Artifacts, Anecdotes and Ancient Architecture

Accounting for Decades of Technical Testing Debt

WOPR21 Experience Report

Dan Downing

Mentora
A Forsythe Company

Context

Problem

- Global provider of *value-added network* services is experiencing severe outages in their sub-standard US datacenter

Solution

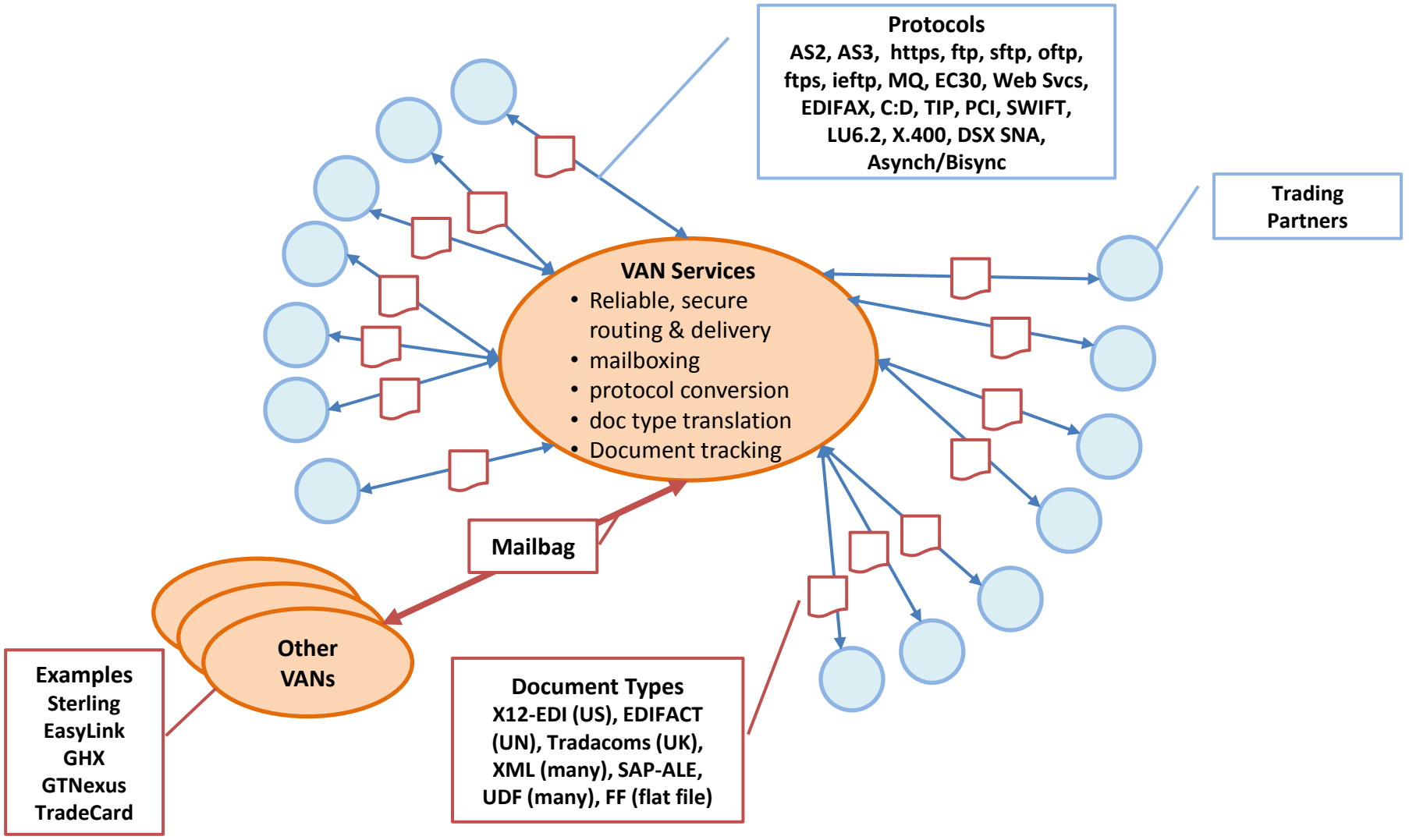
- Build out a pair of new “active-active” datacenters in Tier 1 hosting facilities, then migrate services and customers

Testing Project

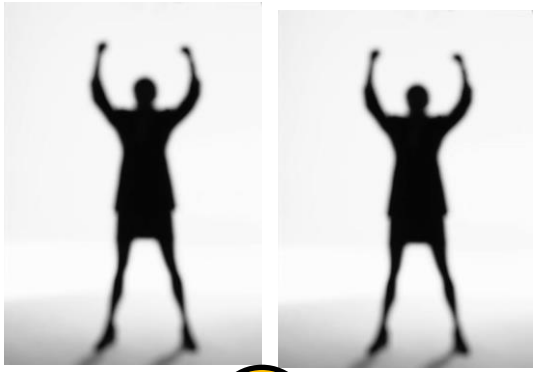
- Validate end-to-end functionality, performance, high-availability and reliability of the new data centers

Business: Value Added Network Services

Facilitate supply-chain trading partners' electronic communications



The Testing Team



QA Mngr (1)

Project Plan

1. Understand services and how customers use them
2. Understand services topology being migrated to the new DC
3. Understand processing and specs of end-to-end flows
4. Identify high volume flows
5. Understand manual testing, available tools

Discover

6. Design performance testing approach
7. Build tests
8. Compute Target Volume
9. Configure tests for Target Volume
10. Define test scenarios
11. Instrument environment to measure key resources

Design, Build

12. Execute tests
13. Analyze results, diagnose / fix, re-test

Execute, Analyze, Fix

Technical Testing Debt Elements

- **40-years of ad hoc evolution**
  - Products evolved from the mid-1970's; lacking cohesive architecture
  - Proliferation of "one-off" versions for large customers
  - Inconsistent "lingo": Evolved over time, acquisitions
- **New services topology**
  - Designs evolving in real-time – very few people really know
- **Testing Practices**
  - Most testing is product-release focused; no "macro, end-to-end" perspective
- **Fragmented Subject Matter Expertise**
  - Who knows what? Where are they?
  - To enroll SMEs, must establish credibility
- **Processing volumes**
  - Metrics limited, coarse: "KCs and ICs per node per day"
  - Not reliable: Contain much double-counting of "internal document hops"
  - No breakdowns by doc type, protocol , translated vs. pass-through

Hence, “Archaeological Dig”

