

Opportunities for NDN in Augmented Reality

Jeff Burke

NDNComm 3
March 23, 2017

Purpose of this brief talk

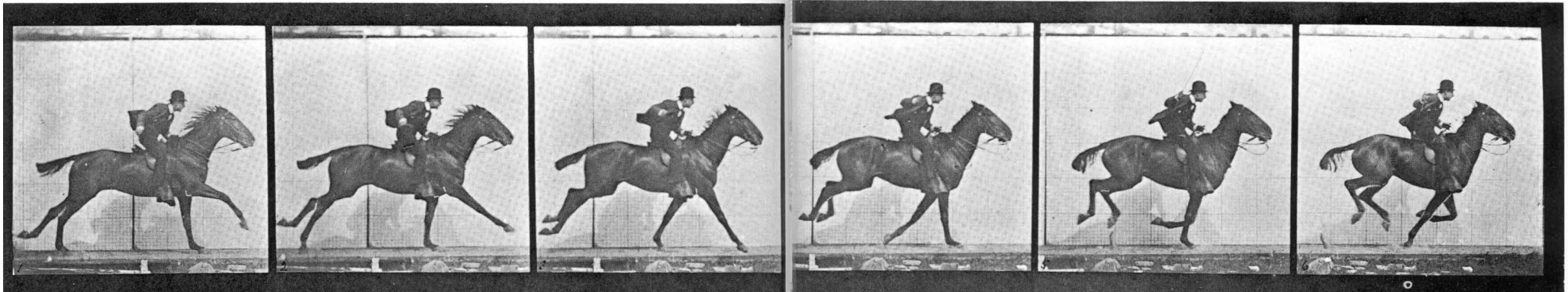
Continue previous individual conversations about augmented reality, and a brief discussion at the NIST workshop.

Outline application areas / emerging media delivery forms to encourage the NDN community to broaden beyond 2D video delivery applications as drivers.

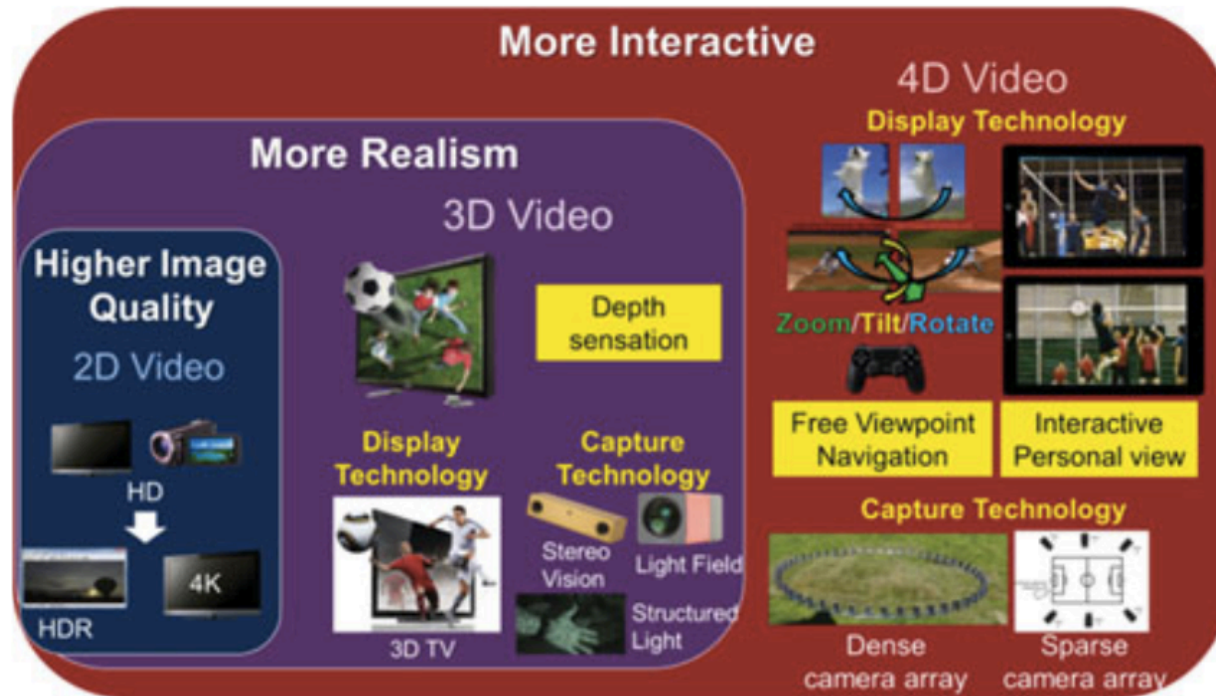
Suggest directions we will be going in at UCLA REMAP as part of upcoming (hopefully) funded programs.

