

Coding norms & parallelization

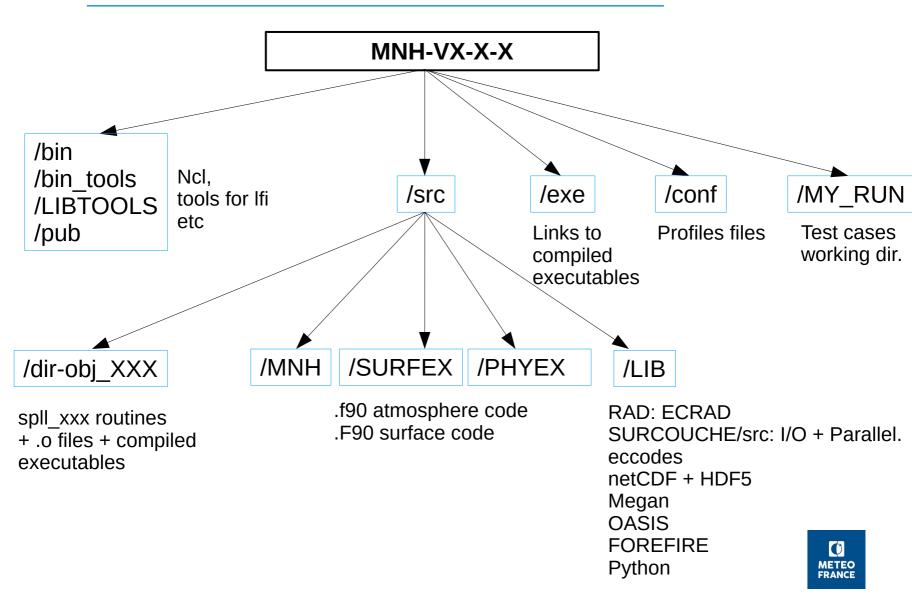
Méso-NH tutorial 12-15 November 2024

Outline

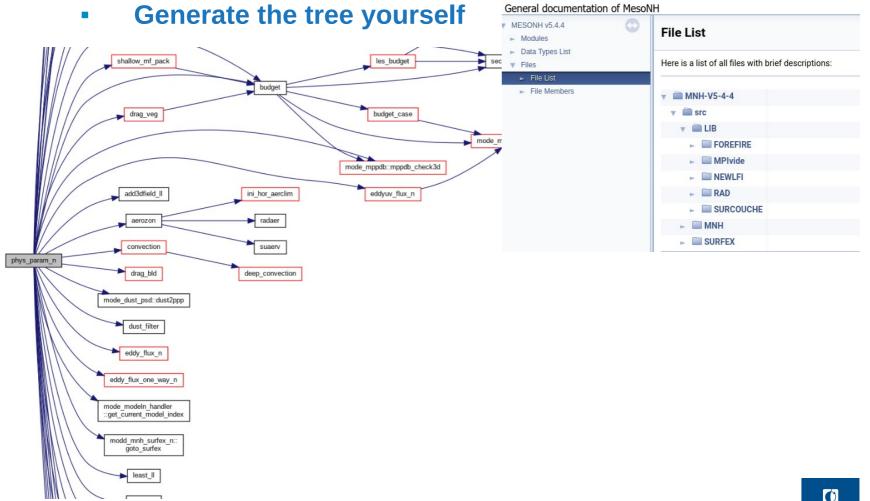
- How to read the code? (users)
 - Tree
 - Fortran 95 norms
 - MesoNH norms
- How to develop in Meso-NH? (developers)
 - Basic rules
 - ► GIT
 - Specificities for PHYEX
- Parallelization
 - Principles
 - Large grids and scalability
 - Reproducibility
- I/O
- FAQ



Tree of the PACK



Tree of the code: Doxygen





Read the code

Fortran 95

- up to 132 characters per line
- ► No blank line (use !)
- Continuation character &
- ► CODE IN CAPITAL LETTERS! comments in small letters

Variable declaration

- ► Global: used everywhere, declaration in module
- Dummy argument in routine/function

e.g.: REAL, DIMENSION(:,:), INTENT(INOUT) :: PUT ! u-wind at time t

- ► **Local**: only known in the routine
- All variables must be declared: IMPLICIT NONE



Read the code: name of variables

DOCTOR norm convention

Type Status	INTEGER	REAL	LOGICAL	CHARACTER	TYPE
Global	N	X	L (not LP)	С	T (not TP,TS,TZ)
Dummy argument	K	P (not PP)	O	Н	TP
Local	l (not IS)	Z (not ZS)	G (not GS)	Y (not YS, YP)	TZ
Loop control	J (not JP)	-	-	-	-



