





#### Ladies and Gentlemen, we can rebuild things. We have the technology. We have the capability to make the world's best computing environments. ADAC will lead the way. Better than before. Better. Stronger. Faster.

# **GPU Usage Collection**

ADAC Tokyo Nicholas P. Cardo, CSCS February 16, 2018



# **Overview of the Situation**

- The Problem
  - Thousands of batch jobs run each day, but did they make use of any GPUs?
    - How effectively were they used?
- The Objective
  - Create a mechanism to capture indicators of GPU usage per batch job
  - Record the captured data with job data
  - Available to the user as well as system level reporting
  - Leverage existing technologies where possible
- Three Design Elements
  - Data collecting
  - Data recording
  - Data reporting









ADAC - GPU Statistics 3

ETH zürich

#### **Design Overview**

# Sound familiar? RUR





ADAC - GPU Statistics 4

**ETH** zürich

# Implementation – Putting the pieces together.

- Cray RUR enhanced to support Native Slurm (no ALPS)
- Data collection started via prolog
- Data collection terminated via epilog
- Data stored in Slurm account record (AdminComment) as JSON
- Data can be extracted with sacct or MySQL





ADAC - GPU Statistics 5



# **GPU Statistics (NVML)**

This is the beginning... More data available with newer devices.

- Statistics Captured
  - gpusecs Total GPU seconds used (accumulated across all nodes)
  - maxgpusecs Maximum GPU seconds used on a node (maximum of all nodes)
  - maxmem Maximum GPU memory consumed on a node (maximum of all nodes)
  - summem Total GPU memory consumed (accumulated across all nodes)
- Stored as JSON in AdminComment field in Slurm accounting





## **Batch Job Summary Report**

atch Job Summary Report						Written to the end of a job's star		
Batch Jo	b Summary Re	eport for Job "t	est1" (60	51962) oi	n dom			Sout file
	Submit	Eliç	Jible		Start	End	Elapsed	Timelimit
2018-02-	-01T07:49:21	2018-02-01T07:4	9:21 2018	3-02-01T	07 <b>:</b> 49 <b>:</b> 22	2018-02-01T07:49:30	00:00:08	00:15:00
Username	e Account	Partition	NNodes	Energy				
cardo	csstaff	normal	1	757	joules			
gpusecs	maxgpusecs	maxmem		summem				
4	4	303038464	3(	03038464				
Scratch	File System	Files	Quota	 a				
/scratch /scratch	n/snx1600tds n/snx3000tds	 1 1	1000000	- ) )				





# **Next Steps**

- To be installed on Piz Daint February 21<sup>st</sup>
- Meaningful reports need to be developed
- Application support already investigating how to link with XALT
  - Linking to a GPU enabled library doesn't mean the GPU was used
- Pass RUR enhancements back to Cray <--</li>
- Work with SchedMD to enhance accounting record
- Differentiate between no GPU usage and no GPU on a node

Give back to the community











Thank you for your attention.

