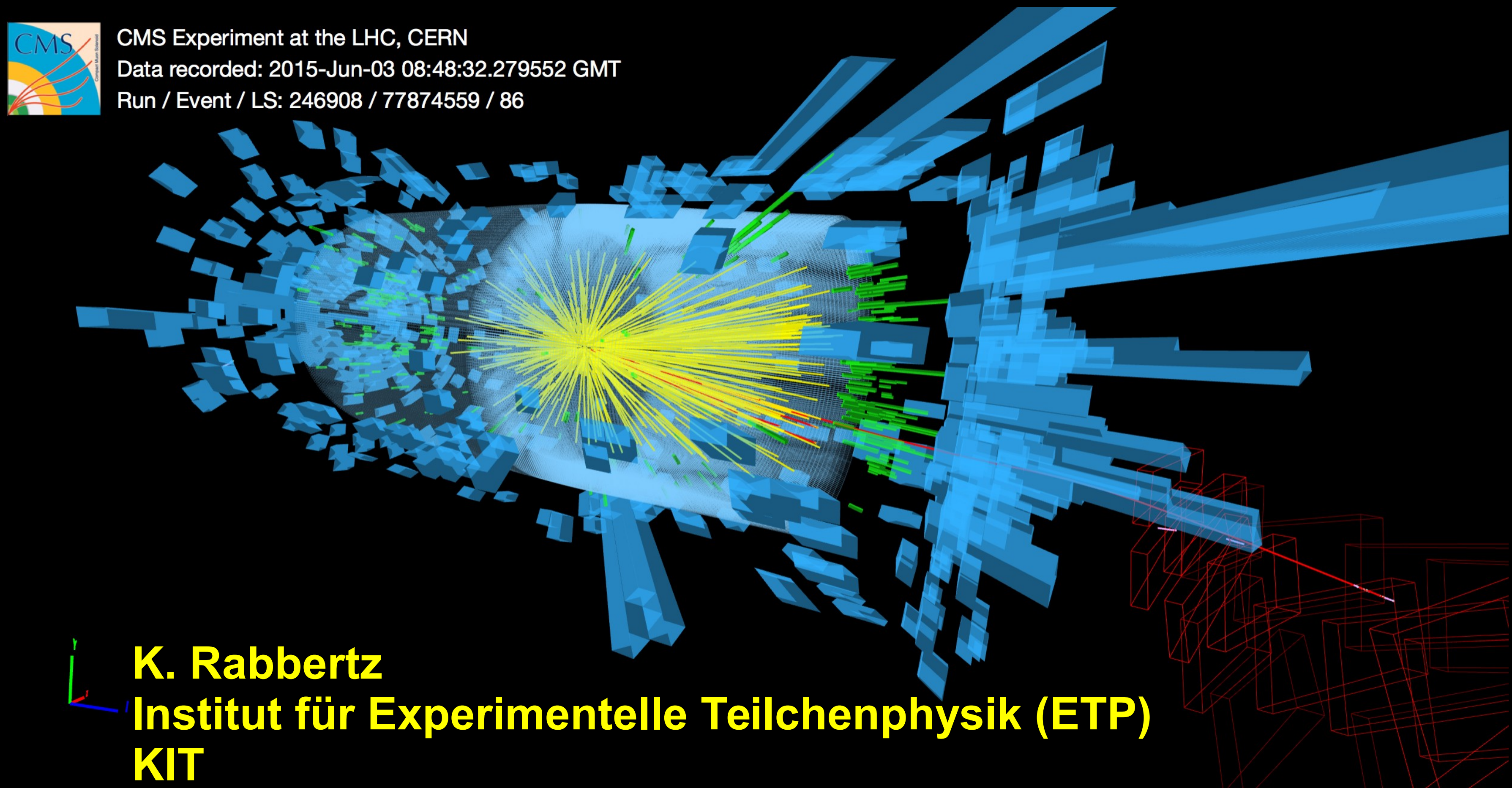
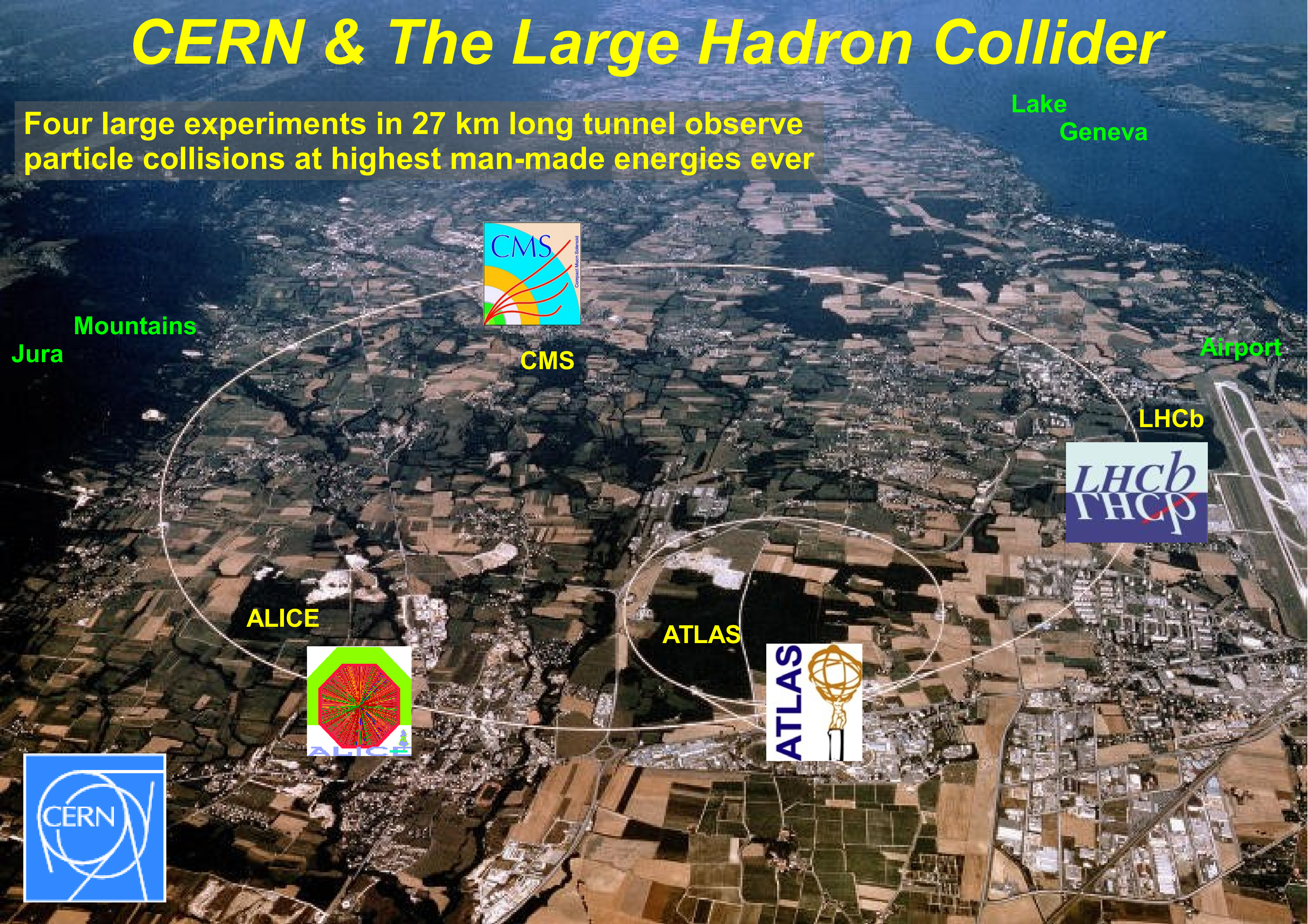
High precision predictions for particle collisions at the Large Hadron Collider

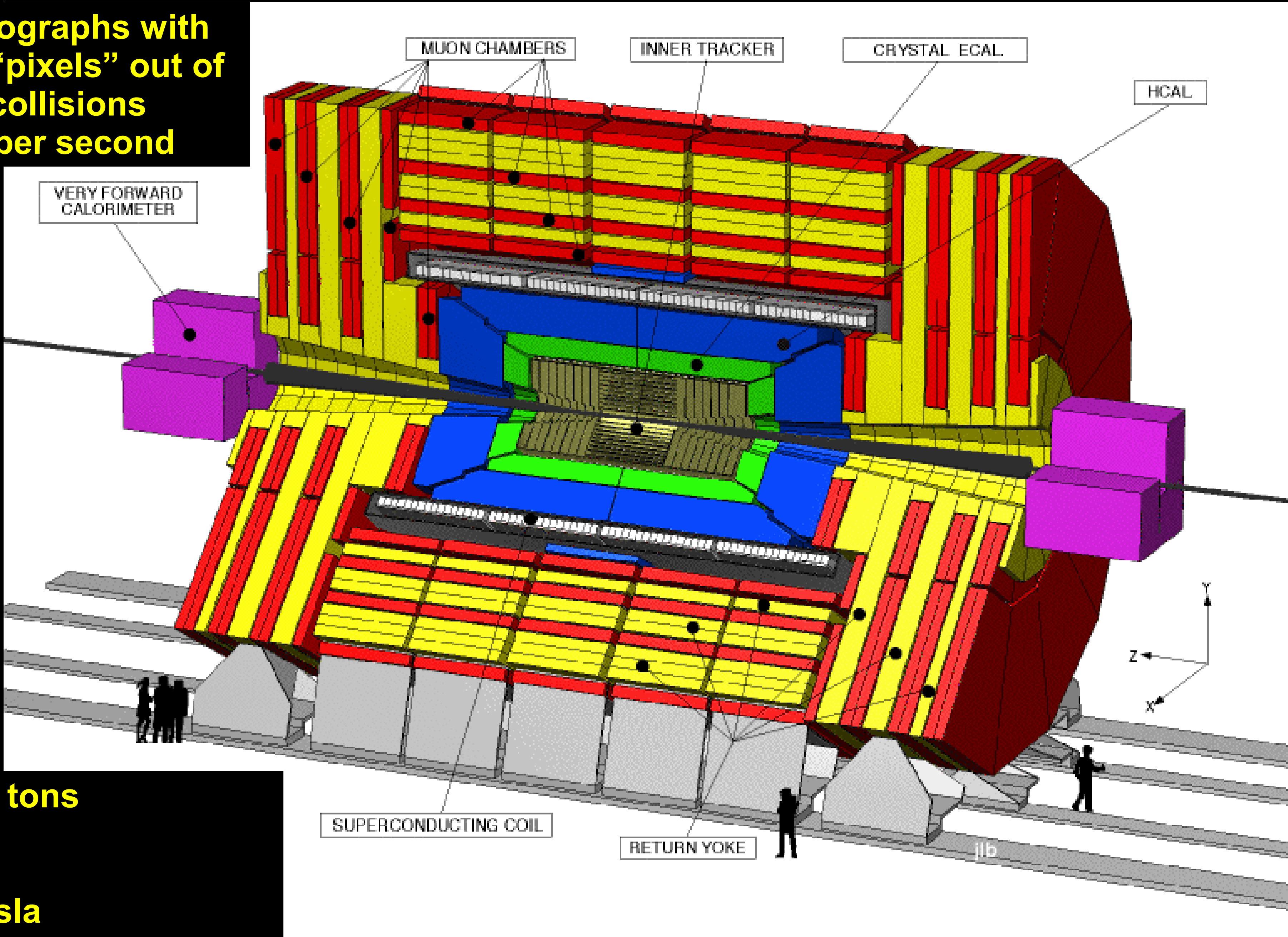


K. Rabbertz
Institut für Experimentelle Teilchenphysik (ETP)
KIT


CERN & The Large Hadron Collider
Four large experiments in 27 km long tunnel observe
particle collisions at highest man-made energies ever
Lake
Geneva
Jura
Mountains
CMS
Airport
LHCb
ALICE
ATLAS
CERN

