
ZE: Alfred Wegener Institute Helmholtz Centre

Forderkennzeichen: 16HPC017

Vorhabenbezeichnung : Enabling dynamic and intelligent workflows in the future EuroHPC ecosystem

Laufzeit des Vorhabens: 01.01.2021- 28.02.2024

Berichtszeitraum: 01.01.2021- 28.02.2024

Kurzbericht

Im Laufe dieses Projekts wurden bedeutende Fortschritte bei der Entwicklung und Implementierung eines innovativen Earth System Model (ESM) Workflows erzielt. Dieses Projekt konzentrierte sich auf die Verbesserung der Fähigkeiten von Klimasimulationen durch die Integration neuer High-Performance-Computing (HPC) Technologien.

Das Projekt begann mit einer gründlichen Analyse und Systematisierung der bestehenden ESM-Workflow-Komponenten. Wir identifizierten die wesentlichen Software- und Infrastrukturanforderungen, die zur effektiven Unterstützung dieser Komponenten erforderlich sind. Diese grundlegende Arbeit ermöglichte es uns, eine Reihe relativer Metriken zu definieren, um Verbesserungen und den Gesamterfolg der Entwicklung des eFlows4HPC-Workflows zu messen, insbesondere im Hinblick auf die angestrebten Anwendungsfälle. Aus dieser Analyse konnten wir spezifische Anforderungen für den Software-Stack und die notwendigen Infrastrukturanpassungen ableiten, die während der gesamten Projektlaufzeit erforderlich waren.

Eines der ersten und wichtigsten Ergebnisse war die Entwicklung einer umfassenden Designarchitektur für den ESM-spezifischen HPC-Workflow. In dieser Architektur wurden alle Aktivitäten, Komponenten und ihre Interaktionen innerhalb des Arbeitsablaufs detailliert beschrieben, so dass ein klarer Plan für die nachfolgenden Implementierungsphasen entstand. Diese Entwurfsphase legte den Grundstein für die Entwicklung eines dynamischen und reaktionsschnellen ESM-Workflows, mit besonderem Augenmerk auf der Ermöglichung von Echtzeit-Datenanalysen während der Simulationsläufe.

Im weiteren Verlauf des Projekts konzentrierten wir uns auf die Implementierung der Workflow-Komponenten, beginnend mit der Meereskomponente des ESM, FESOM2. In dieser Phase ging es darum, externe Software wie Hecuba zu integrieren, die eine Schnittstelle zur Cassandra-Datenbank bildet, um die dynamische Datenanalyse zu erleichtern. Diese Integration ermöglichte den Zugriff auf Daten, die während der Modellläufe erzeugt wurden, und deren Analyse in Echtzeit, was eine bedeutende Innovation in den ESM-Workflows darstellt.

Die Umstellung von serieller auf parallele IO war eine wichtige Errungenschaft in dieser Phase. Indem wir es parallelen Prozessen innerhalb des Modells ermöglichten, Daten unabhängig voneinander zu senden, konnten wir die für die Datenverarbeitung benötigte Zeit und die Speicherressourcen erheblich reduzieren. Diese Umstellung war für die Verarbeitung der großen Datenmengen, die der ESM generiert, von entscheidender Bedeutung und stellte sicher, dass der Workflow auch in großem Maßstab effizient arbeiten kann.

Die Integration von Hecuba in FESOM2 war ein weiterer wichtiger Meilenstein. Diese Integration ergänzte das traditionelle File-IO-System durch einen fortschrittlicheren Mechanismus, der Daten direkt an eine Cassandra-Datenbank übertrug. Die Umstellung auf Cassandra stellte jedoch eine Herausforderung für die Datenanalyse dar, insbesondere weil sich die Datenformate von den

standardmäßigen NetCDF-Dateien unterscheiden, die normalerweise in FESOM2 verwendet werden. Um dieses Problem zu lösen, haben wir Python-Wrapper für Cassandra-Daten entwickelt, um den Analyseprozess zu rationalisieren und ihn kompatibel mit unseren bevorzugten Analysetools zu machen.

Während des gesamten Projekts wurden erhebliche Verbesserungen und Fehlerkorrekturen an der Hecuba-Software vorgenommen, um ihre Stabilität und Leistung zu erhöhen. Diese Verbesserungen waren entscheidend für den nahtlosen Betrieb des integrierten Arbeitsablaufs. Aufbauend auf diesen Fortschritten entwickelten wir einen Prototyp für dynamische Arbeitsabläufe mit PyCOMPSS und FESOM2, der erfolgreich auf dem Marenostrom 4 System getestet wurde. Dieser Prototyp war ein bedeutender Schritt nach vorn und demonstrierte die Machbarkeit einer dynamischen Echtzeit-Datenanalyse innerhalb von ESM-Workflows.

Im letzten Jahr des Projekts konzentrierten wir uns auf die Verfeinerung und Vervollständigung des dynamischen ESM-Workflows. Wir implementierten eine In-Memory-Online-Analysefunktion für in Cassandra gespeicherte Daten, die die Verarbeitung und Analyse von Simulationsdaten in Echtzeit ermöglicht. Außerdem wurde eine Komponente für maschinelles Lernen eingeführt, die Anomalien in den Daten erkennen und die Größe der Zeitschritte in FESOM2 dynamisch anpassen soll. Diese adaptive Fähigkeit verbesserte die Stabilität und Genauigkeit der Simulationen erheblich.

Darüber hinaus haben wir einen Feedback-Mechanismus in den PyCOMPSS-Workflow-Manager implementiert, der es ermöglicht, Simulationen auf der Grundlage der Ergebnisse der Online-Datenanalyse beizubehalten oder zu beschneiden. Dieses Maß an Dynamik und Reaktionsfähigkeit im Arbeitsablauf ist eine neue Entwicklung, die mehr Kontrolle und Flexibilität bei der Verwaltung von ESM-Simulationen bietet.

Eine weitere wichtige Errungenschaft war die Fertigstellung verschiedener IO-Backends für FESOM, die mehrere Optionen für die Datenspeicherung bieten, darunter in der Cassandra-Datenbank, in Cloud-Speicherlösungen wie Swift-OpenStack und S3 sowie im Chunk-basierten Zarr-Format auf Festplatte. Dank dieser Flexibilität kann der Arbeitsablauf auf verschiedene Computerumgebungen und Speicheranforderungen zugeschnitten werden.

