

# Отказоустойчивость коммуникационных сетей со свойствами малого мира

А. Демичев, В. Ильин, А. Крюков, С. Поляков

*НИИЯФ МГУ и НИЦ “Курчатовский институт”*

*Национальный Суперкомпьютерный Форум*

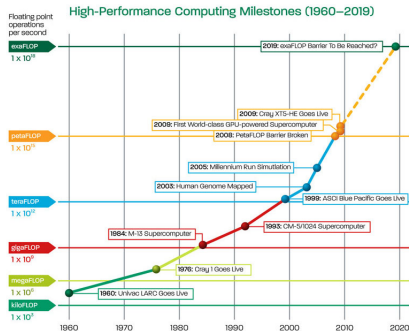
*НСКФ-2013*

*Россия, Переславль-Залесский,  
ИПС имени А.К. Айламазяна РАН,  
26-29 ноября 2013 года*

*Работа частично финансируется РФФИ, грант 12-07-00408-а*

# Суперкомпьютеры следующего поколения

- ▶ Появление реальных вычислительных систем следующего поколения - с производительностью порядка эксафлопса ( $10^{18}$  FLOPS) - ожидается не ранее 2018–2020 года
- ▶ однако принципы их построения начинают интенсивно разрабатываться уже сейчас
- ▶ на пути к их построению предстоит решить ряд сложных научно-технических задач
- ▶ выработать принципиально новые решения для их архитектуры и аппаратной реализации.



# Суперкомпьютеры следующего поколения: коммуникационная сеть

- ▶ Коммуникационная сеть - одна из важнейших составляющих любого суперкомпьютера
  - ▶ в первую очередь определяет возможность увеличения числа вычислительных узлов, что необходимо для достижения желаемой производительности
- ▶ Как ожидается, коммуникационная сеть суперкомпьютеров экзафлопного уровня будет объединять  $> 100\,000$  вычислительных узлов (“end points”)
- ▶  $\Rightarrow$  разработка коммуникационных сетей с хорошими свойствами масштабируемости и возможностью эффективно обслуживать огромное число вычислительных узлов



## В данной работе

- ▶ исследована устойчивость важнейших характеристик стохастических и регулярных (детерминистских) коммуникационных сетей *с малым средним расстоянием между узлами* при выходе из строя части узлов
  - ▶ сети “малого мира”
- ▶ в случае стохастических сетей используется алгоритм с оптимальными значениями числа перемычек и параметра распределения их длин
- ▶ в качестве регулярных сетей рассмотрена iBT-сеть (Interlaced Bypass Torus Networks), обладающая наилучшими характеристиками в классе сетей, построенных на основе детерминистских алгоритмов
- ▶ показано, что в широком диапазоне значений относительного числа вышедших из строя узлов рассмотренные сети являются весьма устойчивыми к выходу узлов из строя, причем iBT-сети ведут себя несколько лучше, чем стохастические сети

# Исходные принципы подхода

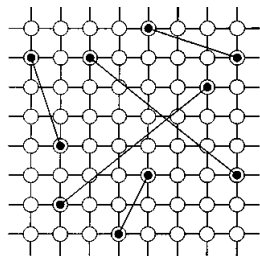
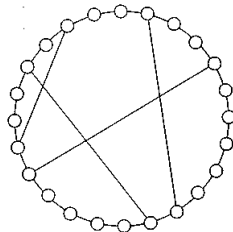
- ▶ Сеть должна обладать свойствами “малого мира”, а именно, медленным (логарифмическим) ростом среднего расстояния между узлами с ростом их числа
  - ▶  $\implies$  масштабируемость
- ▶ Сеть должна иметь базовую структуру  $D$ -мерной регулярной решетки
  - ▶ Например, для моделирования  $D$ -мерных объектов
- ▶ Эта базовая структура должна иметь дополнительные связи таким образом, чтобы могли достаточно оптимальным образом решаться и общие задачи  $\rightarrow$  малая средняя сетевая длина коммуникаций, оптимальная навигация и маршрутизация по сети
  - ▶ Желательно, чтобы характеристики новых связей зависели от варьируемых параметров для возможности оптимизации

# Стохастические решеточные сети с перемычками (РСП) - модификация алгоритма Уоттса-Строгача (Watts-Strogatz, 1998))

1. исходный объект -  $D$ -мерная регулярная решетка с топологией тора, содержащая  $N = L^D$  узлов ;
2. последовательно перебираются все узлы решетки и с вероятностью  $0 < p \leq 1$  к каждому узлу  $i$  подсоединяют первый конец перемычки;
3. второй конец перемычки выбирается случайно с вероятностью

