Total time taken: 20s + 45s + 10s + 4(30s) + 45s = 4m15s ≈ 4m40s

Actual time taken: 30s + 1m30s + 10s + 4(30s) + 1min + 7(10s)= 6m20s ≈ 6m

**Introduction**

***Time***: **20 secs**

**[*SCREEN*]**: Full screen

Points to note: [For script content]

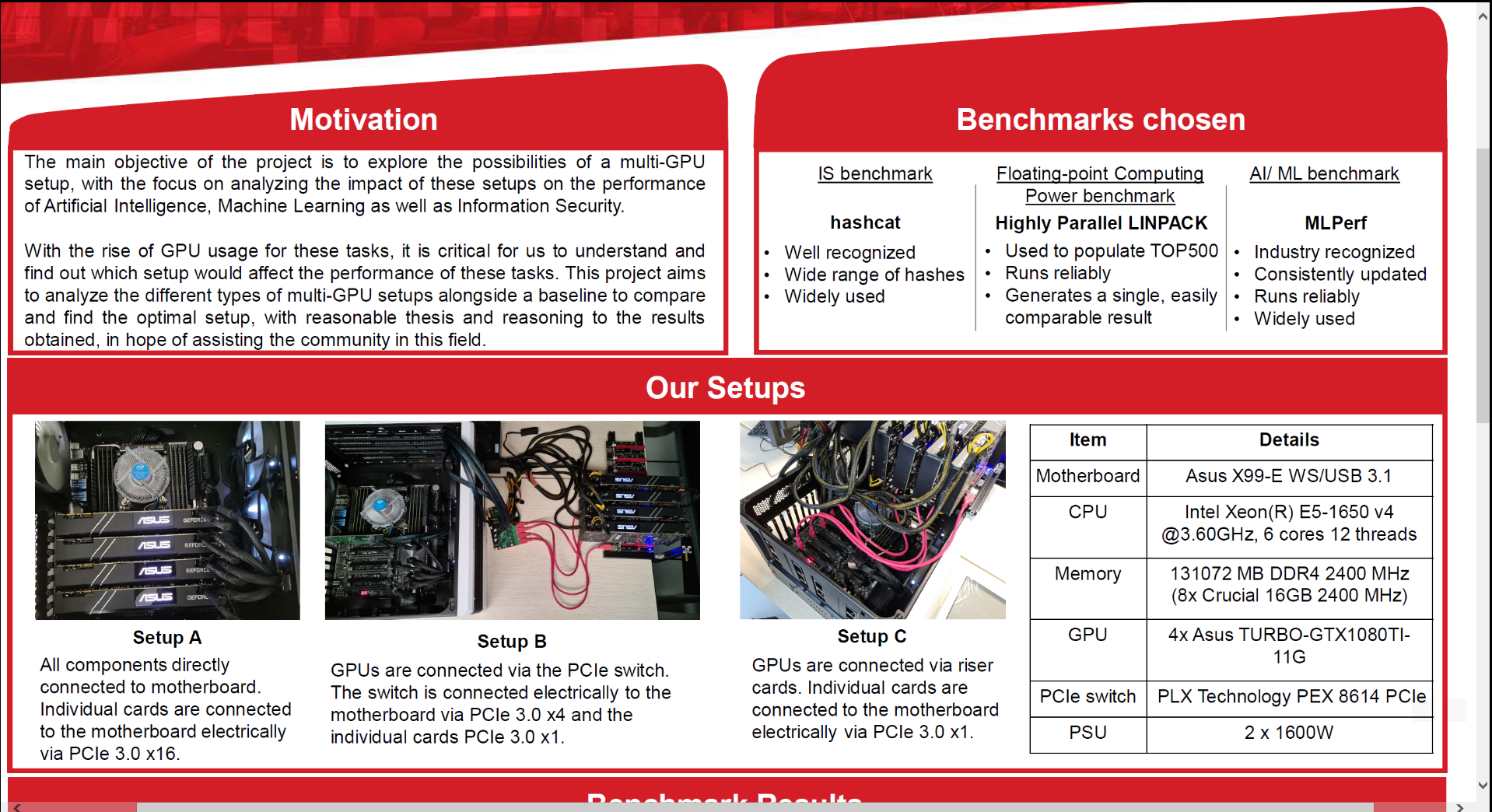
* Acad sup, Industry sup
* Project group
* Name of project