Nachdem alle Komponenten integriert waren, führten wir eine umfassende Bewertung des dynamischen ESM-Workflows durch und bestätigten seine Effektivität und Zuverlässigkeit. Der Workflow wurde dann in das Workflow Repository und das HPCWaaS-System integriert, so dass er von Alien4Cloud aus gestartet werden kann. Diese Fähigkeit wurde sowohl auf dem Marenostrom 4 System bei BSC als auch auf dem Levante System bei DKRZ erfolgreich getestet.

Dieses Projekt hat erfolgreich einen innovativen dynamischen ESM-Workflow entwickelt und implementiert, der den Bereich der Klimamodellierung erheblich vorantreibt. Durch die Integration fortschrittlicher HPC-Technologien und Echtzeit-Datenanalysefähigkeiten haben wir einen Arbeitsablauf geschaffen, der nicht nur die Effizienz und Genauigkeit von Klimasimulationen erhöht, sondern auch den Weg für zukünftige Innovationen in der Erdsystemmodellierung bereitet. Die Ergebnisse dieses Projekts stellen einen Fortschritt in unserer Fähigkeit dar, komplexe Klimaphänomene zu verstehen und vorherzusagen, was letztendlich zu einer fundierteren Entscheidungsfindung angesichts der globalen Umweltprobleme beiträgt.

Während des gesamten Projekts haben wir eng mit Partnern des Barcelona Supercomputing Center (BSC, Spanien) und des Euro-Mediterranean Center on Climate Change (CMCC, Italien) zusammengearbeitet. Wir tauschten uns auch aktiv mit Partnern aus anderen Institutionen aus, die am eFlows4HPC-Projekt teilnehmen.

Teil II: Eingehende Darstellung

Entwicklung von Anforderungen an den eFlows4HPC-Software-Stack aus Pillar II und Bewertungsmetriken.

In einem ersten Schritt analysieren wir die Anforderungen und definieren die Bewertungsmetriken, die für die Entwicklung von Hochleistungs-Workflows im Kontext von Erdsystemmodellen (ESM) erforderlich sind. Das Projekt zielt darauf ab, HPC-Workflows durch die Integration innovativer Komponenten wie dynamische Datenanalyse und KI-gestützte Diagnostik in ESM-Simulationen zu verbessern und damit letztlich deren Effizienz, Skalierbarkeit und Anpassungsfähigkeit zu erhöhen. Daher lag der Schwerpunkt auf der Definition von Metriken für diese Komponenten.

Wir geben einen umfassenden Überblick über einen typischen ESM-Arbeitsablauf, der mehrere Phasen umfasst, wie z.B. die Vorbereitung von Rechengittern, Anfangsbedingungen, Antriebsdaten, Modellkompilierung und Ausführung. Auf diese Phasen folgen die Datenausgabe, die Nachbearbeitung und die Analyse, die mit der Archivierung und Verteilung der Daten endet. Diese Arbeitsabläufe sind komplex und rechenintensiv, so dass eine robuste Infrastruktur erforderlich ist, die große Datenmengen, intensive E/A-Operationen und die komplizierten Abhängigkeiten zwischen den verschiedenen Komponenten bewältigen kann.

Die dynamische Diagnosekomponente, die das AWI im Rahmen des eFlows4HPC-Projekts eingeführt hat, wurde entwickelt, um spezifische Herausforderungen in traditionellen ESM-Workflows zu bewältigen. Die Komponente ermöglicht das Beschneiden von Ensemblemitgliedern während der Laufzeit und optimiert so die Nutzung von Rechen- und Speicherressourcen. Dieser dynamische Ansatz ermöglicht eine bessere Ressourcenzuweisung, indem er unnötige Simulationen eliminiert und so die Gesamteffizienz steigert.

Wir definieren spezifische Software- und Infrastrukturanforderungen, die zur Unterstützung des ESM-Workflows mit innovativen Komponenten erforderlich sind. Dazu gehören die Verfügbarkeit von HPC-orientierten Programmiersprachen, spezielle Software-Abhängigkeiten und wesentliche Infrastrukturelemente wie schnelle Netzwerkkommunikation, reaktionsschnelle Dateisysteme und umfangreiche

Speicherkapazitäten. Außerdem besteht ein Bedarf an Fehlertoleranz, Portabilität über verschiedene HPC-Infrastrukturen hinweg und Integration mit Langzeitspeicherlösungen.

Um die Effektivität des entwickelten Workflows zu bewerten, wird eine Reihe von Metriken vorgeschlagen. Diese Metriken decken verschiedene Aspekte ab, darunter Entwicklung und Wartung, Zugänglichkeit und Bereitstellung, Datenverwaltung, Zuverlässigkeit, Leistung, Skalierbarkeit, Energieverbrauch und Kosten. So werden beispielsweise Metriken wie Lines of Code (LoC), Grad der Portabilität (DoP) und Deployment Time (DT) vorgeschlagen, um die Leichtigkeit der Entwicklung und des Einsatzes in verschiedenen Umgebungen zu bewerten. Leistungsmetriken wie Ausführungszeit (ET), Beschleunigung (SU) und E/A-Zeit (IOT) sind entscheidend, um sicherzustellen, dass die Workflows effizient und skalierbar bleiben. Wir führen auch ESM-Workflow-spezifische Metriken ein, wie die Genauigkeit der Ergebnisse (AR) und Simulierte Jahre pro Tag (SYPD), um die wissenschaftliche Validität und den Durchsatz von ESM-Simulationen zu messen.

Die in der Anfangsphase definierten Anforderungen und Metriken waren die Richtschnur für das Design und die Implementierung der eFlows4HPC-Softwarearchitektur. Das ultimative Ziel war es, ein flexibles, skalierbares und effizientes System zur Verwaltung von Arbeitsabläufen zu schaffen, das in verschiedenen HPC-Umgebungen eingesetzt werden kann und so eine ehrgeizigere wissenschaftliche Forschung in der Klimamodellierung und anderen verwandten Bereichen ermöglicht. In den nächsten Phasen des Projekts werden diese Anforderungen in konkrete Softwarelösungen umgesetzt und ihre Effektivität wird am Ende des Projekts anhand der definierten Metriken streng evaluiert werden.

Entwurf des Anwendungsfalls ESM Dynamischer Arbeitsablauf

In der nächsten Phase des Projekts konzentrierten wir uns auf die Definition des architektonischen Rahmens und des Workflow-Designs, die für die Implementierung des Anwendungsfalls "Dynamischer ESM-Workflow" erforderlich sind. Die Arbeit konzentrierte sich darauf, sicherzustellen, dass die Softwarekomponenten effizient strukturiert und orchestriert sind, um die komplexen Aufgaben zu bewältigen, die mit diesem Workflow verbunden sind.

