

Spectrum Scale at Deutscher Wetterdienst

Fulfilling high throughput super-computing storage needs
with NEC GxFS Storage Appliances

Spectrum Scale User Meeting, March 8 2017, Ehningen, Germany

Manuel Reiter, DWD

Thanks to T. Bernecker, NEC, and T. Hanisch, DWD, for contributed slides and graphics!





Deutscher Wetterdienst (DWD)

is subordinate to the

Federal Ministry of Transport and Digital Infrastructure (BMVI)

- founded as per 11 November 1952
- legal basis: Law on the Deutscher Wetterdienst of 10 September 1998

Our overall task

Protection of life and property from weather and climate-related risks and provision of support to the general public and the economy by means of reliable meteorological services



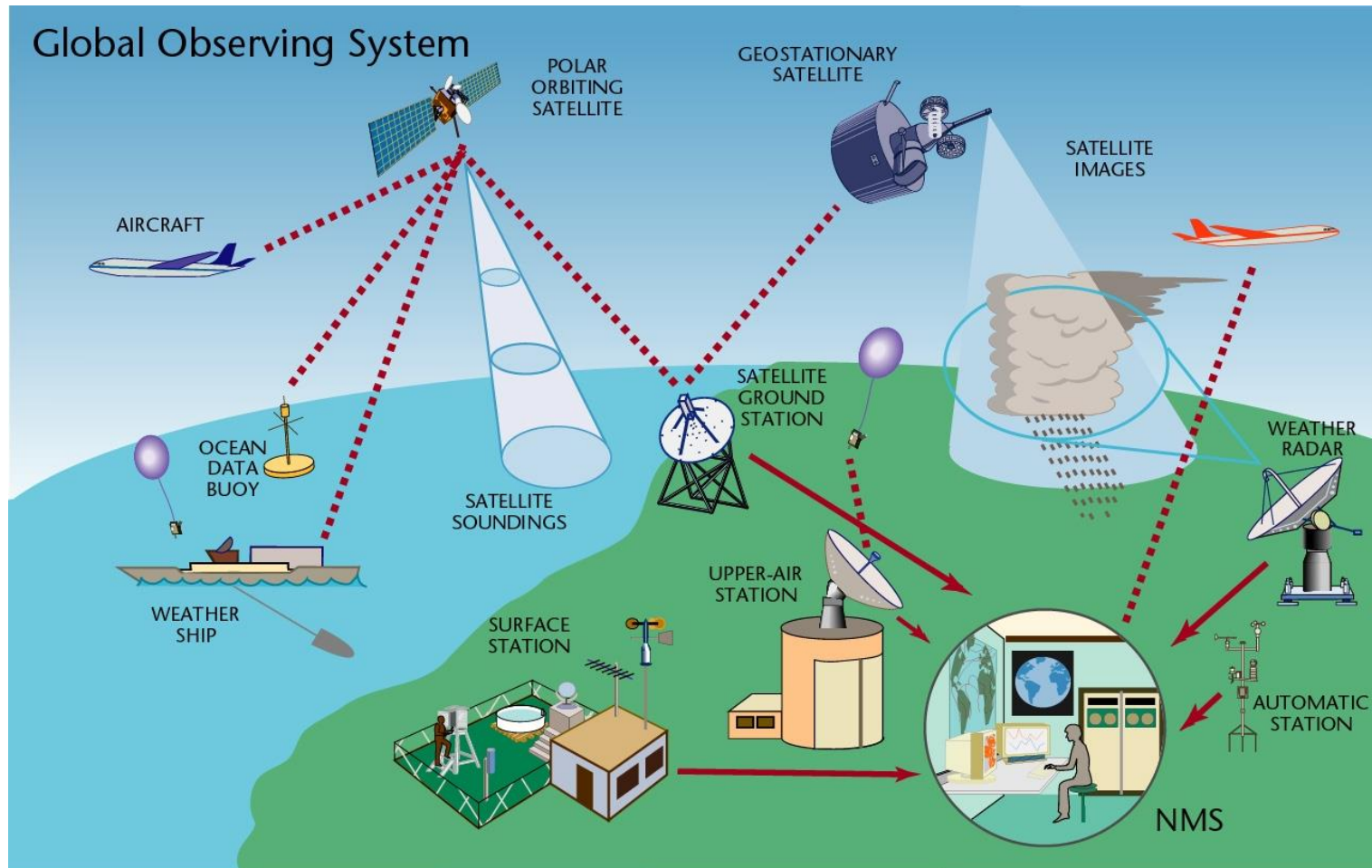
DWD throughout Germany



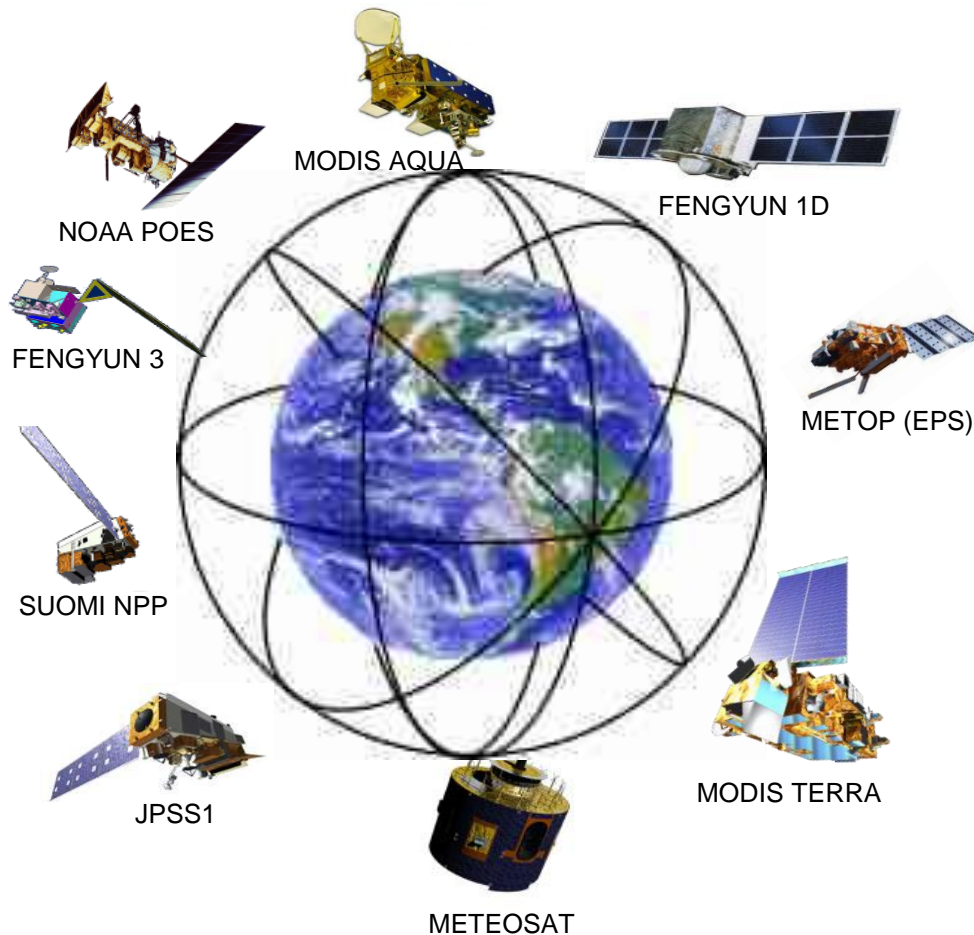
- 2500 staff members
- Headquarters in Offenbach am Main
- 6 branch offices (Hamburg, Potsdam, Leipzig, Essen, Stuttgart, Munich), partly with more than 100 staff members
- 5 regional climate offices providing consultancy services in the field of climate and environment
- 5 MET -and 3 agrometeorological advisory centres
- 182 main weather stations
- 48 radioactivity measuring stations
- 1,784 secondary weather and precipitation stations
- 1,184 phenological observation sites
- 17 weather radar sites in Germany
- 2 meteorological observatories
- 18 automatic shipboard weather stations
- 2 staffed main weather stations aboard research ships
- 461 weather reporting stations on board merchant ships
- 4 automated shipboard aerological stations (ASA)
- 9 upper-air stations with around 7,000 radiosonde launches per year and 3 mobile measuring units
- Use of several meteorological satellites



