



Integrating NoSQL & Hadoop on AWS

CPSC 8620: Database Management Systems Design Research Project

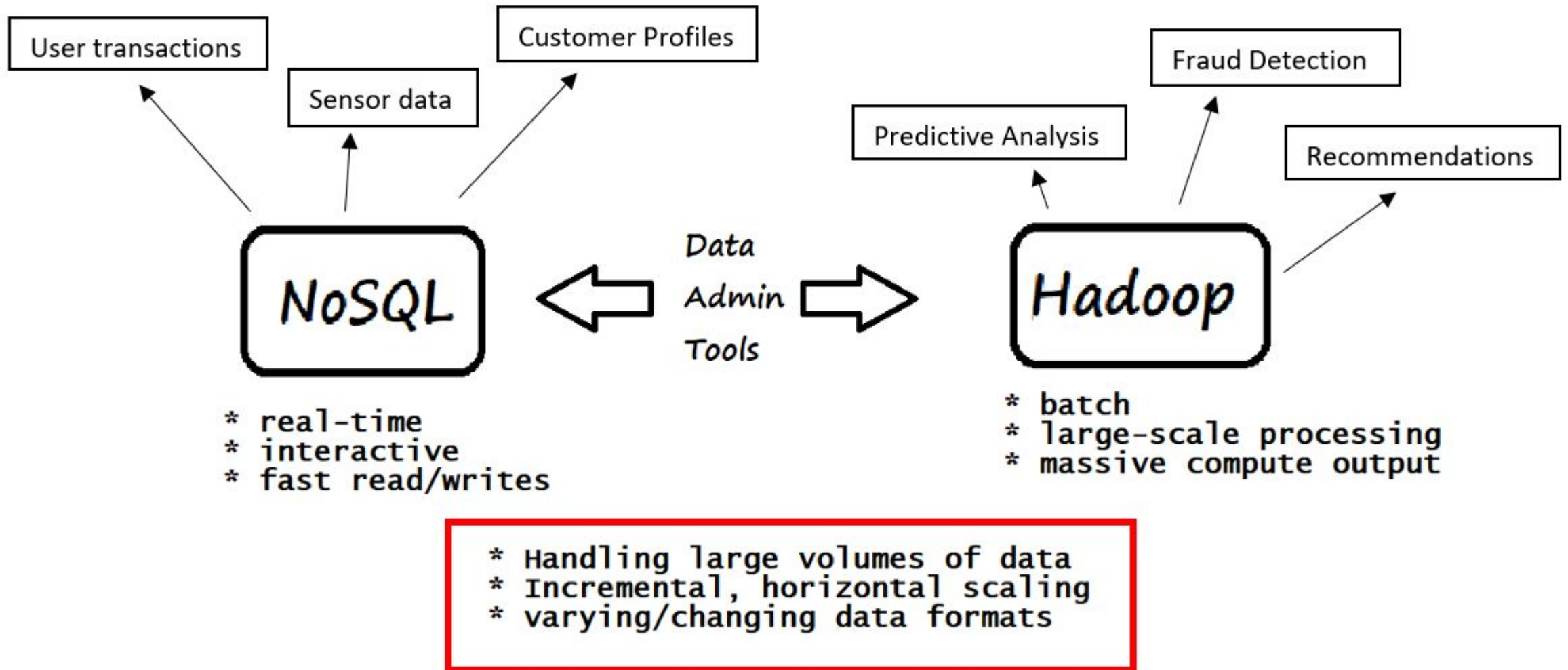
December 3, 2019

Clemson University

Team Name: Team TV

Submitted by: Tamanna Baig | Vimal Vakeel

Proposal



Initial Approach

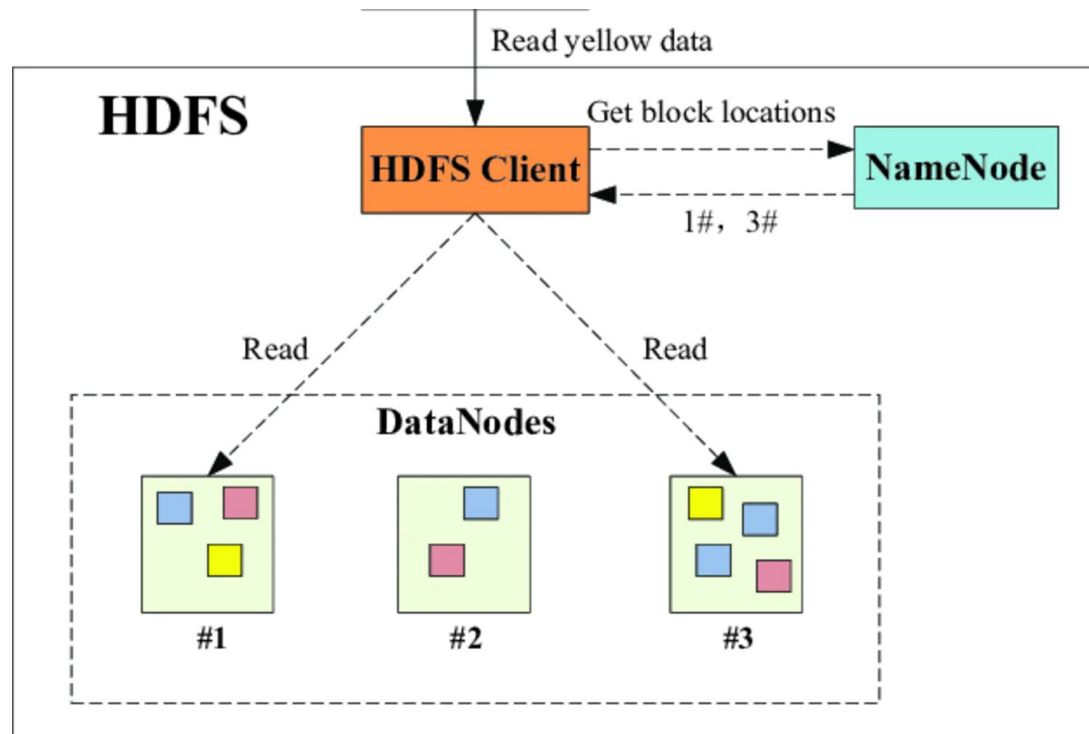
- Deploying myHadoop test environment on Clemson's Cypress cluster
- Designing Docker instances and the deploy on Clemson's Palmetto



AWS Educate to setup EC2 instances for Hadoop, MongoDB, Cassandra

Complete Setup - Hadoop

<input type="checkbox"/>	Name	Instance ID	Instance Type	Availability Zone	Instance State	Status Checks	Alarm	Public DNS (IPv4)
<input checked="" type="checkbox"/>	namenode	i-025921e68508059...	t2.micro	us-east-1b	running	2/2 checks ...	No...	ec2-3-95-165-148.compu
<input type="checkbox"/>	datanode1	i-07a0c89fc670847e8	t2.micro	us-east-1b	running	2/2 checks ...	No...	ec2-18-234-241-76.comp
<input type="checkbox"/>	datanode2	i-092f0854cefb0a1b0	t2.micro	us-east-1b	running	2/2 checks ...	No...	ec2-54-89-241-181.comp
<input type="checkbox"/>	datanode3	i-0dd4037fefbf56310	t2.micro	us-east-1b	running	2/2 checks ...	No...	ec2-3-94-55-212.computr



Setup- MongoDB

<input type="checkbox"/>	Name ▾	Instance ID ▾	Instance Type ▾	Availability Zone ▾	Instance State ▴	Status Checks ▾	Alarm \$	Public DNS (IPv4)
<input checked="" type="checkbox"/>	MongoDB	i-0515d973aa78cf561	t2.micro	us-east-1d	 running	 2/2 checks ...	No... 	ec2-35-173-180-190.com

- MongoDB-Hadoop was connected using “Mongo-Hadoop” driver provided by MongoDB Inc.
- Benchmark - YCSB client

