IFS components on NEC SX Aurora TSUBASA

CloudSC and ecRAD

Fatemeh Pouyan, Ioan Hadade, Olivier Marsden

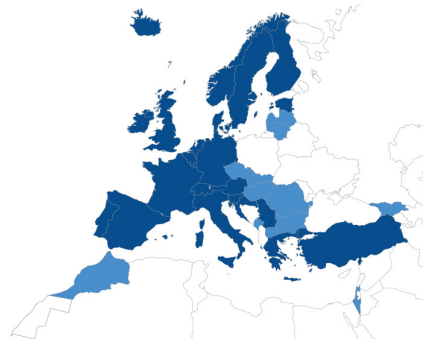
fatemeh.pouyan@ecmwf.int

European Centre for Medium-Range Weather Forecasts

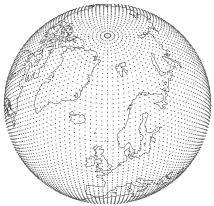
ECMWF

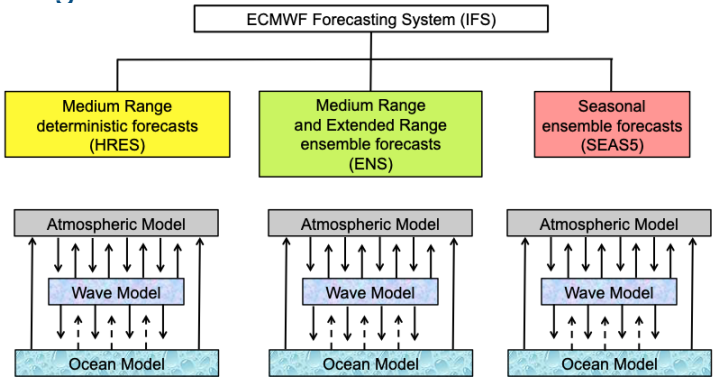
The European Centre for Medium-Range Weather Forecasts

- Independent intergovernmental organisation
- Established in 1975, today supported by **23 member** and **12 cooperating states**
- Headquarters in **Reading (UK)**, data centre in **Bologna (IT)**, Offices in **Bonn (DE)**
- Research institute and 24/7 operational service:
  - produce and disseminate NWP
  - operate meteorological data archive
  - implement Copernicus services CAMS and C3S
  - provide computing resources to member states



Operational configurations





Forecast period	10 days	15 days / 46 days	7 months / 13 months
Ensemble members	1	51	51 / 15
Frequency	2x daily	2x daily / 2x weekly	monthly / quarterly
Resolution	TCo1279L137 (9km)	TCo639L137 (18km) / TCo319L137 (36km)	TCo319L91 (36km)

CLOUDSC

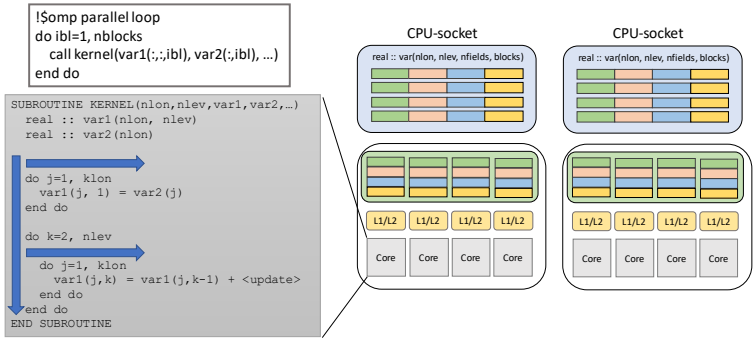
CLOUDSC : the IFS cloud scheme

- is one of the 7 physical process parameterizations in the IFS
- describes the effects of cloud microphysics on humidity-related variables in an atmospheric column
- is the largest single-scope bottleneck in an atmospheric IFS time-step

