

# Improving performance of SLAV model for medium range weather prediction



Фадеев Р.Ю., Толстых М.А.,  
Гойман Г.С., Шашкин В.В.

ИВМ РАН

Гидрометцентр России

МФТИ

e-mail: [rost.fadeev@gmail.com](mailto:rost.fadeev@gmail.com)

23.09.2023, Суперкомпьютерные дни в России

# Ключевая идея

Ускорение вычислений за счет дополнительных к вычислительным MPI-процессов, основным назначением которых является работа с файловой системой.

*Первая реализация:*

*M. Tolstykh, G. Goyman, R. Fadeev et al. Structure and Algorithms of SL-AV Atmosphere Model Parallel Program Complex // Lobachevskii J Math. 179, 2018. Vol. 39. P. 587–595.*

Особенность вычислительной системы:

Расчет прогноза погоды происходит на выделенной части Cray XC-40 (ГВЦ Росгидромета). Файловая система Lustre – общая для всех задач (в т.ч. для обработки данных).

Мотивация:

Среднесрочный прогноз погоды на 24 часа должен рассчитываться за не более чем 20 минут.

# ПЛАВ – это многомасштабное моделирование атмосферы

1. Одна программная реализация, разные конфигурации.

2. Принципиальные конфигурации ПЛАВ:

ПЛАВ072L96 - долгосрочный прогноз (оперативно);

- субсезонный прогноз (оперативно);

- ансамблевый среднесрочный (оперативно);

ПЛАВ20

- ансамблевый среднесрочный (в разработке);

ПЛАВ10

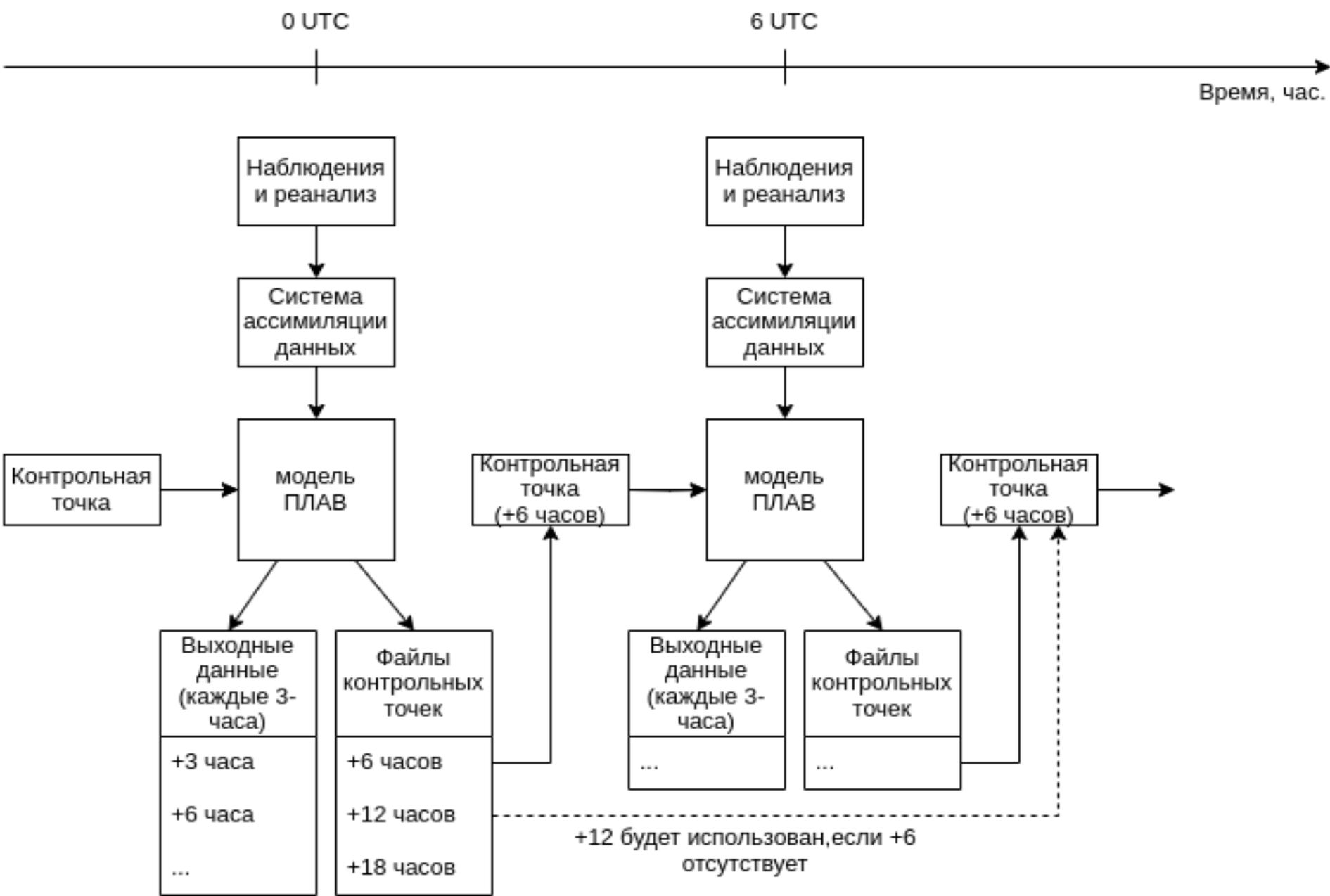
- среднесрочный прогноз высокой детализации (оперативно);

- краткосрочный прогноз для гражданской авиации (оперативно)

3. Сопутствующие технологии: генерация ансамбля, усвоение почвенных характеристик, блок диагностики, представление результатов и др.