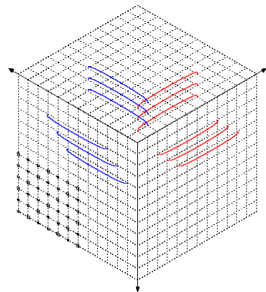
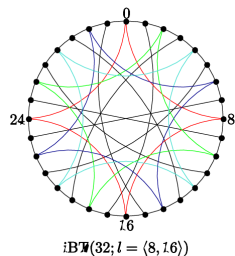
$$P(r) \sim r^{-\alpha}$$

- функция решеточного расстояния  $r$  между узлами

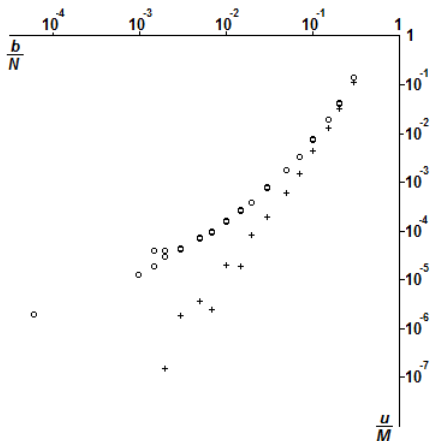


# Детерминистская сеть iBT (Interlaced Bypass Torus Networks) (P.Zhang, R.Powell, Y. Deng, 2010))

1. исходный объект -  $D$ -мерная регулярная решетка с топологией тора
2. для каждого узла вычисляются две характеристики:
  - ▶ номер измерения, вдоль которого направлены перемычки, выходящие из данного узла:
  - ▶ длина перемычек (из predetermined набора длин), которые выходят из данного узла;
3. к каждому узлу присоединяются две перемычки (в двух направлениях измерения) в соответствии с вычисленными характеристиками



## Устойчивость сетей при выходе из строя части узлов

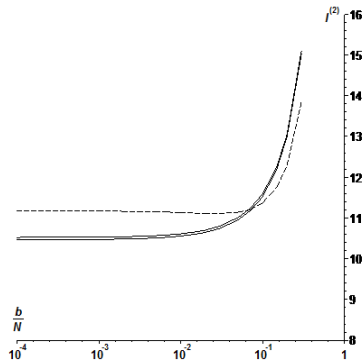


- ▶ Зависимость доли  $u/M$  недоставленных при выбранном алгоритме сообщений от доли вышедших из строя узлов (линейный размер всех сетей  $L = 128$ )
  - ▶ в стохастических сетях:  $\circ$
  - ▶ в сетях iVT:  $+$
- ▶ при относительно небольшой доле неисправных узлов в iVT-сетях оказывается меньше потерянных сообщений, а при достижении доли неисправных узлов  $\sim 10\%$  относительное число потерянных сообщений сравнивается для обоих типов сетей



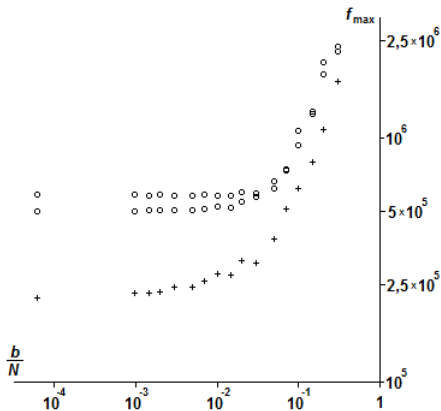
## Зависимость среднего навигационного расстояния $\ell^{(2)}$ от доли числа вышедших из строя узлов $b/N$

- ▶ в стохастических сетях при  $\alpha = 1$  (сплошные линии соответствуют лучшим образцам из ансамбля в 100 экземпляров)
- ▶ в iBT-сетях при  $s_1 = 8$ ,  $s_2 = 32$  (прерывистая линия)
- ▶ оба лучших экземпляра стохастических сетей и iBT-сеть ведут себя примерно одинаково
  - ▶ пока доля вышедших из строя узлов  $\lesssim 10\%$ , средняя навигационная длина практически не меняется
  - ▶ когда доля вышедших из строя узлов  $\gtrsim 10\%$ , средняя навигационная длина резко возрастает.



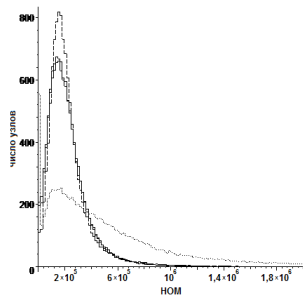
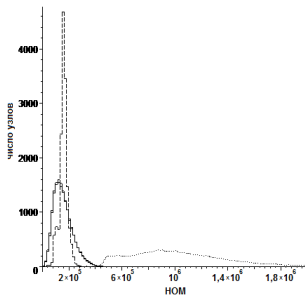
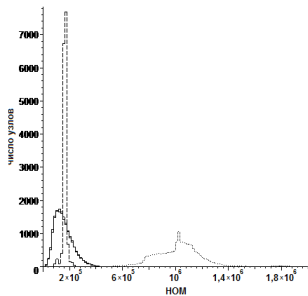
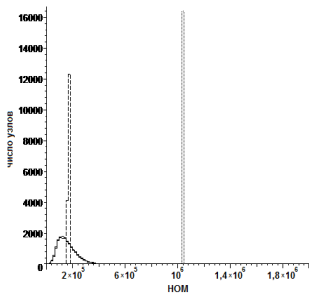
# Устойчивость максимальной нагрузки относительно маршрутизации (МНОМ; forwarding index) от доли выведенных из строя узлов

