Monitoring Client Experience

Bruce F. Bahlmann
MediaOne

Abstract

The customer demand for quality Internet access is prompting a change in the way Internet information services (or high-speed Internet Service – HSD) will be engineered, maintained, and marketed in the future. As a result, traditional measurement applications of Internet service will give rise to more sophisticated applications which focus on customer experience and quality.

Best Effort Quality

Today's Internet service is not very sophisticated when it comes to quality control. Almost magically, from a customer's perspective, Internet service continues to work and "seems" fairly reliable. Behind the scenes it's a constant flurry of activity where projects to upgrade circuits, equipment, and software abound. This constant change creates many challenges operationally to achieve standardization and yet get the most out of the equipment in service. As a result, quality control & maximizing return on investment are in constant conflict with the deployment of new services and stabilization of the network build out.

Providers of Internet service in this space reside in a very competitive market that requires them to constantly seek new affiliates. They use terms like "redundancy" and "high availability" to lure affiliates into signing service contracts with them. However, the language of many of these service contracts all but define the level of service or overall quality of service individual customers can expect. The service contract merely binds the affiliate to the Internet provider in exchange for Internet provider making its best effort at providing a quality Internet experience for the affiliates' customers. Sometimes the service agreement is entirely void of service quality related measurables. In this case, an additional document called a service level agreement (or SLA) is necessary to define the level of service the affiliate can expect for signing on with a particular Internet provider. Ideally, this document represents a compromise between the type of service the Internet provider is willing to supply and what the affiliate is willing to accept. For reasons that will be explained later, it is often difficult (if not impossible) to reach a compromise on the SLA. As a result many SLAs go unsigned – which means the Internet provider is under no agreement to provide a particular level/quality of service to the affiliate.

Today's Quality "Guarantee"

In cases where a service level agreement is signed the Internet provider is "bound" to the level of service expected in return by the document as well as the affiliate. Within a typical SLA are several points of interest. Notably, the "Key Performance Indicators" and the "Network Services Conformance" sections provide the operational parameters that the Internet provider has committed to supplying. Key performance indicators are focused on response to outages or escalations where network services conformance is concerned with availability. The rest of this document will focus on the subject of availability.

Availability

One of the commonly used terms with regard to providing Internet service is "Availability". Availability is defined as capable of being obtained and/or accessible

for use. Internet providers use the word availability to signify the amount of reliability they intend to provide with respect to various services they supply. Availability is typically defined in terms of percent (%) with higher percents equating to higher reliability.

The availability projections within the SLA are usually based on the Internet provider's "best effort" to measure the accessibility of the services they provide. One of most common tools in use today to measure availability is ping. The ping application communicates with Internet hosts to determine their operational status. For example if a host is operational (or "up") it is reported as "alive" by the ping application. If the host is not operational (or "down"), it reports "no response" or "request time out" by the ping application. Although the ping application is a useful operational tool on the Internet, it is not a very reliable means of measuring availability. For example, the host may be up but the application (or service) supplied by the host could be down. In this case the availability is reported incorrectly. As a result, there is a difference between application availability (measured via the application's client) and host availability (measured via ping).

Unfortunately, availability is not mentioned with respect to service windows used by the Internet provider. A service window is when Internet providers perform necessary installations, changes, and upgrades. The service agreement dictates certain times and days as potential service windows that Internet providers may use to maintain service. These days and times are usually coordinated with affiliates so the affiliate and its customers understand any resulting down time during the window. The ease at which service windows can be scheduled

and the fact that service windows rarely count against service reliability affords the Internet provider a means of constantly changing the system it uses to provide Internet service. The resulting constant change along with the lack of affiliate-initiated acceptance and/or safeguards prevents quality control from being achieved.

