

# Производительность Osmosis в операциях обрезки по регионам РФ

Результаты тестов

[Обсудить в форуме](#) Комментариев — 2

[Osmosis](#) - программа на Java используемая для операций ввода/вывода для данных OSM XML создаваемых в рамках проекта [OpenStreetMap](#).

Эта статья демонстрирует результаты тестирования производительности Osmosis в операциях экстракции фрагмента данных (обрезки). Эта статья не является руководством пользователя Osmosis, данный вопрос достаточно подробно описан на вики OSM.

Для тестов использовался последняя на момент написания статьи сборка Osmosis ([0.33](#)). Тестирование производилось на сервере под управлением FreeBSD 7.1-STABLE, 8Gb RAM, 2x4 CPU (Xeon Quad 2.1 ГГц). Запуск osmosis осуществлялся в 6 параллельных потоков.

В качестве исходного использовался фрагмент общей БД (planet-latest) на территорию РФ (--bounding-box top=82.2 left=26.8 bottom=41.8 right=180), размер файла 3.2 Gb, компрессии нет. Для обрезки использовались файлы обрезки POLY созданные на базе административных границ из VMap0 ([скачать файлы обрезки](#)).

Основной задачей тестов было установить работоспособность режима completeways. Результаты обрезки обновляемые ежедневно по методу PI можно скачать [здесь](#). С подробным описанием всех настроек можно ознакомиться на [странице Osmosis](#).

## Тесты

Производились следующие тесты:

**PD** - обрезка полигональными объектами с параметрами по-умолчанию. Командная строка:

```
osmosis --read-xml file=rus.osm --bounding-polygon file=regname.poly --write-xml file=outputreg.osm
```

**PDB** - обрезка полигональными объектами с параметрами по-умолчанию, режим индексирования узлов - bitset. Командная строка:

```
osmosis --read-xml file=rus.osm --bounding-polygon file=regname.poly used-node idTrackerType=BitSet --write-xml file=outputreg.osm
```

**P** - обрезка полигональными объектами с параметрами по-умолчанию, отключен парсинг дат. Командная строка:

```
osmosis --read-xml file=rus.osm enableDateParsing=no --bounding-polygon file=regname.poly --write-xml file=outputreg.osm
```

**PI** - обрезка полигональными объектами с корректной обрезкой выходящих за границы полигона обрезки объектов. Командная строка:

```
osmosis --read-xml file=rus.osm --bounding-polygon file=regname.poly clipIncompleteEntities=true --write-xml file=outputreg.osm
```

**PC** - обрезка полигональными объектами с добавлением выходящих за границы полигона обрезки объектов. Командная строка:

```
osmosis --read-xml file=rus.osm --bounding-polygon file=regname.poly completeWays=yes --write-xml file=outputreg.osm
```

**PCB** - обрезка полигональными объектами с добавлением выходящих за границы полигона обрезки объектов и указанием режима индексирования узлов BitSet в ограничивающем полигоне. Командная строка:

```
osmosis --read-xml file=rus.osm --bounding-polygon file=regname.poly completeWays=yes  
idTrackerType=BitSet --write-xml file=outputreg.osm
```

**PCCB** - обрезка полигональными объектами с добавлением выходящих за границы полигона обрезки объектов и отношений и указанием режима индексирования узлов BitSet в ограничивающем полигоне. Командная строка:

```
osmosis --read-xml file=rus.osm --bounding-polygon file=regname.poly completeWays=yes  
completeReations=yes idTrackerType=BitSet --write-xml file=outputreg.osm
```

**MD, M** - эквивалентны PD и P, но в качестве полигонов обрезки используются прямоугольные охваты (количество узлов в полигоне обертки = 4).

**NodesPoly** - количество узлов в файлах обрезки.

**Nodes** - примерное количество узлов в обрезанных регионах по алгоритму PD.

## Выводы

1. Производительность osmosis при обрезке с completeways=yes - удовлетворительная, при правильном указании параметров, но занимает примерно в два раза больше времени.
2. Предварительная обрезка небольшим районом и последующая обрезка точным контуром с с completeways=yes - практически не изменяет скорости операции. Результаты этого теста не приводятся, так как он проводился не только на отдельном, но достаточно большом районе (kirov).
3. Режим индексирования узлов bitset не дает ускорения, ни с completeways=yes, ни без (результаты не показаны), если --use-node указывается глобально, а не для --bp.
4. Обрезка полигоном с сотнями или тысячами узлов или прямоугольным охватом при прочих настройках по умолчанию может дать прирост производительности в несколько раз, но только для больших областей, для небольших эффект может быть обратным.
5. Прирост при отключение парсинга дат пренебрежимо мал.

## Результаты тестов

[Скачать](#) таблицу. Время в минутах.