Setup- Apache Cassandra

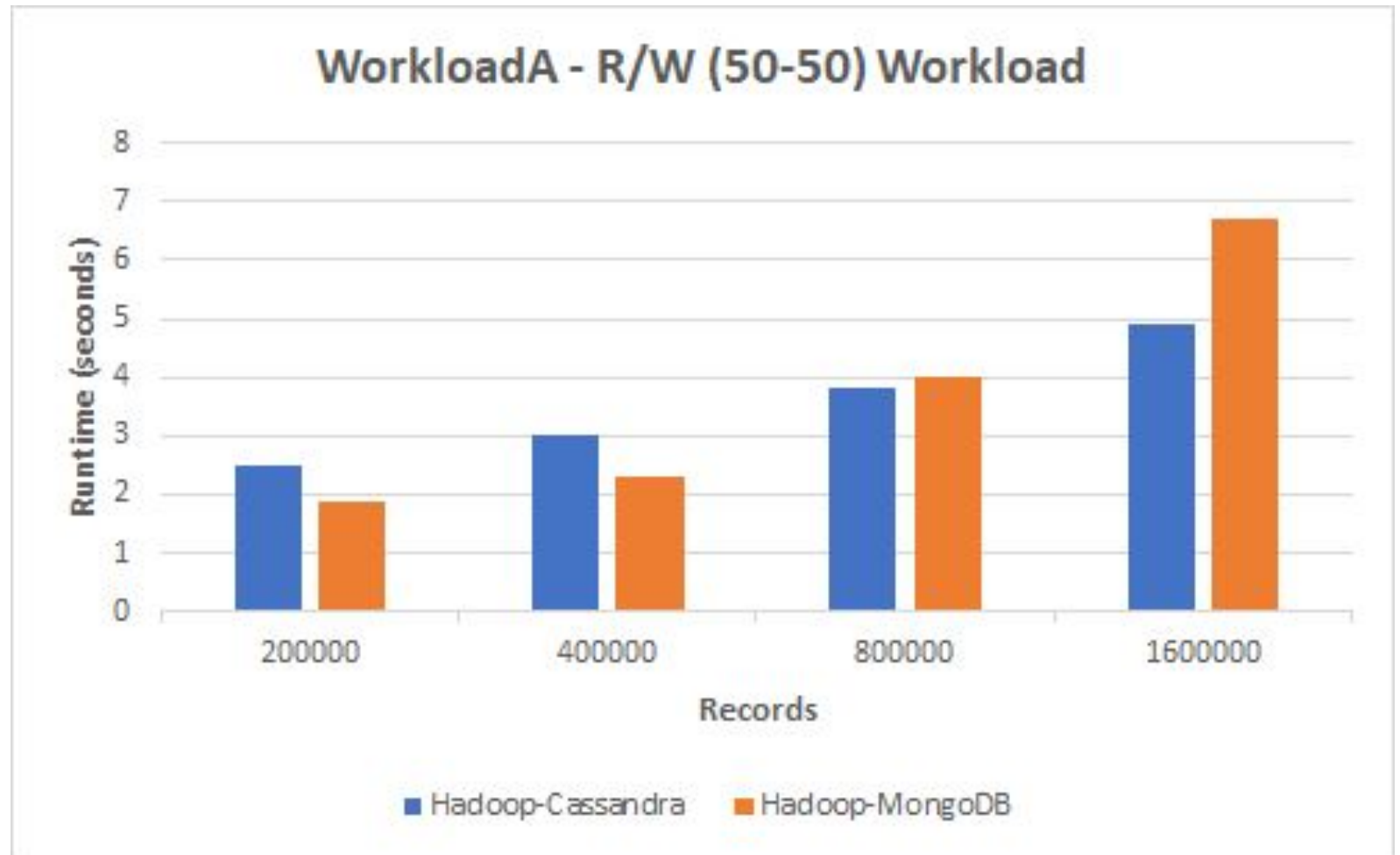
<input type="checkbox"/>	Name	Instance ID	Instance Type	Availability Zone	Instance State	Status Checks	Alarm Status	Public DNS (IPv4)
<input checked="" type="checkbox"/>	Cassandra-2	i-0a3d0684ffa8e0c24	t2.micro	us-east-1d	 running	 2/2 checks ...	None	 ec2-18-208-216-206.cc
<input type="checkbox"/>	Namenode	i-000aeb9a2c885c971	t2.micro	us-east-1c	 stopped		None	
<input type="checkbox"/>	Datanode1	i-083f737aa8c716f4a	t2.micro	us-east-1c	 stopped		None	

- Cassandra-Hadoop Datastax ODBC driver
- Benchmark - YCSB Client

Results

Workload A - This is a balanced workload. With 50-50 read-write transactions.

