Vránics Dávid Ferenc, Lovas Róbert, Kardos Péter, Bottyán Zsolt, Palik Mátyás

WRF[1] BENCHMARK MEASUREMENTS AND COST COMPARISON. VIRTUALIZED ENVIRONMENT VERSUS PHYSICAL HARDWARE

The authors performed Weather Research and Forecasting model benchmark measurements on a wide variety of computer platforms while keeping track of the associated costs. The test executions took place in cloud environments, on dedicated, physical servers and personal computers for reference. The unified measurement framework and the use of software container technology ensure the comparability of the results. The derived secondary data supports the planning of resources for the research project, and makes it possible to predict the required computing performance for later tasks during the research progress. The article details the setup and results of the measurements, while explaining the used technology and model. The results show that for smaller scale applications, cloud computing provides a less costly alternative to physical servers, while on a larger scale, usage of a dedicated physical server is advised.

Keywords: *WRF, benchmark, cost, cloud, container, hardware, comparison*

THE GOALS OF THE MEASUREMENT AND COMPARISON

„The Weather Research and Forecasting (WRF) Model is a next-generation mesoscale numerical weather prediction system designed for both atmospheric research and operational forecasting needs." [1]

Our research subproject, codenamed "UAS_ENVIRON" under project "Increasing and integrating the interdisciplinary scientific potential relating to aviation safety into the international research network at the National University of Public Service - VOLARE" aims at providing a safe and reliable framework for flight support and control systems in case of unmanned aerial flight. One of the focal areas is the meteorological prediction of flight conditions and collection of weather data.

The precedents for this research include setting up a meteorological support system [2] and database [3] for UAVs[2]. Later on, a prototype setup of WRF and weather data collecting UAVs was successfully used for sounding the planetary boundary layer [4].

The current trends of computing technology indicate that cloud, virtualization together with container technologies are going to be the next wave of innovation at several application areas. Cloud providers offer the same performance at an ever cheaper price, while increasing the available rentable capacity. They usually even provide free trial for a limited time period, while renting and configuring a virtual server takes only a few clicks, then the server is ready to boot in even a couple minutes.

Our hypothesis is that there is a point, until a well-scaling distributed application – such as a WRF instance – is cheaper to run in cloud environment, than obtaining, configuring and maintaining a physical server of similar configuration. For the cost estimations, we assume a 3-year

[1] Weather Research and Forecasting
[2] Unmanned Aerial Vehicles

computer system lifetime, and the same contract period for clouds as for physical servers, since it is obviously not feasible to buy and configure the hardware just for a one hour measurement.

The performance data presented in this article is measured on the actual systems described in the later chapters.

WRF BENCHMARK METRICS

The output of the benchmark script lists the most important metrics measured, for example:

```
items:      149
max:        26.893060
min:        2.911170
sum:        495.158710
mean:       3.323213
mean/max:   0.123571
```

The *item* count is the number of time steps processed. *Max* and *min* represent the maximum and minimum time in seconds taken by processing a time step, while *sum* represents the total processing time of all items. The average time per time step is represented by the *mean* value. This is the sum divided by the item number.

Additionally, the average GigaFLOPS value[3] can be determined by dividing the total operation count value of the benchmark – defined in billion floating point operations – by the mean value.

Simulation speed is the ratio of the model time step to the measured average time per time step.

The most significant metrics are the mean and the sum. As the item count is constant through our measurements, we will use the sum value for representation of performance [5].

MEASUREMENT SETUP

Docker concept

Docker is the world's leading open-source software container platform [6]. It simplifies software dependency handling, and ensures portability between different hardware, platforms, operating systems and architectures while supporting secure and agile deployment of new features.

For the purposes of benchmark measurement and the follow-up result comparison, the most important factor is portability, which simplifies setting up the environment on a wide variety of host machines in physical and cloud environments.

Docker images encapsulate environment settings and implement software dependencies (e.g. binaries and libraries) through inheriting other images. Figure 1 presents a comparison between traditional operating system virtualization and Docker software container technology. Docker also provides a simple command line interface to manage, download (pull) and create new images by Docker engine but further sophisticated tools are also available for complex, workflow-oriented and

[3] Billion floating point operations per second

orchestrated usage scenarios, such as the Occopus cloud and container orchestrator tool [7].

A related work on performance measurement compares high performance computing resources in cloud and physical environment, with and without utilizing the Docker software container technology [8]. The results show that the performance loss caused by the utilization of Docker is 5-10%, negligible compared to the 10-15× improvement in deployment time. The comparison shows that the expected performance of cloud resources is slightly lower than the performance of physical systems.

