

CMS ANALYSIS JOBS AT US AND BRAZILIAN TIER-2 SITES

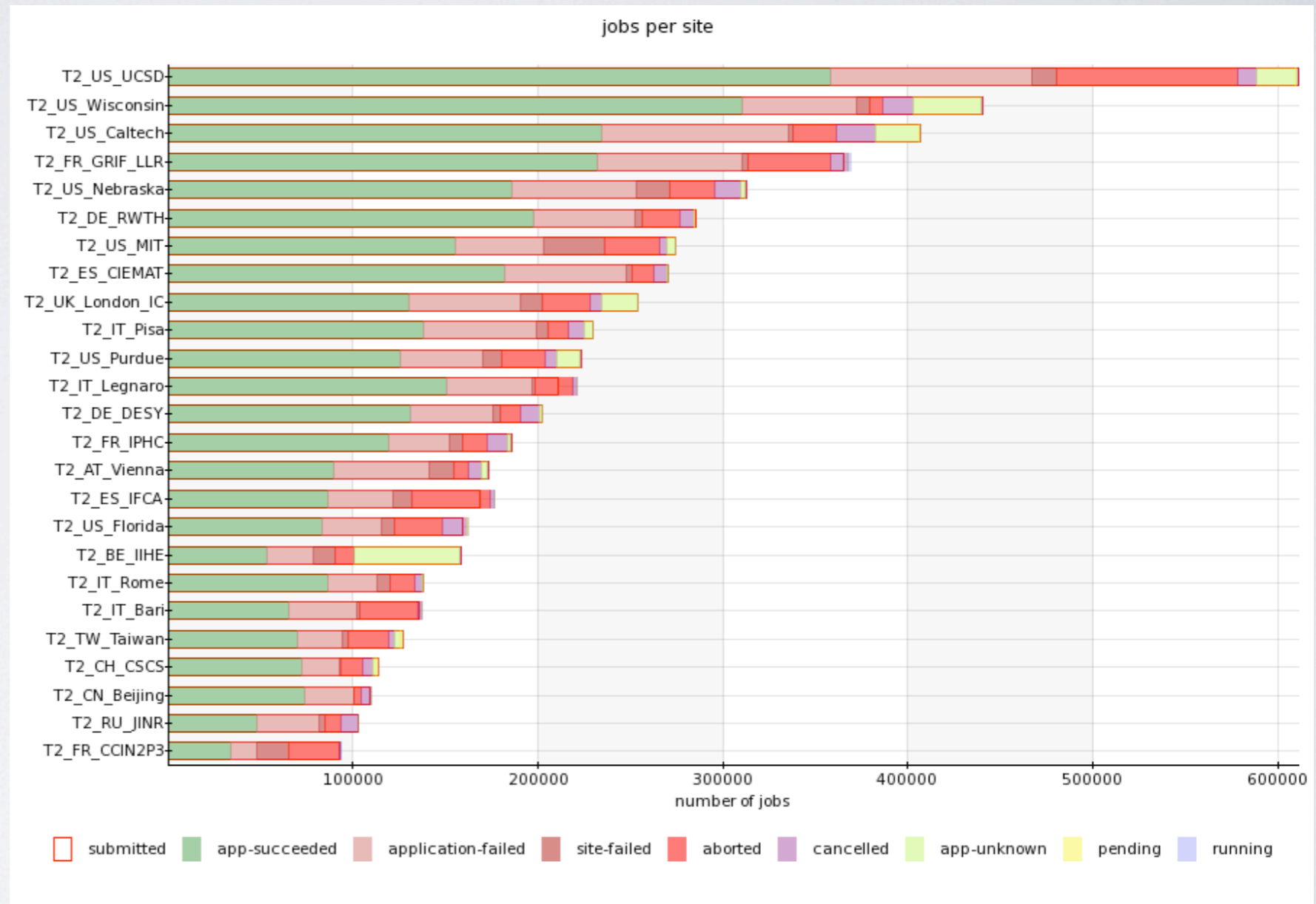
JAMES LETTS

**T2_US_UCSD,
DDT & ANALYSIS-OPS**

**CMS DASHBOARD STATISTICS
FROM OCT 1, 2009 THROUGH
FEB 23, 2010**

CMS ANALYSIS JOBS

- FROM OCT 2009 THRU FEB 2010
- OVER 2.5 MILLION JOBS SUBMITTED TO US & BRAZILIAN TIER-2's, 38% OF WORLD-WIDE TOTAL



CMS ANALYSIS JOBS AT T2_US

	SUCCESS RATE	CPU SEC	JOB SLOTS
T2_US	72%	9.6E+9 29%	1049 27%
ALL T2	70%	3.26E+10	3873

- SUCCESS RATES ARE TYPICAL OF CMS JOBS.
- PERCENTAGE OF TOTAL RESOURCES DELIVERED WORLD-WIDE IS SIGNIFICANT

CMS ANALYSIS JOBS - ERRORS

SITE FAIL

- 6.5% OF JOBS AT T2_US AND T2_BR
- WORLD-WIDE, 4.6%
- MAJOR SITE-FAIL MODE:
 - 72% WERE ERROR 8020 FILE OPEN ERROR (WORLD-WIDE)

APP FAIL

- 23% OF JOBS AT T2_US AND T2_BR
- WORLD-WIDE, 25%
- MAJOR APP-FAIL MODE:
 - STAGE OUT: 53%, I.E. ABOUT 12% OF U.S. JOBS FAIL REMOTE STAGE OUT

RESPONSES TO TICKETS

- RESPONSE TIME TO ANALYSISOPS SAVANNAH TICKETS

- OCTX HAD 13 ANALYSISOPS TICKETS FOR US/BR SITES DURING 7 DAYS (MOSTLY FILE READ PROBLEMS, ERROR 8020)

- MEDIAN INITIAL RESPONSE TIME TO THE TICKET OF OUR SITES DURING OCTX WAS 2H50M

- CONSIDERING THAT MOST TICKETS WERE ORIGINATED DURING WORKING HOURS AT CERN, THIS IS IMPRESSIVE.

- SHORTEST RESPONSE WAS 5 MINS

- LONGEST WAS 82 DAYS (ONLY TICKET OVER 13H, FELL OFF THE RADAR...)

- ONLY 3 TICKETS SINCE OCTX, RESPONSES: 19M, 2H14M, 81H5M.