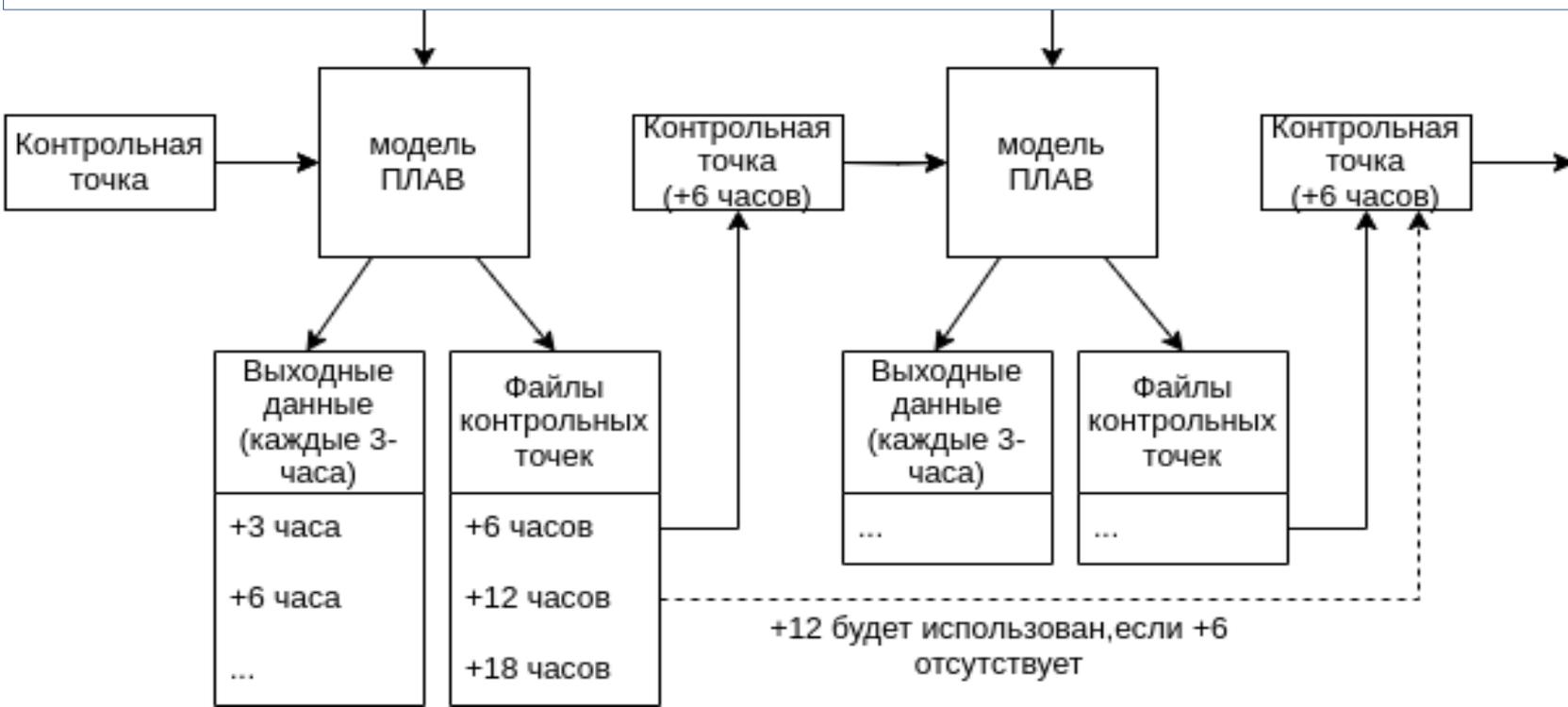


# Технология прогноза на основе ПЛАВ



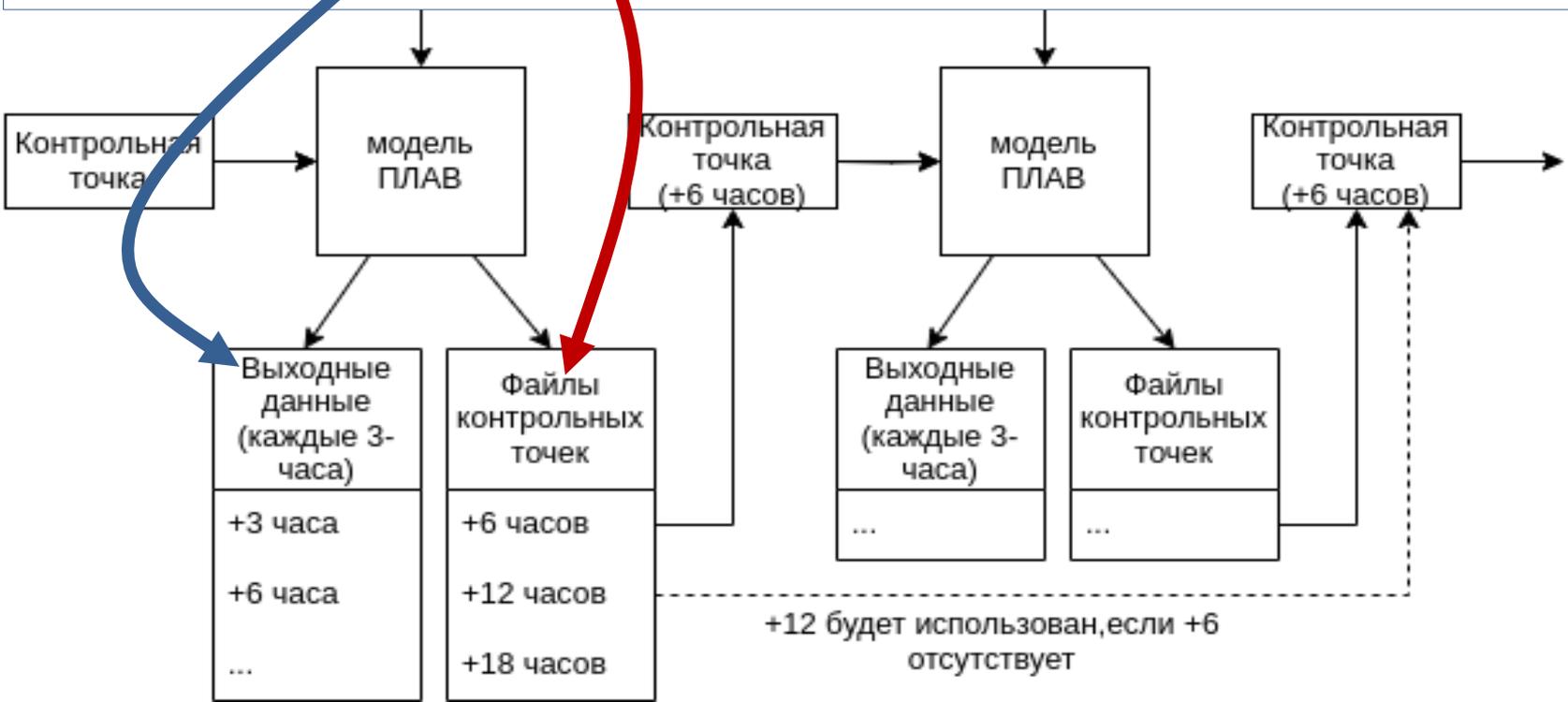
# Технология прогноза на основе ПЛАВ

Идентификатор модели	ПЛАВ072L96	ПЛАВ20	ПЛАВ10
Число узлов по долготе	400 (0.9°)	1600 (0.225°)	3600 (0.1°)
Число узлов по широте	251 (0.72°)	865	1945
Число уровней по вертикали	96	51	104
Шаг по времени, с.	1440	450	270
Число шагов по времени в сутки	60	192	320
Размер файла с контрольной точкой, Гб.	0.3	<b>2.8</b>	<b>22.7</b>
Размер сжатого файла с контрольной точкой, Гб.	0.3	1.9	9.9
Размер файла с выходной продукцией, Гб,	-	0.7	3.3



# Технология прогноза на основе ПЛАВ

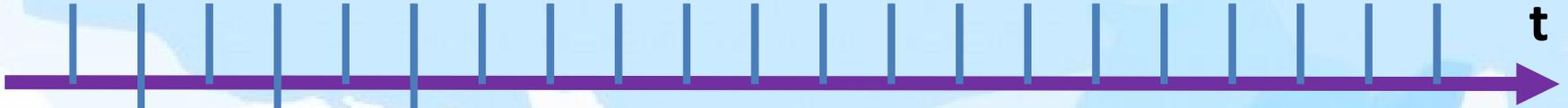
Идентификатор модели	ПЛАВ072L96	ПЛАВ20	ПЛАВ10
Число узлов по долготе	400 (0.9°)	1600 (0.225°)	3600 (0.1°)
Число узлов по широте	251 (0.72°)	865	1945
Число уровней по вертикали	96	51	104
Шаг по времени, с.	1440	450	270
Число шагов по времени в сутки	60	192	320
Размер файла с контрольной точкой, Гб.	0.3	<b>2.8</b>	<b>22.7</b>
Размер сжатого файла с контрольной точкой, Гб.	0.3	1.9	9.9
Размер файла с выходной продукцией, Гб,	-	0.7	3.3



