«Некоторые фантастические прогнозы будущего ИТ»

Непейвода Н. Н..
ИПС имени А.К.Айламазяна РАН

20.11.2012, Ижевск
«Былое и думы»

Прежде всего, вычислительные системы резко разойдутся на два класса: системы для орангов и системы для людей. Орангутанги успешно освоили общение с ipadами, обменивались изображениями и музыкой.

Люди решают задачи и делают дела, оранги самовыражаются и развлекаются. Говорить буду для людей.

Классы решаемых задач:

1. Вычисления и представление их результатов
2. Информационный поиск и отображение его результатов
3. Поиск решений и их отображение и модификция
4. Ведение и поддержка баз данных и связанных с ними процессов обработки данных
5. Высший дизайн и искусство

Что такое «высокопроизводительные вычисления» и «суперкомпьютер»?

- **Суперкомпьютер** (англ. supercomputer, СуперЭВМ) — вычислительная машина, значительно превосходящая по своим техническим параметрам большинство существующих компьютеров — **ru.wikipedia.org**
- Мировой рейтинг Top500: **www.top500.org**
- **Производительность** — число операций в секунду
- **Пиковая** и **реальная** производительность (на задаче …)
  - **КПД**
- Тест **Linpack**
  - Другие тесты и другие рейтинги: www.topcrunch.org, **www.graph500.org**

Что такое «высокопроизводительные вычисления» и «суперкомпьютер»?

- Суперкомпьютер == высокопроизводительная система
- Машина для самых сложных задач — значительно превосходит по техническим параметрам большинство существующих компьютеров:
  - много операций в секунду
  - много памяти
  - высокие скорости внутренних обменов данными (bandwidth, latency, massage rate)
- ***Быстрое развитие отрасли, закон Мура***
  - удвоение за ~18 месяцев — 1000 раз за 11 лет
  - Mflops (1975, 106), Gflops (1986, 109),
  - Tflops (1997, 1012), Pflops (2008, 1015)
  - Eflops (2019, 1018), Zflops (2040, 1021)

Роль и место суперкомпьютерных технологий

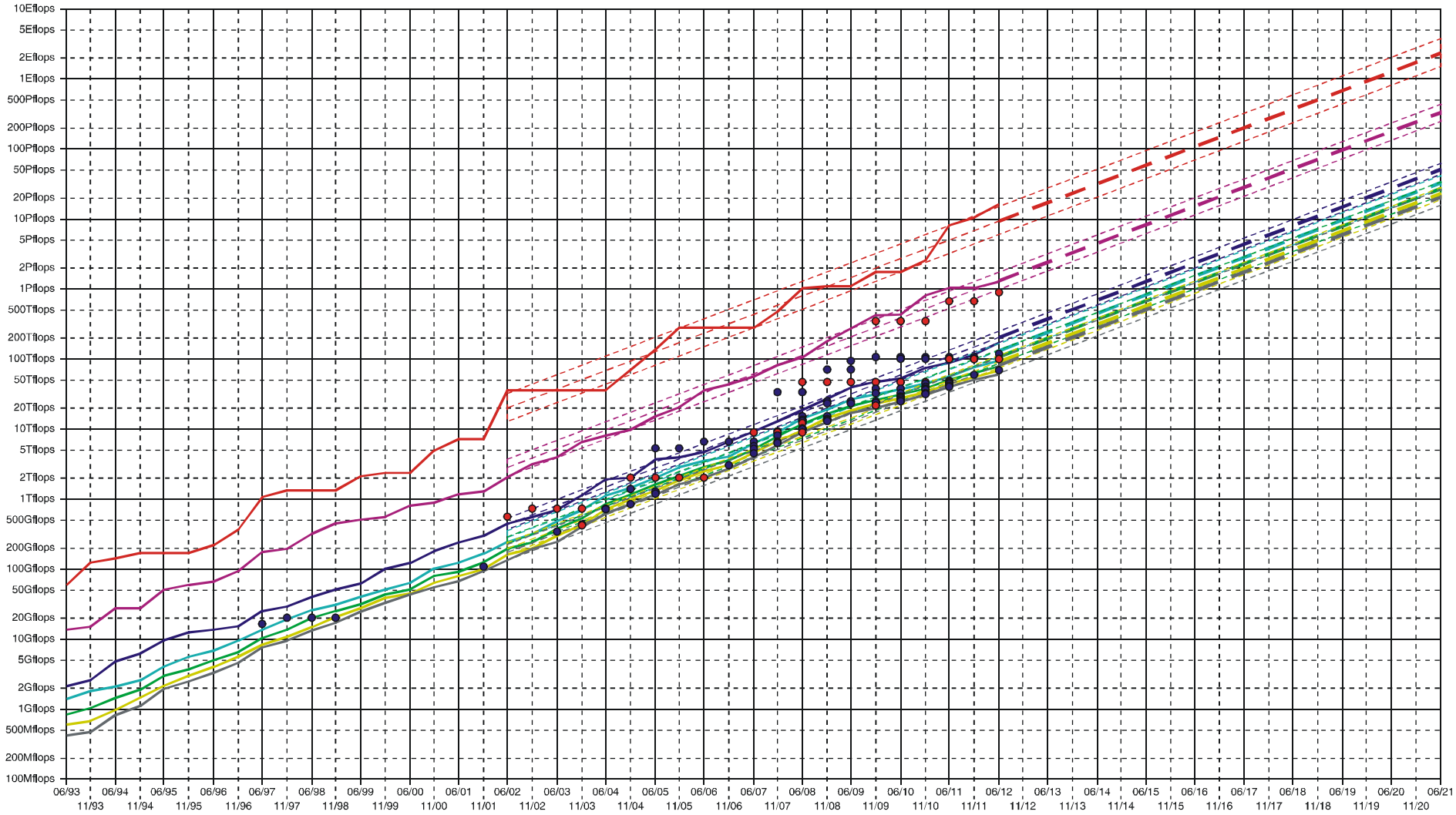
- «Технологии, таланты и деньги доступны многим странам. Поэтому США стоит перед лицом беспрецедентны зарубежных экономических конкурентов.
  Страна, желающая победить в конкуренции, должна победить в вычислениях».
  Дебора Винс-Смит, Президент Совета по конкурентоспособности США.
  4 Jun 2004, opening «Supercharging Innovation, Competitiveness: HPC Conference»

Вычислительные системы сверхвысокой производительности?

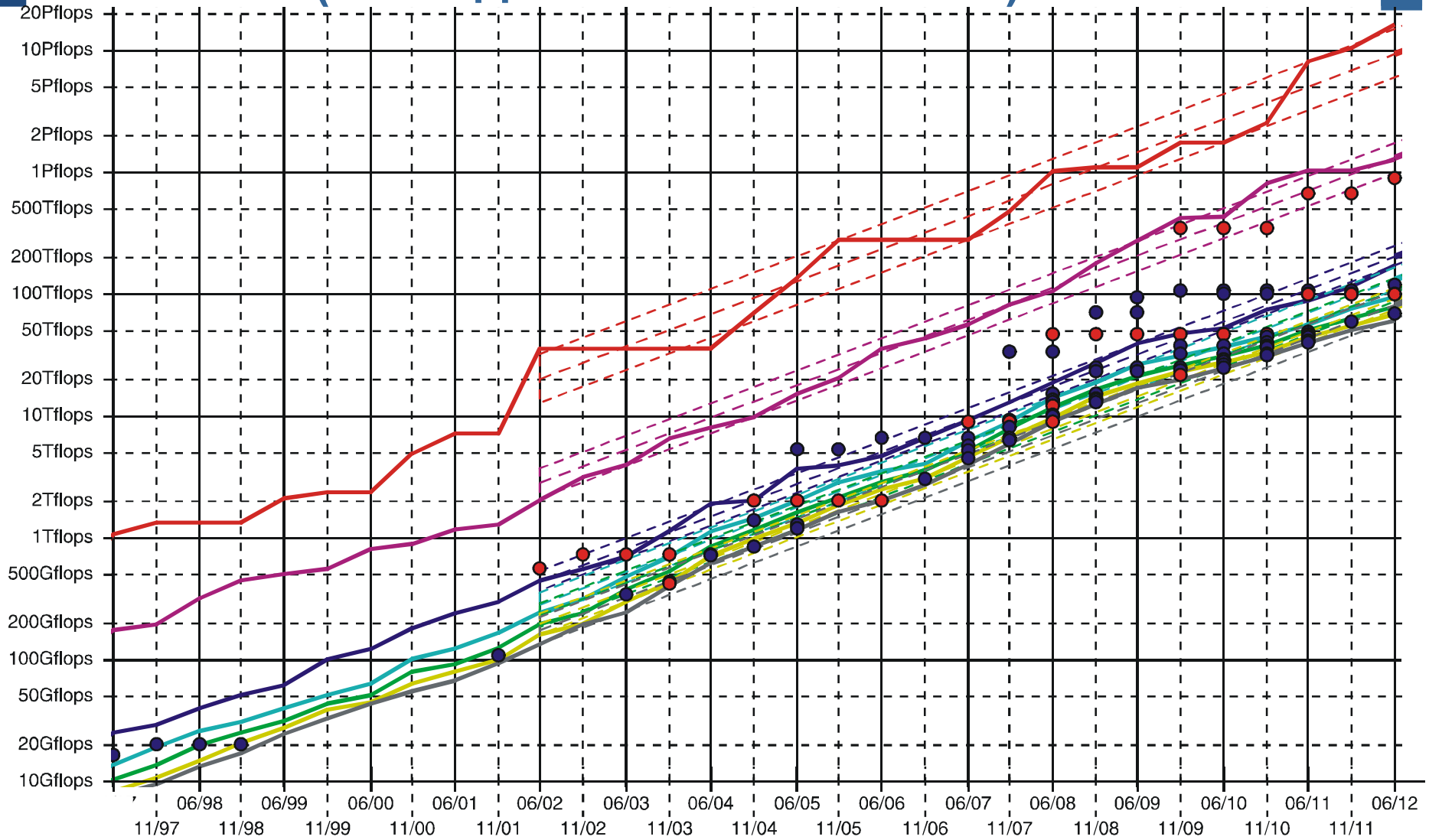
Вычислительная система высокой производительности (суперкомпьютер) — вычислительная машина, значительно превосходящая по своим техническим параметрам большинство существующих компьютеров — ***Top500***

- ***Вычислительная система сверхвысокой производительности*** — вычислительная машина, значительно превосходящая по своим техническим параметрам большинство существующих суперкомпьютеров. То есть, это суперкомпьютер, соответствующий уровню ***Top1–10***

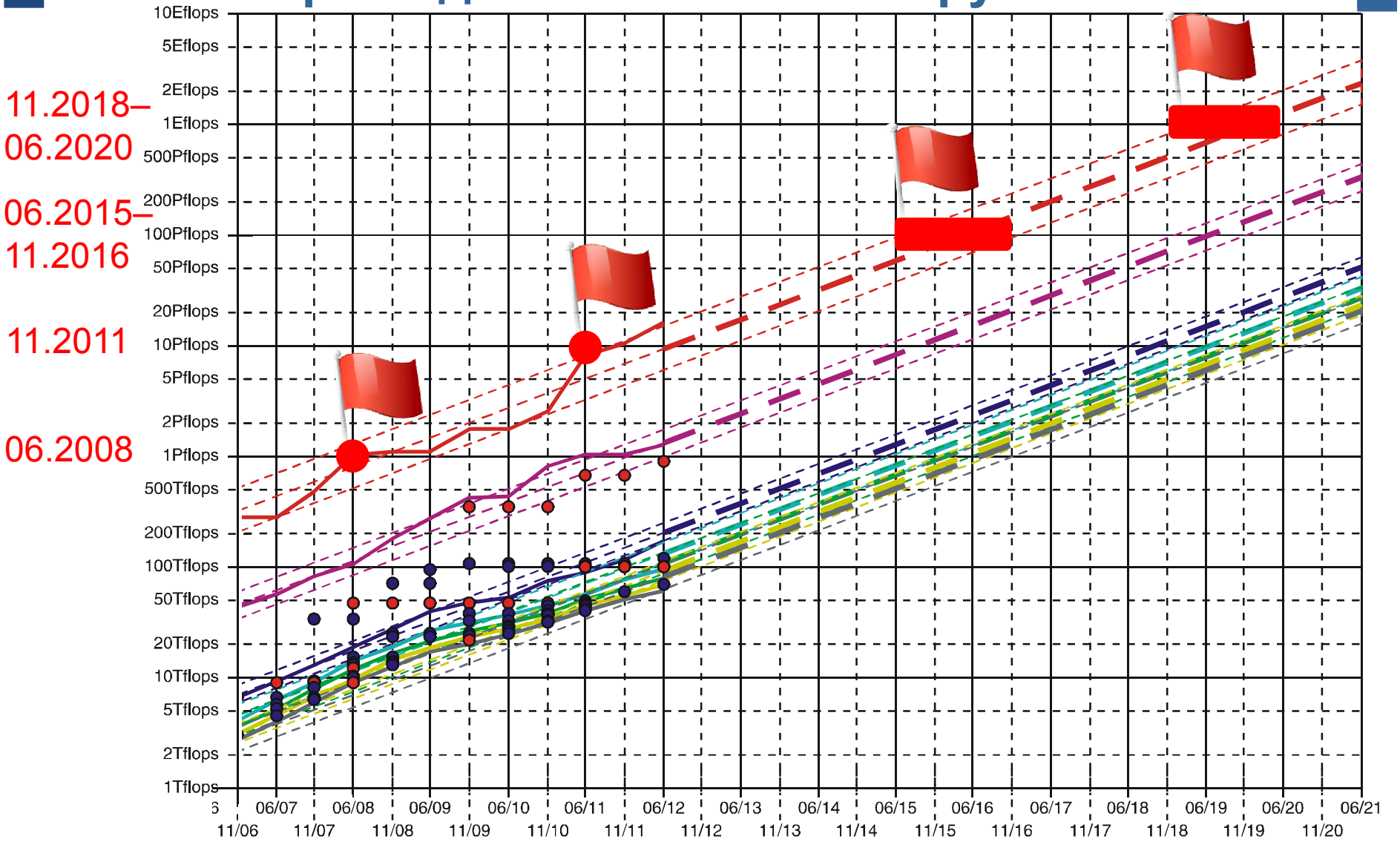
Развитие суперкомпьютерной отрасли в мире и в России (взгляд с птичьего полета)



Развитие суперкомпьютерной отрасли в России (взгляд с птичьего полета)

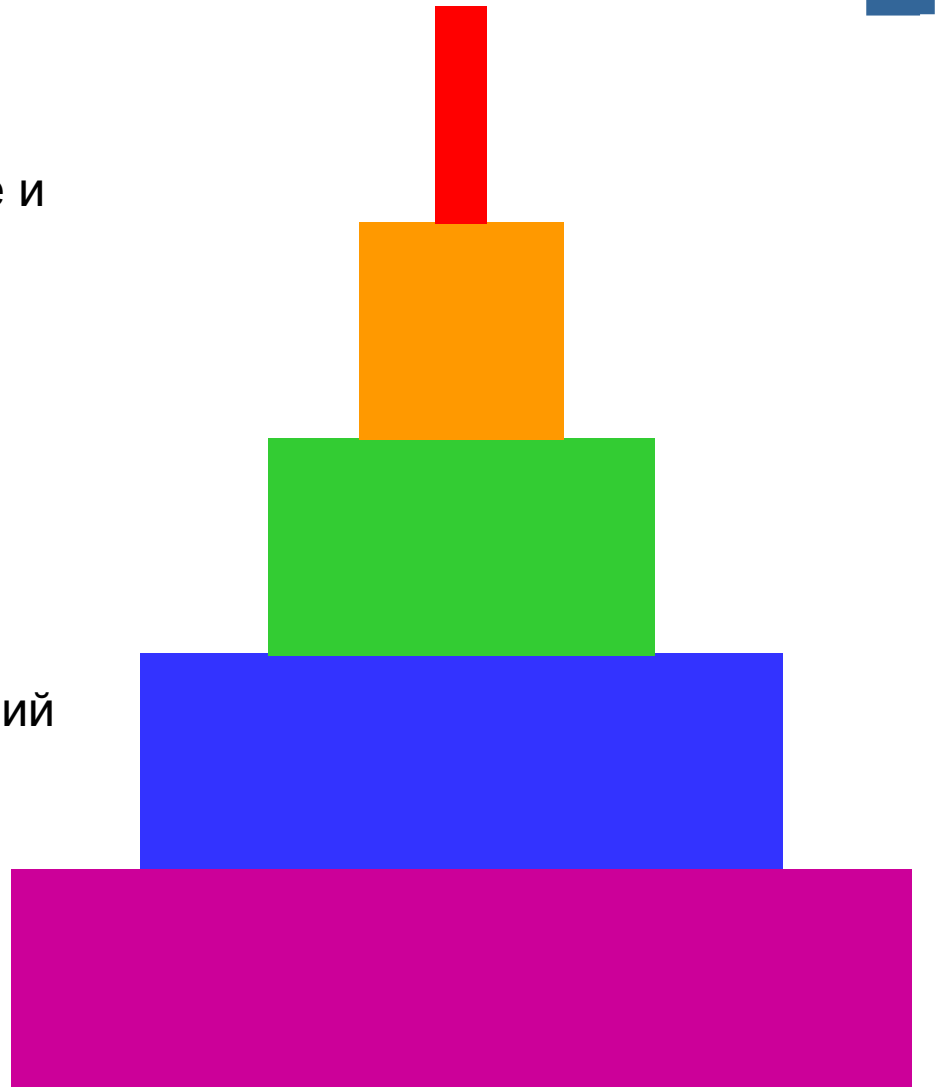


Развитие суперкомпьютерной отрасли
Планомерное движение к новым рубежам

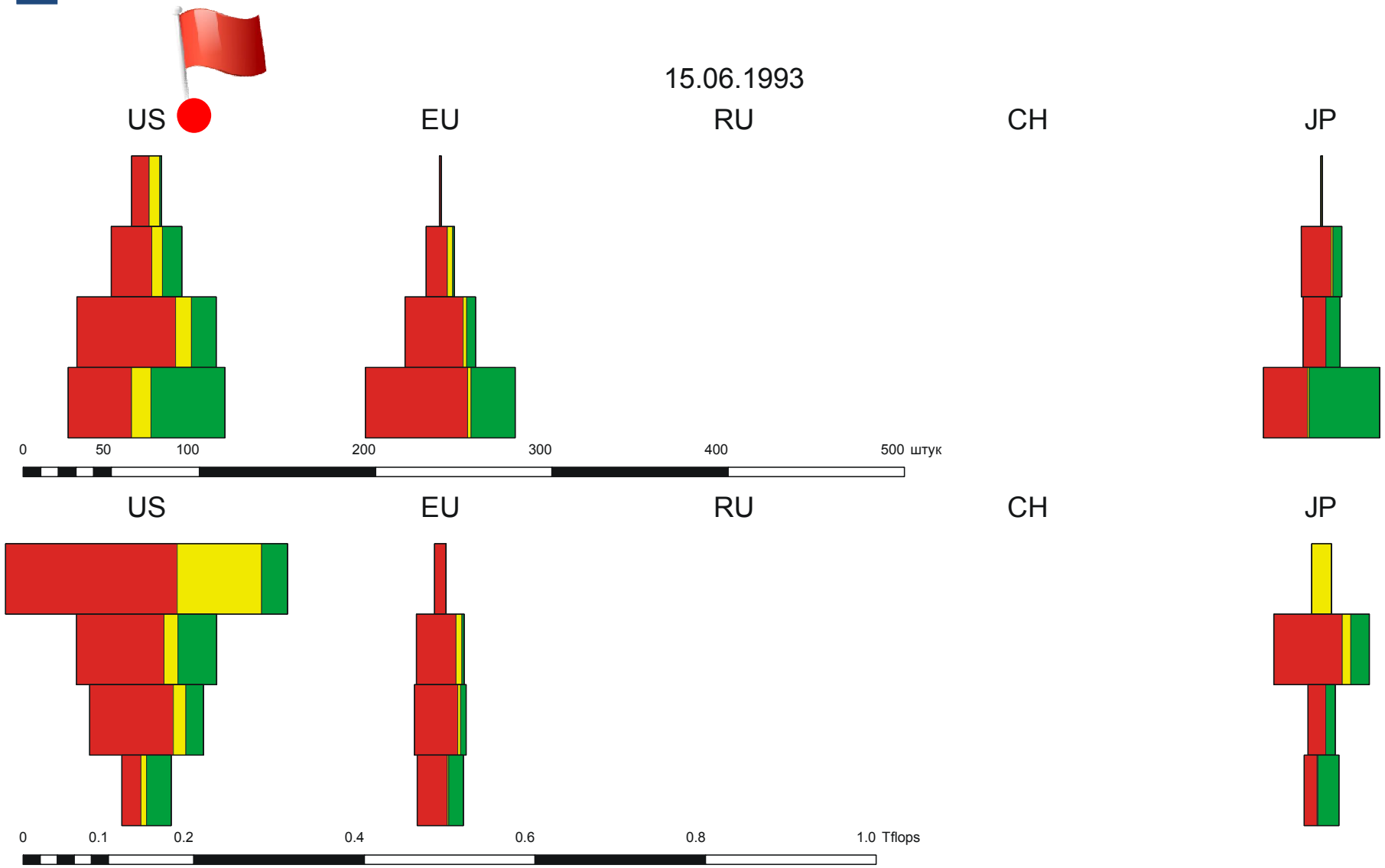


Иерархия киберинфраструктуры

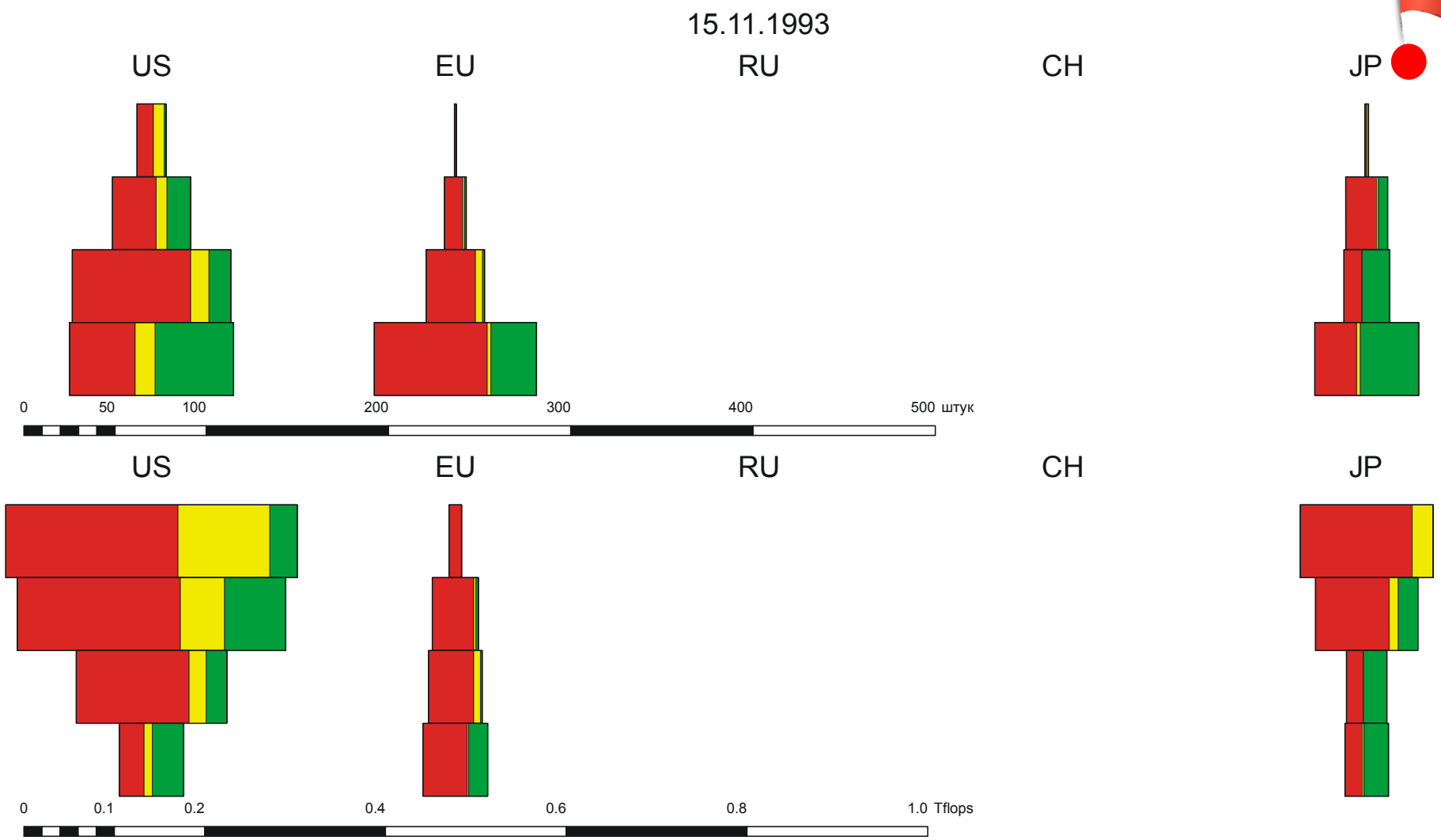
- **#1–#20** Крупнейшие национальные центры
- **#21–#100** Крупнейшие региональные и отраслевые центры
- **#101–250** Крупные региональные и корпоративные центры
- **#251–#500** Центры предприятий и научных учреждений
- СуперЭВМ небольших исследовательских компаний, лабораторий и научных подразделений