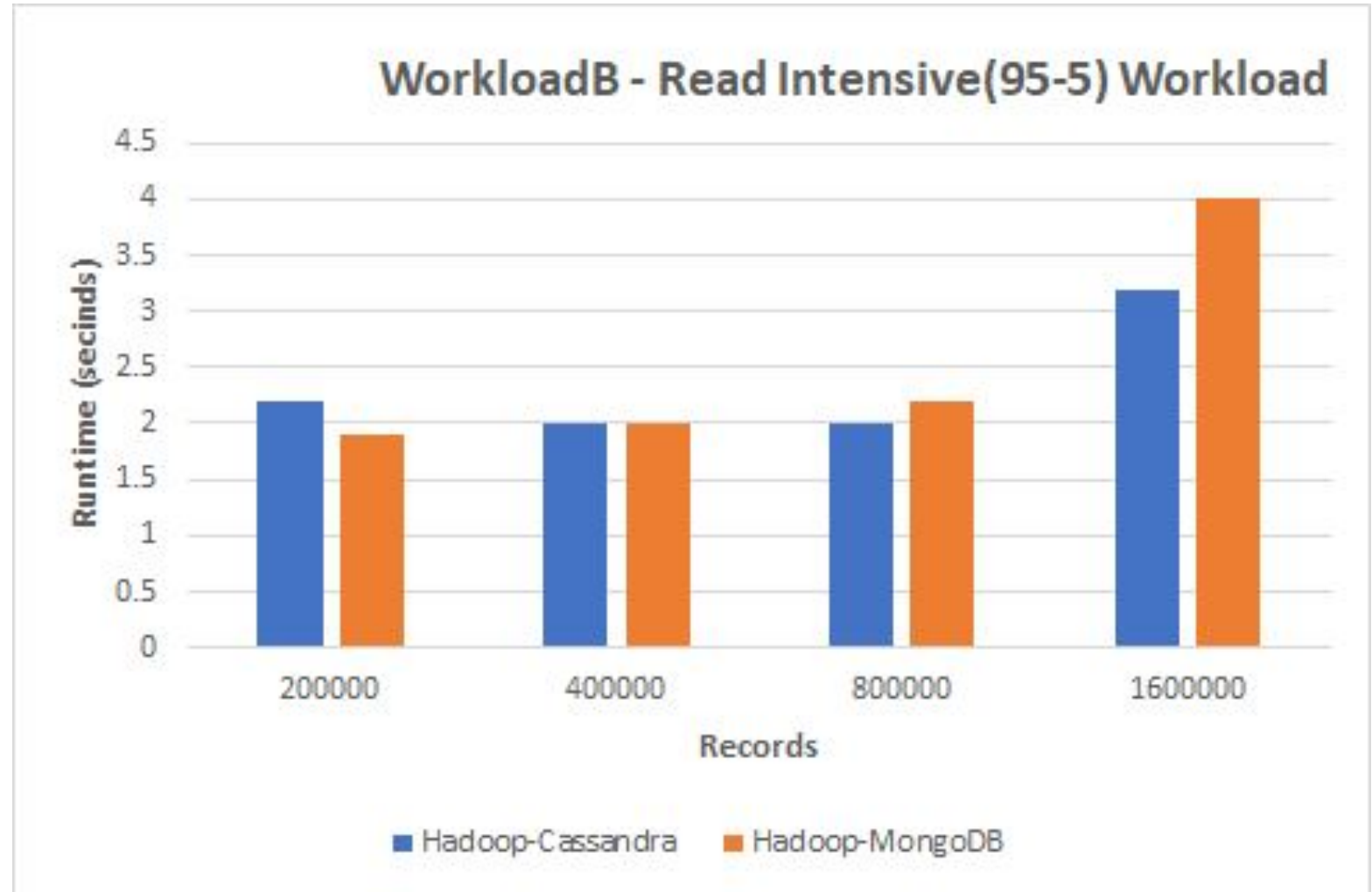
- H-M performs better than H-C with smaller records.
- As the number of records gradually increases the performance of H-M starts to drop, and H-C starts to process faster.
- Integrating with MongoDB when the dataset is of a relatively smaller size would be ideal



Results

