# CSCS Site Update

Slurm User Group 2016
Nicholas P. Cardo, CSCS
September 26, 2016

"Someone calling themselves a customer says they want something called service."
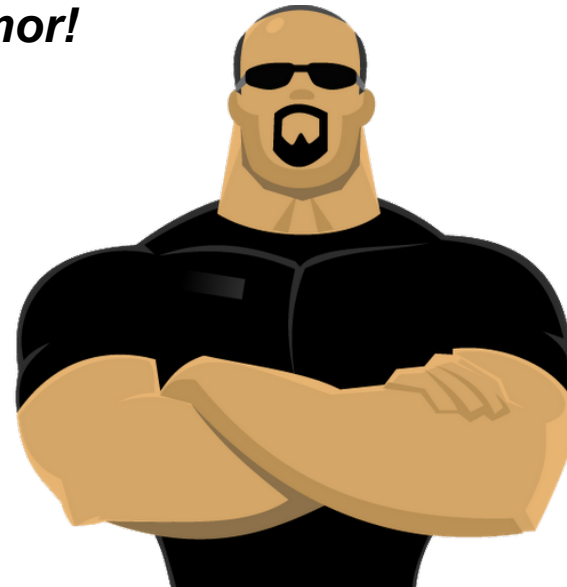
# Agenda

- Who We Are
- Trends
- Slurm @ CSCS
- Thoughts

*Lessons Learned*

**WARNING: There will be audience participation and a poor attempt at humor!**

# Centro Svizzero di Calcolo Scientifico (CSCS)

*"Founded in 1991, CSCS, the Swiss National Supercomputing Centre, develops and provides the key supercomputing capabilities required to solve important problems to science and/or society. The centre enables world-class research with a scientific user lab that is available to domestic and international researchers through a transparent, peer-reviewed allocation process. CSCS's resources are open to academia, and are available as well to users from industry and the business sector. The centre is operated by ETH Zurich and is located in Lugano.*

*CSCS is a unit of the Swiss Federal Institute of Technology in Zurich (ETH Zurich). Since October 1st, 2008, CSCS has been part of the division headed by the Vice president for personnel, resources and infrastructures of ETH Zurich."*

Source: http://www.cscs.ch/about/index.html

# The Past Year *(consolidation & migration)*



One Big System with many diverse customers and lots of unique requirements.

workload

workload

workload

"CSCS's resources are open to academia, and are available as well to users from industry and the business sector."