Киберинфраструктура ведущих стран

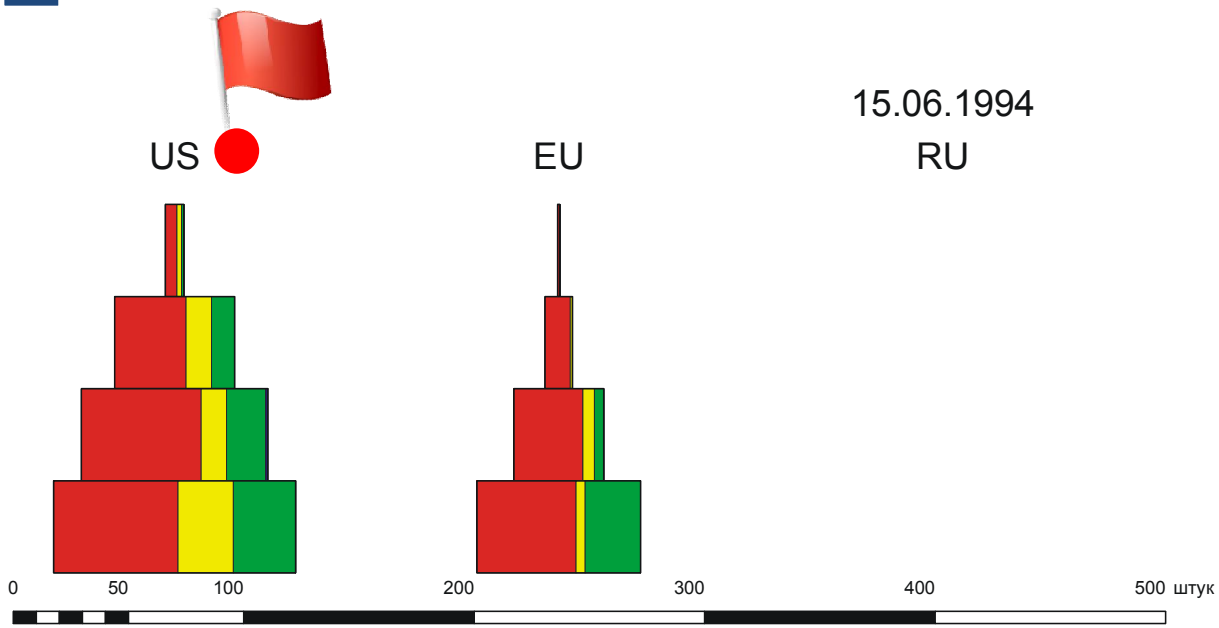


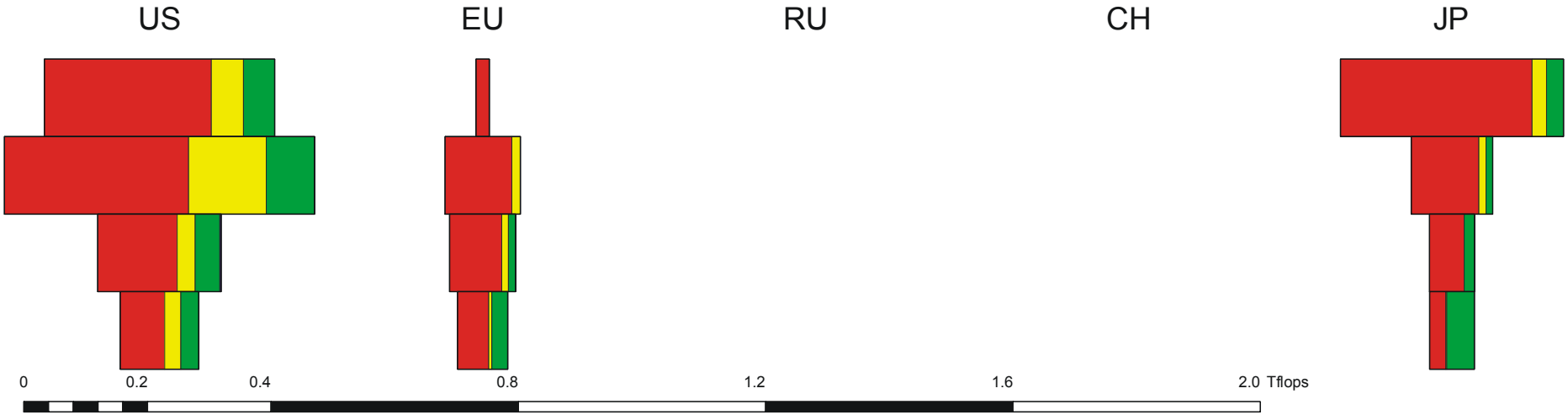
Киберинфраструктура ведущих стран



Киберинфраструктура ведущих стран

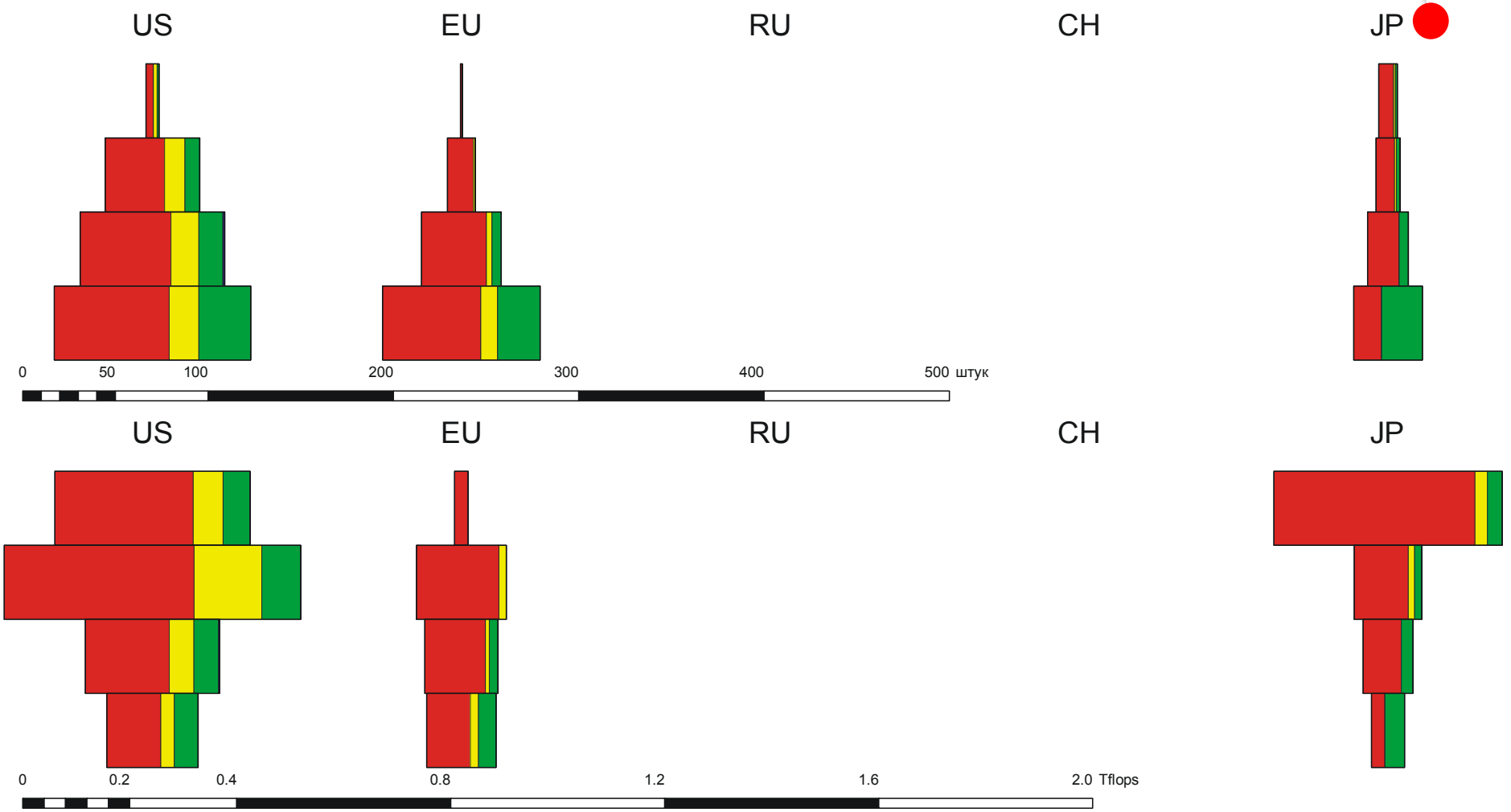
15.06.1994



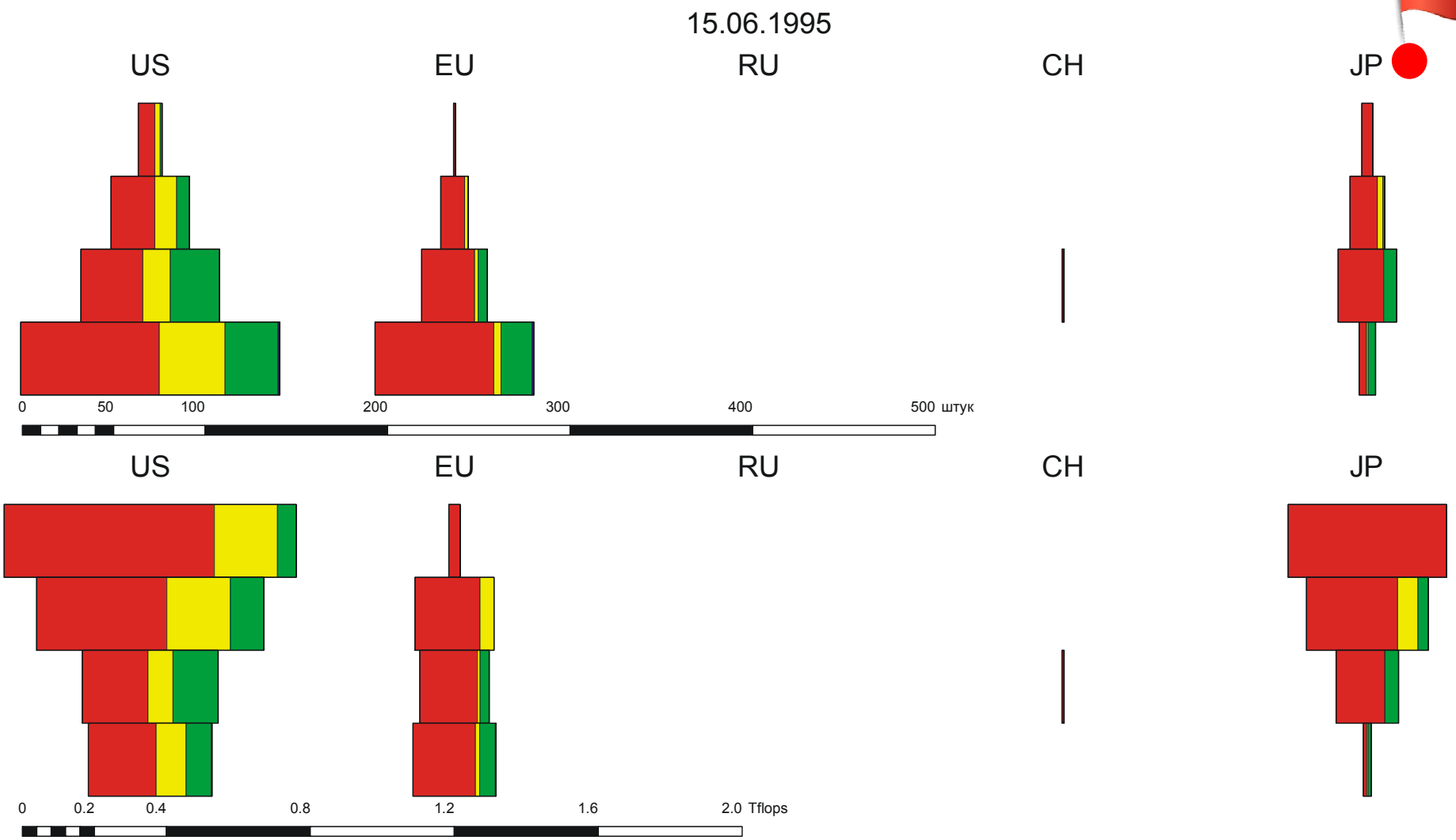


Киберинфраструктура ведущих стран

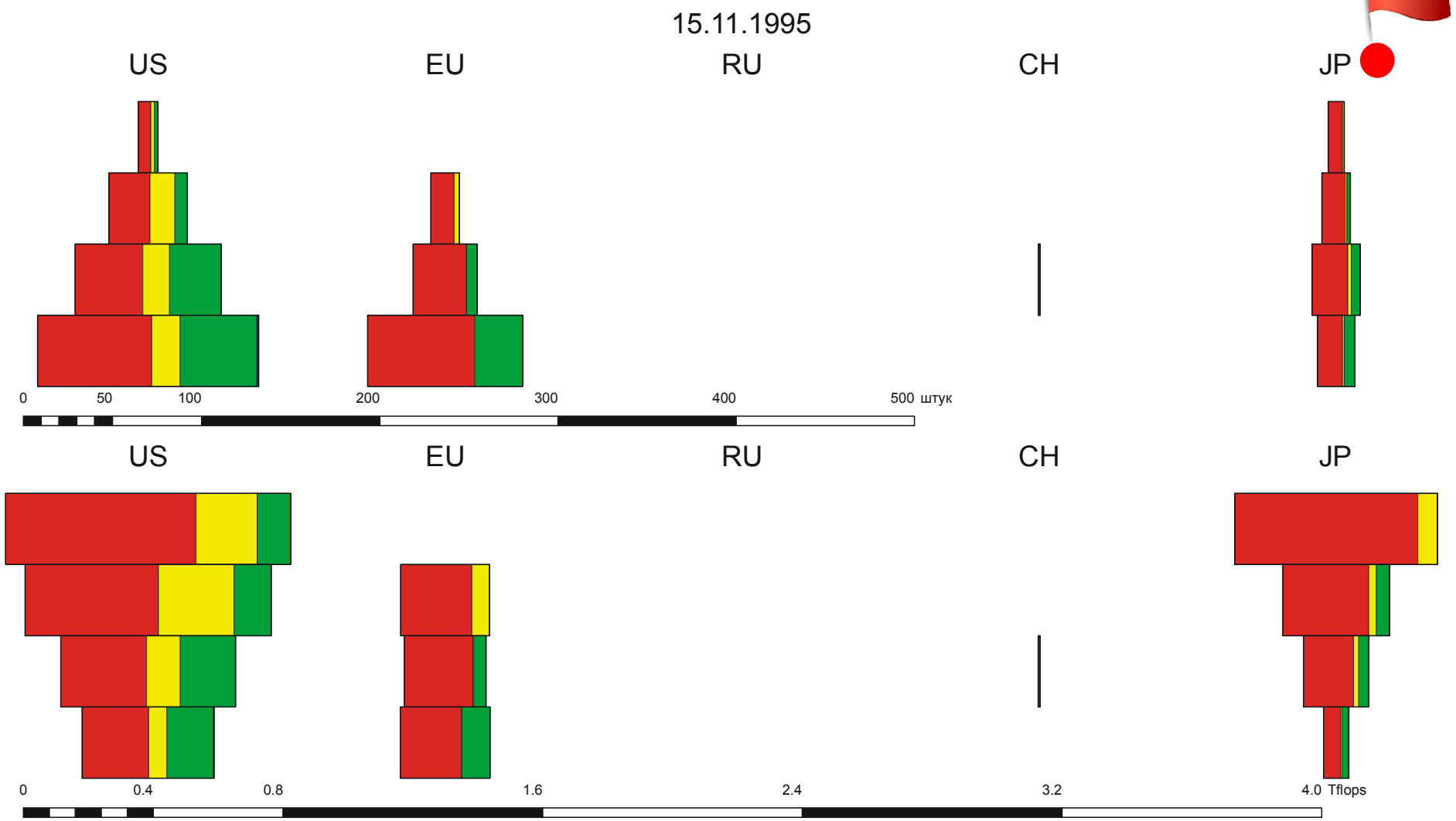
15.11.1994



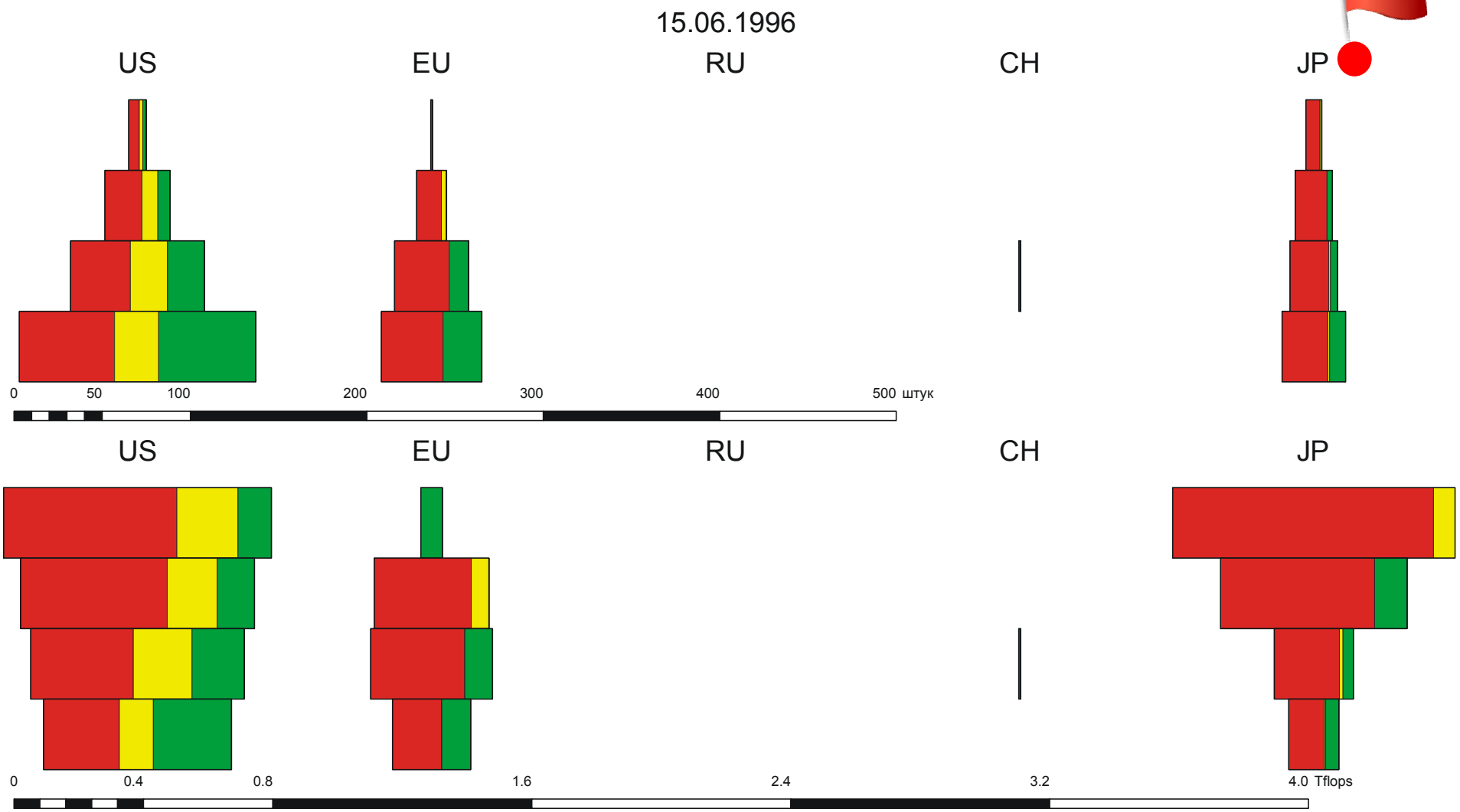
Киберинфраструктура ведущих стран



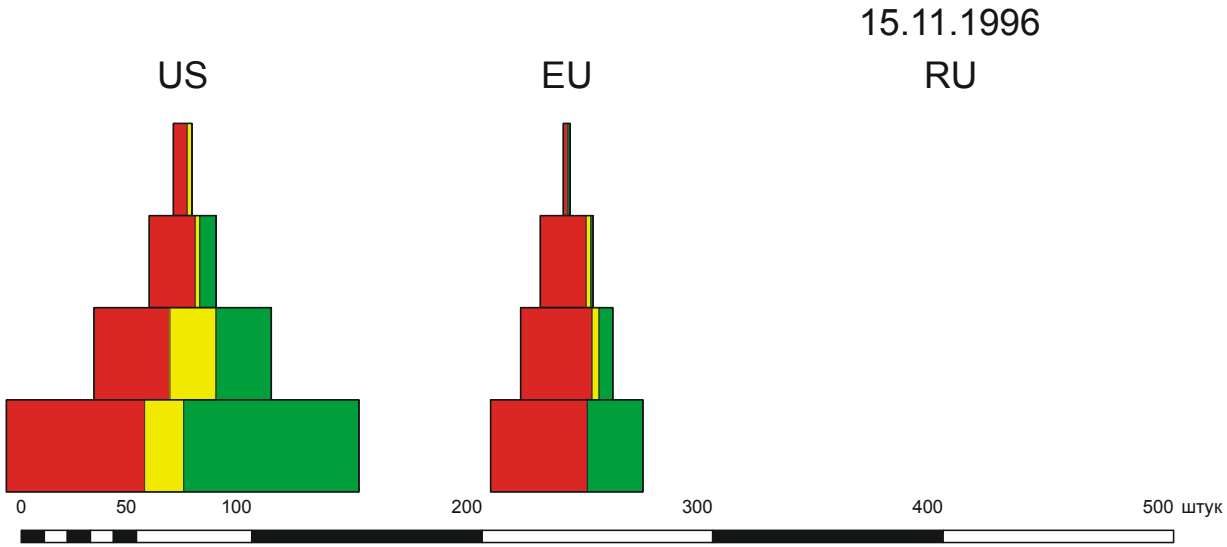
Киберинфраструктура ведущих стран

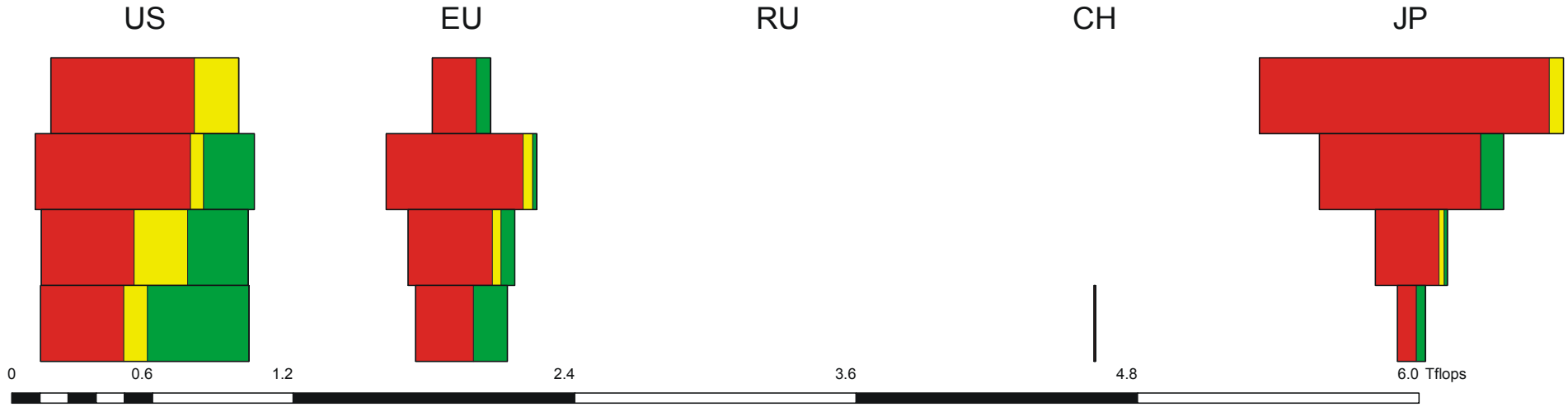


Киберинфраструктура ведущих стран

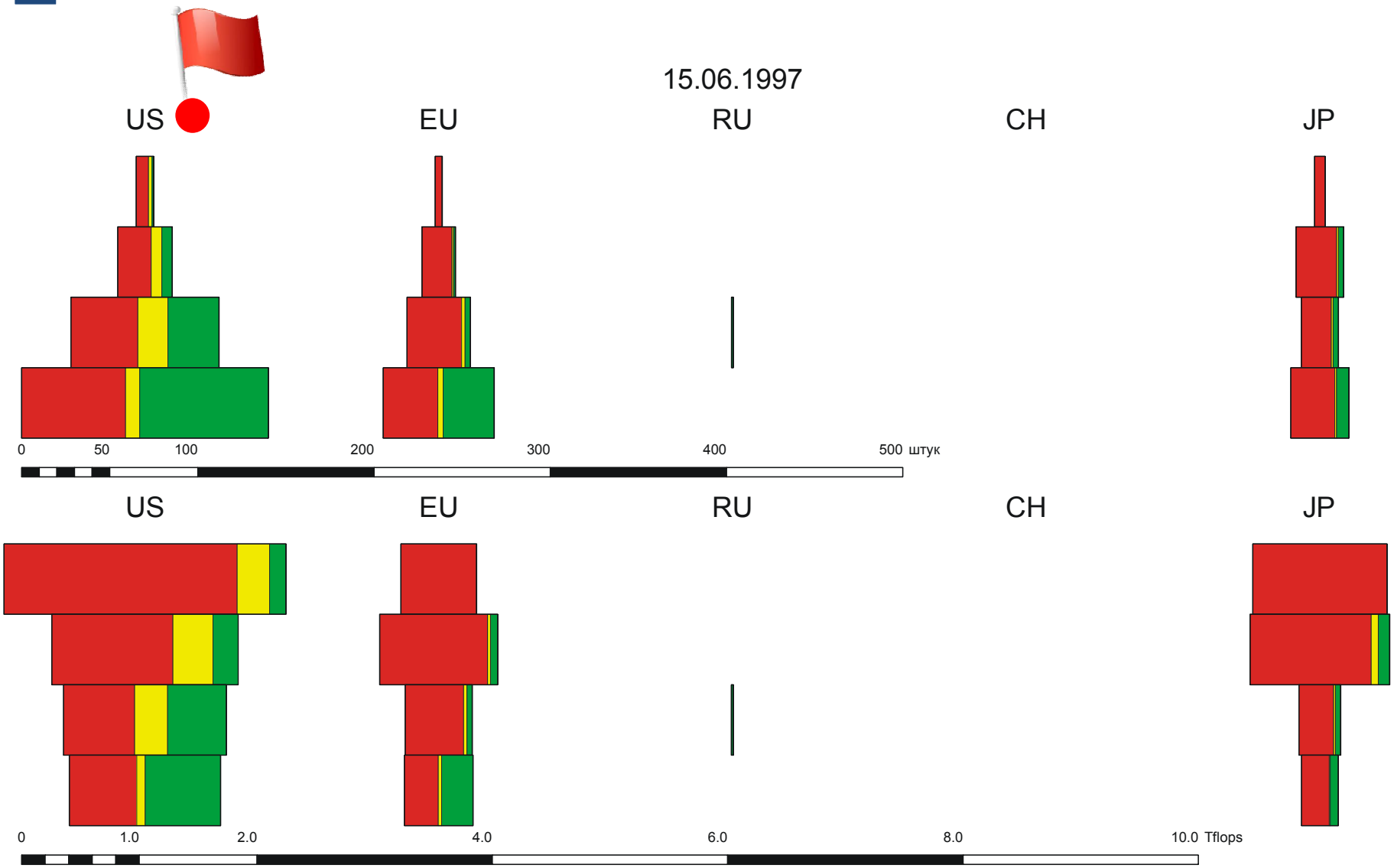


Киберинфраструктура ведущих стран



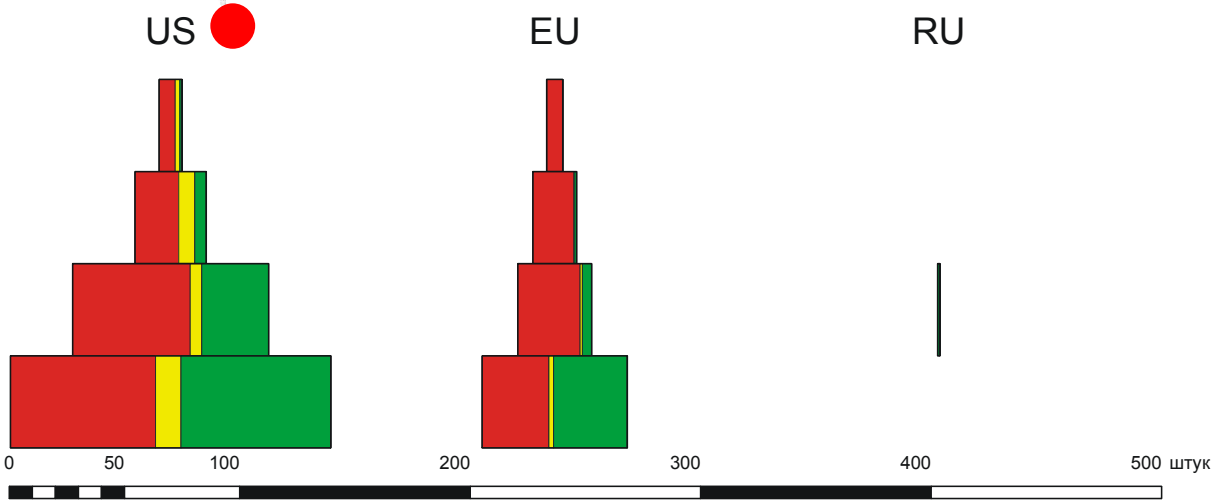


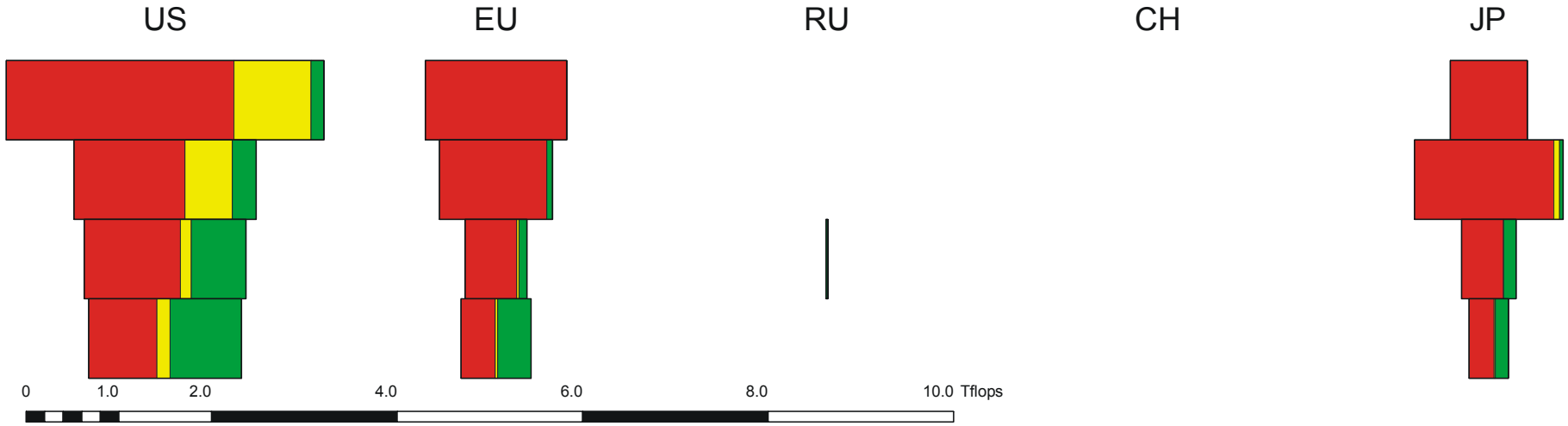
Киберинфраструктура ведущих стран



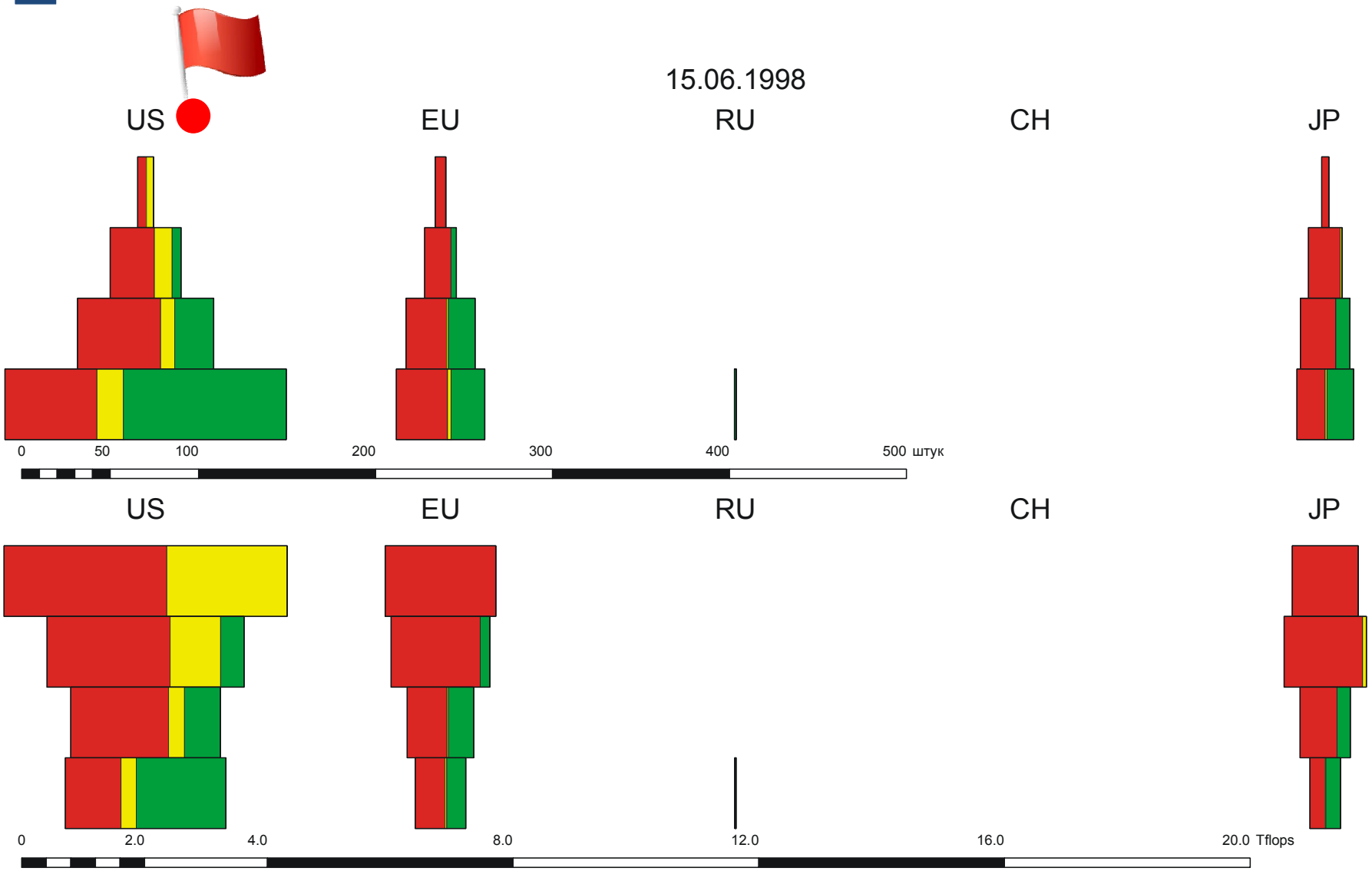
