



**Procedura aperta ai sensi dell'art 60 comma 1, del D.lgs n. 50/2016 finalizzata
all'acquisizione dei servizi necessari alle attività di calcolo ad elevate prestazioni
Numero gara 8242858 CIG 8857443132**

Capitolato speciale di appalto

Indice generale

ART 1 - Premessa – scopo e inquadramento del progetto	3
1.1. Introduzione	3
1.2. Sistema modellistico	3
1.3. Sistema di calcolo ad elevate prestazioni	5
1.3.1. CLUSTER ARPAP	5
1.3.2. CLUSTER PLX	5
1.3.3. CLUSTER GALILEO (2015-2018).....	6
1.3.4. CLUSTER GALILEO (2018-2021).....	7
ART 2 - Oggetto dell'appalto	8
ART 3 - Descrizione dei servizi oggetto dell'appalto	8
3.1. Nodo equivalente.....	8
3.2. Unità di servizio	8
3.2.1. Unità di servizio in dedicato (USD):	8
3.2.2. Unità di servizio a consumo (USC):	8
3.3. Ore di calcolo.....	8
3.4. Utenze e accesso	9
3.5. Suite operative delle catene modellistiche	9
3.5.1. Modellistica per la qualità dell'aria	9
3.5.2. Modellistica meteorologica	10
3.6. Apparati.....	10
3.7. Supporto all'utenza.....	10
3.8. Supporto specialistico.....	10
3.9. Supporto alla formazione.....	11
ART 4 - Caratteristiche tecniche delle dotazioni ed apparati correlati al servizio	11
4.1. Ambiente	11
4.1.1. Sistema operativo.....	11
4.1.2. Linguaggi e compilatori	11
4.1.3. HPC workload manager and job scheduler.....	12
4.1.4. Software environment manager	12
4.1.5. Collaborative development environment (CDE).....	12
4.1.6. Librerie.....	12
4.2. Processori.....	12
4.3. Aree di archiviazione	13
4.3.1. area di lavoro in dedicato	13
4.3.2. area di repository in dedicato	13
4.3.3. area scratch condivisa tra le utenze.....	13
4.4. Accessibilità	13
4.5. Continuità del servizio	14
ART 5 - Workflow sistema modellistico attualmente operativo	14
5.1. Introduzione	14
5.2. Scheduler	15
5.3. Catena modellistica principale (progn_g2_ope).....	16
5.3.1. Sope_to_g2	16
5.3.2. Pope_to_g2	17
5.4. Catena modellistica ad alta risoluzione (progn_g3_ope)	18
ART 6 - Supporto e formazione	19
6.1. Supporto specialistico.....	19
6.1.1. Modifica e correzione codici	19
6.1.2. Implementazione nuove funzionalità	19
6.1.3. Parallelizzazione codici	19
6.1.4. Sviluppo ed ingegnerizzazione di soluzioni personalizzate	19
6.2. Supporto alla formazione.....	19
6.2.1. Linguaggi di programmazione	20
6.2.2. Tecniche di programmazione	20
6.2.3. Build System tools e package management tools	20
6.2.4. Operative environment.....	20

ART 1 - PREMESSA – SCOPO E INQUADRAMENTO DEL PROGETTO

1.1. Introduzione

Il Dipartimento Tematico Rischi Naturali ed Ambientali svolge attività di modellistica numerica per applicazioni in campo ambientale, con impiego di competenze specialistiche e risorse computazionali dedicate.

A partire dall'anno 2005 viene svolta, attività di modellistica di dispersione e diffusione di inquinanti in atmosfera a supporto dei soggetti istituzionali preposti al controllo della qualità dell'aria nella Regione, attraverso un sistema modellistico complesso, in grado di effettuare simulazioni multi-scala di fenomeni di trasporto, dispersione, trasformazione chimica e deposizione degli inquinanti. A partire dal sistema modellistico è stata realizzata ed implementata una catena operativa prognostica, finalizzata alla produzione e trasmissione di previsioni di concentrazioni di inquinanti sul territorio regionale (concentrazioni al suolo dei principali inquinanti, indici sintetici dello stato di qualità dell'aria su area estesa, stime di concentrazione dei principali inquinanti su base comunale), in modalità operativa.

Nel corso degli anni successivi il sistema modellistico è stato oggetto di continue evoluzioni e miglioramenti per mantenerlo allo stato dell'arte. In particolare, la parallelizzazione del modello CTM (Chemistry Transport Model) *FARM* con l'introduzione di paradigmi OpenMP e MPI per la parallelizzazione su sistemi di calcolo a memoria condivisa e memoria distribuita, ha permesso all'Agenzia, a partire dal triennio 2010-2011-2012, l'acquisizione ed utilizzo di un sistema di calcolo parallelo ad elevate prestazioni.

Nella Convenzione Triennale con il Dipartimento di Protezione Civile 2009-2012 si è delineata l'esigenza di avviare la fase operativa della catena modellistica meteorologica COSMO in configurazione di assimilazione in continuo di dati osservati ad elevata risoluzione, per la produzione di simulazioni di analisi sul Nord Italia, e per la realizzazione di simulazioni previsionali a breve scadenza da attivarsi in caso di condizioni meteorologiche avverse. Essendo la modellistica meteorologica propedeutica all'operatività della modellistica dispersiva per la qualità dell'aria, ed avendo la modellistica meteorologica di riferimento (Direttiva P.C.M. del 27/02/2004) un valore per applicazioni in campo ambientale, il *sistema di calcolo ad elevate prestazioni* acquisito ed utilizzato per la modellistica per la qualità dell'aria diviene lo strumento per implementare anche la modellistica meteorologica per lo svolgimento delle attività tecniche della Struttura, in qualità di Centro di Competenza del Dipartimento di Protezione Civile.

Nel successivo triennio 2013-2014-2015, per affrontare i problemi dell'obsolescenza delle risorse computazionali e per adeguarsi alle crescenti esigenze delle applicazioni della modellistica numerica, si procede a trasformare gradualmente la composizione delle dotazioni di calcolo inizialmente acquisite. Con un progressivo affiancamento si passa da una configurazione iniziale con risorse di calcolo di proprietà, e quindi di uso esclusivo dell'Agenzia, gestite attraverso servizi di *housing*, di assistenza sistemistica e supporto specialistico, verso una configurazione di servizio complessivo, in cui non viene acquisita la proprietà delle risorse di calcolo ma l'utilizzo, esclusivo o privilegiato, in quantità dinamicamente adeguata alle richieste e per il tempo di effettivo utilizzo. I servizi di supporto e assistenza specialistica, comunque presenti, vengono interamente spostati dall'area sistemistica all'area applicativa.

Alla conclusione del secondo triennio di operatività, il sistema di calcolo raggiunge la piena operatività nella configurazione di *servizio di messa a disposizione delle risorse di calcolo* (dimensionalmente corrispondente alle esigenze dell'Agenzia), in cui viene sfruttata l'elasticità e l'intrinseca dinamicità del servizio per introdurre contrattualmente la disponibilità di ulteriori risorse *on-demand*, da attivare su richiesta, in caso di esigenze particolari (legate alle attività condotte ovvero progetti nazionali ed internazionali di particolare interesse).

In tale configurazione i servizi di calcolo ad elevate prestazioni necessari al Dipartimento Rischi Naturali e Ambientali di Arpa Piemonte per lo svolgimento delle attività di modellistica numerica con applicazioni in campo ambientale, sono rinnovati per i successivi trienni 2016-2019 e 2019-2021.

1.2. Sistema modellistico

Il sistema modellistico fa parte, insieme alla rete di monitoraggio della qualità dell'aria ed all'Inventario Regionale delle Emissioni in Atmosfera (*IREA*), degli strumenti previsti dal Programma

di Valutazione della Qualità dell'aria della Regione Piemonte, redatto ai sensi dell'art. 5, comma 6, del d.lgs 155/2010 ed approvato con Deliberazione della Giunta Regionale 30 dicembre 2019, n. 24-903:

- E' lo strumento con il quale sono state realizzate le valutazioni tecniche alla base della redazione del Piano regionale di qualità dell'aria (PRQA) della Regione Piemonte, approvato dal Consiglio regionale, con DCR 25 marzo 2019, n. 364-6854.
- E' lo strumento alla base del semaforo antismog, un meccanismo di attivazione di misure temporanee per la limitazione delle emissioni in condizioni emergenziali, introdotto con Deliberazione della Giunta Regionale n. 9-2916 del 26 febbraio 2021.
- Il sistema modellistico è alla base delle attività condotte dal Dipartimento Rischi Naturali ed Ambientali in numerosi progetti nazionali ed internazionali quali:
 - CLIMAERA (<https://www.climaera.eu>), progetto nato nell'ambito del programma di cooperazione transfrontaliera (Interreg) Italia-Francia Alcotra, su un periodo di tre anni (2017 - 2020);
 - MITIMPACT (<http://www.mitimpact.com/it/438-2/>), progetto nato nell'ambito del programma di cooperazione transfrontaliera (Interreg) Italia-Francia Alcotra, su un periodo di tre anni (2017 - 2020);
 - PREPAIR (<https://www.lifeprepare.eu/>), Po Regions Engaged to Policies of Air, progetto LIFE15 IPE IT 013 PREPAIR (CUP E69D17000580008), durata dal 2017 al 2024; **in particolare nel triennio 2022-2024 si prevede di utilizzare ai fini delle attività progettuali il servizio oggetto del presente capitolato per un tempo pari almeno 6 mesi.**
 - programma MIRROR COPERNICUS, nell'ambito della partecipazione del SNPA (Sistema Nazionale per la Protezione dell'Ambiente), durata 2021-2024.

