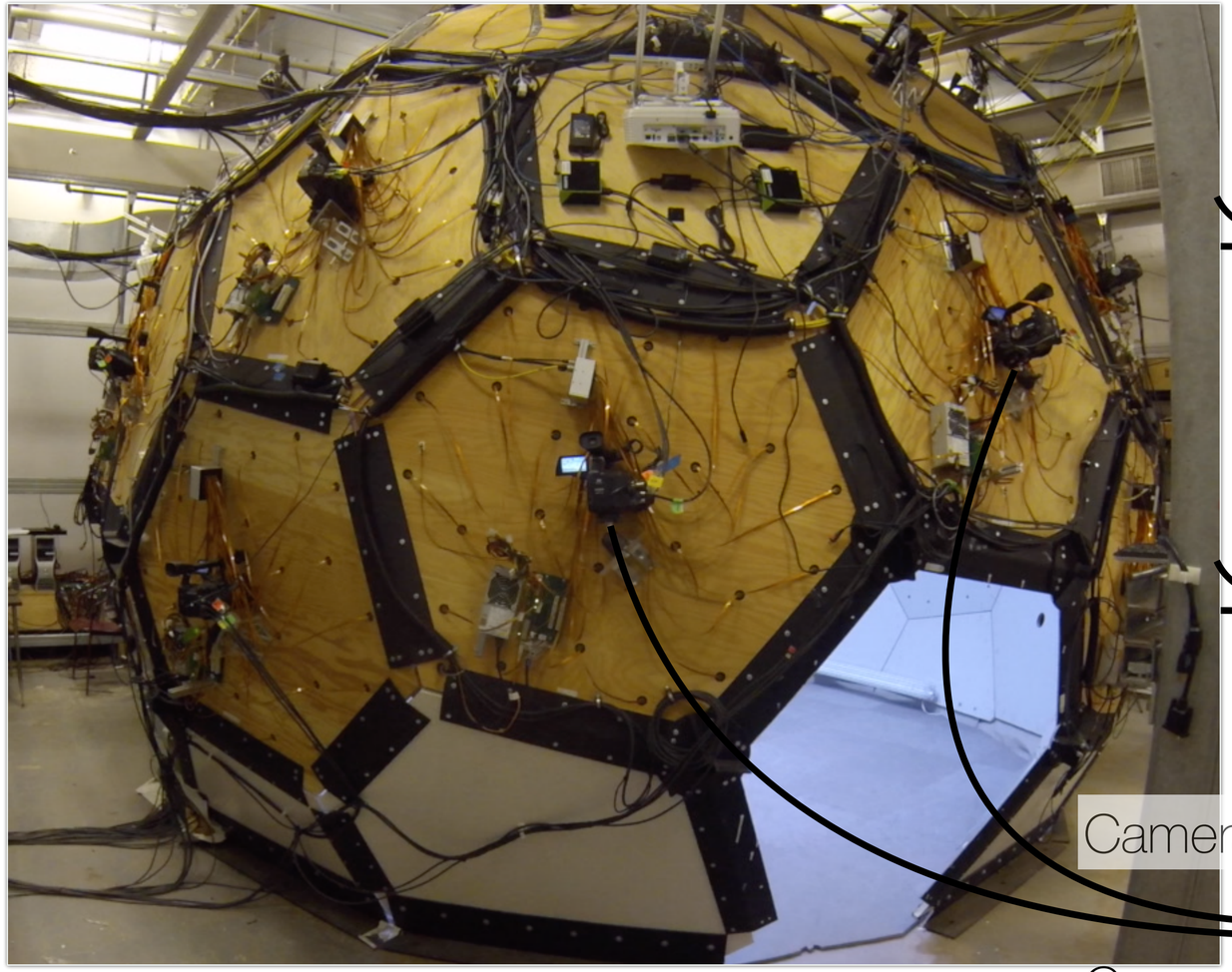
Option 1: Asynchronized Initiation Without Preview





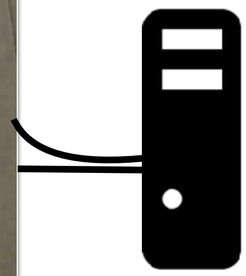
Capture Software

Option 1: Asynchronized Initiation Without Preview



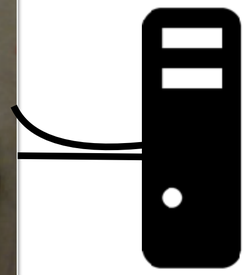
Camera 1

Camera 2



Thread 1

Thread 2

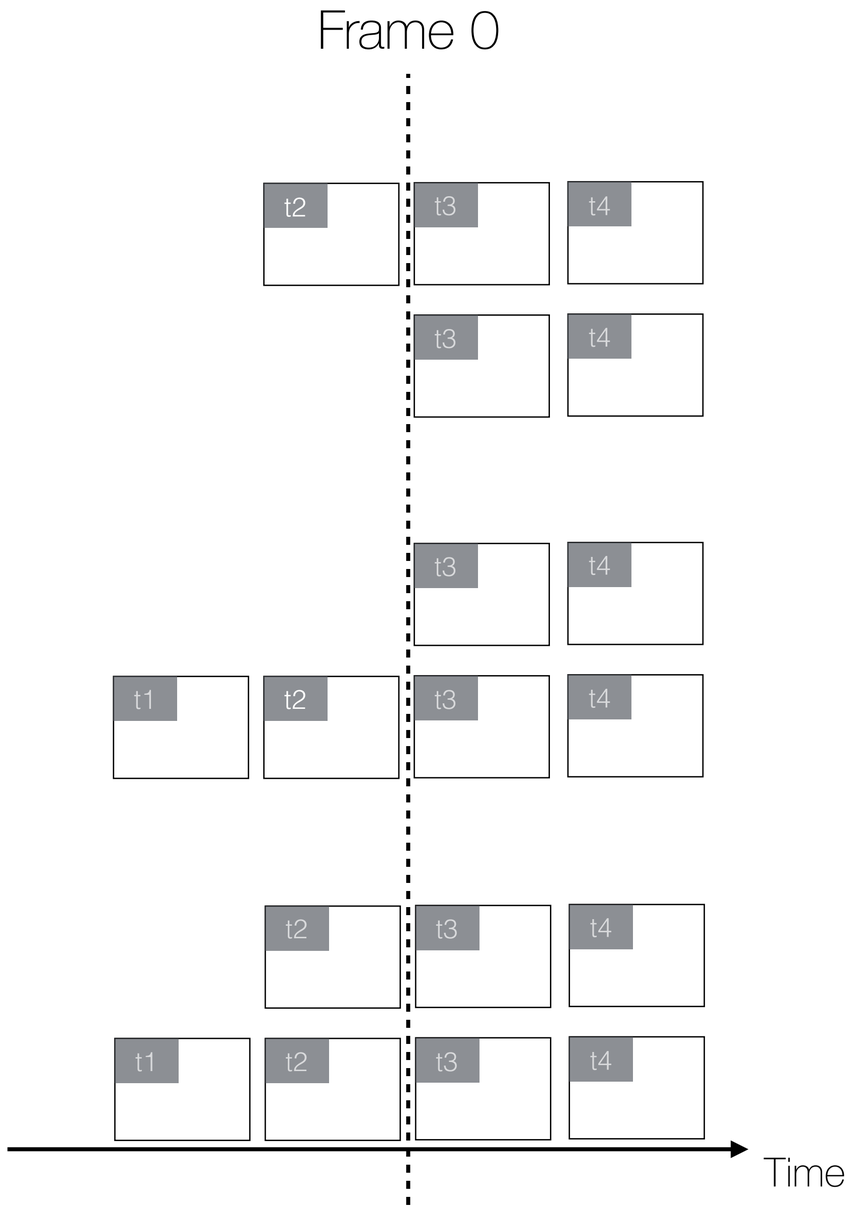


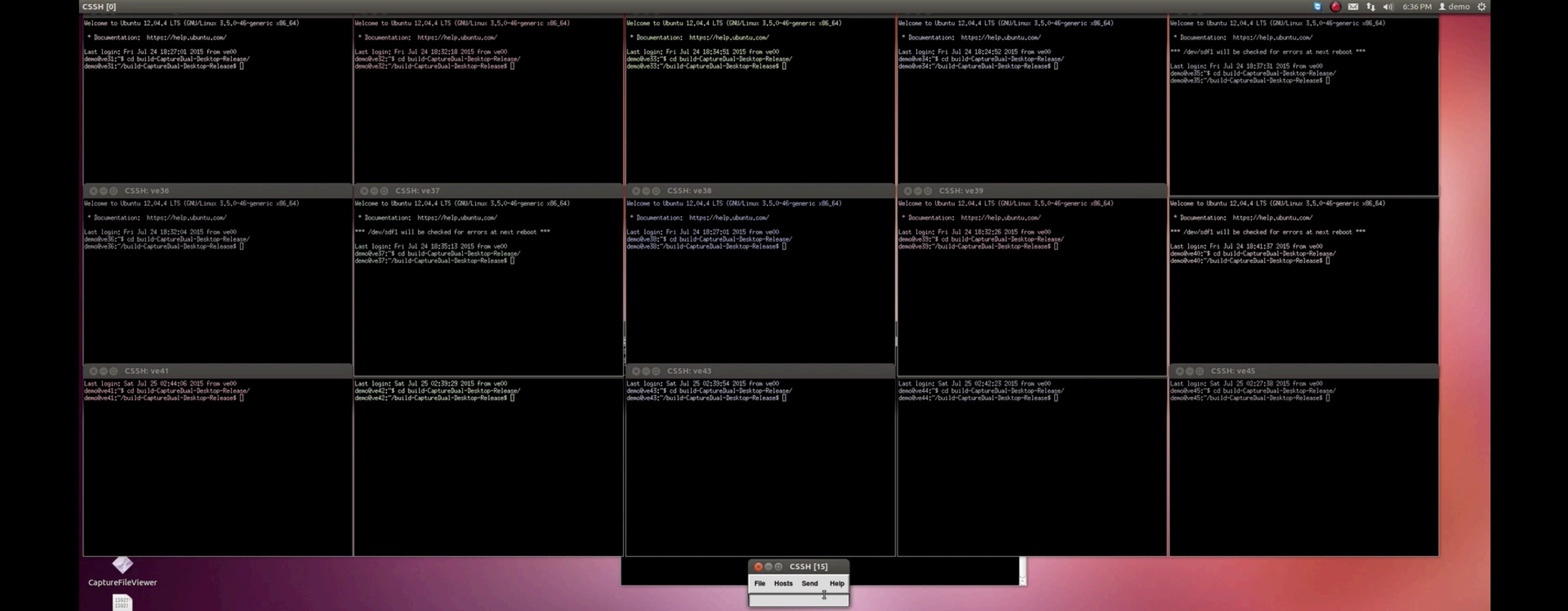
Thread 1

Thread 2

Thread 1

Thread 2

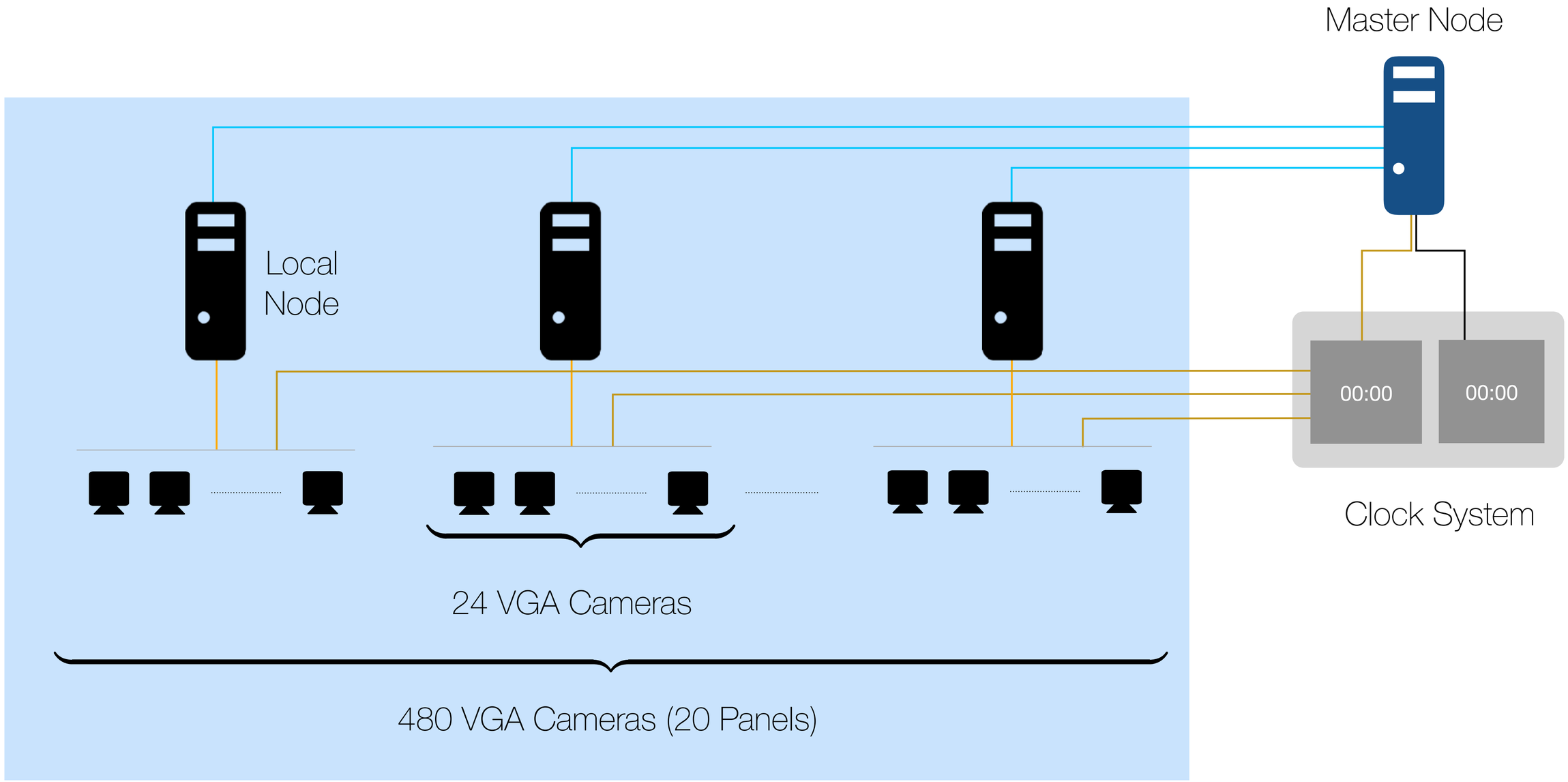




We cannot see the data during the capture

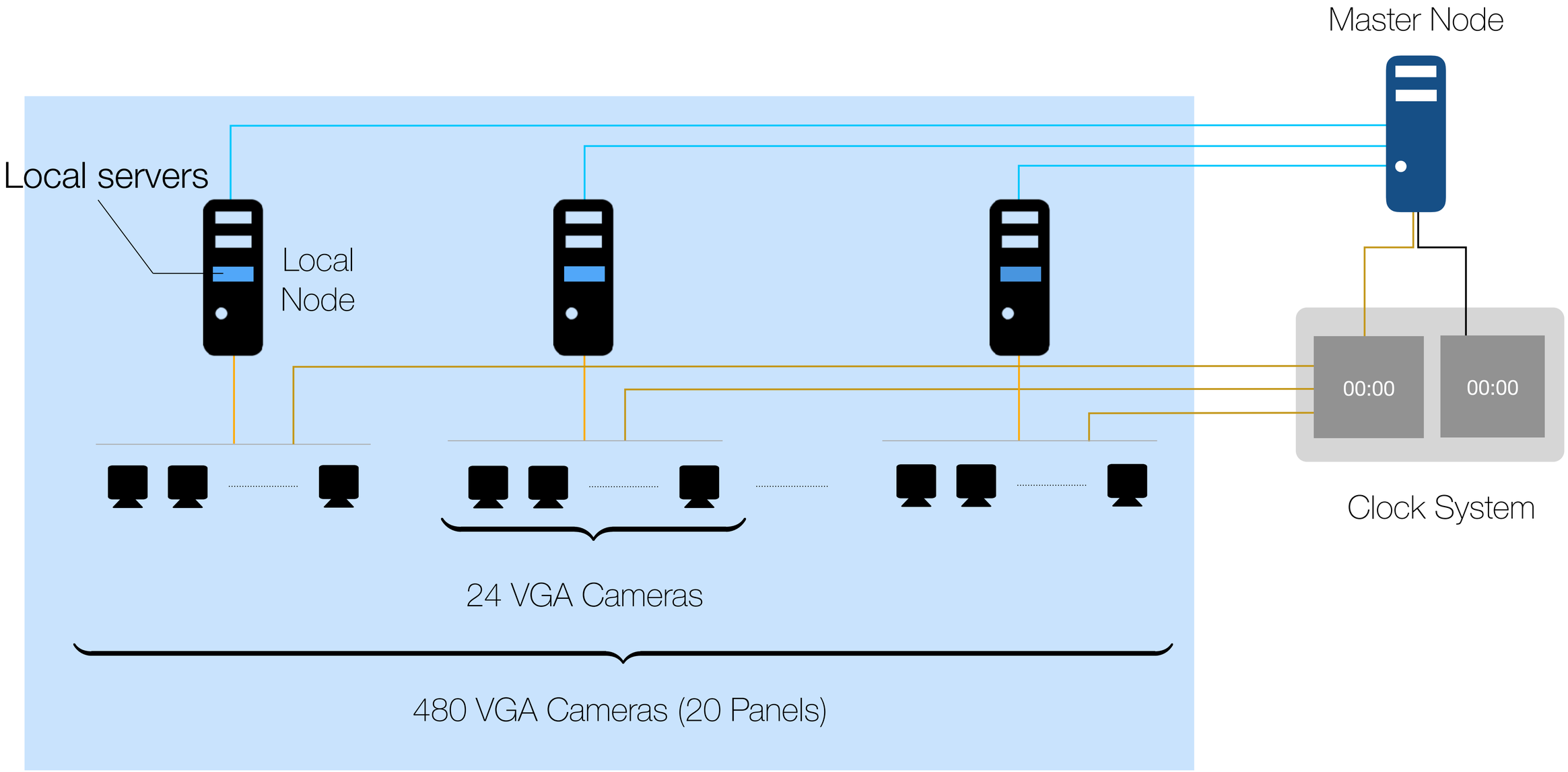
Capture Software

Option 2: Synchronized Initiation With Preview



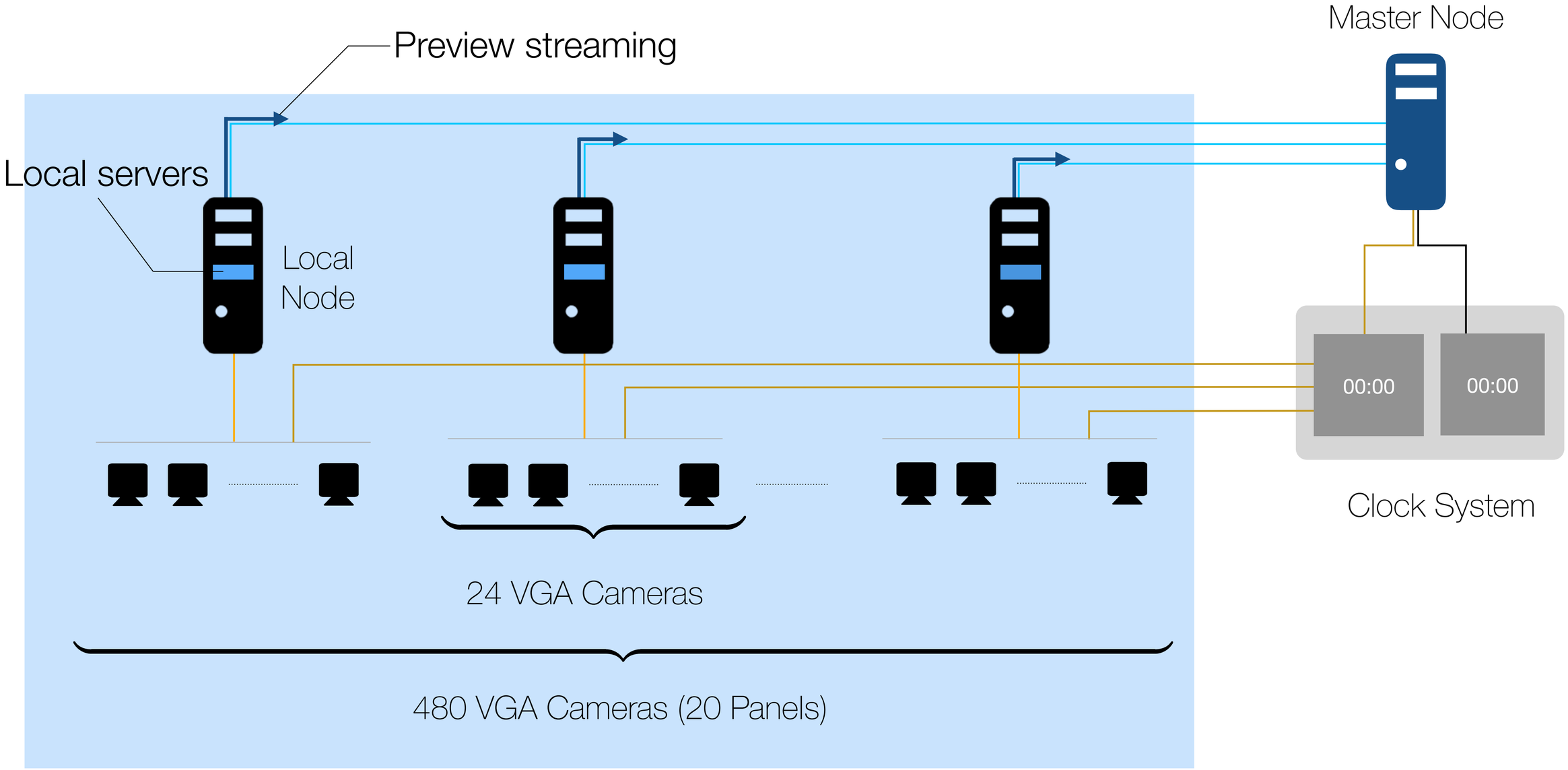
Capture Software

Option 2: Synchronized Initiation With Preview



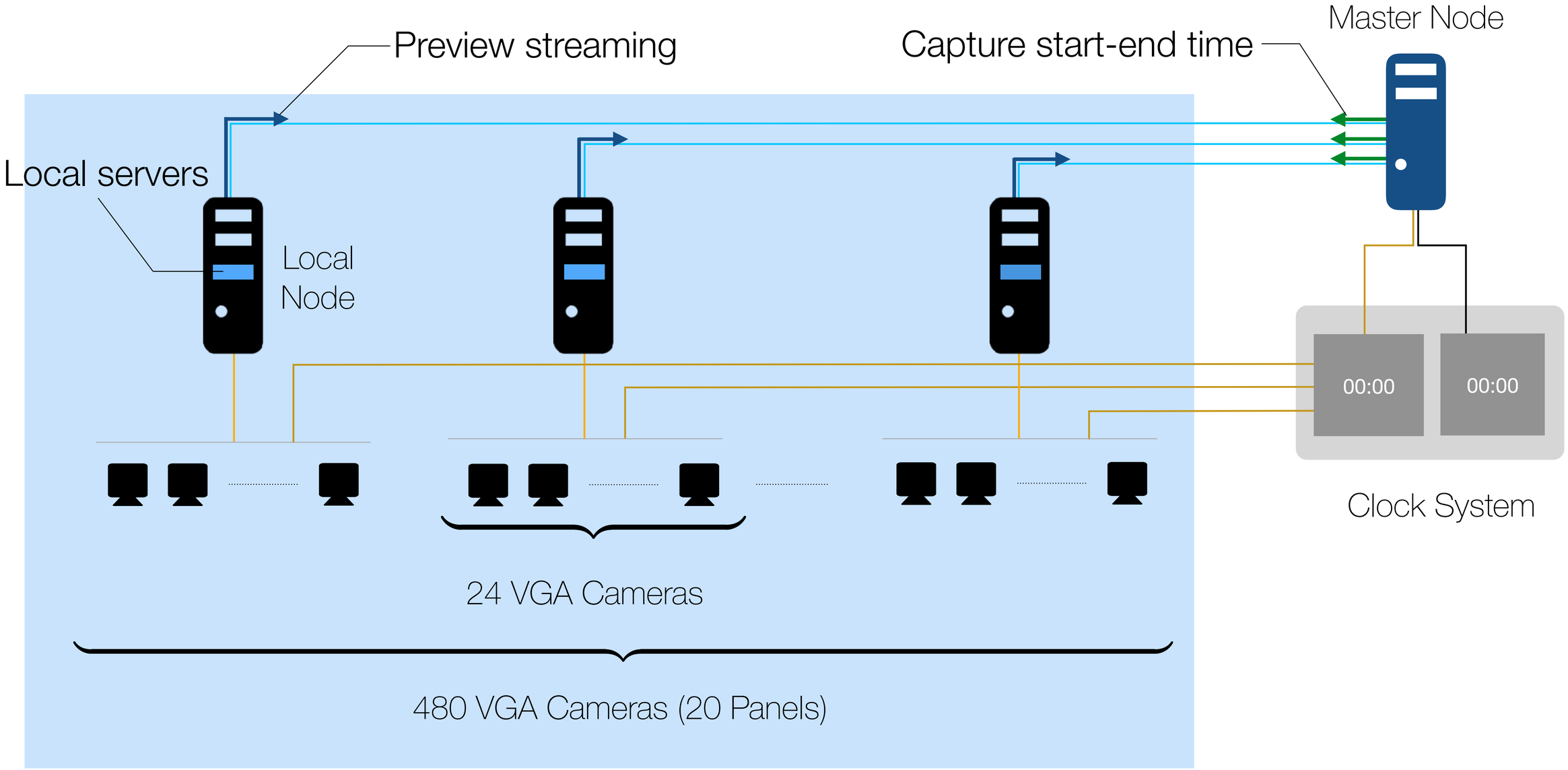
Capture Software

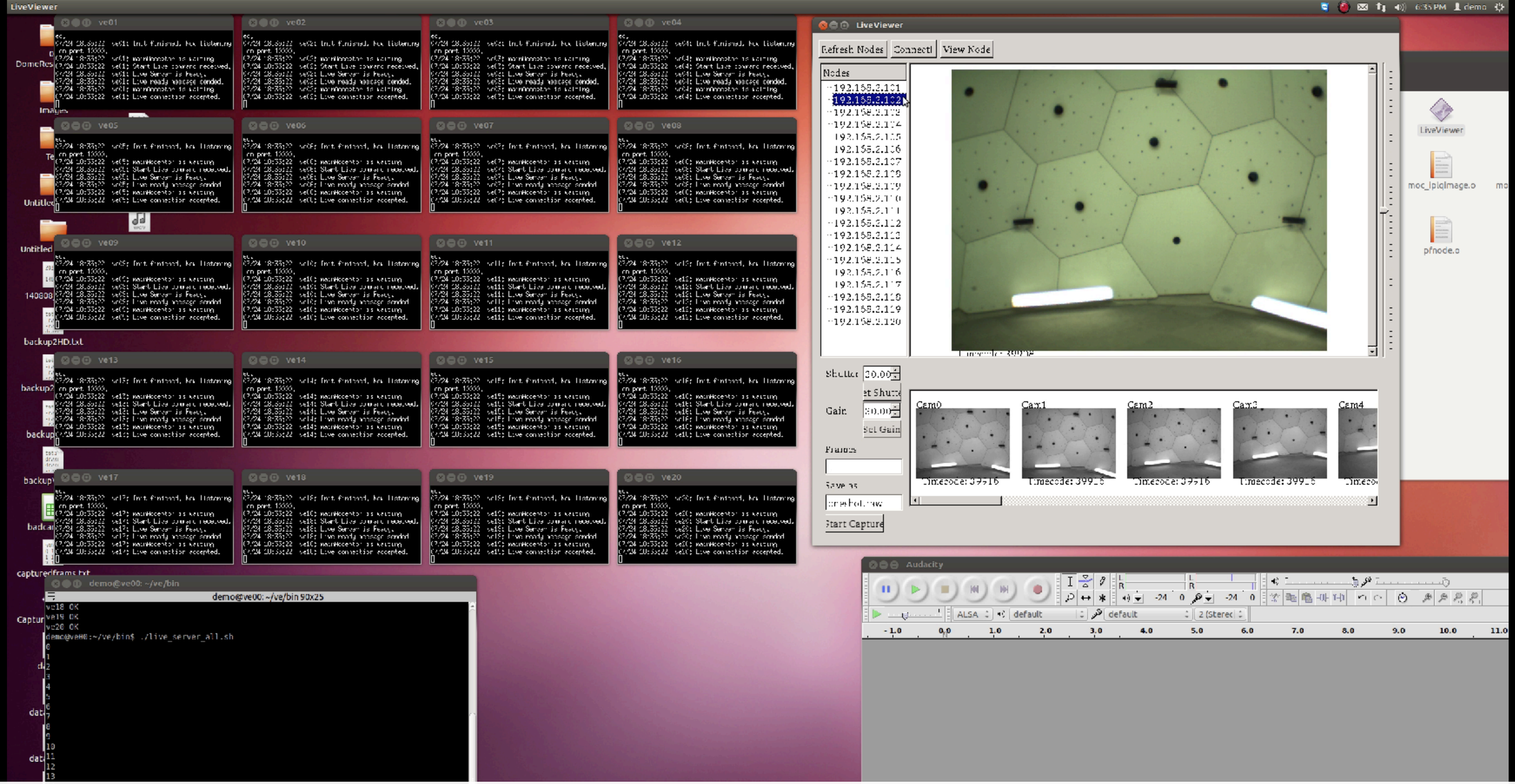
Option 2: Synchronized Initiation With Preview



