DZone Refcardz

Brought to you by:
StrongLoop

CONTENTS INCLUDE:

Node.js

Server-Side JavaScript for Backends, API Servers and Web Apps

WHAT IS NODE?

The official description according to the nodejs.org website is as follows:

"A platform built on Chrome's JavaScript runtime for easily building fast, scalable network applications."

Translation: Node runs on top of Google's open source JavaScript engine called V8. It's written in C++ and is used in Google's Chrome browser. It's fast!

"Uses an event-driven, non-blocking I/O model that makes it lightweight and efficient."

Translation: Developing distributed, multi-threaded applications using traditionally synchronous languages can be complex and daunting, Node leverages JavaScript's asynchronous programming style via the use of event loops with callbacks to make applications naturally fast, efficient, and non-blocking. If you know JavaScript, you already know quite a bit about Node!

"Perfect for data-intensive real-time applications that run across distributed devices."

Translation: Many application performance problems stem from being I/O bound. Because Node is designed to be non-blocking and event driven when manipulating data, reading files or accessing APIs, it's ideally suited to be distributed across multiple process and machines in a network. Popular uses for Node include web servers, API gateways and backends for mobile applications.

Because it's not limited to one connection per thread like most web server architectures, Node scales to many thousands of concurrent connections. This makes it perfect for writing Mobile and Internet of Things APIs which must interact with many devices in small increments, often holding open a connection while the device connects over a slow network.

NODE IS JAVASCRIPT ON THE SERVER

Node allows developers to write server-side applications in JavaScript. Server-side applications perform tasks that aren't suitably performed on the client, like processing and persisting data or files, plus tasks like connecting to other networked servers, serving web pages and pushing notifications. Seeing that JavaScript is an incredibly popular language with web and mobile front-end developers, the ability to use this same skill to program server-side tasks, in theory, increases a developer's productivity. It may also reduce the need for separate languages or code bases between front-end and backend applications.

HOW DOES NODE WORK?

Synchronous vs asynchronous programming

C and Java traditionally use synchronous I/O, which means time is wasted waiting. You can get around this by writing multithreaded programs, but for some developers, writing these types of applications in a distributed networking environment can be daunting. Of course there is also the issue of the number of threads a system can actually spawn. Node by contrast is a single-threaded way of programming evented, non-blocking, asynchronous I/O applications.

Synchronous vs asynchronous: by analogy

In order to understand non-blocking I/O, let's picture a common scenario. Suppose you are at a restaurant with friends.

A typical experience at a restaurant would be something like this:

- You sit at a table and the server grabs your drink order.
- The server goes back to the bar and passes your order to a bartender.
- While the bartender is working on your drink, the server moves on to grab another table's drink order.
- The server goes back to the bar and passes along the other table's order.
- Before the server brings back your drinks, you order some food.
- Server passes your food order to the kitchen.
- Your drinks are ready now, so the server picks them up and brings them back to your table.
- The other table's drinks are ready, so the server picks them up and takes them to the other table.
- Finally your food is ready, so server picks it up and brings it back to your table.

Basically every interaction with the server follows the same pattern. First, you order something. Then, the server goes on to process your order and return it to you when it's ready. Once the order is handed off to the bar or kitchen, the server is free to get new orders or to deliver previous orders that are completed. Notice that at no point in time is the server doing more than one thing. They can only process one request at a time. This is how non-blocking Node.js applications work. In Node, your application code is like a restaurant server processing orders, and the bar/kitchen is the operating system handling your I/O calls.

Your single-threaded JavaScript application is responsible for all the processing up to the moment it requires I/O. Then, it hands the work off to the operating system, which takes care of processing the rest. Back to our restaurant example, if every time the server got an order request they had to wait for the bar/kitchen to finish before taking the next request, then the service for this restaurant would be very slow and customers would most likely be unsatisfied. This is how blocking I/O works.

