

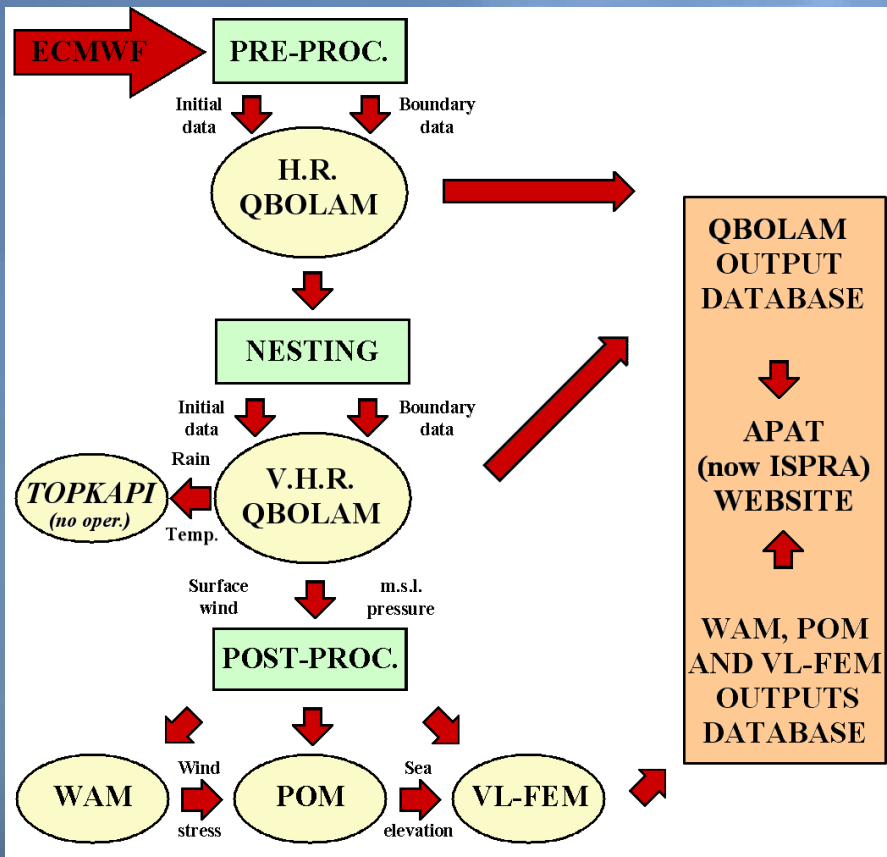
L'aggiornamento e la parallelizzazione del modello BOLAM nel Sistema Idro-Meteo-Mare



Parte I – Il contratto di ricerca ISPRA – ISAC-CNR

Marco Casaioli , Stefano Mariani – ACQ-MON Sett. Idrologia

COME VOI BEN SAPETE...