# Работа с файловой системой в процессе счета ПЛАВ

## Выходные файлы

3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48



6 12 18

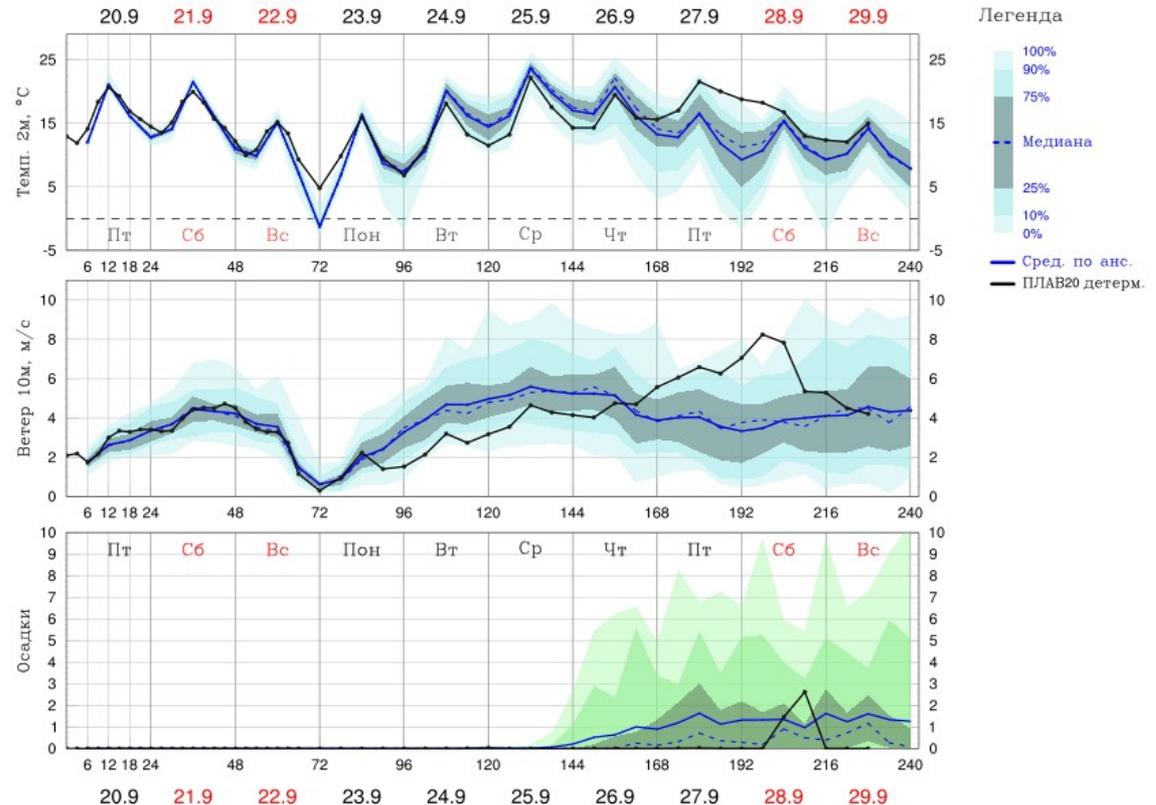
## Контрольные точки

Метеостанция МОСКВА, ВДНХ (27612)

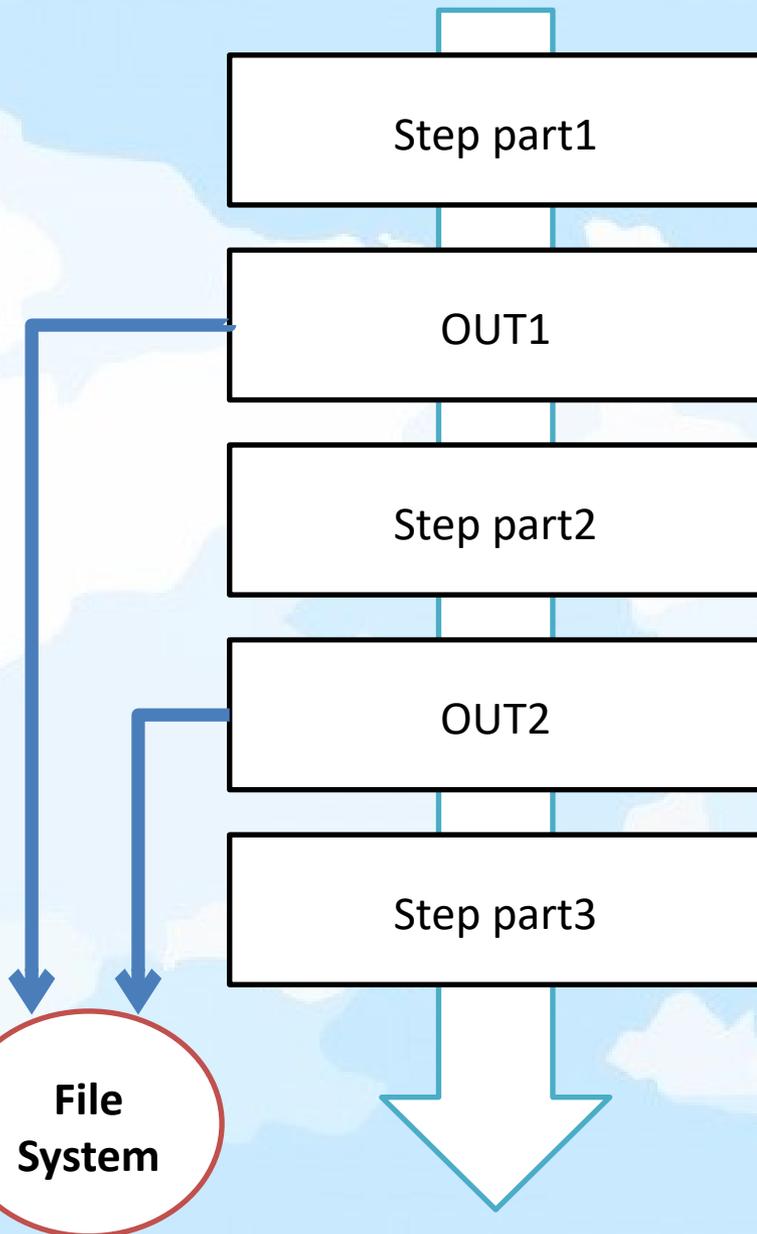
Прогноз от 0 ВСВ 20.9.2024 на 240 час.