Event Loop concurrency model

Node leverages a browser-style currency model on the server. As we all know, JavaScript was originally designed for the browser where events are things like mouse movements and clicks. Moved to the server, this same model allows for the idea of an event loop for server events such as network requests. In a nutshell, JavaScript waits for an *event* and whenever that *event* happens, a *callback* function occurs.
For example, your browser is constantly looping waiting for events like clicks or mouse-overs to occur, but this listening for events doesn't block the browser from performing other tasks. On the server this might mean that instead of a program waiting to return a response until it queries databases,

accesses files or connects to various APIs, it immediately moves on to the next unit of work until the event returns with whatever response was asked of it. Instead of blocking entire programs waiting for I/O to complete, the event loop allows applications to move on and wait for events in order to continue the flow of the program. In this way Node achieves multitasking more efficiently than using threads.

Event Loop analogy

Think of event loops as some delivering mail. They collect the letters or events from the post office (server). These letters can be equated to events or incoming requests that need to be handled i.e. delivered. The letter carrier goes to every mailbox in his area and delivers the letters/events to the destination mailboxes. These destination mailboxes can be equated to JavaScript functions or downstream.

The postman does not wait at the mailbox to receive a reply. When the user of the mailbox responds with a letter, on his routes, he picks it up. Every mailbox has a separate route and routes here can be thought of as the callback. Every incoming letter/request has a callback associated, so that a response can be sent anytime when ready (asynchronously) using the callback routes.

Event Loop code example

Let's look at a simple example of asynchronously reading a file into a buffer. This is a two step process in which first there is a request to read the file, then a callback to handle the file buffer (or error) from the asynchronous file read.

```
var fs = require('fs');
fs.readFile('my_file.txt', function (err, data) {
            if (err) throw err;
            console.log(data);
});
```

The second argument to **readFile** is a callback function that runs after the file is read. The request to read the file goes through Node bindings to **libuv**. Then **libuv** gives the task of reading the file to a thread. When the thread completes reading the file into the buffer, the result goes to V8. It then goes through the Node Bindings in the form of a callback with the buffer. In the callback shown the data argument has the buffer with the file data.

Example of an HTTP server using Node:

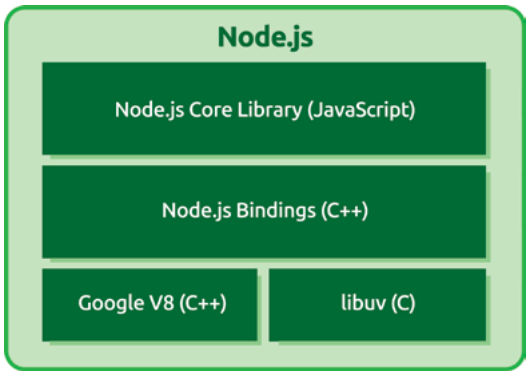
```
var http = require('http');

http.createServer(
 function (request, response) {
   response.writeHead(200, {'Content-Type': 'text/plain'});
   response.end('Hello World\n');
 }
).listen(8080);

console.log('Server running at http://localhost:8080/');
```

ARCHITECTURE

There are four building blocks that constitute Node. First, Node encapsulates libuv to handle asynchronous events and Google's V8 to provide a run-time for for JavaScript. Libuv is what abstracts away all of the underlying network and file system functionality on both Windows and POSIX-based systems like Linux, Mac OSX and Unix. The core functionality of Node, modules like **Assert, HTTP, Crypto** etc., reside in a core library written in JavaScript. The **Node Bindings** provide the glue connecting these technologies to each other and to the operating system.



WHAT ARE THE PERFORMANCE CHARACTERISTICS OF NODE?

Everyone knows benchmarks are a specific measurement and don't account for all cases. Certainly, what and how you measure matters a lot. But there's one thing we can all agree on: at high levels of concurrency (thousands of connections) your server needs to become asynchronous and non-blocking. We could have finished that sentence with IO, but the issue is that if any part of your server code blocks, you're going to need a thread. At these levels of concurrency, you can't go about creating threads for every connection. So, the whole code path needs to be non-blocking and async, not just the IO layer. This is where Node excels.

