AWS re:Invent

NOV. 29 – DEC. 3, 2021 | LAS VEGAS, NV

NFX204

Large-scale distributed training of media ML models with Amazon FSx

Lingyi Liu
Netflix

Mahdi Kalayeh
Netflix

Nagendra Kamath
Netflix

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Outline

Media ML applications in Netflix

Self-supervised video representation learning

Distributed GPU training architecture and challenges

Scalable data loading with Amazon FSx for Lustre

Result analysis

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Media ML applications in Netflix

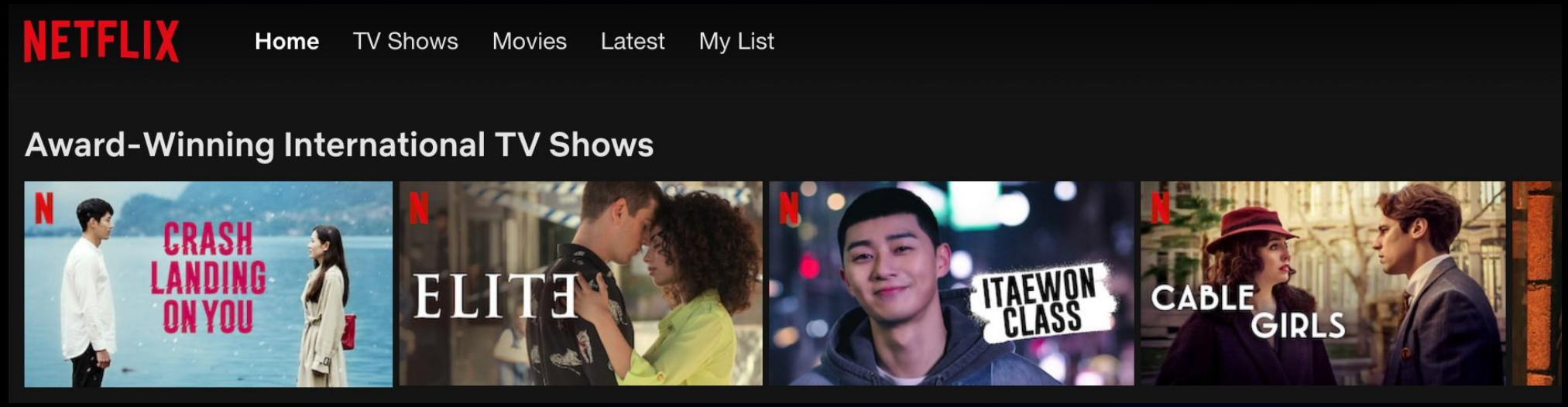
- Goal: optimizing a visual representation of our content
- Develop computer vision (CV), audio, and NLP algorithms/tools
  - Aid artwork generation
  - Aid video trailer generation
  - Enhance productivity
    - Creative process
    - Post-production, such as quality control and VFX
  - …

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Media ML applications in Netflix

- Example I: retrieve character-focused video thumbnail candidates
  - Assistive tool to narrow down the search space for creatives



Paper: Character-focused Video Thumbnail Retrieval, Shervin Ardeshir, Nagendra Kamath, Hossein Taghavi, ICML 2020 ML4MD

