# Designing FSO Links for Flexible Inter-Rack Networking

Our goal in this section is to present the design of FSO transceiver modules that have the needed properties for the design of a flexible inter-rack fabric, i.e., (i) size, power and cost effectiveness, (ii) ability to provide 10-100Gbps data rate, (iii) ability to align with the needed precision, (iv) ability to steer precisely and with low latency. The conventional use case for FSO has been terrestrial long distance (miles) or even satellite and space communications (1000s of miles or more) []. These use cases have just a single link, as opposed to a very large network of laser links that Firefly requires. As such commodity FSO systems are not designed to optimize on size, power or cost as aggressively.[[1]](#footnote-1) In addition, while alignment issues are indeed addressed, the above use cases do not require fast steering ability. Thus, new FSO systems must be engineered grounds up for use in Firefly– a challenge that we undertake in the proposed project. Of note is that some of the significant challenges faced in outdoor optical propagation (e.g., beam path variations due to scattering from fog or dust) largely disappear in datacenters thus presenting a different set of design tradeoffs. A basic theme of our design approach is using of commodity components as much as possible and building an optimized system around them. This makes an entire end-to-end demonstration and evaluation possible in realistic datacenter scenarios (described in Section XXX) within the timeline of the project. While this prototype will be of small scale, we articulate below a pathway to design large-scale scalable systems.

## 3.1 Design for Low Cost, Size, Power and Reliability

An FSO communication link consists of three basic components: i) a *transmitter* (TX) that is a modulated laser source (typically infrared), ii) a *receiver* (RX) that is a high-speed optical detector along with a demodulator, iii) the *propagation channel* represents themedium and path that the light traverses between TX and RX. Instead of a first-principles design of TX and RX, we propose to repurpose commodity optical *small form-factor pluggable* (SFP) transceivers [] that are typically used to interface optical fibers with (electrical) packet switches. The advantage of this approach is that we can exploit the large commodity market of optical SFPs that already operate at 10+ Gbps. The optical SFPs are very small in size (state size??) and already contain very reliable, field-tested TX and RX components with widespread deployment. They are also indeed low power (state power??). In any case, datacenters often use them regardless of the network architecture, whenever optical fibers need to be used to support high data rates at distances that standard copper cables are unable to carry [].

### Optical Design

The optical SFP is designed to interface directly to an optical fiber[[2]](#footnote-2) that constantly re-confines the laser beam within the narrow fiber by internal reflections. Without the fiber, the beam launches into free space and *diverges* in a cone as it propagates – quickly losing power. Divergence can be arrested using a suitably designed collimating lens on the optical path near the TX. This makes the laser beams roughly parallel, now diverging at a very slow rate – in the order of milli-radians or less, which often could be ignored except at the extreme distances in the datacenter. Finally a similar lens near the RX focuses the beam back onto the detector. Now, from basic optics an inverse relationship exists between the diameter of the propagating laser beam at the so called ‘beam waist’ (the narrowest part of the beam near source) and the rate at which it diverges (divergence angle). Thus the optical design is critical so that a beam of the right diameter can be formed. While larger beam waist are better to keep the divergence negligible this also requires a larger lens (size issue); smaller diameters on the other hand may make the beam diverge too quickly for it to be strong enough at the extreme distances within the datacenter (say, ~100m). This tradeoff is crucial in designing the FSO links in Firefly and influences the optical design. Beam diameter also influences alignment challenges that we will describe now.

### Alignment

The optical detector at RX gathers energy proportional to its size (in the order of XX for SFPs). Ideally, we do not want the beam to have a much larger diameter at the RX than the detector size lest only a very small fraction of the power will be captured that may not be enough for reliable detection. Thus, both large diameter beam waist and large divergences are poor design choices. On the other hand, large diameter helps in alignment as small shifts in the optical path will have negligible effect in the detector so long as the total energy received at the detector is above the detection threshold. This helps recovering from rack vibrations, jarring effects due to maintenance, optical path drifts due to air temperature variations, and other similar issues. If the beam diameter itself is not sufficiently large to account for these small positioning small corrections of the TX and RX will be needed. Piezoelectric positioners or thermal heaters with a high-thermal-expansion material can provide these needed precision corrections at reasonable costs. The feedback needed to these corrections can be obtained from the DOM (digital optical monitoring) support already present in the optical SFP standard and carried on the I2C bus via the connectors on the module. If corrections on the RX side is insufficient and the TX side needs to be adjusted as well, the RF-based control mechanism will be used (discussed in Section 6).