Киберинфраструктура ведущих стран

15.11.1997

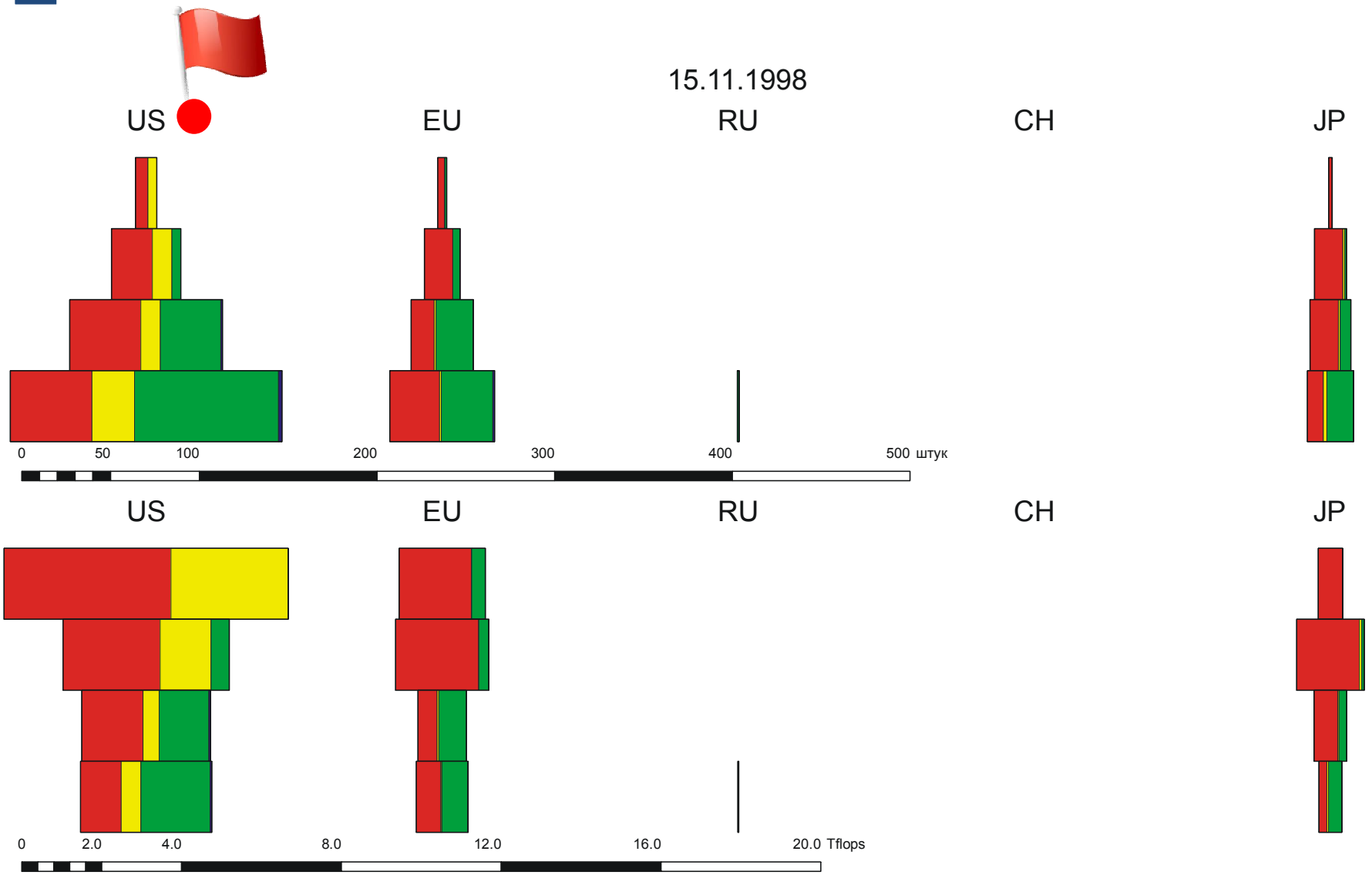




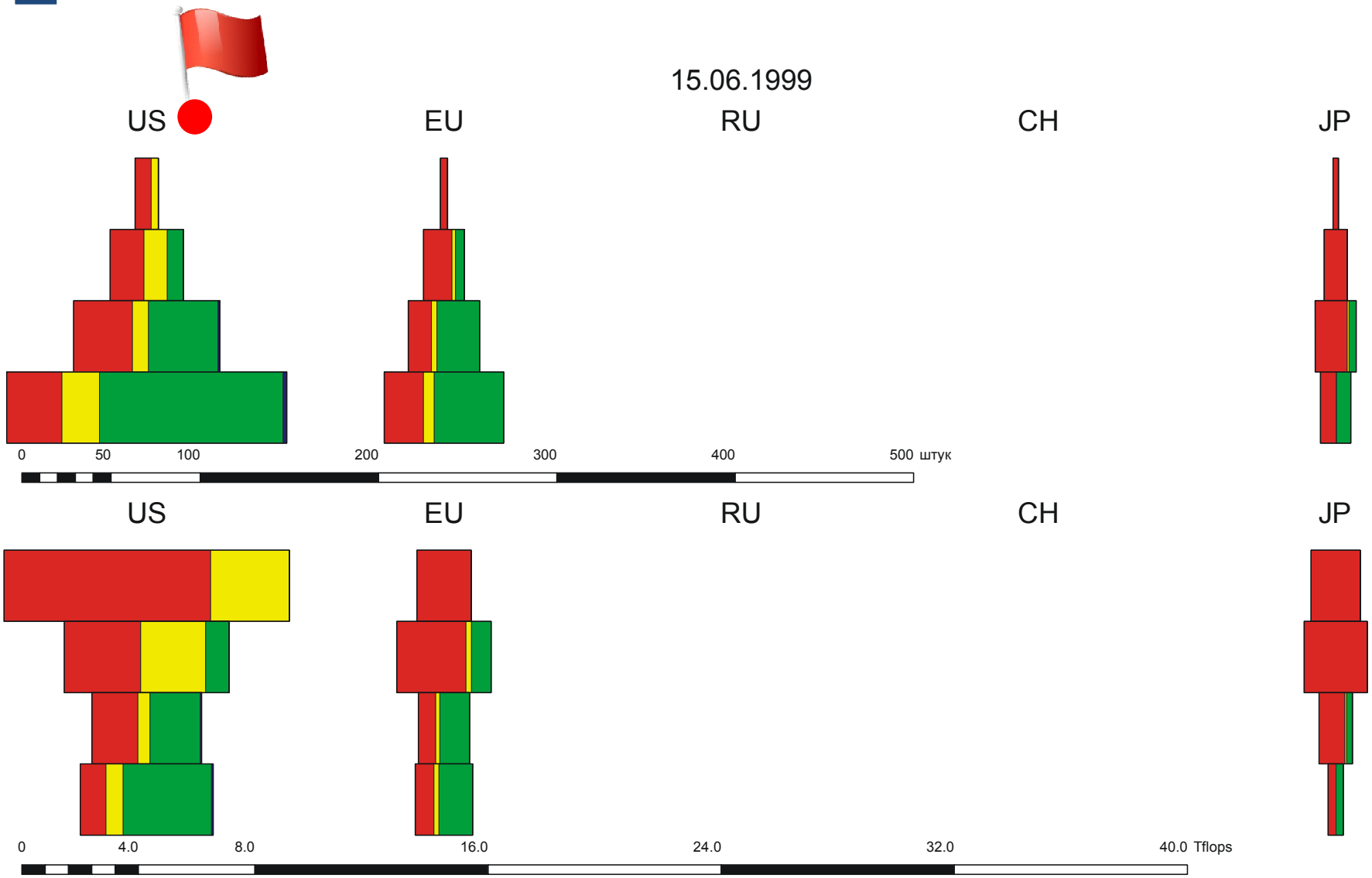
Киберинфраструктура ведущих стран



Киберинфраструктура ведущих стран

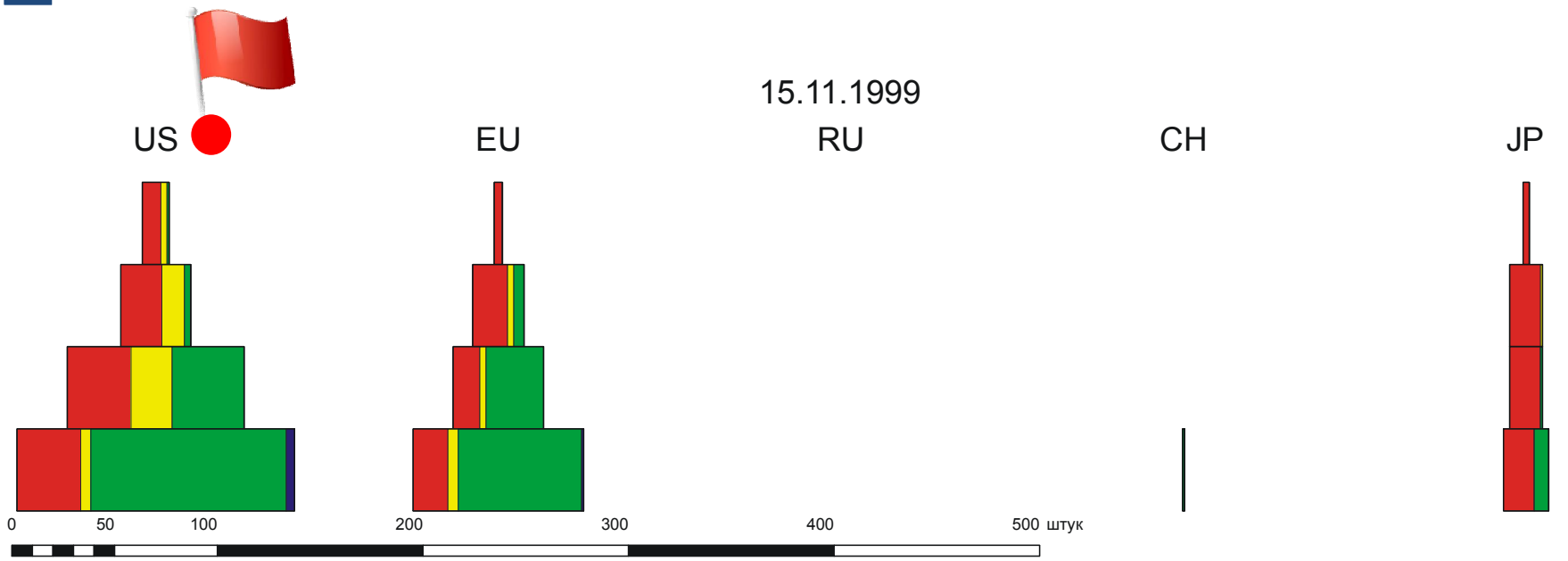


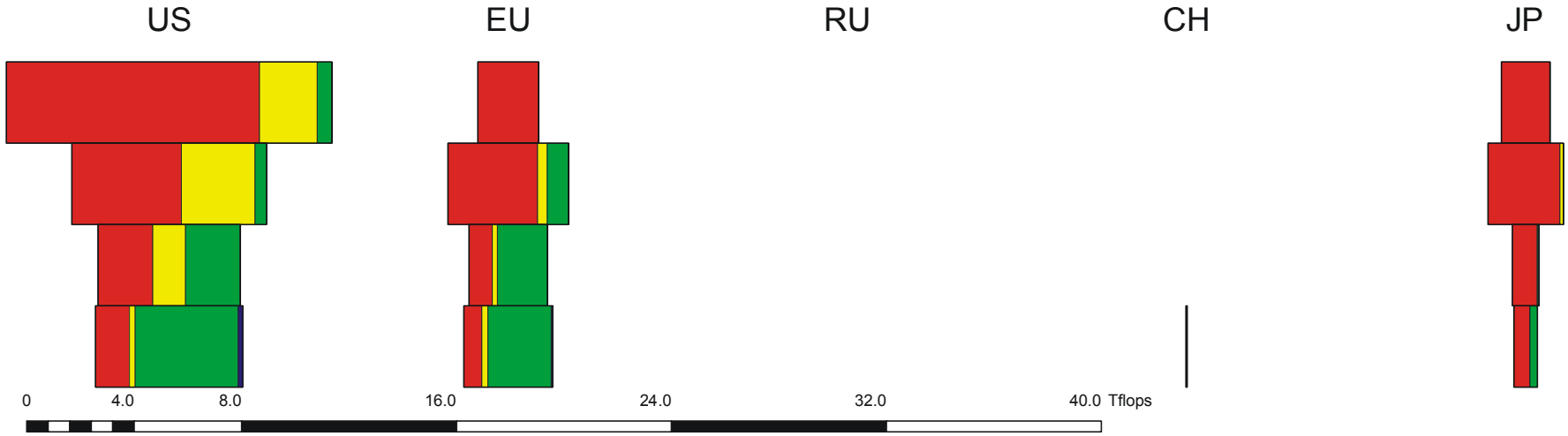
Киберинфраструктура ведущих стран



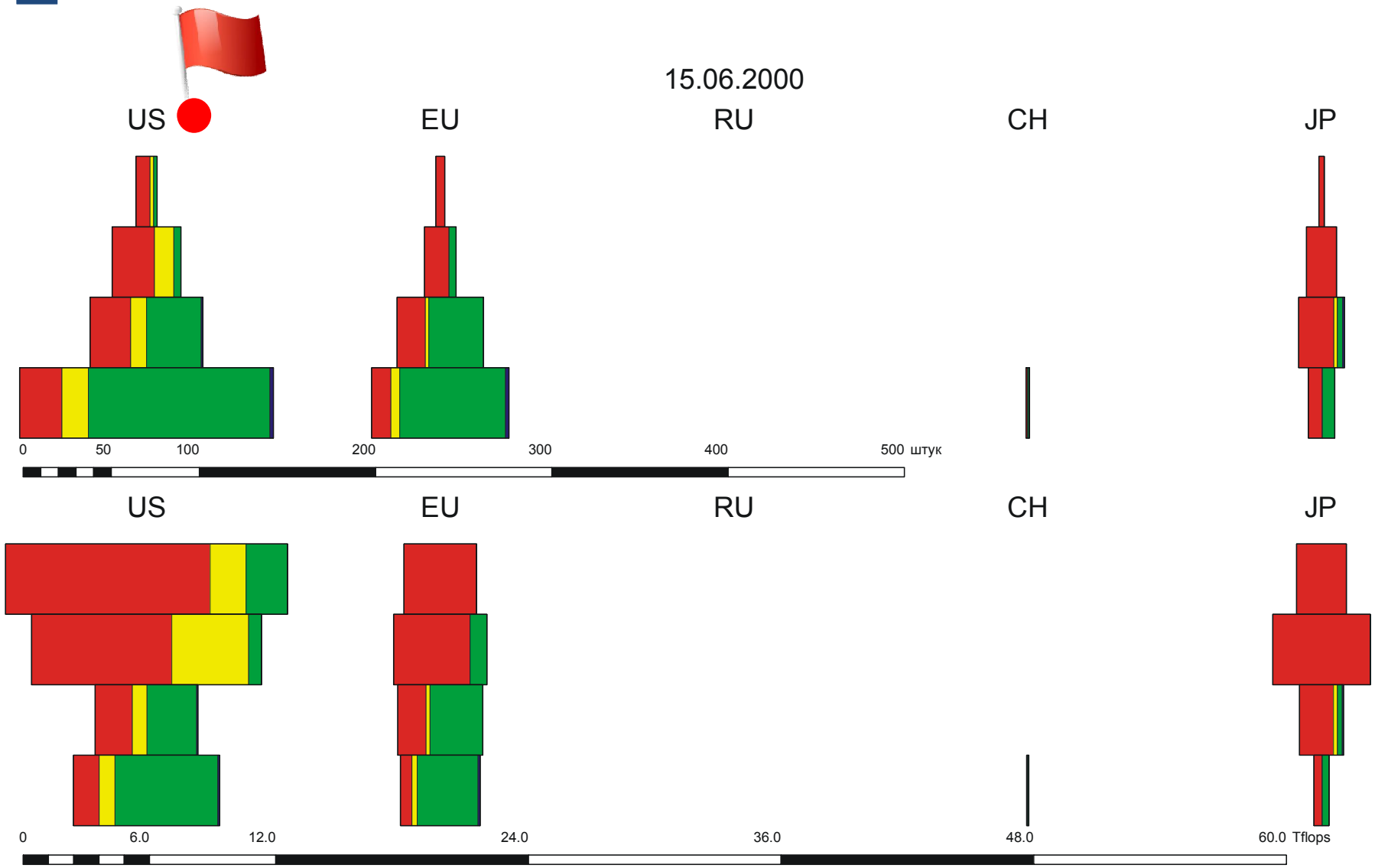
Киберинфраструктура ведущих стран

15.11.1999

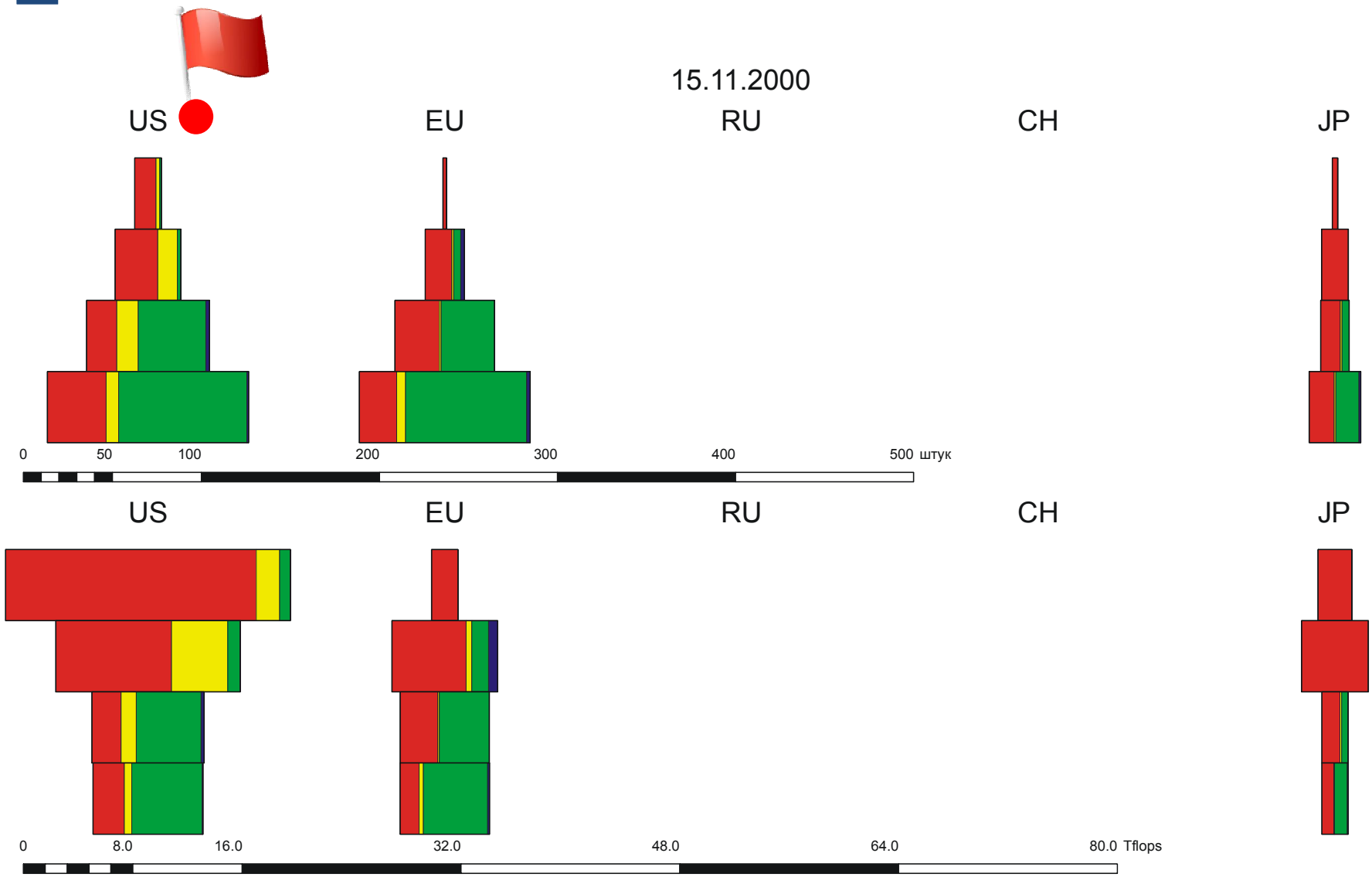




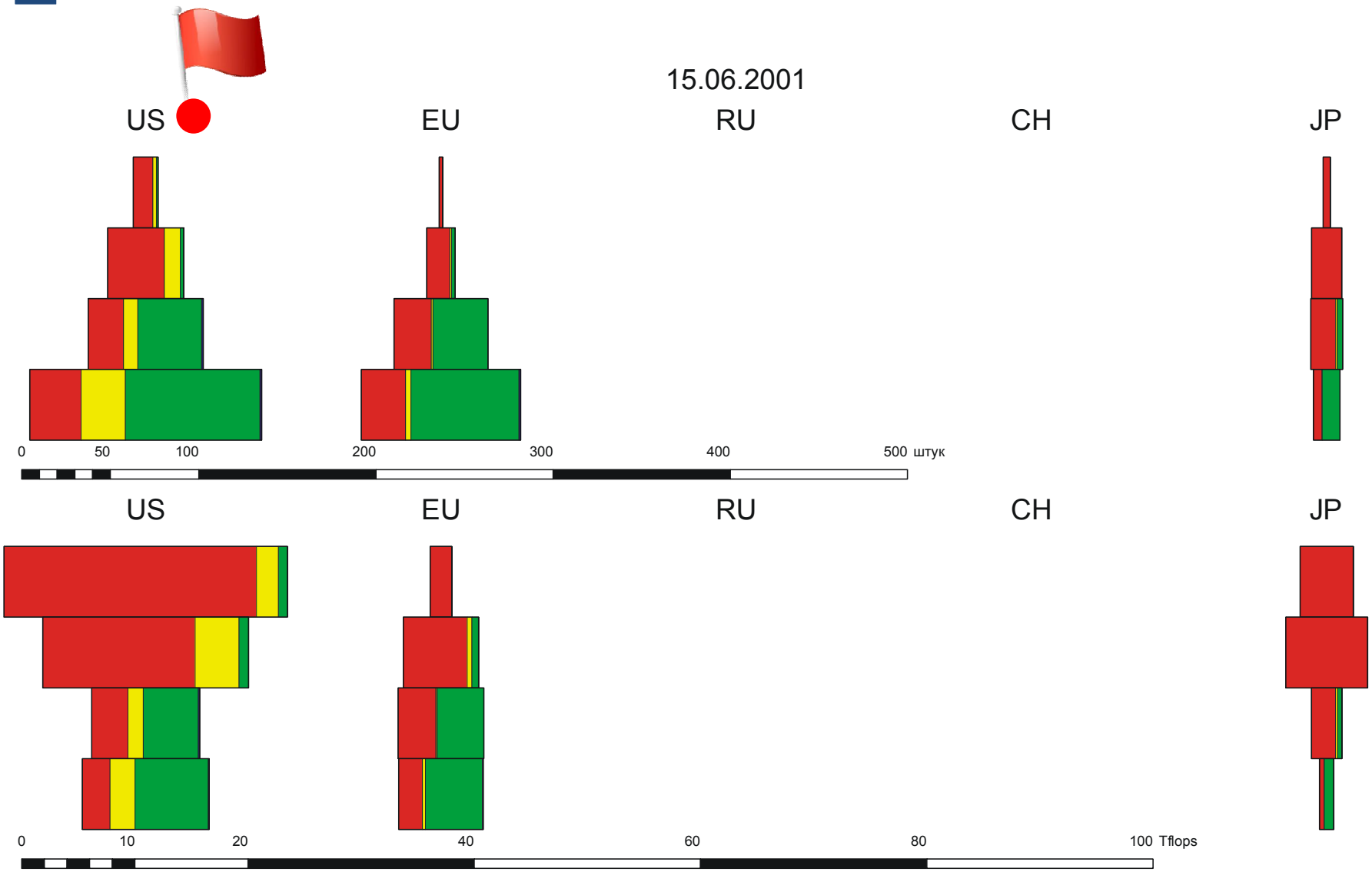
Киберинфраструктура ведущих стран



Киберинфраструктура ведущих стран

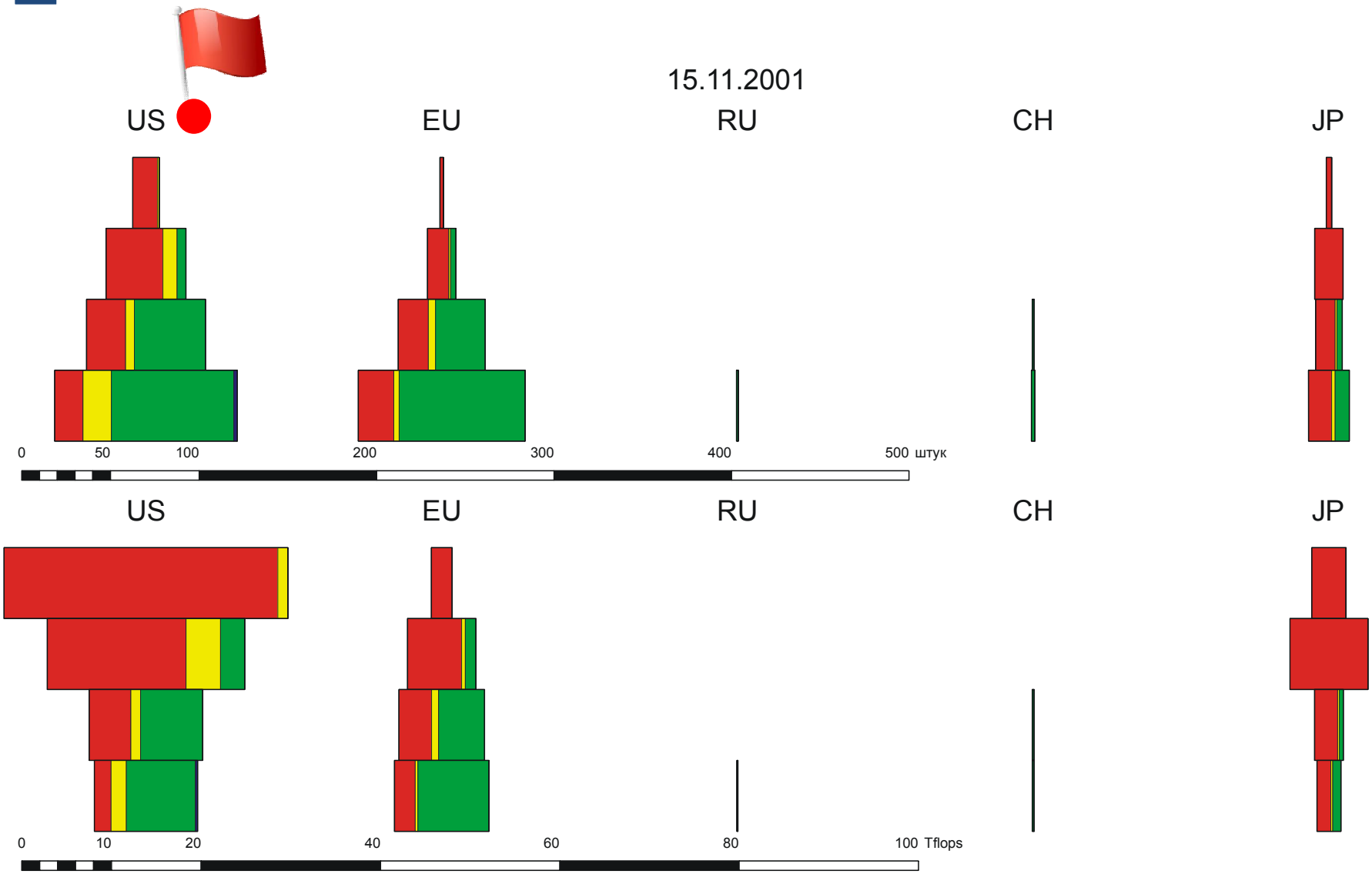


Киберинфраструктура ведущих стран

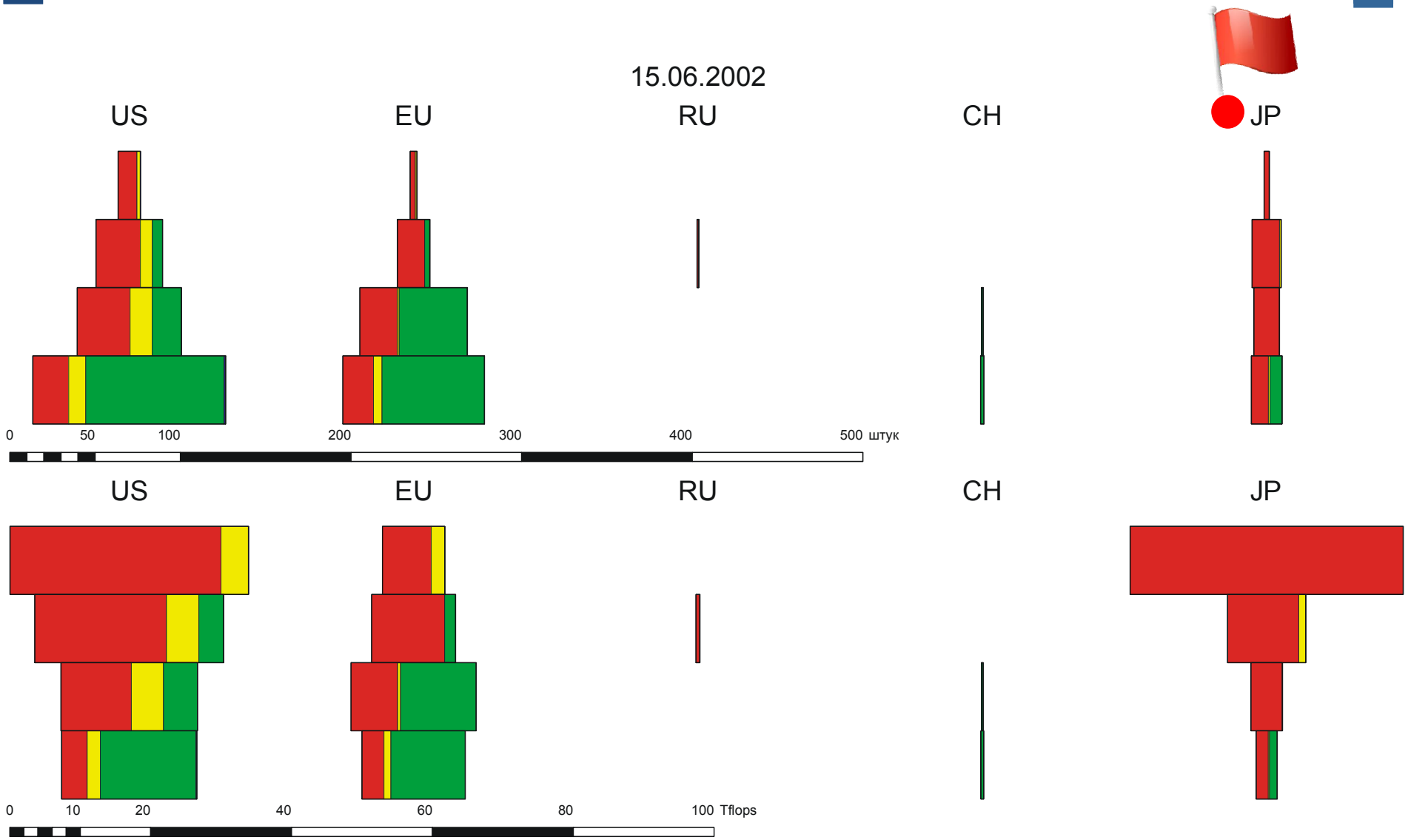


Киберинфраструктура ведущих стран

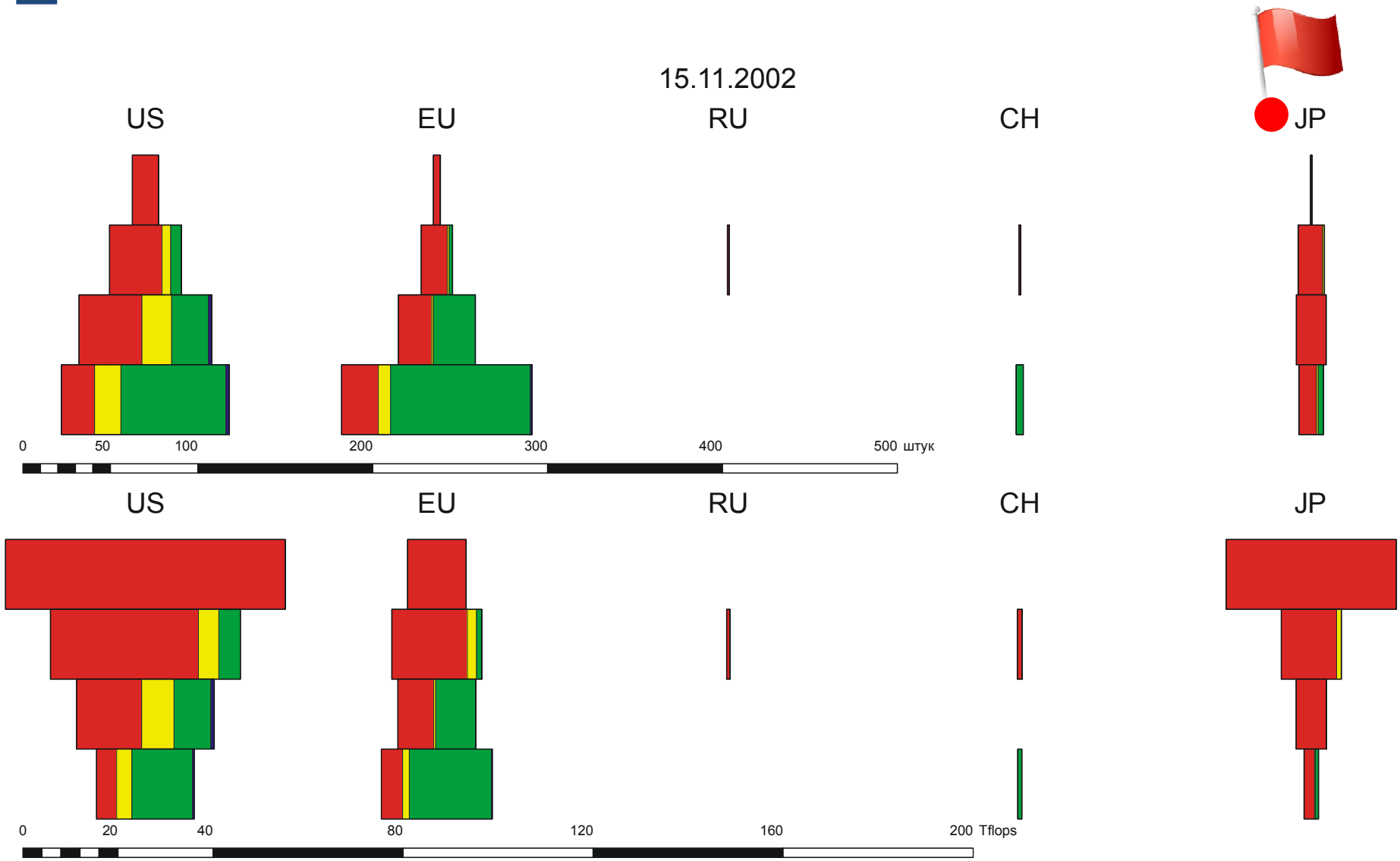
15.11.2001



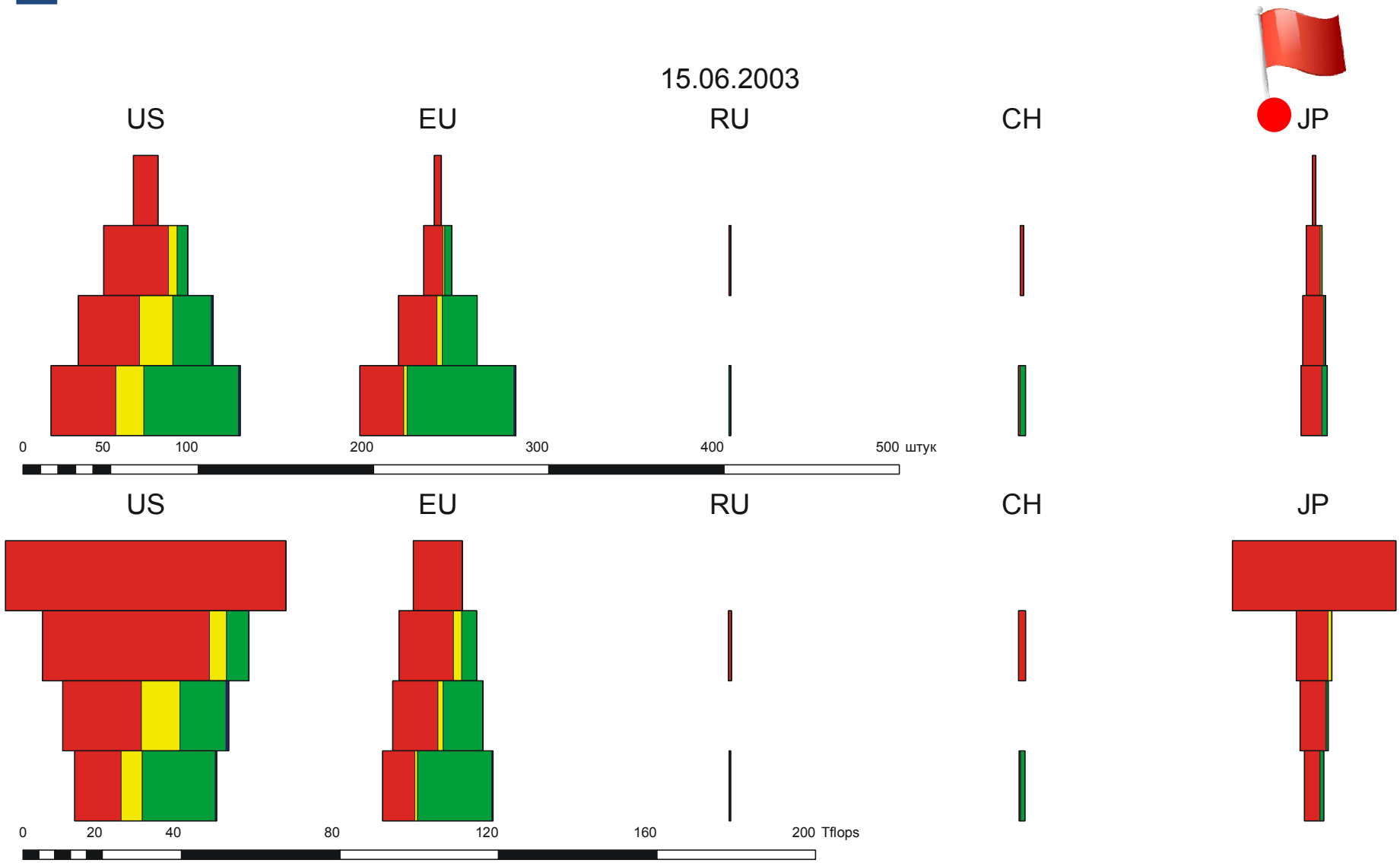
Киберинфраструктура ведущих стран



