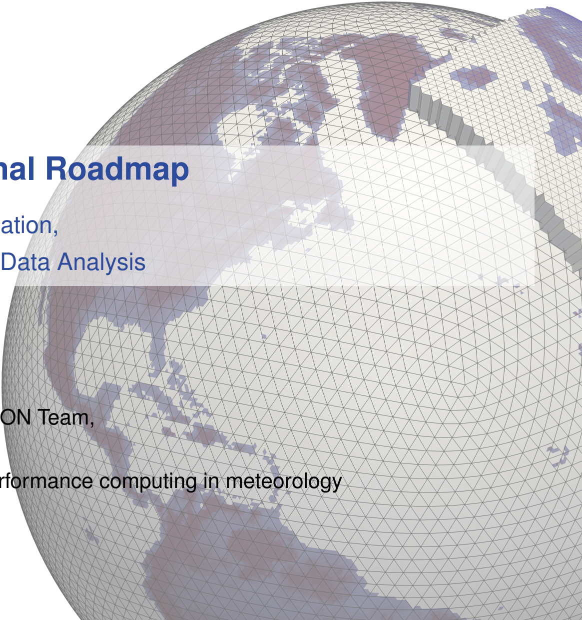
Deutscher Wetterdienst
Wetter und Klima aus einer Hand

DWD’s Operational Roadmap

Implications for Computation, Data Management and Data Analysis



Florian Prill, DWD + the ICON Team,

18th Workshop on high performance computing in meteorology
September 24 – 28, 2018

Talk Outline – DWD's Operational Roadmap

- Evolution of DWD's numerical weather prediction suite (NWP)

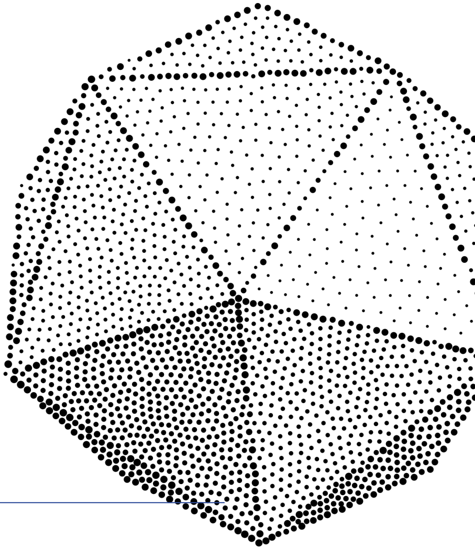
"Equation 1"

ICON-NWP = Model Code + User Community

- DWD NWP strategy for the mid-term future

"Equation 2"

HPC = Computation + Data Management



Evolution of DWD's Numerical Weather Prediction Suite

Operational NWP at DWD

24/7 operations in Offenbach, Germany.

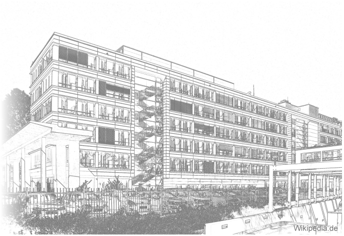
- History of DWD's numerical weather prediction dates back to 1966.
- Numerous meteorological and climatological services.
- Daily provision of boundary data for ~ 21 partners, driving regional models.

Redundantly installed **HPC system**:

Cray XC40: Intel HSW/BDW

41 472 + 34 560 Cores
1.46 + 1.22 PFLOPS

+ Linux commodity cluster for pre- and post-processing.



Operational NWP at DWD

24/7 operations in Offenbach, Germany.

- History of DWD's numerical weather prediction dates back to 1966.
- Numerous meteorological and climatological services.
- Daily provision of boundary data for ~ 21 partners, driving regional models.

