NOTED

Network Optimized Transfer of Experimental Data

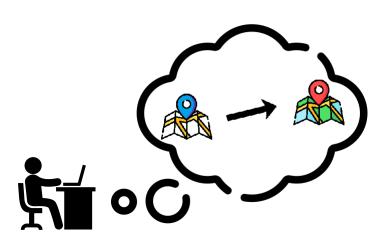
Joanna Waczyńska

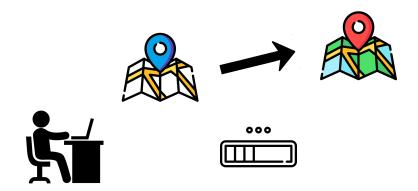
October 12, 2020



Outline

- Problem
- Solution
- Understanding the problem
- Tech side
- Operation of the project of the p
- 6 Analyze
- Conclusion v1
- Real-Life Examples
- Conclusion v2

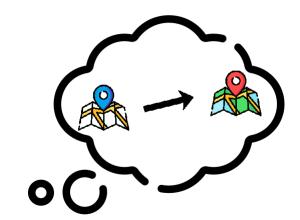




What happened?



Is it problem with network overloading?







History of transfer presented in FTS

Success rate

Timestamp	Decision	Running	Queue	(last 1min)	Throughput	EMA	Diff	Explanation
2020-09- 14T12:59:14Z	100	102	780	100.00%	970.26 MiB/s	960.22 MiB/s	Θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2020-09- 14T12:58:06Z	100	102	798	100.00%	980.63 MiB/s	959.10 MiB/s	θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2020-09- 14T12:56:59Z	100	109	836	100.00%	984.99 MiB/s	956.71 MiB/s	θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2020-09- 14T12:55:51Z	100	99	842	100.00%	988.25 MiB/s	953.56 MiB/s	θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2020-09- 14T12:54:44Z	100	103	485	100.00%	987.42 MiB/s	949.71 MiB/s	θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2020-09- 14T12:53:35Z	100	100	506	100.00%	990.96 MiB/s	945.52 MiB/s	Θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2020-09- 14T12:52:28Z	100	102	540	100.00%	995.23 MiB/s	940.47 MiB/s	0	Good link efficiency, current average throughput is larger than the preceding average. Hit uprange limit
2020-09- 14T12:51:21Z	100	101	578	100.00%	992.11 MiB/s	934.39 MiB/s	0	throughput is than the preceding average. Hit uses ange limit
								\downarrow \downarrow

NOTED Network Optimized Transfer of Experimental Data

FTS - File Transfer Service at CERN

Waczyńska

Icons made by Freepik from www.flaticon.com

5 / 65

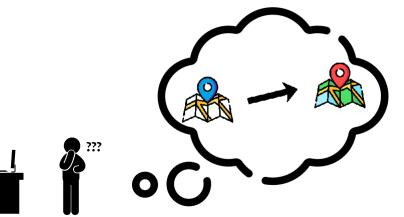
History of transfer presented in FTS

	Running	Queue	(last 1min)	Throughput	EMA	Diff	Explanation
100	102	780	100.00%	970.26 MiB/s	960.22 MiB/s	θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
100	102	798	100.00%	980.63 MiB/s	959.10 MiB/s	θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
100	109	836	100.00%	984.99 MiB/s	956.71 MiB/s	θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
100	99			988.25 MiB/s	953.56 MiB/s	θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
100	103	485 bulk	100.00%	987.42 MiB/s	949.71 MiB/s	θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
100	100	506	100.00%	990.96 MiB/s	945.52 MiB/s	Θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
100	102	540	100.00%	995.23 MiB/s	940.47 MiB/s	0	Good link efficiency, current average throughput is larger than the preceding average Hit upper range limit
100	101	578	100.00%	992.11 MiB/s	934.39 MiB/s	θ	does ffficiency, current average through the starger than the preceding average the upper range limit
	100 100 100 100	100 102 100 109 100 99 100 103 100 100	100 102 798 100 109 836 100 99 842 New black 100 103 485 K 100 100 506	100 102 798 100.00% 100 109 836 100.00% 100 99 842 \rightarrow{\text{N}}{\text{E}} 100.00% 100 103 485 \rightarrow{\text{E}}{\text{E}} 100.00% 100 100 506 100.00%	100 102 798 100.00% 980.63 M18/s 100 109 836 100.00% 984.99 M18/s 100 99 842 N 100.00% 988.25 M18/s 100 103 485 N 100.00% 987.42 M18/s 100 100 506 100.00% 990.96 M18/s 100 102 540 100.00% 995.23 M18/s	100 102 798 100.00% 980.63 MiB/s 959.10 MiB/s 100 109 836 100.00% 984.99 MiB/s 956.71 MiB/s 100 99 842 N 100.00% 988.25 MiB/s 953.56 MiB/s 100 103 485 N 100.00% 987.42 MiB/s 949.71 MiB/s 100 100 506 100.00% 990.96 MiB/s 945.52 MiB/s 100 102 540 100.00% 995.23 MiB/s 940.47 MiB/s	100 102 798 100.00% 980.63 M1B/s 959.10 M1B/s 0 100 109 836 100.00% 984.99 M1B/s 956.71 M1B/s 0 100 99 842 100.00% 988.25 M1B/s 953.56 M1B/s 0 100 103 485 100.00% 987.42 M1B/s 949.71 M1B/s 0 100 100 506 100.00% 990.96 M1B/s 945.52 M1B/s 0 100 102 540 100.00% 995.23 M1B/s 940.47 M1B/s 0

