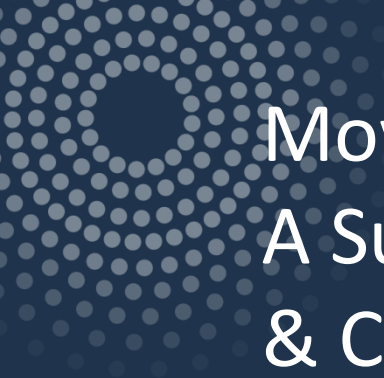# Presenter

Tony Brown, Production Security & Engineering R&D
SAS Institute Inc.

Tony Brown works in the SAS PS&E Performance Lab. His duties include researching, studying and testing hardware and software interaction. The goal of this work is to ensure maximum performance for SAS Software and Solutions, and assist Customers with performance infrastructure.

Moving from SAS 9 to Viya
A Success Guide to Viya Infrastructure & Configuration

Tony Brown

SAS Performance Lab – Production Security & Engineering R&D

SAS

THE POWER TO KNOW®

Table of Contents

- SAS 9 Vs. Viya – Logical and Physical Differences

- Viya General Infrastructure– Components Servers and Services

- Most Common Performance Challenges

- Infrastructure Performance Best Practices
  - Recommended Data Flow To/From Storage
  - Recommended Data Flow To CAS_Disk_Cache
  - Recommended Operations Resource Levels
  - Job Resource Requirements

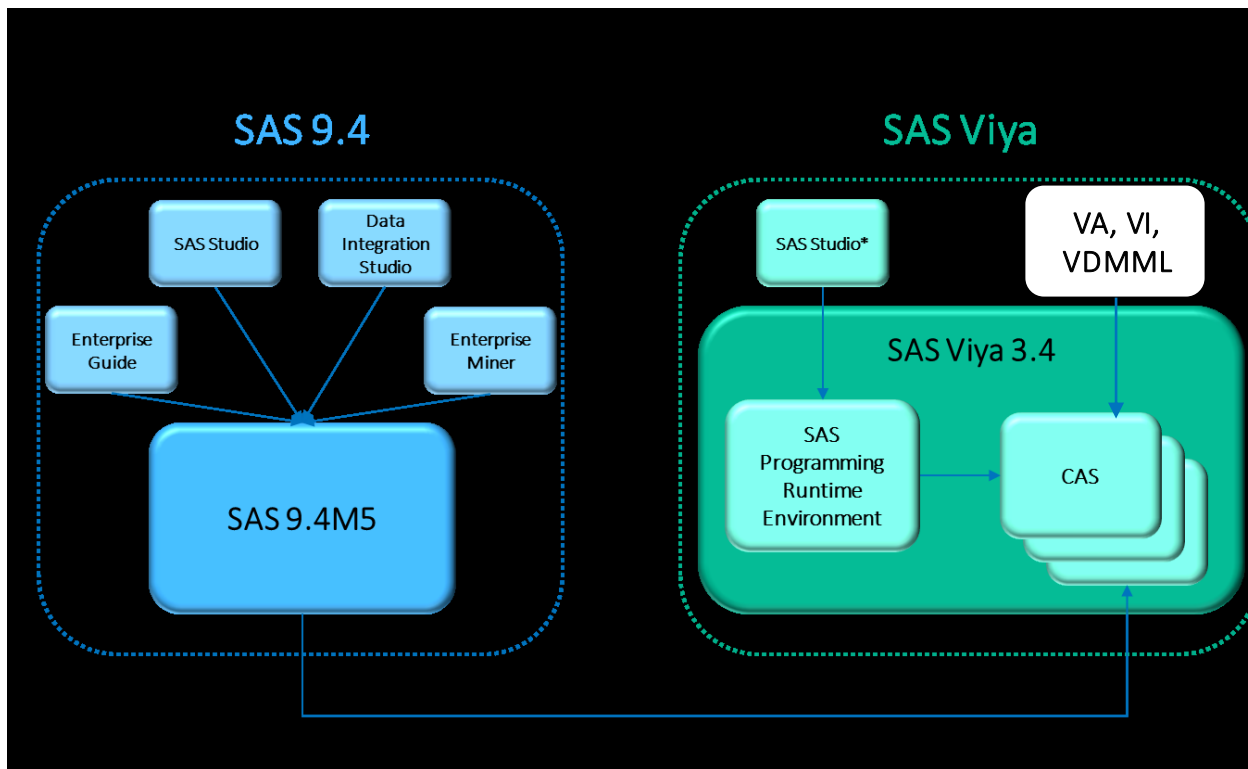- Virtual and Cloud Infrastructure Translation