LCG Savannah

Logged in as jletts

My Incoming Tickets

My Items

My Groups

My Account Conf

Logout

CMS Computing Infrastructure Support - Support: Browse Items

Main Docs Support Source Code Tasks News

(-) Display Criteria

Item ID	Summary	Category	Open/Closed	Priority	Assigned To	Submitted By	Submitted On
#112019	Possible Software Publish Issue or Black Analysis	Analysis	Open	5 - Normal	cmscompinfrasup- t1usfnal	fi...	2010-02-24 11:48
#112816	CRAB job stageout rate need to be	Analysis	Open	5 - Normal	cmscompinfrasup- t2usnebraska	bockjoo	2010-02-17 08:03
#112717	jobs at FNAL LPC: Command not found: client not installed or wrong environment	Analysis Operations	Closed	5 - Normal	cmscompinfrasup- t1usfnal	dlevans	2010-02-12 02:25
#112439	use MMS phedex node for /store/result	Analysis	Closed	5 - Normal	kress	belforte	2010-01-30 13:13
#112400	Many CRAB jobs failing with FileOpenError 8020 at T2_US_Nebraska	Analysis Operations	Closed	5 - Normal	cmscompinfrasup- t2usnebraska	jletts	2010-01-27 11:35
#111931	Problem with CMSSW_3_2_2	Analysis	Closed	5 - Normal	pfeiffer	szillasi	2009-12-16 12:22
#111881	figure out why jobs are sent to Lyon T2 instead of T1 etc.	Analysis Operations	Open	5 - Normal	belforte	belforte	2009-12-14 08:21
#111502	Subscribed data sets with no replicas at	Analysis Operations	Closed	5 - Normal	cmscompinfrasup- datatransfer	jletts	2009-11-25 10:47

Search

in Projects

Hosted Projects

Register New Project

Full List

Contributors Wanted

Statistics

Site Help

User Docs

User Docs: In Depth

Guide

Get Support

CRAB JOB STAGE OUT

- SINGLE LARGEST SOURCE OF JOB ERROR IS STAGE OUT FAILURE.
- **CRAB_2_7_1** WILL HAVE LOCAL STAGE OUT FALLBACK, MUCH LIKE PA NOW HAS.
- FALLBACK **/store/temp/user** DEFINED BY TFC
- STAGE OUT METHOD DEFINED BY THE **site-local-config.xml**

TESTS OF /STORE/TEMP/USER

- NOW PART OF THE SAM TESTS

- <http://tinyurl.com/yzeov23>

- TESTS WRITABILITY OF /store/temp/user AT ALL SITES

T2_IT_Pisa	SRMv2	cmsdcache.pi.infn.it	ok	ok	ok
		gridsrm.pi.infn.it	error	warn	error
T2_IT_Rome	SRMv2	cmsrm-se01.roma1.infn.it	ok	ok	ok
T2_KR_KNU	SRMv2	cluster142.knu.ac.kr	ok	ok	ok
T2_PL_Warsaw	SRMv2	se.polgrid.pl	ok	ok	ok
T2_PT_LIP_Lisbon	SRMv2	srm01.lip.pt	ok	ok	ok
T2_RU_IHEP	SRMv2	se0003-m45.ihep.su	ok	ok	ok
T2_RU_IHEP	SRMv2	se0003-m45.ihep.su	ok	ok	ok
T2_RU_JINR	SRMv2	lcsedc01.jinr.ru	ok	ok	ok
T2_RU_PNPI	SRMv2	cluster.pnpi.nw.ru	ok	ok	ok
T2_RU_RRC_KI	SRMv2	se.grid.kiae.ru	ok	ok	error
T2_RU_RRC_KI	SRMv2	se.grid.kiae.ru	ok	ok	error
T2_TR_METU	SRMv2	eymir.grid.metu.edu.tr	ok	ok	ok
T2_TW_Taiwan	SRMv2	f-dpm001.grid.sinica.edu.tw	ok	ok	ok
T2_UK_London_Brunel	SRMv2	dc2-grid-64.brunel.ac.uk	ok	ok	ok
T2_US_Caltech	SRMv2	cit-se.ultralight.org	ok	ok	ok
T2_US_Florida	SRMv2	srm.ihepa.ufl.edu	ok	ok	ok
		srmb.ihepa.ufl.edu	ok	ok	ok
T2_US_MIT	SRMv2	se01.cmsaf.mit.edu	ok	ok	ok
T2_US_Nebraska	SRMv2	dcache07.unl.edu	ok	ok	ok
		srm.unl.edu	ok	ok	ok
T2_US_Purdue	SRMv2	srm-dcache.rcac.purdue.edu	ok	ok	ok
T2_US_UCSD	SRMv2	bsrm-1.t2.ucsd.edu	ok	ok	ok
T2_US_Wisconsin	SRMv2	cmssrm.hep.wisc.edu	ok	ok	ok

TESTS OF /STORE/TEMP/USER

- ANALYSIS OPS TEAM IS ALSO TESTING ACCESS TO THE **/store/temp/user** AREAS.