CLOUDSC

- the cloud scheme is representative of a significant fraction of atmospheric computational cost - the physical parameterisations
  - no horizontal dependencies in the physical processes
  - horizontal direction, first dimension of storage and work arrays, used for vectorization
  - no indirect addressing in horizontal dimension
- each vertical "column" involves small-array solve with loop-carried dependency

CLOUDSC

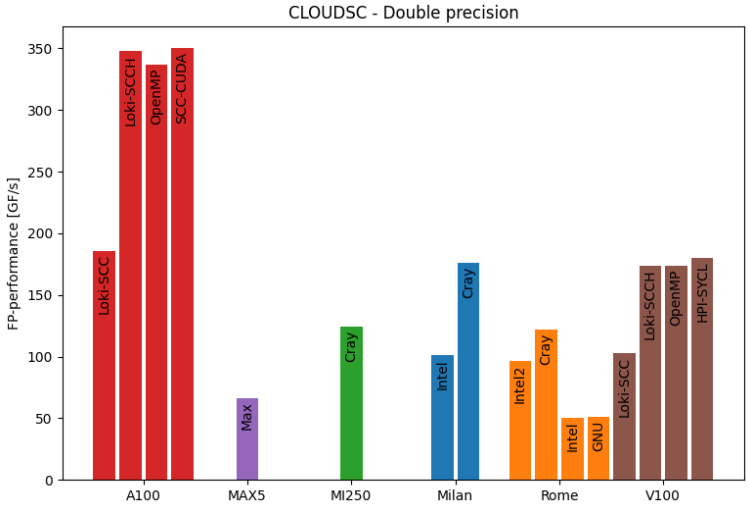

!$omp parallel loop
do ibl=1, nblocks
call kernel(var1(:,:,ibl), var2(:,ibl), …)
end do
SUBROUTINE KERNEL(nlon,nlev,var1,var2,…)
real :: var1(nlon, nlev)
real :: var2(nlon)
do j=1, klon
var1(j, 1) = var2(j)
end do
do k=2, nlev
do j=1, klon
var1(j,k) = var1(j,k-1) + <update>
end do
end do
END SUBROUTINE
CPU-socket
CPU-socket
real :: var(nlon, nlev, nfields, blocks)
real :: var(nlon, nlev, nfields, blocks)
L1/L2
Core

CLOUDSC

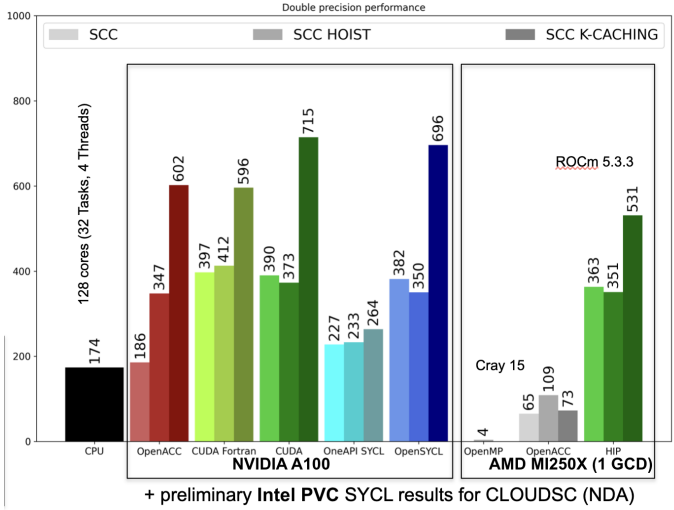
- CLOUDSC was one of the original ESCAPE[1] "dwarfs"
- more recently, released open-source on github under Apache-2 licence: `https://github.com/ecmwf-ifs/dwarf-p-cloudsc`
- thanks to EU project work (ESCAPE, EUROEXA, EUPEX, DestinationEarth) and open sourcing, has been ported to, and tested on, many platforms
- CPUs (Intel, AMD, ARM, A64FX, ...), GPUs (Nvidia, AMD, Intel, Apple), FPGA
- was used as playground to develop source-to-source framework with which physics code is being compile-time targetted for GPUs

