

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Schlussbericht

Verbund: 05H2021 - Run 3 von LHCb am LHC

Zuwendungsempfänger: Rheinische Friedrich-Wilhelms-Universität Bonn
Projektleitung: Prof. Dr. Sebastian Neubert
E-Mail: neubert@hiskp.uni-bonn.de
Förderkennzeichen: 05H21PDCL1
Förderzeitraum: 01.07.2021 - 30.06.2024
Zuwendung: 308.531,51 €
Projektträger: Projektträger DESY

Zusätzlicher Kontakt: Wallraf@verwaltung.uni-bonn.de
Zusätzlicher Name: Christine Wallraf

Genutzte Großgeräte:	Labor	Gerät	Experiment
	CERN	LHC	LHCb
Diplomarbeiten:	0		
Dissertationen:	1		
Habilitationen:	0		
Referierte Publikationen:	3		
Andere Veröffentlichungen:	0		
Patente:	0		
Bachelorarbeiten:	5		
Masterarbeiten:	7		
Staatsexamen:	0		

Dieser Bericht wurde beim Projektträger über einen individuellen Online-Zugang vom Projektleiter eingereicht und am 20.01.2025 09:54 für eine Veröffentlichung freigegeben.

Schlussbericht

Zuwendungsempfänger: Rheinische Friedrich-Wilhelms Universität Bonn

Projektleitung: Prof. Dr. Sebastian Neubert

Verbund: Verbundprojekt 05H2021 (ErUM-FSP T04)

Thema: Run 3 von LHCb am LHC: Betrieb der Realtime Analyse und Datenmanagement

Zusammenfassung

Das LHCb Experiment am Large Hadron Collider am CERN untersucht die Physik der Starken und Schwachen Wechselwirkungen mit dem Ziel unser Verständnis für den Aufbau der Materie und Ihrer Wechselwirkungen im Bereich der Flavorphysik zu vertiefen. Hierfür wurde der Detektor in den vergangenen Jahren mit neuen Subsystemen ausgestattet. Eine kritische und innovative Komponente dieses Upgrades ist das vollständig in Software implementierte Triggersystem, genannt Realtime Analysis RTA. Die Bonner LHCb Gruppe beteiligt sich mit Arbeiten zur Datenverarbeitung an der Inbetriebnahme dieses neuen Systems. Hierzu wurden Interfaces entwickelt um den Zugriff von Forschenden auf die im Experiment gesammelten Daten effizient und flexibel zu gestalten. Im Rahmen des Projektes wurde ein neuer zukunftsweisender Ansatz zur Verfügbarmachung der Daten im Open Data Format entwickelt und erste Prototypen dieses Systems implementiert.

Bericht

1 Aufgabenstellung und Voraussetzungen, unter denen das Vorhaben durchgeführt wurde

Das LHCb Experiment am Large Hadron Collider (LHC) am CERN in Genf hat zum Ziel die fundamentalen Wechselwirkungen zwischen Materieteilchen zu erforschen und nach neuen Phänomenen auf kleinsten Skalen zu suchen. Das Experiment ist zu diesem Zweck optimiert auf die Detektion von Hadronen mit schweren Quarks, die in Proton-Proton Kollisionen am LHC erzeugt werden. Die Zerfälle von Beauty und Charm Hadronen werden in einem Magnetischen Spektrometer vermessen, welches den Rapiditäts-Bereich von $|\eta|=2-5$ abdeckt, und mit Subdetektoren zur Spurrekonstruktion geladener Teilchen, Kalorimetrie, Teilchenidentifikation und einem Müonen-System instrumentiert ist. Das Experiment nimmt seit 2010 Daten.

Das Physik-Program von LHCb umfasst Präzisionsmessung in der Flavor-Physik, die Bestimmung von Parametern des Standardmodells der Elementarteilchenphysik mit höchster Präzision, die Suche nach seltenen Zerfällen und neuen bisher unbekanntem Teilchen, das Studium von Asymmetrien im Flavorsektor des Standardmodells und Untersuchungen der starken Wechselwirkung, insbesondere die Spektroskopie exotischer Hadronen im Charm und Beauty Sektor. Die Gruppe in Bonn konzentriert sich in ihrer Arbeit auf letzteres Forschungsfeld. Sie ist Teil der internationalen LHCb Kollaboration, welche zum Berichtszeitpunkt 1695 Forschende von 98 Instituten aus 24 Ländern umfasst.