FTS - File Transfer Service at CERN

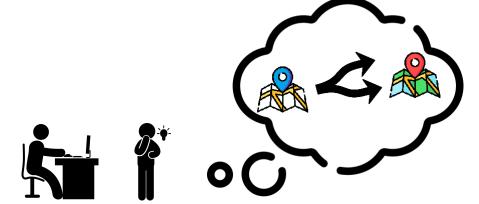
Waczyńska

Idea



Idea

Waczyńska



Timestamp	Decision	Running	Queue	Success rate (last 1min)	Throughput	EMA	Diff	Explanation
2020-09- 14T12:59:14Z	100	102	780	100.00%	970.26 MiB/s	960.22 MiB/s	Θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2020-09- 14T12:58:06Z	100	102	798	100.00%	980.63 MiB/s	959.10 MiB/s	Θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2020-09- 14T12:56:59Z	100	109	836	100.00%	984.99 MiB/s	956.71 MiB/s	Θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2020-09- 14T12:55:51Z	100	99	842	100.00%	988.25 MiB/s	953.56 MiB/s	Θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2020-09- 14T12:54:44Z	100	103	485	100.00%	987.42 MiB/s	949.71 MiB/s	Θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2020-09- 14T12:53:35Z	100	100	506	100.00%	990.96 MiB/s	945.52 MiB/s	Θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2020-09- 14T12:52:28Z	100	102	540	100.00%	995.23 MiB/s	940.47 MiB/s	θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2020-09- 14T12:51:21Z	100	101	578	100.00%	992.11 MiB/s	934.39 MiB/s	θ	throughput is than the preceding average. Hit uses ange limit

FTS - File Transfer Service at CERN

Icons made by Freepik from www.flaticon.com

NOTED Network Optimized Transfer of Experimental Data

Waczyńska

2020-09- 14T12:59:14Z	100	102	780	100.00%	90.26 MiB/s	960.22 MiB/s	Θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2020-09- 14T12:58:06Z	100	102	798	100.00%	980.63 MiB/s	959.10 MiB/s	θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2020-09- 14T12:56:59Z	100	109	836	100.00%	984.99 MiB/s	956.71 MiB/s	θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2020-09- 14T12:55:51Z	100	99	842	100.00%	988.25 MiB/s	953.56 MiB/s	θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2020-09- 14T12:54:44Z	100	103	485	100.00%	987.42 MiB/s	949.71 MiB/s	θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2020-09- 14T12:53:35Z	100	100	506	100.00%	990.96 MiB/s	945.52 MiB/s	θ	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2020-09- 14T12:52:28Z	100	102	540	100.00%	995.23 MiB/s	940.47 MiB/s	0	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2020-09- 14T12:51:21Z	100	101	578	100.00%	902.11 MiB/s	934.39 MiB/s	Θ	does V efficiency, current average through s larger than the preceding average it upper range limit
					,			

EMA

Diff Explanation

Throughput

Success rate

Queue

(last 1min)

FTS - File Transfer Service at CERN

Icons made by Freepik from www.flaticon.com

Timestamp

Decision

Running

				Success rate				
 Timestamp	Decision	Running	Queue	(last 1min)	Throughput	EMA	Diff	Explanation
2020-09- 14T12:59:14Z	2	2	1063	0.00%	0 bytes/s	0 bytes/s	θ	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:52:45Z	2	6	1053	0.00%	0 bytes/s	θ bytes/s	θ	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:46:16Z	2	4	1034	0.00%	0 bytes/s	0 bytes/s	θ	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:40:16Z	2	6	1044	0.00%	0 bytes/s	0 bytes/s	θ	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:33:54Z	2	12	1027	0.00%	0 bytes/s	θ bytes/s	θ	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:27:23Z	2	4	1034	0.00%	0 bytes/s	θ bytes/s	Θ	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:20:21Z	2	6	1014	0.00%	0 bytes/s	0 bytes/s	θ	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:13:06Z	2	2	1045	0.00%	0 bytes/s	θ bytes/s	θ	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:08:06Z	2	4	1060	0.00%	0 bytes/s	0 bytes/s	θ	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:02:49Z	2	6	1053	0.00%	0 bytes/s	0 bytes/s	٩	Bad lin effic y, no changes. Hit
2020-09- 14T11:56:21Z	2	3	963	0.00%	0 bytes/s	θ bytes/s	Ө	Rad link effic lover range li
								ᡮ