[1] the ETP4HPC ESCAPE project `http://www.hpc-escape.eu`

CLOUDSC - performance taster



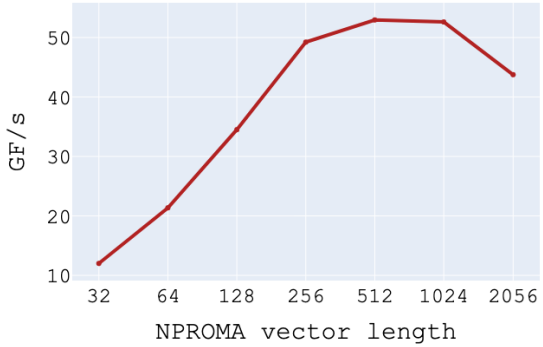
CLOUDSC - latest taste updates!



+ preliminary **Intel PVC** SYCL results for CLOUDSC (NDA)

CLOUDSC - NEC SX-Aurora TSUBASA

- preliminary porting very easy : platform mostly "just works"
- investigation of effect of vector length (NPROMA)



CLOUDSC - NEC SX-Aurora TSUBASA

- Our SX-AT 20B accelerator is nominally 2.45 TFLOPs DP - low fraction of peak

CLOUDSC - NEC SX-Aurora TSUBASA

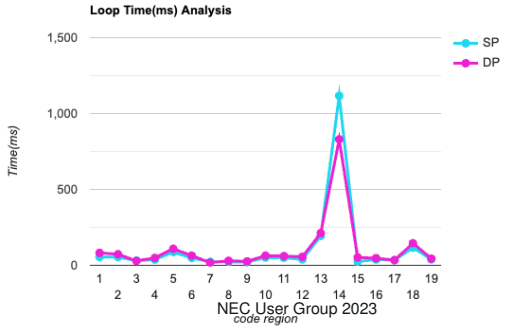
- Our SX-AT 20B accelerator is nominally 2.45 TFLOPs DP - low fraction of peak
- ftrace - NPROMA 256

FREQUENCY	EXCLUSIVE TIME[sec](%)	AVER.TIME [msec]	MOPS	MFLOPS	V.OP RATIO	AVER. V.LEN	VECTOR TIME	1CACHE MISS	CPU PORT CONF	VLD LLC HIT	E.%	PROC.NAME
640	3.259(79.2)	5.092	36360.0	16546.6	98.16	254.1	2.573	0.125	0.002	90.13		CLOUDSC
80	0.410(10.0)	5.130	36086.8	16422.3	98.16	254.1	0.325	0.016	0.000	89.92		-thread0
80	0.409(9.9)	5.112	36216.6	16481.4	98.16	254.1	0.323	0.016	0.000	90.52		-thread1
80	0.405(9.8)	5.059	36594.5	16653.3	98.16	254.1	0.319	0.015	0.000	90.09		-thread2
80	0.405(9.8)	5.059	36595.0	16653.6	98.16	254.1	0.319	0.015	0.000	90.43		-thread3
80	0.405(9.8)	5.057	36610.2	16660.5	98.16	254.1	0.319	0.015	0.000	89.83		-thread4
80	0.404(9.8)	5.056	36620.4	16665.1	98.16	254.1	0.319	0.015	0.000	89.78		-thread5
80	0.410(10.0)	5.130	36086.3	16422.0	98.16	254.1	0.324	0.016	0.000	90.11		-thread6
80	0.410(10.0)	5.131	36083.9	16421.0	98.16	254.1	0.324	0.016	0.000	90.40		-thread7
19	0.338(8.2)	17.807	19749.7	9250.4	98.69	229.1	0.336	0.001	0.000	55.76		VALIDATE_MOD::VALIDATE_R2
400	0.157(3.8)	0.393	16242.7	0.0	88.05	72.7	0.144	0.004	0.000	99.99		EXPAND_MOD::EXPAND_R2$1

