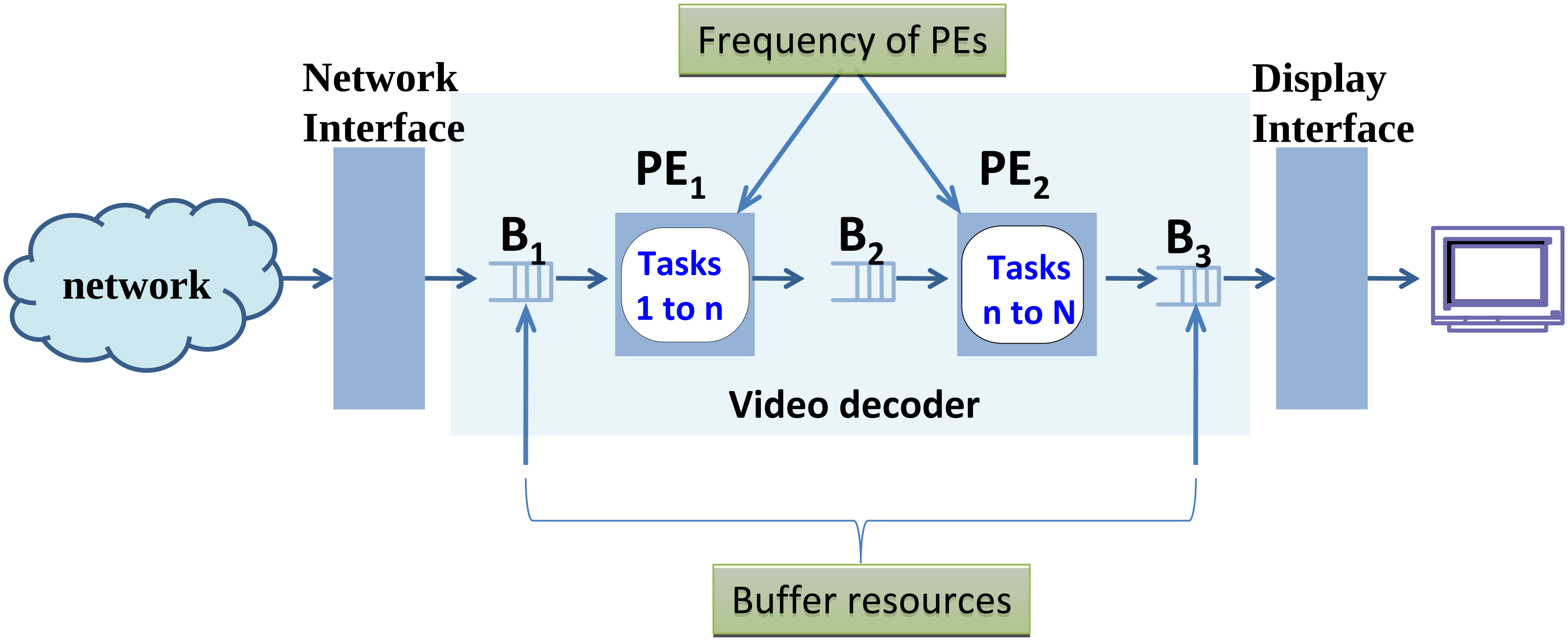
Fast Hybrid Simulation for Accurate Decoded Video Quality Assessment on MPSoC Platforms with Resource Constraints

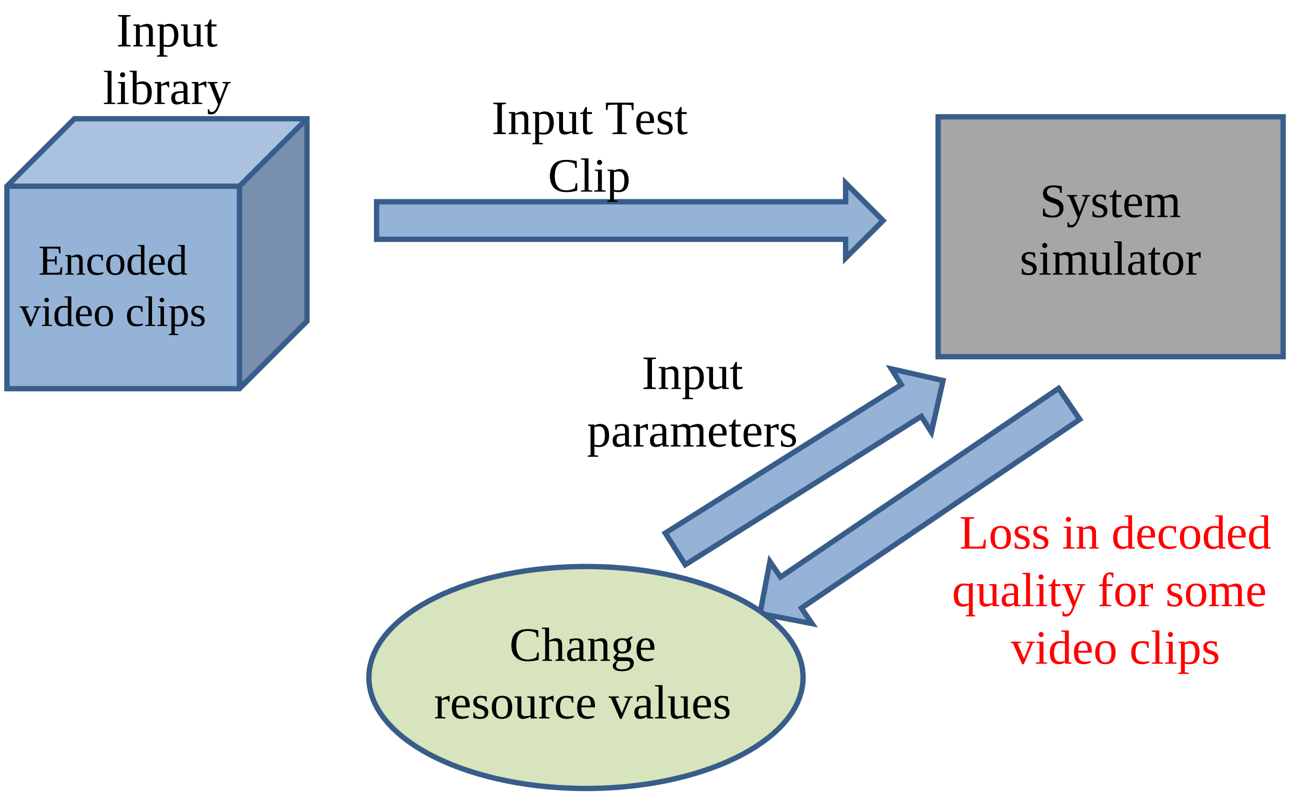
Deepak Gangadharan and Roger Zimmermann
Department of Computer Science, National University of Singapore
and
Samarjit Chakraborty
Institute for Real-Time Computer Systems, TU Munich