The CMS Detector

~1000 3D photographs with
~ 100 million "pixels" out of
> 40 million collisions
per second

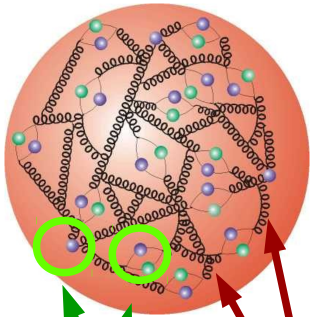


Weight: 14000 tons
Length: 21 m
Height: 15 m
Magnet: 4 Tesla

Look inside Matter

View perpendicular to particle beams

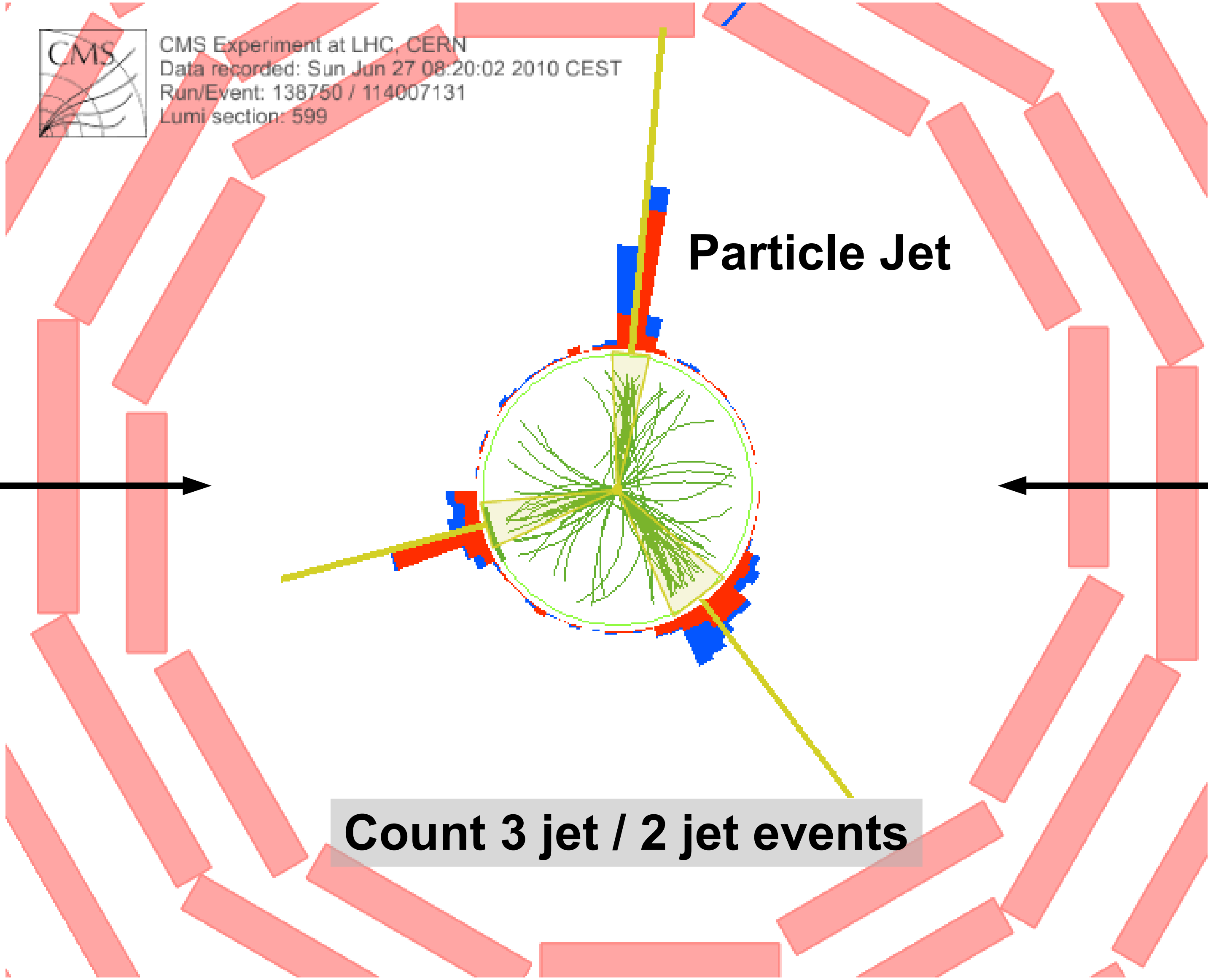
Proton



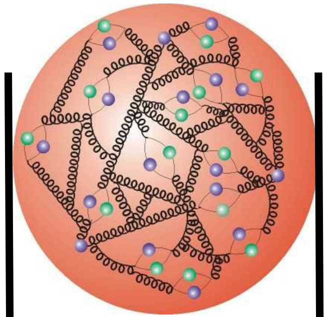
Quarks

Gluons

resolved up to 10^{-19} m



Proton

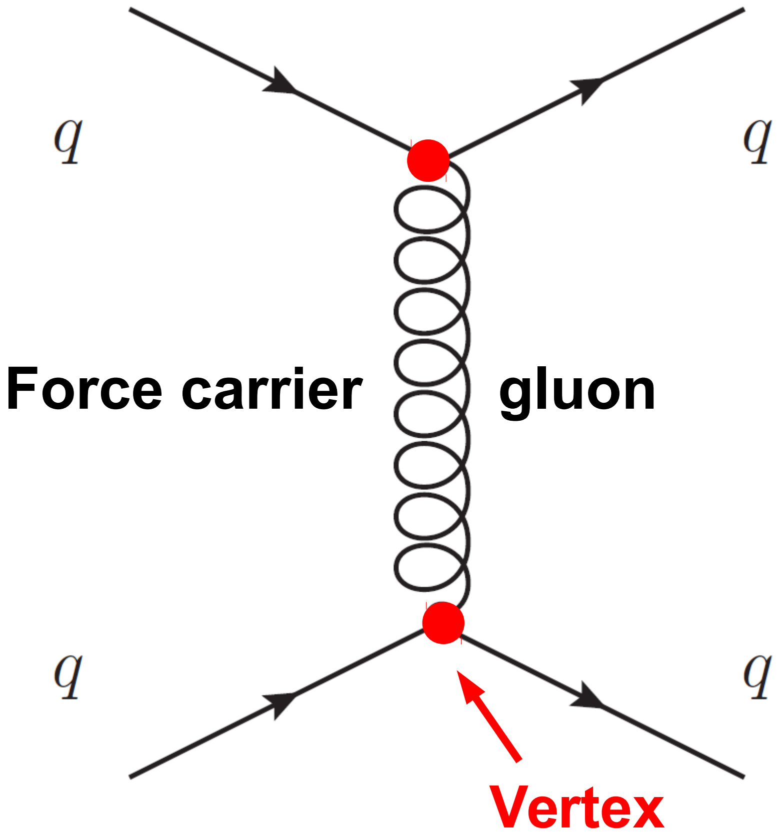


~ 10^{-15} m

Measure the nuclear/strong force inside protons

Theoretical sketch of a quark-quark collision

Two incoming quarks



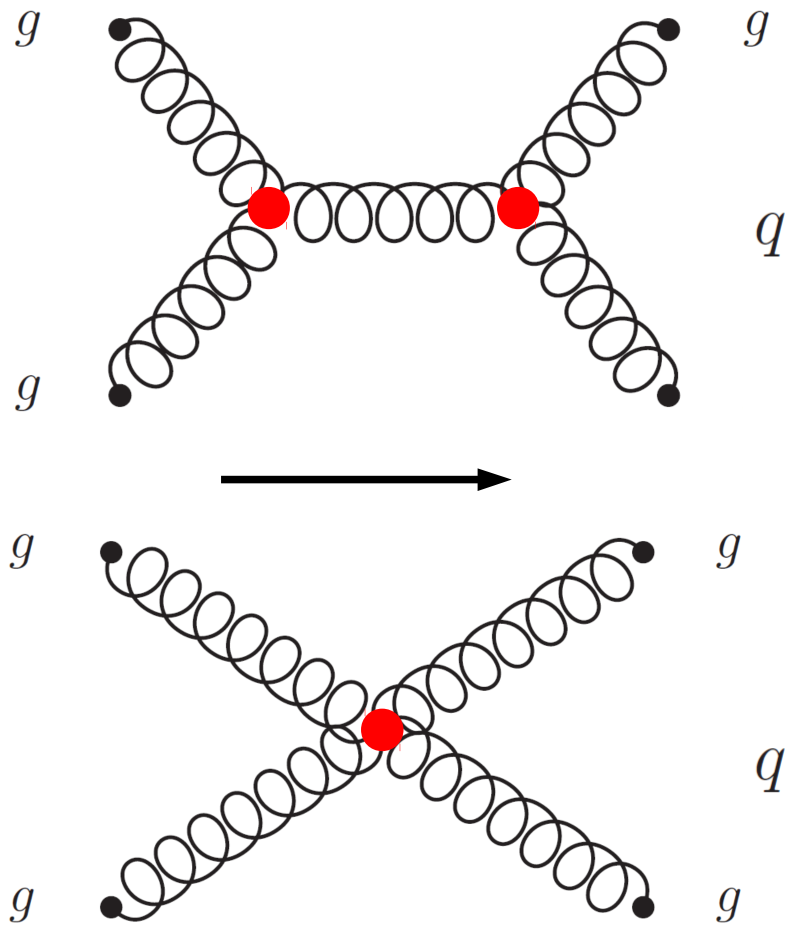
Two outgoing quarks

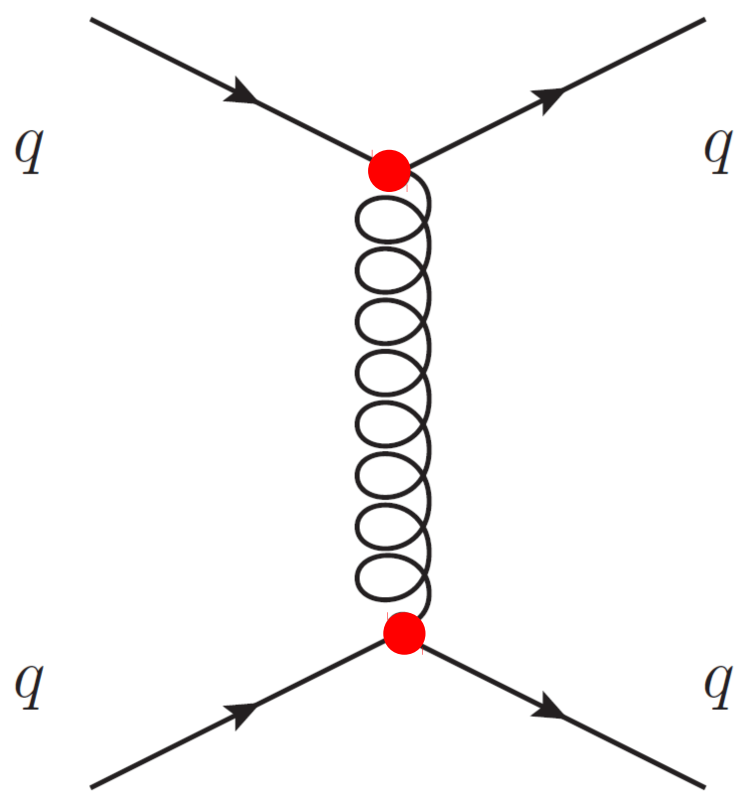
Measurable as two jets

R. Feynman (Nobel-Prize 1965): Prescription for quantitative result!

Precision Theory I.

Two-jet production with gluons!





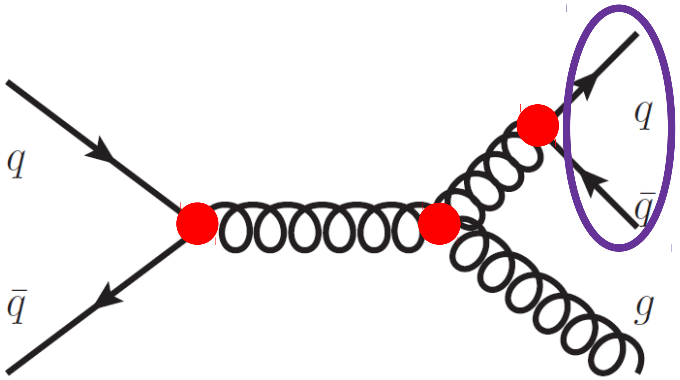