Capture Software

Option 2: Synchronized Initiation With Preview

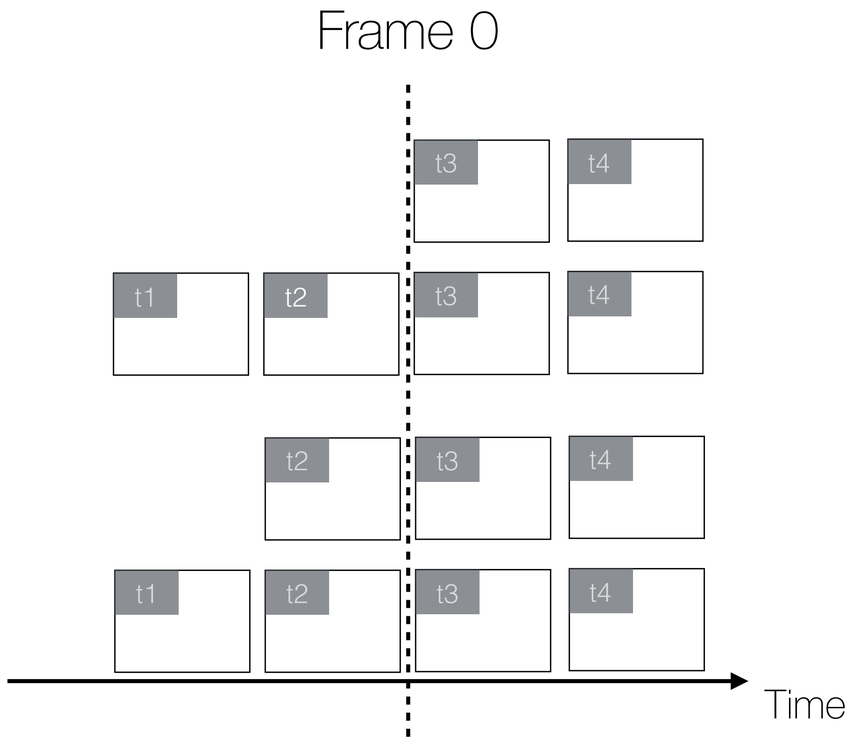


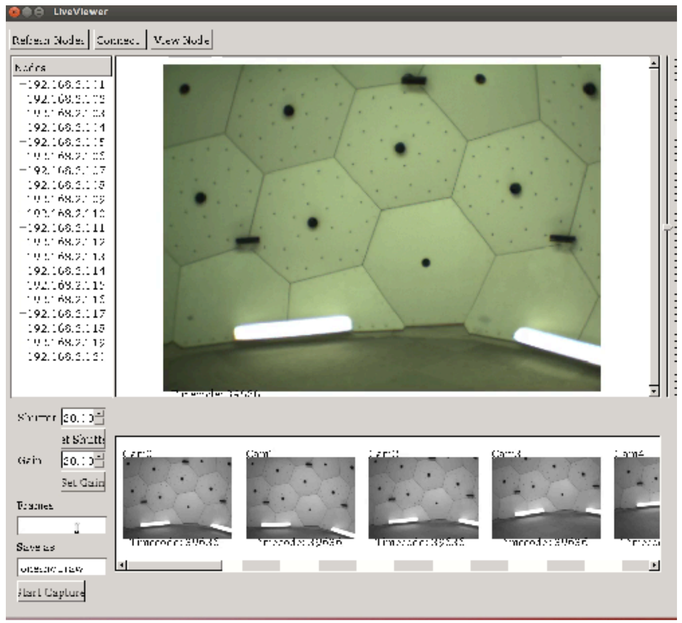


- Pre-visualization
- Aligned start time

Capture Software

Take Home Messages





Timestamped frames allow
a posteriori time alignment

Better capture software may
provide conveniences
(e.g., preview)

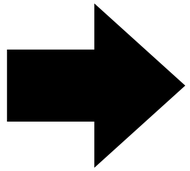
Data/Storage

Data / Storage

Why Do We Need To Consider This?

We would prefer,

- Higher resolution
- Higher frame rate
- More viewpoints

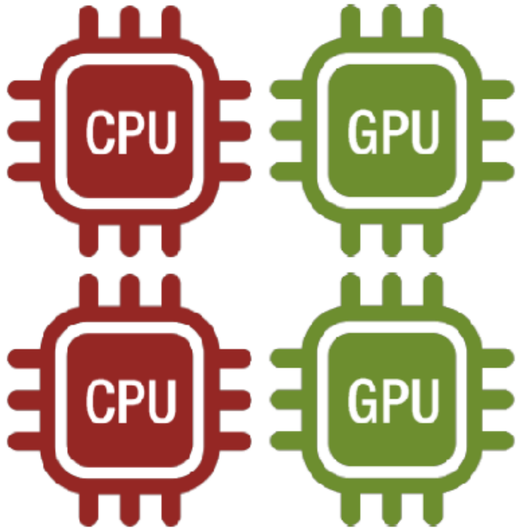


- Large size
- Hard to save

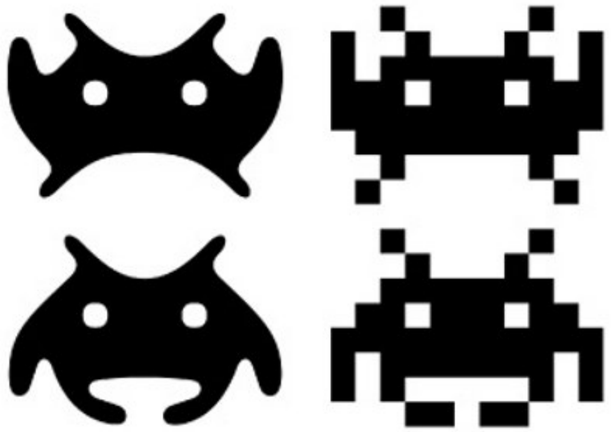
In Panoptic Studio,
8.85 GB/sec = 531 GB/min

Data / Storage

Options?



- Do not save images
- Real-time system

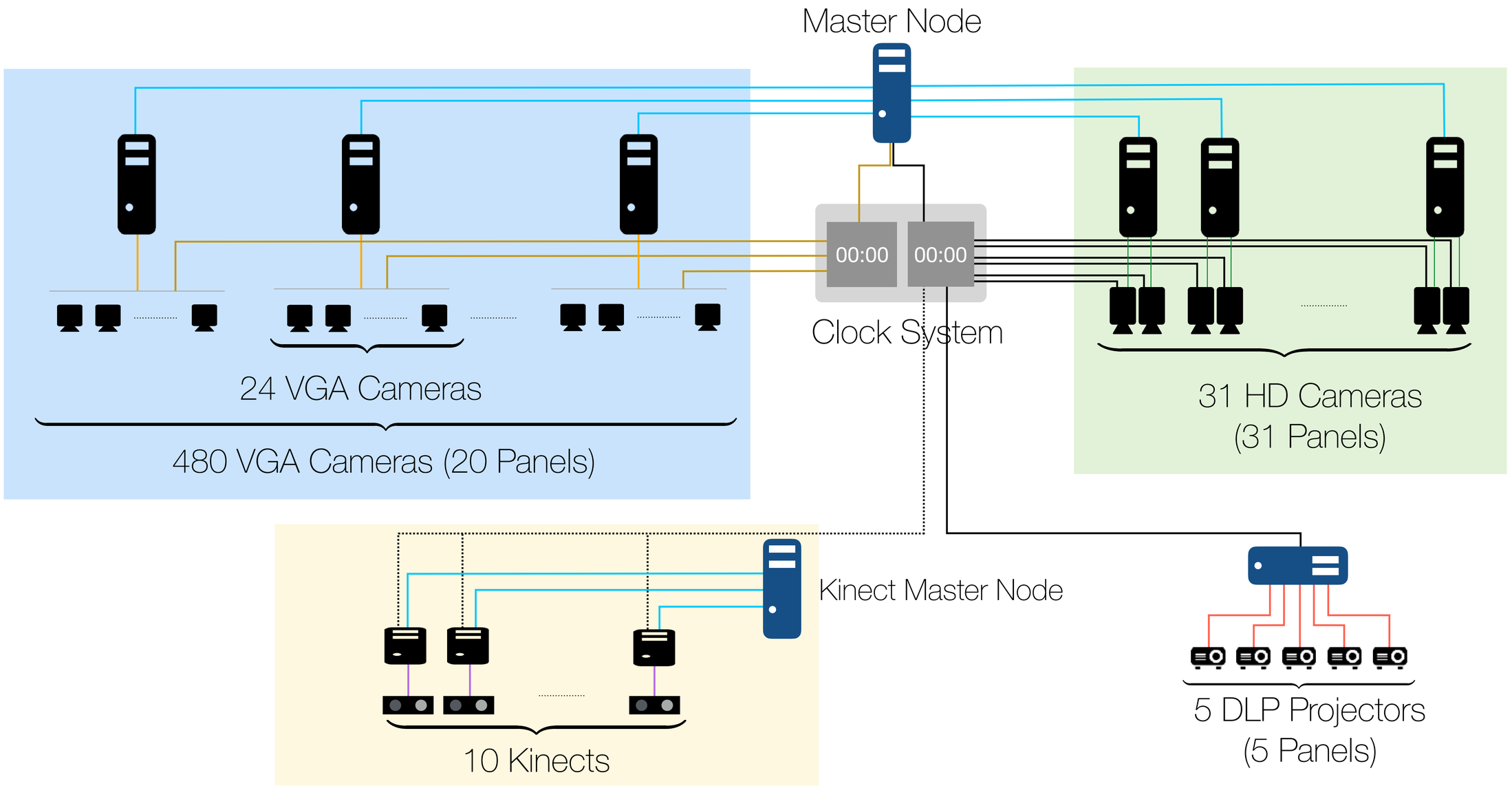


- Save compressed data

- Save raw data

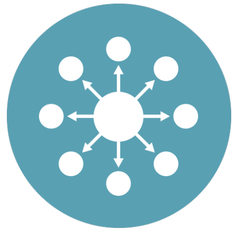
Data / Storage

Our Solution



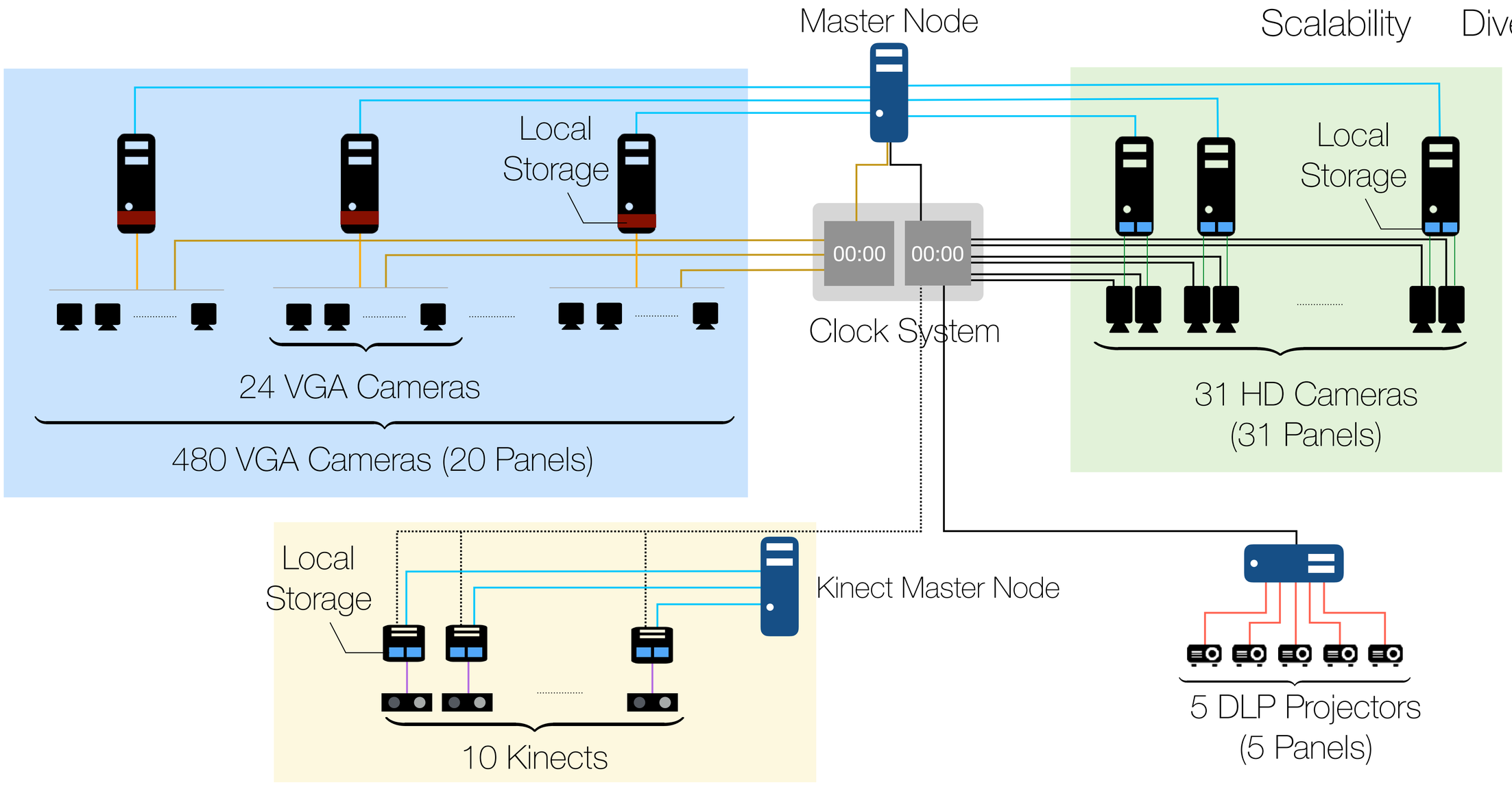
Data / Storage

Our Solution: Local Storage Systems



Scalability

Diversity



Data / Storage

How Big Is The Panoptic Studio Data?

VGA (24 cameras): **184 MB/s** x 20

640 x 480 x 1 Byte/pixel x 24 cameras x 25 frame/sec

HD: **120 MB/s** x 31

1920x1080 x 2 Byte/pixel x 30 frames/sec

Kinect: **145 MB/s** x 10

1920 x 1080 x 2 Byte/pixel x 30 fps
+ depth 512 x 424 x 2 Byte/pixel x 30 fps (12.4 MB/s)
+ IR 512 x 424 x 2 Byte/pixel x 30 fps (12.4 MB/s)
+ body keypoints, audio

Note: average HDD's writing speed: **80-160** MB/s

Data / Storage

How Big Is The Panoptic Studio Data?

VGA (24 cameras): **184 MB/s** x 20

640 x 480 x 1 Byte/pixel x 24 cameras x 25 frame/sec

RAID 0

Why not SSD? ⏱ 💵

SSD (1TB) can capture 90 minutes

HD: **120 MB/s** x 31

1920x1080 x 2 Byte/pixel x 30 frames/sec

Kinect: **145 MB/s** x 10

1920 x 1080 x 2 Byte/pixel x 30 fps

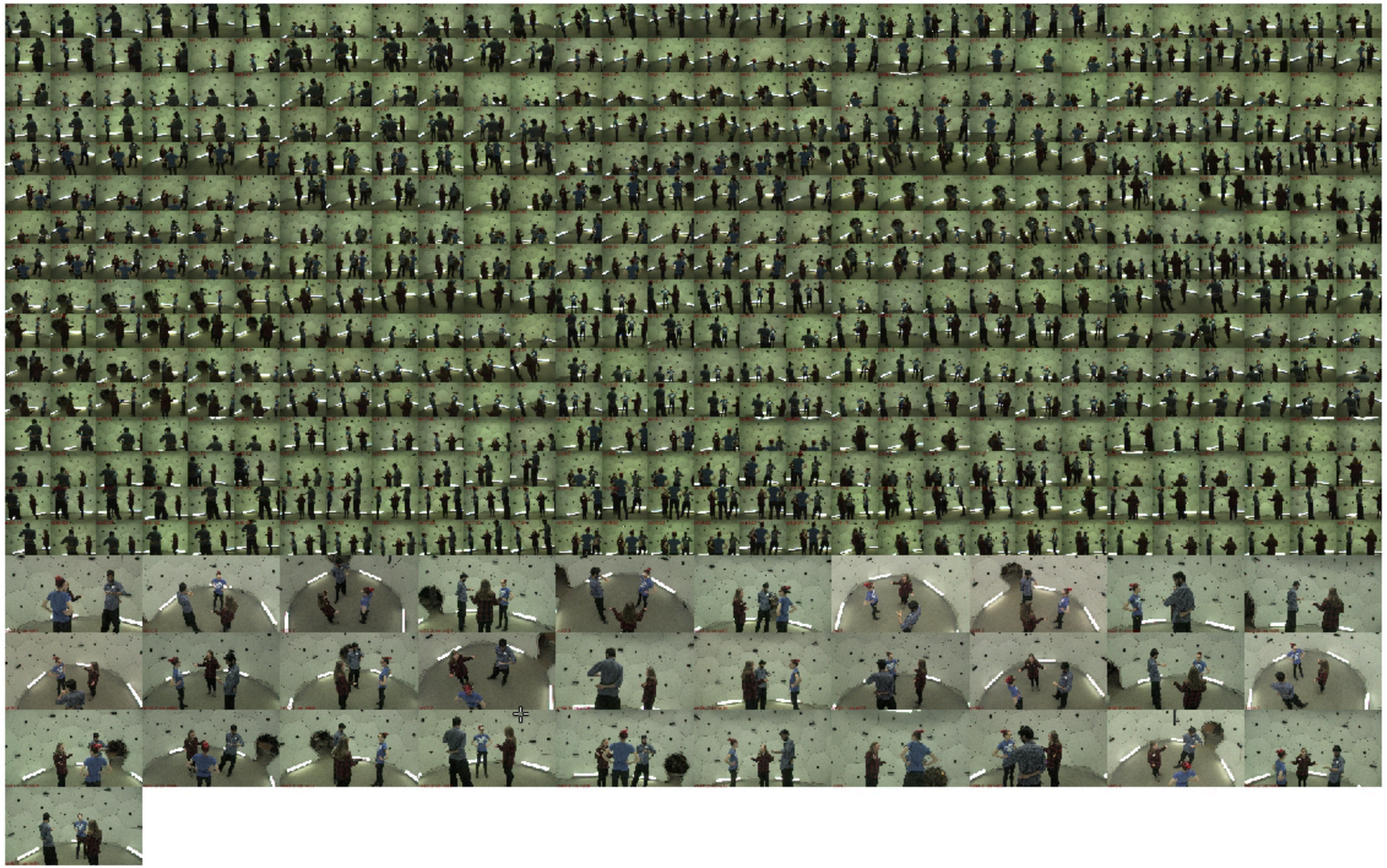
+ depth 512 x 424 x 2 Byte/pixel x 30 fps (12.4 MB/s)

+ IR 512 x 424 x 2 Byte/pixel x 30 fps (12.4 MB/s)

+ body keypoints, audio

Note: average HDD's writing speed: **80-160** MB/s

CaptureFileViewer



	File Name
1	/rcapture/ve01/1704...
2	/rcapture/ve02/1704...
3	/rcapture/ve03/1704...
4	/rcapture/ve04/1704...
5	/rcapture/ve05/1704...
6	/rcapture/ve06/1704...
7	/rcapture/ve07/1704...
8	/rcapture/ve08/1704...
9	/rcapture/ve09/1704...
10	/rcapture/ve10/1704...
11	/rcapture/ve11/1704...
12	/rcapture/ve12/1704...
13	/rcapture/ve13/1704...
14	/rcapture/ve14/1704...
15	/rcapture/ve15/1704...
16	/rcapture/ve16/1704...
17	/rcapture/ve17/1704...
18	/rcapture/ve18/1704...
19	/rcapture/ve19/1704...
20	/rcapture/ve20/1704...
21	/rcaptureHD/ve31-1/...
22	/rcaptureHD/ve31-2/...
23	/rcaptureHD/ve32-1/...
24	/rcaptureHD/ve32-2/...
25	/rcaptureHD/ve33-1/...
26	/rcaptureHD/ve33-2/...
27	/rcaptureHD/ve34-1/...
28	/rcaptureHD/ve34-2/...
29	/rcaptureHD/ve35-1/...
30	/rcaptureHD/ve35-2/...
31	/rcaptureHD/ve35-1/...

Select Fps

25 10 Bac Def

1704CV_juggling_a1

Load reaptured(VGA)

Load reaptured(HD)

Add File

Remove File

Refresh View

Refresh View By Panel

Check camera status

Save Frames 1 Frame Index 1

BMP JPG

Create cam folders

Frame Extraction

Save Sync Info

Create demo video

CAMS 5

FRAMES 11591 Go 0 13799

Storage System

Long-Term Storage

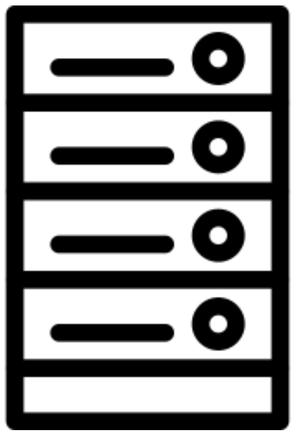
- NAS 12 x 8TB (Synology)
- 1 NAS: 88 TB (with RAID5) = 166 minutes
- $29 / min
  $1,700 + $260 x 12 = $4,820 / 166 min
- Currently Panoptic Studio has about 1 PB data

Storage System

Take Home Messages

Data size vs Writing speed

Required capacity

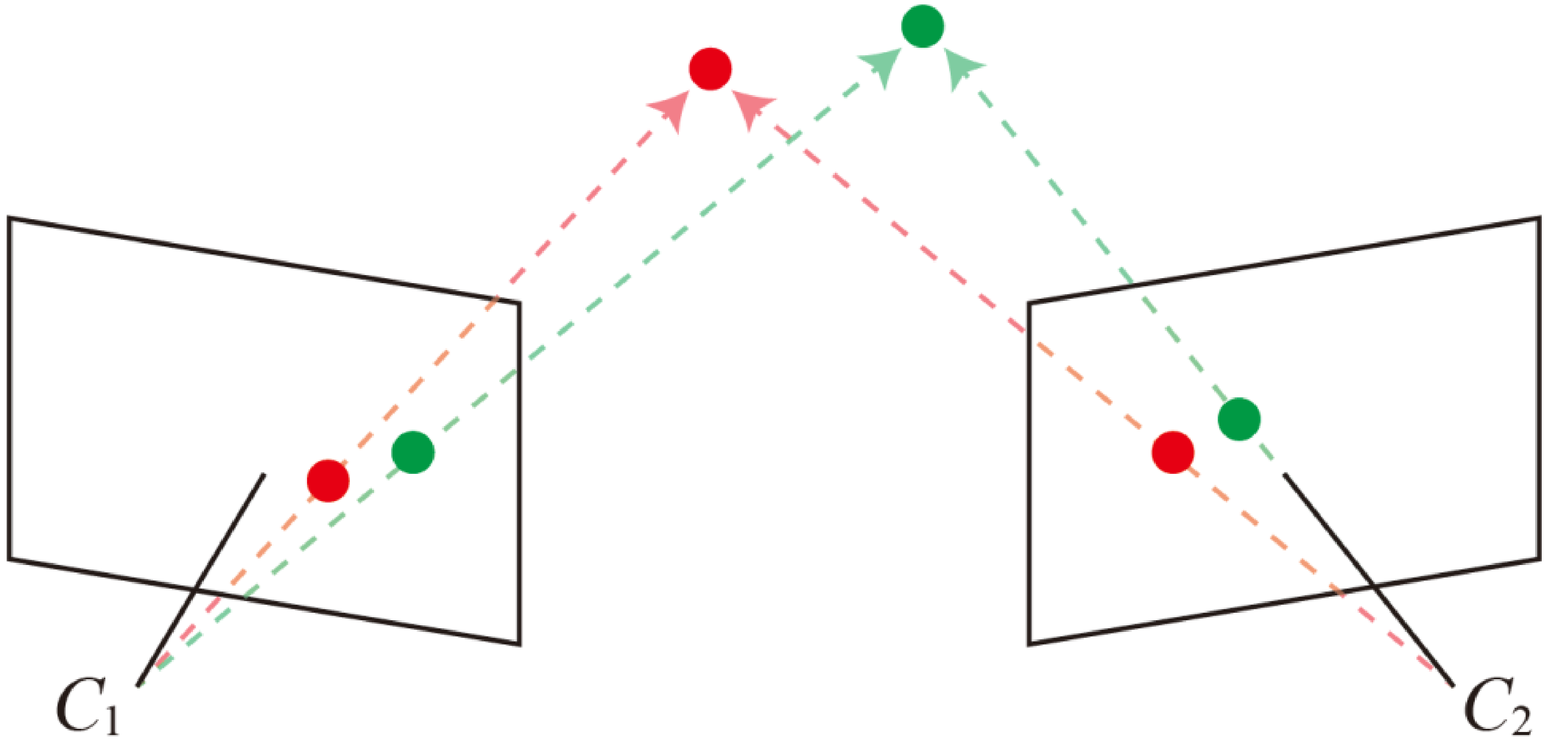


≠ Capture storage
Long-term storage

Camera Calibration

Camera Calibration

Why Do We Need This?