SAS 9 to Viya

The Elementary Version of What is Different
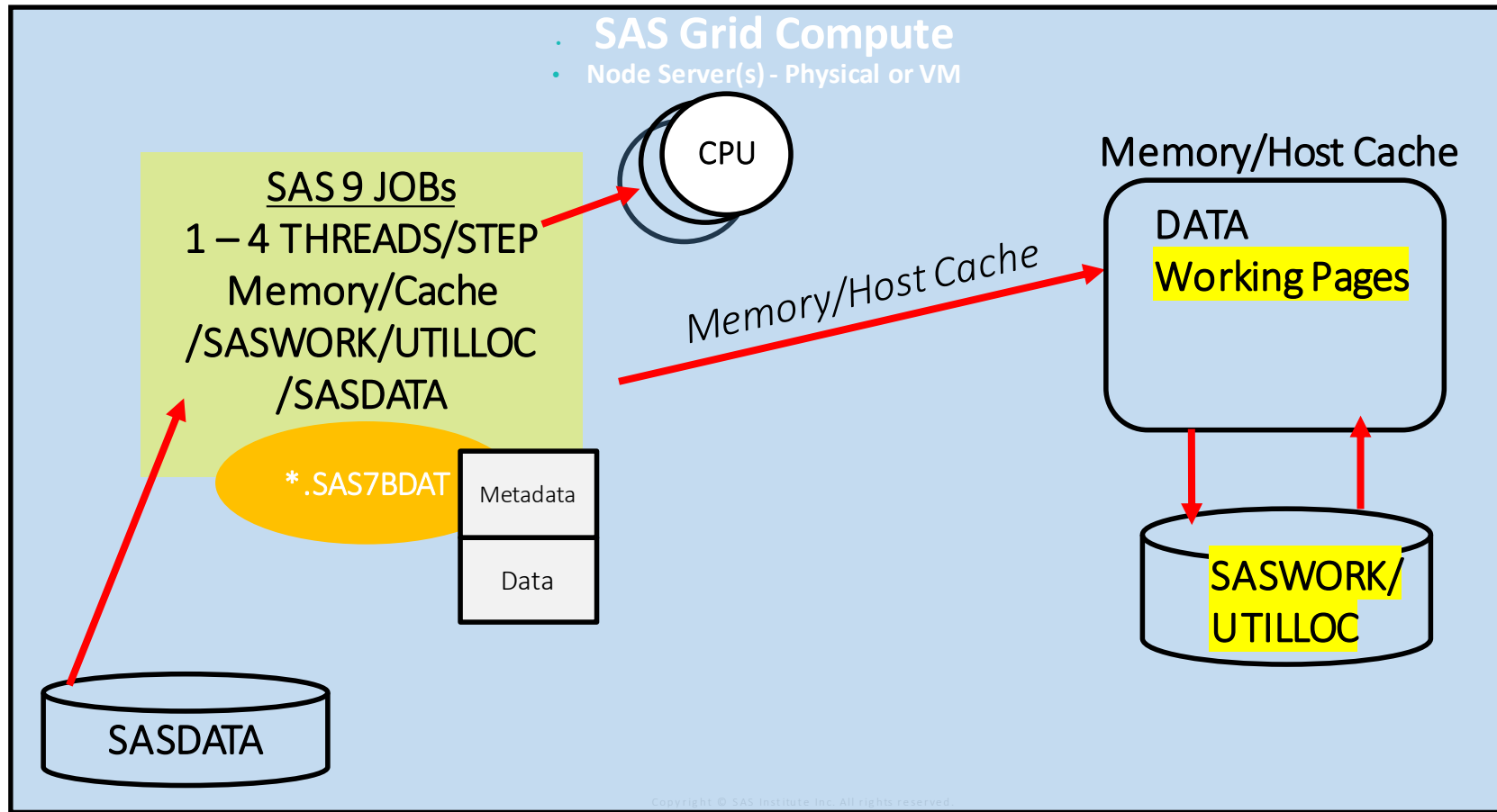
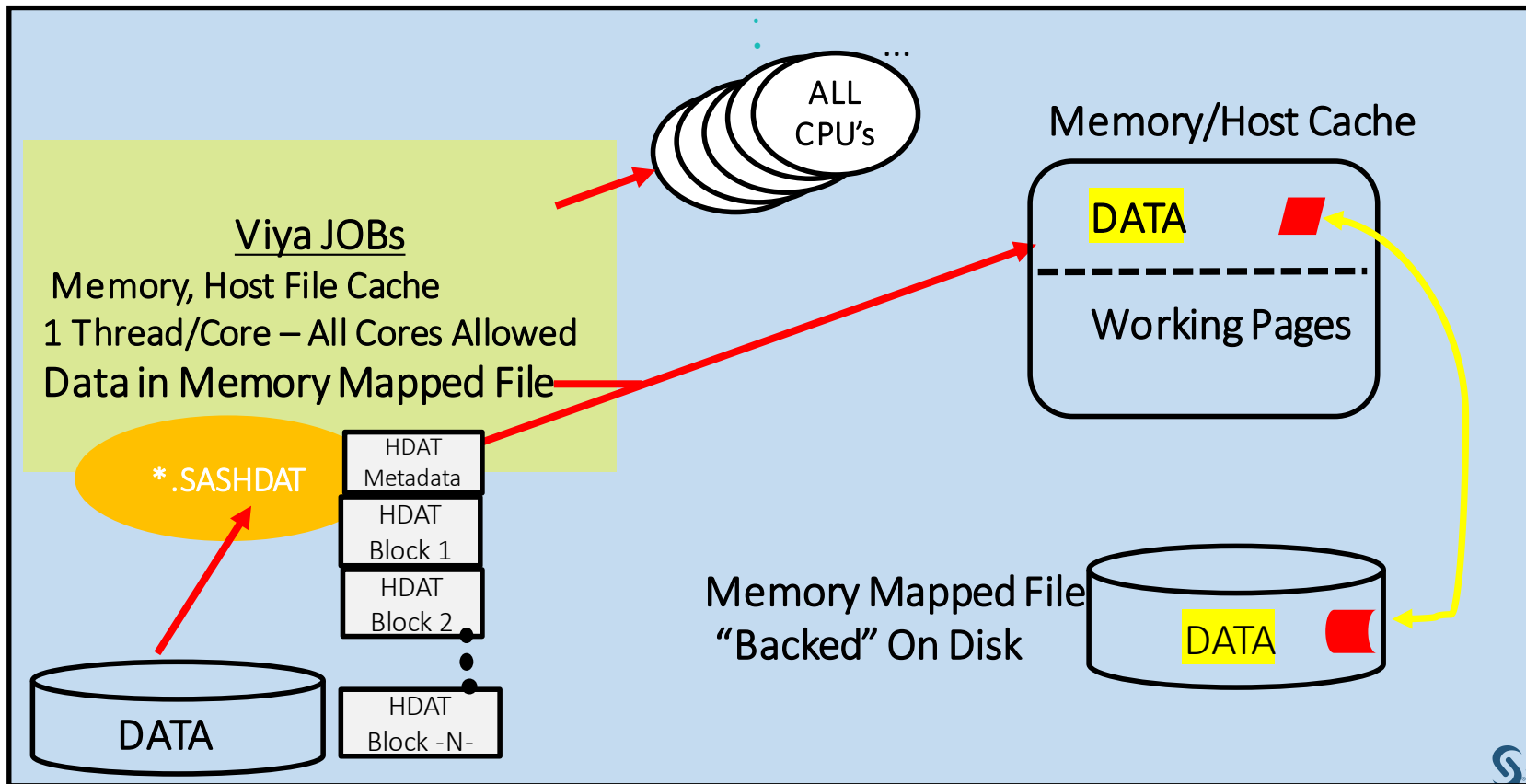# Overview
SAS 9 vs. Viya Quick Look

*Most SAS Viya clients (e.g. SAS Visual Data Mining and Machine Learning) can call CAS directly.

SAS 9 Flow

SAS Grid Compute
- Node Server(s) - Physical or VM

CPU

Memory/Host Cache

SAS 9 JOBs
1 – 4 THREADS/STEP
Memory/Cache
/SASWORK/UTILLOC
/SASDATA

*.SAS7BDAT

Metadata

Data

Memory/Host Cache

DATA
Working Pages

SASWORK/
UTILLOC

SASDATA

Viya CAS Flow

ALL CPU's

Memory/Host Cache

DATA

Working Pages

Viya JOBs
Memory, Host File Cache
1 Thread/Core – All Cores Allowed
Data in Memory Mapped File

