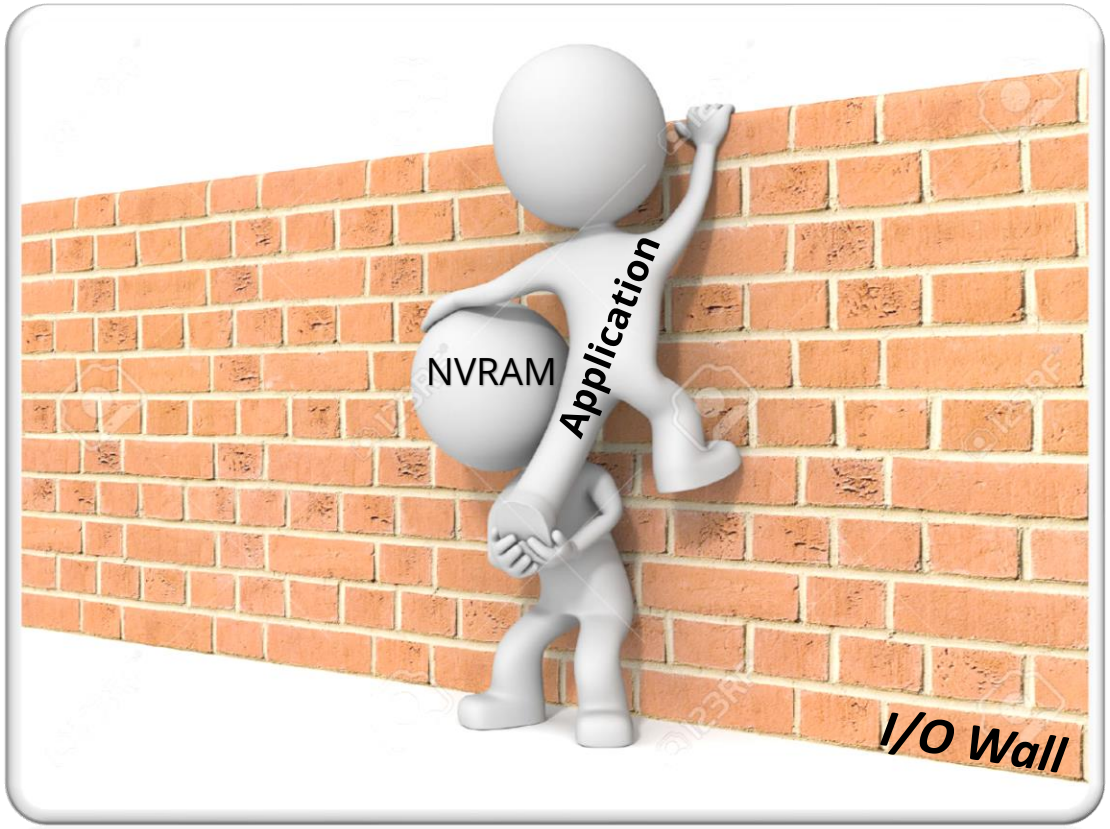
A Study of NVRAM Performance Variability under Concurrent I/O Accesses

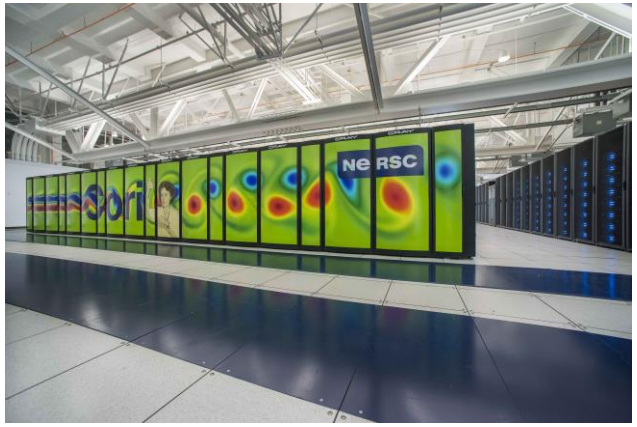
Anthony Kougkas - akougkas@hawk.iit.edu, Hariharan Devarajan, Xian-He Sun

Introduction

- Improvement in the speed of disk-based storage systems has been much slower than that of memory, creating a significant performance gap, the so called **I/O wall.**
- Non-Volatile RAM (NVRAM) offers lower latency and higher bandwidth and can help alleviate the issue.



NVRAM in the wild!



	Capacity	Bandwidth
RAM	expensive	cheap
NVRAM	**OK**	**OK**
DISK	cheap	expensive

In this study…

- We aim to explore NVRAM's performance characteristics
  - Bandwidth and latency
  - Concurrent access
  - Sequential vs Random
- ***Medium Sensitivity:***

$$Medium - Sensitivity = \frac{Application\ Concurrency}{Medium's\ Access\ Lanes} \times \frac{Reduction\ in\ BW\ (observed)}{Maximum\ BW\ (theoretical)}$$

- Is defined as the rate at which each medium experiences a bandwidth reduction due to concurrent accesses competing for the resource.