- **Learn lingo, dialects, people**
  - By “immersion”, but over the phone, webex
  - Developed glossary of terms and synonyms
- **Uncover “bones” and piece them together**
  - Trial-and-error “strawmen” created / destroyed / improved
- **Synthesize learning, draw maps of the “new world”**
  - End-to-end flows, bridging new and old data centers
- **Devise approach to quantify traffic volume by category**
  - Reference data, assumptions, parameterized Excel formulas
  - Summaries by category, breakdowns by flow
- **Learn, adapt existing tools for performance testing**
  - Have LoadRunner, but does not support protocol range
  - Learn home-grown file replication and injection tools
- **Lower expectations, lengthen timeframes**
  - Our own, our customer’s

Learning Challenge: Geographical Fragmentation

London UK
MS Comms Dev

Amstelveen NE
– EMEA DC

Richmond CA –
Dev/QA

Michigan
AI Dev

Brook Park OH
Old DC

New York
Legacy Eng'g
Support

Gaithersburg MD HQ
– MS, TGO Dev/QA

Memphis TN
TGMS Dev

Lithia Springs
GA – New DC

Irvine CA –
Biz Dev/QA

Allen TX –
New DC

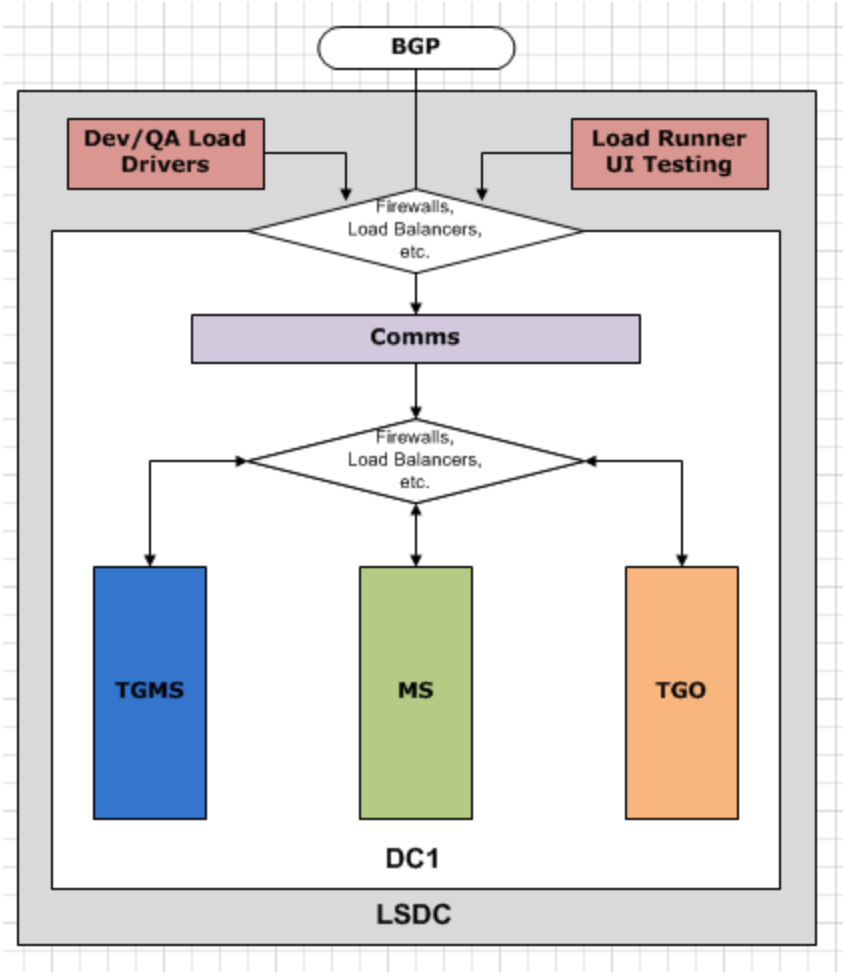
Tampa FL
IEFE Dev/QA

Dallas TX
TLe Dev/QA

Bangalore
SMG Dev/All QA

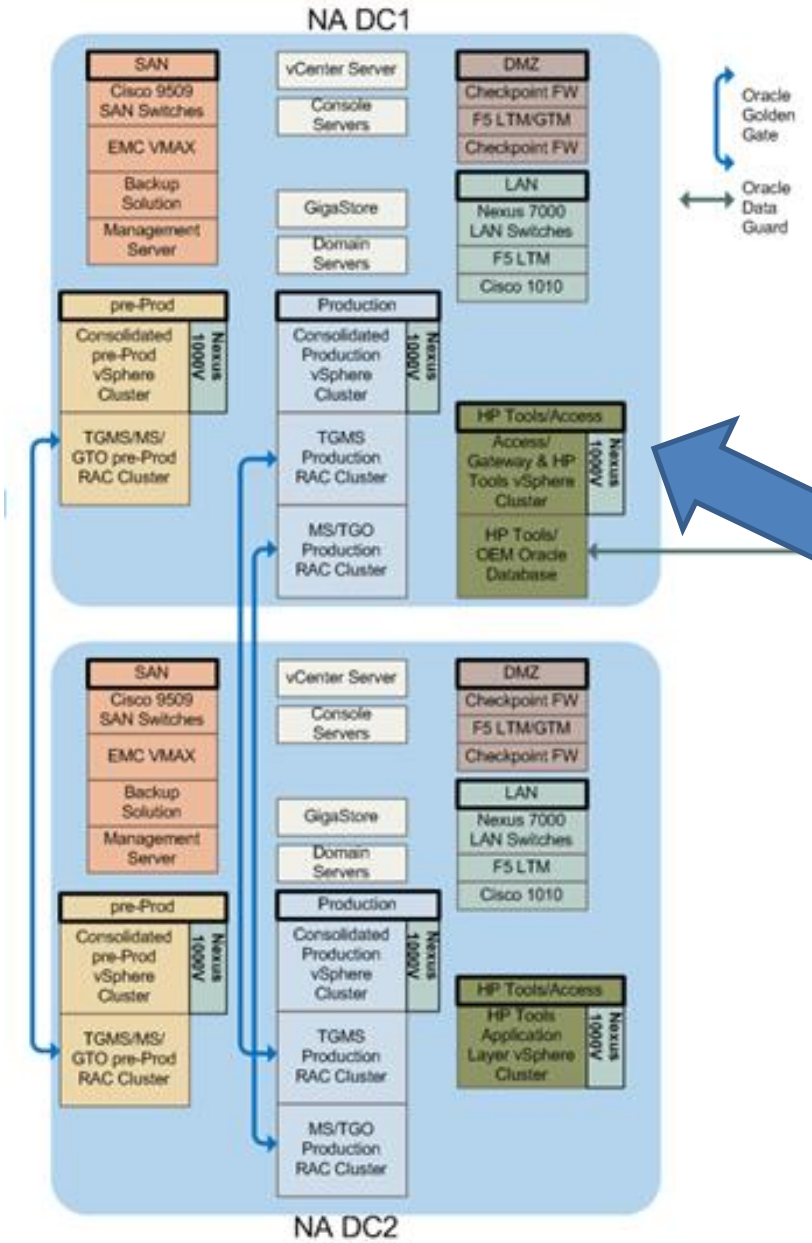
8.5 hrs apart; 3-hr "working window" 7-10am EST

