# On the Possibilities of Multi-Core Processor Use for Real-Time Forecast of Dangerous Convective Phenomena

N.O. Raba[1], E.N. Stankova[1,2]

[1]Saint-Petersburg State University
199034 St.-Petersburg, Russia
[2]Institute for High Performance Computing and Integrated Systems,
199397 St.-Petersburg, Russia
no13@inbox.ru, lena@csa.ru

**Abstract.** We discuss the possibilities of use of the new generation of desktops for solution of one of the most important problems of weather forecasting: real-time prediction of thunderstorms, hails and rain storms. The phenomena are associated with development of intensive convection and are considered as the most dangerous weather conditions. The most perspective way of the phenomena forecast is computer modeling. Small dimensional models (1 - D and 1.5 - D) are the only available to be effectively use in local weather centers and airports for real-time forecasting. We have developed one of such models: 1.5 - D convective cloud model with the detailed description of microphysical processes and have investigated the possibilities of its parallelization on multi-core processors with the different number of cores. The results of the investigations have shown that speed up of cloud evolution calculation can reached the value of 3 if 4 parallelization threads are used.

**Keywords:** multi-core processors, parallelization, thread, numerical model, real-time weather forecast, convective cloud

## 1 Introduction

Climate and weather forecast is among the so called grand-challenge scientific problems, which need for their solution high-performance computer facilities. All the main weather centers in the world are equipped with powerful clusters and supercomputers. But one should take into account that weather forecast in not only the prediction of wind and pressure fields which are the output of the so called regional models and general circulation models of atmosphere, but also the prediction of local dangerous convective phenomena, such as thunderstorms, hails and rain storms. Forecast of rain, hail and thunderstorm is usually provided in rather small weather centers and airports which have modest financial resources and are not able to buy expensive supercomputers or even clusters. Ordinary desktops are the only computational resources that are available. The problem is even more complicated as the forecast should be real-time and it should take no more than one an hour to provide it. As a consequence experts of such local centers have to provide forecast with the help of simple methods and models. Up to now forecast of the dangerous convective phenomena in the airports of CIS (Commonwealth of Independent States) countries is provided with the help of semi-empirical methods and very simple 1-D stationary cloud models. It is evident that requirement of real-time forecast in combination with modest computational resources will not allow using elaborated 2 - D and 3 - D models in such centers. But appearance of desktops with multi-core processors open the possibility of applying elaborated 1-D cloud models with detailed description of microphysical processes. The only requirement is proper use of multi-core processor facilities by means of parallelization.

We have developed 1.5-D convective cloud model with detailed description of microphysical processes and have investigated possibilities of its effective use for real-time forecast of cloud parameters. Calculations have been provided with the help of multi-core processors of different types and different core numbers.

The so called space parallelization in conjunction with the multi-thread technology has been used. The results have shown that speed up of cloud evolution calculation can reached the value of 3 if 4 parallelization threads are used.

## 2 Model description

In the model the region of convective flow is represented by two concentric cylinders [1]. The inner cylinder (with constant radius a) corresponds to the updraft flow region (cloudy region) and the outer cylinder (with constant radius b) – to the surrounding downdraft flow region (cloudless) (Fig.1)

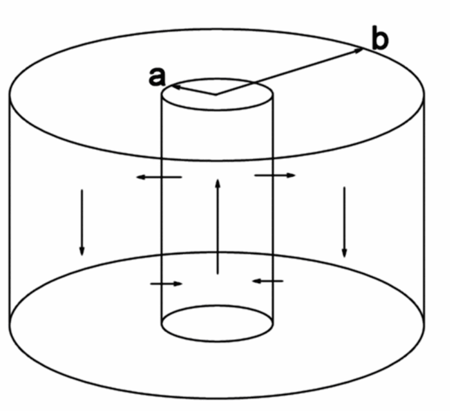

**Fig. 1**. The scheme of up and down flows

The model is 1.5-dimensional with the detailed description of warm (i.e. without the ice phase) microphysical processes. The term 1.5 – dimensional means the following: though all cloud variables are represented with mean values averaged over the horizontal cross section of the cloud, fluxes in and out of the inner cylinder borders are taken into account.

