Review of
The ESCAPE project: Energy-efficient Scalable Algorithms for Weather Prediction at Exascale
by A. Mueller et al.

General comment:
In this manuscript an overview about the achievements in the ESCAPE project is given. The main concept is explained, some of the developments are explained in details and finally some tests are mentioned.
Although I think that this manuscript is a valuable contribution for GMD, I cannot recommend to accept the manuscript in the actual state. The manuscript must be revised in a substantial way before it can be considered again. Therefore I recommend major revision of the manuscript. In the following I will explain my concerns.

Major issues

1. Balance of the manuscript:
   The manuscript is very long and not really balanced. Some parts are explained in details, as e.g. the development of the MPDATA dwarf, but some parts are just mentioned. Especially for the very shortly explained parts, there are very often references to technical reports, i.e. documentation which is generally not peer reviewed. Although there are some performance tests, there is only one figure showing a test for atmospheric flows, and also this test is only marginally described.

   I would recommend to significantly reorganize the manuscript, maybe also considering to split the manuscript into three parts: First, an overview part, where mostly the concept and the new architecture can be explained in a concise way. Second, a model description part, i.e. a detailed description of the different parts of the model, especially of the parts, which are contained in the technical memoranda but not described in peer-reviewed literature. Third, a part dedicated to test cases for atmospheric flows - and maybe also clouds and radiation, since these parts are also included into the model.

   Especially test cases of atmospheric flows would be very interesting, since it is not clear if all the new models represent the atmospheric flow and other atmospheric phenomena in a physically consistent way. Therefore I highly recommend to use well-documented test cases for atmospheric flows, as e.g. Jablonowski & Williamson (2006). It would be interesting to see also tests for clouds and radiation, although I am not really aware of large scale tests, beyond the standard tests as e.g. Weismann & Klemp (1982).

2. Selection of the dwarfs:
   It is not really clear how and why the different dwarfs were chosen. Although I think that this is a well chosen sample of possible models, it should be justified much better. Especially, the choice of the shallow water model is not really clear, because no real results of this model are shown in the manuscript. Therefore, I recommend to describe the choice of the models is a clearer way.

Minor issues:
Cost model:
The benefit of the cost model is not really clear to me. It is introduced in a comparable length as the dwarfs, but it is not really clear why this is so important for the whole manuscript, justifying a large part in the appendix.

References

Jablonowski, C., and D. L. Williamson, 2006: A Baroclinic Instability Test Case for Atmospheric Model Dynamical Cores, Quart. J. Roy. Met. Soc., 132, 2943-2975

Weisman, M. and J. Klemp, 1982: The Dependence of Numerically Simulated Convective Storms on Vertical Wind Shear and Buoyancy. Mon. Wea. Rev., 110, 504-520.