

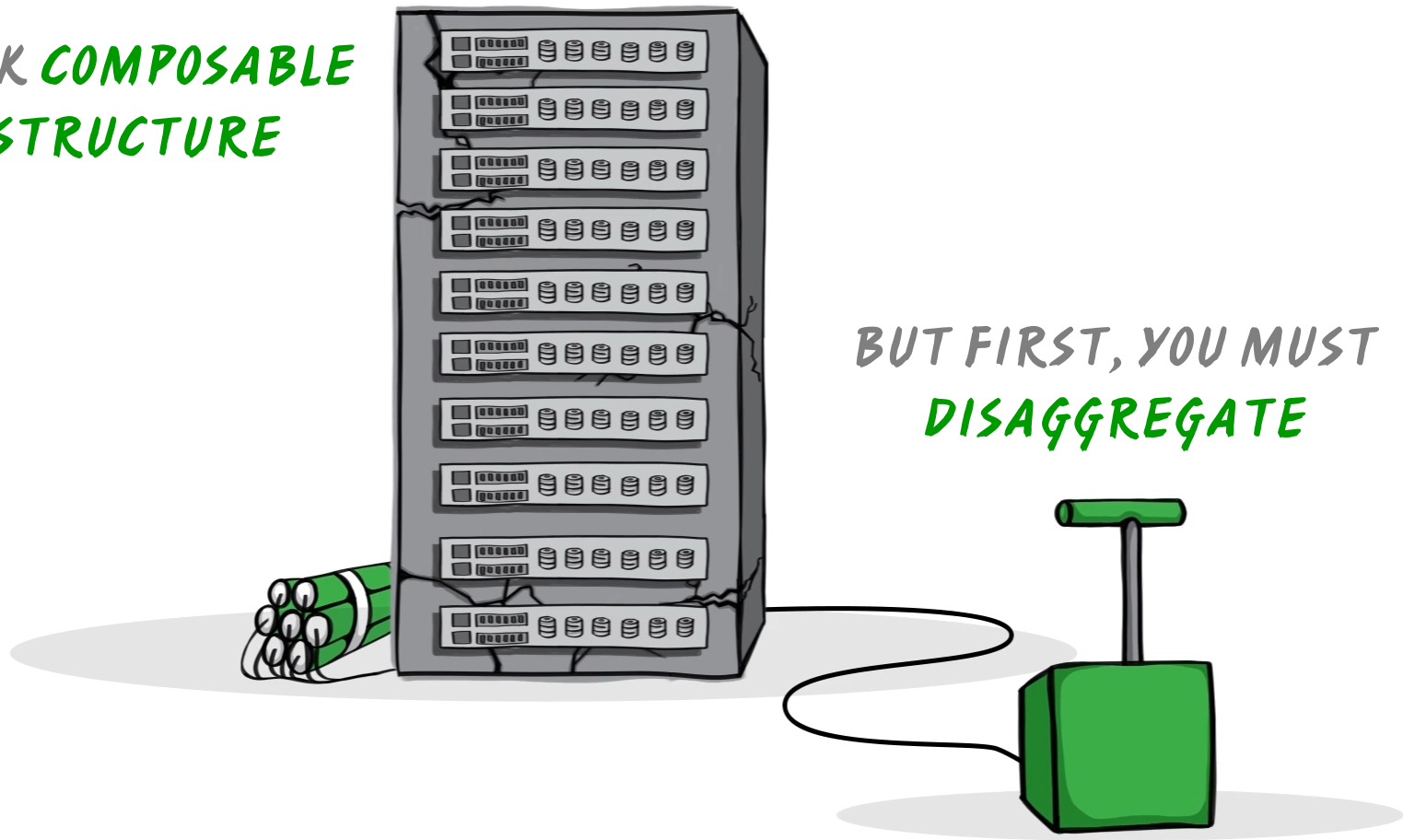


Flash Memory Summit

Flash Considerations for Software Composable Infrastructure

Brian Pawlowski
CTO, DriveScale Inc.