Here's a collection of articles and blogs concerning Node performance:

Node Performance and Benchmarks

WHAT IS NODE GOOD FOR?

Web Applications

Node is becoming popular for web application because web applications are now slowly shifting from purely server-side rendering to single-page applications to optimize the user experience on the client side.

Reasons why:

- Single page applications have the MVC paradigm self-contained within the browser so that the only server side interaction that is required can be through an efficient API for RPC invocation of server side functions and data behind the firewall or in the cloud
- Node's rich ecosystem of npm modules allows you to build web applications front to back with the relative ease of a scripting language that is already ubiquitously understood on the front end
- Single Page Applications and Node are all built on the common dynamic scripting language of JavaScript

Examples of frameworks for Node:

- Express
- Sails.js
- Compound.js
- Flatiron.js
- Derby.js
- Socketstream.js
- Meteor
- Tower.js

Mobile Backends

Node is popular for backends, especially those required by mobile applications. As an I/O library at its heart, Node's ease of use has been applied toward the classic enterprise application use case to be able to gather and normalize existing data and services.

Reasons why:

- As the shift toward hybrid mobile applications becomes more dominant in the enterprise, the re-use of code written in JavaScript on the client side can be leveraged on the server
- Node's rich ecosystem has almost every underlying driver or connector to enterprise data sources such as RDBMS, Files, NoSQL, etc. that would be of interest to mobile clients
- Node's use of JavaScript as a scripting language makes it easy to normalize data into mobile APIs

Examples of mobile backends built in Node:

- Parse (Proprietary)
- LoopBack (Open source)
- FeedHenry (Proprietary)
- Appcelerator Cloud Services (Proprietary)

API Servers

Node is popular for backends, especially those required by mobile applications.

Reasons why:

- Node utilizes JSON as the content-type for data modeling and for the data payload itself. This lightweight format is already evolving to become the most dominant standard for REST APIs
- Node's rich ecosystem consists of asynchronous libraries that can be easily utilized to handle massive concurrency for the API use case

Examples of open source API Servers built in Node:

- Restify
- Deployd
- LoopBack
- actionhero.js

HOW DO I INSTALL NODE?

The good news is that installers exist for a variety of platforms including Windows, Mac OS X, Linux, SunOS – and of course you can compile it yourself from source. Official downloads are available from the nodejs.org website: http://nodejs.org/download/

HOW CAN I MAKE NODE USEFUL?

What is npm?

Node Package Manager ("npm") is the command-line package manager for Node that manages dependencies for your application. npmjs.org is the public repository where you can obtain and publish modules.

How does npm work?

In order for your Node application to be useful, it is going to need things like libraries, web and testing frameworks, data-connectivity, parsers and other functionality. You enable this functionality by installing specific modules via npm.

There's nothing to install to start using npm if you are already running Node v0.6.3 or higher.

You can install any package with this command:

```
$ npm install <name of module>
```

Some popular and most used modules include:

express
A fast, unopinionated, minimalist web framework for Node. Express aims to provide small, robust tooling for HTTP servers, making it a great solution for single page applications, web sites, hybrids, or public HTTP APIs.

async
Async is a utility module which provides straightforward, powerful functions for working with asynchronous JavaScript. Although originally designed for use with Node, it can also be used directly in the browser. Async provides around 20 functions that include the usual 'functional' suspects (map, reduce, filter, each…) in addition to your async function.

request
A simplified HTTP request client. It supports HTTPS and follows redirects by default.

grunt
A JavaScript task runner that helps automate tasks. Grunt can perform repetitive tasks like minification, compilation, unit testing, linting, etc. The Grunt ecosystem is also quite large with hundreds of plugins to choose from. You can find the listing of plugins here.

socket.io
Socket.io makes WebSockets and real-time possible in all browsers and provides built-in multiplexing, horizontal scalability, automatic JSON encoding/decoding, and more.

mongoose
A MongoDB object modeling tool designed to work in an asynchronous environment. It includes built-in type casting, validation, query building, business logic hooks and more, out of the box.

