

ДОПУСТИМОСТЬ ПОСТРОЕНИЯ ВЫЧИСЛИТЕЛЬНОГО КЛАСТЕРА ДЛЯ БЕЛОК-БЕЛКОВОГО ДОКИНГА ИЗ СТАРЫХ ANDROID-УСТРОЙСТВ

FEASIBILITY OF A COMPUTING CLUSTER BASED ON OLD ANDROID DEVICES FOR PROTEIN-PROTEIN DOCKING

А. Д. Казмерчук^{1, 2}, С. Шахаб^{1, 2}

A. D. Kazmerchuk^{1, 2}, S. Shahab^{1, 2}

¹Белорусский государственный университет, МГЭИ им. А. Д. Сахарова БГУ, г. Минск, Республика Беларусь

²Учреждение образования «Международный государственный экологический институт имени А. Д. Сахарова» Белорусского государственного университета, МГЭИ им. А. Д. Сахарова БГУ, г. Минск, Республика Беларусь
bantkaz566@gmail.com, siyamakshahab@mail.ru

¹Belarusian State University, BSU, Minsk, the Republic of Belarus

²International Sakharov Environmental Institute of Belarusian State University, ISEU BSU, Minsk, the Republic of Belarus

Кратко рассмотрены экономические, экологические аспекты, аспекты масштабирования и быстродействия кластерной системы, построенной на старых устройствах под управлением ОС Android, для массивно параллельной задачи белок-белкового докинга.

In this report we briefly evaluate economic, ecological, scaling and computing issues for a hypothetical computing cluster built from old Android devices, that we propose for protein-protein docking test running cluster.

Ключевые слова: вычислительная система, кластер, электронные отходы, андроид, протеин-протеиновый докинг.

Keywords: computing system, cluster, e-waste, Android, protein-protein docking.

<https://doi.org/10.46646/SAKH-2023-2-280-283>

Введение. Принимая во внимание то, что даже опыты 18-летней давности по массивному скринингу белок-белковых комплексов в вычислительных экспериментах по протеин-протеиновому докингу (ППД) проводились на суперкомпьютерах или кластерах с количеством рабочих арифметических логических вычислительных устройств (АЛВУ), исчисляемых сотнями [1], становится ясно, что в современных реалиях, если целью стоит построение действительно быстро функционирующей докинг-системы, количество АЛВУ в них должно быть не меньшим.

Доступ в Беларуси к суперкомпьютерным системам ограничен. Для проведения лишь одного опыта в рамках суперкомпьютера БГУ, требуется предварительная регистрация, работа только через интерфейс *PuTTY*, что вызывает неудобства, так как работа через удалённый терминал не даёт полного контроля над вычислительной системой. Такие опыты одинарны, что полностью побеждает само научное представление эксперимента (научный эксперимент в идеале должен с легкостью повторяться в одинаковых условиях многократно), еще также и потому, что программному обеспечению обычно требуется множество багфиксов. Исправлять ошибки ПО, используя описанный подход, превращается в неоправданно трудоёмкую процедуру.

Построить новый суперкомпьютер для исследователя не представляется возможным. Суперкомпьютеры и их построение требуют огромных финансовых затрат [2]. Само построение суперкомпьютера – это отдельная нетривиальная область информатики и инженерии с ненулевым порогом входа. Таким образом, путь построения собственного суперкомпьютера для исследователя не доступен.

Вторым решением может быть использование облачных вычислений. Однако данный путь также сопряжён с крупными финансовыми затратами на аренду облачной инфраструктуры, и эти затраты могут меняться по желанию поставщика облачной инфраструктуры. А для проведения действительно массивных экспериментов по докингу, затраты на масштабирование существенно возрастают.

Именно поэтому в среде энтузиастов, увлечённых различными видами массивно параллельных вычислений (майнинг, промптинг и обучение нейросетей) стандартным решением является вычислительная автономия на собственной видеокарте (ГП), высокопроизводительных ЦП или на специализированных вычислительных процессорах. Стоимость таких ЦП и ГП также исчисляется обычно сотнями белорусских рублей, или, если удобнее, сотнями и тысячами условных единиц. Например, стоимость 1 многоядерного (от 16 до 64 ядер) процессора *AMD EPYC*, часто используемых в суперкомпьютерах, начинается приблизительно с 1000 условных единиц.

Авторы предлагают иной подход, исходя, прежде всего, из экономической и экологической стороны вопроса, суть которого заключается в построении вычислительного кластера или его прототипа на старых и ненужных смартфонах на ОС *Android* (Андроид, андроид).

Чтобы понять, насколько это допустимо в целом, нужно рассмотреть следующие вопросы:

621892432. Какие попытки подобных построений наблюдались в прошлом?