LET'S TALK **COMPOSABLE**
INFRASTRUCTURE



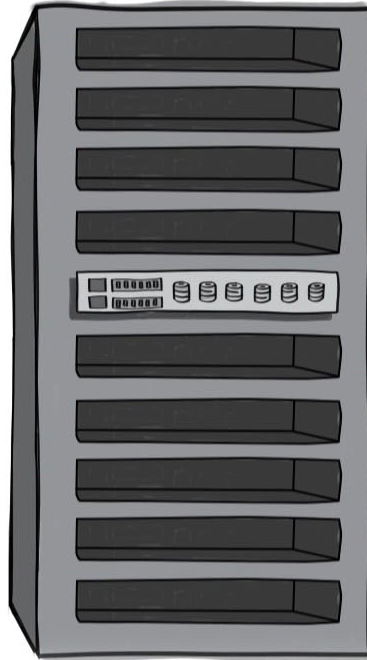
BUT FIRST, YOU MUST
DISAGGREGATE



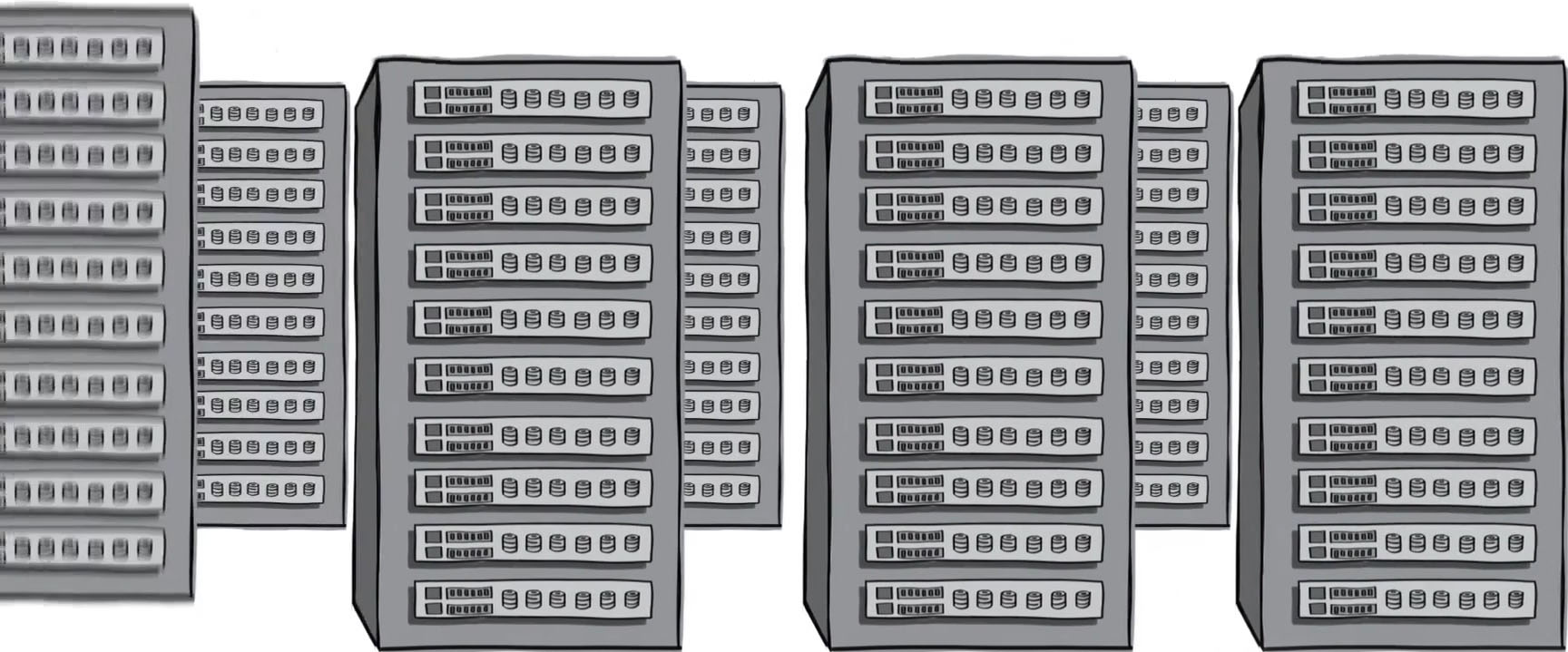
WHAT IS IT?

- **SOFTWARE COMPOSABLE INFRASTRUCTURE**
 - **DISAGGREGATE COMPUTE, NETWORKING, AND STORAGE INTO BASIC BUILDING BLOCKS**
 - **THEN COMPOSE OPTIMIZED “PHYSICAL” NODES INTO SECURE CLUSTERS FOR SCALE-OUT APPS UNDER SOFTWARE CONTROL**
- **WITH AN EMPHASIS ON BIG DATA AND EMERGING ML AND AI APPS**
- **NOT INTERESTED YET IN CPU/MEMORY/GPU DISAGGREGATION**

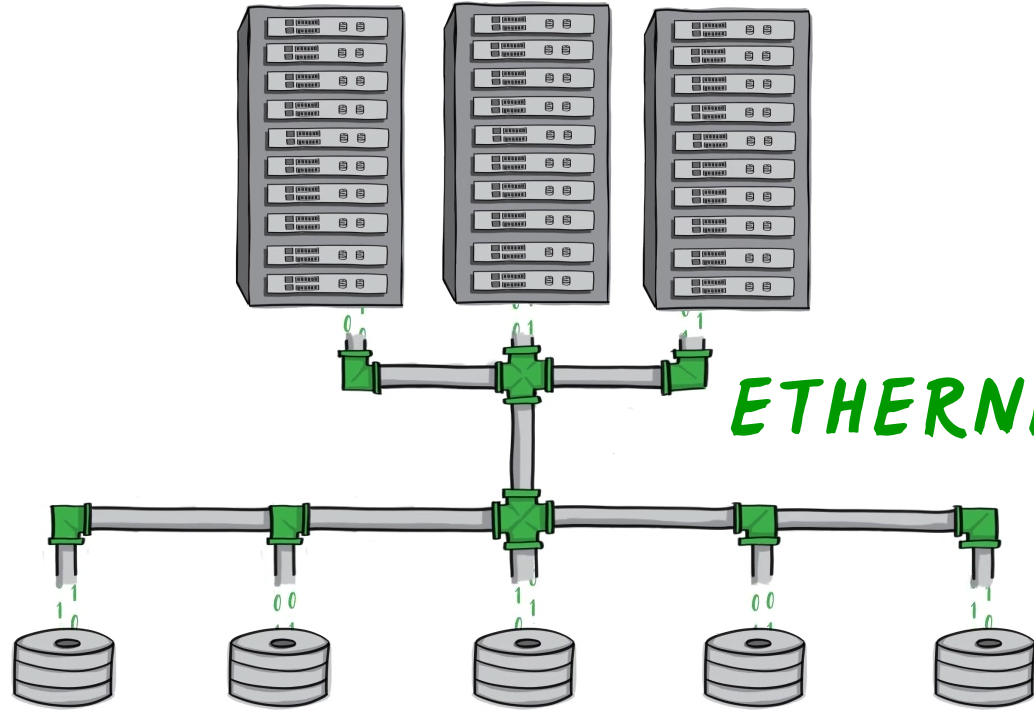
BUT LET'S GO THROUGH A STORY



A SIMPLE SERVER, TO START



A SIMPLE SERVER? PERHAPS NOT...



