Safely Overclocking Flash I/O in SSDs

Kai Zhao, SanDisk

Rino Mellini, PMC-Sierra

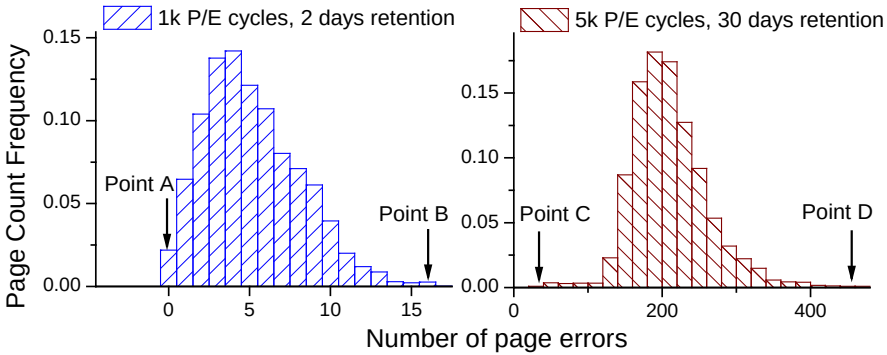
Tong Zhang, Rensselaer Polytechnic Institute

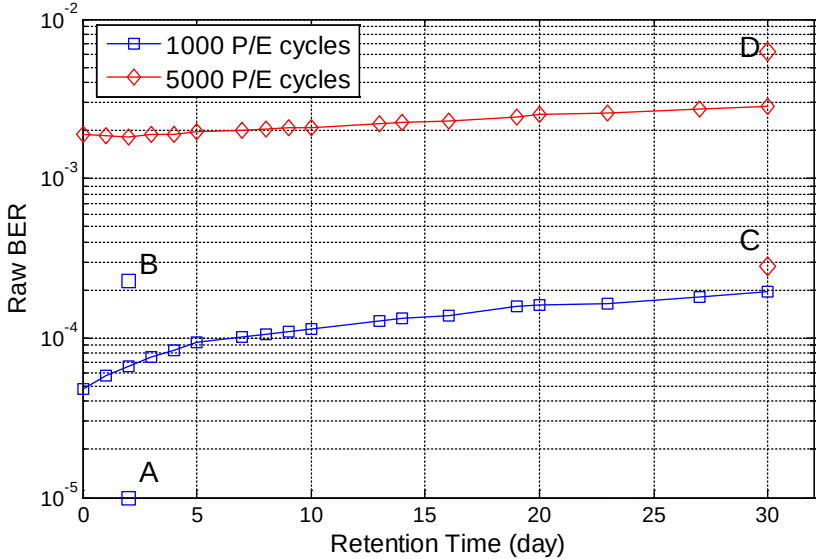
Outline

- Motivation – ECC capability may be under utilized
- Benefit – improve system performance, especially for read operations
- Reliability – leverage LDPC codes to mitigate the impact of overlocking

Page Reliability Variation

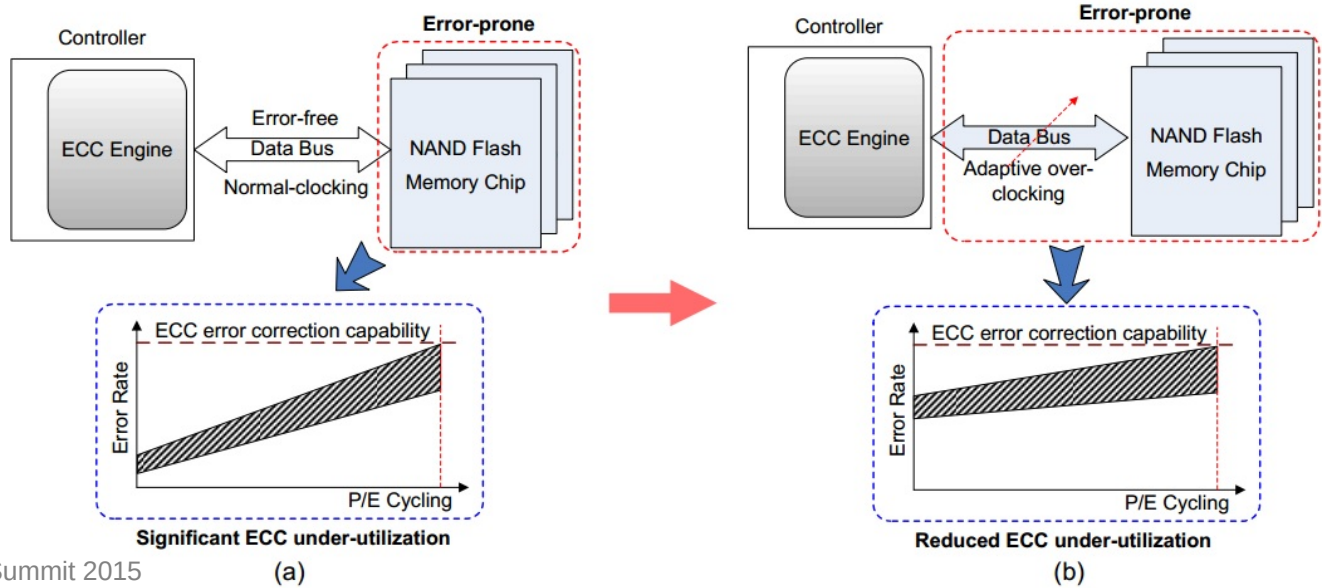
- Larger reliability variations for chips even from the same batch
- Raw BER varies dramatically under different P/E cycles and with different retention time