Initial Topology: “World Is Flat”



- Single DC view
- App is “3 black boxes”
- End-to-End flows are self contained

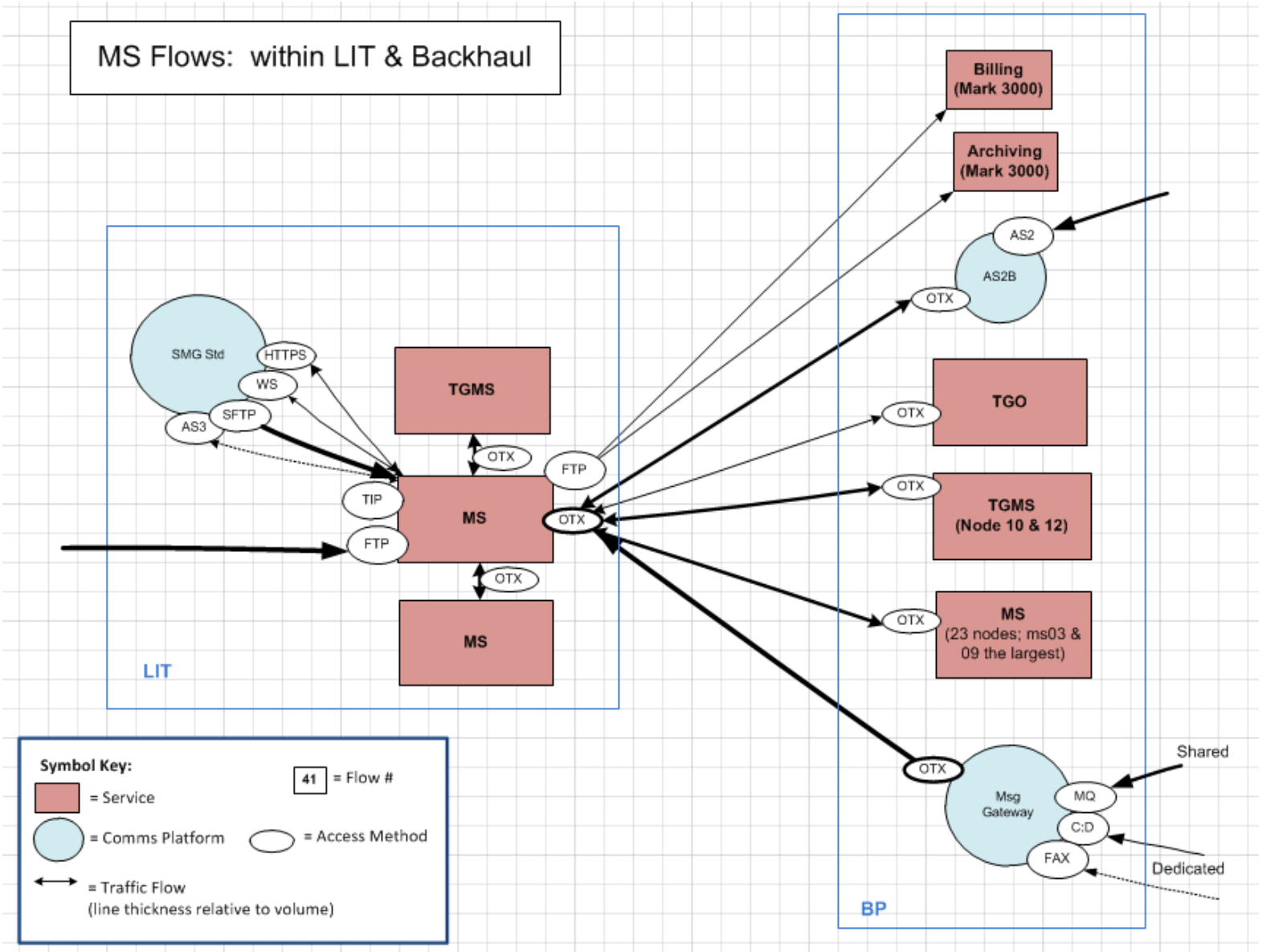
World is Really Round: Active-Active Dual Datacenters



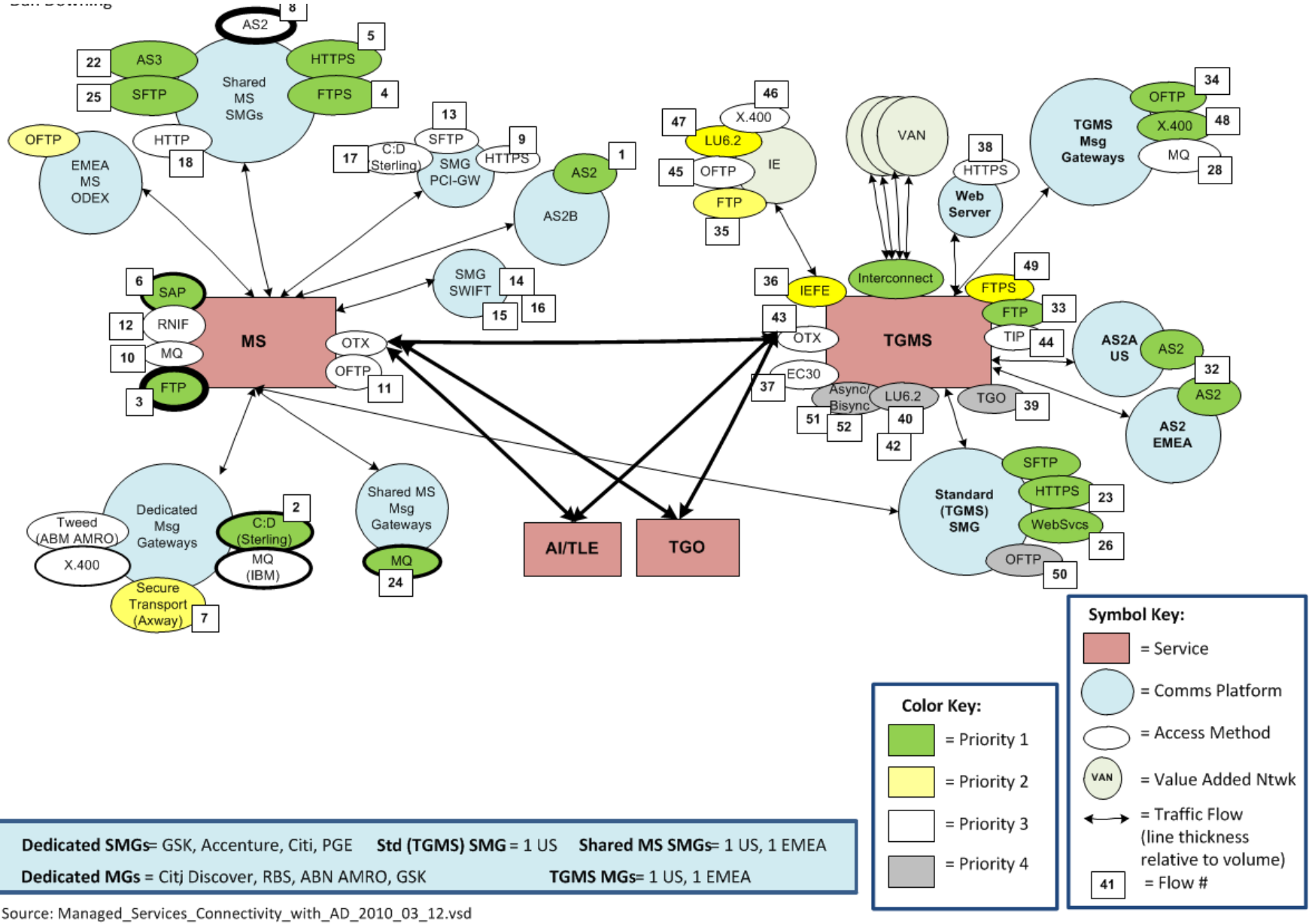
- **OOPS: Need to test HA across DCs!**
- **Black boxes turning whiter**
- **Some End-to-End flows "backhaul" to old DC**