Not just AR...

Emerging Delivery Technologies



Evolution of Video (not to mention “CG” imagery)



From: Lee, Chuen-Chien, Ali Tabatabai, and Kenji Tashiro. "Free viewpoint video (FVV) survey and future research direction." *APSIPA Transactions on Signal and Information Processing* 4 (2015): e15.

Immersive Video, Free Viewpoint Television, Volumetric Capture

Source: NTT

Record a scene from multiple angles and enable interpolation between views; or, use a single spherical camera for immersive capture.

Enable interactive viewpoint selection by the viewer.

e.g., Work by NTT, Sony; volumetric capture by 8i, others. First comedy series in 360 released by Hulu in 2016.

Tanimoto, Masayuki. "Free-viewpoint television." *Image and Geometry Processing for 3-D Cinematography*. Springer Berlin Heidelberg, 2010. 53-76.

3/22/17

Opportunities for NDN in Augmented Reality

Watch game like an actual spectator
C-17
Free Viewpoint Video Synthesis and Delivery
Personalized Video Playback from Arbitrary Viewpoint
NTT
The system synthesizes and delivers videos according to users' requests. It can provide video from an arbitrary view in real time even if no actual camera exists at the viewpoint. The video is generated from the output of multiple sensors including a video camera, RGBD sensor, and range sensor.

Capturing the scene using multiple sensors
Real-time synthesis of free-viewpoint video

RGB/depth sensor
360° range sensor
360° spherical camera

Features

- Real-time synthesis and delivery of free-viewpoint video, which is from an arbitrary viewpoint and arbitrary view direction.
- The system employs the international standard for free-viewpoint video coding. The standard, which NTT has driven forward, suits situations where a large number of viewers watch content simultaneously.
- Improved sensors potentially derive better quality of free-view video synthesis. The application of the method to sports such as soccer that are played on very wide fields will be investigated as future work.

Application Scenarios

- While users watch sports in a stadium, they can replay the game in their own portable devices from arbitrary viewpoints.
- This free-viewpoint playback can be done at home or in public viewings.
- For a broadcasting scenario, commentators can use arbitrary viewpoints to explain what is happening in the play when a real camera view cannot get a good view.

Free-viewpoint playback in stadium on personal devices such as tablets!
Realtime free viewpoint playback even when watching TV in living room!

Contact: yu.kawan@ntt.net.jp
Copyright © 2015 NTT. All Rights Reserved.

FAST COMPANY
LEADERSHIP MAGAZINE MOST INNOVATIVE COMPANIES MOST CREATIV
Why Volumetric VR Is The Real Future Of Virtual Reality
With 8i's technology, it's possible to walk around a human subject in a VR experience. That makes it more immersive, more real, say experts.

8i

Contemporary broadcast



3/22/17



Opportunities for NDN in Augmented Reality



Francis Ford Coppola brings experimental 'Live Cinema' workshop to UCLA



Francis Ford Coppola speaks at a news conference presenting his experimental Live Cinema workshop at the UCLA School of Theater, Film and Television. (courtesy of UCLA School of Theater, Film and Television)

By **Mark Olsen**

JULY 23, 2016, 6:25 PM

“I am still a boy scientist in my heart,” Francis Ford Coppola said to a group that had gathered at a UCLA theater to hear the legendary director talk about his latest project, an ambitious process he calls “Live Cinema.”

Spatially Selectable Video / Vertical Editing

“Up until now...[motion] picture editors have thought almost exclusively in the **horizontal direction**.

The question to be answered was simply, “What’s next?”
...that’s complicated enough—there are a tremendous number of **options** in the construction of a film.

In the future, that number is going to become **even more cosmic** because film editors will have to start thinking vertically as well, which is to say: “**What can I edit within the frame?**”

-Walter Murch, In the Blink of an Eye, 2011

Gandhi, Vineet, and Rémi Ronfard. "A computational framework for vertical video editing." 4th Workshop on Intelligent Camera Control, Cinematography and Editing. 2015.

D'Acunto, Lucia, et al. "Using MPEG DASH SRD for zoomable and navigable video." Proceedings of the 7th International Conference on Multimedia Systems. ACM, 2016.