*.SASHDAT

HDAT Metadata

HDAT Block 1

HDAT Block 2

HDAT Block -N-

DATA

Memory Mapped File "Backed" On Disk

DATA

Serial SAS9 vs. Parallel Viya on CAS

Single Host

SERIAL

SAS 9

| DATA | DATA | SORT | DATA | SORT | DATA | REG |

SAS JOB

CPU Threads

SINGLE NODE

Parallel

VIYA/CAS

DATA — SAS JOB — SMP Host

DATA — SAS JOB — MPP Host

DATA — SAS JOB — MPP Host

DATA — SAS JOB — MPP Host

DATA — SAS JOB — MPP Host

CPU Threads

1 To MANY NODES
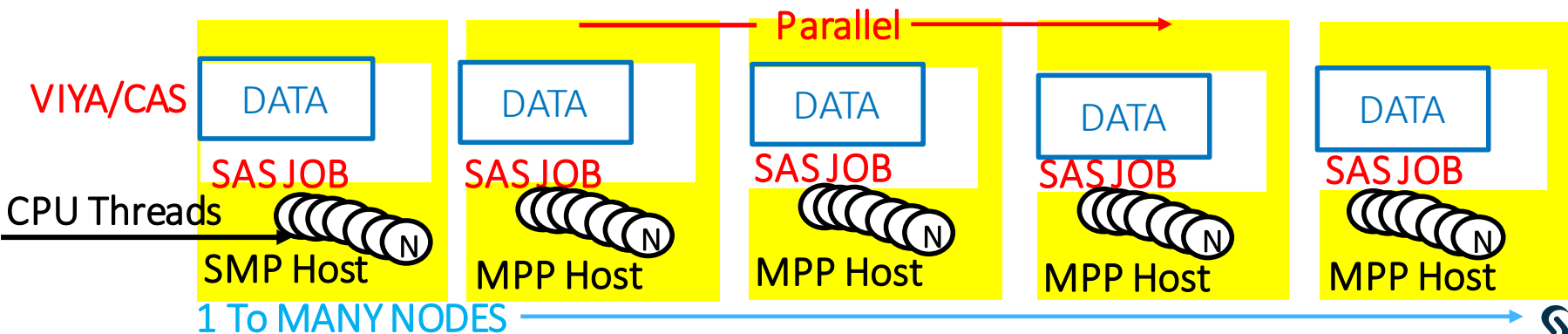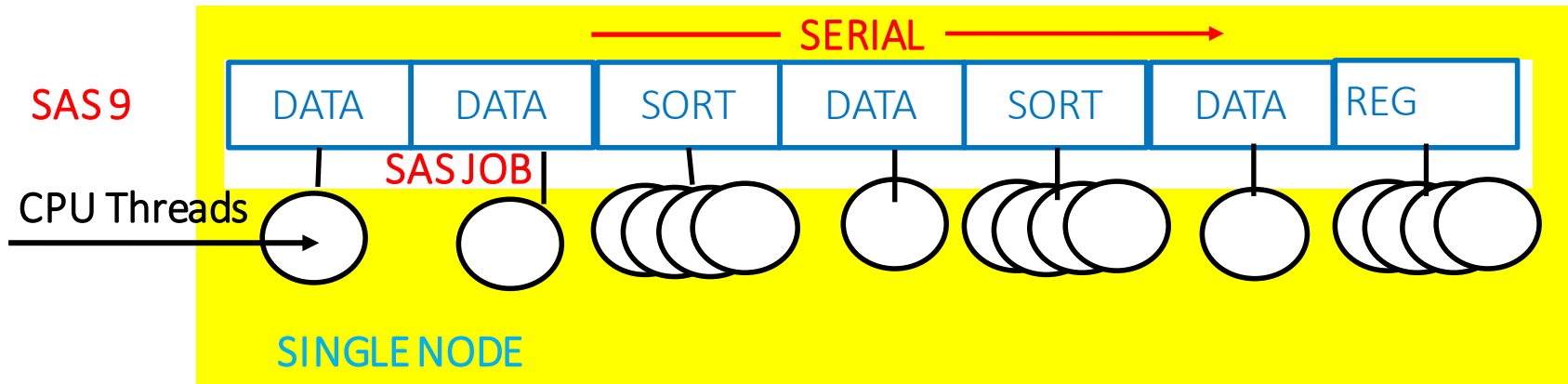
Loading Data into Viya – Serial vs. Parallel DNFS

CAS Controller

CAS Worker

CAS Worker

CAS Worker

.sashdat

DNFS Parallel Load

Can be very high instantaneous demand load against storage

Serial Load

Storage Solution
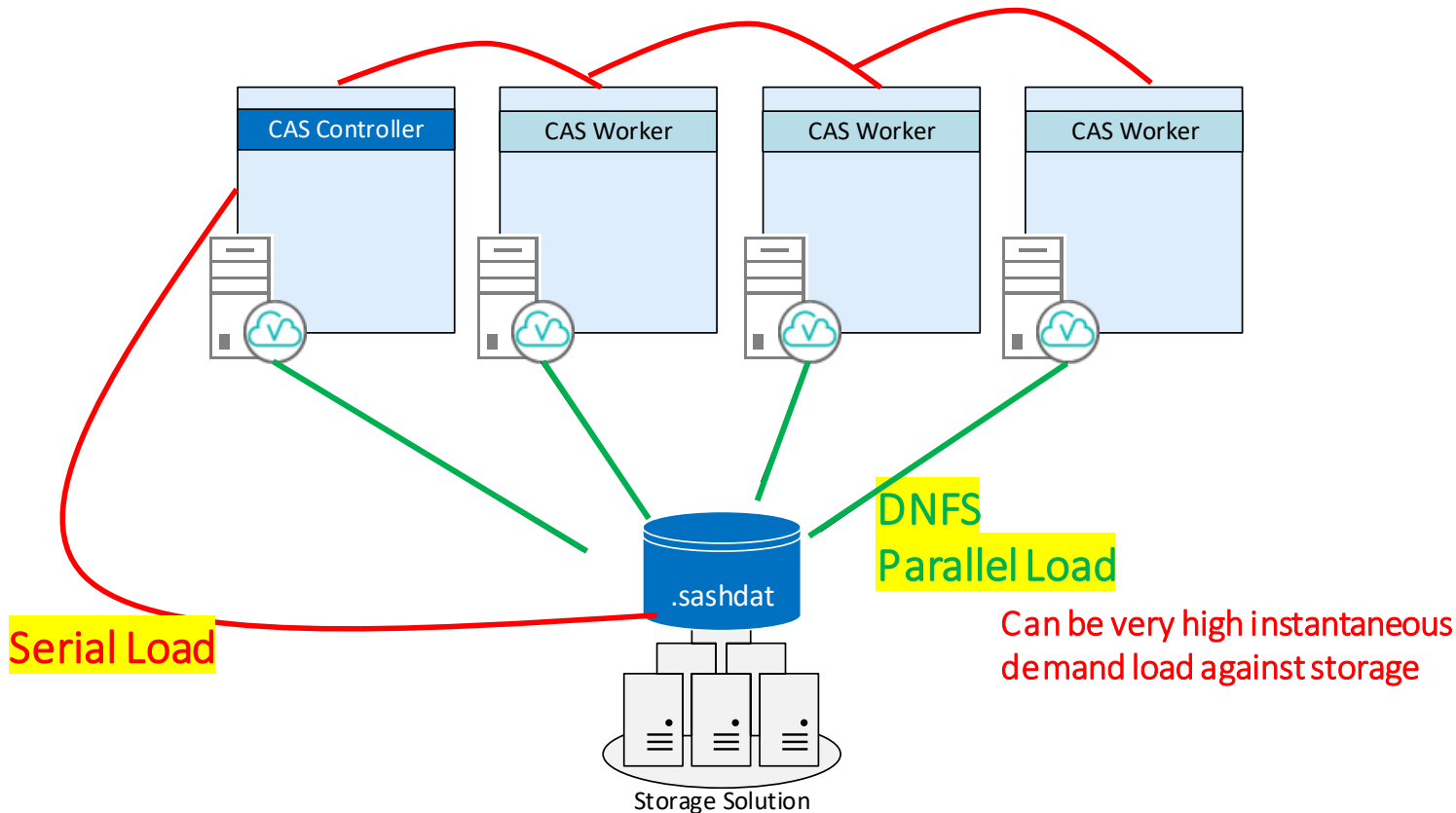
Infrastructure – Servers and Services