621892433. Является ли уровень вложений в такой кластер допустимым

621889312. Как построение такого кластера сказывается на экологии?

621891792. Насколько высокопроизводительной будет данная система

621892352. Какие навыки требуются для масштабирования данного решения? Насколько просто масштабировать такое решение.

Попытки построения таких кластеров в прошлом. В основном, идея построения многопроцессорных систем на старых смартфонах родилась из юзерленда (*user land*), пользовательского пространства, и *DIY*-культуры, когда пользователи мобильных устройств стали обращать внимание на то, что, со временем, в их домовладениях накапливается большое количество электронного мусора (*e-waste*), включая множество старых, но в том числе и рабочих смартфонов.

Основных путей для избавления от данных устройств четыре.

1. Устройства можно выбросить. По личному опыту, это, вероятно, превалирующее решение в обществе. Более 5 миллиардов смартфонов ежегодно по всему миру оказываются на свалке [5].

2. Также распространен вариант 2 – оставить устройство лежать на полке – потому что и выкинуть жалко, и применить некуда.

3. Меньшинство пользователей прибегает к варианту 3, пытаясь выручить хоть какие-либо деньги с перепродажи таких телефонов. Нередко цены на такие телефоны составляют от 1 белорусского рубля, а также можно встретить предложения об отдаче таких телефонов даром.

3а. Отдача телефонов даром может служить как подвариантом варианта 2, так и подвариантом варианта 3 – либо владелец хотел бы продать устройства, но ясно понимает, что время, потраченное на продажу, дороже, чем вероятная стоимость продажи, либо он хочет отдать устройство тем, кому оно может пригодиться.

4. Четвертым вариантом избавления от старых устройств является нишевый вариант: переназначение старых андроид-устройств. Вариантов переназначения устройств – множество, например, при быстром поиске в поисковой системе *Google* тут же можно найти попытки построения *Plex*-сервера на старых смартфонах фирмы *SAMSUNG*, попытки использования старых смартфонов вместо программаторов в программировании умного дома, мобильный почтовый сервер, аналог *RAID*-массива, и множество других. Чтобы оценить множество вариантов переназначения старых андроид-смартфонов (ПСАС), можно воспользоваться системой *Google*, и запросом «*old smartphone repurposing*».

Сложно сказать, откуда эта идея появилась в юзерленде. Среди вероятных вариантов происхождения ПСАС авторы предполагают накопление критической массы смартфонов у пользователей, или же подорожание одноплатных компьютеров *Raspberry PI*, что вынудило сообщество разработчиков и энтузиастов *Raspberry PI* искать альтернативные аналогичные системы.

ПСАС в академической среде. Единственным экспериментом по кластерингу андроид-смартфонов, о котором можно найти упоминания в сети Интернет, является эксперимент группы Себастиана Шильдта, Феликса Бюшинга и Ларса Вольфа из Политехнического Университета Брауншвайга, Германия, под названием *DroidCluster* [3]. Данный эксперимент был проведен в 2012 году. Судя по сохранившимся слайдам презентации, которые данная группа предоставила на конференции *PhoneCom* в 2012 году в Китае, эта идея возникла у них больше как развлечение, чем серьезное мероприятие.

Во время эксперимента группа соединила свои шесть смартфонов в кластер в локальной сети. 4 из 6 смартфонов были одноядерными, с частотой менее 1 ГГц (чипы *MTK*), 1 смартфон был двухядерным на чипе *Tegra 2* (частота около 1 ГГц x 2), а самый производительный смартфон содержал чип *TI OMAP* с двумя ядрами и частотой 1.2 ГГц на 1 ядро. Поверх системы андроид, исследователи установили виртуализированную версию ОС *Debian Linux*, установили пакет *MPI* на виртуализированный *Debian*, и провели некоторые вычисления, о сущности которых данных в их презентации не сохранилось. Система показала линейный рост мощности при масштабировании и весьма умеренные 188.3 МФлопс. Однако группа сделала вывод о перспективности данного направления, особенно акцентировав внимание на том, что современные (на 2012 год) андроид-смартфоны более производительны, чем настольные компьютеры 2002 года.

Полученный мини-кластер исследователи назвали *DroidCluster*. Признавая все предшествующие данному тексту наработки в данной области, авторы предлагают потенциально рассматриваемый кластер назвать *ULDC* (*UserLand + DroidCluster*), или просто *DroidCluster* (ДроидКластер).

Следует отметить, что мотивацией группы Шильдта, Бюшинга, Вольфа послужило их неудовлетворение относительно времени простаивания (времени простоя) типичного пользовательского андроид-смартфона в карманах пользователей в то время, как эти устройства, по их мнению, обладают значительными вычислительными характеристиками.