Points to be said in video in order:

1. Hi Professors, our project group number for ITP is IS-10 and our group have been working on the project ‘Design and Performance Analysis of a Portable Multi-GPU Unit for Artificial Intelligence, Machine Learning and Information Security’
2. The academic supervisor for this project is Assistant Professor Goh Weihan and the industry supervisor is Arthur Loo Wee Yeong

**Motivation - Benchmarks - Setups**

***Time***: **45 secs**

**[*SCREEN*]**: 

Points to note: [For script content]

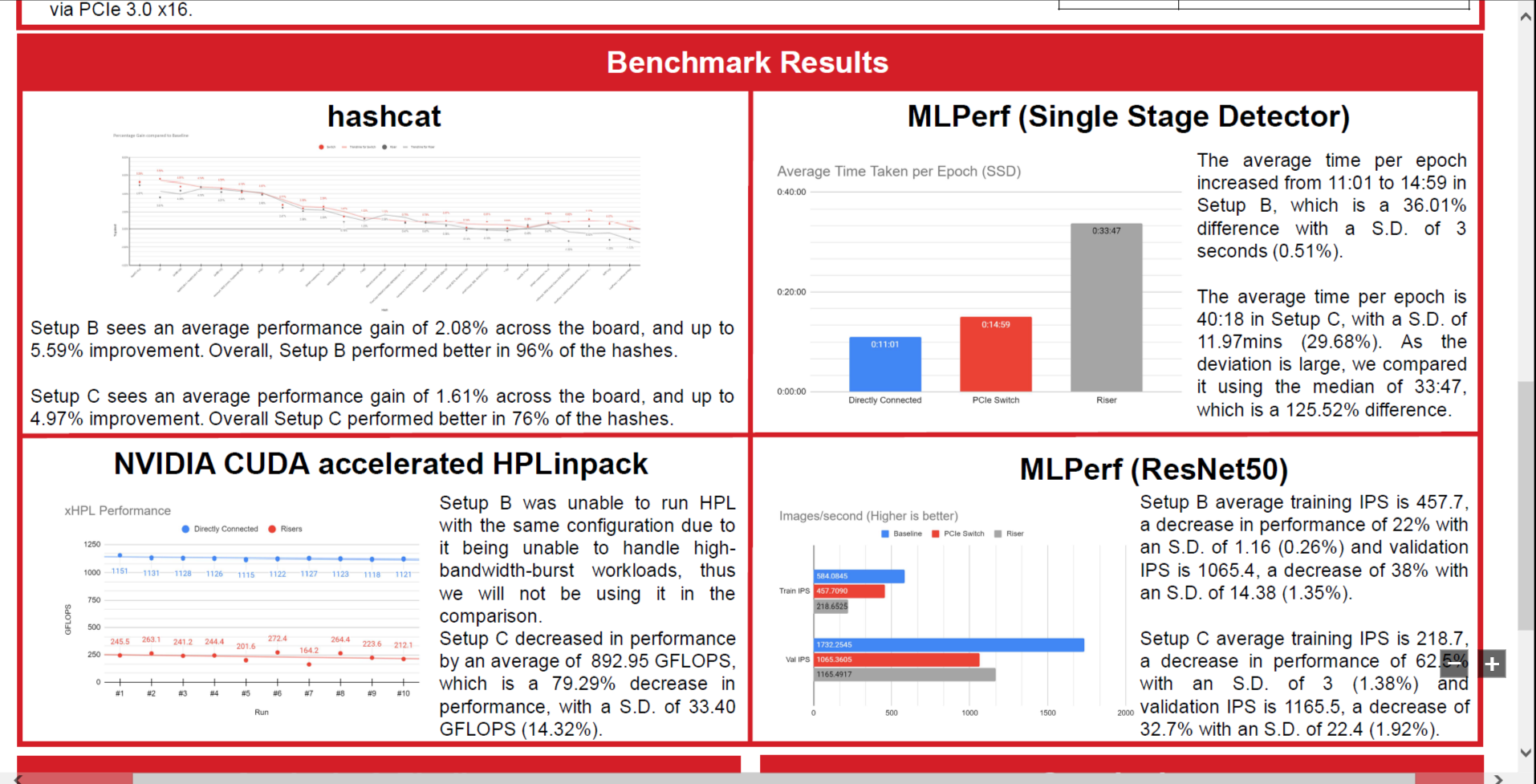
* State project aim
* State that we have included hardware specs of our system [no need to specify in video]
* Briefly state the 3 setups
* State the benchmarks we used
* Mention setup A is baseline across all benchmarks