Viya General Infrastructure
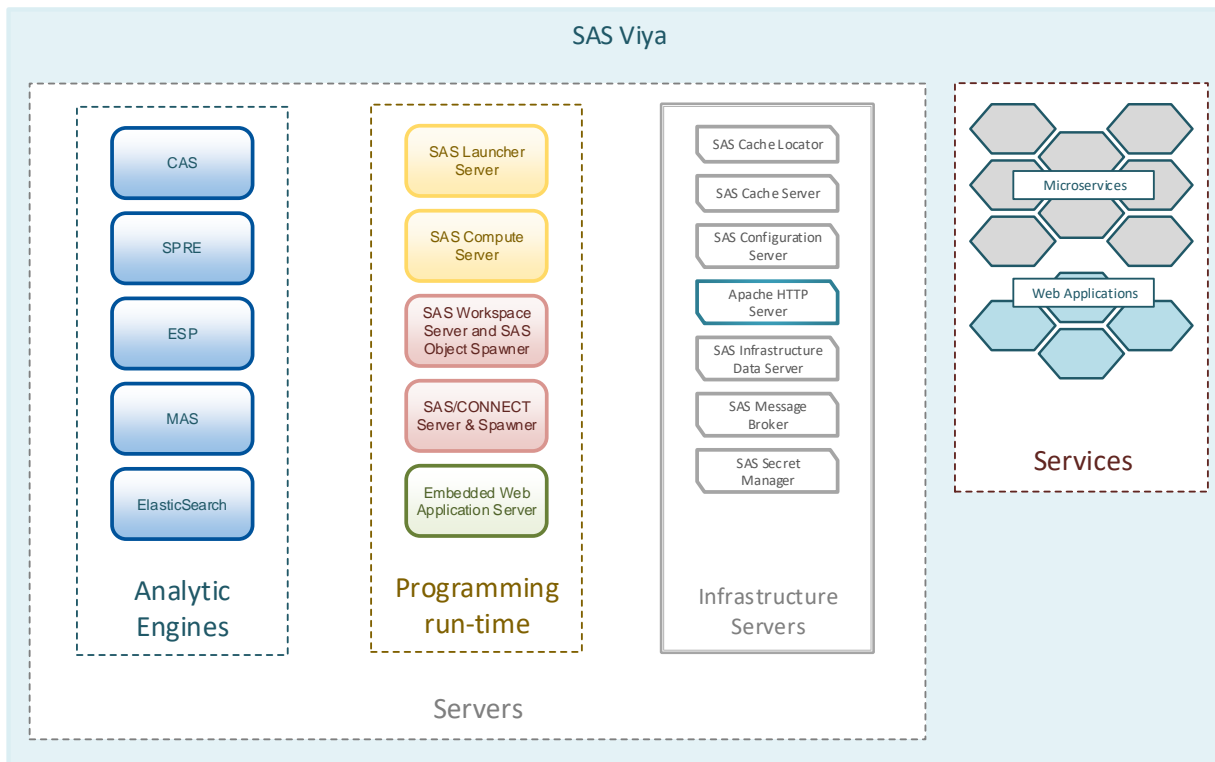
SAS Viya

Analytic Engines
- CAS
- SPRE
- ESP
- MAS
- ElasticSearch

Programming run-time
- SAS Launcher Server
- SAS Compute Server
- SAS Workspace Server and SAS Object Spawner
- SAS/CONNECT Server & Spawner
- Embedded Web Application Server

Infrastructure Servers
- SAS Cache Locator
- SAS Cache Server
- SAS Configuration Server
- Apache HTTP Server
- SAS Infrastructure Data Server
- SAS Message Broker
- SAS Secret Manager