Мотивация авторов данного тезиса другая. Если в 2012 году смартфоны общего назначения на системе андроид представляли из себя новинку, и исследователи пытались осознать их вычислительный потенциал (в духе

проходивших на то время массовых экспериментов волонтерских вычислений под общим суффиксом **@HOME*), то авторов интересует то, что через 10 лет после идеи о ДроидКластере, современная масса андроид-смартфонов представляет из себя потенциальный или уже существующий электронный мусор (*e-waste*), который никуда не используется.

Что представляет из себя типичный современный андроид-смартфон. С точки зрения данного текста, авторы будут понимать под современностью 2017 год, так как телефоны приблизительно из этого года на этап написания в 2023 году активно пополняют фонд электронного мусора.

Современный на то время андроид-смартфон – это устройство с 2 – 8-ядерным процессором общего назначения на архитектуре ARM или x86 (*Intel Atom*), в некоторых случаях оснащенное дополнительно графическим процессором (ГП) общего пользования на несколько ядер, имеет около 16 Гб флеш-памяти для хранения и около 2–4 Гб оперативной памяти. Следует отметить, что новопроизведенные ноутбуки для пользовательских задач в 2023 году имеют сравнимую конфигурацию.

Другими словами, современный на то время телефон представляет из себя высокопроизводительный компьютер общего назначения.

Вопрос об андроид-планшетах. В целом, андроид-планшеты зарекомендовали себя, как андроид-телефоны с менее удобным форм-фактором, и также активно пополняют фонд электронного мусора на момент написания. Нередки случаи, когда андроид-планшеты выбрасываются в мусор или продаются оптом за символическую цену как нерабочие. Как показывает практика, многие из этих планшетов на самом деле являются рабочими, но требующими специальных процедур для зарядки «глубоко» разряженной литиевой батареи. Как и андроид-смартфоны, планшеты 2010-х годов представляют из себя высокопроизводительные компьютеры общего назначения. Как и андроид-смартфоны, андроид-планшеты используют ОС Андроид, а это значит, что андроид-планшеты можно также использовать для расширения предлагаемого ДроидКластера.

Куда в Беларуси идёт электронный мусор? По опыту авторов, пользователи смартфонов, которые предпочитают выбросить свой старый смартфон, не сортируют этот мусор отдельно, и он отправляется на свалки, где закапывается. Несмотря на то, что в Беларуси существуют частные компании, стремящиеся сократить количество электронного мусора (электронных отходов), а также несмотря на то, что Беларусь занимает лидирующее место в СНГ по переработке электронных отходов, распространены ситуации, когда такой мусор не сортируется владельцами, а оказывается на свалках [4,5]

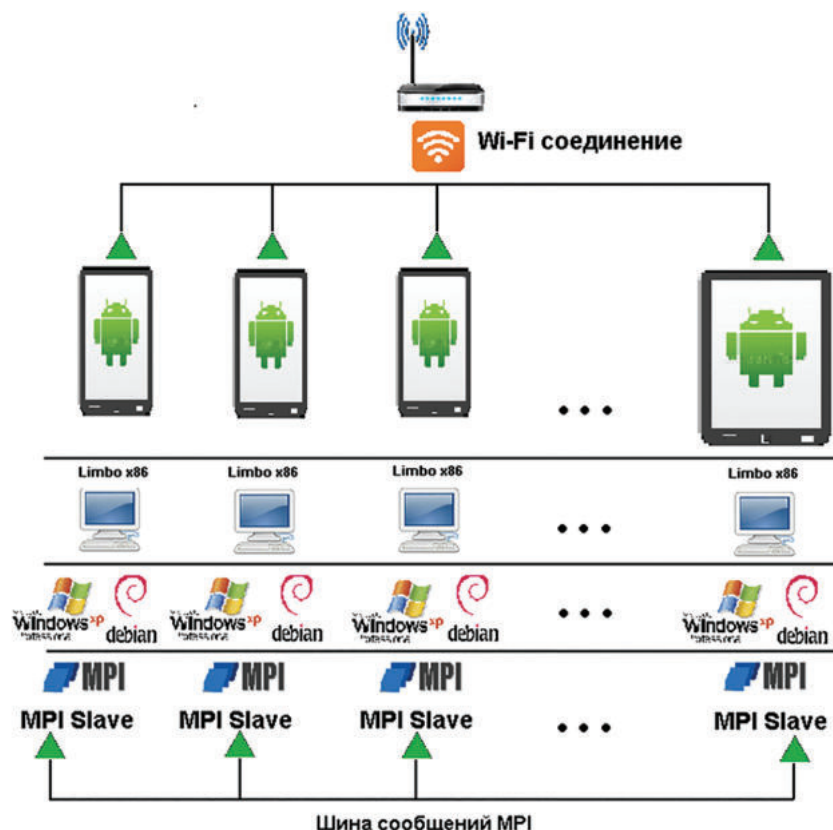
Вопрос о допустимости iOS-устройств в предлагаемый кластер. Система *iOS*, представленная на устройствах *iPhone* и *iPad* фирмы *Apple* бинарно несовместима с системой *Android*. Это значит, что для масштабирования ДроидКластера за счёт устройств *Apple* потребуется отдельное ПО. Это увеличивает издержки на содержание кластера. Более того, как показывает практика, устройства *Apple* имеют дополнительную неутилитарную функцию – являются предметами статуса – и, как следствие, реже пополняют фонд электронного мусора.