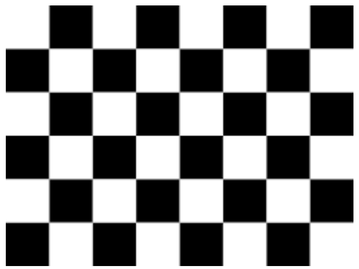
Intrinsic: $\mathbf{K}$

Extrinsic: $\mathbf{R}\ \mathbf{t}$

Lens distortion: k_1, k_2, p_1, p_2, p_3

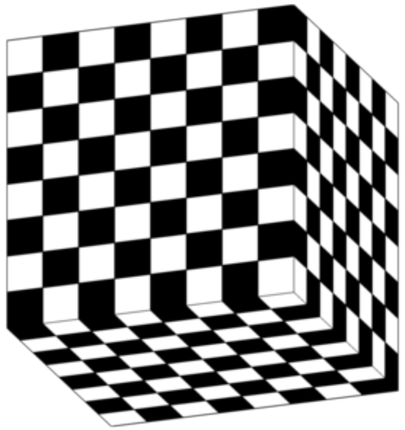
Camera Calibration

Space of Options



2D Checker Board

Wand (Light, Ball)

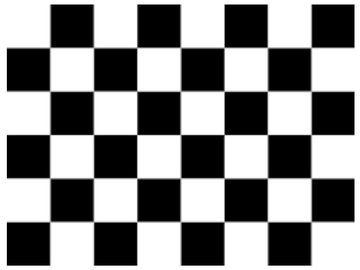


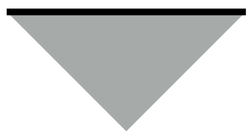
3D Pattern

Structure from Motion

Camera Calibration

A Single Camera Example

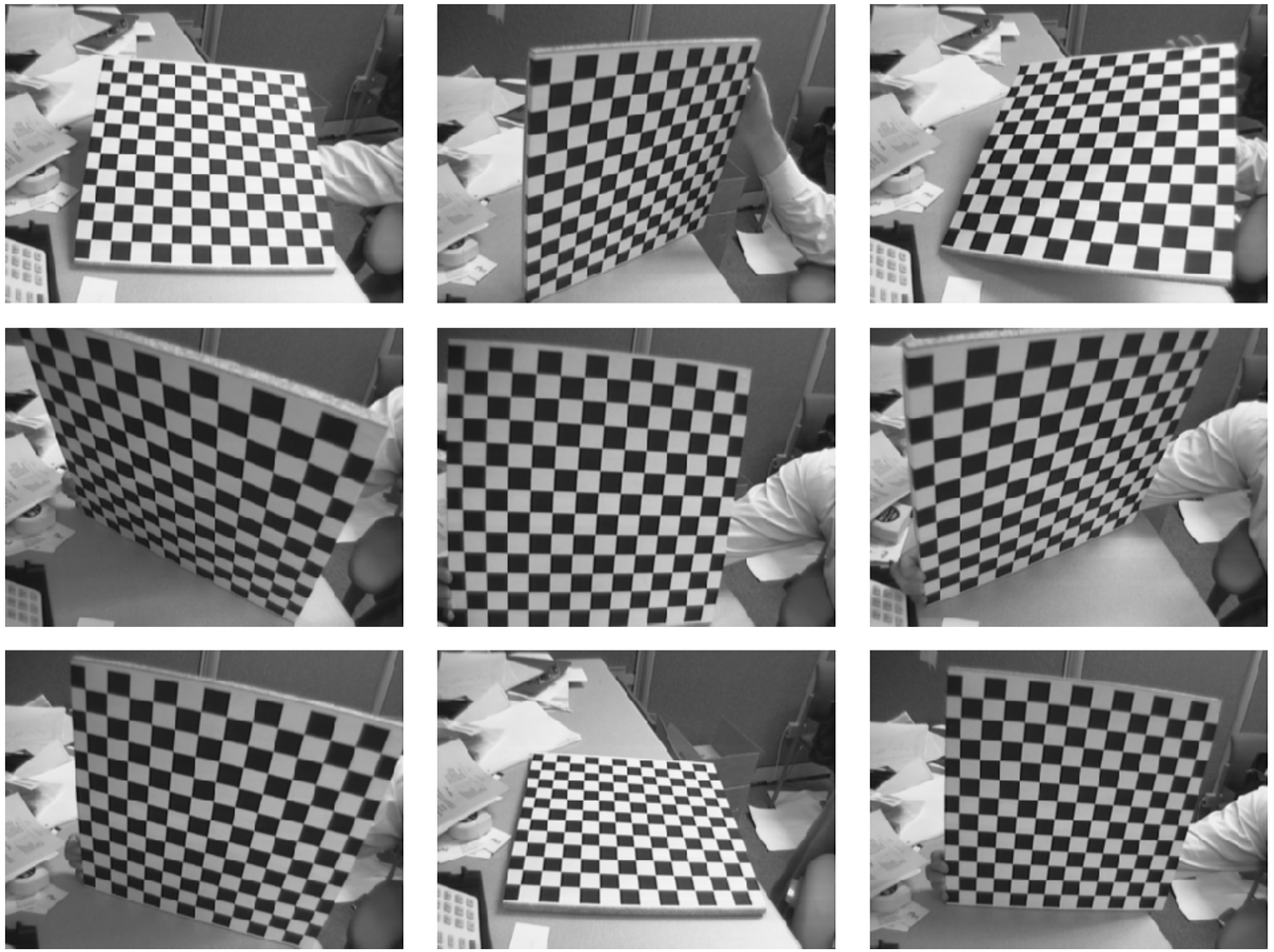




Intrinsic: $\mathbf{K}$

Extrinsic: $\mathbf{R}\,\mathbf{t}$

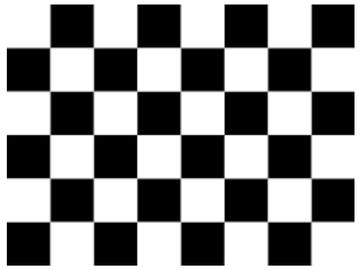
Lens distortion: k_1, k_2, p_1, p_2, p_3

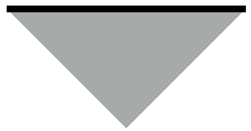


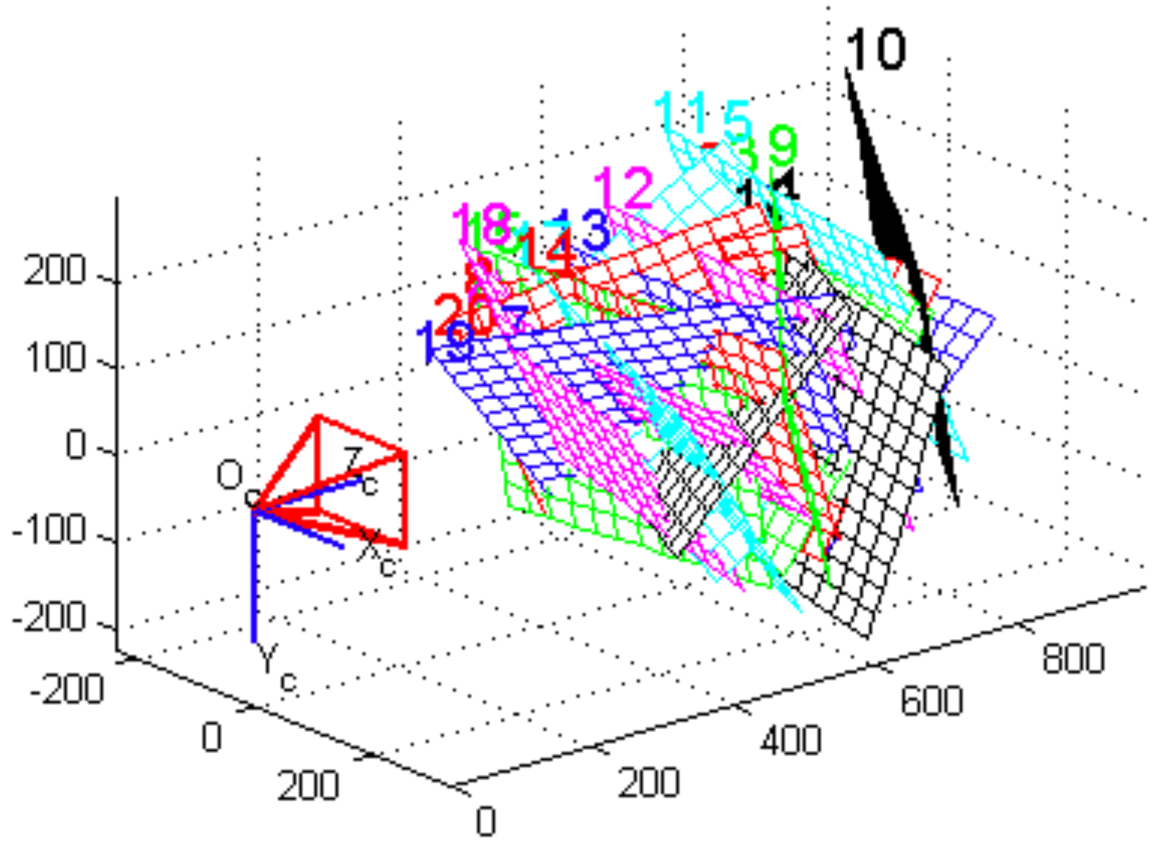
Caltech Camera Calibration Toolbox

Camera Calibration

A Single Camera Example







Intrinsic: $\mathbf{K}$

Extrinsic: $\mathbf{R}\,\mathbf{t}$

Lens distortion: k_1, k_2, p_1, p_2, p_3

Caltech Camera Calibration Toolbox

A Planar Pattern To Calibrate Panoptic Studio?

Accuracy

- Good accuracy

Scalability

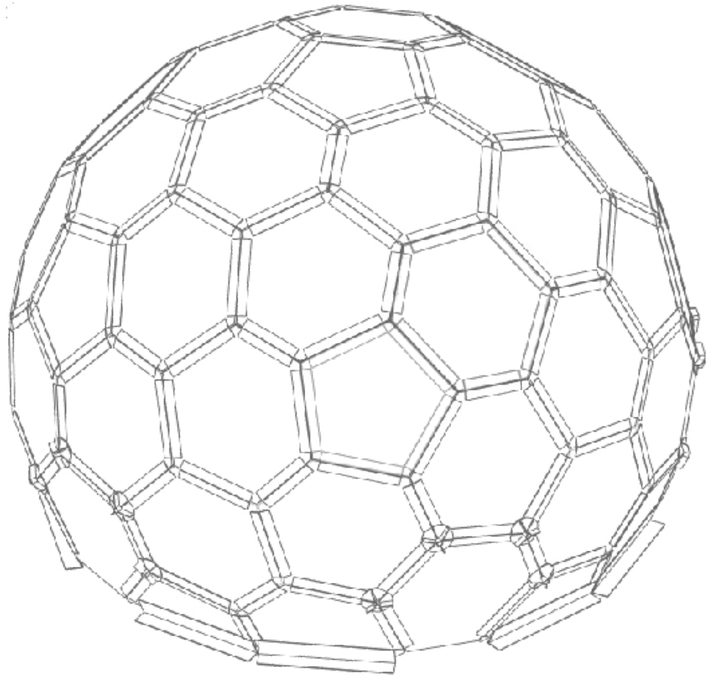
- Need to cover many locations

Diversity

- Problem for sensors with different frame rates (VGA 25fps, HD 30fps, Kinects 30fps)

Our Solution For Camera Calibration

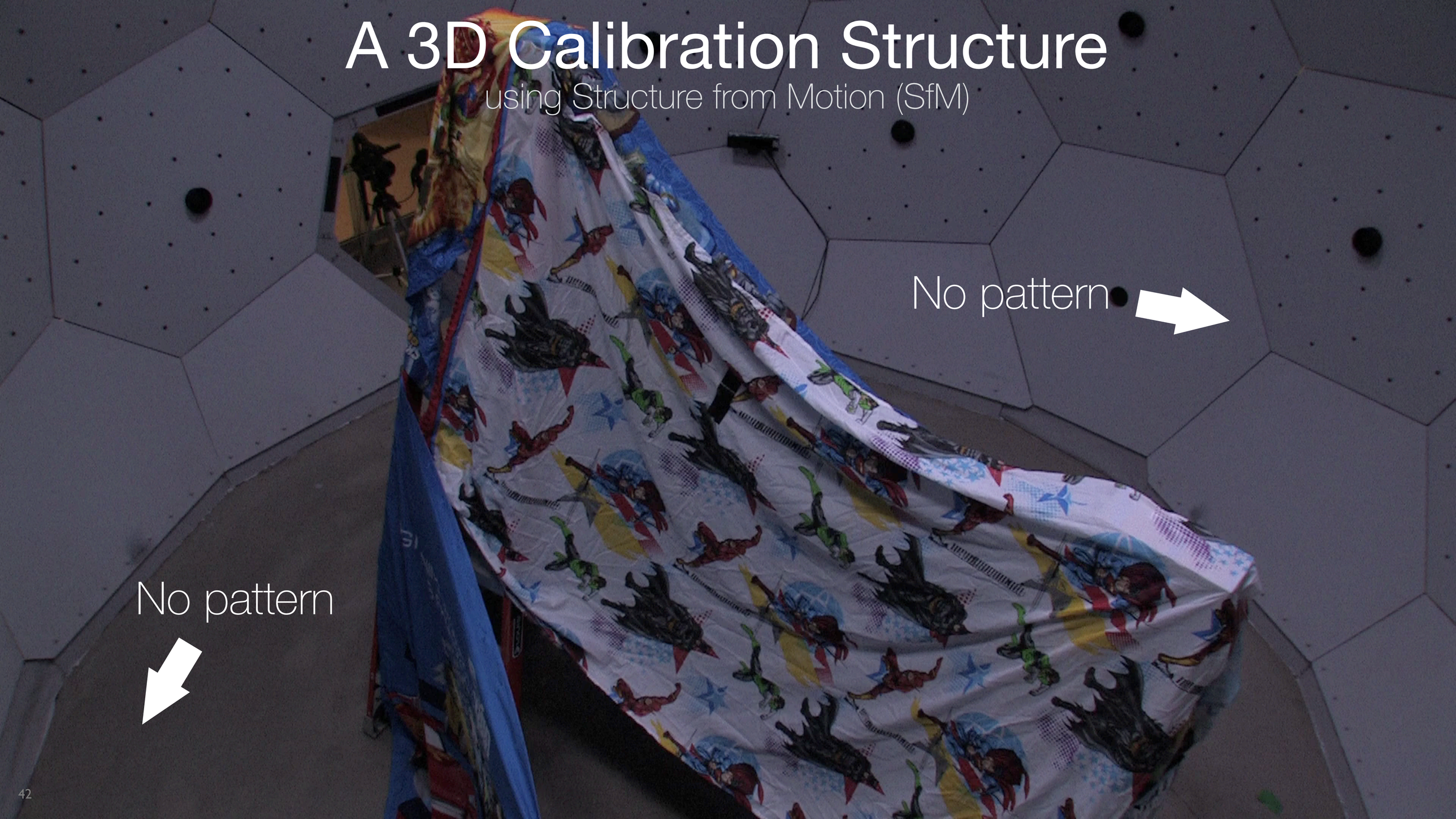
How To Calibrate Panoptic Studio Cameras



- **Different types of cameras**
  - No perfect sync among different types
  - Pattern should be stationary
- **A large working volume**
  - Pattern should cover as much space as possible
- **500+ cameras**
  - Fully automatic method is needed
  - Avoid any image selections

A 3D Calibration Structure

using Structure from Motion (SfM)



Panoptic Studio Camera Calibration

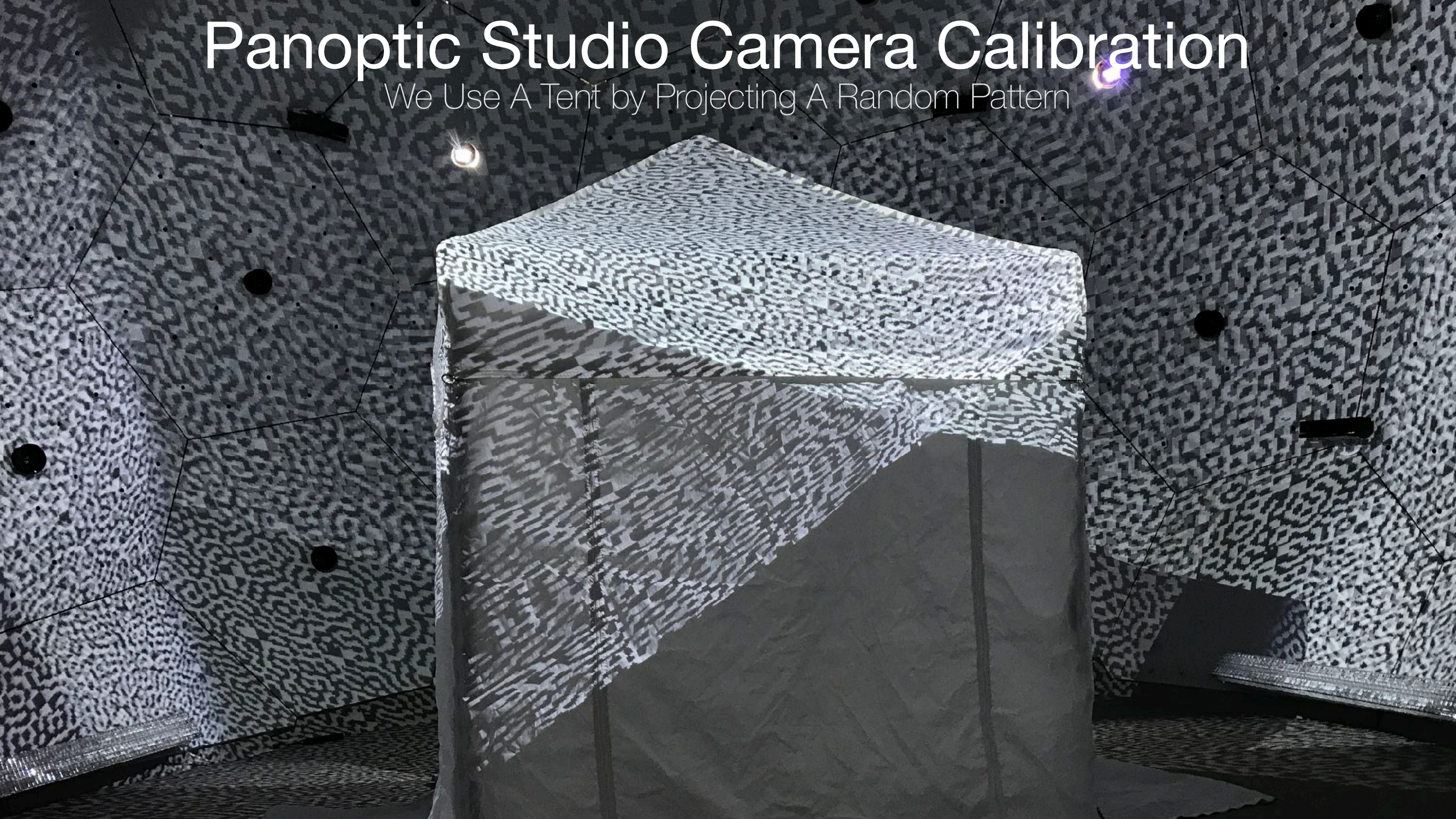
We Use A Tent

Panoptic Studio Camera Calibration

We Use A Tent

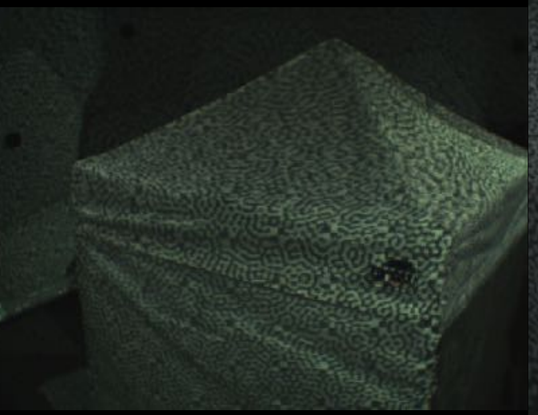
Panoptic Studio Camera Calibration

We Use A Tent by Projecting A Random Pattern

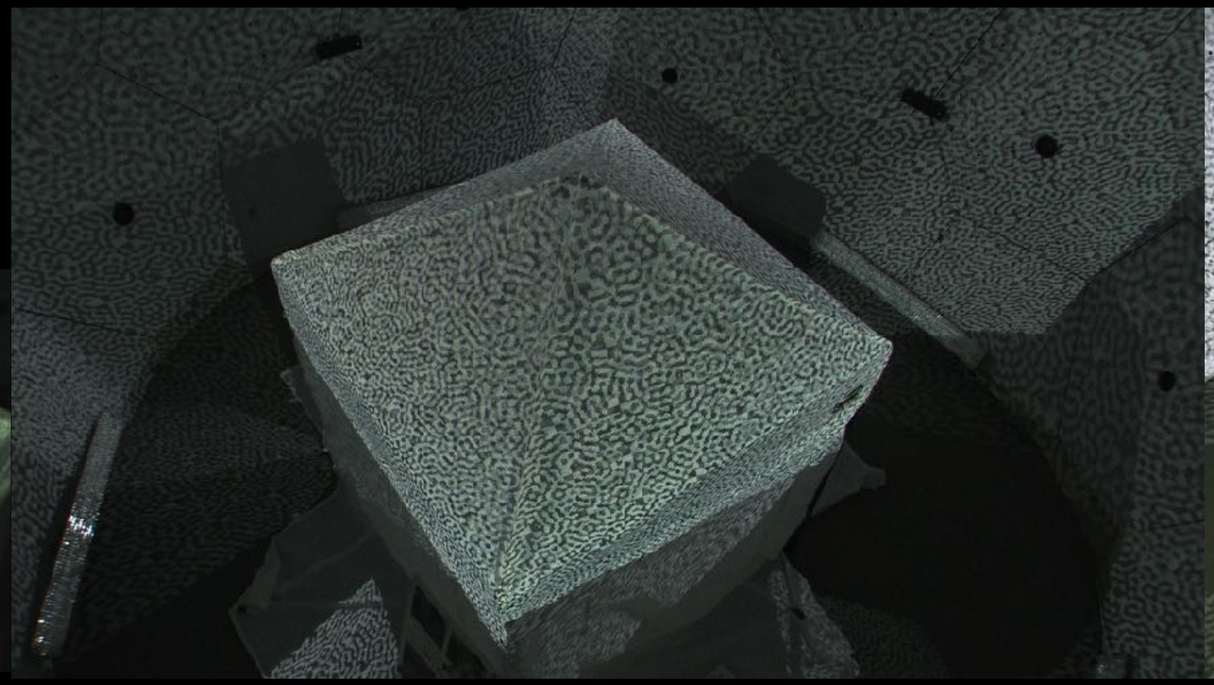


Panoptic Studio Camera Calibration

We Use A Tent by Projecting A Random Pattern



VGA

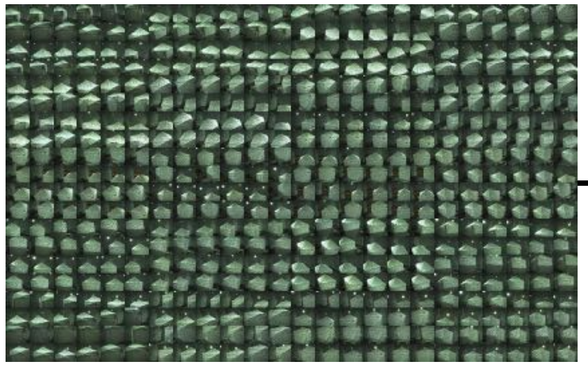


HD

Kinect Color

Calibration Pipeline

Based on Structure from Motion (SfM)



Capture

→ Matching → SfM

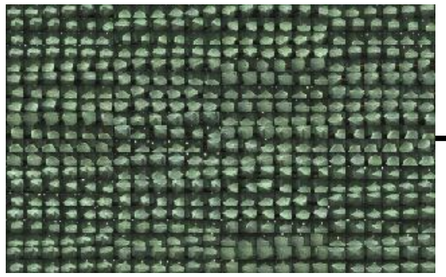
Intrinsic: $\mathbf{K}$
Extrinsic: $\mathbf{R}\,\mathbf{t}$
Lens distortion: k_1

→ Bundle Adjustment →

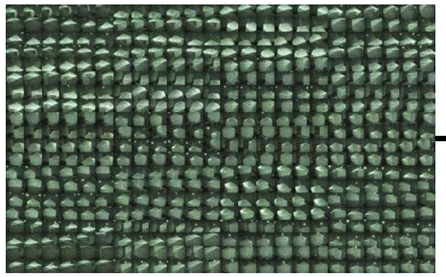
Intrinsic: $\mathbf{K}$
Extrinsic: $\mathbf{R}\,\mathbf{t}$
Lens distortion:
k_1, k_2, p_1, p_2, p_3

Calibration Pipeline

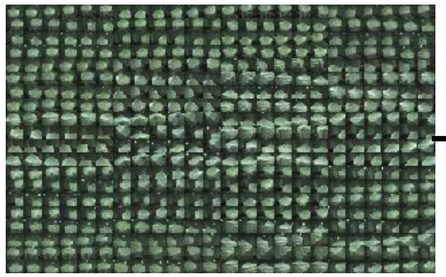
Based on Structure from Motion (SfM)



Capture 1



Capture 2



Capture 3

Matching

Matching

Matching

Merging

SfM

Intrinsic: $\mathbf{K}$
Extrinsic: $\mathbf{R}\,\mathbf{t}$
Lens distortion: k_1

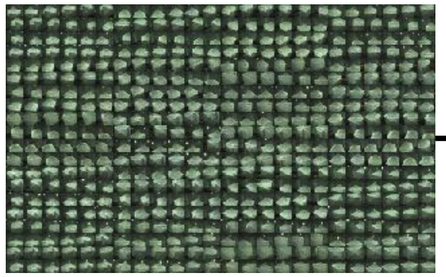
Bundle Adjustment

Intrinsic: $\mathbf{K}$
Extrinsic: $\mathbf{R}\,\mathbf{t}$
Lens distortion:

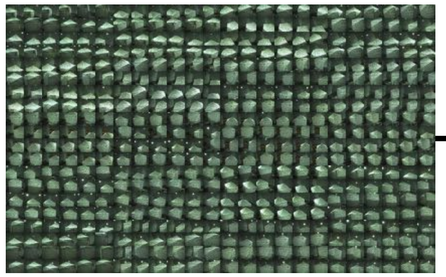
k_1, k_2, p_1, p_2, p_3

Calibration Pipeline

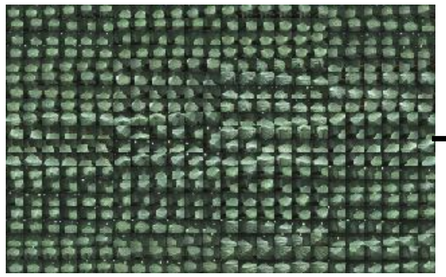
Based on Structure from Motion (SfM)



Capture 1



Capture 2



Capture 3

Matching

Matching

Matching

Merging

SfM

Intrinsic: $\mathbf{K}$
Extrinsic: $\mathbf{R}\,\mathbf{t}$
Lens distortion: k_1

Bundle Adjustment

Intrinsic: $\mathbf{K}$
Extrinsic: $\mathbf{R}\,\mathbf{t}$
Lens distortion:

k_1, k_2, p_1, p_2, p_3

Panoptic Studio Camera Calibration

Discussion

- Less accurate keypoint localization
- Scale factor is missing

- 5 minutes for capture
- Fully automatic process

- Applicable for different types of sensors (except Kinect depth)

Camera Calibration

Take Home Messages

Three Basic Steps for Calibration:

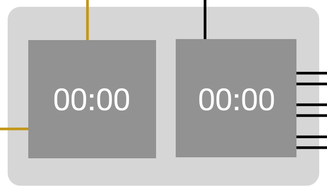
- Identify & match correspondences
  - Keypoint localization
  - Coverage of working volume
  - Practicality
- Initialization (camera matrices and 3D points)
- Bundle adjustment
  - VisualSfM + custom codes
  - COLMAP (https://colmap.github.io/)