In a datacenter, the space above the racks is a natural choice for laser propagation as this space is roughly free from obstruction. The FSO devices will be anchored on the top of the rack and connected to ToR switch ports. The devices themselves will either be staggered so that they themselves do not present obstructions or mirrors in the ceiling (and possibly additional mirrors on the beam path for redirection) will be used to avoid obstructions. Overall, we anticipate that a single FSO assembly including the alignment and beam redirection machinery can be put together within about 3”x8” footprint such as that a few tens of such devices can be packed on the top of a rack. A critical task in the project will be designing the optical elements (e.g., mirrors, lenses) and their precise positioning such that they operate for wide range of distances, e.g., few meters for neighboring racks or over 100m for distant racks in a large data center.

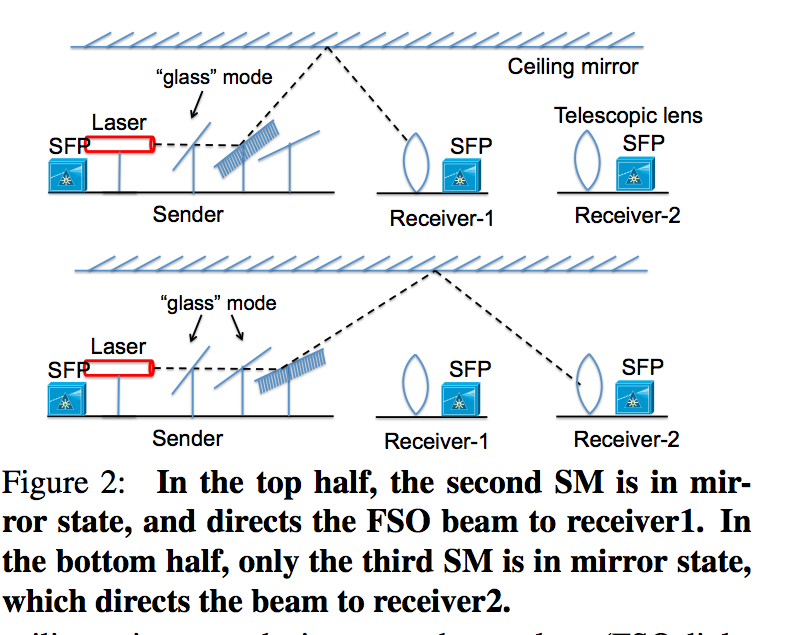
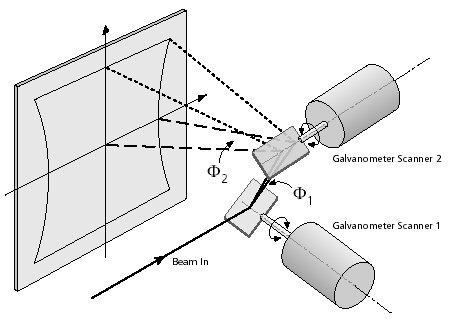
## 3.2 Reconfigurability via Beam Path Redirection

In Firefly the ToR switches are to be interconnected via the FSO links to create an FSO-based inter-rack fabric. To achieve reconfigurability in this fabric, the beam from the TX must fundamentally be able to redirect between multiple possible RXs on top of other racks. We have investigated two mechanisms to achieve this, one via switchable mirrors and the other via beam steering. They both present certain tradeoffs. The impact of such tradeoffs in the overall performance is unclear without a careful analysis. Such an evaluation will be part of our work. We anticipate that we will pick one or a combination design in our final prototype and in future research.