# Slurm @ CSCS

Each System
Accounts in Their
Own Tables

Central Slurm Database
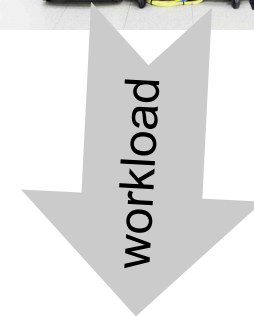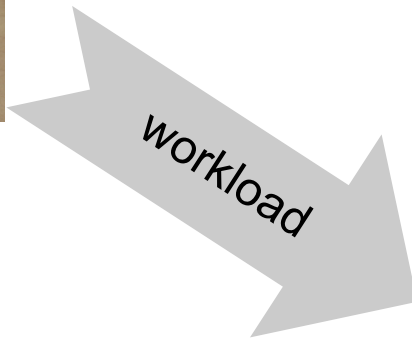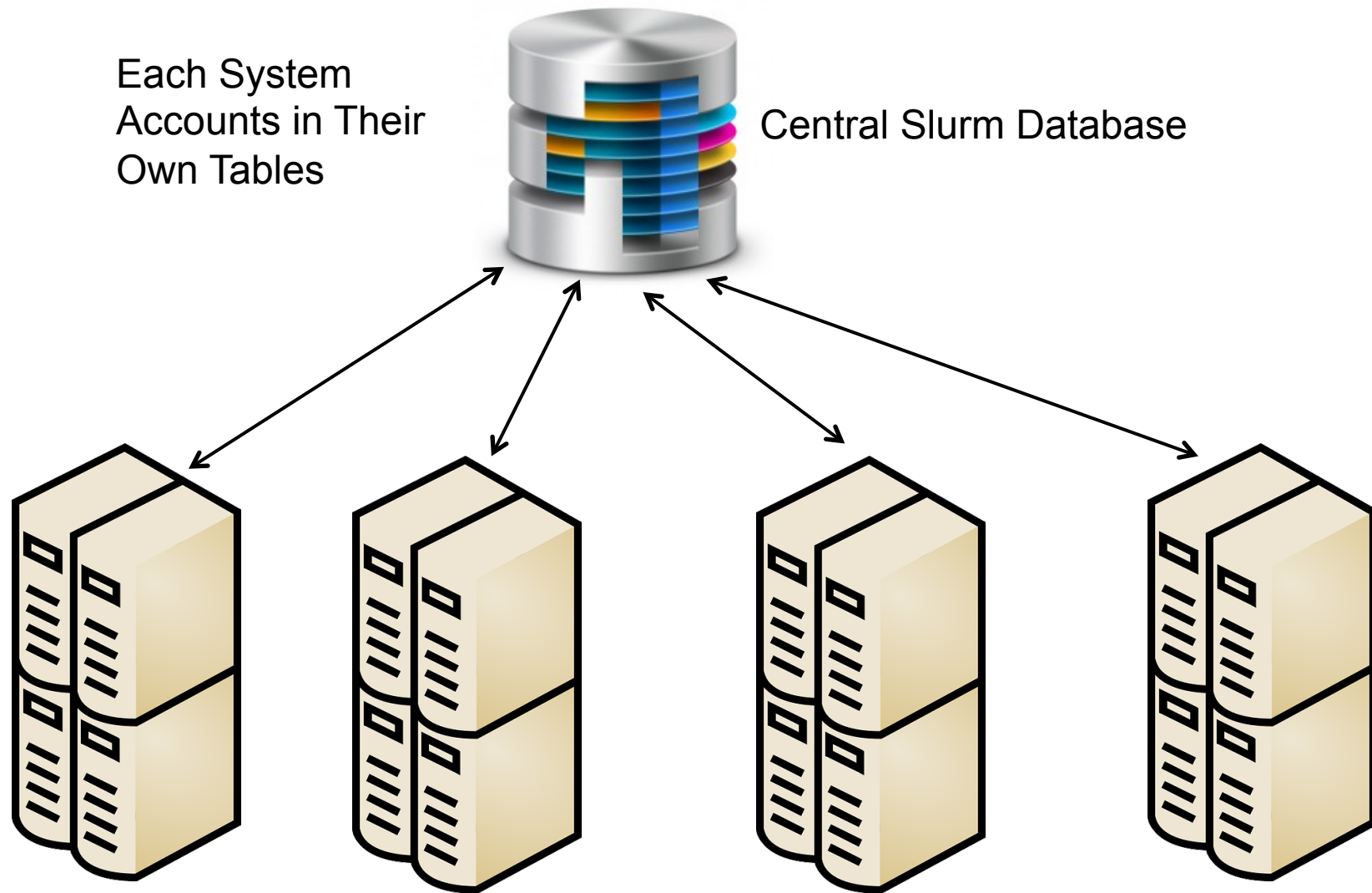
# Challenges

- Can't have per system defaults, only 1 per association.

- Consolidation of customers brings complex requirements.
  - Sometimes competing requirements.

- Keeping it fair to all.

- Keeping it simple for the users.

- Diagnosing problems.

- Explaining why…
  - *Why isn't **my** job running?*
  - *Why did **that** job start before mine?*
  - *Why are there **idle** nodes?*
  - *Why did **my** job fail?*
  - ***But my jobs are more important!***

Let's have some fun and everyone please stand. If you have received the question or have been told the statement, please be seated.

cscs

ETHzürich

# Increasing Complexity

- Single system with multiple compute architectures.

- Single system with nodes of different memory sizes.

- Single system with nodes of different memory speeds.

- Single system with nodes of different memory sizes.

- Single system with Dedicated AND Shared compute nodes.

- Single system where not all users have access to all compute resources.

- Large volume of jobs.

- Large and Small scale jobs.