The ratio of the area of cross section of inner cylinder to the area of cross section of outer ring-shaped cylinder is equal to

$$K_{ab} = a^2 / (b^2 — a^2) \tag{1}$$

In generalized form the equations for vertical velocity, temperature and mixing ratios of water vapour and cloud droplets inside the inner (equation 2) and outer (equation 3) cylinders can be written as follows:

$$\frac{\partial \phi_{in}}{\partial t} = -w_{in} \frac{\partial \phi_{in}}{\partial z} - \frac{2\alpha^2}{a} | w_{in} - w_{out} | (\phi_{in} - \phi_{out}) + \frac{2}{a} U_a (\phi_{in} - \phi_a) + \frac{1}{\rho_{a_0}} \frac{\partial}{\partial z} K_f \frac{\partial \phi_{in}}{\partial z} + F_{\phi_{in}} - A_{\phi_{in}} + G_{\phi_{in}}, \tag{2}$$

$$\frac{\partial \phi_{out}}{\partial t} = -w_{out} \frac{\partial \phi_{out}}{\partial z} - \frac{2\alpha^2}{a} | w_{in} - w_{out} | (\phi_{out} - \phi_{in}) + \frac{2}{a} U_a (\phi_{out} - \phi_a) + \frac{1}{\rho_{a_0}} \frac{\partial}{\partial z} K_f \frac{\partial \phi_{out}}{\partial z} + F_{\phi_{out}} - A_{\phi_{out}} + G_{\phi_{out}}, \tag{3}$$

Where the variables with subscripts 'in' and 'out' relate to the values, averaged over the inner and outer cylinders consequently. $\phi$ can take the values of vertical velocity $w$, temperature $T$, mixing ration of water vapor $Q_v$ and mixing ratio of cloud droplets in the $i$-th drop-size interval $Q_{ci}$. $t$ and $z$ are independent variables (time and height consequently), $\alpha$ is the coefficient for lateral eddy mixing through the periphery of the cloud, $U_a$ is determined by the equation of mass continuity under assumption of incompressibility which is given as $\frac{2u_a}{a} + \frac{1}{\rho_{a_0}} \frac{\partial(\rho_{a_0} w_{in})}{\partial z} = 0$ , $\rho_{a_0}$ is density of the atmospheric air, $K_f$ is the turbulent viscosity coefficient.

Concrete form of the terms $F_\phi, A_\phi, G_\phi$ depends upon the meaning of $\phi$.

$F_w = \frac{g(T_v - T_{v0})}{T_{v0}} - gQ_c$, where $T_v$ is the virtual temperature, $T_{v0}$ - is the virtual temperature averaged over the cross sections of the both cylinders, $Q_c$ is the total mixing ratio of cloud drops $Q_c = \sum_{i=1}^{N-1} Q_{ci}$, $F_T, F_{Q_v}, F_{Q_{c_i}}$ describe the input of the microphysical process of condensation into the change of temperature and mixing rations of water vapor and cloud droplets in the $i$-th drop-size interval consequently. $A_w = A_{Q_v} = 0$, $A_T = -\gamma_a w / T$, where $\gamma_a$ is the dry adiabatic lapse rate, $A_{Q_{ci}} = V_{di} \frac{\partial}{\partial z} (Q_{ci})$, where $V_{di}$ - is the value of cloud drop terminal velocity of in the $i$-th drop-size interval, $G_w = 0$, $G_T, G_{Q_v}, G_{Q_{c_i}}$ describe the input of the microphysical process of evaporation into the change of temperature and mixing rations of water vapor and cloud droplets in the *i*-th drop-size interval consequently.

The detailed description of the dynamical part of the model is presented in [2].