Servers

Services
- Microservices
- Web Applications

Viya General Infrastructure
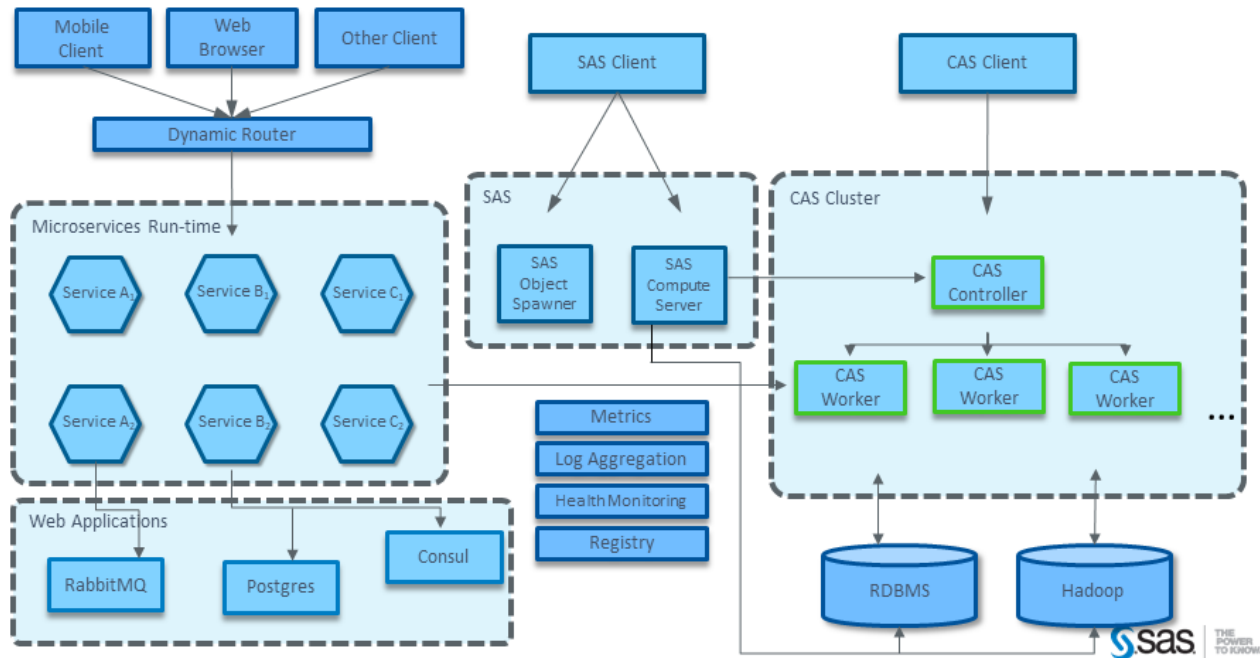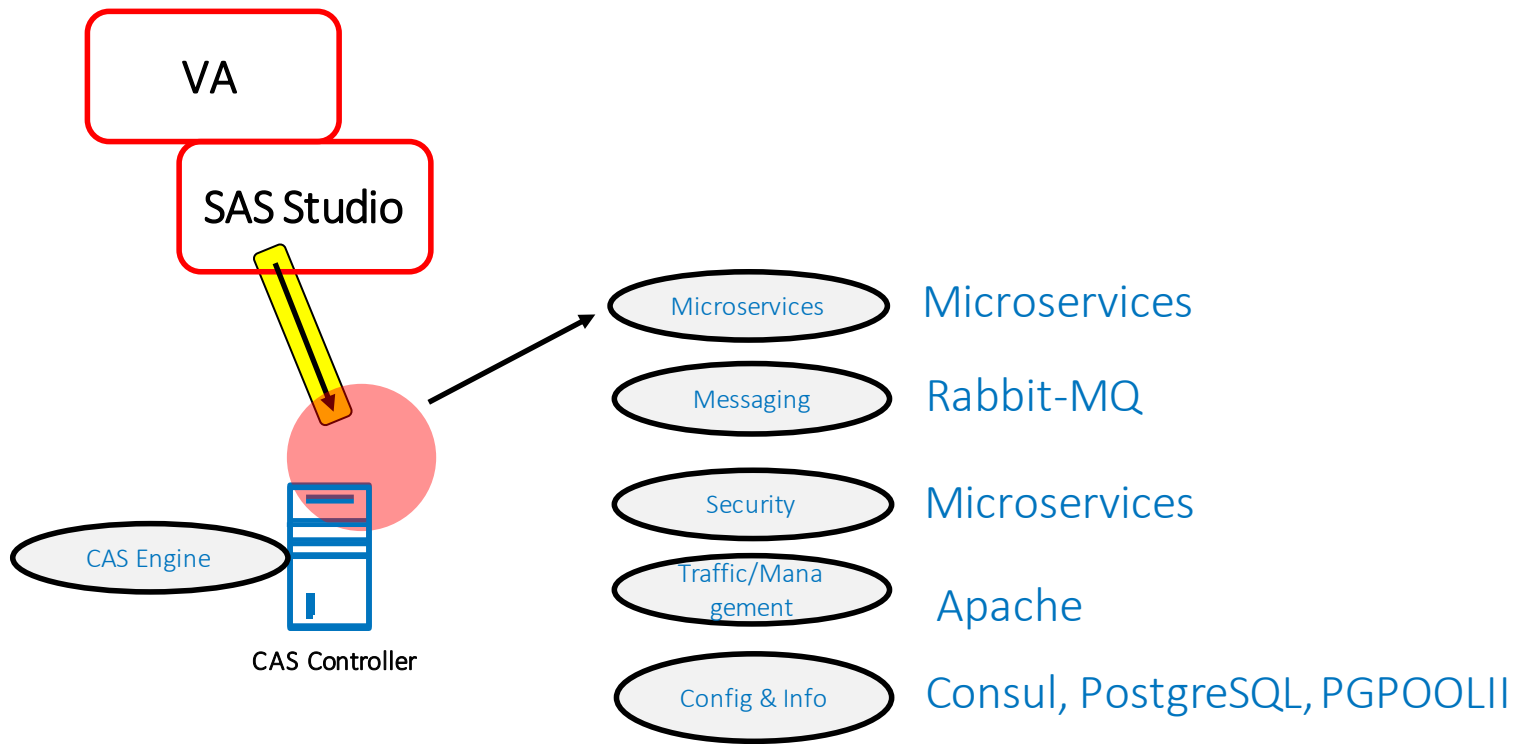
Performance Challenges

Most Common Performance Challenges

- Slow Performance Largely Due To:
  - Combining Server Roles on a single host (Microservices/Controller/SPRE)
  - Inappropriate hardware architecture purchase and deploy vs. recommended
    - CPU, Memory, Storage, OS, Tuning, Network, Virtualization Design
  - Host hardware and network IO bandwidth/latency insufficiency
  - Microservices resource insufficiency, request queueing on Improperly Tuned NIC cards, etc.
  - Data Management Issues:
    - Bandwidth Location of Source Data
    - Encoding Conversions/CEDA Ignored
    - Shared File Systems Behavior
    - Poor Coding Practices Converting from SAS 9 to CAS
    - Poor design of data handling for Model Building
  - Crossing Data Centers or Firewalls (Availability Zones in Cloud)