Read the code: name of variables

Example: dummy argument variables

```
INTEGER,
                        INTENT(IN)
                                      :: KKA
                                                       !near ground array index
                                      :: KKU
                                                       !uppest atmosphere array index
INTEGER,
                        INTENT(IN)
INTEGER,
                        INTENT(IN)
                                    :: KKL
                                                       !vert. levels type 1=MNH -1=AR0
                                                       ! number of moist var.
                        INTENT(IN)
                                      :: KRR
INTEGER,
                                                       ! number of liquid water var.
INTEGER,
                        INTENT(IN)
                                     :: KRRL
                                      :: KRRI
INTEGER,
                        INTENT(IN)
                                                       ! number of ice water var.
                                                       ! switch for syncronous
LOGICAL,
                        INTENT(IN)
                                          OCLOSE OUT
                                                       ! file opening
LOGICAL,
                        INTENT(IN)
                                      :: OTURB FLX
                                                       ! switch to write the
                                  ! turbulent fluxes in the syncronous FM-file
CHARACTER(LEN=80),
                        INTENT(IN)
                                      :: HTURBDIM
                                                       ! dimensionality of the
                                                        ! turbulence scheme
CHARACTER(LEN=80),
                                                       ! type of Third Order Moment
                                          MOTH
                        INTENT(IN)
REAL,
                                      :: PIMPL, PEXPL! Coef. for temporal disc.
                        INTENT(IN)
                                                       ! Double Time Step
REAL,
                        INTENT(IN)
                                      :: PTSTEP
TYPE (TFILEDATA).
                        INTENT(IN)
                                      :: TPFILE
                                                       ! Output file
                                          PDZZ, PDXX, PDYY, PDZX, PDZY
REAL, DIMENSION(:,:,:), INTENT(IN)
                                                       ! Metric coefficients
REAL, DIMENSION(:,:),
                                          PDIRCOSZW
                                                        ! Director Cosinus of the
                        INTENT(IN)
                                                       ! normal to the ground surface
REAL, DIMENSION(:,:,:), INTENT(IN)
                                      :: PZZ
                                                       ! altitudes
```



Read the code: modules

Module = units of compilation

```
MODULE name_mod

declarations

END MODULE name_mod
```

Use in another subroutine: USE name_mod

4 types

- MODD_xxx: variables Declarations
 - ► For MODD, we encourage to select the variables needed USE MODD_xxx, **ONLY**: var1, var2
- MODN xxx: Namelist declaration
- MODE_xxx: Functions
- MODI_xxx: included in the routine Interface (check at the call that arguments are correct) ⇒ will be removed soon (created at compilation)



Read the code: modules

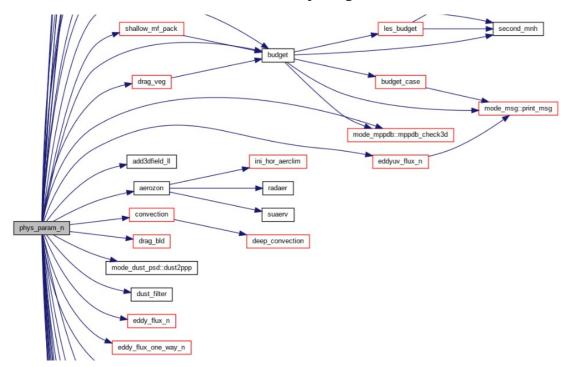
Examples

- MODD: modd blank.f90
- MODN: modn_backup.f90
- MODE: mode_thermo.f90
- MODI: spll_modi_phys_param_n.f90 (in / src/dir-obj..../MASTER/)
 - created from phys_paramn.f90 (/src/MNH)



Read the code: _n files

- xxxn files are for grid-nesting
 - e.g. phys_paramn.f90, modn_advn.f90
 - Routines called X times for X sub-models
 - Routines without _n cannot use modules with _n
 - Variables relative to the horizontal domain (i,j) must be introduced by dummy arguments





Outline

- How to read the code? (users)
 - Tree
 - Fortran 95 norms
 - MesoNH norms
- How to develop in Meso-NH? (developers)
 - Basic rules
 - ► GIT
 - Specificities for PHYEX
- Parallelization
 - Principles
 - Large grids and scalability
 - Reproducibility
- FAQ



Develop in MNH: good practices

Array

Dynamical allocation

```
REAL, DIMENSION(:,:,:), ALLOCATABLE :: ZZS
ALLOCATE(ZZS(IIU,IJU))
IF(ALLOCATED(ZZS)) ...
DEALLOCATE(ZZS)
```