FTS - File Transfer Service at CERN

2020-09- 14T12:59:14Z 2020-09- 14T12:52:45Z 2020-09- 14T12:46:16Z 2020-09- 14T12:40:16Z	2 2 2	6 4	1063 1053 1034	0.00%	0 bytes/s 0 bytes/s	0 bytes/s	0	Bad link efficiency, no changes. Hit lower range limit Bad link efficiency, no changes. Hit lower range limit
14T12:52:45Z 2020-09- 14T12:46:16Z 2020-09-	2	4	1034		• •	0 bytes/s	Θ	
14T12:46:16Z 2020-09-				0.00%	0 bytes/s			
	2	6		1	,	0 bytes/s	θ	Bad link efficiency, no changes. Hit lower range limit
			1044	0.00%	0 bytes/s	θ bytes/s	0	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:33:54Z	2	12	1027	0.00%	0 bytes/s	θ bytes/s	θ	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:27:23Z	2	4	1034	0.00%	0 bytes/s	θ bytes/s	θ	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:20:21Z	2	6	1014	0.00%	0 bytes/s	θ bytes/s	θ	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:13:06Z	2	2	1045	0.00%	0 bytes/s	θ bytes/s	θ	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:08:06Z	2	4	1060	0.00%	0 bytes/s	θ bytes/s	θ	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:02:49Z	2	6	1053	0.00%	0 bytes/s	θ bytes/s	0	Bad link effici no changes. Hit lower range l
2020-09- 14T11:56:21Z	2	3	963	0.00%	0 bytes/s	θ bytes/s	θ	ink vac' y, no changes. Hit
							Í	1511

FTS - File Transfer Service at CERN

				Success rate				
Timestamp	Decision	Running	Queue	(last 1min)	Throughput	EMA	Diff	Explanation
2020-09- 14T12:59:14Z	2	2	1063	0.00%	0 bytes/s	θ bytes/s	Θ	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:52:45Z	2	6	1053	0.00%	0 bytes/s	0 bytes/s	Θ	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:46:16Z	2	4	1034	0.00%) bytes/s	0 bytes/s	Θ	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:40:16Z	2	6	1044	0.00%	0 bytes/s	θ bytes/s	0	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:33:54Z	2	12	1027	0.00%	0 bytes/s	θ bytes/s	Θ	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:27:23Z	2	4	1034	0.00%	0 bytes/s	θ bytes/s	Θ	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:20:21Z	2	6	1014	0.00%	0 bytes/s	θ bytes/s	Θ	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:13:06Z	2	2	1045	0.00%	6 bytes/s	θ bytes/s	Θ	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:08:06Z	2	4	1060	0.00%	0 bytes/s	θ bytes/s	0	Bad link efficiency, no changes. Hit lower range limit
2020-09- 14T12:02:49Z	2	6	1053	0.00%	0 bytes/s	0 bytes/s	0	ad 1. efficiency, no changes. Hit
2020-09- 14T11:56:21Z	2	3	963	0.00%	0 bytes/s	θ bytes/s	Θ	Bad ficiency, no changes. Hit
								4

FTS - File Transfer Service at CERN

Icons made by Freepik from www.flaticon.com NOTED Network Optimized Transfer of Experimental Data Waczyńska

Polling time!



Timestamp	Decision	Running	Queue	lmin)	Throughput	EMA	Diff	Explanation
2820-86- 92T07:43:24Z	17	18	1733	58.09%	5.44 MiB/s	11.83 MiB/s	-1	Bad link efficiency, no changes
2020-06- 02T07:39:24Z	18	19	1726	28.08%	5.80 MiB/s	12.54 MiB/s	-1	Bad link efficiency
2020-06- 02T07:35:51Z	19	21	1719	39.08%	8.99 M1B/s	13.29 MiB/s	-1	Bad link efficiency
2020-06- 02T07:32:20Z	20	21	1715	68.98%	12.27 MiB/s	13.76 MiB/s	-1	Bad link efficiency
2020-06- 02T07:28:52Z	21	22	1711	64.08%	14.57 MiB/s	13.93 MiB/s	-1	Bad link effic. cy
2020-06- 02T07:25:34Z	22	22	1705	80.00%	15.97 M1B/s	13.86 MiB/s	2	Good link efficiency erage throughput is larger than the prec
2020-06- 02T07:22:19Z	20	24	1784	44.08%	15.51 MiB/s	13.62 MiB/s	-1	Bad link efficiency

Success rate (last

Icons made by Freepik from www.flaticon.com

4

Problematic situations



Timestamp	Decision	Running	Queue	lmin)	Throughput	EMA	Diff	Explanation
2020-06- 02T07:43:24Z	17	18	1733	58.08%	5.44 MiB/s	11.83 MiB/s	-1	Bad link efficiency, no changes
2020-06- 02T07:39:24Z	18	19	1726	20.00%	5.80 M1B/s	12.54 MiB/s	-1	Bad link efficiency
2020-06- 02T07:35:51Z	19	21	1719	39.08%	8.99 M1B/s	13.29 MiB/s	-1	Bad link efficiency
2820-86- 92T07:32:20Z	20	21	1715	68.08%	12.27 MiB/s	13.76 MiB/s	-1	Bad link efficiency
2020-06- 02T07:28:52Z	21	22	1711	64.09%	14.57 MiB/s	13.93 MiB/s	-1	Bad link effic. y
2020-06- 02T07:25:34Z	22	22	1705	80.00%	15.97 M1B/s	13.86 MiB/s	2	Good link efficiency terage throughput is larger than the prec
2020-06- 02T07:22:19Z	20	24	1704	44.08%	15.51 M1B/s	13.62 MiB/s	-1	Bad link efficiency

Success rate (last

Active transfers

- Average of the last quarter success Rate is higher than the set limit (now 74%), or the last success Rate was higher than the set limit (now 74%).
- Last decision was higher than 2.
- The last report about transfer was observed in the last 15 minutes.

What do we need to take action?