Высота 147 м.  
55.83° с.ш., 37.62° в.д.  
Москва

модель  
ПЛАВ072L96  
Размер ансамбля 40



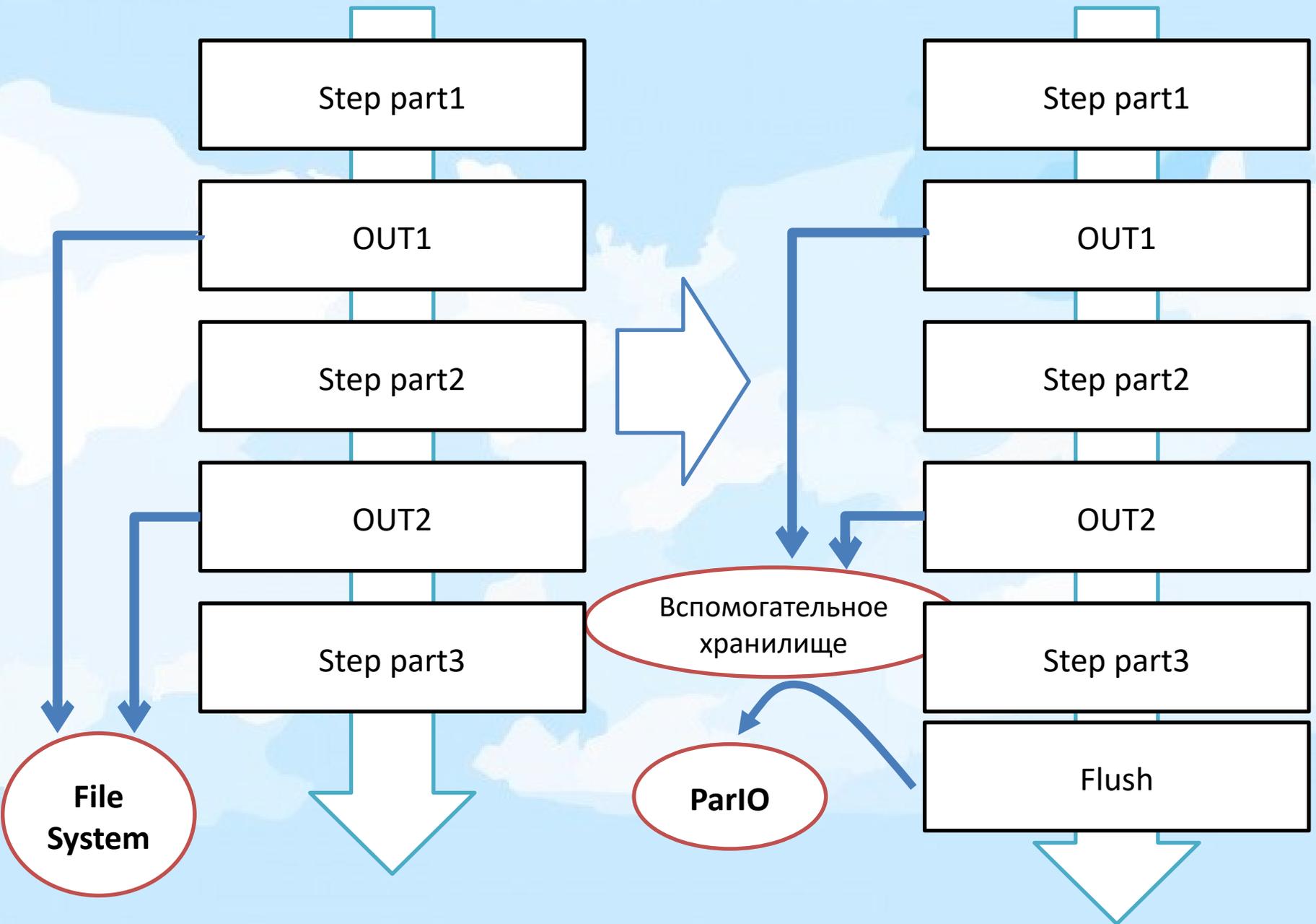
# Организация процедуры расчета 1 шага по времени



Запись выходной продукции осуществляется «на месте» по мере готовности данных.

Работа с файловой системой – блокирующая вычисления операция.

# Организация процедуры расчета 1 шага по времени



# Организация процедуры расчета 1 шага по времени

Цель: минимальное изменение/ветвление программного кода.

Вместо записи данных – сохраняем (одинарная точность).

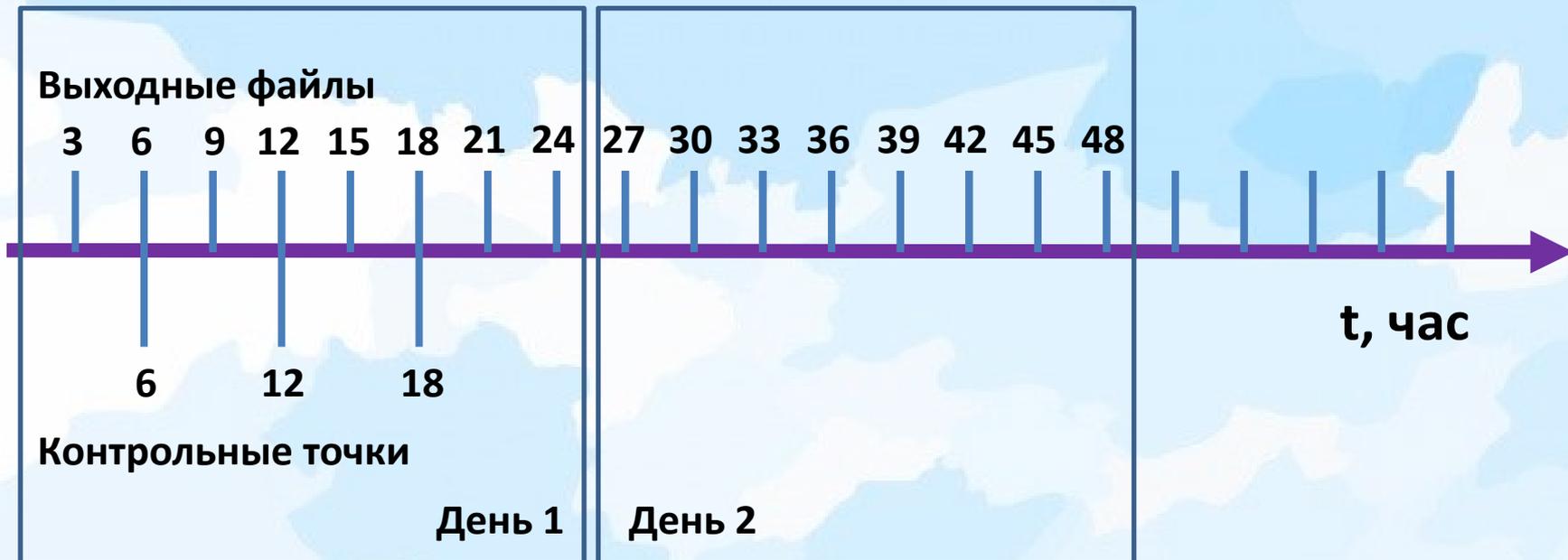
В конце шага по времени выполняем дополнительную процедуру упаковки и выгрузки данных в ParIO.

Получение данных и их запись в ParIO – разные процедуры.



# Численные эксперименты

Основные характеристики качества: время расчета первого и второго модельных дней. Изучаются модели: ПЛАВ10 и ПЛАВ20.