Two outgoing quarks

→

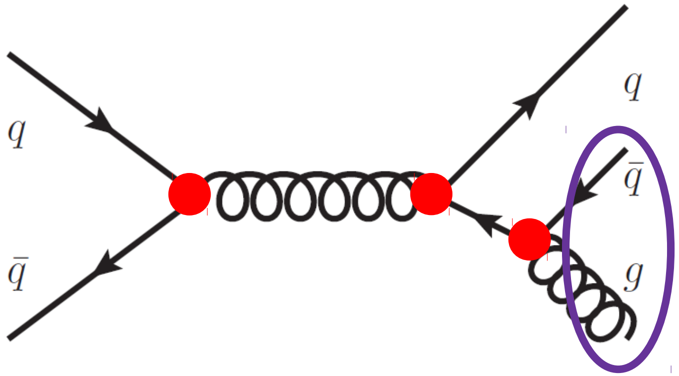
Measurable as two jets

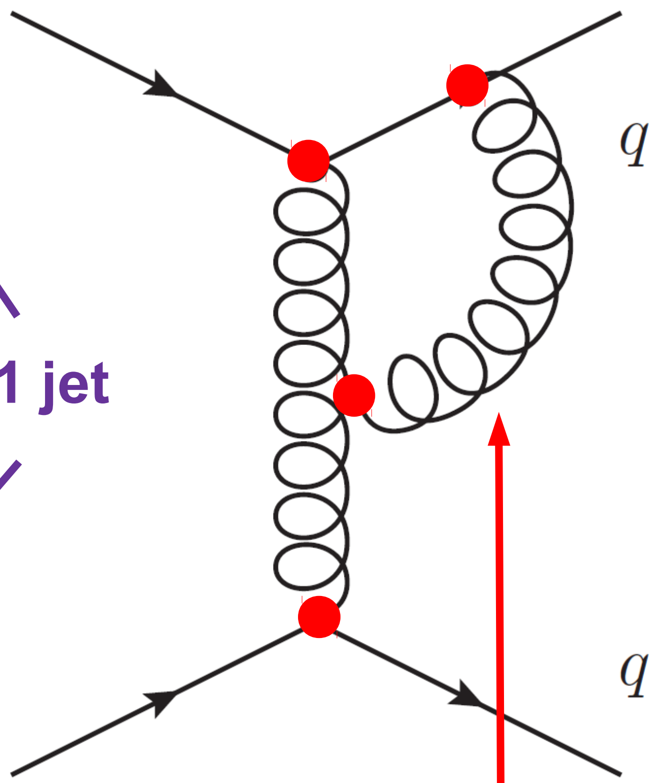
Reaction probability proportional to number of vertices
Leading order (LO) 1970s: Uncertainty estimate: 50-100%
Computing time: O(1h)

More vertices → next-to-leading order!



unresolved = 1 jet





Two outgoing quarks

Measurable as two jets

Particularly difficult: Diagrams with loops!

Next-to-leading order (NLO) 1990s: Uncertainty estimate: 5%

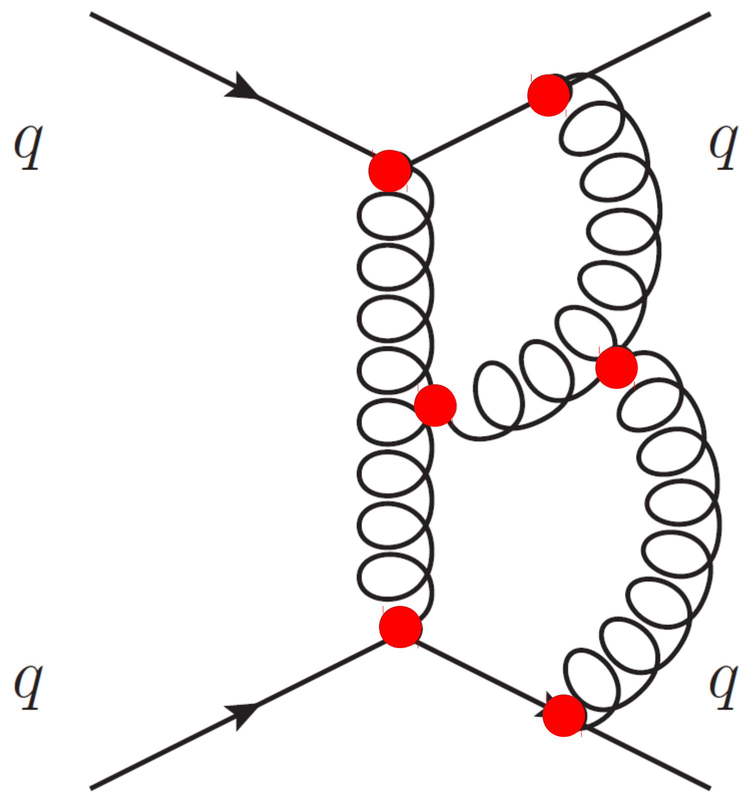
Computing time: O(1000 h)

Precision Theory III.

Many VERY complicated diagrams → next-to-next-to-leading order!

Two incoming quarks

→



Two outgoing quarks

→

Measurable as two jets

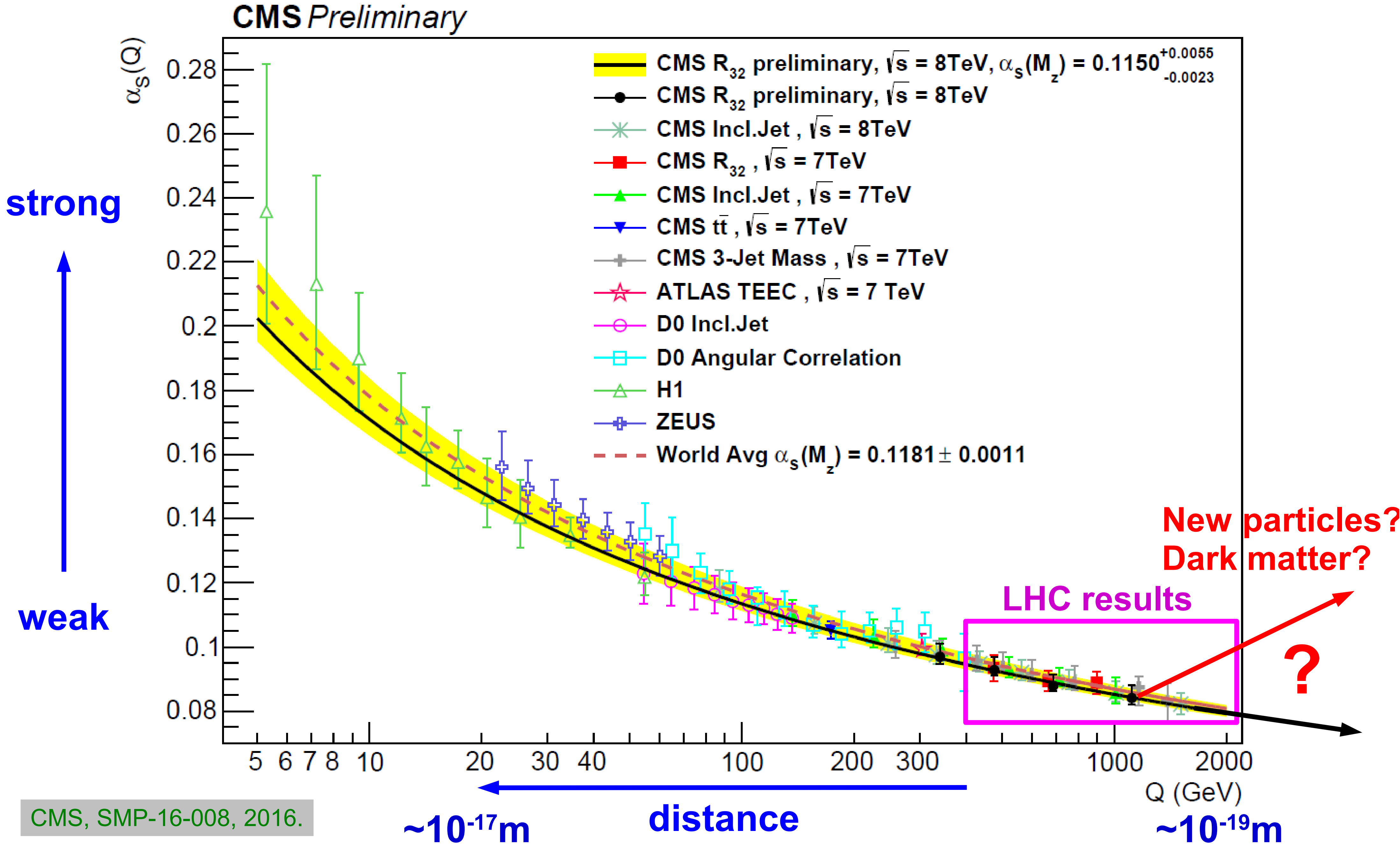
J. Currie, T. Gehrmann, A. Gehrmann-de Ridder, N. Glover, A. Huss, J. Pires, arXiv:1705.10271.