- AS PART OF THE VALIDATION AND TESTING OF **CRAB_2_7_1_pre11**

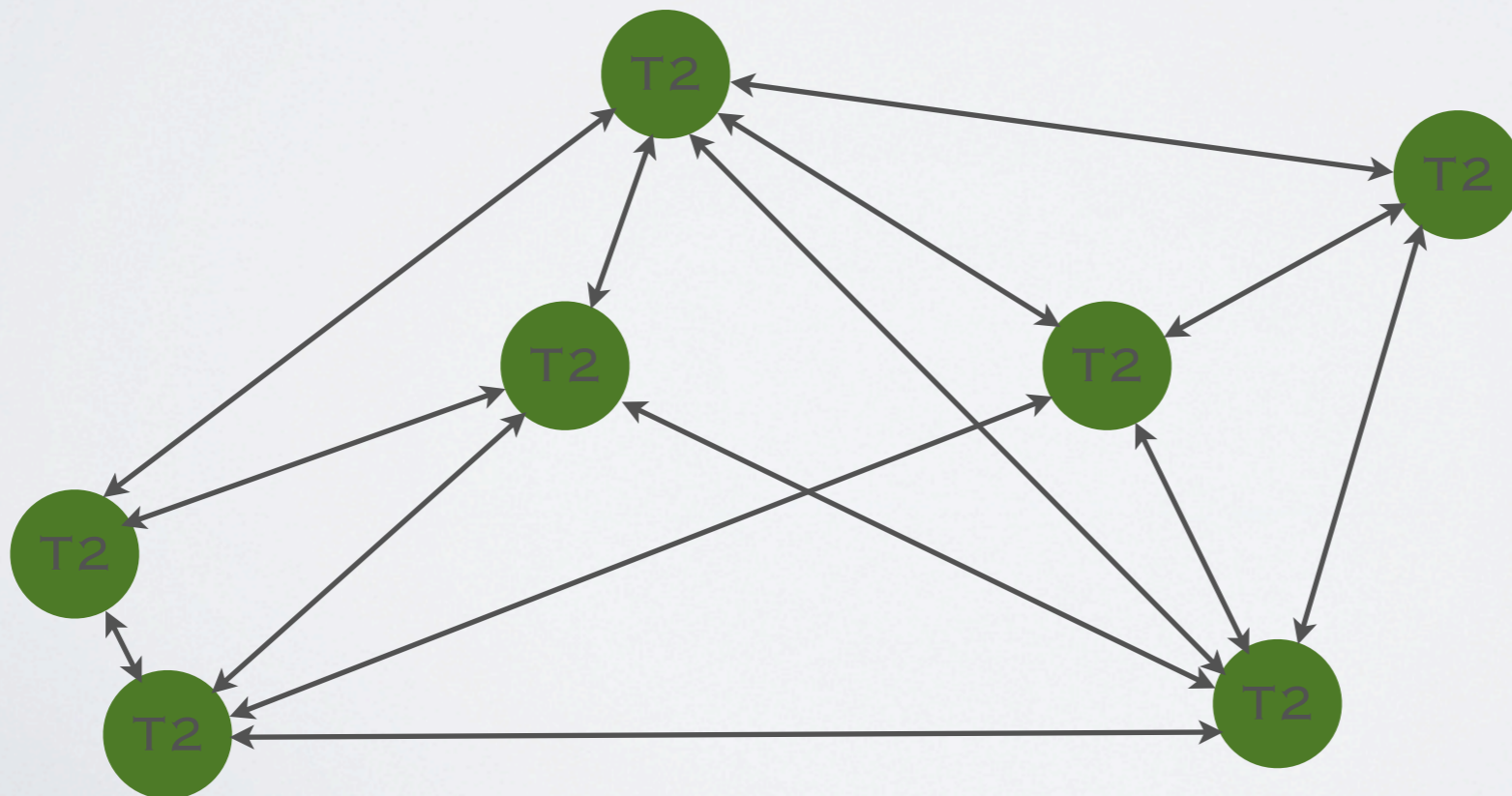
- WHILE THESE TEST ACCESS TO **/store/temp/user**, WHAT ABOUT THE DATA TRANSFERS TO THE “CORRECT” SE?

- EVENTUALLY WILL BE PROVIDED BY: **crab_getoutput**

T2_IT_Pisa	SRMv2	cmsdcache.pi.infn.it	ok	ok	ok
T2_IT_Rome	SRMv2	gridsrm.pi.infn.it	error	warn	error
T2_KR_KNU	SRMv2	cluster142.knu.ac.kr	ok	ok	ok
T2_PT_LIP_Lisbon	SRMv2	srm01.lip.pt	ok	ok	ok
T2_RU_IHEP	SRMv2	se0003.m45.ihep.su	ok	ok	ok
T2_RU_JINR	SRMv2	lcsedc01.jinr.ru	ok	ok	ok
T2_RU_PNPI	SRMv2	cluster.pnpi.nw.ru	ok	ok	ok
T2_RU_KI	SRMv2	se.grid.kiae.ru	ok	ok	error
T2_RU_SINP	SRMv2	grid128.sinp.msu.ru	ok	ok	error
T2_TR_METU	SRMv2	eymir.grid.metu.edu.tr	ok	ok	ok
T2_UK_London_Brunel	SRMv2	dc2-grid-64.brunel.ac.uk	ok	ok	ok
T2_US_Caltech	SRMv2	cit-se.ultralight.org	ok	ok	ok
T2_US_Florida	SRMv2	srm.ihepa.ufl.edu	ok	ok	ok
		srmb.ihepa.ufl.edu	ok	ok	ok
T2_US_MIT	SRMv2	se01.cmsaf.mit.edu	ok	ok	ok
T2_US_Nebraska	SRMv2	srm.unl.edu	ok	ok	ok
T2_US_Purdue	SRMv2	srm-dcache.rcac.purdue.edu	ok	ok	ok
T2_US_UCSD	SRMv2	bsrm-1.t2.ucsd.edu	ok	ok	ok
T2_US_Wisconsin	SRMv2	cmssrm.hep.wisc.edu	ok	ok	ok

T2-T2 TRANSFER TESTS WITH LCG TOOLS

- DEDICATED DATA TRANSFER STUDY BETWEEN ALL TIER-2 SITES
- MIMIC **CRAB** STAGE OUT, WHICH USES **lcg-cp** AT END OF JOB
- WRITE A 20MB FILE TO **/store/temp** AT EVERY TIER-2
- ATTEMPT TO COPY THIS FILE TO EVERY OTHER TIER-2 USING **lcg-cp**



T2-T2 TRANSFER TESTS WITH LCG TOOLS

- 35 TIER-2 SITES CONSIDERED (WHEREVER I HAD WRITE PERMISSION TO **/store/temp**)
- INCLUDES ALL OF OUR SITES IN THE US AND BRAZIL
- TIER-1, TIER-3 NOT CONSIDERED
- 1 190 POSSIBLE LINKS