Il sistema modellistico è una rete integrata di codici di calcolo, che utilizza in ingresso molteplici basi dati, che vengono elaborate da componenti specifiche del sistema, al fine di produrre il set di informazioni necessarie alle simulazioni dei modelli di qualità dell'aria. Le componenti principali del sistema possono essere schematicamente raccolte come:

- modulo territoriale: per la determinazione dei parametri geofisici e territoriali e la costruzione di campi utilizzati dagli ulteriori moduli;
- modulo meteorologico: per la produzione, a partire da informazioni di base (osservazioni al suolo ed in quota, campi tridimensionali, previsti o analizzati, di modelli meteorologici a mesoscala), dei campi bidimensionali delle variabili di turbolenza e dei campi tridimensionali delle variabili meteorologiche, sui domini di calcolo utilizzati dai modelli di qualità dell'aria; il modulo meteorologico può anche essere costituito, a seconda dell'applicazione, da un modello meteorologico numerico quale W.R.F (The Weather Research and Forecasting Model);
- modulo delle emissioni: per la costruzione dei campi bidimensionali orari dei ratei di emissione per tutte le specie chimiche trattate dai modelli di qualità dell'aria, a partire dalle informazioni aggregate contenute negli inventari delle emissioni disponibili;
- modulo per le condizioni iniziali ed al contorno: per la predisposizione dei campi di condizioni iniziali ed al contorno sui domini di simulazione, a partire da osservazioni di parametri chimici o da campi tridimensionali di concentrazioni simulati da modelli a mesoscala;
- modello di qualità dell'aria: per la generazione di campi tridimensionali di concentrazioni di inquinanti in atmosfera ed al suolo. Vengono simulati meccanismi di trasporto, dispersione, deposizione (secca e umida) e trasformazione chimica. Il modello di qualità dell'aria può operare su domini più o meno estesi e con risoluzioni molto elevate. Il modello di qualità dell'aria può essere di tipo euleriano (nello specifico FARM – Flexible Air Regionale Model) oppure lagrangiano (in particolare SPRAY – modello Lagrangiano a particelle);
- modulo di post-elaborazione: per l'elaborazione dei campi tridimensionali di concentrazione e la generazione di campi aggregati di indicatori, mappe tematiche e per la valutazione della capacità e prestazioni del sistema modellistico;
- modulo di data fusion: per effettuare l'integrazione, in modalità di analisi, dei campi di concentrazione prodotti dal modello euleriano con i dati osservati di qualità dell'aria ai fini di ottenere la migliore stima dello stato di qualità dell'aria;

- modulo per il source apportionment: utilizzato solo in modalità diagnostica non operativa, modulo che permette di quantificare, tramite un particolare utilizzo del modello euleriano, la stima dell'impatto di specifiche sorgenti (o insieme di sorgenti) sulla qualità dell'aria.

1.3. Sistema di calcolo ad elevate prestazioni

1.3.1. CLUSTER ARPAP

Nel corso del primo triennio di attività (2010-2011-2012) sono stati acquisiti gli apparati hardware per la realizzazione di un sistema di calcolo ad elevate prestazioni, cluster ARPAP, di proprietà ed utilizzo esclusivo di Arpa Piemonte, che è stato installato presso CINECA. La configurazione del Cluster ARPAP può essere schematizzata come segue:

- nodi di management: 1 HP DL380 GS
 - CPU = 1 E5405;
 - RAM = 4GB;
- nodi di calcolo: 32 HP DL 160 G6
 - CPU = E5520 quad core 2,26 Ghz;
 - RAM = 8 GB;
- apparati di interconnessione:
 - switch di tipo A: 2 Voltaire 4X DDR 24p INT Mngd Switch – infiniband;
 - switch di tipo B: 1 Switch gigabit ethernet HP ProCurve 2848 – infiniband;
 - switch di tipo C: 1 Switch gigabit ethernet HP ProCurve 2650 – infiniband;
- risorse di archiviazione: 1 NetApp FAS2020 12 TB;
- Sistema operativo: Red Hat RHEL 5.3;
- Compilatori: GNU 4.5, PGI 10.4, Intel 11.1, python 2.6.3, OpenMPI;
- HPC workload manager and job scheduler: PBSPro 10.1;
- Software environment manager: Modules Environment 3.1.6.

1.3.2. CLUSTER PLX

Nel triennio successivo la configurazione del Cluster ARPAP è rimasta invariata. Non sono stati apportati interventi di aggiornamento e di potenziamento ma ci si è limitati all'ordinaria manutenzione. Sono state progressivamente affiancate ulteriori risorse di calcolo, acquisendo come servizi l'accesso e l'utilizzo di risorse di calcolo di CINECA. In particolare, sono state utilizzate risorse di calcolo del Cluster PLX, un cluster condiviso, che espone pertanto risorse contemporaneamente verso molteplici utenti. Il principio di risorse ad uso esclusivo viene quindi superato dal concetto di risorse dedicate, in cui non vengono riservati specifici nodi di calcolo a determinati utenti (i nodi sono interscambiabili e vengono gestiti dal sistema secondo una logica di ottimizzazione delle performance) ma viene garantito l'utilizzo di un definito set di risorse (numero di core di calcolo per tempo di utilizzo) entro i limiti previsti contrattualmente per ciascuno degli utenti. Nel profilo dei differenti utenti viene definita la rispettiva priorità di accesso alle risorse; gli eventuali tempi di attesa (in seguito code) tra il momento della sottomissione di una applicazione (nel seguito job) e l'effettiva esecuzione, sono il prodotto della priorità dell'utente per l'onerosità (ovvero la richiesta di risorse da allocare) del job in questione (job particolarmente onerosi potranno essere sottomessi, indipendentemente dalla priorità dell'utente, solo se la disponibilità di risorse del sistema sarà adeguata).

- Descrizione: Super computer PLX-GPU, modello IBM PLX (iDataPlex DX360M3), con architettura Linux Infiniband Cluster.
- Nodi di calcolo: sono complessivamente disponibili ventotto (28) nodi di calcolo (14 nodi il primo anno, 7 nodi il secondo anno, 7 nodi il terzo anno), con le seguenti caratteristiche:
 - 2 six-cores Intel Westmere 2.40 GHz per nodo, ossia 12 core/nodo;
 - dotazione RAM 48 GB/node, ovvero 4GB/core;
- apparati di interconnessione: Infiniband with 4x QDR switches;

- risorse di archiviazione: vengono messe a disposizione, per tutta la durata del servizio, sia risorse dedicate, ad uso esclusivo delle utenze di Arpa collegate al servizio in oggetto, sia risorse in condivisione con le altre utenze del cluster:
 - *area di archiviazione dedicata* della dimensione di ventisei (26) Tb, sottoposta a backup giornaliero;
 - *area SCRATCH in condivisione* con gli altri utenti del cluster di calcolo, a supporto della realizzazione delle attività di calcolo dei sistemi modellistici.
- Sistema operativo: Red Hat RHEL 5.6;
- Compilatori: Fortran F90, C, C++ (Intel, PGI e GNU suites);
- HPC workload manager and job scheduler: PBSPro 10.1;
- Software environment manager: Modules Environment 3.1.6;

Nel corso del triennio il sistema modellistico viene installato e configurato sul Cluster *PLX* e l'operatività viene gradualmente migrata sulle nuove risorse di calcolo. Il Cluster *ARPAP* rimane dedicato prevalentemente ad attività di sviluppo e test, fino alla conclusione del relativo contratto, al termine del quale viene dismesso.

La possibilità di accedere in maniera dinamica alle risorse di un cluster di calcolo consente di modulare l'impiego di risorse sulla base delle effettive richieste dei jobs sottomessi, con un consumo (numero di core di calcolo per tempo di utilizzo) ritagliato sulle specifiche applicazioni. Nello stesso tempo si offre la possibilità di incrementare le risorse disponibili in occasione di specifiche esigenze. Si prevede pertanto di poter accedere, ad integrazione delle risorse dedicate, a risorse on-demand, attivate su specifica richiesta.

1.3.3. CLUSTER GALILEO (2015-2018)

Nel corso dell'anno 2015 le politiche di gestione ed evoluzione dei propri cluster di calcolo da parte di *CINECA* decretano il progressivo abbandono del Cluster *PLX* a favore del Cluster *GALILEO* (architettura 2-4 volte più veloce).

