

Best Practices for Configuration on Symmetrix DMX and Symmetrix V-Max Series with Enginuity

John Adams

EMC WORLD
ORLANDO 2009

Objectives

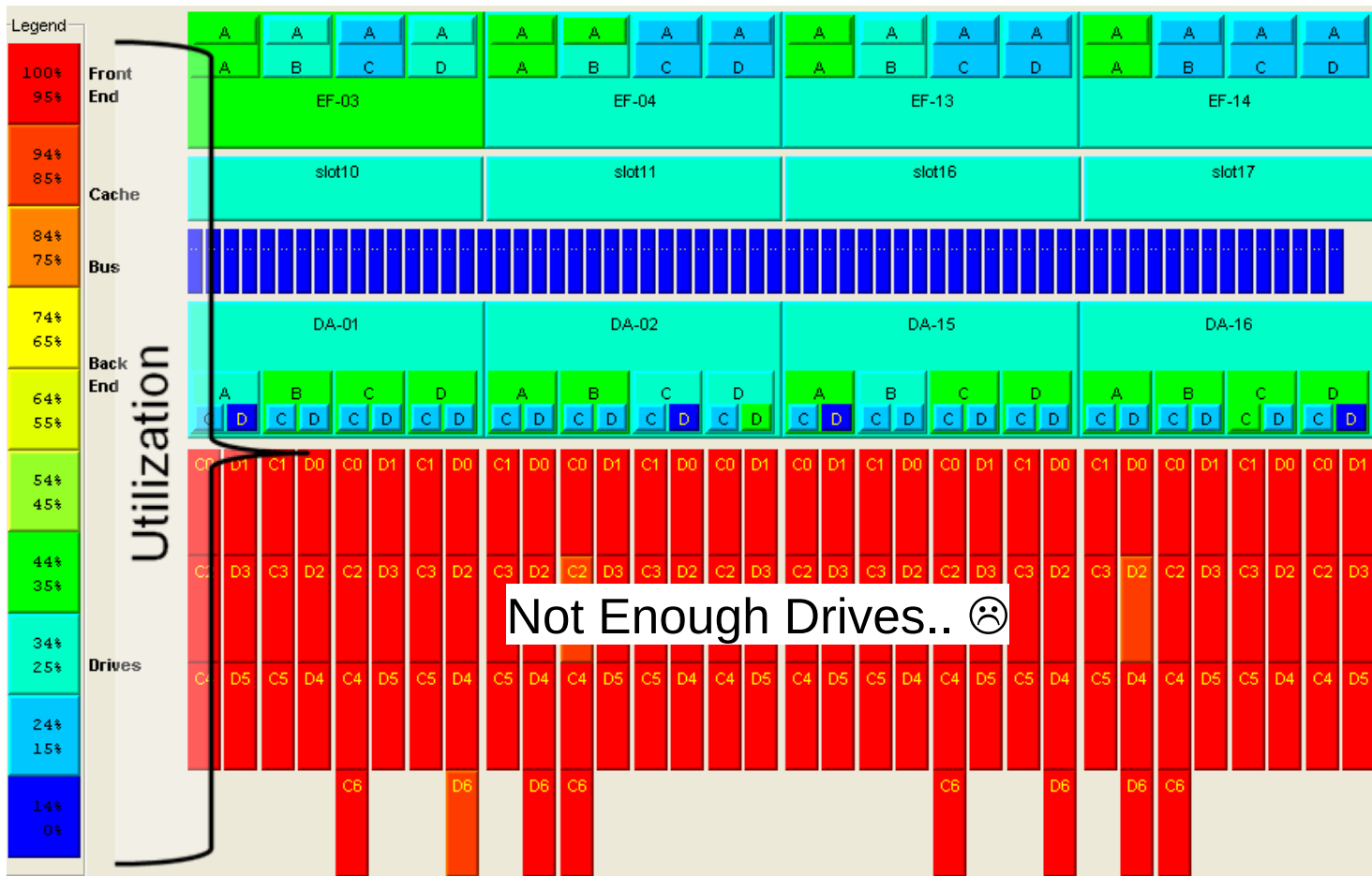


At the end of this session, you will be able to

- Understand Symmetrix Configuration and Enginuity features that may affect performance and data layout decisions
- Plan configuration decisions more effectively to obtain optimal performance from your Symmetrix
- Have some "Rules of Thumb" to use to make choices in configurations

- Basic Rules of Thumb for Configurations
 - Configure Enough Resources
 - Utilize Resources Evenly
 - Simple Data Layouts Are Better
- Common configuration questions covering:
 - Sizing system
 - Optimum hyper size/number, queue depth, number of Meta members
 - Meta Volumes
 - Vs. host based striping
 - Striped or Concatenated
 - Stripe size
 - Segregate volumes
 - Disk type and RAID protection
 - Cache
 - Utilizations

Configure enough resources for your workload



Front End
(Hosts)

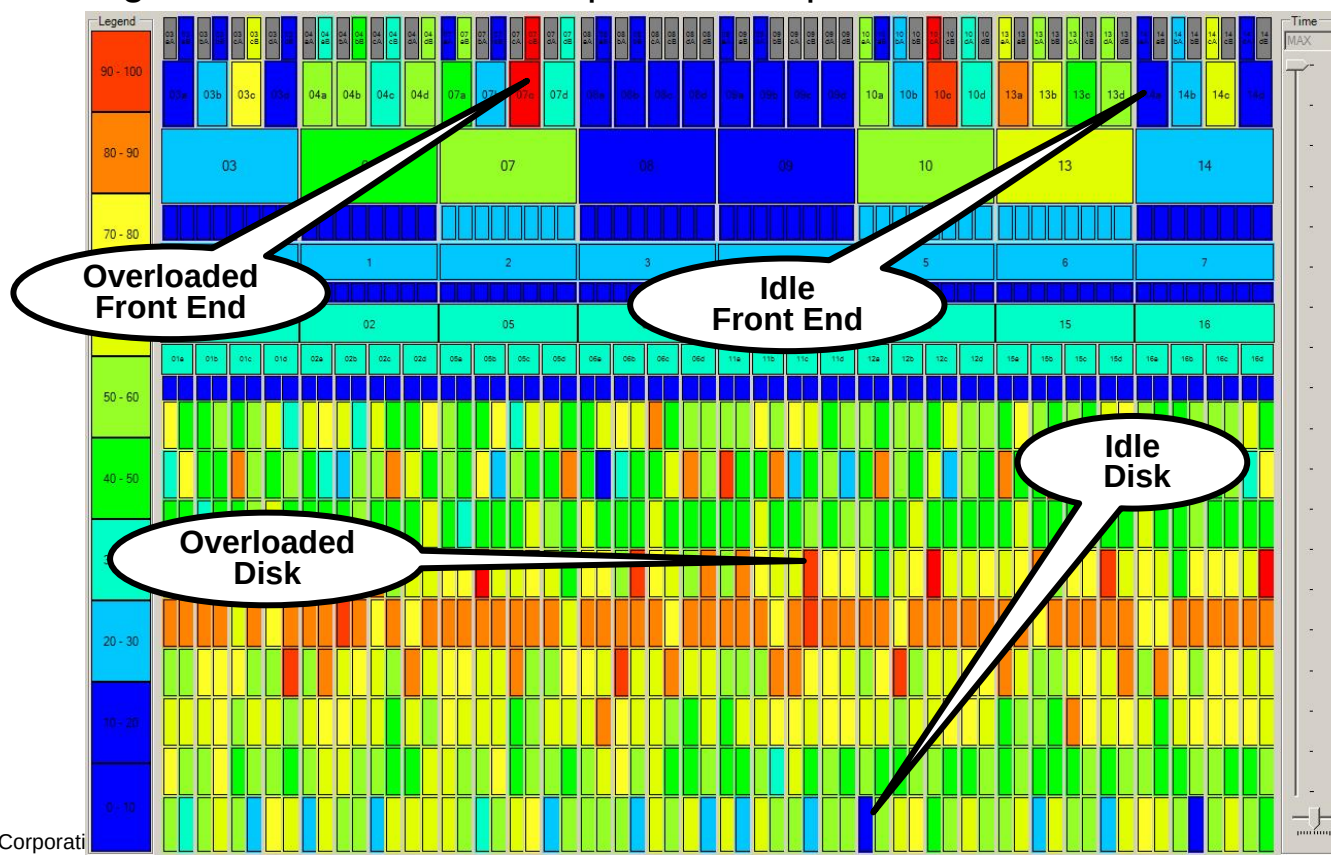
Cache &
Matrix

Back End
(DAs)

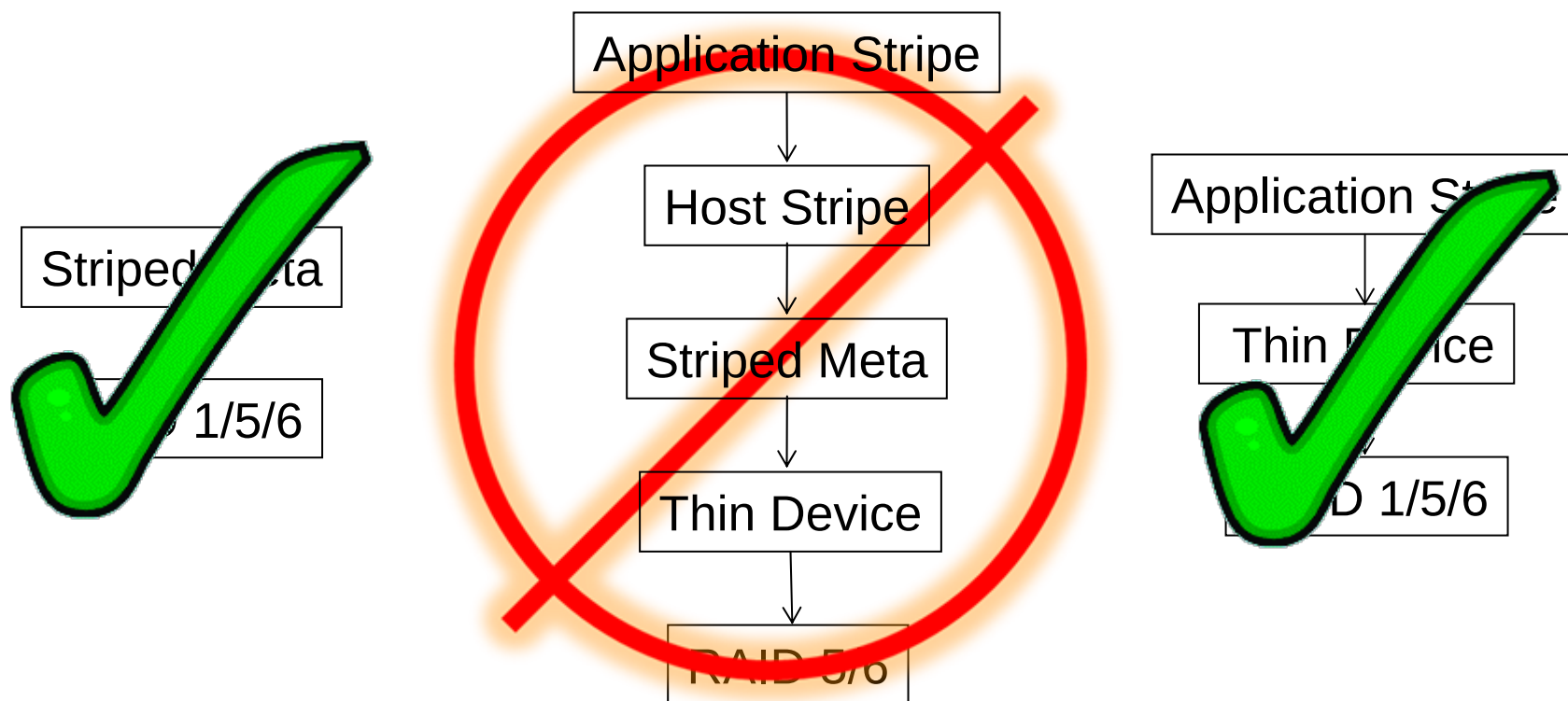
Disks

Use those resources evenly

- You get the best overall performance by spreading the workload across all available components
- This includes, Front-End directors, Back-End directors and disks
 - Path management software can help the FE, Optimizer the BE



Simple is usually better



Commonly asked questions



- What size system do I need?
 - Front-End slices/boards, Back-End slices/boards
- What is the optimum hyper size or number of hypers on a disk?
- What is the optimum queue depth?
- What is the optimum number of members in a Meta Volume?
- Should I use Meta Volumes or host based striping? Or both?
- Striped or Concatenated Metas?
- What stripe size should I choose?
- Should I segregate my “X” volumes from my “Y” volumes?
 - Clone targets from Source volumes, etc.
 - On different disks
 - On different DAs
- What disk class should I choose?
- What RAID protection should I choose?
- How much cache do I need?
- How do I know when my system/components is/are reaching maximum utilizations?