T2-T2 TRANSFER TESTS WITH LCG TOOLS

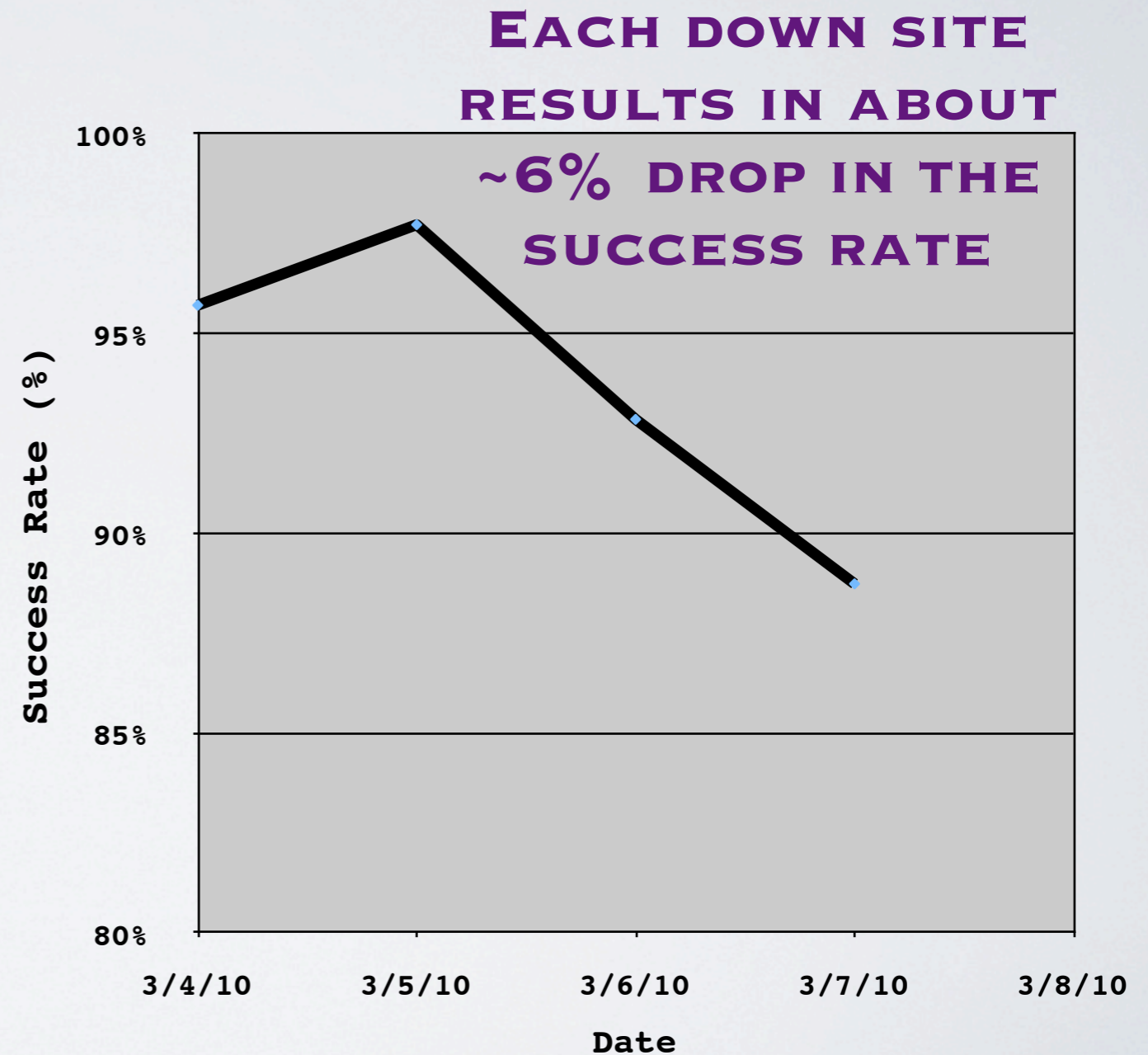
- SOMEWHAT DIFFERENT FROM **CRAB** STAGE OUT
- THIRD-PARTY TRANSFER DONE CENTRALLY, NOT FROM WN AT THE SOURCE SITE
- STILL A GOOD “PROOF OF PRINCIPLE” WHETHER LCG TOOLS CAN MOVE FILES AT ALL BETWEEN SITES, AND WITH WHAT SUCCESS RATE
- EVENTUALLY NEED AN “OFFLINE” JOB OUTPUT TRANSFER TOOL AFTER LOCAL STAGE OUT:
`crab -getoutput`

T2-T2 TRANSFER TESTS WITH LCG TOOLS

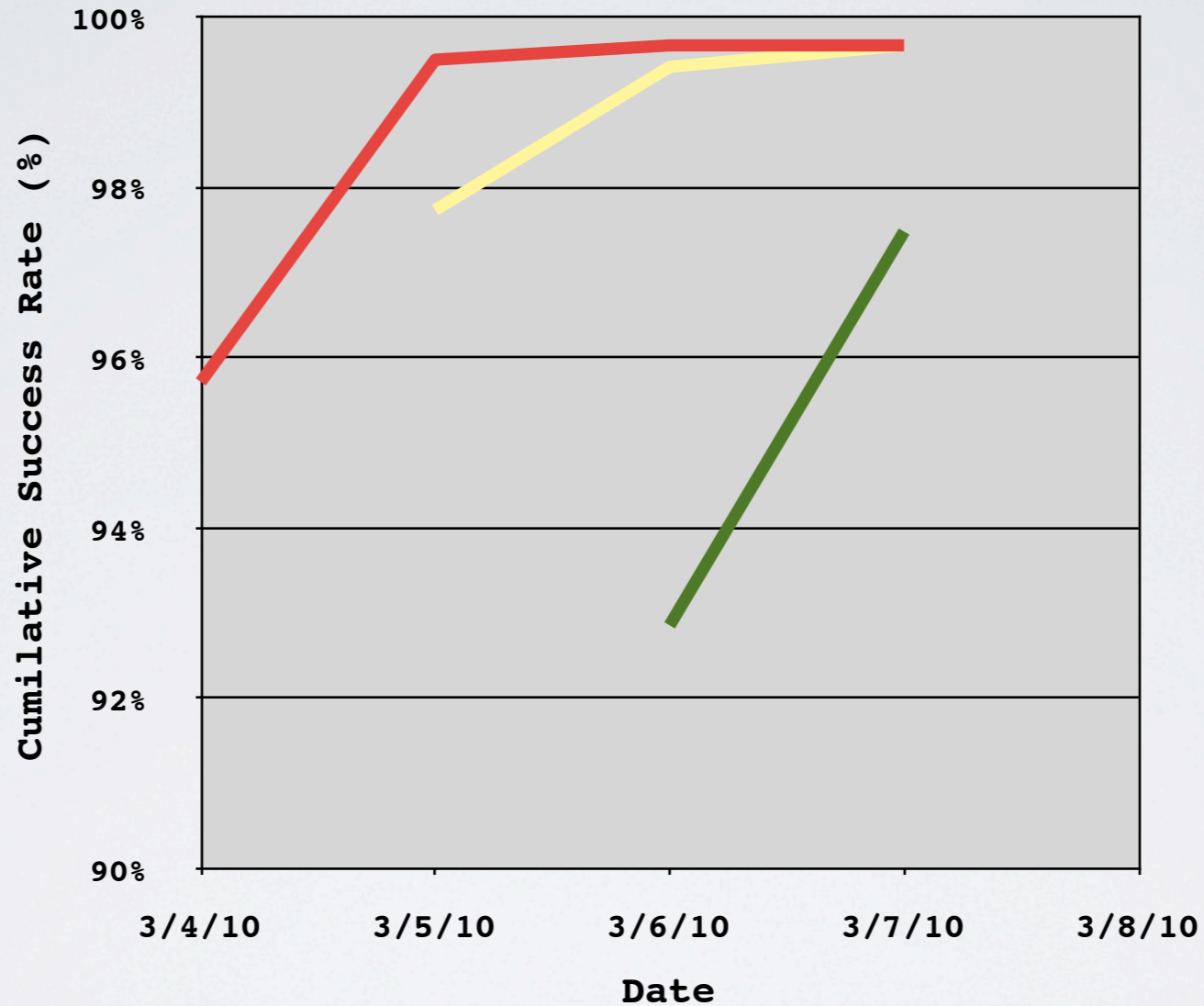
- ALL US AND BRAZILIAN TIER-2 SITES WERE TESTED
- GENERAL RESULTS (FOR ALL SITES):
 - **90±5% OF COPY JOBS SUCCEEDED ON THE FIRST ATTEMPT**
CONSISTENT WITH ABOUT A 12% FAILURE RATE OF JOBS DUE TO STAGE OUT, AS WE SAW EARLIER
 - **99% SUCCEEDED BY THE SECOND ATTEMPT ~1 DAY LATER**
SUGGESTS THAT “TRYING AGAIN LATER” IS A GOOD MODEL FOR **CRAB** STAGE OUT.
 - MOST ERRORS CAUSED BY A FEW TIER-2 SITES BEING UNAVAILABLE TEMPORARILY.
 - OTHER ERRORS ARE SEEMINGLY RANDOM, INDEPENDENT.