Particularly difficult: Diagrams with two loops!

NNLO: More than 20 years later 2017: Uncertainty estimate: 1%

Computing time: O(250000 h)

The Strong Force



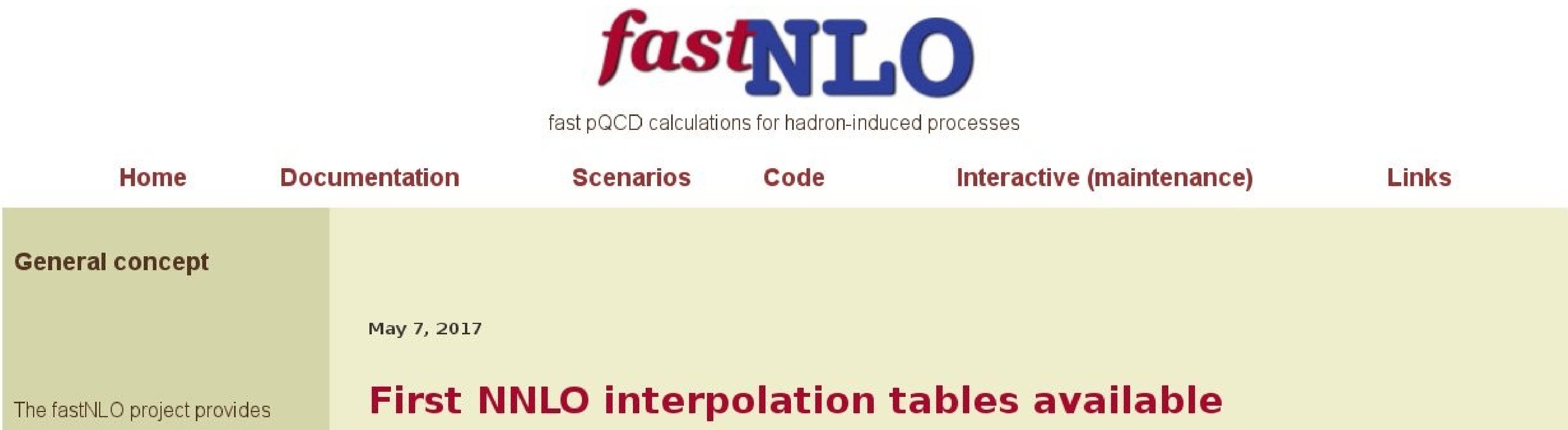
CMS, SMP-16-008, 2016.

Comparison of Theory to Data

- **Parametric fits: Need hundreds of computations for varying inputs**
- **Not possible with 250 kh computing time!**
- **Solution: Interpolation grids to avoid repetitions**
- **Two packages available APPLgrid & fastNLO**

APPLgrid, Carli et al., Eur. Phys. J. C, 2010, 66, 503.
fastNLO, Britzger et al., arXiv:0609285, 1208.3641.

- **We use fastNLO, i.a. developed at ETP** https://fastnlo.hepforge.org



Calculation Workflow

- **1. Preprocessing:** Check of interpolation quality
  - Short test jobs — O(10 h)
- **2. NNLOJET Warm-up:** Vegas integration optimisation
  - 1 long (multi-core) job per process (→ bwUniCluster at KIT) — O(100 h)
- **3. Interpolation Warm-up:** Adapt interpolation grids to phase space
  - Only phase space provided from NNLOJET → significant speed-up — O(100 h)
- **4. Interpolation grid production:**
  - Thousands of parallel jobs (→ bwForCluster NEMO at Freiburg) — O(250 kh)
- **5. Postprocessing:** Statistical evaluation and combination of all produced grids …
  - Combination scripts/programs — O(100 h)
- **6. Present final results** — 20 min :-)

Step 4: Mass Production

- **NNLOJET + APPLfast**
  - **Massive parallelised computing on Virtual Machines with 24h lifetime**

Job Type	# Jobs	Events / Job	Runtime / Job	# Events	Total Output	Total Runtime
LO	10	140 M	20.6 h	1.4 G	24 MB	206 h
NLO-R	200	6 M	19.0 h	1.2 G	1.3 GB	3800 h
NLO-V	200	5 M	21.2 h	1.0 G	1.2 GB	4240 h
NNLO-RRa	5000	60 k	22.5 h	0.3 G	26 GB	112500 h
NNLO-RRb	5000	40 k	20.3 h	0.2 G	27 GB	101500 h
NNLO-RV	1000	200 k	19.8 h	0.2 G	6.4 GB	19800 h
NNLO-VV	300	4 M	20.5 h	1.2 G	2.0 GB	6150 h
Total	11710	---	---	5.5 G	64 GB	248196 h

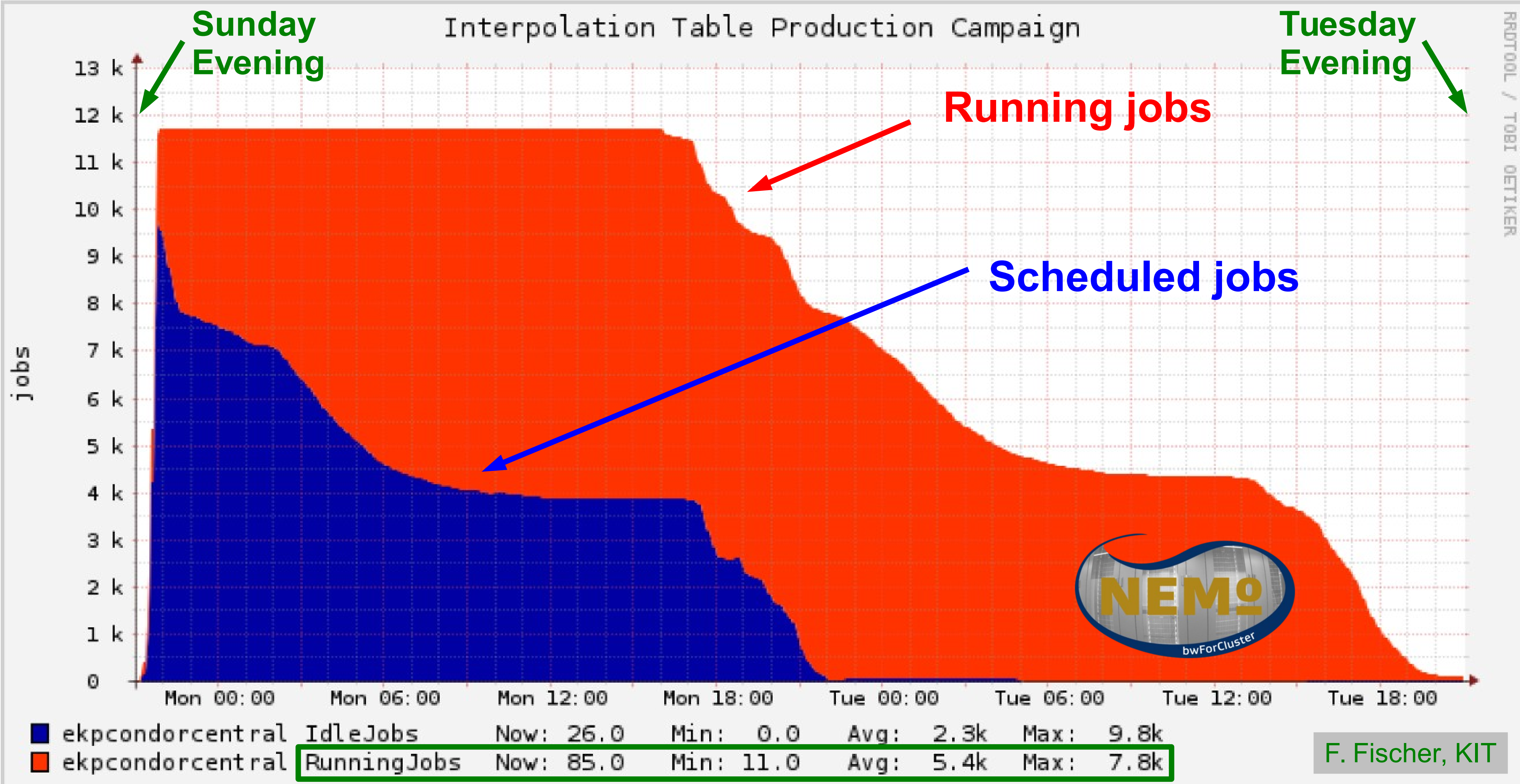
Production output: O(100 GB)
Final output: O(100 MB)

CPU time: O(250 kh)