What size system do I need?



- Each resource has a limit of I/Os per second and MBs per second
 - Disk Drives
 - Back-End controllers (DAs)
 - Front-End controllers (Fibre, FICON, GigE)
 - RDF controllers
 - “Slices” in this presentation refer to Front-End, Back-End and RDF CPU complexes
- Configure enough components to support your workload peaks
 - Tools available to help you determine how many you need
- Use those resources as uniformly as possible
 - Skewed workloads are limited by most heavily utilized components

What size system do I need?

Front-End connections



- Go wide before you go deep
 - Use all the “zero” ports on director first and then the “one” ports
 - Spread across directors first, then on the same director
- Each front-end port/slice has performance limits (see table below)
- Two active ports on one FA slice do not generally do more IO/s
- Front-end slices have different performance for different workloads
- For example (all with small I/Os on one “slice” Normalized to Read Hit)

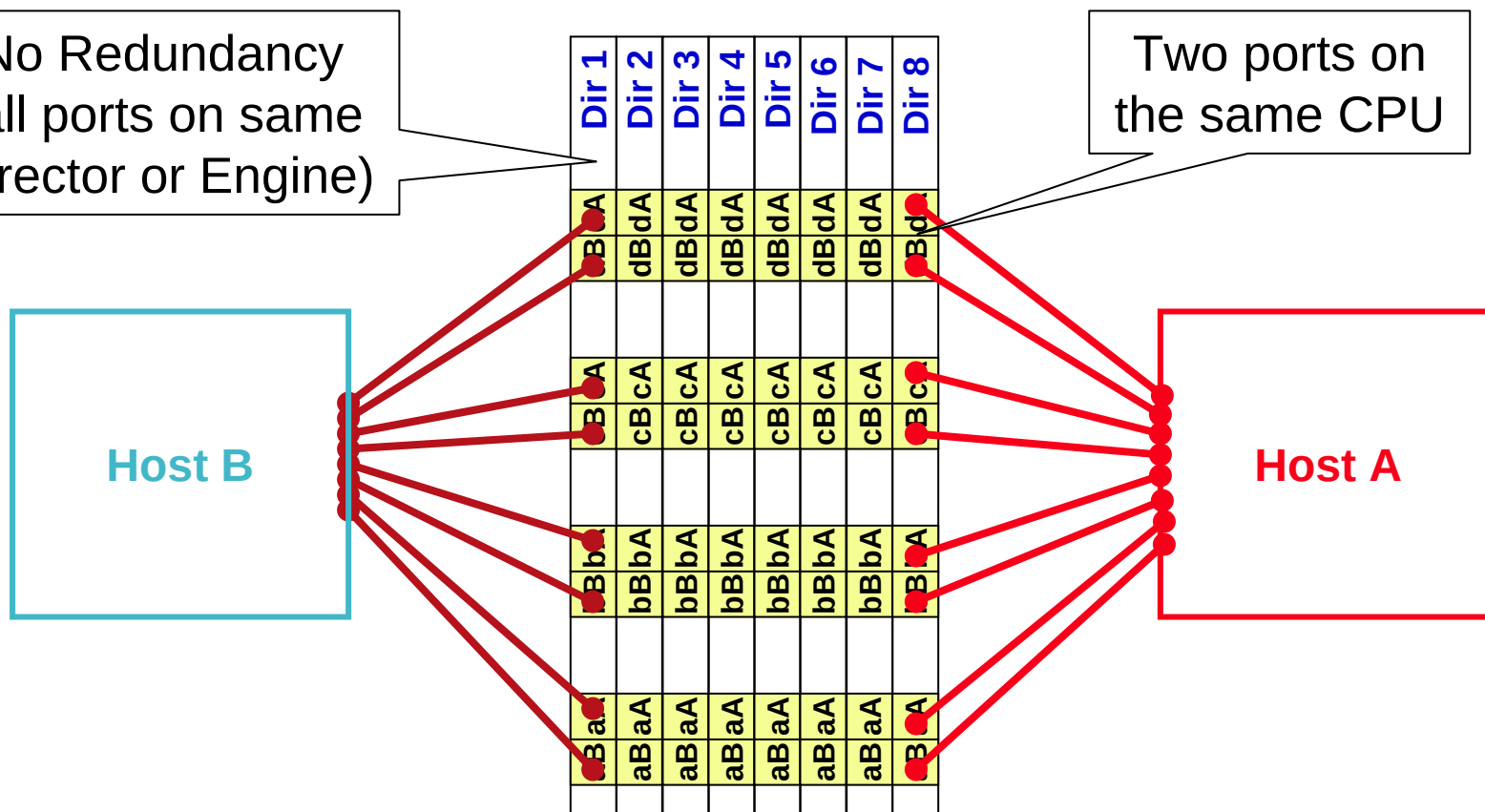
FA Slice Performance	Symmetrix DMX-4 System	Symmetrix V-Max System
Random Read Hit IO/s	1	1
Random Read Miss IO/s	1/2	1/2
Random Write Hit IO/s	1/2	1/2

- You also get increased availability if you spread wide first

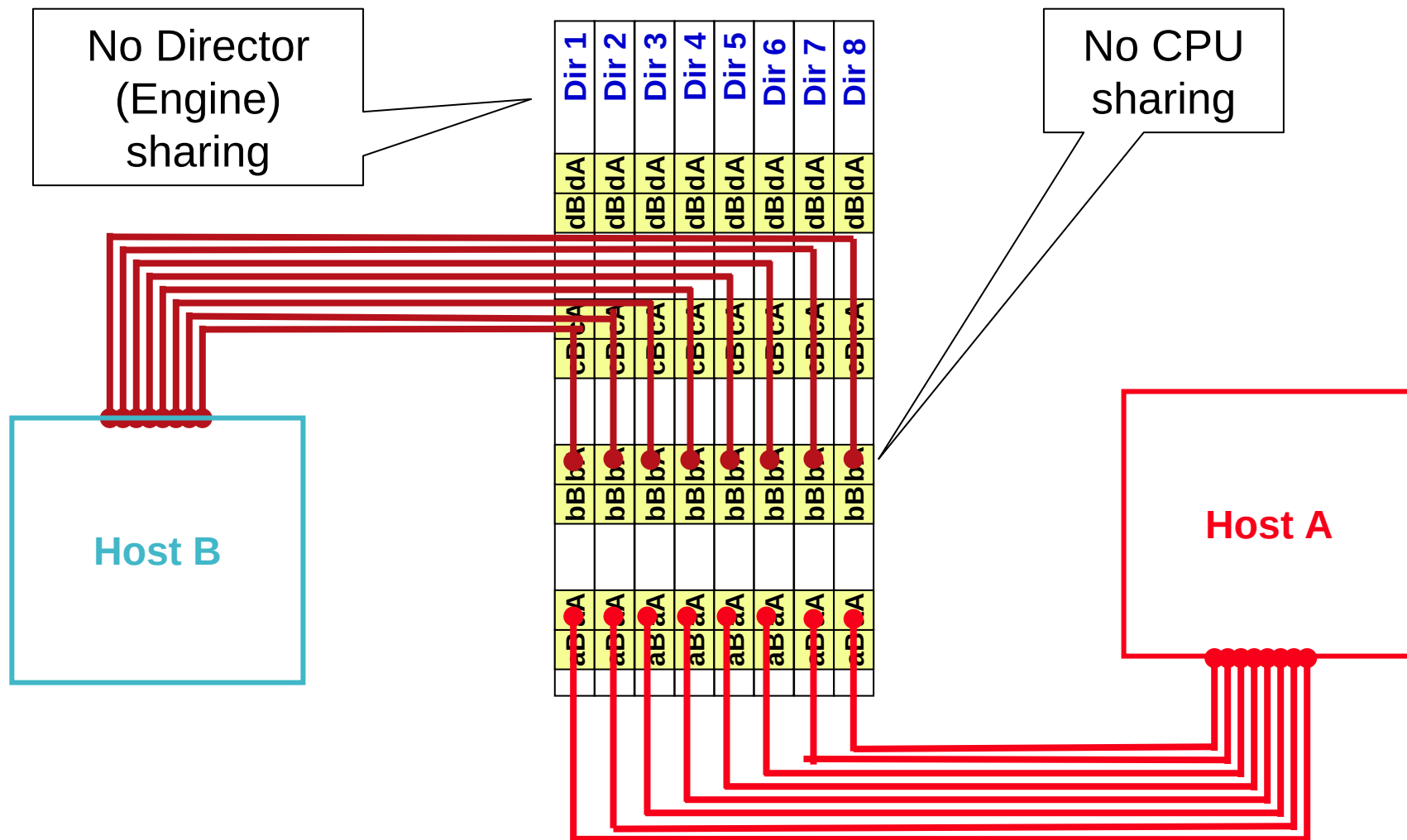
Worst Connection: Two hosts with eight Fibre Channel adapters each

No Redundancy
(all ports on same
Director or Engine)

Two ports on
the same CPU



Best Connection: Two hosts with eight Fibre Channel adapters each

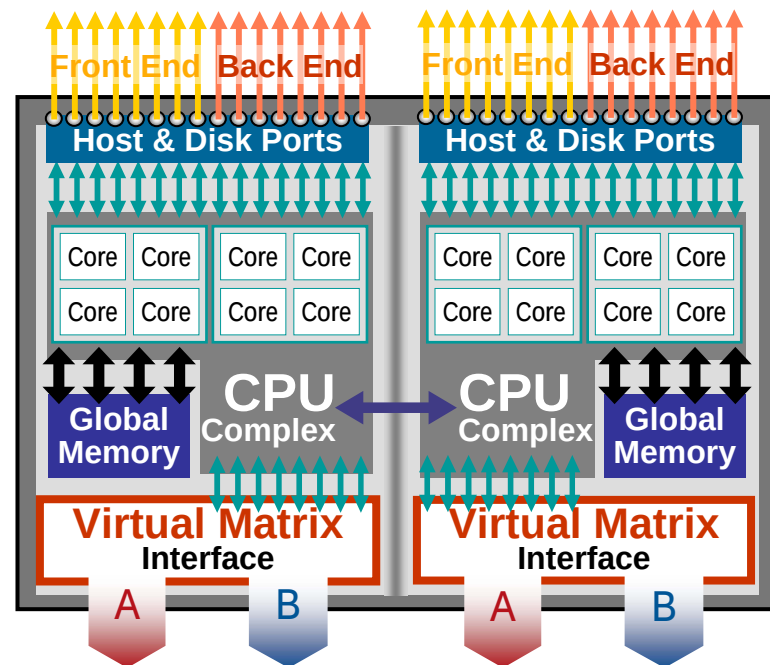


What size system do I need?

