

Low-latency networks for trading infrastructure

shaping tomorrow with you

FUJITSU

Low-latency networks for trading infrastructure



Information Technology (IT) departments across a range of industries are concerned about the amount of latency in their network infrastructures. Finance, however, is utterly fixated on the latency metric.

In the rapidly evolving world of algorithmic trading, the gap in time between when an instruction is issued and when it is executed across a network is a point of particular hypersensitivity. Along key trading routes, latency reduced to mere nanoseconds is emerging as simple table-stakes of competition in the financial industry. It's a challenging standard that demands evaluation and optimization of end-to-end infrastructures and processes—spanning all of a firm's connections to brokers, co-location providers, dealers, exchanges, hedge funds, information feeds and any other instance where lapses can be identified and eliminated. Figure 1 provides a conceptual view of the many connection points involved in such an evaluation.

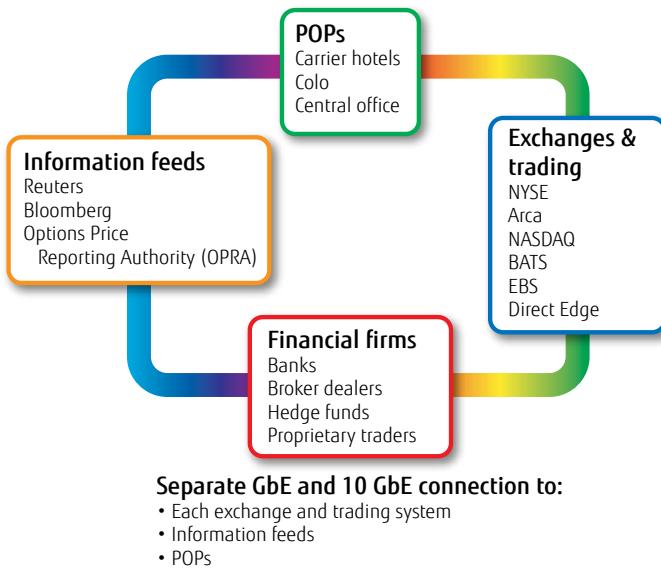


Figure 1: Metro connection points in low-latency networks

The evaporation of tolerance for latency has influenced the entire operational strategies of some financial institutions. Whereas firms once simply contracted with carriers for site-to-site connectivity over shared networks, many now demand dedicated fiber-optic infrastructures for their most important routes and most time-sensitive applications. Even the electronics used to "light" the fiber networks now command scrutiny. Traditional approaches to common optical transport functions from transmitter to receiver can yield levels of latency that were considered negligible only a couple of years ago but are untenable by today's standards.

Financial networking's hypersensitivity to latency

In High-Frequency Trading (HFT) and other forms of algorithmic trading, buy and sell orders are automated by computer-run simulation programs that interpret and react to various inputs of externally generated intelligence (price/volume market data, labor statistics, etc.) The programs must be able to receive information and deliver instructions incredibly quickly, because the first financial institutions that get their trade requests to market get the best prices. This race has grown more heated as algorithmic trading has assumed an expanding role in the financial industry, fueling order flow. HFT, which accounted for close to 75% of all U.S. equity trading in 2009, is dependent on the lowest achievable levels of latency. Arbitrage strategies based on low latency produce annual aggregate profits in excess of \$21 billion.

Traditionally, the strategy to improve a firm's algorithmic-trading posture has been to task mathematicians with tweaking the computer models in hopes of more precisely predicting future market events. A firm might have further upgraded the servers and switches on which the computer programs run.

But deploying the highest-capacity, screaming-fast router or multi-core processor computer blades will not suffice today. Contracting with a carrier for the highest bandwidth connection to an exchange is not sufficient either. The relentless pace of innovation in low-latency networking has rapidly raised the competitive stakes in algorithmic trading, to the point that success in contemporary finance forgives latencies no greater than nanoseconds. And that fact has heaped tremendous, unprecedented pressure on the managers of financial networks to deliver even the most modest of reductions in delay anywhere in the ecosystem of equipment and processes that enable a firm's algorithmic-trading applications. This challenge demands uncovering and eliminating tiny traces of latency in the building-to-building fiber-optic transport connections themselves.