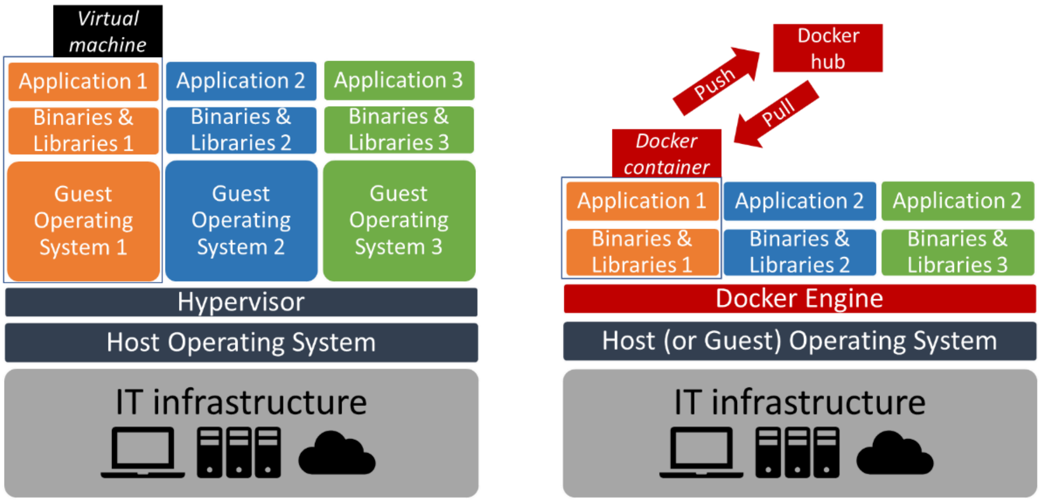

Figure 1. Comparison of traditional operating system virtualization with Docker software container technology including Docker hub for publishing and storing images (figure is the authors' own work)

Actual Docker image setup

Our Docker image contains WRF version 3.7.1 (August 14, 2015), compiled with gcc version 5.3.1 20160413 (Ubuntu 5.3.1-14ubuntu2.1), based on operating system Ubuntu 16.04 LTS. The prepared image is available on the official Docker hub as *andrewid/wrf_benchmark*.

WRF setup

A benchmark setup is used to measure and compare performance of systems based on a common indicator. To ensure comparability, the same WRF input and parameters are used. The indicator is usually derived from the execution time of the benchmark.

The WRF setup for this benchmark consists of a 48-hour forecast time, 12 km horizontal resolution on a 425 by 300 grid with 35 vertical levels case over the Continental U.S. (CONUS) domain on October 24, 2001 with 72 seconds model time step. The time period for the actual benchmark measurement is 3 hours, starting from October 25, 2001 00Z. This input data is available online. [9]

The actual item count is 150 in the benchmark, but the first one is discarded because it contains initialization and input/output operations [5].

The measured operation count for this benchmark is 30.1 billion floating point operations.

Machine setup

Docker ensures the portability between hosts. WRF version, compiler and its version is identical through the benchmarked machines.

A system is a specific instance of a platform. For example, based on Windows platform, different systems can be set up which differ in operating system versions, available number of cores or memory.

The most notable parameters of a system are the following:

1. name of system (product name, hostname, institution);
2. operating system and version;
3. processor: manufacturer, type and speed; include cache sizes if known;
4. cores per socket and sockets per node;
5. main memory per core;
6. interconnect: type (e.g. Infiniband, Gigabit Ethernet), product name, and network topology (if known) [5].

During the measurements, the following computer systems were examined:

Name of system	OS and version	Processor	Cores	Main memory	Other relevant information	Price in EUR/hrs
Google Cloud	CentOs 7.3	vCPU (VM instance)	8*, 16, 18*, 20*, 22*, 24* cores	32 GB	5 measurements and pricing on 16 CPU, only 1 on 8, 18, 20, 22, 24	0.472
MTA Cloud (SZTAKI & Wigner)		vCPU / Intel(R) Xeon(R) CPU E5-2640 v3 @ 2.60 GHz	2, 4, 8 cores	8 GB	m1.xlarge, KVM, currently free, pricing to be determined	0.000
Microsoft Azure F4S*	CentOs 7.3	vCPU (VM instance)	4 cores	2 GB/core	F4S type VM, local SSD	0.210
Microsoft Azure DS3-V2	CentOs 7.3	vCPU (VM instance)	4 cores	3.5 GB/core	DS3_V2 type VM, local SSD	0.310
Scaleway bare metal*		Intel(R) Atom(TM) CPU C2550 @ 2.40 GHz	4 cores (dedicated)	2 GB/core	C2S (only one measurement, no significant difference from VM)	0.024
Scaleway virtual machine		vCPU / Intel(R) Atom(TM) CPU C2750 @ 2.40 GHz	4 cores	1 GB/core	VC1M type VM	0.012
Dell Latitude E6540	Ubuntu 14.04.5	Intel(R) Core(TM) i7-4600M CPU @ 2.90 GHz, 4096 KB L3 cache, HT	2 core/1 socket (4 core with HT)	4 GB/core DDR3	2000 EUR price with 3 years factory warranty as expected lifetime => 0.076 EUR/hrs; 0,14362 kWh adapter consumption, ~40 HUF/kWh = 0.130 EUR/kWh => 0.019 EUR/hrs power; server room, networking,	0.095

Name of system	OS and version	Processor	Cores	Main memory	Other relevant information	Price in EUR/hrs
					maintenance not included, it is an off-the-shelf laptop	
Meteor24*		Intel Xeon E5645 (HT enabled) @ 2.40 GHz	6 core/2 socket (12 core with HT/socket = 24 core)			No estimate
Home PC*		Intel i7-4500U (HT enabled) @ 1.80 GHz	2 core/1 socket (4 core with HT)			No estimate
Cloud.hu X5670*		Intel Xeon E5670 @ 2.93 GHz	16 cores		52 HUF/hrs	0.168
Cloud.hu X5650*		Intel Xeon E5650 @ 2.67 GHz	8, 16 cores		35 HUF/hrs, 52 HUF/hrs	0.168
Server with 4xE7-4870*		Intel Xeon E7-4870 @ 2.4 GHz (HT enabled)	10 cores/4 socket (80 cores with HT) max; 20, 40, 44 tested			No estimate
RackForest 2xE5-2620v4*		Intel Xeon E5-2620v4 @ 2.1 GHz (HT enabled)	8, 16 cores (16, 32 cores with HT)	16 GB	61,595 HUF/mon, 730 hrs/mon, 309 HUF=1 EUR	0.273
RackForest 1xE3-1230v5*		Intel Xeon E3-1230v5 @ 3.40 GHz (HT enabled)	4 cores (8 cores with HT)	8 GB	33,655 HUF/mon, 730 hrs/mon, 309 HUF=1 EUR	0.149

Table 1. Available data of tested computer systems and pricing

The first five columns of Table 1 describe the system setup, while the last two describe and estimate the Euro/hours maintenance cost on a three years' time period, if sufficient data is available.