Symmetrix V-Max system “Locality”



- Symmetrix V-Max system directors have memory on the director board
 - In some cases, data can be accessed without going over the Virtual Matrix
- For data in global memory, we can do a local read if the data is in the local director, and the data is the primary copy
 - Each mirrored slot has a primary and a secondary location

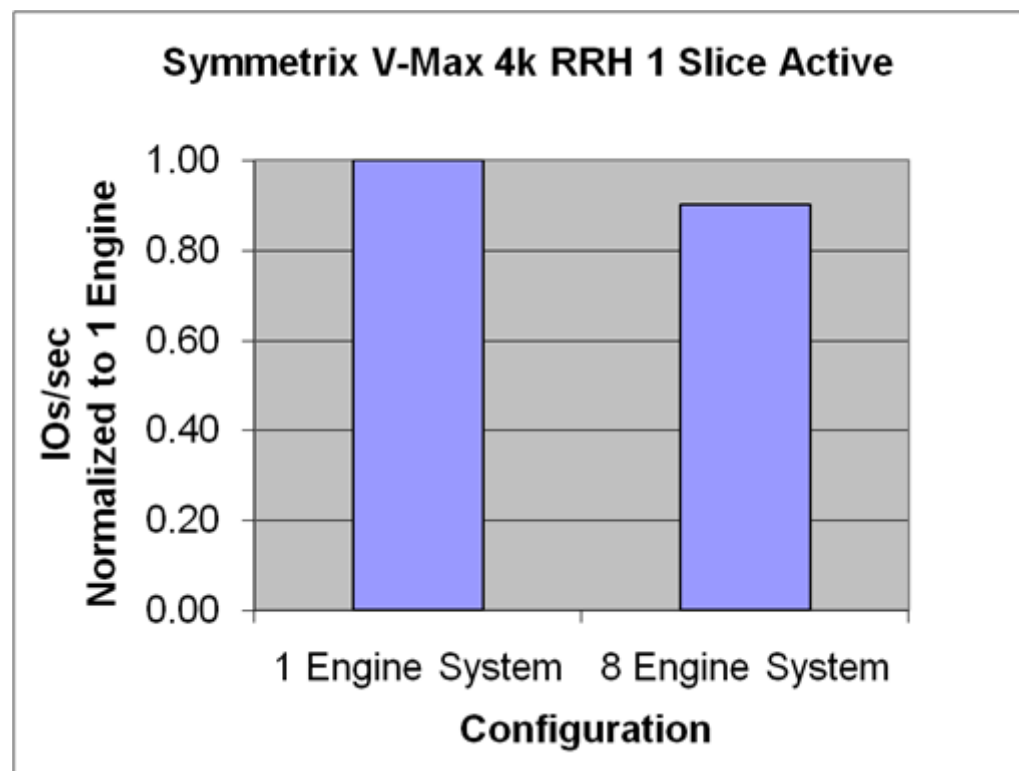


What size system do I need?

Locality with Read Hits on Symmetrix V-Max system



- 1 Active Slice on a 1 Engine System vs.
- 1 Active Slice on a 8 Engine System
 - 50/50 chance doing a local read on a single Engine system
 - 1 in 16 chance doing a local read on a 8 Engine system

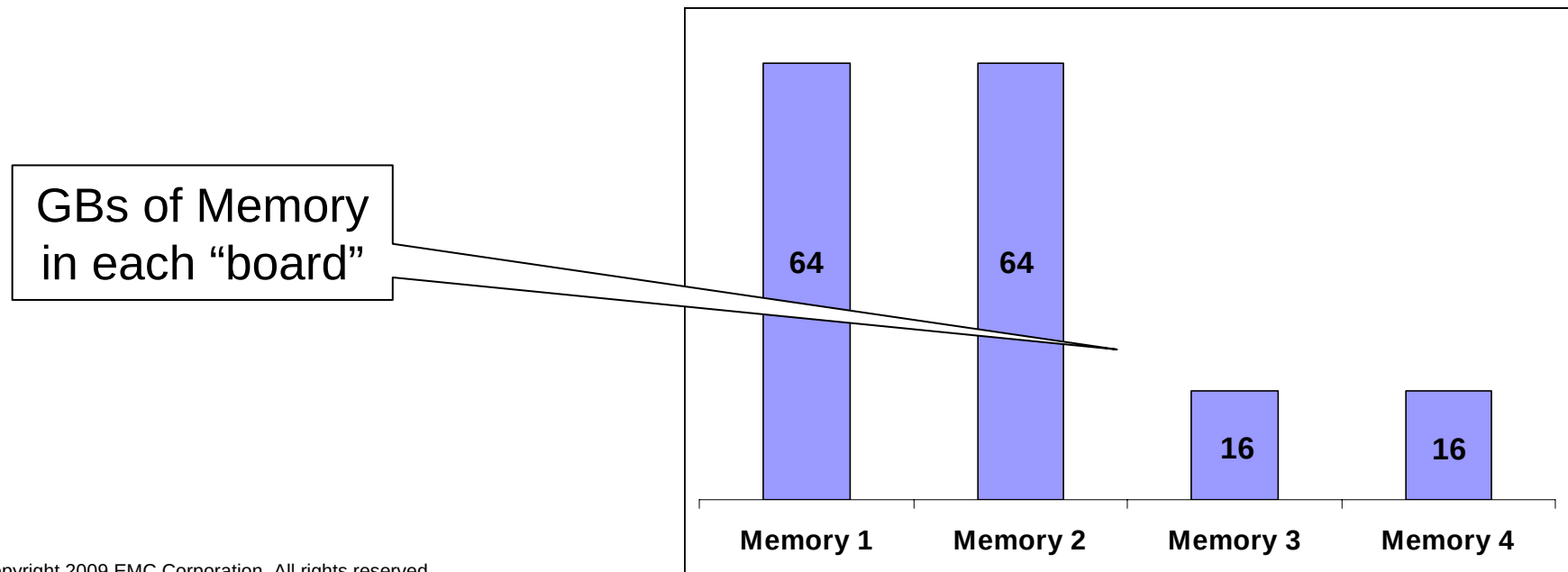


What size system do I need?

Memory considerations



- Symmetrix V-Max system or Symmetrix DMX system memory boards or engines should be the same size for maximum GB/s throughput
- Imbalance will make little or no difference with small I/O workloads
- Imbalance will create more accesses to the boards or engines with the larger amount of memory
 - Creating skewed distribution over the hardware resources



What size system do I need?

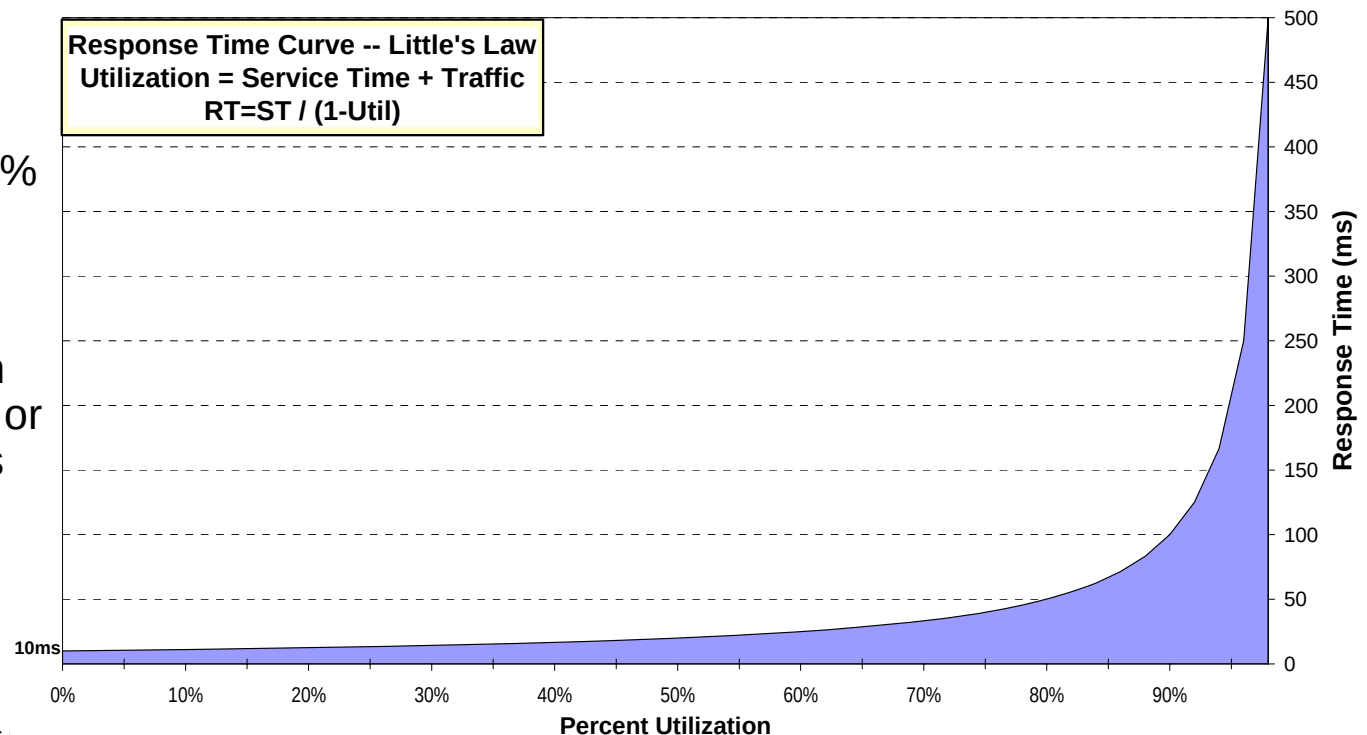
Utilization limits



- Back-end director slices and disks have performance limits as well
- It is important to understand the mixture of your workload and configure enough resources so that the array doesn't get close to 100% utilization

—As a rule of thumb, a limit of no more than 50% utilization is good if response time is very critical

—A higher utilization can be tolerated if only IO/s or total throughput matters



What size system do I need?