Data Link Overclocking

- Worst-case based ECC is under utilization in most of the time
- We could trade reliability (when raw BER is low) for performance
- One possible way – Overclocking on data link



Potential Benefits

- Total latency = Operation (read/program) latency +

  Data transfer latency + ECC decoding latency

$$\tau_{phy-acc} = \tau_{op} + \tau_{xfer} + \tau_{ECC}$$

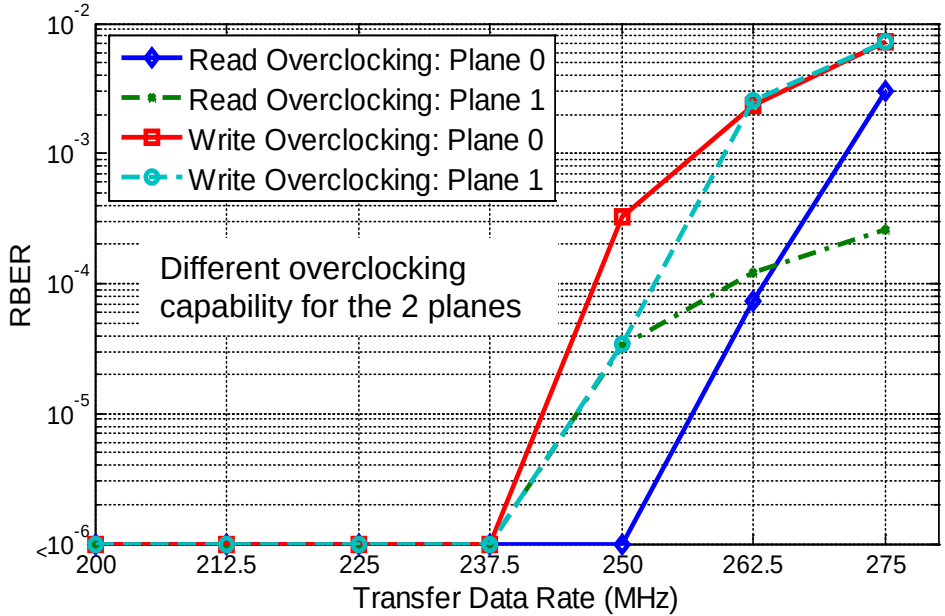
- ECC decoding latency is relatively small
- Transfer latency is comparable to read latency
- Smaller transfer latency could significantly improve read performance

Practical Considerations

- Characterize errors caused by overclocking
  - *Errors in the overclocked data link*
  - *Errors in the NAND flash cell array*
- Overclocking for read and/or write path
  - *Evaluate read/write path separately*
  - *Permanent errors on write path*
- Adaptation of overclocking in late lifetime of the drive
  - *Read retry*
  - *Multiple data transfer rate*