Im vorliegenden Projekt beteiligte sich die Bonner LHCb Gruppe an der Entwicklung von Software-Paketen, die der Verarbeitung der gesammelten Daten im Real Time Analysis (RTA) System und den nachfolgenden Datenaufbereitungsschritten dienen. Die Gruppe war Ende 2019 durch die Berufung Prof. Neuberts auf eine W2 TT Professur nach Bonn neu etabliert worden. Es handelt sich um den ersten Beitrag der Gruppe im LHCb Forschungsverbund.

2 Wissenschaftlicher und technischer Stand, an den angeknüpft wurde

In den Jahren 2019-2022 wurde der LHCb Detektor im Rahmen eines Upgrades mit neuen Komponenten ausgestattet und dann für den Run 3 des LHC in 2022 wieder in Betrieb genommen. Dieses Upgrade ermöglicht es, Daten bei einer höheren instantanen Luminosität aufzuzeichnen. Eine entscheidende Komponente der neuen Detektors ist der erstmals vollständig in Software implementierte Trigger, der es erlaubt komplexe Physik-Signaturen in Echtzeit (daher der Name „Realtime Analyse“ RTA) zu rekonstruieren und zu selektieren. Dadurch gewinnt man gegenüber traditionellen Triggersystemen deutlich an Flexibilität und Effizienz bei der Suche nach komplexen Physik-Signaturen, welche essentiell für das Physik-Program sind. So erlaubt der RTA Ansatz zum Beispiel die Echtzeit-Suche nach CKM-unterdrückten Zerfällen.

Der Software Trigger wurde im Zuge des Upgrade Projektes in weiten Teilen neu entwickelt und implementiert. Auch die sich anschließende Datenverarbeitungskette musste dem neuen Konzept angepasst werden. Die Arbeiten der Bonner Gruppe gliedern sich in diese Entwicklungen ein.

3 Planung und Ablauf des Vorhabens sowie Kooperation mit Dritten

Die Beiträge zur Datenverarbeitungssoftware wurden in enger Zusammenarbeit mit den entsprechenden Projektarbeitsgruppen in der LHCb Kollaboration, insbesondere das RTA Projekt und die DPA (Data processing and analysis) Arbeitsgruppe geplant und durchgeführt. Die Projektmitarbeitenden aus Bonn nahmen an wöchentlichen virtuellen Treffen der LHCb Entwicklerteams teil und präsentierten ihren Arbeitsfortschritt auf Kollaborationstreffen.

Gefördert wurde das Arbeitspaket A2 „Run 3 von LHCb am LHC: Betrieb der Realtime

Analyse und Datenmanagement“. Dieses Paket war weiter in vier Aktivitäten unterteilt. In A2.1 wurde ein Daten Interface zwischen dem RTA System und den Analyse-Teams entwickelt. Ziel war es hier eine wohldefinierte Schnittstelle und benutzerfreundliche Werkzeuge zu schaffen, die es erlauben auf die vom RTA System produzierten Daten mit Hilfe der zentralen LHCb Computing Infrastruktur zuzugreifen. Die Gruppe beteiligte sich an der Entwicklung des entsprechenden „Analysis Production“ Systems. Dieses System bietet die Möglichkeit Datenabfragen und die Produktion von gefilterten Ntuplen in zentral gemanagten Produktionen auf dem WLCG laufen zu lassen. Diese Zentralisierung des Datenzugriffs ermöglicht erst die effiziente Analyse der großen Datenmengen, die in Run 3 erwartet werden.