Wir begannen mit der Festigung der Architektur des eFlows4HPC-Software-Stacks, der eine entscheidende Rolle bei der Gestaltung und Implementierung der für den Dynamic ESM-Workflow erforderlichen Softwarekomponenten spielt. Die Architektur umfasst zwei Hauptkomponenten: einen Workflow-Manager (YORC) und einen Task-Manager (PyCOMPSs). Der Workflow-Manager, der eine erweiterte Tosca-Spezifikation verwendet, orchestriert den gesamten Arbeitsablauf, einschließlich Ressourcenzuweisung, Datenlogistik und Aufgabenkoordination. Der Aufgabenmanager, der als Python-Anwendung implementiert ist, führt spezifische Aufgaben aus, die im Rahmen des Workflows definiert wurden, und nutzt dabei die von PyCOMPS bereitgestellten parallelen Verarbeitungsmöglichkeiten. Diese Architektur wurde so konzipiert, dass sie modular und skalierbar ist, so dass sie sich an verschiedene ESM-Szenarien anpassen und Rechenressourcen effizient verwalten kann.

Bei der Gestaltung des Workflows wurde eine klare Trennung zwischen der High-Level-Orchestrierung und der Low-Level-Aufgabenausführung vorgenommen, um eine robuste und flexible Implementierung zu ermöglichen. Der High-Level-Workflow, der von YORC verwaltet wird, ist für die Bereitstellung und Initialisierung von Softwarekomponenten und Datensätzen verantwortlich und stellt sicher, dass alle erforderlichen Ressourcen für die Simulationen zur Verfügung stehen. Der Low-Level-Workflow, der von PyCOMPSs verwaltet wird, kümmert sich um die Ausführung der Simulationen selbst und nutzt dabei die parallelen Verarbeitungsmöglichkeiten der zugrunde liegenden HPC-Infrastruktur voll aus. Darüber hinaus wurden umfassende Überwachungsmechanismen entwickelt, um sicherzustellen, dass die Simulationen wie erwartet ablaufen und um dynamische Anpassungen auf der Grundlage von Echtzeit-Datenanalysen zu ermöglichen.

Darüber hinaus wurden fortschrittliche Datenlogistikdienste entwickelt, um die Bewegung und Umwandlung von Daten innerhalb des Workflows zu verwalten. Dadurch wird sichergestellt, dass die Daten verfügbar sind, wann und wo sie benötigt werden, ohne dass es zu Engpässen oder Verzögerungen im Simulationsprozess kommt. Die Integration neuartiger Tools und Techniken, wie z.B. innovative Speicherlösungen und datengesteuerte Aufgaben, wurde ebenfalls untersucht, um die Effizienz und Effektivität des Arbeitsablaufs zu verbessern.

Insgesamt wurde mit der Arbeit in dieser Phase eine solide Grundlage für die Implementierung des Dynamic ESM Workflow Use Case geschaffen, die sicherstellt, dass die Architektur für die Anforderungen der Hochleistungs-Erdsystemmodellierung geeignet ist. Die hier getroffenen Entscheidungen dienen als Richtschnur für die Entwicklung der Softwarekomponenten in den nachfolgenden Phasen, wobei der Schwerpunkt auf der Maximierung von Leistung, Skalierbarkeit und Anpassungsfähigkeit liegt.

Erste Iteration der Softwareversion

In dieser Phase des eFlows4HPC-Projekts lag der Schwerpunkt auf der Entwicklung und Implementierung des Dynamic Earth System Model (ESM) Workflow innerhalb von Pillar II. Ziel war es, die ESM-Ausführung auf HPC-Infrastrukturen durch die Integration von dynamischen Workflow-Komponenten, KI-gesteuerten Methoden für das Ensemble Pruning und fortschrittlichen I/O-Technologien zu verbessern, um letztendlich die Nutzung von Rechenressourcen zu optimieren und die Effizienz von Klimamodell-Simulationen zu steigern.

Der dynamische ESM-Workflow befasst sich mit mehreren kritischen Herausforderungen, die mit der Ausführung komplexer ESMs auf Supercomputern verbunden sind. Zu diesen Herausforderungen gehören die manuelle Verwaltung verschiedener Workflow-Schritte (z.B. Modellinitialisierung, Ausführung und Nachbearbeitung) und die Ineffizienz, die durch die Einbeziehung von Ensemblemitgliedern verursacht wird, die keine wertvollen wissenschaftlichen Ergebnisse liefern.

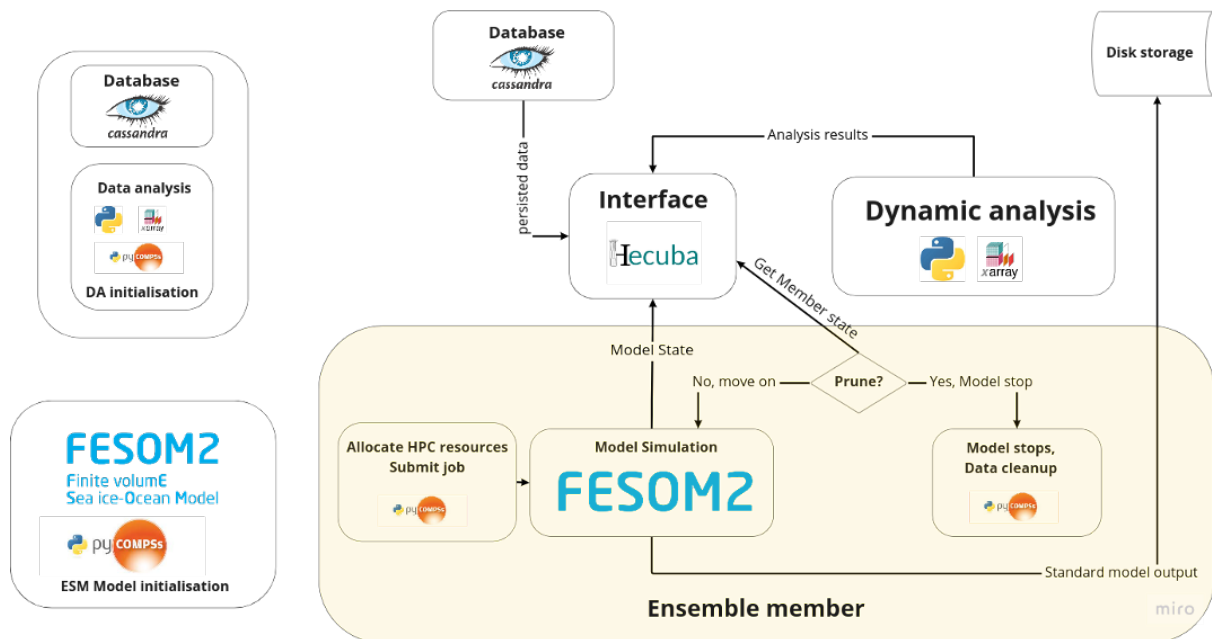


Abbildung. 1 ESM Dynamischer ESM-Workflow mit dem Fluss von Daten und Ereignissen