	NodesPoly	Nodes	PD	PDB	P	PI	PC	PCB	PCCB	MD	M
adygeya	275	186868	2,52	6,51	2,47	2,86	14,2	7,61	5,94	2,78	2,44
altay	1623	67987	2,68	7,2	2,55	2,64	18	6,64	6,24	3,77	2,43
altayskiy	1839	139993	2,97	7,24	3,04	3,19	93,3	7,87	6,58	2,45	2,5
amur	5481	59930	3,15	6,77	3,19	3,24	22,6	8,97	5,8	2,47	2,38
arkhan	2100	79140	2,81	6,32	2,79	3,1	33,3	7,17	6,74	2,49	2,39
astrakhan	1090	20808	2,43	6,39	2,37	2,45	7,8	6,37	6,45	2,36	2,49
bashkir	2945	155642	3,64	8	3,61	3,94	25,7	11,1	8,2	2,48	2,51
belgorod	1188	98304	2,67	5,92	2,56	2,63	49,2	9,21	5,87	2,41	2,51
bryansk	1690	37422	2,61	5,83	2,52	2,58	8,71	9,73	6,18	2,5	2,38
buryat	4264	60982	5,67	9,2	5,63	6,2	25,9	13,1	6,88	2,55	2,53
chechen	665	9548	2,44	5,57	2,32	2,3	7,02	6,85	6,33	2,43	2,39
chel	3110	294647	4,02	7,32	3,88	3,82	542	7,87	7,91	2,6	2,57
chukot	3601		2,37	6,48	2,87		29,1	11	6,13		
chuvash	801	16638	2,41	6,04	2,38	2,38	8,22	7,04	6,13	2,45	2,42
dagestan	1428	10260	2,37	6,38	2,34	2,67	5,83	7,5	5,65	2,69	2,34
evrey	1044	18670	2,48	7,13	2,43	2,47	6,5	7,13	5,99	3,98	2,39
ingush	332	9541	2,34	6,78	2,34	2,34	6,42	7,31	5,44	2,55	2,39
irkutsk	8280	415039	10,4	14,6	10,2	10,6	1140	20,2	8,99	2,66	2,53
ivanov	921	19285	2,42	6,58	2,43	2,49	7,12	8,93	6,19	2,38	2,46