Accuracy of measurement

In case of cloud infrastructures, the overprovisioning of resources and the multi-tenancy may cause unpredictable loads on the virtualized CPUs, network, etc. That is why the measurements have been repeated on such systems.

MEASUREMENT RESULTS

WRF Scalability

Our repeated test runs have shown, however, that the difference between the repeated measurements' results is non-significant, and the values are representing the actual system under test quite prominently. For example, in case of the 4 vCPU Scaleway machine, three consecutive results provided *16.517*, *16.530* and *16.523* as the *mean* value. Based on this experience, some measurements were not run repeatedly. These results are marked below with an asterisk (*) symbol and are only measured once.

The mean values from the repeated measurements were used in all the other cases.

System	max	min	sum	mean	mean/max
Google Cloud* (24 vCPU)	37.189	1.669	299.467	2.010	0.054
Google Cloud* (22 vCPU)	36.126	1.739	314.213	2.109	0.058
Google Cloud* (20 vCPU)	34.495	1.802	320.354	2.150	0.062
Google Cloud* (18 vCPU)	32.553	1.966	344.769	2.314	0.071
Google Cloud (16 vCPU)	37.247	2.126	386.162	2.592	0.070
Google Cloud* (8 vCPU)	40.477	3.353	590.536	3.963	0.098
Meteor24* (24 CPU)	7.474	4.900	763.965	5.127	0.686
MTA Sztaki (8 vCPU)	27.633	2.870	492.018	3.302	0.120
MTA Sztaki (4 vCPU)	34.626	5.027	902.782	6.059	0.175
MTA Sztaki (2 vCPU)	38.675	8.799	1479.297	9.928	0.257
MS Azure DS3-V2 (4 vCPU)	53.195	5.525	935.723	6.280	0.118
MS Azure F4S* (4 vCPU)	52.330	5.452	918.367	6.164	0.118
Dell Latitude E6540 4 CPU	54.299	5.746	963.313	6.465	0.119
Dell Latitude E6540 3 CPU	56.459	6.538	1131.684	7.595	0.135
Dell Latitude E6540 2 CPU	59.800	7.096	1180.622	7.924	0.133
Dell Latitude E6540 1 CPU	50.983	11.764	1906.052	12.792	0.251
Home PC* (4 CPU)	37.673	9.009	1551.132	10.410	0.276
Scaleway* (4 CPU)	67.352	15.248	2490.115	16.712	0.248
Scaleway (4 vCPU)	66.261	15.025	2461.995	16.523	0.249
Cloud.hu X5670* (16 vCPU)	16.986	3.664	667.207	4.478	0.264
Cloud.hu X5650* (16 vCPU)	38.476	4.905	841.519	5.648	0.147
Cloud.hu X5650* (8 vCPU)	33.485	7.426	1265.526	8.493	0.254
Server with 4xE7-4870* (44 core)	2.926	1.550	254.659	1.709	0.584
Server with 4xE7-4870* (40 core)	24.358	1.614	284.540	1.910	0.078
Server with 4xE7-4870* (20 core)	22.752	2.545	430.434	2.889	0.127
RackForest with 2xE5-2620v4* (32 core)	20.338	1.166	204.863	1.375	0.068
RackForest with 2xE5-2620v4* (16 core)	15.986	2.036	342.232	2.297	0.144
RackForest with 1xE3-1230v5* (8 core)	24.411	4.815	762.793	5.119	0.210

Table 2. WRF performance data

The *sum* values and *core (thread) numbers* are represented on the following chart for each system of Table 1, with the core numbers specified and values measured in Table 2.