Disk drives



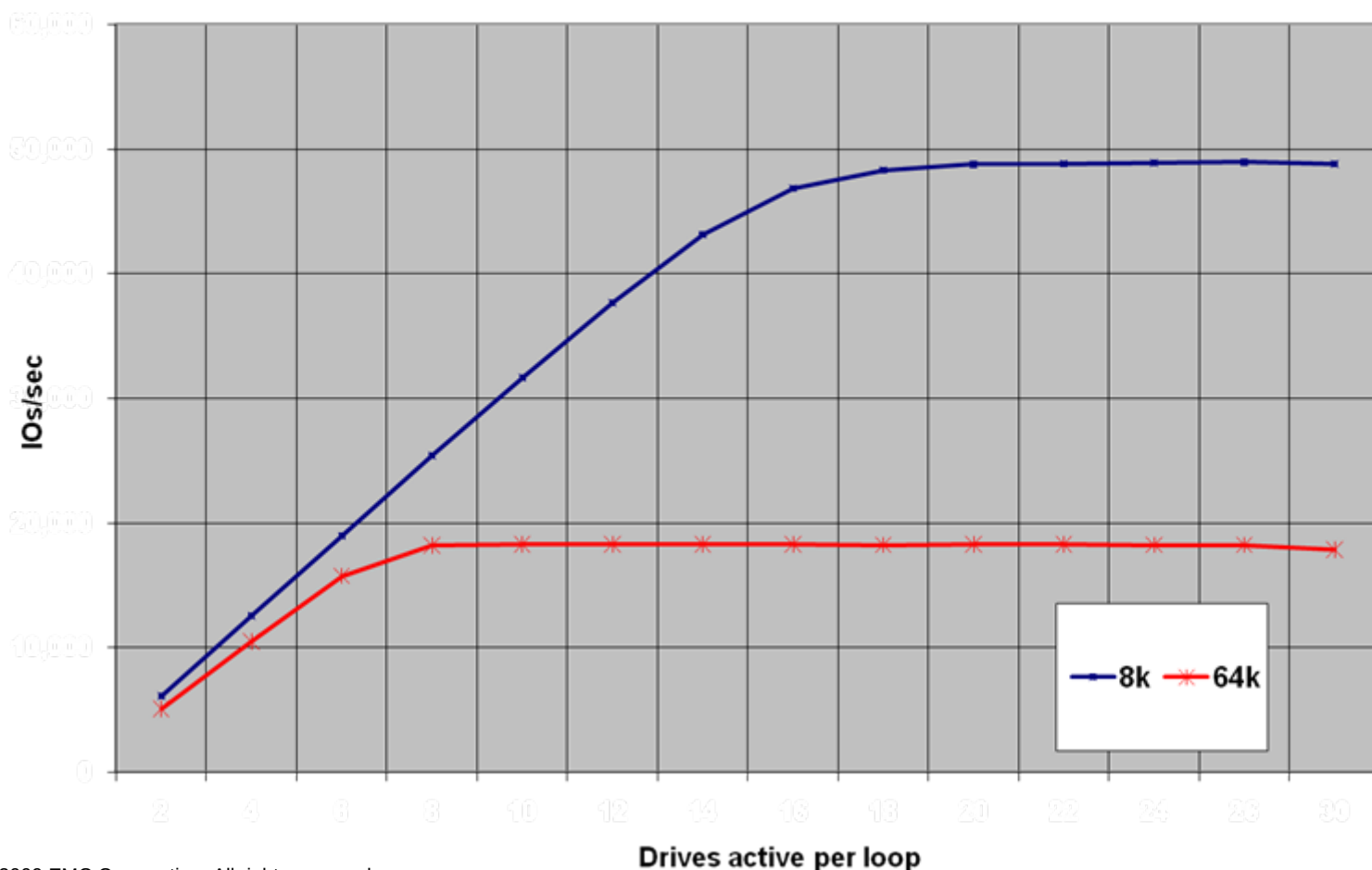
- How many disk drives do I need?
- Maximum performance limits scale linearly as you add drives until some director limit is reached
 - With small I/O sizes, the limit reached is the CPU limit
 - With larger I/O sizes, we can reach a throughput limit in the “plumbing” instead

What size system do I need?

Drive scaling 1 slice and whole Engine

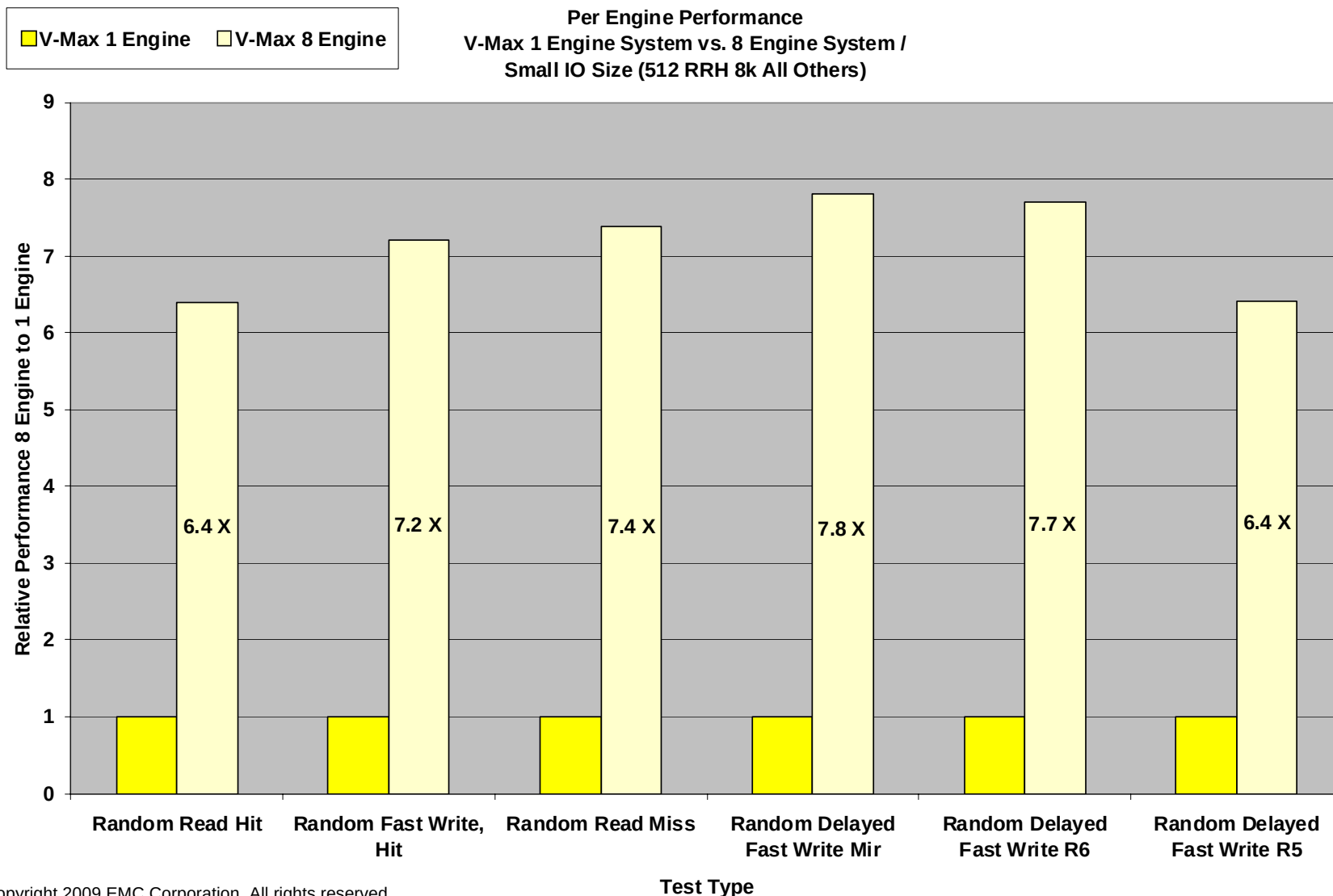


Random Read Miss Scaling
Symmetrix V-Max 1 Engine System
5874.121 Mirrored Devices



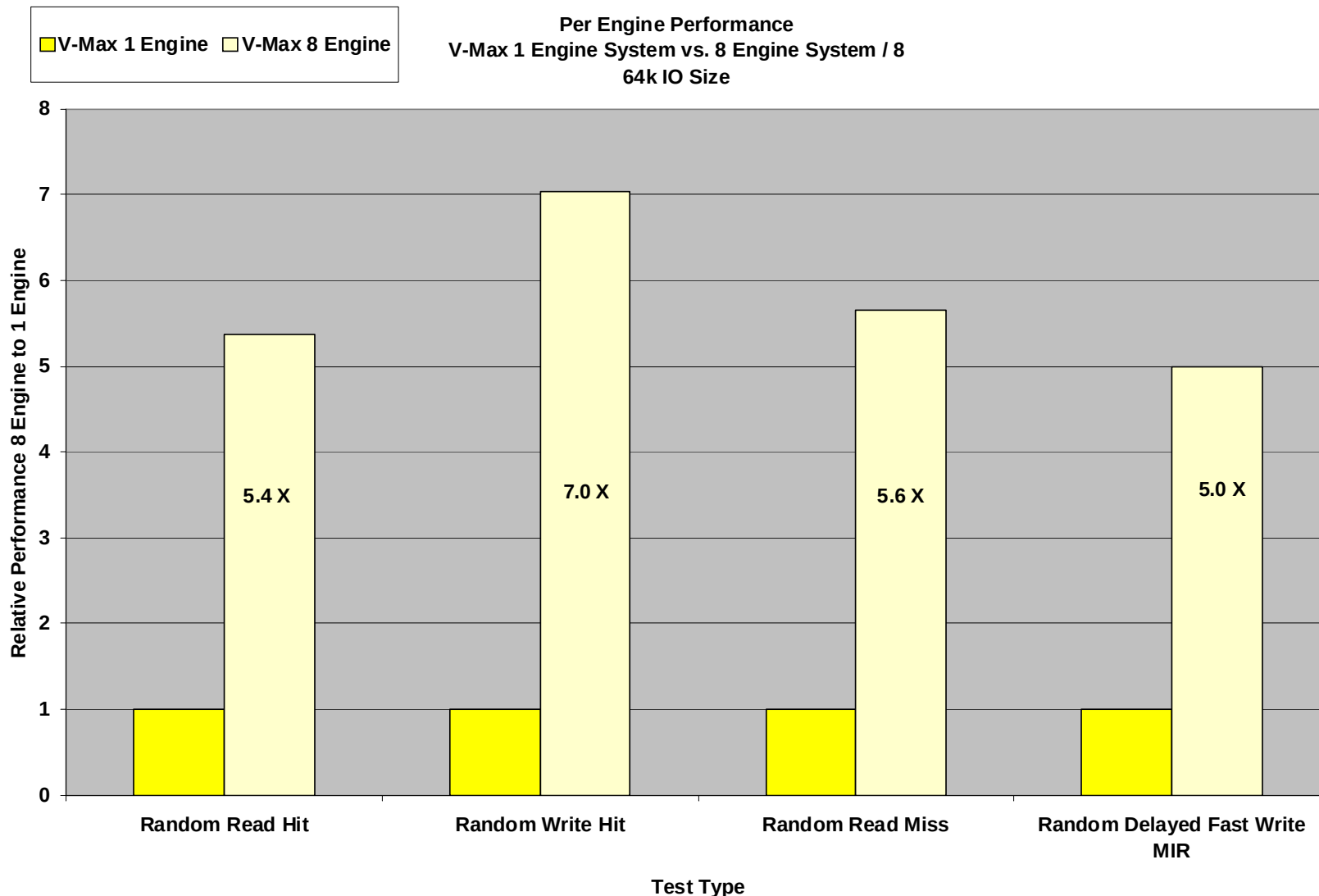
What size system do I need?

Engine scaling



What size system do I need?