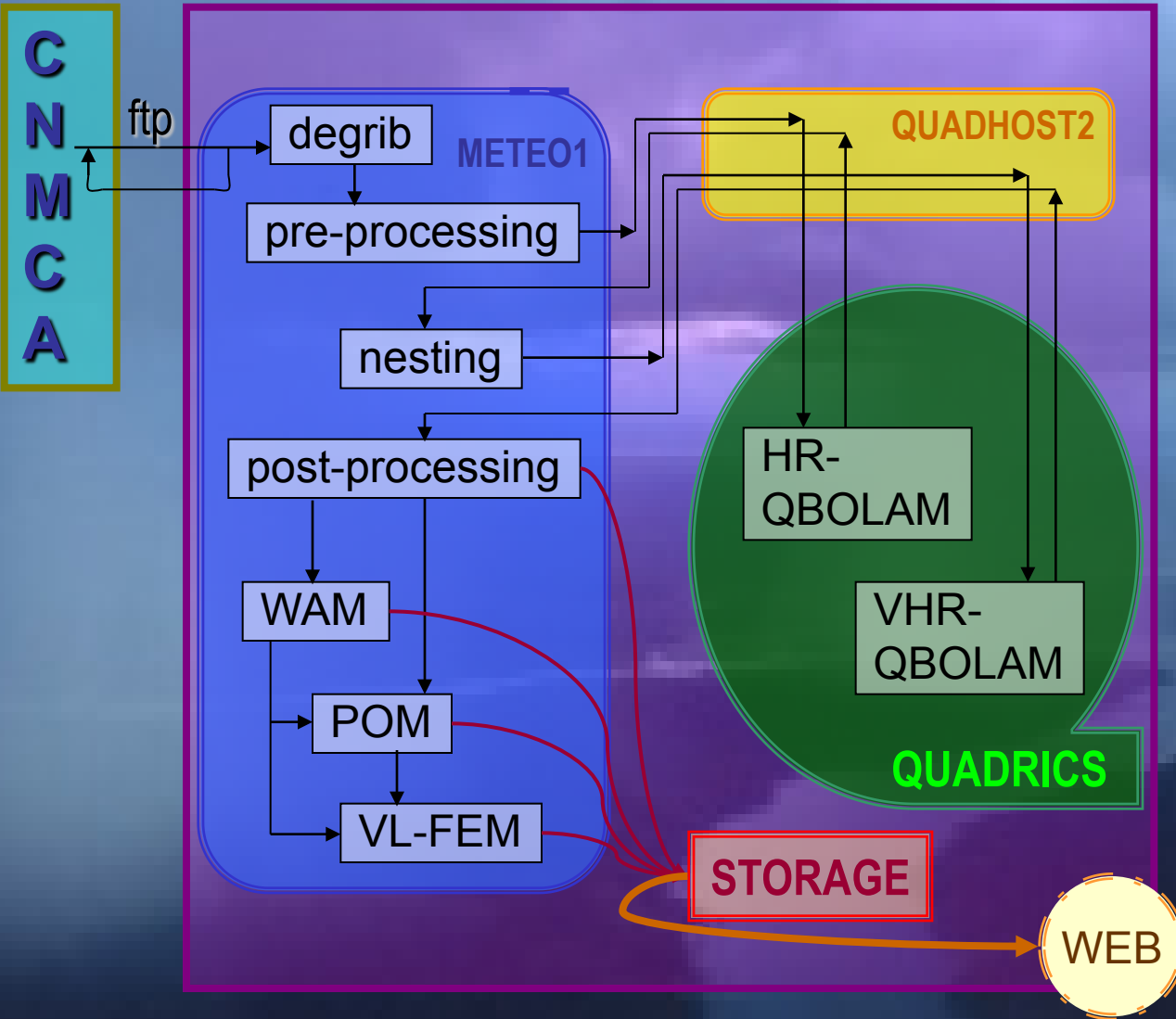
Il SIMM è un sistema integrato per la previsione idro-meteorologica e marina nel Bacino del Mediterraneo, con lo scopo di rappresentare i processi dinamici dalla scala globale alla scala locale di moto.

Questo sistema nasce dalla collaborazione di differenti istituzioni italiane di ricerca e di servizio: DSTN; ENEA; e CNR.

Il sistema è costituito da quattro modelli in cascata: QBOLAM; WAM; POM; e VL-FEM. Le condizioni iniziali e al contorno per il modello meteo sono fornite dall'analisi e la previsione del modello globale ECMWF.

“SIMM produced the first systematic, integrated hydro-meteorological and sea-state forecasts over the entire Mediterranean area, bridging from planetary to local scales of atmospheric motion” (Speranza et al, 2007)