T2-T2 TRANSFER TESTS WITH LCG TOOLS

- ANY GIVEN ATTEMPT: 85-95% SUCCESS. SOME ELEMENT OF RANDOMNESS IN THE FAILURES.
- HOWEVER, LARGE DAY-TO-DAY SHIFTS CAUSED BY SITES BEING UNAVAILABLE FOR ALL OR PART OF THE TEST:
 - THU MAR 4: T2_UK_London_IC
 - FRI MAR 5: **NO PROBLEM SITES**
 - SAT MAR 6: T2_BR_UERJ AND T2_BR_SPRACE
 - SUN MAR 7: T2_IT_Rome, T2_BR_UERJ AND T2_BR_SPRACE

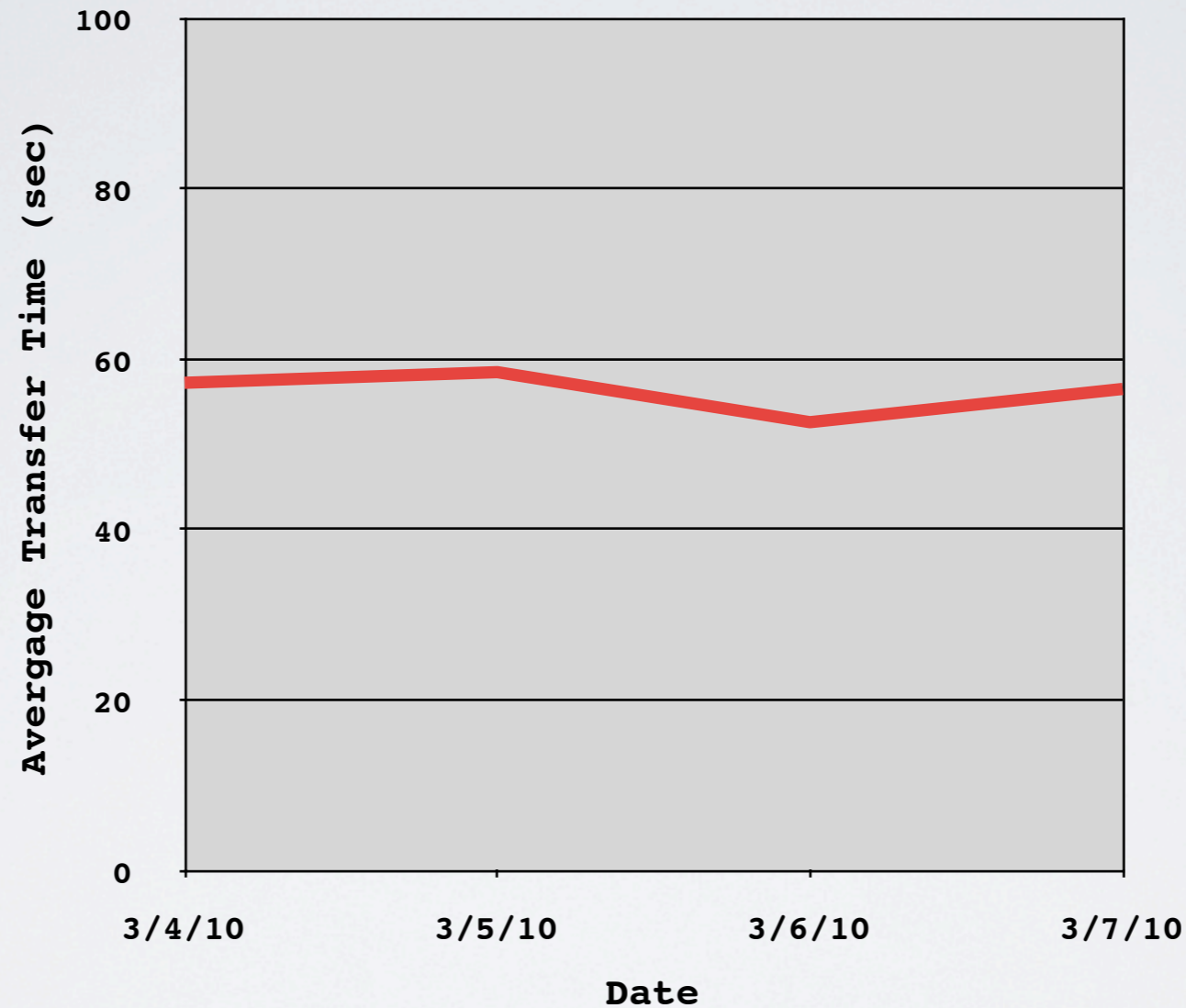


T2-T2 TRANSFER TESTS WITH LCG TOOLS



- SINCE MOST NON-RANDOM FAILURES ARE DUE TO TEMPORARY SITE UNAVAILABILITY, RE-TRYING A DAY LATER USUALLY RESULTS IN SUCCESS

T2-T2 TRANSFER TESTS WITH LCG TOOLS



- AVERAGE TRANSFER TIME (INCLUDING OVERHEAD): 55 SEC
- AVERAGE TRANSFER TIME WITHIN US/BR: 38 SEC

T2-T2 TRANSFER TESTS WITH LCG TOOLS

- GENERALLY FAILURE WAS DUE TO A SITE BEING UNAVAILABLE, RATHER THAN A PROBLEM ON A SPECIFIC LINK.
- ONLY 4 LINKS WORLD-WIDE NEVER WORKED AFTER SEVERAL ATTEMPTS (ALL TO OR FROM U.S.):
 - TWO FAILURES WERE DUE TO NO NETWORK PATH IN EITHER DIRECTION BETWEEN T2_US_Caltech AND T2_UA_KIPT, UNDERSTOOD ISSUE.
 - OTHER PROBLEMS: BOTH T2_US_Caltech AND T2_US_Purdue TO T2_FR_GRIF_IRFU

IMPLICATIONS FOR PHEDEX LINK COMMISSIONING

- IMPLICATION FOR PhEDEx LINK COMMISSIONING?
- CONSIDER THAT ONLY 50% OF T2-T2 LINKS WORK “OUT OF THE BOX” ON INITIAL LoadTest ATTEMPTS, WHILE 99% OF PATHS WORK WITH A SIMPLE SMALL FILE TRANSFER USING **lcg-cp**.
- ARE WE MAKING THE JOB OF COMMISSIONING PhEDEx LINKS MORE DIFFICULT FOR OURSELVES BY NOT SETTING THINGS UP AT EACH SITE BEFORE TESTING?