Motivation (Typical MPSoC architecture for video decoders)



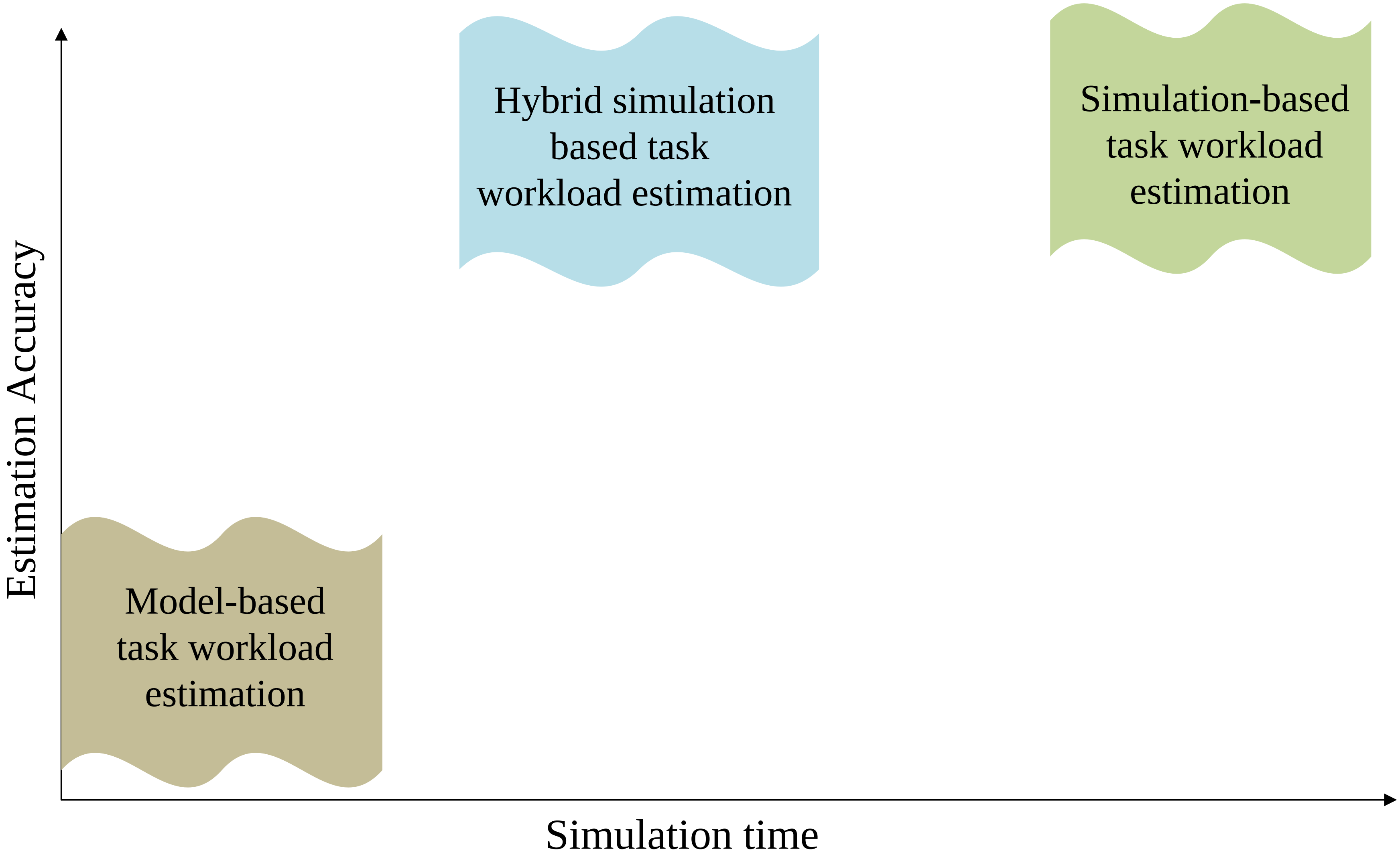
- Mobile devices with video decoders have resource (frequency, buffer etc) constraints
- Resources can be determined based on the output quality requirement
- Necessity for an efficient framework to explore the trade-off between resources and display quality

Motivation (Conventional method - Design space exploration of resources for no loss in quality)



- System should satisfy the “no quality loss” constraint for all test video clips
- A model of the system running a decoder application is used in the system simulator
- All video clips simulated to obtain decoder task workload values
- High simulation time!!

Motivation (Reducing simulation time)



Organization

- Related Work
- Overview of hybrid simulation-based quality assessment framework
- MPEG2/MPEG4 workload models
- Frame discard strategy
- Experimental results

Related Work

- Very little work on studying the trade-offs between scarce system resources (frequency and buffer) and decoded video quality important for mobile devices
- Yanhong et al. [1] investigate trade-offs between quality of MPEG-4 decoded video and processor frequency

