# Machine Features

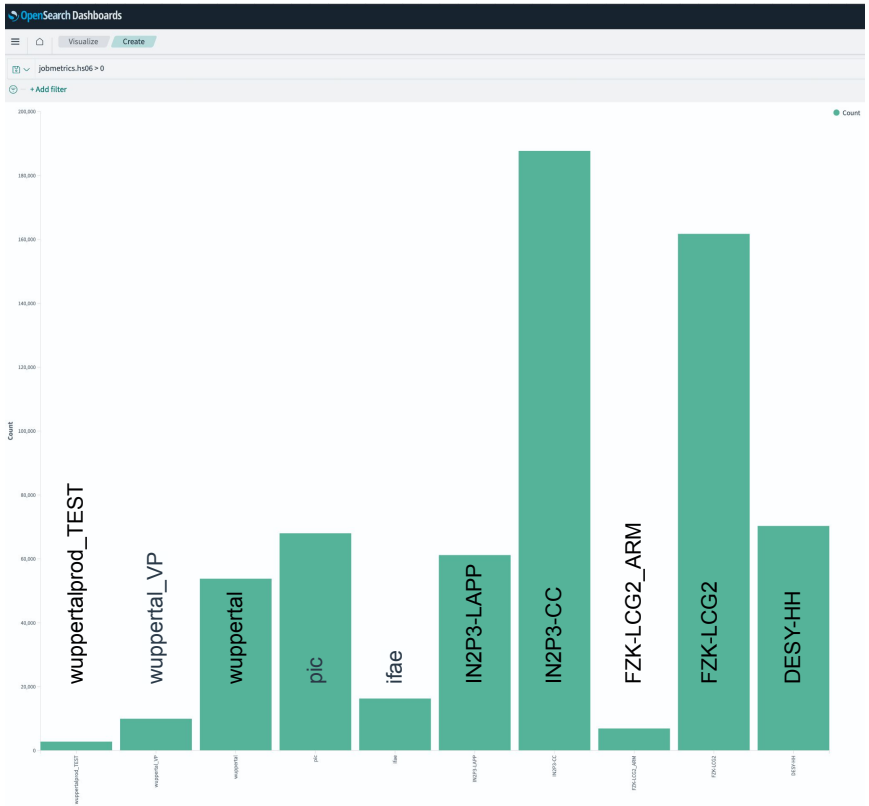
P. Nilsson (BNL)

# Machine Features

- Original proposal: https://hepsoftwarefoundation.org/notes/HSF-TN-2016-02.pdf
  - I.e. there are also “Job Features”, but unclear if used by anyone
- In case local environment variable `MACHINEFEATURES` is set, pilot attempts to pull several published variables relevant for the worker node it runs on
  - E.g. on DESY-HH: `MACHINEFEATURES=/etc/batch/machinefeatures/`
- For benchmark reporting, pilot uses ‘`hs06`’ and ‘`total_cpu`’ variables and reports with job metrics
  - Benchmark value is corrected for core count like so:
    - *`reported_hs06 = hs06 * corecount * perf_scale / total_cpu`*
    - *`perf_scale`* is currently set to 1, but likely needs to be updated to a more realistic number
- For graceful shutdown of pilot, ‘`shutdowntime`’ *can be* used
  - If set, pilot will start monitoring the shutdown time once per minute and will terminate pilot ten minutes before the desired shutdown time

# Queues using machine features

- DESY-HH using hyperthreading
  - **MF/Pilot:** `reported_hs06=116.07 (3250) total_cpu=224 corecount=8 perf_scale=1`
    - where 3250 is the original `hs06`
  - **Pilot:** `found 112 cores (56 cores per socket, 2 sockets)`
    - -> “total CPU” = 112 or 224?
- Wuppertal not using hyperthreading
  - **MF/Pilot:** `reported_hs06=27.19 (1740) total_cpu=64 corecount=1 perf_scale=1`
  - **Pilot:** `found 64 cores (32 cores per socket, 2 sockets)`
    - -> “total CPU” = 64
- Better to factor out hyperthreading from the reported hs06?




From Tania

# Misleading Total CPU

- From original machine features proposal

  **total_cpu** Number of processors which may be allocated to jobs. Typically the number of processors seen by the operating system on one worker node (that is the number of "processor :" lines in /proc/cpuinfo on Linux), but potentially set to more or less than this for performance reasons. (Use case 3.)

  - I.e. with 'processor' they mean 'core', hyperthreading not taken into account - also, number of "processor" lines in /proc/cpuinfo might be affected in case of virtual machines / virtual CPUs
- Pilot calculates `total_cpu` by counting reported "Core(s) per socket" and "Socket(s)" in output from lscpu command
  - Can identify if hyperthreading is used, and if so compensate `reported_hs06` (divide by 2)

# All machine features

**total_cpu** Number of processors which may be allocated to jobs. Typically the number of processors seen by the operating system on one worker node (that is the number of "processor :" lines in /proc/cpuinfo on Linux), but potentially set to more or less than this for performance reasons. (Use case 3.)

**hs06** Total HS06 rating of the full machine in its current setup. HS06 is measured following the HEPiX recommendations [3], with HS06 benchmarks run in parallel, one for each processor which may be allocated to jobs. (Use case 3.)

**shutdowntime** Shutdown time for the machine as a UNIX time stamp in seconds. The value is dynamic and optional. If the file is missing, no shutdown is foreseen. (Use case 4.)

**grace_secs** If the resource provider announces a shutdown time to the jobs on this host, that time will not be less than grace_secs seconds after the moment the shutdown time is set. This allows jobs to begin packages of work knowing that there will be sufficient time for them to be completed even if a shutdown time is announced. This value is required if a shutdown time will be set or changed which will affect any jobs which have already started on this host.