- Reliability.

# Rising To The Challenge!

- ## Step 1: Panic!
  - Scream
  - Shave your head
  - Get a tattoo

- ## Step 2: Accept your fate!
  - Take a deep breath and count to 10.

- ## Step 3: Who Cares!
  - Ok, reality sets in.  You have customers, support staff, and lets not forget bosses.

- ## Step 4: Analyze it!
  - Break it apart and solve the pieces.

- ## Step 5: Celebrate!
  - Revel in your success and enjoy a cold beer!

Sir Henry Royce
1863 - 1933

*"Strive for perfection in everything you do. Take the best that exists and make it better. When it does not exist, design it. Accept nothing nearly right or good enough."*

CSCS

ETH zürich

# Slurm Version(s)

- Moving from 15.08 to 16.05.

- Some legacy systems running 14.11. These systems are in the process of being phased out.

- *Note that older versions are not compatible with 16.05.*

- We upgrade the database first, so that we can rotate through the systems to upgrade them when appropriate.
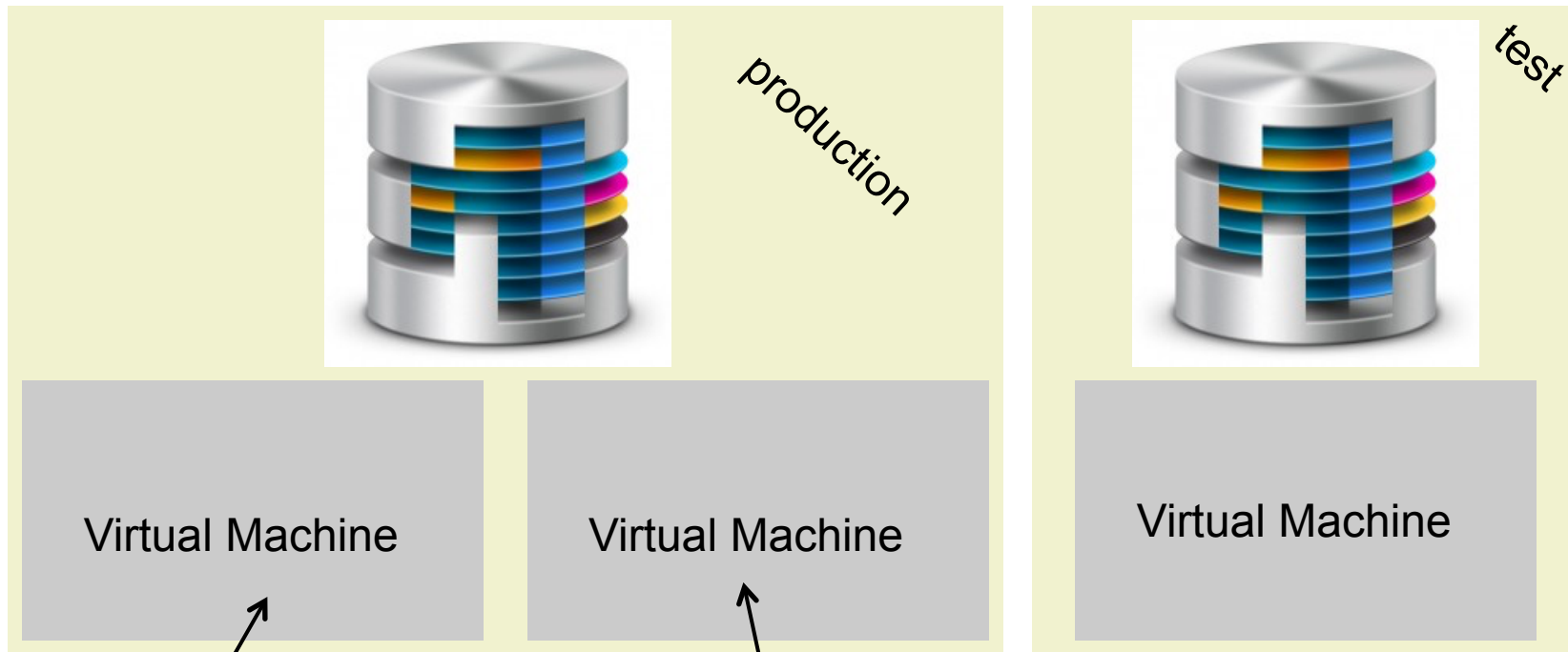
*We have realized that we do NOT have a documented strategy in place for keeping the systems up-to-date with Slurm releases and in sync with each other.*