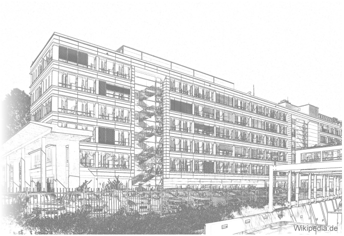
Redundantly installed **HPC system**:

Cray XC40: Intel HSW/BDW

41 472 + 34 560 Cores
1.46 + 1.22 PFLOPS

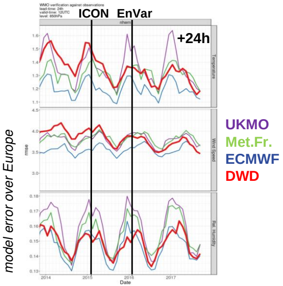
+ Linux commodity cluster for pre- and post-processing.

Upcoming procurement Q4/2018 !



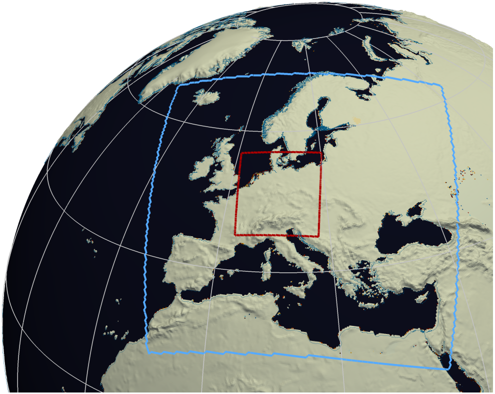
The Past: The NWP Process Chain Until 2015

- **GME:** 35 km height, 20 km mesh size global weather forecasts ≤ 7 days.
- **COSMO-EU** and **COSMO-DE**: regional components with 7 km mesh size (EU), 2.8 km mesh size (DE: Germany).
- **3dVar** variational data assimilation.



Example:

GME 20 km / L60:
≈ 88.5 million grid points.



Modernization Since 2015

2015 Introduction of the **ICON model**.
ICON = ICON = ICOsahedral Nonhydrostatic model.

ICON replaces both GME and COSMO-EU:
13 km / 6.5 km mesh size, up to 5 km height.

2016 **En-Var hybrid data assimilation** employing a first-guess ensemble, 40 km / 6.5 km mesh size.

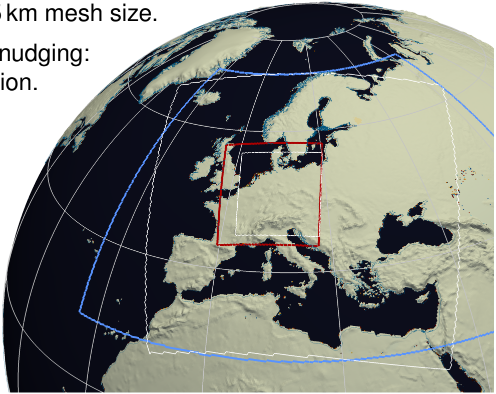
2017 Replacement for regional model nudging:
KENDA ensemble data assimilation.

2018 **COSMO-D2**: enlarged domain with 2.2 km, 65 levels.

2018 Introduction of **ICON ensemble**
40 km globally, 20 km nested.

Example:

13 km global × 90 levels
+ 6.5 km Europe nest × 60 levels
≈ 305 unknowns for a 3D variable.



ICON = ICOsahedral Nonhydrostatic Model

$$
\begin{aligned}
\partial_t v_n \quad &+ (\zeta + f)\, v_t \quad + \partial_n K + w\,\partial_z v_n \;=\; -c_{pd}\theta_v \partial_n \pi \\
\partial_t w \quad &+ \mathbf{v}_n \cdot \nabla w \quad + w\,\partial_z w \;=\; -c_{pd}\theta_v \partial_z \pi - g \\
\partial_t \rho \quad &+ \nabla \cdot (\mathbf{v}\rho) \;=\; 0 \\
\partial_t(\rho\theta_v) &+ \nabla \cdot (\mathbf{v}\rho\theta_v) \;=\; 0
\end{aligned}
$$

(v_n, w, ρ, θ_v: prognostic variables)

- v_n, w: velocity components
- ρ: density
- θ_v: virtual potential temperature
- K: horizontal kinetic energy
- ζ: vertical vorticity component
- π: Exner function

Joint development project of DWD and **Max-Planck-Institute for Meteorology**.