ETHERNET RULES

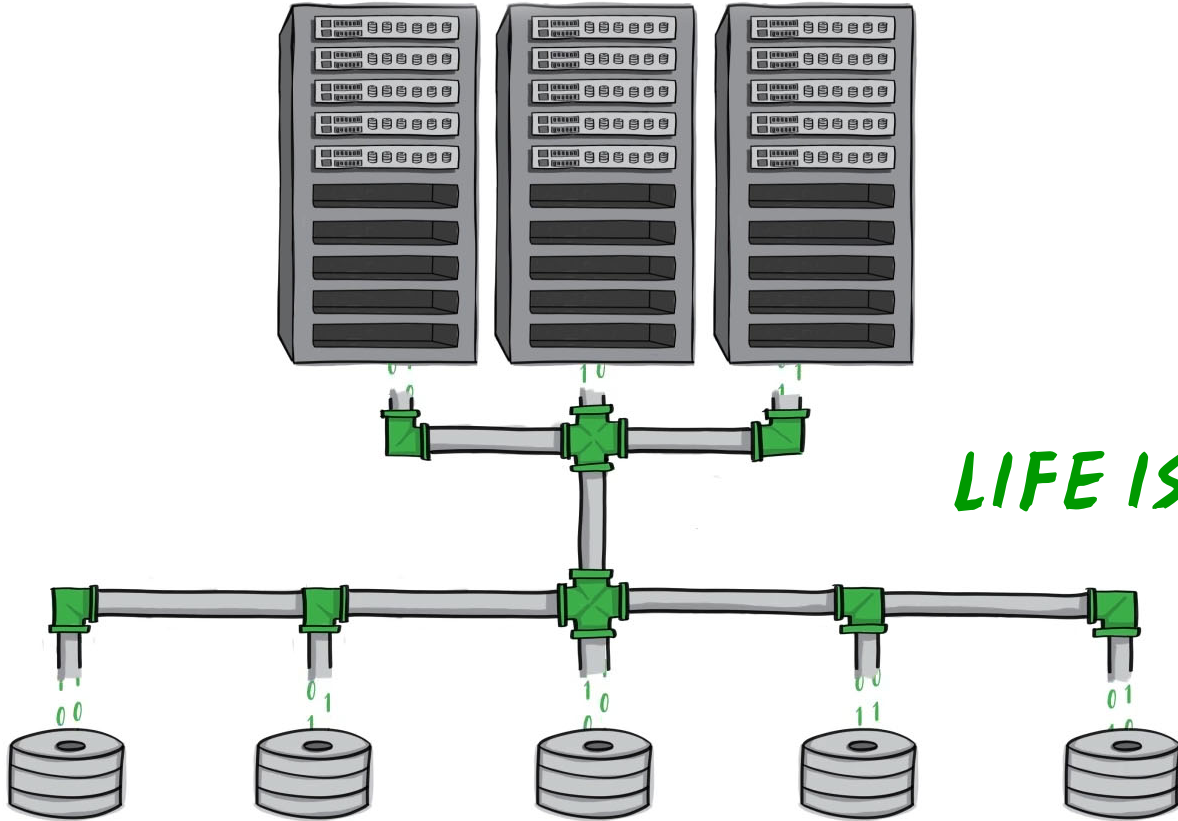
NETWORKED STORAGE



LESSONS LEARNED

- **SEPARATING COMPUTE AND STORAGE**
 - **INCREASES UTILIZATION**
 - **ENABLES SKU LEVEL LIFECYCLE MANAGEMENT**
- **SAN VS. NAS WARS LEFT ONE LASTING IMPRESSION – ETHERNET WINS**
 - **SCALES BETTER**
 - **COMMON COMMODITY NETWORKING REDUCES OVERALL COST AND OPERATIONAL COMPLEXITY**

VIRTUALIZATION

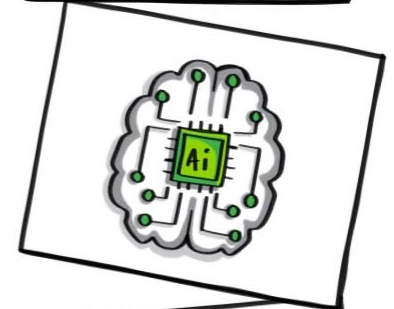
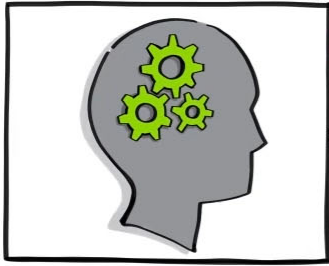
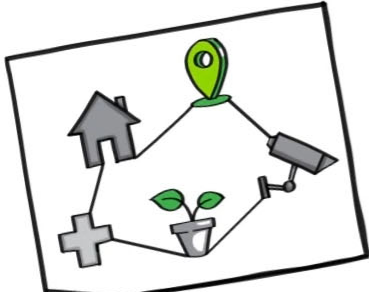


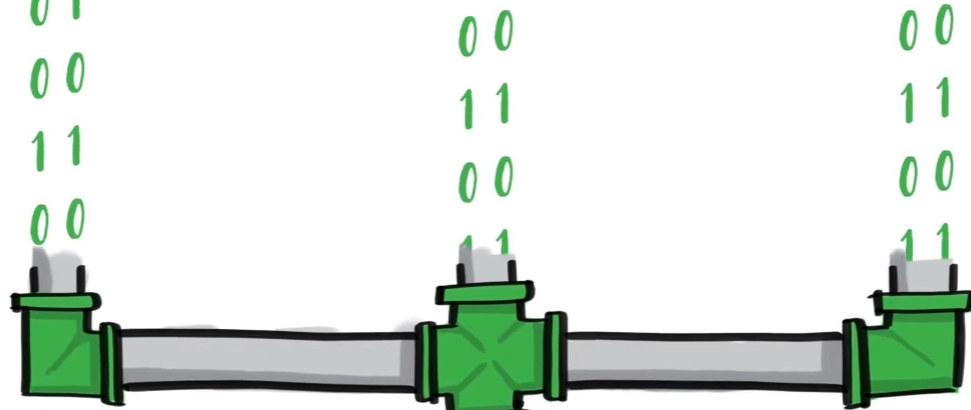


LESSONS LEARNED

- **EFFECTIVE RESOURCE UTILIZATION (HERE CPU) WAS KEY TO COST EFFICIENCY (REDUCE OVERPROVISIONING)**
 - **CORE COUNTS GREW WHEN CLOCK SPEED WALL WAS HIT – LEGACY APPS DID NOT RESPOND LEAVING CPU CYCLES ON THE TABLE**
 - **SAFE MULTITENANCY ALLOWED CO-EXISTING APPLICATIONS**
- **SECONDARY EFFECT WAS PROFOUND**
 - **EMERGENCE OF SOFTWARE-BASED APPLICATION PROVISIONING**
 - **COMMODITIZATION OF COMPUTE NODES**
- **RECLAIM INCREASINGLY SCARCE REAL ESTATE IN DATA CENTER**