- Descrizione: super computer Tier-1 modello IBM NeXtScale con architettura Linux Infiniband Cluster. Viene introdotto nel gennaio 2015 (completamente operativo dal 02 febbraio 2015) e figura tra i più potenti super computer a disposizione dei ricercatori italiani ed europei in ambito privato e pubblico. È disponibile per la ricerca in ambito europeo come Tier-1 system dell'infrastruttura PRACE (www.prace-project.eu). Il super computer *GALILEO* è usato principalmente per sviluppare ed eseguire applicazioni con architetture ibride, sfruttando le applicazioni software nel campo della fluidodinamica computazionale, della scienza dei materiali, delle scienze ambientali e della geofisica.
- nodi di calcolo: disponibili 7 nodi di calcolo con le seguenti caratteristiche:
 - 2 x8-cores Intel Haswell 2.40 Ghz per nodo ossia 16 core per nodo per un totale complessivo di 112 core;
 - dotazione RAM 128 GB/node, ovvero 8 GB/core;
- apparati di interconnessione: Infiniband with 4x QDR switches;
- risorse di archiviazione: vengono messe a disposizione fino alla conclusione del contratto di servizio sia risorse di archiviazione dedicate, riservate alle utenze di Arpa, sia risorse di archiviazione condivise tra gli utenti del cluster:
 - area *WORK* di utilizzo riservato, con dimensione di un (1) Tb e sottoposta a backup giornaliero. In caso di necessità le dimensioni assegnate possono essere aumentate;
 - area *DRES* di utilizzo riservato, con dimensione di cinque (5) Tb sottoposta a backup giornaliero. L'area di archiviazione in questione non è un'area locale del server ma una risorsa condivisa, accessibile con i comuni protocolli di comunicazione in rete. In caso di necessità le dimensioni assegnate possono essere aumentate;
 - area *SCRATCH* in condivisione con gli altri utenti del cluster di calcolo, a supporto della realizzazione delle attività di calcolo dei sistemi modellistici.
- Sistema operativo: CentOS 7.0;
- Compilatori: Fortran F90, C, C++ (Intel, PGI e GNU suites);

- HPC workload manager and job scheduler: PBSPro 13.0.0;
- Software environment manager: Modules Environment 3.1.6;

Il passaggio al Cluster *GALILEO* ha determinato la necessità di migrare il sistema modellistico verso il nuovo ambiente di calcolo, simile nell'impostazione al precedente ma con sensibili differenze nella configurazione e nelle procedure di gestione.

1.3.4. CLUSTER GALILEO (2018-2021)

Nel gennaio 2018, *CINECA* ha effettuato importanti aggiornamenti sul Cluster *GALILEO*, che è stato riconfigurato con i nodi Intel Xeon E5-2697 v4 (Broadwell) derivati dall'esperienza sul Cluster *Marconi*. In conseguenza, le risorse dedicate ad ARPA Piemonte sono diventate:

- nodi di calcolo: disponibili 4 nodi di calcolo con le seguenti caratteristiche:
 - 2 x 18-cores Intel Xeon E5-2697 v4 at 2.30 GHz + 2 nVidia K80 GPUs, ossia 36 core per nodo per un totale complessivo di 144 core;
 - dotazione RAM 128 GB/node, ovvero 8 GB/core;
- apparati di interconnessione: Intel OmniPath (100Gb/s) high-performance network;
- risorse di archiviazione: le stesse risorse di archiviazione previste nella precedente configurazione, ma con un aumento, alle stesse condizioni contrattuali, delle dimensioni dell'area *WORK* a (3) TB.
- Sistema operativo: CentOS 7.4;
- Compilatori: Fortran F90, C, C++ (Intel, PGI e GNU suites);
- HPC workload manager and job scheduler: SLURM 2002;
- Software environment manager: Modules Environment 3.2.10.

Il passaggio alla nuova versione del Cluster *GALILEO* ha determinato la necessità di migrare il sistema modellistico verso il nuovo ambiente di calcolo, simile nell'impostazione al precedente ma con sensibili differenze nella configurazione.

Nel corso del 2021, ultimo anno di attività del contratto attualmente in essere, il Cluster *GALILEO* è uscito dalla produzione il 15 marzo, rimanendo attivo solo per le attività di ARPA Piemonte e pochi altri utenti, in attesa dell'entrata in produzione del nuovo cluster HPC *GALILEO100* prevista entro il mese di luglio 2021 e sul quale dovranno migrare tutti i servizi dedicati ad ARPA Piemonte. Tra le dotazioni previste per il Cluster *Galileo100*:

- nodi di calcolo:
 - 564 *computing server*
 - 348 *standard server*:
 - 2x CPU 8260 Intel Cascade Lake, 24 core, 2.4 GHz
 - 384 GB RAM DDR4 2933MT/s
 - 480 GB SSD
 - 180 *Data processing server*:
 - 2x CPU 8260 Intel Cascade Lake, 24 core, 2.4 GHz
 - 384 GB RAM DDR4 2933MT/s
 - 2 TB SSD
 - 1,5 TB Intel Optan
 - Interconnessione: Infiniband 100Gbs
- nodi di calcolo cloud:
 - 77 *computing server OpenStack*
 - Equipaggiati con 2x CPU 8260 Intel Cascade Lake, 24cores, 2.4 GHz
 - 768 GB RAM DDR4 2933MT/s
 - 2 TB SSD1 PB CEPH storage ad alte prestazioni dedicato (fullNVM/SSD)
 - Interconnessione: Ethernet 100Gbs

- risorse di archiviazione:
 - 20,5 PB di *actives torage*, accessibile sia dalla partizione *cloud*, che dai server HPC
 - 5 PB di *fast storage* con funzionalità di *fast caching* di tutto il traffico I/O
- HPC workload manager and job scheduler: SLRUM;
- Software environment manager: Spack.

ART 2 - OGGETTO DELL'APPALTO

L'oggetto del presente appalto è l'acquisizione dei **servizi necessari alle attività di calcolo ad elevate prestazioni necessarie al Dipartimento tematico Rischi Naturali e Ambientali di Arpa Piemonte per lo svolgimento delle attività di modellistica numerica per applicazioni in campo ambientale, con impiego di competenze specialistiche e risorse computazionali dedicate.**

I servizi oggetto del presente appalto hanno durata di **trentasei (36) mesi a partire dalla data di attivazione del contratto, rinnovabile di ulteriori trentasei (36) mesi.** I servizi si configurano come *risorse dedicate* erogate tramite le risorse ed il personale dell'Aggiudicatario, con la possibilità di integrazione attraverso l'attivazione di servizi *on demand* e prevedono l'accesso e l'utilizzo degli apparati che ne consentono l'erogazione.

La descrizione e la definizione dei servizi e degli apparati coinvolti sono contenute rispettivamente negli articoli 3, 4, 5 e 6 del presente capitolato speciale d'appalto.

ART 3 - DESCRIZIONE DEI SERVIZI OGGETTO DELL'APPALTO

Nel presente paragrafo vengono esposti i termini, le definizioni e la descrizione dei servizi che sono oggetto del presente appalto.

3.1. Nodo equivalente

Viene definito *nodo equivalente* un nodo dedicato al calcolo, che soddisfi specifici requisiti di potenza e velocità, e che corrisponda ad una determinata disponibilità di core fisici per il calcolo. Un *nodo equivalente* può essere un nodo di calcolo reale di un cluster, oppure un nodo logico, costituito dalla composizione di più nodi reali oppure da una frazione di un nodo reale. Le richieste sulla disponibilità di core di calcolo per i *nodi equivalenti* sono determinate dalle caratteristiche di parallelizzazione dei codici del sistema modellistico, che sfrutta sia direttive a memoria condivisa che direttive a memoria distribuita. Le caratteristiche richieste di potenza, velocità e disponibilità di core per un nodo equivalente sono specificate nel successivo paragrafo 3.2.

3.2. Unità di servizio

La quantificazione dei servizi richiesti viene espressa attraverso l'introduzione della definizione di *unità di servizio (US)*:

3.2.1. UNITÀ DI SERVIZIO IN DEDICATO (USD):

I servizi erogati in dedicato vengono quantificati secondo la seguente metrica:

$$1 \text{ USD} == 1 \text{ nodo equivalente} / \text{ mese}$$

3.2.2. UNITÀ DI SERVIZIO A CONSUMO (USC):

Per fare fronte ad esigenze di consulenza o supporto specialistico od altrimenti a esigenze straordinarie o comunque non preventivabili di risorse di calcolo aggiuntive, relative a periodi circoscritti e limitati o riconducibili a specifiche attività progettuali è possibile attivare delle *unità di servizio a consumo*, che vengono rendicontate e fatturate con periodicità predefinita (ogni tre o ogni sei mesi). Le risorse utilizzate *on-demand* vengono quantificate secondo la seguente metrica:

$$1 \text{ USC} == 1 \text{ nodo equivalente} / \text{ mese}$$

3.3. Ore di calcolo

Si definisce *ore di calcolo* il tempo effettivo di utilizzo dei *core di calcolo* sui clusters. Le *ore di calcolo* costituiscono la metrica per la quantificazione del consumo da parte delle attività dei singoli utenti e dei singoli *jobs* sottomessi, e vengono rendicontate dai sistemi di management dei cluster.

Sono valide le relazioni:

$$1 \text{ USD} == 11520 \text{ ore di calcolo}$$

1 USC == 11520 ore di calcolo

Il riepilogo del consumo delle *ore di calcolo* deve essere oggetto di opportuna reportistica, redatta ed inviata periodicamente (frequenza almeno mensile) al personale di Arpa coinvolto nel servizio, a cura dell'Aggiudicatario. Modalità e tempistiche della reportistica vengono definite in collaborazione con il personale di Arpa coinvolto nel servizio.