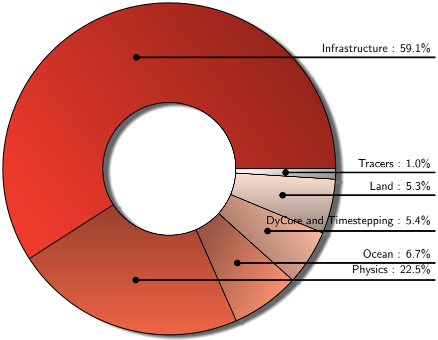
Close collaboration with the **Karlsruhe Institute of Technology** and the **DKRZ**.

- Nonhydrostatic dynamical core on an icosahedral-triangular Arakawa C-grid with mass-related quantities at cell circumcenters.
- Local mass conservation and mass-consistent transport.
- Two-way nesting with capability for multiple overlapping nests.
- Limited area mode also available.

The ICON Model: Building Blocks

Components: **weather**, **climate**, **ocean**, **land**.

- $\sim$ 25 active developers.
- Built to run on x86-based MPPs; **scales to** $\mathcal{O}\left(10^4+\right)$ **cores**.
- system layer encapsulated in C libraries: *e.g. I/O, communication, calendar.*
- 300 000 (logical) lines of **Fortran code**.
- Looooong life cycle: some parts of the ICON model $\sim$ 1986.

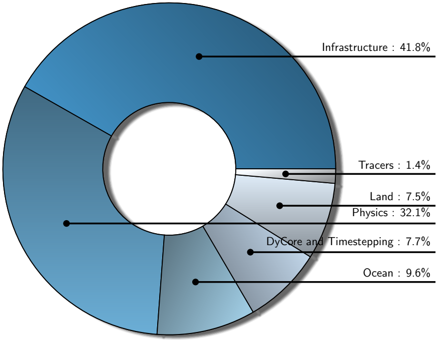


(ICON code portions)

The ICON Model: Building Blocks

Components: **weather**, **climate**, **ocean**, **land**.

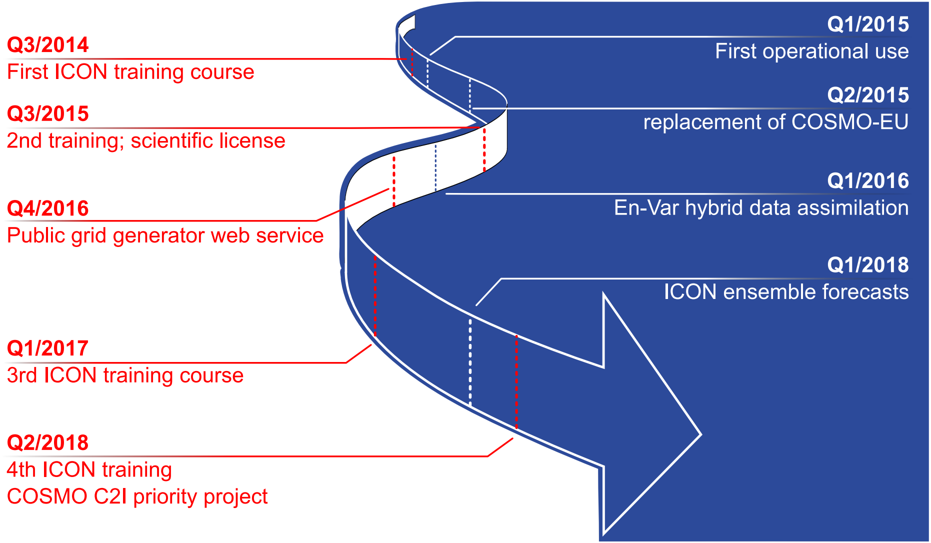
- $\sim$ 25 active developers.
- Built to run on x86-based MPPs; **scales to** $\mathcal{O}\left(10^4+\right)$ **cores**.
- system layer encapsulated in C libraries: *e.g. I/O, communication, calendar.*
- 300 000 (logical) lines of **Fortran code**.
- Looooong life cycle: some parts of the ICON model $\sim$ 1986.