### Approach I: Dedicated-Alignment — Switchable Mirrors

In a *dedicated alignment* approach each beam path is manually oriented during statically and remains fixed during normal operation. The particular beam path is then selected among several possible beam paths. One approach to do this is to use *switchable mirrors* (SMs) made from a special liquid crystal material that can be electrically controlled to rapidly switch between reflection (mirror) and transparent (glass) states [].

Referring to Figure SM\_FIG(a), each FSO device will be equipped with multiple SMs. Each SM will be pre-aligned to a dedicated beam path. The desired link is established by placing one of the SMs on the TX in the mirror state and the other SMs in the transparent state. An analogous arrangement will be made at the other end (not shown). At any instant, only a subset of candidate links is active (one per FSO) based on the SMs’ states. A mirror fastened to the ceiling redirects the beam back to the receiving rack, making efficient use of the above-rack space while minimizing interference. The pre-alignment for each SM is done statically at configuration time. This is directed by pre-configuration topology design (Section 4). A custom, hand-held mechanical adjusting mechanism can be used that will be attached temporarily to SM assembly and adjust each mirror along the calculated orientation.

 [](http://www.google.com/url?sa=i&rct=j&q=&esrc=s&frm=1&source=images&cd=&cad=rja&docid=mfsr3XEzoZtjjM&tbnid=rJ136oICtaHt1M:&ved=0CAUQjRw&url=http://www.zamisel.com/SSpostavka2.html&ei=uLVUUvePDOT54AOCjoFw&bvm=bv.53760139,d.dmg&psig=AFQjCNFpHBYLx9VALecmxk7PiwXM_sipNw&ust=1381369648042364)

**Figure SM\_FIG: Beam re-direction approaches. (a) A dedicated-alignment beam re-direction approach. (Top) The second SM is in reflection mode, redirecting the beam to Receiver 1. (Bottom) The second mirror is now transparent and the third mirror is in reflection mode, directing the beam to Receiver 2. (b) Beam steering with GM. Two mirrors direct incident beam into a rectangular cone.**

### Design Approach II: Beam Steering — Galvo Mirrors

In *beam steering*, the optical system can dynamically re-direct the beam on command and in real time. This requires computer-controlled movable optics. Among the many possible candidates, one promising approach is to use a *galvo mirror* (GM) system. Shown in Figure SM\_FIG (b), two computer-controlled, motorized mirrors are mounted at right angles direct incident beam into a rectangular cone. The (fixed) incident beam can thus be directed into a rectangular cone under computer control. Commercially available systems [] can provide a cone half angle ( and ) of ±20°, for a total rectangular cone angle of 40° in both directions. A typical pointing accuracy is within 15 µrad [], resulting in a beam positioning precision within 1.5mm for beam paths of up to 100m. GM provides some advantages relative to SM. Since GM can steer on a continuous scale no additional alignment mechanism is necessary as in SM. Also, GM can re-direct the beam to any number of FSOs within the cone in contrast to SMs that provide only a small number of beam paths. The re-direction latencies are however comparable or better in GMs. However, the limitation of GMs is a limited steering angle that makes the network topology dependent on layout geometry and rack locations. To address this pre-alignment can be used just like SMs, or a third servo and mirror combination can be used to increase the angle. Commercially available GMs are also more expensive than SM. They may not have the form factor we desire, but a custom design is certainly possible to achieve custom angles and form factors. Exploring such design choices will be part of our study.