3.4. Utenze e accesso

Devono essere fornite le utenze sui cluster di calcolo (in seguito definite *users*), nel numero e con le caratteristiche necessarie alla implementazione ed al completo utilizzo dei sistemi modellistici.

Gli *users* devono essere forniti in numero non inferiore a quattro (4), e devono essere abilitati all'accesso ai nodi di login dei cluster, nei limiti e con le prerogative esposte nel successivo capitolo **ART.4**. Ciascuno degli *users* deve disporre di una *home* personalizzata di dimensione non inferiore a cento Gigabyte (100 Gb), con le consuete dotazioni di sicurezza e privacy (protezione dei contenuti, modifica e recupero delle credenziali di accesso, tracciamento delle connessioni e dei tempi di connessione). Tutti gli *users* devono avere accesso completo a tutte le risorse hardware e software previste nella fornitura del presente appalto (nodi di calcolo, aree di storage condivise o private, compilatori, librerie e moduli software, sistemi di management e schedulazione dei *jobs* e delle *code*, etc..).

Gli *users* devono essere personali, per ciascuno dei tecnici Arpa impiegati nelle attività, affiancati da ulteriori *utenze operative* (in seguito *operative users*), con privilegi opportunamente limitati, dedicate allo svolgimento di specifiche mansioni (esecuzione delle procedure operative, monitoraggio dei processi e delle risorse, reportistica, alert...).

Deve essere possibile, previo accordo tecnico con gli amministratori dei clusters, la definizione di regole di protezione e/o condivisione (creazione e condivisione di gruppi, regole ACL, etc..) delle risorse, in maniera pubblica oppure privata tra gli *users*. Deve essere possibile, previo accordo tecnico con gli amministratori dei clusters, la creazione di progetti e/o sotto-progetti, in cui organizzare le attività svolte, con la possibilità di assegnare gli *users* a uno o più progetti / sotto-progetti.

Si richiede la possibilità di individuare e configurare uno o più *users* cui assegnare specifiche autorizzazioni per la gestione delle risorse e dei progetti e sotto-progetti.

La disponibilità degli *users* personali e operativi sui cluster di calcolo viene garantita 7gg 24h, salvo attività ordinarie o straordinarie legate alla gestione e manutenzione degli apparati, che vengono tempestivamente comunicate.

3.5. Suite operative delle catene modellistiche

Il presente appalto prevede l'acquisizione e la messa in opera dei servizi per il calcolo ad elevate prestazioni dedicati alla modellistica numerica ambientale, tramite la disponibilità delle risorse e l'implementazione degli ambienti dedicati alle attività operative. Coprono particolare importanza le *suite operative* che consentano l'esecuzione, il monitoraggio ed il controllo delle catene modellistiche.

3.5.1. MODELLISTICA PER LA QUALITÀ DELL'ARIA

La catena modellistica per la qualità dell'aria viene descritta nelle sue componenti e nel dettaglio del suo *workflow* nel successivo capitolo **ART.5**. Attualmente è implementata una *suite operativa* della catena presso l'Aggiudicatario del servizio in essere, con le seguenti caratteristiche:

- utilizzo del linguaggio di programmazione python per l'avvio coordinato (rispetto delle precedenze di esecuzione e dei prerequisiti di funzionamento) dei moduli della modellistica, che utilizzano il linguaggio di programmazione Fortran;
- integrazione con *software environment manager* e *workload manager and job scheduler* per la configurazione dell'ambiente di esecuzione e per la sottomissione e gestione dei *job*, sia seriali che paralleli, dei moduli modellistici;

Il presente appalto richiede, sulla base delle caratteristiche della catena modellistica e delle risorse attualmente in uso, la realizzazione e l'implementazione di un'opportuna *suite operativa*:

- utilizzo del linguaggio di programmazione python (versione ≥ 3.9);
- gestione del *workflow* della catena modellistica per il controllo delle precedenze, per la verifica dei prerequisiti e per la sottomissione dei moduli modellistici; si propone l'utilizzo di strumenti

specifici per tale funzionalità e si citano a puro titolo di esempio alcuni prodotti: *FireWorks*, *Pyflow*, *Apache Airflow*, *Apache Taverna*. È facoltà dell'aggiudicatario proporre ed adottare soluzioni tecnicamente differenti, nel rispetto dei requisiti richiesti;

- lo sviluppo della *suite operativa* deve essere coordinato in ogni sua fase con il personale di Arpa coinvolto nelle attività e condotto con modalità *training on the job*;
- è richiesto l'utilizzo di *collaborative development environment* (si citano come esempio *GitHub*, *GitLab*) per lo sviluppo e la manutenzione dei codici della suite operativa;

In caso di adozione di dotazioni alternative a quanto proposto o citato, l'Aggiudicatario è tenuto, qualora ritenuto necessario dal personale Arpa coinvolto nelle attività, ad effettuare una specifica formazione del personale, relativamente alle dotazioni adottate.

3.5.2. MODELLISTICA METEOROLOGICA

La catena modellistica meteorologica non prevede una suite operativa ma la possibilità di eseguire ri-analisi e ri-forecast di casi studio di interesse nazionale, individuati in accordo con il DPC. Due sono i modelli utilizzati a tale scopo, *COSMO* e *ICON*, il cui *workflow* comprende anche i moduli *INT2LM* (*COSMO*) e le *DWD ICON Tools* (*REMAP*, *RESUB* e *fieldextra*). Sono necessari *software environment manager*, *workload manager* e *job scheduler* per la configurazione dell'ambiente di esecuzione e per la sottomissione e gestione dei *job*, sia seriali che paralleli, dei vari moduli modellistici.

Il presente appalto richiede un supporto per l'installazione dei moduli modellistici citati.

3.6. **Apparati**

Il presente appalto prevede i servizi di accesso agli apparati che erogano i servizi di calcolo in oggetto. Gli apparati devono essere accessibili ed operativi 7gg 24h, salvo attività ordinarie o straordinarie legate alla gestione e manutenzione dei sistemi o degli apparati, che vengono tempestivamente comunicate. La descrizione degli apparati è rimandata al successivo *ART.4*.

Relativamente alle risorse di storage, dei singoli *users* o condivise tra gli *users* nell'ambito di progetti e sotto-progetti, sono previsti servizi di *backup* (integrali e/o differenziali), con periodicità almeno giornaliera e *tempo di ritenzione* pari all'intero periodo di estensione del presente appalto.

3.7. **Supporto all'utenza**

Viene richiesta la garanzia di supporto di tipo sistemistico e sistemistico-applicativo nelle attività di predisposizione dei sistemi modellistici sui *clusters* di calcolo, nelle attività di utilizzo operativo, nelle manutenzioni ordinarie ed evolutive dei sistemi per il periodo di durata del presente appalto. In particolare, vengono garantiti i seguenti interventi:

- supporto in attività di implementazione, configurazione e avvio dei sistemi modellistici, a seguito di nuove attività, porting, manutenzioni ordinarie o straordinarie;
- supporto nell'utilizzo, configurazione e personalizzazione di strumenti di sviluppo ed analisi dei codici, e di strumenti di management e schedulazione di *jobs* e *code*;
- supporto nelle attività di monitoraggio, di analisi e di verifica del funzionamento dei sistemi modellistici nella fase operativa, per l'individuazione di malfunzionamenti o difetti di implementazione, inefficienza dei codici o dei sistemi implementati;
- supporto e consulenza nella correzione di malfunzionamenti, difetti ed inefficienze nella implementazione e configurazione dei codici e dei sistemi modellistici.

Vengono definite le procedure e le modalità di attivazione dei servizi all'utenza da parte del personale di Arpa Piemonte interessato. Il servizio viene erogato dal personale dell'ente appaltatore nei giorni feriali e nell'orario di ufficio, compatibilmente con la disponibilità del personale e l'onerosità delle richieste; modalità e tempistiche di erogazione del servizio vengono concordate tra il personale dell'Agenzia interessato ed il personale dell'ente appaltatore.

L'assistenza all'utenza, in tutte le sue forme, verbali o scritte, dirette od indirette, compreso tutto il materiale afferente (documentazione, manualistica, esercitazioni, reportistica, etc.) indipendentemente dal supporto tramite cui viene fornita, **deve sempre essere erogata tassativamente in lingua italiana.**

3.8. **Supporto specialistico**

È possibile avvalersi di un supporto specialistico negli ambiti di:

- modifica e correzione codici;
- implementazione di nuove funzionalità;
- parallelizzazione dei codici;
- sviluppo ed ingegnerizzazione di soluzioni personalizzate per la realizzazione ed automatizzazione di attività;

I servizi di supporto specialistico vengono richiesti ed attivati *on-demand* e vengono rendicontati a consumo. Nella quantificazione e rendicontazione di tali servizi si applica la metrica derivante dalla seguente uguaglianza:

$$1 \text{ giornata-supporto} == 1 \text{ USC}$$

3.9. Supporto alla formazione

È possibile avvalersi di un supporto alla formazione, in *modalità di training on the job* per l'approfondimento di competenze già acquisite da parte del personale di Arpa Piemonte, ed in *modalità di corso/lezione*, per l'introduzione di nuove competenze. La descrizione delle attività di formazione è contenuta nel successivo ART.5.