Uncovering latency in optical networks

Enterprises can undertake a number of technology improvements to improve latency: adoption of faster central processing units (CPUs) and network interface controllers (NICs), accelerated middleware appliances, and low-latency switches. Firms that require additional improvements in order to achieve ultra-low latency and optimize transaction rates can turn next to three common sources of delay in optical transport (Figure 2):

- Fiber delay: the length of the actual fiber cables
- Proximity delay: the physical distance to the fiber connection
- Equipment delay: the processing speed of the network equipment

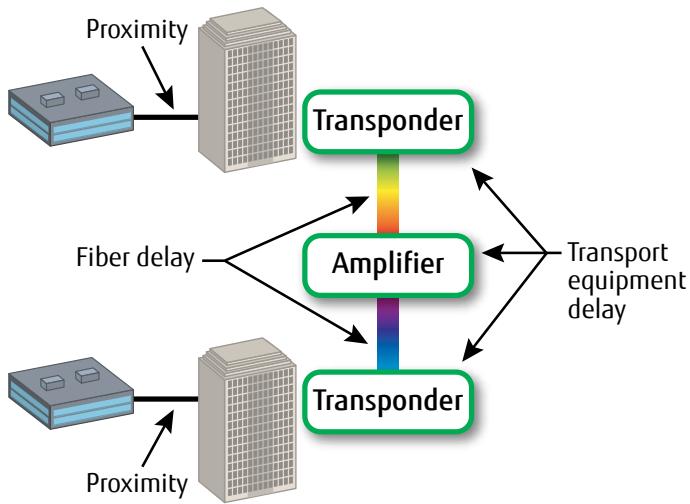


Figure 2: Sources of latency in optical transport networks

Fiber delay

It's a simple fact: the longer the physical route, the greater the time it takes for traffic to get from one end to the other. Cut roughly eight inches of optical fiber between two connected locations, and you cut a nanosecond of transport latency between the two sites. The competitive advantage of minimizing fiber delay can be significant for enterprises that conduct transactions in major cities, because the fiber routes within metropolitan areas often snake up and down manholes, across streets and along whatever available easements the fiber installer could secure. Enterprises requiring ultra-low latency must closely evaluate the directness or circuitousness of the routes available for dedicated infrastructures from dark-fiber providers.

Proximity delay

The space near a fiber access point is a finite quantity, and the firm that can locate its equipment closer to the fiber access point than can another will—all else being equal—realize competitive advantage in the speed with which transactions can be processed. Given that the real estate near fiber-junction points is in such high demand, some enterprises reach out to lease space from co-location providers.

Equipment delay

Finally, there's the actual speed of the optical equipment across the end-to-end transport network. When an enterprise's acceptable level of latency is measured in full seconds, the delays injected by the gear that traffic encounters along a fiber network are probably not of consequence. But it is incorrect to believe that data in a glass network actually moves at the speed of light. Every device introduces some amount of delay, and, therefore, any non-optimized transport equipment across the information path can thwart business models that are based on the most challenging network latency limits.

Eliminating these sources of equipment delay is trickier because different devices carry out key, common optical network functions in very different ways, yielding very different degrees of latency.

A deeper look at transport latency

Color conversion, amplification, dispersion compensation and regeneration are all common optical-networking functions that are carried out among different optical-networking solutions. The enterprise that fails to understand how its Wavelength Division Multiplexing (WDM) gear accomplishes these four basic functions risks adding significantly to the time it takes for network instructions to be executed after being issued. And, in latency-sensitive algorithmic trading, a firm's whole business strategy could be jeopardized by ill-fitting implementations of necessary optical-transport functions across the fiber strands linking key locations.