Testbed

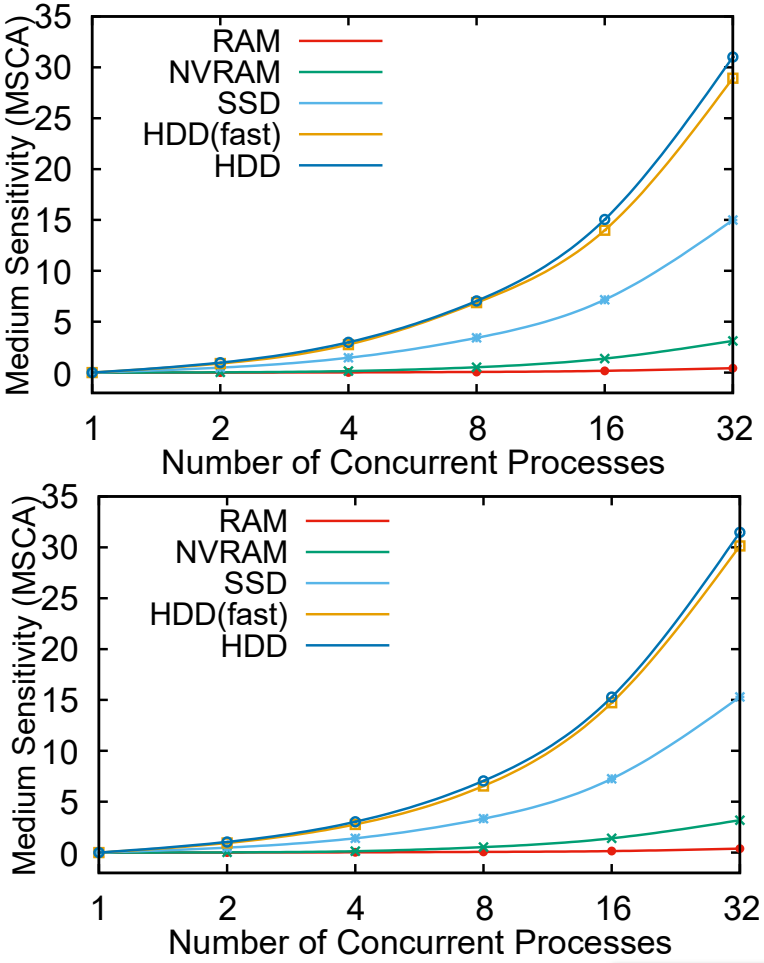
- Chameleon System
  - Storage Hierarchy Appliance
  - Centos 7.1, Mpich 3.2, OrangeFS 2.9.6
- Synthetic benchmark
  - 2 GB total I/O
  - 64 MB requests
  - File-per-process
  - Sequential

Device	RAM	NVRAM	SSD	HDD fast	HDD
Model	M386A4G40DM0	Intel DC P3700	Intel DC S3610	Seagate ST600MP0005	Seagate ST9250610NS
Connection	DDR4 2133Mhz	PCIe Gen3 x8	SATA 6Gb/s	12Gb/s SAS	SATA 6Gb/s
Capacity	512 GB(32GBx16)	1 TB	1.6 TB	600 GB	2.4 TB
Latency	13.5 ns	20 µs	55-66 µs	2 ms	4.16 ms
RPM	-	-	-	15000	7200
Buffer	-	-	-	128 MB	64 MB
Max Read BW	65000 MB/s	2800 MB/s	550 MB/s	215 MB/s	115 MB/s
Max Write BW	59000 MB/s	1900 MB/s	500 MB/s	185 MB/s	95 MB/s

Results

Observations:

- RAM is the least sensitive to concurrent access with only a small bandwidth reduction.
- Spinning drives and SSD are bound from the SATA controller
- However, SSD has some degree of internal parallelism (e.g., 2 channels) and thus is somewhere in the middle.
- **NVRAM** demonstrated behavior similar with RAM mostly because the NVMe controller uses the PCIe gen3 bus with x8 lanes and internal parallelism.



Q&A


рахмат
danke
謝謝
ngiyabonga
teşekkür ederim
Баярлалаа
spas
mersi
kia ora
barka
welalin
tack
спасибо
faafetai lava
vinaka
спасибі
blagodaram
dank je
misaotra
matondo
paldies
grazzi
mahalo
tapadh leat
thank you
gracias
kiitos
dankie
nanni
nandri
dhanyavad
хвала
asante
manana
obrigada
tenki
murakoze
maururu
köszönöm
hvala
gràcie
bayarlalaa
bedankt
enkosi
akun dankon
aciū
djiere dieuf
tau
mochchakkeram
дякую
mamnun
chokrane
dziękuje
chnorakaloutioun
gratias ago
gràcies
sulpáy
taiku
go raibh maith agat
sobodi
děkuji
obrigado
sagolun
sukriya
kop khun krap
ありがとう
arigatô
takk
dakujem
trugarez
mèsi
didi madloba
najis tuke
tanemirt
grazie
kam sah hamnida
terima kasih
rahmet
diolch
dhanyavadagalu
shukriya
mercé
мерси
rahmat
তোমাকে ধন্যবাদ
감사합니다
xièxie
ευχαριστώ
merci