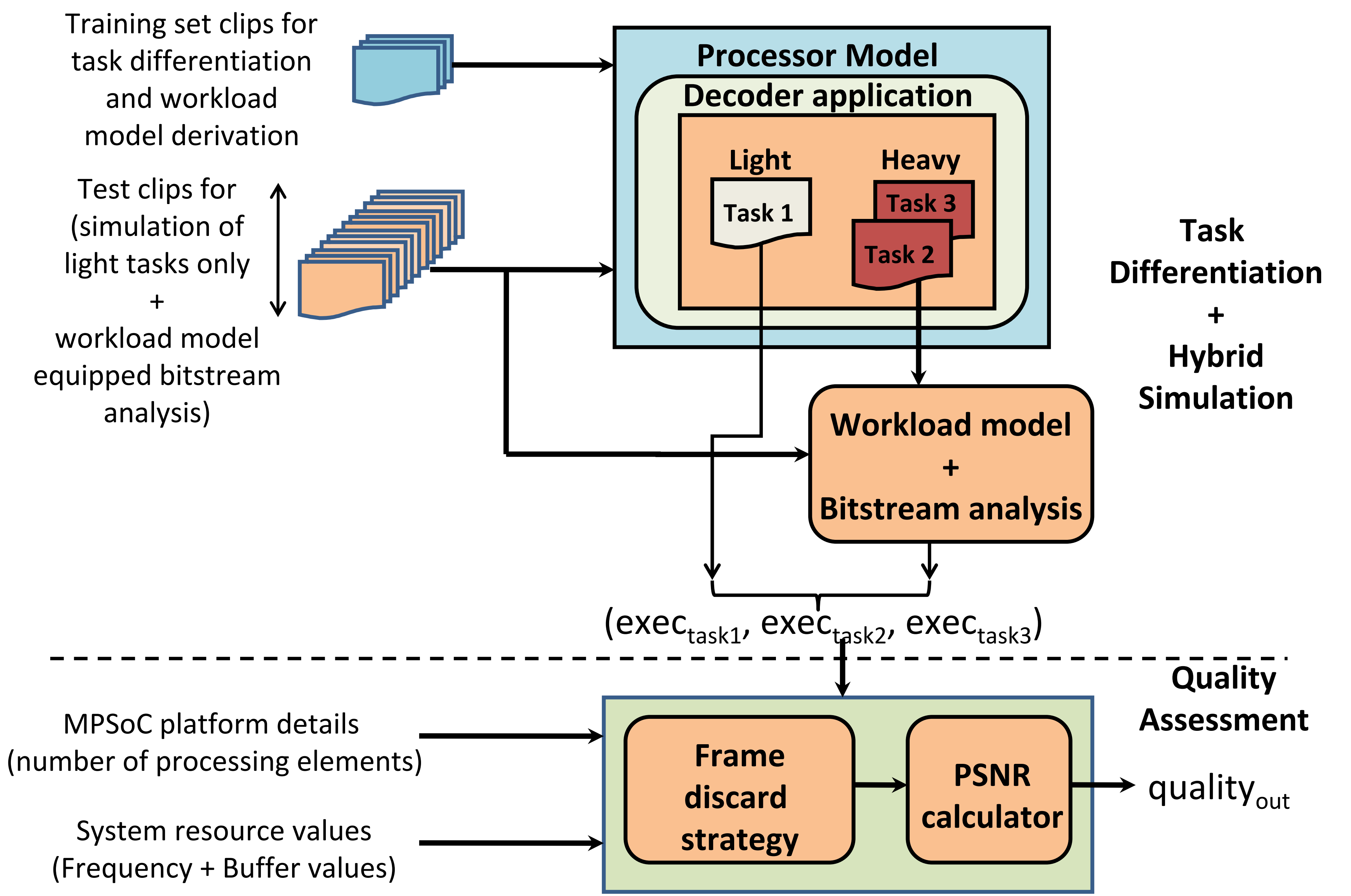
  - Use expensive simplescalar simulations to find task workload values
- Koumaras et al. [2] propose an end-to-end video quality prediction framework with packet losses in a network transmission scenario

  - Cannot be adapted to mobile devices because PSNR value estimation is not accurate

[1] L. Yanhong, S. Chakraborty, O. W. Tsang, A. Gupta, and S. Mohan. **Workload characterization and cost-quality tradeoffs in mpeg-4 decoding on resource-constrained devices**. In *3rd IEEE Workshop on Embedded Systems for Real-Time Multimedia (ESTIMedia), 2005.*

[2] H. Koumaras, C. H. Lin, C. K. Shieh, and A. Kourtis. **A framework for end-to-end video quality prediction of mpeg video**. *Journal of Visual Communication and Image Representation,* 21(2):139–154, 2010.

Overview of our framework



Step 1 - Task Differentiation

- Tasks in a decoder application are differentiated based on a **lightness** property of the tasks (using training set video clips)

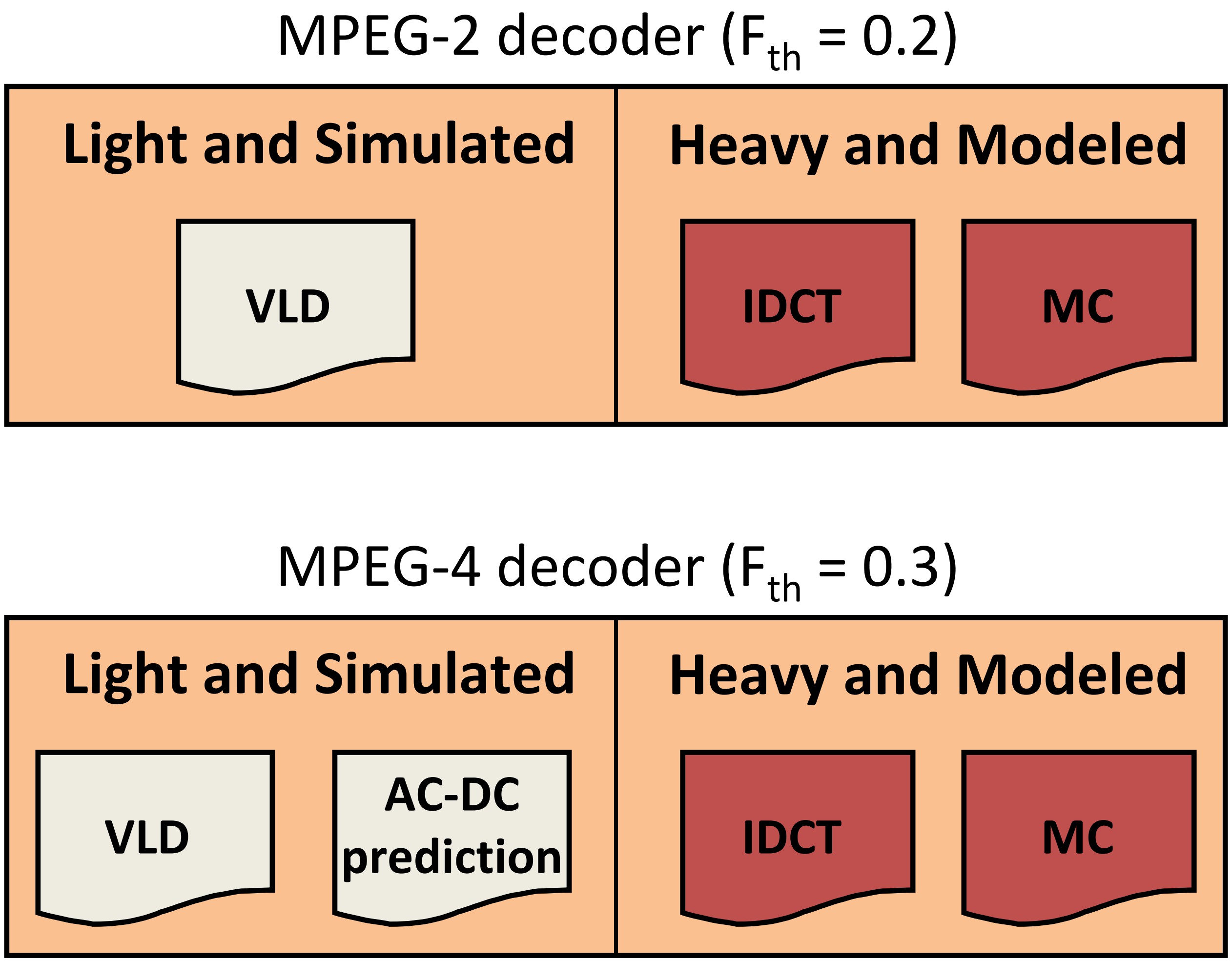
- For a task t_k, ;k where $1 \cdot k \cdot N_T$, **lightness** property is determined by