# Slurm Data Reliability



production

test

Virtual Machine

Virtual Machine

Virtual Machine

AccountingStorageHost          AccountingStorageBackupHost

# Slurm Control Daemon

**BackupAddr**
The name that **BackupController** should be referred to in establishing a communications path. This name will be used as an argument to the gethostbyname() function for identification. For example, "elx0000" might be used to designate the Ethernet address for node "lx0000". By default the **BackupAddr** will be identical in value to **BackupController**.

**BackupController**
The short, or long, name of the machine where Slurm control functions are to be executed in the event that **ControlMachine** fails (i.e. the name returned by the command "hostname -s"). This node may also be used as a compute server if so desired. It will come into service as a controller only upon the failure of ControlMachine and will revert to a "standby" mode when the ControlMachine becomes available once again.

**ControlAddr**
Name that **ControlMachine** should be referred to in establishing a communications path. This name will be used as an argument to the gethostbyname() function for identification. For example, "elx0000" might be used to designate the Ethernet address for node "lx0000". By default the **ControlAddr** will be identical in value to **ControlMachine**.

**ControlMachine**
The short, or long, hostname of the machine where Slurm control functions are executed (i.e. the name returned by the command "hostname -s"). This value must be specified. In order to support some high availability architectures, multiple hostnames may be listed with comma separators and one **ControlAddr** must be specified. The high availability system must insure that the slurmctld daemon is running on only one of these hosts at a time. See the **RELOCATING CONTROLLERS** section if you change this.
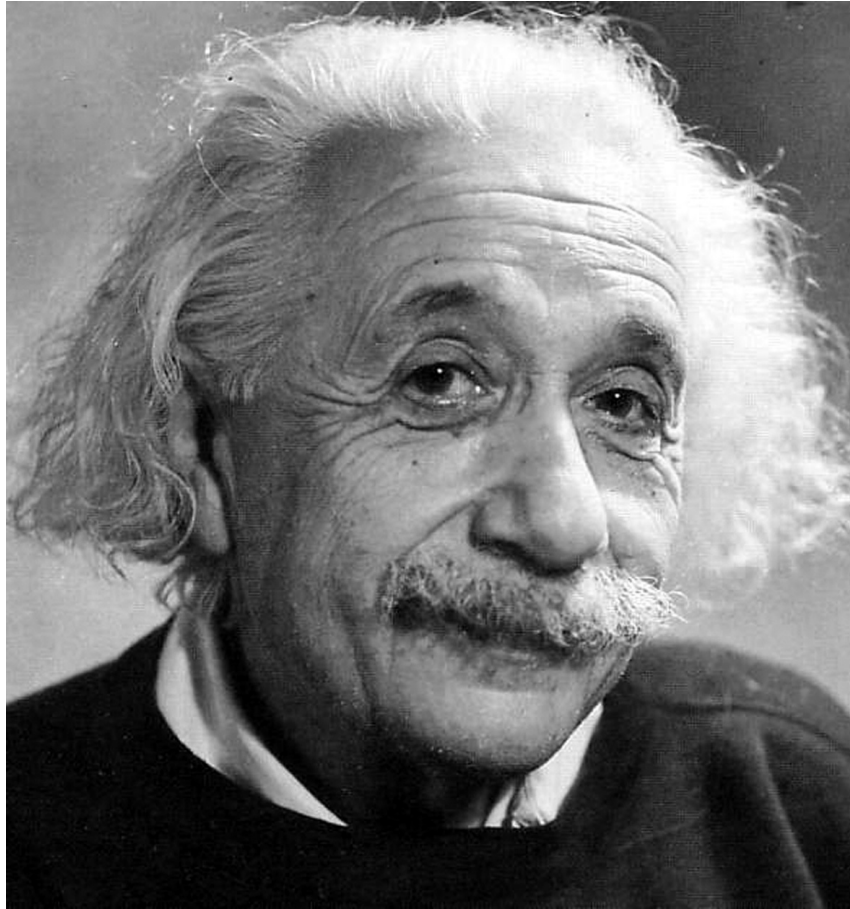
Source: http://slurm.schedmd.com/slurm.conf.html

## Configure Primary and Backup Control Daemon Nodes

# What do Users Want?

- An easy to use User Interface.

