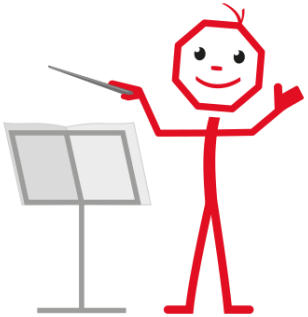


# CWG10
# Control, Configuration and Monitoring

Status and plans for Control, Configuration and Monitoring

16 December 2014

ALICE $O^2$ Asian Workshop 2014@Pusan


# Outline

- Motivation
- A brief overview of data taking operations
- Lessons learned from Run 1
- CCM Overview
- Performance tests
- Next steps

# Motivation

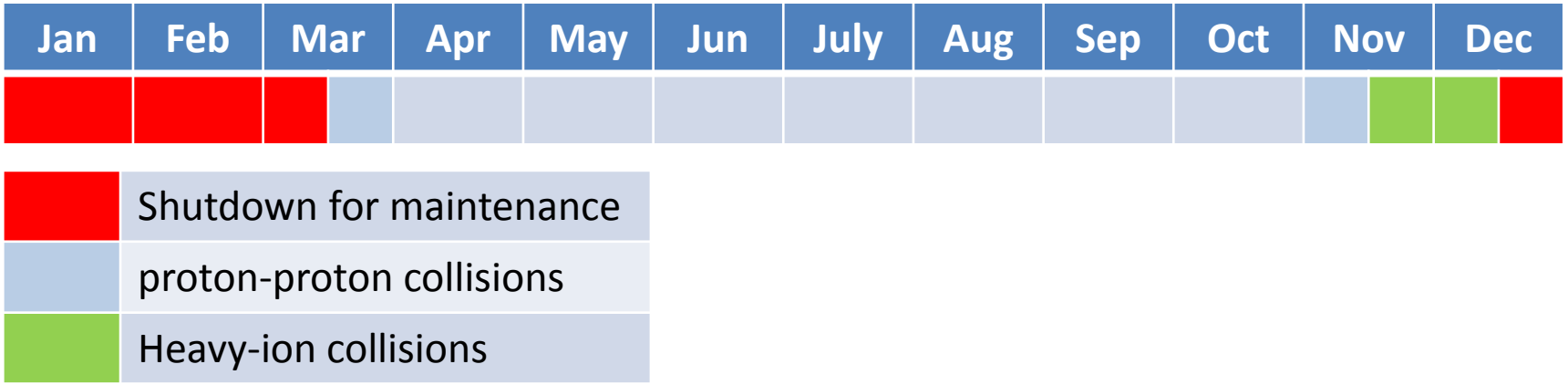
- Why do we need a Control System ?
  - Start and stop processes
  - Sequence of operations, synchronization
  - External systems
  - Automation
- Why do we need a Configuration System ?
  - Configure processes
- Why do we need a Monitoring System ?
  - Detect abnormal conditions
  - Automation