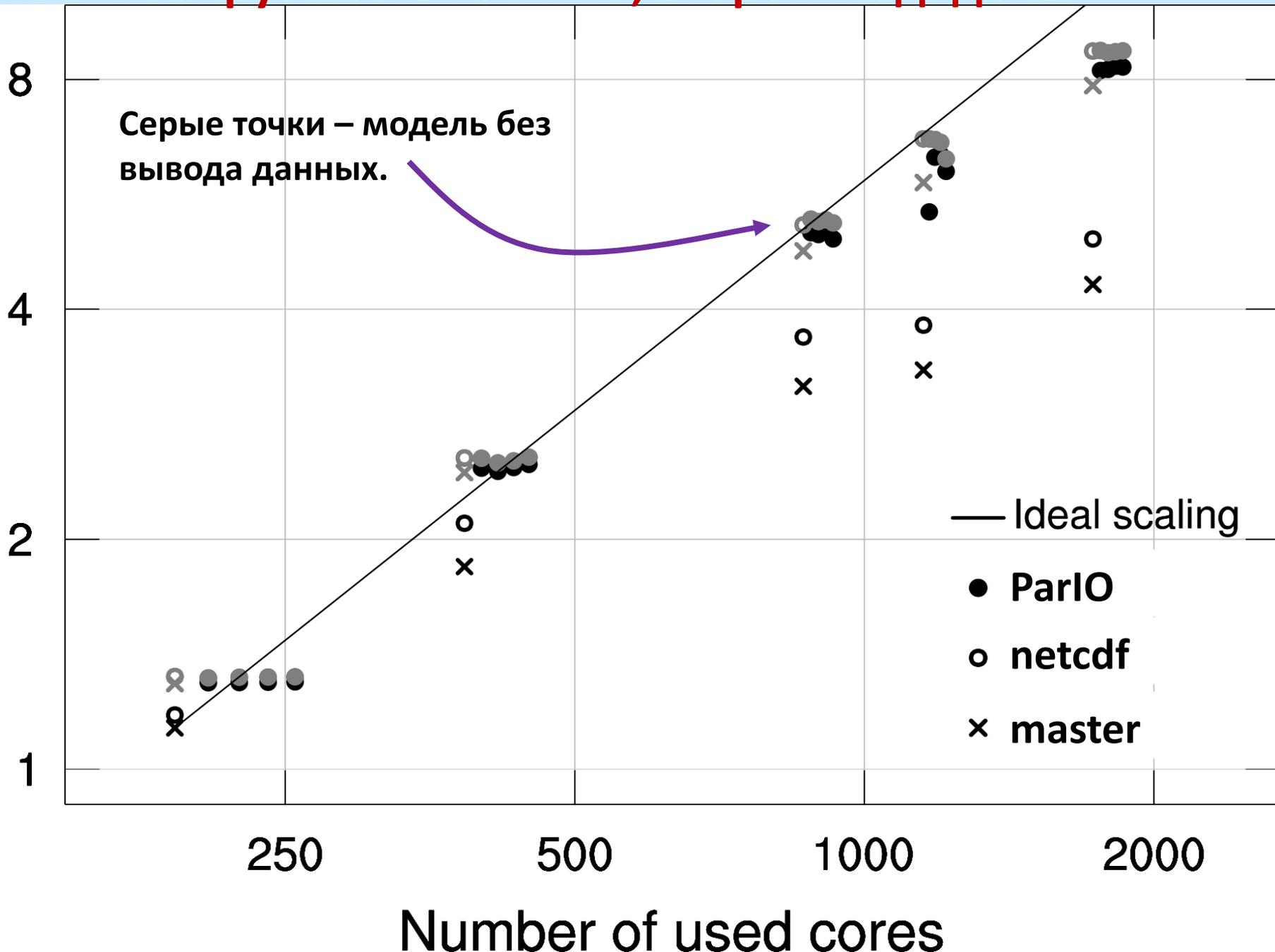


Исследуем зависимость времени расчета от размера MPI-коммуникатора ПЛАВ и ParIO.



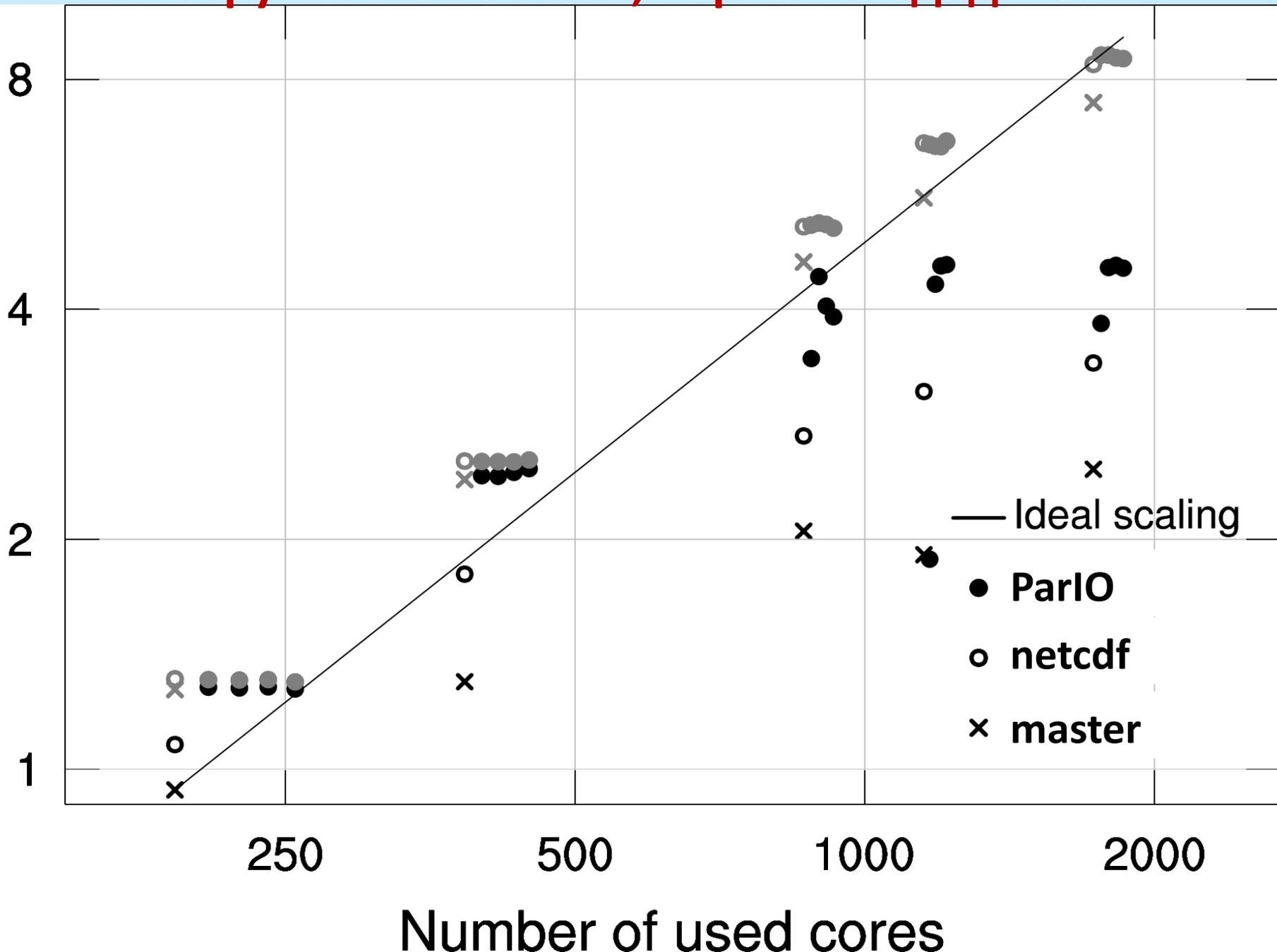
# Масштабируемость ПЛАВ20, второй мод. день.

Simulated days per 10 min.



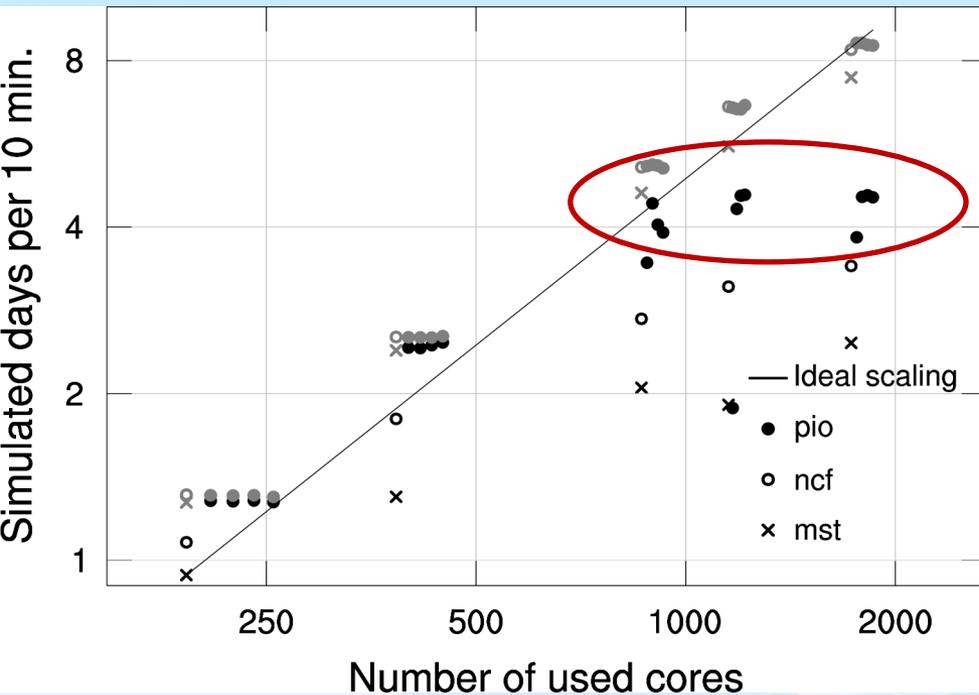
# Масштабируемость ПЛАВ20, первый мод. день.