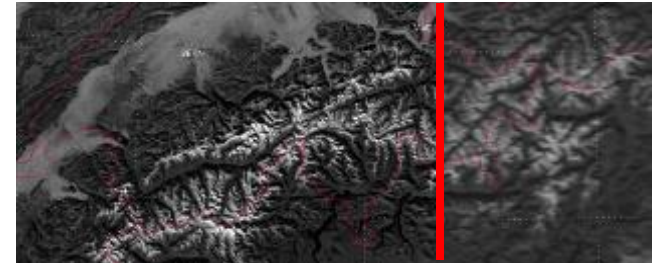
Meteorological input data



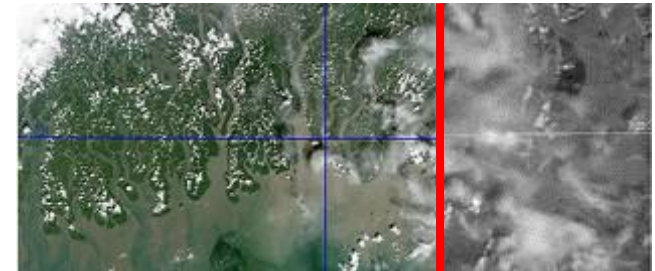
Weather Satellites



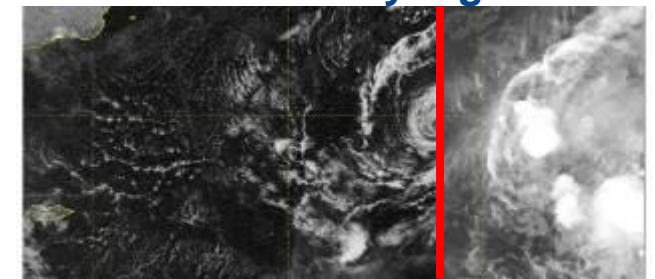
Higher resolution



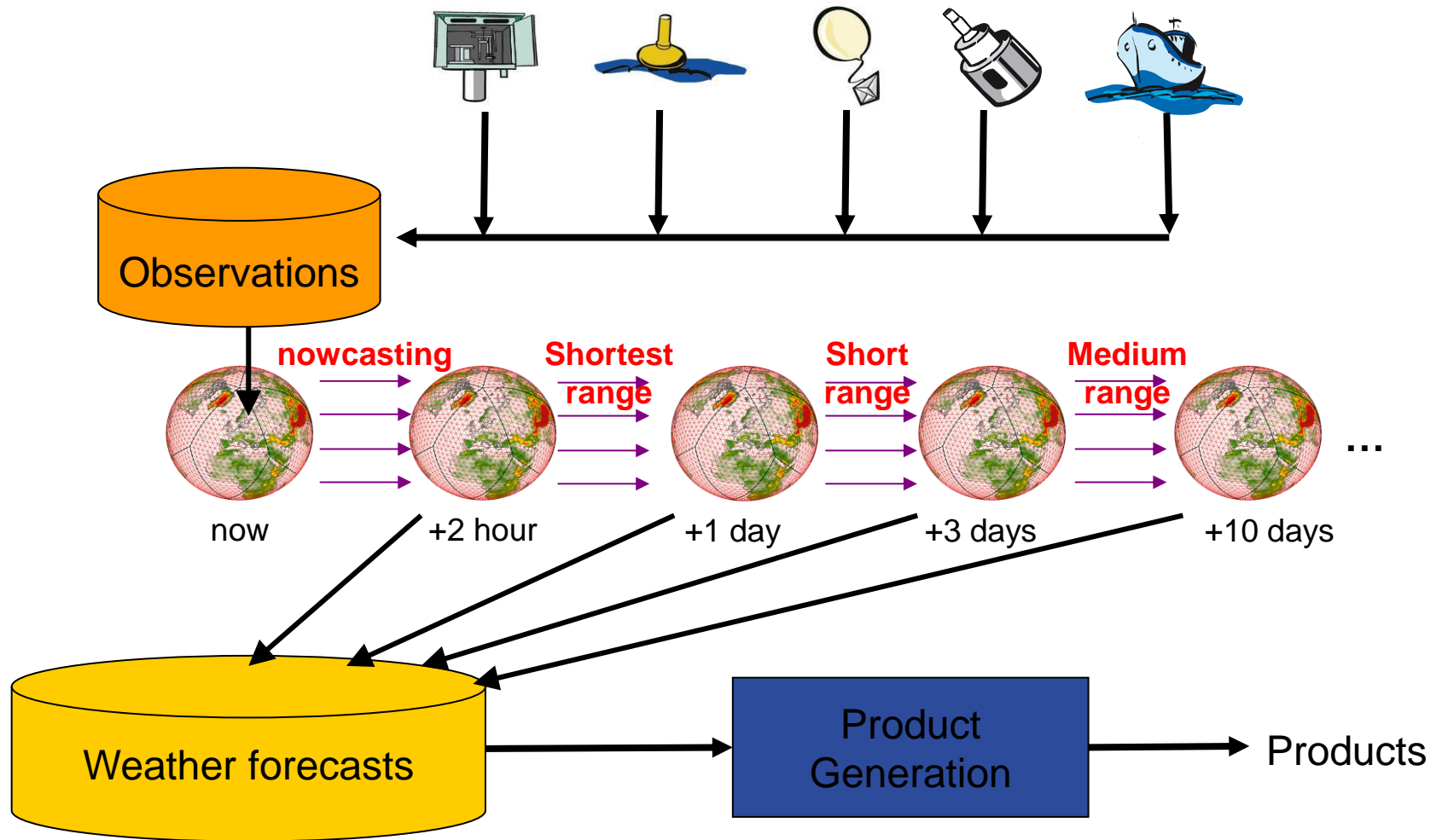
Color



Visible channel by night



From observation to meteorological product



The **deterministic** NWP system at DWD

Deutscher Wetterdienst
Wetter und Klima aus einer Hand



Global model **ICON**

Grid spacing: **13** km

Layers: **90**

Forecast range:

180h at 00, 12 UTC

120h at 06, 18 UTC

30h at 03, 09, 15, 21 UTC

1 grid element: 173 km²

ICON-EU Nest

Grid spacing: **6.5** km

Layers: **60**

Forecast range:

120h at 00, 06, 12, 18 UTC

30h at 03, 09, 15, 21 UTC

1 grid element: 43 km²

COSMO-DE

Grid spacing: **2.8** km

Layers: **50**

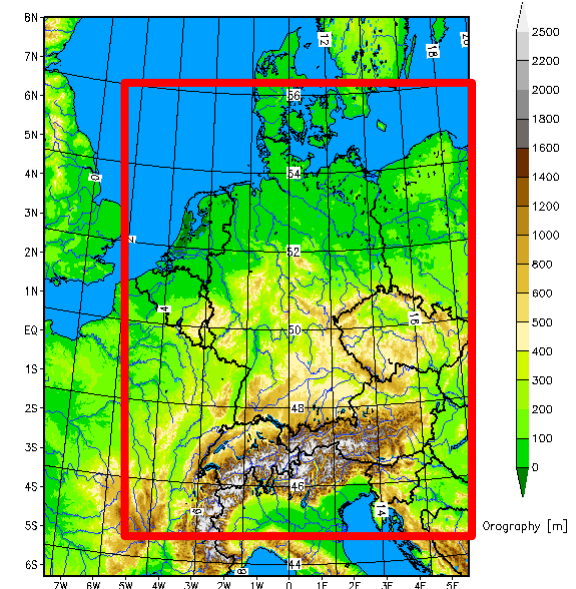
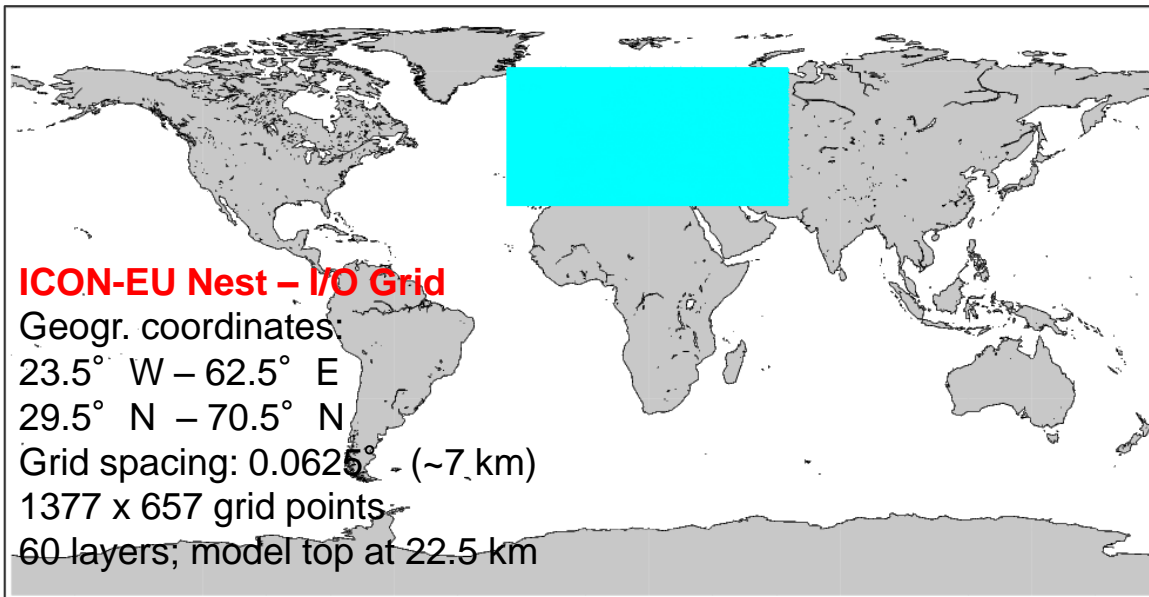
Forecast range:

27h / **45h** at 00, **03**, 06, 09,

12, 15, 18, 21 UTC

421x461 grid points

1 grid element: 8 km²



The probabilistic NWP system at DWD

Deutscher Wetterdienst
Wetter und Klima aus einer Hand



ICON-EPS: M40

Grid spacing: 40 km

Layers: 90

Forecast range:

180h at 00, 12 UTC

120h at 06, 18 UTC

30h at 03, 09, 15, 21 UTC

1 grid element: 1638 km²

ICON-EU-EPS Nest

Grid spacing: 20 km

Layers: 60

Forecast range:

120h at 00, 06, 12, 18 UTC

30h at 03, 09, 15, 21 UTC

1 grid element: 407 km²

COSMO-DE-EPS: M20

Grid spacing: 2.8 km

Layers: 50

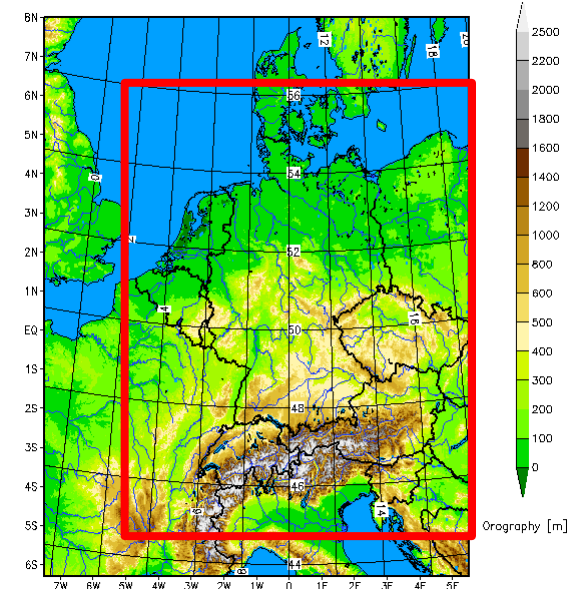
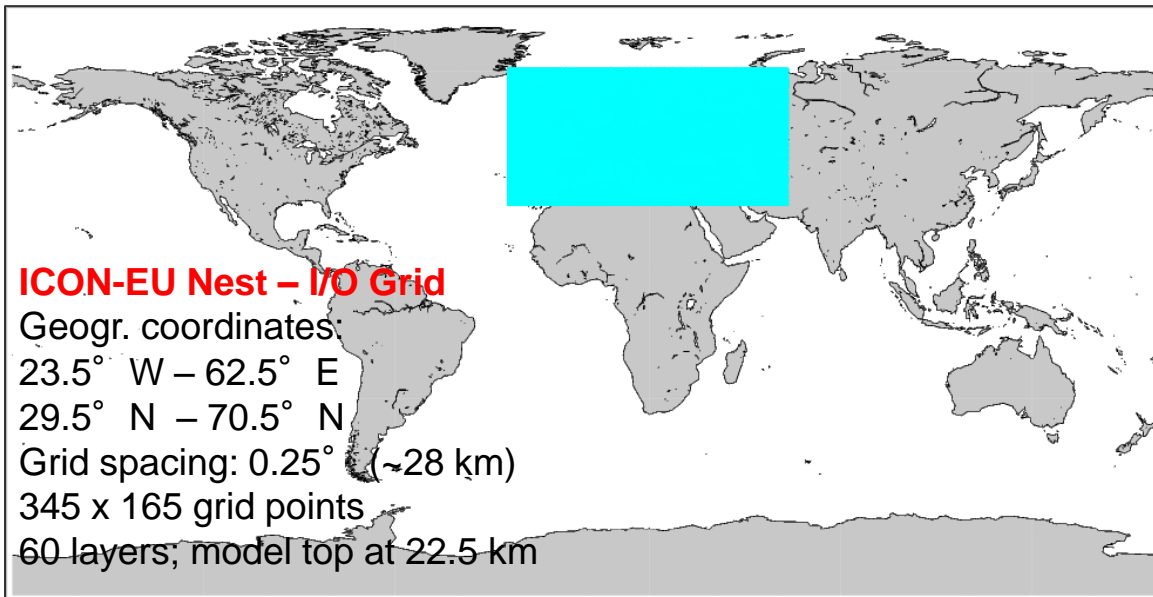
Forecast range:

27h / 45h at 00, 03, 06, 09,

12, 15, 18, 21 UTC

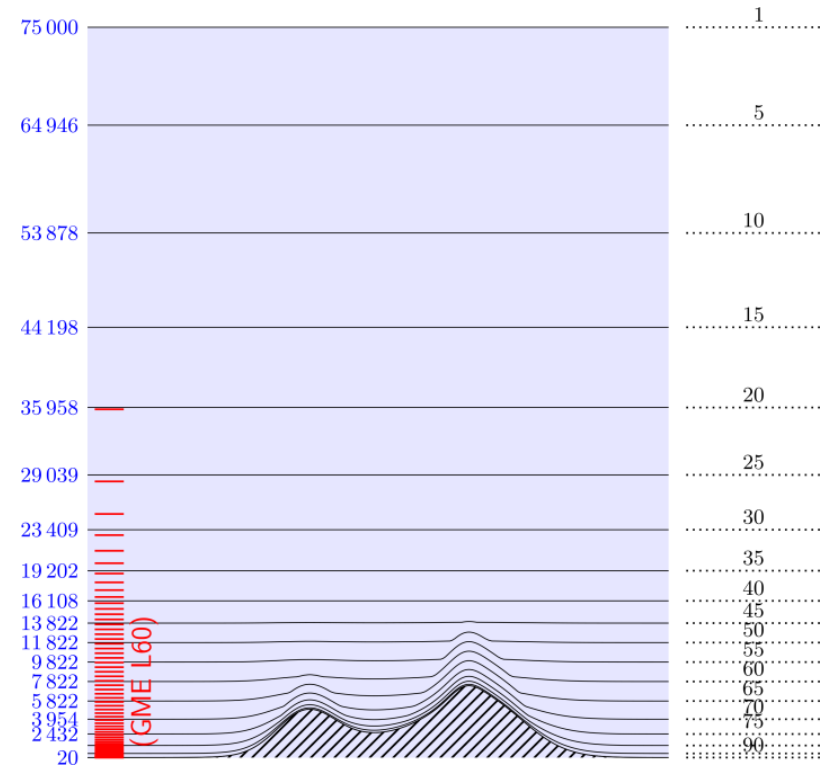
421x461 grid points

1 grid element: 8 km²



ICON and ICON-EU Nest

Deutscher Wetterdienst
Wetter und Klima aus einer Hand

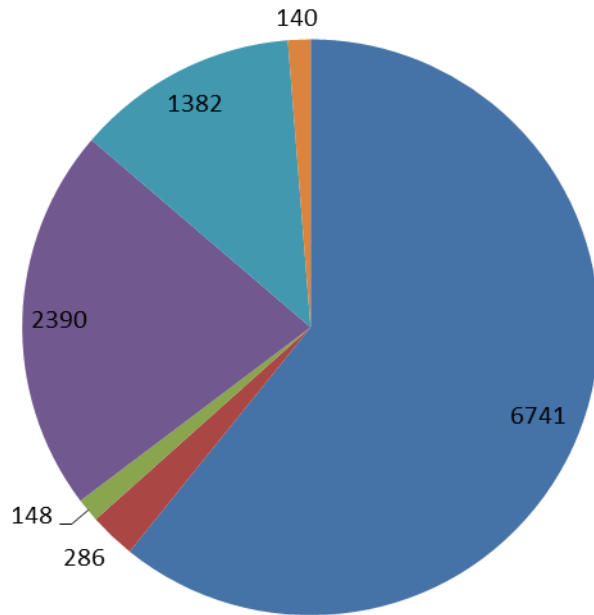


20.01.2015: ICON-global
21.07.2015: ICON-EU Nest

(Grid spacing 13 km, 90 layers)
(Grid spacing 6.5 km, 60 layers)

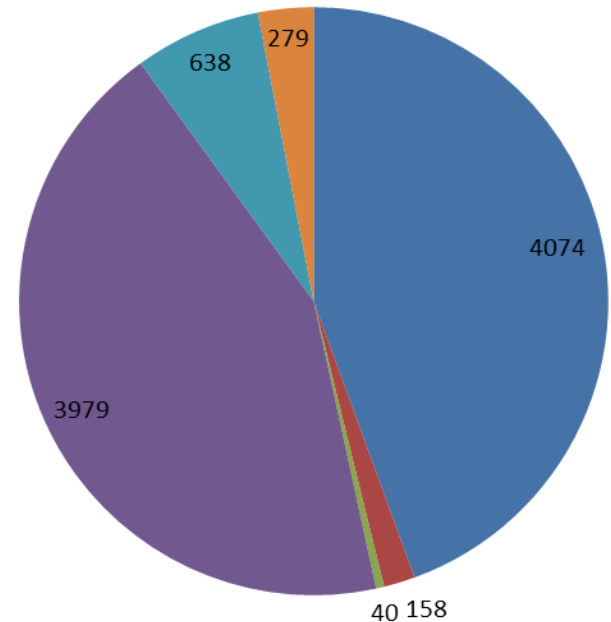


NWP data production (GByte / day)