Miscellaneous

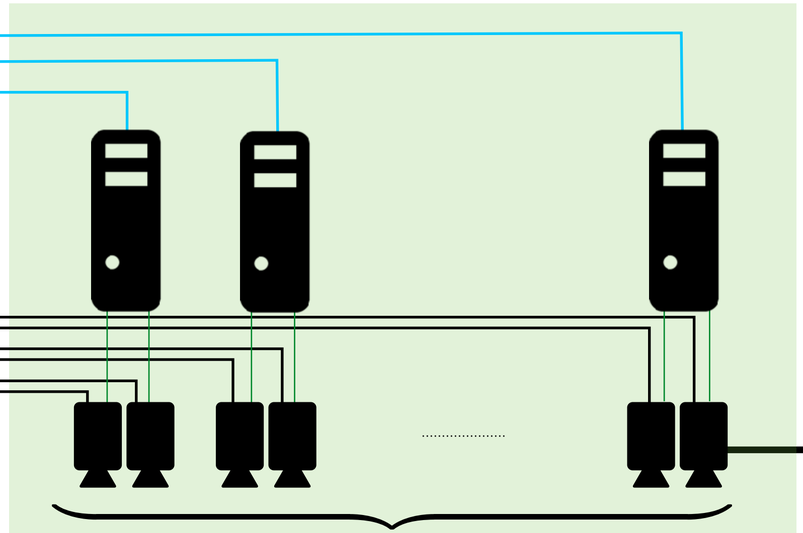
Audio Capture

Timestamped by HD Cameras

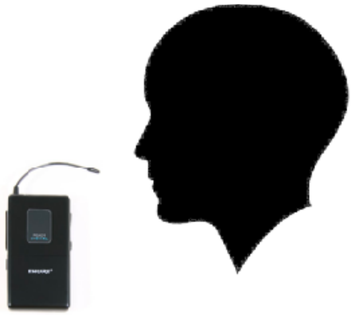
Master Node



Clock System



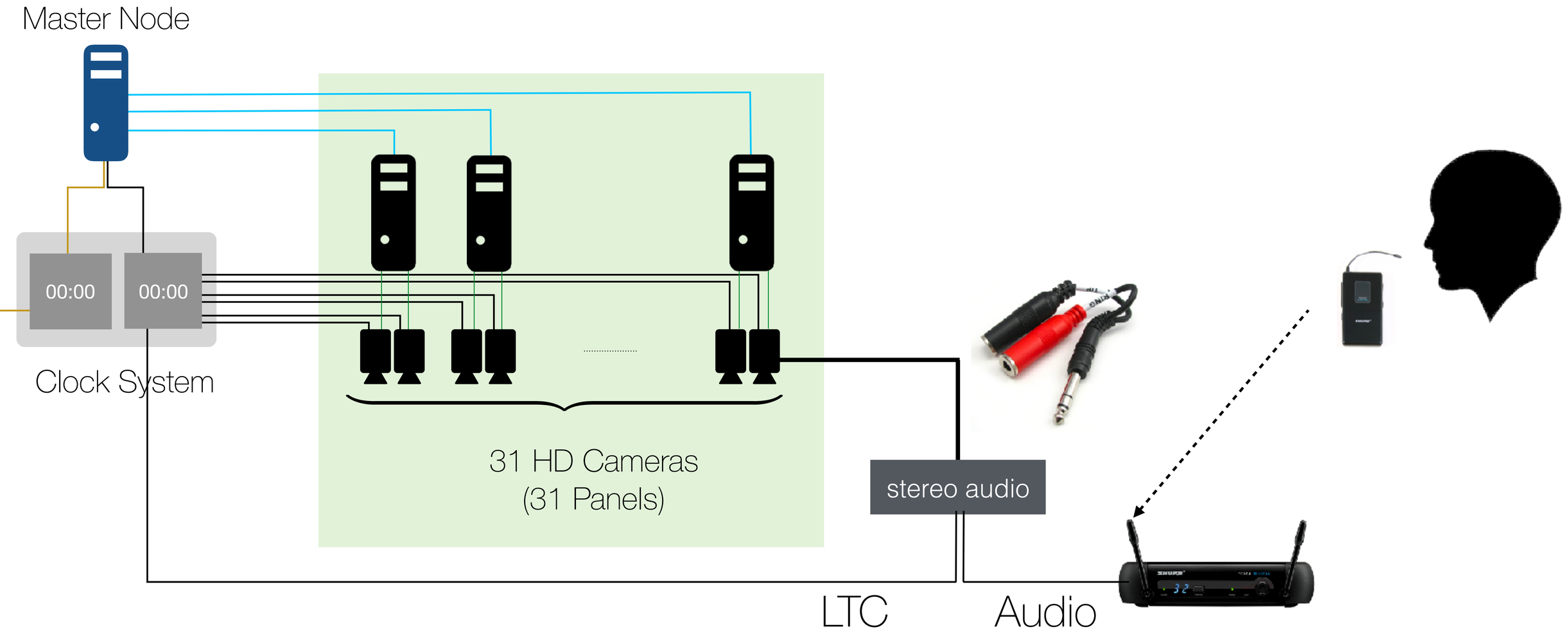
31 HD Cameras
(31 Panels)



Currently 15 wireless + 3 fixed microphones

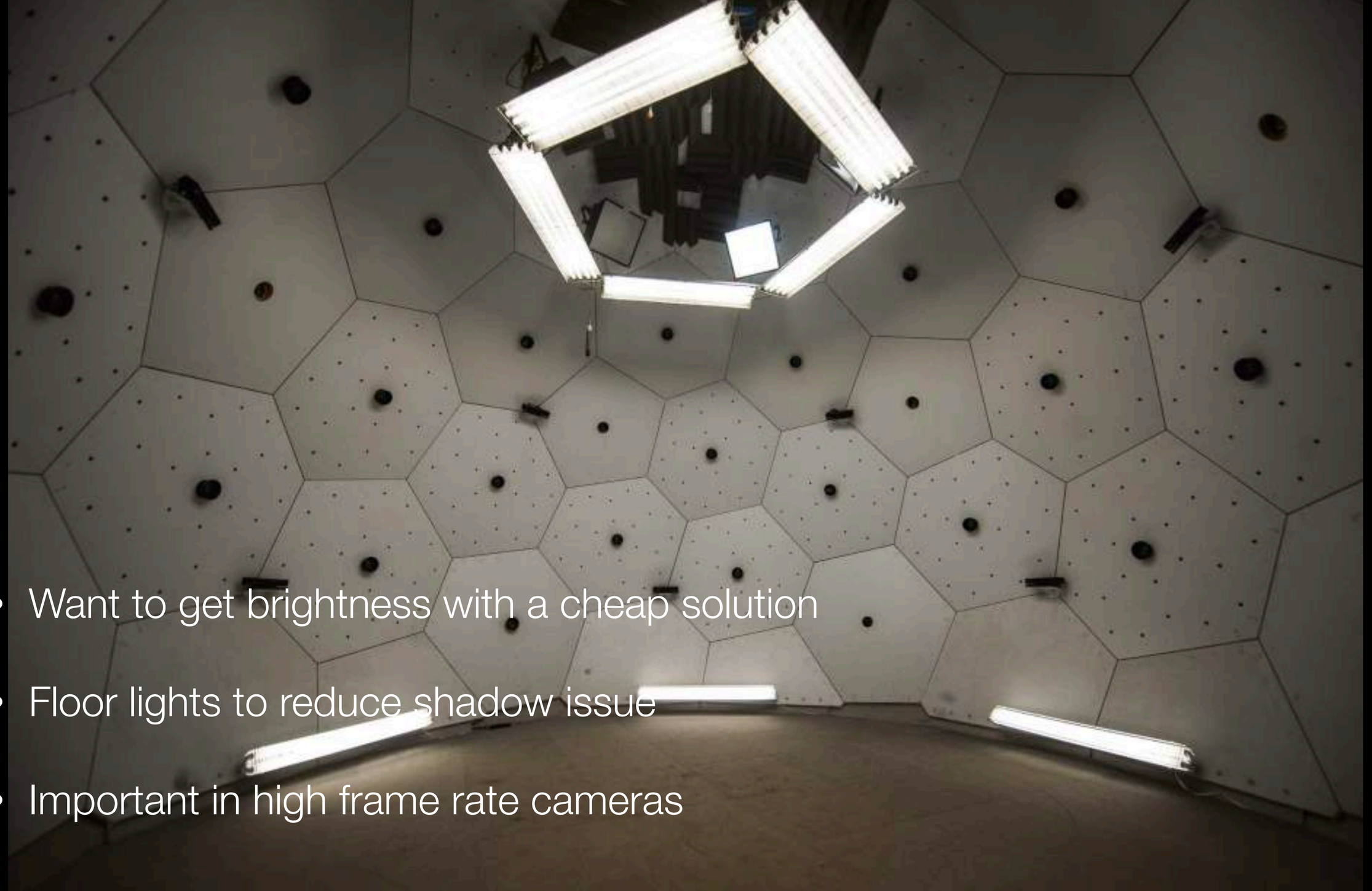
Audio Capture

Timestamped by HD Cameras



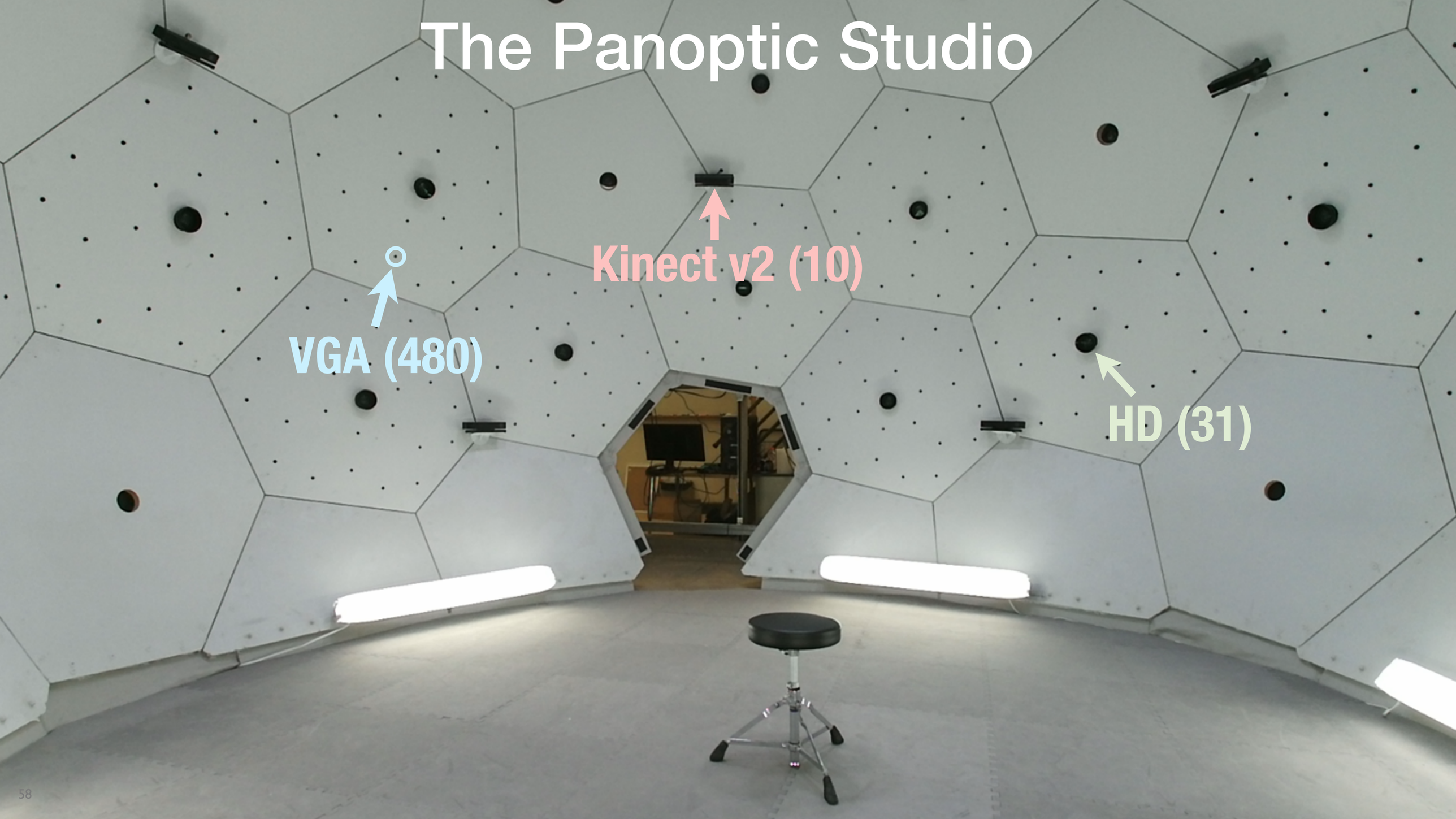
Lighting

We originally didn't plan for this



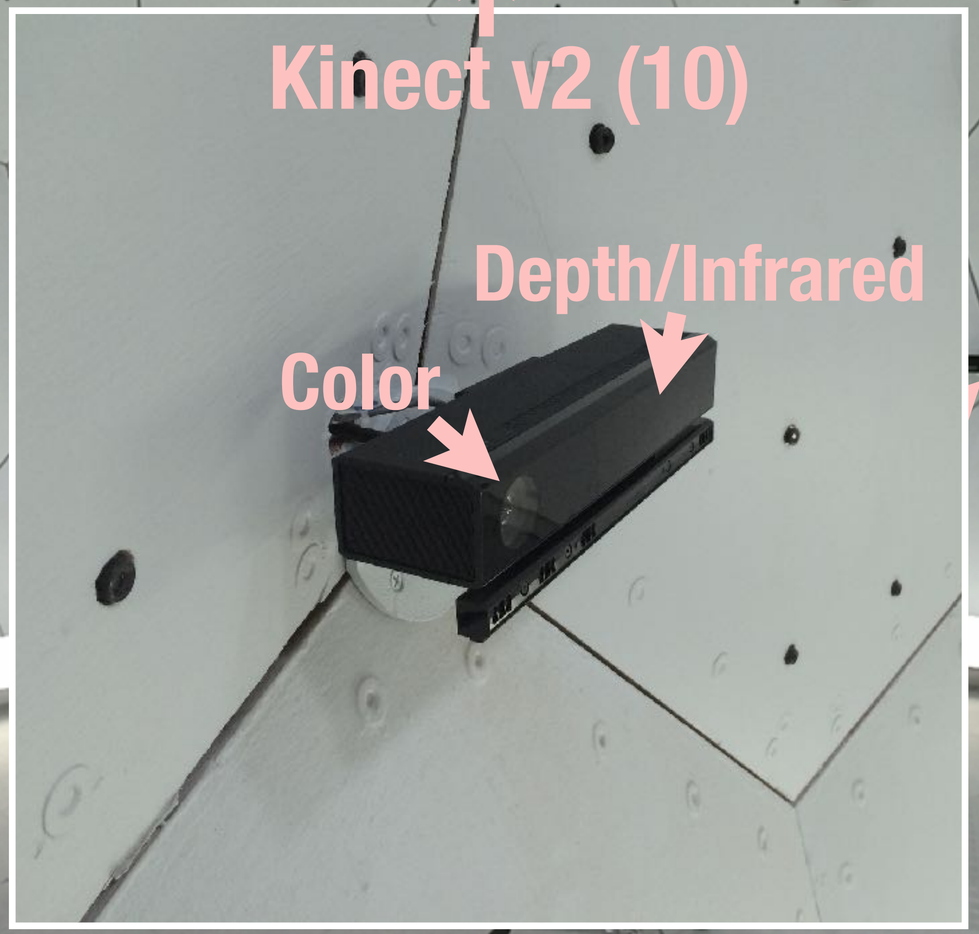
- Want to get brightness with a cheap solution
- Floor lights to reduce shadow issue
- Important in high frame rate cameras

Multiview RGB-D: Kinect Subsystem

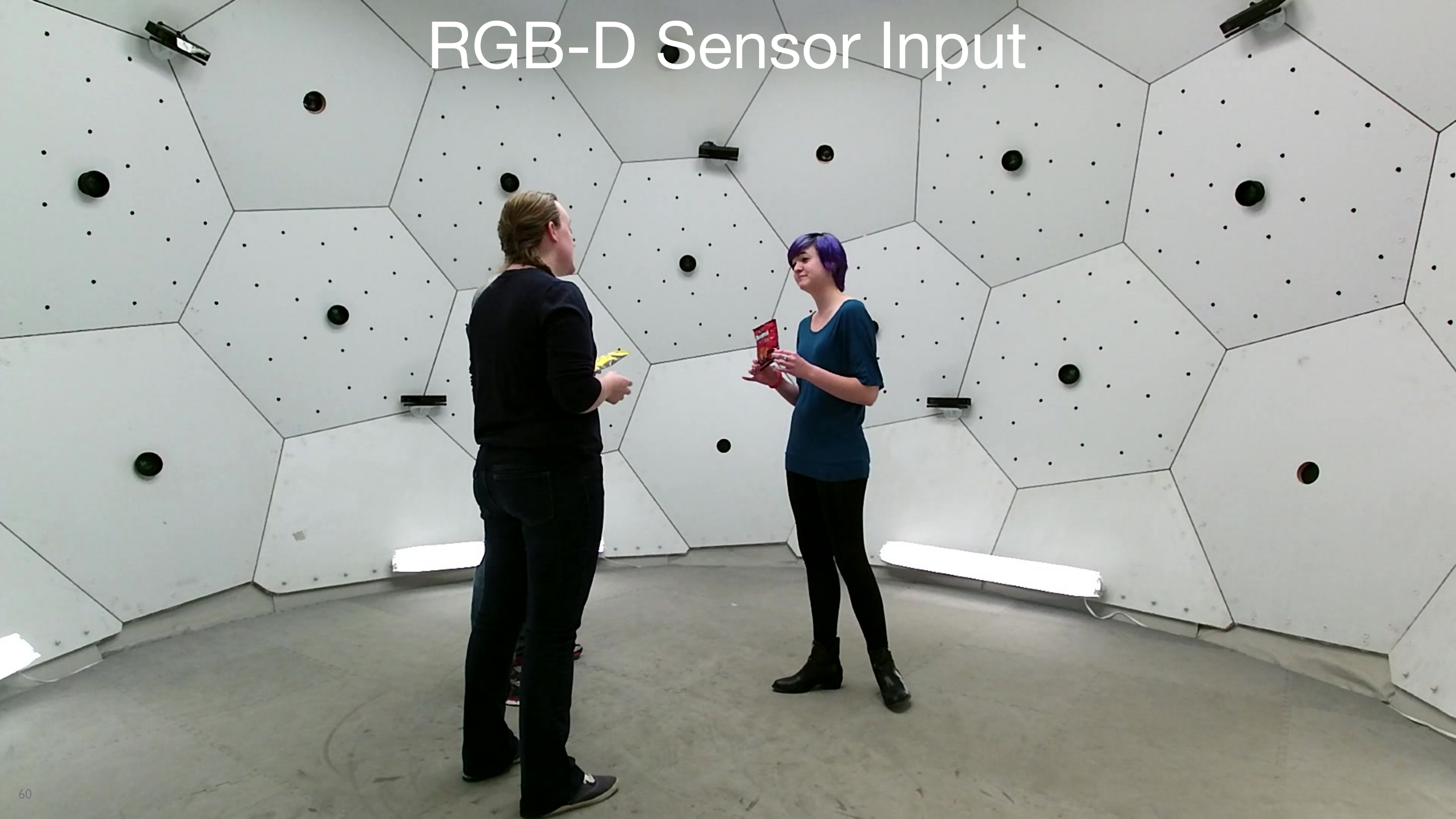

The Panoptic Studio
Kinect v2 (10)
VGA (480)
HD (31)