Wichtigste Errungenschaften:

1. die Vereinfachung der HPC-Ausführung:

- Die COMPSs-Laufzeitumgebung wurde eingesetzt, um die Komplexität von HPC-Umgebungen zu abstrahieren und so eine rationalisierte Ausführung von ESM-Workflows zu ermöglichen. Dank dieser Abstraktion kann derselbe Workflow in verschiedenen HPC-Infrastrukturen ausgeführt werden, ohne dass die Benutzer über umfassende Kenntnisse der jeweiligen HPC-Architektur verfügen müssen. Die Implementierung konzentrierte sich auf die Automatisierung der Verwaltung von Ressourcen, der Auftragsplanung und der Datenverarbeitung, was es den Benutzern erleichtert, komplexe Simulationen ohne manuelle Eingriffe auszuführen.

2. KI-gesteuertes Ensemble Pruning:

- Ein bedeutender Fortschritt war die Entwicklung von KI-gesteuerten Methoden zum dynamischen Pruning von Ensemblemitgliedern während der Modellausführung. Bei diesem Ansatz wurde ein Pruning-Mechanismus implementiert, der den Zustand jedes Ensemblemitglieds regelmäßig anhand von Scheindaten überprüft und so den Prozess der Identifizierung und des Ausscheidens von Mitgliedern simuliert, die wahrscheinlich keine wertvollen Ergebnisse liefern. Der Pruning-Mechanismus wurde in das PyCOMPSs-Framework integriert und ermöglicht die automatische Löschung irrelevanter Ensemble-Mitglieder, wodurch Rechenzeit und Speicherressourcen gespart werden. Diese Funktion ist entscheidend für die Optimierung der Ressourcennutzung bei groß angelegten Simulationen.

3. Integration fortschrittlicher I/O-Technologien:

- Das Projekt untersuchte und implementierte die Integration des verteilten Key-Value-Speichers Hecuba mit dem FESOM2 Ozeanmodell. Diese Integration wurde durch die Entwicklung eines Fortran-Wrappers erreicht, der es dem FESOM2-Modell ermöglicht, über C/C++-APIs mit der Hecuba-Datenbank zu kommunizieren. Der Wrapper erleichtert die Speicherung von Modellausgabedaten in Cassandra, einer skalierbaren NoSQL-Datenbank, die eine effiziente Online-Datenverarbeitung und Echtzeitdiagnosen während der Modellausführung ermöglicht.
- Bei der anfänglichen Implementierung wurde ein Arbeitsablauf eingerichtet, bei dem die Modellausgabedaten direkt in Hecuba geschrieben werden, wodurch die traditionelle dateibasierte E/A umgangen wird. Dadurch wurde nicht nur die Geschwindigkeit des Datenzugriffs verbessert, sondern auch die Grundlage für künftige dynamische Analyseaufgaben geschaffen, die mit in Hecuba gespeicherten Live-Modelldaten arbeiten werden. Die Verwendung von Mocked-Daten in dieser Phase ermöglichte das Testen und die Validierung der Integration, wobei geplant ist, das vollständige Datenmodell und die dynamische Analyse in späteren Phasen zu implementieren.

4. die Workflow-Ausführung und Aufgaben-Orchestrierung:

- Der dynamische ESM-Workflow wurde mit PyCOMPSs ausgeführt, einem Python-basierten Task-Management-Framework, das die Ausführung von Modellsimulationen auf HPC-Ressourcen orchestriert. Der Workflow wurde so konzipiert, dass er mehrere Ensemble-Mitglieder handhaben kann, die jeweils parallel ausgeführt werden, wobei die einzelnen Simulationen in kleinere Teile zerlegt werden, die nacheinander verarbeitet werden können. Der Task Scheduler in PyCOMPSs verwaltete die Abhängigkeiten zwischen den Aufgaben und sorgte so für eine effiziente parallele Ausführung.
- Der Arbeitsablauf enthielt auch einen Checkpoint-Mechanismus, der eine Zwischenbewertung des Fortschritts der einzelnen Ensemblemitglieder ermöglichte. Wenn ein Mitglied auf der Grundlage der KI-gesteuerten Beschneidungskriterien als irrelevant eingestuft wurde, löste der Workflow die Stornierung nachfolgender Aufgaben für dieses Mitglied aus, um die Verschwendung von Rechenressourcen zu verhindern.

In dieser Phase des eFlows4HPC-Projekts wurden die grundlegenden Elemente des dynamischen ESM-Workflows erfolgreich eingeführt und sein Potenzial zur deutlichen Verbesserung der Effizienz und Skalierbarkeit von Klimamodell-Simulationen auf HPC-Infrastrukturen demonstriert. Die Fortschritte bei der Orchestrierung von Aufgaben, der Ressourcenoptimierung und der Datenverwaltung dienen als robuste Plattform für weitere Verbesserungen in den nachfolgenden Phasen.

[Zweite Iteration der Softwareversion](#)

In dieser Phase des eFlows4HPC-Projekts lag unser Schwerpunkt weiterhin auf der Weiterentwicklung des Dynamic Earth System Model (ESM) Workflow. Die Arbeit konzentrierte sich auf die Verbesserung der in Iteration 1 entwickelten Fähigkeiten, die Integration des gekoppelten OpenIFS-FESOM2-Modells in den Workflow und die

Optimierung der Leistung und Modularität des Workflows für den Einsatz auf HPC-Systemen, insbesondere mit MareNostrum 4.