Light field Photography

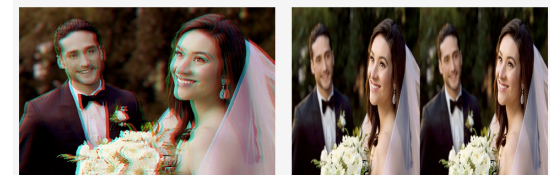
Captures the light field from a scene using (for example) a microlens array, rather than a conventional single lenses.

Enables computational photography techniques for focus, depth of field, orientation, and other changes after the fact, depending on the camera.

e.g., Lytro, OTOY, Stanford University Computer Graphics Laboratory

Ng, Ren, et al. "Light field photography with a hand-held plenoptic camera." Computer Science Technical Report CSTR 2.11 (2005): 1-11.

Source: Lytro.com

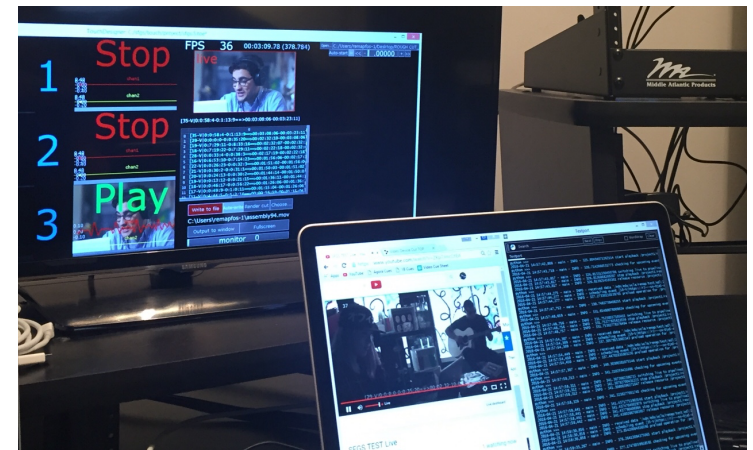


Dynamic Video Assembly

Real-time or just-in-time assembly of video frames for live streaming based on based on story choices.

e.g., Work by Google Creative Labs, REMAP, Interlude.fm. (And many others in a fully computer-generated context.)

Literat, Ioana. "The work of art in the age of mediated participation: Crowdsourced art and collective creativity." *International Journal of Communication* 6 (2012): 23.



Not just delivering rectangles any more

- Video source switching
- Spatial and temporal ROIs
- Viewpoint (within a scene)
- Focus and depth of field (for example)
- Story-related choices

Not shown:

- CG-generated scenes
- Real-time feature detection
- Real-time inpainting, etc.
- Payment and membership information.