Предлагаемая архитектура кластера. Для системы Андроид существуют портированные программные средства виртуализации из систем *Linux* и *Windows*. Предлагается развернуть на устройствах андроид ПО *Limbo x86* с установленной на них нетребовательной к ресурсам ОС, для которой существует портированная версия библиотеки для распределенных вычислений *MPI* (например, *Windows XP* или *Debian Linux*). Далее, следует учитывать, что в *Limbo x86* есть опция настройки *WLAN*-сети хост-устройства (то есть андроид-устройства) в качестве *LAN/Ethernet*-сети виртуальной машины *Limbo x86*. Далее следует увязать все полученные виртуальные устройства через *WLAN*-сеть *Wi-Fi*-роутера, реализовать распределенное *MPI*-приложение и применить его на полученной сети. Примерная архитектура ДроидКластера изображена на иллюстрации 1.

Масштабируемость системы. Для того, чтобы добавить дополнительное АЛВУ на базе ОС Андроид в ДроидКластер, требуется установить на него *Limbo x86*, перенести работающий образ системы с *MPI* с уже настроенного смартфона из кластера, и включить его во *WLAN*-сеть ДроидКластера. Таким образом, масштабирование системы тривиально.

Быстродействие. Исходя из того, что не существует общей теории расчета быстродействия ПО на современных АЛВУ общего назначения, из-за их большого разнообразия, авторы отмечают, что для массивно распараллеливаемых задач вроде макромолекулярного докинга, решающую роль играет количество АЛВУ на кластер. Исходя из этих рассуждений, ДроидКластер может как служить отдельным вычислительным кластером, так и быть вычислительным дополнением к уже существующему *MPI*-кластеру.

Выводы. Авторы предлагают построение вычислительного кластера на старых андроид-смартфонах, который является адекватным решением согласно критериям экономичности, экологичности и масштабируемости, а также является относительно допустимым решением по быстродействию при написании алгоритмов массивно параллельных задач, в том числе задачи макромолекулярного докинга.



ЛИТЕРАТУРА

1. Andrey Tovchigrechko, Ilya A. Vakser: GRAMM-X public web server for protein–protein docking. *Nucleic Acids Research*, Volume 34, Issue suppl_2, 1 July 2006, Pages W310–W314, <https://doi.org/10.1093/nar/gkl206> [Электронный ресурс]. Режим доступа: https://academic.oup.com/nar/article/34/suppl_2/W310/2505594
2. Shaun Franks, Johnathan Yerby: CREATING A LOW-COST SUPERCOMPUTER WITH RASPBERRY PI. Middle Georgia State College, 2014 [Электронный ресурс]. Режим доступа: https://www.researchgate.net/publication/330855193_Creating_a_Low-cost_Supercomputer_with_Raspberry_PI
3. Felix Buesching, Sebastian Schildt, Lars Wolf: DroidCluster: Towards Smartphone Cluster Computing - The Streets are Paved with Potential Computer Clusters [Электронный ресурс] / 2012, DOI:10.1109/ICDCSW.2012.59. Conference: 32nd International Conference on Distributed Computing Systems Workshops (ICDCS 2012 Workshops). Режим доступа: https://www.researchgate.net/publication/230851531_DroidCluster_Towards_Smartphone_Cluster_Computing_-_The_Streets_are_Paved_with_Potential_Computer_Clusters
4. Репортаж о сборе и переработке электроники в ООО «БелВторСырье» «Новое утро» на РТР-Беларусь [Электронный ресурс] / РТР-Беларусь, СТБ, Минск, 2020. Режим доступа: <https://www.youtube.com/watch?v=5I9dDkTXU6U>
5. Безмусорное будущее [Электронный ресурс] / АТН, Минск, 2023. Режим доступа: <https://www.youtube.com/watch?v=2uG0sx1DzQs>