Points to be said in video in order:

1. The aim of our project is to analyse the 3 different multi-GPU setups to find the most optimal setup among the 3 - after a thorough analysis of the results, backed with reasonable explanations with the hope of contributing to the community.
2. We have included the hardware specs of our system on the right side. Besides the PCIe switch, the rest are consistent across all our setups. Note that our multi-GPU system consists of 4x Asus Turbo-GTX1080TI GPUs.
3. Setup A consists of all components being directly connected to the motherboard. Individual cards are connected electrically via PCIe 3.0 x16.
4. In Setup B, the GPUs are connected externally via the PCIe switch. The switch is then connected electrically to the motherboard via a PCIe 3.0 x4 and the individual cards via PCIe 3.0 x1.
5. Setup C has the GPUs connected via riser cards. The riser cards are connected electrically to the motherboard via PCIe 3.0 x1.
6. The benchmarks we have successfully implemented to be able to run on our system target the domains of IS & AI/ML. They are hashcat, Highly Parallel Linpack and SSD (also known as Single Stage Detector) and ResNet50.
7. Both SSD and ResNet50 are benchmarks within MLPerf.
8. We run each benchmark more than once.[State reason for multiple runs?/Justify need for multiple runs?] Hashcat ran 10 times, HPlinpack 10 times, SSD ran 5 times and ResNet50 ran 3 times.
9. Note that we denote Setup A as our baseline when comparing the benchmark results.

**Benchmark Results**

***Time***: **10 secs**

**[*SCREEN*]**: 

Points to note: [For script content]

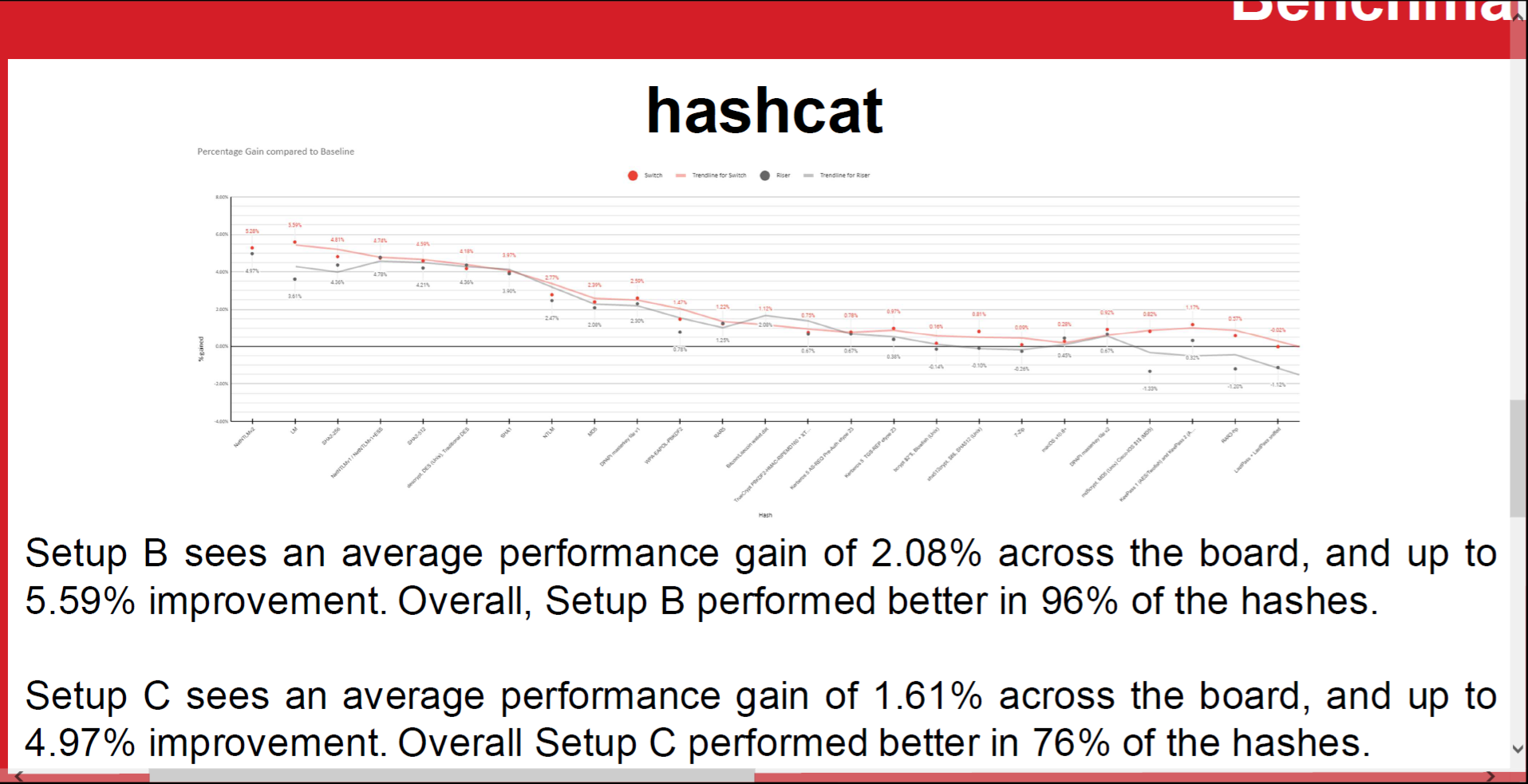
* State that
  + Performed multiple runs to ensure [reason]
  + Tabulated results from those into graphs
  + Perform analysis from given data