Surprisingly, the Internet provider often does the only monitoring of availability levels to measure its compliance established in the SLA. The Internet provider supplies this because the affiliate does not always have the means to do this on their own. However, SLA's typically do not stipulate the type of monitoring (application availability or host availability) they require. In absence of any specific request for monitoring method, host availability is likely reported as the default as it's the easiest to obtain. As a result the monitoring data reported by the Internet provider often does not reflect the actual availability seen from a typical affiliates' customer perspective. This allows the Internet provider to maintain compliance with the SLA while providing a level of service and quality of service that is actually lower. Unless the SLA is re-written to dictate the monitoring method used by the Internet provider the SLA will not represent any guarantee for level/quality of service.

Future Quality Guarantees

Fortunately, customers are beginning to sense some differences among providers of Internet service. Although price is still the biggest factor, level and quality of service are moving up fast on the importance scale. The result of customer initiated preference for quality and reliability is changing what the affiliates' need to satisfy their customers. These changes will include such things as

incentives and penalties enforced on Internet providers as they strive to meet minimum service levels dictated by the SLA.

Incentives & Penalties

To address this affiliates are realizing that availability of Internet service impacts their bottom line. The impact that availability has on things like call volume, truck rolls, and higher sales is not known at this time. However, the ability to compare availability with call volume, look for trends, and establish some relationships between the two is gaining interest from affiliates. At the time of this writing, it seems reasonable to expect that there is a relationship between call volume and availability. It is projected that further analysis could potentially derive a cost factor per customer that is absorbed by the affiliate as a result of lowered availability. Additionally, the cost calculated could in turn be used to establish minimum availability levels an affiliate will accept. Thus having a tool that could provide affiliates with up to the minute calculations on availability could help them better understand the relationships between availability and support costs and reduce the burden that lower availability has on affiliates' bottom line.

Providing motivations to Internet providers is the key to establishing realistic minimum application service levels. Obtaining the history of an Internet provider's performance, one can establish the average service availability level provided. This average availability level could then be used to drive the affiliate's required service availability levels. Combine this with impact studies above could result in the affiliate providing incentives for the Internet provider to perform above their required service availability – such has a kick back premium per customer. Like-wise, service availability levels below the required levels would result in service discounts per customer (to enable the affiliate to recover the added support costs that were the result of lower availability levels caused by the Internet provider). Providing these kinds of incentives and penalties would allow availability to be treated equally with other methods of evaluating an Internet provider's performance. The reality of the matter is that Internet providers need this level of information to make informed decisions of service upgrades and network build out.

Informed Decisions

Having the application availability information provides affiliates with the means to make informed decisions regarding escalation of calls to the Internet provider's tier two services, scheduling of service calls, and acceptance of system upgrades. In fact, this information could actually drive affiliate requests for specific application performance upgrades in some cases. Making informed decisions is the key to cost savings and reduction in outages caused by unnecessary upgrades.

Informed decisions also enable significant cost savings to Internet providers. By using availability information systems could "truly" be scaled in concert with demand thus eliminating costs of over engineering solutions for under populated areas. This would enable targeting of capital expenditures to areas of need (a type of scratch where it itches approach towards network upgrades) would permit Internet providers with a means of controlling costs and increased operational efficiency.

The ability to make informed decisions would also enable both parties to collect historical information needed for capacity planning.

Historical Data

Historical data enables the most efficient use of resources to solve problems. Establishing such things as "baselines" and "peaks" allows vendors to build products that handle the kind of beating that an Internet service in this space demands. Today vendors cannot fully understand the dynamics of the operational environment to build products that can withstand the punishment of taking on all the Internet can dish out. As a result, new products are forced to burn-in while in production mode rather than in less service-impacting mode. The ability to collect application-based data is the single largest factor impacting the collection of historical data.

Application Monitoring

An application called a client experience monitor (CEM) has proven potential to provide affiliates with the information they need to quantify the level of service they receive from Internet providers and guide future agreements for continued service. A working prototype of the CEM is explained as well as a snapshot of the data that has been collected.

