CS4400: INTERNET APPLICATIONS (SCALABLE COMPUTING), D-STREAM
NAME: SUDHANSH MEHTA (16340820)

NILE: A PROGRAMMABLE MONITORING COPROCESSOR (IEEE COMPUTER ARCHITECTURE LETTERS)

Key Contributions/findings: Hardware Performance Counters (HPC) limited in event monitoring capabilities; Nile Match Unit: either programmed by a user process to monitor own execution or by supervisor process to monitor process with lower permissions; Nile's Event- Action Model eliminates need for dedicated hardware to monitor tasks; Extend Linux to support Nile and alter task_struct in Linux kernel to save/restore Nile's Configuration of each process; Performance overhead of Nile due to frequent cache requests & storing/recovering to/from OS.
Key Technology Insights: Modified RISC-V core with a commit log to expose information about instruction execution; User Programmable Nile Match Units with shared memory for handling events; Out-of-Order processor: less power overhead, reduced performance overhead, non-conventional instruction queue methodology; Time-Multiplexing to monitor multiple events using a single Nile Match Unit; Detecting Stack Buffer overflow as an application of Shadow stack.
Key insights into Scalable Computing: Commit Log enables us to Decouple program execution from monitoring process; Out-of-Order processors improves scalability avoiding stalls (in processor cycles) compared to In-order processor; Decoupling the event monitoring process using multiple Match Units watching out for their specific events to occur; Scaling the Issue Queue (queue holding pending instructions: source operands not ready): larger selection of instructions to allow scheduler to find more Independent instructions; Cache latency(need arises with increasing number of cache requests) of Nile: to be dealt with writing to higher levels in memory hierarchy.

DEEP BIMODAL REGRESSION OF APPARENT PERSONALITY TRAITS FROM SHORT VIDEO SEQUENCES (IEEE TRANSACTIONS ON AFFECTIVE COMPUTING)

Key Contributions/findings: Modifying Convolutional Neural Networks (CNN) for exploiting important Visual cues (for Visual Modality); Choice of Non-Convex loss functions as optimal on human labeled data (reduces label noise); Personality Trait Prediction for test video (during testing phase): average scores of images/frames as the score of the video; Applying Ensemble Method (Ensemble of the Visual and Audio Modalities) for getting the final regression scores; Average of the 5-dimensional vector (one for each of Five Traits) as the final ensemble result (final score).
Key Technology Insights: Five Factor Model (Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism) for Personality Analysis; Apparent Personality Analysis (APA) as a choice over psychological tests (time and funds); Combining Convolutional Neural Networks (CNN) with LSTM Networks (Long Short Term Memory) for optimal representation of speech signal from raw audio data (for Audio Modality); Epoch Fusion Model (average of the results obtained after different epochs) as the prediction of Visual modality; Using t-distributed Stochastic Neighborhood Embedding for embedding high dimensional results in a 2-D map.
Key insights into Scalable Computing: Down-Sampling images/frames in a video from 450 to 100 to reduce computational cost (for CNN); Using DAN (modification of traditional CNN) for reducing dimensionality of final feature & accelerated model training (as no traditional fully connected layers); Audio Modality: training of linear regressor by Torch (framework provides easy to use Multi-GPU support and parallelizing Neural Networks); Scaling Synaptic Weights: sufficient spiking and subsequent effective learning; Algorithm-System-Co-design for scalability of multi GPU system(e.g. : changing data's physical locations: reduced communication overhead).