Workload B - This is a Read-Intense workload. With 95-5 read-write transactions.

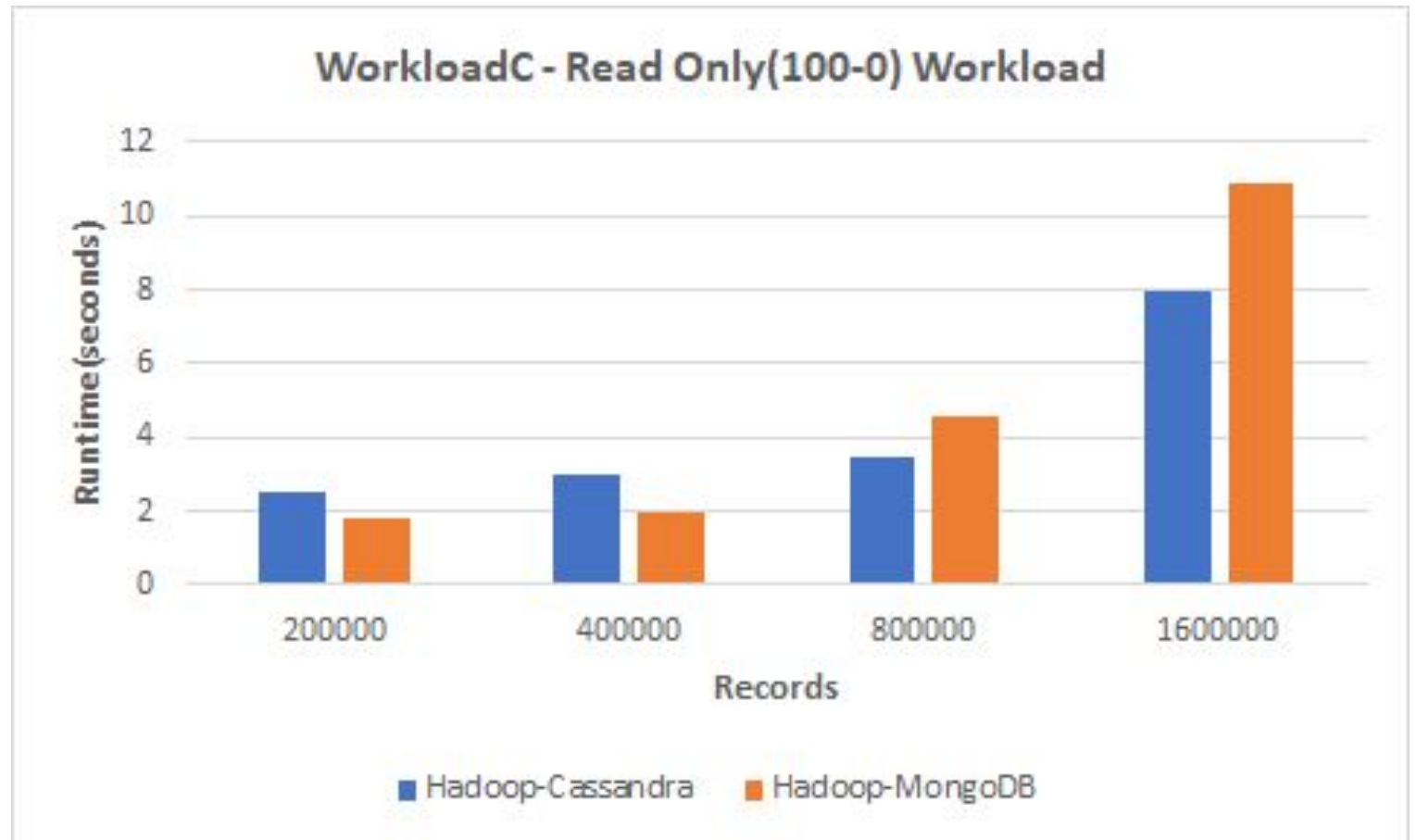
- The performance of both the NoSQL databases looks to be almost the same.
- MongoDB performs badly as the number of records are increasing.
- Integrating with MongoDB when the dataset is of a relatively smaller size would be ideal