► Use (:,:,:) for readability and compilation optimization→ matrix computation

```
ZVMOD(:,:) = SQRT(ZU(:,:)**2 + ZV(:,:)**2)
```

Pointer

► Declaration + Initialization (with NULLIFY() or => NULL())
CHARACTER(LEN=8), DIMENSION(:), POINTER :: NAME
NAME=>NULL()



Do not duplicate code

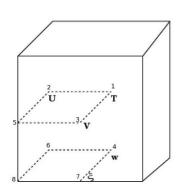
- If you use twice a code ⇒ create a routine/function
 - A large number of routines is OK: avoid « CONTAINS »
 - ► Avoid code duplication, ~1 million lines
 - Can be used by other developers later
- Check existing functions (in mode_)
 - Mode_thermo (r_v from RH, P, θ_v saturation, etc)
 - mode datetime: handling time
 - mode_xxx: for each dvpmnt type
 - shuman.f90: operators (mean, gradient)



Discrete operators

- Mean & finite difference USE MODI_SHUMAN (shuman.f90)
- e.g. $\bar{\alpha}^z$ MXF, MXM, MYF, MYM, MZF, MZM: means
 - Means in X,Y or Z directions for a variable at Flux or Mass point
- $\frac{\partial \alpha}{\partial \alpha}$ DXF, DXM, DYF, DYM, DZM, DZF
- e.g. $\frac{\partial z}{\partial z}$ Finite difference in X,Y or Z directions for a variable at Flux or Mass point

Gradients USE MODI_GRADIENT_M



- GX_M_M, GY_M_M, GZ_M_M
- ► Gradients along X,Y or Z for a Mass variable, results at Mass point
- GX_M_U , GY_M_V , GZ_M_W
- Gradients along X,Y or Z for a Mass variable, results at flux U,V or W point



Available free-to-use variables: modd_blankn

For testing

```
REAL, SAVE
                        :: XDUMMY1, XDUMMY2, XDUMMY3, XDUMMY4, &
                          XDUMMY5, XDUMMY6, XDUMMY7, XDUMMY8
    INTEGER, SAVE
                        :: NDUMMY1, NDUMMY2, NDUMMY3, NDUMMY4, &
                          NDUMMY5, NDUMMY6, NDUMMY7, NDUMMY8
    LOGICAL, SAVE
                        :: LDUMMY1, LDUMMY2, LDUMMY3, LDUMMY4, &
                           LDUMMY5, LDUMMY6, LDUMMY7, LDUMMY8
CHARACTER(LEN=80), SAVE :: CDUMMY1, CDUMMY2, CDUMMY3, CDUMMY4, &
                           CDUMMY5, CDUMMY6, CDUMMY7, CDUMMY8
             SAVE, DIMENSION(JPDUMMY) :: XDUMMY
    INTEGER, SAVE, DIMENSION(JPDUMMY) :: NDUMMY
    LOGICAL. SAVE, DIMENSION(JPDUMMY) :: LDUMMY
CHARACTER(LEN=80), SAVE, DIMENSION(JPDUMMY) :: CDUMMY
    END MODULE MODD BLANKn
```

Use in your code

```
USE MODD_BLANKn, ONLY: XDUMMY1
MYVAR(:,:) = MYVAR(:,:)*XDUMMY1
```

In namelist

&NAM_BLANKn XDUMMY1=1.0 /



Create a new routine: format

```
!MNH LIC Copyright 1998-2020 CNRS, Meteo-France and Universite Paul Sabatier
!MNH LIC This is part of the Meso-NH software governed by the CeCILL-C licence
!MNH LIC version 1. See LICENSE, CeCILL-C V1-en.txt and CeCILL-C V1-fr.txt
!MNH LIC for details. version 1.
     MODULE MODI READ ALL DATA GRIB CASE
     INTERFACE
SUBROUTINE READ ALL DATA GRIB CASE(HFILE, TPPRE REAL1, HGRIB, TPPGDFILE,
                   PTIME HORI, KVERB, ODUMMY REAL
USE MODD IO 11, ONLY: TFILEDATA
CHARACTER(LEN=4), INTENT(IN) :: HFILE
                                             ! which file ('ATMO','ATM1' or 'CHEM')
TYPE(TFILEDATA), POINTER, INTENT(INOUT) :: TPPRE REAL1 ! PRE REAL1 file
CHARACTER(LEN=28), INTENT(IN)
                                :: HGRIB
                                              ! name of the GRIB file
                                             ! physiographic data file
TYPE(TFILEDATA), INTENT(IN) :: TPPGDFILE
                                             ! verbosity level
INTEGER,
                  INTENT(IN)
                                :: KVERB
                                :: ODUMMY REAL ! flag to interpolate dummy fields
LOGICAL,
                  INTENT(IN)
                  INTENT(INOUT) :: PTIME HORI ! time spent in hor. interpolations
REAL,
END SUBROUTINE READ ALL DATA GRIB CASE
END INTERFACE
END MODULE MODI READ ALL DATA GRIB CASE
```