It is well-known now that in order to predict cloud evolution characteristics properly one must use drop size dependent theories that is to include the equation describing the evolution of the number density function of the cloud drops $f$ into the system of cloud equations. Function $f = f(\vec{x}, m, t)$, where m is drop mass, varies in a given space point ($\vec{x}$) due to the processes of advection, sedimentation, turbulent mixing, condensation, nucleation and collection.

For the numerical solution of the equation it is necessary to select discrete points $m_\mathrm{i}$ ($i$ = 0,..., $N$, $m_0 = 0$) along the $m$ axis to define drop size intervals or bins. Then one can replace the stochastic collection equation by the set of equations for $M_i$ - mass fraction in the mass interval defined by:

$$M_i = \int\limits_{m_{i-1/2}}^{m_{i+1/2}} mf(m)dm, \tag{4}$$

$i$ = 1,..., $N$ - 1. So an equation for $\frac{\partial M_i}{\partial t}$ can be written in the following form:

$$\frac{\partial M_i}{\partial t} + \nabla \cdot \left[ \left( \vec{V} + \vec{V}_{di} \right) M_i \right] = \nabla \cdot (K_f \nabla M_i) + J^n \delta_{i1} + C_i + S_i^+ - S_i^- \tag{5}$$

where $\vec{V}$ is velocity vector, $\vec{V}_d$ is terminal velocity of the drop, $K_\mathrm{f}$ is turbulent diffusion coefficient, $C_i$ is rate of condensation (the growth of the drop due to the diffusion of water vapour) of the particle with mass $m$, $J^\mathrm{n}$ is rate of nucleation: rate of formation of the droplet of mass which belongs to the first drop mass interval.

$$\vec{V}_{d_i} = \frac{\int\limits_{m_{i-1/2}}^{m_{i+1/2}} \vec{V}_d(m) m f(m) dm}{\int\limits_{m_{i-1/2}}^{m_{i+1/2}} mf(m)dm} \approx \vec{V}_d(m_i), \tag{6}$$

$$C_i = \int\limits_{m_{i-1/2}}^{m_{i+1/2}} m \frac{\partial \dot{m} f(m)}{\partial m} dm,$$

Terms $S_i^+ - S_i^-$ characterize the process of collection: particle growth due to the collision of the drops with each other.

## 3 Numerical Scheme of the Model

The method of physical process splitting is used for solution of the system of the equations. Only dynamical processes are taken into account at the first stage.

Equations are numerically integrated using a finite difference method. Forward-upstream scheme is used. Vertical velocity is averaged over two grid points (point below is taken if $w \geq 0$ or point above if $w<0$). The final values are obtained on the second stage after completion of the microphysical processes calculation. A time step $\Delta t$ of 1 sec and a height interval $\Delta z$ of 200 m are used.
The height of the cylinder is 15 km. The temperature at the ground surface is 298K. The temperature laps rate is 9,8 K/km up to 2 km and is 6,3 K/km from 2 km to 10 km. The temperature is constant above 10 km. The relative humidity is 100% at the ground and decreases with lapse rate of 5%/km up to the top of cylinder. Initial contents of cloud droplets ($Q_c$) is equal to zero at all levels. Vertical (w) and radial ($u_a$) velocities and $Q_c$ are assumed to be 0 at the top and at the bottom boundaries of the cylinder. $Q_r$ and $Q_i$ are equal to zero at the top boundary. The initial disturbance of vertical velocity in the inner cylinder below 2 km is given as

$$w = \Delta w \cdot z \cdot (2 - z) \tag{7}$$

where $\Delta w$ is taken as 1 m/sec. The coefficient for lateral eddy mixing is 0,1. The vertical eddy diffusion coefficient equals to 100 $m^2$/sec.
Numerical scheme similar to that used in [3] was used for calculation of nucleation and condensation processes and similar to that used in [4] for calculation of collection.