Points to be said in video in order:

1. Results are consolidated after performing multiple runs on the individual benchmarks.
2. They are then tabulated into graphs to highlight the differences across the 3 setups.
3. After which, we perform analysis of the given data and do research to justify those differences.

**Hashcat**

***Time***: **30 secs**

**[*SCREEN*]**: 

Points to note: [For script content]

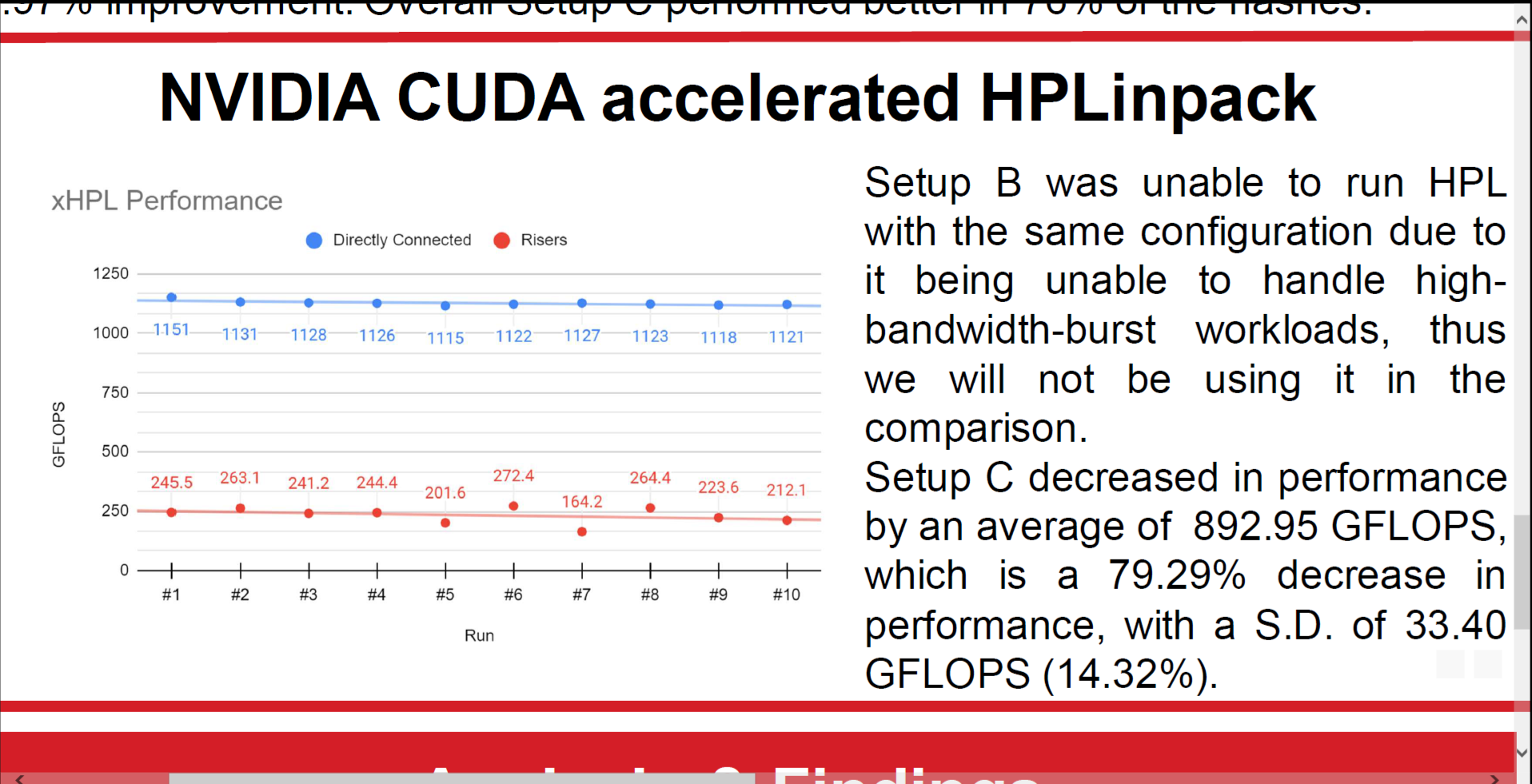
* Compare results against baseline. Take 5 highest percentage improvement against the baseline
* State diff setups avg perf gain
* Both setups perform better across % of hashes