CLOUDSC - NEC SX-Aurora TSUBASA

Optimisation efforts

- hot-spot profiling with Ftrace API
- main per-loop results :
  - SP is not very efficient out of the box for this code (many loops get no more than 10% improvement)
  - single loop dominates the cost



CLOUDSC - NEC SX-Aurora TSUBASA

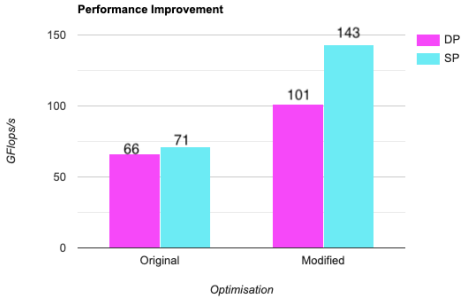
- Main culprit in expensive region is a sorting loop

```
DO JN=1,NCLV
  DO JL=KIDIA,KFDIA
    IF (LLINDEX1(JL,JN) .AND. ZRATIO(JL,JN)<ZMIN(JL)) THEN
      IORDER(JL,JM)=JN
      ZMIN(JL)=ZRATIO(JL,JN)
    ENDIF
  ENDDO
ENDDO
```

CLOUDSC - NEC SX-Aurora TSUBASA

- Main culprit in expensive region is a sorting loop

```
DO JN=1,NCLV
  DO JL=KIDIA,KFDIA
    IF (LLINDEX1(JL,JN) .AND. ZRATIO(JL,JN)<ZMIN(JL)) THEN
      IORDER(JL,JM)=JN
      ZMIN(JL)=ZRATIO(JL,JN)
    ENDIF
  ENDDO
ENDDO
```



CLOUDSC - NEC SX-Aurora TSUBASA

Data structure ordering

- input-output array ordering
  (HOR, VERT, FLD, BLOCK) vs *(HOR, FLD, VERT, BLOCK)*
- ordering of local arrays with NCLV dimension
- almost no effect at all!

CLOUDSC - NEC SX-Aurora TSUBASA

Performance update (thank you NEC DE !)

CLOUDSC - NEC SX-Aurora TSUBASA

Performance update (thank you NEC DE !)

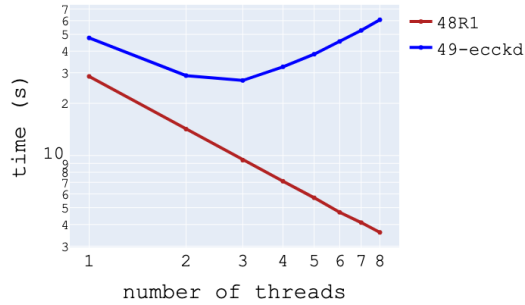
	VE 2	VE 3 (compatibility)	VE 3 (native)
VE SKU	20B	30A	30A
thread count	8	16	16
GF/s DP	103.3	199.6	205.3
GF/s SP	142.8	277.6	290.5

ecRAD - NEC SX-Aurora TSUBASA

- radiation scheme developed by R Hogan at ECMWF
- used by a number of academic research groups and weather centres
- now released open-source : `https://github.com/ecmwf-ifs/ecrad`
- stand-alone execution mode, stand-alone coding style
- Work has just started

ecRAD - NEC SX-Aurora TSUBASA

threads	48R1	49R1 - ecckd
1	28.6	47.7
2	14.2	28.9
3	9.4	27.1
4	7.1	32.5
5	5.7	38.4
6	4.7	45.5
7	4.1	52.7
8	3.6	60.7



ecRAD - NEC SX-Aurora TSUBASA

Effect of -ftrace compiler option