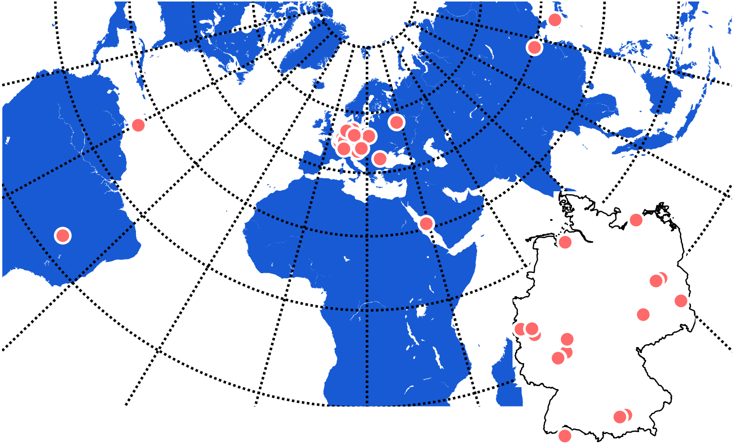
(ICON, Fortran part)

DWD
Max-Planck-Institut für Meteorologie
Karlsruhe Institute of Technology
DKRZ DEUTSCHES KLIMARECHENZENTRUM

ICON-NWP = Model Code + USER COMMUNITY



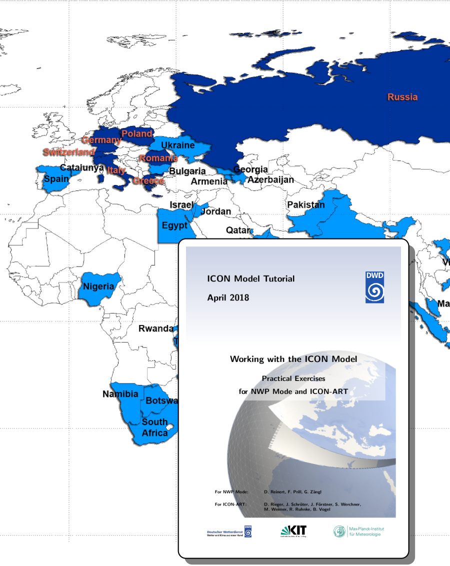
ICON Licensees Map



Alfred-Wegener-Institut, Bremerhaven
BTU Cottbus-Senftenberg
Caribbean Institute for Meteorology and Hydrology
Chinese Academy of Meteorological Sciences (CAMS)
DLR-Institut für Physik der Atmosphäre Oberpfaffenhofen
Forschungszentrum Jülich GmbH
Freie Universität Berlin
Goethe Universität Frankfurt am Main
Hydrometcentre of Russia
IBL Software Engineering, Bratislava
IMGW-PIB, Warsaw
INMET -Instituto Nacional de Meteorologia, Brasilia
Institut für Geographie, Justus-Liebig-Universität
Institut für Geophysik und Meteorologie, Uni. zu Köln
Institut für Meteorologie, Universität Leipzig
Institut für Ostseeforschung Warnemünde
Institut für Physik der Atmosphäre, JGU Mainz
King Abdulaziz University (KAU)
Leibniz Institut für Troposphärenforschung e.V. (TROPOS)
Lomonosov Moscow State University
Luxembourg Institute of Science and Technology (LIST)
Max-Planck-Institut für Chemie, Mainz
Meteo Romania: National Meteorological Administration
MeteoSwiss, Zurich-Airport
Meteorologisches Institut, LMU München
Potsdam Institute for Climate Impact Research
Pukyong National University
Rheinisches Institut für Umweltforschung a. d. Uni. zu Köln
University of Bonn, Meteorological Institute
Wegener Center for Climate and Global Change, Graz

Consortium for Small Scale Modelling.