Old DC remains for shared services not ready to migrate

Early End-to-End Flows Map



…then the Full Comms Topology Emerges

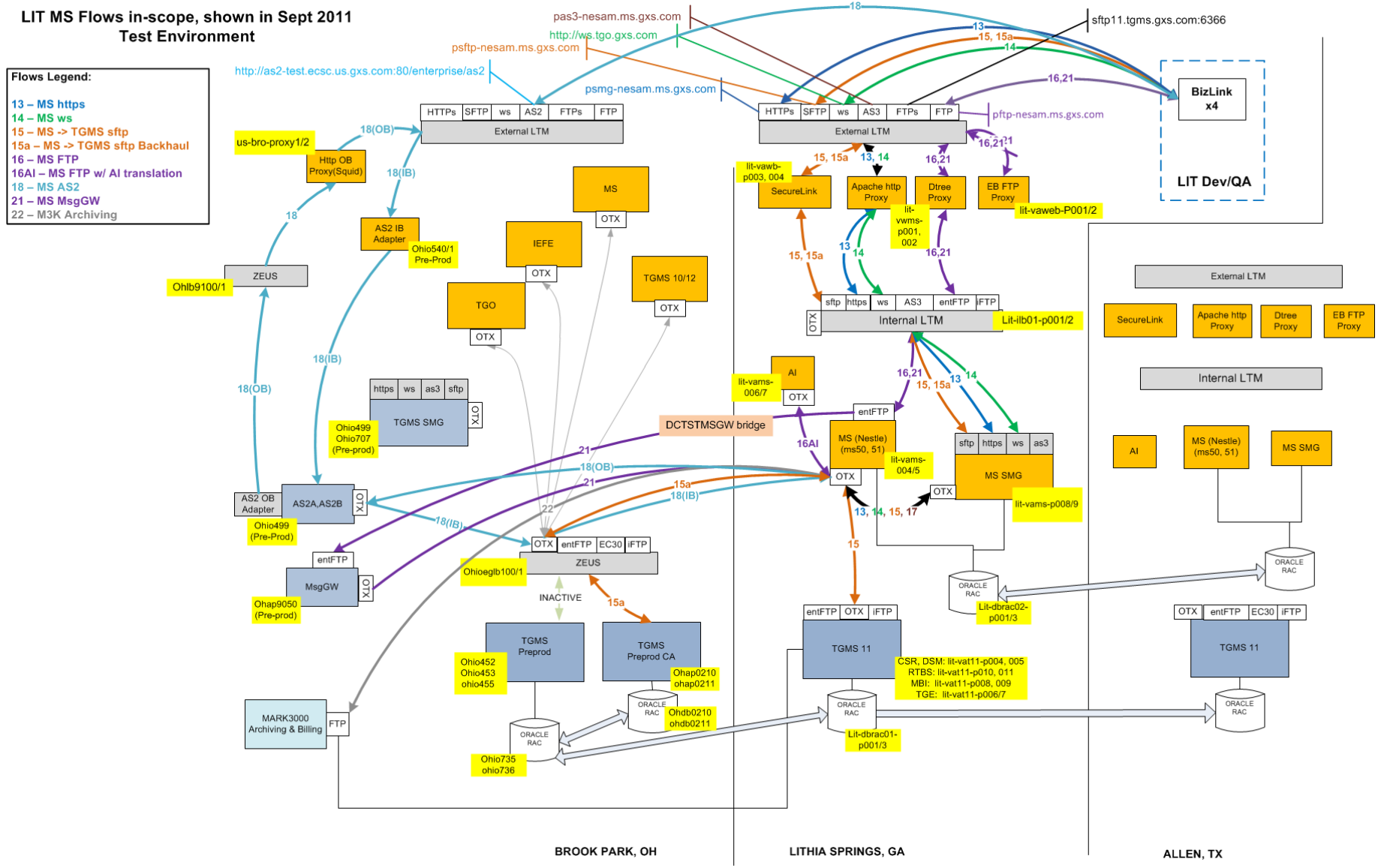


Dedicated SMGs= GSK, Accenture, Citi, PGE **Std (TGMS) SMG** = 1 US **Shared MS SMGs**= 1 US, 1 EMEA