Wichtigste Errungenschaften:

1. entwicklung und einatz des Workflows:

- Wir haben den ESM-Workflow unter Verwendung des FESOM2-Ozeanmodells mit Unterstützung von Alien4Cloud erfolgreich auf MareNostrum 4 entwickelt und implementiert. Diese Entwicklung umfasste die Implementierung einer containerisierten Workflow-Umgebung innerhalb des eFlows4HPC-Software-Stacks, die eine effizientere Ausführung und Verwaltung der Workflow-Komponenten ermöglicht.
- Die Integration des gekoppelten OpenIFS-FESOM2-Modells (AWI-CM) in den Workflow war ein wichtiger Meilenstein. Diese Integration wurde zunächst mit Konfigurationen getestet, bei denen nur das Ozeanmodell verwendet wurde. Später wurde die Integration auf eine vollständige Kopplung mit der atmosphärischen Komponente ausgeweitet, die auf 336 Kernen läuft, die auf die verschiedenen Modellkomponenten verteilt sind.

2. Verbesserungen von Hecuba und PyCOMPSs:

- Das Hecuba-Framework wurde weiterentwickelt, um die Spezifikation der Datenmodelle zu vereinfachen, die für die Aufnahme von Klimamodelldaten in die Cassandra-Datenbank erforderlich sind. Diese Verbesserung war entscheidend für die Verwaltung der großen Datensätze, die durch die gekoppelten ESM-Simulationen erzeugt wurden.
- PyCOMPSs wurde so erweitert, dass es das Pruning von Ensemble-Simulationsmitgliedern während der Laufzeit unterstützt. Diese Funktion ermöglicht es dem Arbeitsablauf, nicht wertvolle Ensemblemitglieder auf der Grundlage vordefinierter Kriterien dynamisch zu entfernen und so die Nutzung der Rechenressourcen zu optimieren. Der Pruning-Mechanismus ist zwar bereits vorhanden, aber die weiteren Verfeinerungen werden sich auf seine Anwendung für die Modellabstimmung und andere fortgeschrittene Anwendungen konzentrieren.

3. die Ausführung und Verwaltung des Workflows:

- Der ESM-Workflow wurde als hybrider Batch-Job unter Verwendung von SLURM und PyCOMPSs ausgeführt. Der Workflow wurde so konfiguriert, dass er mehrere Ensemble-Mitglieder parallel ausführt, mit regelmäßigen Kontrollpunkten, um die Leistung und Relevanz der einzelnen Mitglieder zu bewerten. Der Arbeitsablauf wurde so angepasst, dass er die Ergebnisse dynamischer Analyseaufgaben verarbeiten kann, was die Entscheidungsfindung für das Ensemble Pruning in Echtzeit ermöglicht.
- Es wurde ein umfassendes Initialisierungsskript entwickelt, das die Konfiguration für die verschiedenen Komponenten des gekoppelten Modells enthält. Dieses Skript automatisiert die Vorbereitung von Arbeitsverzeichnissen, Modell-Namenslisten und Ausgabeverzeichnissen und stellt sicher, dass der Workflow mit minimalem manuellem Aufwand zur Ausführung bereit ist.

4. HPC Workflow as a Service (HPCWaaS) Anwendung:

- Eine wichtige Errungenschaft in dieser Phase war die Erstellung einer HPCWaaS-Anwendung unter Verwendung von Alien4Cloud, die die Bereitstellung und Verwaltung des ESM-Workflows über eine Webschnittstelle ermöglicht. Diese Anwendung modularisiert die Workflow-Komponenten und orchestriert ihre Ausführung auf HPC-Systemen. Die Anwendung stützt sich auf eine Tosca-Vorlage zur Definition der Ausführungsumgebung, die derzeit fest kodierte Pfade und Einstellungen enthält. In der nächsten Phase werden diese parametrisiert, um die Wiederverwendbarkeit und Portabilität der Anwendung über verschiedene HPC-Umgebungen hinweg zu verbessern.
- Die Tosca-basierte Anwendung erleichtert die nahtlose Integration mit MareNostrum 4 und möglicherweise anderen HPC-Plattformen, indem sie die von Alien4Cloud bereitgestellten Sicherheits- und Anmeldeungsmanagementfunktionen nutzt. Benutzer können Workflows mit SSH-Schlüsseln, die in einem Tresor gespeichert sind, sicher ausführen und so sicherstellen, dass sensible Informationen angemessen behandelt werden.

Fertigstellung des Arbeitsablaufs und Auswertung der Ergebnisse

In dieser letzten Phase des Projekts konzentrierten wir uns auf die Verfeinerung des dynamischen Erdsystemmodells (ESM). In dieser Phase ging es um die Bewältigung der besonderen Herausforderungen, die ESM-Simulationen auf HPC-Infrastrukturen mit sich bringen, wobei der Schwerpunkt auf der Verbesserung der Datenverarbeitung, der E/A-Effizienz und der On-the-fly-Datenanalyse lag. Die Arbeit zielte darauf ab, sicherzustellen, dass sich der Arbeitsablauf während der Ausführung des Modells dynamisch anpassen kann, wodurch die Ressourcennutzung optimiert und die Simulationseffizienz insgesamt verbessert wird.

Endgültige Version des Arbeitsablaufs

Die endgültige Version des Dynamic ESM-Simulationsworkflows umfasst mehrere wichtige Teile (siehe auch Abb. 2):

- **Simulationskonfiguration:** Die Benutzer geben Parameter wie die HPC-Plattform, die Startdaten und die Ressourcenzuweisung an. Diese Parameter werden von Bash Shell- und Python-Skripten validiert, die auch die Fortran-Namenslisten des Modells vorverarbeiten.
- **Ausführung der Simulation:** Der Workflow startet das FESOM2-Modell auf der HPC-Plattform unter Verwendung der angegebenen Konfiguration. Das Modell gibt Daten an eine von Hecuba verwaltete Cassandra-Datenbank aus, was eine effiziente In-Memory-Analyse ermöglicht.
- **Ensemble Pruning:** Der Arbeitsablauf umfasst eine PyCOMPSs-Aufgabe, die den Status jedes Ensemblemitglieds analysiert und entscheidet, ob es beschnitten werden sollte. Diese Entscheidung basiert auf der Stabilität der Simulation. Instabile Mitglieder werden in der Cassandra-Datenbank markiert und anschließend gestoppt.

Wir haben Testsimulationen mit MareNostrum 4 durchgeführt, bei denen verschiedene Konfigurationen verwendet wurden, darunter Meshes mit niedriger und hoher Auflösung. Die Ergebnisse zeigen, dass der Workflow in der Lage ist, umfangreiche Simulationen effizient zu verwalten und dabei die In-Memory-Analyse und das dynamische Pruning zur Optimierung der Ressourcennutzung zu nutzen.