Früh im Projekt stellte sich heraus, dass aufbauend auf den Entwicklungen zu den Analysis Productions weitere Anwendungen möglich werden. Mit Hinblick auf zukünftige Veröffentlichungen der LHCb Daten im Open Data Verfahren wurde mit dem Ntuple-Wizard ein für die Teilchenphysik völlig neues Konzept des Datenzugriffs entwickelt. In 2023 wurde das Konzept-Papier zum Ntuple-Wizard veröffentlicht [Comput Softw Big Sci 7(2023)6], welches in der Open Science Community auf grosses Interesse gestossen ist und zum Beispiel auf der internationalen Konferenz CHEP 2023 in Norfolk, Virginia, USA vorgestellt wurde. Der Ntuple Wizard eröffnet einen neuen Weg die LHCb Daten im Open Data Verfahren zur Verfügung zu stellen, welcher die Erstellung eigener Open Data Replikas der Daten – und den damit verbundenen zusätzlich benötigten Speicherplatz – vermeidet. Zudem hat das Verfahren den Vorteil dass externen Nutzern der Daten ein User-Interface zum Erstellen von Daten-Anfragen zur Verfügung gestellt wird, welches die Daten erst im eigentlichen Sinne für Nicht-Experten erschliesst. Dieses Web-Interface wurde im Laufe des Projektes entwickelt und ist inzwischen in einer Beta-Testphase in Betrieb genommen worden. Auch eine Integration des Ntuple-Wizard mit dem CERN Open Data Portal konnte im Laufe des Projektes begonnen werden. Mit Projektabschluss ging dieses System in eine Alpha-Testphase und ist nun als Werkzeug für den Release der LHCb Daten aus Run 2 vorgesehen. Diese Entwicklung erfolgte in enger Kooperation mit der LHCb Gruppe aus Michigan und dem CERN Open Data Team.

In der Aktivität A2.2 RTA Qualitätssicherung wurde die Infrastruktur zur automatischen Generierung der Software-Dokumentation mit dem Doxygen Tool wieder instand gesetzt. Diese Dokumentation ist kritisch für die effiziente Arbeit am einem so komplexen Software Projekt wie dem RTA und der LHCb Datenverarbeitungssoftware. Hier war ursprünglich noch ein System zur Automatischen Dokumentation der Trigger-Selektionen vorgesehen. Dieses Feature wurde aber durch die schnelle Evolution des RTA Software-Frameworks im Laufe des Projektes mehrfach zurück gestellt und konnte im Projektzeitraum parallel zur Inbetriebnahme des RTA Systems nicht mehr umgesetzt werden. Die frei werdenden Ressourcen kamen den anderen Aktivitäten im Arbeitspaket zu gute.

In A2.3 wurde mit dem sog. Sprucing ein System in Betrieb genommen, welches die RTA Triggersoftware im Offline-Modus verwendet um bereits aufgezeichnete Daten nocheinmal zu analysieren. Dieses Konzept ermöglicht es neue Selektionen zu verwirklichen, die zum Zeitpunkt der Datennahme noch nicht angedacht waren. Zunächst arbeitete die Gruppe am Monitoring-System, mit dem die Sprucing Prozesse während der Datenverarbeitung überwacht werden. Der Schwerpunkt verlagerte sich im Verlauf des Projektes vom Monitoring hin zu den File-Summary-Records FSR. Für die die Gruppe in 2023 die Verantwortung übernahm. Die FSR sind Meta-Daten, die die von der Trigger-Software generierten Dateien selbst-beschreibend machen. Hier war eine Neuentwicklung notwendig um der Evolution des RTA-Systems rechnung zu tragen. Ein funktionaler Prototyp der neuen FSR wurde in 2023 implementiert und erfolgreich getestet und ging in 2024 in Produktion. Dieses System wird im Nachfolgeprojekt fortgehend weiter entwickelt. Zum Berichtszeitpunkt wurden insgesamt 10 Sprucing Kampagnen mit der entwickelten Software durchgeführt.

In der Aktivität A2.4 Datenmanagement: Reproduzierbarkeit von LHCb Analyse-Workflows wurde ein Template entwickelt, welches die Automatisierung von LHCb Analyse-Workflows auf der REANA Infrastruktur ermöglicht. Das Template verwendet das SNAKEMAKE Workflow System, welches

inzwischen in LHCb zunehmend für die Enduser-Analysen in Verwendung ist. Zudem enthält das Template Docker-Files und die von REANA benötigten Meta-Daten, die als Ausgangspunkt dienen um Analyse-Workflows mit dem Analyse-Code und zugehörigem Software Stack reproduzierbar zu bewahren. Das Template steht der LHCb Kollaboration zur Verfügung. Es sei an dieser Stelle aber erwähnt, dass sich dieses Schema bisher nicht in der Praxis durchgesetzt hat. Als Hürde erweist sich hier der zusätzliche Arbeitsaufwand der Präservation, der im Rahmen typischer Doktorarbeiten kaum leistbar ist. Das Problemfeld geriet im Zuge der Inbetriebnahme des LHCb Ugrades etwas in den Hintergrund. Wir erwarten eine Wiederkehr dieses Themas, wenn die Datenanalyse der Run 3 Daten im vollem Schwung ist.