The Panoptic Studio

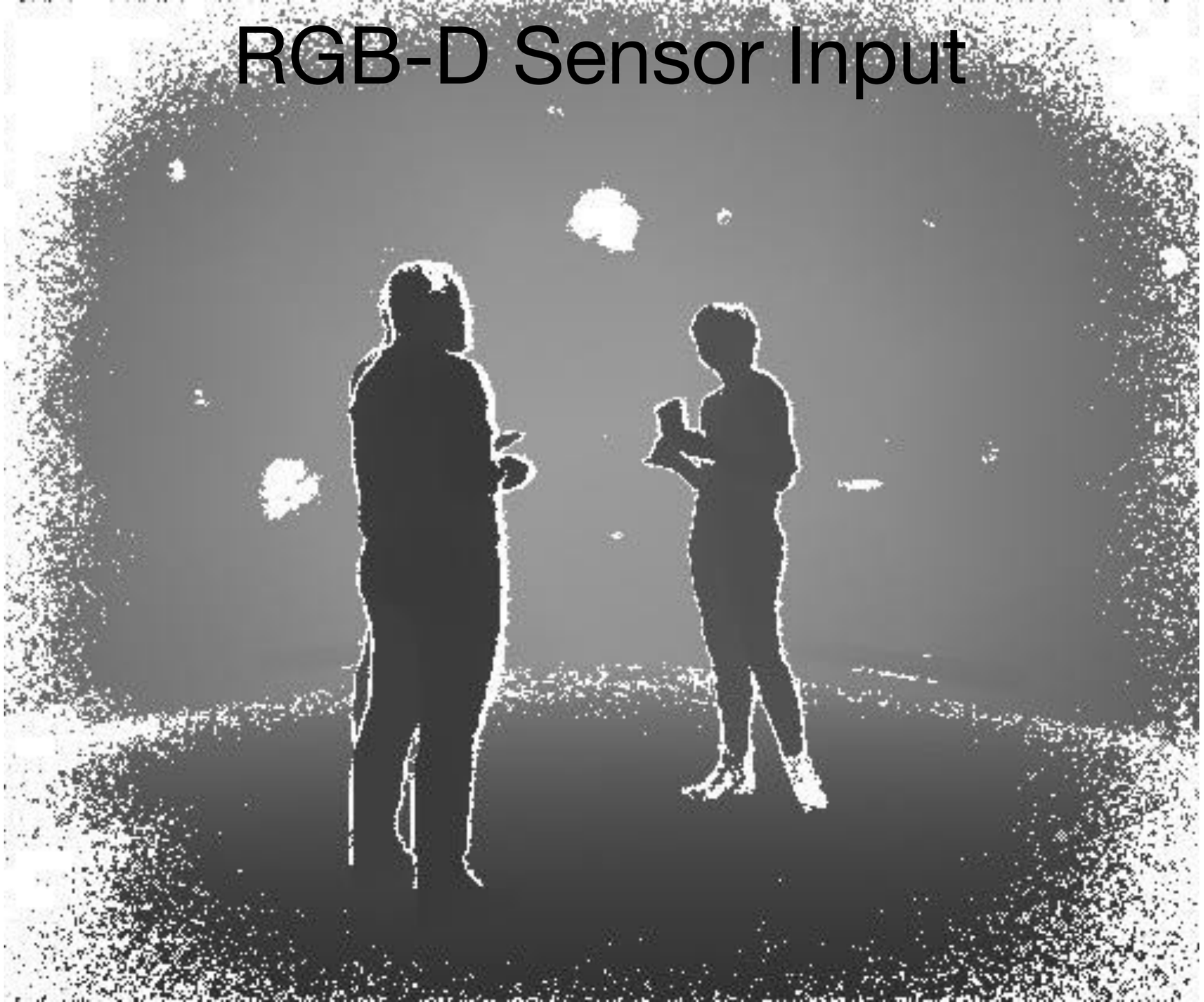
10 Kinect Subsystem

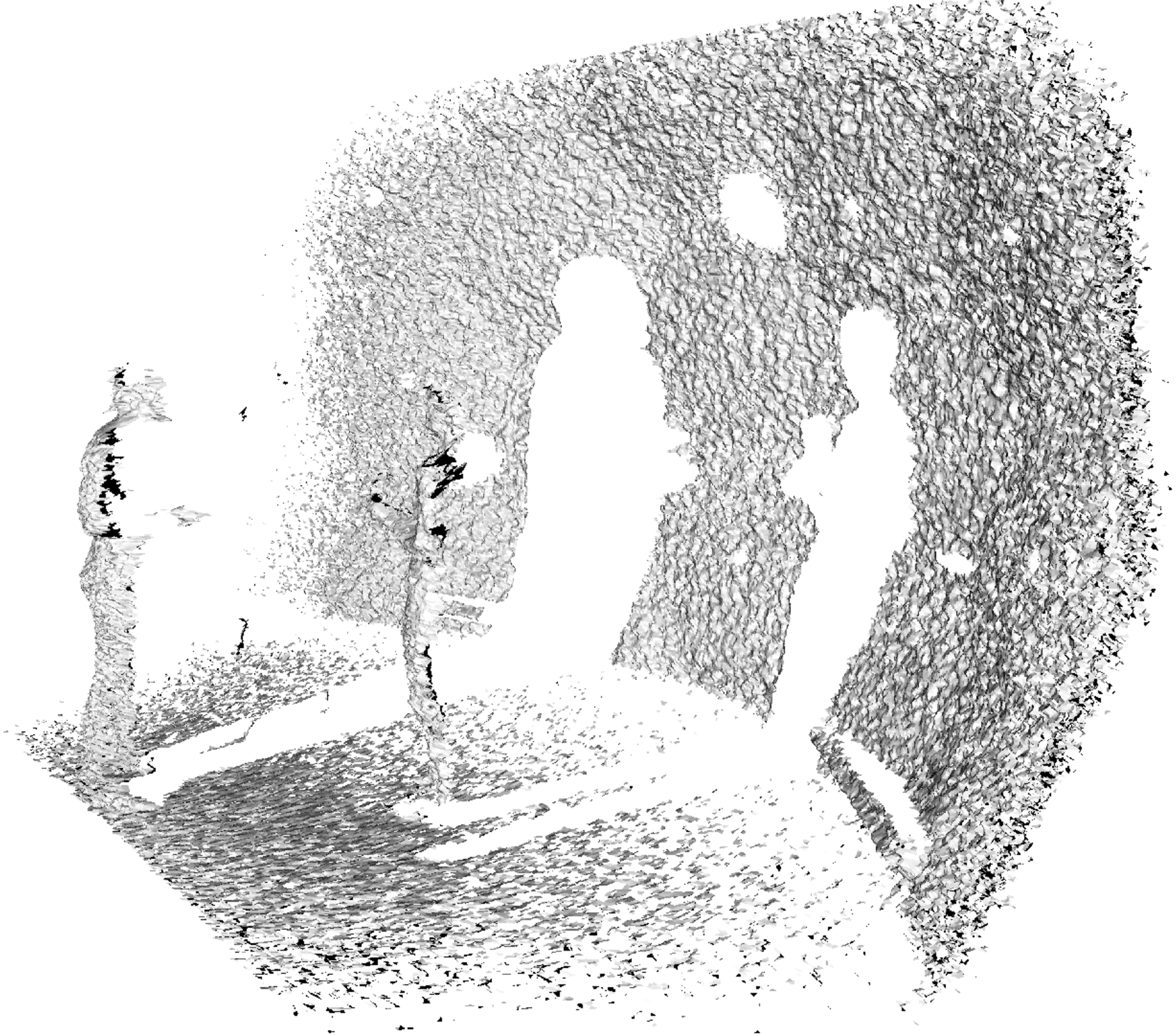


RGB-D Sensor Input



RGB-D Sensor Input

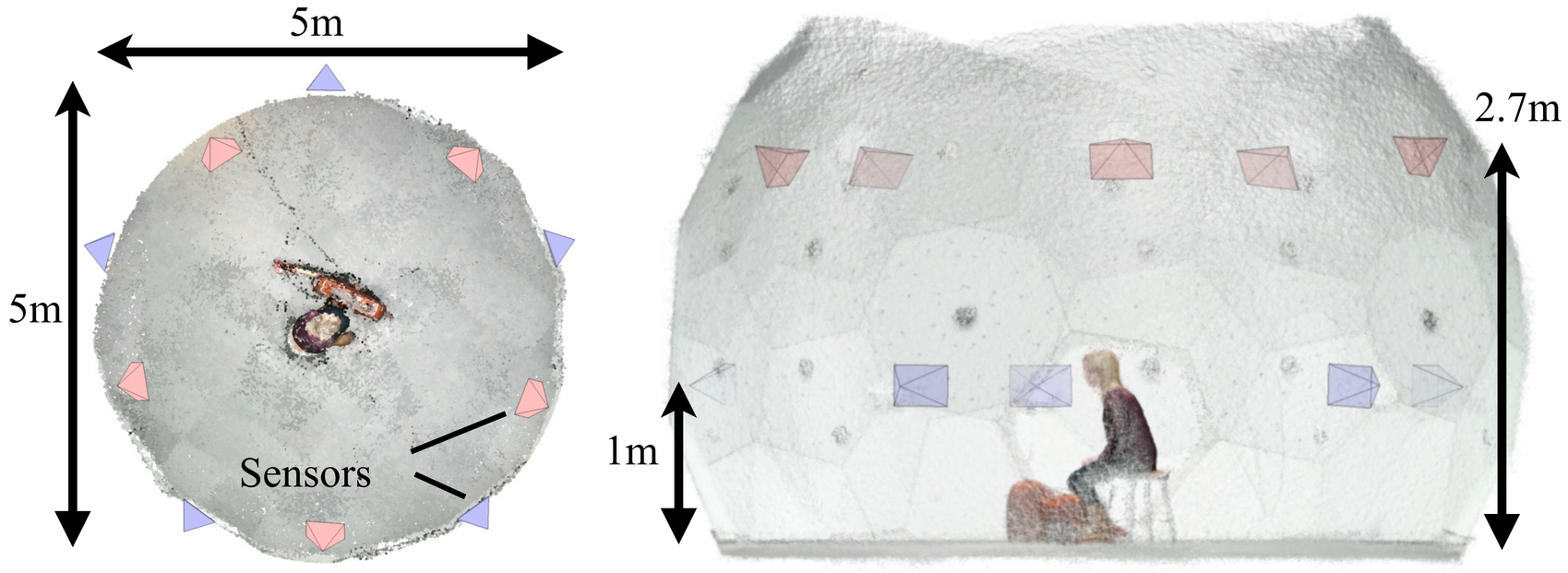




Ten RGB-D Sensors

Sensor Placement

Sensor Placement



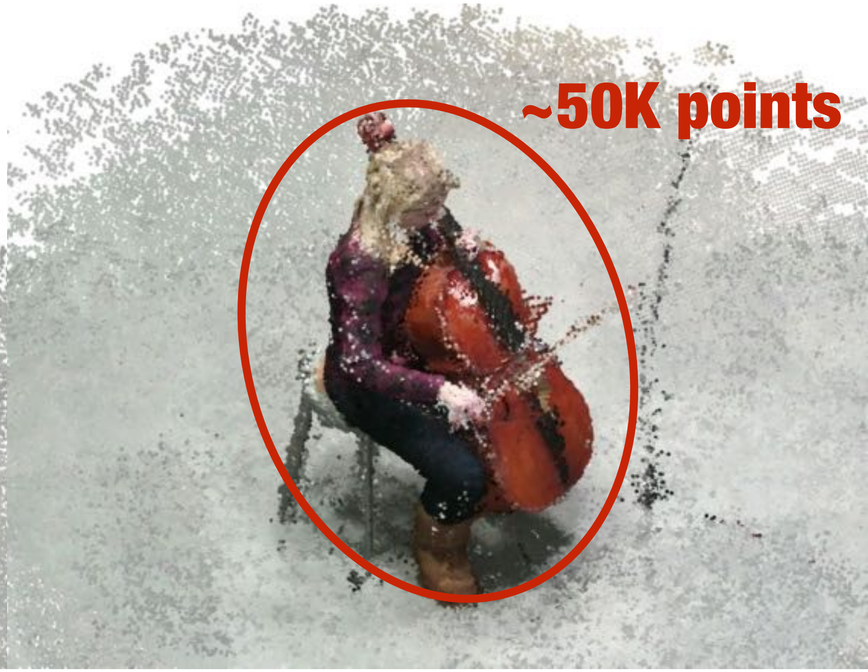
Spatial Sampling

~2M points

One

~10M points

Five

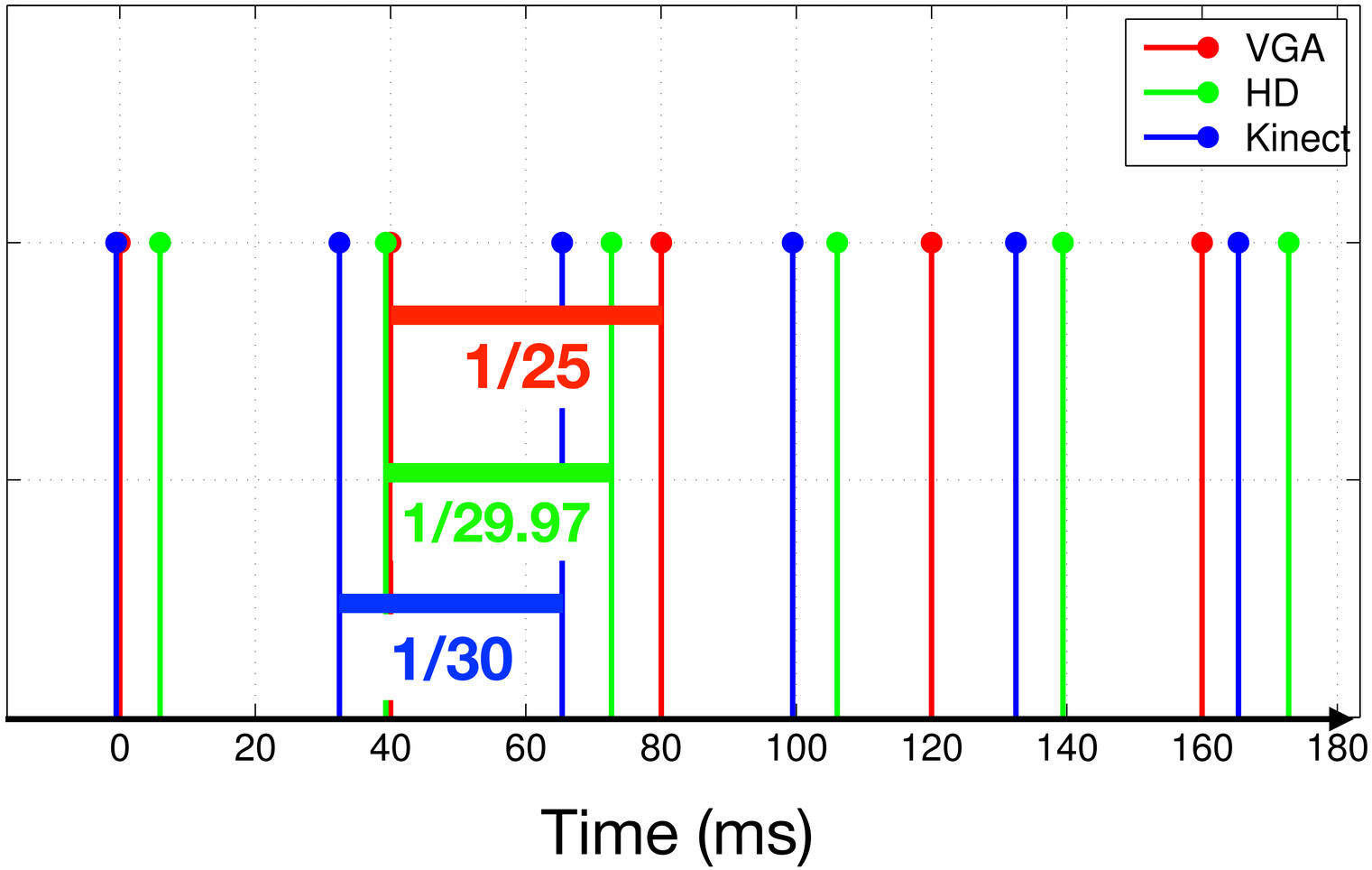


~20M points

Ten

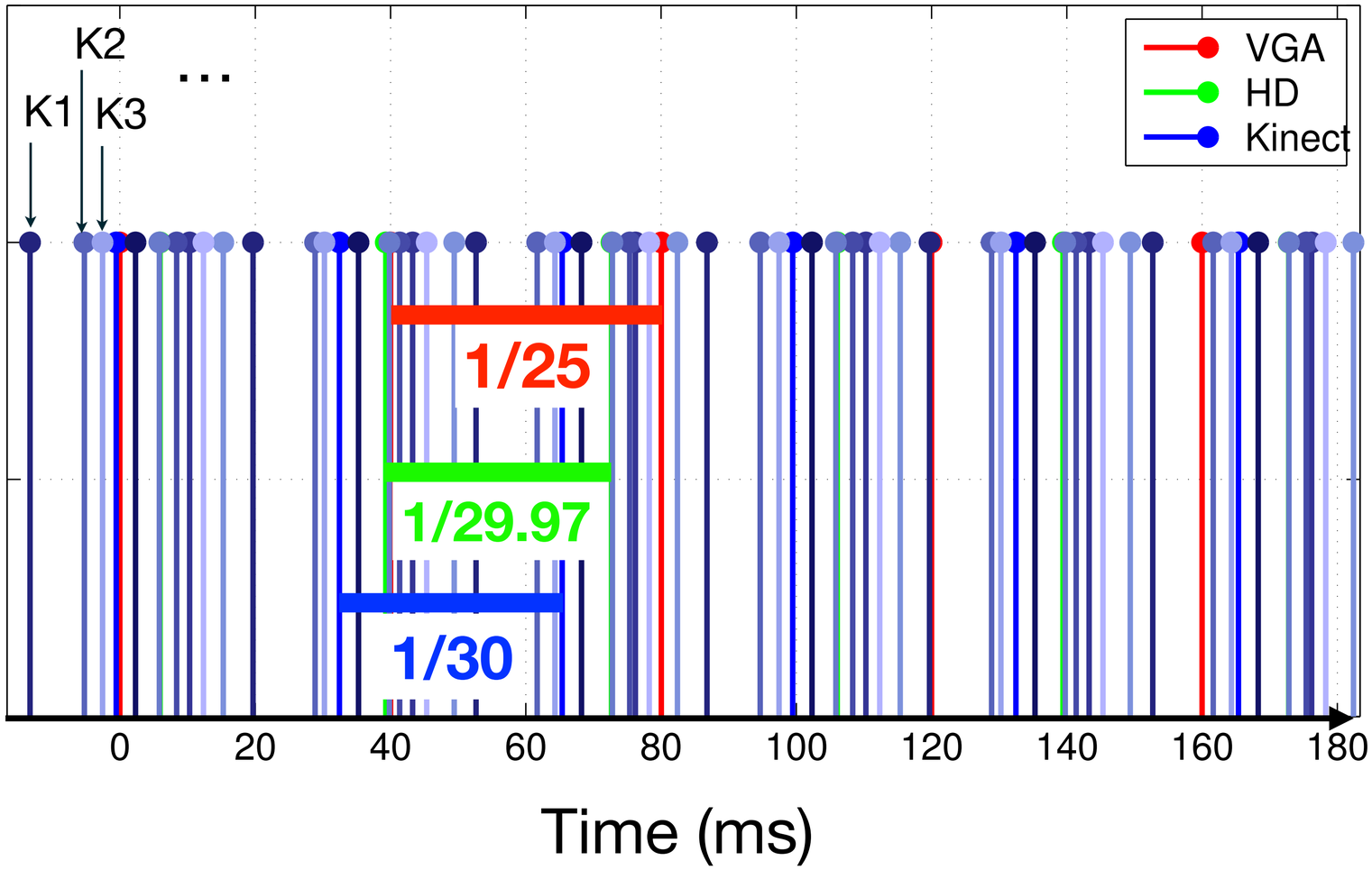
Synchronizing Diverse Sensors

Set Signals on a Common Timebase



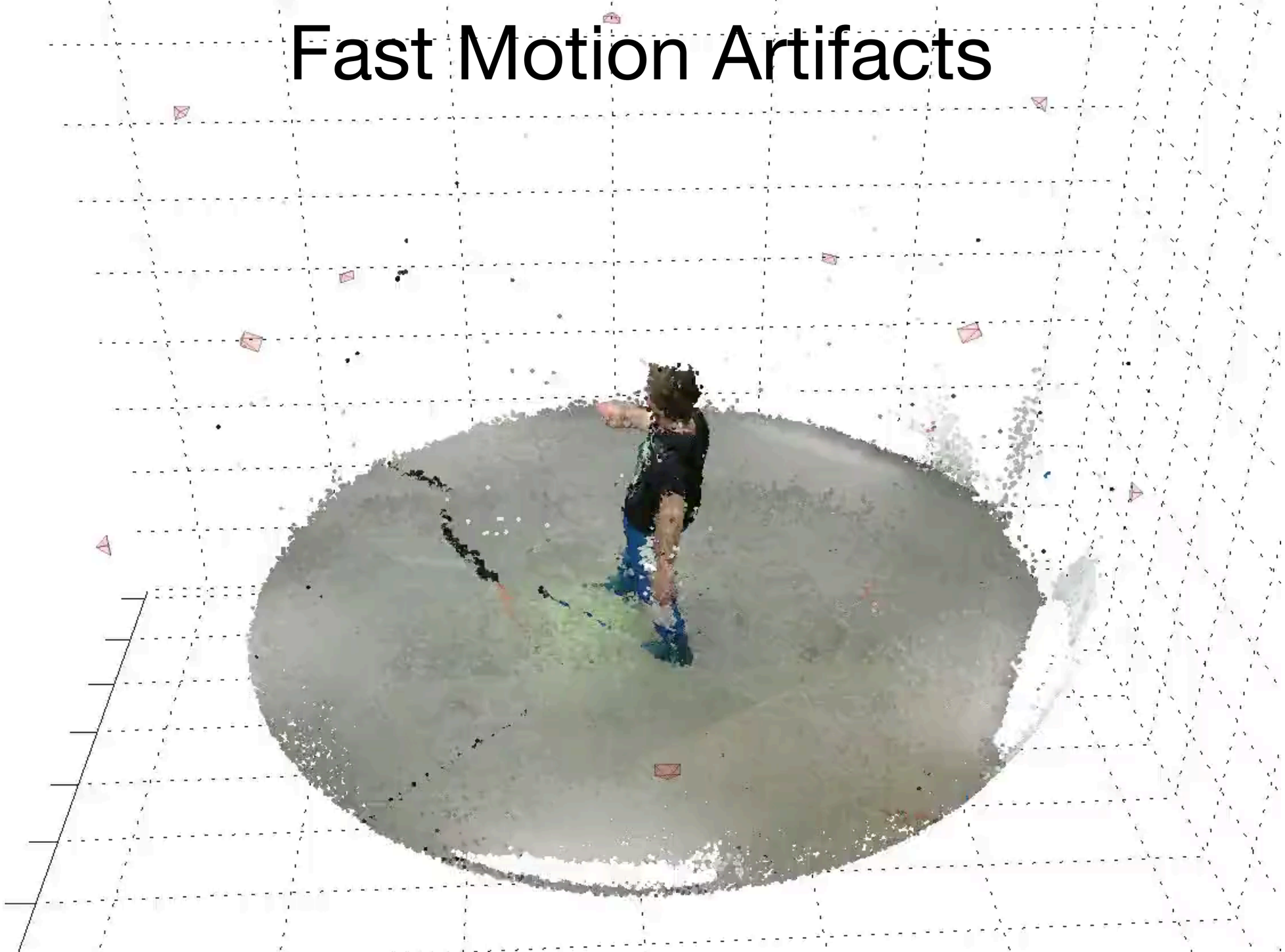
Synchronizing Diverse Sensors

Set Signals on a Common Timebase



Spatio-Temporal Sampling

Fast Motion Artifacts

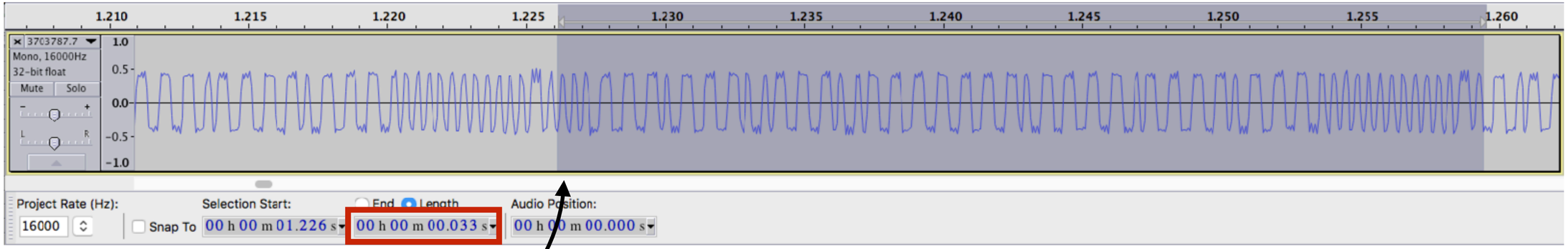



TESTED OK

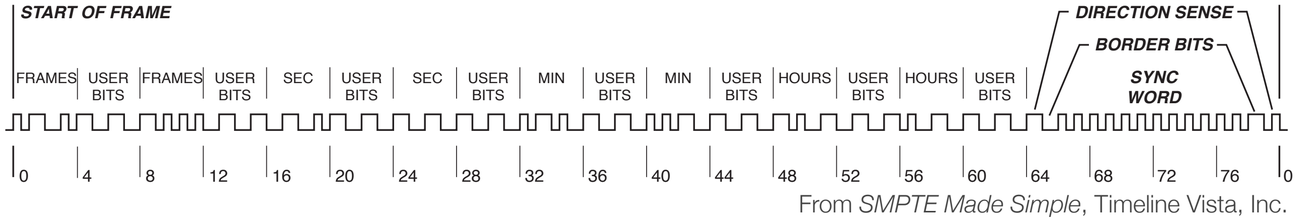

Kinect_Breakout Rev 1.1
P53168
©Tim Godisart

LTC Timecode

Decoding



From production cameras 1/29.97



From *SMPTE Made Simple*, Timeline Vista, Inc.

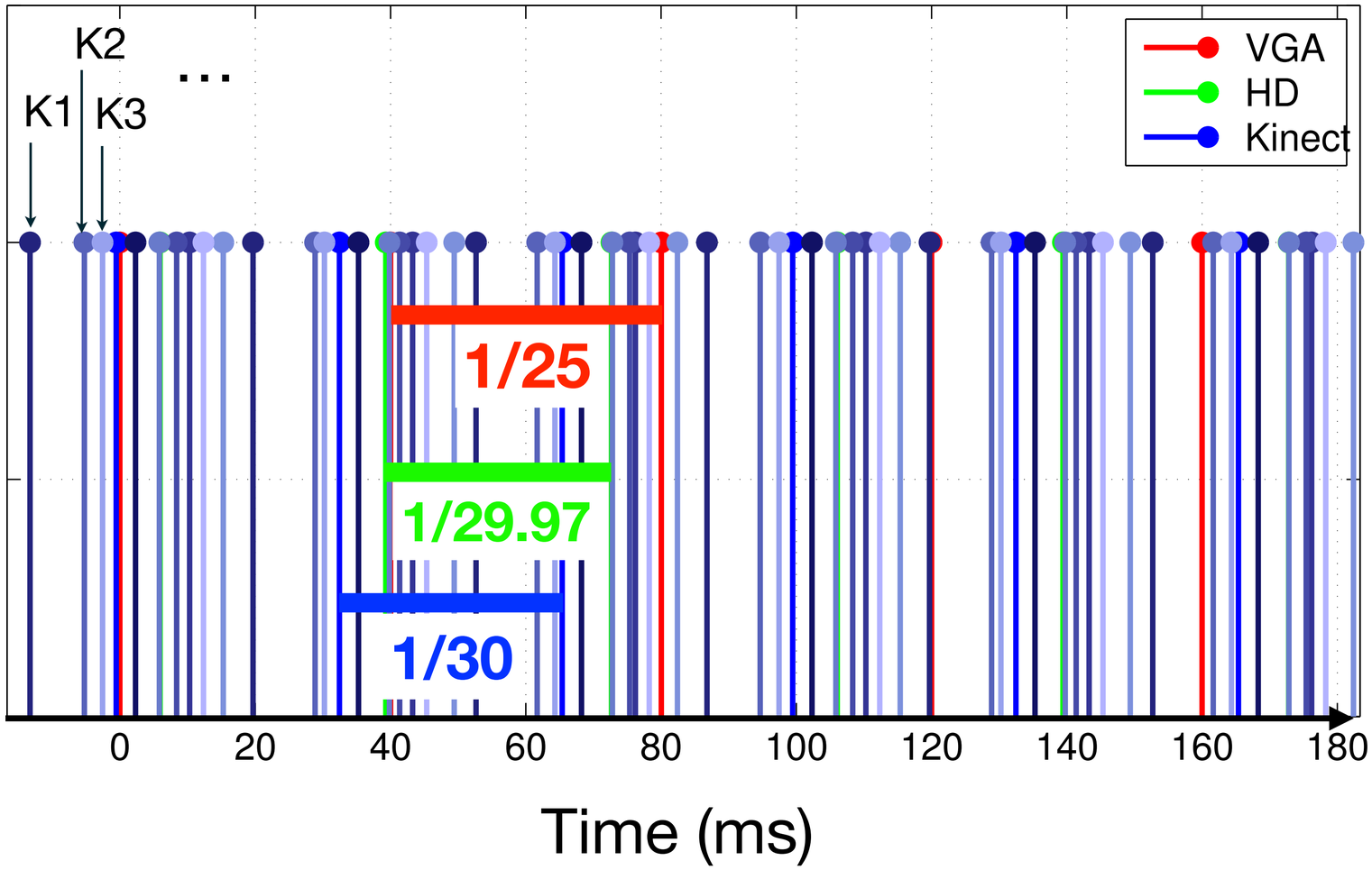
Starting sample

```
04:10:54:28
HH:MM:SS:FF = frame number
```

SMPTE standard, many decoders exist:
http://x42.github.io/libltc/

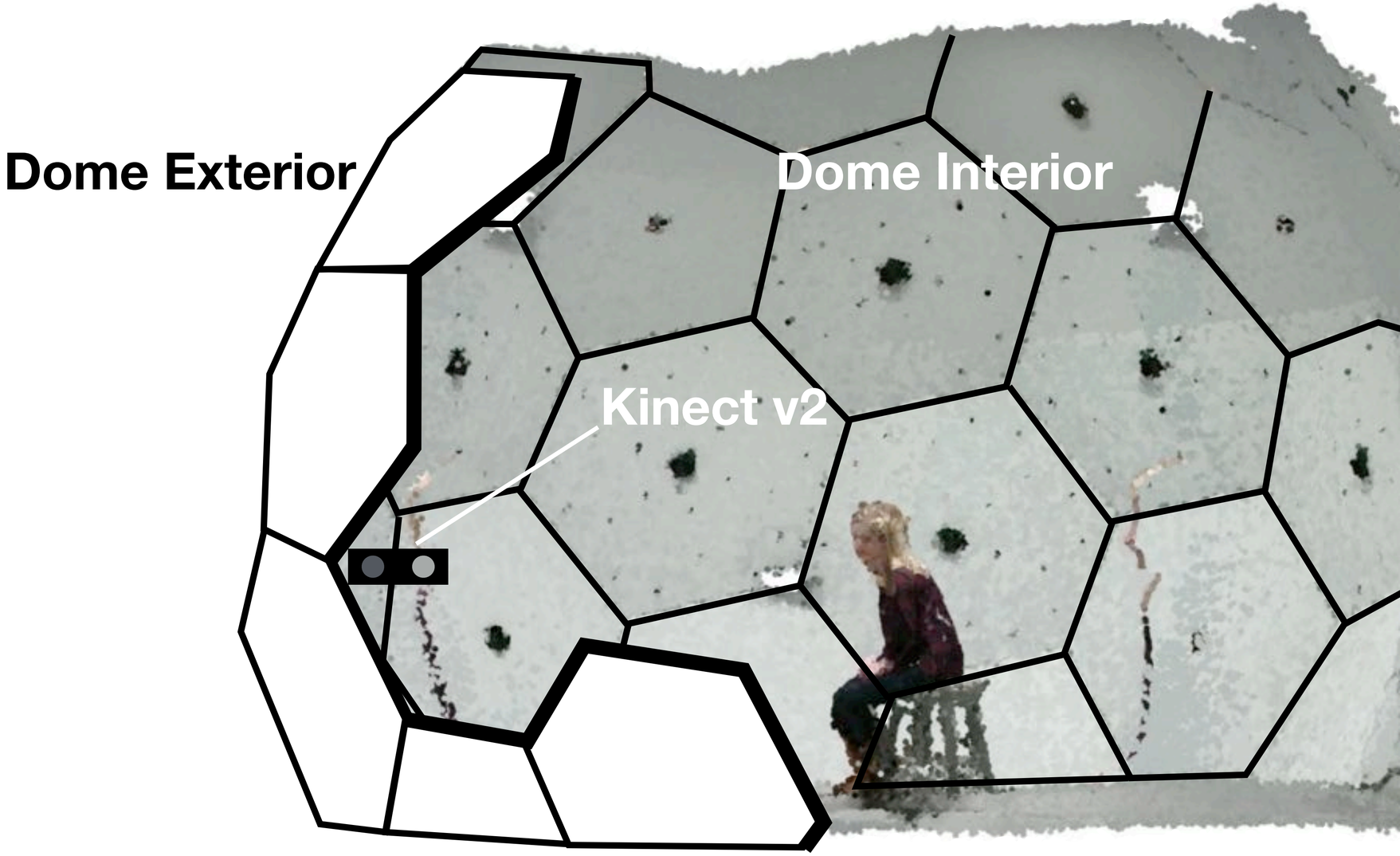
Synchronizing Diverse Sensors

Set Signals on a Common Timebase



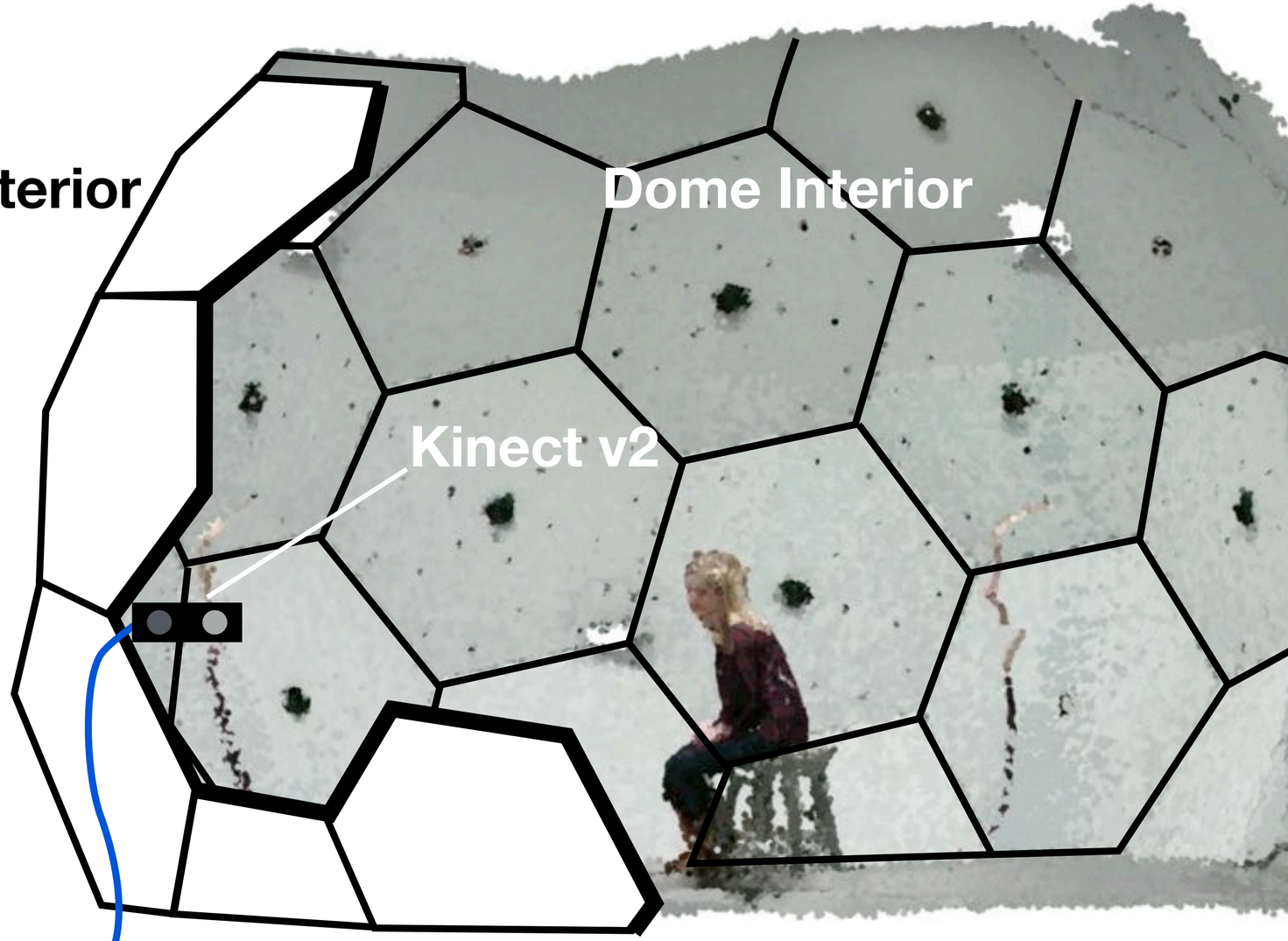
Kinect Capture System

Kinect Capture System



Kinect Capture System

LTC Timecode



Clock System
(LTC)