Operational NWP
13334 GByte / day

- ICON 1
- COSMO-EU 2
- COSMO-DE 3
- ICON-EDA/EPS 4
- COSMO-DE-EPS 5
- COSMO-BC-EPS 6

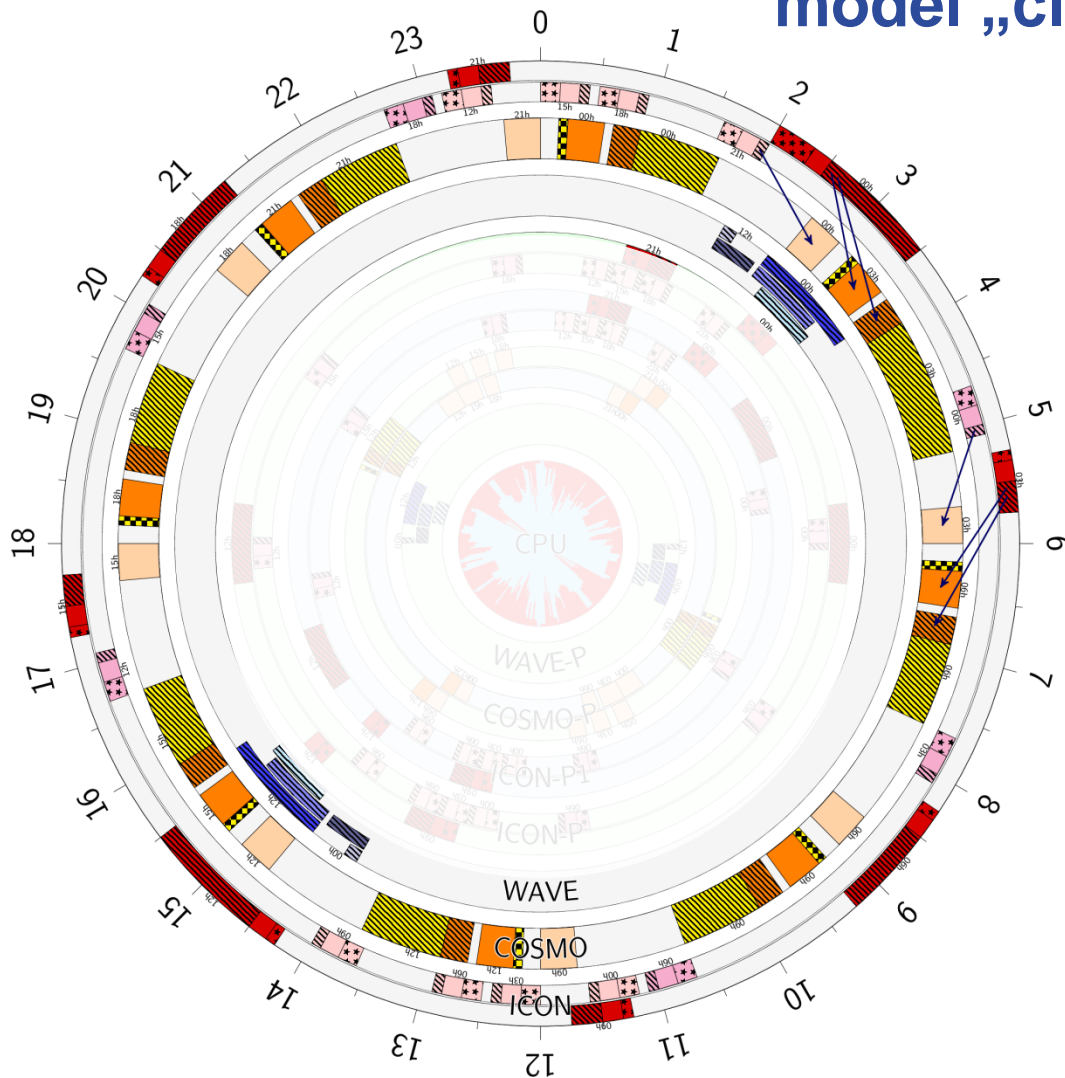


Pre-operational NWP
9168 GByte / day

+ observations ~800 Gbyte / day, expected to increase drastically

model „clock”

Operational timetable of the DWD model suite

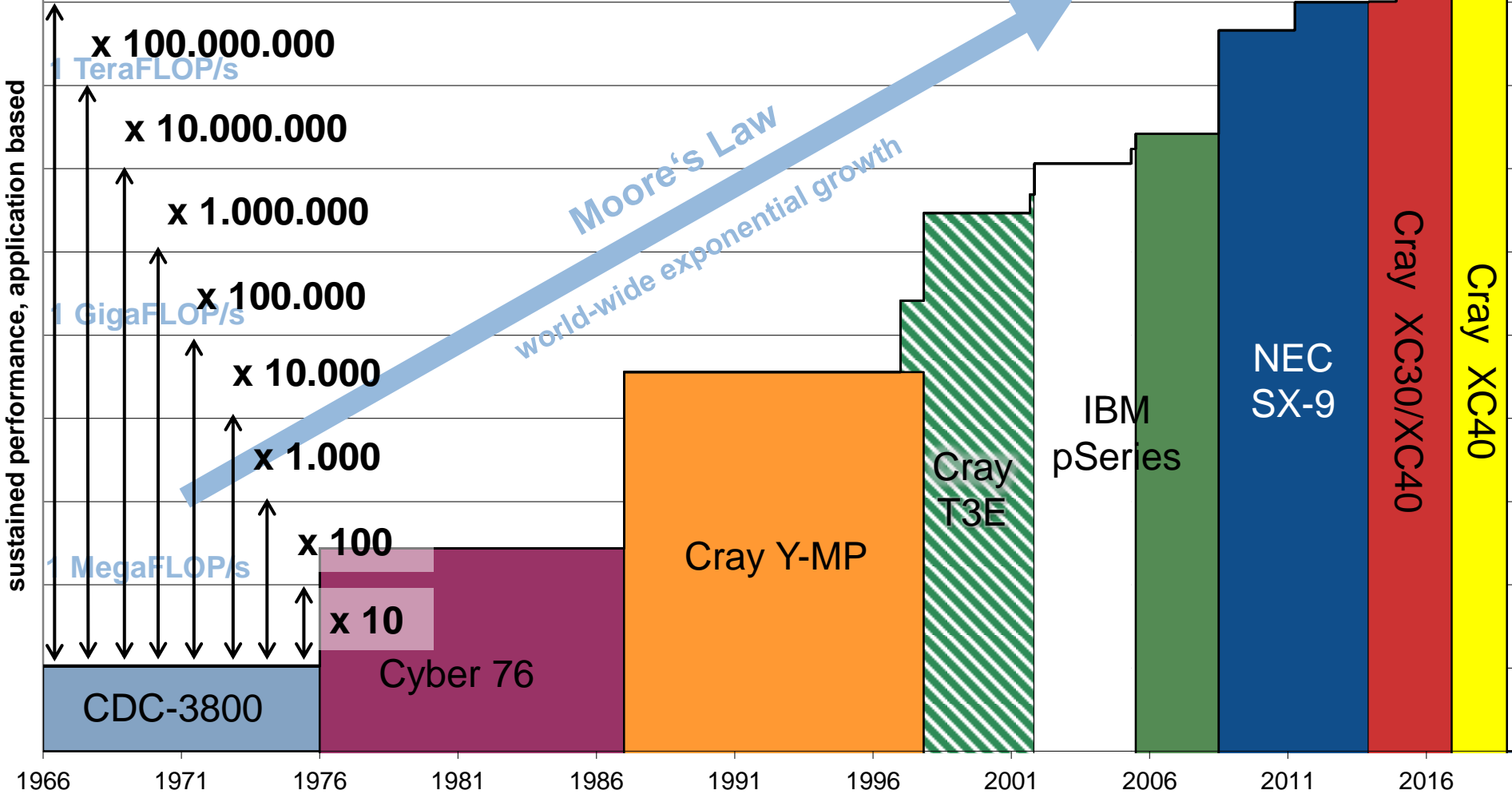


- ICON, COSMO: analysis
- ICON: analysis - serial part
- ICON, COSMO: forecast
- COSMO-DE-EPS: interpolation
- WAVE (GWAM, EWAM, CWAM)
- main run
- pre-assimilation
- assimilation
- 00..23 real time [UTC]
- 00h, 03h, .. model time [UTC]
- P, P1 parallel suite