Color conversion

In color conversion, traffic signals are converted, or "transponded," from gray to a specific light color (wavelength). Frequently, multiple colors are aggregated, or "muxponded" into a single high-speed channel across a fiber-optic network. Electrical Time Division Multiplexing (TDM) techniques might incur milliseconds of latency. Examples are:

- Optical Channel Data Unit (ODU) encapsulation and thin film filters
- Forward Error Correction (FEC) algorithms
- Performance monitoring
- Protocol conversion
- Clock recovery,

Best-in-class, purpose-built transponders and muxponders have emerged that produce latency in only the high nanoseconds to low microseconds.

Amplification

Traffic signals naturally weaken as they travel mile after mile along an optical fiber network. Optical amplifiers are designed to offset this weakening effect by boosting the signal. The Erbium-Doped Fiber Amplifier (EDFA) is one of the most commonly deployed types of amplifier device. But, because delay of hundreds of nanoseconds is introduced by each EDFA on an optical route (even microseconds in the case of some high-gain, dual-stage EDFAs), the most latency-sensitive enterprises must instead turn to alternative, optimized architectures that do not rely on integrated doped fiber spools or that add significant "noise," for example. High amounts of noise and low gain figures both require higher numbers of amplifiers, which increases latency. Recently introduced innovative Raman amplifier technologies do not require the addition of supplementary fiber, leading to a shorter optical path and more efficient transmission. Raman amplifiers also deliver low-latency inline amplification from within the fiber itself.

Dispersion compensation

Signals carrying data traffic also degrade because of "chromatic dispersion," a phenomenon in which usually high-speed optical channels smear into a rainbow of colors and bleed into neighboring traffic streams, which adversely affects performance. Dispersion compensating fiber (DCF) can successfully reverse that effect, but the long spools of special fiber required to accomplish this objective create unwanted latency. In addition, when improperly located, DCFs demand additional amplifiers in a fiber network—producing even more delay. Alternative remedies that leverage Fiber Bragg Gratings (FBGs) can provide dispersion compensation with negligible amounts of latency.

Regeneration

Signal regeneration is another common optical network function. Variants of the feature are designed to offset signal degradation over longer links, but any of the traditional implementations of electrical regeneration can introduce delay as well as jitter. Techniques that entail FEC termination and setup, or non-optimized routing and cabling among filters and cards, for example, can inject hundreds of microseconds of latency. Purpose-built, low-latency approaches to regeneration slash this delay into the nanoseconds.

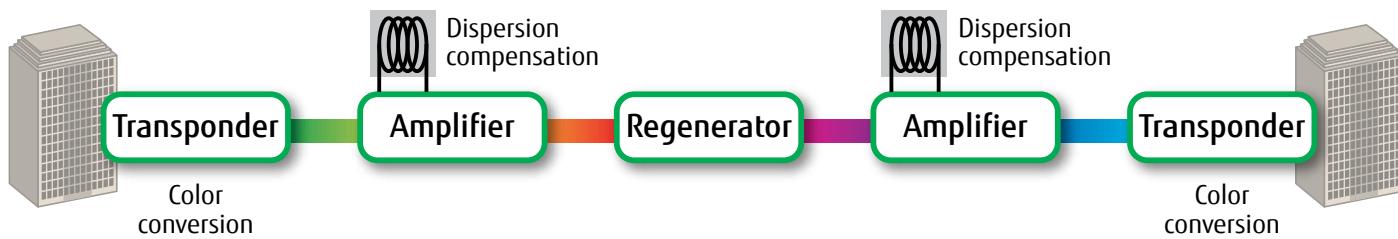


Figure 3: Optical networking functions that can increase latency

Fujitsu network solutions for low-latency trading

Constructed expressly for the job of low-latency networking, the Fujitsu FLASHWAVE 7420 WDM platform delivers breakthrough performance for algorithmic-trading applications. Only the Fujitsu platform—employing transparent wavelength conversion and purpose-engineered techniques for inline amplification, dispersion compensation and signal regeneration—enables financial network managers to eliminate so many sources of optical transport latency end-to-end across infrastructures. Not only does this provide a trading firm with the fastest-possible connectivity along trading routes, the FLASHWAVE 7420 platform goes further to offer in-service latency measurements that help an organization keep algorithms optimally tuned to exploit quickly opening and closing windows of market opportunity.