BUSINESSES BEGAN COLLECTING
MORE DATA





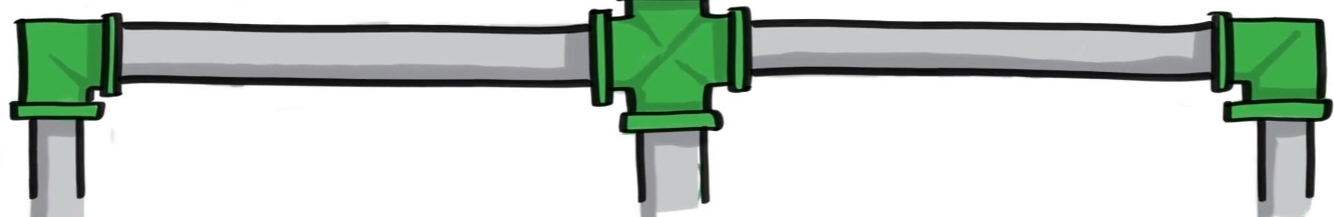
01
00
11
00

00
11
00
11

00
11
00
11

*CONVENTIONAL
APPROACHES*

*STOPPED
WORKING*

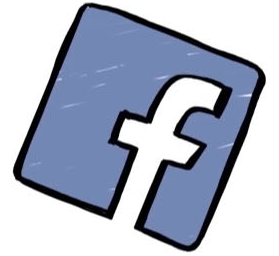


11
00
11

11
00
11

11
00
11

Google

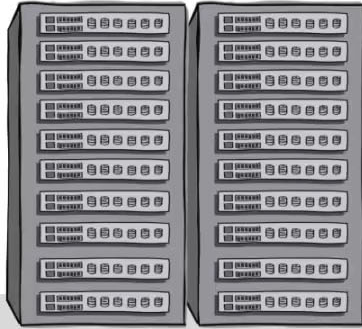


BIG DATA

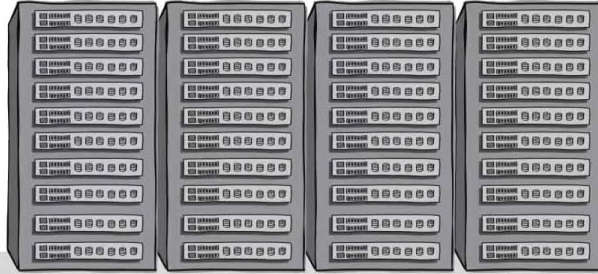


amazon

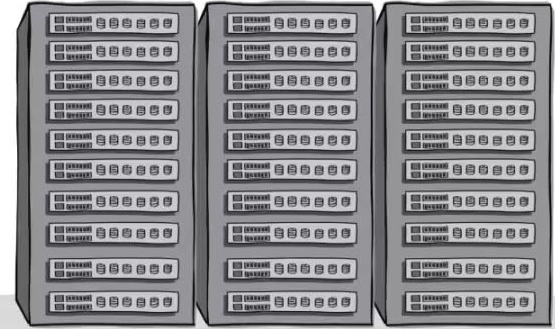
WORKLOAD 1



WORKLOAD 2

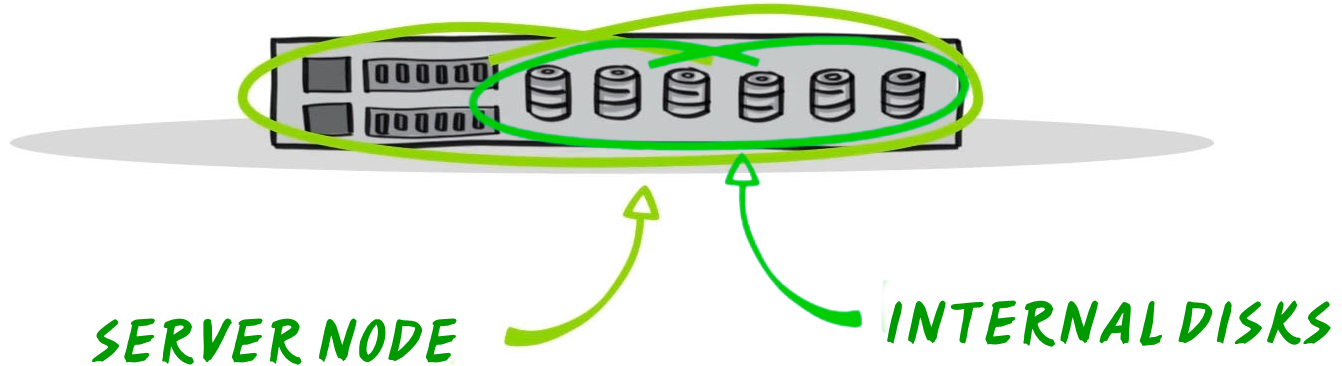


WORKLOAD 3



DEPLOYING PHYSICAL COMMODITY SERVERS
FOR EACH OF THEIR WORKLOADS

"SCALE-OUT" ARCHITECTURE

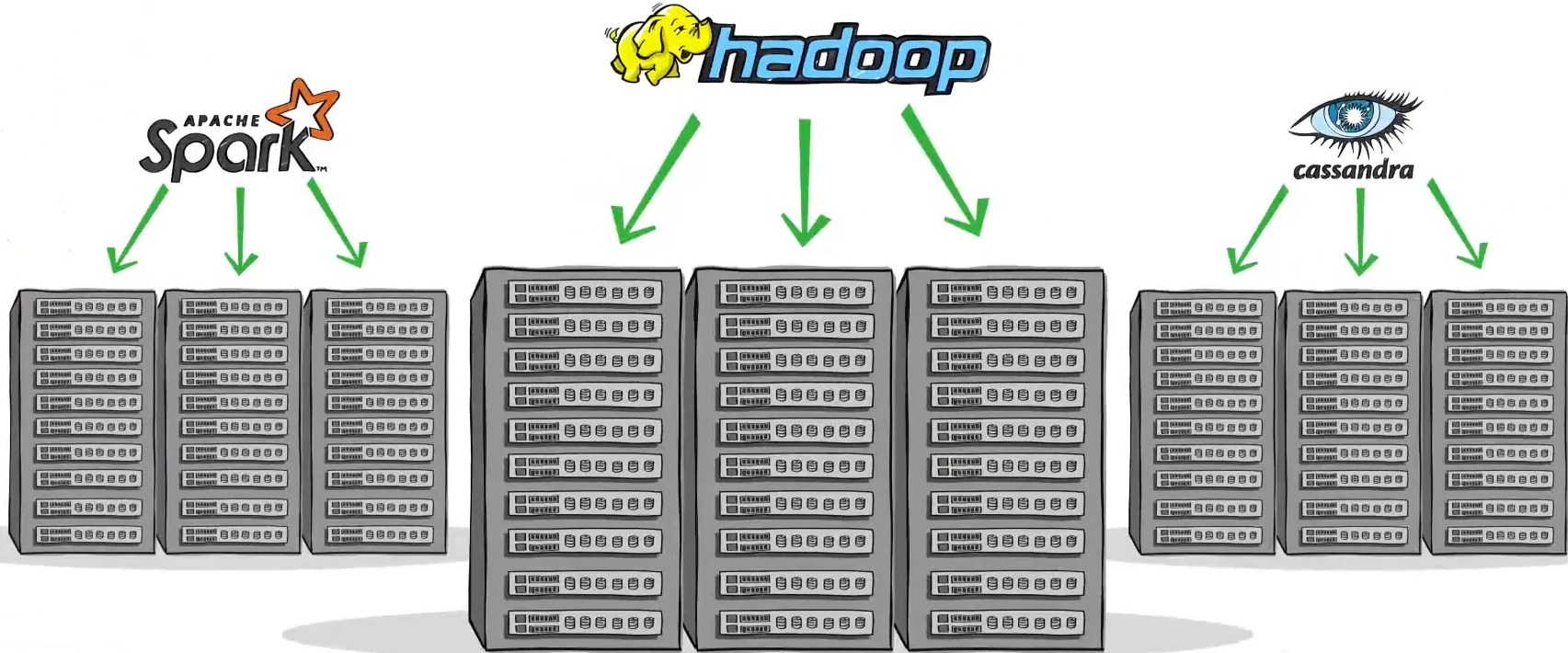




LESSONS LEARNED

- **BIG DATA APPS PARTITIONED THE WORK AND SCALED-OUT**
 - **A NEW CLASS OF BARE METAL COMMODITY APPLICATION ARCHITECTURE**