Engine scaling

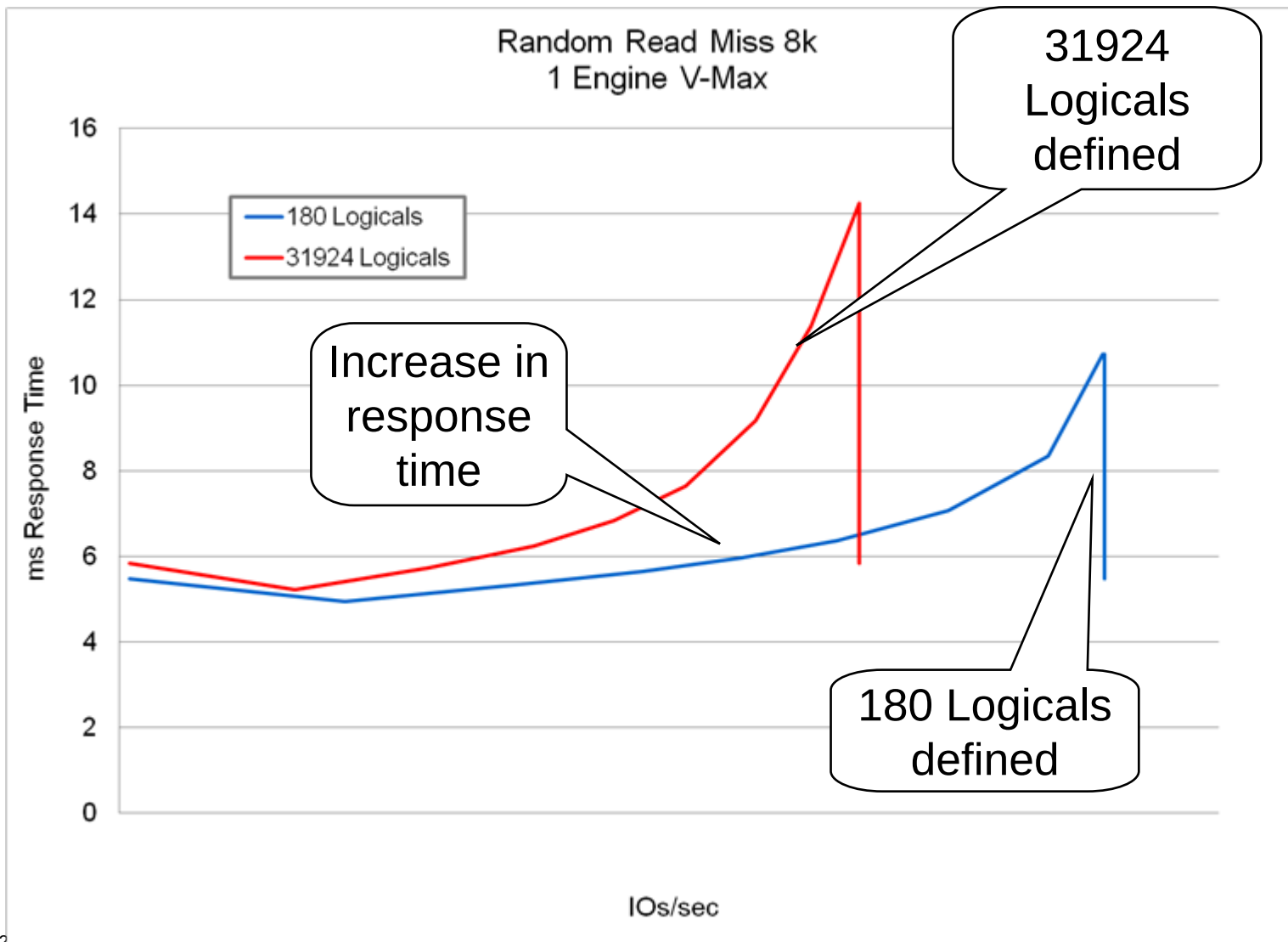


What is the optimum size of a hyper or numbers per disk drive?



- As a general rule of thumb, fewer larger hypers will give better overall system performance
 - There is system overhead to manage a logical volume, so it makes sense that more logical volumes could lead to more overhead
- Frequently legacy hyper size is carried forward because of migration strategy
 - However the same hyper size that made sense 10 years ago on a 9 GB disk drive, doesn't make sense on a 1 TB disk drive today

Symmetrix DMX System – What is the optimum size of a hyper or numbers per disk?

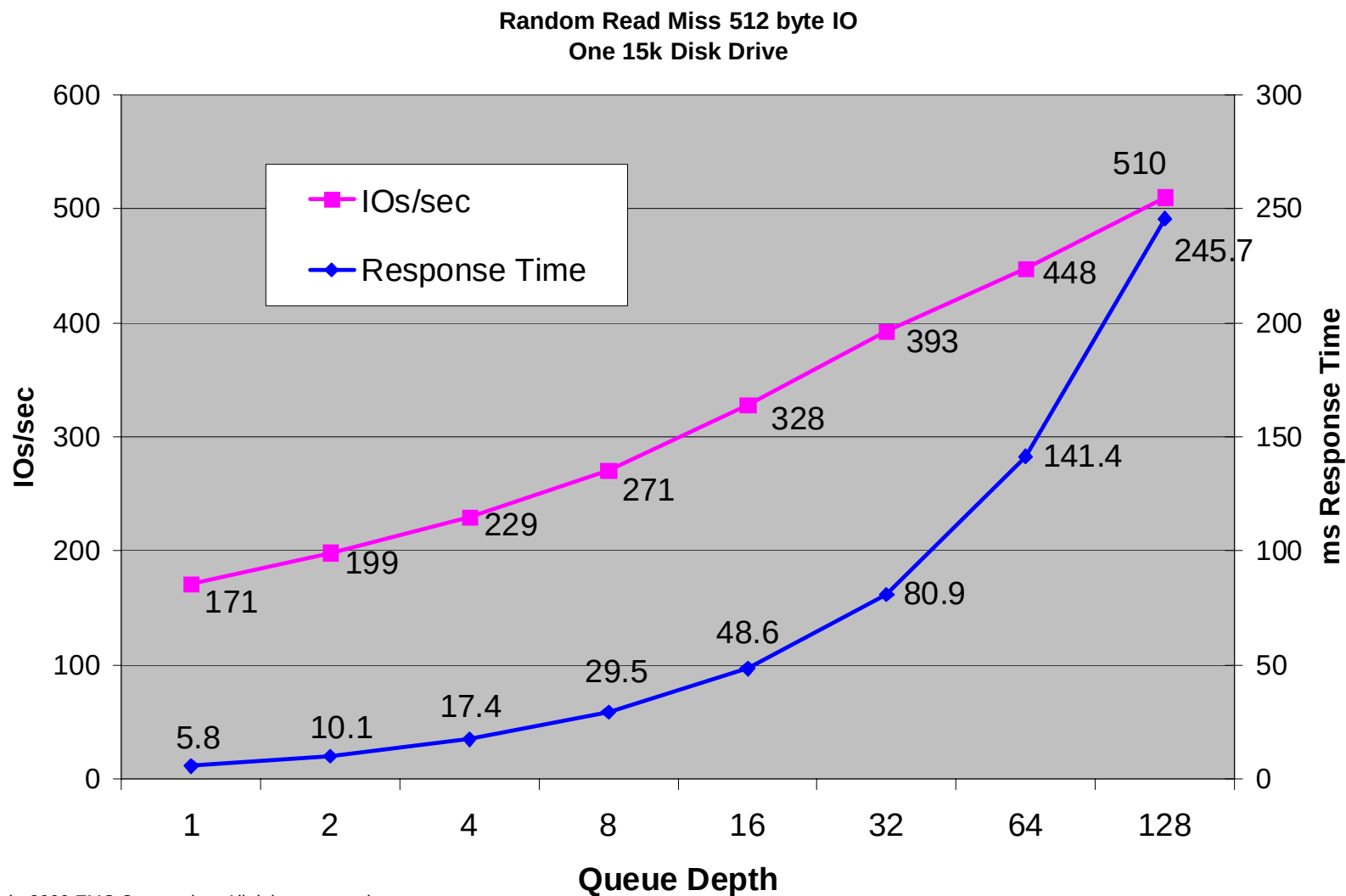


What is the optimum queue depth?



- To get the maximum performance from disk drives and the system, some queuing is necessary
- Single threaded, or 1 I/O at a time, the I/O rate is simply the inverse of the service time
 - For a 5.8 ms service time your maximum IO/s is 172 per second
 - For the same drive with 128 I/Os queued we can get nearly 500 IO/s
 - However the response time is about 250 ms at this I/O rate
- We need 1-4 I/Os queued to the disk to achieve the maximum throughput with reasonable latencies
 - Lower queue lengths if response time is CRITICAL
 - Higher if total I/O rate is more important than response time

What is the optimum queue depth?



What is the optimum number of members in a Meta Volume?



- Meta member counts should be “reasonable”
 - Enginuity supports up to a 255 member meta
 - That would create a single ~7.5 TB volume (Symmetrix DMX-2 systems or below) ,
 - Or a ~15 TB volume (Symmetrix DMX-3/4 systems)
 - Or a 60 TB volume (Symmetrix V-Max systems)
 - This is not what we consider reasonable
- Reasonable sizes for meta member counts are something like
 - 4, 8, 16 or 32
- Even numbers are preferred
 - Powers of 2 fit nicely into the Symmetrix system back-end configurations
- Getting enough I/O into a very large meta can be a problem
 - 32-way RAID 5 7+1 Meta Volume would need at least 256 I/Os queued to have 1 I/O per physical disk

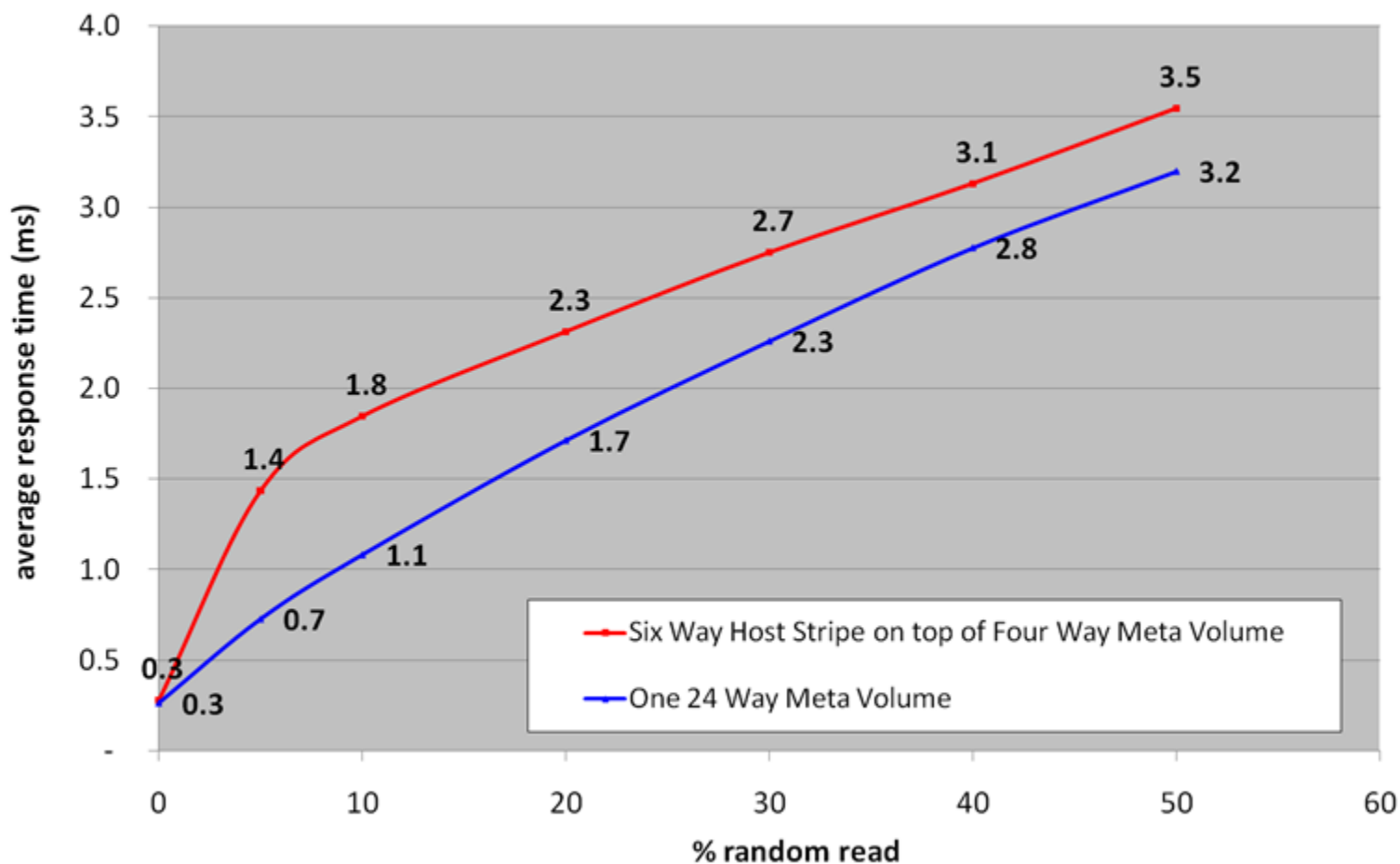
Should I use Meta Volumes or host-based striping? Or both?