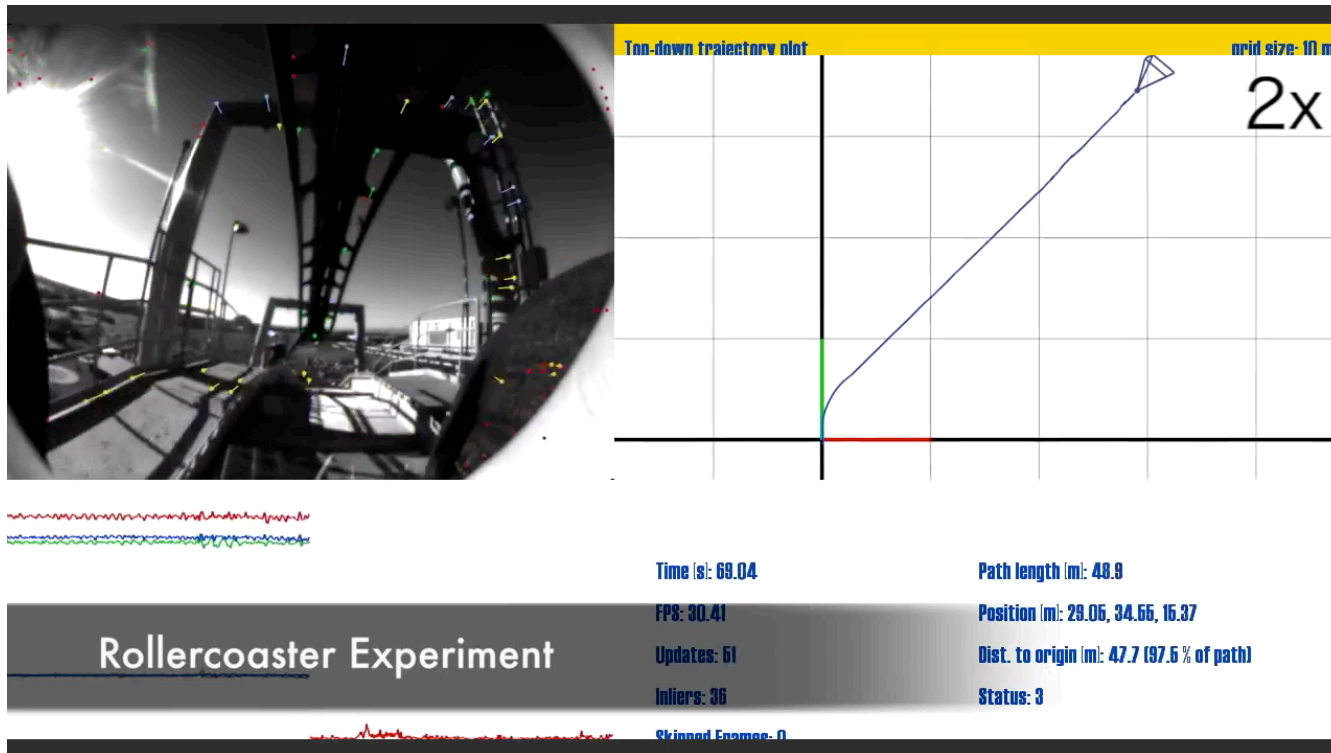
Augmented (and Virtual) Reality

All the above and more...

- Necessarily mobile and low-latency
- Continuous query for relevant content
- Multiple content publishers, variety of trust relationships
- Integrate a variety of data sources, including IoT
- Open-ended interactions

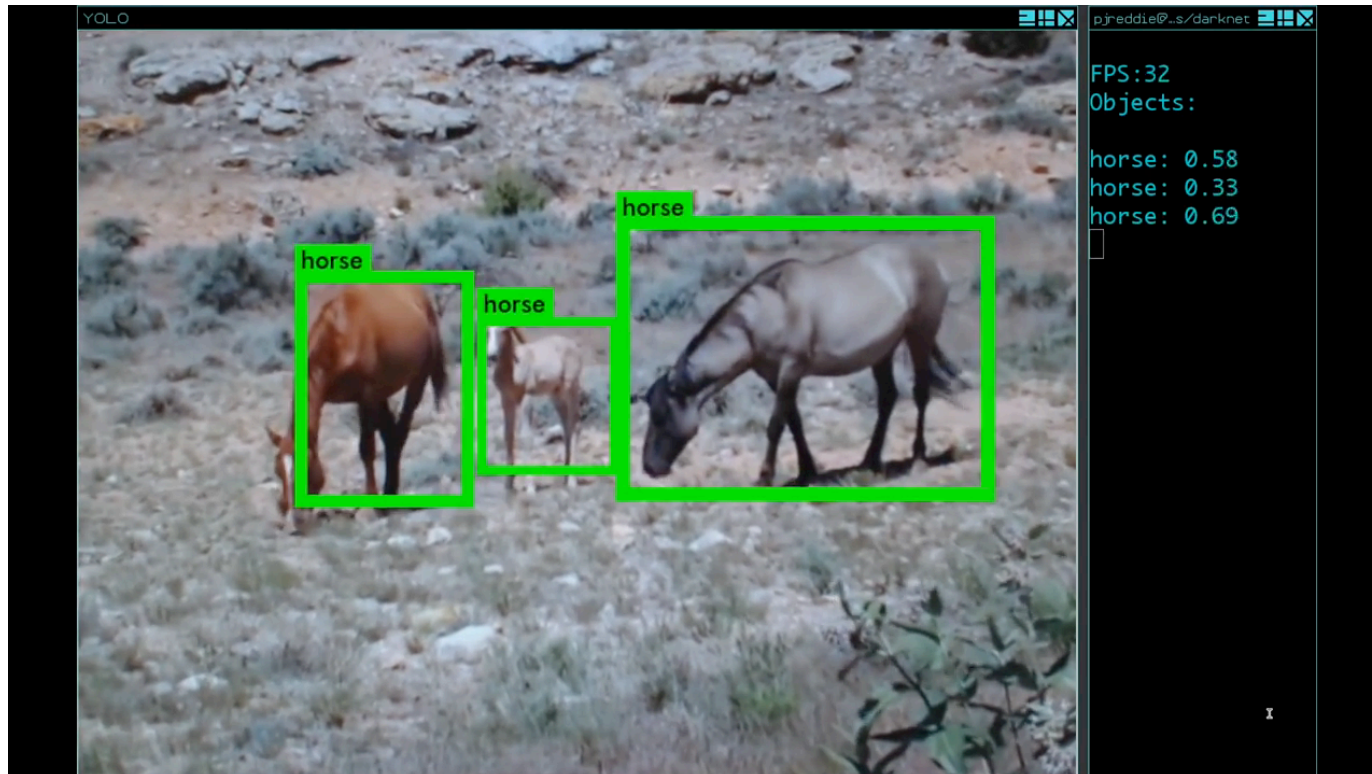


Real-time context (edge-supported?)