$$T_k \cdot F_{th} \text{ £ } T_{tot}$$

- Light tasks and other tasks which cannot be accurately modeled are simulated

- Accurate task workload models are derived for the remaining tasks

MPEG-2/MPEG-4 decoder task differentiation



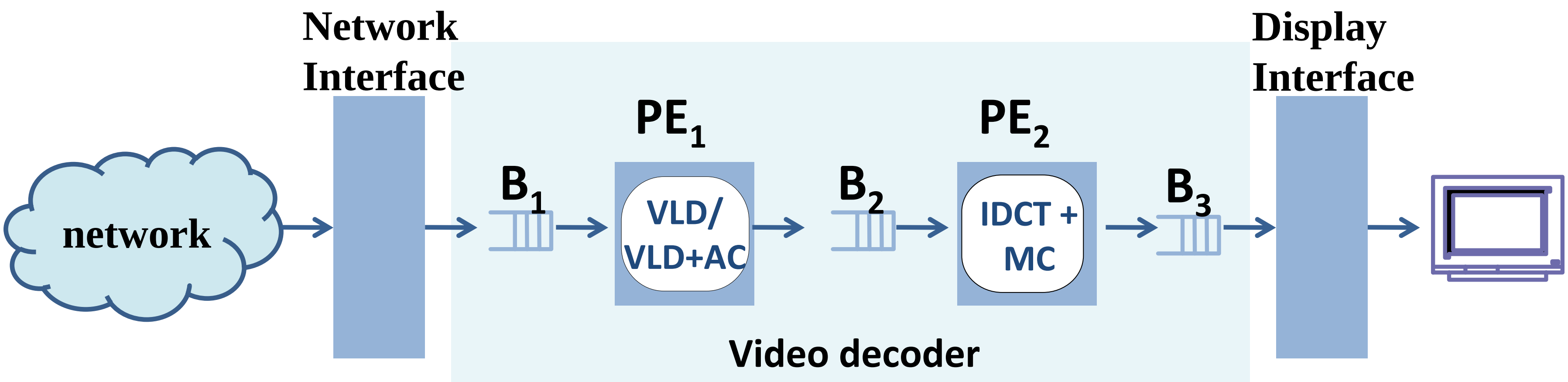
Step 2 - Hybrid Simulation

- Workload of heavy tasks (IDCT+MC for both MPEG-2 and MPEG-4 decoder) from Step 1 are estimated by extracting certain easily obtainable parameters from the video stream

- Accurate workload models are used for IDCT and MC (accuracy measured in terms of a **frame drop deviation** condition)

- Workload of light tasks from Step 1 are estimated by simulation as deriving accurate workload model is difficult

Frame Drop Deviation (FDD)



B_1 – Input buffer, **B_2** – Intermediate buffer, **B_3** - Playout buffer

- $b_{sim}(t) = r_{f1} + B\_bot_{sim}$ and $b_{model}(t) = s_{f2} + B\_bot_{model}$, ; t
  r_{f1} – r no. of MBs of (f1)-th frame at the top of the buffer and 0 ‰ r ‰ FSIZE
  s_{f2} has the same interpretation as r_{f1}

- According to FDD, everytime a buffer overflow occurs, FDD = 0 => |f1 – f2| = 0 => f1 = f2

Workload model for heavy tasks in MPEG-2/MPEG-4

- Video clips from training set used to derive the workload models
- Workload models were developed for portable ISA (PISA).
- VLD task workload **accuracy** is difficult to achieve by using a model in both MPEG-2 and MPEG-4. AC-DC prediction workload accuracy is also difficult to attain in MPEG-4.

MPEG-2 MC workload model

- MC workload depends on 6 parameters of the MPEG-2 frame namely

  1) Y component's x-dimension is HALF-PIXEL

  2) Y component's y-dimension is HALF-PIXEL

  3 & 4) U or V component's x and y-dimension is HALF-PIXEL

  5) forward or backward motion compensation is required

  6) motion compensation window is 16x8 or 16x16

- Depending on the frame type, the MC routine is called with different parameters

- Look-up table with 64 processor workload values is used to model the MC workload

  $Workload_{MC}$ = LUT(x), where x is a 6-bit value

MPEG-2 IDCT workload model

- IDCT workload depends on the number and position of non zero IDCT coefficients in the 8x8 block structure of MB

- To adhere to FDD condition, certain shortcut conditions of fast IDCT implementations are also taken into consideration

- The IDCT workload model proposed is

  Workload_i = basis_i + ® £ n_sig_i - ¯ £ scnt_i, ; i

  Workload_{IDCT} = $\sum_{\iota}$ Workload_i, 1 · i · 6

- From training set simulations
  basis_i = 1965/1852/708/595
  ® = 113 and ¯ = 143

Frame discard strategy

- Multimedia frames are required to be discarded when the resources are insufficient to process them
- Frame discard strategy – Simple mechanism of dropping entire frame if one incoming MB cannot be accommodated in the buffer
- Dropped frames are replaced by previous accepted frames

Frame discard strategy (...contd)