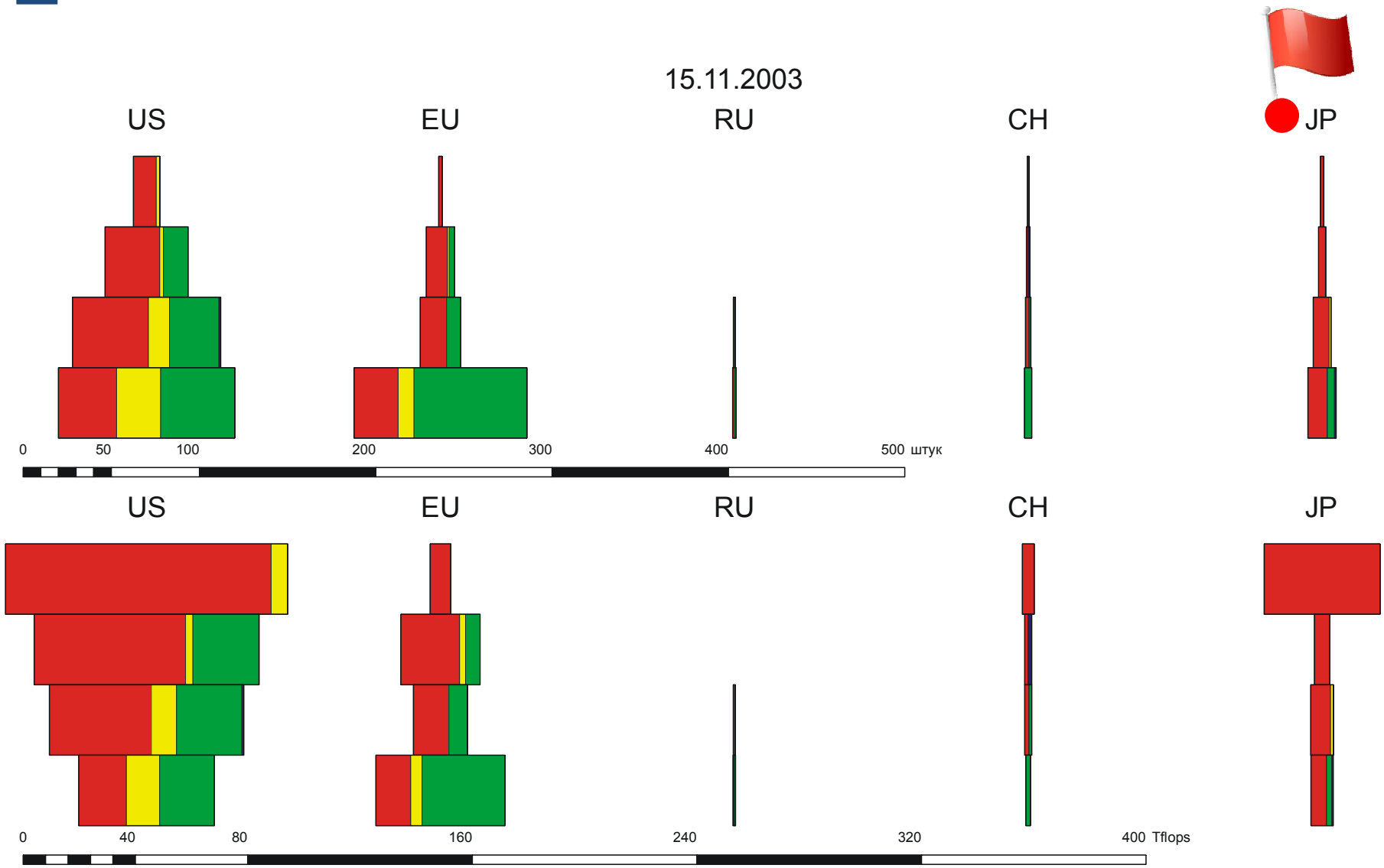
Киберинфраструктура ведущих стран



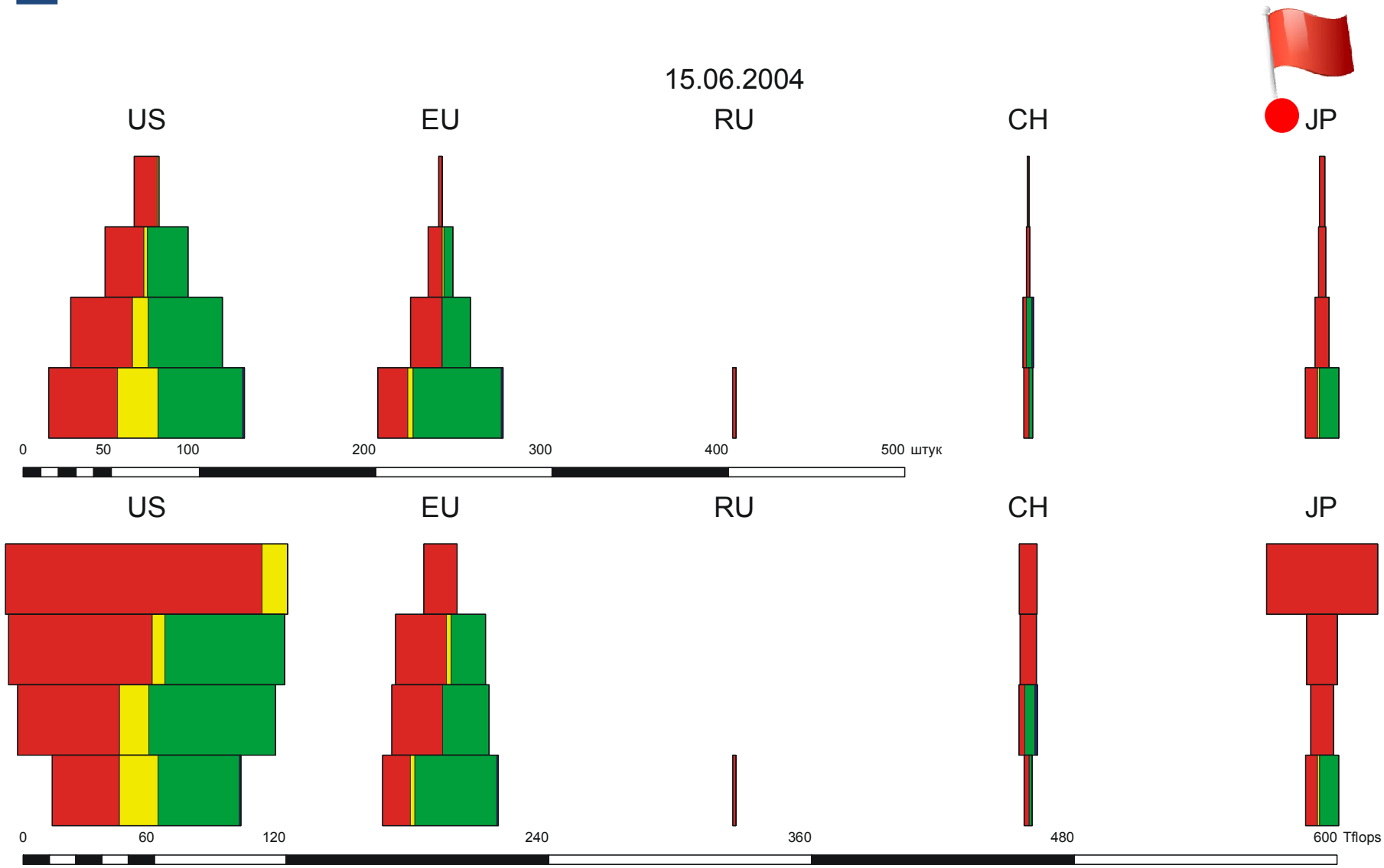
Киберинфраструктура ведущих стран



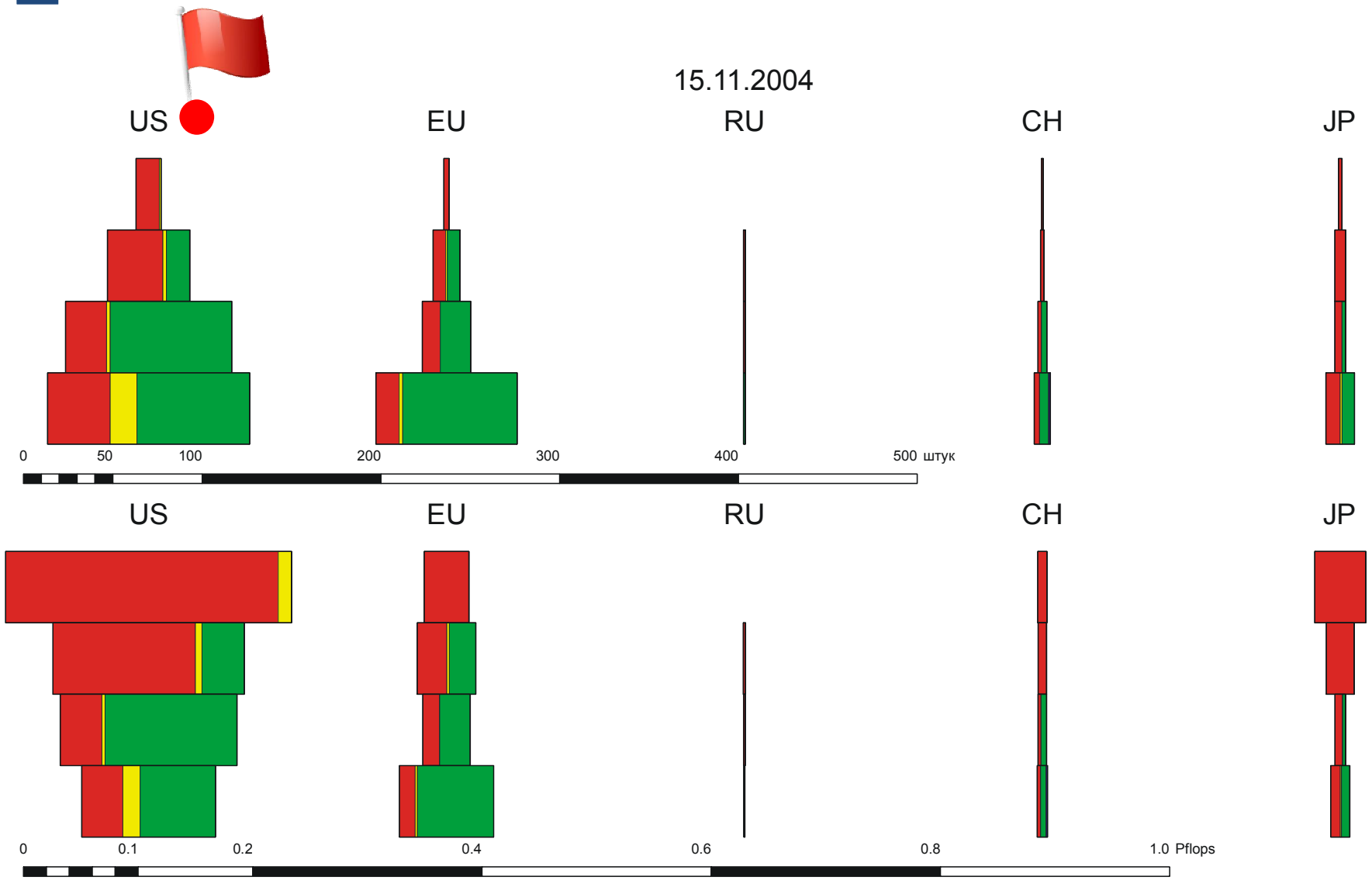
Киберинфраструктура ведущих стран



Киберинфраструктура ведущих стран

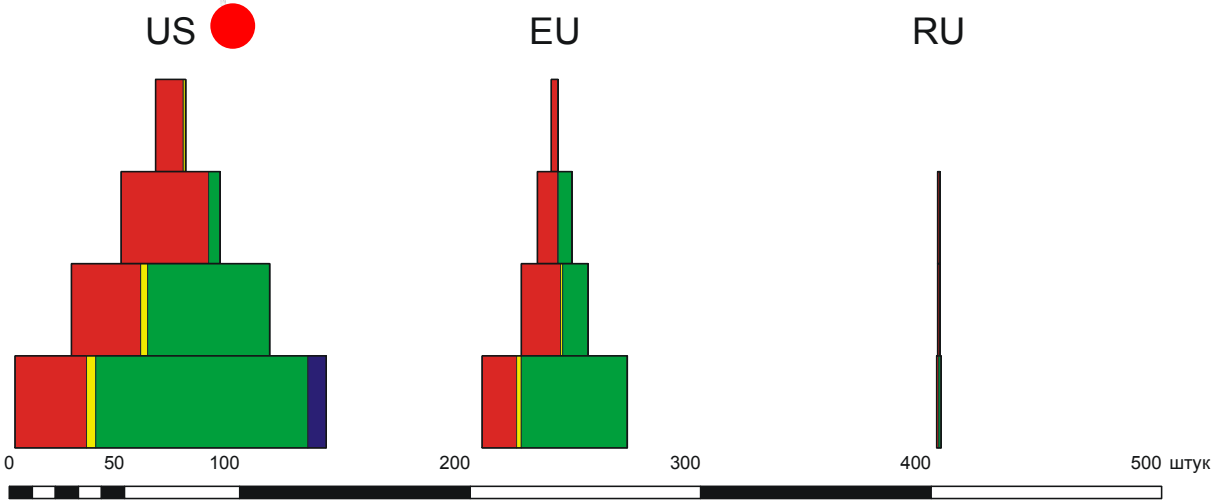


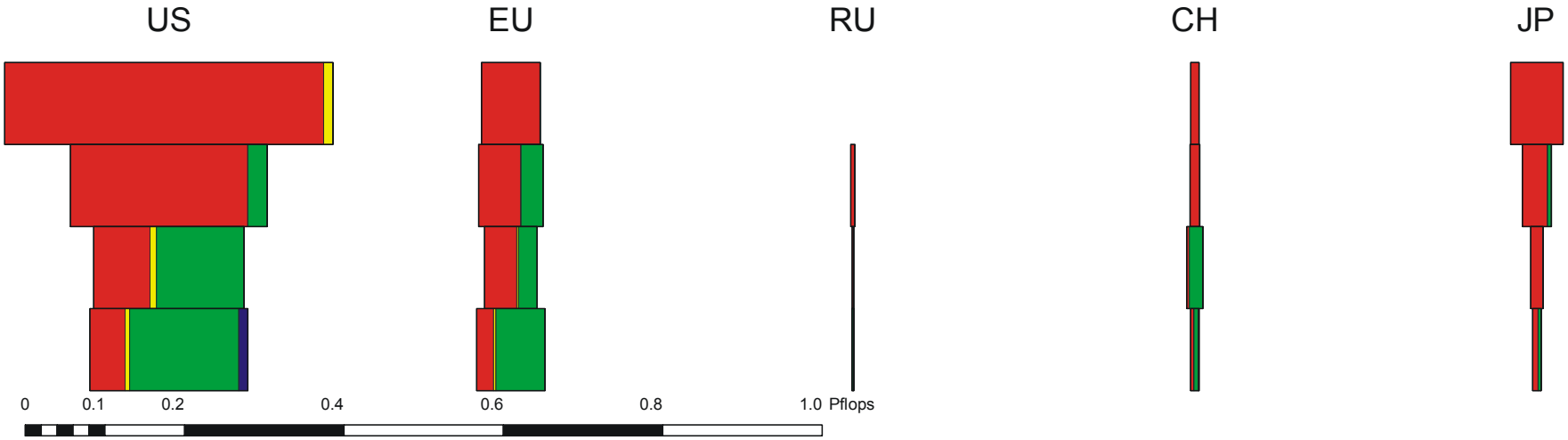
Киберинфраструктура ведущих стран



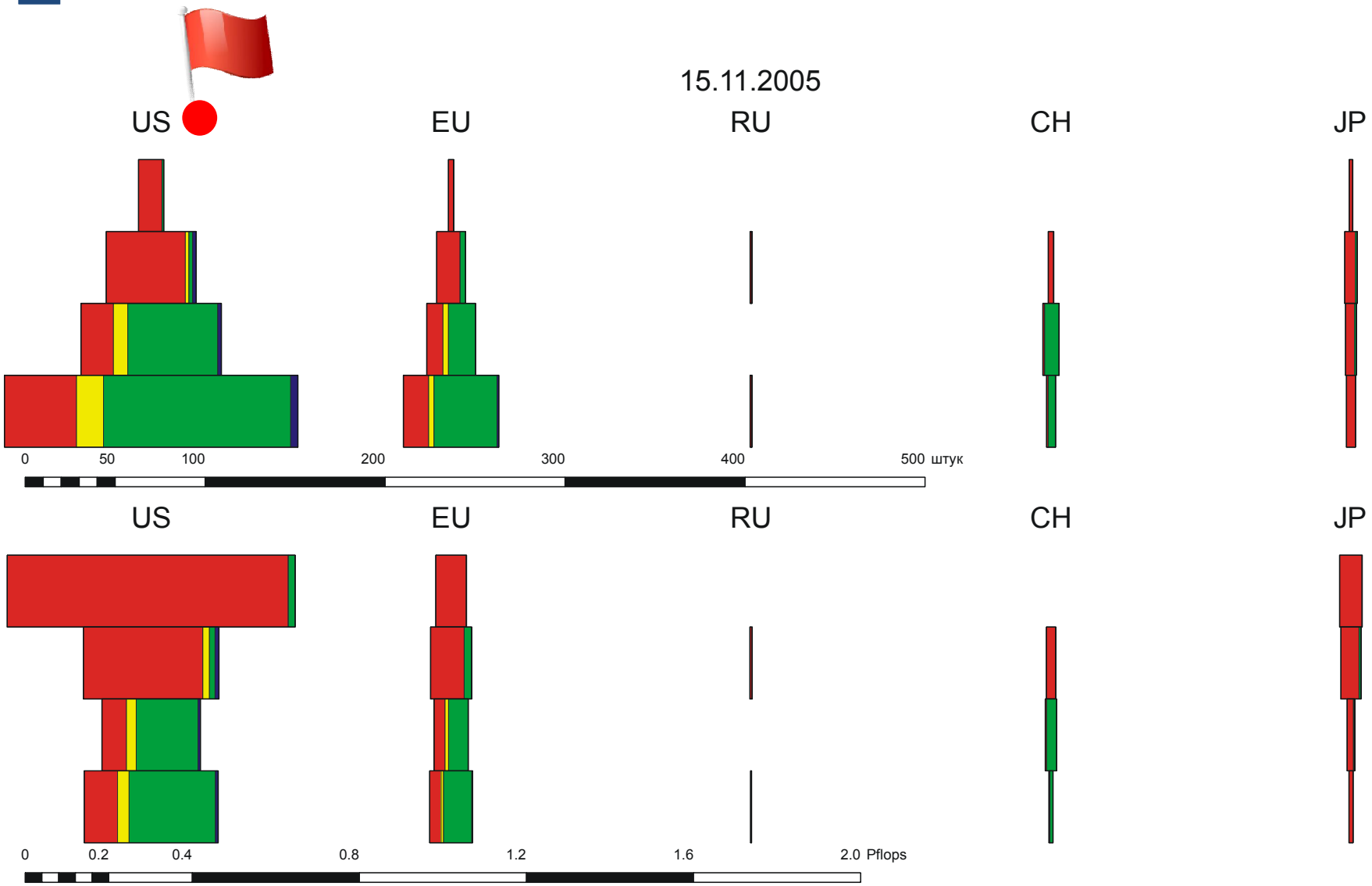
Киберинфраструктура ведущих стран

15.06.2005

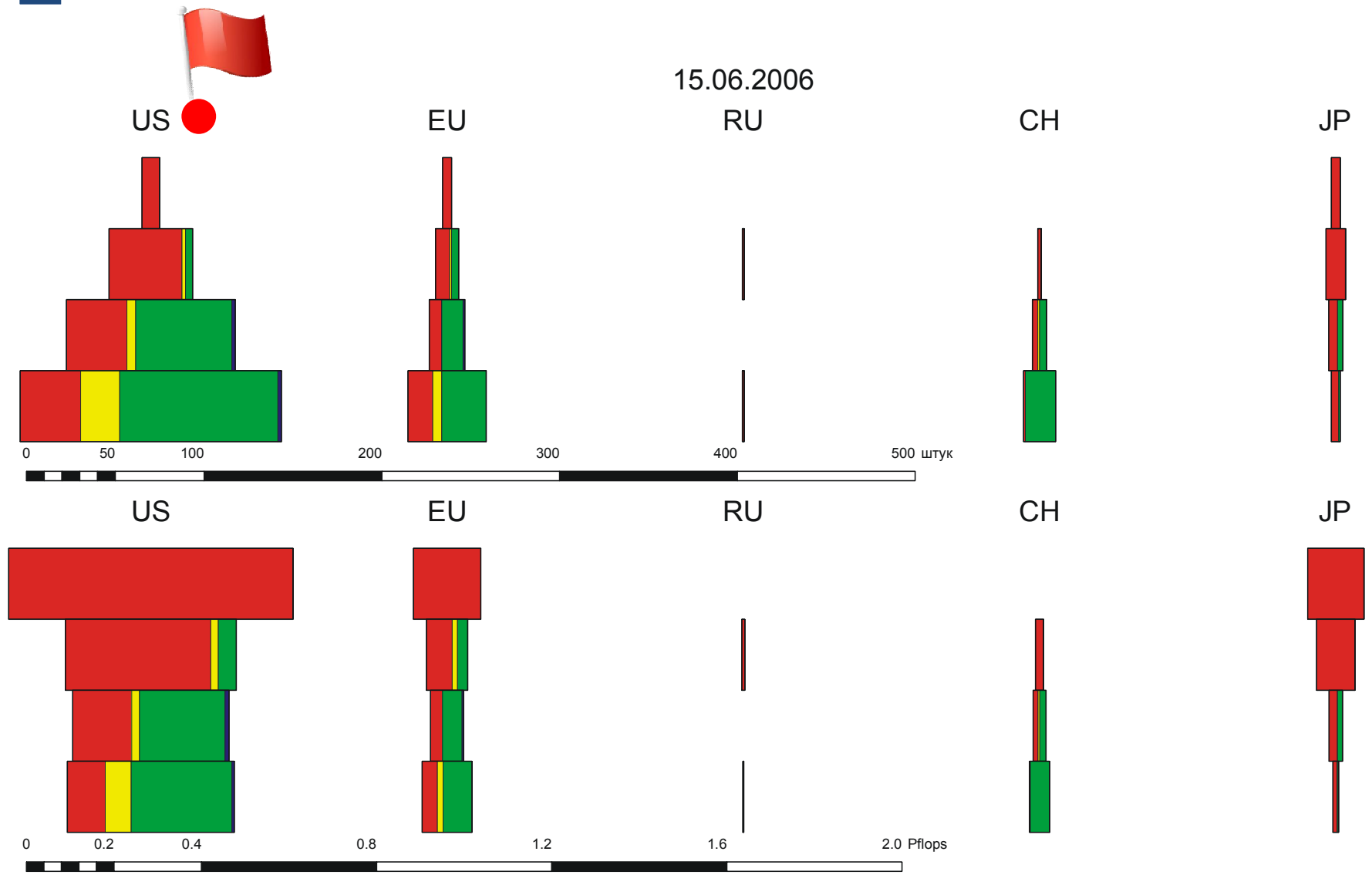




Киберинфраструктура ведущих стран

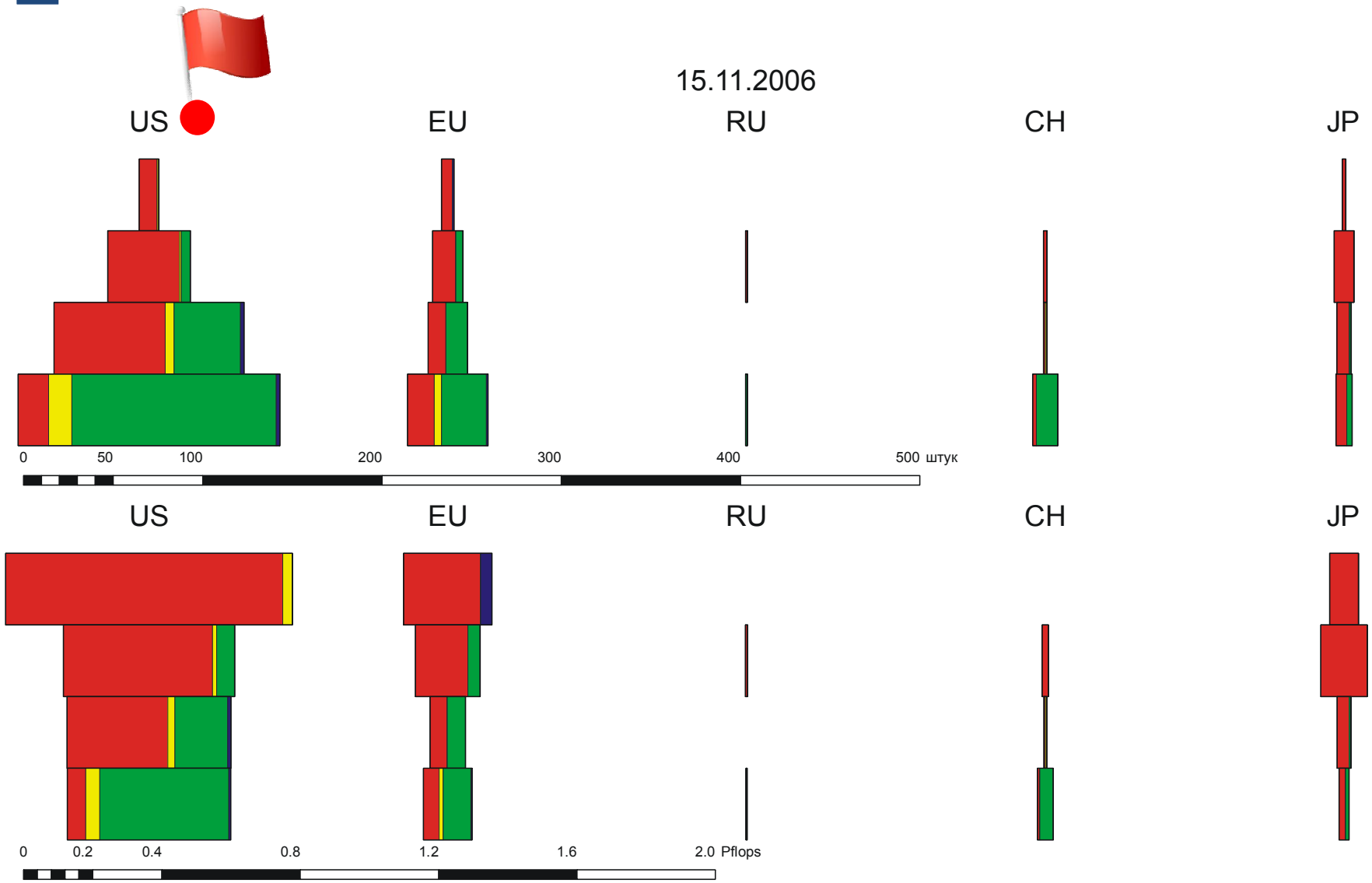


Киберинфраструктура ведущих стран

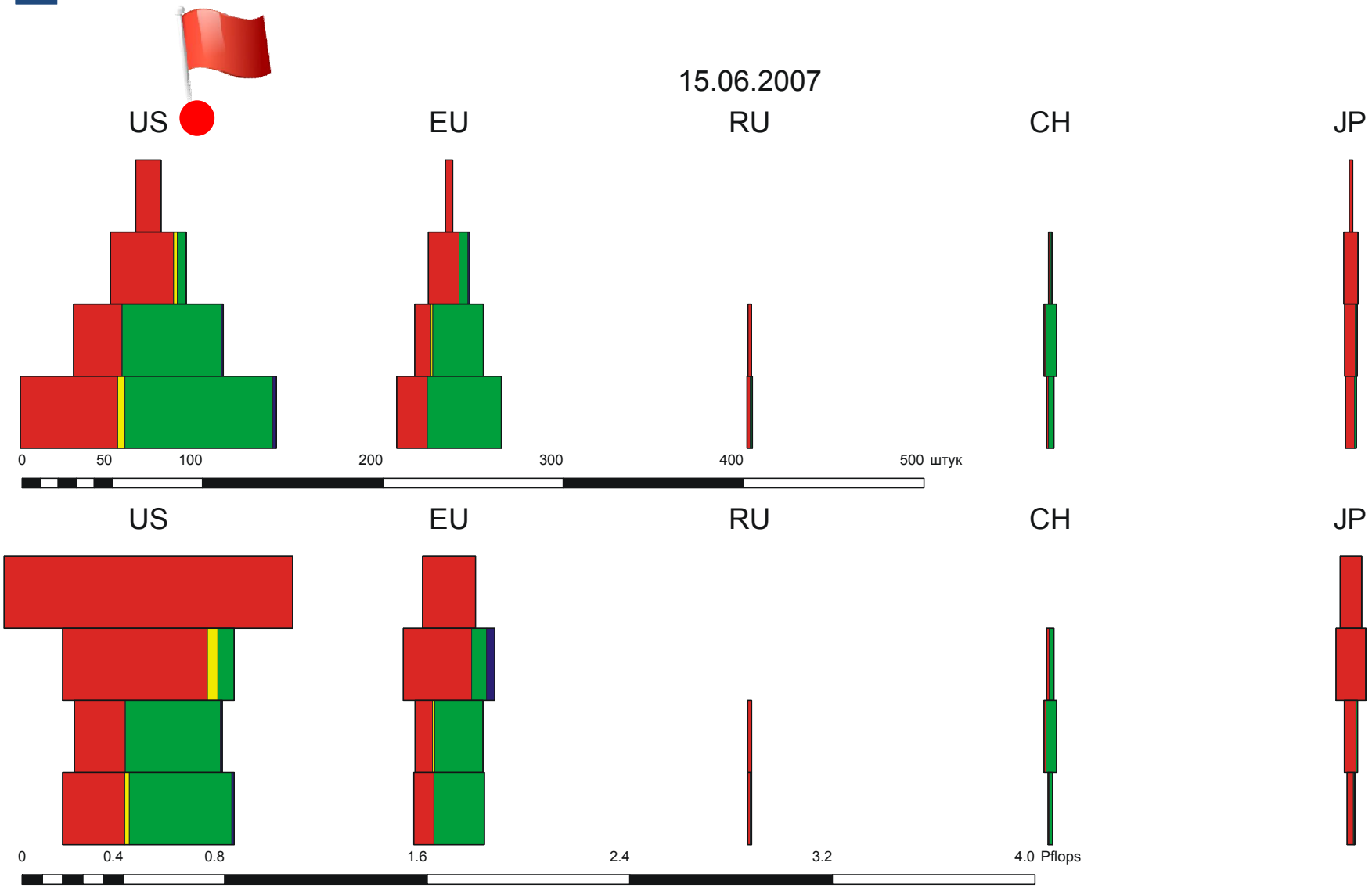


Киберинфраструктура ведущих стран

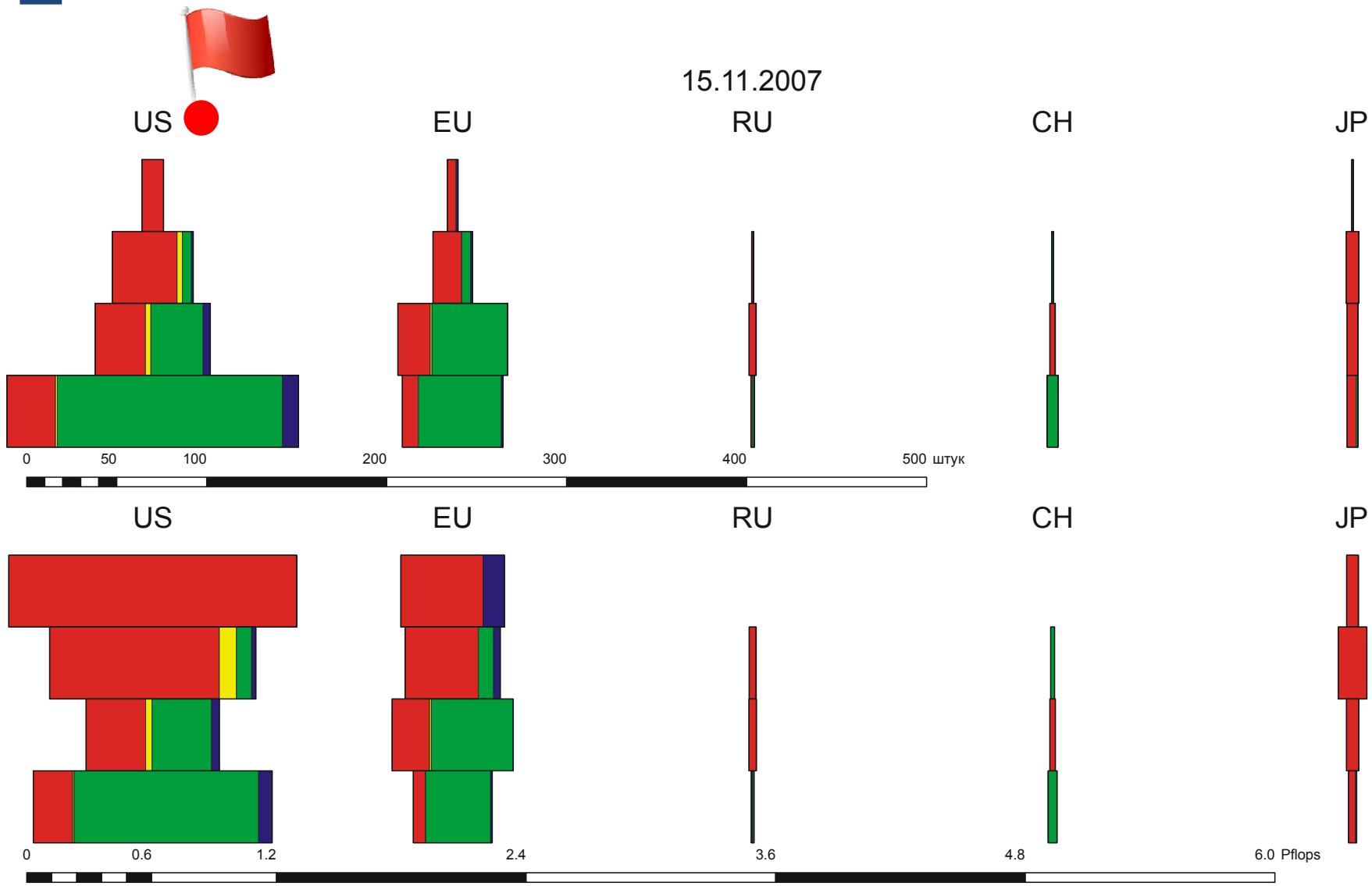
15.11.2006



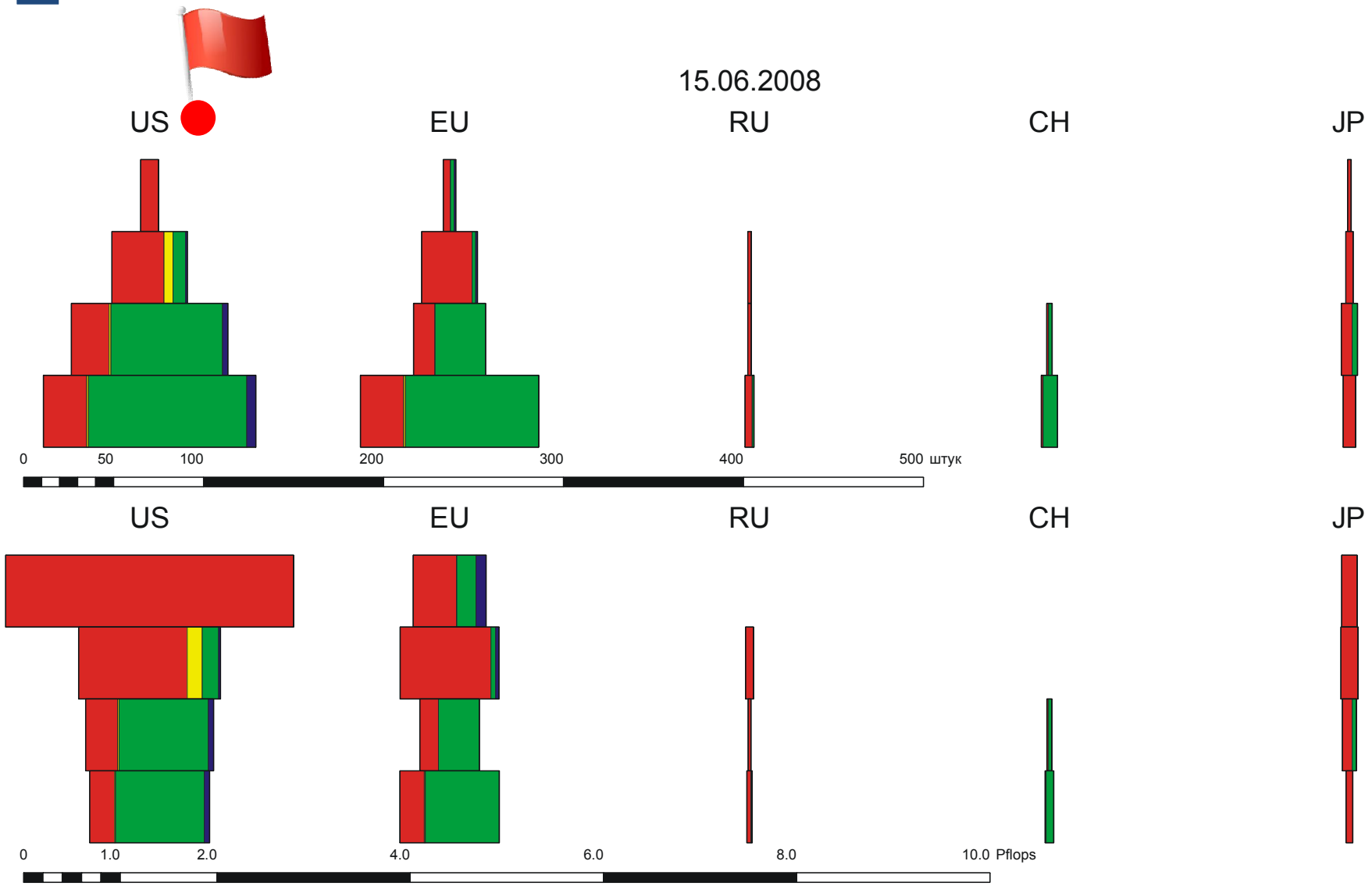
Киберинфраструктура ведущих стран