Google Tango @ GTC 2015

Real-time context (edge-supported?)



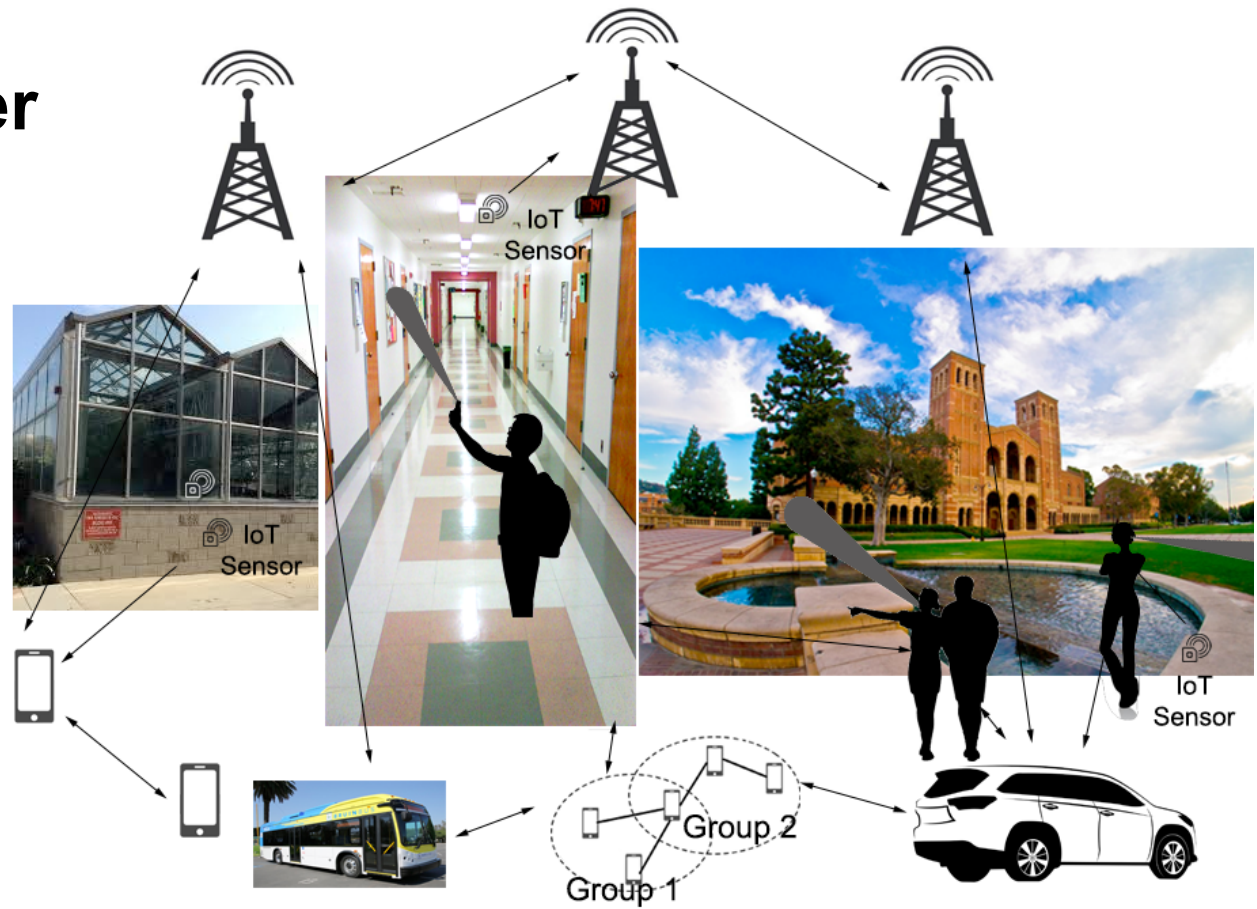
Darknet YOLO (You Only Look Once)

Opportunity for NDN

What NDN *should* be able to do

- Highly granular, post-web content delivery
 - Enable viewpoint-/context-dependent, bandwidth-adaptive applications
 - Direct name-based access to content grains with fewer roundtrips, leverage ultra low latency promised in 5G
- Enable every mobile user to publish content (and context) to many consumers without hitting the cloud first.
 - We need this for delivery *context*, not just live streaming your day.
- Leverage multiple radios in current and future mobile devices easily.
- Provide a consistent mechanism for access on-demand and stored content, services, edge processing, as well as access to IoT and other data.
- Name-based access control offers potential for efficiencies.