Results

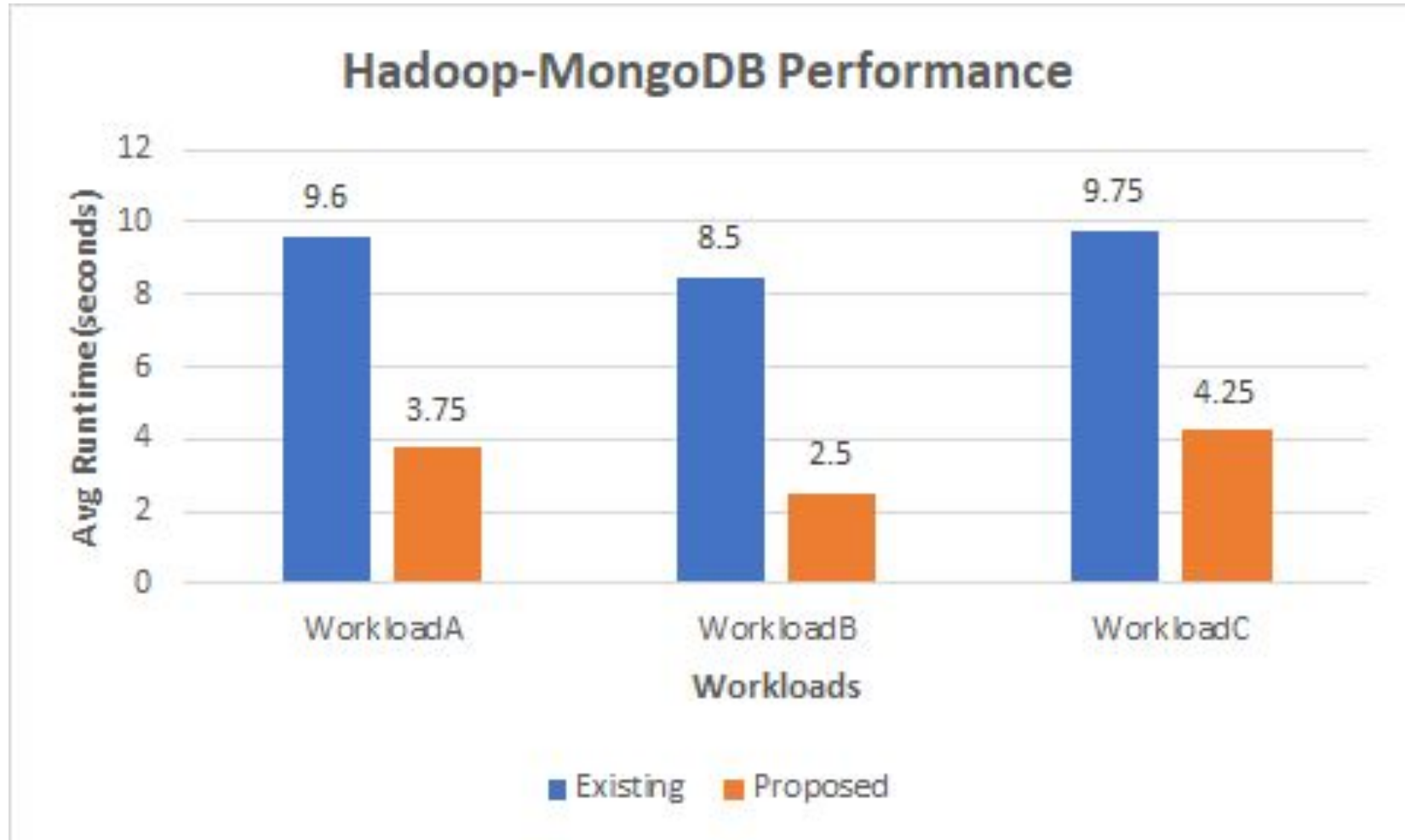
Workload C - This is a Read-only workload. With 100% read transactions.

