

## GridKa School 2015

September 07-11

Thomas Hartmann | 2015.Sep.07

# GridKa School 2015

## Big Data Virtualization Modern Programming

## GridKa School 2015

September 07-11

Thomas Hartmann | 2015.Sep.11

# GridKa School 2015

## Big Data Virtualization Modern Programming

# GridKa School 2015

Big Data Virtualization Modern Programming

## GridKa School 2015 in numbers

- 149 participants
- 17 plenary talks
- 21 workshops
- 450 VMs on 800 cores  
(+ ??? GPUs)
- 1.5TB memory
- 10TB storage
- 240l+ coffee/tea  
+2.5kg of coffee beans
- 108l+ water
- 48kg fruits
- 15kg cookies
- ?? soft drinks, tarte  
flambes,...
- ~ 1000 cups

# GridKa School 2015

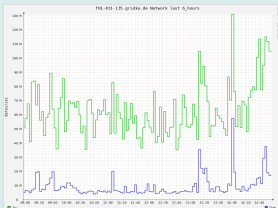
## Big Data Virtualization Modern Programming

### GridKa School 2015 in numbers

- 149 participants
- 17 plenary talks
- 21 workshops
- 450 VMs on 800 cores  
(+ ??? GPUs)
- 1.5TB memory
- 10TB storage
- 240l+ coffee/tea  
+2.5kg of coffee beans
- 108l+ water
- 48kg fruits
- 15kg cookies
- ?? soft drinks, tarte  
flambes,...
- ~ 1000 cups  
**>2/3 coffee cups...**

## School cloud

you made good use of the school HVs and VMs



### All Hypervisors

#### Hypervisor Summary



VCPU Usage  
Used 896 of 800



Memory Usage  
Used 1.5TB of 1.5TB

### Overview

#### Limit Summary



Instances  
Used 419 of 450



VCPUs  
Used 889 of 900



RAM  
Used 1.4TB of 2.1TB



Floating IPs  
Allocated 424 of 1,000

# Many talks and workshops

105.5h of workshops, 12.75h of plenary talks

### Discover data

Kibana is a visualization engine for Elasticsearch. The screenshot shows a dashboard with several bar charts and a table of data, illustrating its capabilities for data discovery and visualization.

### What is a container?

A photograph of several stacked shipping containers in various colors (yellow, red, blue, white), used as a metaphor for containers in computing.

### Modern CPU Evolution

- Moore's law continues for now!
- So transistor density doubles approximately every 18 months.
- This used to mean that computers were about x2 faster every 2 years.
- But not anymore - hardly any increases now in clock speed.
- So little increase in single threaded performance.

The graph shows two trends: a blue line for transistor density (which continues to rise) and a red line for clock speed (which has flattened out). A note at the bottom says "With some signs of slowing, however!"

### IPv6 issues for security/network teams

- Control IPv6 if not using it
- Use Dual stack and avoid use of tunnels wherever possible
- Drop packets containing the flag 0 and unknown option headers
- Deny packets that do not follow rules for extension headers
- Filter IPv6 packets that enter and leave your network
- Restrict who can send messages to multicast group addresses
- Create an Address management plan
- Create a Security Policy for IPv6 (same as IPv4)
- Block unnecessary ICMPv6
- Protect against DoS, DoS, ND and DHCP attacks
  - Use SEND and authenticated DHCPv6
  - NDPoN and NDPIXD on critical segments
- Check/modify all security monitoring, logging and parsing tools

### Georedundancy

The diagram illustrates georedundancy by showing a central data center connected to several other data centers located in different geographic regions, ensuring data availability and disaster recovery.

### Huge Data Sets in Science

The diagram shows a data pipeline starting from sensors, passing through a data bus, to a data store, and then to various analysis and visualization tools. A Venn diagram on the right highlights the intersection of Big Data, Cloud, and Analytics.

### DLC use-cases: the story of a file

The diagram shows a timeline with key milestones: Created, Accept/Reject Deadline, Main analysis complete, Public embargo ends, Anticipated end of interest, and End of life. Below the timeline, specific actions like 'Change Sub' and 'Divide data' are marked.

### DWD's weather models

The map shows three weather model domains: COSMO-DE (2.2 km), ICON-EU (6.5 km), and ICON (13 km). Each domain is associated with specific parameters like grid spacing, vertical levels, forecast range, and run time.

Model	Grid spacing	Vertical levels	Forecast range	Run time per day
COSMO-DE (EPE)	2.2 km	26	17 hours	20
ICON-EU	6.5 km	26	17 hours	20
ICON	13 km	26	17 hours	20

Daily deterministic output: ~2.5 TByte  
Daily probabilistic output: ~3.5 TByte

### Computing: The FORTRAN heritage...

A photograph of a computer terminal displaying a FORTRAN program listing, highlighting the historical significance of the language in scientific computing.

# Many talks and workshops

118.25h in total: ~5d of school



**GridKa School 2015**  
Big Data Virtualization Modern Programming

hope you had fun



**GridKa School 2015**  
Big Data Virtualization Modern Programming



# Many thanks to YOU for GKS2015



**GridKa School 2015**  
Big Data Virtualization Modern Programming

We had fun with YOU

Have a nice trip home

# GridKa School 2015

Big Data Virtualization Modern Programming

GridKa School 2015  
Big Data Virtualization Modern Programming

We had fun with YOU

Have a nice trip home  
and see you next year

# GridKa School 2016

Big Data Virtualization Modern Programming

GridKa School 2015  
Big Data Virtualization Modern Programming