Eine Herausforderung zu Beginn des Projektes war ein unvohergesehener Personalmangel. Daher verzögerte sich der Start der Arbeit an den Aktivitäten A2.1-2.3 um ca ein Jahr. Dennoch konnten die wesentlichen Ziele des Projektes in der verbleibenden Zeit erreicht werden.

Neben den im Rahmen von ErUM geförderten Aktivitäten trug die Bonner Gruppe weiterhin durch die folgenden Beiträge zu LHCb bei. In 2022 wurde mit dem Analysis Lifecycle Management Tool ALCM, eine neue Datenbank zur Verwaltung der Arbeitsprozesse im Begutachtungs und Veröffentlichungsprozess von LHCb Publikationen in Betrieb genommen. Das System wurde von S.Neubert und F.Redl konzipiert und von einem Team aus Software-Ingenieuren am CERN implementiert. Alle LHCb Publikationen und öffentliche Plots nutzen heute dieses System. Im Dezember 2023 wurde der gesamte LHCb Datensatz aus Run 1 auf dem CERN Open Data Portal veröffentlicht. Diese Release wurde im Rahmen einer Doktorarbeit in Bonn vorbereitet und in Zusammenarbeit mit dem CERN IT Team durchgeführt. Die veröffentlichten Datensätze wurde detailliert mit allen Selektionscuts aus der Offlineselktion dokumentiert. Zu diesem Release gehört auch ein Glossar aus mehreren hundert Spezialbegriffen, die zum Verständnis der Dokumentation notwendig sind. Die Gruppe ist führend an den Vorbereitungen für ein zukünftiges Upgrade II beteiligt und hat sich als erste Gruppe in Deutschland dem MightyTracker Konsortium angeschlossen. In Bonn werden insbesondere Tests der HV-CMOS Sensoren aufgebaut, die für diesen zukünftigen Detektor vorgesehen sind. Dazu wurde eine dedizierte Ausleseelektronik entwickelt und am Bonner Isochronen-Zyklotron Strahlenhärte Tests durchgeführt. Ende Mai 2023 richtete die Gruppe den LHCb MightyTracker Workshop in Bonn aus. Dr Klaas Padeken übernahm die Koodination der MightyTracker Sensor AG und war Teil des Editing Teams für das Upgrade II Descoping Document.

Im Bereich Outreach entwickelte die Gruppe eine neue LHCb Masterclass mit dem Thema Hadronenspektroskopie. Dies geschah unter Einbeziehung von Lehramtsstudierenden. Als Teil dieser Arbeit entstand ein 3D-gedrucktes Quark Puzzle, welches das Konzept des Farbeinschlusses bei Hadronen begreifbar macht. Dieses Puzzle wurde unter anderem auf er GamesCom24 Messe präsentiert. Die 3D-Dateien zum selbsdruck wurden Open Source veröffentlicht.

Das Projekt wurde im Verbund ErUM-FSP T04 zusammen mit den Unis aus Heidelberg, Dortmund, Aachen und Rostock, sowie dem Max Planck Institut Heidelberg durchgeführt. Die Bonner Gruppe nahm regelmässig an den Veranstaltungen des Verbundes teil, darunter die FSP Treffen im Herbst (2021 in Dortmund, 2022 in Heidelberg,) und die Doktorandenworkshops in Meinerzhagen und Neckazimmern jährlich im Frühjahr. Die FSP Meetings in 2020 und in 2023 richtete unsere Gruppe in Bonn aus. Die gemeinsamen Forschungsaktivitäten fanden im wesentlichen in den LHCb Arbeitsgruppen statt, an denen Doktoranden und Postdocs wöchentlich via Videokonferenz teilnehmen. Auch die LHCb Kollaborationstreffen wurden zur Vernetzung genutzt. Intensive Zusammenarbeit im Verbund gab es auch bei der Vorbereitung der Ulgrade II Aktivitäten, so fanden mehrere Bestrahlungskampagnen am Bonner Zyklotron gemeinsam mit den Kollegne aus Heidelberg und Dortmund statt. Das Bonner Hardware Meeting fand unter Zuschaltung der Partner in Heidelberg

und Dortmund jeden Mittwoch statt. Auch Teststrahlzeiten am DESY wurden im Rahmen dieser Aktivitäten gemeinsam geplant und durchgeführt.