WHAT IS NEW IN NODE V0.12?

Round-robin clustering

Prior to v0.12, Node's round-robin functionality didn't distribute incoming connections evenly (although everyone expected it to). Node used an old technique which just about all web servers have used at one time or another. The way it typically worked was: a developer would bring up a server and start a few processes that would be made ready to accept new connections. Under the hood, when a new connection was required, all the processes would race to accept the connection.
In theory, this sounded great and should have scaled well, but in practice most operating systems, specifically Linux-based systems, tried to defeat this scheme.

Instead of every available process being considered for a connection, there was the tendency to pick the same process each the time. This meant that the load-balancing scheme didn't work as efficiently as it could. Now, with the new round-robin scheme implemented in v0.12, the master process accepts all the connections and it decides which worker gets to send a response.

For example:

```
var cluster = require('cluster');

// This is the default:
cluster.schedulingPolicy = cluster.SCHED_RR;
// .. or Set this before calling other cluster functions.
cluster.schedulingPolicy = cluster.SCHED_NONE;

// Spawn as many workers as there are CPUs in the system.
for (var i = 0, n = os.cpus().length; i < n; i += 1)
 cluster.fork();
```

Note that the new round-robin scheme is on by default on all operating systems except for Windows. To learn more about the clustering feature, read the official docs and a technical blog by Node contributor Ben Noordhuis.

Multi-Context: running multiple instances in a single process

An oft-requested feature was the ability to embed Node.js in other applications, particularly in ways that let it integrate with other event loops and ("while you're at it") with support for multiple Node execution contexts: that is, the ability to have multiple instances of Node co-exist peacefully within the same process. Imagine a phone or network switch where it is performing routing logic for multiple connections, but in a single process and you're not far off.

In Node v0.12 you can now use multiple execution contexts from within the same event loop. Don't worry, from a user perspective, there are no visible changes, everything still works like before. But if you are an embedder or a native add-on author you should read a technical blog by Ben Noordhuis on how it works.

spawnSynch execSync

Although it might be a bit odd to think about adding these synchronous features to Node, many developers have been implementing various hacks to get precisely this type of synchronous behavior. **execSync** and **spawnSync** work by running a process, then blocking while it runs, and then exiting when the function returns. The motivation for putting this feature into v0.12 was that an increasing number of developers were using Node for more than just writing servers.

For example, many developers are using Node as a replacement for shell scripting. Grunt is a perfect example of this. Grunt is like make, in that it allows you to run small tasks to set things up on your operating system. Under the hood it relies heavily on shelljs. Shelljs goes to great lengths to emulate execsyncc although Node doesn't actually support it.

For example:

```
var child_process = require('child_process');
var fs = require('fs');

function execSync(command) {
 // Run the command in a subshell
 child_process.exec(command + ' 2>&1 1>output && echo done! > done');

 // Block the event loop until the command has executed.
 while (!fs.existsSync('done')) {
   // Do nothing
 }

 // Read the output
 var output = fs.readFileSync('output');

 // Delete the output and done files
 fs.unlinkSync('output');
 fs.unlinkSync('done');

 return output;
}
```

The above code isn't particularly efficient. As of v0.12, the execSync and spawnSync APIs are supported. How it works under the hood is: a nested loop is spawned in libuv, which only reads the output and sleeps when nothing is happening. For example:

```
var spawnSync = require('child_process').spawnSync;

var result = spawnSync('cat',
                       ['-'],
                       { input: 'hello world!',
                         encoding: 'utf8' });

console.log('exit code: %d', result.exitCode);
console.log('output: %s', result.stdout);
console.log('error: %s', result.stderr);
```

To learn more about the spawnSync and execSync APIs, read this technical blog by Node contributor Bert Belder.