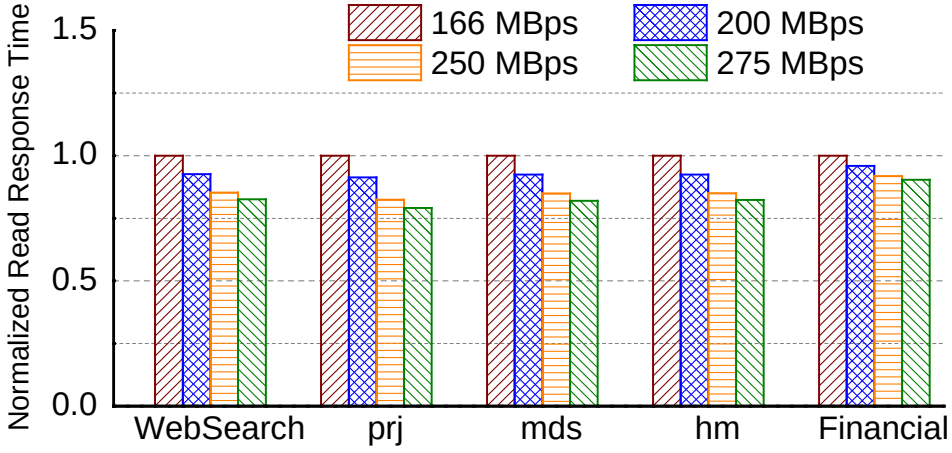
Feasibility of Overclocking

- ONFI 2.2 compliant chips supporting data transfer rate up to 166 MBPS according to the datasheet
- Data link overclocked up to 275 MBPS

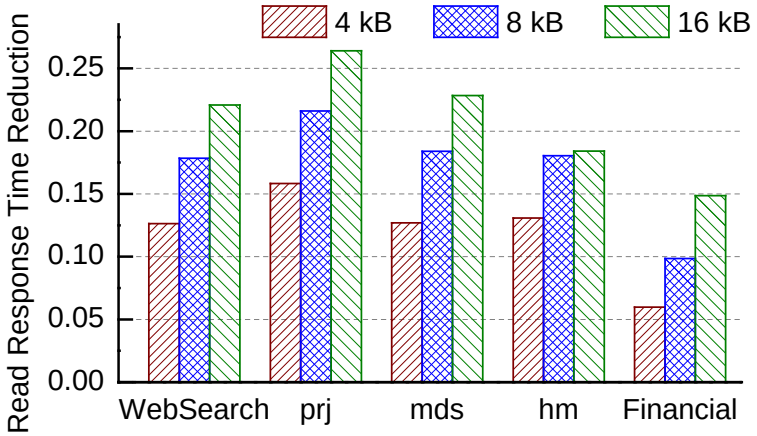


System Performance for Read Overclocking

- Read response time, page size is 8 kB
- Data link overclocked up to 275 MBPS



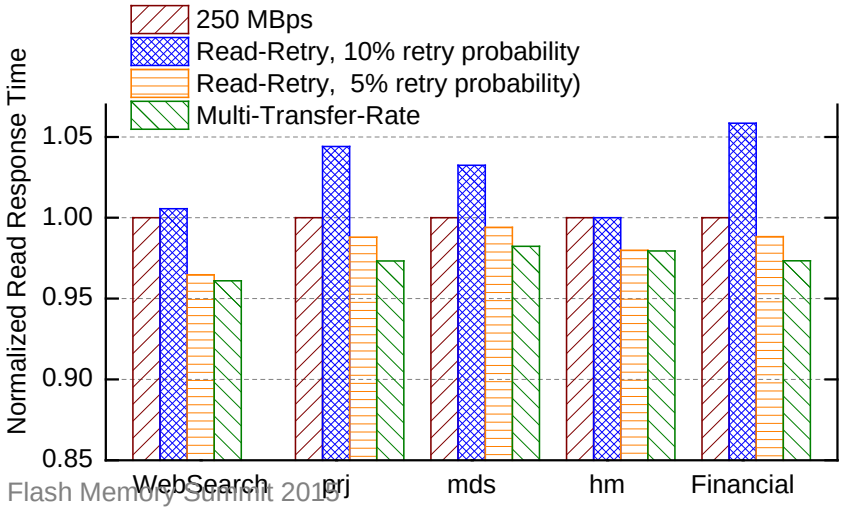
- Comparison of average read response time reduction with respect to page size
- Data transfer rate is 275 MBPS