00:00

Kinect Capture System

Kinect for Windows Adapter

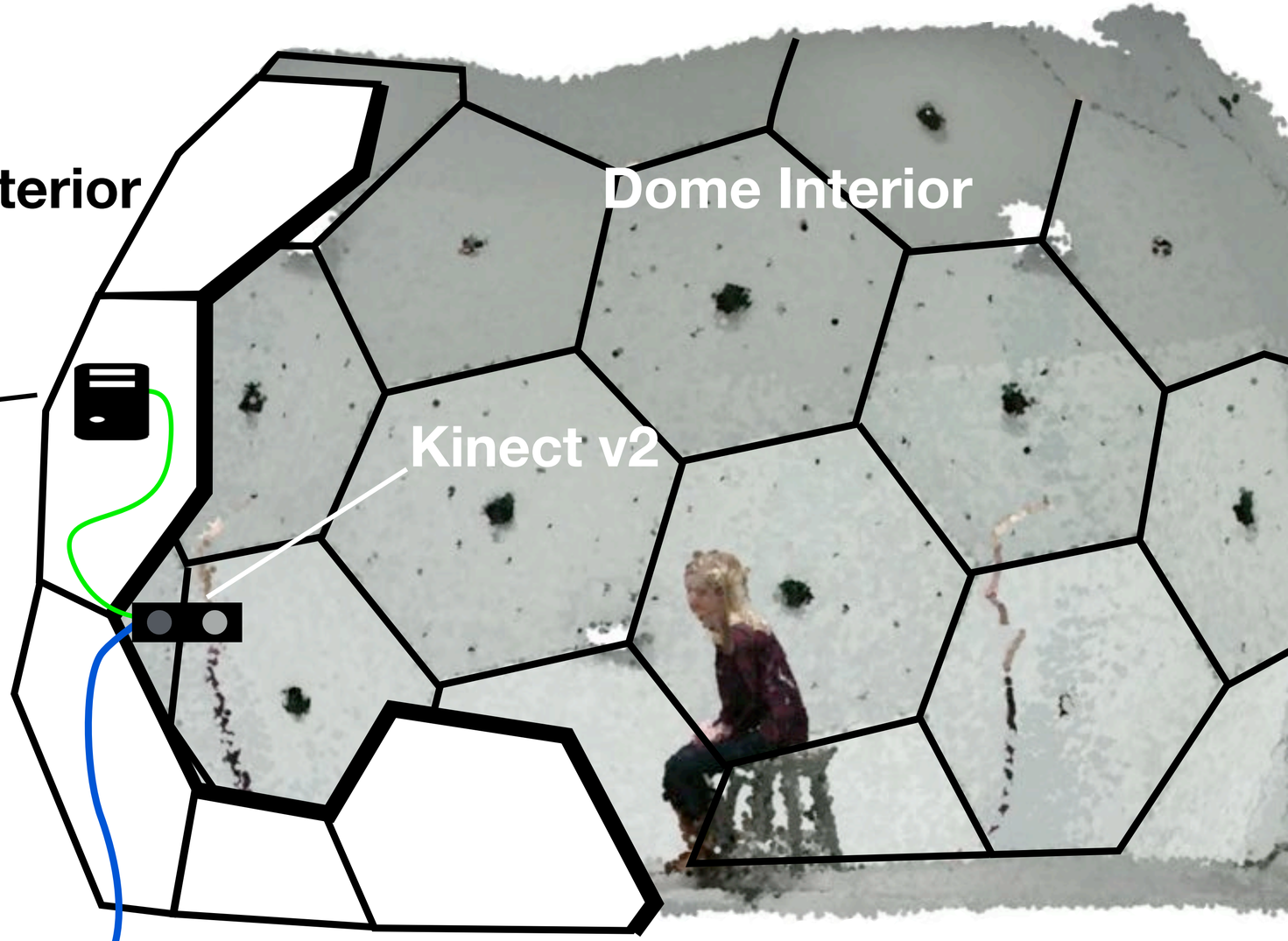
LTC Timecode

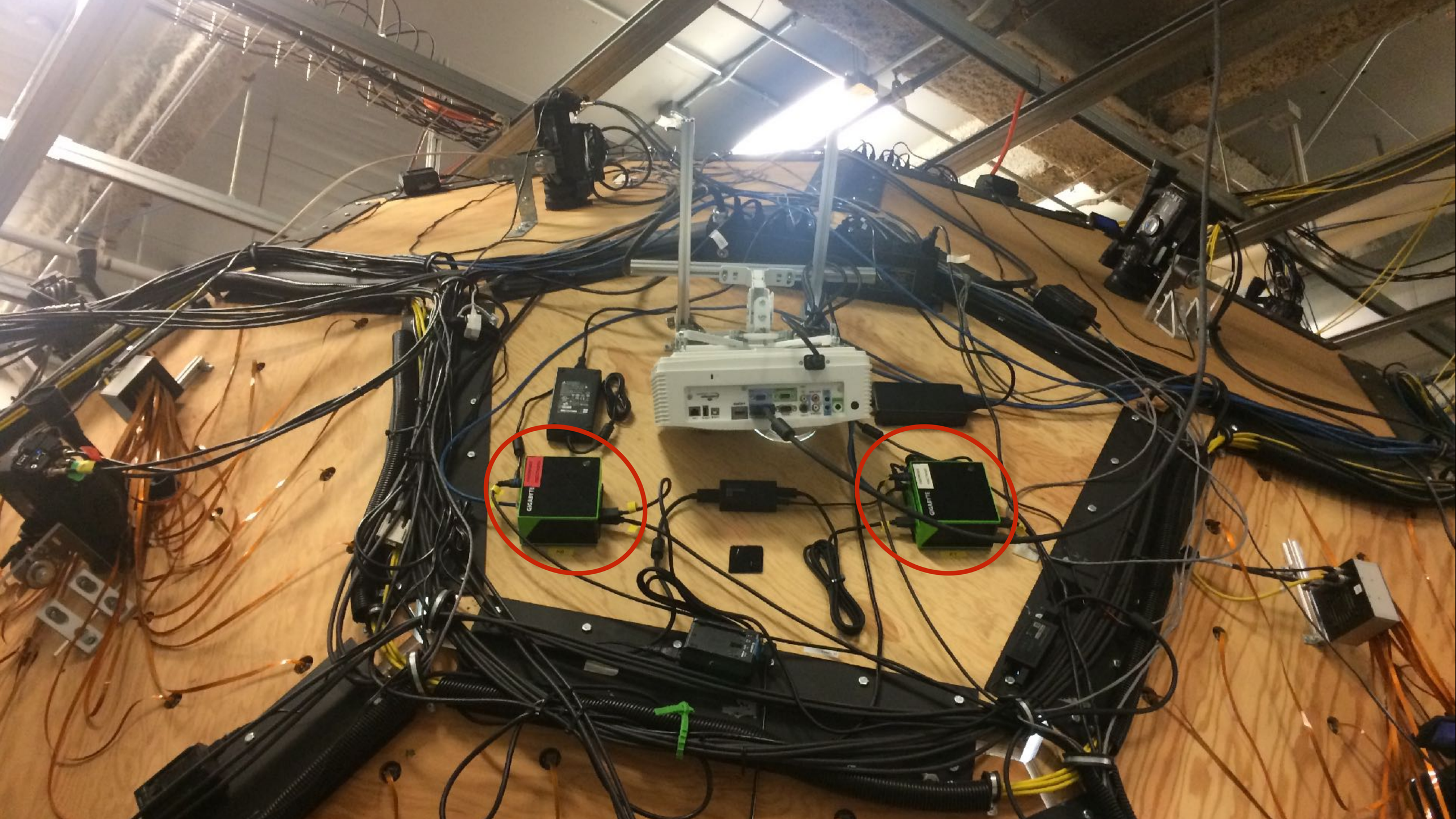
Gigabyte Brix

Capture Node

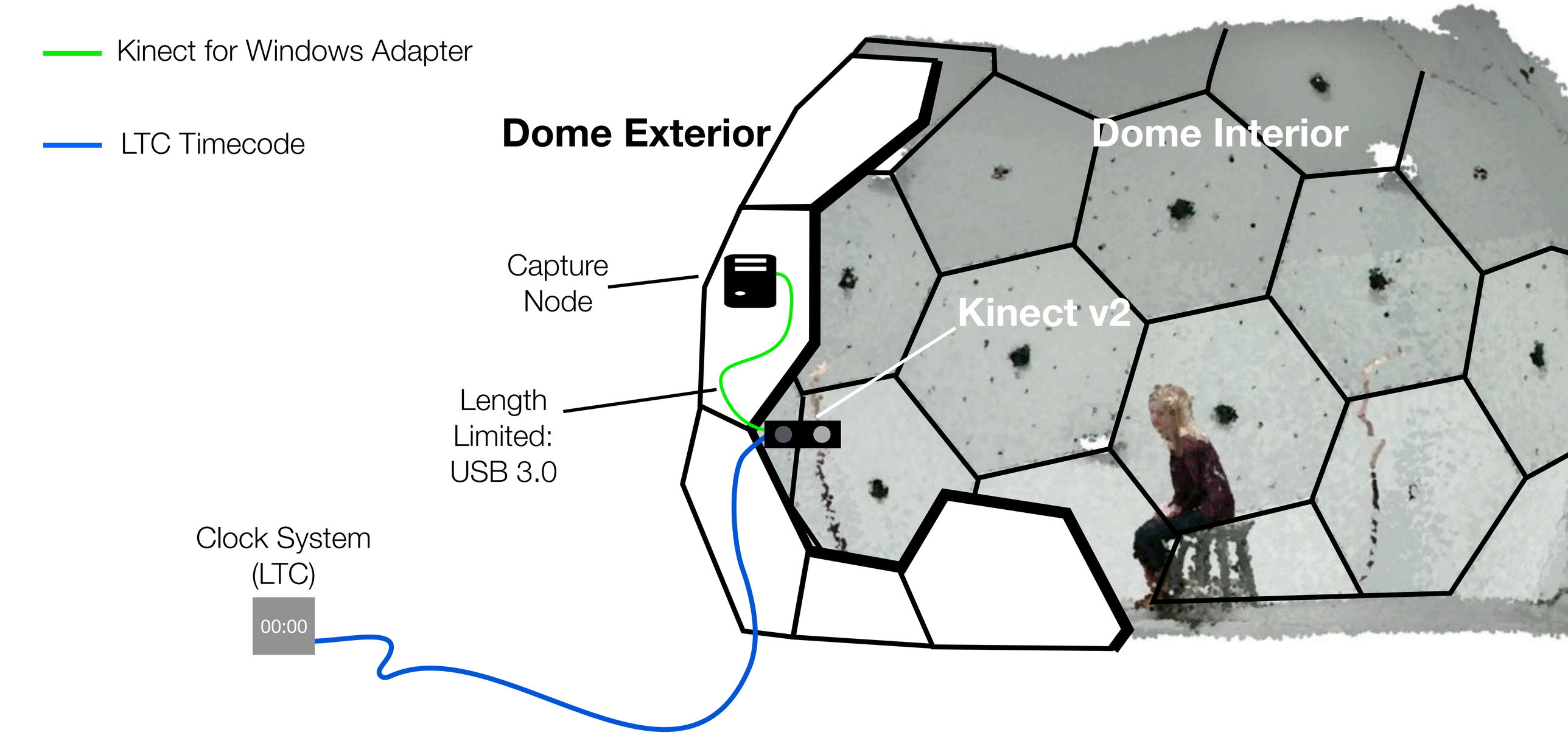
Clock System (LTC)