- DWD is a member of the COSMO community which provides the regional weather models for numerous national MetServices.
- e. g. Switzerland, Italy, Greece, Poland, Romania, Russia, Israel.
- 20 years of experience with limited-area modelling.
- **C2I-Project 2018-2021**: Ensure a smooth transition from the COSMO model to ICON-NWP for all partners.



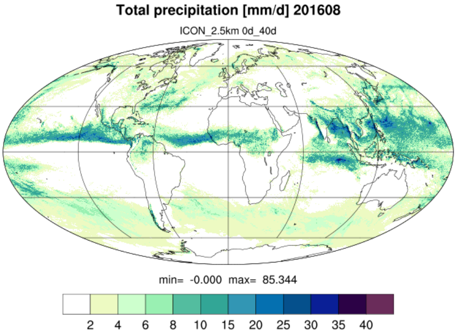
ICON R02B10 DYAMOND Simulation

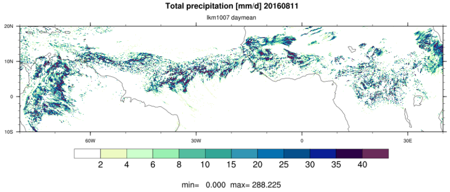
Setup as operational NWP,
40 days simulated.

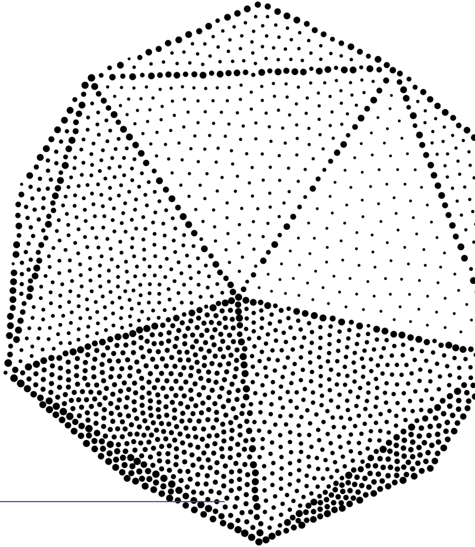
83.9 mio columns, 90 levels.

- minus convection, gwd, sso.
- $\Delta t = 22.5\,\mathrm{s}$.
- run on DKRZ "`mistral`":
  $\sim$ 6 days/day on 540 nodes,
  24 Haswell cores.
- Output: 25 TB.

[MPI-M; D. Klocke, DWD]







DWD NWP Strategy
for the Mid-Term Future

SINFONY Project: 2019 – 2022

Seamless **IN**tegrated **FO**recasti**N**g s**Y**stem:

Product created as a blend of nowcasting and very short-range forecasts .

Update 5 – 15 min — *Rapid Update Cycle (hourly)*

First steps:

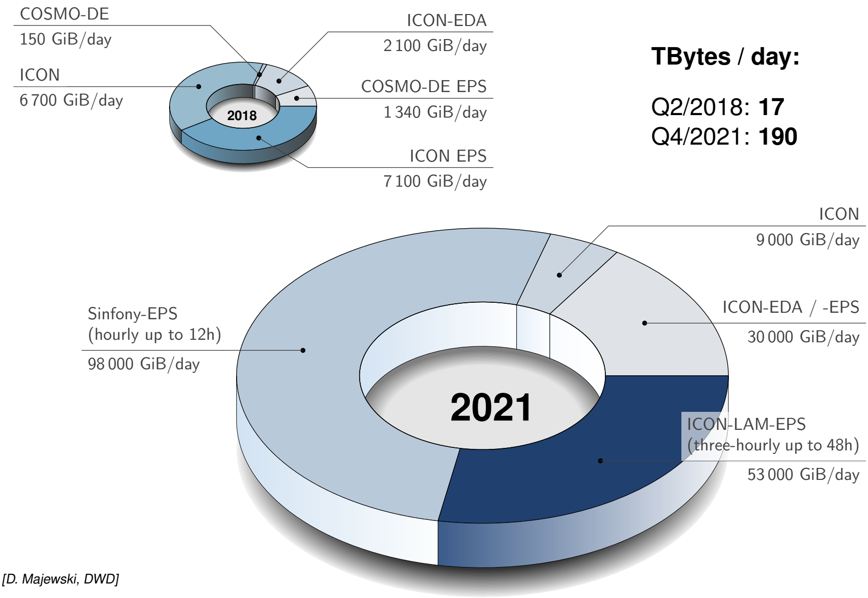
- ICON deterministic: bigger size of nest ICON-EU → ICON-EU(NA)^2.
- ICON ensemble 26 km / 13 km / 6.5 km 40 and 250 members (EPS / data ass.).

Major upgrade:

- Rapid Update Cycle: 12 h ensemble forecast based on very short data cut off with hourly update.
- ICON limited area model ensemble with 2 km / 1 km grid size; update every 3 h.

[D. Majewski, DWD]

NWP Data Production: Expected Growth 2018 – 2021

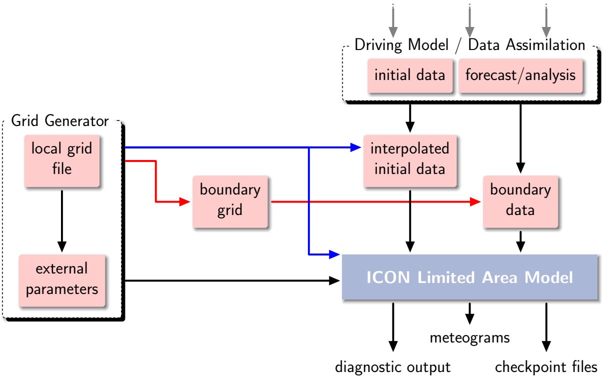


TBytes / day:

Q2/2018: **17**
Q4/2021: **190**

[D. Majewski, DWD]

HPC = Computation + DATA MANAGEMENT



Data-oriented perspective on the NWP process chain:

Processing and storage of high-resolution data: Critical issue!

- Bypass database, discard intermediate products.
- In-situ processing without involving storage resources.

Data Management: Provenance

Data provenance is crucial for the reproducibility and the analysis of defects in the computational geosciences.

One important building block: **Track the computational meshes** and the external parameter sets through all transformations of the scientific workflow.

GRIB2 meta-data: Two independent 128 bit key/value pairs:

`UUIDOfHGrid`	reference to the horizontal (triangular, 2D) grid.
`UUIDOfVGrid`	reference to the vertical (three-dimensional) grid.

Requirements for ICON fingerprints

- repeatable and processor-independent ✓
- should reliably detect
  - small defects (example: bit flips) ✓
  - permutations (example: reordered cells) ✓
  - large differences (example: new topography data) ✓

Parallel Fingerprint Calculation

Fast calculation, but should not consume global arrays.
Computing fingerprints in parallel sounds paradoxical . . .

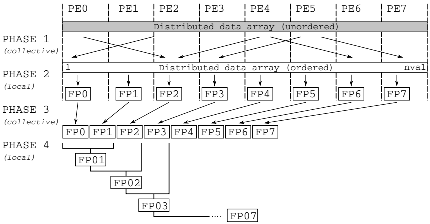
- **independent data-parallel** computation, but
- result is **sensitive to ordering and bit flips** by definition.