IMPLICATIONS FOR PHEDEX LINK COMMISSIONING

- **SHOULD WE BE DEBUGGING SITES RATHER THAN LINKS?**
 - PROPER READ PERMISSIONS - SOURCE SITE
 - PROPER DOWNLOAD AGENT CONFIGURATION - DESTINATION SITE
 - SPEED OF THE LINK - THIS IS THE MOST DIFFICULT ASPECT TO DEBUG - TIMEOUTS, THROUGHPUT, ETC. - FOCUS OF DDT?

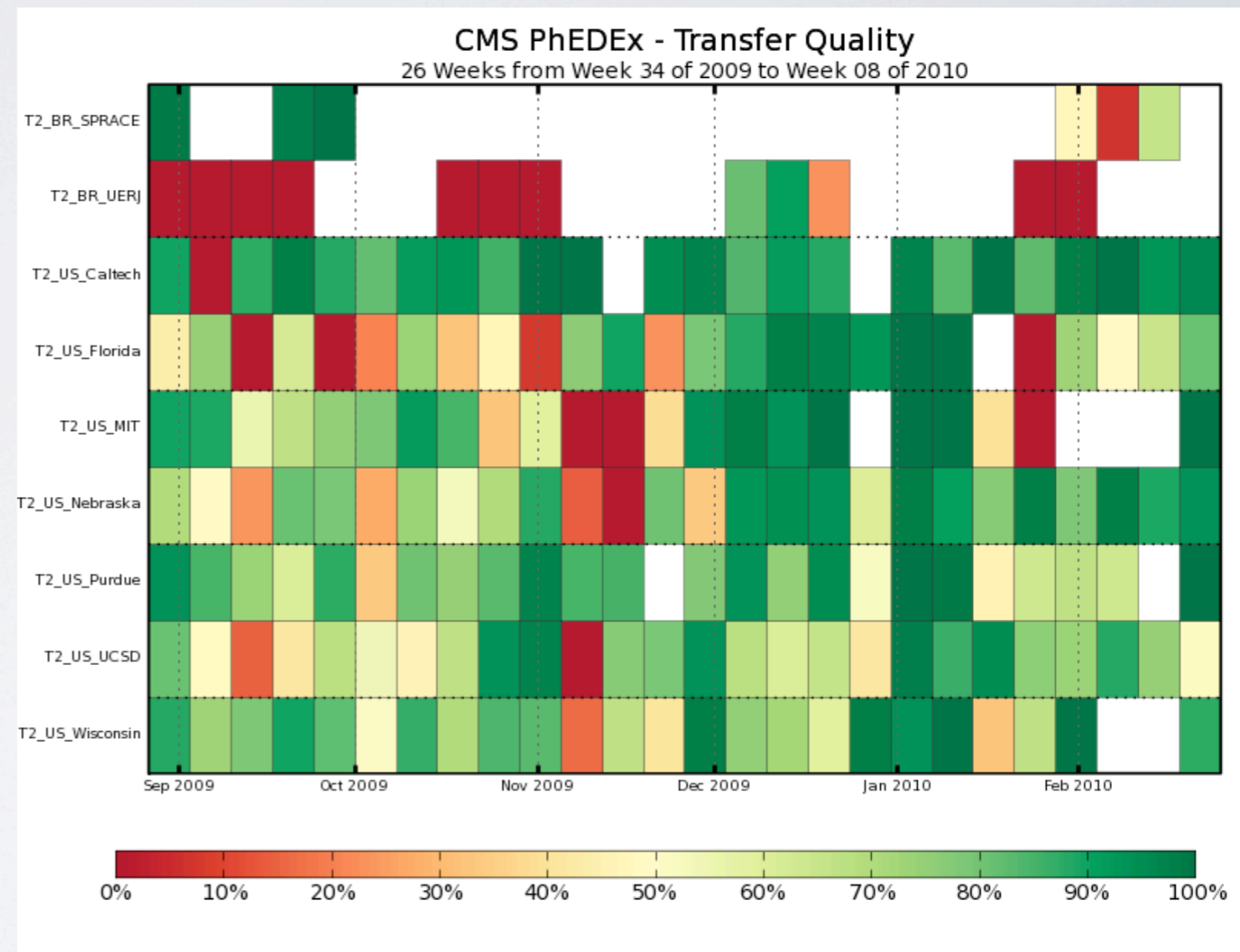
PHEDEX DATA TRANSFER LINKS

- ALL T2_BR_* & T2_US_* CROSS-LINKS WERE COMMISSIONED DURING THE SUMMER OF 2009 IN A COORDINATED EFFORT
- 61 LINKS OUT TO OTHER 29 TIER-2 (23% OF TOTAL)
- 70 LINKS FROM OTHER TIER-2 (27% OF TOTAL)

From Node ↓ To Node →	T2_BR_SPRACE	T2_BR_UERJ	T2_US_Caltech	T2_US_Florida	T2_US_MIT	T2_US_Nebraska	T2_US_Purdue	T2_US_UCSD	T2_US_Wisconsin
T2_BR_SPRACE	█								
T2_BR_UERJ		█							
T2_US_Caltech			█						
T2_US_Florida				█					
T2_US_MIT					█				
T2_US_Nebraska						█			
T2_US_Purdue							█		
T2_US_UCSD								█	
T2_US_Wisconsin									█