00:00




GIGABYTE
GIGABYTE

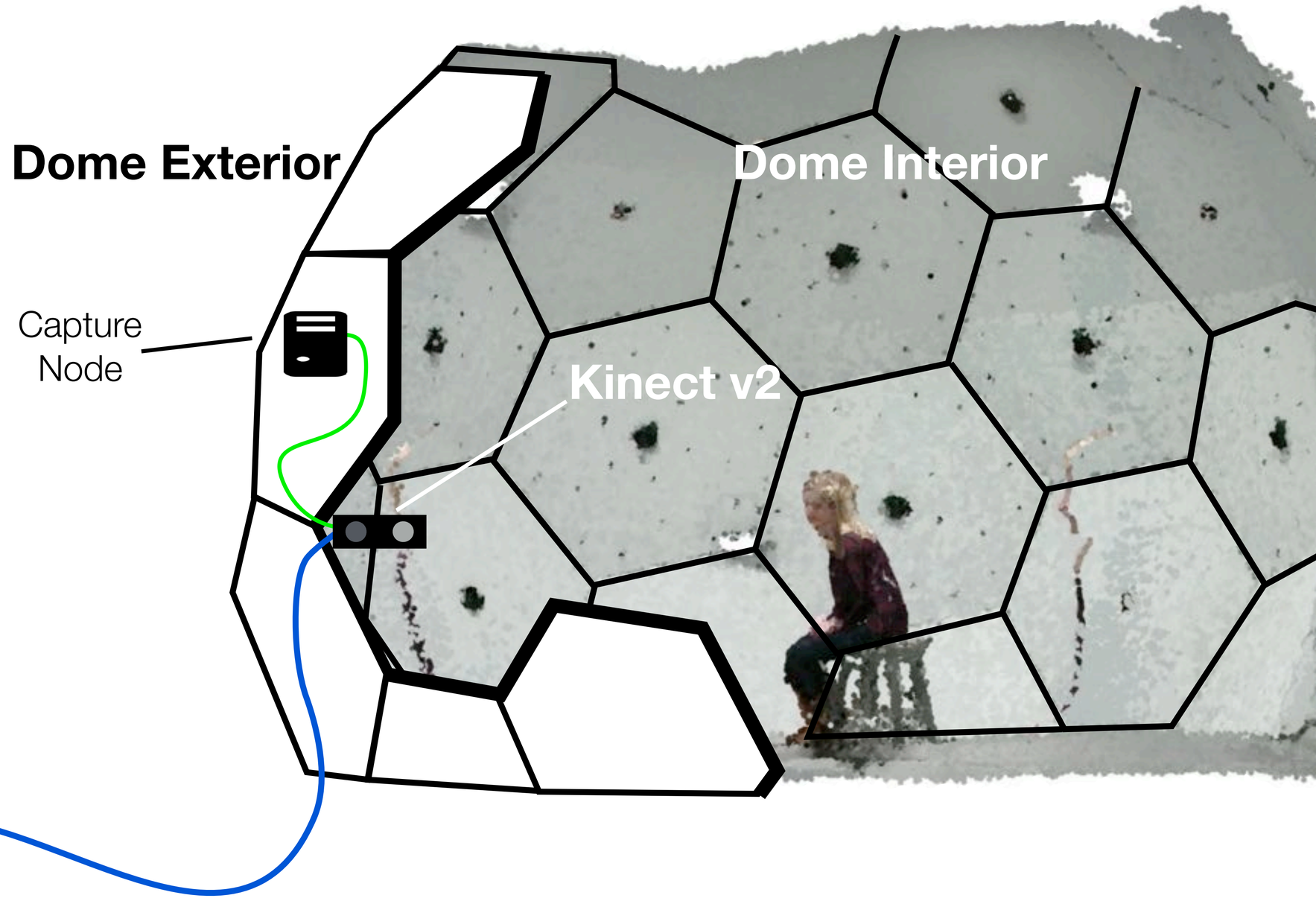
Kinect Capture System



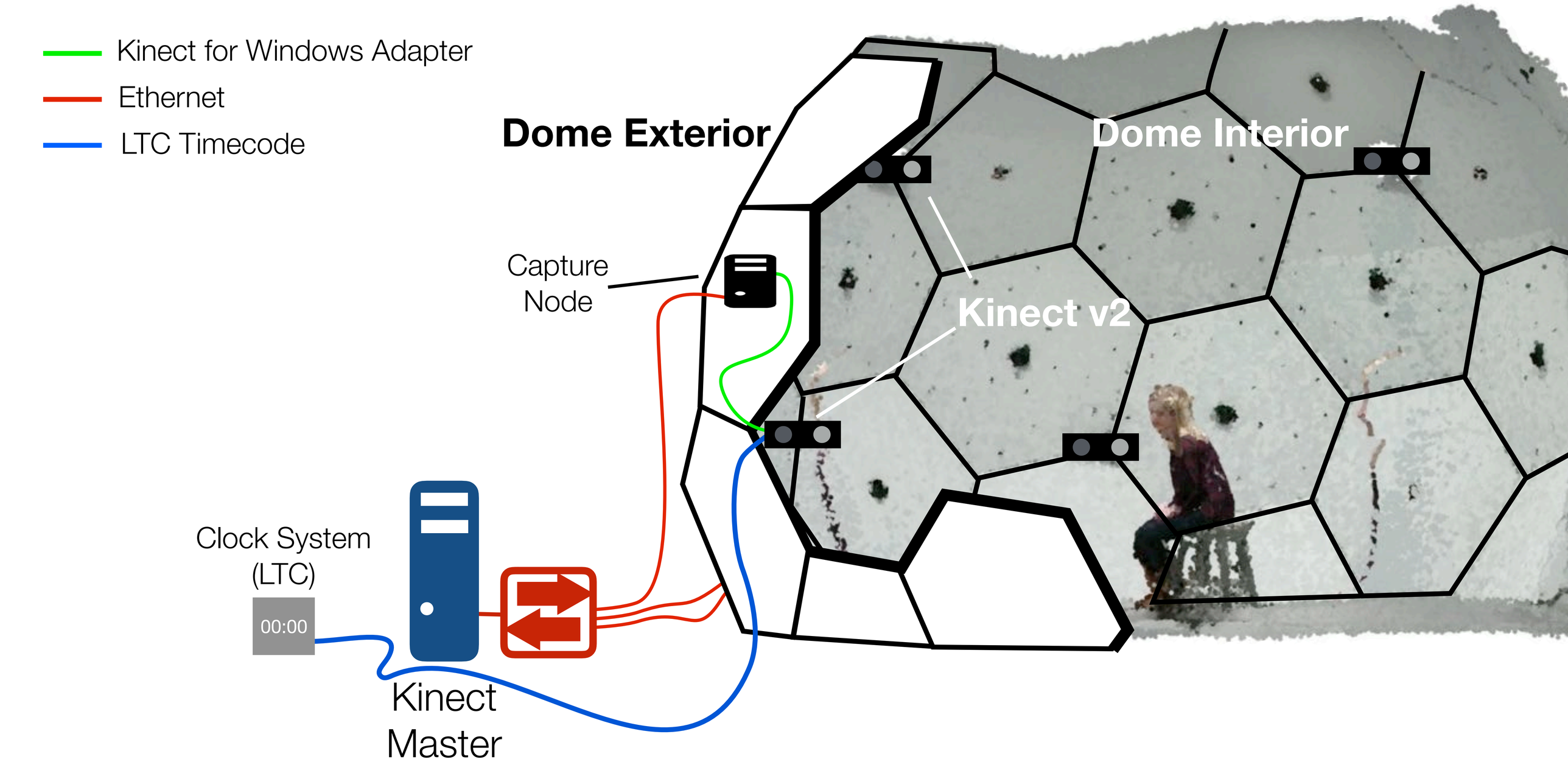
Kinect Capture System

Kinect for Windows Adapter

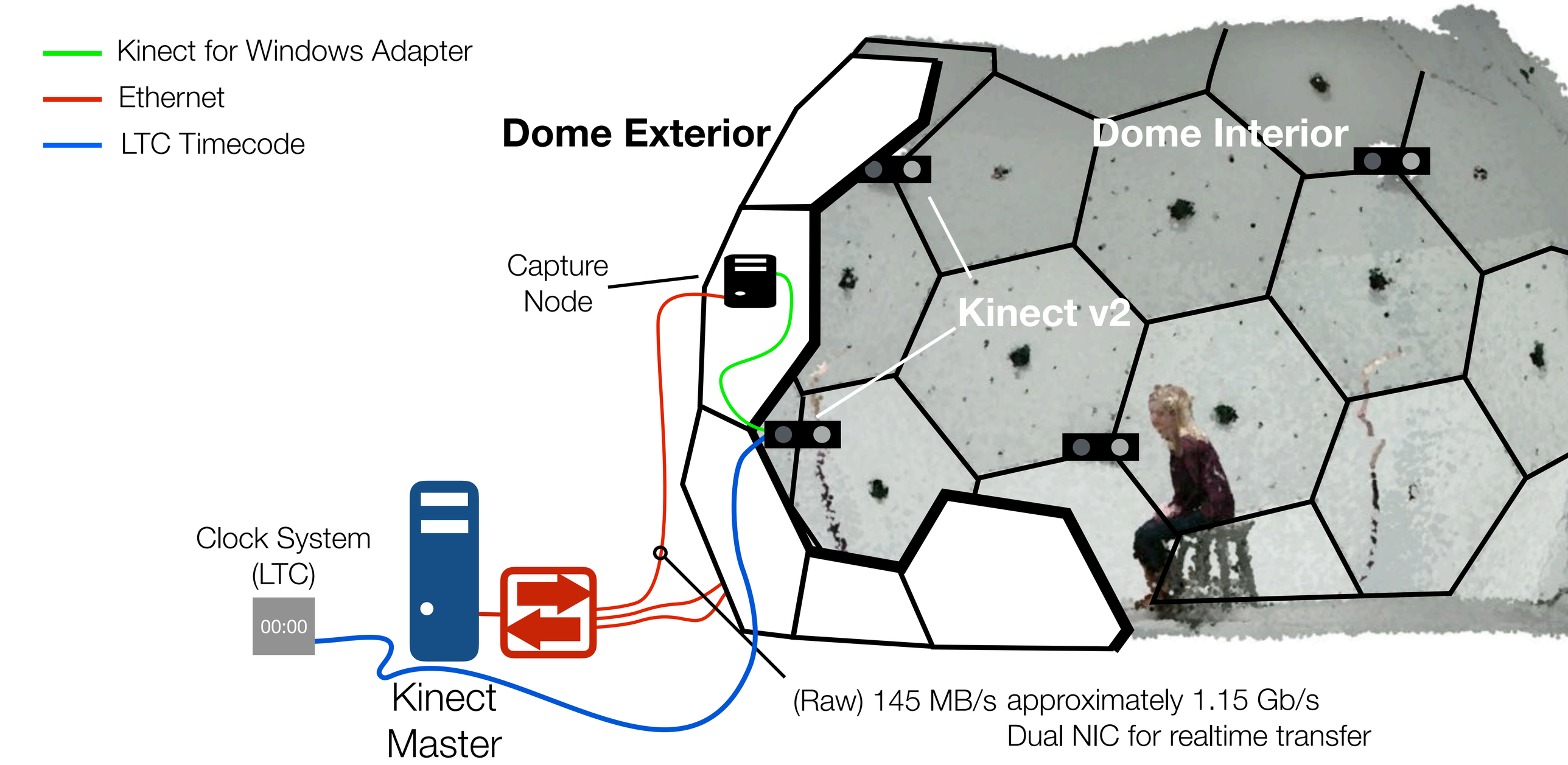
LTC Timecode



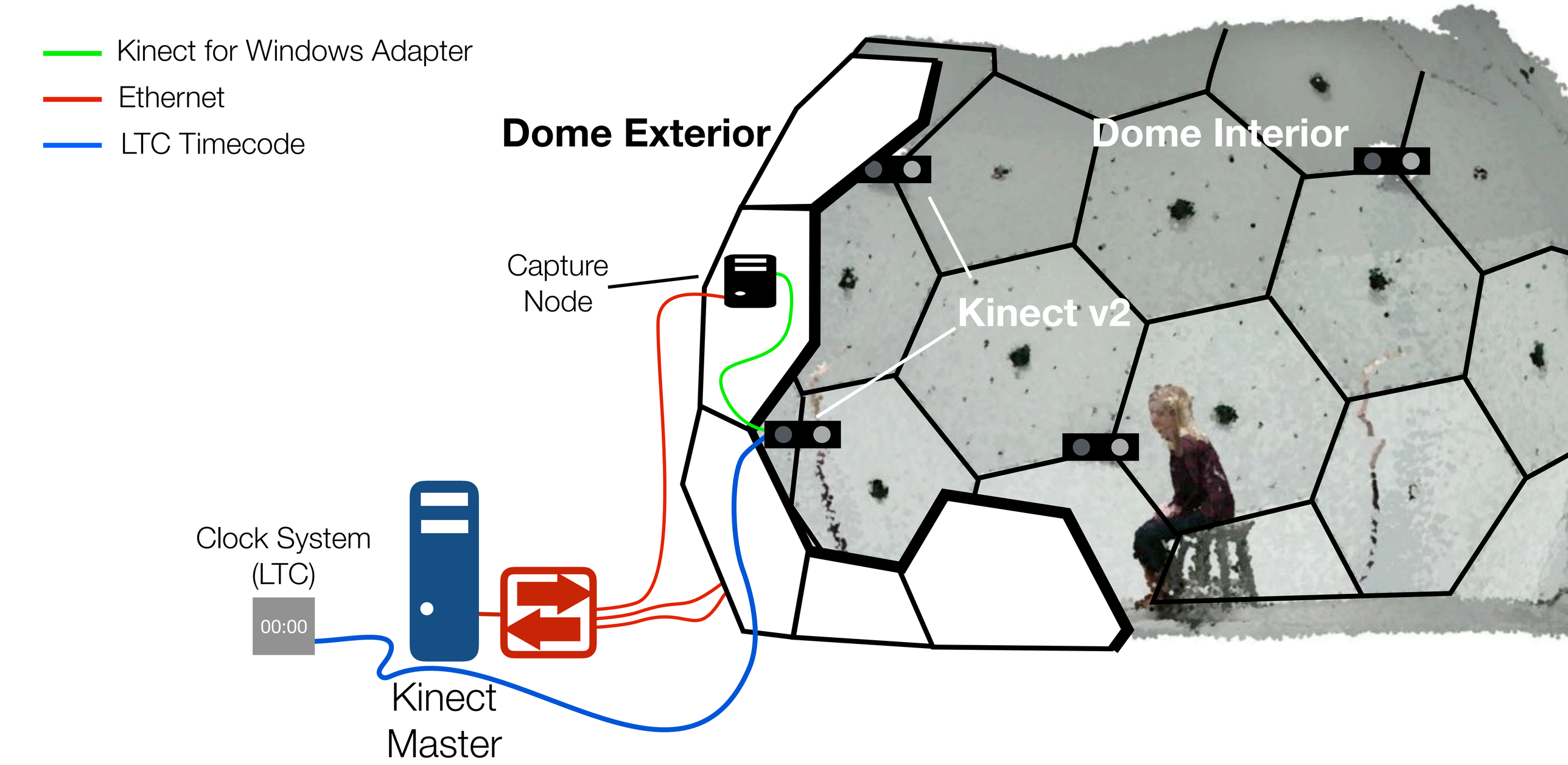
Kinect Capture System



Kinect Capture System

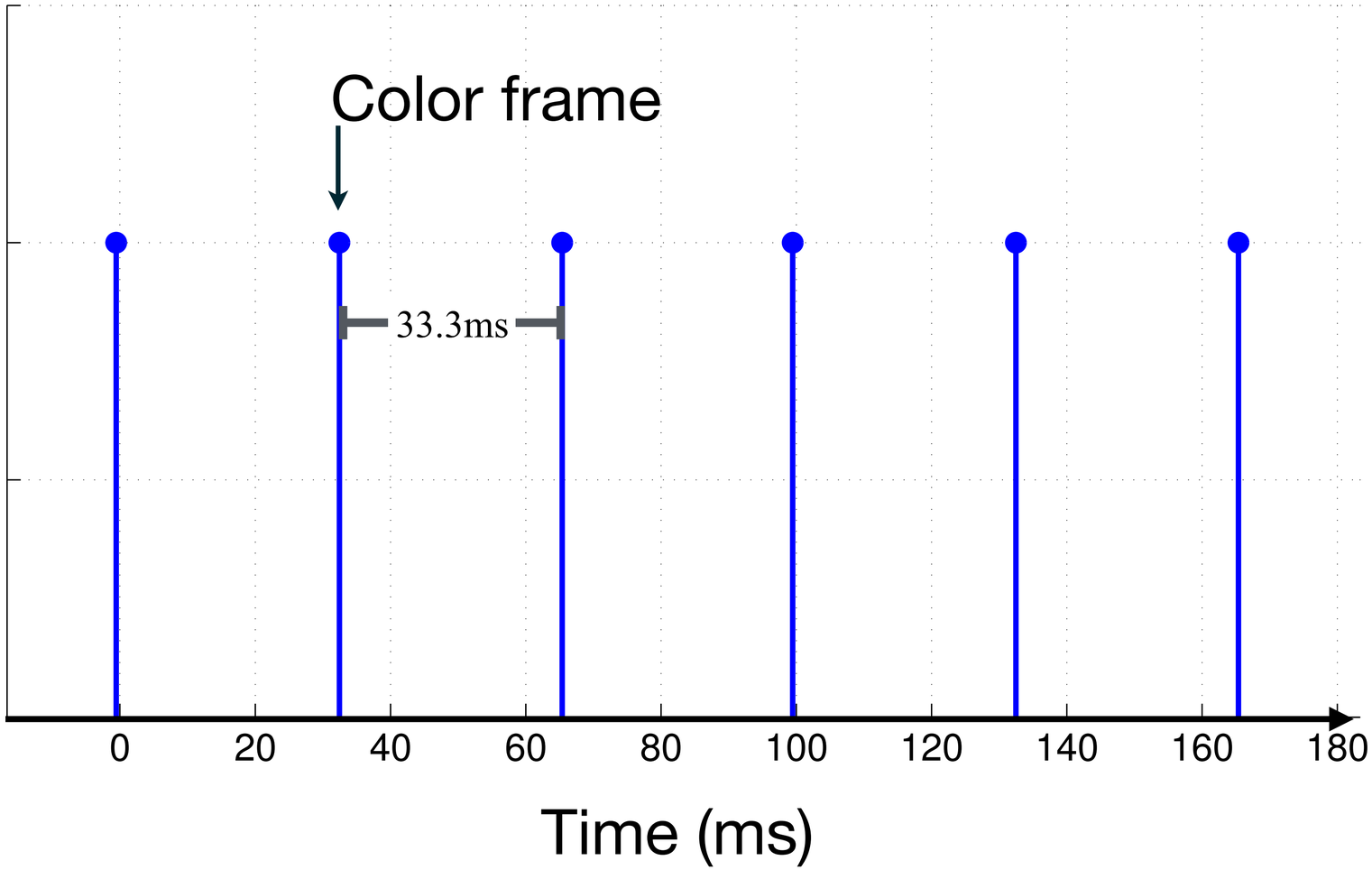


Kinect Capture System



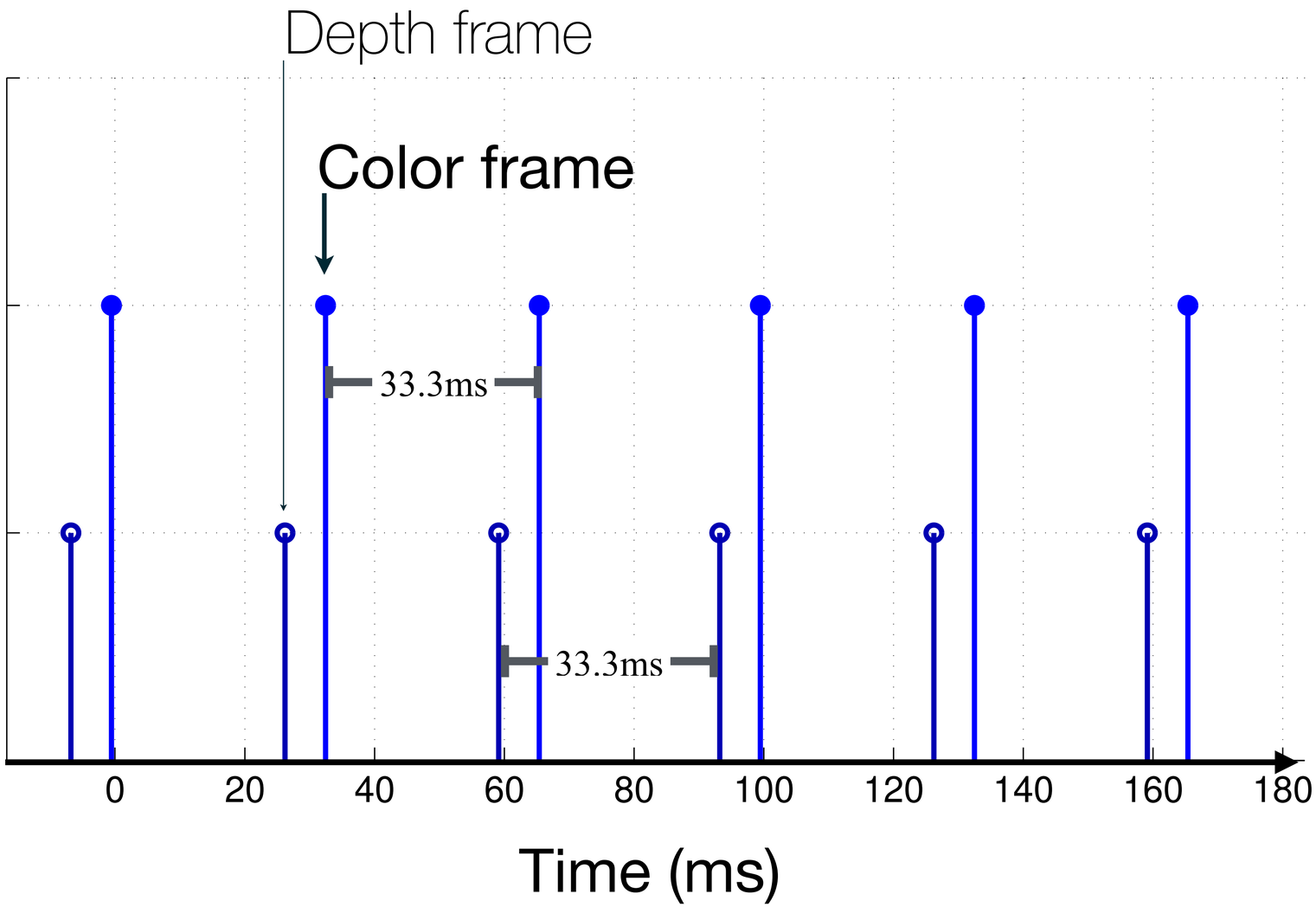
RGB+Depth Relative Timing

What Does Frame Start Mean?



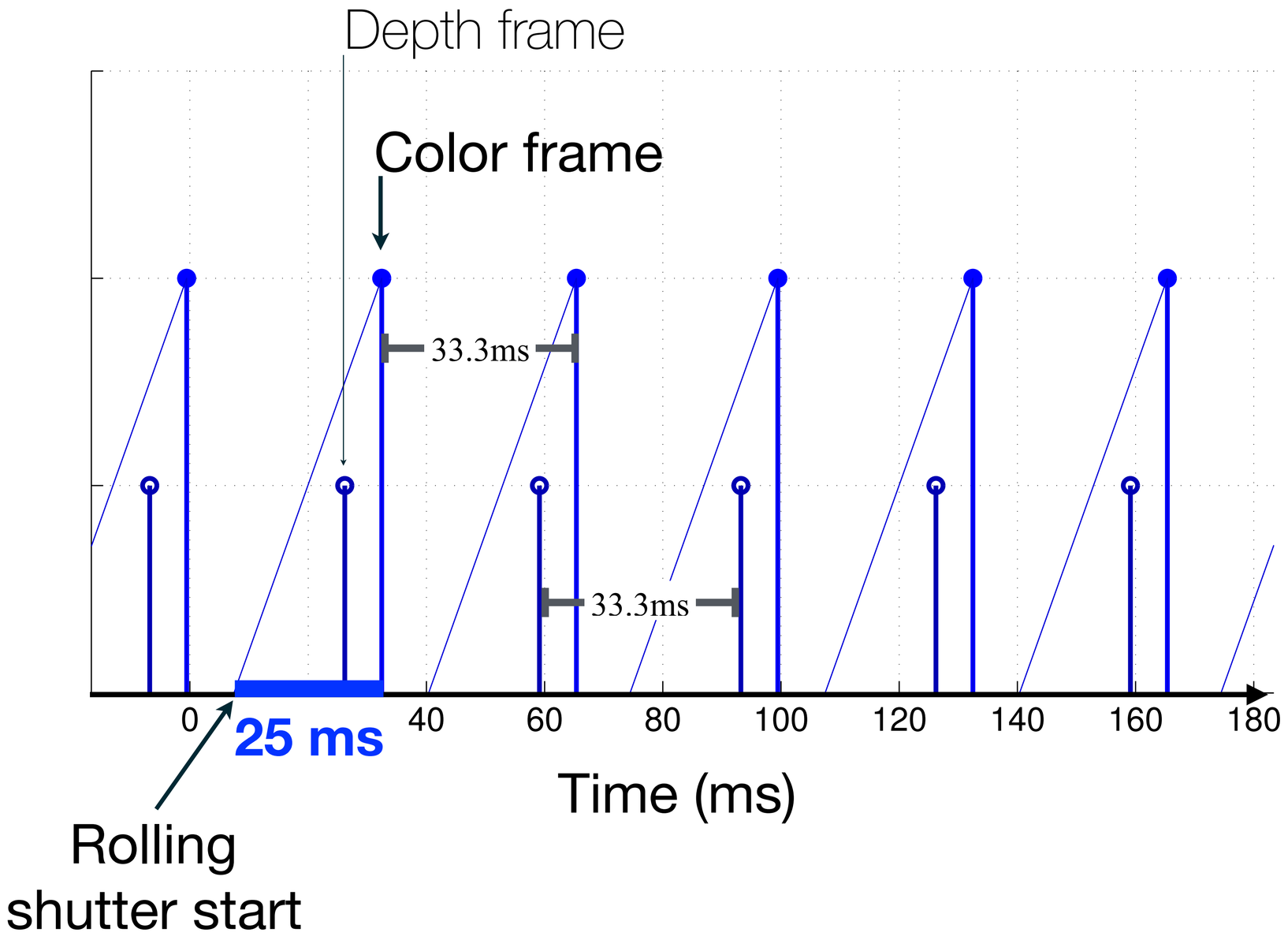
RGB+Depth Relative Timing

What Does Frame Start Mean?

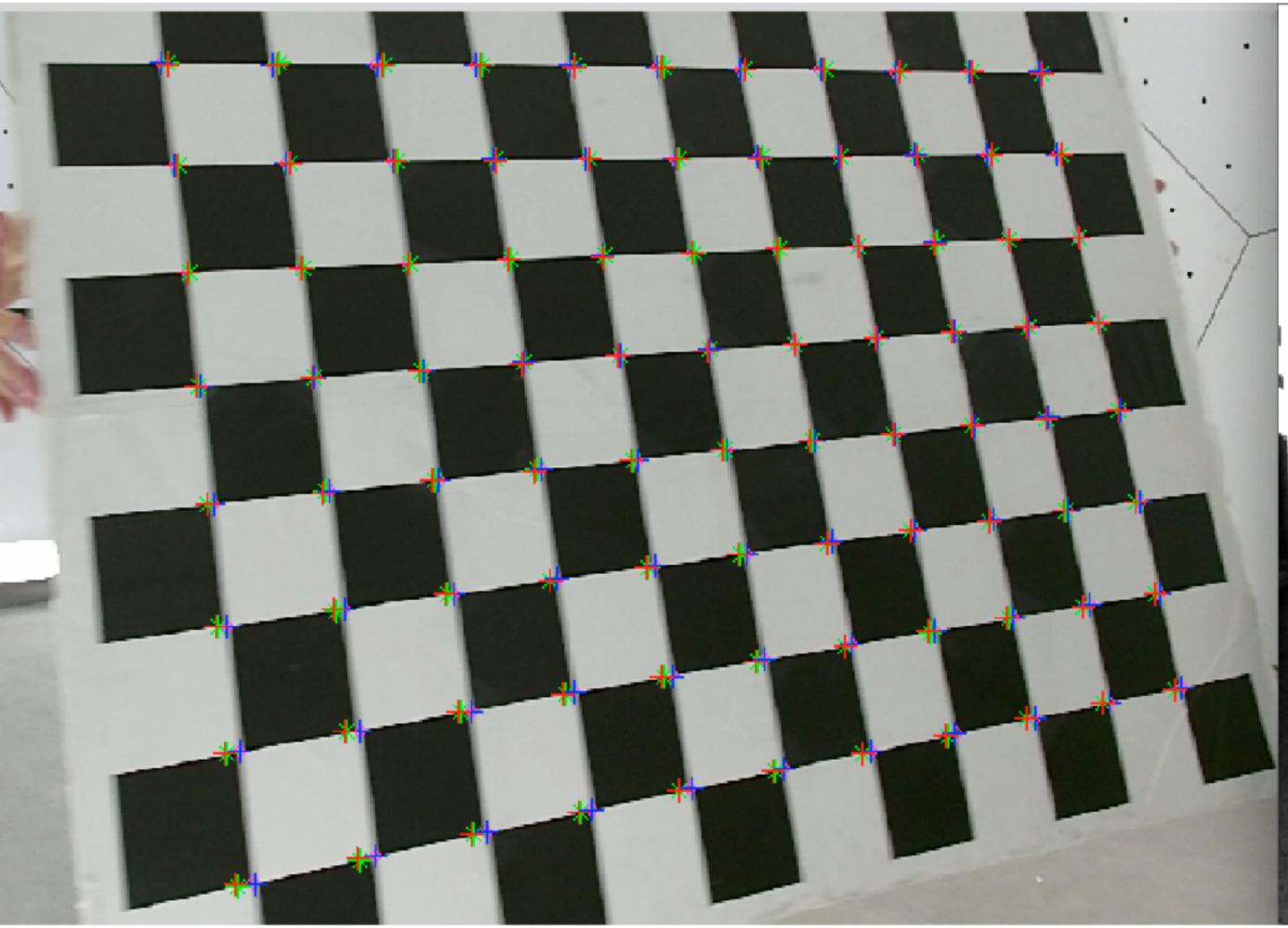


RGB+Depth Relative Timing

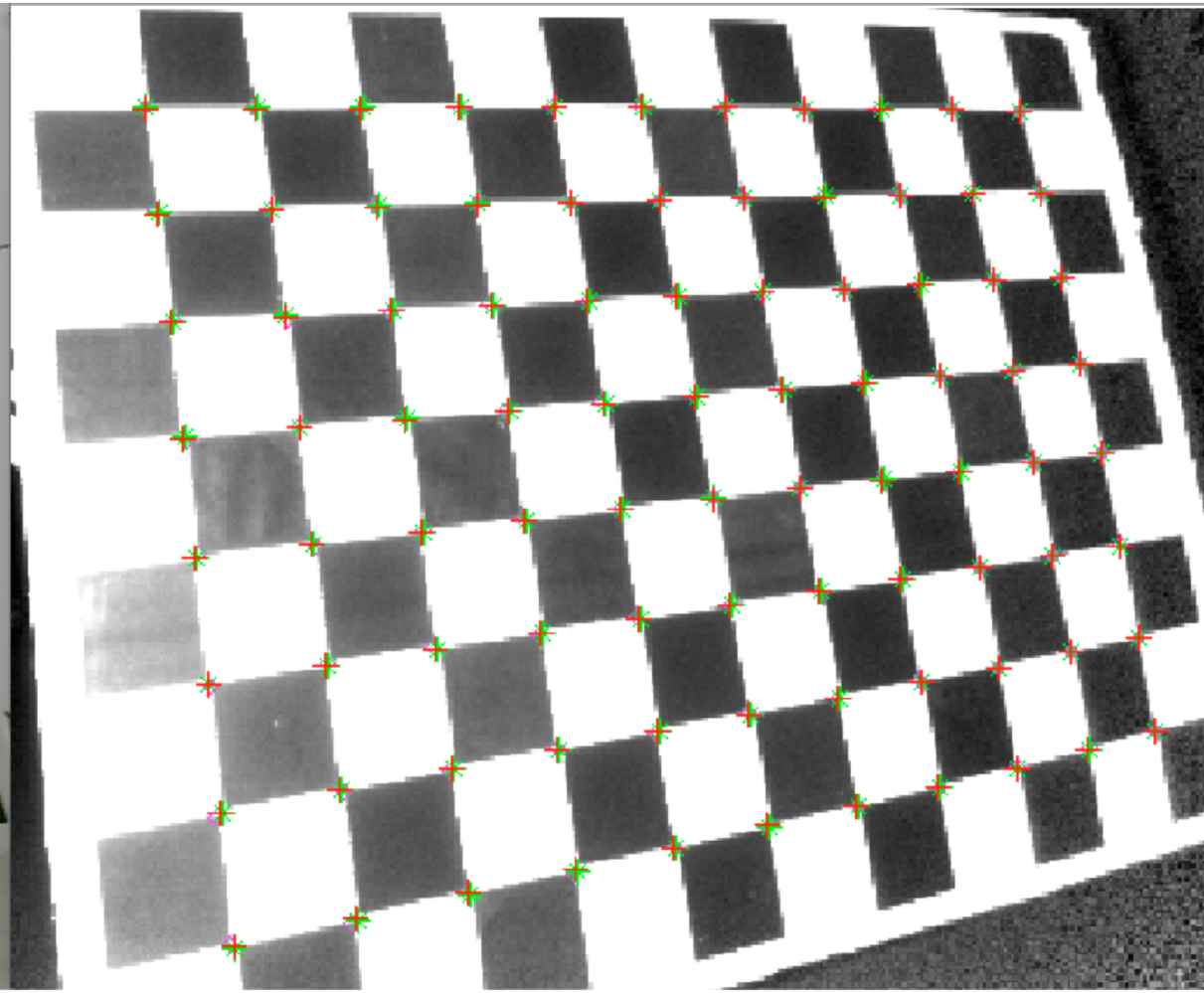
What Does Frame Start Mean?



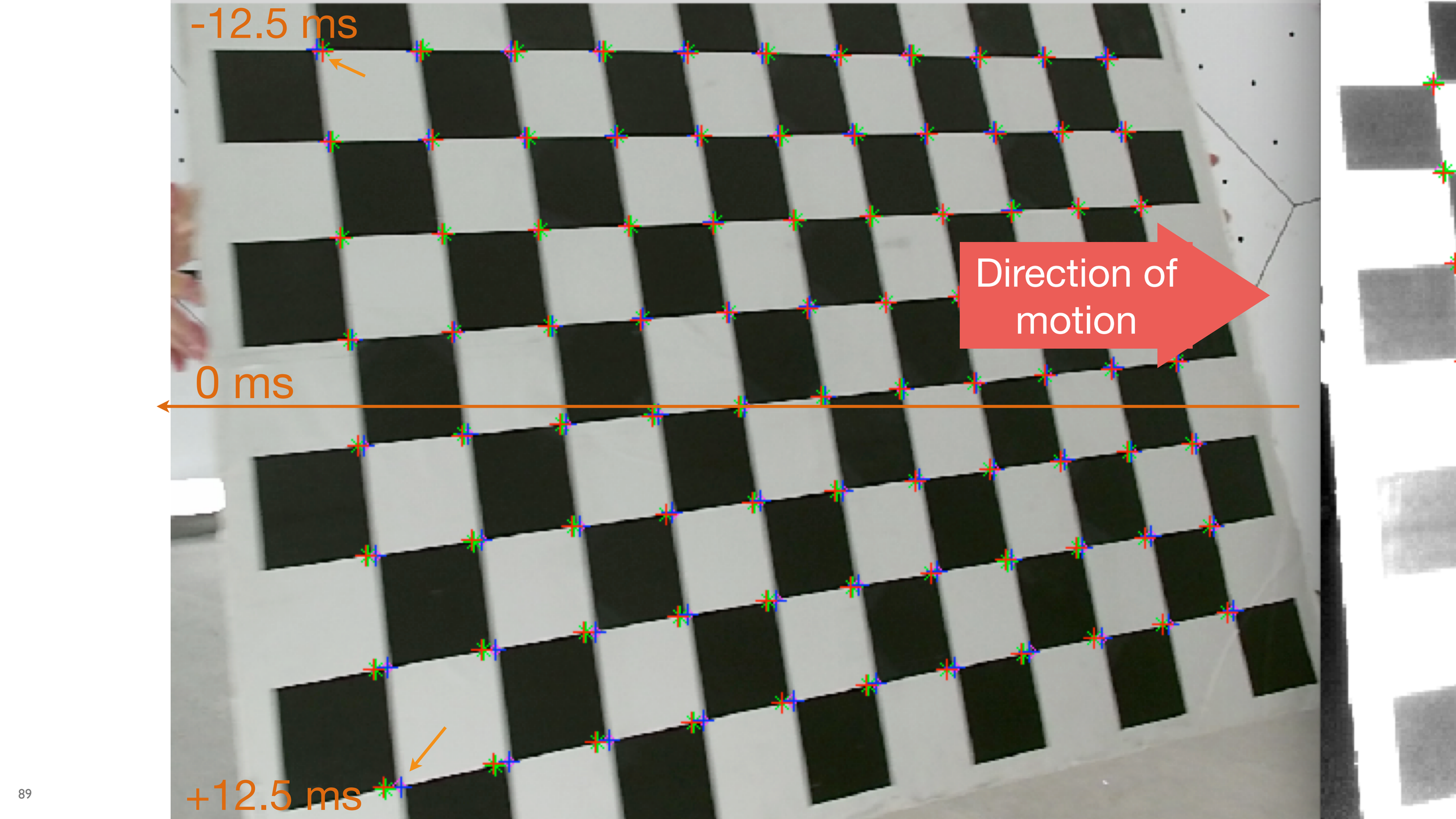
Rolling Shutter Effects

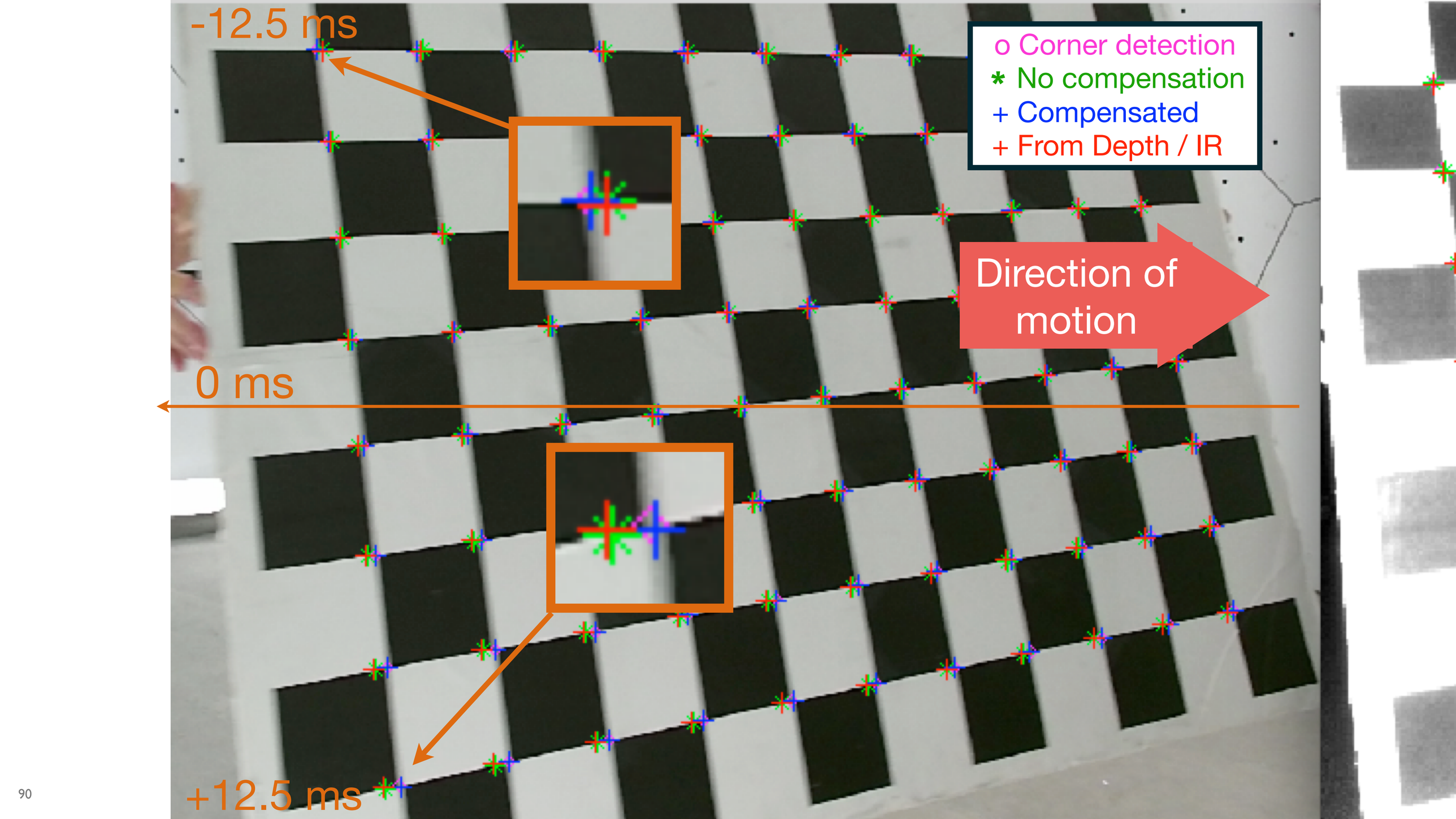


Color

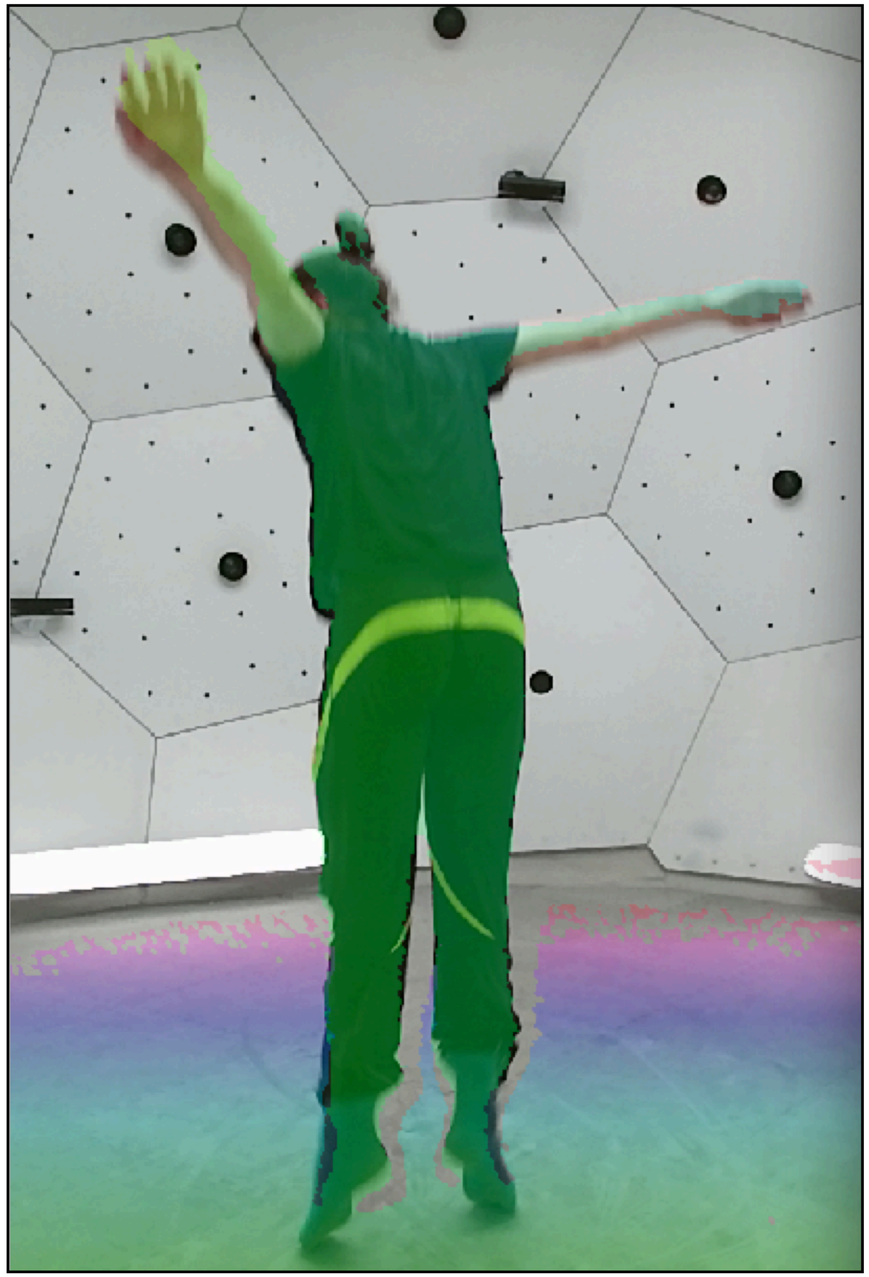


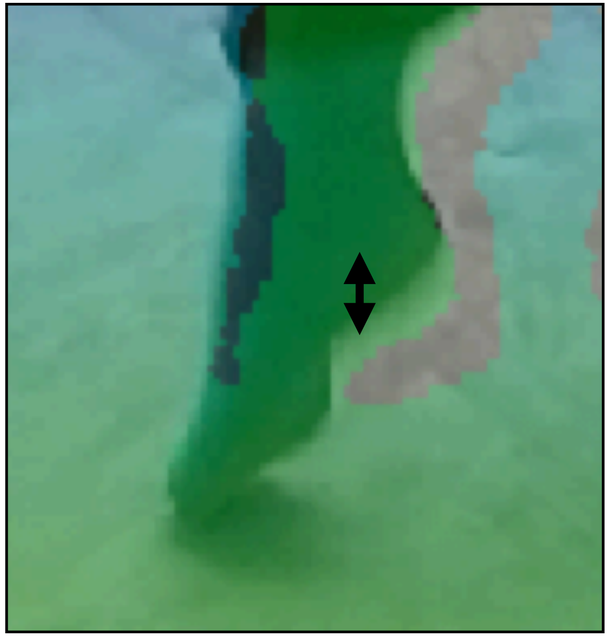
IR


-12.5 ms
Direction of motion
0 ms
+12.5 ms


-12.5 ms
o Corner detection
* No compensation
+ Compensated
+ From Depth / IR
Direction of motion
0 ms
+12.5 ms