## 4 Comparison of Microphysical and Dynamical Process Impact into Model Time Calculation

Each cloud model consists of two main blocks: dynamical one, which is responsible for calculation of velocity components and further transport of temperature and bulk characteristics of water vapor and cloud droplets, and microphysical block, which is responsible for calculation of distribution function of cloud droplets.
Dynamical block in 1-D and 1.5-D cloud droplets is rather simple and does not demand essential computational resources. Calculations have shown that it takes only 15 seconds of computer time to obtain full set of dynamical characteristics of model cloud, evaluating during 60 min. The result is quite acceptable for the needs of real-time forecast. But availability of only dynamical characteristics is not enough for qualitative forecast of a thunderstorm. For this aim we need to obtain data about time evolution of cloud droplet vertical profile, i.e. to calculate space and time characteristics of droplet distribution function provided by microphysical block of the model.
Distribution function calculation demands new mesh generation which should define drop mass intervals or bins in each node of the dynamical space mesh. If dynamical mesh consists of the N nodes, appearance of microphysical block results in increasing of a total number of calculations at least by factor of $N \cdot N_1^2$, where $N_1$ is a number of bins. Taking into account that $N_1 \geq N$ the number of required calculations increases tremendously and so it takes already 30 seconds of computer time to calculate 60 minute cloud evolution cycle if $N_1 \approx N$, 90 seconds if $N_1 = 2N$ and 1100 seconds if

$N_1 = 3N$. We should note that in order to obtain acceptable microphysical cloud characteristics one should take $N_l$ no less than $2 \cdot N$. So we can see that addition of the microphysical block for calculation of only one distribution function increases total calculation time by factor of 3. If one adds distribution function not only for cloud droplets but for different forms of ice crystals, hail and graupel, total calculation time becomes quite unacceptable for real-time forecast. And the necessity of parallelization technique use becomes quite evident.

## 5 Parallelization Model

Numerical scheme for the dynamical part of the model is an explicit one. So we can easily calculate all dynamical characteristics of the cloud at a time step "n+1" if we know them in each node of the mesh at a time step "n". And though to calculate dynamical characteristic in a mesh node "i" we should know corresponding characteristic in a neighbor mesh node "i-1", or "i+1" we can easily do this as all necessary values have been already calculated in the previous time step. That is why we can use space parallelization [5-8] for our problem solution.
For this purpose we divide computational region of the model into several subsections (Fig.2)

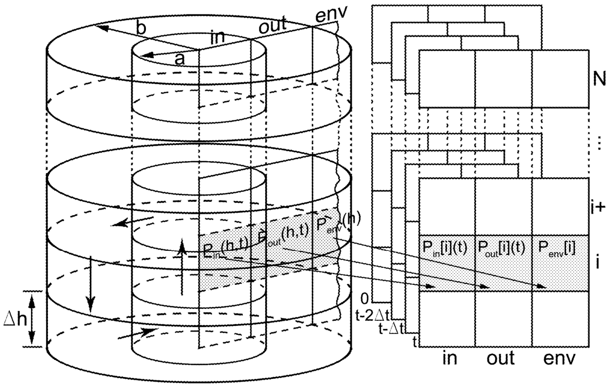

**Fig. 2**. Parallelization scheme of the model

Each subsection represents a cylinder of the height Δh and includes parts of the inner and outer cylinders as well as a part of the environment at rest.
Two methods of parallel calculations have been used. The first one supposes parallelization of only microphysical processes as the most time consuming. Dynamical part of the model is calculated within the entire computational region (big cylinder) while microphysical part is calculated in parallel within space subsections.

The second method implies parallelization of both dynamical and microphysical model blocks, so all cloud characteristics are calculated within each space subsection.

Multi-thread technology was used to realize parallelization methodology. Threads are created, and the data calculated on the previous time step is passed to the threads. Each thread implements calculation within definite mesh nodes. The transfer to the next time step is implemented when all threads fulfill their calculations.