Infrastructure Performance Best Practices

§sas

A CAS Job Relies on a Lot of Software Infrastructure…

VA

SAS Studio

CAS Engine

CAS Controller

Microservices → Microservices

Messaging → Rabbit-MQ

Security → Microservices

Traffic/Management → Apache

Config & Info → Consul, PostgreSQL, PGPOOLII

Viya Infrastructure Environment
Servers, Software, Networks, Storage

NIC

NETWORK & SWITCHES

NIC

NIC

SERVERS

FC

DATA

CPU
Memory
IO (Data in/out, Cache, CAS_Disk_Cache)
Communications

SOFTWARE

Microservices
Messaging
Config & Mgmt
Security
Traffic
Logging/Health
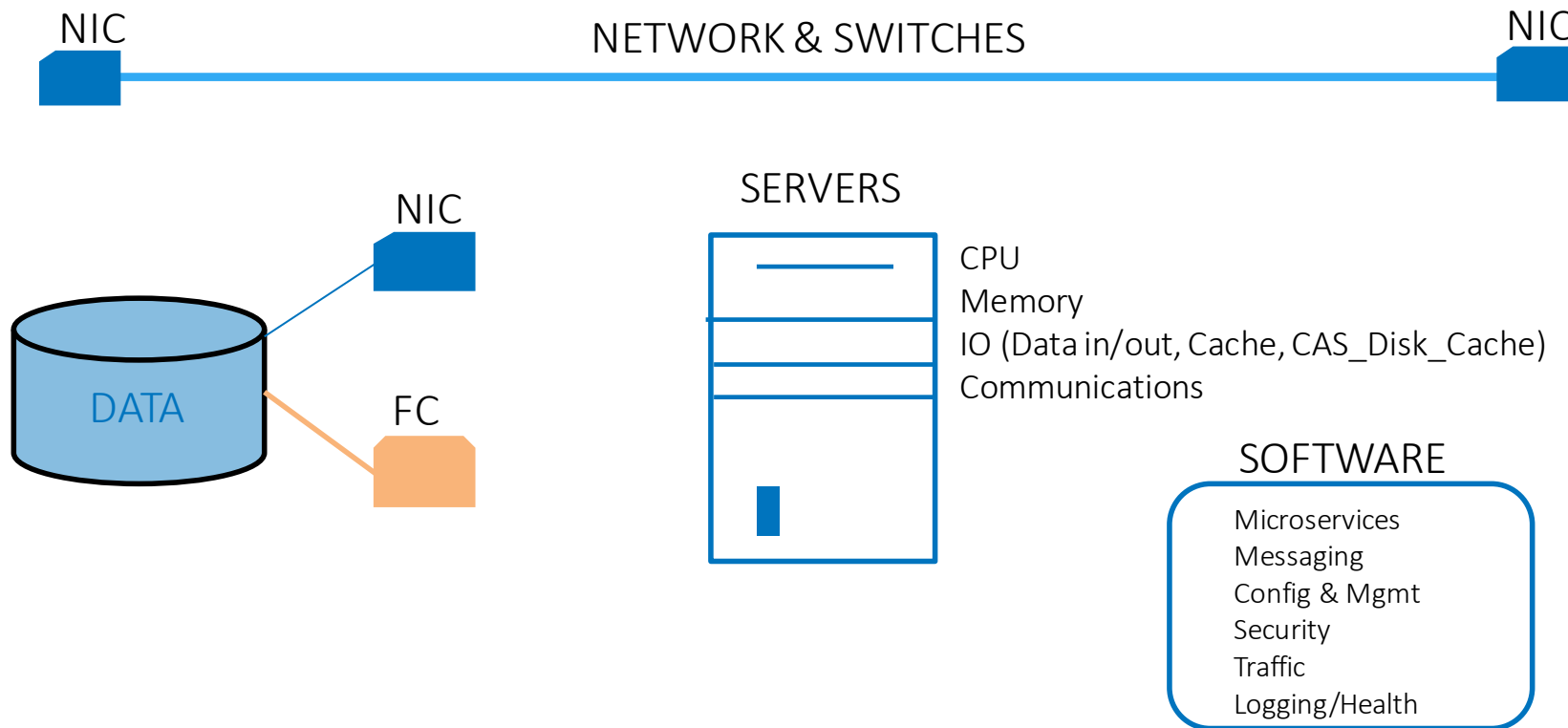
Viya Infrastructure Environment
Recommended DATA Flow From/To Storage
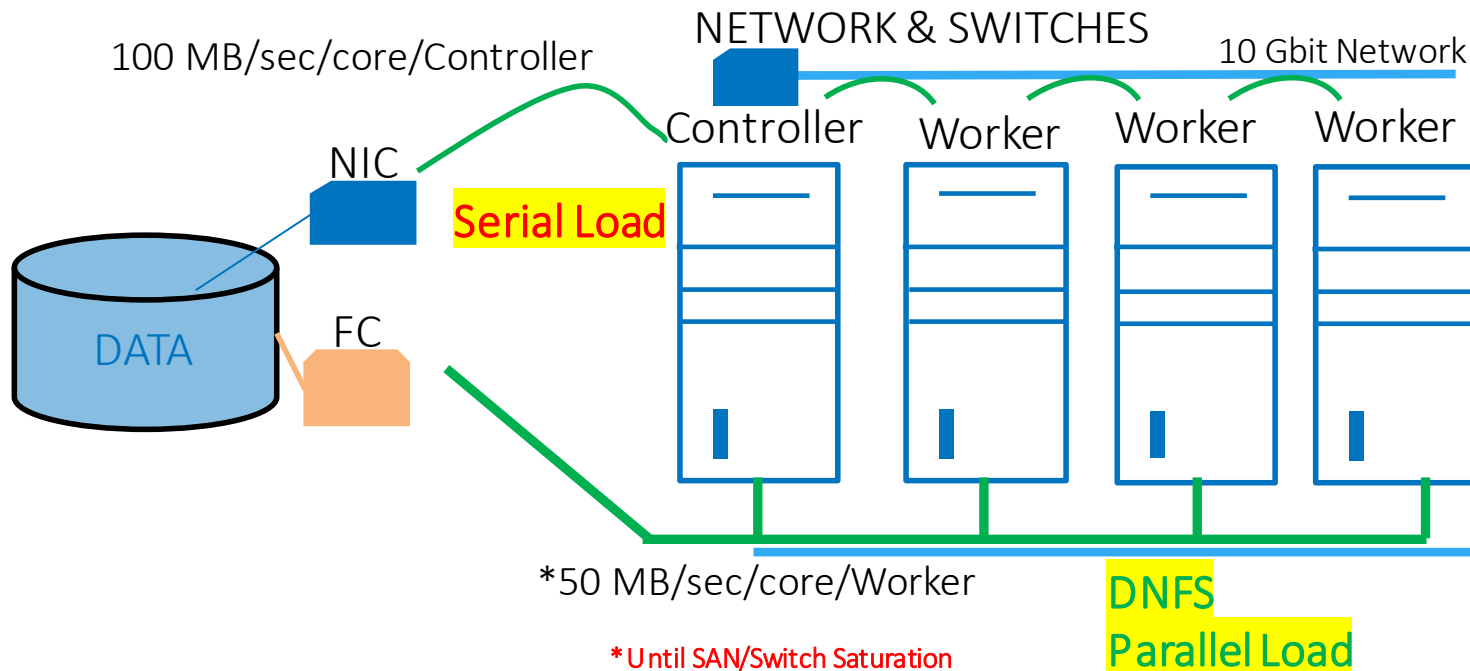
NETWORK & SWITCHES

100 MB/sec/core/Controller

10 Gbit Network

Controller   Worker   Worker   Worker

NIC

Serial Load

DATA

FC

*50 MB/sec/core/Worker

DNFS
Parallel Load

*Until SAN/Switch Saturation

Viya Infrastructure Environment

Recommended DATA Flow to CAS_Disk_Cache Backing Store

Memory Mapped File

CAS_Disk_Cache – High Bandwidth/Low Latency

100 MB/sec /core/Worker

On-Host Storage Devices

Striped LV – SSD/NVME is Best
High Bandwidth, Low Latency on
Random Page Being Returned.....