- H-M performs better than H-C with smaller records.
- As the number of records gradually increases the performance of H-M starts to drop, and H-C starts to process faster.
- Integrating with MongoDB when the dataset is of a relatively smaller size would be ideal



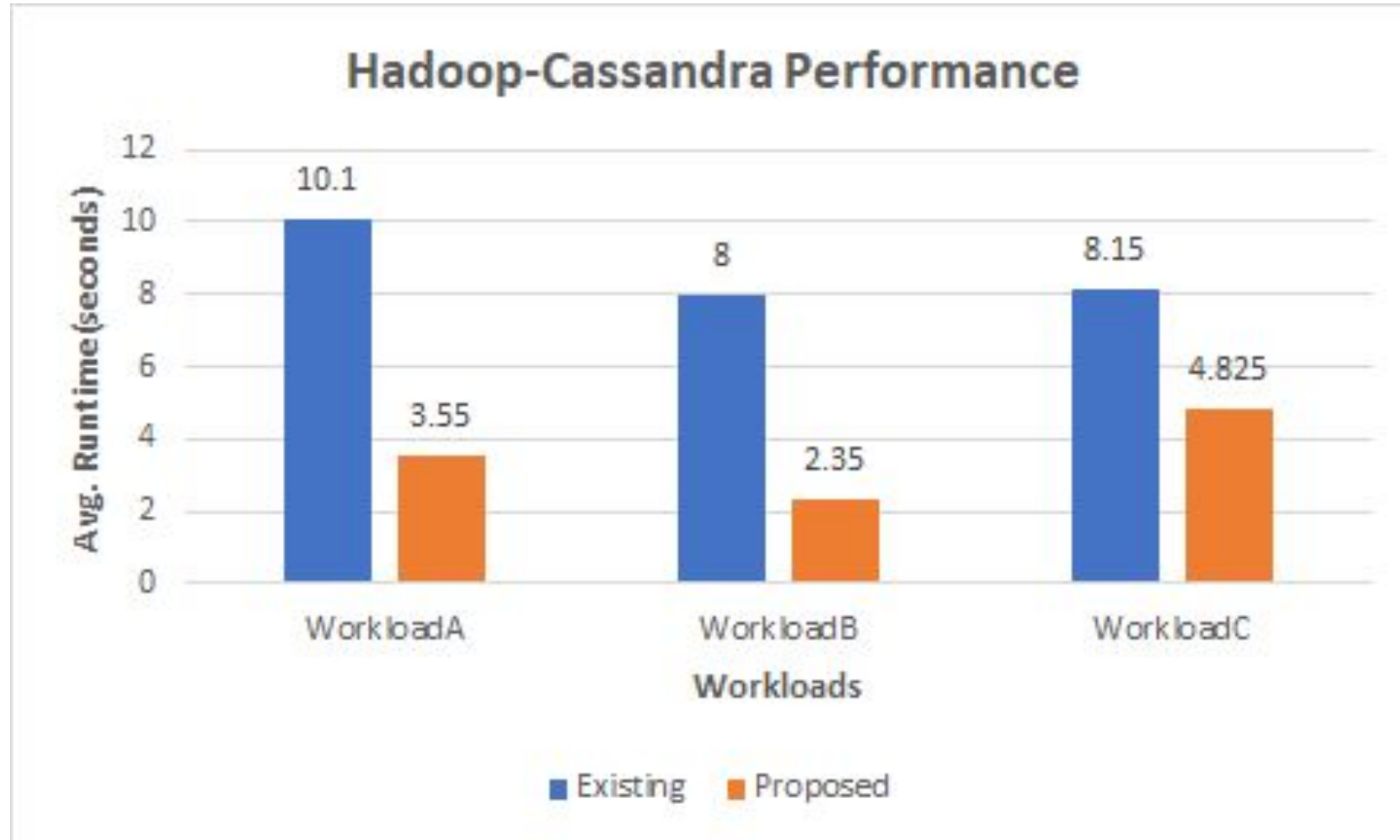
Comparison with Existing System

Existing System was installed using VMs



Comparison with Existing System

Existing System was installed using VMs



WHY DO WHALES JUMP
WHY ARE WITCHES GREEN
WHY ARE THERE MIRRORS ABOVE BEDS
WHY DO I SAY UH

WHY IS SEA SALT BETTER
WHY ARE THERE TREES IN THE MIDDLE OF FIELDS
WHY IS THERE NOT A POKEMON MMO
WHY IS THERE LAUGHING IN TV SHOWS
WHY ARE THERE DOORS ON THE FREEWAY
WHY ARE THERE SO MANY SNOCOSTEXE RUNNING
WHY AREN'T THERE ANY COUNTRIES IN ANTARCTICA
WHY ARE THERE SCARY SOUNDS IN MINECRAFT
WHY IS THERE KICKING IN MY STOMACH
WHY ARE THERE TWO SLASHES AFTER HTTP