Client Experience Monitor Prototype

The goal of the CEM is to regularly perform "client-like" tasks. The CEM is responsible for storing application response results along-side "traditional" availability tests (pings - which are performed in parallel) – see Performance Figure below for sample data collected. This data will enable separate CEM tools to produce periodic reports to summarize compliance with service level agreement, and produce a client experience rating based on the responsiveness of the applications supplied by the Internet provider

Design Goals & Hypotheses

It is projected that a delta exists between up time (from a client's perspective) and application availability reported by the Internet provider. The delta will be the result

Performance:

Last Polled: @09:48:04 Tue Sep 14 1999

Service:	IP:	Status:	Ave Ping Response Time				Ave Application Response Time:			
			Day:	Month:	Year:	Overall:	Day:	Month:	Year:	Overall:
dns1	209.32.160.10	UP	0.078	0.064	0.063	0.063	0.147	0.142	0.139	0.139
dns2	209.32.160.11	UP	0.067	0.059	0.060	0.060	0.145	0.139	0.140	0.140
dns3	24.31.3.8	UP	0.053	0.048	0.048	0.048	0.134	0.123	0.122	0.122
dns4	24.31.3.9	UP	0.057	0.069	0.068	0.068	0.093	0.095	0.095	0.095
news	24.31.3.15	UP	0.082	0.083	0.083	0.083	0.263	0.317	0.316	0.316
ntp1	24.31.3.8	UP	0.070	0.076	0.076	0.076	0.397	0.401	0.401	0.401
pop3	24.128.1.94	UP	0.158	0.162	0.162	0.162	0.862	0.784	0.783	0.783
tftp	24.31.3.8	UP	0.048	0.052	0.051	0.051	0.276	0.284	0.284	0.284
web2	24.31.3.10	UP	0.041	0.036	0.036	0.036	0.141	0.144	0.144	0.144

All response times are listed in seconds...

of degradation in application performance to a point where it is unacceptable to the client (or noticeably impacts its ability to use the service). During these periods of degradation the application availability will remain unchanged when in actuality, the application is "effectively down" from a client's perspective.

It is also projected that a relationship between call volume and application availability exits. The increase in call volume as a result of a decrease in

Internet provider due to the resulting increase in load. Instead, the application is "effectively down" much longer from the client's perspective – see Availability figure below for sample data.

The CEM and its data will seek to provide affiliates with a reliable means to monitor the Internet provider's compliance with the SLA. Monitoring of client experience will strive to eliminate potential bottlenecks or single points of failure to provide the most accurate measurement possible. The CME

Availability:

Service:	Ping Avaiability				Application Availability:			
	Day:	Month:	Year:	Overall:	Day:	Month:	Year:	Overall:
dns1	100%	100%	100%	100%	99.14%	99.31%	99.33%	99.33%
dns2	100%	100%	100%	100%	100%	99.86%	99.86%	99.86%
dns3	100%	99.99%	99.99%	99.99%	100%	99.92%	99.92%	99.92%
dns4	100%	99.96%	99.97%	99.97%	100%	99.97%	99.98%	99.98%
news	100%	99.99%	99.99%	99.99%	100%	99.94%	99.93%	99.93%
ntp1	100%	100%	100%	100%	100%	100%	100%	100%
pop3	99.83%	99.18%	99.17%	99.17%	97.96%	98.62%	98.62%	98.62%
tftp	100%	99.99%	99.99%	99.99%	100%	100%	100%	100%
web2	100%	100%	100%	100%	100%	100%	100%	100%

availability would provide evidence of an additional metric that must be considered with respect to the SLA as its currently absorbed by the affiliate. The prototype actually generated data that allowed availability to be plotted but call volume data was not yet available at the time this document was written.

Additionally, it is projected that during application outages the availability of these applications will fail to depict the actual accessibility of resources provided by the

will also seek to establish a range of "acceptable" client experience ratings. This range is expected to raise the bar on the Internet provider's application performance to account for quantifiable demands by the affiliate for higher service quality and capacity.