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

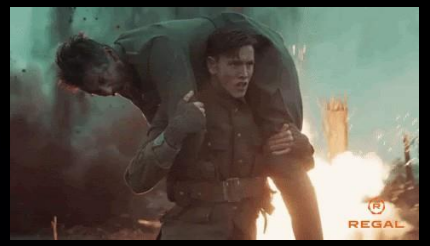
Media ML applications in Netflix

Example II: learning trailer moments

Moments in “Action” trailer







Moments in “Romance” trailer







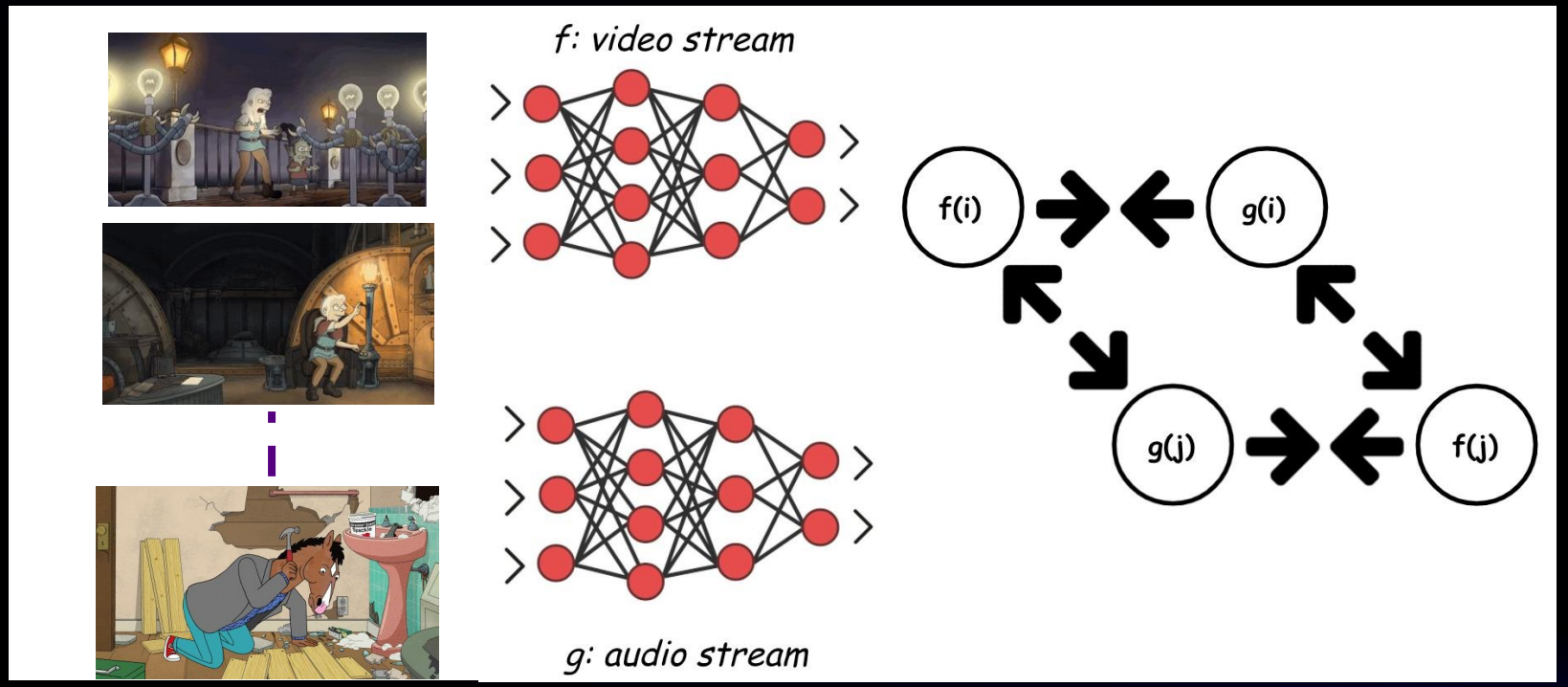
Paper: Learning Trailer Moments in Full-Length Movies with Co-Contrastive Attention, Lezi Wang, Dong Liu, Rohit Puri, Dimitris N. Metaxas, ECCV 2020

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Self-supervised video representation learning

Learning a general-purpose, content-driven representation of Netflix video assets through self-supervision while preserving temporal granularity