Points to be said in video in order:

1. Comparing the results against the baseline, the graph displays the hashes of both setups B and C. Here, we display by order of highest to lowest % improvement from left to right.
2. Setup B shows an avg performance gain of 2.08-5.59% while Setup C shows an avg performance gain of 1.61-4.97%.
3. Setup B performed 96% better than the baseline while setup C performed 76% better than the baseline.

**HPLinpack**

***Time***: **30 secs**

**[*SCREEN*]**: 

Points to note: [For script content]

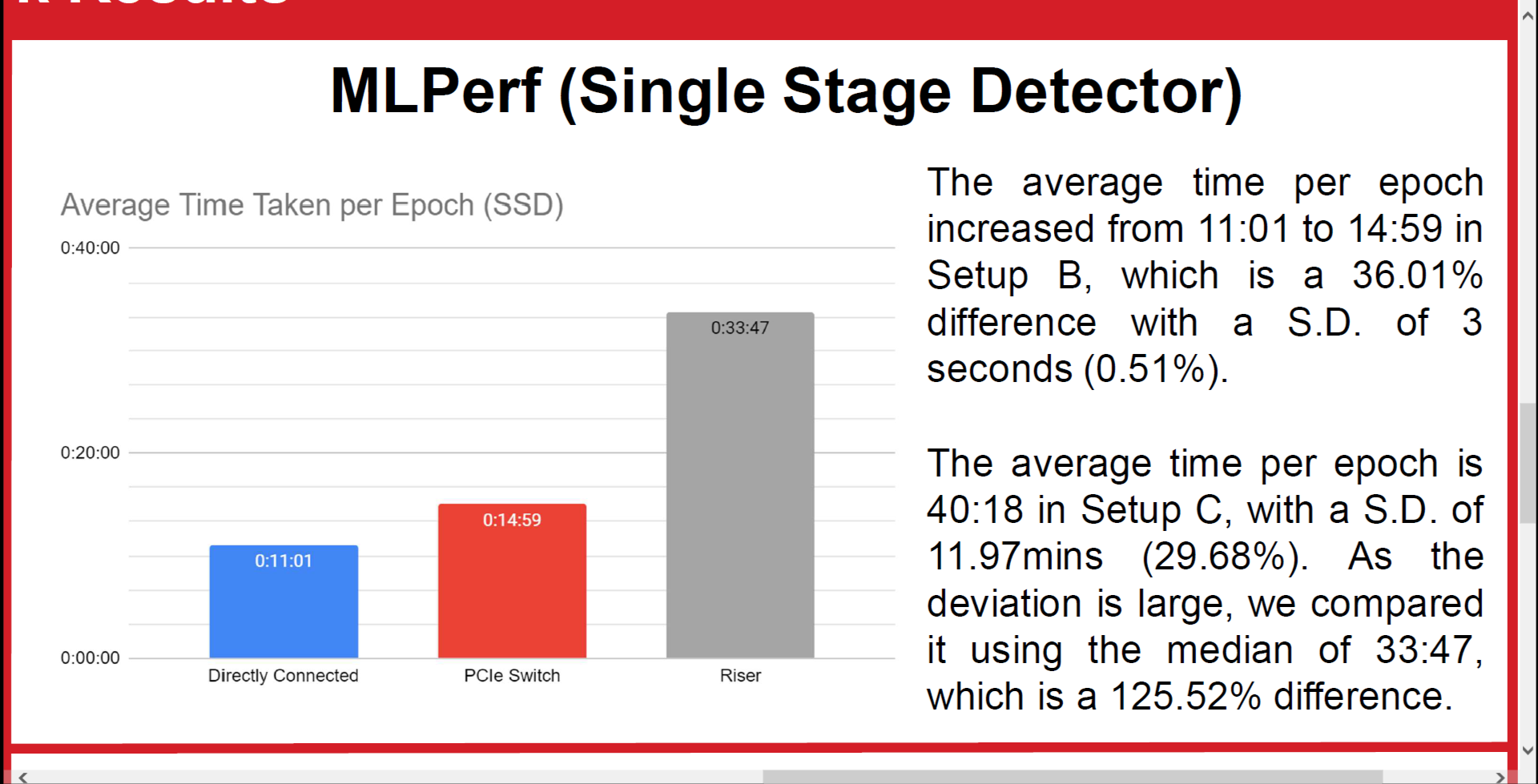
* A vs C. State B hardware limitation prob..
* B not included in the comparison
* C decreased perf by an avg of 892.95 GFLOPS - 79.29% decrease in perf

Points to be said in video in order:

1. For HPLinpack, the graph here only shows Setup C, which is the Riser Card-based setup against the baseline.
2. Setup B was not able to run since it cannot handle high-bandwidth-burst workloads and therefore won’t be included in the comparison.
3. From the graph, it shows that Setup C produced results worse than the baseline by an average of 892.95 GFLOPS. This is equal to about 79.29% decrease in performance. The standard deviation is 33.40 GFLOPS equivalent to 14.32%.[State what is a good range for standard deviation?]

**SSD**

***Time***: **30 secs**

**[*SCREEN*]**: 

Points to note: [For script content]