- Source
- Destination
- Vo
- History queue:
 - Time
 - Information about files (Submitted and Active)
 - Information about success rate and throughput



- Sitename
- SeName
- IP prefixes (IPv4, IPv6)



FTS - File Transfer Service at CERN

CRIC - Computing Resource information Catalog

FTS - Tech side





μ ./jq

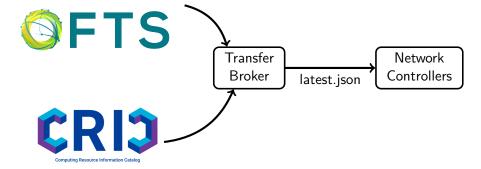
CRIC - Tech side



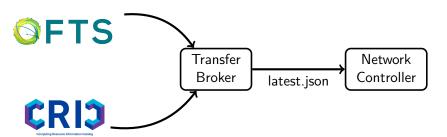


Project

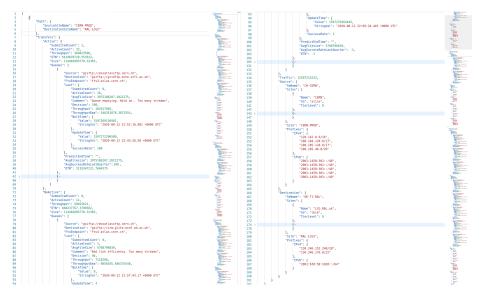
Web site that publishes the json file with all the active transfers, so network controllers can take optimization actions on their network.

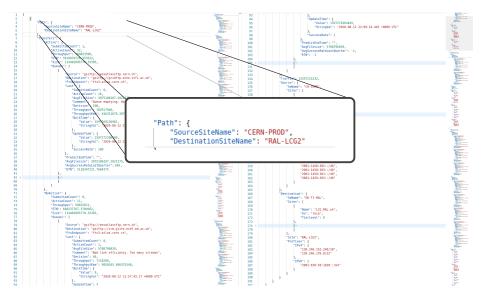


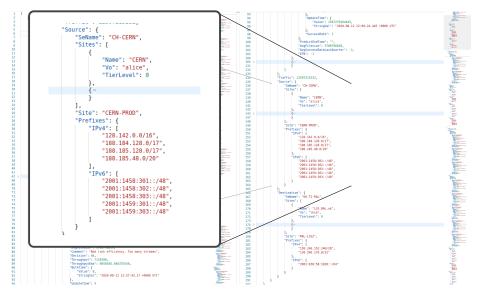
My previous work

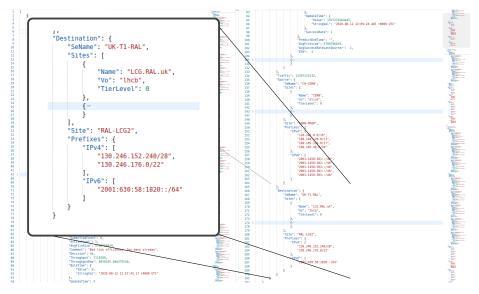


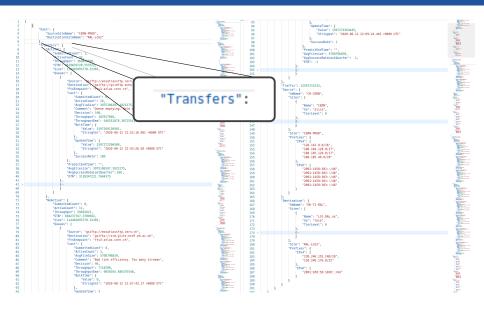
- Obtain transfer start & size information about transfers.
- Find large data transfers.
- Identification transfers.
- Analyze data and add extra estimators.