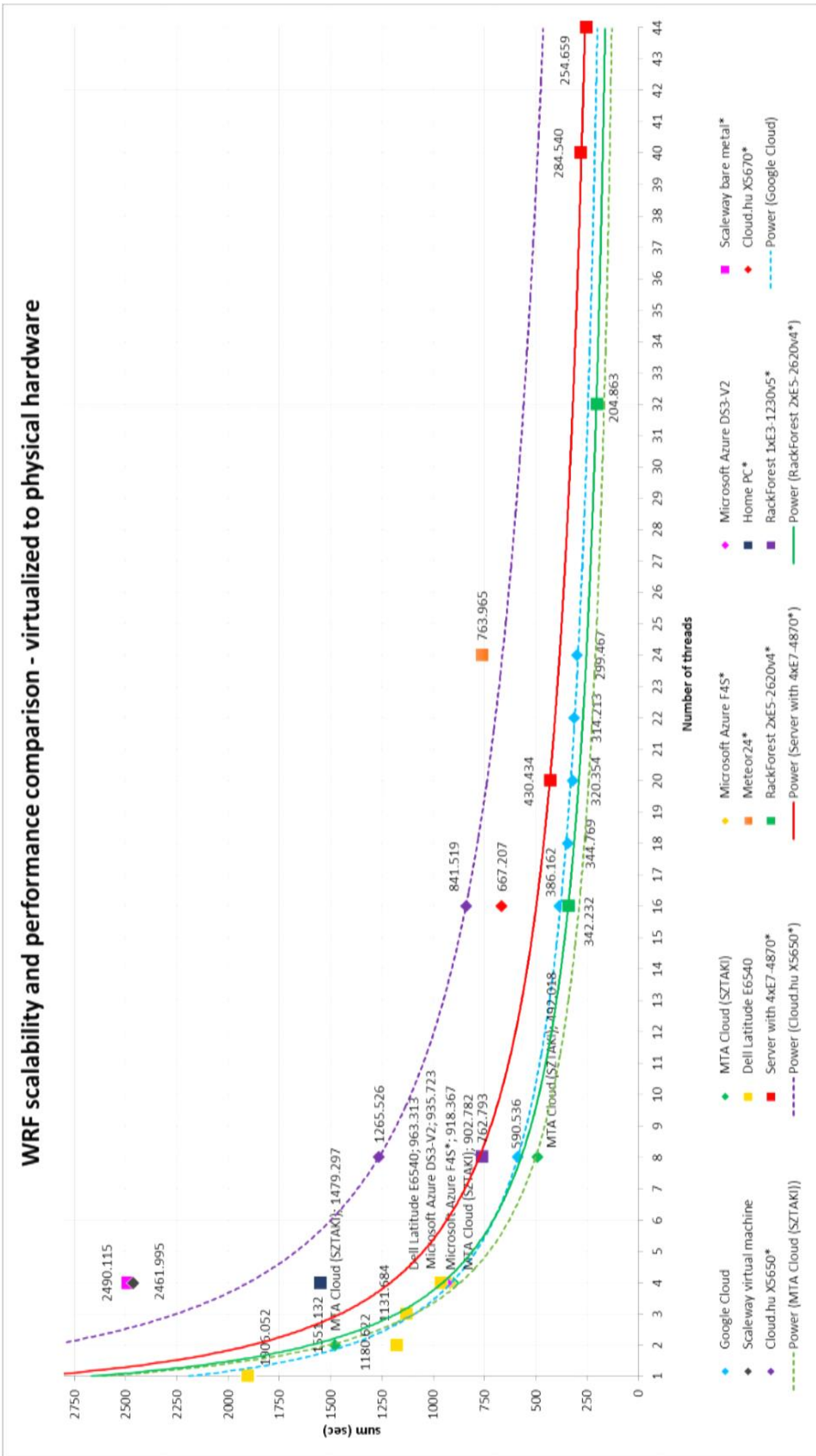

Figure 2. WRF performance data

The performance characteristics are showing a nearly hyperbolic pattern. As the diagram is representing the sum value in seconds on the *y* axis instead of simulation speed or GFLOP/second value, a hyperbolic function is interpolated onto the points instead of a logarithmic one, as they are expected to never have a coordinate with $y \leq 0.0$. That would mean the test was executed in 0.0 or less seconds. The hyperbolic function is represented with a solid line in case of physical servers, and with a dashed line in case of virtual servers.

The measured data shows some odd values that may need some explanation.

The Scaleway machines have shown a significantly lower performance than the others. The main reason for this is that these machines utilize Intel Atom CPUs, which sacrifice computing performance for better, lower energy consumption.

The measurement of the Dell notebook shows an odd curve between the 1 to 4 core values, a higher 2-core or lower 3-core value would be expected to match the expected trend. This may be WRF code specific, as 7 different runs on this same machine followed the same pattern.

Figure 2 shows the hyperbolic trend lines stretched onto the measured points.

More data and diagrams can be found on the benchmark website's results page. [10]

Physical hardware versus virtual hardware performance comparison

The results also show that virtualized services keep up with the physical competition in sense of performance and scalability.

Figure 2 displays physical server data with "■" symbols, while virtual servers are represented with "♦" symbols. The trend lines show that some cloud service providers (dashed trend lines) perform just slightly worse than physical servers (solid trend lines), some were even measured as performing better. The Scaleway bare metal versus virtual machine 4-core data shows that the virtual machine performed even slightly better than the close-physical counterpart.

The 4-core notebook and desktop PC data sits between the performance trends of two measured cloud providers, while multiple cloud providers are extremely close, just slightly faster than them in case of 4-core measurements.

Cost comparison of cloud service providers and physical hardware

In 2014, the Wigner Data Center and the Institute for Computer Science and Control (SZTAKI) initiated the MTA Cloud project together as a joint effort to establish a federated community Cloud for supporting the research activities of the further mostly non-IT specialized member institutes of the Hungarian Academy of Sciences. The recently (Q2/2016) opened OpenStack and Docker container based cloud infrastructure combines resources from Wigner and SZTAKI relying on the nationwide academic internet backbone and other federated services, e.g. eduGain and HEXXA for authentication and authorization. The total capacity of the two deployed cloud sites is 1160 virtualized CPU cores with 3.3 TB memory and 564 TB storage facility (to be extended in 2017). Currently, there is no charge for the MTA Cloud users but a special payment model is to be introduced soon. The cost comparison charts show MTA Cloud with 0.0 cost because of this. The following chart is based on the values in Table 1.