# Team

- CERN
- KMUTT, Thailand
  - See next presentation by Khanasin for an update

# A brief overview of data taking operations

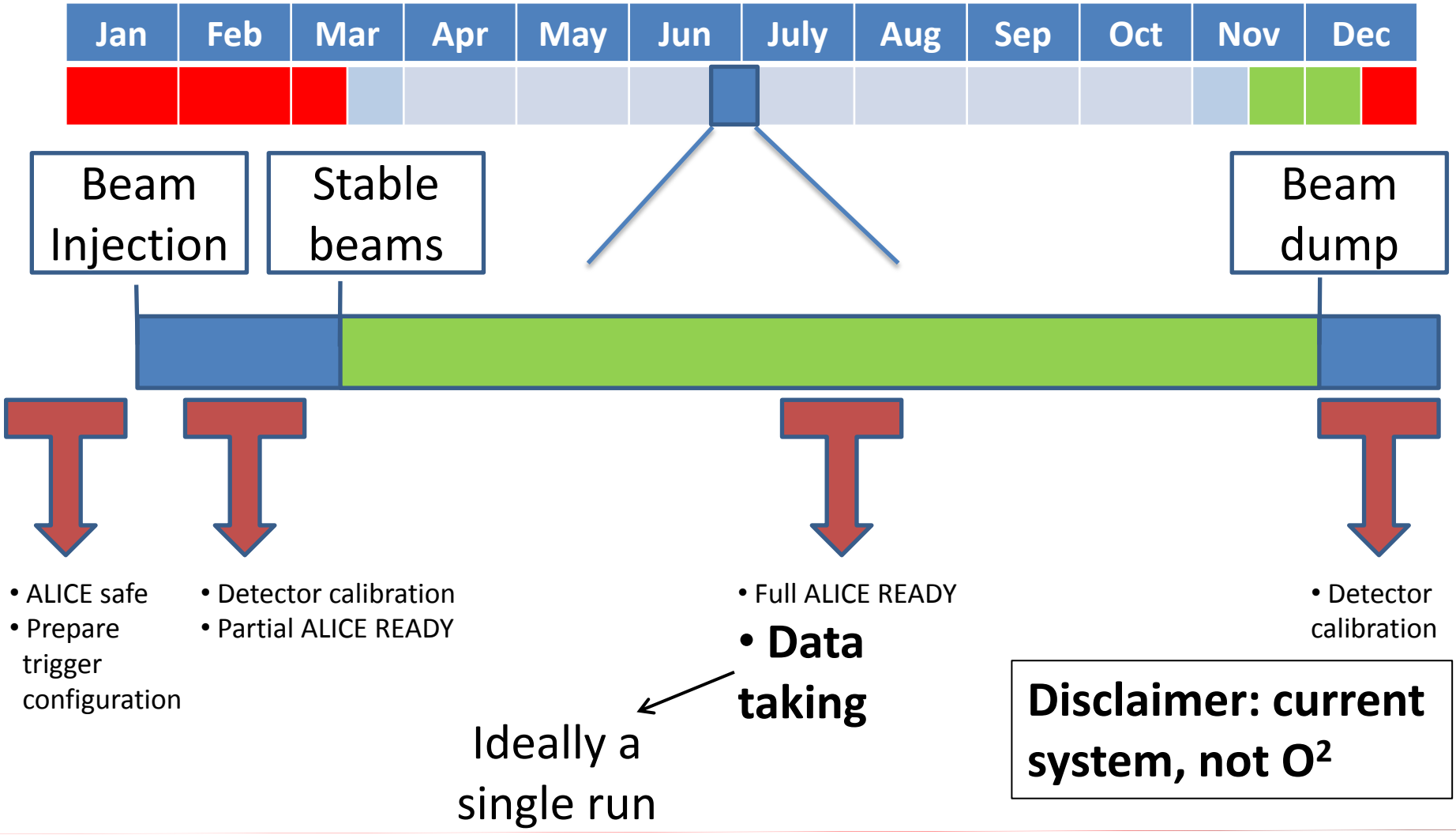
- A typical LHC year




**Disclaimer: current system, not $O^2$**

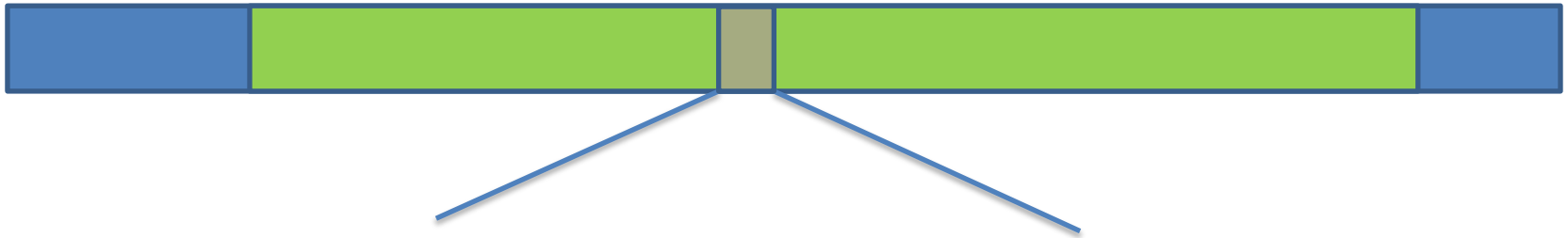
# A brief overview of data taking operations

- A typical LHC Fill (up to 30 hours)