As each launch of the thread demands definite time, the number of threads should be diminished in order to decrease computational overheads. It is optimal to launch N threads for N core processor or 2N threads if the cores optimize 2 threads implementation, while using Hyper-Threading technology for example.

As at each time step processor should wait for completion of implementation of all threads, the problem of load balancing appears to be challenging. It is not easy to find the solution because calculation of cloud characteristics in different subsections demands quite different time due to the fact that it is not necessary to obtain microphysical characteristics in the mesh subsections where cloud droplets are absent and relative humidity is less than 100%. Special procedure of mesh subsection redistribution was used to obtain equal time of thread implementation. The procedure implies calculation in neighboring subsections in different threads and provides acceptable level of load balancing.

It should be noted that some parts of the model program, such as creation and launch of the thread, calculation of boundary characteristics are calculated in single-thread regime.

## 6 Calculation Results

The results of numerical simulation show that the model is capable to describe warm rain processes in convective clouds under various vertical distributions of temperature and relative humidity of the outer atmosphere. The model reproduces evolution of vertical velocity, mixing ration of cloud droplets and cloud droplet spectrum in time and space. It can predict maximum and minimum values of the above mentioned dynamical and microphysical characteristics and besides the values of the height of a cloud base and upper boundary, precipitation rate and total quantity of the rainfall. All that characteristics are of major value for prediction of dangerous convective cloud phenomena such as thunderstorms, hails and rain storms.

Besides numerical experiments targeted to obtain physical results essential attention has been paid for investigation of calculation effectiveness of the model and especially for investigation the effectiveness of parallelization.

Three types of processors were used for model calculations: K1(Core 2 Duo 6400, 2.13 GHz, 2.5 GB, 2 cores), K2 (Core 2 Quad Q8200, 2.33 GHz, 2.5 GB, 4 cores), K3 (Core 2 Quad Q6600, 2.4 GHz, 2.0 GB ,4 cores). Calculations were provided for different number of bins (drop mass intervals), different number of threads and the two methods of parallel calculations (parallelization of only microphysical processes and parallelization of both microphysical and dynamical processes). The results are presented in the tables 1-4.

**Table 1.** Calculation time (seconds) of 1 hr model cloud evolution obtained with the help of different types of processors (K1, K2, K3). Parallelization of only microphysical processes is considered. 4 threads are used. $N^1$ – is the number of bins

| $N^1$ | 50 | 70 | 100 | 150 | 250 |
|---|---|---|---|---|---|
| K1 | 5,52 | 7,29 | 10,05 | 15,36 | 28,39 |
| K2 | 4,19 | 4,94 | 6,16 | 8,78 | 15,02 |
| K3 | 4,09 | 4,78 | 6,14 | 8,70 | 14,84 |

The results presented in the table 1 show that parallelization with the help of 4 core processors is more efficient than with the help of 2 core processor. Efficiency increases with increasing of the number of bins. Time difference between 2 and 4 core processors is about 20% in case of $N^1$=50 and is about 50% in case of $N^1$=250.

**Table 2.** Calculation time (seconds) of 1.5 min model cloud evolution obtained with the help of different number of threads (NTh) (processor K3). Parallelization of only microphysical processes is considered. $N^1$ – is the number of bins.

| $N^1$ | 50 | 70 | 100 |
|---|---|---|---|
| NTh = 1 | 15,60 | 36,19 | 89,14 |
| NTh = 2 | 9,17 | 19,09 | 46,84 |
| NTh = 3 | 13,05 | 17,09 | 34,80 |
| NTh = 4 | 22,56 | 24,25 | 31,66 |

The results presented in the table 2 show that thread number influence is depended on the bin number ($N^1$). At the smallest value of $N^1$ the most effective is 2 threads using, at the biggest value 4 threads using is the most effective.