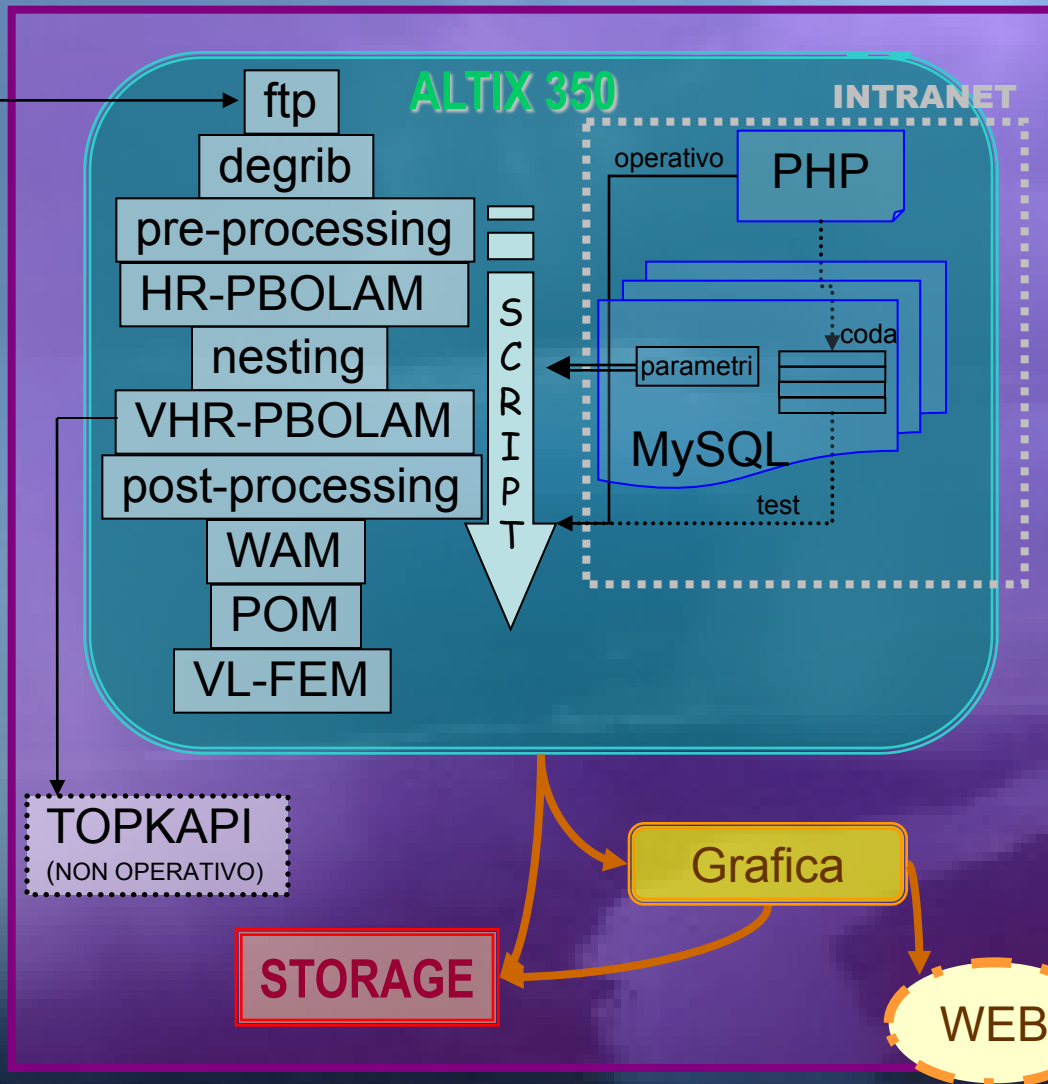
Implementazione 2000 presso il DSTN



QUADRICS: calcolatore superparallelo (128 CPU) velocissimo per l'epoca. Richiede programmazione in linguaggio proprietario (TAO) e macchina host. L'architettura sincrona (SIMD) dà risultati ottimali sulla "dinamica" ma pone severi vincoli sugli schemi di parametrizzazione. In particolare la convezione e la radiazione sono stati semplificati rispetto alla versione originale (1996) di BOLAM.