# A brief overview of data taking operations

- A typical ALICE run



**Start-of-Run**

- Config detectors electronics
- Start online systems
- Store data taking conditions

**Data taking**

- Readout
- Event building
- Online data monitoring
- Online calibration data

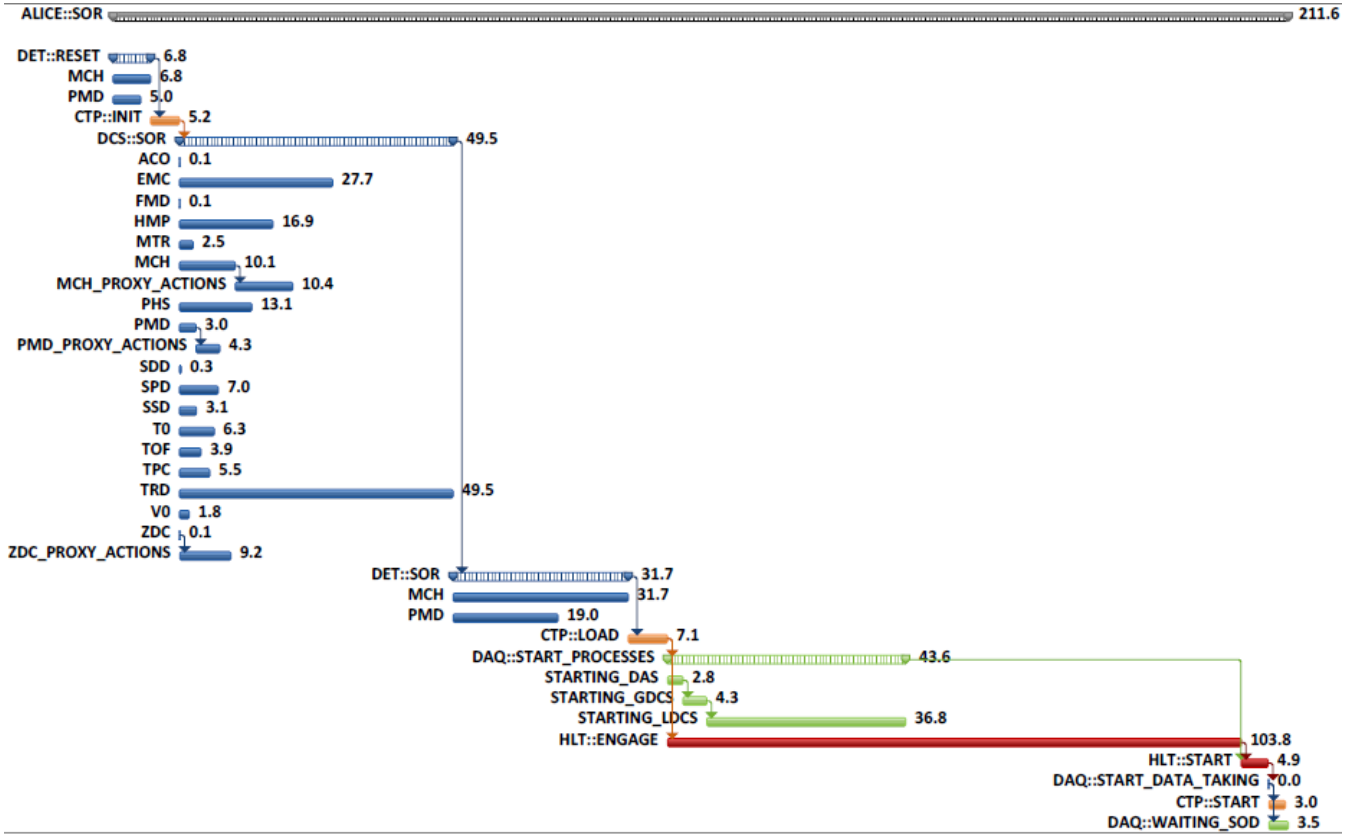
**End-of-Run**

- Export data taking conditions and calibration data to Offline
- Stop online systems

**Disclaimer: current system, not O²**

# A brief overview of data taking operations

- Run 1 SOR sequence (high level)




**Disclaimer: current system, not O²**

# Lessons learned from Run 1 (2010-2013)

- Must be fast when changing run
  - More runs than expected
  - Not everything needs to be restarted