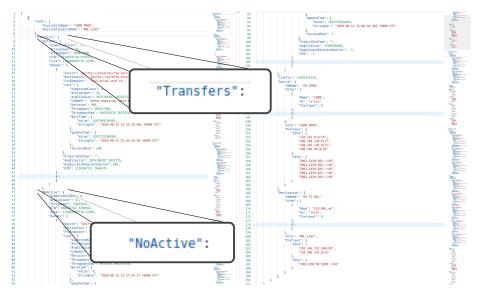


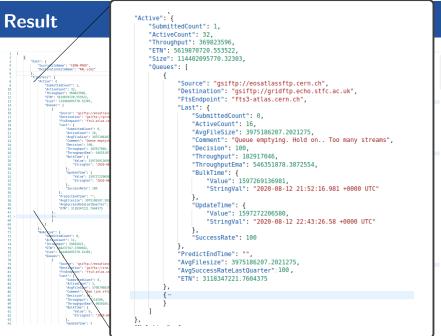








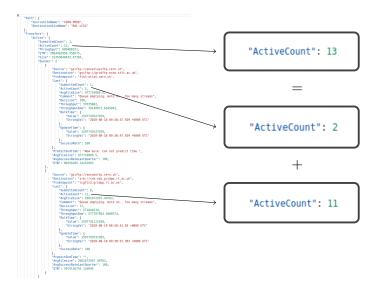




Many transfers were matched on the same Path

```
"Path": 4
    "SourceSiteName": "CERN-PROD".
     "DestinationSiteName": "RAL-LCG2"
Transfers": {
   "Active": {
        "SubmittedCount": 1.
        "ActiveCount": 13,
        "Throughput": 609488321
        "ETN": 3994492950.358575
        "Queues": [
                "Source": "gsiftp://eosatlassftp.cern.ch",
                "Destination": "gsiftp://gridftp.echo.stfc.ac.uk",
                 "FtsEndpoint": "fts3-atlas.cern.ch",
                "Last": (
                   "SubmittedCount": 1.
                    "ActiveCount": 2,
                    "AvgFileSize": 977719868.5,
                    "Comment": "Queue emptying. Hold on.. Too many streams",
                    "Decision": 189.
                    "Throughput": 37635802,
                    "ThroughputFma": 39140072.6185891.
                    "BulkTime": {
                       "Value": 1597743527929.
                        "StringVal": "2020-08-18 09:38:47,979 +0000 UTC"
                        "StringVal": "2020-08-18 09:38:47.929 +0000 UTC"
                    ).
"SuccessRate": 100
                 "PredictEndTime": "New bulk. Can not predict time.",
                "AvgFilesize": 977719060.5.
                "AvgSuccessRateLastQuarter": 100.
                "ETN": 80356285.24252993
                "Source": "gsiftp://eoscmsftp.cern.ch",
                "Destination": "srm://srm-cms.gridpp.rl.ac.uk",
                 "FtsEndpoint": "lcgfts3.gridpp.rl.ac.uk",
                    "SubmittedCount": 0.
                    "ActiveCount": 11.
                    "AvgFileSize": 2681672937.49763.
                    "Comment": "Queue emptying. Hold on.. Too many streams",
                    "Decision": 17.
                    "Throughput": 571844519.
                    "ThroughputEma": 577707964.5600574,
                    "BulkTime": {
                       "Value": 1597741121030,
                        "StringVal": "2020-08-18 08:58:41.03 +0000 UTC"
                   ),
"UpdateTime": {
    "Value": 1597743532303,
    ""2020-08-1
                        "StringVal": "2020-08-18 09:38:52.303 +0000 UTC"
                    }.
"SuccessRate": 100
                PredictEndTime": "",
                "AvgFilesize": 2681672937.49763,
                "AvgSuccessRateLastQuarter": 100,
               "ETN": 3914136745.116045
```

Many transfers were matched on the same Path



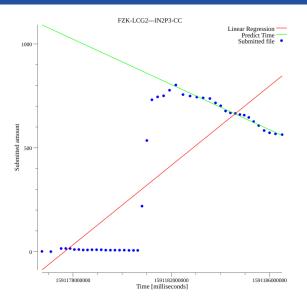
Queue decreases



Why would we want to predict the finish time?

Can we predict the finish time?

Smooth and large data transfers



Problem Solution Understanding the problem Tech side Project Analyze Conclusion v1 Real-Life Examples Conclusion v2

Clear situation

Details for srm://atlassrm-kit.gridka.de → srm://ccsrm.in2p3.fr Q



First Previous 1 Next Last

Timestamp	Decision	Running	Queue	ess rate (last 1min)	Throughput	EMA	Diff	Explanation
2020-06- 03T12:19:27Z	100	168	551	100.00%	448.34 MiB/s	371.37 MiB/s	1	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2820-86- 03T12:15:89Z	99	97	563	198.98%	442.40 MiB/s	362.82 MiB/s	2	Good link efficiency, current average throughput is larger than the preceding average
2020-06- 03T12:10:41Z	97	96	567	100.00%	436.65 MiB/s	353.98 MiB/s	2	Good link efficiency, current average throughput is larger than the preceding average
2020-06- 03T12:06:39Z	95	93	572	100.00%	436.42 M1B/s	344.79 M1B/s	2	Good link efficiency, current average throughput is larger than the preceding average
2020-06- 03T12:02:57Z	93	91	583	100.00%	428.05 MiB/s	334.61 MiB/s	2	Good link efficiency, c t average throughput is larger than the receds erage
2020-06- 03T11:59:17Z	91	92	697	100.08%	422.95 MiB/s	324.23 MiB/s	2	Good link effici
2020-06- 03T11:55:49Z	89	87	626	100.00%	414.37 MiB/s	313.26 MiB/s	2	Good link efficiency larger than the prec
2020-06-	87	86	646	100.00%	408.32 M1B/s	302.03 M1B/s	2	Good link efficie

Icons made by Freepik from www.flaticon.com

4

Problem Solution Understanding the problem Tech side Project Analyze Conclusion v1 Real-Life Examples Conclusion v2

Clear situation





First Previous 1 Next Last hroughput was increasing

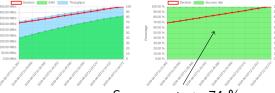
Timestamp	Decision	Running	Suce Queue	ess rate (last lmin)	Throughput	EMA	Diff	Explanation
2820-86- 03T12:19:27Z	100	198	551	100.00%	448.34 MiB/s	371.37 MiB/s	1	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2820-86- 83T12:15:89Z	99	97	563	100.00%	442.40 MiB/s	362.82 MiB/s	2	Good link efficiency, current average throughput is larger than the preceding average
2020-06- 03T12:10:41Z	97	96	567	100.00%	436.65 MiB/s	353.98 MiB/s	2	Good link efficiency, current average throughput is larger than the preceding average
2020-06- 03T12:06:39Z	95	93	572	100.00%	436.42 M18/s	344.79 M1B/s	2	Good link efficiency, current average throughput is larger than the preceding average
2020-06- 03T12:02:57Z	93	91	583	100.08%	428.05 MiB/s	334.61 MiB/s	2	Good link efficiency, currence average throughput is larger than the loceding age
2820-86- 03T11:59:17Z	91	92	697	100.08%	422.95 MiB/s	324.23 MiB/s	2	Good link efficient erage throughput is larger than the precedi
2020-06- 03T11:55:49Z	89	87	626	100.00%	414.37 MiB/s	313.26 MiB/s	2	Good link efficiency, a rage throughput is larger than the preced
2020-06-	87	86	646	100.08%	408.32 M18/s	302.03 M1B/s	2	Good link efficienc