- **OLD WAYS WERE ANATHEMA TO THE NEW BIG DATA APPS**
 - **E.G. HADOOP'S STORAGE STACK OPTIMIZED FOR SEQUENTIAL I/O ENABLING LOCAL DISK**
 - **LEGACY NETWORKED STORAGE ACTUALLY DECREASED PERFORMANCE AND INCREASED COST**
 - **VIRTUALIZATION UNNECESSARY**
- **REQ'D DURABILITY OF THE DATA IS DAYS NOT YEARS**

THERE IS NOT A "SINGLE" BIG DATA APPLICATION



BUT COMMON TO ALL WAS ETHERNET-BASED SCALE-OUT

EACH APPLICATION HAS DIFFERENT REQUIREMENTS

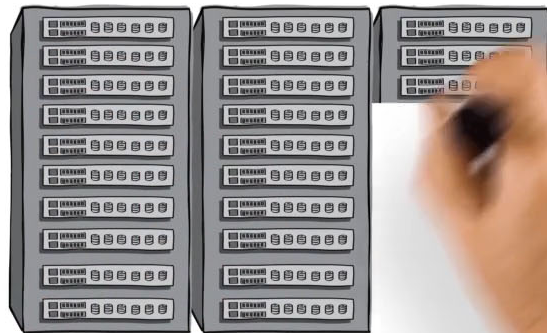
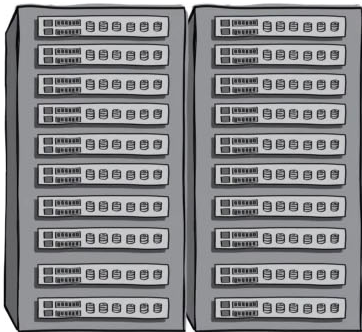
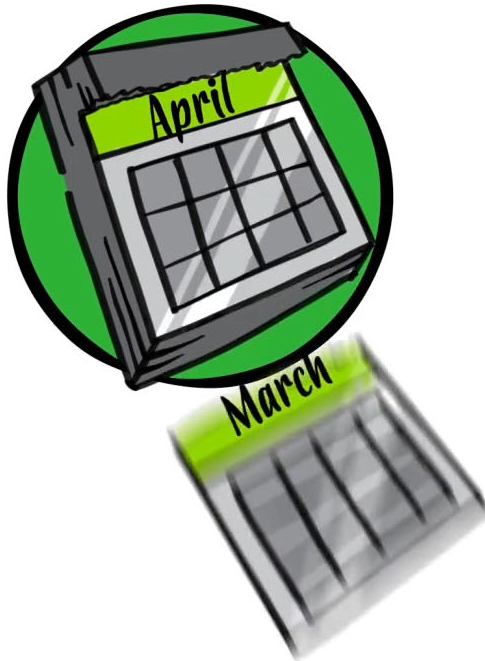
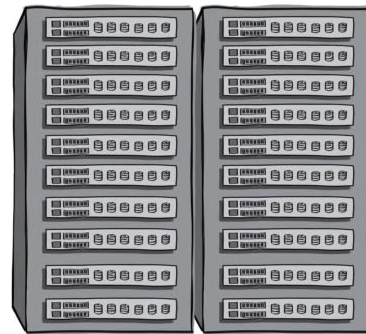
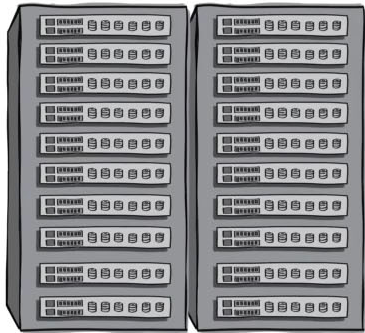
WORKLOAD 1

WORKLOAD 2

WORKLOAD 3

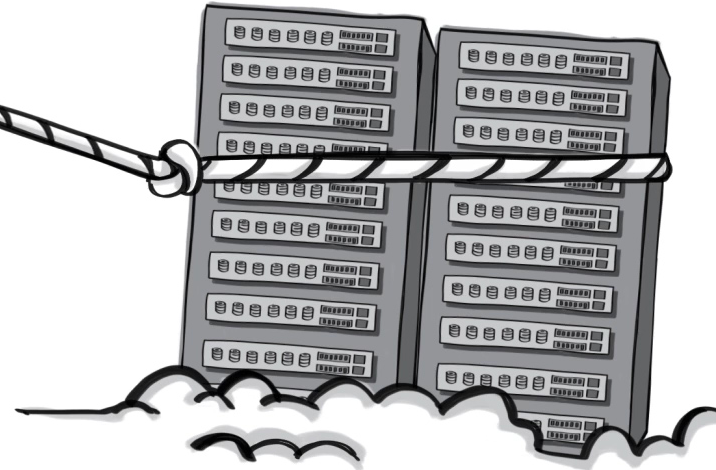
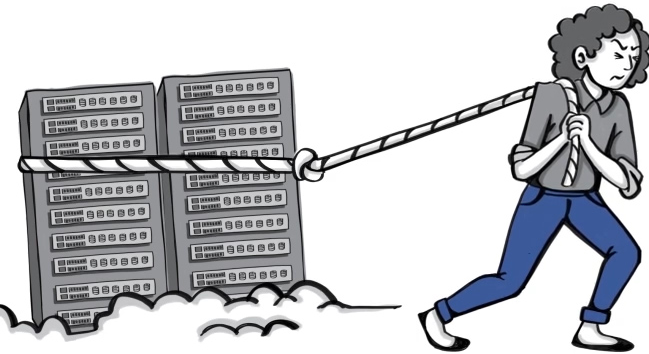