4 Verwendung der Zuwendung

Aus den bewilligten Mitteln wurden ab Oktober 2022 zwei Doktorandenstellen finanziert. Die hierfür benötigten Personalmittel stellen den Löwenanteil der Projektkosten dar. Dazu kamen Reisemittel mit denen Reisen zu Kollaborationstreffen oder zum Schichtdienst ans CERN finanziert wurden. Weitere Reisemittel wurden zur Teilnahme an Konferenzen und Workshops verwendet.

5 Erzielte Ergebnisse mit Gegenüberstellung der vereinbarten Ziele

Die wesentlichen Ergebnisse des Projektes sind

- Beiträge zum Dateninterface: das Analysis Productions System ist in Betrieb genommen worden. Dieses System fungiert als Standardwerkzeug für den Zugriff auf die Daten aus Run 3 (und vorheriger Runs).
- Der Ntuple-Wizard als benutzerfreundliches Interface wurde implementiert und befindet sich in einer Beta-Test Phase. Dieses innovative Werkzeug wurde im Verlaufe des Projekts neu konzipiert und geht über die ursprüngliche Zielsetzung des Projektes hinaus.
- Sprucing: Das System die RTA-Software offline über die Daten laufen zu lassen um neue Selektionen, die zum Zeitpunkt der Datennahme noch nicht implementiert waren, zu ermöglichen wurde in Betrieb genommen. Zum Zeitpunkt des Berichts sind bereits 10 Sprucing-Kampagnen mit diesem Tool durchgeführt worden.
- Ein neues File-Summary-Record Format wurde in der RTA Software implementiert und getestet. Die Notwendigkeit diese Meta-Daten zu überarbeiten ergab sich erst während des Projektes, gliedert sich aber gut an die Arbeiten zum Sprucing an.
- Ein Template zur Automatisierung von Analyseworkflows auf der REANA Infrastruktur des CERN wurde implementiert und steht der Kollaboration zur Verfügung.
- Die automatisch generierte Softwaredokumentation (Dokygen) für die LHCb Software wurde in Stand gesetzt.
- Im Projektzeitraum leitete die Bonner Gruppe maßgeblich die Arbeiten zur Veröffentlichung der LHCb Daten im Open Data Verfahren. Der komplette Datensatz aus Run 1 wurde im Dezember 2023 auf dem Open Data Portal des CERN der Öffentlichkeit bereit gestellt. Die Arbeiten für dieses Release wurden aus dem Emmy Noether Projekt des Projektleiters und Institutsmitteln finanziert.

Damit wurden die Ziele des Projektes erreicht. Einzig die automatische Dokumentation der RTA Selektionen konnte nicht im Projektzeitraum umgesetzt werden. Hier verhinderte eine schnelle Evolution des Selektionsframeworks unter dem Druck der Inbetriebnahme einen Abschluss. Diese Arbeiten werden im Nachfolgeprojekt angegangen.

6 Notwendigkeit und Angemessenheit der geleisteten Arbeit

Die Beiträge zur Datenverarbeitungssoftware, insbesondere die Arbeiten am Sprucing, an den Analysis Productions und dem Ntuple-Wizard sind notwendig für den reibungslosen Betrieb des LHCb Experiments in Run 3. Sie stellen integrale Bestandteile der Datenverarbeitungskette dar und stellen sicher, dass die vom RTA System aufgezeichneten Daten effizient und flexibel von den Analysten weiter verwendet werden können. Die zusätzlichen Arbeiten das Dateninterface auch für den Gebrauch im Open Data Verfahren vorzubereiten, waren dabei durch die starken Synergieeffekte vollkommen angemessen.

7 Voraussichtlicher Nutzen, insbesondere Verwertbarkeit der Ergebnisse

Alle Komponenten an denen die Bonner Gruppe in diesem Projekt mitgewirkt hat sind im aktuellen Betrieb des LHCb Experiments im Einsatz. In 2024 wurde eine integrierte Luminosität von 9.56 fb^{-1} aufgenommen. Dies entspricht etwas mehr als dem gesamten Datensatz aus allen Datennahme Perioden der Jahre 2011-2018. Die gesammelten Daten werden in den kommenden Jahren ausgewertet. Erste Untersuchungen zeigen, dass das RTA System und der damit verbundene Verzicht auf einen Hardware-Trigger tatsächlich eine Verbesserung der Triggereffizienz für Hadronische Zerfälle um ca einen Faktor drei erreicht. Verbunden mit der hohen Luminosität kann somit bei weiterem planmässigen Betrieb des LHC in Run 3 (bis 2026) vorraussichtlich der Umfang der Datensätze in diesen Prozessen um bis zu eine Größenordnung erhöht werden. Dadurch ergeben sich vielversprechende Analysemöglichkeiten über das gesammte Physikprogramm hinweg.