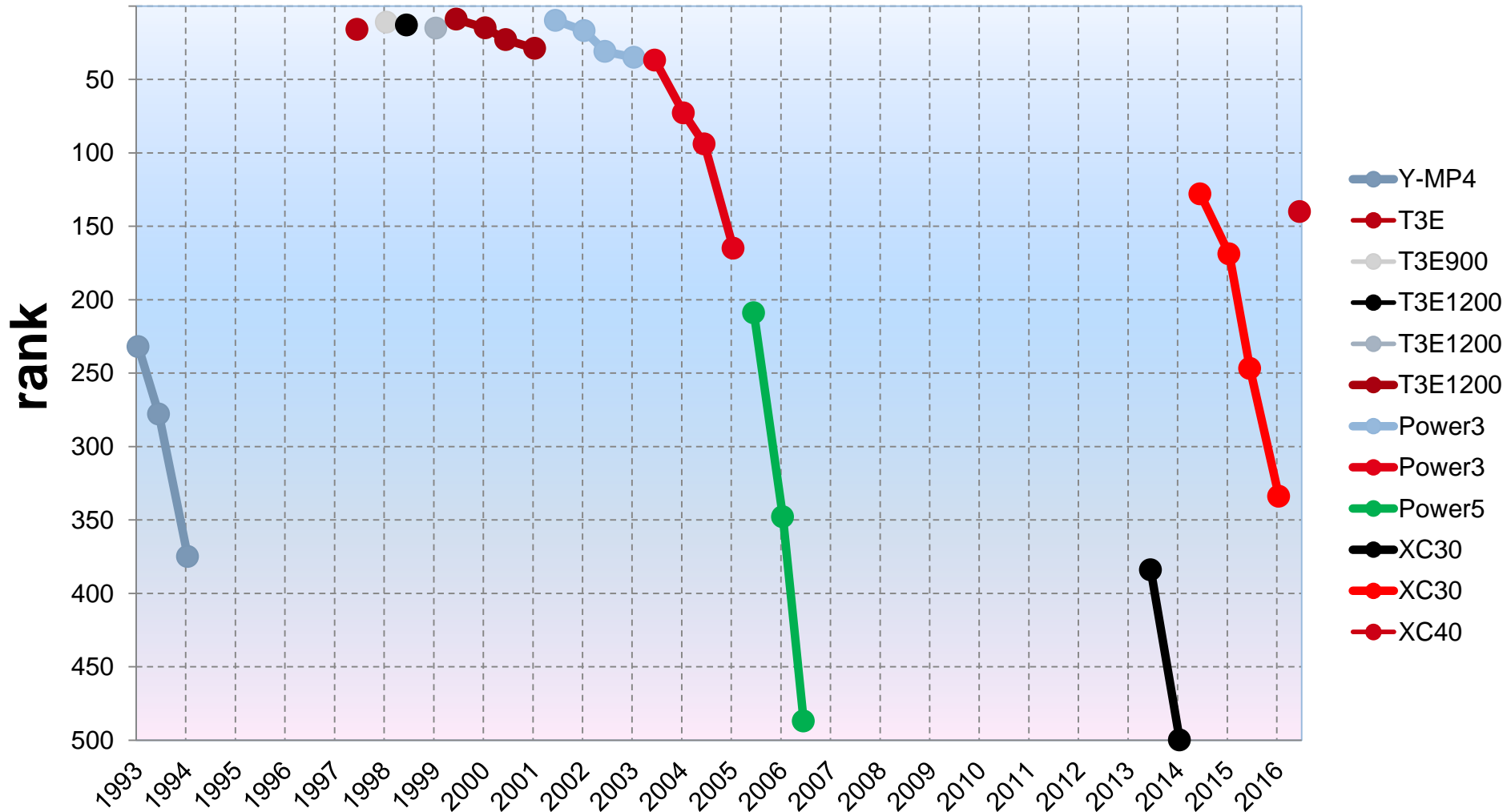




HPC performance at DWD



DWD in TOP500 list




DMRZ (Offenbach) - Computing Center

Computing hall WEST


Computing hall EAST

HPC production



Megware 34 node 768 cores 6656 GiB main memory	Cray XC40 432+544 nodes 29952 cores 122 TiB main memory	Global filesystems Cray Sonexion/ Panasas 2.8 PiB+120 TiB
--	---	---

HPC research



Megware 66 nodes 1528 cores 10752 GiB main memory	Cray XC40 432+544 nodes 29952 cores 122 TiB main memory	Global filesystems Cray Sonexion/ Panasas 5.5 PiB+170 TiB
---	---	---

Archive system Oracle/IBM-HPSS



Oracle STK SL8500
2 tape silos
20000 storing positions
60 tape drives
35 PiB (25 PiB net)