CHANGE IS THE ONLY
CONSTANT IN
BIG DATA





**BARE METAL CAN BE AN EXPENSIVE
AND HEAVY LOAD**



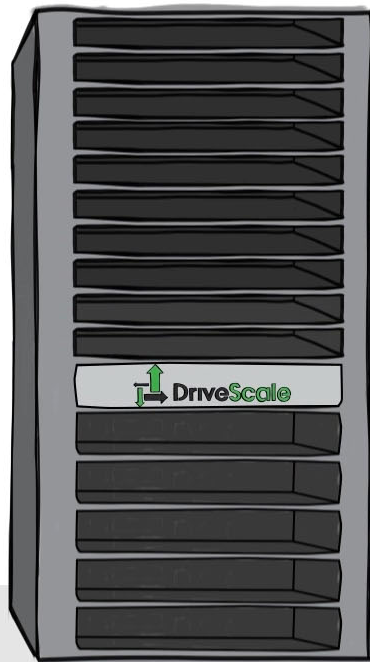
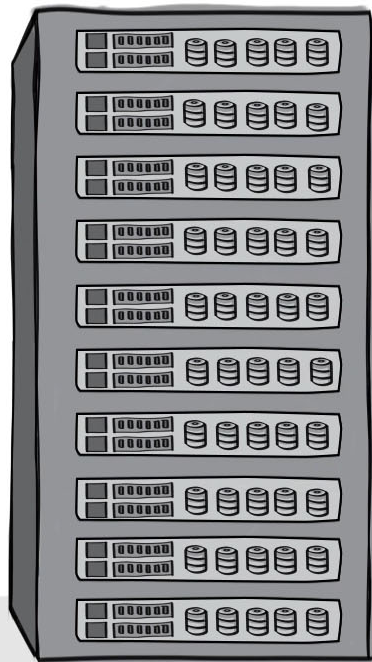


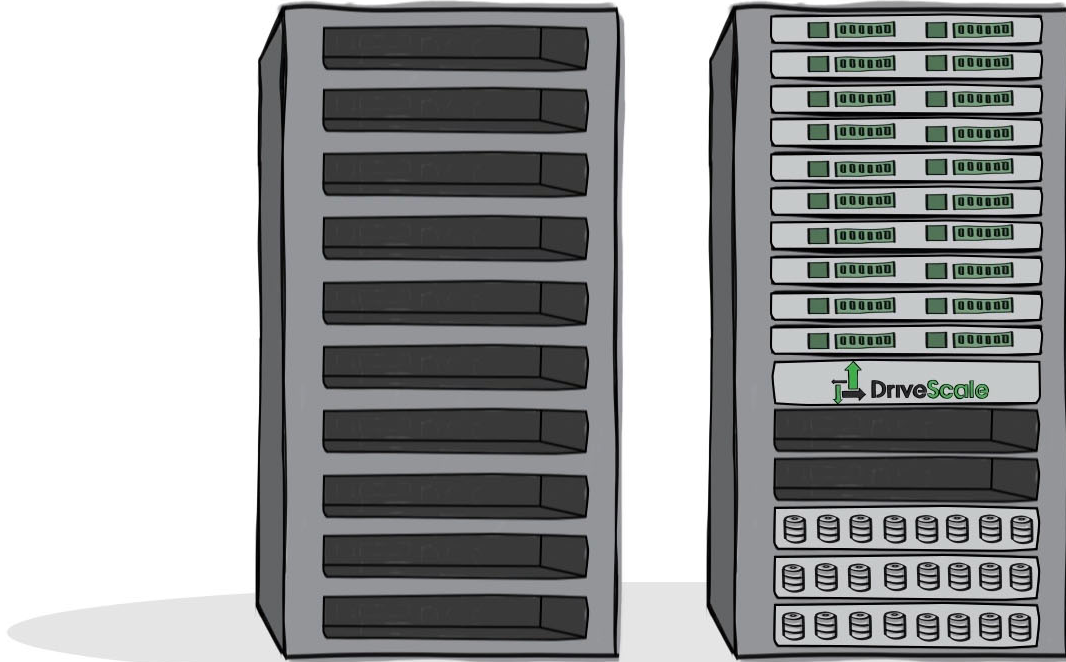
LESSONS FORGOTTEN

- THERE IS NO SUCH THING AS A SINGLE COMMODITY SERVER IN THE TRADITIONAL DATA CENTER WORLD
 - IT WAS HADOOP AND THEN TEN MORE APPS EACH WITH THEIR OWN UNIQUE COMPUTE/STORAGE RATIO REQUIREMENTS
- RESOURCE UTILIZATION DROPPED, OVER-PROVISIONING ROSE
- PEOPLE LOST THE RAPID DEPLOYMENT AND APPLICATION MANAGEMENT THAT CAME WITH VIRTUALIZATION

MOVE FROM
THIS...

THICK FIXED
CLIENTS/HCI





TO THIS!