Die Integration des Ntuple-Wizards mit dem Open Data Portal des CERN wird es in Zukunft aufbauend auf dem Analysis-Productions System erlauben neue Wegen in der Bereitstellung von Datenprodukten im Open Data Verfahren zu gehen. Das Thema Open Data wurde im Verlaufe des Projektes als spannendes Zukunftsthema identifiziert und wird ebenso von den hier angestossenen Entwicklungen profitieren. Der Release der LHCb Daten aus Run 1 ist ein erster Meilenstein. In 2024 wurden die Grundlagen für den Release der Run 2 Daten gelegt, welcher den Ntuple-Wizard als zentrale Infrastruktur nutzen soll.

8 Während der Durchführung des Vorhabens dem Zuwendungsempfänger bekannt gewordenen Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen

Keine.

9 Erfolgte und geplante Veröffentlichungen der Ergebnisse

9.1 Referierte Publikationen (z. B. in Fachzeitschriften oder -büchern und referierte Konferenzproceedings)

Die LHCb Kollaboration hat im Berichtszeitraum ca 150 Fachartikel veröffentlicht. Hier berichten wir nur die für die Bonner Arbeitsgruppe entscheidenden Publikationen.

Roel Aaij et al. Observation of $\Lambda_b^0 \rightarrow \Lambda_c^+ D^{(*)0} K$ and $\Lambda_b^0 \rightarrow \Lambda_c^+ D_s^{*-}$ decays.
Eur. Phys. J. C,84(6):575, 2024.

Christine A. Aidala, Christopher Burr, Marco Cattaneo, Dillon S. Fitzgerald, Adam Morris, Sebastian Neubert, and Donijor Tropmann. Ntuple Wizard: An Application to Access Large-Scale Open Data from LHCb.
Comput. Softw. Big Sci., 7(1):6, 2023.

Hannah Schmitz, Lucas Dittmann, Sebastian Neubert, and Klaas Padeken. Development of

a new Readout System for HVCMOS Sensors for the LHCb Mighty Tracker.
Nucl. Instrum. Meth. A, 1065:169491, 2024.

9.2 Andere Veröffentlichungen (z. B. Konferenzbeiträge wie Vorträge und Poster, unreferierte Proceedings, Conference Notes)

Framework TDR for the LHCb Upgrade II: Opportunities in flavour physics, and beyond, in the HL-LHC era. 2021.

LHCb Upgrade II Scoping Document CERN-LHCC-2024-010 LHCb-TDR-026, 2024

CERN Open Science Policy. Technical report, CERN, Geneva, 2022.

LHCb Analysis Lifecycle Management CERN-LHCb-INT-2021-007

Klaas Padeken. The LHCb Mighty Tracker. PoS, Pixel2022:084, 2023

LHCb Run I data release <https://lhcb-outreach.web.cern.ch/2023/12/20/lhcb-releases-the-entire-run-i-dataset/>

9.3 Abschlussarbeiten (Bachelor, Master, Diplom, Staatsexamen, Promotion, Habilitation)

Doktorarbeit:

Mindaugas Sarpis Dissertation finished December 2023 "Pentaquark Search in $\Lambda_b^0 \rightarrow \Lambda_c^+ D^0 K$ Decays and LHCb Open Data Release"

Masterarbeiten:

Hannah Schmitz Tracking Performance Studies for a Future LHCb Tracking Detector, 13.10.21

Abbhay Meta Search for Pentaquark States in $\Lambda_b^0 \rightarrow \Sigma_c D^- K^-$ decays at LHCb, 14.9.21

Ellinor Eckstein Calibrating Multiparticle Decays using Normalizing Flows, 19.08.22