2 – 3x RAM
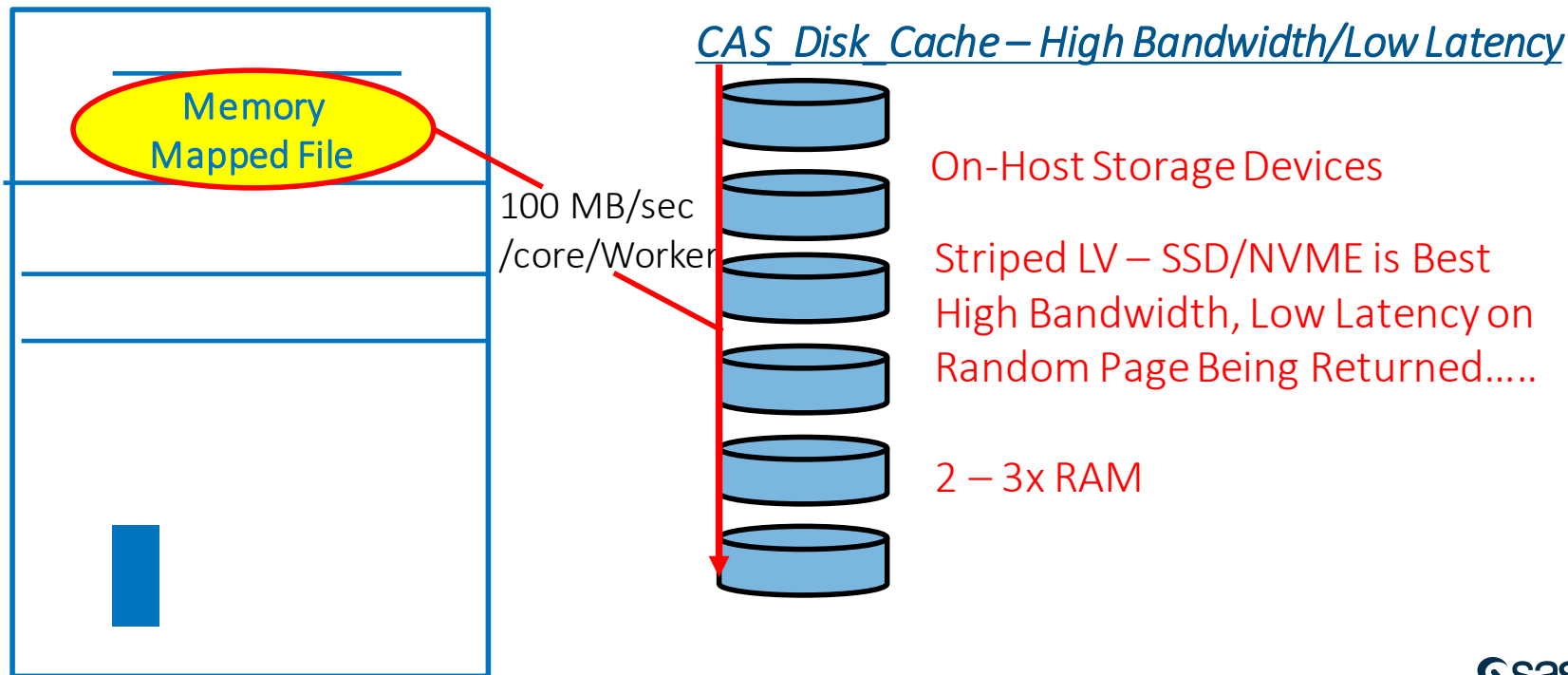
Viya Infrastructure Environment
Recommended Operational Resources

NIC NETWORK LATENCY & Stability!!! **10 Gbit MINIMUM** **NIC**

CAS Nodes MicroServices Rabbit-MQ Consul Postgres

*VERY LOW LATENCY <1 ms

VERY LOW LATENCY low ms

PGPOOL
EV
Security
Logging/Health

Do Not Share Host Resources for These……

- Private Network Segments
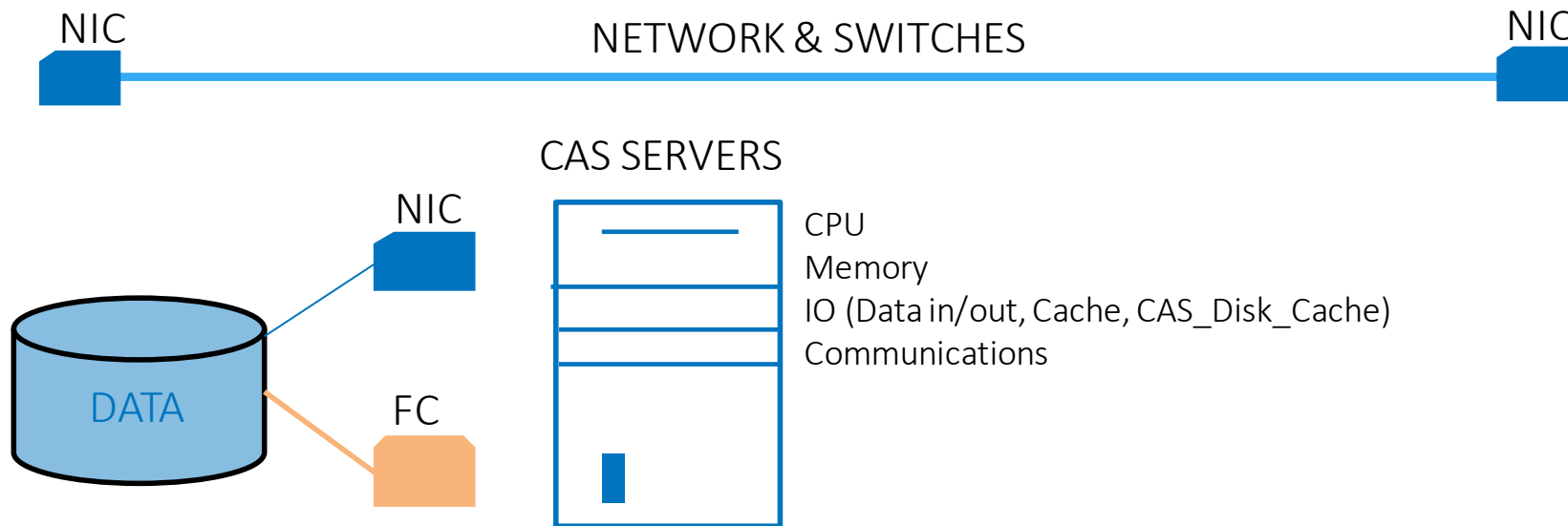- No Hypervisor Level Disruptions
- No Noisy Neighbors
- Low Latency Path to *logging
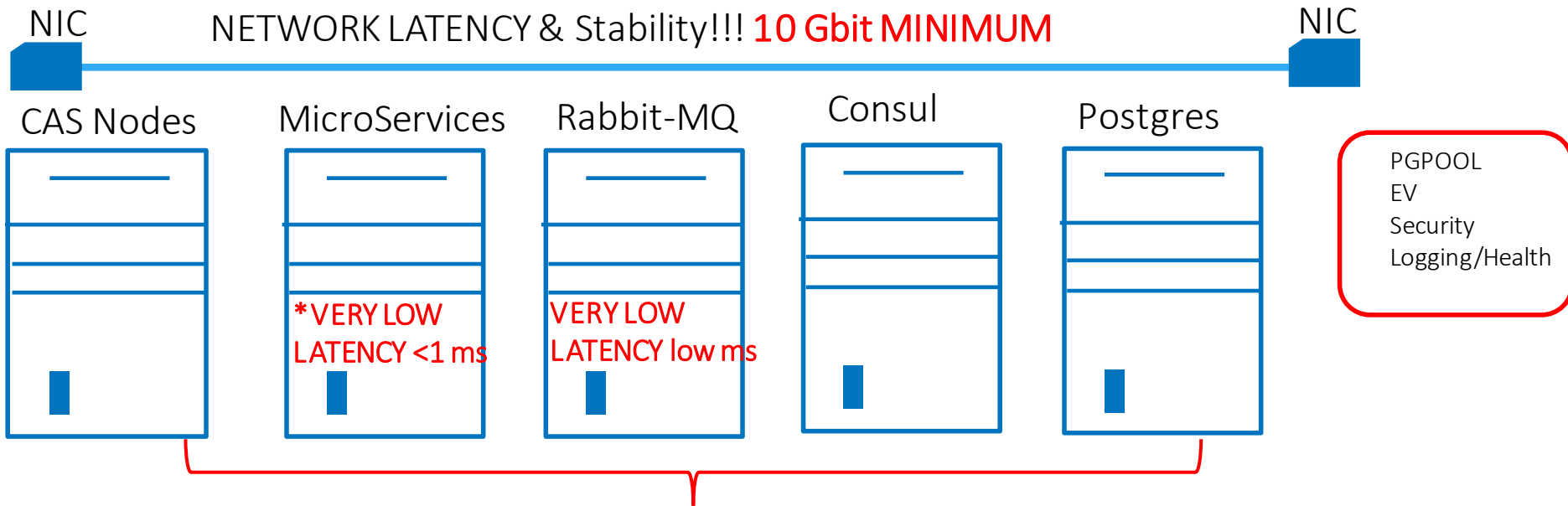
*/opt/sas/viya/config/varlog/~) perform best under 500 microsecond response times for microservice calls/functions

