ION
ASSET ARCHITECTURE

Research

# Eurex T7 Timings 3

06 October 2013

Dennis Lohfert

## Introduction

In a previous note, we examined the new performance characteristics of the new architecture of Eurex, known as T7 (also known as NTA). At the time of publication, we had just seen the first upgrade iteration in this architecture, the upgrade to Version 1.1, which hit rack shelves on 01 July. There was not enough data to meaningfully analyze the effects of this upgrade, so we are due for a catch up. Readers are encouraged to refer to the last issue in order to reference some of the concepts, particularly as far as the meaning of various measurements is concerned.

Version 1.1 was primarily a performance-tuning update, introducing a few enhancements on the functionality side, but most importantly increasing throughput and decreasing latency.

The Eurex engineering team has been maintaining an aggressive update schedule with Version 1.2 being deployed on 09 September 2013 and Version 2.0 being slated for release just two months later at the end of November. This is an unusually frequent update schedule for a derivatives exchange and is an example just how much development work is being poured into the platform currently. It is definitely paying off, as we will see in a moment.

We can analyze and visualize the improvements in different ways. Perhaps the most relevant and most obvious is looking at how transit times are affected. To do so, we collected the statistics spanning 2.5 million DAX Futures orders that we sent between the time T7 became the new platform for interesting futures contracts and the end of September, a period spanning about 150 calendar days. 99.99% of these orders were passive limit orders, that is they were intended to provide liquidity at the time of sending.

## In Numbers

In order to calculate the effect of the various optimizations we calculate the ratio of various transit times before and after the upgrade. In the table below we can see that all numbers are improved and some dramatically more so than others. This is because there are two performance factors that have been improved simultaneously: base latency and throughput. Different parts of the curve are impacted to different degrees. For example, the 99$^{th}$ percentile improvement is primarily a reflection of improved throughput capacity (almost 100% more throughput, which cuts the transit time at the 99$^{th}$ percentile in half).

| | Core Transit Time | | | | Gateway Transit Time | | | |
|---|---|---|---|---|---|---|---|---|
| | Avg | Min | 50% | 99% | Avg | Min | 50% | 99% |
| T7 Version 1.0 | 134 | 60 | 84 | 842 | 350 | 210 | 293 | 1202 |
| T7 Version 1.1 | 97 | 37 | 68 | 454 | 275 | 153 | 244 | 707 |
| Reduction | 28% | 38% | 19% | 46% | 21% | 27% | 18% | 41% |

timings are in microseconds

If we look at the official numbers across the entire exchange show below, we see numbers that closely match ours, except for the 99th percentile. This makes sense because our order set is much more tied to market activity. As such, they are caught more by the throughput part of the equation in many cases (as opposed to just hitting the "steady state" at random).

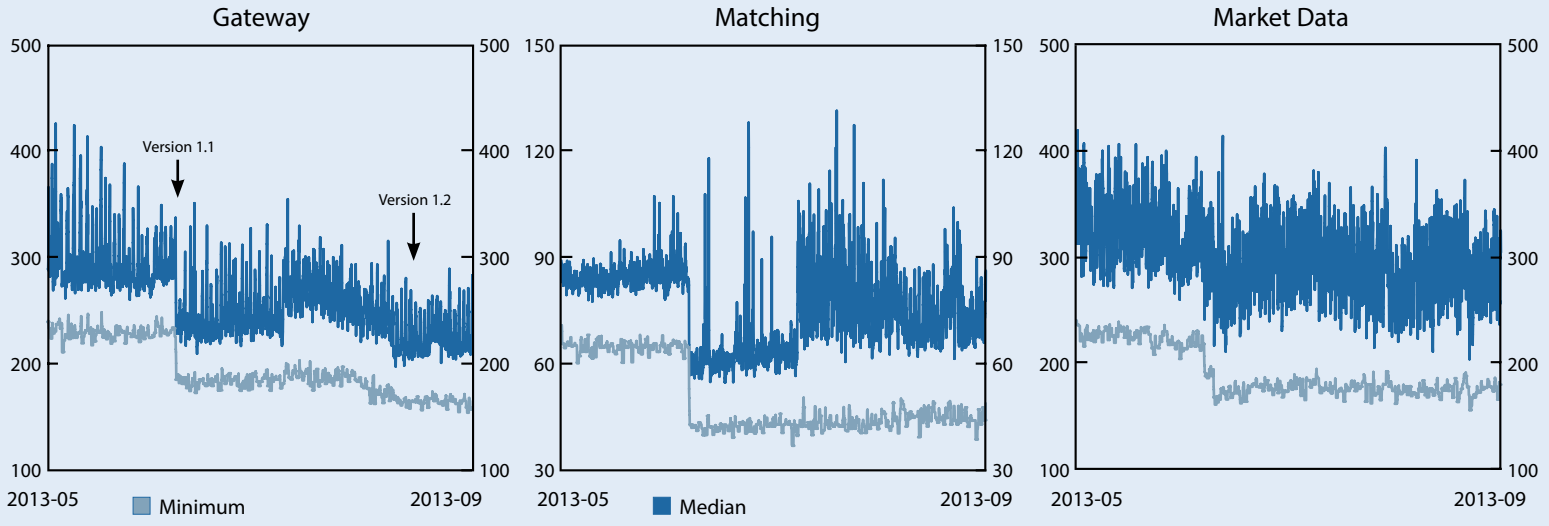
| | Core Transit Time | | | Gateway Transit Time | | |
|---|---|---|---|---|---|---|
| | Avg | 50% | 99% | Avg | 50% | 99% |
| Reduction | 27% | 25% | 34% | 21% | 19% | 24% |