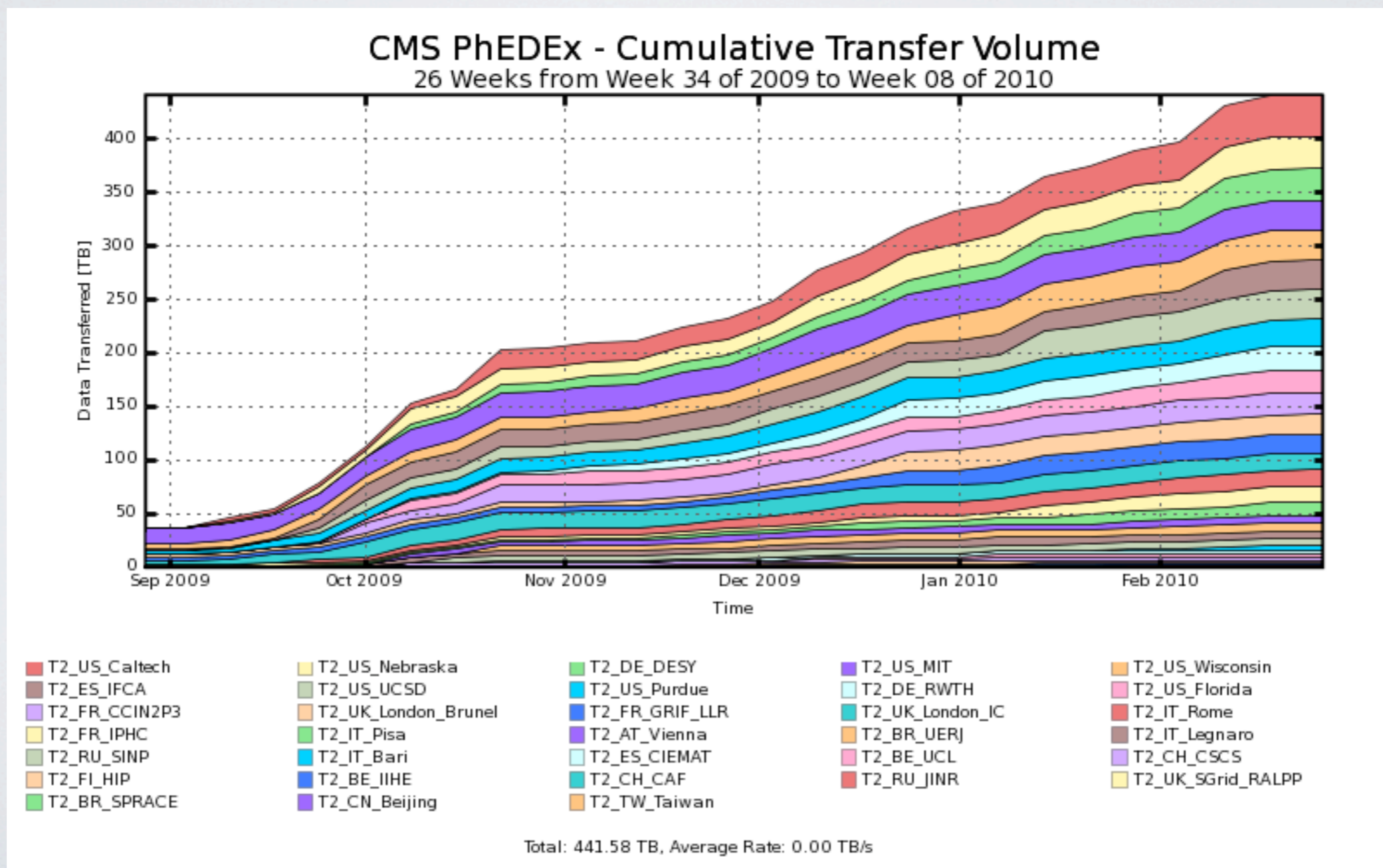
SERVING DATA TO OTHER TIER-2 SITES

- WITH COMMISSIONED TIER-2 LINKS, DATA CAN BE SERVED TO OTHER TIER-2 SITES
- BOTH OFFICIAL DATA AND /STORE/RESULTS FOR PHYSICS GROUPS
- FIRST COMMISSIONING FOCUS WAS ON SITES ASSOCIATED IN PHYSICS GROUPS)
- **GENERALLY GOOD TRANSFER QUALITY!**