I servizi di supporto alla formazione vengono richiesti ed attivati *on-demand* e vengono rendicontati a consumo. Nella quantificazione e rendicontazione di tali servizi si applica la metrica derivante dalla seguente uguaglianza:

$$1 \text{ giornata-supporto} == 1 \text{ USC}$$

La formazione al personale dell'Agenzia, compreso tutto il materiale afferente (documentazione, manualistica, esercitazioni, reportistica, etc.), indipendentemente dal supporto tramite cui viene fornita, **deve sempre essere erogata tassativamente in lingua italiana**. Può fare eccezione il materiale non predisposto specificamente per la formazione in oggetto ovvero il materiale predisposto da terzi (libri di testo, libri di esercitazione, tutorials, video-tutorials, etc.), ritenuto interessante ed inerente all'argomento oggetto della formazione ed indicato a completamento ed approfondimento delle tematiche trattate, solo se viene utilizzata la lingua inglese come lingua alternativa.

ART 4 - CARATTERISTICHE TECNICHE DELLE DOTAZIONI ED APPARATI CORRELATI AL SERVIZIO

4.1. Ambiente

4.1.1. SISTEMA OPERATIVO

è richiesto un sistema operativo Linux. Il personale dell'Agenzia nel corso degli anni ha sviluppato esperienza e competenza nell'ambito delle distribuzioni Red Hat, CentOS, Fedora che vengono indicate come scelta preferenziale. Sono considerate distribuzioni alternative, su indicazione dell'Aggiudicatario, se riconosciute equivalenti e/o migliorative. In caso di adozione di distribuzioni alternative l'Aggiudicatario è tenuto, qualora sia ritenuto necessario dal personale Arpa coinvolto nelle attività, ad effettuare una specifica formazione del personale, relativamente alla distribuzione individuata.

4.1.2. LINGUAGGI E COMPILATORI

Vengono utilizzati prevalentemente i linguaggi di programmazione Fortran (Fortran 77, Fortran 90, Fortran 95), C e Python, e sono richiesti i compilatori:

- compilatori C; *GNU, Intel, PGI* (almeno uno tra *Intel* e *PGI*);
- compilatori Fortran: *GNU, Intel, PGI* (almeno uno tra *Intel* e *PGI*);
- Python: *python* release $\geq 2.7.9$ e *python3* release ≥ 3.9

È inoltre richiesta la presenza e la disponibilità dei seguenti tools:

- *Build System tools* e *package management tools*: in particolare *autotools, Cmake, RPM, setuptools*;
- *Parallel environment*: in particolare *MPI* e *OpenMPI*;
- *Tools per debugging e profiling*: *GDB*, tools proprietari di *Intel* e *PGI*;

Sono prese in considerazione dotazioni alternative su indicazione dell'Aggiudicatario, se riconosciute migliorative o strettamente necessarie. In caso di adozione di dotazioni alternative

l'Aggiudicatario è tenuto ad effettuare una specifica formazione del personale, relativamente alle dotazioni adottate.

4.1.3. HPC WORKLOAD MANAGER AND JOB SCHEDULER

Viene richiesta la disponibilità di sistemi di *workload manager and job scheduler* per la sottomissione dei *job* e l'amministrazione dei processi in esecuzione. In particolare, il personale Arpa ha sviluppato esperienze e competenze nell'utilizzo di *PBSPRO* e *SLURM Workload Manager*. Sono prese in considerazione soluzioni alternative su indicazione dell'Aggiudicatario, se riconosciute equivalenti e/o migliorative. In caso di adozione di dotazioni alternative l'Aggiudicatario è tenuto, qualora ritenuto necessario dal personale Arpa coinvolto nelle attività, ad effettuare una specifica formazione del personale, relativamente alle dotazioni adottate.

4.1.4. SOFTWARE ENVIRONMENT MANAGER

È richiesta la disponibilità di un sistema di *environment manager* per gestire la configurazione di esecuzione delle catene modellistiche in differenti situazioni. Il personale Arpa, nel corso delle precedenti esperienze di appalto ha maturato competenza nell'utilizzo di *Modules environment*. Sono prese in considerazione soluzioni alternative su indicazione dell'Aggiudicatario, se riconosciute equivalenti e/o migliorative. In caso di adozione di dotazioni alternative l'Aggiudicatario è tenuto, qualora ritenuto necessario dal personale Arpa coinvolto nelle attività, ad effettuare una specifica formazione del personale, relativamente alle dotazioni adottate.

4.1.5. COLLABORATIVE DEVELOPMENT ENVIRONMENT (CDE)

È richiesta la disponibilità di sistemi *CDE* per il repository e per la gestione e manutenzione degli sviluppi e dei codici dei moduli modellistici. I *CDE* devono assolvere al minimo alla richiesta delle seguenti funzionalità:

- *version control system*;
- *bug tracking system*;
- *todo list*;
- *document management system*;
- Si citano a titolo di esempio *GitHub* e *GitLab*; si valutano le soluzioni proposte dall'Aggiudicatario che è tenuto ad effettuare una specifica formazione del personale di Arpa coinvolto nell'attività, relativamente alle soluzioni adottate.

I sistemi *CDE* implementati devono essere accessibili dalle utenze del personale di Arpa coinvolto nelle attività, internamente ed esternamente ai sistemi dell'aggiudicatario, devono altresì essere privati, ovvero accessibili in maniera esclusiva al personale autorizzato e secondo le *policy* definite e condivise dal personale Arpa.

4.1.6. LIBRERIE

si richiede in particolare la disponibilità delle librerie *NetCDF*, *HDF5*, *gdal*, *proj4*, *numpy*, *scipy*, *pandas*, *grib_api*, *udunits*. Nel caso in cui sia ritenuto inopportuno inserire tali librerie nelle dotazioni dell'ambiente condiviso degli utenti dei clusters di calcolo, deve essere consentito al personale Arpa coinvolto nelle attività di effettuare le installazioni in ambiente privato, e deve essere fornito il necessario supporto per l'effettuazione delle installazioni.

Nel caso in cui, durante la durata del contratto di appalto, si rendesse opportuno, per motivi tecnici, per motivi strategici, o a seguito di decisioni interne del soggetto Aggiudicatario, provvedere a modifiche sostanziali della configurazione iniziale dell'ambiente, l'Aggiudicatario è tenuto, qualora ritenuto necessario dal personale Arpa coinvolto nelle attività, ad effettuare una specifica formazione del personale, relativamente alle modifiche introdotte.

4.2. Processori

Il numero e tipologia dei processori messi a disposizione devono essere sufficienti a garantire l'effettuazione delle attività operative della Struttura entro i tempi previsti e consentire attività di test e verifica dei codici sviluppati ed implementati. In particolare, numero e tipologia dei processori devono garantire le massime prestazioni per i sistemi modellistici implementati, in considerazione delle tipologie di parallelizzazione utilizzate:

modellistica dispersiva: paradigma *MPI* (memoria distribuita), paradigma *OpenMP* (memoria condivisa), implementazione ibrida;

modellistica meteorologica: paradigma *MPI*, paradigma *OpenMP*.

Il numero e la tipologia dei processori utilizzati devono quantitativamente costituire una soluzione confrontabile e/o migliorativa delle attuali dotazioni disponibili su *Cluster GALILEO100* presso CINECA:

- 24 cores 8260 IntelCascadeLake 2.40 Ghz per nodo;
- dotazione RAM 384 GB DDR4 2933MT/s /node;

L'accesso e l'utilizzo delle risorse di calcolo, laddove non possa essere riservato ed esclusivo per le utenze ARPA facenti riferimento al presente appalto, in ragione della natura intrinseca di un cluster di calcolo condiviso tra utenti differenti, deve essere garantito e prioritario, e comunque tale da minimizzare le code nella sottomissione di *jobs*, e di escludere situazioni di indisponibilità delle risorse:

- le code di sottomissione non devono superare l'intervallo di alcuni minuti (massimo 20 minuti in riferimento ai *run operativi* dei sistemi modellistici);
- la richiesta di disponibilità elevata di risorse (entro i limiti definiti dal contratto di servizio) non deve aumentare le code di sottomissione dei *jobs*;

L'Aggiudicatario è tenuto ad effettuare, in collaborazione con il personale di Arpa coinvolto nelle attività, le prove necessarie e utili ad individuare la configurazione ottimale (numero e tipologia di nodi/*cores*, utilizzo e gestione della memoria distribuita/condivisa, configurazione dell'ambiente di utilizzo) per l'esecuzione operativa del modello di qualità dell'aria *FARM*.

4.3. Aree di archiviazione

Si richiedono risorse di archiviazione a supporto delle attività di modellistica condotte, con le seguenti caratteristiche:

4.3.1. AREA DI LAVORO IN DEDICATO

deve essere costituita da risorse del cluster di calcolo, accessibile come risorsa locale durante l'esecuzione dei codici di calcolo. L'area di lavoro in dedicato deve essere accessibile a tutti gli *users* e riservata a queste sole utenze; deve quindi essere possibile, previo accordo tecnico con gli amministratori dei *clusters*, la definizione di regole di protezione e/o condivisione (creazione e condivisione di gruppi, regole ACL, etc..) delle risorse, in maniera pubblica oppure privata tra gli *users*.