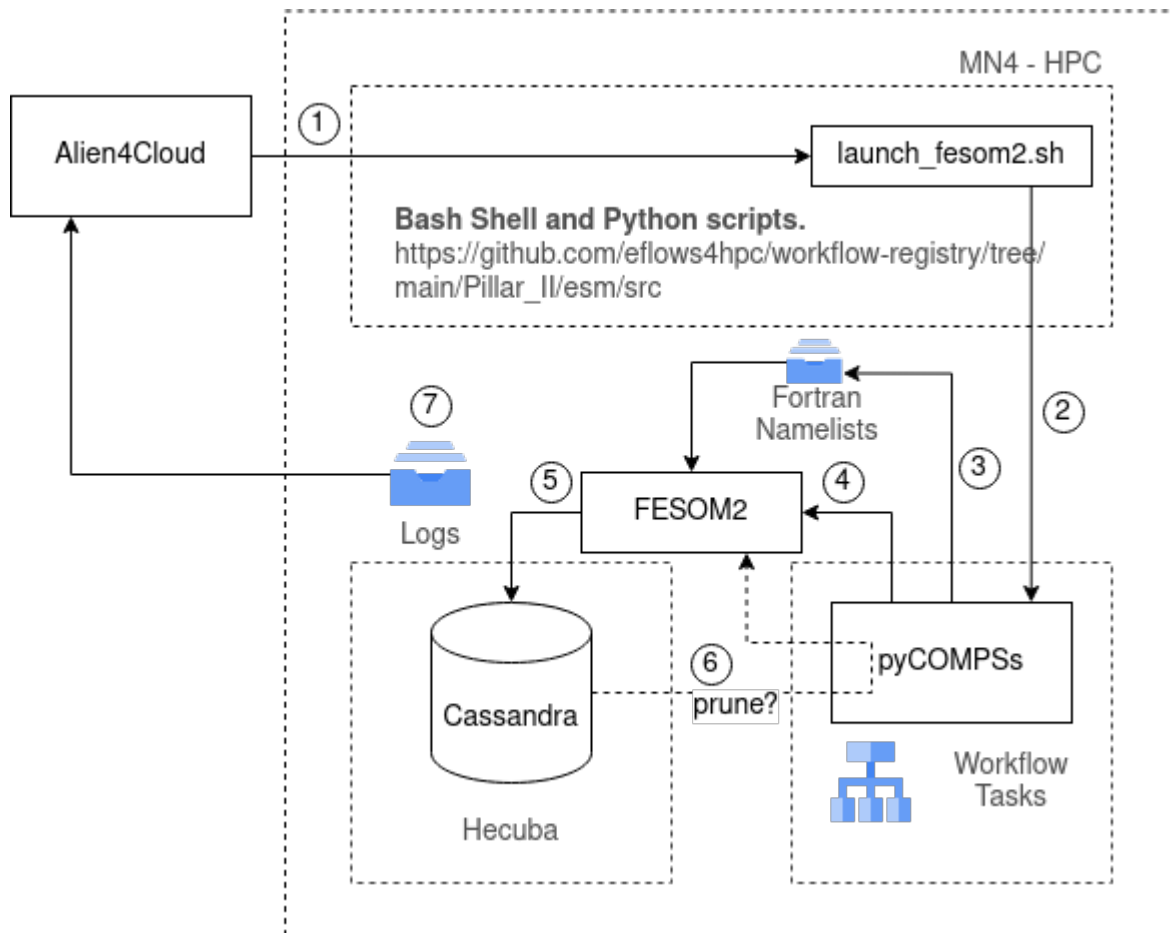


Abbildung 2 - Illustration aller Teile des dynamischen ESM-Workflows

Validierung des Arbeitsablaufs

Zu Beginn des Projekts haben wir eine umfassende Reihe von Anforderungen und Metriken festgelegt, die die Entwicklung des ESM-Workflows leiten sollten. Diese Anforderungen waren wichtig, um sicherzustellen, dass der Workflow die Anforderungen fortgeschrittener Klimasimulationen erfüllt und gleichzeitig an verschiedene HPC-Umgebungen anpassbar ist. Zu den Anforderungen gehörten Aspekte wie Ausführungsstabilität, Portabilität, Workflow-Management und die Integration von KI-Techniken für das Ensemble Pruning.

Um den Workflow zu validieren, entwickelten wir einen Rahmen, der sowohl funktionale als auch nicht-funktionale Metriken umfasste. Diese Metriken bildeten eine quantitative Grundlage für die Bewertung der Leistung, Skalierbarkeit und Ressourcennutzung des Workflows. Zu den wichtigsten Metriken gehörten Ausführungszeit, E/A-Effizienz, Energieverbrauch und die Genauigkeit der Simulationsergebnisse.

Der dynamische ESM-Simulationsworkflow wurde anhand der zu Beginn des Projekts festgelegten Anforderungen gründlich evaluiert. Diese Anforderungen umfassten sowohl funktionale als auch nicht-funktionale Aspekte, die für den Erfolg des Workflows entscheidend sind. Die Bewertung hat gezeigt, dass der Arbeitsablauf die Erwartungen in mehreren Schlüsselbereichen erfüllt oder sogar übertroffen hat.

Erstens wurde die **Ausführungsstabilität** durch die Integration von PyCOMPS für die Fehlertoleranz und die inhärenten Checkpointing-Funktionen von FESOM2 erreicht. Diese Funktionen sorgten dafür, dass sich der Arbeitsablauf von Fehlern erholen konnte, ohne dass es zu größeren Fortschrittsverlusten kam. Im Falle eines Fehlers führte das PyCOMPSs-Framework die fehlgeschlagenen Aufgaben erneut aus, während FESOM2 die Simulationen ab dem letzten gespeicherten Checkpoint wieder aufnahm und so die Kontinuität und Robustheit der Ausführung aufrechterhielt.

Was die **Portabilität** betrifft, so wurde der Workflow erfolgreich mit Singularity containerisiert, wodurch er auf verschiedenen HPC-Infrastrukturen eingesetzt werden konnte. Diese Portabilität wurde durch erfolgreiche Einsätze auf MareNostrum 4 und dem Levante HPC-System bestätigt. Der Containerisierungsansatz in Kombination mit der parametrisierten Ausführung stellte sicher, dass der Workflow problemlos in verschiedene Rechenumgebungen übertragen und dort ausgeführt werden konnte, wodurch die Anforderung der Portabilität erfüllt wurde.