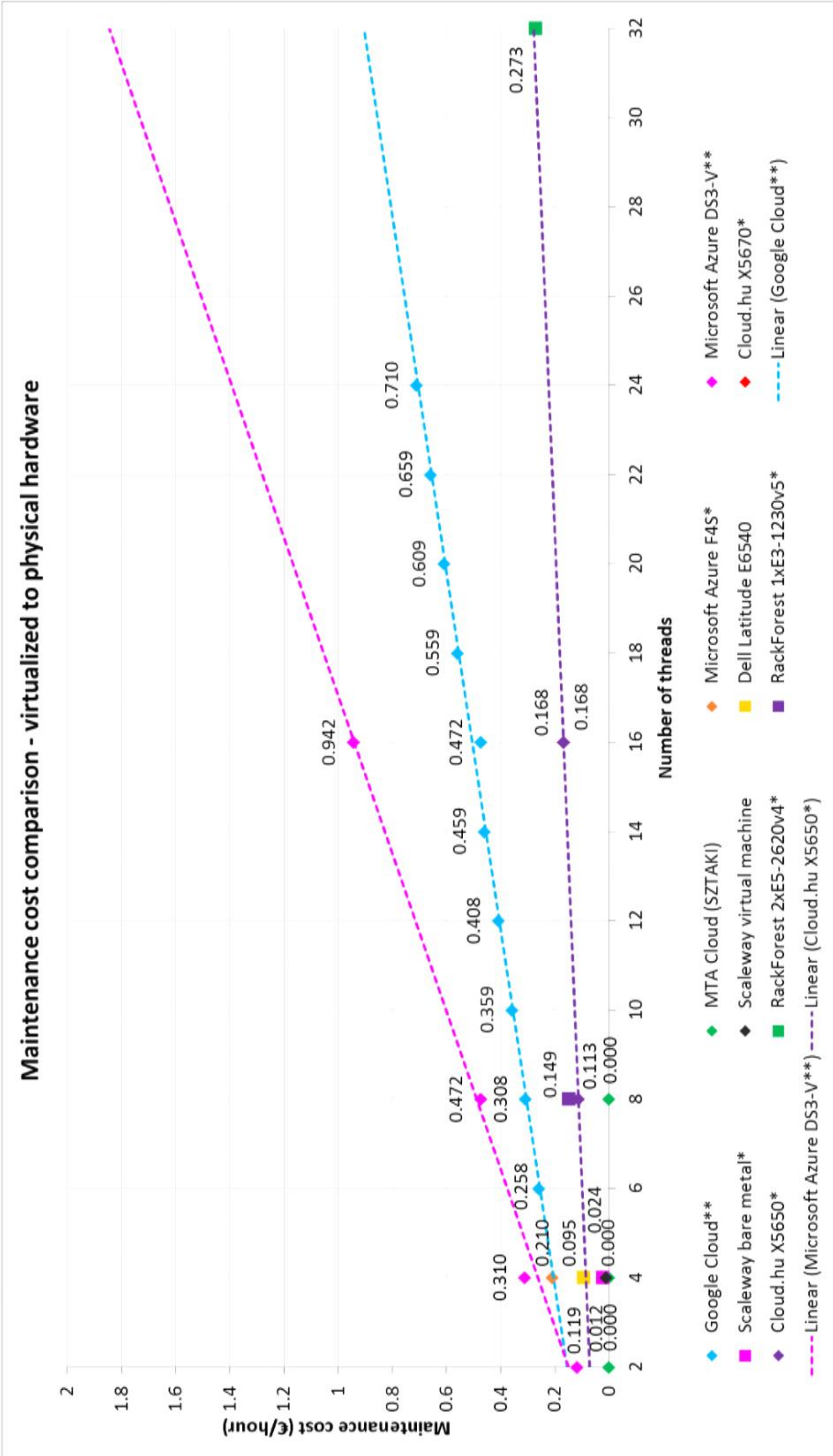

Figure 3. Maintenance cost values

For this diagram, the “*” still indicates that a host was only measured once. However, “**” means that some cost data are not the representing the actual cost experienced during the measurements, but are offered configurations/packages from the provider. These are present just to indicate the cost trend (which is linear, based on the data), and are omitted from the later results, as they do not have actual performance measurements related. The actual values are off the trend on this diagram for Microsoft Azure DS3-V, which includes extra price for storage, and for Google Cloud, where the configuration had a discount at the time compared to the online prices.

In case of cloud service providers, the maintenance cost is very straightforward to calculate, they usually charge for their services in a per-hour or per-month basis.

Physical server maintenance costs, however, are much harder to estimate, because it includes varying factors like power consumption, heating-cooling of server rooms, unexpected break-down, and operator/administrator cost. While some of these factors can be used for calculation with their maximum values, the end result will still be a rough estimation. For this reason we do not provide cost data for the most unreliably estimable cases.

Some cloud providers, like Microsoft [11], Google [12] and Amazon [13] provide detailed TCO[4] calculators partly to cope with this problem, partly to show that cloud services are cheaper as a 3-year server investment. Our experience is that these calculators are not applicable directly for several countries (including Hungary) where e.g. the cost of labor force and electricity differ significantly from the US territories. Therefore, their results are not comparable because of such applied assumptions. For this reason we combined the performance measurement results with the hourly maintenance cost to determine the outcome of our hypothesis.

Combined results

Ultimately, based on the measured data, it is possible to calculate the cost (in Euro) of a computational unit (in TFLOP), using the following formula:

$$C_P = \frac{C_M * t_{sum}}{O} \tag{1}$$

Where C_P is performance cost, C_M is maintenance cost, t_{sum} is the measured total execution time, and O is the total floating point operation count.

If we multiply maintenance cost (which we have in Euro/hour, so we divided it by 3600 to bring it to Euro/seconds), with the measured sum value (which we indicated in seconds) then we get the actual cost of the benchmark run in Euro.

As noted on the WRF benchmark homepage [5], the measured operation count for this benchmark is 30.1 GFLOP. If we divide the calculated Euro cost for a benchmark run by this value, we will get the cost for a GFLOP in case of WRF in Euro/GFLOP.

These values are then converted to Euro/TFLOP by multiplying them with 1000 for the sake of human-readability.