Licence

Module Interface

With declaration of dummy arguments



```
SUBROUTINE READ ALL DATA GRIB CASE(HFILE, TPPRE REAL1, HGRIB, TPPGDFILE,
                     PTIME HORI, KVERB, ODUMMY REAL
       Self doc of the routine
      *READ ALL DATA GRIB CASE* - reads data for the initialization of real cases.
     PURPOSE
                                                                                              PURPOSE
     This routine reads the two input files :
       The PGD which is closed after reading
       The GRIB file
     Projection is read in READ LFIFM PGD (MODD GRID).
     Grid and definition of large domain are read in PGD file and Grib files.
     The PGD files are also read in READ LFIFM PGD.
     The PGD file is closed.
     The MESO-NH domain is defined from PRE REAL1.nam inputs in SET SUBDOMAIN CEP.
     Vertical grid is defined in READ VER GRID.
     PGD fields are stored on MESO-NH domain (in TRUNC PGD).
                                                                                              METHOD
     METHOD
   0. Declarations
     1. Declaration of arguments
     Declaration of local variables

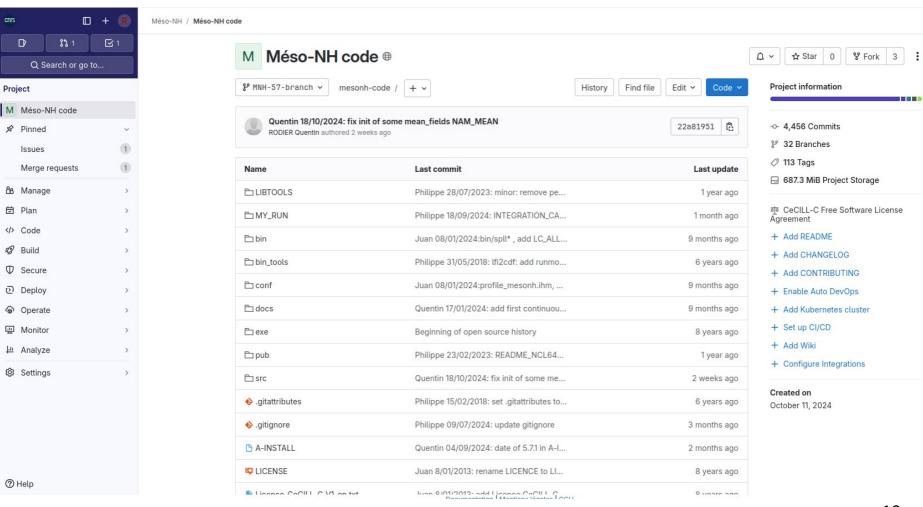
    Read PGD file

     1. Domain restriction
     2. Coordinate conversion to lat, lon system
   Read Grib fields
   3. Vertical grid
   4. Free all temporary allocations
11
11
                                                                                              EXTERNAL routines
     EXTERNAL
                                                                                              and functions used
11
     subroutine READ LFIFM PGD
                                : to read PGD file
                                : to define the horizontal MESO-NH domain.
11
     subroutine SET SUBDOMAIN
11
     subroutine READ VER GRID
                                : to read the vertical grid in namelist file.
11
                                 : horizontal bilinear interpolation
     subroutine HORIBL
11
                                : projection from conformal to lat,lon
     subroutine XYTOLATLON
11
11
     Module
               MODI SET SUBDOMAIN
                                     : interface for subroutine SET SUBDOMAIN
11
     Module
               MODI READ VER GRID
                                   : interface for subroutine READ VER GRID
11
     Module
               MODI HORIBL
                                     : interface for subroutine HORIBL
11
     Module
               MODI XYTOLATLON
                                     : interface for subroutine XYTOLATLON
11
                                                                                              IMPLICIT ARGUMENTS
11
     IMPLICIT ARGUMENTS
11
11
                                                                                                                            O
       Module MODD CONF
                            : contains configuration variables for all models.
          NVERB : verbosity level for output-listing
                                                                                                                          METEO
11
                                                                                                                          FRANCE
                            : contains logical unit names for all models
11
       Module MODD LUNIT
11
          CLUOUTO : name of output-listing