**Run 2: Fast SOR/EOR**

- Must be flexible
  - Not every problem needs to stop a run

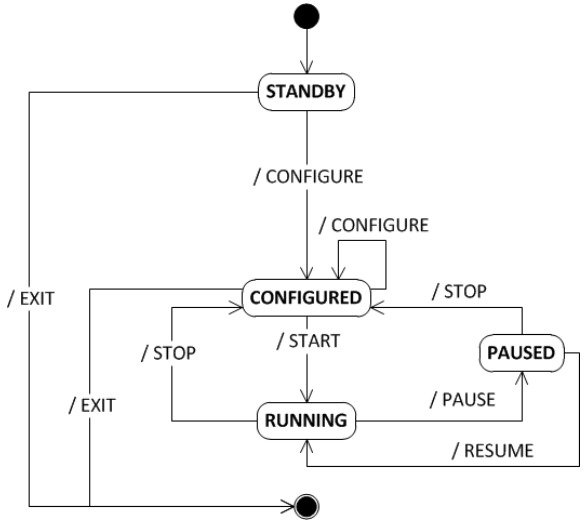
**Run 2: Pause and Recover**

- Must monitor everything
  - Data flow monitoring

**Run 2: MAD**

# Control in O² - Overview

- Process Management
  - Start/stop processes
  - Send commands to processes (CONFIGURE, PAUSE/RESUME, etc.)
  - Estimated: O(100k) processes




- Task Management
  - Ensure that actions are executed in the correct order
- Automation
  - Automatically recover from errors
  - Automatically react to internal events (e.g. need more EPNs), external events (e.g. start of LHC collisions)

# Control in $O^2$ - Notes

- Includes processes from online and offline
- Must control both synchronous and asynchronous tasks
- Cannot be seen as a batch system
  - Bound to external events (e.g. start of collisions)
  - Sequence of operations, synchronization points
  - Low latency very important

# Configuration in $O^2$ - Overview

- Configuration distribution
  - Provide processes with needed configuration parameters
- Dynamic process (re)configuration
  - Essential to achieve fast run transition
- O(1GB) of configuration data

# Monitoring in $O^2$ - Overview

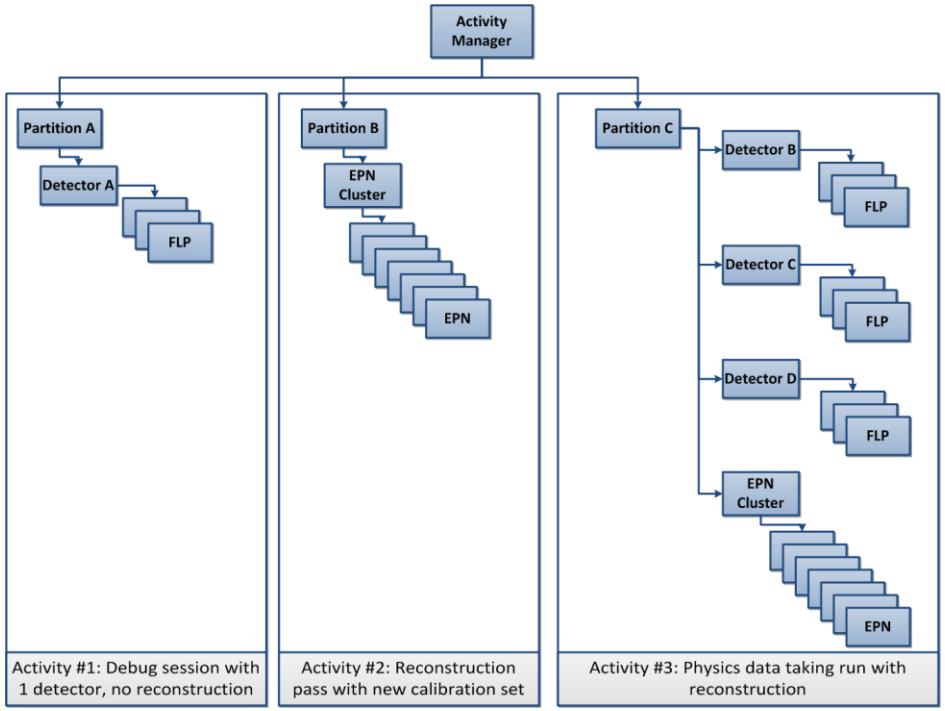
- Data collection and archival
  - System monitoring (CPU, memory, I/O, etc.)
  - Application monitoring (data rates, link backpressure, internal buffer status, etc.)
  - O(600KHz) of monitoring data
- Alarms and action triggering
  - Support shift crew, experts
  - Feedback to Control system