Piet Nogga First Observation of the Decay $B^0 \rightarrow \psi(2S)K^+K^-$ and amplitude analysis of the decay $B_s^0 \rightarrow \psi(2S)K^+K^-$ with the LHCb Experiment, 01.08.22

Kai Habermann Amplitude Fits in the Decay $\Lambda_b^0 \rightarrow \Lambda_c^+ D^0 K$ using TensorFlow2.0 and JAX, 15.01.23

Niclas Sommerfeld Design and Commissioning of a Silicon Chip Timing Setup, 13.10.2023

Can-Deniz Arslan Developments for the LHCb Mighty Tracker Readout System aiming at the Characterization of MAPS, 29.11.23

Bachelor Arbeiten:

Henry Schumacher A flexible cooling system for HV-MAPS sensor tests

Elena Päßgen A Reference Library for Resonance Lineshapes, 23.3.2023

Amelia Carina de Lope Fend Simulation studies for a Sexaquark search at LHCb, 22.11.2023

Stefan Harst - Lehramt Entwicklung einer LHCb Masterclass zum Thema
Hadronenspektroskopie, 15.11.2021

Lukas Exner - Lehramt Didaktische Elementarisierung von Hadronenspektroskopie im Rahmen einer
Netzwerk Teilchenwelt-Masterclass, 25.9.2023

Kurzbericht

- öffentlich -

Zuwendungsempfänger: Rheinische Friedrich-Wilhelms Universität Bonn

Projektleitung: Prof. Dr. Sebastian Neubert

Verbund: Verbundprojekt 05H2021 (ErUM-FSP T04)

Thema: Run 3 von LHCb am LHC: Betrieb der Realtime Analyse und Datenmanagement

1. Ziel und Inhalt des Projektes

Das LHCb Experiment am Large Hadron Collider (LHC) am CERN in Genf hat zum Ziel die fundamentalen Wechselwirkungen zwischen Materieteilchen zu erforschen und nach neuen Phänomenen auf kleinsten Skalen zu suchen. Das Experiment ist zu diesem Zweck optimiert auf die Detektion von Hadronen mit schweren Quarks, die in Proton-Proton Kollisionen am LHC erzeugt werden, und führt Präzisionsmessungen der Zerfallsparameter dieser schweren Hadronen durch. Bei diesen Zerfällen wandelt die schwache Wechselwirkung schwere Quarks in leichtere Quarks um. Das Physikprogramm umfasst hierbei unter anderem die Bestimmung von Parametern des Standardmodells der Elementarteilchenphysik mit höchster Präzision, sowie die Suche nach seltenen Zerfällen, die im Standardmodell stark unterdrückt und daher besonders sensitiv auf neue Physik sind. Quarks nehmen zudem an der starken Wechselwirkung teil und treten nur als gebundene Systeme in Hadronen auf. Mit dem LHCb Detektor wurden bereits zahlreiche neue Quark-Systeme entdeckt, darunter auch vollkommen neue Konfigurationen, wie Tetraquarks und Pentaquarks. Aus dem Spektrum dieser Zustände werden Informationen über die noch wenig verstandenen Bindungskräfte in Hadronen extrahiert. Die Gruppe in Bonn konzentriert sich in ihrer Arbeit auf letzteres Forschungsfeld.

In den Jahren 2019-2022 wurde der LHCb Detektor mit neuen Komponenten ausgestattet und dann für den sog. Run 3 des LHC wieder in Betrieb genommen. Dieses Upgrade des Experiments ermöglicht es, Daten mit einer höheren Rate aufzuzeichnen. Eine entscheidende Komponente der neuen Systems ist der erstmals vollständig in Software implementierte Trigger, der es erlaubt komplexe Physik-Signaturen in Echtzeit (daher der Name „Realtime Analyse“ RTA) zu rekonstruieren und zu selektieren. Im vorliegenden Projekt wurden Software-Pakete entwickelt, die der Verarbeitung der gesammelten Daten dienen. Die Bonner Gruppe trug zum sog. Sprucing und zu den sog. Analysis Productions bei. Beide Systeme dienen dazu aus dem mit dem RTA System aufgezeichneten Datenstrom, die für eine spezifische Fragestellung interessanten Daten heraus zu filtern und so aufzubereiten, dass Forschende sich auf die Interpretation dieser Daten konzentrieren können. Diese Entwicklungen fanden im Rahmen der koordinierten Software-Entwicklung in der LHCb Kollaboration statt.