Profiling APIs

Prior to v0.12, profiling could only be enabled at startup and heap dumps required a native add-on. In the latest release, APIs have been added to address these issues so that if you wanted to monitor gc behavior for example, you could use the following:

```
var v8 = require('v8');

v8.cpuProfiler.setSamplingInterval(1);
v8.cpuProfiler.start();

v8.on('gc', function() {
 console.log('Garbage collection just happened!');
});
```

To learn more about the profiling APIs, you can read this technical blog by Node contributor Ben Noordhuis that dives deep into the performance optimizations in the latest release.

NODE API GUIDE

Below is a list of the most commonly used Node APIs. For a complete list and for an APIs current state of stability or development, please consult the Node API documentation.

Buffer
Functions for manipulating, creating and consuming octet streams, which you may encounter when dealing with TCP streams or the file system. Raw data is stored in instances of the Buffer class. A Buffer is similar to an array of integers but corresponds to a raw memory allocation outside the V8 heap. A Buffer cannot be resized.

Child Process
Functions for spawning new processes and handling their input and output. Node provides a tri-directional popen(3) facility through the child_process module.

Cluster
A single instance of Node runs in a single thread. To take advantage of multi-core systems, the user will sometimes want to launch a cluster of Node processes to handle the load. The cluster module allows you to easily create child processes that all share server ports.

Crypto
Functions for dealing with secure credentials that you might use in an HTTPS connection. The crypto module offers a way of encapsulating secure credentials to be used as part of a secure HTTPS net or http connection. It also offers a set of wrappers for OpenSSL's hash, hmac, cipher, decipher, sign and verify methods.

Debugger
You can access the V8 engine's debugger with Node's built-in client and use it to debug your own scripts. Just launch Node with the debug argument (**node debug server.js**). A more feature-filled alternative debugger is node-inspector. It leverages Google's Blink DevTools, allows you to navigate source files, set breakpoints and edit variables and object properties, among other things.

Events
Contains the **EventEmitter** class used by many other Node objects. Events defines the API for attaching and removing event listeners and interacting with them. Typically, event names are represented by a camel-cased string; however, there aren't any strict restrictions on case, as any string will be accepted. Functions can then be attached to objects, to be executed when an event is emitted. These functions are called **listeners**. Inside a listener function, the object is the **EventEmitter** that the listener was attached to. All EventEmitters emit the event **newListener** (when new listeners are added) and **removeListener** (when a listener is removed).

To access the EventEmitter class use:

```
require('events').EventEmitter.


emitter.on(event, listener) adds a listener to the end of the listeners array
for the specified event. For example:


server.on('connection', function (stream) {
 console.log('someone connected!');
});
```

This calls returns emitter, which means that calls can be chained.

Globals
Globals allow for objects to be available in all modules. (Except where noted in the documentation.)

HTTP
This is the most important and most used module for a web developer. It allows you to create HTTP servers and make them listen on a given port. It also contains the **request** and **response** objects that hold information about incoming requests and outgoing responses. You also use this to make HTTP requests from your application and do things with their responses. HTTP message headers are represented by an object like this:

```
{ 'content-length': '123',
 'content-type': 'text/plain',
 'connection': 'keep-alive',
 'accept': '*/*' }
```

In order to support the full spectrum of possible HTTP applications, Node's HTTP API is very low-level. It deals with stream handling and message parsing only. It parses a message into headers and body but it does not parse the actual headers or the body.

Modules
Node has a simple module loading system. In Node, files and modules are in one-to-one correspondence. As an example, foo.js loads the module circle.js in the same directory.

The contents of foo.js:

```
var circle = require('./circle.js');
console.log( 'The area of a circle of radius 4 is '
          + circle.area(4));
```

The contents of circle.js:

```
var PI = Math.PI;

exports.area = function (r) {
 return PI * r * r;
};

exports.circumference = function (r) {
 return 2 * PI * r;
};
```

The module circle.js has exported the functions **area()** and **circumference()**. To add functions and objects to the root of your module, you can add them to the special exports object. Variables local to the module will be private, as though the module was wrapped in a function. In this example the variable PI is private to circle.js.