L'area di lavoro deve avere dimensione non inferiore ai cinque Tera Byte (5 Tb) e deve essere sottoposta a servizio di *backup* e di *restore* dei contenuti, con una frequenza almeno quotidiana, e con tempi di ritenzione pari all'intera durata del servizio;

4.3.2. AREA DI REPOSITORY IN DEDICATO

è costituita da un'area di archiviazione di dimensione non inferiore ai venti Tera Byte (20 Tb), e deve essere sottoposta a servizio di *backup* e di *restore* dei contenuti, con una frequenza almeno quotidiana, e con tempi di ritenzione pari all'intera durata del servizio.

L'area di *repository* deve essere accessibile a tutti gli *users* ma non si richiede l'accesso operativo come risorse locale durante l'esecuzione dei codici. L'accesso all'area di *repository* deve esser riservato alle utenze di Arpa, e deve essere possibile, previo accordo tecnico con gli amministratori dei *clusters*, la definizione di regole di protezione e/o condivisione (creazione e condivisione di gruppi, regole ACL, etc..) delle risorse, in maniera pubblica oppure privata tra gli *users*.

4.3.3. AREA SCRATCH CONDIVISA TRA LE UTENZE

deve essere riservata agli *users* un'area di *SCRATCH* tra le risorse condivise tra gli utenti del cluster, che sia di supporto funzionale nelle fasi di esecuzione dei sistemi modellistici. Non sono richiesti servizi di *backup* e non sono prefissate dimensioni.

4.4. Accessibilità

Gli *users* sul cluster di calcolo devono essere acceduti via internet in maniera sicura, attraverso il riconoscimento delle credenziali di sicurezza fornite. Se previsto dalle *policy* di sicurezza dell'Aggiudicatario, l'accesso agli *users* sui cluster di calcolo può essere limitato a specifici indirizzi IP. L'accesso agli *users* via internet deve avvenire attraverso protocolli di rete sicuri (SSH). Gli *users* devono essere abilitati, nel rispetto delle *policy* di sicurezza dell'Aggiudicatario, ad accedere alle risorse internet.

4.5. Continuità del servizio

Il servizio di calcolo ad alte prestazioni, in tutte le sue componenti e relativamente a tutte le risorse che ne fanno parte, deve essere erogato in alta affidabilità e deve essere disponibile in modalità continuativa 7gg 24h:

- le attività di manutenzione ordinaria e/o straordinaria che determinino una limitazione delle funzionalità del cluster di calcolo devono essere programmate con ampio anticipo e tempestivamente comunicate al personale dell'Agenzia;
- interruzioni impreviste e *fault* dei servizi non devono impedire l'accesso agli apparati ed alle dotazioni e non devono di fatto impedire l'operatività dei sistemi modellistici per un intervallo continuativo di tempo superiore alle quarantotto (48) ore;
- il numero complessivo delle interruzioni registrate, programmate o dovute ad imprevisti, non deve eccedere il numero di dodici (12) nel corso di un anno;
- in caso di interruzioni impreviste, *fault* dei servizi e malfunzionamenti che compromettano l'operatività dei sistemi modellistici sul cluster di calcolo, l'Aggiudicatario è tenuto a fornire l'assistenza ed il supporto specialistico al personale Arpa, per il ripristino tempestivo della completa funzionalità dei sistemi.
- In caso di interruzioni maggiori di quarantotto (48) ore, il contraente deve essere in grado di fornire risorse di calcolo alternative equivalenti, ossia in grado di soddisfare completamente tutti i requisiti del contratto.

ART 5 - WORKFLOW SISTEMA MODELLISTICO ATTUALMENTE OPERATIVO

5.1. Introduzione

Il sistema modellistico di ARPA Piemonte attualmente implementato ed operativo sul cluster HPC (*High Performance Computing*) GALILEO del dipartimento S.C.A.I (*Super Computing Applications and Innovation*) di CINECA si compone di differenti catene operative:

- la catena modellistica principale (denominata nel seguito *progn_g2_ope*), operativa sul dominio sovra regionale e sul dominio target regionale (rispettivamente in blu ed in nero nella successiva Figura 1) descritta a livello di *workflow* nel paragrafo 5.3;
- la catena modellistica ad alta risoluzione (*progn_g3_ope*), attualmente operativa sull'area metropolitana di Torino e sulla provincia di Novara (in rosso nella successiva figura), descritta a livello di *workflow* nel paragrafo 5.4.

La catena *progn_g2_ope* è indipendente dalla catena *progn_g3_ope*; quest'ultima necessita invece degli output della precedente per poter essere eseguita.

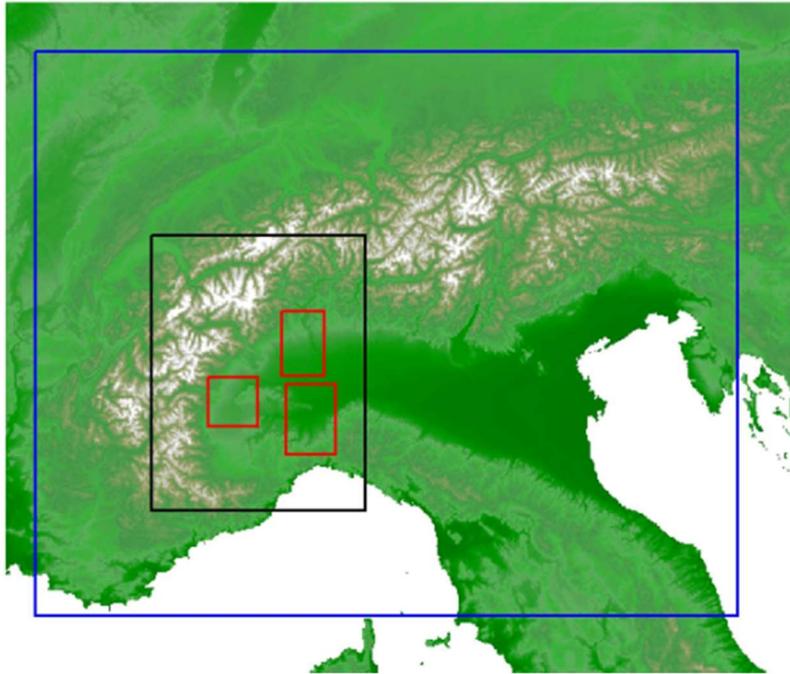


Figura 1

La gestione delle catene *progn_g2_ope* e *progn_g3_ope* avviene tramite una procedura python, implementata e sviluppata presso il Dipartimento Rischi Ambientali e Naturali di ARPA Piemonte; la procedura è la medesima per tutte e catene, ma viene utilizzata con ambiente di esecuzione e configurazioni differenti.

La procedura gestionale viene eseguita con *jobs* gestiti da *SLURM WORLOAD MANAGER* (nell'attuale configurazione del cluster HPC *GALILEO*). Ciascuna catena modellistica è eseguita tramite due differenti *jobs*:

- *Sope* che gestisce tutti i processi seriali relativi alla acquisizione ed elaborazione dei differenti dati di ingresso al modello di qualità dell'aria FARM;
- *Pope*, che gestisce i processi paralleli (ibridi OpenMP-MPI) relativi alla esecuzione del modello di qualità dell'aria FARM, alla post elaborazione ed archiviazione dei risultati.

I *jobs* sopra indicati hanno priorità nell'esecuzione e vengono avviati senza tempo di attesa in coda, tranne che in caso di contemporaneo utilizzo di tutte le risorse di calcolo a disposizione di ARPA Piemonte sul cluster *GALILEO*.

La schedulazione ed il lancio dei processi avviene quotidianamente tramite un *job* di schedulazione, *progn_scheduler.job*, il cui funzionamento è descritto nel successivo paragrafo 5.2.

5.2. Scheduler

Nel diagramma di Figura 2 è schematizzato il flusso di lavoro dello scheduler dei processi *progn_scheduler.job*:

- tra ora di inizio schedulazione (04:00, primo avvio del job) ed ora di fine schedulazione (14:00), il *job* si avvia e legge dal file di configurazione i processi da lanciare, nella configurazione attuale i processi sono sei (*Sope_to_g2*, *Pope_to_g2* per la catena *progn_g2_ope*; *Sope_to_g3*, *Pope_to_g3*, *Sope_no_g3*, *Pope_no_g3* per le catene *progn_g3_ope*);
- per ciascuno dei processi sopraelencati lo scheduler:
 - controlla se il processo è già stato eseguito con esito positivo;
 - controlla se il processo è in esecuzione o è in coda;
 se nessuna delle precedenti condizioni è verificata allora il processo viene lanciato tramite *workload manager*;
- fino a quando dal controllo non risulta che tutti i processi elencati nel file di configurazione sono terminati con esito positivo, *progn_scheduler.job* si risottomette automaticamente ogni 20 minuti;
- quando tutti i processi sono terminati correttamente o è stata superata l'ora di fine schedulazione, *progn_scheduler.job* si rischedula per le 04:00 del giorno successivo.