Simulated days per 10 min.

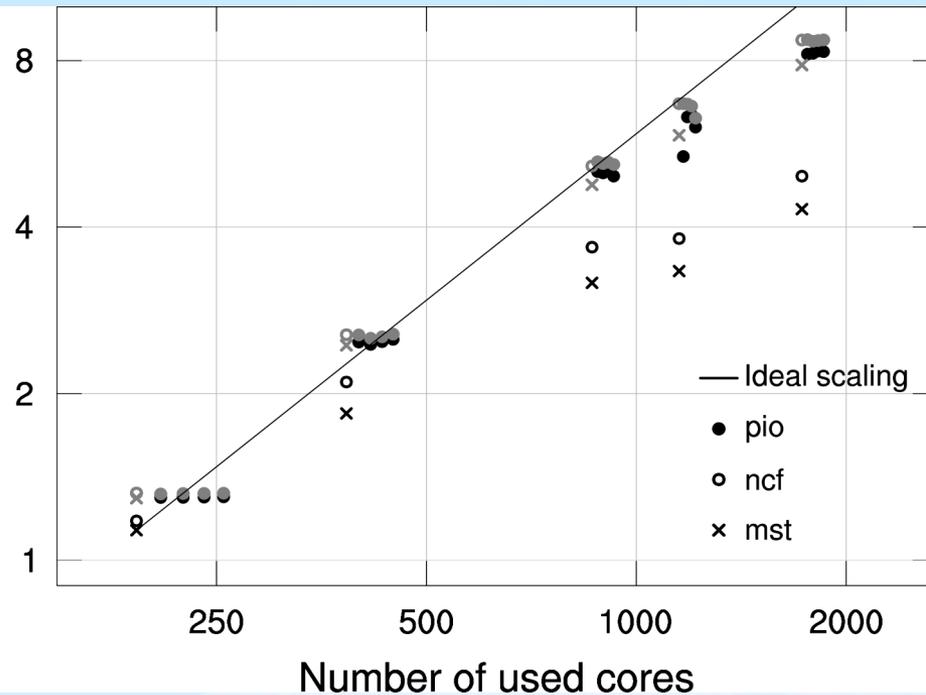


# Масштабируемость ПЛАВ20

Первый модельный день.



Второй модельный день.



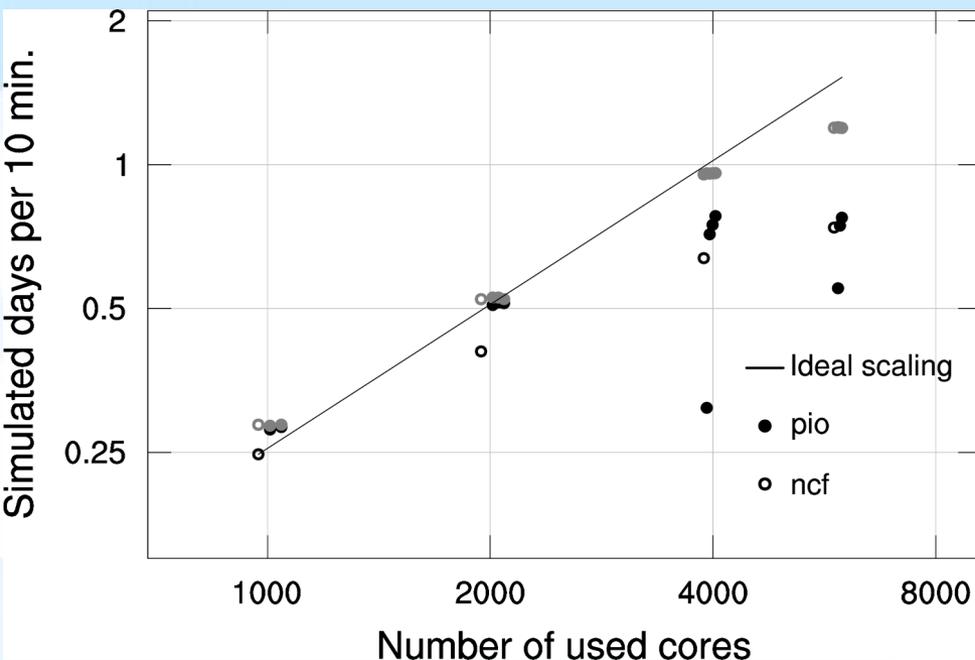
Новое узкое место:

производительность файловой системы + особенности использования ParIO.

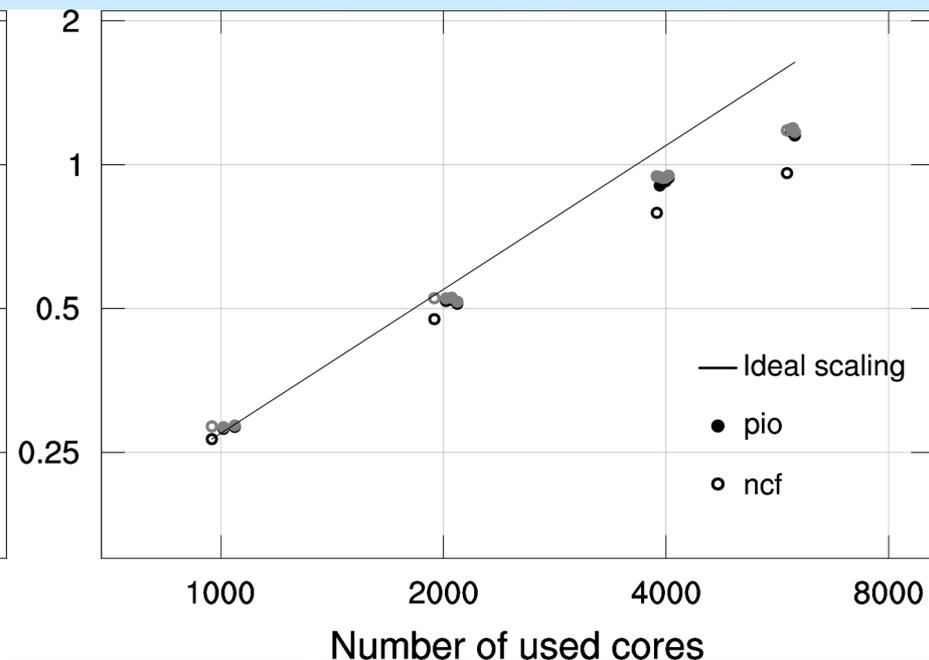
Ранее для оперативного среднесрочного прогноза погоды применялась версия ПЛАВ20 в конфигурации, использующая 768 процессорных дер.

# Масштабируемость ПЛАВ10

Первый модельный день.



Второй модельный день.