                                                                                                                                   17
11
       Module MODD PGDDIM
                            : contains dimension of PGD fields
```

```
REFERENCE
11
11
       Book 1 : Informations on ISBA model (soil moisture)
                                                                                                    REFERENCE (scientific
       "Encoding and decoding Grib data", John D.Chambers, ECMWF, October 95
11
11
       "A guide to Grib", John D.Stackpole, National weather service, March 94
                                                                                                    doc or publication)
...
11
     AUTHOR
11
11
                                                                                                    Author(s)
11
       J. Pettre and V. Bousquet
11
11
     MODIFICATIONS
11
       Original
                   20/11/98
                   15/03/99 (V. Masson) phasing with new PGD fields
                   28/05/99 (V. Bousquet) bug in wind interpolated variable for
11
                   31/05/99 (V. Masson) set pressure points (given on a regular grid at ECMWF)
                             on orography points (assuming the last are included in the former)
11
11
                   08/03/2018 (P.Wautelet) replace ADD FORECAST TO DATE by DATETIME CORRECTDATE
   Philippe Wautelet: 05/2016-04/2018: new data structures and calls for I/O
                                                                                                     Modifications
11
          Pergaud : 2018 add GFS
                    01/2019 (G.Delautier via Q.Rodier) for GRIB2 ARPEGE and AROME from EPYGRAM
11
       Bielli S. 02/2019 Sea salt : significant sea wave height influences salt emission; 5 salt
  P. Wautelet 14/03/2019: correct ZWS when variable not present in file
  Q. Rodier 27/01/2020: switch of GRIB number ID for Orograppy and hydrometeors in ARPEGE/AROME
 Q. Rodier 21/04/2020: correction GFS u and v wind component written in the right vertical ord
! Q. Rodier 02/09/2020 : Read and interpol geopotential height for interpolation on isobaric su
       DECLARATIONS
                                                                                                    0. Declarations
                                                                                                     Use modules
USE MODE DATETIME
USE MODE FM, ONLY: IO FILE CLOSE ll
USE MODE IO 11, ONLY: UPCASE
USE MODI READ HGRID n
USE MODD FIELD n, ONLY: XZWS, XZWS DEFAULT
USE MODD IO 11, ONLY: TFILEDATA
                                                                                                     Declarations
USE GRIB API
                                                                                                     of arguments
IMPLICIT NONE
!* 0.1. Declaration of arguments
                                                                                                                              O
                                                                                                                             METEO
CHARACTER(LEN=4), INTENT(IN) :: HFILE ! which file ('ATM0', 'ATM1' or 'CHEM')
TYPE(TFILEDATA), POINTER, INTENT(INOUT) :: TPPRE REAL! PRE REAL1 file
                                                                                                                                     18
CHARACTER(LEN=28), INTENT(IN) :: HGRIB
                                             ! name of the GRIB file
```