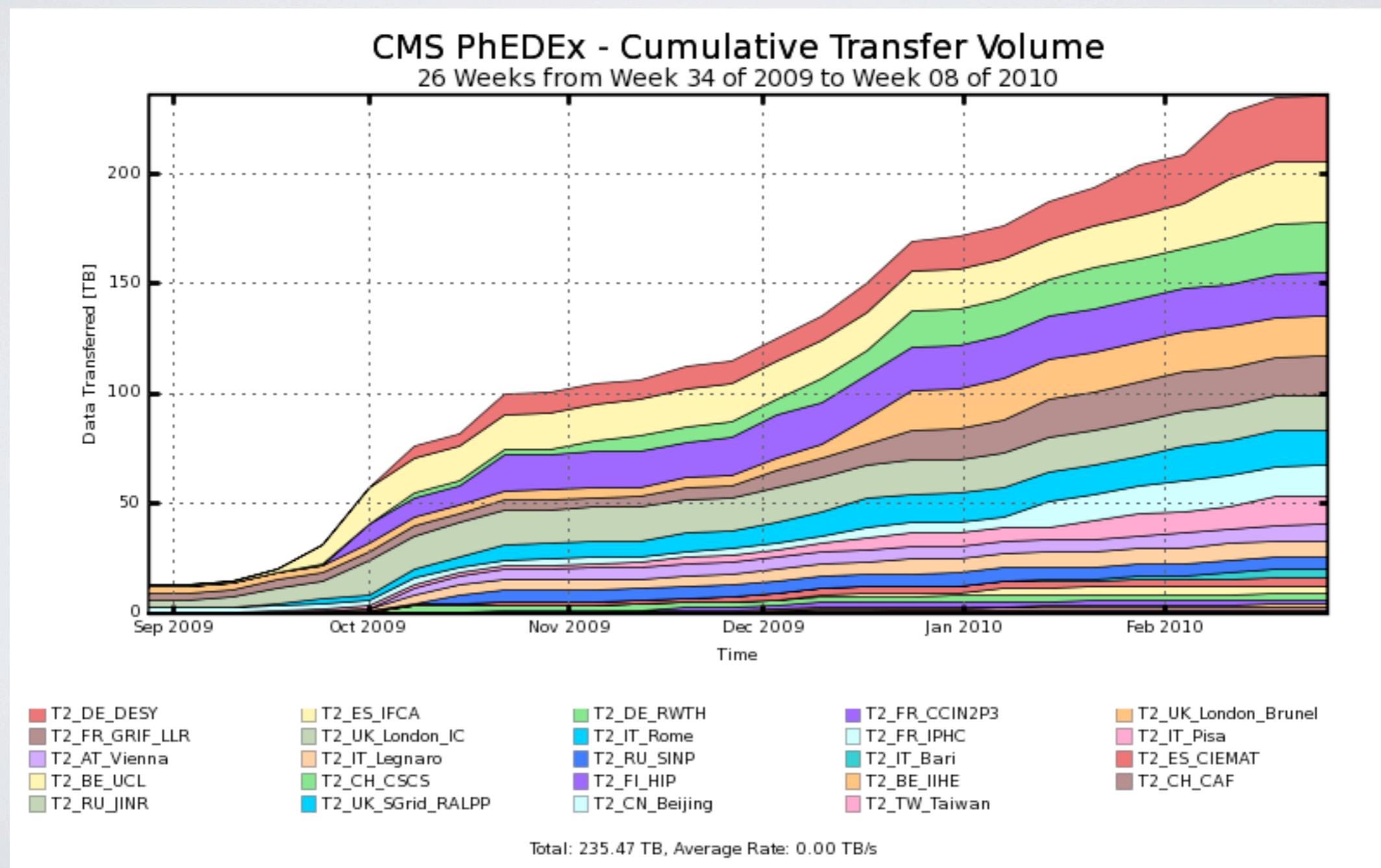
SERVING DATA TO THE WORLD

- T2_US AND T2_BR HAVE SERVED ABOUT 330 TB TO OTHER TIER-2 SITES (INCLUDING EACH OTHER'S SITES) SINCE OCT 1, 2009.



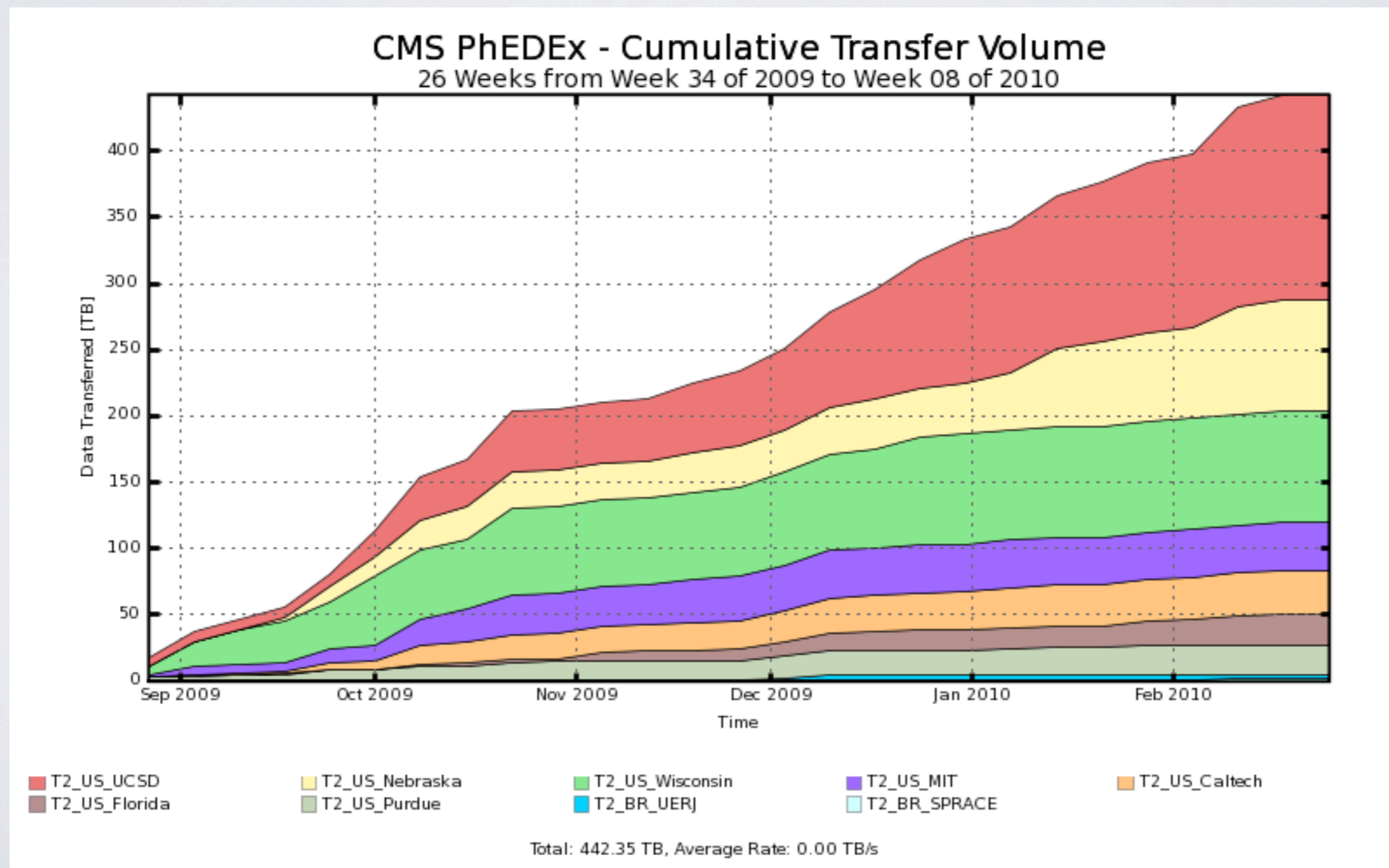
SERVING DATA TO THE WORLD

- ABOUT 50% OF THE DATA WE SERVE IS TO TIER-2 SITES OUTSIDE OF OUR GROUP



SERVING DATA TO TIER-2 SITES

- EVERY ONE OF OUR SITES IS PARTICIPATING IN THIS SERVICE
- OVER 50% OF ALL DATA SERVED WORLD-WIDE BY A TIER-2 TO ANOTHER TIER-2 IS DONE BY OUR SITES



SUMMARY

- US AND BRAZILIAN TIER-2 SITES DELIVERED ABOUT 30% OF ANALYSIS CPU, AND RAN ALMOST 40% OF CMS ANALYSIS JOBS, SINCE OCT 1ST 2009.
- OTHER CRAB JOB STATISTICS AT OUR SITES ARE TYPICAL OF TIER-2 SITES WORLD-WIDE (JOB SUCCESS, CPU/WC, ETC.), EXCEPT FOR SOME SITE-FAIL PERCENTAGES WHICH CAN BE IMPROVED.
- EXCELLENT INITIAL RESPONSE TIMES TO ANALYSIS OPS SAVANNAH TICKETS DURING OCTX
- OUR TIER-2 SITES SERVED OVER 330 TB OF DATA TO OTHER TIER-2 SITES AROUND THE WORLD, REPRESENTING OVER 50% OF THE TOTAL T2-T2 TRAFFIC.