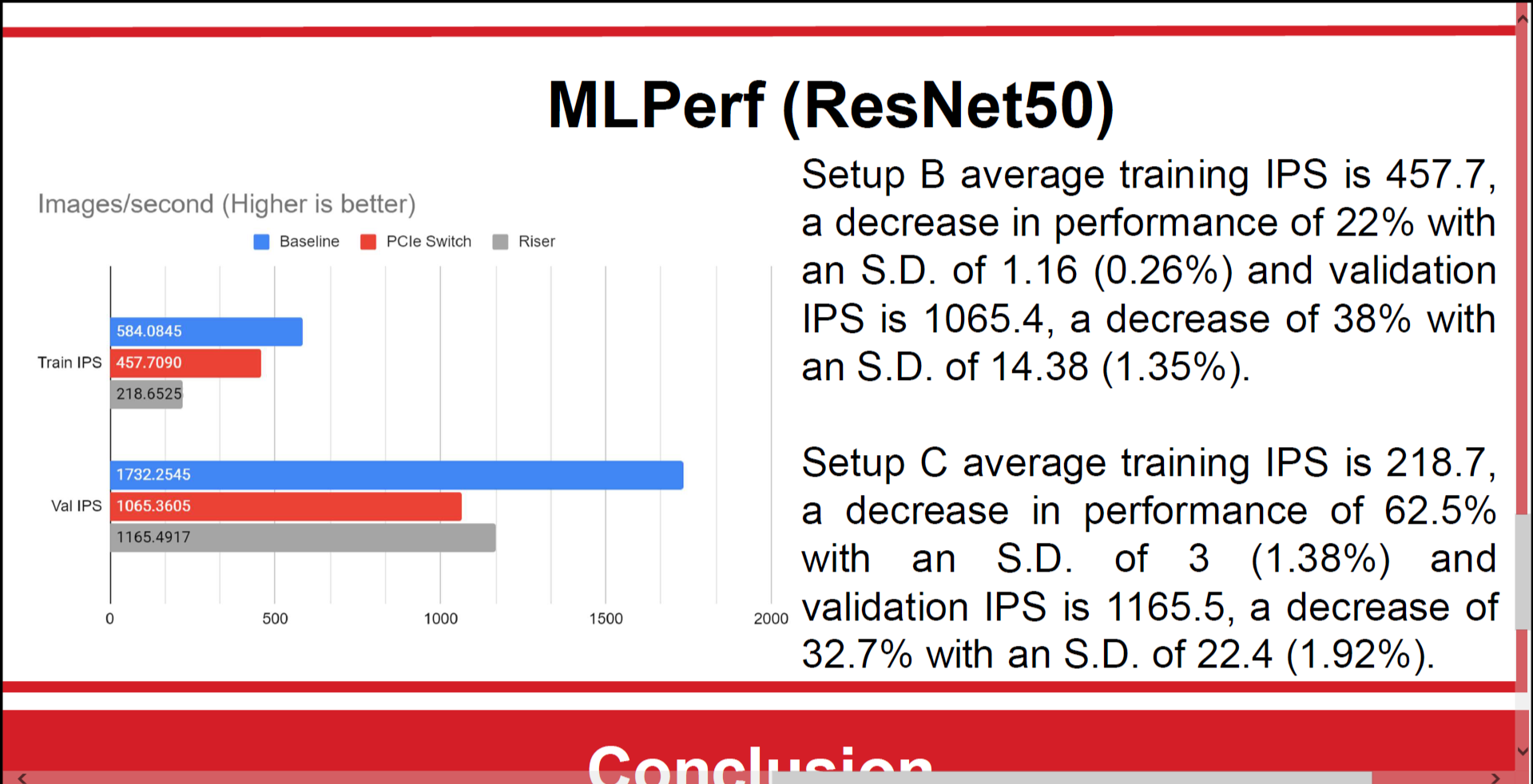
* Mention
  + Quality target
  + Avg time taken per epoch + %
  + S.D + %
* [Mention crashes?] -- Setup C proved to be unstable and takes a much longer time to run

Points to be said in video in order:

1. For SSD, the benchmark goes through multiple epochs to reach a certain quality target. The graph shows the average time taken per epoch for all 3 setups. The baseline achieved the best time with an average of 11:01 mins per epoch while Setups B and C took longer with Setup B at 14:59 mins, (this is an increase by 36.01%) and a low standard deviation of 3 seconds or 0.51%.
2. Setup C is seen to perform much slower than the baseline with an average time of 40:18 mins per epoch with a standard deviation of 11.97 mins equivalent to 29.68%.
3. Because of this large standard deviation, comparison is done instead using the median value which gives an average time of 33:47 mins per epoch and this is equivalent to 125.52% increase from the baseline.

**ResNet**

***Time***: **30 secs**

**[*SCREEN*]**: 

Points to note: [For script content]