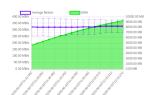
Icons made by Freepik from www.flaticon.com

4

Clear situation







First Previous 1 Next Last

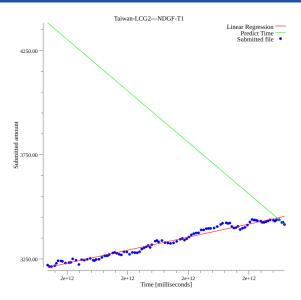
Success rate > 74 %

Timestamp	Decision	Running	Succ Queue	ess rate (last 1min)	Throughput	EMA	Diff	Explanation
2020-06- 03T12:19:27Z	100	168	551	100.00%	448.34 MiB/s	371.37 MiB/s	1	Good link efficiency, current average throughput is larger than the preceding average. Hit upper range limit
2828-86- 03T12:15:89Z	99	97	563	100.00%	442.48 MiB/s	362.82 MiB/s	2	Good link efficiency, current average throughput is larger than the preceding average
2020-06- 03T12:10:41Z	97	96	567	100.00%	436.65 MiB/s	353.98 MiB/s	2	Good link efficiency, current average throughput is larger than the preceding average
2020-06- 03T12:06:39Z	95	93	572	100.00%	436.42 M1B/s	344.79 M1B/s	2	Good link efficiency, current average throughput is larger than the preceding average
2820-86- 03T12:02:57Z	93	91	583	100.00%	428.05 MiB/s	334.61 MiB/s	2	Good link efficiency, currence average throughput is larger than the loceding age
2820-86- 83T11:59:17Z	91	92	687	100.00%	422.95 MiB/s	324.23 MiB/s	2	Good link efficient erage throughput is larger than the precedi
2020-06- 03T11:55:49Z	89	87	626	100.00%	414.37 MiB/s	313.26 MiB/s	2	Good link efficiency, a rage throughput is larger than the preced
2020-06-	87	86	646	100.00%	408.32 M1B/s	302.03 M1B/s	2	Good link efficiency or a

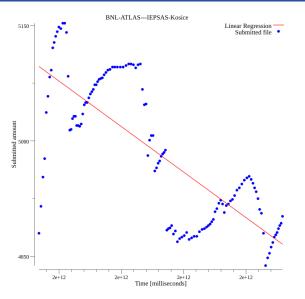
Icons made by Freepik from www.flaticon.com

4

Can we predict finish time?

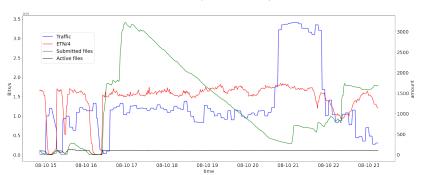


Can we predict the finish time?



Can we predict the finish time?

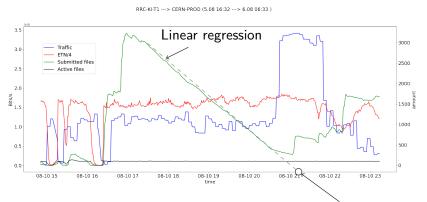




ETN (Estimated traffic network) = $\frac{c*\text{Last.AvgFileSize*Last.ActiveCount}}{\text{time rines last report follows}}$, c = const.

We can predict the queue finish time,

but it doesn't mean we can predict when the transfer will finish in this way.



Predicted queue finish time

We can predict the queue finish time,

but it doesn't mean we can predict when the transfer will finish in this way.



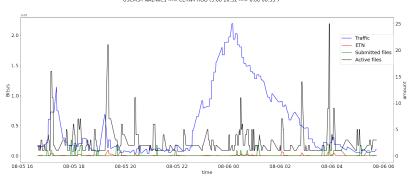
Problem Solution Understanding the problem Tech side Project Analyze Conclusion v1 Real-Life Examples Conclusion v2

We can predict the queue finish time,

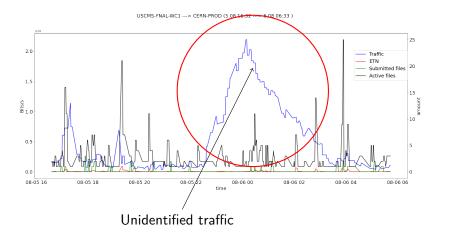
but it doesn't mean we can predict when the transfer will finish in this way.



USCMS-FNAL-WC1 --- > CERN-PROD (5.08 16:32 --- > 6.08 06:33)



Sometimes we can not "describe" the traffic



Conclusion (Transfer broker)

• If we add an extra path, we can optimize the transfer time.

Conclusion (Transfer broker)

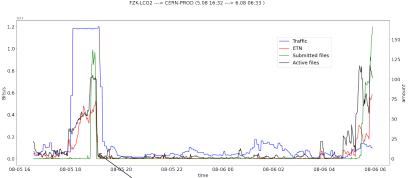
- If we add an extra path, we can optimize the transfer time.
- Using information from FTS we are able to find huge transfers. Transfers can change during observation. Hence, knowing what parameters mean is important to decide how to react.