- Easy to understand queue structure.

- Fast turnaround.

- No Problems, No Muss, No Fuss, Plain and Simple.
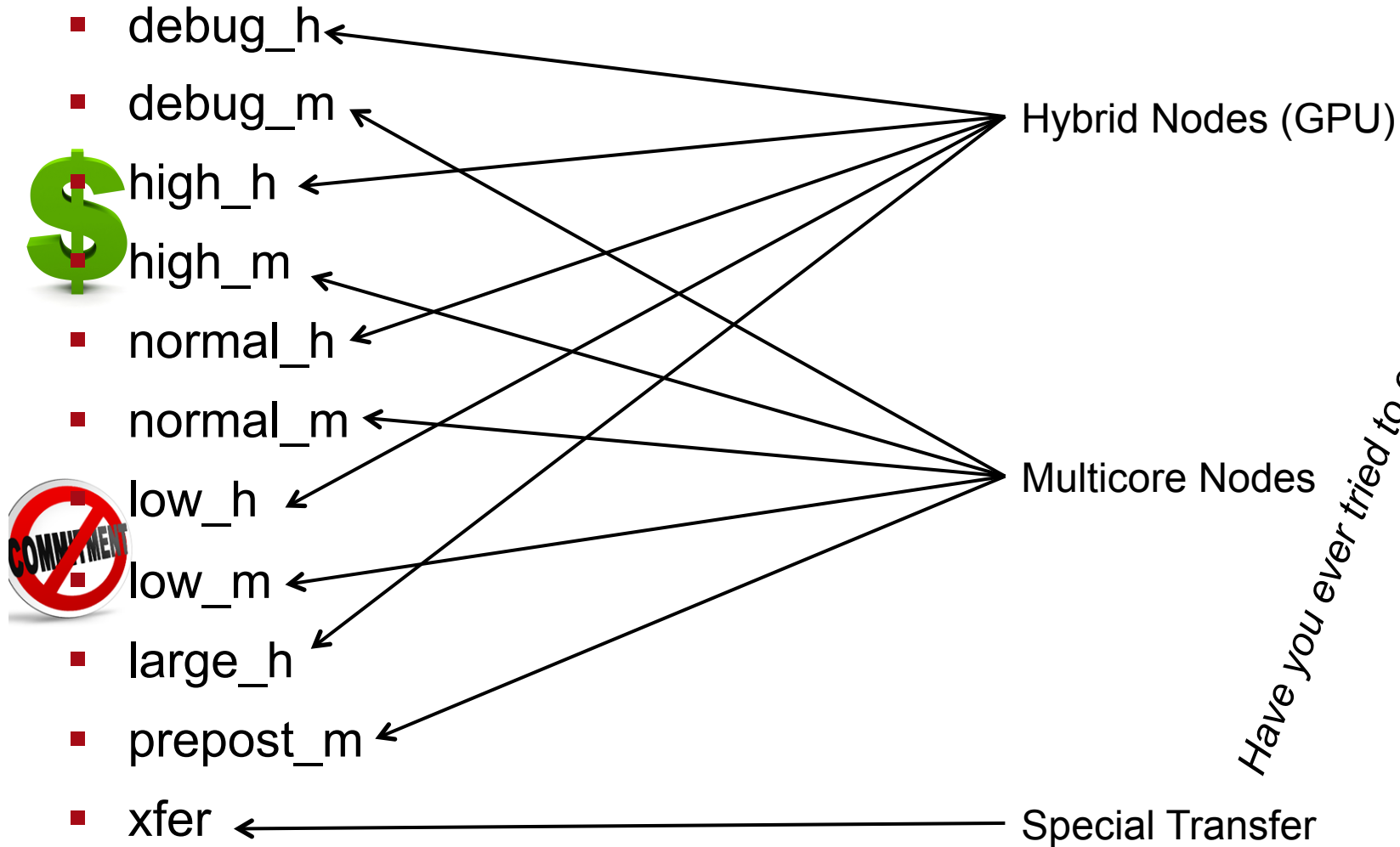
Albert Einstein