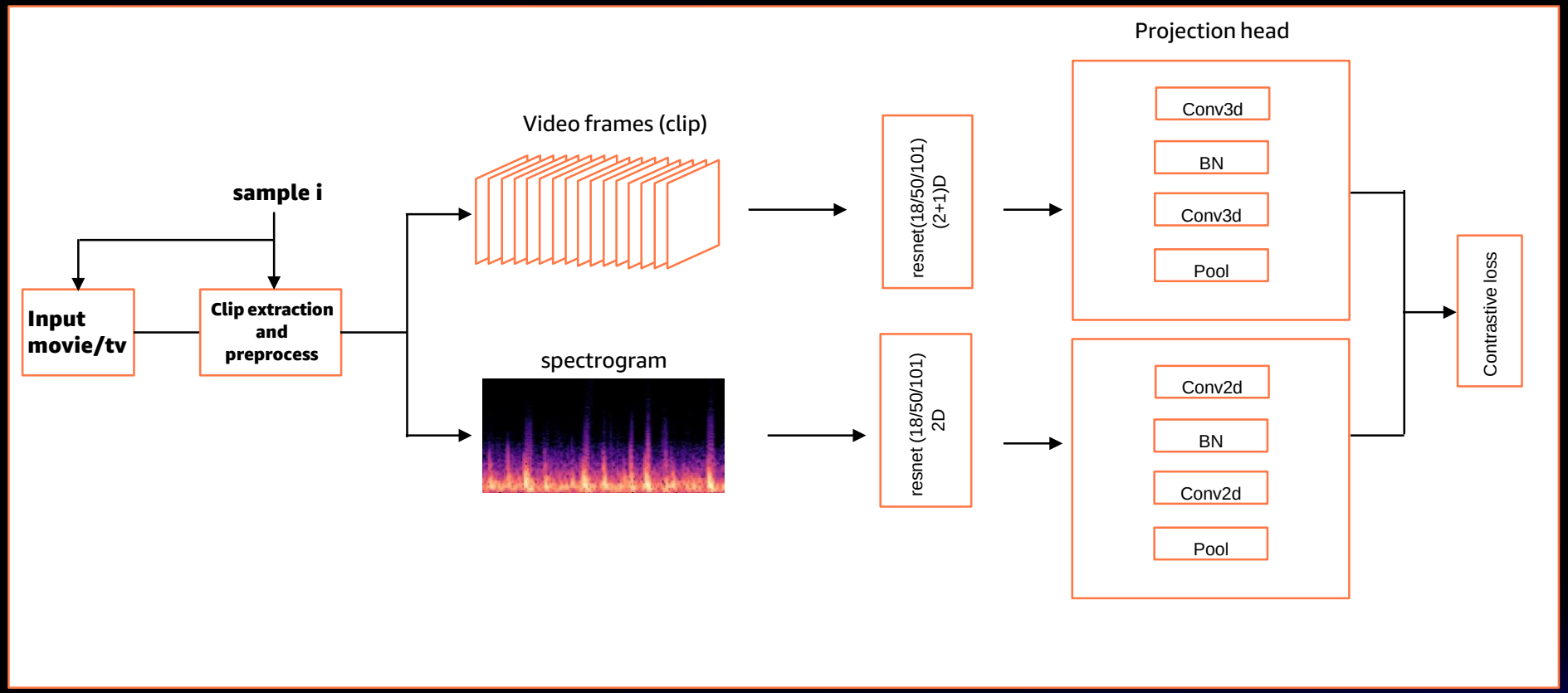


Paper: Watching Too Much Television is Good: Self-Supervised Audio-Visual Representation Learning from Movies and TV Shows.
Mahdi M. Kalayeh, Nagendra Kamath, Lingyi Liu, Ashok Chandrashekar: CoRR abs/2106.08513

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

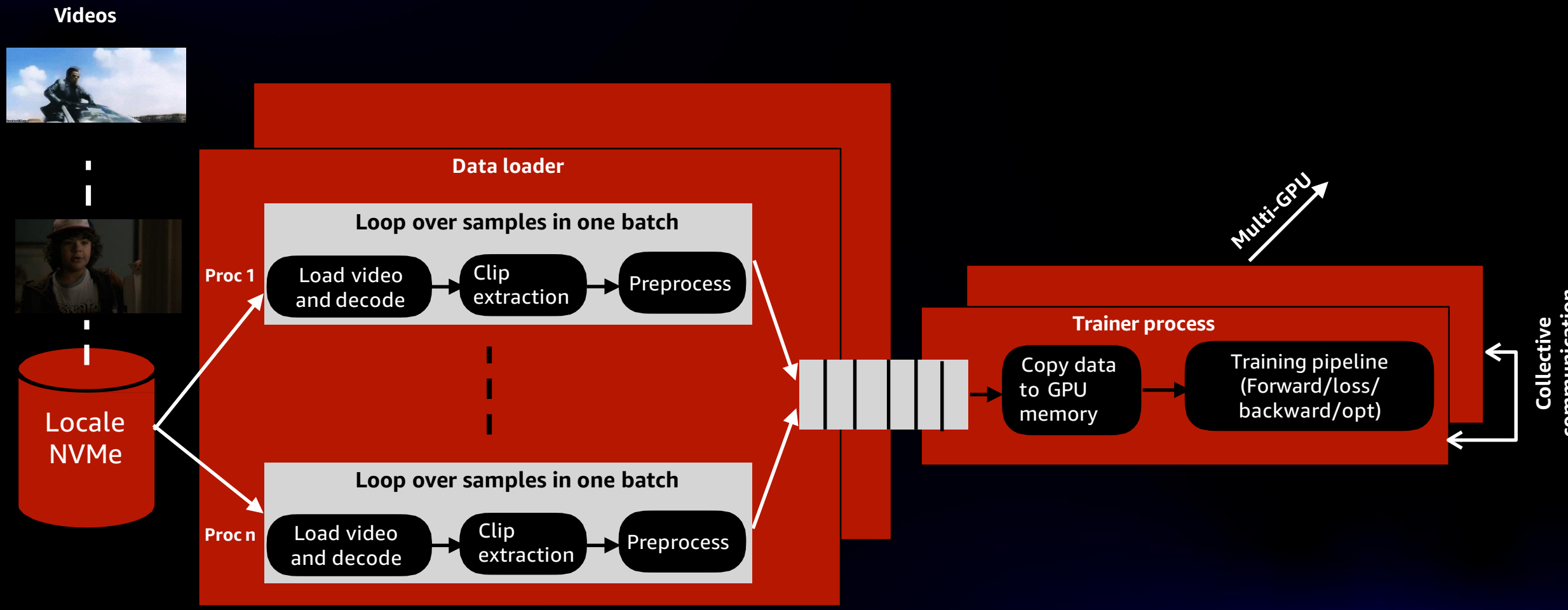
Self-supervised video representation learning