- Rule 1: An I-frame drop is followed by drop of an entire group of pictures (GOP). No. of drops will be a minimum of GOP length
- Rule 2: A P-frame drop is followed by drop of the subsequent frames in the GOP. No. of drops will depend on the position of P-frame in the GOP
- Rule 3: A B-frame drop does not require any other frame drop
- Rule 4: A frame is dropped if any of its MBs causes overflow or it is dependent on a previously dropped frame

PSNR calculation

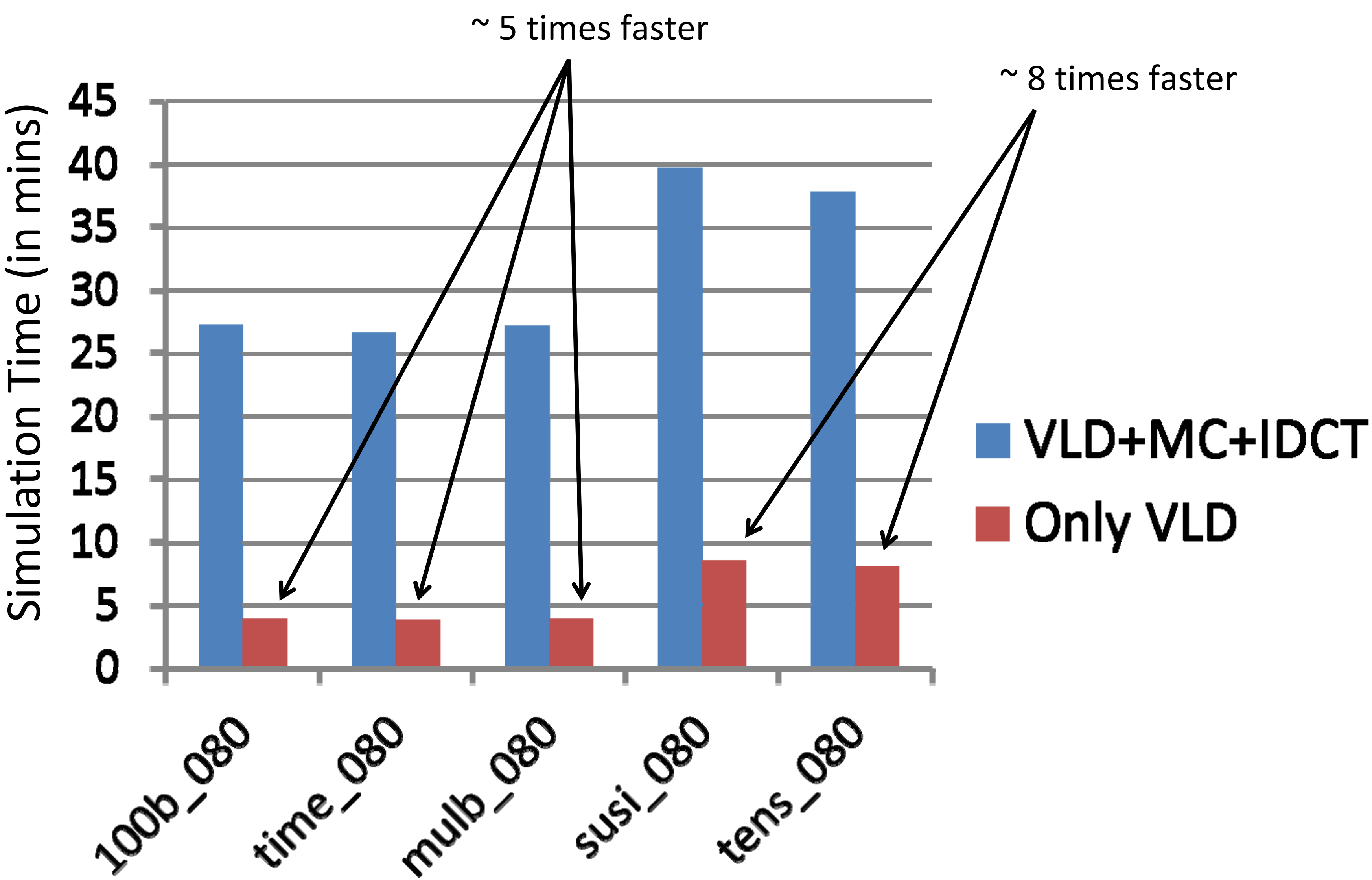
- A difference based PSNR (quality metric) is calculated as the original video frames are not available
- Noise – Mean square error (MSE) between the actual pixel values of dropped frame and pixel values of last accepted frame in display order
- PSNR value calculated using

$$psnr = 10 \times log_{10} \frac{(255 \times 255)}{\frac{(MSE\_r + MSE\_g + MSE\_b)}{(3 \times N_{tot} \times W \times H)}}$$