- ▶ стохастические сети: ○;  
iBT-сети: +
- ▶ до значений  $\approx 10\%$  в случае стохастических и iBT-сетей МНОМ практически не меняется (нагрузка для iBT-сетей несколько ниже)
- ▶ при достижении доли неисправных узлов  $\gtrsim 10\%$  максимальная нагрузка  $f_{max}$  резко возрастает
  - ▶ аналогично поведению средней навигационной длины



# Распределение нагрузок (1/2)

- ▶ стохастич. (сплошные линии), iVT (штрих), тор без перемычек (пунктир)
- ▶ слева направо и сверху вниз:  $b = 0, 0,01N, 0,05N, 0,3N$



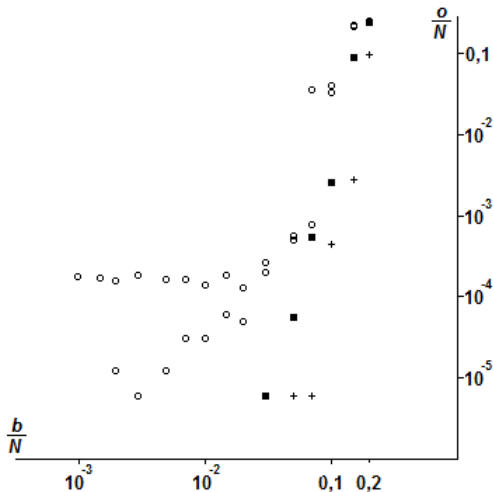
## Распределение нагрузок (НОМ) (2/2)

- ▶ стохастические сети — распределение нагрузки на узлы является достаточно широким
  - ▶ нежелательное свойство
  - ▶ фиксация величины степеней узлов в стохастических сетях практически не влияет на ширину распределения нагрузок, стохастичность расположения перемычек оказывается достаточной для того, чтобы узлы испытывали существенно разную нагрузку
- ▶ решетки без перемычек:
  - ▶ для полностью исправной решетки нагрузка одинакова, но существенно выше, чем для сетей малого мира (большое среднее расстояние)
  - ▶ при выходе из строя даже небольшой части узлов ширина распределения нагрузки резко увеличивается ⇒ значительная часть узлов может испытывать существенные перегрузки;
  - ▶ ⇒ сети малого мира не только обеспечивают меньшую временную задержку при передаче сообщений (малое среднее расстояние), но и являются существенно более устойчивыми при выходе из строя части узлов сети

## Устойчивость сетей при каскадных отключениях (1/2)

- ▶ доля перегруженных узлов от доли первоначально испорченных узлов

- ▶ стохастич.: ○,  
рег. стох. (степень 6): ■,  
iBT: +
- ▶ при проектировании коммуникационной сети маршрутизаторы выбираются с  $k$ -кратным запасом прочности по отношению к iBT-сетям:  
 $f_{th} = 3f_{max}^{(iBT)}$



## Устойчивость сетей при каскадных отключениях (2/2)

- ▶ Результаты показывают, что с точки зрения возможности каскадных отключений оба типа сетей ведут себя примерно одинаково:
  - ▶ до значений доли вышедших из строя узлов  $\sim 10\% \div 15\%$  каскадных отключений нет
  - ▶ выше этих значений происходит резкий рост перегруженных, и как следствие, вышедших из строя узлов
  - ▶ iBT-сети являются даже несколько более устойчивыми к каскадным отключениям, чем стохастические сети
  - ▶ фиксация величины степеней узлов в стохастических сетях не оказывает существенного влияния на каскадные процессы

## Заключение

- ▶ В целом результаты работы показывают
  - ▶ в отличие от регулярных решеток с топологией тора рассмотренные сети со свойствами “малого мира” в широком диапазоне значений относительного числа  $b/N$  вышедших из строя узлов сети являются весьма устойчивыми к выходу узлов из строя.
  - ▶ iBT-сети, построенные на основе детерминистского алгоритма, ведут себя несколько лучше, чем стохастические сети
- ▶ в данной работе рассмотрены только статические (топологические) характеристики сетей — среднее навигационное расстояние (то есть расстояние при выбранной маршрутизации) и нагрузка относительно маршрутизации
  - ▶ планируется исследовать динамические модели передачи сообщений в рассматриваемых коммуникационных сетях, в частности, на основе теории систем массового обслуживания
- ▶ Работа частично финансируется РФФИ, грант 12-07-00408-а