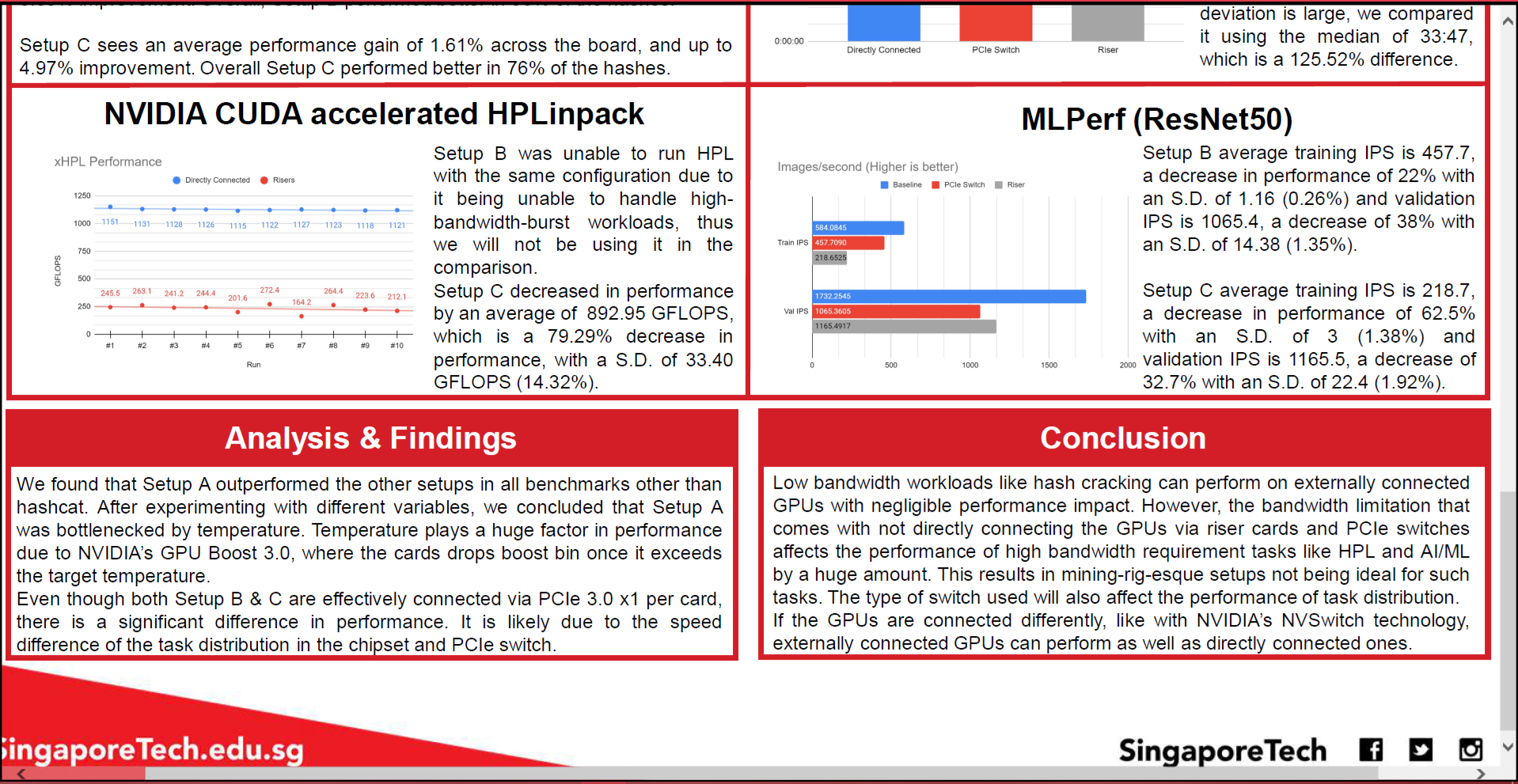
* Training benchmark - training perf result - training/train IPS (metric)
* Inference benchmark - inference perf result - validation/val IPS (metric)

Points to be said in video in order:

1. The ResNet50 benchmark of MLPerf is capable of benchmarking training and inference.
2. The graph displays the results for training and inference for each setup.
3. Setup B has an avg training IPS (IPS stands for Images per Second) of 457.7 which is a 22% decrease in performance and a standard deviation of 1.16 or 0.26%. Setup B inference produces a validation IPS of 1065.4 which is also a decrease in performance by 38% and a standard deviation of 14.38 or 1.35%.
4. Also a decrease in performance, Setup C produced an average training IPS of 218.7, a clear decrease of performance by 62.5%, with a standard deviation of 3 or 1.38%. Its inference results show that its validation IPS is also lower than the baseline with a value of 1165.5, which is a decrease of 32.7%. Its standard deviation is 22.4 which is about a spread of only 1.92%.

**Analysis & Findings + Conclusion?**

***Time***: **45 secs**

**[*SCREEN*]**: 

Points to note: [For script content]

* Read off poster for analysis
  + [Anything to add?]
* Read off poster for conclusion
  + [Anything to add?]

Points to be said in video in order:

1. [Read off poster for analysis and findings]
   1. After a thorough analysis, we found….
2. [Read off poster for conclusion]
   1. In this project, we conclude that...