THIN CLIENTS
AND OPTIMIZED
STORAGE

POOLS OF
DISAGGREGATED
COMPONENTS

SOFTWARE COMPOSABLE INFRASTRUCTURE

SECURE VIRTUAL CLUSTERS

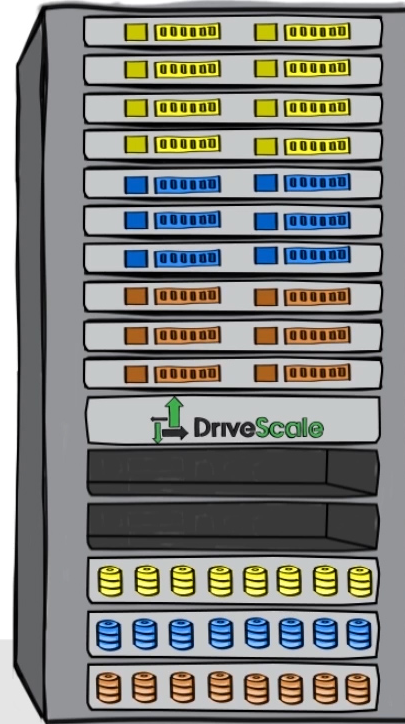


cassandra

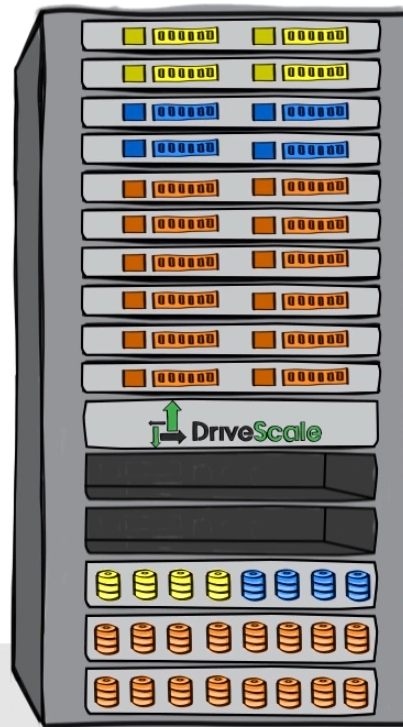


hadoop

APACHE
Spark[™]



EASILY RECONFIGURE BARE METAL CLUSTERS THROUGH SOFTWARE



SOFTWARE ORCHESTRATION
RETURNS TO A VIRTUALIZATION
EXPERIENCE

QUICKLY
SPIN UP NEW APPLICATION CLUSTERS

MOVE RESOURCES
BETWEEN CLUSTERS

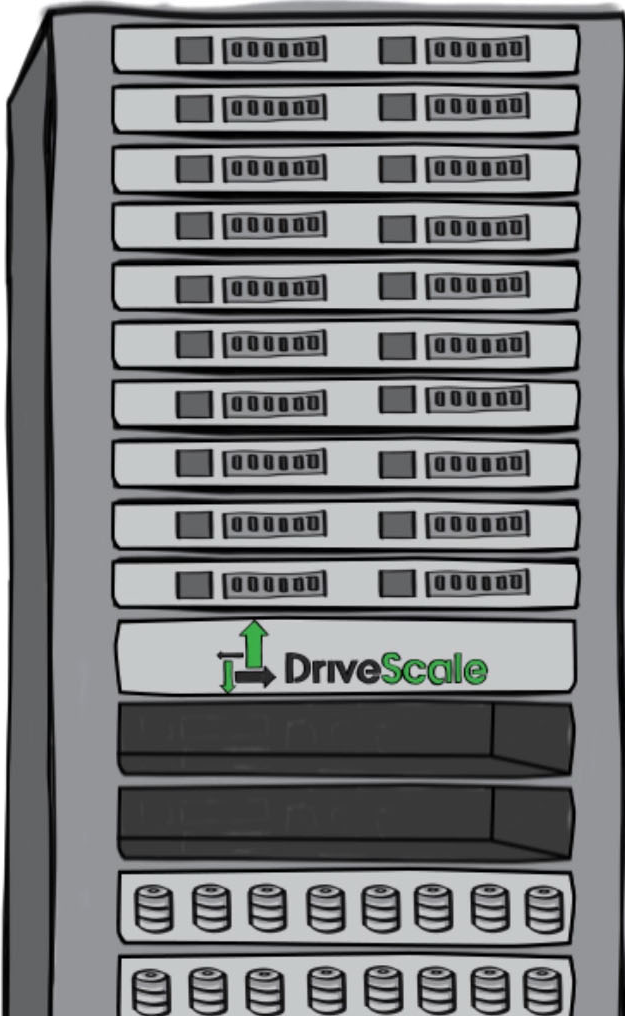
RECONFIGURE ON-DEMAND
EXISTING CLUSTERS

REPLACE
FAILED NODES BY REASSIGNING STORAGE

UPGRADE SEAMLESSLY
TO NEW COMPUTE NODES

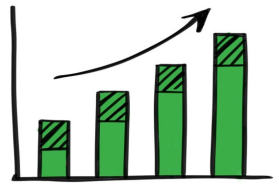
MINUTES!
~~WEEKS~~





COMPUTE

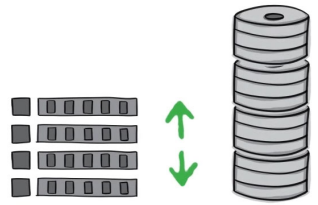
STORAGE



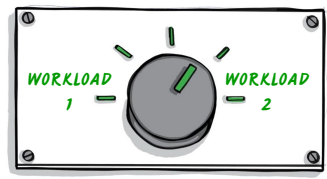
IMPROVE
RESOURCE UTILIZATION



AVOID
OVER PROVISIONING



SCALE RESOURCES
INDEPENDENTLY



MOVE RESOURCES
BETWEEN
WORKLOADS



SOFTWARE COMPOSABLE INFRASTRUCTURE

- REDUCED COST THROUGH CAPEX AND OPEX SAVINGS
 - OPERATIONAL EFFICIENCY GAINS OVER STATIC NODE/HCI
 - ELIMINATES OVER- AND UNDER-PROVISIONING
 - SEPARATES LIFECYCLE MANAGEMENT OF DISAGGREGATED COMPONENTS
- BIG DATA IS RIPE FOR OPTIMIZATION (GREENFIELD/REFRESH)
 - EACH APP HAS STORAGE STACK / MANAGES OWN DATA REDUCTION
 - RATHER THAN LEGACY VM OR ENTERPRISE DEPLOYMENTS
- EFFORTS SUCH AS OPENFLEX PROMISE AN ENHANCED PLUG-N-PLAY SOFTWARE ORCHESTRATED FUTURE