Viya Infrastructure Environment
Job Resources

NIC   NETWORK & SWITCHES   NIC

CAS SERVERS

NIC

FC

DATA

CPU
Memory
IO (Data in/out, Cache, CAS_Disk_Cache)
Communications

Job Load – Its all in Memory!
Follow Your Sizing Guidelines!!
Use Your Cgroups!!

Additional Performance Best Practices

Host Resources
Viya Server Host Resource Provisioning

- OS

  - RHEL 7.7 or Later. Preferred – RHEL 7 Latest Version and Errata. RHEL 8 is Not Vetted for Viya 3.x Yet.

  - RHEL on Power - Only supported on RHEL for IBM Power LE (POWER9) 7.6 (sometimes referred to as RHEL 7.6alt). SAS Viya cannot be installed on RHEL 7.6.3.10 classic as it does not fully support the POWER9 instruction set.

  - In General, follow the SAS on Red Hat 6 & 7 paper mentioned in Host Tuning on a subsequent slide, an exception will be the following:

    - vm.min_free_kbytes = ~1GB

    - vm.dirty_background_ratio = 3

    - vm.dirty_ratio = 10

    - RHEL turn huge transparent page off

    - RHEL turn memory defragmentation off

    - Multi-NIC cards best practices – point to: Slow NIC – DNS, Fast NIC – internal communications

    - Viya Ulimits Tuning - On Red Hat Linux deployments, there are recommended beginning ulimit settings for Viya Nodes. The settings reside in the /etc/security/limits.conf file.
      ```
      sas soft  nofile  150000
      sas hard nofile  150000
      sas soft  nproc 100000
      sas hard nproc 100000
      sas soft  stack 10240
      sas hard stack 10240
      ```

      *Note: You might need to adjust additional Linux operating system settings in order to support these recommended ulimit settings.
      Semaphore Settings
      On Red Hat Linux and SUSE Linux deployments, there are recommended semaphore settings. The settings reside in the /etc/sysctlconf file.
      kernel.sem=512 32000 100 1024
      net.core.somaxconn=2048 for SEMMSL, SEMMNS, SEMOPM, and SEMMNI

Host Resources
Viya Server Host Resource Provisioning

- CPU
  - Sized Per EEC Sizing. Prefer 2 Socket Host Architectures for Bare Metal. Fastest Processor you can afford. Skylake makes most jobs run in about half the time as Broadwell/Haswell. CAS Controller and Worker nodes should have the same chipset model.

- Memory
  - The minimum Memory Amount for Any Viya Server (CAS Controller/Worker, Services, SPRE, Web Services) is 8 GB/core. Memory Sizing for CAS Workers is sized to hold data of incoming Job Table Loads in Memory Mapped Files. Memory installations can range from 256 GB/Worker Node to 1+ TB/Worker Node. CAS Workers should have the same RAM provisioning.
  - Set the cas.memorysize option to 80% of RAM on the CAS workers.

- CAS_Disk_Cache – Backing Store for Memory Mapped Files
  - This on-device extension of the memory mapped files is typically an XFS or EXT4 file system, it cannot be placed on shared file systems such as GPFS, GFS2, Veritas VxCFS, etc. The devices must provide the highest performance you can afford (preferably a stripe of 4 – 6 or more (based on workload) High Speed Write Intensive SSD or NVMe drives, local to the Host Bus. Device head typically exceeds capacity needed to gain device bandwidth aggregation. Capacity for this file system can range from 1.5x RAM to 2x RAM. SAS Studio recommends 3x RAM, and SAS Visual Forecasting recommends 10x RAM. The absolute minimum bandwidth it must support is 100 MB/sec/Core on the host node it is attached to. Faster is better! CAS Workers should have the same CAS_Disk_Cache provisioning.

- Microservices Separation
  - Never Combine Microservices Server with any other role, consider separating PETs (rabbitmq; consul; postgres) and Microservices if PET becomes too noisy/latent for its neighbors

Virtual & Cloud Provisioning

- Underlying cloud infrastructure must match or exceed minimum hardware requirements for bare-metal EEC Hardware Estimate Guidelines.

- Ensure throughput for CAS_DISK_CACHE devices is high throughput/low latency – using front-facing multiple NVMe devices works best on cloud servers

- In a distributed deployment, each CAS worker should be configured identically; provisioning like instance types of Worker nodes is optimal

- Network bandwidth between CAS nodes should be high throughput (=> 10 Gbps) guaranteed (not maximum available)

- Deployment requires sufficient memory from EEC Sizing (not specific to vendor offering)

- It is Optimal to separate stateless and stateful services onto their own cloud servers

- Keep Source Data in Cloud Storage

- Keep Cloud Resources in Same Availability Zone

- Various host instance types will be needed for the different Viya Server roles – SPRE, Services, Web Services, CAS Controller, CAS Workers, etc.

- For VMWare – Hypervisor Level Activities can Derail Services Stability – Follow the VM Guide

Q&A