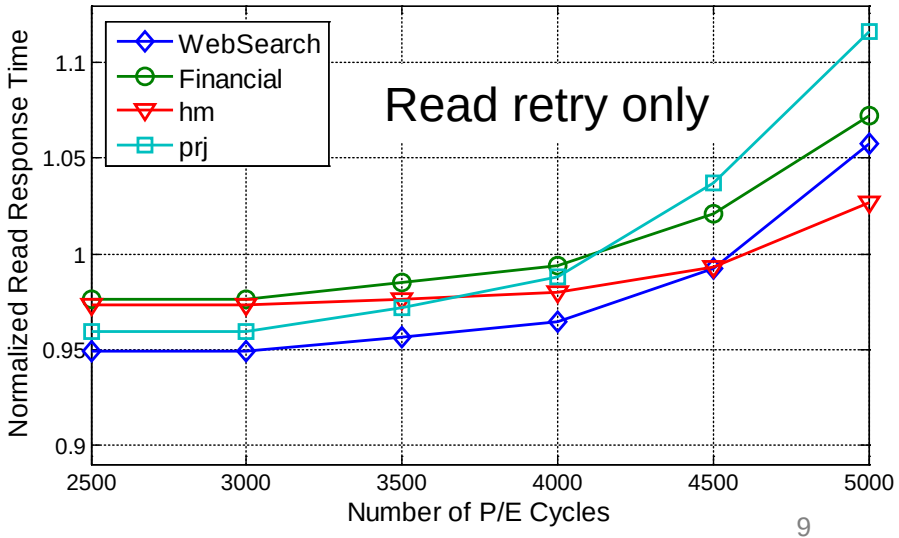


Simulator: DiskSim with SSD addon
Traces: MSR Cambridge trace, UMASS trace repository

Late Lifetime Performance

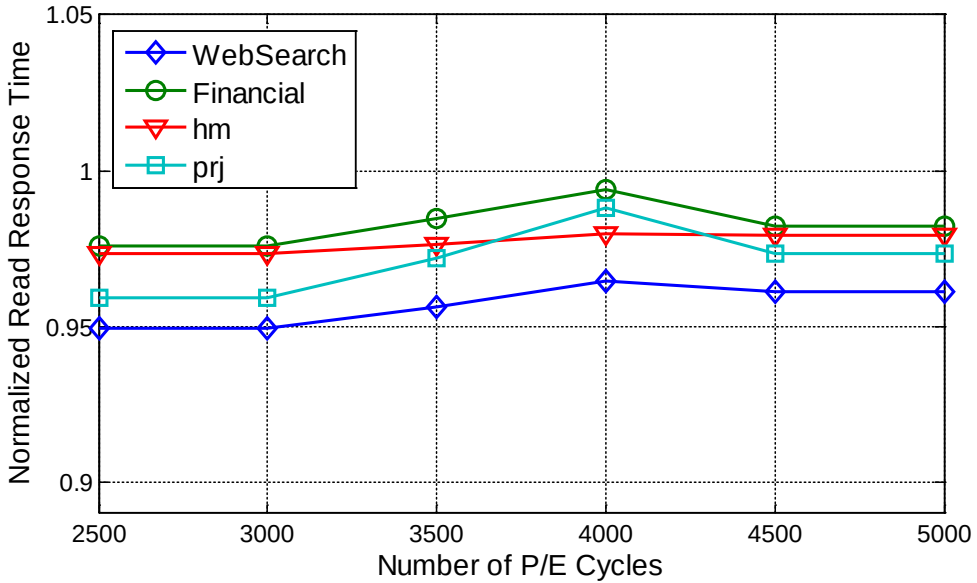
- Average read response time under different adaption strategies
- Read retry vs. multi-transfer-rate for different planes
  - Read retry doesn't sense the flash array, just re-transfer data
  - Direct lower transfer speed for the worse plane





Late Lifetime Performance

- Average read response time when combining read retry and multi-transfer-rate
- After 4000 P/E cycles, only lower the transfer rate for Plane 0