Dataset and CNN model



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Distributed GPU training architecture and challenges



Multi-GPU training architecture for media models (V0)

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Distributed GPU training architecture and challenges

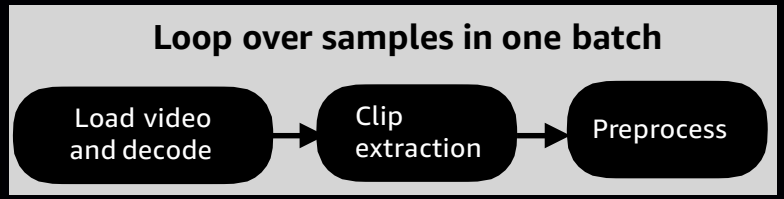
- Bigger dataset (10–100 TB)
  - More movies/TVs in our catalog
  - P4dn NVMe: 8000 GiB
  - Slow rotational drive
- Multi-node distributed training support
  - Duplicate dataset across node
  - Dataset shuffle

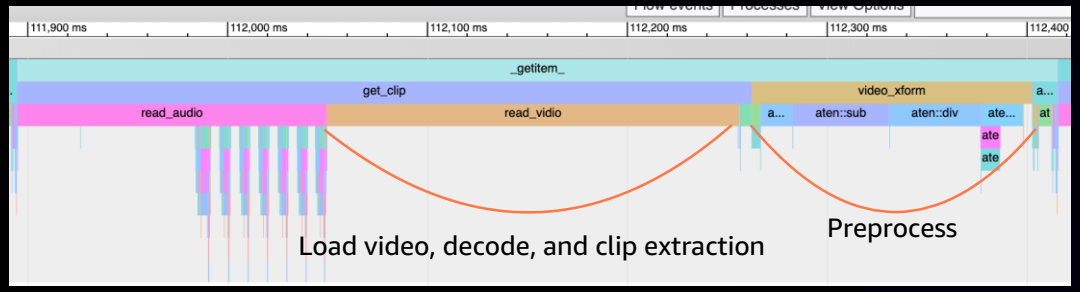
aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

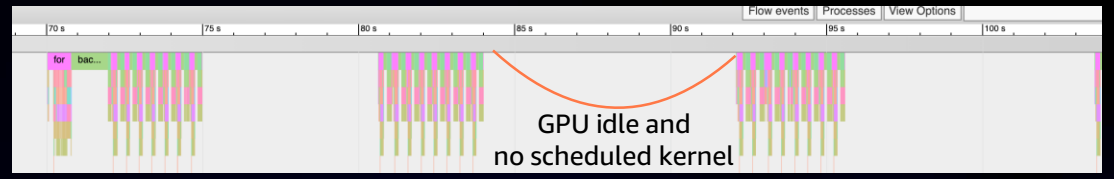
Distributed GPU training architecture and challenges

- Low GPU utilization
  - GPUs are mostly waiting for data





Data loader process profiling



Trainer process profiling

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Distributed GPU training architecture and challenges

- Computation (✓)
  - High-end GPUs: V100/A100
  - High performance operator library by TF or PyTorch
- GPU communication (✓)
  - NVLink and Elastic Fabric Adapter (EFA)
  - Fast all-reduce and other collection comm by NCCL
- Fast I/O to ingest the data into GPU memory (✗)
  - Typically, 5–8 GB/s ingest throughput per node with 8 GPUs
  - 100 TB data
  - Shuffling support

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Scalable data loading with Amazon FSx for Lustre