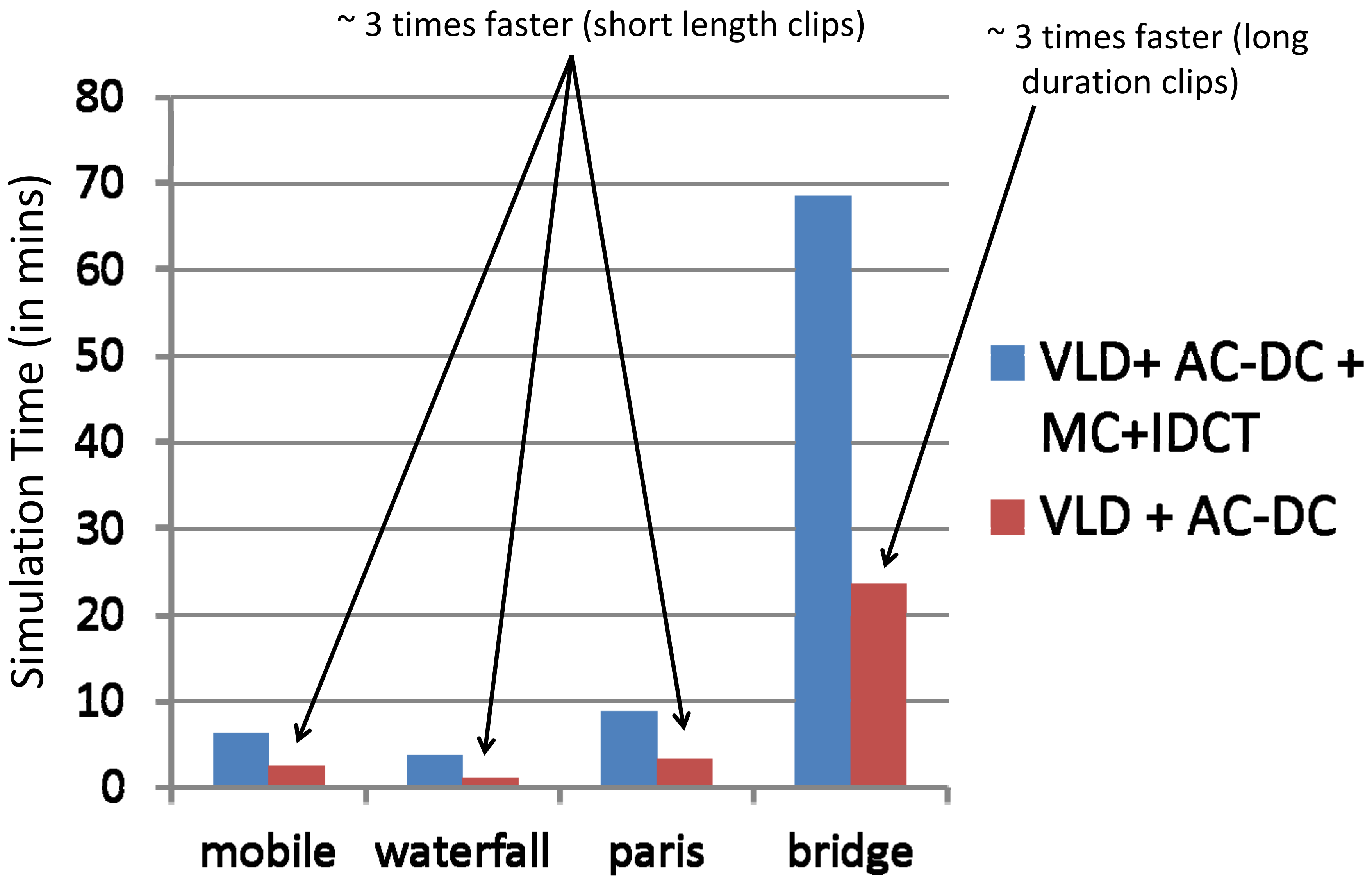
$$(MSE\_r)_n \quad = \quad \sum_{w=0}^{W-1} \sum_{h=0}^{H-1} \left(r_d(h,w,n) - r_c(h,w,n)\right)^2$$

$$MSE\_r \; = \; \sum_{n=0}^{N_{drop}-1} (MSE\_r)_n$$

Experimental Results
(Speed of Hybrid Simulation for MPEG-2)



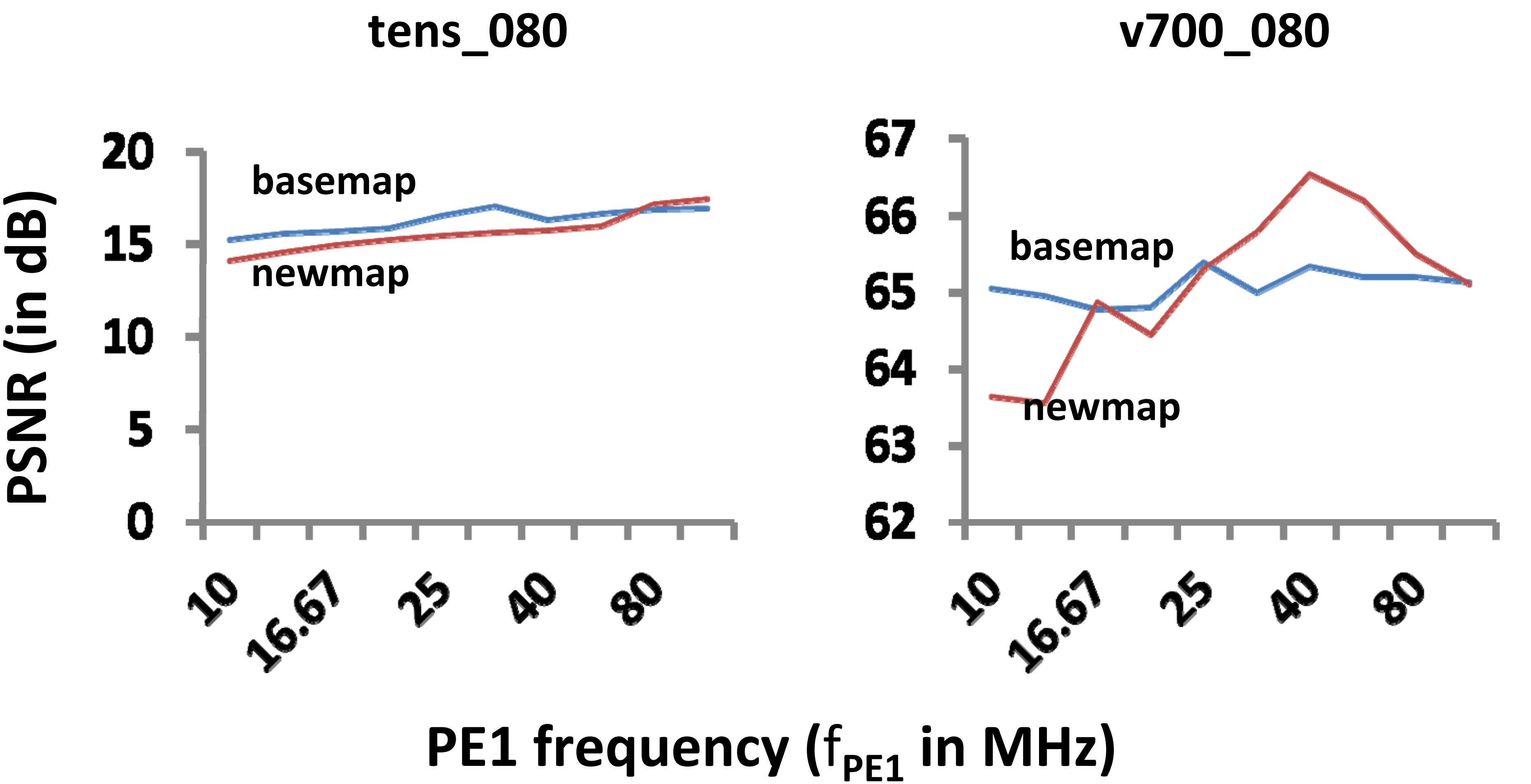
Experimental Results
(Speed of Hybrid Simulation for MPEG-4)