Implementazione 2006 presso il Dip.to Acque di APAT

C
N
M
C
A



Porting di tutta la catena operativa su ALTIX 350 (8 CPU Intel ITANIUM):

- Tempi di calcolo ridotti (es.: “figlio” da 2h a 17’)
- DB MySQL per i parametri ed interfaccia web (PHP) per lancio dei run (operativi o di test) e gestione code
- Storage: 2TB in RAID-5
- Assenza limitazioni allo sviluppo del codice meteo (architettura MIMD)
- Possibili lunghe sequenze di run (reforecasting)
- Sperimentazione modello idrologico (TOPKAPI)

PBOLAM 2006 (F. Valentinotti – QUADRICS) differisce da QBOLAM principalmente per la scrittura ottimizzata per l’architettura dell’ALTIX 350 (ITANIUM)

Interfacce di controllo web e database dei parametri

- Interfaccia PHP (Intranet) per la programmazione della sequenza operativo e per il lancio di simulazioni di test (con gestione automatica delle code)
- Controllo sulla sequenza, sulla durata, sul n. di processori e su altri parametri
- Il set completo dei parametri è gestito tramite database MySQL (controllo completo sui processi / archivio dei job / accesso riservato)

The screenshot displays the Quadrics Poseidon System web interface. The main content area shows the 'Parameters' table structure in the 'Bolam' database. The table has the following columns:

Campo	Tipo	Attributi	Null	Predefinito	Extra	Azione
<input type="checkbox"/> Tag	char(30)		Si	NULL		[Edit] [Delete] [Insert] [Update] [Refresh]
<input type="checkbox"/> SimulProdFlag	char(1)		Si	NULL		[Edit] [Delete] [Insert] [Update] [Refresh]
<input type="checkbox"/> ProductionEnable	int(1)		Si	NULL		[Edit] [Delete] [Insert] [Update] [Refresh]
<input type="checkbox"/> TimeSimulation	char(4)		Si	NULL		[Edit] [Delete] [Insert] [Update] [Refresh]
<input type="checkbox"/> PostFixNameSimulation	char(30)		Si	NULL		[Edit] [Delete] [Insert] [Update] [Refresh]
<input type="checkbox"/> JobStartTime	char(4)		Si	NULL		[Edit] [Delete] [Insert] [Update] [Refresh]
<input type="checkbox"/> JobEndTime	char(4)		Si	NULL		[Edit] [Delete] [Insert] [Update] [Refresh]
<input type="checkbox"/> PrefixPath	char(250)		Si	NULL		[Edit] [Delete] [Insert] [Update] [Refresh]

Below the table, there are options to 'Visualizza per stampa' and 'Proponi la struttura della tabella'. The 'Aggiungi' field is set to 1 campo(). The 'Dopo' dropdown is set to 'Tag'. The 'Esegui' button is visible.

At the bottom, there are statistics for the table:

Indici:		Spazio utilizzato:		Statistiche righe:			
Nome chiave	Tipo	Cardinalità	Azione	Tipo	Utilizzo	Istruzioni	Valore
Tag	UNIQUE	21	[Edit] [Delete]	Dati	6.909 Bytes	Formato	fisso
Crea un indice su 1 columns [Esegui]				Indice	2.048 Bytes	Righe	21
				Totale	8.957 Bytes	Lunghezza riga	329
						Dimensione riga	427 Bytes
						Creazione	24 Ago, 2006 at 07:26 PM
						Ultimo cambiamento	29 Apr, 2009 at 07:47 AM
						Ultimo controllo	23 Gen, 2009 at 07:57 AM

II BOLAM 2007 (seriale)