Furthermore, the FLASHWAVE 7420 system delivers flexible, cost-effective, protocol-agnostic support for all of the leading services in play across financial networks, such as Ethernet, InfiniBand, Fibre Channel, Fibre Channel over Ethernet (FCoE)/Data Center Bridging (DCB) and Fiber Connection (FICON). The Fujitsu solution also offers flexible scalability (fast expansion of up to 80 transmission channels of service over two strands of fiber), cost-effective power and space efficiency, unmatched distance support of fiber spans of up to 2,000 kilometers, physical-layer encryption and optical-line monitoring, and a safety net of best-in-breed technology, personalized network monitoring and application support to ensure highest reliability. Beyond the industry-leading FLASHWAVE 7420 platform, Fujitsu offers financial network managers valuable consultancy in designing the optimal, lowest-latency solution for their needs. Fujitsu also offers an array of latency-optimized cards, and ease-of-use support (interoperability tests, implementation and setup, surveillance and operation, and key partnerships). Together, these convey to financial providers a sustainable competitive advantage in low-latency trading—available uniquely from Fujitsu.

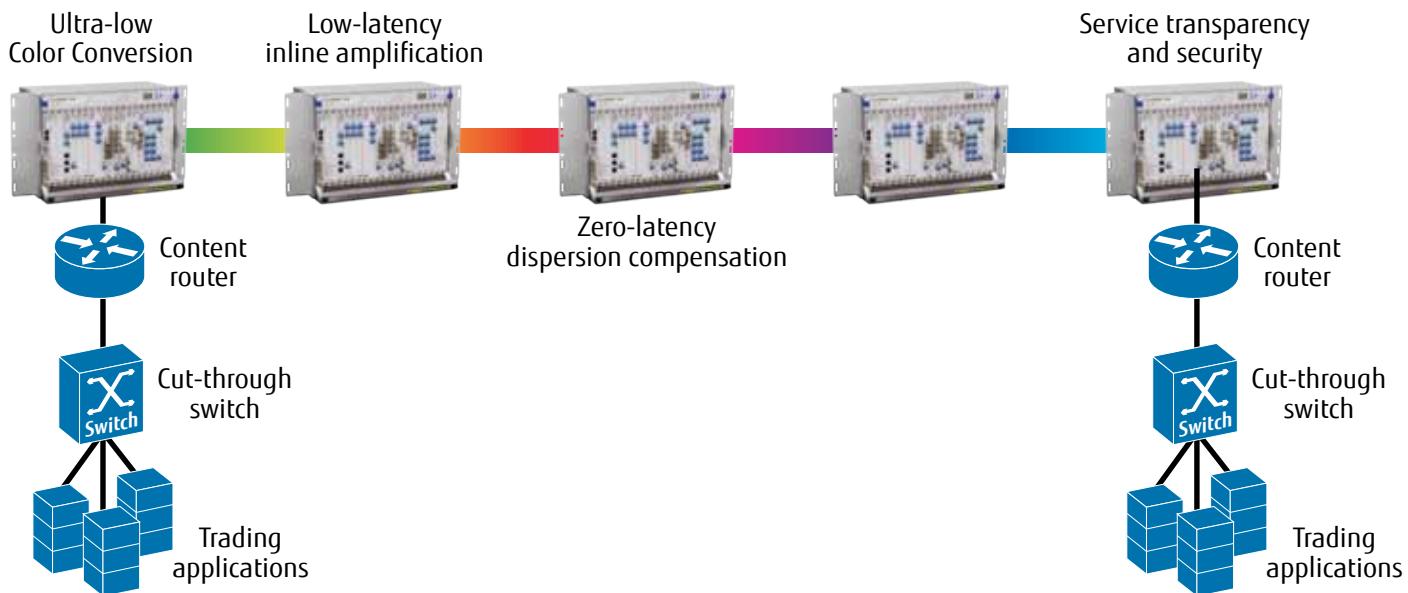


Figure 4: End-to-end low-latency advantage