- Try to avoid too many levels of striping (plaid)
 - One or two levels is enough, more can make any workload look random, even highly sequential ones
- One large meta volume may outperform several smaller meta volumes that are grouped in a host stripe
- In many cases, host-based striping is preferred over meta volumes
- However meta volumes can reduce complexity at the host level

Striping continued

1 Single 24Way Meta vs 6 Host x 4way Meta
Average Read Response Time (100% read workload)



Striped or Concatenated Metas?



- In most cases, striped meta volumes will give you better performance than concatenated ones
 - Because they reside on more spindles.
 - There maybe some exceptions.
 - If you don't have enough drives for all the meta members to be on separate drives (wrapping) you might want to consider concatenated.
 - If you plan to re-stripe many meta volumes again at the host, you should consider concatenated.
 - If you are making very large RAID 5 or RAID 6 metas, and your workload is largely sequential, concatenated may also be a good choice.
- Concatenated Meta Volumes can be placed on the same RAID group
 - Can create a reasonable emulation of a native large volume on Symmetrix DMX systems or Symmetrix V-Max systems
- Don't place striped meta volumes on the same RAID group (wrapped)
- We recommend concatenated Meta Volumes for Virtual Provisioning
 - (Thin Meta Volumes)

What stripe size should I chose?



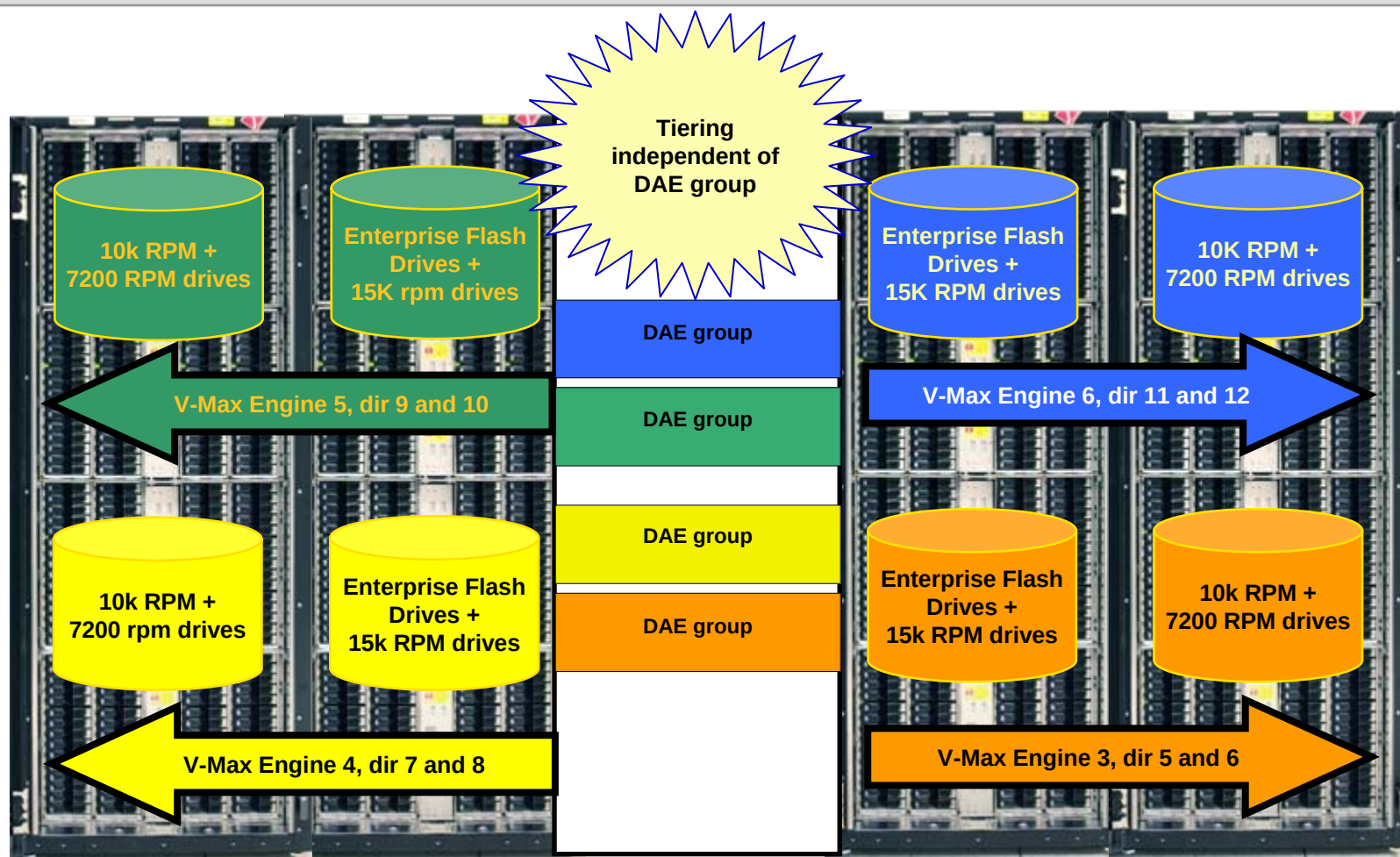
- For most host based striping 128k or 256k is good
 - This is in actual capacity, if stripe is in 'blocks' you need to divide by block size
- May want to consider a smaller stripe size for database logs
 - 64k or maybe smaller as advised by a Symmetrix performance guru

Should I segregate my “X” volumes from my “Y” volumes?

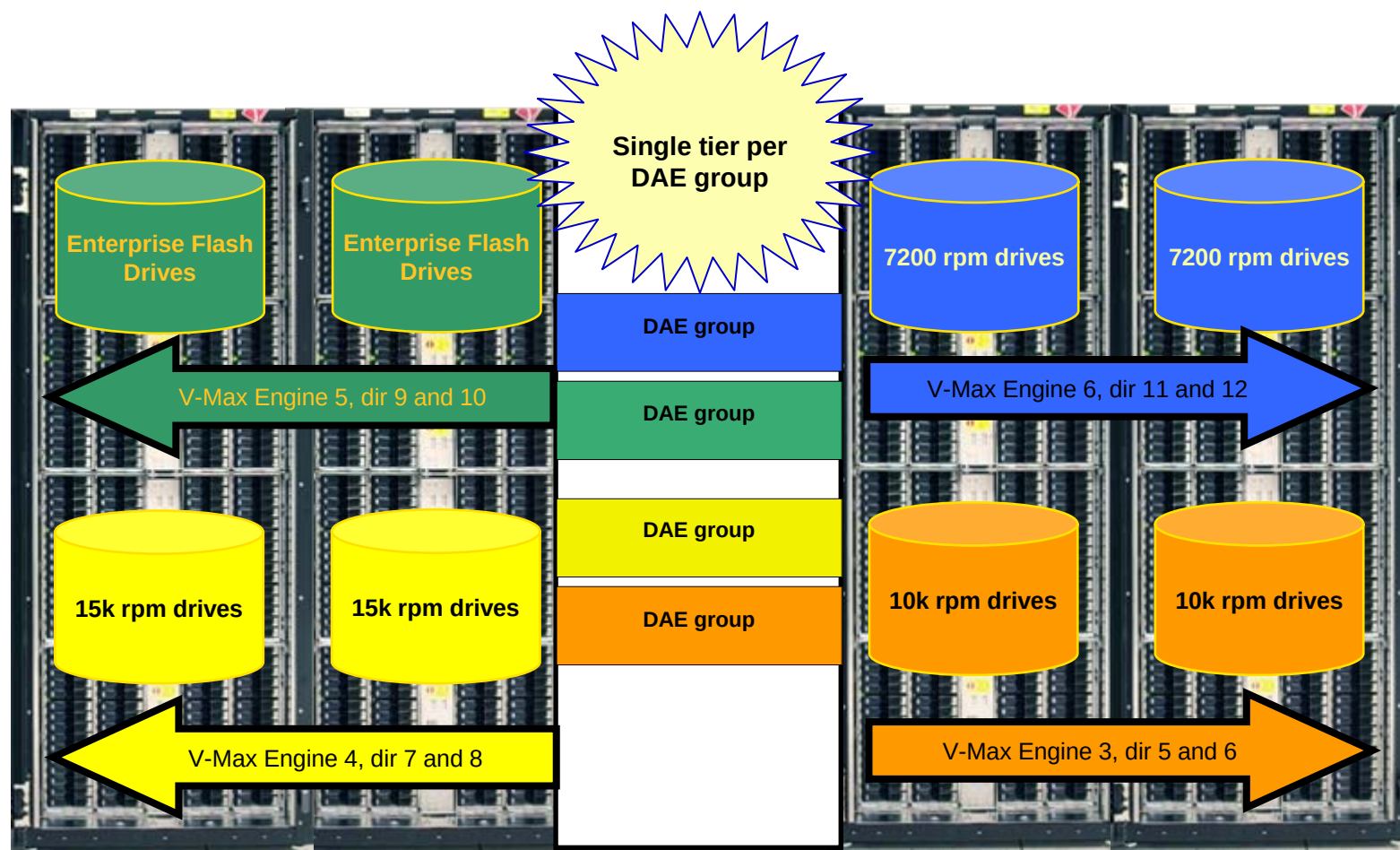


- For the most ***optimal*** system performance, you should not segregate applications/BCVs/Clones onto separate physical disks/DAs or Engines
- For the most ***predictable*** performance, you should segregate applications/BCVs/Clones onto separate physical disks, DAs or Engines
- You should not segregate if the workloads on the various sets of volumes in question are not concurrent with one another, even if you desire the most predictable performance

General Rule Of Thumb: Tiers should share DA resources



Some customers may need segregated tiers



What disk drive class should I choose?