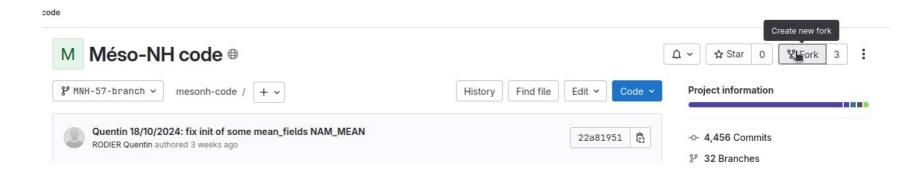
Versioning with GIT

- We use a gitlab instance (KODA)
- git clone https://src.koda.cnrs.fr/mesonh/mesonh-code.git



Versioning with GIT: how to contribute

- You do have a JANUS (CNRS) id
- Create a fork from the official repo



On your fork :

- Create your own branch
- Develop inside, commit
- From time to time, merge from the official branch
- Once ready, make a merge-request



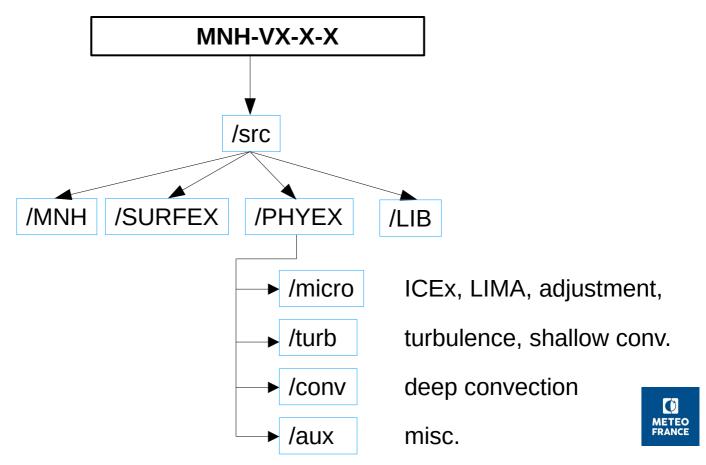
Versioning with GIT: how to contribute

- You do not have a JANUS (CNRS) id
- Send a token request at mesonhsupport@obs-mip.fr
- You will have written permission on your own branch only
 - Develop inside, commit
 - From time to time, merge from the official branch
 - Once ready, send us an email or @tag us on the repo



PHYsics EXternalized

- Keep the same files for Méso-NH and AROME
- Adaptation to GPU baseline (for AROME)
- https://github.com/UMR-CNRM/PHYEX



PHYsics EXternalized

- From 5.6.0, specific norms (more strict):
 - no allocatables : use automatic arrays,
 - dimensions of dummy argument arrays are explicit (no (:,:,:))
 except variables declared with the PARAMETER attribute
 - no variable from modules (e.g. MODD) can be used variables must be added in argument at the interface of PHYEX.
 - no functions returning arrays : use a subroutine.
 - CONTAINS subroutine included in a subroutine must not declare local variables. Declare these variables in the first subroutine.
 - horizontal dimensions are packed into one dimension (NI*NJ)



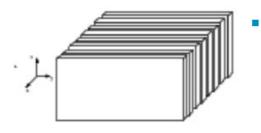
Outline

- How to read the code? (users)
 - Tree
 - Fortran 95 norms
 - MesoNH norms
- How to develop in Meso-NH? (developers)
 - Basic rules
 - ► GIT
 - Specificities for PHYEX
- Parallelization
 - Principles
 - Large grids and scalability
 - Reproductibility
- FAQ



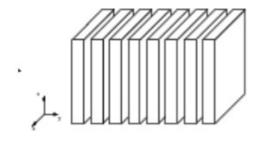
Parallelization

• All the steps of Meso-NH are parallelized (except SPECTRE) mpirun -np \${MPI_TASKS_TOTAL} -ppn \${MPI_TASKS_PER_NODE}

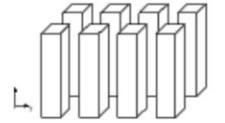


3 types of decomposition

CSPLIT='YSPLITTING' (default option)Adapted to vector machines



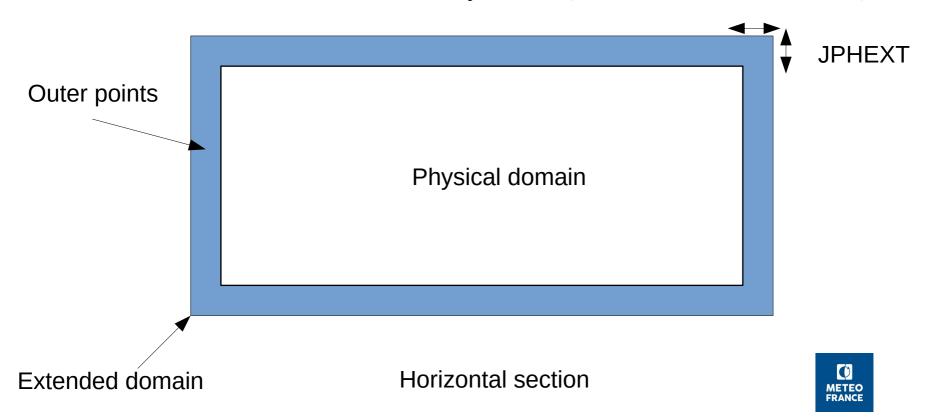
CSPLIT='XSPLITTING'Adapted to vector machines



Adapted to scalar machines with a lot of processors

Physical and extended domain

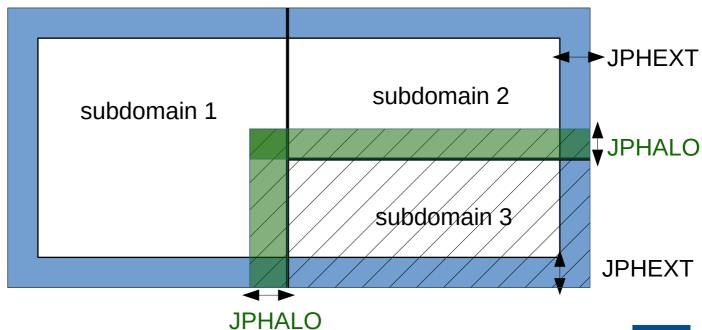
- To deal with LBC, the model arrays are over-dimensioned
 - along vertical dim. by JPVEXT points
 - along horizontal dim. by JPHEXT points
 - ▶ JPHEXT=JPVEXT=1 by default (JPHEXT=3 for WENO5 in CYCL)