Calculated on BwForCluster NEMO in Freiburg

Production Campaign

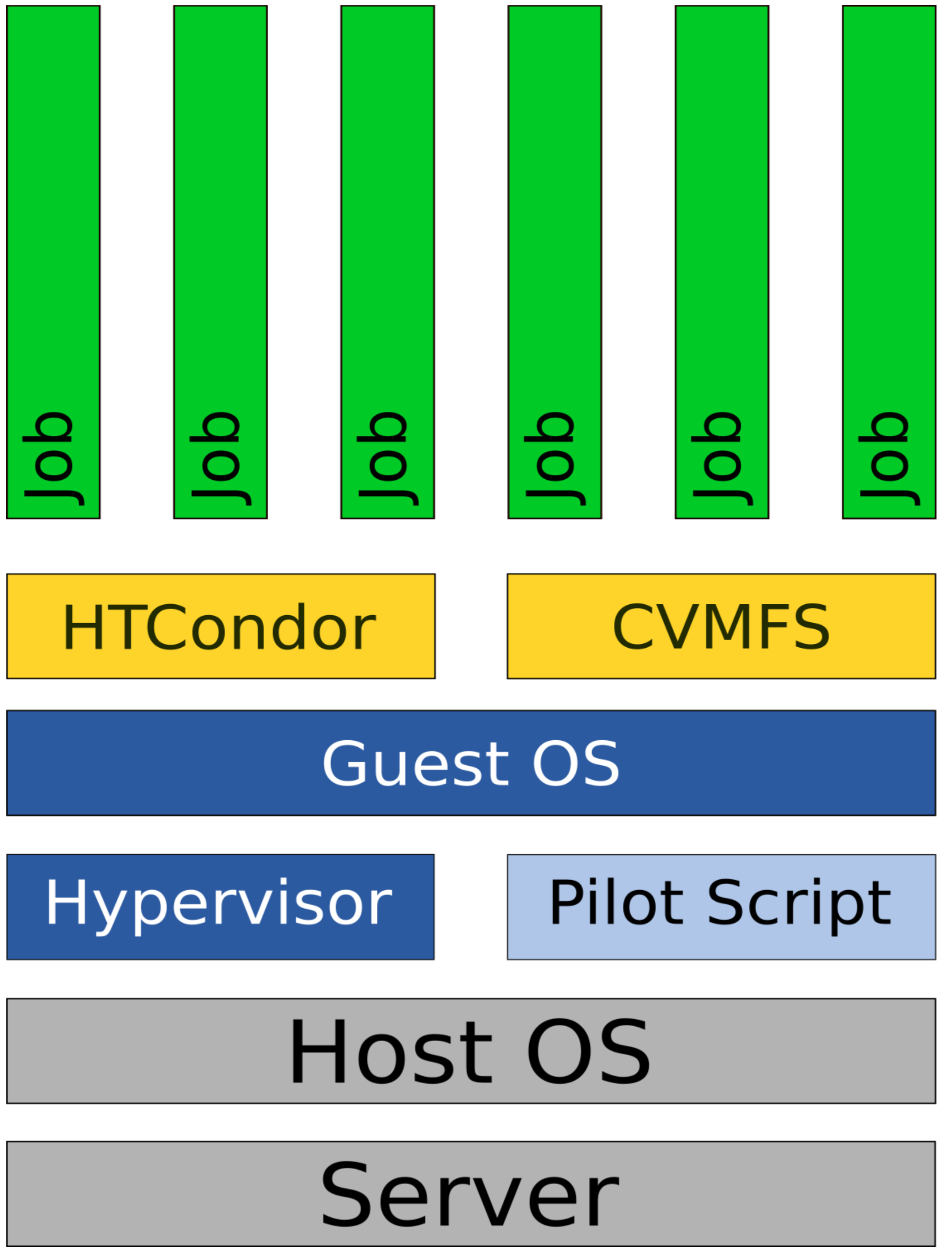
Optimised scenario: Finished in two days with 7800 parallel jobs at maximum!



Software Environment

Many specific software packages needed!

Require specific OS



Access software via CERN Virtual Machine File System (CVMFS)

Run on Scientific Linux 6 (SLC6 based on RedHat 6)

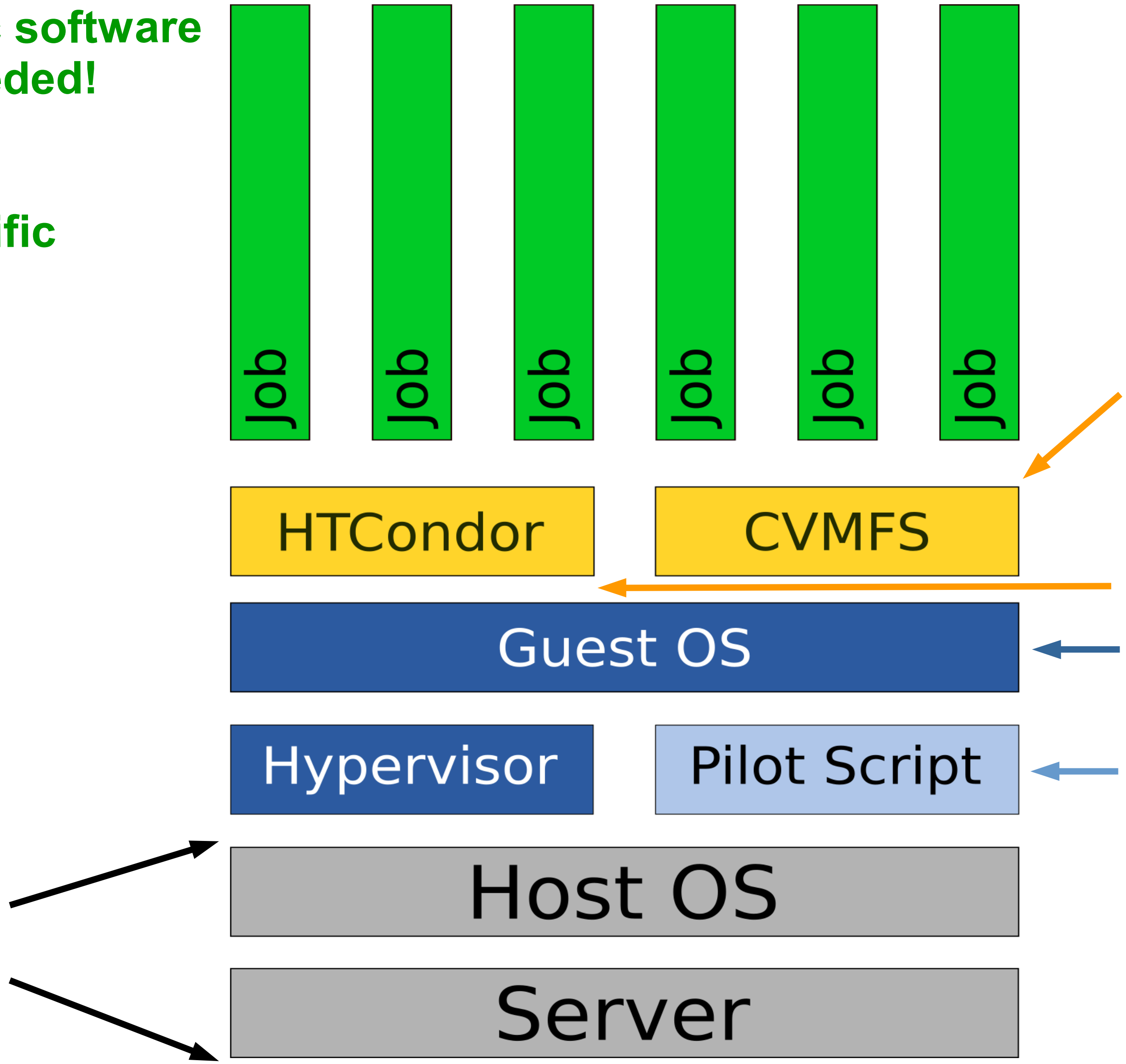
Solution:
Virtualisation

Available on HPC centre

Software Environment

Many specific software packages needed!

Require specific OS



Jobs access software via CVMFS

Batch system (HTCondor) starts job in VM

Compatible guest OS (SLC6)

Request VM & reserve resources

Available on HPC centre

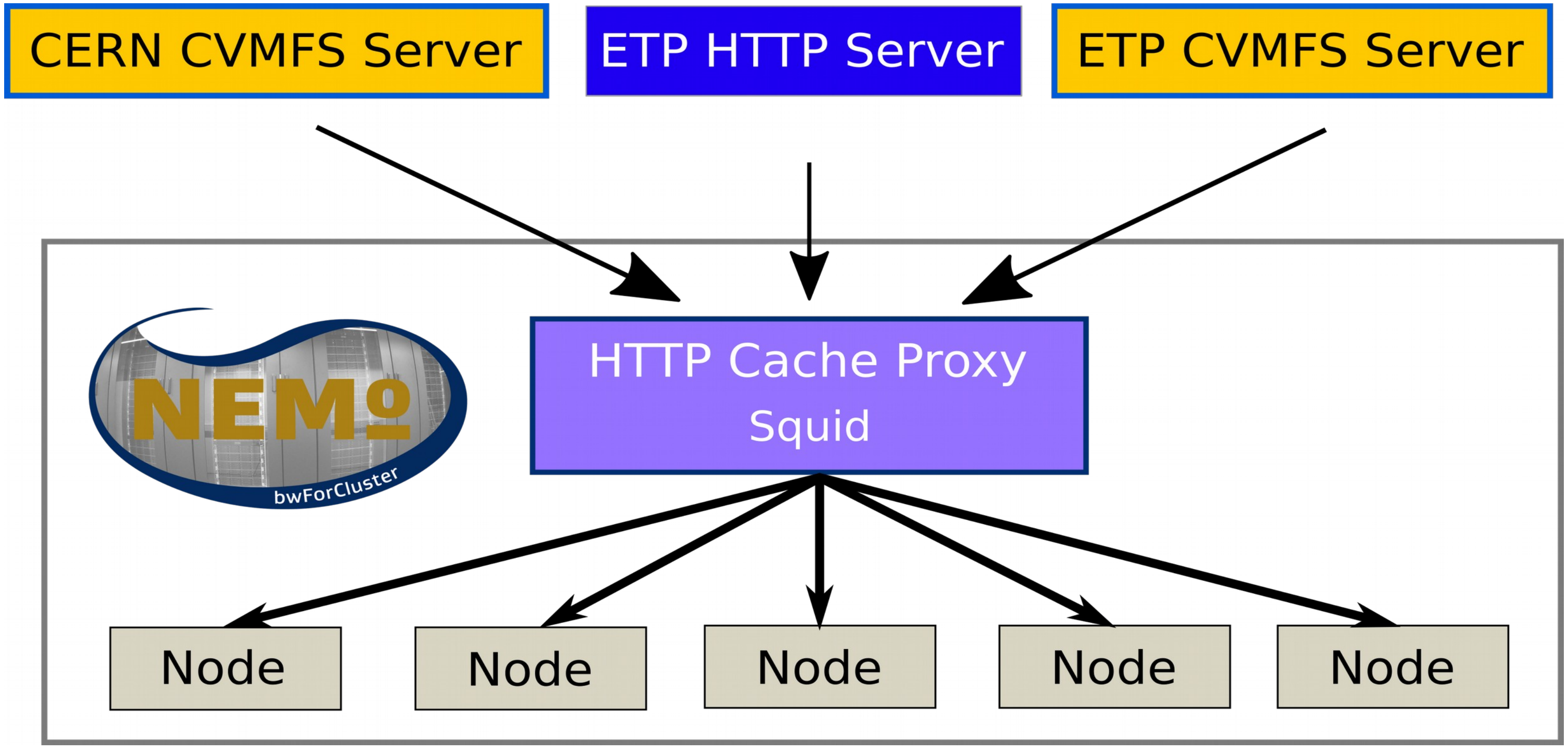
Software Distribution

From CERN:

- Read-only file system
- HTTP based protocol
- Proxy caches files from server

Provided by us (ETP):

- ETP CVMFS Server to provide our own software
- ETP HTTP Server for files with short life-cycle