ICON: **Rabin's fingerprinting method** (1981)

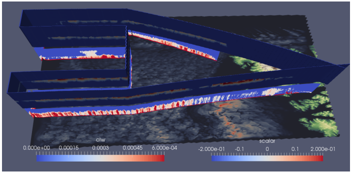
$$f(A) := A(t) \bmod p(t)$$

allows concatenation and thus parallelization:

$$f(\text{concat}(A,B)) = f(f(A) * f(t^l)) + f(B)$$



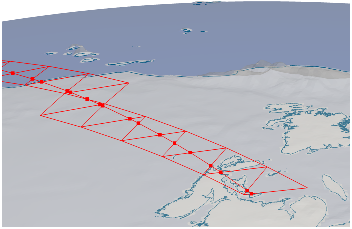
Data Management: In-Situ Reduction



cloud liquid water `clw`, ICON model

Example: Virtual Flight Tracks
Obtain diagnostic output along a virtual flight track.

- Interpretation of weather radar imagery during flight campaigns.
- Dumping and processing of all the data calculated during the simulation would consume too much time and storage.



[V. Schemann, Uni. Cologne; T. Göcke, DWD]

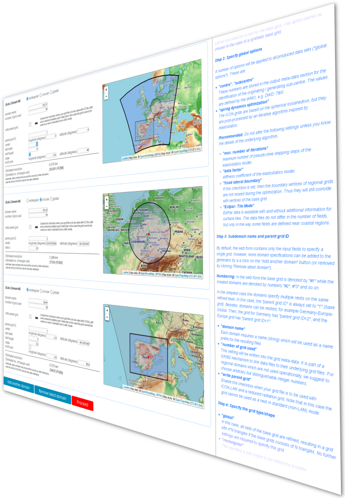
ICON Grid Generator and External Parameters Web Service

In addition to the grid, real-data ICON runs require external parameter files:

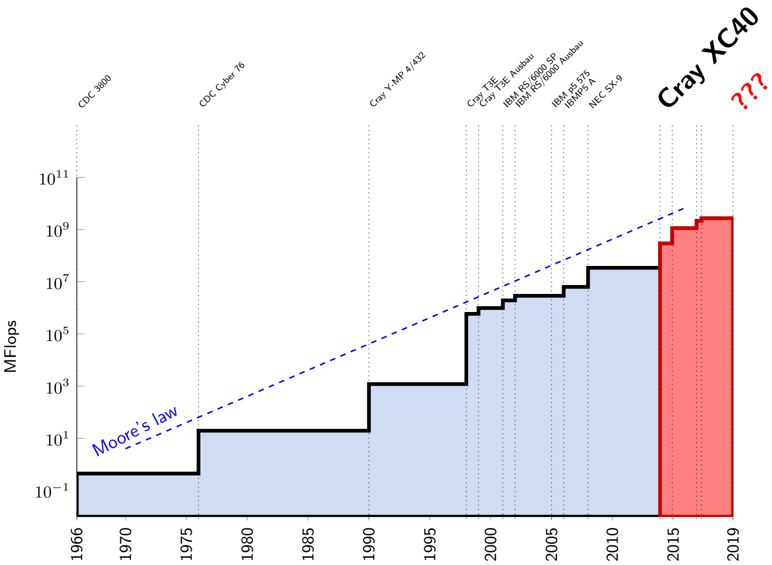
- land-sea mask,
- orography,
- soil type,

and other geographical data sets.

Web service with live-preview to generate own grids and external data.