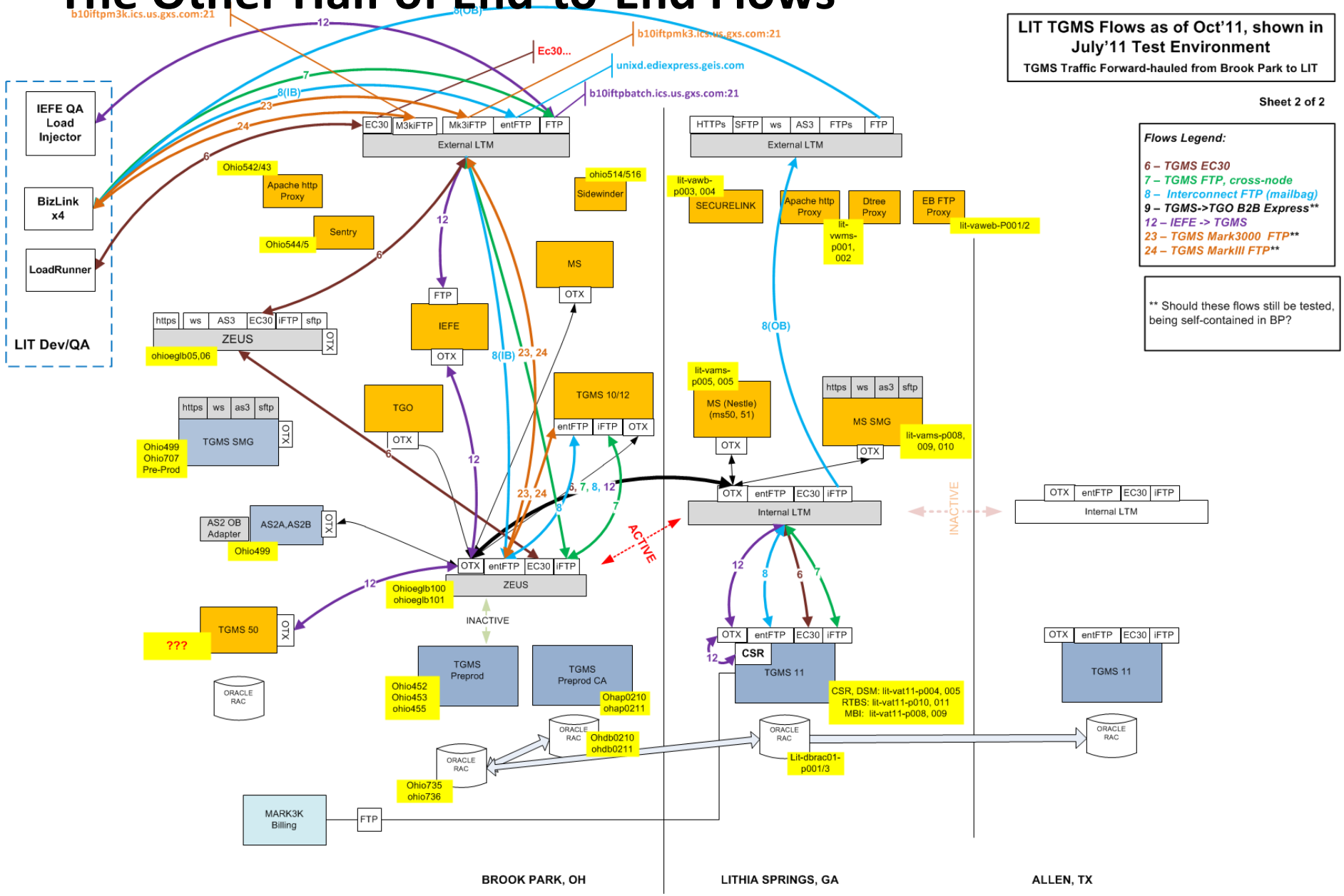
Dedicated MGs = Citj Discover, RBS, ABN AMRO, GSK **TGMS MGs**= 1 US, 1 EMEA

Source: Managed_Services_Connectivity_with_AD_2010_03_12.vsd

Finally, Real End-to-End Flows Map!



The Other Half of End-to-End Flows



LIT TGMS Flows as of Oct'11, shown in July'11 Test Environment

TGMS Traffic Forward-hauled from Brook Park to LIT

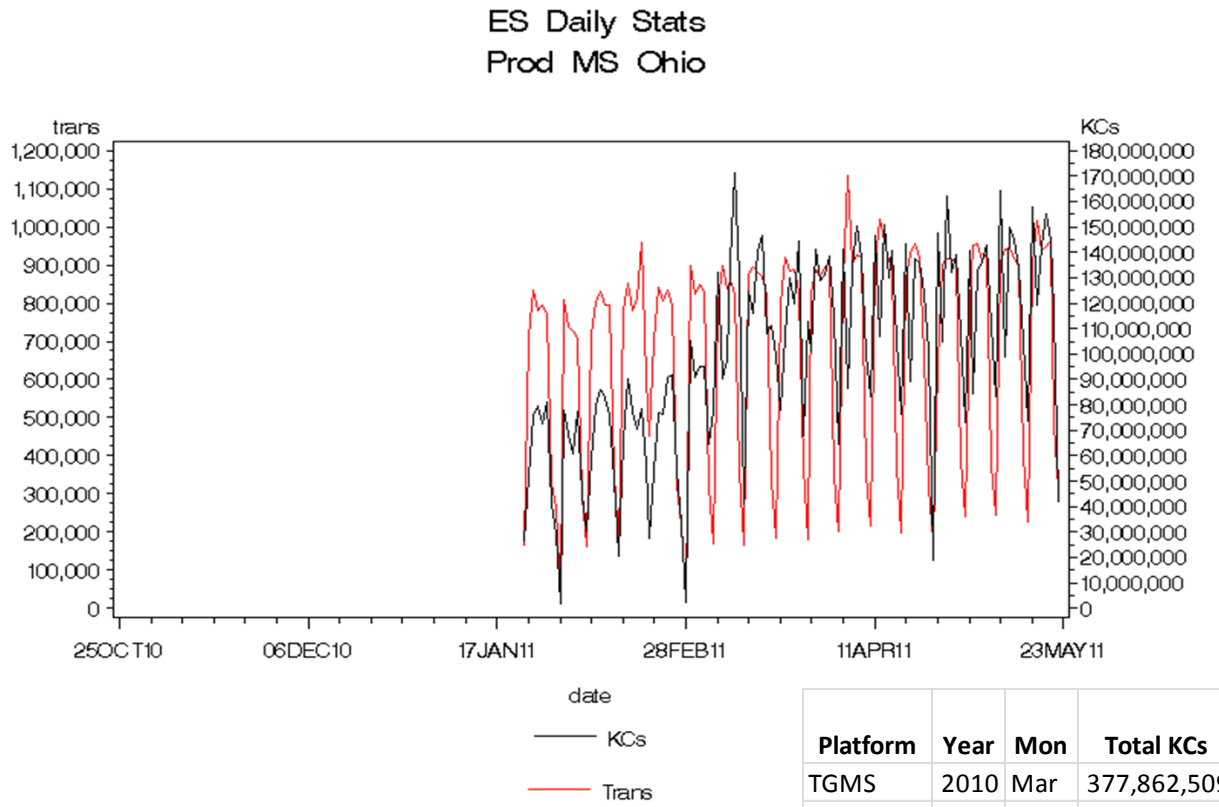
Sheet 2 of 2

Flows Legend:

6 – TGMS EC30
7 – TGMS FTP, cross-node
8 – Interconnect FTP (mailbag)
*9 – TGMS->TGO B2B Express***
12 – IEFE -> TGMS
*23 – TGMS Mark3000 FTP***
*24 – TGMS MarkIII FTP***

** Should these flows still be tested, being self-contained in BP?

Production Traffic Metrics: Our Starting Points



- **No breakdowns by protocol, doc types (never mind flow!)**
- **Coarse, double-counted metrics**
- **Ground-breaking target load estimation**

Platform	Year	Mon	Total KCs	Arrival KCs	Arrival Ics	Depart KCs	Depart Ics
TGMS	2010	Mar	377,862,509	186,565,432	25,259,529	191,297,077	28,963,448
TGMS	2010	Apr	337,623,606	170,921,491	24,161,135	166,702,115	27,743,909
TGMS	2010	May	349,378,933	174,650,629	24,701,141	174,728,304	28,514,551
TGMS	2010	Jun	395,338,951	193,947,719	27,967,182	201,391,232	32,096,284
TGMS	2010	Jul	397,798,329	197,727,677	26,642,475	200,070,652	30,149,739
TGMS	2010	Aug	363,634,304	183,897,561	25,847,732	179,736,743	28,867,599
TGMS	2010	Sep	403,320,020	200,652,217	28,231,358	202,667,803	32,225,608
TGMS	2010	Oct	412,948,954	202,117,482	27,674,858	210,831,472	31,427,861
TGMS	2010	Nov	390,729,033	164,681,777	20,737,899	226,047,256	32,479,328
TGMS	2010	Dec	368,327,216	155,755,917	20,516,365	212,571,299	32,286,270
TGMS	2011	Jan	357,528,393	150,228,622	19,297,674	207,299,771	31,448,098
TGMS	2011	Feb	364,885,134	161,845,900	21,842,444	203,039,234	31,525,001

Assumptions, Number Crunching, Sanity Checks Gets us To Here

	Target GB/day	Source
MS Target	**50**	**Volumes - MS03* (May 2'11)**
entFTP	8.2	guesstimate
SMG	0.7	Volumes - SMG sheet
AS2B	1.4	Volumes - Msg GW sheet
AS2 EMEA	0.7	Volumes - Msg GW sheet
Shared Msg GW(Bayer, Mars, Whirlpool)	1.1	Volumes - Msg GW sheet
Dedicated Msg GW (Citi, Discover)	7.9	Volumes - Msg GW sheet
M3000 archiving	30	S. Eckert
Total - MS	**50.0**	
TGMS Target	**7.5**	**Volumes - TGMS (wk May 8,'11)**
US (Node 11, sent+rcvd)	1.9	Volumes - TGMS sheet
EMEA (Node 12, sent+rcvd)	1.5	Volumes - TGMS sheet
US (Node 10, sent+rcvd)	4.0	Volumes - TGMS sheet
By Protocol		
AS2A	4.2	Volumes - Msg GW sheet
TGMS SMG	0.5	Volumes - Msg GW sheet
Msg GW (MQ)	1.0	Volumes - Msg GW sheet
ftp	1.9	N10+11+12 minus AS2A, SMG, Msg GW
Total - TGMS	**13.1**	
TGO Target	**0.7**	**Volumes - TGO sheet**
TGO Online UI	0.7	Volumes - TGO sheet
Total - TGO	**0.7**	
Total	**64**	**GB/hour**