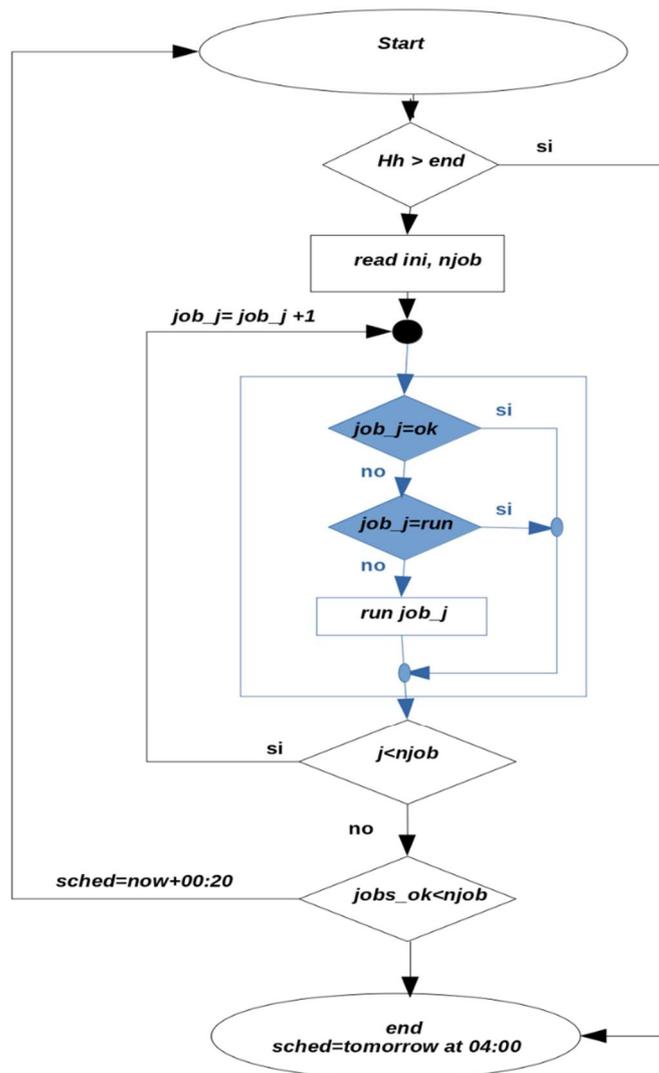


Figura 2. workflow dello schedulatore progn_scheduler.job

5.3. Catena modellistica principale (progn_g2_ope)

Come detto nel precedente paragrafo 5.1 la catena modellistica *progn_g2_ope* è eseguita tramite due *jobs*, *Sope_to_g2* per i processi seriali e *Pope_to_g2* per i processi paralleli.

5.3.1. SOPE TO G2

Nel dettaglio *Sope_to_g2* esegue in sequenza i seguenti moduli, secondo lo schema riportato in Figura 3:

1. *emma*:

processa i dati dell'inventario delle emissioni e prepara i campi di emissione in input al modello di qualità dell'aria per tutte le scadenze di simulazione (72 ore);

2. *getPrevAir*:

acquisisce i dati di previsione di qualità dell'aria messi a disposizione dal sistema Prev'Air (www.prevail.org) su un'utenza riservata ad ARPA Piemonte tramite protocollo http;

3. *chimere2farm*:

processa i dati del sistema Prev'Air per la costruzione delle *boundary conditions* (BC) necessarie al modello di qualità dell'aria per tutte le scadenze di simulazione (72 ore). Nel caso in cui l'acquisizione dal sistema Prev'Air non vada a buon fine entro un'ora prefissata (08:00, ora locale), le BC vengono elaborate a partire da un dataset climatologico predisposto da ARPA Piemonte e residente in locale su *GALILEO*;

4. *getCOSMO3D*:

acquisisce i volumi 3D del modello COSMO-I5, modello di riferimento del *DIPARTIMENTO DI PROTEZIONE CIVILE*; i dati di COSMO vengono acquisiti direttamente su *GALILEO* dalle utenze operative di CINECA/ARPAE che eseguono quotidianamente il modello;

5. *grb2gap/gap/surfproMegan*:

processa per tutte le scadenze di simulazione (72 ore) i dati di COSMO-I5 per costruire i dati di ingresso che vanno a costituire il driver meteorologico (variabili dinamiche e termiche – *gap* - e di turbolenza - *surfproMegan*) per il modello di qualità dell'aria.

I moduli sono indipendenti tra loro; vengono eseguiti tutti a seguito di controlli preliminari che non siano già stati eseguiti con successo; per il *modulo 3* è previsto un dataset di backup nel caso in cui i dati del *modulo 2* non siano disponibili, mentre il *modulo 5* viene eseguito solo se il *modulo 4* è terminato con esito positivo, altrimenti (esito negativo del *modulo 4*, dati di COSMO-I5 non disponibili), il *modulo 5* non può essere eseguito. Se tutti i moduli terminano con esito positivo, il *job* termina con esito positivo, altrimenti termina con messaggio di *warning* o di *error*.

5.3.2. POPE TO G2

Pope_to_g2 ha il compito di eseguire il modello di qualità dell'aria (*FARM*), effettuare le necessarie post-elaborazioni e procedere all'archiviazione dei dati prodotti. *Sope_to_g2* lavora in blocco su tutto il periodo di simulazione mentre *Pope_to_g2* lavora con una logica che privilegia i giorni di previsione; i moduli vengono eseguiti secondo la sequenza nel seguito descritta (e rappresentata in Figura 3) per ogni singolo giorno di previsione e per entrambi i domini su cui è applicata la catena modellistica, ovvero prima tutti i moduli per il primo giorno su tutti i domini, poi per il secondo giorno su tutti i domini ed infine per il terzo su tutti i domini:

1. *FARM*:

esegue il controllo sulla presenza degli output del modello di qualità dell'aria per il giorno in corso:

- nel caso siano presenti il modulo si interrompe con buon esito ed il *job* passa al modulo successivo;
- nel caso non siano presenti, il modulo esegue il check sulla disponibilità di tutti i files di input necessari al modello di qualità dell'aria:
 - in caso di esito positivo lancia in modalità parallela ibrida (MPI-OpenMP) il modello di qualità dell'aria *FARM*;
 - in caso di esito negativo, il *job* si interrompe con messaggio di *error*;

2. *farm2bc*:

esegue il controllo sulla presenza degli output del modulo per il giorno in corso:

- nel caso siano presenti il modulo si interrompe con buon esito ed il *job* passa al modulo successivo;
- nel caso non siano presenti, il modulo esegue il controllo sulla disponibilità dei dati di output del modello di qualità dell'aria:
 - in caso di esito positivo, i dati di concentrazione tridimensionali vengono processati per creare le condizioni iniziali ed al contorno per i domini ad alta risoluzione delle catene *progn_g3_ope* (o su domini configurabili dall'utente);
 - in caso di esito negativo, il modulo si interrompe con messaggio di *warning* ed si passa al modulo successivo;

3. *slicer*:

esegue il controllo sulla presenza degli output del modulo per il giorno in corso:

- nel caso siano presenti il modulo si interrompe con buon esito ed il *job* passa al modulo successivo;
- nel caso non siano presenti, il modulo esegue il controllo sulla disponibilità dei dati di output del modello di qualità dell'aria e dei dati di input meteorologici:
 - in caso di esito positivo, i dati meteorologici e di concentrazione tridimensionali vengono processati per creare archivi bidimensionali con le sole variabili di interesse;
 - in caso di esito negativo, il *job* si interrompe con messaggio di *error*;

4. *PostElab*:

esegue il controllo sulla presenza degli output del modulo per il giorno in corso:

- nel caso siano presenti il modulo si interrompe con buon esito ed il *job* passa al modulo successivo;
- nel caso non siano presenti il modulo effettua il controllo sulla disponibilità dei dati prodotti dal modulo 3 relativamente alle concentrazioni di qualità dell'aria:
- in caso di esito positivo il modulo provvede a creare gli indicatori di legge previsti dalle configurazioni della catena modellistica;
- in caso di esito negativo, il *job* si interrompe con messaggi di *error*;

5. *archive*:

esegue il controllo sull'avvenuta archiviazione dei dati della corsa della catena modellistica:

- se il risultato è positivo, il modulo si interrompe ed il *job* termina con esito positivo;
- se il risultato è negativo, il modulo esamina la disponibilità di tutti i dati prodotti dai precedenti moduli 2, 3, 4:
- in caso di esito positivo, il modulo provvede ad archiviare i dati che devono essere conservati (secondo quanto specificato in fase di configurazione del modulo) ed il *job* termina con esito positivo;
- in caso di esito negativo, il *job* termina con un messaggio di *error*.