2. Ablauf und Ergebnisse des Vorhabens

In den Jahren 2022 bis Mitte 2024 wurden im Rahmen des Vorhabens in Bonn zwei DoktorandInnen gefördert. Beide waren in die internationalen Entwicklungsteams des RTA bzw DPA (Data processing and analysis) Subprojektes in LHCb eingebunden, nahmen

regelmässig an virtuellen Teambesprechungen teil, präsentierten ihre Arbeit auf Kollaborationstreffen und leisteten Schichtdienst zum Betrieb des LHCb Experiments in 2023 und 24.

Die Idee des Sprucing Programms ist es, die RTA Software, die online in Echtzeit die Daten selektiert in einem zweiten Schritt nocheinmal offline über dedizierte Datenströme laufen zu lassen. Diese Entwicklung ermöglicht es neue, innovative Fragestellungen an die Daten zu implementieren, die zum Zeitpunkt der Datennahme noch nicht explizit vorgesehen waren. Der Bonner Beitrag konzentrierte sich hier zunächst darauf einen Monitoring Mechanismus zu implementieren mit dem eindeutig dokumentiert wird, welche Selektionsalgorithmen im Sprucing wie viele Ressourcen brauchen und welchen Output liefern. Ein erster Prototyp dieses Monitorings wurde im Rahmen des Projektes erfolgreich implementiert. Danach übernahm die Gruppe der Verantwortung über die sog. File Summary Records, Metadaten, die den Inhalt der im RTA System erzeugten Dateien intern dokumentieren. Diese Komponente wurde in 2023 erfolgreich implementiert und getestet und kommt inzwischen routinemässig in der Verarbeitung der Daten aus 2023 und 2024 zum Einsatz.

Die Analysis Productions stellen ein Interface zwischen der RTA-Software und den Endnutzern der Daten dar. Das System erlaubt die zentrale Aufbereitung der Daten für die Physikanalyse. Dabei können Daten nach den Bedürfnissen der Analysten gefiltert und in ein einfacheres Datenformat (typischerweise sog N-Tuple) konvertiert werden, welches dann zur Interpretation individuell weiter verarbeitet wird. Die Bonner Gruppe trug zur Automatisierung dieser Datenverarbeitungs-Abläufe bei. Bereits zu Beginn des Projektes wurde klar, dass sich dieses System hervorragend eignet um die Daten auch Dritten – also Nicht-Experten – zur Verfügung zu stellen. Und so konzentrierten sich die weiteren Entwicklungen im Folgenden auf ein innovatives Tool – Den sog. Ntuple-Wizard, der es Forschenden erlaubt ohne Spezialkenntnisse der LHCb Software, mit Hilfe eines gut dokumentierten Web-Interfaces Anfragen an die Daten zu formulieren, die dann auf dem Analysis Productions Service abgearbeitet werden. Dieses System wurde im Verlauf des Projektes voll entwickelt und ging Anfang 2024 in die Beta-Test Phase um Feedback von Analysten der LHCb Kollaboration einzuholen. Zudem wurde das System mit dem CERN Open Data Portal integriert und soll in Zukunft die LHCb Daten für die Forschungsgemeinschaft offen und FAIR zugänglich machen. Diese Integration wird nun auch im Nachfolgeprojekt weiter vorangetrieben.

3. Darstellung der wesentlichen Ergebnisse und deren konkreter Nutzen sowie ggf. die Zusammenarbeit mit anderen Forschungseinrichtungen

Im Rahmen der Bonner Beiträge zum LHCb Experiment wurden Software-Tools entwickelt, die es ermöglichen, die grossen, mit dem RTA System aufgezeichneten Datensätze für die Physikanalyse aufzubereiten und den Analysten zur Verfügung zu stellen. Die entwickelten Systeme sind wichtige Komponenten im Betrieb der gesamten Datennahme und Verarbeitung. Sie kommen sowohl den Forschenden in LHCb als auch, im Rahmen der LHCb Open Data Initiative, der weiteren Forschungsgemeinschaft zugute. Die Bonner Gruppe leitete auch den Release der gesamten LHCb Daten aus den Jahren 2011-2012 im Open Data Verfahren. Im Dezember 2023 wurden so in Zusammenarbeit mit dem CERN Open Data Portal ca 900 TB an Daten für die Öffentlichkeit zugänglich gemacht.