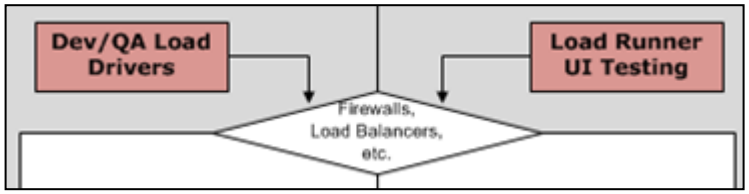
Finally, Target Volumes By Flow

Color Key:
Derived
Computed
Actual

** From March'11 Production volumes computations

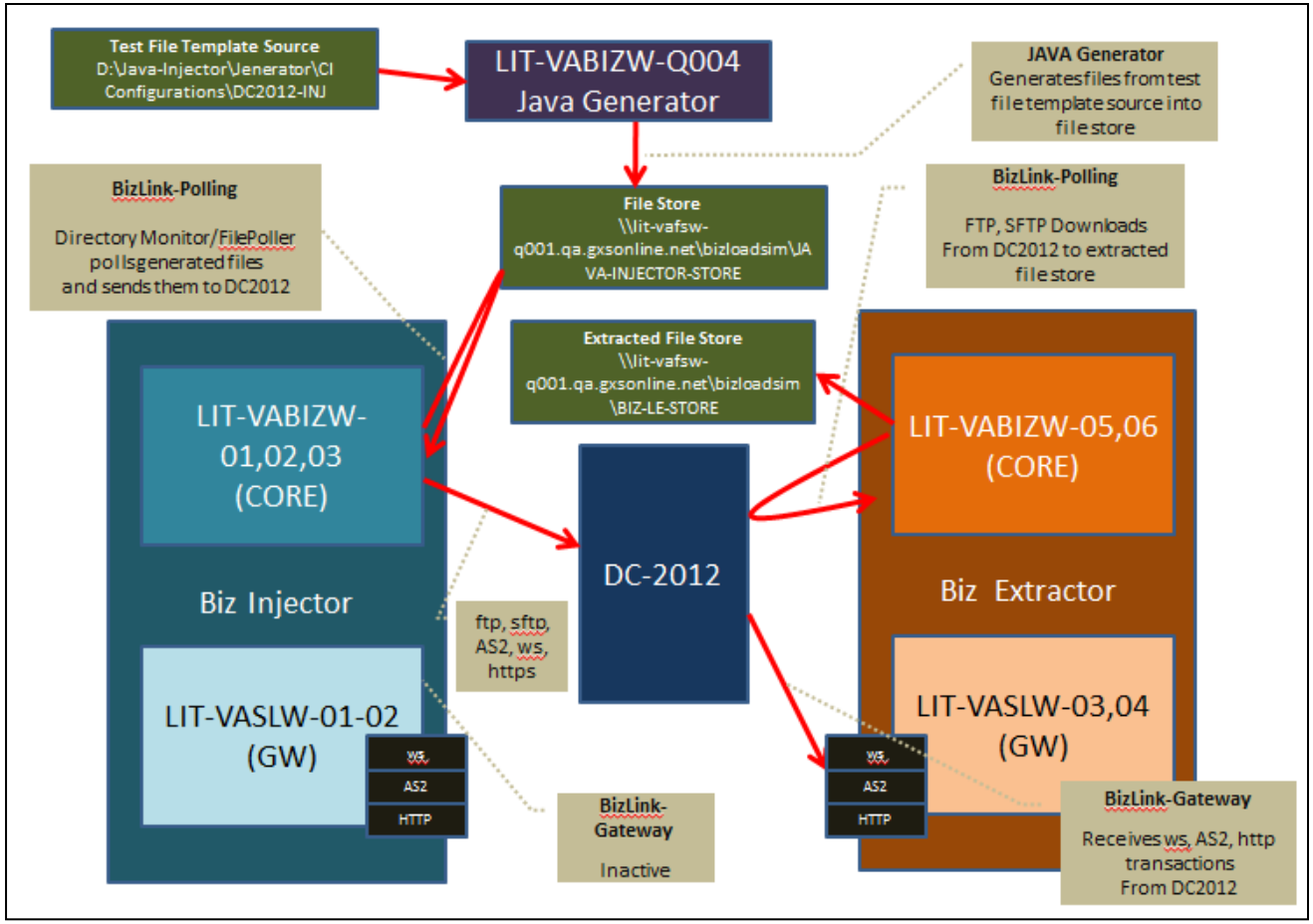
			100% Target Load						
Flow#	Flow - Existing	Transp.	Doc/ hr**	Doc/ min	KB/ doc	KB/hr	Urls (external LB)	Sender IDs	Rcvr IDs
13	DC2012-INJ|Flow_13_MS_HTTPS|Receiver	HTTP	244	4	22	5368	psmg-nesam.ms.gxs.com -> ms50 (Americas) psmg-nesao.ms.gxs.com -> ms51 (Other) 150.105.216.59 --> Allen via Ext LB and Apache Proxy **pweb-nesam.ms.gxs.com <-- 9/8: (Hassan) use this one!**	SMG_HTTP_S1 / PASSWORD	SMG_HTTP_R / PASSWORD
14	DC2012-INJ|Flow_14_MS_WS|Receiver	HTTP	87	1	20	1740	upload: https://pws.tgo.gxs.com/smg/ws_comm/upload?wsdl download: https://pws.tgo.gxs.com/smg/ws_comm/download?wsdl	SMG_WS_S1 / PASSWORD	SMG_WS_R1 / PASSWORD
15	DC2012-INJ|Flow_15_MS_TGMS_sftp_cross_srvc	SFTP	768	13	22	16896	psftp-nesam.ms.gxs.com	MS502TGMS / PASSWORD	MS Rcvr: TGMS2MS50 / PASSWORD TGMS Rcvr: AAC07184 / HA5SJZGT
15A	DC2012-INJ|Flow_15_MS_TGMS_sftp_cross_srvc (with backhaul)	SFTP	768	13	22	16896	psftp-nesam.ms.gxs.com	MS502TGMSBP / PASSWORD	TGMS Rcvr: tgmsftp4 / xxx AAB27353@pre-iftp.am.gxsics.com G24RFXNA
16	DC2012-INJ|Flow_16_MS_FTP_Enterprise|Receiver	FTP	4271	71	1	4271	Lithia: **pftp-nesam.ms.gxs.com** -> external ltm for send **pdctst.gtm.gxsonline.net** ->internal ltm for download Allen: 150.105.192.216	TEST1FTP / PASSWORD	TEST2FTP / PASSWORD
16AI	DC2012-INJ|Flow_16AI_MS_FTP_Enterprise|Receiver	FTP	4271	71	1	4271	pftp-nesam.ms.gxs.com	001234567 / PASSWORD	006928576 / PASSWORD
18	DC2012-INJ|Flow_18_MS_AS2B_backhaul|Receiver	AS2	1103	18	21	23163	http://as2-test.ecsc.us.gxs.com:80/enterprise/as2	77AS2USER	MS50USER
21	DC2012-INJ|Flow_21_MS_MsgGW_ftp_backhaul|Receiver	FTP	1232	21	1	1232	pdctst.gtm.gxsonline.net	MGTEST1 / PASSWORD	MGTEST2 / PASSWORD
22	MS, M3000 Archiving - FTP (backhaul)	FTP				1200000	Steve Eckert batch scripts	n/a	n/a
	Totals		**12761**	**213**		**1273837**	**KB/hr**		
	18	**smg**	**1099**	**per hr**		**1274**	**MB/hr**		
		ftp	**9774**	**per hr**		**15286**	**MB/day**		