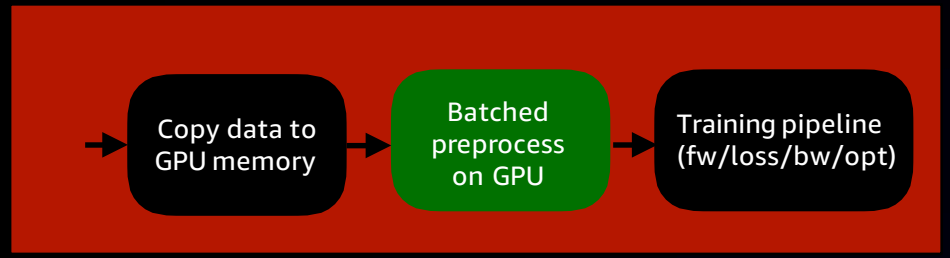
- Revisited data loading pipeline toward saturating GPUs
  - Precomputing all video/audio clips
    - Save video loading/decoding time
      - Minor changes to data loader
      - {idx} → video_clip_{idx}.mp4
    - Reused in training with different hyper parameters
    - Retaining unchanged computations in precompute
      - read_video, decode_video, extract_clip, resample

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Scalable data loading with Amazon FSx for Lustre

- Revisited data loading pipeline toward saturating GPUs
  - Offloading remaining preprocess to GPU
    - Data augmentation
    - Batching transform
    - Put it at the beginning of training pipeline

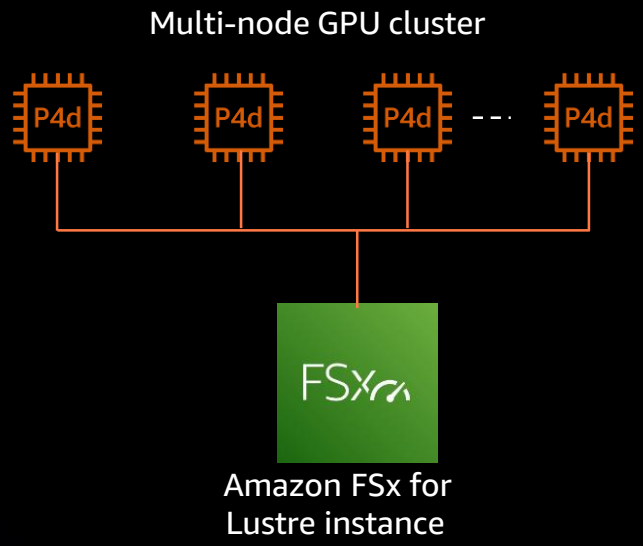


Q: how to efficiently store/access these audio/video clips?

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Scalable data loading with Amazon FSx for Lustre

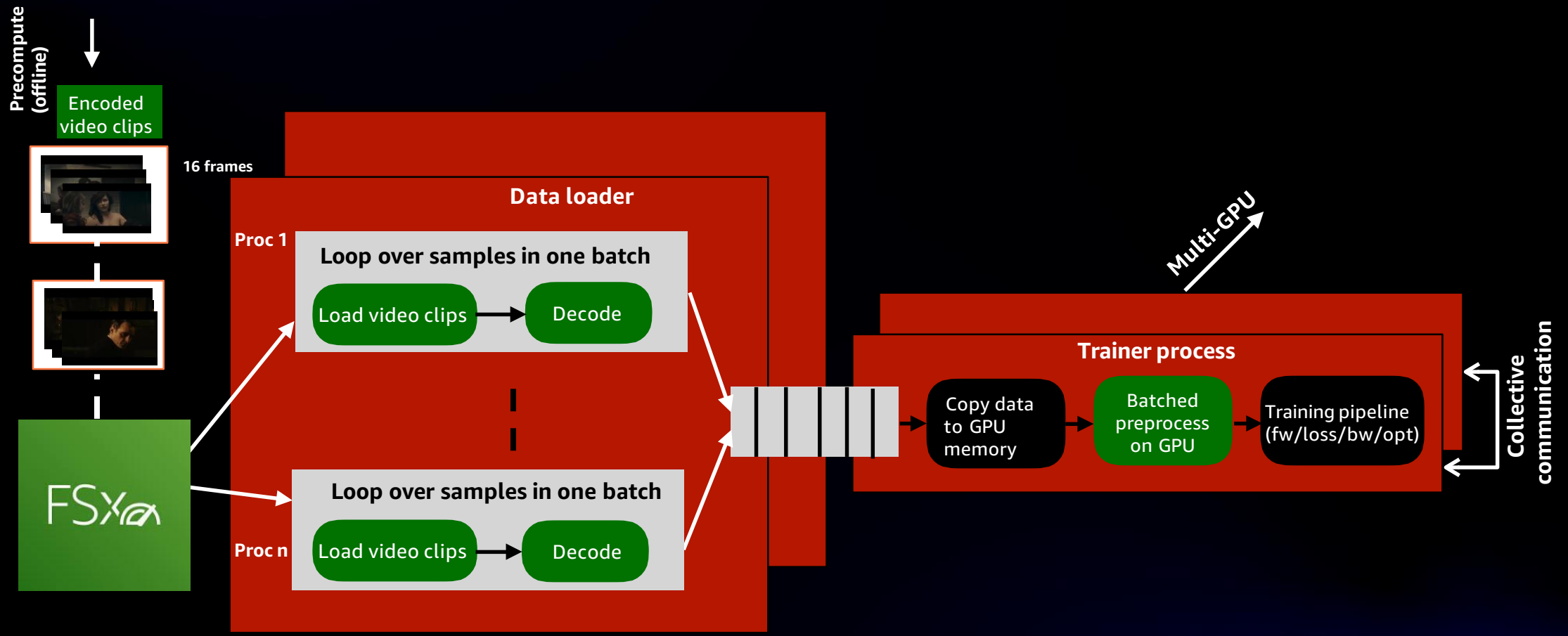
- Amazon FSx for Lustre
  - High performance network file system
  - Massively scalable performance
- Storing all the precomputed clips in FSx for Lustre instance
- Scalable loading data from multi-node GPU machines
- Automatic support of data shuffling for different epochs
- **Encoding video clips to optimize storage**
  - Per node, 96 CPUs to decode clips during training time



aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Scalable data loading with Amazon FSx for Lustre



Multi-GPU training architecture for media models with Amazon FSx for Lustre (v1)

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Scalable data loading with Amazon FSx for Lustre

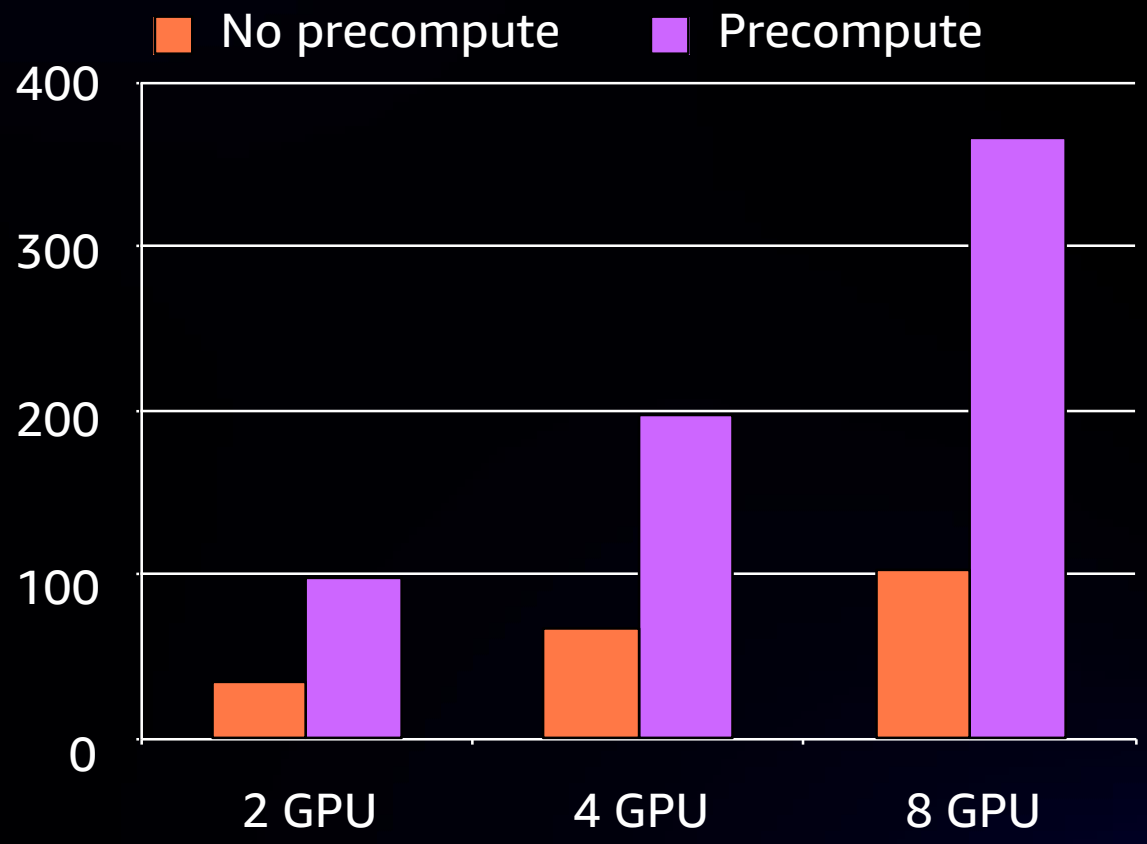
- Client performance tuning for FSx for Lustre
  - Read/write pattern
    - Each clip file is read once, cannot be reused across epoch
    - Write clip file is also once in precompute
  - Tune max_cache_mb to reduce client memory usage
    - Important for writing during precompute
  - Tune max_rpcs_in_flight to increase concurrent RPCs
  - Watching read/write stats such as offset, RPC stream, and I/O size
  - Try data compression to further save storage cost