### Design Approach III: Integrating Emerging Technologies

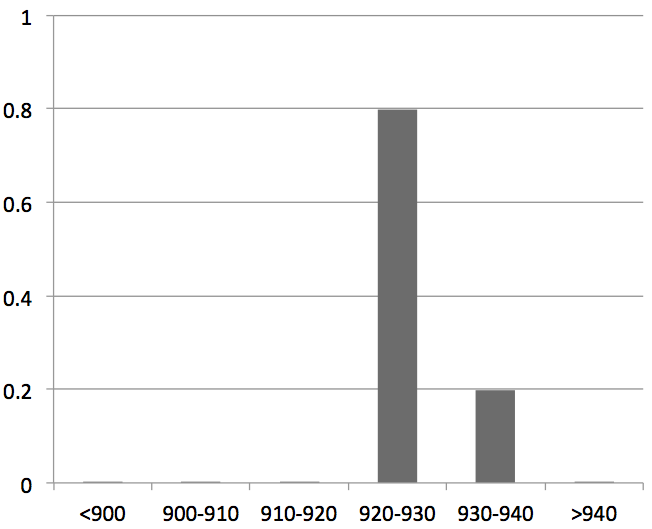
The previous design approaches are based on technology available today, to provide a guaranteed path forward. During the project, however, an exhaustive search of emerging and less-known technologies will be conducted to explore additional options in the photonics and optoelectronics communities for beam steering, alignment, and range. Examples include MEMS-based switchable mirrors [], low-cost, high range-of-motion servos, low-cost, high-performance aspheric focusing optics (as used in phone cameras), and custom-designed systems built in-house.

## 

**Figure OPTICS\_LAYOUT: Experimental prototype showing FSO communication using SFP. (THE PIC WILL CHANGE SHOWING THE ENTIRE LINK)**

## Preliminary Work: Proof-of-Concept Prototyping

A proof-of-concept prototype has been developed to demonstrate free space operation using SFPs. See Figure OPTICS\_LAYOUT. It uses a pair of 1Gbps SFPs using 1310nm laser. Instead of launching the beam directly from the SFP we launch from a single mode optical fiber that is connected to the TX SFP on one end with the other end terminating in free space. Due to the narrow 8-10 μm fiber diameter the initial beam divergence is very large. An achromatic doublet lens is used to collimate the beam to a roughly 4mm diameter waist with the fiber tip positioned at the focal point of the lens. An optical bench and translating mounts help in the positioning. The collimated beam propagates up to a distance of 7.5m where an identical lens re-focuses beam on the detector of the RX SFP. Since the SFP we used uses two separate optical paths (for duplex operation), the return link is closed using a regular fiber. This way standard network protocols can be run to characterize the free space link. Two laptops are connected to the SFPs via media converters and TCP throughput experiments (with forward direction going over free space) are run for over 30 hours continuously. See the distribution of per sec throughputs in Figure THROUGHPUT that demonstrates an extremely stable link. It does not show any statistical difference from the ‘wired’ case (both links using fiber). The quality of link is also studied when the TX-RX are misaligned. The TCP throughput is stable up to a transverse shift of 0.7mm showing a great promise in addressing alignment issues. Beyond this the throughput drops sharply going to zero within another 0.1mm.



**Figure THROUGHPUT: Distribution of per-sec TCP throughput (in Mbps) for the 1Gbps FSO link.**

We have also built a proof-of-concept prototype to evaluate the viability of switchable mirrors. Here, we have used a 12” x 15” switchable mirror (SM) from Kentoptronics [] tuned for the IR spectrum; and normal mirrors. The switching latency of the SM is found to be around 250 msec. Because the switching latency is proportional to the SM’s surface area [], we estimate a <5 msec latency for a small (1” x 1”) SM we propose to use. Finally, we have confirmed that the FSO beam can be reflected from conventional mirrors with no loss in TCP throughputs even after multiple reflections.

1. Significantly powerful laser sources are needed anyway to cover long distances. Typical commercially available systems for terrestrial applications [] are roughly 2 cubic feet, costs $5-10K for a single link and consumes XX watts. They indeed include elaborate alignment machinery suitable for long distance outdoor communications including recovery from building swaying, various atmospheric effects etc. [↑](#footnote-ref-1)
2. Typically, a fiber pair for bidirectional communications. Single fiber SFPs [] are also possible where the laser and detector at each end point are interfaced to the same fiber with two directions operating at two different wavelengths. As would be apparent later, use of single fiber SFPs make our approach easier as only one optical path needs to be designed per FSO transceiver. [↑](#footnote-ref-2)