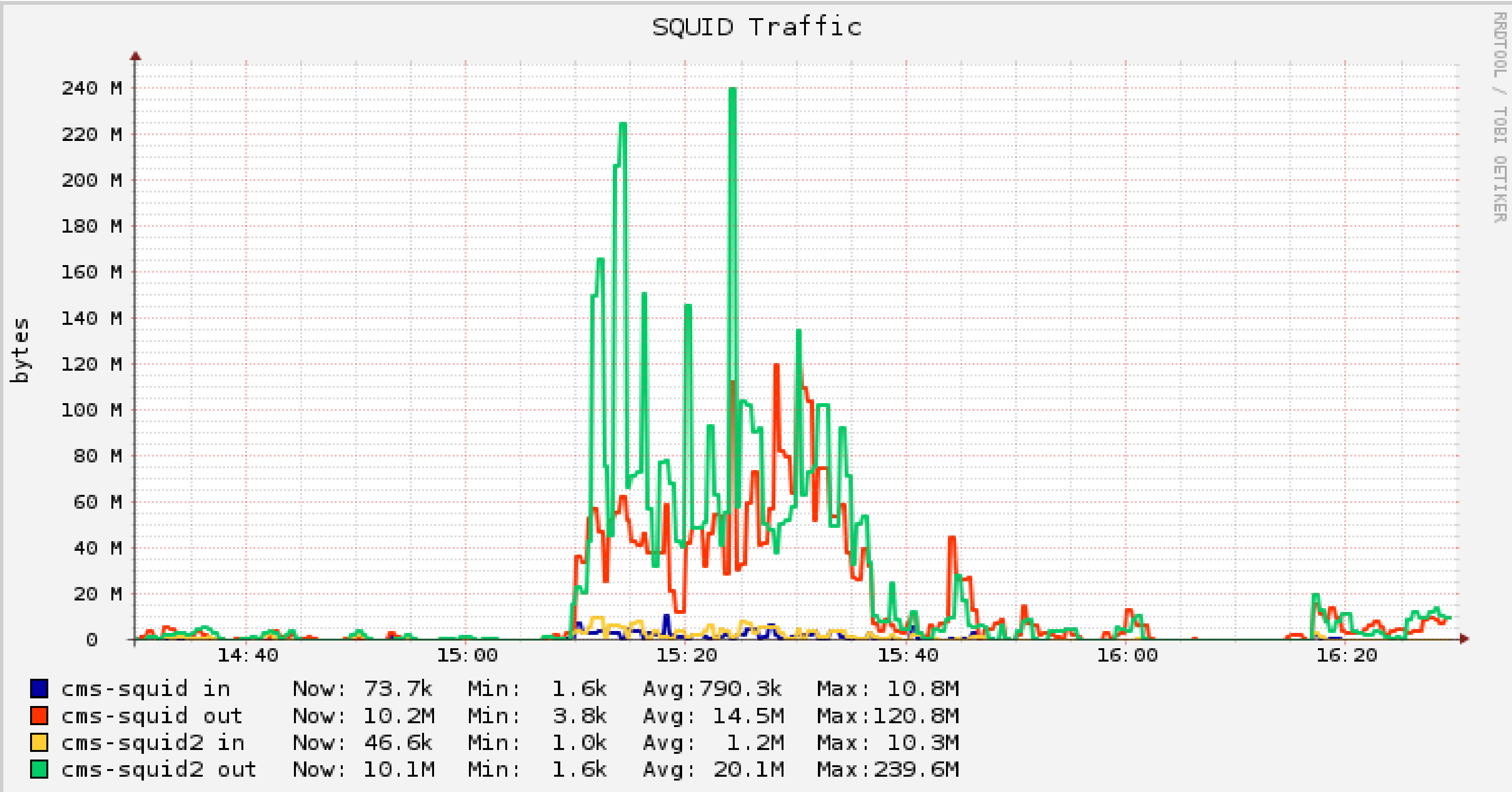
Software Distribution

From CERN:

- Read-only file system
- HTTP based protocol
- Proxy caches files from server

Provided by us (ETP):

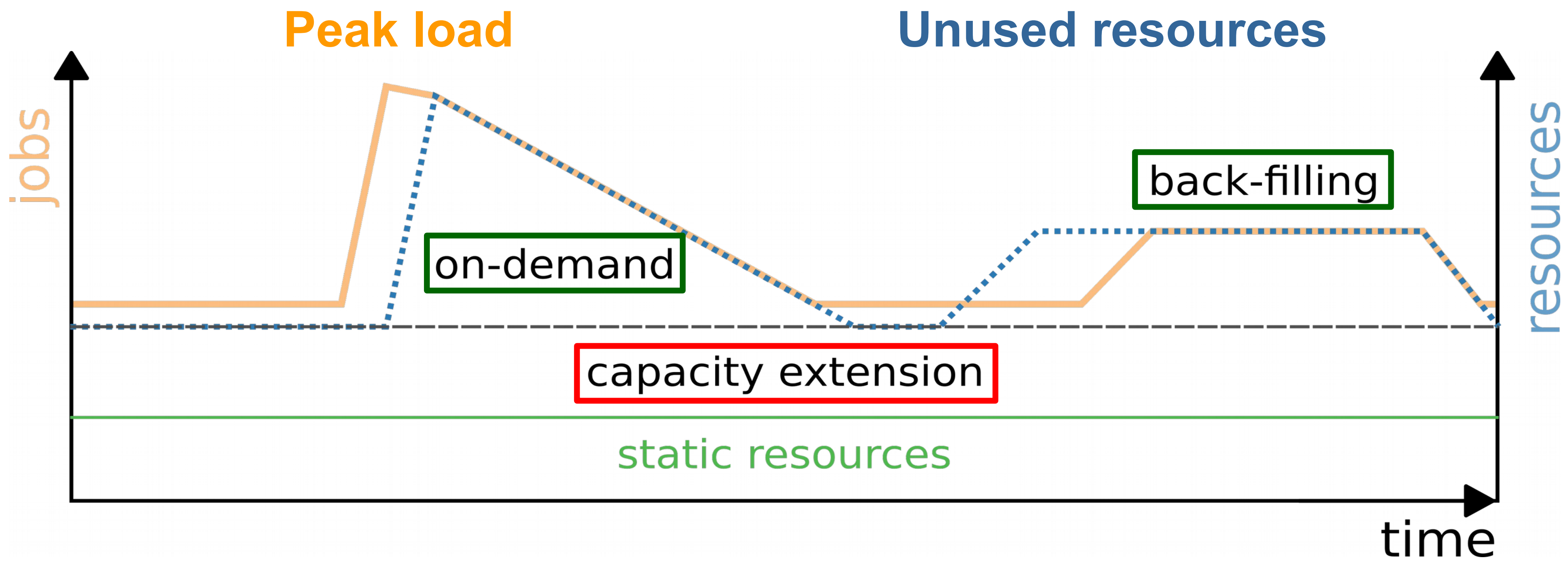
- ETP CVMFS Server to provide our own software
- ETP HTTP Server for files with short life-cycle



Drastic reduction of incoming traffic

Resource Provisioning Models

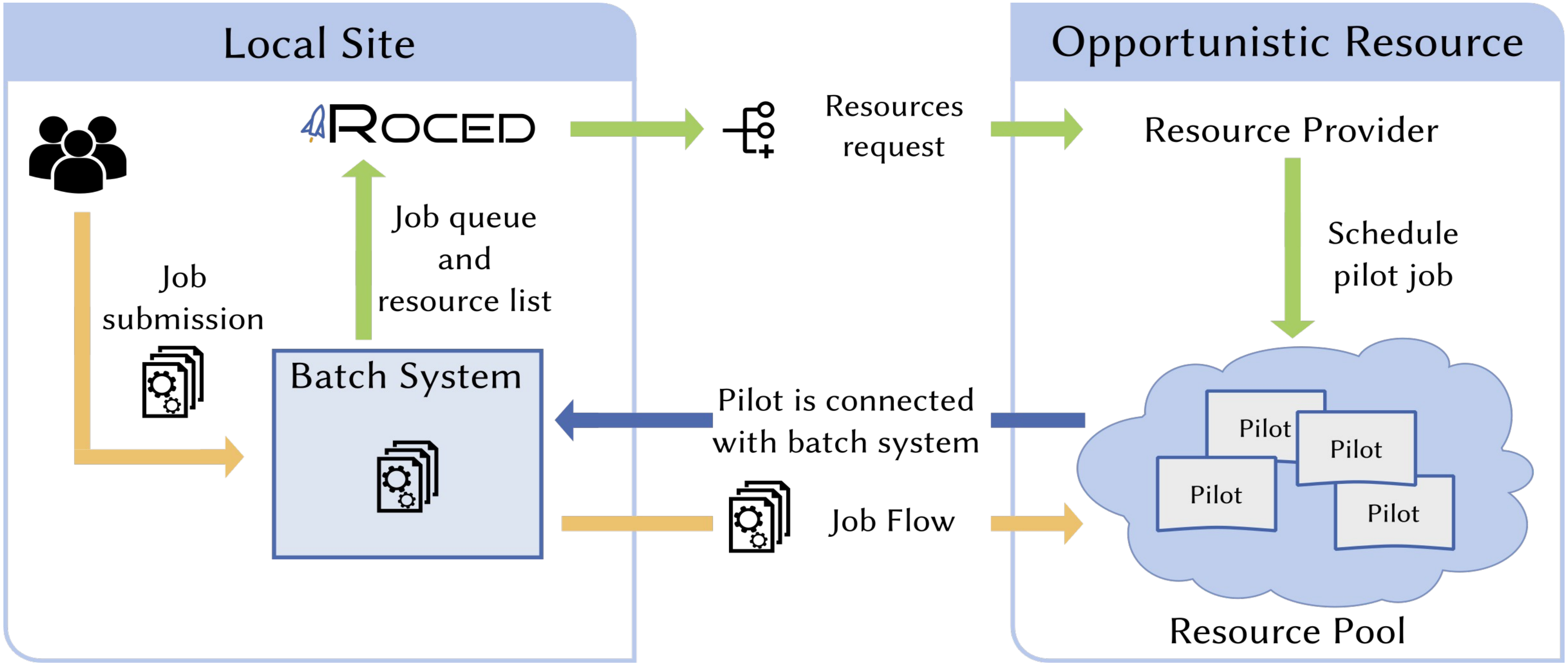
- Resource dependent provisioning of opportunistic resources
  - Simple but unflexible: Constant capacity extension
  - On-demand booking for job peak loads
  - Back-filling of unused resources
- Resource scheduler for dynamic resource usage and controlling