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Result analysis

Throughput: number of samples ingested by the training system

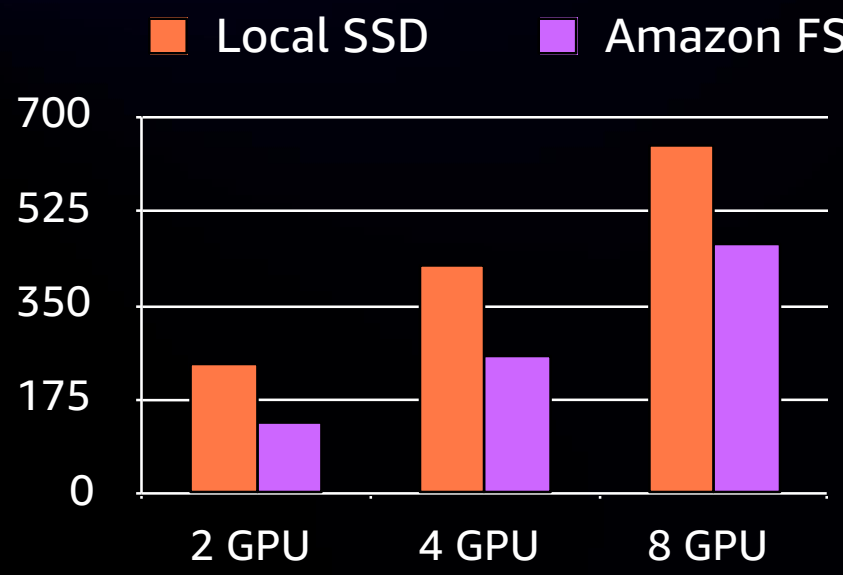


3–4x throughput improvement with precompute & Amazon FSx
{A100, 5 loaders per GPU}

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Result analysis

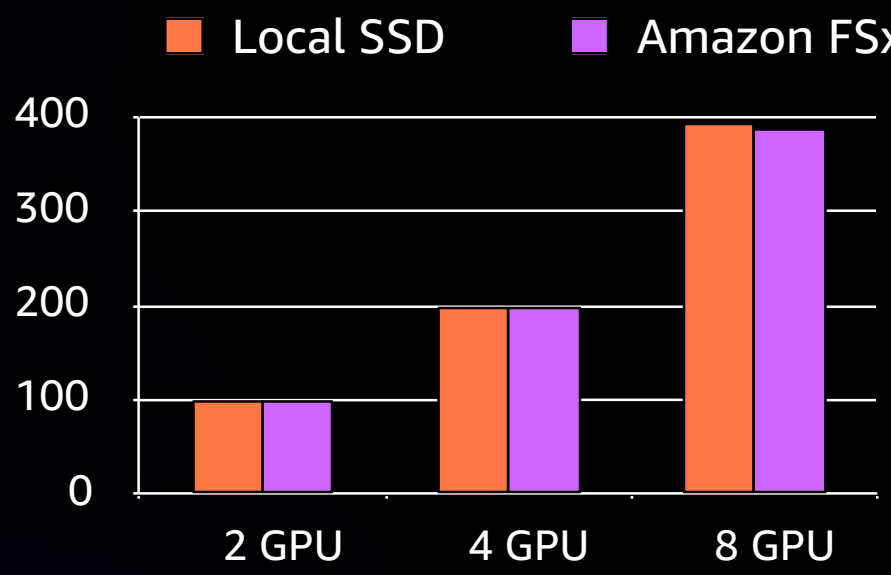
Case I: data I/O bottleneck (small model)



Local NVMe disk versus Amazon FSx
{A100, 5 loaders per GPU}

- Amazon FSx could scale linearly for workload with I/O bottleneck
- Amazon FSx is more scalable to parallel reading
  - 4 GPU: 60% of local SSD
  - 8 GPU: 71% of local SSD