[4]Total Cost of Ownership

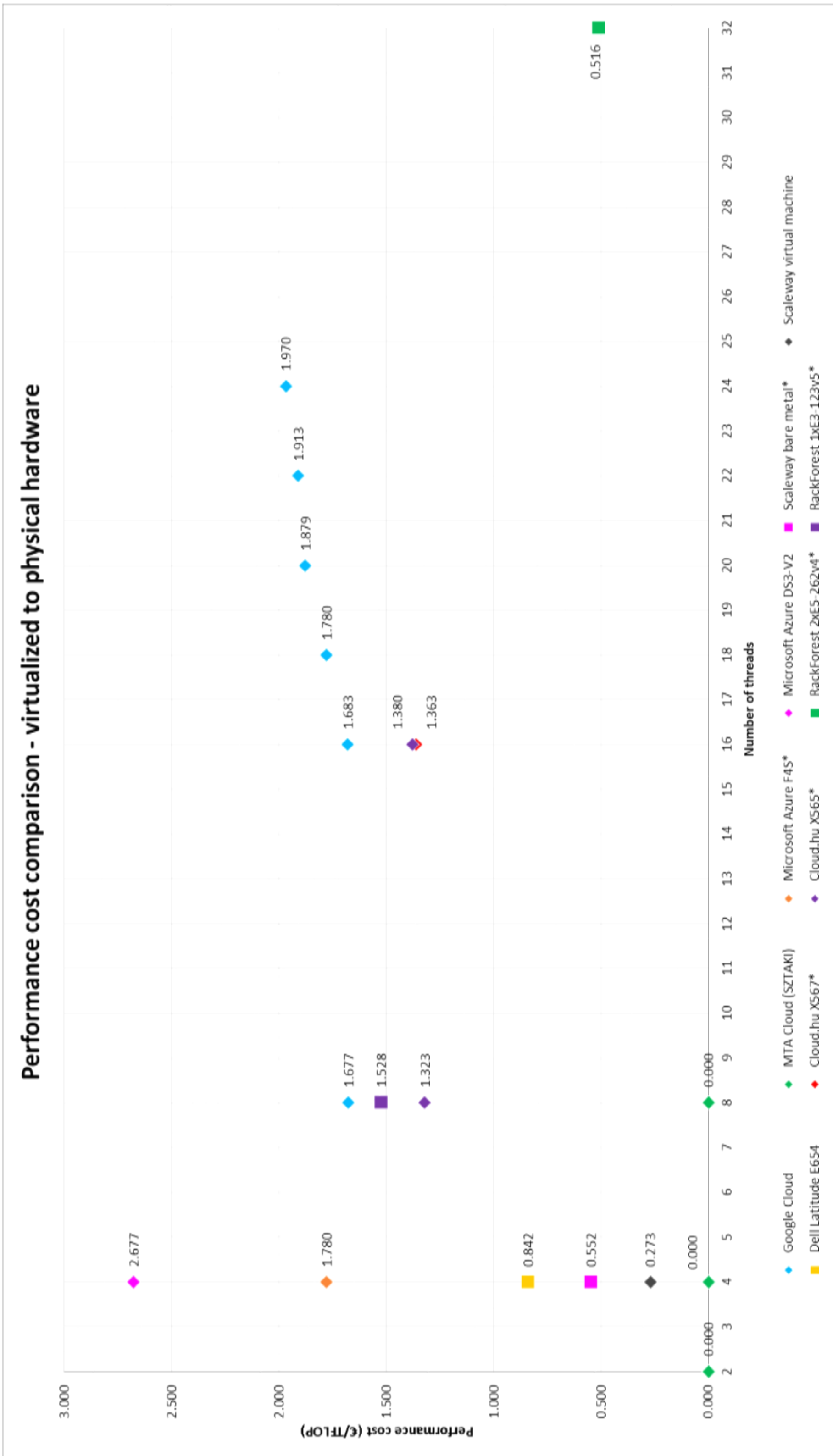

Figure 4. Performance cost values

Figure 4 visualizes the values from Table 3. For each thread number, the lowest value is the cheapest; meaning it costs less to run the same WRF model with the same parameters on a computer system that is closer to 0.0 on the *y* axis than the ones above it.

Note again, that MTA Cloud does not have a comprehensive and final payment model yet, therefore its cost is still 0.0 on the diagram.

System / number of threads	2	4	8	16	18	20	22	24	32
Google Cloud			1.677	1.683	1.780	1.879	1.913	1.970	
MTA Cloud (SZTAKI)	0.000	0.000	0.000						
Microsoft Azure F4S*		1.780							
Microsoft Azure DS3-V2		2.677							
Scaleway bare metal*		0.552							
Scaleway virtual machine		0.273							
Dell Latitude E654		0.842							
Cloud.hu X567*				1.363					
Cloud.hu X565*			1.323	1.380					
RackForest 2xE5-262v4*									0.516
RackForest 1xE3-123v5*			1.528						

Table 3. Performance and cost combined, €/TFLOP

An interesting finding is that the Scaleway performance was the worst measured, but still because of the extremely low pricing it is the most cost effective system to run WRF instances on 4 threads. This may be possible because of the relatively low power consumption of the Intel Atom processors, and the aggressive pricing strategy of Scaleway.

The Microsoft Azure solutions however prove to be the costliest, most probably because of their grade of additional services and built-in (but actually not used) support costs.

In between these values sits the business-grade Dell notebook, its estimated cost only contains the one-time hardware price (3-year warranty included) and the maximum power consumption. Other costs are excluded from the estimation, as it is an off-the-shelf notebook.