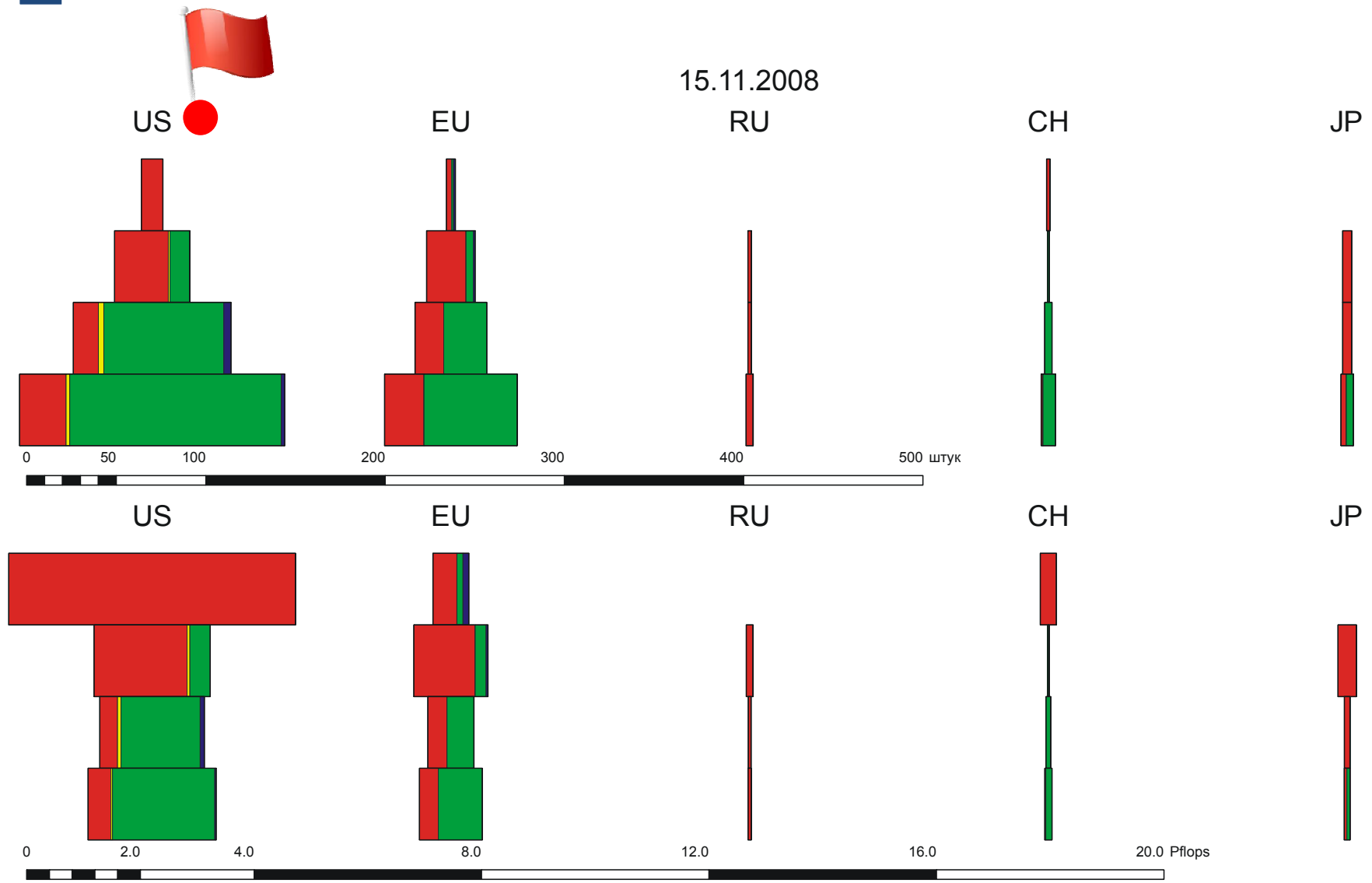
Киберинфраструктура ведущих стран



Киберинфраструктура ведущих стран

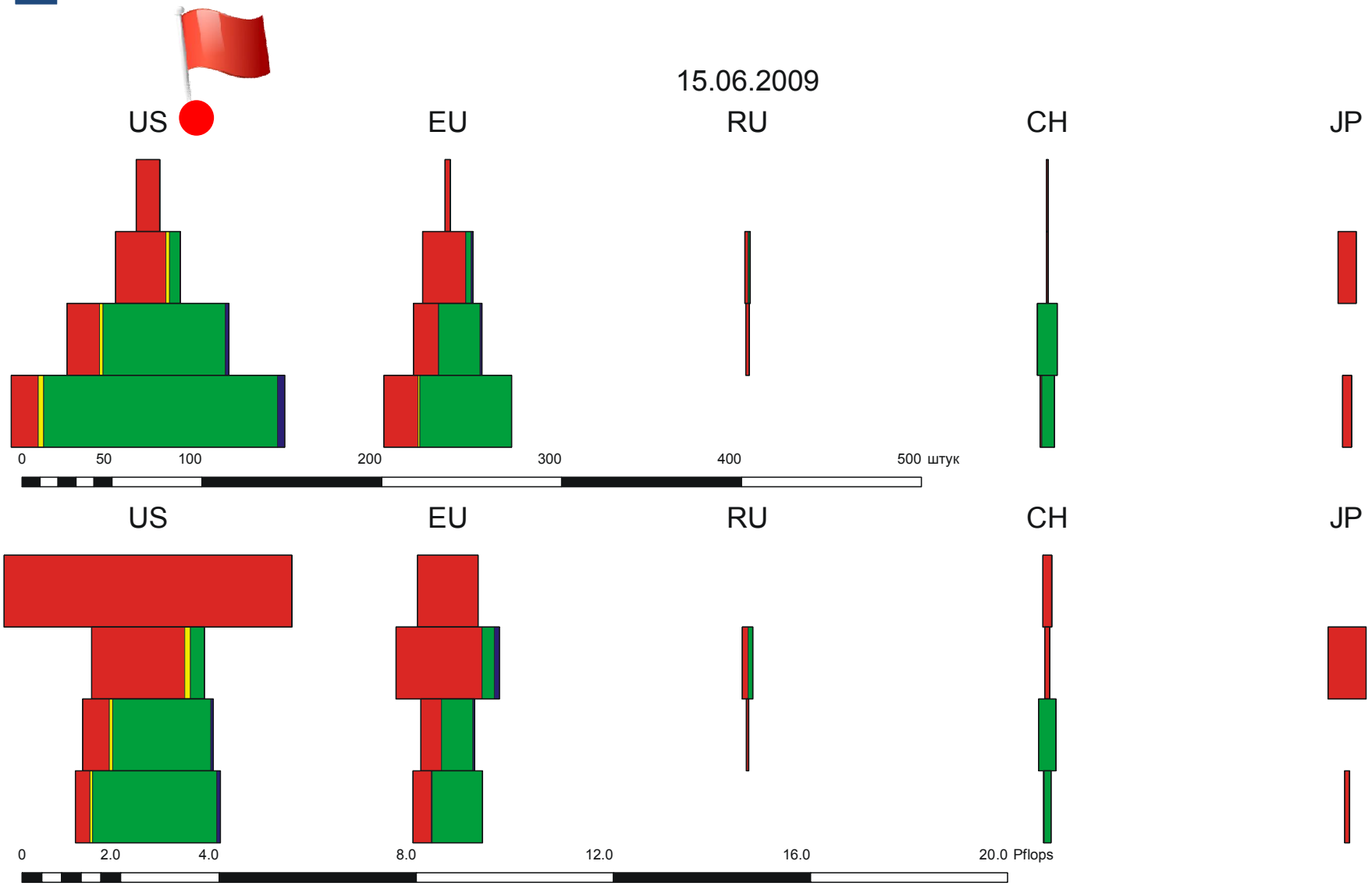


Киберинфраструктура ведущих стран

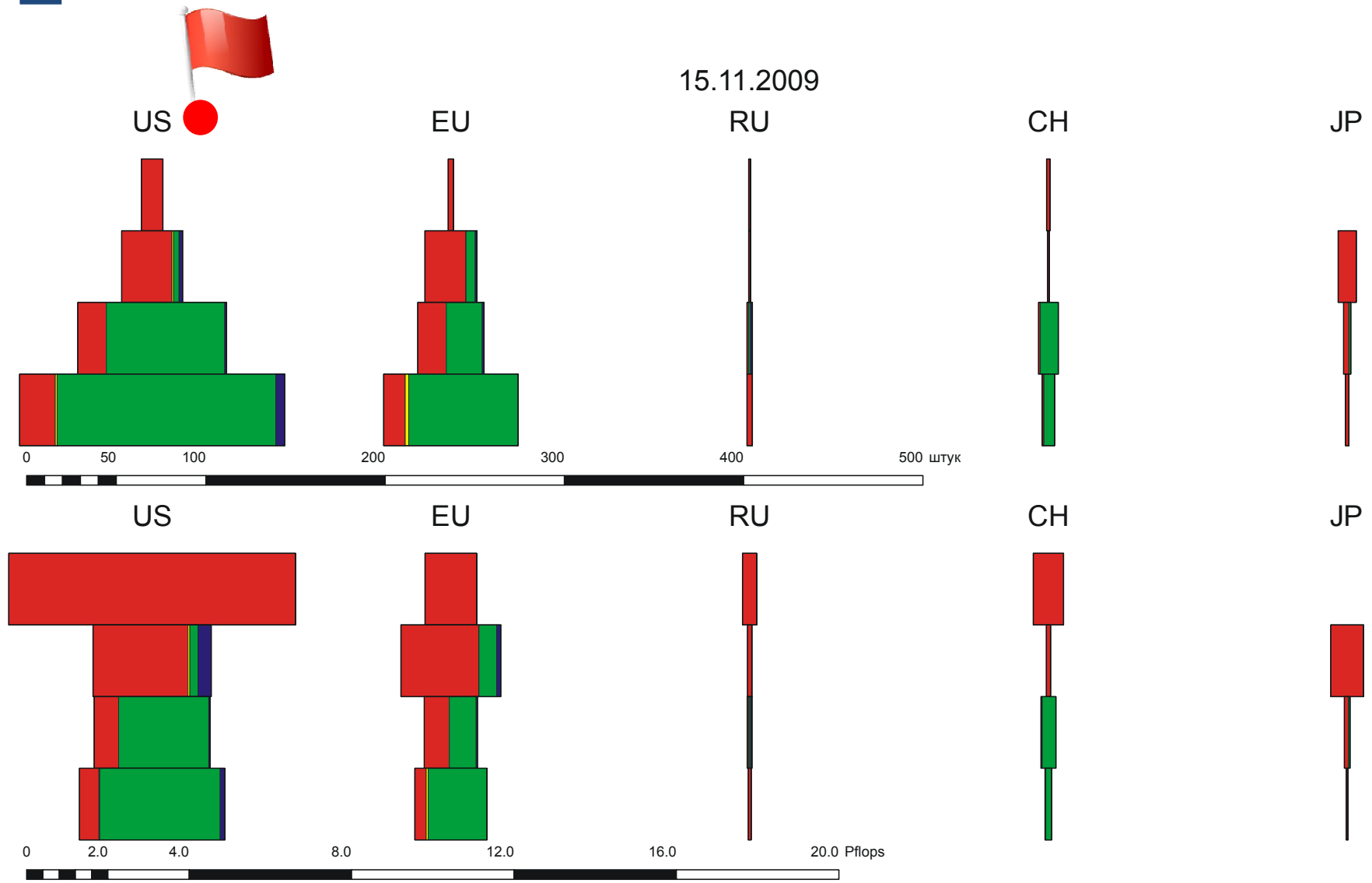


Киберинфраструктура ведущих стран

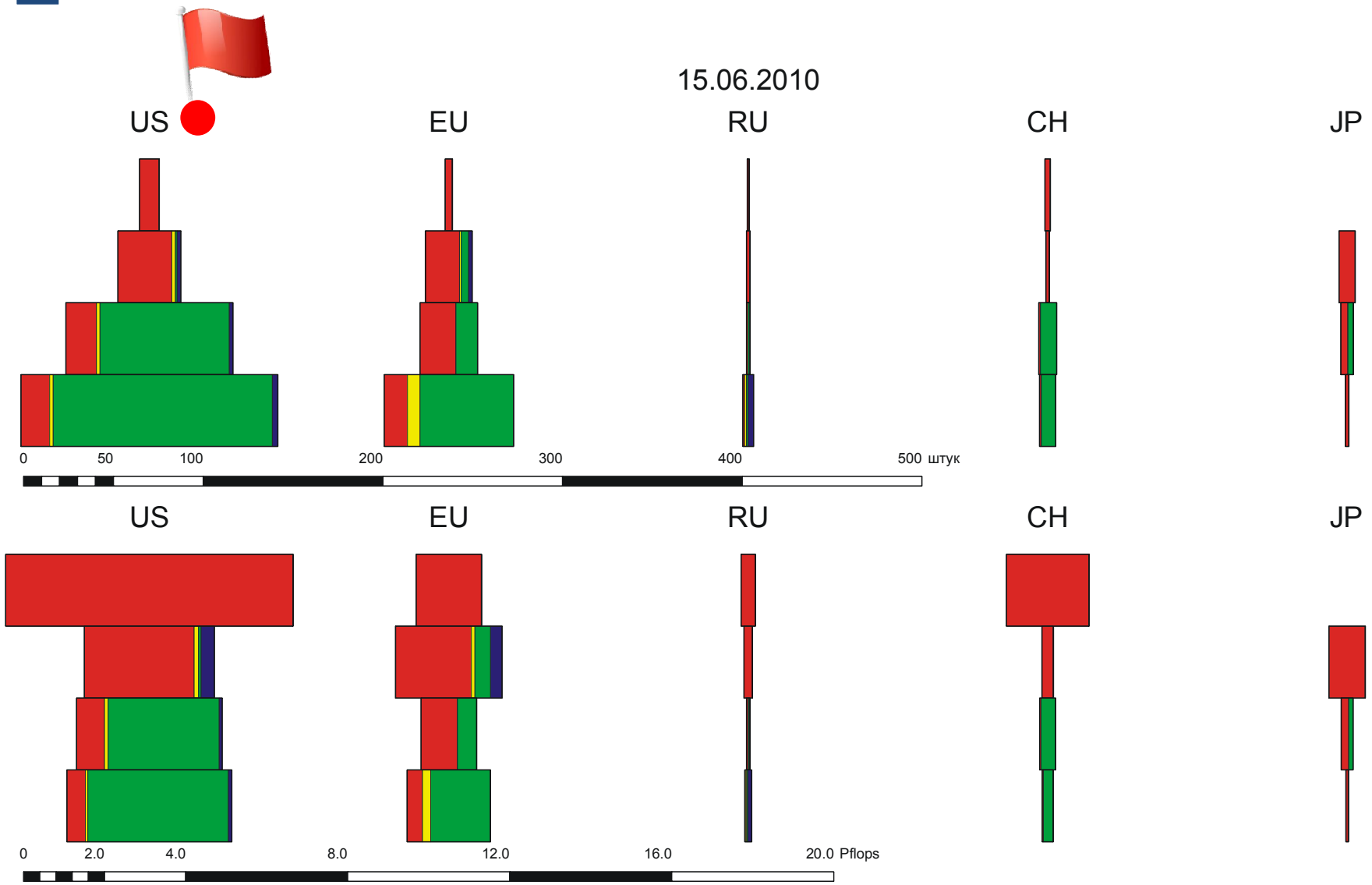
15.06.2009



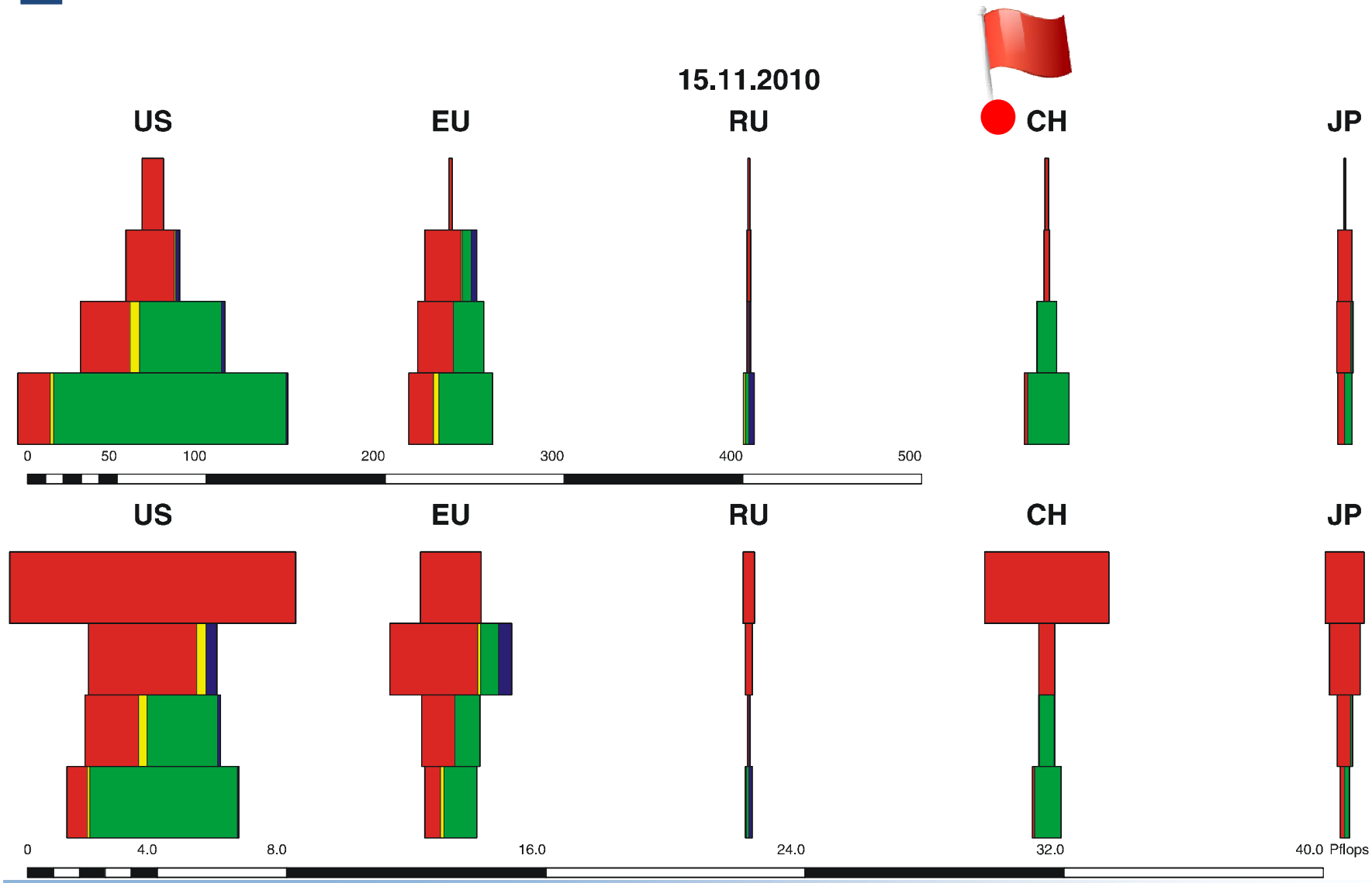
Киберинфраструктура ведущих стран



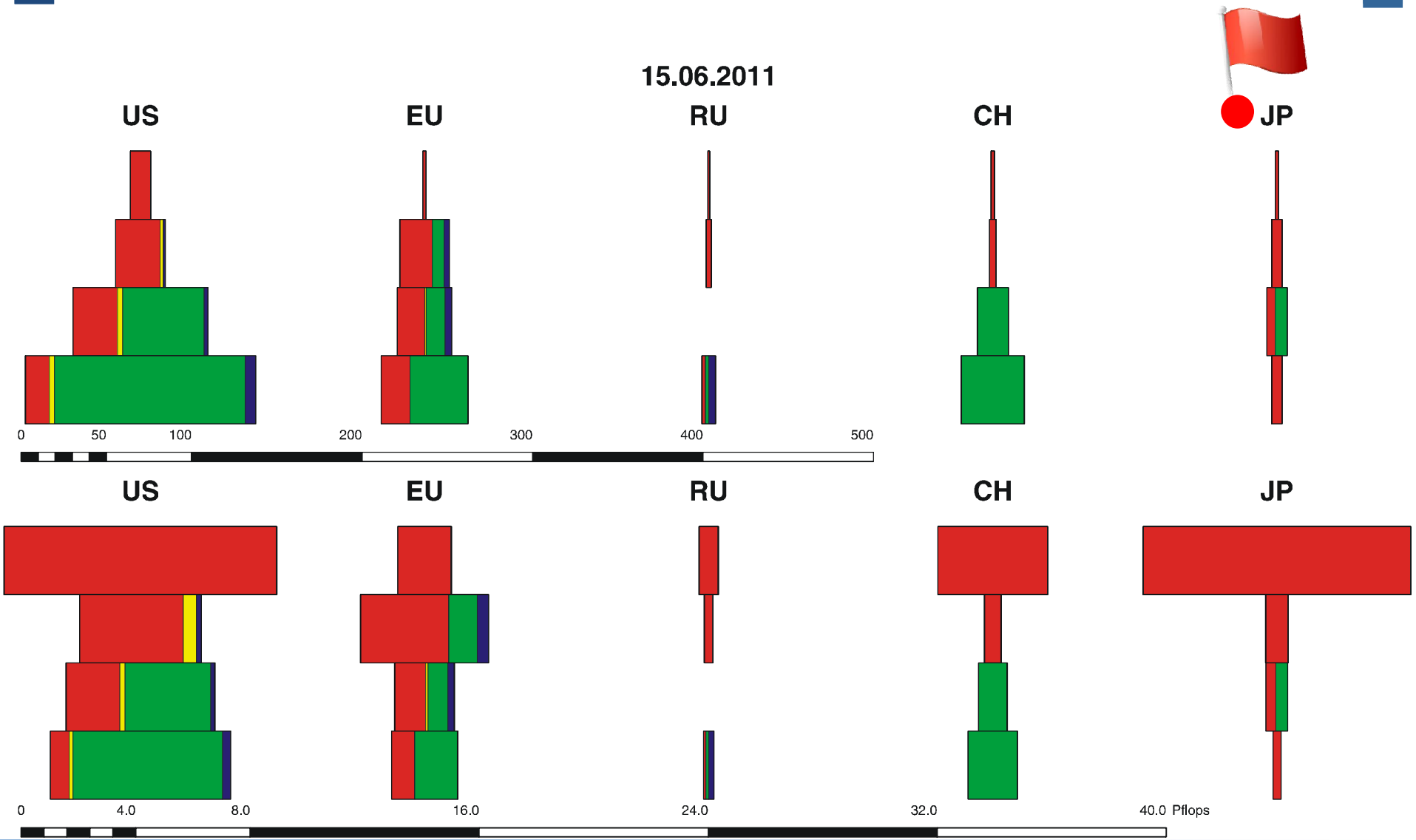
Киберинфраструктура ведущих стран



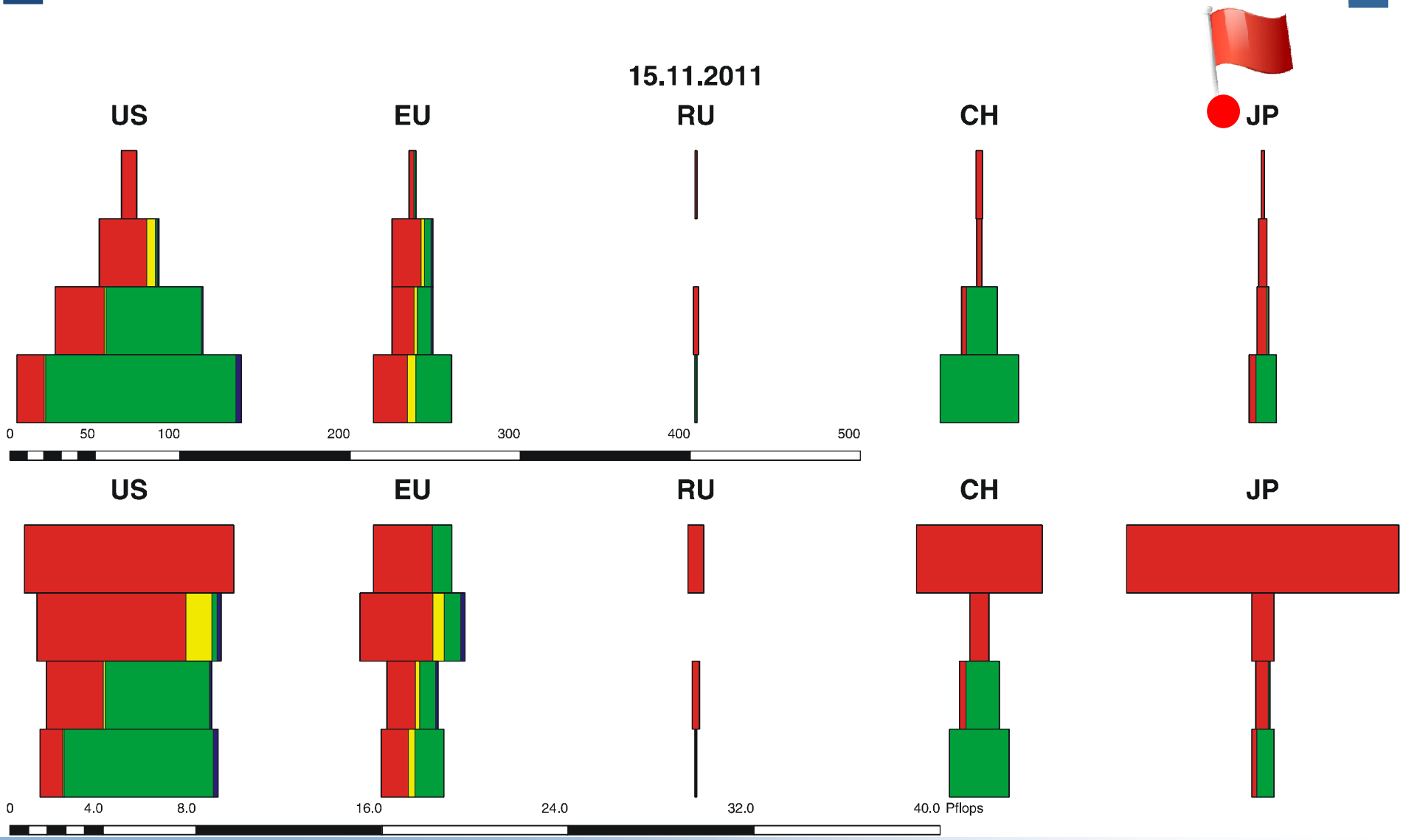
Киберинфраструктура ведущих стран



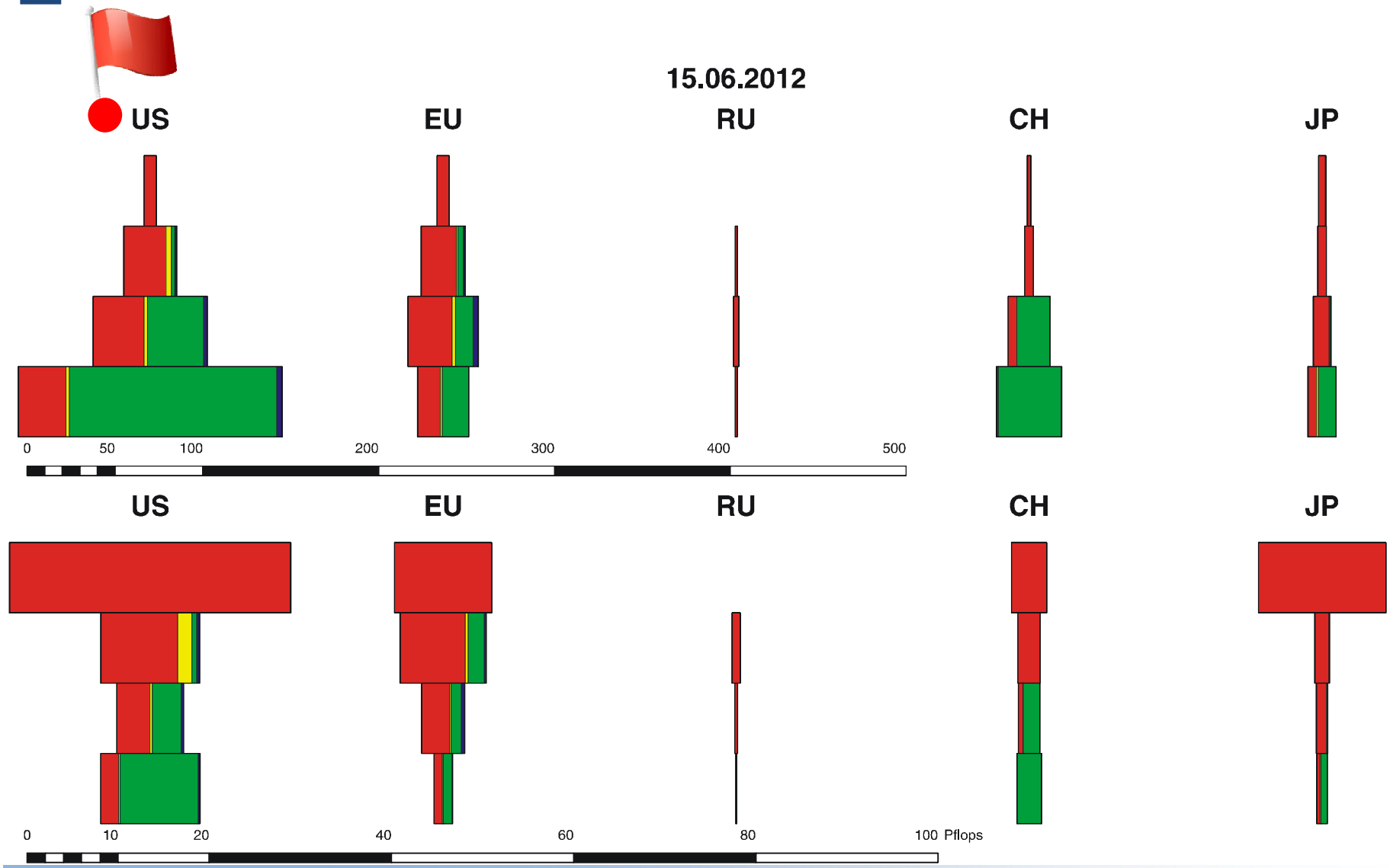
Киберинфраструктура ведущих стран



Киберинфраструктура ведущих стран



Киберинфраструктура ведущих стран



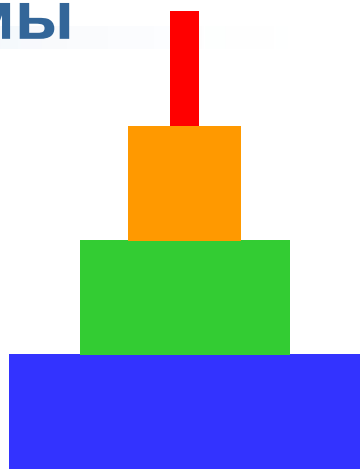
Страны Top1

1993	1994	1995	1996
1997	1998	1999	2000
2001	2002	2003	2004
2005	2006	2007	2008