Fast Motion Artifacts

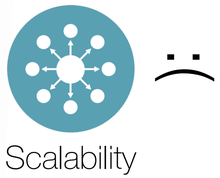




Kinect Synchronization

Take Home Messages

- Off-the-shelf RGBD sensors are cheap, but you'll have to deal with synchronization & interference



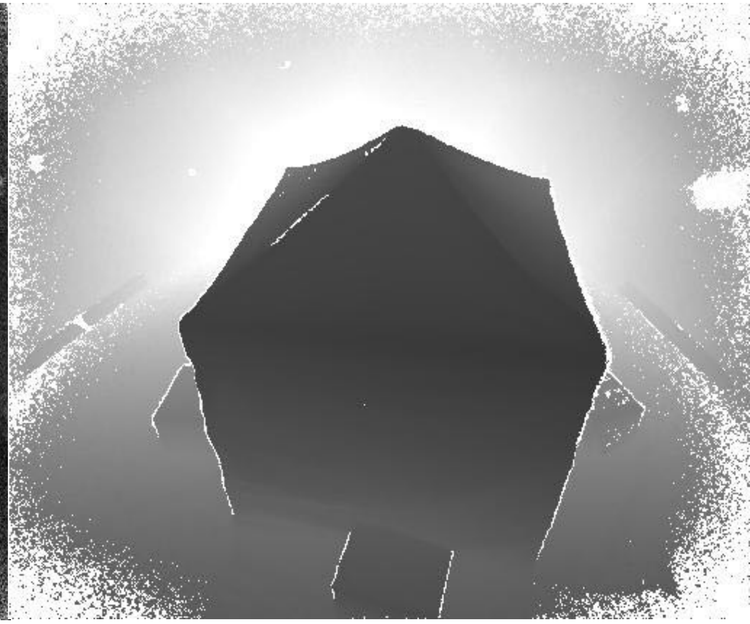
- Modifying hardware is very impractical: PTP as an alternative

- Quick way to get 3D reconstruction

Kinect Calibration

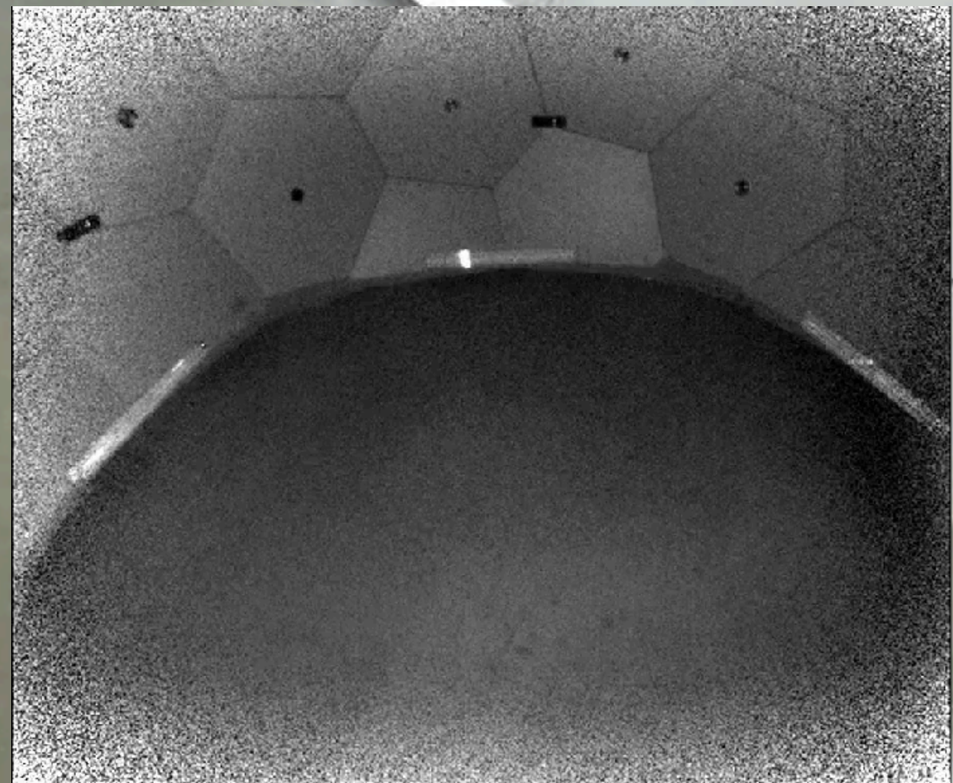
Color

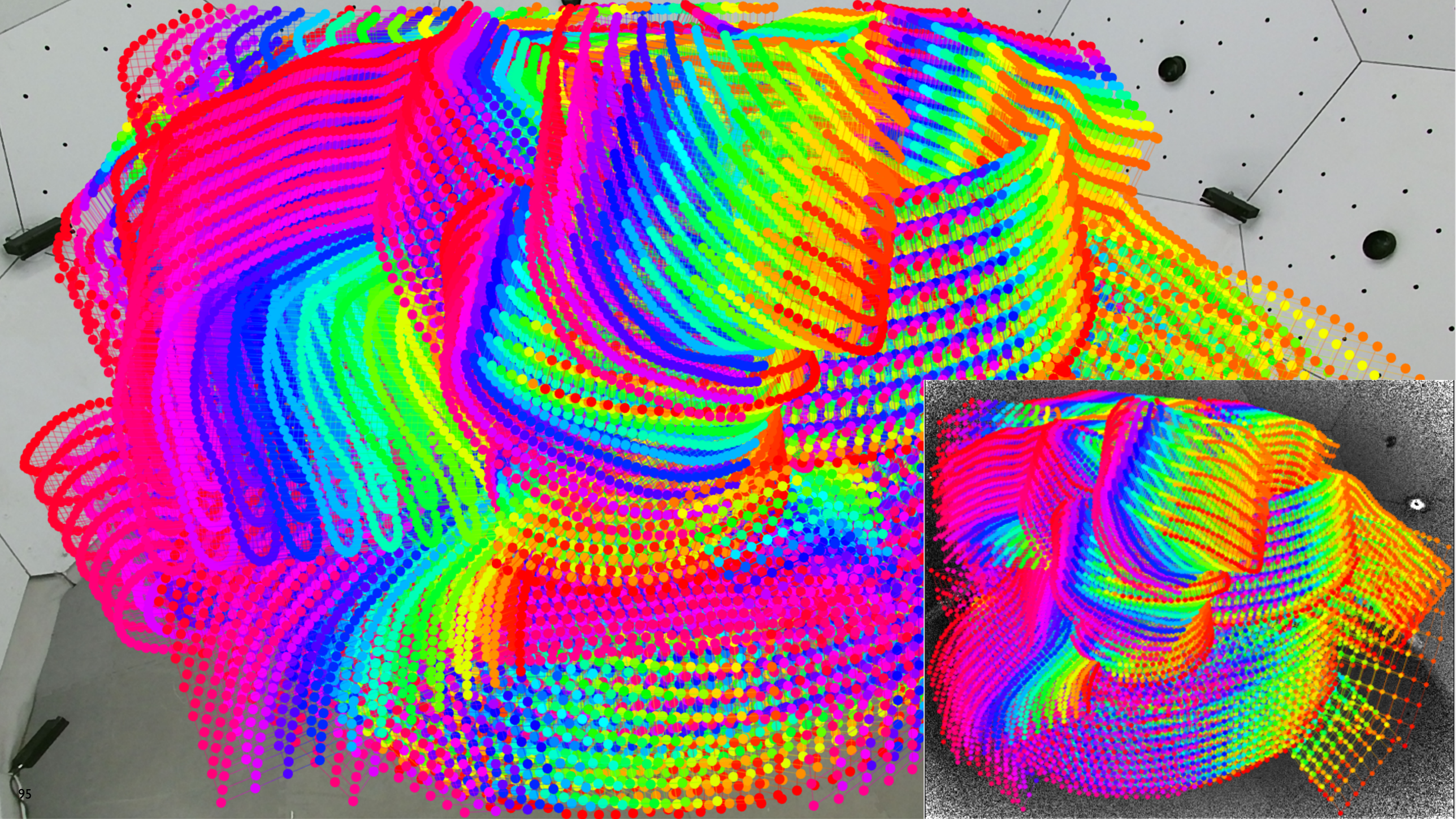
IR

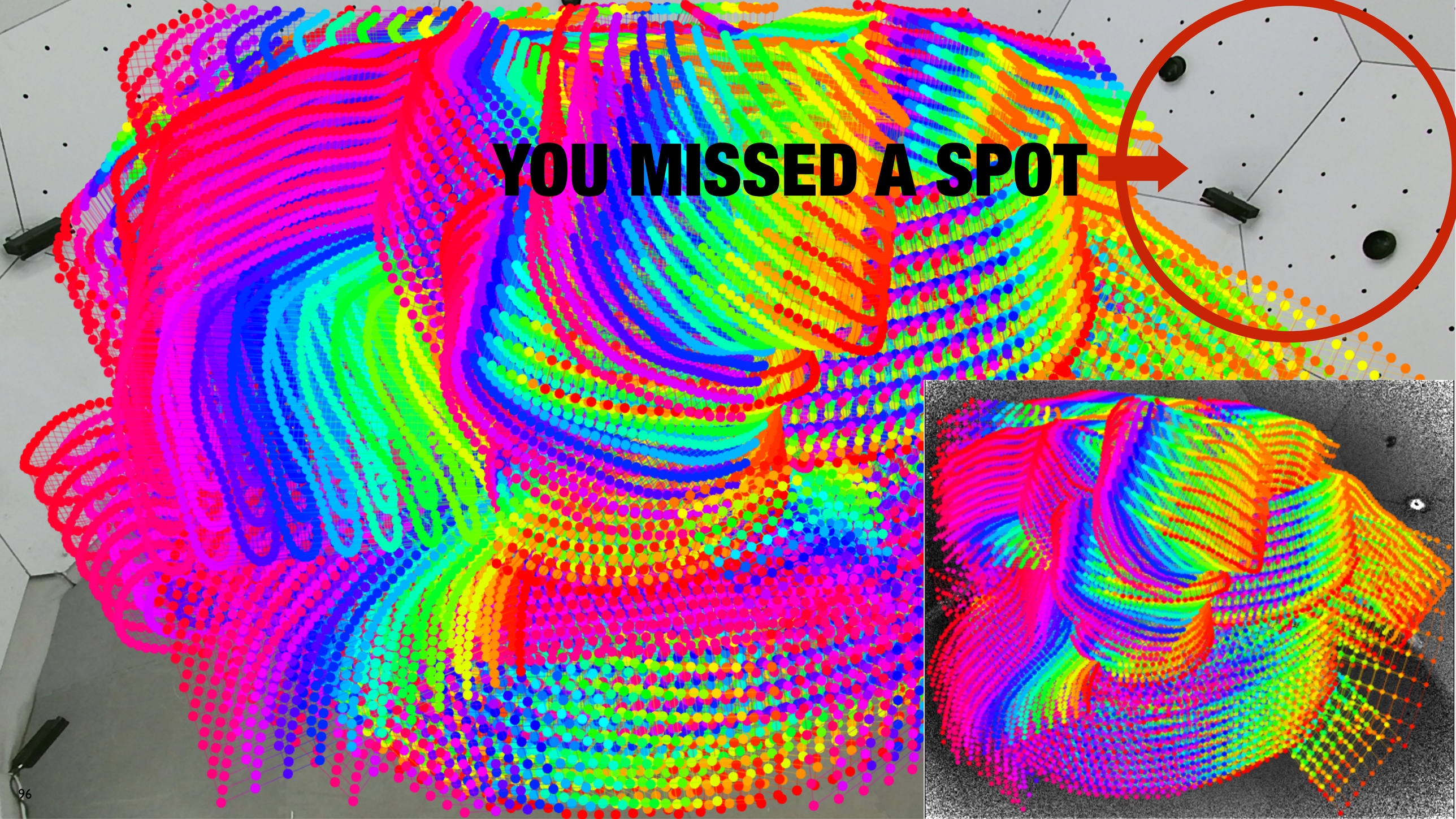


Depth

... Back To The Checkerboard






YOU MISSED A SPOT

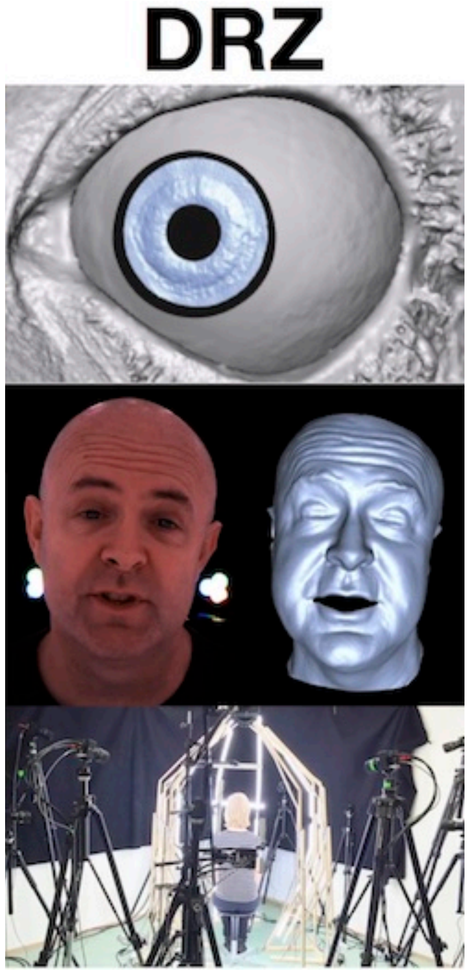
Kinect Calibration

Take Home Messages

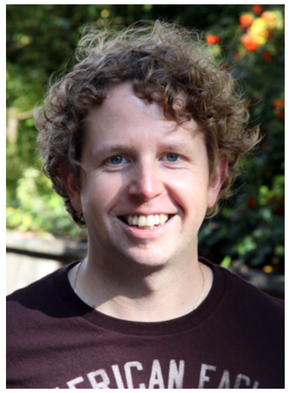
- Combining RGB+IR+Depth will require specialized calibration
- Static checkerboard to calibrate intrinsics and extrinsics between color & depth per Kinect, SfM between sensors
- Multiple RGB-D sensors may produce additional noise and interferences

Break/Questions

After the break:



Thabo Beeler

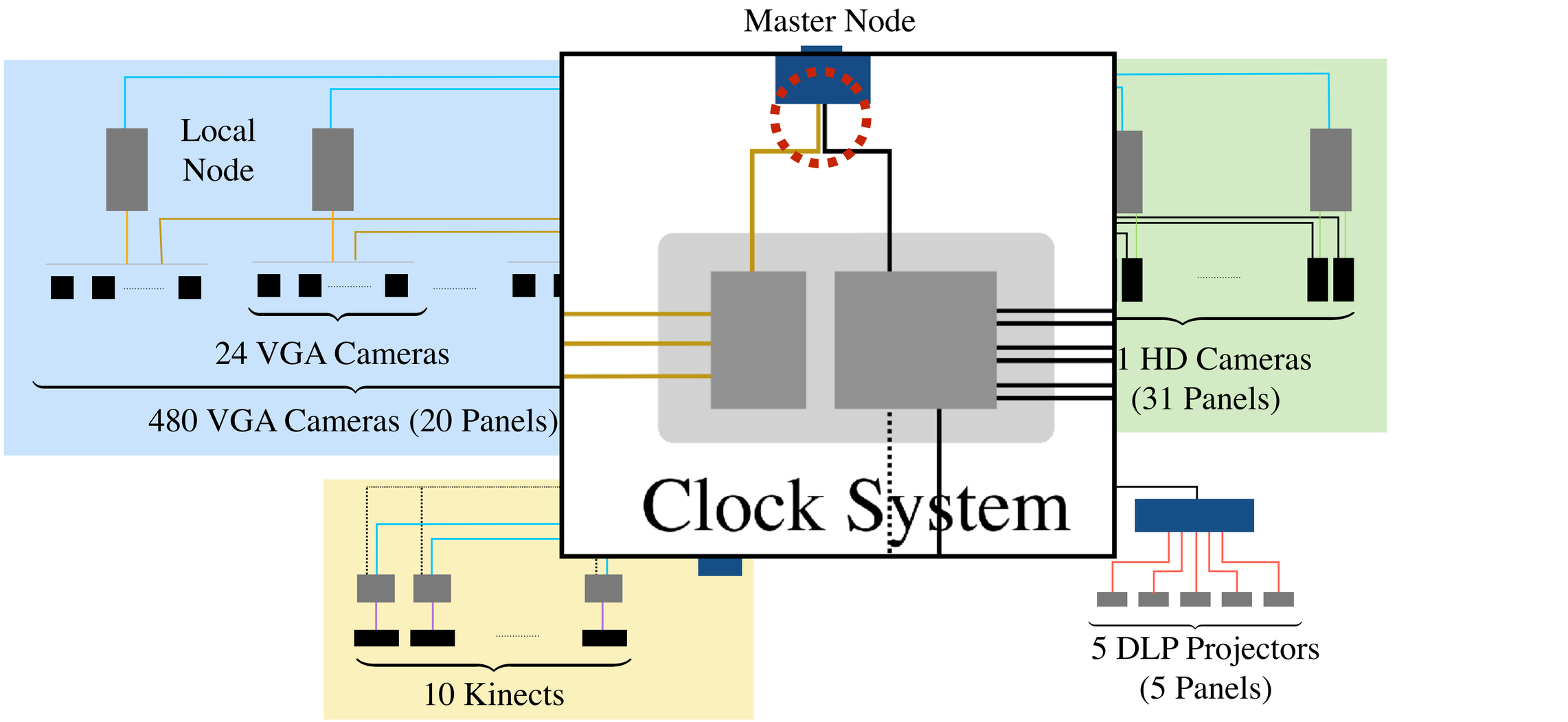


Derek Bradley

Backup Slides

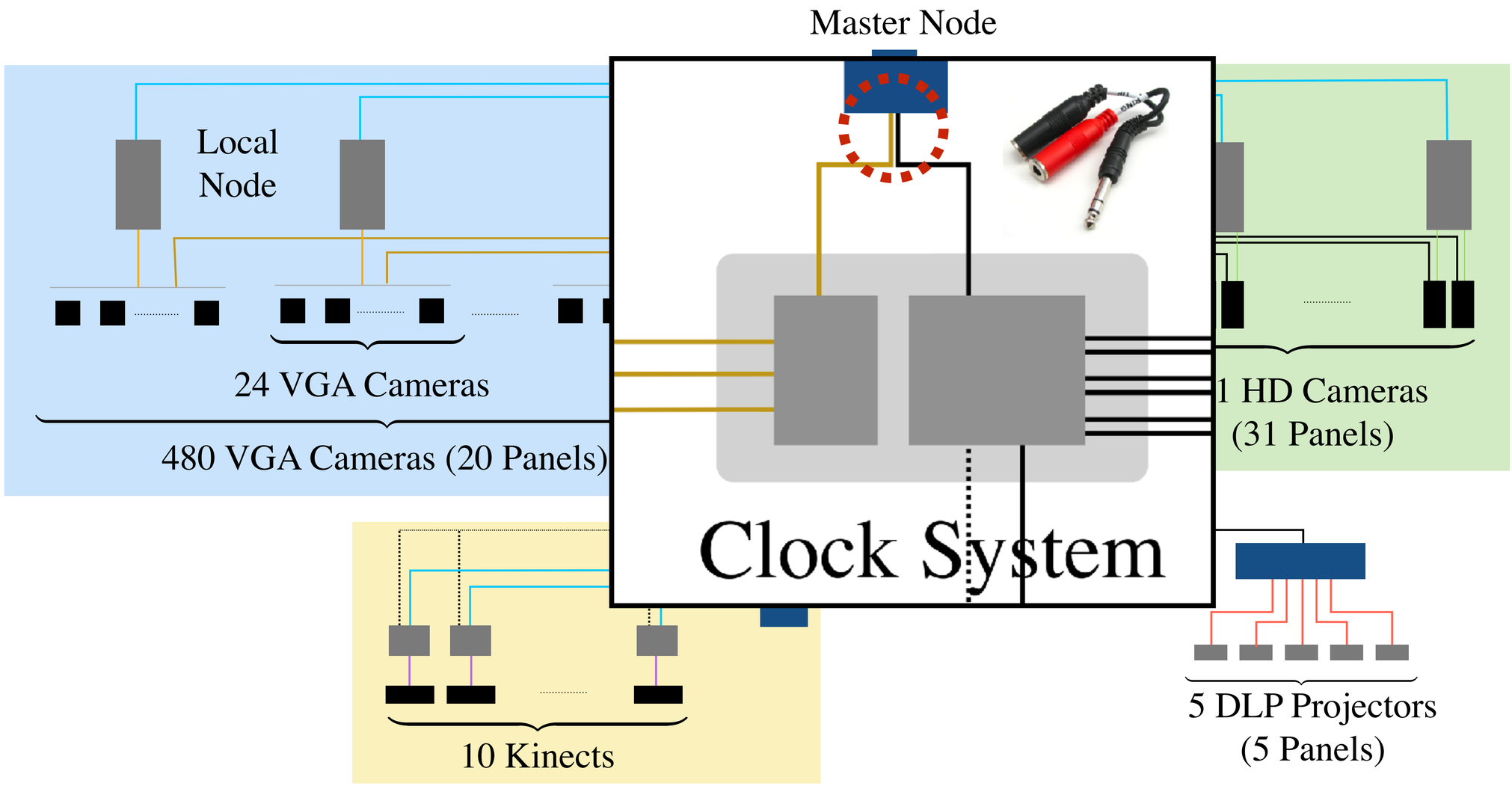
Panoptic Studio Architecture

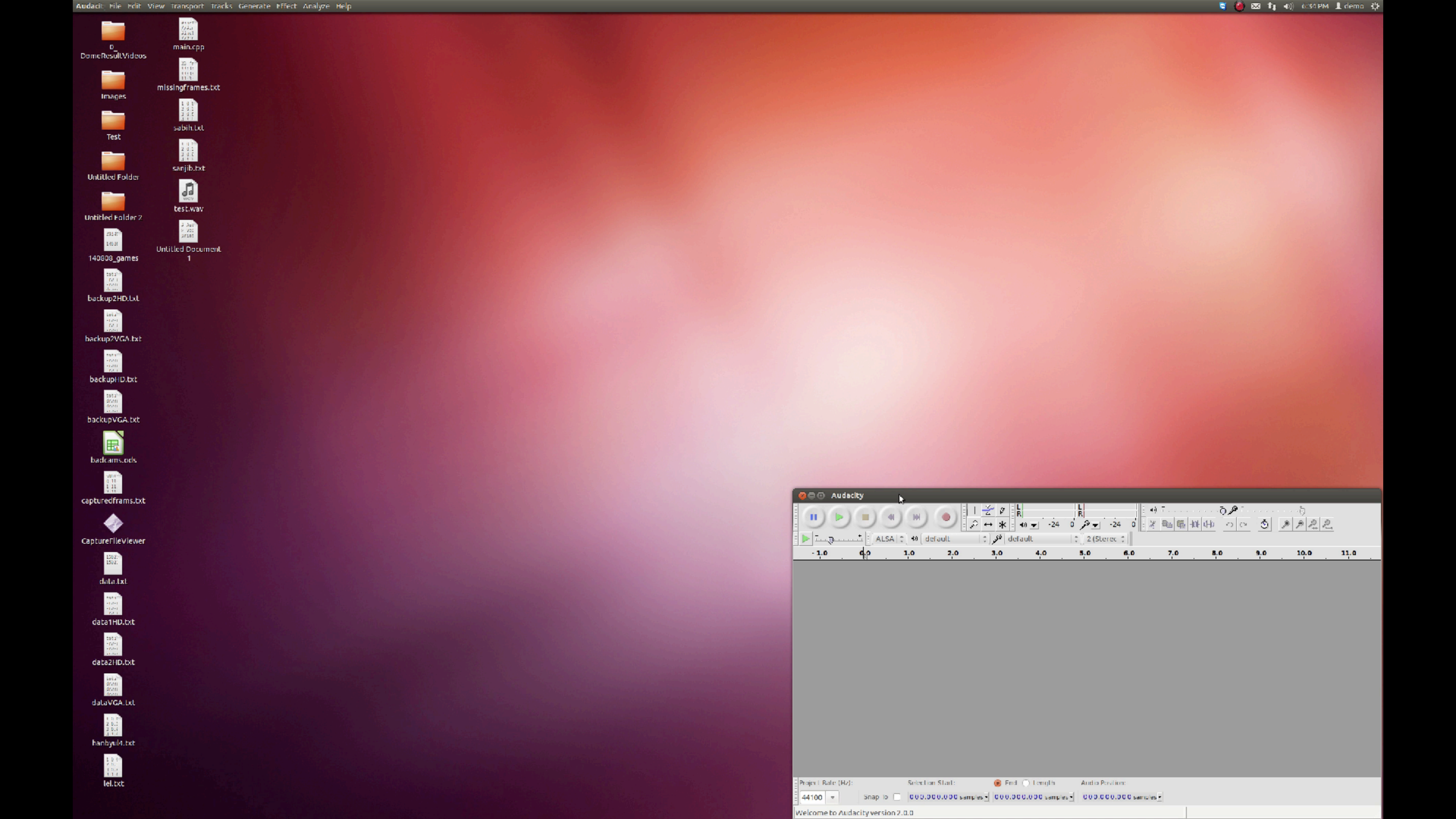
Time Alignment Across Subsystems



Panoptic Studio Architecture

Time Alignment Across Subsystems




Audacity File Edit View Transport Tracks Generate Effect Analyze Help
6:34 PM demo
0_DomeResultVideos
Images
Test
Untitled Folder
Untitled Folder 2
140808_games
backup2HD.txt
backup2VGA.txt
backupHD.txt
backupVGA.txt
badcams.ods
capturedframs.txt
CaptureFileViewer
data.txt
data1HD.txt
data2HD.txt
dataVGA.txt
hanhyul4.txt
lel.txt
main.cpp
missingframes.txt
sabih.txt
sanjib.txt
test.wav
Untitled Document 1
Audacity
ALSA
default
default
2 (Stereo
Welcome to Audacity version 2.0.0