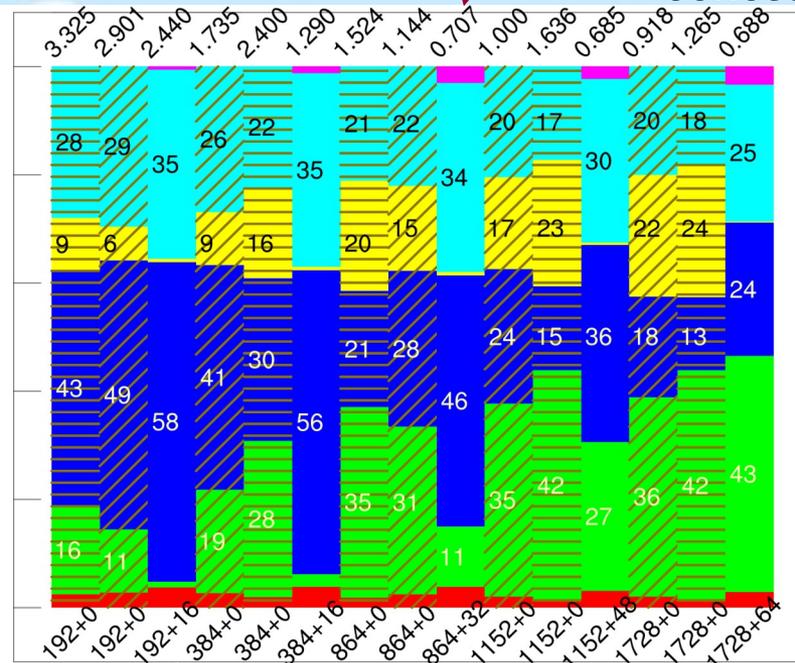
Для оперативного среднесрочного прогноза погоды применяется версия ПЛАВ10 в конфигурации, использующая 3960 процессорных ядер.

# ПЛАВ20

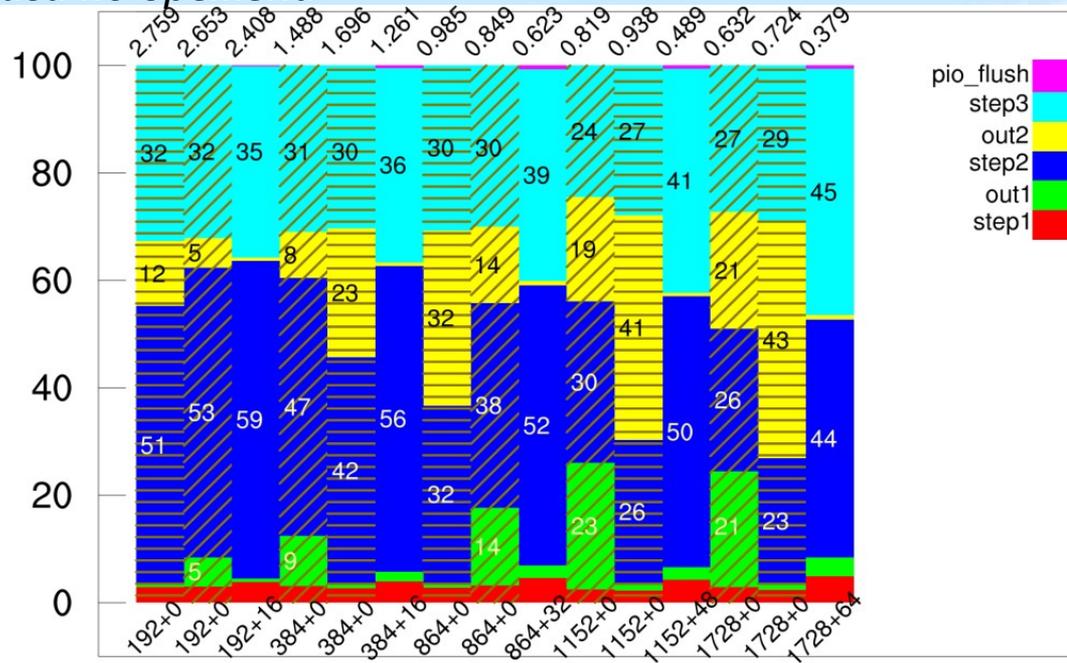
Первый модельный день.

Второй модельный день.

Среднее время расчета  
одного шага по времени



slav20 configurations



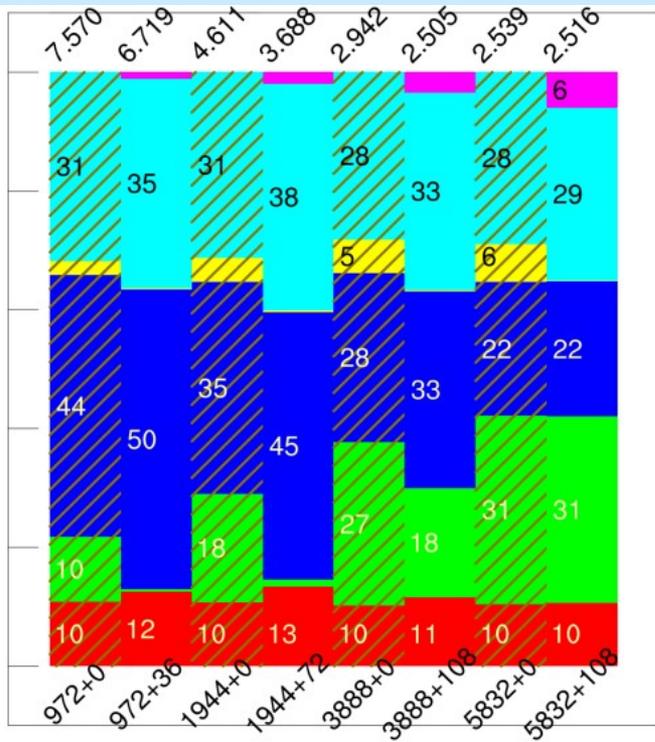
slav20 configurations

Число ядер ПЛАВ + число ядер ParIO

# ПЛАВ10

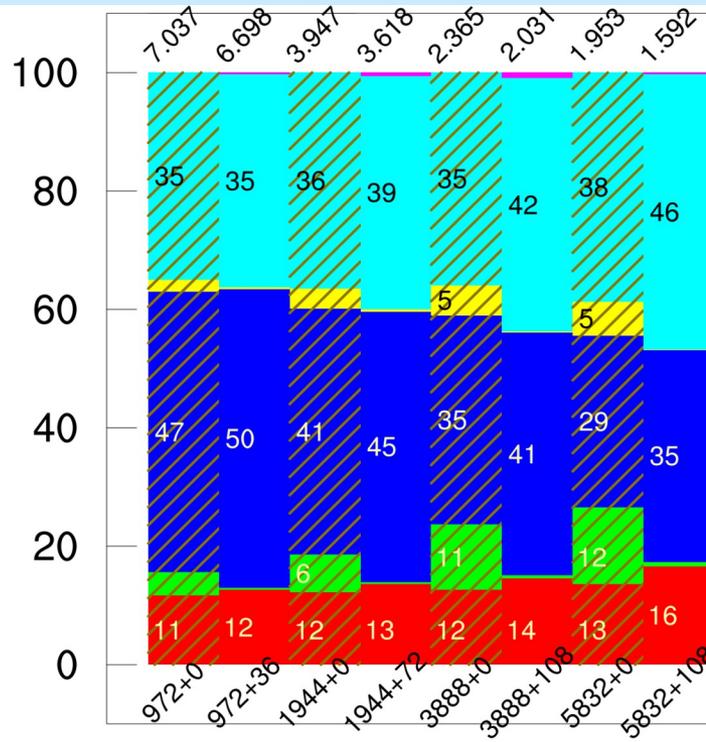
Первый модельный день.

Второй модельный день.



slav10 configur.