The design of the CEM is based purely on a "proof of concept". The goal of building the prototype is to demonstrate a working CEM and collect sample data for analysis and hypothesis confirmation. The prototype will also provide direction for follow-on work

and serve as an example for future efforts and/or spin-off projects.

Results

After running the prototype for six weeks several data points were realized. Most importantly all hypothesis were confirmed.

1. *Degradation of performance*
   Several instances of performance degradation occurred during the CEM prototype trial. They included DNS round trip times exceeding one second, POP3 connects exceeding 30 seconds, and NEWS requiring more than one minute to download a single article. These minimums where established arbitrarily and not based on actual tests with customers to determine their acceptance.
2. *Increase in call volume during reduced availability*
   Increased call volumes were confirmed by phoning call center during perceived outages. Each time the call center confirmed that call volume had increased during the duration of the perceived outage. No further effort was spent to determine the actual increase.
3. *Difference between application and host availability*
   Any type of host outage rarely accompanied these periods of degradation. Based on the calculations of the prototype tool the "effective outage" was more nearly three times that of any perceived host outage.

Learned Results

Interestingly, ping outages seemed to follow periods of application outages. It was like the host became overwhelmed with the application demands and went down from a ping perspective. Then a short time later came back up however was bombarded by requests from clients thus resulting in another application outage.

The tool also exposed several configuration errors made by the Internet provider. For example, several DNS servers lay idle while one DNS server seemed to be handling a majority of the requests by clients. This could be simply fixed on the Internet provider's DHCP server if only they had detailed performance information on each application and they were attentive at optimizing the use of every component in their system.

Further Study

It is likely that call volume may vary by application and the extent of the outage. Meaning, some applications may cause more calls than others, just as some outages are more extensive than others. Further study is needed to establish specific relationships between various applications and their various outage tendencies before any kind of penalty can be established for such an outage.

Forward

Since the affiliate is ultimately responsible for providing the service (or seen in the eyes of the customer as responsible for sustaining reliable Internet service), it must seek ways to provide the highest quality service possible. One of the best ways to provide reliable service would be to pass along these requirements to their Internet provider. The following suggests some ways to accomplish this:

- Establish some means of confirming the quality and reliability of the service supplied by the Internet provider.

- Establish motivations for the Internet provider to seek the highest availability possible.
- Dictate terms such that the affiliate will conduct its own service level verification and share this with the Internet provider as a means to allow it to maintain the terms of the contract
- Provide customers with access to current status of various applications, scheduled outage windows, etc.
- Provide the data needed to make more informative decisions regarding handling customer trouble calls and coordinating requested upgrades by the Internet provider.
- Strive to negotiate every service window rather than opening the gates for constant change.

Providing reliable Internet service helps the affiliate in the following ways:

- Increased availability (higher reliability) means lower trouble calls and potentially fewer truck rolls. Every call answered that is trouble related is potentially one less sales call answered.
- Increased availability means higher customer confidence in providing Internet service via cable TV lines and thus opens doors for sales in new markets.
- Increased availability also means more satisfied customers which translates into greater demand.

Consideration of client performance as a driving factor for application availability levels has not yet reached the main stream and "quality" features such as availability and reliability play a limited role in today's customer selection of an Internet information service. However, as customer's choices of Internet access become more equal in terms of speed, capability, price, and flexibility, "quality" will be what differentiates one Internet Information service from another.

As the market for Internet service shifts gears to begin focusing on quality, affiliates need to be ready to quantify the service levels they want to provide. Work at home customers will be one of the first to demand the highest possible levels of service and will likely compare various options before buying. Having access to up to the minute service levels will enable marketing to go after these highly demanding customers. Thus there is a need for such a tool or system to drive up service availability levels and empower affiliates continued growth in the future.