Parallel decomposition

Horizontal decomposition

- Each subdomain is allocated on a different core
- ► Finite difference requires data along the border of adjacent physical subdomains ⇒ extended domain
- Overlap area = HALO of size JPHALO=1 (by default)

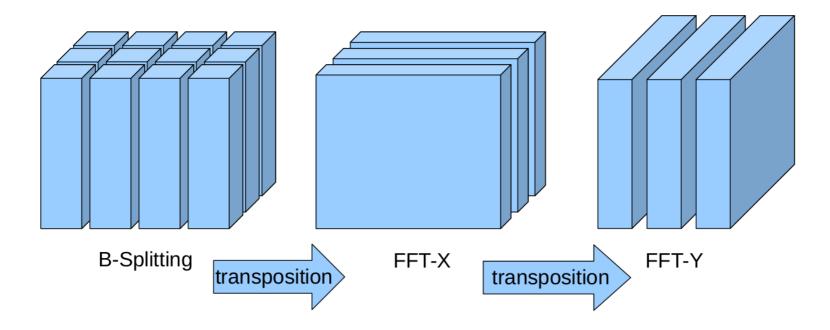


Extended

Extended subdomain

Large grids

Before, the pressure solver was limiting (CPRESOPT='CRESI')

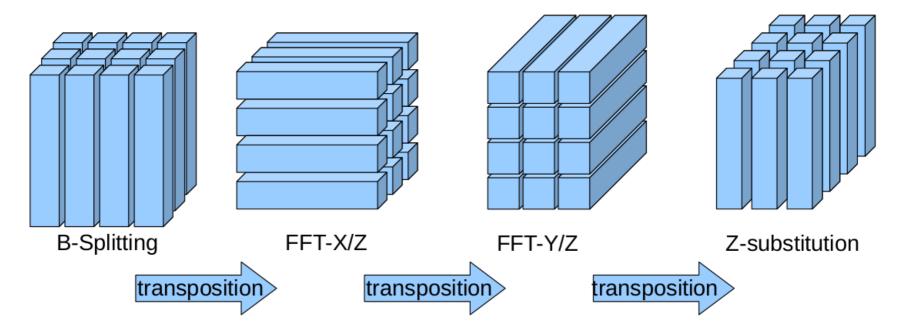


Maximum cores = min (dimX, dimY)



Large grids

Now, parallelization of the pressure solver in Z (CPRESOPT='ZRESI')



- Maximum cores = dimX * dimY
- Will be the default value of CPRESOPT (not yet)



Large grids: optimization

Large grids

- With BSPLITTING, try to set up a nb of cores such as the decomposition will give square subdomains.
- Avoid prime numbers

Scalability

The scalability is good down to subdomains size of 8x8 pts

For square subdomains, we recommend to use maximum (SmallDim/8)² cores

with SmallDim = the smallest dimension within all domains

- Example: 1st domain of 256x256x80 points
 2nd domain of 128x128x80 points
 - \Rightarrow Maximum nb of cores = $(128/8)^2$ = 256 cores



Online doc (books and guides)