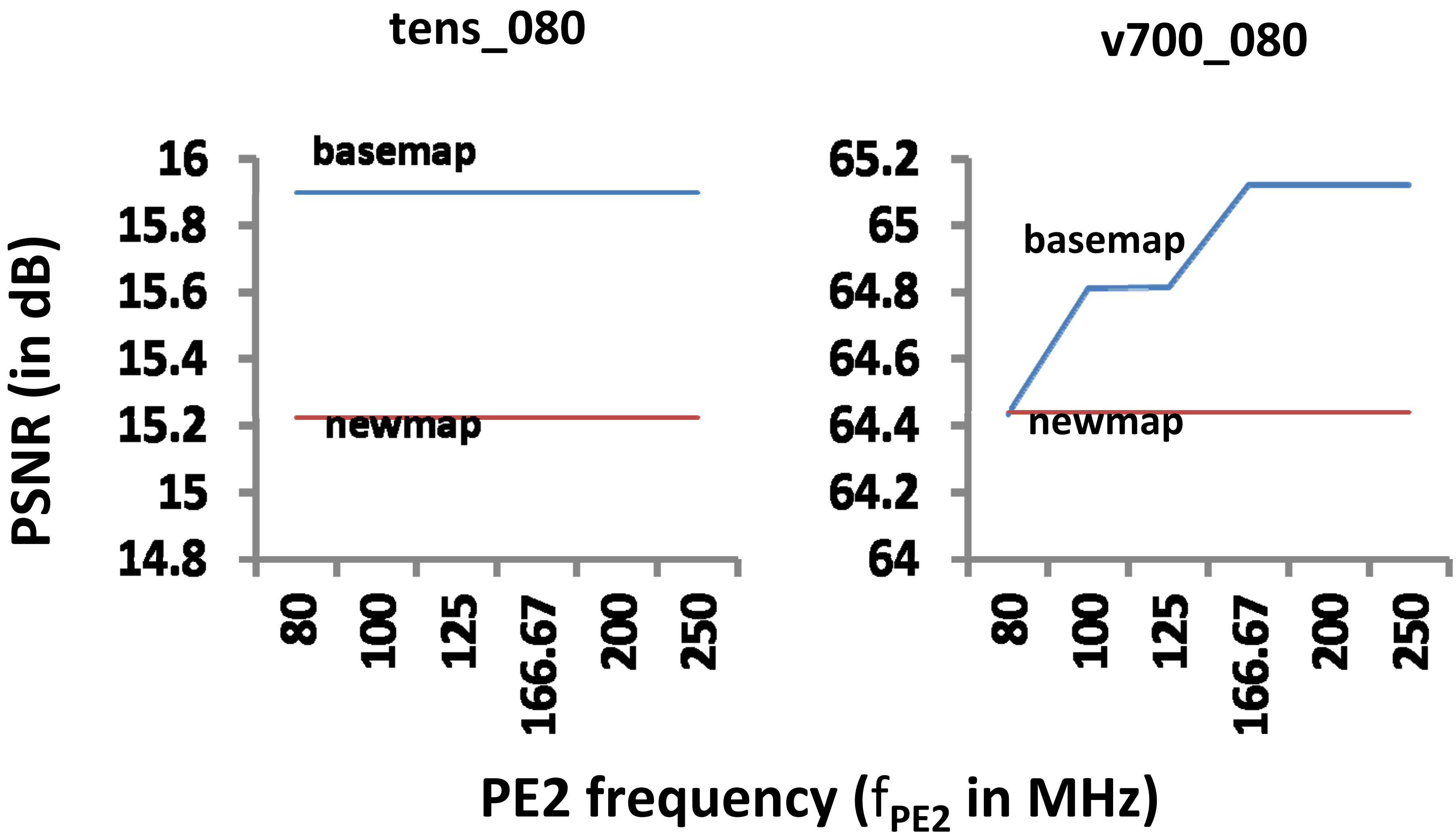
Experimental Results (Quality Estimation Accuracy)

- PSNR values obtained using model-based workload values were verified using simulation-based workload values
- PSNR values computed for 1500 resource combinations of f_{PE1}, f_{PE2}, B_1 and B_2
- PSNR values found for the *basemap* as well as *newmap* task mapping configurations
- *basemap* – 98% accuracy for motion video clips. Only § 0.3% deviation in PSNR values for the remaining 2%. 100% accuracy for the still videos.
- *newmap* – 100% accuracy in PSNR estimation for all clips

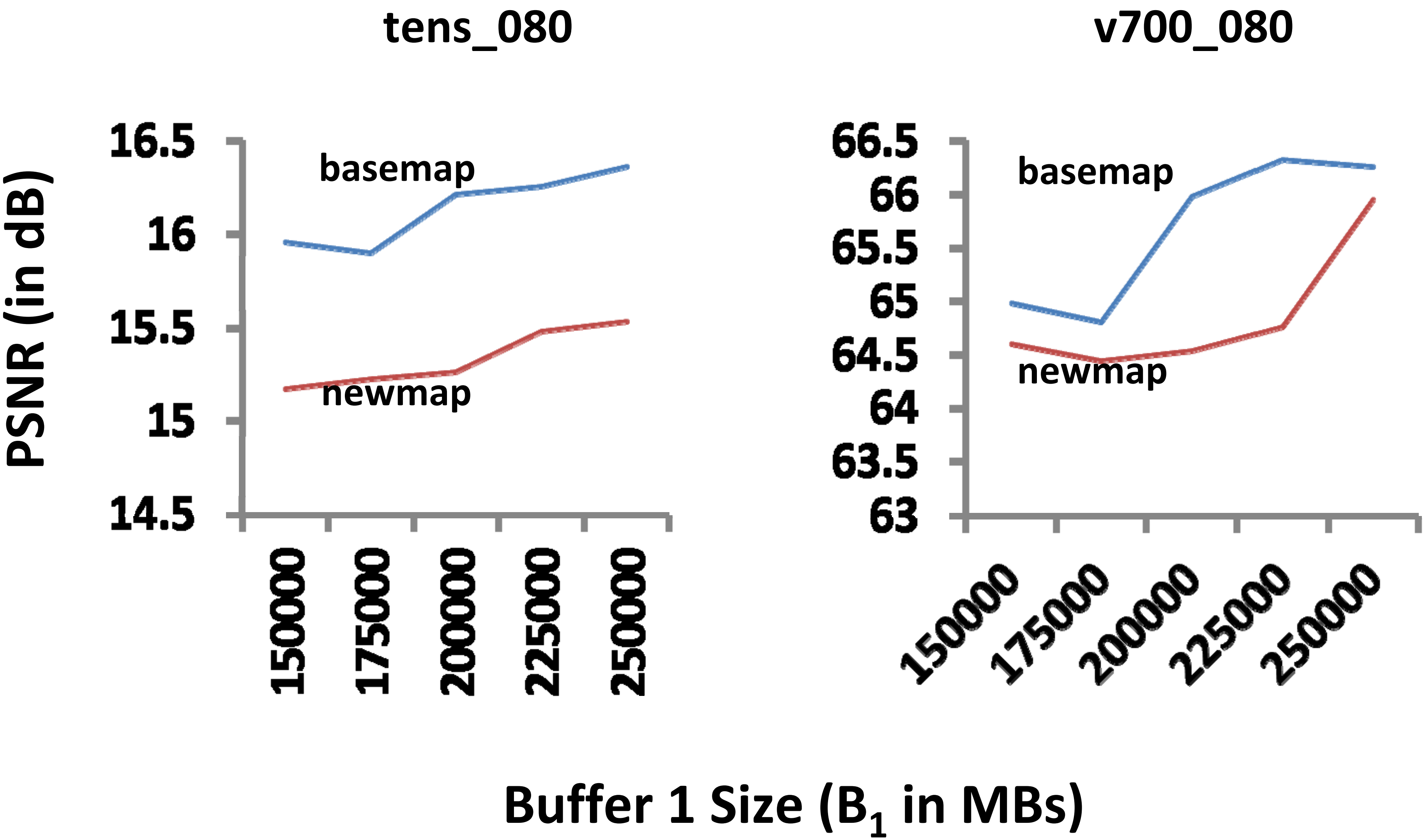
Experimental Results (PSNR vs f_{PE1} trade-off)