DWD HPC Infrastructure: Past, Current

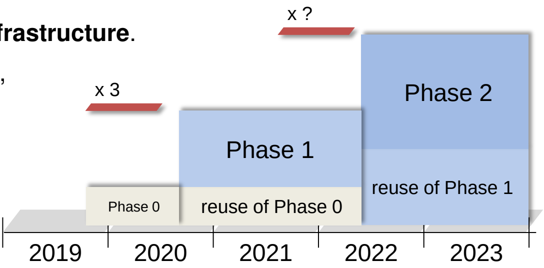


Project HPC19

Renewal of DWD's central production infrastructure:

Tender 1 (HPC)	2x HPC, 2x Storage, Housing
Tender 2 (Data Management)	2x Data Management System, 2x Storage

- Increase operational computing power in two steps.
- Adaptation to increased storage requirements.
- **Two systems: operations** / **research**
  research system always 20% larger
  than production system.
- Establish a **georedundant infrastructure**.
- **Benchmarks** based on ICON, COSMO, data assimilation.



HPC = Data Management + COMPUTATION

The “old” computational challenges remain . . .

code balance

ratio of memory data traffic to arithmetic work

usually much larger than

machine balance

ratio of memory bandwidth to peak arithmetic performance

For **ICON** this means (e.g. in iterative stencil loops)

x86 architectures:	Hybrid MPI + OpenMP parallelization of all components, efficient use of caches is important.
GPGPUs:	Higher floating point peak performance, but host-device memory latency.

PASC ENIAC project 2017 – 2020

ENIAC = Enabling the ICON model on heterogeneous Architectures.
OpenACC port + DSL “GridTools” (encoding stencil info in C++ data types).

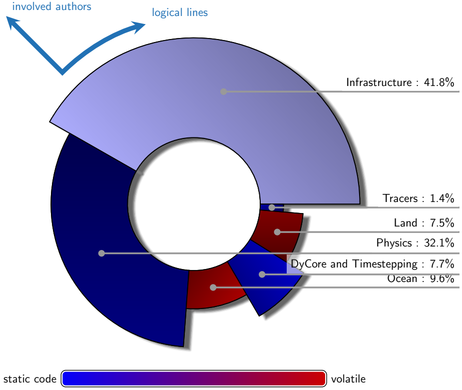
What is “Code Infrastructure”?

Characterized by relatively low arithmetic intensity? - No.

Examples:

Internal post-processing, interpolation, I/O layer, asynchronous communication.

- Many developers involved!
- Neither “static” nor “volatile”.
- Tendency to grow and duplicate.
- Much larger code portions than it used to be in the past.



High technical complexity of code:

Strong focus on massively parallel computer architectures, efficient data management and community adaptability.

“Equation 1”

ICON-NWP = MODEL CODE + USER COMMUNITY

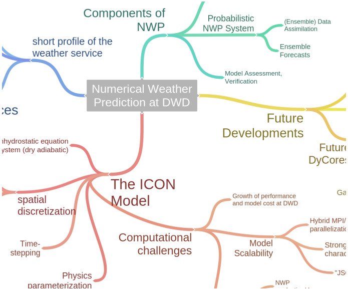
“Equation 2”

HPC = COMPUTATION + DATA MANAGEMENT

Challenges in the
DWD model chain 2019 – 2022:

- modelling
- computational complexity
- data handling migraines

accelerators won’t help us here!



Florian Prill

Met. Analyse und Modellierung
Deutscher Wetterdienst
e-mail: Florian.Prill@dwd.de

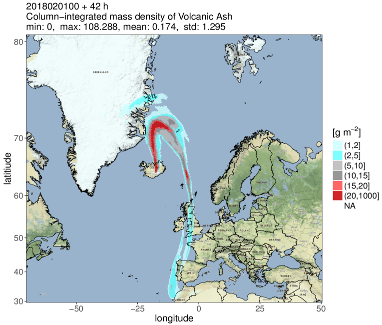
ICON-ART
Operational emergency system

Operational model capable of computing the spread of mineral dust, volcanic ash, and radioactive particles.

- Introduction 2019 – 2022.
- global ICON-ART simulation with nests over Europe: 26 / 13 / 6.5 km grid size.
- E. g. prediction of impact on air transportation.

Regional pollen forecast

ICON-ART simulations over Europe, Germany: 6.5 km grid size.



[J. Förstner, DWD]