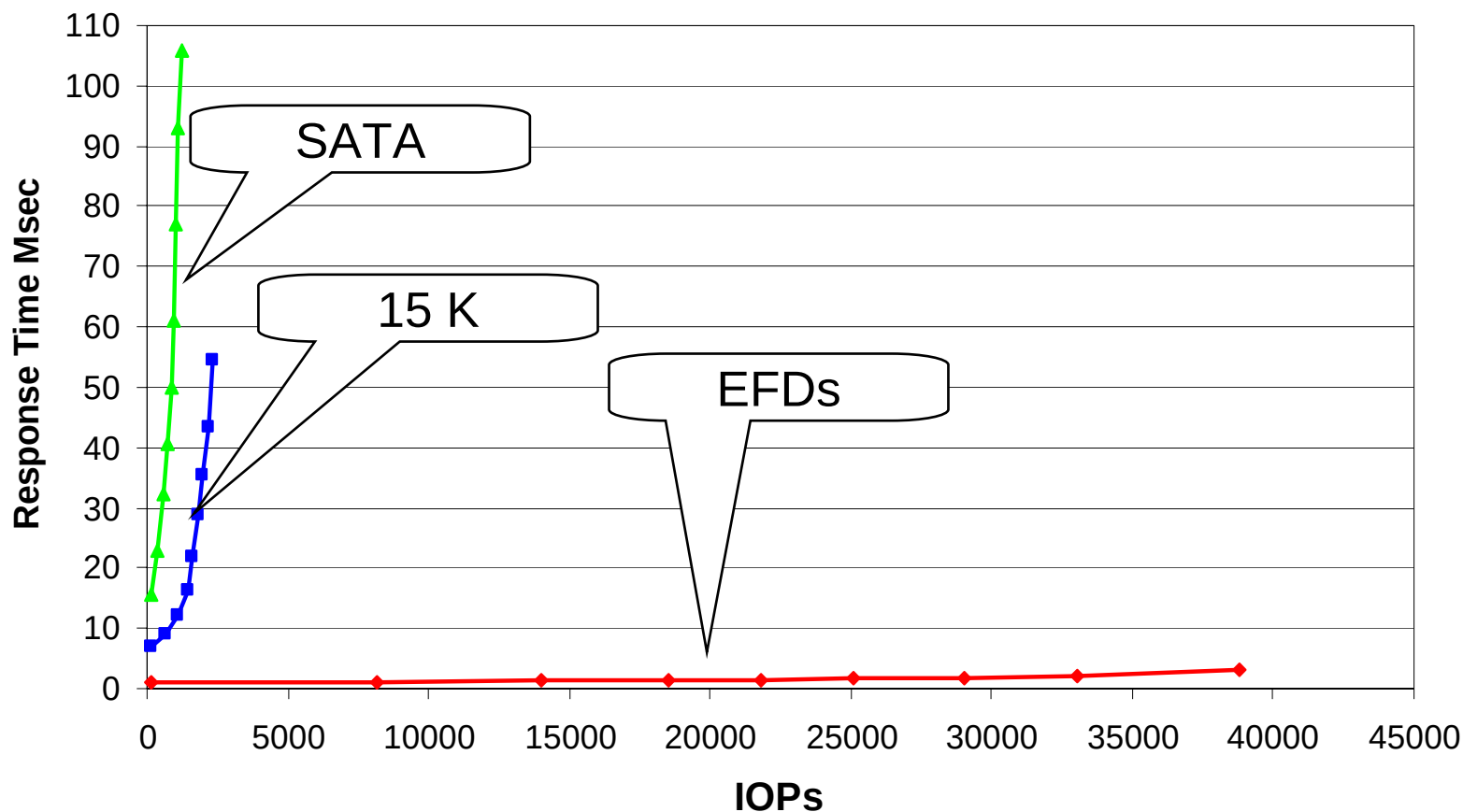


- Enterprise Flash Drives (EFDs) provide the best response time and maximum I/O rate of all the drives
- 15K FC drives provide the best response time for backend reads of all the spinning disks and are about 30% faster for small random reads than FC 10K drives
- 10K FC drives provide slightly longer response times than 15K FC drives
- 7200 RPM drives (SATA) have even higher response times, and are not as good at doing many things at once as the Fiber Channel drives
 - They can be just as good as Fibre drives for large block sequential transfers
 - Just as long as there are not a lot of them running concurrently on the disk

What disk class should I choose?

SATA vs. 15 K FC vs. EFDs

100% Random Read Miss 8KB
One Drive per DA Processor - 8 processors

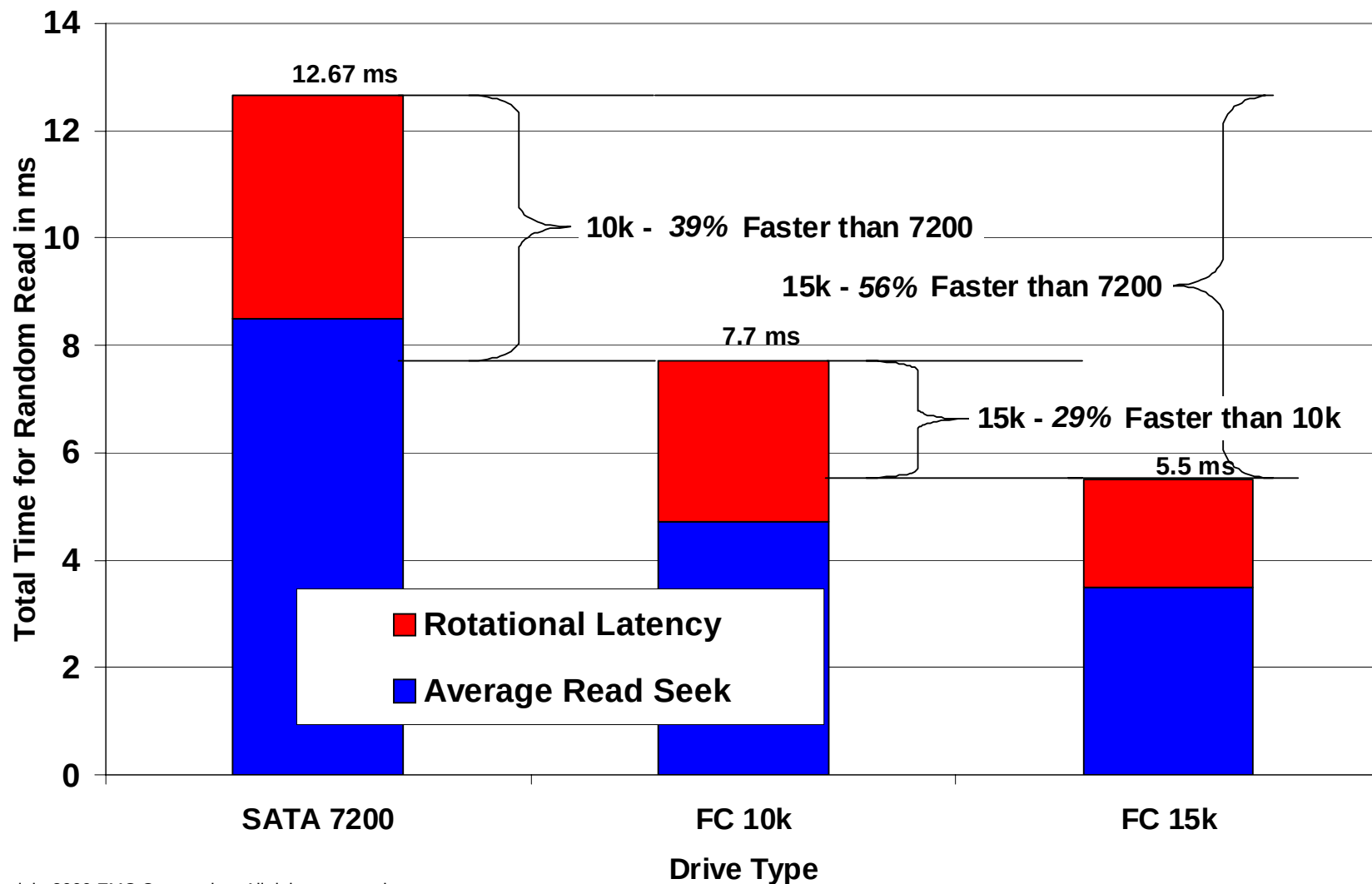


What disk class should I choose?

Random Read – spinning disks



Random Read Time

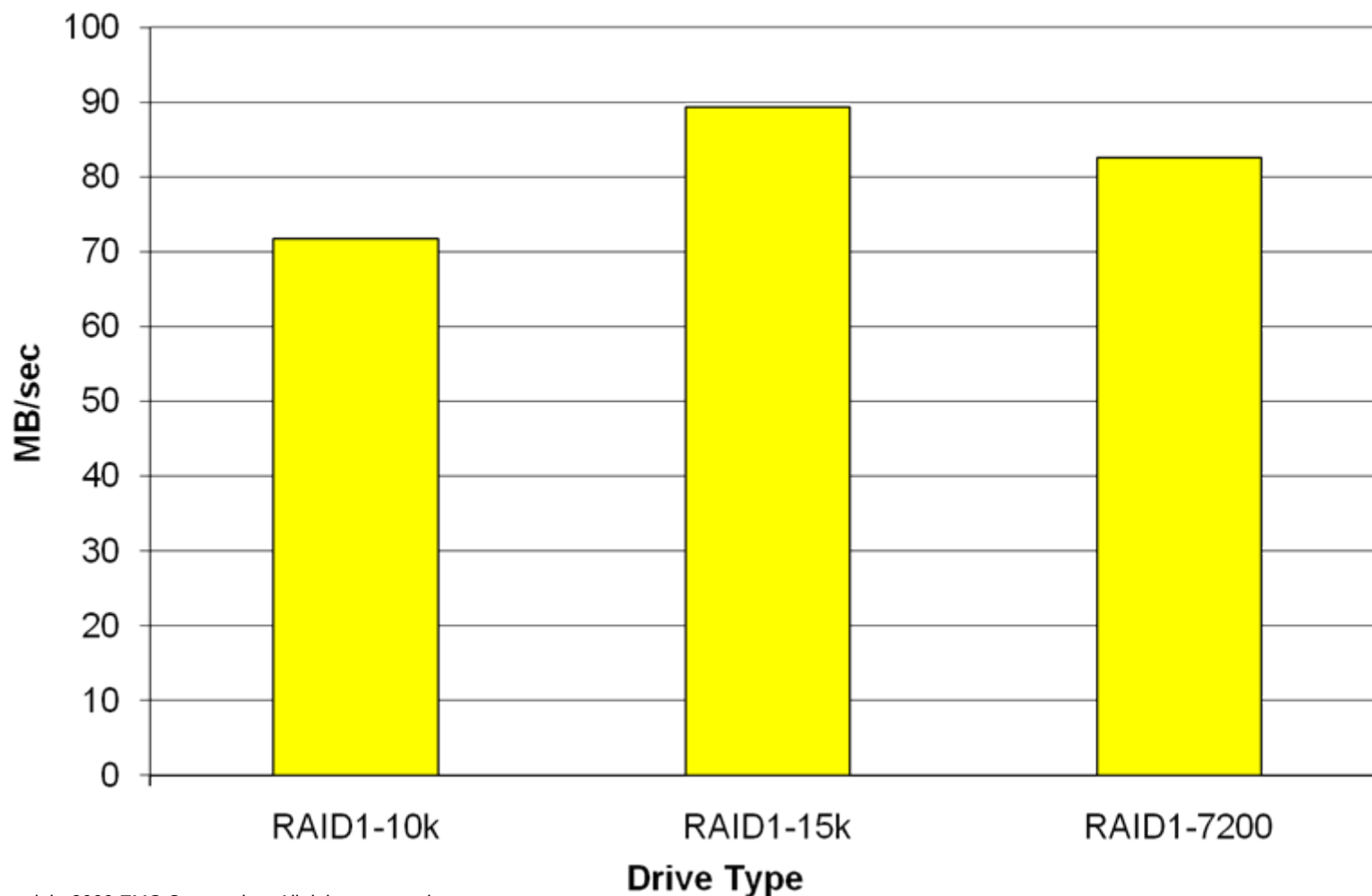


What disk class should I choose?