kabardin	548	21224	2,39	6,55	2,32	2,43	7,48	6,5	5,97	2,38	2,38
kalinin	895		2,33	6,07	2,45	2,32	7,24	8,64	6,29	2,34	2,3
kalmyk	1440	5591	2,4	6,53	2,45	2,41	6,98	7,85	5,82	2,37	2,42
kaluzh	1216	68416	2,57	6,9	2,53	2,52	16,9	7,82	5,81	2,58	2,43
kamch	4186	113053	3,35	6,72	3,21	3,19	71,5	7,62	5,44	2,48	2,49
karach	690	46363	2,44	6,32	2,44	2,42	9,49	6,88	6,58	2,46	2,36
karel	1753	747183	4,99	8,73	4,93	5,76	2220	10,6	8,33	2,74	2,69
kemerovo	1662	85893	2,78	6,46	2,77	2,86	29,8	8,37	6,84	2,62	2,46
khabar	6903	143526	4,79	8,68	4,77	4,93	118	7,28	6,81	2,46	2,5
khakas	1113	37405	2,5	5,86	2,46	2,47	10,2	7,58	6,14	2,45	2,45
khanty	3143	46308	2,91	6,22	2,79	3,14	16,4	8,02	5,93	2,63	2,37
kirov	2592	107391	3,04	6,48	3,08	3,23	48,4	11,4	7,5	2,34	2,58
komi	2587	97879	3,24	6,75	3,25	3,87	38,9	7,61	6,09	2,52	2,47
kostrom	1635	28085	2,49	5,74	2,39	2,43	7,43	6,21	5,91	2,39	2,42
krasnodar	2038	471317	4	7,23	4,02	3,95	777	11,2	9,01	2,48	2,58
krasnoyarsk	7762	468826	21,7	24,2	21,9	24,9	957	38,6	14,1	2,84	2,92
kurgan	1796	17098	2,49	6,1	2,48	2,53	6,66	8,68	5,94	2,41	2,48
kursk	1235	108055	2,68	6,66	2,67	2,63	10,2	7,8	6,43	2,47	2,45
leningrad	1701	640291	4,38	8,42	4,5	4,84		8,92	6,35	2,79	2,75
lipetsk	982	21507	2,54	5,65	2,48	2,46		6,55	5,36	2,4	2,36
magadan	3615	21507	3,25	6,45	2,88	2,9		6,01	5,52	2,4	2,44
mariyel	1099	29435	2,53	5,78	2,45	2,49		6,37	5,59	2,49	2,48
mordov	1338	13579	2,52	6,27	2,45	2,65	6	7,49	5,66	2,56	2,39
mosobl	1536	987687	5,32	9,01	5,15	5,34	6700	11	9,95	2,8	2,96
murmansk	1992	293136	3,54	8,19	3,5	3,66	232	8,53	7,09	4,06	2,51
nenec	1015	144979	2,92	6,3	2,98	3,18	133	9,69	6,71	2,52	2,56
nizhegorod	1854	100113	2,81	6,36	2,78	2,82	52,7	10,1	6,04	2,47	2,38
novgorod	1566	215100	3	6,61	3,06	3,27	199	9,42	6,98	2,5	2,5
novosib	2030	104029	2,96	6,75	2,95	3,08	51,8	7,77	7,82	2,49	2,45
omsk	2002	54111	2,58	6,15	2,6	2,58	21,7	7,74	6	2,41	2,42
orenburg	3866	103797	4,16	8,21	4,11	3,89	28,8	7,72	7,81	2,59	2,46
orlovsk	1032	16520	2,34	5,7	2,36	2,36	14,7	7,39	5,6	2,47	2,47
osetiya	669	9543	2,48	5,82	2,33	2,34	6,61	9,3	6,17	2,41	2,4
penz	1286	55060	2,47	5,99	2,47	2,52	19,6	7,04	6,68	2,49	2,4
perm	2177	150118	3,07	6,66	3,11	3,11	117	11,2	7,46	2,52	2,53
prim	335	82003	3,05	6,47	2,92	2,9	29,8	8,88	6,23	2,42	2,46
pskov	1753	147417	3	6,39	3,05	3,07	44	7,81	6,76	2,56	2,57
rostov	2598	70082	3,04	6,74	2,99	3,13	15,4	7,47	5,9	2,5	2,54
ryazan	1246	51089	2,58	6,61	2,62	2,58	13,8	7,07	6,31	3,83	2,41
sakhalin	27	59474	2,37	7,21	2,37	2,58	17,9	9,15	5,75	2,8	2,78
samar	1371	407033	3,39	7,88	3,5	3,44	991	8,99	7,44	2,48	2,57
saratov	2148	32290	2,52	5,89	2,63	2,59	10,3	9,8	5,79	2,42	2,35
smol	1909	65523	2,73	6,33	2,62	2,8	8,71	10,9	6,64	2,52	2,35
stavrop	991	27658	2,5	6,16	2,42	2,47	7,59	9,7	5,57	2,4	2,45
sverdl	2189	274849	3,79	7,53	3,76	4,09	297	12,4	8,5	2,55	2,57
tambov	950	11522	2,43	6,66	2,38	2,37	6,55	6,81	5,5	2,35	2,39
tatar	2481	69856	2,87	6,03	2,78	2,74	20,8	6,95	6,14	2,54	2,48
tomsk	1507	88581	2,68	5,99	2,68	2,8	33	6,19	5,77	2,46	2,48
tul	998	48851	2,58	6,26	2,5	2,65	8,77	7,53	5,86	2,46	2,39
tumen	1696	107134	2,76	6,89	2,78	2,96	52,8	7,9	6,59	2,32	2,48
tver	2146	180355	5,1	9,39	5,11	5,63	89,3	10,8	10	2,72	2,78
tyva	2782	41649	2,81	7,25	2,67	2,87	9,98	7,03	6,52	3,92	2,42

udmurt	1552	34089	2,55	6,13	2,57	2,6	10,1	8,81	5,82	2,55	2,36
ulyan	1288	24085	3,05	7,4	3	3,2	8,21	7,34	6,9	2,53	2,52
vladimir	1216	101163	2,73	6,06	2,63	2,81	20,7	10,2	6,81	2,51	2,45
volgograd	1705	63547	2,55	5,97	2,49	2,59	25	9,98	5,88	2,46	2,4
vologda	2510	197257	3,06	6,55	3,21	3,34	222	8,1	6,88	2,47	2,55
voronezh	1286	90157	2,64	6,01	2,52	2,6	46,8	7,71	5,91	2,52	2,52
yakut	9350	590590	14,1	18,1	14,1	15,4	1980	24,2	20,4	2,7	2,67
yamal	4134	166419	4,28	7,58	4,28	4,91	152	10,5	8,22	2,46	2,47
yarosl	1133	201458	2,8	6,13	2,71	2,99	195	8,69	6,01	2,45	2,47
zabaikal	6361	62449	3,55	6,75	3,48	3,46	21,4	7,84	5,88	2,44	2,43
Всего			<b>284,76</b>	<b>594,73</b>	<b>282,55</b>	<b>294,3</b>	<b>18344,89</b>	<b>733,41</b>	<b>545,47</b>	<b>212,51</b>	<b>203,37</b>

[Обсудить в форуме](#) Комментариев — 2

## Ссылки по теме

- [Данные OpenStreetMap по регионам РФ в форматах shape и OSM XML](#)
- [Файлы регионов обрезки POLY и их создание в QGIS](#)

Последнее обновление: April 30 2011

Дата создания: 22.11.2009

Автор(ы): [Максим Дубинин](#)