Die **integrierte Workflow-Verwaltung** war eine weitere wichtige Anforderung, die vollständig erfüllt wurde. PyCOMPS spielte eine entscheidende Rolle bei der Verwaltung von Aufgabenabhängigkeiten, parallelen Simulationen und der Verarbeitung von Batch-Jobs. Es bot eine kohärente Umgebung für die Ausführung komplexer Workflows und stellte sicher, dass Simulationen und Analyseaufgaben effektiv koordiniert wurden. Durch den Einsatz von Alien4Cloud wurde die Workflow-Verwaltung weiter optimiert, da die Benutzer die Ausführung über eine benutzerfreundliche Weboberfläche steuern und überwachen können.

Der Workflow beinhaltete auch eine **Integration mit einer Langzeitspeicherung**, die für die Archivierung von Simulationsergebnissen und die Einhaltung der FAIR (Findable, Accessible, Interoperable, Reusable)-Datenprinzipien unerlässlich war. Die Simulationsergebnisse wurden in einer Cassandra-Datenbank gespeichert, und es besteht auch die Möglichkeit, die Daten in einen Cloud-Speicher zu übertragen. Dieser Ansatz erleichterte die langfristige Aufbewahrung der Ergebnisse und unterstützte ihre Wiederverwendung in nachgelagerten Anwendungen.

Die **Anpassungsfähigkeit des Workflows** wurde durch die Flexibilität der Workflow-Konfiguration demonstriert. Der Workflow ermöglichte die einfache Änderung von Parametern wie Modellauflösung und Datenvolumen, um sich an unterschiedliche Simulationsanforderungen anzupassen. Diese Anpassungsfähigkeit war ausschlaggebend dafür, dass der Arbeitsablauf eine Vielzahl von Szenarien bewältigen konnte, einschließlich Änderungen der Berechnungsanforderungen und des Bedarfs an Modellen mit höherer Auflösung.

Der Zugriff auf **In-Memory-Zwischenergebnisse** war ein wesentliches Merkmal des Dynamic ESM Workflows. Der Workflow ermöglichte einen effizienten Zugriff auf In-Memory-Daten mithilfe von Hecuba- und Dask-Arrays, so dass eine Echtzeitanalyse während der Simulationsausführung möglich war. Diese Fähigkeit ermöglichte fliegende Anpassungen der Simulation und verbesserte die Gesamteffizienz und Genauigkeit der Ergebnisse.

Der Workflow erfüllte auch die Anforderung der **KI-Integration für das Ensemble Pruning**. KI-Techniken wurden eingesetzt, um die Simulationszeitschritte dynamisch anzupassen und Ensemble-Mitglieder auf der Grundlage ihrer Stabilität zu beschneiden. Dieser Ansatz optimierte die Nutzung von Rechenressourcen, indem er unnötige Simulationen eliminierte und so die Ausführungszeit und den Energieverbrauch reduzierte.

Darüber hinaus wurden die **ML/DL-Fähigkeiten** des Workflows durch die In-Memory-Datendarstellung verbessert, was die Integration von Methoden des maschinellen Lernens und des Deep Learning erleichterte. Diese Fähigkeit ermöglichte die Anwendung fortschrittlicher Analysen auf Simulationsergebnisse, wodurch die Genauigkeit und Effizienz des Workflows weiter verbessert wurde.

Die Anforderung an die **Unterstützung von High-Performance-Computing** wurde durch die Fähigkeit des Workflows, effizient über verschiedene HPC-Umgebungen hinweg zu skalieren, erfüllt. Der containerisierte Workflow wurde auf mehreren Plattformen getestet und zeigte, dass er seine Leistung und Skalierbarkeit beibehalten kann. Ergänzt wurde dies durch die vordefinierten Umgebungen von FESOM2, die eine reibungslose Ausführung auf verschiedenen HPC-Systemen gewährleisten.

Schließlich unterstützte der Workflow die **Analyse mehrerer Mitglieder** und ermöglichte die gleichzeitige Ausführung mehrerer Ensemblemitglieder. Die Ergebnisse dieser Simulationen waren als Xarray-Datensätze zugänglich und ermöglichten eine umfassende Analyse über verschiedene Ensemble-Mitglieder hinweg. Diese Funktion war entscheidend für die Durchführung groß angelegter Klimasimulationen und stellte sicher, dass der Arbeitsablauf komplexe, mehrdimensionale Daten verarbeiten konnte.

Insgesamt erfüllte der Dynamic ESM-Workflow erfolgreich die umfassenden Anforderungen und bewies seine Robustheit, Flexibilität und Effizienz bei der Ausführung von Klimasimulationen in großem Maßstab auf HPC-Infrastrukturen.

Der Workflow wurde anhand von vordefinierten Metriken bewertet und zeigte eine starke Leistung und Effizienz. Er enthält etwa 2.951 Zeilen Code und erreicht eine 100%ige Portabilität durch Containerisierung. Die Bereitstellungszeit lag unter zwei Minuten und die Ausführungszeit betrug durchschnittlich 160 Sekunden auf MareNostrum 4 mit 144 Kernen. Der Workflow skaliert linear bis zu 8.000 Kernen, wobei E/A-Operationen 55% der Ausführungszeit ausmachen. Die Fehlertoleranz wurde vollständig über PyCOMPSs und das Checkpointing von FESOM2 implementiert. Der Arbeitsablauf verbrauchte 0,02 Kernstunden pro CPU und 2.532 W an Energie, wobei durch dynamisches Pruning erhebliche Einsparungen erzielt wurden. Der Simulationsdurchsatz erreichte 60 simulierte Jahre pro Tag (SYPD) für Konfigurationen mit niedriger Auflösung und 6 SYPD für Konfigurationen mit

hoher Auflösung, was die Effizienz und Skalierbarkeit des Workflows in Hochleistungs-Rechenumgebungen bestätigt.

Der Arbeitsablauf wurde von Experten des Max-Planck-Instituts für Meteorologie (MPI-M) und des Deutschen Klimarechenzentrums (DKRZ) extern evaluiert. Die Bewertung hob die erfolgreiche Integration der In-Memory-Datenanalyse und die Zusammenarbeit zwischen AWI und BSC hervor. Die Experten äußerten sich positiv über die Fähigkeit des Workflows, die Projektziele zu erreichen, und machten Vorschläge für weitere Verbesserungen, wie z.B. die Verfeinerung der Reaktion des Modells auf kritische Ereignisse und die Untersuchung der Verwendung des Zarr-Formats für die Datenverarbeitung.