 7.1 Initialization ■ 7.1.1 **SET DIM II** ■ 7.1.2 **SET JP II** 7.1.3 SET DAD II 7.1.4 SET LB X II 7.1.5 SET XRATIO II 7.1.6 SET XOR II 7.1.7 SET XEND II 7.1.8 INI PARA II ■ 7.1.9 END PARA II o 7.2 Halo 7.2.1 ADD2DFIELD II, ADD3DFIELD II 7.2.2 ADD1DFIELD II 7.2.3 DEL2DFIELD II, DEL3DFIELD II 7.2.4 ADD FIELD2 II 7.2.5 DEL FIELD2 II 7.2.6 UPDATE HALO II To update the halo with the values computed by the neighboring subdomains ■ 7.2.7 **UPDATE 1DHALO II** ■ 7.2.8 **UPDATE HALO2 II** 7.2.9 UPDATE BOUNDARIES II 7.3 Data distribution ■ 7.3.1 **REMAP 2WAY X II** 7.3.2 REMAP X Y II 7.3.3 REMAP Y X II 7.3.4 REMAP X 2WAY II 7.3.5 EXTRACT II ■ 7.3.6 **GET SLICE II** 7.4 Domain informations ■ 7.4.1 **GET DIM EXT II** ➤ To get the dimension of the extended subdomain ■ 7.4.2 **GET DIM PHYS II** To get the dimension of the physical subdomain 7.4.3 GET OR II 7.4.4 GET GLOBALDIMS II 7.4.5 GET INDICE II To get the origin and end coordinates of the physical local subdomain 7.4.6 GET PHYSICAL II 7.4.7 GET INTERSECTION II • 7.4.8 LNORTH II, LWEST II, LSOUTH II, LEAST II Returns a boolean which is True if the process is situated at the north... o 7.5 Min Max of the physical domain ■ 7.5.1 **GMAXLOC I** 7.5.2 GMINLOC II ■ 7.5.3 MAX II, MĪN II 7.6.1 REDUCESUM II 7.6.2 SUM DIM1 II, SUM DIM2 II 7.6.3 **SUM 1DFIELD II** O ■ 7.6.4 **SUM 1D II** 7.6.5 SUM 2D II

 7.6.6 SUM 3D II 7.6.7 SUMMASK II

7.6.8 SUMMASKCOMP II



Reproducibility

Repro = result independent of the number of processes

- On the same machine, same compiler, with an optimization level of DEBUG or O2
- i.e. Repro. not garanteed between different machines, compilers or in O3 (aggressive compiler optimization level)

Help us keep the code reproducible

- Do not use functions such as ALL, ANY, COUNT, MAXVAL, MINVAL, SUM, MINLOC, MAXLOC ⇒ local applications
- Use the parallelized versions _II (ex. MAX II, SUM 3D II)
- Avoid anticipated exit of a loop (EXIT, CYCLE, RETURN)

General view: replace

« if the process has converged for every point, we stop » (EXIT, CYCLE, RETURN)

By

« we compute on all the points where the process has not converged yet » (WHERE)

A checking tool exists (MPPDB_CHECK): ask us for help



Input / Outputs

- Jupyter Notebook tutorial and documentation
 - github.com/**PhilippeWautelet**/2022-mesonh-io-lecture/blob/main/io_mnh.ipynb
- Files structure (dimensions, netCDF groups, category)
- Main namelists and options (backup, frequent outputs, budgets, LES budgets, etc)
- How to :
 - write/read new variables
 - Handle metadata



- Large grid / High resolution
 - 1) Insufficient virtual memory (often at PREP_REAL_CASE or MESONH)

The RSSMax value returns the memory used on the most consuming node (must be < 256GB on Belenos/Taranis)

- Increase the number of nodes or decrease the number of cores by node
- Check you have ulimit -s unlimited on bash



Large grid / High resolution

2) INCREASE YOUR HALO_PREP IN NAM_PREP_SURF_ATM (at PREP_REAL_CASE)

Some points lack data and are too far away from other points. Please define a higher halo value in &NAM_PGDFILE NHALO=xxx

(at PREP_PGD)

Problem in the extra/interpolation of surface fields

SURFEX needs more point of the same COVER to interpolate prognostic fields data

Increase the NHALO (maximum = max(dim) / 2)

&NAM_PREP_SURF_ATM NHALO_PREP for prep &NAM_PGDFILE_NHALO / for pgd

Warning: increasing the NHALO increases significally the duration of the PGD/PREP

In last cases, set NHALO/NHALO_PREP=0 : SURFEX will interpolate by using the whole domain field instead of the local field + NHALO. If not enough data is found, you must change your domain coverage



- Large grid / High resolution
 - 3) Error occurs in MPI_BSEND or MPI_ERR_BUFFER (often at PREP_REAL_CASE or MESONH)

Problem in the MPI buffer size

The process #0 does not have enough memory to write one 3D field

- Increase the MPI buffer size (40 by default)
 & NAM_CONFZ MPI_BUFFER_SIZE=200 /
- Recommended value = 2 * size of a 3D field in MB

MPI_BUFFER_SIZE =
$$2 * N_x * N_y * N_z * 8 / 10^6$$



Small grid

Use a very small number of processes, tempting to use a lot with HPC

- ► SPAWNING usually works with 1 process
- ► MESONH: try first on one node and a few processes

Read warnings in

- OUTPUT_LISTING files
- .eo job outputs

Do not hesitate to report bugs

mesonhsupport@obs-mip.fr