Sequential Reads – spinning disks



1 Disk - Single Thread 128k Sequential Read



What RAID protection should I choose?

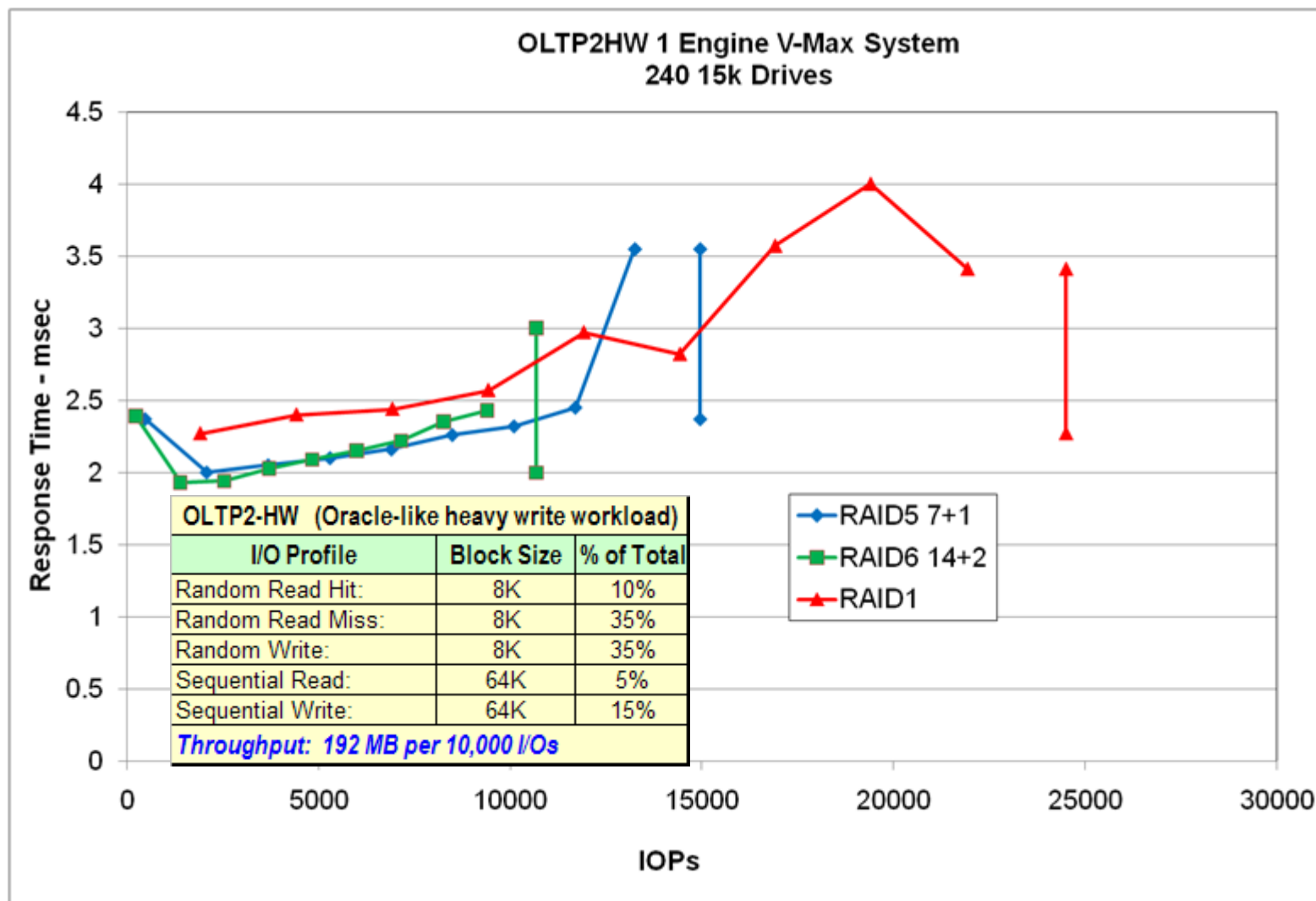


- Performance of reads similar across all protection types
 - If you are comparing the same number of drives!
- Major difference is with random write performance
 - Mirrored: 1 Host Write = 2 Writes
 - RAID 5: 1 Host Write = 2 Reads + 2 Writes
 - RAID 6: 1 Host Write = 3 Reads + 3 Writes
- Cost is also a factor
 - RAID 5/6 are best at 12.5% or 25% protection overhead
 - RAID 1 has a 50% protection overhead

Protection	Random Write Performance	Resiliency
Mirrored	Best	Better
RAID 5	Better	Good
RAID 6	Good	Best

What RAID protection should I choose?

Spindle comparison

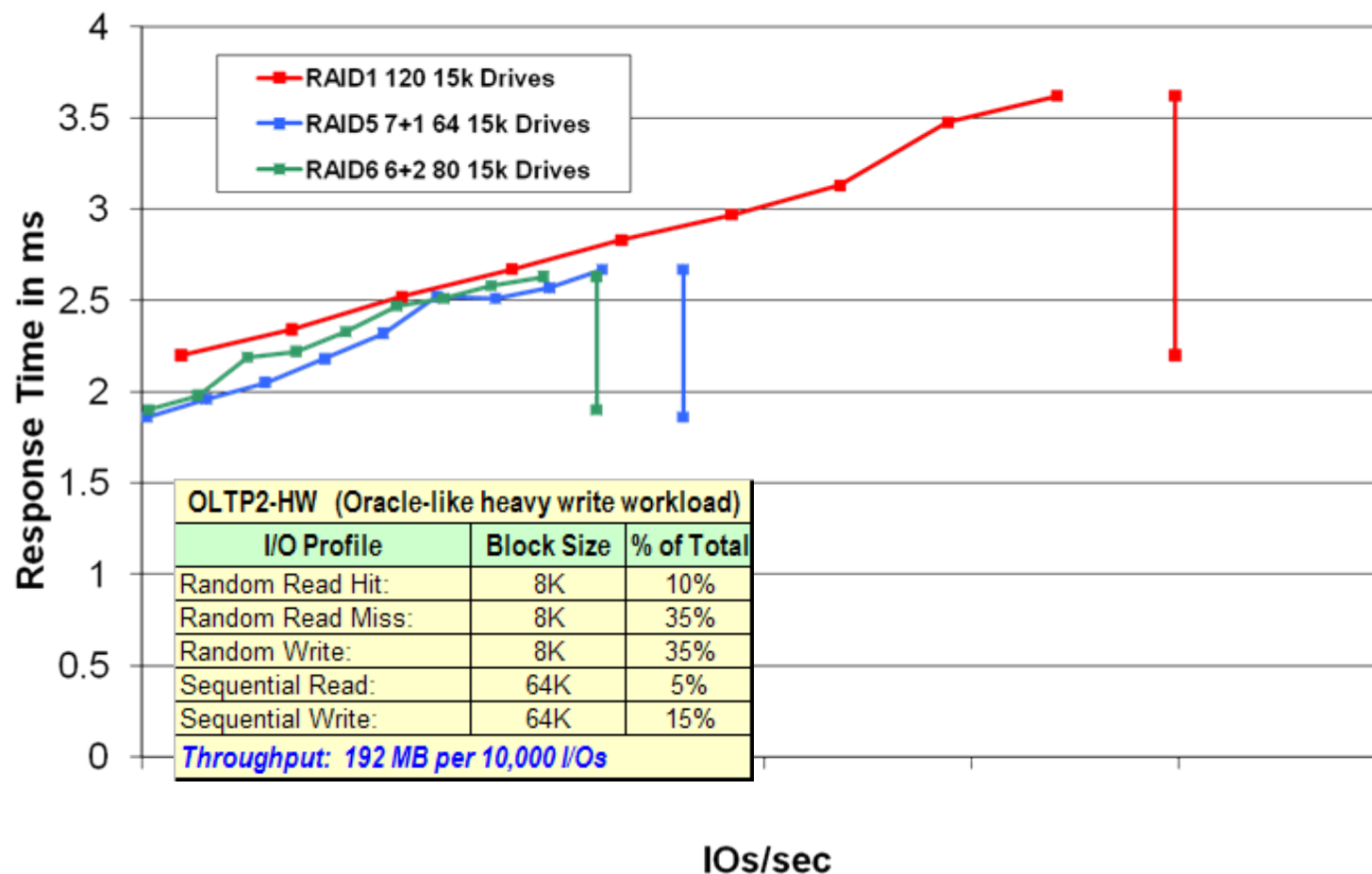


What RAID protection should I choose?

Capacity comparison



OLTP2HW
Symmetrix V-Max 1 Engine 5874.120 28GB Cache



How much cache do I need?



- There are minimum cache requirements to support the capacity
- These also include a “performance” requirement
 - To support some minimum fall through time at the limits
- Easiest method is to utilize the Dynamic Cache Partitioning What If (DCPwi) tool
 - Put like devices together in cache partitions
 - Start analysis mode and collect DCP stats
 - Use DCPwi tool to see the effect of cache on overall hit rates
 - DCPwi tool is a Symmetrix SPEED Guru only tool
- Only other method is to collect very L-O-N-G front end CDB traces
 - These can be used for cache analysis
 - This analysis will be quite accurate for the given period, given a long enough trace
 - This process can be very time consuming
 - Is only accurate for the trace collected

How do I know when I'm getting close to limits?



- Watch for growth trends in your workload
 - SPA was designed specifically for this purpose
- Look out for increasing response time
 - Host based tools (iostat, sar, RMF)
- Monitor utilization metrics in WLA/STP
 - Some metrics may not be perfect, but watching trends is still valuable
- Better to be proactive than waiting to hit the “wall”
- Any utilizations well over 50% should be considered a possible source of future issues with growth

Three most important points for optimal Symmetrix system performance

- Configure Enough Resources
- Utilize Resources Evenly
- Simple Data Layouts Are Better

Questions???

Related Sessions at EMC World



- Achieving the Best Performance from Virtual Provisioning on Symmetrix
- Architecture and Performance of the Symmetrix DMX Series and Symmetrix V-Max Series with Enginuity
- Aspects of Performance in a Symmetrix TimeFinder Environment
- Aspects of Remote Replication Performance in Symmetrix DMX and Symmetrix V-Max Series with Enginuity
- Capacity and Performance Planning for Symmetrix (Co-presented with Microsoft adCenter SAN Engineering)
- Four Site Migration: A Customer Success Story
- SRDF/A Network Performance Analysis and Troubleshooting
- Symmetrix Performance and Capacity Planning with Symmetrix Performance Analyzer (SPA)
- Symmetrix Tiered Storage Best Practices and The Power of FAST

EMC²

where information lives[®]