IBM X3650 M5
9 nodes
184 cores
480 GiB main memory
2 PiB disk storage

Data server production



5 x SUN X5-4
5 x 72 cores
5 x 1 TiB main memory
3.3 PiB disk storage

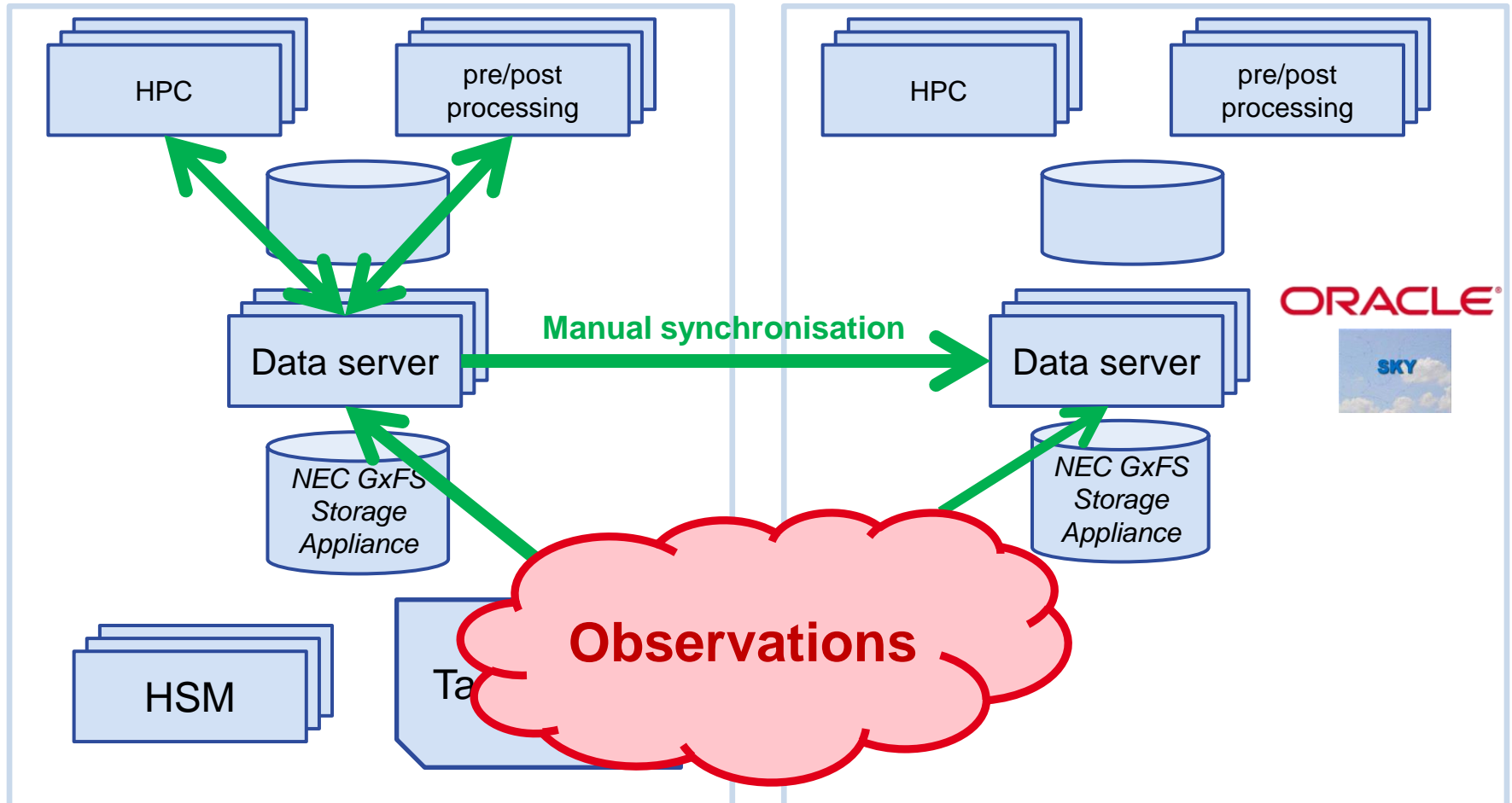
Data server research



5 x SUN X5-4
5 x 72 cores
5 x 1 TiB main memory
3.5 PiB disk storage

10 GBit

Data flow at DWD

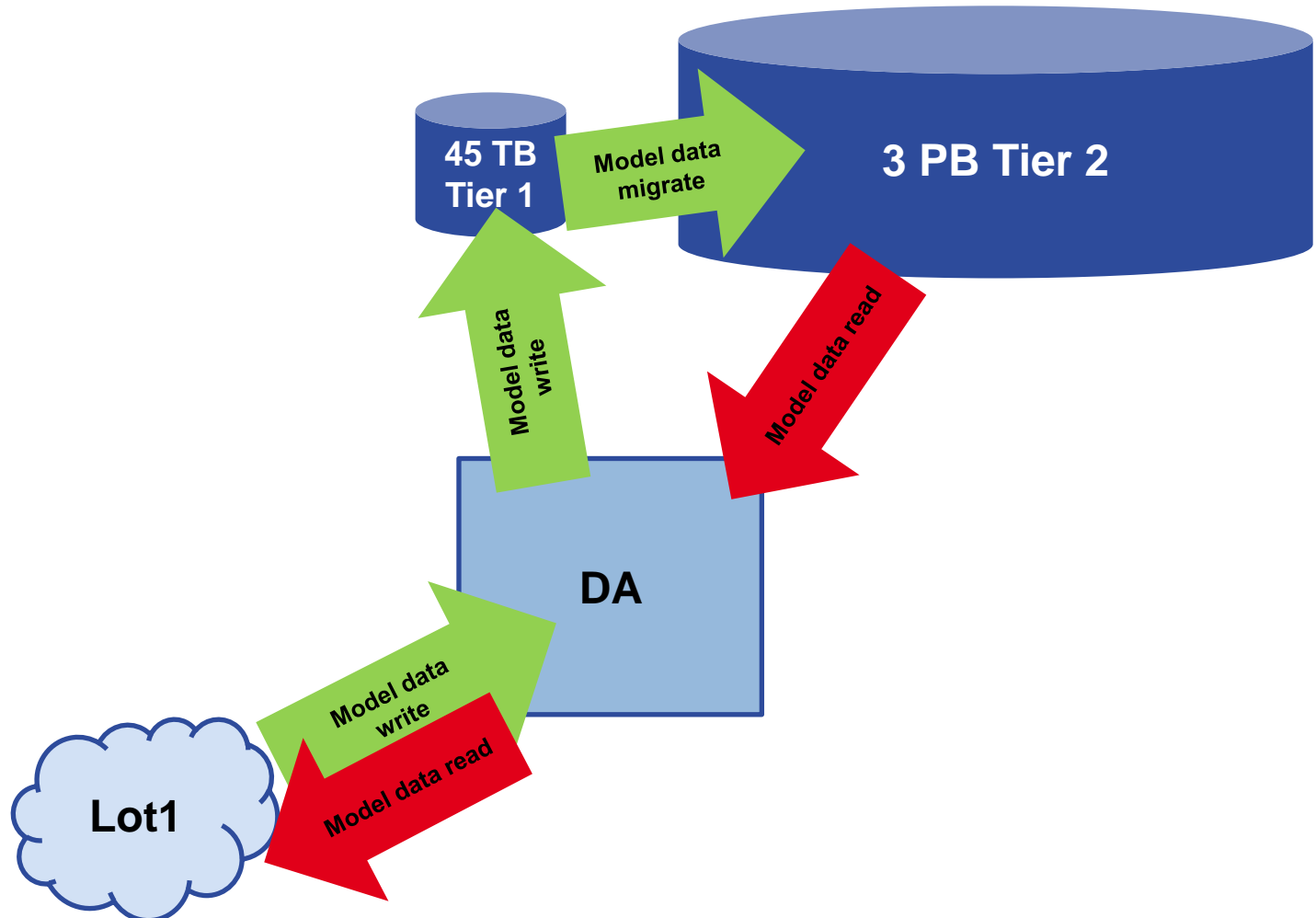


Requirements for our data servers

- **99.9% availability (99% per system)**
- **>3 PB storage capacity**
- **>36 GB/s file system performance**
- **>20 GB/s network performance**
- **Server performance benchmarks based on actual system usage (Oracle, SKY)**



Tiered Storage (data flow schematics)



NEC GxFS Storage Appliance / building block approach

- flexible approach (scalable in capacity and performance)
- internal redundancies
- powered by IBM Spectrum Scale[®]
- ~ 6–12 GB/s throughput per building block, dependent on I/O patterns, > 36 GB/s demonstrated out-of-the-box
- near 100% scaling when using multiple building blocks
- 4 clustered building blocks per hall
- redundancies on all levels to minimize downtimes



building block architecture



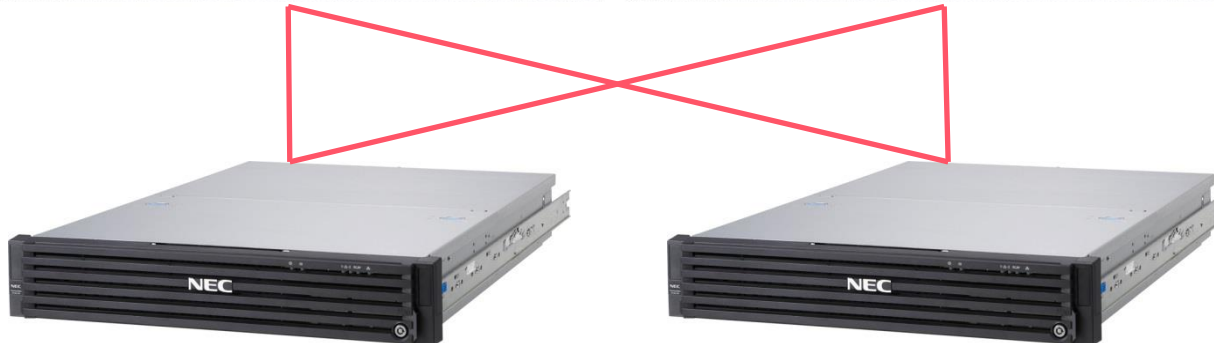
NEC SNA660-I

- redundant controllers
- redundant IB ports per controller
- 224 Gbit/s throughput / Array
- ~ 8 TB SSD capacity
- ~ 660 TB NLSAS capacity
- 2 arrays per building block