8-thread values show that the 8-core RackForest physical server is between Google Cloud and Cloud.hu performance cost.

Two 16-core configurations on Cloud.hu are very close to each other, while Google Cloud is costlier. It is also showing an increasing trend in performance cost with more cores.

The 32-cores RackForest data shows that with so many parallel threads it is still expected to be less expensive to rent an actual physical server than to contract a cloud provider for a virtual machine with a similar configuration.

Related works

Grid computing can be considered as a predecessor of cloud computing from several aspects, and WRF modelling has been benchmarked on Grid computing platforms, including the German D-GRID infrastructure, before the rising of cloud computing and container based platforms in the area of high-performance applications. [14]

Later some other widespread cloud computing platform have been investigated, including Amazon, but these studies focused particularly on multi-VM and MPI executions of WRF. [15][16]

Docker container and partly the Amazon (EC2) based execution of WRF models have been already investigated by NCAR in order to avoid software dependencies, to improve education and research activities, and also to allow the reproducibility of simulations. [17]

However, the related works did not provide detailed benchmark results focusing on cost factors on various (mostly cloud based) platforms, and they covered only the most prominent Grid and cloud providers. Our studies attempted to overcome on these limitations, and involved computational resources e.g. from different European cloud providers (such as Scaleway, Cloud.hu, and MTA Cloud) and cost analysis as well.

Beside this project, more than 20 research teams have started utilizing the MTA Cloud in 2016 with no or little experiences with advanced cloud usage scenarios such as multi-VM deployment. The presented Docker based WRF simulation together with its benchmark serves as a valuable use case for the further development of SZTAKI's Occopus cloud and container orchestrator tool [7] as a part of MTA Cloud.

CONCLUSION

The measurements were successfully executed and evaluated on multiple hosts, making it possible to compare and publish the results with precisely estimated cost in most cases.

Our hypothesis stands: for less threaded or short, occasional measurements the cloud service providers usually offer the same WRF performance at lower costs, while for higher scaled scenarios, physical servers are the less costly option if we assume continuous, and long-term load on them. For example, in case of 4 threaded measurements, the cost of performance for a laptop is around four times more expensive compared to the service of the cheapest commercial cloud provider. Meanwhile, based on the trends in case of 32 core measurements, the cost of performance for a physical server is expected to be 4-5× less costly compared to the cloud providers.

Still, the actual point where we can say that the cost advantage turns from virtualized to physical hardware would be very hard to determine. This is due to the varying factors during the measurements and the limited or missing cost data for some hosts.

The Docker setup is already reused during our latest research with different WRF cases, the results and cost estimation may also be interesting to other meteorological research projects that are using applications similar to WRF for modeling weather.

REFERENCES

[1] The Weather Research & Forecasting Model, http://www.wrf-model.org/index.php (22.03.2017.)

[2] Hadobács K., Vidnyánszky Z., Bottyán Zs., Wantuch F., Tuba Z.: A pilóta nélküli légijárművek meteorológiai támogató rendszerének kialakítása és alkalmazhatóságának bemutatása esettanulmányokon keresztül. *Repüléstudományi Közlemények*, XXV 2 (2013), 405–421.

[3] Bottyán Zs., Wantuch F., Tuba Z., Hadobács K., Jámbor K.: Repülésmeteorológiai klíma adatbázis kialakítása az UAV–k komplex meteorológiai támogató rendszeréhez. *Repüléstudományi Közlemények*, XXIV 3 (2012), 11–28.

[4] Szabó Z., Istenes Z., Gyöngyösi Z., Bottyán Zs., Weidinger T., Balczó M.: A planetáris határréteg szondázása pilótanélküli repülő eszközzel (UAV). *Repüléstudományi Közlemények*, XXV 2 (2013), 422–434.

[5] WRF V3 Parallel Benchmark Page, http://www2.mmm.ucar.edu/wrf/WG2/benchv3/ (22.03.2017.)

[6] Mouat, A.: *Using Docker: Developing and Deploying Software with Containers*. Sebastopol: O'Reilly Media, 2015.

[7] Kacsuk P., Kecskeméti G., Kertész A., Németh Zs., Kovács J., Farkas Z.: Infrastructure Aware Scientific Workflows and Infrastructure Aware Workflow Managers. *Science Gateways. Journal of Grid Computing,* 14 (2016), 641.

[8] Muhtaroglu, N., Kolcu, B., Ari, I.: Testing Performance of Application Containers in the Cloud with HPC Loads, in Iványi, P., Topping, B. H. V., Várady, G. (Editors), *Proceedings of the Fifth International Conference on Parallel, Distributed, Grid and Cloud Computing for Engineering*. Civil-Comp Press, Stirlingshire, UK, Paper 17, 2017. doi:10.4203/ccp.111.17

[9] conus12km_data_v3, http://www2.mmm.ucar.edu/WG2bench/conus12km_data_v3 (22.03.2017.)

[10] WRF 12 Kilometer CONUS Benchmarks, http://www2.mmm.ucar.edu/wrf/WG2/benchv3/12KM_Results_20100414percore.htm (22.03.2017.)

[11] Total Cost of Ownership (TCO) Calculator PREVIEW, https://www.tco.microsoft.com/ (22.03.2017.)

[12] TCO Pricing Calculator, https://cloud.google.com/pricing/tco/ (22.03.2017.)

[13] AWS Total Cost of Ownership (TCO) Calculator, https://awstcocalculator.com/ (22.03.2017.)