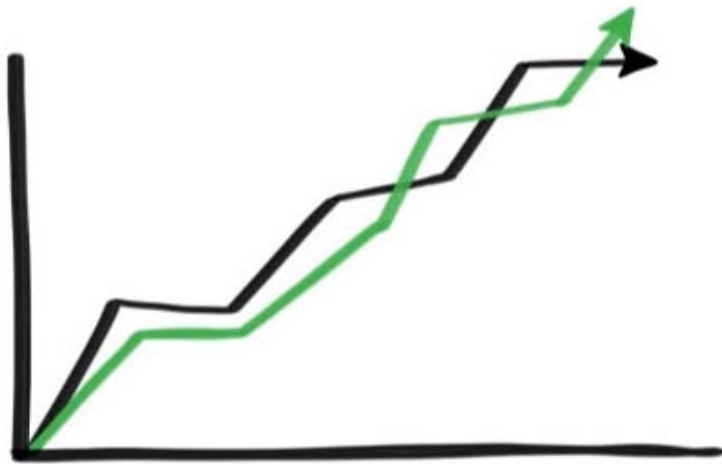


Flash Memory Summit

FLASH CONSIDERATIONS FOR SCI

- ***WHAT IS SPECIAL ABOUT FLASH (SOLID STATE STORAGE)***
- ***HOW CAN SOFTWARE COMPOSABLE INFRASTRUCTURE HELP***

FLASH IS FOR PERFORMANCE

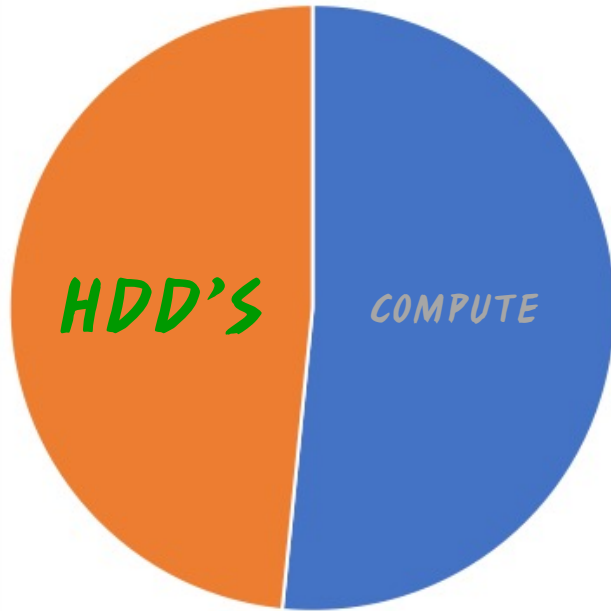


■ COMPOSABLE FLASH DRIVE

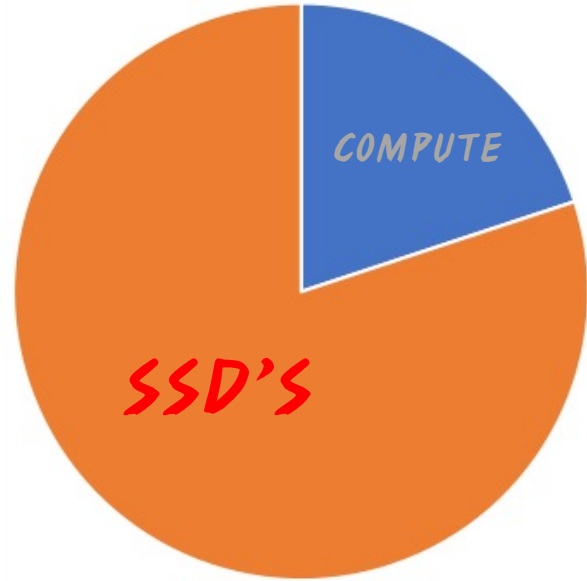
■ INTERNAL FLASH DRIVE

- NATIVE NVME DRIVES IN COMPOSABLE INFRASTRUCTURE CAN BE FULLY UTILIZED
- 100GBE FABRIC
- OPTIMIZED ISCSI STACK AND MULTIPLEX
- KILL UNNEEDED "DATA MGMT" FEATURES IN DATA PATH FOR BIG DATA APPS
- NVME OVER FABRIC – NOW ROCE V2
LATER NVME OVER TCP (ETHERNET WON)

THE PROBLEM TO SOLVE USE FLASH EFFICIENTLY



\$32,000



\$84,000

ROUGH CALCULATIONS, 24 X 2TB DISK RACK SERVER, DUAL SOCKET

ATTACKING OVERPROVISIONING

- RIGHT-SIZING FLASH FOR APPLICATIONS REMAINS YOUR BIGGEST LEVER
 - TRADITIONAL CAPTURED DAS RESTRICTS COMPUTE/SSD RATIO AT INITIAL DEPLOYMENT
 - DISAGGREGATE FLASH TO ALLOW (RE)COMPOSABILITY
 - ADD'LY CARVE AND SHARE SINGLE SSDS AMONGST APPLICATIONS
- EXPLOIT COST EFFICIENT FLASH FORM FACTORS AND TECHNOLOGIES
 - HDD FORM FACTOR FIXED / FLASH EMINENTLY MALLEABLE
- RADICALLY REDUCE YOUR COMPUTE COSTS TOO
 - "TIME SHARE" COMPUTE BETWEEN APPLICATIONS
 - "BURST" PROVISION ADDITIONAL COMPUTE ON A TEMPORARY BASIS

OPERATIONAL EFFICIENCIES

- REFRESH COMPUTE AND FLASH SEPARATELY
 - YOU CAN'T AFFORD TO THROW AWAY SSDS WITH COMPUTE
- PAY-AS-YOU-GROW FLASH AS APPLICATION DEMANDS CHANGE
 - AVOID UP FRONT OVERPROVISIONING OF CAPTURED SSD
- MONITOR AND WEAR LEVEL GLOBALLY ACROSS POOLS OF FLASH ACROSS YOUR APPLICATIONS
- OPTIMIZE DATA CENTER REAL ESTATE
- MANAGE MORE WITH LESS THROUGH AUTOMATION



RECAP

- SOFTWARE COMPOSABLE INFRASTRUCTURE REDUCES COSTS
- INCREASES AGILITY IN DEPLOYING BIG DATA APPLICATIONS
- SEPARATION OF COMPUTE AND STORAGE REDUCES COST TO DEPLOY FLASH
 - PROVISION JUST THE RIGHT AMOUNT OF FLASH
 - DELIVER LOCAL FLASH PERFORMANCE OVER THE NETWORK
 - MANAGE COMMODITY THIN COMPUTE CLIENTS SEPARATE FROM EXPENSIVE FLASH
 - COST-EFFECTIVE DENSE FORM FACTORS AT THE STORAGE LAYER



Flash Memory Summit

QUESTIONS?