Mellanox SB7800

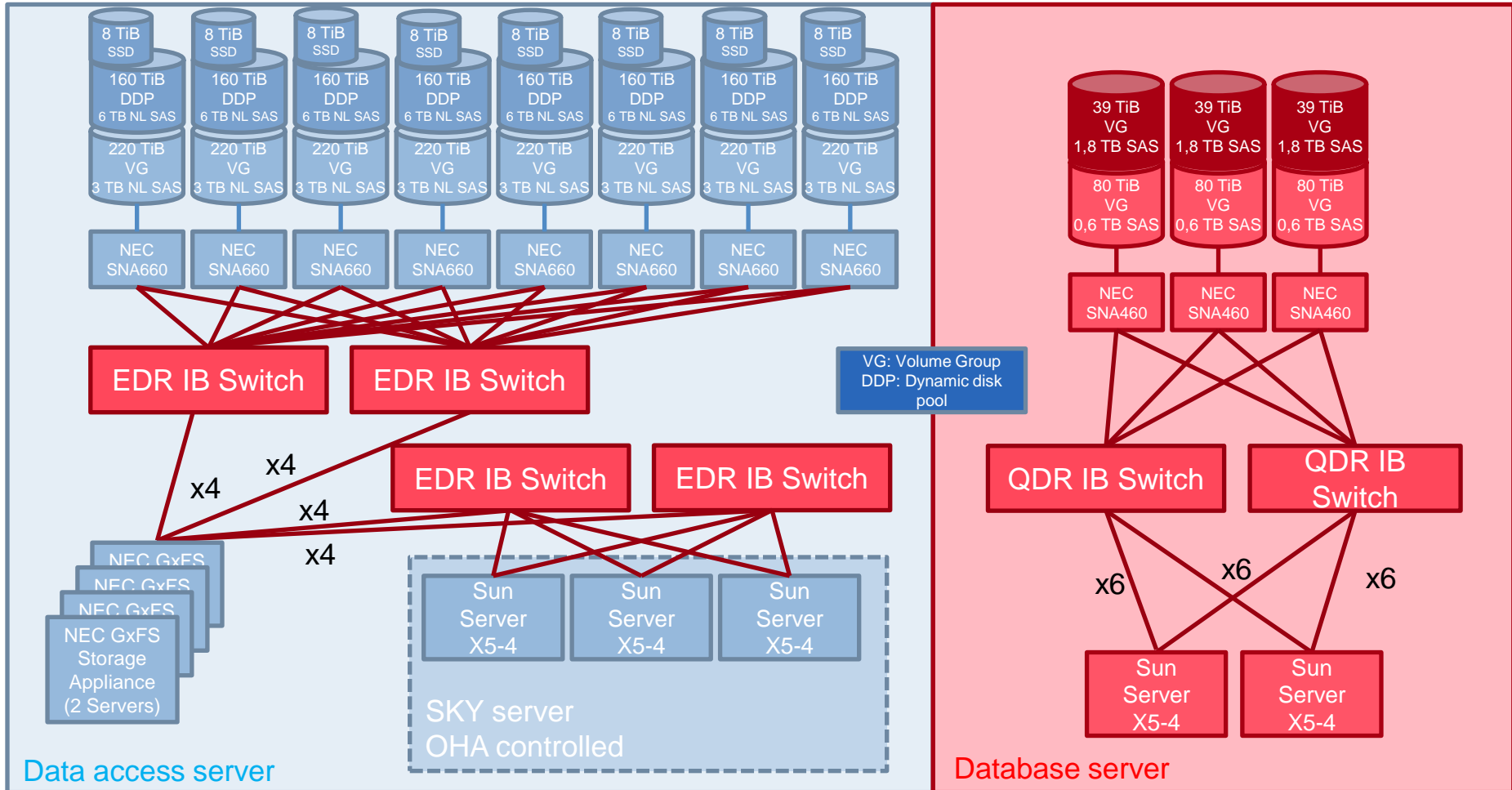
- 36 EDR ports per switch
- built-in subnet manager
- SDR/DDR/QDR/FDR/EDR support



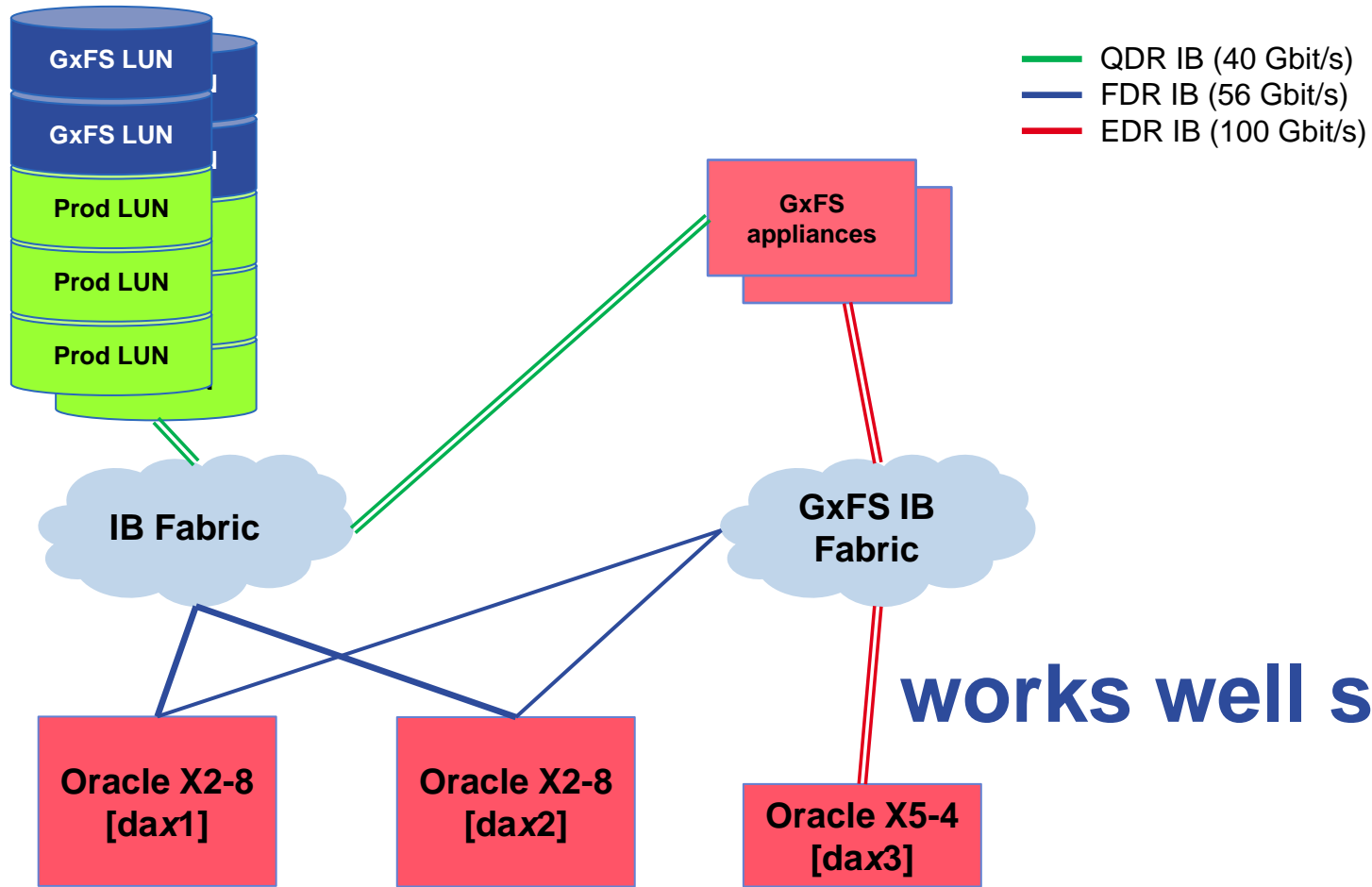
NEC Express 5800

- 2 socket Broadwell system
- 128 GB RAM
- 2x IB FDR HIC
- 2x IB EDR HIC

data server architecture



data migration (schematics)



works well so far!

Thank you!

Questions? Comments?

Manuel.Reiter@dwd.de