[14] Ploski, J., Scherp, G., Petroliagkis, T., Hasselbring, W.: Grid-based deployment and performance measurement of the Weather Research & Forecasting model. *Future Generation Computer Systems,* 25 3 (2009), 346–350.

[15] Fenn, M., Holmes, J., Nucciarone, J.: A Performance and Cost Analysis of the Amazon Elastic Compute Cloud (EC2) Cluster Compute Instance. White Paper, 10 Aug 2010.

[16] Duran-Limon, H. A., Flores-Contreras, J., Parlavantzas, N., Zhao, M., Meulenert-Peña, A.: Efficient execution of the WRF model and other HPC applications in the cloud. *Earth Science Informatics,* 9 (2016), 365.

[17] Hacker, J., Exby, J., Gill, D., Jimenez, I., Maltzahn, C., See, T., Mullendore, G., Fossel, K.: A containerized mesoscale model and analysis toolkit to accelerate classroom learning, collaborative research, and uncertainty quantification. *Bulletin of the American Meteorological Society*, 2016. (in press)

WRF ÖSSZEHASONLÍTÓ MÉRÉSEK A TELJESÍTMÉNY ÉS KÖLTSÉG FÜGGVÉNYÉBEN VIRTUALIZÁLT ILLETVE FIZIKAI HARDVEREK ESETÉN

A szerzők számítógépes rendszerek széles palettáján végeztek teljesítményméréseket az időjárás kutató és előrejelző (WRF) modell használatával, miközben a kapcsolódó költségeket is nyomon követték. A tesztesetek lefuttatására felhő környezetben és dedikált fizikai kiszolgálókon, illetve viszonyításként személyi számítógépeken is sor került. Egységes mérési keretrendszer és a szoftver konténer technológia alkalmazása biztosítja az eredmények összevethetőségét. A származtatott eredmények segítik a kutató projekt erőforrásainak tervezését, illetve lehetővé teszik a későbbi feladatokhoz szükséges számítási kapacitás becslését. A cikk részletezi az alkalmazott beállításokat és kapott eredményeket, miközben kitér az alkalmazott technológia és modell sajátosságaira. Az eredmények fényében azt mondhatjuk, kevesebb párhuzamos szál esetén inkább megéri felhőszolgáltatást bérelni, míg több párhuzamos szál esetén érdemes dedikált fizikai szervert fenntartani.

Kulcsszavak: *WRF, teljesítménymérés, költség, felhő, konténer, hardver, összehasonlítás*

Vránics Dávid Ferenc (MSc) doktoranduszhallgató, informatikus Nemzeti Közszolgálati Egyetem Hadtudományi és Honvédtisztképző Kar Katonai Műszaki Doktori Iskola Biztonságtechnika kutatásterület vranicsd@gmail.com orcid.org/0000-0003-0637-476X	Vránics Dávid Ferenc (MSc) PhD Student, computer scientist National University of Public Service Faculty of Military Science and Officer Training Doctoral School of Military Engineering Security Technology vranicsd@gmail.com orcid.org/0000-0003-0637-476X
Lovas Róbert (PhD) tudományos főmunkatárs Magyar Tudományos Akadémia Számítástechnikai és Automatizálási Kutatóintézet robert.lovas@sztaki.mta.hu orcid.org/0000-0001-9409-2855	Lovas Róbert (PhD) Senior research fellow Hungarian Academy of Sciences Institute for Computer Science and Control robert.lovas@sztaki.mta.hu orcid.org/0000-0001-9409-2855
Kardos Péter (MSc) részlegvezető, meteorológus Hungarocontrol Magyar Légiforgalmi Szolgálat Zrt. Repülőtéri Meteorológiai Részleg Peter.Kardos@hungarocontrol.hu orcid.org/0000-0001-8857-4102	Kardos Péter (MSc) Head of unit, meteorologist Hungarocontrol Hungarian Air Navigation ServicesLtd. Aerodrome Meteorological Unit Peter.Kardos@hungarocontrol.hu orcid.org/0000-0001-8857-4102
Bottyán Zsolt (PhD) egyetemi docens Nemzeti Közszolgálati Egyetem Hadtudományi és Honvédtisztképző Kar Katonai Repülő Intézet Repülésirányító és Repülő-hajózó Tanszék bottyan.zsolt@uni-nke.hu orcid.org/0000-0003-0729-2774	Bottyán Zsolt (PhD) Associate professor National University of Public Service Faculty of Military Science and Officer Training Institute of Military Aviation Department of Aerospace Controller and Pilot Training bottyan.zsolt@uni-nke.hu orcid.org/0000-0003-0729-2774
Palik Mátyás (PhD) intézetigazgató, egyetemi docens Nemzeti Közszolgálati Egyetem Hadtudományi és Honvédtisztképző Kar Katonai Repülő Intézet palik.matyas@uni-nke.hu orcid.org/0000-0002-2304-372X	Palik Mátyás (PhD) Director of institute, associate professor National University of Public Service Faculty of Military Science and Officer Training Institute of Military Aviation palik.matyas@uni-nke.hu orcid.org/0000-0002-2304-372X

This work was supported by the European Regional Development Fund (GINOP 2.3.2-15-2016-00007, “Increasing and integrating the interdisciplinary scientific potential relating to aviation safety into the international research network at the National University of Public Service - VOLARE”). On behalf of Project VOLARE we thank for the usage of MTA Cloud (https://cloud.mta.hu/) that significantly helped us achieving the results published in this paper. The project was realised through the assistance of the European Union, and co-financed by the European Regional Development Fund.

http://www.repulestudomany.hu/folyoirat/2017_2/2017-2-19-0400_Vranics_David_Ferenc_et_al.pdf