In dieser letzten Phase des Projekts haben wir den dynamischen ESM-Workflow erfolgreich weiterentwickelt und wichtige Meilensteine in Bezug auf Robustheit, Übertragbarkeit und Effizienz erreicht. Die Fähigkeit des Workflows, Ensemble-Mitglieder dynamisch zu beschneiden und Analysen in Echtzeit durchzuführen, stellt eine bedeutende Verbesserung auf dem Gebiet der Klimamodellierung dar. Die positive externe Bewertung und die Ausrichtung des Workflows auf die Projektziele unterstreichen sein Potenzial für eine breitere Anwendung in HPC-Umgebungen. Die nächsten Schritte werden sich darauf konzentrieren, die Komponenten des Workflows weiter zu optimieren, zusätzliche KI-Techniken zu integrieren und seine Anwendbarkeit auf andere Modelle und Infrastrukturen zu erweitern.

Zusammenfassung des Verwertungsplans

Im Rahmen des eFlows4HPC-Projekts hat das AWI bedeutende Beiträge geleistet, wobei der Schwerpunkt auf der Entwicklung von dynamischen Workflows für die Erdsystemmodellierung (ESM) lag. Zu den wichtigsten technischen Fortschritten gehören die Optimierung des I/O-Durchsatzes bei hochauflösenden Simulationen, die Schaffung eines flexiblen I/O-Formats und einer Backend-Architektur im FESOM2-Modell sowie die Automatisierung von Ensemblesimulationen, Modellabstimmung und Datennachbearbeitung. Diese Innovationen, insbesondere in Zusammenarbeit mit BSC, haben zu einer effizienteren Skalierung des Modells geführt und die Flexibilität und Anpassbarkeit der Datenspeicherung und -abfrage verbessert, was für die Klima- und Ozeanografiemodellierung von entscheidender Bedeutung ist. Die Containerisierung von FESOM und anderen Workflow-Komponenten wurde ebenfalls priorisiert, um die Portabilität des Modells über verschiedene HPC-Umgebungen hinweg sicherzustellen.

Die Nutzung dieser Entwicklungen ist tief in die Ozean- und Klimamodellierungssysteme des AWI integriert, wobei die Ergebnisse durch den Open-Source-Charakter dieser Modelle direkt an die Nutzer weitergegeben werden. Insbesondere das IFS/FESOM-Modell, eine Schlüsselkomponente der digitalen Zwillinge im DestinE-Projekt, profitiert direkt von den Fortschritten des eFlows4HPC-Projekts. Diese Beiträge tragen entscheidend dazu bei, die doppelten Ziele des Green Deal und der Digitalen Strategie der Europäischen Kommission zu erreichen, indem sie hochwertiges, maßgeschneidertes Wissen für eine evidenzbasierte Entscheidungsfindung liefern. Die Verbesserungen beim I/O-Durchsatz und die neue Backend-Architektur von eFlows4HPC werden bereits in Pilotsimulationen für DestinE eingesetzt und werden auch weiterhin die I/O-Strategie des Projekts prägen.

Darüber hinaus wird das IFS/FESOM-Modell, das durch eFlows4HPC verbessert wurde, in anderen hochauflösenden Modellierungsprojekten wie nextGEMS (Horizon 2020), EERIE (Horizon Europe) und WarmWorld (BMBF) eingesetzt. Diese Projekte zielen darauf ab, die Darstellung der Entwicklung des globalen Klimas auf der Hundertjahres-Skala, einschließlich der Variabilität und der Extreme, zu verbessern. Die Automatisierung der Ensemble-Generierung und die innovativen Techniken zur Modellabstimmung, die in eFlows4HPC entwickelt wurden, sind in diesem Zusammenhang besonders wertvoll, da sie die Herausforderungen der Parameterabstimmung in hochauflösenden Modellen angehen. Die für die Modellcontainerisierung entwickelten Technologien werden auch eine entscheidende Rolle bei der Verbesserung der Übertragbarkeit von Klimamodellen auf immer komplexere und heterogenere HPC-Systeme spielen, einschließlich der Systeme im Rahmen des European High Performance Computing Joint Undertaking (EuroHPC JU).

Insgesamt werden die Fortschritte bei den dynamischen ESM-Workflows von eFlows4HPC die Fähigkeiten und die Anpassungsfähigkeit von Klimamodellen erheblich verbessern, ihre Effektivität in verschiedenen HPC-Umgebungen sicherstellen und zu wichtigen wissenschaftlichen und politischen Zielen auf europäischer Ebene beitragen.

Verbreitung der Ergebnisse

Ejarque, Jorge, Rosa M. Badia, Loïc Albertin, Giovanni Aloisio, Enrico Baglione, Yolanda Becerra, Stefan Boschert, Julian R. Berlin, Alessandro D'Anca, Donatello Elia, François Exertier, Sandro Fiore, José Flich, Arnau Folch, Steven J. Gibbons, Nikolay Koldunov, Francesc Lordan, Stefano Lorito, Finn Løvholt, Jorge Macías, Fabrizio Marozzo, Alberto Micheleni, Marisol Monterrubio-Velasco, Marta Pienkowska, Josep de la Puente, Anna Queralt, Enrique S. Quintana-Ortí, Juan E. Rodríguez, Fabrizio Romano, Riccardo Rossi, Jędrzej Rybicki, Mirosław Kupczyk, Jacopo Selva, Domenico Talia, Roberto Tonini, Paolo Trunfio, Manuela Volpe, "Enabling dynamic and intelligent workflows for HPC, data analytics, and AI convergence." *Future generation computer systems* 134 (2022): 414-429.

Cheedela, Suvarchal K., Nikolay Koldunov, Dmitry Sidorenko, Yolanda Becerra, Julian R. Berlin, Juan J. Costa, and Thomas Jung. "Development of Efficient IO Strategies and Workflows for High-Resolution Climate Simulations." In *AGU Fall Meeting Abstracts*, vol. 2023, no. 407, pp. IN51B-0407. 2023.

ESiWACE3 hackathon: HPC workflows for climate models, October 17, 2023, CSC, Espoo (Finland). "Overview of Pillar II workflows ESM Dynamic Workflow for Climate Simulations" Alessandro Danca (CMCC), Suvarchal Cheedela (AWI).