**Table 3.** Calculation time (seconds) of 1 hr model cloud evolution obtained with the help of different types of processors (K1, K2). Parallelization of both microphysical and dynamical processes is considered. 4 threads are used. $N^1$ – is the number of bins.

| $N^1$ | 50 | 70 | 100 | 150 | 250 |
|---|---|---|---|---|---|
| K1 | 5,14 | 6,73 | 9,44 | 14,30 | 22,62 |
| K2 | 3,90 | 4,51 | 5,64 | 7,66 | 12,86 |

The results presented in the table 3 prove the above conclusion that 4 core processor is more efficient that 2 core. If we compare data in the tables 2 and 3 we can see that dynamical process parallelization contributes not so much in calculation time decrease (less than 10%). So the most time consuming part of the model is microphysical block which should be parallelized first of all.

And at last we present data (table 4) which characterize computational time of sequential algorithm. Comparison of the results in tables 4, 2 and 3 shows that speed up of cloud evolution calculation varies from 1,5 up to 3,0 dependent upon bin number.

**Table 4.** Calculation time (seconds) of 1 hr model cloud evolution obtained with the help of different types of processors (K1, K2). Without Parallelization. $N^1$ – is the number of bins.

| $N^1$ | 50 | 70 | 100 | 150 | 250 |
|---|---|---|---|---|---|
| K1 | 7,86 | 10,61 | 15,52 | 23,70 | 43,00 |
| K2 | 6,27 | 8,80 | 12,47 | 19,38 | 37,36 |

Speed up is less than 4 (the number of threads) because of the time spent on thread creation and launch as well as on operations which should be provided in one thread regime.

## 7 Conclusions

1.5-D convective cloud model with detailed description of microphysical processes is presented in the paper. Possibilities of the model parallelization on multi-core processors with the different number of cores have been investigated. It is shown that parallelization with the help of 4 core processors is more efficient that with the help of 2 core processors. Multi-thread technology was used for realization of parallel algorithm. It is obtained that the number of threads should be equal or should be 2 times more than the number of processor cores. Comparison of the calculation results of sequential and parallel algorithms shows that speed up can vary from 1,5 to 3,0 in case of 4 parallel threads use. Investigation shows that use of rather complex numerical models for real-time forecast of dangerous convective phenomena is possible in case of realization of model parallelization on multi-core processors.

## References

1. Asai T., Kasahara A.: A Theoretical Study of the Compensating Downward Motions Associated with Cumulus Clouds. Journal of the Atmospheric Sciences, vol. 24, pp. 487-497 (1967)
2. Raba N.O. Stankova E.N. Research of influence of compensating descending flow on cloud's life cycle by means of 1.5-dimensional model with 2 cylinders. Proceedings of MGO, V.559, pp. 192-209 (2009) (in Russian).
3. Khain A., Pokrovsky A., Pinsky M.: Simulation of Effects of Atmospheric Aerosols on Deep Turbulent Convective Clouds Using a Spectral Microphysics Mixed-Phase Cumulus Cloud Model. Part I: Model Description and Possible Applications. Journal of the Atmospheric Sciences, vol. 61, pp. 2963-2982 (2004)
4. Stankova E.N., Zatevakhin M.A. The modified Kovetz and Olund method for the numerical solution of stochastic coalescence equation. Proceedings 12th International Conference on Clouds and Precipitation, Zurich, 19-23 August 1996, pp.921-923.
5. Voevodin V.V. Informational structure of sequential programs. Russ. J. of Num. aAn. and Math . Modelling.V.10, N 3, pp. 279-286 (1995).
6. Voevodin V.V. Mathematical foundations of parallel computing. World Scientific Publishing Co., Series in computer science. V. 33, P. 343 (1992)

7. Bogdanov A.V., Korkhov V.V., Mareev V.V., Stankova E.N. Architectures and topologies of multiprocessor computational systems, P.176 (2004,) (in Russian)
8. Babb R.G. (eds) Programming Parallel Processors. Addison-Wesly Publishing Company (1988)