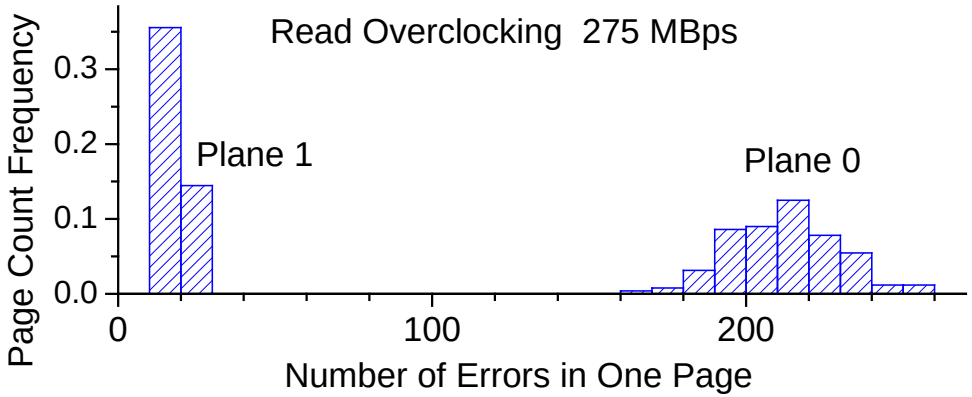


Link Error Mitigation using LDPC

- Effectiveness of the proposed overclocking design strategy depends on how well the ECC can handle the extra errors
- Leveraging LDPC codes and data link error characteristics to
  - Improving the tolerance to overclocking induced errors, which can lead to lower data re-transfer rate
  - Reducing the average number of LDPC decoding iterations, which can lead to lower power consumption and higher decoding throughput
- Two design techniques
  - Overclocking-aware LLR Calculation
  - Inter-plane interleaving

Inter-plane Interleaving

- Each plane has its own page register and I/O circuitry
- Performance after overclocking may be varying
- In our measurement, Plane 0 is much more vulnerable to overclocking-induced I/O errors

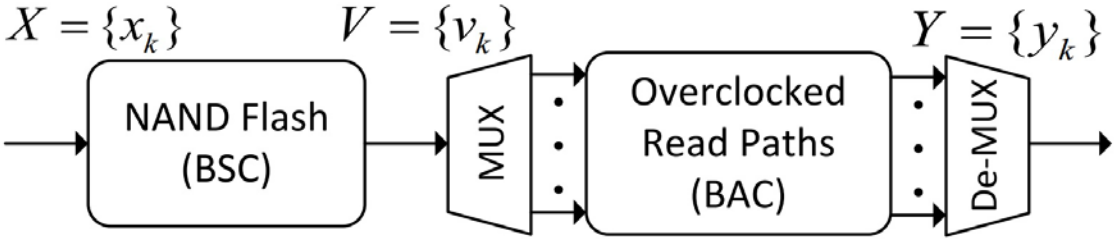


Overclocking-aware LLR Calculation

- The input of LDPC code decoding is the estimated log-likelihood ratio (LLR)

$$\mathrm{LLR}(x_k) = \frac{p(x_k = 0|y_k)}{p(x_k = 1|y_k)}$$

- Binary Asymmetric Channel (BAC). Different error characteristics for each wire and plane