*"Make it as simple as possible, but not any simpler."*

# Queue Structure

Multifactor In Use

- debug_h
- debug_m
- high_h
- high_m
- normal_h
- normal_m
- low_h
- low_m
- large_h
- prepost_m
- xfer

Hybrid Nodes (GPU)

Multicore Nodes

Special Transfer

*Have you ever tried to explain Fairshare to a user?*

# What About Those Large Memory Nodes?

- There are 3 options:
    1. Overlapping partitions.
    2. Fixed partitions.
    3. Specify memory constraints.

- They are mixed in with the normal, high, and large nodes.

- The memory defaults are for the regular memory size nodes.

- Users must specify memory constraints for the batch job.

"Software is a great combination between artistry and engineering."

Bill Gates

# Thoughts…

The sacct utility has been found to be very powerful, but very difficult for users to understand and use on their own.  As a result we have had to create additional utilities to present information to the users.

Users need an easy to use and interpret utility in order for them to understand how they managed to consume their hours.

One area lacking, but the raw data exists, is scheduler efficiency metrics.  The basic metrics of batch systems, with respect to job processing are:  Queue Wait Time and Job Turnaround Time.  Another type of metric is how long jobs are waiting while Ineligible.  Also compute node utilization.

The API documentation could be expanded a bit.  The online documentation is very limited and in order to successfully use the power of the API, you need to read the man pages and header files.  Having a real API manual would be nice to have.
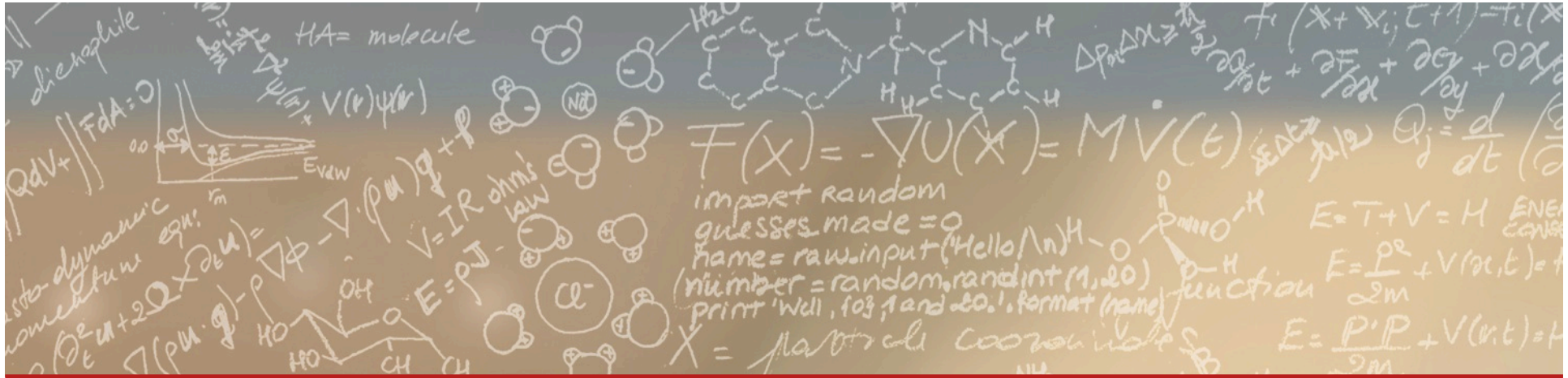
Native Slurm on Cray does not properly handle admindown or suspend nodes reliably.  There seems to be a disconnect between Cray and Slurm.

cscs

*ETH*zürich

# Thank you for your attention.