Time-step fraction, %



slav10 configurations

- pio\_flush
- step3
- out2
- step2
- out1
- step1

# Особенности использования ParIO

```
aprun \  
-n 4 -d 8 -N 4 ./pario post : \  
-n 4 -d 8 -N 4 ./pario ctrp : \  
-n 216 -d 4 -N 8 --mpmd-env OMP_NUM_THREADS=4 ./slm20L51v.exe \  
2024 01 30 00 > $cdir/slm$slmid.out
```

*ParIO запускается в виде отдельных исполняемых файлов в дополнение к ПЛАВ.*

*Преимущества такого подхода:*

*Не нужно вносить изменения в конфигурационный файл: способ работы с файловой системой регулируется способом запуска расчетной модели.*

*Индивидуализация параметров окружения для каждой исполняемой программы.*

# Особенности использования ParIO

```
call MPI_AllGatherV(...)
```

```
call MPI_Comm_split(..)
```

```
Do i
```

```
  pda(i) % tag_range = (/ 40000, 60000 /)
```

```
End do
```

```
call MPI_Comm_split(..)
```

```
call pda(i) % pio % init(type = trim(type),
```

```
  & tag_range = pda(i) % tag_range,
```

```
  & pio_cw = pda(i) % pio_cw, local_cw = pda(i) % child_cw,
```

```
  & master_np = pda(i) % master_np, hybrid_np_min = hybrid_np_min,
```

```
  & hybrid_np_max = hybrid_np_max, child_np = pda(i) % child_np,
```

```
  & mem_use = mem_use, mem_tot = mem_tot, print_lev = 1)
```

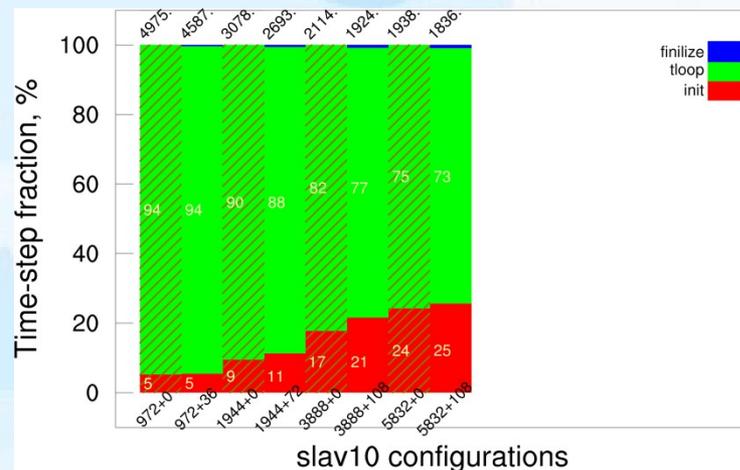
# Ускорение вычислений за счет оптимизации работы с файловой системой. Выводы.

Запись данных с помощью внешних делегатов позволяет существенно повысить скорость вычислений при сравнительно небольшом увеличении числа задействованных моделью ресурсов (2-3%).

Наилучший результат достигается в случае использования нескольких узкоспециализированных делегатов.

Спасибо за внимание!

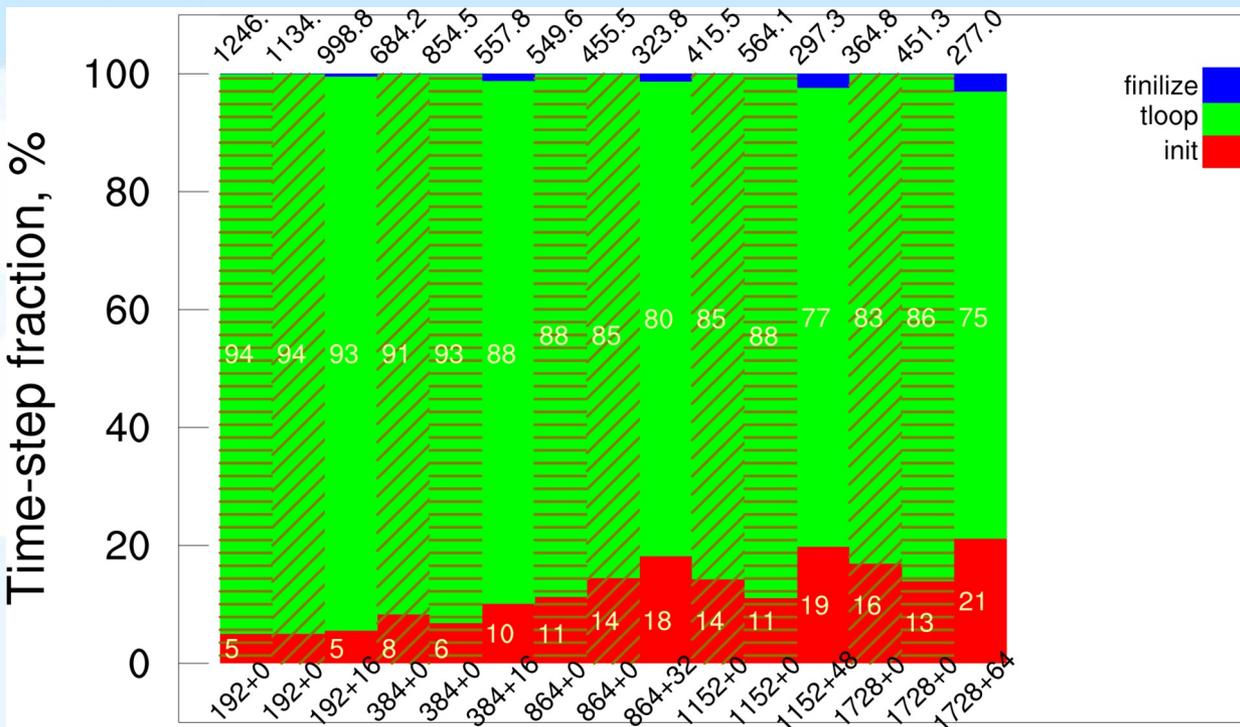
Ростислав Юрьевич Фадеев  
ИВМ РАН, Гидрометцентр России  
[rost.fadeev@gmail.com](mailto:rost.fadeev@gmail.com)  
Telegram: @rostfa



Задача следующего уровня:

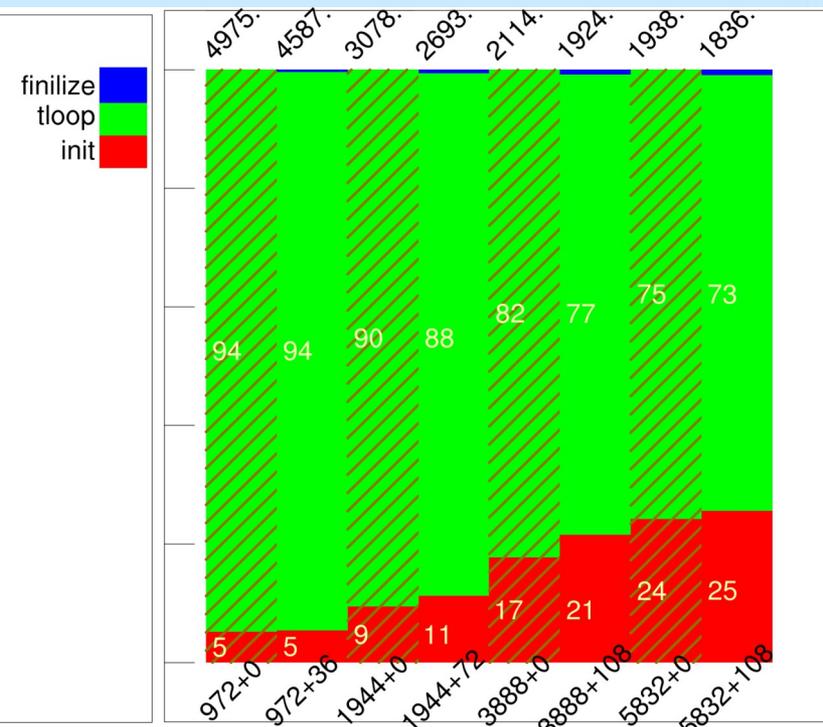
Считывание моделью начальных и других данных занимает существенную часть времени расчетов.

## ПЛАВ20



slav20 configurations

## ПЛАВ10



slav10 configura

Чтение моделью начальных и других данных занимает существенную часть времени расчетов.