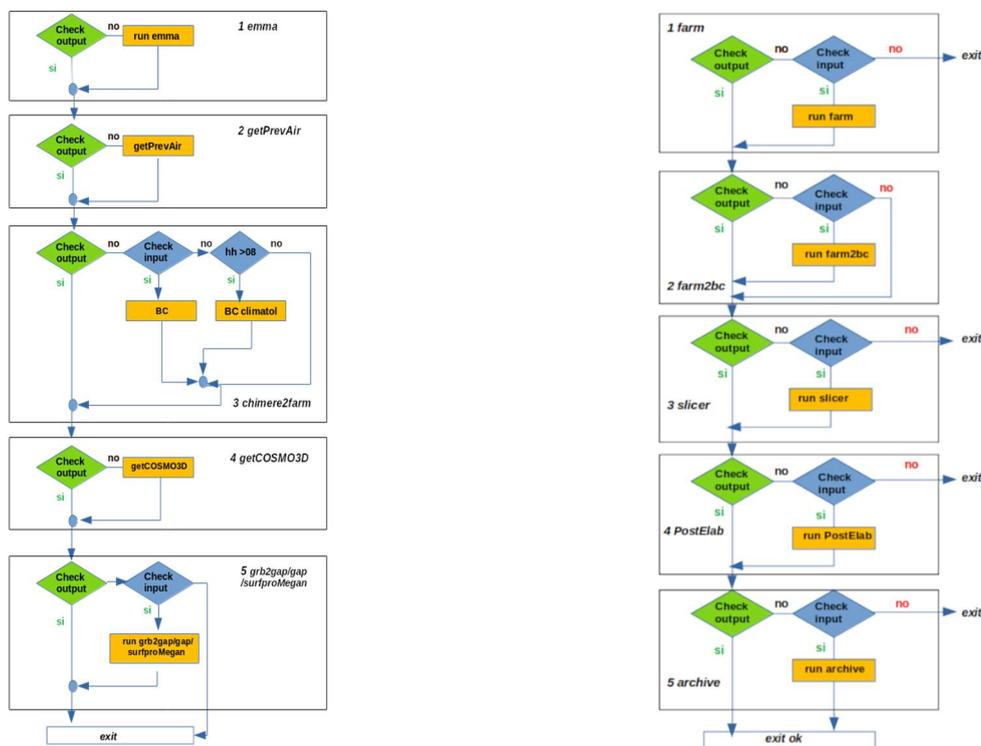


Figura 3. A sinistra workflow di *Sope_to_g2*; a destra workflow di *Pope_to_g2*.

Pope_to_g2 non produce risultati se *Sope_to_g2* non è andato a buon fine: in particolare l'assenza, o la non completezza, dei dati meteorologici di COSMO-I5 - per i quali non è disponibile attualmente un backup a differenza di quanto avviene con la banca dati Prev'Air per le BC - è il principale motivo di *fault* del *job*. Infine, va osservato come ciascun modulo di *Pope_to_g2* sia dipendente dal buon esito del modulo precedente (ad eccezione del modulo 2, da cui non dipendono i successivi).

5.4. Catena modellistica ad alta risoluzione (*progn_g3_ope*)

Rispetto alla catena modellistica principale *progn_g2_ope*, la catena modellistica ad alta risoluzione *progn_g3_ope* può essere eseguita contemporaneamente e con processi tra loro indipendenti su differenti domini di simulazione (progetti *prj*): per ciascun progetto la catena è eseguita tramite due *jobs*, *Sope_prj_g3* per i processi seriali e *Pope_prj_g3* per i processi paralleli. Per entrambi i *jobs* lo schema di funzionamento ed il workflow sono i medesimi di quanto descritto al paragrafo 5.3, con insequenti accorgimenti:

- nel *job Sope_prj_g3*, rispetto a *Sope_to_g2*, non vengono attivati i moduli 2, 3, 4;

– nel *job Pope_prj_g3*, rispetto a *Pope_to_g2*, non viene attivato il modulo 2.

ART 6 - SUPPORTO E FORMAZIONE

La Stazione appaltante si riserva il diritto, per l'intero arco di durata del presente appalto, di attivare *on-demand* dei servizi di supporto specialistico e di supporto alla formazione, verso il personale Arpa coinvolto nelle attività del presente appalto, a fronte di specifiche esigenze o per finalità più generali. L'importo complessivo dei servizi di supporto ed in generale dei servizi attivati *on-demand* non può eccedere la cifra del 20% dell'ammontare dell'importo triennale della fornitura. I servizi di supporto all'utenza, in tutte le rispettive forme, verbali o scritte, dirette od indirette, compreso tutto il materiale afferente (documentazione, manualistica, esercitazioni, reportistica, etc.) indipendentemente dal supporto tramite cui vengono forniti, **devono sempre essere erogati tassativamente in lingua italiana**. Può fare eccezione il materiale relativo ai servizi di formazione, se non predisposto specificatamente per la formazione in oggetto ovvero il materiale predisposto da terzi (libri di testo, libri di esercitazione, tutorials, video-tutorials, etc.), ritenuto interessante ed inerente all'argomento oggetto della formazione ed indicato a completamento ed approfondimento delle tematiche trattate, **solo se viene utilizzata la lingua inglese** come lingua alternativa.

6.1. Supporto specialistico

La stazione Appaltante ha facoltà di attivare servizi di supporto specialistico nei seguenti ambiti:

6.1.1. MODIFICA E CORREZIONE CODICI

la richiesta riguarda la modifica e/o correzione di codici attualmente in uso, per i quali siano stati rilevati malfunzionamenti o inefficienze. L'intervento può essere condotto a supporto ed in collaborazione con il personale dell'Agenzia, ovvero completamente demandato all'Aggiudicatario.

6.1.2. IMPLEMENTAZIONE NUOVE FUNZIONALITÀ

la richiesta riguarda lo sviluppo e l'implementazione di nuove funzionalità dei sistemi modellistici in uso, attraverso la modifica dei codici presenti o tramite sviluppo di nuovi codici. L'intervento può essere condotto a supporto ed in collaborazione con il personale dell'Agenzia, ovvero completamente demandato all'Aggiudicatario.

6.1.3. PARALLELIZZAZIONE CODICI

trasformazione di codici seriali in codici paralleli, attraverso introduzione di opportune direttive e/o la ristrutturazione (struttura dei cicli, strutture dati, gestione della memoria, etc...) dei codici medesimi. L'intervento può essere condotto a supporto ed in collaborazione con il personale dell'Agenzia, ovvero completamente demandato all'Aggiudicatario.

6.1.4. SVILUPPO ED INGEGNERIZZAZIONE DI SOLUZIONI PERSONALIZZATE

al fine di introdurre nuove funzionalità o di realizzare operativamente nuove attività da parte del personale Arpa, oppure per rendere più efficaci e di semplice gestione le operazioni già in essere può essere richiesto il supporto specialistico per sviluppare ed ingegnerizzare opportune soluzioni. Gli strumenti tecnici (linguaggi e tecniche di programmazione, tools, etc...) per la realizzazione delle soluzioni personalizzate deve essere concordato e condiviso con il personale di Arpa coinvolto nella gestione del servizio in oggetto, anche tramite specifica formazione.

L'erogazione del servizio di supporto specialistico prevede tempi di organizzazione e predisposizione delle attività, che sono tanto maggiori quanto più la richiesta è duratura (più giorni di intervento) e specializzata. Tempistiche e modalità di intervento vengono concordate di volta in volta con il personale di Arpa.

Tutti gli interventi di supporto specialistico condotti dall'Aggiudicatario devono essere corredati da un'adeguata documentazione:

- breve descrizione dell'intervento condotto, in cui si dia evidenza alle problematiche su cui si richiede di intervenire, alle soluzioni individuate ed al risultato ottenuto;
- documentazione tecnica delle procedure, delle funzionalità, dei sistemi di nuova introduzione.

6.2. Supporto alla formazione

Gli ambiti per cui deve essere garantita la formazione richiesta sono:

6.2.1. LINGUAGGI DI PROGRAMMAZIONE

vengono utilizzati nei sistemi modellistici in uso i linguaggi di programmazione *C*, *C++*, *Fortran* (*F77*, *F90*, *F95*), *Python*, *SQL*, che possono diventare oggetto di approfondimento e di specifica formazione, al pari di ulteriori linguaggi, scoperti di interesse nel corso delle attività condotte.

6.2.2. TECNICHE DI PROGRAMMAZIONE

in particolare, possono essere oggetto di formazione specifica tecniche di *programmazione parallela* e tecniche di *programmazione object oriented*.

6.2.3. BUILD SYSTEM TOOLS E PACKAGE MANAGEMENT TOOLS

possono essere oggetto di approfondimento e di specifica formazione *tools* di supporto alla programmazione ed alla gestione dei codici (si riportano a titolo di esempio *autotools*, *Cmake*, *RPM*, *setuptools*), sistemi di gestione del *versioning* (*CSV*, *SVN*, *Git*) e piattaforme di *collaborative development* (*GitHub*, *GitLab*).

L'erogazione del servizio di supporto alla formazione prevede tempi di organizzazione e predisposizione delle attività, che sono tanto maggiori quanto più la richiesta è duratura (più giorni di intervento) e specializzata. Tempistiche e modalità di intervento vengono concordate di volta in volta con il personale di Arpa.

6.2.4. OPERATIVE ENVIRONMENT

Possono essere oggetto di formazione specifica e di aggiornamento i sistemi adottati nella definizione e configurazione dell'*environment operativo* di funzionamento della *suite operativa*. Tra questi:

- *HPC Workload manager and job scheduler*;
- *Software environment manager*;
- *Collaborative development environment (CDE)*

Tutti gli interventi di supporto alla formazione condotti dall'Aggiudicatario devono essere corredati da un'adeguata documentazione:

- breve descrizione dell'intervento condotto, in cui siano specificati i contenuti della formazione fornita e le modalità;
- documentazione, manualistica e materiale informativo per la realizzazione delle attività di formazione e come ulteriore approfondimento e riferimento di consultazione.