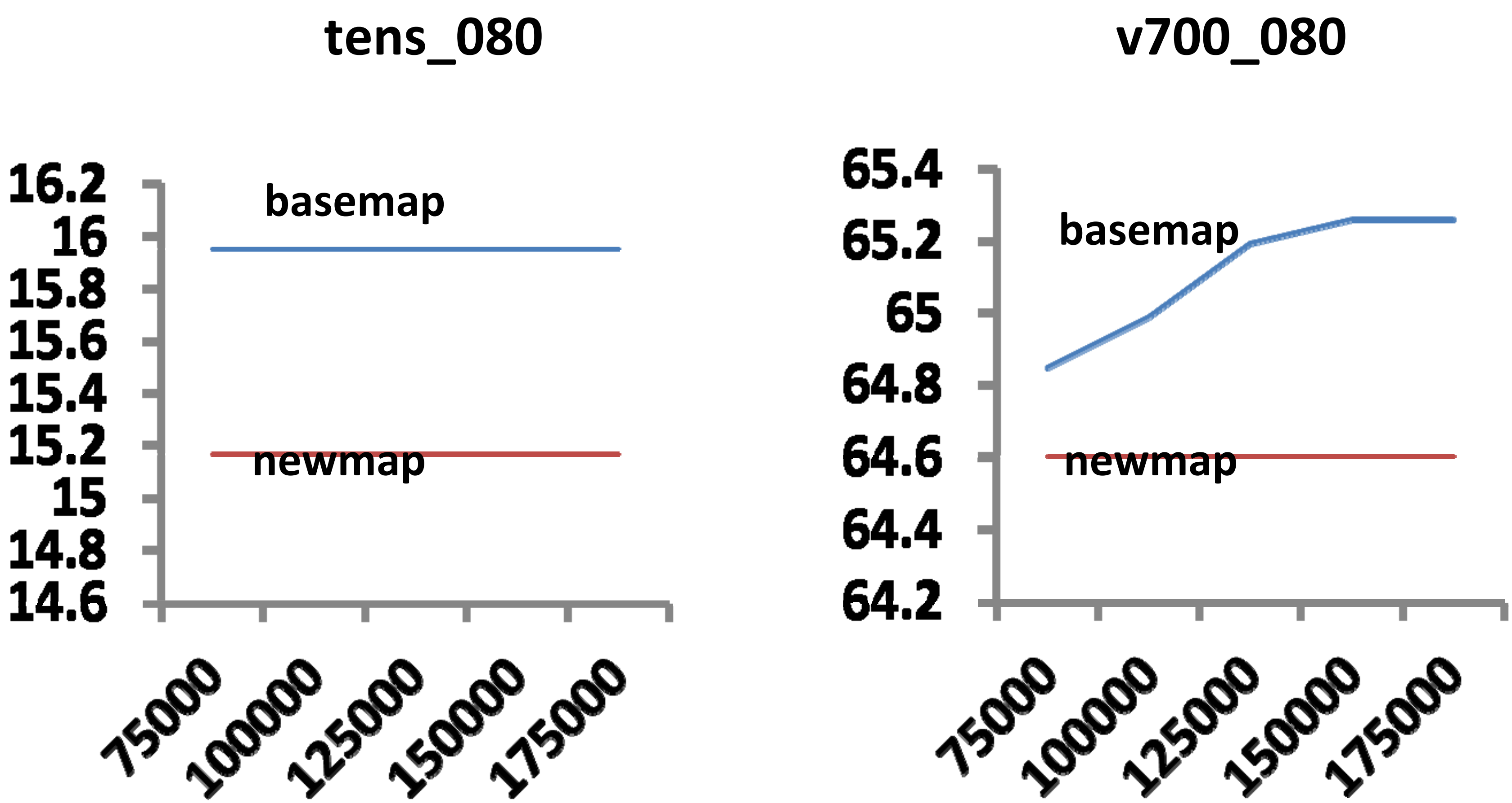
Experimental Results (PSNR vs f_{PE2} trade-off)



Experimental Results (PSNR vs B_1 trade-off)



Experimental Results
(PSNR vs B_2 trade-off)



PSNR (in dB)

Buffer 2 Size (B_2 in MBs)

Concluding Remarks

- **Central focus**:

  To devise a hybrid simulation strategy in order to enable the system designers to rapidly arrive at quantitative estimates of quality degradations for video clips in the context of resource constraints

- Framework also very useful to understand the non-trivial influences of system resources on the quantitative quality degradations

- Can be used to quickly estimate optimal points in resource space for desired quality

- Influence of other components in the framework such as frame discard strategy on the quality estimate can be studied rapidly