Application – Campus browser



Application – Fiction / Entertainment



Entities

- Consumer – content consumer, usually mobile; more than one device, more than one radio, more than one app.
- Publisher – content generator, often mobile; sometimes running on the same device as the consumer.
- (Mobile) Edge – compute and storage capability close to the (mobile) node
- Cloud / Infrastructure – capability further from the edge

Proposal

- Attempt a similar approach for different types of dynamic media.
- Separate and explicitly define path, viewpoint and other context through the content:
 - Major source and publisher selection and viewing choice.
 - Spatial and temporal regions of interest.
 - Focus and depth of field (for example).
 - Content preferences / story choices.
 - Membership, identity, payment information.
- Consumers use sync to offer context, publishers use sync to offer content.

Design Concept (1/4) - Publishing

One or more end-user applications and services:

1. Publishers name the “world” of content, which may include private and dynamic content as well as public and static assets.
2. Consumers and publishers exchange relevant content names hierarchically and asynchronously.

- Challenges:
- Naming strategies for gross content selection – prefix components for providers and content.
 - Progressive / k-nearest-neighbor queries for “nearby” content.
 - Developing a trust model; how to avoid pulling untrusted content.
 - Access control strategies for content that balance publisher, provider, and user needs – not necessarily specific to AR and other emerging content types.

Design Concept (2/4) - Context

1. Consumers name and publish the local context of their viewpoint(s).
 - Path, viewpoint, story choice, preferences.
 - Content type requests, payment offers.
 - Some context (e.g., cameras) may be content.
2. Edge services name and provide external context.
3. Consumers and edges publish context, path, viewpoint asynchronously.

- Challenges:
- Context naming
 - Multi-party exchange of content and context – is this sync?
 - Who gets keys to access what context information?
 - Offering calculation services – leveraging NFN

Design Concept (3/4) - Fetch

1. Consumers assemble the names of content grains and content keys, based on the predicted path and context.
2. Consumers fetch content grains and keys:
 - Prioritizing current needs and mostly likely future path through the content
 - Leveraging names to fetch progressive / scalable content representations

- Challenges:
- Naming strategies for each content type – loosely, the suffix components of the name.
 - Challenges of set / attribute based naming of the content.
 - Access control approaches for content – probably a hybrid.
 - We need to prioritize some content grains over others, any network support for this?

Design Concept (4/4) – Path Mgmt

1. Cloud-based components retrieve context including *potential, fetched, and actual paths* through content.
2. Infrastructure uses machine learning to offer fetching predictions for a given user and path. (Not just for ads any more!)
3. Edge services leverage path predictions to progressively prefetch most likely to be needed content close to the consumers.

- Challenges:
- Conventions for path / viewpoint descriptions for each content type, as well as location, orientation, and other context.
 - Access control decisions.
 - Application-specific challenges in prefetching, outside of the NDN networking context.

Experimental direction

- Build a Unity-based augmented reality app, using NDN-CCL C# library.
 - Implementation challenge: Existing AR toolkits probably not conducive to supporting this approach to networked content.
- Focus on supporting context awareness, not just content.
- Leverage experience with NDN-RTC / Flume for sketching media metadata descriptions and a new approach to sync.
- Expand on (simplify?) name-based access control developed for NDNFit and other apps – focusing on exchange of context / content.
- Lots of conversations needed about places where the term “exchange” is used in the previous slides – can these be built up from the sync primitive?

Smaller Slices towards NDN AR Innovation

- Research approaches for emerging media types
- Pursue network-supported computer vision
- Research like Common Name Library, Consumer/Production API – higher-level abstractions
- Explore context-content exchanges among consumers and publishers
- Access control suitable for “big content” providers
- Naming registration / routing approaches for mobile publishers
- Make “mundane resilience” a priority if we want to see prototypes

Opportunities for NDN in Augmented Reality

Jeff Burke

NDNComm 3
March 23, 2017