# Monitoring in $O^2$ - Notes

- Includes metrics from online and offline
- Includes both low and high frequency metrics
  - Low: every 30 seconds, system metrics
  - High: every second, link status
- Permanent storage will be the limiting factor
  - No need to store everything, can filter “interesting” values

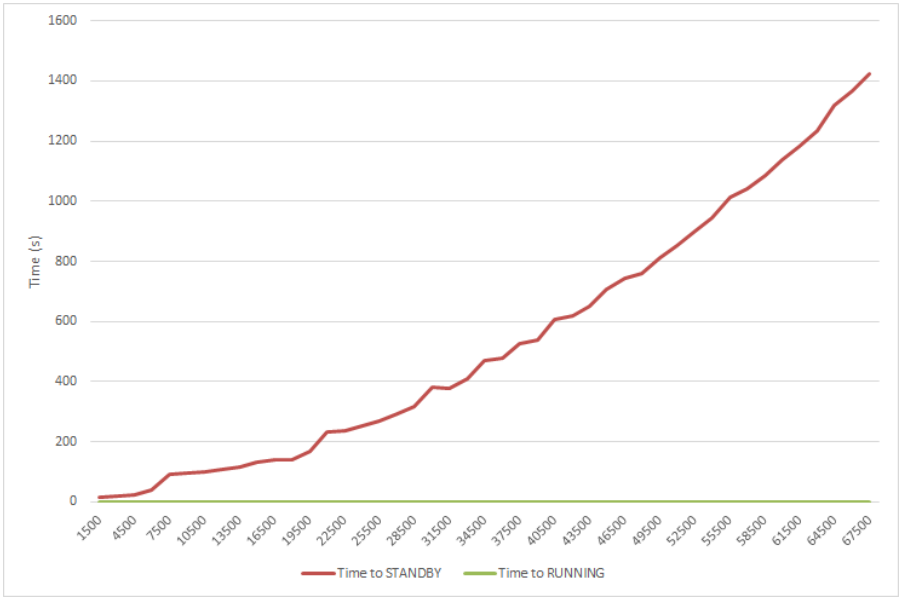
# Performance Tests: Control

- Tool: SMI (State Machine Interface)
- Setup:
  - Level 0 SMI domain:
    Partition CCM
  - Level 1 SMI domain:
    Detector CCMs
    EPN Cluster CCM
  - Level 2 SMI domain:
    FLP CCMs, EPN CCMs
  - Level 2 SMI proxy:
    local process



# Performance Tests: Control

- Setup:
  - 46 hosts
  - 1 Level 0 domain
  - 20 Level 1 domains
  - 1350 Level 2 domains
  - 67500 proxies




- Increase due to initial lookup in DIM DNS
- Conclusion: cannot use in current version

# Performance Tests: Monitoring

- MonALISA + ApMon
- Setup:
  - 10 sender nodes, up to 1000 threads per host (ApMon)
  - 1 MonALISA service, all historical record disabled
- Result: 52 KHz without data loss
- Conclusion: could use 12+ collectors to reach 600 KHz

**By Costin Grigoras**

# Performance Tests: Monitoring

- Zabbix
- Setup:
  - 10 sender nodes, up to 10 processes per host
  - 1 Zabbix Server node, 200 threads, permanent storage disabled (in-memory history enabled)
- Result: 30 KHz without data loss
- Conclusion: could use 20+ collectors to reach 600 KHz

**By Andres Gomez Ramirez**

# Next steps

- Finalise TDR
- Perform more tests:
  - Control: boost library + ZeroMQ
  - Configuration: ZooKeeper
  - Monitoring: MonALISA, Zabbix with permanent storage
- Provide CCM systems for ALFA prototype (CWG13)
- Refine design