2009	2010	2011	2012

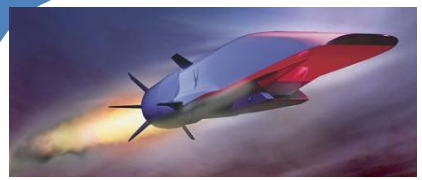
Системы Top1–10 драматически отличаются от всех остальных суперкомпьютеров

Иерархия киберинфраструктуры сверхвысокопроизводительные системы

Системы Top1–10 драматически отличаются от всех остальных суперкомпьютеров



Место в Top500	Rmax (Tflops)
Top1–10	16 325–1 271
Top11–20	1 243–1 035
Top21–100	919–173
Top101–250	173–87
Top251–500	86–61



до 13М

~1М

173–919 км/ч

61–173 км/ч

Top1–10: география

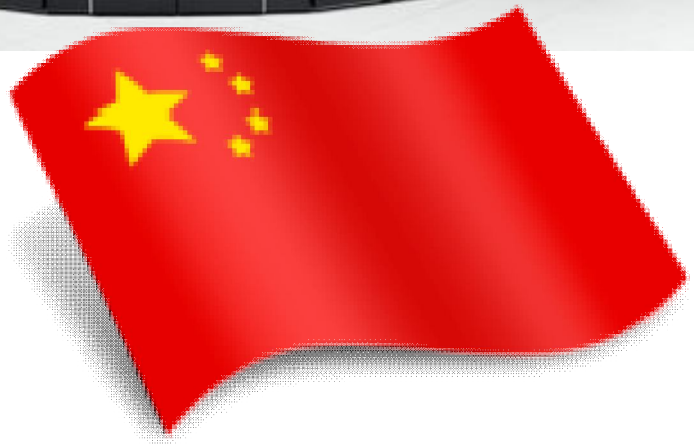
Количество и частота вхождений в Top1–10 за всю историю рейтинга Top500

- ☆ США 256 65.6%
- ☆ Японии 64 16.4%
- ☆ Европа(*) 53 13.6%
- ☆ Китай 13 3.3%
- ☆ Индия 2 0.5%
- ☆ Канада 2 0.5%

(*) Германия (22; 5.6%), Франция (11; 2.8%), Великобритания (10; 2.6%), Испания (5; 1.3%), Нидерланды (2; 0.5%), Швейцария (1; 0.3%), Швеция (1; 0.3%).