Cost-efficient concurrent resource usage

Resource Scheduler: ROCED

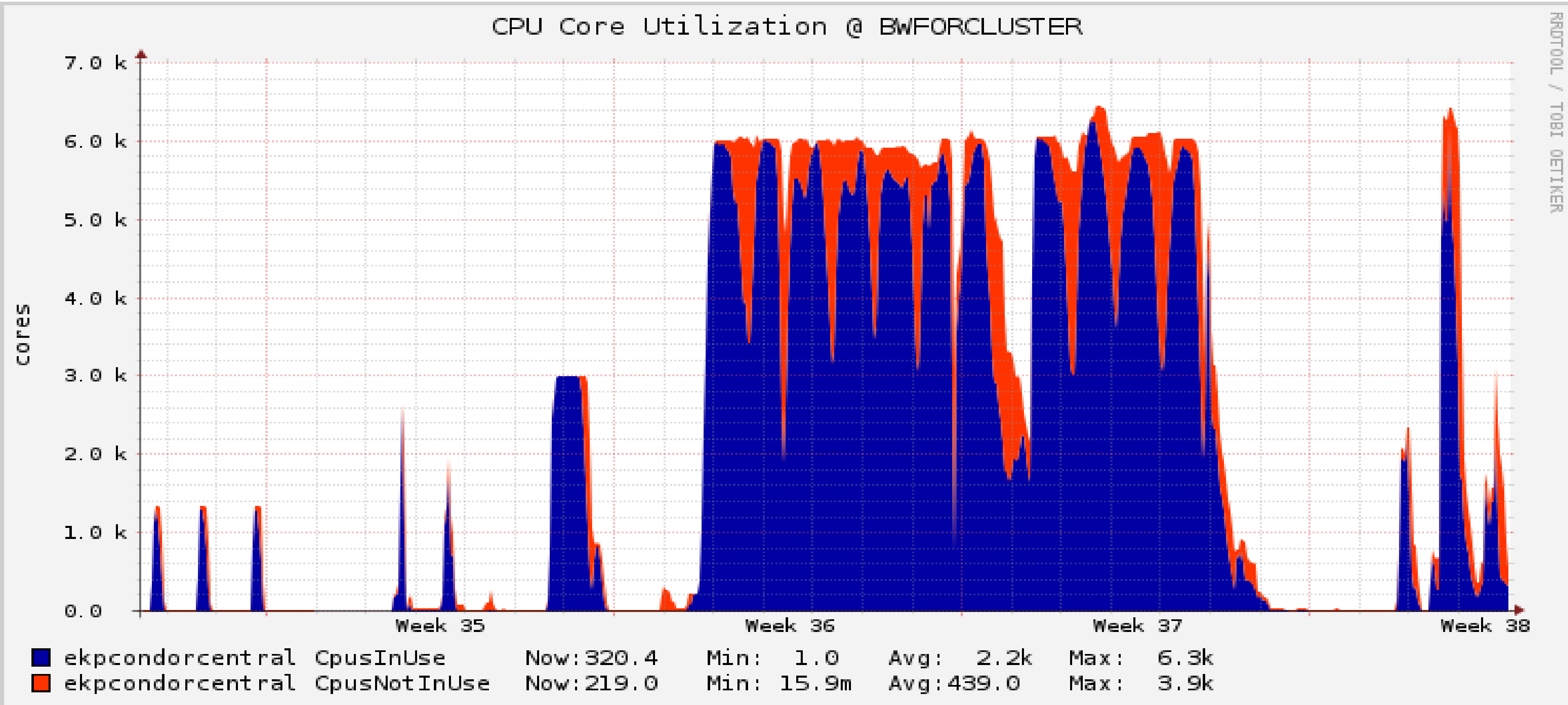
- **Lightweight management solution developed at KIT**
- **Support for multiple batch systems and resource providers**
- https://github.com/roced-scheduler/ROCED



Responsive On-Demand Cloud-enabled Deployment (ROCED)

ROCED at Work on NEMO

- Dynamic on-demand provisioning of VMs
- Integration into local batch system
- Scalability up to 8k cores demonstrated



Summary

- Precise but CPU-intensive predictions available for particle collisions
- Interpolation grids like fastNLO enable their use in comparisons to data to determine fundamental parameters of nature
- Virtualisation and ROCED resource scheduler permit efficient and concurrent usage of opportunistic HPC resources
- Large-scale productions successfully tested on bwForCluster NEMO
- Further improvements are in development
- Ony possible thanks to the Baden-Württemberg HPC support. Thank you!

Thank you for your attention!

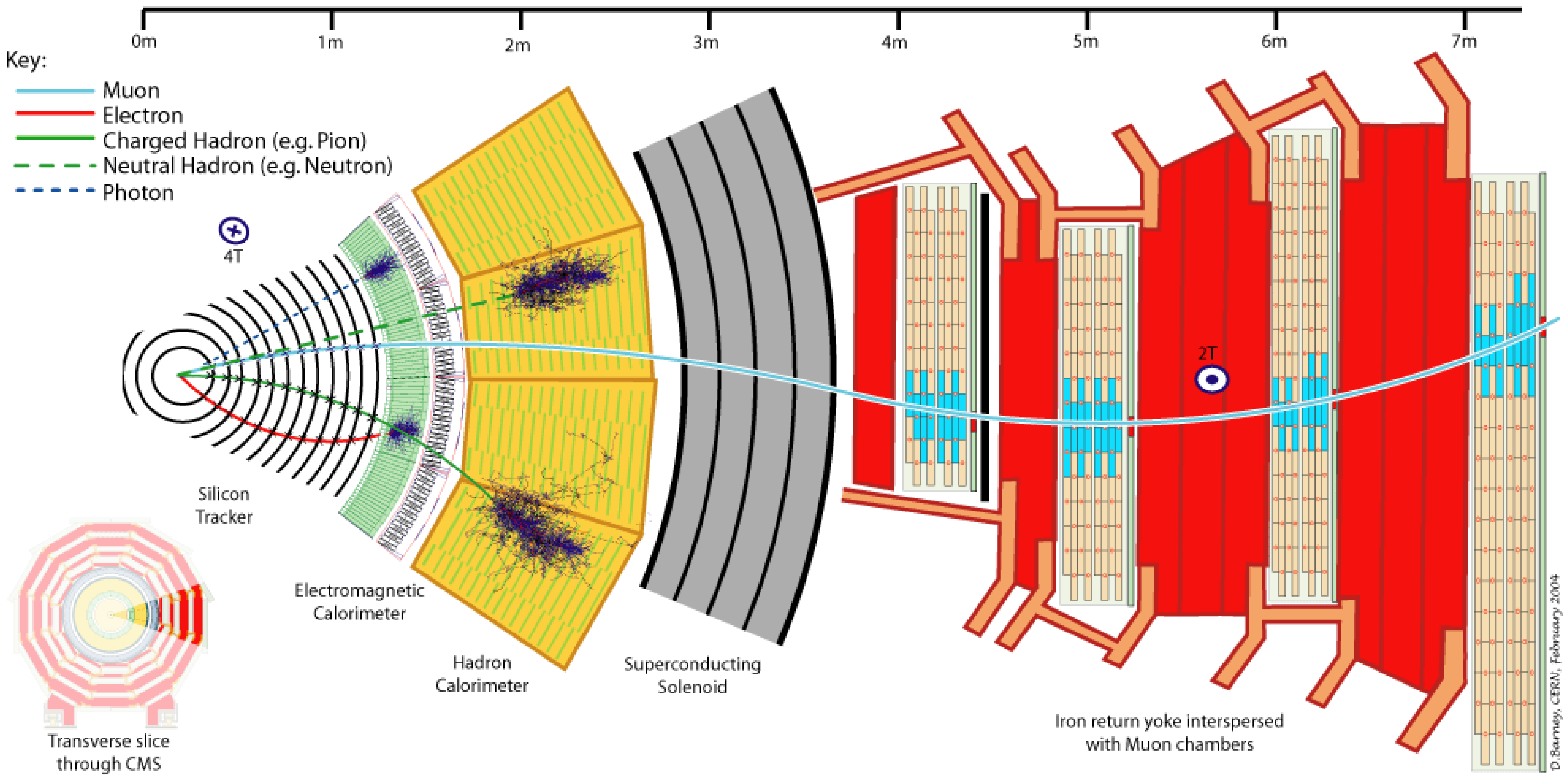
Summary

- Precise but CPU-intensive predictions available for particle collisions
- Interpolation grids like fastNLO enable their use in comparisons to data to determine fundamental parameters of nature
- Virtualisation and ROCED resource scheduler permit efficient and concurrent usage of opportunistic HPC resources
- Large-scale productions successfully tested on bwForCluster NEMO
- Further improvements are in development
- Ony possible thanks to the Baden-Württemberg HPC support. Thank you!

Cordial thanks go to the organisers for the invitation, to my collaborators from NNLOJET, fastNLO & APPLgrid, to the colleagues from bwForCluster at Freiburg and bwUniCluster at KIT for the fantastic support, and to my local colleagues at ETP for all their help and contributions to this talk!