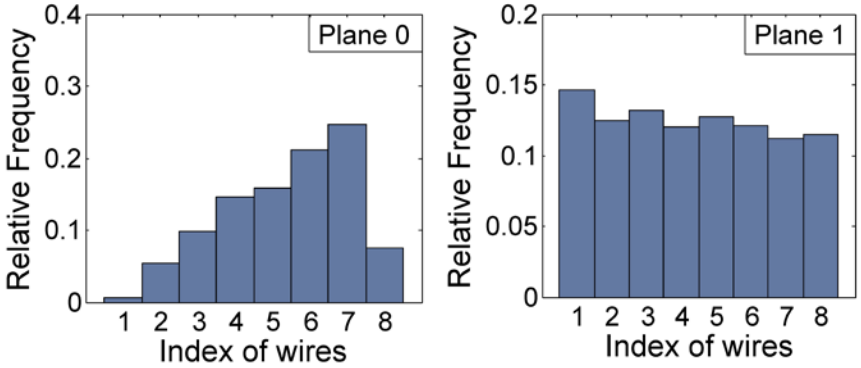


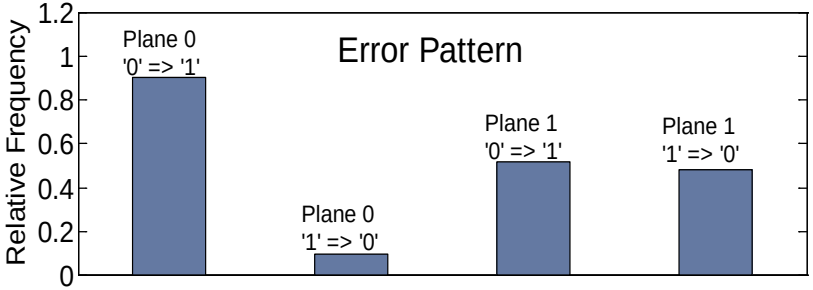
- In the overclocked SSD read channel, LLR can be expressed as

$$\mathrm{LLR}(x_k) = \frac{\Sigma_{i=0}^{l} p_m^{(l)}(y_k|v_k = i)p(v_k = i|x_k = 0,\, c)}{\Sigma_{i=0}^{l} p_m^{(l)}(y_k|v_k = i)p(v_k = i|x_k = 1,\, c)}$$

Error Pattern Characterization

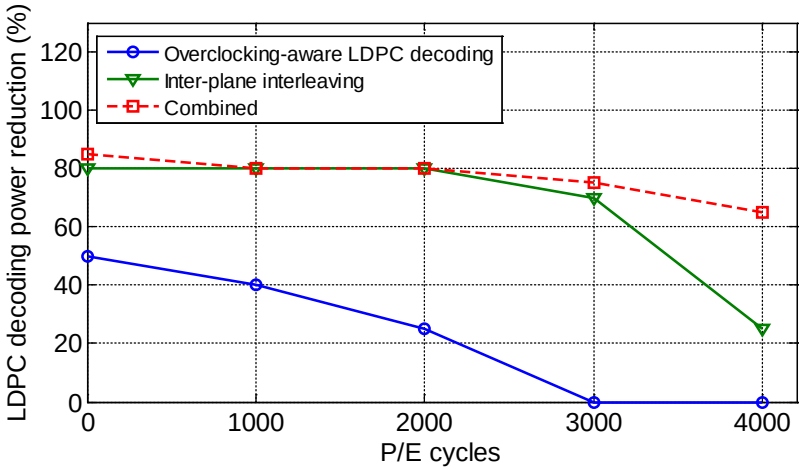
- Plane 0 suffers from more overclocking induced errors, moreover, these errors non-uniformly distributed across the 8-bit bus.
- The pattern ‘0’=>’1’ dominates the errors from Plane 0, while on Plane 1, patterns ‘0’=>’1’ and ‘1’=>’0’ evenly distributed.

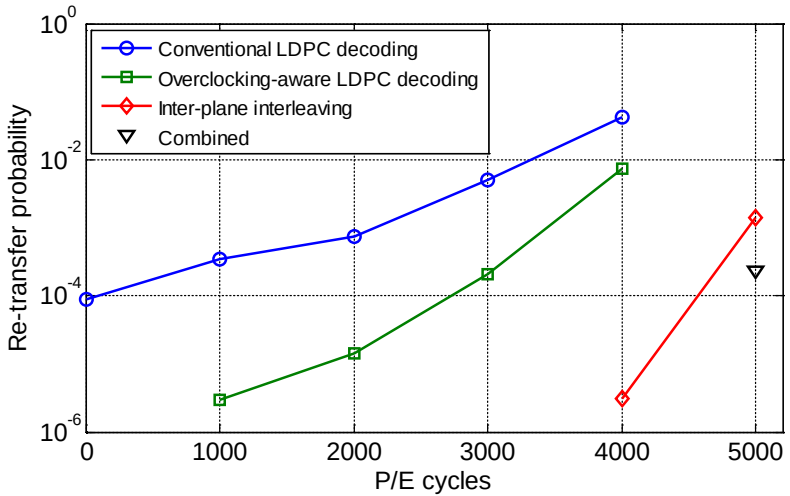




Simulation Results

- LDPC decoding power reduction
- The decoding power is determined by the number of iterations
- Data re-transfer rate reduction
- Data link overclocked to 275 MBPS





Conclusion

- Overclock data link to fully utilize the ECC capability
- Read performance can be significantly improved
- Mitigate the impact of overclocking by leveraging LDPC codes and the error characterization.

Thank you