Китай: Sunway BlueLight MPP

- Сентябрь 2011
- ~8.6К CPU, ~137К ядер
- Водяное охлаждение (только)
- CPU: ShenWei SW1600
  - 64-разрядный RISC
  - 16 ядер, 1,1 ГГц
- ~1 Pflops — 1 МВт, сравните:
  - Tianhe-1A (GPU)
    4,7 Pflops — 4 МВт
  - Jaguar
    2.33 Pfllops — 7 МВт



Ноябрь 2011

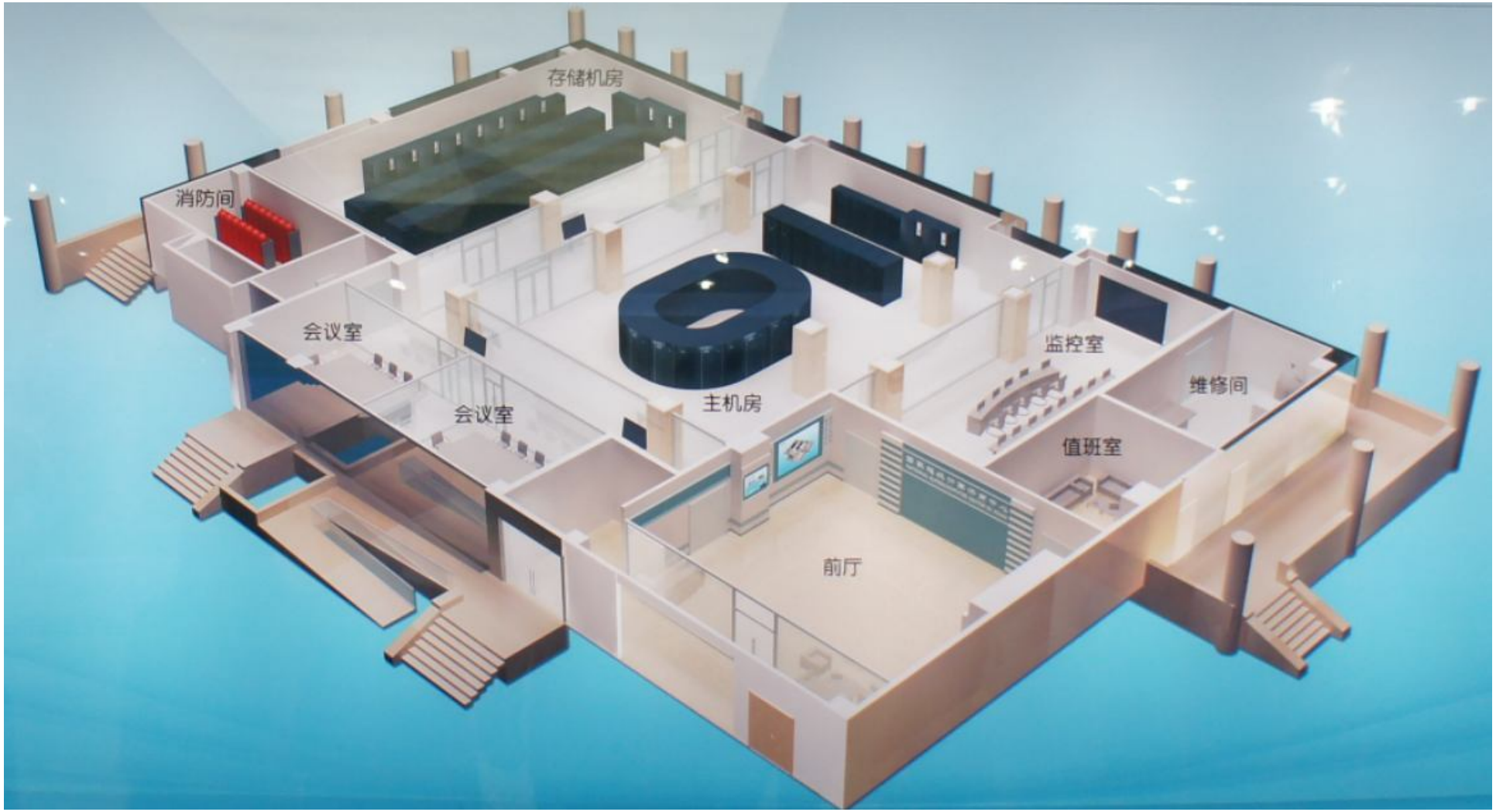
Место	Компьютер	Linpack Tflops	Пик Tflops	КПД	Кватт	Э/эффект. Tflops на Kwatt
14	Sunway Blue Light MPP	796	1,070	74.4%	1,074	0.74
18	Ломоносов	674	1,373	49.1%	2,500	0.27

Отставание

- КПД — в 1.5 раза
- Энергоэффективность — в 2.7 раз

Китай: Sunway BlueLight MPP

Госкорпорация INSPUR

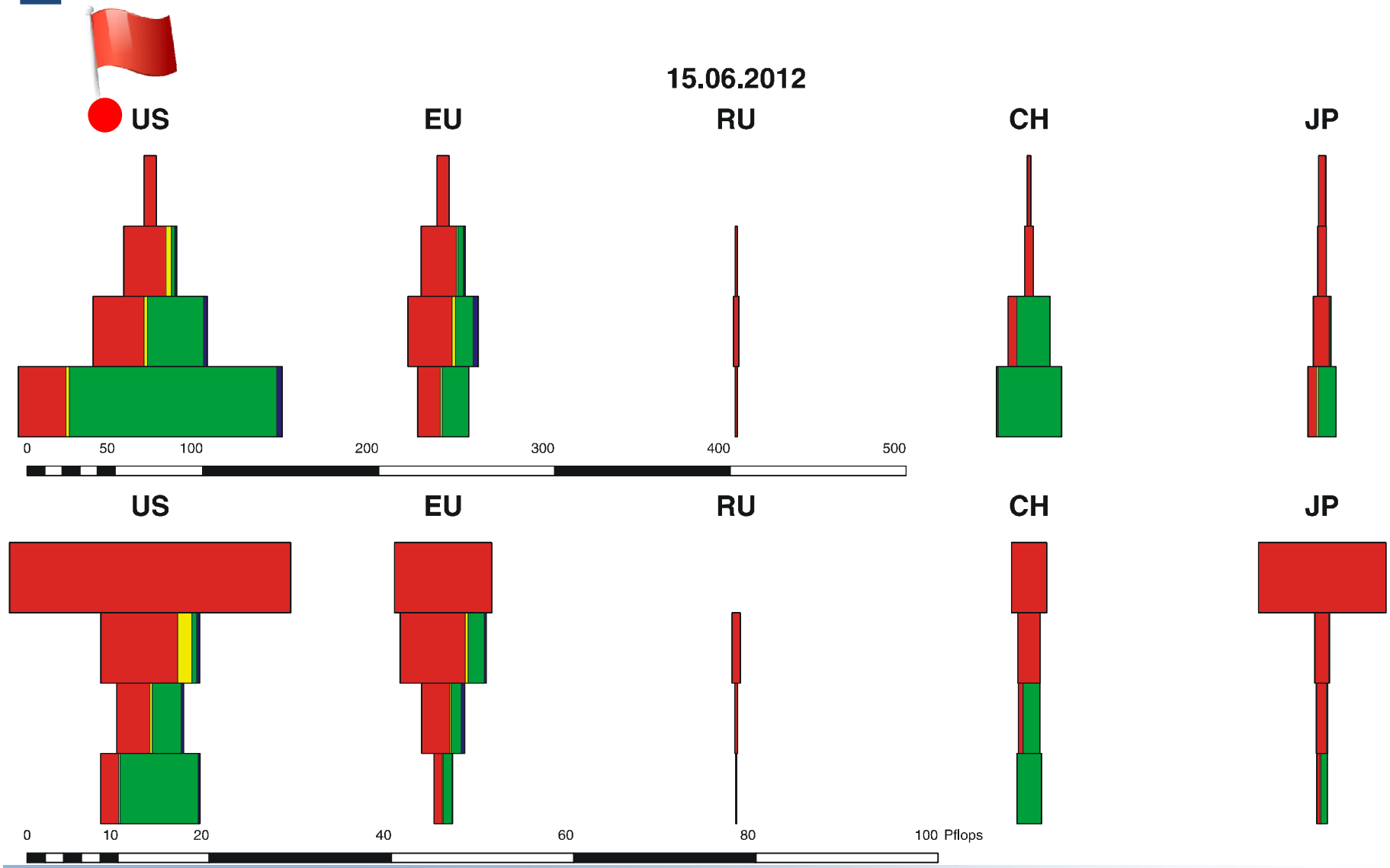


США: Sequoia, IBM BlueGene/Q

- Top1 — Июнь 2012
- 16.3/20.1 Pflops
- ~99К CPU, ~1.6М ядер
- Power BQC 16C
  - 64-разрядный RISC
  - 16 ядер, 1,6 ГГц
- **Интерконнект — свой**
- 7.9 МВт
- Менее 0.5 МВт на 1 Pflops

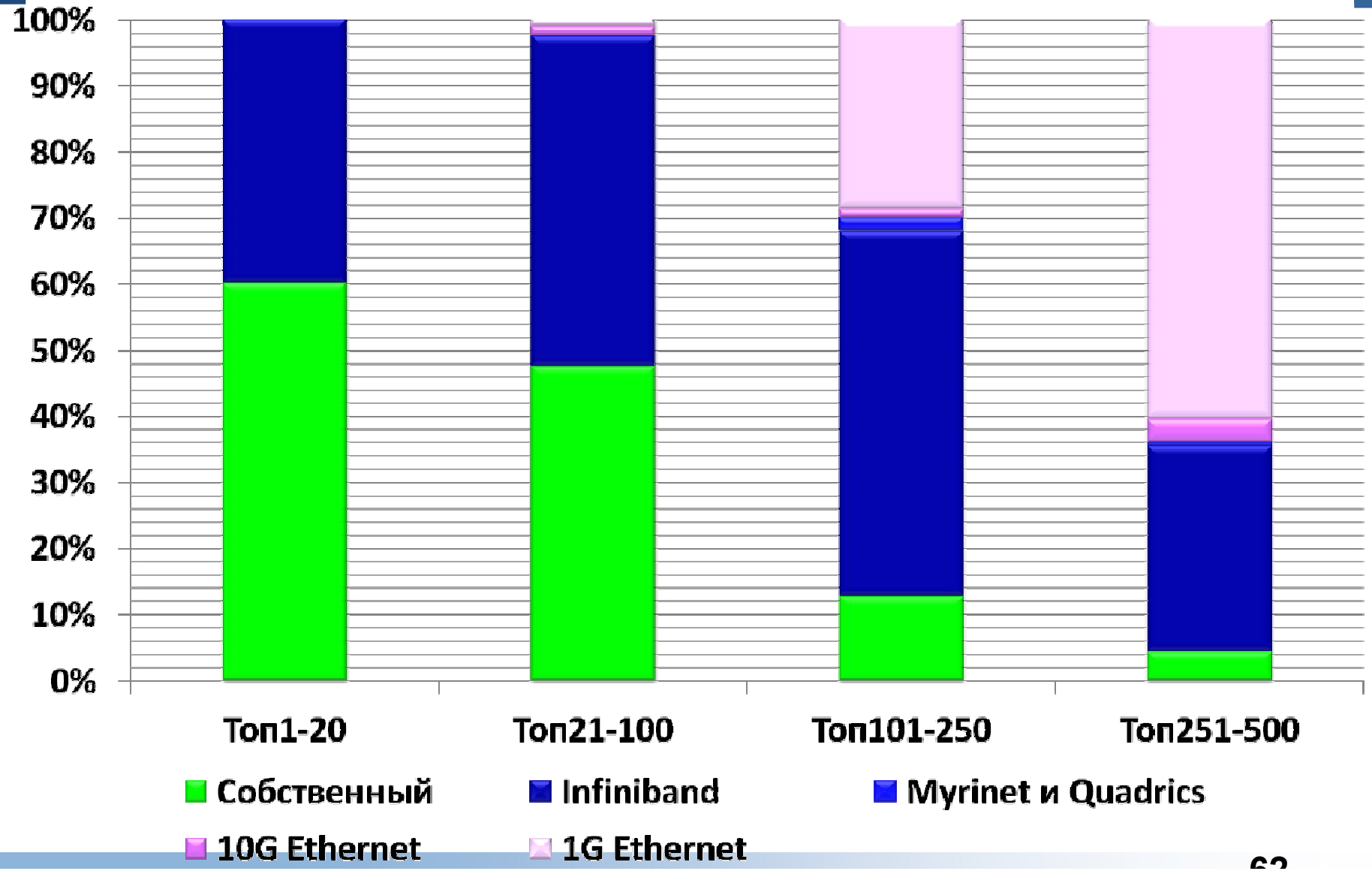
Области использования суперкомпьютеров «Правда, искажающая истину»

Киберинфраструктура ведущих стран



Системы класса «Top1–10» Interconnect «Правда, искажающая истину»

Interconnect в Top500 (июнь, 2012)



Российские разработки собственного интерконнекта

- ***Госкорпорация «Росатом» (РФЯЦ-ВНИИЭФ, Саров)***
  - СМПО — система межпроцессорных обменов
- ***НИЦЭВТ***
  - интерконнект 3D-тор (nD-тор)
- ***ИПМ имени М.В.Келдыша РАН, ООО НПО «Роста», ФГУП «НИИ Квант»***
  - МВС-Экспресс
    - стек ПО SKIF-3D-Torus
- ***ИПС имени А.К.Айламазяна РАН***
  - SKIF-3D-Torus
    - стек ПО SKIF-3D-Torus

SKIF-3D-Torus vs. Infiniband QDR (2009)

	SKIF-3D-Torus	Infiniband QDR
Пропускная способность сети (Гбит/сек/узел)	60	40
Темп выдачи сообщений (MT/s)	14	3
Задержка (мкс)	1.5	1.5

Архитектура стека ПО SKIF-3D-Torus

Прикладная программа

Коммуникационные библиотеки

SKIF-MPI | SKIF-SHMEM | SKIF-ARMCI | SKIF-GASNET

ALT Linux SKIF Cluster

SKIF-Driver

Коммуникационная библиотека нижнего уровня

SkifCh

Маршрутизатор системной коммуникационной сети

SKIF-3D-router (VHDL)

Следите за руками… Intel

- ***Покупки:***
  - CNews: **«Intel купила фирму InfiniBand у QLogic за $125M»**
  - http://ru.intel.com: **«Intel получит лучший суперкомпьютерный интерконнект» — $140M**
    - Приобретение ведущей в отрасли коммутационной технологии компании Cray, связанной с разработками интерконнектов Gemini (и перспективного Aries), для суперкомпьютеров Cray нового поколения —Cascade,— на базе CPU Intel® Xeon®
- ***Суть происходящего***
  - Начало R&D интерконнекта, интегрируемого с QPI

Системы класса «Тор1–10»
Система охлаждения

Ноябрь’2011 Top1–10: супер-HPC

1 — CPU «свой или чужой»
2 — охлаждение

				1	2	Total Cores	Acc. Cores	Rmax	Rpeak	КПД
1	K computer	JP	Fujitsu			705,024		10,510	11,280	93.17%
2	Tianhe-1A	CH	NUDT		?	186,368	100,352	2,566	4,701	54.58%
3	Jaguar	US	Cray			224,162		1,759	2,331	75.46%
4	Nebulae	CH	Dawning		½	120,640	64,960	1,271,000	2,984	42.59%
5	Tsubame 2.0	JP	NEC/HP			73,278	56,994	1,192,000	2,287,630	52.11%
6	Cielo	US	Cray			142,272		1,110,000	1,365,811	81.27%
7	Pleiades	US	SGI			111,104		1,088,000	1,315,328	82.72%
8	Hopper	US	Cray			153,408		1,054,000	1,288,627	81.79%
9	Tera-100	EU FR	Bull SA			138,368		1,050,000	1,254,550	83.70%
10	Roadrunner	US	IBM		?	122,400	110,160	1,042,000	1,375,776	75.74%

Июнь’2012 Top1–10: супер-HPC

1 — CPU «свой или чужой»
2 — охлаждение

				1	2	Total Cores	Acc. Cores	Rmax	Rpeak	КПД
1	Sequoia	US	IBM			1,572,864		16,325	20,133	81.09%
2	K computer	JP	Fujitsu			705,024		10,510	11,280	93.17%
3	Mira	US	IBM			786,432		8,162	10,066	81.09%
4	SuperMUC	EU DE	IBM			147,456		2,897	3,185	90.96%
5	Tianhe-1A	CH	NUDT		?	186,368	100,352	2,566	4,701	54.58%
6	Jaguar	US	Cray			298,592		1,941	2,628	73.87%
7	Fermi	EU IT	IBM			163,840		1,725	2,097	82.28%
8	JuQUEEN	EU DE	IBM			131,072		1,380	1,678	82.28%
9	Curie thin nodes	EU FR	Bull SA			77,184		1,359	1,667	81.52%
10	Nebulae	CH	Dawning		½	120,640	64,960	1,271	2,984	42.59%

Переход от воздушного охлаждения к жидкостному

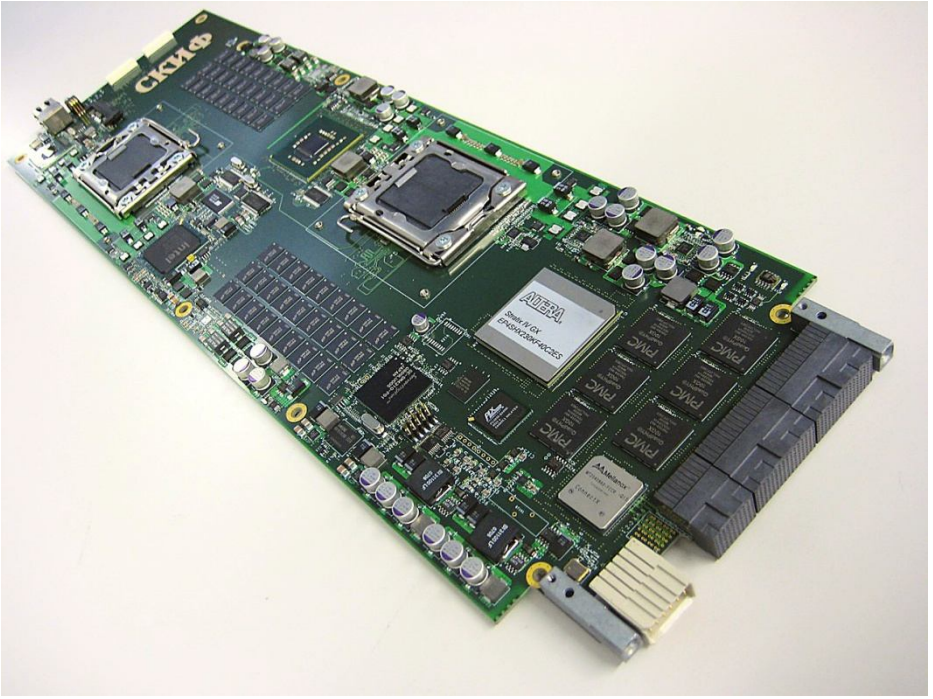
☆ ***Инженерная логика***

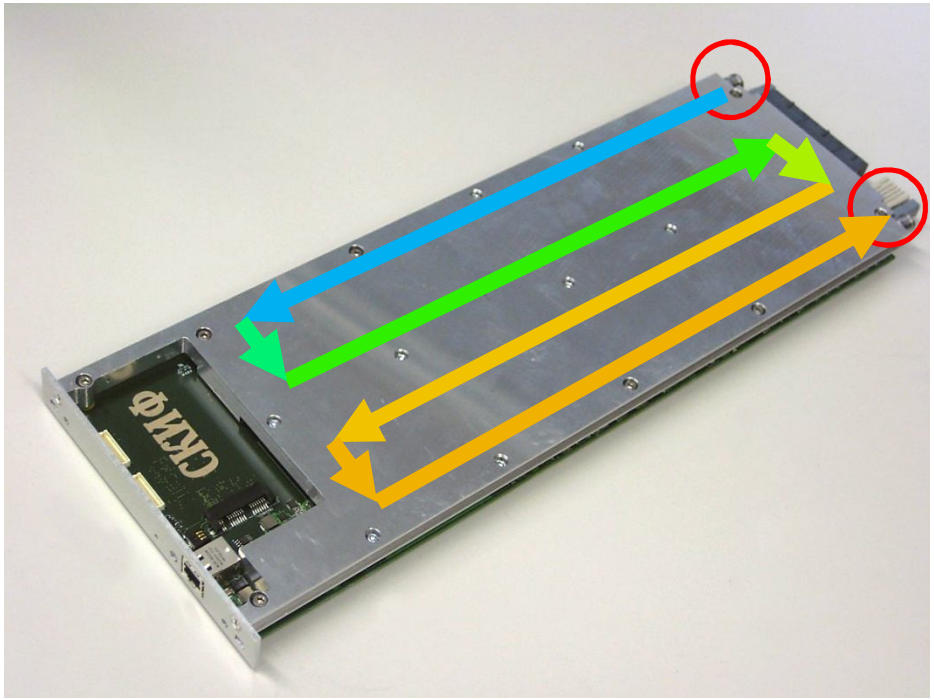
- Нужна большая производительность ➔
- Нужны меньшие задержки ➔
- Нужны более короткие проводники ➔
- Нужна более плотная упаковка электроники ➔
- Будет более высокая объемная плотность тепловыделения ➔
- Нужна более эффективная система охлаждения электроники

☆ ***Эпохи***:

- воздух ➔
- жидкость на уровне шкафа, воздух у печатных плат ➔
- жидкость на уровне печатной платы ➔
- использование фазового перехода ➔

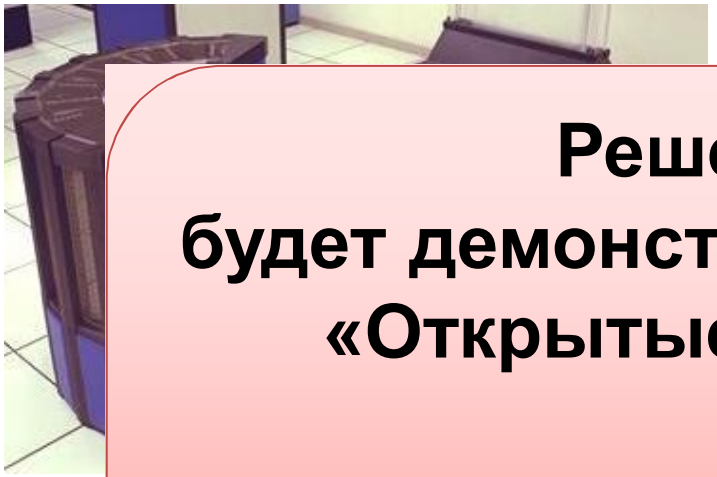
Охлаждающие пластины.
Пример: Узел суперкомпьютера СКИФ-Аврора



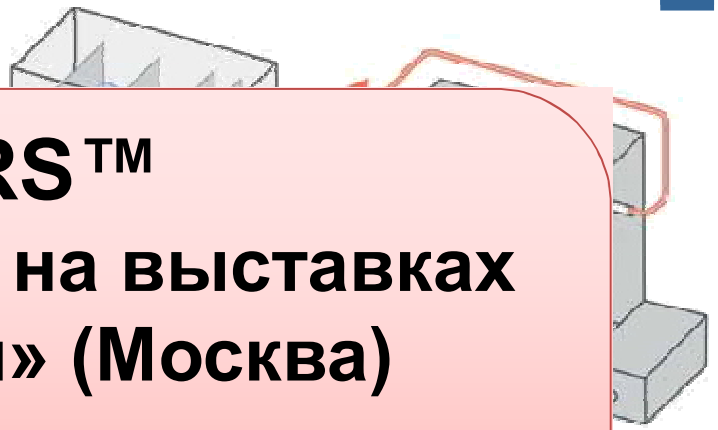


Все мы вышли из воды и должн
туда же вернуться. В эту сре
к/ф «Национальн
особенно
рыбалк

Непосредственное (погружное) жидкостное охлаждение



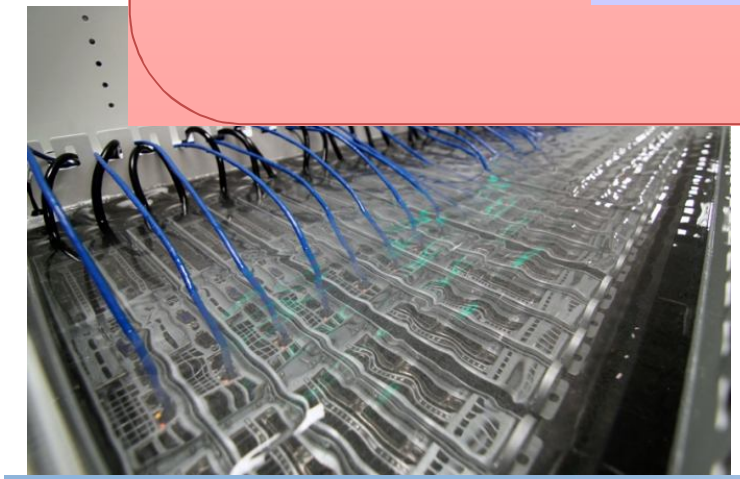
Cray-2, 1985,

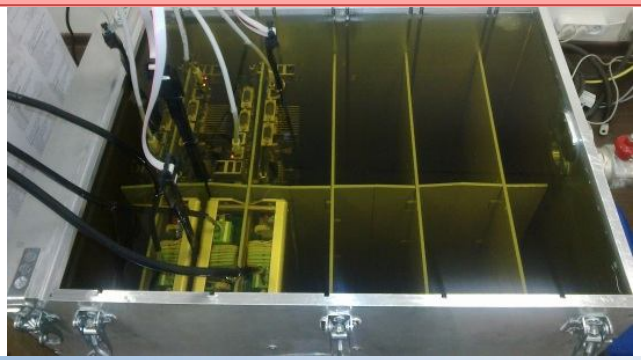


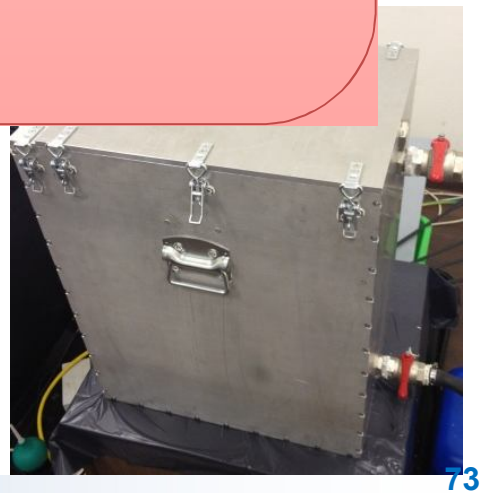
Решение IMMERS™
будет демонстрироваться на выставках
«Открытые инновации» (Москва)
и
НСКФ’2012 (Переславль-Залесский)
www.hpc-platform.ru