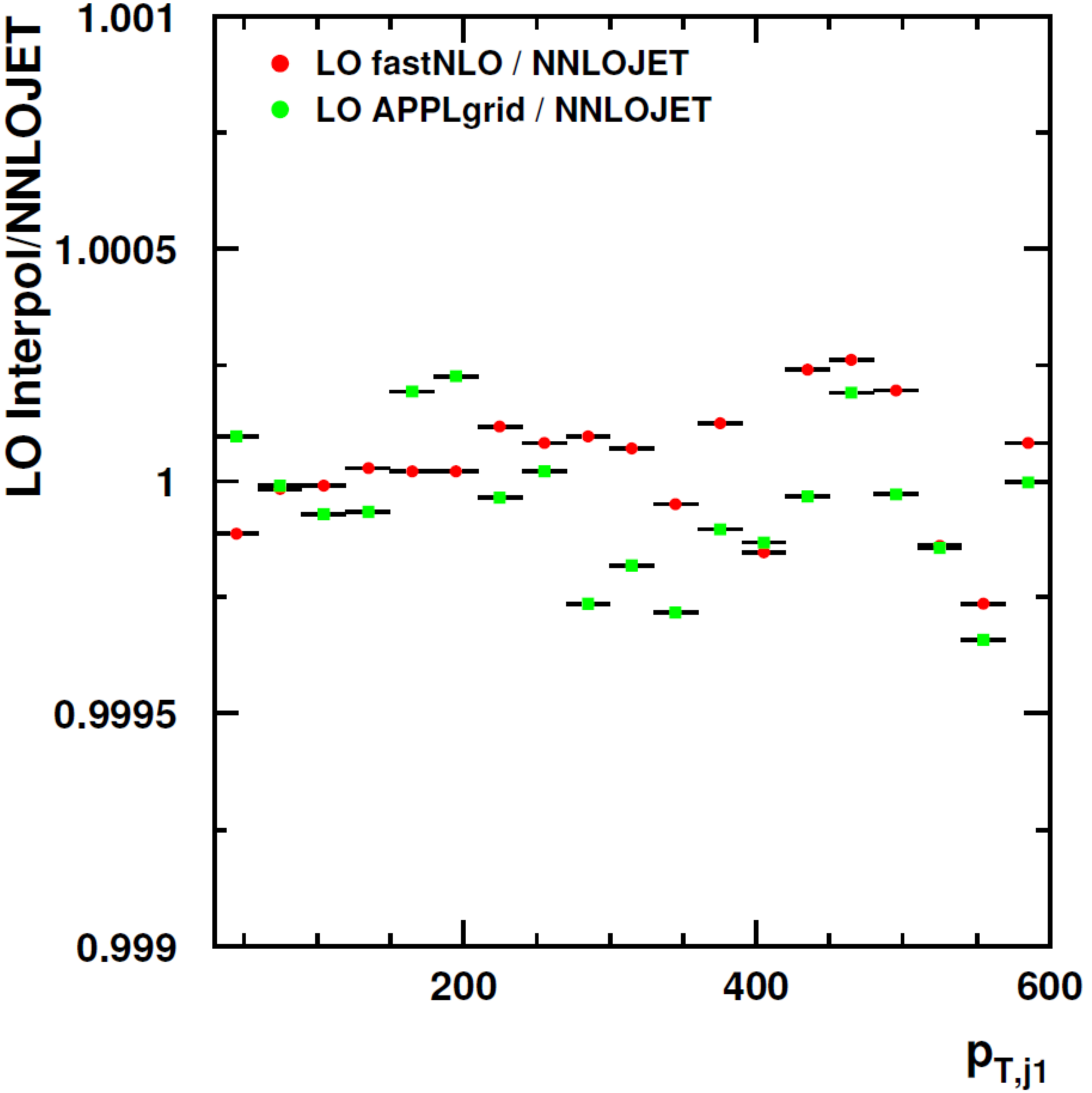
Backup

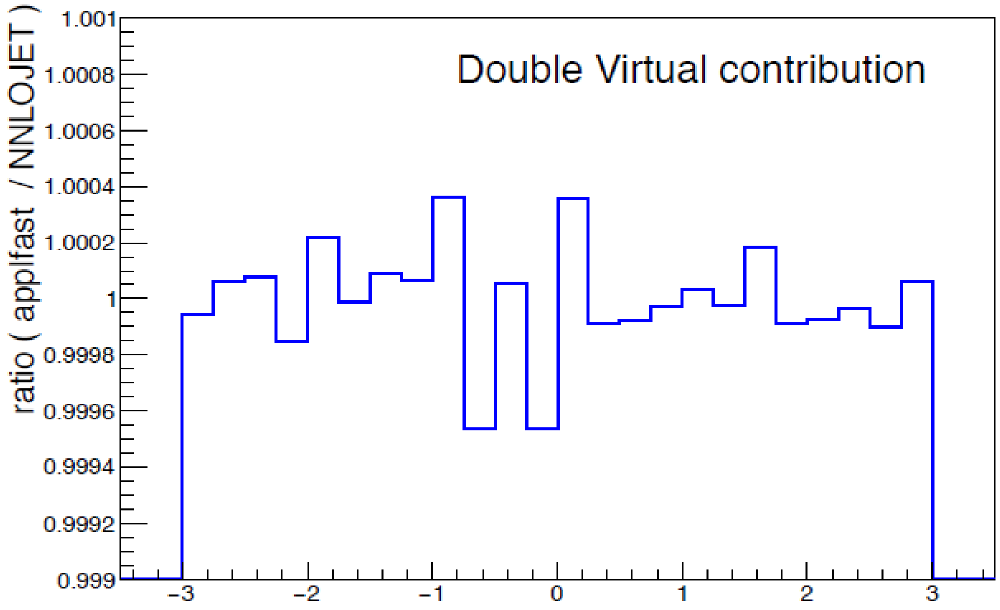
Particle Detection



Step 1: Preprocessing

Z+jet LO Approximation Test
Similar performance at subpermille level





M. Sutton at QCD@LHC

Z+jet Test Setup: $p_{T,j1}$, η_{j1}, y_Z

- E_{cms} = 8 TeV
- $p_{T,jet}$ > 30 GeV
- $|y_{jet}| < 3$
- $|y_{l+,l-}| < 5$
- 80 < M_{l+l-} < 100 GeV
- $\mu_r = \mu_f = M_Z$ = fixed
  (→ no scale interpolation in this test)

Step 2: Vegas Integrations

- **NNLOJET Warm-up:**
  - **Must be one job per process type**
  - **Multi-threading possible**

Job Type	# Jobs	Threads / Job	Events / Job	Runtime / Job	Total Runtime
LO	1	16	32 M	0.35 h	0.35 h
NLO-R	1	16	16 M	1.0 h	1.0 h
NLO-V	1	16	16 M	1.0 h	1.0 h
NNLO-RRa	1	32	5 M	17.5 h	17.5 h
NNLO-RRb	1	32	5 M	20.7 h	20.7 h
NNLO-RV	1	16	8 M	22.4 h	22.4 h
NNLO-VV	1	16	8 M	24.6 h	24.6 h
Total	7	-	-	-	87.6 h

Calculated on BwUniCluster at KIT thanks to Baden-Württemberg High Performance Computing (HPC) support

Step 3: Phase Space Exploration

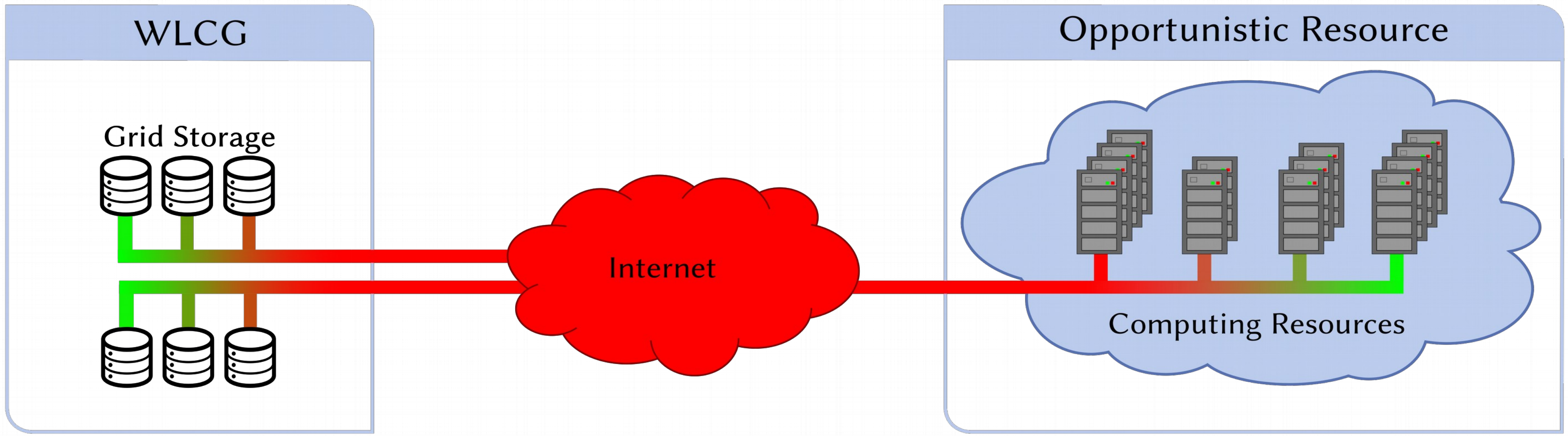
- **APPLfast Warm-up:**
  - NNLOJET is run without CPU-time expensive weight calculation
  - At least 1 job per process needed to determine phase space limits individually
  - Grids created and optimised during warm-up (APPLgrid)
  - Grids created in production step from optimised x and Q-scale limits (fastNLO)
  - Warm-up can be parallelised, if necessary (fastNLO)
  - Presented table used for extensive testing; overkill for normal use

Job Type	# Jobs	Events / Job	Runtime / Job	# Events	Total Runtime
LO	5	500 M	12 h	2.5 G	60 h
NLO-R	5	300 M	18 h	1.5 G	90 h
NLO-V	5	500 M	13 h	2.5 G	65 h
NNLO-RRa	10	50 M	13 h	0.5 G	130 h
NNLO-RRb	10	50 M	15 h	0.5 G	150 h
NNLO-RV	5	300 M	19 h	1.5 G	90 h
NNLO-VV	5	500 M	12 h	2.5 G	60 h
Total	45	---	---	11.5 G	645 h

In this setup most x_{min} limits from LO runs, 3 from higher-order runs.

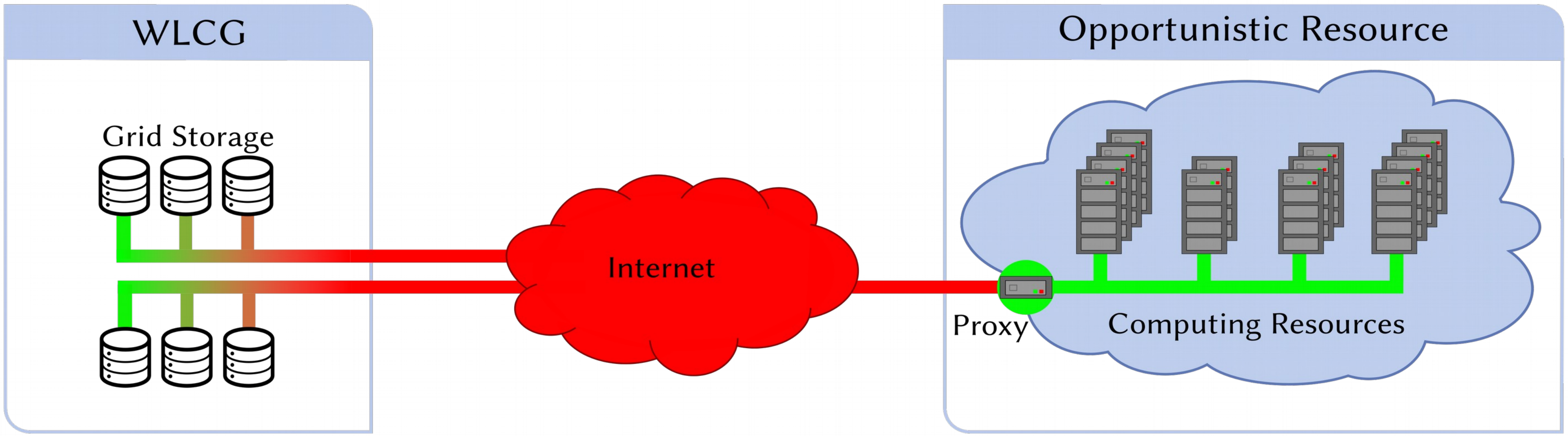
Challenges for Data Intensive Jobs

- Persistent storage only located at specific sites
- Storage performance usually designed for one Grid site
- Network shared as opportunistic resource
- Variable utilisation of storage and network



Challenges for Data Intensive Jobs

- Persistent storage only located at specific sites
- Storage performance usually designed for one Grid site
- Network shared as opportunistic resource
- Variable utilisation of storage and network



Proxy required for file transfer protocol to reduce incoming traffic
Only possible thanks to 20 Gbit link KIT-Freiburg