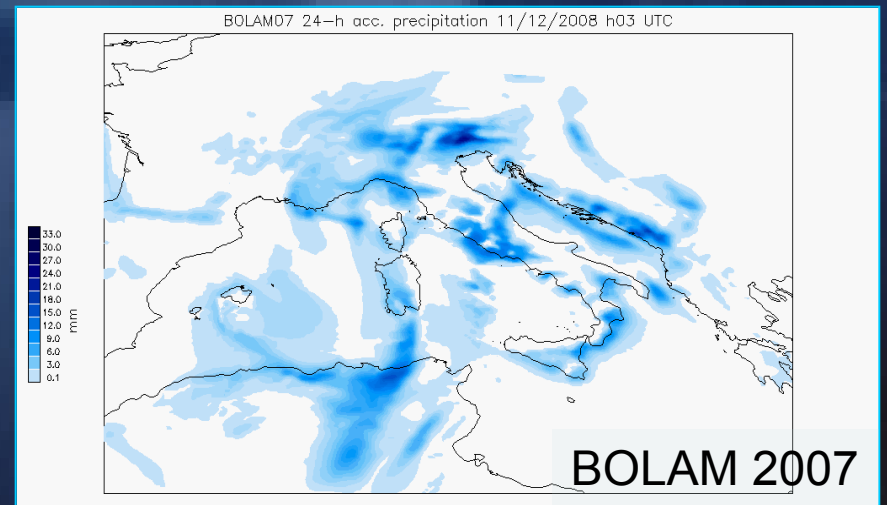
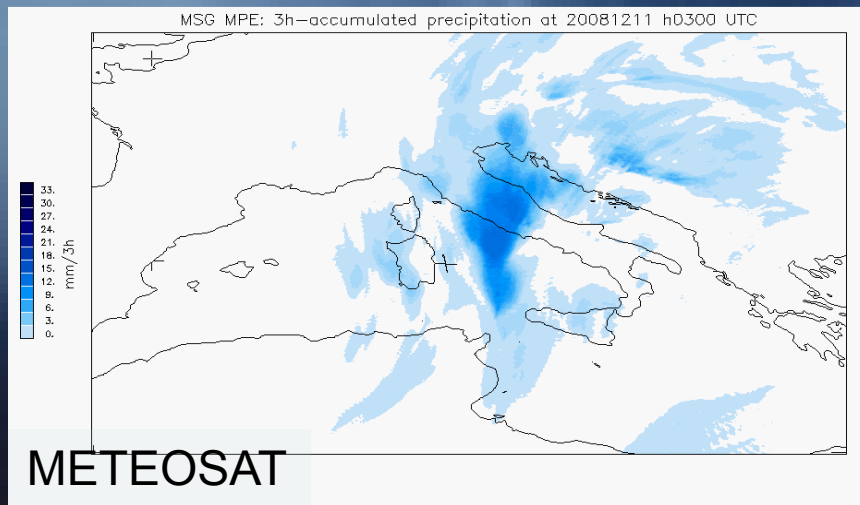
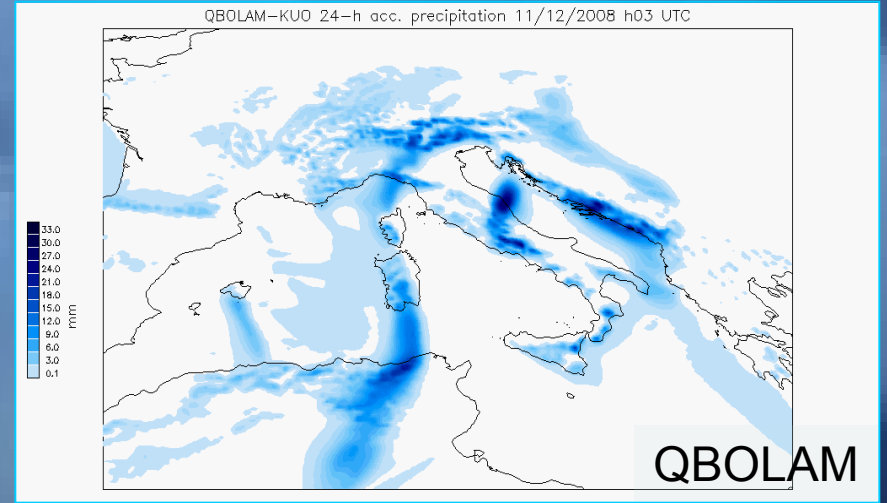
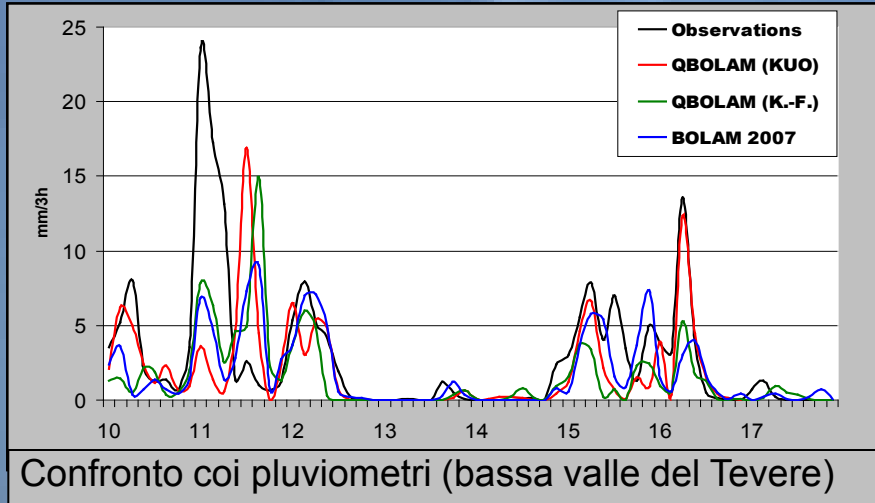
2007: contratto di ricerca con ISAC-CNR Bologna \Rightarrow versione aggiornata di BOLAM (seriale e in configurazione di ricerca) su ALTIX:

- ✓ Nuovo schema di avvezione \Rightarrow modello più stabile \Rightarrow più veloce
 - ✓ Convezione di Kain-Fritsch + trasporto esplicito di 5 idrometeore \Rightarrow miglioramento della previsione della precipitazione
 - ✓ Tre schemi per la radiazione (in QBOLAM: schema semplificato)
 - ✓ Due schemi per la turbolenza: Louis (QBOLAM) ed E-I
 - ✓ Quattro schemi per il suolo (il più semplice è quello di QBOLAM)
- È UTILE per studiare la configurazione ottimale per una successiva parallelizzazione e implementazione nel sistema.
 - Lunghi tempi di esecuzione (monoprocessore; non ottimizzato per ITANIUM) ma questo non è un problema in configurazione di ricerca.

Verifica del BOLAM 2007 (seriale)

Caso studio: piena del Tevere del Dicembre 2008

- Tipologia di evento (ciclone mediterraneo) particolarmente “difficile”
- L'uso del modello aggiornato riduce l'errore di previsione



La parallelizzazione del nuovo BOLAM

Fine 2008: Contratto di ricerca con ISAC-CNR Bologna per:

- Sviluppo di una versione parallela di BOLAM 2007 (o superiore)
- Implementazione operativa su ALTIX all'interno del SIMM, ed in configurazione di ricerca su una workstation (*Enlil*, 2xQuad-Core)
- Adattamento delle interfacce esistenti (MySQL, PHP) per gestire le nuove funzionalità del modello aggiornato
- Inserimento nel sistema di soluzioni sviluppate presso il Dipartimento (in particolare per il post-processing dei dati in uscita)
- Valutazione dei possibili miglioramenti della configurazione del modello meteo nel SIMM (risoluzione, dominio, durata) resi possibili dal nuovo codice, e successiva implementazione della configurazione ottimale
- Training del personale del Dipartimento sul nuovo sistema e il suo uso

FASI DEL PROGETTO

“La ricerca si articola in tre fasi:

Fase 1 Sviluppo di una versione intermedia del codice BOLAM parallelo e implementazione operativa nel SIMM.

Fase 2 Ulteriore sviluppo della versione BOLAM parallela per l’ottimizzazione nel SIMM e test di confronto.

Fase 3 Ottimizzazioni del grigliato di BOLAM parallelo in funzione della potenza di calcolo disponibile, training del personale del Dipartimento Tutela Acque Interne e Marine e fornitura di documentazione tecnica.”