Gree

на







Системы класса «Top1–10»
Суммируя специфику

	← Вычислительные системы сверхвысокой производительности	Остальные суперкомпьютеры (системы высокой производительности) →
	Top1 ← Top10 ← Top20 … Top100 →	Top250→Top500
Решаемые задачи	Наука: фундаментальные и задельные исследования, сложнейшие НИР	Приложения в отраслях реальной экономики, инженерные расчеты, ОКР
Спектр задач	Широкий спектр задач — он характерен для научного поиска	Задачи, специфичные для отрасли
Универсальность системы	Чаще — универсальные системы. Спецвычислители лучше реконфигурируемые	Универсальные системы и спецвычислители — под специфику задач отрасли
Реализация	Только отечественная разработка, возможно на зарубежной элементной базе	Либо отечественная разработка, либо закупка зарубежной системы
Используемые решения	Для многих подсистем нет доступных решений — нужна собственная разработка	Для всех подсистем (и системы в целом) доступны готовые решения
Расходы на	Из бюджета государства	Из средств

Объективный размер бедствия

Ноябрь 2011: У них и у нас

	Количество суперкомпьютеров				
	US	EU	RU	CH	JP
Уровень 1	9	4	1	4	2
Уровень 2	35	24	—	3	5
Уровень 3	87	24	3	21	7
Уровень 4	132	46	1	46	16
Всего	**263**	**98**	**5**	**74**	**30**

	Суммарная производительность суперкомпьютеров на тесте Linpack, Tflops				
	US	EU	RU	CH	JP
Уровень 1	8,988.6	3,367.1	674.1	5,404.6	11,702.0
Уровень 2	7,920.0	4,517.3	—	822.7	958.1
Уровень 3	7,110.4	2,197.6	309.1	1,710.1	611.6
Уровень 4	7,655.4	2,695.9	59.9	2,552.1	952.5
Всего	31,674.5	12,778.0	1,043.1	10,489.5	14,224.1

Июнь 2012: У них и у нас

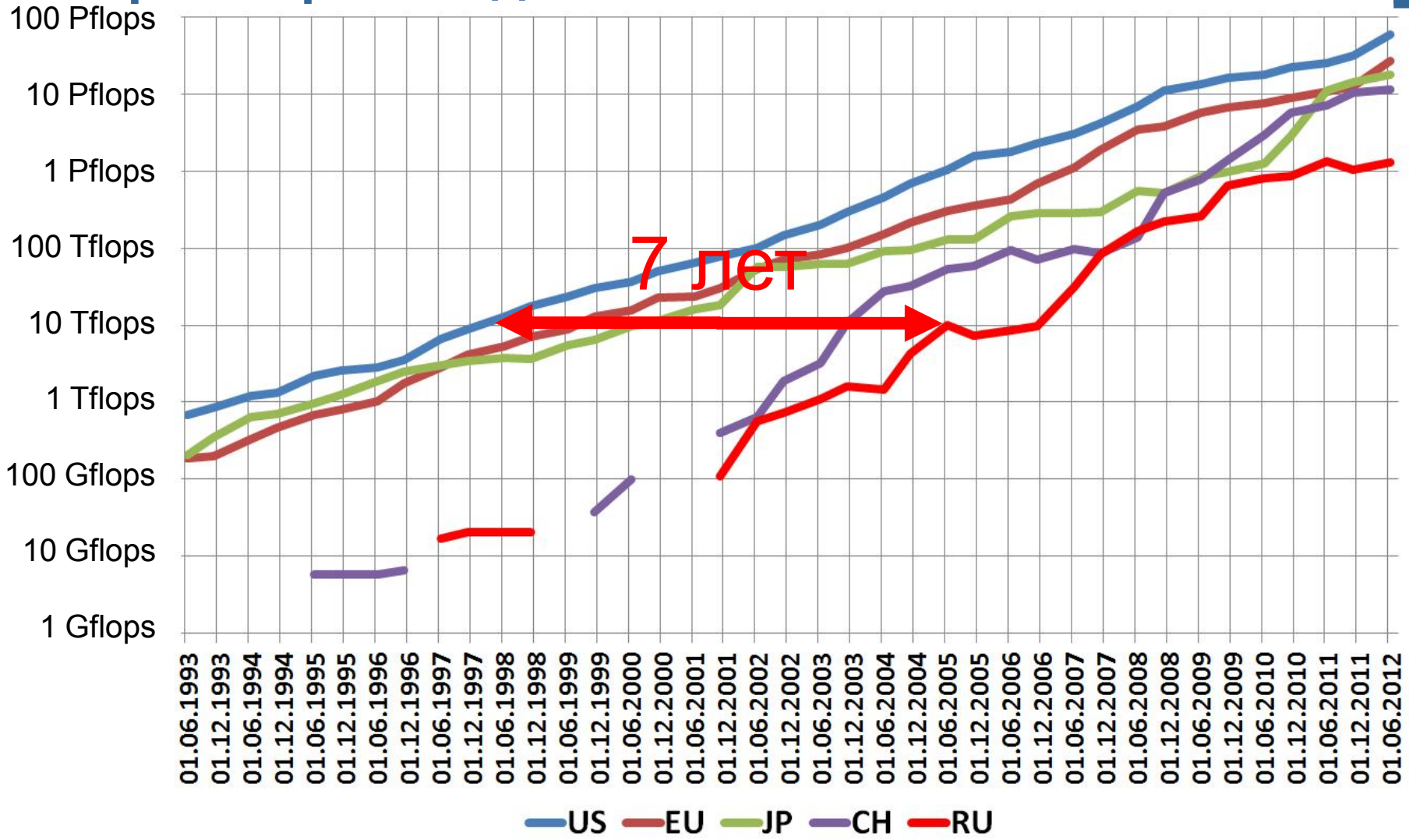
	Количество суперкомпьютеров				
	US	EU	RU	CH	JP
Уровень 1	7	7	0	2	4
Уровень 2	30	25	1	5	5
Уровень 3	65	40	3	24	10
Уровень 4	150	29	1	37	16
Всего	**252**	**101**	**5**	**68**	**35**

	Суммарная производительность суперкомпьютеров на тесте Linpack, Tflops				
	US	EU	RU	CH	JP
Уровень 1	30,877.1	10,655.0	0.0	3,837.0	13,982.0
Уровень 2	10,839.9	9,433.8	901.9	2,456.4	1,614.0
Уровень 3	7,346.7	4,724.3	321.5	2,384.0	1,244.3
Уровень 4	10,911.8	2,045.6	69.6	2,719.7	1,170.9
Всего	59,975.5	26,858.8	1,293.0	11,397.1	18,011.1

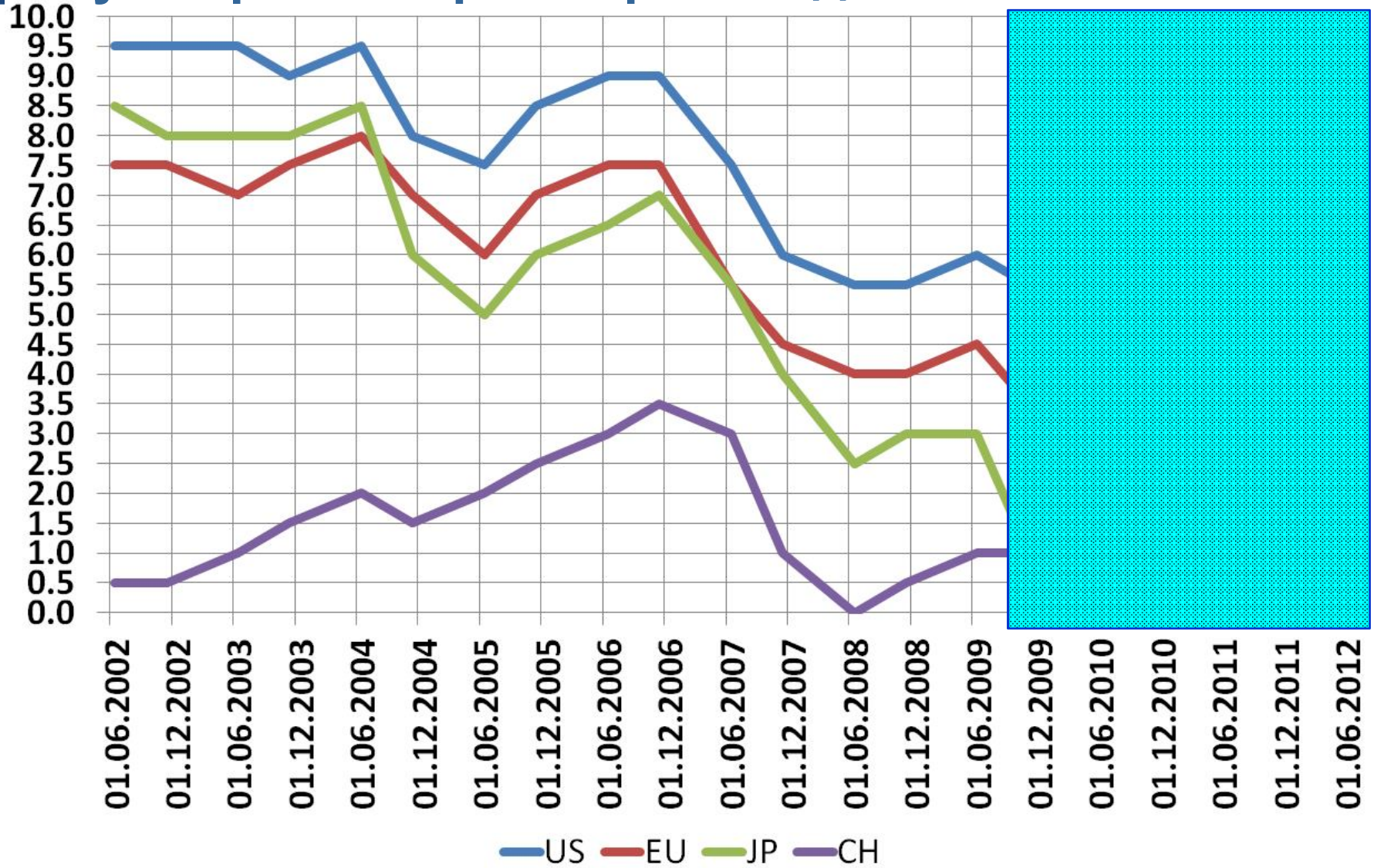
Июнь 2012: Отставание (в разы)

	US	EU	CH	JP
Уровень 1	—	—	—	—
Уровень 2	12.02	10.46	2.72	1.79
Уровень 3	22.85	14.69	7.41	3.87
Уровень 4	156.82	29.40	39.09	16.83
(2012/06) Всего	**46.38**	**20.77**	**8.81**	**13.93**
(2011/11) Всего	**30.37**	**12.25**	**10.06**	**13.64**

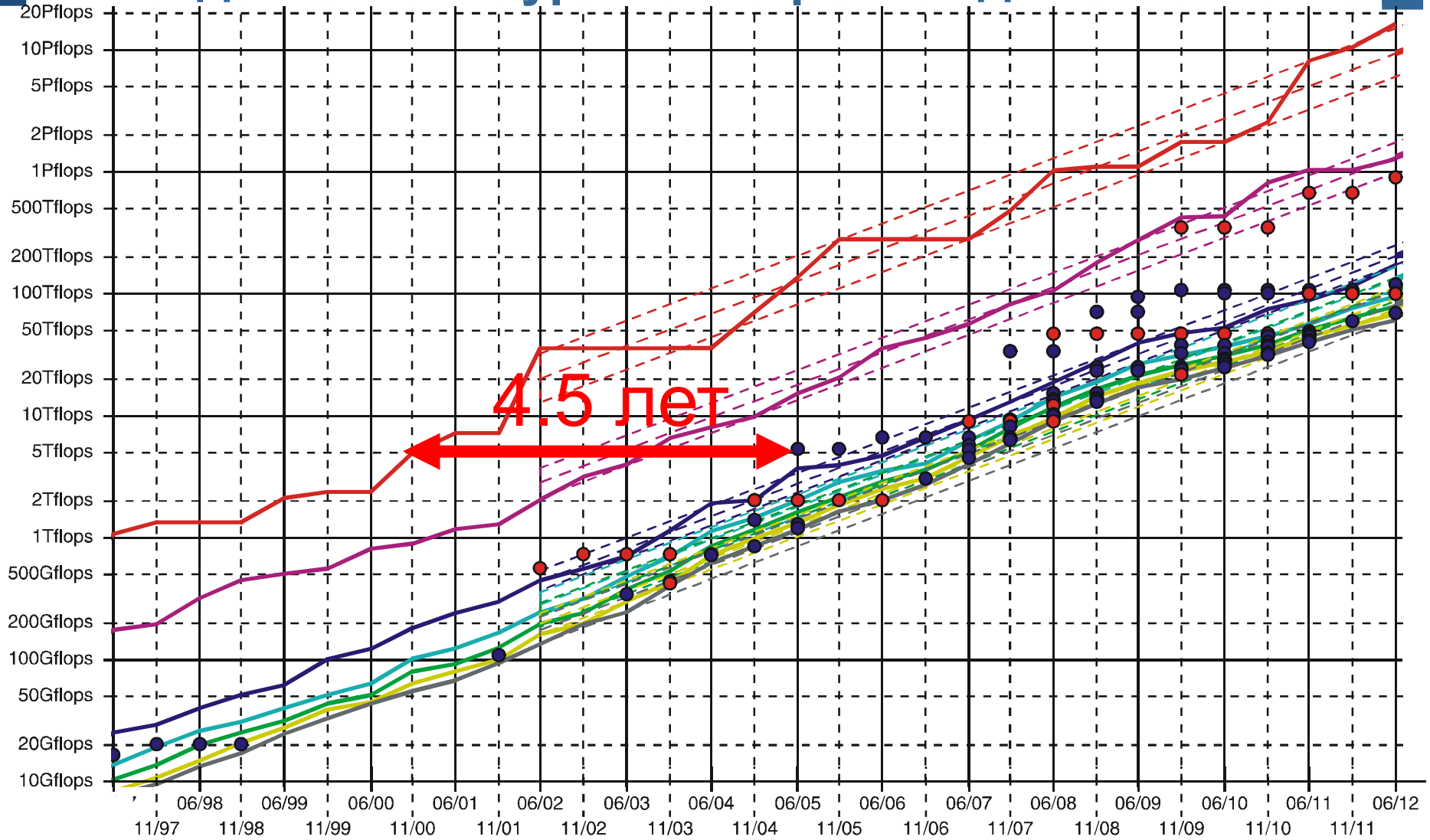
Объективный размер бедствия. Суммарная Linpack-производительность



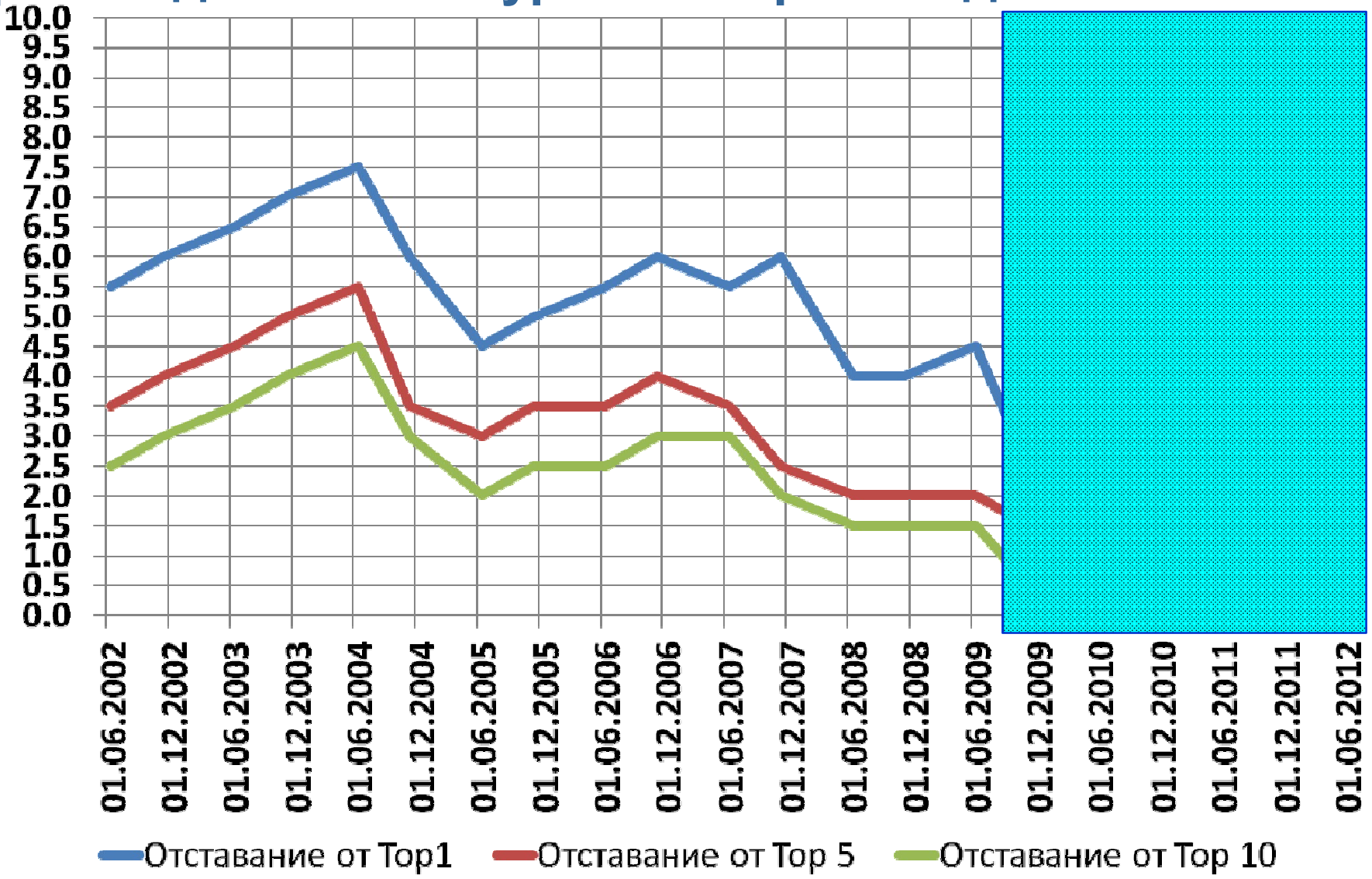
Отставание в годах
Суммарная Linpack-производительность



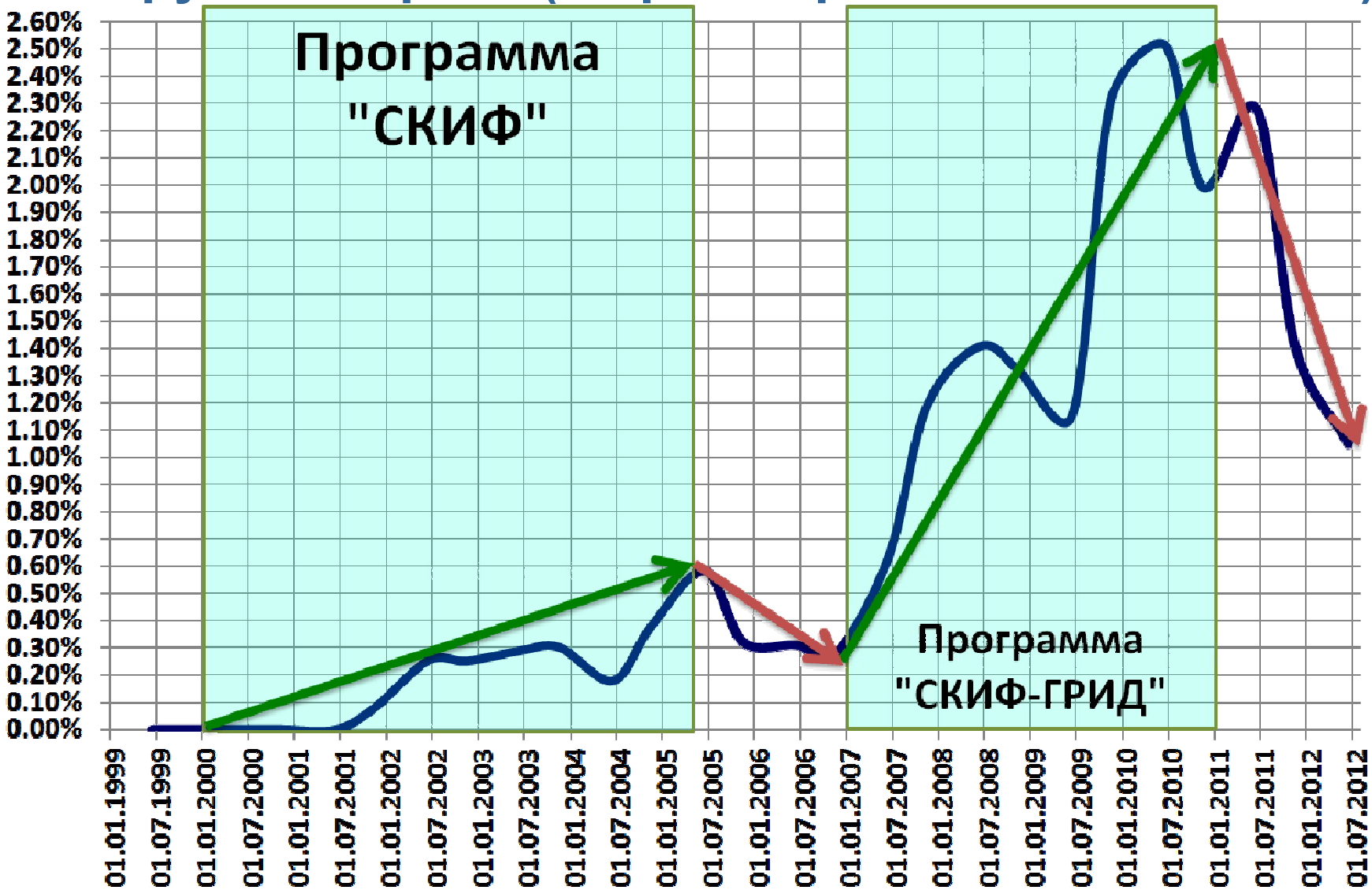
Объективный размер бедствия. Отставание выхода на некий уровень производительности



Объективный размер бедствия. Отставание выхода на некий уровень производительности



Объективный размер бедствия. Доля России и Беларуси в Top500 (Linpack-производительность)



Последний год: пессимизм прессы

- Россия «сдулась» в мировом рейтинге суперкомпьютеров
- Российские суперкомпьютеры провалились в рейтинге энергоэффективности
  - «Ломоносов» потерял 106 позиций в Green500 — 153 место
  - Лучшая российская установка в Green500: СКИФ-Аврора — 109 место
  - Самая худшая: HP, МСЦ РАН — 416 место
- IBM и HP захватили суперкомпьютерный рынок в России
- «Т-Платформы» стали убыточными — 03.04.2012

Суперкомпьютерная отрасль России

Российская академия наук

- За всю историю в различные выпуски мирового рейтинга Топ500 вошли только девять суперкомпьютеров отечественной разработки и ***в восьми случаях из девяти*** в их создании участвовали — а во многих случаях были головными разработчиками,— академические институты
  - ***ИПМ имени М.В. Келдыша РАН и МСЦ РАН*** — суперкомпьютер МВС-1000 (2002);
  - ***ИПС имени А.К. Айламазяна РАН*** — суперкомпьютеры СКИФ К-500 (2003), СКИФ К-1000 (2004), СКИФ-Cyberia (2007), СКИФ МГУ (2008), СКИФ Урал (2008), СКИФ-Аврора модель 4/N (2010), СКИФ-Аврора модель 4/W (2011).

Российская академия наук

§1. В мировой практике самые мощные суперкомпьютеры (Top1–10) в большинстве случаев ориентированы на фундаментальные и задельные исследования. Тем самым, ***оправдано создание такого класса техники именно на базе Российской академии наук***

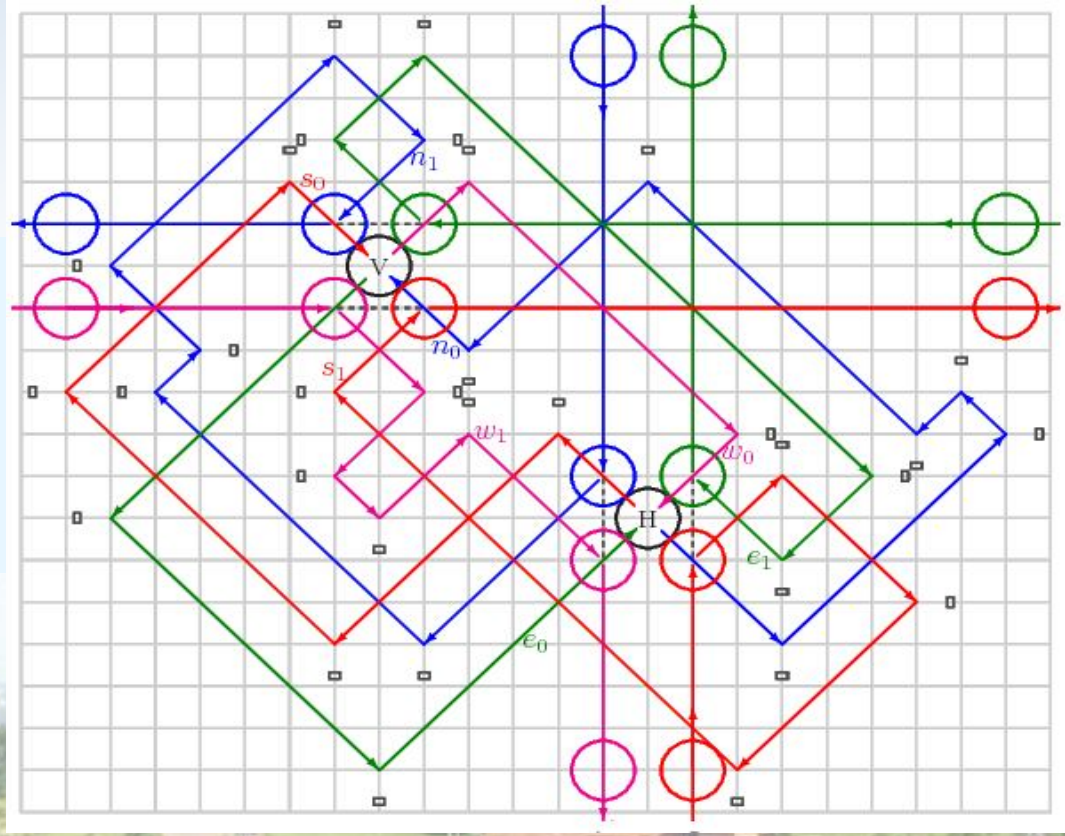
§2. ***Российская академия действительно занимает лидирующие позиции*** в отечественной суперкомпьютерной отрасли, что еще не до конца осознано обществом

§3. ***Системы уровня Top1–10 не могут быть приобретены за рубежом или созданы за счет «отверточной сборки»***.

На что настроены суперкомпьютеры и где не могут быть использованы

- На сеточные модели и визуализация (нефтегаз, метеорология, оборона, ...)
- Сетки даже грубой силой подошли к пределу неадекватности (предположение: непрерывность; на самом деле дискретность, молекулы)
- Страшные затраты на инфраструктуру (коннект, охлаждение)
- Беспомощность в информационном поиске
- Плохая работа со сложными данными
- Две бетонные стенки: предел Ландауэра и скоросмть света
- Вторую на козле не объедкешь; первую можно попытаться

Реверсивные вычисления: обзор мирового опыта



А.Н. Непейвода

Институт Программных Систем

им. А.К.Айламазяна РАН

Принцип Ландауэра (1961) — фон Неймана (1949)

Пусть логическое состояние системы требует вычисления двух других состояний, каждое из которых может принимать N значений. Значит, состояние новой системы будет принимать 2N значений.

Энтропия (S) возрастает на log 2, и с учётом

$$T = \frac{dE}{dS}$$

теряемая энергия составляет

$$E_{diss} \geq T \times k_B \times \ln 2$$

Споры вокруг Ландауэра

Доказательство — Plenio, Vitelli (2001)

Опровержение — Norton (2004)

Доказательство — Andrieux, Gaspard (2007)

Доказательство — Ladyman et al (2007)

Опровержение — Norton (2011)

- Итоги:
- Конец бинарной парадигмы в информатике
- Специализация вместо универсализации в 20%, орангизация в 80%
- Конец эпохи грубой силы
- Конец мыльному пузырю квантовых компьютеров
- Автономизация вместо глобализации

Спасибо за внимание!

? — !