Load Generation Evolves Too…

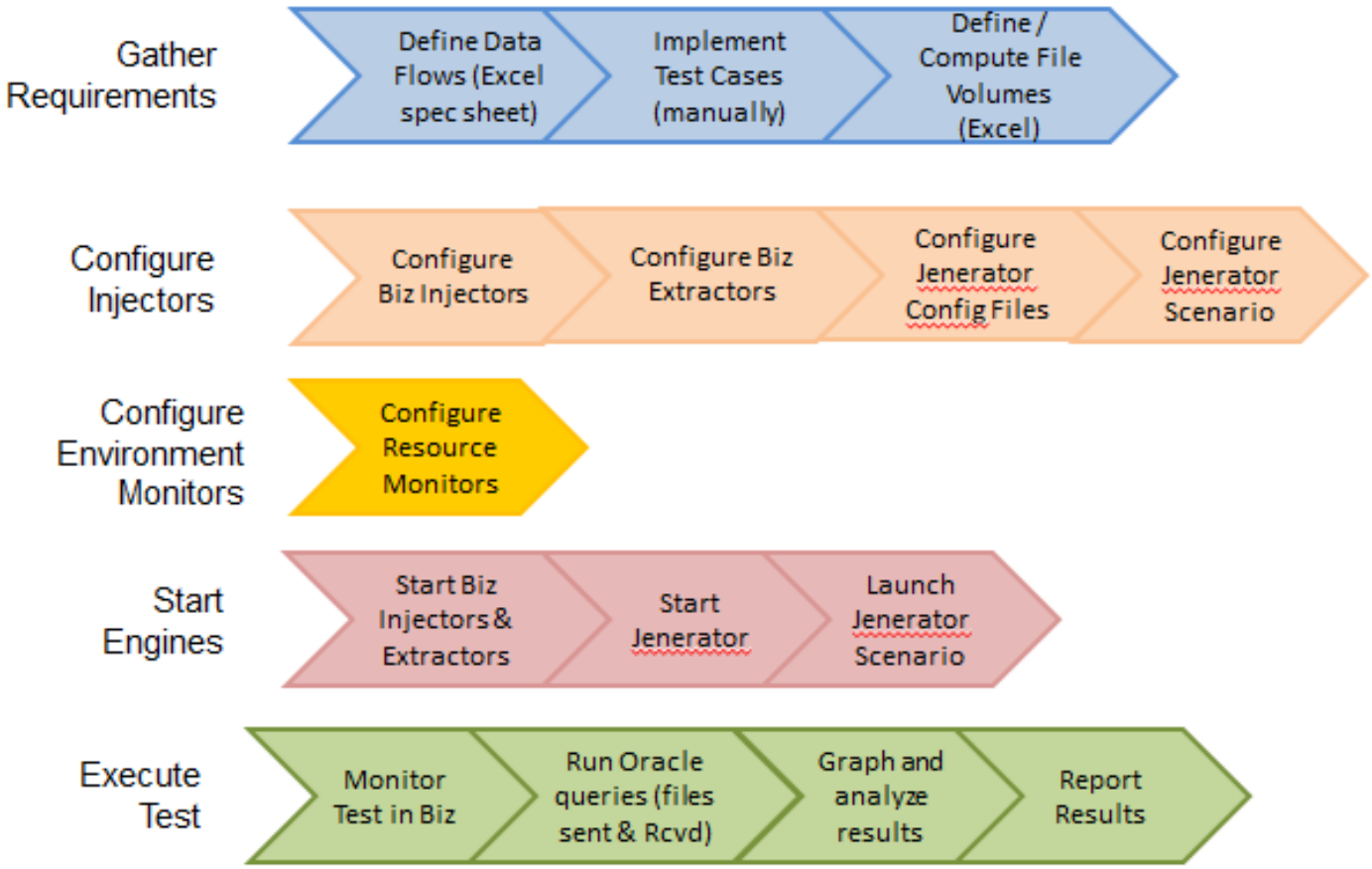


←…from here…

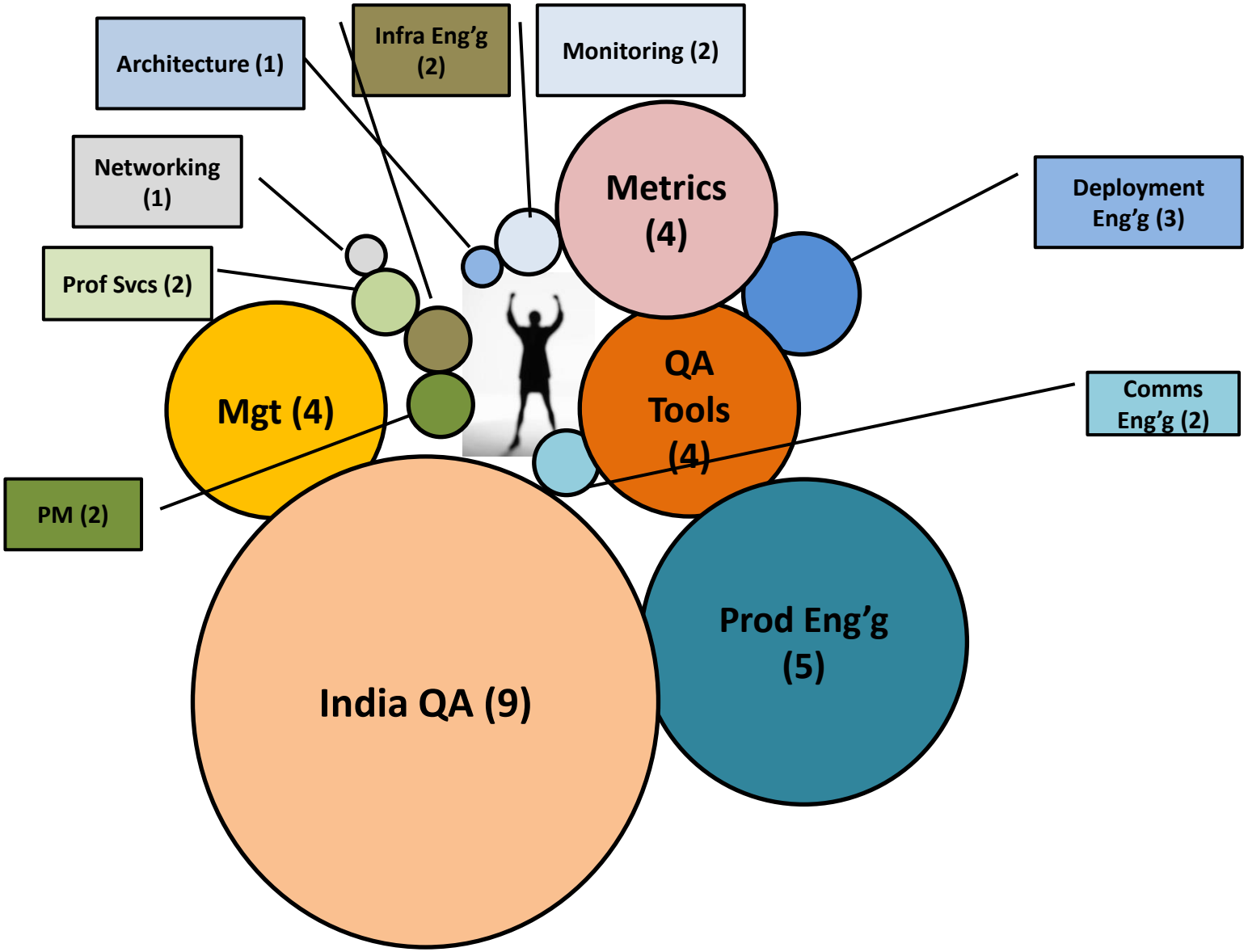
…to here →



…Guided by a new E2E Load Test Process



“Extended Team” Evolved to 40+ people by the end



Testing Technical Debt Retrospective

- **40-years of ad hoc evolution**
  - Understood “old world”, then simplified and validated (iteratively)
- **New Dual-DC plus “Backhaul” Network & Services topology**
  - Modeled, destroyed, remodeled – adjusting test scope accordingly
- **Testing Practices**
  - Learned how done now, invented whole new E2E process and tools
  - Decided what could and couldn’t be tested, measured
- **Fragmented Subject Matter Expertise**
  - Identified people willing to help and willing to teach us the lingo
- **Processing volumes**
  - Found a credible starting point, identified simplifying and estimating assumptions, developed parameterized Excel formulas

Keys To Overcoming

- **Found a couple of smart people willing to help; then built an extended team**
- **Face-to-face week with India QA**
- **Drew maps**
  - Got everyone's attentions, people happy to correct
- **Modeled by series of approximations**
  - Simplified models with supportable simplifying assumptions
  - Models fully traceable to reference points, validated by SMEs
- **Persevered, continuously bringing our stakeholders along**
- **Managed expectations (starting with our own)**