La ricerca è portata avanti dall’ISAC-CNR di Bologna per quanto riguarda le attività a carattere scientifico (modellistica, verifica) ed è da essi affidata al CASPUR di Roma per le attività a carattere informatico (parallelizzazione, ottimizzazione, adattamento delle interfacce e training)

Il coinvolgimento di CASPUR è vantaggioso per l’ISPRA in quanto porta alla formazione di un nucleo di expertise locale relativo al sistema presso un importante centro di calcolo scientifico, tale da poter fornire assistenza anche in futuro.

FASE 1

- Implementata sulle due macchine una versione aggiornata di BOLAM parallela in una direzione (latitudine): PBOLAM 2009-A
 - Una prima ottimizzazione (opzioni di compilazione) ha portato a ridurre i tempi di esecuzione su ALTIX (ulteriore ottimizzazione nella fase 2)
 - Confronto con BOLAM 2007 e tra i risultati ottenuti sulle due macchine: consistenza e valore aggiunto della versione 2009 rispetto a quella 2007
 - Adattamento del sistema di interfacce MySQL-PHP per le sezioni *degrib; padre; nesting; figlio; post-processing*
 - Adattamento del codice per la compatibilità con la struttura del SIMM: produzione di campi superficiali a runtime (in via di completamento)
 - Testing della nuova catena operativa e sua messa in produzione (in corso)
- NB la vecchia catena operativa sarà mantenuta per completare le serie storiche.

FASE 2: Obiettivi - I

- Sviluppo di una versione parallela in longitudine e in latitudine
 - Migliore efficienza computazionale, spec. in caso di incremento del numero di processori
 - Maggiore compatibilità con la versione precedente (es. interfacce)
 - Ottimizzazione del codice rispetto all'architettura di ALTIX (ITANIUM)
 - Gestione della ricompilazione del codice (necessaria in alcuni casi)
- Sviluppo delle interfacce:
 - Possibilità di modificare dall'interfaccia Web i parametri del modello, in particolare per i run di test (ricerca) ma anche, se necessario, per quelli operativi (upgrade del sistema)
 - Ricompilazione tramite interfaccia PHP
 - Adattamento del pre-processing per uso dei dati ECMWF su livelli ibridi

FASE 2 : Obiettivi - II

- Sviluppo del segmento di post-processing:
 - Inserimento del post-processing nella catena operativa, con estrazione dei campi da visualizzare gestita tramite interfaccia PHP
 - Inserimento nel post-processing di schemi di interpolazione (remapping) e di estrazione di campi diagnostici (es. pseudo-water vapor) sviluppati presso il Dipartimento.
 - Uso dell'algoritmo di remapping per mantenere la compatibilità delle uscite di BOLAM coi modelli marini in caso di aumento della risoluzione del modello meteo.
- Testing e verifica del nuovo PBOLAM e del SIMM aggiornato
 - Simulazione di una serie di 3 mesi per la verifica del modello meteo.
 - Verifica dei modelli marini su 3 casi studio di mareggiata /acqua alta
 - Valutazione delle prestazioni del modello a risoluzione più spinta (fino a 6-7 km)

FASE 3 (Ottobre-Gennaio): Obiettivi

- Training del personale del Dipartimento (circa 24h di lezione, 10 persone): conoscenze tecniche riguardanti i codici del modello e loro implementazione e gestione operativa (anche per un eventuale, futuro porting del sistema)
- Definizione della **configurazione ottimale** e sua implementazione operativa
 - Aumento della risoluzione orizzontale. Il modello idrostatico funziona bene ancora a 7-8 km. Il passo dei prodotti ECMWF è già sceso da 50 a 25 km e scenderà a 15 entro fine anno!
 - Ampliamento del dominio di integrazione nelle 4 direzioni. Il modello potrebbe simulare meglio la formazione delle depressioni e fornire previsioni anche su Europa centro-settentrionale e Medio Oriente
 - Estensione della durata della simulazione dalle attuali 48h a 3-5 gg e della frequenza dei run (ogni 12h anziché ogni 24h)

Tutto ciò richiede ovviamente maggiore potenza di calcolo, un incremento del set di dati in ingresso e la disponibilità di uno storage di grandi dimensioni.