Conclusion (Transfer broker)

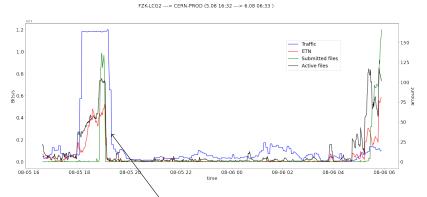
- If we add an extra path, we can optimize the transfer time.
- Using information from FTS we are able to find huge transfers. Transfers can change during observation. Hence, knowing what parameters mean is important to decide how to react.
- We can predict the queue finish time using an easy linear regression model.

Problem Solution Understanding the problem Tech side Project Analyze Conclusion v1 Real-Life Examples Conclusion v2

Is it enough to look only at submitted files?

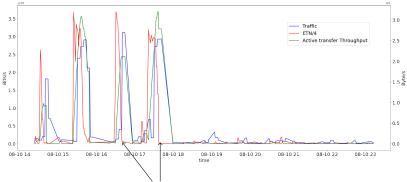


The queue was being observed for only 20 minutes. It started one hour after the transfer which caused increasing traffic. We should consider if we could react faster. Obviously, the queue finished a bit earlier than the transfer. Hence, maybe it is a good idea to focus on the submitted **and** active files (not only on queues).



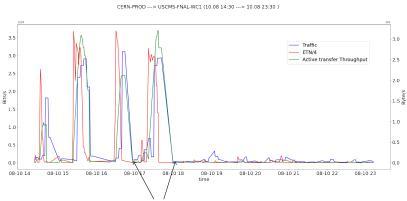
Transfer has finished, and we can observe decreasing traffic





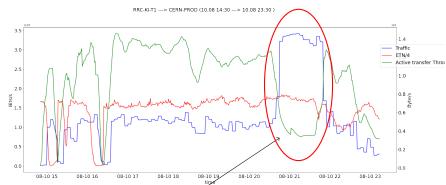
Transfers had finished before traffic decreased.

Coming back to the idea that includes the parameter "throughput"

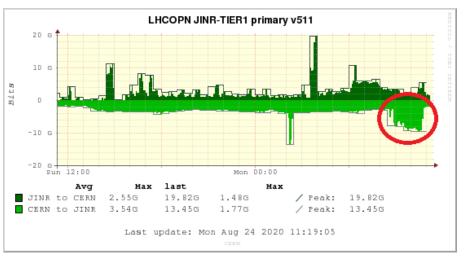


"Throughput" decreased at the same moment as the traffic.

But, if we come back to the previous example...

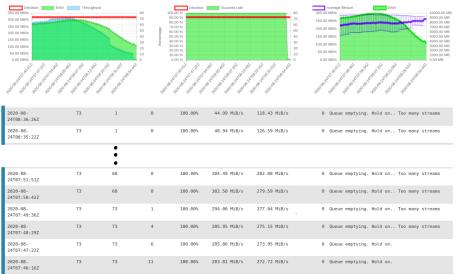


"Throughput" decreased and remained at a lower level concurrently with traffic unidentified in FTS.



It should be a transfer from CERN to JINR which finished at 10:40 a.m.

Details for gsiftp://eoscmsftp.cern.ch → srm://srm-cms-mss.jinr-t1.ru Q Decision EMA Throughput Decision Success rate 90.00 %



NOTED Network Optimized Transfer of Experimental Data

283.87 MiB/s

271.48 MiB/s

0 Queue emptying, Hold on,

2020-08-

24T07:45:08Z

73

74

15

100.00%

${\tt Details \ for \ gsiftp://eoscmsftp.cern.ch \ _ \ srm://srm-cms-mss.jinr-t1.ru} \ ^{\alpha} Throughput" \ was \ decreasing.}$

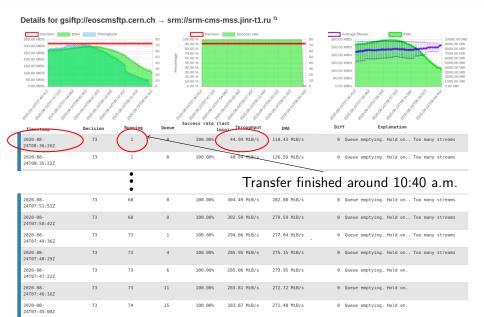


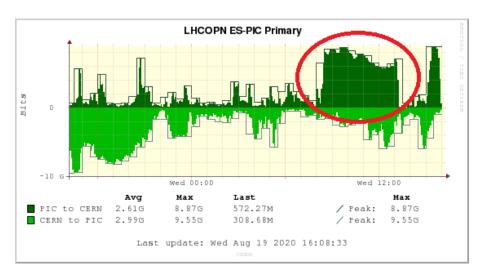
Details for gsiftp://eoscmsftp.cern.ch → srm://srm-cms-mss.jinr-t1.ru Q

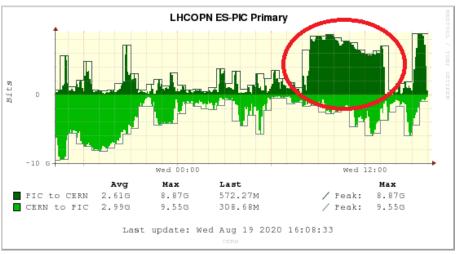


Queue finished at 9:50 a.m. CEST (08:40 a.m. UTC)