Net
Net is one of the most important pieces of functionality in Node core. It allows for the creation of network server objects to listen for connections and act on them. It allows for the reading and writing to sockets. Most of the time, if you're working on web applications, you won't interact with Net directly. Instead you'll use the **HTTP** module to create HTTP-specific servers. If you want to create TCP servers or sockets and interact with them directly, you'll want to work with the Net API.

Process
Used for accessing stdin, stdout, command line arguments, the process ID, environment variables, and other elements of the system related to the currently-executing Node processes. It is an instance of **EventEmitter**. Here's example of listening for **uncaughtException**:

```
process.on('uncaughtException', function(err) {
 console.log('Caught exception: ' + err);
});

setTimeout(function() {
 console.log('This will still run.');
}, 500);

// Intentionally cause an exception, but don't catch it.
nonexistentFunc();
console.log('This will not run.');
```

REPL
Stands for **R**ead-**E**val-**P**rint-**L**oop. You can add a REPL to your own programs just like Node's standalone REPL, which you get when you run node with no arguments. REPL can be used for debugging or testing.

Stream
An abstract interface for streaming data that is implemented by other Node objects, like HTTP server requests, and even stdio. Most of the time you'll want to consult the documentation for the actual object you're working with

rather than looking at the interface definition. Streams are readable, writable, or both. All streams are instances of EventEmitter.

VM
Allows you to compile arbitrary JavaScript code and optionally execute it in a new sandboxed environment immediately, or saved for each client and each server run later.

DEVELOPER TOOLS FOR NODE

Below are some key tools widely adopted in the enterprise and in the community for developing Node applications:

Development Environments

Product/Project	Features/Highlights
WebStorm	- Code analysis - Cross-platform - VCS integration
Sublime Text	- Goto anything - Customizable - Cross-platform
Nide	- Project tree display - Npm integration - Command-line and Mac
Nodeclipse	- Open source - Built on Eclipse
Cloud9 IDE	- Cloud-based - Collaborative - Debug and deploy
IntelliJ	- Node plugin - Code completion - Code analysis

Application Performance Monitoring

Product/Project	Features/Highlights
StrongOps	- Error tracing - Event loop response times - Slowest endpoints
New Relic	- Error rates - Transaction response times - Throughput monitoring
AppDynamics	- Error tracing - Endpoint response time - Historical metrics

Debugging

Product/Project	Features/Highlights
V8 Debugger	- Manual code injection - Breakpoints - Event exception handling
StrongOps – Node Inspector	- Google Blink Dev-Tools based
Cloud9 IDE	- Cloud-based - Code completion - Debug and deploy
WebStorm	- Code analysis - Cross-platform - VCS integration
Nodeclipse	- Code completion - Built-on Eclipse - Tracing and breakpoints
DTrace	- SmartOS only - Transaction tracing

RESOURCES

- StrongLoop **website**
- **StrongLoop** technical blog
- **Official Node website:** nodejs.org
- Node downloads
- Node documentation
- Node on GitHub
- **Official npm website:** npmjs.org
- npm documentation
- Node Linkedin Group

ABOUT THE AUTHOR

Having been programming since he was a little kid, **Bert Belder** got involved with open source when he started to port Node.js to Windows. It got him the core contributor badge, and he hasn't left the project since. In 2012, Bert founded his company StrongLoop together with long-time Node contributor Ben Noordhuis, in an effort to lead Node to world domination.

RECOMMENDED BOOK

This book introduces you to Node, the new web development framework written in JavaScript. You'll learn hands-on how Node makes life easier for experienced JavaScript developers: not only can you work on the front end and back end in the same language, you'll also have more flexibility in choosing how to divide application logic between client and server.

BUY NOW

DZone, Inc.
150 Preston Executive Dr.
Suite 201
Cary, NC 27513

888.678.0399
919.678.0300

Refcardz Feedback Welcome
refcardz@dzone.com

Sponsorship Opportunities
sales@dzone.com

ISBN-13: 978-1-936502-80-6
ISBN-10: 1-936502-80-1
50795
9 781936 502806
$7.95

Version 1.0