WHY ARE THERE CELEBRITIES
WHY DO SNAKES EXIST
WHY DO OYSTERS HAVE PEARLS
WHY ARE DUCKS CALLED DUCKS
WHY DO THEY CALL IT THE CLAP
WHY ARE KYLE AND CARTMAN FRIENDS
WHY IS THERE AN ARROW ON PANG'S HEAD
WHY ARE TEXT MESSAGES BLUE
WHY ARE THERE MUSTACHES ON CLOTHES
WHY ARE THERE MUSTACHES ON CARS
WHY ARE THERE MUSTACHES EVERYWHERE
WHY ARE THERE SO MANY BIRDS IN OHIO
WHY IS THERE SO MUCH RAIN IN OHIO
WHY IS OHIO WEATHER SO WEIRD

WHY ARE THERE MALE AND FEMALE BIKES
WHY ARE THERE BRIDESMAIDS
WHY DO DYING PEOPLE REACH UP
WHY AREN'T THERE WARGOISE, ARTERIES
WHY ARE OLD KLINGONS DIFFERENT

WHY ARE THERE TINY SPIDERS IN MY HOUSE
WHY DO SPIDERS COME INSIDE
WHY ARE THERE HUGE SPIDERS IN MY HOUSE
WHY ARE THERE LOTS OF SPIDERS IN MY HOUSE
WHY ARE THERE SPIDERS IN MY ROOM
WHY ARE THERE SO MANY SPIDERS IN MY ROOM
WHY DO SPIDER BITES ITCH
WHY IS DYING SO SCARY

WHY ARE THERE SQUIRRELS
WHY IS THERE NO GPS IN LAPTOPS
WHY DO KNEES CLICK
WHY AREN'T THERE E. GRADES
WHY IS ISOLATION BAD
WHY DO BOYS LIKE ME
WHY DON'T BOYS LIKE ME
WHY IS THERE ALWAYS A JAWA UPDATE
WHY ARE THERE RED DOTS ON MY THIGHS
WHY IS LYING GOOD

WHY IS PROGRAMMING SO HARD
WHY IS THERE A 0 OHM RESISTOR
WHY DO AMERICANS HATE SOCCER
WHY DO RHYMES SOUND GOOD
WHY DO TREES DIE
WHY IS THERE NO SOUND ON CNN
WHY AREN'T POKEMON REAL
WHY AREN'T BULLETS SHARP
WHY DO DREAMS SEEM SO REAL

WHY DO TESTICLES MOVE
WHY ARE THERE PSYCHICS
WHY ARE HATS SO EXPENSIVE
WHY IS THERE COFFINE IN MY SHAMPOO
WHY DO YOUR BOOBS HURT

WHY AREN'T ECONOMISTS RICH
WHY DO AMERICANS CALL IT SOCCER
WHY ARE MY EARS RINGING
WHY ARE THERE SO MANY AVENGERS
WHY ARE THE AVENGERS FIGHTING THE X MEN
WHY IS WOLVERINE NOT IN THE AVENGERS

WHY ARE THERE ANTS IN MY LAPTOP
WHY IS EARTH TILTED
WHY IS SPACE BLACK
WHY IS OUTER SPACE SO COLD
WHY ARE THERE PYRAMIDS ON THE MOON
WHY IS NASA SHUTTING DOWN

WHY ARE THERE GHOSTS
WHY IS THERE AN OWL IN MY BACKYARD
WHY IS THERE AN OWL OUTSIDE MY WINDOW
WHY IS THERE AN OWL ON THE DOLLAR BILL
WHY DO OWLS ATTACK PEOPLE
WHY ARE AK 47s SO EXPENSIVE
WHY ARE THERE HELICOPTERS CIRCLING MY HOUSE
WHY ARE THERE GODS
WHY ARE THERE TWO SPOOKS

WHY IS MT VESUVIUS THERE
WHY DO THEY SAY T MINUS
WHY ARE THERE OBELISKS
WHY ARE WRESTLERS ALWAYS WET
WHY ARE OCEANS BECOMING MORE ACIDIC
WHY IS ARWEN DYING
WHY AREN'T MY QUAIL LAYING EGGS
WHY AREN'T MY QUAIL EGGS HATCHING
WHY AREN'T THERE ANY FOREIGN MILITARY BASES IN AMERICA