	,						`	,
2020-08- 24T07:51:51Z	73	68		100.00%	304.49 MiB/s	282.08 MiB/s	θ Queue emptying. Hold on Too many	streams
2020-08- 24T07:50:42Z	73	68	θ	100.00%	302.50 MiB/s	279.59 MiB/s	θ Queue emptying. Hold on Too many	streams
2020-08- 24T07:49:36Z	73	73	1	100.00%	294.06 MiB/s	277.04 MiB/s	θ Queue emptying. Hold on Too many	streams
2020-08- 24T07:48:29Z	73	73	4	100.00%	285.95 MiB/s	275.15 MiB/s	θ Queue emptying. Hold on Too many	streams
2020-08- 24T07:47:22Z	73	73	6	100.00%	285.06 MiB/s	273.95 MiB/s	0 Queue emptying. Hold on.	
2020-08- 24T07:46:16Z	73	73	11	100.00%	283.81 MiB/s	272.72 MiB/s	θ Queue emptying. Hold on.	
2020-08- 24T07:45:08Z	73	74	15	100.00%	283.87 MiB/s	271.48 MiB/s	θ Queue emptying. Hold on.	

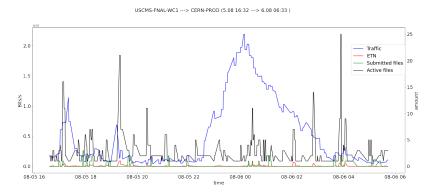






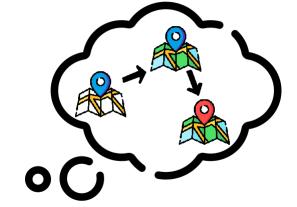
It **should** be a transfer from PIC to CERN which finished at 1:15 p.m.

Sometimes we can not "describe" the traffic. Is it true?



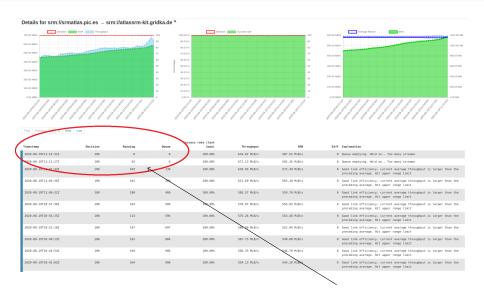
Problem Solution Understanding the problem Tech side Project Analyze Conclusion v1 Real-Life Examples Conclusion v2

Transfer PIC \longrightarrow DE-KIT, could be PIC \longrightarrow CERN \longrightarrow DE-KIT

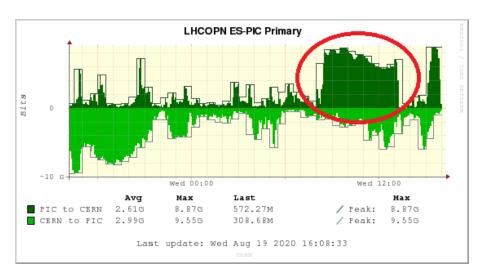




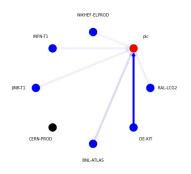


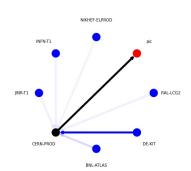


Transfer finished around 1:15 p.m.



Network controller has to know network topology



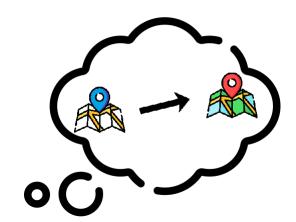


Information about transfers from FTS.

Real transfers.

Problem Solution Understanding the problem Tech side Project Analyze Conclusion v1 Real-Life Examples Conclusion v2

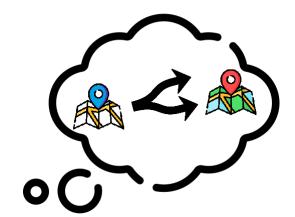
Previous idea







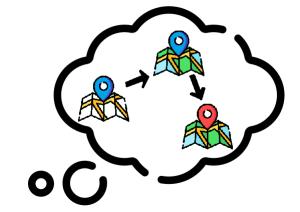
Previous idea







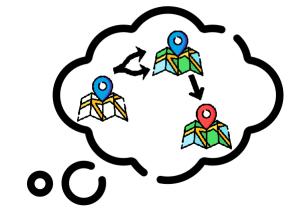
To develop an idea, let's look at the bigger picture







To develop an idea, let's look at the bigger picture







If we add an extra path, we can optimize the transfer time.

• If we add an extra path, we can optimize the transfer time. But the network controller has to know where the path is overloaded.

- If we add an extra path, we can optimize the transfer time. But the network controller has to know where the path is overloaded.
- We have to look at all points in our network topology not only for interesting us Source and destination. It is because we may no have a direct path.

- If we add an extra path, we can optimize the transfer time. But the network controller has to know where the path is overloaded.
- We have to look at all points in our network topology not only for interesting us Source and destination. It is because we may no have a direct path.
- If we want to optimize the network, looking at all parameters (not only queues) could be more effective.

Questions

Contact:

Joanna Waczyńska joanna.waczynska@cern.ch IT-CS-CE

> Edoardo Martelli edoardo.martelli@cern.ch IT-CS-CE

Edward Karavakis edward.Karavakis@cern.ch IT-ST-PDS