Case II: compute bottleneck (deeper model)



Local NVMe disk versus Amazon FSx
{A100, 5 loaders per GPU}

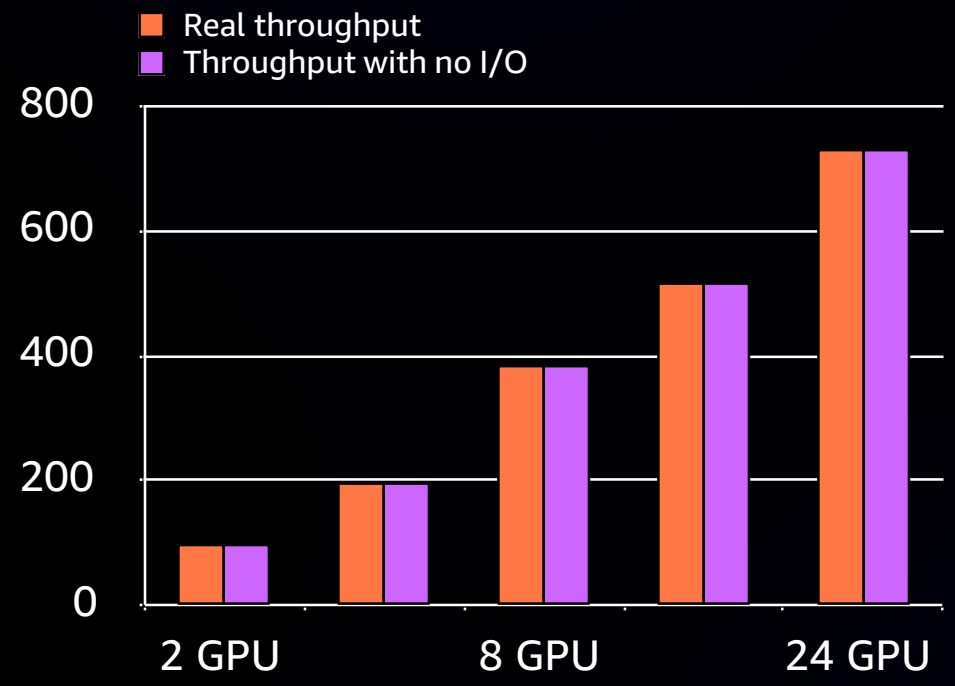
For our media ML workload with model computation bottleneck, Amazon FSx could saturate GPU as the local SSD

aws

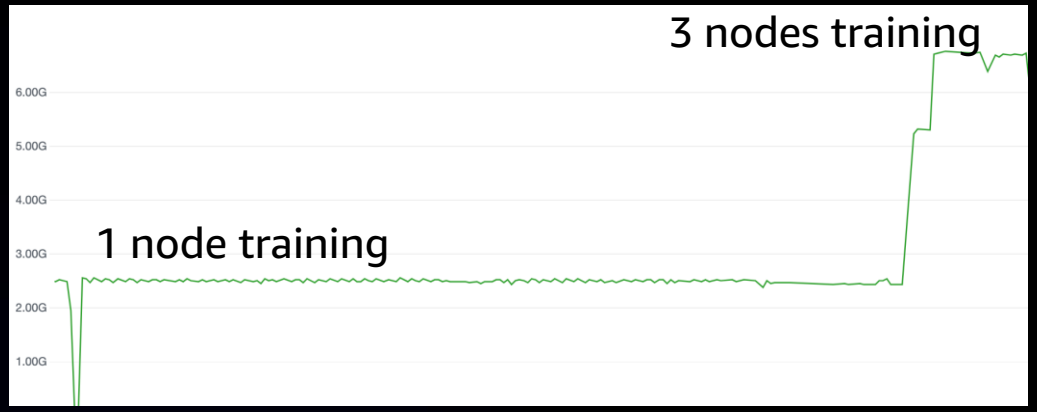
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Result analysis

Scaling multi-node distributed training with Amazon FSx instance



Amazon FSx instance monitoring



- I/O is not bottleneck for training throughput for our workload
- Multi-node scaling efficiency is aligned with PyTorch DDP

Amazon FSx is scalable for multiple clients

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Key takeaways

- Precompute all training samples could resolve the decoding/ preprocessing bottleneck in training of media ML models
- FSx for Lustre provides a scalable solution to store the precomputed samples
- FSx for Lustre is also scalable for multi-node distributed training support

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.