WHY IS SEX SO IMPORTANT
WHY ARE THERE DOGS AFRAID OF FIREWORKS
WHY IS THERE NO KING IN ENGLAND

WHY AREN'T MY ARMS GROWING
WHY AREN'T THERE GUNS IN HARRY POTTER
WHY ARE ULTRASOUNDS IMPORTANT
WHY ARE ULTRASOUND MACHINES EXPENSIVE
WHY IS STEALING WRONG

WHY ARE THERE SLAVES IN THE BIBLE
WHY IS HTTPS CROSSED OUT IN RED
WHY IS THERE A LINE THROUGH HTTPS
WHY IS THERE A RED LINE THROUGH HTTPS ON FACEBOOK
WHY IS HTTPS IMPORTANT

WHY ARE THERE SQUIDS OF CARS
WHY IS THERE PALEGH
WHY ARE THERE SO MANY CROWS IN ROCHESTER
WHY IS PSYCHIC WEAK TO BUG
WHY DO CHILDREN GET CANCER
WHY IS POSEIDON ANGRY WITH ODYSSEUS
WHY IS THERE ICE IN SPACE

WHY ARE THERE GHOSTS
WHY IS THERE AN OWL IN MY BACKYARD
WHY IS THERE AN OWL OUTSIDE MY WINDOW
WHY IS THERE AN OWL ON THE DOLLAR BILL
WHY DO OWLS ATTACK PEOPLE
WHY ARE AK 47s SO EXPENSIVE
WHY ARE THERE HELICOPTERS CIRCLING MY HOUSE
WHY ARE THERE GODS
WHY ARE THERE TWO SPOOKS

WHY IS MT VESUVIUS THERE
WHY DO THEY SAY T MINUS
WHY ARE THERE OBELISKS
WHY ARE WRESTLERS ALWAYS WET
WHY ARE OCEANS BECOMING MORE ACIDIC
WHY IS ARWEN DYING
WHY AREN'T MY QUAIL LAYING EGGS
WHY AREN'T MY QUAIL EGGS HATCHING
WHY AREN'T THERE ANY FOREIGN MILITARY BASES IN AMERICA

WHY IS SEX SO IMPORTANT
WHY ARE THERE DOGS AFRAID OF FIREWORKS
WHY IS THERE NO KING IN ENGLAND

WHY AREN'T MY ARMS GROWING
WHY AREN'T THERE GUNS IN HARRY POTTER
WHY ARE ULTRASOUNDS IMPORTANT
WHY ARE ULTRASOUND MACHINES EXPENSIVE
WHY IS STEALING WRONG

WHY AREN'T MY ARMS GROWING
WHY AREN'T THERE GUNS IN HARRY POTTER
WHY ARE ULTRASOUNDS IMPORTANT
WHY ARE ULTRASOUND MACHINES EXPENSIVE
WHY IS STEALING WRONG

WHY ARE THERE SLAVES IN THE BIBLE
WHY IS HTTPS CROSSED OUT IN RED
WHY IS THERE A LINE THROUGH HTTPS
WHY IS THERE A RED LINE THROUGH HTTPS ON FACEBOOK
WHY IS HTTPS IMPORTANT

WHY AREN'T MY ARMS GROWING
WHY AREN'T THERE GUNS IN HARRY POTTER
WHY ARE ULTRASOUNDS IMPORTANT
WHY ARE ULTRASOUND MACHINES EXPENSIVE
WHY IS STEALING WRONG

WHY AREN'T MY ARMS GROWING
WHY AREN'T THERE GUNS IN HARRY POTTER
WHY ARE ULTRASOUNDS IMPORTANT
WHY ARE ULTRASOUND MACHINES EXPENSIVE
WHY IS STEALING WRONG

WHY AREN'T MY ARMS GROWING
WHY AREN'T THERE GUNS IN HARRY POTTER
WHY ARE ULTRASOUNDS IMPORTANT
WHY ARE ULTRASOUND MACHINES EXPENSIVE
WHY IS STEALING WRONG

WHY AREN'T MY ARMS GROWING
WHY AREN'T THERE GUNS IN HARRY POTTER
WHY ARE ULTRASOUNDS IMPORTANT
WHY ARE ULTRASOUND MACHINES EXPENSIVE
WHY IS STEALING WRONG

WHY AREN'T MY ARMS GROWING
WHY AREN'T THERE GUNS IN HARRY POTTER
WHY ARE ULTRASOUNDS IMPORTANT
WHY ARE ULTRASOUND MACHINES EXPENSIVE
WHY IS STEALING WRONG

WHY AREN'T MY ARMS GROWING
WHY AREN'T THERE GUNS IN HARRY POTTER
WHY ARE ULTRASOUNDS IMPORTANT
WHY ARE ULTRASOUND MACHINES EXPENSIVE
WHY IS STEALING WRONG

QUESTIONS

FOUND IN GOOGLE AUTOCOMPLETE