In any case, the improvements rolled out in July point to a 25% reduction in response time and roughly a doubling of throughput numbers. Both are very respectable numbers and we expect Eurex being able to make additional improvements. As mentioned earlier, they are maintaining an impressive update schedule at the moment. Historically, their engineers have managed time and time again to incrementally squeeze more and more out of their infrastructure.

## Evolution over Time

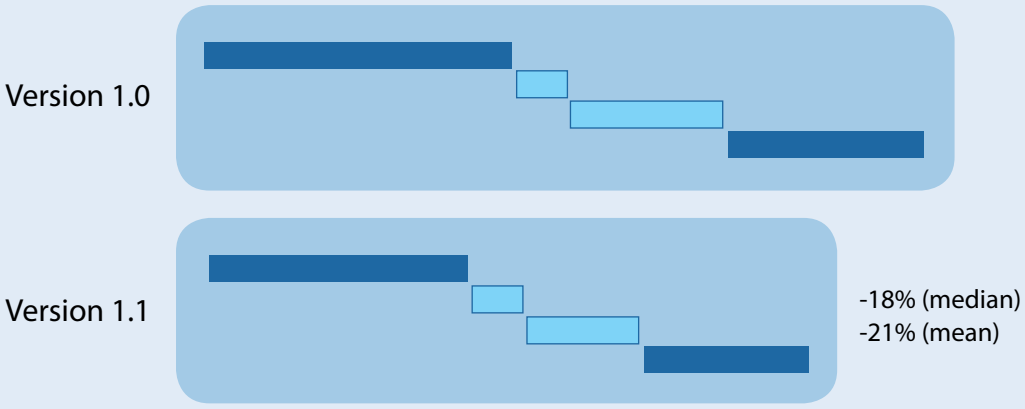
In the figure below, we plot two time series for each measurement of transit time: the minimum over a window of 100 orders and the corresponding median. Minimum transit times, as per our discussion last quarter, have special significance, as they tend to represent the "steady state". It is not what you get most of the time, but it does show what the system is capable of and is much less sensitive to noise (caused by market activity), thus making performance improvements more clearly visible. To re-emphasize: this is not what most of your orders will be seeing, it is mostly a tool to keep tabs on the steady state of the exchange system.

Figure 01: Transit Times over Time



Looking at these graphs, we can easily spot the upgrade to Version 1.1 which brought significant improvements across the entire order chain: Gateways, Matching Engine and Market Data Publishers (there are a few more pieces in the infrastructure but these are what we care about).

Figure 02: Improvement by Stage




## Where to from here?

The public eye has focused on the supposed detrimental effects of high-frequency trading in the last two years. As a result, many of the engineering efforts that have gone into creating new, reliable, deterministic and fast systems have gone largely unnoticed by those not directly involved in trading on these exchanges. These are systems that facilitate the distribution of risk and disseminate timely and accurate pricing data. Some of the engineering has been so successful that it has made fast exchange connectivity essentially a commoditized product. Commoditization tends to be the end result of good, broadly based engineering efforts.

There is ever-increasing regulatory transparency on the trading process. As a consequence of the German High-Frequency Trading Act, which came into effect earlier this year in May, there will be some new features introduced. The next pending release in November (Version 2.0) will for the first time require explicit tagging of all algorithmically generated orders for the benefit of regulators (BaFin primarily, but within the context of ESMA this undoubtedly can be shared with every other regulator).

Additionally, we will see a revised implementation of the order/trade ratio charges, which some may recognize now as “excessive system usage fees”. Regulators feel that excessive quotation activity without any resulting trades is to be curtailed. I don’t necessarily agree with that notion, but one of the hallmarks of good trading is adapting to new environments and changing circumstances.